

Smoothed analysis for low-rank solutions to semidefinite programs in quadratic penalty form

Srinadh Bhojanapalli

TTI Chicago

SRINADH@TTIC.EDU

Nicolas Boumal

Princeton University

NBOUMAL@MATH.PRINCETON.EDU

Prateek Jain

Praneeth Netrapalli

Microsoft Research

PRAJAIN@MICROSOFT.COM

PRANEETH@MICROSOFT.COM

Editors: Sébastien Bubeck, Vianney Perchet and Philippe Rigollet

Abstract

Semidefinite programs (SDP) are important in learning and combinatorial optimization with numerous applications. In pursuit of low-rank solutions and low complexity algorithms, we consider the Burer–Monteiro factorization approach for solving SDPs. For a large class of SDPs, upon random perturbation of the cost matrix, with high probability, we show that all approximate second-order stationary points are approximate global optima for the penalty formulation of appropriately rank-constrained SDPs, as long as the number of constraints scales sub-quadratically with the desired rank. Our result is based on a simple penalty function formulation of the rank-constrained SDP along with a smoothed analysis to avoid worst-case cost matrices. We particularize our results to two applications, namely, Max-Cut and matrix completion.

1. Introduction

Semidefinite programs (SDP) are an important class of optimization problems (Vandenberghe and Boyd, 1996), and are critical to several learning-related tasks, including clustering (Shi and Malik, 2000; Abbe, 2017), matrix completion and regression (Recht et al., 2010; Candès and Recht, 2009), kernel learning (Lanckriet et al., 2004), and sum-of-squares relaxations (Barak et al., 2015).

However, solving SDPs in practice is a challenging task. Consider the following canonical SDP:

$$\begin{aligned} & \underset{X \in \mathbb{R}^{n \times n}}{\text{minimize}} && \langle C, X \rangle \\ & \text{subject to} && \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m, \text{ and } X \succeq 0, \end{aligned} \tag{1}$$

where $C, A_1, \dots, A_m \in \mathbb{R}^{n \times n}$ are symmetric matrices, $\langle A, B \rangle = \text{Tr}(A^T B)$, and X is positive semidefinite. Such problems are convex and can be solved in polynomial time using classical iterative algorithms such as ellipsoid and interior-point methods (Nesterov et al., 1994). However, these algorithms have super-linear complexity (in input size) and tend to scale poorly in practice. As such, they are not well suited for typical learning tasks where both m and n can be fairly large. The two key challenges for these algorithms are: (a) a search space of high dimension on the order of n^2 ; and (b) the need to maintain positive semidefiniteness of the variable matrix X throughout the iterations.

In response to these challenges, [Burer and Monteiro \(2003, 2005\)](#) suggested solving (1) by constraining the search space to matrices of rank at most k , using a parameterization of the form $X = UU^T$ where $U \in \mathbb{R}^{n \times k}$. This reduces the number of variables from $O(n^2)$ to $O(nk)$, and mechanically enforces positive semidefiniteness:

$$\begin{aligned} & \underset{U \in \mathbb{R}^{n \times k}}{\text{minimize}} && \langle C, UU^T \rangle \\ & \text{subject to} && \langle A_i, UU^T \rangle = b_i, \quad i = 1, \dots, m. \end{aligned} \quad (2)$$

This is equivalent to (1) with the additional constraint $\text{rank}(X) \leq k$. This rank constraint is fairly natural, as several SDPs of interest are themselves relaxations of rank-constrained problems. Moreover, [Barvinok \(1995\)](#); [Pataki \(1998\)](#) showed that for every feasible compact SDP, there exists a rank $O(\sqrt{m})$ solution that is also globally optimal. While this ensures that the global optimum of the factored SDP problem (with $k = \Omega(\sqrt{m})$) maps to a global optimum of the original SDP, it is not immediately clear how to solve the factorized problem.

In fact, the factorized problem is a non-convex quadratically constrained quadratic program which in general can be NP-hard. The challenge in solving the problem arises due to the non-convexity as well as due to constraints. In this work, we consider a simple penalty method approach that replaces the constraints with a quadratic penalty in the objective function. In settings where it is not easy to project onto the constraint set (as in problem (1)), it is natural for optimization algorithms to resort to penalty formulations. This is the strategy of popular methods including the interior point, augmented Lagrangian and primal-dual methods. As a first step we consider the simplest such formulation, namely the quadratic penalty method.

The proposed penalty formulation is given by:

$$\underset{X \succeq 0}{\text{minimize}} \quad F_\mu(X) = \langle C, X \rangle + \mu \sum_{i=1}^m (\langle A_i, X \rangle - b_i)^2, \quad (3)$$

where μ is generally a large positive constant. Notice that this is a convex problem. Intuitively, for increasingly large μ , solutions of (3) converge to solutions of (1).

Combining that formulation with the Burer–Monteiro factorization we get:

$$\underset{U \in \mathbb{R}^{n \times k}}{\text{minimize}} \quad L_\mu(U) = \langle C, UU^T \rangle + \mu \sum_{i=1}^m (\langle A_i, UU^T \rangle - b_i)^2. \quad (4)$$

The cost function L_μ is non-convex, and generic optimization algorithms can only guarantee computation of an approximate second-order stationary point (SOSP) ([Cartis et al., 2012](#); [Ge et al., 2015](#)). That is, such algorithms converge to a point U where the gradient of L_μ is small and the Hessian of L_μ is almost positive semidefinite. Such approximate SOSPs need not be close to optimal in general.

As a first contribution, we construct an explicit SDP where a suboptimal SOSP exists even for k as large as $n - 1$. However, we show that such bad SDPs have measure zero among all SDPs when k is large enough. Specifically, we show that if the cost matrix has a small amount of randomness and $\frac{k(k+1)}{2} > m$, then, almost surely, *any* SOSP of L_μ is a global optimum. That is, when $k = \Omega(\sqrt{m})$, for almost all cost matrices C , an SOSP of (4) corresponds to a global optimum. We would like to stress here that for certain non-compact SDPs existence of an SOSP itself is not guaranteed. However, as shown in Section 5, SOSPs exist for several important classes of SDPs.

We next address the question of approximate optimality for approximate SOSPs, as optimization algorithms can only recover approximate SOSPs in polynomial time. Since there is a measure zero set of SDPs with bad SOSPs, there can be a non-zero (but likely small) measure set of SDPs with bad approximate SOSPs. We use *smoothed analysis* to avoid these bad SDPs, by perturbing the objective matrix. We show that for $k = \tilde{\Omega}(\sqrt{m})$ and with high probability upon an appropriately scaled perturbation, any approximate SOSP of L_μ with a perturbed objective and bounded residues is approximately optimal for the perturbed and penalized objective (3). We further discuss settings under which all SOSPs of the penalized problem have bounded norm (and hence bounded residues).

1.1. Main results

The main contributions of this work as follows.

- We propose a simple penalty version of the factored SDP (2) and show that, for almost all cost matrices C , any exact SOSP of the rank-constrained formulation (4) is a global optimum for rank $\Omega(\sqrt{m})$ —see Corollary 4. This result removes the smooth manifold requirement of (Boumal et al., 2016), though it applies to (3), not (1).
- We show that there indeed exists a compact, feasible SDP with a worst-case C for which the penalized, factorized problem admits a suboptimal SOSP (see Theorem 5), even for rank almost as big as the dimension.
- We show that by perturbing the objective function slightly and by performing a smoothed analysis on the resulting problem, we can guarantee with high probability that every approximate SOSP of the perturbed problem is an approximate global optimum of the perturbed and penalized SDP. Hence, we can use standard techniques (Cartis et al., 2012; Ge et al., 2015) to find approximate SOSPs and bound the optimality gap—see Theorem 13.

In summary, we show that for a class of SDPs with bounded solutions, we can find a low-rank solution that is close to global optimality for the penalized objective. Note that finding the smallest rank matrix satisfying a set of linear equations is NP hard (Natarajan, 1995). Our results show how increasing the number of parameters (rank) makes the optimization of this non-convex problem easier. While the extreme case setting $k = n$ makes the constraint trivial, our results show optimality for a non-trivial rank ($\tilde{\Omega}(\sqrt{m})$), and it is an interesting question to understand this trade-off in more detail. We believe that the factorization technique can be leveraged to design faster SDP solvers, and any looseness in the current bounds is an artifact of our proof, which hopefully can be tightened in future works.

1.2. Prior work

Fast solvers for SDPs have garnered interest in the optimization and in the theoretical computer science communities for a long time. Most of the existing results for SDP solvers can be categorized into direct (convex) methods and factorization methods.

Convex methods: Classical techniques such as interior point methods (Nesterov and Nemirovski, 1989, 1988; Alizadeh, 1995) and cutting plane methods (Anstreicher, 2000; Krishnan and Mitchell, 2003) enjoy geometric convergence, but their computational complexity per iteration is high. As a result, it is hard to scale these methods to SDPs with a large number of variables.

With the goal of speeding up the computation, many works have considered: i) a specific and important class of SDPs, namely, SDPs with a trace constraint ($\text{Tr}(X) = 1$), and ii) methods with sub-linear convergence. For these SDPs, [Arora et al. \(2005\)](#) proposed a multiplicative weights method which provides faster techniques for some graph problems, with running time depending on $O(\frac{1}{\epsilon^2})$ and the *width* of the problem. [Hazan \(2008\)](#) proposed a Frank–Wolfe-type algorithm with a complexity of $\tilde{O}(\frac{Z}{\epsilon^{3.5}})$ where Z is the sparsity of C and the A_i ’s. [Garber and Hazan \(2016\)](#); [Garber \(2016\)](#) proposed faster methods that either remove the dependence on Z (sub-linear time), or improve the dependence on ϵ . While these methods improve the per iteration complexity, they still need significant memory as the rank of solutions for these methods is not bounded, and scales at least at the rate of $O(\frac{1}{\epsilon})$. An exception to this is the work by [Yurtsever et al. \(2017\)](#), which uses sketching techniques in combination with conditional gradient method to maintain a low rank representation. However this method is guaranteed to find a low rank optimum only if the conditional gradient method converges to a low rank solution.

Factorization methods: [Burer and Monteiro \(2003, 2005\)](#) proposed a different approach to speed up computations, namely by searching for solutions with smaller rank. Even though all feasible compact SDPs have at least one solution of rank $O(\sqrt{m})$ ([Barvinok, 1995](#); [Pataki, 1998](#)), it is not an easy task to optimize directly on the rank-constrained space because of non-convexity. However, [Burer and Monteiro \(2003, 2005\)](#) showed that any *rank-deficient local minimum* is optimal for the SDP. [Journée et al. \(2010\)](#) extended this result to any rank-deficient SOSP under restrictive conditions on the SDP. However, these results cannot guarantee that SOSPs are rank deficient, or at least that rank-deficient SOSPs can be computed efficiently (or even exist). [Boumal et al. \(2016, 2018b\)](#) address this issue by showing that for a particular class of SDPs satisfying some regularity conditions, and for almost all cost matrices C , any SOSP of the rank-constrained problem with $k = \Omega(\sqrt{m})$ is a global optimum. Later, [Mei et al. \(2017\)](#) showed that for SDPs with elliptic constraints (similar to the Max-Cut SDP), any rank- k SOSP gives a $(1 - \frac{1}{k-1})$ approximation to the optimum value. Both these results are specific to particular classes of SDPs and do not extend to general problems.

In a related setup, [Keshavan et al. \(2010\)](#); [Jain et al. \(2013\)](#) have showed that rank-constrained matrix completion problems can be solved using smart initialization strategies followed by local search methods. Following this, many works have identified interesting statistical conditions under which certain rank-constrained matrix problems have no spurious local minima ([Sun et al., 2016](#); [Bandeira et al., 2016](#); [Ge et al., 2016](#); [Bhojanapalli et al., 2016b](#); [Park et al., 2017](#); [Ge et al., 2017](#); [Zhu et al., 2017](#); [Ge and Ma, 2017](#)). These results are again for specific problems and do not extend to general SDPs.

In contrast, our result holds for a large class of SDPs in penalty form, without strong assumptions on the constraint matrices A_i and for a large class of cost matrices C . We avoid degenerate SDPs with spurious local minima by perturbing the problem and then using a smoothed analysis, which is one of the main contribution of the work.

After the initial preprint of this work became available, we learned about work by [Du and Lee \(2018\)](#), who in a parallel have shown optimality of exact SOSPs using similar techniques as we used in Section 2, but for single hidden layer neural networks with quadratic activations.

Notation

For a smooth function $f(X)$, we refer to first-order stationary points X as FOSPs. Such points satisfy $\nabla f(X) = 0$ (zero gradient). We refer to second-order stationary points at SOSPs. Such points are FOSPs and furthermore satisfy $\nabla^2 f(X) \succeq 0$, i.e., the Hessian is positive semidefinite. The set of symmetric matrices of size n is $\mathcal{S}^{n \times n}$. $\sigma_i(\cdot)$ and $\lambda_i(\cdot)$ denote the i th singular- and eigenvalues respectively, in decreasing order.

2. Exact second-order points typically are optimal

In this section, we study the second-order stationary points of our penalty formulation (4) and show that for “typical” cost matrices C , *exact* SOSPs are optimal for (4) as long as $k = \Omega(\sqrt{m})$.

Our result is based on a simple but powerful argument that has appeared in various forms before, notably in (Burer and Monteiro, 2005). The argument claims that any *rank-deficient* local optimum of (4) (which is really a parameterized version of (3) with a rank constraint) should map to a local optimum of (3) as the constraint $\text{rank}(X) \leq k$ is not active. Since (3) is convex, every local optimum is a global optimum, hence a rank-deficient local optimum of (4) maps to a global optimum of (3). Interestingly, the result holds even if U is just an SOSP rather than a local optimum, something that is readily apparent from the proofs in (Journée et al., 2010), albeit in a restricted setting.

Lemma 1 *Let $f(X)$ be a convex, twice continuously differentiable function of $X \in \mathcal{S}^{n \times n}$. Consider the convex problem*

$$\underset{X \succeq 0}{\text{minimize}} f(X). \quad (5)$$

Now consider the rank-constrained factorized version of the problem:

$$\underset{U \in \mathbb{R}^{n \times k}}{\text{minimize}} g(U) = f(UU^T). \quad (6)$$

If U is an SOSP of (6) with $\text{rank}(U) < k$, then U is a global minimum of (6) and UU^T is a global minimum of (5). (Notice that such a point may not exist in general.)

See Appendix D for a detailed proof.

Thus, (column) rank-deficient SOSPs of (4) are globally optimal and map to global optima of (3). A direct corollary states that non-convexity is benign if $k = n + 1$.

Corollary 2 *Given an SDP in penalized and factorized form (4) with $k > n$, deterministically, any SOSP U is a global optimum, and UU^T is a global optimum for (3).*

Yet, the main goal is to make a statement for small k , so as to reduce the dimensionality of the search space. Unfortunately, in general, SOSPs of non-convex cost functions need not have rank less than k for arbitrary k .

However, the following lemma asserts that, for almost all cost matrices C , provided k grows like \sqrt{m} , all FOSPs (a fortiori, all SOSPs) are rank deficient. Our proof is the same as that of (Boumal et al., 2016, Lemma 9) but the main statement as well as the cost function and conditions on constraints are stated more generally. In particular, unlike the statement in that reference, we do not require that the feasible set of (2) form a smooth manifold.

Lemma 3 Choose k such that $\frac{k(k+1)}{2} > m$. For almost all $C \in \mathcal{S}^{n \times n}$, any FOSP $U \in \mathbb{R}^{n \times k}$ of (4) (if one exists) satisfies $\text{rank}(U) < k$.

See Appendix D for a detailed proof.

These two lemmas lead to an important corollary regarding the factorization approach.

Corollary 4 Given an SDP in penalized and factorized form (4) with k such that $\frac{k(k+1)}{2} > m$, for almost any cost matrix C , deterministically, any SOSP U is a global optimum, and UU^T is a global optimum for (3).

To ensure existence of such solutions, it is necessary to include additional conditions (for example, on the constraints of the SDP.) From (Pataki, 1998; Barvinok, 1995), it is known that SDPs with non-empty, compact search spaces can have a unique solution of rank up to the maximal k such that $\frac{k(k+1)}{2} \leq m$. This indicates that, in general, the condition on k cannot be improved.

These observations lead to the following two natural questions:

1. This result holds only for “typical” C . Is this an artifact of our proof technique, or is it necessary to exclude a zero-measure set of cost matrices C ?
2. This result holds only for *exact* SOSPs, which in general are hard to compute. Numerical methods tend to provide approximate SOSPs only. Can we extend the results to *approximate* SOSPs as well?

The next section answers the first question in the affirmative: there do exist “bad” matrices C for some SDPs, so that any result of the type of Corollary 4 must exclude at least some SDPs. To address the second question, we resort to smoothed analysis, that is, for large classes of SDPs in penalty form, upon perturbing the cost matrix randomly, we show that, with high probability, approximate SOSPs are also good enough to obtain approximately globally optimal solutions of the perturbed problem.

3. Exact second-order points sometimes are suboptimal

Below, we construct an SDP which confirms that it is indeed necessary (in full generality) to exclude some SDPs in Corollary 4, even if k is allowed to grow large.

Pick $n \geq 3$ and set $\epsilon = \sqrt{\frac{6}{n-1}}$. Consider the following $m = n + 1$ constraint matrices in $\mathcal{S}^{n \times n}$:

$$\begin{aligned} A_i &= e_i e_n^\top + e_n e_i^\top, \quad i = 1, \dots, n-1, \\ A_n &= \epsilon \begin{bmatrix} I_{n-1} & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad A_{n+1} = \epsilon \begin{bmatrix} 2I_{n-1} & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned}$$

where $e_i \in \mathbb{R}^n$ is the i th standard basis vector (the i th column of I_n). In words, for $i = 1, \dots, n-1$, each A_i has only two non-zero entries—both equal to one—located in row i of the last column and symmetrically in column i of the last row.

Pair these matrices with the right-hand side vector $b \in \mathbb{R}^m$ defined by

$$b_1 = \dots = b_{n-1} = 0, \quad \text{and} \quad b_n = b_{n+1} = \epsilon \frac{5(n-1)}{3}.$$

Finally, set the cost matrix C to be zero. (A distinct advantage of picking $C = 0$ is that it makes the choice of $\mu > 0$ irrelevant in defining (4).) These prescriptions fully define the SDP (1) and its associated factorized and penalized problem (4), which we can write here as:

$$\underset{U \in \mathbb{R}^{n \times k}}{\text{minimize}} \quad L(U) = \frac{1}{2} \sum_{i=1}^{n+1} (\langle A_i, UU^T \rangle - b_i)^2. \quad (7)$$

Theorem 5 *The SDP defined above admits a global optimum of rank 1. Furthermore, for $k = n - 1$, $\bar{U} \triangleq \begin{bmatrix} I_{n-1} \\ 0 \end{bmatrix}$ is a suboptimal SOSP of (7).*

See Appendix E for a detailed proof of the theorem.

4. Approximate second-order points: smoothed analysis

Recall that Corollary 4 shows that *exact* SOSPs of (4) are optimal for almost all cost matrices C . However, obtaining exact SOSPs is challenging in practice. Standard optimization algorithms such as the trust-region method and the cubic regularization method (Nesterov and Polyak, 2006; Cartis et al., 2012), when run for finitely many iterations, converge to an *approximate* SOSP only, as defined below. All proofs for this section are in Appendix F.

Definition 6 (ϵ -FOSP) *We call U an ϵ -FOSP of a function $f(U)$ if:*

$$\|\nabla f(U)\|_F \leq \epsilon.$$

Definition 7 ((ϵ, γ) -SOSP) *We call U an (ϵ, γ) -SOSP of a function $f(U)$ if:*

$$\|\nabla f(U)\|_F \leq \epsilon \text{ and } \lambda_{\min}(\nabla^2 f(U)) \geq -\gamma\sqrt{\epsilon}.$$

As an extension to Lemma 1—which states rank-deficient *exact* SOSPs are optimal—we now show that approximate SOSPs which are also approximately rank deficient are indeed approximately optimal. To this end, we define the linear operator $\mathcal{A}: \mathcal{S}^{n \times n} \rightarrow \mathbb{R}^m$ with $\mathcal{A}(X)_i = \langle A_i, X \rangle$. We use the following notion of norm for \mathcal{A} :

$$\|\mathcal{A}\| \triangleq \max_{Y \in \mathcal{S}^{n \times n}, \|Y\|_F \leq 1} \|\mathcal{A}(Y)\|_2, \quad \|\mathcal{A}\| = \|\mathcal{A}^*\| \triangleq \max_{y \in \mathbb{R}^m, \|y\|_2 \leq 1} \|\mathcal{A}^*(y)\|_F. \quad (8)$$

Furthermore, we define the residue at a point U to be the vector of constraint violations:

$$\mathbf{r} = \mathbf{r}(U) = \mathbf{r}(UU^T) \triangleq \mathcal{A}(UU^T) - \mathbf{b}. \quad (9)$$

Lemma 8 *Let $U \in \mathbb{R}^{n \times k}$ be an (ϵ, γ) -SOSP of (4) such that $\sigma_k^2(U) \leq \frac{\gamma\sqrt{\epsilon}}{8\mu\|\mathcal{A}\|^2}$. Then,*

$$\lambda_{\min}(C + 2\mu\mathcal{A}^*(\mathbf{r})) \geq -\gamma\sqrt{\epsilon}.$$

Furthermore, if a global optimum \tilde{X} for (3) exists, then the optimality gap obeys:

$$F_\mu(UU^T) - F_\mu(\tilde{X}) \leq \gamma\sqrt{\epsilon} \text{Tr}(\tilde{X}) + \frac{1}{2}\epsilon\|U\|_F.$$

(Once again, we stress that U and \tilde{X} as prescribed may not exist in general.)

To reach a statement about approximate optimality of approximate SOSPs, it remains to show that approximate FOSPs are approximately rank deficient. Such a result would constitute a generalization of Lemma 3. In that lemma, we had to exclude a pathological set of “bad” matrices C . Hence, here too, we expect to encounter difficulties with some C ’s.

For this reason, we resort to a *smoothed analysis*. That is: on the off-chance that the cost matrix C is “bad”, we perturb it with a random Gaussian matrix. We further assume that (a) k is large enough, and (b) approximate FOSPs have bounded residues \mathbf{r} . That residues are indeed bounded is established under special conditions in later subsections.

Theorem 9 *Draw a random matrix G with $G_{ij} \sim \mathcal{N}(0, \sigma_G^2)$ i.i.d. for $i \leq j$ and $G = G^T$. Let $U \in \mathbb{R}^{n \times k}$ be an ϵ -FOSP of (4) with perturbed cost matrix $C + G$. Assume there exists a constant B which only depends on the problem parameters $\mathcal{A}, \mathbf{b}, C$ and μ such that:*

1. *With probability at least $1 - \delta$ on the choice of G , all ϵ -FOSPs of the perturbed problem have bounded residue: $\|\mathbf{r}\|_2 \leq B$, and*
2. *$k \geq 3 \left[\log(\frac{n}{\delta'}) + \sqrt{\text{rank}(\mathcal{A}) \log \left(1 + \frac{8\mu B \|\mathcal{A}\| \sqrt{c_0 n}}{\sigma_G} \right)} \right]$ for some $\delta' \in (0, 1)$, where c_0 is a universal constant.*

Then, with probability at least $1 - \delta - \delta'$,

$$\sigma_k(U) \leq \frac{2\epsilon}{\sigma_G} \frac{\sqrt{c_0 n}}{k}.$$

Crucially, notice that $\text{rank}(\mathcal{A}) \leq m$, so (up to log factors) k is required to grow like \sqrt{m} , as desired.

4.1. Compact SDPs

To leverage Lemma 8 and Theorem 9, we must control the residues at approximate FOSPs of (4). This is delicate in general. In this part, we make the following assumption.

Assumption 10 *The search space $\mathcal{C} = \{X \succeq 0 : \mathcal{A}(X) = \mathbf{b}\}$ of the SDP (1) is non-empty and compact, where $\mathcal{A}: \mathcal{S}^{n \times n} \rightarrow \mathbb{R}^m$ is the linear operator defined by $\mathcal{A}(X)_i = \langle A_i, X \rangle$.*

When this is the case, standard results from (Barvinok, 1995; Pataki, 1998) guarantee the existence of a global optimum of rank r where $\frac{r(r+1)}{2} \leq m$ for the SDP (1)—always. It is reasonable to expect such low-rank solutions might also exist for the penalized problem (3), and that one should be able to compute these by solving the factorized problem (4)—at least, generically. This section is about making these expectations precise in the soft case, where one only computes approximate SOSPs.

A technical necessity in our proofs is to show that FOSPs of (4) have bounded norm. To do this, we need a technical modification of (4). Specifically, consider the following geometric fact.

Proposition 11 *For a given SDP (1), assume \mathcal{C} is non-empty. Then, \mathcal{C} is compact if and only if there exists a positive definite matrix A_0 and a nonnegative real b_0 such that $\langle A_0, X \rangle = b_0$ for all $X \in \mathcal{C}$. Furthermore, unless $\mathcal{C} = \{0\}$, $b_0 > 0$.*

Thus, under Assumption 10, we can rewrite (1) with an explicit redundant constraint involving $A_0 \succ 0$:

$$\begin{aligned} & \underset{X \in \mathcal{S}^{n \times n}}{\text{minimize}} && \langle C, X \rangle \\ & \text{subject to} && \langle A_0, X \rangle = b_0, \\ & && \langle A_i, X \rangle = b_i, \quad i = 1, \dots, m, \text{ and } X \succeq 0. \end{aligned} \quad (10)$$

Accordingly, we define $\tilde{\mathcal{A}}: \mathcal{S}^{n \times n} \rightarrow \mathbb{R}^{m+1}$ and $\tilde{\mathbf{b}} \in \mathbb{R}^{m+1}$ such that $\tilde{\mathcal{A}}(X)_i = \langle A_i, X \rangle$ for $i = 0, \dots, m$, and $\mathcal{C} = \{X \succeq 0 : \tilde{\mathcal{A}}(X) = \tilde{\mathbf{b}}\}$. With the extended residue definition

$$\tilde{\mathbf{r}} = \tilde{\mathbf{r}}(U) = \tilde{\mathbf{r}}(UU^T) \triangleq \tilde{\mathcal{A}}(UU^T) - \tilde{\mathbf{b}}, \quad (11)$$

the associated penalty formulations are:

$$\underset{X \succeq 0}{\text{minimize}} \quad \tilde{F}_\mu(X) = \langle C, X \rangle + \mu \|\tilde{\mathbf{r}}(X)\|_2^2, \quad (12)$$

$$\underset{U \in \mathbb{R}^{n \times k}}{\text{minimize}} \quad \tilde{L}_\mu(U) = \langle C, UU^T \rangle + \mu \|\tilde{\mathbf{r}}(U)\|_2^2. \quad (13)$$

We note that, in full generality, finding (A_0, b_0) as in Proposition 11 may be as hard as solving an SDP, but in practical applications (A_0, b_0) may be easy to determine. In particular, SDPs with a trace constraint satisfy this with $A_0 = I_n$. For example, for the Max-Cut SDP, feasible matrices have constant trace n , so that $A_0 = I_n$ and $b_0 = n$ are suitable.

For this modified formulation, approximate FOSPs have bounded norm and bounded residues.

Lemma 12 *Consider problem (13) with $A_0 \succ 0$ and $b_0 \geq 0$. For any U ,*

$$\|\mathbf{r}\|_2 \leq \|\mathcal{A}\| \|U\|_F^2 + \|\mathbf{b}\|_2, \quad \text{and} \quad \|\tilde{\mathbf{r}}\|_2 \leq \|\tilde{\mathcal{A}}\| \|U\|_F^2 + \|\tilde{\mathbf{b}}\|_2.$$

If U is an ϵ -FOSP and $b_0 > 0$, then

$$\|U\|_F^2 \leq \max \left\{ \left(\frac{\epsilon}{2\mu b_0 \lambda_{\max}(A_0)} \right)^2, \frac{1}{\lambda_{\min}(A_0)^2} \left(\frac{\|C\|_2}{2\mu} + \frac{3}{2} b_0 \lambda_{\max}(A_0) \right) + \frac{\|\mathbf{b}\|_2}{2\lambda_{\min}(A_0)} \right\}.$$

We are now ready to state the main result by connecting Lemma 8 and Theorem 9 via Lemma 12. Let $B_0 \triangleq \|\tilde{\mathcal{A}}\| \max \left\{ \left(\frac{1}{4\mu b_0 \lambda_{\max}(A_0)} \right)^2, \frac{1}{\lambda_{\min}(A_0)^2} \left(\frac{\|C\|_2 + 3\sigma_G \sqrt{n}}{2\mu} + \frac{3}{2} b_0 \lambda_{\max}(A_0) \right) + \frac{\|\mathbf{b}\|_2}{2\lambda_{\min}(A_0)} \right\} + \|\tilde{\mathbf{b}}\|_2$.¹

Theorem 13 (Global optimality.) *Let \tilde{X} be a global optimum of (12). Let $\delta \in (0, 1)$ and c_0 be a universal constant. Draw a random matrix G with $G_{ij} \sim \mathcal{N}(0, \sigma_G^2)$ i.i.d. for $i \leq j$ and $G = G^T$. Let $U \in \mathbb{R}^{n \times k}$ be an (ϵ, γ) -SOSP of (13) with perturbed cost matrix $C + G$ and:*

$$\epsilon \leq \left(\frac{\gamma k^2 \sigma_G^2}{32 c_0 n \mu \|\mathcal{A}\|^2} \right)^{2/3} \quad \text{and} \quad k \geq 3 \left[\log \left(\frac{n}{\delta} \right) + \sqrt{\text{rank}(\mathcal{A}) \log \left(1 + \frac{8\mu B_0 \|\tilde{\mathcal{A}}\| \sqrt{c_0 n}}{\sigma_G} \right)} \right].$$

Then, with probability at least $1 - O(\delta)$ the optimality gap obeys:

$$\tilde{F}_\mu(UU^T) - \tilde{F}_\mu(\tilde{X}) \leq \gamma \sqrt{\epsilon} \text{Tr}(\tilde{X}) + \frac{1}{2} \epsilon \|U\|_F. \quad (14)$$

1. We pick σ_G first and then B_0 , k and ϵ .

This result shows that for compact SDPs (10), for $k = \tilde{\Omega}(\sqrt{m})$, with high probability upon the perturbation, all approximate SOSPs of the perturbed factorized problem are approximately optimal.

Notice that the result requires ϵ smaller than σ_G , which is limiting but unavoidable as there can be SDPs with bad approximate SOSPs. Hence, if we perturb by only a small amount (small σ_G), then we need to find highly accurate SOSPs to avoid these bad approximate SOSPs. Another way to look at the result is to see σ_G as a tentative distance from bad SDPs. Hence, for SDPs far away from these bad problems (higher σ_G), even high ϵ solutions are approximately (ϵ -) optimal.

Setting μ : From equation (14), to get residues of magnitude less than some ϵ (approximate feasibility) at an approximate SOSP U , we need μ to be $O\left(\frac{|\langle C, UU^T \rangle - \langle C, \tilde{X} \rangle| + \gamma\sqrt{\epsilon} \text{Tr}(\tilde{X}) + \frac{1}{2}\epsilon \|U\|_F}{\epsilon}\right)$. Using penalty formulations such as ALM can help us in getting similar approximate feasibility for much smaller values of μ . For general penalty problems we refer to Proposition 2.4 of (Bertsekas, 2014) which gives the relationship between the value of μ and the feasibility of a solution.

4.2. SDPs with positive definite cost

We now consider a second class of SDPs: ones where the cost matrix C is positive definite. The feasible set of these SDPs need not be compact. However, FOSPs for these SDPs are bounded, hence we will be able to show similar results as in Section 4.1. Consider the penalty formulation of the perturbed problem,

$$\underset{U \in \mathbb{R}^{n \times k}}{\text{minimize}} \quad \hat{L}_\mu(U) = \langle C + G, UU^T \rangle + \mu \sum_{i=1}^m (\langle A_i, UU^T \rangle - b_i)^2, \quad (15)$$

where G is a symmetric random matrix with $G_{ij} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma_G^2)$ for $i \leq j$. Let $\hat{F}_\mu(UU^T) = \hat{L}_\mu(U)$. To prove an optimality result for this problem, we first show a residue bound for any ϵ -FOSP of $\hat{L}_\mu(U)$.

Lemma 14 *Consider (15) with a positive definite cost matrix C . Let $\sigma_G \leq \frac{\lambda_{\min}(C)}{6\sqrt{n \log(n/\delta)}}$. Then, with probability at least $1 - \delta$, at any ϵ -FOSP U of (15), the residue obeys:*

$$\|\mathbf{r}\|_2 = \|\mathcal{A}(UU^T) - \mathbf{b}\|_2 \leq \|\mathcal{A}\| \max \left\{ \left(\frac{2\epsilon}{\lambda_{\min}(C)} \right)^2, \frac{2\mu}{\lambda_{\min}(C)} \|\mathbf{b}\|_2^2 \right\} + \|\mathbf{b}\|_2.$$

Using this, we get the following result from Lemma 8 and Theorem 9 along same lines as that of Theorem 13. Let

$$B_0 \triangleq \|\mathcal{A}\| \max \left\{ \left(\frac{1}{\lambda_{\min}(C)} \right)^2, \frac{2\mu}{\lambda_{\min}(C)} \|\mathbf{b}\|_2^2 \right\} + \|\mathbf{b}\|_2.$$

Theorem 15 (Global optimality.) *Let $\delta \in (0, 1)$ and c_0 be a universal constant. Given an SDP (1) with positive definite objective matrix C , let \tilde{X} be a global optimum of the perturbed problem (15), and let $\sigma_G \leq \frac{\lambda_{\min}(C)}{4\sqrt{n \log(n/\delta)}}$. Let U be an (ϵ, γ) -SOSP of the perturbed problem (15) with:*

$$\epsilon \leq \left(\frac{\gamma k^2 \sigma_G^2}{32c_0 n \mu \|\mathcal{A}\|^2} \right)^{2/3} \quad \text{and} \quad k \geq 3 \left[\log \left(\frac{n}{\delta} \right) + \sqrt{\text{rank}(\mathcal{A}) \log \left(1 + \frac{8\mu B_0 \|\mathcal{A}\| \sqrt{c_0 n}}{\sigma_G} \right)} \right].$$

Then, with probability at least $1 - O(\delta)$,

$$\widehat{F}_\mu(UU^T) - \widehat{F}_\mu(\tilde{X}) \leq \gamma\sqrt{\epsilon} \text{Tr}(\tilde{X}) + \frac{1}{2}\epsilon\|U\|_F.$$

This result shows that even though the feasible set of SDP is not compact, as long as the objective is positive definite, all approximate SOSPs of the perturbed objective are approximately optimal (with high probability upon the perturbation). Without the positive definite condition, SDPs can have unbounded solutions (see Section 2.4 of [Gärtner and Matousek \(2012\)](#)). We also require a bound on the magnitude of the perturbation (σ_G), as otherwise the objective ($C + G$) can be indefinite with (too) high probability, which may result in unbounded solutions.

5. Applications

We present applications of our results to two SDPs: Max-Cut (deferred to Section A) and matrix completion, both of which are important problems in the learning domain and have been studied extensively. Interest has grown to develop efficient solvers for these SDPs ([Arora and Kale, 2007](#); [Mei et al., 2017](#); [Hardt, 2013](#); [Bandeira et al., 2016](#)).

This work differs from previous efforts in at least two ways. First, we aim to demonstrate that Burer–Monteiro-style approaches, which are often used in practice, can indeed lead to provably efficient algorithms for general SDPs. We believe that building upon this work, it should be possible to improve the time-complexity guarantees of such factorization-based algorithms. Second, we note that several problems formulated as SDPs in fact necessitate low-rank solutions, for example because of memory concerns (as is the case in matrix completion), and factorization approaches provide a natural means to control rank.

5.1. Matrix Completion

In this section we specialize our results for the matrix completion problem [Candès and Recht \(2009\)](#). The goal of a matrix completion problem is to find a low-rank matrix M using only a small number of its entries, with applications in recommender systems. To ensure that the computed matrix is low-rank and generalizes well, one typically imposes nuclear-norm regularization which leads to the following SDP:

$$\begin{aligned} \min \quad & \text{Tr}(W_1) + \text{Tr}(W_2) \quad \equiv \quad \min \quad \langle I, Z \rangle \\ \text{s. t.} \quad & X_{ij} = M_{ij}, (i, j) \in \mathcal{S} \quad \text{s. t.} \quad \frac{1}{2} \langle e_{i+n}e_{j+n}^T + e_{j+n}e_{i+n}^T, Z \rangle = M_{ij}, (i, j) \in \mathcal{S} \\ & \begin{bmatrix} W_1 & X \\ X^T & W_2 \end{bmatrix} \succeq 0 \quad \quad \quad Z \succeq 0. \end{aligned}$$

Here \mathcal{S} is the set of observed indices of M and $Z \triangleq \begin{bmatrix} W_1 & X \\ X^T & W_2 \end{bmatrix}$. Let

$$\widehat{L}_\mu(U) = \langle I + G, UU^T \rangle + \mu \sum_{i=1}^m \left(\frac{1}{2} \langle e_{i+n}e_{j+n}^T + e_{j+n}e_{i+n}^T, UU^T \rangle - M_{ij} \right)^2 \quad (16)$$

be the corresponding penalty objective. Let $\widehat{F}_\mu(UU^T) = \widehat{L}_\mu(U)$. The objective is positive definite with $\lambda_1(C) = \lambda_n(C) = 1$. Also, since \mathcal{A} is a sub-sampling operator, $\|\mathcal{A}\| \leq 1$. Finally, for

$\epsilon^2 \leq \frac{\mu}{2} \sqrt{\sum_{(i,j) \in \mathcal{S}} M_{ij}^2}$, the residues are bounded by:

$$\|\mathcal{A}\| \max \left\{ \left(\frac{2\epsilon}{\lambda_n(C)} \right)^2, \frac{2\mu}{\lambda_n(C)} \|\mathbf{b}\|_2^2 \right\} + \|\mathbf{b}\|_2 \leq \max 3\mu \sqrt{\sum_{(i,j) \in \mathcal{S}} M_{ij}^2}.$$

Applying Theorem 15 for this setting gives the following corollary.

Corollary 16 *There exists an absolute numerical constant c_2 such that the following holds. With probability greater than $1 - \delta$, every (ϵ, γ) -SOSP U of the perturbed matrix completion problem $\hat{L}_\mu(U)$ (16) with:*

$$\sigma_G \leq \frac{1}{4\sqrt{n \log(n/\delta)}}, \quad \epsilon \leq \frac{1}{c_2} \left(\frac{\gamma |\mathcal{S}| \sigma_G^2}{n\mu} \right)^{2/3}, \quad \text{and} \quad k = \tilde{\Omega} \left(\sqrt{|\mathcal{S}| \log \left(\frac{\mu^2 \sqrt{n} \sqrt{\sum_{(i,j) \in \mathcal{S}} M_{ij}^2}}{\sigma_G} \right)} \right),$$

satisfies $\hat{F}_\mu(UU^T) - \hat{F}_\mu(X^*) \leq \gamma\sqrt{\epsilon} \text{Tr}(X^*) + \frac{1}{2}\epsilon \|U\|_F$, where X^* is a global optimum of $\hat{F}_\mu(X)$.

This result shows that for the matrix completion problem with m observations, for rank $\tilde{\Omega}(\sqrt{m})$, with high probability upon perturbation, any approximate local minimum of the factorized and penalized problem is an approximate global minimum.

Most of the existing results on matrix completion either require strong distribution assumptions on \mathcal{S} and incoherence assumptions on M to recover a low-rank solution (Candès and Recht, 2009; Jain et al., 2013). The standard nuclear norm minimization algorithms are not guaranteed to converge to low-rank solutions without these assumptions, which implies that the entire matrix would need to be stored for prediction which is infeasible in practice. Similarly, generalization error bounds (Foygel and Srebro, 2011) as well as differential privacy guarantees depend on recovery of a low-rank solution.

Our result guarantees finding a solution of rank $\tilde{\Omega}(\sqrt{m})$ without any statistical assumptions on the sampling or the matrix, unlike existing results (Ge et al., 2016), though it involves a random perturbation. The tradeoff is our results does not guarantee finding a lower (potentially a constant) rank solution, even if one exists for a given problem.

6. Conclusions and perspectives

In this paper we considered the Burer–Monteiro factorization to solve SDPs (2) in the penalty form which allows us to find solutions of rank at most k without doing any eigen-decomposition. We established that with a small perturbation to the objective C , with high probability, all approximate local minima are approximately globally optimal for the SDP, provided $k = \tilde{\Omega}(\sqrt{m})$ (which is the right order though constants and dependence on other parameters could certainly be improved). This is achieved through smoothed analysis, which we believe is an appropriate tool to deal with the pathological cases exhibited in Section 3. A natural direction of improvement for the present work is to tackle the more complex ALM formulations, which is expected to help in satisfying constraints more accurately.

Finally, we studied the applicability of our results to two applications: Max-Cut (in the appendix) and matrix completion. While these particularizations do not always improve over the specialized solvers for these problems, we believe that the work done here in studying low-rank parameterization of SDPs will be a helpful step towards building up to faster methods.

Acknowledgments

NB thanks Dustin Mixon for many interesting conversations on the applicability of smoothed analysis to low-rank SDPs. NB was supported in part by NSF award DMS-1719558. SB thanks Alex Dimakis for pointing to the book by Gärtner and Matousek (2012) on SDPs.

References

- Emmanuel Abbe. Community detection and stochastic block models: recent developments. *arXiv preprint arXiv:1703.10146*, 2017.
- Alekh Agarwal, Sahand Negahban, and Martin J Wainwright. Fast global convergence rates of gradient methods for high-dimensional statistical recovery. In *Advances in Neural Information Processing Systems*, pages 37–45, 2010.
- Farid Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5(1):13–51, 1995.
- Kurt M Anstreicher. The volumetric barrier for semidefinite programming. *Mathematics of Operations Research*, 25(3):365–380, 2000.
- Sanjeev Arora and Satyen Kale. A combinatorial, primal-dual approach to semidefinite programs. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 227–236. ACM, 2007.
- Sanjeev Arora, Elad Hazan, and Satyen Kale. Fast algorithms for approximate semidefinite programming using the multiplicative weights update method. In *Foundations of Computer Science, 2005. FOCS 2005. 46th Annual IEEE Symposium on*, pages 339–348. IEEE, 2005.
- A.S. Bandeira, N. Boumal, and V. Voroninski. On the low-rank approach for semidefinite programs arising in synchronization and community detection. In *Proceedings of The 29th Conference on Learning Theory, COLT 2016, New York, NY, June 23–26, 2016*.
- Boaz Barak, Jonathan A Kelner, and David Steurer. Dictionary learning and tensor decomposition via the sum-of-squares method. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 143–151. ACM, 2015.
- A.I. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete & Computational Geometry*, 13(1):189–202, 1995. doi: 10.1007/BF02574037.
- Dimitri P Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.
- R. Bhatia. *Matrix analysis*, volume 169. Springer Science & Business Media, 2013.
- 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.

- N. Boumal, P.-A. Absil, and C. Cartis. Global rates of convergence for nonconvex optimization on manifolds. *IMA Journal of Numerical Analysis*, 2018a. doi: 10.1093/imanum/drx080.
- N. Boumal, V. Voroninski, and A.S. Bandeira. Deterministic guarantees for Burer–Monteiro factorizations of smooth semidefinite programs. *arXiv preprint arXiv:1804.02008*, 2018b.
- Nicolas Boumal, Vlad Voroninski, and Afonso Bandeira. The non-convex burer-monteiro approach works on smooth semidefinite programs. In *Advances in Neural Information Processing Systems*, pages 2757–2765, 2016.
- Samuel Burer and Renato DC Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003.
- Samuel Burer and Renato DC Monteiro. Local minima and convergence in low-rank semidefinite programming. *Mathematical Programming*, 103(3):427–444, 2005.
- Emmanuel J Candès and Benjamin Recht. Exact matrix completion via convex optimization. *Foundations of Computational mathematics*, 9(6):717–772, 2009.
- Coralie 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.
- Simon S Du and Jason D Lee. On the power of over-parametrization in neural networks with quadratic activation. *arXiv preprint arXiv:1803.01206*, 2018.
- 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.
- Rina Foygel and Nathan Srebro. Concentration-based guarantees for low-rank matrix reconstruction. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 315–340, 2011.
- Dan Garber. Faster projection-free convex optimization over the spectrahedron. In *Advances in Neural Information Processing Systems*, pages 874–882, 2016.
- Dan Garber and Elad Hazan. Sublinear time algorithms for approximate semidefinite programming. *Mathematical Programming*, 158(1-2):329–361, 2016.
- Bernd Gärtner and Jiri Matousek. *Approximation algorithms and semidefinite programming*. Springer Science & Business Media, 2012.
- Rong Ge and Tengyu Ma. On the optimization landscape of tensor decompositions. In *Advances in Neural Information Processing Systems*, pages 3656–3666, 2017.
- Rong Ge, Furong Huang, Chi Jin, and Yang Yuan. Escaping from saddle points—online stochastic gradient for tensor decomposition. In *Proceedings of The 28th 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.
- Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6): 1115–1145, 1995.
- Moritz Hardt. Understanding alternating minimization for matrix completion. *arXiv preprint arXiv:1312.0925*, 2013.
- Elad Hazan. Sparse approximate solutions to semidefinite programs. In *LATIN 2008: Theoretical Informatics*, pages 306–316. Springer, 2008.
- U. Helmke and M.A. Shayman. Critical points of matrix least squares distance functions. *Linear Algebra and its Applications*, 215:1–19, 1995. doi: [https://doi.org/10.1016/0024-3795\(93\)00070-G](https://doi.org/10.1016/0024-3795(93)00070-G).
- Prateek Jain, Praneeth Netrapalli, and Sujay Sanghavi. Low-rank matrix completion using alternating minimization. In *Proceedings of the 45th 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. *arXiv preprint arXiv:1703.00887*, 2017a.
- Chi Jin, Praneeth Netrapalli, and Michael I Jordan. Accelerated gradient descent escapes saddle points faster than gradient descent. *arXiv preprint arXiv:1711.10456*, 2017b.
- Michel Journée, Francis Bach, P-A Absil, and Rodolphe Sepulchre. Low-rank optimization on the cone of positive semidefinite matrices. *SIAM Journal on Optimization*, 20(5):2327–2351, 2010.
- Raghuveer H Keshavan, Andrea Montanari, and Sewoong Oh. Matrix completion from a few entries. *Information Theory, IEEE Transactions on*, 56(6):2980–2998, 2010.
- Kartik Krishnan and John E Mitchell. Properties of a cutting plane method for semidefinite programming. 2003.
- Gert RG Lanckriet, Nello Cristianini, Peter Bartlett, Laurent El Ghaoui, and Michael I Jordan. Learning the kernel matrix with semidefinite programming. *Journal of Machine learning research*, 5(Jan):27–72, 2004.
- Song Mei, Theodor Misiakiewicz, Andrea Montanari, and Roberto Imbuzeiro Oliveira. Solving sdps for synchronization and maxcut problems via the grothendieck inequality. In Satyen Kale and Ohad Shamir, editors, *Proceedings of the 2017 Conference on Learning Theory*, volume 65 of *Proceedings of Machine Learning Research*, pages 1476–1515, Amsterdam, Netherlands, 07–10 Jul 2017. PMLR.
- Balas Kausik Natarajan. Sparse approximate solutions to linear systems. *SIAM journal on computing*, 24(2):227–234, 1995.

- Sahand Negahban and Martin J Wainwright. Restricted strong convexity and weighted matrix completion: Optimal bounds with noise. *The Journal of Machine Learning Research*, 13(1): 1665–1697, 2012.
- Yurii Nesterov and Arkadi Nemirovski. Polynomial barrier methods in convex programming. *Ekonom. i Mat. Metody*, 24(6):1084–1091, 1988.
- Yurii Nesterov and Arkadi Nemirovski. *Self-concordant functions and polynomial-time methods in convex programming*. USSR Academy of Sciences, Central Economic & Mathematic Institute, 1989.
- Yurii Nesterov and Boris T Polyak. Cubic regularization of newton method and its global performance. *Mathematical Programming*, 108(1):177–205, 2006.
- Yurii Nesterov, Arkadi Nemirovski, and Yinyu Ye. *Interior-point polynomial algorithms in convex programming*, volume 13. SIAM, 1994.
- Hoi H Nguyen. Random matrices: repulsion in spectrum. *arXiv preprint arXiv:1709.06682*, 2017.
- 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.
- G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of operations research*, 23(2):339–358, 1998. doi: 10.1287/moor.23.2.339.
- 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.
- Jianbo Shi and Jitendra Malik. Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, 22(8):888–905, 2000.
- Ju Sun, Qing Qu, and John Wright. A geometric analysis of phase retrieval. *preprint arXiv:1602.06664*, 2016.
- Ruoyu Sun and Zhi-Quan Luo. Guaranteed matrix completion via non-convex factorization. *arXiv preprint arXiv:1411.8003*, 2014.
- Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM review*, 38(1):49–95, 1996.
- R. Vershynin. *High-dimensional probability*, 2016.
- H. Wolkowicz. Some applications of optimization in matrix theory. *Linear algebra and its applications*, 40:101–118, 1981.
- Alp Yurtsever, Madeleine Udell, Joel Tropp, and Volkan Cevher. Sketchy decisions: Convex low-rank matrix optimization with optimal storage. In *Artificial Intelligence and Statistics*, pages 1188–1196, 2017.

Zhihui Zhu, Qiuwei Li, Gongguo Tang, and Michael B Wakin. Global optimality in low-rank matrix optimization. *arXiv preprint arXiv:1702.07945*, 2017.

Appendix A. Max-Cut

In this section we present our results for the Max-Cut problem which finds applications in clustering related problems. In a seminal paper, [Goemans and Williamson \(1995\)](#) defined the following SDP to solve the Max-Cut problem: $\min_{X \in \mathbb{R}^{n \times n}} \langle C, X \rangle$, s.t. $X_{ii} = 1 \forall 1 \leq i \leq n$, $X \succeq 0$, where n is the number of vertices in the given graph and C is its adjacency matrix. Since the constraint set also satisfies $\text{Tr}(X) = n$, we consider the following penalized, non-convex version of the problem.

$$\hat{L}_\mu(U) \triangleq \langle C + G, UU^\top \rangle + \mu \left(\left(\langle I, UU^\top \rangle - n \right)^2 + \sum_{i=1}^n \left(\langle e_i e_i^\top, UU^\top \rangle - 1 \right)^2 \right), \quad (17)$$

where G is a random symmetric Gaussian matrix. Let $\hat{F}_\mu(UU^\top) = \hat{L}_\mu(U)$. After some simplifying computations, we have the following corollary of Theorem 13.

Corollary 17 *There exists an absolute numerical constant c_1 such that the following holds. With probability greater than $1 - \delta$, every (ϵ, γ) -SOSP U of the perturbed Max-Cut problem $\hat{L}_\mu(U)$ (17) with:*

$$\epsilon \leq \frac{1}{c_1} \left(\frac{\gamma \sigma_G^2}{\mu n} \right)^{2/3}, \quad \text{and} \quad k = \tilde{\Omega} \left(\sqrt{n \log \left(\frac{\mu^2 \sqrt{n}}{\sigma_G} \right)} \right),$$

satisfies $\hat{F}_\mu(UU^\top) - \hat{F}_\mu(X^) \leq \gamma \sqrt{\epsilon} \text{Tr}(X^*) + \frac{1}{2} \epsilon \|U\|_F$, where X^* is a global optimum of $\hat{F}_\mu(X)$.*

The above result states that for the penalized version of the perturbed Max-Cut SDP, the Burer-Monteiro approach finds an approximate global optimum as soon as the factorization rank $k = \tilde{\Omega}(\sqrt{n})$ (with high probability). Existing results for Max-Cut using this approach either only handle exact SOSPs ([Boumal et al., 2016](#)), or require $k = n + 1$ ([Boumal et al., 2018a](#)), or require k that is dependent on $\frac{1}{\epsilon}$ ([Mei et al., 2017](#)). Moreover, complexity per iteration scales only linearly with the number of edges in the graph.

Appendix B. Gradient Descent

In previous sections we have seen that for the perturbed penalty objective (15), under some technical conditions on the SDP, with high probability upon appropriate choice of the parameters, every approximate SOSP is approximately optimal. Second-order methods such as cubic regularization and trust regions ([Nesterov and Polyak, 2006](#); [Cartis et al., 2012](#)) converge to an approximate SOSP in polynomial time. While gradient descent with random initialization can take exponential time to converge to an SOSP ([Du et al., 2017](#)), a recent line of work starting with [Ge et al. \(2015\)](#) has established that perturbed gradient descent (PGD)² converges to an SOSP as efficiently as second-order methods in the worst case, with high probability. In particular we have the following *almost dimension free* convergence rate for PGD from ([Jin et al., 2017a](#)).

2. This is vanilla gradient descent but with additional random noise added to the updates when the gradient magnitude becomes smaller than a threshold.

Theorem 18 (Theorem 3 of Jin et al. (2017a)) *Let f be l -smooth (that is, its gradient is l -Lipschitz) and have a ρ -Lipschitz Hessian. There exists an absolute constant c_{\max} such that, for any $\delta \in (0, 1)$, $\epsilon \leq \frac{l^2}{\rho}$, $\Delta_f \geq f(X_0) - f^*$, and constant $c \leq c_{\max}$, $PGD(X_0, l, \rho, \epsilon, c, \delta, \Delta_f)$ applied to the cost function f outputs a (ρ^2, ϵ) SOSp with probability at least $1 - \delta$ in*

$$O\left(\frac{(f(X_0) - f^*)l}{\epsilon^2} \log^4\left(\frac{nk l \Delta_f}{\epsilon^2 \delta}\right)\right)$$

iterations.

The above theorem requires the function f to be smooth and Hessian-Lipschitz. The next lemma states that the perturbed penalty objective (15) satisfies these requirements—proof in Appendix G.

Lemma 19 *In the region $\{U \in \mathbb{R}^{n \times k} : \|U\|_F \leq \tau\}$ for some $\tau > 0$, the cost function $\hat{L}_\mu(U)$ in (15) is l -smooth and its Hessian is ρ -Lipschitz with:*

- $l \leq 2\|C + G\|_2 + 4\mu\|\mathcal{A}\|\|\mathbf{b}\|_2 + 12\mu\tau^2\|\mathcal{A}\|^2$, and
- $\rho \leq 16\mu\tau\|\mathcal{A}\|^2$.

Here, $\|\mathcal{A}\|$ is as defined in (8). Notice furthermore that, with high probability, $\|G\|_2 \leq 3\sigma_G\sqrt{n}$. In that event, $\|C + G\|_2 \leq \|C\|_2 + 3\sigma_G\sqrt{n}$.

Combining this lemma with the above theorem shows that the perturbed gradient method converges to an (ϵ, ρ^2) SOSp in $\tilde{O}(\frac{1}{\epsilon^2})$ steps (ignoring all other problem parameters). This can be improved to $\tilde{O}(\frac{1}{\epsilon^{1.75}})$ using a variant of Nesterov’s accelerated gradient descent (Jin et al., 2017b). Moreover, if the objective function is (restricted) strongly convex in the vicinity of the local minimum, then we can further improve the rates to $\text{poly log}(\frac{1}{\epsilon})$ (Jin et al., 2017a). This property is satisfied for problems where \mathcal{A} meets either restricted isometry conditions or when \mathcal{A} pertains to a uniform sampling of incoherent matrices (Agarwal et al., 2010; Negahban and Wainwright, 2012; Sun and Luo, 2014). See (Bhojanapalli et al., 2016a) for more discussions on restricted strong convexity close to the global optimum.

The complexity of the algorithm is given by Gradient-Computation-Time \times Number of iterations. Computing the gradient in each iteration requires $O(Zk + nk^2 + mnk)$ arithmetic operations where Z is the number of non-zeros in C and the constraint matrices. For dense problems this becomes $O(mn^2k)$. However, most practical problems tend to have a certain degree of sparsity in the constraint matrices so that the computational complexity of such a method can be significantly smaller than the worst-case bound.

Appendix C. Proof of Lemma 21: lower-bound for smallest singular values

First we state a special case of Corollary 1.17 from (Nguyen, 2017). Let $N_I(X)$, denote the number of eigenvalues of X in the interval I .

Corollary 20 *Let M' be a deterministic symmetric matrix in $\mathcal{S}^{n \times n}$. Let G' be a random symmetric matrix with entries G'_{ij} drawn i.i.d. from $\mathcal{N}(0, 1)$ for $i \geq j$ (in particular, independent of M' .) Then, for given $0 < \gamma < 1$, there exists a constant $c = c(\gamma)$ such that for any $\epsilon > 0$ and $k \geq 1$, with I being the interval, $[-\frac{\epsilon k}{\sqrt{n}}, \frac{\epsilon k}{\sqrt{n}}]$,*

$$\Pr[N_I(M' + G') \geq k] \leq n^k \left(\frac{c\epsilon}{\sqrt{2\pi}}\right)^{(1-\gamma)k^2/2}.$$

We can use the above corollary to prove Lemma 21.

Proof In our case, entries of G have variance σ_G^2 . Thus, set $G = \sigma_G G'$, and set $\bar{M} = \sigma_G M'$. From Corollary 20, we get

$$N_{\sigma_G I}(\bar{M} + G) = N_I(M' + G') < k$$

with probability at least $1 - n^k \left(\frac{c\epsilon}{\sqrt{2\pi}} \right)^{(1-\gamma)k^2/2}$. In this event, $\sigma_{n-(k-1)}(\bar{M} + G) \geq \frac{\epsilon k}{\sqrt{n}} \sigma_G$. Choose $\gamma = \frac{1}{2}$, and $\epsilon = \frac{1}{2c}$. Substituting this we get with probability at least $1 - \exp\left(-\frac{k^2}{8} \log(8\pi) + k \log(n)\right)$ that

$$\sigma_{n-(k-1)}(\bar{M} + G) \geq \frac{k}{2c\sqrt{n}} \sigma_G.$$

Hence, $\sum_{i=1}^k \sigma_{n-(i-1)}(\bar{M} + G)^2 \geq \sigma_{n-(k-1)}(\bar{M} + G)^2 \geq \frac{k^2}{c_0 n} \sigma_G^2$, for some absolute constant $c_0 = 4c^2$. ■

Appendix D. Proofs for Section 2

Proof [Proof of Lemma 1] Necessary and sufficient optimality conditions for (5) are: $\nabla f(X) \succeq 0$ and $\nabla f(X)X = 0$. Let U be an SOSF for (6) with $\text{rank}(U) < k$ and define $X = UU^T$. Then, $\nabla g(U) = 2\nabla f(UU^T)U = 0$ and $\nabla^2 g(U) \succeq 0$. The first statement readily shows that $\nabla f(X)X = 0$. The Hessians of f and g are related by:

$$\frac{1}{2} \nabla^2 g(U)[\dot{U}] = \nabla f(UU^T)\dot{U} + \nabla^2 f(UU^T)[U\dot{U}^T + \dot{U}U^T]U.$$

Since $\text{rank}(U) < k$, there exists a vector $z \in \mathbb{R}^k$ such that $Uz = 0$ and $\|z\|_2 = 1$. For any $x \in \mathbb{R}^n$, set $\dot{U} = xz^T$ so that $U\dot{U}^T + \dot{U}U^T = 0$. Using second-order stationarity of U , we find:

$$0 \leq \frac{1}{2} \left\langle \dot{U}, \nabla^2 g(U)[\dot{U}] \right\rangle = \langle xz^T, \nabla f(UU^T)xz^T \rangle = x^T \nabla f(UU^T)x.$$

This holds for all $x \in \mathbb{R}^n$, hence $\nabla f(UU^T) \succeq 0$ and $X = UU^T$ is optimal for (5). Since (5) is a relaxation of (6), it follows that U is optimal for (6). ■

Proof [Proof of Lemma 3] Let U be any FOSP of (4) and consider the linear operator $\mathcal{A}: \mathcal{S}^{n \times n} \rightarrow \mathbb{R}^m$ defined by $\mathcal{A}(X)_i = \langle A_i, X \rangle$. By first-order stationarity, we have:

$$\nabla L_\mu(U) = 2(C + 2\mu\mathcal{A}^*(\mathcal{A}(UU^T) - b))U = 0.$$

Hence, the nullity of $C + 2\mu\mathcal{A}^*(\mathcal{A}(UU^T) - b)$ (the dimension of its kernel) satisfies:

$$\text{rank}(U) \leq \text{null}(C + 2\mu\mathcal{A}^*(\mathcal{A}(UU^T) - b)) \leq \max_{y \in \mathbb{R}^m} \text{null}(C + \mathcal{A}^*(y)). \quad (18)$$

The maximum over y is indeed attained since the function null takes integer values in $0, \dots, n$. Say the maximum evaluates to ℓ . Then, for some y , $M \triangleq C + \mathcal{A}^*(y)$ has nullity ℓ . Hence,

$$C = M - \mathcal{A}^*(y) \in \mathcal{N}_\ell + \text{im } \mathcal{A}^*,$$

where \mathcal{N}_ℓ is the manifold of symmetric matrices of size n and nullity ℓ , $\text{im } \mathcal{A}^*$ is the range of \mathcal{A}^* and the plus is a set-sum. More generally, assuming the maximum in (18) is p or more, then

$$C \in \mathcal{M}_p \triangleq \bigcup_{\ell=p, \dots, n} \mathcal{N}_\ell + \text{im } \mathcal{A}^*.$$

The manifold \mathcal{N}_ℓ has dimension $\frac{n(n+1)}{2} - \frac{\ell(\ell+1)}{2}$ (Helmke and Shayman, 1995, Prop. 2.1(i)), while $\text{im } \mathcal{A}^*$ has dimension at most m . Hence,

$$\dim \mathcal{M}_p \leq m + \max_{\ell=p, \dots, n} \dim \mathcal{N}_\ell = m + \frac{n(n+1)}{2} - \frac{p(p+1)}{2}.$$

Since C is in $\mathcal{S}^{n \times n}$ and $\dim \mathcal{S}^{n \times n} = \frac{n(n+1)}{2}$, almost no C lives in \mathcal{M}_p if $\dim \mathcal{M}_p < \dim \mathcal{S}^{n \times n}$, which is the case if $\frac{p(p+1)}{2} > m$. Stated differently: $\text{rank}(U) \leq p$, and for almost all $C \in \mathcal{S}^{n \times n}$, $\frac{p(p+1)}{2} \leq m$. To conclude, require that k is strictly larger than any p which satisfies $\frac{p(p+1)}{2} \leq m$. ■

Appendix E. Proofs for Section 3

Proof [Proof of Theorem 5] We first show that the SDP admits exactly one feasible point. Indeed, let $X \succeq 0$ be feasible for the SDP. Then, constraints n and $n+1$ imply $\langle A_{n+1} - A_n, X \rangle = 0$. That is, the trace of the principal submatrix of size $n-1$ of X is zero. Since this submatrix is also positive semidefinite, it is zero. Constraints 1 to $n-1$ further show that all entries but X_{nn} are zero. Finally, constraints n and $n+1$ force $X_{nn} = \frac{5(n-1)}{3}$. This X has rank 1 and is necessarily optimal.

We now show that the proposed \bar{U} is suboptimal for L . To this end, build $\tilde{U} \in \mathbb{R}^{n \times k}$ with the last row having squared 2-norm equal to $\frac{5(n-1)}{3}$, and all other rows are zero. Clearly, $\tilde{U}\tilde{U}^T$ is feasible for the SDP, so that $L(\tilde{U}) = 0$: this is optimal. On the other hand, $L(\bar{U}) = \frac{5}{18}(n-1)^2\epsilon^2 > L(\tilde{U})$.

Finally, we check stationarity of \bar{U} . Let $\mathcal{A}: \mathcal{S}^{n \times n} \rightarrow \mathbb{R}^m$ be the linear operator such that $\mathcal{A}(X)_i = \langle A_i, X \rangle$, and define the residue function $\mathbf{r}(U) = \mathcal{A}(UU^T) - b$. The cost function and its derivatives take the following forms:

$$\begin{aligned} L(U) &= \frac{1}{2} \|\mathbf{r}(U)\|_2^2, \\ \nabla L(U) &= 2\mathcal{A}^*(\mathbf{r}(U))U, \\ \nabla^2 L(U)[\dot{U}] &= 2\mathcal{A}^*(\mathbf{r}(U))\dot{U} + 2\mathcal{A}^*(\mathcal{A}(U\dot{U}^T + \dot{U}U^T))U. \end{aligned}$$

Simple computations show that $\mathcal{A}(\bar{U}\bar{U}^T) = (0, \dots, 0, (n-1)\epsilon, 2(n-1)\epsilon)^T$, so that $\mathcal{A}^*(\mathbf{r}(\bar{U})) = -\frac{n-1}{3}\epsilon^2 \cdot e_n e_n^T$: only the bottom-right entry is non-zero. Consequently, $\nabla L(\bar{U}) = 0$: \bar{U} is an FOSP. To show second-order stationarity, we must also show that $\nabla^2 L(\bar{U})$ is positive semidefinite. That is, we must show the inequalities:

$$0 \leq \langle \dot{U}, \nabla^2 L(\bar{U})[\dot{U}] \rangle = 2 \langle \dot{U}\dot{U}^T, \mathcal{A}^*(\mathbf{r}(\bar{U})) \rangle + \|\mathcal{A}(\bar{U}\dot{U}^T + \dot{U}\bar{U}^T)\|_2^2$$

for all $\dot{U} \in \mathbb{R}^{n \times k}$. Let

$$\dot{U} = \begin{bmatrix} - & \dot{u}_1^T & - \\ & \vdots & \\ - & \dot{u}_n^T & - \end{bmatrix}, \quad \text{with } \dot{u}_1, \dots, \dot{u}_n \in \mathbb{R}^k \text{ arbitrary.}$$

Then, $\mathcal{A}(\bar{U}\dot{U}^T + \dot{U}\bar{U}^T) = (2\dot{u}_n^T, q_1, q_2)^T$ for some values q_1, q_2 , so that:

$$\left\langle \dot{U}, \nabla^2 L(\bar{U})[\dot{U}] \right\rangle = -2\frac{n-1}{3}\epsilon^2 \|\dot{u}_n\|_2^2 + 4\|\dot{u}_n\|_2^2 + q_1^2 + q_2^2 \geq \left(4 - 2\frac{n-1}{3}\epsilon^2\right) \|\dot{u}_n\|_2^2.$$

Under our condition on ϵ , this is indeed always nonnegative: \bar{U} is an SOSp. \blacksquare

Appendix F. Proofs for Section 4

Proof [Proof of Lemma 8] The gradient and Hessian of L_μ (4), with $\mathbf{r} \triangleq \mathbf{r}(U) = \mathcal{A}(UU^T) - b$, are:

$$\nabla L_\mu(U) = 2(C + 2\mu\mathcal{A}^*(\mathbf{r}))U, \quad (19)$$

$$\nabla^2 L_\mu(U)[\dot{U}] = 2(C + 2\mu\mathcal{A}^*(\mathbf{r}))\dot{U} + 4\mu\mathcal{A}^*(\mathcal{A}(\dot{U}U^T + U\dot{U}^T))U. \quad (20)$$

Since U is an (ϵ, γ) -SOSP, it holds for all $\dot{U} \in \mathbb{R}^{n \times k}$ with $\|\dot{U}\|_F = 1$ that:

$$-\frac{\gamma\sqrt{\epsilon}}{2} \leq \frac{1}{2} \left\langle \dot{U}, \nabla^2 L_\mu(U)[\dot{U}] \right\rangle = \left\langle C + 2\mu\mathcal{A}^*(\mathbf{r}), \dot{U}\dot{U}^T \right\rangle + \mu \left\| \mathcal{A}(\dot{U}U^T + U\dot{U}^T) \right\|_2^2. \quad (21)$$

We now construct specific \dot{U} 's to exploit the fact that U is almost rank deficient. Let $z \in \mathbb{R}^k$ be a right singular vector of U such that $\|Uz\|_2 = \sigma_k(U)$ (that is, z is associated to the least singular value of U and $\|z\|_2 = 1$.) For any $x \in \mathbb{R}^n$ with $\|x\|_2 = 1$, introduce $\dot{U} = xz^T$ in (21):

$$-\frac{\gamma\sqrt{\epsilon}}{2} \leq x^T(C + 2\mu\mathcal{A}^*(\mathbf{r}))x + \mu \left\| \mathcal{A}(\dot{U}U^T + U\dot{U}^T) \right\|_2^2.$$

The last term is easily controlled:

$$\left\| \mathcal{A}(\dot{U}U^T + U\dot{U}^T) \right\|_2 \leq 2\|\mathcal{A}\| \|\dot{U}\|_F = 2\|\mathcal{A}\| \|Uz\|_2 = 2\|\mathcal{A}\| \|Uz\|_2 \|x\|_2 = 2\|\mathcal{A}\| \sigma_k(U).$$

Let x be an eigenvector of $C + 2\mu\mathcal{A}^*(\mathbf{r})$ associated to its least eigenvalue and combine the last two statements together with the assumption on $\sigma_k(U)$ to find:

$$\lambda_{\min}(C + 2\mu\mathcal{A}^*(\mathbf{r})) \geq -\frac{\gamma\sqrt{\epsilon}}{2} - 4\mu\|\mathcal{A}\|^2 \sigma_k^2(U) \geq -\gamma\sqrt{\epsilon}. \quad (22)$$

This inequality is key to bound the optimality gap. For this part, we rely on the fact that $L_\mu(U) = F_\mu(UU^T)$ and F_μ is convex on $\mathcal{S}^{n \times n}$ (3). Specifically, let \tilde{X} be a global optimum for F_μ (assuming it exists), and set $X = UU^T$. Then, $\nabla F_\mu(X) = C + 2\mu\mathcal{A}^*(\mathbf{r})$, $\nabla L_\mu(U) = 2\nabla F_\mu(X)U$ and:

$$\begin{aligned} F_\mu(\tilde{X}) - F_\mu(X) &\geq \left\langle \nabla F_\mu(X), \tilde{X} - X \right\rangle = \left\langle C + 2\mu\mathcal{A}^*(\mathbf{r}), \tilde{X} \right\rangle - \frac{1}{2} \left\langle \nabla L_\mu(U), U \right\rangle \\ &\geq -\gamma\sqrt{\epsilon} \text{Tr}(\tilde{X}) - \frac{1}{2}\epsilon \|U\|_F. \end{aligned}$$

In the last step, we used (22) as well as approximate first-order stationarity. \blacksquare

Proof [Proof of Proposition 11] One direction is elementary: if there exists $A_0 \succ 0$ and $b_0 \geq 0$ such that $\langle A_0, X \rangle = b_0$ for all $X \in \mathcal{C}$, then,

$$\forall X \in \mathcal{C}, \quad \text{Tr}(X) = \langle I_n, X \rangle \leq \lambda_{\min}(A_0)^{-1} \langle A_0, X \rangle = \lambda_{\min}(A_0)^{-1} b_0.$$

Thus, the trace of $X \succeq 0$ is bounded, and it follows that \mathcal{C} is compact. Furthermore: if $b_0 = 0$, then $\mathcal{C} = \{0\}$; and if $b_0 > 0$, then $0 \notin \mathcal{C}$.

To prove the other direction, assume \mathcal{C} is non-empty and compact. If $\mathcal{C} = \{0\}$, let $A_0 = I_n$, $b_0 = 0$. Now assume $\mathcal{C} \neq \{0\}$. The SDP comes in a primal-dual pair:

$$\min_{X \in \mathcal{S}^{n \times n}} \langle C, X \rangle \quad \text{s.t.} \quad \mathcal{A}(X) = b, \quad X \succeq 0, \quad (\text{P})$$

$$\max_{y \in \mathbb{R}^m} \langle b, y \rangle \quad \text{s.t.} \quad C \succeq \mathcal{A}^*(y). \quad (\text{D})$$

It is well known that if (D) is infeasible, then (P) is unbounded or infeasible (Wolkowicz, 1981, Thm. 4.1(a)). Since we assume \mathcal{C} is non-empty, this simplifies to: if (D) is infeasible, then (P) is unbounded. The contrapositive states: if (P) is bounded, then (D) is feasible. By our compactness assumption on \mathcal{C} , we know that (P) is bounded for all $C \in \mathcal{S}^{n \times n}$. Thus, (D) is feasible for any C . In particular, take $C = -I_n$: there exists $-y \in \mathbb{R}^m$ such that $A_0 \triangleq \mathcal{A}^*(y) \succeq I_n$. Furthermore,

$$\forall X \in \mathcal{C}, \quad \langle A_0, X \rangle = \langle \mathcal{A}^*(y), X \rangle = \langle y, \mathcal{A}(X) \rangle = \langle y, b \rangle \triangleq b_0.$$

Since there exists $X \neq 0$ in \mathcal{C} , it follows that $b_0 > 0$. ■

Proof [Proof of Theorem 9] Using (19), U is an ϵ -FOSP of the perturbed problem if and only if $\|(M + G)U\|_F \leq \frac{\epsilon}{2}$, where $M = C + 2\mu\mathcal{A}^*(\mathbf{r})$. Let $U = P\Sigma Q^T$ be a thin SVD of U (P is $n \times k$ with orthonormal columns; Q is $k \times k$ orthogonal). Then,

$$\begin{aligned} \|(M + G)U\|_F &= \|(M + G)P\Sigma\|_F \\ &\geq \sigma_k(U) \|(M + G)P\|_F \\ &\geq \sigma_k(U) \sqrt{\sum_{i=1}^k \sigma_{n-(i-1)}(M + G)^2}. \end{aligned}$$

Hence, we control the smallest singular value of U in terms of ϵ and the k smallest singular values of $M + G$:

$$\sigma_k(U) \leq \frac{\epsilon}{2\sqrt{\sum_{i=1}^k \sigma_{n-(i-1)}(M + G)^2}}. \quad (23)$$

The next lemma helps lower-bound the denominator—it follows from Theorem 1.16 and Corollary 1.17 in (Nguyen, 2017); see proof in Appendix C.

Lemma 21 *Let \bar{M} be a fixed symmetric matrix of size n . Let G be a symmetric Gaussian matrix of size n , independent of \bar{M} , with diagonal and upper-triangular entries sampled independently from $\mathcal{N}(0, \sigma_G^2)$. There exists an absolute constant c_0 such that:*

$$\Pr \left[\sum_{i=1}^k \sigma_{n-(i-1)}(\bar{M} + G)^2 < \frac{k^2}{c_0 n} \sigma_G^2 \right] \leq \exp \left(-\frac{k^2}{8} \log(8\pi) + k \log(n) \right).$$

We cannot use Lemma 21 directly, as in our case M is not statistically independent of G . Indeed, M depends on U through the residue $\mathbf{r} = \mathbf{r}(U)$ and U is an ϵ -FOSP: a feature that depends on G . To

resolve this, we cover the set of possible M 's with a net, under the assumption that \mathbf{r} is bounded. Lemma 21 provides a bound for each \bar{M} in this net. This can be extended to hold for all \bar{M} 's in the net simultaneously via a union bound. By taking a sufficiently dense net, we can then infer that M is necessarily close to one of these \bar{M} 's, and conclude.

Let \mathcal{E} be the event (on G) that $\|\mathbf{r}\|_2 \leq B$ for all ϵ -FOSPs of the perturbed problem. Conditioned on \mathcal{E} , we have

$$\|M - C\|_F = 2\mu\|\mathcal{A}^*(\mathbf{r})\|_F \leq 2\mu B\|\mathcal{A}\|,$$

where $\|\mathcal{A}\|$ is defined in (8). As a result, M lies in a ball of center C and radius $2\mu B\|\mathcal{A}\|$ in an affine subspace of dimension $\text{rank}(\mathcal{A})$. A unit-ball in Frobenius norm in d dimensions admits an ε -net of $(1 + 2/\varepsilon)^d$ points (Vershynin, 2016, Cor. 4.2.13). Thus, we can pick a net with $\left(1 + \frac{4\mu B\|\mathcal{A}\|}{\sigma_G} \sqrt{\frac{4c_0 n}{k^2}}\right)^{\text{rank}(\mathcal{A})}$ points in such a way that, independently of \mathbf{r} , there exists a point \bar{M} in the net satisfying:

$$\|\bar{M} - M\|_F \leq \sqrt{\frac{k^2}{4c_0 n}} \sigma_G = \frac{k}{2\sqrt{c_0 n}} \sigma_G. \quad (24)$$

Let $T: \mathcal{S}^{n \times n} \rightarrow \mathbb{R}^k$ be defined by $T_q(A) = (\sigma_{n-q+1}(A), \dots, \sigma_n(A))^T$, that is: T extracts the q smallest singular values of A , in order. Then,

$$\begin{aligned} \|\bar{M} - M\|_F &= \|(\bar{M} + G) - (M + G)\|_F \\ &\geq \|T_n(\bar{M} + G) - T_n(M + G)\|_2 \\ &\geq \|T_k(\bar{M} + G) - T_k(M + G)\|_2 \\ &\geq \|T_k(\bar{M} + G)\|_2 - \|T_k(M + G)\|_2, \end{aligned}$$

where the first inequality follows from (Bhatia, 2013, Ex. IV.3.5). Hence,

$$\sqrt{\sum_{i=1}^k \sigma_{n-(i-1)}(M + G)^2} \geq \sqrt{\sum_{i=1}^k \sigma_{n-(i-1)}(\bar{M} + G)^2} - \|\bar{M} - M\|_F. \quad (25)$$

Now, taking a union bound for \mathcal{E} and for Lemma 21 over each \bar{M} in the net, we get (24) and

$$\sqrt{\sum_{i=1}^k \sigma_{n-(i-1)}(\bar{M} + G)^2} \geq \frac{k}{\sqrt{c_0 n}} \sigma_G \quad (26)$$

with probability at least

$$1 - \exp\left(-\frac{k^2}{8} \log(8\pi) + k \log(n) + \text{rank}(\mathcal{A}) \cdot \log\left(1 + \frac{4\mu B\|\mathcal{A}\|}{\sigma_G} \sqrt{\frac{4c_0 n}{k^2}}\right)\right) - \delta.$$

Inside the log, we can safely replace k with 1, as this only hurts the probability. Then, the result holds with probability at least

$$1 - \exp\left(-\frac{k^2}{8} \log(8\pi) + k \log(n) + \text{rank}(\mathcal{A}) \cdot \log\left(1 + \frac{8\mu B\|\mathcal{A}\|}{\sigma_G} \sqrt{c_0 n}\right)\right) - \delta.$$

We aim to pick k so as to ensure

$$\exp \left(-\frac{k^2}{8} \log(8\pi) + k \log(n) + \text{rank}(\mathcal{A}) \cdot \log \left(1 + \frac{8\mu B \|\mathcal{A}\|}{\sigma_G} \sqrt{c_0 n} \right) \right) \leq \delta'.$$

This is a quadratic condition of the form

$$-ak^2 + bk + c \leq \log(\delta')$$

for some $a, b > 0, c \geq 0$. Since k is positive we get, $k \geq \frac{b + \sqrt{a(c + \log(1/\delta'))}}{a}$, which is satisfied for,

$$k \geq 3 \left[\log \left(\frac{n}{\delta'} \right) + \sqrt{\text{rank}(\mathcal{A}) \log \left(1 + \frac{8\mu B \|\mathcal{A}\| \sqrt{c_0 n}}{\sigma_G} \right)} \right].$$

Combining (23), (24), (25) and (26), we find:

$$\sigma_k(U) \leq \frac{\epsilon}{\sigma_G} \frac{2\sqrt{c_0 n}}{k}$$

with probability at least $1 - \delta - \delta'$. ■

Proof [Proof of Lemma 12] If $U = 0$, the bounds clearly hold: assume $U \neq 0$ in what follows. Using $\nabla \tilde{L}_\mu(U) = 2(C + 2\mu \tilde{\mathcal{A}}^*(\tilde{\mathbf{r}}))U$, the definition of ϵ -FOSP reads:

$$\frac{\epsilon}{2} \geq \left\| \left(C + 2\mu \tilde{\mathcal{A}}^*(\tilde{\mathbf{r}}) \right) U \right\|_F.$$

Combining this with $\|A\|_F \geq \frac{1}{\|B\|_F} \langle A, B \rangle$ for $B \neq 0$ (Cauchy–Schwarz) gives:

$$\frac{\epsilon}{2} \geq \frac{1}{\|U\|_F} \left\langle \left(C + 2\mu \tilde{\mathcal{A}}^*(\tilde{\mathbf{r}}) \right) U, U \right\rangle.$$

This can be further developed as:

$$\begin{aligned} \frac{\epsilon \|U\|_F}{2} &\geq \left\langle C + 2\mu \tilde{\mathcal{A}}^*(\tilde{\mathbf{r}}), UU^T \right\rangle \\ &= \langle C, UU^T \rangle + 2\mu \left\langle \tilde{\mathbf{r}}, \tilde{\mathcal{A}}(UU^T) \right\rangle \\ &= \langle C, UU^T \rangle + 2\mu \left\langle \tilde{\mathcal{A}}(UU^T) - \tilde{\mathbf{b}}, \tilde{\mathcal{A}}(UU^T) \right\rangle. \end{aligned} \tag{27}$$

At this point, we separate the constraint (A_0, b_0) from the rest, using the usual definition for $(\mathcal{A}, \mathbf{b})$ which capture constraints $1, \dots, m$:

$$\begin{aligned} \frac{\epsilon \|U\|_F}{2} &\geq \langle C, UU^T \rangle + 2\mu \left(\langle \mathcal{A}(UU^T) - \mathbf{b}, \mathcal{A}(UU^T) \rangle + (\langle A_0, UU^T \rangle - b_0) \langle A_0, UU^T \rangle \right) \\ &\geq \langle C, UU^T \rangle + 2\mu \left(\|\mathcal{A}(UU^T)\|_2^2 - \|\mathbf{b}\|_2 \|\mathcal{A}(UU^T)\|_2 + (\langle A_0, UU^T \rangle - b_0) \langle A_0, UU^T \rangle \right). \end{aligned}$$

Let $y = \|\mathcal{A}(UU^T)\|_2$. Then the above inequality holds when

$$y^2 - \|\mathbf{b}\|_2 y + \frac{1}{2\mu} \left(\langle C, UU^T \rangle - \frac{\epsilon \|U\|_F}{2} \right) + (\langle A_0, UU^T \rangle - b_0) \langle A_0, UU^T \rangle \leq 0.$$

For this to happen we need the above quadratic to have real roots. This requires:

$$\begin{aligned} \frac{1}{4}\|\mathbf{b}\|_2^2 &\geq \frac{1}{2\mu} \left(\langle C, UU^T \rangle - \frac{\epsilon\|U\|_F}{2} \right) + (\langle A_0, UU^T \rangle - b_0) \langle A_0, UU^T \rangle \\ &\geq \frac{1}{2\mu} \left(-\|CU\|_F \|U\|_F - \frac{\epsilon\|U\|_F}{2} \right) + \lambda_{\min}(A_0)^2 \|U\|_F^4 - b_0 \lambda_{\max}(A_0) \|U\|_F^2 \\ &\geq \lambda_{\min}(A_0)^2 \|U\|_F^4 - \frac{\|C\|_2}{2\mu} \|U\|_F^2 - b_0 \lambda_{\max}(A_0) \|U\|_F^2 - \frac{\epsilon}{4\mu} \|U\|_F, \end{aligned}$$

where we used that for any two matrices A and B , it holds that $\|AB\|_F \leq \|A\|_2 \|B\|_F$. Focus on the last two terms of the last inequality. We distinguish two cases. Either

$$b_0 \lambda_{\max}(A_0) \|U\|_F^2 + \frac{\epsilon}{4\mu} \|U\|_F \geq \frac{3}{2} b_0 \lambda_{\max}(A_0) \|U\|_F^2,$$

in which case $\|U\|_F \leq \frac{\epsilon}{2\mu b_0 \lambda_{\max}(A_0)}$ (assuming $b_0 > 0$). Or the opposite holds, and:

$$\frac{1}{4}\|\mathbf{b}\|_2^2 \geq \lambda_{\min}(A_0)^2 \|U\|_F^4 - \left(\frac{\|C\|_2}{2\mu} + \frac{3}{2} b_0 \lambda_{\max}(A_0) \right) \|U\|_F^2.$$

This is a quadratic inequality in $y = \|U\|_F^2$ of the form $ay^2 - by - c \leq 0$ with coefficients $a > 0$ and $b, c \geq 0$. Such a quadratic always has at least one real root, so that $y \leq \frac{b + \sqrt{b^2 + 4ac}}{2a}$. Furthermore, $\sqrt{b^2 + 4ac} \leq \sqrt{b^2 + (\sqrt{4ac})^2} + 2b\sqrt{4ac} = b + \sqrt{4ac}$. Hence, $y \leq \frac{b}{a} + \sqrt{\frac{c}{a}}$, which means:

$$\|U\|_F^2 \leq \frac{1}{\lambda_{\min}(A_0)^2} \left(\frac{\|C\|_2}{2\mu} + \frac{3}{2} b_0 \lambda_{\max}(A_0) \right) + \frac{\|\mathbf{b}\|_2}{2\lambda_{\min}(A_0)}.$$

Accounting for the two distinguished cases, we find:

$$\|U\|_F^2 \leq \max \left\{ \left(\frac{\epsilon}{2\mu b_0 \lambda_{\max}(A_0)} \right)^2, \frac{1}{\lambda_{\min}(A_0)^2} \left(\frac{\|C\|_2}{2\mu} + \frac{3}{2} b_0 \lambda_{\max}(A_0) \right) + \frac{\|\mathbf{b}\|_2}{2\lambda_{\min}(A_0)} \right\}.$$

We now bound the residues (generically) in terms of $\|U\|_F$, using submultiplicativity for $\|UU^T\|_F \leq \|U\|_F^2$ and the definition of $\|\mathcal{A}\|$ (8):

$$\|\mathbf{r}\|_2 = \|\mathcal{A}(UU^T) - \mathbf{b}\|_2 \leq \|\mathcal{A}\| \|UU^T\|_F + \|\mathbf{b}\|_2 \leq \|\mathcal{A}\| \|U\|_F^2 + \|\mathbf{b}\|_2.$$

Evidently, the same bound holds for $\tilde{\mathcal{A}}, \tilde{\mathbf{b}}, \tilde{\mathbf{r}}$. ■

Proof [Proof of Theorem 13]

By Lemma 12, for a problem perturbed with G , the residues of all ϵ -FOSPs, $\|\tilde{\mathbf{r}}\|_2$, are bounded as:

$$\|\tilde{\mathcal{A}}\| \max \left\{ \left(\frac{\epsilon}{2\mu b_0 \lambda_{\max}(A_0)} \right)^2, \frac{1}{\lambda_{\min}(A_0)^2} \left(\frac{\|C + G\|_2}{2\mu} + \frac{3}{2} b_0 \lambda_{\max}(A_0) \right) + \frac{\|\mathbf{b}\|_2}{2\lambda_{\min}(A_0)} \right\} + \|\tilde{\mathbf{b}}\|_2$$

With probability at least $1 - \delta$, $\|C + G\|_2 \leq \|C\|_2 + 3\sigma_G \left(\sqrt{n} + \sqrt{2 \log(1/\delta)} \right)$. Hence, Theorem 9 applies with this δ and

$$B = \|\tilde{\mathcal{A}}\| \max \left\{ \left(\frac{\epsilon}{2\mu b_0 \lambda_{\max}(A_0)} \right)^2, \frac{1}{\lambda_{\min}(A_0)^2} \left(\frac{\|C\|_2 + 3\sigma_G \sqrt{n}}{2\mu} + \frac{3}{2} b_0 \lambda_{\max}(A_0) \right) + \frac{\|\mathbf{b}\|_2}{2\lambda_{\min}(A_0)} \right\} + \|\tilde{\mathbf{b}}\|_2.$$

Hence, with k as prescribed in that theorem for a given $\delta' = \delta \in (0, 1)$, with probability at least $1 - 2\delta$, it holds that

$$\sigma_k(U) \leq \frac{2\epsilon}{\sigma_G} \frac{\sqrt{c_0 n}}{k}$$

for any ϵ -FOSP. Lemma 8 requires $\sigma_k^2(U) \leq \frac{\gamma \sqrt{\epsilon}}{8\mu \|\mathcal{A}\|^2}$. Hence, we choose: $\epsilon \leq \left(\frac{\gamma k^2 \sigma_G^2}{32 c_0 n \mu \|\mathcal{A}\|^2} \right)^{2/3}$, and with probability at least $1 - 2\delta$ hypothesis of Lemma 8 is satisfied. Let \tilde{X} be a global optimum for \tilde{F}_μ , then the optimality gap obeys:

$$\tilde{F}_\mu(UU^T) - \tilde{F}_\mu(\tilde{X}) \leq \gamma \sqrt{\epsilon} \text{Tr}(\tilde{X}) + \frac{1}{2} \epsilon \|U\|_F.$$

■

F.1. Proof of section 4.2

Proof [Proof of Lemma 14] With probability at least $1 - \delta$, $\sigma_1(G) \leq 3\sigma_G \sqrt{n}$. In that event, for $\sigma_G \leq \frac{\lambda_n(C)}{6\sqrt{n \log(n/\delta)}}$, we have $C + G \succeq \frac{\lambda_n(C)}{2} I$.

U is an ϵ -FOSP of (15) implies $\|2(C + G + 2\mu \mathcal{A}^*(\mathbf{r}))U\|_F \leq \epsilon$.

$$\begin{aligned} \frac{\epsilon}{2} &\geq \|(C + G + 2\mu \mathcal{A}^*(\mathbf{r}))U\|_F \\ &\geq \frac{1}{\|U\|_F} \langle C + G + 2\mu \mathcal{A}^*(\mathbf{r}), UU^T \rangle. \end{aligned}$$

Hence,

$$\begin{aligned} \frac{\epsilon \|U\|_F}{2} &\geq \langle C + G, UU^T \rangle + 2\mu \langle \mathcal{A}^*(\mathbf{r}), UU^T \rangle \\ &\geq \frac{\lambda_n(C)}{2} \|U\|_F^2 + 2\mu \langle \mathbf{r}, \mathcal{A}(UU^T) \rangle \\ &\geq \frac{\lambda_n(C)}{2} \|U\|_F^2 + 2\mu (\|\mathcal{A}(UU^T)\|_2^2 - \|\mathbf{b}\|_2 \|\mathcal{A}(UU^T)\|_2). \end{aligned}$$

The above inequality is a quadratic in $y = \|\mathcal{A}(UU^T)\|_2$: $y^2 - y\|\mathbf{b}\|_2 + \frac{1}{2\mu} \left(\frac{\lambda_n(C)}{2} \|U\|_F^2 - \frac{\epsilon \|U\|_F}{2} \right) \leq 0$. If $\frac{\epsilon \|U\|_F}{2} \geq \frac{\lambda_n(C)}{4} \|U\|_F^2$, then $\|U\|_F \leq \frac{2\epsilon}{\lambda_n(C)}$. Else, for the above inequality to hold we need the quadratic to have real roots.

$$\begin{aligned} \|\mathbf{b}\|_2^2 &\geq 4 \cdot 1 \cdot \frac{1}{2\mu} \left(\frac{\lambda_n(C)}{2} \|U\|_F^2 - \frac{\epsilon \|U\|_F}{2} \right) \\ &\geq \frac{2}{\mu} \frac{\lambda_n(C)}{4} \|U\|_F^2. \end{aligned}$$

The last inequality follows from $\frac{\epsilon\|U\|_F}{2} \leq \frac{\lambda_n(C)}{4}\|U\|_F^2$. Hence, $\|U\|_F^2 \leq \max \left\{ \left(\frac{2\epsilon}{\lambda_n(C)} \right)^2, \frac{2\mu}{\lambda_n(C)} \|\mathbf{b}\|_2^2 \right\}$. Hence,

$$\begin{aligned} \|\mathbf{r}\|_2 = \|\mathcal{A}(UU^T) - \mathbf{b}\|_2 &\leq \|\mathcal{A}(UU^T)\|_2 + \|\mathbf{b}\|_2 \leq \|\mathcal{A}\| \|UU^T\|_F + \|\mathbf{b}\|_2 \\ &\leq \|\mathcal{A}\| \max \left\{ \left(\frac{2\epsilon}{\lambda_n(C)} \right)^2, \frac{2\mu}{\lambda_n(C)} \|\mathbf{b}\|_2^2 \right\} + \|\mathbf{b}\|_2. \end{aligned}$$

■

Appendix G. Proofs for Section B

Proof [Proof of Lemma 19.] We start by showing that the gradient is l -Lipschitz continuous. The gradient is given by:

$$\nabla \widehat{L}_\mu(U) = [2(C + G) + 4\mu\mathcal{A}^*(\mathbf{r})] U,$$

where $\mathbf{r} = \mathbf{r}(U) = \mathcal{A}(UU^T) - \mathbf{b}$. Hence, for $U_1, U_2 \in \mathbb{R}^{n \times k}$, with notation $\mathbf{r}_1 = \mathbf{r}(U_1)$, $\mathbf{r}_2 = \mathbf{r}(U_2)$,

$$\begin{aligned} \left\| \nabla \widehat{L}_\mu(U_1) - \nabla \widehat{L}_\mu(U_2) \right\|_F &\leq \|2(C + G)(U_1 - U_2)\|_F + 4\mu \|\mathcal{A}^*(\mathbf{r}_1)U_1 - \mathcal{A}^*(\mathbf{r}_2)U_2\|_F \\ &\leq 2\|C + G\|_2 \|U_1 - U_2\|_F + 4\mu \|\mathcal{A}^*(\mathbf{r}_1)(U_1 - U_2)\|_F \\ &\quad + 4\mu \|\mathcal{A}^*(\mathbf{r}_1 - \mathbf{r}_2)U_2\|_F \\ &\leq (2\|C + G\|_2 + 4\mu \|\mathcal{A}^*(\mathbf{r}_1)\|_2) \|U_1 - U_2\|_F \\ &\quad + 4\mu \|\mathcal{A}^*(\mathbf{r}_1 - \mathbf{r}_2)U_2\|_F. \end{aligned}$$

This further simplifies using the norm of \mathcal{A} (8): $\|\mathcal{A}^*(\mathbf{r}_1)\|_2 \leq \|\mathcal{A}\| \|\mathbf{r}_1\|_2$ and $\|\mathbf{r}_1\|_2 \leq \|\mathcal{A}\| \|U_1\|_F^2 + \|\mathbf{b}\|_2$, so that if $\|U_1\|_F \leq \tau$:

$$\|\mathcal{A}^*(\mathbf{r}_1)\|_2 \leq (\tau^2 \|\mathcal{A}\| + \|\mathbf{b}\|_2) \|\mathcal{A}\|.$$

Similarly, using $\|U_2\|_F \leq \tau$ as well:

$$\begin{aligned} \|\mathcal{A}^*(\mathbf{r}_1 - \mathbf{r}_2)U_2\|_F &\leq \|\mathcal{A}^*(\mathcal{A}(U_1U_1^T - U_2U_2^T))\|_2 \|U_2\|_F \\ &\leq \tau \|\mathcal{A}\|^2 \|U_1U_1^T - U_2U_2^T\|_F \\ &= \tau \|\mathcal{A}\|^2 \|U_1U_1^T - U_1U_2^T + U_1U_2^T - U_2U_2^T\|_F \\ &\leq \tau \|\mathcal{A}\|^2 (\|U_1(U_1 - U_2)^T\|_F + \|(U_1 - U_2)U_2^T\|_F) \\ &\leq 2\tau^2 \|\mathcal{A}\|^2 \|U_1 - U_2\|_F. \end{aligned} \tag{28}$$

Combining, we find

$$\begin{aligned} \left\| \nabla \widehat{L}_\mu(U_1) - \nabla \widehat{L}_\mu(U_2) \right\|_F &\leq (2\|C + G\|_2 + 4\mu \|\mathcal{A}\| (\tau^2 \|\mathcal{A}\| + \|\mathbf{b}\|_2)) \|U_1 - U_2\|_F \\ &\quad + 8\mu\tau^2 \|\mathcal{A}\|^2 \|U_1 - U_2\|_F, \end{aligned}$$

which establishes the Lipschitz constant for $\nabla \widehat{L}_\mu$.

We now show that the Hessian is ρ -Lipschitz continuous in operator norm, that is, we must show that for any U_1 and U_2 with norms bounded by τ ,

$$\max_{\|\dot{U}\|_F \leq 1} \left\langle \nabla^2 \widehat{L}_\mu(U_1)[\dot{U}] - \nabla^2 \widehat{L}_\mu(U_2)[\dot{U}], \dot{U} \right\rangle \leq \rho \|U_1 - U_2\|_F.$$

Recall from (21) that

$$\left\langle \nabla^2 \widehat{L}_\mu(U)[\dot{U}], \dot{U} \right\rangle = 2 \left\langle C + G + 2\mu \mathcal{A}^*(\mathbf{r}), \dot{U} \dot{U}^T \right\rangle + 2\mu \|\mathcal{A}(U \dot{U}^T + \dot{U} U^T)\|_2^2.$$

Hence,

$$\begin{aligned} & \left\langle \nabla^2 \widehat{L}_\mu(U_1)[\dot{U}], \dot{U} \right\rangle - \left\langle \nabla^2 \widehat{L}_\mu(U_2)[\dot{U}], \dot{U} \right\rangle \\ &= 4\mu \left\langle \mathcal{A}^*(\mathbf{r}_1 - \mathbf{r}_2), \dot{U} \dot{U}^T \right\rangle + 2\mu \left(\|\mathcal{A}(U_1 \dot{U}^T + \dot{U} U_1^T)\|_2^2 - \|\mathcal{A}(U_2 \dot{U}^T + \dot{U} U_2^T)\|_2^2 \right). \end{aligned}$$

On one hand, following the same reasoning as in (28), we have

$$\begin{aligned} \left\langle \mathcal{A}^*(\mathbf{r}_1 - \mathbf{r}_2), \dot{U} \dot{U}^T \right\rangle &\leq \|\mathcal{A}^*(\mathbf{r}_1 - \mathbf{r}_2)\|_F \|\dot{U} \dot{U}^T\|_F \\ &\leq 2\tau \|\mathcal{A}\|^2 \|U_1 - U_2\|_F \|\dot{U}\|_F^2. \end{aligned}$$

On the other hand, using that for any two vectors u, v we have

$$\|u\|_2^2 - \|v\|_2^2 = \langle u + v, u - v \rangle \leq \|u + v\|_2 \|u - v\|_2 \leq (\|u\|_2 + \|v\|_2) \|u - v\|_2,$$

we can find:

$$\|\mathcal{A}(U_1 \dot{U}^T + \dot{U} U_1^T)\|_2^2 - \|\mathcal{A}(U_2 \dot{U}^T + \dot{U} U_2^T)\|_2^2 \leq 4\tau \|\mathcal{A}\|^2 \|U_1 - U_2\|_F \|\dot{U}\|_F^2.$$

For this, we used $\|\mathcal{A}(U \dot{U}^T + \dot{U} U^T)\|_2 \leq \|\mathcal{A}\| \|U \dot{U}^T + \dot{U} U^T\|_F \leq \tau \|\mathcal{A}\| \|\dot{U}\|_F$ when $\|U\|_F \leq \tau$ and

$$\begin{aligned} \|\mathcal{A}(U_1 \dot{U}^T + \dot{U} U_1^T - U_2 \dot{U}^T - \dot{U} U_2^T)\|_2 &\leq \|\mathcal{A}\| \left(\|(U_1 - U_2) \dot{U}^T\|_F + \|\dot{U} (U_1 - U_2)^T\|_F \right) \\ &\leq 2\|\mathcal{A}\| \|\dot{U}\|_F \|U_1 - U_2\|_F. \end{aligned}$$

Overall, this shows $\rho = 16\mu\tau \|\mathcal{A}\|^2$ is an appropriate Lipschitz constant. ■