GENERALIZED POWER CONES: OPTIMAL ERROR BOUNDS AND AUTOMORPHISMS[*]

YING LIN[†], SCOTT B. LINDSTROM[‡], BRUNO F. LOURENÇO[§], AND TING KEI PONG[†]

Abstract. Error bounds are a requisite for trusting or distrusting solutions in an informed way. Until recently, provable error bounds in the absence of constraint qualifications were unattainable for many classes of cones that do not admit projections with known succinct expressions. We build such error bounds for the generalized power cones, using the recently developed framework of one-step facial residual functions. We also show that our error bounds are tight in the sense of that framework. Besides their utility for understanding solution reliability, the error bounds we discover have additional applications to the algebraic structure of the underlying cone, which we describe. In particular we use the error bounds to compute the automorphisms of the generalized power cones, and to identify a set of generalized power cones that are self-dual, irreducible, nonhomogeneous, and perfect.

1. Introduction. In a Euclidean space $\mathcal{E}$, consider the conic feasibility problem:

(Feas)   find $\boldsymbol{x} \in (\mathcal{L} + \boldsymbol{a}) \cap \mathcal{K}$,

where $\mathcal{L}$ is a subspace, $\boldsymbol{a} \in \mathcal{E}$, and $\mathcal{K}$ is a closed convex cone. We desire an upper bound on the distance from an arbitrary $\boldsymbol{x}$ to the feasible region $(\mathcal{L} + \boldsymbol{a}) \cap \mathcal{K}$. The upper bound we seek should depend on the two distances between $\boldsymbol{x}$ and $\mathcal{K}$, and between $\boldsymbol{x}$ and $\mathcal{L} + \boldsymbol{a}$, respectively. Such a guarantee is a kind of *error bound*; error bounds are a fundamental topic in the optimization literature [13, 18, 26, 33, 46] and widely used in the convergence analysis of algorithms. Typically, (Feas) is almost never solved exactly; instead algorithms and solvers often return an approximate solution. Then, error bounds can be used to evaluate the trustworthiness of approximate solutions because they tell us how close they are to the true set of feasible solutions.

In this paper, we consider the case when $\mathcal{K} = \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ is the generalized power cone

$$\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} = \left\{ \boldsymbol{x} = (\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) \in \mathbb{R}^{m+n} \,\middle|\, \|\overline{\boldsymbol{x}}\| \leq \prod_{i=1}^{n} \widetilde{x}_i^{\alpha_i}, \overline{\boldsymbol{x}} \in \mathbb{R}^m, \widetilde{\boldsymbol{x}} \in \mathbb{R}_+^n \right\},$$

[†]Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, People's Republic of China (ying.lin@connect.polyu.hk, tk.pong@polyu.edu.hk).

[‡]Centre for Optimisation and Decision Science, Curtin University, Perth 6102, Australia (scott.lindstrom@curtin.edu.au).

[§]Department of Statistical Inference and Mathematics, Institute of Statistical Mathematics, Tokyo 190-8562, Japan (bruno@ism.ac.jp).

where $m \geq 1$, $n \geq 2$, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in (0, 1)^n$ with $\sum_{i=1}^{n} \alpha_i = 1$, and $\|\overline{\boldsymbol{x}}\|$ denotes the Euclidean norm of $\overline{\boldsymbol{x}}$. In the specific case when $m \geq 1$, $n = 2$, and $\boldsymbol{\alpha} = (1/2, 1/2)$, $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ is isomorphic to a second-order cone, whose worst-case error bound is known to be Hölderian with exponent $1/2$, thanks to the work of Luo and Sturm [25]. The remaining cases, while not as well known as the second-order cone case, admit more direct modeling of certain problems and have found applications in geometric programs, generalized location problems, and portfolio optimization [5, 28]. More broadly, the inclusion of the power cone[1] (and the exponential cone) makes all the convex instances from the MINLPLib2 benchmark library conic representable [24, 27]. This broad utility has motivated the development of self-concordant barriers [5, 43, 36] and the ongoing development of specialized interior point methods [30, 37]. Optimization with the generalized power cones is implemented in commercial and open source solvers like MOSEK, Alfonso, DDS, and Hypatia [6, 15, 34, 28].

One of this paper's main contributions—Theorem 3.10—is a complete error bound analysis for the generalized power cone problem (Feas). The generalized power cone cases pose two significant obstructions to error bound analysis that are not present in the second-order cone case. First, known forms for projections onto generalized power cones do not admit simple representations [12]; second, their facial structure is more complicated. We obviate the first obstruction via the framework of one-step facial residual functions ($\mathbb{1}$-FRFs), which was established in [19, 20]. The second challenge, facial complexity, we tackle directly. In particular, we build $\mathbb{1}$-FRFs for all faces of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. All these $\mathbb{1}$-FRFs are tight in the natural sense of [20]. Consequently, all of the error bounds in Theorem 3.10 are tight in this sense.

While error bounds are typically used in convergence analysis and to evaluate the quality of approximate solutions, our approach via $\mathbb{1}$-FRFs admits a surprising additional application to the algebraic structure of the underlying cone. In order to explain our next results, we recall a few concepts. The *automorphism group* of a cone $\mathcal{K}$ is the set of the bijective linear operators $\boldsymbol{A}$ satisfying $\boldsymbol{A}\mathcal{K} = \mathcal{K}$. A cone is said to be *homogeneous* if its automorphism group acts transitively on its relative interior. We say that a cone is *irreducible* if it is not the direct sum of two nontrivial cones whose spans only intersect at the origin.

Because automorphisms of cones must preserve optimal FRFs (up to positively rescaled shifts), we can use our results to establish the automorphism group for $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ in Theorem 4.1 and compute its dimension in Theorem 4.2.

This is useful because the automorphism group of a closed convex cone $\mathcal{K}$ has important implications for complementarity problems over $\mathcal{K}$; see [9]. In particular, denoting the dual cone of $\mathcal{K}$ by $\mathcal{K}^*$, a complementarity condition of the form "$\boldsymbol{x} \in \mathcal{K}, \boldsymbol{y} \in \mathcal{K}^*, \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$" can be split into a square system of equations if and only if the dimension of the automorphism group of $\mathcal{K}$ is at least $\dim \mathcal{K}$; see [32, Theorem 1]. In this case, $\mathcal{K}$ is said to be a *perfect cone*.

Many of the concrete examples of irreducible perfect cones in the literature correspond to homogeneous cones. In this paper we will show that the generalized power cone is irreducible, perfect (when $m \geq 3$), and, except when it reduces to the second-order cone case, always nonhomogeneous. This gives an interesting example of an irreducible cone with good complementarity properties that is not a homogeneous cone. To summarize, our main contributions are as follows:

1. We completely determine the tightest possible error bounds for the generalized power cone; see Theorem 3.10.

---

[1]This refers to $\mathcal{P}_{1,2}^{\boldsymbol{\alpha}}$.

2. Using our error bounds, we completely determine the automorphism group of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ and discuss some theoretical questions related to homogeneity and perfectness (in the sense of [10, 9]); see section 4.

Although we do not discuss the details, we mention in passing that determining the error bound associated to conic linear systems makes it possible to compute the KL-exponent of certain functions, as done, for example, in [20, section 5.1] using results from [45]. See more on the connection between error bounds, KL-exponents, and convergence rates in [3].

This paper is organized as follows. In section 2, we recall notation and preliminaries. In section 3, we furnish the eponymous error bounds. In section 4, we provide the further application to the algebraic structure of the generalized power cones.

2. Notation and preliminaries. We will use italic letters to represent real scalars, bold lowercase letters to denote vectors, bold uppercase letters to stand for matrices,[2] and calligraphic capital letters for (sub)spaces and sets. Let $\mathcal{E}$ be a finite-dimensional Euclidean space, and let $\mathbb{R}_+$ and $\mathbb{R}_-$ be the set of nonnegative and nonpositive real numbers, respectively. The inner product of $\mathcal{E}$ is denoted by $\langle \cdot, \cdot \rangle$ and the induced norm by $\|\cdot\|$. With that, for $\boldsymbol{x} \in \mathcal{E}$ and a closed convex set $\mathcal{C} \subseteq \mathcal{E}$, we denote the projection of $\boldsymbol{x}$ onto $\mathcal{C}$ by $P_{\mathcal{C}}(\boldsymbol{x})$ so that $P_{\mathcal{C}}(\boldsymbol{x}) = \arg\min_{\boldsymbol{y} \in \mathcal{C}} \|\boldsymbol{x} - \boldsymbol{y}\|$ and the distance between $\boldsymbol{x}$ and $\mathcal{C}$ by $\operatorname{dist}(\boldsymbol{x}, \mathcal{C}) = \inf_{\boldsymbol{y} \in \mathcal{C}} \|\boldsymbol{x} - \boldsymbol{y}\| = \|\boldsymbol{x} - P_{\mathcal{C}}(\boldsymbol{x})\|$. For any $\boldsymbol{x} \in \mathcal{E}$ and $\eta \geq 0$, we denote the ball centered at $\boldsymbol{x}$ with radius η by $\mathcal{B}(\boldsymbol{x}; \eta) := \{\boldsymbol{y} \in \mathcal{E} \mid \|\boldsymbol{y} - \boldsymbol{x}\| \leq \eta\}$; we write $\mathcal{B}(\eta)$ for the ball centered at $\boldsymbol{0}$ with radius η for simplicity. A diagonal matrix with diagonal vector being $\boldsymbol{x}$ is denoted by $\operatorname{Diag}(\boldsymbol{x})$. Meanwhile, we use $\mathcal{C}^{\perp}$ to denote the orthogonal complement of $\mathcal{C}$ and $\boldsymbol{I}_n$ to represent the $n \times n$ identity matrix.

We now recall the definition of Lipschitzian and Hölderian error bounds. Let $\mathcal{C}_1, \mathcal{C}_2 \subseteq \mathcal{E}$ be closed convex sets with $\mathcal{C}_1 \cap \mathcal{C}_2 \neq \emptyset$. We say that $\mathcal{C}_1, \mathcal{C}_2$ satisfy a *uniform Hölderian error bound* with exponent $\gamma \in (0, 1]$ if for every bounded set $\mathcal{B} \subseteq \mathcal{E}$ there exists a constant $\kappa_{\mathcal{B}}$ such that $\operatorname{dist}(\boldsymbol{x}, \mathcal{C}_1 \cap \mathcal{C}_2) \leq \kappa_{\mathcal{B}} \max\{\operatorname{dist}(\boldsymbol{x}, \mathcal{C}_1), \operatorname{dist}(\boldsymbol{x}, \mathcal{C}_2)\}^{\gamma}$ for all $\boldsymbol{x} \in \mathcal{B}$. If $\gamma = 1$, then the error bound is said to be *Lipschitzian*.

Let $\mathcal{K} \subseteq \mathcal{E}$ be a closed convex cone and $\mathcal{K}^*$ its dual cone. We will use $\operatorname{int}\mathcal{K}$, $\operatorname{ri}\mathcal{K}$, $\partial\mathcal{K}$, $\operatorname{span}\mathcal{K}$, $\dim\mathcal{K}$ to denote the interior, relative interior, boundary, linear span, and dimension of $\mathcal{K}$, respectively. If $\mathcal{K} \cap -\mathcal{K} = \{\boldsymbol{0}\}$, we say that $\mathcal{K}$ is *pointed*.

A *face* of $\mathcal{K}$ is a closed convex cone $\mathcal{F}$ satisfying $\mathcal{F} \subseteq \mathcal{K}$ and the property that if $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{K}$ and $\boldsymbol{x} + \boldsymbol{y} \in \mathcal{F}$, then $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{F}$.[3] We write $\mathcal{F} \trianglelefteq \mathcal{K}$ if $\mathcal{F}$ is a face of $\mathcal{K}$ and $\mathcal{F} \trianglelefteq \mathcal{K}$ if $\mathcal{F}$ is a *proper* face of $\mathcal{K}$, i.e., $\mathcal{F} \neq \mathcal{K}$. A face $\mathcal{F}$ is said to be *nontrivial* if $\mathcal{F}$ is proper and $\mathcal{F} \neq \mathcal{K} \cap -\mathcal{K}$. If $\mathcal{F} = \mathcal{K} \cap \{\boldsymbol{z}\}^{\perp}$ for some $\boldsymbol{z} \in \mathcal{K}^*$, $\mathcal{F}$ is called an *exposed face* of $\mathcal{K}$.

The facial structure of the closed convex cone $\mathcal{K}$ is important for deducing *error bounds* for (Feas); see the seminal work of Sturm [38]. Recently, a new framework based on the *facial reduction algorithm* [4, 35, 44] and *one-step facial residual functions (1-FRFs)* [19, Definition 3.4] was proposed for establishing error bounds for (Feas) without requiring any constraint qualifications; see [22, 19, 20]. Next, we present a very brief overview of the framework; for more detailed explanations and the underlying intuition behind the techniques, see [19, 20]. First, we recall the definition of one-step facial residual functions.

DEFINITION 2.1 (one-step facial residual function [19, Definition 3.4]). *Let $\mathcal{K}$ be a closed convex cone and $\boldsymbol{z} \in \mathcal{K}^*$. Suppose that $\psi_{\mathcal{K},\boldsymbol{z}} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies the following:*

---

[2]With an abuse of notation, we use $\boldsymbol{0}$ to denote a zero vector/matrix, whose dimension should be clear from the context.

[3]By convention, we discard the empty face.

(i) $\psi_{\mathcal{K},\boldsymbol{z}}$ *is nonnegative, nondecreasing in each argument, and for every* $t \in \mathbb{R}_+$, $\psi_{\mathcal{K},\boldsymbol{z}}(0,t) = 0$.

(ii) *The following implication holds for any* $\boldsymbol{x} \in \operatorname{span} \mathcal{K}$ *and* $\epsilon \geq 0$:

$$\operatorname{dist}(\boldsymbol{x},\mathcal{K}) \leq \epsilon, \ \langle \boldsymbol{x},\boldsymbol{z} \rangle \leq \epsilon \implies \operatorname{dist}(\boldsymbol{x},\mathcal{K} \cap \{\boldsymbol{z}\}^{\perp}) \leq \psi_{\mathcal{K},\boldsymbol{z}}(\epsilon,\|\boldsymbol{x}\|).$$

Then, $\psi_{\mathcal{K},\boldsymbol{z}}$ *is said to be a one-step facial residual function (*$\mathbb{1}$*-FRF) for* $\mathcal{K}$ *and* $\boldsymbol{z}$.

The basic idea of the aforementioned framework is as follows. Suppose that in each step of the facial reduction algorithm we can find a suitable $\mathbb{1}$-FRF for the "current" face and "next" exposing vector until we reach a face $\mathcal{F}$ such that $\mathcal{F}$ and $\mathcal{L}+\boldsymbol{a}$ satisfy the *partial polyhedral Slater's (PPS) condition* [22, Definition 3].[4] Then, we can construct an error bound for $\mathcal{K}$ and $\mathcal{L}+\boldsymbol{a}$ by composing these residual functions in a specific manner. In this regard, if (Feas) is feasible, we define the *distance to the PPS condition* of (Feas), denoted by $d_{\mathrm{PPS}}(\mathcal{K},\mathcal{L}+\boldsymbol{a})$, as the length *minus one* of the *shortest* chain of faces (among those chains constructed as in [22, Proposition 5]) such that the PPS condition holds for the final face in the chain and $\mathcal{L}+\boldsymbol{a}$.

We end this section with the following lemma, which is useful in the analysis of one-dimensional faces. It will be used repeatedly in our subsequent discussions.

LEMMA 2.2 ([20, Lemma 2.5]). *Let* $\mathcal{K}$ *be a pointed closed convex cone, and let* $\boldsymbol{z} \in \partial\mathcal{K}^* \setminus \{\boldsymbol{0}\}$ *be such that* $\mathcal{F} := \{\boldsymbol{z}\}^{\perp} \cap \mathcal{K}$ *is a one-dimensional proper face of* $\mathcal{K}$. *Let* $\boldsymbol{f} \in \mathcal{K} \setminus \{\boldsymbol{0}\}$ *be such that* $\mathcal{F} = \{t\boldsymbol{f} \,|\, t \geq 0\}$. *Let* $\eta > 0$ *and* $\boldsymbol{v} \in \partial\mathcal{K} \cap B(\eta) \setminus \mathcal{F}$, *and let* $\boldsymbol{w} = P_{\{\boldsymbol{z}\}^{\perp}}(\boldsymbol{v})$ *and* $\boldsymbol{u} = P_{\mathcal{F}}(\boldsymbol{w})$ *with* $\boldsymbol{u} \neq \boldsymbol{w}$. *Then it holds that* $\langle \boldsymbol{f},\boldsymbol{z} \rangle = 0$ *and we have*

$$\|\boldsymbol{v} - \boldsymbol{w}\| = \frac{|\langle \boldsymbol{z},\boldsymbol{v} \rangle|}{\|\boldsymbol{z}\|}, \qquad \|\boldsymbol{u} - \boldsymbol{w}\| = \begin{cases} \left\| \boldsymbol{v} - \frac{\langle \boldsymbol{z},\boldsymbol{v} \rangle}{\|\boldsymbol{z}\|^2}\boldsymbol{z} - \frac{\langle \boldsymbol{f},\boldsymbol{v} \rangle}{\|\boldsymbol{f}\|^2}\boldsymbol{f} \right\| & \text{if } \langle \boldsymbol{f},\boldsymbol{v} \rangle \geq 0, \\ \left\| \boldsymbol{v} - \frac{\langle \boldsymbol{z},\boldsymbol{v} \rangle}{\|\boldsymbol{z}\|^2}\boldsymbol{z} \right\| & \text{otherwise .} \end{cases}$$

Moreover, when $\langle \boldsymbol{f},\boldsymbol{v} \rangle \geq 0$ *(or, equivalently,* $\langle \boldsymbol{f},\boldsymbol{w} \rangle \geq 0$*), we have* $\boldsymbol{u} = P_{\operatorname{span}\mathcal{F}}(\boldsymbol{w})$. *On the other hand, if* $\langle \boldsymbol{f},\boldsymbol{v} \rangle < 0$, *we have* $\boldsymbol{u} = \boldsymbol{0}$.

3. Error bounds for the generalized power cone. We consider the generalized power cone and its dual. Let $m \geq 1$, $n \geq 2$, and $\boldsymbol{\alpha} = (\alpha_1,\dots,\alpha_n) \in (0,1)^n$ with $\sum_{i=1}^{n} \alpha_i = 1$; the generalized power cone $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ and its dual $(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^*$ are given, respectively, by

(3.1)
$$\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} = \left\{ \boldsymbol{x} = (\overline{\boldsymbol{x}},\widetilde{\boldsymbol{x}}) \in \mathbb{R}^{m+n} \,\middle|\, \|\overline{\boldsymbol{x}}\| \leq \prod_{i=1}^{n} \widetilde{x}_i^{\alpha_i},\ \overline{\boldsymbol{x}} \in \mathbb{R}^m,\ \widetilde{\boldsymbol{x}} \in \mathbb{R}_+^n \right\},$$
$$(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* = \left\{ \boldsymbol{z} = (\overline{\boldsymbol{z}},\widetilde{\boldsymbol{z}}) \in \mathbb{R}^{m+n} \,\middle|\, \|\overline{\boldsymbol{z}}\| \leq \prod_{i=1}^{n} \left(\frac{\widetilde{z}_i}{\alpha_i}\right)^{\alpha_i},\ \overline{\boldsymbol{z}} \in \mathbb{R}^m,\ \widetilde{\boldsymbol{z}} \in \mathbb{R}_+^n \right\}.$$

Here, given a vector $\boldsymbol{x} \in \mathbb{R}^{m+n}$, we let $\overline{\boldsymbol{x}} \in \mathbb{R}^m$ be the vector corresponding to its first m entries and let $\widetilde{\boldsymbol{x}} \in \mathbb{R}^n$ be the vector corresponding to its last n entries.

In this section, we will prove the main result of our paper: a complete analysis of the error bounds of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. This will require an analysis of the facial structure of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, which we will do shortly after the following lemmas.

---

[4]We note that this implies a Lipschitzian error bound holds for $\mathcal{F}$ and $\mathcal{L}+\boldsymbol{a}$; see [2, Corollary 3] and the discussion preceding [19, Proposition 2.3].

LEMMA 3.1. *Let $n \geq 2$ and $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n) \in (0,1)^n$ with $\sum_{i=1}^n \alpha_i = 1$. Let $\boldsymbol{\zeta} \in \operatorname{int} \mathbb{R}^n_-$ satisfy $\prod_{i=1}^n (-\zeta_i/\alpha_i)^{\alpha_i} = 1$. Define $\widetilde{\boldsymbol{\zeta}} := -\boldsymbol{\alpha} \circ \boldsymbol{\zeta}^{-1}$, where $\circ$ is the Hadamard product and the inverse is taken componentwise. Then there exist $C > 0$ and $\epsilon > 0$ so that*

$$(3.2) \quad -1 - \langle \boldsymbol{\zeta}, \boldsymbol{\omega} \rangle \geq C \| \boldsymbol{\omega} - \widetilde{\boldsymbol{\zeta}} \|^2 \text{ whenever } \boldsymbol{\omega} \in \operatorname{int} \mathbb{R}^n_+, \| \boldsymbol{\omega} - \widetilde{\boldsymbol{\zeta}} \| \leq \epsilon \text{ and } \prod_{i=1}^n \omega_i^{\alpha_i} = 1.$$

Moreover, for any $\boldsymbol{\omega} \in \operatorname{int} \mathbb{R}^n_+$ satisfying $\prod_{i=1}^n \omega_i^{\alpha_i} = 1$, it holds that $\langle \boldsymbol{\zeta}, \boldsymbol{\omega} \rangle \leq -1$; furthermore, we have $\langle \boldsymbol{\zeta}, \boldsymbol{\omega} \rangle = -1$ if and only if $\boldsymbol{\omega} = \widetilde{\boldsymbol{\zeta}}$.

Proof. For each i, we see from the Taylor series of $\ln(\cdot)$ at $\widetilde{\zeta}_i > 0$ that

$$\ln(\omega_i) = \ln(\widetilde{\zeta}_i) + \widetilde{\zeta}_i^{-1}(\omega_i - \widetilde{\zeta}_i) - \widetilde{\zeta}_i^{-2}(\omega_i - \widetilde{\zeta}_i)^2 + O(|\omega_i - \widetilde{\zeta}_i|^3) \text{ as } \omega_i \to \widetilde{\zeta}_i, \ \omega_i > 0.$$

Thus, there exist $\epsilon_i > 0$ and $c_i > 0$ so that

$$\ln(\widetilde{\zeta}_i) \geq \ln(\omega_i) - \widetilde{\zeta}_i^{-1}(\omega_i - \widetilde{\zeta}_i) + c_i(\omega_i - \widetilde{\zeta}_i)^2 \text{ whenever } |\omega_i - \widetilde{\zeta}_i| \leq \epsilon_i \text{ and } \omega_i > 0.$$

Let $\epsilon := \min_{1 \leq i \leq n} \epsilon_i > 0$. Multiplying both sides of the above inequality by α_i and summing the resulting inequalities from $i = 1$ to n, we see that whenever $\boldsymbol{\omega} \in \operatorname{int} \mathbb{R}^n_+$ satisfies $\| \boldsymbol{\omega} - \widetilde{\boldsymbol{\zeta}} \| \leq \epsilon$ and $\prod_{i=1}^n \omega_i^{\alpha_i} = 1$, we have

$$0 \overset{(a)}{=} \sum_{i=1}^n \alpha_i \ln(\widetilde{\zeta}_i) \geq \sum_{i=1}^n \alpha_i \ln(\omega_i) - \sum_{i=1}^n \alpha_i \widetilde{\zeta}_i^{-1}(\omega_i - \widetilde{\zeta}_i) + \sum_{i=1}^n \alpha_i c_i(\omega_i - \widetilde{\zeta}_i)^2$$

$$\overset{(b)}{=} -\sum_{i=1}^n \alpha_i \widetilde{\zeta}_i^{-1}(\omega_i - \widetilde{\zeta}_i) + \sum_{i=1}^n \alpha_i c_i(\omega_i - \widetilde{\zeta}_i)^2 \overset{(c)}{=} -\sum_{i=1}^n \alpha_i \widetilde{\zeta}_i^{-1}\omega_i + 1 + \sum_{i=1}^n \alpha_i c_i(\omega_i - \widetilde{\zeta}_i)^2$$

$$= \sum_{i=1}^n \zeta_i \omega_i + 1 + \sum_{i=1}^n \alpha_i c_i(\omega_i - \widetilde{\zeta}_i)^2,$$

where (a) and (b) hold because $\prod_{i=1}^n \widetilde{\zeta}_i^{\alpha_i} = \prod_{i=1}^n \omega_i^{\alpha_i} = 1$, (c) uses the fact that $\sum_{i=1}^n \alpha_i = 1$, and the last equality follows from the definition of $\widetilde{\boldsymbol{\zeta}}$. Rearranging the above inequality, we conclude that (3.2) holds with $C = \min_{1 \leq i \leq n} \alpha_i c_i > 0$.

Next, let $\boldsymbol{\omega} \in \operatorname{int} \mathbb{R}^n_+$ satisfy $\prod_{i=1}^n \omega_i^{\alpha_i} = 1$. Then $(-1, \boldsymbol{\omega}) \in \mathcal{P}_{1,n}^{\boldsymbol{\alpha}}$. Recall from the assumption that $(1, -\boldsymbol{\zeta}) \in (\mathcal{P}_{1,n}^{\boldsymbol{\alpha}})^*$. From these we deduce $\langle \boldsymbol{\zeta}, \boldsymbol{\omega} \rangle \leq -1$. If $\langle \boldsymbol{\zeta}, \boldsymbol{\omega} \rangle = -1$, then

$$\sum_{i=1}^n \alpha_i \left(\frac{-\zeta_i}{\alpha_i} \right) \omega_i = \sum_{i=1}^n (-\zeta_i)\omega_i = 1 = \prod_{i=1}^n \omega_i^{\alpha_i} = \prod_{i=1}^n \left(\frac{-\zeta_i}{\alpha_i} \right)^{\alpha_i} \prod_{i=1}^n \omega_i^{\alpha_i}.$$

Taking $\ln$ on both sides of the above equality, we see that

$$\ln \left[\sum_{i=1}^n \alpha_i \left(\frac{-\zeta_i}{\alpha_i} \right) \omega_i \right] = \sum_{i=1}^n \alpha_i \ln \left[\left(\frac{-\zeta_i}{\alpha_i} \right) \omega_i \right].$$

Since $\ln$ is strictly concave and $\alpha_i \in (0,1)$ for all i, we conclude that there exists $c > 0$ so that $\omega_i \cdot (-\zeta_i/\alpha_i) = c$ for all i. This, together with the facts that $\prod_{i=1}^n \omega_i^{\alpha_i} = \prod_{i=1}^n (-\zeta_i/\alpha_i)^{\alpha_i} = 1$ and $\sum_{i=1}^n \alpha_i = 1$, gives $c = 1$. It thus follows that $\boldsymbol{\omega} = \widetilde{\boldsymbol{\zeta}}$. Conversely, it is routine to check that if $\boldsymbol{\omega} = \widetilde{\boldsymbol{\zeta}}$, then $\prod_{i=1}^n \omega_i^{\alpha_i} = 1$ and $\langle \boldsymbol{\zeta}, \boldsymbol{\omega} \rangle = -1$. $\square$

The next lemma is obtained by applying [20, Lemma 4.1] with $p = q = 2$.

LEMMA 3.2. *Let $\boldsymbol{\zeta} \in \mathbb{R}^m$ ($m \geq 1$) satisfy $\|\boldsymbol{\zeta}\| = 1$. Define $\overline{\boldsymbol{\zeta}} := -\boldsymbol{\zeta}$. Then there exist $C > 0$ and $\epsilon > 0$ so that*

(3.3)
$$1 + \langle \boldsymbol{\zeta}, \boldsymbol{w} \rangle \geq C \sum_{i \in I} |w_i - \overline{\zeta}_i|^2 + \frac{1}{2} \sum_{i \notin I} |w_i|^2 \quad \text{whenever } \|\boldsymbol{w} - \overline{\boldsymbol{\zeta}}\| \leq \epsilon \quad \text{and} \quad \|\boldsymbol{w}\| = 1,$$

where $I = \{i \,|\, \overline{\zeta}_i \neq 0\}$. Furthermore, for any $\boldsymbol{w}$ satisfying $\|\boldsymbol{w}\| \leq 1$, it holds that $\langle \boldsymbol{\zeta}, \boldsymbol{w} \rangle \geq -1$, with the equality holding if and only if $\boldsymbol{w} = \overline{\boldsymbol{\zeta}}$.

3.1. The facial structure of $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$. In this subsection, we discuss the faces of $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$. We first characterize the proper nontrivial exposed faces of $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$ in the following proposition.

PROPOSITION 3.3 (proper nontrivial exposed faces of $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$). *Let $\boldsymbol{z} = (\overline{\boldsymbol{z}}, \widetilde{\boldsymbol{z}}) \in \partial(\mathcal{P}^{\boldsymbol{\alpha}}_{m,n})^* \backslash \{\boldsymbol{0}\}$.*
   (i) *If $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$, then $\boldsymbol{z}$ exposes the following one-dimensional face:*

   (3.4)
   $$\mathcal{F}_r := \{\boldsymbol{z}\}^\perp \cap \mathcal{P}^{\boldsymbol{\alpha}}_{m,n} = \{t\boldsymbol{f} \in \mathbb{R}^{m+n} \,|\, t \geq 0\} \text{ with } \boldsymbol{f} = (-\overline{\boldsymbol{z}}/\|\overline{\boldsymbol{z}}\|^2, \boldsymbol{\alpha} \circ \widetilde{\boldsymbol{z}}^{-1}),$$

   where the inverse is taken componentwise.
   (ii) *If $\overline{\boldsymbol{z}} = \boldsymbol{0}$, then $\boldsymbol{z}$ exposes the following face of dimension $n - |\mathcal{I}|$:*

   (3.5) $$\mathcal{F}_{\boldsymbol{z}} := \{\boldsymbol{z}\}^\perp \cap \mathcal{P}^{\boldsymbol{\alpha}}_{m,n} = \{\boldsymbol{x} = (\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) \in \mathbb{R}^{m+n}_+ \,|\, \overline{\boldsymbol{x}} = \boldsymbol{0}, \, \widetilde{x}_i = 0 \,\, \forall i \in \mathcal{I}\},$$

   where $\mathcal{I} := \{i \,|\, \widetilde{z}_i > 0\} \neq \emptyset$ and $|\mathcal{I}|$ denotes the cardinality of $\mathcal{I}$.

Proof. (i) Notice that $\boldsymbol{x} = (\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) \in \{\boldsymbol{z}\}^\perp \cap \mathcal{P}^{\boldsymbol{\alpha}}_{m,n} \backslash \{\boldsymbol{0}\}$ if and only if $\boldsymbol{x} \in \partial \mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$, $\boldsymbol{x} \neq \boldsymbol{0}$, and

(3.6) $$\langle \overline{\boldsymbol{z}}, \overline{\boldsymbol{x}} \rangle + \langle \widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{x}} \rangle = 0.$$

The above relation yields

(3.7) $$\sum_{i=1}^n \widetilde{z}_i \widetilde{x}_i = -\langle \overline{\boldsymbol{z}}, \overline{\boldsymbol{x}} \rangle \leq \|\overline{\boldsymbol{z}}\| \|\overline{\boldsymbol{x}}\| \leq \prod_{i=1}^n \left(\frac{\widetilde{x}_i \widetilde{z}_i}{\alpha_i} \right)^{\alpha_i},$$

where the last inequality follows from the definition of $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$ in (3.1).

Note that $\widetilde{x}_i$ cannot be *all* zero, for otherwise $\overline{\boldsymbol{x}}$ will also be zero since $\boldsymbol{x} \in \partial \mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$, which contradicts $\boldsymbol{x} \neq \boldsymbol{0}$. In addition, we must have $\widetilde{z}_i > 0$ for all i because $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$ and $\boldsymbol{z} \in \partial(\mathcal{P}^{\boldsymbol{\alpha}}_{m,n})^* \backslash \{\boldsymbol{0}\}$. Using these observations, we have $\sum_{i=1}^n \widetilde{z}_i \widetilde{x}_i > 0$. Combining this with (3.7), we deduce that $\widetilde{x}_i \widetilde{z}_i > 0$ for all i. Now we can take ln on both sides of (3.7) to obtain

(3.8) $$\ln \left[\sum_{i=1}^n \alpha_i \left(\frac{\widetilde{x}_i \widetilde{z}_i}{\alpha_i} \right) \right] \leq \alpha_1 \ln \left(\frac{\widetilde{x}_1 \widetilde{z}_1}{\alpha_1} \right) + \cdots + \alpha_n \ln \left(\frac{\widetilde{x}_n \widetilde{z}_n}{\alpha_n} \right).$$

Using this together with the fact that $\ln(\cdot)$ is strictly concave, we deduce that (3.8) holds as an equality. Hence, there exists a constant $c > 0$ so that

(3.9) $$\widetilde{x}_i = c \alpha_i \widetilde{z}_i^{-1} \,\, \forall i = 1, 2, \dots, n.$$

Plugging (3.9) into (3.6), we obtain

$$(3.10) \qquad \langle \overline{\boldsymbol{z}}, \overline{\boldsymbol{x}} \rangle = -\langle \widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{x}} \rangle = -c \sum_{i=1}^{n} \alpha_i = -c.$$

Moreover, using (3.9) and the last relation in (3.7), we see that

$$\|\overline{\boldsymbol{z}}\| \|\overline{\boldsymbol{x}}\| \leq \prod_{i=1}^{n} \left(\frac{\widetilde{x}_i \widetilde{z}_i}{\alpha_i} \right)^{\alpha_i} = c.$$

The two displayed lines above show that $\|\overline{\boldsymbol{z}}\| \|\overline{\boldsymbol{x}}\| = -\langle \overline{\boldsymbol{z}}, \overline{\boldsymbol{x}} \rangle$, which together with $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$ implies that there exists $\kappa > 0$ so that

$$(3.11) \qquad \overline{\boldsymbol{x}} = -\kappa \overline{\boldsymbol{z}}.$$

Plugging (3.11) into (3.10), we obtain that $\kappa = c/\|\overline{\boldsymbol{z}}\|^2$. Using this together with (3.9) and (3.11), we can now conclude that

$$\mathcal{F}_{\mathrm{r}} := \{\boldsymbol{z}\}^{\perp} \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}} = \{t\boldsymbol{f} \in \mathbb{R}^{m+n} \mid t \geq 0\} \text{ with } \boldsymbol{f} = (-\overline{\boldsymbol{z}}/\|\overline{\boldsymbol{z}}\|^2, \boldsymbol{\alpha} \circ \widetilde{\boldsymbol{z}}^{-1}),$$

where the inverse is taken componentwise.

(ii) In this case, $\overline{\boldsymbol{z}} = \boldsymbol{0}$. Then $\mathcal{I} := \{i \mid \widetilde{z}_i > 0\}$ is nonempty because $\boldsymbol{z} \neq \boldsymbol{0}$. Hence, $\boldsymbol{x} = (\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) \in \{\boldsymbol{z}\}^{\perp} \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}} \setminus \{\boldsymbol{0}\}$ if and only if $\boldsymbol{x} \in \partial \mathcal{P}_{m,n}^{\boldsymbol{\alpha}} \setminus \{\boldsymbol{0}\}$ and satisfies

$$\sum_{i \in \mathcal{I}} \widetilde{z}_i \widetilde{x}_i = 0.$$

This means that $\widetilde{x}_i = 0$ whenever $i \in \mathcal{I}$ and hence $\overline{\boldsymbol{x}} = \boldsymbol{0}$. Thus,

$$\mathcal{F}_{\boldsymbol{z}} := \{\boldsymbol{z}\}^{\perp} \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}} = \{\boldsymbol{x} = (\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) \in \mathbb{R}_+^{m+n} \mid \overline{\boldsymbol{x}} = \boldsymbol{0}, \ \widetilde{x}_i = 0 \ \forall i \in \mathcal{I}\}. \qquad \square$$

Having characterized the proper exposed faces of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, we will show that $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ is *projectionally exposed* [4, 39], which means that for every face $\mathcal{F}$ of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ there is a linear operator $\boldsymbol{P}$ satisfying $\boldsymbol{P}(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) = \mathcal{F}$ and $\boldsymbol{P}^2 = \boldsymbol{P}$. In particular, $\boldsymbol{P}$, which depends on $\mathcal{F}$, is a projection that is not necessarily orthogonal. Projectionally exposed cones are both facially exposed [39, Corollary 4.4] and amenable [22, Proposition 9]; see also [23].

PROPOSITION 3.4 (generalized power cones are projectionally exposed). $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ *is projectionally exposed, in particular, all its faces are exposed.*

Proof. Sung and Tam proved in [39, Corollary 4.5] that a sufficient condition for a cone to be projectionally exposed is that all its exposed faces are projectionally exposed. With this in mind, let $\mathcal{F}$ be an exposed face of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. If $\mathcal{F} = \{\boldsymbol{0}\}$ or $\mathcal{F} = \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, then the zero map and the identity map are, respectively, projections mapping $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ to $\mathcal{F}$. Otherwise, $\mathcal{F}$ is a nonzero proper face of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ and is of the form $\{\boldsymbol{z}\}^{\perp} \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ for some $\boldsymbol{z} = (\overline{\boldsymbol{z}}, \widetilde{\boldsymbol{z}}) \in \partial(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* \setminus \{\boldsymbol{0}\}$. By the analysis in cases (i), (ii), we only need to consider two cases.

First, suppose that $\mathcal{F}$ is a one-dimensional face as in (3.4) and let $\boldsymbol{u} \in (\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^*$ be such that $\langle \boldsymbol{f}, \boldsymbol{u} \rangle = 1$. At least one such $\boldsymbol{u}$ exists, since otherwise we would have $\boldsymbol{f} \in ((\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^*)^{\perp} = \{\boldsymbol{0}\}$. Then, $\boldsymbol{P} = \boldsymbol{f}\boldsymbol{u}^{\top}$ satisfies $\boldsymbol{P}^2 = \boldsymbol{P}$ and $\boldsymbol{P}(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) = \mathcal{F}$, as required.

Next, suppose that $\mathcal{F}$ is as in (3.5). Then, we let $\boldsymbol{P}$ be the linear map that maps $(\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}})$ to $(\boldsymbol{0}, \widetilde{\boldsymbol{y}})$ where $\widetilde{y}_i = 0$ if $i \in \mathcal{I}$ and $\widetilde{y}_i = \widetilde{x}_i$ if $i \notin \mathcal{I}$. With that, $\boldsymbol{P}$ is a projection mapping $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ to $\mathcal{F}$. $\qquad \square$

3.2. Deducing error bounds and one-step facial residual functions for $\mathcal{P}_{m,n}^{\alpha}$. We start with the faces $\mathcal{F}_r$ that correspond to a $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\alpha})^* \backslash \{\boldsymbol{0}\}$ with $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$. We have the following result.

THEOREM 3.5. *Let $\boldsymbol{z} = (\overline{\boldsymbol{z}}, \widetilde{\boldsymbol{z}}) \in \partial(\mathcal{P}_{m,n}^{\alpha})^* \backslash \{\boldsymbol{0}\}$ with $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$ and let $\mathcal{F}_r := \{\boldsymbol{z}\}^{\perp} \cap \mathcal{P}_{m,n}^{\alpha}$. Let $\eta > 0$ and define*

$$(3.12) \qquad \gamma_{\boldsymbol{z},\eta} := \inf_{\boldsymbol{v}} \left\{ \frac{\|\boldsymbol{v} - \boldsymbol{w}\|^{\frac{1}{2}}}{\|\boldsymbol{u} - \boldsymbol{w}\|} \;\middle|\; \begin{array}{c} \boldsymbol{v} \in \partial\mathcal{P}_{m,n}^{\alpha} \cap \mathcal{B}(\eta) \backslash \mathcal{F}_r, \; \boldsymbol{w} = P_{\{\boldsymbol{z}\}^{\perp}}(\boldsymbol{v}), \\ \boldsymbol{u} = P_{\mathcal{F}_r}(\boldsymbol{w}), \; \boldsymbol{u} \neq \boldsymbol{w} \end{array} \right\}.$$

Then it holds that $\gamma_{\boldsymbol{z},\eta} \in (0, \infty]$ and that

$$\mathrm{dist}(\boldsymbol{q}, \mathcal{F}_r) \leq \max\{2\sqrt{\eta}, 2\gamma_{\boldsymbol{z},\eta}^{-1}\} \cdot \mathrm{dist}(\boldsymbol{q}, \mathcal{P}_{m,n}^{\alpha})^{\frac{1}{2}} \quad \text{whenever } \boldsymbol{q} \in \{\boldsymbol{z}\}^{\perp} \cap \mathcal{B}(\eta).$$

Proof. Suppose for contradiction that $\gamma_{\boldsymbol{z},\eta} = 0$. Then, in view of [19, Lemma 3.12], there exist $\widehat{\boldsymbol{v}} \in \mathcal{F}_r$ and a sequence $\{\boldsymbol{v}^k\} \subset \partial\mathcal{P}_{m,n}^{\alpha} \cap \mathcal{B}(\eta) \backslash \mathcal{F}_r$ such that

$$(3.13) \qquad \lim_{k \to \infty} \boldsymbol{v}^k = \lim_{k \to \infty} \boldsymbol{w}^k = \widehat{\boldsymbol{v}} \quad \text{and} \quad \lim_{k \to \infty} \frac{\|\boldsymbol{w}^k - \boldsymbol{v}^k\|^{\frac{1}{2}}}{\|\boldsymbol{w}^k - \boldsymbol{u}^k\|} = 0,$$

where $\boldsymbol{w}^k = P_{\{\boldsymbol{z}\}^{\perp}}(\boldsymbol{v}^k)$, $\boldsymbol{u}^k = P_{\mathcal{F}_r}(\boldsymbol{w}^k)$, and $\boldsymbol{u}^k \neq \boldsymbol{w}^k$.

Define, for notational simplicity, $z_0 := \|\overline{\boldsymbol{z}}\|$ and $v_0^k := \|\overline{\boldsymbol{v}}^k\|$. Then, since $\{\boldsymbol{v}^k\} \subset \partial\mathcal{P}_{m,n}^{\alpha}$ and $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\alpha})^*$ with $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$, we have

$$(3.14) \qquad z_0 = \|\overline{\boldsymbol{z}}\| = \prod_{i=1}^{n} \left(\frac{\widetilde{z}_i}{\alpha_i} \right)^{\alpha_i} > 0 \quad \text{and} \quad v_0^k = \|\overline{\boldsymbol{v}}^k\| = \prod_{i=1}^{n} (\widetilde{v}_i^k)^{\alpha_i} \; \forall k.$$

If it holds that $v_0^k = 0$ infinitely often, by passing to a further subsequence, we may assume that $v_0^k = 0$ for all k. Then we have in view of Lemma 2.2 that

$$\|\boldsymbol{v}^k - \boldsymbol{w}^k\| = \frac{1}{\|\boldsymbol{z}\|} |\langle \widetilde{\boldsymbol{z}}, \widetilde{\boldsymbol{v}}^k \rangle| \overset{(a)}{=} \frac{1}{\|\boldsymbol{z}\|} \sum_{i=1}^{n} \widetilde{z}_i \widetilde{v}_i^k \geq \frac{\min_i \widetilde{z}_i}{\|\boldsymbol{z}\|} \|\widetilde{\boldsymbol{v}}^k\|_1 \geq \frac{\min_i \widetilde{z}_i}{\|\boldsymbol{z}\|} \|\widetilde{\boldsymbol{v}}^k\| \overset{(b)}{=} \frac{\min_i \widetilde{z}_i}{\|\boldsymbol{z}\|} \|\boldsymbol{v}^k\|,$$

where (a) holds because $\widetilde{v}_i^k \geq 0$ and $\widetilde{z}_i > 0$ for all i (see (3.14)), and (b) holds since $\|\overline{\boldsymbol{v}}^k\| = 0$. Since $\|\boldsymbol{w}^k - \boldsymbol{u}^k\| = \mathrm{dist}(\boldsymbol{w}^k, \mathcal{F}_r) \leq \|\boldsymbol{w}^k\| \leq \|\boldsymbol{v}^k\|$ as a consequence of the properties of projections, we conclude from this and the above display that $\|\boldsymbol{v}^k - \boldsymbol{w}^k\| \geq \frac{\min_i \widetilde{z}_i}{\|\boldsymbol{z}\|} \|\boldsymbol{w}^k - \boldsymbol{u}^k\|$, contradicting (3.13).

Thus, by considering a further subsequence if necessary, from now on, we assume

$$(3.15) \qquad v_0^k = \|\overline{\boldsymbol{v}}^k\| = \prod_{i=1}^{n} (\widetilde{v}_i^k)^{\alpha_i} > 0 \; \forall k.$$

Using Lemma 2.2, we see that

$$(3.16) \qquad \begin{aligned} \|\boldsymbol{v}^k - \boldsymbol{w}^k\| &= \frac{1}{\|\boldsymbol{z}\|} |\langle \boldsymbol{z}, \boldsymbol{v}^k \rangle| = \frac{1}{\|\boldsymbol{z}\|} \left| \sum_{i=1}^{m} \overline{z}_i \overline{v}_i^k + \sum_{i=1}^{n} \widetilde{z}_i \widetilde{v}_i^k \right| \\ &= \frac{1}{\|\boldsymbol{z}\|} \left| z_0 v_0^k + \sum_{i=1}^{m} \overline{z}_i \overline{v}_i^k - \sum_{i=1}^{n} (-\widetilde{z}_i) \widetilde{v}_i^k - z_0 v_0^k \right| \\ &= \frac{z_0 v_0^k}{\|\boldsymbol{z}\|} \left| 1 + \langle z_0^{-1}\overline{\boldsymbol{z}}, (v_0^k)^{-1}\overline{\boldsymbol{v}}^k \rangle - \langle z_0^{-1}(-\widetilde{\boldsymbol{z}}), (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k \rangle - 1 \right| \\ &= \frac{z_0}{\|\boldsymbol{z}\|} \left(1 + \langle z_0^{-1}\overline{\boldsymbol{z}}, (v_0^k)^{-1}\overline{\boldsymbol{v}}^k \rangle - \langle z_0^{-1}(-\widetilde{\boldsymbol{z}}), (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k \rangle - 1 \right) v_0^k, \end{aligned}$$

where the last equality holds as $\|z_0^{-1}\overline{z}\| = 1$, $\|(v_0^k)^{-1}\overline{v}^k\| = 1$ and $\langle z_0^{-1}\widetilde{z}, (v_0^k)^{-1}\widetilde{v}^k\rangle \geq 1$, thanks to (3.14), (3.15), and Lemma 3.1 applied with $\zeta = -z_0^{-1}\widetilde{z}$.

Let f be defined as in (3.4). We consider two cases:

(I) $\langle f, v^k\rangle \geq 0$ for all sufficiently large k.

(II) $\langle f, v^k\rangle < 0$ infinitely often.

(I) By passing to a further subsequence, we may assume that $\langle f, v^k\rangle \geq 0$ for all k. In this case, if we define

$$Q = I_{m+n} - \frac{zz^\top}{\|z\|^2} - \frac{ff^\top}{\|f\|^2},$$

where f is as in (3.4), then we see from Lemma 2.2 and (3.14) that

$$
\begin{aligned}
\|u^k - w^k\| = \|Qv^k\| &= v_0^k \left\| Q \begin{bmatrix} (v_0^k)^{-1}\overline{v}^k \\ (v_0^k)^{-1}\widetilde{v}^k \end{bmatrix} \right\| \\
&\overset{\text{(a)}}{=} v_0^k \left\| Q \begin{bmatrix} (v_0^k)^{-1}\overline{v}^k \\ (v_0^k)^{-1}\widetilde{v}^k \end{bmatrix} - Q \underbrace{\begin{bmatrix} -z_0^{-1}\overline{z} \\ \alpha \circ (z_0\widetilde{z}^{-1}) \end{bmatrix}}_{z_0 f} \right\| \\
&\leq v_0^k \left[\|(v_0^k)^{-1}\overline{v}^k + z_0^{-1}\overline{z}\| + \|(v_0^k)^{-1}\widetilde{v}^k - \alpha \circ (z_0\widetilde{z}^{-1})\| \right],
\end{aligned}
$$

(3.17)

where (a) holds because $Qf = 0$ (an identity which is clear from the definitions).

Next, in view of (3.14), we can apply Lemma 3.2 to obtain $C_1 > 0$ and $\epsilon_1 \in (0,1)$ so that (3.3) holds with $\overline{\zeta} = -z_0^{-1}\overline{z}$, i.e.,

$$1 + \langle z_0^{-1}\overline{z}, \omega\rangle \geq C_1\|\omega + z_0^{-1}\overline{z}\|^2$$

whenever $\|\omega + z_0^{-1}\overline{z}\| \leq \epsilon_1$ and $\|\omega\| = 1$. On the other hand, in view of the positivity of $1 + \langle z_0^{-1}\overline{z}, \omega\rangle$ when $\|\omega\| = 1$ and $\omega \neq -z_0^{-1}\overline{z}$ (see Lemma 3.2), we know that

$$C_2 := \inf_{\|\omega\|=1} \{1 + \langle z_0^{-1}\overline{z}, \omega\rangle \mid \|\omega + z_0^{-1}\overline{z}\| \geq \epsilon_1\} \in (0, \infty).$$

This together with the fact $\|z_0^{-1}\overline{z}\| = 1$ (see (3.14)) implies that

$$1 + \langle z_0^{-1}\overline{z}, \omega\rangle \geq C_2 \geq 0.25C_2\|\omega + z_0^{-1}\overline{z}\|^2$$

whenever $\|\omega + z_0^{-1}\overline{z}\| \geq \epsilon_1$ and $\|\omega\| = 1$. We thus have (with $C_3 := \min\{C_1, C_2/4\}$)

(3.18) $\qquad 1 + \langle z_0^{-1}\overline{z}, \omega\rangle \geq C_3\|\omega + z_0^{-1}\overline{z}\|^2 \quad \text{whenever} \quad \|\omega\| = 1.$

In addition, noting (3.14) again, we can apply Lemma 3.1 with $\zeta = -z_0^{-1}\widetilde{z} \in \operatorname{int}\mathbb{R}_-^n$ to obtain $C_4 > 0$ and $\epsilon > 0$ so that (3.2) holds with $\widetilde{\zeta} = \alpha \circ (z_0\widetilde{z}^{-1})$, i.e.,

(3.19) $\qquad -1 + \langle z_0^{-1}\widetilde{z}, \omega\rangle \geq C_4\|\omega - \alpha \circ (z_0\widetilde{z}^{-1})\|^2$

whenever $\|\omega - \alpha \circ (z_0\widetilde{z}^{-1})\| \leq \epsilon$, $\omega \in \operatorname{int}\mathbb{R}_+^n$ and $\prod_{i=1}^n \omega_i^{\alpha_i} = 1$.

Furthermore, consider $h : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ defined by

(3.20) $\qquad h(\omega) = \begin{cases} \dfrac{\langle z_0^{-1}\widetilde{z}, \omega\rangle - 1}{\|\omega - \alpha \circ (z_0\widetilde{z}^{-1})\|} & \text{if } \|\omega - \alpha \circ (z_0\widetilde{z}^{-1})\| \geq \epsilon, \, \omega \in \Upsilon, \\ \infty & \text{otherwise,} \end{cases}$

where $\Upsilon = \{\boldsymbol{\omega} \in \mathbb{R}^n_+ \mid \prod_{i=1}^{n} \omega_i^{\alpha_i} = 1\}$. Then we have

$$
\begin{aligned}
\liminf_{\|\boldsymbol{\omega}\| \to \infty} h(\boldsymbol{\omega}) &= \liminf_{\|\boldsymbol{\omega}\| \to \infty, \boldsymbol{\omega} \in \Upsilon} \frac{\langle z_0^{-1}\widetilde{\boldsymbol{z}}, \boldsymbol{\omega} \rangle - 1}{\|\boldsymbol{\omega} - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\|} \\
&\geq \liminf_{\|\boldsymbol{\omega}\| \to \infty, \boldsymbol{\omega} \in \mathbb{R}^n_+} \frac{\langle z_0^{-1}\widetilde{\boldsymbol{z}}, \boldsymbol{\omega} \rangle - 1}{\|\boldsymbol{\omega} - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\|} \overset{\text{(a)}}{\geq} \inf_{\|\boldsymbol{\lambda}\|=1, \boldsymbol{\lambda} \in \mathbb{R}^n_+} \langle z_0^{-1}\widetilde{\boldsymbol{z}}, \boldsymbol{\lambda} \rangle \\
&\overset{\text{(b)}}{\geq} \min_{1 \leq i \leq n} z_0^{-1}\widetilde{z}_i > 0.
\end{aligned}
$$

Here (a) may be verified by multiplying both numerator and denominator of the left side by $1/\|\boldsymbol{\omega}\|$; (b) holds because $z_0^{-1}\widetilde{z}_i > 0$ for all i (see (3.14)). Since h in (3.20) is also lower semicontinuous on any compact set and *is always positive*,[5] it must then hold that $C_5 := \inf h > 0$. In particular, this means that

$$
(3.21) \qquad -1 + \langle z_0^{-1}\widetilde{\boldsymbol{z}}, \boldsymbol{\omega} \rangle \geq C_5 \|\boldsymbol{\omega} - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\|
$$

whenever $\|\boldsymbol{\omega} - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\| \geq \epsilon$, $\boldsymbol{\omega} \in \text{int}\,\mathbb{R}^n_+$ and $\prod_{i=1}^{n} \omega_i^{\alpha_i} = 1$.

By passing to suitable subsequences, we will end up with one of the following two cases:

Case 1: $\|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\| \leq \epsilon$ for all k. Then we have from (3.18) (with $\boldsymbol{\omega} = (v_0^k)^{-1}\overline{\boldsymbol{v}}^k$) and (3.19) (with $\boldsymbol{\omega} = (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k$) that for these k

$$
\begin{aligned}
&1 + \langle z_0^{-1}\overline{\boldsymbol{z}}, (v_0^k)^{-1}\overline{\boldsymbol{v}}^k \rangle - \langle z_0^{-1}(-\widetilde{\boldsymbol{z}}), (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k \rangle - 1 \\
&\qquad \geq \min\{C_3, C_4\}(\|(v_0^k)^{-1}\overline{\boldsymbol{v}}^k + z_0^{-1}\overline{\boldsymbol{z}}\|^2 + \|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\|^2).
\end{aligned}
$$

Combining this with (3.16) and (3.17), we see further that

$$
\begin{aligned}
&\|\boldsymbol{v}^k - \boldsymbol{w}^k\| \\
&\geq \frac{z_0}{\|\boldsymbol{z}\|} \min\{C_3, C_4\}(\|(v_0^k)^{-1}\overline{\boldsymbol{v}}^k + z_0^{-1}\overline{\boldsymbol{z}}\|^2 + \|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\|^2)v_0^k \\
&\geq \frac{z_0}{2\|\boldsymbol{z}\|} \min\{C_3, C_4\}(\|(v_0^k)^{-1}\overline{\boldsymbol{v}}^k + z_0^{-1}\overline{\boldsymbol{z}}\| + \|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\|)^2 v_0^k \\
&\geq \frac{z_0 \min\{C_3, C_4\}}{2\|\boldsymbol{z}\|v_0^k} \|\boldsymbol{u}^k - \boldsymbol{w}^k\|^2 \geq \frac{z_0 \min\{C_3, C_4\}}{2\|\boldsymbol{z}\|\eta} \|\boldsymbol{u}^k - \boldsymbol{w}^k\|^2,
\end{aligned}
$$

where the last inequality holds because $\boldsymbol{v}^k \in \mathcal{B}(\eta)$. The above display contradicts (3.13), and hence Case 1 cannot happen.

Case 2: $\|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\| \geq \epsilon$ for all k. Then we have from (3.18) and (3.21) that for these k

$$
\begin{aligned}
&1 + \langle z_0^{-1}\overline{\boldsymbol{z}}, (v_0^k)^{-1}\overline{\boldsymbol{v}}^k \rangle - \langle z_0^{-1}(-\widetilde{\boldsymbol{z}}), (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k \rangle - 1 \\
&\qquad \geq \min\{C_3, C_5\}(\|(v_0^k)^{-1}\overline{\boldsymbol{v}}^k + z_0^{-1}\overline{\boldsymbol{z}}\|^2 + \|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\|).
\end{aligned}
$$

Using this together with (3.16), we deduce that for all large k,

$$
\begin{aligned}
&\|\boldsymbol{v}^k - \boldsymbol{w}^k\| \\
&\geq \frac{z_0 \min\{C_3, C_5\}}{\|\boldsymbol{z}\|}(\|(v_0^k)^{-1}\overline{\boldsymbol{v}}^k + z_0^{-1}\overline{\boldsymbol{z}}\|^2 + \|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha} \circ (z_0 \widetilde{\boldsymbol{z}}^{-1})\|)v_0^k.
\end{aligned}
$$

___

[5] The positivity can be seen by applying Lemma 3.1 with $\boldsymbol{\zeta} = -z_0^{-1}\widetilde{\boldsymbol{z}}$.

This implies that

$$\|(v_0^k)^{-1}\overline{\boldsymbol{v}}^k + z_0^{-1}\overline{\boldsymbol{z}}\| \leq M_1\sqrt{(v_0^k)^{-1}\|\boldsymbol{v}^k - \boldsymbol{w}^k\|},$$

(3.22)

$$\|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha}\circ(z_0\widetilde{\boldsymbol{z}}^{-1})\| \leq M_1(v_0^k)^{-1}\|\boldsymbol{v}^k - \boldsymbol{w}^k\|,$$

where $M_1 := \max\{(\frac{z_0}{\|\boldsymbol{z}\|}\min\{C_3, C_5\})^{-1}, (\frac{z_0}{\|\boldsymbol{z}\|}\min\{C_3, C_5\})^{-1/2}\}$. Using (3.22) together with (3.17), we obtain that

$$\|\boldsymbol{u}^k - \boldsymbol{w}^k\| \leq M_1 v_0^k \left[\sqrt{(v_0^k)^{-1}\|\boldsymbol{v}^k - \boldsymbol{w}^k\|} + (v_0^k)^{-1}\|\boldsymbol{v}^k - \boldsymbol{w}^k\|\right]$$

(3.23)

$$\overset{(a)}{\leq} M_1\sqrt{\eta}\sqrt{\|\boldsymbol{v}^k - \boldsymbol{w}^k\|} + M_1\|\boldsymbol{v}^k - \boldsymbol{w}^k\|$$

$$\overset{(b)}{\leq} 3M_1\sqrt{\eta}\sqrt{\|\boldsymbol{v}^k - \boldsymbol{w}^k\|},$$

where (a) holds since $\boldsymbol{v}^k \in \mathcal{B}(\eta)$ wherefore $v_0^k \leq \eta$, and (b) holds because $\|\boldsymbol{w}^k\| \leq \|\boldsymbol{v}^k\| \leq \eta$ (because the projection onto $\mathcal{K}$ is nonexpansive and $\boldsymbol{0} \in \mathcal{K}$), wherefore

$$\|\boldsymbol{w}^k - \boldsymbol{v}^k\| = \sqrt{\|\boldsymbol{w}^k - \boldsymbol{v}^k\|}\sqrt{\|\boldsymbol{w}^k - \boldsymbol{v}^k\|} \leq 2\sqrt{\eta}\sqrt{\|\boldsymbol{w}^k - \boldsymbol{v}^k\|}.$$

Altogether, (3.23) contradicts (3.13), and hence Case 2 cannot happen. Summarizing the above discussions, we see that Case (I) cannot happen.

(II) By passing to a further subsequence, we may assume that $\langle \boldsymbol{f}, \boldsymbol{v}^k \rangle < 0$ for all k. This together with the definition of $\boldsymbol{f}$ gives

$$\frac{v_0^k}{z_0}[\langle -z_0^{-1}\overline{\boldsymbol{z}}, (v_0^k)^{-1}\overline{\boldsymbol{v}}^k\rangle + \langle \boldsymbol{\alpha}\circ(z_0\widetilde{\boldsymbol{z}}^{-1}), (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k\rangle] < 0.$$

Since $\frac{v_0^k}{z_0} > 0$, we deduce that $\langle -z_0^{-1}\overline{\boldsymbol{z}}, (v_0^k)^{-1}\overline{\boldsymbol{v}}^k\rangle + \langle \boldsymbol{\alpha}\circ(z_0\widetilde{\boldsymbol{z}}^{-1}), (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k\rangle < 0$ for all k. Then it must hold that

$$\lim_{k\to\infty}\|(v_0^k)^{-1}\overline{\boldsymbol{v}}^k + z_0^{-1}\overline{\boldsymbol{z}}\| + \|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha}\circ(z_0\widetilde{\boldsymbol{z}}^{-1})\| \neq 0;$$

otherwise, we have $(v_0^k)^{-1}\overline{\boldsymbol{v}}^k \to -z_0^{-1}\overline{\boldsymbol{z}}$ and $(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k \to \boldsymbol{\alpha}\circ(z_0\widetilde{\boldsymbol{z}}^{-1})$, which further gives $\langle -z_0^{-1}\overline{\boldsymbol{z}}, (v_0^k)^{-1}\overline{\boldsymbol{v}}^k\rangle + \langle \boldsymbol{\alpha}\circ(z_0\widetilde{\boldsymbol{z}}^{-1}), (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k\rangle \to \|z_0^{-1}\overline{\boldsymbol{z}}\|^2 + \|\boldsymbol{\alpha}\circ(z_0\widetilde{\boldsymbol{z}}^{-1})\|^2 = \|z_0\boldsymbol{f}\|^2 > 0$, a contradiction.

Consequently, there exists $\epsilon > 0$ such that for all sufficiently large k,

(3.24)

$$\|(v_0^k)^{-1}\overline{\boldsymbol{v}}^k + z_0^{-1}\overline{\boldsymbol{z}}\| + \|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k - \boldsymbol{\alpha}\circ(z_0\widetilde{\boldsymbol{z}}^{-1})\| \geq \epsilon.$$

Consider the function $G: \mathbb{R}^{m+n} \to \mathbb{R} \cup \{\infty\}$ defined by

$$G(\boldsymbol{\xi}, \boldsymbol{\omega}) := \begin{cases} \dfrac{|\langle z_0^{-1}\overline{\boldsymbol{z}}, \boldsymbol{\xi}\rangle + \langle z_0^{-1}\widetilde{\boldsymbol{z}}, \boldsymbol{\omega}\rangle|}{\sqrt{1 + \|\boldsymbol{\omega}\|^2}} & \text{if } (\boldsymbol{\xi}, \boldsymbol{\omega}) \in \Xi, \|\boldsymbol{\xi}\| = 1, \text{ and } \boldsymbol{\omega} \in \Upsilon, \\ \infty & \text{otherwise,} \end{cases}$$

where $\Upsilon = \{\boldsymbol{\omega} \in \mathbb{R}_+^n \mid \prod_{i=1}^n \omega_i^{\alpha_i} = 1\}$ and $\Xi = \{(\boldsymbol{\xi}, \boldsymbol{\omega}) \mid \|\boldsymbol{\xi} + z_0^{-1}\overline{\boldsymbol{z}}\| + \|\boldsymbol{\omega} - \boldsymbol{\alpha}\circ(z_0\widetilde{\boldsymbol{z}}^{-1})\| \geq \epsilon\}$. Since $\langle z_0^{-1}\overline{\boldsymbol{z}}, \boldsymbol{\xi}\rangle + \langle z_0^{-1}\widetilde{\boldsymbol{z}}, \boldsymbol{\omega}\rangle = 1 + \langle z_0^{-1}\overline{\boldsymbol{z}}, \boldsymbol{\xi}\rangle - \langle z_0^{-1}(-\widetilde{\boldsymbol{z}}), \boldsymbol{\omega}\rangle - 1$, we see from (3.14),

(3.24), and Lemmas 3.2 and 3.1 that G is never zero. Moreover, it is clearly lower semicontinuous on any compact set, and

$$\liminf_{\|(\boldsymbol{\xi},\boldsymbol{\omega})\|\to\infty} G(\boldsymbol{\xi},\boldsymbol{\omega}) = \liminf_{\|\boldsymbol{\omega}\|\to\infty,\boldsymbol{\omega}\in\Upsilon} \frac{|\langle z_0^{-1}\widetilde{\boldsymbol{z}},\boldsymbol{\omega}\rangle|}{\sqrt{1+\|\boldsymbol{\omega}\|^2}} \overset{(a)}{\geq} \inf_{\|\boldsymbol{\lambda}\|=1,\boldsymbol{\lambda}\in\mathbb{R}_+^n} |\langle z_0^{-1}\widetilde{\boldsymbol{z}},\boldsymbol{\lambda}\rangle| \overset{(b)}{\geq} \min_i |z_0^{-1}\widetilde{z}_i| > 0,$$

where (a) may be verified by multiplying numerator and denominator by $1/\|\boldsymbol{\omega}\|$ and (b) holds since $z_0^{-1}\widetilde{z}_i > 0$ for all i. Thus, $C_6 := \inf G > 0$, and we have for all large k,

$$\begin{aligned}
\frac{\|\boldsymbol{v}^k - \boldsymbol{w}^k\|}{\|\boldsymbol{u}^k - \boldsymbol{w}^k\|} &\overset{(a)}{=} \frac{\|\boldsymbol{v}^k - \boldsymbol{w}^k\|}{\|\boldsymbol{w}^k\|} \overset{(b)}{\geq} \frac{\|\boldsymbol{v}^k - \boldsymbol{w}^k\|}{\|\boldsymbol{v}^k\|} \\
&\overset{(c)}{=} \frac{z_0}{\|\boldsymbol{z}\|} \frac{|\langle z_0^{-1}\overline{\boldsymbol{z}}, (v_0^k)^{-1}\overline{\boldsymbol{v}}^k\rangle + \langle z_0^{-1}\widetilde{\boldsymbol{z}}, (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k\rangle| v_0^k}{\sqrt{(v_0^k)^2 + \|\widetilde{\boldsymbol{v}}^k\|^2}} \\
&= \frac{z_0}{\|\boldsymbol{z}\|} \frac{|\langle z_0^{-1}\overline{\boldsymbol{z}}, (v_0^k)^{-1}\overline{\boldsymbol{v}}^k\rangle + \langle z_0^{-1}\widetilde{\boldsymbol{z}}, (v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k\rangle|}{\sqrt{1 + \|(v_0^k)^{-1}\widetilde{\boldsymbol{v}}^k\|^2}} \overset{(d)}{\geq} \frac{C_6 z_0}{\|\boldsymbol{z}\|},
\end{aligned}$$

where (a) follows from Lemma 2.2, which states that $\boldsymbol{u}^k = \boldsymbol{0}$ in Case (II), (b) holds because the projection onto the cone is nonexpansive and $\boldsymbol{0}$ is in the cone, (c) follows from (3.16), and (d) follows from (3.24), (3.15), and the definitions of G and C_6. The above display contradicts (3.13). Thus, Case (II) also cannot happen.

Summarizing the above, we conclude that (3.13) cannot happen. Thus, in view of [19, Lemma 3.12], we must indeed have $\gamma_{\boldsymbol{z},\eta} \in (0,\infty]$ and that the desired error bound follows from [19, Theorem 3.10]. □

Remark 3.6 (optimality of the error bound in Theorem 3.5). Let $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^*\backslash\{\boldsymbol{0}\}$ with $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$ and let $\mathcal{F}_{\mathrm{r}} := \{\boldsymbol{z}\}^\perp \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. Then necessarily $\widetilde{z}_i > 0$ for all i. Moreover, we also know from the definition that $\alpha_i > 0$ for all i. Now, consider the continuous function $\boldsymbol{q} : (0,\alpha_1) \to \{\boldsymbol{z}\}^\perp$ defined by $\epsilon \mapsto \boldsymbol{q}_\epsilon := (\overline{\boldsymbol{q}}_\epsilon, \widetilde{\boldsymbol{q}}_\epsilon)$ where

$$\overline{\boldsymbol{q}}_\epsilon = -\overline{\boldsymbol{z}}/\|\overline{\boldsymbol{z}}\|^2, \quad (\widetilde{\boldsymbol{q}}_\epsilon)_1 = (\alpha_1 - \epsilon)\widetilde{z}_1^{-1}, \quad (\widetilde{\boldsymbol{q}}_\epsilon)_2 = (\alpha_2 + \epsilon)\widetilde{z}_2^{-1}, \text{ and } (\widetilde{\boldsymbol{q}}_\epsilon)_i = \alpha_i\widetilde{z}_i^{-1} \ \ \forall i \geq 3.$$

Notice that $\boldsymbol{q}_\epsilon$ only differs from the $\boldsymbol{f}$ in (3.4) in two entries. One can check that $\langle \boldsymbol{z}, \boldsymbol{q}_\epsilon\rangle = 0$ and $\boldsymbol{q}_\epsilon \to \boldsymbol{f} \in \mathcal{F}_{\mathrm{r}}\backslash\{\boldsymbol{0}\}$ as $\epsilon \downarrow 0$. Moreover, we have

$$\begin{aligned}
\prod_{i=1}^n (\widetilde{\boldsymbol{q}}_\epsilon)_i^{\alpha_i} &= (\alpha_1 - \epsilon)^{\alpha_1}(\alpha_2 + \epsilon)^{\alpha_2}\widetilde{z}_1^{-\alpha_1}\widetilde{z}_2^{-\alpha_2} \prod_{i=3}^n \left(\frac{\alpha_i}{\widetilde{z}_i}\right)^{\alpha_i} \\
&= \left(1 - \frac{\epsilon}{\alpha_1}\right)^{\alpha_1}\left(1 + \frac{\epsilon}{\alpha_2}\right)^{\alpha_2} \prod_{i=1}^n \left(\frac{\alpha_i}{\widetilde{z}_i}\right)^{\alpha_i} \overset{(a)}{=} \left(1 - \frac{\epsilon}{\alpha_1}\right)^{\alpha_1}\left(1 + \frac{\epsilon}{\alpha_2}\right)^{\alpha_2} \|\overline{\boldsymbol{z}}\|^{-1} \\
&= \left(1 - \frac{\epsilon}{\alpha_1}\right)^{\alpha_1}\left(1 + \frac{\epsilon}{\alpha_2}\right)^{\alpha_2} \|\overline{\boldsymbol{q}}_\epsilon\|,
\end{aligned}$$

where (a) holds because $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^*\backslash\{\boldsymbol{0}\}$ with $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$. In view of this, if we define a continuous function $\boldsymbol{p} : (0,\alpha_1) \to \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ by $\epsilon \mapsto \boldsymbol{p}_\epsilon := (\overline{\boldsymbol{p}}_\epsilon, \widetilde{\boldsymbol{p}}_\epsilon)$ where

$$\overline{\boldsymbol{p}}_\epsilon := -\left(1 - \frac{\epsilon}{\alpha_1}\right)^{\alpha_1}\left(1 + \frac{\epsilon}{\alpha_2}\right)^{\alpha_2} \frac{\overline{\boldsymbol{z}}}{\|\overline{\boldsymbol{z}}\|^2} \quad \text{and} \quad \widetilde{\boldsymbol{p}}_\epsilon := \widetilde{\boldsymbol{q}}_\epsilon,$$

then it is clear that $\boldsymbol{p}_\epsilon \in \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, and we can compute that

$$
\text{(3.25)} \quad
\begin{aligned}
\text{dist}(\boldsymbol{q}_\epsilon, \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) &\leq \|\boldsymbol{q}_\epsilon - \boldsymbol{p}_\epsilon\| = \frac{1}{\|\overline{\boldsymbol{z}}\|} \left| \left(1 - \frac{\epsilon}{\alpha_1}\right)^{\alpha_1} \left(1 + \frac{\epsilon}{\alpha_2}\right)^{\alpha_2} - 1 \right| \\
&= \frac{1}{\|\overline{\boldsymbol{z}}\|} \left| (1 - \epsilon + O(\epsilon^2))(1 + \epsilon + O(\epsilon^2)) - 1 \right| = O(\epsilon^2).
\end{aligned}
$$

Next, we estimate $\text{dist}(\boldsymbol{q}_\epsilon, \mathcal{F}_\mathrm{r})$. Notice that $\langle \boldsymbol{q}_\epsilon, \boldsymbol{f} \rangle > 0$ for all sufficiently small ϵ because $\boldsymbol{q}_\epsilon \to \boldsymbol{f}$. Hence, using the definition of $\mathcal{F}_\mathrm{r}$ and Lemma 2.2, we see that

$$
\text{dist}(\boldsymbol{q}_\epsilon, \mathcal{F}_\mathrm{r})^2 = \left\| \boldsymbol{q}_\epsilon - \frac{\langle \boldsymbol{q}_\epsilon, \boldsymbol{f} \rangle}{\|\boldsymbol{f}\|^2} \boldsymbol{f} \right\|^2 = \|\boldsymbol{q}_\epsilon\|^2 - \frac{(\langle \boldsymbol{q}_\epsilon, \boldsymbol{f} \rangle)^2}{\|\boldsymbol{f}\|^2}.
$$

A direct computation then shows that

$$
\begin{aligned}
\|\boldsymbol{q}_\epsilon\|^2 &= \frac{1}{\|\overline{\boldsymbol{z}}\|^2} + (\alpha_1 - \epsilon)^2 \widetilde{z}_1^{-2} + (\alpha_2 + \epsilon)^2 \widetilde{z}_2^{-2} + \sum_{i=3}^{n} \alpha_i^2 \widetilde{z}_i^{-2} \\
&= \frac{1}{\|\overline{\boldsymbol{z}}\|^2} + \sum_{i=1}^{n} \alpha_i^2 \widetilde{z}_i^{-2} + 2\epsilon(\alpha_2 \widetilde{z}_2^{-2} - \alpha_1 \widetilde{z}_1^{-2}) + \epsilon^2(\widetilde{z}_1^{-2} + \widetilde{z}_2^{-2}) \\
&= \|\boldsymbol{f}\|^2 + 2\epsilon(\alpha_2 \widetilde{z}_2^{-2} - \alpha_1 \widetilde{z}_1^{-2}) + \epsilon^2(\widetilde{z}_1^{-2} + \widetilde{z}_2^{-2}),
\end{aligned}
$$

where the last equality follows from the definition of $\boldsymbol{f}$ in (3.4). Furthermore,

$$
\begin{aligned}
(\langle \boldsymbol{q}_\epsilon, \boldsymbol{f} \rangle)^2 &= \left(\frac{1}{\|\overline{\boldsymbol{z}}\|^2} + \alpha_1(\alpha_1 - \epsilon)\widetilde{z}_1^{-2} + \alpha_2(\alpha_2 + \epsilon)\widetilde{z}_2^{-2} + \sum_{i=3}^{n} \alpha_i^2 \widetilde{z}_i^{-2} \right)^2 \\
&= \left[\|\boldsymbol{f}\|^2 + \epsilon(\alpha_2 \widetilde{z}_2^{-2} - \alpha_1 \widetilde{z}_1^{-2}) \right]^2 \\
&= \|\boldsymbol{f}\|^4 + 2\epsilon \|\boldsymbol{f}\|^2 (\alpha_2 \widetilde{z}_2^{-2} - \alpha_1 \widetilde{z}_1^{-2}) + \epsilon^2 (\alpha_2 \widetilde{z}_2^{-2} - \alpha_1 \widetilde{z}_1^{-2})^2.
\end{aligned}
$$

Combining the above three identities, we deduce further that

$$
\text{(3.26)} \quad
\begin{aligned}
\text{dist}(\boldsymbol{q}_\epsilon, \mathcal{F}_\mathrm{r})^2 &= \epsilon^2 \left(\widetilde{z}_1^{-2} + \widetilde{z}_2^{-2} - \frac{(\alpha_2 \widetilde{z}_2^{-2} - \alpha_1 \widetilde{z}_1^{-2})^2}{\|\boldsymbol{f}\|^2} \right) \\
&\geq \epsilon^2 \left(\widetilde{z}_1^{-2} + \widetilde{z}_2^{-2} - \frac{(\alpha_2 \widetilde{z}_2^{-2} - \alpha_1 \widetilde{z}_1^{-2})^2}{\alpha_2^2 \widetilde{z}_2^{-2} + \alpha_1^2 \widetilde{z}_1^{-2}} \right),
\end{aligned}
$$

where the inequality follows from the definition of $\boldsymbol{f}$. Now, notice that in (3.26), the scalar term is strictly greater than zero, because

$$
(\alpha_2 \widetilde{z}_2^{-2} - \alpha_1 \widetilde{z}_1^{-2})^2 < (\alpha_2 \widetilde{z}_2^{-2} + \alpha_1 \widetilde{z}_1^{-2})^2 \leq (\widetilde{z}_1^{-2} + \widetilde{z}_2^{-2})(\alpha_2^2 \widetilde{z}_2^{-2} + \alpha_1^2 \widetilde{z}_1^{-2}),
$$

where the strict inequality holds because $\alpha_i \widetilde{z}_i^{-2} > 0$ for $i = 1, 2$, and the last inequality follows from the Cauchy–Schwarz inequality. This together with (3.26) shows that $\text{dist}(\boldsymbol{q}_\epsilon, \mathcal{F}_\mathrm{r}) = \Omega(\epsilon)$. Combining this with (3.25), we obtain $\limsup_{\epsilon \downarrow 0} \frac{\text{dist}(\boldsymbol{q}_\epsilon, \mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^{\frac{1}{2}}}{\text{dist}(\boldsymbol{q}_\epsilon, \mathcal{F}_\mathrm{r})} < \infty$. Thus $|\cdot|^{\frac{1}{2}}$ satisfies the asymptotic optimality criterion (cf. [20, Definition 3.1]) for $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ and $\boldsymbol{z}$, which implies that the error bound is optimal in the sense of [20, Theorem 3.2(b)].

We now look at the faces that are exposed by $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* \backslash \{\boldsymbol{0}\}$ with $\overline{\boldsymbol{z}} = \boldsymbol{0}$.

THEOREM 3.7. *Let* $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* \backslash \{\boldsymbol{0}\}$ *with* $\overline{\boldsymbol{z}} = \boldsymbol{0}$ *and let* $\mathcal{F}_{\boldsymbol{z}} := \{\boldsymbol{z}\}^{\perp} \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. *Let* $\mathcal{I} := \{i \,|\, \widetilde{z}_i > 0\},$[6] $\eta > 0$ *and define* $\beta := \sum_{i \in \mathcal{I}} \alpha_i$ *and*

$$(3.27) \qquad \gamma_{\boldsymbol{z},\eta} := \inf_{\boldsymbol{v}} \left\{ \frac{\|\boldsymbol{v} - \boldsymbol{w}\|^{\beta}}{\|\boldsymbol{u} - \boldsymbol{w}\|} \;\middle|\; \begin{array}{c} \boldsymbol{v} \in \partial\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} \cap \mathcal{B}(\eta) \backslash \mathcal{F}_{\boldsymbol{z}}, \, \boldsymbol{w} = P_{\{\boldsymbol{z}\}^{\perp}}(\boldsymbol{v}), \\ \boldsymbol{u} = P_{\mathcal{F}_{\boldsymbol{z}}}(\boldsymbol{w}), \, \boldsymbol{u} \neq \boldsymbol{w} \end{array} \right\}.$$

Then it holds that $\gamma_{\boldsymbol{z},\eta} \in (0, \infty]$ *and that*

$$\mathrm{dist}(\boldsymbol{q}, \mathcal{F}_{\boldsymbol{z}}) \leq \max\{2\eta^{1-\beta}, 2\gamma_{\boldsymbol{z},\eta}^{-1}\} \cdot \mathrm{dist}(\boldsymbol{q}, \mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^{\beta} \quad \text{whenever } \boldsymbol{q} \in \{\boldsymbol{z}\}^{\perp} \cap \mathcal{B}(\eta).$$

Proof. In view of [19, Theorem 3.10], we need only show that $\gamma_{\boldsymbol{z},\eta} > 0$. To that end, let $\boldsymbol{v} \in \partial\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} \cap \mathcal{B}(\eta) \backslash \mathcal{F}_{\boldsymbol{z}}$, $\boldsymbol{w} = P_{\{\boldsymbol{z}\}^{\perp}}(\boldsymbol{v})$, $\boldsymbol{u} = P_{\mathcal{F}_{\boldsymbol{z}}}(\boldsymbol{w})$, and $\boldsymbol{u} \neq \boldsymbol{w}$. Then a direct computation shows that

$$(3.28) \quad \|\boldsymbol{w} - \boldsymbol{v}\| = \frac{1}{\|\boldsymbol{z}\|}|\langle \boldsymbol{z}, \boldsymbol{v}\rangle| \overset{\text{(a)}}{=} \frac{1}{\|\boldsymbol{z}\|}\sum_{i \in \mathcal{I}} \widetilde{z}_i \widetilde{v}_i \overset{\text{(b)}}{\geq} \frac{\min_{i \in \mathcal{I}} \widetilde{z}_i}{\|\boldsymbol{z}\|}\sum_{i \in \mathcal{I}} \widetilde{v}_i \overset{\text{(c)}}{\geq} \frac{\min_{i \in \mathcal{I}} \widetilde{z}_i}{\|\boldsymbol{z}\|}\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|,$$

where (a), (b), and (c) hold because $\widetilde{v}_i \geq 0$ and $\widetilde{z}_i > 0$ for all $i \in \mathcal{I}$, with $\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\| := \sqrt{\sum_{i \in \mathcal{I}} \widetilde{v}_i^2}$ (note that $\mathcal{I} \neq \emptyset$, thanks to $\overline{\boldsymbol{z}} = \boldsymbol{0}$ and $\boldsymbol{z} \neq \boldsymbol{0}$). Next, notice that $\boldsymbol{w} = \boldsymbol{v} - \frac{\langle \boldsymbol{z}, \boldsymbol{v}\rangle}{\|\boldsymbol{z}\|^2}\boldsymbol{z}$. Using this and the definitions of $\boldsymbol{z}$ and $\mathcal{I}$, we deduce that

$$(3.29) \qquad \overline{\boldsymbol{w}} = \overline{\boldsymbol{v}}, \;\; \widetilde{w}_i = \widetilde{v}_i - \frac{\widetilde{z}_i}{\|\boldsymbol{z}\|^2}\left(\sum_{j \in \mathcal{I}} \widetilde{z}_j \widetilde{v}_j\right) \;\; \forall i \in \mathcal{I} \quad \text{and} \quad \widetilde{w}_i = \widetilde{v}_i \geq 0 \;\; \forall i \notin \mathcal{I}.$$

In view of this and the definition of $\mathcal{F}_{\boldsymbol{z}}$ in (3.5), we see that $\widetilde{u}_i = \widetilde{w}_i$ whenever $i \notin \mathcal{I}$, and hence

$$(3.30) \qquad \|\boldsymbol{w} - \boldsymbol{u}\| = \sqrt{\|\overline{\boldsymbol{w}}\|^2 + \sum_{i \in \mathcal{I}} \widetilde{w}_i^2} \leq \sqrt{\|\overline{\boldsymbol{v}}\|^2 + n(1 + \sqrt{n})^2\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|^2},$$

where the inequality follows from (3.29) and the fact that for each $i \in \mathcal{I}$,

$$|\widetilde{w}_i| = \left|\widetilde{v}_i - \frac{\widetilde{z}_i}{\|\boldsymbol{z}\|^2}\left(\sum_{j \in \mathcal{I}} \widetilde{z}_j \widetilde{v}_j\right)\right| \leq \left(1 + \frac{|\widetilde{z}_i|}{\|\boldsymbol{z}\|^2}\sum_{j \in \mathcal{I}} |\widetilde{z}_j|\right)\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|$$
$$\leq \left(1 + \frac{\sqrt{n}|\widetilde{z}_i|}{\|\boldsymbol{z}\|}\right)\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\| \leq (1 + \sqrt{n})\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|.$$

Next, note that we have

$$(3.31) \qquad \|\overline{\boldsymbol{v}}\| = \prod_{i=1}^{n}(\widetilde{v}_i)^{\alpha_i} = \prod_{i \notin \mathcal{I}} \widetilde{v}_i^{\alpha_i} \cdot \prod_{i \in \mathcal{I}} \widetilde{v}_i^{\alpha_i} \leq \prod_{i \notin \mathcal{I}} \eta^{\alpha_i} \cdot \prod_{i \in \mathcal{I}} \|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|^{\alpha_i} = \eta^{1-\beta}\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|^{\beta},$$

where the inequality holds because $\boldsymbol{v} \in \mathcal{B}(\eta)$. Combining (3.28), (3.30), and (3.31), we deduce

$$\|\boldsymbol{w} - \boldsymbol{u}\| \leq \sqrt{\|\overline{\boldsymbol{v}}\|^2 + n(1 + \sqrt{n})^2\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|^2} \leq \|\overline{\boldsymbol{v}}\| + (n + \sqrt{n})\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|$$
$$\leq \eta^{1-\beta}\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|^{\beta} + (n + \sqrt{n})\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\| = (\eta^{1-\beta} + (n + \sqrt{n})\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|^{1-\beta})\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|^{\beta}$$
$$\overset{\text{(a)}}{\leq} \eta^{1-\beta}(n + 1 + \sqrt{n})\|\widetilde{\boldsymbol{v}}_{\mathcal{I}}\|^{\beta} \overset{\text{(b)}}{\leq} \frac{\eta^{1-\beta}(n + 1 + \sqrt{n})\|\boldsymbol{z}\|^{\beta}}{(\min_{i \in \mathcal{I}} \widetilde{z}_i)^{\beta}}\|\boldsymbol{w} - \boldsymbol{v}\|^{\beta}.$$

---

[6]Since $\overline{\boldsymbol{z}} = \boldsymbol{0}$ and $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* \backslash \{\boldsymbol{0}\}$, we must have $\emptyset \neq \mathcal{I} \subsetneq \{1, 2, \ldots, n\}$.

Here (a) holds since $\boldsymbol{v} \in \mathcal{B}(\eta)$ and $\beta \in (0,1)$; (b) is true because of (3.28). Thus, $\gamma_{\boldsymbol{z},\eta} \geq \frac{(\min_{i \in \mathcal{I}} \widetilde{z}_i)^\beta}{\eta^{1-\beta}(n+1+\sqrt{n})\|\boldsymbol{z}\|^\beta} > 0$, and the desired error bound follows from [19, Theorem 3.10]. □

Remark 3.8 (optimality of the error bound in Theorem 3.7). Let $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* \backslash \{\boldsymbol{0}\}$ with $\overline{\boldsymbol{z}} = \boldsymbol{0}$ and let $\mathcal{F}_{\boldsymbol{z}} := \{\boldsymbol{z}\}^\perp \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. Let $\mathcal{I} := \{i \,|\, \widetilde{z}_i > 0\} \neq \emptyset$ and define

$$\beta := \sum_{i \in \mathcal{I}} \alpha_i \in (0,1).$$

Fix any $\boldsymbol{u} \in \mathbb{R}^m$ with $\|\boldsymbol{u}\| = 1$ and define the continuous function $\boldsymbol{q} : (0,1) \to \{\boldsymbol{z}\}^\perp$ by $\epsilon \mapsto \boldsymbol{q}_\epsilon := (\bar{\boldsymbol{q}}_\epsilon, \widetilde{\boldsymbol{q}}_\epsilon)$, where

$$\bar{\boldsymbol{q}}_\epsilon = \epsilon^\beta \boldsymbol{u}, \ (\widetilde{\boldsymbol{q}}_\epsilon)_i = 0 \ \forall i \in \mathcal{I}, \text{ and } (\widetilde{\boldsymbol{q}}_\epsilon)_i = 1 \ \forall i \notin \mathcal{I}.$$

It is clear that for all ϵ, $\langle \boldsymbol{z}, \boldsymbol{q}_\epsilon \rangle = 0$ and $\operatorname{dist}(\boldsymbol{q}_\epsilon, \mathcal{F}_{\boldsymbol{z}}) \to 0$ as $\epsilon \downarrow 0$. Now, define the function $\boldsymbol{p} : (0,1) \to \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ by $\epsilon \mapsto \boldsymbol{p}_\epsilon := (\bar{\boldsymbol{p}}_\epsilon, \widetilde{\boldsymbol{p}}_\epsilon)$, where

$$\bar{\boldsymbol{p}}_\epsilon = \epsilon^\beta \boldsymbol{u}, \ (\widetilde{\boldsymbol{p}}_\epsilon)_i = \epsilon \ \forall i \in \mathcal{I}, \text{ and } (\widetilde{\boldsymbol{p}}_\epsilon)_i = 1 \ \forall i \notin \mathcal{I}.$$

Clearly $\boldsymbol{p}_\epsilon$ lies in $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, and we have that $\operatorname{dist}(\boldsymbol{q}_\epsilon, \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) \leq \|\boldsymbol{q}_\epsilon - \boldsymbol{p}_\epsilon\| \leq |\mathcal{I}| \cdot \epsilon$. On the other hand, we have in view of (3.5) that $\operatorname{dist}(\boldsymbol{q}_\epsilon, \mathcal{F}_{\boldsymbol{z}}) = \epsilon^\beta > 0$. Hence, $\limsup_{\epsilon \downarrow 0} \frac{\operatorname{dist}(\boldsymbol{q}_\epsilon, \mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^\beta}{\operatorname{dist}(\boldsymbol{q}_\epsilon, \mathcal{F}_{\boldsymbol{z}})} \leq |\mathcal{I}|^\beta < \infty$. Thus $|\cdot|^\beta$ satisfies the asymptotic optimality criterion (cf. [20, Definition 3.1]) for $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ and $\boldsymbol{z}$, which implies that the error bound is optimal in the sense of [20, Theorem 3.2(b)].

Using Theorems 3.5 and 3.7 together with [19, Lemma 3.9], we have the following result concerning $\mathbb{1}$-FRFs.

COROLLARY 3.9. *Consider* $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ *and its dual cone* $(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^*$.
(i) *Let* $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* \backslash \{\boldsymbol{0}\}$ *with* $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$ *and let* $\mathcal{F}_r := \{\boldsymbol{z}\}^\perp \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. *Let* $\gamma_{\boldsymbol{z},t}$ *be defined as in* (3.12). *Then the function* $\psi_{\mathcal{P}_{m,n}^{\boldsymbol{\alpha}},\boldsymbol{z}} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ *given by*

(3.32)
$$\psi_{\mathcal{P}_{m,n}^{\boldsymbol{\alpha}},\boldsymbol{z}}(\epsilon,t) := \max\{\epsilon, \epsilon/\|\boldsymbol{z}\|\} + \max\{2\sqrt{t}, 2\gamma_{\boldsymbol{z},t}^{-1}\}(\epsilon + \max\{\epsilon, \epsilon/\|\boldsymbol{z}\|\})^{\frac{1}{2}}$$

is a one-step facial residual function for $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ *and* $\boldsymbol{z}$.
(ii) *Let* $\boldsymbol{z} \in \partial(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* \backslash \{\boldsymbol{0}\}$ *with* $\overline{\boldsymbol{z}} = \boldsymbol{0}$ *and let* $\mathcal{F}_{\boldsymbol{z}} := \{\boldsymbol{z}\}^\perp \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. *Let* $\gamma_{\boldsymbol{z},t}$ *be defined as in* (3.27), *where* $\beta := \sum_{i : \widetilde{z}_i > 0} \alpha_i$. *Then the function* $\psi_{\mathcal{P}_{m,n}^{\boldsymbol{\alpha}},\boldsymbol{z}} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ *given by*

(3.33)
$$\psi_{\mathcal{P}_{m,n}^{\boldsymbol{\alpha}},\boldsymbol{z}}(\epsilon,t) := \max\{\epsilon, \epsilon/\|\boldsymbol{z}\|\} + \max\{2t^{1-\beta}, 2\gamma_{\boldsymbol{z},t}^{-1}\}(\epsilon + \max\{\epsilon, \epsilon/\|\boldsymbol{z}\|\})^\beta$$

is a one-step facial residual function for $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ *and* $\boldsymbol{z}$.

We now collect these results to show the tight error bounds for $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$.

THEOREM 3.10 (error bounds for the generalized power cone and their optimality). *Consider* $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ *and its dual cone* $(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^*$. *Let* $\mathcal{L} \subseteq \mathbb{R}^{m+n}$ *be a subspace and* $\boldsymbol{a} \in \mathbb{R}^{m+n}$ *be given. Suppose that* $(\mathcal{L} + \boldsymbol{a}) \cap \mathcal{P}_{m,n}^{\boldsymbol{\alpha}} \neq \emptyset$. *Then the following items hold:*
(i) $d_{\mathrm{PPS}}(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}, \mathcal{L} + \boldsymbol{a}) \leq 1$.
(ii) *If* $d_{\mathrm{PPS}}(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}, \mathcal{L} + \boldsymbol{a}) = 0$, *then a Lipschitzian error bound holds.*

(iii) *If $d_{\mathrm{PPS}}(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}, \mathcal{L} + \boldsymbol{a}) = 1$, consider the chain of faces $\mathcal{F} \subsetneq \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ with length being* 2.
  (a) *If $\mathcal{F} = \mathcal{F}_r$, then a Hölderian error bound with exponent 1/2 holds.*
  (b) *If $\mathcal{F} = \mathcal{F}_{\boldsymbol{z}}$ with $\boldsymbol{z} \in (\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* \cap \mathcal{L}^\perp \cap \{\boldsymbol{a}\}^\perp$, then a Hölderian error bound with exponent $\beta := \sum_{i : \widetilde{z}_i > 0} \alpha_i$ holds.*
  (c) *If $\mathcal{F} = \{\boldsymbol{0}\}$, then a Lipschitzian error bound holds.*
(iv) *All these error bounds are the best in the sense stated in* [20, *Theorem* 3.2(b)].

Proof. As is shown in section 3.1, all the proper exposed faces of the generalized power cone are polyhedral. Then the process of facial reduction needs at most one step to reach the PPS condition. Hence, $d_{\mathrm{PPS}}(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}, \mathcal{L} + \boldsymbol{a}) \leq 1$. This shows item (i).

If $d_{\mathrm{PPS}}(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}, \mathcal{L} + \boldsymbol{a}) = 0$, i.e., (Feas) satisfies the PPS condition, then by [2, Corollary 3], a Lipschitzian error bound holds. This shows item (ii).

Next, let $d_{\mathrm{PPS}}(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}, \mathcal{L} + \boldsymbol{a}) = 1$; i.e., we need one step to reach the PPS condition. In this case, the error bound depends on the exposed face $\mathcal{F}$ that contains the feasible region. If $\mathcal{F} = \mathcal{F}_{\mathrm{r}}$, then by Corollary 3.9 (i), we conclude that a Hölderian error bound with exponent 1/2 holds. Remark 3.6 implies that $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$ satisfies the asymptotic optimality criterion for $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ and $\boldsymbol{z}$ with $\overline{\boldsymbol{z}} \neq \boldsymbol{0}$. Hence, by [20, Theorem 3.2], the obtained Hölderian error bound with exponent 1/2 is the best error bound.

If $\mathcal{F} = \mathcal{F}_{\boldsymbol{z}}$ with $\boldsymbol{z} \in (\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})^* \cap \mathcal{L}^\perp \cap \{\boldsymbol{a}\}^\perp$, then using Corollary 3.9(ii), we conclude that a Hölderian error bound with exponent $\beta := \sum_{i \in \mathcal{I}} \alpha_i$ holds, where $\mathcal{I} = \{i \mid \widetilde{z}_i > 0\}$. The optimality of this error bound comes from Remark 3.8 and [20, Theorem 3.2]. If $\mathcal{F} = \{\boldsymbol{0}\}$, which means the feasible region is $\{\boldsymbol{0}\}$, then a Lipschitzian error bound holds automatically and it is naturally tight; see [22, Proposition 27]. $\square$

4. Application: Self-duality, homogeneity, irreducibility, and perfectness of $\boldsymbol{\mathcal{P}_{m,n}^{\alpha}}$. In this section, we consider the *self-duality, homogeneity, irreducibility*, and *perfectness* of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. We first briefly explain the importance of those questions.

In what follows, we need the following concepts. We will denote by $\mathrm{Aut}(\mathcal{K})$ the group of automorphisms of $\mathcal{K}$ which are the linear bijections $\boldsymbol{M} : \mathcal{E} \to \mathcal{E}$ such that $\boldsymbol{M}\mathcal{K} = \mathcal{K}$. Then, the Lie algebra of $\mathrm{Aut}(\mathcal{K})$ denoted by $\mathrm{Lie}\,\mathrm{Aut}(\mathcal{K})$ corresponds to the linear maps $\boldsymbol{L}$ for which $e^{t\boldsymbol{L}} \in \mathrm{Aut}(\mathcal{K})$ for all $t \in \mathbb{R}$ or, equivalently, is the tangent space at the identity element when $\mathrm{Aut}(\mathcal{K})$ is seen as a Lie group.

Recall that a cone $\mathcal{K}$ is called *self-dual* if there exists a positive definite matrix $\boldsymbol{Q}$ such that $\boldsymbol{Q}\mathcal{K} = \mathcal{K}^*$. This is equivalent to the existence of *some* inner product under which $\mathcal{K}$ becomes self-dual, e.g., [14, Proposition 1]. A cone is *homogeneous* if for every $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{ri}\,\mathcal{K}$ there is a matrix $\boldsymbol{A} \in \mathrm{Aut}(\mathcal{K})$ such that $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$. A homogeneous and self-dual cone is called *symmetric* [7].

If a closed convex cone $\mathcal{K}$ can be expressed as a direct sum of two nonempty and nontrivial sets $\mathcal{K}_1, \mathcal{K}_2 \subset \mathcal{K}$, i.e., $\mathcal{K} = \mathcal{K}_1 + \mathcal{K}_2$ with $\mathcal{K}_1 \neq \{\boldsymbol{0}\}, \mathcal{K}_2 \neq \{\boldsymbol{0}\}$ and $\mathrm{span}(\mathcal{K}_1) \cap \mathrm{span}(\mathcal{K}_2) = \{\boldsymbol{0}\}$, then $\mathcal{K}$ is said to be *reducible*; it might not be immediately obvious, but this forces $\mathcal{K}_1$ and $\mathcal{K}_2$ to be convex cones; see, e.g., [21, Lemma 3.2]. Otherwise, $\mathcal{K}$ is said to be *irreducible* or *indecomposable*; see, e.g., [21, 1, 10].

4.1. Some theoretical context. It is relatively recent that the power cone has been a subject of research in optimization. However, the power cone was first considered in the 1950s by Max Koecher in the context of the so-called *domains of positivity*; see [16]. More precisely, Koecher proposed a family of 3D cones in [16, section 11,d)] which corresponds to $\mathcal{P}_{1,2}^{\boldsymbol{\alpha}}$, with $\alpha \in (0,1)$. After that, the power cone languished in relative obscurity inside the optimization community, although it was discussed briefly in [42] and in [41] under the name *Koecher cone.* As indicated in

the introduction, several works helped to revitalize the interest in power cones by showcasing modeling applications, algorithms, and software [5, 28, 37, 15, 34, 6].

When the power cone is bundled together in the class of "nonsymmetric cones," it might be interesting to take a step back and understand two points: (a) how exactly the power cone fails to be symmetric, and (b) why one should care about this.

Starting from the latter, what is special about symmetric cones is that they are supported by a powerful theory of *Jordan algebras* [7]. Being a symmetric cone is a *very* favorable property that was heavily exploited to develop efficient primal-dual interior point algorithms; see, e.g., [8]. However, being a symmetric cone is also restrictive for it is known that, up to linear isomorphism, each symmetric cone is a direct product of only five types of cones. The most remarkable examples of symmetric cones are the $\mathbb{R}^n_+$, the real symmetric positive semidefinite matrices $\mathcal{S}^n_+$, the second-order cone, and the direct products of these three.

As for item (a), examining (3.1), we immediately see that the dual of $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$ under the Euclidean inner product is just $\boldsymbol{D}\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$, where $\boldsymbol{D}$ is a diagonal matrix with positive entries, so $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$ is indeed self-dual in the sense above. Thus the only gap between $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$ and the class of symmetric cones is the homogeneity.

Given that being symmetric is very advantageous, one may reasonably wonder if the family of cones $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$ parametrized by $\boldsymbol{\alpha}$ and m and n are indeed nonhomogeneous in general. To the best of our knowledge, although it is well known (e.g., see comments in [42, section 4]) that $\mathcal{P}^{\boldsymbol{\alpha}}_{1,2}$ is nonhomogeneous except when $\boldsymbol{\alpha} = (1/2, 1/2)$, there is no result on the generalized power cone regarding which combination of the parameters m, n, and $\boldsymbol{\alpha}$ leads to homogeneity or not. We fill this gap with Theorem 4.2 and Corollary 4.4, which tells us precisely which of the generalized power cones are homogeneous or not.

We also completely determine the automorphism group of $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$. While this may seem an esoteric question, the automorphism group of a cone $\mathcal{K}$ is intimately connected to complementarity questions over $\mathcal{K}$. For example, it is known that $\boldsymbol{L}$ belongs to the Lie algebra of $\mathrm{Aut}(\mathcal{K})$ if and only if the following implication holds:

$$\boldsymbol{x} \in \mathcal{K}, \boldsymbol{y} \in \mathcal{K}^*, \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0 \Rightarrow \langle \boldsymbol{L}\boldsymbol{x}, \boldsymbol{y} \rangle = 0;$$

see [9]. If a cone has "enough" automorphisms, then a complementarity problem can be rewritten as a square system using the matrices from the Lie algebra of $\mathrm{Aut}(\mathcal{K})$. In particular, when the dimension of $\mathrm{Aut}(\mathcal{K})$ is at least $\dim \mathcal{K}$, then the cone is said to be *perfect*; see [9, page 5] and [32, Theorem 1]. An example of this phenomenon is how the conditions $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n_+, \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$ imply n equations $x_i y_i = 0$, which is useful in several contexts.

The quantity $\dim \mathrm{Aut} \mathcal{K}$ is called the *Lyapunov rank of $\mathcal{K}$* [9, 10] and is additive with respect to direct sums [9, Proposition 1]. Since any cone can be written as a direct sum of irreducible cones, it becomes important to identify *which* irreducible cones are perfect.

It is interesting to note that many of the examples of irreducible perfect cones in the literature (e.g., [9, 10, 32]) seem to be homogeneous. In addition, every homogeneous cone is perfect, which follows by known results about Lie groups; e.g., see [17, Theorem 21.20] or section 2 in [31], which summarizes useful results. The final observation we will make in this paper is that, surprisingly, for some choices of parameters, $\mathcal{P}^{\boldsymbol{\alpha}}_{m,n}$ is perfect but nonhomogeneous; see Corollary 4.4. We note that in [40], Sznajder showed that there are choices of parameters for which the so-called *extended second-order cone* is irreducible and perfect. This corresponds to a family

of cones proposed by Németh and Zhang that contains the second-order cones [29]. However, as far as we know, the homogeneity of those cones (or the lack thereof) was not discussed in general.

4.2. Automorphisms of the generalized power cone. In this subsection, we will prove our main results regarding $\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$. The basic strategy is simple: if $\boldsymbol{A} \in \mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$, then $\boldsymbol{A}$ must map a face $\mathcal{F}_1$ of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ to another face $\mathcal{F}_2$ of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ with the same properties such as the dimension. More than that, the optimal exponents associated to FRFs of $\mathcal{F}_1$ and $\mathcal{F}_2$ must be the same. These conditions impose enough restrictions on $\boldsymbol{A}$ that we are able to completely determine its shape. Note that when $n = 2$ and $\boldsymbol{\alpha} = (1/2, 1/2)$, $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ is isomorphic to the second-order cone, whose automorphism group is well known. Below, we focus on the complementary cases.

THEOREM 4.1 (automorphisms of $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$). *For $m \geq 1, n > 2$, and any $\boldsymbol{\alpha} \in (0,1)^n$ such that $\sum_{i=1}^n \alpha_i = 1$, or for $m \geq 1, n = 2$, and any $\boldsymbol{\alpha} \in (0,1)^2$ such that $\alpha_1 \neq \alpha_2$ and $\alpha_1 + \alpha_2 = 1$, it holds that $\boldsymbol{A} \in \mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$ if and only if*

$$(4.1) \qquad \boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{E} \end{bmatrix}$$

for some (invertible) generalized permutation matrix[7] $\boldsymbol{E} \in \mathbb{R}^{n \times n}$ with positive nonzero entries and invertible matrix $\boldsymbol{B} \in \mathbb{R}^{m \times m}$ satisfying $\|\boldsymbol{B}\boldsymbol{x}\| = \prod_{k=1}^n (E_{k,l_k})^{\alpha_{l_k}} \|\boldsymbol{x}\|$ for all $\boldsymbol{x} \in \mathbb{R}^m$, where E_{k,l_k} is the nonzero element in the kth row of $\boldsymbol{E}$ and $\alpha_{l_k} = \alpha_k$.

Proof. Suppose that there exists a matrix

$$\boldsymbol{A} := \begin{bmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{D} & \boldsymbol{E} \end{bmatrix} \text{ with } \boldsymbol{B} \in \mathbb{R}^{m \times m},\, \boldsymbol{C} \in \mathbb{R}^{m \times n},\, \boldsymbol{D} \in \mathbb{R}^{n \times m},\, \boldsymbol{E} \in \mathbb{R}^{n \times n}$$

such that $\boldsymbol{A}\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} = \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$.

First note that the entries of $\boldsymbol{E}$ must all be nonnegative, for if the (i,j)th entry was negative, then we could pick a vector $\boldsymbol{q} := (\boldsymbol{0}, \boldsymbol{c}) \in \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ with $c_j = 1$ and $c_k = 0$ for $k \neq j$, wherefore $\boldsymbol{A}\boldsymbol{q} \notin \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, which is a contradiction.

Additionally, such a matrix $\boldsymbol{A}$ must be invertible, and if ψ is an FRF for a face $\mathcal{F}_1 \trianglelefteq \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, then $\boldsymbol{A}$ must map $\mathcal{F}_1$ onto a face $\mathcal{F}_2 \trianglelefteq \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ which has the same dimension and admits an FRF that is a positively rescaled shift of ψ; see [22, Proposition 17].

Observe from section 3.1 that the generalized power cone has two types of faces defined in (3.4) and (3.5) (denoted by $\mathcal{F}_r$ and $\mathcal{F}_{\boldsymbol{z}}$ respectively, with an abuse of notation) with the corresponding (optimal) $\mathbb{1}$-FRFs in (3.32) and (3.33), respectively. We also notice that the dimension of the faces of the first type is 1, while the dimension of a face of the second type is $n - |\mathcal{I}|$. These lead to the following observations:

(I) Given an $\mathcal{I}$ with $\beta_{\mathcal{I}} := \sum_{i \in \mathcal{I}} \alpha_i$, if $|\mathcal{I}| < n - 1$, i.e., the dimension of the corresponding face is larger than 1, then $\boldsymbol{A}$ must map the face associated with $\mathcal{I}$ to a face associated with an $\bar{\mathcal{I}}$ where $|\mathcal{I}| = |\bar{\mathcal{I}}|$ and $\beta_{\bar{\mathcal{I}}} = \beta_{\mathcal{I}}$.

(II) In the case when $n = 2$, since we assumed $\alpha_1 \neq \alpha_2$ and thus $\alpha_1 \neq 1/2$, $\boldsymbol{A}$ cannot map a one-dimensional face of type $\mathcal{F}_{\boldsymbol{z}}$ (whose FRF admits an optimal exponent of α_1 or α_2) to one of type $\mathcal{F}_r$ (whose FRF admits an optimal exponent of $1/2$).

Thus, a face of type $\mathcal{F}_{\boldsymbol{z}}$ with $|\mathcal{I}| = 1$ must be mapped to a face of the same type. From now on, for each $k \in \{1, 2, \ldots, n\}$, we let i_k and l_k be such that $\boldsymbol{A}\mathcal{F}_{\{k\}} =$

---

[7]A generalized permutation matrix is a matrix where in each column and each row there is exactly one nonzero entry.

$\mathcal{F}_{\{i_k\}}$ and $\boldsymbol{A}\mathcal{F}_{\{l_k\}} = \mathcal{F}_{\{k\}}$, where $\mathcal{F}_{\{k\}}$ denotes the face of type $\mathcal{F}_{\boldsymbol{z}}$ associated with $\mathcal{I} = \{k\}$. We deduce immediately from the above discussions that $\{1, 2, \ldots, n\} = \{i_1, i_2, \ldots, i_n\} = \{l_1, l_2, \ldots, l_n\}$ and $\alpha_k = \alpha_{i_k} = \alpha_{l_k}$.

Now, fix any $k \in \{1, 2, \ldots, n\}$. Then for any $\widetilde{\boldsymbol{x}}_{\mathcal{I}} := (c_1, \ldots, c_{k-1}, 0, c_{k+1}, \ldots, c_n)$ with $c_i > 0$ for all $i \neq k$, it must hold that $\boldsymbol{A}$ maps $\boldsymbol{x}_{\mathcal{I}} := (\boldsymbol{0}, \widetilde{\boldsymbol{x}}_{\mathcal{I}})$ to some $\boldsymbol{x}_{\hat{\mathcal{I}}} := (\boldsymbol{0}, \widetilde{\boldsymbol{x}}_{\hat{\mathcal{I}}})$ with $\hat{\mathcal{I}} = \{i_k\}$, $\alpha_k = \alpha_{i_k}$, and $(\widetilde{\boldsymbol{x}}_{\hat{\mathcal{I}}})_{i_k} = 0$. Thus,

$$\begin{bmatrix} \boldsymbol{B} & \boldsymbol{C} \\ \boldsymbol{D} & \boldsymbol{E} \end{bmatrix} \begin{bmatrix} \boldsymbol{0} \\ \widetilde{\boldsymbol{x}}_{\mathcal{I}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ \widetilde{\boldsymbol{x}}_{\hat{\mathcal{I}}} \end{bmatrix}.$$

Therefore, we have $\boldsymbol{C}\widetilde{\boldsymbol{x}}_{\mathcal{I}} = \boldsymbol{0}$. This together with the arbitrariness of $c_i > 0$ shows that all except possibly the kth column of $\boldsymbol{C}$ are $\boldsymbol{0}$. Since k is arbitrary, then we conclude that $\boldsymbol{C} = \boldsymbol{0}$.

Next, notice that we also have $\boldsymbol{E}\widetilde{\boldsymbol{x}}_{\mathcal{I}} = \widetilde{\boldsymbol{x}}_{\hat{\mathcal{I}}}$. Since $(\widetilde{\boldsymbol{x}}_{\hat{\mathcal{I}}})_{i_k} = 0$, we see that $E_{i_k}\widetilde{\boldsymbol{x}}_{\mathcal{I}} = 0$, where E_{i_k} is the i_kth row of $\boldsymbol{E}$. Using again the arbitrariness of $c_i > 0$ in the definition of $\widetilde{\boldsymbol{x}}_{\mathcal{I}}$, we conclude that all entries of E_{i_k} are 0 except possibly for the kth entry, i.e., E_{i_k} has only one possibly nonzero entry and that entry is nonnegative. From the arbitrariness of k and the fact that $\{i_1, i_2, \ldots, i_n\} = \{1, 2, \ldots, n\}$, we immediately obtain that every entry of the i_kth row $\boldsymbol{E}$ has all of its entries equal to zero except possibly for the kth, which is nonnegative.

Taking into account the fact that $\boldsymbol{A}$ is invertible and $\boldsymbol{C} = \boldsymbol{0}$, we know that none of the columns of $\boldsymbol{E}$ can be identically zero, and so we altogether have that each of the rows and columns of $\boldsymbol{E}$ consists of one strictly positive entry, with all other entries identically zero.[8]

We next claim that $\boldsymbol{A}$ must map faces of type $\mathcal{F}_{\mathrm{r}}$ to a face of type $\mathcal{F}_{\mathrm{r}}$. Since $\boldsymbol{A}$ must permute faces whose FRFs admit the same optimal exponent, we only need to consider the extreme case that there exists a face of type $\mathcal{F}_{\boldsymbol{z}}$ corresponding to an $\mathcal{I} := \{1, 2, \ldots, i-1, i+1, \ldots, n\}$ for some i (i.e., the dimension of the corresponding face is 1) with $\beta_{\mathcal{I}} = 1/2$, and argue that $\boldsymbol{A}$ cannot map $\mathcal{F}_{\mathrm{r}}$ onto such $\mathcal{F}_{\boldsymbol{z}}$. Suppose for contradiction that this happens; then there must exist $\boldsymbol{x} = (\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}})$ in some face of type $\mathcal{F}_{\mathrm{r}}$ with $\overline{\boldsymbol{x}} \neq \boldsymbol{0}$ and $\widetilde{x}_i > 0$ for all i such that

$$\begin{bmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{D} & \boldsymbol{E} \end{bmatrix} \begin{bmatrix} \overline{\boldsymbol{x}} \\ \widetilde{\boldsymbol{x}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{0} \\ e_i \end{bmatrix},$$

where $\boldsymbol{e}_i \in \mathbb{R}^n$ is the vector whose elements are all zero except for the ith element being 1. However, this cannot happen because $\boldsymbol{B}\overline{\boldsymbol{x}} = \boldsymbol{0}$ and the invertibility of $\boldsymbol{B}$ (a consequence of invertibility of $\boldsymbol{A}$) implies $\overline{\boldsymbol{x}} = \boldsymbol{0}$, leading to a contradiction. Hence, $\boldsymbol{A}$ must map faces of type $\mathcal{F}_{\mathrm{r}}$ onto a face of type $\mathcal{F}_{\mathrm{r}}$.

Thus, for any $\boldsymbol{x} = (\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}})$ in one of the type $\mathcal{F}_{\mathrm{r}}$ faces with $\overline{\boldsymbol{x}} \neq \boldsymbol{0}$, $\min_i\{\widetilde{x}_i\} > 0$, and $\|\overline{\boldsymbol{x}}\| = \prod_{i=1}^n \widetilde{x}_i^{\alpha_i}$, there must be $\boldsymbol{y} = (\overline{\boldsymbol{y}}, \widetilde{\boldsymbol{y}})$ in one of the type $\mathcal{F}_{\mathrm{r}}$ faces with $\overline{\boldsymbol{y}} \neq \boldsymbol{0}$, $\min_i\{\widetilde{y}_i\} > 0$, and $\|\overline{\boldsymbol{y}}\| = \prod_{i=1}^n \widetilde{y}_i^{\alpha_i}$ such that

$$\begin{bmatrix} \boldsymbol{B} & \boldsymbol{0} \\ \boldsymbol{D} & \boldsymbol{E} \end{bmatrix} \begin{bmatrix} \overline{\boldsymbol{x}} \\ \widetilde{\boldsymbol{x}} \end{bmatrix} = \begin{bmatrix} \overline{\boldsymbol{y}} \\ \widetilde{\boldsymbol{y}} \end{bmatrix}.$$

Recall that there is exactly one nonzero element in each row of $\boldsymbol{E}$, and this element is positive. From the definition of l_k, this nonzero element is E_{k,l_k}; see footnote 8.

---

[8]Then, we have shown that $E_{s,r} \neq 0$ if and only if $(s, r) = (i_k, k)$ for some $k \in \{1, 2, \ldots, n\}$ (or equivalently $(s, r) = (k, l_k)$ for some $k \in \{1, 2, \ldots, n\}$).

Fix any j and $k \in \{1, \ldots, n\}$. Pick any $(\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) \in \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ such that $\overline{\boldsymbol{x}} = \boldsymbol{e}_j$ and $\prod_{i=1}^n \widetilde{x}_i^{\alpha_i} = 1$. For any $t > 0$, one can check that $(t^{\alpha_{l_k}} \overline{\boldsymbol{x}}, \widetilde{x}_1, \ldots, t\widetilde{x}_{l_k}, \ldots, \widetilde{x}_n) \in \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ belongs to a face of type $\mathcal{F}_{\mathrm{r}}$. Thus, there exists $(\overline{\boldsymbol{y}}, \widetilde{\boldsymbol{y}})$ such that

$$t^{\alpha_{l_k}} \boldsymbol{B} \boldsymbol{e}_j = \overline{\boldsymbol{y}} \quad \text{and} \quad t^{\alpha_{l_k}} D_{k,j} + t E_{k,l_k} \widetilde{x}_{l_k} = \widetilde{y}_k > 0.$$

The second relation implies that $D_{k,j} + t^{1-\alpha_{l_k}} E_{k,l_k} \widetilde{x}_{l_k} > 0$. Letting $t \downarrow 0$, we conclude that $D_{k,j} \geq 0$. As the choices of j and k were arbitrary, we see that all entries of $\boldsymbol{D}$ are nonnegative. Considering $\overline{\boldsymbol{x}} = -\boldsymbol{e}_j$, a similar argument shows that all entries of $\boldsymbol{D}$ are nonpositive. Hence, $\boldsymbol{D} = \boldsymbol{0}$.

Now, for any $\overline{\boldsymbol{x}} \in \mathbb{R}^m$, pick any $(\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) \in \partial \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. Then there exists $(\overline{\boldsymbol{y}}, \widetilde{\boldsymbol{y}}) \in \partial \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ so that[9]

$$\boldsymbol{B} \overline{\boldsymbol{x}} = \overline{\boldsymbol{y}} \quad \text{and} \quad E_{k,l_k} \widetilde{x}_{l_k} = \widetilde{y}_k \quad \text{for } k = 1, 2, \ldots, n.$$

Thus,

$$\|\boldsymbol{B}\overline{\boldsymbol{x}}\| = \|\overline{\boldsymbol{y}}\| = \prod_{k=1}^n \widetilde{y}_k^{\alpha_k} = \prod_{k=1}^n (E_{k,l_k} \widetilde{x}_{l_k})^{\alpha_k} \stackrel{\text{(a)}}{=} \prod_{k=1}^n (E_{k,l_k} \widetilde{x}_{l_k})^{\alpha_{l_k}}$$

$$= \prod_{k=1}^n E_{k,l_k}^{\alpha_{l_k}} \prod_{i=1}^n \widetilde{x}_i^{\alpha_i} = \prod_{k=1}^n E_{k,l_k}^{\alpha_{l_k}} \|\overline{\boldsymbol{x}}\|,$$

where (a) holds as $\alpha_k = \alpha_{l_k}$ for all k. The above shows the necessity of the form in (4.1).

Conversely, if $\boldsymbol{A}$ is a matrix of the form (4.1), then $\boldsymbol{A}$ must be invertible since $\boldsymbol{B}$ and $\boldsymbol{E}$ are invertible. For any $\boldsymbol{x} = (\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) \in \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, we have $\boldsymbol{A}\boldsymbol{x} = (\boldsymbol{B}\overline{\boldsymbol{x}}, \boldsymbol{E}\widetilde{\boldsymbol{x}})$. Hence,

$$\|\boldsymbol{B}\overline{\boldsymbol{x}}\| = \prod_{k=1}^n E_{k,l_k}^{\alpha_{l_k}} \|\overline{\boldsymbol{x}}\| \leq \prod_{k=1}^n E_{k,l_k}^{\alpha_{l_k}} \prod_{i=1}^n \widetilde{x}_i^{\alpha_i} = \prod_{k=1}^n (E_{k,l_k} \widetilde{x}_{l_k})^{\alpha_{l_k}},$$

where the last equality holds as $\{1, \ldots, n\} = \{l_1, \ldots, l_n\}$. This implies $\boldsymbol{A}\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} \subseteq \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$.

We claim

$$(4.2) \quad \text{(i)} \ \left(\boldsymbol{E}^{-1}\right)_{i,j} = \begin{cases} 0, & E_{j,i} = 0, \\ \frac{1}{E_{j,i}}, & E_{j,i} \neq 0; \end{cases} \quad \text{(ii)} \ \|\boldsymbol{B}^{-1}\boldsymbol{x}\| = \prod_{k=1}^n E_{k,l_k}^{-\alpha_{l_k}} \|\boldsymbol{x}\| \quad \forall \boldsymbol{x} \in \mathbb{R}^m.$$

Granting these, we have that for any $\boldsymbol{x} = (\overline{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) \in \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, $\boldsymbol{A}^{-1}\boldsymbol{x} = (\boldsymbol{B}^{-1}\overline{\boldsymbol{x}}, \boldsymbol{E}^{-1}\widetilde{\boldsymbol{x}})$ satisfies

$$\prod_{i=1}^n (\boldsymbol{E}^{-1}\widetilde{\boldsymbol{x}})_i^{\alpha_i} = \prod_{i=1}^n \left(\sum_{j=1}^n (\boldsymbol{E}^{-1})_{i,j} \widetilde{x}_j\right)^{\alpha_i} \stackrel{\text{(a)}}{=} \prod_{k=1}^n \left((\boldsymbol{E}^{-1})_{l_k,k} \widetilde{x}_k\right)^{\alpha_{l_k}} = \prod_{k=1}^n (E_{k,l_k}^{-1} \widetilde{x}_k)^{\alpha_{l_k}}$$

$$= \prod_{k=1}^n E_{k,l_k}^{-\alpha_{l_k}} \prod_{i=1}^n \widetilde{x}_i^{\alpha_{l_i}} \stackrel{\text{(b)}}{=} \prod_{k=1}^n E_{k,l_k}^{-\alpha_{l_k}} \prod_{i=1}^n \widetilde{x}_i^{\alpha_i} \geq \prod_{k=1}^n E_{k,l_k}^{-\alpha_{l_k}} \|\overline{\boldsymbol{x}}\| \stackrel{\text{(c)}}{=} \|\boldsymbol{B}^{-1}\overline{\boldsymbol{x}}\|,$$

where (a) is true thanks to the fact that in the sum there is only one nonzero term, which comes from identity (i) and footnote 8; (b) holds because $\alpha_k = \alpha_{l_k}$ for all k;

_____

[9]Such a $\boldsymbol{y}$ exists because $\boldsymbol{A}$ is invertible and $\boldsymbol{A}\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} = \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$, which implies $\boldsymbol{A}\mathrm{ri}\,\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} = \mathrm{ri}\,\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ and $\boldsymbol{A}\partial\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} = \partial\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$.

(c) comes from identity (ii). Hence, $\boldsymbol{A}^{-1}\boldsymbol{x} \in \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$. This implies $\boldsymbol{A}\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} \supseteq \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ and consequently $\boldsymbol{A}\mathcal{P}_{m,n}^{\boldsymbol{\alpha}} = \mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$.

Now, it remains to show (4.2). Since $\boldsymbol{E}$ is a generalized permutation matrix with all nonzero elements being positive, then we immediately have (i) from $\boldsymbol{E}\boldsymbol{E}^{-1} = \boldsymbol{I}_n$. Recall that, by assumption, $\|\boldsymbol{B}\boldsymbol{x}\| = \prod_{k=1}^{n} E_{k,l_k}^{\alpha_{l_k}}\|\boldsymbol{x}\|$ for any $\boldsymbol{x} \in \mathbb{R}^m$ and $\boldsymbol{B}$ is invertible. Using these, we can deduce (ii) in (4.2) as follows: for any $\boldsymbol{x} \in \mathbb{R}^m$,

$$\|\boldsymbol{x}\| = \|\boldsymbol{B}\boldsymbol{B}^{-1}\boldsymbol{x}\| = \prod_{k=1}^{n} E_{k,l_k}^{\alpha_{l_k}}\|\boldsymbol{B}^{-1}\boldsymbol{x}\|. \qquad\qquad \square$$

The next theorem is about the dimension of $\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$.

THEOREM 4.2. *Let $m \geq 1$, $n \geq 2$, and $\boldsymbol{\alpha} \in (0,1)^n$ such that $\sum_{i=1}^{n} \alpha_i = 1$; then we have the following statements about $\dim\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$:*
   (i) *If $m \geq 1$, $n = 2$, and $\boldsymbol{\alpha} := (1/2, 1/2)$, then $\dim\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) = (m^2 + 3m + 4)/2$.*
   (ii) *If $m \geq 1$, $n > 2$, and $\sum_{i=1}^{n} \alpha_i = 1$ or $m \geq 1$, $n = 2$, $\alpha_1 \neq \alpha_2$, and $\alpha_1 + \alpha_2 = 1$, then*

$$(4.3)\qquad \mathrm{Lie}\,\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) = \left\{ \begin{bmatrix} \boldsymbol{G} & \boldsymbol{0} \\ \boldsymbol{0} & \mathrm{Diag}(\boldsymbol{h}) \end{bmatrix} \,\middle|\, \begin{array}{l} \boldsymbol{G} + \boldsymbol{G}^{\top} = 2\boldsymbol{\alpha}^{\top}h\boldsymbol{I}_m, \\ \boldsymbol{G} \in \mathbb{R}^{m \times m},\, \boldsymbol{h} \in \mathbb{R}^n \end{array} \right\}.$$

Hence, $\dim\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) = \dim\mathrm{Lie}\,\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) = n + m(m-1)/2$.

Proof. (i) If $m \geq 1$, $n = 2$, and $\boldsymbol{\alpha} := (1/2, 1/2)$, then $\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}$ is isomorphic to a second-order cone; see [28, section 3.1.2]. Hence, we know from [9, page 12 (v)] that

$$\dim\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) = \frac{(m+2)^2 - m}{2} = \frac{m^2 + 3m + 4}{2}.$$

(ii) By [11, Corollary 3.45], $\dim\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}}) = \dim\mathrm{Lie}\,\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$. This in addition to [11, Corollary 3.46] shows that it suffices to calculate the dimension of the tangent space at the identity of $\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$ to obtain $\dim\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$.

First, we compute $\mathrm{Lie}\,\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$ and for that we consider an arbitrary continuously differentiable curve $\boldsymbol{F} : (-1,1) \to \mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$ with $\boldsymbol{F}(0) = \boldsymbol{I}_{m+n}$ and $\boldsymbol{F}(t) \in \mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$ for any $t \in (-1,1)$. We further denote

$$\boldsymbol{F}(t) = \begin{bmatrix} \boldsymbol{G}_t & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{H}_t \end{bmatrix} \quad \text{and} \quad \dot{\boldsymbol{F}}(t) = \begin{bmatrix} \dot{\boldsymbol{G}}_t & \boldsymbol{0} \\ \boldsymbol{0} & \dot{\boldsymbol{H}}_t \end{bmatrix},$$

where $\boldsymbol{G}_t \in \mathbb{R}^{m \times m}$ and $\boldsymbol{H}_t \in \mathbb{R}^{n \times n}$ are both invertible; $\boldsymbol{G}_0 = \boldsymbol{I}_m$, $\boldsymbol{H}_0 = \boldsymbol{I}_n$; $\boldsymbol{H}_t$ is a generalized permutation matrix with all nonzero elements being strictly positive (which we assume, by suitably shrinking the neighborhood of definition of $\boldsymbol{F}$ and reparameterizing, to be only nonzero along the diagonal); $\dot{\boldsymbol{F}}(0)$ lies in the tangent space of $\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$ at $\boldsymbol{I}$, that is,

$$(4.4)\qquad\qquad \dot{\boldsymbol{F}}(0) = \begin{bmatrix} \dot{\boldsymbol{G}}_0 & \boldsymbol{0} \\ \boldsymbol{0} & \dot{\boldsymbol{H}}_0 \end{bmatrix} \in \mathrm{Lie}\,\mathrm{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}});$$

and $\dot{\boldsymbol{G}}_t$ and $\dot{\boldsymbol{H}}_t$ refer to the componentwise derivative of $\boldsymbol{G}$ and $\boldsymbol{H}$ with respect to t, respectively.

Since $\boldsymbol{H}_t$ and $\dot{\boldsymbol{H}}_t$ are diagonal, we let $\boldsymbol{h}_t$ and $\dot{\boldsymbol{h}}_t$ be the diagonal vectors of $\boldsymbol{H}_t$ and $\dot{\boldsymbol{H}}_t$, respectively, i.e., $\boldsymbol{H}_t = \mathrm{Diag}(\boldsymbol{h}_t)$ and $\dot{\boldsymbol{H}}_t = \mathrm{Diag}(\dot{\boldsymbol{h}}_t)$. We also let h_t^k

and $\dot{h}_t^k$ denote the kth element of the vectors $\boldsymbol{h}_t$ and $\dot{\boldsymbol{h}}_t$, respectively. Then, from Theorem 4.1,

$$(4.5) \qquad \|\boldsymbol{G}_t\boldsymbol{x}\|^2 = \prod_{k=1}^n (h_t^k)^{2\alpha_k}\|\boldsymbol{x}\|^2 \quad \forall \boldsymbol{x} \in \mathbb{R}^m, \, \forall t \in (-1, 1).$$

Differentiating[10] both sides of (4.5) with respect to t, we can obtain

$$2\boldsymbol{x}^\top \boldsymbol{G}_t^\top \dot{\boldsymbol{G}}_t \boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{x} \sum_{k=1}^n 2\alpha_k \left(h_t^k\right)^{2\alpha_k-1} \dot{h}_t^k \prod_{j \neq k} \left(h_t^j\right)^{2\alpha_j} = \boldsymbol{x}^\top \boldsymbol{x} \sum_{k=1}^n 2\frac{\alpha_k}{h_t^k} \dot{h}_t^k \prod_{j=1}^n \left(h_t^j\right)^{2\alpha_j}$$

$$= 2\left(\boldsymbol{x}^\top \boldsymbol{x} \prod_{j=1}^n (h_t^j)^{2\alpha_j}\right) \sum_{k=1}^n \frac{\alpha_k}{h_t^k} \dot{h}_t^k \overset{\text{(a)}}{=} 2\boldsymbol{x}^\top \boldsymbol{G}_t^\top \boldsymbol{G}_t \boldsymbol{x} \left(\boldsymbol{\alpha} \circ (\boldsymbol{h}_t)^{-1}\right)^\top \dot{\boldsymbol{h}}_t,$$

where the inverse is taken componentwise, and the rest of (a) comes from (4.5). Notice that $\left(\boldsymbol{\alpha} \circ (\boldsymbol{h}_t)^{-1}\right)^\top \dot{\boldsymbol{h}}_t$ is a scalar, and by rearranging terms, one has

$$\boldsymbol{x}^\top \left[\boldsymbol{G}_t^\top \dot{\boldsymbol{G}}_t - \left(\boldsymbol{\alpha} \circ (\boldsymbol{h}_t)^{-1}\right)^\top \dot{\boldsymbol{h}}_t \boldsymbol{G}_t^\top \boldsymbol{G}_t\right] \boldsymbol{x} = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^m, \, \forall t \in (-1, 1).$$

Letting $t = 0$ and recalling $\boldsymbol{G}_0 = \boldsymbol{I}_m, \boldsymbol{H}_0 = \boldsymbol{I}_n$, we have

$$(4.6) \qquad \boldsymbol{x}^\top \left(\dot{\boldsymbol{G}}_0 - \boldsymbol{\alpha}^\top \dot{\boldsymbol{h}}_0 I_m\right) \boldsymbol{x} = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^m.$$

Recall that $2\boldsymbol{x}^\top \dot{\boldsymbol{G}}_0 \boldsymbol{x} = \boldsymbol{x}^\top (\dot{\boldsymbol{G}}_0 + \dot{\boldsymbol{G}}_0^\top)\boldsymbol{x}$. We can thus rewrite (4.6) as

$$\boldsymbol{x}^\top \left(\dot{\boldsymbol{G}}_0 + \dot{\boldsymbol{G}}_0^\top - 2\boldsymbol{\alpha}^\top \dot{\boldsymbol{h}}_0 I_m\right) \boldsymbol{x} = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^m.$$

Since the matrix in the parentheses is zero, the above display implies that

$$\dot{\boldsymbol{G}}_0 + \dot{\boldsymbol{G}}_0^\top = 2\boldsymbol{\alpha}^\top \dot{\boldsymbol{h}}_0 I_m.$$

The above derivation and (4.4) show that any matrix in $\text{Lie Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$ satisfies the above display.

Conversely, suppose that $\boldsymbol{G}$ and $\text{Diag}(\boldsymbol{h})$ are such that $\boldsymbol{G} + \boldsymbol{G}^\top = 2\boldsymbol{\alpha}^\top \boldsymbol{h} I_m$ and $\boldsymbol{U} := \begin{bmatrix} \boldsymbol{G} & \boldsymbol{0} \\ \boldsymbol{0} & \text{Diag}(\boldsymbol{h}) \end{bmatrix}$. We need to show that the matrix exponential $e^{t\boldsymbol{U}}$ belongs to $\text{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$ for every $t \in \mathbb{R}$. To this end, recalling that $e^{\boldsymbol{X}+\boldsymbol{Y}} = e^{\boldsymbol{X}}e^{\boldsymbol{Y}}$ if $\boldsymbol{X}\boldsymbol{Y} = \boldsymbol{Y}\boldsymbol{X}$, we have

$$e^{t\boldsymbol{G}} = e^{2t\boldsymbol{\alpha}^\top \boldsymbol{h} I_m - t\boldsymbol{G}^\top} = e^{2t\boldsymbol{\alpha}^\top \boldsymbol{h} I_m} e^{-t\boldsymbol{G}^\top} = e^{2t\boldsymbol{\alpha}^\top \boldsymbol{h}} e^{-t\boldsymbol{G}^\top},$$

since $2t\boldsymbol{\alpha}^\top \boldsymbol{h} \boldsymbol{I}_m$ and $-t\boldsymbol{G}^\top$ commute. This shows that $(e^{t\boldsymbol{G}})^\top e^{t\boldsymbol{G}} = e^{t\boldsymbol{G}^\top} e^{t\boldsymbol{G}} = e^{2t\boldsymbol{\alpha}^\top \boldsymbol{h}} \boldsymbol{I}_m$, i.e., $e^{t\boldsymbol{G}}$ is an orthogonal matrix multiplied by the scalar $e^{t\boldsymbol{\alpha}^\top \boldsymbol{h}}$. Then

$$(4.7) \qquad \|e^{t\boldsymbol{G}}\boldsymbol{x}\| = e^{t\boldsymbol{\alpha}^\top \boldsymbol{h}}\|\boldsymbol{x}\| = e^{\sum_{i=1}^n th_i\alpha_i}\|\boldsymbol{x}\| = \prod_{i=1}^n (e^{th_i})^{\alpha_i}\|\boldsymbol{x}\| \quad \forall \boldsymbol{x} \in \mathbb{R}^m.$$

Since

$$e^{t\boldsymbol{U}} = \begin{bmatrix} e^{t\boldsymbol{G}} & \boldsymbol{0} \\ \boldsymbol{0} & e^{\text{Diag}(t\boldsymbol{h})} \end{bmatrix} = \begin{bmatrix} e^{t\boldsymbol{G}} & \boldsymbol{0} \\ \boldsymbol{0} & \text{Diag}(e^{t\boldsymbol{h}}) \end{bmatrix},$$

where $e^{t\boldsymbol{h}}$ corresponds to the vector such that its ith component is e^{th_i} and h_i is the ith component of $\boldsymbol{h}$, we conclude from (4.7) and Theorem 4.1 that $e^{t\boldsymbol{U}} \in \text{Aut}\,(\mathcal{P}_{m,n}^{\boldsymbol{\alpha}})$.

Finally, a direct computation shows that the dimension of the right-hand side of (4.3) is $n + m(m-1)/2$, which is just the claimed dimension. $\square$

---

[10]This calculation simply uses the chain rule to differentiate $(h_t^k)^{2\alpha_k}$ for a given k, and then applies the product rule for the product over all k.

4.3. Homogeneity, irreducibility, and perfectness of generalized power cone. In this subsection, we will use Theorem 4.1 to prove the homogeneity, irreducibility, and perfectness of $\mathcal{P}_{m,n}^{\alpha}$. Before moving on, we recall the following lemma.

LEMMA 4.3.
(i) *If a closed convex pointed cone $\mathcal{K}$ is reducible, i.e., $\mathcal{K}$ is a direct sum of two nonempty, nontrivial sets $\mathcal{K}_1$ and $\mathcal{K}_2$, then we have $\mathcal{K}_1 \trianglelefteq \mathcal{K}$, $\mathcal{K}_2 \trianglelefteq \mathcal{K}$, and $\dim(\mathcal{K}) = \dim(\mathcal{K}_1) + \dim(\mathcal{K}_2)$.*
(ii) *A proper cone $\mathcal{K} \subseteq \mathbb{R}^p$ is perfect if and only if $\dim \operatorname{Lie} \operatorname{Aut}(\mathcal{K}) \geq p$.*

Proof.
(i) The fact that $\mathcal{K}_1$ and $\mathcal{K}_2$ are faces is well known; see [21, Lemma 3.2]. The conclusion on dimensions follows directly from the definition of direct sum.
(ii) This fact comes from [32, Theorem 1] and the first display on [9, page 4]. □

Using Lemma 4.3 and Theorems 4.1 and 4.2, we have the following corollary.

COROLLARY 4.4. *Let $m \geq 1$, $n \geq 2$, and $\boldsymbol{\alpha} \in (0,1)^n$ such that $\sum_{i=1}^{n} \alpha_i = 1$; then the following statements hold for the generalized power cone $\mathcal{P}_{m,n}^{\alpha}$:*
(i) *$\mathcal{P}_{m,n}^{\alpha}$ is irreducible.*
(ii) *If $m \geq 1, n = 2$, and $\boldsymbol{\alpha} := (1/2, 1/2)$, then $\mathcal{P}_{m,n}^{\alpha}$ is homogeneous and perfect.*
(iii) *If $m \geq 1$, $n > 2$, and $\sum_{i=1}^{n} \alpha_i = 1$ or $m \geq 1$, $n = 2$, $\alpha_1 \neq \alpha_2$, and $\alpha_1 + \alpha_2 = 1$, then $\mathcal{P}_{m,n}^{\alpha}$ is nonhomogeneous. In addition, if $1 \leq m \leq 2$, then $\mathcal{P}_{m,n}^{\alpha}$ is not perfect; if $m \geq 3$, then $\mathcal{P}_{m,n}^{\alpha}$ is perfect.*

Proof.
(i) Recall that the two types of faces of $\mathcal{P}_{m,n}^{\alpha}$ are defined as in (3.4) and (3.5), with dimensions being 1 and $n - |\mathcal{I}|$, respectively. Since $\mathcal{I} \neq \emptyset$ and so $|\mathcal{I}| \geq 1$, for any possible pair of nontrivial faces $\mathcal{F}_1$ and $\mathcal{F}_2$ of $\mathcal{P}_{m,n}^{\alpha}$, we have $\dim(\mathcal{F}_1) + \dim(\mathcal{F}_2) < m + n = \dim(\mathcal{P}_{m,n}^{\alpha})$. This together with Lemma 4.3(i) shows that $\mathcal{P}_{m,n}^{\alpha}$ is irreducible.
(ii) If $m \geq 1, n = 2$, and $\boldsymbol{\alpha} := (1/2, 1/2)$, $\mathcal{P}_{m,n}^{\alpha}$ is isomorphic to a second-order cone and so is homogeneous; see, for example, [28, section 3.1.2]. The perfectness holds by Theorem 4.2(i) and Lemma 4.3(ii).
(iii) Take any $m \geq 1, n > 2$ with any $\boldsymbol{\alpha} \in (0,1)^n$ such that $\sum_{i=1}^{n} \alpha_i = 1$ or $m \geq 1, n = 2$ with any $\boldsymbol{\alpha} \in (0,1)^2$ such that $\alpha_1 \neq \alpha_2$, and consider $\boldsymbol{x} = (\mathbf{0}, \widetilde{\boldsymbol{x}}) \in \operatorname{ri} \mathcal{P}_{m,n}^{\alpha}$ and $\boldsymbol{y} = (\overline{\boldsymbol{y}}, \widetilde{\boldsymbol{y}}) \in \operatorname{ri} \mathcal{P}_{m,n}^{\alpha}$, where $\min_i\{\widetilde{x}_i\} > 0$, $\min_i\{\widetilde{y}\} > 0$ and $\overline{\boldsymbol{y}} \neq \mathbf{0}$, $\|\overline{\boldsymbol{y}}\| < \prod_{i=1}^{n} \widetilde{y}_i^{\alpha_i}$. Using (4.1), for all $\boldsymbol{A}$ such that $\boldsymbol{A}\mathcal{P}_{m,n}^{\alpha} = \mathcal{P}_{m,n}^{\alpha}$, we have $\boldsymbol{A}\boldsymbol{x} \neq \boldsymbol{y}$ because $\boldsymbol{B}\mathbf{0} = \mathbf{0} \neq \overline{\boldsymbol{y}}$ for all possible $\boldsymbol{B}$. Then by definition $\mathcal{P}_{m,n}^{\alpha}$ with $m \geq 1$, $n = 2$, and $\sum_{i=1}^{n} \alpha_i = 1$ or $m \geq 1$, $n = 2$, $\alpha_1 \neq \alpha_2$, and $\alpha_1 + \alpha_2 = 1$ is nonhomogeneous. By Theorem 4.2(ii), we have $\dim \operatorname{Lie} \operatorname{Aut}(\mathcal{P}_{m,n}^{\alpha}) = n + \frac{m(m-1)}{2} \geq m + n$ if and only if $m \geq 3$. The conclusion concerning perfectness now follows from this and Lemma 4.3(ii). □

Acknowledgment. We thank the referees for their comments, which helped to improve the paper.

REFERENCES

[1] G. P. BARKER, *Theory of cones*, Linear Algebra Appl., 39 (1981), pp. 263–291.
[2] H. H. BAUSCHKE, J. M. BORWEIN, AND W. LI, *Strong conical hull intersection property, bounded linear regularity, Jameson's property (G), and error bounds in convex optimization*, Math. Program., 86 (1999), pp. 135–160.

[3] J. BOLTE, T. P. NGUYEN, J. PEYPOUQUET, AND B. W. SUTER, *From error bounds to the complexity of first-order descent methods for convex functions*, Math. Program., 165 (2017), pp. 471–507.

[4] J. M. BORWEIN AND H. WOLKOWICZ, *Regularizing the abstract convex program*, J. Math. Anal. Appl., 83 (1981), pp. 495–530.

[5] R. CHARES, *Cones and Interior-point Algorithms for Structured Convex Optimization Involving Powers and Exponentials*, Ph.D. thesis, Université Catholique de Louvain, 2009.

[6] C. COEY, L. KAPELEVICH, AND J. P. VIELMA, *Solving natural conic formulations with Hypatia.jl*, INFORMS J. Comput., 34 (2022), pp. 2686–2699.

[7] J. FARAUT AND A. KORÁNYI, *Analysis on Symmetric Cones*, Oxford Math. Monogr., Clarendon Press, Oxford, 1994.

[8] L. FAYBUSOVICH, *Several Jordan-algebraic aspects of optimization*, Optimization, 57 (2008), pp. 379–393.

[9] M. S. GOWDA AND J. TAO, *On the bilinearity rank of a proper cone and Lyapunov-like transformations*, Math. Program., 147 (2014), pp. 155–170.

[10] M. S. GOWDA AND D. TROTT, *On the irreducibility, Lyapunov rank, and automorphisms of special Bishop Phelps cones*, J. Math. Anal. Appl., 419 (2014), pp. 172–184.

[11] B. C. HALL, *Lie Groups, Lie Algebras, and Representations*, 2nd ed., Grad. Texts in Math. 222, Springer, Cham, 2015.

[12] L. T. K. HIEN, *Differential properties of Euclidean projection onto power cone*, Math. Methods Oper. Res., 82 (2015), pp. 265–284.

[13] A. J. HOFFMAN, *On approximate solutions of systems of linear inequalities*, J. Res. Natl. Bur. Stand., 49 (1952), pp. 263–265.

[14] M. ITO AND B. F. LOURENÇO, *The automorphism group and the non-self-duality of p-cones*, J. Math. Anal. Appl., 471 (2019), pp. 392–410.

[15] M. KARIMI AND L. TUNÇEL, *Domain-Driven Solver (DDS) Version 2.1: A MATLAB-based software package for convex optimization problems in domain-driven form*, Math. Prog. Comp., 16 (2024), pp. 37–92.

[16] M. KOECHER, *Positivitätsbereiche im R^n*, Amer. J. Math., 79 (1957), pp. 575–596.

[17] J. M. LEE, *Introduction to Smooth Manifolds*, 2nd ed., Grad. Texts in Math. 218, Springer, New York, 2012.

[18] A. S. LEWIS AND J.-S. PANG, *Error bounds for convex inequality systems*, in Generalized Convexity, Generalized Monotonicity: Recent Results, Springer, 1998, pp. 75–110.

[19] S. B. LINDSTROM, B. F. LOURENÇO, AND T. K. PONG, *Error bounds, facial residual functions and applications to the exponential cone*, Math. Program., 200 (2023), pp. 229–278.

[20] S. B. LINDSTROM, B. F. LOURENÇO, AND T. K. PONG, *Optimal Error Bounds in the Absence of Constraint Qualifications with Applications to the p-Cones and Beyond*, preprint, https://arxiv.org/abs/2109.11729, 2021.

[21] R. LOEWY AND H. SCHNEIDER, *Indecomposable cones*, Linear Algebra Appl., 11 (1975), pp. 235–245.

[22] B. F. LOURENÇO, *Amenable cones: Error bounds without constraint qualifications*, Math. Program., 186 (2021), pp. 1–48.

[23] B. F. LOURENÇO, V. ROSHCHINA, AND J. SAUNDERSON, *Amenable cones are particularly nice*, SIAM J. Optim., 32 (2022), pp. 2347–2375, https://doi.org/10.1137/20M138466X.

[24] M. LUBIN, E. YAMANGIL, R. BENT, AND J. P. VIELMA, *Extended formulations in mixed-integer convex programming*, in International Conference on Integer Programming and Combinatorial Optimization, Q. Louveaux and M. Skutella, eds., Springer, Cham, 2016, pp. 102–113.

[25] Z.-Q. LUO AND J. F. STURM, *Error analysis*, in Handbook of Semidefinite Programming, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., Kluwer Academic Publishers, 2000.

[26] Z.-Q. LUO AND P. TSENG, *Error bounds and convergence analysis of feasible descent methods: A general approach*, Ann. Oper. Res., 46 (1993), pp. 157–178.

[27] MINLPLib, *A Library of Mixed-Integer and Continuous Nonlinear Programming Instances*, 2022, https://www.minlplib.org/.

[28] MOSEK ApS, *MOSEK Modeling Cookbook*, Release 3.3.0, 2022, https://docs.mosek.com/modeling-cookbook/index.html.

[29] S. Z. NÉMETH AND G. ZHANG, *Extended Lorentz cones and mixed complementarity problems*, J. Global Optim., 62 (2015), pp. 443–457.

[30] Y. NESTEROV, *Towards non-symmetric conic optimization*, Optim. Methods Softw., 27 (2012), pp. 893–917.

[31] M. ORLITZKY, *Tight bounds on Lyapunov rank*, Optim. Lett., 16 (2022), pp. 723–728.

[32] M. ORLITZKY AND M. S. GOWDA, *An improved bound for the Lyapunov rank of a proper cone*, Optim. Lett., 10 (2016), pp. 11–17.

[33] J.-S. PANG, *Error bounds in mathematical programming*, Math. Program., 79 (1997), pp. 299–332.

[34] D. PAPP AND S. YILDIZ, *Alfonso: Matlab package for nonsymmetric conic optimization*, INFORMS J. Comput., 34 (2022), pp. 11–19.

[35] G. PATAKI, *Strong duality in conic linear programming: Facial reduction and extended duals*, in Computational and Analytical Mathematics, Springer Proc. Math. Statist. 50, Springer, New York, 2013, pp. 613–634.

[36] S. ROY AND L. XIAO, *On self-concordant barriers for generalized power cones*, Optim. Lett., 16 (2022), pp. 681–694.

[37] A. SKAJAA AND Y. YE, *A homogeneous interior-point algorithm for nonsymmetric convex conic optimization*, Math. Program., 150 (2015), pp. 391–422.

[38] J. F. STURM, *Error bounds for linear matrix inequalities*, SIAM J. Optim., 10 (2000), pp. 1228–1248, https://doi.org/10.1137/S1052623498338606.

[39] C.-H. SUNG AND B.-S. TAM, *A study of projectionally exposed cones*, Linear Algebra Appl., 139 (1990), pp. 225–252.

[40] R. SZNAJDER, *The Lyapunov rank of extended second order cones*, J. Global Optim., 66 (2016), pp. 585–593.

[41] V. A. TRUONG AND L. TUNÇEL, *Geometry of homogeneous convex cones, duality mapping, and optimal self-concordant barriers*, Math. Program., 100 (2003), pp. 295–316.

[42] L. TUNÇEL AND S. XU, *On homogeneous convex cones, the Carathéodory number, and the duality mapping*, Math. Oper. Res., 26 (2001), pp. 234–247.

[43] L. TUNÇEL AND A. NEMIROVSKI, *Self-concordant barriers for convex approximations of structured convex sets*, Found. Comput. Math., 10 (2010), pp. 485–525.

[44] H. WAKI AND M. MURAMATSU, *Facial reduction algorithms for conic optimization problems*, J. Optim. Theory Appl., 158 (2013), pp. 188–215.

[45] P. YU, G. LI, AND T. K. PONG, *Kurdyka–Łojasiewicz exponent via inf-projection*, Found. Comput. Math., 22 (2022), pp. 1171–1217.

[46] Z. ZHOU AND A. M.-C. SO, *A unified approach to error bounds for structured convex optimization problems*, Math. Program., 165 (2017), pp. 689–728.