FULL LENGTH PAPER

Series A

Error bounds, facial residual functions and applications to the exponential cone

Scott B. Lindstrom[1] · Bruno F. Lourenço[2] · Ting Kei Pong[3]

Abstract

We construct a general framework for deriving error bounds for conic feasibility problems. In particular, our approach allows one to work with cones that fail to be amenable or even to have computable projections, two previously challenging barriers. For the purpose, we first show how error bounds may be constructed using objects called *one-step facial residual functions*. Then, we develop several tools to compute these facial residual functions even in the absence of closed form expressions for the projections onto the cones. We demonstrate the use and power of our results by computing tight error bounds for the exponential cone feasibility problem. Interestingly, we discover a natural example for which the tightest error bound is related to the Boltzmann–Shannon entropy. We were also able to produce an example of sets for which a Hölderian error bound holds but the supremum of the set of admissible exponents is not itself an admissible exponent.

Keywords Error bounds · Facial residual functions · Exponential cone

✉ Scott B. Lindstrom
   scott.lindstrom@curtin.edu.au

   Bruno F. Lourenço
   bruno@ism.ac.jp

   Ting Kei Pong
   tk.pong@polyu.edu.hk

[1] Centre for Optimisation and Decision Science, Curtin University, Bentley, Australia

[2] Department of Statistical Inference and Mathematics, Institute of Statistical Mathematics, Tokyo, Japan

[3] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, People's Republic of China

🖄 Springer

1 Introduction

Our main object of interest is the following convex conic feasibility problem:

$$\text{find} \quad \mathbf{x} \in (\mathcal{L} + \mathbf{a}) \cap \mathcal{K}, \qquad \text{(Feas)}$$

where $\mathcal{L}$ is a subspace contained in some finite-dimensional real Euclidean space $\mathcal{E}$, $\mathbf{a} \in \mathcal{E}$ and $\mathcal{K} \subseteq \mathcal{E}$ is a closed convex cone. For a discussion of some applications and algorithms for (Feas) see [22]. See also [5] for a broader analysis of convex feasibility problems. We also recall that a *conic linear program* (CLP) is the problem of minimizing a linear function subject to a constraint of the form described in (Feas). In addition, when the optimal set of a CLP is non-empty it can be written as the intersection of a cone with an affine set. This provides yet another motivation for analyzing (Feas): to better understand feasible regions and optimal sets of conic linear programs. Here, our main interest is in obtaining *error bounds* for (Feas). That is, assuming $(\mathcal{L} + \mathbf{a}) \cap \mathcal{K} \neq \emptyset$, we want an inequality that, given some arbitrary $\mathbf{x} \in \mathcal{E}$, relates the *individual distances* $d(\mathbf{x}, \mathcal{L} + \mathbf{a})$, $d(\mathbf{x}, \mathcal{K})$ to the *distance to the intersection* $d(\mathbf{x}, (\mathcal{L} + \mathbf{a}) \cap \mathcal{K})$. Considering that $\mathcal{E}$ is equipped with some norm $\| \cdot \|$ induced by some inner product $\langle \cdot, \cdot \rangle$, we recall that the *distance function to a convex set C* is defined as follows:

$$d(\mathbf{x}, C) := \inf_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|.$$

When $\mathcal{K}$ is a polyhedral cone, the classical Hoffman's error bound [23] gives a relatively complete picture of the way that the individual distances relate to the distance to the intersection. If $\mathcal{K}$ is not polyhedral, but $\mathcal{L} + \mathbf{a}$ intersects $\mathcal{K}$ in a sufficiently well-behaved fashion (say, for example, when $\mathcal{L} + \mathbf{a}$ intersects ri $\mathcal{K}$, the relative interior of $\mathcal{K}$; see Proposition 2.2), we may still expect "good" error bounds to hold, e.g., [6, Corollary 3]. However, checking whether $\mathcal{L} + \mathbf{a}$ intersects ri $\mathcal{K}$ is not necessarily a trivial task; and, in general, $(\mathcal{L} + \mathbf{a}) \cap$ ri $\mathcal{K}$ can be void.

Here, we focus on error bound results that *do not* require any assumption on the way that the affine space $\mathcal{L} + \mathbf{a}$ intersects $\mathcal{K}$. So, for example, we want results that are valid even if, say, $\mathcal{L} + \mathbf{a}$ fails to intersect the relative interior of $\mathcal{K}$. Inspired by Sturm's pioneering work on error bounds for positive semidefinite systems [50], the class of *amenable cones* was proposed in [34] and it was shown that the following three ingredients can be used to obtain general error bounds for (Feas): (i) amenable cones, (ii) facial reduction [13, 45, 52] and (iii) the so-called facial residual functions (FRFs) [34, Definition 16].

In this paper, we will show that, in fact, it is possible to obtain error bounds for (Feas) by using the so-called *one-step facial residual functions* directly in combination with facial reduction. It is fair to say that computing the facial residual functions is the

most critical step in obtaining error bounds for (Feas). We will demonstrate techniques that are readily adaptable for the purpose.

All the techniques discussed here will be showcased with error bounds for the so-called *exponential cone* which is defined as follows[1]:

$$K_{\exp} := \left\{ (x, y, z) \in \mathbb{R}^3 \mid y > 0, z \geq y e^{x/y} \right\} \cup \{ (x, y, z) \mid x \leq 0, z \geq 0, y = 0 \}.$$

Put succinctly, the exponential cone is the closure of the epigraph of the perspective function of $z = e^x$. It is quite useful in entropy optimization, see [15]. Furthermore, it is also implemented in the MOSEK package, see [17], [37, Chapter 5], and the many modelling examples in Sect. 5.4 therein. There are several other solvers that either support the exponential cone or convex sets closely related to it [16, 25, 39, 41]. See also [20] for an algorithm for projecting onto the exponential cone. So convex optimization with exponential cones is widely available even if, as of this writing, it is not as widespread as, say, semidefinite programming.

The exponential cone $K_{\exp}$ appears, at a glance, to be simple. However, it possesses a very intricate geometric structure that illustrates a wide range of challenges practitioners may face in computing error bounds. First of all, being non-facially-exposed, it is not amenable, so the theory developed in [34] does not directly apply to it. Another difficulty is that not many analytical tools have been developed to deal with the projection operator onto $K_{\exp}$ (compared with, for example, the projection operator onto PSD cones) which is only implicitly specified. Until now, these issues have made challenging the establishment of error bounds for objects like $K_{\exp}$, many of which are of growing interest in the mathematical programming community.

Our research is at the intersection of two topics: *error bounds* and the *facial structure of cones*. General information on the former can be found, for example, in [26, 40]. Classically, there seems to be a focus on the so-called *Hölderian error bounds* (see also [27–29]) but we will see in this paper that non Hölderian behavior can still appear even in relatively natural settings such as conic feasibility problems associated to the exponential cone.

Facts on the facial structure of convex cones can be found, for example, in [3, 4, 42]. We recall that a cone is said to be *facially exposed* if each face arises as the intersection of the whole cone with some supporting hyperplane. Stronger forms of facial exposedness have also been studied to some extent, here are some examples: projectional exposedness [13, 51], niceness [44, 46], tangential exposedness [47], amenability [34]. See also [36] for a comparison between a few different types of facial exposedness. These notions are useful in many topics, e.g.: regularization of convex programs and extended duals [13, 32, 45], studying the closure of certain linear images [32, 43], lifts of convex sets [21] and error bounds [34]. However, as can be seen in Fig. 1, the exponential cone is not even a facially exposed cone, so none of the aforementioned notions apply (in particular, the face $\mathcal{F}_{ne} := \{(0, 0, z) \mid z \geq 0\}$ is not exposed). This was one of the motivations for looking beyond facial exposedness and

---

[1] Our notation for the exponential cone coincides with the one in [39]. However, the x, y, z variables might appear permuted in other papers, e.g., [37, 41, 49].
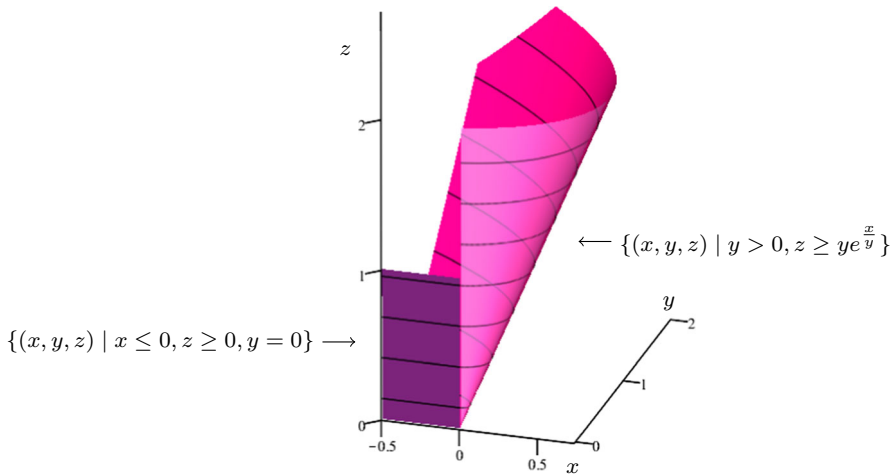
Fig. 1 The exponential cone is the union of the two labelled sets

In the figure: $\longleftarrow \{(x, y, z) \mid y > 0, z \geq ye^{\frac{x}{y}}\}$

$\{(x, y, z) \mid x \leq 0, z \geq 0, y = 0\} \longrightarrow$

developing a framework for deriving error bounds for feasibility problems associated to general closed convex cones.

1.1 Outline and results

The goal of this paper is to build a robust framework that may be used to obtain error bounds for previously inaccessible cones, and to demonstrate the use of this framework by applying it to fully describe error bounds for (Feas) with $\mathcal{K} = K_{\exp}$.

In Sect. 2, we recall preliminaries. New contributions begin in Sect. 3. We first recall some rules for chains of faces and the diamond composition. Then we show how error bounds may be constructed using objects known as one-step facial residual functions. In Sect. 3.1, we build our general framework for constructing one-step facial residual functions. Our key result, Theorem 3.10, obviates the need of computing explicitly the projection onto the cone. Instead, we make use of the parametrization of the boundary of the cone and projections onto the proper *faces* of a cone: thus, our approach is advantageous when these projections are easier to analyze than the projection onto the whole cone itself. We emphasize that *all* of the results of Sect. 3 are applicable to a *general closed convex cone*.

In Sect. 4, we use our new framework to fully describe error bounds for (Feas) with $\mathcal{K} = K_{\exp}$. This was previously a problem lacking a clear strategy, because all projections onto $K_{\exp}$ are implicitly specified. However, having obviated the need to project onto $K_{\exp}$, we successfully obtain all the necessary FRFs, partly because it is easier to project onto the *proper faces* of $K_{\exp}$ than to project onto $K_{\exp}$ itself. Surprisingly, we discover that different collections of faces and exposing hyperplanes admit very different FRFs. In Sect. 4.2.1, we show that for the unique 2-dimensional face, any exponent in (0, 1) may be used to build a valid FRF, while the supremum over all the admissible exponents *cannot* be. Furthermore, a better FRF for the 2D

face can be obtained if we go beyond Hölderian error bounds and consider a so-called *entropic error bound* which uses a modified Boltzmann-Shannon entropy function, see Theorem 4.2. The curious discoveries continue; for infinitely many 1-dimensional faces, the FRF, and the final error bound, feature exponent $1/2$. For the final outstanding 1-dimensional exposed face, the FRF, and the final error bound, are Lipschitzian for all exposing hyperplanes except exactly one, for which *no exponent* will suffice. However, for this exceptional case, our framework *still* successfully finds an FRF, which is logarithmic in character (Corollary 4.11). Consequentially, the system consisting of $\{(0, 0, 1)\}^{\perp}$ and $K_{\exp}$ possesses a kind of "logarithmic error bound" (see Example 4.20) instead of a Hölderian error bound. In Theorems 4.13 and 4.17, we give explicit error bounds by using our FRFs and the suite of tools we developed in Sect. 3. We also show that the error bound given in Theorem 4.13 is tight, see Remark 4.14.

These findings about the exponential cone are surprising, since we are not aware of other objects having this litany of odd behaviour hidden in their structure all at once.[2] One possible reason for the absence of previous reports on these phenomena might have been the sheer absence of tools for obtaining error bounds for general cones. In this sense, we believe that the machinery developed in Sect. 3 might be a reasonable first step towards filling this gap. In Sect. 4.4, we document additional odd consequences and connections to other concepts, with particular relevance to the *Kurdyka-Łojasiewicz (KL) property* [1, 2, 8–10, 30]. In particular, we have two sets satisfying a Hölderian error bound for every $\gamma \in (0, 1)$ but the supremum of allowable exponents is not allowable. Consequently, one obtains a function with the KL property with exponent α for any $\alpha \in (1/2, 1)$ at the origin, but not for $\alpha = 1/2$. We conclude in Sect. 5.

2 Preliminaries

We recall that $\mathcal{E}$ denotes an arbitrary finite-dimensional real Euclidean space. We will adopt the following convention, vectors will be boldfaced while scalars will use normal typeface. For example, if $\mathbf{p} \in \mathbb{R}^3$, we write $\mathbf{p} = (p_x, p_y, p_z)$, where $p_x, p_y, p_z \in \mathbb{R}$.

We denote by $B(\eta)$ the closed ball of radius η centered at the origin, i.e., $B(\eta) = \{\mathbf{x} \in \mathcal{E} \mid \|\mathbf{x}\| \leq \eta\}$. Let $C \subseteq \mathcal{E}$ be a convex set. We denote the relative interior and the linear span of C by ri C and span C, respectively. We also denote the boundary of C by ∂C, and cl C is the closure of C. We denote the projection operator onto C by P_C, so that $P_C(\mathbf{x}) = \operatorname{argmin}_{\mathbf{y} \in C} \|\mathbf{x} - \mathbf{y}\|$. Given closed convex sets $C_1, C_2 \subseteq \mathcal{E}$, we note the following properties of the projection operator

$$\mathrm{d}(\mathbf{x}, C_1) \leq \mathrm{d}(\mathbf{x}, C_2) + \mathrm{d}(P_{C_2}(\mathbf{x}), C_1) \tag{2.1}$$

$$\mathrm{d}(P_{C_2}(\mathbf{x}), C_1) \leq \mathrm{d}(\mathbf{x}, C_2) + \mathrm{d}(\mathbf{x}, C_1). \tag{2.2}$$

---

[2] To be fair, the exponential function is the classical example of a non-semialgebraic analytic function. Given that semialgebraicity is connected to the KL property (which is related to error bounds), one may argue that it is not *that* surprising that the exponential cone has its share of quirks. Nevertheless, given how natural the exponential cone is, the amount of quirks is still somewhat surprising.

2.1 Cones and their faces

Let $\mathcal{K}$ be a closed convex cone. We say that $\mathcal{K}$ is *pointed* if $\mathcal{K} \cap -\mathcal{K} = \{\mathbf{0}\}$. The dimension of $\mathcal{K}$ is denoted by $\dim(\mathcal{K})$ and is the dimension of the linear subspace spanned by $\mathcal{K}$. A *face* of $\mathcal{K}$ is a closed convex cone $\mathcal{F}$ satisfying $\mathcal{F} \subseteq \mathcal{K}$ and the following property

$$\mathbf{x}, \mathbf{y} \in \mathcal{K}, \mathbf{x} + \mathbf{y} \in \mathcal{F} \Rightarrow \mathbf{x}, \mathbf{y} \in \mathcal{F}.$$

In this case, we write $\mathcal{F} \trianglelefteq \mathcal{K}$. We say that $\mathcal{F}$ is *proper* if $\mathcal{F} \neq \mathcal{K}$. A face is said to be *nontrivial* if $\mathcal{F} \neq \mathcal{K}$ and $\mathcal{F} \neq \mathcal{K} \cap -\mathcal{K}$. In particular, if $\mathcal{K}$ is pointed (as is the case of the exponential cone), a nontrivial face is neither $\mathcal{K}$ nor $\{\mathbf{0}\}$. Next, let $\mathcal{K}^*$ denote the dual cone of $\mathcal{K}$, i.e., $\mathcal{K}^* = \{\mathbf{z} \in \mathcal{E} \mid \langle \mathbf{x}, \mathbf{z} \rangle \geq 0, \forall \mathbf{x} \in \mathcal{K}\}$. We say that $\mathcal{F}$ is an *exposed face* if there exists $\mathbf{z} \in \mathcal{K}^*$ such that $\mathcal{F} = \mathcal{K} \cap \{\mathbf{z}\}^{\perp}$.

A *chain of faces* of $\mathcal{K}$ is a sequence of faces satisfying $\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1$ such that each $\mathcal{F}_i$ is a face of $\mathcal{K}$ and the inclusions $\mathcal{F}_{i+1} \subsetneq \mathcal{F}_i$ are all proper. The length of the chain is defined to be ℓ. With that, we define the *distance to polyhedrality of* $\mathcal{K}$ as the length *minus one* of the longest chain of faces of $\mathcal{K}$ such that $\mathcal{F}_\ell$ is polyhedral and $\mathcal{F}_i$ is not polyhedral for $i < \ell$, see [35, Sect. 5.1]. We denote the distance to polyhedrality by $\ell_{\mathrm{poly}}(\mathcal{K})$.

2.2 Lipschitzian and Hölderian error bounds

In this subsection, suppose that $C_1, \ldots, C_\ell \subseteq \mathcal{E}$ are convex sets with nonempty intersection. We recall the following definitions.

Definition 2.1 (*Hölderian and Lipschitzian error bounds*) The sets $C_1, \ldots, C_\ell$ are said to satisfy a *Hölderian error bound* if for every bounded set $B \subseteq \mathcal{E}$ there exist some $\kappa_B > 0$ and an exponent $\gamma_B \in (0, 1]$ such that

$$\mathrm{d}(\mathbf{x}, \cap_{i=1}^{\ell} C_i) \leq \kappa_B \max_{1 \leq i \leq \ell} \mathrm{d}(\mathbf{x}, C_i)^{\gamma_B}, \qquad \forall \, \mathbf{x} \in B.$$

If we can take the same $\gamma_B = \gamma \in (0, 1]$ for all B, then we say that the bound is *uniform*. If the bound is uniform with $\gamma = 1$, we call it a *Lipschitzian error bound*.

We note that the concepts in Definition 2.1 also have different names throughout the literature. When $C_1, \ldots, C_\ell$ satisfy a Hölderian error bound it is said that they satisfy *bounded Hölder regularity*, e.g., see [11, Definition 2.2]. When a Lipschitzian error bound holds, $C_1, \ldots, C_\ell$ are said to satisfy *bounded linear regularity*, see [5, Sect. 5] or [6]. Bounded linear regularity is also closely related to the notion of *subtransversality* [24, Definition 7.5].

Hölderian and Lipschitzian error bounds will appear frequently in our results, but we also encounter non-Hölderian bounds as in Theorems 4.2 and 4.10. Next, we recall the following result which ensures a Lipschitzian error bound holds between families of convex sets when a constraint qualification is satisfied.

Proposition 2.2 [6, Corollary 3] *Let $C_1, \ldots, C_\ell \subseteq \mathcal{E}$ be convex sets such that $C_1, \ldots, C_k$ are polyhedral. If*

$$\left(\bigcap_{i=1}^{k} C_i \right) \cap \left(\bigcap_{j=k+1}^{\ell} ri\, C_j \right) \neq \emptyset,$$

then for every bounded set B there exists $\kappa_B > 0$ such that

$$d\left(\boldsymbol{x}, \cap_{i=1}^{\ell} C_i \right) \leq \kappa_B \left(\max_{1 \leq i \leq \ell} d(\boldsymbol{x}, C_i) \right), \quad \forall \boldsymbol{x} \in B.$$

In view of (Feas), we say that *Slater's condition* is satisfied if $(\mathcal{L} + \mathbf{a}) \cap ri\, \mathcal{K} \neq \emptyset$. If $\mathcal{K}$ can be written as $\mathcal{K}^1 \times \mathcal{K}^2 \subseteq \mathcal{E}^1 \times \mathcal{E}^2$, where $\mathcal{E}^1$ and $\mathcal{E}^2$ are real Euclidean spaces and $\mathcal{K}^1 \subseteq \mathcal{E}^1$ is polyhedral, we say that the *partial polyhedral Slater's (PPS) condition* is satisfied if

$$(\mathcal{L} + \mathbf{a}) \cap (\mathcal{K}^1 \times (ri\, \mathcal{K}^2)) \neq \emptyset. \tag{2.3}$$

Adding a dummy coordinate, if necessary, we can see Slater's condition as a particular case of the PPS condition. By convention, we consider that the PPS condition is satisfied for (Feas) if one of the following is satisfied: 1) $\mathcal{L} + \mathbf{a}$ intersects $ri\, \mathcal{K}$; 2) $(\mathcal{L} + \mathbf{a}) \cap \mathcal{K} \neq \emptyset$ and $\mathcal{K}$ is polyhedral; or 3) $\mathcal{K}$ can be written as a direct product $\mathcal{K}^1 \times \mathcal{K}^2$ where $\mathcal{K}^1$ is polyhedral and (2.3) is satisfied.

Noting that $(\mathcal{L} + \mathbf{a}) \cap (\mathcal{K}^1 \times (ri\, \mathcal{K}^2)) = (\mathcal{L} + \mathbf{a}) \cap (\mathcal{K}^1 \times \mathcal{E}^2) \cap (\mathcal{E}^1 \times (ri\, \mathcal{K}^2))$, we deduce the following result from Proposition 2.2.

Proposition 2.3 (Error bound under PPS condition) *Suppose that* (Feas) *satisfies the partial polyhedral Slater's condition. Then, for every bounded set B there exists $\kappa_B > 0$ such that*

$$d(\boldsymbol{x}, (\mathcal{L} + \boldsymbol{a}) \cap \mathcal{K}) \leq \kappa_B \max\{d(\boldsymbol{x}, \mathcal{K}), d(\boldsymbol{x}, \mathcal{L} + \boldsymbol{a})\}, \quad \forall \boldsymbol{x} \in B.$$

We recall that for $a, b \in \mathbb{R}_+$ we have $a + b \leq 2 \max\{a, b\} \leq 2(a + b)$, so Propositions 2.2 and 2.3 can also be equivalently stated in terms of sums of distances.

3 Facial residual functions and error bounds

In this section, we discuss a strategy for obtaining error bounds for the conic linear system (Feas) based on the so-called *facial residual functions* that were introduced in [34]. In contrast to [34], we will not require that $\mathcal{K}$ be amenable.

The motivation for our approach is as follows. If it were the case that (Feas) satisfies some constraint qualification, we would have a Lipschitzian error bound per Proposition 2.3, see also [6] for other sufficient conditions. Unfortunately, this does not happen in general. However, as long as (Feas) is feasible, there is always a face

of $\mathcal{K}$ that contains the feasible region of (Feas) and for which a constraint qualification holds. The error bound computation essentially boils down to understanding how to compute the distance to this special face. The first result towards our goal is the following.

Proposition 3.1 (An error bound when a face satisfying a CQ is known) *Suppose that* (Feas) *is feasible and let* $\mathcal{F} \trianglelefteq \mathcal{K}$ *be a face such that*

(a) $\mathcal{F}$ *contains* $\mathcal{K} \cap (\mathcal{L} + \boldsymbol{a})$.
(b) $\{\mathcal{F}, \mathcal{L} + \boldsymbol{a}\}$ *satisfies the PPS condition.*[3]

Then, for every bounded set B, there exists $\kappa_B > 0$ *such that*

$$d(\boldsymbol{x}, \mathcal{K} \cap (\mathcal{L} + \boldsymbol{a})) \leq \kappa_B (d(\boldsymbol{x}, \mathcal{F}) + d(\boldsymbol{x}, \mathcal{L} + \boldsymbol{a})), \qquad \forall \boldsymbol{x} \in B.$$

Proof Since $\mathcal{F}$ is a face of $\mathcal{K}$, assumption (a) implies $\mathcal{K} \cap (\mathcal{L} + \mathbf{a}) = \mathcal{F} \cap (\mathcal{L} + \mathbf{a})$. Then, the result follows from assumption (b) and Proposition 2.3. □

From Proposition 3.1 we see that the key to obtaining an error bound for the system (Feas) is to find a face $\mathcal{F} \trianglelefteq \mathcal{K}$ satisfying (a), (b) *and* we must know how to estimate the quantity $d(\mathbf{x}, \mathcal{F})$ from the available information $d(\mathbf{x}, \mathcal{K})$ and $d(\mathbf{x}, \mathcal{L} + \mathbf{a})$.

This is where we will make use of facial reduction and facial residual functions. The former will help us find $\mathcal{F}$ and the latter will be instrumental in upper bounding $d(\mathbf{x}, \mathcal{F})$. First, we recall below a result that follows from the analysis of the FRA-poly facial reduction algorithm developed in [35].

Proposition 3.2 [34, Proposition 5][4] *Let* $\mathcal{K} = \mathcal{K}^1 \times \cdots \times \mathcal{K}^s$, *where each* $\mathcal{K}^i$ *is a closed convex cone. Suppose* (Feas) *is feasible. Then there is a chain of faces*

$$\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1 = \mathcal{K} \tag{3.1}$$

of length ℓ *and vectors* $\{z_1, \ldots, z_{\ell-1}\}$ *satisfying the following properties.*

(i) $\ell - 1 \leq \sum_{i=1}^{s} \ell_{poly}(\mathcal{K}^i) \leq \dim \mathcal{K}$.
(ii) For all $i \in \{1, \ldots, \ell - 1\}$, *we have*

$$z_i \in \mathcal{F}_i^* \cap \mathcal{L}^\perp \cap \{\boldsymbol{a}\}^\perp \quad and \quad \mathcal{F}_{i+1} = \mathcal{F}_i \cap \{z_i\}^\perp.$$

(iii) $\mathcal{F}_\ell \cap (\mathcal{L} + \boldsymbol{a}) = \mathcal{K} \cap (\mathcal{L} + \boldsymbol{a})$ *and* $\{\mathcal{F}_\ell, \mathcal{L} + \boldsymbol{a}\}$ *satisfies the PPS condition.*

In view of Proposition 3.2, we define the *distance to the PPS condition* $d_{\text{PPS}}(\mathcal{K}, \mathcal{L} + \mathbf{a})$ as the length *minus one* of the shortest chain of faces (as in (3.1)) satisfying items (ii) and (iii) in Proposition 3.2. For example, if (Feas) satisfies the PPS condition, we have $d_{\text{PPS}}(\mathcal{K}, \mathcal{L} + \mathbf{a}) = 0$.

Next, we recall the definition of facial residual functions from [34, Definition 16].

---

[3] As a reminder, the PPS condition is, by convention, a shorthand for three closely related conditions, see remarks after (2.3).

[4] Although [34, Proposition 5] was originally stated for pointed cones, it holds for general closed convex cones. Indeed, its proof only relies on [35, Proposition 8], which holds for general closed convex cones.

Definition 3.3 (Facial residual function[5]) Let $\mathcal{K}$ be a closed convex cone, $\mathcal{F} \unlhd \mathcal{K}$ be a face, and let $\mathbf{z} \in \mathcal{F}^*$. Suppose that $\psi_{\mathcal{F},\mathbf{z}} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies the following properties:

(i) $\psi_{\mathcal{F},\mathbf{z}}$ is nonnegative, monotone nondecreasing in each argument and $\psi_{\mathcal{F},\mathbf{z}}(0, t) = 0$ for every $t \in \mathbb{R}_+$.
(ii) The following implication holds for any $\mathbf{x} \in \operatorname{span} \mathcal{K}$ and any $\epsilon \geq 0$:

$$d(\mathbf{x}, \mathcal{K}) \leq \epsilon, \quad \langle \mathbf{x}, \mathbf{z} \rangle \leq \epsilon, \quad d(\mathbf{x}, \operatorname{span} \mathcal{F}) \leq \epsilon \quad \Rightarrow \quad d(\mathbf{x}, \mathcal{F} \cap \{\mathbf{z}\}^{\perp}) \leq \psi_{\mathcal{F},\mathbf{z}}(\epsilon, \|\mathbf{x}\|).$$

Then, $\psi_{\mathcal{F},\mathbf{z}}$ is said to be *a facial residual function for $\mathcal{F}$ and $\mathbf{z}$ with respect to $\mathcal{K}$*.

Definition 3.3, in its most general form, represents "two-steps" along the facial structure of a cone: we have a cone $\mathcal{K}$, a face $\mathcal{F}$ (which could be different from $\mathcal{K}$) and a third face defined by $\mathcal{F} \cap \{\mathbf{z}\}^{\perp}$. In this work, however, we will be focused on the following special case of Definition 3.3.

Definition 3.4 (One-step facial residual function (1-FRF)) Let $\mathcal{K}$ be a closed convex cone and $\mathbf{z} \in \mathcal{K}^*$. A function $\psi_{\mathcal{K},\mathbf{z}} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is called a one-step facial residual function (1-FRF) for $\mathcal{K}$ and $\mathbf{z}$ if it is a facial residual function of $\mathcal{K}$ and $\mathbf{z}$ with respect to $\mathcal{K}$. That is, $\psi_{\mathcal{K},\mathbf{z}}$ satisfies item (i) of Definition 3.3 and for every $\mathbf{x} \in \operatorname{span} \mathcal{K}$ and any $\epsilon \geq 0$:

$$d(\mathbf{x}, \mathcal{K}) \leq \epsilon, \quad \langle \mathbf{x}, \mathbf{z} \rangle \leq \epsilon \quad \Rightarrow \quad d(\mathbf{x}, \mathcal{K} \cap \{\mathbf{z}\}^{\perp}) \leq \psi_{\mathcal{K},\mathbf{z}}(\epsilon, \|\mathbf{x}\|).$$

Remark 3.5 (Concerning the implication in Definition 3.4) In view of the monotonicity of $\psi_{\mathcal{K},\mathbf{z}}$, the implication in Definition 3.4 can be equivalently and more succinctly written as

$$d(\mathbf{x}, \mathcal{K} \cap \{\mathbf{z}\}^{\perp}) \leq \psi_{\mathcal{K},\mathbf{z}}(\max\{d(\mathbf{x}, \mathcal{K}), \langle \mathbf{x}, \mathbf{z} \rangle\}, \|\mathbf{x}\|), \quad \forall \mathbf{x} \in \operatorname{span} \mathcal{K}.$$

The *unfolded* form presented in Definition 3.4 is more handy in our discussions and analysis below.

Facial residual functions always exist (see [34, Sect. 3.2] for the case of pointed cones, although the argument holds in general), but their computation is often nontrivial. Next, we review a few examples.

Example 3.6 (Examples of facial residual functions) If $\mathcal{K}$ is a symmetric cone (i.e., a self-dual homogeneous cone, see [18, 19]), then given $\mathcal{F} \unlhd \mathcal{K}$ and $\mathbf{z} \in \mathcal{F}^*$, there exists a $\kappa > 0$ such that $\psi_{\mathcal{F},\mathbf{z}}(\epsilon, t) := \kappa\epsilon + \kappa\sqrt{\epsilon t}$ is a one-step facial residual function for $\mathcal{F}$ and $\mathbf{z}$, see [34, Theorem 35].

If $\mathcal{K}$ is a polyhedral cone, the function $\psi_{\mathcal{F},\mathbf{z}}(\epsilon, t) := \kappa\epsilon$ can be taken instead, with no dependency on t, see [34, Proposition 18].

---

[5] The only difference between Definition 3.3 and the definition of facial residual functions in [34, Definition 16] is that we added the "with respect to $\mathcal{K}$" part, to emphasize the dependency on $\mathcal{K}$.

Moving on, we say that a function $\tilde{\psi}_{\mathcal{F},\mathbf{z}}$ is a *positively rescaled shift of* $\psi_{\mathcal{F},\mathbf{z}}$ if there are positive constants M_1, M_2, M_3 and nonnegative constant M_4 such that

$$\tilde{\psi}_{\mathcal{F},\mathbf{z}}(\epsilon, t) = M_3 \psi_{\mathcal{F},\mathbf{z}}(M_1\epsilon, M_2 t) + M_4\epsilon. \tag{3.2}$$

This is a generalization of the notion of positive rescaling in [34], which sets $M_4 = 0$. We also need to compose facial residual functions in a special manner. Let $f : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ and $g : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ be functions. We define the *diamond composition* $f \diamond g$ to be the function satisfying

$$(f \diamond g)(a, b) = f(a + g(a, b), b), \qquad \forall a, b \in \mathbb{R}_+. \tag{3.3}$$

Note that the above composition is not associative in general. When we have functions $f_i : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+, i = 1, \ldots, m$ with $m \geq 3$, we define $f_m \diamond \cdots \diamond f_1$ inductively as the function φ_m such that

$$\varphi_i := f_i \diamond \varphi_{i-1}, \qquad i \in \{2, \ldots, m\}$$
$$\varphi_1 := f_1.$$

With that, we have $f_m \diamond f_{m-1} \diamond \cdots \diamond f_2 \diamond f_1 := f_m \diamond (f_{m-1} \diamond (\cdots \diamond (f_2 \diamond f_1)))$.

The following lemma, which holds for a general closed convex cone $\mathcal{K}$, shows how (positively rescaled shifts of) one-step facial residual functions for the *faces of* $\mathcal{K}$ can be combined via the diamond composition to derive useful bounds on the distance to faces. A version of it was proved in [34, Lemma 22], which required the cones to be pointed and made use of general (i.e., not necessarily one-step) facial residual functions with respect to $\mathcal{K}$. This is a subtle, but very crucial difference which will allows us to relax the assumptions in [34].

Lemma 3.7 (Diamond composing facial residual functions) *Suppose* (Feas) *is feasible and let*

$$\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1 = \mathcal{K}$$

be a chain of faces of $\mathcal{K}$ together with $z_i \in \mathcal{F}_i^* \cap \mathcal{L}^\perp \cap \{a\}^\perp$ such that $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{z_i\}^\perp$, for $i = 1, \ldots, \ell - 1$. For each i, let ψ_i be a $\mathbb{1}$-FRF for $\mathcal{F}_i$ and z_i. Then, there is a positively rescaled shift of ψ_i (still denoted as ψ_i by an abuse of notation) so that for every $x \in \mathcal{E}$ and $\epsilon \geq 0$:

$$d(x, \mathcal{K}) \leq \epsilon, \quad d(x, \mathcal{L} + a) \leq \epsilon \quad \Rightarrow \quad d(x, \mathcal{F}_\ell) \leq \varphi(\epsilon, \|x\|),$$

where $\varphi = \psi_{\ell-1} \diamond \cdots \diamond \psi_1$, if $\ell \geq 2$. If $\ell = 1$, we let φ be the function satisfying $\varphi(\epsilon, t) = \epsilon$.

Proof For $\ell = 1$, we have $\mathcal{F}_\ell = \mathcal{K}$, so the lemma follows immediately. Now, we consider the case $\ell \geq 2$. First we note that $\mathcal{L} + a$ is contained in all the $\{z_i\}^\perp$ for

$i = 1, \ldots, \ell - 1$. Since the distance of $\mathbf{x} \in \mathcal{E}$ to $\{\mathbf{z}_i\}^\perp$ is given by $\frac{|\langle \mathbf{x}, \mathbf{z}_i \rangle|}{\|\mathbf{z}_i\|}$, we have the following chain of implications

$$d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq \epsilon \quad \Rightarrow \quad d(\mathbf{x}, \{\mathbf{z}_i\}^\perp) \leq \epsilon \quad \Rightarrow \quad \langle \mathbf{x}, \mathbf{z}_i \rangle \leq \epsilon \|\mathbf{z}_i\|. \tag{3.4}$$

Next, we proceed by induction. If $\ell = 2$, we have that ψ_1 is a one-step facial residual function for $\mathcal{K}$ and $\mathbf{z}_1$. By Definition 3.4, we have

$$\mathbf{y} \in \operatorname{span} \mathcal{K}, \quad d(\mathbf{y}, \mathcal{K}) \leq \epsilon, \quad \langle \mathbf{y}, \mathbf{z}_1 \rangle \leq \epsilon \quad \Rightarrow \quad d(\mathbf{y}, \mathcal{F}_2) \leq \psi_1(\epsilon, \|\mathbf{y}\|).$$

In view of (3.4) and the monotonicity of ψ_1, we see further that

$$\mathbf{y} \in \operatorname{span} \mathcal{K}, \ d(\mathbf{y}, \mathcal{K}) \leq \epsilon, \ d(\mathbf{y}, \mathcal{L} + \mathbf{a}) \leq \epsilon \Rightarrow d(\mathbf{y}, \mathcal{F}_2) \leq \psi_1(\epsilon(1 + \|\mathbf{z}_1\|), \|\mathbf{y}\|). \tag{3.5}$$

Now, suppose that $\mathbf{x} \in \mathcal{E}$ and $\epsilon \geq 0$ are such that $d(\mathbf{x}, \mathcal{K}) \leq \epsilon$ and $d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq \epsilon$. Let $\hat{\mathbf{x}} := P_{\operatorname{span} \mathcal{K}}(\mathbf{x})$. Since $\mathcal{K} \subseteq \operatorname{span} \mathcal{K}$, we have $d(\mathbf{x}, \operatorname{span} \mathcal{K}) \leq d(\mathbf{x}, \mathcal{K})$ and, in view of (2.2), we have that

$$d(\hat{\mathbf{x}}, \mathcal{K}) \leq d(\mathbf{x}, \operatorname{span} \mathcal{K}) + d(\mathbf{x}, \mathcal{K}) \leq 2\epsilon, \tag{3.6}$$
$$d(\hat{\mathbf{x}}, \mathcal{L} + \mathbf{a}) \leq d(\mathbf{x}, \operatorname{span} \mathcal{K}) + d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq 2\epsilon.$$

From (2.1), (3.5) and (3.6) we obtain

$$d(\mathbf{x}, \mathcal{F}_2) \leq d(\mathbf{x}, \operatorname{span} \mathcal{K}) + d(\hat{\mathbf{x}}, \mathcal{F}_2) \leq \epsilon + \psi_1(2\epsilon(1 + \|\mathbf{z}_1\|), \|\hat{\mathbf{x}}\|)$$
$$\leq \epsilon + \psi_1(2\epsilon(1 + \|\mathbf{z}_1\|), \|\mathbf{x}\|),$$

where the last inequality follows from the monotonicity of ψ_1 and the fact that $\|\hat{\mathbf{x}}\| \leq \|\mathbf{x}\|$. This proves the lemma for chains of length $\ell = 2$ because the function mapping (ϵ, t) to $\epsilon + \psi_1(2\epsilon(1 + \|\mathbf{z}_1\|), t)$ is a positively rescaled shift of ψ_1.

Now, suppose that the lemma holds for chains of length $\hat{\ell}$ and consider a chain of length $\hat{\ell} + 1$. By the induction hypothesis, we have

$$d(\mathbf{x}, \mathcal{K}) \leq \epsilon, \quad d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq \epsilon \quad \Rightarrow \quad d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}) \leq \varphi(\epsilon, \|\mathbf{x}\|), \tag{3.7}$$

where $\varphi = \psi_{\hat{\ell}-1} \diamondsuit \cdots \diamondsuit \psi_1$ and the ψ_i are (positively rescaled shifts of) one-step facial residual functions. By the definition of $\psi_{\hat{\ell}}$ as a one-step facial residual function and using (3.4), we may positively rescale $\psi_{\hat{\ell}}$ (still denoted as $\psi_{\hat{\ell}}$ by an abuse of notation) so that for $\mathbf{y} \in \operatorname{span} \mathcal{F}_{\hat{\ell}}$ and $\hat{\epsilon} \geq 0$, the following implication holds:

$$d(\mathbf{y}, \mathcal{F}_{\hat{\ell}}) \leq \hat{\epsilon}, \quad d(\mathbf{y}, \mathcal{L} + \mathbf{a}) \leq \hat{\epsilon} \quad \Rightarrow \quad d(\mathbf{y}, \mathcal{F}_{\hat{\ell}+1}) \leq \psi_{\hat{\ell}}(\hat{\epsilon}, \|\mathbf{y}\|). \tag{3.8}$$

Now, suppose that $\mathbf{x} \in \mathcal{E}$ and $\epsilon \geq 0$ satisfy $d(\mathbf{x}, \mathcal{K}) \leq \epsilon$ and $d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq \epsilon$. Let $\hat{\mathbf{x}} := P_{\operatorname{span} \mathcal{F}_{\hat{\ell}}}(\mathbf{x})$. As before, since $\mathcal{F}_{\hat{\ell}} \subseteq \operatorname{span} \mathcal{F}_{\hat{\ell}}$, we have $d(\mathbf{x}, \operatorname{span} \mathcal{F}_{\hat{\ell}}) \leq d(\mathbf{x}, \mathcal{F}_{\hat{\ell}})$

and, in view of (2.2), we have

$$d(\hat{\mathbf{x}}, \mathcal{F}_{\hat{\ell}}) \leq d(\mathbf{x}, \text{span } \mathcal{F}_{\hat{\ell}}) + d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}) \leq 2d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}) \leq 2d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}) + \epsilon,$$
$$d(\hat{\mathbf{x}}, \mathcal{L} + \mathbf{a}) \leq d(\mathbf{x}, \text{span } \mathcal{F}_{\hat{\ell}}) + d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}) + \epsilon \leq 2d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}) + \epsilon.$$
$$(3.9)$$

Let $\hat{\psi}_{\hat{\ell}}$ be such that $\hat{\psi}_{\hat{\ell}}(s, t) := s + \psi_{\hat{\ell}}(2s, t)$, so that $\hat{\psi}_{\hat{\ell}}$ is a positively rescaled shift of $\psi_{\hat{\ell}}$. Then, (3.9) together with (3.8) and (2.1) gives

$$d(\mathbf{x}, \mathcal{F}_{\hat{\ell}+1}) \leq d(\mathbf{x}, \text{span } \mathcal{F}_{\hat{\ell}}) + d(\hat{\mathbf{x}}, \mathcal{F}_{\hat{\ell}+1}) \leq d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}) + d(\hat{\mathbf{x}}, \mathcal{F}_{\hat{\ell}+1})$$

$$\leq d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}) + \psi_{\hat{\ell}}(\epsilon + 2d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}), \|\hat{\mathbf{x}}\|) \overset{(a)}{\leq} d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}) + \psi_{\hat{\ell}}(2\epsilon + 2d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}), \|\mathbf{x}\|)$$

$$\leq \hat{\psi}_{\hat{\ell}}(\epsilon + d(\mathbf{x}, \mathcal{F}_{\hat{\ell}}), \|\mathbf{x}\|) \overset{(b)}{\leq} \hat{\psi}_{\hat{\ell}}(\epsilon + \varphi(\epsilon, \|\mathbf{x}\|), \|\mathbf{x}\|) = (\hat{\psi}_{\hat{\ell}} \diamond \varphi)(\epsilon, \|\mathbf{x}\|),$$

where (a) follows from the monotonicity of $\psi_{\hat{\ell}}$ and the fact that $\|\hat{\mathbf{x}}\| \leq \|\mathbf{x}\|$, and (b) follows from (3.7) and the monotonicity of $\hat{\psi}_{\hat{\ell}}$. This completes the proof. $\qquad\square$

We now have all the pieces to state an error bound result for (Feas) that does not require any constraint qualifications.

Theorem 3.8 (Error bound based on $\mathbb{1}$-*FRF*s) *Suppose* (Feas) *is feasible and let*

$$\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1 = \mathcal{K}$$

be a chain of faces of $\mathcal{K}$ together with $z_i \in \mathcal{F}_i^* \cap \mathcal{L}^\perp \cap \{\mathbf{a}\}^\perp$ such that $\{\mathcal{F}_\ell, \mathcal{L} + \mathbf{a}\}$ satisfies the PPS condition and $\mathcal{F}_{i+1} = \mathcal{F}_i \cap \{z_i\}^\perp$ for every i. For $i = 1, \ldots, \ell - 1$, let ψ_i be a $\mathbb{1}$-FRF for $\mathcal{F}_i$ and z_i.

Then, there is a suitable positively rescaled shift of the ψ_i (still denoted as ψ_i by an abuse of notation) such that for any bounded set B there is a positive constant κ_B (depending on $B, \mathcal{L}, \mathbf{a}, \mathcal{F}_\ell$) such that

$$\mathbf{x} \in B, d(\mathbf{x}, \mathcal{K}) \leq \epsilon, d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq \epsilon \Rightarrow d(\mathbf{x}, (\mathcal{L} + \mathbf{a}) \cap \mathcal{K}) \leq \kappa_B(\epsilon + \varphi(\epsilon, M)),$$

where $M = \sup_{\mathbf{x} \in B} \|\mathbf{x}\|$, $\varphi = \psi_{\ell-1} \diamond \cdots \diamond \psi_1$, if $\ell \geq 2$. If $\ell = 1$, we let φ be the function satisfying $\varphi(\epsilon, M) = \epsilon$.

Proof The case $\ell = 1$ follows from Proposition 3.1, by taking $\mathcal{F} = \mathcal{F}_1$. Now, suppose $\ell \geq 2$. We apply Lemma 3.7, which tells us that, after positively rescaling and shifting the ψ_i, we have:

$$d(\mathbf{x}, \mathcal{K}) \leq \epsilon, \quad d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq \epsilon \implies d(\mathbf{x}, \mathcal{F}_\ell) \leq \varphi(\epsilon, \|\mathbf{x}\|),$$

where $\varphi = \psi_{\ell-1} \diamond \cdots \diamond \psi_1$. In particular, since $\|\mathbf{x}\| \leq M$ for $\mathbf{x} \in B$ we have

$$d(\mathbf{x}, \mathcal{K}) \leq \epsilon, \quad d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq \epsilon \implies d(\mathbf{x}, \mathcal{F}_\ell) \leq \varphi(\epsilon, M), \qquad \forall \mathbf{x} \in B \ (3.10)$$

By assumption, $\{\mathcal{F}_\ell, \mathcal{L} + \mathbf{a}\}$ satisfies the PPS condition. We invoke Proposition 3.1 to find $\kappa_B > 0$ such that

$$d(\mathbf{x}, \mathcal{K} \cap (\mathcal{L} + \mathbf{a})) \leq \kappa_B (d(\mathbf{x}, \mathcal{F}_\ell) + d(\mathbf{x}, \mathcal{L} + \mathbf{a})), \qquad \forall \mathbf{x} \in B. \qquad (3.11)$$

Combining (3.10), (3.11), we conclude that if $\mathbf{x} \in B$ and $\epsilon \geq 0$ satisfy $d(\mathbf{x}, \mathcal{K}) \leq \epsilon$ and $d(\mathbf{x}, \mathcal{L} + \mathbf{a}) \leq \epsilon$, then we have $d(\mathbf{x}, (\mathcal{L} + \mathbf{a}) \cap \mathcal{K}) \leq \kappa_B(\epsilon + \varphi(\epsilon, M))$. This completes the proof. $\qquad \square$

Theorem 3.8 is an improvement over [34, Theorem 23] because it removes the amenability assumption. Furthermore, it shows that it is enough to determine the one-step facial residual functions for $\mathcal{K}$ and its faces, whereas [34, Theorem 23] may require all possible facial residual functions related to $\mathcal{K}$ and its faces. Nevertheless, Theorem 3.8 is still an abstract error bound result; whether some concrete inequality can be written down depends on obtaining a formula for the φ function. To do so, it would require finding expressions for the one-step facial residual functions. In the next subsections, we will address this challenge.

3.1 How to compute one-step facial residual functions?

In this section, we present some general tools for computing one-step facial residual functions.

Lemma 3.9 (𝟙-*FRF from error bound*) *Suppose that $\mathcal{K}$ is a closed convex cone and let $z \in \mathcal{K}^*$ be such that $\mathcal{F} = \{z\}^\perp \cap \mathcal{K}$ is a proper face of $\mathcal{K}$. Let $\mathfrak{g} : \mathbb{R}_+ \to \mathbb{R}_+$ be monotone nondecreasing with $\mathfrak{g}(0) = 0$, and let $\kappa_{z,\mathfrak{s}}$ be a finite monotone nondecreasing nonnegative function in $\mathfrak{s} \in \mathbb{R}_+$ such that*

$$d(\boldsymbol{q}, \mathcal{F}) \leq \kappa_{z, \|\boldsymbol{q}\|} \mathfrak{g}(d(\boldsymbol{q}, \mathcal{K})) \text{ whenever } \boldsymbol{q} \in \{z\}^\perp. \qquad (3.12)$$

Define the function $\psi_{\mathcal{K}, z} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ by

$$\psi_{\mathcal{K}, z}(s, t) := \max\{s, s/\|z\|\} + \kappa_{z,t} \mathfrak{g}(s + \max\{s, s/\|z\|\}).$$

Then we have

$$d(\boldsymbol{p}, \mathcal{F}) \leq \psi_{\mathcal{K}, z}(\epsilon, \|\boldsymbol{p}\|) \quad \text{whenever} \quad d(\boldsymbol{p}, \mathcal{K}) \leq \epsilon \text{ and } \langle \boldsymbol{p}, z \rangle \leq \epsilon. \qquad (3.13)$$

Moreover, $\psi_{\mathcal{K}, z}$ is a 𝟙-FRF for $\mathcal{K}$ and z.

Proof Suppose that $d(\mathbf{p}, \mathcal{K}) \leq \epsilon$ and $\langle \mathbf{p}, \mathbf{z} \rangle \leq \epsilon$. We first claim that

$$d(\mathbf{p}, \{\mathbf{z}\}^\perp) \leq \max\{\epsilon, \epsilon/\|\mathbf{z}\|\}. \qquad (3.14)$$

This can be shown as follows. Since $\mathbf{z} \in \mathcal{K}^*$, we have $\langle \mathbf{p} + P_\mathcal{K}(\mathbf{p}) - \mathbf{p}, \mathbf{z} \rangle \geq 0$ and

$$\langle \mathbf{p}, \mathbf{z} \rangle \geq -\langle P_\mathcal{K}(\mathbf{p}) - \mathbf{p}, \mathbf{z} \rangle \geq -\epsilon \|\mathbf{z}\|.$$

We conclude that $|\langle \mathbf{p}, \mathbf{z} \rangle| \leq \max\{\epsilon\|\mathbf{z}\|, \epsilon\}$. This, in combination with $d(\mathbf{p}, \{\mathbf{z}\}^{\perp}) = |\langle \mathbf{p}, \mathbf{z} \rangle|/\|\mathbf{z}\|$, leads to (3.14).

Next, let $\mathbf{q} := P_{\{\mathbf{z}\}^{\perp}} \mathbf{p}$. Then we have that

$$
\begin{aligned}
d(\mathbf{p}, \mathcal{F}) \leq \|\mathbf{p} - \mathbf{q}\| + d(\mathbf{q}, \mathcal{F}) &\overset{(a)}{\leq} \max\{\epsilon, \epsilon/\|\mathbf{z}\|\} + d(\mathbf{q}, \mathcal{F}) \\
&\overset{(b)}{\leq} \max\{\epsilon, \epsilon/\|\mathbf{z}\|\} + \kappa_{\mathbf{z}, \|\mathbf{q}\|} \mathfrak{g}\left(d(\mathbf{q}, \mathcal{K})\right) \\
&\overset{(c)}{\leq} \max\{\epsilon, \epsilon/\|\mathbf{z}\|\} + \kappa_{\mathbf{z}, \|\mathbf{p}\|} \mathfrak{g}\left(d(\mathbf{q}, \mathcal{K})\right) \\
&\overset{(d)}{\leq} \max\{\epsilon, \epsilon/\|\mathbf{z}\|\} + \kappa_{\mathbf{z}, \|\mathbf{p}\|} \mathfrak{g}\left(\epsilon + \max\{\epsilon, \epsilon/\|\mathbf{z}\|\}\right),
\end{aligned}
$$

where (a) follows from (3.14), (b) is a consequence of (3.12), (c) holds because $\|\mathbf{q}\| = \|P_{\{\mathbf{z}\}^{\perp}} \mathbf{p}\| \leq \|\mathbf{p}\|$ so that $\kappa_{\mathbf{z}, \|\mathbf{q}\|} \leq \kappa_{\mathbf{z}, \|\mathbf{p}\|}$, and (d) holds because $\mathfrak{g}$ is monotone nondecreasing and

$$
d(\mathbf{q}, \mathcal{K}) \leq d(\mathbf{p}, \mathcal{K}) + \|\mathbf{q} - \mathbf{p}\| \leq \epsilon + \max\{\epsilon, \epsilon/\|\mathbf{z}\|\};
$$

here, the second inequality follows from (3.14) and the assumption that $d(\mathbf{p}, \mathcal{K}) \leq \epsilon$. This proves (3.13). Finally, notice that $\psi_{\mathcal{K}, \mathbf{z}}$ is nonnegative, monotone nondecreasing in each argument, and that $\psi_{\mathcal{K}, \mathbf{z}}(0, t) = 0$ for every $t \in \mathbb{R}_{+}$. Hence, $\psi_{\mathcal{K}, \mathbf{z}}$ is a one-step facial residual function for $\mathcal{K}$ and $\mathbf{z}$. □

In view of Lemma 3.9, one may construct one-step facial residual functions after establishing the error bound (3.12). In the next theorem, we present a characterization for the existence of such an error bound. Our result is based on the quantity (3.15) defined below being *nonzero*. Note that this quantity *does not* explicitly involve projections onto $\mathcal{K}$; this enables us to work with the exponential cone later, whose projections do not seem to have simple expressions. Figure 2 provides a geometric interpretation of (3.15).

Theorem 3.10 (Characterization of the existence of error bounds) *Suppose that $\mathcal{K}$ is a closed convex cone and let $\mathbf{z} \in \mathcal{K}^{*}$ be such that $\mathcal{F} = \{\mathbf{z}\}^{\perp} \cap \mathcal{K}$ is a nontrivial exposed face of $\mathcal{K}$. Let $\eta \geq 0$, $\alpha \in (0, 1]$ and let $\mathfrak{g} : \mathbb{R}_{+} \to \mathbb{R}_{+}$ be monotone nondecreasing with $\mathfrak{g}(0) = 0$ and $\mathfrak{g} \geq |\cdot|^{\alpha}$. Define*

$$
\gamma_{\mathbf{z}, \eta} := \inf_{\mathbf{v}} \left\{ \frac{\mathfrak{g}(\|\mathbf{w} - \mathbf{v}\|)}{\|\mathbf{w} - \mathbf{u}\|} \;\middle|\; \mathbf{v} \in \partial \mathcal{K} \cap B(\eta) \backslash \mathcal{F}, \; \mathbf{w} = P_{\{\mathbf{z}\}^{\perp}} \mathbf{v}, \; \mathbf{u} = P_{\mathcal{F}} \mathbf{w}, \; \mathbf{w} \neq \mathbf{u} \right\}.
\tag{3.15}
$$

Then the following statements hold.

(i) If $\gamma_{\mathbf{z}, \eta} \in (0, \infty]$, then it holds that

$$
d(\mathbf{q}, \mathcal{F}) \leq \kappa_{\mathbf{z}, \eta} \mathfrak{g}(d(\mathbf{q}, \mathcal{K})) \quad \text{whenever } \mathbf{q} \in \{\mathbf{z}\}^{\perp} \cap B(\eta),
\tag{3.16}
$$

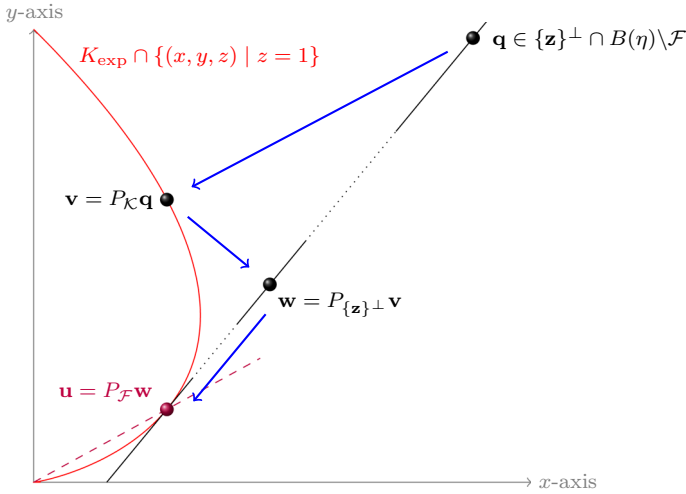*where $\kappa_{\mathbf{z}, \eta} := \max\left\{ 2\eta^{1-\alpha}, 2\gamma_{\mathbf{z}, \eta}^{-1} \right\} < \infty$.*

Fig. 2 Theorem 3.10 shows that we may replace the problem of showing that $\mathfrak{g}(\|\mathbf{q} - P_\mathcal{K}\mathbf{q}\|)/\|\mathbf{q} - P_\mathcal{F}\mathbf{q}\|$ is uniformly bounded away from zero for all $\mathbf{q} \in \{\mathbf{z}\}^\perp \cap B(\eta) \setminus \mathcal{F}$, with the equivalent problem of showing that $\mathfrak{g}(\|\mathbf{v} - P_{\{\mathbf{z}\}^\perp}\mathbf{v}\|)/\|P_{\{\mathbf{z}\}^\perp}\mathbf{v} - P_\mathcal{F} \circ P_{\{\mathbf{z}\}^\perp}\mathbf{v}\|$ is uniformly bounded away from zero for all $\mathbf{v} \in \partial\mathcal{K} \cap B(\eta) \setminus \mathcal{F}$ with $P_{\{\mathbf{z}\}^\perp}\mathbf{v} \neq P_\mathcal{F} \circ P_{\{\mathbf{z}\}^\perp}\mathbf{v}$. This second problem can sometimes be easier to deal with, because it obviates the need to project onto $\mathcal{K}$ and projects onto the nontrivial exposed face $\mathcal{F}$ instead. For describing a possibly higher dimensional problem in 2D, we represent $\{\mathbf{z}\}^\perp$ with a line, $\mathcal{K}$ with a 2D slice, and $\mathcal{F}$ with a dot; of course, this is an oversimplification, since $\mathbf{q}, \mathbf{w}, \mathbf{u}$ are *not* generically colinear, nor would any of the points necessarily lie in the same horizontal slice. The scenario shown is meant to suggest intuition, but it is not a plausible configuration of points

(ii) If there exists $\kappa_B \in (0, \infty)$ so that

$$d(q, \mathcal{F}) \leq \kappa_B \mathfrak{g}(d(q, \mathcal{K})) \quad whenever \ q \in \{z\}^\perp \cap B(\eta), \tag{3.17}$$

then $\gamma_{z,\eta} \in (0, \infty]$.

Proof We first consider item (i). If $\eta = 0$ or $\mathbf{q} \in \mathcal{F}$, the result is vacuously true, so let $\eta > 0$ and $\mathbf{q} \in \{\mathbf{z}\}^\perp \cap B(\eta) \setminus \mathcal{F}$. Then $\mathbf{q} \notin \mathcal{K}$ because $\mathcal{F} = \{\mathbf{z}\}^\perp \cap \mathcal{K}$. Define

$$\mathbf{v} = P_\mathcal{K}\mathbf{q}, \quad \mathbf{w} = P_{\{\mathbf{z}\}^\perp}\mathbf{v}, \quad \text{and} \quad \mathbf{u} = P_\mathcal{F}\mathbf{w}.$$

Then $\mathbf{v} \in \partial\mathcal{K} \cap B(\eta)$ because $\mathbf{q} \notin \mathcal{K}$ and $\|\mathbf{q}\| \leq \eta$. If $\mathbf{v} \in \mathcal{F}$, then we have $d(\mathbf{q}, \mathcal{F}) = d(\mathbf{q}, \mathcal{K})$ and hence

$$d(\mathbf{q}, \mathcal{F}) = d(\mathbf{q}, \mathcal{K})^{1-\alpha} \cdot d(\mathbf{q}, \mathcal{K})^\alpha \leq \eta^{1-\alpha} d(\mathbf{q}, \mathcal{K})^\alpha \leq \kappa_{\mathbf{z},\eta} \mathfrak{g}(d(\mathbf{q}, \mathcal{K})),$$

where the first inequality holds because $\|\mathbf{q}\| \leq \eta$, and the last inequality follows from the definitions of $\mathfrak{g}$ and $\kappa_{\mathbf{z},\eta}$. Thus, from now on, we assume that $\mathbf{v} \in \partial\mathcal{K} \cap B(\eta) \setminus \mathcal{F}$.

Next, since $\mathbf{w} = P_{\{\mathbf{z}\}^\perp}\mathbf{v}$, it holds that $\mathbf{v} - \mathbf{w} \in \{\mathbf{z}\}^{\perp\perp}$ and hence $\|\mathbf{q} - \mathbf{v}\|^2 = \|\mathbf{q} - \mathbf{w}\|^2 + \|\mathbf{w} - \mathbf{v}\|^2$. In particular, we have

$$d(\mathbf{q}, \mathcal{K}) = \|\mathbf{q} - \mathbf{v}\| \geq \max\{\|\mathbf{v} - \mathbf{w}\|, \|\mathbf{q} - \mathbf{w}\|\}, \tag{3.18}$$

where the equality follows from the definition of **v**. Now, to establish (3.16), we consider two cases.

(I) $d(\mathbf{q}, \mathcal{F}) \leq 2d(\mathbf{w}, \mathcal{F})$;
(II) $d(\mathbf{q}, \mathcal{F}) > 2d(\mathbf{w}, \mathcal{F})$.

(I): In this case, we have from $\mathbf{u} = P_{\mathcal{F}}\mathbf{w}$ and $\mathbf{q} \notin \mathcal{F}$ that

$$2\|\mathbf{w} - \mathbf{u}\| = 2d(\mathbf{w}, \mathcal{F}) \geq d(\mathbf{q}, \mathcal{F}) > 0, \tag{3.19}$$

where the first inequality follows from the assumption in this case (I). Hence,

$$\frac{1}{\kappa_{\mathbf{z},\eta}} \overset{(a)}{\leq} \frac{1}{2}\gamma_{\mathbf{z},\eta} \overset{(b)}{\leq} \frac{\mathfrak{g}(\|\mathbf{w} - \mathbf{v}\|)}{2\|\mathbf{w} - \mathbf{u}\|} \overset{(c)}{\leq} \frac{\mathfrak{g}(d(\mathbf{q}, \mathcal{K}))}{2\|\mathbf{w} - \mathbf{u}\|} \overset{(d)}{\leq} \frac{\mathfrak{g}(d(\mathbf{q}, \mathcal{K}))}{d(\mathbf{q}, \mathcal{F})}, \tag{3.20}$$

where (a) is true by the definition of $\kappa_{\mathbf{z},\eta}$, (b) uses the condition that $\mathbf{v} \in \partial\mathcal{K} \cap B(\eta)\backslash\mathcal{F}$, (3.19) and the definition of $\gamma_{\mathbf{z},\eta}$, (c) is true by (3.18) and the monotonicity of $\mathfrak{g}$, and (d) follows from (3.19). This concludes case (I).[6]

(II): Using the triangle inequality, we have

$$2d(\mathbf{q}, \mathcal{F}) \leq 2\|\mathbf{q} - \mathbf{w}\| + 2d(\mathbf{w}, \mathcal{F}) < 2\|\mathbf{q} - \mathbf{w}\| + d(\mathbf{q}, \mathcal{F}),$$

where the strict inequality follows from the condition for this case (II). Consequently, we have $d(\mathbf{q}, \mathcal{F}) \leq 2\|\mathbf{q} - \mathbf{w}\|$. Combining this with (3.18), we deduce further that

$$\begin{aligned} d(\mathbf{q}, \mathcal{F}) &\leq 2\|\mathbf{q} - \mathbf{w}\| \leq 2\max\{\|\mathbf{v} - \mathbf{w}\|, \|\mathbf{q} - \mathbf{w}\|\} \\ &\leq 2d(\mathbf{q}, \mathcal{K}) = 2d(\mathbf{q}, \mathcal{K})^{1-\alpha} \cdot d(\mathbf{q}, \mathcal{K})^{\alpha} \\ &\leq 2\eta^{1-\alpha}d(\mathbf{q}, \mathcal{K})^{\alpha} \leq \kappa_{\mathbf{z},\eta}\mathfrak{g}(d(\mathbf{q}, \mathcal{K})), \end{aligned}$$

where the fourth inequality holds because $\|\mathbf{q}\| \leq \eta$, and the last inequality follows from the definitions of $\mathfrak{g}$ and $\kappa_{\mathbf{z},\eta}$. This proves item (i).

We next consider item (ii). Again, the result is vacuously true if $\eta = 0$, so let $\eta > 0$. Let $\mathbf{v} \in \partial\mathcal{K} \cap B(\eta)\backslash\mathcal{F}$, $\mathbf{w} = P_{\{\mathbf{z}\}^{\perp}}\mathbf{v}$ and $\mathbf{u} = P_{\mathcal{F}}\mathbf{w}$ with $\mathbf{w} \neq \mathbf{u}$. Then $\mathbf{w} \in B(\eta)$, and we have in view of (3.17) that

$$\|\mathbf{w} - \mathbf{u}\| \overset{(a)}{=} d(\mathbf{w}, \mathcal{F}) \overset{(b)}{\leq} \kappa_B\mathfrak{g}(d(\mathbf{w}, \mathcal{K})) \overset{(c)}{\leq} \kappa_B\mathfrak{g}(\|\mathbf{w} - \mathbf{v}\|),$$

where (a) holds because $\mathbf{u} = P_{\mathcal{F}}\mathbf{w}$, (b) holds because of (3.17), $\mathbf{w} \in \{\mathbf{z}\}^{\perp}$ and $\|\mathbf{w}\| \leq \eta$, and (c) is true because $\mathfrak{g}$ is monotone nondecreasing and $\mathbf{v} \in \mathcal{K}$. Thus, we have $\gamma_{\mathbf{z},\eta} \geq 1/\kappa_B > 0$. This completes the proof. $\qquad\square$

Remark 3.11 (About $\kappa_{\mathbf{z},\eta}$ and $\gamma_{\mathbf{z},\eta}^{-1}$) As η increases, the infimum in (3.15) is taken over a larger region, so $\gamma_{\mathbf{z},\eta}$ does not increase. Accordingly, $\gamma_{\mathbf{z},\eta}^{-1}$ does not decrease when η increases. Therefore, the $\kappa_{\mathbf{z},\eta}$ and $\gamma_{\mathbf{z},\eta}^{-1}$ considered in Theorem 3.10 are monotone

---

[6] In particular, in view of (3.20), we see that this case only happens when $\gamma_{\mathbf{z},\eta} < \infty$.

nondecreasing as functions of η when $\mathbf{z}$ is fixed. We are also using the convention that $1/\infty = 0$ so that $\kappa_{\mathbf{z},\eta} = 2\eta^{1-\alpha}$ when $\gamma_{\mathbf{z},\eta} = \infty$.

Thus, to establish an error bound as in (3.16), it suffices to show that $\gamma_{\mathbf{z},\eta} \in (0, \infty]$ for the choice of $\mathfrak{g}$ and $\eta \geq 0$. Clearly, $\gamma_{\mathbf{z},0} = \infty$. The next lemma allows us to check whether $\gamma_{\mathbf{z},\eta} \in (0, \infty]$ for an $\eta > 0$ by considering *convergent* sequences.

Lemma 3.12 *Suppose that $\mathcal{K}$ is a closed convex cone and let $\mathbf{z} \in \mathcal{K}^*$ be such that $\mathcal{F} = \{\mathbf{z}\}^\perp \cap \mathcal{K}$ is a nontrivial exposed face of $\mathcal{K}$. Let $\eta > 0$, $\alpha \in (0, 1]$ and let $\mathfrak{g} : \mathbb{R}_+ \to \mathbb{R}_+$ be monotone nondecreasing with $\mathfrak{g}(0) = 0$ and $\mathfrak{g} \geq |\cdot|^\alpha$. Let $\gamma_{\mathbf{z},\eta}$ be defined as in (3.15). If $\gamma_{\mathbf{z},\eta} = 0$, then there exist $\bar{\mathbf{v}} \in \mathcal{F}$ and a sequence $\{\mathbf{v}^k\} \subset \partial\mathcal{K} \cap B(\eta)\backslash\mathcal{F}$ such that*

$$\lim_{k\to\infty} \mathbf{v}^k = \lim_{k\to\infty} \mathbf{w}^k = \bar{\mathbf{v}} \tag{3.21a}$$

$$\text{and} \quad \lim_{k\to\infty} \frac{\mathfrak{g}(\|\mathbf{w}^k - \mathbf{v}^k\|)}{\|\mathbf{w}^k - \mathbf{u}^k\|} = 0, \tag{3.21b}$$

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

Proof Suppose that $\gamma_{\mathbf{z},\eta} = 0$. Then, by the definition of infimum, there exists a sequence $\{\mathbf{v}^k\} \subset \partial\mathcal{K} \cap B(\eta)\backslash\mathcal{F}$ such that

$$\lim_{k\to\infty} \frac{\mathfrak{g}(\|\mathbf{w}^k - \mathbf{v}^k\|)}{\|\mathbf{w}^k - \mathbf{u}^k\|} = 0, \tag{3.22}$$

where $\mathbf{w}^k = P_{\{\mathbf{z}\}^\perp}\mathbf{v}^k$, $\mathbf{u}^k = P_\mathcal{F}\mathbf{w}^k$ and $\mathbf{w}^k \neq \mathbf{u}^k$. Since $\{\mathbf{v}^k\} \subset B(\eta)$, by passing to a convergent subsequence if necessary, we may assume without loss of generality that

$$\lim_{k\to\infty} \mathbf{v}^k = \bar{\mathbf{v}} \tag{3.23}$$

for some $\bar{\mathbf{v}} \in \mathcal{K} \cap B(\eta)$. In addition, since $0 \in \mathcal{F} \subseteq \{\mathbf{z}\}^\perp$, and projections onto closed convex sets are nonexpansive, we see that $\{\mathbf{w}^k\} \subset B(\eta)$ and $\{\mathbf{u}^k\} \subset B(\eta)$, and hence the sequence $\{\|\mathbf{w}^k - \mathbf{u}^k\|\}$ is bounded. Then we can conclude from (3.22) and the assumption $\mathfrak{g} \geq |\cdot|^\alpha$ that

$$\lim_{k\to\infty} \|\mathbf{w}^k - \mathbf{v}^k\| = 0. \tag{3.24}$$

Now (3.24), (3.23), and the triangle inequality give $\mathbf{w}^k \to \bar{\mathbf{v}}$. Since $\{\mathbf{w}^k\} \subset \{\mathbf{z}\}^\perp$, it then follows that $\bar{\mathbf{v}} \in \{\mathbf{z}\}^\perp$. Thus, $\bar{\mathbf{v}} \in \{\mathbf{z}\}^\perp \cap \mathcal{K} = \mathcal{F}$. This completes the proof. $\qquad\square$

Let $\mathcal{K}$ be a closed convex cone. Lemma 3.9, Theorem 3.10 and Lemma 3.12 are tools to obtain one-step facial residual functions for $\mathcal{K}$. These are exactly the kind of facial residual functions needed in the abstract error bound result, Theorem 3.8. We conclude this subsection with a result that connects the one-step facial residual functions of a product cone and those of its constituent cones, which is useful for deriving error bounds for product cones.

Proposition 3.13 (**1**-*FRF for products*) *Let* $\mathcal{K}^i \subseteq \mathcal{E}^i$ *be closed convex cones for every* $i \in \{1, \ldots, m\}$ *and let* $\mathcal{K} = \mathcal{K}^1 \times \cdots \times \mathcal{K}^m$. *Let* $\mathcal{F} \unlhd \mathcal{K}$, $z \in \mathcal{F}^*$ *and suppose that* $\mathcal{F} = \mathcal{F}^1 \times \cdots \times \mathcal{F}^m$ *with* $\mathcal{F}^i \unlhd \mathcal{K}^i$ *for every* $i \in \{1, \ldots, m\}$. *Write* $z = (z_1, \ldots, z_m)$ *with* $z_i \in (\mathcal{F}^i)^*$.

For every i, *let* $\psi_{\mathcal{F}^i, z_i}$ *be a* **1**-*FRF for* $\mathcal{F}^i$ *and* z_i. *Then, there exists a* $\kappa > 0$ *such that the function* $\psi_{\mathcal{F}, z}$ *satisfying*

$$\psi_{\mathcal{F}, z}(\epsilon, t) = \sum_{i=1}^m \psi_{\mathcal{F}^i, z_i}(\kappa \epsilon, t)$$

is a **1**-*FRF for* $\mathcal{F}$ *and* z.

Proof Suppose that $\mathbf{x} \in \operatorname{span} \mathcal{F}$ and $\epsilon \geq 0$ satisfy the inequalities

$$\mathrm{d}(\mathbf{x}, \mathcal{F}) \leq \epsilon, \quad \langle \mathbf{x}, \mathbf{z} \rangle \leq \epsilon.$$

We note that

$$\mathcal{F} \cap \{\mathbf{z}\}^\perp = (\mathcal{F}^1 \cap \{\mathbf{z}_1\}^\perp) \times \cdots \times (\mathcal{F}^m \cap \{\mathbf{z}_m\}^\perp),$$

and that for every $i \in \{1, \ldots, m\}$,

$$\mathrm{d}(\mathbf{x}_i, \mathcal{F}^i) \leq \mathrm{d}(\mathbf{x}, \mathcal{F}) \leq \epsilon. \tag{3.25}$$

Since $\mathbf{z}_i \in (\mathcal{F}^i)^*$, we have from (3.25) that

$$0 \leq \langle \mathbf{z}_i, P_{\mathcal{F}^i}(\mathbf{x}_i) \rangle = \langle \mathbf{z}_i, P_{\mathcal{F}^i}(\mathbf{x}_i) - \mathbf{x}_i + \mathbf{x}_i \rangle \leq \epsilon \|\mathbf{z}_i\| + \langle \mathbf{z}_i, \mathbf{x}_i \rangle. \tag{3.26}$$

Using (3.26) for all i and recalling that $\langle \mathbf{z}, \mathbf{x} \rangle \leq \epsilon$, we have

$$\langle (\mathbf{z}_1, \ldots, \mathbf{z}_m), (P_{\mathcal{F}^1}(\mathbf{x}_1), \ldots, P_{\mathcal{F}^m}(\mathbf{x}_m)) \rangle \leq \sum_{i=1}^m [\epsilon \|\mathbf{z}_i\| + \langle \mathbf{z}_i, \mathbf{x}_i \rangle] \leq \hat{\kappa} \epsilon, \tag{3.27}$$

where $\hat{\kappa} = 1 + \sum_{i=1}^m \|\mathbf{z}_i\|$. Since $\langle \mathbf{z}_i, P_{\mathcal{F}^i}(\mathbf{x}_i) \rangle \geq 0$ for $i \in \{1, \ldots, m\}$, from (3.27) we obtain

$$\langle \mathbf{z}_i, P_{\mathcal{F}^i}(\mathbf{x}_i) \rangle \leq \hat{\kappa} \epsilon, \quad i \in \{1, \ldots, m\}. \tag{3.28}$$

This implies that for $i \in \{1, \ldots, m\}$ we have

$$\langle \mathbf{z}_i, \mathbf{x}_i \rangle = \langle \mathbf{z}_i, \mathbf{x}_i - P_{\mathcal{F}^i}(\mathbf{x}_i) + P_{\mathcal{F}^i}(\mathbf{x}_i) \rangle \leq \epsilon \|\mathbf{z}\| + \hat{\kappa} \epsilon, \tag{3.29}$$

where the inequality follows from (3.25) and (3.28). Now, recapitulating, the facial residual function $\psi_{\mathcal{F}^i, z_i}$ has the property that if $\gamma_1, \gamma_2 \in \mathbb{R}_+$ then the relations

$$\mathbf{y}_i \in \operatorname{span} \mathcal{F}^i, \quad \mathrm{d}(\mathbf{y}_i, \mathcal{F}^i) \leq \gamma_1, \quad \langle \mathbf{y}_i, \mathbf{z}_i \rangle \leq \gamma_2$$

imply $d(\mathbf{y}_i, \mathcal{F}^i \cap \{\mathbf{z}_i\}^{\perp}) \leq \psi_{\mathcal{F}^i, \mathbf{z}_i}(\max_{1 \leq j \leq 2} \{\gamma_j\}, \|\mathbf{y}_i\|)$. Therefore, from (3.25), (3.29) and the monotonicity of $\psi_{\mathcal{F}^i, \mathbf{z}_i}$, we have upon recalling $\mathbf{x} \in \text{span} \, \mathcal{F}$ that

$$d(\mathbf{x}_i, \mathcal{F}^i \cap \{\mathbf{z}_i\}^{\perp}) \leq \psi_{\mathcal{F}^i, \mathbf{z}_i}(\max\{1, \hat{\kappa} + \|\mathbf{z}\|\}\epsilon, \|\mathbf{x}_i\|). \tag{3.30}$$

Finally, from (3.30), we conclude that

$$d(\mathbf{x}, \mathcal{F} \cap \{\mathbf{z}\}^{\perp}) \leq \sum_{i=1}^{m} d(\mathbf{x}_i, \mathcal{F}^i \cap \{\mathbf{z}_i\}^{\perp}) \leq \sum_{i=1}^{m} \psi_{\mathcal{F}^i, \mathbf{z}_i}(\max\{1, \hat{\kappa} + \|\mathbf{z}\|\}\epsilon, \|\mathbf{x}\|),$$

where we also used the monotonicity of $\psi_{\mathcal{F}^i, \mathbf{z}_i}$ for the last inequality. This completes the proof. $\square$

4 The exponential cone

In this section, we will use all the techniques developed so far to obtain error bounds for the 3D exponential cone $K_{\exp}$. We will start with a study of its facial structure in Sect. 4.1, then we will compute its one-step facial residual functions in Sect. 4.2. Finally, error bounds will be presented in Sect. 4.3. In Sect. 4.4, we summarize odd behaviour found in the facial structure of the exponential cone.

4.1 Facial structure

Recall that the exponential cone is defined as follows:

$$K_{\exp} := \left\{(x, y, z) \mid y > 0, z \geq ye^{x/y}\right\} \cup \{(x, y, z) \mid x \leq 0, z \geq 0, y = 0\}. \tag{4.1}$$

Its dual cone is given by

$$K_{\exp}^* := \left\{(x, y, z) \mid x < 0, ez \geq -xe^{y/x}\right\} \cup \{(x, y, z) \mid x = 0, z \geq 0, y \geq 0\}.$$

It may therefore be readily seen that $K_{\exp}^*$ is a scaled and rotated version of $K_{\exp}$. In this subsection, we will describe the nontrivial faces of $K_{\exp}$; see Fig. 3. We will show that we have the following types of nontrivial faces:

(a) infinitely many exposed extreme rays (1D faces) parametrized by $\beta \in \mathbb{R}$ as follows:

$$\mathcal{F}_{\beta} := \left\{\left(-\beta y + y, y, e^{1-\beta}y\right) \, \middle| \, y \in [0, \infty)\right\}. \tag{4.2}$$

(b) a single "exceptional" exposed extreme ray denoted by $\mathcal{F}_{\infty}$:

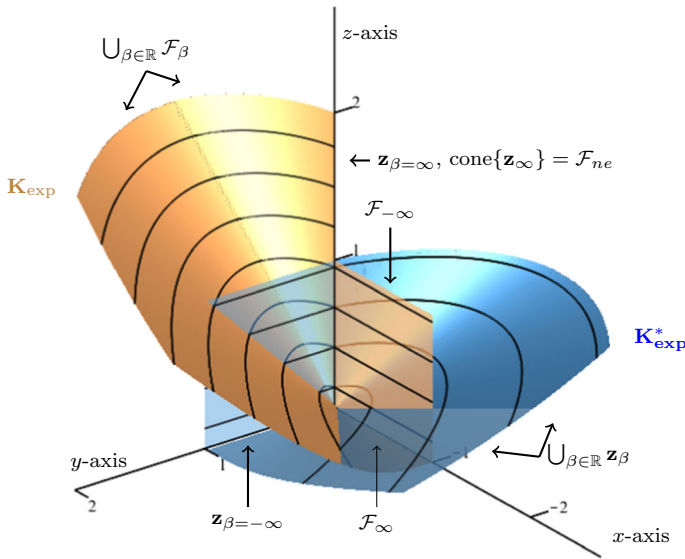$$\mathcal{F}_{\infty} := \{(x, 0, 0) \mid x \leq 0\}. \tag{4.3}$$

Fig. 3 The exponential cone and its dual, with faces and exposing vectors labeled according to our index β

(c) a single non-exposed extreme ray denoted by $\mathcal{F}_{ne}$:

$$\mathcal{F}_{ne} := \{(0, 0, z) \mid z \geq 0\}. \tag{4.4}$$

(d) a single 2D exposed face denoted by $\mathcal{F}_{-\infty}$:

$$\mathcal{F}_{-\infty} := \{(x, y, z) \mid x \leq 0, z \geq 0, y = 0\}, \tag{4.5}$$

where we note that $\mathcal{F}_{\infty}$ and $\mathcal{F}_{ne}$ are the extreme rays of $\mathcal{F}_{-\infty}$.

Notice that except for the case (c), all faces are exposed and thus arise as an intersection $\{\mathbf{z}\}^{\perp} \cap K_{\exp}$ for some $\mathbf{z} \in K_{\exp}^*$. To establish the above characterization, we start by examining how the components of $\mathbf{z}$ determine the corresponding exposed face.

4.1.1 Exposed faces

Let $\mathbf{z} \in K_{\exp}^*$ be such that $\{\mathbf{z}\}^{\perp} \cap K_{\exp}$ is a nontrivial face of $K_{\exp}$. Then $\mathbf{z} \neq 0$ and $\mathbf{z} \in \partial K_{\exp}^*$. We consider the following cases.

$\underline{z_x < 0}$: Since $\mathbf{z} \in \partial K_{\exp}^*$, we must have $z_z e = -z_x e^{\frac{z_y}{z_x}}$ and hence

$$\mathbf{z} = (z_x, z_y, -z_x e^{\frac{z_y}{z_x} - 1}). \tag{4.6}$$

Since $z_x \neq 0$, we see that $\mathbf{q} \in \{\mathbf{z}\}^{\perp}$ if and only if

$$q_x + q_y \left(\frac{z_y}{z_x}\right) - q_z e^{\frac{z_y}{z_x} - 1} = 0. \tag{4.7}$$

Solving (4.7) for q_z and letting $\beta := \frac{z_y}{z_x}$ to simplify the exposition, we have

$$q_z = e^{1 - \frac{z_y}{z_x}} \left(q_x + q_y \cdot \frac{z_y}{z_x} \right) = e^{1-\beta} \left(q_x + q_y \beta \right)$$

$$\text{with } \beta := \frac{z_y}{z_x} \in (-\infty, \infty). \tag{4.8}$$

Thus, we obtain that $\{\mathbf{z}\}^\perp = \{(x, y, e^{1-\beta}(x + y\beta)) \mid x, y \in \mathbb{R}\}$. Combining this with the definition of $K_{\exp}$ and the fact that $\{\mathbf{z}\}^\perp$ is a supporting hyperplane (so that $K_{\exp} \cap \{\mathbf{z}\}^\perp = \partial K_{\exp} \cap \{\mathbf{z}\}^\perp$) yields

$$
\begin{aligned}
K_{\exp} \cap \{\mathbf{z}\}^\perp &= \partial K_{\exp} \cap \{\mathbf{z}\}^\perp \\
&= \left\{ \left(x, y, e^{1-\beta}(x + y\beta)\right) \mid e^{1-\beta}(x + y\beta) = y e^{\frac{x}{y}}, y > 0 \right\} \\
&\quad \bigcup \left\{ \left(x, y, e^{1-\beta}(x + y\beta)\right) \mid x \le 0, e^{1-\beta}(x + y\beta) \ge 0, y = 0 \right\} \\
&= \left\{ \left(x, y, e^{1-\beta}(x + y\beta)\right) \mid e^{1-\beta}(x + y\beta) = y e^{\frac{x}{y}}, y > 0 \right\} \cup \{0\}. \quad (4.9)
\end{aligned}
$$

We now refine the above characterization in the next proposition.

Proposition 4.1 (Characterization of $\mathcal{F}_\beta$, $\beta \in \mathbb{R}$) *Let* $z \in K_{\exp}^*$ *satisfy* $z = (z_x, z_y, z_z)$, *where* $z_z e = -z_x e^{\frac{z_y}{z_x}}$ *and* $z_x < 0$. *Define* $\beta = \frac{z_y}{z_x}$ *as in* (4.8) *and let* $\mathcal{F}_\beta := K_{\exp} \cap \{\mathbf{z}\}^\perp$. *Then*

$$\mathcal{F}_\beta = \left\{ \left(-\beta y + y, y, e^{1-\beta} y \right) \mid y \in [0, \infty) \right\}.$$

Proof Let $\Omega := \{ (-\beta y + y, y, e^{1-\beta} y) \mid y \in [0, \infty) \}$. In view of (4.9), we can check that $\Omega \subseteq \mathcal{F}_\beta$. To prove the converse inclusion, pick any $\mathbf{q} = (x, y, e^{1-\beta}(x + y\beta)) \in \mathcal{F}_\beta$. We need to show that $\mathbf{q} \in \Omega$.

To this end, we note from (4.9) that if $y = 0$, then necessarily $\mathbf{q} = \mathbf{0}$ and consequently $\mathbf{q} \in \Omega$. On the other hand, if $y > 0$, then (4.9) gives $y e^{x/y} = (x + \beta y) e^{1-\beta}$. Then we have the following chain of equivalences:

$$
\begin{aligned}
& y e^{x/y} = (x + \beta y) e^{1-\beta} \\
\Longleftrightarrow\ & -e^{-1} = -(x/y + \beta) e^{-(x/y + \beta)} \\
\overset{(a)}{\Longleftrightarrow}\ & -x/y - \beta = -1 \\
\Longleftrightarrow\ & x = y - y\beta,
\end{aligned}
\tag{4.10}
$$

where (a) follows from the fact that the function $t \mapsto t e^t$ is strictly increasing on $[-1, \infty)$. Plugging the last expression back into $\mathbf{q}$, we may compute

$$q_z = e^{1-\beta}(x + y\beta) = e^{1-\beta}(y - y\beta + y\beta) = y e^{1-\beta}. \tag{4.11}$$

Altogether, (4.10), (4.11) together with $y > 0$ yield

$$\mathbf{q} = \left(y - \beta y, y, y e^{1-\beta}\right) \in \Omega.$$

This completes the proof. □

Next, we move on to the two remaining cases.

$z_x = 0, z_z > 0$: Notice that $\mathbf{q} \in K_{\exp}$ means that $q_y \geq 0$ and $q_z \geq 0$. Since $z_z > 0$ and $z_y \geq 0$, in order to have $\mathbf{q} \in \{\mathbf{z}\}^{\perp}$, we must have $q_z = 0$. The the definition of $K_{\exp}$ also forces $q_y = 0$ and hence

$$\{\mathbf{z}\}^{\perp} \cap K_{\exp} = \{(x, 0, 0) \mid x \leq 0\} =: \mathcal{F}_{\infty}. \tag{4.12}$$

This one-dimensional face is exposed by any hyperplane with normal vectors coming from the set $\{(0, z_y, z_z) : z_y \geq 0, z_z > 0)\}$.

$z_x = 0, z_z = 0$: In this case, we have $z_y > 0$. In order to have $\mathbf{q} \in \{\mathbf{z}\}^{\perp}$, we must have $q_y = 0$. Thus

$$\{\mathbf{z}\}^{\perp} \cap K_{\exp} = \{(x, y, z) \mid x \leq 0, z \geq 0, y = 0\} =: \mathcal{F}_{-\infty}, \tag{4.13}$$

which is the unique two-dimensional face of $K_{\exp}$.

4.1.2 The single non-exposed face and completeness of the classification

The face $\mathcal{F}_{ne}$ is non-exposed because, as shown in Proposition 4.1, (4.12) and (4.13), it never arises as an intersection of the form $\{\mathbf{z}\}^{\perp} \cap K_{\exp}$, for $\mathbf{z} \in K_{\exp}^{*}$.

We now show that all nontrivial faces of $K_{\exp}$ were accounted for in (4.2), (4.3), (4.4), (4.5). First of all, by the discussion in Sect. 4.1.1, all nontrivial exposed faces must be among the ones in (4.2), (4.3) and (4.5). So, let $\mathcal{F}$ be a non-exposed face of $K_{\exp}$. Then, it must be contained in a nontrivial exposed face of $K_{\exp}$.[7] Therefore, $\mathcal{F}$ must be a proper face of the unique 2D face (4.5). This implies that $\mathcal{F}$ is one of the extreme rays of (4.5): $\mathcal{F}_{\infty}$ or $\mathcal{F}_{ne}$. By assumption, $\mathcal{F}$ is non-exposed, so it must be $\mathcal{F}_{ne}$.

4.2 One-step facial residual functions

In this subsection, we will use the machinery developed in Sect. 3 to obtain the one-step facial residual functions for $K_{\exp}$.

Let us first discuss how the discoveries were originally made, and how that process motivated the development of the framework we built in Sect. 3. The FRFs proven here were initially found by using the characterizations of Theorem 3.10 and Lemma 3.12 together with numerical experiments. Specifically, we used *Maple*[8] to numerically

---

[7] This is a general fact. A proper face of a closed convex cone is contained in a proper exposed face, e.g., [13, Proposition 3.6].

[8] Though one could use any suitable computer algebra package.

Springer

evaluate limits of relevant sequences (3.21), as well as plotting lower dimensional slices of the function $\mathbf{v} \mapsto \mathfrak{g}(\|\mathbf{v} - \mathbf{w}\|)/\|\mathbf{w} - \mathbf{u}\|$, where $\mathbf{w}$ and $\mathbf{u}$ are defined similarly as in (3.15).

A natural question is whether it might be simpler to change coordinates and work with the nearly equivalent $\mathbf{w} \mapsto \mathfrak{g}(\|\mathbf{v} - \mathbf{w}\|)/\|\mathbf{w} - \mathbf{u}\|$, since $\mathbf{w} \in \{\mathbf{z}\}^{\perp}$. However, $P_{\{\mathbf{z}\}^{\perp}}^{-1}\{\mathbf{w}\} \cap \partial \mathcal{K}$ may contain multiple points, which creates many challenges. We encountered an example of this when working with the exponential cone, where the change of coordinates from $\mathbf{v}$ to $\mathbf{w}$ necessitates the introduction of the two real branches of the Lambert $\mathcal{W}$ function (see, for example, [7, 12, 14] or [48] for the closely related Wright Omega function). With terrible effort, one can use such a parametrization to prove the FRFs for $\mathcal{F}_\beta$, $\beta \in [-\infty, \infty] \setminus \{\hat{\beta} := -\mathcal{W}_{\text{principal}}(2e^{-2})/2\}$. However, the change of branches inhibits proving the result for the exceptional number $\hat{\beta}$. The change of variables to $\mathbf{v}$ cures this problem by obviating the need for a branch function in the analysis; see [31] for additional details. This is why we present Theorem 3.10 in terms of $\mathbf{v}$. Computational investigation also pointed to the path of proof, though the proof we present may be understood without the aid of a computer.

4.2.1 $\mathcal{F}_{-\infty}$: the unique 2D face

Recall the unique 2D face of $K_{\exp}$:

$$\mathcal{F}_{-\infty} := \{(x, y, z) \mid x \leq 0, z \geq 0, y = 0\}.$$

Define the piecewise modified Boltzmann–Shannon entropy $\mathfrak{g}_{-\infty} : \mathbb{R}_+ \to \mathbb{R}_+$ as follows:

$$\mathfrak{g}_{-\infty}(t) := \begin{cases} 0 & \text{if } t = 0, \\ -t \ln(t) & \text{if } t \in \left(0, 1/e^2\right], \\ t + \frac{1}{e^2} & \text{if } t > 1/e^2. \end{cases} \tag{4.14}$$

For more on its usefulness in optimization, see, for example, [7, 14]. We note that $\mathfrak{g}_{-\infty}$ is monotone increasing and there exists $L \geq 1$ such that the following inequalities hold for every $t \in \mathbb{R}_+$ and $M > 0$:[9]

$$|t| \leq \mathfrak{g}_{-\infty}(t), \quad \mathfrak{g}_{-\infty}(2t) \leq L\mathfrak{g}_{-\infty}(t),$$
$$\mathfrak{g}_{-\infty}(Mt) \leq L^{1+|\log_2(M)|}\mathfrak{g}_{-\infty}(t). \tag{4.15}$$

With that, we prove in the next theorem that $\gamma_{\mathbf{z},\eta}$ is positive for $\mathcal{F}_{-\infty}$, which implies that an *entropic error bound* holds.

Theorem 4.2 (Entropic error bound concerning $\mathcal{F}_{-\infty}$) *Let* $z \in K_{\exp}^*$ *with* $z_x = z_z = 0$ *and* $z_y > 0$ *so that* $\{\mathbf{z}\}^{\perp} \cap K_{\exp} = \mathcal{F}_{-\infty}$ *is the two-dimensional face of* $K_{\exp}$. *Let* $\eta > 0$ *and let* $\gamma_{\mathbf{z},\eta}$ *be defined as in* (3.15) *with* $\mathfrak{g} = \mathfrak{g}_{-\infty}$ *in* (4.14). *Then* $\gamma_{\mathbf{z},\eta} \in (0, \infty]$ *and*

---

[9] The third relation in (4.15) is derived from the second relation and the monotonicity of $\mathfrak{g}_{-\infty}$ as follows:
$\mathfrak{g}_{-\infty}(Mt) = \mathfrak{g}_{-\infty}(2^{\log_2 M} t) \leq \mathfrak{g}_{-\infty}(2^{\lceil \log_2 M \rceil} t) \leq L^{\lceil \log_2 M \rceil} \mathfrak{g}_{-\infty}(t) \leq L^{1+|\log_2 M|} \mathfrak{g}_{-\infty}(t).$

$$d(\boldsymbol{q}, \mathcal{F}_{-\infty}) \leq \max\{2, 2\gamma_{z,\eta}^{-1}\} \cdot \mathfrak{g}_{-\infty}(d(\boldsymbol{q}, K_{\exp}))$$

$$\text{whenever } \boldsymbol{q} \in \{\boldsymbol{z}\}^{\perp} \cap B(\eta). \tag{4.16}$$

Proof In view of Lemma 3.12, take any $\bar{\mathbf{v}} \in \mathcal{F}_{-\infty}$ and a sequence $\{\mathbf{v}^k\} \subset \partial K_{\exp} \cap B(\eta) \backslash \mathcal{F}_{-\infty}$ such that

$$\lim_{k\to\infty} \mathbf{v}^k = \lim_{k\to\infty} \mathbf{w}^k = \bar{\mathbf{v}}, \tag{4.17}$$

where $\mathbf{w}^k = P_{\{\boldsymbol{z}\}^{\perp}} \mathbf{v}^k$, $\mathbf{u}^k = P_{\mathcal{F}_{-\infty}} \mathbf{w}^k$, and $\mathbf{w}^k \neq \mathbf{u}^k$. We will show that (3.21b) does not hold for $\mathfrak{g} = \mathfrak{g}_{-\infty}$.

Since $\mathbf{v}^k \notin \mathcal{F}_{-\infty}$, in view of (4.1) and (4.13), we have $v_y^k > 0$ and

$$\mathbf{v}^k = (v_x^k, v_y^k, v_y^k e^{v_x^k/v_y^k}) = (v_y^k \ln(v_z^k/v_y^k), v_y^k, v_z^k), \tag{4.18}$$

where the second representation is obtained by solving for v_x^k from $v_z^k = v_y^k e^{v_x^k/v_y^k} > 0$. Using the second representation in (4.18), we then have

$$\mathbf{w}^k = (v_y^k \ln(v_z^k/v_y^k), 0, v_z^k) \text{ and } \mathbf{u}^k = (0, 0, v_z^k); \tag{4.19}$$

here, we made use of the fact that $\mathbf{w}^k \neq \mathbf{u}^k$, which implies that $v_y^k \ln(v_z^k/v_y^k) > 0$ and thus the resulting formula for $\mathbf{u}^k$. In addition, we also note from $v_y^k \ln(v_z^k/v_y^k) > 0$ (and $v_y^k > 0$) that

$$v_z^k > v_y^k > 0. \tag{4.20}$$

Furthermore, since $\bar{\mathbf{v}} \in \mathcal{F}_{-\infty}$, we see from (4.13) and (4.17) that

$$\lim_{k\to\infty} v_y^k = 0. \tag{4.21}$$

Now, using (4.18), (4.19), (4.21) and the definition of $\mathfrak{g}_{-\infty}$, we see that for k sufficiently large,

$$\frac{\mathfrak{g}_{-\infty}(\|\mathbf{v}^k - \mathbf{w}^k\|)}{\|\mathbf{u}^k - \mathbf{w}^k\|} = \frac{-v_y^k \ln(v_y^k)}{v_y^k \ln(v_z^k/v_y^k)} = \frac{-\ln(v_y^k)}{\ln(v_z^k) - \ln(v_y^k)}. \tag{4.22}$$

We will show that (3.21b) does not hold for $\mathfrak{g} = \mathfrak{g}_{-\infty}$ in each of the following cases.

(I) $\bar{v}_z > 0$.
(II) $\bar{v}_z = 0$.

(I): In this case, we deduce from (4.21) and (4.22) that

$$\lim_{k\to\infty} \frac{\mathfrak{g}_{-\infty}(\|\mathbf{v}^k - \mathbf{w}^k\|)}{\|\mathbf{u}^k - \mathbf{w}^k\|} = 1.$$

Thus (3.21b) does not hold for $\mathfrak{g} = \mathfrak{g}_{-\infty}$.

(II): By passing to a subsequence if necessary, we may assume that $v_z^k < 1$ for all k. This together with (4.20) gives $\frac{\ln(v_z^k)}{\ln(v_y^k)} \in (0, 1)$ for all k. Thus, we conclude from (4.22) that for all k,

$$\frac{\mathfrak{g}_{-\infty}(\|\mathbf{v}^k - \mathbf{w}^k\|)}{\|\mathbf{u}^k - \mathbf{w}^k\|} = \frac{-\ln(v_y^k)}{\ln(v_z^k) - \ln(v_y^k)} = \frac{1}{1 - \frac{\ln(v_z^k)}{\ln(v_y^k)}} > 1.$$

Consequently, (3.21b) also fails for $\mathfrak{g} = \mathfrak{g}_{-\infty}$ in this case.

Having shown that (3.21b) does not hold for $\mathfrak{g} = \mathfrak{g}_{-\infty}$ in any case, we conclude by Lemma 3.12 that $\gamma_{\mathbf{z}, \eta} \in (0, \infty]$. With that, (4.16) follows from Theorem 3.10 and (4.15). $\qquad\square$

Using Theorem 4.2, we can also show weaker Hölderian error bounds.

Corollary 4.3 *Let* $z \in K_{\exp}^*$ *with* $z_x = z_z = 0$ *and* $z_y > 0$ *so that* $\{z\}^\perp \cap K_{\exp} = \mathcal{F}_{-\infty}$ *is the two-dimensional face of* $K_{\exp}$. *Let* $\eta > 0$, $\alpha \in (0, 1)$, *and* $\gamma_{z, \eta}$ *be as in* (3.15) *with* $\mathfrak{g} = |\cdot|^\alpha$. *Then* $\gamma_{z, \eta} \in (0, \infty]$ *and*

$$d(\boldsymbol{q}, \mathcal{F}_{-\infty}) \leq \max\{2\eta^{1-\alpha}, 2\gamma_{z, \eta}^{-1}\} \cdot d(\boldsymbol{q}, K_{\exp})^\alpha \quad \text{whenever } \boldsymbol{q} \in \{z\}^\perp \cap B(\eta).$$

Proof Suppose that $\gamma_{\mathbf{z}, \eta} = 0$ and let sequences $\{\mathbf{v}^k\}$, $\{\mathbf{w}^k\}$, $\{\mathbf{u}^k\}$ be as in Lemma 3.12. Then $\mathbf{v}^k \neq \mathbf{w}^k$ for all k because $\{\mathbf{v}^k\} \subset K_{\exp} \backslash \mathcal{F}_{-\infty}$, $\{\mathbf{w}^k\} \subset \{z\}^\perp$, and $\mathcal{F}_{-\infty} = K_{\exp} \cap \{z\}^\perp$. Since $\mathfrak{g}_{-\infty}(t)/|t|^\alpha \downarrow 0$ as $t \downarrow 0$ we have

$$\liminf_{k \to \infty} \frac{\mathfrak{g}_{-\infty}(\|\mathbf{w}^k - \mathbf{v}^k\|)}{\|\mathbf{w}^k - \mathbf{u}^k\|} = \liminf_{k \to \infty} \frac{\mathfrak{g}_{-\infty}(\|\mathbf{w}^k - \mathbf{v}^k\|)}{\|\mathbf{w}^k - \mathbf{v}^k\|^\alpha} \frac{\|\mathbf{w}^k - \mathbf{v}^k\|^\alpha}{\|\mathbf{w}^k - \mathbf{u}^k\|} = 0,$$

which contradicts Theorem 4.2 because the quantity in (3.15) should be positive for $\mathfrak{g} = \mathfrak{g}_{-\infty}$. $\qquad\square$

Recalling (4.15), we obtain one-step facial residual functions using Theorem 4.2 and Corollary 4.3 in combination with Theorem 3.10, Remark 3.11 and Lemma 3.9.

Corollary 4.4 ($\mathbb{1}$-*FRF concerning* $\mathcal{F}_{-\infty}$) *Let* $z \in K_{\exp}^*$ *be such that* $\{z\}^\perp \cap K_{\exp} = \mathcal{F}_{-\infty}$ *is the two-dimensional face of* $K_{\exp}$. *Let* $\mathfrak{g} = \mathfrak{g}_{-\infty}$ *in* (4.14) *or* $\mathfrak{g} = |\cdot|^\alpha$ *for* $\alpha \in (0, 1)$.

Let $\kappa_{z, t}$ *be defined as in* (3.16). *Then the function* $\psi_{\mathcal{K}, z} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ *given by*

$$\psi_{\mathcal{K}, z}(\epsilon, t) := \max\{\epsilon, \epsilon/\|z\|\} + \kappa_{z, t}\mathfrak{g}(\epsilon + \max\{\epsilon, \epsilon/\|z\|\})$$

is a $\mathbb{1}$-*FRF for* $K_{\exp}$ *and* z. *In particular, there exist* $\kappa > 0$ *and a nonnegative monotone nondecreasing function* $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ *such that the function* $\hat{\psi}_{\mathcal{K}, z}$ *given by* $\hat{\psi}_{\mathcal{K}, z}(\epsilon, t) := \kappa\epsilon + \rho(t)\mathfrak{g}(\epsilon)$ *is a* $\mathbb{1}$-*FRF for* $K_{\exp}$ *and* z.

4.2.2 $\mathcal{F}_\beta$: the family of one-dimensional faces $\beta \in \mathbb{R}$

Recall from Proposition 4.1 that for each $\beta \in \mathbb{R}$,

$$\mathcal{F}_\beta := \left\{ \left(-\beta y + y, \, y, \, e^{1-\beta} y \right) \, \middle| \, y \in [0, \infty) \right\}$$

is a one-dimensional face of $K_{\exp}$. We will now show that for $\mathcal{F}_\beta$, $\beta \in \mathbb{R}$, the $\gamma_{\mathbf{z}, \eta}$ defined in Theorem 3.10 is positive when $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$. Our discussion will be centered around the following quantities, which were also defined and used in the proof of Theorem 3.10. Specifically, for $\mathbf{z} \in K_{\exp}^*$ such that $\mathcal{F}_\beta = K_{\exp} \cap \{\mathbf{z}\}^\perp$, we let $\mathbf{v} \in \partial K_{\exp} \cap B(\eta) \backslash \mathcal{F}_\beta$ and define

$$\mathbf{w} := P_{\{\mathbf{z}\}^\perp} \mathbf{v} \quad \text{and} \quad \mathbf{u} := P_{\mathcal{F}_\beta} \mathbf{w}. \tag{4.23}$$

We first note the following three important vectors:

$$\widehat{\mathbf{z}} := \begin{bmatrix} 1 \\ \beta \\ -e^{\beta-1} \end{bmatrix}, \quad \widehat{\mathbf{f}} = \begin{bmatrix} 1-\beta \\ 1 \\ e^{1-\beta} \end{bmatrix}, \quad \widehat{\mathbf{p}} = \begin{bmatrix} \beta e^{1-\beta} + e^{\beta-1} \\ -e^{1-\beta} - (1-\beta)e^{\beta-1} \\ \beta^2 - \beta + 1 \end{bmatrix}. \tag{4.24}$$

Note that $\widehat{\mathbf{z}}$ is parallel to $\mathbf{z}$ in (4.6) (recall that $z_x < 0$ for $\mathcal{F}_\beta$, where $\beta := \frac{z_y}{z_x} \in \mathbb{R}$), $\mathcal{F}_\beta$ is the conic hull of $\{\widehat{\mathbf{f}}\}$ according to Proposition 4.1, $\langle \widehat{\mathbf{z}}, \widehat{\mathbf{f}} \rangle = 0$ and $\widehat{\mathbf{p}} = \widehat{\mathbf{z}} \times \widehat{\mathbf{f}} \neq \mathbf{0}$. These three *nonzero* vectors form a mutually orthogonal set. The next lemma represents $\|\mathbf{u} - \mathbf{w}\|$ and $\|\mathbf{w} - \mathbf{v}\|$ in terms of inner products of $\mathbf{v}$ with these vectors, whenever possible.

Lemma 4.5 *Let $\beta \in \mathbb{R}$ and $z \in K_{\exp}^*$ with $z_x < 0$ such that $\mathcal{F}_\beta = \{z\}^\perp \cap K_{\exp}$ is a one-dimensional face of $K_{\exp}$. Let $\eta > 0$, $\mathbf{v} \in \partial K_{\exp} \cap B(\eta) \backslash \mathcal{F}_\beta$ and define $\mathbf{w}$ and $\mathbf{u}$ as in (4.23). Let $\{\widehat{z}, \widehat{f}, \widehat{p}\}$ be as in (4.24). Then*

$$\|\mathbf{w} - \mathbf{v}\| = \frac{|\langle \widehat{\mathbf{z}}, \mathbf{v} \rangle|}{\|\widehat{\mathbf{z}}\|} \quad \text{and} \quad \|\mathbf{w} - \mathbf{u}\| = \begin{cases} \frac{|\langle \widehat{\mathbf{p}}, \mathbf{v} \rangle|}{\|\widehat{\mathbf{p}}\|} & \text{if } \langle \widehat{\mathbf{f}}, \mathbf{v} \rangle \geq 0, \\ \|\mathbf{w}\| & \text{otherwise.} \end{cases}$$

Moreover, when $\langle \widehat{\mathbf{f}}, \mathbf{v} \rangle \geq 0$, we have $\mathbf{u} = P_{\mathrm{span} \, \mathcal{F}_\beta} \mathbf{w}$.

Proof Since $\{\widehat{\mathbf{z}}, \widehat{\mathbf{f}}, \widehat{\mathbf{p}}\}$ is orthogonal, one can decompose $\mathbf{v}$ as

$$\mathbf{v} = \lambda_1 \widehat{\mathbf{z}} + \lambda_2 \widehat{\mathbf{f}} + \lambda_3 \widehat{\mathbf{p}}, \tag{4.25}$$

with

$$\lambda_1 = \langle \widehat{\mathbf{z}}, \mathbf{v} \rangle / \|\widehat{\mathbf{z}}\|^2, \quad \lambda_2 = \langle \widehat{\mathbf{f}}, \mathbf{v} \rangle / \|\widehat{\mathbf{f}}\|^2 \quad \text{and} \quad \lambda_3 = \langle \widehat{\mathbf{p}}, \mathbf{v} \rangle / \|\widehat{\mathbf{p}}\|^2. \tag{4.26}$$

Also, since $\widehat{\mathbf{z}}$ is parallel to $\mathbf{z}$, we must have $\mathbf{w} = \lambda_2 \widehat{\mathbf{f}} + \lambda_3 \widehat{\mathbf{p}}$. Thus, it holds that $\|\mathbf{w} - \mathbf{v}\| = |\lambda_1| \|\widehat{\mathbf{z}}\|$ and the first conclusion follows from this and (4.26).

Next, we have $\mathbf{u} = \hat{t}\,\widehat{\mathbf{f}}$, where

$$\hat{t} = \mathrm{argmin}_{t \geq 0}\|\mathbf{w} - t\widehat{\mathbf{f}}\| = \begin{cases} \frac{\langle \mathbf{w}, \widehat{\mathbf{f}}\rangle}{\|\widehat{\mathbf{f}}\|^2} & \text{if } \langle \widehat{\mathbf{f}}, \mathbf{w}\rangle \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

Moreover, observe from (4.25) that $\langle \widehat{\mathbf{f}}, \mathbf{w}\rangle = \langle \widehat{\mathbf{f}}, \mathbf{v} - \lambda_1 \widehat{\mathbf{z}}\rangle = \langle \widehat{\mathbf{f}}, \mathbf{v}\rangle$. These mean that when $\langle \widehat{\mathbf{f}}, \mathbf{v}\rangle < 0$, we have $\mathbf{u} = 0$, while when $\langle \widehat{\mathbf{f}}, \mathbf{v}\rangle \geq 0$, we have $\mathbf{u} = \frac{\langle \mathbf{w}, \widehat{\mathbf{f}}\rangle}{\|\widehat{\mathbf{f}}\|^2}\widehat{\mathbf{f}} = P_{\mathrm{span}\,\mathcal{F}_\beta}\mathbf{w}$ and

$$\|\mathbf{w} - \mathbf{u}\| = \left\|\mathbf{w} - \frac{\langle \mathbf{w}, \widehat{\mathbf{f}}\rangle}{\|\widehat{\mathbf{f}}\|^2}\widehat{\mathbf{f}}\right\| = |\lambda_3|\|\widehat{\mathbf{p}}\| = |\langle \widehat{\mathbf{p}}, \mathbf{v}\rangle|/\|\widehat{\mathbf{p}}\|,$$

where the second and the third equalities follow from (4.25), (4.26), and the fact that $\mathbf{w} = \lambda_2 \widehat{\mathbf{f}} + \lambda_3 \widehat{\mathbf{p}}$. This completes the proof. $\qquad\square$

We now prove our main theorem in this section.

Theorem 4.6 (Hölderian error bound concerning $\mathcal{F}_\beta$, $\beta \in \mathbb{R}$) *Let $\beta \in \mathbb{R}$ and $z \in K_{\exp}^*$ with $z_x < 0$ such that $\mathcal{F}_\beta = \{z\}^\perp \cap K_{\exp}$ is a one-dimensional face of $K_{\exp}$. Let $\eta > 0$ and let $\gamma_{z,\eta}$ be defined as in (3.15) with $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$. Then $\gamma_{z,\eta} \in (0, \infty]$ and*

$$d(q, \mathcal{F}_\beta) \leq \max\left\{2\sqrt{\eta}, 2\gamma_{z,\eta}^{-1}\right\} \cdot d(q, K_{\exp})^{\frac{1}{2}} \quad \text{whenever } q \in \{z\}^\perp \cap B(\eta).$$

Proof In view of Lemma 3.12, take any $\bar{\mathbf{v}} \in \mathcal{F}_\beta$ and a sequence $\{\mathbf{v}^k\} \subset \partial K_{\exp} \cap B(\eta)\backslash\mathcal{F}_\beta$ such that

$$\lim_{k\to\infty}\mathbf{v}^k = \lim_{k\to\infty}\mathbf{w}^k = \bar{\mathbf{v}}, \tag{4.27}$$

where $\mathbf{w}^k = P_{\{z\}^\perp}\mathbf{v}^k$, $\mathbf{u}^k = P_{\mathcal{F}_\beta}\mathbf{w}^k$, and $\mathbf{w}^k \neq \mathbf{u}^k$. We will show that (3.21b) does not hold for $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$.

We first suppose that $\mathbf{v}^k \in \mathcal{F}_{-\infty}$ infinitely often. By extracting a subsequence if necessary, we may assume that $\mathbf{v}^k \in \mathcal{F}_{-\infty}$ for all k. From the definition of $\mathcal{F}_{-\infty}$ in (4.13), we have $v_x^k \leq 0$, $v_y^k = 0$ and $v_z^k \geq 0$. Thus, recalling the definition of $\widehat{\mathbf{z}}$ in (4.24), it holds that

$$\begin{aligned} |\langle \widehat{\mathbf{z}}, \mathbf{v}^k\rangle| &= |v_x^k - v_z^k e^{\beta-1}| = -v_x^k + v_z^k e^{\beta-1} = |v_x^k| + e^{\beta-1}|v_z^k| \\ &\geq \min\{1, e^{\beta-1}\}\sqrt{|v_x^k|^2 + |v_z^k|^2} = \min\{1, e^{\beta-1}\}\|\mathbf{v}^k\|, \end{aligned} \tag{4.28}$$

where the last equality holds because $v_y^k = 0$. Next, using properties of projections onto subspaces, we have $\|\mathbf{w}^k\| \leq \|\mathbf{v}^k\|$. This together with Lemma 4.5 and (4.28) shows that

$$\|\mathbf{w}^k - \mathbf{u}^k\| = \mathrm{d}(\mathbf{w}^k, \mathcal{F}_\beta) \le \|\mathbf{w}^k\| \le \|\mathbf{v}^k\| \le \frac{1}{\min\{1, e^{\beta-1}\}}|\langle \widehat{\mathbf{z}}, \mathbf{v}^k \rangle|$$

$$= \frac{\|\widehat{\mathbf{z}}\|}{\min\{1, e^{\beta-1}\}}\|\mathbf{w}^k - \mathbf{v}^k\|.$$

Since $\mathbf{w}^k \to \bar{\mathbf{v}}$ and $\mathbf{v}^k \to \bar{\mathbf{v}}$, the above display shows that (3.21b) does not hold for $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$ in this case.

Next, suppose that $\mathbf{v}^k \notin \mathcal{F}_{-\infty}$ for all large k instead. By passing to a subsequence, we assume that $\mathbf{v}^k \notin \mathcal{F}_{-\infty}$ for all k. In view of (4.1) and (4.13), this means in particular that

$$\mathbf{v}^k = (v_x^k, v_y^k, v_y^k e^{v_x^k/v_y^k}) \quad \text{and} \quad v_y^k > 0 \ \text{ for all } k. \tag{4.29}$$

We consider two cases and show that (3.21b) does not hold for $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$ in either of them:

(I) $\langle \widehat{\mathbf{f}}, \mathbf{v}^k \rangle \ge 0$ infinitely often;
(II) $\langle \widehat{\mathbf{f}}, \mathbf{v}^k \rangle < 0$ for all large k.

(I): Since $\langle \widehat{\mathbf{f}}, \mathbf{v}^k \rangle \ge 0$ infinitely often, by extracting a subsequence if necessary, we assume that $\langle \widehat{\mathbf{f}}, \mathbf{v}^k \rangle \ge 0$ for all k. Now, consider the following functions:

$$h_1(\zeta) := \zeta + \beta - e^{\beta+\zeta-1},$$
$$h_2(\zeta) := (\beta e^{1-\beta} + e^{\beta-1})\zeta - e^{1-\beta} - (1 - \beta)e^{\beta-1} + (\beta^2 - \beta + 1)e^\zeta.$$

Using these functions, Lemma 4.5, (4.24) and (4.29), one can see immediately that

$$\|\mathbf{w}^k - \mathbf{v}^k\| = \frac{|\langle \widehat{\mathbf{z}}, \mathbf{v}^k \rangle|}{\|\widehat{\mathbf{z}}\|} = \frac{v_y^k |h_1(v_x^k/v_y^k)|}{\|\widehat{\mathbf{z}}\|},$$

$$\|\mathbf{w}^k - \mathbf{u}^k\| = \frac{|\langle \widehat{\mathbf{p}}, \mathbf{v}^k \rangle|}{\|\widehat{\mathbf{p}}\|} = \frac{v_y^k |h_2(v_x^k/v_y^k)|}{\|\widehat{\mathbf{p}}\|}. \tag{4.30}$$

Note that h_1 is zero **if and only if** $\zeta = 1 - \beta$. Furthermore, we have $h_1'(1 - \beta) = 0$ and $h_1''(1 - \beta) = -1$. Then, considering the Taylor expansion of h_1 around $1 - \beta$ we have

$$h_1(\zeta) = 1 + (\zeta + \beta - 1) - e^{\beta+\zeta-1} = -\frac{(\zeta + \beta - 1)^2}{2} + O(|\zeta + \beta - 1|^3) \quad \text{as } \zeta \to 1 - \beta.$$

Also, one can check that $h_2(1 - \beta) = 0$ and that

$$h_2'(1 - \beta) = \beta e^{1-\beta} + e^{\beta-1} + (\beta^2 - \beta + 1)e^{1-\beta} = e^{\beta-1} + (\beta^2 + 1)e^{1-\beta} > 0.$$

Therefore, we have the following Taylor expansion of h_2 around $1 - \beta$:

$$h_2(\zeta) = (e^{\beta-1} + (\beta^2 + 1)e^{1-\beta})(\zeta + \beta - 1)$$
$$+ O(|\zeta + \beta - 1|^2)$$
$$\text{as } \zeta \to 1 - \beta. \tag{4.31}$$

Thus, using the Taylor expansions of h_1 and h_2 at $1 - \beta$ we have

$$\lim_{\zeta \to 1-\beta} \frac{|h_1(\zeta)|^{\frac{1}{2}}}{|h_2(\zeta)|} = \frac{1}{\sqrt{2}(e^{\beta-1} + (\beta^2 + 1)e^{1-\beta})} > 0. \tag{4.32}$$

Hence, there exist $C_h > 0$ and $\epsilon > 0$ so that

$$|h_1(\zeta)|^{\frac{1}{2}} \geq C_h |h_2(\zeta)| \quad \text{whenever } |\zeta - (1 - \beta)| \leq \epsilon. \tag{4.33}$$

Next, consider the following function[10]

$$H(\zeta) := \begin{cases} \frac{|h_1(\zeta)|}{|h_2(\zeta)|} & \text{if } |\zeta - (1 - \beta)| \geq \epsilon, h_2(\zeta) \neq 0, \\ \infty & \text{otherwise.} \end{cases}$$

Then it is easy to check that H is proper closed and is never zero. Moreover, by direct computation, we have $\lim_{\zeta \to \infty} H(\zeta) = \frac{e^{\beta-1}}{\beta^2 - \beta + 1} > 0$ and

$$\lim_{\zeta \to -\infty} H(\zeta) = \begin{cases} |\beta e^{1-\beta} + e^{\beta-1}|^{-1} & \text{if } \beta e^{1-\beta} + e^{\beta-1} \neq 0, \\ \infty & \text{otherwise.} \end{cases}$$

Thus, we deduce that $\inf H > 0$.

Now, if it happens that $|v_x^k/v_y^k - (1-\beta)| > \epsilon$ for all large k, upon letting $\zeta_k := v_x^k/v_y^k$, we have from (4.30) that for all large k,

$$\frac{\|\mathbf{w}^k - \mathbf{v}^k\|}{\|\mathbf{w}^k - \mathbf{u}^k\|} = \frac{\|\widehat{\mathbf{p}}\|}{\|\widehat{\mathbf{z}}\|} \frac{|h_1(\zeta_k)|}{|h_2(\zeta_k)|} = \frac{\|\widehat{\mathbf{p}}\|}{\|\widehat{\mathbf{z}}\|} H(\zeta_k) \geq \frac{\|\widehat{\mathbf{p}}\|}{\|\widehat{\mathbf{z}}\|} \inf H > 0, \tag{4.34}$$

where the second equality holds because of the definition of H and the facts that $\mathbf{w}^k \neq \mathbf{u}^k$ (so that $h_2(\zeta_k) \neq 0$ by (4.30)) and $|v_x^k/v_y^k - (1 - \beta)| > \epsilon$ for all large k.

On the other hand, if it holds that $|v_x^k/v_y^k - (1-\beta)| \leq \epsilon$ infinitely often, then by extracting a further subsequence, we may assume that $|v_x^k/v_y^k - (1 - \beta)| \leq \epsilon$ for all k. Upon letting $\zeta_k := v_x^k/v_y^k$, we have from (4.30) that

$$\frac{\|\mathbf{w}^k - \mathbf{u}^k\|}{\|\mathbf{w}^k - \mathbf{v}^k\|^{\frac{1}{2}}} = \frac{\sqrt{v_y^k \|\widehat{\mathbf{z}}\|}}{\|\widehat{\mathbf{p}}\|} \frac{|h_2(\zeta_k)|}{|h_1(\zeta_k)|^{\frac{1}{2}}} \leq \frac{\sqrt{v_y^k \|\widehat{\mathbf{z}}\|}}{C_h \|\widehat{\mathbf{p}}\|} \leq \frac{\sqrt{\eta \|\widehat{\mathbf{z}}\|}}{C_h \|\widehat{\mathbf{p}}\|}, \tag{4.35}$$

---

[10] Notice that this function is well defined because h_1 is zero only at $1 - \beta$ and thus we will not end up with $\frac{0}{0}$.

where the first inequality holds thanks to $|v_x^k/v_y^k - (1 - \beta)| \le \epsilon$ for all k, (4.33) and the fact that $\mathbf{w}^k \ne \mathbf{u}^k$ (so that $h_2(\zeta_k) \ne 0$ and hence $h_1(\zeta_k) \ne 0$), and the second inequality holds because $\mathbf{v}^k \in B(\eta)$.

Using (4.34) and (4.35) together with (4.27), we see that (3.21b) does not hold for $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$. This concludes case (I).

(II): By passing to a subsequence, we may assume that $\langle \widehat{\mathbf{f}}, \mathbf{v}^k \rangle < 0$ for all k. Then we see from (4.24) and (4.29) that

$$(1 - \beta)\frac{v_x^k}{v_y^k} + 1 + e^{1-\beta}e^{v_x^k/v_y^k} = \frac{1}{v_y^k}\langle \widehat{\mathbf{f}}, \mathbf{v}^k \rangle < 0.$$

Using this together with the fact that $(1 - \beta)^2 + 1 + e^{2(1-\beta)} > 0$, we deduce that there exists $\epsilon > 0$ so that

$$\left| \frac{v_x^k}{v_y^k} - (1 - \beta) \right| \ge \epsilon \quad \text{for all } k. \tag{4.36}$$

Now, consider the following function

$$G(\zeta) := \frac{|\zeta + \beta - e^{\beta+\zeta-1}|}{\sqrt{\zeta^2 + 1 + e^{2\zeta}}}.$$

Then G is continuous and is zero if and only if $\zeta = 1 - \beta$. Moreover, by direct computation, we have $\lim_{\zeta \to \infty} G(\zeta) = e^{\beta-1} > 0$ and $\lim_{\zeta \to -\infty} G(\zeta) = 1 > 0$. Thus, it follows that

$$\underline{G} := \inf_{|\zeta+\beta-1| \ge \epsilon} G(\zeta) > 0. \tag{4.37}$$

Finally, since $\langle \widehat{\mathbf{f}}, \mathbf{v}^k \rangle < 0$ for all k, we see that

$$\frac{\|\mathbf{w}^k - \mathbf{v}^k\|}{\|\mathbf{w}^k - \mathbf{u}^k\|} \overset{(a)}{\ge} \frac{\|\mathbf{w}^k - \mathbf{v}^k\|}{\|\mathbf{v}^k\|} \overset{(b)}{=} \frac{|\widehat{z}_x v_x^k + \widehat{z}_y v_y^k + \widehat{z}_z v_y^k e^{v_x^k/v_y^k}|}{\|\widehat{\mathbf{z}}\|\sqrt{(v_x^k)^2 + (v_y^k)^2 + (v_y^k)^2 e^{2v_x^k/v_y^k}}}$$

$$\overset{(c)}{=} \frac{|\widehat{z}_x \zeta_k + \widehat{z}_y + \widehat{z}_z e^{\zeta_k}|}{\|\widehat{\mathbf{z}}\|\sqrt{(\zeta_k)^2 + 1 + e^{2\zeta_k}}} \overset{(d)}{=} \frac{1}{\|\widehat{\mathbf{z}}\|} G(\zeta_k) \overset{(e)}{\ge} \frac{1}{\|\widehat{\mathbf{z}}\|} \underline{G} > 0,$$

where (a) follows from $\|\mathbf{w}^k - \mathbf{u}^k\| = \|\mathbf{w}^k\|$ (see Lemma 4.5) and $\|\mathbf{w}^k\| \le \|\mathbf{v}^k\|$ (because $\mathbf{w}^k$ is the projection of $\mathbf{v}^k$ onto a subspace), (b) follows from Lemma 4.5 and (4.29), (c) holds because $v_y^k > 0$ (see (4.29)) and we defined $\zeta_k := v_x^k/v_y^k$, (d) follows from (4.24) and the definition of G, and (e) follows from (4.36) and (4.37). The above together with (4.27) shows that (3.21b) does not hold for $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$, which is what we wanted to show in case (II).

Summarizing the above cases, we conclude that there does not exist the sequence $\{\mathbf{v}^k\}$ and its associates so that (3.21b) holds for $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$. By Lemma 3.12, it must then hold that $\gamma_{\mathbf{z},\eta} \in (0, \infty]$ and we have the desired error bound in view of Theorem 3.10. This completes the proof. $\qquad\square$

Combining Theorems 4.6, 3.10 and Lemma 3.9, and using the observation that $\gamma_{\mathbf{z},0} = \infty$ (see (3.15)), we obtain a one-step facial residual function in the following corollary.

Corollary 4.7 ($\mathbb{1}$-*FRF concerning $\mathcal{F}_\beta$, $\beta \in \mathbb{R}$*) *Let $\beta \in \mathbb{R}$ and $\mathbf{z} \in K^*_{\exp}$ with $z_x < 0$ such that $\mathcal{F}_\beta = \{\mathbf{z}\}^\perp \cap K_{\exp}$ is a one-dimensional face of $K_{\exp}$. Let $\kappa_{z,t}$ be defined as in (3.16) with $\mathfrak{g} = |\cdot|^{\frac{1}{2}}$. Then the function $\psi_{\mathcal{K},z} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ given by*

$$\psi_{\mathcal{K},z}(\epsilon, t) := \max\{\epsilon, \epsilon/\|z\|\} + \kappa_{z,t}(\epsilon + \max\{\epsilon, \epsilon/\|z\|\})^{\frac{1}{2}}$$

is a $\mathbb{1}$-FRF for $K_{\exp}$. In particular, there exist $\kappa > 0$ and a nonnegative monotone nondecreasing function $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ such that the function $\hat{\psi}_{\mathcal{K},z}$ given by $\hat{\psi}_{\mathcal{K},z}(\epsilon, t) := \kappa\epsilon + \rho(t)\sqrt{\epsilon}$ is a $\mathbb{1}$-FRF for $K_{\exp}$ and z.

4.2.3 $\mathcal{F}_\infty$: the exceptional one-dimensional face

Recall the special one-dimensional face of $K_{\exp}$ defined by

$$\mathcal{F}_\infty := \{(x, 0, 0) \mid x \leq 0\}.$$

We first show that we have a Lipschitz error bound for any exposing normal vectors $\mathbf{z} = (0, z_y, z_z)$ with $z_y > 0$ and $z_z > 0$.

Theorem 4.8 (Lipschitz error bound concerning $\mathcal{F}_\infty$) *Let $\mathbf{z} \in K^*_{\exp}$ with $z_x = 0$, $z_y > 0$ and $z_z > 0$ so that $\{\mathbf{z}\}^\perp \cap K_{\exp} = \mathcal{F}_\infty$. Let $\eta > 0$ and let $\gamma_{z,\eta}$ be defined as in (3.15) with $\mathfrak{g} = |\cdot|$. Then $\gamma_{z,\eta} \in (0, \infty]$ and*

$$d(\mathbf{q}, \mathcal{F}_\infty) \leq \max\{2, 2\gamma_{z,\eta}^{-1}\} \cdot d(\mathbf{q}, K_{\exp}) \quad \text{whenever } \mathbf{q} \in \{\mathbf{z}\}^\perp \cap B(\eta).$$

Proof Without loss of generality, upon scaling, we may assume that $\mathbf{z} = (0, a, 1)$ for some $a > 0$. Similarly as in the proof of Theorem 4.6, we will consider the following vectors:

$$\widetilde{\mathbf{z}} := \begin{bmatrix} 0 \\ a \\ 1 \end{bmatrix}, \quad \widetilde{\mathbf{f}} := \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}, \quad \widetilde{\mathbf{p}} := \begin{bmatrix} 0 \\ 1 \\ -a \end{bmatrix}.$$

Here, $\mathcal{F}_\infty$ is the conical hull of $\widetilde{\mathbf{f}}$ (see (4.12)), and $\widetilde{\mathbf{p}}$ is constructed so that $\{\widetilde{\mathbf{z}}, \widetilde{\mathbf{f}}, \widetilde{\mathbf{p}}\}$ is orthogonal.

Now, let $\mathbf{v} \in \partial K_{\exp} \cap B(\eta) \backslash \mathcal{F}_\infty$, $\mathbf{w} = P_{\{\mathbf{z}\}^\perp} \mathbf{v}$ and $\mathbf{u} = P_{\mathcal{F}_\infty} \mathbf{w}$ with $\mathbf{u} \neq \mathbf{w}$. Then, as in Lemma 4.5, by decomposing $\mathbf{v}$ as a linear combination of $\{\widetilde{\mathbf{z}}, \widetilde{\mathbf{f}}, \widetilde{\mathbf{p}}\}$, we have

$$\|\mathbf{w} - \mathbf{v}\| = \frac{|\langle \widetilde{\mathbf{z}}, \mathbf{v} \rangle|}{\|\widetilde{\mathbf{z}}\|} \quad \text{and} \quad \|\mathbf{w} - \mathbf{u}\| = \begin{cases} \frac{|\langle \widetilde{\mathbf{p}}, \mathbf{v} \rangle|}{\|\widetilde{\mathbf{p}}\|} & \text{if } \langle \widetilde{\mathbf{f}}, \mathbf{v} \rangle \geq 0, \\ \|\mathbf{w}\| & \text{otherwise.} \end{cases} \tag{4.38}$$

We consider the following cases for estimating $\gamma_{\mathbf{z}, \eta}$.

(I) $\mathbf{v} \in \mathcal{F}_{-\infty} \backslash \mathcal{F}_\infty$;
(II) $\mathbf{v} \notin \mathcal{F}_{-\infty}$ with $v_x \leq 0$;
(III) $\mathbf{v} \notin \mathcal{F}_{-\infty}$ with $v_x > 0$.

(I): In this case, $\mathbf{v} = (v_x, 0, v_z)$ with $v_x \leq 0 \leq v_z$; see (4.13). Then $\langle \widetilde{\mathbf{f}}, \mathbf{v} \rangle = -v_x \geq 0$ and $|\langle \widetilde{\mathbf{z}}, \mathbf{v} \rangle| = |v_z| = \frac{1}{a}|\langle \widetilde{\mathbf{p}}, \mathbf{v} \rangle|$. Consequently, we have from (4.38) that

$$\|\mathbf{w} - \mathbf{v}\| = \frac{|\langle \widetilde{\mathbf{z}}, \mathbf{v} \rangle|}{\|\widetilde{\mathbf{z}}\|} = \frac{|\langle \widetilde{\mathbf{p}}, \mathbf{v} \rangle|}{a\|\widetilde{\mathbf{z}}\|} = \frac{\|\widetilde{\mathbf{p}}\|}{a\|\widetilde{\mathbf{z}}\|}\|\mathbf{w} - \mathbf{u}\|.$$

(II): In this case, in view of (4.1) and (4.13), we have $\mathbf{v} = (v_x, v_y, v_y e^{v_x/v_y})$ with $v_x \leq 0$ and $v_y > 0$. Then $\langle \widetilde{\mathbf{f}}, \mathbf{v} \rangle = -v_x \geq 0$. Moreover, since $v_y > 0$, we have

$$\begin{aligned}
|\langle \widetilde{\mathbf{z}}, \mathbf{v} \rangle| &= |av_y + v_y e^{v_x/v_y}| = av_y + v_y e^{v_x/v_y} \geq \min\{1, a\}(v_y + v_y e^{v_x/v_y}) \\
&\geq \frac{\min\{1, a\}}{\max\{1, a\}}(v_y + av_y e^{v_x/v_y}) \geq \frac{\min\{1, a\}}{\max\{1, a\}}|v_y - av_y e^{v_x/v_y}| \\
&= \frac{\min\{1, a\}}{\max\{1, a\}}|\langle \widetilde{\mathbf{p}}, \mathbf{v} \rangle|.
\end{aligned}$$

Using (4.38), we then obtain that $\|\mathbf{w} - \mathbf{v}\| \geq \frac{\min\{1,a\}\|\widetilde{\mathbf{p}}\|}{\max\{1,a\}\|\widetilde{\mathbf{z}}\|}\|\mathbf{w} - \mathbf{u}\|$.

(III): In this case, in view of (4.1) and (4.13), $\mathbf{v} = (v_x, v_y, v_y e^{v_x/v_y})$ with $v_x > 0$ and $v_y > 0$. Then $\langle \widetilde{\mathbf{f}}, \mathbf{v} \rangle = -v_x < 0$ and hence $\|\mathbf{w} - \mathbf{u}\| = \|\mathbf{w}\| \leq \|\mathbf{v}\|$, where the equality follows from (4.38) and the inequality holds because $\mathbf{w}$ is the projection of $\mathbf{v}$ onto a subspace. Since $v_y > 0$, we have

$$\begin{aligned}
|\langle \widetilde{\mathbf{z}}, \mathbf{v} \rangle| &= |av_y + v_y e^{v_x/v_y}| = av_y + 0.5v_y e^{v_x/v_y} + 0.5v_y e^{v_x/v_y} \\
&\overset{(a)}{\geq} av_y + 0.5v_y(1 + v_x/v_y) + 0.5v_y e^{v_x/v_y} \\
&\geq 0.5v_y + 0.5v_x + 0.5v_y e^{v_x/v_y} \\
&= 0.5\|\mathbf{v}\|_1,
\end{aligned}$$

where we used $v_y > 0$ and $e^t \geq 1 + t$ for all t in (a) and $\|\mathbf{v}\|_1$ denotes the 1-norm of $\mathbf{v}$. Combining this with (4.38) and the fact that $\|\mathbf{w}\| \leq \|\mathbf{v}\|$, we see that

$$\|\mathbf{w} - \mathbf{v}\| = \frac{|\langle \widetilde{\mathbf{z}}, \mathbf{v} \rangle|}{\|\widetilde{\mathbf{z}}\|} \geq \frac{\|\mathbf{v}\|_1}{2\|\widetilde{\mathbf{z}}\|} \geq \frac{\|\mathbf{v}\|}{2\|\widetilde{\mathbf{z}}\|} \geq \frac{\|\mathbf{w}\|}{2\|\widetilde{\mathbf{z}}\|} = \frac{\|\mathbf{w} - \mathbf{u}\|}{2\|\widetilde{\mathbf{z}}\|}.$$

Summarizing the three cases, we conclude that

$$\gamma_{\mathbf{z},\eta} \geq \min \left\{ \frac{\|\widetilde{\mathbf{p}}\|}{a\|\widetilde{\mathbf{z}}\|}, \frac{\min\{1,a\}\|\widetilde{\mathbf{p}}\|}{\max\{1,a\}\|\widetilde{\mathbf{z}}\|}, \frac{1}{2\|\widetilde{\mathbf{z}}\|} \right\} > 0.$$

In view of Theorem 3.10, we have the desired error bound. This completes the proof.
□

We next turn to the supporting hyperplane defined by $\mathbf{z} = (0, 0, z_z)$ for some $z_z > 0$ and so $\{\mathbf{z}\}^\perp$ is the xy-plane. The following lemma demonstrates that the Hölderian-type error bound in the form of (3.16) with $\mathfrak{g} = |\cdot|^\alpha$ for some $\alpha \in (0, 1]$ no longer works in this case.

Lemma 4.9 (Nonexistence of Hölderian error bounds) *Let* $z \in K_{\exp}^*$ *with* $z_x = z_y = 0$ *and* $z_z > 0$ *so that* $\{z\}^\perp \cap K_{\exp} = \mathcal{F}_\infty$. *Let* $\alpha \in (0, 1]$ *and* $\eta > 0$. *Then*

$$\inf_q \left\{ \frac{d(q, K_{\exp})^\alpha}{d(q, \mathcal{F}_\infty)} \;\middle|\; q \in \{z\}^\perp \cap B(\eta) \backslash \mathcal{F}_\infty \right\} = 0.$$

Proof For each $k \in \mathbb{N}$, let $\mathbf{q}^k := (-\frac{\eta}{2}, \frac{\eta}{2k}, 0)$. Then $\mathbf{q}^k \in \{z\}^\perp \cap B(\eta) \backslash \mathcal{F}_\infty$ and we have $d(\mathbf{q}^k, \mathcal{F}_\infty) = \frac{\eta}{2k}$. Moreover, since $(q_x^k, q_y^k, q_y^k e^{q_x^k/q_y^k}) \in K_{\exp}$, we have $d(\mathbf{q}^k, K_{\exp}) \leq q_y^k e^{q_x^k/q_y^k} = \frac{\eta}{2k} e^{-k}$. Then it holds that

$$\frac{d(\mathbf{q}^k, K_{\exp})^\alpha}{d(\mathbf{q}^k, \mathcal{F}_\infty)} \leq \frac{\eta^{\alpha-1}}{2^{\alpha-1}} k^{1-\alpha} e^{-\alpha k} \to 0 \quad \text{as } k \to \infty$$

since $\alpha \in (0, 1]$. This completes the proof.
□

Since a zero-at-zero monotone nondecreasing function of the form $(\cdot)^\alpha$ no longer works, we opt for the following function $\mathfrak{g}_\infty : \mathbb{R}_+ \to \mathbb{R}_+$ that grows faster around $t = 0$:

$$\mathfrak{g}_\infty(t) := \begin{cases} 0 & \text{if } t = 0, \\ -\frac{1}{\ln(t)} & \text{if } 0 < t \leq \frac{1}{e^2}, \\ \frac{1}{4} + \frac{1}{4}e^2 t & \text{if } t > \frac{1}{e^2}. \end{cases} \tag{4.39}$$

Similar to $\mathfrak{g}_{-\infty}$ in (4.14), $\mathfrak{g}_\infty$ is monotone increasing and there exists a constant $\widehat{L} \geq 1$ such that the following inequalities hold for every $t \in \mathbb{R}_+$ and $M > 0$:

$$|t| \leq \mathfrak{g}_\infty(t), \quad \mathfrak{g}_\infty(2t) \leq \widehat{L}\mathfrak{g}_\infty(t), \quad \mathfrak{g}_\infty(Mt) \leq \widehat{L}^{1+|\log_2(M)|}\mathfrak{g}_\infty(t). \tag{4.40}$$

We next show that error bounds in the form of (3.16) holds for $\mathbf{z} = (0, 0, z_z)$, $z_z > 0$, if we use $\mathfrak{g}_\infty$.

Theorem 4.10 (Log-type error bound concerning $\mathcal{F}_\infty$) *Let $z \in K_{\exp}^*$ with $z_x = z_y = 0$ and $z_z > 0$ so that $\{z\}^\perp \cap K_{\exp} = \mathcal{F}_\infty$. Let $\eta > 0$ and let $\gamma_{z,\eta}$ be defined as in (3.15) with $\mathfrak{g} = \mathfrak{g}_\infty$ in (4.39). Then $\gamma_{z,\eta} \in (0, \infty]$ and*

$$d(q, \mathcal{F}_\infty) \le \max\{2, 2\gamma_{z,\eta}^{-1}\} \cdot \mathfrak{g}_\infty(d(q, K_{\exp})) \quad whenever \ q \in \{z\}^\perp \cap B(\eta).$$

Proof Take $\bar{v} \in \mathcal{F}_\infty$ and a sequence $\{v^k\} \subset \partial K_{\exp} \cap B(\eta) \backslash \mathcal{F}_\infty$ such that

$$\lim_{k \to \infty} v^k = \lim_{k \to \infty} w^k = \bar{v},$$

where $w^k = P_{\{z\}^\perp} v^k$, $u^k = P_{\mathcal{F}_\infty} w^k$, and $w^k \ne u^k$. Since $w^k \ne u^k$, in view of (4.12) and (4.13), we must have $v^k \notin \mathcal{F}_{-\infty}$. Then, from (4.1) and (4.12), we have

$$\begin{aligned}
v^k &= (v_x^k, v_y^k, v_y^k e^{v_x^k/v_y^k}) \text{ with } v_y^k > 0, \\
w^k &= (v_x^k, v_y^k, 0) \text{ and } u^k = (\min\{v_x^k, 0\}, 0, 0).
\end{aligned} \tag{4.41}$$

Since $w^k \to \bar{v}$ and $v^k \to \bar{v}$, without loss of generality, by passing to a subsequence if necessary, we assume in addition that $\|w^k - v^k\| \le e^{-2}$ for all k. From (4.41) we conclude that $v^k \ne w^k$, hence $\mathfrak{g}_\infty(\|w^k - v^k\|) = -(\ln \|w^k - v^k\|)^{-1}$.

We consider the following two cases in order to show that (3.21b) does not hold for $\mathfrak{g} = \mathfrak{g}_\infty$:

(I) $\bar{v} \ne 0$;
(II) $\bar{v} = 0$.

(I): In this case, we have $\bar{v} = (\bar{v}_x, 0, 0)$ for some $\bar{v}_x < 0$. This implies that $v_x^k < 0$ for all large k. Hence, we have from (4.41) that for all large k,

$$\begin{aligned}
\frac{\mathfrak{g}_\infty(\|w^k - v^k\|)}{\|w^k - u^k\|} = \frac{-(\ln \|w^k - v^k\|)^{-1}}{\|w^k - u^k\|} &= -\frac{1}{v_y^k \left(v_x^k/v_y^k + \ln v_y^k\right)} \\
&= -\frac{1}{v_y^k \ln v_y^k + v_x^k} \to -\frac{1}{\bar{v}_x} > 0
\end{aligned}$$

since $v_y^k \to 0$ and $v_x^k \to \bar{v}_x < 0$. This shows that (3.21b) does not hold for $\mathfrak{g} = \mathfrak{g}_\infty$.

(II): If $v_x^k \le 0$ infinitely often, by extracting a subsequence, we assume that $v_x^k \le 0$ for all k. Since $w^k \ne u^k$ (and $w^k \ne v^k$), we note from (4.41) that

$$-\frac{1}{v_y^k \ln v_y^k + v_x^k} = \frac{-(\ln \|w^k - v^k\|)^{-1}}{\|w^k - u^k\|} \in (0, \infty) \text{ for all } k.$$

Since $\{-(v_y^k \ln v_y^k + v_x^k)\}$ is a positive sequence and it converges to zero as $(v_x^k, v_y^k) \to 0$, it follows that $\lim_{k \to \infty} \frac{-(\ln \|w^k - v^k\|)^{-1}}{\|w^k - u^k\|} = \infty$. This shows that (3.21b) does not hold for $\mathfrak{g} = \mathfrak{g}_\infty$.

Now, it remains to consider the case that $v_x^k > 0$ for all large k. By passing to a subsequence if necessary, we assume that $v_x^k > 0$ for all k. By solving for v_x^k from $v_z^k = v_y^k e^{v_x^k/v_y^k} > 0$ and noting (4.41), we obtain that

$$
\begin{aligned}
\mathbf{v}^k &= (v_y^k \ln(v_z^k/v_y^k), v_y^k, v_z^k) \text{ with } v_y^k > 0, \\
\mathbf{w}^k &= (v_y^k \ln(v_z^k/v_y^k), v_y^k, 0) \text{ and } \mathbf{u}^k = (0, 0, 0).
\end{aligned}
\tag{4.42}
$$

Also, we note from $v_x^k = v_y^k \ln(v_z^k/v_y^k) > 0$, $v_y^k > 0$ and the monotonicity of $\ln(\cdot)$ that for all k,

$$
v_z^k > v_y^k > 0.
\tag{4.43}
$$

Next consider the function $h(t) := \frac{1}{t}\sqrt{1 + (\ln t)^2}$ on $[1, \infty)$. Then h is continuous and positive. Since $h(1) = 1$ and $\lim_{t\to\infty} h(t) = 0$, there exists M_h such that $h(t) \le M_h$ for all $t \ge 1$. Now, using (4.42), we have, upon defining $t_k := v_z^k/v_y^k$ that

$$
\begin{aligned}
\frac{\|\mathbf{w}^k - \mathbf{u}^k\|}{-(\ln\|\mathbf{w}^k - \mathbf{v}^k\|)^{-1}} &= \frac{v_y^k \sqrt{1 + [\ln(v_z^k/v_y^k)]^2}}{-(\ln v_z^k)^{-1}} = -v_y^k \sqrt{1 + [\ln(v_z^k/v_y^k)]^2} \ln v_z^k \\
&\overset{(a)}{=} -\frac{v_y^k}{v_z^k}\sqrt{1 + \left[\ln\left(\frac{v_z^k}{v_y^k}\right)\right]^2} v_z^k \ln v_z^k \overset{(b)}{=} -h(t_k) v_z^k \ln v_z^k \\
&\overset{(c)}{\le} -M_h v_z^k \ln v_z^k,
\end{aligned}
$$

where the division by v_z^k in (a) is legitimate because $v_z^k > 0$, (b) follows from the definition of h and the fact that $t_k > 1$ (see (4.43)), and (c) holds because of the definition of M_h and the fact that $-\ln v_z^k > 0$ (thanks to $v_z^k = \|\mathbf{w}^k - \mathbf{v}^k\| \le e^{-2}$). Since $v_z^k \to 0$, it then follows that $\left\{\frac{\|\mathbf{w}^k - \mathbf{u}^k\|}{-(\ln\|\mathbf{w}^k - \mathbf{v}^k\|)^{-1}}\right\}$ is a positive sequence that converges to zero. Thus, $\lim_{k\to\infty} \frac{-(\ln\|\mathbf{w}^k - \mathbf{v}^k\|)^{-1}}{\|\mathbf{w}^k - \mathbf{u}^k\|} = \infty$, which again shows that (3.21b) does not hold for $\mathfrak{g} = \mathfrak{g}_\infty$.

Having shown that (3.21b) does not hold for $\mathfrak{g} = \mathfrak{g}_\infty$, in view of Lemma 3.12, we must have $\gamma_{\mathbf{z},\eta} \in (0, \infty]$. Then the result follows from Theorem 3.10 and (4.40). $\qquad\square$

Combining Theorem 4.8, Theorem 4.10 and Lemma 3.9, and noting (4.40) and $\gamma_{\mathbf{z},0} = \infty$ (see (3.15)), we can now summarize the one-step facial residual functions derived in this section in the following corollary.

Corollary 4.11 (𝟙-*FRF concerning* $\mathcal{F}_\infty$) *Let* $\mathbf{z} \in K_{\exp}^*$ *with* $z_x = 0$ *and* $\{\mathbf{z}\}^\perp \cap K_{\exp} = \mathcal{F}_\infty$.

(i) *In the case when* $z_y > 0$, *let* $\kappa_{\mathbf{z},t}$ *be defined as in* (3.16) *with* $\mathfrak{g} = |\cdot|$. *Then the function* $\psi_{\mathcal{K},\mathbf{z}} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ *given by*

$$
\psi_{\mathcal{K},\mathbf{z}}(\epsilon, t) := \max\{\epsilon, \epsilon/\|\mathbf{z}\|\} + \kappa_{\mathbf{z},t}(\epsilon + \max\{\epsilon, \epsilon/\|\mathbf{z}\|\})
$$

is a $\mathbb{1}$-FRF for $K_{\exp}$ and z. In particular, there exist $\kappa > 0$ and a nonnegative monotone nondecreasing function $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ such that the function $\hat{\psi}_{\mathcal{K},z}$ given by $\hat{\psi}_{\mathcal{K},z}(\epsilon, t):=\kappa\epsilon + \rho(t)\epsilon$ is a $\mathbb{1}$-FRF for $K_{\exp}$ and z.

(ii) *In the case when $z_y = 0$, let $\kappa_{z,t}$ be defined as in* (3.16) *with* $\mathfrak{g} = \mathfrak{g}_\infty$ *in* (4.39). *Then the function* $\psi_{\mathcal{K},z} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ *given by*

$$\psi_{\mathcal{K},z}(\epsilon, t) := \max\{\epsilon, \epsilon/\|z\|\} + \kappa_{z,t}\mathfrak{g}_\infty(\epsilon + \max\{\epsilon, \epsilon/\|z\|\})$$

is a $\mathbb{1}$-FRF for $K_{\exp}$ and z. In particular, there exist $\kappa > 0$ and a nonnegative monotone nondecreasing function $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ such that the function $\hat{\psi}_{\mathcal{K},z}$ given by $\hat{\psi}_{\mathcal{K},z}(\epsilon, t):=\kappa\epsilon + \rho(t)\mathfrak{g}_\infty(\epsilon)$ is a $\mathbb{1}$-FRF for $K_{\exp}$ and z.

4.2.4 The non-exposed face $\mathcal{F}_{ne}$

Recall the unique non-exposed face of $K_{\exp}$:

$$\mathcal{F}_{ne} := \{(0, 0, z) \mid z \geq 0\}.$$

In this subsection, we take a look at $\mathcal{F}_{ne}$. Note that $\mathcal{F}_{ne}$ is an exposed face of $\mathcal{F}_{-\infty}$, which is polyhedral. This observation leads immediately to the following corollary, which also follows from [34, Proposition 18] by letting $\mathcal{F}:=\mathcal{K}:=\mathcal{F}_{-\infty}$ therein. We omit the proof for brevity.

Corollary 4.12 ($\mathbb{1}$-*FRF for $\mathcal{F}_{ne}$*) *Let $z \in \mathcal{F}^*_{-\infty}$ be such that $\mathcal{F}_{ne} = \mathcal{F}_{-\infty} \cap \{z\}^\perp$. Then there exists $\kappa > 0$ such that*

$$\psi_{\mathcal{F}_{-\infty},z}(\epsilon, t):=\kappa\epsilon$$

is a $\mathbb{1}$-FRF for $\mathcal{F}_{-\infty}$ and z.

4.3 Error bounds

In this subsection, we return to the feasibility problem (Feas) and consider the case where $\mathcal{K} = K_{\exp}$. We now have all the tools for obtaining error bounds. Recalling Definition 2.1, we can state the following result.

Theorem 4.13 (*Error bounds for* (Feas) *with* $\mathcal{K} = K_{\exp}$) *Let $\mathcal{L} \subseteq \mathbb{R}^3$ be a subspace and $a \in \mathbb{R}^3$ such that $(\mathcal{L} + a) \cap K_{\exp} \neq \emptyset$. Then the following items hold.*

(i) *The distance to the PPS condition of $\{K_{\exp}, \mathcal{L}+a\}$ satisfies $d_{PPS}(K_{\exp}, \mathcal{L}+a) \leq 1$.*

(ii) *If $d_{PPS}(K_{\exp}, \mathcal{L}+a) = 0$, then $K_{\exp}$ and $\mathcal{L}+a$ satisfy a Lipschitzian error bound.*

(iii) *Suppose $d_{PPS}(K_{\exp}, \mathcal{L} + a) = 1$ and let $\mathcal{F} \subsetneq K_{\exp}$ be a chain of faces of length 2 satisfying items* (ii) *and* (iii) *of Proposition* 3.2. *We have the following possibilities.*

 (a) *If $\mathcal{F} = \mathcal{F}_{-\infty}$ then $K_{\exp}$ and $\mathcal{L}+a$ satisfy an entropic error bound as in* (4.44). *In addition, for all $\alpha \in (0, 1)$, a uniform Hölderian error bound with exponent α holds.*

(b) If $\mathcal{F} = \mathcal{F}_\beta$, with $\beta \in \mathbb{R}$, then $K_{\exp}$ and $\mathcal{L} + a$ satisfy a uniform Hölderian error bound with exponent $1/2$.

(c) Suppose that $\mathcal{F} = \mathcal{F}_\infty$. If there exists $z \in K_{\exp}^* \cap \mathcal{L}^\perp \cap \{a\}^\perp$ with $z_x = 0$, $z_y > 0$ and $z_z > 0$ then $K_{\exp}$ and $\mathcal{L} + a$ satisfy a Lipschitzian error bound. Otherwise, $K_{\exp}$ and $\mathcal{L} + a$ satisfy a log-type error bound as in (4.45).

(d) If $\mathcal{F} = \{0\}$, then $K_{\exp}$ and $\mathcal{L} + a$ satisfy a Lipschitzian error bound.

Proof (i): All proper faces of $K_{\exp}$ are polyhedral, therefore $\ell_{\text{poly}}(K_{\exp}) = 1$. By item (i) of Proposition 3.2, there exists a chain of length 2 satisfying item (iii) of Proposition 3.2. Therefore, $d_{\text{PPS}}(K_{\exp}, \mathcal{L} + a) \leq 1$.

(ii): If $d_{\text{PPS}}(K_{\exp}, \mathcal{L} + a) = 0$, it is because $\{K_{\exp}, \mathcal{L} + a\}$ satisfies the PPS condition, which implies a Lipschitzian error bound by Proposition 2.2.

(iii): Next, suppose $d_{\text{PPS}}(K_{\exp}, \mathcal{L} + a) = 1$ and let $\mathcal{F} \subsetneq K_{\exp}$ be a chain of faces of length 2 satisfying items (ii) and (iii) of Proposition 3.2, together with $z \in K_{\exp}^* \cap \mathcal{L}^\perp \cap \{a\}^\perp$ such that

$$\mathcal{F} = K_{\exp} \cap \{z\}^\perp.$$

Since positively scaling z does not affect the chain of faces, we may assume that $\|z\| = 1$. Also, in what follows, for simplicity, we define

$$\widehat{d}(x) := \max\{d(x, \mathcal{L} + a), d(x, K_{\exp})\}.$$

Then, we prove each item by applying Theorem 3.8 with the corresponding facial residual function.

(a) If $\mathcal{F} = \mathcal{F}_{-\infty}$, the one-step facial residual functions are given by Corollary 4.4. First we consider the case where $\mathfrak{g} = \mathfrak{g}_{-\infty}$ and we have

$$\psi_{\mathcal{K},z}(\epsilon, t) := \epsilon + \kappa_{z,t}\mathfrak{g}_{-\infty}(2\epsilon),$$

where $\mathfrak{g}_{-\infty}$ is as in (4.14). Then, if ψ is a positively rescaled shift of $\psi_{\mathcal{K},z}$, using the monotonicity of $\mathfrak{g}_{-\infty}$ and of $\kappa_{z,t}$ as a function of t, we conclude that there exists $\widehat{M} > 0$ such that

$$\psi(\epsilon, t) \leq \widehat{M}\epsilon + \widehat{M}\kappa_{z,\widehat{M}t}\mathfrak{g}_{-\infty}(\widehat{M}\epsilon).$$

Invoking Theorem 3.8, using the monotonicity of all functions involved in the definition of ψ and recalling (4.15), we conclude that for every bounded set B, there exists $\kappa_B > 0$

$$d\left(x, (\mathcal{L} + a) \cap K_{\exp}\right) \leq \kappa_B \mathfrak{g}_{-\infty}(\widehat{d}(x)), \qquad \forall x \in B, \tag{4.44}$$

which shows that an entropic error bound holds.

Next, we consider the case $\mathfrak{g} = |\cdot|^\alpha$. Given $\alpha \in (0, 1)$, we have the following one-step facial residual function:

$$\psi_{\mathcal{K},\mathbf{z}}(\epsilon, t) := \epsilon + \kappa_{\mathbf{z},t} 2^\alpha \epsilon^\alpha,$$

where $\kappa_{\mathbf{z},t}$ is defined as in (3.16). Invoking Theorem 3.8, we conclude that for every bounded set B, there exists $\kappa_B > 0$ such that

$$\mathrm{d}\left(\mathbf{x}, (\mathcal{L} + \mathbf{a}) \cap K_{\exp}\right) \le \kappa_B \widehat{\mathrm{d}}(\mathbf{x}) + \kappa_B \widehat{\mathrm{d}}(\mathbf{x})^\alpha, \qquad \forall \mathbf{x} \in B,$$

In addition, for $\mathbf{x} \in B$, we have $\widehat{\mathrm{d}}(\mathbf{x}) \le \widehat{\mathrm{d}}(\mathbf{x})^\alpha M$, where $M = \sup_{\mathbf{x} \in B} \widehat{\mathrm{d}}(\mathbf{x})^{1-\alpha}$. In conclusion, for $\kappa = 2\kappa_B \max\{M, 1\}$, we have

$$\mathrm{d}\left(\mathbf{x}, (\mathcal{L} + \mathbf{a}) \cap K_{\exp}\right) \le \kappa \widehat{\mathrm{d}}(\mathbf{x})^\alpha, \qquad \forall \mathbf{x} \in B.$$

That is, a uniform Hölderian error bound holds with exponent α.

(b) If $\mathcal{F} = \mathcal{F}_\beta$, with $\beta \in \mathbb{R}$, then the one-step facial residual function is given by Corollary 4.7, that is, we have

$$\psi_{\mathcal{K},\mathbf{z}}(\epsilon, t) := \epsilon + \kappa_{\mathbf{z},t} \sqrt{2} \epsilon^{1/2},$$

Then, following the same argument as in the second half of item (a), we conclude that a uniform Hölderian error bound holds with exponent $1/2$.

(c) If $\mathcal{F} = \mathcal{F}_\infty$, the one-step facial residual functions are given by Corollary 4.11 and they depend on $\mathbf{z}$. Since $\mathcal{F} = \mathcal{F}_\infty$, we must have $z_x = 0$ and $z_z > 0$, see Sect. 4.1.1.

The deciding factor is whether z_y is positive or zero. If $z_y > 0$, then we have the following one-step facial residual function:

$$\psi_{\mathcal{K},\mathbf{z}}(\epsilon, t) := (1 + 2\kappa_{\mathbf{z},t})\epsilon,$$

where $\kappa_{\mathbf{z},t}$ is defined as in (3.16). In this case, analogously to items (a) and (b) we have a Lipschitzian error bound.

If $z_y = 0$, we have

$$\psi_{\mathcal{K},\mathbf{z}}(\epsilon, t) := \epsilon + \kappa_{\mathbf{z},t} \mathfrak{g}_\infty(2\epsilon),$$

where $\mathfrak{g}_\infty$ is as in (4.39). Analogous to the proof of item (a) but making use of (4.40) in place of (4.15), we conclude that for every bounded set B, there exists $\kappa_B > 0$ such that

$$\mathrm{d}\left(\mathbf{x}, (\mathcal{L} + \mathbf{a}) \cap K_{\exp}\right) \le \kappa_B \mathfrak{g}_\infty(\widehat{\mathrm{d}}(\mathbf{x}))), \qquad \forall \mathbf{x} \in B. \tag{4.45}$$

(d) See [34, Proposition 27].

$\square$

Remark 4.14 (Tightness of Theorem 4.13) We will argue that Theorem 4.13 is tight by showing that for every situation described in item (iii), there is a specific choice of $\mathcal{L}$ and a sequence $\{\mathbf{w}^k\}$ in $\mathcal{L} \setminus K_{\exp}$ with $d(\mathbf{w}^k, K_{\exp}) \to 0$ along which the corresponding error bound for $K_{\exp}$ and $\mathcal{L}$ is off by at most a multiplicative constant.

(a) Let $\mathcal{L} = \operatorname{span} \mathcal{F}_{-\infty} = \{(x, y, z) \mid y = 0\}$ (see (4.5)) and consider the sequence $\{\mathbf{w}^k\}$ where $\mathbf{w}^k = ((1/(k+1)) \ln(k+1), 0, 1)$, for every $k \in \mathbb{N}$. Then, $\mathcal{L} \cap K_{\exp} = \mathcal{F}_{-\infty}$ and we are under the conditions of item (iii)(a) of Theorem 4.13. Since $\{\mathbf{w}^k\} =: B \subseteq \mathcal{L}$, there exists $\kappa_B > 0$ such that

$$d\left(\mathbf{w}^k, \mathcal{L} \cap K_{\exp}\right) \leq \kappa_B \mathfrak{g}_{-\infty}(d(\mathbf{w}^k, K_{\exp})), \quad \forall k \in \mathbb{N}.$$

Then, the projection of $\mathbf{w}^k$ onto $\mathcal{F}_{-\infty}$ is given by $(0, 0, 1)$. Therefore,

$$\frac{\ln(k+1)}{k+1} = d(\mathbf{w}^k, \mathcal{L} \cap K_{\exp}) \leq \kappa_B \mathfrak{g}_{-\infty}(d(\mathbf{w}^k, K_{\exp})).$$

Let $\mathbf{v}^k = ((1/(k+1)) \ln(k+1), 1/(k+1), 1)$ for every k. Then, we have $\mathbf{v}^k \in K_{\exp}$. Therefore, $d(\mathbf{w}^k, K_{\exp}) \leq 1/(k+1)$. In view of the definition of $\mathfrak{g}_{-\infty}$ (see (4.14)), we conclude that for large enough k we have

$$\frac{\ln(k+1)}{k+1} = d(\mathbf{w}^k, \mathcal{L} \cap K_{\exp}) \leq \kappa_B \mathfrak{g}_{-\infty}(d(\mathbf{w}^k, K_{\exp})) \leq \kappa_B \frac{\ln(k+1)}{k+1}.$$

Thus, it holds that for all sufficiently large k,

$$1 \leq \frac{d(\mathbf{w}^k, \mathcal{L} \cap K_{\exp})}{\mathfrak{g}_{-\infty}(d(\mathbf{w}^k, K_{\exp}))} \leq \kappa_B.$$

Consequently, for any given nonnegative function $\mathfrak{g} : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{t \downarrow 0} \frac{\mathfrak{g}(t)}{\mathfrak{g}_{-\infty}(t)} = 0$, we have upon noting $d(\mathbf{w}^k, K_{\exp}) \to 0$ that

$$\frac{d(\mathbf{w}^k, \mathcal{L} \cap K_{\exp})}{\mathfrak{g}(d(\mathbf{w}^k, K_{\exp}))} = \frac{d(\mathbf{w}^k, \mathcal{L} \cap K_{\exp})}{\mathfrak{g}_{-\infty}(d(\mathbf{w}^k, K_{\exp}))} \frac{\mathfrak{g}_{-\infty}(d(\mathbf{w}^k, K_{\exp}))}{\mathfrak{g}(d(\mathbf{w}^k, K_{\exp}))} \to \infty,$$

which shows that the choice of $\mathfrak{g}_{-\infty}$ in the error bound is tight.

(b) Let $\beta \in \mathbb{R}$ and let $\widehat{\mathbf{z}}, \widehat{\mathbf{p}}$ and $\widehat{\mathbf{f}}$ be as in (4.24). Let $\mathcal{L} = \{\mathbf{z}\}^\perp$ with $z_x < 0$ such that $K_{\exp} \cap \mathcal{L} = \mathcal{F}_\beta$. We are then under the conditions of item (iii)(b) of Theorem 4.13. We consider the following sequences

$$\mathbf{v}^k = \begin{bmatrix} 1 - \beta + 1/k \\ 1 \\ e^{1-\beta+1/k} \end{bmatrix}, \quad \mathbf{w}^k = P_{\{\mathbf{z}\}^\perp} \mathbf{v}^k, \quad \mathbf{u}^k = P_{\mathcal{F}_\beta} \mathbf{w}^k.$$

For every k we have $\mathbf{v}^k \in \partial K_{\exp} \setminus \mathcal{F}_\beta$, and $\mathbf{v}^k \neq \mathbf{w}^k$ (because otherwise, we would have $\mathbf{v}^k \in K_{\exp} \cap \{\mathbf{z}\}^\perp = \mathcal{F}_\beta$). In addition, we have $\mathbf{v}^k \to \widehat{\mathbf{f}}$ and, since $\widehat{\mathbf{f}} \in \mathcal{F}_\beta$, we have $\mathbf{w}^k \to \widehat{\mathbf{f}}$ as well.

Next, notice that we have $\langle \widehat{\mathbf{f}}, \mathbf{v}^k \rangle \geq 0$ for k sufficiently large and $|v_x^k/v_y^k - (1 - \beta)| \to 0$. Then, following the computations outlined in case (I) of the proof of Theorem 4.6 and letting $\zeta_k := v_x^k/v_y^k$, we have from (4.30) and (4.31) that $h_2(\zeta_k) \neq 0$ for all large k (hence, $\mathbf{w}^k \neq \mathbf{u}^k$ for all large k), and that

$$L_\beta := \lim_{k \to \infty} \frac{\|\mathbf{w}^k - \mathbf{v}^k\|^{\frac{1}{2}}}{\|\mathbf{w}^k - \mathbf{u}^k\|} = \lim_{k \to \infty} \frac{\|\widehat{\mathbf{p}}\|}{\|\widehat{\mathbf{z}}\|^{\frac{1}{2}}} \frac{|h_1(\zeta_k)|^{\frac{1}{2}}}{|h_2(\zeta_k)|}$$
$$= \frac{\|\widehat{\mathbf{p}}\|}{\|\widehat{\mathbf{z}}\|^{\frac{1}{2}}} \frac{1}{\sqrt{2}(e^{\beta-1} + (\beta^2 + 1)e^{1-\beta})} \in (0, \infty),$$

(4.46)

where the latter equality is from (4.32). On the other hand, from item (iii)(b) of Theorem 4.13, for $B := \{\mathbf{w}^k\}$, there exists $\kappa_B > 0$ such that for all $k \in \mathbb{N}$,

$$\|\mathbf{w}^k - \mathbf{u}^k\| = \mathrm{d}(\mathbf{w}^k, \mathcal{L} \cap K_{\exp}) \leq \kappa_B \mathrm{d}(\mathbf{w}^k, K_{\exp})^{\frac{1}{2}} \leq \kappa_B \|\mathbf{w}^k - \mathbf{v}^k\|^{\frac{1}{2}}.$$

However, from (4.46), for large enough k, we have $\|\mathbf{w}^k - \mathbf{u}^k\| \geq 1/(2L_\beta)\|\mathbf{w}^k - \mathbf{v}^k\|^{\frac{1}{2}}$. Therefore, for large enough k we have

$$\frac{1}{2L_\beta} \|\mathbf{w}^k - \mathbf{v}^k\|^{\frac{1}{2}} \leq \mathrm{d}(\mathbf{w}^k, \mathcal{L} \cap K_{\exp}) \leq \kappa_B \mathrm{d}(\mathbf{w}^k, K_{\exp})^{\frac{1}{2}} \leq \kappa_B \|\mathbf{w}^k - \mathbf{v}^k\|^{\frac{1}{2}}.$$

Consequently, it holds that for all large enough k,

$$\frac{1}{2L_\beta} \leq \frac{\mathrm{d}(\mathbf{w}^k, \mathcal{L} \cap K_{\exp})}{\mathrm{d}(\mathbf{w}^k, K_{\exp})^{\frac{1}{2}}} \leq \kappa_B.$$

Arguing similarly as in case (a), we can also conclude that the choice of $|\cdot|^{\frac{1}{2}}$ in the error bound is tight.

(c) Let $\mathbf{z} = (0, 0, 1)$ and $\mathcal{L} = \{(x, y, 0) \mid x, y \in \mathbb{R}\} = \{\mathbf{z}\}^\perp$. Then, from (4.12), we have $\mathcal{L} \cap K_{\exp} = \mathcal{F}_\infty$. We are then under case (iii)(c) of Theorem 4.13. Because there is no $\widehat{\mathbf{z}} \in \mathcal{L}^\perp$ with $\widehat{z}_y > 0$, we have a log-type error bound as in (4.45). We proceed as in item (a) using sequences such that $\mathbf{w}^k = (-1, 1/k, 0)$, $\mathbf{v}^k = (-1, 1/k, (1/k)e^{-k})$, $\mathbf{u}^k = (-1, 0, 0)$, for every k. Note that $\mathbf{w}^k \in \mathcal{L}$, $\mathbf{v}^k \in K_{\exp}$ and $P_{\mathcal{F}_\infty}(\mathbf{w}^k) = \mathbf{u}^k$, for every k. Therefore, there exists $\kappa_B > 0$ such that

$$\frac{1}{k} = \mathrm{d}(\mathbf{w}^k, \mathcal{L} \cap K_{\exp}) \leq \kappa_B \mathfrak{g}_\infty(\mathrm{d}(\mathbf{w}^k, K_{\exp})) \leq \kappa_B \mathfrak{g}_\infty\left(\frac{1}{ke^k}\right), \quad \forall k \in \mathbb{N}.$$

In view of the definition of $\mathfrak{g}_\infty$ (see (4.39)), there exists $L > 0$ such that for large enough k we have

$$\frac{1}{k} = \mathrm{d}(\mathbf{w}^k, \mathcal{L} \cap K_{\exp}) \leq \kappa_B \mathfrak{g}_\infty(\mathrm{d}(\mathbf{w}^k, K_{\exp})) \leq \frac{L}{k}.$$

Consequently, it holds that for all large enough k,

$$\frac{\kappa_B}{L} \le \frac{d(\mathbf{w}^k, \mathcal{L} \cap K_{\exp})}{\mathfrak{g}_\infty(d(\mathbf{w}^k, K_{\exp}))} \le \kappa_B.$$

Arguing similarly as in case (a), we conclude that the choice of $\mathfrak{g}_\infty$ is tight.

Note that a Lipschitz error bound is always tight up to a constant, because $d(\mathbf{x}, \mathcal{K} \cap (\mathcal{L}+\mathbf{a})) \ge \max\{d(\mathbf{x}, \mathcal{K}), d(\mathbf{x}, \mathcal{L}+\mathbf{a})\}$. Therefore, the error bounds in items (ii), (iii)(d) and in the first half of (iii)(c) are tight.

Sometimes we may need to consider direct products of multiple copies of $K_{\exp}$ in order to model certain problems, i.e., our problem of interest could have the following shape:

$$\text{find} \quad \mathbf{x} \in (\mathcal{L} + \mathbf{a}) \cap \mathcal{K},$$

where $\mathcal{K} = K_{\exp} \times \cdots \times K_{\exp}$ is a direct product of m exponential cones.

Fortunately, we already have all the tools required to extend Theorem 4.13 and compute error bounds for this case too. We recall that the faces of a direct product of cones are direct products of the faces of the individual cones.[11] Therefore, using Proposition 3.13, we are able to compute all the necessary one-step facial residual functions for $\mathcal{K}$. Once they are obtained we can invoke Theorem 3.8. Unfortunately, there is quite a number of different cases one must consider, so we cannot give a concise statement of an all-encompassing tight error bound result.

We will, however, given an error bound result under the following *simplifying assumption of non-exceptionality* or SANE.

Assumption 4.15 (SANE: simplifying assumption of non-exceptionality) Suppose (Feas) is feasible with $\mathcal{K} = K_{\exp} \times \cdots \times K_{\exp}$ being a direct product of m exponential cones. We say that $\mathcal{K}$ and $\mathcal{L}+\mathbf{a}$ satisfy the *simplifying assumption of non-exceptionality* (SANE) if there exists a chain of faces $\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1 = \mathcal{K}$ as in Proposition 3.2 with $\ell - 1 = d_{\text{PPS}}(\mathcal{K}, \mathcal{L} + \mathbf{a})$ such that for all i, the exceptional face $\mathcal{F}_\infty$ of $K_{\exp}$ never appears as one of the blocks of $\mathcal{F}_i$.

Remark 4.16 (SANE is not unreasonable) In many modelling applications of the exponential cone presented in [37, Chapter 5], translating to our notation, the $\mathbf{y}$ variable is fixed to be 1 in (4.1). For example, the hypograph of the logarithm function "$x \le \ln(z)$" can be represented as the constraint "$(x, y, z) \in K_{\exp} \cap (\mathcal{L} + \mathbf{a})$", where $\mathcal{L}+\mathbf{a} = \{(x, y, z) \mid y = 1\}$. Because the y variable is fixed to be 1, the feasible region does not intersect the 2D face $\mathcal{F}_{-\infty}$ nor its subfaces $\mathcal{F}_\infty$ and $\mathcal{F}_{ne}$. In particular, SANE is satisfied. More generally, if $\mathcal{K}$ is a direct product of exponential cones and the affine

---

[11] Here is a sketch of the proof. If $\mathcal{F}^1 \trianglelefteq \mathcal{K}^1$, $\mathcal{F}^2 \trianglelefteq \mathcal{K}^2$, then the definition of face implies that $\mathcal{F}^1 \times \mathcal{F}^2 \trianglelefteq \mathcal{K}^1 \times \mathcal{K}^2$. For the converse, let $\mathcal{F} \trianglelefteq \mathcal{K}^1 \times \mathcal{K}^2$ and let $\mathcal{F}^1$, $\mathcal{F}^2$ be the projection of $\mathcal{F}$ onto the first and second variables, respectively. Suppose that $\mathbf{x}, \mathbf{y} \in \mathcal{K}^1$ are such that $\mathbf{x} + \mathbf{y} \in \mathcal{F}^1$. Then, $(\mathbf{x} + \mathbf{y}, \mathbf{z}) \in \mathcal{F}$ for some $\mathbf{z} \in \mathcal{K}^2$. Since $(\mathbf{x}+\mathbf{y}, \mathbf{z}) = (\mathbf{x}, \mathbf{z}/2)+(\mathbf{y}, \mathbf{z}/2)$ and $\mathcal{F}$ is a face, we conclude that $(\mathbf{x}, \mathbf{z}/2), (\mathbf{y}, \mathbf{z}/2) \in \mathcal{F}$ and $\mathbf{x}, \mathbf{y} \in \mathcal{F}^1$. Therefore $\mathcal{F}^1 \trianglelefteq \mathcal{K}^1$ and, similarly, $\mathcal{F}^2 \trianglelefteq \mathcal{K}^2$. Then, the equality $\mathcal{F} = \mathcal{F}^1 \times \mathcal{F}^2$ is proven using the definition of face and the fact that $(\mathbf{x}, \mathbf{z}) = (\mathbf{x}, 0) + (0, \mathbf{z})$.

space $\mathcal{L} + \mathbf{a}$ is such that the **y** components of each block are fixed positive constants, then $\mathcal{K}$ and $\mathcal{L} + \mathbf{a}$ satisfy SANE.

On the other hand, problems involving the relative entropy $D(x, y) := x \ln(x/y)$ are often modelled as "minimize t" subject to "$(-t, x, y) \in K_{\exp}$" and additional constraints. We could also have sums so that the problem is of the form "minimize $\sum t_i$" subject to "$(-t_i, x_i, y_i) \in K_{\exp}$" and additional constraints. In those cases, it seems that it could happen that SANE is not satisfied.

Under SANE, we can state the following result.

Theorem 4.17 (Error bounds for direct products of exponential cones) *Suppose* (Feas) *is feasible with* $\mathcal{K} = K_{\exp} \times \cdots \times K_{\exp}$ *being a direct product of m exponential cones. Then the following hold.*

   (i) The distance to the PPS condition of $\{\mathcal{K}, \mathcal{L} + \mathbf{a}\}$ *satisfies* $d_{PPS}(\mathcal{K}, \mathcal{L} + \mathbf{a}) \leq m$.
  (ii) If SANE is satisfied, then $\mathcal{K}$ *and* $\mathcal{L} + \mathbf{a}$ *satisfy a uniform Hölderian error bound with exponent* $2^{-d_{PPS}(K_{\exp}, \mathcal{L} + \mathbf{a})}$.

Proof (i): All proper faces of $K_{\exp}$ are polyhedral, therefore $\ell_{\text{poly}}(K_{\exp}) = 1$. By item (i) of Proposition 3.2, there exists a chain of length ℓ satisfying item (iii) of Proposition 3.2 such that $\ell - 1 \leq m$. Therefore, $d_{PPS}(\mathcal{K}, \mathcal{L} + \mathbf{a}) \leq \ell - 1 \leq m$.

(ii): If SANE is satisfied, then there exists a chain $\mathcal{F}_\ell \subsetneq \cdots \subsetneq \mathcal{F}_1 = \mathcal{K}$ of length $\ell \leq m + 1$ as in Proposition 3.2, together with the corresponding $\mathbf{z}_1, \ldots, \mathbf{z}_{\ell-1}$. Also, the exceptional face $\mathcal{F}_\infty$ never appears as one of the blocks of the $\mathcal{F}_i$.

In what follows, for simplicity, we define

$$\widehat{d}(\mathbf{x}) := \max\{d(\mathbf{x}, \mathcal{L} + \mathbf{a}), d(\mathbf{x}, \mathcal{K})\}.$$

Then, we invoke Theorem 3.8, which implies that given a bounded set B, there exists a constant $\kappa_B > 0$ such that

$$d(\mathbf{x}, (\mathcal{L} + \mathbf{a}) \cap \mathcal{K}) \leq \kappa_B(\widehat{d}(\mathbf{x}) + \varphi(\widehat{d}(\mathbf{x}), M)), \tag{4.47}$$

where $M = \sup_{\mathbf{x} \in B} \|\mathbf{x}\|$ and there are two cases for φ. If $\ell = 1$, φ is the function such that $\varphi(\epsilon, M) = \epsilon$. If $\ell \geq 2$, we have $\varphi = \psi_{\ell-1} \diamond \cdots \diamond \psi_1$, where ψ_i is a (suitable positively rescaled shift of a) one-step facial residual function for $\mathcal{F}_i$ and $\mathbf{z}_i$. In the former case, the PPS condition is satisfied, we have a Lipschitzian error bound and we are done. We therefore assume that the latter case occurs with $\ell - 1 = d_{PPS}(\mathcal{K}, \mathcal{L} + \mathbf{a})$.

First, we compute the one-step facial residual functions for each $\mathcal{F}_i$. In order to do that, we recall that each $\mathcal{F}_i$ is a direct product $\mathcal{F}_i^1 \times \cdots \times \mathcal{F}_i^m$ where each $\mathcal{F}_i^j$ is a face of $K_{\exp}$, excluding $\mathcal{F}_\infty$ by SANE. Therefore, a one-step facial residual function for $\mathcal{F}_i^j$ can be obtained from Corollary 4.4, 4.7 or 4.12. In particular, taking the worst[12] case in consideration, and taking the maximum of the facial residual functions, there exists a nonnegative monotone nondecreasing function $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ such that the

---

[12] $\sqrt{\cdot}$ is "worse" than $\mathfrak{g}_{-\infty}$ in that, near zero, $\sqrt{t} \geq \mathfrak{g}_{-\infty}(t)$. The function $\mathfrak{g}_\infty$ need not be considered because, by SANE, $\mathcal{F}_\infty$ never appears.

function ψ given by

$$\psi(\epsilon, t) := \rho(t)\epsilon + \rho(t)\sqrt{\epsilon}$$

is a one-step facial residual function for each $\mathcal{F}_i^j$. In what follows, in order to simplify the notation, we define $\hat{\mathfrak{g}}(t) := \sqrt{t}$. Also, for every j, we use $\hat{\mathfrak{g}}_j$ to denote the composition of $\hat{\mathfrak{g}}$ with itself j-times, i.e.,

$$\hat{\mathfrak{g}}_j = \underbrace{\hat{\mathfrak{g}} \circ \cdots \circ \hat{\mathfrak{g}}}_{j \text{ times}}; \tag{4.48}$$

and we set $\hat{\mathfrak{g}}_0$ to be the identity map.

Using the above notation and Proposition 3.13, we conclude the existence of a nonnegative monotone nondecreasing function $\sigma : \mathbb{R}_+ \to \mathbb{R}_+$ such that the function ψ_i given by

$$\psi_i(\epsilon, t) := \sigma(t)\epsilon + \sigma(t)\hat{\mathfrak{g}}(\epsilon)$$

is a one-step facial residual function for $\mathcal{F}_i$ and $\mathbf{z}_i$. Therefore, for $\mathbf{x} \in B$, we have

$$\psi_i(\epsilon, \|\mathbf{x}\|) \le \sigma(M)\epsilon + \sigma(M)\hat{\mathfrak{g}}(\epsilon) = \psi_i(\epsilon, M), \tag{4.49}$$

where $M = \sup_{\mathbf{x} \in B} \|\mathbf{x}\|$.

Next we are going to make a series of arguments related to the following informal principle: over a bounded set only the terms $\hat{\mathfrak{g}}_j$ with largest j matter. We start by noting that for any $\mathbf{x} \in B$ and any $0 \le k \le j \le \ell$,

$$\hat{\mathfrak{g}}_k(\widehat{\mathrm{d}}(\mathbf{x})) = \widehat{\mathrm{d}}(\mathbf{x})^{2^{-k}} = \widehat{\mathrm{d}}(\mathbf{x})^{(2^{-k}-2^{-j})}\widehat{\mathrm{d}}(\mathbf{x})^{2^{-j}} \le \hat{\kappa}_{j,k}\widehat{\mathrm{d}}(\mathbf{x})^{2^{-j}} \le \hat{\kappa}\hat{\mathfrak{g}}_j(\widehat{\mathrm{d}}(\mathbf{x})), \tag{4.50}$$

where $\hat{\kappa}_{j,k} := \sup_{\mathbf{x} \in B} \widehat{\mathrm{d}}(\mathbf{x})^{(2^{-k}-2^{-j})} < \infty$ because $\mathbf{x} \mapsto \widehat{\mathrm{d}}(\mathbf{x})^{(2^{-k}-2^{-j})}$ is continuous, and $\hat{\kappa} := \max_{0 \le k \le j \le \ell} \hat{\kappa}_{j,k}$.

Now, let $\varphi_j := \psi_j \diamond \cdots \diamond \psi_1$, where $\diamond$ is the diamond composition defined in (3.3). We will show by induction that for every $j \le \ell - 1$ there exists κ_j such that

$$\varphi_j(\widehat{\mathrm{d}}(\mathbf{x}), M) \le \kappa_j\hat{\mathfrak{g}}_j(\widehat{\mathrm{d}}(\mathbf{x})), \qquad \forall \mathbf{x} \in B. \tag{4.51}$$

For $j = 1$, it follows directly from (4.49) and (4.50). Now, suppose that the claim is valid for some j such that $j + 1 \le \ell - 1$. By the inductive hypothesis, we have

$$\begin{aligned}
\varphi_{j+1}(\widehat{\mathrm{d}}(\mathbf{x}), M) &= \psi_{j+1}(\widehat{\mathrm{d}}(\mathbf{x}) + \varphi_j(\widehat{\mathrm{d}}(\mathbf{x}), M), M) \\
&\le \psi_{j+1}(\widehat{\mathrm{d}}(\mathbf{x}) + \kappa_j\hat{\mathfrak{g}}_j(\widehat{\mathrm{d}}(\mathbf{x})), M) \\
&\le \psi_{j+1}(\tilde{\kappa}_j\hat{\mathfrak{g}}_j(\widehat{\mathrm{d}}(\mathbf{x})), M),
\end{aligned} \tag{4.52}$$

where $\tilde{\kappa}_j := 2 \max\{\hat{\kappa}, \kappa_j\}$ and the last inequality follows from (4.50). Then, we plug $\epsilon = \tilde{\kappa}_j \hat{\mathfrak{g}}_j(\widehat{\mathrm{d}}(\mathbf{x}))$ in (4.49) to obtain

$$
\begin{aligned}
\psi_{j+1}(\tilde{\kappa}_j \hat{\mathfrak{g}}_j(\widehat{\mathrm{d}}(\mathbf{x})), M) &= \sigma(M)\tilde{\kappa}_j \hat{\mathfrak{g}}_j(\widehat{\mathrm{d}}(\mathbf{x})) + \sigma(M)\hat{\mathfrak{g}}(\tilde{\kappa}_j \hat{\mathfrak{g}}_j(\widehat{\mathrm{d}}(\mathbf{x}))) \\
&= \sigma(M)\tilde{\kappa}_j \hat{\mathfrak{g}}_j(\widehat{\mathrm{d}}(\mathbf{x})) + \sigma(M)\sqrt{\tilde{\kappa}_j}\hat{\mathfrak{g}}_{j+1}(\widehat{\mathrm{d}}(\mathbf{x})) \\
&\leq \sigma(M)(\tilde{\kappa}_j \hat{\kappa} + \sqrt{\tilde{\kappa}_j})\hat{\mathfrak{g}}_{j+1}(\widehat{\mathrm{d}}(\mathbf{x})),
\end{aligned}
\tag{4.53}
$$

where the last inequality follows from (4.50). Combining (4.52) and (4.53) concludes the induction proof. In particular, (4.51) is valid for $j = \ell - 1$. Then, taking into account some positive rescaling and shifting (see (3.2)) and adjusting constants, from (4.47), (4.51) and (4.50) we deduce that there exists $\kappa > 0$ such that

$$
\mathrm{d}\,(\mathbf{x}, (\mathcal{L} + \mathbf{a}) \cap \mathcal{K}) \leq \kappa \hat{\mathfrak{g}}_{\ell-1}(\widehat{\mathrm{d}}(\mathbf{x})), \qquad \forall \mathbf{x} \in B
$$

with $\hat{\mathfrak{g}}_{\ell-1}$ as in (4.48). To complete the proof, we recall that $d_{\mathrm{PPS}}(\mathcal{K}, \mathcal{L} + \mathbf{a}) = \ell - 1$. $\qquad\square$

Remark 4.18 (Variants of Theorem 4.17) Theorem 4.17 is not tight and admits variants that are somewhat cumbersome to describe precisely. For example, the $\mathfrak{g}_{-\infty}$ function was not taken into account explicitly but simply "relaxed" to $t \mapsto \sqrt{t}$.

Going for greater generality, we can also drop the SANE assumption altogether and try to be as tight as our analysis permits when dealing with possibly inSANE instances. Although there are several possibilities one must consider, the overall strategy is the same as outlined in the proof of Theorem 4.17: invoke Theorem 3.8, fix a bounded set B, pick a chain of faces as in Proposition 3.2 and upper bound the diamond composition of facial residual function as in (4.51). Intuitively, whenever sums of function compositions appear, only the "higher" compositions matter. However, the analysis must consider the possibility of $\mathfrak{g}_{-\infty}$ or $\mathfrak{g}_{\infty}$ appearing. After this is done, it is just a matter to plug this upper bound into (4.47).

We conclude this subsection with an application. In [11], among other results, the authors showed that when a Hölderian error bound holds, it is possible to derive the convergence rate of several algorithms from the exponent of the error bound. As a consequence, Theorems 4.13 and 4.17 allow us to apply some of their results (e.g., [11, Corollary 3.8]) to the conic feasibility problem with exponential cones, *whenever a Hölderian error bound holds*. For non-Hölderian error bounds appearing in Theorem 4.13, different techniques are necessary, such as the ones discussed in [33] for deriving convergence rates under more general error bounds.

4.4 Miscellaneous odd behavior and connections to other notions

In this final subsection, we collect several instances of pathological behaviour that can be found inside the facial structure of the exponential cone.

Example 4.19 (Hölderian bounds and the non-attainment of admissible exponents) We recall Definition 2.1 and we consider the special case of two closed convex sets C_1, C_2 with non-empty intersection. We say that $\gamma \in (0, 1]$ is an *admissible exponent* for C_1, C_2 if C_1 and C_2 satisfy a uniform Hölderian error bound with exponent γ. It turns out that the supremum of the set of admissible exponents is not itself admissible. In particular, if $C_1 = K_{\exp}$ and $C_2 = \operatorname{span} \mathcal{F}_{-\infty}$, then we see from Corollary 4.3 that $C_1 \cap C_2 = \mathcal{F}_{-\infty}$ and that C_1 and C_2 satisfy a uniform Hölderian error bound for all $\gamma \in (0, 1)$; however, in view of the sequence constructed in Remark 4.14(a), the exponent cannot be chosen to be $\gamma = 1$.

In fact, from Theorem 4.13 and Remark 4.14(a), C_1 and C_2 satisfy an entropic error bound which is tight and is, in a sense, better than any Hölderian error bound with $\gamma \in (0, 1)$ but worse than a Lipschitzian error bound.

Example 4.20 (Non-Hölderian error bound) The facial structure of $K_{\exp}$ can be used to derive an example of two sets that provably do not have a Hölderian error bound. Let $C_1 = K_{\exp}$ and $C_2 = \{\mathbf{z}\}^\perp$, where $z_x = z_y = 0$ and $z_z = 1$ so that $C_1 \cap C_2 = \mathcal{F}_\infty$. Then, for every $\eta > 0$ and every $\alpha \in (0, 1]$, there is no constant $\kappa > 0$ such that

$$\mathrm{d}(\mathbf{x}, \mathcal{F}_\infty) \leq \kappa \max\{\mathrm{d}(\mathbf{x}, K_{\exp})^\alpha, \mathrm{d}(\mathbf{x}, \{\mathbf{z}\}^\perp)^\alpha\}, \qquad \forall\, \mathbf{x} \in B(\eta).$$

This is because if there were such a positive κ, the infimum in Lemma 4.9 would be positive, which it is not. This shows that C_1 and C_2 do not have a Hölderian error bound. However, as seen in Theorem 4.10, C_1 and C_2 have a log-type error bound. In particular if $\mathbf{q} \in B(\eta)$, using (2.1), (2.2) and Theorem 4.10, we have

$$\begin{aligned}
\mathrm{d}(\mathbf{q}, \mathcal{F}_\infty) &\leq \mathrm{d}(\mathbf{q}, \{\mathbf{z}\}^\perp) + \mathrm{d}(P_{\{\mathbf{z}\}^\perp}(\mathbf{q}), \mathcal{F}_\infty) \\
&\leq \mathrm{d}(\mathbf{q}, \{\mathbf{z}\}^\perp) + \max\{2, 2\gamma_{\mathbf{z},\eta}^{-1}\}\mathfrak{g}_\infty(\mathrm{d}(P_{\{\mathbf{z}\}^\perp}(\mathbf{q}), K_{\exp})) \\
&\leq \widehat{\mathrm{d}}(\mathbf{q}) + \max\{2, 2\gamma_{\mathbf{z},\eta}^{-1}\}\mathfrak{g}_\infty(2\widehat{\mathrm{d}}(\mathbf{q})),
\end{aligned} \tag{4.54}$$

where $\widehat{\mathrm{d}}(\mathbf{q}) := \max\{\mathrm{d}(\mathbf{q}, K_{\exp}), \mathrm{d}(\mathbf{q}, \{\mathbf{z}\}^\perp)\}$ and in the last inequality we used the monotonicity of $\mathfrak{g}_\infty$.

Let $C_1, \cdots, C_m$ be closed convex sets having nonempty intersection and let $C := \cap_{i=1}^m C_i$. Following [33], we say that $\varphi : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is a *consistent error bound function (CEBF)* for $C_1, \ldots, C_m$ if the following inequality holds

$$\mathrm{d}(\mathbf{x},\, C) \leq \varphi\left(\max_{1 \leq i \leq m} \mathrm{d}(\mathbf{x}, C_i),\, \|\mathbf{x}\|\right) \quad \forall\, \mathbf{x} \in \mathcal{E};$$

and the following technical conditions are satisfied for every $a, b \in \mathbb{R}_+$: $\varphi(\cdot, b)$ is monotone nondecreasing, right-continuous at 0 and $\varphi(0, b) = 0$; $\varphi(a, \cdot)$ is mononotone nondecreasing. CEBFs are a framework for expressing error bounds and can be used in the convergence analysis of algorithms for convex feasibility problems, see [33, Sects. 3 and 4]. For example, $C_1, \ldots, C_m$ satisfy a Hölderian error bound (Definition 2.1) if and only if these sets admit a CEBF of the format

$\varphi(a, b) := \rho(b) \max\{a, a^{\gamma(b)}\}$, where $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ and $\gamma : \mathbb{R}_+ \to (0, 1]$ are monotone nondecreasing functions [33, Theorem 3.4].

We remark that in Example 4.20, although the sets C_1, C_2 do not satisfy a Hölderian error bound, the log-type error bound displayed therein is covered under the framework of consistent error bound functions. This is because $\mathfrak{g}_\infty$ is a continuous monotone nondecreasing function and $\gamma_{\mathbf{z},\eta}^{-1}$ is monotone nondecreasing as a function of η (Remark 3.11). Therefore, in view of (4.54), the function given by $\varphi(a, b) := a + \max\{2, 2\gamma_{\mathbf{z},b}^{-1}\}\mathfrak{g}_\infty(2a)$ is a CEBF for C_1 and C_2.

By the way, it seems conceivable that many of our results in Sect. 3.1 can be adapted to derive CEBFs for arbitrary convex sets. Specifically, Lemma 3.9, Theorem 3.10, and Lemma 3.12 only rely on convexity rather than on the more specific structure of cones.

Next, we will see that we can also adapt Examples 4.19 and 4.20 to find instances of odd behavior of the so-called *Kurdyka-Łojasiewicz (KL) property* [1, 2, 8–10, 30]. First, we recall some notations and definitions. Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper closed convex extended-real-valued function. We denote by $\operatorname{dom}\partial f$ the set of points for which the subdifferential $\partial f(\mathbf{x})$ is non-empty and by $[a < f < b]$ the set of $\mathbf{x}$ such that $a < f(\mathbf{x}) < b$. As in [10, Sect. 2.3], we define for $r_0 \in (0, \infty)$ the set

$$\mathcal{K}(0, r_0) := \{\phi \in C[0, r_0) \cap C^1(0, r_0) \mid \phi \text{ is concave, } \phi(0) = 0, \ \phi'(r) > 0 \ \forall r \in (0, r_0)\}.$$

Let $B(\mathbf{x}, \epsilon)$ denote the closed ball of radius $\epsilon > 0$ centered at $\mathbf{x}$. With that, we say that f satisfies the KL property at $\mathbf{x} \in \operatorname{dom}\partial f$ if there exist $r_0 \in (0, \infty)$, $\epsilon > 0$ and $\phi \in \mathcal{K}(0, r_0)$ such that for all $\mathbf{y} \in B(\mathbf{x}, \epsilon) \cap [f(\mathbf{x}) < f < f(\mathbf{x}) + r_0]$ we have

$$\phi'(f(\mathbf{y}) - f(\mathbf{x}))\mathrm{d}(0, \partial f(\mathbf{y})) \geq 1.$$

In particular, as in [30], we say that f satisfies the *KL property with exponent* $\alpha \in [0, 1)$ at $\mathbf{x} \in \operatorname{dom}\partial f$, if ϕ can be taken to be $\phi(t) = ct^{1-\alpha}$ for some positive constant c. Next, we need a result which is a corollary of [10, Theorem 5].

Proposition 4.21 *Let $C_1, C_2 \subseteq \mathbb{R}^n$ be closed convex sets with $C_1 \cap C_2 \neq \emptyset$. Define $f : \mathbb{R}^n \to \mathbb{R}$ as*

$$f(\mathbf{y}) = d(\mathbf{y}, C_1)^2 + d(\mathbf{y}, C_2)^2.$$

Let $\mathbf{x} \in C_1 \cap C_2$, $\gamma \in (0, 1]$. Then, there exist $\kappa > 0$ and $\epsilon > 0$ such that

$$d(\mathbf{y}, C_1 \cap C_2) \leq \kappa \max\{d(\mathbf{y}, C_1), d(\mathbf{y}, C_2)\}^\gamma, \qquad \forall \mathbf{y} \in B(\mathbf{x}, \epsilon) \qquad (4.55)$$

if and only if f satisfies the KL property with exponent $1 - \gamma/2$ at $\mathbf{x}$.

Proof Note that $\inf f = 0$ and $\operatorname{argmin} f = C_1 \cap C_2$. Furthermore, (4.55) is equivalent to the existence of $\kappa' > 0$ and $\epsilon > 0$ such that

$$\mathrm{d}(\mathbf{y}, \operatorname{argmin} f) \leq \varphi(f(\mathbf{y})), \qquad \forall \mathbf{y} \in B(\mathbf{x}, \epsilon),$$

where φ is the function given by $\varphi(r) = \kappa' r^{\gamma/2}$. With that, the result follows from [10, Theorem 5]. □

Example 4.22 (Examples in the KL world) In Example 4.19, we have two sets C_1, C_2 satisfying a uniform Hölderian error bound for $\gamma \in (0, 1)$ but not for $\gamma = 1$. Because C_1 and C_2 are cones and the corresponding distance functions are positively homogeneous, this implies that for $\mathbf{0} \in C_1 \cap C_2$, a Lipschitzian error bound never holds at any neighbourhood of $\mathbf{0}$. That is, given $\eta > 0$, there is no $\kappa > 0$ such that

$$\mathrm{d}(\mathbf{y}, C_1 \cap C_2) \leq \kappa \max\{\mathrm{d}(\mathbf{y}, C_1), \mathrm{d}(\mathbf{y}, C_2)\}, \qquad \forall \mathbf{y} \in B(\eta)$$

holds. Consequently, the function f in Proposition 4.21 satisfies the KL property with exponent α for any $\alpha \in (1/2, 1)$ at the origin, but not for $\alpha = 1/2$. To the best of our knowledge, this is the first explicitly constructed function in the literature such that the infimum of KL exponents at a point is not itself a KL exponent.

Similarly, from Example 4.20 we obtain C_1, C_2 for which (4.55) does not hold for $\mathbf{0} \in C_1 \cap C_2$ with any chosen $\kappa, \varepsilon > 0, \ \gamma \in (0, 1]$. Thus from Proposition 4.21 we obtain a function f that does not satisfy the KL property with exponent $\beta \in [1/2, 1)$ at the origin. Since a function satisfying the KL property with exponent $\alpha \in [0, 1)$ at an $\mathbf{x} \in \mathrm{dom}\partial f$ necessarily satisfies it with exponent β for any $\beta \in [\alpha, 1)$ at $\mathbf{x}$, we see that this f does not satisfy the KL property with any exponent at the origin. On passing, we would like to point out that there are functions known in the literature that fail to satisfy the KL property; e.g., [9, Example 1].

5 Concluding remarks

In this work, we presented an extension of the results of [34] and showed how to obtain error bounds for conic linear systems using one-step facial residual functions and facial reduction (Theorem 3.8) even when the underlying cone is not amenable. Related to facial residual functions, we also developed techniques that aid in their computation; see Sect. 3.1. Finally, all techniques and results developed in Sect. 3 were used in some shape or form in order to obtain error bounds for the exponential cone in Sect. 4. Our new framework unlocks analysis for cones not reachable with the techniques developed in [34]; these include cones that are not facially exposed, as well as cones for which the projection operator has no simple closed form or is only implicitly specified. These were, until now, significant barriers against error bound analysis for many cones of interest.

As future work, we are planning to use the techniques developed in this paper to analyze and obtain error bounds for some of these other cones that have been previously unapproachable. Potential examples include the cone of $n \times n$ completely positive matrices and its dual, the cone of $n \times n$ copositive matrices. The former is not facially exposed when $n \geq 5$ (see [53]) and the latter is not facially exposed when $n \geq 2$. It would be interesting to clarify how far error bound problems for these cones can be tackled by our framework. Or, more ambitiously, we could try to obtain some of the facial residual functions and some error bound results. Of course, a significant

challenge is that their facial structure is not completely understood, but we believe that even partial results for general n or complete results for specific values of n would be relevant and, possibly, quite non-trivial. Finally, as suggested by one of the reviewers, our framework may be enriched by investigating further geometric interpretations of the key quantity $\gamma_{\mathbf{z},\eta}$ in (3.15), beyond Fig. 2. For instance, it will be interesting to see whether the positivity of $\gamma_{\mathbf{z},\eta}$ is related to some generalization of the angle condition in [38], which was originally proposed for the study of Lipschitz error bounds.

References

1. Attouch, H., Bolte, J., Redont, P., Soubeyran, A.: Proximal alternating minimization and projection methods for nonconvex problems: an approach based on the Kurdyka-Łojasiewicz inequality. Math. Oper. Res. **35**, 438–457 (2010)
2. Attouch, H., Bolte, J., Svaiter, B.F.: Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss–Seidel methods. Math. Program. **137**, 92–129 (2013)
3. Barker, G.P.: The lattice of faces of a finite dimensional cone. Linear Algebra Appl. **7**(1), 71–82 (1973)
4. Barker, G.P.: Theory of cones. Linear Algebra Appl. **39**, 263–291 (1981)
5. Bauschke, H.H., Borwein, J.M.: On projection algorithms for solving convex feasibility problems. SIAM Rev. **38**(3), 367–426 (1996)
6. Bauschke, H.H., Borwein, J.M., Li, W.: Strong conical hull intersection property, bounded linear regularity, Jameson's property (G), and error bounds in convex optimization. Math. Program. **86**(1), 135–160 (1999)
7. Bauschke, H.H., Lindstrom, S.B.: Proximal averages for minimization of entropy functionals. (2020) arXiv preprint arXiv:1807.08878
8. Bolte, J., Daniilidis, A., Lewis, A.: The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. SIAM J. Optim. **17**, 1205–1223 (2007)
9. Bolte, J., Daniilidis, A., Lewis, A., Shiota, M.: Clarke subgradients of stratifiable functions. SIAM J. Optim. **18**, 556–572 (2007)
10. Bolte, J., Nguyen, T.P., Peypouquet, J., Suter, B.W.: From error bounds to the complexity of first-order descent methods for convex functions. Math. Program. **165**(2), 471–507 (2017)
11. Borwein, J.M., Li, G., Tam, M.K.: Convergence rate analysis for averaged fixed point iterations in common fixed point problems. SIAM J. Optim. **27**(1), 1–33 (2017)
12. Borwein, J.M., Lindstrom, S.B.: Meetings with Lambert W and other special functions in optimization and analysis. Pure and App. Func. Anal. **1**(3), 361–396 (2016)
13. Borwein, J.M., Wolkowicz, H.: Regularizing the abstract convex program. J. Math. Anal. Appl. **83**(2), 495–530 (1981)
14. Burachik, R.S., Dao, M.N., Lindstrom, S.B.: The generalized Bregman distance. (2020) arXiv preprint arXiv:1909.08206
15. Chandrasekaran, V., Shah, P.: Relative entropy optimization and its applications. Math. Program. **161**(1), 1–32 (2017)

16. Coey, C., Kapelevich, L., Vielma, J.P.: Solving natural conic formulations with hypatia.jl. ArXiv e-prints (2021)
17. Dahl, J., Andersen, E.D.: A primal-dual interior-point algorithm for nonsymmetric exponential-cone optimization. Mathematical Programming (2021)
18. Faraut, J., Korányi, A.: Analysis on Symmetric Cones. Oxford Mathematical Monographs. Clarendon Press, Oxford (1994)
19. Faybusovich, L.: Several Jordan-algebraic aspects of optimization. Optimization **57**(3), 379–393 (2008)
20. Friberg, H.A.: Projection onto the exponential cone: a univariate root-finding problem. Optimization Online (2021)
21. Gouveia, J., Parrilo, P.A., Thomas, R.R.: Lifts of convex sets and cone factorizations. Math. Oper. Res. **38**(2), 248–264 (2013)
22. Henrion, D., Malick, J.: Projection methods for conic feasibility problems: applications to polynomial sum-of-squares decompositions. Optim. Methods Softw. **26**(1), 23–46 (2011)
23. Hoffman, A.J.: On approximate solutions of systems of linear inequalities. J. Res. Natl. Bur. Stand. **49**(4), 263–265 (1952)
24. Ioffe, A.D.: Variational Analysis of Regular Mappings: Theory and Applications. Springer Monographs in Mathematics. Springer, Berlin (2017)
25. Karimi, M., Tunçel, L.: Domain-Driven Solver (DDS) Version 2.0: a MATLAB-based software package for convex optimization problems in domain-driven form. ArXiv e-prints (2019)
26. Lewis, A.S., Pang, J.S.: Error bounds for convex inequality systems. In: Crouzeix, J.P., Martínez-Legaz, J.E., Volle, M. (eds.) Generalized Convexity, Generalized Monotonicity: Recent Results, pp. 75–110. Springer, US (1998)
27. Li, G.: On the asymptotically well behaved functions and global error bound for convex polynomials. SIAM J. Optim. **20**(4), 1923–1943 (2010)
28. Li, G.: Global error bounds for piecewise convex polynomials. Math. Program. **137**(1), 37–64 (2013)
29. Li, G., Mordukhovich, B.S., Phạm, T.S.: New fractional error bounds for polynomial systems with applications to Hölderian stability in optimization and spectral theory of tensors. Math. Program. **153**(2), 333–362 (2015)
30. Li, G., Pong, T.K.: Calculus of the exponent of Kurdyka–Łojasiewicz inequality and its applications to linear convergence of first-order methods. Found. Comput. Math. **18**, 1199–1232 (2018)
31. Lindstrom, S.B.: The art of modern homo habilis mathematicus, or: What would Jon Borwein do? In: B. Sriraman (ed.) Handbook of the Mathematics of the Arts and Sciences. Springer (2020)
32. Liu, M., Pataki, G.: Exact duals and short certificates of infeasibility and weak infeasibility in conic linear programming. Math. Program. **167**(2), 435–480 (2018)
33. Liu, T., Lourenço, B.F.: Convergence analysis under consistent error bounds. Found. Comput. Math. (2022)
34. Lourenço, B.F.: Amenable cones: error bounds without constraint qualifications. Math. Program. **186**, 1–48 (2021)
35. Lourenço, B.F., Muramatsu, M., Tsuchiya, T.: Facial reduction and partial polyhedrality. SIAM J. Optim. **28**(3), 2304–2326 (2018)
36. Lourenço, B.F., Roshchina, V., Saunderson, J.: Amenable cones are particularly nice. SIAM J. Optim. **32**(3) (2022). https://doi.org/10.1137/20M138466X
37. MOSEK ApS: MOSEK Modeling Cookbook Release 3.2.2 (2020). https://docs.mosek.com/modeling-cookbook/index.html
38. Ng, K.F., Yang, W.H.: Error bounds for abstract linear inequality systems. SIAM J. Optim. **13**(1), 24–43 (2002)
39. O'Donoghue, B., Chu, E., Parikh, N., Boyd, S.: Conic optimization via operator splitting and homogeneous self-dual embedding. J. Optim. Theory Appl. **169**(3), 1042–1068 (2016)
40. Pang, J.S.: Error bounds in mathematical programming. Math. Program. **79**(1), 299–332 (1997)
41. Papp, D., Yıldız, S.: Alfonso: Matlab package for nonsymmetric conic optimization. To be appear in INFORMS Journal on Computing
42. Pataki, G.: The geometry of semidefinite programming. In: H. Wolkowicz, R. Saigal, L. Vandenberghe (eds.) Handbook of semidefinite programming: theory, algorithms, and applications. Kluwer, online version at http://www.unc.edu/~pataki/papers/chapter.pdf (2000)
43. Pataki, G.: On the closedness of the linear image of a closed convex cone. Math. Oper. Res. **32**(2), 395–412 (2007)

44. Pataki, G.: On the connection of facially exposed and nice cones. J. Math. Anal. Appl. **400**(1), 211–221 (2013)
45. Pataki, G.: Strong duality in conic linear programming: facial reduction and extended duals. In: Computational and Analytical Mathematics, vol. 50, pp. 613–634. Springer, New York (2013)
46. Roshchina, V.: Facially exposed cones are not always nice. SIAM J. Optim. **24**(1), 257–268 (2014)
47. Roshchina, V., Tunçel, L.: Facially dual complete (nice) cones and lexicographic tangents. SIAM J. Optim. **29**(3), 2363–2387 (2019)
48. Serrano, S.A.: Algorithms for unsymmetric cone optimization and an implementation for problems with the exponential cone. Ph.D. thesis, Stanford University (2015)
49. Skajaa, A., Ye, Y.: A homogeneous interior-point algorithm for nonsymmetric convex conic optimization. Math. Program. **150**(2), 391–422 (2015)
50. Sturm, J.F.: Error bounds for linear matrix inequalities. SIAM J. Optim. **10**(4), 1228–1248 (2000)
51. Sung, C.H., Tam, B.S.: A study of projectionally exposed cones. Linear Algebra Appl. **139**, 225–252 (1990)
52. Waki, H., Muramatsu, M.: Facial reduction algorithms for conic optimization problems. J. Optim. Theory Appl. **158**(1), 188–215 (2013)
53. Zhang, Q.: Completely positive cones: are they facially exposed? Linear Algebra Appl. **558**, 195–204 (2018)