Convergence analysis of an Inexact Infeasible Interior Point method for Semidefinite Programming

CONVERGENCE ANALYSIS OF AN INEXACT INFEASIBLE INTERIOR POINT METHOD FOR SEMIDEFINITE PROGRAMMING*

STEFANIA BELLAVIA† AND SANDRA PIERACCINI‡

Abstract. In this paper we present an extension to SDP of the well known infeasible Interior Point method for linear programming of Kojima, Megiddo and Mizuno (*A primal-dual infeasible-interior-point algorithm for Linear Programming*, Math. Progr., 1993). The extension developed here allows the use of inexact search directions; i.e., the linear systems defining the search directions can be solved with an accuracy that increases as the solution is approached. A convergence analysis is carried out and the global convergence of the method is proved.

Key words. Inexact interior point; Semidefinite programming; Global convergence.

This is the authors' post-print version of an article published on *Computational Optimization and Applications*, Volume 29, Number 3, pp. 289-313, DOI:10.1023/B:COAP.0000044183.43473.34.[1]

1. Introduction. In this paper we consider the Semidefinite Program (SDP)

$$\begin{aligned} \min_{X \in \mathcal{S}^n} \ & C \bullet X \\ \text{s.t. } & A_i \bullet X = b_i, \ i = 1, ..., m, \\ & X \succeq 0, \end{aligned} \tag{1.1}$$

where $\mathcal{S}^n$ denotes the set of $n \times n$ symmetric matrices; $C \in \mathcal{S}^n$, $A_i \in \mathcal{S}^n$ and $b_i \in \mathbb{R}$, $i = 1, \dots, m$; $C \bullet X = \mathrm{tr}(CX)$ and $X \succeq 0$ means that X is positive semidefinite. We assume that $A_i, i = 1, \dots, m$, are linearly independent.

Under certain assumptions, X^* is solution of (1.1) if and only if there exist a vector $y^* \in \mathbb{R}^m$ and a $n \times n$ symmetric matrix S^* such that (X^*, y^*, S^*) is a solution of the following constrained nonlinear system:

$$\sum_{i=1}^{m} y_i A_i + S - C = 0, \tag{1.2}$$

$$A_i \bullet X - b_i = 0, \quad i = 1, ..., m, \tag{1.3}$$

$$XS = 0, \tag{1.4}$$

$$S \succeq 0, \qquad X \succeq 0, \tag{1.5}$$

where the $n \times n$ symmetric matrix S and the vector $y \in \mathbb{R}^m$ are the dual variables [2, 14]. The zero on the right hand side of the equations in (1.2)-(1.4) means that every entry on the left-hand side is zero.

*Work supported by MURST, Rome, Italy, through "Cofinanziamenti Programmi di Ricerca Scientifica di Interesse Nazionale".

†Dipartimento di Energetica "S. Stecco", University of Florence, via C. Lombroso 6/17, 50134 Florence, Italy, e-mail: `bellavia@ciro.de.unifi.it`.

‡Dipartimento di Matematica, Politecnico di Torino, c.so Duca degli Abruzzi 24, 10129 Torino, Italy, e-mail: `sandra.pieraccini@polito.it`.

SDP arises in a wide variety of areas as optimal control, structural optimization, pattern recognition, eigenvalue optimization. See the survey paper [26] for some instances. Several authors have discussed generalizations of Interior Point methods for linear programming to the context of Semidefinite Programming. An important paper in this direction is due to Zhang [28], where some key results are given which provide a convenient tool for studying extension from LP to SDP. A rather comprehensive list of references to papers dealing with IP methods for SDP can be found in [24]. Regarding the analysis of infeasible Interior Point methods for SDP, important results were obtained in [12, 14, 21, 22, 28]. More precisely Zhang [28] extended to SDP the long step infeasible Interior Point method for horizontal linear complementarity problems given in [27]; in [14] Kojima, Shindoh and Hara presented an infeasible potential reduction method for SDP; Kojima, Shida and Shindoh [12] and Potra and Sheng [21] independently analyzed generalizations to SDP of the Mizuno-Todd-Ye predictor corrector method [17] for infeasible starting points. In these latter papers also the local superlinear convergence is studied and in [22] the convergence results obtained in [21] are improved.

In this paper, we discuss an extension to SDP of the infeasible Interior Point method for LP given by Kojima, Megiddo, Mizuno [11]. This latter method allows the use of arbitrary starting points and appeared to be an efficient approach in solving LP problems. In the sequel we will refer to this method as the KMM method. The extension we develop here is obtained employing the HRVW/KSH/M search direction [8, 14, 18] and the resulting method, due to its simple structure, can be easily modified in order to use inexact search directions. In fact, at each iteration of an Interior Point method for SDP a linear system must be solved. More precisely, the computation of the search direction can be reduced to the solution of a linear system (the Schur complement equation) $M\Delta y = d$, where $d \in \mathbb{R}^m$ and M is a symmetric and positive definite $m \times m$ matrix. Computing M and solving the system is an expensive task; further, the matrix M is generally dense even if the data matrices of the original problem are sparse. In this context, when m is large, iterative linear solvers, and among them Krylov subspace methods, may be useful for several reasons. In fact, Krylov methods only require the action of the matrix M onto a vector v and not the matrix itself. Therefore, the whole matrix M does not need to be stored and the sparsity of the original problem can be exploited. Moreover, if we are far from a solution it may be unnecessary to compute the search directions with a high accuracy. As a result it can be convenient to use iterative methods for solving the linear systems with an accuracy which increases as far as we get closer to the solution.

Interior Point methods that at each iteration compute an approximate solution of the linear system are called Inexact (or Truncated) Interior Point methods. For these methods a crucial question is how approximately the linear systems can be solved without loosing the good convergence properties of the exact counterpart. In the context of LP problems and complementarity problems many inexact Interior Point methods have been proposed [3, 4, 5, 6, 7, 15, 16, 19, 20]. In particular, in [3, 15] inexact variants of the KMM algorithm have been studied. These two methods mainly differ in the accuracy requirement in the solution of the linear systems; the criterion used in [15] has the drawback that if an iterate happens to be primal-feasible then all the forthcoming iterates remain feasible and the linear systems must be solved exactly, loosing the advantage of inexact computation.

The issue of inexact computation in Interior Point methods for SDP was investigated in [13], [25] and [29]. In [13] an inexact predictor corrector infeasible Interior

Point method is devised and some complexity results for this method are proved. In [25] numerical results are shown for the method proposed in [13]: the linear systems are solved with the conjugate residual method and a decomposition strategy is designed in order to overcome the difficulty to build a good preconditioner for the matrix M. It should be noted that the method in [13] requires at each iteration the solution of a further $m \times m$ linear system whose coefficient matrix is constant throughout the iterations. In [25] is pointed out that this matrix is sparse for many large scale problems and can be factored at reasonable cost, but for some classes of problems (such as those arising from control theory) the matrix is dense and the solution of the additional linear systems can be very expensive to compute and it should be performed by an iterative solver in order to avoid memory problems. In this latter case the advantage of having a constant coefficient matrix is obviously lost. In [29] the infeasible Interior Point method for SDP proposed by Zhang in [28] is modified in order to allow the use of inexact search directions.

Here we present an inexact Interior Point method for SDP that does allow the linear systems to be solved to a low accuracy when the current iterate is far from the solution. More precisely, the accuracy requirement in the solution of the linear systems is related only to the quantity $X \bullet S$. Therefore, even if the iterates happen to be primal feasible, the linear systems can be solved inexactly. Under the assumption that the iterates are bounded, we prove that the method is globally convergent. We note that our convergence results are consistent with those obtained in [15] and [3] for inexact variants of the KMM algorithm for LP, under similar assumptions.

Regarding the exact counterpart of our method, following the general framework given in [28], we prove that it is polynomially convergent without imposing any assumptions on the boundedness of the iterates.

We would like to remark that also Zhou and Toh in [29] use an accuracy requirement in the solution of the linear systems that is related only to the complementarity gap. Anyway, they use a requirement which is tighter than our, in order to preserve the same complexity of the exact method given in [28]. Finally, it should be noted that our method does not require the computation of a further linear system as in [13].

The paper is organized as follows. In section 2 the general class of methods we are involved with is reviewed. In section 3 the inexact search directions are introduced and our Inexact Interior Point method is given. The convergence analysis of the method is carried out in section 4 and in section 5 some practical issues are discussed. Finally, since in this paper we will frequently use Kronecker products, we report in the Appendix some useful results related to Kronecker products as long as some technical results on eigenvalues.

Notation. We use $\mathcal{S}^n_+$ ($\mathcal{S}^n_{++}$) to denote the set of the $n \times n$ positive semidefinite (definite) symmetric matrices. Sometimes the abbreviation s.p.d. will be used as shorthand for "symmetric positive definite".

Given $A \in \mathbb{R}^{p \times p}$, $\lambda_i(A)$ are the eigenvalues of A and $\Lambda(A)$ denotes the spectrum of A. Given $A \in \mathbb{R}^{p \times q}$, $\sigma_i(A)$ are the singular values of A. Throughout the rest of the paper we assume that the eigenvalues of a symmetric matrix A and the singular values of any matrix A are ordered as follows:

$$\lambda_1(A) \geq \lambda_2(A) \geq ... \geq \lambda_p(A), \qquad \sigma_1(A) \geq \sigma_2(A) \geq ... \geq \sigma_{\min\{p,q\}}(A).$$

The notations $X(\alpha)$, $y(\alpha)$, $S(\alpha)$, $\mu(\alpha)$ are used for $X(\alpha) = X + \alpha\Delta X$, $y(\alpha) = y + \alpha\Delta y$, $S(\alpha) = S + \alpha\Delta S$, $\mu(\alpha) = X(\alpha) \bullet S(\alpha)/n$.

The euclidean norm and the Frobenious norm are denoted by $\|\cdot\|$ and $\|\cdot\|_F$, respectively.

For any $m \times n$ matrix A, vec$A$ denotes the mn-vector obtained by stacking the columns of A one by one from the first to the last, while for a given mn-vector v, mat$v$ denotes the $m \times n$ matrix such that vec(mat$v$) $= v$.

2. Preliminary discussion. In this section, first we describe the general infeasible Interior Point path-following framework for SDP. Second, we specify our choices for the search direction and the centrality conditions.

The standard Interior Point method for LP is obtained by applying a perturbed Newton method to the nonlinear system defined by the optimality conditions. However, unlike the corresponding equations for linear programming, the nonlinear system (1.2)-(1.4) is not "quite" square. In fact, the function defined by the left-hand side of (1.2)-(1.4) maps $\mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ into $\mathcal{S}^n \times \mathbb{R}^m \times \mathbb{R}^{n\times n}$ since the product of two symmetric matrices is not necessarily symmetric. Consequently, the domain and the range of this function are not the same spaces, and Newton-type methods cannot be straightforward applied since they do not preserve the symmetry of the matrix X. To apply Newton-type algorithms it is previously necessary to symmetrize (1.4) so that the left-hand side of the resulting equivalent nonlinear system gives a function that maps $\mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ into itself [23]. Zhang [28] introduced a general symmetrization scheme based on the operator $\mathcal{H}_P : \mathbb{R}^{n\times n} \mapsto \mathcal{S}^n$ defined as

$$\mathcal{H}_P(A) = \frac{1}{2}(PAP^{-1} + (PAP^{-1})^T) \quad \forall A \in \mathbb{R}^{n\times n}, \tag{2.1}$$

where $P \in \mathbb{R}^{n\times n}$ is some nonsingular matrix. In [28] the author has shown that for any given nonsingular matrix P the system (1.2)-(1.5) is equivalent to the following nonlinear system:

$$\begin{aligned}
&\sum_{i=1}^m y_i A_i + S - C = 0\\
&A_i \bullet X - b_i = 0, \quad i = 1, \dots, m\\
&\mathcal{H}_P(XS) = 0,\\
&S \succeq 0, \qquad X \succeq 0,
\end{aligned}$$

to which Newton-type methods can be applied.

Taking into account the above considerations, an infeasible Interior Point path-following method, starting form an initial guess $(X_0, y_0, S_0) \in \mathcal{S}^n_{++} \times \mathbb{R}^m \times \mathcal{S}^n_{++}$ (a common choice is $X_0 = S_0 = \rho I$, $\rho > 0$ and $y_0 = 0$) generates a sequence of iterates $\{(X_k, y_k, S_k)\}$ belonging to $\mathcal{S}^n_{++} \times \mathbb{R}^m \times \mathcal{S}^n_{++}$. The sequence $\{(X_k, y_k, S_k)\}$ is defined by

$$\left\{ \begin{array}{rcl} X_{k+1} &=& X_k + \alpha_k \Delta X\\ y_{k+1} &=& y_k + \alpha_k \Delta y\\ S_{k+1} &=& S_k + \alpha_k \Delta S \end{array}\right. \tag{2.2}$$

where $(\Delta X, \Delta y, \Delta S) \in \mathcal{S}^n \times \mathbb{R}^m \times \mathcal{S}^n$ is the solution of the following linear system:

$$\sum_{i=1}^m \Delta y_i A_i + \Delta S = -\sum_{i=1}^m (y_i)_k A_i - S_k + C \tag{2.3}$$

$$A_i \bullet \Delta X = -A_i \bullet X_k + b_i, \quad i = 1, \dots, m \tag{2.4}$$

$$\mathcal{H}_P(\Delta X S_k + X_k \Delta S) = \sigma_k \mu_k I - \mathcal{H}_P(X_k S_k) \tag{2.5}$$

and $\alpha_k \in (0,1]$ is chosen in order to satisfy suitable centrality conditions and a decrease condition on the merit function $X \bullet S$.

In (2.5) $\sigma_k \in (0,1)$ is the centering parameter and

$$\mu_k = (X_k \bullet S_k)/n.$$

Note that (2.3) implies that ΔS is symmetric while ΔX does enjoy symmetry thanks to the introduction of the symmetrization operator. This way the symmetry of X_k and S_k is maintained.

Different choices of the matrix P lead to different search directions with different properties and drawbacks [23, 24]. Throughout the paper, we use the matrix $P = S^{1/2}$ proposed in [8, 14, 18], that gives rise to the so called HRVW/KSH/M search direction. ¿From now on, we drop the subscript P from $\mathcal{H}_{\mathcal{P}}$ when it is clear from the context.

The previous linear system can be written in a more compact form in the following way:

$$\mathcal{A}^T \Delta y + \mathrm{vec}\Delta S = \mathrm{vec}R_d^{(k)} \tag{2.6}$$

$$\mathcal{A}\mathrm{vec}\Delta X = r_p^{(k)} \tag{2.7}$$

$$E_k \mathrm{vec}\Delta X + F_k \mathrm{vec}\Delta S = \mathrm{vec}R_c^{(k)}, \tag{2.8}$$

where $\mathcal{A}^T := (\mathrm{vec}A_1, \mathrm{vec}A_2, ..., \mathrm{vec}A_m) \in \mathbb{R}^{n^2 \times m}$ and

$$\begin{aligned}
\mathrm{vec}R_d^{(k)} &= -(\mathcal{A}^T y_k + \mathrm{vec}S_k - \mathrm{vec}C) \\
r_p^{(k)} &= -(\mathcal{A}\mathrm{vec}X_k - b) \\
R_c^{(k)} &= 2(\sigma_k \mu_k S_k - S_k X_k S_k) \\
E_k &= 2S_k \otimes S_k \\
F_k &= S_k X_k \otimes I + I \otimes S_k X_k.
\end{aligned}$$

For sake of simplicity in the following the iteration index k of $r_p^{(k)}$, $R_d^{(k)}$ and $R_c^{(k)}$ is omitted whenever it can be inferred from the context. ¿From (2.6) we can easily get $\mathrm{vec}\Delta S$; replacing it in (2.8) we obtain

$$\begin{aligned}
\mathcal{A}\mathrm{vec}\Delta X &= r_p \\
E_k \mathrm{vec}\Delta X + F_k(\mathrm{vec}R_d - \mathcal{A}^T \Delta y) &= \mathrm{vec}R_c.
\end{aligned}$$

Now, getting $\mathrm{vec}\Delta X$ from the second equation and replacing it in the first one, we have

$$\mathcal{A}E_k^{-1} F_k \mathcal{A}^T \Delta y = r_p - \mathcal{A}E_k^{-1}\mathrm{vec}R_c + \mathcal{A}E_k^{-1} F_k \mathrm{vec}R_d. \tag{2.9}$$

For convenience, let $M_k = \mathcal{A}E_k^{-1} F_k \mathcal{A}^T$.

Therefore, we can solve the reduced system (2.9) of dimension m obtaining Δy; then ΔX and ΔS are computed via

$$\mathrm{vec}\Delta S = \mathrm{vec}C - \mathrm{vec}S_k - \mathcal{A}^T(y_k + \Delta y) \tag{2.10}$$

and

$$\mathrm{vec}\Delta X = E_k^{-1}\left(\mathrm{vec}R_c - F_k \mathrm{vec}\Delta S\right). \tag{2.11}$$

Concerning the matrix M_k, in [28, Proposition 2.1] it is proved that as long as X_k and S_k are s.p.d and $\mathcal{A}$ is full rank, M_k is s.p.d. as well. As a consequence, for solving the linear system (2.9) we can rely on methods as Cholesky factorization for a direct approach or Conjugate Gradient method for an iterative approach.

We consider the following centrality conditions:

$$\gamma_1 \mu_k \le \Lambda(\mathcal{H}(X_k S_k)) \tag{2.12}$$

$$X_k \bullet S_k \ge \gamma_2 \|r_p\| \tag{2.13}$$

$$X_k \bullet S_k \ge \gamma_3 \|\text{vec} R_d\|, \tag{2.14}$$

where, given $\hat{\gamma} \in (0,1)$, the constants γ_1, γ_2 and γ_3 are defined by

$$\gamma_1 = \min\left\{\hat{\gamma}, \frac{\lambda_n(\mathcal{H}(X_0 S_0))}{\mu_0}\right\}, \quad \gamma_2 = \min\left\{\hat{\gamma}, \frac{X_0 \bullet S_0}{\|r_p^{(0)}\|}\right\}, \quad \gamma_3 = \frac{X_0 \bullet S_0}{\|\text{vec} R_d^{(0)}\|}. \tag{2.15}$$

Then, the step α_k in (2.2) is chosen in such a way that $(X_{k+1}, y_{k+1}, S_{k+1})$ satisfies (2.12)-(2.14) and the following decrease condition on the merit function $X \bullet S$:

$$X_{k+1} \bullet S_{k+1} \le (1 - \alpha_k(1-\beta)) X_k \bullet S_k, \tag{2.16}$$

where $\beta \in (0,1)$. This way, the quantity $X_k \bullet S_k$ is driven to zero and due to conditions (2.13) and (2.14) also $\|r_p\|$ and $\|\text{vec} R_d\|$ are pushed to zero.

Conditions (2.12)-(2.14) and the decrease condition (2.16) are the generalization to SDP of the centrality conditions and the decrease condition used in [11]. Therefore, an Interior Point method for SDP that uses (2.12)-(2.14) and (2.16), turns out to be an extension of the KMM method to SDP.

3. Inexact search directions. At each iteration of an Interior Point method for SDP the linear system (2.9) must be solved. As pointed out in the introduction, when m is large it may be convenient to use an iterative solver to compute the solution of the linear system inexactly.

Here we propose an Inexact Interior Point method for SDP that uses conditions (2.12)-(2.14), (2.16) and the matrix $P = S^{1/2}$ in the symmetrization operator defined by (2.1). In our approach we consider inexact search directions $(\Delta X, \Delta y, \Delta S)$, where Δy satisfies

$$\mathcal{A} E_k^{-1} F_k \mathcal{A}^T \Delta y = r_p - \mathcal{A} E_k^{-1} \text{vec} R_c + \mathcal{A} E_k^{-1} F_k \text{vec} R_d + r_k \tag{3.1}$$

with

$$\|r_k\| \le \eta_k X_k \bullet S_k, \qquad \eta_k \in (0,1), \tag{3.2}$$

and ΔX and ΔS are then computed via (2.10) and (2.11). As it is usual in the context of inexact methods, we will refer to the vector r_k as the residual vector and to the parameter η_k as the forcing term. The forcing term is used to control the level of accuracy in solving the linear systems. We would like to remark that the criterion (3.2) allows us to solve the linear systems to a very low accuracy when the iterates are far from the solution, i.e. when $X_k \bullet S_k$ is large.

Note that solving (3.1) instead of (2.9) corresponds to replace equation (2.7) in the full dimensional system (2.6)-(2.8) with

$$\mathcal{A} \text{vec} \Delta X = r_p + r_k. \tag{3.3}$$

The method outlined above is formally defined as follows.

IIP-SDP Method (Inexact Interior Point Method for SDP)

0. Given $\sigma_0 \in (0,1)$, $\varepsilon > 0$, $\hat{\gamma} \in (0,1)$, $\beta \in (0,1)$, $\eta_0 \in [0,1)$, $X_0 \in \mathcal{S}^n_{++}$, $S_0 \in \mathcal{S}^n_{++}$, $y_0 \in \mathbb{R}^m$.
1. Set $k = 0$ and compute γ_1, γ_2, γ_3 from (2.15).
2. Repeat until $X_k \bullet S_k \le \varepsilon$:
   - 2.1. Solve system (3.1) with $\|r_k\| \le \eta_k X_k \bullet S_k$, for Δy and compute $\Delta X, \Delta S$ via (2.10) and (2.11).
   - 2.2. Compute the new iterate $(X_{k+1}, y_{k+1}, S_{k+1})$ by (2.2) using a step α_k such that $(X_{k+1}, y_{k+1}, S_{k+1})$ satisfies conditions (2.12)-(2.14) and (2.16).
   - 2.3. Choose σ_{k+1} and η_{k+1}.
   - 2.4. Set $k = k + 1$.
3. Stop

Clearly, if $\eta_k = 0$ is taken at each iteration the previous method reduces to an exact infeasible IP method. In other words it turns out to be an extension of the KMM method to SDP where inexact computation is not used. In the following we will refer to this variant as the exact counterpart of the IIP-SDP Method and we will call it the EIP-SDP Method (Exact Interior Point Method for SDP).

We remark that the KMM algorithm uses a slightly different approach to compute the step α_k. More precisely different primal and dual steplenght are allowed and the decrease condition (2.16) is tested also using a scalar $\hat{\beta}$ greater than β in order to try to take larger steps. Anyway, it is easy to see that the convergence results proved in the next section remain valid also for a variant of the IIP-SDP Method that employs the above mentioned features of the KMM algorithm.

Finally, it may be interesting to compare our accuracy requirement (3.2) with the one used in the method proposed by Zhou and Toh in [29]. In their approach, the residual norm must decrease at least $\mathrm{O}(\sqrt{n}\log n\ )$ time faster than $X_k \bullet S_k$. Recall that the forcing term η_k determines how fast the residual norm need to be decreased with respect to $X_k \bullet S_k$. Hence, the forcing terms used in [29] depend on n and this may lead to small forcing terms when n is large. That is, the larger is the dimension, the more accurately the linear systems are solved. On the contrary, our approach does not suffer of this drawback since the forcing terms are not related to the dimension n. On the other hand, the algorithm given in [29] achieves $O(n^2 \log(1/\epsilon))$ complexity.

4. Convergence. In this section we analyze the convergence properties of the IIP-SDP Method and of its exact counterpart.

First, we show that if the sequences $\{\sigma_k\}$ and $\{\eta_k\}$ are chosen in such a way that $\beta - \sigma_k > 0$ and $\sigma_k - \gamma_2\eta_k > 0$, then the sequence of the iterates is well defined, i.e. at each iteration it is possible to satisfy conditions (2.12)-(2.14) and (2.16), and the generated matrices X_k and S_k are s.p.d. for all $k > 0$.

Second, we analyze the convergence behavior of the EIP-SDP Method. We prove that, if the standard starting point $X_0 = S_0 = \rho I$, $\rho > 0$ and $y_0 = 0$ is used and the sequence $\{\sigma_k\}$ is chosen in such a way that $\beta - \sigma_k > \theta_1$, for $\theta_1 > 0$, then the EIP-SDP Method takes no more than $\mathrm{O}(n^{2.5} \ln \frac{1}{\varepsilon})$ iterations to reach $X_k \bullet S_k \le \varepsilon$.

Finally, we focus on the inexact method. Under the following assumption:

A1. The sequence $\{(X_k, S_k)\}$ is bounded, i.e. there exist a constant c_1 such that $\|X_k\|, \|S_k\| \le c_1$, for every $k > 0$, we prove that the sequence $\{X_k \bullet S_k\}$ generated by the IIP-SDP Method with $\varepsilon = 0$, converges to zero, provided that the constant β and the sequences $\{\sigma_k\}$ and $\{\eta_k\}$ are chosen in such a way that $\beta - \sigma_k > \theta_1$ and $\sigma_k - \gamma_2\eta_k > \theta_2$ for some $\theta_1, \theta_2 > 0$.

We remark that the convergence of the complementarity gap to zero implies that all the accumulation points (X^*, y^*, S^*) of $\{(X_k, y_k, S_k)\}$ solve the problem (1.2)-(1.5). This latter observation is due to the fact that $X \bullet S = 0$ if and only if $XS = 0$ (see Lemma 2.9 of [1]) and to conditions (2.13)-(2.14).

Before starting to analyze the convergence of the method, we prove the following technical Lemma.

Lemma 4.1 *Let P be any nonsingular matrix and $A, B \in \mathcal{S}^n$. Then*

$$|A \bullet B| \le n\|\mathcal{H}_P(AB)\|_F.$$

Proof. First note that given any matrix $C \in \mathbb{R}^{n\times n}$, we have

$$\mathrm{tr}(\mathcal{H}_P(C)) = \frac{1}{2}\mathrm{tr}\big(PCP^{-1}\big) + \frac{1}{2}\mathrm{tr}\big([PCP^{-1}]^T\big) = \mathrm{tr}\big(PCP^{-1}\big) = \mathrm{tr}(C)\,. \tag{4.1}$$

Then, using (4.1) we get

$$\begin{aligned}
|A \bullet B| = |\mathrm{tr}(AB)| = |\mathrm{tr}(\mathcal{H}_P(AB))| &= \left|\sum_{i=1}^{n} \lambda_i(\mathcal{H}_P(AB))\right| \\
&\le \sum_{i=1}^{n} |\lambda_i(\mathcal{H}_P(AB))| \le n \max_{1\le i\le n} |\lambda_i(\mathcal{H}_P(AB))| \\
&= n\|\mathcal{H}_P(AB)\| \le n\|\mathcal{H}_P(AB)\|_F
\end{aligned}$$

and the thesis follows. □

The next result show that if (X_k, y_k, S_k) satisfies (2.12)-(2.14) and (2.16) then there exists $\hat\alpha_k \in (0, 1]$ such that, for all $\alpha \in [0, \hat\alpha_k]$, $(X_k(\alpha), y_k(\alpha), S_k(\alpha))$ satisfies

$$\gamma_1\mu_k(\alpha) \le \Lambda(\mathcal{H}_{S_k^{1/2}}(X_k(\alpha)S_k(\alpha))) \tag{4.2}$$
$$X_k(\alpha) \bullet S_k(\alpha) \ge \gamma_2\|r_p(\alpha)\| \tag{4.3}$$
$$X_k(\alpha) \bullet S_k(\alpha) \ge \gamma_3\|\mathrm{vec}R_d(\alpha)\|, \tag{4.4}$$
$$X_k(\alpha) \bullet S_k(\alpha) \le (1 - \alpha(1 - \beta))X_k \bullet S_k \tag{4.5}$$

where

$$r_p(\alpha) = b - \mathcal{A}\mathrm{vec}X(\alpha) \quad \text{and} \quad \mathrm{vec}R_d(\alpha) = \mathrm{vec}C - \mathrm{vec}S(\alpha) - \mathcal{A}^T y(\alpha).$$

Proposition 4.1 *Assume that the triplet (X_k, y_k, S_k) is such that conditions* (2.12)-(2.14) *and* (2.16) *are satisfied and let $(\Delta X, \Delta y, \Delta S)$ be a solution of* (3.3)*,* (2.6) *and* (2.8)*. If*

$$\|r_k\| \le \eta_k X_k \bullet S_k, \qquad \sigma_k - \gamma_2\eta_k > 0, \qquad \beta > \sigma_k, \tag{4.6}$$

then there exists $\hat\alpha_k \in (0, 1]$ such that for all $\alpha \in [0, \hat\alpha_k]$ the triplet $(X_k(\alpha), y_k(\alpha), S_k(\alpha))$ satisfies conditions (4.2)-(4.5)*.*

Proof. For sake of simplicity we drop the matrix $S_k^{1/2}$ from $\mathcal{H}_{S_k^{1/2}}$ and we omit the iteration index k in the proof. First, let us show that there exists $\hat{\alpha}_1 \in (0,1]$ such that

$$\lambda_n(\mathcal{H}(X(\alpha)S(\alpha))) - \gamma_1\mu(\alpha) \geq 0 \qquad \forall \alpha \in [0, \hat{\alpha}_1].$$

Note that equation (2.8) in the linear system is equivalent to

$$\mathcal{H}(XS + \Delta XS + X\Delta S) = \sigma\mu I, \tag{4.7}$$

from which, using (4.1), it follows that

$$\mathrm{tr}(\Delta XS + X\Delta S) = \mathrm{tr}(\sigma\mu I - XS)\,. \tag{4.8}$$

Hence we have

$$\begin{aligned} X(\alpha) \bullet S(\alpha) &= \mathrm{tr}((X + \alpha\Delta X)(S + \alpha\Delta S)) \\ &= \mathrm{tr}\big(XS + \alpha\Delta XS + \alpha X\Delta S + \alpha^2\Delta X\Delta S\big) \\ &= \mathrm{tr}(XS) + \alpha\mathrm{tr}(\Delta XS + X\Delta S) + \alpha^2\mathrm{tr}(\Delta X\Delta S) \\ &= \mathrm{tr}(XS) + \alpha\mathrm{tr}(\sigma\mu I - XS) + \alpha^2\mathrm{tr}(\Delta X\Delta S) \\ &= X \bullet S + \alpha(\sigma - 1)X \bullet S + \alpha^2\mathrm{tr}(\Delta X\Delta S) \\ &= (1 - \alpha + \alpha\sigma)X \bullet S + \alpha^2\Delta X \bullet \Delta S. \end{aligned} \tag{4.9}$$

Further, taking into account (4.7), it can be easily seen that

$$\mathcal{H}(X(\alpha)S(\alpha)) = (1-\alpha)\mathcal{H}(XS) + \alpha\sigma\mu I + \alpha^2\mathcal{H}(\Delta X\Delta S). \tag{4.10}$$

Hence, from (4.9), relation (6.1) in the Appendix and the fact that $\lambda_n(\mathcal{H}(XS)) \geq \gamma_1\mu$, we derive

$$\begin{aligned} \lambda_n(\mathcal{H}(X(\alpha)S(\alpha))) - \gamma_1\mu(\alpha) &= \lambda_n\left((1-\alpha)\mathcal{H}(XS) + \alpha\sigma\mu I + \alpha^2\mathcal{H}(\Delta X\Delta S)\right) \\ &\quad -\gamma_1\left((1-\alpha+\alpha\sigma)\mu + \frac{\alpha^2}{n}\Delta X \bullet \Delta S\right) \\ &\geq (1-\alpha)\lambda_n(\mathcal{H}(XS)) + \alpha\sigma\mu - \alpha^2\|\mathcal{H}(\Delta X\Delta S)\|_F \\ &\quad -\gamma_1(1-\alpha)\mu - \gamma_1\alpha\sigma\mu - \gamma_1\frac{\alpha^2}{n}\Delta X \bullet \Delta S \\ &\geq (1-\alpha)\gamma_1\mu + \alpha\sigma\mu - \alpha^2\|\mathcal{H}(\Delta X\Delta S)\|_F \\ &\quad -\gamma_1(1-\alpha)\mu - \gamma_1\alpha\sigma\mu - \gamma_1\frac{\alpha^2}{n}\Delta X \bullet \Delta S \\ &= (1-\gamma_1)\alpha\sigma\mu - \alpha^2(\|\mathcal{H}(\Delta X\Delta S)\|_F + \frac{\gamma_1}{n}\Delta X \bullet \Delta S) \qquad (4.11) \\ &\geq (1-\gamma_1)\alpha\sigma\mu - \alpha^2\left(\|\mathcal{H}(\Delta X\Delta S)\|_F + \frac{\gamma_1}{n}\left|\Delta X \bullet \Delta S\right|\right) \\ &\geq (1-\gamma_1)\alpha\sigma\mu - \alpha^2(1+\gamma_1)\|\mathcal{H}(\Delta X\Delta S)\|_F. \end{aligned}$$

The latter inequality follows from Lemma 4.1. Then, since $\gamma_1 \leq \hat{\gamma} < 1$, taking

$$\hat{\alpha}_1 = \frac{(1-\gamma_1)\sigma\mu}{(1+\gamma_1)\|\mathcal{H}(\Delta X\Delta S)\|_F},$$

we can state that condition (4.2) is satisfied for all $\alpha \in [0, \hat{\alpha}_1]$.

Now let us show that there exist $\hat{\alpha}_2$ such that for all $\alpha \in [0, \hat{\alpha}_2]$ condition (4.3) is satisfied. Taking into account that $\|r\| \leq \eta X \bullet S$ and $X \bullet S \geq \gamma_2 \|r_p\|$ and using (4.9) and (3.3) we get

$$
\begin{aligned}
X(\alpha) \bullet S(\alpha) - \gamma_2 \|r_p(\alpha)\| &= (1 - \alpha + \alpha\sigma) X \bullet S + \alpha^2 \Delta X \bullet \Delta S \\
&\quad - \gamma_2 \|b - \mathcal{A}\text{vec}X - \alpha \mathcal{A}\text{vec}\Delta X\| \\
&= (1 - \alpha + \alpha\sigma) X \bullet S + \alpha^2 \Delta X \bullet \Delta S \\
&\quad - \gamma_2 \|b - \mathcal{A}\text{vec}X - \alpha(b - \mathcal{A}\text{vec}X + r)\| \\
&\geq (1 - \alpha + \alpha\sigma) X \bullet S + \alpha^2 \Delta X \bullet \Delta S \\
&\quad - \gamma_2 (1 - \alpha) \|b - \mathcal{A}\text{vec}X\| - \gamma_2 \alpha \|r\| \\
&\geq \alpha\sigma X \bullet S + \alpha^2 \Delta X \bullet \Delta S - \gamma_2 \alpha \eta X \bullet S \\
&\geq \alpha(\sigma - \gamma_2 \eta) X \bullet S - \alpha^2 \left| \Delta X \bullet \Delta S \right|.
\end{aligned}
$$

Since $\sigma - \gamma_2 \eta > 0$ by hypothesis, it follows that (4.3) is satisfied for every $\alpha \in [0, \hat{\alpha}_2]$, with

$$
\hat{\alpha}_2 = \frac{(\sigma - \gamma_2 \eta) X \bullet S}{|\Delta X \bullet \Delta S|}.
$$

Now, let us consider condition (4.4). We use (2.6) and, once again, (4.9). Then, taking into account that $X \bullet S \geq \gamma_3 \|\text{vec}R_d\|$, we get

$$
\begin{aligned}
X(\alpha) \bullet S(\alpha) - \gamma_3 \|\text{vec}R_d(\alpha)\| &= (1 - \alpha + \alpha\sigma) X \bullet S + \alpha^2 \Delta X \bullet \Delta S \\
&\quad - \gamma_3 \| - \text{vec}S - \alpha \text{vec}\Delta S + \text{vec}C - \mathcal{A}^T y - \alpha \mathcal{A}^T \Delta y\| \\
&= (1 - \alpha + \alpha\sigma) X \bullet S + \alpha^2 \Delta X \bullet \Delta S \\
&\quad - \gamma_3 \| - \text{vec}S - \mathcal{A}^T y + \text{vec}C - \alpha(\text{vec}C - \text{vec}S - \mathcal{A}^T y)\| \\
&\geq \alpha\sigma X \bullet S - \alpha^2 \left| \Delta X \bullet \Delta S \right|.
\end{aligned}
$$

Then, taking

$$
\hat{\alpha}_3 = \frac{\sigma X \bullet S}{|\Delta X \bullet \Delta S|}, \tag{4.12}
$$

condition (4.4) is satisfied for all $\alpha \in [0, \hat{\alpha}_3]$.

Finally, let us turn to condition (4.5). Using again (4.9), we get

$$
\begin{aligned}
(1 - \alpha(1 - \beta)) X \bullet S - X(\alpha) \bullet S(\alpha) &= (1 - \alpha(1 - \beta)) X \bullet S - (1 - \alpha + \alpha\sigma) X \bullet S \\
&\quad - \alpha^2 \Delta X \bullet \Delta S \\
&= \alpha(\beta - \sigma) X \bullet S - \alpha^2 \Delta X \bullet \Delta S \\
&\geq \alpha(\beta - \sigma) X \bullet S - \alpha^2 \left| \Delta X \bullet \Delta S \right|.
\end{aligned}
$$

Then, since $\beta > \sigma$ by hypothesis, we can conclude that condition (4.5) holds for all $\alpha \in [0, \hat{\alpha}_4]$, where $\hat{\alpha}_4$ is given by:

$$
\hat{\alpha}_4 = \frac{(\beta - \sigma) X \bullet S}{|\Delta X \bullet \Delta S|}. \tag{4.13}
$$

Finally, the thesis follows with

$$
\hat{\alpha}_k = \min\left\{1, X \bullet S \min\left\{\frac{1 - \gamma_1}{1 + \gamma_1} \frac{\sigma}{n \|\mathcal{H}(\Delta X \Delta S)\|_F}, \frac{\sigma - \gamma_2 \eta}{|\Delta X \bullet \Delta S|}, \frac{\sigma}{|\Delta X \bullet \Delta S|}, \frac{\beta - \sigma}{|\Delta X \bullet \Delta S|}\right\}\right\}. \quad \blacksquare \tag{4.14}
$$

□

Note that, as pointed out by Zhang in [28], the computation of a steplenght $\tilde{\alpha}_1^k$ such that (4.2) is satisfied for any $\alpha \in [0, \tilde{\alpha}_1^k]$ can be performed without calculating the eigenvalues of $\mathcal{H}_{S_k^{1/2}}(X_k(\alpha)S_k(\alpha))$, due to (4.11).

The following Proposition, which is essentially [28, Lemma 4.7], ensures that the matrices $X_k(\alpha)$ and $S_k(\alpha)$, with $\alpha \in [0, \hat{\alpha}_k]$, are s.p.d.

Proposition 4.2 *Assume that the hypotheses of Proposition 4.1 are satisfied and let $\hat{\alpha}_k$ be defined as in* (4.14)*. Then for all $\alpha \in [0, \hat{\alpha}_k]$ the matrices $X_k(\alpha)$ and $S_k(\alpha)$ are s.p.d., unless $\hat{\alpha}_k = 1$ and $X_k(1) \bullet S_k(1) = 0$ in which case $(X_k(1), y_k(1), S_k(1))$ is a solution to* (1.2)-(1.5).

Proof. Using (4.9) we get:

$$\begin{aligned} &X_k(\alpha) \bullet S_k(\alpha) - (1-\alpha) X_k \bullet S_k \\ &= (1 - \alpha + \alpha\sigma_k) X_k \bullet S_k + \alpha^2 \Delta X \bullet \Delta S - (1-\alpha) X_k \bullet S_k \\ &\geq \alpha\sigma_k X_k \bullet S_k - \alpha^2 |\Delta X \bullet \Delta S|. \end{aligned}$$

Then, the inequality

$$X_k(\alpha) \bullet S_k(\alpha) \geq (1-\alpha) X_k \bullet S_k \tag{4.15}$$

holds for any $\alpha \in [0, \hat{\alpha}_3]$, with $\hat{\alpha}_3$ defined by (4.12); therefore, it holds for any $\alpha \in [0, \hat{\alpha}_k]$. Then we can proceed as in the proof of [28, Lemma 4.7] and the thesis follows. □

¿From the previous results it follows that step 2.2 of Method IIP-SDP is well defined. In fact, we recall from Proposition 4.1 that, taking $\alpha \in (0, \hat{\alpha}_k]$, the point $(X_{k+1}, y_{k+1}, S_{k+1}) = (X_k(\alpha), y_k(\alpha), S_k(\alpha))$ satisfies conditions (4.2)-(4.5); further, Proposition 4.2 guarantees that X_{k+1} and S_{k+1} are s.p.d. Note that conditions (4.3), (4.4) and (4.5) imply that the point $(X_{k+1}, y_{k+1}, S_{k+1})$ also satisfies conditions (2.13), (2.14) and (2.16), respectively. Concerning condition (2.12), note that at the $(k+1)$-th iteration, the scaling matrix used in (2.1) is $S_{k+1}^{1/2}$ and in general $\mathcal{H}_{S_{k+1}^{1/2}}(X_{k+1}S_{k+1})$ is different from $\mathcal{H}_{S_k^{1/2}}(X_{k+1}S_{k+1})$. Anyway, Lemma 4.2 in [28] states that if S_k and S_{k+1} are s.p.d. then the following inequality holds:

$$\lambda_n(\mathcal{H}_{S_k^{1/2}}(X_{k+1}S_{k+1})) \leq \Lambda(\mathcal{H}_{S_{k+1}^{1/2}}(X_{k+1}S_{k+1})).$$

Hence, we can conclude that the point $(X_{k+1}, y_{k+1}, S_{k+1})$ satisfies condition (2.12) as well.

Now, we focus on the convergence behavior of the sequence $\{X_k \bullet S_k\}$. Here we report some observations and results of [28] which are useful in several places in the analysis that follows. First, in [28, Proposition 2.1] it is shown that the matrix

$$E_k^{-1} F_k = \frac{(X_k \otimes S_k^{-1} + S_k^{-1} \otimes X_k)}{2}$$

is s.p.d.

Further, in this section we make use of the matrix $\hat{S}_k = F_k E_k^T$ which originally appeared in [28]. Regarding this matrix the following result holds.

Lemma 4.2 *([28, Proposition 2.3]) The matrix $\hat{S}_k = F_k E_k^T$ is s.p.d. and can be written as*

$$\hat{S}_k = E_k^{1/2} \hat{F}_k E_k^{1/2} \tag{4.16}$$

where $\hat{F}_k$ is s.p.d. and is given by

$$\hat{F}_k = E_k^{-1/2} F_k E_k^{1/2} = S_k^{1/2} X_k S_k^{1/2} \otimes I + I \otimes S_k^{1/2} X_k S_k^{1/2}. \tag{4.17}$$

Remark 4.1 Note that $X_k S_k$, $S_k X_k$, $X_k^{1/2} S_k X_k^{1/2}$, and $S_k^{1/2} X_k S_k^{1/2}$ are similar matrices. Moreover, note also that

$$\mathcal{H}(X_k S_k) = \frac{1}{2}(S_k^{1/2} X_k S_k S_k^{-1/2} + (S_k^{1/2} X_k S_k S_k^{-1/2})^T) = S_k^{1/2} X_k S_k^{1/2}; \tag{4.18}$$

this yields the similarity between $\mathcal{H}(X_k S_k)$ and $X_k S_k$. All this things considered, for sake of brevity, from now on we will denote the eigenvalues of these matrices as λ_i^k, $i = 1, \ldots, n$. Finally, using Lemma 6.1, part (g), it is straightforward noting that the eigenvalues of F_k and $\hat{F}_k$ are given by $\lambda_i^k + \lambda_j^k$, for $i, j = 1, .., n$.

We also make repeated use of the following matrices:

$$D_k = \hat{S}_k^{-1/2} F_k = \hat{S}_k^{1/2} E_k^{-T}, \qquad D_k^{-T} = \hat{S}_k^{-1/2} E_k = \hat{S}_k^{1/2} F_k^{-T}. \tag{4.19}$$

They will play a key role in the development of this analysis. First, note that from the definition of the matrix D_k we have

$$\begin{aligned} D_k^T D_k &= (\hat{S}_k^{-1/2} F_k)^T \hat{S}_k^{1/2} E_k^{-T} \\ &= F_k^T (\hat{S}_k^{-1/2})^T \hat{S}_k^{1/2} E_k^{-T} \\ &= F_k^T E_k^{-T}. \end{aligned}$$

Then, the symmetry of $E_k^{-1} F_k$ yields

$$D_k^T D_k = E_k^{-1} F_k. \tag{4.20}$$

Moreover, from Lemma 3.3 in [28] we get the following inequality:

$$\|\mathcal{H}(\Delta X \Delta S)\|_F \leq \frac{1}{2}\sqrt{\frac{\lambda_1^k}{\lambda_n^k}}(\|D_k^{-T}\mathrm{vec}\Delta X\|^2 + \|D_k \mathrm{vec}\Delta S\|^2). \tag{4.21}$$

Further, we note that

$$|\Delta X \bullet \Delta S| \leq \|D_k^{-T}\mathrm{vec}\Delta X\| \|D_k \mathrm{vec}\Delta S\| \leq \frac{1}{2}(\|D_k^{-T}\mathrm{vec}\Delta X\|^2 + \|D_k \mathrm{vec}\Delta S\|^2). \tag{4.22}$$

Finally, let us introduce the quantity

$$\nu_k = \prod_{j=0}^{k} (1 - \alpha_j).$$

Note that

$$\mathrm{vec} R_d^{(k+1)} = \nu_k \mathrm{vec} R_d^{(0)} \tag{4.23}$$

and that (4.15) yields

$$\mu_{k+1} \geq \nu_k \mu_0. \tag{4.24}$$

4.1. Convergence of the EIP-SDP Method. Let us consider the exact version of our method. In this case, the global convergence can be proved following the lines given by Zhang in [28] under the assumption that the following conditions hold:

$$X_0 - X_* \succeq 0, \quad S_0 - S_* \succeq 0, \quad X_0 = S_0 = \rho I,$$

where (X_*, y_*, S_*) is a solution of (1.2)-(1.5) and

$$\rho \geq \frac{\mathrm{tr}(X_*) + \mathrm{tr}(S_*)}{n}.$$

Note that the relation $r_p^{(k+1)} = \nu_k r_p^{(0)}$ is satisfied, then Lemma 3.5 of [28] can be used in our context. Therefore, the following result holds.

Lemma 4.3 *Let (X_k, y_k, S_k) be the k-th iterate generated by the EIP-SDP Method, (X_0, y_0, S_0) be the initial guess, (X^*, y^*, S^*) be a solution of the problem. Then,*

$$\|D_k^{-T}\mathrm{vec}\Delta X\|^2 + \|D_k \mathrm{vec}\Delta S\|^2 \leq \left(\xi_k + \sqrt{\xi_k^2 + \zeta_k}\right)^2,$$

where

$$\xi_k = \nu_{k-1}(\|D_k^{-T}\mathrm{vec}(X_0 - X^*)\| + \|D_k \mathrm{vec}(S_0 - S^*)\|), \tag{4.25}$$

$$\zeta_k = \|\hat{S}_k^{-1/2} R_c\|_F^2 + 2\nu_{k-1}^2 (X_0 - X^*) \bullet (S_0 - S^*). \tag{4.26}$$

Concerning the boundedness of ξ_k and ζ_k, since (4.24) and (2.12) hold, we can apply Lemma 3.6 and Lemma 3.7 in [28] obtaining

$$\xi_k \leq 3n\sqrt{\frac{\mu_k}{\gamma_1}} \tag{4.27}$$

and

$$\zeta_k \leq (3 - 2\sigma_k + \sigma_k^2/\gamma_1) n \mu_k. \tag{4.28}$$

Then, Lemma 3.8 in [28] shows that the following inequality holds

$$\|D_k^{-T}\mathrm{vec}\Delta X\|^2 + \|D_k \mathrm{vec}\Delta S\|^2 \leq \frac{38 n^2 \mu_k}{\gamma_1} \tag{4.29}$$

if $n \geq 4$.

Finally, using the above results we can prove the following theorem:

Theorem 4.1 *Let γ_1, γ_2, γ_3 be as in (2.15). Let (X_*, y_*, S_*) be a solution of (1.2)-(1.5). Assume that the constant β and the sequence $\{\sigma_k\}$ are such that $\beta - \sigma_k > \theta_1 > 0$. Further, assume that there exists $\bar{\sigma} > 0$ such that $\sigma_k > \bar{\sigma}$, for all $k > 0$. Then the EIP-SDP Method with $\varepsilon > 0$ terminates in a finite number of iterations. In particular it needs no more than $O[n^{2.5}\ln(X_0 \bullet S_0/\varepsilon)]$ iterations to achieve $X_k \bullet S_k \leq \varepsilon$.*

Proof. Note that

$$\sqrt{\frac{\lambda_1^k}{\lambda_n^k}} \leq \sqrt{\frac{X_k \bullet S_k}{\gamma_1 X_k \bullet S_k / n}} = \sqrt{\frac{n}{\gamma_1}}. \tag{4.30}$$

Then, using (4.21), (4.22) and the bound (4.29), from (4.14) we get:

$$\hat{\alpha}_k \geq \bar{\alpha} = \min\left\{1, \frac{\gamma_1\sqrt{\gamma_1}}{19n^{2.5}} \min\left\{\frac{1-\gamma_1}{1+\gamma_1}\bar{\sigma}, \theta_1\right\}\right\}. \tag{4.31}$$

Moreover, from Proposition 4.1 we have that

$$X_{k+1} \bullet S_{k+1} \leq (1 - \bar{\alpha}(1-\beta)) X_k \bullet S_k. \tag{4.32}$$

Then, the duality gap is reduced at each iteration at least of a factor of

$$1 - (1-\beta)\frac{\gamma_1\sqrt{\gamma_1}}{19n^{2.5}} \min\left\{\frac{1-\gamma_1}{1+\gamma_1}\bar{\sigma}, \theta_1\right\}$$

and this implies that $X_k \bullet S_k \leq \varepsilon$ in at most $O[n^{2.5}\ln(X_0 \bullet S_0/\varepsilon)]$ iterations. □

4.2. Convergence of the IIP-SDP Method. Now, let us consider the inexact case. We cannot directly apply the convergence theory given in [28], because, unlike the exact case, it is not true that $r_p^{(k+1)} = (1-\alpha_k) r_p^{(k)}$. In fact, we have

$$r_p^{(k+1)} = (1-\alpha_k) r_p^{(k)} - \alpha_k r_k$$

and in order to prove the global convergence of the sequence $\{X_k \bullet S_k\}$ to zero, we need to assume the boundedness of the iterates (assumption A1). The proof of our convergence result requires some intermediate stages concerning the boundedness of $\|E_k^{-1}F_k\|$, $\|M_k^{-1}\|$, and $\|\mathcal{H}(\Delta X \Delta S)\|_F$. For sake of simplicity we omit the iteration index k in the proofs of these results.

First note that, recalling Lemma 6.1, part (f) in the Appendix, we have

$$\|E_k\| = 2\|S_k \otimes S_k\| = 2\|S_k\|^2 \leq 2c_1^2; \tag{4.33}$$

similarly, from Lemma 6.1, part (a) we have

$$\|E_k^{-1}\| = \frac{1}{2}\|S_k^{-1} \otimes S_k^{-1}\| = \frac{1}{2}\|S_k^{-1}\|^2. \tag{4.34}$$

Proposition 4.3 *Let assumption A1 be satisfied and let $\{(X_k, y_k, S_k)\}$ be the sequence generated by the IIP-SDP Method. Then, the following inequalities hold:*

$$\|E_k^{-1}F_k\| \leq \frac{c_1^2 n}{\gamma_1 X_k \bullet S_k}, \tag{4.35}$$

$$\|(E_k^{-1}F_k)^{-1}\| \leq \frac{c_1^2 n}{\gamma_1 X_k \bullet S_k}. \tag{4.36}$$

Proof. Let us consider the matrix $E^{-1}F$. We have

$$\begin{aligned}
E^{-1}F &= \frac{1}{2}(X \otimes S^{-1} + S^{-1} \otimes X) \\
&= \frac{1}{2}(X^{1/2} \otimes X^{1/2})(X^{1/2} \otimes X^{-1/2}S^{-1} + X^{-1/2}S^{-1} \otimes X^{1/2}) \\
&= \frac{1}{2}(X^{1/2} \otimes X^{1/2})(I \otimes X^{-1/2}S^{-1}X^{-1/2} + X^{-1/2}S^{-1}X^{-1/2} \otimes I)(X^{1/2} \otimes X^{1/2}).
\end{aligned}$$

This yields

$$\begin{aligned}
\|E^{-1}F\| &\leq \frac{1}{2}\|X^{1/2} \otimes X^{1/2}\|^2(\|I \otimes X^{-1/2}S^{-1}X^{-1/2}\| + \|X^{-1/2}S^{-1}X^{-1/2} \otimes I\|) \\
&= \|X^{1/2}\|^4\|X^{-1/2}S^{-1}X^{-1/2}\| \\
&\leq c_1^2\|X^{-1/2}S^{-1}X^{-1/2}\| \\
&= c_1^2\frac{1}{\lambda_n}.
\end{aligned}$$

Therefore, from (2.12), we get (4.35).

Now, we turn our attention to $\|(E^{-1}F)^{-1}\|$. At this regard, we recall that the matrix $E^{-1}F$ is s.p.d. Let v be an eigenvector of $E^{-1}F$ corresponding to the eigenvalue $\lambda_{n^2}(E^{-1}F)$ and assume that $\|v\| = 1$. From the definition of $\hat{F}$ (see (4.17)) it follows that the relation $E^{-1}F = E^{-1/2}\hat{F}E^{-1/2}$ holds. Then we get

$$\lambda_{n^2}(E^{-1}F) = \|E^{-1/2}\hat{F}E^{-1/2}v\| \geq \frac{\|\hat{F}E^{-1/2}v\|}{\|E^{1/2}\|} \geq \frac{\|v\|}{\|E^{1/2}\|^2\|\hat{F}^{-1}\|} = \frac{\lambda_{n^2}(\hat{F})}{\|E\|}.$$

Thus, from Remark 4.1, (2.12) and (4.33) we get

$$\lambda_{n^2}(E^{-1}F) \geq \frac{\gamma_1 X \bullet S}{c_1^2 n},$$

and this yields (4.36).

□

Proposition 4.4 *Let assumption A1 be satisfied and* $\{(X_k, y_k, S_k)\}$ *be the sequence generated by the IIP-SDP Method. Then,*

$$\|M_k^{-1}\| \leq \frac{c_1^2 n}{\gamma_1 X_k \bullet S_k \sigma_m^2(\mathcal{A})}. \tag{4.37}$$

Proof. Let $\hat{\mathcal{A}} = \mathcal{A}E^{-1/2}$. Then, from the definition of $\hat{F}$ (see (4.17)), it follows

$$M = \mathcal{A}E^{-1}F\mathcal{A}^T = \mathcal{A}E^{-1/2}E^{-1/2}FE^{1/2}E^{-1/2}\mathcal{A}^T = \hat{\mathcal{A}}\hat{F}\hat{\mathcal{A}}^T.$$

Let v be an eigenvector of M corresponding to the eigenvalue $\lambda_m(M)$ and assume that $\|v\| = 1$. Further, let $y = \hat{\mathcal{A}}^T v$; we have

$$\lambda_m(M) = v^T\mathcal{A}E^{-1}F\mathcal{A}^T v = v^T\hat{\mathcal{A}}\hat{F}\hat{\mathcal{A}}^T v = y^T\hat{F}y.$$

Since $\hat{F}$ is s.p.d. we have

$$\lambda_m(M) \geq \lambda_{n^2}(\hat{F})\|y\|^2;$$

Then, Remark 4.1 and (2.12) yield

$$\begin{aligned}\lambda_m(M) &\geq 2\gamma_1 \frac{X \bullet S}{n} \|y\|^2 = 2\gamma_1 \frac{X \bullet S}{n} \|\hat{\mathcal{A}}^T v\|^2 \\ &= 2\gamma_1 \frac{X \bullet S}{n} \|E^{-1/2} \mathcal{A}^T v\|^2. \end{aligned} \tag{4.38}$$

Let $\mathcal{A} = U \Sigma V^T$ be the singular value decomposition of $\mathcal{A}$, i.e. $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n^2\times n^2}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{m\times n^2}$ is given by $\Sigma = (\Sigma_1 \ 0)$, with $\Sigma_1 = \text{diag}(\sigma_i(\mathcal{A})) \in \mathbb{R}^{m\times m}$. Then,

$$\|E^{-1/2} \mathcal{A}^T v\| = \|E^{-1/2} V \Sigma^T U^T v\| = \|E^{-1/2} V \Sigma^T q\|,$$

with $q = U^T v$; therefore $\|q\| = 1$. Since $\mathcal{A}$ has full row rank, Σ_1 is nonsingular. Further, $\|\Sigma^T q\| = \|\Sigma_1 q\|$. Hence,

$$\begin{aligned}\|E^{-1/2} \mathcal{A}^T v\| &= \|E^{-1/2} V \Sigma^T q\| \geq \frac{\|V \Sigma^T q\|}{\|E^{1/2}\|} = \frac{\|\Sigma^T q\|}{\|E^{1/2}\|} = \frac{\|\Sigma_1 q\|}{\|E^{1/2}\|} \\ &\geq \frac{\|q\|}{\|E^{1/2}\| \|\Sigma_1^{-1}\|} = \frac{\sigma_m(\mathcal{A})}{\|E^{1/2}\|}. \end{aligned}$$

Then, using (4.33) we get $\|E^{-1/2} \mathcal{A}^T v\| \geq \frac{\sigma_m(\mathcal{A})}{\sqrt{2} c_1}$ and therefore by (4.38) we can conclude that

$$\lambda_m(M) \geq \frac{\gamma_1 X \bullet S \sigma_m^2(\mathcal{A})}{n\, c_1^2},$$

and this completes the proof. □

Proposition 4.5 *Let assumption A1 be satisfied and let $\{(X_k, y_k, S_k)\}$ be the sequence generated by the IIP-SDP Method. Then,*

$$\|S_k^{-1}\| \leq \frac{c_1 n \sqrt{n}}{\gamma_1 X_k \bullet S_k}. \tag{4.39}$$

Proof. Let us consider the quantity $\|D \text{vec} I\|^2$. Using (4.16), (4.19) and Lemma 6.1, part (d), we get:

$$\begin{aligned}\|D \text{vec} I\|^2 &= (\text{vec} I)^T E^{-1} \hat{S}^{1/2} \hat{S}^{1/2} E^{-1} \text{vec} I \\ &= (\text{vec} I)^T E^{-1/2} \hat{F} E^{-1/2} \text{vec} I \\ &= \frac{1}{2} (\text{vec}(S^{-1/2} I S^{-1/2}))^T \hat{F} \text{vec}(S^{-1/2} I S^{-1/2}) \\ &\geq \frac{1}{2} \lambda_{n^2}(\hat{F}) \|\text{vec} S^{-1}\|^2. \end{aligned}$$

Now, Remark 4.1 and (2.12) yield

$$\|D \text{vec} I\|^2 \geq \lambda_n \|\text{vec} S^{-1}\|^2 \geq \frac{\gamma_1 X \bullet S}{n} \|\text{vec} S^{-1}\|^2. \tag{4.40}$$

Moreover, from (4.20) and (4.35) it follows

$$\begin{aligned}\|D\text{vec}I\|^2 &= (\text{vec}I)^T D^T D\text{vec}I\\ &= (\text{vec}I)^T E^{-1} F\text{vec}I\\ &\le \|E^{-1}F\text{vec}I\|\|\text{vec}I\|\\ &\le \|E^{-1}F\|\|\text{vec}I\|^2\\ &\le \frac{c_1^2 n}{\gamma_1 X\bullet S}n.\end{aligned}$$

Then from (4.40) we obtain

$$\|\text{vec}S^{-1}\| \le \sqrt{\frac{n}{\gamma_1 X\bullet S}}\|D\text{vec}I\| \le \frac{c_1 n\sqrt{n}}{\gamma_1 X\bullet S} \tag{4.41}$$

and, since $\|S^{-1}\| \le \|S^{-1}\|_F = \|\text{vec}S^{-1}\|$, (4.39) holds. □

Proposition 4.6 *Let assumption A1 be satisfied and $\{(X_k, y_k, S_k)\}$ be the sequence generated by the IIP-SDP Method. Assume that for a fixed $\tilde{\varepsilon} > 0$ we have $X_k \bullet S_k \ge \tilde{\varepsilon}$ for all $k > 0$. Then there exists a constant ω such that at each iteration $\|\mathcal{H}(\Delta X\Delta S)\|_F \le \omega$.*

Proof. We will prove that there exists a constant $\omega > 0$ such that

$$\frac{1}{2}\sqrt{\frac{\lambda_1}{\lambda_n}}(\|D^{-T}\text{vec}\Delta X\|^2 + \|D\text{vec}\Delta S\|^2) \le \omega. \tag{4.42}$$

Then, recalling (4.21), the thesis will follow. Thus, we need to bound each element in the left hand side of (4.42).

Using (2.6), (2.8) and the definition (4.19) of the matrix D, we get

$$\begin{aligned}D^{-T}\text{vec}\Delta X &= D^{-T}(E^{-1}\text{vec}R_c - E^{-1}F\text{vec}R_d + E^{-1}F\mathcal{A}^T\Delta y)\\ &= \hat{S}^{-1/2}EE^{-1}F\mathcal{A}^T\Delta y - \hat{S}^{-1/2}EE^{-1}F\text{vec}R_d\\ &\quad +\hat{S}^{-1/2}EE^{-1}\text{vec}(2(\sigma\mu S - SXS))\\ &= \hat{S}^{-1/2}F\mathcal{A}^T\Delta y - \hat{S}^{-1/2}F\text{vec}R_d + 2\sigma\mu\hat{S}^{-1/2}\text{vec}S - 2D^{-T}E^{-1}\text{vec}(SXS)\\ &= D\mathcal{A}^T\Delta y - D\text{vec}R_d + 2\sigma\mu\hat{S}^{-1/2}\text{vec}S - D^{-T}\text{vec}X.\end{aligned}$$

Regarding $D\text{vec}\Delta S$ we have from (2.6)

$$D\text{vec}\Delta S = D\text{vec}R_d - D\mathcal{A}^T\Delta y.$$

Then, we have

$$\|D^{-T}\text{vec}\Delta X\| \le \|D\mathcal{A}^T\Delta y\| + \|D\text{vec}R_d\| + 2\sigma\mu\|\hat{S}^{-1/2}\text{vec}S\| + \|D^{-T}\text{vec}X\| \tag{4.43}$$

and

$$\|D\text{vec}\Delta S\| \le \|D\text{vec}R_d\| + \|D\mathcal{A}^T\Delta y\|. \tag{4.44}$$

Now, we proceed to bound every single term in the right hand sides of (4.43) and (4.44). Let us begin from $2\sigma\mu\|\hat{S}^{-1/2}\text{vec}S\|$. Relation (4.16) and Lemma 6.1, part (d), yield:

$$\begin{aligned}\|\hat{S}^{-1/2}\text{vec}S\|^2 &= (\text{vec}S)^T\hat{S}^{-1}\text{vec}S\\ &= (\text{vec}S)^T E^{-1/2}\hat{F}^{-1}E^{-1/2}\text{vec}S\\ &= \frac{1}{\sqrt{2}}(\text{vec}S)^T E^{-1/2}\hat{F}^{-1}\text{vec}I.\end{aligned}$$

Then, from (4.34), the bound (4.39), Remark 4.1 and (2.12) we get

$$\begin{aligned}\|\hat{S}^{-1/2}\text{vec}S\|^2 &\leq \frac{1}{\sqrt{2}}\|\text{vec}S\|\|E^{-1/2}\|\|\hat{F}^{-1}\|\|\text{vec}I\|\\ &\leq \sqrt{n}c_1\frac{1}{2}\frac{c_1 n\sqrt{n}}{\gamma_1 X\bullet S}\frac{n}{2\gamma_1 X\bullet S}\sqrt{n} = \frac{1}{4}\frac{c_1^2 n^3\sqrt{n}}{(\gamma_1 X\bullet S)^2}.\end{aligned}$$

Therefore we have

$$\begin{aligned}2\sigma\mu\|\hat{S}^{-1/2}\text{vec}S\| &\leq \sigma\mu\frac{c_1 n n^{3/4}}{\gamma_1 X\bullet S}\\ &= \frac{\sigma c_1 n^{3/4}}{\gamma_1}.\end{aligned} \tag{4.45}$$

Concerning the term $\|D^{-T}\text{vec}X\|$, proceeding as before, taking into account that $D^{-1}D^{-T} = (E^{-1}F)^{-1}$ and (4.36) we get:

$$\|D^{-T}\text{vec}X\| \leq \|\text{vec}X\|\|(E^{-1}F)^{-1}\|^{1/2} \leq \frac{c_1^2 n}{\sqrt{\gamma_1 X\bullet S}}. \tag{4.46}$$

Now, let us consider $\|D\text{vec}R_d\|$. Again, using (4.20) we get

$$\|D\text{vec}R_d\| \leq \|\text{vec}R_d\|\|E^{-1}F\|^{1/2}.$$

Therefore, from (4.23), (4.24) and (4.35) we obtain

$$\|D\text{vec}R_d\| \leq \frac{c_1\sqrt{nX\bullet S}}{\sqrt{\gamma_1}X_0\bullet S_0}\|\text{vec}R_d^{(0)}\|. \tag{4.47}$$

Concerning the term $\|D\mathcal{A}^T\Delta y\|$ we have, recalling (4.20),

$$\|D\mathcal{A}^T\Delta y\|^2 = \Delta y^T\mathcal{A}D^T D\mathcal{A}^T\Delta y = \Delta y^T\mathcal{A}E^{-1}F\mathcal{A}^T\Delta y \leq \|\Delta y\|\|M\Delta y\|. \tag{4.48}$$

Let us consider system (3.1). We have

$$\begin{aligned}M\Delta y &= r_p - \mathcal{A}E^{-1}\text{vec}R_c + \mathcal{A}E^{-1}F\text{vec}R_d + r\\ &= b - \mathcal{A}\text{vec}X - \mathcal{A}E^{-1}\text{vec}(2(\sigma\mu S - SXS)) + \mathcal{A}E^{-1}F\text{vec}R_d + r\\ &= b - \mathcal{A}\text{vec}X - 2\sigma\mu\mathcal{A}E^{-1}\text{vec}S + 2\mathcal{A}E^{-1}\text{vec}(SXS) + \mathcal{A}E^{-1}F\text{vec}R_d + r\\ &= b - \mathcal{A}\text{vec}X - \sigma\mu\mathcal{A}\text{vec}S^{-1} + \mathcal{A}\text{vec}X - \mathcal{A}E^{-1}F\text{vec}R_d + r\\ &= b - \sigma\mu\mathcal{A}\text{vec}S^{-1} - \mathcal{A}E^{-1}F\text{vec}R_d + r.\end{aligned}$$

Thus, by (3.2), the inequalities (4.23) and (4.24), and the bounds (4.35) and (4.41),

$$
\begin{aligned}
\|M\Delta y\| &\leq \|b - \sigma\mu\mathcal{A}\mathrm{vec}S^{-1} - \mathcal{A}E^{-1}F\mathrm{vec}R_d + r\| \\
&\leq \|b\| + \sigma\mu\sigma_1(\mathcal{A})\frac{c_1 n\sqrt{n}}{\gamma_1 X \bullet S} + \sigma_1(\mathcal{A})\|E^{-1}F\|\|\mathrm{vec}R_d\| + \|r\| \\
&\leq \|b\| + \sigma\sigma_1(\mathcal{A})\frac{c_1\sqrt{n}}{\gamma_1} + \sigma_1(\mathcal{A})\frac{c_1^2 n}{\gamma_1 X \bullet S}\frac{X \bullet S}{X_0 \bullet S_0}\|\mathrm{vec}R_d^{(0)}\| + \eta X \bullet S \\
&\leq \|b\| + \sigma_1(\mathcal{A})\frac{c_1^2 n}{\gamma_1}\frac{\|\mathrm{vec}R_d^{(0)}\|}{X_0 \bullet S_0} + \sigma\sigma_1(\mathcal{A})\frac{c_1\sqrt{n}}{\gamma_1} + \eta X \bullet S. \qquad (4.49)
\end{aligned}
$$

Thus, by (4.37) we have

$$
\begin{aligned}
\|\Delta y\| &\leq \frac{c_1^2 n}{\gamma_1 X \bullet S \sigma_m^2(\mathcal{A})}\left(\|b\| + \frac{\sigma_1(\mathcal{A})c_1^2 n}{\gamma_1}\frac{\|\mathrm{vec}R_d^{(0)}\|}{X_0 \bullet S_0} + \sigma\sigma_1(\mathcal{A})\frac{c_1\sqrt{n}}{\gamma_1} + \eta X \bullet S\right) \\
&\leq \frac{c_1^2 n}{\gamma_1 X \bullet S \sigma_m^2(\mathcal{A})}\left(\|b\| + \frac{\sigma_1(\mathcal{A})c_1^2 n}{\gamma_1}\frac{\|\mathrm{vec}R_d^{(0)}\|}{X_0 \bullet S_0} + \sigma\sigma_1(\mathcal{A})\frac{c_1\sqrt{n}}{\gamma_1}\right) + \eta\frac{c_1^2 n}{\gamma_1\sigma_m^2(\mathcal{A})} \qquad (4.50)
\end{aligned}
$$

Summarizing, from (4.21) and bounds (4.30) and (4.45)-(4.50) it follows that if $X_k \bullet S_k \geq \tilde{\varepsilon}$, there exists a constant ω such that $\|\mathcal{H}(\Delta X \Delta S)\|_F \leq \omega$. □

With these preliminary results at hand we are now ready to prove the following result.

Theorem 4.2 *Let γ_1, γ_2, γ_3 be as in* (2.15). *Let assumption A1 be satisfied. Assume that the constant β and the sequences $\{\eta_k\}$ and $\{\sigma_k\}$ are such that $\beta - \sigma_k > \theta_1 > 0$ and $\sigma_k - \gamma_2\eta_k > \theta_2 > 0$, for all $k > 0$. Further, assume that σ_k is bounded away from zero whenever $X_k \bullet S_k \not\to 0$. Then the sequence $\{X_k \bullet S_k\}$ generated by the IIP-SDP Method with $\varepsilon = 0$ converges to 0.*

Proof. The sequence $X_k \bullet S_k$ is monotone decreasing and bounded, therefore it is convergent. Assume that $X_k \bullet S_k \to \tilde{\varepsilon} > 0$. Note that, from our assumptions, it follows that there exists $\bar{\sigma}$ such that $\sigma_k \geq \bar{\sigma}$, for all $k > 0$. ¿From Proposition 4.1 we have that

$$X_{k+1} \bullet S_{k+1} \leq (1 - \hat{\alpha}_k(1-\beta))X_k \bullet S_k, \qquad (4.51)$$

with $\hat{\alpha}_k$ given by (4.14). Furthermore, under our assumptions, by Proposition 4.6 it follows that $\hat{\alpha}_k \geq \bar{\alpha}$, where $\bar{\alpha}$ is independent of k and given by:

$$\bar{\alpha} = \min\left\{1, \frac{\tilde{\varepsilon}}{n\omega}\min\left\{\frac{1-\gamma_1}{1+\gamma_1}\bar{\sigma}, \theta_1, \theta_2\right\}\right\}.$$

Then $\hat{\alpha}_k$ is bounded away from zero, and this along with (4.51) implies that $X_k \bullet S_k \to 0$, which is a contradiction. □

5. Practical considerations. In this section, we will spend a few words on the computation of the matrix-vector products $M_k w$, for some vector $w \in \mathbb{R}^m$. This is a crucial implementation issue for an inexact method. In fact, iterative solvers for linear systems only require the action of the coefficient matrix onto a vector w and this computation must be carried out efficiently.

For sake of simplicity we omit the index k. The computation of the product Mw, requires computing the product $z = E^{-1}Fv$, for some $v \in \mathbb{R}^{n^2}$. This can be done without building E and then solving the n^2-dimensional linear system $Ez = Fv$. Recall that $E^{-1} = \frac{1}{2}S^{-1} \otimes S^{-1}$ and $E^{-1}F = \frac{1}{2}\left((X \otimes S^{-1}) + (S^{-1} \otimes X)\right)$. Let us consider first products of the kind $\hat{z} = (X \otimes S^{-1})v$. Let $V \in \mathbb{R}^{n \times n}$ be such that $V = \text{mat} v$, and let $V_{i\cdot}$ and $V_{\cdot j}$ denote the i-th row and j-th column of V, respectively. It is straightforward to note that

$$(X \otimes S^{-1})v = \begin{bmatrix} \sum_{j=1}^n X_{1j} S^{-1} V_{\cdot j} \\ \vdots \\ \sum_{j=1}^n X_{nj} S^{-1} V_{\cdot j} \end{bmatrix};$$

therefore we can solve the n-dimensional linear systems $S\hat{z}^{(l)} = V_{\cdot l}$, for $l = 1, ..., n$, and then build $(X \otimes S^{-1})v$ through linear combinations of the vectors $\hat{z}^{(l)}$.

Now, let us turn to products of the kind $\bar{z} = (S^{-1} \otimes X)v$. We have

$$(S^{-1} \otimes X)v = \begin{bmatrix} \sum_{j=1}^n S^{-1}_{1j} X V_{\cdot j} \\ \vdots \\ \sum_{j=1}^n S^{-1}_{nj} X V_{\cdot j} \end{bmatrix}.$$

Let $G \in \mathbb{R}^{n \times n}$ be such that $G_{\cdot j} = XV_{\cdot j}$ for $j = 1, ..., n$ and let $\bar{Z} = \text{mat}\bar{z}$. We have, for $i = 1, ..., n$,

$$\bar{Z}_{\cdot i} = \sum_{j=1}^n S^{-1}_{ij} G_{\cdot j}$$

and therefore, for $i, l = 1, ..., n$,

$$\bar{Z}_{li} = \sum_{j=1}^n S^{-1}_{ij} G_{lj}.$$

As a consequence, for each $l = 1, \ldots, n$, the vector $\bar{Z}^T_{l\cdot}$ is nothing but the solution of the linear system $S\bar{Z}^T_{l\cdot} = G^T_{l\cdot}$. In conclusion, to compute the matrix-vector product $E^{-1}Fv$, we mainly need to solve $2n$ linear systems of dimension n, with coefficient matrix always equal to S, and therefore s.p.d.

6. Appendix. In this section we collect some known results on eigenvalues and Kronecker products [9, 10].

Theorem 6.1 *(Weyl's Theorem) Let A, B be two given $n \times n$ symmetric matrices. Then, for all $k = 1, ..., n$, we have*

$$\lambda_k(A) + \lambda_n(B) \le \lambda_k(A + B) \le \lambda_k(A) + \lambda_1(B).$$

Remark 6.1 Recalling that $\mathrm{tr}(B) = \sum \lambda_i(B)$, for any symmetric matrix B we have $\|B\|_F^2 = \mathrm{tr}\big(B^T B\big) = \mathrm{tr}\big(B^2\big) = \sum_{i=1}^n |\lambda_i(B)|^2$; hence, $|\lambda_i(B)| \le \|B\|_F$. From this and from Theorem 6.1 we have, for any two symmetric matrices A and B,

$$\lambda_n(A+B) \ge \lambda_n(A) + \lambda_n(B) \ge \lambda_n(A) - |\lambda_n(B)| \ge \lambda_n(A) - \|B\|_F. \tag{6.1}$$

Lemma 6.1 *Let A, B, C, D be matrices of proper dimensions. Then the following properties hold:*

(a) $A \otimes B$ is nonsingular iff A and B are nonsingular; moreover, $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
(b) $(A \otimes B)(C \otimes D) = AC \otimes BD$.
(c) $(A \otimes B)^T = A^T \otimes B^T$.
(d) $(A \otimes B)\mathrm{vec}X = \mathrm{vec}(BXA^T)$.
(e) $\Lambda(A \otimes B) = \{\lambda_i(A)\lambda_j(B), i, j = 1, ..., n\}$.
(f) $\|A \otimes B\| = \|A\|\|B\|$.
(g) $\Lambda(A \otimes I + I \otimes B) = \{\lambda_i(A) + \lambda_j(B), i, j = 1, ..., n\}$.

Acknowledegments The authors are grateful to an anonymous referee for pointing out an error in a previous version of the paper.

REFERENCES

[1] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM Journal on Optimization, 5 (1995), pp. 13–51.
[2] F. Alizadeh, J.-P. Haeberly, and M. L. Overton, *Primal-Dual Interior-Point Methods for Semidefinite Programming: Convergence Rates, Stability and Numerical Results*, SIAM Journal on Optimization, 8 (1998), pp. 746–768.
[3] V. Baryamureeba and T. Steihaug, *On the Convergence of an Inexact Primal-Dual Interior Point Method for Linear Programming*, Tech. Rep. 188, Department of Informatics, University of Bergen, Bergen, Norway, 2000.
[4] S. Bellavia, *An Inexact Interior Point method*, Journal of Optimization Theory and Applications, 96 (1998).
[5] S. Bellavia and M. Macconi, *An Inexact Interior Point method for monotone NCP*, Optimization Methods and Software, 11&12 (1999), pp. 211–241.
[6] J. Bonnans, C. Pola, and R. Rébai, *Perturbed Path Following Interior Point Algorithms* , Optimization Methods and Software, 11&12 (1999), pp. 183–210.
[7] R. W. Freund, F. Jarre, and S. Mizuno, *Convergence of a class of inexact interior point algorithms for linear programming*, Mathematics of Operations Research, 24 (1999), pp. 50–71.
[8] C. Helmberg, F. Rendl, R. Vanderbei, and H. Wolkowicz, *An Interior Point method for Semidefinite Programming* , SIAM Journal on Optimization, 6 (1996), pp. 342–361.
[9] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
[10] ———, *Topics in Matrix Analysis*, Cambridge University Press, 1991.
[11] M. Kojima, N. Megiddo, and S. Mizuno, *A primal-dual infeasible-interior-point algorithm for Linear Programming*, Mathematical Programming, 61 (1993), pp. 263–280.
[12] M. Kojima, M. Shida, and S. Shindoh, *Local Convergence of predictor-corrector infeasible Interior-Point algorithms for SDPs and SDLCPs*, Mathematical Programming (Series A), 80 (1998), pp. 129–160.
[13] ———, *Search Directions in the SDP and the Monotone SDLCP: Generalization and Inexact Computation*, Mathematical Programming, 85 (1999), pp. 51–80.
[14] M. Kojima, S. Shindoh, and S. Hara, *Interior-Point methods for the Monotone Semidefinite Linear Complementarity Problem in symmetric matrices*, SIAM Journal on Optimization, 7 (1997), pp. 86–125.
[15] J. Korzak, *Convergence Analysis of Inexact Infeasible-Interior-Point Algorithms for Solving Linear Programming Problems*, SIAM Journal on Optimization, 11 (2000), pp. 133–148.
[16] S. Mizuno and F.Jarre, *Global and polynomial-time convergence of an infeasible interior point algorithm using inexact computation*, Mathematical Programming, 84 (1999), pp. 357–373.

[17] S. Mizuno, M. Todd, and Y. Ye, *On adaptive-step primal-dual Interior-Point algorithms for linear programming*, Mathematics of Operations Research, 18 (1993), pp. 964–981.

[18] R. Monteiro, *Primal-Dual path following algorithms for Semidefinite Programming*, SIAM Journal on Optimization, 7 (1997), pp. 663–678.

[19] L. Portugal, L. Fernandes, and J. Judice, *A truncated Newton interior-point algorithm for the solution of a multicommodity spatial equilibrium model*, in Complentarity and Variational Problems, M. Ferris and J.-S. Pang, eds., SIAM, 1997, pp. 315–344.

[20] L. Portugal, M. Resende, G. Veiga, and J. Judice, *A truncated primal-infeasible dual-feasible network interior point method*, Networks, 35 (2000), pp. 91–108.

[21] F. A. Potra and R. Sheng, *A superlinearly convergent primal-dual infeasible Interior-Point Algorithm for Semidefinite Programming*, SIAM Journal on Optimization, 8 (1998), pp. 1007–1098.

[22] ———, *Superlinear convergence of Interior-Point Algorithms for Semidefinite Programming*, Journal of Optimization Theory and Applications, 99 (1998), pp. 103–119.

[23] F. A. Potra and S. Wright, *Interior-Point Methods*, Journal of Computational and Applied Mathematics, 124 (2000), pp. 281–302.

[24] M. Todd, *Semidefinite Optimization*, Acta Numerica, 10 (2001), pp. 515–560.

[25] K.-C. Toh and M. Kojima, *Solving Some Large Scale Semidefinite Programs via the Conjugate Residual Method*, SIAM Journal on Optimization, 12 (2002), pp. 669–691.

[26] L. Vandenberghe and S. Boyd, *Semidefinite Programming*, SIAM Review, 38 (1996), pp. 49–95.

[27] Y. Zhang, *On the convergence of a class of infeasible Interior-Point Algorithms for the horizontal linear complementarity problems*, SIAM Journal on Optimization, 4 (1994), pp. 208–227.

[28] ———, *On Extending some Primal-Dual Interior-Point Algorithms from Linear Programming to Semidefinite Programming*, SIAM Journal on Optimization, 8 (1998), pp. 365–386.

[29] G. Zhou and K.-C. Toh, *Polynomiality of an Inexact Infeasible Interior Point Algorithm for Semidefinite Programming*. to appear in Mathematical Programming.