Indian J Pure Appl Math (2026) 57:1532–1550
https://doi.org/10.1007/s13226-024-00718-w

ORIGINAL RESEARCH

A primal-dual interior-point method with full-Newton step for semidefinite optimization

Imene Touil

Received: 13 October 2023 / Accepted: 4 November 2024 / Published online: 17 November 2024

Abstract The aim of this paper is to derive a fresh set of search directions for a semidefinite programming problem using Darvay's technique. The algorithm introduced employs exclusively the full Nesterov–Todd (NT) step in each iteration. We initially establish the local quadratic convergence of the algorithm and subsequently demonstrate that the upper bound for worst-case iterations of the corresponding new algorithm is $\mathscr{O}\left(q^3\sqrt{n}\log\frac{n}{\epsilon}\right)$, where $q \in \mathbb{N}$ signifies the small-update method. Notably, this bound aligns with the current state-of-the-art iteration bounds. Finally, some numerical tests are presented on the developed algorithm.

Keywords Semidefinite optimization · Primal-dual interior-point methods · Full-NT step · Complexity analysis

Mathematics Subject Classification 90C22 · 90C51

1 Introduction

Following Karmarkar's seminal work [16], interior-point methods (IPMs) have emerged as a potent technique for solving linear optimization (LO) problems [21–23]. These methods have found successful extensions to broader categories of optimization, encompassing linear complementarity problems (LCP) [13,24], semidefinite optimization (SDO) [6,11,14,19,25,26,28,30] and second-order cone optimization (SOCO) [2,3]. IPMs stand out as not only highly effective in practical applications but also offering polynomial-time complexity. The choice of search directions holds a pivotal role within these methodologies. In 2002, Peng et al. [21,22] introduced the concept of self-regular functions within the context of LO, subsequently extending this approach to SDO to yield a novel class of primal-dual search directions. In the subsequent year, Darvay [7] introduced an innovative technique for generating search directions in LO. This approach involves a modification to the system that defines the central path, based on transforming the centrality equation $xs = \mu e$ into the new equation $\psi(\frac{xs}{\mu}) = \psi(e)$, where $\psi(t) = \sqrt{t},\ \forall t > 0$. Darvay established that the algorithm based on this technique exhibits an iteration bound of $\mathscr{O}(\sqrt{n}\log\frac{n}{\epsilon})$ for the small-update method, coinciding with the best-known iteration bound.

Researchers have further developed Darvay's technique to cater to various type of optimization problems, including convex quadratic optimization (CQO) [1], SDO [28], SOCO [29], and the linear complementarity problem (LCP) [8,10].

Recently, Darvay and Takács [9] introduced a novel transformation applied to the centering equations $xs = \mu e$ of the central path. This transformation aims to derive a fresh set of search directions for IPMs within the context

Communicated by G.D. Veerappa Gowda.

I. Touil (✉)
LMPA, Department of Mathematics, Mohammed Seddik Ben Yahia University, Jijel, Algeria
E-mail: touilimene@univ-jijel.dz

LMPA, Department of Mathematics, Mohammed Seddik Ben Yahia University, Jijel, Algeria

of LO. Instead of employing the conventional form $xs = \mu e$, their reformulation employs $\psi(\frac{xs}{\mu}) = \psi(\sqrt{\frac{xs}{\mu}})$, where $\psi(t) = t^2,\ \forall t > \frac{1}{\sqrt{2}}$.

More recently, Kheirfam and Nasrollahi [18] extended the technique proposed by Darvay and Takács [9] to encompass additional case $\psi(t) = t^q$ for $q \geq 2$ in the context of LO problems. Following this, Guerra [11] extended this approach to the case $\psi(t) = t^q$ with $q \geq 2$ for SDO problems.

The same authors, in a separate work [17], further extended the technique proposed by Darvay [7] to the case $\psi(t) = t^{\frac{q}{2}}$ with $q \in \mathbb{N}$ for LO problems. Recently, Grimes and Achache [10] extended this approach to the case $\psi(t) = t^{\frac{5}{2}}$ for monotone LCP problems to improve the numerical results of the interior-point algorithm (IPA) designed for LO.

In this manuscript, we expand upon the technique proposed by Darvay [7] by considering the case $\psi(t) = t^{\frac{q}{2}},\ q \in \mathbb{N}$ within the context of SDO problems.

The paper is structured as follows:

In Sect. 2, we provide a recap of the SDO problem along with the associated assumptions. We then proceed to derive the Newton search direction for the central path, utilizing the transformation introduced by Darvay. Moving to Sect. 3, we present the general form of the full-Newton step algorithm. In Sect. 4, we demonstrate the local quadratic convergence of the algorithm. Furthermore, we establish that it capable to solve the SDO problem within polynomial-time. In Sect. 5, we present some numerical experiments. Finally, we conclude the paper in Sect. 6 with some closing remarks.

1.1 Notations and terminology

Throughout this paper, various notations are utilized, and they are explained below:

The set of all $(n \times n)$ matrices with real entries is denoted as $\mathbb{R}^{n\times n}$. Given a matrix $M \in \mathbb{R}^{n\times n}$, M^T signifies the transpose of M. The matrix $M \in \mathbb{R}^{n\times n}$ is called skew-symmetric if $M = -M^T$. $\mathbb{S}^n_+$ ($\mathbb{S}^n_{++}$) denotes the set of positive semidefinite (positive definite) matrices within the real space of $(n \times n)$ symmetric matrices $\mathbb{S}^n$. For $M \in \mathbb{S}^n$, $M \succeq 0$ ($M \succ 0$) indicates that M is a positive semidefinite (positive definite) matrix. The inner product of two matrices A and B in $\mathbb{S}^n$ is the trace of their product: $\langle A, B\rangle = \mathrm{tr}(AB) = \sum_{i,j=1}^n a_{ij}b_{ij}$. For any matrix $M \in \mathbb{S}^n$, $\lambda_i(M),\ i = 1,\ldots,n$, represent the eigenvalues of M, $\lambda_{\max}(M)$ stands for its largest eigenvalue, and $\lambda_{\min}(M)$ is its smallest eigenvalue. The diagonal matrix with diagonal entries $\lambda_i(M),\ i = 1,\ldots,n$, is denoted as $\mathrm{diag}(\lambda_1(M),\ldots,\lambda_n(M))$. Two matrices A and B are called similar ($A \sim B$) if $A = PBP^{-1}$ for some invertible matrix P; additionally, if $A, B \in \mathbb{S}^n$, they share the same eigenvalues. The Frobenius norm of $M \in \mathbb{S}^n$ is $\|M\|_F = \langle M, M\rangle^{\frac{1}{2}} = \sqrt{\sum_{i=1}^n \lambda_i^2(M)}$. For any $M \in \mathbb{S}^n_{++}$, the notation $M^{\frac{1}{2}}$ (or $\sqrt{M}$) denotes its symmetric positive definite square root. Lastly, if $f(x) \geq 0$ is a real-valued function of a nonnegative real variable, $f(x) = \mathscr{O}(x)$ implies $f(x) \leq cx$ for some positive constant c.

2 The SDO problem and preliminaries

In this section, we delve into the SDO problem and lay down the groundwork for our discussion. We also explore its dual problem and provide relevant preliminaries.

2.1 The central path in SDO

We examine the SDO problem in its standard form and explore the concept of the central path.

$$(P)\begin{cases} \min\langle C, X\rangle \\ \mathscr{A}X = b, \\ X \in \mathbb{S}^n_+, \end{cases}$$

and its dual problem

$$(D)\begin{cases} \max b^T y \\ \mathscr{A}^* y + S = C, \\ S \in \mathbb{S}^n_+. \end{cases}$$

The problems (P) and (D) involve the variables $X, S \in \mathbb{S}^n$ and $y \in \mathbb{R}^m$, as well as the given data $b \in \mathbb{R}^m$ and $C \in \mathbb{S}^n$. The linear operator $\mathscr{A} : \mathbb{S}^n \to \mathbb{R}^m$ is defined by:

$$\mathscr{A}X = (\langle A_1, X\rangle, \langle A_2, X\rangle, \ldots, \langle A_m, X\rangle)^T, \tag{2.1}$$

where $A_i \in \mathbb{S}^n$ for $i = 1, \ldots, m$. Its adjoint operator $\mathscr{A}^*$ is defined by

$$\mathscr{A}^* y = \sum_{i=1}^{m} y_i A_i,$$

where $y = (y_1, y_2, \ldots, y_m)^T \in \mathbb{R}^m$.

Hypothesis: Throughout this work, we make the following assumptions:

(H_1): Both problems (P) and (D) satisfy the interior-point condition (IPC), i.e., there exists (X^0, y^0, S^0) such that

$$\mathscr{A}X^0 = b, \quad X^0 \succ 0, \quad \mathscr{A}^* y^0 + S^0 = C, \quad S^0 \succ 0.$$

(H_2): The matrices A_i, $i = 1, \ldots, m$, are linearly independent.

The concept of the central path can be extended to the SDO problem. Without loss of generality, we assume that both (P) and (D) satisfy (H_1).

The optimality conditions, also known as the Karush-Kuhn-Tucker (KKT) conditions, for (P) and (D) are given by

$$\begin{cases} \mathscr{A}X = b, X \succ 0, \\ \mathscr{A}^* y + S = C, S \succ 0, \\ XS = 0. \end{cases} \tag{2.2}$$

The main idea of primal-dual Interior-Point Methods (IPMs) is to replace the last equation in the system (2.2), known as the complementarity condition for (P) and (D), by the parameterized equation $XS = \mu I$, where $\mu > 0$. This yields the following system

$$\begin{cases} \mathscr{A}X = b, X \succ 0, \\ \mathscr{A}^* y + S = C, S \succ 0, \\ XS = \mu I, \quad \mu > 0. \end{cases} \tag{2.3}$$

Under (H_1), the system (2.3) possesses a unique solution denoted by (X_μ, y_μ, S_μ) for each $\mu > 0$ [27]. The collection of all such solutions (X_μ, y_μ, S_μ) with $\mu > 0$ is commonly referred to as the *central path* of the problems (P) and (D). Additionally, as μ approaches zero, the central path converges to a limit. Notably, these limit points satisfy the complementarity condition, leading to optimal solutions for both (P) and (D) (for more details, see, for example, [26,27]).

2.2 The search direction in the context of SDO

In his work [7], Darvay introduced a novel transformation that enables the determination of a new category of search directions for LO problems. This transformation involves replacing the standard centering equation $xs = \mu e$ with $\psi(\frac{xs}{\mu}) = \psi(e)$, where $\psi : (0, \infty) \to (0, \infty)$ is a continuously differentiable and strictly monotone vector function, derived from the function $\psi(t)$ and assumed to be invertible.

Expanding upon the LO case, a similar approach can be taken for the third equation of (2.3). This yields the ensuing system

$$\begin{cases} \mathscr{A}X = b, X \succ 0, \\ \mathscr{A}^* y + S = C, S \succ 0, \\ \psi\left(\frac{XS}{\mu}\right) = \psi(I), \quad \mu > 0. \end{cases} \tag{2.4}$$

2.2.1 Symmetric matrices and matrix function properties

Let's revisit some well-established results concerning symmetric matrices and matrix functions, along with their derivatives [15,20].

Theorem 2.1 (Spectral theorem) *Consider a matrix $A \in \mathbb{R}^{n\times n}$. A is classified as a symmetric matrix if and only if there exists an orthogonal matrix $Q \in \mathbb{R}^{n\times n}$ (i.e., $Q^TQ = I$) such that $A = Q^T diag(\lambda_1(A), \dots, \lambda_n(A))Q$, where I represents the $(n \times n)$ identity matrix.*

As a direct consequence of this theorem, it's evident that $\lambda_i(A^m) = \lambda_i^m(A),\ i = 1, \dots, n$, holds true when $A^m = A \times A \times \cdots \times A$ (m times) belongs to $\mathbb{S}^n_+$, with $m \in \mathbb{N}$.

Let's delve into the definitions and properties of matrix functions.

Definition 2.2 Consider a matrix V that is diagonalizable with the representation

$$V = Q^T diag(\lambda_1(V), \lambda_2(V), \dots, \lambda_n(V))Q,$$

where Q is a nonsingular matrix, and $h(.) : \mathbb{R} \to \mathbb{R}$ is a given function. The matrix function $h(V)$ is defined as

$$h(V) = Q^T diag(h(\lambda_1(V)), h(\lambda_2(V)), \dots, h(\lambda_n(V)))Q. \tag{2.5}$$

This definition establishes how matrix functions are evaluated when applied to diagonalizable matrices.

It's important to note that the matrix Q might not be unique, yet the matrix function $h(V)$ remains well-defined as long as the scalar function $h(t)$ is well-defined for the eigenvalues of V [21,28].

Remark 2.3 In this paper, whenever we refer to the function $h(.)$ and its derivative $h'(.)$ without specific context, it signifies a matrix function when the argument is a matrix. If the argument lies in $\mathbb{R}$, it denotes a univariate function from $\mathbb{R}$ to itself.

Lemma 2.4 (Lemma 2.5 in [28]) *Assuming A and B belong to $\mathbb{S}^n$ and commute, i.e., $AB = BA$, the following holds:*

$$\lambda_i(A + B) = \lambda_i(A) + \lambda_i(B),\ i = 1, \dots, n.$$

Moreover, if $|\lambda_i(B)|$ is sufficiently small, we obtain

$$h(A + B) \approx h(A) + h'(A)B.$$

Subsequently, we employ Newton's method to linearize the system (2.4), resulting in the following system:

$$\begin{cases} \mathscr{A}\Delta X = 0, \\ \mathscr{A}^*\Delta y + \Delta S = 0, \\ \psi\left(\frac{XS}{\mu} + \frac{\Delta XS + X\Delta S + \Delta X\Delta S}{\mu}\right) = \psi(I). \end{cases} \tag{2.6}$$

Here, ΔX, Δy, and ΔS represent the search directions. By utilizing Lemma 2.4, we can express the final equation of (2.6) in the following manner:

$$\psi\left(\frac{XS}{\mu}\right) + \psi'\left(\frac{XS}{\mu}\right)\left(\frac{\Delta XS + X\Delta S}{\mu}\right) = \psi(I),$$

this implies that

$$\Delta XS + X\Delta S = \mu\left(\psi'\left(\frac{XS}{\mu}\right)\right)^{-1}\left(\psi(I) - \psi\left(\frac{XS}{\mu}\right)\right).$$

Then, we obtain the following system:

$$\begin{cases} \mathscr{A}\Delta X = 0, \\ \mathscr{A}^*\Delta y + \Delta S = 0, \\ \Delta X + X\Delta SS^{-1} = \mu\left(\psi'\left(\frac{XS}{\mu}\right)\right)^{-1}\left(\psi(I) - \psi\left(\frac{XS}{\mu}\right)\right)S^{-1}. \end{cases} \tag{2.7}$$

It's evident that $\Delta S \in \mathbb{S}^n$ due to the second equation in (2.7). Nevertheless, ΔX does not necessarily belongs to $\mathbb{S}^n$ in general. To address this issue, various approaches have been proposed to symmetrize the third equation of the system (2.7) in order to ensure that the new system possesses a unique symmetric solution. For further insights, refer to [14,25,30].

In this study, we adopt the Nesterov–Todd (**NT**)-symmetrization scheme introduced in [25]. This choice is motivated by the fact that the **NT** scaling technique aligns the primal and dual variables, specifically X and S, within the same space known as the V-space.

Let's define the matrix

$$P = \left[X^{\frac{1}{2}}\left(X^{\frac{1}{2}}SX^{\frac{1}{2}}\right)^{-\frac{1}{2}}X^{\frac{1}{2}}\right]^{\frac{1}{2}} = \left[S^{-\frac{1}{2}}\left(S^{\frac{1}{2}}XS^{\frac{1}{2}}\right)^{\frac{1}{2}}S^{-\frac{1}{2}}\right]^{\frac{1}{2}}. \tag{2.8}$$

Therefore, we define the matrix V as follows:

$$V = \frac{1}{\sqrt{\mu}}P^{-1}XP^{-1} = \frac{1}{\sqrt{\mu}}PSP. \tag{2.9}$$

It's important to observe that both matrices P and V are symmetric and positive definite. Additionally, we can establish the following relationship:

$$V^2 = \left(\frac{P^{-1}XP^{-1}}{\sqrt{\mu}}\right)\left(\frac{PSP}{\sqrt{\mu}}\right) = \left(\frac{P^{-1}XSP}{\mu}\right). \tag{2.10}$$

Let us further define

$$\overline{A_i} = \frac{1}{\sqrt{\mu}}PA_iP,\ i = 1,\ldots,m,\ D_X = \frac{1}{\sqrt{\mu}}P^{-1}\Delta XP^{-1} \text{ and } D_S = \frac{1}{\sqrt{\mu}}P\Delta SP. \tag{2.11}$$

In the NT-scheme, we substitute $X\Delta SS^{-1}$ by $P^2\Delta S(P^2)^T$. By employing (2.9) and (2.11), we can express the Newton system (2.7), responsible for defining the search directions D_X, Δy and D_S as follows:

$$\begin{cases} \overline{\mathscr{A}}D_X = 0, \\ \overline{\mathscr{A}}^*\Delta y + D_S = 0, \\ D_X + D_S = P_V, \end{cases} \tag{2.12}$$

Here, $\overline{\mathscr{A}}$ represents the operator defined in (2.1), where the matrices A_i are replaced by $\overline{A_i}$ for $i = 1,\ldots,m$. Furthermore, $\overline{\mathscr{A}}^*$ signifies its adjoint, and we have:

$$P_V = \sqrt{\mu}P^{-1}\left(P\psi'(V^2)P^{-1}\right)^{-1}\left(\psi(I) - P\psi(V^2)P^{-1}\right)S^{-1}P^{-1}.$$

It's important to highlight that the selection of the function ψ leads to various values for the matrix P_V, consequently generating distinct search directions.

Let's categorize the various cases of $\psi(t)$ and their corresponding resulting search directions:

- For $\psi(t) = t$, we obtain $P_V = V^{-1} - V$, leading to the classical search direction that has been extensively explored by many researchers (e.g., [4,5,19,21,28]).
- When $\psi(t) = \sqrt{t}$, we have $P_V = 2(I - V)$, which was initially studied by Darvay in [7] for LO problems and later extended by Wang and Bai in [28] for SDO problems.
- In the case of $\psi(t) = t^2$, the result is $P_V = \frac{1}{2}(V - V^3)(2V^2 - I)^{-1}$, a subject studied by Darvay and Takács in [9] for LO problems.
- Lastly, for $\psi(t) = t^q$ with $q \geq 2$, we arrive at $P_V = \frac{2}{q}(V - V^{q+1})(2V^q - I)^{-1}$. This case was explored by Kheirfam and Nasrollahi in [18] for LO problems and later extended to SDO problems by Guerra [11].

These different $\psi(t)$ choices lead to distinct forms of search directions and have been studied extensively in the literature.

In alignment with [1,7,28], this paper delves into the extension of the case $\psi(t) = t^{\frac{q}{2}}$ for all $t > 0$ and $q \in \mathbb{N}$, where $\mathbb{N}$ represents the set of natural numbers. This extension is directed towards addressing SDO problems, leading to the following expression:

$$P_V = \frac{2}{q}(I - V^q)V^{1-q} = \frac{2}{q}(V^{1-q} - V), \text{ with } \lambda_{\min}(V) > 0 \text{ and } q \in \mathbb{N}. \tag{2.13}$$

Let's begin by introducing the following technical lemma:

Lemma 2.5 *Given the previously defined P_V, we can establish the following relationship:*

$$I - \frac{P_V^2}{4} \preceq V^2 + VP_V \preceq I + \left((q-1)^2 - 1\right)\frac{P_V^2}{4}.$$

Proof Referring to (2.13), we find:

$$\begin{aligned} V^2 + VP_V &= V^2 + \frac{2}{q}V(I - V^q)V^{1-q} \\ &= V^q V^{2-q} + \frac{2}{q}(I - V^q)V^{2-q} \\ &= \frac{1}{q}\left[qV^q + 2(I - V^q)\right]V^{2-q} \\ &= \frac{1}{q}\Big[(q-2)V^q + qV^{q-2} - qV^{q-2} + 2I\Big]V^{2-q} \\ &= I + \frac{1}{q}\Big[(q-2)V^q - qV^{q-2} + 2I\Big]V^{2-q}. \end{aligned}$$

By virtue of (2.5) in Definition 2.2, we acquire:

$$V^2 + VP_V - I = h(V) = Q^T diag(h(\lambda_1(V)), h(\lambda_2(V)), \ldots, h(\lambda_n(V)))Q, \tag{2.14}$$

where $h(\lambda_i(V)) = \frac{(q-2)\lambda_i^q(V) - q\lambda_i^{q-2}(V) + 2}{q\lambda_i^{q-2}(V)}$, for $i = 1, \ldots, n$.

Let us define

$$f(t) = (q-2)t^q - qt^{q-2} + 2, \text{ for all } t > 0 \text{ and } q \in \mathbb{N}.$$

This leads us to consider the three distinct cases. Let's analyze the three cases in detail:

Case 1: For $q = 1$, we arrive at $f(t) = -t^{-1}(t-1)^2$, which leads to $h(V) = -\frac{P_V^2}{4}$, this results in the equation:

$$V^2 + VP_V = I - \frac{P_V^2}{4}.$$

Case 2: For $q = 2$, we get $f(t) = 0$, hence $h(V) = 0$, which implies $V^2 + VP_V - I = 0$. Since $\frac{P_V^2}{4} \succ 0$ for $\lambda_i(V) > 0$, we have $V^2 + VP_V \succeq I - \frac{P_V^2}{4}$.

Case 3: When $q \geq 3$, it's easy to verify that $f'(1) = 0$ and $f''(1) = 2q(q-2) > 0$. These conditions indicate that $f(t) \geq f(1) = 0$ for all $t > 0$. Therefore, we conclude: $V^2 + VP_V - I \succeq 0$, gives that $\frac{P_V^2}{4} \succ 0$ for $\lambda_i(V) > 0$, this leads to $V^2 + VP_V \succeq I - \frac{P_V^2}{4}$.

Consequently, we have successfully established the validity of the first inequality.

Furthermore, in order to demonstrate the second inequality, let's introduce the function h as follows:

$$
\begin{aligned}
h(t) &= \frac{f(t)}{qt^{q-2}} = \frac{(q-2)t^q - qt^{q-2} + 2}{qt^{q-2}} \\
&= \frac{(1-t)^2 \left[(q-2)t^{q-2} + 2\sum_{i=2}^{q-1}(q-i)t^{q-i-1}\right]}{qt^{q-2}} \\
&= \frac{(1-t)^2 \left[(q-2)t^{q-2} + 2\sum_{i=2}^{q-1}(q-i)t^{q-i-1}\right] qt^q \left(\sum_{i=0}^{q-1} t^i\right)^2}{q^2 t^{2(q-1)} \left(\sum_{i=0}^{q-1} t^i\right)^2} \\
&= \left(\frac{1-t^q}{qt^{q-1}}\right)^2 \frac{\left[(q-2)t^{q-2} + 2\sum_{i=2}^{q-1}(q-i)t^{q-i-1}\right] qt^q}{\left(\sum_{i=0}^{q-1} t^i\right)^2} \\
&= q\left(\frac{1-t^q}{qt^{q-1}}\right)^2 \left[\frac{(q-2)\left(\sum_{i=0}^{q-1} t^i\right)^2 + 2\sum_{i=2}^{q-1}(q-i)t^{2q-i-1}}{\left(\sum_{i=0}^{q-1} t^i\right)^2} - \frac{2(q-2)\sum_{i=2}^{q} t^{2q-i-1} + (q-2)\left(\sum_{i=0}^{q-1} t^i\right)^2}{\left(\sum_{i=0}^{q-1} t^i\right)^2}\right] \\
&= q\left(\frac{1-t^q}{qt^{q-1}}\right)^2 \left[(q-2) - \frac{2\sum_{i=2}^{q}(i-2)t^{2q-i-1} + (q-2)\left(\sum_{i=0}^{q-1} t^i\right)^2 + 2(q-2)t^{q-1}}{\left(\sum_{i=0}^{q-1} t^i\right)^2}\right] \\
&\le q(q-2)\left(\frac{1-t^q}{qt^{q-1}}\right)^2 = \left((q-1)^2 - 1\right)\left(\frac{1-t^q}{qt^{q-1}}\right)^2.
\end{aligned}
$$

Substituting this latest bound into (2.14), we infer that:

$$
\begin{aligned}
V^2 + VP_V - I &\preceq \left((q-1)^2 - 1\right) Q^T diag\left(\left(\frac{1-\lambda_1^q(V)}{q\lambda_1^{q-1}(V)}\right)^2, \ldots, \left(\frac{1-\lambda_n^q(V)}{q\lambda_n^{q-1}(V)}\right)^2\right) Q. \\
&= \left((q-1)^2 - 1\right)\frac{P_V^2}{4},
\end{aligned}
$$

where the last equality follows from (2.13). Hence, the proof is now complete. □

For the analysis of the interior-point algorithm, let's introduce the norm-based proximity measure $\sigma(X, S; \mu)$ as follows:

$$
\sigma(V) := \sigma(X, S; \mu) = \frac{1}{2}\|D_X + D_S\|_F = \frac{1}{2}\|P_V\|_F = \|\frac{1}{q}(I - V^q)V^{1-q}\|_F. \tag{2.15}
$$

From the initial two equations of system (2.12), it follows that D_X and D_S are orthogonal, specifically $\langle D_X, D_S\rangle = \langle D_S, D_X\rangle = 0$. Subsequently, we can straightforwardly confirm that:

$$
\sigma(V) = 0 \Leftrightarrow V = I \Leftrightarrow D_X = D_S = 0 \Leftrightarrow XS = \mu I,
$$

This means that $(X, y, S) = (X(\mu), y(\mu), S(\mu))$ if and only if $V = I$, implying that ΔX, Δy, and ΔS are all zero. In other words, the new iterate (X_+, y_+, S_+) obtained by taking a full-Newton step is expressed as follows:

$$
(X_+, y_+, S_+) = (X, y, S) + (\Delta X, \Delta y, \Delta S). \tag{2.16}
$$

3 The full-Newton step algorithm

Algorithm 1 presents the general structure of the full-Newton step algorithm.

Algorithm 1 Full-Newton Step Algorithm

Require: Tolerance $\epsilon > 0$; barrier update parameter $\theta \in (0,1)$ (default $\theta = \frac{1}{((2q-1)^3-5q)\sqrt{n}}$); threshold parameter $\tau \in (0,1)$ (default $\tau = \frac{2}{(q-1)^2+2}$); initial iterate (X^0, y^0, S^0) and $\mu^0 = \frac{\langle X^0, S^0\rangle}{n}$ such that $\sigma(X^0, S^0; \mu^0) < \tau$;

begin
$X := X^0;\ y := y^0;\ S := S^0;\ \mu := \mu^0;$
while $\langle X, S\rangle > \epsilon$ **do**
  Compute V and P_V as defined earlier;
  Calculate ΔX, Δy, and ΔS by solving the system (2.12) and using (2.11);
  Update $(X, y, S) := (X, y, S) + (\Delta X, \Delta y, \Delta S)$;
  Update $\mu := (1-\theta)\mu$;
end while
end

4 Analysis of the algorithm

In this section, we will establish that the proposed algorithm exhibits local quadratic convergence and solves the SDO problem in polynomial-time.

For the algorithm analysis, we introduce the notation $Q_V = D_X - D_S$. Utilizing the orthogonality property of D_X and D_S, we derive:

$$\|P_V\|^2 = \|D_X + D_S\|^2 = \|D_X\|^2 + \|D_S\|^2 = \|Q_V\|^2 = 4\sigma(V)^2. \tag{4.1}$$

Moreover, based on the definitions of P_V and Q_V, we can deduce the following relationships:

$$D_X = \frac{P_V + Q_V}{2},\ \ D_S = \frac{P_V - Q_V}{2},\ \ D_{XS} = \frac{D_X D_S + D_S D_X}{2} = \frac{P_V^2 - Q_V^2}{4}. \tag{4.2}$$

Let $0 \le \alpha \le 1$ and define:

$$X(\alpha) = X + \alpha\Delta X \quad \text{and} \quad S(\alpha) = S + \alpha\Delta S. \tag{4.3}$$

We'll refer to the following two useful lemmas (Lemma 6.1 and Lemma 6.3 in [19]), which will be employed in our subsequent analysis.

Lemma 4.1 ([19]) *Suppose $X \succ 0$ and $S \succ 0$. For all $0 \le \overline{\alpha} \le \alpha$, if $\det(X(\alpha)S(\alpha)) > 0$, then $X(\overline{\alpha}) \succ 0$ and $S(\overline{\alpha}) \succ 0$.*

Lemma 4.2 ([19]) *Consider $Q \in \mathbb{S}^n_{++}$ and $M \in \mathbb{R}^{n\times n}$ as skew-symmetric. We find that $\det(Q + M) > 0$. Additionally, if $\lambda_i(Q + M) \in \mathbb{R}$ for $i = 1, \ldots, n$, then the following holds:*

$$0 < \lambda_{\min}(Q) \le \lambda_{\min}(Q + M) \le \lambda_{\max}(Q + M) \le \lambda_{\max}(Q),$$

implying that the matrix $Q + M$ is positive definite (i.e., $Q + M \succ 0$).

These lemmas will prove to be valuable tools in our analysis.

The subsequent lemma confirms the strict feasibility of the new iterates $X(\alpha)$ and $S(\alpha)$ when the condition $\sigma(X, S; \mu) < 1$ is satisfied.

Lemma 4.3 *If $\sigma := \sigma(X, S; \mu) < 1$, then the new iterates X_+ and S_+ are strictly feasible.*

Proof By employing Eqs. (2.11) within Eq. (4.3), we obtain:

$$\begin{aligned}
X(\alpha)S(\alpha) &= XS + \alpha(X\Delta S + \Delta X S) + \alpha^2 \Delta X \Delta S \\
&= \mu P(V^2 + \alpha(D_X V + V D_S) + \alpha^2 D_X D_S)P^{-1} \\
&\sim \mu(V^2 + \alpha(D_X V + V D_S) + \alpha^2 D_X D_S) \\
&= \mu\left(V^2 + \frac{1}{2}\alpha(D_X V + V D_S + V D_X + D_S V) + \frac{1}{2}\alpha^2(D_X D_S + D_S D_X)\right) \\
&\quad + \mu\left(\frac{1}{2}\alpha(D_X V + V D_S - V D_X - D_S V) + \frac{1}{2}\alpha^2(D_X D_S - D_S D_X)\right) \\
&= Q(\alpha) + M(\alpha).
\end{aligned} \tag{4.4}$$

It can be straightforwardly demonstrated that $M(\alpha) = -M(\alpha)^T$. Consequently, by virtue of Lemma 4.2, this implies that $\det(X(\alpha)S(\alpha)) > 0$ if $Q(\alpha) \succ 0$.

Building upon Eq. (4.2) and the implications of Lemma 2.5, we deduce:

$$\begin{aligned}
Q(\alpha) &= \mu\left[V^2 + \frac{1}{2}\alpha(D_X V + V D_S + V D_X + D_S V) + \frac{1}{2}\alpha^2(D_X D_S + D_S D_X)\right] \\
&= \mu\left[V^2 + \frac{1}{2}\alpha(V P_V + P_V V) + \alpha^2 \frac{P_V^2 - Q_V^2}{4}\right] \\
&= \mu\left[V^2 + \alpha V P_V + \alpha^2 \frac{P_V^2 - Q_V^2}{4}\right] \\
&= \mu\left[(1-\alpha)V^2 + \alpha(V^2 + V P_V) + \alpha^2 \frac{P_V^2 - Q_V^2}{4}\right] \\
&\succeq \mu\left[(1-\alpha)V^2 + \alpha\left(I - \frac{P_V^2}{4}\right) + \alpha^2 \frac{P_V^2 - Q_V^2}{4}\right] \\
&= \mu\left[(1-\alpha)V^2 + \alpha\left(I - (1-\alpha)\frac{P_V^2}{4} - \alpha\frac{Q_V^2}{4}\right)\right].
\end{aligned} \tag{4.5}$$

Considering that $0 \le \alpha \le 1$, it follows that $Q(\alpha) \succ 0$ if and only if $I - (1-\alpha)\frac{P_V^2}{4} - \alpha\frac{Q_V^2}{4} \succ 0$.

Proceeding, by applying Eq. (4.1), we obtain:

$$\begin{aligned}
\|(1-\alpha)\frac{P_V^2}{4} + \alpha\frac{Q_V^2}{4}\|_F &\le (1-\alpha)\|\frac{P_V^2}{4}\|_F + \alpha\|\frac{Q_V^2}{4}\|_F \\
&\le (1-\alpha)\frac{\|P_V^2\|_F}{4} + \alpha\frac{\|Q_V^2\|_F}{4} \\
&\le (1-\alpha)\frac{\|P_V\|_F^2}{4} + \alpha\frac{\|Q_V\|_F^2}{4} = \sigma^2 < 1.
\end{aligned}$$

This inequality yields:

$$1 - \left\|(1-\alpha)\frac{P_V^2}{4} + \alpha\frac{Q_V^2}{4}\right\|_F > 0,$$

Hence, for all $0 \le \alpha \le 1$ we can conclude that:

$$\left[(1-\alpha)V^2 + \alpha\left(I - (1-\alpha)\frac{P_V^2}{4} - \alpha\frac{Q_V^2}{4}\right)\right] \succ 0,$$

This implies that $Q(\alpha) \succ 0$, leading to $\det(X(\alpha)S(\alpha)) > 0$ for all $\alpha \in [0, 1]$. Moreover, in accordance with Lemma 4.1, it is evident that $X(1) = X_+ \succ 0$ and $S(1) = S_+ \succ 0$ since $X(0) = X \succ 0$ and $S(0) = S \succ 0$. This concludes the proof of the lemma. □

Lemma 4.4 *Let* $\sigma := \sigma(X, S; \mu) < \frac{2}{(q-1)^2+2}$ *and* $V_+ = \sqrt{\frac{X_+S_+}{\mu}}$, *then*

(1) $\lambda_{\min}(V_+) \geq \sqrt{1-\sigma^2}$.

(2) $\sigma(X_+, S_+; \mu) \leq \dfrac{1-(1-\sigma^2)^{\frac{q}{2}}}{q\sigma^2(1-\sigma^2)^{\frac{q-1}{2}}}\left((q-1)^2+1\right)\sigma^2$.

Proof For the first item, let $\alpha = 1$ in (4.4) and (4.5), we get

$$X_+S_+ = \mu V_+^2 \sim Q(1) + M(1) \succeq \mu\left(I - \frac{Q_V^2}{4}\right) + M(1).$$

From Lemma 4.2, it follows that

$$\lambda_{\min}(V_+^2) = \lambda_{\min}\left(\frac{Q(1)+M(1)}{\mu}\right) \geq \lambda_{\min}\left(\frac{Q(1)}{\mu}\right) \geq \lambda_{\min}\left(I - \frac{Q_V^2}{4}\right).$$

This implies that

$$\lambda_{\min}(V_+^2) \geq \lambda_{\min}\left(I - \frac{Q_V^2}{4}\right) \geq 1 - \lambda_{\max}\left(\frac{Q_V^2}{4}\right) \geq 1 - \left\|\frac{Q_V^2}{4}\right\|_F \geq 1 - \frac{\|Q_V\|_F^2}{4} = 1 - \sigma^2.$$

This proves (1).

For the second item, using (2.15) and the fact that $(I - V_+)V_+^{1-q} = V_+^{1-q}(I - V_+)$, we have

$$\begin{aligned}
\sigma(X_+, S_+; \mu) &= \left\|\frac{1}{q}(I - V_+^q)V_+^{1-q}\right\|_F \\
&= \left\|\frac{1}{q}(I + V_+ + \cdots + V_+^{q-1})(I - V_+)V_+^{1-q}\right\|_F \\
&= \left\|\frac{1}{q}\left(I + V_+ + \cdots + V_+^{q-1}\right)\left(V_+^{1-q}(I + V_+)^{-1}\right)(I - V_+^2)\right\|_F \\
&= \left(\sum_{i=1}^{n}\left(\frac{\left(1 + \lambda_i(V_+) + \cdots + \lambda_i(V_+^{q-1})\right)(1 - \lambda_i(V_+^2))}{q\lambda_i(V_+^{q-1})(1 + \lambda_i(V_+))}\right)^2\right)^{\frac{1}{2}} \qquad (4.6)\\
&\leq \max_{i=1}^{n}\left(\frac{1 + \lambda_i(V_+) + \cdots + \lambda_i(V_+^{q-1})}{q\lambda_i(V_+^{q-1})(1 + \lambda_i(V_+))}\right)\left(\sum_{i=1}^{n}\left(1 - \lambda_i(V_+^2)\right)^2\right)^{\frac{1}{2}} \\
&= \frac{1 + \lambda_{\min}(V_+) + \cdots + \lambda_{\min}^{q-1}(V_+)}{q\lambda_{\min}^{q-1}(V_+)(1 + \lambda_{\min}(V_+))}\|(I - V_+^2)\|_F \\
&= g(\lambda_{\min}(V_+))\|(I - V_+^2)\|_F,
\end{aligned}$$

where the next-to-last equality holds due to the fact that the function g, defined as

$$g(t) = \frac{t^{q-1} + t^{q-2} + \cdots + t + 1}{qt^{q-1}(1+t)} = \frac{1 - t^q}{qt^{q-1}(1 - t^2)}$$

is decreasing on $t > 0$. Hence, from the first item of this lemma, it follows that

$$g(\lambda_{\min}(V_+)) \leq g(\sqrt{1-\sigma^2}) = \frac{1 - (\sqrt{1-\sigma^2})^q}{q\sigma^2(\sqrt{1-\sigma^2})^{q-1}}. \qquad (4.7)$$

On the other hand, from (4.4) with $\alpha = 1$ and Lemma 2.5, we have

$$\begin{aligned}
\mu V_+^2 &= X_+S_+ \sim Q(1) + M(1) \\
&= \mu\left(V^2 + VP_V + \frac{P_V^2 - Q_V^2}{4}\right) + M(1) \\
&\preceq \mu\left(I + \left((q-1)^2 - 1\right)\frac{P_V^2}{4} + \frac{P_V^2}{4} - \frac{Q_V^2}{4}\right) + M(1) \\
&= \mu\left(I + (q-1)^2\frac{P_V^2}{4} - \frac{Q_V^2}{4}\right) + M(1),
\end{aligned} \tag{4.8}$$

this gives

$$V_+^2 - I \preceq (q-1)^2\frac{P_V^2}{4} - \frac{Q_V^2}{4} + \frac{1}{\mu}M(1).$$

Since $M(1)$ is a skew-symmetric matrix, then after some elementary reductions, it follows that

$$\begin{aligned}
\|I - V_+^2\|_F &\leq \left\|-(q-1)^2\frac{P_V^2}{4} + \frac{Q_V^2}{4} - \frac{1}{\mu}M(1)\right\|_F \\
&\leq \left\|-(q-1)^2\frac{P_V^2}{4}\right\|_F + \left\|\frac{Q_V^2}{4} - \frac{1}{\mu}M(1)\right\|_F \\
&\leq (q-1)^2\frac{\|P_V\|_F^2}{4} + \frac{\|Q_V\|_F^2}{4} \\
&= (q-1)^2\frac{\|P_V\|_F^2}{4} + \frac{\|P_V\|_F^2}{4} \\
&= \left((q-1)^2 + 1\right)\sigma^2.
\end{aligned} \tag{4.9}$$

The result holds by substituting (4.7) and (4.9) into (4.6). This completes the proof. □

The following corollary is an immediate consequence of the above lemma which shows the quadratic convergence of full NT-step to the target point $(X(\mu), y(\mu), S(\mu))$.

Corollary 4.5 *If* $\sigma < \frac{2}{(q-1)^2+2}$, *then*

$$\sigma(X_+, S_+; \mu) < \overline{\gamma}(q)\sigma^2.$$

Proof From the above lemma, we have

$$\sigma(X_+, S_+; \mu) \leq \frac{1 - (1-\sigma^2)^{\frac{q}{2}}}{q\sigma^2(1-\sigma^2)^{\frac{q-1}{2}}}\left((q-1)^2 + 1\right)\sigma^2.$$

Let us define

$$g(\sigma) = \frac{1 - (1-\sigma^2)^{\frac{q}{2}}}{q\sigma^2(1-\sigma^2)^{\frac{q-1}{2}}}.$$

It is easy to prove that g is increasing function for all $0 < \sigma < 1$, hence for all $0 < \sigma < \frac{2}{(q-1)^2+2}$. Therefore, after some elementary reductions, we obtain

$$\begin{aligned}
g(\sigma) &\leq g\left(\frac{2}{(q-1)^2+2}\right) \\
&= \frac{\left((q-1)^2+2\right)^{q+1} - (q-1)^q\left((q-1)^2+2\right)\left((q-1)^2+4\right)^{\frac{q}{2}}}{4q(q-1)^{q-1}\left((q-1)^2+4\right)^{\frac{q-1}{2}}} := \gamma(q).
\end{aligned}$$

Hence

$$\sigma(X_+, S_+; \mu) \le \left((q-1)^2 + 1\right)\gamma(q)\sigma^2 := \overline{\gamma}(q)\sigma^2.$$

This proves the corollary. □

The following lemma yields an upper bound on the duality gap after a full NT-step.

Lemma 4.6 *Suppose that* $\sigma < \frac{2}{(q-1)^2+2}$. *Then, after a full NT-step, we have*

$$\langle X_+, S_+\rangle \le \mu\left(n + \frac{4q(q-2)}{((q-1)^2+2)^2}\right).$$

Proof Since $V_+^2 \sim \frac{X_+S_+}{\mu}$, then $\langle X_+, S_+\rangle = \mu \mathrm{tr}(V_+^2)$. From (4.8) with $M(1)$ is a skew-symmetric matrix and using Lemma 2.5, it follows that

$$\begin{aligned}
\langle X_+, S_+\rangle &= \mathrm{tr}\left(\mu(V^2 + VP_V + D_{XS}) + M(1)\right)\\
&= \mathrm{tr}\left[\mu(V^2 + VP_V + D_{XS})\right]\\
&= \mu\left[n + \mathrm{tr}\left(V^2 + VP_V - I + D_{XS}\right)\right]\\
&\le \mu\left[n + \mathrm{tr}\left(((q-1)^2 - 1)\frac{P_V^2}{4}\right) + \mathrm{tr}(D_{XS})\right]\\
&= \mu\left[n + \left((q-1)^2 - 1\right)\mathrm{tr}\left(\frac{P_V^2}{4}\right)\right]\\
&\le \mu\left[n + \left((q-1)^2 - 1\right)\frac{\|P_V\|_F^2}{4}\right]\\
&= \mu\left[n + \left((q-1)^2 - 1\right)\sigma^2\right]\\
&\le \mu\left(n + \frac{4q(q-2)}{\left((q-1)^2+2\right)^2}\right),
\end{aligned}$$

where the last inequality is due to the assumption $\sigma < \frac{2}{(q-1)^2+2}$.

Hence, the proof is completed. □

The next result analyzes the effect on the proximity measure of a full NT-step followed by an update of the parameter μ.

Lemma 4.7 *Let* $\sigma < \frac{2}{(q-1)^2+2}$ *with* $q \neq 1$ *and* $\mu_+ = (1-\theta)\mu$ *with* $0 < \theta < 1$. *Then* $\widehat{V} = \frac{V_+}{\sqrt{1-\theta}} \succ 0$ *and*

$$\sigma(X_+, S_+; \mu_+) \le \frac{\overline{\sigma}^{q-1} + \overline{\theta}\overline{\sigma}^{q-2} + \cdots + \overline{\sigma}\overline{\theta}^{q-2} + \overline{\theta}^{q-1}}{q\overline{\theta}\overline{\sigma}^{q-1}(\overline{\sigma} + \overline{\theta})}\left(\theta\sqrt{n} + \left((q-1)^2+1\right)\sigma^2\right),$$

where $\overline{\sigma} = \sqrt{1-\sigma^2}$ *and* $\overline{\theta} = 1-\theta$.

In addition, if $\theta = \frac{1}{((q-1)^3 - 5q)\sqrt{n}}$ *with* $n \ge 1$, *then* $\sigma(X_+, S_+; \mu_+) < \frac{2}{(q-1)^2+2}$.

Proof By the definition of $\widehat{V}$ and from (1) of Lemma 4.4, we have

$$\begin{aligned}
\lambda_{\min}(\widehat{V}) &= \lambda_{\min}\left(\frac{V_+}{\sqrt{1-\theta}}\right) = \frac{1}{\sqrt{1-\theta}}\lambda_{\min}(V_+)\\
&\ge \sqrt{\frac{1-\sigma^2}{1-\theta}} = \frac{\overline{\sigma}}{\overline{\theta}} \ge \frac{(q-1)^2\left((q-1)^2+4\right)}{\sqrt{1-\theta}\left((q-1)^2+2\right)^2} > 0,
\end{aligned}$$

where the last inequality follows from the fact that $q \neq 1$ and $0 < \theta < 1$. Thus $\widehat{V} \succ 0$, which proves the first part of this lemma.

On the other hand, from (2.15) we have

$$\begin{aligned}
\sigma(X_+, S_+; \mu_+) &= \frac{1}{q} \|(I - \widehat{V}^q)\widehat{V}^{1-q}\|_F \\
&= \frac{1}{q} \left\| \left(I + \widehat{V} + \cdots + \widehat{V}^{q-1}\right)(I - \widehat{V})\widehat{V}^{1-q} \right\|_F \\
&= \left\| \frac{1}{q} \left(I + \widehat{V} + \cdots + \widehat{V}^{q-1}\right)\left(\widehat{V}^{1-q}(I + \widehat{V})^{-1}\right)(I - \widehat{V}^2) \right\|_F \\
&= \left(\sum_{i=1}^{n} \left(\frac{\left(1 + \lambda_i(\widehat{V}) + \cdots + \lambda_i(\widehat{V}^{q-1})\right)(1 - \lambda_i(\widehat{V}^2))}{q\lambda_i(\widehat{V}^{q-1})(1 + \lambda_i(\widehat{V}))} \right)^2 \right)^{\frac{1}{2}} \\
&\leq \max_{i=1}^{n} \left(\frac{1 + \lambda_i(\widehat{V}) + \cdots + \lambda_i(\widehat{V}^{q-1})}{q\lambda_i(\widehat{V}^{q-1})(1 + \lambda_i(\widehat{V}))} \right) \left(\sum_{i=1}^{n} \left(1 - \lambda_i(\widehat{V})^2\right)^2 \right)^{\frac{1}{2}} \\
&= \frac{1 + \lambda_{\min}(\widehat{V}) + \cdots + \lambda_{\min}^{q-1}(\widehat{V})}{q\lambda_{\min}^{q-1}(\widehat{V})(1 + \lambda_{\min}(\widehat{V}))} \|I - \widehat{V}^2\|_F \\
&= g(\lambda_{\min}(\widehat{V}))\|I - \widehat{V}^2\|_F \\
&\leq g(\frac{\overline{\sigma}}{\overline{\theta}})\|I - \widehat{V}^2\|_F,
\end{aligned} \tag{4.10}$$

where the next-to-last equality and the last inequality are due to the fact that $g(t) = \frac{t^{q-1}+t^{q-2}+\cdots+t+1}{qt^{q-1}(1+t)}$ is decreasing on $t > 0$. Substituting $\frac{\overline{\sigma}}{\overline{\theta}}$ and make some simplifications, we get

$$\begin{aligned}
g\left(\frac{\overline{\sigma}}{\overline{\theta}}\right) &= \frac{\left(\frac{\overline{\sigma}}{\overline{\theta}}\right)^{q-1} + \left(\frac{\overline{\sigma}}{\overline{\theta}}\right)^{q-2} + \cdots + \left(\frac{\overline{\sigma}}{\overline{\theta}}\right) + 1}{q\left(\frac{\overline{\sigma}}{\overline{\theta}}\right)^{q-1}\left(1 + \left(\frac{\overline{\sigma}}{\overline{\theta}}\right)\right)} \\
&= \frac{\left(\overline{\sigma}^{q-1} + \overline{\theta}\overline{\sigma}^{q-2} + \cdots + \overline{\sigma}\overline{\theta}^{q-2} + \overline{\theta}^{q-1}\right)}{q\overline{\sigma}^{q-1}(\overline{\sigma} + \overline{\theta})},
\end{aligned} \tag{4.11}$$

Furthermore, from (4.9) we have

$$\begin{aligned}
\|I - \widehat{V}^2\|_F &= \|I - \left(\frac{V_+}{\sqrt{1-\theta}}\right)^2\|_F \\
&= \frac{1}{1-\theta}\|(1-\theta)I - V_+^2\|_F \\
&\leq \frac{1}{1-\theta}\left(\theta\|I\|_F + \|I - V_+^2\|_F\right) \\
&\leq \frac{1}{1-\theta}\left(\theta\sqrt{n} + \left((q-1)^2 + 1\right)\sigma^2\right).
\end{aligned} \tag{4.12}$$

Substituting (4.11) and (4.12) into (4.10) yields the desired result.

To prove the last result of this lemma, let us define

$$h(\overline{\theta}) = \frac{\overline{\sigma}^{q-1} + \overline{\theta}\overline{\sigma}^{q-2} + \cdots + \overline{\sigma}\overline{\theta}^{q-2} + \overline{\theta}^{q-1}}{q\overline{\theta}\overline{\sigma}^{q-1}(\overline{\sigma} + \overline{\theta})} \text{ with } 0 < \overline{\theta} < 1 \text{ and } \overline{\sigma} > 0,$$

the function h is increasing on $0 < \overline{\theta} < 1$ with respect to $\overline{\theta}$, that is

$$h(\overline{\theta}) < h(1) = \frac{\overline{\sigma}^{q-1} + \overline{\sigma}^{q-2} + \cdots + \overline{\sigma} + 1}{q\overline{\sigma}^{q-1}(\overline{\sigma} + 1)}.$$

Now, let us consider the function $l(\overline{\sigma}) = h(1)$, which is decreasing on $]0, +\infty[$ with respect to $\overline{\sigma}$.

From the definition of $\overline{\sigma}$, $\overline{\sigma} = \sqrt{1-\sigma^2} > \sqrt{1 - \frac{4}{\left((q-1)^2+2\right)^2}} := \beta(q)$ and the increasing of $\beta(q)$ for $q \in \mathbb{N}$, we obtain

$$l(\overline{\sigma}) < l(\beta(q)) < l(\beta(\infty)) = l(1) = \frac{1}{2}.$$

Combining these results with the upper bound of $\sigma(X_+, S_+, \mu_+)$ given in (4.10), we obtain

$$\sigma(X_+, S_+, \mu_+) < \frac{1}{2\sqrt{1-\theta}}\left(\theta\sqrt{n} + \left((q-1)^2+1\right)\sigma^2\right).$$

Using $\theta = \frac{1}{((2q-1)^3-5q)\sqrt{n}}$ $(n \geq 1)$ and $\sigma < \frac{2}{(q-1)^2+2}$, we get

$$\begin{aligned}
\sigma(X_+, S_+, \mu_+) &< \frac{1}{2}\left(\frac{1}{\left((2q-1)^3-5q\right) + \frac{4\left((q-1)^2+1\right)}{\left((q-1)^2+2\right)^2}}\right)\sqrt{\frac{(2q-1)^3-5q}{\left((2q-1)^3-5q\right)-1}}\\
&= \frac{1}{2}\left(\frac{2}{(q-1)^2+2}\right)m(q),
\end{aligned}$$

where $m(q) = \left(\frac{(q-1)^2+2}{2\left((2q-1)^3-5q\right)} + \frac{2\left((q-1)^2+1\right)}{(q-1)^2+2}\right)\sqrt{\frac{(2q-1)^3-5q}{\left((2q-1)^3-5q\right)-1}}$, which is increasing for $q \in \mathbb{N}$, moreover $\lim_{q\to+\infty} m(q) = 2$. Consequently,

$$\sigma(X_+, S_+, \mu_+) < \frac{2}{(q-1)^2+2}.$$

Thus, we have completed the proof. □

In the following result, we give an upper bound of the total number of iterations produced by Algorithm 1.

Lemma 4.8 *Suppose that* (X^0, y^0, S^0) *is a strictly feasible point,* $\mu^0 = \frac{\langle X^0, S^0\rangle}{n}$ *and* $\sigma < \frac{2}{(q-1)^2+2}$. *Moreover, let* X^{k+1} *and* S^{k+1} *be the matrices obtained after* $(k+1)$*-th iterate generated by the Algorithm 1. Then, the inequality* $\langle X^{k+1}, S^{k+1}\rangle \leq \epsilon$ *is verified if*

$$k \geq \frac{1}{\theta}\left(\frac{\mu^0\left(n + \frac{4q(q-2)}{\left((q-1)^2+2\right)^2}\right)}{\epsilon}\right).$$

Proof Let $\mu := \mu^k$, then $\mu^k = (1-\theta)^k\mu^0$. From Lemma 4.6, we have

$$\langle X^{k+1}, S^{k+1}\rangle \leq \mu^k\left(n + \frac{4q(q-2)}{\left((q-1)^2+2\right)^2}\right) = (1-\theta)^k\mu^0\left(n + \frac{4q(q-2)}{\left((q-1)^2+2\right)^2}\right).$$

The inequality $\langle X^{k+1}, S^{k+1}\rangle \leq \epsilon$ holds if

$$(1-\theta)^k\mu^0\left(n + \frac{4q(q-2)}{\left((q-1)^2+2\right)^2}\right) \leq \epsilon.$$

Taking logarithms, we can write

$$k\log(1-\theta) \leq \log\epsilon - \log\left(\mu^0\left(n + \frac{4q(q-2)}{\left((q-1)^2+2\right)^2}\right)\right).$$

Using the inequality $\log(1-\theta) \leq -\theta$ for $\theta < 1$, we get

$$-k\theta \leq \log\epsilon - \log\left(\mu^0\left(n + \frac{4q(q-2)}{\left((q-1)^2+2\right)^2}\right)\right),$$

which implies that

$$k\theta \geq \log\left(\mu^0\left(n + \frac{4q(q-2)}{\left((q-1)^2+2\right)^2}\right)\right) - \log\epsilon = \left(\frac{\mu^0\left(n + \frac{4q(q-2)}{\left((q-1)^2+2\right)^2}\right)}{\epsilon}\right).$$

This completes the proof. □

Theorem 4.9 *Let* $\theta = \frac{1}{((2q-1)^3-5q)\sqrt{n}}$, *the Algorithm 1 gives an* ϵ*-optimal solution in at most*

$$\mathscr{O}\left(q^3\sqrt{n}\frac{\log\langle X^0, S^0\rangle}{\epsilon}\right)$$

iterations.

Proof Substituting the values of θ and μ^0 into the above lemma, the proof is completed. □

Corollary 4.10 *If we take* $X^0 = S^0 = I$, *then Algorithm* 1 *performs at most*

$$\mathscr{O}\left(q^3\sqrt{n}\frac{\log n}{\epsilon}\right)$$

iterations, which is the currently best-known iteration bound for small-update IPMs.

5 Numerical tests

In this section, we perform numerical tests to demonstrate the efficiency of Algorithm 1 and the influence of the parameter q. Since the choice of a strictly primal-dual feasible solution (X^0, y^0, S^0) is a crucial factor in any interior-point algorithm, we selected examples from the literature that provide an initial point. For instance, see [11,12]. These examples were implemented using MATLAB R2008b on an Intel Core i3 (1.80 GHz) processor with 4.00 GB of RAM.

We set $\epsilon = 10^{-4}$, $\mu^0 = \frac{\langle X^0, S^0\rangle}{n}$ and the update parameter θ takes the theoretical default value $\theta = \frac{1}{((2q-1)^3-5q)\sqrt{n}}$, where n represents the size of the example and $q \in \{2, 3, 4, 5, 6, 10\}$.

The primal and dual optimal solutions of (P) and (D) are denoted by X^* and (y^*, S^*), respectively. Finally, "Iter" and "CPU" refer to the total number of iterations and the elapsed time in seconds taken by the algorithm to obtain a primal-dual optimal solution, respectively.

The obtained results are compared with the function $\psi(t) = \sqrt{t}$ proposed by Dervay [7] using the theoretical default $\theta = \frac{1}{2\sqrt{n}}$, where n represents the size of the example.

Example 5.1 ([12]) For $m = 3$ and $n = 5$, we define $b = [-2, 2, -2]^T$,

$$A1 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2 & -1 \\ 0 & -1 & 1 & -1 & -2 \end{pmatrix}, \quad A2 = \begin{pmatrix} 0 & 0 & -2 & 2 & 0 \\ 0 & 2 & 1 & 0 & 2 \\ -2 & 1 & -2 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 2 \end{pmatrix},$$

$$A3 = \begin{pmatrix} 2 & 2 & -1 & -1 & 1 \\ 2 & 0 & 2 & 1 & 1 \\ -1 & 2 & 0 & 1 & 0 \\ -1 & 1 & 1 & -2 & 0 \\ 1 & 1 & 0 & 0 & -2 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 3 & -3 & 1 & 1 \\ 3 & 5 & 3 & 1 & 2 \\ -3 & 3 & -1 & 1 & 2 \\ 1 & 1 & 1 & -3 & -1 \\ 1 & 2 & 2 & -1 & -1 \end{pmatrix},$$

the strictly primale-duale feasible solution is given by:

$$X^0 = S^0 = I \text{ and } y^0 = [1, 1, 1]^T.$$

The primale-duale optimal solution obtained is

$$X^* = \begin{pmatrix} 0.0708 & -0.0719 & 0.0158 & 0.0673 & -0.1530 \\ -0.0719 & 0.0735 & -0.0177 & -0.0621 & 0.1661 \\ 0.0158 & -0.0177 & 0.0101 & -0.0095 & -0.0759 \\ 0.0673 & -0.0621 & -0.0095 & 0.1549 & 0.0096 \\ -0.153 & 0.1661 & -0.0759 & 0.0096 & 0.5951 \end{pmatrix},$$

$$\begin{pmatrix} 1.4337 & 0.5753 & -0.0294 & -0.4043 & 0.2169 \\ 0.5753 & 1.0956 & 0.3400 & 0.2169 & -0.1121 \\ -0.0294 & 0.3400 & 1.1875 & 0.2169 & 0.0478 \\ -0.4043 & 0.2169 & 0.2169 & 0.2832 & -0.1415 \\ 0.2169 & -0.1121 & 0.0478 & -0.1415 & 0.0957 \end{pmatrix} \text{ and } y^* = \begin{pmatrix} 0.8585 \\ 1.0937 \\ 0.7831 \end{pmatrix}.$$

The optimal value of both problems (P) and (D) is:

$$\langle C, X^* \rangle = b^T y^* = -1.0956.$$

Example 5.2 ([11]) For $n = 2m$, we define $b(i) = 2$, for $i = 1, \ldots, m$,

$$C = \begin{cases} -1, & \text{if } i = j \leq m, \\ 0, & \text{otherwise,} \end{cases}$$

and the matrices A_i, $i = 1, \ldots, m$ as follows

$$A_k(i, j) = \begin{cases} 1, & \text{if } i = j = k \text{ or } i = j = m + k, \\ 0, & \text{otherwise,} \end{cases}$$

the strictly primale-duale feasible solution is given by:

$$X^0(i, j) = \begin{cases} 2 - \gamma, & \text{if } i = j \leq m, \\ \gamma, & \text{if } i = j > m, \\ 0, & \text{otherwise,} \end{cases} \quad y^0(i) = \frac{-1}{\gamma}, \ i = 1, \ldots, m,$$

and

$$S^0(i, j) = \begin{cases} \frac{1}{\gamma} - 1, & \text{if } i = j \leq m, \\ \frac{1}{\gamma}, & \text{if } i = j > m, \\ 0, & \text{otherwise,} \end{cases}$$

where $\gamma = \frac{4 - \sqrt{8}}{2}$.

The primale-duale optimal solution obtained is:

$$X^*(i, j) = \begin{cases} 2, & \text{if } i = j \leq m, \\ 0, & \text{otherwise,} \end{cases}$$

$y^*(i) = -1$, for $i = 1, \ldots, m$, and

$$S^*(i, j) = \begin{cases} 1, & \text{if } i = j > m, \\ 0, & \text{otherwise,} \end{cases}$$

The optimal value for both problems (P) and (D) is:

$$\langle C, X^* \rangle = b^T y^* = -n.$$

Table 1 Total number of iterations and CPU time for Example 5.1

$\psi(t)$	$t^{\frac{q}{2}}$						$\sqrt{t}$
	q						
	2	3	4	5	6	10	
Itr	98	688	2041	4344	8099	43145	105
CPU	1.7676	10.5225	30.8077	64.7392	137.3579	653.7682	1.8762

Table 2 Total number of iterations and CPU time for Example 5.2

$n/\psi(t)$		$t^{\frac{q}{2}}$						$\sqrt{t}$
		q						
		2	3	4	5	6	10	
10	Itr	24	194	597	1195	2261	12051	64
	CPU	0.9018	6.7553	20.8365	41.6874	75.7552	392.8103	1.0842
20	Itr	28	212	641	1380	2428	12726	102
	CPU	7.0361	52.7200	146.5506	321.2198	604.3106	3.1766e+3	18.0082
30	Itr	32	249	712	1550	2918	14543	128
	CPU	42.7240	328.0901	937.6355	2.0451e+3	3.8245e+3	1.9512e+4	172.8469
40	Itr	40	290	819	1797	3319	17325	149
	CPU	164.2460	2.2779e+3	7.1651e+3	6.5732e+3	2.6352e+4	1.2140e+5	982.2364
50	Itr	56	398	1190	2432	4430	22673	174
	CPU	931.2640	398	2.3009e+4	4.5257e+4	7.0940e+4	5.0850e+5	7.0609e+3

Comments.

• From Tables 1 and 2, we conclude that the algorithm converges very slowly when the value of the update parameter θ is sufficiently small (close to zero). This observation is confirmed with the Default value of θ, which depends on the parameter q on the one hand and on the size of the problem n on the other.

• In all examples, the value $q = 2$ provides the best performance in terms of the number of iterations and CPU time.

In general, the numerical results obtained using the default barrier parameter θ, which is linked to the parameter q, are not good. As q increases, the algorithm's polynomial complexity is lost. This issue arises from the choice of the parametric function associated with the algebraic equivalent transformations (AETs). While introducing a parametric algebraic function along the central path offers a new class of search directions, it leads to numerical inefficiencies in large-scale problems, particularly for large values of q in small-update algorithms.

6 Conclusion

We have extended a primal-dual path-following interior-point algorithm from linear optimization to semidefinite optimization problems, building upon Darvay's technique [7] with a full Nesterov–Todd step. This technique relies on a specific class of AETs applied to the function $\psi(t) = \sqrt{t}$ for all $t > 0$. In this work, we have employed the function $\psi(t) = t^{\frac{q}{2}}$ with $q \in \mathbb{N}$ to derive a new set of search directions. Our analysis establishes that the resulting algorithm achieves the currently best-known iteration bound for small-update interior-point methods, namely, $\mathcal{O}\left(q^3\sqrt{n}\frac{\log n}{\epsilon}\right)$. Furthermore, the analysis closely parallels the approach used in the context of linear optimization [17]. To demonstrate the efficiency of the proposed method, we have provided some numerical results.

However, two crucial issues must be noted regarding Algorithm 1. First, finding an appropriate initial point is not an easy task and remains one of the primary drawbacks of interior-point methods, particularly when applied to large-scale problems. Second, the default barrier parameter θ which is linked to the parameter q. So as q becomes large the numerical results are not good, i.e., we loose the polynomial complexity of the algorithm. This drawback is the choice of the parametric function linked to the AETs. So introducing a parametric algebraic function on the central path certainly, offers a class of new search directions but on the other way is not good numerically for the large scale problems and for large values of the parameter q according to small-update algorithms.

Some interesting topics remain for further works. The extensions to circular cone optimization (CCO) and the elliptic cone optimization deserve to be investigated.

Acknowledgements The author would like to express sincere gratitude to the Editor-in-Chief and the anonymous referees for their valuable comments and suggestions, which have significantly improved the quality of this paper.

Declarations

Conflict of interest The author declares that there are no conflicts of interest regarding the publication of this paper.

References

1. M. Achache (2006) A new primal-dual path-following method for convex quadratic programming. Comp. Appl. Math., 25(1), 97-110
2. F. Alizadeh, D. Goldfarb (2003) Second-order cone programming. Math. Prog., 95(1), 3-51
3. Y. Q. Bai, G. Q. Wang (2007) A new pimal-dual interior-point algorithm for second-order cone optimization based on kernel function. Acta Math. Sin. (Engl. Ser.), 23(11), 2027-2042
4. Y. Q. Bai, G. Lesaja, C. Roos, G. Q. Wang, M. El Ghami (2008) A class of large-update and small-update primal-dual interior-point algorithms for linear optimization. J. optim. Theory Appl., 138(3), 341-359
5. Y. Q. Bai, M. El Ghami, C. Roos (2004) A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization. SIAM J. Optim., 15, 101-128
6. B. K. Choi, G. M. Lee (2009) On complexity analysis of the primal-dual interior-point method for semidefinite optimization problem based on a new proximity function. Nonlinear Anal., 71, 2628-2640
7. Zs. Darvay (2003) New interior-point algorithms in linear programming. Adv. Model. Optim., 5(1), 51-92
8. Zs. Darvay, T. Illés, P. R. Rogó (2022) Predictor-corrector interior-point algorithm for $P_*(\kappa)-$linear complementarity problems based on a new type of algebraic equivalent transformation technique. Eur. j. Oper. Res., 298(1), 25-35
9. Zs. Darvay, P. R. Takács (2018) New method for determining search directions for interior-point algorithms in linear optimization. Optim. Lett., 12, 1099-1116
10. W. Grimes, M. Achache (2023) A path-following interior-point algorithm for monotone LCP based on a modified newton search direction. RAIRO-Oper. Res., 57, 1059-1073
11. L. Guerra (2022) A class of new search directions for full-NT step feasible interior point method in semidefinite optimization. RAIRO-Oper. Res., 56, 3955-3971
12. L. Guerra, M. Achache (2020) A parametric kernel function generating the best known iteration bound for large-update methods for CQSDO. Stat. Optim. Inf. Comput., 8, 876-889
13. F. Gurtuna, C. Petra, F. A. Potra, O. Shevchenko, A. Vancea (2011) Corrector-predictor methods for sufficient linear complementarity problems. Comput. Optim. Appl., 48, 453-485
14. C. Helmberg, F. Rendl, R. J. Vanderbei, H. Wolkowicz (1996) An interior-point method for semidefinite programming. SIAM J. Optim., 6, 342-361
15. R. A. Horn, R. J. Charles (1991) Topics in Matrix Analysis. Cambridge University Press. Cambridge.
16. N. Karmarkar (1984) A new polynomial-time algorithm for linear programming. Combinatorica., 4, 373-395
17. B. Kheirfam, A. Nasrollahi (2018) A full-Newton step interior-point method based on a class of specific algebra transformation. Fundam. Inform., 163, 325-337
18. B. Kheirfam, A. Nasrollahi (2020) An extension for identifying search directions for interior-point methods in linear optimization. Asian-Eur. J. Math., 13(1) 2050014 (12 pages).
19. E. de Klerk (2002) Aspects of semidefinite programming: interior point algorithms and selected applications, Kluwer academic publishers, Dordrecht, The Netherlands
20. H. Lütkepohl (1996) Handbook of Matrices, Humboldt-Universität zu Berlin.
21. J. Peng, C. Roos, T. Terlaky (2002) Self-regular functions and new search directions for linear and semidefinite optimization. Math. Program., 93, 129-171
22. J. Peng, C. Roos, T. Terlaky (2002) Self-regularity: A new Paradigm for Primal-Dual Interior-Point Algorithms. Princeton University Press, Princeton.
23. C. Roos, T. Terlaky, J-Ph. Vial (2005) Theory and algorithms for linear optimization. Springer, New York.
24. U. Schäfer (2004) A linear complementarity problem with a P-matrix, SIAM Rev., 46(2), 189-201
25. M. J. Todd, K. C. Toh, Tütüncü (1998) On the nestrov-todd direction in semidefinite programming. SIAM J. Optim., 8, 769-796
26. I. Touil, D. Benterki, A. Yassine (2017) A feasible primal-dual interior point method for linear semidefinite programming, J. Comput. Appl. Math., 312, 216-230
27. H. Wolkowicz, R. Saigal, L. Vandenberghe (2000) Handbook of Semidefinite Programming, Theory, Algorithms, and Applications, Kluwer Academic Publishers, Boston.
28. G. Q. Wang, Y. Q. Bai (2009) A new primal-dual path-following interior-point algorithm for semidefinite optimization. J. Math. Anal. Appl., 353, 339-349
29. G. Q. Wang, Y. Q. Bai (2009) A primal-dual interior-point algorithm for second-order cone optimization with full Nesterov-Todd step. Appl. Math. & Comput., 215(3), 1047-1061
30. Y. Zhang (1998) On extending some primal-dual algorithms from linear programming to semidefinite programming. SIAM J. Optim., 8, 365-386

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.