A NEW NOTION OF WEIGHTED CENTERS FOR SEMIDEFINITE PROGRAMMING[*]

CHEK BENG CHUA[†]

Abstract. The notion of weighted centers is essential in V-space interior-point algorithms for linear programming. Although there were some successes in generalizing this notion to semidefinite programming via weighted center equations, we still do not have a generalization that preserves two important properties—(1) each choice of weights uniquely determines a pair of primal-dual weighted centers, and (2) the set of all primal-dual weighted centers completely fills up the relative interior of the primal-dual feasible region. This paper presents a new notion of weighted centers for semidefinite programming that possesses both uniqueness and completeness. Furthermore, it is shown that under strict complementarity, these weighted centers converge to weighted centers of optimal faces. Finally, this convergence result is applied to homogeneous cone programming, where the central paths defined by a certain class of optimal barriers for homogeneous cones are shown to converge to analytic centers of optimal faces in the presence of strictly complementary solutions.

Key words. weighted center, semidefinite programming, homogeneous cone programming, facial structure, Cholesky decomposition

AMS subject classifications. 90C22 90C25 90C51

DOI. 10.1137/040613378

1. Introduction. This paper presents a new generalization of the notion of weighted centers from linear programming (LP) to semidefinite programming (SDP). We consider the following primal-dual pair of SDP problems,

$$
\text{(P)} \qquad
\begin{aligned}
\inf \quad & \mathbf{C} \bullet \mathbf{X} \\
\text{subject to} \quad & \mathbf{A}^{(i)} \bullet \mathbf{X} = \mathbf{b}_i, \quad i = 1, \dots, m, \\
& \mathbf{X} \succeq \mathbf{0},
\end{aligned}
$$

and

$$
\text{(D)} \qquad
\begin{aligned}
\sup \quad & \mathbf{b}^T \mathbf{y} \\
\text{subject to} \quad & \mathbf{S} = \mathbf{C} - \sum_{i=1}^{m} \mathbf{A}^{(i)} \mathbf{y}_i, \\
& \mathbf{S} \succeq \mathbf{0},
\end{aligned}
$$

where the $\mathbf{A}^{(i)}$ and $\mathbf{C}$ are symmetric matrices, $\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_m)^T$ and $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_m)^T$ are real m-vectors, $\bullet : (\mathbf{A}, \mathbf{B}) \mapsto tr\mathbf{A}^T\mathbf{B}$ is the trace inner product, and $\mathbf{X} \succeq \mathbf{0}$ means that $\mathbf{X}$ is symmetric and positive semidefinite.

The notion of weighted centers for LP is very useful in interior-point algorithms that use the V-space approach (see [10, 11]). These weighted centers can be characterized in the following two ways:

1. as minimizers of shifted, weighted logarithmic barriers

$$(\mathbf{x}, \mathbf{s}) \in \mathbb{R}^n \times \mathbb{R}^n \mapsto -\sum_{i=1}^n w_i \log \mathbf{x}_i - \sum_{i=1}^n w_i \log \mathbf{s}_i + \mathbf{x}^T \mathbf{s}$$

over the primal-dual feasible region $\{(\mathbf{x}, \mathbf{s}) \in \mathbb{R}^n \times \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}, \ \mathbf{s} = \mathbf{c} - A^T\mathbf{y}, \ \mathbf{y} \in \mathbb{R}^m, \ \mathbf{x} \geq \mathbf{0}, \ \mathbf{s} \geq \mathbf{0}\}$, and

2. as solutions to weighted center equations

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{s} = \mathbf{c} - A^T\mathbf{y} \quad \text{for some } \mathbf{y} \in \mathbb{R}^m,$$
$$\mathbf{x}\mathbf{s} = \mathbf{w}, \quad \mathbf{x} > \mathbf{0}, \quad \text{and} \quad \mathbf{s} > \mathbf{0},$$

where $\mathbf{w} = (w_1, \ldots, w_n)^T$ and $\mathbf{xs}$ denotes the componentwise product of $\mathbf{x}$ and $\mathbf{s}$.

A main obstacle in generalizing weighted centers to SDP is the lack of proper weighted barriers. Nonetheless, there were some successes in generalizing weighted center equations to SDP. Monteiro and Pang [15] considered the weighted Alizadeh–Haeberly–Overton (AHO) centers, where the equation $\mathbf{XS} + \mathbf{SX} = 2\mathbf{W}$ replaces $\mathbf{xs} = \mathbf{w}$. Every symmetric, positive definite matrix $\mathbf{W}$ uniquely determines a weighted AHO center. However, unlike LP, these weighted centers do not fill up the whole relative interior of the primal-dual feasible region, i.e., not every strictly feasible pair of matrices $(\mathbf{X}, \mathbf{S})$ is a pair of weighted AHO centers. Sturm and Zhang [21] considered a different generalization that is based on the Nesterov–Todd (NT) scaling point. This generalization replaces $\mathbf{xs} = \mathbf{w}$ with $\Lambda(\mathbf{XS}) = \mathbf{W}$, where $\Lambda(\mathbf{XS})$ denotes the diagonal matrix with the eigenvalues of $\mathbf{XS}$ on its diagonal, and $\mathbf{W}$ is a positive, diagonal matrix. In contrast with the weighted AHO centers, these weighted NT centers completely fill up the relative interior of the primal-dual feasible region as $\mathbf{W}$ ranges over all positive, diagonal matrices but lacks uniqueness, i.e., the equations may have more than one solution for each positive, diagonal matrix $\mathbf{W}$.

We shall describe an alternative generalization of weighted centers to SDP that possesses both uniqueness and completeness. While this generalization, which is based on Cholesky factors, is similar to a generalization considered by Monteiro and Zanjácomo [17], the main difference lies in the choice of $\mathbf{W}$. In [17], $\mathbf{W}$ is required to be "close" to multiples of the identity matrix in order for the weighted center equation to have a unique solution. On the other hand, we use positive, diagonal matrices $\mathbf{W}$ to ensure uniqueness. By restricting to diagonal matrices, the weighted centers can be characterized as minimizers of certain shifted, weighted logarithmic barriers over the primal-dual feasible region. In each generalization, the collection of weighted centers does not completely fill up the relative interior of the primal-dual feasible region. This drawback can be easily rectified in our generalization by considering orthonormal similarity transformations. Thus, for the first time, we have a notion of weighted centers for SDP that possesses two useful properties—uniqueness and completeness. This lays the foundation for future extensions of V-space algorithms to SDP.

Besides having both uniqueness and completeness, these weighted centers converge to weighted centers of optimal faces under strict complementarity. This generalizes the same property of usual central paths for SDP. Similar results were shown in [12, 13] and by Prieß and Stoer [20] for notions of weighted centers defined by the maps $(\mathbf{X}, \mathbf{S}) \mapsto (\mathbf{XS} + \mathbf{SX})/2$ and $(\mathbf{X}, \mathbf{S}) \mapsto \mathbf{X}^{1/2}\mathbf{S}\mathbf{X}^{1/2}$, respectively.

Yet another reason for considering this generalization is that our weighted centers include the analytic centers defined by a certain class of optimal barriers for homogeneous cones. Consequently, we can apply the above convergence result to ho-

mogeneous cone programming (HCP). Specifically, we show that under strict complementarity, central paths defined by this class of optimal barriers converge to analytic centers of optimal faces.

This paper is organized as follows. The next section starts with some basics of SDP, including a discussion on the facial structures of positive definite cones and the notion of strict complementarity. In section 3, a generalization, based on Cholesky factors, of weighted centers to SDP is presented and a characterization of limit points of these weighted centers under strict complementarity is given. This result is applied to HCP in section 4, where it is shown that central paths defined by a certain class of optimal barriers for HCP converge to analytic centers of optimal faces in the presence of strictly complementary solutions.

Notation and conventions. Throughout this paper, we use the following notation.

The space of symmetric matrices of order n is denoted by $\mathbb{S}^n$ and the cone of symmetric, positive semidefinite (resp., positive definite) matrices of order n is denoted by $\mathbb{S}^n_+$ (resp., $\mathbb{S}^n_{++}$). If $\mathbf{X} \in \mathbb{S}^n$, then the statement $\mathbf{X} \succeq \mathbf{0}$ (resp., $\mathbf{X} \succ \mathbf{0}$) means that $\mathbf{X} \in \mathbb{S}^n_+$ (resp., $\mathbf{X} \in \mathbb{S}^n_{++}$).

For any m-by-n matrix $\mathbf{M}$ and any subsets of indices $I \subset \{1, \ldots, m\}$ and $J \subset \{1, \ldots, n\}$, the submatrix of $\mathbf{M}$ with row indices in I and column indices in J is denoted by $\mathbf{M}_{IJ}$. If $I = \{i\}$ (or $J = \{j\}$) is a singleton, we may also write i (or j) in place of $\{i\}$ (or $\{j\}$).

The identity matrix of appropriate size (in the context used) is denoted by $\mathbf{I}$. For any subset B of positive integer indices, $\mathbf{I}_B$ denotes the 0-1 diagonal matrix of appropriate size with $(\mathbf{I}_B)_{ii} = 1$ if and only if $i \in B$, and $\mathbf{I}^c_B$ denotes $\mathbf{I} - \mathbf{I}_B$.

For each $\mathbf{X} \in \mathbb{S}^n$, $\mathcal{R}(\mathbf{X})$ denotes the range space of $\mathbf{X}$ and $\mathcal{N}(\mathbf{X})$ denotes the null space of $\mathbf{X}$.

For each topological subspace S, relint(S) denotes the relative interior of S and cl(S) denotes the closure of S.

For each sequence $x_1, \ldots, x_n$ of real numbers, $\mathrm{Diag}(x_1, \ldots, x_n)$ denotes the diagonal matrix with $x_1, \ldots, x_n$ on its diagonal.

2. Optimal faces and strict complementarity of SDP. It is well known that each face of $\mathbb{S}^n_+$ can be uniquely identified with a subspace of $\mathbb{R}^n$ as follows: F is a face of $\mathbb{S}^n_+$ if and only if $F = \{\mathbf{X} \in \mathbb{S}^n_+ : \mathcal{R}(\mathbf{X}) \subset \mathbb{V}\}$ for some linear subspace $\mathbb{V} \subset \mathbb{R}^n$. Moreover, for any face $F = \{\mathbf{X} \in \mathbb{S}^n_+ : \mathcal{R}(\mathbf{X}) \subset \mathbb{V}\}$ of $\mathbb{S}^n_+$, $\tilde{\mathbf{X}} \in \mathrm{relint}(F)$ if and only if $\mathcal{R}(\tilde{\mathbf{X}}) = \mathbb{V}$ (see [1]). Thus, matrices in the relative interior of any face of $\mathbb{S}^n_+$ are characterized by having maximal rank among all matrices in the face.

An alternative characterization, based on Cholesky factors, of the relative interior of a face shall now be given.

It is a well-known fact that every symmetric, positive definite matrix $\mathbf{X}$ has a unique Cholesky factor (i.e., a lower triangular matrix $\mathbf{L}$ with nonnegative diagonal satisfying $\mathbf{X} = \mathbf{L}\mathbf{L}^T$). When $\mathbf{X}$ is symmetric and positive semidefinite, it still has a Cholesky factor. However, the Cholesky factor may not be unique when $\mathbf{X}$ is not positive definite. The next proposition shows that we can recover uniqueness by posing an additional condition on $\mathbf{L}$.

PROPOSITION 1. *Every symmetric, positive semidefinite matrix* $\mathbf{X}$ *has a unique Cholesky factor* $\mathbf{L}_{\mathbf{X}}$ *satisfying*

$$(2.1) \qquad\qquad (\mathbf{L}_{\mathbf{X}})_{ii} = 0 \implies (\mathbf{L}_{\mathbf{X}})_{ji} = 0 \ \forall j,$$

i.e., every column of $\mathbf{L}_{\mathbf{X}}$ *either is a zero column or has a positive diagonal entry.*

Proof. Existence. Suppose $\mathbf{X} \in \mathbb{S}_+^n$. We shall prove by induction on n that

(2.2)
$$\forall \mu_k \downarrow 0, \ \forall \{\mathbf{Y}(k)\}_{k=1}^\infty \subset \mathbb{S}_{++}^n \text{ with } \mathbf{Y}(\infty) := \lim_{k\to\infty} \mathbf{Y}(k) \in \mathbb{S}_{++}^n,$$
$$\text{all limit points of } \{\mathbf{L}(k) := \mathbf{L}_{\mathbf{X}+\mu_k\mathbf{Y}(k)}\}_{k=1}^\infty \text{ satisfy (2.1),}$$

where $\mathbf{L}_{\mathbf{X}+\mu_k\mathbf{Y}(k)}$ denotes the unique Cholesky factor of $\mathbf{X} + \mu_k\mathbf{Y}(k) \in \mathbb{S}_{++}^n$. Since the sequence $\{\mathbf{L}(k)\}$ is bounded and hence has at least one limit point, the existence of $\mathbf{L}_{\mathbf{X}}$ follows by taking, say, $\mu_k = 1/k$ and $\mathbf{Y}(k) \equiv \mathbf{I}$.

The case $n = 1$ is trivially true. Suppose that for some $n \geq 1$, (2.2) holds for all $\mathbf{X} \in \mathbb{S}_+^n$. Consider the case $\mathbf{X} \in \mathbb{S}_+^{n+1}$. Let $\mathbf{L}$ denote an arbitrary limit point of $\{\mathbf{L}(k)\}$. By considering a subsequence if necessary, we may assume without any loss of generality that $\lim_{t\to\infty} \mathbf{L}(k) = \mathbf{L}$. We consider two cases.

If $\mathbf{L}_{11} = 0$, then the entries in the first column and row of $\mathbf{X}$ are zeros, whence $\mathbf{L}(k)_{j1} = \sqrt{\mu_k}(\mathbf{Y}(k)_{11})^{-1/2}\mathbf{Y}(k)_{j1}$ for all $j \in \{2, \ldots, n+1\}$. Since $\mathbf{Y}(\infty) \in \mathbb{S}_{++}^n$, it follows that $(\mathbf{Y}(k)_{11})^{-1/2}\mathbf{Y}(k)_{j1} \to (\mathbf{L}_{\mathbf{Y}(\infty)})_{j1}$, whence $\mathbf{L}_{j1} = \lim_{k\to\infty}\mathbf{L}(k)_{j1} = 0$. We then apply (2.2) on the remaining columns and rows of $\mathbf{X}$ and $\mathbf{Y}(k)$ to conclude (2.2) for $\mathbf{X}$.

If $\mathbf{L}_{11} > 0$, then $\mathbf{X}_{11} > 0$ and we only need to show that $\mathbf{L}_{JJ}$ satisfy (2.1) where J denotes the set $\{2, \ldots, n+1\}$. Now

$$\mathbf{L}(k)_{JJ} = \mathbf{L}_{\mathbf{X}_{JJ}+\mu_k\mathbf{Y}(k)_{JJ}-(\mathbf{X}_{J1}+\mu_k\mathbf{Y}(k)_{J1})(\mathbf{X}_{11}+\mu_k\mathbf{Y}(t)_{11})^{-1}(\mathbf{X}_{J1}+\mu_k\mathbf{Y}(k)_{J1})^T}$$
$$= (\mathbf{X}_{11} + \mu_k\mathbf{Y}(k)_{11})^{-1/2}\mathbf{L}_{\hat{\mathbf{X}}+\mu_k\hat{\mathbf{Y}}(k)+\mu_k^2\mathbf{Z}(k)},$$

where

$$\hat{\mathbf{X}} = \mathbf{X}_{11}\mathbf{X}_{JJ} - \mathbf{X}_{J1}\mathbf{X}_{J1}^T \in \mathbb{S}_+^n,$$
$$\hat{\mathbf{Y}}(k) = \mathbf{Y}(k)_{11}\mathbf{X}_{JJ} + \mathbf{X}_{11}\mathbf{Y}(k)_{JJ} - \mathbf{Y}(k)_{J1}(\mathbf{X}_{J1})^T - \mathbf{X}_{J1}(\mathbf{Y}(k)_{J1})^T, \text{ and}$$
$$\mathbf{Z}(k) = \mathbf{Y}(k)_{11}\mathbf{Y}(k)_{JJ} - (\mathbf{Y}(k)_{J1})(\mathbf{Y}(k)_{J1})^T \in \mathbb{S}_{++}^n.$$

Let $\hat{\mathbf{Y}}(\infty)$ denote $\lim_{k\to\infty}\hat{\mathbf{Y}}(k)$. For each $k \in \{1, 2, \ldots, \infty\}$ and each $\mathbf{v}(\neq \mathbf{0}) \in \mathbb{R}^n$,

$$\mathbf{v}^T(\hat{\mathbf{Y}}(k))\mathbf{v}$$
$$= \mathbf{Y}(k)_{11}\mathbf{v}^T\mathbf{X}_{JJ}\mathbf{v} + \mathbf{X}_{11}\mathbf{v}^T\mathbf{Y}(k)_{JJ}\mathbf{v} - 2(\mathbf{v}^T\mathbf{Y}(k)_{J1})(\mathbf{v}^T\mathbf{X}_{J1})$$
$$> \mathbf{Y}(k)_{11}\frac{(\mathbf{v}^T\mathbf{X}_{J1})^2}{\mathbf{X}_{11}} + \mathbf{X}_{11}\frac{(\mathbf{v}^T\mathbf{Y}(k)_{J1})^2}{\mathbf{Y}(k)_{11}} - 2(\mathbf{v}^T\mathbf{Y}(k)_{J1})(\mathbf{v}^T\mathbf{X}_{J1}) \geq 0,$$

where we have used $\hat{\mathbf{Y}}(k) \in \mathbb{S}_{++}^n$, $\mathbf{X} \in \mathbb{S}_+^n$, and $\mathbf{X}_{11} > 0$ in the strict inequality and the arithmetic-geometric mean inequality in the last inequality. Thus we may apply (2.2) to $\hat{\mathbf{X}} \in \mathbb{S}_+^n$ and $\{\hat{\mathbf{Y}}(k) + \mu_k\mathbf{Z}(k)\} \subset \mathbb{S}_{++}^n$ to deduce that $\lim_{k\to\infty}\mathbf{L}_{\hat{\mathbf{X}}+\mu_k\hat{\mathbf{Y}}(k)+\mu_k^2\mathbf{Z}(k)}$ satisfies (2.1). Consequently, $\mathbf{L}_{JJ} = \lim_{k\to\infty}\mathbf{L}(k)_{JJ}$ also satisfies (2.1).

Uniqueness. First, consider the case when $\mathbf{X}$ is a nonnegative diagonal matrix. Let B denote the set of indices of positive diagonal entries of $\mathbf{X}$. Suppose that $\mathbf{L}$ is a Cholesky factor of $\mathbf{X}$ satisfying (2.1). Since $\mathbf{X}_{ii} = 0$ for all $i \notin B$, the ith row of $\mathbf{L}$ must be a row of zeros. Thus, $\mathbf{L}_{ij} = 0$ whenever $i \notin B$ or $j \notin B$. Consequently, $\mathbf{L}_{BB}\mathbf{L}_{BB}^T = \mathbf{X}_{BB}$ is a positive, diagonal matrix. Thus, $\mathbf{L}$ is unique. Now, suppose that $\mathbf{X} \succeq 0$ is arbitrary. Suppose that $\mathbf{L}$ and $\mathbf{L}'$ are Cholesky factors of $\mathbf{X}$ satisfying (2.1). Let B be the set of indices of nonzero columns of $\mathbf{L}$. It is clear that $\mathbf{L}\mathbf{I}_B = \mathbf{L}$,

and thus, $(\mathbf{L} + \mathbf{I}_B^c)\mathbf{I}_B = \mathbf{L}$. Therefore,

$$\begin{aligned}
\mathbf{I}_B &= [(\mathbf{L} + \mathbf{I}_B^c)^{-1}\mathbf{L}][(\mathbf{L} + \mathbf{I}_B^c)^{-1}\mathbf{L}]^T \\
&= (\mathbf{L} + \mathbf{I}_B^c)^{-1}\mathbf{X}(\mathbf{L} + \mathbf{I}_B^c)^{-T} \\
&= [(\mathbf{L} + \mathbf{I}_B^c)^{-1}\mathbf{L}'][(\mathbf{L} + \mathbf{I}_B^c)^{-1}\mathbf{L}']^T.
\end{aligned}$$

Since $((\mathbf{L} + \mathbf{I}_B^c)^{-1}\mathbf{L}')_{ii} = 0 \implies \mathbf{L}'_{ii} = 0 \implies \mathbf{L}'_{ji} = 0 \ \forall j \implies ((\mathbf{L} + \mathbf{I}_B^c)^{-1}\mathbf{L}')_{ji} = 0 \ \forall j$, we have that both $(\mathbf{L} + \mathbf{I}_B^c)^{-1}\mathbf{L}'$ and $(\mathbf{L} + \mathbf{I}_B^c)^{-1}\mathbf{L}$ are Cholesky factors of $\mathbf{I}_B$ satisfying (2.1). Thus, $\mathbf{L} = \mathbf{L}'$. □

In a similar way, we can prove the followinig.

PROPOSITION 2. *Every symmetric, positive semidefinite matrix* $\mathbf{X}$ *has a unique inverse Cholesky factor* $\mathbf{U_X}$ *(i.e., an upper triangular matrix* $\mathbf{U}$ *with nonnegative diagonal satisfying* $\mathbf{X} = \mathbf{U}\mathbf{U}^T$*) satisfying*

$$(2.3) \qquad\qquad (\mathbf{U_X})_{ii} = 0 \implies (\mathbf{U_X})_{ji} = 0 \ \forall j.$$

Henceforth, the unique Cholesky factor of $\mathbf{X}$ that satisfies (2.1) is denoted by $\mathbf{L_X}$, and the unique inverse Cholesky factor of $\mathbf{X}$ that satisfies (2.3) is denoted by $\mathbf{U_X}$.

We now describe the faces of $\mathbb{S}_+^n$ based on these Cholesky factors.

Suppose that F is a face of $\mathbb{S}_+^n$ and $\tilde{\mathbf{X}} \in \mathrm{relint}(F)$ is arbitrary. From the proof of uniqueness, we see that $(\mathbf{L}_{\tilde{\mathbf{X}}} + \mathbf{I}_B^c)^{-1}\tilde{\mathbf{X}}(\mathbf{L}_{\tilde{\mathbf{X}}} + \mathbf{I}_B^c)^{-T} = \mathbf{I}_B$, where B is the set of indices of nonzero columns of $\mathbf{L}_{\tilde{\mathbf{X}}}$. Since $\mathbf{X} \mapsto (\mathbf{L}_{\tilde{\mathbf{X}}} + \mathbf{I}_B^c)^{-1}\mathbf{X}(\mathbf{L}_{\tilde{\mathbf{X}}} + \mathbf{I}_B^c)^{-T}$ is a linear automorphism of $\mathbb{S}_+^n$, it maps F to some face F' of $\mathbb{S}_+^n$ with $\mathbf{I}_B \in \mathrm{relint}(F')$. Therefore, for any $\mathbf{X} \in \mathbb{S}_+^n$, $\mathbf{X} \in F'$ if and only if $\mathcal{R}(\mathbf{X}) \subset \mathcal{R}(\mathbf{I}_B)$, which holds if and only if $(i \notin B) \vee (j \notin B) \implies \mathbf{X}_{ij} = 0$. Consequently,

$$(2.4) \quad F = \{(\mathbf{L}_{\tilde{\mathbf{X}}} + \mathbf{I}_B^c)\mathbf{X}(L_{\tilde{\mathbf{X}}} + \mathbf{I}_B^c)^T : \mathbf{X} \succeq \mathbf{0}, \ (i \notin B) \vee (j \notin B) \implies \mathbf{X}_{ij} = 0\}.$$

From this representation of the face F, we deduce the following.

PROPOSITION 3. *If* F *is a face of* $\mathbb{S}_+^n$, $B = \{i : \exists \mathbf{X} \in F, (\mathbf{L_X})_{ii} \neq 0\}$, *and* $\tilde{\mathbf{X}} \in F$, *then*

1. $(\mathbf{L}_{\tilde{\mathbf{X}}})_{ii} = 0 \ \forall i \notin B$ *and*
2. $\tilde{\mathbf{X}} \in \mathrm{relint}(F) \iff (\mathbf{L}_{\tilde{\mathbf{X}}})_{ii} > 0 \ \forall i \in B$.

Similarly, we can use inverse Cholesky factors to characterize the relative interiors of the faces of $\mathbb{S}_+^n$.

PROPOSITION 4. *If* F *is a face of* $\mathbb{S}_+^n$, $B = \{i : \exists \mathbf{X} \in F, (\mathbf{U_X})_{ii} \neq 0\}$, *and* $\tilde{\mathbf{X}} \in F$, *then*

1. $(\mathbf{U}_{\tilde{\mathbf{X}}})_{ii} = 0 \ \forall i \notin B$ *and*
2. $\tilde{\mathbf{X}} \in \mathrm{relint}(F) \iff (\mathbf{U}_{\tilde{\mathbf{X}}})_{ii} > 0 \ \forall i \in B$.

We now turn our attention to the primal-dual SDP problems.

Let $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ denote the linear operator $\mathbf{X} \mapsto (\mathbf{A}^{(i)} \bullet \mathbf{X})_{i=1}^m$, and let $\mathcal{A}^*$ denote its adjoint operator $\mathbf{y} \mapsto \sum_{i=1}^m \mathbf{A}^{(i)}\mathbf{y}_i$.

We assume the following Slater condition.

ASSUMPTION 5. *There are symmetric, positive definite matrices* $\mathbf{X}$ *and* $\mathbf{S}$ *satisfying* $\mathcal{A}(\mathbf{X}) = \mathbf{b}$, *and* $\mathbf{S} = \mathbf{C} - \mathcal{A}^*(\mathbf{y})$ *for some* $\mathbf{y} \in \mathbb{R}^m$.

This condition implies that the sets of optimal primal and dual solutions are nonempty and bounded and $\tilde{\mathbf{X}}\tilde{\mathbf{S}} = 0$ for any optimal solutions $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{S}}$. The sets of optimal primal and dual solutions are called the *primal optimal face* and the *dual optimal face* respectively, and are denoted by O_p and O_d, respectively. Let F_p and F_d

denote the minimal faces of $\mathbb{S}_+^n$ containing O_p and O_d, respectively. If we take any $\tilde{\mathbf{X}} \in \mathrm{relint}(O_p)$, then $\tilde{\mathbf{X}} \in \mathrm{relint}(F_p)$, and thus

$$O_p = \{\mathbf{X} \in \mathbb{S}_+^n : \mathcal{R}(\mathbf{X}) \subset \mathbb{V}_p, \ \mathcal{A}(\mathbf{X}) = \mathbf{b}\},$$

where $\mathbb{V}_p$ denotes $\mathcal{R}(\tilde{\mathbf{X}})$. Similarly,

$$O_d = \{\mathbf{S} \in \mathbb{S}_+^n : \mathcal{R}(\mathbf{S}) \subset \mathbb{V}_d, \ \mathbf{S} = \mathbf{C} - \mathcal{A}^*(\mathbf{y}), \ \mathbf{y} \in \mathbb{R}^m\},$$

where $\mathbb{V}_d$ denotes $\mathcal{R}(\tilde{\mathbf{S}})$ for any $\tilde{\mathbf{S}} \in \mathrm{relint}(O_d)$.

Let B and N denote the sets $\{i : \exists \mathbf{X} \in F_p, (\mathbf{L_X})_{ii} \neq 0\}$ and $\{i : \exists \mathbf{S} \in F_d, (\mathbf{U_S})_{ii} \neq 0\}$, respectively.

Since the sets O_p and O_d are orthogonal, we have $\mathcal{R}(\mathbf{X}) \subset \mathcal{N}(\mathbf{S})$ and $\mathcal{R}(\mathbf{S}) \subset \mathcal{N}(\mathbf{X})$ for any $(\mathbf{X}, \mathbf{S}) \in O_p \times O_d$. Thus, $\mathbb{V}_p \perp \mathbb{V}_d$. When $\mathbb{V}_p + \mathbb{V}_d = \mathbb{R}^n$, we say that each $(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}) \in \mathrm{relint}(O_p) \times \mathrm{relint}(O_d)$ is a *pair of strictly complementary solutions*. In terms of the index sets B and N, the orthogonality of O_p and O_d implies $B \cap N = \emptyset$ (and thus $|B| + |N| \leq n$), and the existence of strictly complementary solutions can be characterized by $B \cup N = \{1, \ldots, n\}$, i.e., $|B| + |N| = n$. Let T denote the set $\{1, \ldots, n\} \setminus (B \cup N)$ so that $T = \emptyset$ if and only if there are strictly complementary solutions.

We end this section with a useful lemma.

LEMMA 6. *If $\tilde{\mathbf{X}} \in \mathrm{relint}(O_p)$ and $\tilde{\mathbf{S}} \in \mathrm{relint}(O_d)$, then there exists a lower triangular, square matrix $\mathbf{L}(\tilde{\mathbf{X}}, \tilde{\mathbf{S}})$ with positive diagonal such that*

$$\mathbf{L}(\tilde{\mathbf{X}}, \tilde{\mathbf{S}})\tilde{\mathbf{X}}\mathbf{L}(\tilde{\mathbf{X}}, \tilde{\mathbf{S}})^T = \mathbf{I}_B$$

and

$$\mathbf{L}(\tilde{\mathbf{X}}, \tilde{\mathbf{S}})^{-T}\tilde{\mathbf{S}}\mathbf{L}(\tilde{\mathbf{X}}, \tilde{\mathbf{S}})^{-1} = \mathbf{I}_N.$$

Proof. In the proof of uniqueness for Proposition 1, we see that $\mathbf{L}^{-1}\tilde{\mathbf{X}}\mathbf{L}^{-T} = \mathbf{I}_B$, where $\mathbf{L} = \mathbf{L}_{\tilde{\mathbf{X}}} + \mathbf{I}_{N \cup T}$. From the positive semidefiniteness and complementarity of $\mathbf{L}^{-1}\tilde{\mathbf{X}}\mathbf{L}^{-T}$ and $\mathbf{L}^T\tilde{\mathbf{S}}\mathbf{L}$, we conclude that $(\mathbf{L}^T\tilde{\mathbf{S}}\mathbf{L})_{ii} = 0$ whenever $i \in B$. Thus, the ith row of $\mathbf{U}_{\mathbf{L}^T\tilde{\mathbf{S}}\mathbf{L}}$ is a zero row whenever $i \in B$. Consequently, $(\mathbf{U}^T\mathbf{L}^{-1})\tilde{\mathbf{X}}(\mathbf{U}^T\mathbf{L}^{-1})^T = \mathbf{U}^T\mathbf{I}_B\mathbf{U} = \mathbf{I}_B$, where $\mathbf{U} = \mathbf{U}_{\mathbf{L}^T\tilde{\mathbf{S}}\mathbf{L}} + \mathbf{I}_{B \cup T}$. Finally, $(\mathbf{U}^T\mathbf{L}^{-1})^{-T}\tilde{\mathbf{S}}(\mathbf{U}^T\mathbf{L}^{-1})^{-1} = \mathbf{U}^{-1}(\mathbf{L}^T\tilde{\mathbf{S}}\mathbf{L})\mathbf{U}^{-T} = \mathbf{I}_N$. $\square$

3. Weighted centers for SDP. One of the many existing notions of weighted centers for SDP is the weighted centers defined by the following set of equations:

$$\begin{aligned} \mathcal{A}(\mathbf{X}) &= \mathbf{b}, \quad \mathbf{S} = \mathbf{C} - \mathcal{A}^*(\mathbf{y}) \quad \text{for some } \mathbf{y} \in \mathbb{R}^m, \\ \mathbf{L_X}^T\mathbf{S}\mathbf{L_X} &= \mathbf{W}, \quad \mathbf{X} \succ \mathbf{0}, \quad \text{and} \quad \mathbf{S} \succ \mathbf{0}. \end{aligned}$$

(3.1)

Here, the symmetric matrix $\mathbf{W}$ plays the role of the weights. We recover the usual analytic centers by setting $\mathbf{W}$ to a positive multiple of $\mathbf{I}$, in which case any solution is the unique minimizer of a shifted logarithmic determinant barrier, which is strictly convex over the primal-dual feasible region.

When $\mathbf{W}$ is not a positive multiple of $\mathbf{I}$, a result of Monteiro and Zanjácomo [16], which was improved upon by Tunçel and Wolkowicz [22], states that (3.1) has locally unique solutions when $\|\mathbf{W} - \mu\mathbf{I}\|_2 < (\sqrt{3} - 1)\mu$. This result was recently extended by the author and Tunçel [5] to include all $\mathbf{W}$ satisfying $\|\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2} -$

$\mu \mathbf{I}\|_2 < \sqrt{\alpha_{\min}/(2\alpha_{\max})}\mu$ for any diagonal matrix $\mathbf{D} = \text{Diag}(\alpha_1, \ldots, \alpha_n)$ with positive diagonal entries, where $\alpha_{\min}$ and $\alpha_{\max}$ denote $\min\{\alpha_1, \ldots, \alpha_n\}$ and $\max\{\alpha_1, \ldots, \alpha_n\}$, respectively. This extension includes all positive, diagonal matrices $\mathbf{W}$.

In the case when $\mathbf{W}$ is a positive, diagonal matrix, we shall further prove that (3.1) has a (globally) unique solution by showing that any solution is the unique minimizer of some shifted, weighted logarithmic barrier that is strictly convex over the primal-dual feasible region.

3.1. Weighted barriers for semidefinite cones. Fix some arbitrary positive constants $w_1, \ldots, w_n$ and consider the barrier f on the cone $\mathbb{S}_{++}^n$ defined by

$$\mathbf{X} \mapsto -\sum_{i=1}^{n} w_i \log(\mathbf{L_X})_{ii}^2.$$

This is called the *weighted barrier with weights* w_i. The following proposition shows that the weighted barrier is strictly convex.

PROPOSITION 7. *The gradient and Hessian of the weighted barrier f are respectively given by*

$$g(\mathbf{X}) = -\mathbf{L_X}^{-T}\mathbf{W}\mathbf{L_X}^{-1} \quad and \quad H(\mathbf{X}) : \mathbf{V} \mapsto \mathbf{L_X}^{-T}\mathcal{M}^2(\mathbf{L_X}^{-1}\mathbf{V}\mathbf{L_X}^{-T})\mathbf{L_X}^{-1},$$

where $\mathbf{W} = \text{Diag}(w_1, \ldots, w_n)$ *and* $\mathcal{M}$ *denotes the map on* $\mathbb{S}^n$ *defined by* $\mathbf{V}_{ij} \mapsto \sqrt{w_{\min\{i,j\}}}\mathbf{V}_{ij}$.

Proof. For the proof, see [5, section 6]. □

As a consequence, the shifted barrier $f_+ : \mathbf{X} \mapsto f(\mathbf{X}) + \mathbf{C} \bullet \mathbf{X}$ has a unique minimizer $\mathbf{X}$ over the primal feasible region, and $\mathbf{X}$ satisfies the Karush–Kuhn–Tucker conditions

$$\mathcal{A}(\mathbf{X}) = \mathbf{b}, \quad \mathbf{S} = \mathbf{C} - \mathcal{A}^*(\mathbf{y}), \quad and \quad \mathbf{L_X}^T\mathbf{S}\mathbf{L_X} = \mathbf{W}$$

for some $\mathbf{S} \in \mathbb{S}_{++}^n$ and some $\mathbf{y} \in \mathbb{R}^m$. The matrix $\mathbf{S}$ is also uniquely determined and can be characterized as the unique minimizer of the shifted barrier $f_+^* : \mathbf{S} \mapsto f^*(\mathbf{S}) + \bar{\mathbf{X}} \bullet \mathbf{S}$ over the dual feasible region, where

$$f^* : \mathbf{S} \mapsto -\sum_{i=1}^{n} w_i \log(\mathbf{U_S})_{ii}^2$$

is the conjugate functional of f and $\bar{\mathbf{X}}$ is an arbitrary primal feasible solution. Thus, we have proven the following.

THEOREM 8. *The weighted analytic center equation* (3.1) *uniquely determines a pair of solutions* $(\mathbf{X}, \mathbf{S})$ *whenever* $\mathbf{W}$ *is a positive, diagonal matrix.*

Hence, given positive weights $w_1, \ldots, w_n$, we can define the *primal-dual weighted analytic centers* either via the weighted centers equations (3.1), where $\mathbf{W}$ is the diagonal matrix $\text{Diag}(w_1, \ldots, w_n)$, or as minimizers of the shifted barriers $f_+ + f_+^*$ over the primal-dual feasible region.

Unfortunately, unlike the weighted centers for LP, these weighted centers do not fill up the whole relative interior of the primal-dual feasible region, i.e., not all strictly feasible solutions $(\mathbf{X}, \mathbf{S})$ are weighted centers. This drawback can be easily rectified by considering orthonormal similarity transformations on both primal and dual problems.[1]

---

[1]In fact, we can also use general similarity transformations $\mathtt{P} : \mathbf{X} \mapsto \mathbf{P}^{-1}\mathbf{X}\mathbf{P}$, where $\mathbf{P}$ is an invertible matrix, but orthonormal similarity transformations are sufficient for the purpose of this paper.

THEOREM 9. *For each pair of primal-dual strictly feasible solutions* $(\mathbf{X}, \mathbf{S})$, *there exists an orthogonal matrix* $\mathbf{Q}$ *such that under the orthonormal similarity transformation* $\mathcal{Q} : \mathbf{Z} \mapsto \mathbf{Q}^T \mathbf{Z} \mathbf{Q}$ *on both primal and dual problems, the resulting pair of strictly feasible solutions* $(\mathcal{Q}(\mathbf{X}), \mathcal{Q}(\mathbf{S}))$ *is a pair of weighted centers whose weights are eigenvalues of* $\mathbf{X}\mathbf{S}$.

Proof. Consider a Schur decomposition $\mathbf{Q}^T \mathbf{X}\mathbf{S}\mathbf{Q} = \mathbf{L}$ of the product $\mathbf{X}\mathbf{S}$, where $\mathbf{Q}$ is an orthogonal matrix and $\mathbf{L}$ is a lower triangular matrix with eigenvalues of $\mathbf{X}\mathbf{S}$ on its diagonal. Under the orthonormal similarity transformation $\mathcal{Q} : \mathbf{X} \mapsto \mathbf{Q}^T \mathbf{X} \mathbf{Q}$, we see that $\mathcal{Q}(\mathbf{X})\mathcal{Q}(\mathbf{S}) = \mathbf{L}$. Thus, $\mathbf{L}_{\mathcal{Q}(\mathbf{X})}^T \mathcal{Q}(\mathbf{S})\mathbf{L}_{\mathcal{Q}(\mathbf{X})} = \mathbf{L}_{\mathcal{Q}(\mathbf{X})}^{-1}\mathbf{L}\mathbf{L}_{\mathcal{Q}(\mathbf{X})}$ is both symmetric and lower triangular, and hence diagonal. Clearly, this diagonal matrix shares the same diagonal entries with $\mathbf{L}$. □

Therefore, we can obtain a collection of weighted centers that "fills up" the whole interior of the primal-dual feasible region by generalizing the notion of weighted centers to include all primal-dual pairs $(\mathbf{X}, \mathbf{S})$ satisfying

$$\mathbf{A}^{(i)} \bullet \mathbf{X} = \mathbf{b}_i \text{ for } i = 1, \ldots, m,$$

(3.2)
$$\mathbf{S} + \sum_{i=1}^{m} \mathbf{A}^{(i)} \mathbf{y}_i = \mathbf{C} \text{ for some } \mathbf{y} \in \mathbb{R}^m,$$

$$\mathbf{L}_{\mathcal{Q}(\mathbf{X})}^T \mathcal{Q}(\mathbf{S})\mathbf{L}_{\mathcal{Q}(\mathbf{X})} = \mathbf{W}, \quad \mathbf{X} \succ \mathbf{0}, \quad \text{and} \quad \mathbf{S} \succ \mathbf{0},$$

for some orthonormal similarity transformation $\mathcal{Q} : \mathbf{X} \mapsto \mathbf{Q}^T \mathbf{X} \mathbf{Q}$ and some positive, diagonal matrix $\mathbf{W}$. These weighted centers can alternatively be defined as the unique minimizers of the shifted, weighted barriers

$$(\mathbf{X}, \mathbf{S}) \mapsto -\sum_{i=1}^{n} w_i \log(\mathbf{L}_{\mathbf{Q}^T \mathbf{X} \mathbf{Q}})_{ii}^2 - \sum_{i=1}^{n} w_i \log(\mathbf{U}_{\mathbf{Q}^T \mathbf{S} \mathbf{Q}})_{ii}^2 + \mathbf{X} \bullet \mathbf{S}$$

over the primal-dual feasible region, where $\mathbf{Q}$ ranges over all orthogonal matrices of order n and $(w_1, \ldots, w_n)^T$ ranges over all positive n-vectors.

3.2. Target map. A natural and useful consequence of weighted centers is the collection of weighted central paths $\mathcal{WCP}(\mathbf{W}, \mathbf{Q}) := \{(\mathbf{X}(\mu, \mathbf{W}, \mathbf{Q}), \mathbf{S}(\mu, \mathbf{W}, \mathbf{Q})) : \mu > 0\}$, where $(\mathbf{X}(\mu, \mathbf{W}, \mathbf{Q}), \mathbf{S}(\mu, \mathbf{W}, \mathbf{Q})$ is the solution to (3.2) with $\mathbf{W}$ replaced by $\mu \mathbf{W}$.

Since Schur decomposition is generally not unique, we may have two or more weighted central paths passing through the same pair of weighted centers. We address this ambiguity by considering only those Schur decompositions involving lower triangular matrices with diagonal entries in nonincreasing order. In another words, we consider only weighted central paths corresponding to those $\mathbf{W}$ with diagonal entries arranged in nonincreasing order.

Suppose $(\mathbf{X}, \mathbf{S}) = (\mathbf{X}(1, \mathbf{W}, \mathbf{Q}), \mathbf{S}(1, \mathbf{W}, \mathbf{Q}))$ is a pair of weighted centers on the weighted central path $\mathcal{WCP}(\mathbf{W}, \mathbf{Q})$. Consider any weighted central path $\mathcal{WCP}(\mathbf{W}', \mathbf{Q}')$ passing through $(\mathbf{X}, \mathbf{S})$, say, $(\mathbf{X}, \mathbf{S}) = (\mathbf{X}(\mu', \mathbf{W}', \mathbf{Q}'), \mathbf{S}(\mu', \mathbf{W}', \mathbf{Q}'))$ for some $\mu' > 0$. By scaling $\mathbf{W}'$ appropriately, we may assume without any loss of generality that $\mu' = 1$. Now pick an arbitrary pair of weighted centers $(\mathbf{X}(\mu, \mathbf{W}, \mathbf{Q}), \mathbf{S}(\mu, \mathbf{W}, \mathbf{Q}))$ on $\mathcal{WCP}(\mathbf{W}, \mathbf{Q})$. By definition of weighted centers,

$$\mathbf{Q}^T \mathbf{X}(1, \mathbf{W}, \mathbf{Q})\mathbf{S}(1, \mathbf{W}, \mathbf{Q})\mathbf{Q} = \mathbf{L},$$

$$\mathbf{Q}^T \mathbf{X}(\mu, \mathbf{W}, \mathbf{Q})\mathbf{S}(\mu, \mathbf{W}, \mathbf{Q})\mathbf{Q} = \mu\check{\mathbf{L}},$$

and

$$(\mathbf{Q}')^T \mathbf{X}(1, \mathbf{W}', \mathbf{Q}') \mathbf{S}(1, \mathbf{W}', \mathbf{Q}') \mathbf{Q}' = \mathbf{L}'$$

are Schur decompositions, where the diagonal entries of both $\mathbf{L}$ and $\check{\mathbf{L}}$ are precisely the diagonal entries of $\mathbf{W}$ in the same nonincreasing order, and the diagonal entries of $\mathbf{L}'$ are those of $\mathbf{W}'$ in the same order. Since

$$\mathbf{L} = \mathbf{Q}^T \mathbf{X} \mathbf{S} \mathbf{Q} = \mathbf{Q}^T (\mathbf{Q}') \mathbf{L}' (\mathbf{Q}')^T \mathbf{Q}$$

and the diagonal entries of both $\mathbf{L}$ and $\mathbf{L}'$ are in nonincreasing order, it follows from Proposition 21 (Appendix A) that the diagonal entries of $\mathbf{L}$ and $\mathbf{L}'$ (whence those of $\mathbf{W}$ and $\mathbf{W}'$) coincide, and $(\mathbf{Q}')^T \mathbf{Q}$ is a block-diagonal matrix where the size of the ith block is the multiplicity of the ith largest (distinct) eigenvalue of $\mathbf{X} \mathbf{S}$. It then follows from Proposition 21 that

$$\check{\mathbf{L}}' := \mu^{-1} (\mathbf{Q}')^T \mathbf{X}(\mu, \mathbf{W}, \mathbf{Q}) \mathbf{S}(\mu, \mathbf{W}, \mathbf{Q}) \mathbf{Q}' = (\mathbf{Q}')^T \mathbf{Q} \check{\mathbf{L}} \mathbf{Q}^T \mathbf{Q}'$$

is a lower triangular matrix and has diagonal entries in nonincreasing order (and thus shares the same diagonal as $\check{\mathbf{L}}$). Thus

$$(\mathbf{X}(\mu, \mathbf{W}, \mathbf{Q}), \mathbf{S}(\mu, \mathbf{W}, \mathbf{Q})) = (\mathbf{X}(\mu, \mathbf{W}', \mathbf{Q}'), \mathbf{S}(\mu, \mathbf{W}', \mathbf{Q}')) \in \mathcal{WCP}(\mathbf{W}', \mathbf{Q}').$$

Since $(\mathbf{X}(\mu, \mathbf{W}, \mathbf{Q}), \mathbf{S}(\mu, \mathbf{W}, \mathbf{Q}))$ is arbitrary, we conclude that $\mathcal{WCP}(\mathbf{W}, \mathbf{Q}) \subseteq \mathcal{WCP}(\mathbf{W}', \mathbf{Q}')$. Repeating the argument on any arbitrary pair of weighted centers in $\mathcal{WCP}(\mathbf{W}', \mathbf{Q}')$ shows that

$$\mathcal{WCP}(\mathbf{W}, \mathbf{Q}) = \mathcal{WCP}(\mathbf{W}', \mathbf{Q}').$$

We have thus proved the next theorem.

THEOREM 10. *For each pair of primal-dual strictly feasible solutions* $(\mathbf{X}, \mathbf{S})$, *there exists exactly one weighted central path* $\mathcal{WCP}(\mathbf{W}, \mathbf{Q})$ *passing through* $(\mathbf{X}, \mathbf{S})$, *where diagonal entries of* $\mathbf{W}$ *are in nonincreasing order .*

As a direct consequence of the above argument, we have the next theorem.

THEOREM 11. *The map*

$$(\mathbf{X}, \mathbf{S}) \mapsto \mathbf{Q} \mathbf{D} \mathbf{Q}^T,$$

where $\mathbf{Q}^T \mathbf{X} \mathbf{S} \mathbf{Q} = \mathbf{L}$ *is a Schur decomposition of* $\mathbf{X} \mathbf{S}$ *with diagonal entries of* $\mathbf{L}$ *in nonincreasing order, and* $\mathbf{D}$ *is the diagonal matrix sharing the same diagonal with* $\mathbf{L}$, *is a bijection between the set of primal-dual strictly feasible solutions and the cone* $\mathbb{S}^n_{++}$.

Proof. Using Proposition 21 as before, if

$$\mathbf{Q}^T \mathbf{X} \mathbf{S} \mathbf{Q} = \mathbf{L}$$

and

$$(\mathbf{Q}')^T \mathbf{X} \mathbf{S} \mathbf{Q}' = \mathbf{L}'$$

are two Schur decompositions of $\mathbf{X} \mathbf{S}$ where the diagonal entries of both Schur forms $\mathbf{L}$ and $\mathbf{L}'$ are arranged in nonincreasing order, then the diagonal entries of $\mathbf{L}$ and $\mathbf{L}'$ coincide, and $(\mathbf{Q}') \mathbf{Q}$ is a block-diagonal matrix where the size of the ith block is the multiplicity of the ith largest (distinct) eigenvalue of $\mathbf{X} \mathbf{S}$. Consequently

$$\mathbf{Q}^T \mathbf{D} \mathbf{Q} = (\mathbf{Q}') \mathbf{D} \mathbf{Q}',$$

where $\mathbf{D}$ is the diagonal matrix sharing the same diagonal with $\mathbf{L}$ and $\mathbf{L}'$.   □

3.3. Weighted central paths under strict complementarity. The main result in this subsection states that every (primal) weighted central path $\{\mathbf{X}(\mu) : \mu > 0\}$ converges to weighted analytic centers of optimal faces, where $(\mathbf{X}(\mu), \mathbf{S}(\mu))$ is the solution to (3.1) with $\mathbf{W} = \mu\mathrm{Diag}(w_1, \ldots, w_n)$.

We begin by proving a result on the limit points of weighted central paths.

LEMMA 12. *All limit points of the weighted central path lie in the relative interior of the primal optimal face.*

Proof. Suppose that $\mathbf{X}$ is a limit point of the weighted central path. Clearly, from the Karush–Kuhn–Tucker conditions, $\mathbf{X} \in O_p$. So, it suffices to show that $\mathrm{rank}(\mathbf{X}) = |B|$. Let $\{\mathbf{X}(\mu_k)\}_{k=1}^\infty$ be a subsequence converging to $\mathbf{X}$. Since $\{\mathbf{X}(\mu_k)\}$ is bounded, so is $\{\mathbf{L}_{\mathbf{X}(\mu_k)}\}$. So, by choosing a subsequence of $\{\mathbf{X}(\mu_k)\}$ if necessary, we may assume that $\{\mathbf{L}_{\mathbf{X}(\mu_k)}\}$ converges to some lower triangular matrix $\mathbf{L}$. Clearly, $\mathbf{X} = \mathbf{L}\mathbf{L}^T$. Let $\tilde{\mathbf{X}} \in \mathrm{relint}(O_p)$ and $\tilde{\mathbf{S}} \in O_d$ be arbitrary. Now, $(\mathbf{X}(\mu_k) - \tilde{\mathbf{X}}) \bullet (\mathbf{S}(\mu_k) - \tilde{\mathbf{S}}) = 0$, $\mathbf{X}(\mu_k) \bullet \mathbf{S}(\mu_k) = \mu_k \sum_{i=1}^n w_i$, and $\tilde{\mathbf{X}} \bullet \tilde{\mathbf{S}} = 0$ imply that $\mathbf{X}(\mu_k) \bullet \tilde{\mathbf{S}} + \mathbf{S}(\mu_k) \bullet \tilde{\mathbf{X}} = \mu_k \sum_{i=1}^n w_i$. Consequently,

$$\mu_k \sum_{i=1}^n w_i \geq \mathbf{S}(\mu_k) \bullet \tilde{\mathbf{X}}$$
$$= tr[\sqrt{\mu_k}\sqrt{\mathbf{W}}(\mathbf{L}_{\mathbf{X}(\mu_k)})^{-1}\mathbf{L}_{\tilde{\mathbf{X}}}][\sqrt{\mu_k}\sqrt{\mathbf{W}}(\mathbf{L}_{\mathbf{X}(\mu_k)})^{-1}\mathbf{L}_{\tilde{\mathbf{X}}}]^T,$$

from which it follows that $\sqrt{\sum_{i=1}^n w_i} \geq \sqrt{w_i}(\mathbf{L}_{\tilde{\mathbf{X}}})_{ii}/(\mathbf{L}_{\mathbf{X}(\mu_k)})_{ii}$. Since $\tilde{\mathbf{X}}$ lies in the relative interior of F_p, it follows from Proposition 3 that $(\mathbf{L}_{\tilde{\mathbf{X}}})_{ii} > 0$ for all $i \in B$. Thus,

$$\mathbf{L}_{ii} = \lim_{k \to \infty} (\mathbf{L}_{\mathbf{X}(\mu_k)})_{ii} \geq \frac{\sqrt{w_i}(\mathbf{L}_{\tilde{\mathbf{X}}})_{ii}}{\sqrt{\sum_{i=1}^n w_i}} > 0 \quad \forall i \in B.$$

This implies that $\mathrm{rank}(\mathbf{L}) \geq |B|$, and hence $\mathrm{rank}(\mathbf{X}) = |B|$. $\square$

Under strict complementarity, the central path for an SDP problem converges to the analytic center of the optimal face (see [8, 6, 14]). We now generalize this result to the weighted central paths.

Recall from Proposition 3 that for any $\mathbf{X} \in F_p$, $\mathbf{X}$ is in the relative interior of F_p if and only if $(\mathbf{L}_{\mathbf{X}})_{ii} > 0 \ \forall i \in B$. Thus, the functional $f_p : \mathrm{relint}(F_p) \to \mathbb{R}$ defined by

$$\mathbf{X} \mapsto -\sum_{i \in B} w_i \log(\mathbf{L}_{\mathbf{X}})_{ii}^2$$

induces a barrier for the primal optimal face O_p. We shall show that under strict complementarity, every limit point of the weighted central path solves

$$\min\{f_p(\mathbf{X}) : \mathcal{A}(\mathbf{X}) = \mathbf{b}, \ \mathbf{X} \in \mathrm{span}(F_p)\},$$

where $\mathrm{span}(F_p)$ denotes the smallest linear subspace containing F_p.

LEMMA 13. *If the primal-dual pair of SDP problems has strictly complementary solutions, and the subsequence $\{(\mathbf{X}(\mu_k), \mathbf{S}(\mu_k))\}$ converges to $(\mathbf{I}_B, \mathbf{I}_N)$, then*

1. $(\mathbf{L}_{\mathbf{X}(\mu_k)})_{ij} = o(1) \ \forall i \in B, \ j \neq i$, and $(\mathbf{U}_{\mathbf{S}(\mu_k)})_{ij} = o(1) \ \forall i \in N, \ j \neq i$, *and*
2. $(\mathbf{L}_{\mathbf{X}(\mu_k)})_{ij} = o(\sqrt{\mu_k}) \ \forall i \in N, \ j \neq i$, *and* $(\mathbf{U}_{\mathbf{S}(\mu_k)})_{ij} = o(\sqrt{\mu_k}) \ \forall i \in B, \ j \neq i$.

Proof. From $(\mathbf{X}(\mu_k) - \mathbf{I}_B) \bullet (\mathbf{S}(\mu_k) - \mathbf{I}_N) = 0$, it follows that $\mathrm{tr}\mathbf{X}(\mu_k)_N + \mathrm{tr}\mathbf{S}(\mu_k)_B = \mu_k \sum_{i=1}^{n} w_i$. Expanding the left-hand side gives

$$\sum_{\substack{i \in N \\ j \leq i}} (\mathbf{L}_{\mathbf{X}(\mu_k)})_{ij}^2 + \sum_{\substack{i \in B \\ j \geq i}} (\mathbf{U}_{\mathbf{S}(\mu_k)})_{ij}^2$$

$$= \sum_{i \in N} (\mathbf{L}_{\mathbf{X}(\mu_k)})_{ii}^2 + \sum_{i \in B} (\mathbf{U}_{\mathbf{S}(\mu_k)})_{ii}^2 + \sum_{\substack{i \in N \\ j < i}} (\mathbf{L}_{\mathbf{X}(\mu_k)})_{ij}^2 + \sum_{\substack{i \in B \\ j > i}} (\mathbf{U}_{\mathbf{S}(\mu_k)})_{ij}^2.$$

From $(\mathbf{L}_{\mathbf{X}(\mu_k)})^T \mathbf{S}(\mu_k) \mathbf{L}_{\mathbf{X}(\mu_k)} = \mu_k \mathbf{W}$, we get $(\mathbf{U}_{\mathbf{S}(\mu_k)})_{ii}(\mathbf{L}_{\mathbf{X}(\mu_k)})_{ii} = \sqrt{\mu_k w_i}$. Therefore,

$$\sum_{i=1}^{n} w_i = \sum_{i \in B} \frac{w_i}{(\mathbf{L}_{\mathbf{X}(\mu_k)})_{ii}^2} + \sum_{i \in N} \frac{w_i}{(\mathbf{U}_{\mathbf{S}(\mu_k)})_{ii}^2} + \frac{1}{\mu_k} \left(\sum_{\substack{i \in N \\ j < i}} (\mathbf{L}_{\mathbf{X}(\mu_k)})_{ij}^2 + \sum_{\substack{i \in B \\ j > i}} (\mathbf{U}_{\mathbf{S}(\mu_k)})_{ij}^2 \right)$$

$$= \sum_{i \in B} \frac{w_i}{\mathbf{X}(\mu_k)_{ii}} + \sum_{i \in N} \frac{w_i}{\mathbf{S}(\mu_k)_{ii}} + \sum_{\substack{i \in N \\ j < i}} \frac{(\mathbf{L}_{\mathbf{X}(\mu_k)})_{ij}^2}{\mu_k} + \sum_{\substack{i \in B \\ j > i}} \frac{(\mathbf{U}_{\mathbf{S}(\mu_k)})_{ij}^2}{\mu_k}$$

$$+ \sum_{i \in B} \left(\frac{w_i}{(\mathbf{L}_{\mathbf{X}(\mu_k)})_{ii}^2} - \frac{w_i}{\mathbf{X}(\mu_k)_{ii}} \right) + \sum_{i \in N} \left(\frac{w_i}{(\mathbf{U}_{\mathbf{S}(\mu_k)})_{ii}^2} - \frac{w_i}{\mathbf{S}(\mu_k)_{ii}} \right).$$

Since

$$\mathbf{S}(\mu_k)_{ii} = \sum_{j > i} ((\mathbf{U}_{\mathbf{S}(\mu_k)})_{ij})^2 + ((\mathbf{U}_{\mathbf{S}(\mu_k)})_{ii})^2 \geq ((\mathbf{U}_{\mathbf{S}(\mu_k)})_{ii})^2$$

and

$$\mathbf{X}(\mu_k)_{ii} = \sum_{j < i} ((\mathbf{L}_{\mathbf{X}(\mu_k)})_{ij})^2 + ((\mathbf{L}_{\mathbf{X}(\mu_k)})_{ii})^2 \geq ((\mathbf{L}_{\mathbf{X}(\mu_k)})_{ii})^2,$$

the summands in the last two sums are nonnegative. Thus, the right-hand side is at least the sum $\sum_{i \in N} \frac{w_i}{\mathbf{S}(\mu_k)_{ii}} + \sum_{i \in B} \frac{w_i}{\mathbf{X}(\mu_k)_{ii}}$, which, under strict complementarity and the assumptions $\mathbf{X}(\mu_k) \to \mathbf{I}_B$ and $\mathbf{S}(\mu_k) \to \mathbf{I}_N$, converges to the left-hand side as $k \to \infty$. This can occur only when all summands in the last four sums converge to zero. $\quad\square$

We now give the main theorem of this section.

THEOREM 14. *If there are strictly complementary solutions to the primal-dual SDP problems, then the weighted central path for the primal problem converges to the solution of*

(3.3)
$$\min \quad -\sum_{i \in B} w_i \log(\mathbf{L}_{\mathbf{X}})_{ii}^2$$
$$s.t. \quad \mathbf{A}^{(i)} \bullet \mathbf{X} = \mathbf{b}_i, \quad i = 1, \ldots, m,$$
$$\mathbf{X} \in \mathrm{span}(F_p),$$

where F_p is the minimal face of $\mathbb{S}_+^n$ containing the primal optimal face and $B = \{i : \exists \mathbf{X} \in F_p, (\mathbf{L}_{\mathbf{X}})_{ii} \neq 0\}$.

Proof. Suppose $\tilde{\mathbf{X}}$ is an arbitrary limit point of the weighted central path. By Lemma 12, $\tilde{\mathbf{X}} \in \mathrm{relint}(F_p)$. Since $\mathbf{S}(\mu)$ is bounded as $\mu \downarrow 0$, we can choose a sequence

$\{\mu_k\}$ of positive real numbers converging to zero such that $\mathbf{X}(\mu_k) \to \tilde{\mathbf{X}}$, and $\mathbf{S}(\mu_k)$ is convergent with limit $\tilde{\mathbf{S}}$. Let $\mathbf{L}$ denote the matrix $\mathbf{L}(\tilde{\mathbf{X}}, \tilde{\mathbf{S}})$ in the statement of Lemma 6. Since f_p is invariant, up to an additive constant, under the transformation $\mathcal{G} : \mathbf{X} \mapsto \mathbf{L}\mathbf{X}\mathbf{L}^T$, the limit point $\tilde{\mathbf{X}}$ solves (3.3) if and only if $\mathbf{I}_B = \mathcal{G}(\tilde{\mathbf{X}})$ solves

$$(3.4) \qquad \min \left\{ -\sum_{i \in B} w_i \log(\mathbf{L}_{\mathbf{X}})_{ii}^2 : \mathcal{A}(\mathcal{G}^{-1}(\mathbf{X})) = \mathbf{b}, \ \mathbf{X} \in \mathrm{span}(\mathcal{G}(F_p)) \right\}.$$

The matrix $\mathbf{I}_B$ solves (3.4) if and only if the optimality condition

$$\nabla f_p(\mathbf{I}_B) \in \mathrm{span}(\{\mathcal{G}^{-*}(\mathbf{A}^{(1)}), \ldots, \mathcal{G}^{-*}(\mathbf{A}^{(m)})\}) + \mathrm{span}(\mathcal{G}(F_p))^{\perp}$$

holds, where $\mathcal{G}^{-*}$ denotes the adjoint of the inverse of $\mathcal{G}$. Let $\hat{\mathbf{X}}(\mu_k)$ and $\hat{\mathbf{S}}(\mu_k)$ denote $\mathcal{G}(\mathbf{X}(\mu_k))$ and $\mathcal{G}^{-*}(\mathbf{S}(\mu_k))$, respectively. Let $\hat{\mathbf{A}}^{(i)}$ denote $\mathcal{G}^{-*}(\mathbf{A}^{(i)})$. From the description (2.4) of faces of $\mathbb{S}_+^n$ and the assumption that $\mathbf{I}_B \in \mathrm{relint}(\mathcal{G}(F_p))$, we deduce that $\mathrm{span}(\mathcal{G}(F_p))^{\perp} = \{\mathbf{X} \in \mathbb{S}_+^n : \mathbf{X}_{BB} = \mathbf{0}\}$. Thus, the optimality condition is equivalent to

$$\nabla f_p(\mathbf{I}_B)_{BB}(= \mathbf{W}_{BB}) \in \mathrm{span}(\{(\hat{\mathbf{A}}^{(1)})_{BB}, \ldots, (\hat{\mathbf{A}}^{(m)})_{BB}\}).$$

Let $\mathbb{V}$ denote the subspace $\mathrm{span}(\{(\hat{\mathbf{A}}^{(1)})_{BB}, \ldots, (\hat{\mathbf{A}}^{(m)})_{BB}\})$. Since $\hat{\mathbf{S}}(\mu_k) - \mathbf{I}_N = \mathcal{G}^{-*}(\mathbf{S}(\mu_k) - \hat{\mathbf{S}}) \in \mathrm{span}(\{\hat{\mathbf{A}}^{(1)}, \ldots, \hat{\mathbf{A}}^{(m)}\})$, we have that $(\hat{\mathbf{S}}(\mu_k))_{BB} \in \mathbb{V}$. Dividing by μ_k gives

$$\frac{(\hat{\mathbf{S}}(\mu_k))_{BB}}{\mu_k} \in \mathbb{V}.$$

By Lemma 13, $(\mathbf{U}_{(\hat{\mathbf{S}}(\mu_k))})_{ij}/\sqrt{\mu_k} \to 0$ for all $i \in B$ and $j > i$, and $(\mathbf{L}_{(\hat{\mathbf{X}}(\mu_k))})_{ij} \to 0$ for all $i \in B$ and $j < i$. Together with $\sum_{j=1}^{i}(\mathbf{L}_{\hat{\mathbf{X}}(\mu_k)})_{ij}^2 = (\hat{\mathbf{X}}(\mu_k))_{ii} \to 1$ for all $i \in B$, it follows that $(\mathbf{L}_{\hat{\mathbf{X}}(\mu_k)})_{ii} \to 1$. Thus, we deduce from $(\mathbf{U}_{\hat{\mathbf{S}}(\mu_k)})_{ii}(\mathbf{L}_{\hat{\mathbf{X}}(\mu_k)})_{ii} = \sqrt{\mu_k w_i}$ that $(\hat{\mathbf{S}}(\mu_k))_{BB}/\mu_k \to \mathbf{W}_{BB}$. Finally, since $\mathbb{V}$ is closed, the theorem follows. □

4. Application to homogeneous cone programming. In this section, we consider the following primal-dual pair of HCP problems:

$$\begin{aligned} \inf \quad & \mathbf{c}^T\mathbf{x} \\ \text{s.t.} \quad & (\mathbf{a}^{(i)})^T\mathbf{x} = \mathbf{b}_i \quad \text{for } i = 1, \ldots, m, \\ & \mathbf{x} \in \mathrm{cl}(K), \end{aligned}$$

and

$$\begin{aligned} \sup \quad & \mathbf{b}^T\mathbf{y} \\ \text{s.t.} \quad & \mathbf{s} = \mathbf{c} - \sum_{i=1}^{m}(\mathbf{a}^{(i)})\mathbf{y}_i, \\ & \mathbf{s} \in \mathrm{cl}(K^*), \end{aligned}$$

where K is a d-dimensional homogeneous cone (i.e., a pointed, open, convex cone whose group of automorphisms acts transitively on it), $K^* := \{\mathbf{s} : \mathbf{x}^T\mathbf{s} > 0 \ \forall \mathbf{x} \in K\}$ is its dual cone, the $\mathbf{a}^{(i)}$, $\mathbf{c}$, and $\mathbf{x}$ are real d-vectors, and $\mathbf{b} = (\mathbf{b}_1, \ldots, \mathbf{b}_m)^T$ and $\mathbf{y} = (\mathbf{y}_i, \ldots, \mathbf{y}_m)^T$ are real m-vectors.

As before, we assume the following Slater condition.

ASSUMPTION 15. *There exist an* $\mathbf{x} \in K$ *and an* $\mathbf{s} \in K^*$ *satisfying* $(\mathbf{a}^{(i)})^T\mathbf{x} = b_i$ *for* $i = 1, \ldots, m$ *and* $\mathbf{s} = \mathbf{c} - \sum_{i=1}^{m}(\mathbf{a}^{(i)})\mathbf{y}_i$ *for some* $y = (\mathbf{y}_i, \ldots, \mathbf{y}_m)^T \in \mathbb{R}^m$.

It was shown by the author [3] that all homogeneous cones are SDP-representable, i.e., for each homogeneous cone K, there exists a linear map $\mathbf{M} : \mathbb{R}^d \to \mathbb{S}^n$ such that $\mathbf{x} \in K$ if and only if $\mathbf{M}(\mathbf{x}) \succ \mathbf{0}$. Thus, the primal HCP problem can be reformulated as the primal SDP problem

$$
\begin{aligned}
\min \quad & \mathbf{M}^{-*}(\mathbf{c}) \bullet \mathbf{X} \\
\text{s.t.} \quad & \mathbf{M}^{-*}(\mathbf{a}^{(i)}) \bullet \mathbf{X} = b_i, \quad i = 1, \ldots, m, \\
& \mathbf{X} \in \mathbf{M}(\mathbb{R}^d), \\
& \mathbf{X} \succeq \mathbf{0}.
\end{aligned}
$$

Furthermore, it was shown by the author and Tunçel [5] that HCP problems inherit strict complementarity from the corresponding SDP formulations, i.e., a HCP problem has strictly complementary solutions if and only if any SDP reformulation has such solutions. These establish the foundation for applying Theorem 14 to HCP problems.

4.1. SDP-representability of homogeneous cones. Each d-dimensional homogeneous cone K of rank r can be associated with a T-algebra $\mathbb{A} = \bigoplus_{i,j=1}^{r} \mathbb{A}_{ij}$ with involution * such that K is the cone containing elements of the form $\mathbf{ll}^*$, where $\mathbf{l}$ is a lower triangular element with positive diagonal (see [23]). In fact, each $\mathbf{x} \in K$ uniquely determines a lower triangular element $\mathbf{l}$ with positive diagonal such that $\mathbf{x} = \mathbf{ll}^*$. The reader is strongly encouraged to refer to [3] and [23] for more details.

For each $(i,j) \in \{1, \ldots, r\}^2$, let n_{ij} denote the dimension of $\mathbb{A}_{ij}$ as a vector subspace of $\mathbb{A}$ and let $\mathbf{x}_{ij}$ denote the component of $\mathbf{x} \in \mathbb{A}$ in $\mathbb{A}_{ij}$. From the definition of T-algebras, we have $n_{ij} = n_{ji}$ and $n_{ii} = 1$. Also, $\sum_{i=1}^{r} \sum_{j=1}^{i} n_{ij} = d$.

Let $\mathbb{T}$ denote the subspace $\bigoplus_{1 \leq j \leq i \leq r} \mathbb{A}_{ij}$ of lower triangular elements of $\mathbb{A}$. With each $\mathbf{x} \in \mathbb{A}$, we associate the linear operator $\mathcal{M}(\mathbf{x}) : \mathbb{T} \to \mathbb{T}$ defined by $\mathcal{M}(\mathbf{x}) : \mathbf{l} \mapsto \mathrm{Pr}_{\mathbb{T}} \mathbf{xl}$, where $\mathrm{Pr}_{\mathbb{T}}$ denotes the orthogonal projection onto $\mathbb{T}$ under the inner product $\langle \cdot, \cdot \rangle : (\mathbf{x}, \mathbf{y}) \mapsto tr\mathbf{xy}^*$. The author [3] proved that $\mathbf{x} \in K$ if and only if $\mathcal{M}(\mathbf{x})$ is self-adjoint and positive definite. Thus, for any choice of ordered basis $\mathfrak{B}$ for $\mathbb{T}$, the map

$$
(4.1) \qquad\qquad \mathbf{M}_{\mathfrak{B}} : \mathbb{R}^d \to \mathbb{S}^n : \mathbf{x} \mapsto \mathbf{M}_{\mathfrak{B}}(\mathbf{x}),
$$

where $\mathbf{M}_{\mathfrak{B}}(\mathbf{x})$ is the matrix representing $\mathcal{M}(\mathbf{x})$ under $\mathfrak{B}$, is an SDP-representation of K.

Let $\mathbf{l} \in \mathbb{T}$ be arbitrary. Consider the orthogonal decomposition $\bigoplus_{j=1}^{r}\left(\bigoplus_{i=j}^{r} \mathbb{A}_{ij}\right)$ of $\mathbb{T}$ into columns. Fix an arbitrary $j \in \{1, \ldots, r\}$ and consider the restriction of $\mathcal{M}(\mathbf{l})$ to the jth column $\bigoplus_{i=j}^{r} \mathbb{A}_{ij}$. For each $i \in \{j, \ldots, r\}$, let $\mathfrak{B}_{ij}$ denote a basis for $\mathbb{A}_{ij}$. Since $\mathbf{ly}_{ij} = \sum_{k=i}^{r} \mathbf{l}_{ki}\mathbf{y}_{ij} \in \bigoplus_{k=i}^{r} \mathbb{A}_{kj}$ for each $\mathbf{y}_{ij} \in \mathfrak{B}_{ij}$, the operator $\mathbf{y} \mapsto \mathbf{ly}$ on $\bigoplus_{i=j}^{r} \mathbb{A}_{ij}$ is represented by a lower block-triangular matrix $\mathbf{L}^{(j)}$ under the ordered basis $(\mathfrak{B}_{jj}, \ldots, \mathfrak{B}_{rj})$ of $\bigoplus_{i=j}^{r} \mathbb{A}_{ij}$, where elements in each $\mathfrak{B}_{ij}$ are arbitrarily ordered. Furthermore, $\mathrm{Pr}_{\mathbb{A}_{ij}} \mathbf{ly}_{ij} = \rho_i(\mathbf{l})\mathbf{y}_{ij}$ for each $\mathbf{y}_{ij} \in \mathfrak{B}_{ij}$ implies that the $(i-j+1)$st diagonal block in $\mathbf{L}^{(j)}$ is $\rho_i(\mathbf{l})\mathbf{I}$, where $\rho_i(\mathbf{l})$ is the value of the ith entry on the diagonal of $\mathbf{l}$. Thus, $\mathbf{L}^{(j)}$ is in fact a lower triangular matrix with n_{ij} copies of $\rho_i(\mathbf{l})$ on the diagonal for $i = \{j, \ldots, r\}$. Since for each $j \in \{1, \ldots, r\}$, $\mathcal{M}(\mathbf{l})$ maps the jth column $\bigoplus_{i=j}^{r} \mathbb{A}_{ij}$ into itself, it follows that the linear operator $\mathcal{M}(\mathbf{l})$ can be represented by a lower triangular matrix $\mathbf{L}$ with $\sum_{j=1}^{i} n_{ij}$ copies of $\rho_i(\mathbf{l})$ on the diagonal for $i = 1, \ldots, r$.

LEMMA 16. *There exists an ordered basis $\mathfrak{B}$ for $\mathbb{T}$ such that for each $\mathbf{l} \in \mathbb{T}$ with nonnegative diagonal values, the lower triangular matrix $\mathbf{M}_{\mathfrak{B}}(\mathbf{l})$ is a Cholesky factor of the matrix $\mathbf{M}_{\mathfrak{B}}(\mathbf{ll}^*)$. Moreover, the matrix $\mathbf{M}_{\mathfrak{B}}(\mathbf{l})$ has $\sum_{j=1}^{i} n_{ij}$ copies of $\rho_i(\mathbf{l})$ on its diagonal.*

Proof. Let $\mathfrak{B}$ be the ordered basis $(\mathfrak{B}_{11}, \ldots, \mathfrak{B}_{1r}, \mathfrak{B}_{22}, \ldots, \mathfrak{B}_{2r}, \ldots, \mathfrak{B}_{rr})$. It remains to show that $\mathcal{M}(\mathbf{x}) = \mathcal{M}(\mathbf{l}) \circ \mathcal{M}(\mathbf{l})^*$. This is a special case of Proposition 3.4(iii) of [3]. ☐

Henceforth, we shall use the ordered basis in the lemma to define the SDP representation in (4.1), and drop the subscript $\mathfrak{B}$.

4.2. Optimal faces and strict complementarity of HCP. In this subsection, we extend some results in section 2 to the optimal faces of HCP problems. These extensions rely heavily on the appropriate choice of the ordered basis $\mathfrak{B}$ in Lemma 16.

LEMMA 17. *Each $\mathbf{x} \in cl(K)$ has a unique Cholesky factor $\mathbf{l_x}$ (i.e., a lower triangular element $\mathbf{l_x}$ with nonnegative diagonal values such that $\mathbf{x} = \mathbf{ll}^*$) satisfying*

$$(4.2) \qquad \rho_i(\mathbf{l_x}) = 0 \implies (\mathbf{l_x})_{ji} = 0.$$

Proof. Suppose that $\mathbf{x} \in cl(K)$. Therefore $\mathbf{M}(\mathbf{x})$ is symmetric and positive semidefinite. From the proof of existence of Proposition 1, we see that $\mathbf{L}_{\mathbf{M}(\mathbf{x})+\mu\mathbf{I}} \to \mathbf{L}_{\mathbf{M}(\mathbf{x})}$ as $\mu \to 0$. Since $\mathbf{M}(\mathbf{x}) + \mu\mathbf{I} = \mathbf{M}(\mathbf{x} + \mu\mathbf{e})$, where $\mathbf{e}$ is the unit of the T-algebra $\mathbb{A}$, it follows from Lemma 16 that for each positive μ, $\mathbf{L}_{\mathbf{M}(\mathbf{x})+\mu\mathbf{I}} = \mathbf{M}(\mathbf{l}_{\mathbf{x}+\mu\mathbf{e}})$. Consequently, $\mathbf{L}_{\mathbf{M}(\mathbf{x})} = \mathbf{M}(\mathbf{l_x})$, where $\mathbf{l_x} \in \mathbb{T}$ is any limit point of $\{\mathbf{l}_{\mathbf{x}+\mu\mathbf{e}}\}_{\mu>0}$. The limit point $\mathbf{l_x}$ is clearly a Cholesky factor of $\mathbf{x}$. Property (4.2) for $\mathbf{l_x}$ can be deduced from the same property of $\mathbf{L}_{\mathbf{M}(\mathbf{x})}$ in Proposition 1 and the choice of $\mathfrak{B}$ in Lemma 16. Finally, the uniqueness of $\mathbf{l_x}$ follows straightforwardly from the choice of $\mathfrak{B}$ in Lemma 16 and the uniqueness of $\mathbf{L}_{\mathbf{M}(\mathbf{x})}$. ☐

PROPOSITION 18. *If F is a face of K, $B = \{i : \exists \mathbf{x} \in F, \rho_i(\mathbf{l_x}) \neq 0\}$ and $\tilde{\mathbf{x}} \in F$, then*

1. *$\rho_i(\mathbf{l}_{\tilde{\mathbf{x}}}) = 0 \; \forall i \notin B$ and*
2. *$\tilde{\mathbf{x}} \in \text{relint}(F) \iff \rho_i(\mathbf{l}_{\tilde{\mathbf{x}}}) > 0 \; \forall i \in B$.*

Proof. If F is a face of K, then there exists some face F' of $\mathbb{S}_{++}^n$ such that $\mathbf{M}(F) = \mathbf{M}(\mathbb{R}^d) \cap F'$. Thus, using the description (2.4) of F' we may describe F as

$$F = \{((\mathbf{l}_{\tilde{\mathbf{x}}} + \mathbf{e}_B^c)\mathbf{l_x})(\mathbf{l_x^*}(\mathbf{l}_{\tilde{\mathbf{x}}} + \mathbf{e}_B^c)^*) : \mathbf{x} \in cl(K), \; (i \notin B) \vee (j \notin B) \implies \mathbf{x}_{ij} = 0\},$$

where $\tilde{\tilde{\mathbf{x}}} \in \text{relint}(F)$ is arbitrary and $\mathbf{e}_B^c$ denotes the diagonal element of $\mathbb{A}$ with 0-1 diagonal such that $\rho_i(\mathbf{e}_B^c) = 1$ if and only if $i \notin B$. The theorem then follows from this description. ☐

Since every HCP problem can be reformulated as an SDP problem, we may naturally generalize the notion of strict complementarity from SDP to HCP. However, in order for this generalization to be well defined, different SDP reformulations of the same HCP problem should not result in different conclusions on the existence of strictly complementary solutions. Indeed, the author and Tunçel [5] showed that the existence of strictly complementary solutions is independent of the SDP formulation used. Furthermore, this notion of strictly complementary solutions coincides with a more general notion introduced by Pataki [18], which was shown to be a generic property of linear optimization problems over convex cones by Pataki and Tunçel [19].

4.3. Limit points of central paths for HCP. By reformulating HCP problems as SDP problems, any algorithm for SDP translates directly to an algorithm for

HCP. However, from the perspective of theoretical complexity, it is advantageous for algorithms to use optimal barriers for homogeneous cones. In this subsection, we consider a certain class of optimal barriers for homogeneous cones, and we characterize the limit points of the central paths defined by this class of optimal barriers under strict complementarity.

Since each $\mathbf{x} \in K$ uniquely determines a lower triangular element $\mathbf{l}_\mathbf{x}$ with positive diagonal such that $\mathbf{x} = \mathbf{l}_\mathbf{x} \mathbf{l}_\mathbf{x}^*$, the functional $f : K \mapsto \mathbb{R}$ defined by $f : \mathbf{x} \mapsto -\sum_{i=1}^r \log \rho_i(\mathbf{l}_\mathbf{x})^2$ is well defined. Furthermore, it is an r-logarithmically homogeneous, self-concordant barrier for K (see [2]). In fact, we know from a result of Güler and Tunçel [9] that it is optimal for K. We shall now relate this barrier with a weighted barrier of the SDP representation given by (4.1).

For each i, let $J(i)$ denote the set of the indices of the $\bar{n}_i := \sum_{j=1}^i n_{ij}$ copies of $\rho_i(\mathbf{l}_\mathbf{x})$ on the diagonal of $\mathbf{L} = \mathbf{M}(\mathbf{l})$, i.e., $\mathbf{L}_{jj} = \rho_i(\mathbf{l}_\mathbf{x})$ for all $i \in \{1, \ldots, r\}$, all $j \in J(i)$, and all $\mathbf{x} \in K$. Since $\{J(i)\}_{i=1}^r$ is a partition of $\{1, \ldots, n\}$, where $n := \sum_{1 \le j \le i \le r} n_{ij}$, we may define a map $\pi : \{1, \ldots, n\} \to \{1, \ldots, r\}$ such that $j \in J(\pi(j))$ for all $j \in \{1, \ldots, n\}$. For each $i \in \{1, \ldots, r\}$,

$$(4.3) \qquad \log \rho_i(\mathbf{l}_\mathbf{x})^2 = \frac{1}{\bar{n}_i} \bar{n}_i \log \rho_i(\mathbf{l}_\mathbf{x})^2 = \sum_{j \in J(i)} \frac{1}{\bar{n}_{\pi(j)}} \log(\mathbf{L}_{\mathbf{M}(\mathbf{x})})_{jj}^2,$$

from which we deduce that the optimal barrier

$$f(\mathbf{x}) = -\sum_{i=1}^r \log \rho_i(\mathbf{l}_\mathbf{x})^2 = -\sum_{i=1}^n \bar{n}_{\pi(i)}^{-1} \log(\mathbf{L}_{\mathbf{M}(\mathbf{x})})_{ii}^2$$

coincides with the restriction of the weighted barrier for the SDP representation with weights $\bar{n}_{\pi(1)}^{-1}, \ldots, \bar{n}_{\pi(n)}^{-1}$. Consequently, as a corollary to Theorem 14, we have the following.

COROLLARY 19. *If a pair of primal-dual HCP problems has strictly complementary solutions, then the central path converges to the solution of*

$$\min \quad -\sum_{i \in B} \log \rho_i(\mathbf{l}_\mathbf{x})^2$$
$$s.t. \quad (\mathbf{a}^{(i)})^T \mathbf{x} = \mathbf{b}_i, \quad i = 1, \ldots, m,$$
$$\mathbf{x} \in \mathrm{span}(F_p),$$

where F_p is the minimal face of K containing the primal optimal face and $B = \{i : \exists \mathbf{x} \in F_p, \rho_i(\mathbf{l}_\mathbf{x}) \ne 0\}$.

Proof. Since the image of the central path under $\mathbf{M}$ is the path defined by the weighted barrier $\mathbf{X} \mapsto -\sum_{i=1}^n \log \bar{n}_{\pi(i)}^{-1}(\mathbf{L}_\mathbf{X})^2$ for the SDP representation (4.1), it follows from Theorem 14 that when the HCP problem has strictly complementary solutions, the image of the central path under $\mathbf{M}$ converges to the solution of

$$\min \quad -\sum_{i \in J(B)} \bar{n}_{\pi(i)}^{-1} \log(\mathbf{L}_\mathbf{X})_{ii}^2$$
$$s.t. \quad \mathbf{M}^{-*}(\mathbf{a}^{(i)}) \bullet \mathbf{X} = \mathbf{b}_i, \quad i = 1, \ldots, m,$$
$$\mathbf{X} \in \mathbf{M}(\mathbb{R}^d),$$
$$\mathbf{X} \in \mathrm{span}(F_p'),$$

where $J(B)$ denotes $\cup_{i \in B} J(i)$ and F_p' is the face of $\mathbb{S}_{++}^n$ such that $F_p' \cap \mathbf{M}(\mathbb{R}^d) = F_p$. The theorem then follows from (4.3). □

5. Conclusion. We end this paper with some open questions and directions for future research.

1. Since the notion of weighted centers introduced in this paper possesses both uniqueness and completeness, we may use them for future development of V-space algorithms for SDP. One approach is to consider the target map $(\mathbf{X}, \mathbf{S}) \mapsto \mathbf{Q}\mathbf{D}\mathbf{Q}^T$, where $\mathbf{Q}^T\mathbf{X}\mathbf{S}\mathbf{Q} = \mathbf{L}$ is a Schur decomposition of $\mathbf{X}\mathbf{S}$ with diagonal entries of $\mathbf{L}$ arranged in nonincreasing order and $\mathbf{D}$ is a diagonal matrix that shares the same diagonal entries with $\mathbf{L}$, which is a bijection between the primal-dual strictly feasible region and the cone $\mathbb{S}^n_{++}$. Another approach would be to linearly transform the primal-dual problems via the orthonormal similarity transformation $\mathcal{Q} : \mathbf{X} \mapsto \mathbf{Q}^T\mathbf{X}\mathbf{Q}$ so that $\mathbf{L}^T_{\mathcal{Q}(\mathbf{X})}\mathcal{Q}(\mathbf{S})\mathbf{L}_{\mathcal{Q}(\mathbf{X})}$ is diagonal and use the locally injective map $(\mathbf{X}, \mathbf{S}) \mapsto \mathbf{L}^T_{\mathbf{X}}\mathbf{S}\mathbf{L}_{\mathbf{X}}$ as the V-space map.

2. The limit points of weighted centers for SDP were characterized in this paper only under strict complementarity. In the absence of strict complementarity, the limit point of the usual central path can be characterized either as the analytic center of a certain subset of the optimal face (see [6]) or as the unique minimizer of the logarithmic determinant barrier for the optimal face with an additional term (see [7]). Future extensions of these results to weighted central paths would complete the characterization of their limit points.

3. By treating central paths for HCP problems as weighted central paths for the SDP reformulations, any V-space algorithm that follows weighted central paths naturally translates to a primal-dual algorithm that follows central paths of HCP problems. However, without exploiting the structure of homogeneous cones in the analysis of the algorithm, its theoretical complexity will generally be no better than algorithms that follow the usual central path of the SDP reformulation. Thus, some nontrivial work is needed to improve the analysis of these V-space algorithms for HCP.

4. In [12, 13, 20], the analyticity of various notions of weighted central paths were studied. In [4], we investigate the analyticity of the weighted central paths introduced in this paper.

Appendix A. Technical results.

LEMMA 20. *If* $\mathbf{L}$ *is a real, lower triangular, diagonalizable matrix with nonincreasing diagonal entries, and*

$$\mathbf{P}^{-1}\mathbf{L}\mathbf{P} = \mathbf{D}$$

is a diagonalization of $\mathbf{L}$ *where* $\mathbf{D}$ *has nonincreasing diagonal entries, then* $\mathbf{P}$ *is lower block-diagonal where the size of the kth block on the diagonal is the multiplicity of the kth largest diagonal entry of* $\mathbf{L}$.

Proof. We shall prove by induction on the number of distinct diagonal entries of $\mathbf{L}$, which is the number of distinct eigenvalues of $\mathbf{L}$.

When $\mathbf{L}$ has only one distinct eigenvalue, the lemma is trivial.

Suppose that the lemma is true whenever $\mathbf{L}$ has at most p distinct eigenvalues. Consider the case where $\mathbf{L}$ has $p + 1$ distinct eigenvalues. Let m denote the multiplicity of its largest eigenvalue $\lambda_{\max}$. We write all matrices in the 2-by-2 block form $\mathbf{M} = \left[\begin{smallmatrix}\mathbf{M}_{11} & \mathbf{M}_{12} \\ \mathbf{M}_{21} & \mathbf{M}_{22}\end{smallmatrix}\right]$, where $\mathbf{M}_{11}$ is m-by-m. The diagonals of $\mathbf{L}$ and $\mathbf{D}$ (which contain eigenvalues of the similar matrices $\mathbf{L}$ and $\mathbf{D}$) coincide since they are arranged in nonincreasing order. Now

$$\begin{bmatrix} \mathbf{L}_{11} & \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{12} \\ \mathbf{P}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_{12} \\ \mathbf{P}_{22} \end{bmatrix} \mathbf{D}_{22} \implies \mathbf{L}_{11}\mathbf{P}_{12} = \mathbf{P}_{12}\mathbf{D}_{22}$$

implies that each nonzero column of $\mathbf{P}_{12}$ is an eigenvector of $\mathbf{L}_{11}$ whose associated eigenvalue is a diagonal entry of $\mathbf{D}_{22}$. Since the only eigenvalue of $\mathbf{L}_{11}$ is $\lambda_{\max}$ and all diagonal entries of $\mathbf{D}_{22}$ are strictly less than $\lambda_{\max}$, it follows that $\mathbf{P}_{12}$ is a zero matrix. Consequently, it follows from

$$\begin{bmatrix} \mathbf{P}_{11} & \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{L}_{11} & \\ \mathbf{L}_{21} & \mathbf{L}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{11} & \\ \mathbf{P}_{21} & \mathbf{P}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{11} & \\ & \mathbf{D}_{22} \end{bmatrix}$$

that $\mathbf{L}_{22} = \mathbf{P}_{22}\mathbf{D}_{22}\mathbf{P}_{22}^{-1}$ is diagonalizable. Since $\mathbf{L}_{22}$ has p distinct eigenvalues, we may apply the induction hypothesis to conclude that $\mathbf{P}_{22}$ is lower block-diagonal where the size of the kth block on the diagonal is the multiplicity of the kth largest diagonal entry of $\mathbf{L}_{22}$.   ⬚

PROPOSITION 21. *Suppose* $\mathbf{L}$ *is a real, lower triangular, diagonalizable matrix with nonincreasing diagonal entries and* $\mathbf{Q}$ *is an orthogonal matrix of the same size as* $\mathbf{L}$. *Then* $\mathbf{Q}^T\mathbf{L}\mathbf{Q}$ *is a lower triangular matrix with nonincreasing diagonal entries if and only if* $\mathbf{Q}$ *is block-diagonal where the size of the kth block is the multiplicity of the kth largest diagonal entry of* $\mathbf{L}$.

Proof. "only if": Suppose $\mathbf{Q}^T\mathbf{L}\mathbf{Q}$ is lower triangular with nonincreasing diagonal entries. Let $\mathbf{P}^{-1}(\mathbf{Q}^T\mathbf{L}\mathbf{Q})\mathbf{P} = \mathbf{D}$ be a diagonalization of $\mathbf{Q}^T\mathbf{L}\mathbf{Q}$ where $\mathbf{D}$ has nonincreasing diagonal entries, and hence shares the same diagonal as $\mathbf{L}$. It follows from the preceding lemma that $\mathbf{P}$ is lower block-diagonal where the size of the kth block is the multiplicity of the kth largest diagonal entry of $\mathbf{L}$. Since

$$(\mathbf{Q}\mathbf{P})^{-1}\mathbf{L}(\mathbf{Q}\mathbf{P}) = \mathbf{D}$$

is a diagonalization of $\mathbf{L}$, we conclude, using the preceding lemma, that $\mathbf{Q}\mathbf{P}$, whence $\mathbf{Q}$, is lower block-diagonal where the size of the kth block is the multiplicity of the kth largest diagonal entry of $\mathbf{L}$.

"if": Suppose $\mathbf{Q}$ is block-diagonal where the size of the kth block is the multiplicity of the kth largest diagonal entry of $\mathbf{L}$. Let $\mathbf{P}^{-1}\mathbf{L}\mathbf{P} = \mathbf{D}$ be a diagonalization of $\mathbf{L}$ where $\mathbf{D}$ has nonincreasing diagonal entries, and hence shares the same diagonal as $\mathbf{L}$. By the preceding lemma, $\mathbf{P}$, whence $\mathbf{Q}^T\mathbf{P}$, is lower block-diagonal. Thus $\mathbf{Q}^T\mathbf{L}\mathbf{Q} = (\mathbf{Q}^T\mathbf{P})\mathbf{D}(\mathbf{Q}^T\mathbf{P})^{-1}$ is lower block-diagonal. Since the kth diagonal block of $\mathbf{D}$ is a multiple of the identity matrix, so is the kth diagonal block of $(\mathbf{Q}^T\mathbf{P})\mathbf{D}(\mathbf{Q}^T\mathbf{P})^{-1}$. Consequently $\mathbf{Q}^T\mathbf{L}\mathbf{Q} = (\mathbf{Q}^T\mathbf{P})\mathbf{D}(\mathbf{Q}^T\mathbf{P})^{-1}$ is actually lower triangular and shares the same diagonal as $\mathbf{D}$.   ⬚

REFERENCES

[1] G. P. BARKER AND D. CARLSON, *Cones of diagonally dominant matrices*, Pacific J. Math., 57 (1975), pp. 15–32.
[2] C. B. CHUA, *An Algebraic Perspective on Homogeneous Cone Programming, and the Primal-Dual Second Order Cone Approximations Algorithm for Symmetric Cone Programming*, Ph.D. thesis, Cornell University, Ithaca, NY, 2003.
[3] C. B. CHUA, *Relating homogeneous cones and positive definite cones via T-algebras*, SIAM J. Optim., 14 (2003), pp. 500–506.
[4] C. B. CHUA, *Analyticity of Weighted Central Path and Error Bound for Semidefinite Programming*, CORR 2005-15, Department of Combinatorics and Optimization, Faculty of Mathematics, University of Waterloo, Canada, 2005.

[5]   C. B. Chua and L. Tunçel, *Invariance and Efficiency of Convex Representations*, Research
      Report CORR 2004-18, Department of Combinatorics and Optimization, Faculty of Math-
      ematics, University of Waterloo, Canada, 2004.
[6]   E. de Klerk, M. Halická, and C. Roos, *Limiting behavior of the central path in semidefinite
      optimization*, Optim. Methods Softw., 20 (2005), pp. 99–113.
[7]   A. Forsgren and G. Sporre, *Characterization of the Limit Point of the Central Path
      in Semidefinite Programming*, Report trita-mat-2002-os12, Department of Mathematics,
      Royal Institute of Technology, Stockholm, Sweden, 2002.
[8]   D. Goldfarb and K. Scheinberg, *Interior point trajectories in semidefinite programming*,
      SIAM J. Optim., 8 (1998), pp. 871–886.
[9]   O. Güler and L. Tunçel, *Characterization of the barrier parameter of homogeneous convex
      cones*, Math. Program., 81 (1998), pp. 55–76.
[10] B. Jansen, C. Roos, T. Terlaky, and J.-P. Vial, *Primal-dual target-following algorithms
      for linear programming*, Ann. Oper. Res., 62 (1996), pp. 197–231.
[11] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise, *A Unified Approach to Interior Point
      Algorithms for Linear Complementarity Problems*, Lecture Notes in Comput. Sci. 538,
      Springer-Verlag, Berlin, Heidelberg, New York, 1991.
[12] Z. Lu and R. D. C. Monteiro, *Limiting Behavior of the Alizadeh-Haeberly-Overton Weighted
      Paths in Semidefinite Programming*, working paper, School of ISyE, Georgia Tech, 2003.
[13] Z. Lu and R. D. C. Monteiro, *Error bounds and limiting behaviour of weighted paths asso-
      ciated with the SDP map $X^{1/2}SX^{1/2}$*, SIAM J. Optim., 15 (2004), pp. 348–374.
[14] Z.-Q. Luo, J. F. Sturm, and S. Zhang, *Superlinear convergence of a symmetric primal-
      dual path following algorithm for semidefinite programming*, SIAM J. Optim., 8 (1998),
      pp. 59–81.
[15] R. D. C. Monteiro and J.-S. Pang, *On two interior-point mappings for nonlinear semidefinite
      complementarity problems*, Math. Oper. Res., 23 (1998), pp. 39–60.
[16] R. D. C. Monteiro and P. Zanjácomo, *Implementation of primal-dual methods for semidef-
      inite programming based on Monteiro and Tsuchiya Newton directions and their variants*,
      Optim. Methods Softw., 11/12 (1999), pp. 91–140.
[17] R. D. C. Monteiro and P. Zanjácomo, *General interior-point maps and existence of weighted
      paths for nonlinear semidefinite complementarity problems*, Math. Oper. Res., 25 (2000),
      pp. 382–399.
[18] G. Pataki, *Cone LP's and semidefinite programs: Goemetry and a simplex type method*, in
      Proceedings of the Fifth IPCO Conference, W. H. Cunningham, S. T. McCormick, and
      M. Queyranne, eds., Springer-Verlag, New York, 1996, pp. 161–174.
[19] G. Pataki and L. Tunçel, *On the generic properties of convex optimization problems in conic
      form*, Math. Program., 89 (2001), pp. 449–457.
[20] M. Preiß and J. Stoer, *Analysis of infeasible-interior-point paths arising with semidefinite
      linear complementarity problems*, Math. Program., 99 (2004), pp. 499–520.
[21] J. F. Sturm and S. Zhang, *On weighted centers for semidefinite programming*, Eur. J. Oper.
      Res., 126 (2000), pp. 391–407.
[22] L. Tunçel and H. Wolkowicz, *Strengthened existence and uniqueness conditions for search
      directions in semidefinite programming*, Linear Algebra Appl., 400 (2005), pp. 31–60.
[23] E. B. Vinberg, *The theory of convex homogeneous cones*, Tr. Mosk. Mat. Obs., 12 (1963),
      pp. 303–358.