Limiting behavior of the Alizadeh-Haeberly-Overton weighted paths in semidefinite programming

Zhaosong Lu* Renato D. C. Monteiro‡

July 24, 2003

Abstract

This paper studies the limiting behavior of weighted infeasible central paths for semidefinite programming obtained from centrality equations of the form $XS + SX = 2\nu W$, where $W$ is a fixed positive definite matrix and $\nu > 0$ is a parameter, under the assumption that the problem has a strictly complementary primal-dual optimal solution. It is shown that a weighted central path as a function of $\nu$ can be extended analytically beyond 0 and hence that the path and its derivatives converge as $\nu \downarrow 0$. Characterization of the limit points and first-order derivatives of the scaled weighted central path are also provided. We finally derive an error bound on the distance between a point lying in a certain neighborhood of the central path and the set of primal-dual optimal solutions.

Key words: limiting behavior, weighted central path, error bound, semidefinite programming.

AMS 2000 subject classification: 90C22, 90C25, 90C30

1 Introduction

Let $S^n$ denote the space of $n \times n$ real symmetric matrices. We consider the semidefinite programming (SDP) problem

\[
\begin{align*}
\text{minimize} & \quad C \cdot X \\
\text{(P)} & \quad \text{subject to} \quad AX = b, \quad X \succeq 0,
\end{align*}
\]

and its associated dual SDP problem

\[
\begin{align*}
\text{maximize} & \quad b^T y \\
\text{(D)} & \quad \text{subject to} \quad A^* y + S = C, \quad S \succeq 0,
\end{align*}
\]

---

*School of ISyE, Georgia Institute of Technology, Atlanta, Georgia 30332, USA (Email: zhaosong@isye.gatech.edu)
‡School of ISyE, Georgia Tech, Atlanta, Georgia 30332, USA (Email: monteiro@isye.gatech.edu). This author was supported in part by NSF Grants CCR-9902010, CCR-0203113 and INT-9910084 and ONR grant N00014-03-1-0401.
where the data consists of $C \in S^n$, $b \in \mathbb{R}^m$ and a linear operator $A : S^n \to \mathbb{R}^m$, the primal variable is $X \in S^n$, and the dual variable consists of $(S, y) \in S^n \times \mathbb{R}^m$. For a matrix $V \in S^n$, the notation $V \succeq 0$ means that $V$ is positive semidefinite. Given a fixed positive matrix $W \in S^n$, $\Delta b \in \mathbb{R}^m$ and $\Delta C \in S^n$, our interest in this paper is to study the set of solutions of the following system of nonlinear equations parametrized by the parameter $\nu > 0$:

$$AX = b + \nu \Delta b, \quad X \succeq 0,$$

$$A^* y + S = C + \nu \Delta C, \quad S \succeq 0,$$

$$XS + SX = 2\nu W.$$  (5)

Under suitable conditions on $(W, \Delta C, \Delta b)$, it has been shown in Monteiro and Zanjácomo [?] that the above system has a unique solution, denoted by $p(\nu) \equiv (X(\nu), S(\nu), y(\nu))$, for every $\nu > 0$.

Alizadeh et al. [?, ?] are the first to apply the map $(X, S) \in S^n \times S^n \to XS + SX$ to derive a search direction (called AHO direction) and use this direction to propose a primal-dual interior-point algorithm for SDP. Later on, Kojima et al. [?] propose a globally convergent predictor-corrector infeasible-interior-point algorithm for the monotone semidefinite linear complementarity problem using AHO direction, and show its quadratic local convergence under the strict complementarity assumption. Here, we refer to the path $\nu > 0 \to p(\nu)$ as the primal-dual AHO-weighted central path associated with $(P)$ and $(D)$. The main objective of this paper is to analyze the limiting behavior of this path as $\nu \downarrow 0$.

When $(W, \Delta C, \Delta b) = (I, 0, 0)$, the path $\nu > 0 \to p(\nu)$, as shown in Alizadeh et al. [?], is exactly the central path associated with $(P)$ and $(D)$. Properties of the central path have been extensively studied on several papers due to the important role it plays in the development of interior-point algorithms for cone programming, nonlinear programming and complementarity problems. Early works dealing with the well-definedness, differentiability and limiting behavior of weighted central paths in the context of the linear programming and the monotone complementarity problem include [?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?, ?].

Convergence of the central path for an SDP problem towards a primal-dual optimal solution has been established by Kojima et al. [?] (see also Halická et al. [?]) using a deep algebraic geometry result (see Lemma 3.1 of Milnor [?]). Characterization of the limit point of the central path has been obtained by Luo et al. [?] for SDP problems possessing strictly complementary primal-dual optimal solutions. Also, for this special class of SDP problems, Halická [?] has shown that the central path can be extended analytically “beyond” $\nu = 0$. For more general SDP problems, the above issues regarding the central path still remain open although some progress have been made on a few papers. These include Goldfarb and Scheinberg [?] who proved that the limit point of the central path must be a maximally complementary optimal solution and Halická et al. [?] and Sporre and Forsgren [?] who provided a partial characterization of the limit point of the central path as being the analytic center of some convex subset of the optimal solution set.

Generalization of the notion of weighted central paths from linear programming to SDP problems is a delicate issue. While for a linear programming a weighted central path can be characterized as optimal solutions of certain weighted logarithmic barrier problems, this characterizations does not seem to be a good source to obtain a suitable notion of weighted central paths for SDP. Instead, Monteiro and Zanjácomo [?] (see also Monteiro and Pang [?]) work directly with a system consisting of (3), (4) and an equation of the form $\Psi(X, S) = \nu W$, for some suitable map $\Psi : D \subseteq S^n \times S^n \to S^n$, and show that this system has a unique solution for every $\nu > 0$. Special instances of the map $\Psi$ for which the above result applies include the map $(X, S) \to (XS + SX)/2$ and $(X, S) \to X^{1/2} SX^{1/2}$.
Lu and Monteiro [?] have investigated the limiting behavior of the weighted central paths related to the second map above for the special class of SDPs possessing strictly complementary primal-dual optimal solutions. They have shown that a weighted central path as a function of \(\sqrt{\nu}\) can be extended analytically beyond 0.

In this paper, we will be interested in the first map above and its corresponding weighted central path, i.e. the path of solutions of systems of the form (3)-(5). More specifically, we will investigate the asymptotic properties of the weighted central paths for the special class of SDPs possessing strictly complementary primal-dual optimal solutions. We prove in Section 4 that the path \(\nu > 0 \rightarrow p(\nu)\) can be extended analytically beyond 0. As a consequence, we see that a weighted central path converges as \(\Theta(\nu)\). We also characterize the limit point and the first-order derivative of the scaled weighted central path as \(\nu \downarrow 0\). Finally, We derive in Section 5 an upper bound on the distance between a point lying in a certain neighborhood of the central path and the set of primal-dual optimal solutions.

The organization of this paper is as follows. Section 2 introduces the assumptions made throughout the paper. We discusses some properties about the weighted central paths in Section 3. Sections 4 and 5 establish the results mentioned in the previous paragraph.

1.1 Notation

The space of symmetric \(n \times n\) matrices will be denoted by \(S^n\). Given matrices \(X\) and \(Y\) in \(\mathbb{R}^{p \times q}\), the standard inner product is defined by \(X \cdot Y = \text{tr}(X^T Y)\), where \(\text{tr}(\cdot)\) denotes the trace of a matrix. The Euclidean norm and its associated operator norm, i.e., the spectral norm, are both denoted by \(|\cdot|\). The Frobenius norm of a \(p \times q\)-matrix \(X\) is defined as \(\|X\|_F = \sqrt{X \cdot X}\). Given a point \(f\) and a set \(F\) in a finite dimensional normed vector space, the distance from \(f\) to \(F\) is defined as \(\text{dist}(f, F) = \inf_{f \in F} \|f - f\|\). If \(X \in S^n\) is positive semidefinite (resp., definite), we write \(X \succeq 0\) (resp., \(X > 0\)). The cone of positive semidefinite (resp., definite) matrices is denoted by \(S^n_+\) (resp., \(S^n_{++}\)). Either the identity matrix or operator will be denoted by \(I\). The image (or range) space of a linear operator \(A\) will be denoted by \(\text{Im}(A)\); the dimension of the subspace \(\text{Im}(A)\), referred to as the rank of \(A\), will be denoted by \(\text{rank}(A)\). Given a linear operator \(F : E \rightarrow F\) between two finite dimensional inner product spaces \((E, \langle \cdot, \cdot \rangle_E)\) and \((F, \langle \cdot, \cdot \rangle_F)\), its adjoint is the unique operator \(F^* : F \rightarrow E\) satisfying \(\langle F(u), v \rangle_E = \langle u, F^*(v) \rangle_E\) for all \(u \in E\) and \(v \in F\).

If \(\{u(\nu) : \nu > 0\}\) and \(\{v(\nu) : \nu > 0\}\) are real sequences with \(v(\nu) > 0\), then \(u(\nu) = o(v(\nu))\) means that \(\lim_{\nu \rightarrow 0} u(\nu)/v(\nu) = 0\). Given functions \(f : \Omega \rightarrow E\) and \(g : \Omega \rightarrow \mathbb{R}_{++}\), where \(\Omega\) is an arbitrary set and \(E\) is a normed vector space, and a subset \(\Omega \subseteq \Omega\), we write \(f(w) = \Theta(g(w))\) for all \(w \in \Omega\) to mean that \(\|f(w)\| \leq M g(w)\) for all \(w \in \Omega\); moreover, for a function \(U : \Omega \rightarrow \mathbb{R}_{++}\), we write \(U(w) = \Theta(g(w))\) for all \(w \in \Omega\) if \(U(w) = O(g(w))\) and \(U(w)^{-1} = O(1/g(w))\) for all \(w \in \Omega\). The latter condition is equivalent to the existence of a constant \(M > 0\) such that

\[
\frac{1}{M} I \preceq \frac{1}{g(w)} U(w) \preceq MI, \quad \forall w \in \Omega.
\]

2 Preliminaries

In this section, we describe our assumptions that will be used throughout the paper. We also describe the weighted central path that will be the subject of our study in this paper. The conditions for its well-definedness are also stated.
Throughout this paper we will be dealing with the pair of dual SDPs (P) and (D) (see (1) and (2), respectively). Denote the feasible sets of (P) and (D) by \( \mathcal{F}_P \) and \( \mathcal{F}_D \), respectively. Throughout our presentation we make the following assumptions on the pair of problems (P) and (D).

A.1 \( A : \mathbb{S}^n \to \mathbb{R}^m \) is an onto linear operator;

A.2 There exists a pair of strictly complementary primal-dual optimal solution for (P) and (D), that is a triple \((X^*, S^*, y^*) \in \mathcal{F}_P \times \mathcal{F}_D\) satisfying \( X^* S^* = 0 \) and \( X^* + S^* \succ 0 \).

We will assume that Assumptions A.1 and A.2 are in force throughout our presentation. Hence, we will state our results without explicitly mentioning them.

Assumption A.1 is not really crucial for our analysis but it is convenient to ensure that the variables \( S \) and \( y \) are in one-to-one correspondence. We will see that the dual weighted central path can always be defined in the \( S \)-space. The goal of Assumption A.1 is just to ensure that this path can also be extended to the \( y \)-space.

Assumption A.2 is the one that is commonly used in the analysis of superlinear convergence of interior-point algorithms and it plays an important role in our analysis. In fact, it is a very challenging problem to generalize the analysis of this paper to the case where Assumption A.2 is dropped or simply relaxed.

By assumption A.2, since \( X^* S^* = S^* X^* = 0 \), we can diagonalize \( X^* \) and \( S^* \) simultaneously, i.e. find an orthonormal \( P \in \mathbb{R}^{n \times n} \) such that \( P^T X^* P \) and \( P^T S^* P \) are both diagonal. Performing the change of variables \( \hat{X} = P^T X P \) and \((\hat{S}, \hat{y}) = (P^T S P, y)\) on problems (P) and (D) yield another pair of primal and dual SDPs which has a primal-dual optimal solution \((\hat{X}^*, \hat{S}^*, \hat{y}^*)\) such that \( \hat{X}^* \) and \( \hat{S}^* \) are both diagonal. To simplify our notation, we will assume without loss of generality that the original (P) and (D) already have a primal-dual optimal solution \((X^*, S^*, y^*)\) such that

\[
X^* = \begin{bmatrix} \Lambda_B & 0 \\ 0 & 0 \end{bmatrix}, \quad S^* = \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_N \end{bmatrix}, \tag{6}
\]

where \( \Lambda_B \equiv \text{diag}(\lambda_1, \cdots, \lambda_K) \), \( \Lambda_N \equiv \text{diag}(\lambda_{K+1}, \cdots, \lambda_n) \) for some integer \( 0 \leq K \leq n \) and some scalars \( \lambda_i > 0, i = 1, 2, \cdots, n \). Here the subscripts \( B \) and \( N \) signify the “basic” and “nonbasic” subspaces (following the terminology of linear programming). Throughout this paper, the decomposition of any \( n \times n \) matrix \( V \) is always made with respect to the above partition \( B \) and \( N \), namely:

\[
V = \begin{bmatrix} V_B & V_{BN} \\ V_{NB} & V_N \end{bmatrix}.
\]

Notice that \( X \in \mathcal{F}_P \) is an optimal solution of (P) if and only if \( X S^* = 0 \). Hence, by assumption A.2, the primal optimal solution set \( \mathcal{F}_P^* \) is given by

\[
\mathcal{F}_P^* \equiv \{ X \in \mathcal{F}_P : X_{BN} = 0, X_{NB} = 0 \text{ and } X_N = 0 \}.
\]

Analogously, the dual optimal solution set \( \mathcal{F}_D^* \) is given by

\[
\mathcal{F}_D^* \equiv \{ (S, y) \in \mathcal{F}_D : S_{BN} = 0, S_{NB} = 0 \text{ and } S_B = 0 \}.
\]

Define the linear map \( \mathcal{G} : \mathbb{S}^n \times \mathbb{S}^n \times \mathbb{R}^m \to \mathbb{S}^n \times \mathbb{R}^m \) by

\[
\mathcal{G}(X, S, y) \equiv (A^* y + S - C, A X - b) \tag{7}
\]
and the set $\mathcal{G}_{++}$ by
\[ \mathcal{G}_{++} \equiv \mathcal{G}(\mathcal{S}_{++}^n \times \mathcal{S}_{++}^n \times \mathbb{R}^m). \] (8)

Given $(W, \Delta C, \Delta b) \in \mathcal{S}^n \times \mathcal{S}^n \times \mathbb{R}^m$, in this paper we are interested in the solutions of the system of nonlinear equations (3)-(5) parametrized by the parameter $\nu > 0$. The following result gives condition on $(W, \Delta C, \Delta b)$ for system (3)-(5) to have a unique solution for each $\nu > 0$.

**Proposition 2.1** Assume that $(W, \Delta C, \Delta b) \in \mathcal{S}_{++}^n \times \mathcal{G}_{++}$. Then, for any $\nu > 0$, the system (3)-(5) has a unique solution, denoted by $(X(\nu), S(\nu), y(\nu))$. Moreover, the path $\nu > 0 \rightarrow (X(\nu), S(\nu), y(\nu))$ is analytic.

**Proof.** This proposition follows as a direct consequence of Corollary 1 of Monteiro and Zanjácomo [?] applied to the central path map considered in Example 1 of Subsection 2.2 of [?].

For a given $(W, \Delta C, \Delta b) \in \mathcal{S}_{++}^n \times \mathcal{G}_{++}$, the path $\nu > 0 \rightarrow (X(\nu), S(\nu), y(\nu))$ will be referred to as the AHO-weighted central path. For shortness, we also call it the weighted central path. In view of the above proposition, we will assume throughout Sections 3 and 4 that the following condition is true.

**A.3** $(W, \Delta C, \Delta b) \in \mathcal{S}_{++}^n \times \mathcal{G}_{++}$.

## 3 Properties of AHO-weighted central path

We will introduce some properties of the weighted central path. The sizes of the weighted path in terms of blocks are given in this section. The analysis of the limiting behavior of the weighted central path strongly relies on those results. Although Kojima et al. in section 5 of [?] have essentially derived all those results, we derive them here in a different and simpler approach with the aid of Hoffman Lemma [?].

The following result gives some estimates on the sizes of the blocks of $X(\nu)$ and $S(\nu)$.

**Lemma 3.1** For all $\nu > 0$ sufficiently small, we have:
\[
\begin{align*}
X_B(\nu) &= \Theta(1), & X_N(\nu) &= \Theta(\nu), \\
S_B(\nu) &= \Theta(\nu), & S_N(\nu) &= \Theta(1), \\
X_{BN}(\nu) &= \mathcal{O}(\sqrt{\nu}), & S_{BN}(\nu) &= \mathcal{O}(\sqrt{\nu}).
\end{align*}
\] (9)-(11)

As a consequence, the weighted central path $\{(X(\nu), y(\nu), S(\nu))\}$ is bounded and any its accumulation point as $\nu \downarrow 0$ is a strictly complementary primal-dual optimal solution of $(P)$ and $(D)$.

**Proof.** Following the same proof as Lemma 2.2 of Lu and Monteiro [?], we obtain that
\[
X_B(\nu) = \mathcal{O}(1), \quad S_N(\nu) = \mathcal{O}(1), \quad X_N(\nu) = \mathcal{O}(\nu), \quad S_B(\nu) = \mathcal{O}(\nu)
\] (12)

and (11) holds. By Lemma 3.3 of Monteiro [?], we have
\[
\lambda_{\min}(X^{1/2}S X^{1/2}) \geq \lambda_{\min}(SX + XS)/2
\]

5
where $\lambda_{\min}(A)$ is the smallest eigenvalue of $A$. Noting that $\lambda_{\min}(SX + XS)/2 = \nu \lambda_{\min}(W)$, 
we have $X^{1/2}SX^{1/2} \succeq \nu \lambda_{\min}(W)I$. Using this fact together with (12), we can easily follow the proof of Lemma 3.2 of Luo et al. [?] to derive (9) and (10), except that the identity matrix $I$ should be replaced by $\lambda_{\min}(W)I$ throughout their proof.

The following lemma establishes the relationship between $X_{BN}(\nu)$ and $S_{BN}(\nu)$.

**Lemma 3.2** For all $\nu > 0$ sufficiently small, we have:

\[
\|S_{BN}(\nu)\| = \Theta(\|X_{BN}(\nu)\|) + O(\nu), \\
-X_{BN}(\nu) \bullet S_{BN}(\nu) = \Theta(\|X_{BN}(\nu)\|^2) + O(\nu\|X_{BN}(\nu)\|).
\]

**Proof.** For notational convenience, we will use $X$ and $S$ to denote $X(\nu)$ and $S(\nu)$. Let $V = XS$. Using Lemma 3.1, we obtain that

\[
\|V_{NB}\| = \|X_{NB}S_B + X_NS_{NB}\| = O(\nu^{3/2}).
\]

Using (5) and the fact $V = XS$, we have

\[
X_B S_{BN} + X_{BN} S_N + V_{NB}^T = 2\nu W_{BN},
\]

which implies that

\[
S_{BN} = -X_B^{-1}(X_{BN} S_N + V_{NB}^T - 2\nu W_{BN}).
\]

Using this identity, (13) and Lemma 3.1, we see that the first conclusion follows. Again, using them, we obtain that

\[
X_{BN} \bullet S_{BN} = -\text{tr}(X_{BN}^T X_B^{-1} X_{BN} S_N) + O(\nu\|X_{BN}\|),
\]

\[
= -\|X_B^{-1/2} X_{BN} S_N^{1/2}\|_F^2 + O(\nu\|X_{BN}\|),
\]

which together with Lemma 3.1 implies the second conclusion.

We will improve the estimate on the sizes of $X_{BN}(\nu)$ and $S_{BN}(\nu)$ in the following lemma.

**Lemma 3.3** For all $\nu > 0$ sufficiently small, we have:

\[
X_{BN}(\nu) = O(\nu), \quad S_{BN}(\nu) = O(\nu).
\]

**Proof.** Suppose that $X_{BN}(\nu) = O(\nu)$ does not hold. Then there exits a positive sequence $\nu_k \downarrow 0$ as $k \to \infty$ such that $\nu_k = o(\|X_{BN}(\nu_k)\|)$. For convenience, we omit the index $k$ from $\nu_k$ throughout the remaining proof. Then the above identity can be written as $\nu = o(\|X_{BN}(\nu)\|)$, which together with Lemma 3.2 implies that

\[
\|S_{BN}(\nu)\| = \Theta(\|X_{BN}(\nu)\|), \quad (14)
\]

\[
-X_{BN}(\nu) \bullet S_{BN}(\nu) = \Theta(\|X_{BN}(\nu)\|^2). \quad (15)
\]

For any $\nu > 0$, consider the linear system

\[
A(X - X(\nu)) = -\nu \Delta b, \\
X_{BN} - X_{BN}(\nu) = -X_{BN}(\nu), \\
X_N - X_N(\nu) = -X_N(\nu).
\]

(16)
We see that any $X^* \in \mathcal{F}_P^*$ is a feasible solution to this system. Hence, by Hoffman Lemma [7], there exists a sufficiently large constant $\hat{C}$ (independent on $\nu$) such that for any $\nu > 0$, this system has a solution $\tilde{X} \in \mathcal{S}^n$, which satisfies
\[
\|\tilde{X} - X(\nu)\| \leq \hat{C}(\nu\|\Delta b\| + \|X_N(\nu)\| + \|X_B(\nu)\|).
\]
(18)

Analogously, for any $\nu > 0$, there exists $(\tilde{S}, g) \in \mathcal{S}^m \times \mathbb{R}^m$ which satisfies
\[
\mathcal{A}^*(y - g(\nu)) + S - S(\nu) = -\nu \Delta C,
\]
(19)
\[
S_B - S_B(\nu) = -S_B(\nu),
\]
(20)
\[
\hat{C}(\nu\|\Delta C\| + \|S_B(\nu)\| + \|S_B(\nu)\|) \geq \|S - S(\nu)\|.
\]
(21)

Noticing that $(\Delta C, \Delta b) \in \mathcal{G}_{++}$, we have $\Delta C = \mathcal{A}^*y^0 + S^0 - C$ and $\Delta b = \mathcal{A}X^0 - b$ for some $(X^0, S^0, y^0) \in \mathcal{S}^n_+ \times \mathcal{S}^m_+ \times \mathbb{R}^m$. We easily see that, for any given $(X^*, S^*, g^*) \in \mathcal{F}_P^* \times \mathcal{F}_D$, 
\[
\mathcal{A}((\tilde{X} - X(\nu) + \nu(X^0 - X^*)) \bullet (\tilde{S} - S(\nu) + \nu(S^0 - S^*)) \in \text{Im}(\mathcal{A}^*).
\]

Hence, we obtain that
\[
(\tilde{X} - X(\nu) + \nu(X^0 - X^*)) \bullet (\tilde{S} - S(\nu) + \nu(S^0 - S^*)) = 0.
\]
(22)

Note that $\|\tilde{X}_B - X_B(\nu)\| \leq \|\tilde{X} - X(\nu)\|$ and $\|\tilde{S}_N - S_N(\nu)\| \leq \|\tilde{S} - S(\nu)\|$. Using this fact, (22), (16), (17), (19), (20), (18), (21), (14), (15) and Lemma 3.1, we obtain that, for all $\nu > 0$ sufficiently small,
\[
|X_B(\nu) \bullet S_B(\nu)| \leq \bar{C} \nu (\|\tilde{X} - X(\nu)\| + \|\tilde{S} - S(\nu)\|),
\]
\[
\leq \bar{C} \nu (\|X_B(\nu)\| + \|S_B(\nu)\|),
\]
\[
\leq \bar{C} \nu + (|X_B(\nu) \bullet S_B(\nu)|)^{1/2},
\]

where $\tilde{C}, \bar{C}$ and $\bar{C}$ are some constants (independent on $\nu$) and the last inequality follows from (14) and (15). Let $\xi = (|X_B(\nu) \bullet S_B(\nu)|)^{1/2}$. From the last inequality above, we have $\xi^2 \leq \bar{C} \nu + \xi$, which together with the fact $\xi > 0$ implies $\xi \leq (\bar{C} + (\bar{C})^{1/2})\nu/2$. Hence, $\xi = O(\nu)$. Using this result and (15), we obtain $\|X_B(\nu)\| = O(\nu)$, which contradicts with the assumption $\nu = o(\|X_B(\nu)\|)$. Therefore, $X_B(\nu) = O(\nu)$ holds. The proof of $S_B(\nu) = O(\nu)$ directly follows from Lemma 3.2. ■

4 Analyticity of the AHO-weighted central path

In this section we will show that the weighted central path can be extended analytically to $\nu = 0$. We also characterize the limit point and the first-order derivative of the scaled weighted central path as $\nu \downarrow 0$.

For the sake of shortness, it is convenient to introduce the following definition.

**Definition 1** Let $w : (0, \infty) \to E$ be a given function where $E$ is a finite dimensional normed vector space. The function $w$ is said to be analytic at 0 if there exist $\epsilon > 0$ and an analytic function $\psi : (-\epsilon, \epsilon) \to E$ such that $w(t) = \psi(t)$ for all $t \in (0, \epsilon)$. 

7
The following theorem is one of the main results of this section. Its proof will be given at the end of this section.

**Theorem 4.1** The weighted central path \( \nu > 0 \to (X(\nu), S(\nu), y(\nu)) \) is is analytic and also analytic at \( \nu = 0 \). Consequently, the weighted path and all its \( k \)-th order derivatives, \( k \geq 1 \), converge as \( \nu \downarrow 0 \).

A key step towards showing the above result is a reformulation of the weighted central path system as we now discuss. We first define

\[
D_N = \begin{bmatrix} I & 0 \\ 0 & I/\nu \end{bmatrix}.
\]

Then the equation \( XS + SX = 2\nu W \) can be rewritten as

\[
XD_N D_N^{-1} S + SD_N^{-1} D_N X = 2\nu W. \tag{23}
\]

Now, let

\[
U^n = \{ U \in \mathbb{R}^{n \times n} : U_B \in \mathcal{S}^{[B]}, \, U_N \in \mathcal{S}^{[N]}, \, U_{NB} = 0 \},
\]

\[
U^n_{++} = \{ U \in U^n : U_B > 0, \, U_N > 0 \}.
\]

and define \( \mathcal{L} : U^n \to \mathbb{R}^{n \times n} \) as

\[
\mathcal{L}(U) = \begin{bmatrix} 0 & 0 \\ U_{BN}^T & 0 \end{bmatrix}, \quad \forall U \in U^n.
\]

Given any \((X, S) \in \mathcal{S}^n_{++} \times \mathcal{S}^n_{++}\), we define \((\bar{U}, \bar{V}, \bar{X}, \bar{S}) \in \mathcal{U}^n_{++} \times \mathcal{U}^n_{++} \times \mathcal{S}^n \times \mathcal{S}^n\) as follows

\[
\bar{X} = \begin{bmatrix} X_B & X_{BN}/\nu \\ X_{NB}/\nu & X_N/\nu \end{bmatrix}, \quad \bar{S} = \begin{bmatrix} S_B/\nu & S_{BN}/\nu \\ S_{NB}/\nu & S_N \end{bmatrix}, \tag{24}
\]

\[
\bar{U} = \begin{bmatrix} \tilde{X}_B & \tilde{X}_{BN} \\ 0 & \tilde{X}_N \end{bmatrix}, \quad \bar{V} = \begin{bmatrix} \tilde{S}_B & \tilde{S}_{BN} \\ 0 & \tilde{S}_N \end{bmatrix}. \tag{25}
\]

Then we immediately obtain that

\[
XD_N = \begin{bmatrix} \bar{X}_B & \bar{X}_{BN} \\ \nu \bar{X}_N & \bar{X}_N \end{bmatrix} = \bar{U} + \nu \mathcal{L}(\bar{U}) = U(\bar{U}),
\]

\[
\frac{1}{\nu} D_N^{-1} S = \begin{bmatrix} \bar{S}_B & \bar{S}_{BN} \\ \nu \bar{S}_N & \bar{S}_N \end{bmatrix} = \bar{V} + \nu \mathcal{L}(\bar{V}) = U(\bar{V})
\]

where \( U(\nu) \equiv I + \nu \mathcal{L} \). In view of the above identities, for \( \nu > 0 \), (23) is equivalent to

\[
U_\nu(\bar{U}) U_\nu(\bar{V}) + \left( U_\nu(\bar{U}) U_\nu(\bar{V}) \right)^T = 2W. \tag{26}
\]

Accordingly, define \((\bar{U}(\nu), \bar{V}(\nu), \bar{X}(\nu), \bar{S}(\nu))\) with replacing \((X, S)\) in (24) and (25) by \((X(\nu), S(\nu))\), respectively. Proposition 2.1 and the above arguments establish the following key result.
Proposition 4.2 Let \((X^*, S^*, y^*) \in \mathcal{F}_B^* \times \mathcal{F}_B^*\) be given. Then, for every \(\nu > 0\), \((\tilde{U}(\nu), \tilde{V}(\nu), \tilde{X}(\nu), \tilde{S}(\nu))\) is a solution of the system defined by (25), (26) and the linear equations

\[
A \cdot \begin{bmatrix}
\tilde{X}_B - X^*_B & \nu \tilde{X}_{BN} \\
\nu \tilde{X}_{NB} & \nu \tilde{X}_N
\end{bmatrix} = \nu \Delta b, \\
\begin{bmatrix}
\nu \tilde{S}_B & \nu \tilde{S}_{BN} \\
\nu \tilde{S}_{NB} & \tilde{S}_N - S^*_N
\end{bmatrix} \in \nu \Delta C + \text{Im}(A^*).
\]

Moreover, it is analytic for all \(\nu > 0\).

The next result states some basic properties about the accumulation points of \((\tilde{U}(\nu), \tilde{V}(\nu), \tilde{X}(\nu), \tilde{S}(\nu))\) as \(\nu\) approaches 0.

Lemma 4.3 The path \(\nu > 0 \rightarrow (\tilde{U}(\nu), \tilde{V}(\nu), \tilde{X}(\nu), \tilde{S}(\nu))\) remains bounded as \(\nu\) approaches 0 and any accumulation point \((\tilde{U}^*, \tilde{V}^*, \tilde{X}^*, \tilde{S}^*)\) of this path as \(\nu\) approaches 0 is in \(\mathcal{U}_+^n \times \mathcal{U}_+^n \times \mathcal{S}^n \times \mathcal{S}^n\) and satisfies

\[
\tilde{U}^* = \begin{bmatrix}
\tilde{X}_B^* & \tilde{X}_{BN}^* \\
0 & \tilde{X}_N^*
\end{bmatrix}, \quad \tilde{V}^* = \begin{bmatrix}
\tilde{S}_B^* & \tilde{S}_{BN}^* \\
\tilde{S}_{NB}^* & \tilde{S}_N^*
\end{bmatrix},
\]

\[
\tilde{U}^* \tilde{V}^* + (\tilde{U}^* \tilde{V}^*)^T = 2W.
\]

Proof. Relation (24), Lemma 3.1 and Lemma 3.3 imply that \((\tilde{X}(\nu), \tilde{S}(\nu))\) remains bounded as \(\nu\) approaches 0. So does \((\tilde{U}(\nu), \tilde{V}(\nu))\) according to this fact and relation (25). Using (24) and Lemma 3.1, we see that \((\tilde{U}^*, \tilde{V}^*) \in \mathcal{U}_+^n \times \mathcal{U}_+^n\). The remaining proof follows directly from (25) and (26).

Our next goal is to show that the path \(\nu > 0 \rightarrow (\tilde{U}(\nu), \tilde{V}(\nu), \tilde{X}(\nu), \tilde{S}(\nu))\) is analytic at \(\nu = 0\). The basic tool we use to establish this fact is the implicit function theorem applied to a specific system of equations. A first natural candidate for such a system seems to be the one given by (25), (26), (27) and (28). However, the main drawback of this system is that its derivative with respect to \((\tilde{U}, \tilde{V}, \tilde{X}, \tilde{S})\) is generally singular for \(\nu = 0\) (even though for \(\nu > 0\) it is always nonsingular). The main cause for this phenomenon is that the “rank” of the linear equations (27) and (28) changes when \(\nu\) becomes 0.

We will now show how the linear equations (27) and (28) can be reformulated into equivalent linear equations whose rank is constant for every \(\nu \in \mathbb{R}\). First note that the linear operator \(A : S^m \rightarrow \mathbb{R}^m\) can be expressed as

\[
A(X) = A_B(X_B) + A_{BN}(X_{BN}) + A_N(X_N) \equiv (A_B A_{BN} A_N) \begin{bmatrix} X_B \\ X_{BN} \\ X_N \end{bmatrix},
\]

for some linear operators \(A_B : S^{|B|} \rightarrow \mathbb{R}^m, A_{BN} : \mathbb{R}^{|B| \times |N|} \rightarrow \mathbb{R}^m\) and \(A_N : S^{|N|} \rightarrow \mathbb{R}^m\).

A well-known result from linear algebra says that any matrix can be put into row-echelon form after a sequence of elementary row operations. A similar type of argument allows one to establish the following result.
Lemma 4.4 Let $A : S^n \rightarrow \mathbb{R}^m$ be an onto linear operator. Assume that

$$i_1 = \text{rank}(A_B), \quad i_2 = \text{rank}(A_B A_{BN}) - i_1, \quad i_3 = \text{rank}(A) - (i_1 + i_2) = m - (i_1 + i_2).$$

Then there exists an isomorphism $T : \mathbb{R}^m \rightarrow \mathbb{R}^m$ such that

$$(T \circ A)(X) = \begin{pmatrix} A_{11}(X_B) + A_{12}(X_{BN}) + A_{13}(X_N) \\ A_{22}(X_{BN}) + A_{23}(X_N) \\ A_{33}(X_N) \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{pmatrix} \begin{pmatrix} X_B \\ X_{BN} \\ X_N \end{pmatrix},$$

for some linear operators

$$A_{11} : S^{[B]} \rightarrow \mathbb{R}^{i_1}, \quad A_{12} : \mathbb{R}^{[B] \times [N]} \rightarrow \mathbb{R}^{i_1}, \quad A_{13} : S^{[N]} \rightarrow \mathbb{R}^{i_2},$$

$$A_{22} : \mathbb{R}^{[B] \times [N]} \rightarrow \mathbb{R}^{i_2}, \quad A_{33} : S^{[N]} \rightarrow \mathbb{R}^{i_3}$$

such that rank($A_{11}$) = $i_1$, rank($A_{22}$) = $i_2$, rank($A_{33}$) = $i_3$.

We can now reformulate the linear system (27) with the use of Lemma 4.4 as follows. Using Lemma 4.4, we easily see that (27) is equivalent to the linear system

$$\begin{pmatrix} A_{11} & \nu A_{12} & \nu A_{13} \\ 0 & \nu A_{22} & \nu A_{23} \\ 0 & 0 & \nu A_{33} \end{pmatrix} \begin{pmatrix} X_B - X_B^* \\ X_{BN} \\ X_N \end{pmatrix} = \nu \begin{pmatrix} \Delta b_1 \\ \Delta b_2 \\ \Delta b_3 \end{pmatrix},$$

where $(\Delta b_1, \Delta b_2, \Delta b_3) \in \mathbb{R}^{i_1} \times \mathbb{R}^{i_2} \times \mathbb{R}^{i_3}$ and $\Delta b \equiv T(\Delta b)$. Dividing the second and third blocks of rows in the above system by $\nu$, respectively, we obtain the following system

$$\begin{pmatrix} A_{11} & \nu A_{12} & \nu A_{13} \\ 0 & A_{22} & A_{23} \\ 0 & 0 & A_{33} \end{pmatrix} \begin{pmatrix} X_B - X_B^* \\ X_{BN} \\ X_N \end{pmatrix} = \nu \begin{pmatrix} \Delta b_1 \\ \Delta b_2 \\ \Delta b_3 \end{pmatrix}. \quad (32)$$

Note that the linear system (32) is equivalent to (27) for every $\nu > 0$. Hence, $\nabla(\nu)$ satisfies (32) for every $\nu > 0$. A nice feature of (32) is that the operator on its left hand side does not loose full rankness as $\nu$ becomes 0. We state this fact in the following proposition.

Proposition 4.5 Let $A_\nu : S^n \rightarrow \mathbb{R}^m$ be the operator defined on the left hand side of (32). Then, rank($A_\nu$) = $m$ for every $\nu \in \mathbb{R}$.

The linear system (28) can also be reformulated with the aid of Lemma 4.4 as follows. First note that by Lemma 4.4 we have

$$\text{Im}(A^*) = \text{Im} \left[ (T \circ A)^* \right] = \text{Im} \left[ \begin{pmatrix} A_{11}^* & 0 & 0 \\ A_{12}^* & A_{22}^* & 0 \\ A_{13}^* & A_{23}^* & A_{33}^* \end{pmatrix} \right] = \text{Im} \left[ \begin{pmatrix} \nu A_{11}^* & 0 & 0 \\ \nu A_{12}^* & \nu A_{22}^* & 0 \\ \nu A_{13}^* & \nu A_{23}^* & A_{33}^* \end{pmatrix} \right].$$
for every $\nu > 0$. Hence, for every $\nu > 0$ (28) is equivalent to
\[
\begin{pmatrix}
\nu \bar{S}_B \\
\nu \bar{S}_{BN} \\
\bar{S}_N - \bar{S}_N^*
\end{pmatrix}
\in \nu
\begin{pmatrix}
\Delta C_B \\
\Delta C_{BN} \\
\nu \Delta C_N
\end{pmatrix}
+ \text{Im}
\begin{bmatrix}
\nu A_{11}^* & 0 & 0 \\
\nu A_{12}^* & \nu A_{22}^* & 0 \\
\nu A_{13}^* & \nu A_{23}^* & \nu A_{33}^*
\end{bmatrix}.
\]
Dividing the first and second block of rows in the above system by $\nu$, respectively, we obtain the system
\[
\begin{pmatrix}
\bar{S}_B \\
\bar{S}_{BN} \\
\bar{S}_N - \bar{S}_N^*
\end{pmatrix}
\in \begin{pmatrix}
\Delta C_B \\
\Delta C_{BN} \\
\nu \Delta C_N
\end{pmatrix}
+ \text{Im}
\begin{bmatrix}
A_{11}^* & 0 & 0 \\
A_{12}^* & A_{22}^* & 0 \\
A_{13}^* & A_{23}^* & A_{33}^*
\end{bmatrix}, \quad (33)
\]
which is equivalent to (28) for every $\nu > 0$, and hence satisfied by $\bar{S}(\nu)$ for all $\nu > 0$. Let the operator
$B_\nu : \mathbb{R}^m \rightarrow \mathbb{S}^n$ be defined such that the last term on the right hand side of (33) is written as $\text{Im}(B_\nu)$.
Using the definition of $A_\nu$ and $B_\nu$, we conclude that (32) and (33) hold if and only if there exists $\hat{y}$ such that
\[
A_\nu(\bar{X} - X^*) = \begin{pmatrix}
\frac{\nu \tilde{b}_1}{\tilde{b}_2} \\
\frac{\nu \tilde{b}_3}{\Delta b_3}
\end{pmatrix}, \quad B_\nu \hat{y} + (\bar{S} - S^*) = \begin{pmatrix}
\Delta C_B \\
\Delta C_{BN} \\
\nu \Delta C_N
\end{pmatrix}. \quad (34)
\]
We have the following result whose proof follows straightforwardly from the above arguments.

**Proposition 4.6** There exists a curve $\hat{y} : \mathbb{R}_{++} \rightarrow \mathbb{R}^m$ such that $(\bar{U}(\nu), \bar{V}(\nu), \bar{X}(\nu), \bar{S}(\nu), \hat{y}(\nu))$ is a solution of (25), (26) and (34) in $U^1_{++} \times U^1_{++} \times \mathbb{S}^n \times \mathbb{S}^n \times \mathbb{R}^m$ for every $\nu > 0$. Moreover, the path $\nu > 0 \rightarrow (\bar{U}(\nu), \bar{V}(\nu), \bar{X}(\nu), \bar{S}(\nu), \hat{y}(\nu))$ remains bounded as $\nu$ approaches 0 and any of its accumulation points is in $U^1_{++} \times U^1_{++} \times \mathbb{S}^n \times \mathbb{S}^n \times \mathbb{R}^m$.

The path $\nu > 0 \rightarrow (\bar{U}(\nu), \bar{V}(\nu), \bar{X}(\nu), \bar{S}(\nu), \hat{y}(\nu))$ is referred as the scaled weighted central path. The system formed by (25), (26) and (34) is the one which we will use to establish that the scaled weighted central path is analytic at $\nu = 0$. This would follow by the implicit function theorem if we can establish that the Jacobian of this system with $\nu = 0$ with respect to $(\bar{U}, \bar{V}, \bar{X}, \bar{S}, \hat{y})$ is nonsingular as long as $(\bar{U}, \bar{V}) \in U^1_{++} \times U^1_{++}$. The nonsingularity of this Jacobian can be easily seen to be equivalent to showing that $(\bar{\Delta U}, \bar{\Delta V}, \bar{\Delta X}, \bar{\Delta S}, \bar{\Delta \hat{y}}) = (0, 0, 0, 0, 0)$ is the only solution of the following linear system:
\[
\bar{\Delta U} = \begin{bmatrix}
\Delta \bar{X}_B & \Delta \bar{X}_{BN} \\
0 & \Delta \bar{X}_N
\end{bmatrix}, \quad \bar{\Delta V} = \begin{bmatrix}
\Delta \bar{S}_B & \Delta \bar{S}_{BN} \\
0 & \Delta \bar{S}_N
\end{bmatrix}.
\]
\[
\bar{\Delta U}V + \bar{U}\bar{\Delta V} + (\bar{\Delta U}V + \bar{U}\bar{\Delta V})^T = 0,
\]
\[
A_0 \bar{\Delta X} = 0, \quad B_0 \bar{\Delta \hat{y}} + \bar{\Delta S} = 0. \quad (37)
\]

**Lemma 4.7** Assume that $(\bar{U}, \bar{V}) \in U^1_{++} \times U^1_{++}$. Then, the system (35)-(37) has $(\bar{\Delta U}, \bar{\Delta V}, \bar{\Delta X}, \bar{\Delta S}, \bar{\Delta \hat{y}}) = (0, 0, 0, 0, 0)$ as its unique solution.

**Proof.** Using the definition of $A_{\nu}$ and $B_{\nu}$, we see that the equations in (37) are, respectively,
\[
\begin{pmatrix}
A_{11} & 0 & 0 \\
0 & A_{22} & A_{23} \\
0 & 0 & A_{33}
\end{pmatrix} \bar{\Delta X} = 0
\]
and
\[
\begin{pmatrix}
A_{11}^* & 0 & 0 \\
A_{12}^* & A_{22}^* & 0 \\
0 & 0 & A_{33}^*
\end{pmatrix}
(\widetilde{\Delta y}) + \widetilde{\Delta S} = 0.
\] (39)

?From the two equations above, we easily see that
\[
\widetilde{\Delta X}_B \bullet \widetilde{\Delta S}_B = 0, \quad \widetilde{\Delta X}_N \bullet \widetilde{\Delta S}_N = 0.
\]

Further, in view of (35), we obtain that
\[
\widetilde{\Delta U}_B \bullet \widetilde{\Delta V}_B = 0, \quad \widetilde{\Delta U}_N \bullet \widetilde{\Delta V}_N = 0. \tag{40}
\]

Using the fact that \((\widetilde{U}, \widetilde{V}) \in \mathcal{U}_+^n \times \mathcal{U}_+^n\) and \((\widetilde{\Delta U}, \widetilde{\Delta V}) \in \mathcal{U}_+^n \times \mathcal{U}_+^n\), we have \(\widetilde{\Delta UV} + \widetilde{U}\Delta \widetilde{V} \in \mathcal{U}_+^n\), which together with (36) implies that
\[
\widetilde{\Delta UV} + \widetilde{U}\Delta \widetilde{V} = 0.
\]

This equation can be written as
\[
\begin{align*}
\Delta \widetilde{U}_B \widetilde{V}_B + \Delta \widetilde{U}_B \Delta \widetilde{V}_B &= 0, \tag{41} \\
\Delta \widetilde{U}_N \widetilde{V}_N + \Delta \widetilde{U}_N \Delta \widetilde{V}_N &= 0, \tag{42} \\
\Delta \widetilde{U}_B \widetilde{V}_{BN} + \Delta \widetilde{U}_{BN} \widetilde{V}_B + \Delta \widetilde{U}_B \Delta \widetilde{V}_{BN} + \Delta \widetilde{U}_{BN} \Delta \widetilde{V}_B &= 0. \tag{43}
\end{align*}
\]

By virtue of \((\widetilde{U}, \widetilde{V}) \in \mathcal{U}_+^n \times \mathcal{U}_+^n\), we know that \(\Delta \widetilde{U}_B, \Delta \widetilde{V}_B, \Delta \widetilde{V}_N \sim 0\). Multiplying (41) on the left by \((\widetilde{U}_B)^{-1/2}\) and on the right by \((\widetilde{V}_B)^{-1/2}\), squaring both sides of the resulting expression and using (40), we conclude that
\[
\| (\widetilde{U}_B)^{-1/2} \Delta \widetilde{U}_B (\widetilde{V}_B)^{1/2} \|_F = 0, \quad \| (\widetilde{U}_B)^{1/2} \Delta \widetilde{V}_B (\widetilde{V}_B)^{-1/2} \|_F = 0,
\]
from which it follows that \(\Delta \widetilde{U}_B = \Delta \widetilde{V}_B = 0\). Similarly, using (42) and the fact \(\Delta \widetilde{U}_N \bullet \Delta \widetilde{V}_N = 0\), we have \(\Delta \widetilde{U}_N = \Delta \widetilde{V}_N = 0\). Hence, (43) becomes
\[
\Delta \widetilde{U}_{BN} \widetilde{V}_N + \Delta \widetilde{U}_B \Delta \widetilde{V}_{BN} = 0. \tag{44}
\]

According to (35), we also have
\[
\begin{align*}
\Delta \widetilde{X}_B &= \Delta \widetilde{S}_B = 0, \quad \Delta \widetilde{X}_N = \Delta \widetilde{S}_N = 0.
\end{align*}
\]

Using (39) and the fact that \(\Delta \widetilde{S}_B = 0\) and \(A_{11}^*\) is one-to-one, we obtain \(\Delta \widetilde{S}_{BN} \in \text{Im}(A_{22}^*)\). Similarly, we have \(A_{22}^*(\Delta \widetilde{X}_{BN}) = 0\). Hence, we conclude that \(\Delta \widetilde{X}_{BN} \bullet \Delta \widetilde{S}_{BN} = 0\), which together with (35) implies \(\Delta \widetilde{U}_{BN} \bullet \Delta \widetilde{V}_{BN} = 0\). Using this identity and (44), and applying the same argument as above, we obtain that \(\Delta \widetilde{U}_{BN} = \Delta \widetilde{V}_{BN} = 0\). Again, in view of (35), we have \(\Delta \widetilde{X}_{BN} = \Delta \widetilde{S}_{BN} = 0\). Hence, we conclude that
\[
\Delta \widetilde{U} = \Delta \widetilde{V} = \Delta \widetilde{X} = \Delta \widetilde{S} = 0.
\]

Also, \(\Delta \widetilde{y} = 0\) follows from (37) and the fact that \(\Delta \widetilde{S} = 0\) and \(B_0\) is one-to-one.

We are now ready to show that \((\tilde{U}(\nu), \tilde{V}(\nu), \tilde{X}(\nu), \tilde{S}(\nu), \tilde{y}(\nu))\) is analytic for all \(\nu \geq 0\).
Theorem 4.8 Let \((X^*, S^*, y^*) \in \mathcal{F}_P \times \mathcal{F}_D\) be given. There hold:

i) the path \(\nu > 0 \rightarrow \bar{p}(\nu) \equiv (\bar{U}(\nu), \bar{V}(\nu), \bar{X}(\nu), \bar{S}(\nu), \bar{y}(\nu))\) is analytic at 0; consequently, \(\bar{p}(\nu)\) and all its k-th order derivatives, \(k \geq 1\), converge as \(\nu \downarrow 0\);

ii) \((\bar{U}^*, \bar{V}^*, \bar{X}^*, \bar{S}^*, \bar{y}^*) \equiv \lim_{\nu \downarrow 0}(\bar{U}(\nu), \bar{V}(\nu), \bar{X}(\nu), \bar{S}(\nu), \bar{y}(\nu))\) is the unique solution of the system defined by (29), (30) and

\[
A_0(\bar{X} - X^*) = \begin{pmatrix} 0 \\ \Delta b_2 \\ \Delta b_3 \end{pmatrix}, \quad B_0 \bar{y} + (\bar{S} - S^*) = \begin{pmatrix} \Delta C_B \\ \Delta C_{BN} \\ 0 \end{pmatrix}
\]  

in \(U_{++}^n \times U_{++}^n \times S^n \times S^n \times \mathbb{R}^m\);

iii) \((\bar{\delta}U^*, \bar{\delta}V^*, \bar{\delta}X^*, \bar{\delta}S^*, \bar{\delta}y^*) \equiv \lim_{\nu \downarrow 0}(\bar{\delta}U(\nu), \bar{\delta}V(\nu), \bar{\delta}X(\nu), \bar{\delta}S(\nu), \bar{\delta}y(\nu))\) is the unique solution of the linear system defined by

\[
\begin{align*}
\bar{\delta}U & = \begin{bmatrix} \bar{\delta}X_B \\ \bar{\delta}X_{BN} \\ 0 \\ \bar{\delta}X_N \end{bmatrix} \\
\bar{\delta}V & = \begin{bmatrix} \bar{\delta}S_B \\ \bar{\delta}S_{BN} \\ 0 \\ \bar{\delta}S_N \end{bmatrix},
\end{align*}
\]

\[
\bar{\delta}U \bar{V}^* + \bar{U}^* \bar{\delta}V + (\bar{\delta}U \bar{V}^* + \bar{U}^* \bar{\delta}V)^T = - \left[ \mathcal{L}(\bar{U}^*) \bar{V}^* + \bar{U}^* \mathcal{L}(\bar{V}^*) + (\mathcal{L}(\bar{U}^*) \bar{V}^* + \bar{U}^* \mathcal{L}(\bar{V}^*))^T \right],
\]

\[
A_0 \bar{\delta}X = -C_0 \bar{X}^* + \begin{pmatrix} \bar{\Delta}_b 1 \\ 0 \\ 0 \end{pmatrix}, \quad B_0 \bar{\delta}y + \bar{\delta}S = -D_0 \bar{y}^* + \begin{pmatrix} 0 \\ 0 \\ \Delta C_N \end{pmatrix},
\]

where

\[
C_0 \equiv \begin{pmatrix} 0 & A_{12} & A_{13} \\ 0 & 0 & 0 \end{pmatrix}, \quad D_0 \equiv \begin{pmatrix} 0 & 0 & 0 \\ A_{13} & A_{23} & 0 \end{pmatrix}.
\]

Proof. Let \(O = U_{++}^n \times U_{++}^n \times S^n \times S^n \times \mathbb{R}^m\) and \(H(w, \nu) = H(\bar{U}, \bar{V}, \bar{X}, \bar{S}, \bar{y}, \nu)\) be the map defined by system (25), (26) and (34). Indeed, \(H(w, \nu)\) is analytic of \(w\) and \(\nu\). By Proposition 4.6, the path \(w(\nu) = (\bar{U}(\nu), \bar{V}(\nu), \bar{X}(\nu), \bar{S}(\nu), \bar{y}(\nu))\) has an accumulation point \(w^* = (\bar{U}^*, \bar{V}^*, \bar{X}^*, \bar{S}^*, \bar{y}^*)\) in \(O\), which satisfies \(H(w^*, 0) = 0\). By Lemma 4.7, it follows that \(H_w(w^*, 0)\) is nonsingular. In view of implicit function theorem, there exist an \(e > 0\) and an analytical function \(\bar{w}(\nu) = (\bar{U}(\nu), \bar{V}(\nu), \bar{X}(\nu), \bar{S}(\nu), \bar{y}(\nu))\) \(O\) defined on \((-e, e)\) such that \(H(\bar{w}(\nu), \nu) = 0\) for every \(\nu \in (-e, e)\) and \(\bar{w}(0) = w^*\). Hence, it follows that \(\bar{X}_B(0) = \bar{X}_B^* = \bar{U}_B^* > 0\). Similarly, we have that \(\bar{X}_N(0), \bar{S}_B(0), \bar{S}_N(0) > 0\). Now, for \(\nu \in (0, e)\), let

\[
\bar{X}(\nu) \equiv \begin{bmatrix} \bar{X}_B(\nu) \\ \bar{X}_N(\nu) \end{bmatrix}, \quad \bar{S}(\nu) \equiv \begin{bmatrix} \nu \bar{S}_B(\nu) \\ \nu \bar{S}_N(\nu) \end{bmatrix}.
\]

Using the fact that (27) is equivalent to (32) for \(\nu > 0\), we see that \(\bar{X}(\nu)\) satisfies \(AX = b\) for \(\nu \in (0, e)\). Similarly, we have \(\bar{S}(\nu) - S^* \in \text{Im}(A^*)\), which together with the fact \(S^* \in C + \text{Im}(A^*)\), implies \(\bar{S}(\nu) \in C + \text{Im}(A^*)\). Hence, there exists \(\bar{y}(\nu) \in \mathbb{R}^m\) such that \((\bar{S}(\nu), \bar{y}(\nu))\) satisfies \(A^* \bar{y} + S = C\) for \(\nu \in (0, e)\). By virtue of (25), (26) and (49), we see that \((\bar{X}(\nu), \bar{S}(\nu))\) satisfies \(XS + SX = 2\nu W\).
for $\nu \in (0, \epsilon)$. In view of (49) and the fact that $\hat{X}_B(0), \hat{X}_N(0), \hat{S}_B(0), \hat{S}_N(0) \succ 0$, there exists a sufficiently small $0 < \delta < \epsilon$ such that $\hat{X}(\nu) \succ 0$ and $\hat{S}(\nu) \succ 0$ for every $\nu \in (0, \delta)$. Hence, for every $\nu \in (0, \delta)$, $(\hat{X}(\nu), \hat{S}(\nu), \hat{y}(\nu))$ satisfies (3)-(5). By Proposition 2.1, we have $(\hat{X}(\nu), \hat{S}(\nu), \hat{y}(\nu)) = (X(\nu), S(\nu), y(\nu))$ for every $\nu \in (0, \delta)$. According to (24) and (49), we obtain that $\hat{X}(\nu) = \bar{X}(\nu)$ and $\hat{S}(\nu) = \bar{S}(\nu)$ for all $\nu \in (0, \delta)$. Using (18) and the fact that $B_{\nu}$ is one-to-one, we have $\hat{y}(\nu) = \bar{y}(\nu)$ for all $\nu \in (0, \delta)$. Hence, we conclude that $w(\nu) = \bar{w}(\nu)$ for all $\nu \in (0, \delta)$. In term of definition 1, it follows that $i)$ holds. Upon letting $\nu \downarrow 0$ on $H(w(\nu), \nu) = 0$, we easily see that $w^*$ satisfies (29), (30) and (45). The proof of uniqueness follows from the similar argument as above. Indeed, if the system $H(w, 0) = 0$ has two distinct solutions in $U^m_{++} \times S^m$, it is clear that there exists two distinct weighted paths in a small neighborhood of $\nu = 0$. It contradicts with Proposition 2.1. Derivating the identity $H(w(\nu), \nu) = 0$ with respect to $\nu > 0$ and letting $\nu \downarrow 0$, we conclude that $\delta w = \delta w^* \equiv (\delta \hat{U}^*, \delta \bar{V}^*, \delta \bar{X}^*, \delta \bar{S}^*, \delta \bar{y}^*)$ satisfies

$$H'_w(w^*, 0)\delta w = -H'_w(w^*, 0).$$

Statement iii) now follows from the fact that $H'_w(w^*, 0)$ is nonsingular and the latter system is equivalent to (46)-(48).

In the remainder of this paper, we will let $(\hat{U}^*, \hat{V}^*, \hat{X}^*, \hat{S}^*, \hat{y}^*)$ and $(\bar{U}^*, \bar{V}^*, \bar{X}^*, \bar{S}^*, \bar{y}^*)$ denote the limits of $(\hat{U}(\nu), \hat{V}(\nu), \hat{X}(\nu), \hat{S}(\nu), \hat{y}(\nu))$ and $(\bar{U}(\nu), \bar{V}(\nu), \bar{X}(\nu), \bar{S}(\nu), \bar{y}(\nu))$, respectively, as $\nu \downarrow 0$ (as in Theorem 4.8 above). Observe that Theorem 4.8 provides a characterization of $(\hat{U}^*, \hat{V}^*, \hat{X}^*, \hat{S}^*, \hat{y}^*)$ as being the unique solution of a certain system of equations which arises by first performing some transformations to the original weighted central path system, and then setting $\nu = 0$ in the resulting system. Hence, it is reasonable to expect that the linear equations (45) can be entirely described in terms of the original data $(W, A, C, \Delta C, b, \Delta b)$. Indeed, the following result gives this alternative description of (45).

**Theorem 4.9** $(\hat{U}^*, \hat{V}^*, \hat{X}^*, \hat{S}^*)$ is the unique solution of the system given by (29), (30) and the linear equations

$$A_B(\hat{X}_B) = b, \quad [A_{BN} A_N] \begin{bmatrix} \hat{X}_{BN} \\ \hat{X}_N \end{bmatrix} \in \Delta b + \text{Im}(A_B), \quad (50)$$

$$\begin{bmatrix} \hat{S}_B \\ \hat{S}_{BN} \end{bmatrix} \in \begin{bmatrix} \Delta C_B \\ \Delta C_{BN} \end{bmatrix} + \text{Im} \begin{bmatrix} A^*_B \\ A^*_{BN} \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ \hat{S}_N \end{bmatrix} \in C + \text{Im} \begin{bmatrix} A^*_B \\ A^*_{BN} \end{bmatrix} \quad (51)$$

in $U^m_{++} \times S^m_{++} \times S^m_{++}$.

**Proof:** From Theorem 4.8(ii), it suffices to show that (45) is equivalent to (50) and (51). Since the first equation of (45) is the same as (32) with $\nu = 0$, we have that the first equation of (45) holds if and only if

$$A_{11}(\hat{X}_B) = A_{11}(X_B), \quad A_{22}(\hat{X}_{BN}) + A_{23}(\hat{X}_N) = \Delta b_2, \quad A_{33}(\hat{X}_N) = \Delta b_3. \quad (52)$$
By Lemma 4.4, the first identity in (52) can be written as

$$(T \circ \mathcal{A}) \begin{pmatrix} \tilde{X}_B \\ 0 \\ 0 \end{pmatrix} = (T \circ \mathcal{A}) \begin{pmatrix} X_B \\ 0 \\ 0 \end{pmatrix},$$

and hence it is equivalent to $\mathcal{A}_B(\tilde{X}_B) = \mathcal{A}_B(X_B) = b$, in view of relation (31) and the fact that $T$ is an isomorphism. By Lemma 4.4 and the fact that $\mathcal{A}_{11}$ is onto, the second and third identities in (52) hold if and only if

$$(T \circ \mathcal{A}) \begin{pmatrix} \tilde{X}_B \\ \tilde{X}_{BN} \\ \tilde{X}_N \end{pmatrix} = \begin{pmatrix} \tilde{\Delta}b_1 \\ \tilde{\Delta}b_2 \\ \tilde{\Delta}b_3 \end{pmatrix} = T(\Delta b)$$

for some $\tilde{X}_B \in \mathcal{S}^{[B]}$, and hence it is equivalent to $\mathcal{A}_{BN}(\tilde{X}_{BN}) + \mathcal{A}_N(\tilde{X}_N) \in \Delta b + \text{Im}(\mathcal{A}_B)$, in view of (31) and the fact that $T$ is an isomorphism. We have thus shown that the first equation of (45) is equivalent to (50).

The fact that the second equation of (45) holds if and only if (51) holds can be proved in a similar way as above.

The following result gives an alternative characterization of $(\delta U^*, \delta V^*, \delta X^*, \delta S^*)$ involving the original data $(W, \mathcal{A}, C, \Delta C, b, \Delta b)$.

**Theorem 4.10** $(\delta U^*, \delta V^*, \delta X^*, \delta S^*)$ is the unique solution of the linear system of equations (46), (47) and

$$\begin{pmatrix} \mathcal{A}_B & \mathcal{A}_{BN} & \mathcal{A}_N \\ \tilde{\mathcal{A}}_B & \tilde{\mathcal{A}}_{BN} & \mathcal{A}_N \end{pmatrix} \begin{pmatrix} \delta \tilde{X}_B \\ \delta \tilde{X}_{BN} \\ \delta \tilde{X}_N \end{pmatrix} = \Delta b, \quad (\mathcal{A}_{BN} \quad \mathcal{A}_N) \begin{pmatrix} \tilde{\Delta}b_1 \\ \tilde{\Delta}b_2 \\ \tilde{\Delta}b_3 \end{pmatrix} = \Delta b + \text{Im}(\mathcal{A}_B), \quad \begin{pmatrix} \tilde{\mathcal{A}}_B & \tilde{\mathcal{A}}_{BN} & \mathcal{A}_N \end{pmatrix} \begin{pmatrix} \tilde{\mathcal{S}}_B \\ \tilde{\mathcal{S}}_{BN} \\ \tilde{\mathcal{S}}_N \end{pmatrix} \in \Delta C + \text{Im}(\mathcal{A}_B), \quad (\mathcal{A}_{BN} \quad \mathcal{A}_N) \begin{pmatrix} \mathcal{A}_B \\ \mathcal{A}_{BN} \end{pmatrix} \begin{pmatrix} \mathcal{A}_B \\ \mathcal{A}_{BN} \end{pmatrix} \begin{pmatrix} \tilde{\mathcal{S}}_B \\ \tilde{\mathcal{S}}_{BN} \\ \tilde{\mathcal{S}}_N \end{pmatrix} \in \Delta C.$$

**Proof.** From Theorem 4.8(iii), it suffices to show that (48) is equivalent to (53) and (54). Observe that the first equation of (48) can be written as

$$\begin{align*}
\mathcal{A}_{11}(\delta \tilde{X}_B) + \mathcal{A}_{12}(\delta \tilde{X}_{BN}) + \mathcal{A}_{13}(\delta \tilde{X}_N) &= \tilde{\Delta}b_1, \\
\mathcal{A}_{22}(\delta \tilde{X}_{BN}) + \mathcal{A}_{23}(\delta \tilde{X}_N) &= 0, \\
\mathcal{A}_{33}(\delta \tilde{X}_N) &= 0.
\end{align*}$$

Using Lemma 4.4, the fact that $\mathcal{A}_{11}$ is onto and the identities $\mathcal{A}_{22}(\tilde{X}_{BN}) + \mathcal{A}_{23}(\tilde{X}_N) = \Delta b_2$ and $\mathcal{A}_{33}(\tilde{X}_N) = \tilde{\Delta}b_3$ that hold in view of (45), we easily see that the first and last two equations above are respectively equivalent to

$$(T \circ \mathcal{A}) \begin{pmatrix} \delta \tilde{X}_B \\ \delta \tilde{X}_{BN} \\ \delta \tilde{X}_N \end{pmatrix} = \begin{pmatrix} \tilde{\Delta}b_1 \\ \tilde{\Delta}b_2 \\ \tilde{\Delta}b_3 \end{pmatrix}, \quad (T \circ \mathcal{A}) \begin{pmatrix} \delta \tilde{X}_B \\ \delta \tilde{X}_{BN} \\ \delta \tilde{X}_N \end{pmatrix} = T(\Delta b).$$
for some $\tilde{X}_B \in \mathcal{S}^{[B]}$. The latter conditions in turn are equivalent to (53) in view of (31) and the fact that $T$ is an isomorphism.

Using similar arguments as to ones used above, it can be shown that the second equation of (48) holds if and only if (54) holds.

The proof of Theorem 4.1 is now obvious. Indeed, the analyticity of the map $\nu \to (X(\nu), S(\nu))$ follows from (24) and the analyticity of $\nu \to (\tilde{X}(\nu), \tilde{T}(\nu))$. The analyticity of $\nu \to y(\nu)$ follows from the analyticity of $\nu \to S(\nu)$ and Assumption A.1. The last statement of the theorem is obvious.

5 Error bound analysis

In this section we derive an error bound on the distance of a point lying in a certain neighborhood of the central path to the primal-dual optimal set. This error bound has played a crucial role in the local convergence analysis of a primal-dual interior point algorithm proposed by Kojima et al. [?].

For any given nonempty compact set $\mathcal{K} \subset \mathcal{G}_{++}$ and constants $\gamma, \tau > 0$, define

$$
\mathcal{N}(\gamma, \tau, \mathcal{K}) = \{(X, S, y) \in \mathcal{S}^{n+} \times \mathcal{S}^{n+} \times \mathbb{R}^m : G(X, S, y) \in \tau \mathcal{K}, \|(XS + SX)/2 - \tau I\| \leq \gamma \tau\},
$$

where the map $G$ and the set $\mathcal{G}_{++}$ are defined in (7) and (8), respectively.

Observe that the set $\cup_{\tau > 0} \mathcal{N}(\gamma, \tau, \mathcal{K})$ forms a neighborhood of the primal-dual central path. This neighborhood and related ones have once been used in the development of primal-dual interior point algorithms for SDP. For example, see Kojima et al. [?].

The following result gives an error bound on the distance of a point lying in $\mathcal{N}(\gamma, \tau, \mathcal{K})$ to the primal-dual optimal set $\mathcal{F}_P^* \times \mathcal{F}_D^*$. Its proof will be given at the end of this section.

**Theorem 5.1** Let $\gamma \in (0, 1]$ and any nonempty compact set $\mathcal{K} \subset \mathcal{G}_{++}$ be given. Then, there exists a constant $M = M(\gamma, \mathcal{K}) > 0$ such that

$$
\text{dist}((X, S, y), \mathcal{F}_P^* \times \mathcal{F}_D^*) \leq M \tau,
$$

for every $\tau \in (0, 1]$ and $(X, S, y) \in \mathcal{N}(\gamma, \tau, \mathcal{K})$.

In view of Proposition 2.1, for each $(\nu, W, \Delta C, \Delta b) \in (0, \infty) \times \mathcal{S}^{n+} \times \mathcal{G}_{++}$, the system of nonlinear equations (3)-(5) has a unique solution, which in this section we denote by $(X(\nu, W, \Delta C, \Delta b), S(\nu, W, \Delta C, \Delta b), y(\nu, W, \Delta C, \Delta b))$ in order to emphasize and study its dependence on $(W, \Delta C, \Delta b)$. Moreover, in view of Theorem 4.1, the limit

$$
\lim_{\nu \to 0}(X(\nu, W, \Delta C, \Delta b), S(\nu, W, \Delta C, \Delta b), y(\nu, W, \Delta C, \Delta b)),
$$

denoted by $(X(0, W, \Delta C, \Delta b), S(0, W, \Delta C, \Delta b), y(0, W, \Delta C, \Delta b))$, exists for every $(W, \Delta C, \Delta b) \in \mathcal{S}^{n+} \times \mathcal{G}_{++}$. Hence, the functions $X(\cdot, \cdot, \cdot), S(\cdot, \cdot, \cdot)$ and $y(\cdot, \cdot, \cdot)$ are well-defined over the set $(0, \infty) \times \mathcal{S}^{n+} \times \mathcal{G}_{++}$. In an obvious way, we can also define the functions $\tilde{X}(\nu, W, \Delta C, \Delta b), \tilde{S}(\nu, W, \Delta C, \Delta b)$ and $\tilde{y}(\nu, W, \Delta C, \Delta b)$ over the set $(0, \infty) \times \mathcal{S}^{n+} \times \mathcal{G}_{++}$.

It turns out that the above functions are analytic according to the following definition. We say that a function $f : \Omega \subseteq E \to F$, where $E, F$ are two finite dimensional normed vector spaces, is analytic if there exists an open set $\mathcal{O} \subseteq E$ containing $\Omega$ and an analytic function $\tilde{f} : \mathcal{O} \to F$ such that $\tilde{f}$ restricted to $\Omega$ is equal to $f$. 

16
Theorem 5.2 There hold:

i) the map \((\nu, W, \Delta C, \Delta b) \in [0, \infty) \times S^n_+ \times G_+ \rightarrow (\bar{X}(\nu, W, \Delta C, \Delta b), \bar{S}(\nu, W, \Delta C, \Delta b), \bar{y}(\nu, W, \\
\Delta C, \Delta b))\) is analytic;

ii) the map \((\nu, W, \Delta C, \Delta b) \in [0, \infty) \times S^n_+ \times G_+ \rightarrow (X(\nu, W, \Delta C, \Delta b), S(\nu, W, \Delta C, \Delta b), y(\nu, W, \\
\Delta C, \Delta b))\) is analytic.

Proof. The proof of the theorem is identical to the proof of Theorem 4.8 and Theorem 4.1, except that when invoking the implicit function theorem, we should view \((\nu, W, \Delta C, \Delta b)\) as the parameter vector.

Theorem 5.3 Let \(\gamma \in (0, 1]\) be given. Then, for all \((\nu, W, \Delta C, \Delta b) \in [0, 1]\times \mathcal{W}(\gamma) \times \mathcal{K}\), there exists a constant \(M = M(\gamma, \mathcal{K}) > 0\) such that

\[
\| (X(\nu, W, \Delta C, \Delta b), S(\nu, W, \Delta C, \Delta b)) - (X(0, W, \Delta C, \Delta b), S(0, W, \Delta C, \Delta b)) \| \leq M \nu,
\]

where \(\mathcal{W}(\gamma) \equiv \{W \in S^n_+ : \|W - I\| \leq \gamma\}\).

Proof. By the mean value theorem, we have

\[
\| (X(\nu, W, \Delta C, \Delta b), S(\nu, W, \Delta C, \Delta b)) - (X(0, W, \Delta C, \Delta b), S(0, W, \Delta C, \Delta b)) \| \\
\leq \sup_{\theta \in [0, 1]} \| (X'(\theta \nu, W, \Delta C, \Delta b), S'(\theta \nu, W, \Delta C, \Delta b)) \| \nu
\]

By Theorem 5.2(ii) and the fact that \(\mathcal{W}(\gamma) \times \mathcal{K}\) is compact, there exists a constant \(M = M(\gamma, \mathcal{K}) > 0\) such that \(\| (X'(\theta \nu, W, \Delta C, \Delta b), S'(\theta \nu, W, \Delta C, \Delta b)) \| \leq M\) for all \((\theta, \nu, W, \Delta C, \Delta b) \in [0,1] \times [0,1] \times \mathcal{W}(\gamma) \times \mathcal{K}\). Hence, the conclusion follows.

The proof of Theorem 5.1 now follows from Assumption A.1 and Theorem 5.3 with \(\nu = \tau, W = (X S + SX)/(2\tau), (X, S) = (X(\nu, W, \Delta C, \Delta b), S(\nu, W, \Delta C, \Delta b))\) and the fact \((X(0, W, \Delta C, \Delta b), S(0, W, \Delta C, \Delta b), y(0, W, \Delta C, \Delta b)) \in \mathcal{F}_p \times \mathcal{F}_D\).