

A new infeasible interior-point method based on Darvay's technique for symmetric optimization

Behrouz Kheirfam

Published online: 25 October 2013
© Springer Science+Business Media New York 2013

Abstract We present a full Nesterov and Todd step primal-dual infeasible interior-point algorithm for symmetric optimization based on Darvay's technique by using Euclidean Jordan algebras. The search directions are obtained by an equivalent algebraic transformation of the centering equation. The algorithm decreases the duality gap and the feasibility residuals at the same rate. During this algorithm we construct strictly feasible iterates for a sequence of perturbations of the given problem and its dual problem. Each main iteration of the algorithm consists of a feasibility step and some centering steps. The starting point in the first iteration of the algorithm depends on a positive number ξ and it is strictly feasible for a perturbed pair. The 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 to the central path of the new perturbed pair. The algorithm finds an ϵ -optimal solution or detects infeasibility of the given problem. Moreover, we derive the currently best known iteration bound for infeasible interior-point methods.

Keywords Symmetric optimization · Infeasible interior-point methods · Darvay's technique · Euclidean Jordan algebras · Small-update method

1 Introduction

Symmetric optimization (SCO) is to minimize a linear function over the intersection of an affine space and the symmetric cone. Symmetric cones are intimately related to Euclidean Jordan algebras (Faraut and Korányi 1994), and these algebras provide us with a basic toolbox to carry out our analysis. There is an extensive literature on the analysis of interior-point methods (IPMs) for SCO. The basic idea for solving SCO using IPMs is due to Nesterov and Nemirovski (1994). Their method was primarily either primal or dual based. Later on, Nesterov and Todd (1997) proposed interior-point algorithms on a special class of cones called self-scaled cones. The first work connecting Jordan algebras and optimization is due

B. Kheirfam (✉)
Department of Mathematics, Azarbaijan Shahid Madani University, Tabriz, Iran
e-mail: b.kheirfam@azaruniv.edu

to Güler (1996). He observed that the family of the self-scaled cones is identical to the set of symmetric cones for which there exists a complete classification theory. Faybusovich (1997) first extended primal-dual IPMs for semidefinite optimization (SDO) to SCO by using Euclidean Jordan algebras. Muramatsu (2002) presented a commutative class of search directions for SCO and analyzed the complexities of primal-dual IPMs for SCO. Rangarajan (2006) proved the polynomial-time convergence of infeasible IPMs (IIPMs) for conic programming over symmetric cones using a wide neighborhood of the central path for a commutative family of search directions. Schmieta and Alizadeh (2003) introduced primal-dual IPMs for SCO extensively under the framework of Euclidean Jordan algebra. By constructing strictly feasible iterates for a sequence of the perturbed problems of the given problem and its dual problem, Roos (2006) introduced a full-Newton primal-dual IIPM for linear optimization (LO) based on the classical directions. Kheirfam (2011) extended Roos' algorithm for LO to SDO based on a special self-regular proximity. Gu et al. (2011) extended this algorithm to SCO by using Euclidean Jordan algebra. Darvay (2003) proposed a full-Newton step primal-dual path-following interior-point algorithm for LO. The search direction of his algorithm is introduced by using an algebraic equivalent transformation of the centering equation which define the central path and then applying Newton's method for the new system of equations. Recently, Wang and Bai (2012) generalized Darvay's full-Newton step primal-dual path-following interior-point algorithm for LO to SCO by using Euclidean Jordan algebras.

In this paper, by combining Darvay's and Roos' works we propose a full Nesterov and Todd (NT) step primal-dual infeasible interior-point algorithm for SCO by using Euclidean Jordan algebras. At each iteration, we use only full NT-steps which have the advantage that no line searches are needed. By constructing strictly feasible iterates for a sequence of the perturbed problems of the given problem and its dual problem, our algorithm decreases the duality gap and the feasibility residuals at the same rate. Finally, the order of the iteration bound coincides with the currently best known iteration bound for SCO.

The paper is organized as follows: In Sect. 2, we briefly recall the theory of the Euclidean Jordan algebra and their associated symmetric cones. In Sect. 3, after briefly reviewing the concept of the central path for SCO, we recall the Darvay's technique for SCO, the search directions and the quadratical convergence. The primal-dual infeasible interior-point algorithm for SCO is presented in Sect. 4. In Sect. 5, we analyze the algorithm and derive the currently best known iteration bound for IIPMs. Finally, some conclusions and remarks follow in Sect. 6.

2 Euclidean Jordan algebras and symmetric cones

In this section, we introduce Jordan algebras and symmetric cones as well as some of their basic properties, closely following (Schmieta and Alizadeh 2003). These properties serve as our toolbox for the analysis of IIPMs. For a comprehensive study of Jordan algebras, the reader is referred to Faraut and Korányi (1994).

A Jordan algebra \mathcal{J} is a finite dimensional vector space endowed with a bilinear map $\circ : \mathcal{J} \times \mathcal{J} \rightarrow \mathcal{J}$ if for all $x, y \in \mathcal{J}$, $x \circ y = y \circ x$, and $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$, where $x^2 = x \circ x$. A Jordan algebra (\mathcal{J}, \circ) is called Euclidean if there exists a symmetric positive definite quadratic form Q on \mathcal{J} such that $Q(x \circ y, z) = Q(x, y \circ z)$. A Jordan algebra has an identity element, if there exists a unique element $e \in \mathcal{J}$ such that $x \circ e = e \circ x = x$, for all $x \in \mathcal{J}$. The set $\mathcal{K}(\mathcal{J}) := \{x^2 : x \in \mathcal{J}\}$ is called the cone of squares of the Euclidean Jordan algebra $(\mathcal{J}, \circ, \langle \cdot, \cdot \rangle)$. A cone is symmetric if and only if it is the cone of squares of a Euclidean Jordan algebra (Faraut and Korányi 1994).

An element $c \in \mathcal{J}$ is idempotent if $c \circ c = c$. Two elements x and y are orthogonal if $x \circ y = 0$. An idempotent c is primitive if it is nonzero and can not be expressed by sum of two other nonzero idempotents. A set of primitive idempotents $\{c_1, c_2, \dots, c_k\}$ is called a Jordan frame if $c_i \circ c_j = 0$, for any $i \neq j \in \{1, 2, \dots, k\}$ and $\sum_{i=1}^k c_i = e$. For any $x \in \mathcal{J}$, let r be the smallest positive integer such that $\{e, x, x^2, \dots, x^r\}$ is linearly dependent; r is called the degree of x and is denoted by $\text{deg}(x)$. The rank of \mathcal{J} , denoted by $\text{rank}(\mathcal{J})$, is defined as the maximum of $\text{deg}(x)$ over all $x \in \mathcal{J}$.

Theorem 1 (Theorem III.1.2 in Faraut and Korányi 1994) *Let $(\mathcal{J}, \circ, \langle \cdot, \cdot \rangle)$ be a Euclidean Jordan algebra with $\text{rank}(\mathcal{J}) = r$. Then, for any $x \in \mathcal{J}$, there exists a Jordan frame $\{c_1, c_2, \dots, c_r\}$ and real numbers $\lambda_1(x), \lambda_2(x), \dots, \lambda_r(x)$ such that $x = \sum_{i=1}^r \lambda_i(x)c_i$.*

Every $\lambda_i(x)$ is called an eigenvalue of x . We denote $\lambda_{\min}(x)$ ($\lambda_{\max}(x)$) as the minimal (maximal) eigenvalue of x . We can define the following: the square root, $x^{\frac{1}{2}} := \sum_{i=1}^r \sqrt{\lambda_i(x)}c_i$, wherever all $\lambda_i \geq 0$, the inverse, $x^{-1} := \sum_{i=1}^r \lambda_i(x)^{-1}c_i$, wherever all $\lambda_i \neq 0$, the square $x^2 := \sum_{i=1}^r \lambda_i(x)^2c_i$; the trace $\text{tr}(x) := \sum_{i=1}^r \lambda_i(x)$. If x^{-1} is defined, we call x invertible.

Since “ \circ ” is bilinear map, for every $x \in \mathcal{J}$, a linear operator $L(x)$ can be defined such that $L(x)y = x \circ y$ for all $y \in \mathcal{J}$. In particular, $L(x)e = x$ and $L(x)x = x^2$. For each $x \in \mathcal{J}$, define

$$P(x) := 2L(x)^2 - L(x^2),$$

where, $L(x)^2 = L(x)L(x)$. The map $P(x)$ is called the quadratic representation of x . For any $x, y \in \mathcal{J}$, x and y are said to be operator commutable if $L(x)$ and $L(y)$ commute, i.e., $L(x)L(y) = L(y)L(x)$. In other words, x and y are operator commutable if for all $z \in \mathcal{J}$, $x \circ (y \circ z) = y \circ (x \circ z)$ (Schmieta and Alizadeh 2003). For any $x, y \in \mathcal{J}$, the inner product is defined as $\langle x, y \rangle = \text{tr}(x \circ y)$, and the Frobenius norm of x as follows $\|x\|_F = \sqrt{\langle x, x \rangle} = \sqrt{\text{tr}(x^2)} = \sqrt{\sum_{i=1}^r \lambda_i^2(x)}$. Observe that $\|e\|_F = \sqrt{r}$, since identity element e has eigenvalue 1 with multiplicity r . Furthermore, we have

$$\lambda_{\min}(x) \leq \|x\|_F, \quad \lambda_{\max}(x) \leq \|x\|_F \quad \text{and} \quad |\langle x, y \rangle| \leq \|x\|_F \|y\|_F.$$

Lemma 1 (Lemma 3.2 in Faybusovich 2002) *Let $x, s \in \text{int } \mathcal{K}$. Then, there exists a unique $w \in \text{int } \mathcal{K}$ such that*

$$x = P(w)s.$$

Moreover,

$$w = P(x^{\frac{1}{2}})(P(x^{\frac{1}{2}})s)^{-\frac{1}{2}} [= P(s^{-\frac{1}{2}})(P(s^{\frac{1}{2}})x)^{\frac{1}{2}}],$$

where $\text{int } \mathcal{K}$ denotes the interior of the cone \mathcal{K} .

The point w is called the scaling point of x and s . Hence, there exists $\tilde{v} \in \text{int } \mathcal{K}$ such that

$$\tilde{v} = P(w)^{-\frac{1}{2}}x = P(w)^{\frac{1}{2}}s,$$

which is the so-called NT-scaling of \mathcal{J} .

Let $x, y \in \mathcal{J}$. We say that two elements x and y are similar, as denoted by $x \sim y$, if and only if x and y share the same set of eigenvalues. We have $x \in \mathcal{K}$ if and only if $\lambda_i \geq 0$, for all $i = 1, 2, \dots, r$ and $x \in \text{int } \mathcal{K}$ if and only if $\lambda_i > 0$, for all $i = 1, 2, \dots, r$. We also say x is positive semidefinite (positive definite) if $x \in \mathcal{K}$ ($x \in \text{int } \mathcal{K}$).

Theorem 2 (Theorem III.2.1, Proposition III.2.2 in Faraut and Korányi 1994) *Let \mathcal{J} be a Euclidean Jordan algebra. Then $\mathcal{K}(\mathcal{J})$ is a symmetric cone, and is the set of elements x in \mathcal{J} for which $L(x)$ is positive semidefinite. Furthermore, if x is invertible, then $P(x)\text{int}\mathcal{K}(\mathcal{J}) = \text{int}\mathcal{K}(\mathcal{J})$.*

Lemma 2 (Lemma 30 in Schmieta and Alizadeh 2003) *Let $x, s \in \text{int}\mathcal{K}$. Then*

$$\|P(x)^{\frac{1}{2}}s - e\|_F \leq \|x \circ s - e\|_F.$$

Lemma 3 (Theorem 4 in Sturm 2000) *Let $x, s \in \text{int}\mathcal{K}$. Then*

$$\lambda_{\min}(P(x)^{\frac{1}{2}}s) \geq \lambda_{\min}(x \circ s).$$

Lemma 4 (Lemma 14 in Schmieta and Alizadeh 2003) *Let $x, s \in \mathcal{J}$. Then, the eigenvalues of $x + s$ are bounded as follows:*

$$\lambda_{\min}(x + s) \geq \lambda_{\min}(x) + \lambda_{\min}(s) \geq \lambda_{\min}(x) - \|s\|_F.$$

3 Full NT-step based on Darvay’s technique

In this section, we briefly recall the central path for SCO, and Darvay’s technique extended to SCO by Wang and Bai (2012). Consider the following primal and dual problems:

$$\min \quad \{ \langle c, x \rangle : Ax = b, x \in \mathcal{K} \}, \tag{P}$$

and

$$\max \quad \{ b^T y : A^T y + s = c, s \in \mathcal{K} \}, \tag{D}$$

where, c and the rows of A lie in \mathcal{J} , and $b \in R^m$. Moreover, the rows of A are linearly independent. Throughout the paper, we assume that (P) and (D) satisfy the interior point condition (IPC), i.e., there exist $x^0 \in \text{int}\mathcal{K}$ and $s^0 \in \text{int}\mathcal{K}$ such that $Ax^0 = b$ and $A^T y^0 + s^0 = c$. This can be achieved via the so-called homogeneous self-dual embedding (Sturm 2000). Under the IPC, finding an optimal solution of (P) and (D) is equivalent to solving the following system (Schmieta and Alizadeh 2003; Faybusovich 1997):

$$\begin{aligned} Ax &= b, & x &\in \mathcal{K} \\ A^T y + s &= c, & s &\in \mathcal{K} \\ x \circ s &= 0. \end{aligned} \tag{1}$$

The basic idea of primal-dual IPMs is to replace the third equation in (1), the so-called complementary condition for (P) and (D), by the parameterized equation $x \circ s = \mu e$, with $\mu > 0$. Thus, one may consider

$$\begin{aligned} Ax &= b, & x &\in \mathcal{K} \\ A^T y + s &= c, & s &\in \mathcal{K} \\ x \circ s &= \mu e. \end{aligned} \tag{2}$$

For each $\mu > 0$, the parameterized system (2) has a unique solution $(x(\mu), y(\mu), s(\mu))$, is called the μ -center of (P) and (D), and these solutions form a curve parameterized by μ . This curve is called the central path and most IPMs approximately follow the central path

to reach the optimal set. If $\mu \rightarrow 0$, then the limit of the path exists and yields optimal solutions for (P) and (D) (Vieira 2007).

Similarly to the LO case (Darvay 2003), Wang and Bai (2012) replace the standard centering equation $x \circ s = \mu e$ by $\varphi\left(\frac{x \circ s}{\mu}\right) = \varphi(e)$ where $\varphi(\cdot)$ is the vector-valued function induced by the univariate function $\varphi(t)$. Thus, the system (2) becomes

$$\begin{aligned} Ax &= b, & x &\in \mathcal{K} \\ A^T y + s &= c, & s &\in \mathcal{K} \\ \varphi\left(\frac{x \circ s}{\mu}\right) &= \varphi(e). \end{aligned} \tag{3}$$

Applying Newton’s method to system (3), then using Taylor’s theorem to the third equation, leads to

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ x \circ \Delta s + s \circ \Delta x &= \mu \left(\varphi' \left(\frac{x \circ s}{\mu} \right) \right)^{-1} \circ \left(\varphi(e) - \varphi \left(\frac{x \circ s}{\mu} \right) \right). \end{aligned} \tag{4}$$

Due to the fact that $L(x)L(s) \neq L(s)L(x)$, system (4) does not always have a unique solution in $\text{int}\mathcal{K}$. It is well known that this difficulty can be resolved by applying a scaling scheme. This is given in the following lemma.

Lemma 5 (Lemma 28 in Schmieta and Alizadeh 2003) *Let $u \in \text{int}\mathcal{K}$. Then*

$$x \circ s = \mu e \iff P(u)x \circ P(u)^{-1}s = \mu e.$$

Replacing the third equation of the system (3) by

$$\varphi\left(\frac{P(u)x \circ P(u)^{-1}s}{\mu}\right) = \varphi(e),$$

and applying Newton’s method to the resulting system leads us to the following system

$$\begin{aligned} A\Delta x &= 0, \\ A^T \Delta y + \Delta s &= 0, \\ P(u)x \circ P(u)^{-1}\Delta s + P(u)^{-1}s \circ P(u)\Delta x &= \mu \left(\varphi' \left(\frac{P(u)x \circ P(u)^{-1}s}{\mu} \right) \right)^{-1} \circ \left(\varphi(e) - \varphi \left(\frac{P(u)x \circ P(u)^{-1}s}{\mu} \right) \right). \end{aligned} \tag{5}$$

Let $u = w^{-\frac{1}{2}}$, where w is the NT-scaling point of x and s as defined in Lemma 1. We define

$$v := \frac{P(w)^{-\frac{1}{2}}x}{\sqrt{\mu}} \left[= \frac{P(w)^{\frac{1}{2}}s}{\sqrt{\mu}} \right], \tag{6}$$

and

$$\bar{A} := \sqrt{\mu}AP(w)^{\frac{1}{2}}, \quad d_x := \frac{P(w)^{-\frac{1}{2}}\Delta x}{\sqrt{\mu}}, \quad d_s := \frac{P(w)^{\frac{1}{2}}\Delta s}{\sqrt{\mu}}. \tag{7}$$

This enables us to rewrite the system (5), considering $\varphi(t) = \sqrt{t}$, as follows:

$$\begin{aligned} \bar{A}d_x &= 0, \\ \bar{A}^T \frac{\Delta y}{\mu} + d_s &= 0, \\ d_x + d_s &= 2(e - v) := p_v. \end{aligned} \tag{8}$$

The search directions d_x and d_s are obtained by solving (8) so that Δx and Δs are computed via (7). The new iterate is obtained by taking a full NT-step as follows

$$x^+ := x + \Delta x, \quad s^+ := s + \Delta s.$$

For the analysis of the algorithm, we define a norm-based proximity measure $\sigma(x, s; \mu)$ as follows

$$\sigma(v) := \sigma(x, s; \mu) := \frac{\|P_v\|_F}{2} = \|e - v\|_F. \tag{9}$$

We can conclude that

$$\sigma(v) = 0 \iff v = e \iff d_x = d_s = 0 \iff x \circ s = \mu e. \tag{10}$$

Hence, the value of $\sigma(v)$ can be considered as a measure for the distance between the given triple (x, y, s) and the μ -center.

The following lemma is crucial in the analysis of the algorithm, which states that the Newton process is quadratically convergent. We recall it without proof.

Lemma 6 (Lemma 4.4 in Wang and Bai 2012) *Let $\sigma := \sigma(x, s; \mu) < 1$. Then*

$$\sigma(x^+, s^+; \mu) \leq \frac{\sigma^2}{1 + \sqrt{1 - \sigma^2}}.$$

Thus $\sigma(x^+, s^+; \mu) \leq \sigma^2$, which means the quadratical convergence of the algorithm.

The following lemma provides an iteration bound $O(\sqrt{r} \log \frac{\langle x^0, s^0 \rangle}{\epsilon})$ for the algorithm.

Lemma 7 (Theorem 4.1 in Wang and Bai 2012) *Let $\theta = \frac{1}{2\sqrt{r}}$. Then the algorithm requires at most*

$$O\left(\sqrt{r} \log \frac{\langle x^0, s^0 \rangle}{\epsilon}\right),$$

iterations. The output is a primal-dual pair (x, s) satisfying $\langle x, s \rangle \leq \epsilon$.

4 A full NT-step infeasible IPM

In the case of an infeasible method, we call the triple (x, y, s) an ϵ -solution of (P) and (D) if the norms of the residual vectors $b - Ax$ and $c - A^T y - s$ do not exceed ϵ , and also $tr(x \circ s) \leq \epsilon$. In what follows, we present an infeasible-start algorithm that generates an ϵ -solution of (P) and (D), if it exists, or establishes that no such solution exists.

4.1 The perturbed problems

We assume (P) and (D) have an optimal solution (x^*, y^*, s^*) with $tr(x^* \circ s^*) = 0$. As usual for IIPMs, we start the algorithm with $(x^0, y^0, s^0) = \xi(e, 0, e)$ and $\mu^0 = \xi^2$, where ξ is a positive number such that

$$x^* + s^* \preceq_{\mathcal{K}} \xi e, \tag{11}$$

where the Löwner partial ordering $\succeq_{\mathcal{K}}$ of \mathcal{J} defined by a cone \mathcal{K} is defined by $x \succeq_{\mathcal{K}} s$ if $x - s \in \mathcal{K}$. The initial values of the primal and dual residual vectors are $r_p^0 = b - Ax^0$ and $r_d^0 = c - A^T y^0 - s^0$. In general, $r_p^0 \neq 0$ and $r_d^0 \neq 0$. The iterates generated by the algorithm will be infeasible for (P) and (D), but they will be feasible for perturbed versions of (P) and (D) that we introduce below. For any ν with $0 < \nu \leq 1$, we consider the perturbed problem

$$\min \{ (c - \nu r_d^0, x) : b - Ax = \nu r_p^0, x \in \mathcal{K} \}, \tag{P_\nu}$$

and its dual problem

$$\max \{ (b - \nu r_p^0)^T y : c - A^T y - s = \nu r_d^0, s \in \mathcal{K} \}. \tag{D_\nu}$$

Note that if $\nu = 1$, then $x = x^0$ and $(y, s) = (y^0, s^0)$ yield strictly feasible solutions of (P_ν) and (D_ν) , respectively. We conclude that if $\nu = 1$, both (P_ν) and (D_ν) are strictly feasible, which means that both perturbed problems (P_ν) and (D_ν) satisfy the IPC. It should be mentioned that the problems (P_ν) and (D_ν) for LO have been studied in Freund (1996). More generally, one has the following lemma (Lemma 4.1 in Gu et al. 2011).

Lemma 8 *Let (P) and (D) be feasible and $0 < \nu \leq 1$. Then, the perturbed problems (P_ν) and (D_ν) satisfy the IPC.*

Assuming that (P) and (D) are both feasible, it follows from Lemma 8 that the problems (P_ν) and (D_ν) satisfy the IPC, for each $\nu \in (0, 1]$. Then, their central paths exist, meaning that the system

$$\begin{aligned} b - Ax &= \nu r_p^0, & x &\in \mathcal{K}, \\ c - A^T y - s &= \nu r_d^0, & s &\in \mathcal{K}, \\ x \circ s &= \mu e, \end{aligned} \tag{12}$$

has a unique solution, for any $\mu > 0$. For $\nu \in (0, 1]$ and $\mu = \nu \mu^0 = \nu \xi^2$, we denote this unique solution as $(x(\mu, \nu), y(\mu, \nu), s(\mu, \nu))$, where $x(\mu, \nu)$ is the μ -center of (P_ν) and $(y(\mu, \nu), s(\mu, \nu))$ is the μ -center of (D_ν) . The parameters μ and ν will always be in a one-to-one correspondence, according to $\mu = \nu \mu^0 = \nu \xi^2$. For the sake of simplicity, we denote $x(\mu) = x(\mu, \nu)$, $y(\mu) = y(\mu, \nu)$ and $s(\mu) = s(\mu, \nu)$. By taking $\nu = 1$, one has $(x(1), y(1), s(1)) = (x^0, y^0, s^0) = (\xi e, 0, \xi e)$ and $x^0 \circ s^0 = \mu^0 e$. Hence, x^0 is the μ^0 -center of the perturbed problem (P_1) and (y^0, s^0) is the μ^0 -center of the perturbed problem (D_1) .

Similarly to the feasible case, we replace the standard centering equation $x \circ s = \mu e$ by $\varphi\left(\frac{x \circ s}{\mu}\right) = \varphi(e)$. Then, the system of (12) becomes

$$b - Ax = \nu r_p^0, \quad x \in \mathcal{K}, \tag{13}$$

$$c - A^T y - s = \nu r_d^0, \quad s \in \mathcal{K}, \tag{14}$$

$$\varphi\left(\frac{x \circ s}{\mu}\right) = \varphi(e). \tag{15}$$

Algorithm 1 A full NT-step IIPM based on Darvay's technique.

Input:

accuracy parameter $\epsilon > 0$;
 barrier update parameter θ , $0 < \theta < 1$;
 and threshold parameter $0 < \tau \leq h < 1$ (default $h = \frac{1}{\sqrt{2}}$).

begin

$x := \xi e$; $y := 0$; $s := \xi e$; $\mu := \mu^0 = \xi^2$; $v = 1$;

while $\max(\text{tr}(x \circ s), \|r_p\|_F, \|r_d\|_F) > \epsilon$ **do**

begin

 feasibility step:

$(x, y, s) := (x, y, s) + (\Delta^f x, \Delta^f y, \Delta^f s)$;

 update of μ and v :

$\mu := (1 - \theta)\mu$; $v := (1 - \theta)v$;

 centering step:

while $\sigma(x, s; \mu) > \tau$ **do**

begin

$(x, y, s) := (x, y, s) + (\Delta x, \Delta y, \Delta s)$

end

end

end.

4.2 An iteration of our algorithm

We just established that if $v = 1$ and $\mu = \mu^0$, then (x^0, y^0, s^0) is the μ -center of the problems (P_v) and (D_v) . We measure proximity to the μ -center of the perturbed problems by the quantity $\sigma(x, s; \mu)$ as defined in (9).

Initially, we have $\sigma(x, s; \mu) = 0$. In the sequel, we assume that at the start of each iteration, just before the μ - and v -update, $\sigma(x, s; \mu) \leq \tau$, where τ is a positive threshold value. This certainly holds at the start of the first iteration. Since we then have $\sigma(x, s; \mu) = 0$.

Now, we describe one main iteration of our algorithm. The algorithm begins with an infeasible interior-point (x, y, s) such that (x, y, s) is feasible for the perturbed problems (P_v) and (D_v) , with $\mu = v\mu^0$ and such that $\text{tr}(x \circ s) \leq r\mu$ and $\sigma(x, s; \mu) \leq \tau$. We reduce v to $v^+ = (1 - \theta)v$, with $\theta \in (0, 1)$, and find new iterate (x^+, y^+, s^+) that is feasible for the perturbed problems (P_{v^+}) and (D_{v^+}) , and such that $\sigma(x^+, s^+; \mu^+) \leq \tau$. Every iteration consists of a feasibility step, a μ -update and a few centering steps, respectively. First, we find a new point (x^f, y^f, s^f) which is feasible for the perturbed problems with $v^+ := (1 - \theta)v$. Then, μ is decreased to $\mu^+ := (1 - \theta)\mu$. Generally, there is no guarantee that $\sigma(x^f, s^f; \mu^+) \leq \tau$. So, a few centering steps is applied to produce a new point (x^+, y^+, s^+) such that $\sigma(x^+, s^+; \mu^+) \leq \tau$. This process is repeated until the algorithm terminates. We now summarize the steps of the algorithm as Algorithm 1.

5 Analysis of the feasibility step

First, we describe the feasibility step in details. The analysis will follow in the sequel. Suppose that we have strictly feasible iterate (x, y, s) for (P_v) and (D_v) . This means that (x, y, s) satisfies (13) and (14) with $\mu = v\xi^2$. We need displacements $\Delta^f x$, $\Delta^f y$ and $\Delta^f s$ such that

$$x^f := x + \Delta^f x, \quad y^f := y + \Delta^f y, \quad s^f := s + \Delta^f s, \quad (16)$$

are feasible for (P_{v^+}) and (D_{v^+}) . One may easily verify that (x^f, y^f, s^f) satisfies (13) and (14), with v replaced by v^+ , only if the first two equations in the following system are satisfied.

$$\begin{aligned} A\Delta^f x &= \theta v r_p^0, \\ A^T \Delta^f y + \Delta^f s &= \theta v r_d^0 \\ P(w)^{-\frac{1}{2}} x \circ P(w)^{\frac{1}{2}} \Delta^f s + P(w)^{\frac{1}{2}} s \circ P(w)^{-\frac{1}{2}} \Delta^f x & \\ &= \mu \left(\varphi' \left(\frac{P(w)^{-\frac{1}{2}} x \circ P(w)^{\frac{1}{2}} s}{\mu} \right) \right)^{-1} \circ \left(\varphi(e) - \varphi \left(\frac{P(w)^{-\frac{1}{2}} x \circ P(w)^{\frac{1}{2}} s}{\mu} \right) \right). \end{aligned} \tag{17}$$

The third equation is inspired by the third equation in the system (5) that we used to define search directions for the feasible case.

According to (17), after the feasibility step the iterates satisfy the affine equations in (13) and (14), with v replaced by v^+ . The hard part in the analysis will be to guarantee that x^f, s^f are positive and to guarantee that the new iterate satisfies $\sigma(x^f, s^f; \mu^+) \leq h < 1$.

Let (x, y, s) denote the iterate at the start of an iteration with $tr(x \circ s) \leq r\mu$ and $\sigma(x, s; \mu) \leq \tau$. At the start of the first iteration this is certainly true, because $tr(x^0 \circ s^0) = r\mu^0$ and $\sigma(x^0, s^0; \mu^0) = 0$. Define

$$d_x^f := \frac{1}{\sqrt{\mu}} P(w)^{-\frac{1}{2}} \Delta^f x, \quad d_s^f := \frac{1}{\sqrt{\mu}} P(w)^{\frac{1}{2}} \Delta^f s, \tag{18}$$

where w is the NT-scaling point of x and s . One can easily check that the system (17), which defines the search directions $\Delta^f x, \Delta^f y$ and $\Delta^f s$, considering $\varphi(t) = \sqrt{t}$, can be expressed in terms of the scaled search directions d_x^f and d_s^f as follows

$$\begin{aligned} \bar{A} d_x^f &= \theta v r_p^0, \\ \bar{A}^T \frac{\Delta^f y}{\mu} + d_s^f &= \frac{1}{\sqrt{\mu}} \theta v P(w)^{\frac{1}{2}} r_d^0, \\ d_x^f + d_s^f &= 2(e - v) = p_v, \end{aligned} \tag{19}$$

where $\bar{A} = \sqrt{\mu} A P(w)^{\frac{1}{2}}$. To get the search directions $\Delta^f x$ and $\Delta^f s$ in the original x and s -space we use (18), which gives

$$\Delta^f x = \sqrt{\mu} P(w)^{\frac{1}{2}} d_x^f, \quad \Delta^f s = \sqrt{\mu} P(w)^{-\frac{1}{2}} d_s^f.$$

The new iterates are obtained by taking a full step, as given by (16). Hence, we have

$$\begin{aligned} x^f &= x + \Delta^f x = \sqrt{\mu} P(w)^{\frac{1}{2}} (v + d_x^f), \\ s^f &= s + \Delta^f s = \sqrt{\mu} P(w)^{-\frac{1}{2}} (v + d_s^f). \end{aligned}$$

From the third equation in (19) we derive that

$$\begin{aligned} (v + d_x^f) \circ (v + d_s^f) &= v^2 + v \circ (2(e - v)) + d_x^f \circ d_s^f \\ &= v^2 + v \circ p_v + d_x^f \circ d_s^f. \end{aligned} \tag{20}$$

Moreover, from $p_v = 2(e - v)$, we have

$$v + \frac{1}{2} p_v = e \implies v^2 + v \circ p_v = e - \frac{1}{4} p_v^2. \tag{21}$$

Let us introduce the notation

$$\bar{p}_v := d_x^f - d_s^f. \tag{22}$$

In that case, we have

$$d_x^f \circ d_s^f = \frac{p_v^2 - \bar{p}_v^2}{4}. \tag{23}$$

Using (20), (21) and (23), we get

$$(v + d_x^f) \circ (v + d_s^f) = e - \frac{1}{4}p_v^2 + \frac{1}{4}(p_v^2 - \bar{p}_v^2) = e - \frac{1}{4}\bar{p}_v^2. \tag{24}$$

Lemma 9 (Lemma 4.1 in Wang and Bai 2012) *Let $x(\alpha) = x + \alpha \Delta x$, $s(\alpha) = s + \alpha \Delta s$ for $0 \leq \alpha \leq 1$ with $x, s \in \text{int } \mathcal{K}$ and $x(\alpha) \circ s(\alpha) \in \text{int } \mathcal{K}$ for $\alpha \in [0, \bar{\alpha}]$. Then, $x(\bar{\alpha}) \in \text{int } \mathcal{K}$ and $s(\bar{\alpha}) \in \text{int } \mathcal{K}$.*

Since $P(w)^{\frac{1}{2}}$ and $P(w)^{-\frac{1}{2}}$ are automorphisms of $\text{int } \mathcal{K}$, Theorem 2 implies that x^f and s^f belong to $\text{int } \mathcal{K}$ if and only if $v + d_x^f$ and $v + d_s^f$ belong to $\text{int } \mathcal{K}$, respectively. The following lemma shows the strict feasibility of the full NT-step.

Lemma 10 *Let $x \in \text{int } \mathcal{K}$ and $s \in \text{int } \mathcal{K}$. Then the iterates (x^f, y^f, s^f) are strictly feasible if*

$$\left| \lambda_i \left(\frac{\bar{p}_v}{2} \right) \right| < 1, \quad i = 1, 2, \dots, r.$$

Proof Introduce a step length α with $\alpha \in [0, 1]$ and define

$$v_x(\alpha) = v + \alpha d_x^f, \quad v_s(\alpha) = v + \alpha d_s^f.$$

We thus have

$$\begin{aligned} v_x(\alpha) \circ v_s(\alpha) &= v^2 + \alpha v \circ (d_x^f + d_s^f) + \alpha^2 d_x^f \circ d_s^f \\ &= v^2 + \alpha v \circ p_v + \alpha^2 d_x^f \circ d_s^f \\ &= (1 - \alpha)v^2 + \alpha \left(e - \frac{p_v^2}{4} \right) + \frac{\alpha^2}{4} (p_v^2 - \bar{p}_v^2). \end{aligned}$$

If $|\lambda_i(\frac{\bar{p}_v}{2})| < 1$, then we have

$$0 \prec_{\mathcal{K}} e - \frac{1}{4}\bar{p}_v^2 = e - \frac{1}{4}p_v^2 + \frac{1}{4}(p_v^2 - \bar{p}_v^2),$$

and this implies that

$$\frac{1}{4}(p_v^2 - \bar{p}_v^2) \succ_{\mathcal{K}} -e + \frac{1}{4}p_v^2. \tag{25}$$

Therefore, we get for $0 \leq \alpha \leq 1$

$$\begin{aligned} &(1 - \alpha)v^2 + \alpha \left(e - \frac{p_v^2}{4} \right) + \frac{\alpha^2}{4} (p_v^2 - \bar{p}_v^2) \\ &\succ_{\mathcal{K}} (1 - \alpha)v^2 + \alpha \left(e - \frac{p_v^2}{4} \right) + \alpha^2 \left(-e + \frac{1}{4}p_v^2 \right) \\ &= (1 - \alpha)v^2 + \alpha(1 - \alpha) \left(e - \frac{p_v^2}{4} \right) \\ &= (1 - \alpha)^2 v^2 + 2\alpha(1 - \alpha)v \succeq_{\mathcal{K}} 0. \end{aligned} \tag{26}$$

So $v_x(\alpha) \circ v_s(\alpha) \in \text{int } \mathcal{K}$ for $\alpha \in [0, 1]$. Hence, since $x, s \in \text{int } \mathcal{K}$, Lemma 9 implies that $v_x(1) = v + d_x^f \in \text{int } \mathcal{K}$ and $v_s(1) = v + d_s^f \in \text{int } \mathcal{K}$. This completes the proof. \square

In the sequel, we denote

$$\vartheta(v) := \frac{1}{2} \sqrt{\|d_x^f\|_F^2 + \|d_s^f\|_F^2},$$

which implies $\|d_x^f\|_F \leq 2\vartheta(v)$ and $\|d_s^f\|_F \leq 2\vartheta(v)$. Moreover we have

$$\|d_x^f \circ d_s^f\|_F \leq \frac{1}{2} (\|d_x^f\|_F^2 + \|d_s^f\|_F^2) = 2\vartheta(v)^2, \tag{27}$$

$$|\lambda_i(d_x^f \circ d_s^f)| \leq \|d_x^f \circ d_s^f\|_F \leq 2\vartheta(v)^2, \quad i = 1, 2, \dots, r. \tag{28}$$

We proceed by deriving an upper bound for $\sigma(x^f, s^f, \mu^+)$. Recall from definition (9) that

$$\sigma(x^f, s^f; \mu^+) := \sigma(v^f) = \|e - v^f\|_F, \tag{29}$$

where $v^f := \frac{1}{\sqrt{\mu(1-\theta)}} P(w^f)^{-\frac{1}{2}} x^f [= \frac{1}{\sqrt{\mu(1-\theta)}} P(w^f)^{\frac{1}{2}} s^f]$ and w^f is the scaling point of x^f and s^f .

Lemma 11 (Lemma 4.3 in Gu et al. 2011) *One has*

$$\sqrt{1-\theta} v^f \sim [P(v + d_x^f)^{\frac{1}{2}} (v + d_s^f)]^{\frac{1}{2}}.$$

Lemma 12 *If $|\lambda_i(\frac{\bar{p}_v}{2})| < 1, i = 1, 2, \dots, r$. Then*

$$\sigma(v^f) \leq \frac{\sigma(v)^2 + 2\vartheta(v)^2 + \theta\sqrt{r}}{1 - \theta + \sqrt{(1-\theta)(1 - \sigma(v)^2 - 2\vartheta(v)^2)}}.$$

Proof Using Lemma 11, Lemma 2 and (24) we have

$$\begin{aligned} \sigma(v^f) &= \|e - v^f\|_F \leq \frac{1}{1 + \lambda_{\min}(v^f)} \|e - (v^f)^2\|_F \\ &= \frac{1}{1 + \lambda_{\min}(v^f)} \left\| e - P\left(\frac{v + d_x^f}{\sqrt{1-\theta}}\right)^{\frac{1}{2}} \left(\frac{v + d_s^f}{\sqrt{1-\theta}}\right) \right\|_F \\ &\leq \frac{1}{1 + \lambda_{\min}(v^f)} \left\| e - \left(\frac{v + d_x^f}{\sqrt{1-\theta}}\right) \circ \left(\frac{v + d_s^f}{\sqrt{1-\theta}}\right) \right\|_F \\ &= \frac{1}{1 + \lambda_{\min}(v^f)} \left\| e - \frac{e - \frac{1}{4}\bar{p}_v^2}{1-\theta} \right\|_F \\ &= \frac{1}{(1-\theta)(1 + \lambda_{\min}(v^f))} \left\| -\theta e + \frac{1}{4}\bar{p}_v^2 \right\|_F \\ &\leq \frac{1}{(1-\theta)(1 + \lambda_{\min}(v^f))} \left(\theta\sqrt{r} + \left(\frac{\|\bar{p}_v\|_F}{2}\right)^2 \right). \end{aligned} \tag{30}$$

By Lemma 11, Lemma 3, (20), (21), Lemma 4 and (28) we get

$$\lambda_{\min}((v^f)^2) = \lambda_{\min}\left(P\left(\frac{v + d_x^f}{\sqrt{1-\theta}}\right)^{\frac{1}{2}} \left(\frac{v + d_s^f}{\sqrt{1-\theta}}\right)\right)$$

$$\begin{aligned}
 &\geq \lambda_{\min} \left(\left(\frac{v + d_x^f}{\sqrt{1-\theta}} \right) \circ \left(\frac{v + d_s^f}{\sqrt{1-\theta}} \right) \right) \\
 &= \frac{1}{1-\theta} \lambda_{\min} \left(e - \frac{1}{4} p_v^2 + d_x^f \circ d_s^f \right) \\
 &\geq \frac{1}{1-\theta} \left(1 - \frac{1}{4} \|p_v\|_F^2 + \lambda_{\min}(d_x^f \circ d_s^f) \right) \\
 &\geq \frac{1}{1-\theta} (1 - \sigma(v)^2 - 2\vartheta(v)^2). \tag{31}
 \end{aligned}$$

By substituting (31) into (30) and using (9), the third equation of (19) and (27), we get

$$\begin{aligned}
 \sigma(v^f) &\leq \frac{\sqrt{1-\theta}}{(1-\theta)(\sqrt{1-\sigma(v)^2 - 2\vartheta(v)^2} + \sqrt{1-\theta})} \left(\left(\frac{\|\bar{p}_v\|_F}{2} \right)^2 + \theta\sqrt{r} \right) \\
 &= \frac{\sqrt{1-\theta}}{(1-\theta)(\sqrt{1-\sigma(v)^2 - 2\vartheta(v)^2} + \sqrt{1-\theta})} \left(\frac{1}{4} \|d_x^f - d_s^f\|_F^2 + \theta\sqrt{r} \right) \\
 &= \frac{\frac{1}{4} (\|d_x^f + d_s^f\|_F^2 - 4\langle d_x^f, d_s^f \rangle) + \theta\sqrt{r}}{\sqrt{1-\theta}(\sqrt{1-\sigma(v)^2 - 2\vartheta(v)^2} + \sqrt{1-\theta})} \\
 &= \frac{\frac{1}{4} (\|p_v\|_F^2 - 4\langle d_x^f, d_s^f \rangle) + \theta\sqrt{r}}{\sqrt{1-\theta}(\sqrt{1-\sigma(v)^2 - 2\vartheta(v)^2} + \sqrt{1-\theta})} \\
 &= \frac{\sigma(v)^2 - \langle d_x^f, d_s^f \rangle + \theta\sqrt{r}}{\sqrt{1-\theta}(\sqrt{1-\sigma(v)^2 - 2\vartheta(v)^2} + \sqrt{1-\theta})} \\
 &\leq \frac{\sigma(v)^2 + \|d_x^f\|_F \|d_s^f\|_F + \theta\sqrt{r}}{\sqrt{1-\theta}(\sqrt{1-\sigma(v)^2 - 2\vartheta(v)^2} + \sqrt{1-\theta})} \\
 &\leq \frac{\sigma(v)^2 + \frac{1}{2} (\|d_x^f\|_F^2 + \|d_s^f\|_F^2) + \theta\sqrt{r}}{\sqrt{1-\theta}(\sqrt{1-\sigma(v)^2 - 2\vartheta(v)^2} + \sqrt{1-\theta})} \\
 &= \frac{\sigma(v)^2 + 2\vartheta(v)^2 + \theta\sqrt{r}}{\sqrt{1-\theta}(\sqrt{1-\sigma(v)^2 - 2\vartheta(v)^2} + \sqrt{1-\theta})},
 \end{aligned}$$

this completes the proof. □

Because we need to have $\sigma(v^f) \leq h < 1$ (default $h = \frac{1}{\sqrt{2}}$), it follows from Lemma 12 that it suffices to have

$$\frac{\sigma(v)^2 + 2\vartheta(v)^2 + \theta\sqrt{r}}{1-\theta + \sqrt{(1-\theta)(1-\sigma(v)^2 - 2\vartheta(v)^2)}} \leq \frac{1}{\sqrt{2}} < 1. \tag{32}$$

At this stage, we choose

$$\tau = \frac{1}{16}, \quad \theta = \frac{\alpha}{2\sqrt{r}}, \quad \alpha \leq 1. \tag{33}$$

The left-hand side of (32) is monotonically increasing with respect to $\vartheta(v)^2$, then for $r \geq 1$ and $\sigma(v) \leq \tau$, one can verify that

$$\vartheta(v) \leq \frac{1}{2\sqrt{2}} \Rightarrow \sigma(v^f) \leq \frac{1}{\sqrt{2}} < 1. \tag{34}$$

5.1 Upper bound for $\vartheta(v)$

In this section, we obtain an upper bound for $\vartheta(v)$ which will enable us to find a default value for θ . For this purpose, consider the system (19) which defines the search directions $\Delta^f x$, $\Delta^f y$ and $\Delta^f s$ in terms of the scaled search directions d_x^f and d_s^f . By eliminating d_s^f from (19), we get

$$\begin{aligned} \bar{A}d_x^f &= \theta v r_p^0, \\ -\bar{A}^T \frac{\Delta^f y}{\mu} + d_x^f &= p_v - \frac{1}{\sqrt{\mu}} \theta v P(w)^{\frac{1}{2}} r_d^0. \end{aligned} \tag{35}$$

Multiplying the second equation of (35) from left with \bar{A} and using the first equation of (35) it follows that

$$-\frac{\Delta^f y}{\mu} = (\bar{A}\bar{A}^T)^{-1} \bar{A} \left(p_v - \frac{1}{\sqrt{\mu}} \theta v P(w)^{\frac{1}{2}} r_d^0 \right) - \theta v (\bar{A}\bar{A}^T)^{-1} r_p^0.$$

Substitution into the second equation of (35) gives

$$d_x^f = (I - \bar{A}^T (\bar{A}\bar{A}^T)^{-1} \bar{A}) \left(p_v - \frac{1}{\sqrt{\mu}} \theta v P(w)^{\frac{1}{2}} r_d^0 \right) + \theta v \bar{A}^T (\bar{A}\bar{A}^T)^{-1} r_p^0.$$

Let (x^*, y^*, s^*) be the optimal solution satisfying (11). Then, we have

$$r_p^0 = A(x^* - x^0), \quad r_d^0 = A^T(y^* - y^0) + (s^* - s^0),$$

and

$$(I - \bar{A}^T (\bar{A}\bar{A}^T)^{-1} \bar{A}) \bar{A}^T (y^* - y^0) = 0.$$

Hence, it follows that

$$\begin{aligned} d_x^f &= (I - \bar{A}^T (\bar{A}\bar{A}^T)^{-1} \bar{A}) \left(p_v - \frac{\theta v}{\sqrt{\mu}} P(w)^{\frac{1}{2}} (s^* - s^0) \right) \\ &\quad + \frac{\theta v}{\sqrt{\mu}} \bar{A}^T (\bar{A}\bar{A}^T)^{-1} \bar{A} P(w)^{-\frac{1}{2}} (x^* - x^0). \end{aligned}$$

For d_s^f , by using $d_x^f + d_s^f = p_v$, we obtain

$$\begin{aligned} d_s^f &= \frac{\theta v}{\sqrt{\mu}} (I - \bar{A}^T (\bar{A}\bar{A}^T)^{-1} \bar{A}) P(w)^{\frac{1}{2}} (s^* - s^0) \\ &\quad + \bar{A}^T (\bar{A}\bar{A}^T)^{-1} \bar{A} \left(p_v - \frac{\theta v}{\sqrt{\mu}} P(w)^{-\frac{1}{2}} (x^* - x^0) \right). \end{aligned}$$

Therefore, using orthogonality and the Cauchy-Schwartz inequality, after some computations, we obtain

$$\begin{aligned} \|d_x^f\|_F^2 + \|d_s^f\|_F^2 &\leq 2\|p_v\|_F^2 \\ &\quad + \frac{3\theta^2 v^2}{\mu} (\|P(w)^{-\frac{1}{2}}(x^* - x^0)\|_F^2 + \|P(w)^{\frac{1}{2}}(s^* - s^0)\|_F^2) \\ &= 8\sigma(v)^2 + \frac{3\theta^2 v^2}{\mu} (\|P(w)^{-\frac{1}{2}}(x^* - x^0)\|_F^2 + \|P(w)^{\frac{1}{2}}(s^* - s^0)\|_F^2). \end{aligned} \tag{36}$$

Since (x^*, y^*, s^*) is a primal-dual feasible solution we have $x^* \succeq_{\mathcal{K}} 0$ and $s^* \succeq_{\mathcal{K}} 0$. Hence

$$0 \preceq_{\mathcal{K}} x^0 - x^* \preceq_{\mathcal{K}} \xi e, \quad 0 \preceq_{\mathcal{K}} s^0 - s^* \preceq_{\mathcal{K}} \xi e.$$

Using that $P(w)^{\frac{1}{2}}$ is self-adjoint with respect to the inner product $\langle \cdot, \cdot \rangle$ and $P(w)e = w^2$, we may write

$$\begin{aligned} \|P(w)^{\frac{1}{2}}(s^0 - s^*)\|_F^2 &= \langle P(w)(s^0 - s^*), (s^0 - s^*) \rangle \\ &= \langle P(w)(s^0 - s^*), \xi e \rangle - \langle P(w)(s^0 - s^*), \xi e - (s^0 - s^*) \rangle \\ &\leq \langle P(w)(s^0 - s^*), \xi e \rangle = \langle s^0 - s^*, P(w)\xi e \rangle = \xi \langle s^0 - s^*, w^2 \rangle \\ &= \xi \langle \xi e, w^2 \rangle - \xi \langle \xi e - (s^0 - s^*), w^2 \rangle \leq \xi \langle \xi e, w^2 \rangle = \xi^2 \text{tr}(w^2). \end{aligned}$$

In the same way it follows that

$$\|P(w)^{-\frac{1}{2}}(x^0 - x^*)\|_F^2 \leq \xi^2 \text{tr}(w^{-2}).$$

Substitution of the last two inequalities into (36) and use from $\mu = v\xi^2$ gives

$$\begin{aligned} \|d_x^f\|_F^2 + \|d_s^f\|_F^2 &\leq 8\sigma(v)^2 + 3\theta^2 v \text{tr}(w^2 + w^{-2}) \\ &\leq 8\sigma(v)^2 + \frac{3\theta^2}{\xi^2 \lambda_{\min}(v)^2} \text{tr}(x + s)^2 \\ &\leq 8\sigma(v)^2 + \frac{3\theta^2}{\xi^2(1 - \sigma(v))^2} \text{tr}(x + s)^2, \end{aligned} \tag{37}$$

where the second inequality follows by Lemma 4.5 in Gu et al. (2011) and the third inequality follows by (9), because

$$\sigma(v)^2 = \|e - v\|_F^2 = \sum_{i=1}^r (1 - \lambda_i(v))^2,$$

implies that

$$1 - \sigma(v) \leq \lambda_i(v) \leq 1 + \sigma(v), \quad i = 1, 2, \dots, r.$$

Lemma 13 (Lemma 4.6 in Gu et al. 2011) *Let (x, y, s) be feasible for the perturbed problems (P_v) and (D_v) and let $(x^0, y^0, s^0) = (\xi e, 0, \xi e)$ and (x^*, y^*, s^*) be as defined in (11). Then*

$$\text{tr}(x + s) \leq 2\xi r, \tag{38}$$

where $r = \text{tr}(e)$ is the rank of the associated Euclidean Jordan algebra.

Substituting (38) into (36), we obtain

$$\|d_x^f\|_F^2 + \|d_s^f\|_F^2 \leq 8\sigma(v)^2 + \frac{12\theta^2 r^2}{(1 - \sigma(v))^2}. \tag{39}$$

5.2 Value for θ

At this stage, we choose $\tau = \frac{1}{16}$. Since $\sigma(v) \leq \tau = \frac{1}{16}$ and the right-hand-side of (39) is monotonically increasing in $\sigma(v)$, we have

$$\|d_x^f\|_F^2 + \|d_s^f\|_F^2 \leq \frac{1}{32} + \frac{3072\theta^2 r^2}{225}.$$

Using $\theta = \frac{\alpha}{2\sqrt{r}}$, the above relation becomes

$$\|d_x^f\|_F^2 + \|d_s^f\|_F^2 \leq \frac{1}{32} + \frac{768r\alpha^2}{225}. \tag{40}$$

From (34) we know that $\vartheta(v) \leq \frac{1}{2\sqrt{2}}$ is needed in order to have $\sigma(v^f) \leq \frac{1}{\sqrt{2}}$. Due to (40), this will hold if

$$\frac{1}{32} + \frac{768r\alpha^2}{225} \leq \frac{1}{2}.$$

If we take

$$\alpha = \frac{1}{\sqrt{8r}}, \tag{41}$$

the above inequality is satisfied.

5.3 Complexity analysis

We have seen that if at the start of an iteration the iterate satisfies $\sigma(x, s; \mu) \leq \tau$, with $\tau = \frac{1}{16}$, then after the feasibility step, with θ as defined in (33) and α as in (41), the iterate is strictly feasible and satisfies $\sigma(x^f, s^f; \mu^+) \leq \frac{1}{\sqrt{2}}$.

After the feasibility step, we perform a few centering steps in order to get the iterate (x^+, y^+, s^+) which satisfies $\sigma(x^+, s^+; \mu^+) \leq \tau$. By Lemma 6, after k centering steps we will have the iterate (x^+, y^+, s^+) that is still feasible for (P_{v^+}) and (D_{v^+}) and satisfies

$$\sigma(x^+, s^+; \mu^+) \leq \left(\frac{1}{\sqrt{2}}\right)^{2k}.$$

From this, one easily deduces that $\sigma(x^+, s^+; \mu^+) \leq \tau$ will hold after at most

$$1 + \log_2 \log_2 \frac{1}{\tau} = 1 + \log_2 \log_2 16 = 1 + \log_2 4 = 3, \tag{42}$$

centering steps. So, each main iteration consists of one feasibility step and at most 3 centering steps.

In each main iteration both the duality gap and the norms of the residual vectors are reduced by the factor $1 - \theta$. Hence, the total number of main iterations is bounded above by

$$\frac{1}{\theta} \log \frac{\max\{r\xi^2, \|r_p^0\|_F, \|r_d^0\|_F\}}{\epsilon},$$

the above iteration bound was first obtained by Potra (1996), and it is still the best-known iteration bound for IIPMs. Due to (33), (41) and the fact that we need at most 4 inner iterations per main iteration, we may state the main result of the paper.

Theorem 3 *If (P) and (D) are feasible and $\xi > 0$ is such that $x^* + s^* \leq \xi e$ for some optimal solution x^* of (P) and (y^*, s^*) of (D), then after at most*

$$16\sqrt{2} r \log \frac{\max\{r\xi^2, \|r_p^0\|_F, \|r_d^0\|_F\}}{\epsilon},$$

inner iterations, the algorithm finds an ϵ -optimal solution of (P) and (D).

6 Conclusions

In this paper, we presented a primal-dual infeasible interior-point algorithm for SCO by combining Darvay's technique and Roos' algorithm. To our knowledge, this is the first time that a primal-dual infeasible interior-point algorithm, with adaptive combinations of Darvay's technique and Roos' algorithm, has been analyzed for SCO underlying the Euclidean Jordan algebras. The proposed algorithm is closely related to Roos' algorithm. The complexity obtained here over symmetric cones, using the NT direction, coincides with the best bound obtained for IIPMs.

References

- Darvay, Z. (2003). New interior point algorithms in linear programming. *Advanced Modelling and Optimization*, 5(1), 51–92.
- Faraut, J., & Korányi, A. (1994). *Analysis on symmetric cones*. New York: Oxford University Press.
- Faybusovich, L. (1997). Linear systems in Jordan algebras and primal-dual interior-point algorithms. *Journal of Computational and Applied Mathematics*, 86(1), 149–175.
- Faybusovich, L. (2002). A Jordan-algebraic approach to potential-reduction algorithms. *Mathematische Zeitschrift*, 239(1), 117–129.
- Freund, R. M. (1996). An infeasible-start algorithm for linear programming whose complexity depends on the distance from the starting point to the optimal solution. *Annals of Operations Research*, 62, 29–57.
- Gu, G., Zangiabadi, M., & Roos, C. (2011). Full Nesterov-Todd step infeasible interior-point method for symmetric optimization. *European Journal of Operational Research*, 214(3), 473–484.
- Güler, O. (1996). Barrier functions in interior-point methods. *Mathematics of Operations Research*, 21(4), 860–885.
- Kheirfam, B. (2011). A full NT-step infeasible interior-point algorithm for semidefinite optimization based on a self-regular proximity. *ANZIAM Journal*, 53, 48–67.
- Muramatsu, M. (2002). On a commutative class of search directions for linear programming over symmetric cones. *Journal of Optimization Theory and Applications*, 112(3), 595–625.
- Nesterov, Y., & Nemirovski, A. (1994). *Interior-point polynomial algorithms in convex programming*. Philadelphia: SIAM.
- Nesterov, Y. E., & Todd, M. J. (1997). Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research*, 22(1), 1–42.
- Potra, F. A. (1996). An $O(nL)$ infeasible-interior-point algorithm for LCP with quadratic convergence. *Annals of Operations Research*, 62, 81–102.
- Rangarajan, B. K. (2006). Polynomial convergence of infeasible-interior-point methods over symmetric cones. *SIAM Journal on Optimization*, 16(4), 1211–1229.
- Roos, C. (2006). A full-Newton step $O(n)$ infeasible interior-point algorithm for linear optimization. *SIAM Journal on Optimization*, 16(4), 1110–1136.
- Schmieta, S. H., & Alizadeh, F. (2003). Extension of primal-dual interior-point algorithms to symmetric cones. *Mathematical Programming*, 96(3), 409–438.
- Sturm, J. F. (2000). Similarity and other spectral relations for symmetric cones. *Algebra and Applications*, 312(1–3), 135–154.
- Vieira, M. V. C. (2007). *Jordan algebraic approach to symmetric optimization*. Ph.D thesis, Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, The Netherlands.
- Wang, G. Q., & Bai, Y. Q. (2012). A new full Nesterov-Todd step primal-dual path-following interior-point algorithm for symmetric optimization. *Journal of Optimization Theory and Applications*, 154(3), 966–985.