Abstract cutting plane method for ℓ_0 minimization

Adrien Le Franc, Jean-Philippe Chancelier
Michel De Lara, Seta Rakotomandimby

Goal: minimizing ℓ_0 on $\ker A \setminus \{0\}$

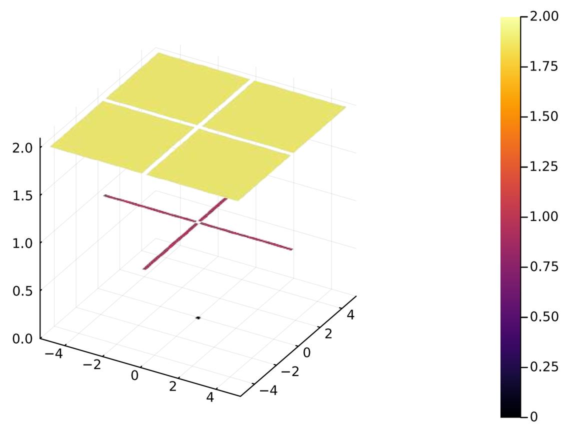
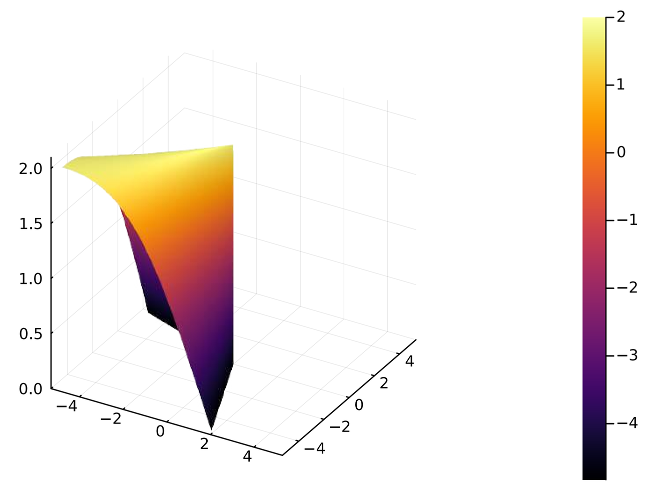


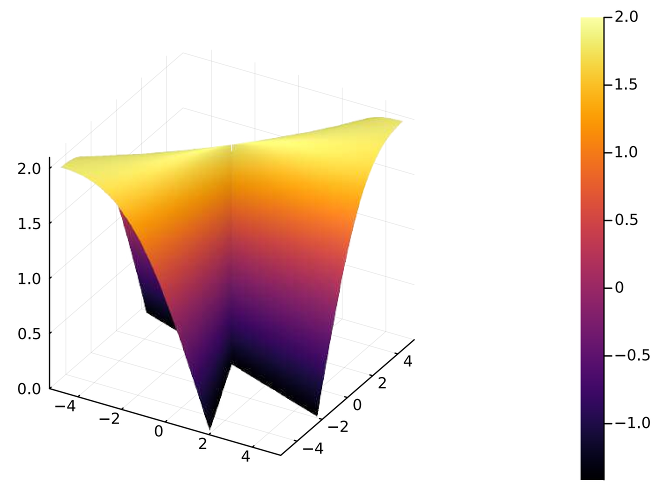
Figure: ℓ_0

$\min \{\ell_0(x) : x \in \mathbb{R}^n \setminus \{0\}, Ax = 0\}$

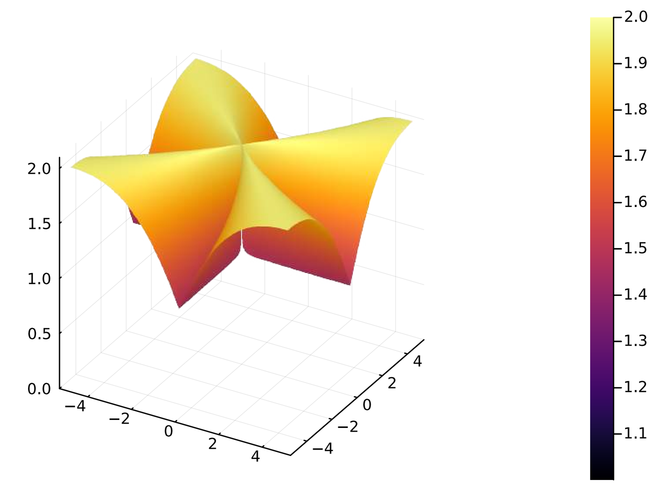
— New **Capra polyhedral approximations** of ℓ_0 using Capra cuts



(a) One diagonal



(b) Two diagonals



(c) All diagonals

— **Algorithm** and **convergence result** for the abstract cutting plane method [1, 4, 3] for a **general coupling** c

We say $\underbrace{\{x^i\}_{i\geq 0}}_{\text{primal iterates}} \subset X$, $\underbrace{\{y^i\}_{i\geq 0}}_{\text{dual iterates}} \subset \mathcal{Y}$ and $\underbrace{\{z^i\}_{i\geq 1}}_{\text{lower bounds}} \subset \mathbb{R}$ are generated by $\mathrm{CP}(X, x^0, f, c, (Y^i)_{i\geq 0}, (E^i)_{i\geq 0})$, if

1. **Initialization.**
   $x^0 \in \underbrace{X}_{\text{optimization set}} \subset \mathcal{X}$
2. ***c*-subgradient selection.**
   $y^i = Y^i(x^i)$, where $Y^i : X \to \mathcal{Y}$ s.t. $\underbrace{Y^i(x) \in \partial_c f(x)}_{c\text{-subgradient selector}}$
3. ***i*-th primal subproblem.**
   $$(x^i, z^i) \in \mathop{\arg\min}_{(x,z)\in\mathcal{X}\times\mathbb{R}} z \text{ s.t. } \begin{cases} x \in X\ ,\ (x,z) \in \overbrace{E^i \subset \mathcal{X}\times\mathbb{R}}^{\text{additional constraints}} \\ z \geq f(x^j) + c(x, y^j) - c(x^j, y^j) \\ \forall j \in [\![0, i-1]\!] \end{cases}$$
4. **Stop condition.** If not satisfied $i := i + 1$. Go to Step 2

Theorem

Let $\mathrm{CP}(X, x^0, f, c, (Y^i)_{i\geq 0}, (E^i)_{i\geq 0})$ be a cutting plane method generating $\{x^i\}_{i\geq 0} \subset X$, $\{y^i\}_{i\geq 0} \subset \mathcal{Y}$ and $\{z^i\}_{i\geq 1} \subset \mathbb{R}$

If

- *$X \subset \mathcal{X}$ is compact and $f : (\mathcal{X}, d) \to \overline{\mathbb{R}}$ is l.s.c. in X*
- *$\partial_c f(x) \neq \emptyset$, $\forall x \in X$*
- *$(\arg\min_X f) \times \{\min_X f\} \subset E^i \subset \mathcal{X} \times \mathbb{R}$, for all $i \in \mathbb{N}$*
- *there exists $M > 0$ such that*
  $|c(x, y) - c(x', y)| \leq Md(x, x')$, $\forall x, x' \in X$
  $\forall y \in \bigcup_{i\in\mathbb{N}} Y^i(X \cap \pi_{\mathcal{X}}(E^i))$

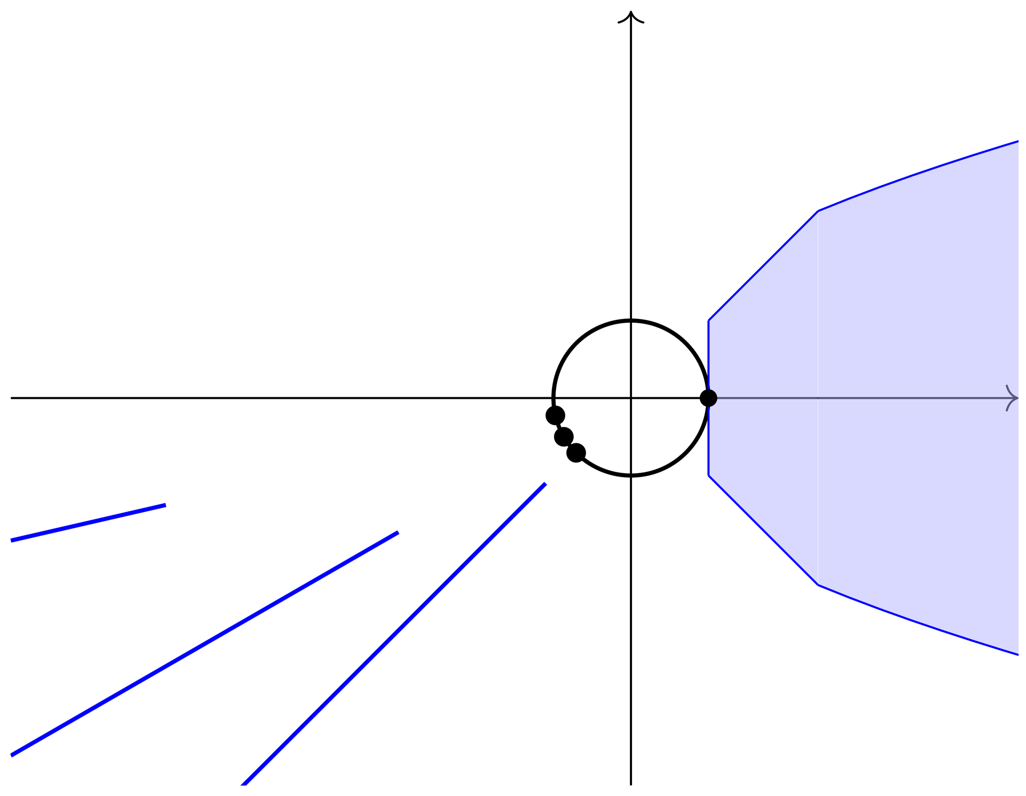
Then

- *$z^i \nearrow \min_X f$*
- *$\{x^i\}_{i\geq 0}$ has a subsequence $\{x^{\nu(i)}\}_{i\geq 0} \xrightarrow[i\to+\infty]{} x^* \in \arg\min_X f$*

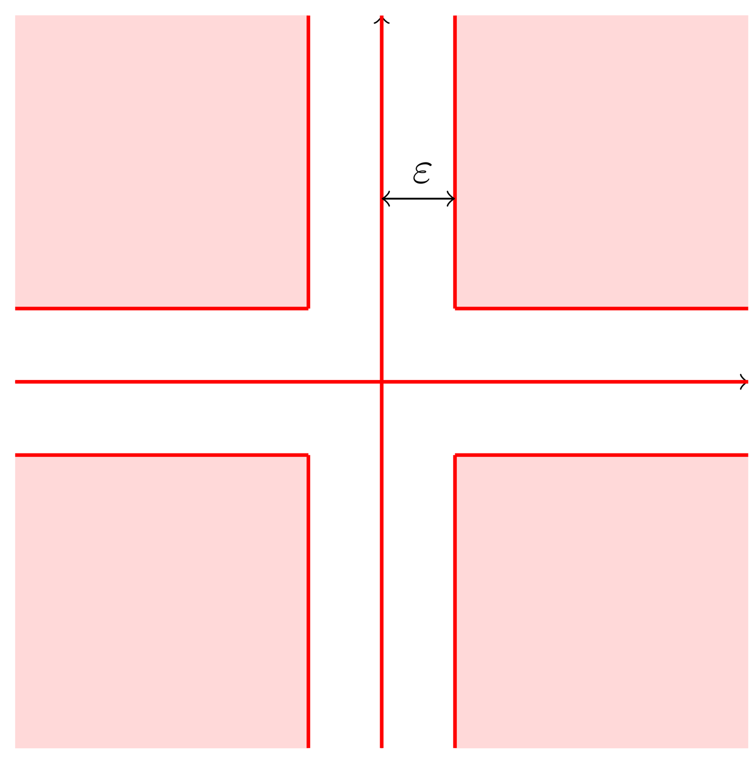
— **Diverging** ¢-subgradients of ℓ_0 near sparse point [2] and proposed solution with **sheath constraints** E

Definition

- For a source norm $\|\cdot\|$, the Capra coupling $¢ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$
  $$¢(x, y) = \left\langle \frac{x}{\|x\|} \mid y \right\rangle \ , \quad \text{where } \frac{0}{0} = 0$$
- Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function and we define its ¢-subdifferential $\partial_¢ f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by
  $$y \in \partial_¢ f(x) \iff ¢(x', y) - f(x') \leq ¢(x, y) - f(x) \quad \forall x' \in \mathbb{R}^n$$

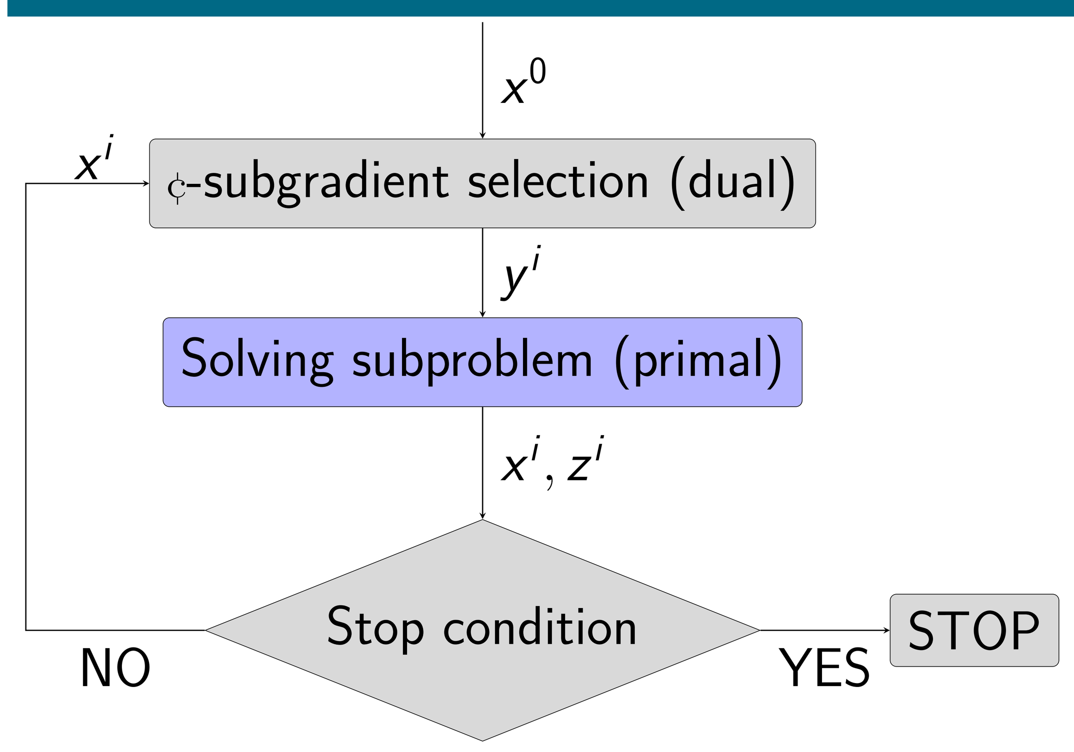


(a) Diverging ¢-subgradients for ℓ_0 near sparse points



(b) Solution: Sheath constraints E

— The **subproblem** of the Capra cutting plane method is a **linear program** on the sphere!

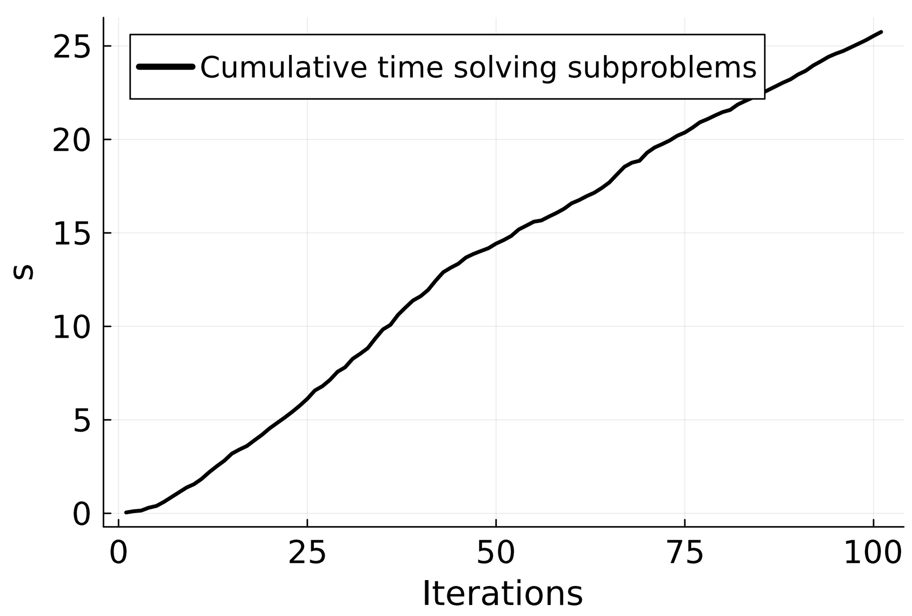


(a) Diagram of the Capra cutting plane method

When X and E are polyhedral (e.g. $X = \{x : Ax = 0\}$)

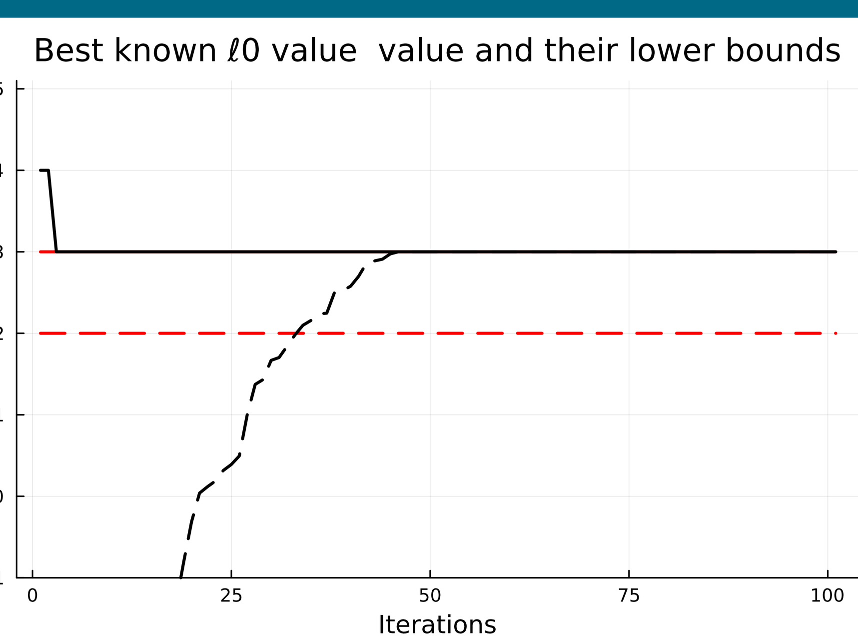
$$\min_{\substack{z\in\mathbb{R} \\ \underbrace{s\in S}_{\text{sphere constraint}}}} z \text{ s.t. } \begin{cases} \underbrace{s \in \mathrm{cone}(X)}_{\text{linear constraints}} \\ \underbrace{(s, z) \in E}_{\text{linear constraints}} \\ z \geq \underbrace{\langle s \mid y^j\rangle + f(x^j) - ¢(x^j, y^j)}_{\text{linear constraint}} \\ \forall j \in [\![0, i-1]\!] \end{cases}$$

(b) Capra cutting plane method subproblem

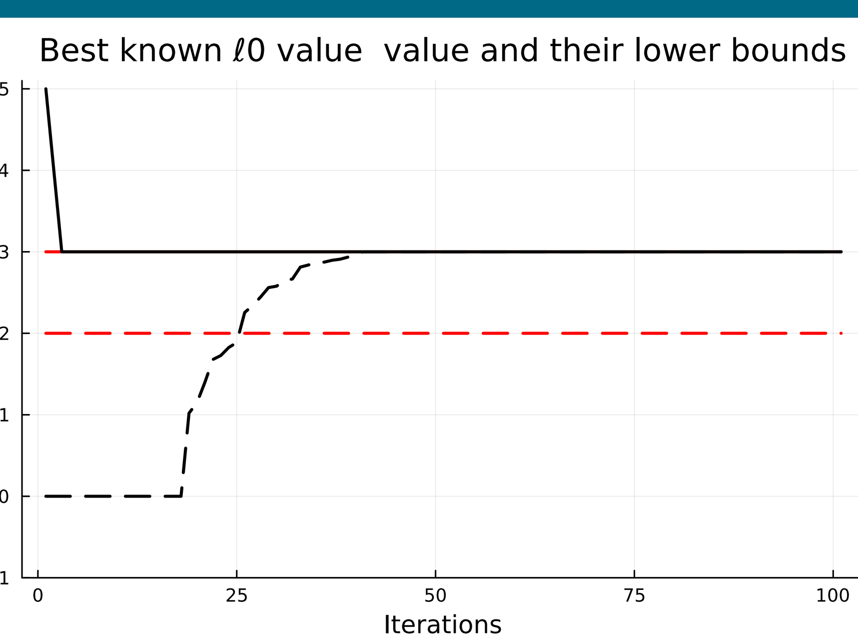


(c) Gurobi cumulative solving time on $A \in \mathbb{R}^{2\times 5}$

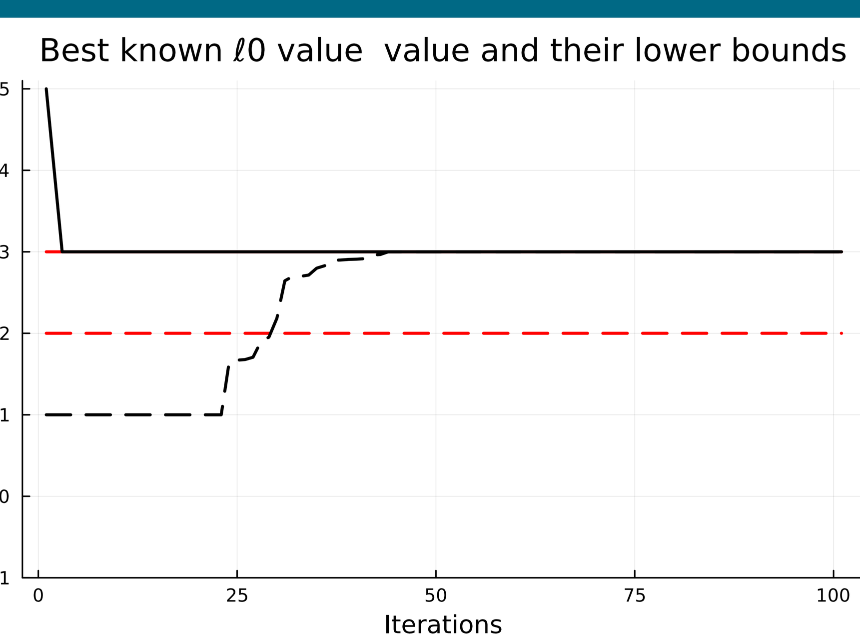
— **Numerical results** for minimizing ℓ_0 in $\ker A \setminus \{0\}$ for a gaussian matrix $A \in \mathbb{R}^{2\times 5}$ with **varying initial cut**



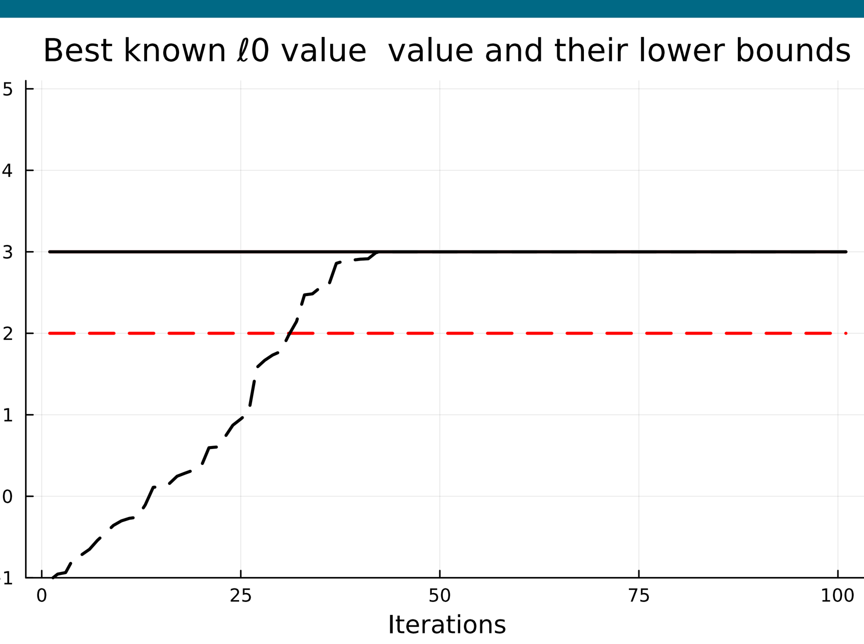
(a) No initial cut



(b) Constant value 0 initial cut



(c) Constant value 1 initial cut



(d) Random initial cut in $[-1, 1]^n$

— References

- J. E. Kelley, Jr. The cutting-plane method for solving convex programs. *Journal of the society for Industrial and Applied Mathematics*, 8(4):703–712, 1960.
- A. Le Franc, J.-P. Chancelier, and M. De Lara. The capra-subdifferential of the l_0 pseudonorm. *Optimization*, pages 1–23, 2022. accepted for publication.
- D. Pallaschke and S. Rolewicz. *Foundations of mathematical optimization*, volume 388 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1997.
- A. Rubinov. *Abstract convexity and global optimization*, volume 44 of *Nonconvex Optimization and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000.