SIAM J. MATRIX ANAL. APPL.
Vol. 46, No. 2, pp. 1091–1116

A SEMIDEFINITE RELAXATION FOR SUMS OF HETEROGENEOUS QUADRATIC FORMS ON THE STIEFEL MANIFOLD*

KYLE GILMAN†, SAMUEL BURER‡, AND LAURA BALZANO†

Abstract. We study the maximization of sums of heterogeneous quadratic forms 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 of the original problem 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 to a simple feasibility problem to certify global optimality of a candidate 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.

Key words. semidefinite programming, heteroscedastic probabilistic principal component analysis, Stiefel manifold optimization

MSC codes. 62H25, 90C22, 90C26, 90C46

DOI. 10.1137/23M1545136

1. Introduction.

This paper studies the problem known in the literature as *the maximization of sums of heterogeneous quadratic forms over the Stiefel manifold*[1] [6, 7, 14, 35]. Specifically, given $d \times d$ symmetric positive semidefinite (PSD) matrices $\mathbf{M}_1, \ldots, \mathbf{M}_k \succeq 0$ for $k < d$, we wish to maximize the convex objective function $\sum_{i=1}^k \mathbf{u}_i' \mathbf{M}_i \mathbf{u}_i$ over the nonconvex constraint that $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_k] \in \mathbb{R}^{d \times k}$ has orthonormal columns:

$$\max_{\mathbf{U} \in \mathrm{St}(k,d)} \sum_{i=1}^k \mathbf{u}_i' \mathbf{M}_i \mathbf{u}_i, \tag{1.1}$$

where $\mathrm{St}(k,d) = \{\mathbf{U} \in \mathbb{R}^{d \times k} : \mathbf{U}'\mathbf{U} = \mathbf{I}_k\}$ denotes the Stiefel manifold. This problem arises in modern signal processing and machine learning applications like heteroscedastic probabilistic principal component analysis (HPPCA) [24, 25], heterogeneous clutter in radar sensing [37], and robust sparse PCA [13]. Each of these applications involves learning a signal subspace for data possessing heterogeneous statistics.

*Received by the editors January 25, 2023; accepted for publication (in revised form) by F. Krahmer October 31, 2024; published electronically April 17, 2025.
https://doi.org/10.1137/23M1545136

Funding: The first and third authors were supported in part by NSF CAREER award CCF-1845076, ARO YIP award W911NF1910027, AFOSR YIP award FA9550-19-1-0026, and NSF BIGDATA award IIS-1838179. The third author was also supported by NSF award CCF-2331590.

†Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI 48109 USA (kgilman@umich.edu, girasole@umich.edu).

‡Department of Business Analytics, University of Iowa, Iowa City, IA 52242-1994 USA (samuel-burer@uiowa.edu).

[1]We note here that "heterogeneous" refers to the fact that the $\mathbf{M}_i$ are distinct and the problem is not separable in each $\mathbf{u}_i$. Indeed, the objective in (1.1) is a homogeneous polynomial in the entries of $\mathbf{U}$ since all terms are degree 2.

In particular, HPPCA [25] 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, we are given L data groups $(\mathbf{Y}_1, \ldots, \mathbf{Y}_L)$, where each $\mathbf{Y}_\ell \in \mathbb{R}^{d \times n_\ell}$ represents a matrix of n_ℓ samples of a d-dimensional signal plus additive Gaussian noise with variance v_ℓ. Using second-order statistics $\mathbf{A}_\ell := \frac{1}{v_\ell}\mathbf{Y}_\ell\mathbf{Y}_\ell' \succeq 0$ 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 $\mathbf{U}$ and can be equivalently recast in the form (1.1) as follows:

(1.2)
$$\max_{\mathbf{U}:\mathbf{U}'\mathbf{U}=\mathbf{I}} \sum_{\ell=1}^{L} \operatorname{tr}(\mathbf{U}'\mathbf{A}_\ell\mathbf{U}\mathbf{W}_\ell) = \max_{\mathbf{U}:\mathbf{U}'\mathbf{U}=\mathbf{I}} \sum_{\ell=1}^{L}\sum_{i=1}^{k} w_{\ell,i}\mathbf{u}_i'\mathbf{A}_\ell\mathbf{u}_i = \max_{\mathbf{U}:\mathbf{U}'\mathbf{U}=\mathbf{I}} \sum_{i=1}^{k} \mathbf{u}_i'\mathbf{M}_i\mathbf{u}_i,$$

where $\mathbf{W}_\ell := \operatorname{diag}(\{w_{\ell,i}\}_{i=1}^k)\ \forall\ \ell$ and $\mathbf{M}_i := \sum_{\ell=1}^{L} w_{\ell,i}\mathbf{A}_\ell\ \forall\ i$. Other sensing problems such as independent component analysis (ICA) [39] and approximate joint diagonalization (AJD) [33] also model data with heterogeneous statistics and optimize objective functions of a similar form, as we discuss in section 3.

For (1.2), the case of a single Brockett cost function ($L = 1$) has a known analytical solution 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. Given that (1.1) is nontrivial and challenging in several ways—nonconvex due to the Stiefel manifold constraint, nonseparable because of the weighted sum of objectives, and not readily solved by singular value or eigenvalue decomposition—many works apply iterative local solvers to (1.1).

However, given the nonconvexity of (1.1), these local approaches do not find a global maximum in general. An alternative approach is to relax problems such as (1.1) 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 an 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 (1.1).

The main contribution of this paper is a novel convex SDP relaxation of (1.1), whose constraint set is related to the Fantope but distinct. By studying this SDP and its optimality criteria, we derive sufficient conditions to certify the global optimality of any 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 dimension. Our work also generalizes existing results for (1.1) 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 (1.1).

Notation. We use boldface uppercase letters $\mathbf{A}$ to denote matrices, boldface lowercase letters $\mathbf{v}$ to denote vectors, and italic lowercase letters c for scalars. We denote

the cone of $d \times d$ symmetric PSD matrices as $\mathbb{S}_+^d$ and use $\mathbf{A} \succeq 0$ to denote an element $\mathbf{A} \in \mathbb{S}_+^d$. We denote the Hermitian transpose of a matrix as $\mathbf{A}'$, the trace of a matrix as $\operatorname{tr}(\mathbf{A})$, and the inner product of matrices (with identical dimensions) $\langle \mathbf{A}, \mathbf{B} \rangle := \operatorname{tr}(\mathbf{A}'\mathbf{B})$. We also make use of the notation $[\mathbf{A}, \mathbf{B}] = 0$ for commuting square matrices $\mathbf{A}$ and $\mathbf{B}$ of the same sizes, which is equivalent to $\mathbf{AB} - \mathbf{BA} = 0$, where here 0 is the zero matrix. The spectral norm of a matrix is denoted by $\|\mathbf{A}\|$, the Frobenius norm by $\|\mathbf{A}\|_{\mathrm{F}}$, and the trace norm by $\|\mathbf{A}\|_{\mathrm{tr}} := \sqrt{\frac{1}{d}\sum_{i,j=1}^{d} |\mathbf{A}_{i,j}|^2} = \frac{1}{\sqrt{d}}\|\mathbf{A}\|_{\mathrm{F}}$. The identity matrix of size d is denoted as $\mathbf{I}_d$. Finally, we denote $[k] := \{1, \ldots, k\}$.

2. SDP relaxation. By relaxing the considered nonconvex problem (1.1) 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 (1.1) 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 (1.1) and the Stiefel manifold constraints as

$$\max_{\mathbf{u}_1,\ldots,\mathbf{u}_k} \operatorname{tr}\left(\sum_{i=1}^{k} \mathbf{M}_i \mathbf{u}_i \mathbf{u}_i'\right) \quad \text{s.t. } \operatorname{tr}(\mathbf{u}_i \mathbf{u}_i') = 1 \quad \forall i \in [k], \quad \operatorname{tr}(\mathbf{u}_j \mathbf{u}_i') = 0 \quad \forall i \neq j. \tag{2.1}$$

Letting $\mathbf{X}_i = \mathbf{u}_i \mathbf{u}_i' \in \mathbb{R}^{d \times d}$, and using the eigenvalue structure of the rank-k projection matrix $\sum_{i=1}^{k} \mathbf{u}_i \mathbf{u}_i'$, this is equivalent to the lifted problem:

$$\begin{aligned} \max_{\mathbf{X}_1,\ldots,\mathbf{X}_k} \operatorname{tr}\left(\sum_{i=1}^{k} \mathbf{M}_i \mathbf{X}_i\right) \quad \text{s.t. } \lambda_j\left(\sum_{i=1}^{k} \mathbf{X}_i\right) \in \{0,1\} \quad \forall j \in [d] \\ \operatorname{tr}(\mathbf{X}_i) = 1, \quad \operatorname{rank}(\mathbf{X}_i) = 1, \quad \mathbf{X}_i \succcurlyeq 0 \quad \forall i \in [k], \end{aligned} \tag{2.2}$$

where $\lambda_j(\cdot)$ indicates the jth 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 [30, 40], we relax the eigenvalue constraint in (2.2) to $0 \preccurlyeq \sum_{i=1}^{k} \mathbf{X}_i \preccurlyeq \mathbf{I}$ and remove the rank constraint, which yields the SDP relaxation we consider throughout the remainder of this work:

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

Note that $0 \preccurlyeq \sum_{i=1}^{k} \mathbf{X}_i$ can be omitted since it is already satisfied when $\mathbf{X}_i \succcurlyeq 0 \ \forall \ i$.

The feasible set of (SDP-P) is closely related to the convex set found in [22, 30, 40] called the *Fantope*. The Fantope is the convex hull of all matrices $\mathbf{UU}' \in \mathbb{R}^{d \times d}$ such that $\mathbf{U} \in \mathbb{R}^{d \times k}$ and $\mathbf{U}'\mathbf{U} = \mathbf{I}$ [20, 31]. Indeed, our relaxation can be viewed as providing a decomposition of the Fantope variable $\mathbf{X} = \mathbf{UU}'$ into the sum $\mathbf{X}_1 + \cdots + \mathbf{X}_k$ such that each $\mathbf{X}_i$ satisfies $\operatorname{tr}(\mathbf{X}_i) = 1$ and $0 \preceq \mathbf{X}_i \preceq \mathbf{I}$. This decomposition allows (SDP-P) to capture the exact form of the objective function, which sums the individual terms $\operatorname{tr}(\mathbf{M}_i \mathbf{X}_i)$.

Precisely, the feasible set of (SDP-P) is a convex relaxation of the set $\{(\mathbf{u}_1\mathbf{u}_1', \ldots, \mathbf{u}_k\mathbf{u}_k') : \mathbf{U}'\mathbf{U} = \mathbf{I}\}$. Naturally, one wonders whether our relaxation always

solves the original nonconvex problem. We show in section SM4.2 in the supplement that it does not, using a counter example that demonstrates our relaxation does not exactly capture the convex hull, which is a necessary condition for the relaxation to be tight for all objectives. 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 $\mathbf{Z}_i \in \mathbb{S}^d_+$ for $i \in [k]$, $\mathbf{Y} \in \mathbb{S}^d_+$, $\boldsymbol{\nu} \in \mathbb{R}^k$, the dual of (SDP-P), which will play a central role in the theory of this paper, is

(SDP-D)

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

The derivation of (SDP-D) in subsection SM2.1 in the supplement follows by standard analysis of the Lagrangian. However, a short proof of weak duality also verifies that (SDP-D) upper bounds (SDP-P):

$$\begin{aligned}\sum_{i=1}^k \operatorname{tr}(\mathbf{M}_i \mathbf{X}_i) &= \sum_{i=1}^k \operatorname{tr}((\mathbf{Y} - \mathbf{Z}_i + \nu_i \mathbf{I})\mathbf{X}_i) \\ &= \operatorname{tr}\left(\mathbf{Y} \sum_{i=1}^k \mathbf{X}_i\right) - \sum_{i=1}^k \operatorname{tr}(\mathbf{Z}_i \mathbf{X}_i) + \sum_{i=1}^k \nu_i \operatorname{tr}(\mathbf{X}_i) \\ &\le \operatorname{tr}(\mathbf{Y}) + \sum_{i=1}^k \nu_i,\end{aligned}$$

where the inequality follows from the constraints in (SDP-P) and (SDP-D). Therefore $p^* \le d^*$. Since the constraint set of (SDP-P) is closed and bounded with nonempty interior, and strong duality holds by the following lemma, then there exists an optimal primal solution to (SDP-P) and optimal dual solution to (SDP-D).

LEMMA 2.1. *If $k < d$, strong duality holds for the SDP relaxation with primal* (SDP-P) *and dual* (SDP-D).

The proof of this lemma follows from Slater's condition and can be found in Appendix A.

We now define the "rank-one property" of a feasible solution of (SDP-P), which allows us to characterize the relationship between optimal solutions of (SDP-P) and optimal solutions of the original nonconvex problem.

DEFINITION 2.2 (rank-one property (ROP)). *A feasible solution to* (SDP-P) *is said to have the ROP if $\mathbf{X}_1, \dots, \mathbf{X}_k$ are all rank-one.*

We note that if a feasible solution has the ROP, the first singular vectors of the $\mathbf{X}_i$ are necessarily mutually orthogonal, and $\sum_i \mathbf{X}_i$ is a rank-k projection matrix, due to the constraint $\sum_i \mathbf{X}_i \preccurlyeq \mathbf{I}$. The following lemma establishes the relationship between the properties of the optimal solutions of (SDP-P) to those of the original nonconvex problem.

LEMMA 2.3. *An optimal solution $\mathbf{X}^* := (\mathbf{X}_1^*, \dots, \mathbf{X}_k^*)$ to the SDP relaxation in* (SDP-P) *is an optimal solution to the original nonconvex problem in* (1.1) *(equivalently* (2.2)*) if and only if $\mathbf{X}^*$ has the ROP.*

The proof of this lemma can be found in Appendix A. The next lemma now relates the properties of the optimal solutions to (SDP-D) to optimal solutions of (SDP-P) with the ROP.

LEMMA 2.4. *If the optimal dual variables* $\mathbf{Z}_i^*$ *for* $i = 1, \ldots, k$ *each have rank* $d-1$*, the optimal solution* $\mathbf{X}^* := (\mathbf{X}_1^*, \ldots, \mathbf{X}_k^*)$ *has the ROP.*

The proof of this result is also in Appendix A, and it follows directly from complementary slackness. This key result, through careful analysis of the dual problem, will later allow us to characterize problem instances with ROP solutions, which by Lemma 2.3, are optimal solutions to the nonconvex problem.

3. Related work. There are a few important related works on the objective in (1.1), as well as many more than can be reviewed here, including ones on eigenvalue/eigenvector problems and their variations, low-rank SDPs, and nonconvex quadratics where $\mathbf{M}_i$ are not PSD. For the curious reader, section SM1 in the supplement provides a more extensive related work section. Here, we focus on the works most directly related to (1.1).

The papers [6, 7, 35] previously investigated the sum of heterogeneous quadratic forms in (1.1). The work in [7] only studied the structure of this problem when the matrices $\mathbf{M}_i$ were commuting. The work in [35] derived sufficient second-order global optimality conditions, but these conditions are difficult to check in general and, for example, do not seem to hold for the heteroscedastic PCA problem. Works such as [27] and [32] consider a very similar problem to (1.2), but without the eigenvalue constraint in (2.2), making their SDP a rank-constrained separable SDP; see also [30, section 4.3]. Pataki [32] 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 $\mathbf{X}_i$, Pataki's bounds do not guarantee rank-one, or even low-rank, optimal solutions.

A recent paper [16] analyzes general sufficient conditions under which an SDP relaxation, which has a rank-one optimal solution, retains a rank-one optimal solution after the perturbation of the objective and/or constraint data. The analysis in [16] does not seem to apply directly to our own work for two reasons: (i) the authors of [16] analyze the basic Shor relaxation, a natural and popular SDP relaxation for quadratically constrained quadratic programs, which we show in subsection SM4.1 in the supplement is trivially not tight in our setting; and (ii) their relaxation has a single-block matrix variable, which is analyzed to be rank-one at optimality, whereas we analyze several blocks $\mathbf{X}_1, \ldots, \mathbf{X}_k$, each of which is rank-one at optimality when the SDP is tight.

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 [38, 40, 42], rather than the sum of multiple matrices as in our setting. The works in [9, 34] show that nonconvex Burer–Monteiro factorizations [15], which solve low-rank SDPs without orthogonality constraints, have no spurious local minima and that approximate second-order stationary points are approximate global optima. Other works have studied algorithms to optimize the nonconvex problem, like those in [11, 12, 13, 25, 37], using minorize-maximize or Riemannian gradient ascent algorithms, which do not come with global optimality guarantees. Our problem also has interesting connections to AJD, which is well-studied and often applied to blind source separation or ICA problems [3, 8, 28, 36, 39]. See section SM1 in 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 [1, 2, 18] to majorization-minimization (MM) algorithms, where [13] applied MM methods to solve (1.1) 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. We seek to fill this gap by proposing a check for global optimality of a local solution.[2] Similar types of problems for running fast probabilistic algorithms and checking whether the candidate solution is the optimal solution to the convex relaxation also appear in [4].

By Lemma 2.3, an optimal solution of (SDP-P) with rank-one matrices $\mathbf{X}_i$ globally solves the original nonconvex problem (1.1). In this section, given a candidate $\overline{\mathbf{U}} = [\overline{\mathbf{u}}_1 \cdots \overline{\mathbf{u}}_k] \in \mathrm{St}(k,d)$ to (1.1), we investigate conditions guaranteeing that the rank-one matrices $\overline{\mathbf{X}}_i = \overline{\mathbf{u}}_i \overline{\mathbf{u}}_i'$, which are feasible for (SDP-P), in fact comprise an optimal solution of (SDP-P), implying that $\overline{\mathbf{U}}$ optimizes (1.1). Similar to [19, 41, 42] for Fantope problems, our results yield a dual SDP certificate to verify the primal optimality of the candidates $\overline{\mathbf{X}}_1, \ldots, \overline{\mathbf{X}}_k$ constructed from a local solution $\overline{\mathbf{U}}$. We show our certificate scales favorably in computation compared to the full SDP, with the most complicated computations of our algorithm requiring us to solve a feasibility problem in k variables with several $d \times d$ linear matrix inequalities (LMIs).

THEOREM 4.1. *Let* $\overline{\mathbf{U}} \in \mathrm{St}(k,d)$, *and let* $\overline{\boldsymbol{\Lambda}} = \mathrm{sym}(\sum_{i=1}^k \overline{\mathbf{U}}' \mathbf{M}_i \overline{\mathbf{U}} \mathbf{E}_i)$, *where* $\mathrm{sym}(\mathbf{A}) := \frac{1}{2}(\mathbf{A} + \mathbf{A}')$, $\mathbf{E}_i \triangleq \mathbf{e}_i \mathbf{e}_i'$, *where* $\mathbf{e}_i$ *is the* i*th standard basis vector in* $\mathbb{R}^k$, *and* $\mathbf{M}_i \succeq 0 \ \forall\ i \in [k]$. *If there exist* $\overline{\boldsymbol{\nu}} = [\overline{\nu}_1 \cdots \overline{\nu}_k] \in \mathbb{R}^k$ *such that*

$$\begin{aligned} \overline{\mathbf{U}}(\overline{\boldsymbol{\Lambda}} - \mathbf{D}_{\overline{\boldsymbol{\nu}}})\overline{\mathbf{U}}' + \overline{\nu}_i \mathbf{I} - \mathbf{M}_i \succeq 0 \quad \forall i = 1, \ldots, k \\ \overline{\boldsymbol{\Lambda}} - \mathbf{D}_{\overline{\boldsymbol{\nu}}} \succeq 0, \end{aligned} \tag{4.1}$$

where $\mathbf{D}_{\overline{\boldsymbol{\nu}}} := \mathrm{diag}(\overline{\nu}_1, \ldots, \overline{\nu}_k)$, *then* $\overline{\mathbf{U}}$ *is a globally optimal solution to the original nonconvex problem* (1.1).

The proof, found in Appendix B, uses the Karush–Kuhn–Tucker (KKT) conditions along with the conditions on $\overline{\boldsymbol{\nu}}$ to construct a dual certificate of SDP optimality. We note that Theorem 4.1 is based on a strong sufficient condition, which in particular implies that any feasible $\overline{\mathbf{U}}$ satisfying (4.1) is a second-order stationary point.

In light of Theorem 4.1, to test whether a candidate $\overline{\mathbf{U}}$ is globally optimal, we simply assess whether system (4.1) is feasible using an LMI solver. If it is indeed feasible, then $\overline{\mathbf{U}}$ is globally optimal. On the other hand, if (4.1) is infeasible, it indicates one of two things: (i) The SDP is not tight, i.e., the SDP strictly upper bounds the original problem. The candidate $\overline{\mathbf{U}}$ may or may not be globally optimal to the original nonconvex problem. (ii) The SDP is tight, but the candidate $\overline{\mathbf{U}}$ is a suboptimal local solution. Section SM4.4 in the supplement also describes an extension of the certificate to the sum of Brocketts with additive linear terms.

It is important to note that Theorem 4.1 implies $\overline{\mathbf{U}}$ is an *exact* second-order stationary point. Since in practice it is not possible to obtain exact stationary points using numerical solvers, one may wonder if Theorem 4.1 can be applied in practice. However, given some $\overline{\mathbf{U}} \in \mathrm{St}(k,d)$ obtained by a solver that only approximately sat-

[2] To be clear, while our work does not guarantee that a local solution is globally optimal, we propose a certificate based on a sufficient condition to check if the local solution is globally optimal.

isfies dual feasibility, we can precisely characterize the suboptimality of this solution. To this end, we provide a corollary to Theorem 4.1, whose proof can be found in Appendix B, where the semidefinite constraints are only approximately satisfied.

COROLLARY 4.2. *Let* $\overline{\mathbf{U}} \in \mathrm{St}(k,d)$ *be a feasible point of* (1.1)*, and let* $\overline{\mathbf{\Lambda}} = \mathrm{sym}(\sum_{i=1}^{k} \overline{\mathbf{U}}'\mathbf{M}_i\overline{\mathbf{U}}\mathbf{E}_i)$*. Let* ϵ^* *be the optimal value of*

$$\min_{\epsilon\geq 0,\, \overline{\boldsymbol{\nu}}\in\mathbb{R}^k} \epsilon \quad \text{s.t.} \quad \overline{\mathbf{U}}(\overline{\mathbf{\Lambda}} - \mathbf{D}_{\overline{\boldsymbol{\nu}}})\overline{\mathbf{U}}' + \overline{\nu}_i\mathbf{I} - \mathbf{M}_i \succeq -\epsilon\mathbf{I} \quad \forall i = 1,\dots,k \qquad \overline{\mathbf{\Lambda}} - \mathbf{D}_{\overline{\boldsymbol{\nu}}} \succeq -\epsilon\mathbf{I}, \tag{4.2}$$

where $\mathbf{D}_{\overline{\boldsymbol{\nu}}} := \mathrm{diag}(\overline{\nu}_1,\dots,\overline{\nu}_k)$*. Then* $\overline{\mathbf{U}}$ *is a near optimal solution to the original nonconvex problem* (1.1) *in the sense that its objective value is bounded below by* $p^* - \epsilon^* d$*.*

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 of (SDP-D). The proposed global certificate in (4.1) significantly reduces the number of variables from $\mathcal{O}(d^2)$ in (SDP-D) (upon eliminating the variables $\mathbf{Z}_i$) to merely k variables in (4.1). Using [5, section 6.6.3] it can be shown that computing the certificate only, based on a given $\overline{\mathbf{U}}$, results in a substantial reduction in flops by a factor of $\mathcal{O}(d^3/k)$ over solving (SDP-D). Subsequently, an MM solver with complexity on par with standard first-order based methods [13], whose cost is $\mathcal{O}(dk^2 + k^3)$ per iteration, combined with our global optimality certificate, is preferable to solving the full relaxation (SDP-P) for large problems. See subsection SM2.3 in the supplement for more details.

4.2. SDP tightness in the close-to jointly diagonalizable 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 assignment problem [7], 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 $\mathbf{X}_i$ also exists 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 CJD and can be made arbitrarily close as the number of data points n grows or as the noise levels diminish or become homoscedastic. This gives strong theoretical support for the tightness of the SDP for the HPPCA problem when n is large or the noise levels are small or close in value.

DEFINITION 4.3 (CJD). *We say that unit spectral-norm, symmetric matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *are CJD if they are almost commuting, that is, when the commuting distance measured by some norm* $\|\cdot\|$*, between* $\mathbf{A}$ *and* $\mathbf{B}$ *is significantly less than* 1*:*

$$\|[\mathbf{A},\mathbf{B}]\| := \|\mathbf{AB} - \mathbf{BA}\| \leq \delta \qquad \text{for some } 0 < \delta \ll 1.$$

The matrices $\mathbf{A}$ and $\mathbf{B}$ are jointly diagonalizable if and only if they commute, i.e., the commuting distance is zero.

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 optimal solutions $\mathbf{X}_i$. All proofs for the results in this subsection are found in Appendix C.

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 PSD since the primal constraint $\operatorname{tr}(\mathbf{X}_i) = 1$ ensures that replacing $\mathbf{M}_i$ by $\mathbf{M}_i + \lambda_i \mathbf{I} \succeq 0$, where λ_i is a positive constant, simply shifts the objective value by λ_i. Thus, we assume $\mathbf{M}_i \succeq 0 \ \forall\ i = 1, \ldots, k$.

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

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

to be our set of admissible coefficient k-tuples. We know that both (SDP-P) and (SDP-D) have interior points for all $\mathbf{c} \in \mathcal{C}$, so that strong duality holds for all $\mathbf{c} \in \mathcal{C}$. The following results draw upon the fact that (SDP-P) is equivalent to a linear program (LP) when $\mathbf{M}_1, \ldots, \mathbf{M}_k$ are jointly diagonalizable, i.e., the problem is a diagonal SDP. While we require the assumption that the equivalent LP in the jointly diagonalizable case has a unique optimal solution, we find this is a reasonable, mild assumption based on [17, Theorem 4], which proves the uniqueness property holds generically for LPs.

LEMMA 4.4. *Let* $\mathbf{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) \in \mathcal{C}$. *If* $\mathbf{M}_i$ *are jointly diagonalizable for* $i = 1, \ldots, k$ *and the associated LP for* (SDP-P) *has a unique optimal solution, then there exists an optimal solution of* (SDP-D) *with* $\operatorname{rank}(\mathbf{Z}_i) \geq d - 1$ *for all* $i = 1, \ldots, k$.

The result follows directly from the Goldman–Tucker theorem on strict complementarity for LPs.

DEFINITION 4.5. *For* $\mathbf{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) \in \mathcal{C}$ *and* $\overline{\mathbf{c}} = (\overline{\mathbf{M}}_1, \ldots, \overline{\mathbf{M}}_k) \in \mathcal{C}$, *define* $\operatorname{dist}(\mathbf{c}, \overline{\mathbf{c}}) \triangleq \max_{i \in [k]} \|\mathbf{M}_i - \overline{\mathbf{M}}_i\|_{\mathrm{tr}}$.

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

THEOREM 4.6. *Let* $\overline{\mathbf{c}} := (\overline{\mathbf{M}}_1, \ldots, \overline{\mathbf{M}}_k) \in \mathcal{C}$ *be given such that* $\overline{\mathbf{M}}_i$, $i = 1, \ldots, k$, *are jointly diagonalizable and the associated LP, which is derived from the diagonal SDP of* (SDP-P) *with objective coefficients* $\overline{\mathbf{c}}$, *has a unique optimal solution. Then there exists a full-dimensional neighborhood* $\overline{\mathcal{C}} \ni \overline{\mathbf{c}}$ *in* $\mathcal{C}$ *such that* (SDP-P) *has the ROP for all* $\mathbf{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) \in \overline{\mathcal{C}}$.

Proof. Using Lemma 4.4, let $\mathbf{y}^0 := (\overline{\mathbf{Y}}, \overline{\mathbf{Z}}_i, \overline{\nu}_i)$ be the optimal solution of the dual problem (SDP-D) for $\overline{\mathbf{c}} = (\overline{\mathbf{M}}_1, \ldots, \overline{\mathbf{M}}_k)$, which has $\operatorname{rank}(\overline{\mathbf{Z}}_i) \geq d - 1 \ \forall\ i$. Proposition C.3 in Appendix C considers a function $y(\mathbf{c}; \mathbf{y}^0)$ that returns the optimal solution of (SDP-D) for $\mathbf{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k)$ closest to $\mathbf{y}^0$ and shows that this function is continuous. It follows that its preimage

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

contains $\overline{\mathbf{c}}$ and is an open set because the set of all $(\mathbf{Y}, \mathbf{Z}_i, \nu_i)$ with $\operatorname{rank}(\mathbf{Z}_i) \geq d - 1$ is an open set. After intersecting with $\mathcal{C}$, we have shown existence of this full-dimensional set $\overline{\mathcal{C}}$. From complementarity of the KKT conditions of the assignment LP, $\operatorname{rank}(\mathbf{Z}_i) = d - 1$ for $i = 1, \ldots, k$. Applying Lemma 2.4 then completes the theorem. ∎

The next corollary shows that for a general tuple of matrices $\mathbf{c} := (\mathbf{M}_1, \ldots, \mathbf{M}_k)$ that are pairwise CJD for small enough δ, (SDP-P) is tight and has the ROP.

In the following results, we will then prove the HPPCA generative model results in $(\mathbf{M}_1, \dots, \mathbf{M}_k)$ being CJD. While these are sufficient conditions, they are by no means necessary, and subsection SM4.4 in the supplement gives an example of $\mathbf{M}_i$ that are *not* CJD but where the convex relaxation has the ROP. It is important to note the results do not quantify an exact δ for (SDP-P) to achieve the ROP, but only the existence of one.

COROLLARY 4.7. *Let $\epsilon > 0$, and let $\mathbf{c} := (\mathbf{M}_1, \dots, \mathbf{M}_k)$ be a tuple of self-adjoint matrices, where $\|[\mathbf{M}_i, \mathbf{M}_j]\|_{\text{tr}} := \|\mathbf{M}_i\mathbf{M}_j - \mathbf{M}_j\mathbf{M}_i\|_{\text{tr}} \le \epsilon \ \forall \ i, j \in [k]$, and assume $\|\mathbf{M}_i\| \le 1 \ \forall \ i \in [k]$. Then there exists a tuple of commuting self-adjoint matrices $\overline{\mathbf{c}} := (\overline{\mathbf{M}}_1, \dots, \overline{\mathbf{M}}_k)$ with $[\overline{\mathbf{M}}_i, \overline{\mathbf{M}}_j] = 0 \ \forall \ i, j \in [k]$ and a $\delta(\epsilon, k) > 0$ such that $\text{dist}(\mathbf{c}, \overline{\mathbf{c}}) \le \delta(\epsilon, k)$ and $\delta(\epsilon, k)$ is a function satisfying $\lim_{\epsilon \to 0} \delta(\epsilon, k) = 0$. Assume the associated LP, which is derived from the diagonal SDP of* (SDP-P) *and is parameterized by $\bar{\mathbf{c}}$, has a unique optimal solution.*

If $\epsilon > 0$ is such that dist$(\mathbf{c}, \overline{\mathbf{c}}) \le \delta(\epsilon, k)$ implies $\mathbf{c} \in \overline{\mathcal{C}}$, where $\overline{\mathcal{C}}$ is given by Theorem 4.6*, (SDP-P) parameterized by $\mathbf{c}$ has the ROP.*

Proof. The result follows from directly applying the extension of Lin's theorem for a tuple of $k \ge 3$ matrices [21, Theorem 3] (see Lemma SM3.13 in the supplement) to $(\mathbf{M}_1, \dots, \mathbf{M}_k)$. ∎

The next corollary gives a similar result, but tailored specifically to problem (1.2).

COROLLARY 4.8. *Let $\epsilon > 0$, and define $\mathbf{c} := (\mathbf{M}_1, \dots, \mathbf{M}_k)$ for* (1.2)*, where $\|[\mathbf{A}_i, \mathbf{A}_j]\|_{\text{tr}} \le \epsilon \ \forall \ i, j \in [L]$, and assume $\|\mathbf{A}_i\| \le 1 \ \forall \ i \in [k]$. Then there exists a tuple of commuting self-adjoint matrices $\overline{\mathbf{c}} := (\overline{\mathbf{M}}_1, \dots, \overline{\mathbf{M}}_k)$ with $[\overline{\mathbf{M}}_i, \overline{\mathbf{M}}_j] = 0 \ \forall \ i, j \in [k]$ and a $\delta(\epsilon, k) > 0$ such that $\text{dist}(\mathbf{c}, \overline{\mathbf{c}}) \le \delta(\epsilon, k) \sum_{\ell=1}^{L} \max_{i \in [k]} w_{\ell,i}$.*

4.2.2. HPPCA possesses the CJD property. Consider the heteroscedastic probabilistic PCA problem in [25] where L data groups of $n_1, \dots, n_L$ samples $(n = \sum_{\ell=1}^{L} n_\ell)$ with known noise variances $v_1, \dots, v_L$, respectively, are generated by the model

$$\mathbf{y}_{\ell,j} = \mathbf{U}\mathbf{\Theta}\mathbf{z}_{\ell,j} + \boldsymbol{\eta}_{\ell,j} \in \mathbb{R}^d \quad \forall \ell \in [L], j \in [n_\ell]. \tag{4.3}$$

Here, $\mathbf{U} \in \text{St}(k, d)$ is a planted subspace, $\mathbf{\Theta} = \text{diag}(\sqrt{\lambda}_1, \dots, \sqrt{\lambda}_k)$ represent the known signal amplitudes, $\mathbf{z}_{\ell,j} \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{I}_k)$ are latent variables, and $\boldsymbol{\eta}_{\ell,j} \overset{\text{iid}}{\sim} \mathcal{N}(\mathbf{0}, v_\ell \mathbf{I}_d)$ are additive Gaussian heteroscedastic noises. Assume that $\lambda_i \ne \lambda_j$ for $i \ne j \in [k]$ and $v_\ell \ne v_m$ for $\ell \ne m \in [L]$. The maximum likelihood problem in [25, equation 3] with respect to $\mathbf{U}$ is then equivalently (1.2) for $\mathbf{A}_\ell = \sum_{j=1}^{n_\ell} \frac{1}{v_\ell} \mathbf{y}_{\ell,j} \mathbf{y}'_{\ell,j}$ 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, the signal-to-noise ratio λ_i / v_ℓ grows, or the variances are close to the median noise variance, the matrices in the HPPCA problem are almost commuting under the spectral norm. The proof is found in Appendix C.

PROPOSITION 4.9. *Let $\mathbf{c} = (\frac{1}{n}\mathbf{M}_1, \dots, \frac{1}{n}\mathbf{M}_k)$ be the (normalized) data matrices of the HPPCA problem. Then there exists commuting $\overline{\mathbf{c}} = (\overline{\mathbf{M}}_1, \dots, \overline{\mathbf{M}}_k)$ (constructed in the proof) and a universal constant $C > 0$ such that for any $\bar{v} \ge 0$ and any $t > 0$, with probability exceeding $1 - e^{-t}$,*

$$\frac{\|\frac{1}{n}\mathbf{M}_i - \overline{\mathbf{M}}_i\|}{\|\overline{\mathbf{M}}_1\|} \le \min \left\{ \sum_{\ell=1}^{L} \frac{\gamma_\ell(\bar{v})}{\frac{\lambda_i}{v_\ell} + 1}, C \frac{\bar{\sigma}_i}{\bar{\sigma}_1} \max \left\{ \sqrt{\frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i} \log d + t}{n}}, \frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i} \log d + t}{n} \log(n) \right\} \right\}, \tag{4.4}$$

where

$$\gamma_\ell(\bar{v}) := \left| \frac{\bar{v}}{v_\ell} - 1 \right|, \qquad \bar{\sigma}_i := \|\overline{\mathbf{M}}_i\| = \sum_{\ell=1}^{L} \frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell}+1} \frac{n_\ell}{n} \left(\frac{\lambda_1}{v_\ell} + 1 \right),$$

$$\bar{\xi}_i := \operatorname{tr}(\overline{\mathbf{M}}_i) = \sum_{\ell=1}^{L} \frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell}+1} \frac{n_\ell}{n} \left(\frac{1}{v_\ell} \sum_{j=1}^{k} \lambda_j + d \right).$$

Remark 4.10. It seems natural to let $\bar{v} = v_{\text{med}} = \min_v \sum_{\ell=1}^{L} |v - v_\ell|$, i.e., the median noise variance, which provides an upper bound for (4.4), i.e.,

$$\min_{\bar{v}} \sum_{\ell=1}^{L} \frac{\gamma_\ell(\bar{v})}{\frac{\lambda_i}{v_\ell}+1} = \min_{\bar{v}} \sum_{\ell=1}^{L} \frac{|\bar{v} - v_\ell|}{\lambda_i + v_\ell} \leq \sum_{\ell=1}^{L} \frac{|v_{\text{med}} - v_\ell|}{\lambda_i + v_\ell}.$$

The proof of Proposition 4.9 in Appendix C analyzes two cases separately, obtaining bounds for the normalized distance under the spectral norm between each $\mathbf{M}_i$ and $\overline{\mathbf{M}}_i$. The final result in (4.4) then takes the minimum of the two bounds. The left argument of the minimum operator in (4.4) reflects the effect of the heterogeneous noise. As all of the variances become close in value to some $\bar{v} \geq 0$, the matrices $\mathbf{M}_i$ become almost commuting, eventually becoming equal when all the variances are equal, i.e., the noise is homogeneous. In addition, the distance depends on the inverse signal-to-noise ratios between the eigenvalues λ_i and variances v_ℓ, so as the noise diminishes, the matrices $\mathbf{M}_i$ also become almost commuting.

The right argument of the minimum operator captures the effects of growing dimension, rank, and sample size using the concentration of the sample covariance matrices for Gaussian random variables [29]. First, the normalized distance between each $\mathbf{M}_i$ and $\overline{\mathbf{M}}_i$ grows as $\mathcal{O}(d \log d)$ and linearly with $\sum_{i=1}^{k} \lambda_i$ (which is related to the rank), as reflected by the terms $\bar{\xi}_i$. Lastly, the bound diminishes as $\mathcal{O}(1/\sqrt{n})$, where n is the total number of data samples; as the sample size grows, the matrices become almost commuting.

5. Numerical experiments. All 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. Since Theorem 4.1 requires an exact stationary point, and in practice, an iterative solver only returns an inexact stationary point, the KKT conditions may not be exactly satisfied. However, in practice, we found using smaller numerical precisions in the SDP and iterative solvers is often sufficient to achieve a numerical certificate, albeit inexact. When computing a first-order stationary point with an iterative solver, we terminate the algorithm when the norm of the gradient on the manifold is less than 10^{-10}. We point the reader to section SM5 in the supplement for further detailed settings.

5.1. Assessing the 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 optimal rank-one $\mathbf{X}_i$, yielding a globally optimal solution of (1.1). Similar experiments for other forms of (1.1), including where $\mathbf{M}_i$ are random low-rank PSD matrices, are found in section SM5 and give similar insights.

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

$\mathbf{v} = [1, 4]$		**Fraction of 100 trials with ROP**			
		$k = 3$	$k = 5$	$k = 7$	$k = 10$
$\mathbf{n} = [5, 20]$	$d = 10$	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
$\mathbf{n} = [20, 80]$	$d = 10$	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
$\mathbf{n} = [100, 400]$	$d = 10$	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

Here, the $\mathbf{M}_i$ were generated according to our HPPCA model in (4.3) where $\mathbf{A}_\ell = \frac{1}{v_\ell}\sum_{i=1}^{n_\ell} \mathbf{y}_{\ell,i}\mathbf{y}'_{\ell,i}$ 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 made $\boldsymbol{\lambda}$ a k-length vector of entries uniformly spaced in the interval $[1, 4]$, and we varied the ambient dimension d, rank k, samples $\mathbf{n} := [n_1, \ldots, n_L]$, and variances $\mathbf{v}$ for both $L = 2$ and $L = 3$ noise groups. Each random instance was generated from a new random draw of $\mathbf{U}$ on the Stiefel manifold, latent variables $\mathbf{z}_{\ell,i}$, and noise vectors $\boldsymbol{\eta}_{\ell,i}$.

Tables 1, 2, and 3 show the results of these experiments for various choices of dimension d, rank k, samples $\mathbf{n}$, and variances $\mathbf{v}$. We solved the SDP for 100 random data instances for $d \leq 50$ and 20 random data instances for $d \geq 100$. The table shows the fraction of trials that resulted in rank-one $\mathbf{X}_i \; \forall \; i = 1, \ldots, k$. We computed the average error of the sorted eigenvalues of each optimal solution $\overline{\mathbf{X}}_i$ to $\mathbf{e}_1$, i.e., $\frac{1}{k}\sum_{i=1}^{k} \|\operatorname{diag}(\boldsymbol{\Sigma}_i) - \mathbf{e}_1\|_2^2$, where $\overline{\mathbf{X}}_i = \mathbf{V}_i\boldsymbol{\Sigma}_i\mathbf{V}'_i$, and counted any trial with error greater than 10^{-5} as not tight.

The SDP solutions possessed the ROP in the vast majority of trials. As we increased the total number of samples n in Tables 1 and 3, the convex relaxation became tight in 100% of the trials, as predicted by the commuting error bound dependency on $\mathcal{O}(1/\sqrt{n})$ in Proposition 4.9. As d or k increased, we generally observed a few instances where the SDP was not tight, which conforms with the theory in Proposition 4.9. As we decreased the spread of the variances, Table 2 shows the fraction of tight instances increased, reaching 100% in the homoscedastic setting, as expected because then all of the $\mathbf{M}_i$ are equal. Likewise, Table 3 shows this behavior for the $L = 3$ case.

5.2. Assessing global optimality of local solutions. In this section, we used the Stiefel majorization-minimization (StMM) solver with a linear majorizer from [13] to obtain a local solution $\overline{\mathbf{U}}_{\mathrm{MM}}$ to (1.1) for various inputs $\mathbf{M}_i$ and used Theorem 4.1 to certify if the local solution is globally optimal or if the certificate fails. For comparison,

TABLE 2

Numerical experiments showing the fractions of trials where the SDP was tight for instances of the HPPCA problem as we varied d, k, and $\mathbf{v}$ using $L = 2$ groups with samples $\mathbf{n} = [10, 40]$ and $\mathbf{n} = [50, 200]$. Due to the large computation time of solving the full SDP for larger values of $d \geq 100$, we only ran 20 independent trials for each experiment setting.

		Fraction of 100 trials with ROP			
$\mathbf{n} = [10, 40]$		$k = 3$	$k = 5$	$k = 7$	$k = 10$
$\mathbf{v} = [1, 1]$	$d = 10$	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
$\mathbf{v} = [1, 2]$	$d = 10$	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
$\mathbf{v} = [1, 3]$	$d = 10$	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

		Fraction of 20 trials with ROP	
$\mathbf{v} = [1, 3]$		$k = 5$	$k = 10$
$\mathbf{n} = [10, 40]$	$d = 100$	1	0.85
	$d = 200$	0.95	0.35
	$d = 300$	0.75	0.35
$\mathbf{n} = [50, 200]$	$d = 100$	1	0.95
	$d = 200$	1	0.8
	$d = 300$	1	0.85

we obtained candidate solutions $\overline{\mathbf{X}}_i$ from the SDP and performed a rank-one SVD of each to form $\overline{\mathbf{U}}_{\text{SDP}}$, i.e.,

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

while measuring how close the solutions are to being rank-one. In the case where the SDP is not tight, the rank-one directions from the $\overline{\mathbf{X}}_i$ will not be orthonormal, so as a heuristic, we projected $\overline{\mathbf{U}}_{\text{SDP}}$ onto the Stiefel manifold by the orthogonal Procrustes solution, denoted by the operator $\mathcal{P}_{\text{St}}(\cdot)$ [13].

5.2.1. Synthetic CJD matrices. To empirically verify our theory from section 4, we generated each $\mathbf{M}_i \in \mathbb{R}^{d\times d}$ to be a diagonally dominant matrix resembling an approximately rank-k 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 constructed $\mathbf{M}_k = \mathbf{D}_k + \mathbf{N}_k$, where $\mathbf{D}_k$ is a diagonal matrix with k nonzero entries drawn uniformly at random from $[0, 1]$, and $\mathbf{N}_k = \frac{1}{10d}\mathbf{S}\mathbf{S}'$ for $\mathbf{S} \in \mathbb{R}^{d\times 10d}$ independent and identically distributed as $\mathcal{N}(0, \sigma\mathbf{I})$ for varying σ. We then generated the remaining $\mathbf{M}_i$ for $i = k-1, \ldots, 1$

TABLE 3

Numerical experiments showing the fractions of trials where the SDP was tight for instances of the HPPCA problem as we varied d, k, $\mathbf{n}$, and v_2 for $L = 3$ groups with noise variances $\mathbf{v} = [1, v_2, 4]$. The left and right tables show the results for $d = 20$ and $d = 50$, respectively. We swept v_2 in a way such that $\mathbf{n} = [20, 20, 60], \mathbf{v} = [1, 4, 4]$ is statistically equivalent to the problem in Table 1 for $\mathbf{n} = [20, 80], \mathbf{v} = [1, 4]$ and, similarly, $\mathbf{n} = [20, 80, 400], \mathbf{v} = [1, 1, 4]$ is statistically equivalent to the problem in Table 1 for $\mathbf{n} = [100, 400], \mathbf{v} = [1, 4]$.

		Fraction of 100 trials with ROP			
$d = 20$, $v_1 = 1$, $v_3 = 4$		$v_2 = 1$	$v_2 = 2$	$v_2 = 3$	$v_2 = 4$
$k = 5$	$\mathbf{n} = [20, 20, 60]$	1	1	1	1
	$\mathbf{n} = [20, 80, 60]$	1	1	1	0.99
	$\mathbf{n} = [20, 80, 200]$	1	1	1	1
	$\mathbf{n} = [20, 20, 400]$	1	1	1	1
	$\mathbf{n} = [20, 80, 400]$	1	1	1	1
	$\mathbf{n} = [100, 100, 400]$	1	1	1	1
	$\mathbf{n} = [200, 200, 400]$	1	1	1	1
$k = 10$	$\mathbf{n} = [20, 20, 60]$	0.99	1	0.99	0.97
	$\mathbf{n} = [20, 80, 60]$	1	1	0.99	0.99
	$\mathbf{n} = [20, 80, 200]$	1	1	1	1
	$\mathbf{n} = [20, 20, 400]$	1	1	1	1
	$\mathbf{n} = [20, 80, 400]$	1	1	1	1
	$\mathbf{n} = [100, 100, 400]$	1	1	1	1
	$\mathbf{n} = [200, 200, 400]$	1	1	1	1

		Fraction of 100 trials with ROP			
$d = 50$, $v_1 = 1$, $v_3 = 4$		$v_2 = 1$	$v_2 = 2$	$v_2 = 3$	$v_2 = 4$
$k = 5$	$\mathbf{n} = [20, 20, 60]$	1	1	0.98	0.96
	$\mathbf{n} = [20, 80, 60]$	1	1	0.99	0.96
	$\mathbf{n} = [20, 80, 200]$	1	1	0.98	0.99
	$\mathbf{n} = [20, 20, 400]$	1	1	1	0.99
	$\mathbf{n} = [20, 80, 400]$	1	1	1	1
	$\mathbf{n} = [100, 100, 400]$	1	1	0.99	1
	$\mathbf{n} = [200, 200, 400]$	1	1	1	1
$k = 10$	$\mathbf{n} = [20, 20, 60]$	1	0.97	0.96	0.92
	$\mathbf{n} = [20, 80, 60]$	1	1	0.98	0.94
	$\mathbf{n} = [20, 80, 200]$	1	0.98	0.99	0.99
	$\mathbf{n} = [20, 20, 400]$	0.99	0.99	1	0.99
	$\mathbf{n} = [20, 80, 400]$	1	1	1	0.99
	$\mathbf{n} = [100, 100, 400]$	1	1	1	1
	$\mathbf{n} = [200, 200, 400]$	1	1	1	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 normalized all by $1/\max_{i\in[k]} \|\mathbf{M}_i\|$ so that $\|\mathbf{M}_i\| \leq 1 \ \forall \ i \in [k]$. With this setup, by varying σ, we swept through a range of commuting distances under the spectral norm, i.e. $\max_{i,j\in[k]} \|\mathbf{M}_i\mathbf{M}_j - \mathbf{M}_j\mathbf{M}_i\|$. For all experiments, we generated problems with parameters $d = 10$ and $k = 3$, and ran StMM for 2,000 maximum iterations or until the norm of the gradient on the manifold was less than 10^{-10}.

Figure 1(a) shows the gap of the objective values between the SDP relaxation (before projection onto the Stiefel) and the nonconvex problem ($p_{\text{SDP}} - p_{\text{StMM}}$) versus

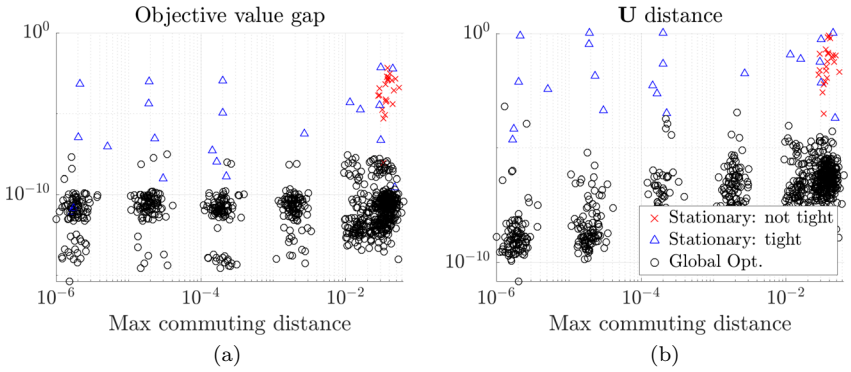


FIG. 1. *Numerical simulations for synthetic CJD matrices for* $d = 10, k = 3$ *with increasing* σ *and* 100 *random problem instances for each setting. As* σ *grows, the max commuting distance grows.*

the commuting distance. Figure 1(b) shows the distance between the two obtained solutions computed as $\frac{1}{\sqrt{k}}\||\overline{\mathbf{U}}'_{\mathrm{StMM}}\overline{\mathbf{U}}_{\mathrm{SDP}}|-\mathbf{I}_k\|_F$ (where $|\cdot|$ denotes taking the elementwise absolute value) versus commuting distance. Figure 3(a) shows the percentage of trials where $\overline{\mathbf{U}}_{\mathrm{StMM}}$ could not be certified globally optimal. Like before, we declared an SDP's solution "tight" if the mean error of its solutions to a rank-one matrix with binary eigenvalues, i.e., $\frac{1}{k}\sum_{i=1}^{k}\|\lambda_{\downarrow}^{(i)} - \mathbf{e}_1\|_2$, was less than 10^{-5}, where $\lambda_{\downarrow}^{(i)}$ denotes the sorted eigenvalues of $\mathbf{X}_i$ in descending order, and $\mathbf{e}_1$ is the first standard basis vector in $\mathbb{R}^d$. Trials with the marker "$\circ$" indicate trials where global optimality was certified. The marker "$\times$" represents trials where $\overline{\mathbf{U}}$ was not certified as globally optimal and the SDP relaxation was not tight; "$\triangle$" markers indicate trials where the SDP was tight, but (4.1) was not satisfied, implying a suboptimal local maximum.

Toward the left of Figure 1(a), with small σ and the $(\mathbf{M}_1, \ldots, \mathbf{M}_k)$ all being very close to commuting, 100% of experiments returned tight rank-one SDP solutions. Notably, there appears to be a sharp cutoff point where this behavior ends, and the SDP relaxation was not tight in a small percentage of cases. While the large majority of trials still admitted a tight convex relaxation, these results empirically corroborate the sufficient conditions derived in Theorem 4.6 and Corollary 4.7.

Where the SDP is tight, Figure 1 shows the StMM solver returned 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 observed 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.

5.2.2. HPPCA. We repeated the experiments just described for $\mathbf{M}_i$ generated by the model in (4.3) for $d = 50$, $\boldsymbol{\lambda} = [4, 3.25, 2.5, 1.75, 1]$, and $L = 2$ noise groups with variances $\mathbf{v} = [1, 4]$. For each of 100 trials, we drew a random model with a different generative $\mathbf{U}$ for sample sizes $\mathbf{n} = [n_1, 4n_1]$, where we swept through increasing values of n_1 on the horizontal axis in Figure 3(b). For each experiment, we normalized the $\mathbf{M}_i$ by the maximum of their spectral norms, and then recorded the results obtained from the SDP and StMM solvers with respect to the computed maximum commuting distance of the $\mathbf{M}_i$ in Figure 2. We ran StMM for a maximum of 10,000 iterations and recorded both the global optimality certification of each StMM run and if the SDP was tight.

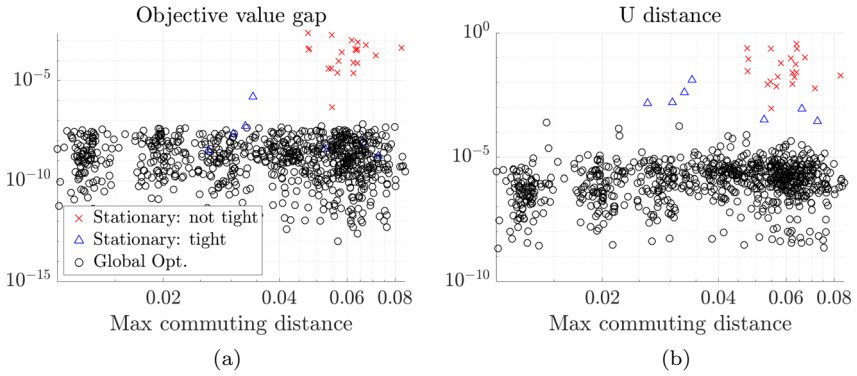
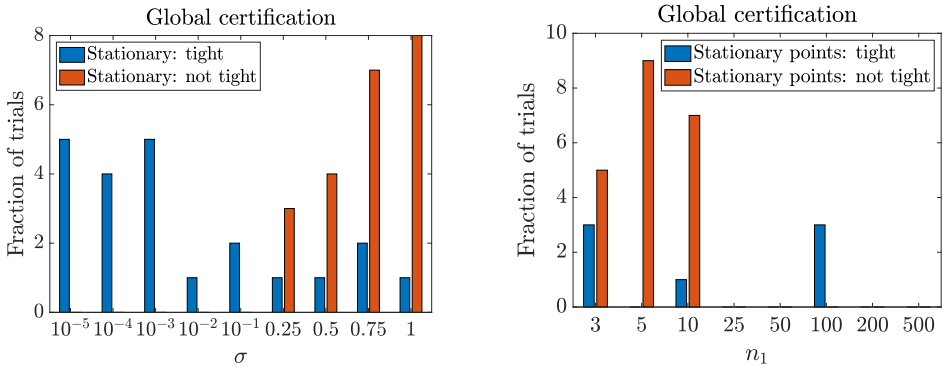


FIG. 2. *Numerical simulations for* $\mathbf{M}_i$ *generated by the HPPCA model in* (4.3) *for* $d = 50$, $k = 5$, *noise variances* $\mathbf{v} = [1, 4]$, *and* $\boldsymbol{\lambda} = [4, 3.25, 2.5, 1.75, 1]$ *with increasing samples* n. *As* n *grows, the max commuting distance gets smaller.*



(a) Synthetic CJD matrices for $d = 10$, $k = 3$ with increasing σ.

(b) Data matrices generated by the HPPCA model in (4.3) with increasing samples $\mathbf{n} = [n_1, 4n_1]$.

FIG. 3. *Percentages of global certification of StMM solutions out of* 100 *trials. The fractions not shown are tight instances certified as global.*

Proposition 4.9 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 was indeed what we observed: as the number of samples increased in Figure 3(b), the maximum commuting distance of the $\mathbf{M}_i$ decreased, i.e., the simulations moved to the left on the horizontal axes of Figure 2(a) and 2(b). In this nearly commuting regime, the SDP obtained tight rank-one $\mathbf{X}_i$ in 100% of the trials, and all of the StMM runs attained 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 a dual certificate was not attained. Also within this regime, several trials of the StMM solver found suboptimal local maxima.

5.3. Computation time. Figure 4 compares the scalability of our SDP relaxation in (SDP-P) to the StMM solver with the global certificate check in (4.1) for synthetically generated HPPCA problems of varying data dimension. We measured the median computation time across 10 independent trials of both algorithms. The experiment strongly demonstrates the computational superiority of the first-order

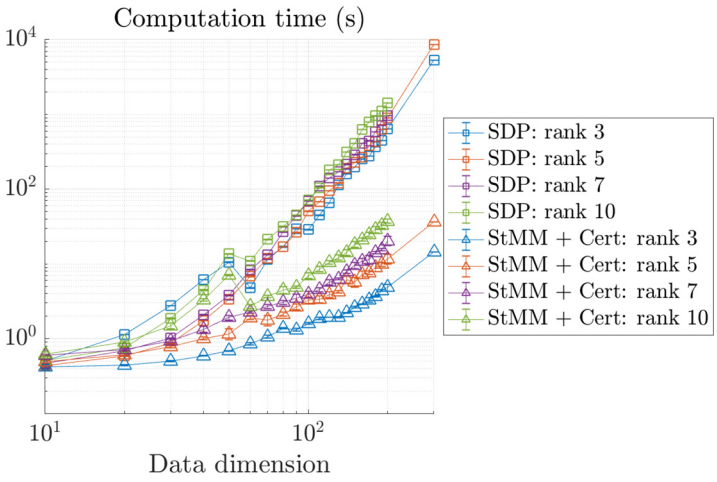


FIG. 4. *Computation time of* (SDP-P) *versus StMM for* 2000 *iterations with global certificate check* (4.1) *for HPPCA problems as the data dimension varies. We used* $\mathbf{v} = [1, 4]$*, and* $\mathbf{n} = [100, 400]$ *and made* $\boldsymbol{\lambda}$ *a* k*-length vector with entries equally spaced in the interval* $[1, 4]$*, where the rank of the model is* k*. Markers indicate the median computation time taken over* 10 *trials, and error bars show the standard deviation. Due to memory and computation limitations for* $d = 300$*, we only performed one timing test for* $k = 3$ *and* $k = 5$*.*

method with our certificate compared to the full SDP, as predicted by the computational complexity analysis in subsection 4.1. StMM+Certificate scaled 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 (1.1) using any choice of solver on the Stiefel manifold and perform a fast check of its terminal output for global optimality.

6. Future work and conclusion. In this work, we proposed a novel SDP relaxation for the sums of heterogeneous quadratic forms 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 guaranteeing the relaxation has the ROP and providing both theoretical and empirical support that a motivating signal processing application—the HPPCA problem—possesses a tight relaxation in many instances.

While the global certificate scales well compared to solving the full SDP, the LMI feasibility program still requires forming and factoring $d \times d$ size matrices, requiring storage of $\mathcal{O}(d^2)$ elements. One potential solution is to apply recent works like [43] to our problem, which use randomized algorithms to reduce the storage and arithmetic costs for scalable semidefinite programming. Further, it remains interesting to prove a sufficient analytical certificate as well as proving more general sufficient conditions on the $\mathbf{M}_i$ that guarantee the ROP. A key future extension is to precisely quantify the size of the region in Theorem 4.6 where the SDP has the ROP.

Another direction for future research would be to generalize Theorem 4.6 or to simplify its proof. While the problem in [16] is distinct from our own for the reasons discussed in section 3, it would be interesting to determine whether the ideas and insights of their theory can be applied in our case.

Appendix A. Proofs of results in section 2.

Proof of Lemma 2.1. The problem is convex and satisfies Slater's condition; see Lemma A.1. Specifically, for optimal primal solutions $\overline{\mathbf{X}}_i$ and optimal dual solutions $\overline{\mathbf{Y}}$, $\overline{\mathbf{Z}}_i$, and $\overline{\nu}_i \ \forall \ i \in [k]$, we have $\langle \mathbf{I} - (\sum_{i=1}^k \overline{\mathbf{X}}_i), \overline{\mathbf{Y}} \rangle = 0$ and therefore $\mathrm{tr}(\overline{\mathbf{Y}}) = \langle \overline{\mathbf{Y}}, \sum_{i=1}^k \overline{\mathbf{X}}_i \rangle$. Then

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

since $\langle \overline{\mathbf{Z}}_i, \overline{\mathbf{X}}_i \rangle = 0$ and $\sum_{i=1}^k \overline{\nu}_i (1 - \mathrm{tr}(\overline{\mathbf{X}}_i)) = 0$. Thus, $p^* = d^*$. ∎

LEMMA A.1. *The primal problem in* (SDP-P) *is strictly feasible for* $k < d$.

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

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

Suppose $\mathbf{X}_i = \frac{1}{d}\mathbf{I} \ \forall \ i$. Then $\mathrm{tr}(\mathbf{X}_i) = 1$ and $\mathbf{X}_i \succ 0 \ \forall \ i$, and $\sum_{i=1}^k \mathbf{X}_i = \frac{k}{d}\mathbf{I}$, satisfying $0 \prec \sum_{i=1}^k \mathbf{X}_i \prec \mathbf{I}$ when $k < d$. ∎

Proof of Lemma 2.3. Since the problem in (SDP-P) has a larger constraint set than (1.1), any solution to (SDP-P) that satisfies the constraints of (1.1) also constitutes a solution to this original nonconvex problem.

For the "if" direction, assume that the optimal $\mathbf{X}_i$ for (SDP-P) have the ROP. Since $\mathrm{tr}(\mathbf{X}_i) = 1$ by definition of (SDP-P), when we decompose $\mathbf{X}_i = \mathbf{u}_i \mathbf{u}_i'$ we have $\mathbf{u}_i$ that are norm-1. In order for $\sum_{i=1}^k \mathbf{X}_i \preccurlyeq \mathbf{I}$, the $\mathbf{u}_i$ 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 (1.1) in the sense that $\mathbf{X}_i = \mathbf{u}_i \mathbf{u}_i'$ gives the optimal $\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \cdots & \mathbf{u}_k \end{bmatrix}$. Then by definition we see that the $\mathbf{X}_i$ have the ROP. ∎

LEMMA A.2. *Suppose* $\mathbf{X}_i$ *for* $i = 1, \ldots, k$ *each have trace* 1 *and satisfy* $\lambda_1(\mathbf{X}_i) = 1$, *and therefore each* $\mathbf{X}_i$ *is rank-one. We decompose* $\mathbf{X}_i = \mathbf{u}_i \mathbf{u}_i'$ *and note that* $\mathbf{u}_i$ *are norm*-1. *Then* $\sum_{i=1}^k \mathbf{X}_i$ *satisfies* $0 \preccurlyeq \sum_{i=1}^k \mathbf{X}_i \preccurlyeq \mathbf{I}$ *if and only if* $\mathbf{u}_i' \mathbf{u}_j = 0 \quad \forall i \neq j$.

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

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

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

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

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

The backward direction is immediate because when $\mathbf{u}_i' \mathbf{u}_j = 0$ for $i \neq j$, $\sum_{i=1}^k \mathbf{u}_i \mathbf{u}_i'$ is the eigenvalue decomposition of $\mathbf{X}$ with k eigenvalues equal to one. ∎

Proof of Lemma 2.4. Suppose $\mathbf{Z}_i$ is rank $d-1$. By complementarity at optimality, we have $\mathbf{Z}_i\mathbf{X}_i = 0 \quad \forall i$, which means $\mathbf{X}_i$ lies in the nullspace of $\mathbf{Z}_i$, which has dimension 1, so each $\mathbf{X}_i$ is rank-one. By primal feasibility, $\mathrm{tr}(\mathbf{X}_i) = 1$, so $\lambda_1(\mathbf{X}_i) = 1 \quad \forall i = 1, \ldots, k$. By Lemma A.2, the optimal solution is an orthogonal projection matrix, and the optimal $\mathbf{X}_i$ are orthogonal. ∎

Appendix B. Proof of Theorem 4.1 and Corollary 4.2.

Proof of Theorem 4.1. By Lemma 2.1, primal and dual feasible solutions of (SDP-P) and (SDP-D), $\overline{\mathbf{X}}_i, \overline{\mathbf{Z}}_i, \overline{\mathbf{Y}}, \overline{\boldsymbol{\nu}}$, are simultaneously optimal if and only if they satisfy the following KKT conditions [10], where the variables and constraints are indexed by $i \in [k]$:

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

$$\text{(KKT-b)} \qquad \overline{\mathbf{Y}} = \mathbf{M}_i + \overline{\mathbf{Z}}_i - \overline{\nu}_i \mathbf{I}, \qquad \overline{\mathbf{Y}} \succeq 0,$$

$$\text{(KKT-c)} \qquad \left\langle \mathbf{I} - \sum_{i=1}^{k} \overline{\mathbf{X}}_i, \overline{\mathbf{Y}} \right\rangle = 0,$$

$$\text{(KKT-d)} \qquad \langle \overline{\mathbf{Z}}_i, \overline{\mathbf{X}}_i \rangle = 0,$$

$$\text{(KKT-e)} \qquad \overline{\mathbf{Z}}_i \succeq 0.$$

Similar to the work in [42], our strategy is then to construct $\overline{\mathbf{X}}_i$ and $\overline{\mathbf{Y}}, \overline{\mathbf{Z}}_i, \overline{\boldsymbol{\nu}}$ satisfying these conditions. Given $\overline{\mathbf{U}}$ and $\overline{\boldsymbol{\nu}}$ in the statement of the theorem, we define $\overline{\mathbf{X}}_i = \overline{\mathbf{u}}_i \overline{\mathbf{u}}_i'$, $\overline{\mathbf{Y}} = \overline{\mathbf{U}}(\overline{\boldsymbol{\Lambda}} - \mathbf{D}_{\overline{\boldsymbol{\nu}}})\overline{\mathbf{U}}'$, and $\overline{\mathbf{Z}}_i = \overline{\mathbf{Y}} + \bar{\nu}_i \mathbf{I} - \mathbf{M}_i$. By construction, $\overline{\mathbf{X}}_i$ satisfy (KKT-a). Also by construction $\overline{\mathbf{Y}} = \mathbf{M}_i + \overline{\mathbf{Z}}_i - \overline{\boldsymbol{\nu}}_i \mathbf{I}$, and the assumption that $\overline{\boldsymbol{\Lambda}} \succeq \mathbf{D}_{\overline{\boldsymbol{\nu}}}$ ensures $\overline{\mathbf{Y}} \succeq 0$ to satisfy (KKT-b). One can also verify that $\langle \mathbf{I} - \overline{\mathbf{X}}, \overline{\mathbf{Y}} \rangle = 0$ by construction, thus satisfying (KKT-c). So it remains to show $\langle \overline{\mathbf{Z}}_i, \overline{\mathbf{X}}_i \rangle = 0$ and $\overline{\mathbf{Z}}_i \succeq 0$.

Moreover, $\overline{\mathbf{Z}}_i \succeq 0$ by the assumption in (4.1), satisfying (KKT-e). We finally verify (KKT-d), i.e., $\langle \overline{\mathbf{Z}}_i, \overline{\mathbf{X}}_i \rangle = 0$, with $\overline{\mathbf{U}} = [\overline{\mathbf{u}}_1 \cdots \overline{\mathbf{u}}_k]$:

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

Remark B.1. Given the fact that the Lagrange multipliers $\overline{\nu}_i$ corresponding to the trace constraints are nonnegative by Lemma B.2, this also implies that $\overline{\boldsymbol{\Lambda}} \succeq 0$. We note that this indeed fulfills a necessary condition for $\overline{\mathbf{U}}$ to be a second-order stationary point by Lemma SM2.1 and Lemma SM2.2 in the supplement.

See subsection SM2.2 in the supplement for additional remarks.

For the following results in this paper, we require a proof that the optimal $\overline{\boldsymbol{\nu}}$ in (SDP-D) are nonnegative.

LEMMA B.2. *Assume all* $\mathbf{M}_i$ *are PSD, and* $k < d$. *Then all* $\nu_i \geq 0$ *at optimality.*

Proof. For a contradiction suppose the optimal $\boldsymbol{\nu}$ has at least one coordinate that is strictly negative. Without loss of generality, let $\nu_1 < 0$ be the smallest (most negative) coordinate of $\boldsymbol{\nu}$, and rewrite the objective in terms of $\mathbf{M}_1$ and eliminating $\mathbf{Y}$ as

$$d^* = \min_{\nu_i, \mathbf{Z}_i} \operatorname{tr}(\mathbf{Z}_1 + \mathbf{M}_1) - d\nu_1 + \sum_{i=1}^{k} \nu_i \tag{B.1}$$

$$\begin{aligned} \text{s.t. } & \mathbf{M}_i + \mathbf{Z}_i \succcurlyeq \nu_i \mathbf{I} \quad \forall i = 1, \dots, k, \\ & \mathbf{M}_1 + \mathbf{Z}_1 - \nu_1 \mathbf{I} = \mathbf{M}_j + \mathbf{Z}_j - \nu_j \mathbf{I} \quad \forall j = 2, \dots, k, \\ & \mathbf{Z}_i \succcurlyeq 0 \quad \forall i = 1, \dots, k. \end{aligned} \tag{B.2}$$

Now consider new variables $\{\tilde{\nu}_i, \tilde{\mathbf{Z}}_i\}_{i=1}^k$, where we let $\tilde{\nu}_1 = 0$, $\tilde{\nu}_i = \nu_i - \nu_1$ for $i = 2, \dots, k$, and leave all the $\mathbf{Z}$ variables unchanged: $\tilde{\mathbf{Z}}_i = \mathbf{Z}_i \ \forall\ i$.

These new variables are still feasible. Certainly $\mathbf{M}_1 + \tilde{\mathbf{Z}}_1 = \mathbf{M}_1 + \mathbf{Z}_1 \succcurlyeq \tilde{\nu}_1 \mathbf{I} = 0$ as both $\mathbf{M}_1, \mathbf{Z}_1$ are PSD. Also $\mathbf{M}_1 + \tilde{\mathbf{Z}}_1 - \tilde{\nu}_1 \mathbf{I} = \mathbf{M}_j + \tilde{\mathbf{Z}}_j - \tilde{\nu}_j \mathbf{I}$, since substituting in, we have $\mathbf{M}_1 + \mathbf{Z}_1 = \mathbf{M}_j + \mathbf{Z}_j - (\nu_j - \nu_1)\mathbf{I}$, which was feasible for the original optimal point. From this last equation note that since $\mathbf{M}_1 + \mathbf{Z}_1 \succcurlyeq 0$, then $\mathbf{M}_j + \mathbf{Z}_j - (\nu_j - \nu_1)\mathbf{I} = \mathbf{M}_j + \tilde{\mathbf{Z}}_j - \tilde{\nu}_j \mathbf{I} \succeq 0$.

However, with the assumption that $k < d$, this yields a contradiction because we have reduced the objective value from

$$\operatorname{tr}(\mathbf{Z}_1 + \mathbf{M}_1) - d\nu_1 + \sum_{i=1}^{k} \nu_i \qquad \text{to} \qquad \operatorname{tr}(\mathbf{Z}_1 + \mathbf{M}_1) - k\nu_1 + \sum_{i=1}^{k} \nu_i \,.$$

Therefore $\nu_i < 0$ cannot be optimal. ∎

Proof of Corollary 4.2. We first argue that this problem attains an optimal solution as follows. We note that (4.2) is feasible by taking $\overline{\boldsymbol{\nu}} = 0$ and ϵ sufficiently large. Next, the optimal value of (4.2) is clearly bounded below by 0. In addition, for any fixed $\bar{\epsilon}$, one can see that the level set of feasible points $(\epsilon, \overline{\boldsymbol{\nu}})$ with $\epsilon \leq \bar{\epsilon}$ is bounded via the constraint $0 \preceq \mathbf{D}_{\overline{\boldsymbol{\nu}}} \preceq \overline{\boldsymbol{\Lambda}} + \epsilon \mathbf{I}$, which in particular bounds each entry of $\overline{\boldsymbol{\nu}}$ from below by Lemma B.2 and from above by the corresponding diagonal entry of $\overline{\boldsymbol{\Lambda}}$. Hence, an optimal solution $(\epsilon^*, \overline{\boldsymbol{\nu}}^*)$ is attained. Let ϵ^* be the unique optimal value of the optimization problem. From this ϵ^*, now we construct a solution to the following approximate KKT conditions [10] of (SDP-P), indexing the variables and constraints by $i \in [k]$:

$$\text{(eps-KKT-a)} \qquad \overline{\mathbf{X}}_i \succeq 0, \quad \sum_{i=1}^{k} \overline{\mathbf{X}}_i \preceq \mathbf{I}, \quad \operatorname{tr}(\overline{\mathbf{X}}_i) = 1,$$

$$\text{(eps-KKT-b)} \qquad \overline{\mathbf{Y}} = \mathbf{M}_i + \overline{\mathbf{Z}}_i - \overline{\nu}_i \mathbf{I}, \qquad \overline{\mathbf{Y}} \succeq -\epsilon^* \mathbf{I},$$

$$\text{(eps-KKT-c)} \qquad \left\langle \mathbf{I} - \sum_{i=1}^{k} \overline{\mathbf{X}}_i, \overline{\mathbf{Y}} \right\rangle = 0,$$

$$\text{(eps-KKT-d)} \qquad \langle \overline{\mathbf{Z}}_i, \overline{\mathbf{X}}_i \rangle = 0,$$

$$\text{(eps-KKT-e)} \qquad \overline{\mathbf{Z}}_i \succeq -\epsilon^* \mathbf{I}.$$

Given a $\overline{\mathbf{U}}$ and optimal $\overline{\boldsymbol{\nu}}$ to (4.2), we define $\overline{\mathbf{X}}_i = \overline{\mathbf{u}}_i \overline{\mathbf{u}}_i'$, $\overline{\mathbf{Y}} = \overline{\mathbf{U}}(\overline{\boldsymbol{\Lambda}} - \mathbf{D}_{\overline{\boldsymbol{\nu}}})\overline{\mathbf{U}}'$, and $\overline{\mathbf{Z}}_i = \overline{\mathbf{Y}} + \overline{\nu}_i \mathbf{I} - \mathbf{M}_i$. By construction, $\overline{\mathbf{X}}_i$ satisfy (eps-KKT-a), and it is clear that $\overline{\mathbf{Y}} = \mathbf{M}_i + \overline{\mathbf{Z}}_i - \overline{\boldsymbol{\nu}}_i \mathbf{I}$ satisfies the first condition in (eps-KKT-b). One can also verify that $\langle \mathbf{I} - \overline{\mathbf{X}}, \overline{\mathbf{Y}} \rangle = 0$ by construction, thus satisfying (eps-KKT-c).

One can easily show that $\overline{\mathbf{\Lambda}} - \mathbf{D}_{\overline{\boldsymbol{\nu}}} \succeq -\epsilon^* \mathbf{I}$ ensures $\overline{\mathbf{Y}} \succeq -\epsilon^* \mathbf{I}$ (eps-KKT-b). Moreover, $\overline{\mathbf{Z}}_i \succeq -\epsilon^* \mathbf{I}$ by the assumption in (4.2), satisfying (eps-KKT-e). Just as we did in the proof of Theorem 4.1 we finally verify (eps-KKT-d), i.e., $\langle \overline{\mathbf{Z}}_i, \overline{\mathbf{X}}_i \rangle = 0$, with $\overline{\mathbf{U}} = [\overline{\mathbf{u}}_1 \cdots \overline{\mathbf{u}}_k]$.

Let us now focus on $\overline{\mathbf{Y}}, \overline{\mathbf{Z}}_i$, which are approximately feasible for the dual problem. By defining $\mathbf{Y} := \overline{\mathbf{Y}} + \epsilon^* \mathbf{I}$, $\mathbf{Z}_i := \overline{\mathbf{Z}}_i + \epsilon^* \mathbf{I}$, $\mathbf{Z} := (\mathbf{Z}_1, \ldots, \mathbf{Z}_k)$, and $\boldsymbol{\nu} := \overline{\boldsymbol{\nu}}$, we recover dual feasibility, i.e., $\mathbf{Y} \succeq 0$ and $\mathbf{Z}_i \succeq 0$. Hence, the duality gap between $\overline{\mathbf{X}}_1, \ldots, \overline{\mathbf{X}}_k$ and $\mathbf{Y}, \mathbf{Z}, \boldsymbol{\nu}$ is nonnegative and, in fact, equals $\epsilon^* d$ due to the approximate KKT system:

$$\begin{aligned}
d(\mathbf{Y}, \mathbf{Z}, \boldsymbol{\nu}) - p(\overline{\mathbf{U}}) &= \operatorname{tr}(\mathbf{Y}) + \sum_i \boldsymbol{\nu}_i - \sum_i \langle \mathbf{M}_i, \overline{\mathbf{X}}_i \rangle \\
&= \operatorname{tr}(\mathbf{Y}) + \sum_i \boldsymbol{\nu}_i - \sum_i \langle \mathbf{Y} - \mathbf{Z}_i + \boldsymbol{\nu}_i \mathbf{I}, \overline{\mathbf{X}}_i \rangle \\
&= \operatorname{tr}(\mathbf{Y}) + \sum_i \boldsymbol{\nu}_i - \left\langle \mathbf{Y}, \sum_i \overline{\mathbf{X}}_i \right\rangle + \sum_i \langle \mathbf{Z}_i, \overline{\mathbf{X}}_i \rangle - \sum_i \boldsymbol{\nu}_i \\
&= \left\langle \mathbf{Y}, \mathbf{I} - \sum_i \overline{\mathbf{X}}_i \right\rangle + \sum_i \langle \mathbf{Z}_i, \overline{\mathbf{X}}_i \rangle \\
&= \left\langle \overline{\mathbf{Y}} + \epsilon^* \mathbf{I}, \mathbf{I} - \sum_i \overline{\mathbf{X}}_i \right\rangle + \sum_i \langle \overline{\mathbf{Z}}_i + \epsilon^* \mathbf{I}, \overline{\mathbf{X}}_i \rangle \\
&= \epsilon^* \operatorname{tr}\left(\mathbf{I} - \sum_i \overline{\mathbf{X}}_i \right) + \epsilon^* \sum_i \operatorname{tr}(\overline{\mathbf{X}}_i) = \epsilon^* \operatorname{tr}(\mathbf{I}) = \epsilon^* d.
\end{aligned}$$

In other words, letting $p(\overline{\mathbf{U}})$ be the primal objective associated with $\overline{\mathbf{U}}$ and $d(\mathbf{Y}, \mathbf{Z}, \boldsymbol{\nu})$ be the dual objective associated with $\mathbf{Y}, \mathbf{Z}, \boldsymbol{\nu}$, we have shown that the duality gap $d(\mathbf{Y}, \mathbf{Z}, \boldsymbol{\nu}) - p(\overline{\mathbf{U}}) = \epsilon^* d$, which implies $p(\overline{\mathbf{U}}) = d(\mathbf{Y}, \mathbf{Z}, \boldsymbol{\nu}) - \epsilon^* d \geq d^*(\mathbf{Y}, \mathbf{Z}, \boldsymbol{\nu}) - \epsilon^* d = p^* - \epsilon^* d$. ∎

Appendix C. Proofs of intermediate results supporting Theorem 4.6. Next, we give general convex analysis results that allow us to prove Theorem 4.6.

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

$$\begin{aligned}
(P; \mathbf{c}) \qquad & p(\mathbf{c}) := \min_{\mathbf{x}} \{ \mathbf{c}' \mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \in \mathcal{K} \}, \\
(D; \mathbf{c}) \qquad & d(\mathbf{c}) := \max_{\mathbf{y}} \{ \mathbf{b}' \mathbf{y} : \mathbf{c} - \mathbf{A}' \mathbf{y} \in \mathcal{K}^* \}.
\end{aligned}$$

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

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

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

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

$$\mathrm{Opt}(D;\mathbf{c}) := \{\mathbf{y} \in \mathrm{Feas}(D;\mathbf{c}) : \mathbf{b}'\mathbf{y} = d(\mathbf{c})\}$$

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

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

$$\mathrm{Extra}(D) := \{\mathbf{y} : \mathbf{f} - \mathbf{E}'\mathbf{y} \geq 0\}$$

satisfies the following.

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

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

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 $\{\mathbf{y}^k \in L^k\}$, every limit point $\bar{\mathbf{y}}$ of the sequence satisfies $\bar{\mathbf{y}} \in \bar{L}$; and (ii) every member $\bar{\mathbf{y}} \in \bar{L}$ is the limit point of some sequence $\{\mathbf{y}^k \in L^k\}$.

LEMMA C.1. *Under Assumptions* C.0.1 *and* C.0.2*, let* $\{\mathbf{c}^k \in \mathcal{C}\} \to \overline{\mathbf{c}}$ *be any convergent sequence. Then*

$$\big\{\mathrm{Feas}(D;\mathbf{c}^k) \cap \mathrm{Extra}(D)\big\} \quad \to \quad \mathrm{Feas}(D;\overline{\mathbf{c}}) \cap \mathrm{Extra}(D).$$

Proof. See subsection SM3.1 in the supplement for the proof. ∎

LEMMA C.2. *Under Assumptions* C.0.1 *and* C.0.2*, let* $\{\mathbf{c}^k \in \mathcal{C}\} \to \overline{\mathbf{c}}$ *be any convergent sequence. Then*

$$\big\{\mathrm{Opt}(D;\mathbf{c}^k)\big\} \quad \to \quad \mathrm{Opt}(D;\overline{\mathbf{c}}).$$

Proof. See subsection SM3.2 in the supplement for the proof. ∎

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

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

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

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

Proof. We must show that, for any convergent $\{\mathbf{c}^k\} \to \overline{\mathbf{c}}$, we also have convergence $\{y(\mathbf{c}^k)\} \to y(\overline{\mathbf{c}})$. This follows because $\{\mathrm{Opt}(D;\mathbf{c}^k)\} \to \mathrm{Opt}(D;\overline{\mathbf{c}})$ by Lemma C.2. ∎

Theorem 4.6 uses Proposition C.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 ROP. For this, we apply

the general theory developed above, specifically Proposition C.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 C.0.1 and C.0.2. In particular, for a fixed, user-specified upper bound $\mu > 0$, we define $\mathcal{C} := \{\mathbf{c} = (\mathbf{M}_1, \ldots, \mathbf{M}_k) : 0 \preceq \mathbf{M}_i \preceq \mu\mathbf{I} \quad \forall\ i = 1, \ldots, k\}$ to be our set of admissible coefficient k-tuples. In addition, we have shown in Lemma B.2 that all $\mathbf{M}_i \succeq 0$ implies that all ν_i are nonnegative at optimality. Thus, we enforce the redundant constraint that $\nu_i \geq 0\ \forall\ i \in [k]$.

We know that both (SDP-P) and (SDP-D) have interior points for all $\mathbf{c} \in \mathcal{C}$, so that strong duality holds. For the dual in particular, the equation $\mu\mathbf{I} = \mathbf{M}_i + ((\mu + \epsilon)\mathbf{I} - \mathbf{M}_i) - \epsilon\mathbf{I}$ shows that, for all $\epsilon > 0$, $\mathbf{Y}(\epsilon) := \mu\mathbf{I},\ \mathbf{Z}(\epsilon)_i := (\mu+\epsilon)\mathbf{I} - \mathbf{M}_i,\ \nu(\epsilon)_i := \epsilon$ 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 C.0.1.

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

We now repeat the discussion leading up to Theorem 4.6 for completeness. The first lemma says that the diagonal problem has dual variables $\mathbf{Z}_i$ such that $\mathrm{rank}(\mathbf{Z}_i) \geq d - 1$, implying that the primal variables $\mathbf{X}_i$ are rank-one.

Proof of Lemma 4.4. 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

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

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

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

Since the primal is an assignment problem, its unique optimal solution has the property that each $\mathrm{diag}(\mathbf{X}_i)$ 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 $\mathrm{diag}(\mathbf{X}_i) + \mathrm{diag}(\mathbf{Z}_i) > 0$ for each i. Hence, there exists a dual optimal solution with $\mathrm{rank}(\mathbf{Z}_i) \geq d - 1$ for each i, as desired. ∎

Proof of Corollary 4.8. We apply Lemma SM3.13 to $(\mathbf{A}_1, \ldots, \mathbf{A}_L)$. Then there exist Hermitian symmetric matrices $\bar{\mathbf{A}}_\ell$ such that $\|[\bar{\mathbf{A}}_\ell, \bar{\mathbf{A}}_m]\|_{\mathrm{tr}} = 0\ \forall\ \ell, m \in [L]$ such that $\|\mathbf{A}_\ell - \bar{\mathbf{A}}_\ell\|_{\mathrm{tr}} \leq \delta(\epsilon, k)\ \forall\ \ell \in [L]$. Let $\overline{\mathbf{M}}_i := \sum_{\ell=1}^L w_{\ell,i}\bar{\mathbf{A}}_\ell$. Then the matrices $\overline{\mathbf{M}}_i$ commute and are jointly diagonalizable:

$$[\overline{\mathbf{M}}_i, \overline{\mathbf{M}}_j] = \overline{\mathbf{M}}_i\overline{\mathbf{M}}_j - \overline{\mathbf{M}}_j\overline{\mathbf{M}}_i = 2\sum_{\ell \neq m}^L w_{\ell,i}w_{m,j}(\bar{\mathbf{A}}_\ell\bar{\mathbf{A}}_m - \bar{\mathbf{A}}_m\bar{\mathbf{A}}_\ell) = 0. \tag{C.1}$$

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

$$\|\mathbf{M}_i - \overline{\mathbf{M}}_i\|_{\text{tr}} = \left\| \sum_{\ell=1}^{L} w_{\ell,i}(\mathbf{A}_\ell - \bar{\mathbf{A}}_\ell) \right\|_{\text{tr}} \leq \sum_{\ell=1}^{L} w_{\ell,i} \|\mathbf{A}_\ell - \bar{\mathbf{A}}_\ell\|_{\text{tr}} \leq \sum_{\ell=1}^{L} w_{\ell,i}\delta(\epsilon,k). \tag{C.2}$$

∎

The following lemma is used in the proof of Proposition 4.9.

LEMMA C.4. *Let $\overline{\mathbf{M}}_i := \mathbb{E}[\frac{1}{n}\mathbf{M}_i] \in \mathbb{R}^{d\times d}$, where the expectation is taken with respect to the normalized data observations, and let $C > 0$ be a universal constant. Then $\|[\overline{\mathbf{M}}_i, \overline{\mathbf{M}}_j]\| = 0$, and with probability at least $1 - e^{-t}$ for $t > 0$,*

$$\frac{\|\frac{1}{n}\mathbf{M}_i - \overline{\mathbf{M}}_i\|}{\|\overline{\mathbf{M}}_1\|} \leq C\frac{\bar{\sigma}_i}{\bar{\sigma}_1} \max\left\{ \sqrt{\frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i}\log d + t}{n}}, \frac{\frac{\bar{\xi}_i}{\bar{\sigma}_i}\log d + t}{n}\log(n) \right\}, \text{ where} \tag{C.3}$$

$$\bar{\sigma}_i = \|\overline{\mathbf{M}}_i\| = \sum_{\ell=1}^{L} \frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell}+1}\frac{n_\ell}{n}\left(\frac{\lambda_1}{v_\ell}+1\right),$$

$$\bar{\xi}_i = \text{tr}(\overline{\mathbf{M}}_i) = \sum_{\ell=1}^{L} \frac{\frac{\lambda_i}{v_\ell}}{\frac{\lambda_i}{v_\ell}+1}\frac{n_\ell}{n}\left(\frac{1}{v_\ell}\sum_{j=1}^{k}\lambda_j + d\right).$$

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

$$\mathbb{E}[\mathbf{M}_i] = \frac{1}{n}\sum_{\ell=1}^{L}\sum_{j=1}^{n_\ell}\mathbb{E}[\tilde{\mathbf{y}}_{\ell,j}\tilde{\mathbf{y}}'_{\ell,j}] = \sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}\left(\frac{1}{v_\ell}\mathbf{U}\boldsymbol{\Theta}^2\mathbf{U}' + \mathbf{I}\right). \tag{C.4}$$

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

$$\mathbb{E}[\mathbf{M}_i] = \mathbf{U}\left(\sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}\left(\frac{1}{v_\ell}\boldsymbol{\Theta}^2 + \mathbf{I}_k\right)\right)\mathbf{U}' + \left(\sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}\right)\mathbf{U}_\perp\mathbf{U}'_\perp \tag{C.5}$$

$$= \begin{bmatrix}\mathbf{U} & \mathbf{U}_\perp\end{bmatrix}\begin{bmatrix}\boldsymbol{\Sigma}_i & 0\\ 0 & \gamma_i\mathbf{I}_{d-k}\end{bmatrix}\begin{bmatrix}\mathbf{U}'\\ \mathbf{U}'_\perp\end{bmatrix}, \tag{C.6}$$

where $\boldsymbol{\Sigma}_i := \sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}(\frac{1}{v_\ell}\boldsymbol{\Theta}^2 + \mathbf{I}_k)$ and $\gamma_i := \sum_{\ell=1}^{L} w_{\ell,i}\frac{n_\ell}{n}$, from which we obtain the expressions for $\bar{\sigma}_i = \|\mathbb{E}[\mathbf{M}_i]\|$ and $\bar{\xi}_i = \text{tr}(\mathbb{E}[\mathbf{M}_i])$. Then invoking Lemma SM3.14 in the supplement to bound the concentration of a normalized sample covariance matrix to its expectation with high probability yields the final result. ∎

Proof of Proposition 4.9. We argue there are two possible sets of commuting $(\overline{\mathbf{M}}_1, \ldots, \overline{\mathbf{M}}_k)$ 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 $\overline{\mathbf{M}}_i$ by the number of samples and the largest spectral norm of the $\overline{\mathbf{M}}_i$, which is equivalent to also normalizing the distance. Using the definition of the weights $w_{\ell,i}$ in HPPCA, it is straightforward to show that $\mathbf{M}_1 \succeq \mathbf{M}_2 \succeq \cdots \succeq \mathbf{M}_k$. Accordingly, we normalize by $1/\|n\overline{\mathbf{M}}_1\|$.

First, if the variances are zero or all the same, i.e., noiseless or homoscedastic noisy data, then all the $\mathbf{M}_i$ are equal. 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 $\overline{\mathbf{M}} := \frac{1}{n}\sum_{\ell=1}^{L} \bar{v}\mathbf{A}_\ell$ for some $\bar{v} \geq 0 \ \forall \ i \in [k]$, where we recall from (4.3) that $\mathbf{A}_\ell = \sum_{j=1}^{n_\ell} \frac{1}{v_\ell}\mathbf{y}_{\ell,j}\mathbf{y}'_{\ell,j}$. Then

(C.7)
$$\frac{\|\frac{1}{n}\mathbf{M}_i - \overline{\mathbf{M}}\|}{\|\overline{\mathbf{M}}\|} = \frac{\frac{\lambda_i}{\lambda_i+\bar{v}}\|\sum_{\ell=1}^{L}\frac{(\bar{v}-v_\ell)/v_\ell}{\lambda_i/v_\ell+1}\mathbf{A}_\ell\|}{\frac{\lambda_i}{\lambda_i+\bar{v}}\sum_{\ell=1}^{L}\|\mathbf{A}_\ell\|} \leq \frac{\sum_{\ell=1}^{L}\frac{|\bar{v}-v_\ell|/v_\ell}{\lambda_i/v_\ell+1}\|\mathbf{A}_\ell\|}{\|\sum_{\ell=1}^{L}\mathbf{A}_\ell\|} \leq \sum_{\ell=1}^{L}\frac{\frac{|\bar{v}-v_\ell|}{v_\ell}}{\frac{\lambda_i}{v_\ell}+1},$$

where the last inequality above results from the fact $\frac{\|\mathbf{A}_\ell\|}{\|\sum_{\ell=1}^{L}\mathbf{A}_\ell\|} \leq 1$ for all $\ell \in [L]$ using Weyl's inequality for symmetric PSD matrices [26]. While the bound above depends on the SNR and the gaps between the variances, 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 $\overline{\mathbf{c}} = (\overline{\mathbf{M}}_1, \ldots, \overline{\mathbf{M}}_k)$ for $\overline{\mathbf{M}}_i := \mathbb{E}[\frac{1}{n}\mathbf{M}_i]$, where the expectation here is with respect to the normalized data generated by the model in (4.3).

Let $\overline{\mathbf{M}}_i := \mathbb{E}[\frac{1}{n}\mathbf{M}_i] \in \mathbb{R}^{d\times d}$, where the expectation is taken with respect to the normalized data observations. Then by Lemma C.4 and taking the minimum with (C.7), we obtain the final result. ∎

Acknowledgments. The authors thank Nicolas Boumal for his helpful discussions, references, and notes relating to dual certificates of low-rank SDPs and manifold optimization. We also thank David Hong and Jeffrey Fessler for their feedback on this paper and their discussions relating to heteroscedastic PPCA. We also mention and give special thanks to Alex Wang who pointed out an error in a previous version of this manuscript and for his discussions on how to correct it. Lastly, we thank the anonymous reviewers whose suggested revisions helped improve the paper.

REFERENCES

[1] T. E. Abrudan, J. Eriksson, and V. Koivunen, *Steepest descent algorithms for optimization under unitary matrix constraint*, IEEE Trans. Signal Process., 56 (2008), pp. 1134–1147, https://doi.org/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.

[3] B. Afsari, *Sensitivity analysis for the problem of matrix joint diagonalization*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1148–1171, https://doi.org/10.1137/060655997.

[4] A. S. Bandeira, *A note on probably certifiably correct algorithms*, C. R. Math., 354 (2016), pp. 329–333, https://doi.org/10.1016/j.crma.2015.11.009.

[5] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MOS-SIAM Series on Optimization, Society for

Industrial and Applied Mathematics, Philadelphia, PA; Mathematical Programming Society, Philadelphia, PA, 2001, https://doi.org/10.1137/1.9780898718829.

[6] O. A. Berezovskyi, *On the lower bound for a quadratic problem on the Stiefel manifold*, Cybernet. Syst. Anal., 44 (2008), pp. 709–715, https://doi.org/10.1007/s10559-008-9038-4.

[7] M. Bolla, G. Michaletzky, G. Tusnády, and M. Ziermann, *Extrema of sums of heterogeneous quadratic forms*, Linear Algebra Appl., 269 (1998), pp. 331–365, https://doi.org/10.1016/S0024-3795(97)00230-9.

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

[9] N. Boumal, V. Voroninski, and A. Bandeira, *The non-convex Burer-Monteiro approach works on smooth semidefinite programs*, in Proceedings of the Conference on in Neural Information Processing Systems, 2016, https://proceedings.neurips.cc/paper/2016/file/3de2334a314a7a72721f1f74a6cb4cee-Paper.pdf.

[10] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, 2004, https://doi.org/10.1017/CBO9780511804441.

[11] A. Breloy, G. Ginolhac, F. Pascal, and P. Forster, *Clutter subspace estimation in low rank heterogeneous noise context*, IEEE Trans. Signal Process., 63 (2015), pp. 2173–2182, https://doi.org/10.1109/TSP.2015.2403284.

[12] A. Breloy, G. Ginolhac, F. Pascal, and P. Forster, *Robust covariance matrix estimation in heterogeneous low rank context*, IEEE Trans. Signal Process., 64 (2016), pp. 5794–5806, https://doi.org/10.1109/TSP.2016.2599494.

[13] A. Breloy, S. Kumar, Y. Sun, and D. P. Palomar, *Majorization-minimization on the Stiefel manifold with application to robust sparse PCA*, IEEE Trans. Signal Process., 69 (2021), pp. 1507–1520, https://doi.org/10.1109/TSP.2021.3058442.

[14] R. W. Brockett, *Least squares matching problems*, Linear Algebra Appl., 122-124 (1989), pp. 761–777, https://doi.org/10.1016/0024-3795(89)90675-7.

[15] S. Burer and R. D. Monteiro, *A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization*, Math. Program., 95 (2003), pp. 329–357, https://doi.org/10.1007/s10107-002-0352-8.

[16] D. Cifuentes, S. Agarwal, P. A. Parrilo, and R. R. Thomas, *On the local stability of semidefinite relaxations*, Math. Program., 193 (2022), pp. 629–663, https://doi.org/10.1007/s10107-021-01696-1.

[17] M. Dür, B. Jargalsaikhan, and G. Still, *Genericity results in linear conic programming—A tour d'horizon*, Math. Oper. Res., 42 (2017), pp. 77–94, https://doi.org/10.1287/moor.2016.0793.

[18] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 303–353, https://doi.org/10.1137/S0895479895290954.

[19] K. Fan, *On a theorem of Weyl concerning eigenvalues of linear transformations*, Proc. Natl. Acad. Sci., 35 (1949), pp. 652–655, https://doi.org/10.1073/pnas.35.11.652.

[20] P. A. Fillmore and J. P. Williams, *Some convexity theorems for matrices*, Glasgow Math. J., 12 (1971), pp. 110–117, https://doi.org/10.1017/S0017089500001221.

[21] N. Filonov and I. Kachkovskiy, *A Hilbert-Schmidt Analog of Huaxin Lin's Theorem*, preprint, arXiv:1008.4002, 2010.

[22] D. Garber and R. Fisher, *Local linear convergence of gradient methods for subspace optimization via strict complementarity*, in Advances in Neural Information Processing Systems, Vol. 35, 2022, pp. 30486–30498.

[23] M. Grant and S. Boyd, *CVX: Matlab Software for Disciplined Convex Programming, version* 2.1, 2014, http://cvxr.com/cvx.

[24] M. Gu, W. Shen, *Generalized probabilistic principal component analysis of correlated data*, J. Mach. Learn. Res., 21 (2020), pp. 1–41.

[25] D. K. Hong, K. Gilman, L. Balzano, and J. A. Fessler, *HePPCAT: Probabilistic PCA for data with heteroscedastic noise*, IEEE Trans. Signal Process., 69 (2021), pp. 4819–4834, https://doi.org/10.1109/TSP.2021.3104979.

[26] R. A. Horn and C. R. Johnson, eds., *Matrix Analysis*, Cambridge, Cambridge University Press, 1985.

[27] Y. Huang and D. P. Palomar, *Rank-constrained separable semidefinite programming with applications to optimal beamforming*, IEEE Trans. Signal Process., 58 (2009), pp. 664–678, https://doi.org/10.1109/TSP.2009.2031732.

[28] M. Kleinsteuber and H. Shen, *Uniqueness analysis of non-unitary matrix joint diagonalization*, IEEE Trans. Signal Process., 61 (2013), pp. 1786–1796, https://doi.org/10.1109/TSP.2013.2242065.

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

[30] Z.-Q. Luo, T.-H. Chang, D. Palomar, and Y. Eldar, *SDP relaxation of homogeneous quadratic optimization: Approximation*, Convex Optim. Signal Process. Commun., 117 (2010).

[31] M. L. Overton and R. S. Womersley, *On the sum of the largest eigenvalues of a symmetric matrix*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 41–45, https://doi.org/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 (1998), pp. 339–358, https://doi.org/10.1287/moor.23.2.339.

[33] D.-T. Pham and M. Congedo, *Least square joint diagonalization of matrices under an intrinsic scale constraint*, in 8th International Conference on Independent Component Analysis and Signal Separation, Lecture Notes in Computer Science 5441, Springer, Paraty, Brazil, 2009, pp. 298–305, https://doi.org/10.1007/978-3-642-00599-2.

[34] T. Pumir, S. Jelassi, and N. Boumal, *Smoothed analysis of the low-rank approach for smooth semidefinite programs*, in Proceedings of the Conference on Neural Information Processing Systems, 2018, https://proceedings.neurips.cc/paper/2018/file/a1d50185e7426cbb0acad1e6ca74b9aa-Paper.pdf.

[35] T. Rapcsák, *On minimization on Stiefel manifolds*, European J. Oper. Res., 143 (2002), pp. 365–376, https://doi.org/10.1016/S0377-2217(02)00329-6.

[36] X. Shi, *Joint approximate diagonalization method*, in Blind Signal Processing: Theory and Practice, Springer, Berlin, Heidelberg, 2011, pp. 175–204, https://doi.org/10.1007/978-3-642-11347-5_8.

[37] Y. Sun, A. Breloy, P. Babu, D. P. Palomar, F. Pascal, and G. Ginolhac, *Low-complexity algorithms for low rank clutter parameters estimation in radar systems*, IEEE Trans. Signal Process., 64 (2016), pp. 1986–1998, https://doi.org/10.1109/TSP.2015.2512535.

[38] U. Tantipongpipat, S. Samadi, M. Singh, J. H. Morgenstern, and S. Vempala, *Multicriteria dimensionality reduction with applications to fairness*, Proceedings of the Conference on Neural Information Processing Systems, 2019.

[39] F. J. Theis, T. P. Cason, and P. A. Absil, *Soft dimension reduction for ICA by joint diagonalization on the Stiefel manifold*, in Independent Component Analysis and Signal Separation, T. Adali, C. Jutten, J. M. T. Romano, and A. K. Barros, eds., Springer, Berlin, Heidelberg, 2009.

[40] V. Q. Vu, J. Cho, J. Lei, and K. Rohe, *Fantope projection and selection: A near-optimal convex relaxation of sparse PCA*, in Proceedings of the Conference on Neural Information Processing Systems, 2013, pp. 2670–2678.

[41] J.-H. Won, T. Zhang, and H. Zhou, *Orthogonal trace-sum maximization: Tightness of the semidefinite relaxation and guarantee of locally optimal solutions*, SIAM J. Optim., 32 (2022), pp. 2180–2207, https://doi.org/10.1137/21M1422707.

[42] J.-H. Won, H. Zhou, and K. Lange, *Orthogonal trace-sum maximization: Applications, local algorithms, and global optimality*, SIAM J. Matrix Anal. Appl., 42 (2021), pp. 859–882, https://doi.org/10.1137/20M1363388.

[43] A. Yurtsever, J. A. Tropp, O. Fercoq, M. Udell, and V. Cevher, *Scalable semidefinite programming*, SIAM J. Math. Data Sci., 3 (2021), pp. 171–200, https://doi.org/10.1137/19M1305045.