Normalizations of factorizations over convex cones and their effects on extension complexity

Adam Brown[1†], Kanstantsin Pashkovich [2*†], Levent Tunçel [2†]

[1]School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, 30318, Georgia, United States.
[2]Department for Combinatorics and Optimization, University of Waterloo, 200 University Avenue West, Waterloo, N2L 3G1, Ontario, Canada.

*Corresponding author(s). E-mail(s): kpashkov@uwaterloo.ca;
Contributing authors: ajmbrown@gatech.edu;
levent.tuncel@uwaterloo.ca;
[†]These authors contributed equally to this work.

Abstract

Factorizations over cones and their duals play central roles for many areas of mathematics and computer science. One of the reasons behind this is the ability to find a representation for various objects using a well-structured family of cones, where the representation is captured by the factorizations over these cones. Several major questions about factorizations over cones remain open even for such well-structured families of cones as non-negative orthants and positive semidefinite cones. Having said that, we possess a far better understanding of factorizations over non-negative orthants and positive semidefinite cones than over other families of cones. One of the key properties that led to this better understanding is the ability to normalize factorizations, i.e., to guarantee that the norms of the vectors involved in the factorizations are bounded in terms of an input and in terms of a constant dependent on the given cone. Our work aims at understanding which cones guarantee that factorizations over them can be normalized, and how this effects extension complexity of polytopes over such cones.

1 Introduction

A convex cone is called *regular*, if it is closed, pointed and has nonempty interior. Let $\mathcal{C}$ be a regular convex cone in $\mathbb{R}^n$, then its *dual* $\mathcal{C}^\star$ is defined as

$$\mathcal{C}^\star := \{x \in \mathbb{R}^n \,:\, \langle x,\, y \rangle \geq 0 \quad \text{for all} \quad y \in \mathcal{C}\}\,.$$

If $\mathcal{C}$ is a regular convex cone then so is its dual cone. *Factorization problem over a cone* (and its dual) is described in [1, Definition 2.2]. Factorizations over cones play an important role for several areas of mathematics and computer science [2], [3], [4], [5],[6], etc. Intuitively, non-negative factorizations (and the related nonnegative rank) allow to compress information about features of various objects to fewer key features. In this spirit, Yannakakis established the connection between the nonnegative rank of slack matrices and "compressing" polytopes. In particular, the work of Yannakakis [7] showed that so-called *linear extension complexity* of a nontrivial polytope equals the *nonnegative rank* of its slack matrix.

Now we know explicit families of 0/1-polytopes that do not admit any linear or positive semidefinite extension of polynomial size [8], [9] and [10]. Surprisingly, for several decades after the work of Yannakakis [7], it was an open question whether every 0/1-polytope in $\mathbb{R}^n$ has linear extension complexity that is polynomial in n. For the first time this question was answered in negative by Rothvoss [11]. The arguments in the work of Rothvoss [11] are based on counting and can be starightforwardly extended to stronger collections of extensions, as long as these collections admit *a normalization of factorizations*.

In their work on positive semidefinite cones, Briët et al. [12] provided a negative answer for a broader class of extensions known as positive semidefinite extensions by introducing a normalization approach. Subsequent research by Fawzi et al. [6] expanded on this by presenting another normalization method using John-Löwner ellipsoids. Averkov et al. [13] similarly employed these ellipsoids to analyze the semidefinite extension complexity of polytopes.

The utility of John-Löwner ellipsoids for normalization has proven versatile, finding applications beyond the analysis of positive semidefinite cone extension complexity, such as in Banach space theory, as mentioned in the work [14, Lemma 4.2]. Building on these foundations, our research aims to extend normalization results to more general convex cones, seeking to broaden our understanding of cone normalizations.

Interestingly, while our arguments in Section 3 may initially appear very distinct, they nonetheless share underlying connections to the problem of constructing structured minimum volume ellipsoids that contain specific sets.

The limitations of extended formulations were studied also in the context of polytopes that are not necessarily 0/1. The normalization of factorizations plays an important role for providing lower bounds on extension complexity for various families

of polytopes. For example, the extension complexity of even very small-dimensional polytopes is not fully understood at the moment. In particular, for polytopes of dimension two, first non-trivial lower bounds for linear extension complexity were provided in [8]: a lower bound of $\Omega(\sqrt{n})$ for a general n-gon and $\Omega(\sqrt{n}/\log(n))$ for an integral n-gon, where the last result was building on the normalization and counting ideas from [11]. There is a gap between these lower bounds and currently best upper bounds on extensions complexity of polygons. Recently, Shitov[15] provided a sublinear upper bound of $O(n^{2/3})$ for linear extension complexity of general n-gons. In some randomized constructions for polytopes with n vertices and with a constant dimension, in [16] it was shown that the linear extension complexity is asymptotically almost surely equal to $\Theta(\sqrt{n})$. Thus, even for polytopes of dimension two, many questions about linear extension complexity remain open.

1.1 Problem of Normalizing Factorizations

Before further discussion, let us formulate the question about normalization of convex cones. Given a regular convex cone $\mathcal{C}$ we say that it is possible to *normalize* factorizations over this cone and its dual with a constant $f_{\mathcal{C}} \in \mathbb{R}_+$ if for every pair of compact sets $\mathcal{A} \subseteq \mathcal{C}$ and $\mathcal{B} \subseteq \mathcal{C}^\star$ there exists a pair of sets $\widetilde{\mathcal{A}} \subseteq \mathcal{C}$ and $\widetilde{\mathcal{B}} \subseteq \mathcal{C}^\star$ and two bijections $g : \mathcal{A} \to \widetilde{\mathcal{A}}$, $q : \mathcal{B} \to \widetilde{\mathcal{B}}$ such that

$$\langle a, b \rangle = \langle g(a), q(b) \rangle \qquad \text{for all} \quad a \in \mathcal{A}, b \in \mathcal{B} \tag{1}$$

$$\|\widetilde{a}\|_2 \le f_{\mathcal{C}} \sup_{a \in \mathcal{A}, b \in \mathcal{B}} \langle a, b \rangle \qquad \text{for all} \quad \widetilde{a} \in \widetilde{\mathcal{A}} \tag{2}$$

$$\|\widetilde{b}\|_2 \le f_{\mathcal{C}} \sup_{a \in \mathcal{A}, b \in \mathcal{B}} \langle a, b \rangle \qquad \text{for all} \quad \widetilde{b} \in \widetilde{\mathcal{B}}. \tag{3}$$

Informally, we will also say that in the above case the sets $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ are being *rescaled*.

In Section 2, we prove that if a regular convex cone admits normalization of factorizations then such a normalization can be achieved via automorphisms of the cone. In Section 3, for indecomposable symmetric cones $\mathcal{C}$, we relate the normalization constant $f_{\mathcal{C}}$ to the Carathéodory number of $\mathcal{C}$. More generally, for indecomposable homogeneous cones $\mathcal{C}$, we bound the normalization constant $f_{\mathcal{C}}$ in terms of the barrier parameter of certain self-concordant barriers for $\mathcal{C}$. While we prove that normalization of factorizations is possible over every homogeneous cone, we show that normalization of factorizations is impossible for even some elementary hyperbolic cones. In Section 4, we discuss the consequences of normalization of factorizations on extension complexity.

2 Normalization and Automorphisms

In this section, we show that if a regular convex cone allows a normalization of factorizations, then there exists a normalization induced by automorphisms of this cone and its dual. First, we show that if we are rescaling sets $\mathcal{A} \subseteq \mathcal{C}$, $\mathcal{B} \subseteq \mathcal{C}^\star$ which span $\mathbb{R}^n$,

then the bijections $g : \mathcal{A} \to \widetilde{\mathcal{A}}$, $q : \mathcal{B} \to \widetilde{\mathcal{B}}$ as in Section 1.1 will always be invertible linear maps.

Lemma 1. *Let $\mathcal{C}$ be a regular convex cone and let $\mathcal{C}^*$ be its dual. Let $\mathcal{A}$ and $\mathcal{B}$ be subsets of $\mathcal{C}$ and $\mathcal{C}^*$ respectively. If $\mathcal{A}$ and $\mathcal{B}$ both span $\mathbb{R}^n$ and there exist two sets $\widetilde{\mathcal{A}} \subseteq \mathcal{C}$ and $\widetilde{\mathcal{B}} \subseteq \mathcal{C}^\star$ and two bijections $g : \mathcal{A} \to \widetilde{\mathcal{A}}$, $q : \mathcal{B} \to \widetilde{\mathcal{B}}$ such that*

$$\langle a,\, b \rangle = \langle g(a),\, q(b) \rangle \qquad for\ all \quad a \in \mathcal{A},\, b \in \mathcal{B},$$

then g and q are invertible linear maps. Moreover, q is the inverse-adjoint of g.

Proof. Let us select n linearly independent vectors $a_1, a_2, \ldots, a_n$ in $\mathcal{A}$ and n linearly independent vectors $b_1, b_2, \ldots, b_n$ in $\mathcal{B}$. Let us construct matrices $R, \widetilde{R}, L, \widetilde{L} \in \mathbb{R}^{n \times n}$, where for each $i \in \{1, 2, \ldots, n\}$ the matrices $R, \widetilde{R}, L$ and $\widetilde{L}$ have as i-th row the vector $a_i^\top$, $g(a_i)^\top$, $b_i^\top$ and $q(b_i)^\top$, respectively. Clearly, the matrices R and L are non-singular. The matrices $\widetilde{R}$ and $\widetilde{L}$ are also non-singular, since $R \cdot L^\top$ is a non-singular matrix and $R \cdot L^\top = \widetilde{R} \cdot \widetilde{L}^\top$.

For all $a \in \mathcal{A}$ we have $L \cdot a = \widetilde{L} \cdot g(a)$, so for all $a \in \mathcal{A}$ we can write $g(a) = \widetilde{L}^{-1} \cdot L \cdot a$, showing that g is a linear map. Analogously, for all $b \in \mathcal{B}$ we have $q(b) = \widetilde{R}^{-1} \cdot R \cdot b$, showing that q is a linear map.

From the formula $R \cdot L^\top = \widetilde{R} \cdot \widetilde{L}^\top$, we can rearrange to get

$$R^{-1} \cdot \widetilde{R} = L^\top \cdot (\widetilde{L})^{-\top} = \left(L^{-1} \cdot \widetilde{L} \right)^{-\top}.$$

Since we noted that $g(a) = \widetilde{L}^{-1} \cdot L \cdot a$ and $q(b) = \widetilde{R}^{-1} \cdot R \cdot b$, we conclude that q is the inverse-adjoint of g. $\square$

Now we show that if any compact subset of a given cone can be normalized, then they can be normalized by automorphisms of the cone itself.

Theorem 2. *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a regular convex cone. Suppose there exists a constant $f_{\mathcal{C}} \in \mathbb{R}_+$ such that for every pair of compact sets $\mathcal{A} \subseteq \mathcal{C}$ and $\mathcal{B} \subseteq \mathcal{C}^\star$, there exist a pair of sets $\widetilde{\mathcal{A}} \subseteq \mathcal{C}$ and $\widetilde{\mathcal{B}} \subseteq \mathcal{C}^\star$ and a pair of bijections $g : \mathcal{A} \to \widetilde{\mathcal{A}}$, $q : \mathcal{B} \to \widetilde{\mathcal{B}}$ such that (1), (2), (3) hold. Then, there exist bijections $g : \mathcal{A} \to \widetilde{\mathcal{A}}$, $q : \mathcal{B} \to \widetilde{\mathcal{B}}$ such that (1), (2), (3) hold and g is an automorphism of $\mathcal{C}$ and q is an automorphism of $\mathcal{C}^\star$.*

Before we start the proof of Theorem 2, let us outline the idea behind it. In the proof, we construct $\mathcal{A}' \subseteq \mathcal{C}$ and $\mathcal{B}' \subseteq \mathcal{C}^\star$ such that $\mathcal{A} \subseteq \mathcal{A}'$, $\mathcal{B} \subseteq \mathcal{B}'$ and

$$\sup_{a \in \mathcal{A}, b \in \mathcal{B}} \langle a,\, b \rangle = \sup_{a' \in \mathcal{A}', b' \in \mathcal{B}'} \langle a',\, b' \rangle.$$

Since we assume that it is possible to normalize factorizations over $\mathcal{C}$ and $\mathcal{C}^\star$ with respect to the constant $f_{\mathcal{C}}$, there will exist a pair of sets $\widetilde{\mathcal{A}}' \subseteq \mathcal{C}$ and $\widetilde{\mathcal{B}}' \subseteq \mathcal{C}^\star$ and bijections $g' : \mathcal{A}' \to \widetilde{\mathcal{A}}'$, $q' : \mathcal{B}' \to \widetilde{\mathcal{B}}'$ such that such that the analogues of (1), (2), (3) hold for these sets. Clearly, then the restriction of g' to $\mathcal{A}$ and the restriction of q' to $\mathcal{B}$ can be used to define a normalization for a factorization corresponding to $\mathcal{A}$ and $\mathcal{B}$. The structure of the constructed $\mathcal{A}'$ and $\mathcal{B}'$ forces the bijections g' and q' to be an automorphism of $\mathcal{C}$ and an automorphism of $\mathcal{C}^\star$, respectively.

Proof of Theorem 2. First of all, let us consider a special case, when $\sup_{a \in \mathcal{A}, b \in \mathcal{B}} \langle a, b \rangle$ equals 0. In this case, both g and q can be defined as $g : x \mapsto \varepsilon \cdot x$ and $q : x \mapsto \varepsilon \cdot x$ for a sufficiently small ε, $\varepsilon > 0$, because both $\mathcal{A}$ and $\mathcal{B}$ are compact and so bounded. Clearly, for every cone such maps g and q are automorphisms. In the rest of the proof, we assume that $\gamma := \sup_{a \in \mathcal{A}, b \in \mathcal{B}} \langle a, b \rangle$ is strictly larger than 0.

Since $\mathcal{A}$ and $\mathcal{B}$ are compact sets, there exists M, $M > 0$ such that

$$\|a\|_2 \leq M \qquad \text{and} \qquad \|b\|_2 \leq M \qquad \text{for all} \quad a \in \mathcal{A}, b \in \mathcal{B}.$$

Now let us define $\mathcal{A}' \subseteq \mathcal{C}$ and $\mathcal{B}' \subseteq \mathcal{C}^\star$

$$\mathcal{A}' := \mathcal{A} \cup \left\{ x \in \mathcal{C} \, : \, \|x\|_2 \leq \min\left\{ \frac{\gamma}{M}, \sqrt{\gamma} \right\} \right\}$$
$$\mathcal{B}' := \mathcal{B} \cup \left\{ x \in \mathcal{C}^\star \, : \, \|x\|_2 \leq \min\left\{ \frac{\gamma}{M}, \sqrt{\gamma} \right\} \right\} .$$

It is straightforward to check that both $\mathcal{A}'$ and $\mathcal{B}'$ are compact sets and that $\sup_{a \in \mathcal{A}, b \in \mathcal{B}} \langle a, b \rangle$ equals $\sup_{a' \in \mathcal{A}', b' \in \mathcal{B}'} \langle a', b' \rangle$.

By assumption, there exist bijections $g' : \mathcal{A}' \to \widetilde{\mathcal{A}}'$ and $q' : \mathcal{B}' \to \widetilde{\mathcal{B}}'$ satisfying the analogues of (1), (2), (3). Since $\mathcal{C}$ is full dimensional the sets $\mathcal{A}'$ and $\mathcal{B}'$ are spanning, so Lemma 1 implies that g' and q' are invertible linear maps.

Now let us show that the bijections g' and q' are automorphisms of $\mathcal{C}$ and $\mathcal{C}^\star$, respectively. For this, it is enough to show $g'(\text{cone}(\mathcal{A}')) = \mathcal{C}$ and $q'(\text{cone}(\mathcal{B}')) = \mathcal{C}^\star$, since $\text{cone}(\mathcal{A}') = \mathcal{C}$ and $\text{cone}(\mathcal{B}') = \mathcal{C}^\star$. By definition, we have $g'(\text{cone}(\mathcal{A}')) \subseteq \mathcal{C}$ and $q'(\text{cone}(\mathcal{B}')) \subseteq \mathcal{C}^\star$. So it is enough to show that $g'(\text{cone}(\mathcal{A}')) \supseteq \mathcal{C}$ and $q'(\text{cone}(\mathcal{B}')) \supseteq \mathcal{C}^\star$. Let us assume, that we have $g'(\text{cone}(\mathcal{A}')) \not\supseteq \mathcal{C}$, i.e. let us assume that there exists $\widetilde{a}' \in \mathcal{C}$ such that $\widetilde{a}'$ is not in $g'(\text{cone}(\mathcal{A}'))$. The preimage $(g')^{-1}(\widetilde{a}')$ has a nonnegative scalar product with all vectors $b' \in \mathcal{B}'$. Since $\text{cone}(\mathcal{B}') = \mathcal{C}^\star$, we have that $(g')^{-1}(\widetilde{a}')$ is in $\mathcal{C}$. However, since $\widetilde{a}'$ is not in $g'(\text{cone}(\mathcal{A}'))$, we know that $(g')^{-1}(\widetilde{a}')$ is not in $\text{cone}(\mathcal{A}') = \mathcal{C}$, contradiction. Analogously, we have $q'(\text{cone}(\mathcal{B}')) \supseteq \mathcal{C}^\star$, finishing the proof. □

The next remark is due to the fact that to normalize a factorization over a Cartesian product of two cones, we can normalize the factorizations over each of these cones separately.

Remark 3. *Let us be given two regular convex cones $\mathcal{C}$ and $\mathcal{D}$, such that it is possible to normalize factorizations with respect to them with constants $f_\mathcal{C}$ and $f_\mathcal{D}$, respectively. Then, for the cone $\mathcal{C} \times \mathcal{D}$ it is possible to normalize factorizations with a constant $\max(f_\mathcal{C}, f_\mathcal{D})$.*

3 Normalization of Cones

In our normalization of the cones, we will construct the scaling automorphisms by utilizing a convex optimization approach via some special classes of strictly convex functions whose domains are the interiors of regular convex cones in consideration. We start with the so-called *ϑ-logarithmically homogeneous self-concordant barrier*

(LHSCB) F (see, Definition 2.3.2 in [17]). Given a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, we define the *Legendre-Fenchel conjugate* of f, $f_* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ as

$$f_*(v) := \sup_{x \in \mathbb{R}^n} \{-\langle v, x \rangle - f(x)\}.$$

In the next lemma, we collect some of the properties of ϑ-LHSCB.

Lemma 4. *Every regular convex cone $\mathcal{C} \subset \mathbb{R}^n$ admits a ϑ-LHSCB F, where $1 \le \vartheta \le n$. Moreover, every ϑ-LHSCB F for $\mathcal{C}$ satisfies the following:*

(a) For every sequence $\{x^{(k)}\} \subset \text{int}(\mathcal{C})$ such that $x^{(k)} \to \bar{x} \in \text{bd}(\mathcal{C})$, we have

$$F\left(x^{(k)}\right) \to +\infty \qquad \text{and} \qquad \left\| \nabla F\left(x^{(k)}\right) \right\| \to +\infty;$$

(b) $\langle -\nabla F(x), h \rangle^2 \le \vartheta \langle \nabla^2 F(x) h, h \rangle$, for all $x \in \text{int}(\mathcal{C})$ and for all $h \in \mathbb{R}^n$;

(c) $\nabla^2 F(x) x = -\nabla F(x)$, for all $x \in \text{int}(\mathcal{C})$;

(d) $\nabla F : \text{int}(\mathcal{C}) \to \text{int}(\mathcal{C}^*)$ is a bijection and ∇F_* is its inverse;

(e) $\nabla^2 F_*(-\nabla F(x)) = \left[\nabla^2 F(x)\right]^{-1}$, for all $x \in \text{int}(\mathcal{C})$;

$\nabla^2 F(-\nabla F_*(u)) = \left[\nabla^2 F_*(u)\right]^{-1}$, *for all $u \in \text{int}(\mathcal{C}^*)$;*

(f) $\langle \nabla^2 F(x) h, h \rangle^{1/2} \le \langle -\nabla F(x), h \rangle$, for all $x \in \text{int}(\mathcal{C})$ and for all $h \in \mathcal{C}$;

(g) $\langle (\nabla^2 F(x))^{-1} h, h \rangle^{1/2} \le \langle x, h \rangle$, for all $x \in \text{int}(\mathcal{C})$, for all $h \in \mathcal{C}^*$.

Proof. The fact that every regular cone in $\mathbb{R}^n$ admits a ϑ-LHSCB F, where $1 \le \vartheta \le n$, follows from Theorem 2.5.1 in [17] and [18]. The properties (a) and (b) are essentially parts of a definition of ϑ-LHSCB. For part (c), see [17, Proposition 2.3.4, eqn. (2.3.12)]. For parts (e) see [19, eqn. (2.11)]. For part (d) see [20, Proposition 5.1]. For parts (f) and (g) see [21, Theorem 1.1 and eqns. (1.17 and 1.18)], as well as [22, Section 4]. $\square$

A regular convex cone $\mathcal{C}$ is called *homogeneous* if the automorphism group of $\mathcal{C}$ acts transitively on the interior of $\mathcal{C}$. That is, for every pair of interior points x, y of $\mathcal{C}$, there exists a nonsingular linear map A such that $A(\mathcal{C}) = \mathcal{C}$ and $Ax = y$. For a recent source on homogeneous cones, see [23].

A regular convex cone is called *self-dual*, if there exists an inner product under which the cone is equal to its dual.

3.1 Symmetric Cones

A regular convex cone is called *symmetric* if it is homogeneous and self-dual. Let $\mathcal{C}$ be a symmetric cone. A *ϑ-self-scaled barrier* for $\mathcal{C}$ is a closed, at least three times continuously differentiable, strictly convex function whose domain is the interior of $\mathcal{C}$,

it is ϑ-logarithmically homogeneous for $\vartheta \geq 1$, and has many additional properties (see Definition 2.1 in [19] also see [24]). In the next lemma, we collect only those properties that we will need from self-scaled barriers in this paper.

Lemma 5. *Every symmetric cone $\mathcal{C} \subset \mathbb{R}^n$ admits a ϑ-self-scaled barrier F, where ϑ is the Carathéodory number of $\mathcal{C}$. Every ϑ-self-scaled barrier F for $\mathcal{C}$ is a ϑ-LHSCB for $\mathcal{C}$. Moreover, every ϑ-self-scaled barrier F for $\mathcal{C}$ satisfies the following:*

(a) $\nabla^2 F(x), \left[\nabla^2 F(x)\right]^{1/2} \in \mathrm{Aut}(\mathcal{C})$ for all $x \in \mathrm{int}(\mathcal{C})$;

(b) for every $u \in \mathcal{C}$ the function $f_u : \mathrm{int}(\mathcal{C}) \to \mathbb{R}$, $f_u(x) := \langle -\nabla F(x), u \rangle$, is convex.

Proof. It was shown in Section 3 of [19] that every symmetric cone $\mathcal{C}$ admits a ϑ-self-scaled barrier, where $\vartheta = O(\dim(\mathcal{C}))$. Theorem 4.1 and Lemma 4.1 of [25] combined with Theorem 5.5 of [26], establish that every symmetric cone $\mathcal{C}$ admits a ϑ-self-scaled barrier, where ϑ is the Carathéodory number of $\mathcal{C}$. By Definition 2.1 of [19], every ϑ-self-scaled barrier for $\mathcal{C}$ is a ϑ-LHSCB for $\mathcal{C}$. Property (a) follows from Theorem 3.1 part (iii) of [19] (establishing that the Hessians yield automorphisms) and Theorem 4.1 from [20] and Theorem 5.5 of [26]. Property (b) is Lemma 3.1 of [19]. $\square$

Theorem 6. *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a symmetric cone and let $\mathcal{A}, \mathcal{B}$ be finite sets such that $\mathcal{A} \subseteq \mathcal{C} \subseteq \mathbb{R}^n$, $\mathcal{B} \subseteq \mathcal{C}^\star \subseteq \mathbb{R}^n$ and that*

$$\langle a, b \rangle \leq \Delta \quad \text{for all} \quad a \in \mathcal{A},\ b \in \mathcal{B}$$

for some given Δ, where $\mathcal{C}^* = \mathcal{C}$ under $\langle \cdot, \cdot \rangle$. Further assume that $\mathrm{cone}(\mathcal{A})$ intersects the interior of $\mathcal{C}$ and $\mathrm{cone}(\mathcal{B})$ intersects the interior of $\mathcal{C}^\star$. Then, there exists a self-adjoint automorphism $x \mapsto Lx$ of $\mathcal{C}$ such that

$$\langle La, La \rangle \leq \vartheta \Delta \qquad \text{and} \qquad \langle L^{-1}b, L^{-1}b \rangle \leq \vartheta \Delta \qquad \text{for all} \quad a \in \mathcal{A}, b \in \mathcal{B}\,,$$

where ϑ is the Carathéodory number of $\mathcal{C}$.

Proof. Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a symmetric cone and $\mathcal{A}, \mathcal{B}$ be given as above. Suppose that both of the sets $\mathrm{cone}(\mathcal{A}) \cap \mathrm{int}(\mathcal{C})$, $\mathrm{cone}(\mathcal{B}) \cap \mathrm{int}(\mathcal{C}^\star)$ are nonempty. Let F be a ϑ-self-scaled barrier for $\mathcal{C}$, where ϑ is the Carathéodory number of $\mathcal{C}$ (such F exists by Lemma 5). Consider the following optimization problem, where $w \in \mathbb{R}^n$ and $t \in \mathbb{R}$ are variables.

$$\begin{aligned}
\inf \quad & t \\
\text{subject to} \quad & \\
& \langle -\nabla F(w), a \rangle \leq t \qquad \text{for all} \quad a \in \mathcal{A} \qquad\qquad \textbf{(SYM1)} \\
& \langle w, b \rangle \leq t \qquad\qquad \text{for all} \quad b \in \mathcal{B} \\
& w \in \mathrm{int}\,\mathcal{C}\,.
\end{aligned}$$

By Lemma 5 part (b), the solution set of each constraint in the first group of constraints is a convex set; the second set of constraints define a polyhedron (in fact,

a polyhedral cone), the third constraint is a convex set constraint and the objective function is linear. Therefore, the optimization problem (**SYM1**) is a convex optimization problem. The problem is not infeasible (pick any w in the interior of $\mathcal{C}$ and set $t := 1 + \max\left\{\max_{a \in \mathcal{A}}\left\{\langle -\nabla F(w), a\rangle\right\}, \max_{b \in \mathcal{B}}\left\{\langle w, b\rangle\right\}\right\}$) and is not unbounded (by definition of the dual cone, $\mathcal{A}$ and $\mathcal{B}$, t is nonnegative for every feasible solution). Hence, the optimal value is finite.

Claim: The optimal value of (**SYM1**) is attained.

Proof of Claim: By Lemma 4, part (a), $\left\|\nabla F\left(w^{(k)}\right)\right\| \to +\infty$, for every sequence in the interior of $\mathcal{C}$ converging to a boundary point of $\mathcal{C}$. Since $\mathrm{cone}(\mathcal{A}) \cap \mathrm{int}(\mathcal{C}) \neq \emptyset$, due to the first group of constraints, for every sequence in the feasible region converging to a boundary point of $\mathcal{C}$, $t \to +\infty$. Thus, the optimal value can only be attained in the interior of $\mathcal{C}$. Hence, there exists a positive integer M such that the closed set $\left\{\begin{pmatrix} w \\ t \end{pmatrix} \in \mathrm{int}(\mathcal{C}) \oplus \mathbb{R}_+ : F(w) \leq M\right\}$ contains all minimizers of (**SYM1**). Since, in addition, $\mathrm{cone}(\mathcal{B}) \cap \mathrm{int}(\mathcal{C}^\star) \neq \emptyset$, the set

$$\left\{\begin{pmatrix} w \\ t \end{pmatrix} \in \mathrm{int}(\mathcal{C}) \oplus \mathbb{R}_+ : \langle w, b\rangle \leq \overline{t} \ \forall b \in \mathcal{B}, \ \ t \leq \overline{t}, \ \ F(w) \leq M\right\},$$

where $\overline{t}$ is the optimal value, is nonempty, compact and contains all optimal solutions. By the above arguments, we proved that the convex optimization problem (**SYM1**) can be written equivalently as the minimization of a linear function over a nonempty, compact set. Therefore, this latter problem as well as (**SYM1**) attain their optimal value in the interior of $\mathcal{C}$. $\diamond$

Note that the feasible solution mentioned before the above claim is a Slater point for (**SYM1**). Thus, by necessary (and sufficient) conditions for optimality, every optimal solution $\left(\overline{w}, \overline{t}\right)$ satisfies the following conditions for some $\lambda \in \mathbb{R}^{\mathcal{A}}$, and $\mu \in \mathbb{R}^{\mathcal{B}}$.

$$\nabla^2 F\left(\overline{w}\right)\left(\sum_{a \in \mathcal{A}} \lambda_a a\right) = \left(\sum_{b \in \mathcal{B}} \mu_b b\right) \tag{4}$$

$$\sum_{a \in \mathcal{A}} \lambda_a + \sum_{b \in \mathcal{B}} \mu_b = 1 \tag{5}$$

$$\langle -\nabla F(\overline{w}), a\rangle \leq \overline{t} \qquad\qquad \text{for all} \quad a \in \mathcal{A} \tag{6}$$

$$\langle \overline{w}, b\rangle \leq \overline{t} \qquad\qquad \text{for all} \quad b \in \mathcal{B} \tag{7}$$

$$\lambda_a \geq 0, \quad \mu_b \geq 0 \qquad\qquad \text{for all} \quad a \in \mathcal{A} \quad \text{and} \quad b \in \mathcal{B} \tag{8}$$

$$\lambda_a > 0 \implies \langle -\nabla F(\overline{w}), a\rangle = \overline{t} \qquad\qquad \text{for all} \quad a \in \mathcal{A} \tag{9}$$

$$\mu_b > 0 \implies \langle \overline{w}, b\rangle = \overline{t} \qquad\qquad \text{for all} \quad b \in \mathcal{B} \tag{10}$$

$$\overline{w} \in \mathrm{int}\,\mathcal{C}. \tag{11}$$

Let us define a' and b' as follows

$$a' := \sum_{a \in \mathcal{A}} \lambda_a a \qquad\qquad \text{and} \qquad\qquad b' := \sum_{b \in \mathcal{B}} \mu_b b\,.$$

Using (11) and Lemma 4 part (c) together with (4) we have

$$\langle -\nabla F(\overline{w}), a' \rangle = \langle \nabla^2 F(\overline{w})\overline{w}, a' \rangle = \langle \overline{w}, \nabla^2 F(\overline{w})a' \rangle = \langle \overline{w}, b' \rangle \,.$$

By (8), (9) and (10), we get

$$\langle -\nabla F(\overline{w}), a' \rangle = \overline{t} \sum_{a \in \mathcal{A}} \lambda_a$$

and

$$\langle \overline{w}, b' \rangle = \overline{t} \sum_{b \in \mathcal{B}} \mu_b \,,$$

and so $\overline{t} \sum_{a \in \mathcal{A}} \lambda_a = \overline{t} \sum_{b \in \mathcal{B}} \mu_b$ and consequently by (5) we have

$$\sum_{a \in \mathcal{A}} \lambda_a = \sum_{b \in \mathcal{B}} \mu_b = 1 \,.$$

Now, we consider another optimization problem (with w and τ as variables) which gets us closer to finding a scaling:

$$\inf \quad \tau$$

subject to

$$\langle \nabla^2 F(w)a, a \rangle \leq \tau \qquad \text{for all} \quad a \in \mathcal{A} \qquad \textbf{(SYM2)}$$
$$\langle \left(\nabla^2 F(w) \right)^{-1} b, b \rangle \leq \tau \qquad \text{for all} \quad b \in \mathcal{B}$$
$$w \in \operatorname{int} \mathcal{C} \,.$$

By Lemma 4 parts (f) and (g), $\left(\overline{w}, \overline{t}^2 \right)$ is a feasible solution to (**SYM2**) with objective value $\overline{t}^2$. By Lemma 4 part (b), we have

$$\vartheta \langle \nabla^2 F(\overline{w})a', a' \rangle \geq \langle -\nabla F(\overline{w}), a' \rangle^2 = \overline{t}^2 \,.$$

Since

$$\langle \nabla^2 F(\overline{w})a', a' \rangle = \langle b', a' \rangle \leq \Delta \,,$$

we have that the objective value of the feasible solution $\left(\overline{w}, \overline{t}^2 \right)$ of (**SYM2**) is at most $\vartheta \Delta$. To finish the proof, we define L as $\left(\nabla^2 F(\overline{w}) \right)^{1/2}$ and use Lemma 5 part (a). $\qquad \square$

The statement of the theorem is invariant under switching $\mathcal{C}$ and $\mathcal{C}^\star$. Indeed, the proof of the theorem also has this primal-dual symmetry. For self-scaled barriers, for the choice of inner-product which results in $\mathcal{C}^\star = \mathcal{C}$, F and F_* may only differ by a constant. As a result, the derivatives of F and F_* are the same.

The above theorem and its proof generalize to the case that $\mathcal{A}$ and $\mathcal{B}$ are not necessarily finite sets, but they are bounded. Next, let us consider the simpler direction

9

(specialization) when $\mathcal{A} = \{a\}$, $\mathcal{B} = \{b\}$ (both sets are singletons). If they satisfy the hypotheses of Theorem 6, then $a \in \text{int}(\mathcal{C})$ and $b \in \text{int}(\mathcal{C}^\star)$. In this case, the "good" feasible solution $(\overline{w}, \overline{t})$ of the optimization problem (**SYM2**) used in the proof above is determined by the unique solution of

$$\nabla^2 F(w)a = b, \quad w \in \text{int}(\mathcal{C}).$$

This unique solution leads to $\nabla^2 F(\overline{w})$ which is called the *Nesterov–Todd scaling* in the area of interior-point methods in convex optimization.

Our optimization problem is trying to find the smallest possible radius $\sqrt{\tau}$ of a pair of primal-dual ellipsoids such that

- ellipsoids are centred at the origin;
- primal ellipsoid contains a, the dual ellipsoid contains b;
- the primal ellipsoids are defined by the Hessians of a self-scaled barrier for $\mathcal{C}$ (and hence in each instance, the dual ellipsoid is defined by the inverse of the specific Hessian chosen for the primal cone, due to the definition of the dual norm).

This optimization problem is a kind of a dual problem to *Todd's largest primal-dual ellipsoids problem* [27]. Todd asks " given interior points a and b for $\mathcal{C}$ and $\mathcal{C}^*$ respectively, what are the *largest* primal-dual ellipsoids centred at a and b respectively and contained in $\mathcal{C}$ and $\mathcal{C}^\star$ respectively?" Here, "largest" means the geometric mean of the radii of the primal and dual ellipsoids is maximized. So, this problem aims to construct largest inner ellipsoidal approximations to primal and dual cones simultaneously, a very useful and fundamental object in theory and algorithms for convex optimization. Todd [27] proves that $\overline{w}$ (mentioned above) yields an optimal solution for his problem.

Let $\kappa(\mathcal{C})$ denote the Carathéodory number and $\ell(\mathcal{C})$ denote the longest chain of nonempty faces of a regular convex cone $\mathcal{C}$. Then, $\kappa(\mathcal{C}) \leq \ell(\mathcal{C}) - 1$ [28]. Interestingly, in Averkov [29] as well as in Saunderson [30] where extension complexity of a regular convex cone in terms of products of "low-complexity" cones is studied the geometric measure of complexity $\ell(\mathcal{C})$ of the representing cones plays a key role. In the case of symmetric cones, $\mathcal{C}$, $\kappa(\mathcal{C}) = \ell(\mathcal{C}) - 1$.

Our results in this subsection and the next provide bounds on the normalization constant $f_{\mathcal{C}}$ which depend on the barrier parameter ϑ. The latter is additive under products of cones while the former is bounded by the maximum of $f_{\mathcal{C}}$ (see, Remark 3). Therefore, our results in this subsection and the next are strongest, when they are applied to indecomposable cones.

3.2 Homogeneous Cones

In the next lemma we collect the properties that we will need from certain ϑ-LHSCB barriers for homogeneous cones.

Lemma 7. *Every homogeneous cone $\mathcal{C} \subset \mathbb{R}^n$ admits a ϑ-logarithmically homogeneous self-concordant barrier F, such that*
(a) $\vartheta \leq n$;

(b) there exists a real-valued function g such that

$$F(L(x)) = F(x) + g(L), \text{ for every } L \in \text{Aut}(\mathcal{C});$$

(c) for every $u \in \mathcal{C}$ the function $f_u : \text{int}(\mathcal{C}) \to \mathbb{R}$, $f_u(x) := \langle -\nabla F(x), u \rangle$, is convex;

(d) for every $x \in \text{int}(\mathcal{C})$, there exists $L \in \text{Aut}(\mathcal{C})$ such that $\nabla^2 F(x) = L^\star L$.

Proof. For parts a, b and c see [31, Section 8 and Theorem 6.1]. For part d see [23, Theorem 5.3]. $\qquad\square$

Theorem 8. *Let $\mathcal{C} \subset \mathbb{R}^n$ be a homogeneous cone. Also let $\mathcal{A}$, $\mathcal{B}$ be finite sets such that $\mathcal{A} \subseteq \mathcal{C}$, $\mathcal{B} \subseteq \mathcal{C}^\star$ and*

$$\langle a, b \rangle \leq \Delta \quad \text{for all} \quad a \in \mathcal{A}, \, b \in \mathcal{B}$$

for some given Δ. Further assume that $\text{cone}(\mathcal{A})$ intersects the interior of $\mathcal{C}$ and $\text{cone}(\mathcal{B})$ intersects the interior of $\mathcal{C}^\star$. Then, there exists an automorphism $x \mapsto Lx$ of $\mathcal{C}$ such that

$$\langle La, La \rangle \leq \vartheta \Delta \qquad \text{and} \qquad \langle L^{-\star}b, L^{-\star}b \rangle \leq \vartheta \Delta \qquad \text{for all} \quad a \in \mathcal{A}, b \in \mathcal{B},$$

where $L^{-\star}$ is inverse-adjoint to L, and ϑ is the parameter of a ϑ-LHSCB satisfying the properties listed in the statement of Lemma 7.

Proof. Let F be a ϑ-LHSCB for $\mathcal{C}$ satisfying the hypothesis (such F exists by Lemma 7). Again by Lemma 7 for every point $w \in \text{int}\,\mathcal{C}$, $\nabla^2 F(w)$ can be represented as $L^\star L$ such that the linear map $x \mapsto Lx$ defines an automorphism of $\mathcal{C}$. Similarly, $\left(\nabla^2 F(W)\right)^{-1}$ can be represented as $L^{-1}L^{-\star}$ where the linear map $x \mapsto L^{-\star}x$ defines an automorphism of $\mathcal{C}^\star$.

Similarly, to the proof of Theorem 6 we can write the optimization problems in variables (w, t) and (w, τ) respectively:

$$
\begin{aligned}
\inf \quad & t \\
\text{subject to} \quad & \\
& \langle -\nabla F(w), a \rangle \leq t \qquad \text{for all} \quad a \in \mathcal{A} \qquad\qquad \textbf{(HOM1)}\\
& \langle w, b \rangle \leq t \qquad\qquad\quad \text{for all} \quad b \in \mathcal{B} \\
& w \in \text{int}\,\mathcal{C}
\end{aligned}
$$

and

$$
\begin{aligned}
\inf \quad & \tau \\
\text{subject to} \quad & \\
& \langle \nabla^2 F(w)a, a \rangle \leq \tau \qquad\qquad \text{for all} \quad a \in \mathcal{A} \qquad\qquad \textbf{(HOM2)}\\
& \langle \left(\nabla^2 F(w)\right)^{-1}b, b \rangle \leq \tau \qquad \text{for all} \quad b \in \mathcal{B} \\
& w \in \text{int}\,\mathcal{C}.
\end{aligned}
$$

11

Also in the case of a homogeneous cone, the program (**HOM2**) is convex (this time by Lemma 7) and therefore we can apply the same analysis as in Theorem 6 to show that the optimal objective value of (**HOM2**) is bounded above by $\vartheta\Delta$. □

The following example demonstrates that the normalization through automorphisms is not applicable to some non-homogeneous cones.

Example 9. *Consider the cone $\mathcal{C} := \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 \leq x_3^2 , x_1 \geq 0 , x_3 \geq 0\}$.*
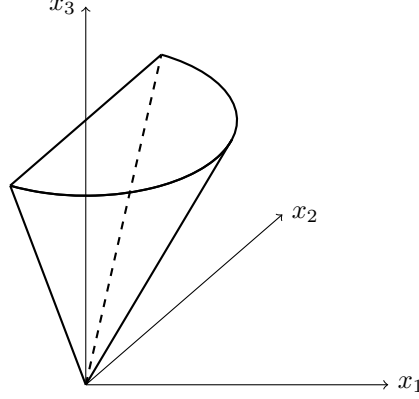
The dual cone $\mathcal{C}^\star$ is not facially exposed and therefore is not homogeneous [32]. So the results in this section are not applicable to the cone $\mathcal{C}$. Note however that $\mathcal{C}$ is a hyperbolic cone (intersection of two well-known hyperbolic cones: a closed half-space and a three-dimensional second-order cone) and its dual $\mathcal{C}^*$ is not [33, Example 4].

Let us show that no appropriate rescaling is possible for $\mathcal{C}$ and $\mathcal{C}^\star$ in some cases. Automorphisms of the cone $\mathcal{C}$ can be parametrized by $\alpha, \beta \in \mathbb{R}$, $\alpha > 0$ as follows

$$x \mapsto \alpha \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\beta^2 + 1} & \beta \\ 0 & \beta & \sqrt{\beta^2 + 1} \end{bmatrix} x \qquad and \qquad x \mapsto \alpha \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sqrt{\beta^2 + 1} & -\beta \\ 0 & \beta & \sqrt{\beta^2 + 1} \end{bmatrix} x$$

with the corresponding automorphisms of the dual cone $\mathcal{C}^\star$

$$x \mapsto \frac{1}{\alpha} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{\beta^2 + 1} & -\beta \\ 0 & -\beta & \sqrt{\beta^2 + 1} \end{bmatrix} x \qquad and \qquad x \mapsto \frac{1}{\alpha} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\sqrt{\beta^2 + 1} & -\beta \\ 0 & \beta & \sqrt{\beta^2 + 1} \end{bmatrix} x .$$

Let us define sets $\mathcal{A}$ and $\mathcal{B}$ as the sets consisting of single elements $(M, 0, M)$ and $(-M, 0, M)$, respectively, where M is a large number. It is straightforward to check that no suitable rescaling of these $\mathcal{A}$ and $\mathcal{B}$ is possible through the above pairs of automorphisms. Therefore, the normalization results of this section cannot be generalized to hyperbolic cones.

Building upon the literature, we reference Remark 2 by Averkov et al. [13], which highlighted the critical role of automorphisms of the positive semidefinite cone in their proof. Our results in the preceding and current sections further demonstrate that automorphisms continue to play a significant role in more general convex cone settings.

4 Rescaling of Cones and Extension Complexity

This section is dedicated to the implications of factorization normalizations on the expressive power of cones. The expressive power of a cone $\mathcal{C} \subseteq \mathcal{R}^n$ will be reflected in the pairs $\mathcal{V}$, $\mathcal{F}$ that admit a factorization over the cone $\mathcal{C}$, where $\langle v, f \rangle \geq 0$ for every $v \in \mathcal{V}$ and $f \in \mathcal{F}$. In other words, we are interested in the pairs $\mathcal{V}$, $\mathcal{F}$ such that for every $v \in \mathcal{V}$ and for every $f \in \mathcal{F}$ there are $a_v \in \mathcal{C}$ and $b_f \in \mathcal{C}^\star$ with $\langle v, f \rangle = \langle a_v, b_f \rangle$ for every $v \in \mathcal{V}$ and for every $f \in \mathcal{F}$.

The intuition behind the study of limitations in the expressive power of a cone $\mathcal{C}$ is as follows. First, we consider a specific collection of pairs $\mathcal{V}$, $\mathcal{F}$, where $\langle v, f \rangle \geq 0$ for every $v \in \mathcal{V}$ and $f \in \mathcal{F}$. We plan to prove that some of the pairs in the collection cannot be factorized using the cone $\mathcal{C}$. Second, for the sake of argument, let us consider a pair $\mathcal{V}$, $\mathcal{F}$ that can be factorized using the cone $\mathcal{C}$. Consider the sets

$$\mathcal{A} := \{a_v \,:\, v \in \mathcal{V}\} \qquad \text{and} \qquad \mathcal{B} := \{b_f \,:\, f \in \mathcal{F}\} \,.$$

We show that there exists a set Γ (which does not depend on $\mathcal{V}$ and $\mathcal{F}$) and there exist at most $2n$ vectors in Γ, such that the whole pair $\mathcal{V}$ and $\mathcal{F}$ can be identified from these $2n$ vectors in Γ. The $2n$ vectors in Γ are obtained by carefully selecting a set of at most $2n$ vectors from $\mathcal{B}$ and then for each of them picking a "closest" vector from Γ. Third, the cardinality of the constructed set Γ depends only on the cone $\mathcal{C}$, in particular, it depends only on the dimension of its ambient space n and the normalization constant $f_{\mathcal{C}}$. Now using the counting argument, for a sufficiently large collection of pairs $\mathcal{V}$, $\mathcal{F}$ we would not be able to identify each of them using at most $2n$ vectors from Γ. This leads to the statement that the cone $\mathcal{C}$ does not have enough of expressive power to capture all of the studied pairs $\mathcal{V}$, $\mathcal{F}$.

4.1 Relationship to Polytopes

Let us now consider how the pairs $\mathcal{V}$, $\mathcal{F}$ appear in the context of polytopes. We assume that a set $\mathcal{V}$ is a subset of $\mathcal{P} \subseteq \mathcal{R}^d$ and $\mathcal{F}$ is a subset of $\mathcal{H} \subseteq \mathcal{R}^d$. For example, $\mathcal{P}$ can correspond to the set of all $0/1$ points in $\mathcal{R}^{d-1}$ and $\mathcal{H}$ can correspond to the set of all possible hyperplanes spanned by $d-1$ affinely independent $0/1$ points in $\mathcal{R}^{d-1}$. The next theorem corresponds to representing the convex hull of $\mathcal{X} \subseteq \{0,1\}^d$ through its outer and inner descriptions, i.e. once as a set of linear inequalities, say $\mathcal{F}$, and as a set of its extreme points, say $\mathcal{V}$. The upper bound on M in the below theorem can be obtained for example through a straightforward application of Cramer's rule.

Theorem 10. *Given a nonnegative integer d, there is a way to assign to each $\mathcal{X} \subseteq \{0,1\}^{d-1}$ a unique $\mathcal{V} \subseteq \{0,1\}^{d-1} \times \{1\}$ and a unique $\mathcal{F} \subseteq \mathbb{Z}^d$ such that*
(i) $\|f\|_\infty \leq M$, $\|v\|_\infty \leq 1$ for all $v \in \mathcal{V}$ and $f \in \mathcal{F}$;
(ii) $\langle v,\, f \rangle \geq 0$ for all $v \in \mathcal{V}$ and $f \in \mathcal{F}$;
(iii) for all $v \in \{0,1\}^{d-1} \times \{1\}$, $v \notin \mathcal{V}$ there exists $f \in \mathcal{F}$ such that $\langle v, f \rangle \leq -1$,
where $M = 2^{d \log(2d)}$.

Similarly, the next theorem corresponds to outer and inner descriptions of cyclic polytopes. We can use Lemma 5 in [8] to estimate the coefficients in an outer description of a cyclic polytope, i.e., to obtain an upper bound on M.

Theorem 11. *Given a nonnegative integer d, there is a way to assign to each subset $\mathcal{X}$ of the set $\left\{ (k, k^2, \ldots, k^{d-1}) \,:\, k \in \{0, \ldots, t\} \right\}$ a unique $\mathcal{V} \subseteq \mathbb{Z}^{d-1} \times \{1\}$ and a unique $\mathcal{F} \subseteq \mathbb{Z}^d$ such that*

(i) $\|f\|_\infty \leq M$, $\|v\|_\infty \leq t^{d-1}$ for all $v \in \mathcal{V}$ and $f \in \mathcal{F}$;

(ii) $\langle v, f \rangle \geq 0$ for all $v \in \mathcal{V}$ and $f \in \mathcal{F}$;

(iii) for all $v \in \mathbb{Z}^{d-1} \times \{1\}$, $v \notin \mathcal{V}$ there exists $f \in \mathcal{F}$ such that $\langle v, f \rangle \leq -1$,

where $M = ((d+1)t^d)^d$.

4.2 Compact Encoding

In this section, we show Lemma 13 that guarantees that under some conditions, the pairs $\mathcal{V}$ and $\mathcal{F}$ can be compactly encoded. Before stating and proving Lemma 13 we need some technical result about "nets". The next lemma gives us a way to round the vectors in one of the factorization parts such that the rounded vectors lie in a set Γ, where Γ depends only on the cone $\mathcal{C}$ and the set Γ has an appropriate bound on its cardinality. Later, ρ in the lemma will be a bound on the norms of the vectors in the factorization over the cone $\mathcal{C}$ and its dual and; ϵ will indicate the precision of rounding. The lemma below follows from a classical result of Rogers [34], also see [35]. The lemma is used for going from exact factorizations to "approximate" factorizations, where one of the factors can use a bounded number of vectors.

Lemma 12. *Fix a dimension $n \geq 3$, radius $\rho > 0$, and $\epsilon > 0$ such that $\rho/\epsilon \geq n$. Then there exists a set $\Gamma := \Gamma(n, \rho, \epsilon) \subseteq \mathbb{R}^n$ such that*

(i) for all $u \in B_0(\rho) := \{x \in \mathbb{R}^n \,:\, \|x\|_2 \leq \rho\}$ there is $\overline{u} \in \Gamma$ such that $\|u - \overline{u}\|_2 \leq \epsilon$, and

(ii) $|\Gamma| \leq \exp(1)(n \ln n + n \ln \ln n + 5n) (\rho/\epsilon)^n$.

The next lemma tells that under some conditions on $\mathcal{V}$, $\mathcal{F}$, the set $\mathcal{V}$ can be reconstructed using at most $2n$ vectors from a sufficiently small Γ, i.e. $\mathcal{V}$ can be reconstructed from a relatively small amount of information.

Lemma 13. *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a regular convex cone that can be normalized with coefficient $f_\mathcal{C}$. Let us have two families of vectors $\mathcal{P}, \mathcal{H} \subseteq \mathbb{R}^d$, nonnegative integer M, and sets $\mathcal{V} \subseteq \mathcal{P}$, $\mathcal{F} \subseteq \mathcal{H}$ such that*

(i) $\|f\|_\infty \leq M$, $\|v\|_\infty \leq 1$ for all $f \in \mathcal{F}$ and $v \in \mathcal{V}$;

(ii) $\langle v, f \rangle \geq 0$ for all $v \in \mathcal{V}$ and $f \in \mathcal{F}$;

(iii) for all $v \in \mathcal{P}$, if $v \notin \mathcal{V}$ then there exists $f^* \in \mathcal{F}$ such that $\langle v, f^* \rangle \leq -1$;

(iv) the pair $\mathcal{V}$, $\mathcal{F}$ can be factorized using the cone $\mathcal{C}$.

Let us define $\rho := \sqrt{(d+1)M} \cdot f_\mathcal{C}$ and $\epsilon := (4(n+d)\rho)^{-1}$. Then for every $i \in \{1, \ldots, n+d\}$ there exists

(1) a vector $f_i \in \mathcal{H}$ such that $\|f_i\|_\infty \leq M$,

(2) a vector $u_{f_i} \in \Gamma := \Gamma(n, \rho, \epsilon)$, where Γ is as in Lemma 12;

such that the following holds

$$\mathcal{V} = \{x \in \mathcal{P} \,:\, \text{there exists } y \in \mathcal{C}^* \cap B_0(\rho) \text{ such that}$$

$$|\langle f_i, x \rangle - \langle u_{f_i}, y \rangle| \leq \frac{1}{4(n+d)} \quad \text{for all } i \in \{1, \ldots, n+d\}\}.$$

Proof. By assumption there is a $\mathcal{C}$-factorization $(a_v)_{v \in \mathcal{V}} \subseteq \mathcal{C}$, $(b_f)_{f \in \mathcal{F}} \subseteq \mathcal{C}^*$ for $\mathcal{V}$ and $\mathcal{F}$ such that for every $v \in \mathcal{V}$ and every $f \in \mathcal{F}$ we have

$$\langle v, f \rangle = \langle a_v, b_f \rangle .$$

Due to normalization of the cone $\mathcal{C}$, we may assume that $\|a_v\|_2 \le \sqrt{dM} \cdot f_\mathcal{C}$ for all $v \in \mathcal{V}$ and that $\|b_f\|_2 \le \sqrt{dM} \cdot f_\mathcal{C}$ for all $f \in \mathcal{F}$. Let us now select a "subsystem of maximum volume", as in [11] and [12]. For linearly independent set of vectors $x_1, \dots, x_k \subseteq \mathbb{R}^n$ denote the k-dimensional parallelepiped volume

$$\mathrm{vol}\left(\sum_{i=1}^{k} \lambda_i x_i \; : \; \lambda_i \in [0,1] \quad \text{for all } i \right) = \det\left((x_i^\top x_j)_{ij} \right)^{1/2} .$$

Clearly, if the vectors $x_1, \dots, x_k \subseteq \mathbb{R}^n$ are dependent the volume is zero. Let us define

$$W := \mathrm{span}\left(\{ (f, b_f) : f \in \mathcal{F} \} \right) ,$$

and let $I' \subseteq \mathcal{F}$ be a subset of size $|I'| = \dim(W)$ such that $\mathrm{vol}(\{ (f, b_f) : f \in I' \})$ is maximized. Note that $|I'| \le n + d$ since the dimension of the linear space W is at most $n + d$.

Now, for each $f \in \mathcal{F}$ let $\overline{b}_f$ be an element of Γ such that

$$\|b_f - \overline{b}_f\|_2 \le \epsilon = \left(4(n+d)\sqrt{(d+1)M} \cdot f_\mathcal{C} \right)^{-1} ,$$

which exists by property (i) of Γ. Let us define the set $\overline{\mathcal{V}}$ as the following set

$$\overline{\mathcal{V}} := \{ v \in \mathcal{P} \; : \text{there exists } y \in \mathcal{C} \cap B_0(\rho) \text{ such that}$$
$$\left| \langle v, f \rangle - \langle \overline{b}_f, y \rangle \right| \le \frac{1}{4(n+d)} \text{ for all } f \in I' \}.$$

In the remaining part of the proof we show that $\overline{\mathcal{V}} = \mathcal{V}$ as desired in the statement of the lemma, by selecting f_i, $i \in \{1, \dots, n+d\}$ to be the vectors in I' (duplicating them, if needed) and setting $u_f := \overline{b}_f$ for $f \in I'$.

First, we show that $\mathcal{V} \subseteq \overline{\mathcal{V}}$. Let us consider some v in $\mathcal{V}$, and then let us show that v is also in $\overline{\mathcal{V}}$. For every $v \in \mathcal{V}$ we have $\|a_v\|_2 \le \rho = \sqrt{(n+1)M} \cdot f_\mathcal{C}$, so for every $f \in I'$ we have

$$\left| \langle v, f \rangle - \langle \overline{b}_f, a_v \rangle \right| = \left| \langle v, f \rangle - \langle a_v, b_f \rangle + \langle b_f - \overline{b}_f, a_v \rangle \right| = \left| \langle b_f - \overline{b}_f, a_v \rangle \right|$$
$$\le \|b_f - \overline{b}_f\|_2 \|a_v\|_2 \le \frac{1}{4(n+d)},$$

where the last line follows from the Cauchy-Schwarz inequality. Thus, by the definition of $\overline{\mathcal{V}}$, we conclude that v lies in $\overline{\mathcal{V}}$.

15

Now we show the reverse inclusion that $\overline{\mathcal{V}} \subseteq \mathcal{V}$. For this, suppose that there exists $v \in \mathcal{P}$ that is not in $\mathcal{V}$ and let us show that v is also not in $\overline{\mathcal{V}}$. Since v is not in $\mathcal{V}$, by (iii) we know that there exists some $f^* \in \mathcal{F}$ such that $\langle f^*, v \rangle \leq -1$.

Now, it may not be the case that f^* is in I'. However, the vector (f^*, b_{f^*}) is in the linear space W, and so we can express (f^*, b_{f^*}) as a linear combination of the vectors (f, b_f), $f \in I'$. In particular, there are unique multipliers $\nu \in \mathbb{R}^{I'}$ such that $(f^*, b_{f^*}) = \sum_{f \in I'} \nu_i (f, b_f)$. Using the choice of subsystem I' as a system with a maximum volume and using Cramer's rule, we have that for every $f' \in I'$

$$|\nu_{f'}| = \frac{\mathrm{vol}(\{(f, b_f) : f \in I' \backslash \{f'\} \cup \{f^*\}\})}{\mathrm{vol}(\{(f, b_f) : f \in I'\})} \leq 1 \,.$$

For every $y \in \mathcal{C} \cap B_0(\rho)$, we have $\langle b_{f^*}, y \rangle \geq 0$ and so we also have

$$1 \leq |-\langle f^*, v \rangle + \langle b_{f^*}, y \rangle| = \left| \sum_{f \in I'} \nu_f \left(-\langle f, v \rangle + \langle b_f, y \rangle \right) \right|$$

$$\leq \sum_{f \in I'} |\nu_f| \, |-\langle f, v \rangle + \langle b_f, y \rangle| \leq (n+d) \max_{f \in I'} |-\langle f, v \rangle + \langle b_f, y \rangle| \,.$$

Also for every $y \in \mathcal{C} \cap B_0(\rho)$ and for every $f \in I'$, we have

$$|-\langle f, v \rangle + \langle b_f, y \rangle| = \left| -\langle f, v \rangle + \langle \overline{b}_f, y \rangle + \langle b_f - \overline{b}_f, y \rangle \right|$$

$$\leq \left| -\langle f, v \rangle + \langle \overline{b}_f, y \rangle \right| + \left| \langle b_f - \overline{b}_f, y \rangle \right|$$

$$\leq \left| -\langle f, v \rangle + \langle \overline{b}_f, y \rangle \right| + \frac{1}{4(n+d)} \,,$$

where the last inequality follows from $\|b_f - \overline{b}_f\|_2 \leq \epsilon = (4(n+d)\rho)^{-1}$ and $\|y\|_2 \leq \rho$. Combining this with the previous inequality

$$\max_{f \in I'} |-\langle f, v \rangle + \langle b_f, y \rangle| \geq 1/(n+d) \,,$$

we get

$$\max_{f \in I'} \left| -\langle f, v \rangle + \langle \overline{b}_f, y \rangle \right| \geq \max_{f \in I'} |-\langle f, v \rangle + \langle b_f, y \rangle| - \frac{1}{4(n+d)}$$

$$\geq \frac{1}{(n+d)} - \frac{1}{4(n+d)} \geq \frac{1}{2(n+d)}$$

which implies that v does not lie in $\overline{\mathcal{V}}$. $\qquad\square$

The next corollary outlines the strategy that is used for utilizing Lemma 13 for establishing lower bounds for extension complexity.

Corollary 14. *Let $\mathcal{C} \subseteq \mathbb{R}^n$ be a regular convex cone that can be normalized with coefficient $f_{\mathcal{C}}$. Let us have two families of vectors $\mathcal{P}, \mathcal{H} \subseteq \mathbb{R}^d$ and a nonnegative integer M. Let us assume that we have a collection of N objects each of which leads to a unique $\mathcal{V} \subseteq \mathcal{P}$ and a unique $\mathcal{F} \subseteq \mathcal{H}$ such that* (i), (ii), (iii), (iv) *in Lemma 13 are satisfied. Let us define $\rho := \sqrt{(n+1)M} \cdot f_{\mathcal{C}}$ and $\epsilon := (4(n+d)\rho)^{-1}$.*

If every above pair $\mathcal{V}, \mathcal{F}$ can be factorized using the cone $\mathcal{C}$ then we have

$$|N| \leq (\Gamma(n, \rho, \epsilon) \cdot |\{f \in \mathcal{H} \,:\, \|f\|_\infty \leq M\}|)^{n+d} \ .$$

Proof. Indeed, by Lemma 13 each $\mathcal{V}$ can be uniquely identified by $n + d$ vectors in $\{f \in \mathcal{H} \,:\, \|f\|_\infty \leq M\} \times \Gamma(n, \rho, \epsilon)$, leading to the desired inequality. $\qquad\square$

4.3 Relationship to Extension Complexity

Using Corollary 14 and Theorem 10, we can show that there are $0/1$ polytopes in $\mathbb{R}^d$ with no extension over a cone $\mathcal{C} \subseteq \mathbb{R}^n$ if both $f_{\mathcal{C}}$ and n are polynomial in n. The statement is equivalent to showing that at least one of the pairs $\mathcal{V}, \mathcal{F}$ from Theorem 10 cannot be factorized using the cone $\mathcal{C}$. The proof of this fact follows the proofs of [11] and [12]. For the sake of contradiction, let us assume that every pair $\mathcal{V}, \mathcal{F}$ as in Theorem 10 can be factorized using a cone $\mathcal{C}$ for each $\mathcal{X} \subseteq \{0,1\}^d$, where $\mathcal{C} \subseteq \mathbb{R}^n$ can be normalized with coefficient $f_{\mathcal{C}}$. Thus by Lemma 12 and Lemma 13, the total number of different possible elements Γ is at most

$$
\begin{aligned}
|\Gamma| &\leq \exp(1)(n \ln n + n \ln \ln n + 5n)\left(\frac{\rho}{\epsilon}\right)^n = \\
&\quad \exp(1)(n \ln n + n \ln \ln n + 5n)\left(4 \cdot \rho^2 \cdot (n+d)\right)^n \leq \\
&\quad \exp(1)(n \ln n + n \ln \ln n + 5n)\left(4 \cdot (d+1)M \cdot f_{\mathcal{C}}^2 \cdot (n+d)\right)^n \leq \\
&\quad n^2 \cdot \left(3 \cdot (d+1)M \cdot f_{\mathcal{C}}^2 \cdot n\right)^n \ .
\end{aligned}
$$

Thus, the total number of different possible $\mathcal{V}$ is at most

$$\left(\left(2 \cdot 2^{d \log(2d)} + 1\right)^{d+1} \cdot n^2 \cdot \left(3 \cdot (d+1)2^{d \log(2d)} \cdot f_{\mathcal{C}}^2 \cdot n\right)^n\right)^{n+d}$$

and so

$$2^{2^d} - 1 \leq \left(\left(2 \cdot 2^{d \log(2d)} + 1\right)^{d+1} \cdot n^2 \cdot \left(3 \cdot (d+1)2^{d \log(2d)} \cdot f_{\mathcal{C}}^2 \cdot n\right)^n\right)^{n+d},$$

showing that $n \cdot f_{\mathcal{C}}$ has to be exponentially large with respect to d in order for $\mathcal{C}$ to lead to an extension of every $0/1$ polytope of dimension d.

Similarly, Corollary 14 and Theorem 11 show that there are cyclic polytopes corresponding to Theorem 11 with no extension over a cone $\mathcal{C} \subseteq \mathbb{R}^n$ if both $f_{\mathcal{C}}$ and n are polynomial in n and t. Lower bounds for cyclic polytopes were studied in [8] and this proof would be analogous to theirs.

References

[1] Gouveia, J., Parrilo, P.A., Thomas, R.R.: Lifts of convex sets and cone factorizations. Mathematics of Operations Research **38**(2), 248–264 (2013) https://doi.org/10.1287/moor.1120.0575

[2] Lee, D.D., Seung, H.S.: Learning the parts of objects by non-negative matrix factorization. Nature **401**(6755), 788–791 (1999)

[3] Devarajan, K.: Nonnegative matrix factorization: an analytical and interpretive tool in computational biology. PLoS computational biology **4**(7), 1000029 (2008)

[4] Ding, C., He, X., Simon, H.D.: On the equivalence of nonnegative matrix factorization and spectral clustering. In: Proceedings of the 2005 SIAM International Conference on Data Mining, pp. 606–610 (2005)

[5] Kubjas, K., Robeva, E., Sturmfels, B.: Fixed points of the EM algorithm and nonnegative rank boundaries. The Annals of Statistics **43**(1), 422–461 (2015)

[6] Fawzi, H., Gouveia, J., Parrilo, P.A., Robinson, R.Z., Thomas, R.R.: Positive semidefinite rank. Mathematical Programming **153**, 133–177 (2015)

[7] Yannakakis, M.: Expressing combinatorial optimization problems by linear programs. Journal of Computer and System Sciences **43**(3), 441–466 (1991) https://doi.org/10.1016/0022-0000(91)90024-Y

[8] Fiorini, S., Rothvoß, T., Tiwary, H.R.: Extended formulations for polygons. Discrete Comput. Geom. **48**(3), 658–668 (2012)

[9] Rothvoß, T.: The matching polytope has exponential extension complexity. Journal of the ACM (JACM) **64**(6), 1–19 (2017)

[10] Lee, J.R., Raghavendra, P., Steurer, D.: Lower bounds on the size of semidefinite programming relaxations. In: Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing, pp. 567–576 (2015)

[11] Rothvoß, T.: Some 0/1 polytopes need exponential size extended formulations. Mathematical Programming **142**, 255–268 (2011)

[12] Briët, J., Dadush, D., Pokutta, S.: On the existence of 0/1 polytopes with high semidefinite extension complexity. Math. Program. **153**(1), 179–199 (2015)

[13] Averkov, G., Kaibel, V., Weltge, S.: Maximum semidefinite and linear extension complexity of families of polytopes. Mathematical Programming **167**, 381–394 (2018)

[14] Linial, N., Mendelson, S., Schechtman, G., Shraibman, A.: Complexity measures of sign matrices. Combinatorica **27**(4), 439–463 (2007)

[15] Shitov, Y.: Sublinear extensions of polygons (2020) arXiv:1412.0728 [math.CO]

[16] Kwan, M., Sauermann, L., Zhao, Y.: Extension complexity of low-dimensional polytopes. Transactions of the American Mathematical Society **375**(6), 4209–4250 (2022)

[17] Nesterov, Y., Nemirovskii, A.: Interior-point Polynomial Algorithms in Convex Programming. SIAM Studies in Applied Mathematics, vol. 13. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, USA (1994)

[18] Lee, Y.T., Yue, M.-C.: Universal barrier is n-self-concordant. Math. Oper. Res. **46**(3), 1129–1148 (2021)

[19] Nesterov, Y., Todd, M.J.: Self-scaled barriers and interior-point methods for convex programming. Math. Oper. Res. **22**(1), 1–42 (1997)

[20] Tunçel, L.: Primal-dual symmetry and scale invariance of interior-point algorithms for convex optimization. Math. Oper. Res. **23**(3), 708–718 (1998)

[21] Nesterov, Y., Tunçel, L.: Local superlinear convergence of polynomial-time interior-point methods for hyperbolicity cone optimization problems. SIAM J. Optim. **26**(1), 139–170 (2016)

[22] Nesterov, Y.: Lectures on Convex Optimization. Springer Optimization and Its Applications, vol. 137. Springer, Cham, Switzerland (2018)

[23] Tunçel, L., Vandenberghe, L.: Linear optimization over homogeneous matrix cones. Acta Numerica **32**, 675–747 (2023)

[24] Nesterov, Y.E., Todd, M.J.: Primal-dual interior-point methods for self-scaled cones. SIAM J. Optim. **8**(2), 324–364 (1998)

[25] Güler, O., Tunçel, L.: Characterization of the barrier parameter of homogeneous convex cones. Math. Programming **81**(1, Ser. A), 55–76 (1998)

[26] Hauser, R.A., Güler, O.: Self-scaled barrier functions on symmetric cones and their classification. Found. Comput. Math. **2**(2), 121–143 (2002)

[27] Todd, M.J.: Largest dual ellipsoids inscribed in dual cones. Math. Program. **117**(1-2, Ser. B), 425–434 (2009)

[28] Ito, M., Lourenço, B.F.: A bound on the Carathéodory number. Linear Algebra Appl. **532**, 347–363 (2017)

[29] Averkov, G.: Optimal size of linear matrix inequalities in semidefinite approaches to polynomial optimization. SIAM J. Appl. Algebra Geom. **3**(1), 128–151 (2019)

[30] Saunderson, J.: Limitations on the expressive power of convex cones without long

chains of faces. SIAM J. Optim. **30**(1), 1033–1047 (2020)

[31] Güler, O.: Hyperbolic polynomials and interior point methods for convex programming. Math. Oper. Res. **22**(2), 350–377 (1997)

[32] Anh Truong, V., Tunçel, L.: Geometry of homogeneous convex cones, duality mapping, and optimal self-concordant barriers. Mathematical Programming **100**(2), 295–316 (2004) https://doi.org/10.1007/s10107-003-0470-y

[33] Chua, C.B., Tunçel, L.: Invariance and efficiency of convex representations. Math. Program. **111**(1-2), 113–140 (2008)

[34] Rogers, C.A.: Covering a sphere with spheres. Mathematika **10**, 157–164 (1963)

[35] Verger-Gaugry, J.-L.: Covering a ball with smaller equal balls in $\mathbb{R}^n$. Discrete Comput. Geom. **33**(1), 143–155 (2005)