A Constraint-reduced Algorithm for Semidefinite Optimization Problems using HKM and AHO directions

Sungwoo Park *

June 2, 2015

Abstract

We develop a new constraint-reduced infeasible *predictor-corrector* interior point method for semidefinite programming, and we prove that it has polynomial global convergence and superlinear local convergence. While the new algorithm uses HKM direction in predictor step, it adopts AHO direction in corrector step to achieve a faster approach to the central path. In contrast to the previous constraint-reduced algorithm, the proposed algorithm can accomplish superlinear convergence without repeated corrector step due to the fast centering effect of AHO direction.

Keywords: Semidefinite programming, Interior point methods, Constraint reduction, Primal dual infeasible, Global and Local convergence.
AMS Classification: 90C22, 65K05, 90C51

1 Introduction

Interior point method (IPM) generally has faster convergence than active set methods, but requires more computational cost to find a search direction for each iteration. Constraint reduction method can be an effective tool to overcome the expensive computation of IPM's. In this work, we propose a new algorithm for semidefinite programming (SDP), constraint-reduced infeasible *predictor-corrector* IPM, and prove its polynomial global convergence and superlinear local convergence.

Many different IPM's for SDP [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15] have been developed by using different search directions. For example, Potra et al. [14] proposed a *predictor-corrector* infeasible IPM algorithm using HKM direction and showed its global and local convergence. Kojima et al. [5] introduced a modified algorithm with repeating corrector step and showed its local convergence without a fast centering assumption that the previous algorithm [14] required. Later, Potra et al. [13] proved local convergence of the modified algorithm without nondegeneracy assumption used in [5]. On the other hand, Kojima et al. [6] introduced a different *predictor-corrector* algorithm using AHO direction, and they proved its quadratic convergence under a nondegeneracy assumption. Inspired by the good centering effect of AHO direction, Potra and Sheng [12] and Ji et al. [4] developed a superlinearly converging algorithm using the HKM direction in the predictor step and a MZ direction with a bounded scaling matrix.

Constraint reduction method has been applied to different constrained optimization problems: linear programming (LP) [16, 17, 18, 19, 20, 21, 22], support vector machine (SVM) [23, 24], and quadratic programming (QP) [25]. More recently, Park and O'Leary [26, 27] developed a constraint-reduced *predictor-corrector* IPM for SDP and proved its polynomial global convergence. Later, Park [28] introduced a modified algorithm with superlinear local convergence, which repeats corrector step to achieve tangential convergence to the central path.

*KCG holdings, 545 Washington BLVD, Jersey City, NJ 07310, USA. swpark81@gmail.com

$\mathcal{S}^n$	The set of $n \times n$ symmetric matrices
$\mathcal{S}^n_+$	The set of $n \times n$ symmetric positive semidefinite matrices
$\mathcal{S}^n_{++}$	The set of $n \times n$ symmetric positive definite matrices
$\boldsymbol{X} \succ \boldsymbol{0}$	A positive definite matrix
$\boldsymbol{X} \succeq \boldsymbol{0}$	A positive semidefinite matrix
$\boldsymbol{A} \bullet \boldsymbol{B} = \operatorname{tr}\left(\boldsymbol{A}\boldsymbol{B}^T\right)$	The dot-product of matrices
$\mu = (\boldsymbol{X} \bullet \boldsymbol{Z})/n$	The duality gap
$\boldsymbol{x} = \operatorname{vec}(\boldsymbol{X})$	The vectorization of a given matrix $\boldsymbol{X}$, A stack of columns of $\boldsymbol{X}^T$
$\operatorname{mat}(\boldsymbol{x})$	The inverse of $\operatorname{vec}(\boldsymbol{X})$
$\operatorname{symm}(\boldsymbol{X}) = \frac{1}{2}(\boldsymbol{X} + \boldsymbol{X}^T)$	The symmetric part of $\boldsymbol{X}$
$\boldsymbol{G} \otimes \boldsymbol{H}$	Kronecker product of matrices $\boldsymbol{G}$ and $\boldsymbol{H}$
$\boldsymbol{G} \circledast \boldsymbol{H}$	Symmetric Kronecker product of matrices $\boldsymbol{G}$ and $\boldsymbol{H}$
$\\|\boldsymbol{A}\\|$	The 2-norm of a matrix $\boldsymbol{A}$
$\\|\boldsymbol{A}\\|_F = (\sum_{ij} a_{ij}^2)^{1/2}$	The Frobenius norm of a matrix $\boldsymbol{A}$
$\boldsymbol{G}^k = o(1)$	$\left\\|\boldsymbol{G}^k\right\\| \to 0$ as $k \to \infty$
$\boldsymbol{G}^k = O(1)$	$\exists\ \Gamma > 0$ such that $\left\\|\boldsymbol{G}^k\right\\| \leq \Gamma$
$\boldsymbol{G}^k = \Omega(1)$	$\exists\ \Gamma > 0$ such that $1/\Gamma \leq \left\\|\boldsymbol{G}^k\right\\| \leq \Gamma$
$\boldsymbol{G}^k = o(\eta_k)$	$\boldsymbol{G}^k/\eta_k = o(1)$
$\boldsymbol{G}^k = O(\eta_k)$	$\boldsymbol{G}^k/\eta_k = O(1)$
$\boldsymbol{G}^k = \Omega(\eta_k)$	$\boldsymbol{G}^k/\eta_k = \Omega(1)$

Table 1: Notation for the SDP.

In this paper, we introduce a new constraint-reduced IPM for SDP by adopting HKM direction in the predictor step and AHO direction in the corrector step, so it is a constraint-reduced version of the algorithm by Potra et al. [12]. Our algorithm has two essential advantages compared to the previous algorithms. First, in contrast to the previous constraint-reduced IPM developed by Park [28], it does not need the repeated corrector steps to achieve superlinear local convergence. Second, the constraint reduction can make the corrector step with AHO direction more practical. Computing AHO direction requires to solve Lyapunov equations for every iteration, so it is computationally more expensive than HKM direction. Constraint reduction effectively avoids solving unnecessary Lyapunov equations by ignoring *inactive* constraints.

The rest of the paper is organized as follows. In Sect. 2, we define a block-constrained SDP and propose constraint-reduced HKM/AHO directions. Sect. 3 presents the new constraint-reduced algorithm guided by new constraint reduction criteria for predictor and corrector steps. We prove its polynomial global convergence and superlinear local convergences in Sect. 4 and Sect. 5. Finally, in Sect. 6, we summarize our results and suggest future studies.

2 Constraint-reduced Directions

In this paper, we discuss primal-dual SDP problems below under the Slater condition. A list of important notation are summarized in Table 2.

$$\text{Primal SDP:} \qquad \min_{\boldsymbol{X}} \boldsymbol{C} \bullet \boldsymbol{X} \quad \text{s.t. } \boldsymbol{A}_i \bullet \boldsymbol{X} = b_i \text{ for } i = 1, \ldots, m, \quad \boldsymbol{X} \succeq \boldsymbol{0}, \tag{1}$$

$$\text{Dual SDP:} \qquad \max_{\boldsymbol{y},\boldsymbol{Z}} \boldsymbol{b}^T \boldsymbol{y} \quad \text{s.t. } \sum_{i=1}^{m} y_i \boldsymbol{A}_i + \boldsymbol{Z} = \boldsymbol{C}, \quad \boldsymbol{Z} \succeq \boldsymbol{0}, \tag{2}$$

where $\boldsymbol{C} \in \mathcal{S}^n$, $\boldsymbol{A}_i \in \mathcal{S}^n$, $\boldsymbol{X} \in \mathcal{S}^n$, and $\boldsymbol{Z} \in \mathcal{S}^n$.

Assumption 1 (Slater condition)**.** *There exists a primal and dual feasible point* $(\mathbf{X}, \mathbf{y}, \mathbf{Z})$ *such that* $\mathbf{X} \succ \mathbf{0}$ *and* $\mathbf{Z} \succ \mathbf{0}$.

Under Slater condition, $(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{Z}) \in \mathcal{S}^n_{++} \times \mathbb{R}^m \times \mathcal{S}^n_{++}$ is a solution of (1) and (2) if and only if

$$\boldsymbol{A}_i \bullet \boldsymbol{X} = b_i \quad \text{for } i = 1, \ldots, m, \tag{3}$$

$$(\sum_{i=1}^{m} y_i \boldsymbol{A}_i) + \boldsymbol{Z} = \boldsymbol{C}, \tag{4}$$

$$\boldsymbol{X}\boldsymbol{Z} = 0. \tag{5}$$

IPM for SDP iteratively searches for such $(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{Z})$ along the central path, and at each iteration, it determines search directions to reduce primal and dual residuals for (3) and (4) or to approach the central path. Many different search directions have been proposed. For example, HKM direction (Helmberg-Rendl-Vanderbei-Wolkowicz/Kojima-Shindoh-Hara/Monteiro) [1, 7, 8], AHO direction (Alizadeh-Haeberly-Overton) [3], and NT direction (Nesterov-Todd direction) [11].

Among different directions, Zhang [15] and Monterio [9] showed that HKM and AHO directions $(\Delta\boldsymbol{X}, \Delta\boldsymbol{y}, \Delta\boldsymbol{Z})$ can be found in a unified framework by solving

$$\begin{aligned} \boldsymbol{A}_i \bullet \Delta\boldsymbol{X} &= r_{pi}, \quad i = 1, \ldots, m, & (6)\\ \left(\sum_{i=1}^{m} \Delta y_i \boldsymbol{A}_i\right) + \Delta\boldsymbol{Z} &= \boldsymbol{R}_d, & (7)\\ \operatorname{symm}\left(\boldsymbol{P}\left(\boldsymbol{X}\Delta\boldsymbol{Z} + \Delta\boldsymbol{X}\boldsymbol{Z}\right)\boldsymbol{P}^{-1}\right) &= \operatorname{symm}\left(\boldsymbol{P}\boldsymbol{R}_c\boldsymbol{P}^{-1}\right), & (8) \end{aligned}$$

where

$$r_{pi} = b_i - \boldsymbol{A}_i\boldsymbol{x}, \quad \boldsymbol{R}_d = \boldsymbol{C} - \boldsymbol{Z} - \left(\sum_{i=1}^{m} y_i \boldsymbol{A}_i\right), \quad \boldsymbol{R}_c = \mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Z}, \text{ and} \tag{9}$$

$$\boldsymbol{P} = \begin{cases} \boldsymbol{Z}^{1/2} & \text{(for HKM)}, \\ \boldsymbol{I} & \text{(for AHO)}. \end{cases} \tag{10}$$

On the other hand, Alizadeh et al. [3] unified HKM and AHO directions in a different manner by using a reduced form equation. For the unified approach, they defined two different types of vectorizations and Kronecker products: the one for HKM and the other for AHO. For $\boldsymbol{G} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{K} \in \mathcal{S}^n$, we define vectorizations as

$$\begin{aligned} \operatorname{nvec}(\boldsymbol{G}) &:= [g_{11}, \ldots, g_{1n}, g_{21}, \ldots, g_{2n}, \ldots, g_{n1}, \ldots, g_{nn}]^T, \\ \operatorname{svec}(\boldsymbol{K}) &:= [k_{11}, \sqrt{2}k_{12}, \ldots, \sqrt{2}k_{1n}, \sqrt{2}k_{22}, \ldots, \sqrt{2}k_{2n}, \ldots, k_{nn}]^T, \end{aligned}$$

where g_{ij} and k_{ij} are elements at i-th row and j-th column. To discuss both directions in a common framework, we let $\operatorname{vec}(\cdot)$ denote $\operatorname{nvec}(\cdot)$ for HKM and $\operatorname{svec}(\cdot)$ for AHO direction, and let $\operatorname{mat}(\cdot)$ denote their corresponding inverse functions. In addition, we define Kronecker products $\boldsymbol{M} \otimes \boldsymbol{N}$ and $\boldsymbol{M} \circledast \boldsymbol{N}$ so that the following identities hold[1]

$$(\boldsymbol{M} \otimes \boldsymbol{N}) \operatorname{nvec}(\boldsymbol{G}) = \operatorname{nvec}\left(\boldsymbol{M}\boldsymbol{G}\boldsymbol{N}^T\right), \tag{11}$$

$$(\boldsymbol{M} \circledast \boldsymbol{N}) \operatorname{svec}(\boldsymbol{K}) = \operatorname{svec}\left(\operatorname{symm}\left(\boldsymbol{M}\boldsymbol{K}\boldsymbol{N}^T\right)\right), \tag{12}$$

[1] See Alizadeh et al. [3, Appendix] for further properties of symmetric Kronecker product. In this paper, we use *row-ordered* vectorization. In column-ordered vectorization, (11) is replaced by

$$(\boldsymbol{M} \otimes \boldsymbol{N}) \operatorname{nvec}(\boldsymbol{G}) = \operatorname{nvec}\left(\boldsymbol{N}\boldsymbol{G}\boldsymbol{M}^T\right).$$

where $\boldsymbol{M} \in \mathbb{R}^{n\times n}$, $\boldsymbol{N} \in \mathbb{R}^{n\times n}$, $\boldsymbol{G} \in \mathbb{R}^{n\times n}$, and $\boldsymbol{K} \in \mathcal{S}^n$. We also define a matrix $\mathcal{A}$, accumulating $\mathrm{vec}\,(\boldsymbol{A}_i)$ as

$$\mathcal{A} = \begin{bmatrix} \mathrm{vec}\,(\boldsymbol{A}_1)^T \\ \vdots \\ \mathrm{vec}\,(\boldsymbol{A}_m)^T \end{bmatrix}.$$

By using the notation, Alizadeh et al. [3] showed that the HKM[2] and AHO directions can be found by solving a reduced form equation for $\Delta \boldsymbol{y}$,

$$(\mathcal{A}\boldsymbol{E}^{-1}\boldsymbol{F}\mathcal{A}^T)\Delta\boldsymbol{y} = \boldsymbol{r}_p + \mathcal{A}\boldsymbol{E}^{-1}(\boldsymbol{F}\boldsymbol{r}_d - \boldsymbol{r}_c), \tag{13}$$

where

$$\begin{aligned}
\boldsymbol{r}_p &= \boldsymbol{b} - \mathcal{A}\boldsymbol{x}, \quad \boldsymbol{r}_d = \boldsymbol{c} - \boldsymbol{z} - \mathcal{A}^T\boldsymbol{y}, \quad \text{and} \\
\boldsymbol{r}_c &= \begin{cases} \mathrm{nvec}\,(\mu\boldsymbol{I} - \boldsymbol{X}\boldsymbol{Z}) & \text{(for HKM)}, \\ \mathrm{svec}\,(\mu\boldsymbol{I} - \mathrm{symm}\,(\boldsymbol{X}\boldsymbol{Z})) & \text{(for AHO)}. \end{cases} \\
\boldsymbol{E} &= \begin{cases} \boldsymbol{I} \otimes \boldsymbol{Z} & \text{(for HKM)}, \\ \boldsymbol{I} \circledast \boldsymbol{Z} & \text{(for AHO)}, \end{cases} \qquad \boldsymbol{F} = \begin{cases} \boldsymbol{X} \otimes \boldsymbol{I} & \text{(for HKM)}, \\ \boldsymbol{X} \circledast \boldsymbol{I} & \text{(for AHO)}. \end{cases}
\end{aligned}$$

In this paper, we especially focus on the SDP whose matrices $\boldsymbol{A}_i$ and $\boldsymbol{C}$ in (1) and (2) have diagonal block structure, so

$$\boldsymbol{A}_i = \begin{bmatrix} \boldsymbol{A}_{i1} & & & \\ & \boldsymbol{A}_{i2} & & \\ & & \ddots & \\ & & & \boldsymbol{A}_{iq} \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_1 & & & \\ & \boldsymbol{C}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{C}_q \end{bmatrix},$$

where q is the number of blocks. Many practical SDP problems have such diagonal block structure, for which we can partition each primal and dual constraints (1) and (2) into q independent block constraints. For the block-constrained SDP, our *predictor-corrector* algorithm adaptively classifies each block constraint into *active* and *inactive* groups for every iteration. In Sect. 3, we will formally define *active* and *inactive* blocks. For a given block-diagonal matrix $\boldsymbol{G} \in \mathbb{R}^{n\times n}$, we let $\boldsymbol{G}_j$ denote the j-the block of $\boldsymbol{G}$ for $j = 1, \ldots, q$. Without loss of generality, we assume that the *active* and *inactive* blocks are sorted along the diagonal. We let $\widehat{\boldsymbol{G}}$ and $\widetilde{\boldsymbol{G}}$ denote *active* block and *inactive* block, and let $\widehat{\boldsymbol{g}}$ and $\widetilde{\boldsymbol{g}}$ denote their vectorization, so

$$\boldsymbol{G} = \begin{bmatrix} \widehat{\boldsymbol{G}} & \boldsymbol{0} \\ \boldsymbol{0} & \widetilde{\boldsymbol{G}} \end{bmatrix}, \quad \boldsymbol{g} = \mathrm{vec}\,(\boldsymbol{G}) = \begin{bmatrix} \widehat{\boldsymbol{g}} \\ \widetilde{\boldsymbol{g}} \end{bmatrix}.$$

We let the vectorization for block-diagonal matrix exclude zero elements in off-diagonal blocks to simplify notation. We will use the block notation above for any applicable block-structured matrices like $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{A}_i$, $\boldsymbol{E}$, and $\boldsymbol{F}$ unless defined otherwise. We also define $\widehat{\mathcal{A}}$ and $\widetilde{\mathcal{A}}$, *active* and *inactive* blocks of $\mathcal{A} = \left[\widehat{\mathcal{A}}, \widetilde{\mathcal{A}}\right]$, and accumulated matrix $\mathcal{A}_j$ of j-th blocks $\boldsymbol{A}_{ij}$ for $j = 1, \ldots, q$, so

$$\widehat{\mathcal{A}} = \begin{bmatrix} \mathrm{vec}\left(\widehat{\boldsymbol{A}}_1\right)^T \\ \vdots \\ \mathrm{vec}\left(\widehat{\boldsymbol{A}}_m\right)^T \end{bmatrix}, \quad \widetilde{\mathcal{A}} = \begin{bmatrix} \mathrm{vec}\left(\widetilde{\boldsymbol{A}}_1\right)^T \\ \vdots \\ \mathrm{vec}\left(\widetilde{\boldsymbol{A}}_m\right)^T \end{bmatrix}, \quad \mathcal{A}_j = \begin{bmatrix} \mathrm{vec}\,(\boldsymbol{A}_{1j})^T \\ \vdots \\ \mathrm{vec}\,(\boldsymbol{A}_{mj})^T \end{bmatrix}.$$

[2]While $\boldsymbol{E} = (\boldsymbol{Z} \otimes \boldsymbol{I})$ and $\boldsymbol{F} = (\boldsymbol{I} \otimes \boldsymbol{X})$ in Alizadeh et al. [3], $\boldsymbol{E} = (\boldsymbol{I} \otimes \boldsymbol{Z})$ and $\boldsymbol{F} = (\boldsymbol{X} \otimes \boldsymbol{I})$ in our equations because we use *row-ordered* vectorization.

With this notation, Schur complement matrix in (13) can be broken down as

$$(\mathcal{A}\boldsymbol{E}^{-1}\boldsymbol{F}\mathcal{A}^T) = (\widehat{\mathcal{A}}\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}}\widehat{\mathcal{A}}^T) + (\widetilde{\mathcal{A}}\widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{F}}\widetilde{\mathcal{A}}^T) = \sum_{j=1}^{m}(\mathcal{A}_j\boldsymbol{E}_j^{-1}\boldsymbol{F}_j\mathcal{A}_j^T).$$

To solve the reduced form equation (13), the construction of Schur complement matrix $(\mathcal{A}\boldsymbol{E}^{-1}\boldsymbol{F}\mathcal{A}^T)$ is the most expensive, which costs $O(mn^3 + m^2n^2)$ operations. Especially, for AHO direction, we need to solve a Lyapunov equation that finds a matrix $\boldsymbol{H}$ such that

$$\boldsymbol{A}_i\boldsymbol{X} + \boldsymbol{X}\boldsymbol{A}_i = \boldsymbol{H}\boldsymbol{Z} + \boldsymbol{Z}\boldsymbol{H},$$

for given $\boldsymbol{A}_i$, $\boldsymbol{X}$, and $\boldsymbol{Z}$, which requires eigenvalue decompositions of $\boldsymbol{X}$ and $\boldsymbol{Z}$ and a series of matrix multiplications. See Alizadeh et al. [3, Appendix]. To save the compuational cost, Park and O'Leary [26] proposed a constraint-reduced HKM direction by replacing $(\mathcal{A}\boldsymbol{E}^{-1}\boldsymbol{F}\mathcal{A}^T)$ with $(\widehat{\mathcal{A}}\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}}\widehat{\mathcal{A}}^T)$. We extend their idea of constraint reduction to AHO direction. So, instead of (13), we propose to solve a constraint-reduced equation for $\Delta\boldsymbol{y}$,

$$\widehat{\boldsymbol{M}}\Delta\boldsymbol{y} = \boldsymbol{r}_p + \mathcal{A}\boldsymbol{E}^{-1}(\boldsymbol{F}\boldsymbol{r}_d - \boldsymbol{r}_c), \tag{14}$$

where

$$\widehat{\boldsymbol{M}} = (\widehat{\mathcal{A}}\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}}\widehat{\mathcal{A}}^T).$$

Then, we obtain constraint-reduced direction $(\Delta\boldsymbol{x}, \Delta\boldsymbol{z})$ and the perturbation term $\Delta\widetilde{\boldsymbol{x}}_\epsilon$ due to the constraint reduction as

$$\Delta\boldsymbol{z} = \boldsymbol{r}_d - \mathcal{A}^T\Delta\boldsymbol{y} \tag{15}$$

$$\Delta\boldsymbol{x} = \begin{bmatrix} \Delta\widehat{\boldsymbol{x}} \\ \Delta\widetilde{\boldsymbol{x}} \end{bmatrix} = \begin{bmatrix} (\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}})\widehat{\mathcal{A}}^T\Delta\boldsymbol{y} & -(\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}})\widehat{\boldsymbol{r}}_d + \widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{r}}_c \\ & -(\widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{F}})\widetilde{\boldsymbol{r}}_d + \widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{r}}_c \end{bmatrix} \tag{16}$$

$$\Delta\widetilde{\boldsymbol{x}}_\epsilon = (\widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{F}})\widetilde{\mathcal{A}}^T\Delta\boldsymbol{y} \tag{17}$$

In contrast to AHO direction, HKM direction needs explicit symmetrization for $\Delta\boldsymbol{X}$ after (16),

$$\Delta\boldsymbol{X} = \mathrm{symm}\left(\mathrm{mat}\left(\Delta\boldsymbol{x}\right)\right).$$

The lemma below explains how the constraint reduction affects the original directions.

Lemma 2.1 (Constraint-reduced directions)**.** *The solution* $(\Delta\mathbf{X}, \Delta\mathbf{y}, \Delta\mathbf{Z})$ *and* $\Delta\widetilde{\mathbf{x}}_\epsilon$ *of (14)–(17) satisfies (6), (7), and*

$$\mathrm{symm}\left(\mathbf{P}\left(\mathbf{X}\Delta\mathbf{Z} + (\Delta\mathbf{X} + \Delta\mathbf{X}_\epsilon)\mathbf{Z}\right)\mathbf{P}^{-1}\right) = \mathrm{symm}\left(\mathbf{P}\mathbf{R}_c\mathbf{P}^{-1}\right), \tag{8*}$$

where $\mathbf{P}$ *is as defined in (10) and*

$$\Delta\mathbf{X}_\epsilon = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathrm{mat}\left(\Delta\widetilde{\mathbf{x}}_\epsilon\right) \end{bmatrix}.$$

Proof. The proof for HKM direction can be found in [27, Lemma 2.1 ff]. In this proof, we only prove the case of AHO direction. Eq. (7) is satified by (15).
By (14) and (16),

$$\begin{aligned} \mathcal{A}\Delta\boldsymbol{x} &= [\widehat{\mathcal{A}}(\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}})\widehat{\mathcal{A}}^T\Delta\boldsymbol{y} - \widehat{\mathcal{A}}(\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}})\widehat{\boldsymbol{r}}_d + \widehat{\mathcal{A}}\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{r}}_c] + [-\widetilde{\mathcal{A}}(\widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{F}})\widetilde{\boldsymbol{r}}_d + \widetilde{\mathcal{A}}\widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{r}}_c] \\ &= \boldsymbol{r}_p + \mathcal{A}\boldsymbol{E}^{-1}(\boldsymbol{F}\boldsymbol{r}_d - \boldsymbol{r}_c) + [-\widehat{\mathcal{A}}(\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}})\widehat{\boldsymbol{r}}_d + \widehat{\mathcal{A}}\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{r}}_c] \\ &\quad + [-\widetilde{\mathcal{A}}(\widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{F}})\widetilde{\boldsymbol{r}}_d + \widetilde{\mathcal{A}}\widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{r}}_c] = \boldsymbol{r}_p, \end{aligned}$$

which implies to (6). In addition, defining $\Delta\breve{\boldsymbol{x}}$ as,

$$\Delta\breve{\boldsymbol{x}} := \begin{bmatrix} \Delta\widehat{\boldsymbol{x}} \\ \Delta\widetilde{\boldsymbol{x}} + \Delta\widetilde{\boldsymbol{x}}_\epsilon \end{bmatrix},$$

by (16) and (17), we have

$$\begin{aligned} \Delta\breve{\boldsymbol{x}} &= \begin{bmatrix} (\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}})\widehat{\mathcal{A}}^T\Delta\boldsymbol{y} - (\widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{F}})\widehat{\boldsymbol{r}}_d + \widehat{\boldsymbol{E}}^{-1}\widehat{\boldsymbol{r}}_c \\ (\widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{F}})\widetilde{\mathcal{A}}^T\Delta\boldsymbol{y} - (\widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{F}})\widetilde{\boldsymbol{r}}_d + \widetilde{\boldsymbol{E}}^{-1}\widetilde{\boldsymbol{r}}_c \end{bmatrix} \\ &= (\boldsymbol{E}^{-1}\boldsymbol{F})\mathcal{A}^T\Delta\boldsymbol{y} - (\boldsymbol{E}^{-1}\boldsymbol{F})\boldsymbol{r}_d + \boldsymbol{E}^{-1}\boldsymbol{r}_c. \end{aligned}$$

So, by (15),

$$\begin{aligned} \boldsymbol{E}\Delta\breve{\boldsymbol{x}} + \boldsymbol{F}\Delta\boldsymbol{z} &= \boldsymbol{F}\mathcal{A}^T\Delta\boldsymbol{y} - \boldsymbol{F}\boldsymbol{r}_d + \boldsymbol{r}_c + \boldsymbol{F}\Delta\boldsymbol{z} \\ &= \boldsymbol{F}(\mathcal{A}^T\Delta\boldsymbol{y} - \boldsymbol{r}_d + \Delta\boldsymbol{z}) + \boldsymbol{r}_c = \boldsymbol{r}_c. \end{aligned}$$

Therefore, by the definition of $\boldsymbol{E}$ and $\boldsymbol{F}$ for AHO direction and (12),

$$\operatorname{symm}\left(\boldsymbol{X}\Delta\boldsymbol{Z} + (\Delta\boldsymbol{X} + \Delta\boldsymbol{X}_\epsilon)\boldsymbol{Z}\right) = \operatorname{symm}\left(\boldsymbol{R}_c\right). \qquad \square$$

□

3 Algorithm

The new constraint-reduced algorithm is presented in Algorithm SDP:Reduced:HKM:AHO. We define a few essential parameters to explain the algorithm. First, the algorithm uses constants α and β such that

$$\frac{\sqrt{2}\beta^2}{1-(\sqrt{2}+1)\beta} + \frac{\beta^2}{2(1-(\sqrt{2}+1)\beta)^2} \le \alpha < \beta < \frac{\beta}{1-(\sqrt{2}+1)\beta} < 1. \tag{18}$$

For example, we can pick $(\alpha, \beta) = (0.185, 0.2)$. To simplify notation, we define two additional constants g and h as

$$g := \frac{\beta}{1-(\sqrt{2}+1)\beta}, \qquad h := \alpha - (\sqrt{2}\beta g + 0.5g^2). \tag{19}$$

Then, we can rewrite (18) as

$$\sqrt{2}\beta g + 0.5g^2 \le \alpha < \beta < g,$$

so $h \ge 0$. Using the constants, we define a function $f(\delta)$ as

$$f(\delta) := g^2\delta^2 + 2g\beta(g-\beta+1)\delta - 2h\beta^2 = g^2(\delta-\delta^-)(\delta-\delta^+), \tag{20}$$

where $\delta^+ \ge 0$ and $\delta^- \le 0$ are the roots of $f(\delta)$. Because $g^2 > 0$ and $h\beta^2 \ge 0$, there exist such roots δ^+ and δ^-.

We define a neighborhood $\mathcal{N}(\gamma, \tau)$ of the central path as

$$\mathcal{N}(\gamma,\tau) := \{(\boldsymbol{X},\boldsymbol{Z}) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++} : \left\|\boldsymbol{Z}^{1/2}\boldsymbol{X}\boldsymbol{Z}^{1/2} - \tau\boldsymbol{I}\right\|_F \le \gamma\tau\}.$$

In addition, for each iterate $(\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{Z}^k)$ and τ_k, we define ϕ_k as

$$\phi_k := \max\left(\frac{\left\|\boldsymbol{X}^k\boldsymbol{Z}^k\right\|}{\sqrt{\tau_k}}, \sqrt{\tau_k}\right).$$

Algorithm SDP:Reduced:HKM:AHO: Primal-Dual Infeasible Constraint-Reduced Predictor-Corrector Algorithm for Block Diagonal SDP with HKM and AHO directions

1. Input : $\mathcal{A}$, $\boldsymbol{b}$, $\boldsymbol{C}$; α and β satisfying (18); δ^+ defined in (20); convergence tolerance τ^*; $\omega \in (0, 0.5)$ for the perturbation bound of the primal direction in the predictor step.

2. Set $\boldsymbol{X}^0 = \boldsymbol{Z}^0 = \rho \boldsymbol{I}$ for $\rho > \max(||\boldsymbol{X}^*||, ||\boldsymbol{Z}^*||)$, and set $\tau = \tau_0 = \mu_0 = (\boldsymbol{X}^0 \bullet \boldsymbol{Z}^0)/n$ so that $(\boldsymbol{X}^0, \boldsymbol{Z}^0) \in \mathcal{N}(\alpha, \tau_0)$.

3. Repeat until $\tau < \tau^*$: For $k = 0, 1, \ldots,$

   (a) **(Predictor step by HKM):** Set $(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{Z}) = (\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{Z}^k)$, $(\boldsymbol{r}_p, \boldsymbol{r}_d, \tau) = (\boldsymbol{r}_p^k, \boldsymbol{r}_d^k, \tau_k)$.

   i. Sort the constraint blocks in decreasing order of $\left\|(\boldsymbol{I} \otimes \boldsymbol{Z}_j)^{-1}(\boldsymbol{X}_j \otimes \boldsymbol{I})\right\|$.

   ii. Initially, $\widehat{\boldsymbol{M}} = \boldsymbol{0}$. For $j = 1, \ldots, p$, until $\widehat{\boldsymbol{M}}$ is full-rank and Condition 1 is satisfied, update $\widehat{\boldsymbol{M}}$ as $\widehat{\boldsymbol{M}} \leftarrow \widehat{\boldsymbol{M}} + \mathcal{A}_j(\boldsymbol{I} \otimes \boldsymbol{Z}_j)^{-1}(\boldsymbol{X}_j \otimes \boldsymbol{I})\mathcal{A}_j^T$.

   iii. Solve (14) for HKM direction with $\widehat{\boldsymbol{M}}$ and $\boldsymbol{r}_c = \text{nvec}\,(-\boldsymbol{X}\boldsymbol{Z})$ to find $(\Delta\boldsymbol{X}, \Delta\boldsymbol{y}, \Delta\boldsymbol{Z})$. Choose a step length $\overline{\theta} \in [\widehat{\theta}, \breve{\theta}]$ defined by (22) and (23),
   $\overline{\boldsymbol{X}} = \boldsymbol{X} + \overline{\theta}\Delta\boldsymbol{X}, \quad \overline{\boldsymbol{y}} = \boldsymbol{y} + \overline{\theta}\Delta\boldsymbol{y}, \quad \overline{\boldsymbol{Z}} = \boldsymbol{Z} + \overline{\theta}\Delta\boldsymbol{Z}.$

   iv. If $\overline{\theta} = 1$, terminate the iteration with optimal solution $(\overline{\boldsymbol{X}}, \overline{\boldsymbol{y}}, \overline{\boldsymbol{Z}})$.

   (b) **(Corrector step by AHO):** Set $\overline{\tau} = (1 - \overline{\theta})\tau$

   i. Sort the constraint blocks in decreasing order of $\left\|(\boldsymbol{I} \circledast \overline{\boldsymbol{Z}}_j)^{-1}(\overline{\boldsymbol{X}}_j \circledast \boldsymbol{I})\right\|$.

   ii. Initially, $\widehat{\boldsymbol{M}} = \boldsymbol{0}$. For $j = 1, \ldots, p$, until $\widehat{\boldsymbol{M}}$ is full-rank and Condition 2 is satisfied, update $\widehat{\boldsymbol{M}}$ as $\widehat{\boldsymbol{M}} \leftarrow \widehat{\boldsymbol{M}} + \mathcal{A}_j(\boldsymbol{I} \circledast \overline{\boldsymbol{Z}}_j)^{-1}(\overline{\boldsymbol{X}}_j \circledast \boldsymbol{I})\mathcal{A}_j^T$.

   iii. Solve (14) for AHO direction with $\widehat{\boldsymbol{M}} = \widehat{\boldsymbol{M}}_C$, $\boldsymbol{r}_p = \boldsymbol{0}$, $\boldsymbol{r}_d = \boldsymbol{0}$, and $\boldsymbol{r}_c = \text{svec}\left(\overline{\tau}\boldsymbol{I} - \text{symm}\left(\overline{\boldsymbol{X}}\;\overline{\boldsymbol{Z}}\right)\right)$ to find $(\Delta\overline{\boldsymbol{X}}, \Delta\overline{\boldsymbol{y}}, \Delta\overline{\boldsymbol{Z}})$. Take a full step as
   $\overline{\boldsymbol{X}}^+ = \overline{\boldsymbol{X}} + \Delta\overline{\boldsymbol{X}}, \quad \overline{\boldsymbol{y}}^+ = \overline{\boldsymbol{y}} + \Delta\overline{\boldsymbol{y}}, \quad \overline{\boldsymbol{Z}}^+ = \overline{\boldsymbol{Z}} + \Delta\overline{\boldsymbol{Z}}.$

   (c) Update $(\boldsymbol{X}^{k+1}, \boldsymbol{y}^{k+1}, \boldsymbol{Z}^{k+1}) = (\overline{\boldsymbol{X}}^+, \overline{\boldsymbol{y}}^+, \overline{\boldsymbol{Z}}^+)$,
   $\boldsymbol{r}_p^{k+1} = \boldsymbol{b} - \mathcal{A}\boldsymbol{x}^{k+1}$, $\boldsymbol{r}_d^{k+1} = \boldsymbol{c} - \boldsymbol{z}^{k+1} - \mathcal{A}^T\boldsymbol{y}^{k+1}$, and $\tau_{k+1} = \overline{\tau}$,

The algorithm searches for constraint-reduced directions by solving (14) for HKM direction in a predictor step and for AHO direction in a corrector step. As Lemma 2.1 shows, the constraint-reduced directions bring about a perturbation terms $\Delta\boldsymbol{X}_\epsilon$ in the complementarity equation (8*). We quantify the perturbation terms in predictor step and corrector step as

$$\delta_\epsilon := \frac{1}{\tau}\left\|\boldsymbol{Z}^{1/2}\Delta\boldsymbol{X}_\epsilon\;\boldsymbol{Z}^{1/2}\right\|_F, \quad \overline{\delta}_\epsilon := \frac{1}{\overline{\tau}}\left\|\overline{\boldsymbol{Z}}^{1/2}\Delta\overline{\boldsymbol{X}}_\epsilon\;\overline{\boldsymbol{Z}}^{1/2}\right\|_F. \tag{21}$$

For the predictor step, we define the following parameters that will provide lower and upper bounds of its step size.

$$\widehat{\theta} = \frac{(\alpha - \beta - \delta_\epsilon) + \sqrt{(\alpha - \beta - \delta_\epsilon)^2 + 4\delta(\beta - \alpha)}}{2\delta}, \tag{22}$$

$$\breve{\theta} = \max\{\tilde{\theta} \in [0,1] : (\boldsymbol{X} + \overline{\theta}\Delta\boldsymbol{X}, \boldsymbol{y} + \overline{\theta}\Delta\boldsymbol{y}, \boldsymbol{Z} + \overline{\theta}\Delta\boldsymbol{Z}) \in \mathcal{N}(\beta, (1 - \overline{\theta})\tau), \forall\overline{\theta} \in [0, \tilde{\theta}]\}, \tag{23}$$

where

$$\delta = \frac{1}{\tau}\left\|\boldsymbol{Z}^{1/2}\Delta\boldsymbol{X}\Delta\boldsymbol{Z}\boldsymbol{Z}^{-1/2}\right\|_F.$$

Based on the parameters, we can present the constraint reduction criteria as follows.

Condition 1 (Requirement on predictor step's perturbation)**.** *For a given positive constant* C_ϵ*,*

$$\delta_\epsilon \le \min(\frac{\omega}{\tau_k}\delta_x, C_\epsilon\phi_k),$$

where $0 < \omega < 0.5$ *and* $\delta_x := \left\|\mathbf{Z}^{1/2}\Delta\mathbf{X}\mathbf{Z}^{1/2}\right\|_F$.

Condition 2 (Requirement on corrector step's perturbation)**.** *For a given positive constant* $\overline{C}_\epsilon$*,*

$$\overline{\delta}_\epsilon \le \min\left(\delta^+, \overline{C}_\epsilon \overline{\tau}^{0.5}\right) \quad and \quad \overline{\delta}_\epsilon < 1 - (2+\sqrt{2})\beta.$$

where δ^+ *is a nonnegative root of (20).*

In both predictor and corrector step, we incrementally constructs $\widehat{\boldsymbol{M}}$, so

$$\begin{aligned}
\widehat{\boldsymbol{M}} &\leftarrow \widehat{\boldsymbol{M}} + \mathcal{A}_j(\boldsymbol{I}\otimes \boldsymbol{Z}_j)^{-1}(\boldsymbol{X}_j \otimes \boldsymbol{I})\mathcal{A}_j^T \quad \text{(for predictor)},\\
\widehat{\boldsymbol{M}} &\leftarrow \widehat{\boldsymbol{M}} + \mathcal{A}_j(\boldsymbol{I}\circledast \overline{\boldsymbol{Z}}_j)^{-1}(\overline{\boldsymbol{X}}_j \circledast \boldsymbol{I})\mathcal{A}_j^T \quad \text{(for corrector)}.
\end{aligned}$$

until the resulting $\widehat{\boldsymbol{M}}$ has a full rank and satisfies the corresponding condition: Condition 1 for the predictor step and Condition 2 for the corrector step. Note that the incremental construction can come up with $\widehat{\boldsymbol{M}}$ satisfying the conditions because δ_ϵ and $\overline{\delta}_\epsilon$ becomes zero by including all blocks in $\widehat{\boldsymbol{M}}$, in which case no constraints are reduced. We say the blocks included in the final $\widehat{\boldsymbol{M}}$ *active* and the remaining blocks *inactive*.

To efficiently construct the matrix $\widehat{\boldsymbol{M}}$, we sort the blocks in an decreasing orders of $\left\|(\boldsymbol{I}\otimes \boldsymbol{Z}_j)^{-1}(\boldsymbol{X}_j \otimes \boldsymbol{I})\right\|$ and $\left\|(\boldsymbol{I}\circledast \overline{\boldsymbol{Z}}_j)^{-1}(\overline{\boldsymbol{X}}_j \circledast \boldsymbol{I})\right\|$ because the greater norm of a block implies to the more contributions to Schur complement matrix. Because we do not need these norms in a high accuracy for sorting the blocks, we can approximate them from the estimates of $\|\boldsymbol{X}_j\|$, $\|\boldsymbol{Z}_j^{-1}\|$, $\|\overline{\boldsymbol{X}}_j\|$, and $\left\|\overline{\boldsymbol{Z}}_j^{-1}\right\|$ to avoid the additional computation.

Once we construct $\widehat{\boldsymbol{M}}$ satisfying the corresponding conditions, we update the iterate as follows. In predictor step, we solve (6), (7), and (8*) for HKM direction $(\Delta\boldsymbol{X}, \Delta\boldsymbol{y}, \Delta\boldsymbol{Z})$ with setting $\mu = 0$, and we update the iterate as

$$(\overline{\boldsymbol{X}}, \overline{\boldsymbol{y}}, \overline{\boldsymbol{Z}}) = (\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{z}) + (1-\overline{\theta})(\Delta\boldsymbol{X}, \Delta\boldsymbol{y}, \Delta\boldsymbol{Z}),$$

where $\widehat{\theta} \le \overline{\theta} \le \breve{\theta}$. Park and O'Leary [27, Lemma 3.2] showed that such a step size $\overline{\theta}$ exists under Condition 1. On the other hand, in corrector step, we solve (6), (7), and (8*) for AHO direction with setting $r_{pi} = 0$, $\boldsymbol{R}_d = \boldsymbol{0}$, and $\mu = \overline{\tau}$. Thus, we find $(\Delta\overline{\boldsymbol{X}}, \Delta\overline{\boldsymbol{y}}, \Delta\overline{\boldsymbol{Z}})$ such that

$$\boldsymbol{A}_i \bullet \Delta\overline{\boldsymbol{X}} = 0, \quad i = 1, \ldots, m, \tag{24}$$

$$\sum_{i=1}^m \Delta\overline{y}_i \boldsymbol{A}_i + \Delta\overline{\boldsymbol{Z}} = \boldsymbol{0}, \tag{25}$$

$$\operatorname{symm}\left(\overline{\boldsymbol{X}}\Delta\overline{\boldsymbol{Z}} + (\Delta\overline{\boldsymbol{X}} + \Delta\overline{\boldsymbol{X}}_\epsilon)\overline{\boldsymbol{Z}}\right) = \overline{\tau}\boldsymbol{I} - \operatorname{symm}\left(\overline{\boldsymbol{X}}\,\overline{\boldsymbol{Z}}\right). \tag{26}$$

Then, we update the iterate as

$$(\boldsymbol{X}^+, \boldsymbol{y}^+, \boldsymbol{Z}^+) = (\overline{\boldsymbol{X}}, \overline{\boldsymbol{y}}, \overline{\boldsymbol{z}}) + (\Delta\overline{\boldsymbol{X}}, \Delta\overline{\boldsymbol{y}}, \Delta\overline{\boldsymbol{Z}}). \tag{27}$$

4 Polynomial Global Convergence

In this section, we show Algorithm SDP:Reduced:HKM:AHO has polynomial global convergence. We define feasible and optimal sets of $(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{Z})$,

$$\begin{aligned}
\mathcal{F}^0 &:= \{(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{Z}) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n : \ (\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{Z}) \text{ satisfies (3) and (4). }\},\\
\mathcal{F} &:= \{(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{Z}) \in \mathcal{F}^0 : (\boldsymbol{X}, \boldsymbol{Z}) \in \mathcal{S}^n_+ \times \mathcal{S}^n_+\},\\
\mathcal{F}^* &:= \{(\boldsymbol{X}, \boldsymbol{y}, \boldsymbol{Z}) \in \mathcal{F} : \boldsymbol{X} \bullet \boldsymbol{Z} = 0\}.
\end{aligned}$$

The following two preliminary lemmas are useful for the convergence analysis.

Lemma 4.1 (Monteiro [8], Lemma 3.3)**.** *Suppose that* $\mathbf{M} \in \mathbb{R}^{p\times p}$ *is a nonsingular matrix. Then, for any* $\mathbf{G} \in \mathcal{S}^p$*, we have*

$$\|\mathbf{G}\|_F \le \left\|\mathrm{symm}\left(\mathbf{M}\mathbf{G}\mathbf{M}^{-1}\right)\right\|_F.$$

Lemma 4.2 (Potra and Sheng [12], Lemma 2.4)**.** *Let* $(\mathbf{X},\mathbf{Z}) \in \mathcal{N}(\gamma,\tau')$ *for some* $\gamma \in [0,\sqrt{2}-1]$ *and* $\tau' > 0$*. Suppose that* $(\mathbf{D}_x, \Delta\mathbf{y}, \mathbf{D}_z) \in (\mathbb{R}^{n\times n}\times\mathbb{R}^m\times\mathbb{R}^{n\times n})$ *is a solution of the linear equations:*

$$\begin{aligned}\mathrm{symm}\left(\mathbf{X}\mathbf{D}_z + \mathbf{D}_x\mathbf{Z}\right) &= \mathrm{symm}\left(\mathbf{H}\right),\\ \mathbf{A}_i \bullet \mathbf{D}_x = 0, \quad i &= 1,\dots,m,\\ \sum_{i=1}^m \Delta y_i \mathbf{A}_i + \mathbf{D}_z &= \mathbf{0},\end{aligned}$$

for some $\mathbf{H}$ *such that* $\mathbf{Z}^{1/2}\mathbf{H}\mathbf{Z}^{-1/2} \in \mathcal{S}^n$*. Then, we have*

$$\left\|\mathbf{Z}^{1/2}(\mathbf{X}\mathbf{D}_z + \mathbf{D}_x\mathbf{Z} - \mathbf{H})\mathbf{Z}^{-1/2}\right\|_F \le \sqrt{2}\gamma\delta'_z \tag{28}$$

$$\delta'\tau'^2 \le \frac{1}{2}\left(\delta_x'^2 + \delta_z'^2\right) \le \frac{\left\|\mathbf{Z}^{1/2}\mathbf{H}\mathbf{Z}^{-1/2}\right\|_F^2}{2(1-(\sqrt{2}+1)\gamma)^2}, \tag{29}$$

where

$$\begin{aligned}\delta' &= \frac{1}{\tau'}\left\|\mathbf{Z}^{1/2}\mathbf{D}_x\mathbf{D}_z\mathbf{Z}^{-1/2}\right\|_F,\\ \delta'_x &= \left\|\mathbf{Z}^{1/2}\mathbf{D}_x\mathbf{Z}^{1/2}\right\|_F,\\ \delta'_z &= \left\|\tau'\mathbf{Z}^{-1/2}\mathbf{D}_z\mathbf{Z}^{-1/2}\right\|_F^2.\end{aligned}$$

Based on the preliminary lemmas, we show how the iterate and its residuals change after the predictor and corrector steps.

Lemma 4.3 (After Predictor Step)**.** *For* $(\mathbf{X},\mathbf{Z}) \in \mathcal{N}(\alpha,\tau)$*, after the predictor step,*

$$\overline{\mathbf{R}}_p = (1-\overline{\theta})\mathbf{R}_p, \quad \overline{\mathbf{R}}_d = (1-\overline{\theta})\mathbf{R}_d, \quad \overline{\tau} = (1-\overline{\theta})\tau.$$

If $\overline{\theta} < 1$*,*

$$(\overline{\mathbf{X}},\overline{\mathbf{Z}}) \in \mathcal{N}(\beta,\overline{\tau}).$$

If $\overline{\theta} = 1$*,* $(\overline{\mathbf{X}},\overline{\mathbf{y}},\overline{\mathbf{Z}})$ *is a solution of the SDP.*

Proof. Park and O'Leary [27, Lemma 3.2] proved this for α and β such that

$$\frac{\beta^2}{2(1-\beta)^2} < \alpha < \beta \le \frac{\beta}{1-\beta} < 1.$$

It can be easily shown that (18) is sufficient to the inequality above. □ □

Lemma 4.4 (After Corrector Step)**.** *For* $(\overline{\mathbf{X}},\overline{\mathbf{Z}}) \in \mathcal{N}(\beta,\overline{\tau})$*, after the predictor step,*

$$(\mathbf{X}^+,\mathbf{Z}^+) \in \mathcal{N}(\alpha,\tau^+), \tag{30}$$

$$\mathbf{R}_p^+ = \overline{\mathbf{R}}_p, \quad \mathbf{R}_d^+ = \overline{\mathbf{R}}_d. \tag{31}$$

where $\overline{\mathbf{R}}_p$ *and* $\overline{\mathbf{R}}_d$ *are as defined in Lemma 4.3.*

Proof. By primal and dual equations (24) and (25), Eq. (31) is evident, so we focus on showing $(\boldsymbol{X}^+, \boldsymbol{Z}^+) \in \mathcal{N}(\alpha, \tau^+)$. By (27),

$$\begin{aligned}(\boldsymbol{X}^+\boldsymbol{Z}^+ - \overline{\tau}\boldsymbol{I}) &= (\overline{\boldsymbol{X}} + \Delta\overline{\boldsymbol{X}})(\overline{\boldsymbol{Z}} + \Delta\overline{\boldsymbol{Z}}) - \overline{\tau}\boldsymbol{I} \\ &= (\overline{\boldsymbol{X}}\ \overline{\boldsymbol{Z}} + \overline{\boldsymbol{X}}\Delta\overline{\boldsymbol{Z}} + \Delta\overline{\boldsymbol{X}}\overline{\boldsymbol{Z}} - \overline{\tau}\boldsymbol{I}) + \Delta\overline{\boldsymbol{X}}\Delta\overline{\boldsymbol{Z}} \\ &= \boldsymbol{J} + \Delta\overline{\boldsymbol{X}}\Delta\overline{\boldsymbol{Z}}, \end{aligned} \tag{32}$$

where

$$\boldsymbol{J} = \overline{\boldsymbol{X}}\ \overline{\boldsymbol{Z}} + \overline{\boldsymbol{X}}\Delta\overline{\boldsymbol{Z}} + \Delta\overline{\boldsymbol{X}}\overline{\boldsymbol{Z}} - \overline{\tau}\boldsymbol{I}.$$

Because $(\overline{\boldsymbol{X}}, \overline{\boldsymbol{Z}}) \in \mathcal{N}(\beta, \overline{\tau})$ and $(\overline{\boldsymbol{X}}, \overline{\boldsymbol{y}}, \overline{\boldsymbol{Z}})$ satisfies (24)–(26), we can use Lemma 4.2 with setting $\boldsymbol{H} = \overline{\tau}\boldsymbol{I} - \overline{\boldsymbol{X}}\ \overline{\boldsymbol{Z}} - \Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}}$, $\gamma = \beta$, and $\tau' = \overline{\tau}$.
By using (21) and $(\overline{\boldsymbol{X}}, \overline{\boldsymbol{Z}}) \in \mathcal{N}(\beta, \overline{\tau})$

$$\begin{aligned}&\left\|\overline{\boldsymbol{Z}}^{1/2}\boldsymbol{H}\overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F = \left\|\overline{\boldsymbol{Z}}^{1/2}(\overline{\tau}\boldsymbol{I} - \overline{\boldsymbol{X}}\ \overline{\boldsymbol{Z}} - \Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}})\overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F \\ &\leq \left\|\overline{\tau}\boldsymbol{I} - \overline{\boldsymbol{Z}}^{1/2}\overline{\boldsymbol{X}}\ \overline{\boldsymbol{Z}}^{1/2}\right\|_F + \left\|\overline{\boldsymbol{Z}}^{1/2}\Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}}^{1/2}\right\|_F = (\beta + \overline{\delta}_\epsilon)\overline{\tau}. \end{aligned} \tag{33}$$

Thus, by Lemma 4.2,

$$\begin{aligned}\left\|\overline{\boldsymbol{Z}}^{1/2}(\boldsymbol{J} + \Delta\overline{\boldsymbol{X}}_\epsilon)\overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F &= \left\|\overline{\boldsymbol{Z}}^{1/2}(\overline{\boldsymbol{X}}\Delta\overline{\boldsymbol{Z}} + \Delta\overline{\boldsymbol{X}}\ \overline{\boldsymbol{Z}} - \overline{\tau}\boldsymbol{I} + \overline{\boldsymbol{X}}\ \overline{\boldsymbol{Z}} + \Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}})\overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F \\ &= \left\|\overline{\boldsymbol{Z}}^{1/2}(\overline{\boldsymbol{X}}\Delta\overline{\boldsymbol{Z}} + \Delta\overline{\boldsymbol{X}}\ \overline{\boldsymbol{Z}} - \boldsymbol{H})\overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F \\ &\leq \sqrt{2}\beta\delta_z' \leq \sqrt{2}\beta\frac{\left\|\overline{\boldsymbol{Z}}^{1/2}\boldsymbol{H}\overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F}{1 - (\sqrt{2}+1)\beta} \quad (\because (28) \text{ and } (29)). \end{aligned}$$

Then, we can rewrite the inequality above as,

$$\begin{aligned}\left\|\overline{\boldsymbol{Z}}^{1/2}\boldsymbol{J}\overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F &\leq \left\|\overline{\boldsymbol{Z}}^{1/2}\Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F + \sqrt{2}\beta\frac{\left\|\overline{\boldsymbol{Z}}^{1/2}\boldsymbol{H}\overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F}{1 - (\sqrt{2}+1)\beta} \\ &\leq \overline{\delta}_\epsilon\overline{\tau} + \frac{\sqrt{2}\beta}{1 - (\sqrt{2}+1)\beta}(\beta + \overline{\delta}_\epsilon)\overline{\tau} \quad (\because (21) \text{ and } (33)) \\ &= \frac{\sqrt{2}\beta^2\overline{\tau}}{1 - (\sqrt{2}+1)\beta} + \frac{1-\beta}{1 - (\sqrt{2}+1)\beta}\overline{\delta}_\epsilon\overline{\tau}. \end{aligned} \tag{34}$$

Again by (29) and (33),

$$\left\|\overline{\boldsymbol{Z}}^{1/2}\Delta\boldsymbol{X}\Delta\boldsymbol{Z}\boldsymbol{Z}^{-1/2}\right\|_F \leq \frac{\left\|\overline{\boldsymbol{Z}}^{1/2}\boldsymbol{H}\overline{\boldsymbol{Z}}^{\ -1/2}\right\|_F^2}{2(1 - (\sqrt{2}+1)\beta)^2\overline{\tau}} \leq \frac{(\beta + \overline{\delta}_\epsilon)^2\overline{\tau}}{2(1 - (\sqrt{2}+1)\beta)^2} \tag{35}$$

Define

$$\boldsymbol{B} := \boldsymbol{Z}^{1/2}(\boldsymbol{X}^+\boldsymbol{Z}^+ - \overline{\tau}\boldsymbol{I})\overline{\boldsymbol{Z}}^{\ -1/2}.$$

By (32), (34), and (35), with recalling definition of g and h in (19),

$$\begin{aligned}
\|\mathrm{symm}\,(\boldsymbol{B})\|_F &\leq \left\|\overline{\boldsymbol{Z}}^{1/2}\boldsymbol{J}\overline{\boldsymbol{Z}}^{-1/2}\right\|_F + \left\|\overline{\boldsymbol{Z}}^{1/2}\Delta\boldsymbol{X}\Delta\boldsymbol{Z}\overline{\boldsymbol{Z}}^{-1/2}\right\|_F \;\; (\because (32)) \\
&\leq \left(\frac{\sqrt{2}\beta^2\overline{\tau}}{1-(\sqrt{2}+1)\beta} + \frac{1-\beta}{1-(\sqrt{2}+1)\beta}\overline{\delta}_\epsilon\overline{\tau}\right) \\
&\quad + \left(\frac{(\beta+\overline{\delta}_\epsilon)^2\overline{\tau}}{2(1-(\sqrt{2}+1)\beta)^2}\right) \;\; (\because (34) \text{ and } (35)) \\
&= \sqrt{2}\overline{\tau}g\beta + \frac{\overline{\tau}g\overline{\delta}_\epsilon(1-\beta)}{\beta} + \frac{g^2\overline{\tau}(\beta+\overline{\delta}_\epsilon)^2}{2\beta^2} \;\; (\because \text{def. of } g \text{ in } (19)) \\
&= \frac{\overline{\tau}(g^2\overline{\delta}_\epsilon^2 + 2g\beta(g-\beta+1)\overline{\delta}_\epsilon - 2hg^2)}{2\beta^2} + \alpha\overline{\tau} \;\; (\because \text{def. of } h \text{ in } (19)).
\end{aligned}$$

By definition of $f(\delta)$ in (20),

$$\|\mathrm{symm}\,(\boldsymbol{B})\|_F \leq \frac{\overline{\tau}}{2\beta^2}f(\overline{\delta}_\epsilon) + \alpha\overline{\tau}.$$

By Condition 2, $\overline{\delta}_\epsilon \in [0, \delta^+]$, so $f(\overline{\delta}_\epsilon) \leq 0$. Thus,

$$\|\mathrm{symm}\,(\boldsymbol{B})\|_F \leq \alpha\overline{\tau}. \tag{36}$$

Again by (29), (33), and Condition 2,

$$\begin{aligned}
\left\|\overline{\boldsymbol{Z}}^{-1/2}\Delta\boldsymbol{Z}\overline{\boldsymbol{Z}}^{-1/2}\right\|_F &\leq \frac{\left\|\overline{\boldsymbol{Z}}^{1/2}\boldsymbol{H}\overline{\boldsymbol{Z}}^{-1/2}\right\|_F}{(1-(\sqrt{2}+1)\beta)\overline{\tau}} \;\; (\because (29)) \\
&\leq \frac{(\beta+\overline{\delta}_\epsilon)\overline{\tau}}{(1-(\sqrt{2}+1)\beta)\overline{\tau}} = \frac{(\beta+\overline{\delta}_\epsilon)}{(1-(\sqrt{2}+1)\beta)} \;\; (\because (33)) \\
&< \frac{(\beta+1-(2+\sqrt{2})\beta)}{(1-(\sqrt{2}+1)\beta)} = 1. \;\; (\because \text{Condition 2})
\end{aligned}$$

Therefore,

$$\begin{aligned}
\boldsymbol{I} + \overline{\boldsymbol{Z}}^{-1/2}\Delta\boldsymbol{Z}\overline{\boldsymbol{Z}}^{-1/2} &\succ \boldsymbol{0} \\
\boldsymbol{Z}^+ = \overline{\boldsymbol{Z}} + \Delta\overline{\boldsymbol{Z}} &\succ \boldsymbol{0},
\end{aligned}$$

so $(\boldsymbol{Z}^+)^{1/2}$ and its inverse exist. We set $\boldsymbol{G}$ and $\boldsymbol{H}$ in Lemma 4.1 as

$$\begin{aligned}
\boldsymbol{G} &= (\boldsymbol{Z}^+)^{1/2}\boldsymbol{X}^+(\boldsymbol{Z}^+)^{1/2} - \overline{\tau}\boldsymbol{I}, \\
\boldsymbol{M} &= \overline{\boldsymbol{Z}}^{1/2}(\boldsymbol{Z}^+)^{-1/2}.
\end{aligned}$$

Note that

$$\boldsymbol{B} = \boldsymbol{M}\boldsymbol{G}\boldsymbol{M}^{-1} = \overline{\boldsymbol{Z}}^{1/2}(\boldsymbol{X}^+\boldsymbol{Z}^+)\overline{\boldsymbol{Z}}^{-1/2} - \overline{\tau}\boldsymbol{I}.$$

Then, by Lemma 4.1 together with (36),

$$\left\|(\boldsymbol{Z}^+)^{1/2}\boldsymbol{X}^+(\boldsymbol{Z}^+)^{1/2} - \overline{\tau}\boldsymbol{I}\right\|_F \leq \|\mathrm{symm}\,(\boldsymbol{B})\|_F \leq \alpha\overline{\tau}.$$

Thus, $\lambda_{\min}((\boldsymbol{Z}^+)^{1/2}\boldsymbol{X}^+(\boldsymbol{Z}^+)^{1/2}) \geq (1-\alpha)\overline{\tau} > 0$, so $(\boldsymbol{Z}^+)^{1/2}\boldsymbol{X}^+(\boldsymbol{Z}^+)^{1/2} \succ \boldsymbol{0}$. Because $\boldsymbol{Z}^+ \succ \boldsymbol{0}$, $\boldsymbol{X}^+ \succ \boldsymbol{0}$. Therefore, $(\boldsymbol{X}^+, \boldsymbol{Z}^+) \in \mathcal{N}(\alpha, \overline{\tau}) = \mathcal{N}(\alpha, \tau^+)$. □ □

Finally, we can show the polynomial global convergence from Lemma 4.3 and 4.4.

Theorem 4.1 (Polynomial global convergence)**.**

$$(\mathbf{X}^k, \mathbf{Z}^k) \in \mathcal{N}(\alpha, \tau_k), \tag{37}$$

$$(\overline{\mathbf{X}}^k, \overline{\mathbf{Z}}^k) \in \mathcal{N}(\beta, (1-\overline{\theta}_k)\tau_k), \tag{38}$$

$$\mathbf{R}_p^k = \psi_k \mathbf{R}_p^0, \quad \mathbf{R}_d^k = \psi_k \mathbf{R}_d^0, \quad \tau_k = \psi_k \tau_0, \tag{39}$$

$$n(1-\alpha)\tau_k \leq \mathbf{X}^k \bullet \mathbf{Z}^k \leq n(1+\alpha)\tau_k. \tag{40}$$

where

$$\psi_k = \prod_{j=0}^{k-1}(1-\overline{\theta}_j).$$

In addition, $\overline{\theta}_k$ *is bounded away from* 0*, so the algorithm polynomially converges.*

Proof. (37)–(39) are direct results by Lemma 4.3 and 4.4. We can show (40) from (37). From the results (37)–(40), we can show $\overline{\theta}_k$ is bounded away from 0. See Park and O'Leary [27, Lemma 3.5–3.8 and Theorem 3.2]. □ □

5 Superlinear Local Convergence

In this section, we prove superlinear convergence of Algorithm SDP:Reduced:HKM:AHO. The proofs are organized as follows: First, under strict complementarity assumption, we show asymptotic bounds of $(\overline{\boldsymbol{X}}, \overline{\boldsymbol{Z}})$ in Lemma 5.1 and $\Delta\overline{\boldsymbol{X}}_\epsilon \overline{\boldsymbol{Z}}$ in Lemma 5.2. Second, based on the asymptotic bounds, Lemma 5.3 and 5.4 present asymptotic bound of $(\Delta\overline{\boldsymbol{X}}, \Delta\overline{\boldsymbol{Z}})$ under nondegeneracy assumption. Finally, Theorem 5.1 proves superlinear local convergence by using the lemmas above. In the proof, we also utilize asymptotic bounds of η_k by Lemma 5.5 and $(1-\overline{\theta}_k)$ by Lemma 5.6.

Assumption 2 (Strict complementarity)**.** *The SDP problem has a strictly complementary solution* $(\mathbf{X}^*, \mathbf{y}^*, \mathbf{Z}^*) \in \mathcal{F}^*$ *such that* $\mathbf{X}^* + \mathbf{Z}^* \succ \mathbf{0}$*.*

Lemma 5.1 (Potra et al. [12], Lemma 3.5)**.** *Under Assumption 2, for the iterate satisfying (37)–(40),*

$$\mathbf{Q}^T(\overline{\mathbf{X}}^{k-1})^{1/2}\mathbf{Q} = \begin{bmatrix} O(1) & O(\sqrt{\tau_k}) \\ O(\sqrt{\tau_k}) & O(\sqrt{\tau_k}) \end{bmatrix}, \quad \mathbf{Q}^T(\overline{\mathbf{X}}^{k-1})^{-1/2}\mathbf{Q} = \begin{bmatrix} O(1) & O(1) \\ O(1) & O(1/\sqrt{\tau_k}) \end{bmatrix},$$
$$\mathbf{Q}^T(\overline{\mathbf{Z}}^{k-1})^{1/2}\mathbf{Q} = \begin{bmatrix} O(\sqrt{\tau_k}) & O(\sqrt{\tau_k}) \\ O(\sqrt{\tau_k}) & O(1) \end{bmatrix}, \quad \mathbf{Q}^T(\overline{\mathbf{Z}}^{k-1})^{-1/2}\mathbf{Q} = \begin{bmatrix} O(1/\sqrt{\tau_k}) & O(1) \\ O(1) & O(1) \end{bmatrix}.$$

Lemma 5.2 (Asymptotic bound of $\Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}}$)**.**

$$\left\|\Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}}\right\|_F = O(\overline{\tau})$$

Proof. By Condition 2 and Lemma 5.1,

$$\begin{aligned}
\left\|\Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}}\right\|_F &\leq \left\|(\boldsymbol{Q}^T\overline{\boldsymbol{Z}}^{\ -1/2}\boldsymbol{Q})(\boldsymbol{Q}^T\overline{\boldsymbol{Z}}^{1/2}\Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}}^{1/2}\boldsymbol{Q})(\boldsymbol{Q}^T\overline{\boldsymbol{Z}}^{1/2}\boldsymbol{Q})\right\|_F \\
&\leq \left\|\boldsymbol{Q}^T\overline{\boldsymbol{Z}}^{\ -1/2}\boldsymbol{Q}\right\|_F \left\|\boldsymbol{Q}^T\overline{\boldsymbol{Z}}^{1/2}\Delta\overline{\boldsymbol{X}}_\epsilon\ \overline{\boldsymbol{Z}}^{1/2}\boldsymbol{Q}\right\|_F \left\|\boldsymbol{Q}^T\overline{\boldsymbol{Z}}^{1/2}\boldsymbol{Q}\right\|_F \\
&= O(1/\sqrt{\overline{\tau}})(\overline{\delta}_\epsilon\overline{\tau})O(1) = O(\overline{\delta}_\epsilon\overline{\tau}^{1/2}) = O(\overline{\tau}). \quad (\because \text{Condition 2}) \qquad \square
\end{aligned}$$

□

In the following lemmas, we show the asymptotic bounds of the corrector step $(\Delta\overline{\boldsymbol{X}}, \Delta\overline{\boldsymbol{Z}})$ under nondegeneracy assumption. For the nondegeneracy assumption, we define a null space $\mathcal{G}_0$ against primal and dual constraints as

$$\begin{aligned}\mathcal{G}_0 = \{(\boldsymbol{U}', \boldsymbol{V}') \in \mathcal{S}^n \times \mathcal{S}^n \quad : \quad & \boldsymbol{A}_i \bullet \boldsymbol{U}' = 0,\ i = 1, \dots, m \\ & \sum_{i=1}^{m} \boldsymbol{w}'_i \boldsymbol{A}_i + \boldsymbol{V}' = \boldsymbol{0},\ \boldsymbol{w}_i \in \mathbb{R}^m\}.\end{aligned}$$

Assumption 3 (Nondegeneracy)**.** $(\mathbf{U}, \mathbf{V}) = \mathbf{0}$ *if* $\mathbf{X}^*\mathbf{V} + \mathbf{U}\mathbf{Z}^* = \mathbf{0}$ *and* $(\mathbf{U}, \mathbf{V}) \in \mathcal{G}_0$

Lemma 5.3 (Kojima et al. [6], Lemma 6.3)**.** *Under Assumption 3, if* $(\mathbf{U}, \mathbf{V}) \in \mathcal{G}_0$ *satisfies*

$$\mathrm{symm}\,(\mathbf{X}^*\mathbf{V} + \mathbf{U}\mathbf{Z}^*) = \mathbf{0},$$

then $(\mathbf{U}, \mathbf{V}) = (\mathbf{0}, \mathbf{0})$.

Lemma 5.4 (Similar to Potra and Sheng [12], Lemma 4.2)**.** *Under Assumption 3*

$$\left\|(\Delta\overline{\mathbf{X}}^{k-1}, \Delta\overline{\mathbf{Z}}^{k-1})\right\|_F = O(\sqrt{\tau_k})$$

Proof. By Lemma 5.1,

$$\left\|\overline{\boldsymbol{X}}^{k-1}\overline{\boldsymbol{Z}}^{k-1} + \overline{\boldsymbol{Z}}^{k-1}\overline{\boldsymbol{X}}^{k-1}\right\|_F = O(\sqrt{\tau_k}).$$

By (24) and (25),

$$\left(\frac{\Delta\overline{\boldsymbol{X}}^{k-1}}{\sqrt{\tau_k}}, \frac{\Delta\overline{\boldsymbol{Z}}^{k-1}}{\sqrt{\tau_k}}\right) \in \mathcal{G}_0.$$

On the other hand, by (26), Lemma 5.1, and Lemma 5.2,

$$\begin{aligned} & \mathrm{symm}\left(\overline{\boldsymbol{X}}^{k-1}\frac{\Delta\overline{\boldsymbol{Z}}^{k-1}}{\sqrt{\tau_k}} + \frac{\Delta\overline{\boldsymbol{X}}^{k-1}}{\sqrt{\tau_k}}\overline{\boldsymbol{Z}}^{k-1}\right) \\ = \; & \mathrm{symm}\left(\tau_k \boldsymbol{I} - \overline{\boldsymbol{X}}^{k-1}\overline{\boldsymbol{Z}}^{k-1} - \Delta\overline{\boldsymbol{X}}_{\epsilon}^{k-1}\overline{\boldsymbol{Z}}^{k-1}\right)/\sqrt{\tau_k} \\ = \; & O(\sqrt{\tau_k}) + O(1) + O(\sqrt{\tau_k}) = O(1). \end{aligned} \tag{41}$$

Assume that $\left\{\frac{(\Delta\overline{\boldsymbol{X}}^{k-1}, \Delta\overline{\boldsymbol{Z}}^{k-1})}{\sqrt{\tau_k}}\right\}$ is not bounded. Then, we can find a subsequence such that

$$\frac{(\Delta\overline{\boldsymbol{X}}^{k-1}, \Delta\overline{\boldsymbol{Z}}^{k-1})}{\sqrt{\tau_k}} \to \infty.$$

We define $\Delta\overline{\boldsymbol{X}}'$ and $\Delta\overline{\boldsymbol{Z}}'$ as

$$\frac{(\Delta\overline{\boldsymbol{X}}^{k-1}, \Delta\overline{\boldsymbol{Z}}^{k-1})}{\left\|(\Delta\overline{\boldsymbol{X}}^{k-1}, \Delta\overline{\boldsymbol{Z}}^{k-1})\right\|_F} \to (\Delta\overline{\boldsymbol{X}}', \Delta\overline{\boldsymbol{Z}}') \neq (\mathbf{0}, \mathbf{0}), \quad (\Delta\overline{\boldsymbol{X}}', \Delta\overline{\boldsymbol{Z}}') \in \mathcal{G}_0.$$

By dividing both sides of (41) by $\left\|(\Delta\overline{\mathbf{X}}_{k-1}, \Delta\overline{\mathbf{Z}}_{k-1})\right\|_F/\sqrt{\tau_k}$ and letting $k \to \infty$ along the subsequence, we have

$$\mathrm{symm}\left(\boldsymbol{X}^*\Delta\overline{\boldsymbol{Z}}' + \Delta\overline{\boldsymbol{X}}'\boldsymbol{Z}^*\right) = \mathbf{0},$$

for $(\Delta\overline{\boldsymbol{X}}', \Delta\overline{\boldsymbol{Z}}') \neq (\mathbf{0}, \mathbf{0})$ and $(\Delta\overline{\boldsymbol{X}}', \Delta\overline{\boldsymbol{Z}}') \in \mathcal{G}_0$. This is a contradiction to Lemma 5.3. □ □

To proceed with the further proofs, we need to define a set $\mathcal{M}$ and a shortest distance η_k of current iterate $(\boldsymbol{X}^k, \boldsymbol{Z}^k)$ to the set $\mathcal{M}$. Let $(\boldsymbol{X}^*, \boldsymbol{y}^*, \boldsymbol{Z}^*)$ be the strict complementary solution, and let $\boldsymbol{Q}$ be the orthogonal matrix whose columns are the eigenvectors of $\boldsymbol{X}^*$ and $\boldsymbol{Z}^*$. We define mutually exclusive sets $\boldsymbol{B}$ and $\boldsymbol{N}$ as

$$\boldsymbol{B} := \{i : \boldsymbol{q}_i^T \boldsymbol{X}^* \boldsymbol{q}_i > 0\}, \qquad \boldsymbol{N} := \{i : \boldsymbol{q}_i^T \boldsymbol{Z}^* \boldsymbol{q}_i > 0\},$$

such that $\boldsymbol{B} \cup \boldsymbol{N} = \{1, 2, \ldots, n\}$ where $\boldsymbol{q}_i$ is the i-th column of $\boldsymbol{Q}$. We also define a set $\mathcal{M}$ of $(\boldsymbol{X}', \boldsymbol{y}', \boldsymbol{Z}')$ as

$$\mathcal{M} := \{(\boldsymbol{X}', \boldsymbol{y}', \boldsymbol{Z}') \in \mathcal{F}^0 : \boldsymbol{q}_i^T \boldsymbol{X}' \boldsymbol{q}_j = 0 \text{ if } i \text{ or } j \in \boldsymbol{N} \text{ and } \boldsymbol{q}_i^T \boldsymbol{Z}' \boldsymbol{q}_j = 0 \text{ if } i \text{ or } j \in \boldsymbol{B}\}.$$

For a given $(\boldsymbol{X}^k, \boldsymbol{y}^k, \boldsymbol{Z}^k)$, we let $(\breve{\boldsymbol{X}}^k, \breve{\boldsymbol{y}}^k, \breve{\boldsymbol{Z}}^k)$ denote a solution[3] of

$$\min_{\boldsymbol{X}', \boldsymbol{y}', \boldsymbol{Z}'} \left\| (\boldsymbol{Z}^k)^{1/2} (\boldsymbol{X}^k - \boldsymbol{X}')(\boldsymbol{Z}^k - \boldsymbol{Z}')(\boldsymbol{Z}^k)^{-1/2} \right\|_F$$

such that $(\boldsymbol{X}', \boldsymbol{y}', \boldsymbol{Z}') \in \mathcal{M}$ and $\left\| [\boldsymbol{X}', \boldsymbol{Z}'] \right\|_F \leq \Gamma$ for a constant Γ where $\left\| [\boldsymbol{X}^k, \boldsymbol{Z}^k] \right\|_F \leq \Gamma$ for $\forall k$. Then, we define η_k as

$$\eta_k := \frac{1}{\tau_k} \left\| (\boldsymbol{Z}^k)^{1/2} (\boldsymbol{X}^k - \breve{\boldsymbol{X}}^k)(\boldsymbol{Z}^k - \breve{\boldsymbol{Z}}^k)(\boldsymbol{Z}^k)^{-1/2} \right\|_F . \tag{42}$$

In Theorem 5.1, we will prove superlinear local convergence by showing that $(1 - \overline{\theta}_k)$ in Theorem 4.1 converges to zero. In the proof, the following two lemmas are used for the asymptotic bounds of η_k and $(1 - \overline{\theta}_k)$.

Lemma 5.5 (Potra et al. [13], Theorem 6.1)**.** *For the iterate satisfying (37)–(40), under Assumption 2, if*

$$\lim_{k \to \infty} \mathbf{X}^k \mathbf{Z}^k / \sqrt{\tau_k} = 0,$$

then

$$\eta_k = O(\phi_k).$$

Lemma 5.6 (Park [28], Lemma 4.6)**.** *For the iterate satisfying (37)–(40), under Assumption 2,*

$$(1 - \overline{\theta}_k) = O(\eta_k + \delta_\epsilon).$$

Theorem 5.1 (Similar to Potra and Sheng [12], Theorem 4.3)**.** *Under Assumption 2 and 3, Algorithm SDP:Reduced:HKM:AHO is superlinear convergent with Q-order at least 1.5.*

Proof. By using (26) and (27), we have

$$\begin{aligned} \text{symm}\left(\boldsymbol{X}^+ \boldsymbol{Z}^+\right) &= \text{symm}\left(\overline{\boldsymbol{X}}\, \overline{\boldsymbol{Z}} + \overline{\boldsymbol{X}} \Delta \overline{\boldsymbol{Z}} + \Delta \overline{\boldsymbol{X}}\, \overline{\boldsymbol{Z}} + \Delta \overline{\boldsymbol{X}} \Delta \overline{\boldsymbol{Z}}\right) \\ &= \text{symm}\left(\overline{\tau} \boldsymbol{I} + \Delta \overline{\boldsymbol{X}} \Delta \overline{\boldsymbol{Z}} - \Delta \overline{\boldsymbol{X}}_\epsilon\, \overline{\boldsymbol{Z}}\right) \end{aligned}$$

By Lemma 5.2 and 5.4, we can obtain the bound of above as

$$\text{symm}\left(\boldsymbol{X}^+ \boldsymbol{Z}^+\right) = O(\overline{\tau}) = O(\tau^+)$$

On the other hand,

$$\begin{aligned} \left\| 2\text{symm}\left(\boldsymbol{X}^+ \boldsymbol{Z}^+\right) \right\|_F^2 &= \left\| \boldsymbol{X}^+ \boldsymbol{Z}^+ + \boldsymbol{Z}^+ \boldsymbol{X}^+ \right\|_F^2 \\ &= 2 \left\| \boldsymbol{X}^+ \boldsymbol{Z}^+ \right\|_F^2 + 2\text{tr}\left((\boldsymbol{X}^+ \boldsymbol{Z}^+)^2\right) \\ &= 2 \left\| \boldsymbol{X}^+ \boldsymbol{Z}^+ \right\|_F^2 + 2 \left\| (\boldsymbol{X}^+)^{1/2} \boldsymbol{Z}^+ (\boldsymbol{X}^+)^{1/2} \right\|_F^2 \\ &\geq 2 \left\| \boldsymbol{X}^+ \boldsymbol{Z}^+ \right\|_F^2 , \end{aligned}$$

[3] See Potra et al. [14, pp18–19] for the existence of the solution.

so

$$\left\|\boldsymbol{X}^{+}\boldsymbol{Z}^{+}\right\|_F \le \sqrt{2}\left\|\mathrm{symm}\left(\boldsymbol{X}^{+}\boldsymbol{Z}^{+}\right)\right\|_F = O(\tau^{+}).$$

Thus,

$$\phi_k = \max\left(\frac{\left\|\boldsymbol{X}^{k}\boldsymbol{Z}^{k}\right\|}{\sqrt{\tau_k}}, \sqrt{\tau_k}\right) = O(\sqrt{\tau_k}),$$

so $\left\|\boldsymbol{X}^{k}\boldsymbol{Z}^{k}\right\|/\sqrt{\tau_k} \to 0$. By Lemma 5.5, Lemma 5.6, and Condition 1,

$$(1-\overline{\theta}_k) = O(\eta_k + \delta_\epsilon) = O(\phi_k) = O(\sqrt{\tau_k}).$$

Therefore, by Theorem 4.1,

$$\tau_{k+1} = (1-\overline{\theta}_k)\tau_k = O(\tau_k^{1.5}). \qquad \square$$

$\square$

6 Conclusions

In this study, we introduced a constraint-reduced *predictor-corrector* IPM, Algorithm SDP:Reduced:HKM:AHO, and proved its polynomial global and superlinear local convergences. By adopting AHO direction in the corrector step, we achieve fast centering effect without repeated corrector steps. From the perspective, the algorithm is a constraint-reduced version of Potra et al. [12]. In addition, the new algorithm has a benefit that we can avoid solving Lyapunov equations of *inactive* blocks for computing AHO direction.

Kojima et al. [6] developed IPM using AHO direction for both predictor and corrector steps, and proved its quadratic convergence. In Sect. 2, we extended the idea of constraint-reduced HKM direction by Pakr and O'Leary [26] to AHO direction. We can consider applying the constraint-reduced AHO direction to the algorithm of Kojima et al. [6]. It will be another breakthrough if we can establish a new constraint-reduced IPM with a quadratic convergence by using the constraint-reduced AHO direction in both predictor and corrector steps.

References

[1] Helmberg, C., Rendl, F., Vanderbei, R.J., Wolkowicz, H.: An interior-point method for semidefinite programming. SIAM J. Opt. **6**(2), 342–361 (1996)

[2] Alizadeh, F., Haeberly, J.P.A., Overton, M.L.: Complementarity and nondegeneracy in semidefinite programming. Math. Prog. **77**(1), 111–128 (1997)

[3] Alizadeh, F., Haeberly, J.P.A., Overton, M.L.: Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results. SIAM J. Opt. **8**(3), 746–768 (1998)

[4] Ji, J., Potra, F.A., Sheng, R.: On the local convergence of a predictor-corrector method for semidefinite programming. SIAM J. Opt. **10**(1), 195–210 (1999)

[5] Kojima, M., Shida, M., Shindoh, S.: Local convergence of predictor-corrector infeasible-interior-point algorithms for SDPs and SDLCPs. Math. Prog. **80**(2), 129–160 (1998)

[6] Kojima, M., Shida, M., Shindoh, S.: A predictor-corrector interior-point algorithm for the semidefinite linear complementarity problem using the Alizadeh–Haeberly–Overton search direction. SIAM J. Opt. **9**(2), 444–465 (1999)

[7] Kojima, M., Shindoh, S., Hara, S.: Interior-point methods for the monotone semidefinite linear complementarity problem in symmetric matrices. SIAM J. Opt. **7**(1), 86–125 (1997)

[8] Monteiro, R.D.C.: Primal-dual path-following algorithms for semidefinite programming. SIAM J. Opt. **7**(3), 663–678 (1997)

[9] Monteiro, R.D.C.: Polynomial convergence of primal-dual algorithms for semidefinite programming based on Monteiro and Zhang family of directions. SIAM J. Opt. **8**(3), 797–812 (1998)

[10] Monteiro, R.D.C., Zhang, Y.: A unified analysis for a class of long-step primal-dual path-following interior point algorithms for semidefinite programming. Math. Prog. **81**(3), 281–299 (1998)

[11] Nesterov, Y.E., Todd, M.J.: Primal-dual interior-point methods for self-scaled cones. SIAM J. Opt. **8**(2), 324–364 (1998)

[12] Potra, F.A., Sheng, R.: Superlinear convergence of a predictor-corrector method for semidefinite programming without shrinking central path neighborhood. Tech. Rep. 91, Reports on Computational Mathematics, Department of Mathematics, Univ. of Iowa (1996)

[13] Potra, F.A., Sheng, R.: Superlinear convergence of interior-point algorithms for semidefinite programming. J. Opt. Theory Appl. **99**(1), 103–119 (1998)

[14] Potra, F.A., Sheng, R.: A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. SIAM J. Opt. **8**(4), 1007–1028 (1998)

[15] Zhang, Y.: On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. SIAM J. Opt. **8**(2), 365–386 (1998)

[16] Dantzig, G.B., Ye, Y.: A build-up interior-point method for linear programming: Affine scaling form. Tech. rep., Stanford Univ. (1991)

[17] Hertog, D.D., Roos, C., Terlaky, T.: Adding and deleting constraints in the path-following method for LP. In: Advances in Optimization and Approximation, vol. 1, pp. 166–185. Springer, New York (1994)

[18] Kaliski, J.A., Ye, Y.: A decomposition variant of the potential reduction algorithm for linear programming. Manage. Sci. **39**, 757–776 (1993)

[19] Tits, A.L., Absil, P.A., Woessner, W.P.: Constraint reduction for linear programs with many inequality constraints. SIAM J. Opt. **17**(1), 119–146 (2006)

[20] Tone, K.: An active-set strategy in an interior point method for linear programming. Math. Prog. **59**(3), 345–360 (1993)

[21] Winternitz, L.B., Nicholls, S.O., Tits, A.L., O'Leary, D.P.: A constraint-reduced variant of Mehrotra's predictor-corrector algorithm. Comput. Optim. Appl. **51**(3), 1001–1036 (2012)

[22] Ye, Y.: An $O(n^3 L)$ potential reduction algorithm for linear programming. Math. Prog. **50**(2), 239–258 (1991)

[23] Jung, J.H., O'Leary, D.P., Tits, A.L.: Adaptive constraint reduction for training support vector machines. Elec. Trans. Numer. Anal. **31**, 156–177 (2008)

[24] Williams, J.A.: The use of preconditioning for training support vector machines. Master's thesis, Applied Mathematics and Scientific Computing Program, Univ. of Maryland, College Park, MD (2008)

[25] Jung, J.H., O'Leary, D.P., Tits, A.L.: Adaptive constraint reduction for convex quadratic programming. Comput. Optim. Appl. **51**(1), 125–157 (2012)

[26] Park, S., O'Leary, D.P.: A polynomial time constraint-reduced algorithm for semidefinite optimization problems. J. Opt. Theory Appl. (2015)

[27] Park, S., O'Leary, D.P.: A polynomial time constraint reduced algorithm for semidefinite optimization problems, with convergence proofs. Tech. rep., Univ. of Maryland, College Park, MD (2015). URL http://www.optimization-online.org/DB_HTML/2013/08/4011.html

[28] Park, S.: A constraint-reduced algorithm for semidefinite optimization problem with superlinear convergence. Tech. rep. (2015). URL http://www.optimization-online.org/DB_HTML/2015/05/4919.html