# Error bound for conic feasibility problems: Case studies on Exponential cone and $p$-cones

Ting Kei Pong
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hong Kong

Talk @ Nanjing University
September 2022

(Joint work with Scott B. Lindstrom and Bruno F. Lourenço)

# Conic programming problem

Conic program: Let $\mathcal{K}$ be a closed convex cone in a Euclidean space $\mathcal{E}$, $c \in \mathcal{E}$, $\mathcal{A}$ be a linear map on $\mathcal{E}$ and $b \in \text{Range}(\mathcal{A})$.

$$\underset{x}{\text{Minimize}}\ \langle c, x \rangle \ \text{ subject to }\ \mathcal{A}x = b, \quad x \in \mathcal{K}.$$

# Conic programming problem

Conic program: Let $\mathcal{K}$ be a closed convex cone in a Euclidean space $\mathcal{E}$, $c \in \mathcal{E}$, $\mathcal{A}$ be a linear map on $\mathcal{E}$ and $b \in \text{Range}(\mathcal{A})$.

$$\underset{x}{\text{Minimize}} \ \langle c, x \rangle \ \text{ subject to } \ \mathcal{A}x = b, \quad x \in \mathcal{K}.$$

Examples of cones:

- $\mathbb{R}^n_+$, $\mathcal{S}^n_+$, second-order cones.

# Conic programming problem

Conic program: Let $\mathcal{K}$ be a closed convex cone in a Euclidean space $\mathcal{E}$, $c \in \mathcal{E}$, $\mathcal{A}$ be a linear map on $\mathcal{E}$ and $b \in \text{Range}(\mathcal{A})$.

$$\underset{x}{\text{Minimize}} \ \langle c, x \rangle \ \text{ subject to } \ \mathcal{A}x = b, \ \ x \in \mathcal{K}.$$

Examples of cones:

- $\mathbb{R}^n_+$, $\mathcal{S}^n_+$, second-order cones.
- Exponential cone $K_{\exp}$:

  $$K_{\exp} := \{x \in \mathbb{R}^3 \mid x_2 > 0, x_3 \geq x_2 e^{x_1/x_2}\} \cup \{(x_1, 0, x_3) \mid x_1 \leq 0, x_3 \geq 0\}.$$

  - ⋆ Epigraph of (the closure of) the perspective function of $z \mapsto \exp(z)$.
  - ⋆ Recent addition to MOSEK and other conic solvers.
  - ⋆ Has applications in relative entropy optimization (Chandrasekaran, Shah '17).

# Conic programming problem cont.

Examples of cones cont.:

- *p*-cones $K_p^{n+1}$ $(p > 1)$:

$$K_p^{n+1} := \{x = (x_0, \overline{x}) \in \mathbb{R}^{n+1} \mid x_0 \geq \|\overline{x}\|_p\}.$$

  - ⋆ Reduces to second-order cone when $p = 2$.
  - ⋆ Widely studied as natural generalization of second-order cones.

# Conic programming problem cont.

Examples of cones cont.:

- *p*-cones $K_p^{n+1}$ $(p > 1)$:

$$K_p^{n+1} := \{x = (x_0, \overline{x}) \in \mathbb{R}^{n+1} \mid x_0 \geq \|\overline{x}\|_p\}.$$

  - ⋆ Reduces to second-order cone when $p = 2$.
  - ⋆ Widely studied as natural generalization of second-order cones.

- Geometric mean cone:

$$K_{\text{geo}} := \{x = (x_0, \overline{x}) \in \mathbb{R} \times \mathbb{R}_+^n \mid x_0 \leq \Pi_{i=1}\overline{x}_i^{1/n}\}.$$

  - ⋆ Available in Hypatia. (Coey, Kapelevich, Vielma '21)
  - ⋆ Arises when modeling *Perron-Frobenius matrix completion* problems. (Agrawal, Diamond, Boyd '19)

# Conic programming problem cont.

Examples of cones cont.:

- $p$-cones $K_p^{n+1}$ ($p > 1$):

$$K_p^{n+1} := \{x = (x_0, \overline{x}) \in \mathbb{R}^{n+1} \mid x_0 \geq \|\overline{x}\|_p\}.$$

  - ⋆ Reduces to second-order cone when $p = 2$.
  - ⋆ Widely studied as natural generalization of second-order cones.

- Geometric mean cone:

$$K_{\text{geo}} := \{x = (x_0, \overline{x}) \in \mathbb{R} \times \mathbb{R}_+^n \mid x_0 \leq \Pi_{i=1} \overline{x}_i^{1/n}\}.$$

  - ⋆ Available in Hypatia. (Coey, Kapelevich, Vielma '21)
  - ⋆ Arises when modeling *Perron-Frobenius matrix completion* problems. (Agrawal, Diamond, Boyd '19)

- Product cones of the above...

# Conic feasibility problem

Conic feasibility problem: Let $\mathcal{K}$ be a closed convex cone, $\mathcal{L}$ be a subspace of a Euclidean space $\mathcal{E}$ and $a \in \mathcal{E}$.

$$\text{Find} \quad x \in \mathcal{K} \cap (\mathcal{L} + a).$$

# Conic feasibility problem

Conic feasibility problem: Let $\mathcal{K}$ be a closed convex cone, $\mathcal{L}$ be a subspace of a Euclidean space $\mathcal{E}$ and $a \in \mathcal{E}$.

$$\text{Find} \quad x \in \mathcal{K} \cap (\mathcal{L} + a).$$

- Focus on the feasible case, i.e., $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$.
- Arises from optimality conditions of conic programs.

# Conic feasibility problem

Conic feasibility problem: Let $\mathcal{K}$ be a closed convex cone, $\mathcal{L}$ be a subspace of a Euclidean space $\mathcal{E}$ and $a \in \mathcal{E}$.

$$\text{Find} \quad x \in \mathcal{K} \cap (\mathcal{L} + a).$$

- Focus on the feasible case, i.e., $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$.
- Arises from optimality conditions of conic programs.

When is a feasibility problem approximately solved?

# Conic feasibility problem

Conic feasibility problem: Let $\mathcal{K}$ be a closed convex cone, $\mathcal{L}$ be a subspace of a Euclidean space $\mathcal{E}$ and $a \in \mathcal{E}$.

$$\text{Find} \quad x \in \mathcal{K} \cap (\mathcal{L} + a).$$

- Focus on the feasible case, i.e., $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$.
- Arises from optimality conditions of conic programs.

When is a feasibility problem approximately solved?

- $\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a))$ is a measure on how “feasible” $x$ is.

# Conic feasibility problem

Conic feasibility problem: Let $\mathcal{K}$ be a closed convex cone, $\mathcal{L}$ be a subspace of a Euclidean space $\mathcal{E}$ and $a \in \mathcal{E}$.

$$\text{Find} \quad x \in \mathcal{K} \cap (\mathcal{L} + a).$$

- Focus on the feasible case, i.e., $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$.
- Arises from optimality conditions of conic programs.

When is a feasibility problem approximately solved?

- $\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a))$ is a measure on how "feasible" $x$ is. Hard to compute!
- Typically, $\mathrm{d}(x, \mathcal{K})$ and $\mathrm{d}(x, \mathcal{L} + a)$ are relatively easier to compute.
- Is $x$ "a good soln." when $\max\{\mathrm{d}(x, \mathcal{K}), \mathrm{d}(x, \mathcal{L} + a)\}$ is small?

# Conic feasibility problem

Conic feasibility problem: Let $\mathcal{K}$ be a closed convex cone, $\mathcal{L}$ be a subspace of a Euclidean space $\mathcal{E}$ and $a \in \mathcal{E}$.

$$\text{Find} \quad x \in \mathcal{K} \cap (\mathcal{L} + a).$$

- Focus on the feasible case, i.e., $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$.
- Arises from optimality conditions of conic programs.

When is a feasibility problem approximately solved?

- $\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a))$ is a measure on how "feasible" $x$ is. Hard to compute!
- Typically, $\mathrm{d}(x, \mathcal{K})$ and $\mathrm{d}(x, \mathcal{L} + a)$ are relatively easier to compute.
- Is $x$ "a good soln." when $\max\{\mathrm{d}(x, \mathcal{K}), \mathrm{d}(x, \mathcal{L} + a)\}$ is small?

Key: Compare the orders of magnitude of $\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a))$ and $\max\{\mathrm{d}(x, \mathcal{K}), \mathrm{d}(x, \mathcal{L} + a)\}$.

# Conic feasibility problem

Conic feasibility problem: Let $\mathcal{K}$ be a closed convex cone, $\mathcal{L}$ be a subspace of a Euclidean space $\mathcal{E}$ and $a \in \mathcal{E}$.

$$\text{Find} \quad x \in \mathcal{K} \cap (\mathcal{L} + a).$$

- Focus on the feasible case, i.e., $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$.
- Arises from optimality conditions of conic programs.

When is a feasibility problem approximately solved?

- $\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a))$ is a measure on how “feasible” $x$ is. Hard to compute!
- Typically, $\mathrm{d}(x, \mathcal{K})$ and $\mathrm{d}(x, \mathcal{L} + a)$ are relatively easier to compute.
- Is $x$ “a good soln.” when $\max\{\mathrm{d}(x, \mathcal{K}), \mathrm{d}(x, \mathcal{L} + a)\}$ is small?

Key: Compare the orders of magnitude of $\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a))$ and $\max\{\mathrm{d}(x, \mathcal{K}), \mathrm{d}(x, \mathcal{L} + a)\}$.

Note: Typically, $(\mathcal{L} + a) \cap \mathrm{ri}\, \mathcal{K} = \emptyset$.

# Error bounds

Definition: Let $\theta \in (0, 1]$. We say that $\{\mathcal{K}, \mathcal{L} + a\}$ satisfies a (uniform) Hölderian error bound with exponent $\theta$ if for every bounded set $B$, there exists $c_B > 0$ such that

$$\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a)) \leq c_B \left(\max\{\mathrm{d}(x, \mathcal{K}), \mathrm{d}(x, \mathcal{L} + a)\}\right)^\theta \quad \forall x \in B.$$

If $\theta = 1$, we say that a Lipschitz error bound holds.

# Error bounds

Definition: Let $\theta \in (0, 1]$. We say that $\{\mathcal{K}, \mathcal{L} + a\}$ satisfies a (uniform) Hölderian error bound with exponent $\theta$ if for every bounded set $B$, there exists $c_B > 0$ such that

$$\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a)) \leq c_B \left(\max\{\mathrm{d}(x, \mathcal{K}), \mathrm{d}(x, \mathcal{L} + a)\}\right)^{\theta} \quad \forall x \in B.$$

If $\theta = 1$, we say that a Lipschitz error bound holds.

Examples:

- If $\mathcal{K}$ is polyhedral, Lipschitz error bound holds. (Hoffman '57)
- If $(\mathcal{L} + a) \cap \mathrm{ri}\,\mathcal{K} \neq \emptyset$, Lipschitz error bound holds. (Bauschke, Borwein '96)

# Error bounds

Definition: Let $\theta \in (0, 1]$. We say that $\{\mathcal{K}, \mathcal{L} + a\}$ satisfies a (uniform) Hölderian error bound with exponent $\theta$ if for every bounded set $B$, there exists $c_B > 0$ such that

$$\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a)) \le c_B \left(\max\{\mathrm{d}(x, \mathcal{K}), \mathrm{d}(x, \mathcal{L} + a)\}\right)^{\theta} \quad \forall x \in B.$$

If $\theta = 1$, we say that a Lipschitz error bound holds.

Examples:

- If $\mathcal{K}$ is polyhedral, Lipschitz error bound holds. (Hoffman '57)
- If $(\mathcal{L} + a) \cap \mathrm{ri}\, \mathcal{K} \neq \emptyset$, Lipschitz error bound holds. (Bauschke, Borwein '96)
- If $\mathcal{K} = \mathcal{S}^n_+$, Hölderian error bound with exponent $2^{-(\ell-1)}$ holds (Sturm '00); here $\ell$ has to do with facial reduction. (Borwein, Wolkowicz '81)

# Faces and facial reduction

Definition: A sub-cone $\mathcal{F} \subseteq \mathcal{K}$ is called

- a face if $x, \ y \in \mathcal{K}$ and $x + y \in \mathcal{F}$ implies $x, \ y \in \mathcal{F}$;
- an exposed face if $\exists \ z \in \mathcal{K}^*$ such that $\mathcal{F} = \mathcal{K} \cap \{z\}^\perp$.

Note: Recall that $\mathcal{K}^* := \{x \mid \langle x, y \rangle \geq 0 \ \ \forall y \in \mathcal{K}\}$.

# Faces and facial reduction

Definition: A sub-cone $\mathcal{F} \subseteq \mathcal{K}$ is called

- a face if $x, \ y \in \mathcal{K}$ and $x + y \in \mathcal{F}$ implies $x, \ y \in \mathcal{F}$;
- an exposed face if $\exists\, z \in \mathcal{K}^*$ such that $\mathcal{F} = \mathcal{K} \cap \{z\}^{\perp}$.

Note: Recall that $\mathcal{K}^* := \{x \mid \langle x, y\rangle \geq 0 \ \ \forall y \in \mathcal{K}\}$.

**Theorem 1.** (Lourenço, Muramatsu, Tsuchiya '18)
Suppose $\mathcal{K} \cap (\mathcal{L} + a) \neq \emptyset$. Then there exists a chain of faces

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

and vectors $\{z_1, \ldots, z_{\ell-1}\}$ satisfying

- For all $i \in \{1, \ldots, \ell - 1\}$,

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

- $\mathcal{F}_\ell \cap (\mathcal{L} + a) = \mathcal{K} \cap (\mathcal{L} + a)$ and $\{\mathcal{F}_\ell, \mathcal{L} + a\}$ satisfies a Lipschitz error bound.

# Sturm's error bounds and facial reduction

Key observation: (Sturm '00, Lourenço '21)
Let $\mathcal{F} \trianglelefteq \mathcal{S}_+^n$ and $z \in \mathcal{F}^*$. Then $\exists\, \kappa > 0$ such that for all $x$,

$$\mathrm{d}(x, \mathcal{F} \cap \{z\}^\perp) \le \kappa\epsilon + \kappa\sqrt{\epsilon\|x\|},$$

where $\epsilon = \max\{\mathrm{d}(x,\mathcal{F}), \mathrm{d}(x,\{z\}^\perp)\}$.

# Sturm's error bounds and facial reduction

Key observation: (Sturm '00, Lourenço '21)
Let $\mathcal{F} \trianglelefteq \mathcal{S}^n_+$ and $z \in \mathcal{F}^*$. Then $\exists\ \kappa > 0$ such that for all $x$,

$$\mathrm{d}(x, \mathcal{F} \cap \{z\}^\perp) \leq \kappa\epsilon + \kappa\sqrt{\epsilon\|x\|},$$

where $\epsilon = \max\{\mathrm{d}(x,\mathcal{F}), \mathrm{d}(x,\{z\}^\perp)\}$.

Error bound for $\{\mathcal{S}^n_+, (\mathcal{L}+a)\}$ follows from induction: For $x \in B$,

$$\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L}+a)) = \mathrm{d}(x, \mathcal{F}_\ell \cap (\mathcal{L}+a)) \leq c_\ell \max\{\mathrm{d}(x,\mathcal{F}_\ell), \mathrm{d}(x,\mathcal{L}+a)\}$$

# Sturm's error bounds and facial reduction

Key observation: (Sturm '00, Lourenço '21)
Let $\mathcal{F} \trianglelefteq \mathcal{S}^n_+$ and $z \in \mathcal{F}^*$. Then $\exists\, \kappa > 0$ such that for all $x$,

$$\mathrm{d}(x, \mathcal{F} \cap \{z\}^\perp) \le \kappa\epsilon + \kappa\sqrt{\epsilon\|x\|},$$

where $\epsilon = \max\{\mathrm{d}(x, \mathcal{F}), \mathrm{d}(x, \{z\}^\perp)\}$.

Error bound for $\{\mathcal{S}^n_+, (\mathcal{L} + a)\}$ follows from induction: For $x \in B$,

$$\begin{aligned}
&\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a)) = \mathrm{d}(x, \mathcal{F}_\ell \cap (\mathcal{L} + a)) \le c_\ell \max\{\mathrm{d}(x, \mathcal{F}_\ell), \mathrm{d}(x, \mathcal{L} + a)\} \\
&\le c_\ell[\mathrm{d}(x, \mathcal{L} + a) + \mathrm{d}(x, \mathcal{F}_{\ell-1} \cap \{z_{\ell-1}\}^\perp)] \\
&\le c_\ell[\mathrm{d}(x, \mathcal{L} + a) + c_{\ell-1}\big(\max\{\mathrm{d}(x, \mathcal{F}_{\ell-1}), \mathrm{d}(x, \{z_{\ell-1}\}^\perp)\} \\
&\qquad + \sqrt{\max\{\mathrm{d}(x, \mathcal{F}_{\ell-1}), \mathrm{d}(x, \{z_{\ell-1}\}^\perp)\}\|x\|}\big)]
\end{aligned}$$

# Sturm's error bounds and facial reduction

Key observation: (Sturm '00, Lourenço '21)
Let $\mathcal{F} \trianglelefteq \mathcal{S}^n_+$ and $z \in \mathcal{F}^*$. Then $\exists\, \kappa > 0$ such that for all $x$,

$$\mathrm{d}(x, \mathcal{F} \cap \{z\}^\perp) \le \kappa\epsilon + \kappa\sqrt{\epsilon\|x\|},$$

where $\epsilon = \max\{\mathrm{d}(x, \mathcal{F}), \mathrm{d}(x, \{z\}^\perp)\}$.

Error bound for $\{\mathcal{S}^n_+, (\mathcal{L} + a)\}$ follows from induction: For $x \in B$,

$$\begin{aligned}
&\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L} + a)) = \mathrm{d}(x, \mathcal{F}_\ell \cap (\mathcal{L} + a)) \le c_\ell \max\{\mathrm{d}(x, \mathcal{F}_\ell), \mathrm{d}(x, \mathcal{L} + a)\} \\
&\le c_\ell[\mathrm{d}(x, \mathcal{L} + a) + \mathrm{d}(x, \mathcal{F}_{\ell-1} \cap \{z_{\ell-1}\}^\perp)] \\
&\le c_\ell[\mathrm{d}(x, \mathcal{L} + a) + c_{\ell-1}\big(\max\{\mathrm{d}(x, \mathcal{F}_{\ell-1}), \mathrm{d}(x, \{z_{\ell-1}\}^\perp)\} \\
&\qquad + \sqrt{\max\{\mathrm{d}(x, \mathcal{F}_{\ell-1}), \mathrm{d}(x, \{z_{\ell-1}\}^\perp)\}\|x\|}\big)] \\
&= O([\max\{\mathrm{d}(x, \mathcal{F}_{\ell-1}), \mathrm{d}(x, \mathcal{L} + a)\}]^{\frac{1}{2}})
\end{aligned}$$

# Sturm's error bounds and facial reduction

Key observation: (Sturm '00, Lourenço '21)
Let $\mathcal{F} \trianglelefteq \mathcal{S}^n_+$ and $z \in \mathcal{F}^*$. Then $\exists\, \kappa > 0$ such that for all $x$,

$$\mathrm{d}(x, \mathcal{F} \cap \{z\}^\perp) \le \kappa\epsilon + \kappa\sqrt{\epsilon\|x\|},$$

where $\epsilon = \max\{\mathrm{d}(x,\mathcal{F}), \mathrm{d}(x,\{z\}^\perp)\}$.

Error bound for $\{\mathcal{S}^n_+, (\mathcal{L}+a)\}$ follows from induction: For $x \in B$,

$$\begin{aligned}
&\mathrm{d}(x, \mathcal{K} \cap (\mathcal{L}+a)) = \mathrm{d}(x, \mathcal{F}_\ell \cap (\mathcal{L}+a)) \le c_\ell \max\{\mathrm{d}(x,\mathcal{F}_\ell), \mathrm{d}(x,\mathcal{L}+a)\}\\
&\le c_\ell[\mathrm{d}(x,\mathcal{L}+a) + \mathrm{d}(x, \mathcal{F}_{\ell-1} \cap \{z_{\ell-1}\}^\perp)]\\
&\le c_\ell[\mathrm{d}(x,\mathcal{L}+a) + c_{\ell-1}\big(\max\{\mathrm{d}(x,\mathcal{F}_{\ell-1}), \mathrm{d}(x,\{z_{\ell-1}\}^\perp)\}\\
&\qquad + \sqrt{\max\{\mathrm{d}(x,\mathcal{F}_{\ell-1}), \mathrm{d}(x,\{z_{\ell-1}\}^\perp)\}\|x\|}\big)]\\
&= O([\max\{\mathrm{d}(x,\mathcal{F}_{\ell-1}), \mathrm{d}(x,\mathcal{L}+a)\}]^{\frac{1}{2}}) = O([\max\{\mathrm{d}(x,\mathcal{K}), \mathrm{d}(x,\mathcal{L}+a)\}]^{\frac{1}{2^{\ell-1}}}).
\end{aligned}$$

# Facial residual function

Definition: (Lourenço '21, Lindstrom, Lourenço, P. '22)
Let $\mathcal{F} \trianglelefteq \mathcal{K}$ and $z \in \mathcal{F}^*$. Suppose $\psi : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies

- $\psi$ is nondecreasing in each argument and $\psi(0, t) = 0 \ \forall \ t \in \mathbb{R}_+$;
- It holds that

$$\mathrm{d}(x, \mathcal{F} \cap \{z\}^\perp) \leq \psi(\max\{\mathrm{d}(x, \mathcal{F}), \mathrm{d}(x, \{z\}^\perp)\}, \|x\|) \quad \forall x \in \operatorname{span} \mathcal{F}.$$

Then $\psi$ is called a 1-step facial residual function for $\mathcal{F}$ and $z$.

# Facial residual function

Definition: (Lourenço '21, Lindstrom, Lourenço, P. '22)
Let $\mathcal{F} \trianglelefteq \mathcal{K}$ and $z \in \mathcal{F}^*$. Suppose $\psi : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies

- $\psi$ is nondecreasing in each argument and $\psi(0, t) = 0 \;\forall\, t \in \mathbb{R}_+$;
- It holds that

$$\mathrm{d}(x, \mathcal{F} \cap \{z\}^\perp) \leq \psi(\max\{\mathrm{d}(x, \mathcal{F}), \mathrm{d}(x, \{z\}^\perp)\}, \|x\|) \quad \forall x \in \operatorname{span} \mathcal{F}.$$

Then $\psi$ is called a 1-step facial residual function for $\mathcal{F}$ and $z$.

Remarks:

- For $\mathcal{K} = \mathcal{S}^n_+$, we have $\psi(s, t) = \kappa \cdot (s + \sqrt{st})$ for some $\kappa > 0$.
- Induction arguments show that error bound can be derived given the face chain from facial reduction and by composing 1-step facial residual functions. (Lindstrom, Lourenço, P. '22)

# Facial residual function

Definition: (Lourenço '21, Lindstrom, Lourenço, P. '22)
Let $\mathcal{F} \trianglelefteq \mathcal{K}$ and $z \in \mathcal{F}^*$. Suppose $\psi : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies

- $\psi$ is nondecreasing in each argument and $\psi(0, t) = 0 \; \forall \; t \in \mathbb{R}_+$;
- It holds that

$$\mathrm{d}(x, \mathcal{F} \cap \{z\}^\perp) \leq \psi(\max\{\mathrm{d}(x, \mathcal{F}), \mathrm{d}(x, \{z\}^\perp)\}, \|x\|) \quad \forall x \in \operatorname{span} \mathcal{F}.$$

Then $\psi$ is called a 1-step facial residual function for $\mathcal{F}$ and $z$.

Remarks:

- For $\mathcal{K} = \mathcal{S}^n_+$, we have $\psi(s, t) = \kappa \cdot (s + \sqrt{st})$ for some $\kappa > 0$.
- Induction arguments show that error bound can be derived given the face chain from facial reduction and by composing 1-step facial residual functions. (Lindstrom, Lourenço, P. '22)
- Two key ingredients: For each nonpolyhedral $\mathcal{F}$,
  - ⋆ Identify all its exposed faces.
  - ⋆ Obtain all 1-step facial residual functions: Depends on $\mathcal{F}$ and $z$.

# Outline

Aim: Case studies on errors bounds for $K_{\exp}$ and $p$-cones ($K_p^{n+1}$).

1. Derive error bounds for $\{K_{\exp}, \mathcal{L} + a\}$.
   - ⋆ Describe all nontrivial exposed faces of exponential cone $K_{\exp}$.
   - ⋆ Find the associated 1-step facial residual functions ($\mathbb{1}$-FRFs).
   - ⋆ Discuss our strategy for computing $\mathbb{1}$-FRFs.
2. Derive error bounds for $\{K_p^{n+1}, \mathcal{L} + a\}$ with $p \in (1, \infty)$.
   - ⋆ Describe all nontrivial exposed faces of $p$-cone $K_p^{n+1}$.
   - ⋆ Find the associated 1-step facial residual functions ($\mathbb{1}$-FRFs).
   - ⋆ Discuss our strategy for computing $\mathbb{1}$-FRFs.

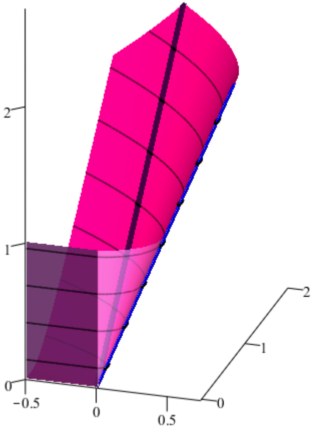
# Outline

Aim: Case studies on errors bounds for $K_{\exp}$ and $p$-cones ($K_p^{n+1}$).

1. Derive error bounds for $\{K_{\exp}, \mathcal{L} + a\}$.
   - ⋆ Describe all nontrivial exposed faces of exponential cone $K_{\exp}$.
   - ⋆ Find the associated 1-step facial residual functions ($\mathbb{1}$-FRFs).
   - ⋆ Discuss our strategy for computing $\mathbb{1}$-FRFs.
2. Derive error bounds for $\{K_p^{n+1}, \mathcal{L} + a\}$ with $p \in (1, \infty)$.
   - ⋆ Describe all nontrivial exposed faces of $p$-cone $K_p^{n+1}$.
   - ⋆ Find the associated 1-step facial residual functions ($\mathbb{1}$-FRFs).
   - ⋆ Discuss our strategy for computing $\mathbb{1}$-FRFs.
3. Applications:
   - ⋆ (Informally) The infimum of Kurdyka-Łojasiewicz (KL) exponents may not be a KL exponent.
   - ⋆ Finding KL exponent for some $p$-norm regularized problems.

# Exponential cone

$$K_{\exp} = \{x \in \mathbb{R}^3 \mid x_2 > 0, x_3 \geq x_2 e^{x_1/x_2}\} \cup \{(x_1, 0, x_3) \mid x_1 \leq 0, x_3 \geq 0\}.$$



# Faces of exponential cone

$$K_{\exp} = \{x \in \mathbb{R}^3 \mid x_2 > 0, x_3 \geq x_2 e^{x_1/x_2}\} \cup \{(x_1, 0, x_3) \mid x_1 \leq 0, x_3 \geq 0\}.$$

Nontrivial exposed faces: (Lindstrom, Lourenço, P. '22)

- (Infinitely many) 1-D face exposed by a unique (up to scaling) $z \in \partial K^*_{\exp} \backslash \{0\}$:

$$\mathcal{F}_\beta := \{(t(1-\beta), t, te^{1-\beta} \mid t \geq 0\} \text{ for each } \beta \in \mathbb{R}.$$

- An exceptional extreme ray exposed by any $z \in \{(0, z_2, z_3) \mid z_2 \geq 0, z_3 > 0\} \subset \partial K^*_{\exp} \backslash \{0\}$:

$$\mathcal{F}_\infty := \{(x_1, 0, 0) \mid x_1 \leq 0\}.$$

- A unique 2-D face exposed uniquely (up to scaling) by $(0, 1, 0)$:

$$\mathcal{F}_{-\infty} := \{(x_1, 0, x_3) \mid x_1 \leq 0, x_3 \geq 0\}.$$

# $\mathbb{1}$-FRFs for exponential cone

$\mathbb{1}$-FRFs: (Lindstrom, Lourenço, P. '22) Assume $\|z\| = 1$.

- (Infinitely many) 1-D face exposed by a unique $z \in \partial K^*_{\exp} \setminus \{0\}$:

$$\psi(s,t) = s + \kappa(t) \cdot \sqrt{2s},$$

  where $\kappa(\cdot)$ is nonnegative nondecreasing.

- The exceptional extreme ray $\mathcal{F}_\infty$:

$$\psi(s,t) = \begin{cases} s + 2\kappa(t)s & \text{if } z_2 > 0, \\ s + \kappa(t) \cdot \mathfrak{g}_\infty(2s) & \text{if } z_2 = 0, \end{cases}$$

  where $\kappa(\cdot)$ is nonnegative nondecreasing, and $\mathfrak{g}_\infty : \mathbb{R}_+ \to \mathbb{R}_+$ is

$$\mathfrak{g}_\infty(s) := \begin{cases} 0 & \text{if } s = 0, \\ -(\ln s)^{-1} & \text{if } 0 < s \le e^{-2}, \\ 0.25(1 + e^2 s) & \text{else.} \end{cases}$$

# $\mathbb{1}$-FRFs for exponential cone cont.

$\mathbb{1}$-FRFs cont.: Assume $\|z\| = 1$.

- The unique 2-D face exposed uniquely by $(0,1,0)$:

$$\psi(s,t) = s + \kappa(t) \cdot \mathfrak{g}_{-\infty}(2s),$$

where $\kappa(\cdot)$ is nonnegative nondecreasing, $\mathfrak{g}_{-\infty} : \mathbb{R}_+ \to \mathbb{R}_+$ is

$$\mathfrak{g}_{-\infty}(s) := \begin{cases} 0 & \text{if } s = 0, \\ -s \ln s & \text{if } 0 < s \le e^{-2}, \\ s + e^{-2} & \text{else.} \end{cases}$$

# $\mathbb{1}$-FRFs for exponential cone cont.

$\mathbb{1}$-FRFs cont.: Assume $\|z\| = 1$.

- The unique 2-D face exposed uniquely by $(0, 1, 0)$:

$$\psi(s,t) = s + \kappa(t) \cdot \mathfrak{g}_{-\infty}(2s),$$

where $\kappa(\cdot)$ is nonnegative nondecreasing, $\mathfrak{g}_{-\infty} : \mathbb{R}_+ \to \mathbb{R}_+$ is

$$\mathfrak{g}_{-\infty}(s) := \begin{cases} 0 & \text{if } s = 0, \\ -s\ln s & \text{if } 0 < s \le e^{-2}, \\ s + e^{-2} & \text{else.} \end{cases}$$

Remark:

- As all nontrivial exposed faces are polyhedral, $\ell \le 2$. The final error bound has the same "order of magnitude" as the $\mathbb{1}$-FRFs.
- As $s \downarrow 0$: $-(\ln s)^{-1} \to 0$ slower than $s^\alpha$ for any $\alpha \in (0, 1]$; $-s\ln s \to 0$ faster than $s^\alpha$ for any $\alpha \in (0, 1)$, but is slower than $s$.
- These $\mathbb{1}$-FRFs are asymptotically the "best". (Lindstrom, Lourenço, P. '22)

# Computing $\mathbb{1}$-FRFs

Difficulties:

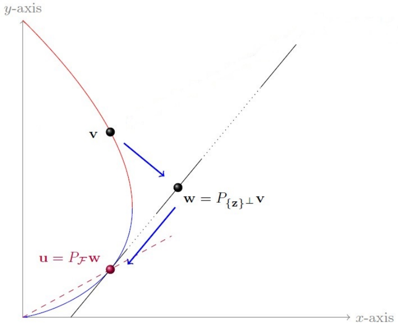
- Need to compare $\mathrm{d}(x, K_{\exp})$, $\mathrm{d}(x, \{z\}^\perp)$, and $\mathrm{d}(x, K_{\exp} \cap \{z\}^\perp)$.
- Projection onto $K_{\exp}$ (and hence $\mathrm{d}(x, K_{\exp})$) does not have an easy-to-analyze analytic form.

# Computing $\mathbb{1}$-FRFs

Difficulties:

- Need to compare $\mathrm{d}(x, K_{\exp})$, $\mathrm{d}(x, \{z\}^{\perp})$, and $\mathrm{d}(x, K_{\exp} \cap \{z\}^{\perp})$.
- Projection onto $K_{\exp}$ (and hence $\mathrm{d}(x, K_{\exp})$) does not have an easy-to-analyze analytic form.

The $v - w - u$ approach:



# Computing $\mathbb{1}$-FRFs cont.

**Theorem 2.** (Lindstrom, Lourenço, P. '22)
Let $\mathcal{K}$ be a closed convex cone and $z \in \mathcal{K}^*$ (with $\|z\| = 1$) be such that $\{z\}^\perp \cap \mathcal{K}$ is a nontrivial exposed face of $\mathcal{K}$.
Let $\eta > 0$, $\alpha \in (0, 1]$ and $\mathfrak{g} : \mathbb{R}_+ \to \mathbb{R}_+$ be nondecreasing and satisfy $\mathfrak{g}(0) = 0$ and $\mathfrak{g} \geq |\cdot|^\alpha$. Consider

$$\gamma_{z,\eta} := \inf_v \left\{ \frac{\mathfrak{g}(\|w - v\|)}{\|w - u\|} \,\middle|\, \begin{array}{c} v \in \partial\mathcal{K} \cap B(\eta) \setminus (\{z\}^\perp \cap \mathcal{K}) \\ w = \mathrm{Proj}_{\{z\}^\perp} v, \;\; u = \mathrm{Proj}_{\{z\}^\perp \cap \mathcal{K}} w, \;\; w \neq u \end{array} \right\}.$$

Suppose that $\gamma_{z,\eta} \in (0, \infty]$. Then

$$\mathrm{d}(x, \{z\}^\perp \cap \mathcal{K}) \leq \kappa_{z,\eta} \mathfrak{g}(\mathrm{d}(x, \mathcal{K})) \quad \forall\, x \in \{z\}^\perp \cap B(0, \eta),$$

where $\kappa_{z,\eta} := \max\{2\eta^{1-\alpha}, 2\gamma_{z,\eta}^{-1}\}$. Moreover,

$$\psi(s, t) := s + \kappa_{z,t}\mathfrak{g}(2s)$$

is a $\mathbb{1}$-FRF of $\mathcal{K}$ and $z$.

# Main idea

We illustrate how **Theorem 2** can be used for computing $\mathbb{1}$-FRF for those (infinitely many) 1-D face $\mathcal{F}_\beta$.

# Main idea

We illustrate how **Theorem 2** can be used for computing $\mathbb{1}$-FRF for those (infinitely many) 1-D face $\mathcal{F}_\beta$.

**Key Lemma 1.** (Lindstrom, Lourenço, P. '22)
Let $\beta \in \mathbb{R}$ and $z \in K_{\exp}^*$ with $z_1 < 0$ such that $\mathcal{F}_\beta = \{z\}^\perp \cap K_{\exp}$. Let $\eta > 0$, $v \in \partial K_{\exp} \cap B(\eta) \backslash F_\beta$, $w = \mathrm{Proj}_{\{z\}^\perp} v$ and $u = \mathrm{Proj}_{\mathcal{F}_\beta} w$. Then

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

where

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

Moreover, when $\langle \hat{f}, v\rangle \geq 0$, we have $u = P_{\mathrm{span} F_\beta} w$.

# Main idea cont.

Consider those (infinitely many) 1-D face $\mathcal{F}_\beta$ exposed by a unique (up to scaling) $z \in \partial K_{\exp}^* \backslash \{0\}$: Let $\|z\| = 1$ and $\mathcal{F}_\beta = \text{cone}\,\{\hat{f}\}$.

# Main idea cont.

Consider those (infinitely many) 1-D face $\mathcal{F}_\beta$ exposed by a unique (up to scaling) $z \in \partial K^*_{\exp}\backslash\{0\}$: Let $\|z\| = 1$ and $\mathcal{F}_\beta = \text{cone}\,\{\hat{f}\}$.

- If $\langle f, v\rangle \geq 0$ and $v \notin \mathcal{F}_{-\infty}$, then $v = (v_1, v_2, v_2 e^{v_1/v_2})$ and

  $$\|w - v\| = \Omega(v_2 \cdot |h_1(v_1/v_2)|) \;\text{ and }\; \|w - u\| = O(v_2 \cdot |h_2(v_1/v_2)|),$$

  where

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

  For those $v_1/v_2$ close to $1-\beta$, the exponent of $1/2$ pops up upon comparing Taylor series at $\zeta = 1-\beta$.
- Other cases can be dealt with similarly.

# Faces of $p$-cone

We focus on the case where $p \in (1, \infty) \backslash \{2\}$ and $n \geq 2$.

$$K_p^{n+1} = \{x = (x_0, \overline{x}) \in \mathbb{R}^{n+1} \mid x_0 \geq \|\overline{x}\|_p\}.$$

It is known that $(K_p^{n+1})^* = K_q^{n+1}$, where $1/p + 1/q = 1$.

# Faces of $p$-cone

We focus on the case where $p \in (1, \infty)\backslash\{2\}$ and $n \geq 2$.

$$K_p^{n+1} = \{x = (x_0, \overline{x}) \in \mathbb{R}^{n+1} \mid x_0 \geq \|\overline{x}\|_p\}.$$

It is known that $(K_p^{n+1})^* = K_q^{n+1}$, where $1/p + 1/q = 1$.

Nontrivial exposed faces: (Lindstrom, Lourenço, P. '22)

- (Infinitely many) 1-D face exposed by a unique (up to scaling) $z := (z_0, \overline{z}) \in \partial K_q^{n+1}\backslash\{0\}$:

  $$\mathcal{F}_z := \{t \cdot f \mid t \geq 0\} = K_p^{n+1} \cap \{z\}^\perp,$$

  where

  $$f := \begin{bmatrix} 1 \\ -\mathrm{sgn}(\overline{z}) \circ |z_0^{-1}\overline{z}|^{q-1} \end{bmatrix},$$

  where the sign, inverse, absolute value and power are taken componentwise.

# $\mathbb{1}$-FRFs for $p$-cone

$\mathbb{1}$-FRFs: (Lindstrom, Lourenço, P. '22) Assume $\|z\| = 1$.

- (Infinitely many) 1-D face exposed by the unique $z \in \partial K_q^{n+1} \backslash \{0\}$:

$$\psi(s,t) = s + \kappa(t) \cdot (2s)^{\alpha_z},$$

where $\kappa(\cdot)$ is nonnegative nondecreasing, and

$$\alpha_z := \begin{cases} \frac{1}{2} & \text{if } |J_z| = n, \\ \frac{1}{p} & \text{if } |J_z| = 1 \text{ and } p < 2, \\ \min\left\{\frac{1}{2}, \frac{1}{p}\right\} & \text{otherwise,} \end{cases}$$

with $J_z := \{i \mid \overline{z}_i \neq 0\}$.

# $\mathbb{1}$-FRFs for $p$-cone

$\mathbb{1}$-FRFs: (Lindstrom, Lourenço, P. '22) Assume $\|z\| = 1$.

- (Infinitely many) 1-D face exposed by the unique $z \in \partial K_q^{n+1} \backslash \{0\}$:

$$\psi(s,t) = s + \kappa(t) \cdot (2s)^{\alpha_z},$$

  where $\kappa(\cdot)$ is nonnegative nondecreasing, and

$$\alpha_z := \begin{cases} \frac{1}{2} & \text{if } |J_z| = n, \\ \frac{1}{p} & \text{if } |J_z| = 1 \text{ and } p < 2, \\ \min\left\{\frac{1}{2}, \frac{1}{p}\right\} & \text{otherwise}, \end{cases}$$

  with $J_z := \{i \mid \bar{z}_i \neq 0\}$.

Remark:

- Since all nontrivial faces are polyhedral, $\ell \leq 2$. Hence, the final error bound has the same "order of magnitude" as the $\mathbb{1}$-FRFs.
- These $\mathbb{1}$-FRFs are asymptotically the "best". (Lindstrom, Lourenço, P. '22)

# Main idea

How do the $f$ and the weird exponent arise?

# Main idea

How do the $f$ and the weird exponent arise?

**Key Lemma 2.** (Lindstrom, Lourenço, P. '22)

Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$ and $\zeta \in \mathbb{R}^n$ satisfy $\|\zeta\|_q = 1$. Define

$$\overline{\zeta} := -\text{sgn}(\zeta) \circ |\zeta|^{q-1},$$

Then $\|\overline{\zeta}\|_p = 1$ and the following statements hold:

- There exist $C > 0$ and $\epsilon > 0$ so that

$$1 + \langle \zeta, \omega \rangle \geq C \sum_{i \in I} |\omega_i - \overline{\zeta}_i|^2 + \frac{1}{p} \sum_{i \notin I} |\omega_i|^p$$

  whenever $\|\omega - \overline{\zeta}\| \leq \epsilon$ and $\|\omega\|_p = 1$, where $I = \{i : \overline{\zeta}_i \neq 0\}$.
- For any $\omega$ satisfying $\|\omega\|_p \leq 1$, it holds that $\langle \zeta, \omega \rangle \geq -1$, with the equality holding if and only if $\omega = \overline{\zeta}$.

# Main idea

How do the $f$ and the weird exponent arise?

**Key Lemma 2.** (Lindstrom, Lourenço, P. '22)

Let $p, q \in (1, \infty)$ be such that $\frac{1}{p} + \frac{1}{q} = 1$ and $\zeta \in \mathbb{R}^n$ satisfy $\|\zeta\|_q = 1$. Define

$$\overline{\zeta} := -\mathrm{sgn}(\zeta) \circ |\zeta|^{q-1},$$

Then $\|\overline{\zeta}\|_p = 1$ and the following statements hold:

- There exist $C > 0$ and $\epsilon > 0$ so that

$$1 + \langle \zeta, \omega \rangle \geq C \sum_{i \in I} |\omega_i - \overline{\zeta}_i|^2 + \frac{1}{p} \sum_{i \notin I} |\omega_i|^p$$

  whenever $\|\omega - \overline{\zeta}\| \leq \epsilon$ and $\|\omega\|_p = 1$, where $I = \{i : \overline{\zeta}_i \neq 0\}$.
- For any $\omega$ satisfying $\|\omega\|_p \leq 1$, it holds that $\langle \zeta, \omega \rangle \geq -1$, with the equality holding if and only if $\omega = \overline{\zeta}$.

We apply the above lemma to $\zeta := z_0^{-1}\overline{z}$ for a $z \in \partial K_q^{n+1} \backslash \{0\}$.

# KL property & exponent

Definition: (Attouch et al. '10, Attouch et al. '13)
Let $f$ be proper closed and $\alpha \in [0, 1)$. The function $f$ is said to have the Kurdyka-Łojasiewicz (KL) property with exponent $\alpha$ at $\bar{x} \in \operatorname{dom} \partial f$ if there exist $c$, $\nu$, $\epsilon > 0$ so that

$$c\,[f(x) - f(\bar{x})]^{\alpha} \leq \mathrm{d}(0, \partial f(x))$$

whenever $x \in \operatorname{dom} \partial f$, $\|x - \bar{x}\| \leq \epsilon$ and $f(\bar{x}) < f(x) < f(\bar{x}) + \nu$.

# KL property & exponent

Definition: (Attouch et al. '10, Attouch et al. '13)
Let $f$ be proper closed and $\alpha \in [0, 1)$. The function $f$ is said to have the Kurdyka-Łojasiewicz (KL) property with exponent $\alpha$ at $\bar{x} \in \mathrm{dom}\, \partial f$ if there exist $c$, $\nu$, $\epsilon > 0$ so that

$$c\,[f(x) - f(\bar{x})]^{\alpha} \leq \mathrm{d}(0, \partial f(x))$$

whenever $x \in \mathrm{dom}\, \partial f$, $\|x - \bar{x}\| \leq \epsilon$ and $f(\bar{x}) < f(x) < f(\bar{x}) + \nu$.

Remarks:

- Proper closed semialgebraic functions are KL functions with exponent $\alpha \in [0, 1)$. (Bolte et al. '07)

# KL property & exponent

**Definition**: (Attouch et al. '10, Attouch et al. '13)
Let $f$ be proper closed and $\alpha \in [0,1)$. The function $f$ is said to have the Kurdyka-Łojasiewicz (KL) property with exponent $\alpha$ at $\bar{x} \in \operatorname{dom}\partial f$ if there exist $c$, $\nu$, $\epsilon > 0$ so that

$$c\,[f(x) - f(\bar{x})]^{\alpha} \le \mathrm{d}(0, \partial f(x))$$

whenever $x \in \operatorname{dom}\partial f$, $\|x - \bar{x}\| \le \epsilon$ and $f(\bar{x}) < f(x) < f(\bar{x}) + \nu$.

Remarks:

- Proper closed semialgebraic functions are KL functions with exponent $\alpha \in [0,1)$. (Bolte et al. '07)
- Let $f$ be proper closed convex and $\bar{x} \in \operatorname{Arg\,min} f$. Then $f$ has KL exponent $\alpha$ at $\bar{x}$ if and only if there exists $\hat{c}$, $\hat{\nu}$, $\hat{\epsilon} > 0$ so that

$$\hat{c} \cdot \mathrm{d}(x, \operatorname{Arg\,min} f) \le (f(x) - \inf f)^{1-\alpha}$$

whenever $x \in \operatorname{dom} f$, $\|x - \bar{x}\| \le \hat{\epsilon}$ and $f(\bar{x}) < f(x) < f(\bar{x}) + \hat{\nu}$. (Bolte et al. '17)

# Infimum of KL exponent

Example: In $K_{\exp}$, recall that a $\mathbb{1}$-FRF for the unique 2-D face $\mathcal{F}_{-\infty}$ is:

$$\psi(s,t) = s + \kappa(t)\cdot \mathfrak{g}_{-\infty}(2s),$$

where $\kappa(\cdot)$ is nonnegative nondecreasing, $\mathfrak{g}_{-\infty} : \mathbb{R}_+ \to \mathbb{R}_+$ is

$$\mathfrak{g}_{-\infty}(s) := \begin{cases} 0 & \text{if } s = 0, \\ -s\ln s & \text{if } 0 < s \le e^{-2}, \\ s + e^{-2} & \text{else.} \end{cases}$$

Moreover, $\mathcal{F}_{-\infty} = K_{\exp} \cap \mathcal{F}_{-\infty}$.

# Infimum of KL exponent

Example: In $K_{\exp}$, recall that a $\mathbb{1}$-FRF for the unique 2-D face $\mathcal{F}_{-\infty}$ is:

$$\psi(s,t) = s + \kappa(t) \cdot \mathfrak{g}_{-\infty}(2s),$$

where $\kappa(\cdot)$ is nonnegative nondecreasing, $\mathfrak{g}_{-\infty} : \mathbb{R}_+ \to \mathbb{R}_+$ is

$$\mathfrak{g}_{-\infty}(s) := \begin{cases} 0 & \text{if } s = 0, \\ -s \ln s & \text{if } 0 < s \leq e^{-2}, \\ s + e^{-2} & \text{else.} \end{cases}$$

Moreover, $\mathcal{F}_{-\infty} = K_{\exp} \cap \mathcal{F}_{-\infty}$. Then there exists $\epsilon \in (0, e^{-2})$ and $c > 0$ such that

$$\mathrm{d}(x, K_{\exp} \cap \mathcal{F}_{-\infty}) \leq c \cdot \mathfrak{g}_{-\infty}(\mathrm{d}(x, K_{\exp})) \;\; \forall x \in \mathcal{F}_{-\infty} \cap B(0, \epsilon).$$

# Infimum of KL exponent

Example: In $K_{\exp}$, recall that a $\mathbb{1}$-FRF for the unique 2-D face $\mathcal{F}_{-\infty}$ is:

$$\psi(s,t) = s + \kappa(t) \cdot \mathfrak{g}_{-\infty}(2s),$$

where $\kappa(\cdot)$ is nonnegative nondecreasing, $\mathfrak{g}_{-\infty} : \mathbb{R}_+ \to \mathbb{R}_+$ is

$$\mathfrak{g}_{-\infty}(s) := \begin{cases} 0 & \text{if } s = 0, \\ -s\ln s & \text{if } 0 < s \le e^{-2}, \\ s + e^{-2} & \text{else.} \end{cases}$$

Moreover, $\mathcal{F}_{-\infty} = K_{\exp} \cap \mathcal{F}_{-\infty}$. Then there exists $\epsilon \in (0, e^{-2})$ and $c > 0$ such that

$$\mathrm{d}(x, K_{\exp} \cap \mathcal{F}_{-\infty}) \le c \cdot \mathfrak{g}_{-\infty}(\mathrm{d}(x, K_{\exp})) \ \ \forall x \in \mathcal{F}_{-\infty} \cap B(0,\epsilon).$$

Since for any $\alpha \in (0,1)$, $-s\ln s \le s^\alpha$ for all sufficiently small $s > 0$, there exists $\epsilon_1 > 0$ such that

$$\mathrm{d}(x, K_{\exp} \cap \mathcal{F}_{-\infty}) \le c \cdot (\mathrm{d}(x, K_{\exp})^2)^{\alpha/2} \ \ \forall x \in \mathcal{F}_{-\infty} \cap B(0,\epsilon_1).$$

# Infimum of KL exponent cont.

Example cont.: Thus, the function

$$f(x) := \mathrm{d}(x, K_{\exp})^2 + \delta_{\mathcal{F}_{-\infty}}(x)$$

has KL exponent $1 - \alpha/2$ at the origin, i.e., any number in $(1/2, 1)$.

# Infimum of KL exponent cont.

Example cont.: Thus, the function

$$f(x) := \mathrm{d}(x, K_{\exp})^2 + \delta_{\mathcal{F}_{-\infty}}(x)$$

has KL exponent $1 - \alpha/2$ at the origin, i.e., any number in $(1/2, 1)$.

However, there cannot be $\epsilon_2 > 0$ and $c_2 > 0$ such that

$$\mathrm{d}(x, K_{\exp} \cap \mathcal{F}_{-\infty}) \leq c_2 \cdot \mathrm{d}(x, K_{\exp}) \;\; \forall x \in \mathcal{F}_{-\infty} \cap B(0, \epsilon_2) :$$

# Infimum of KL exponent cont.

Example cont.: Thus, the function

$$f(x) := \mathrm{d}(x, K_{\exp})^2 + \delta_{\mathcal{F}_{-\infty}}(x)$$

has KL exponent $1 - \alpha/2$ at the origin, i.e., any number in $(1/2, 1)$.

However, there cannot be $\epsilon_2 > 0$ and $c_2 > 0$ such that

$$\mathrm{d}(x, K_{\exp} \cap \mathcal{F}_{-\infty}) \leq c_2 \cdot \mathrm{d}(x, K_{\exp}) \ \ \forall x \in \mathcal{F}_{-\infty} \cap B(0, \epsilon_2) :$$

Indeed, suppose the above holds. Consider

$$x^k = \left(\frac{\ln(k+1)}{k+1}, 0, 1\right) \text{ and } v^k = \left(\frac{\ln(k+1)}{k+1}, \frac{1}{k+1}, 1\right).$$

Then $v^k \in K_{\exp}$.

# Infimum of KL exponent cont.

Example cont.: Thus, the function

$$f(x) := \mathrm{d}(x, K_{\exp})^2 + \delta_{\mathcal{F}_{-\infty}}(x)$$

has KL exponent $1 - \alpha/2$ at the origin, i.e., any number in $(1/2, 1)$.

However, there cannot be $\epsilon_2 > 0$ and $c_2 > 0$ such that

$$\mathrm{d}(x, K_{\exp} \cap \mathcal{F}_{-\infty}) \le c_2 \cdot \mathrm{d}(x, K_{\exp}) \;\; \forall x \in \mathcal{F}_{-\infty} \cap B(0, \epsilon_2) :$$

Indeed, suppose the above holds. Consider

$$x^k = \left(\frac{\ln(k+1)}{k+1}, 0, 1\right) \text{ and } v^k = \left(\frac{\ln(k+1)}{k+1}, \frac{1}{k+1}, 1\right).$$

Then $v^k \in K_{\exp}$. Since $\mathcal{F}_{-\infty} := \{(x_1, 0, x_3) \mid x_1 \le 0, x_3 \ge 0\}$, we have for all large $k$ that

$$\frac{\ln(k+1)}{k+1} = \mathrm{d}(x^k, K_{\exp} \cap \mathcal{F}_{-\infty}) \le c_2 \|x^k - v^k\| = \frac{c_2}{k+1},$$

a contradiction.

# Infimum of KL exponent cont.

Example cont.: Thus, the function

$$f(x) := \mathrm{d}(x, K_{\exp})^2 + \delta_{\mathcal{F}_{-\infty}}(x)$$

has KL exponent $1 - \alpha/2$ at the origin, i.e., any number in $(1/2, 1)$.

However, there cannot be $\epsilon_2 > 0$ and $c_2 > 0$ such that

$$\mathrm{d}(x, K_{\exp} \cap \mathcal{F}_{-\infty}) \leq c_2 \cdot \mathrm{d}(x, K_{\exp}) \;\; \forall x \in \mathcal{F}_{-\infty} \cap B(0, \epsilon_2) :$$

Indeed, suppose the above holds. Consider

$$x^k = \left(\frac{\ln(k+1)}{k+1}, 0, 1\right) \text{ and } v^k = \left(\frac{\ln(k+1)}{k+1}, \frac{1}{k+1}, 1\right).$$

Then $v^k \in K_{\exp}$. Since $\mathcal{F}_{-\infty} := \{(x_1, 0, x_3) \mid x_1 \leq 0, x_3 \geq 0\}$, we have for all large $k$ that

$$\frac{\ln(k+1)}{k+1} = \mathrm{d}(x^k, K_{\exp} \cap \mathcal{F}_{-\infty}) \leq c_2 \|x^k - v^k\| = \frac{c_2}{k+1},$$

a contradiction. $\therefore$ $1/2$ is not a KL exponent at the origin!

# KL exponent of $p$-regularized problems

Consider the $p$-norm regularized problem:

$$\underset{x \in \mathbb{R}^n}{\text{Minimize}} \quad F(x) := \frac{1}{2}\|Ax - b\|^2 + \sum_{i=1}^{s} \lambda_i \|x_i\|_p,$$

where

- $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\lambda_i > 0$ for all $i$.
- The variable $x = (x_1, \ldots, x_s)$ with $x_i \in \mathbb{R}^{n_i}$, $n_i \geq 2$.

# KL exponent of $p$-regularized problems

Consider the $p$-norm regularized problem:

$$\underset{x\in\mathbb{R}^n}{\text{Minimize}} \quad F(x) := \frac{1}{2}\|Ax-b\|^2 + \sum_{i=1}^{s} \lambda_i \|x_i\|_p,$$

where

- $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $\lambda_i > 0$ for all $i$.
- The variable $x = (x_1, \ldots, x_s)$ with $x_i \in \mathbb{R}^{n_i}$, $n_i \geq 2$.
- When $p = 1$ or $\infty$, KL exponent of $F$ is $\frac{1}{2}$ (Tseng, Yun '09).
- When $p = 2$, KL exponent of $F$ is $\frac{1}{2}$ (Tseng' 10).
- When $p \in (1, 2)$, KL exponent of $F$ is $\frac{1}{2}$; when $p \in (2, \infty)$, KL exponent of $F$ cannot be $\frac{1}{2}$ (Zhou, Zhang, So '15).

# KL exponent of $p$-regularized problems

Consider the $p$-norm regularized problem:

$$\underset{x\in\mathbb{R}^n}{\text{Minimize}} \quad F(x) := \frac{1}{2}\|Ax-b\|^2 + \sum_{i=1}^{s} \lambda_i \|x_i\|_p,$$

where

- $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $\lambda_i > 0$ for all $i$.
- The variable $x = (x_1, \ldots, x_s)$ with $x_i \in \mathbb{R}^{n_i}$, $n_i \geq 2$.
- When $p = 1$ or $\infty$, KL exponent of $F$ is $\frac{1}{2}$ (Tseng, Yun '09).
- When $p = 2$, KL exponent of $F$ is $\frac{1}{2}$ (Tseng' 10).
- When $p \in (1, 2)$, KL exponent of $F$ is $\frac{1}{2}$; when $p \in (2, \infty)$, KL exponent of $F$ cannot be $\frac{1}{2}$ (Zhou, Zhang, So '15).
- The above results used the interplay between KL property, growth condition and Luo-Tseng error bound in the convex scenario. (Bolte et al. '17, Drusvyatskiy, Lewis '18)

# KL exponent of $p$-regularized problems

Consider the $p$-norm regularized problem:

$$\underset{x \in \mathbb{R}^n}{\text{Minimize}} \quad F(x) := \frac{1}{2}\|Ax - b\|^2 + \sum_{i=1}^{s} \lambda_i \|x_i\|_p,$$

where

- $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\lambda_i > 0$ for all $i$.
- The variable $x = (x_1, \ldots, x_s)$ with $x_i \in \mathbb{R}^{n_i}$, $n_i \geq 2$.
- When $p = 1$ or $\infty$, KL exponent of $F$ is $\frac{1}{2}$ (Tseng, Yun '09).
- When $p = 2$, KL exponent of $F$ is $\frac{1}{2}$ (Tseng' 10).
- When $p \in (1, 2)$, KL exponent of $F$ is $\frac{1}{2}$; when $p \in (2, \infty)$, KL exponent of $F$ cannot be $\frac{1}{2}$ (Zhou, Zhang, So '15).
- The above results used the interplay between KL property, growth condition and Luo-Tseng error bound in the convex scenario. (Bolte et al. '17, Drusvyatskiy, Lewis '18)
- Denote the optimal value by $\theta$.

# KL exponent of $p$-regularized problems cont.

A conic programming reformulation:

$$\begin{array}{ll} \underset{t,u,w,y,x}{\text{Minimize}} & 0.5t + \sum_{i=1}^{s} \lambda_i y_i \\ \text{subject to} & Ax - w = b, \quad u = 1, \\ & (t,u,w) \in \mathcal{Q}_r^{m+2}, \quad (y_i, x_i) \in K_p^{n_i+1}, \qquad i = 1, \ldots, s. \end{array}$$

where

$$\mathcal{Q}_r^{m+2} := \{(t,u,w) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^m \mid tu \geq \|w\|^2\}.$$

# KL exponent of $p$-regularized problems cont.

A conic programming reformulation:

$$\begin{array}{ll} \underset{t,u,w,y,x}{\text{Minimize}} & 0.5t + \sum_{i=1}^{s} \lambda_i y_i \\ \text{subject to} & Ax - w = b, \quad u = 1, \\ & (t, u, w) \in \mathcal{Q}_r^{m+2}, \quad (y_i, x_i) \in K_p^{n_i+1}, \qquad i = 1, \ldots, s. \end{array}$$

where

$$\mathcal{Q}_r^{m+2} := \{(t, u, w) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^m \mid tu \geq \|w\|^2\}.$$

Notation:

- $\mathbf{v} = (t, u, w, (y_1, x_1), \ldots, (y_s, x_s))$.
- $\mathcal{V} = \{\mathbf{v} \mid 0.5t + \sum_{i=1}^{s} \lambda_i y_i = \theta, u = 1, Ax - w = b\}$.
- $\mathcal{K} = \mathcal{Q}_r^{m+2} \times K_p^{n_i+1} \times \cdots \times K_p^{n_i+1}$.

# KL exponent of $p$-regularized problems cont.

A conic programming reformulation:

$$\begin{array}{ll}\underset{t,u,w,y,x}{\text{Minimize}} & 0.5t + \sum_{i=1}^{s} \lambda_i y_i \\ \text{subject to} & Ax - w = b, \quad u = 1, \\ & (t,u,w) \in \mathcal{Q}_r^{m+2}, \quad (y_i, x_i) \in K_p^{n_i+1}, \qquad i = 1,\ldots,s.\end{array}$$

where

$$\mathcal{Q}_r^{m+2} := \{(t,u,w) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^m \mid tu \geq \|w\|^2\}.$$

Notation:

- $\mathbf{v} = (t, u, w, (y_1, x_1), \ldots, (y_s, x_s))$.
- $\mathcal{V} = \{\mathbf{v} \mid 0.5t + \sum_{i=1}^{s} \lambda_i y_i = \theta, u = 1, Ax - w = b\}$.
- $\mathcal{K} = \mathcal{Q}_r^{m+2} \times K_p^{n_i+1} \times \cdots \times K_p^{n_i+1}$.
- Solution set is $\mathcal{V} \cap \mathcal{K}$.

# KL exponent of $p$-regularized problems cont.

A conic programming reformulation:

$$\begin{array}{ll} \underset{t,u,w,y,x}{\text{Minimize}} & 0.5t + \sum_{i=1}^{s} \lambda_i y_i \\ \text{subject to} & Ax - w = b, \quad u = 1, \\ & (t,u,w) \in \mathcal{Q}_r^{m+2}, \quad (y_i, x_i) \in K_p^{n_i+1}, \qquad i = 1, \ldots, s. \end{array}$$

where

$$\mathcal{Q}_r^{m+2} := \{(t,u,w) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^m \mid tu \geq \|w\|^2\}.$$

Notation:

- $\mathbf{v} = (t, u, w, (y_1, x_1), \ldots, (y_s, x_s))$.
- $\mathcal{V} = \{\mathbf{v} \mid 0.5t + \sum_{i=1}^{s} \lambda_i y_i = \theta, u = 1, Ax - w = b\}$.
- $\mathcal{K} = \mathcal{Q}_r^{m+2} \times K_p^{n_i+1} \times \cdots \times K_p^{n_i+1}$.
- Solution set is $\mathcal{V} \cap \mathcal{K}$.

Idea: Derive error bound for $\{\mathcal{V}, \mathcal{K}\}$ and invoke KL calculus rules on inf-projection (Yu, Li, Pong '22).

# KL exponent of $p$-regularized problems cont.

**Theorem 3.** (Lindstrom, Lourenço, P. '22)
Let $x^*$ be a global minimizer of $F$. Then $F$ satisfies the KL property at $x^*$ with exponent $1 - \alpha$, where

$$\alpha = \min\{0.5, 1/p\}^d \text{ for some } d \leq s + 1.$$

# KL exponent of $p$-regularized problems cont.

**Theorem 3.** (Lindstrom, Lourenço, P. '22)
Let $x^*$ be a global minimizer of $F$. Then $F$ satisfies the KL property at $x^*$ with exponent $1 - \alpha$, where

$$\alpha = \min\{0.5, 1/p\}^d \text{ for some } d \le s + 1.$$

Moreover, for the $v^* \in \mathcal{V} \cap \mathcal{K}$ that corresponds to $x^*$, if it holds that

$$v^* \in \mathrm{ri}(\mathcal{K} \cap \{s^*\}^\perp)$$

for some optimal solution $s^*$ of the dual conic program, then $d \le 1$.

# KL exponent of $p$-regularized problems cont.

**Theorem 3.** (Lindstrom, Lourenço, P. '22)
Let $x^*$ be a global minimizer of $F$. Then $F$ satisfies the KL property at $x^*$ with exponent $1 - \alpha$, where

$$\alpha = \min\{0.5, 1/p\}^d \text{ for some } d \leq s + 1.$$

Moreover, for the $v^* \in \mathcal{V} \cap \mathcal{K}$ that corresponds to $x^*$, if it holds that

$$v^* \in \mathrm{ri}(\mathcal{K} \cap \{s^*\}^\perp)$$

for some optimal solution $s^*$ of the dual conic program, then $d \leq 1$.

Key idea:

- The dominant exponent of the $\mathbb{1}$-FRF of the product cone $\mathcal{K}$ comes from the "worst case exponent" of its constituents. (Lindstrom, Lourenço, P. '22)
- $d + 1 = \ell$, the length of the chain of faces in the facial reduction process.

# Conclusion and future work

Conclusion:

- Error bounds for $\{K_{\exp}, \mathcal{L} + a\}$ and $\{K_p^{n+1}, \mathcal{L} + a\}$ using facial reduction and $\mathbb{1}$-FRFs.
- Techniques for computing $\mathbb{1}$-FRFs.
- Application to the study of KL exponents.

References:


- S. B. Lindstrom, B. F. Lourenço and T. K. Pong.
  *Error bounds, facial residual functions and applications to the exponential cone*.
  To appear in *Math. Program.*
  Available at https://arxiv.org/abs/2010.16391.
- S. B. Lindstrom, B. F. Lourenço and T. K. Pong.
  *Optimal error bounds in the absence of constraint qualifications with applications to the p-cones and beyond*.
  Preprint. Available at https://arxiv.org/abs/2109.11729.


Thanks for coming! ☺