MATHEMATICS OF OPERATIONS RESEARCH
Vol. 25, No. 3, August 2000, pp. 365–380
Printed in U.S.A.

THE VOLUMETRIC BARRIER FOR SEMIDEFINITE PROGRAMMING

KURT M. ANSTREICHER

We consider the volumetric barrier for semidefinite programming, or "generalized" volumetric barrier, as introduced by Nesterov and Nemirovskii. We extend several fundamental properties of the volumetric barrier for a polyhedral set to the semidefinite case. Our analysis facilitates a simplified proof of self-concordance for the semidefinite volumetric barrier, as well as for the combined volumetric-logarithmic barrier for semidefinite programming. For both of these barriers we obtain self-concordance parameters equal to those previously shown to hold in the polyhedral case.

1. Introduction. This paper concerns the volumetric barrier for semidefinite programming. The volumetric barrier for a polyhedral set $\mathscr{P}=\{x \mid Ax \geq c\}$, where A is an $m\times n$ matrix, was introduced by Vaidya (1996). Vaidya used the volumetric barrier in the construction of a cutting plane algorithm for convex programming; see also Anstreicher (1997b, 1999a, 1999b). Subsequently Vaidya and Atkinson (1993) (see also Anstreicher 1997a) used a hybrid combination of the volumetric and logarithmic barriers for $\mathscr{P}$ to construct an $O(m^{1/4}n^{1/4}L)$-iteration algorithm for a linear programming problem defined over $\mathscr{P}$, with integer data of total bit size L. For $m \gg n$ this complexity compares favorably with $O(\sqrt{m}L)$, the best known iteration complexity for methods based on the logarithmic barrier.

Nesterov and Nemirovskii (1994, §5.5) proved self-concordance results for the volumetric, and combined volumetric-logarithmic, barriers that are consistent with the algorithm complexities obtained in Vaidya and Atkinson (1993). In fact Nesterov and Nemirovskii (1994) obtain results for extensions of the volumetric and combined barriers to a set of the form $\mathscr{S}=\{x \mid \sum_{i=1}^{n} x_i A_i \succeq C\}$, where A_i, $i=1,\dots,n$ and C are $m\times m$ symmetric matrices, and $\succeq$ denotes the semidefinite ordering. The set $\mathscr{S}$ is a strict generalization of $\mathscr{P}$, since $\mathscr{P}$ can be represented by using diagonal matrices in the definition of $\mathscr{S}$. Optimization over a set of the form $\mathscr{S}$ is now usually referred to as *semidefinite programming*; see for example Alizadeh (1995) or Vandenberghe and Boyd (1996). It is well known (see Nesterov and Nemirovskii 1994) that an extension of the logarithmic barrier to $\mathscr{S}$ obtains an m-self-concordant barrier. In Nesterov and Nemirovskii (1994) it is also shown that semidefinite extensions of the volumetric, and combined volumetric-logarithmic barrier are $O(\sqrt{mn})$, and $O(\sqrt{mn})$, self-concordant barriers for $\mathscr{S}$, respectively.

The self-concordance proofs in Nesterov and Nemirovskii (1994, §5.5) are extremely technical, and do not obtain the constants that would be needed to actually implement algorithms using the barriers. Simplified proofs of self-concordance for the volumetric and combined barriers for $\mathscr{P}$ are obtained in Anstreicher (1997a). In particular, it is shown these barriers are $225\sqrt{mn}$, and $450\sqrt{mn}$ self-concordant barriers for $\mathscr{P}$, respectively.

Received April 5, 1998; revised December 30, 1999.
AMS 1991 subject classification. Primary: 90C30.
OR/MS subject classification. Primary: Programming/semidefinite, volumetric barrier, semidefinite programming, self-concordance.
Key words. Volumetric barrier, semidefinite programming, self-concordance.

0364-765X/00/2503/365/$05.00
1526-5471 electronic ISSN, © 2000, INFORMS

The proofs of these self-concordance results use a number of fundamental properties of the volumetric barrier established in Anstreicher (1996, 1997a). Unfortunately, however, the analysis of Anstreicher (1997a) does not apply to the more general semidefinite constraint defining $\mathcal{S}$, as considered in Nesterov and Nemirovskii (1994). With the current activity in semidefinite programming the extension of results for the volumetric and combined barriers to $\mathcal{S}$ is of some interest. For example, in Nesterov and Nemirovskii (1994, p. 204) it is argued that with a large number of low-rank quadratic constraints, the combined volumetric-logarithmic barrier applied to a semidefinite formulation obtains a lower complexity than the usual approach of applying the logarithmic barrier directly to the quadratic constraints.

The purpose of this paper is to extend the analysis of the volumetric and combined barriers in Anstreicher (1996, 1997a) to the semidefinite case. This analysis is by necessity somewhat complex, but in the end we obtain semidefinite generalizations for virtually all of the fundamental results in Anstreicher (1996, 1997a). These include:

• The semidefinite generalization of the matrix $Q(x)$ having $Q(x) \preceq \nabla^2 V(x) \preceq 3Q(x)$, where $V(\cdot)$ is the volumetric barrier.

• The semidefinite generalization of the matrix Σ, which in the polyhedral case is the diagonal matrix $\Sigma = \mathrm{Diag}(\sigma)$. Representations of $\nabla V(x)$ and $Q(x)$ in terms of Σ clearly show the relationship with the polyhedral case (see Table 1, at the end of §4).

• Semidefinite generalizations of fundamental inequalities between $Q(x)$ and the Hessian of the logarithmic barrier (see Theorems 4.2 and 4.3).

• Self-concordance results for the volumetric, and combined, barriers identical to those obtained for the polyhedral case. In particular, we prove that these barriers are $225\sqrt{mn}$, and $450\sqrt{mn}$ self-concordant barriers for $\mathcal{S}$, respectively.

The fact that we obtain self-concordance results identical to those previously shown to hold in the polyhedral case is somewhat surprising, because one important element in the analysis here is significantly different than in Anstreicher (1997a). In Anstreicher (1997a), self-concordance is established by proving a relative Lipschitz condition on the Hessian $\nabla^2 V(\cdot)$. This proof is based on Shur product inequalities, and an application of the Gershgorin circle theorem. The use of the Lipschitz condition is attractive because it eliminates the need to explicitly consider the third directional derivatives of the volumetric barrier. We have been unable to extend this proof technique to the semidefinite case, however, and consequently here we explicitly consider the third directional derivatives of $V(\cdot)$. The proof of the main result concerning these third derivatives (Theorem 5.1) is based on properties of Kronecker products. Despite the fact that on this point the analytical techniques used here and in Anstreicher (1997a) are quite different, the final self-concordance results are identical.

An outline of the paper follows. In the next section we briefly consider some mathematical preliminaries. The most significant of these are well-known properties of the Kronecker product, which we use extensively throughout the paper. In §3 we define the logarithmic, volumetric, and combined barriers for $\mathcal{S}$, and state the main self-concordance theorems. The proofs of these results are deferred until §5. Section 4 considers a detailed analysis of the volumetric barrier for $\mathcal{S}$. We first obtain Kronecker product representations for the gradient and Hessian of $V(\cdot)$, which are then used to prove a variety of results generalizing those in Anstreicher (1996, 1997a). Later in the section the matrix Σ is defined, and alternative representations of $\nabla V(x)$ and $Q(x)$ in terms of Σ are obtained (see Table 1). Section 5 considers the proofs of self-concordance for the volumetric and combined barriers. The main work here is to obtain Kronecker product representations for the third directional derivatives of $V(\cdot)$, and then prove a result (Theorem 5.1) relating the third derivatives to $Q(x)$.

2. Preliminaries. In this section we briefly consider several points of linear algebra and matrix calculus that will be required in the sequel. To begin, let A and B be $m \times m$ matrices. We use $\operatorname{tr}(A)$ to denote the trace of A, and $A \cdot B$ to denote the matrix inner product

$$A \cdot B = \sum_{i,j} a_{ij} b_{ij} = \operatorname{tr}(AB^T).$$

Let $\sigma_A \in \Re^m$ denote the vector of singular values of A, that is, the positive square roots of the eigenvalues of $A^T A$. The *Frobenius norm* of A is then $\|A\| = (\operatorname{tr}(A^T A))^{1/2} = \|\sigma_A\|$, and the *spectral norm* is $|A| = \|\sigma_A\|_\infty$. We say that a matrix A is *positive semidefinite* (psd) if A is symmetric, and has all nonnegative eigenvalues. We use $\succeq$ to denote the semidefinite ordering for symmetric matrices: $A \succeq B$ if $A - B$ is psd. For a vector $v \in \Re^n$, $\operatorname{Diag}(v)$ is the $n \times n$ diagonal matrix with $\operatorname{Diag}(v)_{ii} = v_i$ for each i. We will make frequent use of the following elementary properties of $\operatorname{tr}(\cdot)$. Parts (1) and (3) of the following proposition are well known, and parts (2) and (4) follow easily from (1) and (3), respectively.

PROPOSITION 2.1. *Let A and B be $m \times m$ matrices. Then*
(1) $tr(AB) = tr(BA)$;
(2) *If A is symmetric, then* $tr(AB) = tr(AB^T)$;
(3) *If A and B are psd, then $A \cdot B \geq 0$, and $A \cdot B = 0$ if and only if $AB = 0$.*
(4) *If $A \succeq 0$ and $B \succeq C$, then $A \cdot B \geq A \cdot C$.*

Let A and B be $m \times n$, and $k \times l$, matrices, respectively. The *Kronecker product* of A and B, denoted $A \otimes B$, is the $mk \times nl$ block matrix whose i,j block is $a_{ij}B$, $i = 1, \ldots, m$, $j = 1, \ldots, n$. For our purposes it is also very convenient to define a "symmetrized" Kronecker product:

$$A \otimes_S B = \frac{1}{2}(A \otimes B + B \otimes A).$$

For an $m \times n$ matrix A, $\mathbf{vec}(A) \in \Re^{mn}$ is the vector formed by "stacking" the columns of A one atop another, in the natural order. The following properties of the Kronecker product are all well known; see for example Horn and Johnson (1991), except for (2), which follows immediately from (1) and the definition of $\otimes_S$.

PROPOSITION 2.2. *Let A, B, C, and D be conforming matrices. Then*
(1) $(A \otimes B)(C \otimes D) = AC \otimes BD$;
(2) $(A \otimes_S B)(C \otimes_S D) = \frac{1}{2}(AC \otimes_S BD + AD \otimes_S BC)$;
(3) $(A \otimes B)^T = A^T \otimes B^T$;
(4) *If A and B are nonsingular, then $A \otimes B$ is nonsingular, and $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$;*
(5) $\mathbf{vec}(ABC) = (C^T \otimes A)\mathbf{vec}(B)$;
(6) *If A and B are psd, then $A \otimes B$ is psd.*

Lastly we consider two simple matrix calculus results. Let X be a nonsingular matrix with $\det(X) > 0$. Then (see for example Graham 1981, p. 75),

$$\frac{\partial}{\partial x_{ij}} \ln \det(X) = [X^{-1}]_{ji}, \tag{1}$$

and also (see for example Graham 1981, p. 64),

$$\frac{\partial}{\partial x_{ij}} X^{-1} = -X^{-1} e_i e_j^T X^{-1}, \tag{2}$$

where e_i denotes the ith elementary vector.

3. Main results. Let G be a closed convex subset of $\Re^n$, and let $F(\cdot)$ be a C^3, convex mapping from $\mathrm{Int}(G)$ to $\Re$, where $\mathrm{Int}(\cdot)$ denotes interior. Then (Nesterov and Nemirovskii 1994) $F(\cdot)$ is called a *ϑ-self-concordant* barrier for G if $F(\cdot)$ tends to infinity for any sequence approaching a boundary point of G,

$$|D^3F(x)[\xi,\xi,\xi]| \le 2(D^2F(x)[\xi,\xi])^{3/2}$$

for every $x \in \mathrm{Int}(G)$ and $\xi \in \Re^n$, and

$$\sup_{x\in \mathrm{Int}(G)} \{\nabla F(x)\nabla^2 F(x)^{-1}\nabla F(x)^T\} \le \vartheta.$$

As shown by Nesterov and Nemirovskii (1994, Theorem 3.2.1), the existence of a ϑ-self-concordant barrier for G implies that a linear, or convex quadratic, objective can be minimized on G to within a tolerance ε of optimality using $O(\sqrt{\vartheta}|\ln\varepsilon|)$ iterations of Newton's method.

Consider a set $\mathscr{S} \subset \Re^n$ of the form

$$\mathscr{S} = \left\{ x \mid S(x) = \sum_{i=1}^n x_i A_i - C \succeq 0 \right\},$$

where A_i, $i = 1,\dots,n$ and C are $m\times m$ symmetric matrices. We assume throughout that the matrices $\{A_i\}$ are linearly independent, and that a point x with $S(x)\succ 0$ exists. It is then easy to show that $\mathrm{Int}(\mathscr{S}) = \{x \mid S(x)\succ 0\}$. The *logarithmic barrier* for $\mathscr{S}$ is the function

$$f(x) = -\ln\det(S(x)),$$

defined on the interior of $\mathscr{S}$. As shown by Nesterov and Nemirovskii (1994, Proposition 5.4.5). $f(\cdot)$ is an m-self-concordant barrier for $\mathscr{S}$, implying the existence of polynomial-time interior-point algorithms for linear, and convex quadratic, semidefinite programming.

The *volumetric barrier* $V(\cdot)$ for $\mathscr{S}$, as defined in Nesterov and Nemirovskii (1994, §5.5), is the function

$$V(x) = \frac{1}{2}\ln\det(\nabla^2 f(x)).$$

The first main result of the paper is the following improved characterization of the self-concordance of $V(\cdot)$.

THEOREM 3.1. *Let $\mathscr{S} = \{x \mid S(x)\succeq 0\}$, where $S(x) = \sum_{i=1}^n x_iA_i - C$, and each A_i, and C, are $m\times m$ symmetric matrices. Then $225\, m^{1/2}V(\cdot)$ is a ϑ-self-concordant barrier for $\mathscr{S}$, for $\vartheta = 225\, m^{1/2} n$.*

Theorem 3.1 generalizes a result for the polyhedral volumetric barrier (Anstreicher 1997a, Theorem 5.1), and provides an alternative to the semidefinite self-concordance result of Nesterov and Nemirovskii (1994, Theorem 5.5.1). It is worthwhile to note that in fact the analysis in Nesterov and Nemirovskii (1994, §5.5) does not apply directly to the barrier $V(\cdot)$ for $\mathscr{S}$ as given here, because Nesterov and Nemirovskii assume that the "right-hand side" matrix C is zero. In practice, this assumption can be satisfied by extending $\mathscr{S}$ to the cone

$$\mathscr{K} = \left\{ (x_0, x) \mid \sum_{i=1}^n x_iA_i - x_0C \succeq 0 \right\},$$

and then intersecting $\mathscr{K}$ with the linear constraint $x_0 = 1$ to recover $\mathscr{S}$. The analysis in Nesterov and Nemirovskii (1994) would then be applied to the volumetric barrier $\hat{V}(\cdot)$ for $\mathscr{K}$. The advantage of working with $\mathscr{K}$ is that some general results of Nesterov and Nemirovskii can then be applied, because $\hat{V}(\cdot)$ is $(n+1)$-logarithmically homogenous; see Nesterov and Nemirovskii (1994, §2.3.3). (For example Theorem 4.4, required in the analysis of §5, could be replaced by the fact that $\nabla\hat{V}(x)\nabla^2\hat{V}(x)^{-1}\nabla\hat{V}(x)^T = n+1$, from Nesterov and Nemirovskii 1994, Proposition 2.3.4.) Our analysis shows, however, that the homogeneity assumption used in Nesterov and Nemirovskii (1994) is not needed to prove self-concordance for the semidefinite volumetric barrier.

The *combined volumetric-logarithmic barrier* for $\mathscr{S}$ is the function

$$V_\rho(x) = V(x) + \rho f(x),$$

where $V(\cdot)$ is the volumetric barrier, $f(\cdot)$ is the logarithmic barrier, and ρ is a positive scalar. The combined barrier was introduced for polyhedral sets in Vaidya and Atkinson (1993), and extended to semidefinite constraints in Nesterov and Nemirovskii (1994). Our main result on the self-concordance of $V_\rho(\cdot)$ is the following:

THEOREM 3.2. *Let $\mathscr{S} = \{x \mid S(x) \succeq 0\}$, where $S(x) = \sum_{i=1}^n x_iA_i - C$, and each A_i, and C, are $m\times m$ symmetric matrices. Assume that $n<m$, and let $\rho = (n-1)/(m-1)$. Then $225(m/n)^{1/2}V_\rho(\cdot)$ is a ϑ-self-concordant barrier for $\mathscr{S}$, for $\vartheta = 450m^{1/2}n^{1/2}$.*

Theorems 3.1 and 3.2 imply that if $m \gg n$, then the self-concordance parameter ϑ for the volumetric or combined barrier for $\mathscr{S}$ (particularly the latter) can be lower than m, the parameter for the logarithmic barrier. It follows that for $m \gg n$ the complexity of interior-point algorithms for the minimization of a linear, or convex quadratic, function over $\mathscr{S}$ may be improved by utilizing $V(\cdot)$ or $V_\rho(\cdot)$ in place of $f(\cdot)$.

4. The volumetric barrier. Let $f(\cdot)$ be the logarithmic barrier for $\mathscr{S}$, as defined in the previous section. It is well known (see for example Vandenberghe and Boyd 1996) that the first and second partial derivatives of $f(\cdot)$ at an interior point of $\mathscr{S}$ are given by:

$$\frac{\partial f(x)}{\partial x_i} = -\mathrm{tr}(S^{-1}A_i), \qquad \frac{\partial^2 f(x)}{\partial x_i\partial x_j} = \mathrm{tr}(S^{-1}A_iS^{-1}A_j),$$

where throughout we use $S = S(x)$ whenever possible to reduce notation. Let $\mathscr{A}$ be the $m^2\times n$ matrix whose ith column is $\mathbf{vec}(A_i)$. Since

$$\frac{\partial^2 f(x)}{\partial x_i\partial x_j} = \mathbf{vec}(A_i)^T\mathbf{vec}(S^{-1}A_jS^{-1}) = \mathbf{vec}(A_i)^T[S^{-1}\otimes S^{-1}]\mathbf{vec}(A_j),$$

where the second equality uses Proposition 2.2(5), the Hessian matrix $H = H(x) = \nabla^2 f(x)$ can be represented in the form (see Alizadeh 1995)

$$H = \mathscr{A}^T[S^{-1}\otimes S^{-1}]\mathscr{A}. \tag{3}$$

Note that $H = H(x)$ is positive definite under the assumptions that $S = S(x)\succ 0$, and that the matrices $\{A_i\}$ are linearly independent.

Our first goal in this section is to obtain Kronecker product representations for the gradient and Hessian of $V(\cdot)$. To start, it is helpful to compute

$$\frac{\partial S(x)^{-1}}{\partial x_k} = \sum_{i,j}\frac{\partial S(x)^{-1}}{\partial[S(x)]_{ij}}\frac{\partial[S(x)]_{ij}}{\partial x_k} = -\sum_{i,j}S^{-1}e_ie_j^TS^{-1}[A_k]_{ij} = -S^{-1}A_kS^{-1}, \tag{4}$$

where the second equality uses (2). In addition, using (4) and the definitions of $\otimes$ and $\otimes_S$, it is easy to see that

$$\frac{\partial}{\partial x_k}(S(x)^{-1} \otimes S(x)^{-1}) = -2S^{-1} \otimes_S S^{-1}A_kS^{-1}. \tag{5}$$

Now applying the chain rule, (1), and (5), we find that

$$\begin{aligned}
\frac{\partial V(x)}{\partial x_i} &= \frac{1}{2}H^{-1}(x) \cdot \frac{\partial H(x)}{\partial x_i} && (6)\\
&= \frac{1}{2}H^{-1}(x) \cdot \mathscr{A}^T \left[\frac{\partial}{\partial x_i}(S(x)^{-1} \otimes S(x)^{-1})\right] \mathscr{A}\\
&= -H^{-1} \cdot \mathscr{A}^T[S^{-1} \otimes_S S^{-1}A_iS^{-1}]\mathscr{A}\\
&= -(\mathscr{A}H^{-1}\mathscr{A}^T) \cdot (S^{-1}A_iS^{-1} \otimes_S S^{-1})\\
&= -P \cdot (S^{-1/2}A_iS^{-1/2} \otimes_S I), && (7)
\end{aligned}$$

where the last equality uses Proposition 2.2(1), $S^{-1/2}$ is the unique positive definite matrix having $(S^{-1/2})^2 = S^{-1}$, and

$$P = P(S) = [S^{-1/2} \otimes S^{-1/2}]\mathscr{A}(\mathscr{A}^T[S^{-1} \otimes S^{-1}]\mathscr{A})^{-1}\mathscr{A}^T[S^{-1/2} \otimes S^{-1/2}] \tag{8}$$

is the orthogonal projection onto the range of $[S^{-1/2} \otimes S^{-1/2}]\mathscr{A}$. Note that the jth column of $[S^{-1/2} \otimes S^{-1/2}]\mathscr{A}$ is exactly

$$[S^{-1/2} \otimes S^{-1/2}]\mathbf{vec}(A_j) = \mathbf{vec}(S^{-1/2}A_jS^{-1/2}),$$

using Proposition 2.2(5). It follows that P is a representation, as an $m^2 \times m^2$ matrix, of the projection onto the subspace of $R^{m\times m}$ spanned by $\{S^{-1/2}A_jS^{-1/2},\ j = 1, \ldots, n\}$.

We will next compute the second partial derivatives of $V(\cdot)$. To start, using (2) and (5), we obtain

$$\begin{aligned}
\frac{\partial H(x)^{-1}}{\partial x_k} &= \sum_{i,j} \frac{\partial H(x)^{-1}}{\partial [H(x)]_{ij}} \frac{\partial [H(x)]_{ij}}{\partial x_k}\\
&= 2\sum_{i,j}(H^{-1}e_ie_j^TH^{-1})\mathbf{vec}(A_i)^T[S^{-1} \otimes_S S^{-1}A_kS^{-1}]\mathbf{vec}(A_j)\\
&= 2H^{-1}\mathscr{A}^T[S^{-1} \otimes_S S^{-1}A_kS^{-1}]\mathscr{A}H^{-1}. && (9)
\end{aligned}$$

Also, using (4) and the definition of $\otimes_S$, we have

$$\begin{aligned}
&\frac{\partial}{\partial x_j}(S(x)^{-1}A_iS(x)^{-1} \otimes_S S(x)^{-1})\\
&\quad = \left(\frac{\partial}{\partial x_j}S(x)^{-1}A_iS(x)^{-1}\right) \otimes_S S(x)^{-1} + S(x)^{-1}A_iS(x)^{-1} \otimes_S \left(\frac{\partial}{\partial x_j}S(x)^{-1}\right)\\
&\quad = -(S^{-1}A_jS^{-1}A_iS^{-1} + S^{-1}A_iS^{-1}A_jS^{-1}) \otimes_S S^{-1} - S^{-1}A_iS^{-1} \otimes_S S^{-1}A_jS^{-1}. && (10)
\end{aligned}$$

Combining (6), (9), and (10), and using Proposition 2.1(2), we obtain

$$\nabla^2 V(x) = 2Q(x) + R(x) - 2T(x), \tag{11}$$

where $Q = Q(x)$, $R = R(x)$, and $T = T(x)$ are the $n \times n$ matrices having

$$
\begin{aligned}
(12)\qquad Q_{ij} &= \mathscr{A}H^{-1}\mathscr{A}^T \cdot (S^{-1}A_iS^{-1}A_jS^{-1} \otimes_S S^{-1}),\\
R_{ij} &= \mathscr{A}H^{-1}\mathscr{A}^T \cdot (S^{-1}A_iS^{-1} \otimes_S S^{-1}A_jS^{-1}),\\
T_{ij} &= \mathscr{A}H^{-1}\mathscr{A}^T \cdot (S^{-1}A_iS^{-1} \otimes_S S^{-1})\mathscr{A}H^{-1}\mathscr{A}^T(S^{-1}A_jS^{-1} \otimes_S S^{-1}).
\end{aligned}
$$

THEOREM 4.1. *For any x having $S(x) \succ 0$, $0 \preceq Q(x) \preceq \nabla^2 V(x) \preceq 3Q(x)$.*

PROOF. Let $\xi \in \Re^n$, $\xi \neq 0$. Then

$$
(13)\qquad \xi^T Q \xi = \sum_{i,j} Q_{ij}\xi_i\xi_j = \mathscr{A}H^{-1}\mathscr{A}^T \cdot (S^{-1}BS^{-1}BS^{-1} \otimes_S S^{-1}) = P \cdot (\bar{B}^2 \otimes_S I),
$$

where $B = B(\xi) = \sum_{i=1}^n \xi_i A_i$, and $\bar{B} = S^{-1/2}BS^{-1/2}$. Similarly

$$
(14)\qquad \xi^T R \xi = \mathscr{A}H^{-1}\mathscr{A}^T \cdot (S^{-1}BS^{-1} \otimes S^{-1}BS^{-1}) = P \cdot (\bar{B} \otimes \bar{B}),
$$

$$
\begin{aligned}
(15)\qquad \xi^T T \xi &= \mathscr{A}H^{-1}\mathscr{A}^T \cdot (S^{-1}BS^{-1} \otimes_S S^{-1})\mathscr{A}H^{-1}\mathscr{A}^T(S^{-1}BS^{-1} \otimes_S S^{-1})\\
(16)\qquad &= P \cdot (\bar{B} \otimes_S I)P(\bar{B} \otimes_S I).
\end{aligned}
$$

Note that from Proposition 2.1(3) and Proposition 2.2(6) we immediately have $\xi^T Q \xi \geq 0$ and $\xi^T T \xi \geq 0$, since I, P and $\bar{B}^2$ are all psd. Since ξ is arbitrary, it follows that $Q \succeq 0$, $T \succeq 0$. In addition, the fact that P is a projection implies that

$$
(17)\qquad (\bar{B} \otimes_S I)P(\bar{B} \otimes_S I) \preceq (\bar{B} \otimes_S I)(\bar{B} \otimes_S I) = \frac{1}{2}([\bar{B}^2 \otimes_S I] + [\bar{B} \otimes \bar{B}]),
$$

where the last equality uses Proposition 2.2(2). Applying Proposition 2.1(4), we conclude that

$$
P \cdot (\bar{B} \otimes_S I)P(\bar{B} \otimes_S I) \leq \frac{1}{2}P \cdot ([\bar{B}^2 \otimes_S I] + [\bar{B} \otimes \bar{B}]),
$$

which is exactly $\xi^T T \xi \leq (1/2)\xi^T(Q + R)\xi$. Since ξ is arbitrary, we have shown that $T \preceq (1/2)(Q + R)$, which together with (11), $Q \succeq 0$, and $T \succeq 0$ implies that

$$
0 \preceq Q(x) \preceq \nabla^2 V(x) \preceq 2Q(x) + R(x).
$$

To complete the proof we must show that $R(x) \preceq Q(x)$. Let v_i, $i = i, \ldots, m$ be orthonormal eigenvectors of $\bar{B}$, with corresponding eigenvalues λ_i, $i = 1, \ldots, m$. Then (see Horn and Johnson 1991, Theorem 4.4.5) $\bar{B}^2 \otimes_S I$ has orthonormal eigenvectors $v_i \otimes v_j$, $i, j = 1, \ldots, m$, with corresponding eigenvalues $(1/2)(\lambda_i^2 + \lambda_j^2)$, while (see Horn and Johnson 1991, Theorem 4.2.12) $\bar{B} \otimes \bar{B}$ has the same eigenvectors $v_i \otimes v_j$, with corresponding eigenvalues $\lambda_i\lambda_j$. It then follows from $(\lambda_i - \lambda_j)^2 \geq 0$ for each i, j that

$$
(18)\qquad \bar{B}^2 \otimes_S I \succeq \bar{B} \otimes \bar{B},
$$

and Proposition 2.1(4) then implies that $P \cdot (\bar{B}^2 \otimes_S I) \geq P \cdot (\bar{B} \otimes \bar{B})$, which is exactly $\xi^T Q \xi \geq \xi^T R \xi$. Since ξ is arbitrary we have shown that $Q \succeq R$, as required. □

Theorem 4.1 generalizes a similar result (Anstreicher 1996, Theorem A.4) for the polyhedral volumetric barrier. It follows from Theorem 4.1 that $V(\cdot)$ is convex on the interior of $\mathscr{S}$. In the next theorem we demonstrate that in fact $V(\cdot)$ is strictly convex. Theorem 4.2 is also a direct extension of a result for the polyhedral volumetric barrier; see Anstreicher (1996, Theorem A.5).

THEOREM 4.2. *Let x have $S(x) \succ 0$. Then $Q(x) \succeq (1/m)H(x)$.*

PROOF. Let $\xi \in \Re^n$, $\xi \neq 0$. Then

$$\xi^T H\xi = \xi^T \mathscr{A}^T (S^{-1} \otimes S^{-1})\mathscr{A}\xi = \|(S^{-1/2} \otimes S^{-1/2})\mathscr{A}\xi\|^2.$$

Let $B = B(\xi) = \sum_{i=1}^n \xi_i A_i$, and $\bar{B} = S^{-1/2}BS^{-1/2}$. Since $\mathscr{A}\xi = \mathbf{vec}(B)$, we have

$$\xi^T H\xi = \|(S^{-1/2} \otimes S^{-1/2})\mathbf{vec}(B)\|^2 = \|\mathbf{vec}(\bar{B})\|^2 = \|\bar{B}\|^2 = \|\lambda\|^2, \tag{19}$$

where λ_i, $i = 1, \ldots, m$ are the eigenvalues of $\bar{B}$, with corresponding orthonormal eigenvectors v_i, $i = 1, \ldots, m$. As described in the proof of Theorem 4.1, $\bar{B}^2 \otimes_S I$ then has a full set of orthonormal eigenvectors $v_i \otimes v_j$, $i, j = 1, \ldots, m$, with corresponding eigenvalues $(1/2)(\lambda_i^2 + \lambda_j^2)$. It follows from (13) that

$$\xi^T Q\xi = P \cdot \tfrac{1}{2} \sum_{i,j} (\lambda_i^2 + \lambda_j^2)(v_i \otimes v_j)(v_i \otimes v_j)^T \tag{20}$$

$$= \tfrac{1}{2} \sum_{i,j} (\lambda_i^2 + \lambda_j^2)(v_i \otimes v_j)^T P(v_i \otimes v_j)$$

$$\geq \sum_i \lambda_i^2 (v_i \otimes v_i)^T P(v_i \otimes v_i). \tag{21}$$

On the other hand, $P\,\mathbf{vec}(\bar{B}) = \mathbf{vec}(\bar{B})$ implies that

$$\|\lambda\| = \|P\,\mathbf{vec}(\bar{B})\| = \left\|P\,\mathbf{vec}\left(\sum_{i=1}^m \lambda_i v_i v_i^T\right)\right\| = \left\|P \sum_{i=1}^m \lambda_i (v_i \otimes v_i)\right\| \leq \sum_{i=1}^m |\lambda_i| \upsilon_i, \tag{22}$$

where $\upsilon_i = \|P(v_i \otimes v_i)\|$. Then (21) and (22) together imply that $\xi^T Q\xi \geq \sum_{i=1}^m \upsilon_i^2 \lambda_i^2$, and $\sum_{i=1}^m \upsilon_i |\lambda_i| \geq \|\lambda\|$, from which it follows that $\xi^T Q\xi \geq (1/m)\|\lambda\|^2 = (1/m)\xi^T H\xi$. $\square$

For a given $\xi \in \Re^n$, and $\bar{B} = S^{-1/2}(\sum_{i=1}^n \xi_i A_i)S^{-1/2}$, the conclusion of Theorem 4.2 is that

$$\xi^T Q\xi \geq \frac{1}{m}\|\bar{B}\|^2. \tag{23}$$

A strengthening of (23), using $|\bar{B}|$ in place of $\|\bar{B}\|$, is a key element in our analysis of the self-concordance of $V(\cdot)$, in the next section. The next theorem gives a remarkably direct generalization of a result for the polyhedral volumetric barrier; see Anstreicher (1996, Proposition 2.3).

THEOREM 4.3. *Let x have $S = S(x) \succ 0$, $\xi \in \Re^n$, and $\bar{B} = S^{-1/2}(\sum_{i=1}^n \xi_i A_i)S^{-1/2}$. Then $\xi^T Q(x)\xi \geq (2/[1+\sqrt{m}])|\bar{B}|^2$.*

PROOF. Let v_i, $i = 1, \ldots, m$ be orthonormal eigenvectors of $\bar{B}$, with corresponding eigenvalues λ_i. Then $P\,\mathbf{vec}(\bar{B}) = \mathbf{vec}(\bar{B})$ can be written $P\sum_{i=1}^m \lambda_i(v_i \otimes v_i) = \sum_{i=1}^m \lambda_i(v_i \otimes v_i)$, and for any j we have

$$|\lambda_j| = \left|(v_j \otimes v_j)^T \left(P \sum_{i=1}^m \lambda_i (v_i \otimes v_i)\right)\right| \leq \upsilon_j \left(\sum_{i=1}^m |\lambda_i| \upsilon_i\right), \tag{24}$$

where $\upsilon_i = \|P(v_i \otimes v_i)\|$. Without loss of generality (scaling ξ as needed, and re-ordering indices) we may assume that $\|\lambda\|_\infty = |\lambda_1| = 1$. Then (24) implies that

$$\upsilon_1^2 + \upsilon_1 \sum_{i=2}^m |\lambda_i| \upsilon_i \geq 1.$$

Since $\xi^T Q\xi \geq \sum_{i=1}^m \lambda_i^2 v_i^2$, from (21), we are naturally led to consider the optimization problem

$$\begin{aligned} &\min_v \; v_1^2 + \sum_{i=2}^m \lambda_i^2 v_i^2, \\ &\sum_{i=2}^m |\lambda_i| v_i \geq \frac{1}{v_1} - v_1, \qquad (25) \\ &0 \leq v \leq e. \end{aligned}$$

For fixed $0 < v_1 \leq 1$, the constraint in (25) implies that

$$\sum_{i=2}^m \lambda_i^2 v_i^2 \geq \frac{1}{m-1}\left(\frac{1}{v_1} - v_1\right)^2 = \frac{1}{m-1}\left(\frac{1}{v_1^2} - 2 + v_1^2\right),$$

so the objective value in (25) can be no lower than

$$v_1^2 + \frac{1}{m-1}\left(\frac{1}{v_1^2} - 2 + v_1^2\right) = \frac{1}{m-1}\left(mv_1^2 + \frac{1}{v_1^2} - 2\right). \qquad (26)$$

A straightforward differentiation shows that the minimal value for (26) occurs when $v_1^2 = 1/\sqrt{m}$, and the value is then

$$\frac{1}{m-1}\left(mv_1^2 + \frac{1}{v_1^2} - 2\right) = \frac{1}{m-1}(2\sqrt{m} - 2) = \frac{2}{m-1}(\sqrt{m} - 1) = \frac{2}{\sqrt{m} + 1}.$$

We have thus shown that if $|\bar{B}| = 1$, then $\xi^T Q\xi \geq 2/(\sqrt{m} + 1)$. $\square$

Next we will obtain alternative representations of $\nabla V(x)$ and $Q(x)$ that emphasize the connection between the semidefinite volumetric barrier and the volumetric barrier for a polyhedral set. For fixed x with $S = S(x) \succ 0$, let $\bar{A}_i = S^{-1/2} A_i S^{-1/2}$, $i = 1, \ldots, n$. Let U_i, $i = 1, \ldots, n$ be symmetric matrices having $\|U_i\| = 1$ for all i, and $U_i \cdot U_j = 0$, $i \neq j$, such that the linear span of $\{U_i,\ i = 1, \ldots, n\}$ is equal to the span of $\{\bar{A}_i,\ i = 1, \ldots, n\}$. (Such $\{U_i\}$ may be obtained by applying a Gram-Schmidt procedure to $\{\bar{A}_i\}$.) Let U be the $m^2 \times n$ matrix whose ith column is $\mathbf{vec}(U_i)$, and let $\Sigma = \sum_{k=1}^n U_k^2$. Then $P = P(S)$, from (8), can be written in the form $P = UU^T$. It follows, from (7), that

$$\begin{aligned} \frac{\partial V(x)}{\partial x_i} &= -P \cdot (\bar{A}_i \otimes_S I) \\ &= -UU^T \cdot (\bar{A}_i \otimes_S I) \\ &= -\mathrm{tr}(U^T (\bar{A}_i \otimes_S I) U) \\ &= -\frac{1}{2} \sum_{k=1}^n \mathbf{vec}(U_k)^T [(\bar{A}_i \otimes I) + (I \otimes \bar{A}_i)] \mathbf{vec}(U_k) \\ &= -\frac{1}{2} \sum_{k=1}^n \mathbf{vec}(U_k)^T \mathbf{vec}(U_k \bar{A}_i + \bar{A}_i U_k) \\ &= -\frac{1}{2} \sum_{k=1}^n U_k \cdot (U_k \bar{A}_i + \bar{A}_i U_k) \end{aligned}$$

$$= -\bar{A}_i \cdot \left(\sum_{k=1}^{n} U_k^2\right)$$

$$= -\bar{A}_i \cdot \Sigma. \tag{27}$$

Similarly, from (12) we have

$$\begin{aligned}
Q(x)_{ij} &= P \cdot (\bar{A}_i\bar{A}_j \otimes_S I) \\
&= \operatorname{tr}(U^T(\bar{A}_i\bar{A}_j \otimes_S I)U) \\
&= \sum_{k=1}^{n} \mathbf{vec}(U_k)^T(\bar{A}_i\bar{A}_j \otimes_S I)\mathbf{vec}(U_k) \\
&= \frac{1}{2}\sum_{k=1}^{n} \mathbf{vec}(U_k)^T\mathbf{vec}(\bar{A}_i\bar{A}_jU_k + U_k\bar{A}_i\bar{A}_j) \\
&= \frac{1}{2}\sum_{k=1}^{n} U_k \cdot (\bar{A}_i\bar{A}_jU_k + U_k\bar{A}_i\bar{A}_j) \\
&= \frac{1}{2}\sum_{k=1}^{n} \operatorname{tr}(U_k(\bar{A}_j\bar{A}_iU_k + U_k\bar{A}_j\bar{A}_i)) \\
&= \sum_{k=1}^{n} \operatorname{tr}(\bar{A}_iU_k^2\bar{A}_j)
\end{aligned}$$

$$= \operatorname{tr}(\bar{A}_i\Sigma\bar{A}_j). \tag{28}$$

The characterizations of $\nabla V(x)$ and $Q(x)$ given in (27) and (28) are very convenient for the proof of the following theorem, which will be required in the analysis of self-concordance in the next section.

THEOREM 4.4. *Let x have $S(x) \succ 0$. Then $\nabla V(x)Q(x)^{-1}\nabla V(x)^T \leq n$.*

PROOF. From (28) we have

$$\begin{aligned}
Q(x)_{ij} &= \mathbf{vec}(\bar{A}_i)^T\mathbf{vec}(\Sigma\bar{A}_j) \\
&= \mathbf{vec}(\bar{A}_i)^T(I \otimes \Sigma)\mathbf{vec}(\bar{A}_j),
\end{aligned}$$

using Proposition 2.2(5). Letting $\bar{\mathscr{A}}$ be the $m^2 \times n$ matrix whose ith column is $\mathbf{vec}(\bar{A}_i)$, we can then write

$$Q(x) = \bar{\mathscr{A}}^T(I \otimes \Sigma)\bar{\mathscr{A}}. \tag{29}$$

In addition, it follows from (27) that

$$\nabla V(x)^T = -\bar{\mathscr{A}}^T\mathbf{vec}(\Sigma). \tag{30}$$

Combining (29) and (30), we obtain

$$\begin{aligned}
&\nabla V(x)[Q(x)]^{-1}\nabla V(x)^T \\
&\quad = \mathbf{vec}(\Sigma)^T\bar{\mathscr{A}}(\bar{\mathscr{A}}^T(I \otimes \Sigma)\bar{\mathscr{A}})^{-1}\bar{\mathscr{A}}^T\mathbf{vec}(\Sigma) \\
&\quad = \mathbf{vec}(\Sigma^{1/2})^T(I \otimes \Sigma^{1/2})\bar{\mathscr{A}}(\bar{\mathscr{A}}^T(I \otimes \Sigma)\bar{\mathscr{A}})^{-1}\bar{\mathscr{A}}^T(I \otimes \Sigma^{1/2})\mathbf{vec}(\Sigma^{1/2})
\end{aligned}$$

TABLE 1. Comparison of logarithmic and volumetric barriers

	Polyhedral	Semidefinite
Logarithmic	$\nabla f_i = \bar{a}_i^T e$	$\nabla f_i = \mathrm{tr}(\bar{A}_i)$
	$H_{ij} = \bar{a}_i^T \bar{a}_j$	$H_{ij} = \mathrm{tr}(\bar{A}_i \bar{A}_j)$
Volumetric	$\nabla V_i = \bar{a}_i^T \sigma$	$\nabla V_i = \mathrm{tr}(\bar{A}_i \Sigma)$
	$Q_{ij} = \bar{a}_i^T \Sigma \bar{a}_j$	$Q_{ij} = \mathrm{tr}(\bar{A}_i \Sigma \bar{A}_j)$

$$
\begin{aligned}
&\leq \mathbf{vec}(\Sigma^{1/2})^T \mathbf{vec}(\Sigma^{1/2}) \\
&= \mathrm{tr}(\Sigma) \\
&= n,
\end{aligned}
$$

because $\Sigma = \sum_{k=1}^n U_k^2$, and $\mathrm{tr}(U_k^2) = U_k \cdot U_k = 1$ for each k, by construction. □

One final point concerning the matrix Σ is the issue of uniqueness, for a given $S = S(x)$. Since Σ is defined above in terms of $\{U_i\}$, and the $\{U_i\}$ are not unique, it is not at all obvious that Σ is unique. We will now show that Σ is invariant to the choice of $\{U_i\}$, and is therefore unique. To see this, note that by definition

$$
\Sigma_{ij} = \sum_{k=1}^n (U_k^2)_{ij} = \sum_{k=1}^n (U_k)_i^T (U_k)_j, \tag{31}
$$

where $(U_k)_i$ denotes the ith column of U_k (recall that each U_k is symmetric by construction). Let $e_i \in \Re^m$ denote the ith elementary vector, and let I be an $m \times m$ identity matrix. By inspection $[e_i \otimes I]^T U$ is then the $m \times n$ matrix whose kth column is $(U_k)_i$. It follows from (31) that

$$
\Sigma_{ij} = \mathrm{tr}(U^T[e_j \otimes I][e_i \otimes I]^T U) = \mathrm{tr}([e_i \otimes I]^T U U^T [e_j \otimes I]) = \mathrm{tr}([e_i \otimes I]^T P [e_j \otimes I]),
$$

where P is the projection from (8). Since P is uniquely determined by $\{A_i\}$ and $S = S(x)$, Σ is also unique, as claimed.

In the following table we give a summary of first and second order information for the logarithmic and volumetric barriers, for polyhedral and semidefinite constraints. For the polyhedral case we have $s = s(x) = Ax - c$, where A is an $m \times n$ matrix whose ith column is a_i. Given x with $s = s(x) > 0$, we let $S = \mathrm{Diag}(s)$, $\bar{a}_i = S^{-1} a_i$, $P = P(s) = S^{-1}A(A^T S^{-2} A)^{-1} A^T S^{-1}$, $\sigma \in \Re^m$ be the vector whose components are those of the diagonal of P, and $\Sigma = \mathrm{Diag}(\sigma)$. For the volumetric barrier, in both the polyhedral and semidefinite cases, the matrix Q satisfies $Q(x) \preceq \nabla^2 V(x) \preceq 3Q(x)$. Note that, as should be the case, all formulas for the semidefinite case also apply to the polyhedral case, with $\bar{A}_i = \mathrm{Diag}(\bar{a}_i)$.

5. Self-concordance. In this section we obtain proofs for the self-concordance results in Theorems 3.1 and 3.2. We begin with an analysis of the third directional derivatives of $V(\cdot)$. Let x have $S(x) \succ 0$, and $\xi \in \Re^n$. Using (4), (9), and (13), we immediately obtain

$$
\begin{aligned}
\frac{\partial}{\partial x_i} \xi^T Q(x) \xi = {} & 2\mathscr{A} H^{-1} \mathscr{A}^T [S^{-1} A_i S^{-1} \otimes_S S^{-1}] \mathscr{A} H^{-1} \mathscr{A}^T \cdot [S^{-1} B S^{-1} B S^{-1} \otimes_S S^{-1}] \\
& + \mathscr{A} H^{-1} \mathscr{A}^T \cdot \frac{\partial}{\partial x_i} [S^{-1} B S^{-1} B S^{-1} \otimes_S S^{-1}],
\end{aligned} \tag{32}
$$

where $B = B(\xi) = \sum_{i=1}^n \xi_i A_i$. Moreover, from (4) it is immediate that

$$\frac{\partial}{\partial x_i}[S^{-1}BS^{-1}BS^{-1}\otimes_S S^{-1}]$$
$$= -(S^{-1}A_iS^{-1}BS^{-1}BS^{-1} + S^{-1}BS^{-1}A_iS^{-1}BS^{-1} + S^{-1}BS^{-1}BS^{-1}A_iS^{-1})\otimes_S S^{-1}$$
$$-S^{-1}BS^{-1}BS^{-1}\otimes_S S^{-1}A_iS^{-1}. \tag{33}$$

Combining (32) and (33), and using Proposition 2.1(2), we conclude that the first directional derivative of $\xi^T Q(x)\xi$, in the direction ξ, is given by

$$\begin{aligned} D^1\xi^T Q(x)\xi[\xi] &= \sum_{i=1}^n \xi_i \frac{\partial}{\partial x_i}\xi^T Q(x)\xi \\ &= 2\mathscr{A}H^{-1}\mathscr{A}\cdot(S^{-1}BS^{-1}\otimes_S S^{-1})\mathscr{A}H^{-1}\mathscr{A}(S^{-1}BS^{-1}BS^{-1}\otimes_S S^{-1}) \\ &\quad -3\mathscr{A}H^{-1}\mathscr{A}\cdot(S^{-1}BS^{-1}BS^{-1}BS^{-1}\otimes_S S^{-1}) \\ &\quad -\mathscr{A}H^{-1}\mathscr{A}\cdot(S^{-1}BS^{-1}BS^{-1}\otimes_S S^{-1}BS^{-1}) \\ &= 2P\cdot[\bar{B}\otimes_S I]P[\bar{B}^2\otimes_S I] - 3P\cdot[\bar{B}^3\otimes_S I] - P\cdot[\bar{B}^2\otimes_S \bar{B}], \end{aligned} \tag{34}$$

where $\bar{B} = S^{-1/2}BS^{-1/2}$, and P is defined as in (8). Very similar computations, using (14) and (15), result in

$$D^1\xi^T R(x)\xi[\xi] = 2P\cdot[\bar{B}\otimes_S I]P[\bar{B}\otimes\bar{B}] - 4P\cdot[\bar{B}^2\otimes_S \bar{B}], \tag{35}$$

$$\begin{aligned} D^1\xi^T T(x)\xi[\xi] &= 4P\cdot[\bar{B}\otimes_S I]P[\bar{B}\otimes_S I]P[\bar{B}\otimes_S I] - 4P\cdot[\bar{B}\otimes_S I]P[\bar{B}^2\otimes_S I] \\ &\quad -2P\cdot[\bar{B}\otimes_S I]P[\bar{B}\otimes\bar{B}]. \end{aligned} \tag{36}$$

Combining (11) with (34), (35), and (36), we obtain the third directional derivative of $V(\cdot)$:

$$\begin{aligned} D^3V(x)[\xi,\xi,\xi] &= 12P\cdot[\bar{B}\otimes_S I]P[\bar{B}^2\otimes_S I] - 6P\cdot[\bar{B}^3\otimes_S I] - 6P\cdot[\bar{B}^2\otimes_S \bar{B}] \\ &\quad + 6P\cdot[\bar{B}\otimes_S I]P[\bar{B}\otimes\bar{B}] - 8P\cdot[\bar{B}\otimes_S I]P[\bar{B}\otimes_S I]P[\bar{B}\otimes_S I]. \end{aligned} \tag{37}$$

THEOREM 5.1. *Let x have $S = S(x)\succ 0$, $\xi\in\Re^n$, and $\bar{B} = S^{-1/2}(\sum_{i=1}^n \xi_i A_i)S^{-1/2}$. Then $|D^3V(x)[\xi,\xi,\xi]| \le 30|\bar{B}|\xi^T Q(x)\xi$.*

PROOF. Using the fact that

$$[\bar{B}^2\otimes_S I][\bar{B}\otimes_S I] = \tfrac{1}{2}([\bar{B}^3\otimes_S I] + [\bar{B}^2\otimes_S \bar{B}]),$$

from Proposition 2.2(2), (37) can be re-written as

$$\begin{aligned} D^3V(x)[\xi,\xi,\xi] &= P[\bar{B}\otimes_S I]P\cdot(12[\bar{B}^2\otimes_S I] + 6[\bar{B}\otimes\bar{B}] - 8[\bar{B}\otimes_S I]P[\bar{B}\otimes_S I]) \\ &\quad -12P\cdot[\bar{B}^2\otimes_S I][\bar{B}\otimes_S I]. \end{aligned} \tag{38}$$

We will analyze the two terms in (38) separately. First, from (17) we have

$$12[\bar{B}^2\otimes_S I] + 6[\bar{B}\otimes\bar{B}] - 8[\bar{B}\otimes_S I]P[\bar{B}\otimes_S I] \succeq 8[\bar{B}^2\otimes_S I] + 2[\bar{B}\otimes\bar{B}].$$

Using (18), and the similar relationship $[\bar{B}\otimes\bar{B}]\succeq -[\bar{B}^2\otimes_S I]$, it follows that

$$6[\bar{B}^2\otimes_S I] \preceq 12[\bar{B}^2\otimes_S I] + 6[\bar{B}\otimes\bar{B}] - 8[\bar{B}\otimes_S I]P[\bar{B}\otimes_S I] \preceq 18[\bar{B}^2\otimes_S I]. \tag{39}$$

Let λ_i, $i = 1,\ldots,m$ be the eigenvalues of $\bar{B}$. Then (see Horn and Johnson 1991, Theorem 4.4.5) the eigenvalues of $[\bar{B}\otimes_S I]$ are of the form $(1/2)(\lambda_i+\lambda_j)$, $i,j = 1,\ldots,m$, so

$$-|\bar{B}|I \preceq [\bar{B} \otimes_S I] \preceq |\bar{B}|I,$$

$$-|\bar{B}|P \preceq P[\bar{B} \otimes_S I]P \preceq |\bar{B}|P. \tag{40}$$

Using (39), (40), the fact that $[\bar{B}^2 \otimes_S I] \succeq 0$, and Proposition 2.1(4), we then obtain

$$\begin{aligned} &|P[\bar{B} \otimes_S I]P \cdot (12[\bar{B}^2 \otimes_S I] + 6[\bar{B} \otimes \bar{B}] - 8[\bar{B} \otimes_S I]P[\bar{B} \otimes_S I])| \\ &\quad \le 18|\bar{B}|P \cdot [\bar{B}^2 \otimes_S I]. \end{aligned} \tag{41}$$

In addition, the fact that $[\bar{B}^2 \otimes_S I]$ and $[\bar{B} \otimes_S I]$ have the same eigenvectors implies that

$$-|\bar{B}|[\bar{B}^2 \otimes_S I] \preceq [\bar{B}^2 \otimes_S I][\bar{B} \otimes_S I] \preceq |\bar{B}|[\bar{B}^2 \otimes_S I],$$

and therefore

$$|P \cdot [\bar{B}^2 \otimes_S I][\bar{B} \otimes_S I]| \le |\bar{B}|P \cdot [\bar{B}^2 \otimes_S I]. \tag{42}$$

The proof is completed by combining (38), (41), (42), and (13). □

Using Theorem 5.1 we can now prove the first main result of the paper, characterizing the self-concordance of $V(\cdot)$.

PROOF OF THEOREM 3.1. The fact that $V(x) \to \infty$ as x approaches the boundary of $\mathscr{S}$ follows from (3), and the fact that $S(x)$ is singular on the boundary of $\mathscr{S}$. Combining the results of Theorems 4.3 and 5.1, we obtain

$$|D^3V(x)[\xi,\xi,\xi]| \le 30\left(\frac{1+\sqrt{m}}{2}\right)^{1/2}(\xi^TQ(x)\xi)^{3/2} \le 30m^{1/4}(D^2V(x)[\xi,\xi])^{3/2}, \tag{43}$$

using the fact that $\xi^TQ(x)\xi \le \xi^T\nabla^2V(x)\xi = D^2V(x)[\xi,\xi]$, from Theorem 4.1. In addition, $0 \prec Q(x) \preceq \nabla^2V(x)$ implies that $\nabla^2V(x)^{-1} \preceq Q(x)^{-1}$ (see Horn and Johnson 1985, Corollary 7.7.4), so Theorem 4.4 implies that

$$\nabla V(x)\nabla^2V(x)^{-1}\nabla V(x)^T \le n. \tag{44}$$

The proof is completed by noting the effect on (43) and (44) when $V(\cdot)$ is multiplied by the factor 225 $m^{1/2}$. □

Next we consider the self-concordance of the combined volumetric-logarithmic barrier $V_\rho(\cdot)$, as defined in §3. We begin with some well-known properties of the logarithmic barrier $f(\cdot)$.

LEMMA 5.2. *Let x have $S = S(x) \succ 0$, $\xi \in \Re^n$, and $\bar{B} = S^{-1/2}(\sum_{i=1}^n \xi_iA_i)S^{-1/2}$. Then $|D^3f(x)[\xi,\xi,\xi]| \le 2|\bar{B}|\xi^TH(x)\xi$.*

PROOF. From (3) and (5) we obtain

$$\begin{aligned} \frac{\partial H_{ij}(x)}{\partial x_k} &= -2\mathbf{vec}(A_i)^T[S^{-1}A_kS^{-1} \otimes_S S^{-1}]\mathbf{vec}(A_j) \\ &= -2\mathbf{vec}(A_i)^T(S^{-1/2} \otimes S^{-1/2})[S^{-1/2}A_kS^{-1/2} \otimes_S I](S^{-1/2} \otimes S^{-1/2})\mathbf{vec}(A_j) \\ &= -2\mathbf{vec}(\bar{A}_i)^T[\bar{A}_k \otimes_S I]\mathbf{vec}(\bar{A}_j), \end{aligned} \tag{45}$$

where $\bar{A}_i = S^{-1/2}A_iS^{-1/2}$ for each i. It follows easily from (45) that

$$\frac{\partial \xi^TH(x)\xi}{\partial x_k} = -2\mathbf{vec}(\bar{B})^T[\bar{A}_k \otimes_S I]\mathbf{vec}(\bar{B}),$$

and therefore

$$D^3 f(x)[\xi,\xi,\xi] = \sum_{k=1}^{m} \xi_k \frac{\partial \xi^T H(x)\xi}{\partial x_k} = -2\mathbf{vec}(\bar{B})^T[\bar{B} \otimes_S I]\mathbf{vec}(\bar{B}).$$

It is then immediate from the fact that $|\bar{B} \otimes_S I| = |\bar{B}|$ (see the proof of Theorem 5.1) that

$$|D^3 f(x)[\xi,\xi,\xi]| \le 2|\bar{B}|\mathbf{vec}(\bar{B})^T\mathbf{vec}(\bar{B}) = 2|\bar{B}|\,\|\bar{B}\|^2 = 2|\bar{B}|\xi^T H(x)\xi,$$

where the final equality uses (19). $\square$

It follows from Lemma 5.2, the fact that $|\bar{B}| \le \|\bar{B}\|$, and

$$\nabla f(x)(\nabla^2 f(x))^{-1}\nabla f(x)^T = \mathbf{vec}(I)^T \bar{\mathscr{A}}(\bar{\mathscr{A}}^T\bar{\mathscr{A}})^{-1}\bar{\mathscr{A}}\,\mathbf{vec}(I) \le \mathbf{vec}(I)^T\mathbf{vec}(I) = m,$$

that $f(\cdot)$ is an m-self-concordant barrier for $\mathscr{S}$, as shown by Nesterov and Nemirovskii (1994). Using Lemma 5.2 we immediately obtain the following generalization of Theorem 5.1.

Corollary 5.3. *Let x have $S = S(x) \succ 0$, $\xi \in \Re^n$, and $\bar{B} = S^{-1/2}(\sum_{i=1}^n \xi_i A_i)S^{-1/2}$. Then $|D^3 V_\rho(x)[\xi,\xi,\xi]| \le 30|\bar{B}|\xi^T(Q(x) + \rho H(x))\xi$.*

Proof. Combining Theorem 5.1 and Lemma 5.2, we obtain

$$\begin{aligned} |D^3 V_\rho(x)[\xi,\xi,\xi]| &\le |D^3 V(x)[\xi,\xi,\xi]| + \rho|D^3 f(x)[\xi,\xi,\xi]| \\ &\le 30|\bar{B}|\xi^T Q(x)\xi + 2\rho|\bar{B}|\xi^T H(x)\xi \\ &\le 30|\bar{B}|\xi^T(Q(x) + \rho H(x))\xi. \quad \square \end{aligned}$$

Next we require a generalization of Theorem 4.3 that applies with $Q(x)+\rho H(x)$ in place of $Q(x)$. The following theorem obtains a direct extension of a result for the polyhedral combined barrier (Anstreicher 1996, Theorem 3.3) to the semidefinite case. To prove the theorem we will utilize the matrices $\{U_k\}$, as defined in §4, to reduce the theorem to a problem already analyzed in the proof of Anstreicher (1996, Theorem 3.3).

Theorem 5.4. *Let x have $S = S(x) \succ 0$, $\xi \in \Re^n$, and $\bar{B} = S^{-1/2}(\sum_{i=1}^n \xi_i A_i)S^{-1/2}$. Then $\xi^T(Q(x) + \rho H(x))\xi \ge [2\sqrt{\rho(m-1)+1}/(1+\sqrt{m})]|\bar{B}|^2$.*

Proof. Let v_i, $i = 1,\ldots,m$ be orthonormal eigenvectors of $\bar{B}$, with corresponding eigenvalues λ_i. By the definition of $\{U_k\}$, there is a vector $\bar{\xi} \in \Re^n$ such that

$$\bar{B} = \sum_{i=1}^{m} \lambda_i v_i v_i^T = \sum_{k=1}^{n} \bar{\xi}_k U_k, \tag{46}$$

and therefore $\|\bar{B}\| = \|\lambda\| = \|\bar{\xi}\|$. It follows from (46) that for each $i = 1,\ldots,m$,

$$v_i^T \bar{B} v_i = \lambda_i = \sum_{k=1}^{n} \bar{\xi}_k v_i^T U_k v_i,$$

and therefore $\lambda = W\bar{\xi}$, where W is the $m \times n$ matrix with

$$w_{ik} = v_i^T U_k v_i = U_k \cdot v_i v_i^T = (v_i \otimes v_i)^T \mathbf{vec}(U_k). \tag{47}$$

Let U be the $m^2 \times n$ matrix whose kth column is $\mathbf{vec}(U_k)$, and let V be the $m^2 \times m$ matrix whose ith column is $v_i \otimes v_i$. From (47) we can then write $W = V^T U$. Now let w_i denote the ith row of W. Then

$$\|w_i\|^2 = (WW^T)_{ii} = (V^T U U^T V)_{ii} = (v_i \otimes v_i)^T P (v_i \otimes v_i) = \|P(v_i \otimes v_i)\|^2, \tag{48}$$

where P is the projection matrix from (8). Using (19), (21), and (48) we then have

$$\xi^T (Q(x) + \rho H(x))\xi \geq \sum_{i=1}^m \lambda_i^2 \|P(v_i \otimes v_i)\|^2 + \rho\|\lambda\|^2 = \sum_{i=1}^m (w_i^T \bar{\xi})^2 \|w_i\|^2 + \rho\|\bar{\xi}\|^2.$$

Moreover it is clear that $\sum_{i=1}^m (w_i^T \bar{\xi})^2 = \|\lambda\|^2 = \|\bar{\xi}\|^2$, and also $|\bar{B}| = \|\lambda\|_\infty = \|W\bar{\xi}\|_\infty$. We are now exactly in the structure of the proof of Anstreicher (1996, Theorem 3.3), with U of that proof replaced by the matrix W. In that proof it is shown that the solution objective value of the problem

$$\min_{W,\bar{\xi}} \rho\|\bar{\xi}\|^2 + \sum_{i=1}^m (w_i^T \bar{\xi})^2 \|w_i\|^2,$$

$$\sum_{i=1}^m (w_i^T \bar{\xi})^2 = \|\bar{\xi}\|^2,$$

$$\|W\bar{\xi}\|_\infty = 1,$$

can be no lower than

$$\frac{2\sqrt{\rho(m-1)+1}}{1+\sqrt{m}}.$$

It follows that $\xi^T (Q(x) + \rho H(x))\xi \geq [2\sqrt{\rho(m-1)+1}/(1+\sqrt{m})]|\bar{B}|^2$, as claimed. □

The final ingredient needed to prove the self-concordance of $V_\rho(\cdot)$ is the following simple generalization of Theorem 4.4.

THEOREM 5.5. *Let x have $S(x) \succ 0$. Then $\nabla V_\rho(x)[Q(x) + \rho H(x)]^{-1} \nabla V_\rho(x)^T \leq n + \rho m$.*

PROOF. From the representations in Table 1 we easily obtain

$$\nabla V_\rho(x)^T = -\bar{\mathscr{A}}^T \mathbf{vec}(\Sigma + \rho I),$$
$$Q(x) + \rho H(x) = \bar{\mathscr{A}}^T [I \otimes (\Sigma + \rho I)]\bar{\mathscr{A}}.$$

Let $\Sigma_\rho = \Sigma + \rho I$. It follows that

$$\begin{aligned}
&\nabla V_\rho(x)[Q(x) + \rho H(x)]^{-1} \nabla V_\rho(x)^T \\
&\quad = \mathbf{vec}(\Sigma_\rho^{1/2})[I \otimes \Sigma_\rho^{1/2}]\bar{\mathscr{A}}(\bar{\mathscr{A}}^T [I \otimes \Sigma_\rho]\bar{\mathscr{A}})^{-1} \bar{\mathscr{A}}^T [I \otimes \Sigma_\rho^{1/2}]\mathbf{vec}(\Sigma_\rho^{1/2}) \\
&\quad \leq \mathbf{vec}(\Sigma_\rho^{1/2})^T \mathbf{vec}(\Sigma_\rho^{1/2}) \\
&\quad = \mathrm{tr}(\Sigma_\rho) \\
&\quad = n + \rho m. \quad \square
\end{aligned}$$

Using the above results we can now prove the second main result of the paper, characterizing the self-concordance of the combined volumetric-logarithmic barrier for $\mathscr{S}$.

Proof of Theorem 3.2. Combining the results of Corollary 5.3 and Theorem 5.4, with $\rho = (n-1)/(m-1)$, we obtain

$$|D^3 V_\rho(x)[\xi,\xi,\xi]| \le 30 \left(\frac{1+\sqrt{m}}{2\sqrt{n}}\right)^{1/2} (\xi^T(Q(x)+\rho H(x))\xi)^{3/2} \le 30 \left(\frac{m}{n}\right)^{1/4} (D^2 V_\rho(x)[\xi,\xi])^{3/2}, \tag{49}$$

using the fact that $\xi^T Q(x)\xi \le \xi^T \nabla^2 V(x)\xi = D^2 V(x)[\xi,\xi]$, from Theorem 4.1. In addition, $0 \prec Q(x) \preceq \nabla^2 V(x)$ implies that $\nabla^2 V_\rho(x)^{-1} \preceq (Q(x)+\rho H(x))^{-1}$ (see Horn and Johnson 1985, Corollary 7.7.4), so Theorem 5.5 implies that

$$\nabla V_\rho(x) \nabla^2 V_\rho(x)^{-1} \nabla V_\rho(x)^T \le n + m(n-1)/(m-1) < 2n. \tag{50}$$

The proof is completed by noting the effect on (49) and (50) when $V_\rho(\cdot)$ is multiplied by the factor $225(m/n)^{1/2}$. $\square$

References

Alizadeh, F. 1995. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.* **5** 13–51.

Anstreicher, K. M. 1996. Large step volumetric potential reduction algorithms for linear programming. *Ann. Oper. Res.* **62** 521–538.

——. 1997a. Volumetric path following algorithms for linear programming. *Math. Programming* **76** 245–263.

——. 1997b. On Vaidya's volumetric cutting plane method for convex programming. *Math. Oper. Res.* **22** 63–89.

——. 1999a. Towards a practical volumetric cutting plane method for convex programming. *SIAM J. Optim.* **9** 190–206.

——. 1999b. Ellipsoidal approximations of convex sets based on the volumetric barrier. *Math. Oper. Res.* **24** 193–203.

Graham, A. 1981. *Kronecker Products and Matrix Calculus*: *With Applications*. Ellis Horwood, Chichester, U.K.

Horn, R. A., C. R. Johnson, 1985. *Matrix Analysis*. Cambridge University Press, Cambridge, U.K.

——, ——. 1991. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, U.K.

Nesterov, N., A. Nemirovskii. 1994. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, PA.

Vaidya, P. M. 1996. A new algorithm for minimizing convex functions over convex sets. *Math. Programming* **73** 291–341.

——, D. S. Atkinson. 1993. A technique for bounding the number of iterations in path following algorithms. P. M. Pardaos, ed. *Complexity in Numerical Optimization*. World Scientific, Singapore, 462–489.

Vandenberghe, L., S. Boyd. 1996. Semidefinite programming. *SIAM Rev.* **38** 49–95.

K. Anstreicher: Department of Management Sciences, University of Iowa, Iowa City, Iowa 52242