Smoothed analysis of the low-rank approach for smooth semidefinite programs

Thomas Pumir[*]
ORFE Department
Princeton University
tpumir@princeton.edu

Samy Jelassi[*]
ORFE Department
Princeton University
sjelassi@princeton.edu

Nicolas Boumal
Department of Mathematics
Princeton University
nboumal@math.princeton.edu

Abstract

We consider semidefinite programs (SDPs) of size n with equality constraints. In order to overcome scalability issues, Burer and Monteiro proposed a factorized approach based on optimizing over a matrix Y of size $n \times k$ such that $X = YY^*$ is the SDP variable. The advantages of such formulation are twofold: the dimension of the optimization variable is reduced, and positive semidefiniteness is naturally enforced. However, optimization in Y is non-convex. In prior work, it has been shown that, when the constraints on the factorized variable regularly define a smooth manifold, provided k is large enough, for almost all cost matrices, all second-order stationary points (SOSPs) are optimal. Importantly, in practice, one can only compute points which approximately satisfy necessary optimality conditions, leading to the question: are such points also approximately optimal? To answer it, under similar assumptions, we use smoothed analysis to show that approximate SOSPs for a randomly perturbed objective function are approximate global optima, with k scaling like the square root of the number of constraints (up to log factors). Moreover, we bound the optimality gap at the approximate solution of the perturbed problem with respect to the original problem. We particularize our results to an SDP relaxation of phase retrieval.

1 Introduction

We consider semidefinite programs (SDP) over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ of the form:

$$\begin{aligned} \min_{X \in \mathbb{S}^{n\times n}} \quad & \langle C, X\rangle \\ \text{subject to} \quad & \mathcal{A}(X) = b, \\ & X \succeq 0, \end{aligned} \tag{SDP}$$

with $\langle A, B\rangle = \mathrm{Re}[\mathrm{Tr}(A^*B)]$ the Frobenius inner product (A^* is the conjugate-transpose of A), $\mathbb{S}^{n\times n}$ the set of self-adjoint matrices of size n (real symmetric for $\mathbb{R}$, or Hermitian for $\mathbb{C}$), $C \in \mathbb{S}^{n\times n}$ the cost matrix, and $\mathcal{A}\colon \mathbb{S}^{n\times n} \to \mathbb{R}^m$ a linear operator capturing m equality constraints with right hand side $b \in \mathbb{R}^m$: for each i, $\mathcal{A}(X)_i = \langle A_i, X\rangle = b_i$ for given matrices $A_1, \ldots, A_m \in \mathbb{S}^{n\times n}$. The optimization variable X is positive semidefinite. We let $\mathcal{C}$ be the feasible set of (SDP):

$$\mathcal{C} = \left\{X \in \mathbb{S}^{n\times n} : \mathcal{A}(X) = b \text{ and } X \succeq 0\right\}. \tag{1}$$

[*]Equal contribution

32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.

Large-scale SDPs have been proposed for machine learning applications including matrix completion [Candès and Recht, 2009], community detection [Abbé, 2018] and kernel learning [Lanckriet et al., 2004] for $\mathbb{K} = \mathbb{R}$, and in angular synchronization [Singer, 2011] and phase retrieval [Waldspurger et al., 2015] for $\mathbb{K} = \mathbb{C}$. Unfortunately, traditional methods to solve (SDP) do not scale (due to memory and computational requirements), hence the need for alternatives.

In order to address such scalability issues, Burer and Monteiro [2003, 2005] restrict the search to the set of matrices of rank at most k by factorizing X as $X = YY^*$, with $Y \in \mathbb{K}^{n\times k}$. It has been shown that if the search space $\mathcal{C}$ (1) is compact, then (SDP) admits a global optimum of rank at most r, where $\dim \mathbb{S}^{r\times r} \leq m$ [Barvinok, 1995, Pataki, 1998], with $\dim \mathbb{S}^{r\times r} = \frac{r(r+1)}{2}$ for $\mathbb{K} = \mathbb{R}$ and $\dim \mathbb{S}^{r\times r} = r^2$ for $\mathbb{K} = \mathbb{C}$. In other words, restricting $\mathcal{C}$ to the space of matrices with rank at most k with $\dim \mathbb{S}^{k\times k} > m$ does not change the optimal value. This factorization leads to a quadratically constrained quadratic program:

$$\begin{aligned} \min_{Y\in\mathbb{K}^{n\times k}} \quad & \langle C, YY^*\rangle \\ \text{subject to} \quad & \mathcal{A}(YY^*) = b. \end{aligned} \tag{P}$$

Although (P) is in general non-convex because its feasible set

$$\mathcal{M} = \mathcal{M}_k = \left\{Y \in \mathbb{K}^{n\times k} : \mathcal{A}(YY^*) = b\right\} \tag{2}$$

is non-convex, considering (P) instead of the original SDP presents significant advantages: the number of variables is reduced from $O(n^2)$ to $O(nk)$, and the positive semidefiniteness of the matrix is naturally enforced. Solving (P) using local optimization methods is known as the Burer–Monteiro method and yields good results in practice: Kulis et al. [2007] underlined the practical success of such low-rank approaches in particular for maximum variance unfolding and for k-means clustering (see also [Carson et al., 2017]). Their approach is significantly faster and more scalable. However, the non-convexity of (P) means further analysis is needed to determine whether it can be solved to global optimality reliably.

For $\mathbb{K} = \mathbb{R}$, in the case where $\mathcal{M}$ is a compact, smooth manifold (see Assumption 1 below for a precise condition), it has been shown recently that, up to a zero-measure set of cost matrices, second-order stationary points (SOSPs) of (P) are globally optimal provided $\dim \mathbb{S}^{k\times k} > m$ [Boumal et al., 2016, 2018b]. Algorithms such as the Riemannian trust-regions method (RTR) converge globally to SOSPs, but unfortunately they can only guarantee *approximate* satisfaction of second-order optimality conditions in a finite number of iterations [Boumal et al., 2018a].

The aforementioned papers close with a question, crucial in practice: when is it the case that *approximate* SOSPs, which we now call ASOSPs, are approximately optimal? Building on recent proof techniques by Bhojanapalli et al. [2018], we provide some answers here.

Contributions

This paper formulates approximate global optimality conditions holding for (P) and, consequently, for (SDP). Our results rely on the following core assumption as set in [Boumal et al., 2016].

Assumption 1 (Smooth manifold)**.** *For all values of k up to n such that $\mathcal{M}_k$ is non-empty, the constraints on* (P) *defined by $A_1, \ldots, A_m \in \mathbb{S}^{n\times n}$ and $b \in \mathbb{R}^m$ satisfy at least one of the following:*

1. *$\{A_1Y, \ldots, A_mY\}$ are linearly independent in $\mathbb{K}^{n\times k}$ for all $Y \in \mathcal{M}_k$; or*
2. *$\{A_1Y, \ldots, A_mY\}$ span a subspace of constant dimension in $\mathbb{K}^{n\times k}$ for all Y in an open neighborhood of $\mathcal{M}_k$ in $\mathbb{K}^{n\times k}$.*

In [Boumal et al., 2018b], it is shown that (a) if the assumption above is verified for $k = n$, then it automatically holds for all values of $k \leq n$ such that $\mathcal{M}_k$ is non-empty; and (b) for those values of k, the dimension of the subspace spanned by $\{A_1Y, \ldots, A_mY\}$ is independent of k: we call it m'.

When Assumption 1 holds, we refer to problems of the form (SDP) as *smooth* SDPs because $\mathcal{M}$ is then a smooth manifold. Examples of smooth SDPs for $\mathbb{K} = \mathbb{R}$ are given in [Boumal et al., 2018b]. For $\mathbb{K} = \mathbb{C}$, we detail an example in Section 4. Our main theorem is a smooth analysis result (cf. Theorem 3.1 for a more formal statement). An ASOSP is an *approximate* SOSP (a precise definition follows.)

Theorem 1.1 (Informal)**.** *Let Assumption 1 hold and assume $\mathcal{C}$ is compact. Randomly perturb the cost matrix C. With high probability, if $k = \tilde{\Omega}(\sqrt{m})$, any ASOSP $Y \in \mathbb{K}^{n\times k}$ for* (P) *is an approximate global optimum, and $X = YY^*$ is an approximate global optimum for* (SDP) *(with the perturbed C.)*

The high probability proviso is with respect to the perturbation only: if the perturbation is "good", then all ASOSPs are as described in the statement. If $\mathcal{C}$ is compact, then so is $\mathcal{M}$ and known algorithms for optimization on manifolds produce an ASOSP in finite time (with explicit bounds). Theorem 1.1 ensures that, for k large enough and for any cost matrix C, with high probability upon a random perturbation of C, such algorithms produce an approximate global optimum of (P).

Theorem 1.1 is a corollary of two intermediate arguments, developed in Lemmas 3.1 and 3.2:

1. Probabilistic argument (Lemma 3.1): By perturbing the cost matrix in the objective function of (P) with a Gaussian Wigner matrix, with high probability, any approximate first-order stationary point Y of the perturbed problem (P) is almost column-rank deficient.

2. Deterministic argument (Lemma 3.2): If an approximate second-order stationary point Y for (P) is also almost column-rank deficient, then it is an approximate global optimum and $X = YY^*$ is an approximate global optimum for (SDP).

The first argument is motivated by *smoothed analysis* [Spielman and Teng, 2004] and draws heavily on a recent paper by Bhojanapalli et al. [2018]. The latter work introduces smoothed analysis to analyze the performance of the Burer–Monteiro factorization, but it analyzes a quadratically penalized version of the SDP: its solutions do not satisfy constraints exactly. This affords more generality, but, for the special class of smooth SDPs, the present work has the advantage of analyzing an exact formulation. The second argument is a smoothed extension of well-known on-off results [Burer and Monteiro, 2003, 2005, Journee et al., 2010]. Implications of this theorem for a particular SDP are derived in Section 4, with applications to phase retrieval and angular synchronization.

Thus, for smooth SDPs, our results improve upon [Bhojanapalli et al., 2018] in that we address exact-feasibility formulations of the SDP. Our results also improve upon [Boumal et al., 2016] by providing approximate optimality results for approximate second-order points with relaxation rank k scaling only as $\tilde{\Omega}(\sqrt{m})$, whereas the latter reference establishes such results only for $k = n + 1$. Finally, we aim for more generality by covering both real and complex SDPs, and we illustrate the relevance of complex SDPs in Section 4.

Related work

A number of recent works focus on large-scale SDP solvers. Among the direct approaches (which proceed in the convex domain directly), Hazan [2008] introduced a Frank–Wolfe type method for a restricted class of SDPs. Here, the key is that each iteration increases the rank of the solution only by one, so that if only a few iterations are required to reach satisfactory accuracy, then only low dimensional objects need to be manipulated. This line of work was later improved by Laue [2012], Garber [2016] and Garber and Hazan [2016] through hybrid methods. Still, if high accuracy solutions are desired, a large number of iterations will be required, eventually leading to large-rank iterates. In order to overcome such issue, Yurtsever et al. [2017] recently proposed to combine conditional gradient and sketching techniques in order to maintain a low rank representation of the iterates.

Among the low-rank approaches, our work is closest to (and indeed largely builds upon) recent results of Bhojanapalli et al. [2018]. For the real case, they consider a penalized version of problem (SDP) (which we here refer to as (P-SDP)) and its related penalized Burer–Monteiro formulation, here called (P-P). With high probability upon random perturbation of the cost matrix, they show approximate global optimality of ASOSPs for (P-P), assuming k grows with $\sqrt{m}$ and either the SDP is compact or its cost matrix is positive definite. Given that there is a zero-measure set of SDPs where SOSPs may be suboptimal, there can be a small-measure set of SDPs where ASOSPs are not approximately optimal [Bhojanapalli et al., 2018]. In this context, the authors resort to smoothed analysis, in the same way that we do here. One drawback of that work is that the final result does not hold for the original SDP, but for a non-equivalent penalized version of it. This is one of the points we improve here, by focusing on smooth SDPs as defined in [Boumal et al., 2016].

Notation

We use $\mathbb{K}$ to refer to $\mathbb{R}$ or $\mathbb{C}$ when results hold for both fields. For matrices A, B of same size, we use the inner product $\langle A, B\rangle = \mathrm{Re}[\mathrm{Tr}(A^*B)]$, which reduces to $\langle A, B\rangle = \mathrm{Tr}(A^TB)$ in the real case. The associated Frobenius norm is defined as $\|A\| = \sqrt{\langle A, A\rangle}$. For a linear map f between matrix spaces, this yields a subordinate operator norm as $\|f\|_{\mathrm{op}} = \sup_{A\neq 0} \frac{\|f(A)\|}{\|A\|}$. The set of self-adjoint matrices of size n over $\mathbb{K}$, $\mathbb{S}^{n\times n}$, is the set of symmetric matrices for $\mathbb{K} = \mathbb{R}$ or the set of Hermitian matrices for $\mathbb{K} = \mathbb{C}$. We also write $\mathbb{H}^{n\times n}$ to denote $\mathbb{S}^{n\times n}$ for $\mathbb{K} = \mathbb{C}$. A self-adjoint matrix X is positive semidefinite ($X \succeq 0$) if and only if $u^*Xu \geq 0$ for all $u \in \mathbb{K}^n$. Furthermore, I is the identity operator and I_n is the identity matrix of size n. The integer m' is defined after Assumption 1.

2 Geometric framework and near-optimality conditions

In this section, we present properties of the smooth geometry of (P) and approximate global optimality conditions for this problem. In covering these preliminaries, we largely parallel developments in [Boumal et al., 2016]. As argued in that reference, Assumption 1 implies that the search space $\mathcal{M}$ of (P) is a submanifold in $\mathbb{K}^{n\times k}$ of codimension m'. We can associate tangent spaces to a submanifold. Intuitively, the tangent space $\mathrm{T}_Y\mathcal{M}$ to the submanifold $\mathcal{M}$ at a point $Y \in \mathcal{M}$ is a subspace that best approximates $\mathcal{M}$ around Y, when the subspace origin is translated to Y. It is obtained by linearizing the equality constraints.

Lemma 2.1 (Boumal et al. [2018b, Lemma 2.1])**.** *Under Assumption 1, the tangent space at Y to $\mathcal{M}$ (2), denoted by $\mathrm{T}_Y\mathcal{M}$, is:*

$$\begin{aligned}\mathrm{T}_Y\mathcal{M} &= \left\{\dot{Y} \in \mathbb{K}^{n\times k} \colon \mathcal{A}(\dot{Y}Y^* + Y\dot{Y}^*) = 0\right\} \\ &= \left\{\dot{Y} \in \mathbb{K}^{n\times k} \colon \langle A_iY, \dot{Y}\rangle = 0 \textit{ for } i = 1, \ldots, m\right\}. \end{aligned} \tag{3}$$

By equipping each tangent space with a restriction of the inner product $\langle\cdot,\cdot\rangle$, we turn $\mathcal{M}$ into a Riemannian submanifold of $\mathbb{K}^{n\times k}$. We also introduce the orthogonal projector $\mathrm{Proj}_Y \colon \mathbb{K}^{n\times k} \to \mathrm{T}_Y\mathcal{M}$ which, given a matrix $Z \in \mathbb{K}^{n\times k}$, projects it to the tangent space $\mathrm{T}_Y\mathcal{M}$:

$$\mathrm{Proj}_Y Z := \underset{\dot{Y}\in \mathrm{T}_Y\mathcal{M}}{\mathrm{argmin}} \ \|\dot{Y} - Z\|. \tag{4}$$

This projector will be useful to phrase optimality conditions. It is characterized as follows.

Lemma 2.2 (Boumal et al. [2018b, Lemma 2.2])**.** *Under Assumption 1, the orthogonal projector admits the closed form*

$$\mathrm{Proj}_Y Z = Z - \mathcal{A}^*\left(G^\dagger \mathcal{A}(ZY^*)\right)Y,$$

where $\mathcal{A}^* \colon \mathbb{R}^m \to \mathbb{S}^{n\times n}$ is the adjoint of $\mathcal{A}$, G is a Gram matrix defined by $G_{ij} = \langle A_iY, A_jY\rangle$ (it is a function of Y), and $G^\dagger$ denotes the Moore–Penrose pseudo-inverse of G (differentiable in Y).

(See a proof in Appendix A.) To properly state the approximate first- and second-order necessary optimality conditions for (P), we further need the notions of *Riemannian gradient* and *Riemannian Hessian* on the manifold $\mathcal{M}$. We recall that (P) minimizes the function g, defined by

$$g(Y) = \langle CY, Y\rangle, \tag{5}$$

on the manifold $\mathcal{M}$. The Riemannian gradient of g at Y, $\mathrm{grad}\, g(Y)$, is the unique tangent vector at Y such that, for all tangent $\dot{Y}$, $\langle \mathrm{grad}\, g(Y), \dot{Y}\rangle = \langle \nabla g(Y), \dot{Y}\rangle$, with $\nabla g(Y) = 2CY$ the Euclidean (classical) gradient of g evaluated at Y. Intuitively, $\mathrm{grad}\, g(Y)$ is the tangent vector at Y that points in the steepest ascent direction for g as seen from the manifold's perspective. A classical result states that, for Riemannian submanifolds, the Riemannian gradient is given by the projection of the classical gradient to the tangent space [Absil et al., 2008, eq. (3.37)]:

$$\mathrm{grad}\, g(Y) = \mathrm{Proj}_Y(\nabla g(Y)) = 2\left(C - \mathcal{A}^*\left(G^\dagger\mathcal{A}(CYY^*)\right)\right)Y. \tag{6}$$

This leads us to define the matrix $S \in \mathbb{S}^{n\times n}$ which plays a key role to guarantee approximate global optimality for problem (P), as discussed in Section 3:

$$S = S(Y) = C - \mathcal{A}^*(\mu) = C - \sum_{i=1}^m \mu_i A_i, \tag{7}$$

where $\mu = \mu(Y) = G^\dagger \mathcal{A}(CYY^*)$. We can write the Riemannian gradient of g evaluated at Y as

$$\operatorname{grad} g(Y) = 2SY. \tag{8}$$

The Riemannian gradient enables us to define an approximate first-order necessary optimality condition below. To define the approximate second-order necessary optimality condition, we need to introduce the notion of Riemannian Hessian. The Riemannian Hessian of g at Y is a self-adjoint operator on the tangent space at Y obtained as the projection of the derivative of the Riemannian gradient vector field [Absil et al., 2008, eq. (5.15)]. Boumal et al. [2018b] give a closed form expression for the Riemannian Hessian of g at Y:

$$\operatorname{Hess} g(Y)[\dot{Y}] = 2 \cdot \operatorname{Proj}_Y(S\dot{Y}). \tag{9}$$

We can now formally define the approximate necessary optimality conditions for problem (P).

Definition 2.1 (ε_g-FOSP)**.** *$Y \in \mathcal{M}$ is an ε_g–first-order stationary point for* (P) *if the norm of the Riemannian gradient of g at Y almost vanishes, specifically,*

$$\|\operatorname{grad} g(Y)\| = \|2SY\| \leq \varepsilon_g,$$

where S is defined as in equation (7)*.*

Definition 2.2 ($(\varepsilon_g, \varepsilon_H)$-SOSP)**.** *$Y \in \mathcal{M}$ is an $(\varepsilon_g, \varepsilon_H)$–second-order stationary point for* (P) *if it is an ε_g–first-order stationary point and the Riemannian Hessian of g at Y is almost positive semidefinite, specifically,*

$$\forall \dot{Y} \in \mathrm{T}_Y\mathcal{M}, \qquad \frac{1}{2}\left\langle \dot{Y}, \operatorname{Hess} g(Y)[\dot{Y}] \right\rangle = \langle \dot{Y}, S\dot{Y} \rangle \geq -\varepsilon_H \|\dot{Y}\|^2.$$

From these definitions, it is clear that S encapsulates the approximate optimality conditions of problem (P).

3 Approximate second-order points and smoothed analysis

We state our main results formally in this section. As announced, following [Bhojanapalli et al., 2018], we resort to smoothed analysis [Spielman and Teng, 2004]. To this end, we consider perturbations of the cost matrix C of (SDP) by a *Gaussian Wigner matrix*. Intuitively, smoothed analysis tells us how large the variance of the perturbation should be in order to obtain a new SDP which, with high probability, is sufficiently distant from any pathological case. We start by formally defining the notion of Gaussian Wigner matrix, following [Ben Arous and Guionnet, 2010].

Definition 3.1 (Gaussian Wigner matrix)**.** *The random matrix $W = W^*$ in $\mathbb{S}^{n\times n}$ is a* Gaussian Wigner matrix *with variance σ_W^2 if its entries on and above the diagonal are independent, zero-mean Gaussian variables (real or complex depending on context) with variance σ_W^2.*

Besides Assumption 1, another important assumption for our results is that the search space $\mathcal{C}$ (1) of (SDP) is compact. In that scenario, there exists a finite constant R such that

$$\forall X \in \mathcal{C}, \quad \operatorname{Tr}(X) \leq R. \tag{10}$$

Thus, for all $Y \in \mathcal{M}$, $\|Y\|^2 = \operatorname{Tr}(YY^*) \leq R$. Another consequence of compactness of $\mathcal{C}$ is that the operator $\mathcal{A}^* \circ G^\dagger \circ \mathcal{A}$ is uniformly bounded, that is, there exists a finite constant K such that

$$\forall Y \in \mathcal{M}, \quad \|\mathcal{A}^* \circ G^\dagger \circ \mathcal{A}\|_{\mathrm{op}} \leq K, \tag{11}$$

where $G^\dagger$ is a continuous function of Y as in Lemma 2.2. We give explicit expressions for the constants R and K for the case of phase retrieval in Section 4.

We now state the main theorem, whose proof is in Appendix E.

Theorem 3.1. *Let Assumption 1 hold for* (SDP) *with cost matrix $C \in \mathbb{S}^{n\times n}$ and m constraints. Assume $\mathcal{C}$* (1) *is compact, and let R and K be as in* (10) *and* (11)*. Let W be a Gaussian Wigner matrix with variance σ_W^2 and let $\delta \in (0,1)$ be any tolerance. Define κ as:*

$$\kappa = \kappa(R, K, C, n, \sigma_W) = RK\left(\|C\|_{\mathrm{op}} + 3\sigma_W\sqrt{n}\right). \tag{12}$$

There exists a universal constant c_0 such that, if the rank k for the low-rank problem (P) satisfies

$$k \geq 3\left[\log(n) + \sqrt{\log(1/\delta)} + \sqrt{m \cdot \log\left(1 + \frac{6\kappa\sqrt{c_0 n}}{\sigma_W}\right)}\right], \tag{13}$$

then, with probability at least $1 - \delta - e^{-\frac{n}{2}}$ on the random matrix W, any $(\varepsilon_g, \varepsilon_H)$-SOSP $Y \in \mathbb{K}^{n\times k}$ of (P) with perturbed cost matrix $C + W$ has bounded optimality gap:

$$0 \leq g(Y) - f^\star \leq (\varepsilon_H + \varepsilon_g^2 \eta)R + \frac{\varepsilon_g}{2}\sqrt{R}, \tag{14}$$

with g the cost function of (P), $f^\star$ the optimal value of (SDP) (both perturbed), and

$$\eta = \eta(R, K, C, n, m, \sigma_W) = \frac{c_0 n K(2 + KR)^2 \left(\|C\|_{\mathrm{op}} + 3\sigma_W\sqrt{n}\right)}{9m\sigma_W^2 \log\left(1 + \frac{6\kappa\sqrt{c_0 n}}{\sigma_W}\right)}. \tag{15}$$

This result indicates that, as long as the rank k is on the order of $\sqrt{m}$ (up to logarithmic factors), the optimality gap in the *perturbed* problem is small if a sufficiently good *approximate* second-order point is computed. Since (SDP) may admit a unique solution of rank as large as $\Theta(\sqrt{m})$ (see for example [Laurent and Poljak, 1996, Thm. 3.1(ii)] for the Max-Cut SDP), we conclude that the scaling of k with respect to m in Theorem 3.1 is essentially optimal.

There is an incentive to pick σ_W small, since the optimality gap is phrased in terms of the perturbed problem. As expected though, taking σ_W small comes at a price. Specifically, the required rank k scales with $\sqrt{\log(1/\sigma_W)}$, so that a smaller σ_W may require k to be a larger multiple of $\sqrt{m}$. Furthermore, the optimality gap is bounded in terms of η with a dependence in $\varepsilon_g^2/\sigma_W^2$; this may force us to compute more accurate approximate second-order points (smaller ε_g) for a similar guarantee when σ_W is smaller: see also Corollary 3.1 below.

As announced, the theorem rests on two arguments which we now present—a probabilistic one, and a deterministic one:

1. Probabilistic argument: In the smoothed analysis framework, we show, for k large enough, that ε_g-FOSPs of (P) have their smallest singular value near zero, with high probability upon perturbation of C. This implies that such points are almost column-rank deficient.
2. Deterministic argument: If Y is an $(\varepsilon_g, \varepsilon_H)$-SOSP of (P) and it is almost column-rank deficient, then the matrix $S(Y)$ defined in equation (7) is almost positive semidefinite. From there, we can derive a bound on the optimality gap.

Formal statements for both follow, building on the notation in Theorem 3.1. Proofs are in Appendices C and D, with supporting lemmas in Appendix B.

Lemma 3.1. *Let Assumption 1 hold for (SDP). Assume $\mathcal{C}$ (1) is compact. Let W be a Gaussian Wigner matrix with variance σ_W^2 and let $\delta \in (0, 1)$ be any tolerance. There exists a universal constant c_0 such that, if the rank k for the low-rank problem (P) is lower bounded as in (13), then, with probability at least $1 - \delta - e^{-\frac{n}{2}}$ on the random matrix W, we have*

$$\|W\|_{\mathrm{op}} \leq 3\sigma_W\sqrt{n},$$

and furthermore: any ε_g-FOSP $Y \in \mathbb{K}^{n\times k}$ of (P) with perturbed cost matrix $C + W$ satisfies

$$\sigma_k(Y) \leq \frac{\varepsilon_g}{\sigma_W}\frac{\sqrt{c_0 n}}{k},$$

where $\sigma_k(Y)$ is the kth singular value of the matrix Y.

Lemma 3.2. *Let Assumption 1 hold for (SDP) with cost matrix C. Assume $\mathcal{C}$ is compact. Let $Y \in \mathbb{K}^{n\times k}$ be an $(\varepsilon_g, \varepsilon_H)$-SOSP of (P) (for any k). Then, the smallest eigenvalue of $S = S(Y)$ (7) is bounded below as*

$$\lambda_{\min}(S) \geq -\varepsilon_H - \zeta\|C\|_{\mathrm{op}} \cdot \sigma_k^2(Y),$$

where $\zeta = K(2 + KR)^2$ with R, K as in (10) and (11), and $\sigma_k(Y)$ is the kth singular value of Y (it is zero if $k > n$). This holds deterministically for any cost matrix C.

Combining the two above lemmas, the key step in the proof of Theorem 3.1 is to deduce a bound on the optimality gap from a bound on the smallest eigenvalue of S: see Appendix E.

We have shown in Theorem 3.1 that a perturbed version of (P) can be approximately solved to global optimality, with high probability on the perturbation. In the corollary below, we further bound the optimality gap at the approximate solution of the perturbed problem with respect to the original, unperturbed problem. The proof is in Appendix F.

Corollary 3.1. *Assume $\mathcal{C}$ is compact and let R be as defined in (10). Let $X \in \mathcal{C}$ be an approximate solution for (SDP) with perturbed cost matrix $C + W$, so that the optimality gap in the perturbed problem is bounded by ε_f. Let $f^\star$ denote the optimal value of the unperturbed problem (SDP), with cost matrix C. Then, the optimality gap for X with respect to the unperturbed problem is bounded as:*

$$0 \leq \langle C, X\rangle - f^\star \leq \varepsilon_f + 2\|W\|_{\mathrm{op}} R.$$

Under the conditions of Theorem 3.1, with the prescribed probability, ε_f and $\|W\|_{\mathrm{op}}$ can be bounded so that for an $(\varepsilon_g, \varepsilon_H)$-SOSP Y of the perturbed problem (P) we have:

$$0 \leq \langle CY, Y\rangle - f^\star \leq (\varepsilon_H + \varepsilon_g^2 \eta)R + \frac{\varepsilon_g}{2}\sqrt{R} + 6\sigma_W \sqrt{n} R,$$

where η is as defined in (15) and σ_W^2 is the variance of the Wigner perturbation W.

4 Applications

The approximate global optimality results established in the previous section can be applied to deduce guarantees on the quality of ASOSPs of the low-rank factorization for a number of SDPs that appear in machine learning problems. Of particular interest, we focus on the phase retrieval problem. This problem consists in retrieving a signal $z \in \mathbb{C}^d$ from n amplitude measurements $b = |Az| \in \mathbb{R}^n_+$ (the absolute value of vector Az is taken entry-wise). If we can recover the complex phases of Az, then z can be estimated through linear least-squares. Following this approach, Waldspurger et al. [2015] argue that this task can be modeled as the following non-convex problem:

$$\begin{aligned} \min_{u \in \mathbb{C}^n} \quad & u^* C u \\ \text{subject to} \quad & |u_i| = 1, \text{ for } i = 1, \ldots, n, \end{aligned} \tag{PR}$$

where $C = \mathrm{diag}(b)(I - AA^\dagger)\mathrm{diag}(b)$ and $\mathrm{diag}\colon \mathbb{R}^n \to \mathbb{H}^{n\times n}$ maps a vector to the corresponding diagonal matrix. The classical relaxation is to rewrite the above in terms of $X = uu^*$ (lifting) without enforcing $\mathrm{rank}(X) = 1$, leading to a complex SDP which Waldspurger et al. [2015] call PhaseCut:

$$\begin{aligned} \min_{X \in \mathbb{H}^{n\times n}} \quad & \langle C, X\rangle \\ \text{subject to} \quad & \mathrm{diag}(X) = 1, \\ & X \succeq 0. \end{aligned} \tag{PhaseCut}$$

The same SDP relaxation also applies to a problem called angular synchronization [Singer, 2011]. The Burer–Monteiro factorization of (PhaseCut) is an optimization problem over a matrix $Y \in \mathbb{C}^{n\times k}$ as follows:

$$\begin{aligned} \min_{Y \in \mathbb{C}^{n\times k}} \quad & \langle CY, Y\rangle \\ \text{subject to} \quad & \mathrm{diag}(YY^*) = 1. \end{aligned} \tag{PhaseCut-BM}$$

For a feasible Y, each row has unit norm: the search space is a Cartesian product of spheres in $\mathbb{C}^k$, which is a smooth manifold. We can check that Assumption 1 holds for all $k \geq 1$. Furthermore, the feasible space of the SDP is compact. Therefore, Theorem 3.1 applies.

In this setting, $\mathrm{Tr}(X) = n$ for all feasible X, and $\|\mathcal{A}^* \circ G^\dagger \circ \mathcal{A}\|_{\mathrm{op}} = 1$ for all feasible Y (because $G = G(Y) = I_m$ for all feasible Y—see Lemma 2.2—and $\mathcal{A}^* \circ \mathcal{A}$ is an orthogonal projector from Hermitian matrices to diagonal matrices). For this reason, the constants defined in (10) and (11) can be set to $R = n$ and $K = 1$.

As a comparison, Mei et al. [2017] also provide an optimality gap for ASOSPs of (PhaseCut) without perturbation. Their result is more general in the sense that it holds for all possible values of k.

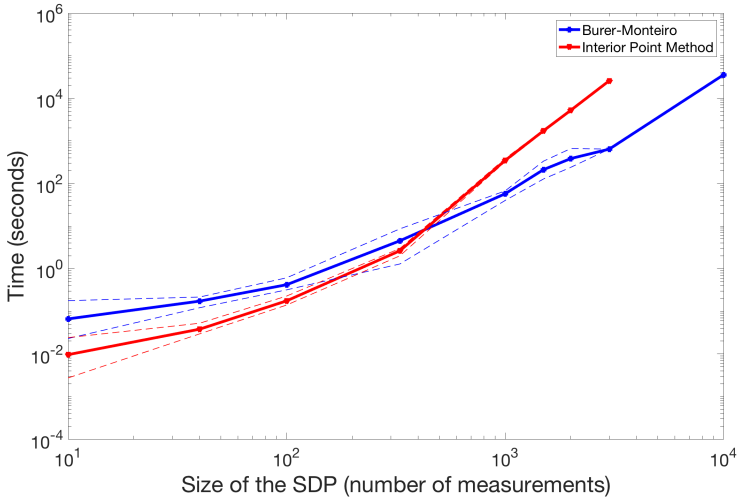

Figure 1: Computation time of the dedicated interior-point method (IPM) and of the Burer–Monteiro approach (BM) to solve (PhaseCut). For increasing values of n (horizontal axis), we display the computation time averaged over four independent realizations of the problem (vertical axis). The smallest and largest observed computation times are represented with dashed lines. At $n = 3000$, BM is about 40 times faster than IPM. For the largest value of n, IPM runs out of memory.

However, when k is large, it does not accurately capture the fact that SOSPs are optimal, thus incurring a larger bound on the optimality gap of ASOSPs. In contrast, our bounds do show that for k large enough, as $\varepsilon_g, \varepsilon_H$ go to zero, the optimality gap goes to zero, with the trade-off that they do so for a perturbed problem (though see Corollary 3.1), with high probability.

Numerical Experiments

We present the empirical performance of the low-rank approach in the case of (PhaseCut). We compare it with a dedicated interior-point method (IPM) implemented by Helmberg et al. [1996] for real SDPs and adapted to phase retrieval as done by Waldspurger et al. [2015]. This adaptation involves splitting the real and the imaginary parts of the variables in (PhaseCut) and forming an equivalent real SDP with double the dimension. The Burer–Monteiro approach (BM) is implemented in complex form directly using Manopt, a toolbox for optimization on manifolds [Boumal et al., 2014]. In particular, a Riemannian Trust-Region method (RTR) is used [Absil et al., 2007]. Theory supports that these methods can return an ASOSP in a finite number of iterations [Boumal et al., 2018a]. We stress that the SDP is *not* perturbed in these experiments: the role of the perturbation in the analysis is to understand why the low-rank approach is so successful in practice despite the existence of pathological cases. In practice, we do not expect to encounter pathological cases.

Our numerical experiment setup is as follows. We seek to recover a signal of dimension d, $z \in \mathbb{C}^d$, from n measurements encoded in the vector $b \in \mathbb{R}^n_+$ such that $b = |Az| + \epsilon$, where $A \in \mathbb{C}^{n\times d}$ is the sensing matrix and $\epsilon \sim \mathcal{N}(0, \mathrm{I}_d)$ is standard Gaussian noise. For the numerical experiments, we generate the vectors z as complex random vectors with i.i.d. standard Gaussian entries, and we randomly generate the complex sensing matrices A also with i.i.d. standard Gaussian entries. We do so for values of d ranging from 10 to 1000, and always for $n = 10d$ (that is, there are 10 magnitude measurements per unknown complex coefficient, which is an oversampling factor of 5.) Lastly, we generate the measurement vectors b as described above and we cap its values from below at 0.01 in order to avoid small (or even negative) magnitude measurements.

For n up to 3000, both methods solve the same problem, and indeed produce the same answer up to small discrepancies. The BM approach is more accurate, at least in satisfying the constraints, and, for $n = 3000$, it is also about 40 times faster than IPM. BM is run with $k = \sqrt{n}$ (rounded up), which is expected to be generically sufficient to include the global optimum of the SDP (as confirmed in practice). For larger values of n, the IPM ran into memory issues and we had to abort the process.

5 Conclusion

We considered the low-rank (or Burer–Monteiro) approach to solve equality-constrained SDPs. Our key assumptions are that (a) the search space of the SDP is compact, and (b) the search space of its low-rank version is smooth (the actual condition is slightly stronger). Under these assumptions, we proved using smoothed analysis that, provided $k = \tilde{\Omega}(\sqrt{m})$ where m is the number of constraints, if the cost matrix is perturbed randomly, with high probability, approximate second-order stationary points of the perturbed low-rank problem map to approximately optimal solutions of the perturbed SDP. We also related optimality gaps in the perturbed SDP to optimality gaps in the original SDP. Finally, we applied this result to an SDP relaxation of phase retrieval (also applicable to angular synchronization).

Acknowledgments

NB is partially supported by NSF award DMS-1719558.

References

E. Abbé. *Community Detection and Stochastic Block Models*, volume 14. Now Publishers inc., 2018. doi:10.1561/0100000067.

P.-A. Absil, C. Baker, and K. Gallivan. Trust-region methods on Riemannian manifolds. *Foundations of Computational Mathematics*, 7(3):303–330, 2007. doi:10.1007/s10208-005-0179-9.

P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2008.

A. Bandeira, N. Boumal, and A. Singer. Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *Mathematical Programming*, 163(1):145–167, 2017. doi:10.1007/s10107-016-1059-6.

A. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete & Computational Geometry*, 13(1):189–202, 1995. doi:10.1007/BF02574037.

G. Ben Arous and A. Guionnet. Wigner matrices. In *Handbook on Random Matrices*, pages 433–451. Oxford University Press, 2010.

R. Bhatia. *Positive definite matrices*. Princeton University Press, 2007.

S. Bhojanapalli, N. Boumal, P. Jain, and P. Netrapalli. Smoothed analysis for low-rank solutions to semidefinite programs in quadratic penalty form. In S. Bubeck, V. Perchet, and P. Rigollet, editors, *Proceedings of the 31st Conference On Learning Theory*, volume 75 of *Proceedings of Machine Learning Research*, pages 3243–3270. PMLR, 06–09 Jul 2018. URL http://proceedings.mlr.press/v75/bhojanapalli18a.html.

N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre. Manopt, a Matlab toolbox for optimization on manifolds. *Journal of Machine Learning Research*, 15:1455–1459, 2014. URL https://www.manopt.org.

N. Boumal, V. Voroninski, and A. Bandeira. The non-convex Burer-Monteiro approach works on smooth semidefinite programs. In *Advances in Neural Information Processing Systems*, pages 2757–2765, 2016.

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. Bandeira. Deterministic guarantees for Burer–Monteiro factorizations of smooth semidefinite programs. *To appear in Communications on Pure and Applied Mathematics*, 2018b.

S. Burer and R. D. Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003.

S. Burer and R. D. Monteiro. Local minima and convergence in low-rank semidefinite programming. *Mathematical Programming*, 103(3):427–444, 2005.

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

T. Carson, D. G. Mixon, and S. Villar. Manifold optimization for k-means clustering. In *2017 International Conference on Sampling Theory and Applications (SampTA)*, pages 73–77, July 2017. doi:10.1109/SAMPTA.2017.8024388.

D. Garber. Faster projection-free convex optimization over the spectrahedron. *Advances in Neural Information Processing Systems*, pages 874–882, 2016.

D. Garber and E. Hazan. Sublinear time algorithms for approximate semidefinite programming. *Mathematical Programming: Series A and B archive*, 158:329–361, 2016.

G. H. Golub and V. Pereyra. The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM Journal on Numerical Analysis*, 10(2):413–432, 1973. doi:10.1137/0710036.

E. Hazan. Sparse approximate solutions to semidefinite programs. *LATIN 2008: Theoretical Informatics*, pages 306–316, 2008.

C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM Journal on Optimization*, 6(2):342–361, 1996.

M. Journee, F. Bach, P.-A. Absil, and R. Sepulchre. Low-rank optimization on the cone of positive semidefinite matrices. *SIAM Journal on Optimization*, 20(5):2327–2351, 2010.

B. Kulis, A. C. Surendran, and J. C. Platt. Fast low-rank semidefinite programming for embedding and clustering. In *Artificial Intelligence and Statistics*, pages 235–242, 2007.

G. R. Lanckriet, N. Cristianini, P. Bartlett, L. E. Ghaoui, and M. I. Jordan. Learning the kernel matrix with semidefinite programming. *Journal of Machine learning research*, 5(Jan):27–72, 2004.

S. Laue. A hybrid algorithm for convex semidefinite optimization. *ICML*, pages 177–184, 2012.

M. Laurent and S. Poljak. On the facial structure of the set of correlation matrices. *SIAM Journal on Matrix Analysis and Applications*, 17(3):530–547, 1996. doi:10.1137/0617031.

S. Mei, T. Misiakiewicz, A. Montanari, and R. I. Oliveira. Solving sdps for synchronization and maxcut problems via the grothendieck inequality. In *Proceedings of the 30th Conference on Learning Theory, COLT 2017, Amsterdam, The Netherlands, 7-10 July 2017*, pages 1476–1515, 2017. URL `http://proceedings.mlr.press/v65/mei17a.html`.

H. H. Nguyen. Random matrices: Overcrowding estimates for the spectrum. *Journal of Functional Analysis*, 275(8):2197–2224, 2018. doi:10.1016/j.jfa.2018.06.010.

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.

P. Rigollet and J.-C. Hütter. High dimensional statistics, 2017. URL `http://www-math.mit.edu/~rigollet/PDFs/RigNotes17.pdf`.

A. Singer. Angular synchronization by eigenvectors and semidefinite programming. *Applied and computational harmonic analysis*, 30(1):20–36, 2011.

D. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *Journal of the ACM*, 51(3):385–463, May 2004. doi:10.1145/990308.990310.

I. Waldspurger, A. d'Aspremont, and S. Mallat. Phase recovery, maxcut and complex semidefinite programming. *Mathematical Programming*, 149(1-2):47–81, 2015.

A. Yurtsever, M. Udell, J. A. Tropp, and V. Cevher. Sketchy decisions: Convex low-rank matrix optimization with optimal storage. *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, 54:1188–1196, 2017.

A Proof of Lemma 2.2

We follow the proof of [Boumal et al., 2018b, Lemma 2.2] and reproduce it here to be self-contained, and also because the reference treats only the real case; writing the proof here explicitly allows to verify that, indeed, all steps go through for the complex case as well.

Orthogonal projection is along the normal space, so that $\mathrm{Proj}_Y Z$ is in $\mathrm{T}_Y\mathcal{M}$ (3) and $Z - \mathrm{Proj}_Y Z$ is in $\mathrm{N}_Y\mathcal{M}$, where the normal space at Y is (using Assumption 1)

$$\mathrm{N}_Y\mathcal{M} = \left\{ Z \in \mathbb{K}^{n\times k} : \langle Z, \dot{Y}\rangle = 0 \; \forall \dot{Y} \in \mathrm{T}_Y\mathcal{M} \right\} = \mathrm{span}\{A_1 Y, \ldots, A_m Y\}. \tag{16}$$

From the latter we infer there exists $\mu \in \mathbb{R}^m$ such that

$$Z - \mathrm{Proj}_Y Z = \sum_{i=1}^m \mu_i A_i Y = \mathcal{A}^*(\mu) Y,$$

since the adjoint of $\mathcal{A}$ is $\mathcal{A}^*(\mu) = \mu_1 A_1 + \cdots + \mu_m A_m$. Multiply on the right by Y^* and apply $\mathcal{A}$ to obtain

$$\mathcal{A}(ZY^*) = \mathcal{A}(\mathcal{A}^*(\mu) YY^*),$$

where we used $\mathcal{A}(\mathrm{Proj}_Y(Z)Y^*) = 0$ since $\mathrm{Proj}_Y(Z) \in \mathrm{T}_Y\mathcal{M}$. The right-hand side expands into

$$\mathcal{A}(\mathcal{A}^*(\mu)YY^*)_i = \left\langle A_i, \sum_{j=1}^m \mu_j A_j YY^* \right\rangle = \sum_{j=1}^m \langle A_i Y, A_j Y\rangle \, \mu_j = (G\mu)_i,$$

where G is a real, positive semidefinite matrix of size m defined by $G_{ij} = \langle A_i Y, A_j Y\rangle$. By construction, this system of equations in μ has at least one solution; we single out $\mu = G^\dagger \mathcal{A}(ZY^*)$, where $G^\dagger$ is the Moore–Penrose pseudo-inverse of G. The function $Y \mapsto G^\dagger$ is continuous and differentiable at $Y \in \mathcal{M}$ provided G has constant rank in an open neighborhood of Y in $\mathbb{K}^{n\times k}$ [Golub and Pereyra, 1973, Thm 4.3], which is the case for all $Y \in \mathcal{M}$ under Assumption 1.

B Lower-bound for smallest singular values

This appendix provides supporting results necessary for Appendix C, which is devoted to the proof of Lemma 3.1. The statements we need are established for $\mathbb{K} = \mathbb{R}$ in [Bhojanapalli et al., 2018, Cor. 5, Lem. 7]. Here we give the corresponding statements for $\mathbb{K} = \mathbb{C}$: the proofs are essentially the same.

We first state a special case of Corollary 1.17 from [Nguyen, 2018]. Here, $N_I(X)$ denotes the number of eigenvalues of $X \in \mathbb{S}^{n\times n}$ in the real interval I. (Note that the reference covers the real case in its main statement, and addresses the complex case later on as a remark.) For Gaussian Wigner matrices, we follow Definition 3.1. Furthermore, $\mathbb{P}\{E\}$ denotes the probability of event E.

Corollary B.1. *Let $\overline{M}$ be a deterministic Hermitian matrix of size n. Let $\overline{W}$ be a Gaussian Wigner matrix with variance 1. Then, for any given $0 < \gamma < 1$, there exists a constant $c = c(\gamma)$ such that for any $\varepsilon > 0$ and $k \geq 1$, with I being the interval $\left[-\frac{\varepsilon k}{\sqrt{n}}, \frac{\varepsilon k}{\sqrt{n}}\right]$,*

$$\mathbb{P}\left\{ N_I(\overline{M} + \overline{W}) \geq k \right\} \leq n^k \left(\frac{c\varepsilon}{\sqrt{2\pi}} \right)^{(1-\gamma)k^2/2}.$$

The next lemma follows easily—the original proof for $\mathbb{K} = \mathbb{R}$ is in [Bhojanapalli et al., 2018, Lem. 7].

Lemma B.1. *Let M be a deterministic Hermitian matrix of size n. Let W be a complex Gaussian Wigner matrix of size n with variance σ_W^2, independent of M. There exists an absolute constant c_0 such that:*

$$\mathbb{P}\left\{ \sum_{i=1}^k \sigma_{n-(i-1)}(M+W)^2 < \frac{k^2\sigma_W^2}{c_0 n} \right\} \leq \exp\left(-\frac{k^2}{8}\log(8\pi) + k\log(n) \right).$$

Proof. In our case, the entries of W have variance $\mathbb{E}[|W_{i,j}|^2] = \sigma_W^2$. Thus, set $W = \sigma_W \overline{W}$ and $M = \sigma_W \overline{M}$. From Corollary B.1, we get

$$N_{\sigma_W I}(M + W) = N_I(\overline{M} + \overline{W}) < k$$

with probability at least $1 - n^k \left(\frac{c\varepsilon}{\sqrt{2\pi}}\right)^{(1-\gamma)k^2/2}$. In this event, $\sigma_{n-(k-1)}(\overline{M} + \overline{W}) \geq \frac{\varepsilon k}{\sqrt{n}} \sigma_W$. With the choices $\gamma = \frac{1}{2}$ and $\varepsilon = \frac{1}{2c}$, we get that

$$\sigma_{n-(k-1)}(\overline{M} + \overline{W}) \geq \frac{k}{2c\sqrt{n}} \sigma_W$$

with probability at least $1 - \exp(-\frac{k^2}{8}\log(8\pi) + k\log(n))$. In that event,

$$\sum_{i=1}^{k} \sigma_{n-(i-1)}(\overline{M} + \overline{W})^2 \geq \sigma_{n-(k-1)}(\overline{M} + \overline{W})^2 \geq \frac{k^2}{c_0 n} \sigma_W^2$$

for some absolute constant $c_0 = 4c^2$. □

C Proof of Lemma 3.1

This section builds on a result from [Bhojanapalli et al., 2018], where a similar statement was made under different assumptions. The proof follows closely the developments therein, with appropriate changes. Using (8), Y is an ε_g-FOSP of the perturbed problem if and only if $\|2SY\| \leq \varepsilon_g$ with $S = M + W$ and

$$M = C - (\mathcal{A}^* \circ G^\dagger \circ \mathcal{A})\left((C + W)YY^*\right). \tag{17}$$

Let $Y = P\Sigma Q^*$ be a thin SVD of Y, where P is $n \times k$ with orthonormal columns (assuming without loss of generality $k \leq n$, as otherwise $\sigma_k(Y) = 0$ deterministically) and Q is $k \times k$ orthogonal. Then,

$$\begin{aligned}
\varepsilon_g \geq \|2SY\| &= \|2(M + W)Y\| \\
&\geq 2\sigma_k(Y)\|(M + W)P\| \\
&\geq 2\sigma_k(Y)\sqrt{\sum_{i=1}^{k} \sigma_{n-(i-1)}(M + W)^2}.
\end{aligned}$$

Thus, we control the smallest singular value of Y in terms of ε and the k smallest singular values of $M + W$:

$$\sigma_k(Y) \leq \frac{\varepsilon_g}{2\sqrt{\sum_{i=1}^{k} \sigma_{n-(i-1)}(M + W)^2}}. \tag{18}$$

Given that M is not statistically independent of W, we are not able to directly apply Lemma B.1. Indeed, M depends on W and on Y, and Y itself is an ε_g-FOSP of the perturbed problem: a feature which depends on W. To tackle this issue, we cover the set of possible Ms with a net. Lemma B.1 provides a bound for each $\bar{M}$ in this net. By union bound, we can extend the lemma for all $\bar{M}$. By taking a sufficiently dense net, we then infer that M is necessarily close to one of these $\bar{M}$'s and conclude.

To this end, we first control $\|M - C\|$ using the definitions of R (10) and K (11):

$$\begin{aligned}
\|M - C\| &= \|\mathcal{A}^* \circ G^\dagger \circ \mathcal{A}\left((C + W)YY^*\right)\| \\
&\leq \|\mathcal{A}^* \circ G^\dagger \circ \mathcal{A}\|_{\mathrm{op}} \|C + W\|_{\mathrm{op}} \|YY^*\| \\
&\leq K(\|C\|_{\mathrm{op}} + \|W\|_{\mathrm{op}})R.
\end{aligned}$$

Since W is a Gaussian Wigner matrix with variance σ_W^2, it is a well known fact (see for instance[2] Part 1 of Appendix A in [Bandeira et al., 2017]) that, with probability at least $1 - e^{-\frac{n}{2}}$,

$$\|W\|_{\mathrm{op}} \leq 3\sigma_W \sqrt{n}. \tag{19}$$

[2]The reference proves the statement for complex matrices with diagonal entries equal to zero. That proof can easily be adapted to the definition of Wigner matrices used in this paper, both real and complex.

Hence, with probability at least $1 - e^{-\frac{n}{2}}$,

$$\|M - C\| \leq RK(\|C\|_{\text{op}} + 3\sigma_W\sqrt{n}) \triangleq \kappa,$$

where we recover κ as defined in (12).

As a result, M lies in a ball of center C and radius κ. Moreover, from (17), we remark that M lives in an affine subspace of dimension $\text{rank}(\mathcal{A}^* \circ G^\dagger \circ \mathcal{A}) = \text{rank}(\mathcal{A})$. A unit ball in Frobenius norm in d dimensions admits an ε-net of $\left(1 + \frac{3}{\varepsilon}\right)^d$ points (see for instance Lemma 1.18 in [Rigollet and Hütter, 2017]).[3] Thus, we pick a $\frac{k\sigma_W}{2\kappa\sqrt{c_0 n}}$-net on the unit ball with $\left(1 + \frac{6\kappa\sqrt{c_0 n}}{k\sigma_W}\right)^{\text{rank}(\mathcal{A})}$ points. Rescaling by a factor κ gives a $\frac{k\sigma_W}{2\sqrt{c_0 n}}$-net of a ball of radius κ centered at zero. Hence, for any M as in (17) there necessarily exists a point $\bar{M}$ in the net satisfying:

$$\|\bar{M} - M\| \leq \frac{k\sigma_W}{2\sqrt{c_0 n}}. \tag{20}$$

Let $T\colon \mathbb{S}^{n\times n} \to \mathbb{R}^k$ be defined by $T_q(A) = (\sigma_{n-q+1}(A), \ldots, \sigma_n(A))^\top$, that is: T extracts the q smallest singular values of A, in order. Then, by using the result from Exercise IV.3.5. in [Bhatia, 2007] in the first inequality,[4] we have:

$$\begin{aligned}
\|\bar{M} - M\| &= \|(\bar{M} + W) - (M + W)\| \\
&= \sqrt{\sum_{i=1}^n \sigma_i^2\left((\bar{M} + W) - (M + W)\right)} \\
&\geq \sqrt{\sum_{i=1}^n \left(\sigma_i(\bar{M} + W) - \sigma_i(M + W)\right)^2} \\
&= \|T_n(\bar{M} + W) - T_n(M + W)\| \\
&\geq \|T_k(\bar{M} + W) - T_k(M + W)\| \\
&\geq \|T_k(\bar{M} + W)\| - \|T_k(M + W)\|,
\end{aligned}$$

where we used the triangular inequality in the last inequality. Thus, rearranging we obtain

$$\sqrt{\sum_{i=1}^k \sigma_{n-(i-1)}(M + W)^2} \geq \sqrt{\sum_{i=1}^k \sigma_{n-(i-1)}(\bar{M} + W)^2} - \|\bar{M} - M\|. \tag{21}$$

Taking a union bound for Lemma B.1 over each $\bar{M}$ in the net, we get that

$$\sqrt{\sum_{i=1}^k \sigma_{n-(i-1)}(\bar{M} + W)^2} \geq \frac{k\sigma_W}{\sqrt{c_0 n}} \tag{22}$$

holds with probability at least

$$1 - \exp\left(-\frac{k^2}{8}\log(8\pi) + k\log(n) + \text{rank}(\mathcal{A}) \cdot \log\left(1 + \frac{6\kappa\sqrt{c_0 n}}{k\sigma_W}\right)\right). \tag{23}$$

Combining (20), (21) and (22), we conclude that

$$\sqrt{\sum_{i=1}^k \sigma_{n-(i-1)}(M + W)^2} \geq \frac{k\sigma_W}{2\sqrt{c_0 n}} \tag{24}$$

[3] The lemma in the reference shows that for any $\varepsilon \in (0, 1)$ the cardinality of one such ε-net is bounded by $(3/\varepsilon)^d$. Furthermore, for $\varepsilon \geq 1$, there is an obvious ε-net of cardinality one, comprising just the origin. Hence, for any $\varepsilon > 0$, it is possible so find an ε-net of cardinality at most $\max\left(1, (3/\varepsilon)^d\right) \leq (1 + 3/\varepsilon)^d$.

[4] The same result can be obtained by using Theorem IV.2.14 in the reference. In this setting, one considers the function $F(A) = -\sum_{i=1}^n \sigma_i^2(A)$; then, use the subadditive property of F, i.e., $F(A + B) \leq F(A) + F(B)$, and define $A = \bar{M} + W$ and $B = -(M + W)$.

holds with probability bounded as in (23). Combining with (18), we obtain

$$\sigma_k(Y) \leq \frac{\varepsilon_g}{\sigma_W} \frac{\sqrt{c_0 n}}{k}$$

as desired. It remains to discuss the probability of success, which we do below.

Inside the log in (23), we can safely replace k with 1, as this only hurts the probability. Then, the result holds with probability at least

$$1 - \exp\left(-\frac{k^2}{8}\log(8\pi) + k\log(n) + \operatorname{rank}(\mathcal{A}) \cdot \log\left(1 + \frac{6\kappa\sqrt{c_0 n}}{\sigma_W}\right)\right).$$

We would like to constrain k such that the exponential part is bounded by δ. In this fashion, taking a union bound with event (19), we will get an overall probability of success of at least $1 - \delta - e^{-\frac{n}{2}}$. Equivalently, k must satisfy the quadratic inequality

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

with $a, b > 0$, $c \geq 0$ defined by $a = \frac{\log(8\pi)}{8}, b = \log(n), c = \operatorname{rank}(\mathcal{A}) \cdot \log\left(1 + \frac{6\kappa\sqrt{c_0 n}}{\sigma_W}\right)$. This quadratic inequality can be rewritten as:

$$ak^2 - bk - c' \geq 0,$$

with $c' = c + \log(1/\delta)$. This quadratic has two distinct real roots, one positive and one negative:

$$\frac{b \pm \sqrt{b^2 + 4ac'}}{2a}.$$

Since k is positive, we deduce that k needs to be larger than the positive root. The latter obeys the following inequality:[5]

$$\frac{b + \sqrt{b^2 + 4ac'}}{2a} \leq \frac{b + b + 2\sqrt{ac'}}{2a} = \frac{b + \sqrt{ac'}}{a} = \frac{1}{a}b + \frac{1}{\sqrt{a}}\sqrt{c'}.$$

Since both $1/a$ and $1/\sqrt{a}$ are smaller than 3, it is sufficient to require

$$k \geq 3\left(b + \sqrt{c + \log(1/\delta)}\right).$$

Assuming $\delta \leq 1$, we can use the inequality in the footnote again and find that it is sufficient to have

$$k \geq 3\left(b + \sqrt{\log(1/\delta)} + \sqrt{c}\right).$$

Plugging in the definitions of b and c, we find the sufficient condition (with $\delta \leq 1$):

$$k \geq 3\left[\log(n) + \sqrt{\log(1/\delta)} + \sqrt{\operatorname{rank}(\mathcal{A}) \cdot \log\left(1 + \frac{6\kappa\sqrt{c_0 n}}{\sigma_W}\right)}\right].$$

Since $\operatorname{rank}(\mathcal{A}) \leq m$, we obtain the desired sufficient bound on k.

D Proof of Lemma 3.2

The Riemannian gradient and Hessian of the objective function g of (P) are respectively given by equations (8) and (9). Since Y is an $(\varepsilon_g, \varepsilon_H)$-SOSP, it holds for all $\dot{Y} \in \mathrm{T}_Y\mathcal{M}$ (3) with $\|\dot{Y}\| = 1$ that:

$$-\varepsilon_H \leq \frac{1}{2}\left\langle \dot{Y}, \operatorname{Hess} g(Y)[\dot{Y}]\right\rangle = \langle \dot{Y}, S\dot{Y}\rangle. \tag{25}$$

Our goal is to show that S is almost positive semidefinite. To this end, we first construct specific $\dot{Y}$'s to exploit the fact that Y is almost rank deficient. Let $z \in \mathbb{K}^k$ be a right singular vector of Y such that $\|Yz\| = \sigma_k(Y)$ and $\|z\| = 1$. For any $x \in \mathbb{K}^n$ with $\|x\| = 1$, we introduce $U = xz^*$. Decompose U

[5] We use that, for any $u, v \geq 0$, $\sqrt{u+v} \leq \sqrt{\sqrt{u}^2 + \sqrt{v}^2 + 2\sqrt{u}\sqrt{v}} = \sqrt{u} + \sqrt{v}$.

in two components: $U = U_T + U_{T^\perp}$, with U_T the component of U in the tangent space $\mathrm{T}_Y\mathcal{M}$ and $U_{T^\perp}$ the orthogonal component in $\mathrm{N}_Y\mathcal{M}$. Given that $\|z\| = 1$, using (25) with $\dot{Y} = U_T$, we have:

$$\begin{aligned}\langle x, Sx\rangle = \langle U, SU\rangle &= \langle U_T, SU_T\rangle + 2\langle U_{T^\perp}, SU_T\rangle + \langle U_{T^\perp}, SU_{T^\perp}\rangle \\ &\geq -\varepsilon_H\|U_T\|^2 + 2\langle U_{T^\perp}, SU_T\rangle + \langle U_{T^\perp}, SU_{T^\perp}\rangle \\ &\geq -\varepsilon_H + 2\langle U_{T^\perp}, SU_T\rangle + \langle U_{T^\perp}, SU_{T^\perp}\rangle \\ &= -\varepsilon_H + 2\langle U_{T^\perp}, SU\rangle - \langle U_{T^\perp}, SU_{T^\perp}\rangle, \end{aligned} \tag{26}$$

where we also used $\|U_T\|^2 \leq \|U\|^2 = 1$. We know by Lemma 2.2 that U_T can be written as:

$$U_T = \mathrm{Proj}_Y U = xz^* - \mathcal{A}^*\left(G^\dagger\mathcal{A}\left(xz^*Y^*\right)\right)Y. \tag{27}$$

Therefore, the component along the normal space, $U_{T^\perp}$, is:

$$U_{T^\perp} = \mathcal{A}^*\left(G^\dagger\mathcal{A}\left(xz^*Y^*\right)\right)Y. \tag{28}$$

Using (28), we can derive an upper bound on $\langle U_{T^\perp}, SU_{T^\perp}\rangle$. Indeed, by Cauchy–Schwarz we obtain:

$$\langle U_{T^\perp}, SU_{T^\perp}\rangle \leq \|U_{T^\perp}\|^2\|S\|_{\mathrm{op}}.$$

From the expression for S in (7) and the definitions of R (10) and K (11), the two factors are easily bounded since $\|xz^*Y^*\| = \|Yz\| = \sigma_k(Y)$:

$$\|U_{T^\perp}\| \leq \|\mathcal{A}^* \circ G^\dagger \circ \mathcal{A}\|_{\mathrm{op}}\|xz^*Y^*\|\|Y\| \leq K\sqrt{R}\cdot\sigma_k(Y),$$

and

$$\begin{aligned}\|S\|_{\mathrm{op}} &\leq \|C\|_{\mathrm{op}} + \|\mathcal{A}^*(G^\dagger\mathcal{A}(CYY^*))\|_{\mathrm{op}} \\ &\leq \|C\|_{\mathrm{op}} + \|\mathcal{A}^* \circ G^\dagger \circ \mathcal{A}\|_{\mathrm{op}}\|CYY^*\| \leq (1+KR)\|C\|_{\mathrm{op}}.\end{aligned}$$

Combining, we find the bound

$$\langle U_{T^\perp}, SU_{T^\perp}\rangle \leq K^2R(1+KR)\|C\|_{\mathrm{op}}\cdot\sigma_k(Y)^2. \tag{29}$$

Through a similar reasoning, we can handle the remaining term in (26). The important step is to make sure $\sigma_k(Y)$ appears quadratically:

$$\begin{aligned}\langle U_{T^\perp}, SU\rangle &= \left\langle \mathcal{A}^*\left(G^\dagger\mathcal{A}\left(xz^*Y^*\right)\right)Y, Sxz^*\right\rangle \\ &= \left\langle (\mathcal{A}^* \circ G^\dagger \circ \mathcal{A})(xz^*Y^*), Sxz^*Y^*\right\rangle \\ &\geq -\|\mathcal{A}^* \circ G^\dagger \circ \mathcal{A}\|_{\mathrm{op}}\|S\|_{\mathrm{op}}\|xz^*Y^*\|^2 \\ &\geq -K(1+KR)\|C\|_{\mathrm{op}}\cdot\sigma_k(Y)^2.\end{aligned} \tag{30}$$

Finally, combining (29) and (30) with (26) yields:

$$\begin{aligned}\langle x, Sx\rangle &\geq -\varepsilon_H - 2K(1+KR)\|C\|_{\mathrm{op}}\cdot\sigma_k(Y)^2 - K^2R(1+KR)\|C\|_{\mathrm{op}}\cdot\sigma_k(Y)^2 \\ &= -\varepsilon_H - K(2+KR)(1+KR)\|C\|_{\mathrm{op}}\cdot\sigma_k(Y)^2 \\ &\geq -\varepsilon_H - K(2+KR)^2\|C\|_{\mathrm{op}}\cdot\sigma_k(Y)^2 \\ &= -\varepsilon_H - \zeta\|C\|_{\mathrm{op}}\cdot\sigma_k(Y)^2,\end{aligned}$$

where ζ is as defined in the lemma statement. This holds for any unit vector x, hence the proof is complete.

E Proof of Theorem 3.1

We now build on Lemmas 3.1 and 3.2 to prove Theorem 3.1. The first part of the argument is fully deterministic: it relates the minimal eigenvalue of S to the optimality gap of the optimization problem.

Let Y be an $(\varepsilon_g, \varepsilon_H)$-SOSP of problem (P) with perturbed cost matrix $\tilde{C} = C + W$. By Lemma 3.2 applied to the perturbed problem,

$$\lambda_{\min}(S) \geq -\varepsilon_H - \zeta\|\tilde{C}\|_{\mathrm{op}}\sigma_k(Y)^2, \tag{31}$$

where ζ is as defined in that lemma, and S is as defined in (7) with cost matrix $\tilde{C}$ instead of C:

$$S(Y) = \tilde{C} - \mathcal{A}^*(\mu(Y)), \text{ and}$$
$$\mu(Y) = G^\dagger \mathcal{A}(\tilde{C}YY^*).$$

Using the definition of $\mathcal{C}$, for all $X^{'} \in \mathcal{C}$ feasible for the problem (SDP),

$$\lambda_{\min}(S) \cdot \mathrm{Tr}(X') \leq \langle S(Y), X' \rangle = \langle \tilde{C}, X' \rangle - \langle \mathcal{A}^*(\mu(Y)), X' \rangle = \langle \tilde{C}, X' \rangle - \langle \mu(Y), b \rangle .$$

In particular

$$\langle \mu(Y), b \rangle = \langle \mu(Y), \mathcal{A}(YY^*) \rangle = \langle \tilde{C} - S(Y), YY^* \rangle = g(Y) - \langle S(Y)Y, Y \rangle .$$

Combining those equations, using $\mathrm{grad}\, g(Y) = 2S(Y)Y$ and taking $X' = X^*$, we find

$$\begin{aligned} 0 \leq g(Y) - f^\star &\leq -\lambda_{\min}(S) \cdot \mathrm{Tr}(X^*) + \frac{1}{2}\langle \mathrm{grad}\, g(Y), Y \rangle \\ &\leq -\lambda_{\min}(S) \cdot \mathrm{Tr}(X^*) + \frac{\varepsilon_g}{2}\|Y\|. \end{aligned}$$

Since $\mathcal{C}$ is compact, we use the definition of R in (10) to get that $\mathrm{Tr}(X^*) \leq R$ and $\|Y\| \leq \sqrt{R}$:

$$\begin{aligned} 0 \leq g(Y) - f^\star &\leq -\lambda_{\min}(S) \cdot R + \frac{\varepsilon_g}{2}\sqrt{R} \\ &\leq \left(\varepsilon_H + \zeta \|\tilde{C}\|_{\mathrm{op}} \sigma_k(Y)^2\right) R + \frac{\varepsilon_g}{2}\sqrt{R}, \end{aligned} \tag{32}$$

where we used (31) in the last step.

We can now turn to the probabilistic part of the proof. Using Lemma 3.1, we have with probability at least $1 - \delta - e^{-\frac{n}{2}}$ that

$$\|W\|_{\mathrm{op}} \leq 3\sigma_W \sqrt{n}, \text{ and}$$
$$\sigma_k(Y) \leq \frac{\varepsilon_g}{\sigma_W} \frac{\sqrt{c_0 n}}{k},$$

and, by assumption,

$$k \geq 3\left[\log(n) + \sqrt{\log(1/\delta)} + \sqrt{m \cdot \log\left(1 + \frac{6\kappa\sqrt{c_0 n}}{\sigma_W}\right)}\right] \geq 3\sqrt{m \cdot \log\left(1 + \frac{6\kappa\sqrt{c_0 n}}{\sigma_W}\right)}.$$

In that event, combining, it follows that

$$\|\tilde{C}\|_{\mathrm{op}} \leq \|C\|_{\mathrm{op}} + 3\sigma_W\sqrt{n}, \text{ and}$$
$$\sigma_k(Y)^2 \leq \varepsilon_g^2 \frac{c_0 n}{9m\sigma_W^2 \log\left(1 + \frac{6\kappa\sqrt{c_0 n}}{\sigma_W}\right)}.$$

Combining with the deterministic result (32), we find that the optimality gap is bounded as

$$0 \leq g(Y) - f^\star \leq \left(\varepsilon_H + \varepsilon_g^2 \eta\right) R + \frac{\varepsilon_g}{2}\sqrt{R},$$

where η is as defined in (15). This concludes the proof.

F Proof of Corollary 3.1

Consider the two following functions:

$$f(C) = \min_{X \in \mathcal{C}} \langle C, X \rangle , \qquad h(C) = \max_{X \in \mathcal{C}} \langle C, X \rangle .$$

By assumption on X,

$$\langle C, X \rangle - \langle -W, X \rangle = \langle (C + W), X \rangle \leq f(C + W) + \varepsilon_f.$$

We can rearrange and get:

$$\langle C, X\rangle \leq f(C+W) + \varepsilon_f + \langle -W, X\rangle \leq f(C+W) + \varepsilon_f + h(-W).$$

Moreover,

$$f(C+W) = \min_{X\in\mathcal{C}} \left(\langle C, X\rangle + \langle W, X\rangle\right) \leq f(C) + h(W).$$

Overall, we get a bound on the optimality gap, using that $f(C) = f^\star$:

$$\langle C, X\rangle - f^\star \leq \varepsilon_f + h(W) + h(-W).$$

To conclude, observe that

$$h(W) = \max_{X\in\mathcal{C}} \langle W, X\rangle \leq \|W\|_{\mathrm{op}} \max_{X\in\mathcal{C}} \mathrm{Tr}(X) \leq \|W\|_{\mathrm{op}} R,$$

where we used that $\mathrm{Tr}(X) \leq R$ for all $X \in \mathcal{C}$. The same bound applies to $h(-W)$.