CrossMark

An extension of Chubanov's algorithm to symmetric cones

Bruno F. Lourenço[1] · **Tomonari Kitahara**[2] ·
Masakazu Muramatsu[3] · **Takashi Tsuchiya**[4]

Abstract In this work we present an extension of Chubanov's algorithm to the case of homogeneous feasibility problems over a symmetric cone $\mathcal{K}$. As in Chubanov's method for linear feasibility problems, the algorithm consists of a *basic procedure* and a step where the solutions are confined to the intersection of a half-space and $\mathcal{K}$. Following an earlier work by Kitahara and Tsuchiya on second order cone feasibility problems, progress is measured through the volumes of those intersections: when they become sufficiently small, we know it is time to stop. We never have to explicitly compute the volumes, it is only necessary to keep track of the reductions between iterations. We show this is enough to obtain concrete upper bounds to the minimum eigenvalues of a scaled version of the original feasibility problem. Another distinguishing feature of our approach is the usage of a spectral norm that takes into

✉ Bruno F. Lourenço
lourenco@st.seikei.ac.jp

Tomonari Kitahara
kitahara.t.ab@m.titech.ac.jp

Masakazu Muramatsu
muramatu@cs.uec.ac.jp

Takashi Tsuchiya
tsuchiya@grips.ac.jp

[1] Department of Computer and Information Science, Faculty of Science and Technology, Seikei University, 3-3-1 Kichijojikitamachi, Musashino-shi, Tokyo 180-8633, Japan

[2] Graduate School of Decision Science and Technology, Tokyo Institute of Technology, 2-12-1-W9-62, Ookayama, Meguro-ku, Tokyo 152-8552, Japan

[3] Department of Computer Science, The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan

[4] National Graduate Institute for Policy Studies, 7-22-1 Roppongi, Minato-ku, Tokyo 106-8677, Japan

account the way that $\mathcal{K}$ is decomposed as simple cones. In several key cases, including semidefinite programming and second order cone programming, these norms make it possible to obtain better complexity bounds for the basic procedure when compared to a recent approach by Peña and Soheili. Finally, in the appendix, we present a translation of the algorithm to the homogeneous feasibility problem in semidefinite programming.

Keywords Symmetric cone · Feasibility problem · Chubanov's method

Mathematics Subject Classification 90C25 · 90C22 · 65K05

1 Introduction

Consider the following feasibility problem:

$$\begin{aligned} \text{find} \quad & x & \text{(P)} \\ \text{subject to} \quad & \mathcal{A}x = 0 \\ & x \in \text{int}\, \mathcal{K}, \end{aligned}$$

where $\mathcal{K}$ is a *symmetric cone* in a finite dimensional space $\mathcal{E}$ of dimension d, int $\mathcal{K}$ denotes the interior of $\mathcal{K}$ and $\mathcal{A}$ is a linear map. The problem (P) is a special case of *symmetric cone programming* (SCPs), a broad framework that contains linear programming (LP), second order cone programming (SOCP) and semidefinite programming (SDP). In particular, among the symmetric cones, we have the nonnegative orthant $\mathbb{R}_+^n$, the second order cone (Lorentz cone) $\mathcal{L}_n$, the positive semidefinite cone $\mathcal{S}_+^n$ and any direct product of those.

Besides providing an unified framework for several important classes of problems, symmetric cones are very rich in structure, making it possible to use several useful linear algebraic concepts such as eigenvalues in far broader context. Typically, symmetric cone programs are solved via interior point methods [21] and the symmetric cone programming framework is robust enough to encompass primal dual interior point methods [6,19,26], which are the pinnacle of the research on interior point methods.

For the special case of linear programming, there are a few well-known algorithms such as the simplex method, the ellipsoid method [12] and, of course, interior point methods. The last two were the only known methods that had polynomial time complexity, until a recent work by Chubanov [3,4]. Chubanov's algorithm is a new polynomial time algorithm for linear feasibility problems (i.e., $\mathcal{K} = \mathbb{R}_+^n$) with promising computational performance. See also the improved version described by Roos [25] and a related work by by Li, Roos and Terlaky [14]. At this point, it seems that Chubanov's approach does not fit into previous families of methods.

Our initial motivation for this work was an attempt to generalize Chubanov's algorithm to feasibility problems over symmetric cones, following the work by Kitahara and Tsuchiya [13] for second order cone feasibility problems. The algorithm in [13] works by finding a succession of half-spaces that confine the solutions of a scaled

version of (P). These half-spaces are found through a "basic procedure" and then they are adjusted to ensure that the volume of the intersection with $\mathcal{K}$ is sufficiently small. The algorithm progresses by keeping track of the volume reductions and it either proves (P) is feasible/infeasible or proves that a scaled version of (P), if feasible, only has solutions that are "very close" to the boundary of the cone $\mathcal{K}$, where "very close" is adjustable via an ϵ parameter. The scaled problem in [13] makes use of the usual "infinity-norm" on $\mathbb{R}^n$. Here we use what is, perhaps, the proper generalization to the Jordan algebraic setting: see the $\|\cdot\|_{1,\infty}$ spectral norm in Sect. 2.1.

As in [13], the algorithm described here has some flavor of the ellipsoid method [12] since we use volumes to measure progress. It also has some kinship to interior point methods and, in fact, a classical bound for self-concordant functions (Lemma 9) makes an appearance when we prove complexity bounds, see also Theorems 10 and 12. In both theorems, we make use of a self-concordant barrier for $\mathcal{K}$ and we explicitly compute a function φ that bounds the volume reduction. Furthermore, we seek the maximum possible volume reduction that our analysis permits. These aspects of our discussion seem to be novel and we note that self-concordant functions have not appeared in previous works [4,13,23]. We remark that Peña and Soheili [23] also describe methods for solving (P) which are partly inspired by Chubanov's work. Later in this section, we will contrast their work with ours.

We now present some of the main ideas. Every symmetric cone can be written as a direct product of symmetric cones $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_\ell$, where each $\mathcal{K}_i$ is a *simple* symmetric cone contained in some finite dimensional space $\mathcal{E}_i$. We also have $\mathcal{E} = \mathcal{E}_1 \times \cdots \times \mathcal{E}_\ell$. The dimension and rank of each $\mathcal{E}_i$ are denoted by d_i, r_i, respectively, and we define $d = d_1 + \cdots + d_\ell$ and $r = r_1 + \cdots + r_\ell$. See Sect. 2 for a review on the necessary notions on symmetric cones. Associated to (P), we have the problem

$$\begin{aligned} \text{find} \quad & y, u & \text{(D)}\\ \text{subject to} \quad & y = \mathcal{A}^\top u \\ & y \in \mathcal{K}, y \neq 0, \end{aligned}$$

Due to the Gordan–Stiemke's theorem (see Corollary 2 in Luo, Sturm and Zhang [15]), (P) is feasible if and only if (D) is infeasible.

Since multiplying a solution of (P) by a positive constant gives rise to another solution of (P), it makes sense to consider a scaled version such as follows:

$$\begin{aligned} \text{find} \quad & x & \text{(P}^{\mathcal{A}}_{\text{scaled}})\\ \text{subject to} \quad & \mathcal{A}x = 0 \\ & \|x\|_{1,\infty} \leq 1 \\ & x \in \text{int}\,\mathcal{K}, \end{aligned}$$

where $\|x\|_{1,\infty}$ denotes the maximum of the "1-spectral norms" of the blocks of x. That is, we have $\|x\|_{1,\infty} = \max\{\|x_1\|_1, \ldots, \|x_\ell\|_1\}$, where $\|x_i\|_1$ is the sum of the absolute values of the eigenvalues of x_i. See Sect. 2.1 for more details. The usage of $\|\cdot\|_{1,\infty}$ might seem arbitrary at first, but it is the key for obtaining good complexity bounds in one of the phases of our algorithm.

Now, suppose that $\epsilon > 0$ is given. We say that x is an ϵ-*feasible* solution to ($P^{\mathcal{A}}_{scaled}$) if x is feasible for ($P^{\mathcal{A}}_{scaled}$) and $\lambda_{min}(x) \geq \epsilon$, where $\lambda_{min}(x)$ is the minimum eigenvalue of x. The algorithm we describe here accomplishes one of the following three goals:

1. to find a feasible solution to (P),
2. to prove (P) is infeasible by finding a solution to (D),
3. to prove that the minimum eigenvalue of any solution to ($P^{\mathcal{A}}_{scaled}$) must be smaller than ϵ.[1]

Denote the set of feasible solutions of ($P^{\mathcal{A}}_{scaled}$) by $\mathcal{F}^{\mathcal{A}}_{scaled}$. Let $\hat{x} = e/r$, where e, r are the identity element and the rank of $\mathcal{E}$, respectively. The approach we describe here is based on the fact that if $\hat{x}$ is not feasible for (P), we can find in a reasonable time through a so-called "basic procedure" a solution to either (P) or (D) or, failing that, for at least one of the blocks $\mathcal{E}_k$ we obtain a half-space $H \subseteq \mathcal{E}_k$ with very special properties.

This half-space can be seen as a certificate that if $x \in \mathcal{F}^{\mathcal{A}}_{scaled}$, then x_k must be close to the boundary of $\mathcal{K}_k$. Furthermore, we get a concrete bound on $\lambda_{min}(x_k)$, see Lemma 16. This is analogous to the observation that in Chubanov's original algorithm, once the basic procedure finishes without finding solutions to either (P) or (D), there is a least one coordinate k for which we know for sure that $x_k \leq 1/2$ holds for all x that are feasible for a scaled version of (P), see the comments after Lemma 2.1 in [4]. In fact, when $\mathcal{K} = \mathbb{R}^n_+$, the constraint $\|x\|_{1,\infty} \leq 1$ in ($P^{\mathcal{A}}_{scaled}$) forces x to belong to the unit cube $[0, 1]^n$, which is exactly the same type of scaling used in [4]. Furthermore, when $\mathcal{K}$ is a direct product of second order cones, if $x \in \mathcal{K}$ then $\|x\|_{1,\infty}$ is twice the largest component of x and thus become the usual (non-spectral) "infinity-norm" times a positive constant, which is the same norm used in [13] up to a positive constant.

More concretely, after the basic procedure (Algorithm 1) ends, we obtain at least one index k together with elements $w_k, v_k \in \mathcal{E}_k$ which define a half-space $H(w_k, v_k) = \{x_k \in \mathcal{E}_k \mid \langle x_k, w_k \rangle \leq \langle w_k, v_k \rangle\}$ such that the following three key properties hold.

(P.1) If $x \in \mathcal{F}^{\mathcal{A}}_{scaled}$, then $x_k \in H(w_k, v_k)$.
(P.2) There is a linear bijection $Q = Q_1 \oplus \cdots \oplus Q_\ell$ such that $Q(\mathcal{K}) = \mathcal{K}$ and $Q_k(H(e_k, e_k/r_k)) = H(w_k, v_k)$.
(P.3) The volume $H(w_k, v_k) \cap \mathcal{K}_k$ is less than the volume of $H(e_k, e_k/r_k) \cap \mathcal{K}_k$ and their ratio is smaller than a positive constant $\delta < 1$.

Once Q is found, we may consider the problem

$$\begin{aligned} \text{find} \quad & x & (P^{\mathcal{A}Q}_{scaled}) \\ \text{subject to} \quad & \mathcal{A}Qx = 0 \\ & \|x\|_{1,\infty} \leq 1 \\ & x \in int\,\mathcal{K}. \end{aligned}$$

---

[1] As in the case of semidefinite matrices, an element of $x \in int\,\mathcal{K}$ with small minimum eigenvalue is very close to the boundary of $\mathcal{K}$. So, while item 3 does not rule out the possibility of (P) being infeasible, it shows that it is very close to being infeasible.

We will discuss (P.1), (P.2) and (P.3) in Sect. 3, but in a nutshell, we construct Q in such a way that if x is feasible for ($\mathrm{P}^{\mathcal{A}}_{\text{scaled}}$), then $Q^{-1}(x)$ is feasible for ($\mathrm{P}^{\mathcal{A}Q}_{\text{scaled}}$). On the other hand, if x is feasible for ($\mathrm{P}^{\mathcal{A}Q}_{\text{scaled}}$), then $Q(x)$ is a feasible solution to the original problem (P).

Since the system ($\mathrm{P}^{\mathcal{A}Q}_{\text{scaled}}$) has the same shape as (P), we may apply the basic procedure again. Then, (P.3) ensures that we are making progress towards confining the k-th block of the feasible region of ($\mathrm{P}^{\mathcal{A}}_{\text{scaled}}$) to a region inside of $\mathcal{K}_k$ with smaller volume. As in Chubanov's algorithm, when $\mathcal{K}_k$ is the half-line $\mathbb{R}_+$, (P.3) ensures that the x_k is contained in an interval $[0, \delta]$, with $\delta < 1$.

The main complexity result are given in Proposition 13 and Theorem 17. Our version of the basic procedure (Algorithm 1) takes at most $\ell^3 r_{\max}^2$ iterations, where $r_{\max} = \max\{r_1, \ldots, r_\ell\}$. The main algorithm (Algorithm 2) takes at most $\frac{r}{\varphi(2)} \log\left(\frac{1}{\epsilon}\right)$ iterations, where $\varphi(2)$ is a positive constant as in Theorems 10 and 12.

Here we highlight some of the keys aspects of our analysis while contrasting with the approach by Peña and Soheili [23].

1. Our analysis take into consideration the block division of $\mathcal{K}$. This is expressed, in part, by our usage of the $\|\cdot\|_{1,\infty}$ spectral norm, which we believe is novel in this context and has advantages even when $\ell = 1$, when it becomes the $\|\cdot\|_1$ spectral norm. As such, the complexity of the basic procedure (Algorithm 1) and the overall complexity is expressed in terms of ℓ and $r_{\max} = \max\{r_1, \ldots, r_\ell\}$. Our version of the basic procedure performs at most $\mathcal{O}(\ell^3 r_{\max}^2)$ iterations (Proposition 13) and is analogous to the "Von Neumann scheme" discussed in [23], which, as the authors remark, when specialized to $\mathcal{K} = \mathbb{R}_+^n$ gives essentially the same basic procedure described by Chubanov. Their version using the "Von Neumann scheme" produces a basic procedure which performs at most $\mathcal{O}(r^4)$ iterations, where $r = r_1 + \cdots + r_\ell$. Although not always, there are at least a few cases of interest where $\ell^3 r_{\max}^2$ is a better bound than r^4.
   Consider, for instance, the case where the rank of the blocks are equal or close to being equal. If we have product of ℓ second order cones, we get a complexity of $\mathcal{O}(\ell^3)$ vs $\mathcal{O}(\ell^4)$. The same bounds hold for the case $\mathcal{K} = \mathbb{R}_+^n$. In the specific case of semidefinite programming over the $n \times n$ matrices, we get $\mathcal{O}(n^2)$ vs $\mathcal{O}(n^4)$. One case for which the bound in [23] is better is when $\ell = 30$, $r_{\max} = 21$ and $r = 50$.
   In [23], the authors also suggest the usage of the so-called "smooth perceptron" [27] which lowers the number of iterations to $\mathcal{O}(r^2)$, although each iteration is more expensive. The approach we describe here can easily be adapted to use the smooth perceptron, thus yielding a basic procedure requiring no more than $\mathcal{O}(\ell^{3/2} r_{\max})$ iterations, which still has better performance in the cases we mentioned, see the remarks after Proposition 14.
2. The progress of the algorithm in [23] is measured through the quantity $\chi = \max\{\det x \mid \mathcal{A}x = 0, x \in \text{int}\,\mathcal{K}, \|x\|^2 = r\}$. The performance of the approach in [23] also depends on χ and this leads to a very elegant complexity analysis. Nevertheless, we think some of the geometric intuition is lost by using χ. We hope to recover that by painting an intuitive picture in terms of volumes and half-spaces.

Furthermore, the algorithm in [23] runs until a solution to (D) or (P) is found, which happens in a finite number of iterations when (P) is feasible. We remark that it is possible that the algorithm in [23] loops infinitely and this is also true for Algorithm 2 of [23], see the comments after Thereom 1 therein. Clearly, in a practical implementation, we also have to introduce some stopping criteria that depends, perhaps, on the machine precision, since it is meaningless to continue computations after the intermediate numbers get smaller than a certain threshold. In this sense, it is of great interest to know what could be said about the eigenvalues of the solutions of (P) or of some scaled version of it, when the algorithm in [23] is stopped prematurely. In contrast, at each iteration of Algorithm 2 we get concrete upper bounds on $\lambda_{\min}(x)$, for $x \in \mathcal{F}_{\text{scaled}}^{\mathcal{A}}$, see Lemma 16. This is possible, in part, due to the volumetric considerations in Proposition 8. We believe a similar discussion would have been useful in [23]. We should remark however, that Theorem 3 in [23] implies that after k iterations of their Algorithm 4, the bound $\chi \leq 1.5^{-k}$ holds.

3. At the end of the basic procedure in [23], a vector z is obtained and only the component with largest eigenvalue takes a role in the actual rescaling of the problem, see Step 4 in Algorithm 4 in [23]. Furthermore, a fixed step size is used throughout the algorithm (the parameter a in Algorithm 4 in [23]).

   In our approach, after the basic procedure ends, we obtain a vector y but do not discard the smaller eigenvalues. We then check for each block $\mathcal{K}_k$, whether y_k affords a volume reduction for that block and if it does, we take the maximum possible step size our analysis permits, see Theorems 10 and 12. This corresponds to Lines 6 and 7 in Algorithm 2 and this behavior also sets our approach apart from the algorithm in [13]. A graph relating the step size ρ and the volume reduction can be seen in Figure 1.

   This article is divided as follows. On Sect. 2 we review some notions related to symmetric cones. On Sect. 3 we discuss several notions regarding volumes in symmetric cones, in particular, we discuss how to confine the blocks of ($\text{P}_{\text{scaled}}^{\mathcal{A}}$) to a region with smaller volume starting from an appropriate y vector. On Sect. 4, we discuss the basic procedure and on Sect. 5 we discuss the main algorithm. Section 6 concludes this paper. In Appendix A, for ease of reference, we restate Algorithms 1 and 2 for the case of semidefinite programming.

2 Preliminary considerations

Here, we review some aspects of the theory of symmetric cones and Euclidean Jordan algebras. More details can be found in the book by Faraut and Korányi [5] and also in the survey article by Faybusovich [8]. Let $\mathcal{E}$ be a finite dimensional Euclidean space equipped with an inner product $\langle \cdot, \cdot \rangle$, and $\circ : \mathcal{E} \times \mathcal{E} \to \mathcal{E}$ be a bilinear map. We say that $(\mathcal{E}, \circ)$ is an Euclidean Jordan algebra if the following properties are satisfied.

1. $y \circ z = z \circ y$,
2. $y \circ (y^2 \circ z) = y^2 \circ (y \circ z)$, where $y^2 = y \circ y$,
3. $\langle y \circ z, w \rangle = \langle y, z \circ w \rangle$,

for all $y, w, z \in \mathcal{E}$. All symmetric cones arise as the cone of squares of some Euclidean Jordan algebra. That is, for a symmetric cone $\mathcal{K}$, we have $\mathcal{K} = \{x \circ x \mid x \in \mathcal{E}\}$, for some Euclidean Jordan algebra $(\mathcal{E}, \circ)$. See Theorems III.2.1 and III.3.1 in [5], for more details. Furthermore, we can assume that $\mathcal{E}$ has an (unique) identity element e satisfying $y \circ e = y$, for all $y \in \mathcal{E}$.

In what follows, we say that c is *idempotent* if $c \circ c = c$. Moreover, c is *primitive* if it is nonzero and there is no way of writing $c = a + b$ with a and b nonzero idempotent elements satisfying $a \circ b = 0$.

Theorem 1 (Spectral theorem, see Theorem III.1.2 in [5]) *Let $(\mathcal{E}, \circ)$ be an Euclidean Jordan Algebra and let $x \in \mathcal{E}$. Then there are:*

(i) *primitive idempotents $c_1, \ldots, c_r$ satisfying*

$$c_i \circ c_j = 0 \qquad \text{for } i \neq j, \tag{1}$$
$$c_i \circ c_i = c_i, \qquad i = 1, \ldots, r, \tag{2}$$
$$c_1 + \cdots + c_r = e, \qquad i = 1, \ldots, r, \tag{3}$$

(ii) *unique real numbers $\lambda_1, \ldots, \lambda_r$ satisfying*

$$x = \sum_{i=1}^{r} \lambda_i c_i. \tag{4}$$

Moreover, r only depends on $\mathcal{E}$.

We say that the $c_1, \ldots, c_r$ in Theorem 1 form a *Jordan Frame* for x. The $\lambda_1, \ldots, \lambda_r$ are the *eigenvalues* of x. We will denote the maximum eigenvalue of x by $\lambda_{\max}(x)$ and the minimum by $\lambda_{\min}(x)$. The r appearing in the theorem is called the *rank* of $\mathcal{E}$ and depends only on the algebra $\mathcal{E}$. The rank of x is the number of nonzero eigenvalues. Given $x \in \mathcal{E}$, we define its trace by

$$\operatorname{tr} x = \lambda_1 + \cdots + \lambda_r,$$

where $\lambda_1, \cdots, \lambda_r$ are the eigenvalues of x. The trace function is linear and can be used to define an inner product for $\mathcal{E}$. So, henceforth, we will assume that the inner product $\langle \cdot, \cdot \rangle$ is defined as follows.

$$\langle x, y \rangle = \operatorname{tr}(x \circ y).$$

When $\mathcal{E}$ is the space of $n \times n$ symmetric matrices, this inner product becomes the usual Frobenius inner product. Note that if c is a primitive idempotent, we have $\langle c, c \rangle = 1$. Furthermore, we have $\langle e, e \rangle = r$. We also have a determinant function defined as

$$\det x = \lambda_1 \times \cdots \times \lambda_r.$$

If the rank of x is equal to r, it admits an inverse element denoted by x^{-1}. We have $x^{-1} = \sum_{i=1}^{r} \lambda_i^{-1} c_i$. We then have $x \circ x^{-1} = e$.

We recall the following properties of $\mathcal{K}$. The results follows from various propositions that appear in [5], such as Proposition III.2.2 and Exercise 3 in Chapter III. See also Equation (10) in [28].

Proposition 2 *Let $x \in \mathcal{E}$.*

(i) $x \in \mathcal{K}$ *if and only if the eigenvalues of x are nonnegative.*
(ii) $x \in \text{int}\, \mathcal{K}$ *if and only if the eigenvalues of x are positive.*

It follows that if $x \in \mathcal{K}$, it admits a square root defined as $\sqrt{x} = \sum_{i=1}^{r} \sqrt{\lambda_i} c_i$. With that, we have $\sqrt{x} \circ \sqrt{x} = x$. We will also write $x^{1/2}, x^{-1/2}$ for $\sqrt{x}$ and $\sqrt{x^{-1}}$, respectively.

An Euclidean Jordan algebra $(\mathcal{E}, \circ)$ is said to be *simple* if it is not possible to decompose it as an orthogonal direct sum $\mathcal{E} = \mathcal{E}_1 \oplus \mathcal{E}_2$ with $\mathcal{E}_1$ and $\mathcal{E}_2$ being themselves nonzero Euclidean Jordan algebras. When $\mathcal{K}$ is the cone of squares of a simple Euclidean Jordan algebras, we will also say that it is a *simple cone*. In this paper, we assume the following decomposition

$$\mathcal{E} = \mathcal{E}_1 \oplus \ldots \oplus \mathcal{E}_\ell$$
$$\mathcal{K} = \mathcal{K}_1 \oplus \ldots \oplus \mathcal{K}_\ell,$$

where the $\mathcal{E}_i$ are simple Euclidean Jordan Algebras of rank r_i and $\mathcal{K}_i$ is the cone of squares of $\mathcal{E}_i$. The dimension of each $\mathcal{E}_i$ is denoted by d_i and we have $d = d_1 + \cdots + d_\ell$ and $r = r_1 + \cdots + r_\ell$. We also have $e = (e_1, \ldots, e_\ell)$, where e_i is the identity in $\mathcal{E}_i$. Note that orthogonality expressed by this decomposition is not only with respect the inner product $\langle \cdot, \cdot \rangle$ but also with respect the Jordan Product $\circ$, meaning that $x_i \circ y_j = 0$ if $x_i \in \mathcal{E}_i, y_j \in \mathcal{E}_j$ with $i \neq j$.

If $x \in \mathcal{E}$, we write x_i for the corresponding i-th block of x. That is, we have $x = (x_1, \ldots, x_\ell)$, where each $x_i \in \mathcal{E}_i$. With this decomposition, the theory described so far can be applied in a blockwise fashion. So, we have, for instance, $x \circ y = (x_1 \circ y_1, \ldots, x_\ell \circ y_\ell)$, $\text{tr}\, x = \sum_{i=1}^{\ell} \text{tr}\, x_i$, $\det x = \prod_{i=1}^{\ell} \det x_i$ and $\lambda_{\min}(x) = \min\{\lambda_{\min}(x_1), \ldots, \lambda_{\min}(x_\ell)\}$.

We also have a so-called *quadratic representation* Q_x defined as the linear map such that

$$Q_x(y) = 2x \circ (x \circ y) - x^2 \circ y.$$

We also have $Q_x(y) = (Q_{x_1}(y_1), \ldots, Q_{x_\ell}(y_\ell))$. Denote by I_k the identity map on $\mathcal{E}_k$, we have $Q_{e_k} = I_k$. The quadratic representation has the following key properties, see Section 1.2 in [28] and Section II.3 together with Proposition III.4.2 in [5].

Proposition 3 *Suppose $x \in \text{int}\, \mathcal{K}$*

(i) Q_x *is a self-adjoint linear map on $\mathcal{E}$ such that $Q_x(\mathcal{K}) = \mathcal{K}$.*
(ii) $Q_x^{-1} = Q_{x^{-1}}$, $Q_x(e) = x^2$, $Q_x(x^{-1}) = x$, $Q_{ax} = a^2 Q_x$ *for $a \in \mathbb{R}$.*
(iii) *if $\mathcal{K}$ is simple then $\det Q_x = (\det x)^{\frac{2d}{r}}$.*

We remark that since Q_x is a self-adjoint linear map over the linear space $\mathcal{E}$, all its eigenvalues are real. In item (iii), the determinant of Q_x is, of course, the product of the eigenvalues of Q_x. Another way of visualizing the determinant is recalling that after identifying $\mathcal{E}$ with some $\mathbb{R}^d$, Q_x is representable as a square matrix S, so the determinant that appears in item (iii) refers to the determinant of that S. See Section 1.2 in the work by Sturm [28] for the relations between eigenvalues of Q_x and x.

2.1 Spectral norms for general Jordan algebras

In this paper, a few different norms will come into play.[2] We will write $\|\cdot\|$ or $\|\cdot\|_2$ for the norm induced by $\langle \cdot, \cdot \rangle$. We have $\|x\| = \sqrt{\mathrm{tr}\,(x^2)} = \sqrt{\lambda_1^2 + \cdots + \lambda_r^2}$. We will also define $\|x\|_1 = |\lambda_1| + \cdots + |\lambda_r|$ and $\|x\|_\infty = \max\{|\lambda_1|, \ldots, |\lambda_r|\} = \max\{\lambda_{\max}(x), -\lambda_{\min}(x)\}$. Note that if $x \in \mathcal{K}$, we have $\|x\|_1 = \langle x, e \rangle$ and $\|x\|_\infty = \lambda_{\max}(x)$.

Recall that $\mathcal{E}$ is not necessarily a simple algebra, but admits a decomposition $\mathcal{E} = \mathcal{E}_1 \oplus \ldots \oplus \mathcal{E}_\ell$, where each $\mathcal{E}_i$ is indeed simple. In this work, we will consider norms that take into account the way that $\mathcal{E}$ is decomposed as simple algebras. We will use $\|x\|_{1,\infty}$ to denote the maximum among the 1-norms of the blocks. That is

$$\|x\|_{1,\infty} = \max\{\|x_1\|_1, \ldots, \|x_\ell\|_1\}.$$

So that when $\ell = 1$, we have $\|x\|_{1,\infty} = \|x\|_1$. Furthermore, when $x \in \mathcal{K}$, we have

$$\|x\|_{1,\infty} = \max\{\langle x_1, e_1 \rangle, \ldots, \langle x_\ell, e_\ell \rangle\}.$$

We also define $\|x\|_{\infty,1} = \|x_1\|_\infty + \cdots + \|x_\ell\|_\infty$.

Recall that if $g(\cdot)$ is an arbitrary norm on $\mathcal{E}$, we define its conjugate norm by $g(x)^* = \sup\{\langle x, y \rangle \mid g(y) \leq 1\}$. With that, we have the generalized Cauchy–Schwarz inequality $\langle x, y \rangle \leq g(x)g(y)^*$. We then have the following result, which is unsurprising but requires a careful proof.

Proposition 4 *Let $y \in \mathcal{E}$. Then*

(i) $\|y\|_1^* = \|y\|_\infty$,
(ii) $\|y\|_{1,\infty}^* = \|y\|_{\infty,1}$,
(iii) $\|y\|_{\infty,1} \leq \sqrt{\ell}\, \|y\|_2$.

Proof (i) By definition, we have

$$\|y\|_1^* = \sup\{\langle y, z \rangle \mid \|z\|_1 \leq 1\}.$$

Let $z \in \mathcal{E}$ write its Jordan decomposition as $z = \sum_{i=1}^r \lambda_i c_i$. Note that the condition $\|z\|_1 \leq 1$ means that $|\lambda_1| + \cdots + |\lambda_r| \leq 1$. Let $\mathcal{I}$ denote the set of primitive idempotents in $\mathcal{E}$. We have

---

$$\|y\|_1^* = \sup_{c_1,\ldots,c_r \in \mathcal{I}} \sup_{|\lambda_1|+\cdots+|\lambda_r| \leq 1} \sum_{i=1}^r \lambda_i \langle y, c_i \rangle$$

$$= \sup_{c_1,\ldots,c_r \in \mathcal{I}} \max\{|\langle y, c_1 \rangle|, \ldots, |\langle y, c_r \rangle|\}$$

$$= \sup_{c \in \mathcal{I}} |\langle y, c \rangle|,$$

where the second equality follows from the fact that $\sum_{i=1}^r \lambda_i \langle y, c_i \rangle$ is the Euclidean inner product between the vector $(\lambda_1, \ldots, \lambda_r)$ and $(\langle y, c_1 \rangle, \ldots, \langle y, c_r \rangle)$ in the $\mathbb{R}^r$ space. In $\mathbb{R}^r$, we know that the usual 1-norm and the usual ∞-norm are conjugated pairs and this is why the equality follows.

Now, $\sup_{c \in \mathcal{I}} |\langle y, c \rangle| = \sup_{c \in \mathcal{I}} \max\{\langle y, c \rangle, -\langle y, c \rangle\}$ holds. Then, a result from Hirzebruch implies that $\sup_{c \in \mathcal{I}} \langle y, c \rangle = \lambda_{\max}(y)$ and $\sup_{c \in \mathcal{I}} \langle -y, c \rangle = -\lambda_{\min}(y)$. See either *Satz* 2.3 in [10], Exercise 4 in Chapter 4 of [5] or Equation (9) in [28]. This shows that $\|y\|_1^* = \|y\|_\infty$.

(ii) By definition, we have

$$\|y\|_{1,\infty}^* = \sup\left\{\sum_{i=1}^\ell \langle y_i, z_i \rangle \mid \|z\|_{1,\infty} \leq 1\right\}.$$

Note that if $\|z\|_{1,\infty} \leq 1$, then $\|z_i\|_1 \leq 1$, for all i. By the generalized Cauchy–Schwarz inequality, we have $\langle y_i, z_i \rangle \leq \|y_i\|_\infty \|z_i\|_1 \leq \|y_i\|_\infty$. Therefore, $\|y\|_{1,\infty}^* \leq \|y\|_{\infty,1}$. To show that $\|y\|_{1,\infty}^* = \|y\|_{\infty,1}$, it is enough to construct z such that $\langle y, z \rangle = \|y\|_{\infty,1}$ and $\|z\|_{1,\infty} \leq 1$. We construct z in a block-wise fashion. First, let $y_i = \sum_{j=1}^{r_i} \lambda_j c_j$ be the Jordan decomposition of $y_i \in \mathcal{E}_i$, where $\lambda_1 \leq \cdots \leq \lambda_{r_i}$. Then $\|y_i\|_\infty = \max\{-\lambda_1, \lambda_{r_i}\}$. We then let z_i be either $c_1, -c_1$ or c_{r_i}, so that $\langle y_i, z_i \rangle = \|y_i\|_\infty$ and $\|z_i\|_1 = 1$. If we construct the blocks of z in this way we have $\langle y, z \rangle = \|y\|_{\infty,1}$ and $\|z\|_{1,\infty} \leq 1$.

(iii) Consider the vector $v = (\|y_1\|_\infty, \ldots, \|y_\ell\|_\infty) \in \mathbb{R}^\ell$. Due to the Cauchy–Schwarz inequality we have

$$\|y\|_{\infty,1} = \langle v, (1, \ldots, 1) \rangle$$

$$\leq \sqrt{\ell}\sqrt{\|y_1\|_\infty^2 + \cdots + \|y_\ell\|_\infty^2}$$

$$\leq \sqrt{\ell}\sqrt{\|y_1\|_2^2 + \cdots + \|y_\ell\|_2^2}$$

$$= \sqrt{\ell}\,\|y\|_2.$$

$\square$

We have the following corollary.

Corollary 5 *Let $x, y \in \mathcal{E}$. The following hold:*

$$\langle y, x \rangle \leq \|y\|_1 \|x\|_\infty \tag{5}$$

$$\langle y, x \rangle \le \|y\|_{1,\infty} \|x\|_{\infty,1} \tag{6}$$

2.2 Examples

There is a classification of the finite simple symmetric cones and here we briefly review some of the more widely used examples. When $\mathcal{E} = \mathcal{S}^n$ is the space of $n \times n$ symmetric matrices we can define $x \circ y$ as $\frac{xy+yx}{2}$. With that, $\mathcal{K}$ is the (simple) cone $\mathcal{S}^n_+$ of $n \times n$ positive semidefinite matrices, which has rank n. We have that $\|\cdot\|$ becomes the Frobenius norm. In this case, the eigenvalues as discussed in Theorem 1 are exactly the same eigenvalues that appear in classical linear algebra. The primitive idempotents correspond to rank 1 matrices with norm 1. So Theorem 1 expresses the fact that any symmetric matrix can be written as linear combination of mutually orthogonal rank 1 matrices. Finally, the quadratic map Q_x is such that $Q_x(y) = xyx$. For $x \in \mathcal{K}$, we have $\|x\|_1 = \operatorname{tr} x$ and $\|x\|_\infty = \lambda_{\max}(x)$, where tr is the usual matrix trace.

Another interesting case is when $\mathcal{E} = \mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{n-1}$ and $\mathbb{R}^{n-1}$ is equipped the usual Euclidean inner product. Let $x \in \mathcal{E}$, we decompose it as $x = (x_1, \overline{x})$, where $x_1 \in \mathbb{R}$ and $\overline{x} \in \mathbb{R}^{n-1}$. We then define $x \circ y = (x_1 y_1 + \langle \overline{x}, \overline{y} \rangle, x_1 \overline{y} + y_1 \overline{x})$. With that, $\mathcal{K}$ becomes the Lorentz cone $\mathcal{L}_n = \{x \in \mathcal{E} \mid x_1^2 \ge x_2^2 + \cdots + x_n^2, x_1 \ge 0\}$. Note that the identity element is $e = (1, 0, \ldots, 0)$. Given $x \in \mathcal{E}$, its eigenvalues and corresponding idempotents are

$$\lambda_1 = x_1 + \|\overline{x}\| \qquad \lambda_2 = x_1 - \|\overline{x}\| \tag{7}$$

$$c_1 = \begin{cases} \frac{1}{2}(1, z) & \|\overline{x}\| = 0 \\ \frac{1}{2}\left(1, \frac{\overline{x}}{\|\overline{x}\|}\right) & \|\overline{x}\| \ne 0 \end{cases} \qquad c_2 = \begin{cases} \frac{1}{2}(1, -z) & \|\overline{x}\| = 0 \\ \frac{1}{2}\left(1, -\frac{\overline{x}}{\|\overline{x}\|}\right) & \|\overline{x}\| \ne 0, \end{cases} \tag{8}$$

where $z \in \mathbb{R}^{n-1}$ is any vector satisfying $\|z\| = 1$. See Example 1 in [8], Section 2 in [31], Section 2 and Proposition 1 in [18] or Equations (23) and (24) in [1] for more details. We have $\det x = x_1^2 - \|\overline{x}\|^2$ and $\operatorname{tr}(x) = 2x_1$. Note that the inner product $\langle x, y \rangle = \operatorname{tr}(x \circ y)$ is twice the inner product on $\mathbb{R}^n$, that is, $\langle x, y \rangle = 2(x_1 y_1 + \cdots + x_n y_n)$. Note that if $x \in \mathcal{K}$, then $\|x\|_1 = 2x_1$ and $\|x\|_\infty = x_1 + \|\overline{x}\|$.

The quadratic representation is given by

$$Q_x = \begin{pmatrix} \|x\|^2 & 2x_1 \overline{x}^\top \\ 2x_1 \overline{x} & \det x \, I_{n-1} + 2\overline{x}\overline{x}^\top \end{pmatrix},$$

where I_{n-1} is the $(n-1) \times (n-1)$ identity matrix and $\overline{x}^\top$ indicates the transpose of the column vector $\overline{x}$. The rank of $\mathcal{L}_n$ is 2.

In second order cone programming, it is common to consider a direct product of ℓ Lorentz cones in which case the rank of $\mathcal{K}$ is 2ℓ. If $x \in \mathcal{E}$, we have $x = (x_1, \ldots, x_\ell)$ and every x_i can be written as $x_i = (x_{i1}, x_{i2})$ with $x_{i1} \in \mathbb{R}$, $x_{i2} \in \mathbb{R}^{d_i-1}$. In this case, if $x \in \mathcal{K}$, then $\|x\|_{1,\infty} = \max\{2x_{11}, \ldots, 2x_{\ell 1}\}$. The condition $x_i \in \mathcal{K}_i$ forces, in particular, x_{i1} to be at least as large as the other components of x_i. This means that $\|x\|_{1,\infty}$ is *twice* the coordinate of x with largest value, and, therefore, is twice the usual (non-spectral) infinity norm, for elements inside the cone $\mathcal{K}$. In particular, for

the case of second order cones, the scaled problem ($\text{P}^{\mathcal{A}}_{\text{scaled}}$) is essentially equivalent to the one considered in [13].

Finally, we remark that $\mathcal{K} = \mathbb{R}^n_+$ is not a simple cone. The Jordan-algebraic way of analyzing it is as a direct product $\mathbb{R}_+ \times \ldots \times \mathbb{R}_+$. Note that $\mathbb{R}_+$ is a simple cone of rank 1. If we take $\mathcal{E} = \mathbb{R}$ and define $x \circ y = (x_1 y_1, \ldots, x_n y_n)$, the corresponding cone of squares is $\mathbb{R}^n_+$. We have $\text{tr}\,(x) = x_1 + \cdots + x_n$, $\det x = x_1 \times \cdots \times x_n$ and $e = (1, \ldots, 1)$. The quadratic representation is given by

$$Q_x = \begin{pmatrix} x_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & x_n^2 \end{pmatrix}.$$

In this case, the eigenvalues of x are its components, so that $\|x\|_\infty = \|x\|_{1,\infty} = \max |x_i|$ and $\|x\|_1 = \|x\|_{\infty,1} = |x_1| + \cdots + |x_n|$.

2.3 Remarks on notation

For a matrix $\mathcal{A}$, we will write $\ker \mathcal{A}$ for its kernel and $\mathcal{A}^*$ for its adjoint. If $x \in \mathcal{E}$, we write x_k for the corresponding block belonging to $\mathcal{E}_k$. We will also write $w_k \in \mathcal{E}_k$ for the elements of $\mathcal{E}_k$, without necessarily assuming that w_k is just a block of some element $w \in \mathcal{E}$. This will appear in Theorem 12 and in Algorithm 2, in order to emphasize the set to which the elements belong.

With this notation in mind, for $w_k, v_k \in \mathcal{E}_k$, we define the half-space $H(w_k, v_k) := \{x_k \in \mathcal{E}_k \mid \langle x_k, w_k \rangle \leq \langle w_k, v_k \rangle\}$. We also write $\text{vol}\,(w_k, v_k)$ for the volume of the region $H(w_k, v_k) \cap \mathcal{K}_k$, which corresponds to the integral $\int_{H(w_k, v_k) \cap \mathcal{K}_k} 1$.

As we are using subindexes to denote the blocks, we will use superindexes to indicate the iteration number in Algorithms 1 and 2, which should not be confused with exponentiation.

3 Volumetric considerations

Let $r_{\max} = \max\{r_1, \ldots, r_\ell\}$. The algorithm we discuss in this work depends on having at each iteration access to an element $y \in \mathcal{K}$ satisfying

$$\|P_{\mathcal{A}} y\| \leq \frac{1}{2 r_{\max} \sqrt{\ell}} \|y\|_{1,\infty},$$

where $P_{\mathcal{A}}$ is the orthogonal projection on the kernel of $\mathcal{A}$. In particular, $P_{\mathcal{A}} x = x$ if $\mathcal{A} x = 0$ and $P_{\mathcal{A}} x = 0$ if and only if there is u such that $x = \mathcal{A}^\top u$. We will, in fact, conduct a more general analysis and suppose that y satisfies

$$\|P_{\mathcal{A}} y\| \leq \frac{1}{\rho r_{\max} \sqrt{\ell}} \|y\|_{1,\infty},$$

for some $\rho > 1$. We are now going to show how to use y to construct w_k, v_k such that $H(w_k, v_k)$ satisfies the properties (P.1), (P.2) and (P.3) described in Sect. 1. We divide our discussion in two parts. In Sect. 3.1, we deal with the case where $\mathcal{K}$ is a simple symmetric cone. Then, we remove the simplicity assumption and do the general case in Sect. 3.2.

3.1 The simple case

Here, we suppose that $\mathcal{K}$ is a simple symmetric cone, so that the dimension of $\mathcal{E}$ is d, $\ell = 1, r_{\max} = r$ and $\|\cdot\|_{1,\infty} = \|\cdot\|_1$. First, we study a few properties of the intersection $H(w, v) \cap \mathcal{K}$, where $H(w, v) = \{x \in \mathcal{E} \mid \langle w, x \rangle \leq \langle w, v \rangle\}$. We denote the volume of $H(w, v) \cap \mathcal{K}$ by vol (w, v). We then have the following results, which are analogous to Proposition 2.2 in [13].

Proposition 6 *Suppose $w \in$ int $\mathcal{K}$. Then,*

$$Q_{w^{-1/2}\sqrt{\langle w,v \rangle}}(H(e, e/r)) = H(w, v) \tag{9}$$

$$\text{vol}(w, v) = \left(\frac{\langle w, v \rangle}{\sqrt[r]{\det w}} \right)^d \text{vol}(e, e/r) \tag{10}$$

Proof For the first equation, suppose that $x \in H(e, e/r)$. Then, by Proposition 3, we have

$$\left\langle w, Q_{w^{-1/2}\sqrt{\langle w,v \rangle}}(x) \right\rangle = \langle w, v \rangle \langle Q_{w^{-1/2}}(w), x \rangle = \langle w, v \rangle \langle e, x \rangle \leq \langle w, v \rangle,$$

which shows that $Q_{w^{-1/2}\sqrt{\langle w,v \rangle}} \in H(w, v)$.

Also by Proposition 3, we have $Q_{w^{-1/2}\sqrt{\langle w,v \rangle}}^{-1} = \frac{1}{\langle w,v \rangle} Q_{w^{1/2}}$. So, suppose that $x \in H(w, v)$.

$$\frac{1}{\langle w, v \rangle} \langle e, Q_{w^{1/2}}(x) \rangle = \frac{1}{\langle w, v \rangle} \langle Q_{w^{1/2}}(e), x \rangle = \frac{1}{\langle w, v \rangle} \langle w, x \rangle \leq 1,$$

which shows that $\frac{1}{\langle w,v \rangle} Q_{w^{1/2}}(x) \in H(e, e/r)$.

For the second equation,

$$\int_{H(w,v) \cap \mathcal{K}} 1 = \int_{Q_{w^{-1/2}\sqrt{\langle w,v \rangle}}(H(e,e/r) \cap \mathcal{K})} 1$$

$$= \int_{H(e,e/r) \cap \mathcal{K}} \left| \det Q_{w^{-1/2}\sqrt{\langle w,v \rangle}} \right|$$

$$= \det Q_{w^{-1/2}\sqrt{\langle w,v \rangle}} \text{vol}(e, e/r)$$

By Proposition 3, we have

$$
\begin{aligned}
\det Q_{w^{-1/2}\sqrt{\langle w,v\rangle}} &= \det\langle w,v\rangle Q_{w^{-1/2}} \\
&= \langle w,v\rangle^d \det Q_{w^{-1/2}} \\
&= \langle w,v\rangle^d \left(\det w^{-1/2}\right)^{\frac{2d}{r}} \\
&= \left(\frac{\langle w,v\rangle}{\sqrt[r]{\det w}}\right)^d .
\end{aligned}
$$

$\square$

The next lemma shows that we can use y to construct a hyperplane $H(w,v)$ which contains $\mathcal{F}^{\mathcal{A}}_{\text{scaled}}$ and has its roots in the work by Chubanov on LPs, see page 692 of [4]. See also, Lemma 3.1 in [13] for the SOCP case.

Lemma 7 *Suppose x is feasible for ($\text{P}^{\mathcal{A}}_{\text{scaled}}$) and that $y \in \mathcal{K}$ satisfies*

$$
\|P_{\mathcal{A}}y\| \le \frac{1}{\rho r}\|y\|_1 ,
$$

then $x \in H(y, e/\rho r)$.

Proof Since $y \in \mathcal{K}$, we have $\|y\|_1 = \langle y,e\rangle$. In addition, since x is feasible for ($\text{P}^{\mathcal{A}}_{\text{scaled}}$) and $\ell = 1$, we have $\|x\|_{1,\infty} = \|x\|_1 \le 1$. Then, by using the Proposition 4 and the inequality in (5), we have

$$
\langle y,x\rangle = \langle y, P_{\mathcal{A}}x\rangle = \langle P_{\mathcal{A}}y, x\rangle \le \|P_{\mathcal{A}}y\|_\infty \|x\|_1 \le \|P_{\mathcal{A}}y\| \le \frac{1}{\rho r}\langle y,e\rangle.
$$

$\square$

Due to Proposition 6, as long as $y \in \operatorname{int}\mathcal{K}$, there is some Q that maps $H(e, e/r)$ bijectively into $H(y, e/\rho r)$. The only problem is that the volume of $H(y, e/\rho r) \cap \mathcal{K}$ might not be sufficiently small. We will now address this issue and try to give some intuition on the choices we will make ahead.

As long as $\mathcal{F}^{\mathcal{A}}_{\text{scaled}}$ is contained in $H(w,v) \cap \mathcal{K}$, we may select any w and v such that the volume of $H(w,v) \cap \mathcal{K}$ is small. However, finding w and v directly is cumbersome. Instead, we use y and settle for the easier goal of finding w, v such that $H(y, e/\rho r) \cap H(e, e/r) \subseteq H(w,v)$, since this is enough to ensure that $\mathcal{F}^{\mathcal{A}}_{\text{scaled}}$ is contained in $H(w,v)$. While doing so, we seek to minimize $\operatorname{vol}(w,v)$. Theorem 22.3 of [24], tells us that $H(y, e/\rho r) \cap H(e, e/r) \subseteq H(w,v)$ if and only if there are $\alpha \ge 0, \beta \ge 0$ such that

$$
\alpha y + \beta e = w \tag{11}
$$

$$
\alpha \frac{\langle y,e\rangle}{\rho r} + \beta \le \langle v,w\rangle. \tag{12}
$$

Furthermore, for fixed $v \in \operatorname{int} \mathcal{K}$, the w that minimizes $\operatorname{vol}(w, v)$ is v^{-1}. This is a consequence of the next proposition, which is analogous to Proposition 2.3 in [13].

Proposition 8 *Fix $v \in \operatorname{int} \mathcal{K}$. Then $w = v^{-1}$ minimizes $\operatorname{vol}(w, v)$. So that*

$$\operatorname{vol}(w, v) = \left(\frac{r}{\sqrt[r]{\det v^{-1}}} \right)^d \operatorname{vol}(e, e/r) = \left(r \sqrt[r]{\det v} \right)^d \operatorname{vol}(e, e/r).$$

Proof It follows from Proposition 6 that minimizing $\operatorname{vol}(w, v)$ is the same as minimizing $\left(\frac{\langle w, v \rangle}{\sqrt[r]{\det w}} \right)^d$. We will minimize, equivalently, $\log \left(\frac{\langle w, v \rangle}{\sqrt[r]{\det w}} \right)$. We have

$$f(w) = \log \left(\frac{\langle w, v \rangle}{\sqrt[r]{\det w}} \right) = \log \langle w, v \rangle - \frac{1}{r} \log \det w.$$

Proposition III.4.2 of [5] tells us that $\nabla \log \det w = w^{-1}$. We then have:

$$\nabla f(w) = \frac{v}{\langle w, v \rangle} - \frac{1}{r} w^{-1}.$$

In order to force $\nabla f(w) = 0$, we may take, for instance, $w = v^{-1}$ so that $\langle w, v \rangle = r$. From Proposition 6, we obtain

$$\operatorname{vol}(w, v) = \left(\frac{r}{\sqrt[r]{\det v^{-1}}} \right)^d \operatorname{vol}(e, e/r) = \left(r \sqrt[r]{\det v} \right)^d \operatorname{vol}(e, e/r).$$

$\square$

From Proposition 8, our quest for a good $H(w, v)$ can be summarized by trying to minimize

$$\det w^{-1} = \det(\alpha y + \beta e)^{-1},$$

subject to $\alpha \geq 0$, $\beta \geq 0$ and $\alpha \frac{\langle y, e \rangle}{\rho r} + \beta \leq r$. Note that there is no benefit in letting the last inequality be inactive. Furthermore, we would rather minimize the logarithm of the function above, which gives a convex problem. So we let $\alpha = \left(\frac{r - \beta}{\langle y, e \rangle} \right) \rho r$ and consider the following problem

$$\inf_{0 \leq \beta \leq r} \quad -\log \det \left(re + \left(\frac{r - \beta}{\langle y, e \rangle} \right) \rho r y + e(\beta - r) \right). \tag{Aux}$$

Of course, there is no need to minimize (Aux) exactly. We only need that $\frac{r}{\sqrt[r]{\det w}} < \delta$, or equivalently, $-r \log r + r \log \delta > -\log \det w$, for some fixed positive constant $\delta < 1$ that does not depend on $\mathcal{A}$ but could, possibly, depend on r and d. In Theorem 10 we present a choice of β that is good enough for our purposes.

We first need the following auxiliary fact, which appears in some form or another in the analysis of many interior point methods dating back to the original paper by Karmakar, see Lemma 4.2 in [11]. It is in fact a consequence of the self-concordance of the $-\log\det(\cdot)$ function and can be derived from the bounds appearing, for instance, in Section 2.1.1 in [20]. Still, for the sake of self-containment, we will not directly use self-concordance in proving the next bound.

Lemma 9 *Let $h \in \mathcal{E}$ be such that $\|h\| < r$. Then*

$$-\log\det(re + h) \leq -r\log r - \frac{\langle h, e\rangle}{r} + \frac{\|h\|^2}{2r(r - \|h\|)}.$$

Proof Note that the case $r = 1$ corresponds to $\mathcal{E} = \mathbb{R}, \mathcal{K} = \mathbb{R}_+, e = 1$ and is the statement that if $\alpha \in \mathbb{R}$ and $-1 < \alpha < 1$, then

$$-\log(1 + \alpha) \leq -\alpha + \frac{\alpha^2}{2(1 - |\alpha|)}, \tag{13}$$

which is a known bound for the logarithm function. See Lemmas 4.1 and 4.2 in [11] or Lemma 3.1 in [7]. From (13) we can prove the general case. Let $\lambda_1, \ldots, \lambda_r$ be the eigenvalues of h. We have the following relations

$$\langle h, e\rangle = \lambda_1 + \cdots + \lambda_r \tag{14}$$

$$\|h\|^2 = \lambda_1^2 + \cdots + \lambda_r^2. \tag{15}$$

Since the corresponding eigenvalues of $re + h$ are $r + \lambda_1, \ldots, r + \lambda_r$, we have

$$-\log\det(re + h) = -r\log r - \sum_{i=1}^{r} \log\left(1 + \frac{\lambda_i}{r}\right). \tag{16}$$

By hypothesis, we have $\|h\|^2 < r^2$, which implies that $|\lambda_i/r|^2 < 1$ and $|\lambda_i/r| < 1$ for every i. Recalling (14) and applying (13) to each term of (16) we get

$$-\log\det(re + h) \leq -r\log r - \frac{\langle h, e\rangle}{r} + \sum_{i=1}^{r} \frac{\lambda_i^2}{2r^2(1 - |\lambda_i/r|)}.$$

Finally, due to (15) and the fact that $1 - |\lambda_i/r| \geq 1 - \|h\|/r > 0$ holds for all i, we have the desired inequality. $\qquad\square$

We are now ready to show how to construct $H(w, v)$ satisfying the properties (P.1), (P.2) (P.3) outlined in Sect. 1.

Theorem 10 *Let $\rho > 1$ and $y \in \mathcal{K}$ be such that $\mathcal{F}^{\mathcal{A}}_{scaled} \subseteq H(y, e/\rho r)$ and $y \neq 0$. Let[3]*

$$\beta = r - \left(\frac{1}{\rho} - \frac{1}{\sqrt{\rho(3\rho - 2)}} \right)$$

$$w = \left(\frac{r - \beta}{\langle y, e \rangle} \right) \rho r y + \beta e$$

$$v = w^{-1}.$$

Then, the following hold:

(i) $\mathcal{F}^{\mathcal{A}}_{scaled} \subseteq H(y, e/\rho r) \cap H(e, e/r) \subseteq H(w, v)$

(ii) $Q_{w^{-1/2}\sqrt{r}}(H(e, e/r)) = H(w, v)$

(iii) $\mathrm{vol}\,(w, v) = \left(\frac{r}{\sqrt[r]{\det w}} \right)^d \mathrm{vol}\,(e, e/r) \leq (\exp(-\varphi(\rho)/r))^d \,\mathrm{vol}\,(e, e/r)$, *where*

$$\varphi(\rho) = 2 - \frac{1}{\rho} - \sqrt{3 - \frac{2}{\rho}}.$$

In particular if $\rho \geq 2$, we have $\mathrm{vol}\,(w, v) < (0.918)^{d/r} \,\mathrm{vol}\,(e, e/r)$.

Proof Following the discussion so far, we consider the problem (Aux).

$$\inf_{0 \leq \beta \leq r} \quad -\log \det \left(re + \left(\frac{r - \beta}{\langle y, e \rangle} \right) \rho r y + e\,(\beta - r) \right). \qquad (Aux)$$

To avoid making it seem as if β appeared from thin air, we will try to show the rationale behind our choice of β. Our task is to show that there is some $0 \leq \beta \leq r$ and some fixed constant $\delta < 1$ for which

$$\frac{r}{\sqrt[r]{\det w}} < \delta.$$

Or, equivalently,

$$-\log \det \left(re + \left(\frac{r - \beta}{\langle y, e \rangle} \right) \rho r y + e\,(\beta - r) \right) \leq -r \log r + r \log \delta.$$

Once we find β and δ we can claim that $\mathrm{vol}\,(w, v) \leq \delta^d \mathrm{vol}\,(e, e/r)$. Furthermore, by construction, we have item (i), since β and $\alpha = \left(\frac{r - \beta}{\langle y, e \rangle} \right) \rho r$ are nonnegative and solve the linear inequalities (11) and (12). Item (ii) is just a consequence of Proposition 6. To start, let $h = \left(\frac{r - \beta}{\langle y, e \rangle} \right) \rho r y + e\,(\beta - r)$. We have

---

[3] If $y \neq 0$ and $y \in \mathcal{K}$ then $\langle y, e \rangle > 0$, since the latter is sum of the eigenvalues of y, which are nonnegative. Furthermore, not all of them are zero, since $y \neq 0$.

$$h = (r - \beta)\left(\frac{\rho r}{\langle y, e \rangle} y - e\right)$$

$$\frac{\langle h, e \rangle}{r} = (r - \beta)(\rho - 1) \tag{17}$$

$$\|h\|^2 = (r - \beta)^2 \left(r(1 - 2\rho) + \rho^2 r^2 \frac{\|y\|^2}{\langle y, e \rangle^2}\right)$$

As $\|y\| \leq \langle y, e \rangle$ and $\rho > 1$, we get the bound

$$\|h\| \leq (r - \beta)\rho r. \tag{18}$$

So in order to apply Lemma 9, it is enough to take $(r - \beta)\rho r < r$. So, suppose that $r - \beta < 1/\rho$. This, together with (18), implies that

$$\frac{1}{r(1 - \rho(r - \beta))} \geq \frac{1}{r - \|h\|}. \tag{19}$$

Then, (17), (18) and (19) together with Lemma 9 imply that:

$$-\log\det(re + h) \leq -r\log r - (r - \beta)(\rho - 1) + \frac{(r - \beta)^2 \rho^2 r^2}{2r^2(1 - \rho(r - \beta))}$$

$$= -r\log r - (r - \beta)(\rho - 1) + \frac{(r - \beta)^2 \rho^2}{2(1 - \rho(r - \beta))}.$$

Let $\psi(z) = -z(\rho - 1) + \frac{z^2 \rho^2}{2(1 - \rho z)}$. A tedious but straightforward calculation[4] shows that in the interval $(0, 1/\rho)$, $\psi(z)$ is minimized at

$$z_\rho = \frac{1}{\rho} - \frac{1}{\sqrt{\rho(3\rho - 2)}}.$$

and for that z_ρ we have $\psi(z_\rho) = \sqrt{3 - \frac{2}{\rho}} + \frac{1}{\rho} - 2$. So, we let $r - \beta = z_\rho$ and conclude that for this choice of β, we have

$$-\log\det(re + h) \leq -r\log r - \left(2 - \frac{1}{\rho} - \sqrt{3 - \frac{2}{\rho}}\right).$$

---

[4] It is easy to verify this computation by using a computer algebra system such as the open-source software `maxima` [17]. Consider the following three `maxima` commands: `f:-z*(p-1)+(z^2*p^2)/(2*(1-p*z));` `zp:solve(diff(f,z,1),z);` `for i:1 thru 2 do disp(combine(expand(ratsimp(substitute(zp[i],z,f)))));`. The first two commands compute z_ρ and the last computes $\psi(z_\rho)$. Note that we discard one of the solutions because it is not in $(0, 1/\rho)$.
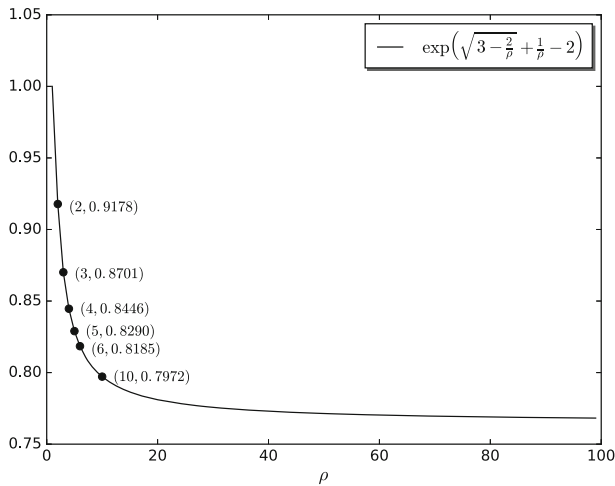
Fig. 1 The function $\exp(-\varphi(\rho))$. The decrease in volume is bounded by $\exp(-\varphi(\rho))^{\frac{d}{r}}$

Therefore, we take δ such that $r \log \delta = \psi(z_\rho)$. Thus we conclude that $\mathrm{vol}\,(w, v) \leq (\exp(-\varphi(\rho))/r))^d \, \mathrm{vol}\,(e, e/r)$, where $\varphi(\rho) = -\psi(z_\rho)$. Examining the first derivative of $-\varphi$, we see that it is a decreasing function. In particular, if $\rho \geq 2$, we have $\exp(-\varphi(\rho)) < 0.918$. $\square$

See Figure 1 for the graph of $\exp(-\varphi(\rho))$.

3.2 The general case

Now, we move on to the general case and we suppose that $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_\ell$, where each $\mathcal{K}_i$ is a simple symmetric cone of rank r_i and dimension d_i. We have $r = r_1 + \cdots + r_\ell$ and $d = d_1 + \cdots + d_\ell$. Suppose that we have found some nonzero $y \in \mathcal{K}$ satisfying

$$\|P_{\mathcal{A}}y\| \leq \frac{1}{2r_{\max}\sqrt{\ell}} \, \|y\|_{1,\infty}\,,$$

where we recall that $r_{\max} = \max\{r_1, \ldots, r_\ell\}$ and that since $y \in \mathcal{K}$, we have $\|y\|_{1,\infty} = \max\{\langle e_1, y_1\rangle, \ldots, \langle e_\ell, y_\ell\rangle\}$, see Sect. 2.1. We have the following observation.

Lemma 11 *Suppose x is feasible for $(P^{\mathcal{A}}_{scaled})$ and that $y \in \mathcal{K}$ satisfies $P_{\mathcal{A}}y \neq 0$ and*

$$\|P_{\mathcal{A}}y\| \leq \frac{1}{2r_{\max}\sqrt{\ell}} \, \|y\|_{1,\infty}\,.$$

For every $k \in \{1, \ldots, \ell\}$, define ρ_k as

$$\rho_k = \frac{\|y_k\|_1}{r_k \, \|P_{\mathcal{A}}y\| \, \sqrt{\ell}}. \tag{20}$$

(i) *Let k be any index such that $\|y\|_{1,\infty} = \|y_k\|_1$, then $\rho_k \geq 2$.*
(ii) *Let x be feasible for $(P^{\mathcal{A}}_{scaled})$, then $x_k \in H(y_k, e_k/\rho_k r_k)$.*

Proof (i) By the definition of ρ_k, we see that $\rho_k \geq 2$ if $\|y_k\|_1 = \|y\|_{1,\infty}$, since $r_k \leq r_{\max}$.

(ii) Since $y \in \mathcal{K}$, we have $\|y_k\|_1 = \langle y_k, e_k \rangle$. Recall that $P_{\mathcal{A}}$ is self-adjoint, so that $\langle y, P_{\mathcal{A}} x \rangle = \langle P_{\mathcal{A}} y, x \rangle$. In addition, since $x \in \mathcal{F}^{\mathcal{A}}_{scaled}$, we have $\|x\|_{1,\infty} \leq 1$. Therefore, using the Proposition 4 together with the generalized Cauchy–Schwarz inequality in (6) we have:

$$\langle y_k, x_k \rangle \leq \langle y, x \rangle = \langle y, P_{\mathcal{A}} x \rangle \leq \|P_{\mathcal{A}} y\|_{\infty,1} \|x\|_{1,\infty} \leq \|P_{\mathcal{A}} y\|_2 \sqrt{\ell}$$
$$= \frac{\|y_k\|_1}{\rho_k r_k} = \frac{\langle y_k, e_k \rangle}{\rho_k r_k}.$$

$\square$

As a reminder, we write vol (w_k, v_k) for the volume of the region $H(w_k, v_k) \cap \mathcal{K}_k$. The idea is that once y is found, we can use y_k and Theorem 10 to find a hyperplane $H(w_k, v_k) \subseteq \mathcal{E}_k$, such that vol (w_k, v_k) is small. This is summarized in the following theorem.

Theorem 12 *Let $y \in \mathcal{K}$ be such that $\|P_{\mathcal{A}} y\| \leq \frac{1}{2 r_{\max} \sqrt{\ell}} \|y\|_{1,\infty}$ and that $P_{\mathcal{A}} y \neq 0$. For all blocks $\mathcal{E}_k$, let ρ_k be as in Eq. (20) and suppose that for some k we have $\rho_k > 1$. Let $w_k, v_k \in \mathcal{K}_k, \beta_k \in \mathbb{R}$ be such that[5]*

$$\beta_k = r_k - \left(\frac{1}{\rho_k} - \frac{1}{\sqrt{\rho_k(3\rho_k - 2)}} \right)$$
$$w_k = \left(\frac{r_k - \beta_k}{\langle y_k, e_k \rangle} \right) \rho_k r_k y_k + \beta_k e_k$$
$$v_k = w_k^{-1}.$$

Then, the following hold:

(i) *If $x \in \mathcal{F}^{\mathcal{A}}_{scaled}$, then $x_k \in H(w_k, v_k)$,*
(ii) *$Q_{w_k^{-1/2} \sqrt{r_k}}(H(e_k, e_k/r_k)) = H(w_k, v_k)$,*
(iii) *vol $(w_k, v_k) = \left(\frac{r_k}{r_k \sqrt[d]{\det w_k}} \right)^{d_k}$ vol $(e_k, e_k/r_k) \leq (\exp(-\varphi(\rho_k)/r_k))^{d_k}$ vol $(e_k, e_k/r_k)$, where*

$$\varphi(\rho_k) = 2 - \frac{1}{\rho} - \sqrt{3 - \frac{2}{\rho}}.$$

In particular if $\rho_k \geq 2$, we have vol $(w_k, v_k) < (0.918)^{d_k/r_k}$ vol $(e_k, e_k/r_k)$.

---

[5] $P_{\mathcal{A}} y \neq 0$ implies $y \neq 0$, which implies $\langle y, e \rangle > 0$, since $y \in \mathcal{K}$. See the footnote in Theorem 10.

Proof Due to Lemma 11, we have that if $x \in \mathcal{F}^{\mathcal{A}}_{\text{scaled}}$, then x_k is confined to $H(y_k, e_k/\rho_k r_k) \cap H(e_k, e_k/r_k)$. Then, the choice of β_k, w_k, v_k is exactly the same as suggested by Theorem 10. This shows that $H(y_k, e_k/\rho_k r_k) \cap H(e_k, e_k/r_k) \subseteq H(w_k, v_k)$, which implies $x_k \in H(w_k, v_k)$, which is item (i). Item (ii) and (iii) are also direct consequences of items (ii) and (iii) of Theorem 10. □

Then, the idea is to replace $\mathcal{A}$ by $\mathcal{A}Q$, where Q is the linear map such that

$$Q(x_1, \ldots, x_k, \ldots, x_\ell) = (Q_1(x_1), \ldots, Q_\ell(x_\ell)),$$

where

$$Q_i = \begin{cases} Q_{w_i^{-1/2}} \sqrt{r_i} & \text{if } \rho_i > 1 \\ I_i & \text{otherwise}, \end{cases}$$

where I_i is the identity map on $\mathcal{E}_i$. If y satisfying Proposition 12 is found, $\rho_k \geq 2$ holds for at least one block. Therefore, for at least one block the decrease in volume is bounded above by a positive constant less than 1. This is the foundation upon which Algorithm 2 in Sect. 5 is constructed.

4 Basic procedure

We remark that the procedure described here is a direct extension of Chubanov's basic procedure in [4], see also Section 3 in [13] for the corresponding discussion on SOCPs. From Sect. 3 and Theorem 12, we see that as long as we are able to find y such that $\|P_{\mathcal{A}} y\| \leq \frac{1}{2r_{\max}\sqrt{\ell}} \|y\|_{1,\infty}$, we can confine one of the blocks of the feasible solutions of ($P^{\mathcal{A}}_{\text{scaled}}$) to a region $H(w_k, v_k) \cap \mathcal{K}_k$ with smaller volume than $H(e_k, e_k/r_k) \cap \mathcal{K}_k$.

The "basic procedure" can be seen as an algorithm for minimizing $\|P_{\mathcal{A}} y\|$, but we early stop it and pass control forward after the threshold $\|P_{\mathcal{A}} y\| \leq \frac{1}{2r_{\max}\sqrt{\ell}} \|y\|_{1,\infty}$ is met. So suppose that we start with some $y \in \text{int } \mathcal{K}$ such that $\langle y, e \rangle = 1$. We then let $z = P_{\mathcal{A}} y$.

If $z = 0$, then $y \in (\ker \mathcal{A})^\perp$, so $y = \mathcal{A}^\top u$ for some u, which means that y is feasible for (D) since $y \in \mathcal{K}$ and y is not zero. If $z \in \text{int } \mathcal{K}$, then since $\mathcal{A}z = 0$, we have that z is feasible for (P). If neither of those criteria are met but $\|z\| \leq \frac{1}{2r_{\max}\sqrt{\ell}} \|y\|_{1,\infty}$ holds, then we stop the algorithm anyway.

Otherwise, we have $z \neq 0$ and $z \notin \text{int } \mathcal{K}$. Because $z \notin \text{int } \mathcal{K}$, there is a nonzero $c \in \mathcal{K}$ such that $\langle z, c \rangle \leq 0$, which is a consequence of standard separation results and of the fact that $\mathcal{K}$ is self-dual. Since any multiple of c will do the job, we can scale it and suppose that $\langle e, c \rangle = 1$. Then, it seems sensible to try to correct y by considering

$$y' = \alpha y + (1 - \alpha)c,$$

where we select α in such a way that $P_{\mathcal{A}} y'$ is as close as possible to the origin, since we wish to minimize $\|P_{\mathcal{A}} y\|$. Let $p = P_{\mathcal{A}} c$. The α that minimizes $\|P_{\mathcal{A}} y'\|$ is

$$\alpha = \frac{\langle p, p - z \rangle}{\|z - p\|^2}.$$

Note that after computing p, we may check whether p meets our stopping criteria. That is, we check whether $p = 0$ (so that c is feasible for (D)), $p \in \mathrm{int}\,\mathcal{K}$ (so that p is feasible for (P)) or $\|p\| \leq \frac{1}{2r_{\max}\sqrt{\ell}} \|c\|_{1,\infty}$. If neither p nor z satisfies our stopping criteria, then $\alpha \in (0, 1)$ so that $y' \in \mathrm{int}\,\mathcal{K}$, see Theorem 6.1 in [24]. Moreover, this update decreases the norm of z, as described by the following proposition.

Proposition 13 *Let:*

  (i) $y \in \mathrm{int}\,\mathcal{K}$, $\langle e, y \rangle = 1$ *and* $z = P_{\mathcal{A}} y$,
 (ii) $c \in \mathcal{K}$ *be such that* $\langle z, c \rangle \leq 0$ *and* $\langle e, c \rangle = 1$,
(iii) $p = P_{\mathcal{A}} c$ *and* $\alpha = \frac{\langle p, p - z \rangle}{\|z - p\|^2}$,
 (iv) $y' = \alpha y + (1 - \alpha)c$.

Suppose that $p \neq 0$ *and* $z \neq 0$. *Then:*

  (a) $\frac{1}{\|P_{\mathcal{A}} y'\|^2} \geq \frac{1}{\|P_{\mathcal{A}} y\|^2} + 1$.
  (b) $y' \in \mathrm{int}\,\mathcal{K}$, $\langle e, y' \rangle = 1$ *and* $\|y'\|_{1,\infty} \geq \frac{1}{\ell}$,

Proof We first prove item (a). By (ii), we have

$$\langle z, p \rangle = \langle P_{\mathcal{A}} z, c \rangle = \langle z, c \rangle \leq 0.$$

This implies that

$$\|z - p\|^2 = \|z\|^2 + \|p\|^2 - 2\langle z, p \rangle > 0,$$

since both z and p are nonzero. We also have

$$1 = \frac{\|z\|^2 - \langle z, p \rangle}{\|z - p\|^2} + \frac{\|p\|^2 - \langle z, p \rangle}{\|z - p\|^2}.$$

We then conclude that the denominator of α is not zero and $\alpha \in (0, 1)$. We then have

$$P_{\mathcal{A}} y' = p + \alpha(z - p).$$

Therefore,

$$\|P_{\mathcal{A}} y'\|^2 = \alpha^2 \|z - p\|^2 + 2\alpha \langle p, z - p \rangle + \|p\|^2.$$

It follows that

$$\|P_{\mathcal{A}}y'\|^2 = \|p\|^2 - \frac{\langle p, z - p\rangle^2 - \langle z, p\rangle^2}{\|z - p\|^2} = \frac{\|z\|^2\|p\|^2 - \langle z, p\rangle^2}{\|z\|^2 + \|p\|^2 - 2\langle z, p\rangle}$$

$$\leq \frac{\|z\|^2\|p\|^2}{\|z\|^2 + \|p\|^2}.$$

Since $P_{\mathcal{A}}$ is a projection matrix, we have $\|p\|^2 \leq \|c\|^2 \leq \langle c, e\rangle^2 \leq 1$. This implies that

$$\frac{1}{\|P_{\mathcal{A}}y'\|^2} \geq \frac{1}{\|z\|^2} + \frac{1}{\|p\|^2} \geq \frac{1}{\|z\|^2} + 1.$$

Now we prove item (b). Since we have shown that $\alpha \in (0, 1)$, y' is a strict convex combination of c and the interior point y, so it is also an interior point. For the same reason, y' also satisfies $\langle e, y'\rangle = 1$, since $\langle e, c\rangle = 1$ and $\langle e, y\rangle = 1$, due to items (i) and (ii). Finally, note that since $y' \in \mathcal{K}$, we have $1 = \langle y', e\rangle = \sum_{i=1}^{\ell} \|y'_i\|_1 \leq \ell \|y'\|_{1,\infty}$, which shows that $\|y'\|_{1,\infty} \geq \frac{1}{\ell}$. $\qquad\square$

We remark that although the increase in $\frac{1}{\|P_{\mathcal{A}}y'\|^2}$ is bounded below by 1 regardless of c, the proof shows that the actual increase is a function of $-\langle z, p\rangle = -\langle z, c\rangle$. Therefore, if we wish to maximize the increase, we have to minimize $\langle z, c\rangle$ subject to $\langle c, e\rangle = 1$ and $c \in \mathcal{K}$. Similarly to the case of positive semidefinite matrices, the optimal value of this problem is $\lambda_{\min}(z)$, which is achieved when c is the idempotent associated to the minimum eigenvalue of z, see, for instance, Equation (9) in [28].

As we are assuming that $\mathcal{K} = \mathcal{K}_1 \times \ldots \times \mathcal{K}_\ell$, the computation of the minimum eigenvalue of z can be done by computing the minimum eigenvalue of each block and then taking the overall minimum. That is,

$$\lambda_{\min}(x) = \min\{\lambda_{\min}(x_1), \ldots, \lambda_{\min}(x_\ell)\}.$$

This is advantageous because if $\mathcal{K}_i$ is either $\mathbb{R}_+$ or $\mathcal{L}_\ell$ the minimum eigenvalue and the corresponding idempotent can be computed exactly, see (7) and (8).

We can now state our version of the basic procedure which encompasses the discussion so far, see Algorithm 1. We remind that superscripts in Algorithms 1, 2, 3 and 4 such as y^i, $\mathcal{A}^i$ denote the iteration number, not exponentiation. We have the following complexity bounds. For the SOCP analogue, see the comments before Section 5 in [13].

Proposition 14 *Algorithm* 1 *needs no more than* $4\ell^3 r_{\max}^2$ *steps before it halts.*

Proof Due to Proposition 13, all the iterates y^i satisfy $\|y^i\|_{1,\infty} \geq 1/\ell$. Therefore, if $\|P_{\mathcal{A}}y^i\| \leq \frac{1}{2r_{\max}\ell^{3/2}}$, we will meet the stopping criteria $\|P_{\mathcal{A}}y^i\| \leq \frac{\|y^i\|_{1,\infty}}{2r_{\max}\sqrt{\ell}}$. Also due to Proposition 13, $\frac{1}{\|P_{\mathcal{A}}y^i\|^2}$ improves by at least one after each iteration. Therefore, after $4\ell^3 r_{\max}^2$ iterations, Algorithm 1 will halt for sure. $\qquad\square$

Algorithm 1: Basic Procedure

Input: $\mathcal{A}$, $P_{\mathcal{A}}$, $y^1 \in \text{int}\,\mathcal{K}$ such that $\langle y^1, e \rangle = 1$
Output:
  1. y such that $\|P_{\mathcal{A}}y\| \leq \frac{1}{2r_{\max}\sqrt{\ell}}\|y\|_{1,\infty}$, or
  2. a solution to (P), or
  3. a solution to (D).

1 $i \leftarrow 1$, $z \leftarrow P_{\mathcal{A}}y^1$.

2 **while** $z \neq 0$ *and* $z \notin \text{int}\,\mathcal{K}$ *and* $\|z\| > \frac{1}{2r_{\max}\sqrt{\ell}}\left\|y^i\right\|_{1,\infty}$ **do**

3 $\quad$ Let c be an idempotent such that $\langle e, c \rangle = 1$ and $\langle z, c \rangle = \lambda_{\min}(z)$.

4 $\quad$ $p \leftarrow P_{\mathcal{A}}c$.

5 $\quad$ **if** $p = 0$ *or* $p \in \text{int}\,\mathcal{K}$ *or* $\|p\| \leq \frac{1}{2r_{\max}\sqrt{\ell}}\|c\|_{1,\infty} = \frac{1}{2r_{\max}\sqrt{\ell}}$ **then**

6 $\quad\quad$ **stop** and **return** c.

7 $\quad$ **else**

8 $\quad\quad$ $y^{i+1} \leftarrow \alpha y^i + (1-\alpha)c$, where $\alpha = \frac{\langle p, p-z \rangle}{\|z-p\|^2}$.

9 $\quad$ Set $i \leftarrow i+1$ and $z \leftarrow P_{\mathcal{A}}y^i$.

10 **return** y^i.

Any algorithm aimed at minimizing $\|P_{\mathcal{A}}y\|$ can serve as a "basic procedure", as long as there is some bound that ensures that we will hit one of the stopping criteria in a finite amount of time. In Section 5 of [23], the authors describe four different algorithms that can serve as basic procedure. The algorithm described here is analogous to the "Von Neumann scheme" which, as the authors remark, is essentially what was used by Chubanov in [4]. Nevertheless among the four, the "smooth perceptron" [27] has the better complexity bound for the number of iterations, although each iteration is more expensive.

In particular, the smooth perceptron is such that if it has not finished at the i-th iteration then, $\|Py^i\|^2 \leq \frac{8}{(t+1)^2}$. So if we use the stopping criteria that y should satisfy $\|P_{\mathcal{A}}y\| \leq \frac{1}{2r_{\max}\sqrt{\ell}}\|y\|_{1,\infty}$, as in the proof of Proposition 1, it is enough to force $\|P_{\mathcal{A}}y\|$ to be smaller than $\frac{1}{2r_{\max}\ell^{3/2}}$, which will happen in no more than $4\sqrt{2}r_{\max}\ell^{3/2}-1$ iterations.

We will now estimate the cost per iteration of Algorithm 1. Recall that the dimension of $\mathcal{E}$ is d, so we can identify $\mathcal{E}$ with some $\mathbb{R}^d$. Then, we can assume that $\mathcal{A}$ is, in fact, a $m \times d$ matrix. For the sake of estimating the computational cost, we will assume that $\mathcal{A}$ is surjective, but we emphasize that the theoretical results proved in this work do not need this assumption. Note that computations performed in both the Basic Procedure (Algorithm 1) and the Main Procedure (Algorithm 2) only depend on the kernel of $\mathcal{A}$.

Assuming that $\mathcal{A}$ is surjective, we may compute $P_{\mathcal{A}}$ naively through the expression $I - \mathcal{A}^*(\mathcal{A}\mathcal{A}^*)^{-1}\mathcal{A}$, where I is the identity operator on $\mathcal{E}$. Following a similar suggestion in Section 6 of [23], we compute the Cholesky decomposition of $(\mathcal{A}\mathcal{A}^*)^{-1}$ so that $(\mathcal{A}\mathcal{A}^*)^{-1} = LL^*$, where L is a $m \times m$ matrix. Then, we store the $m \times d$ matrix $L^*\mathcal{A}$. Note that computing $\mathcal{A}\mathcal{A}^*$ and decomposing its inverse can be done in time proportional to $\mathcal{O}(m^3) + \mathcal{O}(m^2 d)$.

As $P_{\mathcal{A}}$ stays constant throughout the algorithm, we only need to compute it once and then we just have to worry about computing $P_{\mathcal{A}}y^i$ and $P_{\mathcal{A}}c$ at each iteration,

which costs $\mathcal{O}(md)$ if we make efficient use of the matrix $L^*\mathcal{A}$. In our case, $P_\mathcal{A}$ is computed inside Algorithms 1 and 2 receives it as input.

We now list other costs that need to be taken into account: the cost $c_{\text{int } \mathcal{K}}$ of deciding if $z \in \text{int } \mathcal{K}$, the cost $c_{\min}$ of computing the minimum eigenvalue $\lambda_{\min}(x)$ with the corresponding idempotent and the cost c_{norm} of computing the norm of $\|x\|_{1,\infty}$ for elements $x \in \mathcal{K}$.

The cost $\|x\|_{1,\infty}$ is proportional to $\mathcal{O}(d)$ since for $x \in \mathcal{K}$, we have $\|x\|_{1,\infty} = \max\{\langle x_1, e_1 \rangle, \ldots, \langle x_\ell, e_\ell \rangle\}$. The cost $c_{\text{int } \mathcal{K}}$ is majorized by $c_{\min}$, since we can decide whether x lies in interior of $\mathcal{K}$ by checking if $\lambda_{\min}(x) > 0$. So the complexity of the basic procedure is proportional to

$$\mathcal{O}\left(\ell^3 r_{\max}^2 (\max(md, c_{\min}))\right). \tag{21}$$

To conclude this section, we note that the condition "$z \neq 0$" in Line 2 of Algorithm 1 can, optionally, be substituted by "$y^i - z \notin \mathcal{K}$". The motivation is that if $y^i - z \in \mathcal{K}$, then there are two possibilities. The first is $y^i - z \neq 0$ and $y^i - z$ is a solution to (D), since $y^i - z$ belongs to the range of $\mathcal{A}^*$. The second is $y^i - z = 0$ and z is a solution to (P), since y^i belongs to int $\mathcal{K}$ throughout the algorithm. Either case, we can stop the algorithm.

Using the condition "$y^i - z \notin \mathcal{K}$" does not affect the analysis conducted in this section because $y^i - z \notin \mathcal{K}$ implies $z \neq 0$, since $y^i \in \text{int } \mathcal{K}$. Therefore, Propositions 13 and 14 still apply. The advantage is that since it is harder to satisfy than $z \neq 0$, it may lead to less iterations. The disadvantage is that, in order to verify that $y^i - z \notin \mathcal{K}$, it is necessary to do an extra minimum eigenvalue computation per iteration. This, however, does not alter the order complexity.

5 Full-procedure

In this section, we summarize everything we have done so far and present an algorithm for the feasibility problem (P), see Algorithm 2. In essence, we call the basic procedure (Algorithm 1) and if we fail to find a solution to either (P) or (D), we will receive some vector y. Following the discussion in Sect. 3 and in Theorem 12, for at least one index k, we are able to construct a half-space $H(w_k, v_k)$ such that the volume $H(w_k, v_k) \cap \mathcal{K}_k$ is less than $H(e_k, e_k/r_k) \cap \mathcal{K}_k$ and their ratio is bounded above by a constant, namely, $\exp(-\varphi(\rho_k))^{d_k/r_k} \leq (0.918)^{d_k/r_k}$. Then, we construct an automorphism of $\mathcal{K}$ that maps the region $H(e_k, e_k/r_k) \cap \mathcal{K}_k$ to $H(w_k, v_k) \cap \mathcal{K}_k$. Finally, we substitute $\mathcal{A}$ by $\mathcal{A}Q$ and repeat.

In Algorithm 2, we keep track of the volume reduction along the blocks through the ϵ_k variables. Lemma 16 tells us that the ϵ_k are upper bounds for the minimum eigenvalue of the feasible solutions to $(P_{\text{scaled}}^\mathcal{A})$. When the accumulated reduction is sufficiently small, we have a certificate that $(P_{\text{scaled}}^\mathcal{A})$ does not have ϵ-feasible solutions.

Algorithm 17 and the one in [13] have a few differences. First of all, for the same y, we attempt to find volume reductions along all blocks, where in [13], volume reductions only happen for exactly one block per iteration. Furthermore, here we adapt a greedy

---

Algorithm 2: Main Algorithm

Input: $\mathcal{A}, \mathcal{K}, \epsilon$
Output:
  1. a solution to (P), or
  2. a solution to (D), or
  3. "there is no ϵ-feasible solution".

1 $\mathcal{A}^1 \leftarrow \mathcal{A}, i \leftarrow 0$ and $\epsilon_j \leftarrow 0$ for all $j \in \{1, \ldots, \ell\}$.
2 Compute $P_{\mathcal{A}}$ and call the basic procedure (Algorithm 1) with $\mathcal{A}^i$, $P_{\mathcal{A}}$, $\frac{\epsilon}{r}$ and denote its output by y.
3 **if** y *is feasible for* (D) *or* (P) **then**
4   **stop** and **return** y.
5 **for** $k \in \{1, \ldots, \ell\}$ **do**
6 $\rho_k \leftarrow \frac{\|y_k\|_1}{r_k \|P_{\mathcal{A}} y\| \sqrt{\ell}}$
7   **if** $\rho_k > 1$ **then**
8 $\beta \leftarrow r_k - \left(\frac{1}{\rho_k} - \frac{1}{\sqrt{\rho_k(3\rho_k - 2)}} \right)$
9 $w_k \leftarrow \left(\frac{r_k - \beta}{\langle y_k, e_k \rangle} \right) \rho_k r_k y_k + \beta e_k$
10 $Q_k \leftarrow Q_{w_k^{-1/2} \sqrt{r_k}}$
11 $\epsilon_k \leftarrow \epsilon_k + \log r_k - \frac{1}{r_k} \log \det w_k$
12   **if** $\epsilon_k < \log r_k + \log \epsilon$ **then**
13   **stop**, there is no ϵ-feasible solution.
14   **else**
15 $Q_k \leftarrow I_k$
16 Let Q^i be such that for all $x \in \mathcal{E}$ we have $Q^i(x) = (Q_1(x_1), \ldots, Q_\ell(x_\ell))$.
17 $\mathcal{A}^{i+1} \leftarrow \mathcal{A}^i Q^i, i \leftarrow i + 1$. Go to Line 2.

---

approach and following Theorems 10 and 12, we try to reduce the volume as much as possible by selecting ρ_k to be as large as our analysis permits.

We will now prove a few lemmas regarding the correctness of Algorithm 2.

Lemma 15 *For every iteration i, every block $\mathcal{E}_k$ and for all $x \in \mathcal{F}^{\mathcal{A}}_{scaled}$ we have*

$$x_k \in Q_k^1 \times \cdots \times Q_k^i H(e_k, e_k/r_k).$$

Proof We will proceed by induction, so first we prove the result for $i = 1$. According to Line 7 of Algorithm 2, we have that Q_k^1 is either the identity map or $Q_k^1 = Q_{w_k^{-1/2} \sqrt{r_k}}$, where w_k is selected as in Line 9. If it is the former, it is clear that $x_k \in Q_k^1 H(e_k, e_k/r_k)$, since $\langle x_k, e_k \rangle \leq \|x\|_{1,\infty} \leq 1$. If it is the latter, we must have $\rho_k > 1$, in which case, Theorem 12 tells us that $x_k \in H(w_k, w_k^{-1})$ and that $Q_k^1 H(e_k, e_k/r_k) = H(w_k, w_k^{-1})$.

Now suppose the result is true for some $i \geq 1$, we shall prove it also holds for $i + 1$. We first let

$$u = \left(Q^i \right)^{-1} \times \ldots \times \left(Q^1 \right)^{-1} x.$$

Since all the Q^i are automorphisms of $\mathcal{K}$, we have $u \in \text{int}\,\mathcal{K}$. Now, we argue that $\|u\|_{1,\infty} \leq 1$. Since every Q^i is constructed in a block-wise manner, we have

$$u_k = \left(Q_k^i\right)^{-1} \times \ldots \times \left(Q_k^1\right)^{-1} x_k,$$

which, by the induction hypothesis, implies that $u_k \in H(e_k, e_k/r_k)$ for every k. There-fore, $\langle u_k, e_k \rangle \leq 1$ for every k, which implies $\|u\|_{1,\infty} \leq 1$. Then, following Line 17, u is a feasible solution to the problem of finding $\hat{x}$ such that $\mathcal{A}^{i+1}\hat{x} = 0$, $\|\hat{x}\|_{1,\infty} \leq 1$ and $\hat{x} \in \text{int}\,\mathcal{K}$, where

$$\mathcal{A}^{i+1} = \mathcal{A}Q^1 \times \cdots \times Q^i.$$

As before, Q_k^{i+1} is either the identity map or $Q_{w_k^{-1/2}\sqrt{r_k}}$. If it is the former, we are done. If it is the latter, it is because w_k was constructed by applying Theorem 12 to $\mathcal{A}^{i+1}, \mathcal{K}$ and y, where y is obtained at $(i+1)$-th iteration. We conclude that $u_k \in H(w_k, w_k^{-1}) = Q_k^{i+1} H(e_k, e_k/r_k)$. It follows that $(Q_k^{i+1})^{-1} u_k \in H(e_k, e_k/r_k)$. This is equivalent to $x_k \in Q_k^1 \times \cdots \times Q_k^{i+1} H(e_k, e_k/r_k)$. $\quad\square$

The next lemma justifies the stopping criteria in Line 12.

Lemma 16 *For all $x \in \mathcal{F}_{scaled}^{\mathcal{A}}$, for every iteration and for all $k \in \{1, \ldots, \ell\}$*

$$\epsilon_k \geq \log r_k + \log \lambda_{\min}(x_k)$$

holds. In particular, if $\epsilon_k < \log r_k + \log \epsilon$, then $\mathcal{F}_{scaled}^{\mathcal{A}}$ does not have an ϵ-feasible solution.

Proof We first handle the trivial case. At Line 1, ϵ_k is set to zero. Note that if $x \in \mathcal{F}_{scaled}^{\mathcal{A}}$, then $\langle x_k, e_k \rangle \leq 1$, from which it follows that $\lambda_{\min}(x_k) \leq 1/r_k$, since $\langle x_k, e_k \rangle$ is the sum of eigenvalues of x_k. Therefore, $\log r_k + \log \lambda_{\min}(x_k) \leq 0$.

Lemma 15 tells us that at the i-th iteration we have $x_k \in Q_k H(e_k, e_k/r_k)$ for all $x \in \mathcal{F}_{scaled}^{\mathcal{A}}$, where $Q_k = Q_k^1 \times \cdots \times Q_k^i$. Note that $Q_k H(e_k, e_k/r_k)$ is also a half-space. Denote the volume of $(Q_k(H(e_k, e_k/r_k))) \cap \mathcal{K}_k$ by V_k. Then, from Proposition 8, we have $V_k \geq r_k^{d_k}(\sqrt[r_k]{\det x_k})^{d_k} \text{vol}\,(e_k, e_k/r_k)$, for all $x \in \mathcal{F}_{scaled}^{\mathcal{A}}$. As $\det x_k$ is the product of eigenvalues of x_k, we get

$$V_k \geq r_k^{d_k} \lambda_{\min}(x_k)^{d_k} \text{vol}\,(e_k, e_k/r_k), \tag{22}$$

for all $x \in \mathcal{F}_{scaled}^{\mathcal{A}}$. Furthermore, we have $V_k = \text{vol}\,(e_k, e_k/r_k) \prod_{j=1}^{i} \det Q_k^j$ and $\det Q_k^j$ is either 1 or

$$\left(\frac{r_k}{\sqrt[r_k]{\det w_k}}\right)^{d_k}$$

depending on whether at the j-th iteration we had $\rho_k \leq 1$ or $\rho_k > 1$, respectively. Taking the logarithm at both sides of (22), we conclude that

$$\epsilon_k \geq \log r_k + \log \lambda_{\min}(x_k).$$

Therefore, if $\epsilon_k < \log r_k + \log \epsilon$, we get $\epsilon > \lambda_{\min}(x_k) \geq \lambda_{\min}(x)$, for all $x \in \mathcal{F}^{\mathcal{A}}_{\text{scaled}}$. This implies the absence of ϵ-feasible solutions. □

For what follows, we discard the (trivial) case where ϵ is large. Note that if $\epsilon \geq \frac{1}{r_k}$ for some k, then for $x \in \mathcal{F}^{\mathcal{A}}_{\text{scaled}}$ we have $\lambda_{\min}(x) \leq \lambda_{\min}(x_k) \leq \epsilon$, since the sum of the eigenvalues of x_k is less or equal than $1/r_k$. In this case, there is no work to be done as this shows that $\mathcal{F}^{\mathcal{A}}_{\text{scaled}}$ has no ϵ-feasible solution. So in the next result, we suppose that $\epsilon < \frac{1}{r_k}$ holds for every k.

Theorem 17 *Algorithm 2 stops after no more than*

$$\frac{r}{\varphi(2)} \log\left(\frac{1}{\epsilon}\right) - \sum_{k=1}^{\ell} \frac{r_k \log(r_k)}{\varphi(2)}$$

iterations, where $\varphi(\rho)$ is as in Theorem 12. In particular, $\frac{r}{\varphi(2)} \log\left(\frac{1}{\epsilon}\right)$ iterations are enough.

Proof At each iteration, after Algorithm 1 is called we obtain some point y. If y is neither feasible for (P) nor (D), then it satisfies $\|P_{\mathcal{A}}y\| \leq \frac{1}{2r_{\max}\sqrt{\ell}} \|y\|_{1,\infty}$.

Due to item (ii) of Lemma 11, there is at least one index k for which $\rho_k \geq 2$. Then from Theorem 12, we have that

$$\log r_k - \frac{1}{r_k} \log \det w_k \leq -\frac{\varphi(2)}{r_k} < 0.$$

This means that ϵ_k decreases by at least $\frac{\varphi(2)}{r_k}$. Let us say that "an iteration is good for $\mathcal{K}_k$" if at that iteration we have $\rho_k \geq 2$. From Lemma 16, it follows that we need no more than

$$\frac{r_k}{\varphi(2)} \log\left(\frac{1}{\epsilon r_k}\right).$$

"good iterations for $\mathcal{K}_k$" before the stopping criteria in Line 12 is satisfied. In particular, after

$$\sum_{k=1}^{\ell} \frac{r_k}{\varphi(2)} \log\left(\frac{1}{\epsilon}\right) - \frac{r_k \log(r_k)}{\varphi(2)} = \frac{r}{\varphi(2)} \log\left(\frac{1}{\epsilon}\right) - \sum_{k=1}^{\ell} \frac{r_k \log(r_k)}{\varphi(2)}.$$

iterations, we will meet the minimum number of good iterations for at least one cone. We can then discard the negative terms and obtain the bound $\frac{r}{\varphi(2)} \log\left(\frac{1}{\epsilon}\right)$ □

We now take a look at the cost per iteration of Algorithm 2. The two most expensive operations are computing $P_{\mathcal{A}}$, calling the basic procedure (Line 2) and computing the square root of the inverse of w_k (Line 10). As discussed in Sect. 4, the cost of computing the projection $P_{\mathcal{A}}$ is no more than $\mathcal{O}(m^3) + \mathcal{O}(m^2 d)$.

Furthermore, when $\mathcal{K}_k$ is $\mathbb{R}_+, \mathcal{L}_{d_k}$ or $\mathcal{S}_+^{r_k}$, we need respectively no more than $\mathcal{O}(1), \mathcal{O}(d_k), \mathcal{O}(r_k^3) = \mathcal{O}(d_k^{3/2})$ for approximating the inverse of the square root of w_k. So, it seems that in most cases of interest it will cost no more than $\mathcal{O}(d_k^{3/2})$ to perform Line 10. As we might need to compute $w_k^{-1/2}$ for all blocks, this will cost, in total, no more than $\mathcal{O}(d^{3/2})$.

Therefore, the cost of the basic procedure dominates the cost of computing $w_k^{-1/2}$, see (21). In total, we get

$$\mathcal{O}\left(\frac{r}{\varphi(2)} \log\left(\frac{1}{\epsilon}\right)\left(m^3 + m^2 d + \ell^3 r_{\max}^2(\max(md, c_{\min}))\right)\right),$$

where $c_{\min}$ is the cost of computing the minimum eigenvalue of an element $x \in \mathcal{K}$.

6 Final remarks

In this work, we presented a generalization of Chubanov's algorithm to feasibility problems over symmetric cones. A next interesting step would be to try to extend the algorithm to other families of cones. A key aspect of our algorithm is being able to shift the hyperplane $H(w, v) \cap \mathcal{K}$ back to $H(e, e/r)$, with the aid of some appropriately constructed Q_x. The reason why this can always be done is because $\mathcal{K}$ is a *homogeneous cone*, which means that for every two points $x, y \in \operatorname{int} \mathcal{K}$ there is an automorphism Q of $\mathcal{K}$ such that $Q(x) = y$. It seems likely, although nontrivial, that Chubanov's algorithm can be extended to general homogeneous cone in some form or another. However, it less clear whether such an extension is possible for non-homogeneous cones.

From a more practical side, there is a still a lack of computational results for semidefinite programming and second order cone programming. It will be interesting to take a look at whether Chubanov's algorithm could be competitive with IPMs codes when it comes to feasibility problems.

A significant bottleneck in the algorithm presented here is the computation of the projection matrix. Here, we are considering a simple approach where the $P_{\mathcal{A}^i}$ is computed from scratch every time the basic procedure is called. One possible way to address this is using an incremental strategy analogous to what Peña and Soheili suggested in Section 6 of [23]. We plan to address this topic in detail in future works.

There are also a number of intriguing aspects yet to be elucidated. First of all, in our analysis, the self-concordance of the function $-\log\det$ plays a very important role in guaranteeing that the reductions in volume are bounded by a constant. This usage indicates that there could be a deeper connection to interior-point methods (IPMs) then what is suggested by previous discussions.

Related to that, one of the referee's suggested that Chubanov's algorithm and our extension could be related to analytic center cutting plane methods, in particular the

algorithms described by Sun et al. [29] and by Toh et al. [29,30], which solve feasibility problems over the positive semidefinite matrices. These methods were partly inspired by earlier work on linear feasibility problems by Luo and Sun [16] and Goffin and Vial [9]. They were later adapted to second order cone programming by Oskoorouchi and Goffin [22] and extended to symmetric cone programming by Basescu and Mitchell [2]. Here we briefly explain the basic setting and try to explain the similarities and differences between our approach and analytic center methods.

We consider a convex body Γ contained in $\mathcal{K}$ that is only accessible by querying some oracle. Given some $x \in \mathcal{E}$ the oracle either tell us that $x \in \Gamma$ or provide a cut y such that Γ is guaranteed to be contained in the half-space $H_1 = \{z \in \mathcal{E} \mid \langle z, y \rangle \leq \langle x, y \rangle\}$. We also assume that Γ is contained in the "spectral interval" $\Omega_0 = \{z \in \mathcal{E} \mid z \in \mathcal{K}, e - z \in \mathcal{K}\}$ and that it contains some ϵ-ball. Note that these assumptions imply $\Gamma \cap \operatorname{int} \mathcal{K} \neq \emptyset$ and that Γ is full-dimensional. A basic analytic center cutting plane algorithm would proceed by first guessing some initial point x_0. Then, the oracle is queried and either x_0 is feasible or a cut is returned. If a cut is returned, then we let x_1 be the analytic center of $\Omega_1 := \Omega_0 \cap H_0$ or some appropriate approximation. Then, we test x_1 for feasibility and if we receive a cut y_1 from the oracle, we let x_2 be the analytic center of $\Omega_2 := \Omega_0 \cap H_0 \cap H_1$ or some approximation. We then repeat until a feasible solution is found or some other stopping criterion is met. Here, we recall that the analytic center is the minimizer of a function that is constructed using the barrier $-\log \det(\cdot)$, see Section 2 in [29].

As for the similarities and differences, first of all, it seems that the setting is slightly different. For example, neither the feasible region of (P) nor ($\mathrm{P}^{\mathcal{A}}_{\mathrm{scaled}}$) can be expected to contain some ϵ-ball. Nevertheless, the basic procedure could be seen as an oracle for the feasible region of ($\mathrm{P}^{\mathcal{A}}_{\mathrm{scaled}}$), in the sense that it returns a cut y when the procedure is not able to prove or disprove the feasibility of ($\mathrm{P}^{\mathcal{A}}_{\mathrm{scaled}}$). This cut can be used to define half-spaces that contain some of the blocks of the feasible region of ($\mathrm{P}^{\mathcal{A}}_{\mathrm{scaled}}$), see item ($ii$) of Lemma 11. However, the basic procedure is more than a simple oracle and proactively tries to improve the point received as input. Then, after y is found we try to improve it by decreasing a function constructed around $-\log \det(\cdot)$, as in Theorem 12. It might be possible that the update in Theorem 12 could be seen as trying to approximate the analytic center of some set that contains the feasible region of ($\mathrm{P}^{\mathcal{A}}_{\mathrm{scaled}}$), but, at this moment, it is not clear to us the appropriate way of establishing such a connection. Furthermore, the step where $\mathcal{A}^i$ is updated to $\mathcal{A}^i Q^i$ (Line 17 in Algorithm 2) does not seem to have a clear counterpart in analytic center methods. Nevertheless, we think it could be an interesting topic of future research to take a closer look at the relationship between the methods.

A The case of semidefinite programming

For ease of reference, we "translate" here Algorithms 1 and 2 for the case of semidefinite programming, see Algorithms 3 and 4. In this case, we have $\ell = 1$, $r = n$, $\mathcal{K} = \mathcal{S}^n_+$ and $\mathcal{E}$ is the space of $n \times n$ symmetric matrices $\mathcal{S}^n$. We recall that if $y \in \mathcal{S}^n_+$, then $\|y\|_1 = \langle e, y \rangle = \operatorname{tr} y$, so the stopping criteria in Algorithm 1 becomes $\|z\| \leq \frac{1}{2n} \operatorname{tr} y$. We will use I_n to denote the $n \times n$ identity matrix. We will write $y \succeq 0$, if y is positive semidefinite and $y \succ 0$, if y is positive definite.

Algorithm 3: Basic Procedure - SDP

Input: $\mathcal{A}$, $P_\mathcal{A}$, $y^1 \succ 0$ such that $\operatorname{tr} y^1 = 1$
Output:
 1. $y \succeq 0$ such that $\|P_\mathcal{A} y\| \leq \frac{1}{2n} \|y\|_1$, or
 2. a solution to (P), or
 3. a solution to (D).

1 $i \leftarrow 1$, $z \leftarrow P_\mathcal{A} y^1$.
2 **while** $z \neq 0$ *and* $z \not\succ 0$ *and* $\|z\| > \frac{1}{2n} \operatorname{tr} y^i$ **do**
3   Let v be an eigenvector associated to the minimum eigenvalue of z and such that $\|v\| = 1$. Let
     $c \leftarrow vv^\top$.
4      $p \leftarrow P_\mathcal{A} c$.
5   **if** $p = 0$ *or* $p \succ 0$ *or* $\|p\| \leq \frac{1}{2n}$ **then**
6   **stop** and **return** c.
7   **else**
8          $y^{i+1} \leftarrow \alpha y^i + (1 - \alpha)c$, where $\alpha = \frac{\langle p, p-z \rangle}{\|z-p\|^2}$.
9   Set $i \leftarrow i + 1$ and $z \leftarrow P_\mathcal{A} y^i$.
10 **return** y^i.

Algorithm 4: Main Algorithm - SDP

Input: $\mathcal{A}$, $\mathcal{S}^n_+$, ϵ
Output:
 1. a solution to (P), or
 2. a solution to (D), or
 3. "there is no ϵ-feasible solution".

1 $\mathcal{A}^1 \leftarrow \mathcal{A}$, $i \leftarrow 1$ and $\tilde{\epsilon} \leftarrow 0$.
2 Compute $P_\mathcal{A}$ and call the basic procedure (Algorithm 3) with $\mathcal{A}^i$, $P_\mathcal{A}$, $\frac{I_n}{n}$ and denote its output by
   y.
3 **if** y *is feasible for* (D) *or* (P) **then**
4   **stop** and **return** y.
5 $\rho \leftarrow \frac{\operatorname{tr} y}{n \|P_\mathcal{A} y\|}$.
6 $\beta \leftarrow n - \left(\frac{1}{\rho} - \frac{1}{\sqrt{n(3n-2)}} \right)$.
7 $w \leftarrow \left(\frac{n-\beta}{\operatorname{tr} y} \right) \rho n y + \beta I_n$.
8 $\tilde{\epsilon} \leftarrow \tilde{\epsilon} + \log n - \frac{1}{n} \log \det w$.
9 **if** $\tilde{\epsilon} < \log n + \log \epsilon$ **then**
10   **stop**, there is no ϵ-feasible solution.
11 $\mathcal{A}^{i+1} \leftarrow \mathcal{A}^i Q_{w^{-1/2} \sqrt{n}}$, $i \leftarrow i + 1$. Go to Line 2.

Furthermore, for $w \in \mathcal{S}_+^n$, the quadratic map Q_w is the function that takes $x \in \mathcal{S}^n$ to wxw. We do not need, in fact, to explicitly construct the operator Q_w. In particular, if $\mathcal{A}x = 0$ is represented as the set of solutions of m linear equalities $\mathrm{tr}\,(a_1 x) = 0, \ldots, \mathrm{tr}\,(a_m x) = 0$, the first assignment in Line 11 of Algorithm 4 is the same as substituting every a_i by $n(w^{-1/2} a_i w^{-1/2})$. Finally, as remarked at the end of Sect. 4, the condition $z \neq 0$ in Algorithm 3 can be substituted by $y^i - z \nsucceq 0$.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. **95**(1), 3–51 (2003)
2. Basescu, V.L., Mitchell, J.E.: An analytic center cutting plane approach for conic programming. Math. Oper. Res. **33**(3), 529–551 (2008)
3. Chubanov, S.: A strongly polynomial algorithm for linear systems having a binary solution. Math. Program. **134**(2), 533–570 (2012)
4. Chubanov, S.: A polynomial projection algorithm for linear feasibility problems. Math. Program. **153**(2), 687–713 (2015)
5. Faraut, J., Korányi, A.: Analysis on Symmetric Cones. Oxford Mathematical Monographs. Clarendon Press, Oxford (1994)
6. Faybusovich, L.: Euclidean Jordan algebras and interior-point algorithms. Positivity **1**(4), 331–357 (1997)
7. Faybusovich, L.: A Jordan-algebraic approach to potential-reduction algorithms. Math. Z. **239**(1), 117–129 (2002)
8. Faybusovich, L.: Several Jordan-algebraic aspects of optimization. Optimization **57**(3), 379–393 (2008)
9. Goffin, J.-L., Vial, J.-P.: Convex nondifferentiable optimization: a survey focused on the analytic center cutting plane method. Optim. Methods Softw. **17**(5), 805–867 (2002)
10. Hirzerbruch, U.: Der min–max-satz von E. Fischer für formal-reelle Jordan-algebren. Math. Ann. **186**, 65–69 (1970)
11. Karmarkar, N.: A new polynomial-time algorithm for linear programming. Combinatorica **4**(4), 373–395 (1984)
12. Khachiyan, L.: Polynomial algorithms in linear programming. USSR Comput. Math. Math. Phys. **20**(1), 53–72 (1980)
13. Kitahara, T., Tsuchiya, T.: An extension of Chubanov's polynomial-time linear programming algorithm to second-order cone programming. Optim. Methods Softw. (2017). https://doi.org/10.1080/10556788.2017.1382495
14. Li, D., Roos, K., Terlaky, T.: A polynomial column-wise rescaling von Neumann algorithm. Optimization Online. http://www.optimization-online.org/DB_HTML/2015/06/4979.html (June 2015)
15. Luo, Z., Sturm, J.F., Zhang, S.: Duality results for conic convex programming. Technical Report, Econometric Institute, Erasmus University Rotterdam, The Netherlands (1997)
16. Luo, Z.-Q., Sun, J.: A polynomial cutting surfaces algorithm for the convex feasibility problem defined by self-concordant inequalities. Comput. Optim. Appl. **15**(2), 167–191 (2000)
17. Maxima: Maxima, a computer algebra system. Version 5.36.1. http://maxima.sourceforge.net/ (2015)
18. Monteiro, R.D., Tsuchiya, T.: Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions. Math. Program. **88**(1), 61–83 (2000)
19. Muramatsu, M.: On a commutative class of search directions for linear programming over symmetric cones. J. Optim. Theory Appl. **112**(3), 595–625 (2002)
20. Nemirovski, A.S., Todd, M.J.: Interior-point methods for optimization. Acta Numer. **17**, 191–234 (2008)
21. Nesterov, Y., Nemirovskii, A.: Interior-Point Polynomial Algorithms in Convex Programming. Society for Industrial and Applied Mathematics, Philadelphia (1994)
22. Oskoorouchi, M.R., Goffin, J.-L.: An interior point cutting plane method for the convex feasibility problem with second-order cone inequalities. Math. Oper. Res. **30**(1), 127–149 (2005)
23. Peña, J., Soheili, N.: Solving conic systems via projection and rescaling. Mathematical Programming, e-prints (to appear) (Dec. 2015). arXiv:1512.06154
24. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1997)

25. Roos, K.: An improved version of Chubanov's method for solving a homogeneous feasibility problem. Optimization Online. http://www.optimization-online.org/DB_HTML/2016/11/5745.html (2016)
26. Schmieta, S., Alizadeh, F.: Extension of primal-dual interior point algorithms to symmetric cones. Math. Program. **96**(3), 409–438 (2003)
27. Soheili, N., Peña, J.: A smooth perceptron algorithm. SIAM J. Optim. **22**(2), 728–737 (2012)
28. Sturm, J.F.: Similarity and other spectral relations for symmetric cones. Linear Algebra Appl. **312**(1–3), 135–154 (2000)
29. Sun, J., Toh, K.-C., Zhao, G.: An analytic center cutting plane method for semidefinite feasibility problems. Math. Oper. Res. **27**(2), 332–346 (2002)
30. Toh, K.-C., Zhao, G., Sun, J.: A multiple-cut analytic center cutting plane method for semidefinite feasibility problems. SIAM J. Optim. **12**(4), 1126–1146 (2002)
31. Tsuchiya, T.: A convergence analysis of the scaling-invariant primal-dual path-following algorithms for second-order cone programming. Optim. Methods Softw. **11**(1–4), 141–182 (1999)