Rank optimality for the Burer-Monteiro factorization

Irène Waldspurger, Alden Waters

Rank optimality for the Burer-Monteiro factorization

Irène Waldspurger[*] Alden Waters[†]

December 7, 2018

Abstract

When solving large scale semidefinite programs that admit a low-rank solution, a very efficient heuristic is the Burer-Monteiro factorization: Instead of optimizing over the full matrix, one optimizes over its low-rank factors. This strongly reduces the number of variables to optimize, but destroys the convexity of the problem, thus possibly introducing spurious second-order critical points which can prevent local optimization algorithms from finding the solution. Boumal, Voroninski, and Bandeira [2018] have recently shown that, when the size of the factors is of the order of the square root of the number of linear constraints, this does not happen: For almost any cost matrix, second-order critical points are global solutions. In this article, we show that this result is essentially tight: For smaller values of the size, second-order critical points are not generically optimal, even when considering only semidefinite programs with a rank 1 solution.

1 Introduction

We consider a semidefinite program which is an optimization problem of the following form:

$$\begin{aligned} &\text{minimize } \operatorname{Trace}(CX) \\ &\text{such that } \mathcal{A}(X) = b, \\ &X \succeq 0, \end{aligned} \tag{SDP}$$

where the variable X and the fixed matrix C are symmetric, of size $n \times n$, and $\mathcal{A}$ is a linear operator capturing m equality constraints.

Various iterative algorithms have been developed to solve such a problem at a given precision level ϵ, but, in full generality, they tend to be computationally demanding. For instance, each iteration may cost $O((m+n)mn^2)$ arithmetic operations with an interior-point solver, and $O((m+n)n^2)$ with first-order techniques applied to a smoothed version of the problem.

Improvements are possible if $\mathcal{A}$ has some structure that can be exploited, but they often do not suffice to make large-scale semidefinite programs easy to handle. Another fundamental property of semidefinite programs is usually more useful for designing faster algorithms: These programs tend to have a low-rank minimizer (in many applications, there is a minimizer with rank $O(1)$, and, in any case, there is always one with rank $\sim \sqrt{2m}$ [Pataki, 1998, Thm 2.1]). As low-rank matrices can be stored and manipulated in a much more efficient way than full-rank ones, this allows to propose less computationally demanding algorithms.

Frank-Wolfe methods, in particular, can take advantage of this [Jaggi, 2013; Laue, 2012; Yurtsever, Udell, Tropp, and Cevher, 2017]. In this work, we are interested in another approach, the Burer-Monteiro factorization [Burer and Monteiro, 2005]. The principle behind the factorization is that a semidefinite matrix with rank $p \ll n$ can be factorized as

$$X = UU^T,$$

for some $U \in \mathbb{R}^{n\times p}$. If we assume that a low-rank solution X_{opt} to our problem exists, and if $p \geq \operatorname{rank}(X_{opt})$, Problem (SDP) is then equivalent to

$$\begin{aligned} &\text{minimize } \operatorname{Trace}(CUU^T) \\ &\text{such that } \mathcal{A}(UU^T) = b, \\ &U \in \mathbb{R}^{n\times p}. \end{aligned} \tag{Factorized SDP}$$

[*]waldspurger@ceremade.dauphine.fr, CNRS, Université Paris Dauphine and Inria Mokaplan, France
[†]a.m.s.waters@rug.nl, Bernoulli Institute, Rijksuniversiteit Groningen, Groningen, Netherlands

In this factorized problem, the unknown U has np coordinates, much less than the n^2 coordinates of X. Consequently, we can consider running on Problem (Factorized SDP) local optimization algorithms that would be too slow on Problem (SDP). The downside is that, since the factorized problem is not convex, these algorithms are not guaranteed to find a global minimizer. At best, they provably converge to a second-order critical point.

However, this absence of theoretical guarantees does not prevent them from working very well in many applications, even for very small values of p. In [Burer and Monteiro, 2003] and [Journée, Bach, Absil, and Sepulchre, 2010, Section 5], for example, local optimization algorithms are numerically observed to globally solve various instances of Problem (Factorized SDP), at least when $p \sim \sqrt{2m}$, and often for smaller values. Similar results on instances coming from orthogonal synchronization problems can be found in [Boumal, 2015, Section 5] and [Rosen, Carlone, Bandeira, and Leonard, 2016, Section 5]: Here, X_{opt} has rank 3, and is recovered using a Burer-Monteiro factorization with $p = 4$ or $p = 5$. The same behavior occurs for other matricial problems, that do not have the form (SDP), but also admit a low-rank solution [Mishra, Meyer, Bonnabel, and Sepulchre, 2014].

[Bandeira, Boumal, and Voroninski, 2016a] theoretically explains this behavior for instances of (SDP) and (Factorized SDP) coming from $\mathbb{Z}_2$-synchronization and community detection. This article notably establishes, in particular statistical regimes, where Problem (SDP) has a rank-1 solution, that all second-order critical points of Problem (Factorized SDP) with $p = 2$ are global minimizers. Hence, suitable local optimization algorithms globally solve Problem (Factorized SDP). For problems different from, but related to (SDP), similar results can be found for instance in [Ge, Lee, and Ma, 2016; Sun, Qu, and Wright, 2017; Li, Zhu, and Tang, 2018]; these articles show, in specific settings, that all second-order critical points of the Burer-Monteiro factorization are the optimal solution as soon as $p \geq \text{rank}(X_{opt})$.

While these works shed light on several important practical situations, they do not provide a general theory on when local optimization algorithms can solve Problem (Factorized SDP). With no restrictive assumptions, essentially the only result is due to Boumal, Voroninski, and Bandeira [2018][1]: Building on [Burer and Monteiro, 2005] and [Boumal, 2015], these authors show that, under reasonable geometrical hypotheses, all second-order critical points of Problem (Factorized SDP) are global minimizers, for almost any cost matrix C, as soon as

$$\frac{p(p+1)}{2} > m, \tag{1.1}$$

that is $p > \lfloor \sqrt{2m + 1/4} - 1/2 \rfloor$. Extensions of this can be found in [Pumir, Jelassi, and Boumal, 2018] and [Bhojanapalli, Boumal, Jain, and Netrapalli, 2018].

As a result there is a gap in the literature: In all the concrete settings that could be studied, all second-order critical points of Problem (Factorized SDP) are global minimizers as soon as $p \gtrsim \text{rank}(X_{opt})$, but in the general case, the only guarantees at our disposal state that we need p to be at least as large as $\sim \sqrt{2m}$. In many applications, $\text{rank}(X_{opt}) = O(1)$ while m is of the order of n, hence these two estimates are far apart, which makes a huge difference on the computational cost of algorithms that can be certified to be correct.

This naturally raises the question of determining whether the gap can be reduced. The goal of this article is to negatively answer this question, by showing that Inequality (1.1) is essentially optimal. More precisely,

- We show in Theorem 1 that a minor improvement over Inequality 1.1 is possible: Under a stronger geometrical assumption than in [Boumal, Voroninski, and Bandeira, 2018] (but still reasonable), all second-order critical points of Problem (Factorized SDP) are global minimizers, for almost any C, as soon as

$$\frac{p(p+1)}{2} + p > m. \tag{1.2}$$

This improves over Inequality (1.1), but only in the low order terms.

- We show in Theorem 2 (our main result) that, for any $\mathcal{A}, b$, if p is such that

$$\frac{p(p+1)}{2} + pr_* \leq m,$$

[1] [Mei, Misiakiewicz, Montanari, and Oliveira, 2017] also considers a rather general setting, but with a different focus: The authors prove that second-order critical points of Problem (Factorized SDP) provide an approximation of the optimal cost $\langle C, X_{opt} \rangle$ within an error of order $O(1/p)$, for any p; they do not address the question of when these critical points exactly minimize the cost.

where $r_* = \min\{\text{rank}(X), X \succeq 0, \mathcal{A}(X) = b\}$, then there exists a set of cost matrices C with non-zero Lebesgue measure on which Problem (SDP) admits a global minimizer with rank r_*, but Problem (Factorized SDP) has second-order critical points which are not global minimizers.

In particular, for $\mathcal{A}, b$ such that $r_* = 1$ (as is the case in *MaxCut* relaxations, for instance), Inequality (1.2) is exactly optimal.

This means that, without specific assumptions on C, when running a local optimization algorithm on Problem (Factorized SDP) with p smaller than $\sim \sqrt{2m}$, we cannot be sure not to run into a spurious second-order critical point, even if there exists a global minimizer with rank $O(1)$.

Regarding the organization of this article, Section 2 presents the main results: After some definitions in Subsection 2.1, Theorems 1 and 2 are respectively stated in Subsections 2.2 and 2.3. Subsection 2.4 presents applications of these theorems to important families of problems of the form (SDP). The other sections contain the proofs: Theorem 1 is proved in Section 3, and Theorem 2 in Section 4; Section 5 explains how to verify the hypotheses of the theorems for the examples of Subsection 2.4.

1.1 Notations

1.1.1 Notations related to matrices

For any $p, q \in \mathbb{N}^*$, we denote by I_p the $p \times p$ identity matrix, and by $0_{p\times q}$ the zero $p \times q$ matrix.

For any $p \in \mathbb{N}^*$, we denote by $\mathbb{S}^{p\times p}$ the set of real symmetric $p \times p$ matrices, by $\text{Anti}(p)$ the set of antisymmetric $p \times p$ matrices, and by $O(p)$ the set of orthogonal $p \times p$ matrices.

For any n_1, n_2, we equip $\mathbb{R}^{n_1\times n_2}$, the set of $n_1 \times n_2$ matrices, with the usual scalar product:

$$\forall M_1, M_2 \in \mathbb{R}^{n_1\times n_2}, \quad \langle M_1, M_2\rangle \overset{def}{=} \text{Tr}(M_1^T M_2).$$

The same formula also defines a scalar product on $\mathbb{S}^{p\times p}$, for any $p \in \mathbb{N}^*$. In both cases, the associated norm is the Frobenius norm, which we denote by $||.||_F$.

We denote by $|||.|||$ the operator norm on $\mathbb{S}^{p\times p}$:

$$\forall M \in \mathbb{R}^{p\times p}, \quad |||M||| = \sup_{x\in\mathbb{R}^p\setminus\{0\}} \frac{||Mx||_{\ell^2}}{||x||_{\ell^2}}.$$

For any $p \in \mathbb{N}^*$, we define $\text{diag} : \mathbb{R}^{p\times p} \to \mathbb{R}^p$ as the operator which associates to a matrix the vector of its diagonal elements. The dual operator $\text{Diag} : \mathbb{R}^p \to \mathbb{R}^{p\times p}$ associates a vector to a diagonal matrix with diagonal entries that are the coefficients of the vector.

1.1.2 Topological and geometrical notations

For any element x of a metric space, and any positive real number ϵ, we denote $B(x, \epsilon)$ the open ball with radius ϵ, and $\overline{B}(x, \epsilon)$ the closed ball. The unit sphere in $\mathbb{R}^p$, for any $p \in \mathbb{N}^*$, is denoted by S^{p-1}.

When $\mathcal{M}$ is a manifold, and x an element of $\mathcal{M}$, we denote by $T_x\mathcal{M}$ the tangent space of $\mathcal{M}$ at x. Furthermore, when $\mathcal{M}$ is an embedded submanifold of $\mathbb{R}^p$, it inherits the Riemannian structure of $\mathbb{R}^p$. The volume form associated with this structure defines a measure on $\mathcal{M}$, which we let be $\lambda_{\mathcal{M}}$. When $\mathcal{M} = \mathbb{R}^p$, $\lambda_{\mathcal{M}}$ coincides with the Lebesgue measure.

2 Main results

2.1 Definitions

We consider a problem of the following form:

$$\begin{aligned} \text{minimize } & \langle C, X\rangle \\ \text{such that } & X \in \mathbb{S}^{n\times n}, \\ & \mathcal{A}(X) = b, \\ & X \succeq 0. \end{aligned} \tag{SDP}$$

Here, $\mathcal{A} : \mathbb{S}^{n\times n} \to \mathbb{R}^m$ is a fixed linear map, b a fixed element of $\mathbb{R}^m$, and C an element of $\mathbb{S}^{n\times n}$, which is called the *cost matrix*.

We will always denote by $\mathcal{C}$ be the set of feasible points for this problem:

$$\mathcal{C} = \{X \in \mathbb{S}^{n \times n}, \mathcal{A}(X) = b, X \succeq 0\}.$$

As explained in the introduction, if we assume that Problem (SDP) has an optimal solution X_{opt} with rank r, then this solution can be factorized as

$$X_{opt} = V_{opt} V_{opt}^T,$$

where V_{opt} is a $n \times r$ matrix. A reasonable heuristic to solve Problem (SDP) at a (relatively) cheap computational cost is thus to choose some $p \geq r$, and replace Problem (SDP) with its *rank p Burer-Monteiro factorization*:

$$\begin{aligned} \text{minimize } & \langle C, VV^T \rangle \\ \text{such that } & V \in \mathbb{R}^{n \times p}, \\ & \mathcal{A}(VV^T) = b. \end{aligned} \tag{Factorized SDP}$$

We denote by $\mathcal{M}_p$ the set of feasible points for this problem:

$$\mathcal{M}_p = \{V \in \mathbb{R}^{n \times p}, \mathcal{A}(VV^T) = b\}.$$

It is invariant under multiplication by elements of $O(p)$. We assume that it is sufficiently regular so that we can apply smooth optimization algorithms to Problem (Factorized SDP). More precisely, all our results require that $(\mathcal{A}, b)$ is *p-regular*:

Definition 1. *For some $p \in \mathbb{N}^*$, $(\mathcal{A}, b)$ is said to be* p-regular *if, for all $V \in \mathcal{M}_p$, the linear map*

$$\dot{V} \in \mathbb{R}^{n \times p} \to \mathcal{A}(V\dot{V}^T + \dot{V}V^T) \in \mathbb{R}^m$$

is surjective.

This assumption is of the same style as [Boumal, Voroninski, and Bandeira, 2018, Assumption 1.1]. It notably guarantees that $\mathcal{M}_p$ is a submanifold of $\mathbb{R}^{n \times p}$, with dimension

$$\dim(\mathcal{M}_p) = np - m,$$

and whose tangent space at any point V is

$$T_V \mathcal{M}_p = \{\dot{V} \in \mathbb{R}^{n \times p}, \mathcal{A}(V\dot{V}^T + \dot{V}V^T) = 0\}.$$

The scalar product of $\mathbb{R}^{n \times p}$ defines a metric on the manifold $\mathcal{M}_p$, which we then view as a Riemannian manifold. Many algorithms exist for attempting to minimize a smooth function on a Riemannian manifold; a classical reference on this topic is [Absil, Mahony, and Sepulchre, 2009].

However, as said in the introduction, these algorithms are a priori not guaranteed to find a global minimizer of Problem (Factorized SDP), but only (an approximation of) a first-order or second-order critical point of the cost function $V \in \mathcal{M}_p \to \langle C, VV^T \rangle$ [Boumal, Absil, and Cartis, 2016]. These points are defined as follows:

Definition 2. *Let $\mathcal{N}$ be a Riemannian manifold, and $f : \mathcal{N} \to \mathbb{R}$ a smooth function. For any $x_0 \in \mathcal{N}$, we say that x_0 is*

- *a* first-order critical point *of f if*
$$\nabla f(x_0) = 0;$$
- *a* second-order critical point *of f if*
$$\nabla f(x_0) = 0 \quad \textit{and} \quad \mathrm{Hess} f(x_0) \succeq 0.$$

The goal of this article is to study for which values of p the set of second-order critical points actually coincides with the set of global minimizers of Problem (Factorized SDP). Before turning to this objective, we still need to define two technical properties that will appear in the assumptions of our main theorems.

Definition 3. *Let $p \in \mathbb{N}^*$ be such that $(\mathcal{A}, b)$ is p-regular. We say that $\mathcal{M}_p$ is* face regular *if, for almost any $V \in \mathcal{M}_p$, the map*

$$\phi_V : T \in \mathbb{S}^{p\times p} \to \mathcal{A}(VTV^T) \in \mathbb{R}^m \tag{2.1}$$

is injective.

We chose the name "face regular" because when V has rank p, the injectivity of ϕ_V is equivalent to the fact that the face of $\mathcal{C}$ containing V is a singleton.

Definition 4. *Let $p \in \mathbb{N}^*$ be such that $(\mathcal{A}, b)$ is p-regular.*

Let X_0 belong to $\mathcal{C}$, and let r be its rank. Let $U_0 \in \mathbb{R}^{n\times r}$ be[2] *such that*

$$X_0 = U_0 U_0^T.$$

For any $V \in \mathcal{M}_p$, we say that $\mathcal{M}_p$ is X_0-minimally secant *at V if the map*

$$\begin{aligned}\psi_V : (T, R) &\in \mathbb{S}^{p\times p} \times \mathbb{R}^{r\times p} \\ &\to \mathcal{A}\left(\begin{pmatrix} V & U_0 \end{pmatrix} \begin{pmatrix} T \\ R \end{pmatrix} V^T + V \begin{pmatrix} T \\ R \end{pmatrix}^T \begin{pmatrix} V & U_0 \end{pmatrix}^T \right) \in \mathbb{R}^m \end{aligned} \tag{2.2}$$

is injective.

We chose the name "minimally secant" because the injectivity of ψ_V is a sufficient condition for the intersection

$$T_V \mathcal{M}_p \cap \{W \in \mathbb{R}^{n\times p}, \mathrm{Range}(W) \subset \mathrm{Range}(V) + \mathrm{Range}(X_0)\}$$

to exactly equal the set

$$\{VA, A \in \mathrm{Anti}(p)\}$$

(which it always contains, because of the invariance of $\mathcal{M}_p$ under multiplication by elements of $O(p)$).

Remark 1. *By dimensionality arguments, $\mathcal{M}_p$ can only be face regular if*

$$\dim(\mathbb{S}^{p\times p}) = \frac{p(p+1)}{2} \leq m,$$

and X_0-minimally secant at a point V if

$$\dim(\mathbb{S}^{p\times p} \times \mathbb{R}^{r\times p}) = \frac{p(p+1)}{2} + pr \leq m.$$

2.2 Regime where critical points are global minimizers

As previously stated, most smooth optimization algorithms, when applied to Problem (Factorized SDP), are only guaranteed to find a critical point of this problem, and not a global minimizer. Fortunately, Boumal, Voroninski, and Bandeira [2018] have shown that, when p is large enough, second-order critical points are always global minimizers, for almost all cost matrices C. Therefore, algorithms able to find second-order critical points (like, for instance, the trust-region method), actually solve Problem (Factorized SDP) (and hence also Problem (SDP)) to optimality, provided that C is "generic".

A restated version of the theorem by Boumal, Voroninski, and Bandeira [2018], under minor modifications, is the following:

Theorem. *[Boumal, Voroninski, and Bandeira [2018, Theorem 1.4]] Let $p \in \mathbb{N}^*$ be fixed. We assume that*

1. *The set $\mathcal{C}$ of feasible points for Problem (SDP) is compact;*
2. *$(\mathcal{A}, b)$ is p-regular.*

If

$$\frac{p(p+1)}{2} > m, \tag{2.3}$$

then, for almost all cost matrices $C \in \mathbb{S}^{n\times n}$, if $V \in \mathcal{M}_p$ is a second-order critical point of Problem (Factorized SDP), then

[2] The matrix U_0 is in general not unique, but we can fix it without loss of generality: ψ_V is injective for all possible U_0 or for none of them.

- *V is a global minimizer of Problem (Factorized SDP);*
- *$X = VV^T$ is a global minimizer of Problem (SDP).*

In order for this theorem to be applicable, the rank p of the Burer-Monteiro factorization must satisfy Condition (2.3), hence be (strictly) larger than

$$\left\lfloor \sqrt{2m + \frac{1}{4}} - \frac{1}{2} \right\rfloor.$$

It is natural to ask whether this is optimal, or whether the same guarantees also hold for smaller ranks p, allowing further reductions in the computational complexity of solving Problem (Factorized SDP).

Our first result is that Condition (2.3) can be slightly relaxed, at the price of an additional assumption on $\mathcal{M}_p$. This yields the following theorem:

Theorem 1. *Let $p \in \mathbb{N}^*$ be fixed. We assume that*

1. *The set $\mathcal{C}$ of feasible points for Problem (SDP) is compact.*
2. *$(\mathcal{A}, b)$ is p-regular;*
3. *If $\frac{p(p+1)}{2} \leq m$, then $\mathcal{M}_p$ is face regular.*

If

$$\frac{p(p+1)}{2} + p > m,$$

then the same conclusion holds as in the previous theorem.

Compared to the theorem in [Boumal, Voroninski, and Bandeira, 2018], the additional assumption on $\mathcal{M}_p$ is the third one. Let us remark that we expect it to be satisfied in almost all applications. Indeed, a linear map between two vector spaces is generically injective when the dimension of the second space is at least as large as the dimension of the first one. Consequently, when $p(p+1)/2 \leq m$, the map ϕ_V in Definition 3 should be a priori injective for "generic" matrices V.

We note here that the assumption is satisfied in each of the three cases studied in Subsection 2.4, except for a few very particular values of m and p in the case of *Orthogonal-Cut* (Paragraph 2.4.2).

The proof of Theorem 1 is the subject of Section 3.

2.3 Regime where there may be bad critical points

In the previous subsection, we have seen that, when $\frac{p(p+1)}{2} + p > m$, under some hypotheses, solving Problem (Factorized SDP) to second-order criticality is equivalent to solving Problem (SDP) itself. We can now address our main question: How optimal is this result?

Of course, when Problem (SDP) has a unique global minimizer, whose rank r is of the order of $\sqrt{2m}$, the result cannot be significantly improved: $p \geq r$ is a necessary condition for Problems (SDP) and (Factorized SDP) to have the same minimum. However, as discussed in the introduction, in many applications, Problem (SDP) admits a solution with rank $r \ll \sqrt{2m}$, and the Burer-Monteiro factorization is numerically observed to work when $p = O(r)$.

Our main theorem however states that, even if we assume $r \ll \sqrt{2m}$, our previous result is essentially not improvable without additional assumptions on C (under reasonable assumptions on $(\mathcal{A}, b)$): When $\frac{p(p+1)}{2} + pr \leq m$, there exists a set of cost matrices with non-zero Lebesgue measure for which Problem (SDP) has a rank r optimal solution, and Problem (Factorized SDP) nevertheless has non-optimal second-order critical points.

In particular, the inequality $\frac{p(p+1)}{2} + p > m$ in Theorem 1 is exactly optimal when $(\mathcal{A}, b)$ is such that it is possible to choose $r = 1$ in the following theorem.

Theorem 2. *Let $r \in \mathbb{N}^*$ be fixed. Let $p \geq r$ be such that*

$$\frac{p(p+1)}{2} + pr \leq m.$$

We make the following hypotheses:

1. *$\mathcal{C}$ has at least one extreme point with rank r, denoted by X_0;*

2. *$(\mathcal{A}, b)$ is p-regular;*

3. *There exists $V \in \mathcal{M}_p$ such that $\mathcal{M}_p$ is X_0-minimally secant at V.*

Then there exists a subset $\mathcal{E}_{bad}$ of $\mathbb{S}^{n\times n}$ with non-zero Lebesgue measure such that, for any cost matrix $C \in \mathcal{E}_{bad}$,

- *Problem (SDP) has a unique global minimizer, which has rank r.*
- *Problem (Factorized SDP) has at least one second-order critical point Y that is not a global minimizer.*

The proof of this result is in Section 4.

Remark 2. *A slight adaptation of the proof of Theorem 2 would actually allow to draw a slightly stronger condition than what is stated above: "one second-order critical point Y that is not a global minimizer" can be replaced with "one local minimizer Y that is not a global minimizer".*

Remark 3. *The inequalities $\frac{p(p+1)}{2} + p > m$ and $\frac{p(p+1)}{2} + pr \leq m$ in Theorems 1 and 2 are exactly complementary when $r = 1$. But when $r \geq 2$, our results do not provide information on the regime where*

$$\frac{p(p+1)}{2} + p \leq m < \frac{p(p+1)}{2} + pr.$$

As shown through an example in Appendix B, it is possible that the conclusions of Theorem 2 are still valid, but we do not know whether this is always the case, or whether it depends on $p, r, n, m, \mathcal{A}, b$.

2.4 Examples

In this subsection, we apply Theorems 1 and 2 to three important examples of semidefinite programs: *MaxCut* (Paragraph 2.4.1), *Orthogonal-Cut* (Paragraph 2.4.2) and optimization over a product of spheres (Paragraph 2.4.3).

The only difficulty when applying these theorems is to check that their hypotheses are satisfied. A general discussion about this can be found in Section 5, and detailed proofs are available in the appendix.

2.4.1 MaxCut

The most famous instance of a problem with the form (SDP) is probably the *MaxCut* relaxation:

$$\begin{aligned} \text{minimize } & \langle C, X\rangle \\ \text{such that } & X \in \mathbb{S}^{n\times n}, \\ & \operatorname{diag}(X) = 1, \\ & X \succeq 0. \end{aligned} \tag{SDP-Maxcut}$$

This problem was introduced as a relaxation of the "maximum cut" problem, from graph theory [Delorme and Poljak, 1993; Poljak and Rendl, 1995]. It drew considerable attention when Goemans and Williamson [1995] proved that, combined with a suitable rounding procedure, it yields an approximate algorithm for solving this maximum cut problem, with a ratio much closer to 1 than previous methods. It also appears in phase retrieval [Waldspurger, d'Aspremont, and Mallat, 2015] and $\mathbb{Z}/2\mathbb{Z}$ synchronization [Abbe, Bandeira, and Hall, 2016; Bandeira, Boumal, and Voroninski, 2016a] (in which cases the global optimizer of Problem (SDP-Maxcut) is known, both theoretically and numerically, to often have very low rank, typically 1).

When applied to Problem (SDP-Maxcut), Theorems 1 and 2 show that, for almost any cost matrix, the Burer-Monteiro factorization has no non-optimal second-order critical point as soon as

$$\frac{p(p+1)}{2} + p > n \qquad \iff \qquad p > \left\lfloor \sqrt{2n + \frac{9}{4}} - \frac{3}{2} \right\rfloor,$$

and moreover that this result is optimal, even if we restrict ourselves to instances that admit global minimizers with rank 1.

Corollary 1. *If $p \in \mathbb{N}$ is such that*

$$\frac{p(p+1)}{2} + p > n,$$

then, for almost any cost matrix C, all second-order critical points of the Burer-Monteiro factorization of Problem (SDP-Maxcut) are globally optimal.

On the other hand, for any p such that

$$\frac{p(p+1)}{2} + p \leq n,$$

the set of cost matrices admits a subset with non-zero Lebesgue measure on which

- *Problem (SDP-Maxcut) has a unique global minimizer, which has rank 1;*
- *Its Burer-Monteiro factorization with rank p has at least one non-optimal second-order critical point.*

This corollary is a particular case of Corollary 2, in the next paragraph, whose proof is in Appendix D.3.

2.4.2 Orthogonal-Cut

We now consider a generalization of *MaxCut*, coined *Orthogonal-Cut* in [Bandeira, Kennedy, and Singer, 2016b]:

$$\begin{aligned} &\text{minimize } \langle C, X \rangle \\ &\text{such that } X \in \mathbb{S}^{Sd \times Sd}, \\ &\qquad \text{Block}_s(X) = I_d, \forall s = 1, \dots, S, \\ &\qquad X \succeq 0, \end{aligned} \tag{SDP-Orthogonal-Cut}$$

where d, S belong to $\mathbb{N}^*$ (with, typically, $d = 1, 2$ or 3) and, for any $M \in \mathbb{R}^{Sd \times Sd}$, $s \leq S$, we denote by $\text{Block}_s(M)$ the s-th diagonal $d \times d$ block of M. Observe that this is exactly Problem (SDP-Maxcut) when $d = 1$.

Problem (SDP-Orthogonal-Cut) is a natural relaxation of non-convex problems where one wants to find orthogonal matrices $O_1, \dots, O_S \in O(d)$ which optimize a given criterion. It notably has applications in molecular imaging [Wang, Singer, and Wen, 2013], sensor network localization [Cucuringu, Lipman, and Singer, 2012] and ranking [Cucuringu, 2016]. For some theoretical analysis of this semidefinite problem, including conditions under which it admits a low-rank global minimizer, the reader can refer, not only to [Bandeira, Kennedy, and Singer, 2016b], but to [Chaudhury, Khoo, and Singer, 2015], [Rosen, Carlone, Bandeira, and Leonard, 2016] or [Eriksson, Olsson, Kahl, and Chin, 2018], as well.

Problem (SDP-Orthogonal-Cut) is exactly equivalent to

$$\begin{aligned} &\text{minimize } \langle C, X \rangle \\ &\text{such that } X \in \mathbb{S}^{Sd \times Sd}, \\ &\qquad \mathcal{A}(X) = b, \\ &\qquad X \succeq 0, \end{aligned}$$

with

$$\mathcal{A} : X \in \mathbb{S}^{Sd \times Sd} \to (T_{sup}(\text{Block}_1(X)), \dots, T_{sup}(\text{Block}_S(X))) \in \mathbb{R}^{Sd(d+1)/2},$$

and

$$b = (T_{sup}(I_d), \dots, T_{sup}(I_d)) \in \mathbb{R}^{Sd(d+1)/2},$$

where $T_{sup} : \mathbb{R}^{d \times d} \to \mathbb{R}^{d(d+1)/2}$ is the operator that extracts the $\frac{d(d+1)}{2}$ coefficients of the upper triangular part of a matrix.

With these definitions, $(\mathcal{A}, b)$ is p-regular for any $p \in \mathbb{N}$. In particular, $\mathcal{M}_p$ is a manifold (empty when $p < d$, non-empty when $p \geq d$). We apply Theorems 1 and 2 to this problem (restricting ourselves to the values $d = 1, 2, 3$ because these are the ones most often considered in applications). They show, at least[3]

[3] When $p < 2d$, the face regularity assumption does not always hold.

when $p \geq 2d$, that, for almost any cost matrix, the Burer-Monteiro factorization has no non-optimal second-order critical point when

$$\frac{p(p+1)}{2} + p > \frac{Sd(d+1)}{2} \qquad \Longleftrightarrow \qquad p > \left\lfloor \sqrt{Sd(d+1) + \frac{9}{4}} - \frac{3}{2} \right\rfloor, \tag{2.4}$$

but that this is not true, even in the presence of a rank d minimizer, when

$$\begin{aligned} &\frac{p(p+1)}{2} + pd \leq \frac{Sd(d+1)}{2} \\ &\qquad \Longleftrightarrow \qquad p \leq \left\lfloor \sqrt{Sd(d+1) + \left(d + \frac{1}{2}\right)^2} - \left(d + \frac{1}{2}\right) \right\rfloor. \end{aligned} \tag{2.5}$$

Observe that the bounds on p in Equations (2.4) and (2.5) are very close (they differ by at most $d-1$, recalling that, most often, $d \leq 3$), hence this result is almost optimal.

Corollary 2. *Let us assume that $d = 1, 2$ or 3.*

If $p \geq 2d$ is such that

$$\frac{p(p+1)}{2} + p > \frac{Sd(d+1)}{2},$$

then, for almost any cost matrix C, all second-order critical points of the Burer-Monteiro factorization of Problem (SDP-Orthogonal-Cut) are globally optimal.

On the other hand, for any $p \geq d$ such that

$$\frac{p(p+1)}{2} + pd \leq \frac{Sd(d+1)}{2},$$

the set of cost matrices admits a subset with non-zero Lebesgue measure on which

- *Problem (SDP-Orthogonal-Cut) has a unique global optimum, which has rank d;*
- *Its Burer-Monteiro factorization with rank p has at least one non-optimal second-order critical point.*

The proof is in Appendix D.3.

2.4.3 Optimization over a product of spheres

As a final example, let us consider the problem

$$\begin{aligned} &\text{minimize } \langle C, X \rangle \\ &\text{such that } X \in \mathbb{S}^{D \times D}, \\ &\sum_{k=d_1+\dots+d_{s-1}+1}^{d_1+\dots+d_s} X_{k,k} = 1, \forall s = 1, \dots, S, \\ &X \succeq 0, \end{aligned} \tag{SDP-Product}$$

where $S, d_1, \dots, d_S$ belong to $\mathbb{N}^*$, and we set $D = d_1 + \dots + d_S$. This is the natural semidefinite relaxation of problems that consist in minimizing a degree 2 polynomial function on the product of spheres $S^{d_1} \times \dots \times S^{d_S}$.

Problem (SDP-Product) encompasses several important particular cases: when $d_1 = \dots = d_S = 1$, we recover Problem (SDP-Maxcut). When $d_1 = \dots = d_S = 2$, it is equivalent to a complex version of (SDP-Maxcut) (for matrices C of a particular form). When $S = 2$ and $d_2 = 1$, it is the relaxation of a standard trust-region subproblem [Boumal, Voroninski, and Bandeira, 2018, Subsection 5.2]. For general values of $d_1, \dots, d_S$, it is a simplification of the relaxation of optimization problems over an intersection of ellipsoids, which appear in trust-region algorithms for constrained problems [Celis, 1985].

For Problem (SDP-Product), Theorems 1 and 2 guarantee that the Burer-Monteiro factorization has no non-optimal second-order critical point for almost any cost matrix as soon as

$$\frac{p(p+1)}{2} + p > S \qquad \Longleftrightarrow \qquad p > \left\lfloor \sqrt{2S + \frac{9}{4}} - \frac{3}{2} \right\rfloor,$$

and that this is optimal, even when assuming the existence of a global minimizer of rank 1.

Corollary 3. *If $p \in \mathbb{N}$ is such that*

$$\frac{p(p+1)}{2} + p > S,$$

then, for almost any cost matrix C, all second-order critical points of the Burer-Monteiro factorization of Problem (SDP-Product) are globally optimal.

On the other hand, for any $p \in \mathbb{N}^*$ such that

$$\frac{p(p+1)}{2} + p \leq S,$$

the set of cost matrices admits a subset with non-zero Lebesgue measure on which

- *Problem (SDP-Product) has a unique global optimum, which has rank 1;*
- *Its Burer-Monteiro factorization with rank p has at least one non-optimal second-order critical point.*

The proof is in Appendix D.4.

3 Proof of Theorem 1

This section is devoted to the proof of Theorem 1.

Theorem (Theorem 1)**.** *Let $p \in \mathbb{N}^*$ be fixed. We assume that*

1. *The set $\mathcal{C}$ of feasible points for Problem (SDP) is compact.*
2. *$(\mathcal{A}, b)$ is p-regular;*
3. *If $\frac{p(p+1)}{2} \leq m$, then $\mathcal{M}_p$ is face regular.*

If

$$\frac{p(p+1)}{2} + p > m,$$

then, for almost all cost matrices $C \in \mathbb{S}^{n\times n}$, if $V \in \mathcal{M}_p$ is a second-order critical point of Problem (Factorized SDP), then

- *V is a global minimizer of Problem (Factorized SDP);*
- *$X = VV^T$ is a global minimizer of Problem (SDP).*

When $\frac{p(p+1)}{2} > m$, this result is a direct consequence of [Boumal, Voroninski, and Bandeira, 2018, Thm 1.4] so, from now on, we assume that

$$\frac{p(p+1)}{2} \leq m < \frac{p(p+1)}{2} + p. \tag{3.1}$$

Problem (SDP) admits at least one global minimizer with rank p_*, for some p_* such that $\frac{p_*(p_*+1)}{2} \leq m$ ([Pataki, 1998, Thm 2.1]). We have

$$\frac{(p+1)((p+1)+1)}{2} = \frac{p(p+1)}{2} + p + 1 > m,$$

so Problem (SDP) has a minimizer with rank at most p, that can hence be written as VV^T for $V \in \mathcal{M}_p$. Consequently, Problems (SDP) and (Factorized SDP) have the same minimum, and, for any minimizer V of Problem (Factorized SDP), $X = VV^T$ is a minimizer of Problem (SDP). To establish the theorem, it therefore suffices to show the first half of the statement: For almost all cost matrices, Problem (Factorized SDP) has no second-order critical points that are not global minimizers of Problem (Factorized SDP).

Before presenting the details, let us give an overview of the proof.

In [Boumal, Voroninski, and Bandeira, 2018], the proof consists in showing that any cost matrix C for which non-optimal second-order critical points exist can be written as

$$\mathcal{A}^*(\mu) + C_2, \tag{3.2}$$

for some $\mu \in \mathbb{R}^m$ and $C_2 \in \mathbb{S}^{n\times n}$ with rank at most $n-p$. One then studies the "dimension"[4] of the set of matrices with this form, and shows that it is at most

$$\frac{n(n+1)}{2} + \left(m - \frac{p(p+1)}{2}\right),$$

which is strictly smaller than $\dim(\mathbb{S}^{n\times n}) = \frac{n(n+1)}{2}$ when $m < \frac{p(p+1)}{2}$, implying that the set of problematic matrices has Lebesgue measure zero in $\mathbb{S}^{n\times n}$.

Our proof follows the same principle, but slightly refines it:

- First step (Subsection 3.1): We show that, under Condition (3.1), matrices C for which non-optimal second-order critical points exist can be written as in Equation (3.2), with an additional condition on C_2.

  This condition is that there exists a matrix $V \in \mathcal{M}_{p,\text{non inj}}$ with rank p such that $C_2V = 0$. Here, we denote by $\mathcal{M}_{p,\text{non inj}}$ the set of matrices V for which the map ϕ_V in Equation (2.1) is non-injective.

- Second step (Subsection 3.2): Denoting by $\mathcal{E}_{bad}$ the set of matrices with the form described in the first step, we show that it has measure zero.

  Informally, heuristically reasoning in terms of dimensions, $\mathcal{E}_{bad}$ has the following "dimension":

$$\underbrace{\text{“dim”}(\mathcal{M}_{p,\text{non inj}})}_{\substack{\text{Dimension of} \\ \text{possible } V\text{'s}}} - \underbrace{\frac{(p-1)p}{2}}_{\substack{\mathcal{M}_{p,\text{non inj}} \text{ can} \\ \text{be quotiented by } O(p)}} + \underbrace{\frac{(n-p)(n-p+1)}{2}}_{\substack{\text{For each } V,\text{ dimension} \\ \text{of } \{C_2, C_2V=0\}}} + \underbrace{m}_{\substack{\text{Dimension of} \\ \text{possible } \mu\text{'s}}}$$

$$= \frac{n(n+1)}{2} + \text{“dim”}(\mathcal{M}_{p,\text{non inj}}) - \dim(\mathcal{M}_p).$$

  Because of the face regularity assumption, "dim"$(\mathcal{M}_{p,\text{non inj}}) < \dim(\mathcal{M}_p)$, hence this value is strictly smaller than $\frac{n(n+1)}{2} = \dim(\mathbb{S}^{n\times n})$.

3.1 First step

We define

$$\mathcal{E}_{bad} = \{\mathcal{A}^*(\mu) + C_2, \mu \in \mathbb{R}^m, C_2 \in \mathbb{S}^{n\times n}, \exists V \in \mathcal{M}_{p,\text{non inj}} \\ \text{such that } \operatorname{rank}(V) = p \text{ and } C_2V = 0\},$$

and show that any matrix C for which Problem (Factorized SDP) has a non globally optimal second-order critical point belongs to $\mathcal{E}_{bad}$.

Let C be such a matrix. We denote by V a non-optimal second-order critical point. From [Boumal, Voroninski, and Bandeira, 2018, Thm 1.6], $\operatorname{rank}(V) = p$ and the dimension of the face of $\mathcal{C}$ containing VV^T is at least

$$\frac{p(p+1)}{2} - m + p \overset{\text{Eq. (3.1)}}{>} 0.$$

This face is a subset of the face of VV^T in $\{X \in \mathbb{S}^{n\times n}, X \succeq 0\}$, which is $\{VTV^T, T \in \mathbb{S}^{p\times p}, T \succeq 0\}$. As it is not the singleton $\{VV^T\}$, there must then exist $T \in \mathbb{S}^{p\times p}$ such that $VTV^T \neq VV^T$ and VTV^T is an element of $\mathcal{C}$, implying

$$\mathcal{A}(VTV^T) = b = \mathcal{A}(VV^T).$$

From this, we see that ϕ_V (defined in Equation (2.1)) is not injective, so V belongs to $\mathcal{M}_{p,\text{non inj}}$.

As V is a first-order critical point of Problem (Factorized SDP), there exists, from [Boumal, Voroninski, and Bandeira, 2018, Eqs (8) and (11)], $\mu \in \mathbb{R}^m$ such that

$$(C - \mathcal{A}^*(\mu))V = 0.$$

Setting $C_2 = C - \mathcal{A}^*(\mu)$, we see that C belongs to $\mathcal{E}_{bad}$.

[4] We use the word "dimension" to give an intuition of the proof, but it is improper, since the considered set is not a manifold.

3.2 Second step

To establish the theorem, we must now show that $\mathcal{E}_{bad}$ has measure zero in $\mathbb{S}^{n\times n}$. To do this, we need an explicit (local) parametrization of $\mathcal{E}_{bad}$.

First, we observe that, for any $V \in \mathcal{M}_{p,\text{non inj}}$ and $X \in O(p)$,

- VX also belongs to $\mathcal{M}_{p,\text{non inj}}$;
- $\{C_2 \in \mathbb{S}^{n\times n}, C_2 V = 0\} = \{C_2 \in \mathbb{S}^{n\times n}, C_2(VX) = 0\}$.

Consequently, when parametrizing $\mathcal{E}_{bad}$, we do not need to consider all matrices $V \in \mathcal{M}_{p,\text{non inj}}$: We can restrict ourselves to a subset of $\mathcal{M}_{p,\text{non inj}}$ whose orbit under the multiplicative action of $O(p)$ is the full $\mathcal{M}_{p,\text{non inj}}$. The goal of the next proposition (proved in Appendix A) is to adequately define such a subset.

Proposition 1. *There exists a sequence $(\mathcal{M}_p^{(s)})_{s\in\mathbb{N}}$ of submanifolds of $\mathcal{M}_p$ such that*

1. *for any $s \in \mathbb{N}$, $\dim \mathcal{M}_p^{(s)} = \dim \mathcal{M}_p - \frac{p(p-1)}{2}$;*
2. *$\{V \in \mathcal{M}_p, \text{rank}(V) = p\} = \{VX, V \in \bigcup_{s\in\mathbb{N}} \mathcal{M}_p^{(s)}, X \in O(p)\}$;*
3. *for any $s \in \mathbb{N}$, $\mathcal{M}_{p,non\ inj} \cap \mathcal{M}_p^{(s)}$ has measure zero in $\mathcal{M}_p^{(s)}$.*

Then, for any matrix $V \in \mathcal{M}_p$ with rank p, we need an explicit parametrization of $\{C_2 \in \mathbb{S}^{n\times n}, C_2 V = 0\}$. We denote by $Z_V \in \mathbb{R}^{n\times n}$ an invertible matrix such that

$$Z_V V = \begin{pmatrix} I_p \\ 0_{n-p,p} \end{pmatrix}. \tag{3.3}$$

Such a Z_V is not unique, but necessarily exists, because $\text{rank}(V) = p$. With this definition, for any V,

$$\begin{aligned}\{C_2 \in \mathbb{S}^{n\times n}, C_2 V = 0\} &= \{C_2 \in \mathbb{S}^{n\times n}, (Z_V^{-1})^T C_2 Z_V^{-1} Z_V V = 0\} \\ &= \left\{C_2 \in \mathbb{S}^{n\times n}, (Z_V^{-1})^T C_2 Z_V^{-1} \begin{pmatrix} I_p \\ 0_{n-p,p} \end{pmatrix} = 0\right\} \\ &= \left\{Z_V^T \begin{pmatrix} 0_{p,p} & 0_{p,n-p} \\ 0_{n-p,p} & \Gamma \end{pmatrix} Z_V, \Gamma \in \mathbb{S}^{(n-p)\times(n-p)}\right\}. \end{aligned} \tag{3.4}$$

Additionally, in a small enough neighborhood of any matrix, it is possible to define Z_V satisfying Equation (3.3) in such a way that the map $V \to Z_V$ is smooth. Up to splitting the submanifolds $\mathcal{M}_p^{(s)}$ into (a still countable number of) open subsets, we can thus assume that, for each $s \in \mathbb{N}$, there exists a smooth map

$$V \in \mathcal{M}_p^{(s)} \to Z_V^{(s)} \in \mathbb{R}^{n\times n}$$

such that Equation (3.3) holds with $Z_V^{(s)}$ in place of Z_V, for any $V \in \mathcal{M}_p^{(s)}$.

We now present our parametrization of $\mathcal{E}_{bad}$. For any $s \in \mathbb{N}$, we define

$$\begin{aligned} \zeta_s : \quad \mathbb{R}^m \times \mathcal{M}_p^{(s)} \times \mathbb{S}^{(n-p)\times(n-p)} &\to \mathbb{S}^{n\times n} \\ (\mu, V, \Gamma) &\to \mathcal{A}^*(\mu) + Z_V^T \begin{pmatrix} 0_{p,p} & 0_{p,n-p} \\ 0_{n-p,p} & \Gamma \end{pmatrix} Z_V. \end{aligned}$$

With this definition, we have that

$$\mathcal{E}_{bad} \subset \bigcup_{s\in\mathbb{N}} \zeta_s \left(\mathbb{R}^m \times (\mathcal{M}_{p,\text{non inj}} \cap \mathcal{M}_p^{(s)}) \times \mathbb{S}^{(n-p)\times(n-p)}\right). \tag{3.5}$$

Indeed, let us consider any element $C \in \mathcal{E}_{bad}$. We write

$$C = \mathcal{A}^*(\mu) + C_2,$$

with $C_2 V = 0$ for some $V \in \mathcal{M}_{p,\text{non inj}}$ such that $\text{rank}(V) = p$. From Property 2 of Proposition 1, there exists $s \in \mathbb{N}, \tilde{V} \in \mathcal{M}_p^{(s)}$ and $X \in O(p)$ such that

$$V = \tilde{V} X.$$

Because $C_2 V = 0$, we also have $C_2\tilde{V} = 0$, so from Equation (3.4), C_2 can be written as

$$C_2 = Z_{\tilde{V}}^T \begin{pmatrix} 0_{p,p} & 0_{p,n-p} \\ 0_{n-p,p} & \Gamma \end{pmatrix} Z_{\tilde{V}},$$

for some $\Gamma \in \mathbb{S}^{(n-p)\times(n-p)}$.

Consequently, $C = \zeta_s(\mu, \tilde{V}, \Gamma)$. As V belongs to $\mathcal{M}_{p,\text{non inj}}$ and $\tilde{V} = VX^{-1}$, $\tilde{V}$ is also an element of $\mathcal{M}_{p,\text{non inj}}$, so

$$C \in \zeta_s\left(\mathbb{R}^m \times (\mathcal{M}_{p,\text{non inj}} \cap \mathcal{M}_p^{(s)}) \times \mathbb{S}^{(n-p)\times(n-p)}\right).$$

This proves Equation (3.5).

Let us conclude. For any $s \in \mathbb{N}$, $\mathbb{R}^m \times \mathcal{M}_p^{(s)} \times \mathbb{S}^{(n-p)\times(n-p)}$ is a manifold with dimension

$$\begin{aligned}
&\dim(\mathbb{R}^m \times \mathcal{M}_p^{(s)} \times \mathbb{S}^{(n-p)\times(n-p)}) \\
&\qquad = m + \dim \mathcal{M}_p - \frac{p(p-1)}{2} + \frac{(n-p)(n-p+1)}{2} \\
&\qquad = m + np - m - \frac{p(p-1)}{2} + \frac{(n-p)(n-p+1)}{2} \\
&\qquad = \frac{n(n+1)}{2} \\
&\qquad = \dim(\mathbb{S}^{n\times n}).
\end{aligned}$$

From Property 3 in Proposition 1, the set $\mathcal{M}_{p,\text{non inj}} \cap \mathcal{M}_p^{(s)}$ has zero measure in $\mathcal{M}_p^{(s)}$, so $\mathbb{R}^m \times (\mathcal{M}_{p,\text{non inj}} \cap \mathcal{M}_p^{(s)}) \times \mathbb{S}^{(n-p)\times(n-p)}$ is a subset with zero measure of $\mathbb{R}^m \times \mathcal{M}_p^{(s)} \times \mathbb{S}^{(n-p)\times(n-p)}$. As ζ_s is a smooth map from $\mathbb{R}^m \times \mathcal{M}_p^{(s)} \times \mathbb{S}^{(n-p)\times(n-p)}$ to $\mathbb{S}^{n\times n}$ and as $\dim\left(\mathbb{R}^m \times \mathcal{M}_p^{(s)} \times \mathbb{S}^{(n-p)\times(n-p)}\right) = \dim(\mathbb{S}^{n\times n})$, the set

$$\zeta_s\left(\mathbb{R}^m \times (\mathcal{M}_{p,\text{non inj}} \cap \mathcal{M}_p^{(s,t)}) \times \mathbb{S}^{(n-p)\times(n-p)}\right)$$

has measure zero in $\mathbb{S}^{n\times n}$.

Combined with Equation (3.5), this shows that $\mathcal{E}_{bad}$ has zero Lebesgue measure in $\mathbb{S}^{n\times n}$.

4 Proof of Theorem 2

Theorem (Theorem 2)**.** *Let $r \in \mathbb{N}^*$ be fixed. Let $p \geq r$ be such that*

$$\frac{p(p+1)}{2} + pr \leq m.$$

We make the following hypotheses:

1. *$\mathcal{C}$ has at least one extreme point with rank r, that we denote by X_0;*
2. *$(\mathcal{A}, b)$ is p-regular;*
3. *There exists $V \in \mathcal{M}_p$ such that $\mathcal{M}_p$ is X_0-minimally secant at V.*

Then there exists a subset $\mathcal{E}_{bad}$ of $\mathbb{S}^{n\times n}$ with non-zero Lebesgue measure such that, for any cost matrix $C \in \mathcal{E}_{bad}$,

- *Problem (SDP) has a unique global minimizer, which has rank r.*
- *Problem (Factorized SDP) has at least one second-order critical point Y that is not a global minimizer.*

For the whole proof, we fix $X_0 \in \mathbb{S}^{n\times n}, V \in \mathcal{M}_p$ such that

1. X_0 has rank r and is an extreme point of $\mathcal{C}$.
2. $\mathcal{M}_p$ is X_0-minimally secant at V.

The proof is in two parts. First, we show the existence of one cost matrix C for which X_0 is the unique global minimizer of (SDP), and V is a second-order critical point of Problem (Factorized SDP) but not a global minimizer. In the second part, we show that, under some additional "non-degeneracy" conditions, the two properties that we want to be satisfied on the set $\mathcal{E}_{bad}$ (unique global minimizer for Problem (SDP), with rank r, and existence of a second-order critical, but not globally optimal, point for Problem (Factorized SDP)) are stable to perturbations. Hence, since they are satisfied by C, they are satisfied by all cost matrices in a small ball around C, and it suffices to define $\mathcal{E}_{bad}$ as this small ball.

The first part is embodied by the following lemma, proved in Subsection 4.1.

Lemma 1. *There exists a cost matrix $C \in \mathbb{S}^{n\times n}$ such that*

- *Problem (SDP) has a unique global minimizer, which is the rank r matrix X_0.*
- *The matrix V is a second-order critical point of Problem (Factorized SDP), but not a global minimizer.*
- *The cost matrix C can be written as $C = C_1 + \mathcal{A}^*(g_1)$, with $g_1 \in \mathbb{R}^m$ and $C_1 \in \mathbb{S}^{n\times n}$ such that $C_1 \succeq 0, \operatorname{rank}(C_1) = n - r$ and $C_1 X_0 = 0$.*
- *The Hessian of (Factorized SDP) at V has exactly $\frac{p(p-1)}{2}$ zero eigenvalues.*

In this lemma, the first two properties are the most important ones. The third and fourth ones are the previously mentioned "additional non-degeneracy conditions", that are necessary for the second part of the proof to work.

Let us come to the second part of the proof: The two properties required from all cost matrices in $\mathcal{E}_{bad}$ are stable under perturbations in some neighborhood of C. We treat the two properties with two separate lemmas, respectively proved in Subsections 4.2 and 4.3.

Lemma 2. *For some cost matrix C, we assume that Problem (SDP) has X_0 as unique global minimizer. We also assume that C can be written as*

$$C = C_1 + \mathcal{A}^*(g_1),$$

with $g_1 \in \mathbb{R}^m, C_1 \in \mathbb{S}^{n\times n}$ such that $C_1 \succeq 0, \operatorname{rank}(C_1) = n - r$ and $C_1 X_0 = 0$.

Then, for any C' close enough to C, Problem (SDP) (with cost matrix C') also has a unique global minimizer, and this minimizer has rank r.

Lemma 3. *For some cost matrix C, we assume that Problem (Factorized SDP) has a second-order critical point $V \in \mathcal{M}_p$, with rank p, such that the Hessian of (Factorized SDP) at V has exactly $\frac{p(p-1)}{2}$ zero eigenvalues.*

Let η be any positive real number.

For any C' close enough to C, Problem (Factorized SDP) (with cost matrix C') has a second-order critical point in the ball $B(V, \eta)$.

The proof of the theorem is now finished. Indeed, let $C \in \mathbb{S}^{n\times n}$ be as in Lemma 1. As V is not a global minimizer of Problem (Factorized SDP), VV^T is not a global minimizer of Problem (SDP):

$$\langle C, X_0\rangle < \left\langle C, VV^T\right\rangle.$$

Let $\epsilon, \eta > 0$ be such that, for all $C' \in B(C,\epsilon), V' \in B(V,\eta)$,

$$\langle C', X_0\rangle < \left\langle C', V'V'^T\right\rangle. \tag{4.1}$$

From Lemmas 2 and 3, up to replacing ϵ with a smaller value, we can assume that, for any $C' \in B(C,\epsilon)$,

- Problem (SDP) has a unique global minimizer, with rank r;
- Problem (Factorized SDP) has a second-order critical point in the ball $B(V,\eta)$, that we denote by $V_{C'}$.

Consequently, if we show that, for any $C' \in B(C, \epsilon)$, the second-order critical point $V_{C'}$ is not a global minimizer of Problem (Factorized SDP), we can set $\mathcal{E}_{bad} = B(C, \epsilon)$. For any $C' \in B(C, \epsilon)$,

$$\begin{aligned} \min_{V' \in \mathcal{M}_p} \langle C', V'V'^T \rangle &= \min_{X \in \mathcal{C}} \langle C', X \rangle \\ &\leq \langle C', X_0 \rangle \\ &\overset{\text{Eq (4.1)}}{<} \langle C', V_{C'} V_{C'}^T \rangle , \end{aligned}$$

so $V_{C'}$ is indeed not a global minimizer. (The first equality is true because, as $p \geq r$, Problems (SDP) and (Factorized SDP) have the same minimum.)

4.1 Proof of Lemma 1

Lemma (Lemma 1)**.** *There exists a cost matrix $C \in \mathbb{S}^{n\times n}$ such that*

- *Problem (SDP) has a unique global minimizer, which is the rank r matrix X_0.*
- *The matrix V is a second-order critical point of Problem (Factorized SDP), but not a global minimizer.*
- *The cost matrix C can be written as $C = C_1 + \mathcal{A}^*(g_1)$, with $g_1 \in \mathbb{R}^m$ and $C_1 \in \mathbb{S}^{n\times n}$ such that $C_1 \succeq 0, \mathrm{rank}(C_1) = n - r$ and $C_1 X_0 = 0$.*
- *The Hessian of (Factorized SDP) at V has exactly $\frac{p(p-1)}{2}$ zero eigenvalues.*

  Let $U_0 \in \mathbb{R}^{n\times r}$ be such that

$$X_0 = U_0 U_0^T.$$

As X_0 has rank r, U_0 also has rank r.

We start with a technical observation, proved in Paragraph C.1, that is not conceptually important, but will be useful throughout the proof.

Proposition 2. *Because $\mathcal{M}_p$ is X_0-minimally secant at V,*

$$\mathrm{Range}(U_0) \cap \mathrm{Range}(V) = \{0\}$$

and

$$\mathrm{Rank}(V) = p.$$

The construction of C is divided in two parts, respectively presented in Paragraphs 4.1.1 and 4.1.2. In the first part, we show that, to construct $C \in \mathbb{S}^{n\times n}$ satisfying the four required properties, it suffices to construct a vector $g_1 \in \mathbb{R}^m$ and a matrix $C_2 \in \mathbb{S}^{n\times n}$ fulfilling three different properties, that are less intuitive but easier to manipulate. In the second part, we construct such g_1, C_2.

4.1.1 First step

We start with three lemmas, that provide explicit sufficient conditions on C for the following properties to hold:

- X_0 is the unique global minimizer of Problem (SDP) (Lemma 4).
- V is a first-order critical point of Problem (Factorized SDP) (Lemma 5).
- In addition to being a first-order critical point of Problem (Factorized SDP), V is a second-order critical point, and the Hessian matrix has $\frac{p(p-1)}{2}$ zero eigenvalues (Lemma 6).

Their proofs are in Subsections C.2, C.3 and C.4 of the appendix.

Lemma 4. *If $C \in \mathbb{S}^{n\times n}$ can be written as $C = C_1 + \mathcal{A}^*(g_1)$, with*

- *$g_1 \in \mathbb{R}^m$,*
- *$C_1 \in \mathbb{S}^{n\times n}$ such that $C_1 \succeq 0$, $\mathrm{rank}(C_1) = n - r$ and $C_1 U_0 = 0$,*

then X_0 is the unique global minimizer of (SDP).

Lemma 5. *For $C \in \mathbb{S}^{n\times n}$, the matrix V is a first-order critical point of Problem (Factorized SDP) if and only if C can be written in the form*

$$C = C_2 + \mathcal{A}^*(g_2),$$

with $g_2 \in \mathbb{R}^m$ and $C_2 \in \mathbb{S}^{n\times n}$ such that $C_2 V = 0$.

Lemma 6. *For $C \in \mathbb{S}^{n\times n}$, if $V \in \mathcal{M}_p$ is a first-order critical point of Problem (Factorized SDP), then it is a second-order critical point if and only if*

$$\forall \dot{V} \in T_V\mathcal{M}_p, \quad \left\langle C_2, \dot{V}\dot{V}^T \right\rangle \geq 0, \tag{4.2}$$

where C_2 is the unique matrix satisfying the properties of Lemma 5.

When this happens, the Hessian of Problem (Factorized SDP) at V has exactly $\frac{p(p-1)}{2}$ zero eigenvalues if and only if, in Equation (4.2), the equality is attained exactly for matrices $\dot{V}$ of the form

$$\dot{V} = VA, \quad A \in \mathrm{Anti}(p).$$

From Lemma 4, we see that the first property in the statement of Lemma 1 is implied by the third one. Thus, to find C satisfying the four properties of the lemma statement, it is enough to find C satisfying the second, third and fourth properties.

We also observe that, if X_0 is the unique global minimizer of Problem (SDP) and V a second-order critical point of Problem (Factorized SDP), then VV^T is not a global minimizer of (SDP) (since $VV^T \neq X_0$, as a consequence of Proposition 2). Problems (Factorized SDP) and (SDP) have the same minimum (because of the assumption $p \geq r$), so V is not a global minimizer of Problem (Factorized SDP) either. Therefore, in the second property of Lemma 1, we can neglect the fact that V must not be a global minimizer of Problem (Factorized SDP).

To summarize, the four properties of Lemma 1 are equivalent to the following three ones:

- The matrix C can be written as $C = C_1 + \mathcal{A}^*(g_1)$, with $g_1 \in \mathbb{R}^m$ and $C_1 \in \mathbb{S}^{n\times n}$ such that $C_1 \succeq 0, \mathrm{rank}(C_1) = n - r$ and $C_1 X_0 = 0$.
- The matrix V is a second-order critical point of Problem (Factorized SDP).
- The Hessian of (Factorized SDP) at V has exactly $\frac{p(p-1)}{2}$ zero eigenvalues.

From Lemmas 4, 5 and 6, to construct a matrix C with these three properties, it suffices to find C_1, g_1, C_2, g_2 satisfying the conditions appearing in the lemmas and such that, in addition,

$$C_1 + \mathcal{A}^*(g_1) = C_2 + \mathcal{A}^*(g_2).$$

(Indeed, we can then set $C = C_1 + \mathcal{A}^*(g_1)$.)

Without loss of generality, we can[5] impose $g_2 = 0$. Then, we rewrite our goal as finding $g_1 \in \mathbb{R}^m, C_2 \in \mathbb{S}^{n\times n}$ such that

1. $C_1 \stackrel{def}{=} C_2 - \mathcal{A}^*(g_1)$ is a positive matrix with rank $n - r$, such that $C_1 U_0 = 0$;
2. $C_2 V = 0$;
3. for all $\dot{V} \in T_V\mathcal{M}_p$, $\left\langle C_2, \dot{V}\dot{V}^T \right\rangle \geq 0$, with equality if and only if $\dot{V} = VA$ for some $A \in \mathrm{Anti}(p)$.

4.1.2 Second step

We now prove that it is possible to find g_1, C_2 satisfying the three conditions at the end of the previous paragraph.

We begin with a lemma showing that, if we can find a pair (g_1, C_2) satisfying Conditions 1 and 2, then, from this pair, we can construct another one, that satisfies Conditions 1, 2 and 3. Its proof is in Subsection C.5 of the appendix.

[5]Indeed, a 4-tuplet (C_1, g_1, C_2, g_2) satisfies all the desired properties if and only if $(C_1, g_1 - g_2, C_2, 0)$ also satisfies these properties.

Lemma 7. *Let us assume that $g_1 \in \mathbb{R}^m, C_2 \in \mathbb{S}^{n\times n}$ satisfy Conditions 1 and 2.*

Let us define a subspace $E_\perp$ of $\mathbb{R}^n$ as

$$E_\perp = (\mathrm{Range}(U_0) + \mathrm{Range}(V))^\perp,$$

and denote by $P_\perp \in \mathbb{S}^{n\times n}$ the matrix representing the orthogonal projection from $\mathbb{R}^n$ onto $E_\perp$ in the canonical basis.

For any $t \in \mathbb{R}^+$, we set

$$C_{2,t} = C_2 + tP_\perp.$$

Then, for all $t \in \mathbb{R}^+$ large enough, the pair $(g_1, C_{2,t})$ satisfies Conditions 1, 2 and 3.

Finally, we have to prove the existence of g_1, C_2 meeting Conditions 1 and 2. The following lemma, proved in Subsection C.6, provides a sufficient condition on $g_1 \in \mathbb{R}^m$ for C_2 to exist such that (g_1, C_2) meets Conditions 1 and 2.

Lemma 8. *Let g_1 be an element of $\mathbb{R}^m$. If*

(a) $V^T \mathcal{A}^*(g_1)V \prec 0$,

(b) $U_0^T \mathcal{A}^*(g_1)V = 0_{r,p}$,

then there exists $C_2 \in \mathbb{S}^{m\times m}$ such that g_1, C_2 satisfy Conditions 1 and 2.

To conclude, let us show that there exists $g_1 \in \mathbb{R}^m$ satisfying Conditions (a) and (b). We observe that the dual of ψ_V (defined in Equation (2.2)) is

$$\begin{array}{cccc} \psi_V^* & : & \mathbb{R}^m & \to & \mathbb{S}^{p\times p} \times \mathbb{R}^{r\times p} \\ & & g & \to & \left(2V^T \mathcal{A}^*(g)V, 2U_0^T \mathcal{A}^*(g)V\right). \end{array}$$

Since $\mathcal{M}_p$ is X_0-minimally secant at V, ψ_V is injective, so ψ_V^* is surjective. In particular, it means that there exists g_1 such that

$$V^T \mathcal{A}^*(g_1)V = -I_p \qquad \text{and} \qquad U_0^T \mathcal{A}^*(g_1)V = 0_{r,p}.$$

Such a g_1 satisfies Conditions (a) and (b).

4.2 Proof of Lemma 2

Lemma (Lemma 2)**.** *For some cost matrix C, we assume that Problem (SDP) has X_0 as unique global minimizer. We also assume that C can be written as*

$$C = C_1 + \mathcal{A}^*(g_1),$$

with $g_1 \in \mathbb{R}^m$ and $C_1 \in \mathbb{S}^{n\times n}$ such that $C_1 \succeq 0, \mathrm{rank}(C_1) = n - r$ and $C_1 X_0 = 0$.

Then, for any matrix C' close enough to C, Problem (SDP) (with cost matrix C') also has a unique global minimizer, and this minimizer has rank r.

Proof. We start with a technical proposition, showing the existence and continuity of minimizers of Problem (SDP) when the cost matrix is close to C. The proof is in Subsection C.7.

Proposition 3. *For any $\epsilon > 0$, there exists $\zeta > 0$ such that, when $C' \in B(C, \zeta)$,*

- *Problem (SDP) (with cost matrix C') admits at least one minimizer;*
- *all minimizers of Problem (SDP) belong to the ball $B(X_0, \epsilon)$.*

A second technical proposition, proved in Subsection C.8 shows that (because of the p-regularity assumption) Problem (SDP) satisfies Slater's condition. It is useful for the proof, as it allows to characterize the minimizers of Problem (SDP) in terms of Karush-Kuhn-Tucker conditions.

Proposition 4. *Slater's condition is satisfied: the feasible set $\mathcal{C}$ of Problem (SDP) contains a matrix X such that $X \succ 0$.*

To establish the lemma, it suffices to show that, for any sequence $(C'_k)_{k\in\mathbb{N}}$ of cost matrices converging to C, Problem (SDP) with cost matrix C'_k has a unique minimizer, and this minimizer has rank r, as soon as k is large enough.

Let $(C'_k)_{k\in\mathbb{N}}$ be such a sequence, and let us denote by $(X'_k)_{k\in\mathbb{N}}$ an associated sequence of minimizers (it exists from Proposition 3, at least for k large enough). Let us show that, for any k large enough,

$$\text{rank}(X'_k) = r \text{ and } X'_k \text{ is the unique minimizer of Problem (SDP).} \tag{4.3}$$

For any k, since Slater's condition is true from Proposition 4, strong duality holds, and the Karush-Kuhn-Tucker conditions apply at X'_k. It means that C'_k can be written in the form

$$C'_k = D_k + \mathcal{A}^*(h_k),$$

with $h_k \in \mathbb{R}^m$ and $D_k \in \mathbb{S}^{n\times n}$ such that $D_k \succeq 0$ and $D_k X'_k = 0$.

Because $(\mathcal{A}, b)$ is p-regular, one can show that $D_k \overset{k\to+\infty}{\to} C_1$ and $h_k \overset{k\to+\infty}{\to} g_1$. This is what the following lemma says; its proof is in Subsection C.9.

Lemma 9. *When k goes to infinity,*

$$D_k \to C_1 \quad \textit{and} \quad h_k \to g_1.$$

For any k, because $D_k X'_k = 0$,

$$\text{rank}(D_k) + \text{rank}(X'_k) \leq n. \tag{4.4}$$

From Proposition 3, $(X'_k)_{k\in\mathbb{N}}$ converges to X_0, and from Lemma 9, $(D_k)_{k\in\mathbb{N}}$ converges to C_1. In particular, for k large enough,

$$\begin{aligned} &\text{rank}(X'_k) \geq \text{rank}(X_0) = r \\ \text{and} \quad &\text{rank}(D_k) \geq \text{rank}(C_1) = n - r. \end{aligned}$$

Combined with Equation (4.4), this proves that, for k large enough,

$$\text{rank}(X'_k) = r \quad \text{and} \quad \text{rank}(D_k) = n - r.$$

This establishes the first part of Property (4.3). The second part (that the minimizer is unique) is a direct consequence of Lemma 4, with C_1, g_1, X_0 replaced by D_k, h_k, X'_k. □

4.3 Proof of Lemma 3

Lemma (Lemma 3)**.**
For some cost matrix C, we assume that Problem (Factorized SDP) has a second-order critical point $V \in \mathcal{M}_p$, with rank p, such that the Hessian of (Factorized SDP) at V has exactly $\frac{p(p-1)}{2}$ zero eigenvalues.

Let η be any positive real number.

For any C' close enough to C, Problem (Factorized SDP) (with cost matrix C') has a second-order critical point in the ball $B(V, \eta)$.

Proof of Lemma 3. For any cost matrix $C' \in \mathbb{S}^{n\times n}$, let us denote by $f_{C'}$ the cost function appearing in Problem (Factorized SDP):

$$f_{C'} : W \in \mathcal{M}_p \to \langle C', WW^T \rangle \in \mathbb{R}.$$

To establish the lemma, we show that there exists a smooth map g, defined over a neighborhood of C, such that $g(C) = V$ and, for any C', $g(C')$ is a first-order critical point of (Factorized SDP). Once this is done, with a simple continuity argument on $C' \to \text{Hess}f_{C'}(g(C'))$, we show that, for any C', $g(C')$ is actually a second-order critical point. This implies the lemma.

Construction of g such that, for any C', $g(C')$ is a first-order critical point: To construct g, the first intuition is to look at the map

$$(C', W) \in \mathbb{S}^{n\times n} \times \mathcal{M}_p \to \text{grad}f_{C'}(W) \in T_W\mathcal{M}_p,$$

show that its differential along $\mathcal{M}_p$ is invertible at (C, V), and deduce from this and the implicit function theorem that, for any C' close enough to C, there is a W, smoothly depending on C', such that $\text{grad}f_{C'}(W) = 0$.

Actually, because of the invariance properties of $f_{C'}$, $\text{grad}f_{C'}(W)$ has particular properties, formalized in the following proposition (proved in Subsection C.10), that prevent the differential from being invertible.

Proposition 5. *For any cost matrix $C' \in \mathbb{S}^{n\times n}$, for any $W \in \mathcal{M}_p$,*

$$T_W\mathcal{M}_p \supset \{WA, A \in \mathrm{Anti}(p)\},$$

and, denoting $f_{C'}$ the cost function of Problem (Factorized SDP),

$$\mathrm{grad} f_{C'}(W) \in \{WA, A \in \mathrm{Anti}(p)\}^\perp.$$

Additionally, if W is a first-order critical point of $f_{C'}$,

$$\{WA, A \in \mathrm{Anti}(p)\} \subset \mathrm{Ker}(\mathrm{Hess} f_{C'}(W)).$$

When $\mathrm{rank}(W) = p$, this inclusion is an equality if and only if $\mathrm{Hess} f_{C'}(W)$ has exactly $\frac{p(p-1)}{2}$ zero eigenvalues.

To overcome this technical issue, we consider $\widetilde{\mathcal{M}}_p$, an arbitrary submanifold of $\mathcal{M}_p$ containing V, with dimension $\dim(\mathcal{M}_p) - \frac{p(p-1)}{2}$, such that

$$T_V\widetilde{\mathcal{M}}_p = T_V\mathcal{M}_p \cap \{VA, A \in \mathrm{Anti}(p)\}^\perp,$$

and define

$$\begin{array}{rccc} \chi: & \mathbb{S}^{n\times n} \times \widetilde{\mathcal{M}}_p & \to & T_V\widetilde{\mathcal{M}}_p \\ & (C', W) & \to & P_V(\mathrm{grad} f_{C'}(W)), \end{array}$$

where P_V denotes the orthogonal projection from $\mathbb{R}^{n\times p}$ to $T_V\widetilde{\mathcal{M}}_p$ (it is a homeomorphism when restricted to $T_W\mathcal{M}_p \cap \{WA, A \in \mathrm{Anti}(p)\}^\perp$, for any W close enough to V; its role is simply to ensure that χ takes its values in a set that does not depend on W).

With this definition,

$$d_W\chi(C, V) : T_V\widetilde{\mathcal{M}}_p \to T_V\widetilde{\mathcal{M}}_p \text{ is invertible.} \tag{4.5}$$

We admit this fact for the moment, and explain how to conclude the proof.

From the implicit function theorem, there is a neighborhood $\mathcal{V}_C$ of C in $\mathbb{S}^{n\times n}$ and a smooth map $g : \mathcal{V}_C \to \widetilde{\mathcal{M}}_p$ such that

$$g(C) = V;$$
$$\forall C' \in \mathcal{V}_C, \quad \chi(C', g(C')) = 0.$$

From Proposition 5, for any C', $\mathrm{grad} f_{C'}(W)$ belongs to

$$T_W\mathcal{M}_p \cap \{WA, A \in \mathrm{Anti}(p)\}^\perp,$$

and we have said that P_V is a homeomorphism from this space to $T_V\widetilde{\mathcal{M}}_p$ for any W close enough to V, so, provided that $\mathcal{V}_C$ is small enough,

$$\chi(C', g(C')) = 0 \iff \mathrm{grad} f_{C'}(g(C')) = 0.$$

Therefore for any $C' \in \mathcal{V}_C$, $g(C')$ is a first-order critical point of $f_{C'}$.

<u>For any C', $g(C')$ is a second-order critical point</u>: From the last equation in Proposition 5, $\mathrm{Hess} f_{C'}(g(C'))$ has at least $\frac{p(p-1)}{2}$ zero eigenvalues (its kernel contains a $\frac{p(p-1)}{2}$-dimensional space). When C' is close enough to C, it also has at least $\dim(\mathcal{M}_p) - \frac{p(p-1)}{2}$ strictly positive eigenvalues (because $\mathrm{Hess} f_C(V) = \mathrm{Hess} f_C(g(C))$ does). Thus, for any C' close enough to C, all eigenvalues of $\mathrm{Hess} f_{C'}(g(C'))$ are nonnegative: $g(C')$ is second-order critical.

<u>Proof of Property (4.5)</u>: For any $\dot{W} \in T_V\widetilde{\mathcal{M}}_p$,

$$d_W\chi(C, V) \cdot \dot{W} = P_V(\mathrm{Hess} f_C(V) \cdot \dot{W}) = \mathrm{Hess} f_C(V) \cdot \dot{W}.$$

The second equality is because $\mathrm{Hess} f_c(V) : T_V\mathcal{M}_p \to T_V\mathcal{M}_p$ is self-adjoint, so

$$\begin{aligned} \mathrm{Range}(\mathrm{Hess} f_c(V)) &= T_V\mathcal{M}_p \cap (\mathrm{Ker}(\mathrm{Hess} f_c(V)))^\perp \\ &\overset{(\text{Prop } 5)}{=} T_V\mathcal{M}_p \cap \{VA, A \in \mathrm{Anti}(p)\}^\perp \\ &= T_V\widetilde{\mathcal{M}}_p. \end{aligned}$$

Therefore, for any $\dot{W} \in T_V\widetilde{M}_p$, $d_W\chi(C,V)\cdot\dot{W} = 0$ if and only if

$$\dot{W} \in \mathrm{Ker}(\mathrm{Hess} f_C(V)) \overset{\text{(Prop 5)}}{=} \{VA, A \in \mathrm{Anti}(p)\}.$$

From the definition of $T_V\widetilde{M}_p$, this is only possible if $\dot{W} = 0$. Consequently, $d_W\chi(C,V)$ is an injection from $T_V\widetilde{\mathcal{M}}_p$ to itself whence it is a bijection.

□

5 How to apply Theorems 1 and 2

In this section, we discuss the difficulties that may arise when trying to apply Theorem 1 and 2 to concrete examples of problems of the form (SDP), as done in Subsection 2.4.

The main issue is that it can be difficult to check whether the hypotheses of the theorems hold or not. Compactness, p-regularity, and existence of extreme points for $\mathcal{C}$, are typically easy to deal with. For each of the two theorems, problems happen only with the third hypothesis:

- Face regularity of $\mathcal{M}_p$ when $\frac{p(p+1)}{2} \le m$, for Theorem 1;
- Existence of V at which $\mathcal{M}_p$ is X_0-minimally secant, for Theorem 2.

While we have no "recipe" for verifying these hypotheses in the most general setting, we can provide at least partial methods in two situations:

- When one is interested in a problem with fixed dimensions (n, p, m are fixed) (Subsection 5.1),
- When the problem has a block-diagonal structure (Subsection 5.2).

The second situation contains all three examples discussed in Subsection 2.4.

5.1 When the dimensions are fixed

In this subsection, we consider a problem of the form (SDP) with fixed n, p, m, and provide an easy way to numerically determine whether the two hypotheses hold or not.

We rely on the following remark, which says that, under mild assumptions, the maps ϕ_V and ψ_V are either never injective, or injective for almost any $V \in \mathcal{M}_p$. Its proof is in Subsection D.1, and is based on an analyticity argument.

Remark 4. *Let $p \in \mathbb{N}^*$ be fixed. We assume that $(\mathcal{A}, b)$ is p-regular and $\mathcal{M}_p$ is connected and non-empty. Then the following two statements are equivalent:*

1. *There exists $V \in \mathcal{M}_p$ such that ϕ_V is injective.*
2. *For almost any $V \in \mathcal{M}_p$, ϕ_V is injective (that is, $\mathcal{M}_p$ is face regular).*

Additionally, if X_0 is a rank r element of $\mathcal{C}$, there is also an equivalence between the following two properties:

1. *There exists $V \in \mathcal{M}_p$ such that ψ_V is injective (that is, there exists V at which $\mathcal{M}_p$ is X_0-minimally secant).*
2. *For almost any $V \in \mathcal{M}_p$, ψ_V is injective.*

Consequently, to numerically check the face regularity of $\mathcal{M}_p$, it suffices to pick a "generic"[6] element V_0 in $\mathcal{M}_p$ and to determine whether ϕ_{V_0} is injective. If it is, it implies that ϕ_V is injective for almost any $V \in \mathcal{M}_p$, and $\mathcal{M}_p$ is face regular. Otherwise, $\mathcal{M}_p$ is not face regular with probability 1.

The same principle applies to the property of $\mathcal{M}_p$ being X_0-minimally secant at some V.

[6] By "generic", we mean "chosen according to a probability density that is absolutely continuous with respect to the measure $\lambda_{\mathcal{M}_p}$".

5.2 When $\mathcal{A}$ is block-diagonal

In this subsection, we consider the case where the operator $\mathcal{A}$ is block-diagonal.

Formally, let $r \in \mathbb{N}^*$ and $\tilde{b} \in \mathbb{R}^{r(r+1)/2}$ be fixed. For any $d \in \mathbb{N}^*$, we consider a linear operator

$$\tilde{\mathcal{A}}_d : \mathbb{S}^{d\times d} \to \mathbb{R}^{r(r+1)/2}.$$

For any $S \in \mathbb{N}^*$, and S-tuplet $(d_1, \dots, d_S) \in (\mathbb{N}^*)^S$, we define

$$\begin{array}{rccc} \mathcal{A}_{(d_1,\dots,d_S)}: & \mathbb{S}^{n\times n} & \to & \mathbb{R}^{S\frac{r(r+1)}{2}} \\ & M & \to & (\tilde{\mathcal{A}}_{d_1}(\mathrm{Block}_1(M)), \dots, \tilde{\mathcal{A}}_{d_S}(\mathrm{Block}_S(M))), \end{array}$$

where $n = d_1 + \dots + d_S$ and $\mathrm{Block}_1(M), \dots, \mathrm{Block}_S(M)$ are the successive diagonal blocks of M, with sizes $d_1 \times d_1, \dots, d_S \times d_S$.

Setting

$$b_S = (\tilde{b}, \dots, \tilde{b}) \in \mathbb{R}^{S\frac{r(r+1)}{2}},$$

we consider Problem (SDP) with $\mathcal{A} = \mathcal{A}_{(d_1,\dots,d_S)}$ and $b = b_S$. We denote by $\mathcal{C}^{(d_1,\dots,d_S)}$ its feasible set, and, for any p, by $\mathcal{M}_p^{(d_1,\dots,d_S)}$ the feasible set of its rank p Burer-Monteiro factorization (Factorized SDP).

This framework covers all three examples discussed in Subsection 2.4: Problem (SDP-Maxcut) is the case where $r = d_1 = \dots = d_S = 1$, and

$$\tilde{A}_1 = \mathrm{Id}_{\mathbb{R}\to\mathbb{R}} \quad \text{and} \quad \tilde{b} = 1.$$

Problem (SDP-Orthogonal-Cut) corresponds to $r = d_1 = \dots = d_S = d$ and

$$\tilde{A}_d = T_{sup} \quad \text{and} \quad \tilde{b} = T_{sup}(I_d).$$

Finally, Problem (SDP-Product) is the case where $r = 1$, $d_1, \dots, d_S$ may be arbitrary and, for any $d \in \mathbb{N}^*$,

$$\tilde{A}_d = \mathrm{Trace} \quad \text{and} \quad \tilde{b} = 1.$$

In this setting where $\mathcal{A}$ is block-diagonal, an important remark can be made: The face regularity can be deduced from the existence of V at which the minimally secant property holds. Hence, when we want to apply Theorems 1 and 2, it suffices to establish the hypotheses of Theorem 2. If these hold, they imply the ones of Theorem 1 (except for the compactness one, which we have said is generally easy to check). The proof of this remark is in Subsection D.2.

Remark 5. *Let $(d_s)_{s\in\mathbb{N}}$ be a sequence of positive integers. We assume that*

- *for any s, $\mathcal{C}^{(d_s)}$ has an extreme point with rank r;*
- *for any S and $p \geq r$, $(\mathcal{A}^{(d_1,\dots,d_S)}, b_S)$ is p-regular;*
- *for any S and $p \geq 2r$, $\mathcal{M}_p^{(d_1,\dots,d_S)}$ is connected and non-empty.*

Then the first property below implies the second one:

1. *For any $S \in \mathbb{N}^*$ and $p \geq r$ such that*

$$\frac{p(p+1)}{2} + pr \leq S\frac{r(r+1)}{2},$$

there exists X_0 an extreme point of $\mathcal{C}^{(d_1,\dots,d_S)}$ with rank r, and a matrix $V \in \mathcal{M}_p^{(d_1,\dots,d_S)}$ such that $\mathcal{M}_p^{(d_1,\dots,d_S)}$ is X_0-minimally secant at V.

$\Rightarrow$ 2. *For any $S \in \mathbb{N}^*$ and $p \geq 2r$ such that*

$$\frac{p(p+1)}{2} \leq S\frac{r(r+1)}{2},$$

$\mathcal{M}_p^{(d_1,\dots,d_S)}$ is face regular.

As to establishing the minimally secant property, we do not have a general method. For the three problems considered in Subsection 2.4, our proof relies on showing that, for any V, ψ_V is injective if and only if some specific family of matrices (that depends on V) spans $\mathbb{S}^{(p+r)\times(p+r)}$. This second property is easier to manipulate than, directly, the injectivity of ψ_V. We then get the result by explicitly constructing a particular V for which the spanning condition is verified. But there are probably interesting examples of problem of the form (SDP) for which this technique does not work.

Acknowledgments

A. W. acknowledges support by EPSRC grant EP/L01937X/1.

A Proof of Proposition 1

Proposition (Proposition 1)**.** *There exists a sequence* $(\mathcal{M}_p^{(s)})_{s\in\mathbb{N}}$ *of submanifolds of* $\mathcal{M}_p$ *such that*

1. *for any* $s \in \mathbb{N}$, $\dim \mathcal{M}_p^{(s)} = \dim \mathcal{M}_p - \frac{p(p-1)}{2}$*;*
2. $\{V \in \mathcal{M}_p, \mathrm{rank}(V) = p\} = \{VX, V \in \bigcup_{s\in\mathbb{N}} \mathcal{M}_p^{(s)}, X \in O(p)\}$*;*
3. *for any* $s \in \mathbb{N}$, $\mathcal{M}_{p,non\ inj} \cap \mathcal{M}_p^{(s)}$ *has measure zero in* $\mathcal{M}_p^{(s)}$*.*

Proof of Proposition 1. First, we consider an arbitrary element V_0 in $\mathcal{M}_p$, with rank p, and we define $\mathcal{M}_{p,V_0}$, a submanifold of $\mathcal{M}_p$ fulfilling the following four properties:

$$\dim \mathcal{M}_{p,V_0} = \dim \mathcal{M}_p - \frac{p(p-1)}{2}; \tag{A.1a}$$

$$\exists \eta_{V_0} > 0 \text{ such that } \mathcal{M}_p \cap B(V_0, \eta_{V_0}) \subset \{VX, V \in \mathcal{M}_{p,V_0}, X \in O(p)\}; \tag{A.1b}$$

$$\forall V \in \mathcal{M}_{p,V_0}, \mathrm{rank}(V) = p; \tag{A.1c}$$

$$\lambda_{\mathcal{M}_{p,V_0}}(\mathcal{M}_{p,\text{non inj}} \cap \mathcal{M}_{p,V_0}) = 0. \tag{A.1d}$$

Once this is done, we construct the $\mathcal{M}_p^{(s)}$ from the $\mathcal{M}_{p,V_0}$, using a compactness argument, but we first explain the construction of $\mathcal{M}_{p,V_0}$.

As $\mathcal{M}_p$ is stable by multiplication by elements of $O(p)$, it contains $V_0 O(p)$, so

$$T_{V_0}\mathcal{M}_p \supset \{V_0 A, A \in T_{I_p} O(p)\} = \{V_0 A, A \in \mathrm{Anti}(p)\}.$$

Because V_0 is injective (it is a $n \times p$ matrix with rank p), the vector space $\{V_0 A, A \in \mathrm{Anti}(p)\}$ has dimension $\dim \mathrm{Anti}(p) = \frac{p(p-1)}{2}$. Let $T^{\perp}_{V_0,\mathrm{Anti}}\mathcal{M}_p$ be its orthogonal complement in $T_{V_0}\mathcal{M}_p$; it is a $\left(\dim \mathcal{M}_p - \frac{p(p-1)}{2}\right)$-dimensional vector space.

Let $\mathcal{M}_{p,V_0}$ be any submanifold of $\mathcal{M}_p$ with dimension $\left(\dim \mathcal{M}_p - \frac{p(p-1)}{2}\right)$, containing V_0, whose tangent space at V_0 is $T^{\perp}_{V_0,\mathrm{Anti}}\mathcal{M}_p$. By choosing $\mathcal{M}_{p,V_0}$ small enough, we can ensure that all its elements have rank p, so that Property (A.1c) holds.

The map

$$\begin{array}{cccccc} \chi_{V_0}: & \mathcal{M}_{p,V_0} & \times & O(p) & \to & \mathcal{M}_p \\ & V & , & X & \to & VX \end{array}$$

is $\mathcal{C}^\infty$. Because $T_{V_0}\mathcal{M}_{p,V_0} = T^{\perp}_{V_0,\mathrm{Anti}}\mathcal{M}_p$, its differential at (V_0, I_p) is a bijection. From the inverse function theorem, χ_{V_0} thus defines a diffeomorphism from some neighborhood of (V_0, I_p) in $\mathcal{M}_{p,V_0} \times O(p)$ to some neighborhood of V_0 in $\mathcal{M}_p$. In particular, Property (A.1b) is true.

Up to reducing the size of $\mathcal{M}_{p,V_0}$, we can assume that χ_{V_0} is a diffeomorphism on $\mathcal{M}_{p,V_0} \times (O(p) \cap B(I_p, \epsilon))$, for some $\epsilon > 0$.

Let us show Property (A.1d). We observe that, for any $V \in \mathcal{M}_{p,\text{non inj}}$ and $X \in O(p)$, the matrix VX also belongs to $\mathcal{M}_{p,\text{non inj}}$, so, for any $\epsilon > 0$ small enough,

$$\begin{aligned} &\lambda_{\mathcal{M}_p}(\mathcal{M}_{p,\text{non inj}}) \\ &\geq \lambda_{\mathcal{M}_p}\left(\chi_{V_0}\Big((\mathcal{M}_{p,V_0} \cap \mathcal{M}_{p,\text{non inj}}) \times (O(p) \cap B(I_p, \epsilon))\Big)\right) \\ &= \int_{\mathcal{M}_{p,V_0} \times (O(p)\cap B(I_p,\epsilon))} 1_{\mathcal{M}_{p,\text{non inj}}}(V) |\det d\chi_{V_0}(V, X)| d\lambda_{\mathcal{M}_{p,V_0}}(V) d\lambda_{O(p)}(X), \end{aligned}$$

where $1_{\mathcal{M}_{p,\text{non inj}}}$ denotes the characteristic function of $\mathcal{M}_{p,\text{non inj}}$.

As $\mathcal{M}_p$ is face regular, $\lambda_{\mathcal{M}_p}(\mathcal{M}_{p,\text{non inj}}) = 0$. Because $|\det d\chi_{V_0}(V, X)| > 0$ for any $(V, X) \in \mathcal{M}_{p,V_0} \times (O(p) \cap B(I_p, \epsilon))$ (for ϵ small enough), the function $(V, X) \to 1_{\mathcal{M}_{p,\text{non inj}}}(V)$ must then be zero almost everywhere on $\mathcal{M}_{p,V_0} \times (O(p) \cap B(I_p, \epsilon))$, which is equivalent to Property (A.1d).

We have thus constructed $\mathcal{M}_{p,V_0}$ enjoying Properties (A.1a) to (A.1d), as announced. Since V_0 was arbitrary, we can repeat the construction for any $V \in \mathcal{M}_p$ with rank p; it yields a submanifold $\mathcal{M}_{p,V}$ with the same properties. For any V, we fix $\eta_V > 0$ as in Property (A.1c).

Now, let $(V^{(s)})_{s\in\mathbb{N}}$ be a sequence of elements of $\mathcal{M}_p$ with rank p such that

$$\{V \in \mathcal{M}_p, \operatorname{rank}(V) = p\} \subset \bigcup_{s\in\mathbb{N}} \left(\mathcal{M}_p \cap B(V^{(s)}, \eta_{V^{(s)}})\right).$$

(Such a sequence exists because $\{V \in \mathcal{M}_p, \operatorname{rank}(V) = p\}$, as any subset of $\mathbb{R}^{n\times p}$ is Lindelöf.)

For any s, we set

$$\mathcal{M}_p^{(s)} \stackrel{def}{=} \mathcal{M}_{p,V^{(s)}}.$$

The manifolds $\mathcal{M}_p^{(s)}$ each have dimension $\dim \mathcal{M}_p - \frac{p(p-1)}{2}$ (from Property (A.1a)), which is the first property required in the statement of the proposition. Additionally,

$$\begin{aligned}
\{V \in \mathcal{M}_p, \operatorname{rank}(V) = p\} &\subset \bigcup_{s\in\mathbb{N}} \left(\mathcal{M}_p \cap B(V^{(s)}, \eta_{V^{(s)}})\right) \\
&\stackrel{\text{(A.1b)}}{\subset} \bigcup_{s\in\mathbb{N}} \{VX, V \in \mathcal{M}_{p,V^{(s)}}, X \in O(p)\} \\
&= \{VX, V \in \bigcup_{s\in\mathbb{N}} \mathcal{M}_p^{(s)}, X \in O(p)\}.
\end{aligned}$$

The converse inclusion is also true: For any s and any $V \in \mathcal{M}_p^{(s)}, X \in O(p)$, the matrix VX is an element of $\mathcal{M}_p$ (since $\mathcal{M}_p$ is stable by multiplication by elements of $O(p)$), and $\operatorname{rank}(VX) = \operatorname{rank}(V) = p$. Consequently,

$$\{V \in \mathcal{M}_p, \operatorname{rank}(V) = p\} = \{VX, V \in \bigcup_{s\in\mathbb{N}} \mathcal{M}_p^{(s)}, X \in O(p)\},$$

which is the second required property.

This third required property comes from Property (A.1d).

□

B The condition $\frac{p(p+1)}{2} + pr \le m$ is not tight

In this section, as announced in Remark 3, we provide an example where the conclusions of Theorem 2 hold true, but the assumption

$$\frac{p(p+1)}{2} + pr \le m$$

is not satisfied.

Proposition 6. *If $\mathcal{A} = \operatorname{diag}$ and $b = 1_{n,1}$ ("MaxCut case"), when $r = p = 2$ and $m = n = 6$, there exists a subset $\mathcal{E}_{bad}$ of $\mathbb{S}^{n\times n}$ with non-zero Lebesgue measure satisfying the same properties as in Theorem 2.*

Nevertheless, we have that

$$\frac{p(p+1)}{2} + pr = 7 \not\le m.$$

Proof. The proof of Theorem 2, in Section 4, starts by fixing X_0, V such that X_0 is a rank r extreme point of $\mathcal{C}$, and $\mathcal{M}_p$ is X_0-minimally secant at V. Then, it constructs $\mathcal{E}_{bad}$ from these X_0 and V.

When $\frac{p(p+1)}{2} + pr \not\le m$, it is impossible to choose X_0, V such that $\mathcal{M}_p$ is X_0-minimally secant at V (from Remark 1). Hence, the reasoning of Section 4 does not directly apply. However, the only parts of the proof where this minimally secant assumption is used are the second step in the proof of Lemma 1 (Paragraph 4.1.2) and the proof of Proposition 2.

Consequently, to establish that the conclusions of Theorem 2 remain true when $r = p = 2$ and $m = 6$, it suffices to show that, for some careful choice of X_0, V, Proposition 2 and the conclusions of Paragraph 4.1.2 are valid despite the fact that the minimally secant assumption does not hold.

We set (this precise choice was suggested by numerical experiments)

$$U_0 = \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -1 & 0 \\ 1 & 0 \\ -1 & 0 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad X_0 = U_0 U_0^T;$$

$$V = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ 0 & 1 \end{pmatrix}.$$

The matrices X_0 and V have rank 2, and we can check that they satisfy Proposition 2. We can also check that X_0 is an extreme point of $\mathcal{C}$ (it is equivalent to the equality $\{G \in \mathbb{S}^{2\times 2}, \mathrm{diag}(U_0 G U_0^T) = 0\} = \{0\}$, that can be verified by manually solving the involved linear system).

Showing that the conclusions of Paragraph 4.1.2 are valid means showing the existence of $g_1 \in \mathbb{R}^6, C_2 \in \mathbb{S}^{6\times 6}$ such that

1. $C_1 \overset{def}{=} C_2 - \mathrm{Diag}(g_1)$ is a positive matrix with rank $n - r = 4$, such that $C_1 U_0 = 0$;
2. $C_2 V = 0$;
3. for all $\dot{V} \in T_V \mathcal{M}_2$, $\left\langle C_2, \dot{V}\dot{V}^T \right\rangle \geq 0$, with equality if and only if $\dot{V} = VA$ for some $A \in \mathrm{Anti}(2)$.

We set

$$g_1 = \begin{pmatrix} -\sqrt{5} \\ -2 + \frac{3}{\sqrt{5}} \\ -1 \\ -2 \\ 0 \\ 1 \end{pmatrix}$$

and

$$\begin{aligned} C_2 = &(G^{-1})^T \begin{pmatrix} 0_{2,2} & 0_{2,4} \\ 0_{4,2} & \begin{pmatrix} U_0 & e_1 & e_2 \end{pmatrix}^T \mathrm{Diag}(g_1) \begin{pmatrix} U_0 & e_1 & e_2 \end{pmatrix} \end{pmatrix} G^{-1} \\ &+ 20 (G^{-1})^T \begin{pmatrix} 0_{4,4} & 0_{4,2} \\ 0_{2,4} & I_2 \end{pmatrix} G^{-1}. \end{aligned}$$

where e_1, e_2 are the first two vectors of the canonical basis of $\mathbb{R}^{6\times 1}$, and $G = \begin{pmatrix} V & U_0 & e_1 & e_2 \end{pmatrix} \in \mathbb{R}^{6\times 6}$ is the horizontal concatenation of V, U_0, e_1, e_2 (it is not difficult to check that this matrix is invertible).

With this choice, Property 2 is true, because

$$C_2 V = C_2 G \begin{pmatrix} I_2 \\ 0_{4,2} \end{pmatrix} = 0.$$

For Property 1, we observe that, defining $D_1 = \mathrm{Diag}(g_1)$,

$$\begin{aligned}
&C_2 - \mathrm{Diag}(g_1) = \\
&\quad (G^{-1})^T \left(\begin{pmatrix} 0_{2,2} & 0_{2,4} \\ 0_{4,2} & \begin{pmatrix} U_0 & e_1 & e_2 \end{pmatrix}^T D_1 \begin{pmatrix} U_0 & e_1 & e_2 \end{pmatrix} \end{pmatrix} \right. \\
&\qquad\qquad \left. + 20 \begin{pmatrix} 0_{4,4} & 0_{4,2} \\ 0_{2,4} & I_2 \end{pmatrix} - G^T D_1 G \right) G^{-1} \\
&\quad = (G^{-1})^T \begin{pmatrix} -V^T D_1 V & -V^T D_1 U_0 & -V^T D_1 \begin{pmatrix} e_1 & e_2 \end{pmatrix} \\ -U_0^T D_1 V & 0_{2,2} & 0_{2,2} \\ -\begin{pmatrix} e_1 & e_2 \end{pmatrix}^T D_1 V & 0_{2,2} & 20\, I_2 \end{pmatrix} G^{-1} \\
&\quad = (G^{-1})^T \begin{pmatrix} \frac{6}{5} & \frac{6}{5} & 0 & 0 & 0 & 0 \\ \frac{6}{5} & \frac{14}{5} + \frac{2}{\sqrt{5}} & 0 & 0 & \sqrt{5} & 2 - \frac{3}{\sqrt{5}} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sqrt{5} & 0 & 0 & 20 & 0 \\ 0 & 2 - \frac{3}{\sqrt{5}} & 0 & 0 & 0 & 20 \end{pmatrix} G^{-1}.
\end{aligned}$$

From this expression, we see that $C_2 - \mathrm{Diag}(g_1)$ is a positive matrix with rank 4, and also that

$$(C_2 - \mathrm{Diag}(g_1))U_0 = (C_2 - \mathrm{Diag}(g_1))G \begin{pmatrix} 0_{2,2} \\ I_2 \\ 0_{2,2} \end{pmatrix} = 0.$$

Therefore, Property 1 is true.

We finally consider Property 3. Let us define the bilinear form

$$\begin{aligned} q : \quad T_V\mathcal{M}_2 \times T_V\mathcal{M}_2 &\to \mathbb{R} \\ (\dot{V}_1, \dot{V}_2) &\to \left\langle C_2, \dot{V}_1 \dot{V}_2^T \right\rangle. \end{aligned}$$

It contains $V \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ in its kernel: For any $\dot{V}_1 \in T_V\mathcal{M}_2$,

$$\begin{aligned} q\left(\dot{V}_1, V \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)\right) &= \left\langle C_2 V \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), \dot{V}_1 \right\rangle \\ &= \left\langle 0_{6,2}, \dot{V}_1 \right\rangle \\ &= 0. \end{aligned}$$

(The second equality comes from Property 2: $C_2 V = 0$.)

If we show that the matrix associated to q in an orthonormal basis of the 6-dimensional vector space $T_V\mathcal{M}_2$ has 5 strictly positive eigenvalues, then it proves that q is semidefinite positive, with a kernel of dimension 1, equal to $\mathbb{R}V \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, and hence implies Property 3.

It is not pleasant to check by hand that the matrix associated to q has 5 strictly positive eigenvalues. However, since we have an explicit expression for C_2, we can numerically compute the matrix associated to q in the following orthonormal basis of $T_V\mathcal{M}_2$:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

The 5 largest eigenvalues of this matrix are all larger than 0.4, hence strictly positive.

□

C Auxiliary results for the proof of Theorem 2

C.1 Proof of Proposition 2

Proposition (Proposition 2)**.** *Because $\mathcal{M}_p$ is X_0-minimally secant at V,*

$$\mathrm{Range}(U_0) \cap \mathrm{Range}(V) = \{0\}$$

and

$$\mathrm{Rank}(V) = p.$$

Proof of Proposition 2. The two desired properties are simultaneously true if the following implication holds, for all $x_1, \in \mathbb{R}^r, x_2 \in \mathbb{R}^p$:

$$(U_0 x_1 = V x_2) \qquad \Rightarrow \qquad (x_1 = 0 \text{ and } x_2 = 0).$$

Let us prove this implication. Let $x_1 \in \mathbb{R}^r, x_2 \in \mathbb{R}^p$ be such that $U_0 x_1 = V x_2$. We show that $x_1 = 0$ and $x_2 = 0$.

If we set

$$T = x_2 x_2^T \qquad \text{and} \qquad R = -x_1 x_2^T,$$

then $\left(V \; U_0 \right) \left(\begin{smallmatrix} T \\ R \end{smallmatrix} \right) = (V x_2 - U_0 x_1) x_2^T = 0$, so

$$\mathcal{A}\left(\left(V \; U_0 \right) \left(\begin{smallmatrix} T \\ R \end{smallmatrix} \right) V^T + V \left(\begin{smallmatrix} T \\ R \end{smallmatrix} \right)^T \left(V \; U_0 \right)^T \right) = \mathcal{A}(0) = 0.$$

As $\mathcal{M}_p$ is X_0-minimally secant at V, we must have $(T, R) = (0, 0)$, so $x_2 = 0$.

Since $U_0 x_1 = V x_2 = 0$ and U_0 is injective (it is a $n \times r$ matrix with rank r), we also have $x_1 = 0$. □

C.2 Proof of Lemma 4

Lemma (Lemma 4)**.** *If $C \in \mathbb{S}^{n \times n}$ can be written as $C = C_1 + \mathcal{A}^*(g_1)$, with*

- *$g_1 \in \mathbb{R}^m$,*
- *$C_1 \in \mathbb{S}^{n \times n}$ such that $C_1 \succeq 0$, $\mathrm{rank}(C_1) = n - r$ and $C_1 U_0 = 0$,*

then X_0 is the unique global minimizer of (SDP)*.*

Proof of Lemma 4. This lemma is a variation around the well-known optimality conditions of problems of the form (SDP).

Let us assume that $C = C_1 + \mathcal{A}^*(g_1)$, with C_1, g_1 as in the statement.

Let X be any feasible point of Problem (SDP). We have to show that

$$\langle X, C \rangle \geq \langle X_0, C \rangle,$$

with equality if and only if $X = X_0$.

Because $\mathcal{A}(X) = b = \mathcal{A}(X_0)$, we have that

$$\begin{aligned} \langle X - X_0, C \rangle &= \langle X - X_0, C_1 \rangle + \langle X - X_0, \mathcal{A}^*(g_1) \rangle \\ &= \langle X - X_0, C_1 \rangle + \langle \mathcal{A}(X - X_0), g_1 \rangle \\ &= \langle X - X_0, C_1 \rangle . \end{aligned}$$

Additionally, $\langle X_0, C_1 \rangle = \left\langle U_0 U_0^T, C_1 \right\rangle = \langle U_0, C_1 U_0 \rangle = 0$. As $X, C_1 \succeq 0$,

$$\langle X - X_0, C \rangle = \langle X, C_1 \rangle \geq 0.$$

This already shows that $\langle X, C \rangle \geq \langle X_0, C \rangle$. Let us show that the equality holds if and only if $X = X_0$.

The equality is attained if and only if $\langle X, C_1 \rangle = 0$. Because $X, C_1 \succeq 0$, this is equivalent to

$$\mathrm{Range}(X) \subset \mathrm{Ker}(C_1).$$

We have $\mathrm{Ker}(C_1) = \mathrm{Range}(U_0)$: $\mathrm{Range}(U_0) \subset \mathrm{Ker}(C_1)$ since $C_1 U_0 = 0$, and $\dim(\mathrm{Ker}(C_1)) = n - \mathrm{rank}(C_1) = r = \dim(\mathrm{Range}(U_0))$. Therefore, the equality is attained if and only if $\mathrm{Range}(X) \subset \mathrm{Range}(U_0)$, which is equivalent to X being of the form

$$X = U_0 T_X U_0^T,$$

for some positive matrix $T_X \in \mathbb{S}^{r\times r}$.

But if X has this form, then, for all $\epsilon > 0$ small enough,

$$\begin{aligned}(1-t)X_0 + tX = U_0(I_r + t(T_X - I_r))U_0^T \succeq 0 \text{ for all } t \in [-\epsilon; \epsilon],\\ \Rightarrow \quad (1-t)X_0 + tX \in \mathcal{C} \text{ for all } t \in [-\epsilon; \epsilon].\end{aligned}$$

As X_0 is an extreme point of $\mathcal{C}$, this is only possible if $X_0 = X$. □

C.3 Proof of Lemma 5

Lemma (Lemma 5)**.** *For $C \in \mathbb{S}^{n\times n}$, the matrix V is a first-order critical point of Problem* (Factorized SDP) *if and only if C can be written in the form*

$$C = C_2 + \mathcal{A}^*(g_2),$$

with $g_2 \in \mathbb{R}^m$ and $C_2 \in \mathbb{S}^{n\times n}$ such that $C_2 V = 0$.

Proof of Lemma 5. From [Boumal, Voroninski, and Bandeira, 2018, Eq. 7], the gradient of the cost function of (Factorized SDP) at V is

$$2\mathrm{Proj}_V(CV),$$

where $\mathrm{Proj}_V : \mathbb{R}^{n\times p} \to T_V\mathcal{M}_p$ is the orthogonal projection onto the tangent space of $\mathcal{M}_p$ at V.

Consequently, V is a first-order critical point if and only if $\mathrm{Proj}_V(CV) = 0$, that is

$$\begin{aligned}CV &\in (T_V\mathcal{M}_p)^\perp\\ &= \{\dot V \in \mathbb{R}^{n\times p}, \mathcal{A}(\dot V V^T + V\dot V^T) = 0\}^\perp\\ &= \{\dot V \in \mathbb{R}^{n\times p}, \forall g_2 \in \mathbb{R}^m, \left\langle \dot V V^T + V \dot V^T, \mathcal{A}^*(g_2)\right\rangle = 0\}^\perp\\ &= \{\dot V \in \mathbb{R}^{n\times p}, \forall g_2 \in \mathbb{R}^m, \left\langle \dot V, \mathcal{A}^*(g_2)V\right\rangle = 0\}^\perp\\ &= \left(\{\mathcal{A}^*(g_2)V, g_2 \in \mathbb{R}^m\}^\perp\right)^\perp\\ &= \{\mathcal{A}^*(g_2)V, g_2 \in \mathbb{R}^m\}.\end{aligned}$$

The fact that CV can be written as $\mathcal{A}^*(g_2)V$ for some $g_2 \in \mathbb{R}^m$ is equivalent to the fact that C can be written as $C = C_2 + \mathcal{A}^*(g_2)$, with $g_2 \in \mathbb{R}^m$ and $C_2 \in \mathbb{S}^{n\times n}$ such that $C_2 V = 0$.

□

C.4 Proof of Lemma 6

Lemma (Lemma 6)**.** *For $C \in \mathbb{S}^{n\times n}$, if $V \in \mathcal{M}_p$ is a first-order critical point of Problem* (Factorized SDP)*, then it is a second-order critical point if and only if*

$$\forall \dot V \in T_V\mathcal{M}_p, \quad \left\langle C_2, \dot V \dot V^T\right\rangle \geq 0, \tag{C.1}$$

where C_2 is the unique matrix satisfying the properties of Lemma 5.

When this happens, the Hessian of Problem (Factorized SDP) *at V has exactly $\frac{p(p-1)}{2}$ zero eigenvalues if and only if, in Equation* (4.2)*, the equality is attained exactly for matrices $\dot V$ of the form*

$$\dot V = VA, \quad A \in \mathrm{Anti}(p). \tag{C.2}$$

Proof of Lemma 6. Let $C \in \mathbb{S}^{n\times n}$ be a cost matrix such that $V \in \mathcal{M}_p$ is a first-order critical point of (Factorized SDP). We write

$$C = C_2 + \mathcal{A}^*(g_2),$$

with $C_2 V = 0$, as in Lemma 5.

Such a pair (C_2, g_2) is unique: If $C_2 + \mathcal{A}^*(g_2) = C_2' + \mathcal{A}^*(g_2')$, with $C_2 V = C_2' V = 0$, then $\mathcal{A}^*(g_2 - g_2')V = 0$, so

$$g_2 - g_2' \in \left\{ \mathcal{A}(V\dot{V}^T + \dot{V}V^T), \dot{V} \in \mathbb{R}^{n\times p} \right\}^{\perp}.$$

Because $(\mathcal{A}, b)$ is p-regular (see Definition 1), we must then have $g_2 - g_2' = 0$, hence $g_2 = g_2'$ and $C_2 = C_2'$.

Equation (C.1) is equivalent to V being second-order-critical: Let us denote f_C the cost function of Problem (Factorized SDP). For any $\dot{V} \in T_V\mathcal{M}_p$, from [Boumal, Voroninski, and Bandeira, 2018, Eq. 10],

$$\mathrm{Hess} f_C(V) \cdot (\dot{V}, \dot{V}) = 2 \left\langle S\dot{V}, \dot{V} \right\rangle, \tag{C.3}$$

where S is a matrix of the form $S = C - \mathcal{A}^*(\mu)$, for some $\mu \in \mathbb{R}^m$, that satisfies $2SV = \mathrm{grad} f_C(V) = 0$.

We have just seen that there is a unique way to write C as $C_2 + \mathcal{A}^*(g_2)$, with $g_2 \in \mathbb{R}^m$ and $C_2 V = 0$, so $\mu = g_2$ and $S = C_2$.

Now, in view of Equation (C.3), we can conclude that V is second-order critical if and only if, for any $\dot{V} \in T_V\mathcal{M}_p$,

$$\left\langle C_2, \dot{V}\dot{V}^T \right\rangle = \left\langle S\dot{V}, \dot{V} \right\rangle \geq 0.$$

Equation (C.2) is equivalent to the Hessian having exactly $p(p-1)/2$ zero eigenvalues: The fact that $\mathrm{Hess} f_C(V)$ has exactly $p(p-1)/2$ zero eigenvalues is equivalent to

$$\mathrm{Ker}(\mathrm{Hess} f_C(V)) = \{VA, A \in \mathrm{Anti}(p)\}.$$

(See Proposition 5 in Appendix C.10 for the justification of this equality.)

When V is a second-order critical point, in view of Equation (C.3) and since $S = C_2$, this last property is the same as the following equivalence being true:

$$\left(\left\langle C_2, \dot{V}\dot{V}^T \right\rangle = 0 \right) \quad \Longleftrightarrow \quad \left(\dot{V} \in \{VA, A \in \mathrm{Anti}(p)\} \right).$$

□

C.5 Proof of Lemma 7

Lemma (Lemma 7)**.** *Let us assume that $g_1 \in \mathbb{R}^m, C_2 \in \mathbb{S}^{n\times n}$ satisfy Conditions 1 and 2.*

Let us define a subspace $E_\perp$ of $\mathbb{R}^n$ as

$$E_\perp = (\mathrm{Range}(U_0) + \mathrm{Range}(V))^\perp,$$

and denote by $P_\perp \in \mathbb{S}^{n\times n}$ the matrix representing the orthogonal projection from $\mathbb{R}^n$ onto $E_\perp$ in the canonical basis.

For any $t \in \mathbb{R}^+$, we set

$$C_{2,t} = C_2 + tP_\perp.$$

Then, for all $t \in \mathbb{R}^+$ large enough, the pair $(g_1, C_{2,t})$ satisfies Conditions 1, 2 and 3.

Proof of Lemma 7. Let g_1, C_2 satisfy Conditions 1 and 2.

For any $t \in \mathbb{R}^+$, $g_1, C_{2,t}$ also satisfy Conditions 1 and 2: As $C_2 - \mathcal{A}^*(g_1)$ and $P_\perp$ are positive matrices,

$$C_{2,t} - \mathcal{A}^*(g_1) = C_2 - \mathcal{A}^*(g_1) + tP_\perp \succeq 0.$$

The rank of $C_{2,t} - \mathcal{A}^*(g_1)$ is at least as large as the rank of $C_2 - \mathcal{A}^*(g_1)$ (adding a positive matrix to another one cannot reduce the rank). Hence $\mathrm{rank}(C_{2,t} - \mathcal{A}^*(g_1)) \geq n - r$. From the definition of $P_\perp$, $P_\perp U_0 = 0$, so

$$(C_{2,t} - \mathcal{A}^*(g_1))U_0 = (C_2 - \mathcal{A}^*(g_1))U_0 \overset{(\text{Cond. } 1)}{=} 0.$$

As $\mathrm{rank}(U_0) = r$, we must have $\mathrm{rank}(C_{2,t} - \mathcal{A}^*(g_1)) \leq n - r$ so $\mathrm{rank}(C_{2,t} - \mathcal{A}^*(g_1)) = n - r$. Condition 1 holds.

Additionally, $C_{2,t}V = 0$, because $C_2 V = 0$ from Condition 2 and $P_\perp V = 0$ by construction. Condition 2 also holds.

Condition 3 is true when t is large enough: We recall that Condition 3 is

$$\forall \dot{V} \in T_V\mathcal{M}_p, \quad \left\langle C_{2,t}, \dot{V}\dot{V}^T \right\rangle \begin{array}{l} = 0 \text{ if } \dot{V} \in \{VA, A \in \text{Anti}(p)\}, \\ > 0 \text{ if } \dot{V} \notin \{VA, A \in \text{Anti}(p)\}. \end{array}$$

We first prove that this property is implied by the following one:

$$\forall \dot{V} \in \left(T_V\mathcal{M}_p \cap \{VA, A \in \text{Anti}(p)\}^\perp\right) - \{0_{n,p}\}, \quad \left\langle C_{2,t}, \dot{V}\dot{V}^T \right\rangle > 0. \tag{C.4}$$

Indeed, assuming Property (C.4), let us show that Condition 3 is satisfied. Let $\dot{V} \in T_V\mathcal{M}_p$ be arbitrary. We can uniquely write it as

$$\dot{V} = VA + \ddot{V},$$

with $A \in \text{Anti}(p)$ and $\ddot{V} \in \{VA, A \in \text{Anti}(p)\}^\perp$. As VA belongs to $T_V\mathcal{M}_p$ (see Proposition 5), $\ddot{V}$ also does, so from Property (C.4),

$$\left\langle C_{2,t}, \ddot{V}\ddot{V}^T \right\rangle \geq 0,$$

with equality if and only if $\ddot{V} = 0_{n,p}$.

Consequently,

$$\begin{aligned} \left\langle C_{2,t}, \dot{V}\dot{V}^T \right\rangle &= \left\langle C_{2,t}, (VA + 2\ddot{V})(VA)^T \right\rangle + \left\langle C_{2,t}, \ddot{V}\ddot{V} \right\rangle \\ &= \left\langle C_{2,t}V, (VA + 2\ddot{V})A^T \right\rangle + \left\langle C_{2,t}, \ddot{V}\ddot{V} \right\rangle \\ &\overset{\text{(Cond. 2)}}{=} \left\langle C_{2,t}, \ddot{V}\ddot{V} \right\rangle \\ &\geq 0, \end{aligned}$$

with equality if and only if $\ddot{V} = 0$, that is if and only if $\dot{V} = VA$. This is exactly Condition 3.

To conclude, we have to show that Property (C.4) is true for all t large enough. For any $t \geq 0$, we define

$$S_t = \{\dot{V} \in T_V\mathcal{M}_p \cap \{VA, A \in \text{Anti}(p)\}^\perp, ||\dot{V}||_F = 1, \left\langle C_{2,t}, \dot{V}\dot{V}^T \right\rangle \leq 0\},$$

and observe that Property (C.4) is true if and only if $S_t = \emptyset$.

The sets S_t are closed and bounded, hence compact. For any t_1, t_2 such that $t_1 \leq t_2$, $S_{t_2} \subset S_{t_1}$. Therefore, either

$$\bigcap_{t \geq 0} S_t \neq \emptyset, \tag{C.5}$$

or $S_t = \emptyset$ for all t large enough.

Let us show that Equation (C.5) does not hold. Let $\dot{V}$ be any unit-normed element of $T_V\mathcal{M}_p \cap \{VA, A \in \text{Anti}(p)\}^\perp$, and let us show that $\dot{V} \notin \cap_{t \geq 0} S_t$. We make the following observation (proved in Paragraph C.5.1).

Proposition 7. $\text{Range}(\dot{V}) \not\subset \text{Range}(U_0) + \text{Range}(V)$.

As a consequence,

$$\left\langle P_\perp, \dot{V}\dot{V}^T \right\rangle > 0.$$

Indeed, the scalar product of two positive matrices is always nonnegative. It could only be zero if we had $P_\perp\dot{V}\dot{V}^T = 0$, that is

$$\text{Range}(\dot{V}) \subset \text{Ker}(P_\perp) = \text{Range}(U_0) + \text{Range}(V),$$

which cannot be true from the proposition.

Therefore, for $t > 0$ large enough,

$$\begin{aligned} \left\langle C_{2,t}, \dot{V}\dot{V}^T \right\rangle &= \left\langle C_2, \dot{V}\dot{V}^T \right\rangle + t \left\langle P_\perp, \dot{V}\dot{V}^T \right\rangle \\ &> 0. \end{aligned}$$

This implies that $\dot{V} \notin \bigcap_{t \geq 0} S_t$. As $\dot{V}$ was arbitrary, it shows that Equation (C.5) is not true, hence $S_t = \emptyset$ for all t large enough, which established Property (C.4), and thus concludes the proof. □

C.5.1 Proof of Proposition 7

Proposition (Proposition 7)**.** $\mathrm{Range}(\dot{V}) \not\subset \mathrm{Range}(U_0) + \mathrm{Range}(V)$.

Proof of Proposition 7. We assume by contradiction that the inclusion holds. Then $\dot{V}$ can be written as $\left(V \; U_0 \right) \left(\begin{smallmatrix} R_1 \\ R_2 \end{smallmatrix} \right)$, for two matrices $R_1 \in \mathbb{R}^{p\times p}, R_2 \in \mathbb{R}^{r\times p}$. As $\dot{V}$ belongs to the tangent space of $\mathcal{M}_p$ at V,

$$\begin{aligned} 0 &= \mathcal{A}(V\dot{V}^T + \dot{V}V^T) \\ &= \mathcal{A}\left(V \left(\begin{smallmatrix} R_1 \\ R_2 \end{smallmatrix} \right)^T \left(V \; U_0 \right)^T + \left(V \; U_0 \right) \left(\begin{smallmatrix} R_1 \\ R_2 \end{smallmatrix} \right) V^T \right) \\ &= \psi_V \left(\frac{R_1 + R_1^T}{2}, R_2 \right). \end{aligned}$$

(where ψ_V is the mapping defined in Equation (2.2)).

Because $\mathcal{M}_p$ is X_0-minimally secant at V, we then have

$$\frac{R_1 + R_1^T}{2} = 0 \qquad \text{and} \qquad R_2 = 0.$$

Therefore, $\dot{V} = VR_1$, and $R_1 \in \mathrm{Anti}(p)$. This contradicts the fact that $\dot{V}$ is a non-zero element of $\{VA, A \in \mathrm{Anti}(p)\}^\perp$. □

C.6 Proof of Lemma 8

Lemma (Lemma 8)**.** *Let g_1 be an element of $\mathbb{R}^m$. If*

(a) $V^T\mathcal{A}^*(g_1)V \prec 0$,

(b) $U_0^T\mathcal{A}^*(g_1)V = 0_{r,p}$,

then there exists $C_2 \in \mathbb{S}^{m\times m}$ such that g_1, C_2 satisfy Conditions 1 and 2.

Proof of Lemma 8. Let $g_1 \in \mathbb{R}^m$ satisfy the two conditions. Let us show the existence of C_2 such that Conditions 1 and 2 hold.

The proof is in two parts:

- First, we show the existence of $C_2^{(0)} \in \mathbb{S}^{n\times n}$ such that

$$\left(C_2^{(0)} - \mathcal{A}^*(g_1) \right) U_0 = 0 \tag{C.6a}$$

$$\text{and } C_2^{(0)} V = 0. \tag{C.6b}$$

- Then, from $C_2^{(0)}$, we construct C_2 such that Properties (C.6a) and (C.6b) still hold, but in addition,

$$C_2 - \mathcal{A}^*(g_1) \text{ is a positive matrix with rank } n - r, \tag{C.7}$$

  hence g_1, C_2 satisfy Conditions 1 and 2.

First part: Properties (C.6a) and (C.6b) are equivalent to

$$C_2^{(0)} \begin{pmatrix} V & U_0 \end{pmatrix} = \begin{pmatrix} 0_{n,p} & \mathcal{A}^*(g_1)U_0 \end{pmatrix},$$

so the existence of $C_2^{(0)} \in \mathbb{S}^{n\times n}$ satisfying them is implied by the following proposition (proved in Paragraph C.6.1).

Proposition 8. *Let $H \in \mathbb{R}^{n\times(p+r)}$ be a matrix with rank $p + r$. Then*

$$\begin{aligned} &\mathrm{Range}\left(T \in \mathbb{S}^{n\times n} \to TH \in \mathbb{R}^{n\times(p+r)} \right) \\ &\quad = \left\{ J \in \mathbb{R}^{n\times(p+r)} \textit{ such that } H^TJ \textit{ is symmetric} \right\}. \end{aligned}$$

Indeed, the matrix $\left(V \; U_0 \right)$ has rank $p+r$ (because $\mathrm{rank}(U_0) = r$ and from Proposition 2), so we can apply the proposition with $H = \left(V \; U_0 \right)$. Observe that

$$\begin{pmatrix} V & U_0 \end{pmatrix}^T \begin{pmatrix} 0_{n,p} & \mathcal{A}^*(g_1)U_0 \end{pmatrix} = \begin{pmatrix} 0_{p,p} & V^T\mathcal{A}^*(g_1)U_0 \\ 0_{r,p} & U_0^T\mathcal{A}^*(g_1)U_0 \end{pmatrix}$$

is symmetric because of Condition (b). Consequently, $\left(0_{n,p} \; \mathcal{A}^*(g_1)U_0 \right)$ is in the range of the mapping $T \in \mathbb{S}^{n\times n} \to T \left(V \; U_0 \right)$, hence $C_2^{(0)}$ as desired exists.

Second part: Let $C_2^{(0)} \in \mathbb{S}^{n\times n}$ be fixed, satisfying Properties (C.6a) and (C.6b). We must construct C_2 satisfying Property (C.7), in addition to (C.6a) and (C.6b).

Let $P_\perp \in \mathbb{S}^{n\times n}$ be the representation in the canonical basis of $\mathbb{R}^n$ of the orthogonal projector onto

$$(\mathrm{Range}(U_0) + \mathrm{Range}(V))^\perp .$$

For any $t \geq 0$, we define

$$C_2^{(t)} = C_2^{(0)} + tP_\perp.$$

For any $t \geq 0$, since $P_\perp U_0 = 0$ and $P_\perp V = 0$, $C_2^{(t)}$ satisfies Properties (C.6a) and (C.6b). We show that, for t large enough, $C_2^{(t)}$ also satisfies Property (C.7).

To achieve this, it suffices to prove that, for t large enough,

$$\forall x \in \mathbb{R}^n, \quad x^* \left(C_2^{(t)} - \mathcal{A}^*(g_1) \right) x \geq 0, \text{ with equality iff } x \in \mathrm{Range}(U_0). \tag{C.8}$$

Indeed, if Property (C.8) is true, then $C_2^{(t)} - \mathcal{A}^*(g_1)$ is positive, and its rank is

$$n - \dim(\mathrm{Range}(U_0)) = n - r.$$

Let us establish Property (C.8), for all large t. For any $x \in \mathrm{Range}(U_0)$, because of Property (C.6a), whatever the value of t,

$$x^* \left(C_2^{(t)} - \mathcal{A}^*(g_1) \right) x = 0.$$

As a consequence, we only have to show that, when t is large, for all $x \notin \mathrm{Range}(U_0)$,

$$x^* \left(C_2^{(t)} - \mathcal{A}^*(g_1) \right) x > 0. \tag{C.9}$$

Let such a x be fixed. We can write it as

$$x = u_0 + v + P_\perp x,$$

with $u_0 \in \mathrm{Range}(U_0), v \in \mathrm{Range}(V)$, and $v, P_\perp x$ not both zero. For any $t \geq 0$,

$$\begin{aligned}
x^* & \left(C_2^{(t)} - \mathcal{A}^*(g_1) \right) x \\
&= (u_0 + v + P_\perp x)^* \left(C_2^{(t)} - \mathcal{A}^*(g_1) \right) (u_0 + v + P_\perp x) \\
&\overset{(1)}{=} (v + P_\perp x)^* \left(C_2^{(t)} - \mathcal{A}^*(g_1) \right) (v + P_\perp x) \\
&\overset{(2)}{=} (v + P_\perp x)^* \left(C_2^{(0)} - \mathcal{A}^*(g_1) \right) (v + P_\perp x) + t x^* P_\perp^3 x \\
&= (v + P_\perp x)^* \left(C_2^{(0)} - \mathcal{A}^*(g_1) \right) (v + P_\perp x) + t x^* P_\perp^2 x \\
&= v^* \left(C_2^{(0)} - \mathcal{A}^*(g_1) \right) v + 2x^* P_\perp \left(C_2^{(0)} - \mathcal{A}^*(g_1) \right) v \\
&\qquad + x^* P_\perp (C_2^{(0)} - \mathcal{A}^*(g_1) + tI_n) P_\perp x \\
&\overset{(3)}{=} -v^* \mathcal{A}^*(g_1) v - 2x^* P_\perp \mathcal{A}^*(g_1) v + x^* P_\perp (C_2^{(0)} - \mathcal{A}^*(g_1) + tI_n) P_\perp x \\
&\geq \lambda_{min}(-\mathcal{A}^*(g_1)) ||v||^2 - 2|||\mathcal{A}^*(g_1)||| \, ||v|| \, ||P_\perp x|| \\
&\qquad + \lambda_{min}(C_2^{(0)} - \mathcal{A}^*(g_1) + tI_n) ||P_\perp x||^2.
\end{aligned} \tag{C.10}$$

(Equality (1) is because of Property (C.6a), Equality (2) because $P_\perp v = 0$ by definition of $P_\perp$ and v, and Equality (3) because of Property (C.6b).)

From Condition (a), $\lambda_{min}(-\mathcal{A}^*(g_1)) > 0$. The expression obtained in Equation (C.10) is a second-degree polynomial in $||P_\perp x||$. Computing the corresponding discriminant, we see that, if

$$\lambda_{min}(C_2^{(0)} - \mathcal{A}^*(g_1) + tI_n) > \frac{|||\mathcal{A}^*(g_1)|||^2}{\lambda_{\min}(\mathcal{A}^*(g_1))}, \tag{C.11}$$

then (recalling that $||v||$ and $||P_\perp x||$ are not both zero) Expression (C.10) is necessary positive. Therefore, when Equation (C.11) is true, Equation (C.9) is also true, for any $x \notin \text{Range}(U_0)$.

Inequality (C.11) holds for all t large enough. This concludes the proof. □

C.6.1 Proof of Proposition 8

Proposition (Proposition 8)**.** *Let $H \in \mathbb{R}^{n\times(p+r)}$ be a matrix with rank $p + r$. Then*

$$\begin{aligned}\text{Range}&\left(T \in \mathbb{S}^{n\times n} \to TH \in \mathbb{R}^{n\times(p+r)}\right)\\ &= \left\{J \in \mathbb{R}^{n\times(p+r)} \textit{ such that } H^T J \textit{ is symmetric}\right\}.\end{aligned}$$

Proof of Proposition 8. The kernel of $(T \in \mathbb{S}^{n\times n} \to TH \in \mathbb{R}^{n\times(p+r)})$ is

$$\{T \in \mathbb{S}^{n\times n}, \text{Range}(H) \subset \text{Ker}(T)\},$$

which has dimension

$$\frac{(n - \text{rank}(H))(n - \text{rank}(H) + 1)}{2} = \frac{(n - (p + r))(n - (p + r) + 1)}{2}.$$

As a consequence, the range of $(T \in \mathbb{S}^{n\times n} \to TH \in \mathbb{R}^{n\times(p+r)})$ has dimension

$$\begin{aligned}\frac{n(n+1)}{2} &- \frac{(n - (p + r))(n - (p + r) + 1)}{2}\\ &= n(p + r) - \frac{(p + r)(p + r - 1)}{2}.\end{aligned}$$

As it is included in $\left\{J \in \mathbb{R}^{n\times(p+r)} \text{ such that } H^T J \text{ is symmetric}\right\}$, it suffices to show that

$$\begin{aligned}\dim&\left\{J \in \mathbb{R}^{n\times(p+r)} \text{ such that } H^T J \text{ is symmetric}\right\}\\ &= n(p + r) - \frac{(p + r)(p + r - 1)}{2}.\end{aligned}$$

We first assume that H has the form

$$H = \begin{pmatrix}\Lambda\\ 0_{n-(p+r),p+r}\end{pmatrix}, \tag{C.12}$$

for some invertible $\Lambda \in \mathbb{R}^{(p+r)\times(p+r)}$. In this case,

$$\begin{aligned}&\left\{J \in \mathbb{R}^{n\times(p+r)} \text{ such that } H^T J \text{ is symmetric}\right\}\\ &= \left\{\left(\begin{smallmatrix}\Lambda^{T-1}T_1\\ T_2\end{smallmatrix}\right), T_1 \in \mathbb{S}^{(p+r)\times(p+r)}, T_2 \in \mathbb{R}^{n-(p+r),p+r}\right\},\end{aligned}$$

and this latter space has the prescribed dimension.

When H cannot be written as in Equation (C.12), we can at least write H as

$$H = O_1\tilde{H},$$

with $O_1 \in O(n)$ and $\tilde{H}$ as in Equation (C.12). With these notations,

$$\begin{aligned}&\left\{J \in \mathbb{R}^{n\times(p+r)} \text{ such that } H^T J \text{ is symmetric}\right\}\\ &= \left\{O_1 J \text{ with } J \in \mathbb{R}^{n\times(p+r)} \text{ such that } \tilde{H}^T J \text{ is symmetric}\right\},\end{aligned}$$

and, because $J \to O_1 J$ is a bijection from $\mathbb{R}^{n\times(p+r)}$ to itself, the dimension is the same as in the case where H has the form (C.12). □

C.7 Proof of Proposition 3

Proposition (Proposition 3)**.** *For any $\epsilon > 0$, there exists $\zeta > 0$ such that, when $C' \in B(C, \zeta)$,*

- *Problem* (SDP) *(with cost matrix C') admits at least one minimizer;*
- *all minimizers of Problem* (SDP) *belong to the ball $B(X_0, \epsilon)$.*

Proof of Proposition 3. Let $\epsilon > 0$ be fixed.

For any matrix $C' \in \mathbb{S}^{n\times n}$, we denote by $\langle C', .\rangle$ the function $X \in \mathbb{S}^{n\times n} \to \langle C', X\rangle$ (that is, the cost function of Problem (SDP) with cost matrix C').

The proposition is equivalent to the following property: for any sequence $(C'_k)_{k\in\mathbb{N}}$ of cost matrices converging to C, $\langle C'_k, .\rangle$ admits at least one minimizer on $\mathcal{C}$, and all its minimizers belong to $B(X_0, \epsilon)$, as soon as k is large enough.

Let $(C'_k)_{k\in\mathbb{N}}$ be such a sequence. The desired conclusion is true if the following property holds:

$$\text{For all } k \text{ large enough,} \quad \forall X \in \mathcal{C} - B(X_0, \epsilon), \quad \langle C'_k, X_0\rangle < \langle C'_k, X\rangle\,. \tag{C.13}$$

Indeed, in this case, for any k large enough, no element of $\mathcal{C} - B(X_0, \epsilon)$ is a minimizer of $\langle C'_k, .\rangle$. Additionally, any minimizer of $\langle C'_k, .\rangle$ on $\mathcal{C} \cap \overline{B}(X_0, \epsilon)$ (this exists: a continuous function always admits a minimizer on a non-empty compact set) is a minimizer of $\langle C'_k, .\rangle$ on $\mathcal{C}$, so at least one minimizer exists.

Therefore, we only have to show Property (C.13). We assume, by contradiction, that it is not true. Then, up to replacing $(C'_k)_{k\in\mathbb{N}}$ by a subsequence, we can assume that, for any $k \in \mathbb{N}$,

$$\exists X'_k \in \mathcal{C} - B(X_0, \epsilon), \quad \langle C'_k, X_0\rangle \geq \langle C'_k, X'_k\rangle\,. \tag{C.14}$$

For any k, let X'_k be such a matrix.

Up to extracting a subsequence again, we can assume (by compactness), that

$$\left(\frac{X'_k - X_0}{||X'_k - X_0||}\right)_{k\in\mathbb{N}}$$

converges to some unit-normed limit $Z \in \mathbb{S}^{n\times n}$. From Equation (C.14) and because $(C'_k)_{k\in\mathbb{N}}$ converges to C,

$$\langle C, Z\rangle \leq 0.$$

Equivalently,

$$\langle C, X_0 + \epsilon Z\rangle \leq \langle C, X_0\rangle\,. \tag{C.15}$$

Observe that $X_0 + \epsilon Z$ belongs to $\mathcal{C}$: It is the limit of the sequence

$$\left(\left(1 - \frac{\epsilon}{||X'_k - X_0||}\right) X_0 + \frac{\epsilon}{||X'_k - X_0||} X'_k\right)_{k\in\mathbb{N}^*}.$$

Each element of this sequence belongs to $\mathcal{C}$ (X_0 and X'_k do, and $\mathcal{C}$ is convex), and $\mathcal{C}$ is closed, so the limit also belongs to $\mathcal{C}$. Consequently, Equation (C.15) contradicts the fact that X_0 is the unique minimizer of $\langle C, .\rangle$ on $\mathcal{C}$.

□

C.8 Proof of Proposition 4

Proposition (Proposition 4)**.** *Slater's condition is satisfied: the feasible set $\mathcal{C}$ of Problem* (SDP) *contains a matrix X such that $X \succ 0$.*

Proof of Proposition 4. We first present an observation, whose proof is in Paragraph C.8.1 and relies on the p-regularity of $(\mathcal{A}, b)$.

Proposition 9. *There exists no $g \in \mathbb{R}^m - \{0\}$ such that*

$$\mathcal{A}^*(g)X_0 = 0.$$

The principle of the proof is to show that, if Slater's condition is not satisfied, then a non-zero $g \in \mathbb{R}^m$ exists such that $\mathcal{A}^*(g)X_0 = 0$. We thus assume that $\mathcal{C}$ contains no positive matrix. In other words,

$$\{X \in \mathbb{S}^{n\times n}, \mathcal{A}(X) = b\} \cap \{X \in \mathbb{S}^{n\times n}, X \succ 0\} = \emptyset.$$

From a hyperplane separation theorem, there exists $M \in \mathbb{S}^{n\times n} - \{0_{n,n}\}, \mu \in \mathbb{R}$ such that

$$\forall X \in \{X \in \mathbb{S}^{n\times n}, X \succ 0\}, \quad \langle M, X\rangle > \mu \tag{C.16a}$$

$$\text{and } \forall X \in \{X \in \mathbb{S}^{n\times n}, \mathcal{A}(X) = b\}, \quad \langle M, X\rangle \leq \mu. \tag{C.16b}$$

Equation (C.16a) is equivalent to

$$M \succeq 0 \text{ and } \mu \leq 0.$$

Because $\{X \in \mathbb{S}^{n\times n}, \mathcal{A}(X) = b\} = X_0 + \mathrm{Ker}(\mathcal{A}) = X_0 + (\mathrm{Range}(\mathcal{A}^*))^{\perp}$, Equation (C.16b), on the other hand, is equivalent to

$$M \in \mathrm{Range}(\mathcal{A}^*) \text{ and } \langle M, X_0\rangle \leq \mu.$$

Combining the last two equations yields in particular

$$\langle M, X_0\rangle \leq \mu \leq 0.$$

As M and X_0 are semidefinite positive, this is only possible if $MX_0 = 0$. Denoting by $g \in \mathbb{R}^m$ a vector such that $M = \mathcal{A}^*(g)$, we have that

$$\mathcal{A}^*(g)X_0 = 0,$$

which enters in contradiction with Proposition 9.

□

C.8.1 Proof of Proposition 9

Proposition (Proposition 9)**.** *There exists no $g \in \mathbb{R}^m - \{0\}$ such that*

$$\mathcal{A}^*(g)X_0 = 0.$$

Proof of Proposition 9. We proceed by contradiction, and assume that such a g exists.

We write

$$X_0 = V_0V_0^T,$$

with $V_0 \in \mathbb{R}^{n\times p}$. This is possible, because $\mathrm{rank}(X_0) = r$ and $p \geq r$. The matrix V_0 belongs to $\mathcal{M}_p$.

As $\mathrm{Range}(X_0) = \mathrm{Range}(V_0)$, the condition $\mathcal{A}^*(g)X_0 = 0$ is equivalent to

$$\mathcal{A}^*(g)V_0 = 0.$$

For all $\dot{V} \in \mathbb{R}^{n\times p}$,

$$\begin{aligned}\left\langle \mathcal{A}(V_0\dot{V}^T + \dot{V}V_0^T), g\right\rangle &= \left\langle V_0\dot{V}^T + \dot{V}V_0^T, \mathcal{A}^*(g)\right\rangle \\ &= 2\left\langle \dot{V}, \mathcal{A}^*(g)V_0\right\rangle \\ &= 0.\end{aligned}$$

Hence, the linear map $\dot{V} \in \mathbb{R}^{n\times p} \to \mathcal{A}(V_0\dot{V}^T + \dot{V}V_0^T)$ is not surjective. This contradicts the assumption that $(\mathcal{A}, b)$ is p-regular. □

C.9 Proof of Lemma 9

Lemma (Lemma 9)**.** *When k goes to infinity,*

$$D_k \to C_1 \quad \textit{and} \quad h_k \to g_1.$$

Proof of Lemma 9. Because $C'_k = D_k + \mathcal{A}^*(h_k)$ goes to $C = C_1 + \mathcal{A}^*(g_1)$ when k goes to infinity, we have

$$D_k - C_1 + \mathcal{A}^*(h_k - g_1) \overset{k\to+\infty}{\to} 0. \tag{C.17}$$

In particular, if $h_k \overset{k\to+\infty}{\to} g_1$, then $D_k \overset{k\to+\infty}{\to} C_1$, so we only have to show that $(h_k)_{k\in\mathbb{N}}$ converges to g_1.

We proceed by contradiction, and assume that

$$h_k \overset{k\to+\infty}{\not\to} g_1.$$

The principle of the proof is to combine this hypothesis with a compactness argument to construct a non-zero $g \in \mathbb{R}^m$ such that

$$\mathcal{A}^*(g)X_0 = 0, \tag{C.18}$$

which is impossible from Proposition 9.

Up to replacing $(h_k)_{k\in\mathbb{N}}$ by a subsequence, we can assume that $(||h_k - g_1||)_{k\in\mathbb{N}}$ is lower bounded by a (strictly) positive constant, and, by compactness, that

$$\left(\frac{h_k - g_1}{||h_k - g_1||}\right)_{k\in\mathbb{N}}$$

converges to some non-zero limit. Let $g \in \mathbb{R}^m$ be this limit. We show that it satisfies Property (C.18).

From Equation (C.17),

$$\frac{C_1 - D_k}{||h_k - g_1||} \overset{k\to+\infty}{\to} \mathcal{A}^*(g). \tag{C.19}$$

From Proposition 3, $(X'_k)_{k\in\mathbb{N}}$ converges to X_0, so

$$C_1 X'_k \overset{k\to+\infty}{\to} C_1 X_0 = 0,$$

and because $(||h_k - g_1||)_{k\in\mathbb{N}}$ is bounded away from zero, this implies

$$\frac{C_1 X'_k}{||h_k - g_1||} \overset{k\to+\infty}{\to} 0.$$

Recalling that, from the definition of D_k, $D_k X'_k = 0$ for all k, Equation (C.19) yields:

$$\mathcal{A}^*(g)X_0 = \lim_{k\to+\infty}\left(\frac{C_1 - D_k}{||h_k - g_1||}\right) X'_k = 0.$$

This is exactly Equation (C.18). $\square$

C.10 Proof of Proposition 5

Proposition (Proposition 5)**.** *For any cost matrix $C' \in \mathbb{S}^{n\times n}$, for any $W \in \mathcal{M}_p$,*

$$T_W\mathcal{M}_p \supset \{WA, A \in \mathrm{Anti}(p)\},$$

and, denoting $f_{C'}$ the cost function of Problem (Factorized SDP),

$$\mathrm{grad} f_{C'}(W) \in \{WA, A \in \mathrm{Anti}(p)\}^{\perp}.$$

Additionally, if W is a first-order critical point of $f_{C'}$,

$$\{WA, A \in \mathrm{Anti}(p)\} \subset \mathrm{Ker}(\mathrm{Hess} f_{C'}(W)).$$

When $\mathrm{rank}(W) = p$, this inclusion is an equality if and only if $\mathrm{Hess} f_{C'}(W)$ has exactly $\frac{p(p-1)}{2}$ zero eigenvalues.

Proof of Proposition 5. The first relation is because, if W belongs to $\mathcal{M}_p$, then

$$\{WB, B \in O(p)\} \subset \mathcal{M}_p,$$

so since $T_{I_p}O(p) = \mathrm{Anti}(p)$,

$$\{WA, A \in \mathrm{Anti}(p)\} \subset T_W\mathcal{M}_p. \tag{C.20}$$

The second one is because, for any C', W, and any $B \in O(p)$,

$$f_{C'}(WB) = f_{C'}(W).$$

If we differentiate this equality in B at $B = I_p$, we get the following:

$$\forall A \in \mathrm{Anti}(p), \quad \langle \mathrm{grad} f_{C'}(W), WA \rangle = 0.$$

For the last relation, we differentiate in W the following equality, for any $A \in \mathrm{Anti}(p)$:

$$\langle \mathrm{grad} f_{C'}(W), WA \rangle = 0.$$

It yields, for any $\dot{W} \in T_W \mathcal{M}_p$, $A \in \mathrm{Anti}(p)$,

$$\left\langle \mathrm{Hess} f_{C'}(W) \cdot \dot{W}, WA \right\rangle + \left\langle \mathrm{grad} f_{C'}(W), \dot{W} A \right\rangle = 0.$$

When W is first-order critical (that is $\mathrm{grad} f_{C'}(W) = 0$), it means that

$$\mathrm{Range}(\mathrm{Hess} f_{C'}(W)) \subset \{WA, A \in \mathrm{Anti}(p)\}^{\perp}. \tag{C.21}$$

As $\mathrm{Hess} f_{C'}(W)$ is self-adjoint on $T_W \mathcal{M}_p$,

$$\begin{aligned}
\mathrm{Ker}(\mathrm{Hess} f_{C'}(W)) &= T_W \mathcal{M}_p \cap (\mathrm{Range}(\mathrm{Hess} f_{C'}(W)))^{\perp} \\
&\overset{\text{Eq (C.21)}}{\supset} T_W \mathcal{M}_p \cap \{WA, A \in \mathrm{Anti}(p)\} \\
&\overset{\text{Eq (C.20)}}{=} \{WA, A \in \mathrm{Anti}(p)\}.
\end{aligned}$$

Finally, when $\mathrm{rank}(W) = p$,

$$\dim\{WA, A \in \mathrm{Anti}(p)\} = \dim(\mathrm{Anti}(p)) = \frac{p(p-1)}{2},$$

so the inclusion is an equality if and only if

$$\dim\left(\mathrm{Ker}(\mathrm{Hess} f_{C'}(W))\right) = \frac{p(p-1)}{2},$$

or, equivalently, if and only if $\mathrm{Hess} f_{C'}(W)$ has exactly $\frac{p(p-1)}{2}$ zero eigenvalues. □

D Application of Theorems 1 and 2 to examples

D.1 Proof of Remark 4

Remark (Remark 4)**.** *Let $p \in \mathbb{N}^*$ be fixed. We assume that $(\mathcal{A}, b)$ is p-regular and $\mathcal{M}_p$ is connected and non-empty. Then the following two statements are equivalent:*

1. *There exists $V \in \mathcal{M}_p$ such that ϕ_V is injective.*
2. *For almost any $V \in \mathcal{M}_p$, ϕ_V is injective (that is, $\mathcal{M}_p$ is face regular).*

Additionally, if X_0 is a rank r element of $\mathcal{C}$, there is also an equivalence between the following two properties:

1. *There exists $V \in \mathcal{M}_p$ such that ψ_V is injective (that is, there exists V at which $\mathcal{M}_p$ is X_0-minimally secant).*
2. *For almost any $V \in \mathcal{M}_p$, ψ_V is injective.*

Proof of Remark 4. We prove only the equivalence for the first pair of statements; the proof is identical for the second pair.

Since the second property clearly implies the first one, we assume that the first one holds, and show that it implies the second one. Let $V_0 \in \mathcal{M}_p$ be such that ϕ_{V_0} is injective.

Let $\mathcal{M}_p^{inj} \subset \mathcal{M}_p$ be the set of all V's for which ϕ_V is injective. We have to show that

$$\lambda_{\mathcal{M}_p}\left(\mathcal{M}_p \setminus \mathcal{M}_p^{inj}\right) = 0.$$

We define

$$\mathcal{E} = \{V \in \mathcal{M}_p, \exists \epsilon > 0, \lambda_{\mathcal{M}_p}((\mathcal{M}_p \setminus \mathcal{M}_p^{inj}) \cap B(V, \epsilon)) = 0\}.$$

This is an open subset of $\mathcal{M}_p$. We observe that V_0 belongs to $\mathcal{E}$: as ϕ_{V_0} is injective and $V \to \phi_V$ is continuous, ϕ_V is injective for V close enough to V_0, meaning that $(\mathcal{M}_p \setminus \mathcal{M}_p^{inj}) \cap B(V_0, \epsilon) = \emptyset$ for $\epsilon > 0$ small enough. In particular, $\mathcal{E}$ is non-empty.

We now show that $\mathcal{E}$ is closed. Let V_1 belong to the closure of $\mathcal{E}$, and let us show that $V_1 \in \mathcal{E}$.

Since $\mathcal{M}_p$ is the zero set of the analytic map $V \in \mathbb{R}^{n\times p} \to \mathcal{A}(VV^T) - b \in \mathbb{R}^m$, whose differential at any point of $\mathcal{M}_p$ is surjective ($(\mathcal{A}, b)$ is p-regular), $\mathcal{M}_p$ is an analytic submanifold of $\mathbb{R}^{n\times p}$.

For any d_1, d_2, the set of non-injective matrices in $\mathbb{R}^{d_1\times d_2}$ is a real algebraic manifold (that is, the set of common zeros of a finite set of polynomials). Since the matrix representing ϕ_V in any fixed basis of $\mathbb{S}^{p\times p}$ has polynomial coordinates in V, the set of all V's in $\mathbb{R}^{n\times p}$ for which this map is non-injective is also a real algebraic manifold in $\mathbb{R}^{n\times p}$.

Consequently, around V_1, $\mathcal{M}_p \setminus \mathcal{M}_p^{inj}$ is the intersection of the analytic submanifold $\mathcal{M}_p$ with the zero set of a finite number of polynomials. Two cases can then happen:

- First case: $\mathcal{M}_p \setminus \mathcal{M}_p^{inj}$ coincides with $\mathcal{M}_p$ around V_1.
- Second case: The intersection of $\mathcal{M}_p \setminus \mathcal{M}_p^{inj}$ with some neighborhood of V_1 in $\mathcal{M}_p$ has measure zero in $\mathcal{M}_p$.

Because V_1 is in the closure of $\mathcal{E}$, the first case can be excluded. We are then in the second case, which means

$$V_1 \in \mathcal{E}.$$

We have now seen that $\mathcal{E}$ is non-empty, open and closed in $\mathcal{M}_p$. Because $\mathcal{M}_p$ is connected,

$$\mathcal{E} = \mathcal{M}_p.$$

As a consequence, $\mathcal{M}_p - \mathcal{M}_p^{inj}$ has measure zero in the neighborhood of any point of $\mathcal{M}_p$, implying that

$$\lambda_{\mathcal{M}_p}(\mathcal{M}_p \setminus \mathcal{M}_p^{inj}) = 0.$$

□

D.2 Proof of Remark 5

Remark (Remark 5)**.** *Let $(d_s)_{s\in\mathbb{N}}$ be a sequence of positive integers. We assume that*

- *for any s, $\mathcal{C}^{(d_s)}$ has an extreme point with rank r;*
- *for any S and $p \geq r$, $(\mathcal{A}^{(d_1,\dots,d_S)}, b_S)$ is p-regular;*
- *for any S and $p \geq 2r$, $\mathcal{M}_p^{(d_1,\dots,d_S)}$ is connected and non-empty.*

Then the first property below implies the second one:

1. *For any $S \in \mathbb{N}^*$ and $p \geq r$ such that*

$$\frac{p(p+1)}{2} + pr \leq S\frac{r(r+1)}{2},$$

there exists X_0 an extreme point of $\mathcal{C}^{(d_1,\dots,d_S)}$ with rank r, and a matrix $V \in \mathcal{M}_p^{(d_1,\dots,d_S)}$ such that $\mathcal{M}_p^{(d_1,\dots,d_S)}$ is X_0-minimally secant at V.

$\Rightarrow$ 2. *For any $S \in \mathbb{N}^*$ and $p \geq 2r$ such that*

$$\frac{p(p+1)}{2} \leq S\frac{r(r+1)}{2},$$

$\mathcal{M}_p^{(d_1,\dots,d_S)}$ is face regular.

Proof of Remark 5. Let us assume that the first property holds, and show that the second one also does.

Let S and $p \geq 2r$ be such that

$$\frac{p(p+1)}{2} \leq S\frac{r(r+1)}{2}, \tag{D.1}$$

and let us show that $\mathcal{M}_p^{(d_1,\dots,d_S)}$ is face regular. From Remark 4, it is enough to prove that ϕ_W is injective for at least one $W \in \mathcal{M}_p^{(d_1,\dots,d_S)}$. We set

$$p' = p - r \geq r \text{ and } S' = S - 1.$$

From Equation (D.1),

$$\frac{p'(p'+1)}{2} + p'r = \frac{p(p+1)}{2} - \frac{(r+1)r}{2} \leq S'\frac{r(r+1)}{2}.$$

From the first property, there exists X_0 an extreme point of $\mathcal{C}^{(d_1,\dots,d_{S'})}$ with rank r, and a matrix $V \in \mathcal{M}_{p'}^{(d_1,\dots,d_{S'})}$ such that $\mathcal{M}_{p'}^{(d_1,\dots,d_{S'})}$ is X_0-minimally secant at V. Let such X_0, V be fixed. From these two matrices, we now construct $W \in \mathcal{M}_p^{(d_1,\dots,d_S)}$ such that ϕ_W is injective.

Let $U_0 \in \mathbb{R}^{(d_1+\dots+d_{S'})\times r}$ be such that

$$X_0 = U_0U_0^T.$$

Let also $u_0 \in \mathbb{R}^{d_S\times r}$ be such that $u_0u_0^T$ is an extreme point of $\mathcal{C}^{(d_S)}$ with rank r (from our hypotheses, it exists).

We set

$$W = \begin{pmatrix} V/\sqrt{2} & U_0/\sqrt{2} \\ 0_{d_S\times p'} & u_0 \end{pmatrix} \in \mathbb{R}^{n\times p}.$$

This matrix belongs to $\mathcal{M}_p^{(d_1,\dots,d_S)}$: For any $s \leq S-1$, because VV^T and $X_0 = U_0U_0^T$ belong to $\mathcal{C}^{(d_1,\dots,d_{S-1})}$,

$$\tilde{\mathcal{A}}_{d_s}(\mathrm{Block}_s(WW^T)) = \frac{1}{2}\left(\tilde{\mathcal{A}}_{d_s}(\mathrm{Block}_s(VV^T)) + \tilde{\mathcal{A}}_{d_s}(\mathrm{Block}_s(U_0U_0^T))\right) = \tilde{b},$$

and we also have, since $u_0u_0^T$ belongs to $\mathcal{C}^{(d_S)}$,

$$\tilde{\mathcal{A}}_{d_S}(\mathrm{Block}_S(WW^T)) = \tilde{\mathcal{A}}_{d_S}(u_0u_0^T) = \tilde{b}.$$

Let us show that ϕ_W is injective. For any $T \in \mathbb{S}^{p\times p}$, if we decompose T as

$$T = \begin{pmatrix} T_1 & T_2^T \\ T_2 & T_3 \end{pmatrix}, \text{ with } T_1 \in \mathbb{S}^{(p-r)\times(p-r)}, T_2 \in \mathbb{R}^{r\times(p-r)}, T_3 \in \mathbb{S}^{r\times r},$$

we have the following relation:

$$\begin{aligned}
\phi_W(T) &= \mathcal{A}_{(d_1,\dots,d_S)}(WTW^T) \\
&= \Bigg(\frac{1}{2}\mathcal{A}_{(d_1,\dots,d_{S-1})}(VT_1V^T + VT_2^TU_0^T + U_0T_2V^T \\
&\qquad\qquad + U_0T_3U_0^T), \tilde{\mathcal{A}}_{d_S}(u_0T_3u_0^T)\Bigg) \\
&= \Bigg(\frac{1}{2}\mathcal{A}_{(d_1,\dots,d_{S-1})}\Big(\begin{pmatrix} V & U_0 \end{pmatrix}\begin{pmatrix} T_1/2 \\ T_2 \end{pmatrix}V^T + V\begin{pmatrix} T_1/2 \\ T_2 \end{pmatrix}^T\begin{pmatrix} V & U_0 \end{pmatrix}^T \\
&\qquad\qquad + U_0T_3U_0^T\Big), \tilde{\mathcal{A}}_{d_S}(u_0T_3u_0^T)\Bigg) \\
&= \left(\frac{1}{2}\left(\psi_V(T_1/2, T_2) + \mathcal{A}_{(d_1,\dots,d_{S-1})}\left(U_0T_3U_0^T\right)\right), \tilde{\mathcal{A}}_{d_S}(u_0T_3u_0^T)\right).
\end{aligned}$$

Let us assume that $\phi_W(T) = 0$, and show that $T = 0$. First, from the previous expression of $\phi_W(T)$, we see that $\tilde{\mathcal{A}}_{d_S}(u_0T_3u_0^T) = 0$. This implies that $u_0T_3u_0^T = 0$ (otherwise, $u_0u_0^T$ is not an extreme point of $\mathcal{C}^{(d_S)}$) and, as a consequence, that

$$T_3 = 0$$

(because $\mathrm{rank}(u_0) = r$). Therefore, $\psi_V(T_1/2, T_2) = 0$. Since ψ_V is injective ($\mathcal{M}_{p'}^{(d_1,\dots,d_{S'})}$ is X_0-minimally secant at V),

$$T_1 = 0 \text{ and } T_2 = 0.$$

We have thus shown that all subblocks of T are zero, hence $T = 0$. □

D.3 Proof of Corollary 2

Corollary (Corollary 2)**.** *Let us assume that $d = 1, 2$ or 3.*

If $p \geq 2d$ is such that

$$\frac{p(p+1)}{2} + p > \frac{Sd(d+1)}{2},$$

then, for almost any cost matrix C, all second-order critical points of the Burer-Monteiro factorization of Problem (SDP-Orthogonal-Cut) are globally optimal.

On the other hand, for any $p \geq d$ such that

$$\frac{p(p+1)}{2} + pd \leq \frac{Sd(d+1)}{2},$$

the set of cost matrices admits a subset with non-zero Lebesgue measure on which

- *Problem (SDP-Orthogonal-Cut) has a unique global optimum, which has rank d;*
- *Its Burer-Monteiro factorization with rank p has at least one non-optimal second-order critical point.*

Proof of Corollary 2. Let $d \in \{1, 2, 3\}$ be fixed.

We note that $(\mathcal{A}, b)$ is p-regular for any $p \geq d$, and $\mathcal{M}_p$ is connected, unless $p = d$.

First part of the corollary: It is a direct consequence of Theorem 1; we simply have to check that the three hypotheses of this theorem hold, for any p, S satisfying the required inequalities. The feasible set $\mathcal{C}$ is closed and bounded in $\mathbb{S}^{n\times n}$, hence compact, and we have already seen that $(\mathcal{A}, b)$ is p-regular for any $p \geq d$, so the first two hypotheses hold true.

The third one ($\mathcal{M}_p$ is face regular if $\frac{p(p+1)}{2} \leq S\frac{d(d+1)}{2}$) is more delicate. However, from Remark 5, it is implied by the following property: For any $S \in \mathbb{N}^*, p \geq d$ such that

$$\frac{p(p+1)}{2} + pd \leq S\frac{d(d+1)}{2},$$

there exists X_0 an extreme point of $\mathcal{C}$ with rank d, and $V \in \mathcal{M}_p$ such that $\mathcal{M}_p$ is X_0-minimally secant at V.

We will establish this latter property while proving the second part of the corollary, so this concludes the proof of the first part.

Second part of the corollary: We deduce it from Theorem 2; we simply have to check that the hypotheses of this theorem hold true.

Let S and $p \geq d$ be fixed, satisfying the inequality

$$\frac{p(p+1)}{2} + pd \leq \frac{Sd(d+1)}{2}.$$

We have already seen that $(\mathcal{A}, b)$ is p-regular, which is the second hypothesis.

We set

$$U_0 = \begin{pmatrix} I_d \\ \vdots \\ I_d \end{pmatrix} \in \mathbb{R}^{n\times d}$$

and $X_0 = U_0 U_0^T$. Then X_0 is an extreme point of $\mathcal{C}$ with rank d, so the first hypothesis holds.

Let us now check the third hypothesis: We have to find V such that $\mathcal{M}_p$ is X_0-minimally secant at V. We recall that this condition is equivalent to the map ψ_V in Equation (2.2) to be injective. It does not seem obvious to directly establish the injectivity of ψ_V, for a given V, but this condition happens to be implied by another one, easier to manipulate, as stated in the following proposition, whose proof is in Paragraph D.3.1. (The two conditions are actually equivalent, but we only need an implication.)

Proposition 10. *For any $V \in \mathcal{M}_p$ and $k \le S$, we denote by V_k the k-th $d \times p$ block of V. If*

$$\left\{ \begin{pmatrix} 0_{p\times d} \\ I_d \end{pmatrix} A_0 \begin{pmatrix} 0_{d\times p} & I_d \end{pmatrix} + \sum_{k=1}^{S} \begin{pmatrix} V_k^T \\ I_d \end{pmatrix} A_k \begin{pmatrix} V_k & I_d \end{pmatrix}, A_0, \dots, A_S \in \mathbb{S}^{d\times d} \right\} \\ = \mathbb{S}^{(p+d)\times(p+d)}, \tag{D.2}$$

then ψ_V is injective.

In view of this proposition, we simply have to find matrices $V_1, \dots, V_S \in \mathbb{R}^{d\times p}$ for which Equality (D.2) is true, and such that

$$\forall k \le S, \quad V_k V_k^T = I_d.$$

Indeed, setting $V = \begin{pmatrix} V_1 \\ \vdots \\ V_S \end{pmatrix}$ then produces an element of $\mathcal{M}_p$ for which ψ_V is injective.

We did not find a simple construction for $V_1, \dots, V_S$ that would not depend on d. Hence, we present separate constructions for the cases $d = 1$, $d = 2$ and $d = 3$.

Case $d = 1$: Let us denote $e_1, \dots, e_p$ the elements of the canonical basis of $\mathbb{R}^{1\times p}$, and define $V_1, \dots, V_{\frac{p(p+1)}{2}+p}$ such that

$$\left\{V_1, \dots, V_{\frac{p(p+1)}{2}+p}\right\} = \{e_i, i \le p\} \cup \{-e_i, i \le p\} \cup \left\{ \frac{e_i + e_j}{\sqrt{2}}, i < j \le p \right\}.$$

For $\frac{p(p+1)}{2} + p < s \le S$, we set V_s arbitrarily, in such a way that $V_s V_s^T = 1$.

Let us show that Condition (D.2) is satisfied. If we denote $e'_1, \dots, e'_{p+1}$ the canonical basis of $\mathbb{R}^{1\times(p+1)}$, we have

- For all $i \le p$, $e_i'^T e_i' = -\begin{pmatrix} 0_{p\times 1} \\ 1 \end{pmatrix} \begin{pmatrix} 0_{1\times p} & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} e_i^T \\ 1 \end{pmatrix} \begin{pmatrix} e_i & 1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} -e_i^T \\ 1 \end{pmatrix} \begin{pmatrix} -e_i & 1 \end{pmatrix}$.
- $e_{p+1}'^T e_{p+1}' = \begin{pmatrix} 0_{p\times 1} \\ 1 \end{pmatrix} \begin{pmatrix} 0_{1\times p} & 1 \end{pmatrix}$.
- For all $i \le p$, $e_i'^T e_{p+1}' + e_{p+1}'^T e_i' = \frac{1}{2} \begin{pmatrix} e_i^T \\ 1 \end{pmatrix} \begin{pmatrix} e_i & 1 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} -e_i^T \\ 1 \end{pmatrix} \begin{pmatrix} -e_i & 1 \end{pmatrix}$.
- For all $i < j \le p$,

$$\begin{aligned} e_i'^T e_j' + e_j'^T e_i' = {} & 2 \begin{pmatrix} \frac{e_i^T + e_j^T}{\sqrt{2}} \\ 1 \end{pmatrix} \begin{pmatrix} \frac{e_i + e_j}{\sqrt{2}} & 1 \end{pmatrix} \\ & - e_i'^T e_i' - e_j'^T e_j' - 2 e_{p+1}'^T e_{p+1}' \\ & - \sqrt{2} \left(e_i'^T e_{p+1}' + e_{p+1}'^T e_i' \right) - \sqrt{2} \left(e_j'^T e_{p+1}' + e_{p+1}'^T e_j' \right). \end{aligned}$$

This implies that

$$\mathrm{Span} \left(\left\{ \begin{pmatrix} 0_{p\times 1} \\ 1 \end{pmatrix} \begin{pmatrix} 0_{1\times p} & 1 \end{pmatrix} \right\} \cup \left\{ \begin{pmatrix} V_k^T \\ 1 \end{pmatrix} \begin{pmatrix} V_k & 1 \end{pmatrix} \right\}_{k=1,\dots,S} \right)$$

contains $\{e_i'^T e_i'\}_{i \le p+1} \cup \{e_i'^T e_j' + e_j'^T e_i'\}_{i \ne j \le p+2}$, which is a basis of $\mathbb{S}^{(p+1)\times(p+1)}$. Consequently., Condition (D.2) holds.

Case d=2: We will use the following blocks in our construction:

$$G_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad G_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad G_3 = \begin{pmatrix} 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix},$$

$$G_4 = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix}.$$

We distinguish depending on the congruency of p modulo 3.

If $p \equiv 0[3]$, for any $q = 1, \dots, p/3$, we set

$$\begin{aligned} W_q^{(1)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_1 & 0_{2\times(p-3q)} \end{pmatrix}, & W_q^{(2)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_2 & 0_{2\times(p-3q)} \end{pmatrix}, \\ W_q^{(3)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_3 & 0_{2\times(p-3q)} \end{pmatrix}, & W_q^{(4)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_4 & 0_{2\times(p-3q)} \end{pmatrix}. \end{aligned}$$

For any $q, q' \in \{1, \dots, p/3\}$ such that $q < q'$, we set

$$\begin{aligned}
X_{q,q'}^{(1)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_1 & 0_{2\times 3(q'-q-1)} & G_1 & 0_{2\times(p-3q')} \end{pmatrix} / \sqrt{2}, \\
X_{q,q'}^{(2)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_2 & 0_{2\times 3(q'-q-1)} & G_3 & 0_{2\times(p-3q')} \end{pmatrix} / \sqrt{2}, \\
X_{q,q'}^{(3)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_4 & 0_{2\times 3(q'-q-1)} & G_2 & 0_{2\times(p-3q')} \end{pmatrix} / \sqrt{2}.
\end{aligned}$$

Now we choose $V_1, \dots, V_S$ so that $\{V_1, \dots, V_S\}$ contains all the $W_q^{(i)}$ and all the $X_{q,q'}^{(i)}$. It is possible because there are $\frac{p^2+5p}{6}$ such matrices and our assumption on S is precisely that $S \geq \frac{p^2+5p}{6}$.

With this choice, it is tedious but not difficult to check that any matrix $M \in \mathbb{S}^{(p+2)\times(p+2)}$ that belongs to the orthogonal of the vector space on the left-hand side of Condition (D.2) is zero, implying that Condition (D.2) holds. Indeed, if one divides M into $\left(\frac{p}{3}+1\right)^2$ blocks as follows,

$$\begin{array}{cc}
 & \begin{array}{cccc} 3 & & 3 & 2 \end{array} \\
\begin{array}{c} 3\,\{ \\ \\ 3\,\{ \\ 2\,\{ \end{array} & \begin{pmatrix} M_{1,1} & \cdots & M_{1,\frac{p}{3}} & M_{1,\frac{p}{3}+1} \\ \vdots & & \vdots & \vdots \\ M_{\frac{p}{3},1} & \cdots & M_{\frac{p}{3},\frac{p}{3}} & M_{\frac{p}{3},\frac{p}{3}+1} \\ M_{\frac{p}{3}+1,1} & \cdots & M_{\frac{p}{3}+1,\frac{p}{3}} & M_{\frac{p}{3}+1,\frac{p}{3}+1} \end{pmatrix},
\end{array}$$

one immediately sees that $M_{\frac{p}{3}+1,\frac{p}{3}+1} = 0_{2\times 2}$, because of the orthogonality to $\begin{pmatrix} 0_{p\times 2} \\ I_2 \end{pmatrix} A_0 \begin{pmatrix} 0_{2\times p} & I_2 \end{pmatrix}$ for any $A_0 \in \mathbb{S}^{2\times 2}$. Then, for any $q \leq \frac{p}{3}$, using the fact that $W_q^{(1)}, W_q^{(2)}, W_q^{(3)}, W_q^{(4)}$ all belong to $\{V_1, \dots, V_S\}$, hence $\begin{pmatrix} W_q^{(i)T} \\ I_d \end{pmatrix} S \begin{pmatrix} W_q^{(i)} & I_d \end{pmatrix}$ belongs to the set on the left-hand side of Equation (D.2) for any $i \leq 4, S \in \mathbb{S}^{2\times 2}$, one can show that $M_{q,q} = 0_{3\times 3}, M_{\frac{p}{3}+1,q} = 0_{2\times 3}$ and $M_{q,\frac{p}{3}+1} = 0_{3\times 2}$. Finally, for any q, q' such that $q < q' \leq \frac{p}{3}$, using the fact that $X_{q,q'}^{(1)}, X_{q,q'}^{(2)}, X_{q,q'}^{(3)}$ belong to $\{V_1, \dots, V_S\}$, one can prove that $M_{q,q'} = M_{q',q} = 0_{3\times 3}$.

In the case where $p \equiv 1[3]$, we define $W_q^{(1)}, W_q^{(2)}, W_q^{(3)}, W_q^{(4)}$ and $X_{q,q'}^{(1)}, X_{q,q'}^{(2)}, X_{q,q'}^{(3)}$ as previously, for any $q \leq \frac{p-4}{3}$, and any $q < q' \leq \frac{p-4}{3}$. For $q = \frac{p-1}{3}$, we define six matrices $(W_q^{(i)})_{i=1,\dots,6}$ by

$$W_q^{(i)} = \begin{pmatrix} 0_{2\times p-4} & H_i \end{pmatrix},$$

with

$$H_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad H_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad H_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$
$$H_4 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad H_5 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \qquad H_6 = \begin{pmatrix} 0 & 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{3}{5} & \frac{4}{5} & 0 & 0 \end{pmatrix}.$$

And for $q \leq \frac{p-4}{3}, q' = \frac{p-1}{3}$, we define the following four matrices:

$$\begin{aligned}
X_{q,q'}^{(1)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_1 & 0_{2\times p-3q-4} & H_1 \end{pmatrix} / \sqrt{2}, \\
X_{q,q'}^{(2)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_1 & 0_{2\times p-3q-4} & H_3 \end{pmatrix} / \sqrt{2}, \\
X_{q,q'}^{(3)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_2 & 0_{2\times p-3q-4} & H_1 \end{pmatrix} / \sqrt{2}, \\
X_{q,q'}^{(4)} &= \begin{pmatrix} 0_{2\times 3(q-1)} & G_2 & 0_{2\times p-3q-4} & H_3 \end{pmatrix} / \sqrt{2}.
\end{aligned}$$

We choose $V_1, \dots, V_S$ so that $\{V_1, \dots, V_S\}$ contains all $W_q^{(i)}$ and all $X_{q,q'}^{(i)}$. This is possible because, again, $\{W_q^{(i)}, q, i\} \cup \{X_{q,q'}^{(i)}, q, q', i\}$ has cardinality $\frac{p^2+5p}{6} \leq S$. We conclude as previously, the only difference being that we have to split M into $\left(\frac{p-1}{3}+1\right)^2$ blocks, as follows:

$$\begin{array}{cc}
 & \begin{array}{cccc} 3 & & 4 & 2 \end{array} \\
\begin{array}{c} 3\,\{ \\ \\ 4\,\{ \\ 2\,\{ \end{array} & \begin{pmatrix} M_{1,1} & \cdots & M_{1,\frac{p-1}{3}} & M_{1,\frac{p-1}{3}+1} \\ \vdots & & \vdots & \vdots \\ M_{\frac{p-1}{3},1} & \cdots & M_{\frac{p-1}{3},\frac{p-1}{3}} & M_{\frac{p-1}{3},\frac{p-1}{3}+1} \\ M_{\frac{p-1}{3}+1,1} & \cdots & M_{\frac{p-1}{3}+1,\frac{p-1}{3}} & M_{\frac{p-1}{3}+1,\frac{p-1}{3}+1} \end{pmatrix},
\end{array}$$

Finally, in the case where $p \equiv 2[3]$, we define $W_q^{(1)}, W_q^{(2)}, W_q^{(3)}, W_q^{(4)}$ and $X_{q,q'}^{(1)}, X_{q,q'}^{(2)}, X_{q,q'}^{(3)}$ as before for $q \le \frac{p-2}{3}$ and $q < q' \le \frac{p-2}{3}$. For $q = \frac{p+1}{3}$, we define only three matrices $(W_q^{(i)})_{i=1,2,3}$:

$$W_q^{(i)} = \begin{pmatrix} 0_{2\times(p-2)} & J_i \end{pmatrix},$$

with

$$J_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad J_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad J_3 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

For $q \le \frac{p-2}{3}$, $q' = \frac{p+1}{3}$, we set

$$X_{q,q'}^{(1)} = \begin{pmatrix} 0_{2\times 3(q-1)} & G_1 & 0_{2\times p-3q-2} & J_1 \end{pmatrix}/\sqrt{2},$$
$$X_{q,q'}^{(2)} = \begin{pmatrix} 0_{2\times 3(q-1)} & G_2 & 0_{2\times p-3q-2} & J_1 \end{pmatrix}/\sqrt{2}.$$

We conclude as before.

Case $d = 3$: This case can be dealt with in the same way as $d = 2$, but is even more technical. The easiest thing to do (although maybe not the most elegant one) is to distinguish 12 cases, depending on the congruency of p modulo 12. To avoid pages of definitions, we only focus on the case where $p \equiv 0[12]$ and, even in this case, only provide a sketch of proof.

For any $q \le \frac{p}{12}$, we define 19 matrices $(W_q^{(i)})_{i\le 19}$ of size $3 \times p$, by

$$W_q^{(i)} = \begin{pmatrix} 0_{3\times 12(q-1)} & G_i & 0_{3\times(p-12q)} \end{pmatrix},$$

for matrices $G_1, \dots, G_{19} \in \mathbb{R}^{3\times 12}$ suitably chosen[7].

Then, for any $q, q' \le \frac{p}{12}$, with $q < q'$, one defines 24 matrices $(X_{q,q'}^{(i)})_{i\le 24}$

$$X_{q,q'}^{(i)} = \begin{pmatrix} 0_{3\times 12(q-1)} & G_i' & 0_{3\times 12(q'-q-1)} & G_i'' & 0_{3\times(p-12q')} \end{pmatrix},$$

for appropriate G_i', G_i''.

In the exact same way as in the case $d = 2$, one can then check (it is tedious to do by hand, but easy on a computer), by dividing it into $12\times 12, 12\times 3, 3\times 12$ and 3×3 blocks, that any matrix $M \in \mathbb{S}^{(p+3)\times(p+3)}$ belonging to the orthogonal of the vector space on the left-hand side of Condition (D.2) is zero, meaning that Condition (D.2) is satisfied. ☐

D.3.1 Proof of Proposition 10

Proposition (Proposition 10). *For any $V \in \mathcal{M}_p$ and $k \le S$, we denote by V_k the k-th $d\times p$ block of V. If*

$$\left\{ \begin{pmatrix} 0_{p\times d} \\ I_d \end{pmatrix} A_0 \begin{pmatrix} 0_{d\times p} & I_d \end{pmatrix} + \sum_{k=1}^{S} \begin{pmatrix} V_k^T \\ I_d \end{pmatrix} A_k \begin{pmatrix} V_k & I_d \end{pmatrix}, A_0, \dots, A_S \in \mathbb{S}^{d\times d} \right\} = \mathbb{S}^{(p+d)\times(p+d)}, \tag{D.3}$$

then ψ_V is injective.

Proof of Proposition 10. We recall that ψ_V is the map

$$\begin{aligned} \psi_V : (T,R) \in \mathbb{S}^{p\times p} \times \mathbb{R}^{d\times p} &\\ &\to \mathcal{A}\left(\begin{pmatrix} V & U_0 \end{pmatrix} \begin{pmatrix} T \\ R \end{pmatrix} V^T + V \begin{pmatrix} T \\ R \end{pmatrix}^T \begin{pmatrix} V & U_0 \end{pmatrix}^T \right) \in \mathbb{R}^{Sd(d+1)/2} \\ &= \mathcal{A}\left(\begin{pmatrix} V & U_0 \end{pmatrix} \begin{pmatrix} 2T & R^T \\ R & 0 \end{pmatrix} \begin{pmatrix} V^T \\ U_0^T \end{pmatrix} \right). \end{aligned}$$

We define

$$\begin{aligned} \zeta_V : Q \in \mathbb{S}^{(p+d)\times(p+d)} &\\ \to \left(T_{sup}\left(\begin{pmatrix} 0_{d\times p} & I_d \end{pmatrix} Q \begin{pmatrix} 0_{p\times d} \\ I_d \end{pmatrix} \right), \mathcal{A}\left(\begin{pmatrix} V & U_0 \end{pmatrix} Q \begin{pmatrix} V^T \\ U_0^T \end{pmatrix} \right) \right) &\in \mathbb{R}^{(S+1)d(d+1)/2}. \end{aligned}$$

[7]For the same reason as in the proof of Remark 4, a "generic" choice of $G_1, \dots, G_{19}$ such that $G_i G_i^T = I_3$ is actually fine.

For any T, R, $\psi_V(T,R) = 0$ implies $\zeta_V\left(\left(\begin{smallmatrix} 2T & R^T \\ R & 0 \end{smallmatrix}\right)\right) = (0, \psi_V(T,R)) = 0$, so, if ζ_V is injective, then ψ_V also.

The map ζ_V is injective if and only if its dual

$$\zeta_V^* : (A_k)_{0\le k\le S} \in (\mathbb{R}^{d(d+1)/2})^{(S+1)}$$
$$\to \left(\begin{smallmatrix} 0_{p\times d} \\ I_d \end{smallmatrix}\right) T_{sup}^*(A_0) \left(\begin{smallmatrix} 0_{d\times p} & I_d \end{smallmatrix}\right) + \sum_{k=1}^{S} \left(\begin{smallmatrix} V_k^T \\ I_d \end{smallmatrix}\right) T_{sup}^*(A_k) \left(\begin{smallmatrix} V_k & I_d \end{smallmatrix}\right) \in \mathbb{S}^{(p+d)\times(p+d)}$$

is surjective. As T_{sup}^* is surjective, the surjectivity of ζ_V^* is equivalent to Condition (D.3). □

D.4 Proof of Corollary 3

Corollary (Corollary 3)**.** *If $p \in \mathbb{N}$ is such that*

$$\frac{p(p+1)}{2} + p > S,$$

then, for almost any cost matrix C, all second-order critical points of the Burer-Monteiro factorization of Problem (SDP-Product) are globally optimal.

On the other hand, for any $p \in \mathbb{N}^*$ such that

$$\frac{p(p+1)}{2} + p \le S,$$

the set of cost matrices admits a subset with non-zero Lebesgue measure on which

- *Problem (SDP-Product) has a unique global optimum, which has rank 1;*
- *Its Burer-Monteiro factorization with rank p has at least one non-optimal second-order critical point.*

Proof of Corollary 3. The proof of the first part of the corollary is identical to the one in Subsection D.3[8], so we prove only the second part. Let us consider p such that

$$\frac{p(p+1)}{2} + p \le S. \tag{D.4}$$

We apply Theorem 2; we simply have to check that its hypotheses are satisfied. Let us define

$$\forall s \le S, u_{0,s} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{d_s \times 1} \quad \text{and} \quad \tilde{U}_0 = \begin{pmatrix} u_{0,1} \\ \vdots \\ u_{0,S} \end{pmatrix} \in \mathbb{R}^{D\times 1}.$$

We set $X_0 = \tilde{U}_0 \tilde{U}_0^T$. It is an extreme point of $\mathcal{C}$, with rank 1, so the first hypothesis holds true.

The second hypothesis (p-regularity of $(\mathcal{A}, b)$) can be checked in the same way as for *MaxCut*.

Let us turn to the third hypothesis: The existence of $V \in \mathcal{M}_p$ at which $\mathcal{M}_p$ is X_0-minimally secant (that is, the map ψ_V of Equation (2.2) is injective). We are going to deduce it from the reasoning done for *MaxCut* (Corollary 1).

For any $W \in \mathbb{R}^{S\times p}$, we denote ψ_W^{MaxCut} the map

$$\psi_W^{MaxCut} \in (T, R) \in \mathbb{S}^{p\times p} \times \mathbb{R}^{1\times p}$$
$$\to \operatorname{diag}\left(\left(\begin{matrix} W & \begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} \end{matrix}\right) \left(\begin{smallmatrix} T \\ R \end{smallmatrix}\right) W^T + W \left(\begin{smallmatrix} T \\ R \end{smallmatrix}\right)^T \left(\begin{matrix} W & \begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} \end{matrix}\right)^T\right) \in \mathbb{R}^S.$$

While proving Corollary 1 (Subsection D.3, case $d = 1$), we have seen that, in the *MaxCut* case, the third hypothesis of Theorem 2 is satisfied when Inequality (D.4) holds. In other words, there exists $W \in \mathbb{R}^{S\times p}$, such that $\operatorname{diag}(WW^T) = 1$ and ψ_W^{MaxCut} is injective.

[8] Provided that $p \ge 2$. When $p = 1$ (hence also $S = 1$), we have to check the hypotheses of Theorem 1 by hand, but it is not difficult.

Let W be such a matrix. We denote by $W_1, \ldots, W_S$ its rows, and define

$$V = \begin{pmatrix} W_1 \\ 0_{d_1-1,p} \\ \vdots \\ W_S \\ 0_{d_S-1,p} \end{pmatrix}.$$

With this definition, we observe that, for any $(T, R) \in \mathbb{S}^{p\times p} \times \mathbb{R}^{1\times p}$,

$$\psi_V(T, R) = \psi_W^{MaxCut}(T, R).$$

In particular, as ψ_W^{MaxCut} is injective, ψ_V is also injective.

□

References

E. Abbe, A. S. Bandeira, and G. Hall. Exact recovery in the stochastic block model. *Transactions on Information Theory*, 62(1):471–487, 2016.

P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds.* Princeton University Press, 2009.

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 Conference on Computational Learning Theory*, 2016a.

A. S. Bandeira, C. Kennedy, and A. Singer. Approximating the little grothendieck problem over the orthogonal and unitary groups. *Mathematical programming*, 160(1-2):433–475, 2016b.

S. Bhojanapalli, N. Boumal, P. Jain, and P. Netrapalli. Smoothed analysis for low-rank solutions to semidefinite programs in quadratic penalty form. In *Proceedings of the 31st Conference On Learning Theory*, pages 3243–3270, 2018.

N. Boumal. A riemannian low-rank method for optimization over semidefinite matrices with block-diagonal constraints. Technical report, , 2015. http://arxiv.org/abs/1506.00575.

N. Boumal, P.-A. Absil, and C. Cartis. Global rates of convergence for nonconvex optimization on manifolds. *IMA Journal of Numerical Analysis*, 2016.

N. Boumal, V. Voroninski, and A. S. Bandeira. Deterministic guarantees for Burer-Monteiro factorizations of smooth semidefinite programs. *preprint*, 2018. https://arxiv.org/abs/1804.02008.

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

M.R. Celis. *A trust region strategy for nonlinear equality constrained optimization.* PhD thesis, Rice University, 1985.

K. N. Chaudhury, Y. Khoo, and A. Singer. Global registration of multiple point clouds using semidefinite programming. *SIAM Journal on Optimization*, 25(1):468–501, 2015.

M. Cucuringu. Sync-rank: Robust ranking, constrained ranking and rank aggregation via eigenvector and SDP synchronization. *IEEE Transactions on Network Science and Engineering*, 3(1):58–79, 2016.

M. Cucuringu, Y. Lipman, and A. Singer. Sensor network localization by eigenvector synchronization over the euclidean group. *ACM Transactions on Sensor Networks (TOSN)*, 8(3), 2012.

C. Delorme and S. Poljak. Laplacian eigenvalues and the maximum cut problem. *Mathematical Programming*, 62(3):557–574, 1993.

A. Eriksson, C. Olsson, F. Kahl, and T.-J. Chin. Rotation averaging and strong duality. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 127–135, 2018.

R. Ge, J. D. Lee, and T. Ma. Matrix completion has no spurious local minimum. In *to appear in Advances in Neural Information Processing Systems 29*. Curran Associates, Inc., 2016.

M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

M. Jaggi. Revisiting Frank-Wolfe: Projection-free sparse convex optimization. In *Proceedings of the 30th International Conference on Machine Learning*, pages 427–435, 2013.

M. Journée, 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.

S. Laue. A hybrid algorithm for convex semidefinite optimization. In *Proceedings of the 29th International Conference on Machine Learning*, pages 177–184, 2012.

Q. Li, Z. Zhu, and G. Tang. The non-convex geometry of low-rank matrix optimization. *Information and Inference: A Journal of the IMA*, 2018.

S. Mei, T. Misiakiewicz, A. Montanari, and R. I. Oliveira. Solving SDPs for synchronization and maxcut problems via the grothendieck inequality. *preprint*, 2017. https://arxiv.org/abs/1703.08729.

B. Mishra, G. Meyer, S. Bonnabel, and R. Sepulchre. Fixed-rank matrix factorizations and riemannian low-rank optimization. *Computational Statistics*, 29(3-4):591–621, 2014.

Gábor 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.

S. Poljak and F. Rendl. Nonpolyhedral relaxations of graph-bisection problems. *SIAM Journal on Optimization*, 5(3):467–487, 1995.

T. Pumir, S. Jelassi, and N. Boumal. Smoothed analysis of the low-rank approach for smooth semidefinite programs. In *Advances in Neural Information Processing Systems*, page To appear., 2018.

D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard. Se-sync: A certifiably correct algorithm for synchronization over the special euclidean group. *To appear in International journal of robotics research*, 2016.

J. Sun, Q. Qu, and J. Wright. A geometric analysis of phase retrieval. *Foundations of Computational Mathematics*, 2017.

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

L. Wang, A. Singer, and Z. Wen. Orientation determination of cryo-em images using least unsquared deviations. *SIAM journal on imaging sciences*, 6(4):2450–2483, 2013.

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