Journal of Optimization Theory and Applications (2020) 186:806–825
https://doi.org/10.1007/s10957-020-01731-9

Two Relaxation Methods for Rank Minimization Problems

April Sagan[1] · Xin Shen[2] · John E. Mitchell[1]

Received: 8 May 2020 / Accepted: 24 July 2020 / Published online: 8 August 2020

Abstract

The problem of minimizing the rank of a symmetric positive semidefinite matrix subject to constraints can be lifted to give an equivalent semidefinite program with complementarity constraints. The formulation requires two positive semidefinite matrices to be complementary. This is a continuous and nonconvex reformulation of the rank minimization problem. We develop two relaxations and show that constraint qualification holds at any stationary point of either relaxation of the rank minimization problem, and we explore the structure of the local minimizers.

Keywords Constraint qualification · Optimality conditions · Rank minimization · Semidefinite programs with complementarity constraints

Mathematics Subject Classification 90C33 · 90C53

1 Introduction to Rank Minimization Problems

Rank constrained optimization problems have received increasing interest because of their wide application in many fields including statistics, communication and signal processing [1,2]. In this paper, we mainly consider one genre of the problems, whose objective is to minimize the rank of a matrix subject to a given set of con-

---

Communicated by Liqun Qi.

✉ John E. Mitchell
mitchj@rpi.edu
https://www.rpi.edu/~mitchj

April Sagan
sagana@rpi.edu

Xin Shen
shenxinhust@gmail.com

1 Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180, USA

2 Yelp Inc., San Francisco, CA 94105, USA

straints. This class of problems has been considered as computationally challenging because of its nonconvex nature, particularly because the rank function is highly discontinuous. Many methods have been developed previously to solve the problem, including nuclear norm approximation [1,3–5]. The lack of theoretical guarantee for these convex approximations for general problems motivates us to turn to the exact formulation of the rank function, which can be constructed as a mathematical program with semidefinite cone complementarity constraints (SDCMPCC), as shown in Sect. 3.

Analogously to the LPCC formulation for ℓ_0 minimization problem [6], the advantage of the SDCMPCC formulation is that it can be expressed as a smooth nonlinear program; thus it can be solved by general nonlinear programming algorithms. The purpose of this paper is to investigate the structure of the SDCMPCC formulation, especially stationary points for two relaxations of it. In general, a local minimizer of an SDCMPCC problem may not satisfy first order optimality conditions because of the complementarity constraints. We have previously shown [7] that the first order optimality conditions do indeed hold at local minimizers of the SDCMPCC lifting of the rank minimization. In the current paper, we show in Sect. 4 that these first order optimality conditions also hold at local minimizers of two different relaxations of the SDCMPCC lifting. The structure of the KKT points is described in Sect. 5. We show convergence of the sequence of KKT points as the relaxation parameter is reduced in Sect. 6, where we also show that limit points have a minimal structure, which we call nondominated.

2 Previous Work on Rank Minimization

The rank minimization problem we consider has the general form

$$\begin{aligned} &\underset{X \in \mathbb{R}^{m\times n}}{\text{minimize}} \ \operatorname{rank}(X) + \phi(X) \\ &\text{subject to } X \in \mathcal{C} \end{aligned} \tag{1}$$

where $\mathbb{R}^{m\times n}$ is the space of size m by n matrices, and $\mathcal{C}$ is the feasible region for X. The function $\phi(X)$ is assumed to be convex and Lipschitz, with Lipschitz parameter L. The best-known class of rank minimization problems is matrix completion, where the objective is to recover a low rank matrix from a sparse set of measurements [8]; collaborative filtering problems fit within this framework. Another example of matrix completion problems arises in security in power networks: multi-channel phasor measurement units (PMUs) are located on many power lines and they provide information about the status of links many times per second. This information is typically of low rank, and recovery from missing or corrupted data can be performed through the solution of rank minimization problems [9,10]. There has been considerable progress on algorithms for matrix completion, including recent work on fast gradient methods [11,12]. However, it is not clear that these approaches will generalize to other rank minimization problems.

Positive semidefinite rank minimization has been applied to the Euclidean distance geometry problem [13]. Given an incomplete distance matrix between points in $\mathbb{R}^d$, where D_{ij} is the distance squared between two points i and j, we want to reconstruct the location of each point. If the location of the points is given by the matrix $P \in \mathbb{R}^{n\times d}$, we hope to find a matrix $B = PP^T$ such that at each (i, j) where D_{ij} is known, $D_{ij} = B_{ii} + B_{jj} - 2B_{ij}$. We also know that the rank of B is d, and so B can be recovered by a rank minimization problem.

Further applications include the recovery of correlation matrices in statistics [14], the solution of systems of quadratic equations [15], computing a sum of squares representation of a polynomial function [16], minimum-order controller design in control theory, and model order reduction in system identification [17]. For a description of some other rank minimization problems in engineering applications, see the survey article [17].

The convex relaxation to the rank is the nuclear norm, which is defined as the sum of its singular values:

$$||X||^* = \sum_{i=1}^{\min(m,n)} \sigma_i = \text{trace}(\sqrt{X^T X})$$

for a matrix $X \in \mathbb{R}^{m\times n}$. In the relaxed problem, the objective is to find a matrix with the minimal nuclear norm. The nuclear norm is convex and continuous. Many algorithms have been developed previously to find the optimal solution to the nuclear norm minimization problem, including interior point methods [3], singular value thresholding [5], Augmented Lagrangian method [18], proximal gradient method [19], subspace selection method [20] and so on. These methods have been shown to be efficient and robust in solving large scale nuclear norm minimization problems in some applications. Previous works have provided an explanation for the good performance for convex approximation by showing that nuclear norm minimization and rank minimization are equivalent under certain assumptions. Let $M \in \mathbb{R}^{n\times n}$ be a rank r matrix. Consider the matrix completion problem defined as

$$\begin{aligned} &\underset{X\in\mathbb{R}^{n\times n}}{\text{minimize}} \ ||X||^* \\ &\text{subject to } \mathcal{P}_\Omega(X) = \mathcal{P}_\Omega(M) \end{aligned} \tag{2}$$

where $\Omega \subseteq \{1, \ldots, n\} \times \{1, \ldots, n\}$ and the projection operator $\mathcal{P}_\Omega : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ is defined as

$$[\mathcal{P}_\Omega(X)]_{ij} = \begin{cases} X_{ij}, & (i, j) \in \Omega, \\ 0, & (i, j) \notin \Omega. \end{cases}$$

It has been shown that unique minimizer of (2) is M with high probability, if $|\Omega| \geq Cnr \log(n)$, for an absolute constant C, under the assumption that the M is incoherent and the entries in Ω are uniformly sampled from its domain $\{1, \ldots, n\} \times \{1, \ldots, n\}$ [21].

While a very strong result, it should be noted the approach can fail, as in the counterexample in [22] and as noted in [14]. Several different approaches have been proposed to solve the nuclear norm minimization problem, including [4,23,24]. A generalization to non-negative rank can be found in [25].

In order to more closely approximate the rank of a matrix, Fazel et. al. proposed the LogDet heuristic for positive semidefinite rank minimization [26]. Instead of a convex function, the authors use the following smooth, concave function as a surrogate for the rank function:

$$\log(\det(X + \gamma I)) = \sum_{i=1}^{n} \log(\lambda_i(X) + \gamma)$$

where $\lambda_i(X)$ denotes the i^{th} largest eigenvalue of X, and γ is a fixed parameter. While nonconvex, the authors put forwards a Majorize-Minimization (MM) algorithm to find a local optimum. At each iteration, the first order Taylor expansion centered at the previous iterate is solved as a surrogate function. The algorithm is simplified to solving the following reweighted nuclear norm minimization problem at each iteration.

$$X^{(k+1)} = \underset{X}{\text{argmin}} \left\{ \langle W^{(k)}, X \rangle \;:\; \mathcal{A}(X) = b,\; X \succeq 0 \right\}$$

where $W^{(k)} = (X^{(k-1)} + \delta I)^{-1}$. With this reweighting, smaller eigenvalues are weighted more heavily than larger ones.

Nonconvex regularizers for rank minimization explored in recent years include the truncated nuclear norm [27], the Schatten p norm [28], and the Minimax Concave Penalty regularizer [29]. By minimizing functions closely estimating the rank, these methods are more successful in obtaining a low rank matrix from noisy measurements in many rank minimization applications.

Another popular technique is to constrain the rank of the matrix by using the low rank factorization. If we are trying to find a low rank matrix X, we work with matrices U and V and define $X = UV^T$; the number of columns of U and rows of V are chosen to be no greater than some parameter r, so the rank of X can then be no larger than r. Alternating minimization approaches to rank minimization problems include [30–33]; see also [34]. While the iteration complexity solving the nuclear norm relaxation can be $O(n^3)$ for $n \times n$ matrices, alternating direction methods run in near linear time; however, the overall complexity depends on the condition number of the matrix.

3 Semidefinite Cone Complementarity Formulations

3.1 Mathematical Program with Semidefinite Cone Complementarity Constraints

A mathematical program with semidefinite cone complementarity constraints (SDCMPCC) is a special case of a mathematical program with complementarity constraints (MPCC). In SDCMPCC problems the constraints include complementarity between matrices rather than vectors. When the complementarity between matrices is replaced by the complementarity between vectors, the problem turns into a standard MPCC. The general SDCMPCC program takes the following form:

$$\underset{z\in\mathbb{R}^n}{\text{minimize}} \left\{ f(z) : g(z) \le 0,\ h(z) = 0,\ \mathbb{S}^n_+ \ni G(z) \perp H(z) \in \mathbb{S}^n_+ \right\} \tag{3}$$

where $\mathbb{S}_n^+$ denotes the cone of $n \times n$ positive semidefinite matrices. The notation $G(z) \perp H(z)$ means the matrices $G(z)$ and $H(z)$ are perpendicular to each other. For matrices $G(z), H(z) \in \mathbb{S}_n^+$, this is equivalent to saying that the Frobenius inner product of $G(z)$ and $H(z)$ is equal to 0, where the Frobenius inner product of two matrices $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{m\times n}$ is defined as:

$$\langle A, B\rangle = \text{trace}(A^T B)$$

SDCMPCC can be written as a nonlinear semidefinite program. Nonlinear semidefinite programs recently received much attention because of their wide application. Yamashita [35] surveyed numerical methods for solving nonlinear SDP programs, including Augmented Lagrangian methods, sequential SDP methods and primal-dual interior point methods. These methods still have much room for research in both theory and practice, especially when the size of problem goes large. SDCMPCC is special case of a nonlinear SDP program. In addition to the difficulties in general nonlinear semidefinite programming, the complementarity constraints pose challenges to finding the local optimal solutions since the KKT condition may not hold at the local optima. Previous works showed that optimality conditions in MPCC, such as M-stationary, C-Stationary and Strong Stationary, can be generalized into SDCMPCC. Ding et al. [36] discussed various kinds of first order optimality conditions of SDCMPCC and their relationship with each other. Zhang [37] provided analysis on second order optimality conditions of SDCMPCC.

Example 3.1 The KKT conditions do not hold at any local optimum for

$$\begin{aligned} &\min_{x\in\mathbb{R}^3} \quad x_3 \\ &\text{subject to } 0 \preceq \begin{bmatrix} x_1 & x_3 \\ x_3 & x_2 \end{bmatrix} \perp \begin{bmatrix} x_2 & 0 \\ 0 & x_1 \end{bmatrix} \succeq 0 \end{aligned}$$

Optimal solutions are all points with $x_3 = 0$ and either $x_1 = 0$ or $x_2 = 0$. Note that the Slater constraint qualification holds if the orthogonality condition is not imposed. There is no solution to the KKT conditions when $x_3 = 0$.

3.2 Complementarity Formulation

In this section, we will present an exact reformulation of rank minimization problem using semidefinite cone constraints. This variational formulation was due to Li and Qi [38], with important work on optimality conditions derived by Ding et al. [36]. Other recent work on this formulation includes [7,39,40]. We begin with a special case of (1), in which the matrix variable $X \in \mathbb{R}^{n\times n}$ is restricted to be symmetric and positive semidefinite. The special case takes the form:

$$\begin{aligned} &\underset{X\in\mathbb{S}^n}{\text{minimize}} \ \text{rank}(X) + \phi(X) \\ &\text{subject to } X \in \tilde{\mathcal{C}}, \quad X \succeq 0. \end{aligned} \tag{4}$$

Assumption 1 Assume $\tilde{C} = \{X \mid \langle A_i, X\rangle = b_i,\ \forall i = 1, \ldots, m_2\}$ where each $g_i(x), i = 1, \ldots, m_1$ is convex. We assume the Slater CQ holds for $\tilde{C} \cap \mathbb{S}^n_+$, so there exists a point $X^c \in \mathbb{S}^n_{++}$ satisfying $\langle A_i, X\rangle = b_i$ for $i = 1, \ldots, m_2$.

From now on, when we discuss problem (4), we assume the feasible region satisfies Assumption 1.

By introducing an auxiliary variable $U \in \mathbb{R}^{n\times n}$, we can model Problem (4) as an SDCMPCC:

$$\begin{array}{ll} \underset{X\in\mathbb{S}^n}{\text{minimize}} & \phi(X) + n - \langle I, U\rangle \\ \text{subject to} & X \in \tilde{C},\ X \succeq 0 \\ & 0 \preceq X \perp U \succeq 0 \\ & 0 \preceq U \preceq I \end{array} \tag{5}$$

The equivalence between Problem (4) and Problem (5) can be verified by a proper assignment of U for given feasible X. Suppose X has the eigenvalue decomposition $X = P^T \Sigma P$. Let P_0 be the matrix composed of columns in P corresponding to zero eigenvalues. We can set:

$$U = P_0 P_0^T \tag{6}$$

We proved the following results regarding the SDCMPCC (5) in [7]:

Proposition 3.1 [7] *Assume $\phi(X) \equiv 0$. Each (X, U) with X feasible and U given by (6) is a local optimal solution in Problem (5).*

The KKT conditions for problem (5) are:

$$\begin{array}{l} 0 \preceq U \perp -I + \mu X + Y \succeq 0 \\ 0 \preceq X \perp \nabla\phi(X) + \sum \lambda_i A_i + \mu U \succeq 0 \\ 0 \preceq Y \perp I - U \succeq 0 \end{array} \tag{7}$$

where λ, μ and Y are Lagrangian multipliers corresponding to the constraints $A(X) = b$, $\langle X, U\rangle = 0$ and $I - U \succeq 0$ respectively.

Proposition 3.2 [7] *The KKT conditions hold at local optima of Problem (5).*

Proposition 3.3 [7] *Assume $\phi(X) \equiv 0$. Any feasible pair (X, U) with U given by (6) is a KKT stationary point of problem (5).*

The above results shows that when $\phi(X) \equiv 0$, similar to the problem of ℓ_0 minimization [6], there are too many KKT stationary points in the exact SDCMPCC Formulation, and it is very likely that algorithms will terminate at a stationary point that might be far from a global optimal. As we have shown in the complementarity formulation for the ℓ_0 minimization problem, a possible approach to overcome this drawback is to relax the complementarity constraints. In the following sections we would like to investigate whether this approach works for the SDCMPCC formulation. We investigated a penalty method for Problem (5) in [7].

The complementarity formulation can be extended to cases where the matrix variable $X \in \mathbb{R}^{m\times n}$ is neither positive semidefinite nor symmetric. One way to deal with nonsymmetric X is to embed it in a $(m+n)\times(m+n)$ symmetric psd matrix of the same rank [3]. Alternatively, we can minimize the rank of the matrix X^TX [28]. The $n\times n$ matrix X^TX is both symmetric and positive semidefinite and we have the following modified complementarity constraint: $\langle U, X^TX\rangle = 0$ where $U \in \mathbb{S}^{n\times n}$.

4 Relaxation Schemes for SDCMPCC Formulation

In this section and following sections, we present two kinds of relaxation schemes for the original SDCMPCC formulation. The first relaxation scheme has the following form:

$$\begin{aligned} \underset{X\in\mathbb{S}^n}{\text{minimize}}\ & \phi(X)+n-\langle I, U\rangle \\ \text{subject to}\ & X\in\tilde{\mathcal{C}},\ X\succeq 0 \\ & \langle X, U\rangle \le \epsilon \\ & 0\preceq U\preceq I \end{aligned} \tag{8}$$

We denote the above problem as $SDCNLP(\epsilon)$. In the relaxed problem, instead of restricting the Frobenius product of X and U to be 0, we allow it to take a value no greater than ϵ. Since it bounds the trace, or the sum of eigenvalues of X^TU, we call this the *aggregate relaxation scheme*.

When the matrices X and U are both positive semidefinite, $\langle X, U\rangle = 0$ is equivalent to $XU+UX=0$. In the other relaxation scheme, we don't force the matrix product of X and U to be the zero matrix. Instead, we require the maximum eigenvalue of their product to be no larger than a positive parameter δ:

$$\begin{aligned} \underset{X\in\mathbb{S}^n}{\text{minimize}}\ & \phi(X)+n-\langle I, U\rangle \\ \text{subject to}\ & X\in\tilde{\mathcal{C}},\ X\succeq 0 \\ & (X+\gamma I)U+U(X+\gamma I)\preceq 2\delta I \\ & 0\preceq U\preceq I \end{aligned} \tag{9}$$

Here, the parameter γ satisfies $0\le\gamma\le\delta$. The terms in γI serve to tighten the relaxation. We denote the above problem as $SDCNLP1(\gamma,\delta)$. Since we allow the matrix product to have maximum eigenvalue δ, we call this relaxation scheme the *matrix relaxation* of the original SDCMPCC formulation.

4.1 Global Convergence of Relaxed Formulations

We can establish global convergence results for both formulations. The proofs are straightforward and left to the reader.

Proposition 4.1 *Let $\{\epsilon_k\}$ be a sequence that converges to 0 and (X_k, U_k) be a global optimal solution to $SDCNLP(\epsilon)$. If $\tilde{\mathcal{C}}$ is closed, then any limit point of the sequence $\{(X_k, U_k)\}$ is a global optimal solution to the exact SDCMPCC formulation (5).*

Proposition 4.2 *Let $\{\gamma_k, \delta_k\}$ be a sequence that converges to $(0, 0)$ and (X_k, U_k) be a global optimal solution to $SDCNLP1(\gamma_k, \delta_k)$. If $\mathcal{C}$ is closed, then any limit point of* *the sequence* $\{(X_k, U_k)\}$ *is a global optimal solution to the exact SDCMPCC formulation* (5).

4.2 Constraint Qualification of Relaxed Formulations

The relaxed formulations for Rank Minimization problem are nonlinear semidefinite programs. As with the exact SDCMPCC Formulation, we would like to investigate whether algorithms for general nonlinear semidefinite programming problems can be applied to solve the relaxed formulations. As far as we know, most algorithms in nonlinear semidefinite programming use first order KKT stationary conditions as the criteria for terminilation.

Our relaxations (8) and (9) of (5) are examples of nonlinear conic optimization problems of the form

$$\min_{z \in Q \subseteq \mathbb{R}^n} \{f(z) \, : \, G(z) \in K\} \tag{10}$$

where Q is a closed convex set, $K \subseteq \mathbb{R}^m$ is a closed convex cone, $f : \mathbb{R}^n \to \mathbb{R}$, and $G : \mathbb{R}^n \to \mathbb{R}^m$, and f and G are differentiable. The first order necessary conditions for (10) at a point $z^0 \in Q$ can be written

$$G(z^0) \in K, \, \lambda \in K^+, \langle \lambda, z^0 \rangle = 0 \tag{11}$$

where K^+ is the dual cone to K. Robinson's constraint qualification [41] for problem (10) can be written

$$0 \in \operatorname{int}\{G(z^0) + DG(z^0)(Q - z^0) - K\} \tag{12}$$

where $DG(z^0)$ denotes the Gateaux derivative of $G(z)$ evaluated at z^0. Very informally, Robinson's CQ holds if there exists a direction $Q - z^0$ leading to a point that strictly satisfies the linearization of the constraints. As shown in Bonnans and Shapiro [42], the first order necessary conditions must hold at a local minimizer z^0 of (10) if Robinson's constraint qualification (12) holds.

We show below that, if Assumption 1 holds, then the feasible regions of both (8) and (9) satisfy the Robinson CQ (12), so any local minimizer must satisfy the first order optimality conditions.

4.2.1 Robinson's CQ for the Aggregate Relaxation

Let $(\bar{X}, \bar{U})$ be a feasible solution to (8). Robinson's CQ holds at this point if we can find symmetric matrices (X^d, U^d) satisfying

$$\bar{X} + X^d \in \mathbb{S}^n_{++} \tag{13a}$$

$$\langle A_i, X^d \rangle = 0,\ i = 1, \ldots, m_2 \tag{13b}$$

$$g_i(\bar{X} + X^d) < 0,\ i = 1, \ldots, m_1 \tag{13c}$$

$$0 \prec \bar{U} + U^d \prec I \tag{13d}$$

$$\langle \bar{X}, \bar{U} \rangle + \langle X^d, \bar{U} \rangle + \langle \bar{X}, U^d \rangle < \epsilon \tag{13e}$$

From Assumption 1, there exists X^c satisfying Slater for $\tilde{C} \cap \mathbb{S}^n_+$, so (13a)–(13c) hold for $X^d = \alpha^X(X^c - \bar{X})$ for $0 < \alpha^X \leq 1$. Since $\bar{U} \in \mathbb{S}^n_+$, it has an eigendecomposition

$$\bar{U} = QDQ^T := \begin{bmatrix} Q_0 & Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & \Lambda \end{bmatrix} \begin{bmatrix} Q_0^T \\ Q_1^T \\ Q_2^T \end{bmatrix} \tag{14}$$

where Λ is a diagonal matrix, with entries strictly between 0 and 1. Our direction U^d is obtained by modifying the diagonal blocks of D. In particular, we construct symmetric matrices U_0, U_1, and U_2 so that

$$U^d = \begin{bmatrix} Q_0 & Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} U_0 & 0 & 0 \\ 0 & U_1 & 0 \\ 0 & 0 & U_2 \end{bmatrix} \begin{bmatrix} Q_0^T \\ Q_1^T \\ Q_2^T \end{bmatrix}. \tag{15}$$

The matrix U_0 must be positive definite and U_1 must be negative definite. The construction first requires a factorization of $\bar{X}$:

$$\text{let } M := Q^T \bar{X} Q, \ \text{ so } \bar{X} = \begin{bmatrix} Q_0 & Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} M_0 & . & . \\ . & M_1 & . \\ . & . & M_2 \end{bmatrix} \begin{bmatrix} Q_0^T \\ Q_1^T \\ Q_2^T \end{bmatrix} \tag{16}$$

where M_0, M_1, and M_2 are the diagonal blocks of M.

Proposition 4.3 *Robinson's CQ holds for the aggregate relaxation* (8) *for any* $\epsilon > 0$ *at any feasible point* $(\bar{X}, \bar{U})$.

Proof We choose

$$U_0 = \alpha_0^U I,\ U_1 = -\alpha_1^U I,\ U_2 = -\alpha_2^U M_2$$

with $\alpha_0^U, \alpha_1^U, \alpha_2^U > 0$. Note that

$$\begin{aligned} \langle X^d, \bar{U} \rangle + \langle \bar{X}, U^d \rangle &= \alpha^X \langle X^c - \bar{X}, \bar{U} \rangle + \langle M_0, U_0 \rangle + \langle M_1, U_1 \rangle + \langle M_2, U_2 \rangle \\ &= \alpha^X \langle X^c - \bar{X}, \bar{U} \rangle + \alpha_0^U \operatorname{trace}(M_0) \\ &\quad -\alpha_1^U \operatorname{trace}(M_1) - \alpha_2^U \langle M_2, M_2 \rangle. \end{aligned}$$

We have three cases.

1. $\langle M_2, M_2\rangle > 0$: In this case, we first select $\alpha_2^U > 0$ while ensuring that $\Lambda - \alpha_2^U M_2$ is positive definite. It is then straightforward to select α with $0 < \alpha^X, \alpha_0^U, \alpha_1^U < 1$ so that $\langle X^d, \bar{U}\rangle + \langle \bar{X}, U^d\rangle < 0$ and so (13) is satisfied.
2. $\langle M_2, M_2\rangle = 0$, $\text{trace}(M_1) > 0$: We take $\alpha_2^U = 0$ and choose $0 < \alpha_1^U < 1$. It is then straightforward to select $0 < \alpha^X, \alpha_0^U < 1$ and the result follows as in the first case.
3. $\langle M_2, M_2\rangle = 0$, $\text{trace}(M_1) = 0$: In this case, $M_1 = 0$ and $M_2 = 0$ so $\langle \bar{X}, \bar{U}\rangle = 0 < \epsilon$. Thus, we can pick $\alpha_2^U = 0$ and small positive values for α^X, α_0^U, and α_1^U so that (13) is satisfied. □

4.2.2 Robinson's CQ for the Matrix Relaxation

Let $(\bar{X}, \bar{U})$ be a feasible solution to (9). Robinson's CQ holds at this point if we can find symmetric matrices (X^d, U^d) satisfying

$$\bar{X} + X^d \in \mathbb{S}^n_{++} \tag{17a}$$

$$\langle A_i, X^d\rangle = 0,\ i = 1, \ldots, m_2 \tag{17b}$$

$$g_i(\bar{X} + X^d) < 0,\ i = 1, \ldots, m_1 \tag{17c}$$

$$0 \prec \bar{U} + U^d \prec I \tag{17d}$$

$$\begin{aligned}&\left(\bar{X} + \gamma I\right)\bar{U} + \bar{U}\left(\bar{X} + \gamma I\right) + X^d\bar{U}\\ &\quad + \bar{U}X^d + \left(\bar{X} + \gamma I\right)U^d + U^d\left(\bar{X} + \gamma I\right) \prec 2\delta I\end{aligned} \tag{17e}$$

Note that $(\bar{X}, \nu\bar{U})$ is feasible in (9) for $0 \leq \nu \leq 1$. We can exploit this in construction of a feasible solution to (17).

Proposition 4.4 *Robinson's CQ holds for the matrix relaxation (9) for any $\delta > 0$ at any feasible point $(\bar{X}, \bar{U})$.*

Proof In the notation of (14) and (15), we construct the direction U^d by taking

$$U_0 = \alpha_0^U I, \quad U_1 = -0.5I, \quad U_2 = -0.5\Lambda,$$

for an appropriate small positive constant α_0^U. Note that from (14), we have

$$U^d = \alpha_0^U Q_0 Q_0^T - 0.5\bar{U}.$$

We take the X-direction as $X^d = \alpha^X(X^c - \bar{X})$ for some α^X. Requirements (17a)–(17d) are satisfied provided $0 < \alpha^X \leq 1$ and $0 < \alpha_0^U < 1$. Note that

$$
\begin{aligned}
&\left(\bar{X}+\gamma I\right)\bar{U}+\bar{U}\left(\bar{X}+\gamma I\right)+X^{d}\bar{U}+\bar{U}X^{d}+\left(\bar{X}+\gamma I\right)U^{d}+U^{d}\left(\bar{X}+\gamma I\right)\\
&\quad=\left(\bar{X}+\gamma I\right)\bar{U}+\bar{U}\left(\bar{X}+\gamma I\right)+\alpha^{X}\left((X^{c}-\bar{X})\bar{U}+\bar{U}(X^{c}-\bar{X})\right)\\
&\qquad+\alpha_0^{U}\left(\left(\bar{X}+\gamma I\right)Q_0Q_0^{T}+Q_0Q_0^{T}\left(\bar{X}+\gamma I\right)\right)\\
&\qquad-0.5\left(\bar{X}+\gamma I\right)\bar{U}+\bar{U}\left(\bar{X}+\gamma I\right)\\
&\quad=0.5\left(\left(\bar{X}+\gamma I\right)\bar{U}+\bar{U}\left(\bar{X}+\gamma I\right)\right)+\alpha^{X}\left((X^{c}-\bar{X})\bar{U}+\bar{U}(X^{c}-\bar{X})\right)\\
&\qquad+\alpha_0^{U}\left(\left(\bar{X}+\gamma I\right)Q_0Q_0^{T}+Q_0Q_0^{T}\left(\bar{X}+\gamma I\right)\right)\\
&\quad\preceq\delta I+\alpha^{X}\left((X^{c}-\bar{X})\bar{U}+\bar{U}(X^{c}-\bar{X})\right)\\
&\qquad+\alpha_0^{U}\left(\left(\bar{X}+\gamma I\right)Q_0Q_0^{T}+Q_0Q_0^{T}\left(\bar{X}+\gamma I\right)\right)\\
&\quad\text{from the feasibility of }(\bar{X},\bar{U})\text{in (9)}\\
&\quad\prec 2\delta I
\end{aligned}
$$

for sufficiently small positive values for α^X and α_0^U. □

5 Local Optimality Condition of Relaxed Formulations

Bonnans and Shapiro [42] gave a description of first order optimality condition in semidefinite programming. Since Robinson's CQ holds at any feasible solution of problem (8) under Assumption 1, the first order KKT conditions must hold at any local optimum. Let $(\bar{X}, \bar{U})$ be a local minimizer to either of the relaxations. We show in this section that either the matrices $\bar{X}$ and $\bar{U}$ are simultaneously diagonalizable, or there exists another feasible solution that is at least as good as $(\bar{X}, \bar{U})$ which is simultaneously diagonalizable. Further, we explore the structure of the KKT points arising in each relaxation.

5.1 The Aggregate Relaxation

Proposition 5.1 *Let $(\bar{X}, \bar{U})$ be a local minimizer of the aggregate relaxation (8). If $\bar{X}$ and $\bar{U}$ are not simultaneously diagonalizable then there exists another matrix $\hat{U}$ satisfying*

1. *$(\bar{X}, \hat{U})$ is feasible in (8).*
2. *$\bar{X}$ and $\hat{U}$ are simultaneously diagonalizable.*
3. *$\langle I, \hat{U}\rangle = \langle I, \bar{U}\rangle$.*

Proof The matrix $\bar{U}$ must be a global minimizer to the convex SDP

$$
\begin{array}{ll}
\min_U & -\langle I, U\rangle \\
\text{subject to} & \langle \bar{X}, U\rangle \le \epsilon \\
& 0 \preceq U \preceq I
\end{array}
\tag{18}
$$

Since $\bar{X} \in \mathbb{S}^n_+$, we can diagonalize it, so $\bar{X} = PDP^T$ for a diagonal matrix D and an orthogonal matrix P. Note that $\langle \bar{X}, U\rangle = \langle D, P^T U P\rangle$ so problem (18) is equivalent to the problem

$$\begin{array}{ll} \min_V & -\langle I, V\rangle \\ \text{subject to} & \langle D, V\rangle \leq \epsilon \\ & 0 \preceq V \preceq I \end{array} \tag{19}$$

with the correspondence $V \leftrightarrow P^T U P$, and exploiting the orthogonality of P. Since I and D are both diagonal matrices, there exists an optimal solution $\hat{V}$ to (19) that is diagonal. The matrix $\hat{U} := P\hat{V}P^T$ is then optimal to (18) and satisfies the three listed criteria. □

The KKT condition in the aggregate relaxed formulation (8) works in a similar way with some thresholding methods. The objective function not only counts the number of 0 eigenvalues, but also the number of eigenvalues whose sum is below a certain threshold. Let $X \in \mathcal{C}$ be feasible in the aggregate relaxation problem, with eigendecompostion

$$X = \sum_{i=1}^{n} \sigma_i^X v_i v_i^T \quad \text{with } \sigma_1^X \leq \ldots \leq \sigma_n^X. \tag{20}$$

Let l be the index such that $\sum_{i=1}^{l} \sigma_i^X \leq \epsilon$, and $\sum_{i=1}^{l+1} \sigma_i^X > \epsilon$. An optimal solution U to problem (18) has the same eigenvectors as X and eigenvalues given by

$$\sigma_j^U = \begin{cases} 1, & j \leq l, \\ \frac{\epsilon - \sum_{i=1}^{l} \sigma_i}{\sigma_{l+1}^X}, & j = l+1, \\ 0, & j > l+1. \end{cases} \tag{21}$$

The optimal solution V to (19) corresponding to this solution U is to take $V_{jj} = \sigma_j^U$ for $1 \leq j \leq n$. Note that $0 \leq \sigma_{l+1}^U < 1$ and that $0 \leq l \leq n-1$.

Note that we must have if $V_{jk} = 0$ for all $1 \leq j < k \leq l+1$, since otherwise the largest eigenvalue of V would be strictly larger than one. If either $j > l+1$ or $k > l+1$ then we must have $V_{jk} = 0$ from the requirement that $V \succeq 0$. It follows that when the eigenvalue σ_{l+1}^X has multiplicity 1, this choice of V is the unique solution to (19). In particular, it is clear from linear programming that the diagonal entries of V must be as specified.

When the eigenvalue σ_{l+1}^X has multiplicity > 1, there are multiple optimal solutions to (19) and (18). However, they all correspond to taking different eigenbases for the eigenspace corresponding to σ_{l+1}^X, so again X and U are simultaneously diagonalizable. Hence we have the following theorem.

Theorem 5.1 *If $(\bar{X}, \bar{U})$ is a local minimizer to (8) then $\bar{X}$ and $\bar{U}$ are simultaneously diagonalizable.*

The KKT local optimality conditions take the following form for the aggregate relaxed formulation:

$$\begin{aligned} 0 &\preceq U \perp -I + \mu X + Y \succeq 0 \\ 0 &\preceq X \perp \nabla\phi(X) + \sum \lambda_i A_i + \mu U \succeq 0 \\ 0 &\preceq Y \perp I - U \succeq 0 \\ 0 &\leq \mu \perp \epsilon - \langle U, X\rangle \geq 0 \end{aligned} \tag{22}$$

for $X \in \tilde{\mathcal{C}}$, where λ, μ and Y are Lagrangian multipliers corresponding to the constraints $A(X) = b$, $\langle X, U\rangle \leq \epsilon$ and $I - U \succeq 0$ respectively. If V is not an optimal solution to (19) then there exists a feasible direction in the U variables which strictly improves the linear part of the objective function for (8). The following proposition then follows.

Proposition 5.2 *If $(\bar{X}, \bar{U})$ is a stationary point to (8) then $\bar{X}$ and $\bar{U}$ are simultaneously diagonalizable. Further, if the eigenvalue $\sigma^U_{l+1} > 0$ then the KKT multiplier in (22) is $\mu = 1/\sigma^X_{l+1}$.*

Proof The complementarity in KKT condition (22) implies that $v_{l+1}^T Y v_{l+1} = 0$ and (since $\sigma^U_{l+1} > 0$) also $v_{l+1}^T(-I + \mu\bar{X} + Y)v_{l+1} = 0$. It follows that $v_{l+1}^T \bar{X} v_{l+1} = \frac{1}{\mu}$, which is to say that $\mu = \frac{1}{\sigma^X_{l+1}}$. □

5.2 The Matrix Relaxation

Proposition 5.3 *Let $(\bar{X}, \bar{U})$ be a local minimizer of the matrix relaxation (9). If $\bar{X}$ and $\bar{U}$ are not simultaneously diagonalizable then there exists another matrix $\hat{U}$ satisfying*

1. *$(\bar{X}, \hat{U})$ is feasible in (9).*
2. *$\bar{X}$ and $\hat{U}$ are simultaneously diagonalizable.*
3. *$\langle I, \hat{U}\rangle = \langle I, \bar{U}\rangle$.*

Proof The matrix $\bar{U}$ must be a global minimizer to the convex SDP

$$\begin{array}{ll} \min_U & -\langle I, U\rangle \\ \text{subject to} & \left(\bar{X} + \gamma I\right) U + U \left(\bar{X} + \gamma I\right) \preceq 2\delta I \\ & 0 \preceq U \preceq I \end{array} \tag{23}$$

Since $\bar{X} \in \mathbb{S}^n_+$, we can diagonalize it, so $\bar{X} = PDP^T$ for a diagonal matrix D and an orthogonal matrix P. Let

$$\tilde{D} := D + \gamma I, \tag{24}$$

so for $\gamma > 0$ we have the diagonal entries $\tilde{D}_{ii} > 0$ for $i = 1, \ldots, n$, and $\tilde{D}$ is invertible. Note that problem (23) is equivalent to the problem

$$\begin{array}{ll} \min_V & -\langle I, V\rangle \\ \text{subject to} & \tilde{D}V + V\tilde{D} \preceq 2\delta I \\ & 0 \preceq V \preceq I \end{array} \tag{25}$$

with the correspondence $V \leftrightarrow P^T U P$, and exploiting the orthogonality of P. The diagonal entries of the matrix product $\tilde{D}V + V\tilde{D}$ are

$$\left(\tilde{D}V + V\tilde{D}\right)_{ii} = 2\tilde{D}_{ii}V_{ii} \quad \text{for } i = 1, \ldots, n.$$

Hence any feasible solution to (25) must satisfy

$$V_{ii} \le \min\left\{1, \frac{\delta}{\tilde{D}_{ii}}\right\}, \qquad \text{for } i = 1, \ldots, n,$$

so the optimal value of (23) and (25) is bounded below by

$$-\langle I, V\rangle \ge -\sum_{i=1}^{n} \min\left\{1, \frac{\delta}{\tilde{D}_{ii}}\right\}.$$

This value is achieved by the diagonal matrix $\hat{V}$, where

$$\hat{V}_{ii} = \min\left\{1, \frac{\delta}{\tilde{D}_{ii}}\right\}, \qquad \text{for } i = 1, \ldots, n. \tag{26}$$

Hence problem (23) is solved by the matrix

$$\hat{U} := P\hat{V}P^T \tag{27}$$

which satisfies the requirements of the proposition. □

We can eliminate U from the matrix relaxation (9) when $\gamma = \delta > 0$. In particular, we can set $U = \delta(X + \delta I)^{-1}$:

Corollary 5.1 *Let* $(\bar{X}, \bar{U})$ *be a feasible solution to* (9) *with* $\gamma = \delta > 0$. *Then* $\left(\bar{X}, \delta\left(\bar{X} + \delta I\right)^{-1}\right)$ *is also feasible in* (9) *and* $\langle I, \bar{U}\rangle \le \langle I, \delta\left(\bar{X} + \delta I\right)^{-1}\rangle$.

Proof Note that from (24) we have $\bar{D}_{ii} \ge \delta$, so from (26) we have $\hat{V}_{ii} = \delta/\bar{D}_{ii}$ for $i = 1, \ldots, n$, so $\hat{V} = \delta\bar{D}^{-1}$. From (27), we then obtain

$$\hat{U} = P\hat{V}P^T = \delta P\bar{D}^{-1}P^T = \delta\left(\bar{X} + \delta I\right)^{-1},$$

as required. □

It follows from this corollary that we can eliminate the variable U from the matrix relaxation (9) when $\gamma = \delta > 0$, obtaining the equivalent nonlinear nonconvex SDP:

$$\begin{aligned} &\underset{X \in \mathbb{S}^n}{\text{minimize}}\ \phi(X) + n - \langle I, \delta(X + \delta I)^{-1}\rangle \\ &\text{subject to } X \in \tilde{\mathcal{C}},\ X \succeq 0. \end{aligned} \tag{28}$$

Corollary 5.2 *Assume $\gamma = \delta > 0$. If there exists an optimal solution $(\bar{X}, \bar{U})$ to (9) then there exists an optimal solution to the problem*

$$\begin{aligned} &\underset{X \in \mathbb{S}^n}{\text{minimize}} \;\; \phi(X) + n - \langle I, U \rangle \\ &\text{subject to} \;\; X \in \tilde{\mathcal{C}},\; X \succeq 0 \\ &\qquad\qquad (X + \gamma I)\, U + U\, (X + \gamma I) \preceq 2\delta I \\ &\qquad\qquad 0 \preceq U \end{aligned} \tag{29}$$

that satisfies $U \preceq I$.

Proof When $\gamma = \delta > 0$, the proof of Proposition 2 does not require the constraint $U \preceq I$. Hence the solution $\left(\bar{X}, \delta\left(\bar{X} + \delta I\right)^{-1}\right)$ also solves the problem without the constraint $U \preceq I$. □

It is straightforward to construct examples, where there exist optimal solutions to the matrix relaxation that are not simultaneously diagonalizable.

Example 5.1 Let

$$\tilde{\mathcal{C}} \cap \mathbb{S}^3_+ = \left\{ \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\} \quad \text{and} \quad \bar{U} = \begin{bmatrix} \frac{1}{4} & \bar{U}_{12} & 0 \\ \bar{U}_{12} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{with } \bar{U}_{12} \text{ unspecified.}$$

Let $\gamma = \delta = 1$. The set of optimal solutions U to (9) includes all symmetric matrices of the form $\bar{U}$ satisfying $0 \preceq \bar{U} \preceq I$. Note that this example shows that the constraint $U \preceq I$ is not redundant even if $\delta = \gamma$. For example, taking $U_{12} = \frac{2}{3}$ satisfies $(X + I)U + U(X + I) \preceq 2I$, $X \succeq 0$, and $U \succeq 0$, but does not satisfy $U \preceq I$. From Corollary 5.2, there exists an optimal solution to (29) where U and X are simultaneously diagonalizable (namely, set all the off-diagonal entries to zero), and for that solution we do obtain $U \preceq I$.

We now consider KKT stationary points of the matrix relaxation (9). Under Assumption 1, any local optimum to the matrix relaxation (9) must satisfy the following KKT conditions:

$$\begin{aligned} &0 \preceq U \;\perp\; -I + \Omega\,(X + \gamma I) + (X + \gamma I)\,\Omega + Y \succeq 0 \\ &0 \preceq X \;\perp\; \nabla\phi(X) + \sum \lambda_i A_i + \Omega U + U \Omega \succeq 0 \\ &0 \preceq Y \;\perp\; I - U \succeq 0 \\ &0 \preceq \Omega \;\perp\; 2\delta I - U\,(X + \gamma I) - (X + \gamma I)\, U \succeq 0 \end{aligned} \tag{30}$$

for $X \in \tilde{\mathcal{C}}$, where λ and Y corresponds to the same constraints as in the formulation (8) and Ω represents the multiplier corresponding to the constraint $2\delta I - U\,(X + \gamma I) - (X + \gamma I)\, U \succeq 0$. Given $\bar{X} \in \tilde{\mathcal{C}} \cap \mathbb{S}^n_+$ with eigenvalue decomposition

$$\bar{X} = \sum \sigma_i^X v_i\, v_i^T, \tag{31}$$

it follows from Proposition 5.3 that we can construct a feasible $U = \sum \sigma_i^U v_i v_i^T$ where σ^U takes the value:

$$\sigma_i^U = \begin{cases} \dfrac{\delta}{\sigma_i^X + \gamma}, & \text{if } \sigma_i^X + \gamma \geq \delta \\ 1, & \text{if } \sigma_i^X + \gamma < \delta \end{cases} \tag{32}$$

Further, this choice of U is a global optimal solution to the problem (23), where $\bar{X}$ is fixed.

6 Local Convergence of KKT Stationary Points

In this section, we show that limit points of KKT stationary points of the relaxation scheme are KKT stationary points of the SDCMPCC formulation. We first consider the aggregate relaxation (8). Given a sequence of relaxation parameters $\epsilon^k \to 0$, we obtain a sequence of KKT points (X_k, U_k) with KKT multpliers μ_k. As we show below in Proposition 6.1, any limit point of the sequence where the multipliers do not diverge is a KKT point of (5). However, Proposition 5.2 opens up the possibility that the KKT multipliers might diverge while the iterates (X_k, U_k) converge.

Proposition 6.1 *Let (X_k, U_k) be a local optimum of the relaxed formulation (8) with relaxation parameter $\{\epsilon_k\}$ and with KKT multipliers (Y_k, μ_k) in the conditions (22). Any limit point of a subsequence $\{(X_k, U_k, Y_k, \mu_k)\}$ as $\epsilon_k \to 0$ is a KKT stationary point of the exact SDCMPCC formulation (5).*

Proof Since the multipliers converge, the KKT conditions (7) are satisfied in the limit. □

Convergence of KKT stationary point can be established for the matrix relaxation formulation (9).

Proposition 6.2 *Let (X_k, U_k) be a local optimum of the relaxed formulation (9) with relaxation parameter $\{\delta_k\}$, and $(\lambda_k, \Omega_k, Y_k)$ be the corresponding Lagrangian multipliers. As $\delta_k \to 0$, any limit point $(\bar{X}, \bar{U}, \bar{\lambda})$ of the sequence $\{(X_k, U_k, \lambda_k)\}$ is a KKT stationary point of the exact SDCMPCC Formulation.*

Proof The feasibility of the limit point can be verified by the continuity of the function $XU + UX$. Since the limit point of $\{(X_k, U_k, \lambda_k)\}$ is bounded, it can be shown that Ω_k and Y_k are also bounded and there is a convergent subsequence with the limit point $(\bar{\Omega}, \bar{Y})$. By the closedness of the semidefinite cone we have:

$$\sum \bar{\lambda}_i A_i + \bar{\Omega} \bar{U} + \bar{U} \bar{\Omega} \succeq 0$$

By continuity the complementarity constraints in KKT condition holds. Thus $(\bar{X}, \bar{U}, \bar{\lambda})$ is a KKT stationary point of the SDCMPCC formulation. □

We further investigate the property of the limit points of local optima of the positive semidefinite relaxation. In the complementarity formulation for ℓ_0 minimization, we showed in [6] that any limit point of local optimal solutions of the complementarity formulation has to be nondominated. We defined a feasible point x^* to be nondominated if there is no other feasible point $\bar{x}$ with

$$|\bar{x}_i| \leq |x_i^*| \ \forall i, \quad \text{with strict inequality for at least one component.}$$

Equivalently, there is no other vector $\bar{x}$ with the same support as x^* with $|\bar{x}|$ smaller than $|x^*|$. Note that the convergence proof does not require an assumption of the restricted isometry property.

We extend this concept to the semidefinite case by working with eigenvalues instead of components. In semidefinite programming, we define nondominated points as:

Definition 6.1 Given a polyhedral subset of the cone of semidefinite matrices $\mathcal{C} = \{X \in \mathbb{S}_+^n | \langle A_i, X \rangle \geq b_i, i = 1, \ldots, m\}$, a matrix $X \in \mathcal{C}$ is called nondominated in $\mathcal{C}$ if there does not exist $Y \in \mathcal{C}$ such that $Y \preceq X$ and $Y \neq X$.

We show that the sequence of solutions to the matrix relaxation converges to a nondominated solution to the original complementarity formulation (5).

Proposition 6.3 *Assume the feasible region of* X *is the intersection of a polyhedron and the cone of positive semidefinite matrices. Assume* $\phi(X) = 0 \, \forall X \in \tilde{\mathcal{C}}$. *Let* $\gamma = \delta$ *in* (9). *The limit point* $(\bar{X}, \bar{U})$ *of local optimal solutions* (X_k, U_k) *to the matrix relaxation formulation* (9) *as* $\delta \to 0$ *must be nondominated.*

Proof The proposition can be validated by contradiction. Assume that the limit point is dominated, then there exists a nontrivial direction $dX \succeq 0$ such that $\bar{X} - dX$ is feasible. We can start with showing that when k is large enough, X_k is dominated.

Suppose $\bar{X}$ has the eigenvalue decomposition:

$$\bar{X} = P^T \begin{bmatrix} D & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P$$

where D is a diagonal matrix with strictly positive diagonal entries. Since $\bar{X} - dX \succeq 0$ and $dX \succeq 0$, there exists a positive semidefinite matrix M with

$$\bar{X} - dX = P^T \begin{bmatrix} M & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} P$$

where $0 \preceq M \preceq D$. By scaling dX, we can assume M is positive definite and for any constraint with $\langle A_i, \bar{X} \rangle > b_i$ we have $\langle A_i, \bar{X} - dX \rangle > b_i$. Let $\beta_k \Delta X_k = X_k - \bar{X}$, with $||\Delta X_k|| = 1$ and $\beta_k \to 0$. The matrix ΔX_k can be written as:

$$\Delta X_k = P^T \begin{bmatrix} G_k & H_k & 0 \\ H_k^T & \Lambda_k & 0 \\ 0 & 0 & 0 \end{bmatrix} P$$

where Λ_k is positive definite. The matrix $\bar{X} - dX + \beta_k \Delta X_k$ can be written as:

$$\bar{X} - dX + \beta_k \Delta X_k = P^T \begin{bmatrix} M + \beta_k G_k & \beta_k H_k & 0 \\ \beta_k H_k^T & \beta_k \Lambda_k & 0 \\ 0 & 0 & 0 \end{bmatrix} P$$

which is p.s.d since the Schur complement $\beta_k \Lambda_k - \beta_k^2 H_k^T (M + \beta_k G_k)^{-1} H_k$ is p.s.d when k is large enough. It can also be easily verified that all the linear constraints hold at $\bar{X} - dX + \beta_k \Delta X_k$ when k is large enough.

Note that $X_k - dX = \bar{X} - dX + \beta_k \Delta X_k \succeq 0$, so X_k is dominated by $X_k - dX$. From (32), the eigenvalues of the auxiliary matrix $\tilde{U}_k$ corresponding to the feasible matrix $X_k - dX$ are

$$\sigma_i^{\tilde{U}_k} = \frac{\delta}{\sigma_i^{X_k - dX} + \gamma} \geq \frac{\delta}{\sigma_i^{X_k} + \gamma} = \sigma_i^{U_k}, \quad \text{with strict inequality for at least one eigenvalue.}$$

Given the assignment of $\tilde{U}_k$, the objective value in (5) corresponding to $X_k - dX$ must be strictly less than that given by X_k, which contradicts the local optimality of (X_k, U_k). □

Note that we cannot extend the nondominated results to the aggregate relaxed formulation. The limit point of the local optima of the aggregate formulation might be dominated.

7 Conclusions

We have investigated two relaxations of the SDCMPCC approach to rank minimization problems. The first relaxation imposes a positive upper bound on the Frobenius inner product of two matrices required to be complementary in the exact solution. The second relaxation exploits the PSD structure more fully through a matrix inequality, which ties together the eigenspaces of the matrices. We showed that Robinson's constraint qualification holds for both relaxations, and hence any local minimizer satisfies the first order necessary conditions. Limit points to the second relaxation have a non-dominated structure. The matrix relaxation also allows an equivalent reformulation (28) without the complementarity variables, and computational results with a variant of this reformulation are contained in a forthcoming paper [43].

Acknowledgements This work was supported in part by the Air Force Office of Sponsored Research under Grants FA9550-08-1-0081 and FA9550-11-1-0260 and by the National Science Foundation under Grant Numbers CMMI-1334327 and DMS-1736326.

References

1. Fazel, M., Hindi, H., Boyd, S.P.: A rank minimization heuristic with application to minimum order system approximation. In: Proceedings of the 2001 American Control Conference, June 2001. IEEE, pp. 4734–4739 (2001). https://faculty.washington.edu/mfazel/nucnorm.html

2. Srebro, N., Jaakkola, T.: Weighted low-rank approximations. In: International Conference on Machine Learning. Atlanta, GA, pp. 720–727 (2003)
3. Liu, Z., Vandenberghe, L.: Interior-point method for nuclear norm approximation with application to system identification. SIAM J. Matrix Anal. Appl. **31**(3), 1235–1256 (2009)
4. Ma, S., Goldfarb, D., Chen, L.: Fixed point and Bregman iterative methods for matrix rank minimization. Math. Program. **128**(1–2), 321–353 (2011)
5. Cai, J.F., Candès, E.J., Shen, Z.: A singular value thresholding algorithm for matrix completion. SIAM J. Optim. **20**(4), 1956–1982 (2010)
6. Feng, M., Mitchell, J.E., Pang, J., Shen, X., Wächter, A.: Complementarity formulations of ℓ_0-norm optimization problems. Pac. J. Optim. **14**(2), 273–305 (2018)
7. Shen, X., Mitchell, J.E.: A penalty method for rank minimization in symmetric matrices. Comput. Optim. Appl. **71**(2), 353–380 (2018). https://doi.org/10.1007/s10589-018-0010-6
8. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. Found. Comput. Math. **9**(6), 717–772 (2009)
9. Gao, P., Wang, M., Ghiocel, S.G., Chow, J.H., Fardanesh, B., Stephopoulos, G.: Missing data recovery by exploiting low-dimensionality in power systems synchrophasor measurements. IEEE Trans. Power Syst. **31**(2), 1006–1013 (2016)
10. Gao, P., Wang, M., Ghiocel, S.G., Chow, J.H., Fardanesh, B., Stephopoulos, G., Razanousky, M.P.: Identification of successive "unobservable" cyber data attacks in power systems. IEEE Trans. Signal Process. **64**(21), 5557–5570 (2016)
11. Jain, P., Netrapalli, P.: Fast exact matrix completion with finite samples. Technical report, Microsoft Research, India (2014)
12. Cherapanamjeri, Y., Gupta, K., Jain, P.: Nearly-optimal robust matrix completion. Technical report, Microsoft Research, India (2016)
13. Alfakih, A.Y., Anjos, M.F., Piccialli, V., Wolkowicz, H.: Euclidean distance matrices, semidefinite programming, and sensor network localization. Port. Math. **68**(1), 53–102 (2011)
14. Miao, W., Pan, S., Sun, D.: A rank-corrected procedure for matrix completion with fixed basis coefficients. Math. Program. **159**, 289–338 (2016)
15. Zhao, Y.B.: An approximation theory of matrix rank minimization and its application to quadratic equations. Linear Algebra Appl. **437**, 77–93 (2013)
16. Recht, B., Fazel, M., Parrilo, P.A.: Guaranteed minimum-rank solutions of linear matrix inequalities via nuclear norm minimization. SIAM Rev. **52**(3), 471–501 (2010)
17. Fazel, M., Hindi, H., Boyd, S.: Rank minimization and applications in system theory. In: Proceedings of the 2004 American Control Conference, Boston, Massachusetts, June 30–Jul 2, 2004, pp. 3273–3278 (2004)
18. Lin, Z., Chen, M., Ma, Y.: The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. Technical report, Perception and Decision Lab, University of Illinois, Urbana-Champaign, IL (2010)
19. Liu, Y.J., Sun, D., Toh, K.C.: An implementable proximal point algorithmic framework for nuclear norm minimization. Math. Program. **133**(1–2), 399–436 (2012)
20. Hsieh, C.J., Olsen, P.: Nuclear norm minimization via active subspace selection. In: International Conference on Machine Learning. Beijing, China, pp. 575–583 (2014)
21. Candès, E.J., Tao, T.: The power of convex relaxation: near-optimal matrix completion. IEEE Trans. Inf. Theory **56**(5), 2053–2080 (2009)
22. Zhang, H., Lin, Z., Zhang, C.: A counterexample for the validity of using nuclear norm as a convex surrogate of rank. In: Machine Learning and Knowledge Discovery in Databases. Springer, Berlin, Germany, pp. 226–241 (2013)
23. Goldfarb, D., Ma, S., Wen, Z.: Solving low-rank matrix completion problems efficiently. In: Forty-Seventh Annual Allerton Conference, Allerton House, UIUC, Illinois, USA, September 30–October 2, pp. 1013–1020 (2009)
24. Luss, R., Teboulle, M.: Conditional gradient algorithms for rank-one matrix approximations with a sparsity constraint. SIAM Rev. **55**(1), 65–98 (2013)
25. Fawzi, H., Parrilo, P.A.: Lower bounds on nonnegative rank via nonnegative nuclear norms. Math. Program. **153**(1), 41–66 (2015)
26. Fazel, M., Hindi, H., Boyd, S.: Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices. In: Proceedings of the 2003 American Control Conference, pp. 2156–2162 (2003)

27. Zhang, D., Hu, Y., Ye, J., Li, X., He, X.: Matrix completion by truncated nuclear norm regularization. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2192–2199 (2012). https://doi.org/10.1109/CVPR.2012.6247927
28. Mohan, K., Fazel, M.: Reweighted nuclear norm minimization with application to system identification. In: Proceedings of the American Control Conference. IEEE, Baltimore, MD, pp. 2953–2959 (2010)
29. Wang, S., Liu, D., Zhang, Z.: Nonconvex relaxation approaches to robust matrix recovery. In: Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence, IJCAI'13. AAAI Press, pp. 1764–1770 (2013). http://dl.acm.org/citation.cfm?id=2540128.2540381
30. Keshavan, R.H., Montanari, A., Oh, S.: Matrix completion from a few entries. IEEE Trans. Inf. Theory **56**(6), 2980–2998 (2010)
31. Li, X., Ling, S., Strohmer, T., Wei, K.: Rapid, robust, and reliable blind deconvolution via nonconvex optimization. Appl. Comput. Harmonic Anal. **47**(3), 893–934 (2019)
32. Sun, R., Luo, Z.Q.: Guaranteed matrix completion via nonconvex factorization. In: 2015 IEEE 56th Annual Symposium on Foundations of Computer Science, pp. 270–289 (2015)
33. Tanner, J., Wei, K.: Low rank matrix completion by alternating steepest descent methods. Appl. Comput. Harmonic Anal. **40**, 417–429 (2016)
34. Wei, K., Cai, J.F., Chan, T.F., Leung, S.: Guarantees of Reimannian optimization for low rank matrix recovery. SIAM J. Matrix Anal. Appl. **37**(3), 1198–1222 (2016)
35. Yamashita, H., Yabe, H.: A survey of numerical methods for nonlinear semidefinite programming. J. Oper. Res. Soc. Jpn. **58**(1), 24–60 (2015)
36. Ding, C., Sun, D., Ye, J.: First order optimality conditions for mathematical programs with semidefinite cone complementarity constraints. Math. Program. **147**(1–2), 539–579 (2014)
37. Zhang, Y., Wu, J., Zhang, L.: First order necessary optimality conditions for mathematical programs with second-order cone complementarity constraints. J. Glob. Optim. **63**(2), 253–279 (2015)
38. Li, Q., Qi, H.D.: A sequential semismooth Newton method for the nearest low-rank correlation matrix problem. SIAM J. Optim. **21**, 1641–1666 (2011)
39. Liu, Y., Bi, S., Pan, S.: Equivalent Lipschitz surrogates for zero-norm and rank optimization problems. Technical report, School of Mathematics, GuangDong University of Technology, Guangzhou, China (2018)
40. Yuan, G., Ghanem, B.: A proximal alternating direction method for semi-definite rank minimization. In: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, vol. AAAI-16, pp. 2300–2308 (2016)
41. Robinson, S.M.: Stability theory for systems of inequalities. Part II: differentiable nonlinear systems. SIAM J. Numer. Anal. **13**(4), 497–513 (1976)
42. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York (2000)
43. Sagan, A., Mitchell, J.E.: Alternating minimization for generalized nonconvex relaxations to rank minmization. Technical report, Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, NY 12180 (2020)