A Semidefinite Relaxation for Sums of Heterogeneous Quadratics on the Stiefel Manifold

Kyle Gilman* Sam Burer† Laura Balzano*

May 30, 2022

Abstract

We study the maximization of sums of heterogeneous quadratic functions over the Stiefel manifold, a nonconvex problem that arises in several modern signal processing and machine learning applications such as heteroscedastic probabilistic principal component analysis (HPPCA). In this work, we derive a novel semidefinite program (SDP) relaxation and study a few of its theoretical properties. We prove a global optimality certificate for the original nonconvex problem via a dual certificate, which leads us to propose a simple feasibility problem to certify global optimality of a candidate local solution on the Stiefel manifold. In addition, our relaxation reduces to an assignment linear program for jointly diagonalizable problems and is therefore known to be tight in that case. We generalize this result to show that it is also tight for close-to jointly diagonalizable problems, and we show that the HPPCA problem has this characteristic. Numerical results validate our global optimality certificate and sufficient conditions for when the SDP is tight in various problem settings.

Contents

*Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI, 48105 USA (email: kgilman@umich.edu). K. Gilman and L. Balzano were supported in part by ARO YIP award W911NF1910027, AFOSR YIP award FA9550-19-1-0026, and NSF BIGDATA award IIS-1838179.

†Department of Business Analytics, University of Iowa , Iowa City, IA 52242 USA.

1 Introduction

This paper studies the problem known in the literature as *the maximization of sums of heterogeneous quadratic functions over the Stiefel manifold*[1] [15; 7; 35; 6]. Specifically, given $d \times d$-size symmetric positive semidefinite (PSD) matrices $\mathbf{M}_1, \ldots, \mathbf{M}_k \succeq 0$ for $k \leq d$, we wish to maximize the convex objective function $\sum_{i=1}^k \boldsymbol{u}_i' \mathbf{M}_i \boldsymbol{u}_i$ over the nonconvex constraint that $\boldsymbol{U} = [\boldsymbol{u}_1 \cdots \boldsymbol{u}_k] \in \mathbb{R}^{d \times k}$ has orthonormal columns:

$$\max_{\boldsymbol{U} \in \mathrm{St}(k,d)} \sum_{i=1}^k \boldsymbol{u}_i' \mathbf{M}_i \boldsymbol{u}_i, \tag{1}$$

where $\mathrm{St}(k,d) = \{\boldsymbol{U} \in \mathbb{R}^{d \times k} : \boldsymbol{U}'\boldsymbol{U} = \mathbf{I}_k\}$ denotes the Stiefel manifold. This problem arises in modern signal processing and machine learning applications like heteroscedastic probabilistic principal component analysis (HPPCA) [24], heterogeneous clutter in radar sensing [38], and robust sparse PCA [14]. Each of these applications involves learning a signal subspace for data possessing heterogeneous statistics.

In particular, HPPCA [24] models data collected from sources of varying quality with different additive noise variances, and estimates the best approximating low-dimensional subspace by maximizing the likelihood, providing superior estimation compared to standard PCA. Specifically, given L data groups $[\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_L]$ with $\boldsymbol{Y}_\ell \in \mathbb{R}^{d \times n_\ell}$, second-order statistics $\mathbf{A}_\ell := \frac{1}{v_\ell} \boldsymbol{Y}_\ell \boldsymbol{Y}_\ell' \succeq 0$ for $\ell \in [L]$, and known positive weights $w_{\ell,i}$ for $(\ell, i) \in [L] \times [k]$, a subproblem of HPPCA involves optimizing the sum of Brockett cost functions [2, Section 4.8] with respect to a k-dimensional orthonormal basis $\boldsymbol{U}$, and can be equivalently recast in the form (1) as follows:

$$\max_{\boldsymbol{U}:\boldsymbol{U}'\boldsymbol{U}=\mathbf{I}} \sum_{\ell=1}^L \mathrm{tr}(\boldsymbol{U}' \mathbf{A}_\ell \boldsymbol{U} \boldsymbol{W}_\ell) = \max_{\boldsymbol{U}:\boldsymbol{U}'\boldsymbol{U}=\mathbf{I}} \sum_{\ell=1}^L \sum_{i=1}^k \boldsymbol{u}_i' w_{\ell,i} \mathbf{A}_\ell \boldsymbol{u}_i = \max_{\boldsymbol{U}:\boldsymbol{U}'\boldsymbol{U}=\mathbf{I}} \sum_{i=1}^k \boldsymbol{u}_i' \mathbf{M}_i \boldsymbol{u}_i, \tag{2}$$

where $\boldsymbol{W}_\ell := \mathrm{diag}(\{w_{\ell,i}\}_{i=1}^k)$ for all ℓ and $\mathbf{M}_i := \sum_{\ell=1}^L w_{\ell,i} \mathbf{A}_\ell$ for all i. Other sensing problems such as independent component analysis (ICA) [40] and approximate joint diagonalization (AJD) [33] also model

[1] We note here that "heterogeneous" refers to the fact the $\mathbf{M}_i$ are distinct and the problem is not separable in each $\boldsymbol{u}_i$. Indeed, the objective in (1) is a homogeneous polynomial in the entries of $\boldsymbol{U}$ since all terms are degree 2.

data with heterogeneous statistics and optimize objective functions of a similar form, as we discuss more in Section 3.

For (2), the case of a single Brockett cost function ($L = 1$) has known analytical solutions obtained by the SVD or eigendecomposition [2, Section 4.8], whereas analytical solutions are not known for $L \geq 2$. Indeed, for $L \geq 2$ and general $\mathbf{A}_\ell$, few, if any, guarantees for optimal recovery exist except in special cases, such as when the constructed $\mathbf{M}_i$ commute [7]. Generally speaking, existing theory only gives restrictive sufficient conditions for global optimality that are typically difficult to check in practice; see [35], for example. Given that (1) is nontrivial and challenging in several ways—nonconvex due to the Stiefel manifold constraint, non-separable because of the weighted sum of objectives, and not readily solved by singular value or eigenvalue decomposition—many works apply iterative local solvers to the nonconvex problem (1).

However, given the nonconvexity of (1), it can be unclear whether these local approaches find a global maximum. An alternative approach is to relax problems such as (1) to a semidefinite program (SDP), allowing the use of standard convex solvers. While the SDP has stronger optimality guarantees, the challenge is then to derive conditions under which the SDP is tight, i.e., returns the optimal solution to the original nonconvex problem. SDP relaxations such as the "Fantope" [20; 31] exist for solving PCA-like problems, but to the best of our knowledge, no previous convex methods exist to solve (1).

The main contribution of this paper is a novel convex SDP relaxation of (1), whose constraint set is related to the Fantope but distinctly unique. By studying this SDP and its optimality criteria, we derive sufficient conditions to certify the global optimality of a local stationary point, e.g., a candidate solution obtained from any iterative solver for the nonconvex problem. We then propose a straightforward method to certify global optimality by solving a much smaller SDP feasibility problem that scales favorably with the problem dimensions. Our work also generalizes existing results for (1) with commuting matrices to the case with "almost commuting" matrices, showing that as long as the data matrices are within an open neighborhood of a commuting tuple of data matrices (to be defined precisely in Section 4.2), the SDP is tight and provably recovers an optimal solution of (1).

Notation We use italic, boldface, upper case letters $\boldsymbol{A}$ to denote matrices, $\boldsymbol{v}$ to denote vectors, and c for scalars. We denote the cone of $d \times d$-size symmetric positive semidefinite matrices as $\mathbb{S}^d_+$, and use $\boldsymbol{A} \succeq 0$ to denote an element $\boldsymbol{A} \in \mathbb{S}^d_+$. We denote the Hermitian transpose of a matrix as $\boldsymbol{A}'$, the trace of a matrix as $\mathrm{tr}(\boldsymbol{A})$, and the inner product of matrices (with compatible inner dimensions) $\langle \boldsymbol{A}, \boldsymbol{B} \rangle := \mathrm{tr}(\boldsymbol{A}'\boldsymbol{B})$. The spectral norm of a matrix is denoted by $\|\boldsymbol{A}\|$ and the Frobenius norm by $\|\boldsymbol{A}\|_F$. The identity matrix of size d is denoted as $\mathbf{I}_d$. Notation $i \in [k]$ means $i = 1, \ldots, k$.

2 Semidefinite program relaxation

By relaxing the considered nonconvex problem (1) to a convex one, the well-established principles of convex optimization permit us to study when an optimal solution of the SDP relaxation recovers a global maximum of (1) and importantly, when a given local stationary point is a global maximum. After re-expressing the original problem using equivalent constraints, we lift the variables into the cone of PSD matrices, relax the nonconvex constraints to convex surrogates, and obtain an SDP.

First, we begin by slightly rewriting (1) and the Stiefel manifold constraints as

$$\max_{\boldsymbol{u}_1,\ldots,\boldsymbol{u}_k} \mathrm{tr}\left(\sum_{i=1}^k \mathbf{M}_i \boldsymbol{u}_i \boldsymbol{u}_i'\right) \quad \text{s.t. } \mathrm{tr}\left(\boldsymbol{u}_i \boldsymbol{u}_i'\right) = 1 \quad \forall i \in [k], \quad \mathrm{tr}\left(\boldsymbol{u}_j \boldsymbol{u}_i'\right) = 0 \quad \forall i \neq j. \tag{3}$$

Letting $\boldsymbol{X}_i = \boldsymbol{u}_i \boldsymbol{u}_i' \in \mathbb{R}^{d \times d}$, this is equivalent to the lifted problem:

$$\begin{aligned} \max_{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k} \mathrm{tr}\left(\sum_{i=1}^k \mathbf{M}_i \boldsymbol{X}_i\right) \quad \text{s.t. } & \lambda_j\left(\sum_{i=1}^k \boldsymbol{X}_i\right) \in \{0,1\} \quad \forall j \in [d] \\ & \mathrm{tr}(\boldsymbol{X}_i) = 1, \quad \mathrm{rank}(\boldsymbol{X}_i) = 1, \quad \boldsymbol{X}_i \succcurlyeq 0 \quad \forall i \in [k], \end{aligned} \tag{4}$$

where $\lambda_j(\cdot)$ indicates the j-th eigenvalue of its argument. Note that this problem is nonconvex due to the rank constraint and the constraint that the eigenvalues are binary. Similar to the relaxations in [41; 30], we relax the eigenvalue constraint in (4) to $0 \preccurlyeq \sum_{i=1}^k \boldsymbol{X}_i \preccurlyeq \mathbf{I}$ and remove the rank constraint, which yields the SDP relaxation we consider throughout the remainder of this work.

$$p^* = \min_{\boldsymbol{X}_1,\ldots,\boldsymbol{X}_k} \quad -\operatorname{tr}\left(\sum_{i=1}^k \mathbf{M}_i \boldsymbol{X}_i\right) \quad \text{s.t. } \sum_{i=1}^k \boldsymbol{X}_i \preccurlyeq \mathbf{I}, \quad \operatorname{tr}(\boldsymbol{X}_i) = 1, \quad \boldsymbol{X}_i \succcurlyeq 0 \quad i = 1,\ldots,k. \tag{SDP-P}$$

Note that $0 \preccurlyeq \sum_{i=1}^k \boldsymbol{X}_i$ can be omitted since it is already satisfied when $\boldsymbol{X}_i \succcurlyeq 0$ for all i. The feasible set of (SDP-P) is closely related to the convex set found in [41; 30; 21] called the *Fantope*. The Fantope is the convex hull of all matrices $\boldsymbol{U}\boldsymbol{U}' \in \mathbb{R}^{d\times d}$ such that $\boldsymbol{U} \in \mathbb{R}^{d\times k}$ and $\boldsymbol{U}'\boldsymbol{U} = \mathbf{I}$ [20; 31]. Indeed, our relaxation can be viewed as providing a decomposition of the Fantope variable $\boldsymbol{X}$ into the sum $\boldsymbol{X}_1 + \cdots + \boldsymbol{X}_k$ such that each $\boldsymbol{X}_i$ satisfies $\operatorname{tr}(\boldsymbol{X}_i) = 1$ and $0 \preceq \boldsymbol{X}_i \preceq \mathbf{I}$. This decomposition allows (SDP-P) to capture the exact form of the objective function, which sums the individual terms $\operatorname{tr}(\mathbf{M}_i\boldsymbol{X}_i)$. Precisely, the feasible set of (SDP-P) is a convex relaxation of the set $\{(\boldsymbol{u}_1\boldsymbol{u}_1', \ldots, \boldsymbol{u}_k\boldsymbol{u}_k') : \boldsymbol{U}'\boldsymbol{U} = \mathbf{I}\}$. Naturally, one wonders whether our relaxation always solves the original nonconvex problem. We show in Appendix C that it does not, using a counter-example. Our work therefore studies this SDP in two ways: First, we provide a global optimality certificate. Second, we study a class of "close-to jointly diagonalizable" problem instances, which includes the heteroscedastic PCA problem, and show that the SDP is tight for this class.

For dual variables $\boldsymbol{Z}_i \in \mathbb{S}_+^d$, $\boldsymbol{Y} \in \mathbb{S}_+^d$, $\boldsymbol{\nu} \in \mathbb{R}^k$, the dual of (SDP-P), which will play a central role in the theory of this paper, is

$$d^* = \min_{\boldsymbol{Y},\boldsymbol{Z}_i,\boldsymbol{\nu}} \operatorname{tr}(\boldsymbol{Y}) + \sum_{i=1}^k \nu_i \quad \text{s.t. } \boldsymbol{Y} \succcurlyeq 0, \quad \boldsymbol{Y} = \mathbf{M}_i + \boldsymbol{Z}_i - \nu_i\mathbf{I}, \quad \boldsymbol{Z}_i \succcurlyeq 0 \quad \forall i \in [k]. \tag{SDP-D}$$

Definition 2.1 (Rank-one property (ROP))**.** *A solution to* (SDP-P) *is said to have the rank-one property if $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_k$ are all rank-one.*

We note that if a solution has the rank-one property, the first singular vectors of the $\boldsymbol{X}_i$ are necessarily mutually orthogonal, and $\sum_i \boldsymbol{X}_i$ is a projection matrix, due to the constraint $\sum_i \boldsymbol{X}_i \preccurlyeq \mathbf{I}$.

2.1 Strong duality and equivalence of optimizers

Our theoretical results require the following lemmas, whose proofs can be found in Appendix B.

Lemma 2.2. *Strong duality holds for the SDP relaxation with primal* (SDP-P) *and dual* (SDP-D)*.*

Lemma 2.3. *The solution to the SDP relaxation in* (SDP-P) *is the optimal solution to the original nonconvex problem in* (1) *(equivalently* (4)*) if and only if the optimal $\boldsymbol{X}_i$ have the rank-one property.*

Lemma 2.4. *Assume $\mathbf{M}_i$ are PSD. Then the optimal $\nu_i \geq 0$.*

Lemma 2.5. *If the optimal dual variables $\boldsymbol{Z}_i$ for $i = 1, \ldots, k$ are each rank $d-1$, the optimal solution variables $\boldsymbol{X}_i$ have the rank-one property.*

3 Related work

There are a few important related works on the objective in (1), as well as many more related works than can be reviewed here, including ones on eigenvalue/eigenvector problems and the many variations thereof, low-rank SDPs, nonconvex quadratics where $\mathbf{M}_i$ are not PSD, etc. Appendix A in the supplement provides an extensive related work section. Here, we focus on the works most directly related to (1).

Bolla et al. [7], Rapcsák [35], and Berezovskyi [6] previously investigated the sum of heterogeneous quadratics in (1). The work in [7] only studied the structure of this problem when the matrices $\mathbf{M}_i$ were

commuting. Rapcsák [35] derived sufficient second-order global optimality conditions, but these conditions are difficult to check in general and do not seem to hold for the heteroscedastic PCA problem. Works such as Huang and Palomar [25] and Pataki [32] consider a very similar problem to (2), but without the eigenvalue constraint in (4), making their SDP a rank-constrained separable SDP; see also Luo et al. [30, Section 4.3]. Pataki studied upper bounds on the rank of optimal solutions of general SDPs, but in the case of (SDP-P), since our problem introduces the additional constraint summing the $\boldsymbol{X}_i$, Pataki's bounds do not guarantee rank-one, or even low-rank, optimal solutions.

Recent works have also studied convex relaxations of PCA and other low-rank subspace problems that seek to bound the eigenvalues of a single matrix [41; 39; 44], rather than the sum of multiple matrices as in our setting. The works in [10; 34] prove global convergence of nonconvex Burer–Monteiro factorization [16] approaches to solve low-rank semidefinite programs without orthogonality constraints. Other works have studied optimizers of the nonconvex problem, like those in [14; 13; 38; 12; 24], using minorize-maximize or Riemannian gradient ascent algorithms, which do not come with optimality guarantees. Our problem also has interesting connections to approximate joint diagonalization (AJD), which is well-studied and often applied to blind source separation or independent component analysis (ICA) problems [40; 8; 26; 3; 37]. See Appendix A of the supplement for further details.

4 Theoretical Results

4.1 Dual certificate of the SDP

In practical settings for high-dimensional data, a variety of iterative local methods are often applied to solve nonconvex problems over the Stiefel manifold, from gradient ascent by geodesics [2; 18; 1] and more recently majorization-minimization (MM) algorithms, where Breloy et al. [14] applied MM methods to solve (1) with guarantees of convergence to a stationary point. While the computational complexity and memory requirements of these solvers scale well, their obtained solutions lack any global optimality guarantees. Our contribution seeks to fill this gap by proposing a check for global optimality of a local solution.

By Lemma 2.3, an optimal solution of (SDP-P) with rank-one matrices $\boldsymbol{X}_i$ globally solves the original nonconvex problem (1). In this section, given a candidate stationary point $\bar{\boldsymbol{U}} = [\bar{\boldsymbol{u}}_1 \cdots \bar{\boldsymbol{u}}_k] \in \mathrm{St}(k,d)$ to (1), we investigate conditions guaranteeing that the rank-one matrices $\bar{\boldsymbol{X}}_i = \bar{\boldsymbol{u}}_i \bar{\boldsymbol{u}}_i'$, which are feasible for (SDP-P), in fact comprise an optimal solution of (SDP-P), implying that $\bar{\boldsymbol{U}}$ optimizes (1). Similar to [43; 19] for Fantope problems, our results yield a dual SDP certificate to verify the primal optimality of the candidates $\bar{\boldsymbol{X}}_1, \ldots, \bar{\boldsymbol{X}}_k$ constructed from a local solution $\bar{\boldsymbol{U}}$. We show our certificate scales favorably in computation over the full SDP, with the most complicated computations of our algorithm requiring us to solve a feasibility problem in k variables with $d \times d$-size linear matrix inequalities (LMI).

Theorem 4.1. *Let $\bar{\boldsymbol{U}} \in \mathrm{St}(k,d)$ be a local maximizer to* (1)*, and let $\bar{\boldsymbol{\Lambda}} = \sum_{i=1}^k \bar{\boldsymbol{U}}' \mathbf{M}_i \bar{\boldsymbol{U}} \mathbf{E}_i$, where $\mathbf{E}_i \triangleq \boldsymbol{e}_i \boldsymbol{e}_i'$ and $\boldsymbol{e}_i$ is the i^{th} standard basis vector in $\mathbb{R}^k$. If there exist $\bar{\boldsymbol{\nu}} = [\bar{\nu}_1 \cdots \bar{\nu}_k] \in \mathbb{R}^k_+$ such that*

$$\begin{aligned} \bar{\boldsymbol{U}}(\bar{\boldsymbol{\Lambda}} - \mathbf{D}_{\bar{\boldsymbol{\nu}}})\bar{\boldsymbol{U}}' + \bar{\nu}_i \mathbf{I} - \mathbf{M}_i \succeq 0 \quad \forall i = 1, \ldots, k \\ \bar{\boldsymbol{\Lambda}} - \mathbf{D}_{\bar{\boldsymbol{\nu}}} \succeq 0, \end{aligned} \tag{5}$$

where $\mathbf{D}_{\bar{\boldsymbol{\nu}}} := \mathit{diag}(\bar{\nu}_1, \ldots, \bar{\nu}_k)$, then $\bar{\boldsymbol{U}}$ is an optimal solution to (SDP-P) *and a globally optimal solution to the original nonconvex problem* (1)*.*

In light of Theorem 4.1, to test whether a candidate stationary point $\bar{\boldsymbol{U}}$ is globally optimal, we simply assess whether system (5) is feasible using an LMI solver. If it is indeed feasible, then $\bar{\boldsymbol{U}}$ is globally optimal. On the other hand, if (5) is infeasible, it indicates one of two things: 1) The SDP does not return tight, rank-one solutions $\bar{\boldsymbol{X}}_i$, and the SDP strictly upper bounds the original problem on the Stiefel manifold. The candidate stationary point $\bar{\boldsymbol{U}}$ may or may not be globally optimal to the original nonconvex problem. 2) The SDP is tight, but the candidate stationary point $\bar{\boldsymbol{U}}$ is a suboptimal local maximum. The proof, found in Appendix D, constructs dual variables and verifies the KKT conditions. Appendix H also describes an extension of the certificate to the sum of Brocketts with additive linear terms.

Arithmetic complexity While SDP relaxations of nonconvex optimization problems can provide strong provable guarantees, their practicality can be limited by the time and space required to solve them, particularly when using off-the-shelf interior-point solvers, which in our case require $\mathcal{O}(d^3)$ [5] storage and floating point operations (flops) per iteration. In contrast, the SDP relaxation admits improved practical tools to transfer theoretical guarantees to the nonconvex setting. The proposed global certificate significantly reduces the number of variables from $\mathcal{O}(d^2)$ in (SDP-D) (upon eliminating the variables $\boldsymbol{Z}_i$) to merely k variables in (5). Using [4, Section 6.6.3] it can be shown that computing the certificate only, with a given $\bar{\boldsymbol{U}}$, results in a substantial reduction in flops by a factor of $\mathcal{O}(d^3/k)$ over solving (SDP-D). Subsequently, a first-order MM solver [14], whose cost is $\mathcal{O}(dk^2 + k^3)$ per iteration, combined with our global optimality certificate, is an obvious preference to solving the full SDP in (SDP-P) for large problems. See Appendix D.1 for more details.

4.2 SDP tightness in the close-to jointly diagonalizable (CJD) case

While Section 4.1 provides a technique to certify the global optimality of a solution to the nonconvex problem, the check will fail if the point is not globally optimal or if the SDP is not tight. General conditions on $\mathbf{M}_i$ that guarantee tightness of (SDP-P) are still not known. However, when the matrices $\mathbf{M}_i$ are jointly diagonalizable, our problem reduces to a linear programming (LP) assignment problem, and by standard LP theory, a solution with rank-one $\mathbf{X}_i$ exists and the SDP (or equivalent LP) is a tight relaxation [7]. Our next major contribution is to show that a solution with rank-one $\mathbf{X}_i$ exists also for cases that are *close-to jointly diagonalizable* (CJD). We first give a continuity result showing there is a neighborhood around the diagonal case for which (SDP-P) is still tight. Then we show that for the HPPCA problem, the matrices $\mathbf{M}_i$ are close-to diagonalizable and can be made arbitrarily close as the number of data points n grows or the noise levels diminish. This gives strong support for tightness of the SDP for the HPPCA problem when n is large or the noise levels are small.

Definition 4.2 (Close-to jointly diagonalizable (CJD))**.** *We say that unit spectral-norm matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are CJD if they are almost commuting, that is, when the commuting distance $\boldsymbol{A}$ and $\boldsymbol{B}$, $\|[\boldsymbol{A}, \boldsymbol{B}]\| := \|\boldsymbol{AB} - \boldsymbol{BA}\| \leq \delta$ for $0 < \delta \ll 1$.*

4.2.1 Continuity and tightness in the CJD case

In this section, we employ a technical continuity result for the dual feasible set to conclude that there is a neighborhood of problem instances around every diagonal instance for which (SDP-P) gives rank-one solutions $\boldsymbol{X}_i$. All proofs for results in this subsection are found in Appendix E.

Given a k-tuple of symmetric matrices $(\mathbf{M}_1, \ldots, \mathbf{M}_k)$, our primal-dual pair is given by (SDP-P) and (SDP-D). Note that, without loss of generality, we may assume each $\mathbf{M}_i$ is positive semidefinite since the primal constraint $\mathrm{tr}(\boldsymbol{X}_i) = 1$ ensures that replacing $\mathbf{M}_i$ by $\mathbf{M}_i + \lambda_i \boldsymbol{I} \succeq 0$, where λ_i is a positive constant, simply shifts the objective value by λ_i. In addition, we have shown in Lemma 2.4 that $\mathbf{M}_i \succeq 0$ implies ν_i is nonnegative at dual optimality. So we assume $\mathbf{M}_i \succeq 0$ and enforce $\nu_i \geq 0$ for all $i = 1, \ldots, k$.

For a fixed, user-specified upper bound $\mu > 0$, we define the closed convex set

$$\mathcal{C} := \{\boldsymbol{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) : 0 \preceq \mathbf{M}_i \preceq \mu \boldsymbol{I} \quad \forall\, i = 1, \ldots, k\},$$

to be our set of admissible coefficient k-tuples. We know that both (SDP-P) and (SDP-D) have interior points for all $\boldsymbol{c} \in \mathcal{C}$, so that strong duality holds.

Lemma 4.3. *Let $\boldsymbol{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) \in \mathcal{C}$. If $\mathbf{M}_i$ are jointly diagonalizable for $i = 1, \ldots, k$ and (SDP-P) has a unique optimal solution, then there exists an optimal solution of (SDP-D) with rank$(\boldsymbol{Z}_i) \geq d - 1$ for all $i = 1, \ldots, k$.*

Definition 4.4. *For $\boldsymbol{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) \in \mathcal{C}$ and $\bar{\boldsymbol{c}} = (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k) \in \mathcal{C}$, define dist$(\boldsymbol{c}, \bar{\boldsymbol{c}}) \triangleq \max_{i \in [k]} \|\mathbf{M}_i - \bar{\mathbf{M}}_i\|$.*

We are now ready to state our main result in this subsection.

Theorem 4.5. *Let $\bar{\boldsymbol{c}} := (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k) \in \mathcal{C}$ be given such that $\bar{\mathbf{M}}_i$, $i = 1, \ldots, k$, are jointly diagonalizable and the primal problem (SDP-P) with objective coefficients $\bar{\boldsymbol{c}}$ has a unique optimal solution. Then there exists a full-dimensional neighborhood $\bar{\mathcal{C}} \ni \bar{c}$ in $\mathcal{C}$ such that (SDP-P) has the rank-one property for all $\boldsymbol{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) \in \bar{\mathcal{C}}$.*

Proof Sketch. Using Lemma 4.3, let $\boldsymbol{y}^0 := (\bar{\boldsymbol{Y}}, \bar{\boldsymbol{Z}}_i, \bar{\nu}_i)$ be the optimal solution of the dual problem (SDP-D) for $\bar{\boldsymbol{c}} = (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k)$, which has $\operatorname{rank}(\bar{\boldsymbol{Z}}_i) \geq d-1$ for all i. Proposition E.3 in Appendix E considers a function $y(\boldsymbol{c}; \boldsymbol{y}^0)$ that returns the optimal solution of (SDP-D) for $\boldsymbol{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k)$ closest to $\boldsymbol{y}^0$, and shows that function is continuous. It follows that its preimage

$$y^{-1}\left(\{(\boldsymbol{Y}, \boldsymbol{Z}_i, \nu_i) : \operatorname{rank}(\boldsymbol{Z}_i) \geq d-1 \ \ \forall\ i\}\right)$$

contains $\bar{\boldsymbol{c}}$ and is an open set because the set of all $(\boldsymbol{Y}, \boldsymbol{Z}_i, \nu_i)$ with $\operatorname{rank}(\boldsymbol{Z}_i) \geq d-1$ is an open set. After intersecting with $\mathcal{C}$, we have shown existence of this full-dimensional set $\bar{\mathcal{C}}$. From complementarity of the KKT conditions of the assignment LP, $\operatorname{rank}(\boldsymbol{Z}_i) = d-1$ for $i = 1, \ldots, k$. Applying Lemma 2.5 then completes the theorem. □

The next corollary shows that for a general tuple of almost commuting matrices $\boldsymbol{c} := (\mathbf{M}_1, \ldots, \mathbf{M}_k)$ that are CJD, (SDP-P) is tight and has the rank-one property. In the following results, we will then prove the HPPCA generative model for $\mathbf{M}_i$ results in a problem that is CJD. While these are sufficient conditions, they are by no means necessary, and Appendix F gives an example of $\mathbf{M}_i$ that are *not* CJD but where the convex relaxation has the rank-one property.

Corollary 4.5.1. *Assume $\|\mathbf{M}_i\| \leq 1$ for all $i \in [k]$, and let $\boldsymbol{c} := (\mathbf{M}_1, \ldots, \mathbf{M}_k)$. Suppose $\|[\mathbf{M}_i, \mathbf{M}_j]\| := \|\mathbf{M}_i\mathbf{M}_j - \mathbf{M}_j\mathbf{M}_i\| \leq \delta$ for all $i, j \in [k]$. Then there exists $\bar{\boldsymbol{c}} := (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k)$ of commuting, jointly-diagonalizable matrices such that $\|[\bar{\mathbf{M}}_i, \bar{\mathbf{M}}_j]\| = 0$ for all $i, j \in [k]$ where dist$(\boldsymbol{c}, \bar{\boldsymbol{c}}) \leq \mathcal{O}(\epsilon(\delta))$ and $\epsilon(\delta)$ is a function satisfying $\lim_{\delta \to 0} \epsilon(\delta) = 0$. If δ is small enough, there exists $\epsilon > \epsilon(\delta) > 0$ such that dist$(\boldsymbol{c}, \bar{\boldsymbol{c}}) \leq \epsilon$ implies $\boldsymbol{c} \in \bar{\mathcal{C}}$ and* (SDP-P) *has the rank-one property.*

4.2.2 HPPCA is CJD

Consider the heteroscedastic probabilistic PCA problem in [24] where L data groups of $n_1, \ldots, n_L$ samples ($n = \sum_{\ell=1}^{L} n_\ell$) with known noise variances $v_1, \ldots, v_L$ respectively are generated by the model

$$\boldsymbol{y}_{\ell,i} = \boldsymbol{U}\boldsymbol{\Theta}\boldsymbol{z}_{\ell,i} + \boldsymbol{\eta}_{\ell,i} \in \mathbb{R}^d \quad \forall \ell \in [L], i \in [n_\ell]. \tag{6}$$

Here, $\boldsymbol{U} \in \operatorname{St}(k, d)$ is a planted subspace, $\boldsymbol{\Theta} = \operatorname{diag}(\sqrt{\lambda}_1, \ldots, \sqrt{\lambda}_k)$ represent the known signal amplitudes, $\boldsymbol{z}_{\ell,i} \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{I}_k)$ are latent variables, and $\boldsymbol{\eta}_{\ell,i} \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, v_\ell \mathbf{I}_d)$ are additive Gaussian heteroscedastic noises. Assume that $\lambda_i \neq \lambda_j$ for $i \neq j \in [k]$ and $v_\ell \neq v_m$ for $\ell \neq m \in [L]$. The maximum likelihood problem in [24, Equation 3] with respect to $\boldsymbol{U}$ is then equivalently (2) for $\mathbf{A}_\ell = \sum_{i=1}^{n_\ell} \frac{1}{v_\ell} \boldsymbol{y}_{\ell,i} \boldsymbol{y}'_{\ell,i}$ and $w_{\ell,i} = \frac{\lambda_i}{\lambda_i + v_\ell} \in (0, 1]$. Our next result says that, as the number of samples n grows or the signal-to-noise ratio λ_i / v_ℓ grows, the matrices in the HPPCA problem are almost commuting. The proof is found in Appendix E.

Proposition 4.6. *Let $\boldsymbol{c} = (\frac{1}{n}\mathbf{M}_1, \ldots, \frac{1}{n}\mathbf{M}_k)$ be the (normalized) data matrices of the HPPCA problem. Then there exists commuting $\bar{\boldsymbol{c}} = (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k)$ (constructed in the proof) such that, for a universal constant $C > 0$ and with probability exceeding $1 - e^{-t}$ for $t > 0$,*

$$\frac{\|\mathbf{M}_i - \bar{\mathbf{M}}_i\|}{\|\bar{\mathbf{M}}_1\|} \leq \min\left\{ \sum_{\ell=1}^{L} \frac{1}{\frac{\lambda_i}{v_\ell} + 1}, C\frac{\bar{\sigma}_i}{\bar{\sigma}_1} \max\left\{ \sqrt{\frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i}\log d + t}{n}}, \frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i}\log d + t}{n} \log(n) \right\} \right\}, \quad \textit{where} \tag{7}$$

$$\bar{\sigma}_i = \|\bar{\mathbf{M}}_i\| = \sum_{\ell=1}^{L} \frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell} + 1} \frac{n_\ell}{n} \left(\frac{\lambda_1}{v_\ell} + 1 \right) \quad \textit{and} \quad \bar{\xi}_i = \operatorname{tr}(\bar{\mathbf{M}}_i) = \sum_{\ell=1}^{L} \frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell} + 1} \frac{n_\ell}{n} \left(\frac{1}{v_\ell} \sum_{i=1}^{k} \lambda_i + d \right).$$

5 Numerical experiments

All of the following numerical experiments were computed using MATLAB R2018a on a MacBook Pro with a 2.6 GHz 6-Core Intel Core i7 processor. When solving SDPs, we use the SDPT3 solver of the CVX package in MATLAB [23]. All code necessary to reproduce our experiments is available at `https://github.com/kgilman/Sums-of-Heterogeneous-Quadratics`. When executing each algorithm in practice, we remark that the results may vary with the choice of user specified numerical tolerances and other settings. We point the reader to Appendix G for detailed settings.

5.1 Assessing the rank-one property (ROP)

In this subsection, we conduct experiments showing that, for many random instances of the HPPCA application, the SDP relaxation (SDP-P) is tight with rank-one orthonormal primal solutions $\boldsymbol{X}_i$, yielding a globally optimal solution of (1). Similar experiments for other forms of (1), including where $\mathbf{M}_i$ are random low-rank PSD matrices, are found in Appendix G and give similar insights. Here, the $\mathbf{M}_i$ are generated according to our HPPCA model in (6) where $\mathbf{A}_\ell = \frac{1}{v_\ell}\sum_{i=1}^{n_\ell} \boldsymbol{y}_{\ell,i}\boldsymbol{y}'_{\ell,i}$ and weight matrices $\mathbf{W}_\ell$ are calculated as $\mathbf{W}_\ell = \operatorname{diag}(w_{\ell,1},\ldots,w_{\ell,k})$ for $w_{\ell,i} = \lambda_i/(\lambda_i + v_\ell)$. We make $\boldsymbol{\lambda}$ a k-length vector of entries uniformly spaced in the interval $[1,4]$, and vary ambient dimension d, rank k, samples $\boldsymbol{n}$, and variances $\boldsymbol{v}$ for $L = 2$ noise groups. Each random instance is generated from a new random draw of $\boldsymbol{U}$ on the Stiefel manifold, latent variables $\boldsymbol{z}_{\ell,i}$, and noise vectors $\boldsymbol{\eta}_{\ell,i}$.

Tables 1 and 2 show the results of these experiments for various choices of dimension d and rank k. We solve the SDP for 100 random data instances in Matlab CVX. The table shows the fraction of trials that result in rank-one $\boldsymbol{X}_i$ with first eigenvectors orthogonal across $i = 1,\ldots,k$. We compute the average error of the sorted eigenvalues of each $\bar{\boldsymbol{X}}_i$ to $\boldsymbol{e}_1$, i.e. $\frac{1}{k}\sum_{i=1}^{k} \|\operatorname{diag}(\boldsymbol{\Sigma}_i) - \boldsymbol{e}_1\|_2^2$ where $\bar{\boldsymbol{X}}_i = \boldsymbol{V}_i\boldsymbol{\Sigma}_i\boldsymbol{V}_i'$, and count any trial with error greater than 10^{-5} as not tight. The SDP solutions possess the ROP in the vast majority of trials. As the $\mathbf{M}_i$ concentrate to be almost commuting with increased sample sizes, the convex relaxation becomes tight in 100% of the trials, as shown in Table 1. Similarly, as we decrease the spread of the variances, Table 2 shows the fraction of tight instances increases, reaching 100% in the homoscedastic setting, as expected.

		Fraction of 100 trials with ROP			
		$k=3$	$k=5$	$k=7$	$k=10$
$\boldsymbol{n}=[5,20]$	$d=10$	1	0.99	1	1
	$d=20$	1	0.98	0.98	0.99
	$d=30$	0.99	0.93	0.98	0.97
	$d=40$	0.98	0.91	0.99	0.98
	$d=50$	0.97	0.95	0.96	0.98
$\boldsymbol{n}=[20,80]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	1	1	1	0.98
	$d=40$	1	1	0.97	0.95
	$d=50$	1	0.98	0.98	0.97
$\boldsymbol{n}=[100,400]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	1	1	1	1
	$d=40$	1	1	1	1
	$d=50$	1	1	1	1

Table 1: **(HPPCA)** Numerical experiments showing the fraction of trials where the SDP was tight for instances of the HPPCA problem as we vary d, k, and $\boldsymbol{n}$ using $L = 2$ groups with noise variances $\boldsymbol{v} = [1,4]$.

		Fraction of 100 trials with ROP			
		$k=3$	$k=5$	$k=7$	$k=10$
$\boldsymbol{v}=[1,1]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	1	1	1	1
	$d=40$	1	1	1	1
	$d=50$	1	1	1	1
$\boldsymbol{v}=[1,2]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	1	0.98	1	1
	$d=40$	1	1	0.99	1
	$d=50$	1	1	1	0.99
$\boldsymbol{v}=[1,3]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	0.99	0.99	0.97	0.99
	$d=40$	1	0.98	0.97	0.99
	$d=50$	1	0.97	0.96	0.98

Table 2: **(HPPCA)** Numerical experiments showing the fraction of trials where the SDP was tight for instances of the HPPCA problem as we vary d, k, and $\boldsymbol{v}$ using $L = 2$ groups with samples $\boldsymbol{n} = [10,40]$.

5.2 Assessing global optimality of local solutions

In this section, we use the Stiefel majorization-minimization (StMM) solver with a linear majorizer from Breloy et al. [14] to obtain a local solution $\bar{\boldsymbol{U}}_{\mathrm{MM}}$ to (1) for various inputs $\mathbf{M}_i$ and use Theorem 4.1 to certify the local solution either as globally optimal or as a non-optimal stationary point. For comparison, we obtain candidate solutions $\bar{\boldsymbol{X}}_i$ from the SDP and perform a rank-one SVD of each to form $\bar{\boldsymbol{U}}_{\mathrm{SDP}}$, i.e.

$$\bar{\boldsymbol{U}}_{\mathrm{SDP}} = \mathcal{P}_{\mathrm{St}}([\bar{\boldsymbol{u}}_1 \cdots \bar{\boldsymbol{u}}_k]), \quad \bar{\boldsymbol{u}}_i = \operatorname*{argmax}_{\boldsymbol{u}:\|\boldsymbol{u}\|_2=1} \boldsymbol{u}' \bar{\boldsymbol{X}}_i \boldsymbol{u},$$

while measuring how close the solutions are to being rank-one. In the case the SDP is not tight, the rank-one directions of the $\boldsymbol{X}_i$ will not be orthonormal, so as a heuristic, we project $\bar{\boldsymbol{U}}_{\mathrm{SDP}}$ onto the Stiefel manifold by the orthogonal Procrustes solution, denoted by the operator $\mathcal{P}_{\mathrm{St}}(\cdot)$ [14].

Synthetic CJD matrices To empirically verify our theory from Section 4, we generate each $\mathbf{M}_i \in \mathbb{R}^{d\times d}$ to be a diagonally dominant matrix resembling an approximately rank-r sample covariance matrix, such that, in a similar manner to HPPCA, $\mathbf{M}_1 \succeq \mathbf{M}_2 \succeq \cdots \succeq \mathbf{M}_k \succeq 0$. Specifically, we first construct $\mathbf{M}_k = \mathbf{D}_k + \mathbf{N}_k$, where $\mathbf{D}_k$ is a diagonal matrix with r nonzero entries drawn uniformly at random from $[0, 1]$, and $\mathbf{N}_k = \frac{1}{10d}\mathbf{S}\mathbf{S}'$ for $\mathbf{S} \in \mathbb{R}^{d\times 10d}$ whose entries are drawn i.i.d. as $\mathcal{N}(0, \sigma\mathbf{I})$. We then generate the remaining $\mathbf{M}_i$ for $i = k-1, \ldots, 1$ as $\mathbf{M}_i = \mathbf{M}_{i-1} + \mathbf{D}_i + \mathbf{N}_i$ with new random draws of $\mathbf{D}_i$ and $\mathbf{N}_i$ and normalize all by $1/\max_{i\in[k]} \|\mathbf{M}_i\|$ so that $\|\mathbf{M}_i\| \leq 1$. With this setup, when we sweep σ, we sweep through a range of commuting distances, i.e. $\max_{i,j\in[k]} \|\mathbf{M}_i\mathbf{M}_j - \mathbf{M}_j\mathbf{M}_i\|$. For all experiments, we generate problems with parameters $d = 10$, $k = 3$, $r = 3$, and run StMM for 2000 maximum iterations or until the norm of the gradient on the manifold is less than 10^{-10}.

Fig. 1a shows the gap of the objective values between the SDP relaxation (before projection onto the Stiefel) and the nonconvex problem ($p_{\mathrm{SDP}} - p_{\mathrm{StMM}}$) versus the commuting distance. Fig. 1b shows the distance between the two obtained solutions computed as $\frac{1}{\sqrt{k}}\||\bar{\boldsymbol{U}}'_{\mathrm{StMM}}\bar{\boldsymbol{U}}_{\mathrm{SDP}}| - \mathbf{I}\|_F$ (where $|\cdot|$ denotes taking the elementwise absolute value) versus commuting distance. Fig. 1c shows the percentage of trials where $\bar{\boldsymbol{U}}_{\mathrm{StMM}}$ could not be certified globally optimal. Like before, we declare an SDP's solution "tight" if the mean error of its solutions to rank-one approximations is less than 10^{-5}. Trials with the marker "$\circ$" indicate trials where global optimality was certified. The marker "$\times$" represents trials where $\bar{\boldsymbol{U}}$ was not certified as globally optimal and the SDP relaxation was not tight; "$\triangle$" markers indicate trials where the SDP was tight, but (5) was not satisfied, implying a suboptimal local maximum.

Towards the left of Fig. 1a, with small σ and the $(\mathbf{M}_1, \ldots, \mathbf{M}_k)$ all being very close to commuting, 100% of experiments return tight rank-one SDP solutions. Interestingly, there appears to be a sharp cut-off point where this behavior ends, and the SDP relaxation is not tight in a small percentage of cases. While the large majority of trials still admit a tight convex relaxation, these results empirically corroborate the sufficient conditions derived in Theorem 4.5 and Corollary 4.5.1.

Where the SDP is tight, Fig. 1 shows the StMM solver returns the globally optimal solution in more than 95% of the problem instances. Indicated by the "$\triangle$" markers, the remaining cases can only be certified as stationary points, implying a local maximum was found. Indeed, we observe a correspondence between trials with both large objective value gap and distance of the candidate solution to the globally optimal solution returned by the SDP.

HPPCA We repeat the experiments just described for $\mathbf{M}_i$ generated by the model in (6) for $d = 50$, $\boldsymbol{\lambda} = [4, 3.25, 2.5, 1.75, 1]$, and $L = 2$ noise groups with variances $\boldsymbol{v} = [1, 4]$. We draw 100 random models with a different generative $\boldsymbol{U}$ for sample sizes $\boldsymbol{n} = [n_1, 4n_1]$, where we sweep through increasing values of n_1 on the horizontal axis in Fig. 2c, drawing 100 different random data initializations. For each experiment, we normalize the $\mathbf{M}_i$ by the maximum of their spectral norms, and then record the results obtained from the SDP and StMM solvers with respect to the computed maximum commuting distance of the $\mathbf{M}_i$ in Fig. 2. We run StMM for a maximum of 10,000 iterations, and record whether the SDP was tight and the global optimality certification of each StMM run.

Proposition 4.6 suggests that, even with poor SNR like in this example, as the number of data samples increases, the $\mathbf{M}_i$ should concentrate to be nearly commuting. This is indeed what we observe: as the number

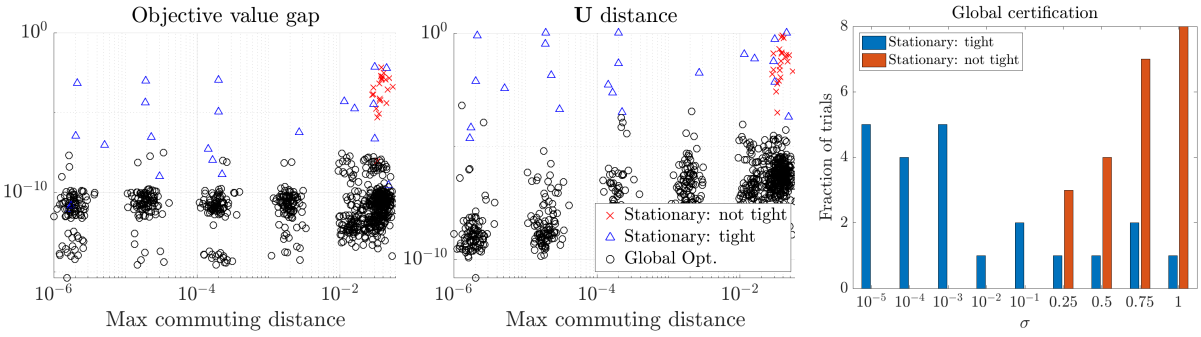

Figure 1: Numerical simulations for synthetic CJD matrices for $d = 10, k = 3$ with increasing σ.

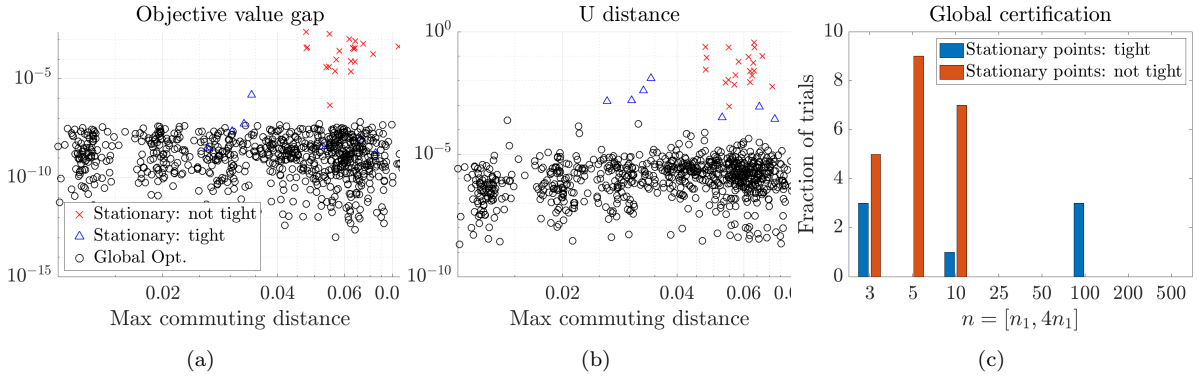

Figure 2: Numerical simulations for $\mathbf{M}_i$ generated by the HPPCA model in (6) for $d = 50$, $k = 5$, noise variances $[1, 4]$, and $\boldsymbol{\lambda} = [4, 3.25, 2.5, 1.75, 1]$ with increasing samples n.

of samples increases, the maximum commuting distance on the horizontal axis of Figs. 2a and 2b decreases. In this nearly-commuting regime, the SDP obtains tight rank-one $\boldsymbol{X}_i$ in 100% of the trials, and interestingly, all of the StMM runs attain the global maximum, suggesting a seemingly benign nonconvex landscape. In contrast, we observed several trials in the low-sample setting where the SDP failed to be tight and the dual certificate was null. Also within this regime, several trials of the StMM solver find suboptimal local maxima.

5.3 Computation time

Fig. 3 compares the scalability of our SDP relaxation in (SDP-P) to the StMM solver with the global certificate check in (5) for synthetically generated HPPCA problems of varying data dimension. We measure the median computation time across 10 independent trials of both algorithms. The experiment strongly demonstrates the computational superiority of the first-order method with our certificate compared to the full SDP. StMM+Certificate scales nearly 60 times better in computation time for the largest dimension with $k = 3$ and 15 times for $k = 10$, while offering a crucial theoretical guarantee to a nonconvex problem that may contain spurious local maxima. Thus, we can solve the nonconvex problem posed in (1) using any choice of solver on the Stiefel manifold and perform a fast check of its terminal output for global optimality.

6 Future Work & Conclusion

In this work, we proposed a novel SDP relaxation for the sums of heterogeneous quadratics problem, from which we derived a global optimality certificate to check a local solution of a nonconvex program. Our other major contribution proved a continuity result showing sufficient conditions when the relaxation has the ROP

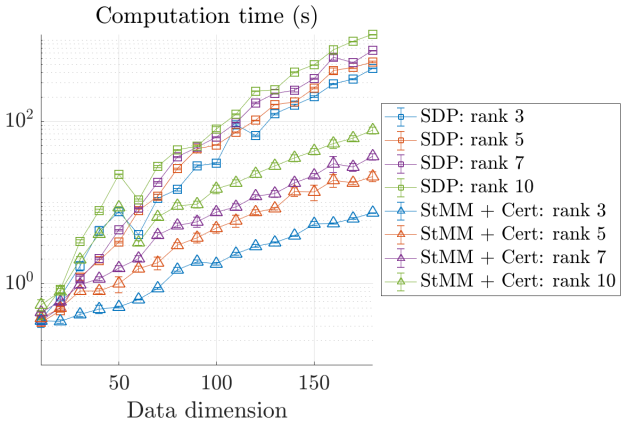

Figure 3: Computation time of (SDP-P) versus StMM for 2000 iterations with global certificate check (5) for HPPCA problems as the data dimension varies. We use $\boldsymbol{v} = [1, 4]$, and $\boldsymbol{n} = [100, 400]$ and make $\boldsymbol{\lambda}$ a k-length vector with entries equally spaced in the interval $[1, 4]$. Markers indicate the median computation time taken over 10 trials, and error bars show the standard deviation.

and provided both theoretical and empirical support that one motivating signal processing application–the HPPCA problem–possesses a tight relaxation in many instances.

While the global certificate check we propose scales well compared to solving the full SDP, the LMI feasiblity program still requires forming and factoring $d \times d$ size matrices, requiring storage of $\mathcal{O}(d^2)$ elements. One exciting possibility is to apply recent works like [45] to our problem, which use randomized algorithms to reduce the storage and arithmetic costs for scalable semidefinite programming. Further, it remains a strong interest to prove a sufficient analytical certificate, in addition to proving more general sufficient conditions on the $\mathbf{M}_i$ to guarantee the ROP.

While we hope the work herein has a positive impact in HPPCA applications like air quality monitoring [24] or medical imaging, we acknowledge the potential for dimensionality reduction algorithms to yield disparate reconstruction errors on populations within a dataset, such as PCA on labeled faces in the wild data set (LFW), which returns higher reconstruction error for women than men even with equal population ratios in the dataset [36]; also see [39].

Acknowledgments

The authors would like to thank Nicolas Boumal for his helpful discussions, references, and notes relating to dual certificates of low-rank SDP's and manifold optimization. They would also like to thank David Hong and Jeffrey Fessler for their feedback on this paper and their discussions relating to heteroscedastic PPCA.

References

[1] Traian E. Abrudan, Jan Eriksson, and Visa Koivunen. Steepest descent algorithms for optimization under unitary matrix constraint. *IEEE Transactions on Signal Processing*, 56(3):1134–1147, 2008. doi: 10.1109/TSP.2007.908999.

[2] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008. ISBN 978-0-691-13298-3.

[3] Bijan Afsari. Sensitivity analysis for the problem of matrix joint diagonalization. *SIAM Journal on Matrix Analysis and Applications*, 30(3):1148–1171, 2008. doi: 10.1137/060655997. URL https://doi.org/10.1137/060655997.

[4] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on Modern Convex Optimization*. Society for Industrial and Applied Mathematics, 2001. doi: 10.1137/1.9780898718829. URL https://epubs.siam.org/doi/abs/10.1137/1.9780898718829.

[5] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on modern convex optimization*. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Programming Society (MPS), Philadelphia, PA, 2001. ISBN 0-89871-491-5. doi: 10.1137/1.9780898718829. URL https://doi-org.proxy.lib.uiowa.edu/10.1137/1.9780898718829. Analysis, algorithms, and engineering applications.

[6] O. A. Berezovskyi. On the lower bound for a quadratic problem on the Stiefel manifold. *Cybernetics and Sys. Anal.*, 44(5):709–715, September 2008. ISSN 1060-0396. doi: 10.1007/s10559-008-9038-4. URL https://doi.org/10.1007/s10559-008-9038-4.

[7] Marianna Bolla, György Michaletzky, Gábor Tusnády, and Margit Ziermann. Extrema of sums of heterogeneous quadratic forms. *Linear Algebra and its Applications*, 269(1):331 – 365, 1998. ISSN 0024-3795. doi: https://doi.org/10.1016/S0024-3795(97)00230-9. URL http://www.sciencedirect.com/science/article/pii/S0024379597002309.

[8] F. Bouchard, J. Malick, and M. Congedo. Riemannian optimization and approximate joint diagonalization for blind source separation. *IEEE Transactions on Signal Processing*, 66(8):2041–2054, 2018. doi: 10.1109/TSP.2018.2795539.

[9] Florent Bouchard, Bijan Afsari, Jérôme Malick, and Marco Congedo. Approximate joint diagonalization with Riemannian optimization on the general linear group. *SIAM Journal on Matrix Analysis and Applications*, 41, 01 2019. doi: 10.1137/18M1232838.

[10] Nicolas Boumal, Vlad Voroninski, and Afonso Bandeira. The non-convex Burer-Monteiro approach works on smooth semidefinite programs. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. URL https://proceedings.neurips.cc/paper/2016/file/3de2334a314a7a72721f1f74a6cb4cee-Paper.pdf.

[11] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. doi: 10.1017/CBO9780511804441.

[12] Arnaud Breloy, Guillaume Ginolhac, Frédéric Pascal, and Philippe Forster. Clutter subspace estimation in low rank heterogeneous noise context. *IEEE Transactions on Signal Processing*, 63(9):2173–2182, 2015. doi: 10.1109/TSP.2015.2403284.

[13] Arnaud Breloy, Guillaume Ginolhac, Frédéric Pascal, and Philippe Forster. Robust covariance matrix estimation in heterogeneous low rank context. *IEEE Transactions on Signal Processing*, 64(22):5794–5806, 2016. doi: 10.1109/TSP.2016.2599494.

[14] Arnaud Breloy, Sandeep Kumar, Ying Sun, and Daniel P. Palomar. Majorization-minimization on the Stiefel manifold with application to robust sparse PCA. *IEEE Transactions on Signal Processing*, 69:1507–1520, 2021. doi: 10.1109/TSP.2021.3058442.

[15] Roger W Brockett. Least squares matching problems. *Linear algebra and its applications*, 122:761–777, 1989.

[16] Samuel Burer and Renato DC Monteiro. A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization. *Mathematical Programming*, 95(2):329–357, 2003.

[17] Samuel Burer, Kurt M. Anstreicher, and Mirjam Dür. The difference between 5×5 doubly nonnegative and completely positive matrices. *Linear Algebra Appl.*, 431(9):1539–1552, 2009. ISSN 0024-3795. doi: 10.1016/j.laa.2009.05.021.

[18] Alan Edelman, Tomás A Arias, and Steven T Smith. The geometry of algorithms with orthogonality constraints. *SIAM journal on Matrix Analysis and Applications*, 20(2):303–353, 1998.

[19] Ky Fan. On a theorem of Weyl concerning eigenvalues of linear transformations. *Proceedings of the National Academy of Sciences*, 35(11):652–655, 1949. doi: 10.1073/pnas.35.11.652. URL `https://www.pnas.org/doi/abs/10.1073/pnas.35.11.652`.

[20] P. A. Fillmore and J. P. Williams. Some convexity theorems for matrices. *Glasgow Math. J.*, 12:110–117, 1971. ISSN 0017-0895. doi: 10.1017/S0017089500001221.

[21] Dan Garber and Ron Fisher. Efficient algorithms for high-dimensional convex subspace optimization via strict complementarity. *arXiv preprint arXiv:2202.04020*, 2022.

[22] Klaus Glashoff and Michael M Bronstein. Almost-commuting matrices are almost jointly diagonalizable. *arXiv preprint arXiv:1305.2135*, 2013.

[23] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. `http://cvxr.com/cvx`, March 2014.

[24] David Ke Hong, Kyle Gilman, Laura Balzano, and Jeffrey A. Fessler. HePPCAT: Probabilistic PCA for data with heteroscedastic noise. *IEEE Transactions on Signal Processing*, 69:4819–4834, 2021.

[25] Yongwei Huang and Daniel P Palomar. Rank-constrained separable semidefinite programming with applications to optimal beamforming. *IEEE Transactions on Signal Processing*, 58(2):664–678, 2009.

[26] M. Kleinsteuber and H. Shen. Uniqueness analysis of non-unitary matrix joint diagonalization. *IEEE Transactions on Signal Processing*, 61(7):1786–1796, 2013. doi: 10.1109/TSP.2013.2242065.

[27] Vladimir Koltchinskii and Karim Lounici. Concentration inequalities and moment bounds for sample covariance operators. *Bernoulli*, 23(1):110–133, 2017. ISSN 13507265. doi: 10.3150/15-BEJ730.

[28] Terry A Loring and Adam PW Sørensen. Almost commuting self-adjoint matrices: the real and self-dual cases. *Reviews in Mathematical Physics*, 28(07):1650017, 2016.

[29] Karim Lounici. High-dimensional covariance matrix estimation with missing observations. *Bernoulli*, 20(3):1029 – 1058, 2014. doi: 10.3150/12-BEJ487. URL `https://doi.org/10.3150/12-BEJ487`.

[30] Zhi-Quan Luo, Tsung-Hui Chang, DP Palomar, and YC Eldar. SDP relaxation of homogeneous quadratic optimization: approximation. *Convex Optimization in Signal Processing and Communications*, page 117, 2010.

[31] Michael L. Overton and Robert S. Womersley. On the sum of the largest eigenvalues of a symmetric matrix. *SIAM J. Matrix Anal. Appl.*, 13(1):41–45, 1992. ISSN 0895-4798. doi: 10.1137/0613006.

[32] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Math. Oper. Res.*, 23:339–358, 1998.

[33] Dinh-Tuan Pham and Marco Congedo. Least square joint diagonalization of matrices under an intrinsic scale constraint. In Tülay Adali, Christian Jutten, João Marcos Travassos Romano, and Allan Kardec Barros, editors, *ICA 2009 - 8th International Conference on Independent Component Analysis and Signal Separation*, volume 5441 of *Lecture Notes in Computer Science*, pages 298–305, Paraty, Brazil, February 2009. Springer. doi: 10.1007/978-3-642-00599-2_38. URL `https://hal.archives-ouvertes.fr/hal-00371941`.

[34] Thomas Pumir, Samy Jelassi, and Nicolas Boumal. Smoothed analysis of the low-rank approach for smooth semidefinite programs. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 31. Curran Associates, Inc., 2018. URL https://proceedings.neurips.cc/paper/2018/file/a1d50185e7426cbb0acad1e6ca74b9aa-Paper.pdf.

[35] Tamás Rapcsák. On minimization on Stiefel manifolds. *European Journal of Operational Research*, 143 (2):365–376, 2002.

[36] Samira Samadi, Uthaipon Tantipongpipat, Jamie H Morgenstern, Mohit Singh, and Santosh Vempala. The price of fair pca: One extra dimension. *Advances in neural information processing systems*, 31, 2018.

[37] Xizhi Shi. *Joint Approximate Diagonalization Method*, pages 175–204. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011. ISBN 978-3-642-11347-5. doi: 10.1007/978-3-642-11347-5_8. URL https://doi.org/10.1007/978-3-642-11347-5_8.

[38] Ying Sun, Arnaud Breloy, Prabhu Babu, Daniel P. Palomar, Frédéric Pascal, and Guillaume Ginolhac. Low-complexity algorithms for low rank clutter parameters estimation in radar systems. *IEEE Transactions on Signal Processing*, 64(8):1986–1998, 2016. doi: 10.1109/TSP.2015.2512535.

[39] Uthaipon Tantipongpipat, Samira Samadi, Mohit Singh, Jamie H Morgenstern, and Santosh Vempala. Multi-criteria dimensionality reduction with applications to fairness. *Advances in neural information processing systems*, 32, 2019.

[40] Fabian J. Theis, Thomas P. Cason, and P. A. Absil. Soft dimension reduction for ICA by joint diagonalization on the Stiefel manifold. In Tülay Adali, Christian Jutten, João Marcos Travassos Romano, and Allan Kardec Barros, editors, *Independent Component Analysis and Signal Separation*, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[41] Vincent Q Vu, Juhee Cho, Jing Lei, and Karl Rohe. Fantope projection and selection: A near-optimal convex relaxation of sparse PCA. In *Advances in neural information processing systems*, pages 2670–2678, 2013.

[42] Hermann Weyl. Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung). *Mathematische Annalen*, 71 (4):441–479, 1912. doi: 10.1007/BF01456804. URL https://doi.org/10.1007/BF01456804.

[43] Joong-Ho Won, Teng Zhang, and Hua Zhou. Orthogonal trace-sum maximization: Tightness of the semidefinite relaxation and guarantee of locally optimal solutions. *arXiv preprint arXiv:2110.05701*, 2021.

[44] Joong-Ho Won, Hua Zhou, and Kenneth Lange. Orthogonal trace-sum maximization: Applications, local algorithms, and global optimality. *SIAM Journal on Matrix Analysis and Applications*, 42(2): 859–882, 2021.

[45] Alp Yurtsever, Joel A Tropp, Olivier Fercoq, Madeleine Udell, and Volkan Cevher. Scalable semidefinite programming. *SIAM Journal on Mathematics of Data Science*, 3(1):171–200, 2021.

A Related Work

In this extended related work discussion, we first describe works very closely related to our problem in (1), and then describe works more generally related to SDP relaxations of rank or orthogonality constrained problems.

Bolla et al. [7], Rapcsák [35], and Berezovskyi [6] also previously investigated the sum of heterogeneous quadratics in (1). The work in [7] only studied the structure of this problem for some special cases where all of the matrices $\mathbf{M}_i$ were either equal, diagonal, or commuting. Rapcsák [35] derived sufficient second-order global optimality conditions for the Hessian of the Lagrangian. However, these conditions are generally difficult to check in practice since the Hessian scales quickly with sizes $kd \times kd$ and, in general, is not PSD over the entire space—hence requiring verification of its positive semidefiniteness restricted to vectors on the manifold tangent space. Berezovskyi [6] proved that the dual Lagrangian bound is exact for the case of Boolean problem variables.

Works such as Huang and Palomar [25] and Pataki [32] consider a very similar problem to (2), but without the constraint in (4), making their SDP a rank-constrained separable SDP; see also Luo et al. [30, Section 4.3]. Pataki studied upper bounds on the rank of optimal solutions of general SDPs, but in the case of (SDP-P), since our problem introduces the additional constraint summing the $\boldsymbol{X}_i$, Pataki's bounds do not guarantee rank-1, or even low-rank, optimal solutions.

Our problem also has interesting connections to the well-studied problem in the literature of approximate joint diagonalization (AJD), which is often applied to blind source separation or independent component analysis (ICA) problems [40; 8; 26; 3; 37]. Given a set of symmetric PSD matrices that represent second order data statistics, one seeks the matrix, usually constrained to lie in the set of orthogonal or invertible matrices, that jointly diagonalizes the set of matrices optimally, albeit approximately. When all matrices in the set commute, the diagonalizer is simply the shared eigenspace, but often in practice, due to noise, finite samples, or numerical errors, the set does not commute and can only be approximately diagonalized.

Expanding our matrix variable $\boldsymbol{U} \in \mathbb{R}^{d\times k}$ to a full basis $\boldsymbol{U} \in \mathbb{R}^{d\times d}$, problem (2) is equivalent to

$$\min_{\boldsymbol{U}\in\mathbb{R}^{d\times d}:\boldsymbol{U}'\boldsymbol{U}=\boldsymbol{U}\boldsymbol{U}'=\mathbf{I}} \sum_{\ell=1}^{L} \frac{1}{2}\|\boldsymbol{U}'\mathbf{A}_\ell\boldsymbol{U} - \boldsymbol{W}_\ell\|_F^2 + C, \tag{8}$$

where $\boldsymbol{W}_\ell = \operatorname{diag}(w_{\ell,1}, \ldots, w_{\ell,k}, 0, \ldots, 0) \succeq 0$, and C is a constant. The objective functions in Pham and Congedo [33, Equation 4] and Bouchard et al. [9, Equation 8], given a fixed diagonal matrix, bear great similarity to ours, with the difference being that the diagonal matrix $\boldsymbol{W}_\ell$ above is not a function of $\boldsymbol{U}$, making it a distinct problem from AJD. However, if the diagonal matrix is fixed, then AJD simplifies to (8). Accordingly, problems (2) and (8) can be loosely interpreted as finding the $\boldsymbol{U}$ that best approximately jointly diagonalizes the data second-order statistics $\boldsymbol{A}_\ell$ to each $\boldsymbol{W}_\ell$. The AJD literature often employs Riemannian manifold optimization to solve the chosen objective function iteratively. To the best of our knowledge, no work has yet shown an analytical solution beyond the case when all the matrices commute nor proven global optimality criteria for these nonconvex programs.

The works in [10; 34] prove global convergence of nonconvex Burer–Monteiro factorization approaches to solve low-rank semidefinite programs, but these are distinct from our problem in which the columns of the orthonormal basis are constrained together in (4). Other works have studied optimizers to the nonconvex problem, like those in [14; 13; 38; 12], using minorize-maximize or Riemannian gradient ascent algorithms. While efficient and scalable, these methods do not have global optimality guarantees beyond proof of convergence to a critical point. Recent works have also studied convex relaxations of PCA and other low-rank subspace problems that bound the eigenvalues of a single matrix [41; 39; 44], rather than the sum of multiple matrices as in our setting. Won et al. [44, 43] study the SDP relaxation of maximizing the sum of traces of matrix quadratic forms on a product of Stiefel manifolds using the Fantope and propose a global optimality certificate. We emphasize their problem pertains to optimizing a trace sum over multiple orthonormal bases, each on a different Stiefel manifold, whereas our problem separates over the columns of a single basis on the Stiefel and is completely distinct from theirs. Extending the theory of the dual certificate from Fan [19] to the orthogonal trace maximization problem, they propose a simple way to test the global optimality of a given stationary point from an iterative solver of the nonconvex problem. Then in [43], the

same authors prove that for an additive noise model with small noise, their SDP relaxation is tight, and the solution of the nonconvex problem is globally optimal with high probability.

Many works study SDP relaxations of low-rank problems without Fantope constraints, a few of which we highlight here. The works in [A2; A3; A4] study SDP relaxations of Burer-Monteiro factorizations for optimization problems with multiple linear constraints. From the local properties of candidate solutions, they devise dual certificates to check for global optimality. [A5; A6] show for low-rank SDPs with rank-r and m linear constraints, no spurious local minima exist if $(r+1)(r+2)/2 > m+1$; [A6] also proves convergence of the nonconvex Burer-Monteiro factorization to the optimal SDP solution, with [A7] strengthening this result, showing such algorithms converge provably in polynomial time, given that $r \gtrsim \sqrt{2(1+\eta)m}$ for any fixed constant $\eta > 0$.

Similar to our work, the authors in [A1] seek to recover multiple rank-one matrices, in their case for the overcomplete ICA problem. They solve separate SDP relaxations for each atom of the dictionary, using a deflation method to find the atoms in succession. In contrast, our work estimates all of the rank-one matrices simultaneously, and requires that their first principal components form an orthonormal basis, whereas the dictionary atoms in ICA are only constrained to be unit-norm.

B Proofs of Section 2

B.1 Derivation of (SDP-D)

The Lagrangian function of (SDP-P), with dual variables $\boldsymbol{\nu} \in \mathbb{R}^k, \boldsymbol{Y} \in \mathbb{S}^d_+, \boldsymbol{Z}_i \in \mathbb{S}^d_+$ for $i = 1, \ldots, k$, is

$$\begin{aligned} &\mathcal{L}(\boldsymbol{X}_i, \boldsymbol{\nu}, \boldsymbol{Y}, \boldsymbol{Z}_i) = \\ &- \operatorname{tr}\left(\sum_{i=1}^k \mathbf{M}_i \boldsymbol{X}_i\right) - \sum_{i=1}^k \nu_i (1 - \operatorname{tr}(\boldsymbol{X}_i)) - \operatorname{tr}\left(\boldsymbol{Y}\left(\mathbf{I} - \sum_{i=1}^k \boldsymbol{X}_i\right)\right) - \sum_{i=1}^k \operatorname{tr}(\boldsymbol{Z}_i \boldsymbol{X}_i), \end{aligned} \tag{9}$$

for which the dual function is

$$g(\boldsymbol{Y}, \boldsymbol{Z}_i, \boldsymbol{\nu}) = \inf_{\boldsymbol{X}_i} \mathcal{L}(\boldsymbol{X}_i, \boldsymbol{\nu}, \boldsymbol{Y}, \boldsymbol{Z}_i) = \begin{cases} -\operatorname{tr}(\boldsymbol{Y}) - \sum_{i=1}^k \boldsymbol{\nu}_i & \text{s.t. } \boldsymbol{Y} = \mathbf{M}_i + \boldsymbol{Z}_i - \nu_i \mathbf{I} \quad \forall i \in [k] \\ -\infty & \text{otherwise.} \end{cases} \tag{10}$$

This yields the dual problem

$$\max_{\boldsymbol{Y}, \boldsymbol{Z}_i, \boldsymbol{\nu}} g(\boldsymbol{Y}, \boldsymbol{Z}_i, \boldsymbol{\nu}) \quad \text{s.t. } \boldsymbol{Y} \succeq 0, \quad \boldsymbol{Z}_i \succeq 0, \quad \boldsymbol{Y} = \mathbf{M}_i + \boldsymbol{Z}_i - \nu_i \mathbf{I}, \quad \forall i \in [k]. \tag{11}$$

Lemma B.1 (Restatement of Lemma 2.2)**.** *Strong duality holds for the SDP relaxation with primal* (SDP-P) *and dual* (SDP-D)*.*

Proof of Lemma B.1/Lemma 2.2. The problem is convex and satisfies Slater's condition, see Lemma B.2. Specifically, at optimality we have $\langle \mathbf{I} - (\sum_{i=1}^k \bar{\boldsymbol{X}}_i), \bar{\boldsymbol{Y}} \rangle = 0$ and therefore $\operatorname{tr}(\bar{\boldsymbol{Y}}) = \langle \bar{\boldsymbol{Y}}, \sum_{i=1}^k \bar{\boldsymbol{X}}_i \rangle$. Then

$$d^* = -\left\langle \sum_{i=1}^k \mathbf{M}_i + \bar{\boldsymbol{Z}}_i + \bar{\nu}_i \mathbf{I}, \bar{\boldsymbol{X}}_i \right\rangle + \sum_{i=1}^k \bar{\nu}_i = -\operatorname{tr}\left(\sum_{i=1}^k \mathbf{M}_i \bar{\boldsymbol{X}}_i\right),$$

since $\langle \bar{\boldsymbol{Z}}_i, \bar{\boldsymbol{X}}_i \rangle = 0$ and $\sum_{i=1}^k \bar{\nu}_i (1 - \operatorname{tr}(\bar{\boldsymbol{X}}_i)) = 0$. Thus, $p^* = d^*$. □

Lemma B.2. *The primal problem in* (SDP-P) *is strictly feasible.*

Proof. To be strictly feasible we must have $\boldsymbol{X}_i$, $i = 1, \ldots, k$ such that

$$0 \prec \sum_{i=1}^k \boldsymbol{X}_i \prec \mathbf{I}, \quad \operatorname{tr}(\boldsymbol{X}_i) = 1, \quad \boldsymbol{X}_i \succ 0, \quad i = 1, \ldots, k$$

Suppose $\boldsymbol{X}_i = \frac{1}{d}\mathbf{I}$ for all i. Then $\mathrm{tr}(\boldsymbol{X}_i) = 1$ and $\boldsymbol{X}_i \succ 0$ for all i, and $\sum_{i=1}^k \boldsymbol{X}_i = \frac{k}{d}\mathbf{I}$, satisfying $0 \prec \sum_{i=1}^k \boldsymbol{X}_i \prec \mathbf{I}$. □

Lemma B.3 (Restatement of Lemma 2.3)**.** *The solution to the SDP relaxation in* (SDP-P) *is the optimal solution to the original nonconvex problem in* (1) *(equivalently* (4)*) if and only if the optimal* $\boldsymbol{X}_i$ *have the rank-one property.*

Proof of Lemma B.3/Lemma 2.3. Since the problem in (SDP-P) has a larger constraint set than (1), any solution to (SDP-P) that satisfies the constraints of (1) is also a solution to this original nonconvex problem.

For the "if" direction assume that the optimal $\boldsymbol{X}_i$ for (SDP-P) have the rank-one property. Since $\mathrm{tr}(\boldsymbol{X}_i) = 1$ by definition of (SDP-P), when we decompose $\boldsymbol{X}_i = \boldsymbol{u}_i\boldsymbol{u}_i'$ we have $\boldsymbol{u}_i$ that are norm-1. In order for $\sum_{i=1}^k \boldsymbol{X}_i \preccurlyeq \mathbf{I}$, the $\boldsymbol{u}_i$ must be orthogonal. For the "only if" direction, assume that the solution to the SDP relaxation in (SDP-P) is the optimal solution to the original nonconvex problem in (1). The constraint in (1) that the columns of $\boldsymbol{U}$ are orthonormal implies that $\boldsymbol{X}_i$ must have the rank-one property. □

Lemma B.4 (Restatement of Lemma 2.4)**.** *Assume* $\mathbf{M}_i$ *are PSD. Then the optimal* $\nu_i \geq 0$.

Proof of Lemma B.4/Lemma 2.4. Another equivalent formulation of the dual problem eliminates $\boldsymbol{Y}$:

$$d^* = \min_{\nu_i, \boldsymbol{Z}_i} \frac{1}{k}\sum_{i=1}^k \{\mathrm{tr}(\boldsymbol{Z}_i + \mathbf{M}_i) - (d-k)\nu_i\} \tag{12}$$

$$\begin{aligned} \text{s.t. } & \mathbf{M}_i + \boldsymbol{Z}_i \succcurlyeq \nu_i\mathbf{I} \quad \forall i = 1, \ldots, k \\ & \mathbf{M}_i + \boldsymbol{Z}_i - \nu_i\mathbf{I} = \mathbf{M}_j + \boldsymbol{Z}_j - \nu_j\mathbf{I} \quad \forall i, j = 1, \ldots, k \\ & \boldsymbol{Z}_i \succcurlyeq 0 \end{aligned} \tag{13}$$

Suppose otherwise that $\nu_i < 0$. Then $d^* > 0$ since $\mathrm{tr}(\boldsymbol{Z}_i + \mathbf{M}_i) \geq 0$. But if instead we let $\nu_i = 0$, $\boldsymbol{Z}_i = \mathbf{M}_i = 0$, we have a feasible solution where $d^* = 0$. □

Lemma B.5. *Suppose* $\boldsymbol{X}_i$ *for* $i = 1, \ldots, k$ *are each trace 1 and each has* $\lambda_1(\boldsymbol{X}_i) = 1$*, and therefore each* $\boldsymbol{X}_i$ *is rank 1. We decompose* $\boldsymbol{X}_i = \boldsymbol{u}_i\boldsymbol{u}_i'$ *and note that* $\boldsymbol{u}_i$ *are norm-1. Then* $\sum_{i=1}^k \boldsymbol{X}_i$ *satisfies*

$$0 \preccurlyeq \sum_{i=1}^k \boldsymbol{X}_i \preccurlyeq \mathbf{I}$$

if and only if

$$\boldsymbol{u}_i'\boldsymbol{u}_j = 0 \quad \forall i \neq j \; .$$

Proof. Forward direction: Suppose $\boldsymbol{X} = \sum_{i=1}^k \boldsymbol{X}_i$ has eigenvalues in $[0, 1]$ and $\mathrm{tr}(\boldsymbol{X}) = k$. Since $\mathrm{rank}(\boldsymbol{X}) \leq k$ by subadditivity of rank, this implies both that $\boldsymbol{X}$ is rank-k and its eigenvalues are either zero or one. Note then that

$$\mathrm{tr}(\boldsymbol{X}\boldsymbol{X}') = k = \mathrm{tr}\left((\sum_i \boldsymbol{u}_i\boldsymbol{u}_i')(\sum_i \boldsymbol{u}_i\boldsymbol{u}_i')\right) = \sum_i (\boldsymbol{u}_i'\boldsymbol{u}_i)^2 + \mathrm{tr}\left(2\sum_{i\neq j}(\boldsymbol{u}_i'\boldsymbol{u}_j)^2\right) \; .$$

Since $\boldsymbol{u}_i$ are norm-1 then the sum $\sum_i (\boldsymbol{u}_i'\boldsymbol{u}_i)^2 = k$. This means

$$\mathrm{tr}\left(2\sum_{i\neq j}(\boldsymbol{u}_i'\boldsymbol{u}_j)^2\right) = 0 \; ,$$

which is true if and only if $\boldsymbol{u}_i'\boldsymbol{u}_j = 0$.

The backward direction is direct because when $\boldsymbol{u}_i'\boldsymbol{u}_j = 0$ for $i \neq j$, $\sum_{i=1}^k \boldsymbol{u}_i\boldsymbol{u}_i'$ is the singular value decomposition of $\boldsymbol{X}$ with k eigenvalues equal to one. □

Lemma B.6 (Restatement of Lemma 2.5)**.** *If the optimal dual variables $\boldsymbol{Z}_i$ for $i = 1, \ldots, k$ are each rank $d-1$, the optimal solution variables $\boldsymbol{X}_i$ have the rank-one property.*

Proof of Lemma B.6/Lemma 2.5. Suppose $\boldsymbol{Z}_i$ is rank $d-1$. By complementarity at optimality, we have $\boldsymbol{Z}_i \boldsymbol{X}_i = 0 \quad \forall i$, which means $\boldsymbol{X}_i$ lies in the nullspace of $\boldsymbol{Z}_i$, which has dimension 1, so each $\boldsymbol{X}_i$ is rank-1. By primal feasibility, $\mathrm{tr}(\boldsymbol{X}_i) = 1$, so $\lambda_1(\boldsymbol{X}_i) = 1 \quad \forall i = 1, \ldots, k$. By Lemma B.5, the optimal solution is an orthogonal projection matrix, and the optimal $\boldsymbol{X}_i$ are orthogonal. □

C Counterexample for Convex-Hull Result

By construction, the feasible set of (SDP-P) is a convex relaxation of the set

$$\{(\boldsymbol{u}_1 \boldsymbol{u}_1', \ldots, \boldsymbol{u}_k \boldsymbol{u}_k') : \boldsymbol{U}'\boldsymbol{U} = \mathbf{I}\}. \tag{14}$$

Given its relationship with the Fantope, a natural question is whether our relaxation captures the convex hull of (14), which would guarantee that our SDP relaxation is always exact. We prove here that this is not the case. Even so, there might exist sufficient conditions on $(\mathbf{M}_1, \ldots, \mathbf{M}_k)$ guaranteeing that the relaxation is exact. We do not explore such sufficient conditions in this subsection.

So let us prove formally that the feasible set of (SDP-P) in general does not capture the convex hull of (14). Specifically, we claim that, for $d = 4$ and $k = 2$, the matrix $\boldsymbol{X} = \boldsymbol{X}_1 + \boldsymbol{X}_2$ given by

$$\boldsymbol{X}_1 := \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and

$$\boldsymbol{X}_2 := \frac{1}{12} \begin{pmatrix} 3 & 1 & 3 & 1 \\ 1 & 3 & 1 & 3 \\ 3 & 1 & 3 & 1 \\ 1 & 3 & 1 & 3 \end{pmatrix}$$

cannot be a strict convex combination of feasible points $\boldsymbol{X}^{(j)} = \boldsymbol{X}_1^{(j)} + \boldsymbol{X}_2^{(j)}$ for some $j = 1, \ldots, J$ such that every $\boldsymbol{X}_i^{(j)}$ is rank-1. Said differently, $(\boldsymbol{X}_1, \boldsymbol{X}_2)$ cannot be a strict convex combination of elements of (14). Note that $\mathrm{rank}(\boldsymbol{X}_1) = \mathrm{rank}(\boldsymbol{X}_2) = 2$, so that $(\boldsymbol{X}_1, \boldsymbol{X}_2)$ itself is not an element of (14). In addition, it is easy to verify that $\mathrm{rank}(\boldsymbol{X}) = 4$ and $\lambda_{\max}[\boldsymbol{X}] = 1$. Our argument is based on the following proposition, whose contrapositive states that $(\boldsymbol{X}_1, \boldsymbol{X}_2)$ cannot be a strict convex combination because $\mathrm{rank}(\boldsymbol{X}) = 4$.

Proposition C.1. *Let $d \geq k = 2$ be given. Suppose $\boldsymbol{X} = \boldsymbol{X}_1 + \boldsymbol{X}_2$ is feasible for* (SDP-P) *such that:*

- *$(\boldsymbol{X}_1, \boldsymbol{X}_2)$ is a strict convex combination of points in* (14)*, i.e., for some integer $J \geq 2$, there exist positive scalars $\lambda_1, \ldots, \lambda_J$ and Stiefel matrices*

  $$\boldsymbol{U}^{(j)} := \begin{pmatrix} \boldsymbol{u}_1^{(j)} & \boldsymbol{u}_2^{(j)} \end{pmatrix} \in \mathbb{R}^{d \times 2} \quad \forall\, j = 1, \ldots, J$$

  such that

  $$(\boldsymbol{X}_1, \boldsymbol{X}_2) = \sum_{j=1}^{J} \lambda_j \left(\boldsymbol{u}_1^{(j)} (\boldsymbol{u}_1^{(j)})', \boldsymbol{u}_2^{(j)} (\boldsymbol{u}_2^{(j)})' \right), \qquad \sum_{j=1}^{J} \lambda_j = 1;$$

- *$\mathrm{rank}(\boldsymbol{X}_1) = \mathrm{rank}(\boldsymbol{X}_2) = 2$;*
- *$\lambda_{\max}[\boldsymbol{X}] = 1$.*

Then $\mathrm{rank}(\boldsymbol{X}) \leq 3$.

Proof. For each $i = 1, 2$, the equation

$$\boldsymbol{X}_i = \sum_{j=1}^{J} \lambda_j \boldsymbol{u}_i^{(j)} (\boldsymbol{u}_i^{(j)})' = \begin{pmatrix} \sqrt{\lambda_1}\boldsymbol{u}_i^{(1)} & \cdots & \sqrt{\lambda_J}\,\boldsymbol{u}_i^{(J)} \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1}\boldsymbol{u}_i^{(1)} & \cdots & \sqrt{\lambda_J}\,\boldsymbol{u}_i^{(J)} \end{pmatrix}' \tag{15}$$

ensures $\text{Range}(\boldsymbol{X}_i) = \text{Span}\{\boldsymbol{u}_i^{(1)}, \ldots, \boldsymbol{u}_i^{(J)}\}$; see Lemma 1 of [17] for example. Keeping in mind that $\text{rank}(\boldsymbol{X}_1) = \text{rank}(\boldsymbol{X}_2) = 2$ by assumption, we claim that, without loss of generality, we can reorder $j = 1, \ldots, J$ such that $\text{Range}(\boldsymbol{X}_i) = \text{Span}(\{\boldsymbol{u}_i^{(1)}, \boldsymbol{u}_i^{(2)}\})$ for both i simultaneously. In other words, we claim that each $\boldsymbol{X}_i$ "gets its rank" from the vectors $\boldsymbol{u}_i^{(1)}$ and $\boldsymbol{u}_i^{(2)}$.

If $J = 2$, the claim is obvious. If $J > 2$, first reorder the indices $\{1, \ldots, J\}$ such that $\text{Range}(\boldsymbol{X}_1) = \text{Span}(\{\boldsymbol{u}_1^{(1)}, \boldsymbol{u}_1^{(2)}\})$. If the claim holds for this new ordering, we are done. Otherwise, we can further reorder $\{3, \ldots, J\}$ such that

$$\begin{aligned} &\text{Range}(\boldsymbol{X}_1) = \text{Span}(\{\boldsymbol{u}_1^{(1)}, \boldsymbol{u}_1^{(2)}, \boldsymbol{u}_1^{(3)}\}) \quad \text{with} \quad \boldsymbol{u}_1^{(1)} \nparallel \boldsymbol{u}_1^{(2)} \\ &\text{Range}(\boldsymbol{X}_2) = \text{Span}(\{\boldsymbol{u}_2^{(1)}, \boldsymbol{u}_2^{(2)}, \boldsymbol{u}_2^{(3)}\}) \quad \text{with} \quad \boldsymbol{u}_2^{(1)} \parallel \boldsymbol{u}_2^{(2)} \text{ and } \boldsymbol{u}_2^{(1)} \nparallel \boldsymbol{u}_2^{(3)}. \end{aligned}$$

We now consider two exhaustive subcases. First, if $\boldsymbol{u}_1^{(1)} \nparallel \boldsymbol{u}_1^{(3)}$, then we see that $\boldsymbol{X}_1$ gets its rank from $\boldsymbol{u}_1^{(1)}$, $\boldsymbol{u}_1^{(3)}$ and $\boldsymbol{X}_2$ gets its rank from $\boldsymbol{u}_2^{(1)}$, $\boldsymbol{u}_2^{(3)}$. So by another reordering of $\{1, 2, 3\}$, the claim is proved. The second subcase $\boldsymbol{u}_1^{(2)} \nparallel \boldsymbol{u}_1^{(3)}$ is similar.

With the claim proven, define $\boldsymbol{W}_i := \text{Span}\{\boldsymbol{u}_i^{(1)}, \boldsymbol{u}_i^{(2)}\} = \text{Span}\{\boldsymbol{u}_i^{(1)}, \ldots, \boldsymbol{u}_i^{(J)}\}$ for both $i = 1, 2$. By adding the equations (15) for $i = 1, 2$, we also have

$$\text{rank}(\boldsymbol{X}) = \dim(\boldsymbol{W}_1 + \boldsymbol{W}_2) = \dim(\text{Span}\{\boldsymbol{u}_1^{(1)}, \boldsymbol{u}_1^{(2)}, \boldsymbol{u}_2^{(1)}, \boldsymbol{u}_2^{(2)}\}).$$

Next, let $\boldsymbol{v}$ be a maximum eigenvector of $\boldsymbol{X}$ with $\|\boldsymbol{v}\| = 1$ by definition. Also, for each j, define $\boldsymbol{V}_j := \text{Span}\{\boldsymbol{u}_1^{(j)}, \boldsymbol{u}_2^{(j)}\} = \text{Range}(\boldsymbol{U}^{(j)})$, and let

$$\alpha_j := (\boldsymbol{v}^T \boldsymbol{u}_1^{(j)})^2 + (\boldsymbol{v}^T \boldsymbol{u}_2^{(j)})^2 \leq 1$$

be the squared norm of the projection of $\boldsymbol{v}$ onto $\boldsymbol{V}_j$. We have

$$1 = \boldsymbol{v}^T \boldsymbol{X} \boldsymbol{v} = \sum_{j=1}^{J} \lambda_j \left((\boldsymbol{v}^T \boldsymbol{u}_1^{(j)})^2 + (\boldsymbol{v}^T \boldsymbol{u}_2^{(j)})^2 \right) = \sum_{j=1}^{J} \lambda_j \alpha_j.$$

Since each $\alpha_j \leq 1$ and since $\boldsymbol{\lambda}$ is a convex combination, it follows that $\alpha_j = 1$ for all j, which then implies $\boldsymbol{v} \in \boldsymbol{V}_j$ for all j, i.e., $\boldsymbol{v} \in \boldsymbol{V}_1 \cap \boldsymbol{V}_2$.

Finally, we have $\boldsymbol{W}_1 + \boldsymbol{W}_2 = \boldsymbol{V}_1 + \boldsymbol{V}_2$ because both Minkowski sums span the four vectors $\boldsymbol{u}_i^{(j)}$ for $i = 1, 2$ and $j = 1, 2$. Hence,

$$\begin{aligned} \text{rank}(\boldsymbol{X}) &= \dim(\boldsymbol{W}_1 + \boldsymbol{W}_2) \\ &= \dim(\boldsymbol{V}_1 + \boldsymbol{V}_2) = \dim(\boldsymbol{V}_1) + \dim(\boldsymbol{V}_2) - \dim(\boldsymbol{V}_1 \cap \boldsymbol{V}_2) \\ &\leq 2 + 2 - 1 = 3. \end{aligned}$$

where the inequality follows because $\boldsymbol{v} \in \boldsymbol{V}_1 \cap \boldsymbol{V}_2$. □

D Proof of Theorem 4.1

Lemma D.1. *Let $\mathcal{F}(\boldsymbol{U})$ denote the objective function with respect to $\boldsymbol{U}$ in (1) over $St(k,d)$. A point $\boldsymbol{U} \in St(k,d)$ is a local maximum of $\mathcal{F}$ if*

$$\boldsymbol{\Lambda} = \boldsymbol{U}' \sum_{i=1}^{k} \mathbf{M}_i \boldsymbol{U} \mathbf{E}_i \quad \textit{is symmetric, and}$$

$$-\sum_{i=1}^{k} \langle \dot{\boldsymbol{U}}, \mathbf{M}_i \dot{\boldsymbol{U}} \mathbf{E}_i \rangle + \langle \dot{\boldsymbol{U}} \boldsymbol{\Lambda}, \dot{\boldsymbol{U}} \rangle \geq 0 \quad \forall \dot{\boldsymbol{U}} \in \mathbb{R}^{d \times k} \textit{ such that } \dot{\boldsymbol{U}}'\boldsymbol{U} + \boldsymbol{U}'\dot{\boldsymbol{U}} = 0.$$

Proof. Taking $\bar{\mathcal{F}}$ to be the quadratic function in (1) over Euclidean space, the Euclidean gradient $\boldsymbol{\nabla}\bar{\mathcal{F}}(\boldsymbol{U}) = 2\sum_{i=1}^{k} \mathbf{M}_i \boldsymbol{U} \boldsymbol{E}_i = 2\begin{bmatrix}\mathbf{M}_1\boldsymbol{u}_1 & \cdots & \mathbf{M}_k\boldsymbol{u}_k\end{bmatrix}$, where $\mathbf{E}_i := \boldsymbol{e}_i\boldsymbol{e}_i' \in \mathbb{R}^{k\times k}$ and $\boldsymbol{e}_i$ is the i^{th} standard basis vector in $\mathbb{R}^k$. The Euclidean Hessian can also easily be derived as $\boldsymbol{\nabla}^2\bar{\mathcal{F}}(\boldsymbol{U})[\dot{\boldsymbol{U}}] = 2\sum_{i=1}^{k} \mathbf{M}_i \dot{\boldsymbol{U}} \boldsymbol{E}_i$. Restricting $\bar{\mathcal{F}}$ to the Stiefel manifold, let $\mathcal{F} := \bar{\mathcal{F}}|_{\mathrm{St}(k,d)}$. If $\boldsymbol{U} \in \mathrm{St}(k,d)$ is a local maximizer of (1), then

$$\boldsymbol{\nabla}\mathcal{F}(\boldsymbol{U}) = 0 \quad \text{and} \quad \boldsymbol{\nabla}^2\mathcal{F}(\boldsymbol{U}) \preceq 0, \tag{16}$$

where $\boldsymbol{\nabla}\mathcal{F}$ and $\boldsymbol{\nabla}^2\mathcal{F}$ denote the Riemannian gradient and Hessian of $\mathcal{F}$, respectively.

From [2], the gradient on the manifold for local maximizer $\boldsymbol{U}$ satisfies

$$\begin{aligned}
\boldsymbol{\nabla}\mathcal{F} = \boldsymbol{\nabla}\bar{\mathcal{F}} - \boldsymbol{U}\mathrm{sym}(\boldsymbol{U}'\boldsymbol{\nabla}\bar{\mathcal{F}}(\boldsymbol{U})) &= (\mathbf{I} - \boldsymbol{U}\boldsymbol{U}')\boldsymbol{\nabla}\bar{\mathcal{F}} + \boldsymbol{U}\mathrm{skew}(\boldsymbol{U}'\boldsymbol{\nabla}\bar{\mathcal{F}}), &(17)\\
&= 2(\mathbf{I} - \boldsymbol{U}\boldsymbol{U}')\sum_{i=1}^{k} \mathbf{M}_i\boldsymbol{U}\boldsymbol{E}_i + \boldsymbol{U}\sum_{i=1}^{k}[\boldsymbol{U}'\mathbf{M}_i\boldsymbol{U}, \boldsymbol{E}_i] &(18)\\
&= 0 &(19)
\end{aligned}$$

where $\mathrm{sym}(\boldsymbol{A}) = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}')$, $\mathrm{skew}(\boldsymbol{A}) = \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}')$, and $[\boldsymbol{A}, \boldsymbol{B}] = \boldsymbol{A}\boldsymbol{B} - \boldsymbol{B}\boldsymbol{A}$. We note the left and right expressions of the Riemannian gradient in (18) lie in the orthogonal complement of $\mathrm{Span}(\boldsymbol{U})$ and the $\mathrm{Span}(\boldsymbol{U})$, respectively, so $\boldsymbol{\nabla}\mathcal{F}$ vanishes if and only if $(\mathbf{I} - \boldsymbol{U}\boldsymbol{U}')\boldsymbol{\nabla}\bar{\mathcal{F}} = 0$, and $\sum_{i=1}^{k}[\boldsymbol{U}'\mathbf{M}_i\boldsymbol{U}, \boldsymbol{E}_i] = 0$, implying $\boldsymbol{U}'\boldsymbol{\nabla}\bar{\mathcal{F}} = \boldsymbol{\nabla}\bar{\mathcal{F}}'\boldsymbol{U}$. Letting $\boldsymbol{\Lambda} := \mathrm{sym}(\boldsymbol{U}'\boldsymbol{\nabla}\bar{\mathcal{F}})$, this also implies

$$\boldsymbol{U}\boldsymbol{\Lambda} = \boldsymbol{\nabla}\bar{\mathcal{F}} = \sum_{i=1}^{k} \mathbf{M}_i\boldsymbol{U}\boldsymbol{E}_i, \tag{20}$$

and multiplying both sides by $\boldsymbol{U}'$ yields the expression for $\boldsymbol{\Lambda}$, which is symmetric as shown above so we can drop the $\mathrm{sym}(\cdot)$ operator.

It can be shown the Riemannian Hessian is negative semidefinite if and only if

$$\langle \dot{\boldsymbol{U}}, \boldsymbol{\nabla}^2\bar{\mathcal{F}}(\boldsymbol{U})[\dot{\boldsymbol{U}}] - \dot{\boldsymbol{U}}\boldsymbol{\Lambda}\rangle \leq 0 \tag{21}$$

for all $\dot{\boldsymbol{U}} \in T_{\boldsymbol{U}}\mathrm{St}(d,k)$, where $T_{\boldsymbol{U}}\mathrm{St}(d,k)$ is the tangent space of the Stiefel manifold, i.e. the set $T_{\boldsymbol{U}}\mathrm{St}(d,k) = \{\dot{\boldsymbol{U}} \in \mathbb{R}^{d\times k} : \boldsymbol{U}'\dot{\boldsymbol{U}} + \dot{\boldsymbol{U}}'\boldsymbol{U} = 0\}$. Plugging in the expressions for $\boldsymbol{\Lambda}$ and the Hessian of $\bar{\mathcal{F}}$ yield the main result.

□

The following lemma is adapted from [7, Corollary 4.2]

Lemma D.2. *Let $\bar{\boldsymbol{U}} \in \mathbb{R}^{d\times k}$ be a local maximum of (1). Then $\bar{\boldsymbol{\Lambda}} = \bar{\boldsymbol{U}}'\sum_{i=1}^{k} \mathbf{M}_i\bar{\boldsymbol{U}}\mathbf{E}_i$ is positive semidefinite.*

Proof. Since $k < d$, there exists a unit vector $\boldsymbol{z}$ in the span of $\bar{\boldsymbol{U}}_\perp \in \mathbb{R}^{d\times d-k}$ where $\bar{\boldsymbol{U}}'\bar{\boldsymbol{U}}_\perp = 0$. Let $\boldsymbol{a} = [a_1, \ldots, a_k]' \in \mathbb{R}^k$ be an arbitrary nonzero vector. Let $\dot{\boldsymbol{U}} := \boldsymbol{z}\boldsymbol{a}'$, and let $\bar{\boldsymbol{\Lambda}} = \sum_{i=1}^{k} \bar{\boldsymbol{U}}'\mathbf{M}_i\bar{\boldsymbol{U}}\mathbf{E}_i$. Then

clearly $\bar{\boldsymbol{U}}'\dot{\boldsymbol{U}} = 0$, and $\bar{\boldsymbol{U}}'\dot{\boldsymbol{U}} + \dot{\boldsymbol{U}}'\bar{\boldsymbol{U}} = 0$, so the second-order stationary necessary condition in Lemma D.1 applies:

$$\boldsymbol{a}'\bar{\boldsymbol{\Lambda}}\boldsymbol{a} = \langle \dot{\boldsymbol{U}}\bar{\boldsymbol{\Lambda}}, \dot{\boldsymbol{U}}\rangle \geq \sum_{i=1}^{k} \langle \dot{\boldsymbol{U}}, \mathbf{M}_i\dot{\boldsymbol{U}}\mathbf{E}_i\rangle = \sum_{i=1}^{k} (a_i)^2 \boldsymbol{z}'\mathbf{M}_i\boldsymbol{z} \geq 0. \tag{22}$$

Therefore, since $\boldsymbol{a}'\bar{\boldsymbol{\Lambda}}\boldsymbol{a} \geq 0$ for arbitrary $\boldsymbol{a}$, $\bar{\boldsymbol{\Lambda}}$ is positive semidefinite. □

Theorem D.3 (Restatement of Theorem 4.1)**.** *Let $\bar{\boldsymbol{U}} \in \mathrm{St}(k,d)$ be a local maximizer to* (1)*, and let $\bar{\boldsymbol{\Lambda}} = \sum_{i=1}^{k} \bar{\boldsymbol{U}}'\mathbf{M}_i\bar{\boldsymbol{U}}\mathbf{E}_i$, where $\mathbf{E}_i \triangleq \boldsymbol{e}_i\boldsymbol{e}_i'$ and $\boldsymbol{e}_i$ is the i^{th} standard basis vector in $\mathbb{R}^k$. If there exist $\bar{\boldsymbol{\nu}} = [\bar{\nu}_1 \cdots \bar{\nu}_k] \in \mathbb{R}^k_+$ such that*

$$\begin{aligned} \bar{\boldsymbol{U}}(\bar{\boldsymbol{\Lambda}} - \mathbf{D}_{\bar{\boldsymbol{\nu}}})\bar{\boldsymbol{U}}' + \bar{\nu}_i\mathbf{I} - \mathbf{M}_i \succeq 0 \quad \forall i = 1,\ldots,k \\ \bar{\boldsymbol{\Lambda}} - \mathbf{D}_{\bar{\boldsymbol{\nu}}} \succeq 0, \end{aligned} \tag{23}$$

where $\mathbf{D}_{\bar{\boldsymbol{\nu}}} := diag(\bar{\nu}_1,\ldots,\bar{\nu}_k)$, then $\bar{\boldsymbol{U}}$ is an optimal solution to (SDP-P) *and a globally optimal solution to the original nonconvex problem* (1)*.*

Proof of Theorem D.3/ Theorem 4.1. By Lemma 2.2, primal and dual feasible solutions of (SDP-P) and (SDP-D), $\bar{\boldsymbol{X}}_i, \bar{\boldsymbol{Z}}_i, \bar{\boldsymbol{Y}}, \bar{\boldsymbol{\nu}}$, are simultaneously optimal if and only if they satisfy the following Karush-Kuhn Tucker (KKT) conditions [11], where the variables and constraints are indexed by $i \in [k]$:

$$\bar{\boldsymbol{X}}_i \succeq 0, \quad \sum_{i=1}^{k} \bar{\boldsymbol{X}}_i \preceq \mathbf{I}, \quad \mathrm{tr}(\bar{\boldsymbol{X}}_i) = 1 \tag{KKT-a}$$

$$\bar{\boldsymbol{Y}} = \mathbf{M}_i + \bar{\boldsymbol{Z}}_i - \bar{\nu}_i\mathbf{I}, \qquad \bar{\boldsymbol{Y}} \succeq 0 \tag{KKT-b}$$

$$\langle \mathbf{I} - \sum_{i=1}^{k} \bar{\boldsymbol{X}}_i, \bar{\boldsymbol{Y}}\rangle = 0 \tag{KKT-c}$$

$$\langle \bar{\boldsymbol{Z}}_i, \bar{\boldsymbol{X}}_i\rangle = 0 \tag{KKT-d}$$

$$\bar{\boldsymbol{Z}}_i \succeq 0. \tag{KKT-e}$$

Similar to the work in [44], our strategy is then to construct $\bar{\boldsymbol{X}}_i$ and $\bar{\boldsymbol{Y}}, \bar{\boldsymbol{Z}}_i, \bar{\boldsymbol{\nu}}$ satisfying these conditions. Given $\bar{\boldsymbol{U}}$ and $\bar{\boldsymbol{\nu}}$ in the statement of the theorem, we define $\bar{\boldsymbol{X}}_i = \bar{\boldsymbol{u}}_i\bar{\boldsymbol{u}}_i'$, $\bar{\boldsymbol{Y}} = \bar{\boldsymbol{U}}(\bar{\boldsymbol{\Lambda}} - \mathbf{D}_{\bar{\boldsymbol{\nu}}})\bar{\boldsymbol{U}}'$, and $\bar{\boldsymbol{Z}}_i = \bar{\boldsymbol{Y}} + \bar{\nu}_i\mathbf{I} - \mathbf{M}_i$. By construction, $\bar{\boldsymbol{X}}_i$ satisfy (KKT-a), and it is clear that $\bar{\boldsymbol{Y}} = \mathbf{M}_i + \bar{\boldsymbol{Z}}_i - \bar{\boldsymbol{\nu}}_i\mathbf{I}$ satisfies (KKT-b). One can also verify that $\langle \mathbf{I} - \bar{\boldsymbol{X}}, \bar{\boldsymbol{Y}}\rangle = 0$ by construction, thus satisfying (KKT-c). So it remains to show $\bar{\boldsymbol{Y}} \succeq 0$, $\langle \bar{\boldsymbol{Z}}_i, \bar{\boldsymbol{X}}_i\rangle = 0$, and $\bar{\boldsymbol{Z}}_i \succeq 0$.

The assumption that $\bar{\boldsymbol{\Lambda}} \succeq \mathbf{D}_{\bar{\boldsymbol{\nu}}}$ ensures $\bar{\boldsymbol{Y}} \succeq 0$ (KKT-b). We note that we have shown in Lemma D.1 and Lemma D.2 that $\bar{\boldsymbol{\Lambda}}$ is symmetric PSD, which is a necessary condition for this assumption to hold, given the fact that the Lagrange multipliers $\bar{\nu}_i$ corresponding to the trace constraints are nonnegative by Lemma 2.4.

Moreover, $\bar{\boldsymbol{Z}}_i \succeq 0$ by the assumption in (5), satisfying (KKT-e). We finally verify (KKT-d), i.e. $\langle \bar{\boldsymbol{Z}}_i, \bar{\boldsymbol{X}}_i\rangle = 0$, with $\bar{\boldsymbol{U}} = [\bar{\boldsymbol{u}}_1 \cdots \bar{\boldsymbol{u}}_k]$:

$$\begin{aligned} \langle \bar{\boldsymbol{Z}}_i, \bar{\boldsymbol{X}}_i\rangle &= \langle \bar{\boldsymbol{Y}} + \bar{\nu}_i\mathbf{I} - \mathbf{M}_i, \bar{\boldsymbol{X}}_i\rangle = \langle \bar{\boldsymbol{U}}(\bar{\boldsymbol{\Lambda}} - \mathbf{D}_{\bar{\boldsymbol{\nu}}})\bar{\boldsymbol{U}}' + \bar{\nu}_i\mathbf{I} - \mathbf{M}_i, \bar{\boldsymbol{u}}_i\bar{\boldsymbol{u}}_i'\rangle \\ &= \bar{\boldsymbol{u}}_i'\bar{\boldsymbol{U}}\bar{\boldsymbol{U}}'\sum_{j=1}^{k} \mathbf{M}_j\bar{\boldsymbol{U}}\mathbf{E}_j\bar{\boldsymbol{U}}'\bar{\boldsymbol{u}}_i - \bar{\boldsymbol{u}}_i'\bar{\boldsymbol{U}}\mathbf{D}_{\bar{\boldsymbol{\nu}}}\bar{\boldsymbol{U}}'\bar{\boldsymbol{u}}_i + \bar{\nu}_i - \bar{\boldsymbol{u}}_i'\mathbf{M}_i\bar{\boldsymbol{u}}_i \\ &= \boldsymbol{e}_i'\bar{\boldsymbol{U}}'\sum_{j=1}^{k} \mathbf{M}_j\bar{\boldsymbol{u}}_j\boldsymbol{e}_j'\boldsymbol{e}_i - \boldsymbol{e}_i'\mathbf{D}_{\bar{\boldsymbol{\nu}}}\boldsymbol{e}_i + \bar{\nu}_i - \bar{\boldsymbol{u}}_i'\mathbf{M}_i\bar{\boldsymbol{u}}_i \\ &= \bar{\boldsymbol{u}}_i'\mathbf{M}_i\bar{\boldsymbol{u}}_i - \bar{\nu}_i + \bar{\nu}_i - \bar{\boldsymbol{u}}_i'\mathbf{M}_i\bar{\boldsymbol{u}}_i = 0. \end{aligned}$$

□

One may ask if there is an analytical way to verify the dual variables $\bar{\boldsymbol{Y}}$ and $\bar{\boldsymbol{Z}}_i$ are PSD without computing the LMI feasibility problem in (5). While it is possible to derive sufficient upper bounds on the feasible $\bar{\nu}_i$ to guarantee $\bar{\boldsymbol{\Lambda}} \succeq \mathbf{D}_{\bar{\boldsymbol{\nu}}}$ so that $\bar{\boldsymbol{Y}} \succeq 0$, this is insufficient to certify $\bar{\boldsymbol{Z}}_i \succeq 0$ based on these bounds alone. This is in contrast to [44]; their particular dual certificate matrix is monotone in the Lagrange multipliers (analogous to our $\bar{\nu}_i$), so it is sufficient to test the positive semidefiniteness of the certificate matrix using the analytical upper bounds. Let $\bar{\boldsymbol{U}}_{\perp_i}$ denote an orthonormal basis for $\text{Span}(\mathbf{I} - \bar{\boldsymbol{u}}_i\bar{\boldsymbol{u}}_i')$. Here, since $\bar{\boldsymbol{Z}}_i = \bar{\boldsymbol{U}}\bar{\boldsymbol{\Lambda}}\bar{\boldsymbol{U}}' - \sum_{j\neq i}^{k} \bar{\nu}_j\bar{\boldsymbol{u}}_j\bar{\boldsymbol{u}}_j' + \bar{\nu}_i\bar{\boldsymbol{U}}_{\perp_i}\bar{\boldsymbol{U}}_{\perp_i}' - \mathbf{M}_i$, each $\bar{\boldsymbol{Z}}_i$ is monotone in $\bar{\nu}_i$ but not in $\bar{\nu}_j$ for $j \neq i$. Therefore, there is tension between inflating $\bar{\nu}_i$ and guaranteeing all the $\bar{\boldsymbol{Z}}_i$ are PSD. As such, an analytical solution to check that $\bar{\boldsymbol{\Lambda}} \succeq \mathbf{D}_{\bar{\boldsymbol{\nu}}}$ and the $\bar{\boldsymbol{Z}}_i$ are PSD remains unknown, requiring computation of the LMI feasibility problem in (5).

D.1 Arithmetic Complexity - more details

While SDP relaxations of nonconvex optimization problems can provide strong provable guarantees, their practicality can be limited by the time and space required to solve them, particularly when using off-the-shelf interior-point solvers. Interior-point methods are provably polynomial-time, but in our case the number of floating point operations and the storage per iteration to solve (SDP-P) both grow as $\mathcal{O}(d^3)$ [5], which practically limits d to be in the few hundreds.

On the other hand, the study of the SDP relaxation admits improved practical tools to transfer theoretical guarantees to the nonconvex setting; that is, to investigate when the convex relaxation is tight, and if it is, when a candidate solution of the nonconvex problem is globally optimal. In comparison to the dual problem of the SDP (SDP-D) (upon eliminating the variables $\boldsymbol{Z}_i$), the proposed global certificate significantly reduces the number of variables from $\mathcal{O}(d^2)$ to merely k variables. Precisely, the total computational savings can be shown using [4, Section 6.6.3], for which (SDP-D) scales in arithmetic complexity as $\mathcal{O}((kd)^{1/2}kd^6)$ floating point operations (flops) and the certificate scales by $\mathcal{O}((kd)^{1/2}k^2d^3)$ flops, showing a substantial reduction by a factor of $\mathcal{O}(d^3/k)$ flops. Subsequently, a first order MM solver in [14], whose cost is $\mathcal{O}(dk^2+k^3)$ per iteration, combined with our global optimality certificate is an obvious preference to solving the full SDP in (SDP-P) for large problems. Given the global certificate tool in Theorem 4.1, if (1) has a tight convex relaxation, we can reliably and cheaply certify the terminal output of a first order solver with possibly fewer restarts and without resorting to heuristics in nonconvex optimization, which commonly entails computing many multiple algorithm runs from different initializations and taking the solution with the best objective value.

E Proof of Theorem 4.5

We start this section by giving general convex analysis results that allow us to prove Theorem 4.5.

Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a closed, convex set. For all $\boldsymbol{c} \in \mathcal{C}$, consider a primal-dual pair of linear conic programs parameterized by c:

$$p(\boldsymbol{c}) := \min_{\boldsymbol{x}}\{\boldsymbol{c}^T\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \in \mathcal{K}\} \tag{$P;\boldsymbol{c}$}$$

$$d(\boldsymbol{c}) := \max_{\boldsymbol{y}}\{\boldsymbol{b}^T\boldsymbol{y} : \boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{y} \in \mathcal{K}^*\} \tag{$D;\boldsymbol{c}$}$$

Here, the data $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{R}^m$ are fixed; $\mathcal{K} \subseteq \mathbb{R}^n$ is a closed, convex cone; and $\mathcal{K}^* := \{\boldsymbol{s} \in \mathbb{R}^n : \boldsymbol{s}^T\boldsymbol{x} \geq 0 \ \forall\ \boldsymbol{x} \in \mathcal{K}\}$ is its polar dual. We imagine, in particular, that $\mathcal{K}$ is a direct product of a nonnegative orthant, second-order cones, and positive semidefinite cones, corresponding to linear, second-order-cone, and semidefinite programming.

Define $\text{Feas}(P) := \{\boldsymbol{x} \in \mathcal{K} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ and $\text{Feas}(D;\boldsymbol{c}) := \{\boldsymbol{y} : \boldsymbol{c} - \boldsymbol{A}^T\boldsymbol{y} \in \mathcal{K}^*\}$ to be the feasible sets of $(P;\boldsymbol{c})$ and $(D;\boldsymbol{c})$, respectively. We assume:

Assumption E.0.1. *Feas(P) is interior feasible, and Feas$(D;\boldsymbol{c})$ is interior feasible for all $\boldsymbol{c} \in \mathcal{C}$.*

Then, for all $\boldsymbol{c}$, strong duality holds between $(P;\boldsymbol{c})$ and $(D;\boldsymbol{c})$ in the sense that $p(\boldsymbol{c}) = d(\boldsymbol{c})$ and both $p(\boldsymbol{c})$ and $d(\boldsymbol{c})$ are attained in their respective problems. Accordingly, we also define

$$\mathrm{Opt}(D;\boldsymbol{c}) := \{\boldsymbol{y} \in \mathrm{Feas}(D;\boldsymbol{c}) : \boldsymbol{b}^T\boldsymbol{y} = d(\boldsymbol{c})\}$$

to be the nonempty, dual optimal solution set for each $\boldsymbol{c} \in \mathcal{C}$.

In addition, we assume the existence of linear constraints $\boldsymbol{f} - \boldsymbol{E}^T\boldsymbol{y} \geq 0$, independent of $\boldsymbol{c}$, such that

$$\mathrm{Extra}(D) := \{\boldsymbol{y} : \boldsymbol{f} - \boldsymbol{E}^T\boldsymbol{y} \geq 0\}$$

satisfies:

Assumption E.0.2. *For all $\boldsymbol{c} \in \mathcal{C}$, Feas$(D;\boldsymbol{c}) \cap$ Extra(D) is interior feasible and bounded, and Opt$(D;\boldsymbol{c}) \subseteq$ Extra(D).*

In words, irrespective of $\boldsymbol{c}$, the extra constraints $\boldsymbol{f} - \boldsymbol{E}^T\boldsymbol{y} \geq 0$ bound the dual feasible set without cutting off any optimal solutions and while still maintaining interior, including interiority with respect to $\boldsymbol{f} - \boldsymbol{E}^T\boldsymbol{y} \geq 0$. Note also that Assumption E.0.2 implies the recession cone of $\mathrm{Feas}(D;\boldsymbol{c}) \cap \mathrm{Extra}(D)$ is trivial for (and independent of) all $\boldsymbol{c}$, i.e., $\{\Delta\boldsymbol{y} : -\boldsymbol{A}^T\Delta\boldsymbol{y} \in \mathcal{K}^*, -\boldsymbol{E}^T\Delta\boldsymbol{y} \geq 0\} = \{0\}$.

We first prove a continuity result related to the dual feasible set, in which we use the following definition of a convergent sequence of bounded sets in Euclidean space: a sequence of bounded sets $\{L^k\}$ converges to a bounded set $\bar{L}$, written $\{L^k\} \to \bar{L}$, if and only if: (i) given any sequence $\{\boldsymbol{y}^k \in L^k\}$, every limit point $\bar{\boldsymbol{y}}$ of the sequence satisfies $\bar{\boldsymbol{y}} \in \bar{L}$; and (ii) every member $\bar{\boldsymbol{y}} \in \bar{L}$ is the limit point of some sequence $\{\boldsymbol{y}^k \in L^k\}$.

Lemma E.1. *Under Assumptions E.0.1 and E.0.2, let $\{\boldsymbol{c}^k \in \mathcal{C}\} \to \bar{\boldsymbol{c}}$ be any convergent sequence. Then*

$$\big\{\mathit{Feas}(D;\boldsymbol{c}^k) \cap \mathit{Extra}(D)\big\} \quad \to \quad \mathit{Feas}(D;\bar{\boldsymbol{c}}) \cap \mathit{Extra}(D).$$

Proof. For notational convenience, define $L^k := \mathrm{Feas}(D;\boldsymbol{c}^k) \cap \mathrm{Extra}(D)$ and $\bar{L} := \mathrm{Feas}(D;\bar{\boldsymbol{c}}) \cap \mathrm{Extra}(D)$. Note that L^k and $\bar{L}$ are bounded with interior by Assumption E.0.2. We wish to show $\{L^k\} \to \bar{L}$.

We first note that any sequence $\{\boldsymbol{y}^k \in L^k\}$ must be bounded. If not, then $\{\Delta\boldsymbol{y}^k := \boldsymbol{y}^k/\|\boldsymbol{y}^k\|\}$ is a bounded sequence satisfying

$$\|\Delta\boldsymbol{y}^k\| = 1, \quad \frac{\boldsymbol{c}^k}{\|\boldsymbol{y}^k\|} - \boldsymbol{A}^T\Delta\boldsymbol{y}^k \in \mathcal{K}^*, \quad \frac{\boldsymbol{f}}{\|\boldsymbol{y}^k\|} - \boldsymbol{E}^T\Delta\boldsymbol{y}^k \geq 0$$

and hence has a limit point $\Delta\boldsymbol{y}$ satisfying

$$\Delta\boldsymbol{y} \neq 0, \quad -\boldsymbol{A}^T\Delta\boldsymbol{y} \in \mathcal{K}^*, \quad -\boldsymbol{E}^T\Delta\boldsymbol{y} \geq 0,$$

but this is a contradiction by the discussion after the statement of Assumption E.0.2. We thus conclude that any sequence $\{\boldsymbol{y}^k \in L^k\}$ has a limit point.

Appealing to the definition of the convergence of sets stated before the lemma, we first let $\bar{\boldsymbol{y}}$ be a limit point of any $\{\boldsymbol{y}^k \in L^k\}$ and prove that $\bar{\boldsymbol{y}} \in \bar{L}$. Since

$$\boldsymbol{c}^k - \boldsymbol{A}^T\boldsymbol{y}^k \in \mathcal{K}^*, \quad \boldsymbol{f} - \boldsymbol{E}^T\boldsymbol{y}^k \geq 0$$

for all k, by taking the limit of $\{\boldsymbol{c}^k\}$ and $\{\boldsymbol{y}^k\}$, we have $\bar{\boldsymbol{c}} - \boldsymbol{A}^T\bar{\boldsymbol{y}} \in \mathcal{K}^*$ and $\boldsymbol{f} - \boldsymbol{E}^T\bar{\boldsymbol{y}} \geq 0$ so that indeed $\bar{\boldsymbol{y}} \in \bar{L}$.

Next, we must show that every $\bar{\boldsymbol{y}} \in \bar{L}$ is the limit point of some sequence $\{\boldsymbol{y}^k \in L^k\}$. For this proof, define

$$\kappa(\bar{\boldsymbol{y}}) := \min\{k : \bar{\boldsymbol{y}} \in L^\ell \ \ \forall\ \ell \geq k\},$$

i.e., $\kappa(\bar{\boldsymbol{y}})$ is the smallest k such that $\bar{\boldsymbol{y}}$ is a member of every set in the tail $L^k, L^{k+1}, L^{k+2}, \ldots$. By convention, if there exists no such k, we set $\kappa(\bar{\boldsymbol{y}}) = \infty$.

Let us first consider the case $\bar{\boldsymbol{y}} \in \text{int}(\bar{L})$. We claim $\kappa(\bar{\boldsymbol{y}}) < \infty$, so that setting $\boldsymbol{y}^k = \bar{\boldsymbol{y}}$ for all $k \geq \kappa(\bar{\boldsymbol{y}})$ yields the desired sequence converging to $\bar{\boldsymbol{y}}$. Indeed, as $\bar{\boldsymbol{y}}$ satisfies $\bar{\boldsymbol{c}} - \boldsymbol{A}^T\bar{\boldsymbol{y}} \in \text{int}(\mathcal{K}^*)$ and $\boldsymbol{f} - \boldsymbol{E}^T\bar{\boldsymbol{y}} > 0$, the equation

$$\boldsymbol{c}^k - \boldsymbol{A}^T\bar{\boldsymbol{y}} = \left(\bar{\boldsymbol{c}} - \boldsymbol{A}^T\bar{\boldsymbol{y}}\right) + \left(\boldsymbol{c}^k - \bar{\boldsymbol{c}}\right)$$

shows that $\{\boldsymbol{c}^k - \boldsymbol{A}^T\bar{\boldsymbol{y}}\}$ equals $\bar{\boldsymbol{c}} - A^T\bar{\boldsymbol{y}} \in \text{int}(\mathcal{K}^*)$ plus the vanishing sequence $\{\boldsymbol{c}^k - \bar{\boldsymbol{c}}\}$. Hence its tail is contained in $\text{int}(\mathcal{K}^*)$, thus proving $\kappa(\bar{\boldsymbol{y}}) < \infty$, as desired.

Now we consider the case $\bar{\boldsymbol{y}} \in \text{bd}(\bar{L})$. Let $\boldsymbol{y}^0 \in \text{int}(\bar{L})$ be arbitrary, so that $\kappa(\boldsymbol{y}^0) < \infty$ by the previous paragraph. For a second index $\ell = 1, 2, \ldots$, define $\boldsymbol{z}^\ell := (1/\ell)\boldsymbol{y}^0 + (1 - 1/\ell)\bar{\boldsymbol{y}} \in \text{int}(\bar{L})$. Clearly, $\kappa(\boldsymbol{z}^\ell) < \infty$ for all ℓ and $\{\boldsymbol{z}^\ell\} \to \bar{\boldsymbol{y}}$. We then construct the desired sequence $\{\boldsymbol{y}^k \in L^k\}$ converging to $\bar{\boldsymbol{y}}$ as follows. First, set

$$\begin{aligned} k_1 &:= \kappa(\boldsymbol{z}^1) = \kappa(\boldsymbol{y}^0) \\ k_\ell &:= \max\{k_{\ell-1} + 1, \kappa(\boldsymbol{z}^\ell)\} \quad \forall\ \ell = 2, 3, \ldots \end{aligned}$$

and then, for all ℓ and for all $k \in [k_\ell, k_{\ell+1} - 1]$, define $\boldsymbol{y}^k := \boldsymbol{z}^\ell$. Essentially, $\{\boldsymbol{y}^k\}$ is the sequence $\{\boldsymbol{z}^\ell\}$, except with entries repeated to ensure $\boldsymbol{y}^k$ is in fact a member of L^k for all k. Hence, $\{\boldsymbol{y}^k\}$ converges to $\bar{\boldsymbol{y}}$ as desired. □

We now specialize Lemma E.1 to the dual optimality set.

Lemma E.2. *Under Assumptions E.0.1 and E.0.2, let $\{\boldsymbol{c}^k \in \mathcal{C}\} \to \bar{\boldsymbol{c}}$ be any convergent sequence. Then*

$$\left\{ \mathit{Opt}(D; \boldsymbol{c}^k) \right\} \quad \to \quad \mathit{Opt}(D; \bar{\boldsymbol{c}}).$$

Proof. Lemma E.1 establishes $\{\text{Feas}(D; \boldsymbol{c}^k) \cap \text{Extra}(D)\} \to \text{Feas}(D; \bar{\boldsymbol{c}}) \cap \text{Extra}(D)$. Because problems $(D; \boldsymbol{c}^k)$ and $(D; \bar{\boldsymbol{c}})$ share the same objective $\boldsymbol{b}^T\boldsymbol{y}$, Assumption E.0.2 and Lemma E.1 imply $\{d(\boldsymbol{c}^k)\} \to d(\bar{\boldsymbol{c}})$. Hence, the sequence of hyperplanes $\{\{\boldsymbol{y} : \boldsymbol{b}^T\boldsymbol{y} = d(\boldsymbol{c}^k)\}\}$ converges to $\{\boldsymbol{y} : \boldsymbol{b}^T\boldsymbol{y} = d(\bar{\boldsymbol{c}})\}$, and so $\{\text{Opt}(D; \boldsymbol{c}^k)\} \to \text{Opt}(D; \bar{\boldsymbol{c}})$ by intersecting the two sequences of convergent sets. □

Finally, for given $\boldsymbol{c} \in \mathcal{C}$ and fixed $\boldsymbol{y}^0 \in \mathbb{R}^m$, we define the function

$$y(\boldsymbol{c}) := y(\boldsymbol{c}; \boldsymbol{y}^0) = \text{argmin}\{\|\boldsymbol{y} - \boldsymbol{y}^0\| : \boldsymbol{y} \in \text{Opt}(D; \boldsymbol{c})\},$$

i.e., $y(\boldsymbol{c})$ equals the point in $\text{Opt}(D; \boldsymbol{c})$, which is closest to $\boldsymbol{y}^0$. Since $\text{Opt}(D; \boldsymbol{c})$ is closed and convex, $y(\boldsymbol{c})$ is well defined. We next use Lemma E.2 to show that that $y(\boldsymbol{c})$ is continuous in $\boldsymbol{c}$.

Proposition E.3. *Under the Assumptions E.0.1 and E.0.2, given $\boldsymbol{y}^0 \in \mathbb{R}^m$, the function $y(\boldsymbol{c}) := y(\boldsymbol{c}; \boldsymbol{y}^0)$ is continuous in $\boldsymbol{c}$.*

Proof. We must show that, for any convergent $\{\boldsymbol{c}^k\} \to \bar{\boldsymbol{c}}$, we also have convergence $\{y(\boldsymbol{c}^k)\} \to y(\bar{\boldsymbol{c}})$. This follows because $\{\text{Opt}(D; \boldsymbol{c}^k)\} \to \text{Opt}(D; \bar{\boldsymbol{c}})$ by Lemma E.2. □

Theorem 4.5 uses Proposition E.3 in its proof. Here we discuss how the primal-dual pair SDP-P-(SDP-D) satisfy the assumptions for the proposition. We would like to establish conditions under which (SDP-P) has the rank-1 property. For this, we apply the general theory developed above, specifically Proposition E.3. To show that the general theory applies, we must define the closed, convex set $\mathcal{C}$, which contains the set of admissible objective matrices/coefficients $(\mathbf{M}_1, \ldots, \mathbf{M}_k)$ and which satisfies Assumptions E.0.1 and E.0.2. In particular, for a fixed, user-specified upper bound $\mu > 0$, we define

$$\mathcal{C} := \{c = (\mathbf{M}_1, \ldots, \mathbf{M}_k) : 0 \preceq \mathbf{M}_i \preceq \mu\boldsymbol{I} \quad \forall\ i = 1, \ldots, k\},$$

to be our set of admissible coefficient k-tuples.

We know that both (SDP-P) and (SDP-D) have interior points for all $c \in \mathcal{C}$, so that strong duality holds. For the dual in particular, the equation $\mu\boldsymbol{I} = \boldsymbol{M}_i + ((\mu + \epsilon)\boldsymbol{I} - \boldsymbol{M}_i) - \epsilon\boldsymbol{I}$ shows that, for all $\epsilon > 0$,

$$\boldsymbol{Y}(\epsilon) := \mu\boldsymbol{I}, \qquad \boldsymbol{Z}(\epsilon)_i := (\mu + \epsilon)\boldsymbol{I} - \mathbf{M}_i, \qquad \nu(\epsilon)_i := \epsilon$$

is interior feasible with objective value $d\mu + k\epsilon$. In particular, the redundant constraint $\boldsymbol{\nu} \geq 0$ is satisfied strictly. This verifies Assumption E.0.1.

We next verify Assumption E.0.2. Since the objective value just mentioned is independent of $c = (\mathbf{M}_1, \ldots, \mathbf{M}_k)$, we can take $\epsilon = 1$ and enforce the extra constraint $\operatorname{tr}(\boldsymbol{Y}) + \sum_{i=1}^k \nu_i \leq d\mu + k$ without cutting off any dual optimal solutions and while still maintaining interior. In particular, the solution $(\boldsymbol{Y}(\frac{1}{2}), \boldsymbol{Z}(\frac{1}{2})_i, \boldsymbol{\nu}(\frac{1}{2})_i)$ corresponding to $\epsilon = \frac{1}{2}$ satisfies the new, extra constraint strictly. Finally, note that $\operatorname{tr}(\boldsymbol{Y}) + \sum_i \nu_i \leq d\mu + k$ bounds $\boldsymbol{Y}$ and $\boldsymbol{\nu}$ in the presence of the constraints $\boldsymbol{Y} \succeq 0$ and $\boldsymbol{\nu} \geq 0$, and consequently the constraint $\boldsymbol{Z}_i = \boldsymbol{Y} - \mathbf{M}_i + \nu_i \mathbf{I}$ bounds $\boldsymbol{Z}_i$ for each i.

We now repeat the discussion leading up to Theorem 4.5 for completeness. The first lemma says that the diagonal problem has dual variables $\boldsymbol{Z}_i$ such that $\operatorname{rank}(\boldsymbol{Z}_i) \geq d - 1$, implying that the primal variables $\boldsymbol{X}_i$ are rank-1.

Lemma E.4 (Restatement of Lemma 4.3)**.** *Let $c = (\mathbf{M}_1, \ldots, \mathbf{M}_k) \in \mathcal{C}$. If $\mathbf{M}_i$ is jointly diagonalizable for each $i = 1, \ldots, k$ and (SDP-P) has a unique optimal solution, then there exists an optimal solution of (SDP-D) with rank$(\boldsymbol{Z}_i) \geq d - 1$ for all $i = 1, \ldots, k$.*

Proof of Lemma E.4/Lemma 4.3. Because of the jointly diagonalizable property, we may assume without loss of generality that each $\mathbf{M}_i$ is diagonal. So (SDP-P) is equivalent to the assignment LP

$$\max \left\{ \sum_{i=1}^k \operatorname{diag}(\mathbf{M}_i)' \operatorname{diag}(\boldsymbol{X}_i) : \begin{array}{l} \boldsymbol{e}' \operatorname{diag}(\boldsymbol{X}_i) = 1,\ \operatorname{diag}(\boldsymbol{X}_i) \geq 0 \ \ \forall\ i = 1, \ldots, k \\ \sum_{i=1}^k \operatorname{diag}(\boldsymbol{X}_i) \leq \boldsymbol{e} \end{array} \right\},$$

where $\boldsymbol{e}$ is the vector of all ones, and (SDP-D) is equivalent to the LP

$$\min \left\{ \boldsymbol{e}' \operatorname{diag}(\boldsymbol{Y}) + \sum_{i=1}^k \nu_i : \begin{array}{l} \operatorname{diag}(\boldsymbol{Y}) = \operatorname{diag}(\mathbf{M}_i) + \operatorname{diag}(\boldsymbol{Z}_i) - \nu_i \boldsymbol{e} \ \ \forall\ i = 1, \ldots, k \\ \operatorname{diag}(\boldsymbol{Z}_i) \geq 0 \ \ \forall\ i = 1, \ldots, k, \quad \operatorname{diag}(\boldsymbol{Y}) \geq 0 \end{array} \right\}.$$

Since the primal is an assignment problem, its unique optimal solution has the property that each $\operatorname{diag}(\boldsymbol{X}_i)$ is a standard basis vector (i.e., each has a single entry equal to 1 and all other entries equal to 0). By the Goldman-Tucker strict complementarity theorem for LP, there exists an optimal primal-dual pair such that $\operatorname{diag}(\boldsymbol{X}_i) + \operatorname{diag}(\boldsymbol{Z}_i) > 0$ for each i. Hence, there exists a dual optimal solution with $\operatorname{rank}(\boldsymbol{Z}_i) \geq d - 1$ for each i, as desired. □

Definition E.5 (Restatement of Definition 4.4)**.** *For $\boldsymbol{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) \in \mathcal{C}$ and $\bar{\boldsymbol{c}} = (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k) \in \mathcal{C}$, define*

$$\mathit{dist}(\boldsymbol{c}, \bar{\boldsymbol{c}}) \triangleq \max_{i \in [k]} \|\mathbf{M}_i - \bar{\mathbf{M}}_i\|.$$

We now repeat the continuity result and give a more precise proof.

Theorem E.6 (Restatement of Theorem 4.5)**.** *Let $\bar{\boldsymbol{c}} := (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k) \in \mathcal{C}$ be jointly diagonalizable such that the primal problem (SDP-P) with objective coefficients $\bar{\boldsymbol{c}}$ has a unique optimal solution. Then there exists an full-dimensional neighborhood of $\bar{\mathcal{C}} \ni \bar{c}$ in $\mathcal{C}$ such that (SDP-P) has the rank-1 property for all $\boldsymbol{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) \in \bar{\mathcal{C}}$.*

Proof of Theorem E.6/Theorem 4.5. Using Lemma 4.3, let $\boldsymbol{y}^0 := (\bar{\boldsymbol{Y}}, \bar{\boldsymbol{Z}}_i, \bar{\nu}_i)$ be the optimal solution of the dual problem (SDP-D) for $\bar{\boldsymbol{c}} = (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k)$, which has $\operatorname{rank}(\bar{\boldsymbol{Z}}_i) \geq d - 1$ for all i. Then by Proposition E.3, the function $y(\boldsymbol{c}) := y(\boldsymbol{c}; \boldsymbol{y}^0)$, which returns the optimal solution of (SDP-D) for $\boldsymbol{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k)$ closest to $\boldsymbol{y}^0$, is continuous in $\boldsymbol{c}$. It follows that the preimage

$$y^{-1}\left(\{(\boldsymbol{Y}, \boldsymbol{Z}_i, \nu_i) : \operatorname{rank}(\boldsymbol{Z}_i) \geq d - 1 \ \ \forall\ i\}\right)$$

contains $\bar{\boldsymbol{c}}$ and is an open set because the set of all $(\boldsymbol{Y}, \boldsymbol{Z}_i, \nu_i)$ with $\operatorname{rank}(\boldsymbol{Z}_i) \geq d - 1$ is an open set. After intersecting with $\mathcal{C}$, this full-dimensional set $\bar{\mathcal{C}}$ proves the theorem via the complementarity of the KKT conditions of the assignment LP, $\operatorname{rank}(\boldsymbol{Z}_i) = d - 1$ for $i = 1, \ldots, k$, and Lemma 2.5. □

The next corollary is a slightly more general version of Corollary 4.5.1, which shows that for a general tuple of almost commuting matrices $\boldsymbol{c} := (\mathbf{M}_1, \ldots, \mathbf{M}_k)$ that are CJD, (SDP-P) is tight and has the rank-1 property. We first state Lin's theorem, which we use in the proof.

Lemma E.7. *Lin's Theorem [28; 22]: For all $\epsilon > 0$ there exists a $\delta > 0$ such that if $\|[\mathbf{A}, \mathbf{B}]\|_2 := \|\mathbf{AB} - \mathbf{BA}\|_2 \leq \delta$ for Hermitian symmetric matrices $\mathbf{A}$ and $\mathbf{B}$ where $\|\mathbf{A}\| \leq 1$ and $\|\mathbf{B}\| \leq 1$, then there exist Hermitian symmetric, commuting matrices $\widetilde{\mathbf{A}}$ and $\widetilde{\mathbf{B}}$ in $\mathbb{R}^{d\times d}$ such that $\|[\tilde{\mathbf{A}}, \tilde{\mathbf{B}}]\| = 0$ and $\|\mathbf{A} - \tilde{\mathbf{A}}\|_2 \leq \epsilon$ and $\|\mathbf{B} - \tilde{\mathbf{B}}\|_2 \leq \epsilon$.*

Corollary E.7.1 (Restatement of Corollary 4.5.1 with an additional conclusion)**.** *Assume $\|\mathbf{M}_i\| \leq 1$ for all $i \in [k]$, and let $\boldsymbol{c} := (\mathbf{M}_1, \ldots, \mathbf{M}_k)$. Suppose $\|[\mathbf{M}_i, \mathbf{M}_j]\| := \|\mathbf{M}_i\mathbf{M}_j - \mathbf{M}_j\mathbf{M}_i\| \leq \delta$ for all $i, j \in [k]$. Then there exists $\bar{\boldsymbol{c}} := (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k)$ of commuting, jointly-diagonalizable matrices such that $\|[\bar{\mathbf{M}}_i, \bar{\mathbf{M}}_j]\| = 0$ for all $i, j \in [k]$ where $dist(\boldsymbol{c}, \bar{\boldsymbol{c}}) \leq \mathcal{O}(\epsilon(\delta))$ and $\epsilon(\delta)$ is a function satisfying $\lim_{\delta\to 0} \epsilon(\delta) = 0$. If δ is small enough, there exists $\epsilon > \epsilon(\delta) > 0$ such that $dist(\boldsymbol{c}, \bar{\boldsymbol{c}}) \leq \epsilon$ implies $\boldsymbol{c} \in \bar{\mathcal{C}}$ and* (SDP-P) *has the rank-one property.*

Further, for the problem in (2)*, assume $\|\mathbf{A}_\ell\| \leq 1$ for all $\ell \in [L]$. If $\|[\mathbf{A}_\ell, \mathbf{A}_m]\| \leq \delta$ for all $\ell, m \in [L]$ for some $\delta > 0$, then there exists a $\bar{\boldsymbol{c}} \in \mathcal{C}$ such that $dist(\boldsymbol{c}, \bar{\boldsymbol{c}}) \leq \mathcal{O}(\sum_{\ell=1}^{L} w_{\ell,i}\epsilon(\delta))$.*

Proof of Corollary E.7.1 / Corollary 4.5.1. The general result follows from directly applying Lemma E.7 to each $\mathbf{M}_i$, and for the instance of problem (2), we apply Lemma E.7 to each $\mathbf{A}_\ell$. Then there exist Hermitian symmetric matrices $\bar{\mathbf{A}}_\ell$ such that $\|[\bar{\mathbf{A}}_\ell, \bar{\mathbf{A}}_m]\| = 0$ for all $\ell, m \in [L]$ such that $\|\mathbf{A}_\ell - \bar{\mathbf{A}}_\ell\| \leq \epsilon(\delta)$ for all $\ell \in [L]$. Let $\bar{\mathbf{M}}_i := \sum_{\ell=1}^{L} w_{\ell,i}\bar{\mathbf{A}}_\ell$. Then the matrices $\bar{\mathbf{M}}_i$ commute and are jointly diagonalizable:

$$[\bar{\mathbf{M}}_i, \bar{\mathbf{M}}_j] = \bar{\mathbf{M}}_i\bar{\mathbf{M}}_j - \bar{\mathbf{M}}_j\bar{\mathbf{M}}_i = 2\sum_{\ell\neq m}^{L} w_{\ell,i}w_{m,j}(\bar{\mathbf{A}}_\ell\bar{\mathbf{A}}_m - \bar{\mathbf{A}}_m\bar{\mathbf{A}}_\ell) = 0. \tag{24}$$

Now we measure the distance between each $\mathbf{M}_i$ and $\bar{\mathbf{M}}_i$:

$$\|\mathbf{M}_i - \bar{\mathbf{M}}_i\| = \|\sum_{\ell=1}^{L} w_{\ell,i}(\mathbf{A}_\ell - \bar{\mathbf{A}}_\ell)\| \leq \sum_{\ell=1}^{L} w_{\ell,i}\|\mathbf{A}_\ell - \bar{\mathbf{A}}_\ell\| \leq \sum_{\ell=1}^{L} w_{\ell,i}\epsilon(\delta). \tag{25}$$

□

The next lemma is from Koltchinskii and Lounici [27] and also [29].

Lemma E.8. *Let $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n \subseteq \mathbb{R}^d$ be i.i.d. centered Gaussian random variables in a separable Banach space with covariance operator $\boldsymbol{\Sigma}$ and sample covariance $\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} \boldsymbol{y}_i\boldsymbol{y}_i'$. Then with some constant $C > 0$ and with probability at least $1 - e^{-t}$ for $t > 0$,*

$$\|\hat{\boldsymbol{\Sigma}} - \boldsymbol{\Sigma}\| \leq C\|\boldsymbol{\Sigma}\| \max\left\{\sqrt{\frac{\tilde{r}(\boldsymbol{\Sigma})\log d + t}{n}}, \frac{(\tilde{r}(\boldsymbol{\Sigma})\log d + t)\log n}{n}\right\},$$

where $\tilde{r}(\boldsymbol{\Sigma}) := \mathrm{tr}(\boldsymbol{\Sigma})/\|\boldsymbol{\Sigma}\|$.

Lemma E.9. *Let $\bar{\mathbf{M}}_i := \mathbb{E}[\mathbf{M}_i] \in \mathbb{R}^{d\times d}$, where the expectation is taken with respect to the data observations, and let $C > 0$ be a universal constant. Normalize each $\mathbf{M}_i$ by $1/n$. Then $\|[\bar{\mathbf{M}}_i, \bar{\mathbf{M}}_j]\| = 0$, and with probability at least $1 - e^{-t}$ for $t > 0$*

$$\frac{\|\mathbf{M}_i - \bar{\mathbf{M}}_i\|}{\|\bar{\mathbf{M}}_1\|} \leq C\frac{\bar{\sigma}_i}{\bar{\sigma}_1} \max\left\{\sqrt{\frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i}\log d + t}{n}}, \frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i}\log d + t}{n}\log(n)\right\}, \quad \textit{where} \tag{26}$$

$$\bar{\sigma}_i = \|\bar{\mathbf{M}}_i\| = \sum_{\ell=1}^{L} \frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell} + 1}\frac{n_\ell}{n}(\frac{\lambda_1}{v_\ell} + 1) \quad \textit{and} \quad \bar{\xi}_i = \mathrm{tr}(\bar{\mathbf{M}}_i) = \sum_{\ell=1}^{L} \frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell} + 1}\frac{n_\ell}{n}(\frac{1}{v_\ell}\sum_{i=1}^{k}\lambda_i + d).$$

Proof. Let $\tilde{\boldsymbol{y}}_{\ell,i} := \sqrt{\frac{w_{\ell,i}}{v_\ell}}\boldsymbol{y}_{\ell,i}$ be a rescaling of the data vectors. Then $\tilde{\boldsymbol{y}}_{\ell,i} \overset{iid}{\sim} \mathcal{N}(\mathbf{0}, w_{\ell,i}(\frac{1}{v_\ell}\boldsymbol{U}\boldsymbol{\Theta}^2\boldsymbol{U}' + \mathbf{I}))$. After rescaling, $\mathbf{M}_i = \frac{1}{n}\sum_{\ell=1}^{L}\sum_{j=1}^{n_\ell}\tilde{\boldsymbol{y}}_{\ell,j}\tilde{\boldsymbol{y}}'_{\ell,j}$. Taking the expectation over the data, we have

$$\mathbb{E}[\mathbf{M}_i] = \frac{1}{n}\sum_{\ell=1}^{L}\sum_{j=1}^{n_\ell}\mathbb{E}[\tilde{\boldsymbol{y}}_{m,j}\tilde{\boldsymbol{y}}'_{m,j}|m=\ell] = \sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}\left(\frac{1}{v_\ell}\boldsymbol{U}\boldsymbol{\Theta}^2\boldsymbol{U}' + \mathbf{I}\right). \tag{27}$$

Let $\boldsymbol{U}_\perp \in \mathbb{R}^{d\times d-k}$ be an orthonormal basis spanning the orthogonal complement of Span($\boldsymbol{U}$). Noting that $\mathbf{I} = \boldsymbol{U}\boldsymbol{U}' + \boldsymbol{U}_\perp\boldsymbol{U}'_\perp$, rewrite $\mathbb{E}[\mathbf{M}_i]$ in terms of its eigendecomposition by

$$\mathbb{E}[\mathbf{M}_i] = \boldsymbol{U}\left(\sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}\left(\frac{1}{v_\ell}\boldsymbol{\Theta}^2 + \mathbf{I}_k\right)\right)\boldsymbol{U}' + \left(\sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}\right)\boldsymbol{U}_\perp\boldsymbol{U}'_\perp \tag{28}$$

$$= \begin{bmatrix}\boldsymbol{U} & \boldsymbol{U}_\perp\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma} & 0\\ 0 & \gamma\mathbf{I}_{d-k}\end{bmatrix}\begin{bmatrix}\boldsymbol{U}'\\ \boldsymbol{U}'_\perp,\end{bmatrix} \tag{29}$$

where $\boldsymbol{\Sigma} := \sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}\left(\frac{1}{v_\ell}\boldsymbol{\Theta}^2 + \mathbf{I}_k\right)$ and $\gamma := \sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}$, from which we obtain the expressions for $\bar{\sigma}_i = \|\mathbb{E}[\mathbf{M}_i]\|$ and $\bar{\xi}_i = \mathrm{tr}(\mathbb{E}[\mathbf{M}_i])$. Then invoking Lemma E.8 to bound the concentration of a sample covariance matrix to its expectation with high probability yields the final result. □

Proposition E.10 (Restatement of Proposition 4.6)**.** *Let $\boldsymbol{c} = (\frac{1}{n}\mathbf{M}_1, \ldots, \frac{1}{n}\mathbf{M}_k)$ be the (normalized) data matrices of the HPPCA problem. Then there exists commuting $\bar{\boldsymbol{c}} = (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k)$ (constructed in the proof) such that, for a universal constant $C > 0$ and with probability exceeding $1 - e^{-t}$ for $t > 0$,*

$$\frac{\|\mathbf{M}_i - \bar{\mathbf{M}}_i\|}{\|\bar{\mathbf{M}}_1\|} \le \min\left\{\sum_{\ell=1}^{L}\frac{1}{\frac{\lambda_i}{v_\ell}+1}, C\frac{\bar{\sigma}_i}{\bar{\sigma}_1}\max\left\{\sqrt{\frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i}\log d + t}{n}}, \frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i}\log d + t}{n}\log(n)\right\}\right\}, \quad \textit{where} \tag{30}$$

$$\bar{\sigma}_i = \|\bar{\mathbf{M}}_i\| = \sum_{\ell=1}^{L}\frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell}+1}\frac{n_\ell}{n}\left(\frac{\lambda_1}{v_\ell}+1\right) \quad \textit{and} \quad \bar{\xi}_i = \mathrm{tr}(\bar{\mathbf{M}}_i) = \sum_{\ell=1}^{L}\frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell}+1}\frac{n_\ell}{n}\left(\frac{1}{v_\ell}\sum_{i=1}^{k}\lambda_i + d\right).$$

Proof of Proposition E.10/Proposition 4.6. We argue there are two possible sets of commuting $(\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k)$ that $(\mathbf{M}_1, \ldots, \mathbf{M}_k)$ can converge to, depending on the signal to noise ratios $\frac{\lambda_i}{v_\ell}$ and the number of samples n.

Consider that we can scale all the $\mathbf{M}_i$ in (SDP-P) by a positive scalar constant without changing the optimal solution. Since all the $\mathbf{M}_i$ can be arbitrarily scaled in this manner, and thereby changing any distance measure, we will choose to normalize the matrices $\mathbf{M}_i$ and $\bar{\mathbf{M}}_i$ by the number of samples and the largest spectral norm of the $\bar{\mathbf{M}}_i$, which is equivalent to also normalizing the distance. For the HPPCA application, since $\bar{\mathbf{M}}_1 \succeq \bar{\mathbf{M}}_2 \ldots \succeq \bar{\mathbf{M}}_k$, we normalize by $1/\|n\bar{\mathbf{M}}_1\|$.

First, if the variances are zero or all the same, i.e. noiseless or homoscedastic noisy data, then all the $\mathbf{M}_i$ commute. Otherwise, in the case where each SNR λ_i/v_ℓ of the i^{th} components is large or close to the same value for all $\ell \in [L]$, the weights $w_{\ell,i} = \frac{\lambda_i/v_\ell}{\lambda_i/v_\ell+1}$ are very close to 1 or some constant less than 1, respectively. Therefore, let $\bar{\mathbf{M}} := \frac{1}{n}\sum_{\ell=1}^{L}\mathbf{A}_\ell$ for all $i \in [k]$. Then

$$\frac{\|\mathbf{M}_i - \bar{\mathbf{M}}\|}{\|\bar{\mathbf{M}}\|} = \frac{\left\|\sum_{\ell=1}^{L}(w_{\ell,i}-1)\mathbf{A}_\ell\right\|}{\|\sum_{\ell=1}^{L}\mathbf{A}_\ell\|} \le \frac{\sum_{\ell=1}^{L}\frac{1}{\frac{\lambda_i}{v_\ell}+1}\|\mathbf{A}_\ell\|}{\|\sum_{\ell=1}^{L}\mathbf{A}_\ell\|} \le \sum_{\ell=1}^{L}\frac{1}{\frac{\lambda_i}{v_\ell}+1}, \tag{31}$$

the last inequality above results from the fact $\frac{\|\mathbf{A}_\ell\|}{\|\sum_{\ell=1}^{L}\mathbf{A}_\ell\|} \le 1$ for all $\ell \in [L]$ using Weyl's inequality for symmetric PSD matrices [42].

While the bound above depends on the SNR, it fails to capture the effects of the sample sizes, which also play an important role in how close the $\mathbf{M}_i$ are to commuting. Even in the case where the variances

are larger and more heterogeneous, since the $\mathbf{M}_i$ form a weighted sum of sample covariance matrices, given enough samples, they should concentrate to their respective sample covariance matrices, which commute between $i, j \in [k]$. We show exactly this using the concentration of sample covariances to their expectation in [29], and choose $\bar{\boldsymbol{c}} = (\bar{\mathbf{M}}_1, \ldots, \bar{\mathbf{M}}_k)$ for $\bar{\mathbf{M}}_i := \mathbb{E}[\mathbf{M}_i]$, where the expectation here is with respect to the data generated by the model in (6).

Let $\bar{\mathbf{M}}_i := \mathbb{E}[\mathbf{M}_i] \in \mathbb{R}^{d \times d}$, where the expectation is taken with respect to the data observations. Then by Lemma E.9 and taking the minimum with (31), we obtain the final result.

□

F Example of SDP with rank-one solutions, but $\mathbf{M}_i$ that are not almost commuting

In our paper, we give sufficient conditions for when the SDP returns rank-one orthogonal primal solutions in the case the $\mathbf{M}_i$ matrices almost commute. However, this is not a necessary condition, and we give a counter-example here.

Proposition F.1. *Construct $\mathbf{M}_i$ for $i = 1, \ldots, k$ as follows for given length-d vectors $\boldsymbol{v}_i$, $i = 1, \ldots, k$:*

$$\begin{aligned}
\mathbf{M}_1 &= \boldsymbol{v}_1\boldsymbol{v}_1' + \boldsymbol{v}_2\boldsymbol{v}_2' + \cdots + \boldsymbol{v}_k\boldsymbol{v}_k' \\
\mathbf{M}_2 &= \qquad\quad \boldsymbol{v}_2\boldsymbol{v}_2' + \cdots + \boldsymbol{v}_k\boldsymbol{v}_k' \\
&\quad\ \vdots \\
\mathbf{M}_k &= \qquad\qquad\qquad\quad \boldsymbol{v}_k\boldsymbol{v}_k'
\end{aligned}$$

such that $\mathbf{M}_1 \succeq \mathbf{M}_2 \succeq \cdots \succeq \mathbf{M}_1 \succeq 0$. Let $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ be an orthonormal basis for Span$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ such that, for all $i = 1, \ldots, k$, Span$\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_i\}$ = Span$\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_i\}$. Then $\mathbf{M}_i$ for $i = 1, \ldots, k$ need not be almost commuting, and $(\bar{\boldsymbol{X}}_i, \ldots, \bar{\boldsymbol{X}}_k) = (\boldsymbol{u}_1\boldsymbol{u}_1', \ldots, \boldsymbol{u}_k\boldsymbol{u}_k')$ is the optimal SDP solution with optimal value $p = \operatorname{tr}(\mathbf{M}_1)$.

Proof. $\bar{\boldsymbol{X}}_i$ are clearly feasible with objective value

$$\begin{aligned}
p &= \langle \mathbf{M}_1, \boldsymbol{u}_1\boldsymbol{u}_1' \rangle + \langle \mathbf{M}_2, \boldsymbol{u}_2\boldsymbol{u}_2' \rangle + \cdots + \langle \mathbf{M}_k, \boldsymbol{u}_k\boldsymbol{u}_k' \rangle && (32) \\
&= \sum_{i=1}^{k} (\boldsymbol{v}_i'\boldsymbol{u}_1)^2 + \sum_{i=2}^{k} (\boldsymbol{v}_i'\boldsymbol{u}_2)^2 + \cdots + \sum_{i=k-1}^{k} (\boldsymbol{v}_i'\boldsymbol{u}_{k-1})^2 + (\boldsymbol{v}_k'\boldsymbol{u}_k)^2 && (33) \\
&= \sum_{i=1}^{k} \|\boldsymbol{v}_i\|_2^2 = \operatorname{tr}(\mathbf{M}_1). && (34)
\end{aligned}$$

For any feasible solution, we have

$$\sum_{i=1}^{k} \langle \mathbf{M}_i, \boldsymbol{X}_i \rangle \leq \sum_{i=1}^{k} \langle \mathbf{M}_1, \boldsymbol{X}_i \rangle = \langle \mathbf{M}_1, \sum_{i=1}^{k} \boldsymbol{X}_i \rangle \leq \langle \mathbf{M}_1, \mathbf{I} \rangle = \operatorname{tr}(\mathbf{M}_1),$$

since $\mathbf{M}_1 \succcurlyeq \mathbf{M}_i$ for all i and $\sum_{i=1}^{k} \boldsymbol{X}_i \preccurlyeq \mathbf{I}$. So $\bar{\boldsymbol{X}}_i$ are optimal.

We next consider a rank-2 case to show the $\mathbf{M}_i$ need not be almost commuting. From the construction above, represent $\mathbf{M}_1 = \boldsymbol{v}_1\boldsymbol{v}_1' + \boldsymbol{v}_2\boldsymbol{v}_2'$ and $\mathbf{M}_2 = \boldsymbol{v}_2\boldsymbol{v}_2'$ for some $\boldsymbol{v}_1 = \gamma\boldsymbol{u}_1$ and $\boldsymbol{v}_2 = \alpha\boldsymbol{u}_1 + \beta\boldsymbol{u}_2$ and coefficients γ, α, β. It is easy to show $\|\mathbf{M}_1\mathbf{M}_2 - \mathbf{M}_2\mathbf{M}_1\|_2 = \boldsymbol{v}_2'\boldsymbol{v}_1 \|\boldsymbol{v}_1\boldsymbol{v}_2' - \boldsymbol{v}_2\boldsymbol{v}_1'\|_2 = \gamma^2\alpha\beta\|\boldsymbol{u}_1\boldsymbol{u}_2' - \boldsymbol{u}_2\boldsymbol{u}_1'\|_2 \leq \gamma^2\alpha\beta(\|\boldsymbol{u}_1\boldsymbol{u}_2'\|_2 + \|\boldsymbol{u}_2\boldsymbol{u}_1'\|_2) = 2\gamma^2\alpha\beta$. Since there could exist $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ such that the commutator norm is as large as $2\gamma^2\alpha\beta$, and unless one or more of the coefficients is small, then $\mathbf{M}_1$ and $\mathbf{M}_2$ need not be almost commuting. □

G Extended Experiments

G.1 Assessing the ROP: random PSD $\mathbf{M}_i$

For $\mathbf{M}_i$ that are random PSD matrices of rank k, we generate the matrix $\boldsymbol{A} \in \mathbb{R}^{d\times k}$ with i.i.d. Gaussian samples and compute $\mathbf{M}_i = \boldsymbol{A}\boldsymbol{A}'$.

		Fraction of 100 trials with ROP			
		$k=3$	$k=5$	$k=7$	$k=10$
RandPSD	$d=10$	0.97	0.61	0.3	0.14
	$d=20$	0.92	0.48	0.13	0
	$d=30$	0.93	0.53	0.14	0
	$d=40$	0.92	0.45	0.04	0
	$d=50$	0.95	0.53	0.05	0

Table 3: Numerical experiments showing the percentage of trials where the SDP was tight for random synthetic PSD $\mathbf{M}_i$.

G.2 Assessing the ROP: HPPCA

Table 4 and Table 5 display the full experiment results of their abbreviated versions–Table 1 and Table 2–in Section 5 of the main paper.

G.3 Assessing global optimality of local solutions

Further experiment details For 100 random experiments of each choice of σ, we obtain candidate solutions $\bar{\boldsymbol{X}}_i$ from the SDP and perform a rank-one SVD of each to form $\bar{\boldsymbol{U}}_{\text{SDP}}$, i.e.

$$\bar{\boldsymbol{U}}_{\text{SDP}} = [\bar{\boldsymbol{u}}_1 \cdots \bar{\boldsymbol{u}}_k], \quad \bar{\boldsymbol{u}}_i = \underset{\boldsymbol{u}:\|\boldsymbol{u}\|_2=1}{\text{argmax}}\ \boldsymbol{u}'\bar{\boldsymbol{X}}_i\boldsymbol{u},$$

while measuring how close the solutions are to being rank-1. In the case the SDP is not tight, the rank-1 directions of the $\boldsymbol{X}_i$ will not be orthonormal, so as a heuristic, we project $\bar{\boldsymbol{U}}_{\text{SDP}}$ onto the Stiefel manifold by its QR decomposition. For comparison, we use the Stiefel majorization-minimization (StMM) solver with a linear majorizer [14] to obtain a candidate solution $\bar{\boldsymbol{U}}_{\text{MM}}$ and use Theorem 4.1 to certify it either as a globally optimal or as a stationary point.

When executing each algorithm in practice, we remark that the results may vary with the choice of user specified numerical tolerances and other settings. For the StMM algorithm, we choose a random initialization of $\boldsymbol{U}$ each trial and run the algorithm either for specified maximum number of iterations or until the gradient on the Stiefel manifold is less than some tolerance threshold; here we set `tol` $= 10^{-10}$. Using MATLAB's CVX implementation to solve (SDP-P) and (5), we found setting `cvx_precision` to `high` guarantees the best results for returning tight solutions and verifying global optimality. However, iterates of the StMM algorithm that converge close to a tight SDP solution may still not be sufficient for the feasibility LMI to return a positive certificate if the solution is not numerically optimal to a high level of precision.

H Extension to the sum of Brocketts with linear terms

Given coefficient matrices and vectors $\{(\mathbf{M}_i, \boldsymbol{c}_i)\}_{i=1}^k$, suppose the problem in (1) is augmented with linear terms giving the following optimization problem that appears in [14]:

$$\max_{\boldsymbol{U}:\boldsymbol{U}'\boldsymbol{U}=\mathbf{I}} \sum_{i=1}^{k} \boldsymbol{u}_i'\mathbf{M}_i\boldsymbol{u}_i + \boldsymbol{c}_i'\boldsymbol{u}_i. \tag{35}$$

		Fraction of 100 trials with ROP			
		$k=3$	$k=5$	$k=7$	$k=10$
$\boldsymbol{n}=[5,20]$	$d=10$	1	0.99	1	1
	$d=20$	1	0.98	0.98	0.99
	$d=30$	0.99	0.93	0.98	0.97
	$d=40$	0.98	0.91	0.99	0.98
	$d=50$	0.97	0.95	0.96	0.98
$\boldsymbol{n}=[10,40]$	$d=10$	1	1	0.99	1
	$d=20$	1	1	0.98	0.99
	$d=30$	1	0.99	0.99	0.96
	$d=40$	0.98	0.97	0.92	0.96
	$d=50$	0.99	0.96	0.98	0.88
$\boldsymbol{n}=[20,80]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	1	1	1	0.98
	$d=40$	1	1	0.97	0.95
	$d=50$	1	0.98	0.98	0.97
$\boldsymbol{n}=[50,200]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	1	1	1	1
	$d=40$	1	1	0.99	1
	$d=50$	1	1	0.98	1
$\boldsymbol{n}=[100,400]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	1	1	1	1
	$d=40$	1	1	1	1
	$d=50$	1	1	1	1

Table 4: **(HPPCA)** Numerical experiments showing the percentage of trials where the SDP was tight for instances of the HPPCA problem as we vary d, k, and $\boldsymbol{n}$ using $L=2$ groups with noise variances $\boldsymbol{v}=[1,4]$.

		Fraction of 100 trials with ROP			
		$k=3$	$k=5$	$k=7$	$k=10$
$\boldsymbol{v}=[1,1]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	1	1	1	1
	$d=40$	1	1	1	1
	$d=50$	1	1	1	1
$\boldsymbol{v}=[1,2]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	1	0.98	1	1
	$d=40$	1	1	0.99	1
	$d=50$	1	1	1	0.99
$\boldsymbol{v}=[1,3]$	$d=10$	1	1	1	1
	$d=20$	1	1	1	1
	$d=30$	0.99	0.99	0.97	0.99
	$d=40$	1	0.98	0.97	0.99
	$d=50$	1	0.97	0.96	0.98
$\boldsymbol{v}=[1,4]$	$d=10$	1	1	0.99	1
	$d=20$	1	1	0.98	0.99
	$d=30$	1	0.99	0.99	0.96
	$d=40$	0.98	0.97	0.92	0.96
	$d=50$	0.99	0.96	0.98	0.88

Table 5: **(HPPCA)** Numerical experiments showing the percentage of trials where the SDP was tight for instances of the HPPCA problem as we vary d, k, and $\boldsymbol{v}$ using $L=2$ groups with samples $\boldsymbol{n}=[10,40]$.

It is then easy to see that for the matrices

$$\tilde{\mathbf{M}}_i := \begin{bmatrix} \mathbf{M}_i & \boldsymbol{c}_i \\ \boldsymbol{c}_i' & 0 \end{bmatrix}, \qquad \tilde{\boldsymbol{X}}_i := \begin{bmatrix} \boldsymbol{X}_i & \boldsymbol{u}_i \\ \boldsymbol{u}_i' & 1 \end{bmatrix}, \qquad \boldsymbol{X}_i := \boldsymbol{u}_i \boldsymbol{u}_i' \tag{36}$$

that $\sum_{i=1}^k \boldsymbol{u}_i' \mathbf{M}_i \boldsymbol{u}_i + \boldsymbol{c}_i' \boldsymbol{u}_i = \langle \tilde{\mathbf{M}}_i, \tilde{\boldsymbol{X}}_i \rangle$. Define $\mathbf{A} := [\mathbf{I}_d \ \mathbf{0}_d']' \in \mathbb{R}^{(d+1)\times d}$ and $\boldsymbol{e}_{d+1}$ to be the d+1-standard basis vector in $\mathbb{R}^{d+1}$. Extending (SDP-P), we obtain a generalized relaxation for the problem with linear terms:

$$\max_{\tilde{\boldsymbol{X}}_i} \sum_{i=1}^k \langle \tilde{\mathbf{M}}_i, \tilde{\boldsymbol{X}}_i \rangle \tag{37}$$

$$\text{s.t.} \quad \mathbf{A}' \sum_{i=1}^k \tilde{\boldsymbol{X}}_i \mathbf{A} \preccurlyeq \mathbf{I} \tag{38}$$

$$\langle \mathbf{A}\mathbf{A}', \tilde{\boldsymbol{X}}_i \rangle = 1, \qquad \boldsymbol{e}_{d+1}' \tilde{\boldsymbol{X}}_i \boldsymbol{e}_{d+1} = 1 \qquad \tilde{\boldsymbol{X}}_i \succeq 0. \tag{39}$$

By the Schur complement, the constraint $\tilde{\boldsymbol{X}}_i \succeq 0$ guarantees that $\boldsymbol{X}_i - \boldsymbol{u}_i \boldsymbol{u}_i' \succeq 0$ and therefore also $\boldsymbol{X}_i \succeq 0$. The linear operator $\mathbf{A}$ acts to impose the relevant Fantope-like constraints onto the top-left $d \times d$ positions of the primal variables, and the added constraint on the $(d+1, d+1)^{\text{th}}$ element of each $\tilde{\boldsymbol{X}}_i$ forces it to be 1. For dual variables $\tilde{\boldsymbol{Z}}_i \in \mathbb{S}_+^{d+1}$, $\boldsymbol{Y} \in \mathbb{S}_+^d$, $\boldsymbol{\nu} \in \mathbb{R}^k$, and $\xi \in \mathbb{R}$, the KKT conditions are

$$\tilde{\boldsymbol{X}}_i \succeq 0, \quad \mathbf{A}' \sum_{i=1}^k \bar{\boldsymbol{X}}_i \mathbf{A} \preceq \mathbf{I}, \quad \langle \mathbf{A}\mathbf{A}', \tilde{\boldsymbol{X}}_i \rangle = 1, \quad \boldsymbol{e}_{d+1}' \tilde{\boldsymbol{X}}_i \boldsymbol{e}_{d+1} = 1 \tag{40}$$

$$\mathbf{A}\boldsymbol{Y}\mathbf{A}' = \tilde{\mathbf{M}}_i + \tilde{\boldsymbol{Z}}_i - \bar{\nu}_i \mathbf{A}\mathbf{A}' - \xi \boldsymbol{e}_{d+1}\boldsymbol{e}_{d+1}', \qquad \boldsymbol{Y} \succeq 0 \tag{41}$$

$$\langle \mathbf{I} - \mathbf{A}' \sum_{i=1}^k \bar{\boldsymbol{X}}_i \mathbf{A}, \boldsymbol{Y} \rangle = 0 \tag{42}$$

$$\langle \tilde{\boldsymbol{Z}}_i, \tilde{\boldsymbol{X}}_i \rangle = 0 \tag{43}$$

$$\tilde{\boldsymbol{Z}}_i \succeq 0, \tag{44}$$

which, in fact, are the same KKT conditions as before. If we denote $\boldsymbol{Z}_i := \mathbf{A}' \tilde{\boldsymbol{Z}}_i \mathbf{A}$ to be the top $d+1 \times d+1$ positions of $\tilde{\boldsymbol{Z}}_i$, multiplying (41) by $\mathbf{A}'$ on the left and $\mathbf{A}$ on the right gives back exactly (KKT-b) for the relaxation in (SDP-P).

References

[A1] Podosinnikova, A., Perry, A., Wein, A., Bach, F., D'Aspremont, A. & Sontag, D. Overcomplete independent component analysis via SDP. *The 22nd International Conference On Artificial Intelligence And Statistics*. pp. 2583-2592 (2019)

[A2] Journée, M., Bach, F., Absil, P. & Sepulchre, R. Low-rank optimization on the cone of positive semidefinite matrices. *SIAM Journal On Optimization*. **20**, 2327-2351 (2010)

[A3] Boumal, N., Voroninski, V. & Bandeira, A. Deterministic Guarantees for Burer-Monteiro Factorizations of Smooth Semidefinite Programs. *Communications On Pure And Applied Mathematics*. **73**, 581-608 (2020)

[A4] Bandeira, A., Boumal, N. & Singer, A. Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. *Mathematical Programming*. **163**, 145-167 (2017)

[A5] Zhou, F. & Low, S. Conditions for Exact Convex Relaxation and No Spurious Local Optima. *IEEE Transactions On Control Of Network Systems*. (2021)

[A6] Burer, S. & Monteiro, R. Local minima and convergence in low-rank semidefinite programming. *Mathematical Programming.* **103**, 427-444 (2005)

[A7] Cifuentes, D. & Moitra, A. Polynomial time guarantees for the Burer-Monteiro method. *ArXiv Preprint ArXiv:1912.01745.* (2019)