RUTCOR
RESEARCH
REPORT

Associative Algebras, Symmetric Cones and Polynomial Time Interior Point Algorithms

S. H. Schmieta[a] F. Alizadeh[b]

RRR 17-98, June 1998

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/~rrr

[a]RUTCOR, Rutgers University, e-mail: schmieta@rutcor.rutgers.edu
[b]RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854-8003, e-mail: alizadeh@rutcor.rutgers.edu

Rutcor Research Report
RRR 17-98, June 1998

Associative Algebras, Symmetric Cones and Polynomial Time Interior Point Algorithms

S. H. Schmieta F. Alizadeh

Abstract. We present a generalization of Monteiro's polynomiality proof of a large class of primal-dual interior point algorithms for semidefinite programs. We show that this proof, essentially verbatim, applies to all optimization problems over almost all symmetric cones, that is those cones that can be derived from classes of normed associative algebras and certain Jordan algebras obtained from them. As a consequence, we show that Monteiro's proof can be extended to convex quadratically constrained quadratic optimization problems. We also extend the notion of Zhang family of algorithms, and show that it can be applied to all symmetric cones, in particular the quadratic cone.

Acknowledgements: Research supported in part by the U.S. National Science Foundation grant CCR-9501941 and Office of Naval Research contract number N00014-96-1-0704

1 Introduction

Consider the primal and dual pair of *cone LP* optimization problems in the standard form:

$$\text{(P)} \quad \min\{\langle c, x\rangle : \; Ax = b, \; x \geq_{\mathcal{K}} 0\} \tag{1}$$

$$\text{(D)} \quad \max\{\langle b, y\rangle : \; A^T y + s = c, \; s \geq_{\mathcal{K}^*} 0\} \tag{2}$$

Here $\mathcal{K}$ is a closed, pointed convex cone and $\mathcal{K}^*$ is its dual:

$$\mathcal{K}^* = \{x : \text{for every } z \in \mathcal{K} \; \langle z, x\rangle \geq 0\}.$$

The notation $x \geq_{\mathcal{K}} y$ means that $x - y \in \mathcal{K}$, $x >_{\mathcal{K}} y$ means $x - y \in \mathcal{K}^+$ the interior of $\mathcal{K}$.

It is well-known and easy to see that all convex optimization problems can be formulated as cone LPs. Moreover, if there are strictly feasible primal and dual points–those feasible points that lie in $\mathcal{K}^+$ for the primal or $\mathcal{K}^{*+}$ for the dual–then the corresponding optimal values coincide. This duality theory gives rise to a complementary slackness theorem: A set of generally nonlinear equations which, along with primal and dual feasibility, in principle determine the optimal solution. It is this property that is exploited in all primal-dual algorithms, including those based on interior point methods.

In this paper we focus on the special case where $\mathcal{K}$ is a symmetric cone, that is $\mathcal{K} = \mathcal{K}^*$ under an appropriate inner product and homogeneous: For every pair of points $x, z \in \mathcal{K}^+$ there is a linear transformation A where $A(x) = z$ and $A(\mathcal{K}) = \mathcal{K}$. Optimization over symmetric cones contains as a special case linear programming (LP), semidefinite programming (SDP) (that is a cone-LP where $\mathcal{K}$ is the set of positive semidefinite real symmetric matrices), and convex quadratically constrained quadratic programming (QCQP)–the latter problem can be expressed as a cone LP where $\mathcal{K}$ is the quadratic cone[1]. We show that the polynomial-time complexity of a large class of primal-dual algorithms proved by Monteiro in [Mon96] for the special case of semidefinite programming extends verbatim to optimization problems over symmetric cones. The most interesting corollary of this observation is that Monteiro's proof also works for quadratic cones. More generally we present a framework that proofs and algorithms dealing with SDP can be extended directly to all symmetric cones.

A related approach based on properties of Jordan algebras has been proposed by Faybusovich in [Fay97a, Fay97b] and Tsuchiya [Tsu97, Tsu98] (the latter concentrates on applying Jordan algebra methods to QCQP problems.) Faybusovich [Fay97a] works exclusively with Jordan algebra and its multiplication and develops his complexity analysis within this confine. Tsuchiya uses a similar approach but focuses on the QCQP part. He manages to extend the Monteiro-Zhang analysis of SDP [MZ96] to QCQP problems. He also studies several neighborhoods and examines the so-called short, semi-long and long step path-following methods. Neither of these works deal with the entire Zhang family, or analyze the "pure Newton" methods of $XS + SX$ of [AHO96] or $X^{1/2}SX^{1/2}$ of [MT96, Tse96] (see below for definitions).

[1] Also known as the second order cone, the Lorenz cone and the ice cream cone.

Our approach is more general than those of Faybusovich and Tsuchiya in that we do not rely solely on the Jordan algebra operations. Rather we assume that the euclidean Jordan algebra associated with our symmetric cone is derived from an associative algebra whose properties are used prominently in our development. This approach seems to be necessary in some contexts, for instance in defining the Zhang family and in extending Monteiro's analysis [Mon96] as well as the analysis of Monteiro and Tsuchiya [MT96] and Tseng [Tse96].

We should note that there are Jordan Algebras that are not derived from any associative algebra. In case of euclidean Jordan algebras all except one special class are derivable from such associative algebras, though; the only hold-out is the 27-dimensional algebra of 3×3 hermitian matrices over octonions. It is known that this Jordan algebra is not derivable form any associative algebra and thus our results do not extend directly to cones which are partly constructed by such positive semidefinite matrices. (Schafer attributes this result to A. A. Albert, see [Sch66] page 104.)

Symmetric cones are completely characterized by the following cones:

1. The quadratic cone
$$\mathcal{Q}_n = \left\{ x = (x_0, x_1, \ldots, x_n) : x_0 \geq \sqrt{\sum_{i=1}^{n} x_i^2} \right\};$$
2. the set of $n \times n$ real, symmetric, positive semidefinite matrices;
3. the set of $n \times n$ hermitian positive semidefinite matrices made of complex numbers;
4. the set of $n \times n$ positive semidefinite matrices with quaternion entries;
5. the set of 3×3 positive semidefinite matrices with octonion entries;
6. any cone $\mathcal{K}_1 \times \mathcal{K}_2$ where $\mathcal{K}_1$ and $\mathcal{K}_2$ are themselves symmetric cones.

In view of our earlier remark regarding the 3×3 positive semidefinite matrices with octonion entries we will exclude their consideration in this paper. Herein whenever we use the term symmetric cone, we mean those that are not constructed by such matrices. Likewise, our reference to euclidean Jordan algebras excludes those that are at least partially constructed by the 3×3 hermitian matrices on octonions. See the text of Faraut and Korány [FK94] for a complete treatment of symmetric cones and their relationship to euclidean Jordan algebras.

In the remainder of this section we outline a brief background. To our knowledge it was O. Güler who first brought to the attention of optimization researchers the connection between symmetric cones and extensions of semidefinite programming. He also observed that the theory of self-concordant barriers developed by Nesterov and Nemirovskii [NN94] becomes especially simple for this class of cones.

Faybusovich was not the first person who studied primal-dual algorithms in the context of symmetric cones. It was Nesterov and Todd who in [NT94, NT95] developed the theory of self-scaled barriers and self-scaled cones and using the Nesterov-Nemirovskii theory developed a new class of primal-dual algorithms which is an extension of Monteiro-Adler [MA89] and Kojima, Mizuno and Yoshise [KMY89] methods for linear programming. They also prove a $\sqrt{r}$ iteration bound where r is the self-concordance parameter of the underlying cone. It turns out however that the class of self-scaled cones is identical to the class of symmetric cones, and Nesterov-Todd algorithm is applicable only to cone LP over such cones.

In the special cases of semidefinite programming and convex quadratically constrained quadratic programming there has been an extensive work on primal-dual algorithms in recent years. Earlier, in [Ali95] a simple approach was proposed to extend linear programming interior point algorithms and their polynomiality proofs to semidefinite programming by a word-by-word substitution. A similar technique was proposed by Nemirovskii and Scheinberg for QCQP problems [NS96]. These extensions apply to most potential reduction methods or primal only or dual only algorithms, but do not work with primal-dual algorithms that generalize the Monteiro-Adler LP method. Such an extension relaxes the complementarity condition $XS = 0$ in SDP to $XS = \mu I$ (or $SX = \mu I$ or many other forms). These methods yield different directions because in general $XS \neq SX$. In addition to the Nesterov-Todd method, Kojima, Shindoh and Hara prove a $\sqrt{n}$ iteration bound for SDP over $n \times n$ real symmetric matrices. We refer to these algorithms as the XS and the SX methods, although elsewhere they are referred to as the HRWV/KSH/M method as [HRVW96] and [Mon97] also proposed essentially the same techniques. Alizadeh, Haeberly and Overton in [AHO96] proposed the $XS+SX$ method for SDP which used $XS+SX = 2\mu I$ as the relaxed complementarity condition and applied a "pure Newton" type method to it. This algorithm has excellent numerical properties at the expense of more complex operations in each iteration. Zhang proposed a wider class of primal-dual methods which used $PXSP^{-1} + P^{-T}SXP^{T} = 2\mu I$ as the relaxed complementarity conditions for any nonsingular matrix P. It is easy to verify that choosing $P = I$ yields the $XS + SX$ method, $P = S^{1/2}$ yields the XS method and $P = X^{-1/2}$ yields the SX method. Todd, Toh and Tütüncü in [TTT96] showed that the Nesterov-Todd method is also obtained by a special choice of P (see §4.3). Then Monteiro [Mon96] proved a $\sqrt{n}$ iteration bound on all algorithms of the Zhang family, thus generalizing the Monteiro-Adler proof and many of its extensions to SDP.

In [MT96] Monteiro and Tsuchiya analyze another class related to writing the relaxed complementarity relations as $X^{1/2}SX^{1/2} = \mu I$ and $S^{1/2}XS^{1/2} = \mu I$; they prove polynomial bounds on the number of iterations for various short and long-step algorithms based on these directions. Tseng also proposes a similar class of directions. We do not deal with these methods here. Nonetheless, it should be obvious how to extend such algorithms and their analysis to symmetric cones using the techniques developed here. In particular one can find analogs of these methods for the QCQP problem.

In this paper we extend Monteiro's proof and as a result develop a complete analog of the Zhang family and show that a $\sqrt{r}$ bound on iteration count holds for symmetric cones. We also show how to specialize our general method to the quadratic cone and obtain $\sqrt{r}$ iteration bound for the QCQP algorithm developed by Adler and Alizadeh [AA95] and further studied by the authors in [AS97]; this method is exactly the analog of the $XS + SX$

method for the QCQP problem. Before we proceed we should remind the reader that all cone LP problems over symmetric cones can be expressed by well-known reductions as semidefinite programs over symmetric positive semidefinite matrices with real entries. These reductions do not affect the complexity results achieved for positive semidefinite Hermitian matrices with complex or quaternion entries. Expressing QCQP as a semidefinite program *does* deteriorate the complexity results though, if it is done in the usual way. Let us elaborate. Suppose

$$\mathcal{Q} = \mathcal{Q}_1 \times \mathcal{Q}_2 \times \cdots \times \mathcal{Q}_r$$

and each $\mathcal{Q}_i$ is an (n_i+1)-dimensional quadratic cone. It is known that $x \in \mathcal{Q}_i$ iff the matrix $x_0 I + \bar{x}e_0^T + e_0\bar{x}^T$ is positive semidefinite[2]. Thus every complexity result for SDP directly extends to QCQP by this transformation. However the bounds obtained in this manner are not optimal: This extension implies that the iteration complexity of cone LP over $\mathcal{Q}$ is $\sqrt{n}$ where $n = n_1 + \ldots + n_r$. But the self-concordance theory of Nesterov and Nemirovskii suggests that the iteration complexity should be $\sqrt{r}$. We show a way of embedding the Jordan algebra associated with the quadratic cone in certain associative algebras known as Clifford algebras so that the general complexity results for cone LP over arbitrary symmetric cones extends to the quadratic cone with optimal bound on the iteration count.

In section 2 we build the necessary tools for such analysis. In section 3 we present the extension of Zhang family to symmetric cone optimization and derive the Newton method in this context. Then we present a word-by-word extension of Monteiro's proof of $\sqrt{r}$ bound on the number of iterations. In section 4 we present reductions of the general theory to semidefinite programming over Hermitian positive semidefinite matrices with complex, quaternion or octonion entries. More interestingly we also present the reduction to the quadratic cone by using Clifford algebras as our associative algebra–see [Shi88] for an introduction to this topic. Finally in section 5 we make some concluding remarks.

2 Definitions and Notation

In this section we build the theory of associative algebras and Jordan algebras derived from them. Our development is quite concrete. Since we are concerned only with euclidean algebras we restrict ourselves to subsets of $n \times n$ matrices with real entries. There is no loss in generality as all euclidean algebras considered here have such matrix representation [FK94]. We are concerned with two subsets of matrices.

A set $\mathcal{M}$ of $n \times n$ matrices with real entries together with the usual matrix multiplication forms a *euclidean associative algebra* if

1. $X \in \mathcal{M} \Rightarrow \alpha X \in \mathcal{M} \quad \forall \alpha \in \mathfrak{R}$,
2. $X, Y \in \mathcal{M} \Rightarrow X + Y \in \mathcal{M}$,
3. $X, Y \in \mathcal{M} \Rightarrow XY \in \mathcal{M}$,

[2] Here if $x = (x_0, x_1, \ldots, x_{n_i})$ then $\bar{x} = (0, x_1, \ldots, x_{n_i})$. Also $e_0 = (1, 0, \ldots, 0)$.

4. $X \in \mathcal{M} \Rightarrow X^T \in \mathcal{M}$,
5. $X \in \mathcal{M}$ and X is invertible $\Rightarrow X^{-1} \in \mathcal{M}$

We now define a new binary operation "$\circ$" on the symmetric matrices in $\mathcal{M}$:

$$X \circ Y \stackrel{\text{def}}{=} \frac{1}{2}(XY + YX). \tag{3}$$

A set of symmetric matrices $\mathcal{S}$ in $\mathcal{M}$ together with the "$\circ$" operation forms a *euclidean Jordan algebra derived from* $\mathcal{M}$ if

1. $X \in \mathcal{S} \Rightarrow \alpha X \in \mathcal{S} \quad \forall \alpha \in \mathfrak{R}$,
2. $X, Y \in \mathcal{S} \Rightarrow X + Y \in \mathcal{S}$,
3. $X, Y \in \mathcal{S} \Rightarrow X \circ Y \in \mathcal{S}$,
4. $X \in \mathcal{S}$ and X is invertible $\Rightarrow X^{-1} \in \mathcal{S}$,
5. if $X \in \mathcal{S}$ and there is a symmetric matrix Z such that $X = Z \circ Z$ then $Z \in \mathcal{S}$.

When $X = Z \circ Z$, then X is positive semidefinite, furthermore, there is a unique positive semidefinite Z where $Z \circ Z = X$, in this case we say that $Z = X^{1/2}$.

Remark 1 The sets $\mathcal{M}$ and $\mathcal{S}$ are vector spaces over $\mathfrak{R}$ and therefore are closed. Hence they are closed under taking $f(\cdot)$ for every analytic function f on $\mathcal{M}$ or $\mathcal{S}$, in particular X^{-1} and $X^{1/2}$. Thus requirements 5. in the definition of associative algebras and 4. and 5. of the Jordan algebras are technically redundant.

The positive semidefinite matrices in $\mathcal{S}$ form a cone which we denote by $\mathcal{K}$. $\mathcal{K}^+$ is the set of positive definite matrices in $\mathcal{S}$. The following fact is fundamental and its proof can be found in [FK94].

Proposition 1 *Every symmetric cone $\mathcal{K} \subseteq \mathfrak{R}^n$ is the cone of positive semidefinite elements of some euclidean Jordan Algebra $\mathcal{S}$ which is a subspace of $\mathfrak{R}^n$.*

Define on $\mathcal{M}$ the scalar product

$$\langle X, Y \rangle \stackrel{\text{def}}{=} k \cdot \mathrm{Tr}(XY^T), \tag{4}$$

i.e. $\langle \cdot, \cdot \rangle$ is a multiple of the standard scalar product for matrices. Also define

$$\|X\|_F = \langle X, X \rangle, \ \|X\|_2 = \lambda_1(XX^T).$$

Thus $\|\cdot\|_F$ is a multiple of the standard Frobenius norm. $\|\cdot\|_2$ is the usual 2-norm for matrices. Both norms clearly satisfy the standard norm properties, as well as

$$\|XY\|_F \leq \|X\|_F \cdot \|Y\|_2 . \tag{5}$$

We assume that k is chosen so that $\|X\|_2 \leq \|X\|_F$ for all $X \in M$, hence we have

$$\|XY\|_F \leq \|X\|_F \cdot \|Y\|_F . \tag{6}$$

Remark 2 If $\mathcal{M}$ is the set of all matrices then necessarily $k \geq 1$ so that $\|X\|_2 \leq \|X\|_F$ is satisfied. But for other types of associative algebras it is possible to set k to much smaller numbers and still have this inequality satisfied. This observation leads to lower iteration complexity especially in the case of QCQP.

3 Zhang family and Newton's method

Consider the primal-dual pair of cone-LP problems

$$(P) \qquad \min\{\langle C, X\rangle \mid \langle A_i, X\rangle = b_i, i = 1, \dots, m, X \in \mathcal{K},\} \tag{7}$$

and

$$(D) \qquad \max\{b^T y \mid \sum_{i=1}^{m} y_i A_i + S = C, S \in \mathcal{K}, y \in \mathfrak{R}^m\}, \tag{8}$$

where $\mathcal{K}$ is the symmetric cone associated with $\mathcal{S}$, and $C, A_i \in \mathcal{S}$, for $i = 1, \dots, m$ and $b \in \mathfrak{R}^m$ are given data. We call $X \in \mathcal{K}$ primal feasible if $\langle A_i, X\rangle = b_i$ for $i = 1, \dots, m$. $(S, y) \in \mathcal{K} \times \mathfrak{R}^m$ is called dual feasible if $\sum_{i=1}^m y_i A_i + S = C$. The sets of interior feasible solutions are

$$\mathcal{F}^0(P) \stackrel{\text{def}}{=} \{X \in \mathcal{S} \mid \langle A_i, X\rangle = b_i, i = 1, \dots, m, X \in \mathcal{K}^+\},$$
$$\mathcal{F}^0(D) \stackrel{\text{def}}{=} \{(S, y) \in \mathcal{S} \times \mathfrak{R}^m \mid \sum_{i=1}^{m} y_i A_i + S = C, S \in \mathcal{K}^+\}.$$

The complementary slackness theorem for this pair of cone LP's is given by $XS = 0$ and thus the system

$$\begin{array}{l} \langle A_i, X\rangle = b_i \qquad i = 1, \dots, m \\ \sum_{i=1}^m y_i A_i + S = C \\ XS = 0 \\ X, S \in \mathcal{K}, \quad y \in \mathfrak{R}^m \end{array} \tag{9}$$

completely determines the primal and dual optimal solutions in the absence of various degeneracies [AS97]. The purpose of this section is to show that we can solve this system via Newton's method and obtain polynomial convergence for primal-dual interior-point methods. To this end we establish that Monteiro's proof [Mon96] can be applied to our setting with a slight modification.

The left hand side of the above system, viewed as a function of (X, S, y), maps $\mathcal{S} \times \mathcal{S} \times \mathfrak{R}^m$ into $\mathcal{M} \times \mathcal{S} \times \mathfrak{R}^m$. As these two spaces are different, we get an over-determined system and thus we cannot apply Newton's method directly. We may symmetrize $XS = 0$. In the spirit of Zhang [Zha95] we introduce a *similar symmetrization* operator $H_P : \mathcal{M} \mapsto \mathcal{S}$ defined as

$$H_P(M) \stackrel{\text{def}}{=} \frac{1}{2}\left(PMP^{-1} + (PMP^{-1})^T\right), \forall M \in \mathcal{M}. \tag{10}$$

$P \in \mathcal{M}$ has to be nonsingular and chosen in such a way that PXP^T and $P^{-T}SP^{-1}$ are in $\mathcal{S}$ for $X, S \in \mathcal{S}$. By letting $X = I$ we see that this implies $W = PP^T$ is in $\mathcal{S}$ and by proposition 2.1 in [MZ96] there is no loss of generality in restricting P to $\mathcal{K}^+$.

It has been shown by Zhang [Zha95] that

$$H_P(M) = \tau I \iff M = \tau I,$$

for any nonsingular matrix P, any $\tau \in \mathfrak{R}$ and any matrix M with real spectrum. The last condition is satisfied for $M = XS$ with $X, S \in \mathcal{K}$. As a consequence, (9) is equivalent to the system

$$\begin{array}{ll} \langle A_i, X \rangle = b_i & i = 1, \dots, m \\ \sum_{i=1}^m y_i A_i + S = C & \\ H_P(XS) = 0 & \end{array} \tag{11}$$

to which Newton-type methods can be applied.

If we drop the condition that $X \in \mathcal{K}$ in the primal problem (P) and add the barrier $-\sigma\mu \ln \det X$ to the primal objective, and then apply optimality conditions to the resulting optimization problem we obtain a perturbed system with $H_P(XS) = \sigma\mu I$. Similarly, if the condition $S \in \mathcal{K}$ is dropped from the dual problem (D) and the term $-\sigma\mu \ln \det S$ is added to the dual objective then applying optimality condition results in the same perturbed system, see [AS97]. We now apply the Newton method to this system which leads us to the linear system

$$\begin{array}{ll} \langle A_i, \Delta X \rangle = b_i - \langle A_i, X \rangle & i = 1, \dots, m \\ \sum_{i=1}^m \Delta y_i A_i + \Delta S = C - S - \sum_{i=1}^m y_i A_i & \\ H_P(\Delta X S + X \Delta S) = \sigma\mu I - H_P(XS), & \end{array} \tag{12}$$

where $(\Delta X, \Delta S, \Delta y) \in \mathcal{S} \times \mathcal{S} \times \mathfrak{R}^m$, $\sigma \in [0, 1]$ is a centering parameter and $\mu = \langle X, S \rangle / r$ ($r \stackrel{\text{def}}{=} \langle I, I \rangle$) is the normalized duality gap.

The centrality measure for $(X, S) \in \mathcal{K} \times \mathcal{K}$ is defined as

$$d(X, S) \stackrel{\text{def}}{=} \left\| X^{1/2} S X^{1/2} - \mu I \right\|_F = \sqrt{k \sum_{i=1}^n (\lambda_i(XS) - \mu)^2}. \tag{13}$$

Given a constant $\gamma \in (0, 1)$, we denote by $\mathcal{N}_F(\gamma)$ the following neighborhood of the central path

$$\mathcal{N}_F(\gamma) \stackrel{\text{def}}{=} \{(X, S, y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D) \mid d(X, S) \leq \gamma\mu\}. \tag{14}$$

Below we assume that $(X, S, y) \in \mathcal{F}^0(P) \times \mathcal{F}^0(D)$, $P \in \mathcal{M}$ is a nonsingular matrix satisfying the above conditions, and that $(\Delta X, \Delta S, \Delta y)$ is a solution to (12) for some $\sigma \in [0, 1]$. We adopt Monteiro's notation and define for arbitrary $\alpha \in \mathfrak{R}$:

$$\tilde{X} \stackrel{\text{def}}{=} PXP^T, \qquad \tilde{S} \stackrel{\text{def}}{=} P^{-T}SP^{-1}, \tag{15}$$
$$\widetilde{\Delta X} \stackrel{\text{def}}{=} P\Delta XP^T, \qquad \widetilde{\Delta S} \stackrel{\text{def}}{=} P^{-T}\Delta SP^{-1}, \tag{16}$$
$$X(\alpha) \stackrel{\text{def}}{=} X + \alpha\Delta X, \qquad S(\alpha) \stackrel{\text{def}}{=} S + \alpha\Delta S, \qquad y(a) \stackrel{\text{def}}{=} y + \alpha\Delta y, \tag{17}$$
$$\tilde{X}(\alpha) \stackrel{\text{def}}{=} PX(\alpha)P^T, \qquad \tilde{S}(\alpha) \stackrel{\text{def}}{=} P^{-T}S(\alpha)P^{-1}, \tag{18}$$
$$W_x \stackrel{\text{def}}{=} \tilde{X}^{-1/2}\left[\widetilde{\Delta X}\tilde{S} + \tilde{X}\widetilde{\Delta S} + \tilde{X}\tilde{S} - \sigma\mu I\right]\tilde{X}^{1/2}, \tag{19}$$
$$\mu \stackrel{\text{def}}{=} \frac{\langle X, S\rangle}{r} = \frac{\langle \tilde{X}, \tilde{S}\rangle}{r}, \qquad \mu(\alpha) \stackrel{\text{def}}{=} \frac{\langle X(\alpha), S(\alpha)\rangle}{r} = \frac{\langle \tilde{X}(\alpha), \tilde{S}(\alpha)\rangle}{r}. \tag{20}$$

3.1 Monteiro's proof

Here we restate the technical results of section 3 in [Mon96].

Lemma 1 *Suppose that X and S are in $\mathcal{K}^+$ and P is invertible. Then*

(a) $d(\tilde{X}, \tilde{S}) = d(X, S)$,

(b) $d(X, S) \leq \|H_P(XS - \mu I)\|_F$, *with equality holding if* $PXSP^{-1} \in \mathcal{S}$.

Proof: By lemma 2.1 in [Mon96] the two claims hold for arbitrary symmetric, positive definite matrices X, S and invertible matrices P. ■

Lemma 2 *We have*

$$H_{\tilde{X}^{1/2}}(W_x) = 0, \tag{21}$$
$$\langle \widetilde{\Delta X}, \widetilde{\Delta S}\rangle = \langle \Delta X, \Delta S\rangle = 0. \tag{22}$$

Proof: (21) follows directly from (12) and (19). (22) follows from (12) and the fact that (X, S, y) is primal-dual feasible. ■

Lemma 3 *For every $\alpha \in \mathfrak{R}$, we have*

$$\mu(\alpha) = (1 - \alpha + \sigma\alpha)\mu, \tag{23}$$
$$\tilde{X}^{-1/2}\left[X(\alpha)S(\alpha) - \mu(\alpha)I\right]\tilde{X}^{1/2} = (1-\alpha)(\tilde{X}^{1/2}S\tilde{X}^{1/2} - \mu I) + \alpha W_x + \tag{24}$$
$$+ \alpha^2\tilde{X}^{-1/2}\widetilde{\Delta X}\widetilde{\Delta S}\tilde{X}^{1/2}. \tag{25}$$

Proof: The proof for Lemma 3.4 in [Mon96] adjusted for $\langle I, I\rangle = r$ applies here. ■

Lemma 4 *Let $W \in \mathcal{M}$ be such that $H_Q(W) = 0$ for some nonsingular $Q \in \mathcal{M}$. Then*

$$\|H_Q(W)\|_F \leq \frac{1}{2} \left\|W - W^T\right\|_F, \tag{26}$$

$$\|W\|_F \leq \frac{\sqrt{2}}{2} \left\|W - W^T\right\|_F. \tag{27}$$

In particular, if $W = U_1 + U_2$, for some $U_1 \in \mathcal{S}$ and arbitrary $U_2 \in \mathcal{M}$, then

$$\|W\|_F \leq \sqrt{2} \|U_2\|_F. \tag{28}$$

Proof: Multiplying all inequalities by $1/\sqrt{k}$ we obtain Lemma 3.5 of [Mon96] which is true for arbitrary W and nonsingular Q. ∎

To simplify notation define

$$\Phi_\theta \overset{\text{def}}{=} \tilde{S}^{1/2}\tilde{X}^{1/2} - \theta\mu\tilde{S}^{-1/2}\tilde{X}^{-1/2}.$$

Lemma 5 *For every $\theta \in \mathfrak{R}$, we have*

$$\|W_x\|_F \leq \sqrt{2}\delta_x \|\Phi_\theta\|_F \tag{29}$$

and

$$\begin{aligned} \left\|\tilde{X}^{-1/2}\left[\tilde{X}(\alpha)\tilde{S}(\alpha) - \mu(\alpha)I\right]\tilde{X}^{1/2}\right\|_F \leq & (1-\alpha)d(X,S) + \alpha^2\delta_x\delta_y \\ & + \alpha\sqrt{2}\delta_x \|\Phi_\theta\|_F, \end{aligned} \tag{30}$$

for all $\alpha \in \mathfrak{R}$, where

$$\delta_x \overset{\text{def}}{=} \left\|\tilde{X}^{-1/2}\widetilde{\Delta X}\tilde{S}^{1/2}\right\|_F, \qquad \delta_s \overset{\text{def}}{=} \left\|\tilde{S}^{-1/2}\widetilde{\Delta S}\tilde{X}^{1/2}\right\|_F. \tag{31}$$

Proof: Identical to lemma 3.6 in [Mon96]. ∎

Lemma 6 *If $d(X,S) \leq \gamma\mu$ for some $\gamma \in (0,1)$, then*

$$\left\|\tilde{X}^{-1/2}\tilde{S}^{-1/2}\right\|_2^2 \leq \frac{1}{(1-\gamma)\mu} \tag{32}$$

$$\|\Phi_\theta\|_F^2 \leq \frac{\gamma^2 + (1-\theta)^2 r}{1-\gamma}\mu. \tag{33}$$

Proof: Identical to lemma 3.7 in [Mon96] with two modifications. In (33) we exchange r for n in the original proof. In addition we have to show that $\tilde{X}^{1/2}\tilde{S}X^{1/2} - \mu I$ and I are orthogonal with our new definition of μ.

$$\begin{aligned} \langle \tilde{X}^{1/2}\tilde{S}\tilde{X}^{1/2} - \mu I, I\rangle &= \langle \tilde{X}^{1/2}\tilde{S}\tilde{X}^{1/2}, I\rangle - \mu\langle I, I\rangle \\ &= \langle X, S\rangle - r\mu = 0, \end{aligned}$$

∎

Lemma 7 *If* $(X, S, y) \in \mathcal{N}_F(\gamma)$ *for some* $\gamma > 0$ *satisfying*

$$2\sqrt{2}\frac{\gamma}{1-\gamma} \leq 1, \tag{34}$$

then

$$\max\{\delta_x, \delta_s\} \leq 2\left\|\Phi_\sigma\right\|_F.$$

Proof: Identical to lemma 3.8 in [Mon96]. ∎

Lemma 8 *If* $(X, S, y) \in \mathcal{N}_F(\gamma)$ *for some* $\gamma > 0$ *satisfying* (34), *then for every* $\alpha \in [0, 1]$ *we have*

$$\begin{aligned}&\left\|\tilde{X}^{-1/2}\left[\tilde{X}(\alpha)\tilde{S}(\alpha) - \mu(\alpha)I\right]\tilde{X}^{1/2}\right\|_F \\ &\qquad \leq \left[(1-\alpha)\gamma + 2\sqrt{2}\alpha\frac{\gamma[\gamma^2 + (1-\sigma)^2 r]^{1/2}}{1-\gamma} + 4\alpha^2\frac{\gamma^2 + (1-\sigma)^2 r}{1-\gamma}\right]\mu\end{aligned} \tag{35}$$

Proof: As lemma 3.9 in [Mon96], this is an immediate consequence of the above lemmas. ∎

Theorem 1 *Let* $\gamma, \Gamma \in (0, 1)$ *and* $\delta \in [0, \sqrt{r}]$ *be constants satisfying*

$$\frac{7(\gamma^2 + \delta^2)}{1-\gamma} \leq \left(1 - \frac{\delta}{\sqrt{r}}\right)\Gamma, \qquad \frac{2\sqrt{2}\gamma}{1-\gamma} \leq 1. \tag{36}$$

Suppose that $(X, S, y) \in \mathcal{N}_F(\gamma)$ *and let* $(\Delta X, \Delta S, \Delta y)$ *denote the solution of system* (12), *where*

$$\sigma \stackrel{\text{def}}{=} 1 - \frac{\delta}{\sqrt{r}},$$

and $P \in \mathcal{M}$ *satisfies the conditions defined in section 3. Then,*

(a) $(\hat{X}, \hat{S}, \hat{y}) \stackrel{\text{def}}{=} (X + \Delta X, S + \Delta S, y + \Delta y) \in \mathcal{N}_F(\Gamma)$,

(b) $\langle \hat{X}, \hat{S} \rangle = (1 - \delta/\sqrt{r})\langle X, S \rangle$.

Proof: Identical to theorem 4.1 of [Mon96] replacing r for n. ∎

An immediate consequence of this theorem is that the short-step path following algorithm as described in section 4.1 of [Mon96] has iteration complexity $\mathcal{O}(\sqrt{r})$ for each constant reduction in duality gap. For the predictor-corrector algorithm of section 4.2 of [Mon96], we obtain again iteration complexity $\mathcal{O}(\sqrt{r})$.

4 Application to symmetric cones

4.1 Semidefinite cones of real, complex and quaternion entries

By choosing $\mathcal{M}$ to be the set of all $n \times n$-matrices and $\mathcal{S}$ the set of all real symmetric matrices in $\mathcal{M}$ we see that $\mathcal{K}$ is the cone of positive semidefinite matrices and $(P), (D)$ is the pair of standard primal/dual semidefinite programming problems. Letting $k = 1$ in (4) we obtain $r = n$ and hence regain exactly Monteiro's result in [Mon96].

We can extend this easily to positive semidefinite hermitian matrices:

1. For matrices with complex entries, note that a complex number $x = x_r + ix_i$ can be represented by

$$\begin{pmatrix} x_r & -x_i \\ x_i & x_r \end{pmatrix}.$$

   Hence an $n \times n$-matrix with complex entries can be represented by a $2n \times 2n$-matrix consisting of blocks of the above type. If we take $\mathcal{M}$ to be the set of such matrices and $\mathcal{S}$ the set of symmetric matrices in $\mathcal{M}$, we see that $\mathcal{S}$ represents the hermitian matrices and $\mathcal{K}$ is a representation of the cone of positive semidefinite hermitian matrices. Since for every $X \in \mathcal{S}$, each eigenvalue of X has an even multiplicity, we can choose $k = 1/2$ and still maintain the required norm properties. This way we obtain $r = n$ and the iteration complexity for the above algorithms is $\mathcal{O}(\sqrt{n})$.

2. For matrices with quaternion entries, we can reduce to the previous case using the fact that the quaternion algebra is isomorphic to the algebra of 2×2 complex matrices of the form

$$\begin{pmatrix} x & y \\ -\bar{y} & \bar{x} \end{pmatrix}.$$

   See for example [FK94]. For the reduction to complex matrices we can again choose $k = 1/2$. Together with the previous example we obtain a $4n \times 4n$ real matrix representation for an $n \times n$ quaternion matrix with $k = 1/4$. Hence $r = n$.

Another easy extension is to multiple variables $X_i, S_i,\ i = 1, \dots, l$ each in their own symmetric cone. From the matrix representations $\mathcal{M}_i$ for each block we build a block diagonal matrix

$$M = \begin{pmatrix} M_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & M_l \end{pmatrix}, \quad M_i \in \mathcal{M}_i,$$

Letting $\mathcal{M} = \{M \mid M_i \in \mathcal{M}_i\}$ and defining the scalar product as the sum of the scalar products of each block, we obtain

$$r = \sum_{i=1}^{l} r_i.$$

As an immediate consequence we have that the iteration complexity for quadratically constrained quadratic programming with l blocks is $\mathcal{O}(\sqrt{l})$ for both short-step and predictor-corrector algorithms. But first we should show that one block QCQP has $\mathcal{O}(1)$ iteration complexity under these algorithms.

4.2 Clifford algebra and Embedding the quadratic cone

We now embed the quadratic cone in a Jordan algebra derived from a special kind of associative algebras. Again the construction here is quite concrete. See [Shi88] for a detailed treatment.

Denote by N the set $\{1, \ldots, n\}$.

Definition 1 *The Clifford algebra $\mathcal{C}_n$ is a 2^n-dimensional algebra which is generated by the set of elements $\{e, b_1, b_2, \ldots, b_n\}$ having the property*

$$e^2 = e, \qquad eb_i = b_i e = b_i, \qquad b_i^2 = e, \tag{37}$$

$$b_i b_j = -b_j b_i \quad \textit{for } i \neq j,\ i, j \in N. \tag{38}$$

Its basis consists of the 2^n products

$$b_I \stackrel{\text{def}}{=} e \cdot b_{i_1} \cdots b_{i_r}, \quad 1 \leq i_1 < \cdots < i_r \leq n, I = \bigcup_{j=1}^{r} i_j, I \subseteq N \tag{39}$$

A base element b_I is called *simple* if $|I| \leq 1$. It is called *even* iff $b_I^2 = e$ and *odd* iff $b_I^2 = -e$. We interchangeably write $x \in \mathcal{C}_n$ as

$$x = \sum_{I \subseteq N} x_I b_I \tag{40}$$

or

$$x = (x_e, x_{b_1}, \ldots, x_N). \tag{41}$$

A vector x is called *symmetric* if $x_I \neq 0 \Rightarrow b_I$ even. Similarly x is called simple if $x_I \neq 0 \Rightarrow b_I$ simple. Clearly every simple x is symmetric. Denote by x' the transpose of x in the Clifford algebra sense, i.e. x' is obtained from x by changing the sign for all coefficients of odd base elements. Hence x is symmetric if and only if $x = x'$.

Definition 2 *To each $x \in \mathcal{C}_n$ we associate a $2^n \times 2^n$-matrix $L(x)$ by*

$$L(x)y = xy \quad \forall y \in \mathcal{C}_n. \tag{42}$$

This is the regular matrix representation of the Clifford algebra. Since multiplication in the Clifford algebra is associative, we have

$$L(xy) = L(x)L(y).$$

Definition 3 *For $x \in \mathcal{C}_n$ define the trace of x as*

$$\mathrm{tr}(x) = 2x_e = \frac{1}{2^{n-1}}\mathrm{Tr}(L(x)),$$

where $\mathrm{Tr}(L(x))$ is the usual matrix trace.

For $x, y \in \mathcal{C}_n$ define the scalar product between x and y as

$$\langle x, y\rangle = \mathrm{tr}(xy').$$

Lemma 9 *The matrix $L(x)$ has the following properties:*

(i) $L(e) = I$,

(ii) $L(x)$ *is symmetric except for the sign: We have* $L(x)_{ij} = L(x)_{ji}$ *if* $b_i b_j$ *is even, and* $L(x)_{ij} = -L(x)_{ji}$ *if* $b_i b_j$ *is odd. Hence* x *is symmetric iff* $L(x)$ *is symmetric and* $L(x') = L(x)^T$,

(iii) If $x \neq 0$ *is symmetric then* $L(x)$ *has two eigenvalues* λ_1 *and* λ_2.

$$\lambda_1 = x_e + \|\bar{x}\| \qquad \text{and} \qquad \lambda_2 = x_e - \|\bar{x}\|, \tag{43}$$

where $\bar{x}$ *is* x *with* $x_e = 0$.

We call λ_1 *and* λ_2 *the eigenvalues of* x *and denote them by* $\lambda_1(x), \lambda_2(x)$. *This will also be used in the case where* $L(x)$ *is similar to a symmetric matrix.*

Proof:

(i) Since $L(e)y = ey = y$ for all $y \in \mathcal{C}_n$, the claim follows.

(ii) Let b_k be the unique base element satisfying $b_i = |b_k b_j|$. Without loss of generality assume that $b_i = b_k b_j$. Multiplying both sides by b_k, we obtain

$$b_k b_i = b_k^2 b_j.$$

Hence $L(x)_{ij} = L(x)_{ji} = x_k$ if b_k is even and $L(x)_{ij} = -L(x)_{ji} = x_k$ if b_k is odd.

(iii) Let w be $\bar{x}/(2\,\|\bar{x}\|_2)$, $v_1 = 1/2e + w$, and $v_2 = 1/2e - w$. Observe that $w^2 = 1/4e$. For if we have a term $(w_j w_k)b_j b_k$ that contributes to w_i^2 then there is also $(w_k w_j)b_k b_j = -(w_j w_k)b_j b_k$. Since b_j and b_k are not e the only terms remaining are of the form $w_j^2 b_j^2 = w_j^2 e$. So $w^2 = \|w\|_2^2\, e$ and since the norm of w is $1/2$ by construction, the claim follows.

We have

$$\begin{aligned}
L(x)v_1 &= xv_1 = x(\frac{1}{2}e + w) \\
&= (x_e e + 2\,\|\bar{x}\|_2\, w)(\frac{1}{2}e + w) \\
&= \frac{1}{2}x_e e + \|\bar{x}\|_2\, w + x_e w + 2\,\|\bar{x}\|_2\, w^2 \\
&= (x_e + \|\bar{x}\|_2)(\frac{1}{2}e + w) \\
&= \lambda_1 v_1 \\
L(x)v_2 &= x(\frac{1}{2}e - w) \\
&= \frac{1}{2}x_e e + 2\,\|\bar{x}\|_2\, w - x_e w - 2\,\|\bar{x}\|_2\, w^2 \\
&= (x_e - \|\bar{x}\|_2)(\frac{1}{2}e - w) \\
&= \lambda_2 v_2.
\end{aligned}$$

So v_1 and v_2 are eigenvectors for λ_1 and λ_2. Furthermore $x = \lambda_1 v_1 + \lambda_2 v_2$ and $v_1 + v_2 = e$.

To show that these are the only eigenvalues, we construct a set of eigenvectors for each eigenvalue such that the sets are orthogonal and together span the whole $\mathfrak{R}^{2^n}$. Notice that v_1 and v_2 are orthogonal as $v_1^T v_2 = 1/4 - w^T w = 0$. We can assume that λ_1 and λ_2 are distinct because otherwise x is a multiple of e and hence the only eigenvalue is x_e with multiplicity 2^n.

Let b be an arbitrary vector. Define $v_{b1} = xb - \lambda_2 b$ and $v_{b2} = xb - \lambda_1 b$. Then v_{b1} is an eigenvector for λ_1 and v_{b2} is an eigenvector for λ_2.

$$\begin{aligned}
xv_{b1} &= x^2 b - \lambda_2 x b \\
&= (\lambda_1^2 v_1 + \lambda_2^2)b - \lambda_2(\lambda_1 v_1 + \lambda_2 v_2)b \\
&= \lambda_1(\lambda_1 - \lambda_2)v_1 b \\
&= \lambda_1(\lambda_1 v_1 + \lambda_2 v_2 - \lambda_2(v_1 + v_2))b \\
&= \lambda_1(x - \lambda_2 e)b \\
&= \lambda_1(xb - \lambda_2 b) = \lambda_1 v_{b2} \\
xv_{b2} &= x^2 b - \lambda_1 x b \\
&= (\lambda_1^2 v_1 + \lambda_2^2)b - \lambda_1(\lambda_1 v_1 + \lambda_2 v_2)b \\
&= \lambda_2(\lambda_2 - \lambda_1)v_2 b \\
&= \lambda_2(\lambda_1 v_1 + \lambda_2 v_2 - \lambda_1(v_1 + v_2))b \\
&= \lambda_2(x - \lambda_1 e)b \\
&= \lambda_2(xb - \lambda_1 b) = \lambda_2 v_{b1}
\end{aligned}$$

Choosing for b the 2^n base elements we obtain 2^n eigenvectors for each eigenvalue which together span the whole space since $v_{b_1} - v_{b_2} = (\lambda_1 - \lambda_2)b$. It remains to show that the two sets are orthogonal. So let b and c denote base elements (not necessarily distinct). Then

$$\begin{aligned}
\langle v_{b1}, v_{c2}\rangle &= \langle xb - \lambda_2 b, xc - \lambda_1 c\rangle \\
&= \langle (\frac{1}{2}(\lambda_1 + \lambda_2)e + (\lambda_1 - \lambda_2)w - \lambda_2 e)b, (\frac{1}{2}(\lambda_1 + \lambda_2)e + (\lambda_2 - \lambda_2)w - \lambda_1 e)c\rangle \\
&= \langle (\frac{1}{2}(\lambda_1 - \lambda_2)e + (\lambda_1 - \lambda_2)w)b, (\frac{1}{2}(\lambda_2 - \lambda_1)e + (\lambda_1 - \lambda_2)w)c\rangle \\
&= -(\lambda_1 - \lambda_2)^2 \langle (\frac{1}{2}e + w)b, (\frac{1}{2}e - w)c\rangle \\
&= -(\lambda_1 - \lambda_2)^2 \left[\frac{1}{4}\langle b, c\rangle + \frac{1}{2}\langle wb, c\rangle - \frac{1}{2}\langle b, wc\rangle - \langle wb, wc\rangle\right] \\
&= -(\lambda_1 - \lambda_2)^2 \left[\frac{1}{4}\langle e, cb'\rangle + \frac{1}{2}\langle w, cb'\rangle - \frac{1}{2}\langle w, bc'\rangle - \langle w, wcb'\rangle\right]
\end{aligned}$$

Denote $\alpha = cb'$. Now we have three cases:

(a) $\alpha = \pm e$. Without loss of generality assume $\alpha = e$. Then

$$\begin{aligned}
-\frac{\langle v_{b1}, v_{c2}\rangle}{(\lambda_1 - \lambda_2)^2} &= \frac{1}{4}\langle e, e\rangle + \frac{1}{2}\langle w, e\rangle - \frac{1}{2}\langle w, e\rangle - \langle w, we\rangle \\
&= \frac{1}{4}\langle e, e\rangle - \langle w, w\rangle \\
&= \frac{1}{4} - \frac{1}{4} = 0.
\end{aligned}$$

(b) $\alpha = \alpha', \alpha \neq \pm e$. Then

$$\begin{aligned}
-\frac{\langle v_{b1}, v_{c2}\rangle}{(\lambda_1 - \lambda_2)^2} &= \frac{1}{4}\langle e, \alpha\rangle + \frac{1}{2}\langle w, \alpha\rangle - \frac{1}{2}\langle w, \alpha\rangle - \langle w, w\alpha\rangle \\
&= \frac{1}{4}\langle e, \alpha\rangle - \langle w, w\alpha\rangle \\
&= \frac{1}{2}(\mathrm{tr}(\alpha) - \mathrm{tr}(w\alpha w')) = 0.
\end{aligned}$$

α is a base element and hence it's trace is zero. Also $ww' = w^2 = \frac{1}{4}e$, so $4w' = w^{-1}$. Since the trace is invariant under similarity transformations $\mathrm{tr}(w\alpha w') = 0$.

(c) $\alpha = -\alpha'$. Then

$$-\frac{\langle v_{b1}, v_{c2}\rangle}{(\lambda_1-\lambda_2)^2} = \frac{1}{4}\langle e,\alpha\rangle + \frac{1}{2}\langle w,\alpha\rangle + \frac{1}{2}\langle w,\alpha\rangle - \langle w, w\alpha\rangle$$
$$= \frac{1}{4}\langle e,\alpha\rangle - \langle w, w\alpha\rangle$$
$$= \frac{1}{2}(\mathrm{tr}(\alpha) - \mathrm{tr}(w\alpha w')) = 0.$$

The second equality follows from the fact that α is odd whereas w is symmetric. Hence $\mathrm{tr}(w\alpha') = 0$. The rest is as in the previous case.

■

We call a symmetric $x \in \mathcal{C}_n$ positive definite if $\lambda_i > 0$, positive semidefinite (psd.) if $\lambda_i \geq 0$. If x is psd. we define $x^{1/2}$ as $\sqrt{\lambda_1}v_1 + \sqrt{\lambda_2}v_2$. Clearly $x^{1/2}x^{1/2} = x$. Also $L(x)$ is positive (semi-)definite if x is positive (semi-)definite.

A few observations on $L(x)$ are in order. The set of matrices $L(x), \forall x \in \mathcal{C}_n$ is closed under addition, transposition, scalar and matrix multiplication, and inversion. Since the equation $b_I \cdot x = b_J$ has a unique base element as solution for arbitrary I and J, each $\pm x_I$ occurs exactly 2^n times in $L(x)$. Hence $2x^T y = 2^{1-n}\mathrm{Tr}(L(x)L(y)^T)$. If x and y are simple, then $L(x)L(y) + L(y)L(x)$ is also simple, because the sum is symmetric and can contain only nonzero-coefficients for b_I with $|I| \leq 2$. As b_I is odd for $|I| = 2$ we must have $|I| \leq 1$.

Lemma 10 *Analogous to matrices we call the eigenvalues of xx' the squares of the singular values of x and denote them by $\sigma_1(x)^2$ and $\sigma_2(x)^2$. Now we can define the Frobenius and the euclidean norm as*

$$\|x\|_F \stackrel{\text{def}}{=} \sqrt{\sigma(x)_1^2 + \sigma(x)_2^2}$$
$$\|x\|_2 \stackrel{\text{def}}{=} \sqrt{\sigma(x)_{\max}^2}.$$

These two norms satisfy all conditions for matrix norms and also

$$\|xy\|_2 \leq \|x\|_2 \cdot \|y\|_2$$
$$\|xy\|_F \leq \|x\|_F \cdot \|y\|_2$$
$$\|xy\|_F \leq \|x\|_F \cdot \|y\|_F .$$

Proof: First observe that $\|x\|_F = \|L(x)\|_F / \sqrt{2^{n-1}}$ and $\|x\|_2 = \|L(x)\|_2$. As the general conditions for norms are scaling invariant, we only have to prove the three listed inequalities. The first two are immediate from our observations. The last follows from the second and

$$\|y\|_F = \sqrt{\sigma(y)_1^2 + \sigma(y)^2} \geq \sqrt{\sigma(y)_{\max}} = \|y\|_2 .$$

■

Notice that $\|x\|_F$ is closely related to the standard euclidean vector norm $\|\cdot\|$. The coefficient for e in xx' is $x^T x$, so $\|x\|_F = \sqrt{2x^T x} = \sqrt{2}\|x\|$.

4.3 Iteration complexity for the quadratic cone

The standard quadratic cone programming problem with only one block has the form

$$\begin{array}{ll} \min & \langle c, x\rangle \\ \text{s.t.} & Ax = b \\ & x \in \mathcal{Q} \end{array} . \tag{44}$$

With the results from section 4.2 we see that this is equivalent to

$$\begin{array}{ll} \min & \langle L(c), L(x)\rangle \\ \text{s.t.} & \langle L(a_i), L(x)\rangle = 2^n b_i \quad i = 1, \ldots, m \\ & L(x) \text{ is simple and positive semidefinite} \end{array} , \tag{45}$$

where a_i, c, and x are extended to simple vectors in $\mathcal{C}_n$. Hence we can formulate QCQP in our setting by choosing

$$\mathcal{M} = \{L(x) \mid x \in \mathcal{C}_n\}$$

and

$$\mathcal{S} = \{L(x) \mid x \in \mathcal{C}_n, x \text{ simple}\}.$$

In section 4.2 we have shown that these two sets satisfy all requirements and that we can choose k to be 2^{1-n}. Hence $r = 2$ and we obtain the expected $\mathcal{O}(1)$ iteration complexity.

As was pointed out on page 7 we can assume that $P \in \mathcal{K}^+$. In the case of SDP every such P satisfies $PXP^T \in \mathcal{S}$ and $P^{-T}SP^{-1} \in \mathcal{S}$. It turns out that this is also the case for QCQP. From lemma 9 (iii) we see that if P is simple then so is P^{-1}. In the products PXP^T and $P^{-T}SP^{-1}$ only base elements b_I with nonzero coefficients have $|I| \leq 3$. As both products are clearly symmetric and base elements with $|I| = 2, 3$ are odd, the products are simple. So every member of the Zhang family for SDP has a direct analog in QCQP. Choosing $P = I$ yields the Adler-Alizadeh "pure Newton" direction which is an analog of the $XS+SX$ method in SDP. The XS and SX methods are obtained by choosing $P = S^{1/2}$ and $P = X^{-1/2}$. And the Nesterov-Todd direction uses

$$P = W^{-1/2}, \qquad W = X^{1/2}(X^{1/2}SX^{1/2})^{-1/2}X^{1/2}.$$

The convergence result for the short-step algorithm over quadratic cones with P chosen as in the XS and Nesterov-Todd method has previously been obtained by Tsuchiya [Tsu97],[Tsu98].

A further consideration for QCQP is the computational cost of the operations involving P in an implementation. To obtain iterations that are polynomial time computable, one has to avoid the matrix representation $L(x)$ of an element and work directly with the vector. For the algorithms mentioned above the basic operations are computing UVU for $U, V \in \mathcal{S}$, $H_P(X, S)$ for $P \in \mathcal{S}$, $X, S \in \mathcal{K}$, as well as inversion and square root of elements of $\mathcal{S}$. For the remainder of this section we use the following convention: $u = (u_0, \bar{u}) \in \mathfrak{R}^{n+1}$ is a vector and U is its matrix representation in the Clifford algebra; thus U is a $2^n \times 2^n$ matrix. Similar convention is used for V, X, S, P and $H_P(\cdot)$. By lemma 9 we have that if x is simple, then x^{-1} and $x^{1/2}$ can be computed in linear time. Further, if we can compute uvu efficiently, we can restrict ourselves to $h_e(x,s)$ since we can obtain $h_p(x,s)$ as $h_e(pxp', p'^{-1}sp^{-1})$.

Denote by $P(u)$ the matrix that maps v to uvu. Let $\gamma(x) = x_0^2 - \bar{x}^T\bar{x}$. Then

$$P(u) = \begin{pmatrix} u^Tu & 2u_0\bar{u}^T \\ 2u_0\bar{u} & \gamma(u)I + 2\bar{u}\bar{u}^T \end{pmatrix}.$$

$P(u)$ is known as the quadratic representation of the Jordan algebra represented by $\mathcal{S}$. Another way of writing it is

$$P(u) = 2uu^T + \gamma(u)I - 2\gamma(u)e_0e_0^T, \quad e_0 = (1, 0, \dots, 0)^T.$$

In [Tsu98] the quadratic representation is used to analyze the XS and NT methods for QCQP with a slightly different notation. $G(s)$ is used for $P(s^{1/2})$ and $H(x,s)$ instead of $P(w^{-1/2})$.

As $\gamma(u)$ is computable in linear time, this shows that uvu can also computed in linear time. $h_e(x,s)$ is a well known operator in QCQP and can be computed as

$$h_e(x,s) = \mathrm{Arw}(x)s, \qquad \mathrm{Arw}(x) = x_0 I + e_0\bar{x}^T + \bar{x}e_0^T,$$

which in turn can be computed in linear time. So choosing $P \in \mathcal{S}$ not only guarantees that $H_P(XS) \in \mathcal{S}$ but also that all necessary operations for a primal-dual interior-point algorithm can be implemented efficiently.

5 Conclusion

In this paper we showed that Monteiro's $\sqrt{n}$ bound on iteration count of short-step and predictor-corrector path-following algorithms over the Zhang family of relaxations extends verbatim to optimization over the symmetric cones. This result in fact applies to just about any known algorithm for semidefinite programming. For instance the high order convergence results of Luo, Sturm and Zhang [LSZ96] and Potra and Sheng [PS95] also apply to this more general class. Furthermore algorithms that are not in the Zhang family, such as those of Tseng [Tse96] and Monteiro and Tsuchiya [MT96] can also be extended to symmetric cones and in particular applied to QCQP problems using techniques of this paper. In other words the implications of our approach go beyond Monteiro's analysis; any proof or algorithm in SDP that does not use particular properties not extendable to euclidean associative algebras actually applies to optimization over symmetric cones, and via the Clifford algebra construction to QCQP. Hence it seems more economical to develop interior point (or other algorithms) for symmetric cones (instead of just for the real semidefinite cone) and get similar results for the quadratic cone as an immediate byproduct.

A problem that remains open is whether pure Jordan algebra techniques can be applied to prove polynomiality of the $XS + SX$ or other "pure Newton" methods such $X^{1/2}SX^{1/2}$. If so, such a proof will fill the gap for the 3×3 hermitian matrices on octonions as well.

References

[AA95] I. Adler and F. Alizadeh. Primal-Dual Interior Point Algorithms for Convex Quadratically Constrained and Semidefinite Optimization Problems. Technical Report RRR-111-95, RUTCOR, Rutgers University, 1995.

[AHO96] F. Alizadeh, J. P. Haeberly, and M. L. Overton. Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results. *SIAM J. Optim.*, 1996. To appear.

[Ali95] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM J. Optim.*, 5(1):13–51, 1995.

[AS97] F. Alizadeh and S. Schmieta. Optimization with Semidefinite, Quadratic and Linear Constraints. Technical Report rrr 23-97, RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854-8803, November 1997. Available at URL: `http://rutcor.rutgers.edu/pub/rrr/reports97/23.ps.gz`.

[Fay97a] L. Faybusovich. Euclidean Jordan Algebras and Interior-Point Algorithms. In *Positivity I*, pages 331–357. Kluwer Academic Pubishers, 1997.

[Fay97b] L. Faybusovich. Linear systems in Jordan algebras and primal-Dual interior point algorithms. *Journal of Computational and Applied Mathematics*, 86:149–175, 1997.

[FK94] J. Faraut and A. Korányi. *Analysis on Symmetric Cones.* Oxford University Press, Oxford, UK, 1994.

[HRVW96] C. Helmberg, F. Rendl, R. J. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM J. Optim.*, 6:342–361, 1996.

[KMY89] M. Kojima, S. Mizuni, and A. Yoshise. A primal-dual interior-point algorithm for linear programming. In N. Megiddo, editor, *Progress in Mathematical Programming*, Berlin, 1989. Springer Verlag.

[LSZ96] Z. Q. Luo, J. F. Sturm, and S. Zhang. Superlinear convergence of a symmetric primal-dual path-following algorithm for semidefinite programming. Technical Report 9607, Econometric Institute, Erasmus University, Notterdam, The Netherlands, 1996.

[MA89] R. C. Monteiro and I. Adler. Interior path following primal-dual algorithms. Part I: Linear programming. *Math. Programming*, 44:27–41, 1989.

[Mon96] R. Monteiro. Polynomial Convergence of Primal-Dual Algorithms for Semidefinite Programming Based on Monteiro and Zhang Family of Directions. submitted, 1996.

[Mon97] R. C. Monteiro. Primal-dual path-following algorithms for semidefinite programming. *SIAM J. Optim.*, 7:663–678, 1997.

[MT96] R. Monteiro and T. Tsuchiya. Polynomial Convergence of a New Family of Primal-Dual Algorithms for Semidefinite Programming. Technical Report 627, The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106, Japan, November 1996.

[MZ96] R. D. C. Monteiro and Y. Zhang. A unified analysis for a class of path-following primal-dual interior-point algorithms for semidefinite programming. submitted, 1996.

[NN94] Y. Nesterov and A. Nemirovskii. *Interior Point Polynomial Methods in Convex Programming: Theory and Applications.* Society for Industrial and Applied Mathematics, Philadelphia, 1994.

[NS96] A. Nemirovskii and K. Scheinberg. Extension of Karmarkar's Algorithm onto convex quadratically constrained quadratic programming. *Math. Programming*, 72:273–289, 1996.

[NT94] Y. E. Nesterov and M. J. Todd. Self-Scaled Cones and Interior-Point Methods in Nonlinear Programming. *Math. of Oper. Res.*, April 1994. submitted.

[NT95] Y. E. Nesterov and M. J. Todd. Primal-Dual Interior-Point Methods for Self-Scaled Cones. submitted, 1995.

[PS95] F. A. Potra and R. Sheng. A superlinearly convergent primal-dual infeasible-interior-point algorithm for semidefinite programming. Technical Report 78, Department of Mathematics, The university of Iowa, Iowa City, IA, 1995.

[Sch66] R. D. Schafer. *An Introuduction to Nonassociative Algebras.* Academic Press, New York, 1966.

[Shi88] T. Shimpuku. *Symmetric Algberas by Direct Product of Clifford Algebra.* Seibunsha Publishers, Osaka, Japan, 1988.

[Tse96] P. Tseng. Search directions and convergence analysis of some infeasible path-flowing methods for the monotone semi-definite LCP. Technical Report , Department of Mathemaics, University of Washington, Seatle, WA 98195, 1996.

[Tsu97] Takashi Tsuchiya. A Polynomial Primal-Dual Path-Following Algorithm for Second-Order Cone Programming. Technical Report Research Memorandum No. 649, The Institute of Statistical Mathematics, Tokyo, Japan, 1997.

[Tsu98] Takashi Tsuchiya. A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. Technical Report Research Memorandum No. 664, The Institute of Statistical Mathematics, Tokyo, Japan, 1998.

[TTT96] M. J. Todd, K. C. Toh, and R. H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. Technical Report 829, School of Operations Research and Industrial Engineering, Cornell University, 1996.

[Zha95] Y. Zhang. On extending primal-dual interior-point algorithms from linear programming to semidefinite programming. Technical Report TR 95-20, Dept. of Math/Stat, University of Maryland, Baltimore County, MD 21228, USA, 1995.