Numer Algor (2009) 52:225–255
DOI 10.1007/s11075-009-9270-7

ORIGINAL PAPER

A new full-Newton step $O(n)$ infeasible interior-point algorithm for semidefinite optimization

H. Mansouri · C. Roos

Received: 16 July 2008 / Accepted: 1 February 2009 /
Published online: 26 February 2009

Abstract Interior-point methods for semidefinite optimization have been studied intensively, due to their polynomial complexity and practical efficiency. Recently, the second author designed a primal-dual infeasible interior-point algorithm with the currently best iteration bound for linear optimization problems. Since the algorithm uses only full Newton steps, it has the advantage that no line-searches are needed. In this paper we extend the algorithm to semidefinite optimization. The algorithm constructs strictly feasible iterates for a sequence of perturbations of the given problem and its dual problem, close to their central paths. Two types of full-Newton steps are used, feasibility steps and (ordinary) centering steps, respectively. The algorithm starts from strictly feasible iterates of a perturbed pair, on its central path, and feasibility steps find strictly feasible iterates for the next perturbed pair. By using centering steps for the new perturbed pair, we obtain strictly feasible iterates close enough to the central path of the new perturbed pair. The starting point depends on a positive number ζ. The algorithm terminates either by finding an ε-solution or by detecting that the primal-dual problem pair has no optimal solution (X^*, y^*, S^*) with vanishing duality gap such that the eigenvalues of X^* and S^*

H. Mansouri
Department of Applied Mathematics, Shahrekord University,
P.O. Box 115, Shahrekord, Iran

H. Mansouri (✉) · C. Roos
Department of Electrical Engineering, Mathematics and Computer Science,
Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands
e-mail: H.Mansouri@tudelft.nl

C. Roos
e-mail: C.Roos@tudelft.nl

do not exceed ζ. The iteration bound coincides with the currently best iteration bound for semidefinite optimization problems.

Keywords Semidefinite optimization · Infeasible interior-point method · Primal-dual method · Polynomial complexity

Mathematics Subject Classifications (2000) 90C05 · 90C51

1 Introduction

Semidefinite optimization (SDO) problems are convex optimization problems over the intersection of an affine set and the cone of positive semidefinite matrices. SDO arises in many scientific and engineering fields. For applications in system and control theory we refer to [6, 7] and for applications in combinatorial optimization to [1, 2, 13, 19, 26, 39, 48]. SDO also has been utilized in solving polynomial optimization problems [11, 25, 38]. For recent applications in engineering problems see, e.g., [5, 24, 59].

Semidefinite optimization has recently attracted active research from the interior-point methods (IPMs) community. A rather comprehensive list of references for theory and applications in this field can be found in [1, 2, 8, 12, 15, 18, 37, 51].

Primal-dual IPMs have proven to belong to the most efficient methods in linear optimization (LO), and many polynomiality results exist for these methods. The first primal-dual interior-point methods for LO were constructed by Megiddo [32], Monteiro and Adler [35], Tanabe [49] and Kojima et al. [21].

Extending methods for LO to SDO has been successful in many cases. See, e.g., [23, 36, 53, 55]. For example Nesterov and Todd [36] showed that the primal-dual algorithm for LO maintains its theoretical efficiency when the nonnegativity constraints in LO are replaced by a convex cone, as long as the cone is homogeneous and self-dual, or in the terminology of Nesterov and Todd, as long as the cone is self-scaled. Self-scaled cones are cones that have a self-scaled barrier; the non-negative orthant and the cone of positive semidefinite matrices are special cases.

Recently, Peng et al. [40, 41] designed primal-dual feasible IPMs for LO by using self-regular functions and also extended the approach to SDO. The complexity bounds obtained by these authors are $O\left(\sqrt{n}\right)\log\frac{n}{\varepsilon}$ and $O\left(\sqrt{n}\log n\right)\log\frac{n}{\varepsilon}$, for small-update methods and large-update methods, respectively, which are currently the best known bounds. Bai et al. [3] and Bai and Roos [4] introduced some new kernel functions and designed primal-dual feasible IPMs for LO and they also extended their algorithms to SDO successfully [54].

The methods mentioned in the preceding two paragraphs are so-called feasible IPMs. Feasible IPMs start with a strictly feasible interior point and maintain feasibility during the solution process. It is not at all trivial how to find an initial feasible interior point, however. One method to overcome this problem is to use the homogeneous embedding model as presented first for LO by Ye et al. [56] and extended to SDO by De Klerk et al. [10] and independently also by Luo et al. [27]. The homogeneous embedding method has been implemented in the well known solver SeDuMi.[1] It combines the primal and dual problem in an intelligent way in one (much) larger self-dual model, whose solution can be obtained by any feasible IPM, leaving much freedom to the user in the choice of the initial iterates. Other SDO solvers use a different approach, they use a socalled infeasible IPM.

Infeasible IPMs (IIPMs) start with an arbitrary positive point and feasibility is reached as optimality is approached. The choice of the starting point in IIPMs is crucial for the performance. Lustig [28] and Tanabe [50] were the first to present IIPMs for LO. Kojima et al. [20] were the first who proved the global convergence of a primal-dual IIPM for LO. Zhang [57] was the first who presented a primal-dual IIPM with polynomial iteration complexity $O\left(n^2 \log \frac{1}{\varepsilon}\right)$ for LO and he extended it to SDO [58]. Mizuno [33] and Potra [42, 43] also introduced a primal-dual IIPM for LO with polynomial iteration complexity $O\left(n \log \frac{1}{\varepsilon}\right)$.

Kojima et al. [22] and Potra and Sheng [44] independently analyzed a generalization to SDO of the Mizuno-Todd-Ye predictor corrector method [34] for infeasible starting points and they proved that the complexity of their algorithm is

$$O\left(n \log \frac{\max\left\{\mathbf{Tr}\left(X^0 S^0\right),\ \left\|R_c^0\right\|,\ \left\|r_b^0\right\|\right\}}{\varepsilon}\right). \tag{1}$$

Here R_c^0 and r_b^0 denote the initial values of the primal and dual residuals as defined in Section 4.3. It is assumed in this result that there exist optimal solutions X^* and $(y^*,\ S^*)$ for the primal and dual problems (P) and (D) as defined in Section 3.1 such that $\|X^*\| \le \zeta$ and $\|S^*\| \le \zeta$ and the starting point is $\left(X^0,\ y^0,\ S^0\right) = \zeta\,(I,\ 0,\ I)$, where I denotes the identity matrix.

The second author, in [45], designed a new IIPM algorithm for LO. This algorithm uses intermediate problems. These problems are suitable perturbations of the given problems (P) and (D) so that at any stage the iterates are strictly feasible for the current perturbed problem pair. In each iteration the size of the perturbation decreases at the same speed as the barrier parameter μ. When μ changes to a smaller value, the perturbed problem pair corresponding to μ changes, and hence also the current central path. The iterates are kept feasible for the new perturbed problem pair and close to its central path. To achieve this the algorithm uses a so-called *feasibility step*. This step serves to get iterates that are strictly feasible for the new perturbed problem pair and belong to the region of quadratic convergence of its μ^+-centers, where μ^+ is

[1]The name of this solver reveals this feature: Self-Dual Minimization.

the barrier parameter after updating. Now the algorithm can start from the point obtained in the feasibility step and perform a few centering steps to obtain iterates that are close enough to the μ^+-center of the new perturbed problem pair. This process continues until the algorithm finds an ε-solution or detects that the primal-dual problem pair has no optimal solutions with zero duality gap.

In this paper, we discuss an extension to SDO of the just described algorithm. We show that the techniques that have been developed in the field of feasible full-Newton step IPMs for SDO are sufficient to get a full-Newton step IIPM whose complexity is given by (1), which is currently still the best known complexity of IIPMs. Compared with other IIPMs the ideas underlying our algorithm are quite simple and its analysis is elegant. Moreover, since it does not use line searches, it can be easily implemented.

The paper is organized as follows: in Section 3 we first recall some tools for the analysis of a feasible IPM for SDO that we use also in the analysis of IIPMs proposed in this paper. In Section 4 we describe a primal-dual IIPM method for SDO. The analysis of the feasibility step of our method, the most tedious part of the analysis, is carried out in Section 5. In Section 5.5 we will derive the complexity bound for our algorithm. Finally, some concluding remarks follow in Section 6.

2 Notations

Some notations used throughout the paper are as follows. The superscript T denotes transpose. $\mathbf{R}^n$, $\mathbf{R}^n_+$ and $\mathbf{R}^n_{++}$ denote the set of vectors with n components, the set of nonnegative vectors and the set of positive vectors, respectively. For any $x = (x_1;\ x_2;\ \ldots;\ x_n) \in \mathbf{R}^n$, $x_{\min} = \min(x_1;\ x_2;\ \ldots;\ x_n)$ and $x_{\max} = \max(x_1;\ x_2;\ \ldots;\ x_n)$. $\mathbf{R}^{m\times n}$ is the space of all $m \times n$ matrices. $\mathbf{S}^n$, $\mathbf{S}^n_+$ and $\mathbf{S}^n_{++}$ denote the cone of symmetric, symmetric positive semidefinite and symmetric positive definite $n \times n$ matrices, respectively. $\mathcal{P}$ and $\mathcal{D}$ denote the feasible sets of the primal and dual problem respectively. The relative interior of a convex set $\mathcal{C}$ is denoted as $\mathrm{ri}(\mathcal{C})$. I denotes $n \times n$ identity matrix. We use the classical Löwner partial order $\succeq$ for symmetric matrices. So $A \succeq B\,(A \succ B)$ means that $A - B$ is positive semidefinite (positive definite). The sign $\sim$ denotes similarity of two matrices. The matrix inner product is defined by $A \bullet B = \mathbf{Tr}\left(A^T B\right)$. For any symmetric positive definite matrix $Q \in \mathbf{S}^n_{++}$, the expression $Q^{\frac{1}{2}}$ denotes the symmetric square root of Q. For any symmetric matrix G, $\lambda_{\min}(G)$ $(\lambda_{\max}(G))$ denotes the minimal (maximal) eigenvalue of G. When λ is vector we denote the diagonal matrix $\mathrm{diag}(\lambda)$ with entries λ_i by Λ. For any $V \in \mathbf{S}^n_{++}$, we denote by $\lambda(V)$ the vector of eigenvalues of V arranged in non-increasing order, that is, $\lambda_{\max}(V) = \lambda_1(V) \geq \lambda_2(V) \geq \ldots \geq \lambda_n(V) = \lambda_{\min}(V)$. The Frobenius matrix norm is given by $\|U\|^2 := \sum_{i=1}^m \sum_{j=1}^n U_{ij}^2 = \mathbf{Tr}\left(U^T U\right)$. For any $p \times q$ matrix A, $\mathbf{vec}(A)$ denotes the pq-vector obtained by stacking the columns of A. The Kronecker product of two matrices A and B is denoted by $A \otimes B$ (we refer to [16] for a comprehensive treatment on Kronecker products and related topics).

3 Preliminaries

3.1 The SDO problem

In this section we first introduce the SDO problem and the assumptions that will be used in the paper. We consider the standard form of the SDO problem:

$$\begin{aligned} (P) \qquad & \min \quad C \bullet X \\ & s.t \quad A_i \bullet X = b_i, \quad i = 1, 2, \ldots, m, \quad X \succeq 0, \end{aligned}$$

and its dual:

$$\begin{aligned} (D) \qquad & \max \quad b^T y \\ & s.t \sum_{i=1}^{m} y_i \quad A_i + S = C, \quad S \succeq 0, \end{aligned}$$

where each $A_i \in \mathbf{S}^n$, $b \in \mathbf{R}^m$, and $C \in \mathbf{S}^n$. Without loss of generality we assume that the matrices A_i are linearly independent.

It is worth noting that the duality theory of SDO is weaker than that of LO. Like in LO we have the weak duality property: for any $(X, \; y, \; S) \in \mathcal{P} \times \mathcal{D}$ we have

$$C \bullet X - b^T y = \mathbf{Tr}\left(\left(S + \sum_{i=1}^{m} y_i A_i\right) X\right) - \sum_{i=1}^{m} y_i \mathbf{Tr}\,(A_i X) = \mathbf{Tr}\,(SX) \geq 0,$$

where the inequality follows from $X \succeq 0$ and $S \succeq 0$ (see Lemma 5.3). In other words, the duality gap is nonnegative for any feasible primal-dual pair. As a consequence, feasible solutions $(X, \; y, \; S)$ with zero duality gap are optimal. If (P) and/or (D) have optimal solutions, then their optimal values are denotes as p^* and d^*, respectively. The optimal sets for (P) and (D) are then denoted as follows:

$$\mathcal{P}^* := \left\{X \in \mathcal{P} | C \bullet X = p^*\right\} \text{ and } \mathcal{D}^* := \left\{(y, \; S) \in \mathcal{D} | b^T y = d^*\right\}.$$

A problem (P) (resp. (D)) is called solvable if $\mathcal{P}^*$ (resp. $\mathcal{D}^*$) is nonempty.

The duality properties in SDO are less simple than LO. Recall that an LO problem may be feasible or infeasible. If it is feasible then it is either unbounded or bounded. In case it is bounded it is solvable and the dual problem is solvable as well, with zero duality gap. If it is infeasible then its dual is either unbounded or infeasible. So, for a single problem there are 3 possibilities: a problem is either solvable, unbounded or infeasible.

For an SDO problem, however, the situation is less simple, since a problem that is feasible and bounded is not necessarily solvable. So for a single problem there are now four cases: a problem is solvable, feasible and not solvable, unbounded or infeasible. As a consequence, for a primal-dual pair of SDO problems there are much more possible situations than in the LO case. For example, a problem may be solvable, whereas its dual is unsolvable. Also both problems can be solvable but with positive duality gap. We assume that there exists optimal solutions with zero duality gap, i.e., we assume that the set

$$\mathcal{F}^* := \{(X,\ y,\ S) \in \mathcal{P} \times \mathcal{D} : \quad X \bullet S = 0\}$$

is nonempty.

3.2 The central path for SDO

We assume that (P) and (D) are strictly feasible, i.e., there exist $X \in \mathcal{P}$, $S \in \mathcal{D}$ with $X \succ 0$, $S \succ 0$. It is well known that under this assumption both problems are solvable and the optimality conditions for (P) and (D) can be written as follows.

$$\begin{aligned} A_i \bullet X &= b_i, \quad i = 1, 2, \ldots, m,\ X \succeq 0, \\ \sum_{i=1}^{m} y_i A_i + S &= C, \quad S \succeq 0, \\ XS &= 0. \end{aligned} \tag{2}$$

The basic idea of primal-dual IPMs is to replace the above complementarity condition $XS = 0$ by the parameterized equation $XS = \mu I$. Then we get the following system:

$$\begin{aligned} A_i \bullet X &= b_i, \quad i = 1, 2, \ldots, m,\ X \succeq 0, \\ \sum_{i=1}^{m} y_i A_i + S &= C, \quad S \succeq 0, \\ XS &= \mu I, \end{aligned} \tag{3}$$

where I denotes the $n \times n$ identity matrix and $\mu > 0$. It is well known that the system (3) has a unique solution, denoted $(X(\mu),\ y(\mu),\ S(\mu))$, and that the limit $\lim_{\mu \to 0} (X(\mu),\ y(\mu),\ S(\mu))$ exists and is a solution of system (2) (e.g., see [23]). The set of all solutions $(X(\mu),\ y(\mu),\ S(\mu))$ with $\mu > 0$ is known as the central path.

To obtain a search direction for IPMs the usual approach is to use Newton's method and to linearize (3). However, the resulting system may yield as a solution a search direction ΔX which is not symmetric (ΔS is automatically symmetric). Since we want ΔX to be a symmetric matrix, one must "symmetrize" the linearization of the complementary equation. Based on different symmetrization schemes, several search directions have been proposed, as presented in [15, 23, 36, 47, 52, 58]. In this paper, we use the direction determined by the following system:

$$\begin{aligned} A_i \bullet \Delta X &= 0, \quad i = 1, 2, \ldots, m, \\ \sum_{i=1}^{m} \Delta y_i A_i + \Delta S &= 0, \\ \Delta X + P \Delta S P^T &= \mu S^{-1} - X, \end{aligned} \tag{4}$$

where

$$P := X^{\frac{1}{2}} \left\{ X^{\frac{1}{2}} S X^{\frac{1}{2}} \right\}^{\frac{-1}{2}} X^{\frac{1}{2}} = S^{\frac{-1}{2}} \left\{ S^{\frac{1}{2}} X S^{\frac{1}{2}} \right\}^{\frac{1}{2}} S^{\frac{-1}{2}}. \tag{5}$$

This direction was introduced by Nesterov and Todd [36] and is called the NT-direction, or simply the Newton-direction, see [52]. The system (4) has a unique solution [55] and, obviously, ΔX and ΔS are symmetric. We will refer to the assignment

$$\left(X^+, \ y^+, \ S^+\right) := (X + \Delta X, \ y + \Delta y, \ S + \Delta S)$$

as a full NT step.

The first two equations in the system (4) imply that ΔX and ΔS are orthogonal, i.e.,

$$\mathbf{Tr}\,(\Delta X \Delta S) = 0.$$

A consequence is the following result which makes clear that after a full NT-step the duality gap attains its target value.

Lemma 3.1 (Cf. Corollary 7.1 in [9]) *Let* $(X, \ S) \in \mathrm{ri}\,(\mathcal{P} \times \mathcal{D})$ *and* $\mu > 0$. *Then*

$$\mathbf{Tr}\left(X^+ S^+\right) = n\mu.$$

Let $D = P^{\frac{1}{2}}$, where $P^{\frac{1}{2}}$ denotes the positive semidefinite square root of P. Then D can be used to scale X and S to the same matrix V, namely [9, 47, 54]:

$$V := \frac{1}{\sqrt{\mu}} D^{-1} X D^{-1} = \frac{1}{\sqrt{\mu}} DSD. \tag{6}$$

It follows that

$$V^2 = \frac{1}{\mu} D^{-1} XSD. \tag{7}$$

Note that the matrices D and V are symmetric and positive definite. let us further define

$$\bar{A}_i := \frac{1}{\sqrt{\mu}} D A_i D, \quad i = 1, 2, \ldots, m,$$

and

$$D_X := \frac{1}{\sqrt{\mu}} D^{-1} \Delta X D^{-1}; \quad D_S := \frac{1}{\sqrt{\mu}} D \Delta S D.$$

Then (4) can be written as follows

$$\begin{aligned}\bar{A}_i \bullet D_X &= 0, \quad i = 1, 2, \ldots, m,\\ \sum_{i=1}^{m} \Delta y_i \bar{A}_i + D_S &= 0,\\ D_X + D_S &= V^{-1} - V. \end{aligned} \tag{8}$$

The first two equations in this system imply that D_X and D_S are orthogonal:

$$\mathbf{Tr}\,(D_X D_S) = 0.$$

Hence, using the third equation in (8) we obtain

$$\|D_X\|^2 + \|D_S\|^2 = \left\|V^{-1} - V\right\|^2. \tag{9}$$

This implies that $D_X,\ D_S$ are both zero if and only if $V^{-1} - V = 0$. In this case, X and S satisfy $XS = \mu I$, which indicates that X and S are the μ-centers. Thus, we can use $\left\|V^{-1} - V\right\|$ as the quantity to measure closeness to μ-centers. We define

$$\delta\,(X,\ S,\ \mu) := \delta\,(V) := \frac{1}{2}\left\|V^{-1} - V\right\|. \tag{10}$$

Note that for the special case of LO, V is a diagonal matrix and this proximity measure becomes the same as used in [17].

A pair $(X,\ S) \in \mathcal{P} \times \mathcal{D}$, is called an ϵ-solution of (P) and (D) if $\mathbf{Tr}\,(XS) \le \epsilon$. Assume that a pair $(X,\ S) \in \mathrm{ri}\,(\mathcal{P} \times \mathcal{D})$ is given that is 'close to' $(X\,(\mu)\,,\ S\,(\mu))$, for some $\mu = \mu^0$. Then one finds an ϵ-solution after $\mathrm{O}\big(\sqrt{n}\,\log \frac{n\mu^0}{\epsilon}\big)$ iterations of the algorithm in Fig. 1.

The following lemmas are crucial in the analysis of the algorithm. We recall them without proof. They describe the effect of a μ-update and of a full NT step on $\delta\,(X,\ S,\ \mu)$.

Lemma 3.2 (Lemma 7.5 in [9]) *Let* $(X,\ S) \in \mathrm{ri}\,(\mathcal{P} \times \mathcal{D})$, $n\mu = \mathbf{Tr}\,(XS)$ *and* $\delta := \delta\,(X,\ S,\ \mu)$. *If* $\mu^+ = (1-\theta)\,\mu$ *for* $0 < \theta < 1$, *then one has*

$$\delta\,(X,\ S,\ \mu)^2 = \frac{n\theta^2}{4\,(1-\theta)} + (1-\theta)\,\delta^2.$$

Lemma 3.3 (Lemma 7.4 in [9]) *If* $\delta := \delta\,(X,\ S,\ \mu) \le 1$, *then the full NT step is feasible, i.e.,* X^+ *and* S^+ *are feasible. Moreover, if* $\delta < 1$, *then*

$$\delta\left(X^+,\ S^+,\ \mu\right) \le \frac{\delta^2}{\sqrt{2\left(1-\delta^2\right)}}.$$

Corollary 3.4 (Corollary 7.2 in [9]) *If* $\delta\,(X,\ S,\ \mu) < \frac{1}{\sqrt{2}}$, *then* $\delta\left(X^+,\ S^+,\ \mu\right) < \delta\,(X,\ S,\ \mu)^2$.

Primal-Dual Feasible IPM

Input:
Accuracy parameter $\varepsilon > 0$;
barrier update parameter θ, $0 < \theta < 1$;
feasible pair $\left(X^0,\ S^0\right)$ and $\mu^0 > 0$ such that $\delta\left(X^0,\ S^0,\ \mu^0\right) \leq \frac{1}{2}$.

begin
$X := X^0;\ S := S^0;\ \mu := \mu^0;$
while **Tr** $(XS) \geq \varepsilon$ **do**
begin
update of μ:
$\mu := (1-\theta)\mu;$
centering step:
$(X,\ S) := (X,\ S) + (\Delta X,\ \Delta S);$
end
end

Fig. 1 Feasible full-Newton-step algorithm

This corollary implies that the Newton-process is quadratically convergent when started close to the μ-center.

The following result establishes a polynomial iteration bound of the above described algorithm; it easily follows from the above lemmas.

Theorem 3.5 (Theorem 7.1 in [9]) *If* $\theta = \frac{1}{2\sqrt{n}}$, *then the algorithm requires at most*

$$2\sqrt{n}\log\frac{n\mu^0}{\varepsilon}$$

iterations. The output is a primal-dual pair $(X,\ S)$ *such that* **Tr** $(XS) \leq \varepsilon$.

4 Infeasible full-Newton step IPM

In this section we present an infeasible-start interior-point algorithm that generates an ε-solution (X^*, y^*, S^*) of (P) and (D), or establishes that no such solution exists with vanishing duality gap and such that the eigenvalues of X^* and S^* do not exceed a prescribed number ζ.

4.1 The perturbed problems

We start with choosing arbitrarily $X^0 \succ 0$ and y^0, $S^0 \succ 0$ such that $X^0 S^0 = \mu^0 I$ for some (positive) number μ^0. For any ν with $0 < \nu \leq 1$ we consider the perturbed problem (P_ν), defined by

$$(P_\nu) \quad \min \left\{ \left(C - \nu \left(C - \sum_{i=1}^{m} y_i^0 A_i - S^0 \right) \right) \bullet X : A_i \bullet X \right.$$
$$\left. = b_i - \nu \left(b_i - A_i \bullet X^0 \right),\ X \succeq 0 \right\},$$

and its dual problem (D_ν), which is given by

$$(D_\nu) \max \left\{ \sum_{i=1}^{m} \left(b_i - \nu \left(b_i - A_i \bullet X^0 \right) \right) y_i : \sum_{i=1}^{m} y_i A_i + S \right.$$
$$\left. = C - \nu \left(C - \sum_{i=1}^{m} y_i^0 A_i - S^0 \right),\ S \succeq 0 \right\}.$$

Note that if $\nu = 1$ then $X = X^0$ yields a strictly feasible solution of (P_ν), and $(y, S) = (y^0, S^0)$ a strictly feasible solution of (D_ν). We conclude that if $\nu = 1$ then (P_ν) and (D_ν) are strictly feasible.

Lemma 4.1 *Let the original problems,* (P) *and* (D)*, be feasible. Then for each* ν *satisfying* $0 < \nu \leq 1$ *the perturbed problems* (P_ν) *and* (D_ν) *are strictly feasible.*

Proof Suppose that (P) and (D) are feasible. Let $\bar{X}$ be feasible solution of (P) and $(\bar{y}, \bar{S})$ a feasible solution of (D). Then $A_i \bullet \bar{X} = b_i$ $i = 1, \ldots, m$ and $\sum_{i=1}^{m} \bar{y}_i A_i + \bar{S} = C$, with $\bar{X} \succeq 0$ and $\bar{S} \succeq 0$. Now let $0 < \nu \leq 1$, and consider

$$X = (1-\nu)\ \bar{X} + \nu X^0, \quad y = (1-\nu)\ \bar{y} + \nu y^0, \quad S = (1-\nu)\ \bar{S} + \nu S^0.$$

One has for all $i = 1, \ldots, m$ that,

$$\begin{aligned} A_i \bullet X &= A_i \bullet \left((1-\nu)\ \bar{X} + \nu X^0 \right) \\ &= (1-\nu)\, A_i \bullet \bar{X} + \nu A_i \bullet X^0 \\ &= b_i - \nu \left(b_i - A_i \bullet X^0 \right), \end{aligned}$$

showing that X is feasible for (P_ν). Similarly,

$$\begin{aligned} \sum_{i=1}^{m} y_i A_i + S &= (1-\nu) \left(\sum_{i=1}^{m} \bar{y}_i A_i + \bar{S} \right) + \nu \left(\sum_{i=1}^{m} y_i^0 A_i + S^0 \right) \\ &= (1-\nu)\, C + \nu \left(\sum_{i=1}^{m} y_i^0 A_i + S^0 \right) = C - \nu \left(C - \sum_{i=1}^{m} y_i^0 A_i - S^0 \right), \end{aligned}$$

showing that (y, S) is feasible for (D_ν). Since $\nu > 0$, X and S are positive definite, thus this proves that (P_ν) and (D_ν) are strictly feasible. □

4.2 Central path of the perturbed problems

We assume that (P) and (D) are feasible. Letting $0 < \nu \leq 1$, Lemma 4.1 implies that the problems (P_ν) and (D_ν) are strictly feasible, and hence their central paths exist. This means that the system

$$b_i - A_i \bullet X = \nu \left(b_i - A_i \bullet X^0\right), \quad i = 1, 2, \ldots, m, \qquad X \succeq 0 \tag{11}$$

$$C - \sum_{i=1}^{m} y_i A_i - S = \nu \left(C - \sum_{i=1}^{m} y_i^0 A_i - S^0\right), \quad S \succeq 0$$

$$XS = \mu I. \tag{12}$$

has a unique solution, for every $\mu > 0$. In the sequel this unique solution is denoted as $(X(\mu,\, \nu),\, y(\mu,\, \nu),\, S(\mu,\, \nu))$. These are the μ-centers of the perturbed problems (P_ν) and (D_ν).

Note that since $X^0 S^0 = \mu^0 I$, X^0 is the μ^0-center of the perturbed problem (P_1) and (y^0, S^0) the μ^0-center of (D_1). In other words, $(X(\mu^0, 1),\, y(\mu^0, 1),\, S(\mu^0, 1)) = (X^0, y^0, S^0)$. In the sequel we will always have $\mu = \nu\, \mu^0$, and we will accordingly denote $(X(\mu,\, \nu),\, y(\mu,\, \nu),\, S(\mu,\, \nu))$ simply as $(X(\nu),\, y(\nu),\, S(\nu))$.

4.3 An iteration of our algorithm

We just established that if $\nu = 1$ and $\mu = \mu^0$, then $X = X^0$ and $(y, S) = \left(y^0, S^0\right)$ are the μ-center of (P_ν) and (D_ν) respectively. These are our initial iterates.

We measure proximity to the μ-center of the perturbed problems by the quantity $\delta\,(X,\, S,\, \mu)$ as defined in (9). So, initially we have $\delta\,(X,\, S,\, \mu) = 0$. In the sequel we assume that at the start of each iteration, just before the μ-update, $\delta\,(X,\, S,\, \mu)$ is smaller than or equal to a (small) threshold value $\tau > 0$. So this is certainly true at the start of the first iteration.

Now we describe one iteration of our algorithm. Suppose that for some $\mu \in (0,\, \mu^0]$ we have $X,\; y$ and S satisfying the feasibility conditions (11) and (12) for $\nu = \frac{\mu}{\mu^0}$, and such that $\mathbf{Tr}\,(XS) = n\mu$ and $\delta\,(X,\; S,\; \mu) \leq \tau$. We reduce μ to $\mu^+ = (1 - \theta)\,\mu$, with $\theta \in (0,\; 1)$, and find new iterates $X^+,\; y^+$ and S^+ that satisfy (11) and (12), with μ replaced by μ^+ and ν by $\nu^+ = \frac{\mu^+}{\mu^0}$, and such that $\mathbf{Tr}\,(XS) = n\mu^+$ and $\delta\left(X^+,\; S^+,\; \mu^+\right) \leq \tau$. Note that $\nu^+ = (1 - \theta)\,\nu$.

To be more precise, this is achieved as follows. Each main iteration consists of a feasibility step and a few centering steps. The feasibility step serves to get iterates $\left(X^f,\; y^f,\; S^f\right)$ that are strictly feasible for (P_{ν^+}) and (D_{ν^+}), and moreover belong to the region of quadratic convergence of their μ^+-centers $\left(X\left(\nu^+\right), y\left(\nu^+\right), S\left(\nu^+\right)\right)$, in other words $\delta\left(X^f, S^f, \mu^+\right) \leq \frac{1}{\sqrt{2}}$. Since the triple $\left(X^f, y^f, S^f\right)$ is strictly feasible for (P_{ν^+}) and (D_{ν^+}), we can perform a few centering steps starting at $\left(X^f, y^f, S^f\right)$ targeting at the μ^+-centers of (P_{ν^+}) and (D_{ν^+}) and obtain iterates $\left(X^+, y^+, S^+\right)$ that are feasible for (P_{ν^+}) and (D_{ν^+}) and such that $\delta\left(X^+, S^+, \mu^+\right) \leq \tau$.

Before describing the feasibility step it will be convenient to introduce some new notations. We denote the initial values of the primal and dual residuals as r_b^0 and R_c^0, respectively:

$$\left(r_b^0\right)_i = b_i - A_i \bullet X^0, \quad i = 1, \ldots, m, \tag{13}$$

$$R_c^0 = C - \sum_{i=1}^m y_i^0 A_i - S^0. \tag{14}$$

Using these notations the feasibility conditions for (P_ν) and (D_ν) are

$$A_i \bullet X = b_i - \nu\left(r_b^0\right)_i, \quad i = 1, \ldots, m, \quad X \succeq 0,$$
$$\sum_{i=1}^m y_i A_i + S = C - \nu R_c^0, \quad S \succeq 0$$

and for (P_{ν^+}) and (D_{ν^+}) the feasibility conditions are

$$A_i \bullet X = b_i - \nu^+\left(r_b^0\right)_i, \quad i = 1, \ldots, m, \quad X \succeq 0,$$
$$\sum_{i=1}^m y_i A_i + S = C - \nu^+ R_c^0, \quad S \succeq 0.$$

Now suppose that (X, y, S) is feasible for (P_ν) and (D_ν). For finding iterates that are feasible for (P_ν) and (D_ν) we need search directions $\Delta^f X$, $\Delta^f y$ and $\Delta^f S$ such that

$$A_i \bullet \left(X + \Delta^f X\right) = b_i - \nu^+\left(r_b^0\right)_i, \quad i = 1, \ldots, m,$$
$$\sum_{i=1}^m \left(y_i + \Delta^f y\right) A_i + \left(S + \Delta^f S\right) = C - \nu^+ R_c^0.$$

Since X and (y, S) are feasible for (P_ν) and (D_ν) respectively, it follows that $\Delta^f X$, $\Delta^f y$ and $\Delta^f S$ should satisfy

$$A_i \bullet \Delta^f X = (b_i - A_i \bullet X) - \nu^+\left(r_b^0\right)_i = \nu\left(r_b^0\right)_i - \nu^+\left(r_b^0\right)_i = \theta\nu\left(r_b^0\right)_i,$$
$$\sum_{i=1}^m \Delta^f y_i A_i + \Delta^f S = \left(C - \sum_{i=1}^m y_i A_i - S\right) - \nu^+ R_c^0 = \nu R_c^0 - \nu^+ R_c^0 = \theta\nu R_c^0.$$

Therefore, the following system is used to define $\Delta^f X,\ \Delta^f y$ and $\Delta^f S$:

$$A_i \bullet \Delta^f X = \theta\nu\left(r_b^0\right)_i, \quad i = 1, \ldots, m, \tag{15}$$

$$\sum_{i=1}^{m} \Delta^f y_i A_i + \Delta^f S = \theta\nu R_c^0, \tag{16}$$

$$\Delta^f X + P\Delta^f S P^T = \mu S^{-1} - X, \tag{17}$$

where we used the NT-'trick' to symmetrize $\Delta^f X$ with P as defined in (5). After the feasibility step the iterates are given by

$$X^f = X + \Delta^f X, \tag{18}$$
$$y^f = y + \Delta^f y,$$
$$S^f = S + \Delta^f S. \tag{19}$$

By definition, after the feasibility step the iterates satisfy the affine equations in (11) and (12), with $\nu = \nu^+$. The hard part in the analysis will be to guarantee that X^f and S^f are positive definite and satisfy $\delta\left(X^f,\ S^f,\ \mu^+\right) \leq \frac{1}{\sqrt{2}}$.

After the feasibility step we perform centering steps in order to get iterates $\left(X^+,\ y^+,\ S^+\right)$ that satisfy $\mathbf{Tr}\left(X^+S^+\right) = n\mu^+$ and $\delta\left(X^+,\ S^+,\ \mu^+\right) \leq \tau$. By using Corollary 3.4, the required number of centering steps can easily be obtained. Indeed, assuming $\delta = \delta\left(X^f,\ S^f,\ \mu^+\right) \leq \frac{1}{\sqrt{2}}$, after k centering steps we will have iterates $\left(X^+,\ y^+,\ S^+\right)$ that are still feasible for (P_{ν^+}) and (D_{ν^+}) that satisfy

$$\delta\left(X^+,\ S^+,\ \mu^+\right) \leq \left(\frac{1}{\sqrt{2}}\right)^{2^k}.$$

Just as in the linear case [45] this implies that after at most

$$\log_2\left(\log_2\frac{1}{\tau^2}\right) \tag{20}$$

centering steps we have $\delta\left(X^+,\ S^+,\ \mu^+\right) \leq \tau$.

4.4 The algorithm

A formal description of the algorithm is given in Fig. 2, where r_b and R_c denote the primal and dual residuals, respectively. One may easily verify after each iteration the residuals and the duality gap are reduced by a factor $1-\theta$. The algorithm stops if the norms of residual vectors and the duality gap are less than the accuracy parameter ε.

Primal-Dual Infeasible IPM

Input:
Accuracy parameter $\varepsilon > 0$;
barrier update parameter $\theta,\ 0 < \theta < 1$;
threshold parameter $\tau,\ 0 < \tau \leq \frac{1}{\sqrt{2}}$;
$X^0 \succ 0,\ S^0 \succ 0,\ y^0 = 0$ and $\mu^0 > 0$ such that $X^0 S^0 = \mu^0 I$.

begin
$X := X^0,\ S := S^0,\ y := y^0;\ \mu := \mu^0$;
while $\max\left(\mathbf{Tr}\left(XS\right),\ \|r_b\|,\ \|R_c\|\right) \geq \varepsilon$ **do**
begin
feasibility step:
$(X,\ y,\ S) := (X,\ y,\ S) + \left(\Delta^f X,\ \Delta^f y,\ \Delta^f S\right)$;
μ-update:
$\mu := (1-\theta)\mu$;
centering steps:
while $\delta\left(X,\ S,\ \mu\right) \geq \tau$ **do**
begin
$(X,\ y,\ S) := (X,\ y,\ S) + (\Delta X,\ \Delta y,\ \Delta S)$;
end
end
end

Fig. 2 Infeasible full-Newton-step algorithm

5 An analysis of the algorithm

Let $X,\ y$ and S denote the iterates at the start of an iteration with $\mathbf{Tr}\left(XS\right) = n\mu$ and $\delta\left(X,\ S,\ \mu\right) \leq \tau$. Recall that at the start of first iteration this is certainly true, because $\mathbf{Tr}\left(X^0 S^0\right) = n\mu^0$ and $\delta\left(X^0,\ S^0,\ \mu^0\right) = 0$.

Before dealing with the analysis of the algorithm we recall some lemmas that will be needed.

Lemma 5.1 (Lemma A.1 in [9]) *Let $Q \in \mathbf{S}^n_{++}$, and let $M \in \mathbf{R}^{n\times n}$ be skew-symmetric (i.e., $M = -M^T$). Then $\det\left(Q + M\right) > 0$. Moreover, if the eigenvalues of $Q + M$ are real then*

$$0 < \lambda_{\min}\left(Q\right) \leq \lambda_{\min}\left(Q + M\right) \leq \lambda_{\max}\left(Q + M\right) \leq \lambda_{\max}\left(Q\right).$$

Lemma 5.2 (Lemma 1.2.4 in [14]) *Let $A,\ B \in \mathbf{S}^n_+$. Then we have the following inequalities:*

$$\lambda_{\min}(A)\,\lambda_{\max}(B) \le \lambda_{\min}(A)\,\mathbf{Tr}(B) \le \mathbf{Tr}(AB) \\ \le \lambda_{\max}(A)\,\mathbf{Tr}(B) \le n\lambda_{\max}(A)\,\lambda_{\max}(B).$$

Lemma 5.3 (Theorem A.4 in [9]) *Let $A \in \mathbf{S}^n_{++}$ and $B \in \mathbf{S}^n_{++}$. Then all the eigenvalues of AB are real and positive.*

5.1 The effect of the feasibility step and the choice of θ

As we established in Section 4.3, the feasibility step generates new iterates X^f, y^f and S^f that satisfy the feasibility equations for (P_{ν^+}) and (D_{ν^+}). A crucial element in the analysis is to show that after the feasibility step $\delta\left(X^f,\ S^f,\ \mu^+\right) \le \frac{1}{\sqrt{2}}$, i.e., that the new iterates are within the region where the Newton process targeting at the μ^+-centers of (P_{ν^+}) and (D_{ν^+}) is quadratically convergent.

We define

$$D^f_X := \frac{1}{\sqrt{\mu}} D^{-1}\Delta^f X D^{-1}, \qquad D^f_S := \frac{1}{\sqrt{\mu}} D\Delta^f S D,$$
$$\left(V^f\right)^2 := \frac{1}{\mu} D^{-1} X^f S^f D, \tag{21}$$

with D as defined in Section 3.2. We can now rewrite (15), (16) and (17) as follows.

$$DA_iD \bullet D^f_X = \frac{1}{\sqrt{\mu}}\theta\nu\left(r^0_b\right)_i, \quad i = 1,\ \ldots,\ m,$$
$$\sum_{i=1}^m \frac{\Delta y_i}{\sqrt{\mu}} DA_iD + D^f_S = \frac{1}{\sqrt{\mu}}\theta\nu D R^0_c D,$$
$$D^f_X + D^f_S = V^{-1} - V. \tag{22}$$

From the third equation in (22) we obtain, by multiplying both side from the left with V,

$$VD^f_X + VD^f_S = I - V^2. \tag{23}$$

Using (6), (18), (19) and (21), we obtain

$$X^f = X + \Delta^f X = \sqrt{\mu} D\left(V + D^f_X\right) D,$$
$$S^f = S + \Delta^f S = \sqrt{\mu} D^{-1}\left(V + D^f_S\right) D^{-1}.$$

Therefore

$$X^f S^f = \mu D\left(V + D^f_X\right)\left(V + D^f_S\right) D^{-1}.$$

The last matrix is similar to $\mu(V + D_X^f)(V + D_S^f)$. Thus we have

$$X^f S^f \sim \mu \left(V + D_X^f\right) \left(V + D_S^f\right).$$

To simplify the notation in the sequel we introduce

$$D_{XS}^f := \frac{1}{2} \left(D_X^f D_S^f + D_S^f D_X^f\right), \tag{24}$$

and

$$M := \left(D_X^f V - V D_S^f\right) + \frac{1}{2} \left(D_X^f D_S^f - D_S^f D_X^f\right). \tag{25}$$

Note that D_{XS}^f is symmetric and M is skew-symmetric.

Now we may write, using (23),

$$\begin{aligned} \left(V + D_X^f\right) \left(V + D_S^f\right) &= V^2 + V D_S^f + D_X^f V + D_X^f D_S^f \\ &= I - V D_X^f + D_X^f V + D_X^f D_S^f. \end{aligned}$$

By subtracting and adding $\frac{1}{2} D_S^f D_X^f$ to the last expression we get

$$\begin{aligned} \left(V + D_X^f\right) \left(V + D_S^f\right) &= I + \frac{1}{2} \left(D_X^f D_S^f + D_S^f D_X^f\right) + \left(D_X^f V - V D_X^f\right) \\ &\quad + \frac{1}{2} \left(D_X^f D_S^f - D_S^f D_X^f\right), \\ &= I + D_{XS}^f + M. \end{aligned}$$

Using (24) and (25) we obtain

$$X^f S^f \sim \mu \left(I + D_{XS}^f + M\right). \tag{26}$$

Lemma 5.4 *Let* $X \succ 0$ *and* $S \succ 0$*. Then the iterates* $\left(X^f, y^f, S^f\right)$ *are strictly feasible if*

$$I + D_{XS}^f \succ 0.$$

Proof For the proof we introduce a step length $\alpha \in [0, 1]$, and we define

$$X^\alpha = X + \alpha \Delta^f X, \quad y^\alpha = y + \alpha \Delta^f y, \quad S^\alpha = S + \alpha \Delta^f S.$$

We then have $X^0 = X$, $X^1 = X^f$ and similar relations for y and S. Obviously $\det\left(X^0 S^0\right) = \det\left(\mu^0 I\right) = \left(\mu^0\right)^n > 0$. Our aim is to show that the determinant of $X^\alpha S^\alpha$ remains positive for all $\alpha \le 1$. We may write

$$\begin{aligned} \frac{X^\alpha S^\alpha}{\mu} &\sim \left(V + \alpha D_X^f\right) \left(V + \alpha D_S^f\right) \\ &= \left(V^2 + \alpha \left(D_X^f V + V D_S^f\right) + \alpha^2 D_X^f D_S^f\right) \\ &= \left(V^2 + \alpha \left(V D_X^f + V D_S^f\right) + \alpha \left(D_X^f V - V D_X^f\right) + \alpha^2 D_X^f D_S^f\right). \end{aligned}$$

Using (23) we get

$$\frac{X^{\alpha}S^{\alpha}}{\mu} = \left(V^2 + \alpha\left(I - V^2\right) + \alpha\left(D_X^f V - V D_X^f\right) + \alpha^2 D_X^f D_S^f\right).$$

By subtracting and adding $\alpha^2 I$, $\alpha^2\left(D_X^f V - V D_X^f\right)$ and $\frac{1}{2}\left(D_X^f D_S^f + D_S^f D_X^f\right)$ to the right hand side of the above equality we obtain

$$\frac{X^{\alpha}S^{\alpha}}{\mu} = \left((1-\alpha)V^2 + \alpha(1-\alpha)I + \alpha^2\left(D_{XS}^f + I\right)\right)$$
$$+\alpha\left((1-\alpha)\left(D_X^f V - V D_X^f\right) + \alpha M\right).$$

The matrix $(1-\alpha)\left(D_X^f V - V D_X^f\right) + \alpha M$ is skew-symmetric for $\alpha \in [0, 1]$. Lemma 5.1 therefore implies that the determinant of $X^{\alpha}S^{\alpha}$ will be positive if the symmetric matrix

$$(1-\alpha)V^2 + \alpha(1-\alpha)I + \alpha^2\left(D_{XS}^f + I\right)$$

is positive definite. The latter is true for $0 \le \alpha \le 1$. So $X^{\alpha}S^{\alpha}$ has positive determinant for $\alpha \in [0, 1]$. This implies that X^{α} and S^{α} are nonsingular for $\alpha \in [0, 1]$. Since X^0 and S^0 are positive definite and since X^{α} and S^{α} depend continuously on α, it follows that X^1 and S^1 are positive definite as well. This completes the proof. □

Corollary 5.5 *The iterates $\left(X^f, y^f, S^f\right)$ are certainly strictly feasible if*

$$\left|\lambda_i\left(D_{XS}^f\right)\right| < 1, \textit{ for } \quad i = 1, \ldots, n.$$

Proof By Lemma 5.4, X^f and S^f are strictly feasible if $I + D_{XS}^f \succ 0$. Since the last inequality certainly holds if $\left|\lambda_i\left(D_{XS}^f\right)\right| < 1$, for $i = 1, \ldots, n$, the corollary follows. □

In the sequel we denote

$$\omega(V) := \frac{1}{2}\sqrt{\left\|D_X^f\right\|^2 + \left\|D_S^f\right\|^2}, \tag{27}$$

which implies $\left\|D_X^f\right\| \le 2\omega(V)$ and $\left\|D_S^f\right\| \le 2\omega(V)$. Here we recall from [29] two properties of the Frobenius matrix norm. One has for any symmetric $n \times n$ matrix A:

$$\|A\|^2 = \mathbf{Tr}\left(A^2\right) = \sum_{i=1}^{n}\lambda_i\left(A^2\right) = \sum_{i=1}^{n}\lambda_i(A)^2.$$

This norm is also sub-multiplicative, i.e., for any two square matrices A and B,

$$\|AB\| \le \|A\|\,\|B\|.$$

By using these properties and (27) we have

$$\left\| D^f_{XS} \right\| \leq \left\| D^f_X \right\| \left\| D^f_S \right\| \leq \frac{1}{2} \left(\left\| D^f_X \right\|^2 + \left\| D^f_S \right\|^2 \right) = 2\omega(V)^2, \tag{28}$$

$$\left| \lambda_i \left(D^f_{XS} \right) \right| \leq \left\| D^f_{XS} \right\| \leq 2\omega(V)^2, \qquad i = 1, \ldots, n. \tag{29}$$

Lemma 5.6 *If* $\omega(V) < \frac{1}{\sqrt{2}}$ *then the iterates* $\left(X^f, y^f, S^f\right)$ *are strictly feasible.*

Proof Let $\omega(V) < \frac{1}{\sqrt{2}}$. Then (29) implies that $\left|\lambda_i\left(D^f_{XS}\right)\right| < 1$, for $i = 1, \ldots, n$. By Corollary 5.5 this implies the statement in the lemma. □

Assuming $\omega(V) < \frac{1}{\sqrt{2}}$, which guarantees strict feasibility of the iterates $\left(X^f, y^f, S^f\right)$, we proceed by deriving an upper bound for $\delta\left(X^f, S^f, \mu^+\right)$. Recall from definition (10) that

$$\delta\left(X^f, S^f, \mu^+\right) = \frac{1}{2} \left\| \left(V^f\right)^{-1} - V^f \right\|, \tag{30}$$

with $\left(V^f\right)^2$ as defined in (21). In the sequel we denote $\delta\left(X^f, S^f, \mu^+\right)$ also by shortly by $\delta\left(V^f\right)$. We proceed to find an upper bound for $\delta\left(V^f\right)$ in terms of $\omega(V)$. To this end we need some technical results which give information on the eigenvalues and the norm of V^f.

Lemma 5.7 *One has*

$$\lambda_{\min}\left(\left(V^f\right)^2\right) \geq \frac{1}{1-\theta}\left(1 - 2\omega(V)^2\right).$$

Proof Using (26), after division of both sides by $\mu^+ = (1-\theta)\mu$ we get

$$\left(V^f\right)^2 \sim \frac{\mu\left(I + D^f_{XS} + M\right)}{\mu^+} = \frac{I + D^f_{XS} + M}{1-\theta}. \tag{31}$$

It follows that

$$\lambda_{\min}\left(\left(V^f\right)^2\right) = \frac{1}{1-\theta}\lambda_{\min}\left(I + D^f_{XS} + M\right).$$

Since M is skew-symmetric, Lemma 5.1 implies

$$\begin{aligned}\lambda_{\min}\left(\left(V^f\right)^2\right) &\geq \frac{1}{1-\theta}\lambda_{\min}\left(I + D^f_{XS}\right)\\ &= \frac{1}{1-\theta}\left(1 + \lambda_{\min}\left(D^f_{XS}\right)\right).\end{aligned}$$

Substitution of the bound for $\left|\lambda_{\min}\left(D^f_{XS}\right)\right|$ in (28) yields

$$\lambda_{\min}\left(\left(V^f\right)^2\right) \geq \frac{1}{1-\theta}\left(1 - 2\omega(V)^2\right),$$

which completes the proof. □

Lemma 5.8 *One has*

$$\left\| I - \left(V^f\right)^2 \right\| \leq \frac{2\omega\,(V)^2 + \theta\sqrt{n}}{1-\theta}.$$

Proof Using (31) and properties of the Frobenius norm we have

$$\begin{aligned}
\left\| I - (V^f)^2 \right\|^2 &= \sum_{i=1}^{n} \left(\frac{\lambda_i\Big(I + D^f_{XS} + M\Big)}{1-\theta} - 1 \right)^2 \\
&= \frac{1}{(1-\theta)^2} \sum_{i=1}^{n} \left(\lambda_i \left(I + D^f_{XS} + M\right) - 1 + \theta\right)^2 \\
&= \frac{1}{(1-\theta)^2} \sum_{i=1}^{n} \left(\lambda_i \left(D^f_{XS} + M\right) + \theta\right)^2 \\
&= \frac{1}{(1-\theta)^2} \left(n\theta^2 + \sum_{i=1}^{n} \left(\lambda_i \left(D^f_{XS} + M\right)\right)^2 + 2\theta \sum_{i=1}^{n} \lambda_i \left(D^f_{XS} + M\right) \right)
\end{aligned}$$

Since $\left(\lambda_i(D^f_{XS} + M)\right)^2 = \lambda_i\left((D^f_{XS} + M)^2\right)$, for each i, we obtain

$$\left\| I - (V^f)^2 \right\|^2 = \frac{1}{(1-\theta)^2} \left(n\theta^2 + \mathbf{Tr}\left(\left(D^f_{XS} + M\right)^2\right) + 2\theta\,\mathbf{Tr}\left(D^f_{XS} + M\right) \right). \tag{32}$$

Using the skew-symmetry of M we obtain $\mathbf{Tr}\left(D^f_{XS} + M\right) = \mathbf{Tr}\left(D^f_{XS}\right)$ and

$$\mathbf{Tr}\left(\left(D^f_{XS} + M\right)^2\right) = \mathbf{Tr}\left(\left(D^f_{XS}\right)^2 + MD^f_{XS} + D^f_{XS}M - MM^T\right).$$

Since $MD^f_{XS} + D^f_{XS}M$ is skew-symmetric we obtain

$$\mathbf{Tr}\left(\left(D^f_{XS} + M\right)^2\right) = \mathbf{Tr}\left(\left(D^f_{XS}\right)^2 - MM^T\right) \leq \mathbf{Tr}\left(\left(D^f_{XS}\right)^2\right) = \left\| D^f_{XS} \right\|^2,$$

where the inequality follows since the matrix MM^T is positive semidefinite. Substitution in (32) gives

$$\left\| I - (V^f)^2 \right\|^2 \leq \frac{1}{(1-\theta)^2} \left(n\theta^2 + \left\| D^f_{XS} \right\|^2 + 2\theta\,\mathbf{Tr}\left(D^f_{XS}\right) \right).$$

Now let $\lambda\left(D^f_{XS}\right)$ be the vector consisting of the eigenvalues of D^f_{XS}. Using the Cauchy-Schwartz inequality and (28) we get

$$\mathbf{Tr}\left(D^f_{XS}\right) = \sum_{i=1}^{n} \lambda_i\left(D^f_{XS}\right) = e^T\lambda\left(D^f_{XS}\right) \leq \|e\| \left\|\lambda\left(D^f_{XS}\right)\right\|$$
$$= \|e\| \left\|D^f_{XS}\right\| \leq 2\sqrt{n}\,\omega\left(V\right)^2.$$

Substitution gives, also using (28),

$$\left\|I - \left(V^f\right)^2\right\|^2 \leq \frac{1}{(1-\theta)^2}\left(n\theta^2 + 4\omega\left(V\right)^4 + 4\theta\sqrt{n}\,\omega\left(V\right)^2\right) = \left(\frac{2\omega\left(V\right)^2 + \theta\sqrt{n}}{1-\theta}\right)^2.$$

which implies the lemma. □

Lemma 5.9 *Let $\omega\left(V\right) < \frac{1}{\sqrt{2}}$. Then one has*

$$2\delta\left(V^f\right) \leq \frac{2\,\omega\left(V\right)^2 + \theta\,\sqrt{n}}{\sqrt{(1-\theta)\left(1 - 2\omega\left(V\right)^2\right)}}.$$

Proof We may write, using (30),

$$2\delta\left(V^f\right) = \left\|V^f - \left(V^f\right)^{-1}\right\| = \left\|\left(V^f\right)^{-1}\left(I - \left(V^f\right)^2\right)\right\|$$
$$\leq \lambda_{\max}\left(\left(V^f\right)^{-1}\right)\left\|I - \left(V^f\right)^2\right\| = \frac{1}{\lambda_{\min}\left(\left(V^f\right)\right)}\left\|I - \left(V^f\right)^2\right\|.$$

Using the bounds in Lemma 5.7 and Lemma 5.8 the lemma follows. □

Recall from Section 4.3 that we need to have $\delta\left(V^f\right) \leq \frac{1}{\sqrt{2}}$. By Lemma 5.9 it suffices for this that

$$\frac{2\,\omega\left(V\right)^2 + \theta\,\sqrt{n}}{\sqrt{(1-\theta)\left(1 - 2\omega\left(V\right)^2\right)}} \leq \sqrt{2}. \tag{33}$$

Lemma 5.10 *Let $\omega\left(V\right) \leq \frac{1}{2}$ and*

$$\theta = \frac{\alpha}{2\left(\sqrt{n}+1\right)}, \qquad 0 \leq \alpha \leq 1. \tag{34}$$

Then the iterates $\left(X^f,\ y^f,\ S^f\right)$ are strictly feasible and $\delta\left(V^f\right) \leq \frac{1}{\sqrt{2}}$.

Proof Due to Lemma 5.6 and $\omega\left(V\right) \leq \frac{1}{2}$, the iterates $\left(X^f,\ y^f,\ S^f\right)$ are strictly feasible. We just established that if inequality (33) is satisfied then

$\delta\left(V^f\right) \leq 1/\sqrt{2}$. The left hand side in (33) is monotonically increasing in $\omega(V)$. By substituting $\omega(V) = \frac{1}{2}$, the inequality (33) reduces to

$$\frac{\frac{1}{2} + \theta\sqrt{n}}{\sqrt{\frac{1}{2}(1-\theta)}} \leq \sqrt{2},$$

which is equivalent to

$$4n\theta^2 + 4\left(\sqrt{n}+1\right)\theta - 3 \leq 0.$$

The largest possible value of θ satisfying this inequality is given by

$$\begin{aligned}\theta &= \frac{3}{2\left(\sqrt{n}+1+\sqrt{\left(\sqrt{n}+1\right)^2+3n}\right)} \\ &\geq \frac{3}{2\left(\sqrt{n}+1+\sqrt{\left(\sqrt{n}+1\right)^2+3\left(\sqrt{n}+1\right)^2}\right)} = \frac{1}{2\left(\sqrt{n}+1\right)},\end{aligned}$$

which is in agreement with (34). Thus the lemma has been proved. □

5.2 An upper bound for $\omega(V)$

As became clear in (22), the system (15)–(17), which defines the search directions $\Delta^f X$, $\Delta^f y$ and $\Delta^f S$, can be expressed in terms of scaled search directions D_X^f and D_S^f. We define the linear space $\mathcal{L}$ as follows:

$$\mathcal{L} := \left\{\xi \in \mathbf{S}^n \ : \ DA_iD \bullet \xi = 0, \quad i = 1, \ldots, m\right\}.$$

Using the linear space $\mathcal{L}$, it is clear from the first equation in (22) that the affine space

$$\left\{\xi \in \mathbf{S}^n \ : \ DA_iD \bullet \xi = \frac{1}{\sqrt{\mu}}\theta\nu\left(r_b^0\right)_i, \ i = 1, \ldots, m\right\}$$

equals $D_X^f + \mathcal{L}$. By the second equation in system (22), we have $D_S^f \in \frac{1}{\sqrt{\mu}}\theta\nu DR_c^0D + \mathcal{L}^\perp$. Since $\mathcal{L} \cap \mathcal{L}^\perp = \{0\}$, the spaces $D_X^f + \mathcal{L}$ and $D_S^f + \mathcal{L}^\perp$ meet in a unique matrix. This matrix is denoted below by Q.

Lemma 5.11 *Let Q be the (unique) matrix in the intersection of the affine spaces $D_X + \mathcal{L}$ and $D_S + \mathcal{L}^\perp$. Then*

$$2\omega(V) \leq \sqrt{\|Q\|^2 + \left(\|Q\|^2 + 2\delta(V)\right)^2}.$$

Proof The proof is similar to the proof of Lemma 5.6 in [45], and is therefore omitted. □

From Lemma 5.10 we know that we want to have $\omega(V) \le \frac{1}{2}$ because then $\delta\left(V^+\right) \le \frac{1}{\sqrt{2}}$. Due to Lemma 5.11 this will hold if $\|Q\|$ satisfies

$$\|Q\| + (\|Q\| + 2\delta(V))^2 \le 1. \tag{35}$$

5.3 An upper bound for $\|Q\|$

Recall from Lemma 5.11 that Q is the unique solution of the system

$$DA_iD \bullet Q = \frac{1}{\sqrt{\mu}}\theta\nu\left(r_b^0\right)_i, \quad i = 1, \ldots, m,$$

$$\sum_{i=1}^{m} \frac{\xi_i}{\sqrt{\mu}} DA_iD + Q = \frac{1}{\sqrt{\mu}}\theta\nu DR_c^0 D. \tag{36}$$

We proceed to finding an upper bound for $\|Q\|$. As will become clear below, especially in the proofs of Lemma 5.12 and Lemma 5.15, it will be convenient to choose the initial iterates $\left(X^0,\ y^0,\ S^0\right)$ as follows:

$$X^0 = S^0 = \zeta I, \quad y^0 = 0, \quad \mu^0 = \zeta^2, \tag{37}$$

where $\zeta > 0$ is such that

$$X^* + S^* \preceq \zeta I, \tag{38}$$

for some $(X^*,\ y^*,\ S^*) \in \mathcal{F}^*$. It may be noted that this choice of the initial iterates has become usual for infeasible IPMs for SDO. See, e.g., [22, 42, 44].

For the moment, let us write

$$(r_b)_i = \theta\nu\left(r_b^0\right)_i, \quad i = 1, 2, \ldots, m, \quad R_c = \theta\nu R_c^0,$$

and let r_b be the vector $((r_b)_1;\ (r_b)_2;\ \ldots;\ (r_b)_m)$. For any two matrices E (of size $m \times n$) and F (of size $p \times q$) the Kronecker product $E \otimes F$ is the $mp \times nq$ block matrix

$$E \otimes F = \begin{bmatrix} E_{11}F & \cdots & E_{1n}F \\ \vdots & \ddots & \vdots \\ E_{m1}F & \cdots & E_{mn}F \end{bmatrix}.$$

We recall from [29] some properties of Kronecker product and the operator $\mathbf{vec}(\cdot)$ that are useful for our purpose. These properties are

(a) $(E \otimes F)^T = E^T \otimes F^T$.

(b) If E and F are square and nonsingular, then

$$(E \otimes F)^{-1} = E^{-1} \otimes F^{-1}.$$

(c) For any $E\,(m \times n)$, $F\,(n \times r)$ and $H\,(r \times s)$, we have

$$\mathbf{vec}(EHF) = \left(F^T \otimes E\right)\mathbf{vec}(H).$$

By using these properties and the definition of the inner product of two matrices the system (36) can be rewritten as follows:

$$\mathbf{vec}\,(A_i)^T\,(D \otimes D)\,\mathbf{vec}\,(Q) = \frac{1}{\sqrt{\mu}}\,(r_b)_i\,, \quad i = 1,\, \ldots,\, m,$$

$$\sum_{i=1}^{m} \frac{\xi_i}{\sqrt{\mu}}\,(D \otimes D)\,\mathbf{vec}\,(A_i) + \mathbf{vec}\,(Q) = \frac{1}{\sqrt{\mu}}\,(D \otimes D)\,\mathbf{vec}\,(R_c)\,. \tag{39}$$

Let $\mathcal{A}^T = \left[\mathbf{vec}\,(A_1)\;\; \mathbf{vec}\,(A_2)\;\; \ldots\;\; \mathbf{vec}\,(A_m)\right]$ and $\xi = (\xi_1;\; \xi_2;\; \ldots;\; \xi_m)$. One may easily verify that we can rewrite the system (39) as follows:

$$\mathcal{A}\,(D \otimes D)\,\mathbf{vec}\,(Q) = \frac{1}{\sqrt{\mu}} r_b\,,$$

$$(D \otimes D)\,\mathcal{A}^T \frac{\xi}{\sqrt{\mu}} + \mathbf{vec}\,(Q) = \frac{1}{\sqrt{\mu}}\,(D \otimes D)\,\mathbf{vec}\,(R_c)\,, \tag{40}$$

Lemma 5.12 *With* $\left(X^0,\, y^0,\, S^0\right)$ *as defined in* (37) *and* (38), *we have*

$$\|Q\| \leq \theta\sqrt{\nu\,\mathbf{Tr}\left(P^2 + P^{-2}\right)}. \tag{41}$$

Proof Replacing $\mathcal{A}$, $D \otimes D$ and $\mathbf{vec}\,(Q)$ in system (40) by A, D and q, respectively, yields exactly the same system as in the proof of Lemma 5.7 in [45]. By using similar arguments as there, we obtain the following result:

$$\sqrt{\mu}\,\|\mathbf{vec}\,(Q)\| \leq \theta\nu\sqrt{\left\|\mathbf{vec}\left(D\left(S^0 - \bar{S}\right)D\right)\right\|^2 + \left\|\mathbf{vec}\left(D^{-1}\left(X^0 - \bar{X}\right)D^{-1}\right)\right\|^2},$$

where $\bar{X}$, $\bar{y}$ and $\bar{S}$ satisfy

$$\mathcal{A}\mathbf{vec}\left(\bar{X}\right) = b\,,$$

$$\mathcal{A}^T\bar{y} + \mathbf{vec}\left(\bar{S}\right) = \mathbf{vec}\,(C)\,. \tag{42}$$

Using $\|\mathbf{vec}\,(U)\| = \|U\|$ for any matrix U, we obtain

$$\sqrt{\mu}\,\|Q\| \leq \theta\nu\sqrt{\left\|\left(D\left(S^0 - \bar{S}\right)D\right)\right\|^2 + \left\|D^{-1}\left(X^0 - \bar{X}\right)D^{-1}\right\|^2}. \tag{43}$$

We are still free to choose $\bar{X}$ and $\bar{S}$, such that (42) is satisfied. We use $\bar{X} = X^*$ and $\bar{S} = S^*$, with X^* and S^* as in (38). Then we have

$$0 \preceq X^0 - \bar{X} = X^0 - X^* \preceq \zeta I, \quad 0 \preceq S^0 - \bar{S} = S^0 - S^* \preceq \zeta I.$$

It follows that

$$\left\|D\left(S^0 - \bar{S}\right)D\right\|^2 \leq \zeta^2\left\|D^2\right\|^2 = \zeta^2\,\|P\|^2 = \zeta^2\mathbf{Tr}\left(P^2\right),$$

where we used Lemma 5.2 and $D = P^{\frac{1}{2}}$. In the same way it follows that

$$\left\|D^{-1}\left(X^0 - \bar{X}\right)D^{-1}\right\|^2 \leq \zeta^2\mathbf{Tr}\left(P^{-2}\right).$$

Substituting the last inequalities and $\mu = \nu\mu^0 = \nu\zeta^2$ into (43) gives

$$\|Q\| \leq \theta \sqrt{\nu\,\mathbf{Tr}\left(P^2 + P^{-2}\right)},$$

proving the lemma. □

Lemma 5.13 *With* $\left(X^0,\ y^0,\ S^0\right)$ *as defined in* (37) *and* (38), *we have*

$$\|Q\| \leq \frac{\theta}{\zeta\,\lambda_{\min}(V)} \mathbf{Tr}\,(X + S)\,. \tag{44}$$

Proof Using (5) and Lemma 5.2 we have

$$\begin{aligned}
\mathbf{Tr}\left(P^2\right) &= \mathbf{Tr}\left(X^{\frac12}\left(X^{\frac12}SX^{\frac12}\right)^{\frac{-1}{2}} X \left(X^{\frac12}SX^{\frac12}\right)^{\frac{-1}{2}} X^{\frac12}\right)\\
&= \mathbf{Tr}\left(X\left(X^{\frac12}SX^{\frac12}\right)^{\frac{-1}{2}} X \left(X^{\frac12}SX^{\frac12}\right)^{\frac{-1}{2}}\right)\\
&\leq \frac{1}{\lambda_{\min}\left(\left(X^{\frac12}SX^{\frac12}\right)^{\frac12}\right)} \mathbf{Tr}\left(X^2\left(X^{\frac12}SX^{\frac12}\right)^{\frac{-1}{2}}\right)\\
&\leq \frac{\mathbf{Tr}\left(X^2\right)}{\lambda_{\min}\left(X^{\frac12}SX^{\frac12}\right)} = \frac{\mathbf{Tr}\left(X^2\right)}{\mu\lambda_{\min}\left(V^2\right)},
\end{aligned}$$

where for the last equality we used $V^2 \sim \frac{X^{\frac12}SX^{\frac12}}{\mu} \sim \frac{XS}{\mu}$. In the same way we get

$$\mathbf{Tr}\left(P^{-2}\right) \leq \frac{\mathbf{Tr}\left(S^2\right)}{\mu\lambda_{\min}\left(V^2\right)}.$$

Thus we obtain

$$\sqrt{\mathbf{Tr}\left(P^2 + P^{-2}\right)} \leq \frac{1}{\lambda_{\min}(V)} \sqrt{\frac{\mathbf{Tr}\left(X^2 + S^2\right)}{\mu}}. \tag{45}$$

Moreover, by the positive definiteness of X and S and Lemma 5.3 it follows that

$$\mathbf{Tr}\left(X^2 + S^2\right) \leq \mathbf{Tr}\left(X^2 + S^2 + XS + SX\right) = \mathbf{Tr}\left((X+S)^2\right) \leq \mathbf{Tr}\,(X+S)^2\,. \tag{46}$$

Substituting (45) and (46) in (41) gives

$$\|Q\| \leq \frac{\theta}{\lambda_{\min}(V)} \sqrt{\frac{\nu}{\mu}\mathbf{Tr}\,(X+S)^2}.$$

Since $\mu = \nu\mu^0 = \nu\zeta^2$, the lemma follows. □

5.4 Some bounds for $\mathbf{Tr}\,(X+S)$ and $\lambda_{\min}(V)$. The choice of τ and α

Let X be feasible for (P_ν) and $(y,\ S)$ for (D_ν). We need to find an upper bound for $\mathbf{Tr}\,(X+S)$ and lower bound on the eigenvalues of V as defined in (6). We can rewrite $\delta(V)$ in (10) as follows:

$$\begin{aligned}
4\delta(V)^2 &= \left\|V - V^{-1}\right\|^2 \\
&= \mathbf{Tr}\left(\left(V - V^{-1}\right)^T\left(V - V^{-1}\right)\right) \\
&= \mathbf{Tr}\left(V^2 - 2I + V^{-2}\right) \\
&= \sum_{i=1}^{n}\left(\lambda_i(V)^2 - 2 + \frac{1}{\lambda_i(V)^2}\right) \\
&= \sum_{i=1}^{n}\left(\lambda_i(V) - \frac{1}{\lambda_i(V)}\right)^2.
\end{aligned} \tag{47}$$

Using this one easily derives the following result, which we state without further proof.

Lemma 5.14 (Cf. Lemma II.60 in [46]) *Let $\delta = \delta(V)$ be given by (47). Then*

$$\frac{1}{\rho(\delta)} \le \lambda_i(V) \le \rho(\delta), \tag{48}$$

where

$$\rho(\delta) := \delta + \sqrt{1+\delta^2}. \tag{49}$$

Lemma 5.15 *Let X and $(y,\ S)$ be feasible for the perturbed problems (P_ν) and (D_ν) respectively and let $\left(X^0,\ y^0,\ S^0\right)$ and $(X^*,\ y^*,\ S^*) \in \mathcal{F}^*$ be as defined in (37) and (38). Then we have*

$$\nu\zeta\,\mathbf{Tr}\,(X+S) \le S \bullet X + \nu n\zeta^2.$$

Proof Let

$$\begin{aligned}
X' &= X - \nu X^0 - (1-\nu)X^*, \\
y' &= y - \nu y^0 - (1-\nu)y^*, \\
S' &= S - \nu S^0 - (1-\nu)S^*.
\end{aligned}$$

From (13), (14) and definition of perturbed problems (P_ν) and (D_ν), it is easily seen that $\left(X',\ y',\ S'\right)$ satisfies

$$\begin{aligned}
A_i \bullet X' &= 0, \quad i = 1,\ \ldots,\ m, \\
\sum_{i=1}^{m} y_i' A_i + S' &= 0.
\end{aligned}$$

This implies $X' \bullet S' = 0$, i.e.,

$$\left(X - \nu X^0 - (1-\nu)\, X^*\right) \bullet \left(S - \nu S^0 - (1-\nu)\, S^*\right) = 0.$$

By expanding the last equality and using the fact that $X^* \bullet S^* = 0$ we obtain

$$\nu\left(S^0 \bullet X + X^0 \bullet S\right) = S \bullet X + \nu^2 S^0 \bullet X^0 + \nu\,(1-\nu)\left(S^0 \bullet X^* + X^0 \bullet S^*\right) - (1-\nu)\left(S \bullet X^* + S^* \bullet X\right).$$

Since $\left(X^0,\ y^0,\ S^0\right)$ are as in (37) we have

$$S^0 \bullet X + X^0 \bullet S = \zeta\, \mathbf{Tr}\,(X+S), \qquad S^0 \bullet X^0 = n\zeta^2,$$
$$S^0 \bullet X^* + X^0 \bullet S^* = \zeta\, \mathbf{Tr}\left(X^* + S^*\right).$$

Due to (38) we have $\mathbf{Tr}\,(X^* + S^*) \leq n\zeta$. Substitution gives

$$\begin{aligned}
\nu\zeta\,\mathbf{Tr}\,(X+S) &= S \bullet X + \nu^2 n\zeta^2 + \nu\,(1-\nu)\,\zeta\,\mathbf{Tr}\left(X^* + S^*\right) - (1-\nu)\left(S \bullet X^* + S^* \bullet X\right)\\
&\leq S \bullet X + \nu^2 n\zeta^2 + \nu\,(1-\nu)\, n\zeta^2 - (1-\nu)\left(S \bullet X^* + S^* \bullet X\right)\\
&= S \bullet X + \nu n\zeta^2 - (1-\nu)\left(S \bullet X^* + S^* \bullet X\right)\\
&\leq S \bullet X + \nu n\zeta^2,
\end{aligned}$$

where the last inequality is due to the fact $S \bullet X^* + S^* \bullet X \geq 0$. Hence the proof is complete. □

Lemma 5.16 *Using the same notations as in Lemma* 5.15*, one has*

$$\mathbf{Tr}\,(X+S) \leq \left(\rho\,(\delta)^2 + 1\right) n\zeta, \tag{50}$$

where $\rho\,(\delta)$ *as defined in* (49).

Proof Dividing both sides of the inequality in Lemma 5.15 by $\nu\zeta$, while using that $\mu = \nu\zeta^2$, we get

$$Tr\,(X+S) \leq \mathbf{Tr}\left(\frac{SX}{\mu}\right)\zeta + n\zeta.$$

Hence it suffices for the proof if we show that

$$\mathbf{Tr}\left(\frac{SX}{\mu}\right) \leq n\rho(\delta)^2.$$

Since $SX \sim \mu V^2$, which is immediately clear from (7), the left-hand side equals $\mathbf{Tr}(V^2)$. So we can rewrite the last inequality as

$$\sum_{i=1}^{n} \lambda_i(V)^2 \leq n\rho(\delta)^2.$$

But this inequality is an immediate consequence of Lemma 5.14. Hence the proof is complete. □

By substituting (48) and (50) into (44) we get

$$\|Q\| \leq n\theta\rho(\delta)\left(1+\rho(\delta)^2\right).$$

At this stage we choose

$$\tau = \frac{1}{8}. \tag{51}$$

Since $\delta \leq \tau = \frac{1}{8}$ and $\rho(\delta)$ is monotonically increasing in δ, we have

$$\|Q\| \leq n\theta\rho(\delta)\left(1+\rho(\delta)^2\right) \leq n\theta\rho\left(\frac{1}{8}\right)\left(1+\rho\left(\frac{1}{8}\right)^2\right) = 2.586\,n\,\theta.$$

By using $\theta = \frac{\alpha}{2(\sqrt{n}+1)}$ (see Lemma 5.10) we obtain the following upper bound for the norm of Q:

$$\|Q\| \leq \frac{2.586\,n\,\alpha}{2\left(\sqrt{n}+1\right)}. \tag{52}$$

In (35) we found that in order to have $\delta\left(V^f\right) \leq \frac{1}{\sqrt{2}}$, we should have $\|Q\|^2 + (\|Q\| + 2\delta(V))^2 \leq 1$. Therefore, since $\delta(V) \leq \tau = \frac{1}{8}$, it suffices if Q satisfies $\|Q\|^2 + \left(\|Q\| + \frac{1}{4}\right)^2 \leq 1$. The latter holds if $\|Q\| \leq 0.57097$. Hence, using (52) we obtain that $\delta\left(V^f\right) \leq \frac{1}{\sqrt{2}}$ certainly holds if

$$\frac{2.586\,n\,\alpha}{2\left(\sqrt{n}+1\right)} \leq 0.57097.$$

From this we deduce that by taking

$$\alpha = \frac{2\left(\sqrt{n}+1\right)}{5\,n}, \tag{53}$$

it is guaranteed that $\delta\left(V^f\right) \leq \frac{1}{\sqrt{2}}$.

5.5 Complexity

In the previous sections we have found that if at the start of an iteration the iterates satisfy $\delta(X, S; \mu) \leq \tau$, with τ as defined in (51), and θ as in (34), and with taking α as in (53), then after the feasibility step the iterates satisfy $\delta(X, S; \mu^+) \leq 1/\sqrt{2}$.

According to (20), at most

$$\log_2\left(\log_2 \frac{1}{\tau^2}\right) = \log_2\left(\log_2 64\right) \leq 3$$

centering steps suffice to get iterates that satisfy $\delta(X, S; \mu^+) \leq \tau$. So each iteration consists of one feasibility step and at most 3 centering steps. In each iteration both the duality gap and the norms of the residual vectors are reduced by the factor $1-\theta$. Hence, using $X^0 \bullet S^0 = n\zeta^2$, the total number of iterations is bounded above by

$$\frac{1}{\theta} \log \frac{\max\left\{n\zeta^2,\ \left\|r_b^0\right\|,\ \left\|R_c^0\right\|\right\}}{\varepsilon}.$$

Due to (34) and (53) we have

$$\theta = \frac{\alpha}{2\left(\sqrt{n}+1\right)} = \frac{1}{5n}.$$

Hence the total number of inner iterations is bounded above by

$$20n \log \frac{\max\left\{n\zeta^2,\ \left\|r_b^0\right\|,\ \left\|R_c^0\right\|\right\}}{\varepsilon}.$$

Note that the order of this bound is the same as the bound in [30, 45] for LO. We may state without further proof our main result.

Theorem 5.17 *If* (P) *and* (D) *have optimal solutions* $(X^*,\ y^*,\ S^*) \in \mathcal{F}^*$ *such that* $X^* + S^* \preceq \zeta I$*, then after at most*

$$20n \log \frac{\max\left\{n\zeta^2,\ \left\|r_b^0\right\|,\ \left\|R_c^0\right\|\right\}}{\varepsilon}.$$

iterations the algorithm finds an ε*-solution of* (P) *and* (D)*.*

The above theorem gives a convergence result under the assumption that (P) and (D) have optimal solutions (X^*, y^*, S^*) with zero duality gap and such that the eigenvalues of X^* and S^* do not exceed ζ. One might ask what happens if this condition is not satisfied.

Our analysis of the algorithm has made clear that as long as we have $\delta(X, S; \mu^+) \leq 1/\sqrt{2}$ after each feasibility step then the algorithm will generate an ε-solution of (P) and (D), and the number of iterations will be as given in Theorem 5.17. So, if during the execution of the algorithm it happens that after the feasibility step $\delta(X, S; \mu^+) > 1/\sqrt{2}$, then we must conclude that there exists no optimal solutions $(X^*,\ y^*,\ S^*)$ with zero duality gap such that the eigenvalues of X^* and S^* do not exceed ζ. In that case one might rerun the algorithm with larger values of ζ. If this does not help, then eventually one should realize that (P) and/or (D) do not have optimal solutions at all, or they have optimal solutions with positive duality gap.

6 Concluding remarks

We extended the full-Newton infeasible interior-point algorithm for LO as developed in [45], to SDO. We obtained an iteration bound of the same order as in the LO case [30, 45]. See also [31].

Concerning the practical performance of our algorithm, one should realize that it is a common feature of IPMs for LO and SDO with the best known iteration bounds that their practical performance is close to their theoretical performance. In practice this means that the practical performance must be increased to become competitive with existing solvers. For this several options are available. One option is the use of large-update schemes for the barrier parameter, i.e., the use of larger values of θ. This can be achieved by noting that the value of θ used in this paper guarantees that after the feasibility step one has $\delta(X, S, \mu^+) \leq \frac{1}{\sqrt{2}}$, but this value is a rather pessimistic estimate that indeed will work in all cases. In most iterations, however, much larger values are possible, while still keeping $\delta(X, S, \mu^+) \leq \frac{1}{\sqrt{2}}$. Numerical experiments show that when maximizing θ with respect to this property, the number of iterations reduces drastically and the algorithm becomes really competitive. This goes without worsening the theoretical iteration bound.

Some other interesting topics remain for further research. First, the search direction used in this paper is based on the NT-symmetrization scheme and it is natural to ask if other symmetrization schemes can be used. Second, similar as we discussed for the linear case in [30] we might replace the equation $\Delta^f X + P\Delta^f SP = \mu I - XS$ in the system of the feasibility step by either

$$\Delta^f X + P\Delta^f SP^T = 0 \tag{54}$$

or

$$\Delta^f X + P\Delta^f SP^T = (1-\theta)\,\mu S^{-1} - X.$$

In [30] we used (54) which made the analysis of the feasibility step for LO easier than in [45]. For the SDO case we leave it to the future to analyze a full-Newton step method based on these search directions, but it seems unlikely that this will lead to a better iteration bound.

References

1. Alizadeh, F.: Combinatorial optimization with interior point methods and semi-definite matrices. Ph.D. thesis, University of Minnesota, Minneapolis (1991)
2. Alizadeh, F.: Interior point methods in semidefinite programming with applications to combinatorial optimization. SIAM J. Optim. **5**(1), 13–51 (1995)
3. Bai, Y.Q., El Ghami, M., Roos, C.: A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization. SIAM J. Optim. **15**(1), 101–128 (2004)
4. Bai, Y.Q., Roos, C.: A primal-dual interior-point method based on a new kernel function with linear growth rate. In: Caccetta, L., Rehbock, V. (eds.) Industrial Optimisation. Proceedings of Symposium on Industrial Optimisation and the 9TH Australian Optimisation Day, vol. 1, pp. 15–28. Curtin University of Technology, Australia, Western Australian Centre of Excellence in Industrial Optimisation, Perth (2005)
5. Ben-Tal, A., Nemirovski, A.: Lectures on Modern Convex Optimization, MPS/SIAM Series on Optimization. Analysis, Algorithms, and Engineering Applications. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2001)
6. Boyd, S.E., El Ghaoui, L., Feron, E., Balakrishnan, V.: Linear Matrix Inequalities in System and Control Theory. SIAM Studies in Applied Mathematics. SIAM, Philadelphia, USA (1994)
7. Boyd, S.E., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
8. Boyd, S.E., Vandenberghe, L.: Semidefinite programming. SIAM Rev. **38**(1), 49–96 (1996)

9. De Klerk, E.: Aspects of Semidefinite Programming. Applied Optimization, vol. 65. Kluwer Academic, Dordrecht (2002)
10. De Klerk, E., Roos, C., Terlaky, T.: Initialization in semidefinite programming via a self–dual skew-symmetric embedding. Oper. Res. Lett. **20**, 213–221 (1997)
11. Elabwabi, G.: Topics in global optimization using semidefinite optimization. Ph.D. thesis, Faculty of Mathematics and Computer Science, TU Delft, NL–2628 CD Delft (2007)
12. Haeberley, J.-P.A., Alizadeh, F., Overton, M.L.: Complementarity and nondegeneracy in semidefinite programming. Math. Program. **77**, 111–128 (1997)
13. Goemans, M.X., Williamson, D.P.: Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. J. Assoc. Comput. Mach. **42**(6), 1115–1145 (1995)
14. Helmberg, C.: Semidefinite programming for combinatorial optimization. Technical Report 00-34, Konrad-Zuse-Zentrum für Informationstechink Berlin, Takustraße 7, D-14195 Berlin (2000)
15. Helmberg, C., Rendl, F., Vanderbei, R.J., Wolkowicz, H.: An interior point method for semidefinite programming. SIAM J. Optim. **6**, 342–361 (1994)
16. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991)
17. Jansen, B., Roos, C., Terlaky, T., Vial, J.-Ph.: Primal-dual algorithms for linear programming based on the logarithmic barrier method. J. Optim. Theory Appl. **83**, 1–26 (1994)
18. Jarre, F.: An interior-point method for minimizing the largest eigenvalue of a linear combination of matrices. SIAM J. Control Optim. **31**, 1360–1377 (1993)
19. Karisch, E.S., Rendl, F., Clausen, J.: Solving graph bisection problems with semidefinite programming. INFORMS J. Comput. **12**(3), 177–191 (2000)
20. Kojima, M., Megiddo, N., Mizuno, S.: A primal-dual infeasible-interior-point algorithm for linear programming. Math. Program. **61**, 263–280 (1993)
21. Kojima, M., Mizuno, S., Yoshise, A.: A primal-dual interior point algorithm for linear programming. In: Megiddo, N. (ed.) Progress in Mathematical Programming: Interior Point and Related Methods, pp. 29–47. Springer, New York (1989)
22. Kojima, M., Shida, M., Shindoh, M.: Local convergence of predictor-corrector infeasible interior-point algorithm for SDPs and SDLCPs. Math. Program. Ser. A **80**, 129–160 (1998)
23. Kojima, M., Shindoh, M., Hara, S.: Interior point methods for the monotone linear complementarity problem in symmetric matrices. SIAM J. Optim. **7**, 86–125 (1997)
24. Lasserre, J.B., Prieto-Rumeau, T.: SDP vs. LP relaxations for the moment approach in some performance evaluation problems. Stoch. Models **20**, 439–456 (2004)
25. Lasserre, J.B.: Global optimization with polynomials and the problems of moments. SIAM J. Optim. **11**, 796–817 (2001)
26. Lisser, A., Rendl, F.: Graph partitioning using linear and semidefinite programming. Math. Program. Ser. B **95**(1), 91–101 (2003); ISMP 2000, Part 3 (Atlanta, GA)
27. Luo, Z.Q., Sturm, J.F., Zhang, S.: Conic convex programming and self-dual embedding. Optim. Methods Softw. **14**(3), 196–218 (2000)
28. Lustig, I.J.: Feasibile issues in a primal-dual interior point method for linear programming. Math. Program. **49**, 145–162 (1990/91)
29. Lütkepohl, H.: Handbook of Matrices. Wiley, Chichester (1996)
30. Mansouri, H., Roos, C.: Simplified $O(n)$ infeasible interior-point algorithm for linear optimization using full-Newton step. Optim. Methods Softw. **22**(3), 519–530 (2007)
31. Mansouri, H.: Full-newton step interior-point methods for conic optimization. Ph.D. thesis, Faculty of Mathematics and Computer Science, TU Delft, NL–2628 CD Delft (2008)
32. Megiddo, N.: Pathways to the optimal set in linear programming. In: Megiddo, N. (ed.) Progress in Mathematical Programming: Interior Point and Related Methods, pp. 131–158. Springer, New York (1989)
33. Mizuno, S.: Polynomiality of infeasible interior point algorithms for linear programming. Math. Program. **67**, 109–119 (1994)
34. Mizuno, S., Todd, M.J., Ye, Y.: On adaptive-step primal-dual interior-point algorithms for linear programming. Math. Oper. Res. **18**, 964–981 (1993)
35. Monteiro, R.D.C., Adler, I.: Interior-path following primal-dual algorithms: part I: Linear programming. Math. Program. **44**, 27–41 (1989)

36. Nesterov, Y., Todd, M.J.: Primal-dual interior point methods for self-scaled cones. SIAM J. Optim. **8**(2), 324–364 (1998)
37. Nesterov, Y.E., Nemirovski, A.: Interior Point Polynomial Methods in Convex Programming: Theory and Algorithms. SIAM Publications. SIAM, Philadelphia (1993)
38. Parrilo, P.A.: Semidefinite programming relaxations for semialgebraic problems. Math. Program. **96**, 293–320 (2003)
39. Peña, J., Vera, J., Zuluaga, L.F.: Computing the stability number of a graph via linear and semidefinite programming. SIAM J. Optim. **18**(1), 87–105 (2007) (electronic)
40. Peng, J., Roos, C., Terlaky, T.: Self-regular functions and new search directions for linear and semidefinite optimization. Math. Program. Ser. A **93**(1), 129–171 (2002)
41. Peng, J., Roos, C., Terlaky, T.: Self-Regularity. A New Paradigm for Primal-Dual Interior-Point Algorithms. Princeton University Press, Princeton (2002)
42. Potra, F.A.: A quadratically convergent predictor-corrector method for solving linear programs from infeasible starting points. Math. Program. Ser. A **67**(3), 383–406 (1994)
43. Potra, F.A.: An infeasible-interior-point predictor-corrector algorithm for linear programming. SIAM J. Optim. **6**(1), 19–32 (1996)
44. Potra, F.A., Sheng, R.: A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. SIAM J. Optim. **8**, 1007–1028 (1998)
45. Roos, C.: A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization. SIAM J. Optim. **16**(4), 1110–1136 (2006)
46. Roos, C., Terlaky, T., Vial, J.-Ph.: Theory and Algorithms for Linear Optimization. An Interior-Point Approach. Wiley, Chichester (1997)
47. Sturm, J.F., Zhang, S.: Symmetric primal-dual path following algorithms for semidefinite programming. Appl. Numer. Math. **29**(3), 301–315 (1999)
48. Suzuka, A., Ryuhei, M., Yoshise, A., Tomomi, M.: Semidefinite programming based approaches to home-away assignment problems in sports scheduling. Lect. Notes Comput. Sci. **3521**, 95–103 (2005)
49. Tanabe, K.: Centered Newton method for mathematical programming. Syst. Model. Optim. **113**, 197–206 (1988)
50. Tanabe, K.: Centered Newton method for linear programming: interior and 'exterior' point method (in Janpanese). In: Tone, K. (ed.) New Methods for Linear Programming. vol. 3, pp. 98–100. The Institute of Statistical Mathematics, Tokyo (1990)
51. Todd, M.J.: Semidefinite optimization. Acta Numer. **10**, 515–560 (2001)
52. Todd, M.J., Toh, K.C., Tutuncu, R.H.: On the Nesterov-Todd direction in semidefinite programming. SIAM J. Optim. **8**(3), 769–796 (1998)
53. Vandenberghe, L., Boyd, S.: A primal-dual potential reduction method for problems involving matrix inequalities. Math. Program. Ser. B **69**, 205–236 (1995)
54. Wang, G.Q., Bai, Y.Q., Roos, C.: Primal-dual interior-point algorithms for semidefinite optimization based on a simple kernel function. J. Math. Model. Algorithms **4**, 409–433 (2005)
55. Wolkowicz, H., Saigal, R., Vandenberghe, L.: Handbook of Semidefinite Programming. International Series in Operations Research & Management Science, vol. 27. Kluwer Academic, Boston (2000)
56. Ye, Y., Todd, M.J., Mizuno, S.: An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm. Math. Oper. Res. **19**, 53–67 (1994)
57. Zhang, Y.: On the convergence of a class of infeasible-interior-point methods for the horizontal linear complementarity problem. SIAM J. Optim. **4**, 208–227 (1994)
58. Zhang, Y.: On extending primal-dual interior-point algorithms from linear programming to semidefinite programming. SIAM J. Optim. **8**(2), 365–386 (1998)
59. Zhao, Z., Braams, B.J., Fukuda, M., Overton, M.L., Percus, J.K.: The reduced density matrix method for electronic structure calculations and role of three-index representability. J. Chem. Phys. **120**, 2095–2104 (2004)