ANSI SEMINAR - 2020

ON HIDDEN CONVEXITY IN CHANCE-CONSTRAINED PROBLEMS

Yassine LAGUEL[★] - Joint work with J. Malick[▲] and W. Van Ackooij[◆]

[★] Université Grenoble Alpes - [▲] CNRS - [◆] EDF R&D

A few words about me...

- PhD. Student from Grenoble - France
    - Member of DAO Team, from the lab. Jean Kuntzmann
    - Supervised by **J. Malick** (CNRS, Grenoble)
    - I like Hiking and playing music.
- Work on several topics related to Optimization Under Uncertainty
    - Risk-averse optimization
    - Chance constrained Optimization
    - Distributed Optimization

A few words about me...

- PhD. Student from Grenoble - France
  - Member of DAO Team, from the lab. Jean Kuntzmann
  - Supervised by **J. Malick** (CNRS, Grenoble)
  - I like Hiking and playing music.
- Work on several topics related to Optimization Under Uncertainty
  - Risk-averse optimization
  - Chance constrained Optimization **Today's topic!**
  - Distributed Optimization

Collaboration with

CNRS

J. MALICK

EDF R&D

W. VAN ACKOOIJ

Optimizing under Uncertainty via Chance Constraints

- A chance constraint is a constraint of the form:

$$\mathbb{P}[g(x, \xi) \leq 0] \geq p$$

Optimizing under Uncertainty via Chance Constraints

- A chance constraint is a constraint of the form:

Decision variable $x \in \mathbb{R}^d$

Uncertainty $\xi : \Omega \to \mathbb{R}^m$

$$\mathbb{P}[g(x, \xi) \leq 0] \geq p$$

Coupling function $g : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^k$

Safety probability level $p \in [0, 1)$

Optimizing under Uncertainty via Chance Constraints

- A chance constraint is a constraint of the form:

Decision variable $x \in \mathbb{R}^d$

Uncertainty $\xi : \Omega \to \mathbb{R}^m$

$$\mathbb{P}[g(x,\xi) \leq 0] \geq p$$

Coupling function $g : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^k$

Safety probability level $p \in [0,1)$

- Applications in various contexts
  - Energy management
  - Telecommunications
  - Chemistry

- Chance constrained problems are difficult:
  - non-convex
  - non-smooth

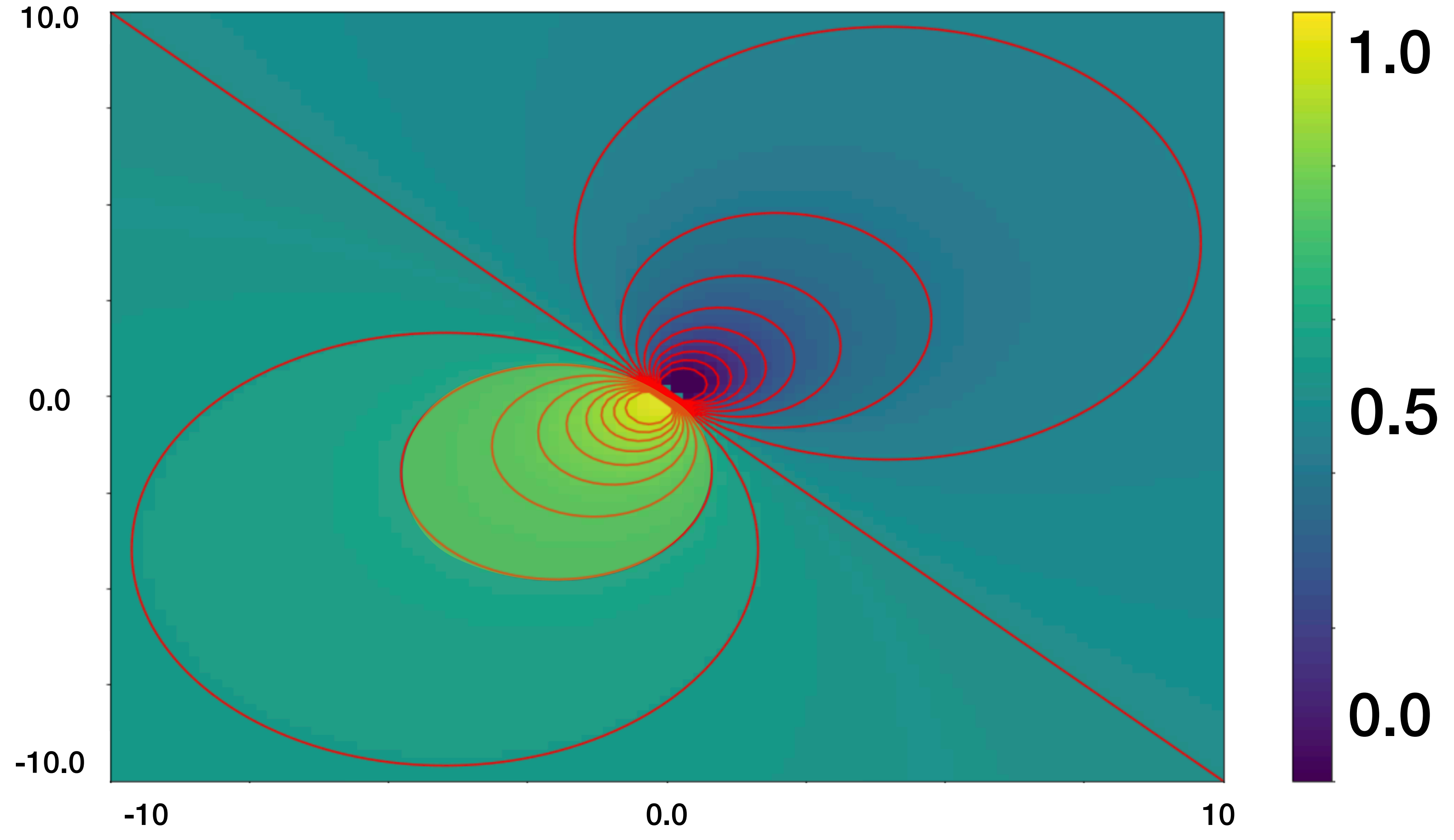
Non-convexity of Chance Constraints

- A classical result

  [Kataoka 1963]

  - Take $g : (x, \xi) \mapsto x^\top \xi$
  - Take $\xi \sim \mathcal{N}(\mu, \Sigma)$

Eventual Convexity on 2D Example

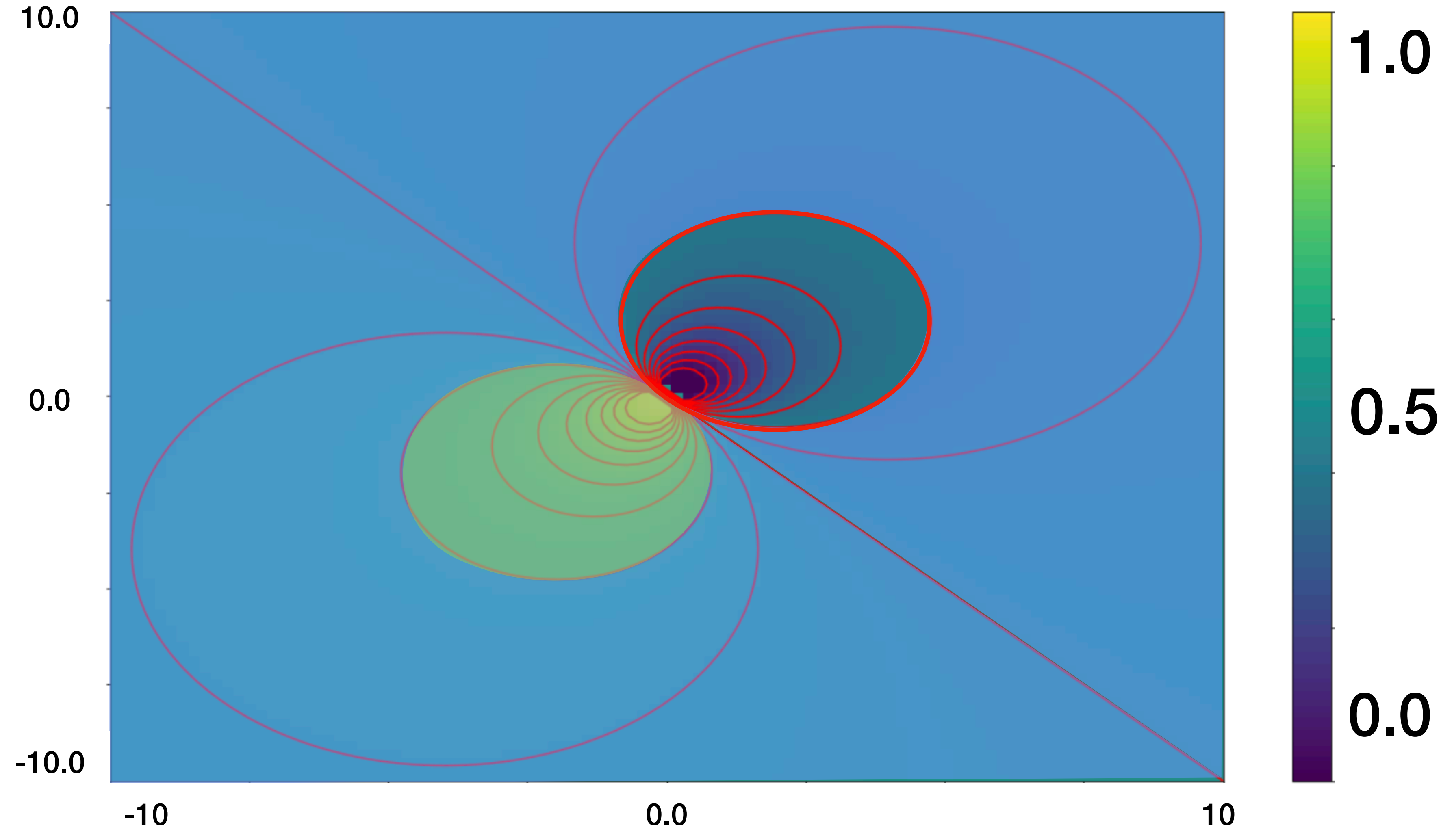


$$\mu = \mathbb{E}[\xi] = (1, 1)^\top, \mathbb{V}ar(\xi) = I_2$$

Non-convexity of Chance Constraints

- A classical result
  [Kataoka 1963]
  - Take $g : (x, \xi) \mapsto x^\top \xi$
  - Take $\xi \sim \mathcal{N}(\mu, \Sigma)$

Eventual Convexity on 2D Example

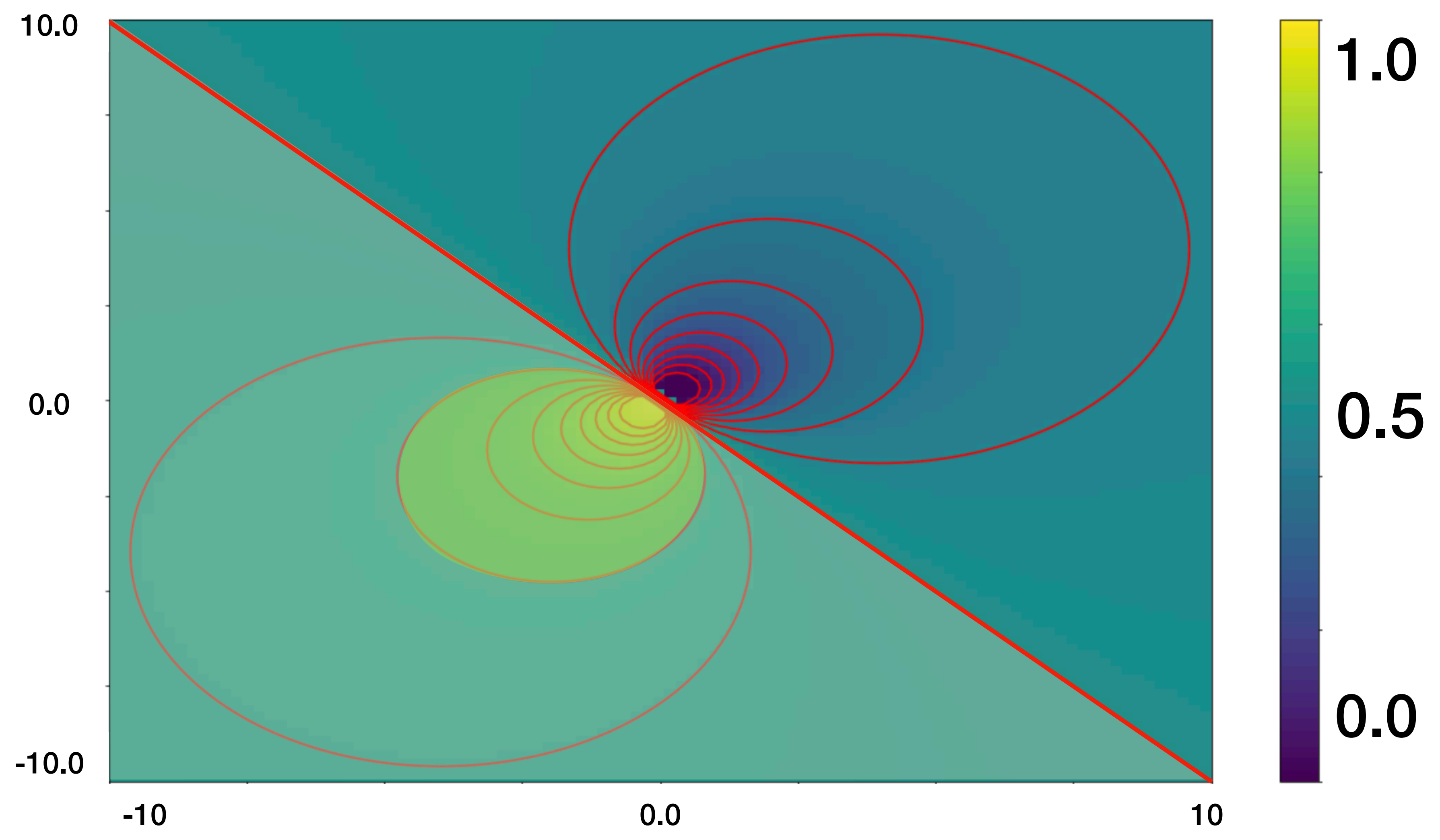


$$\mu = \mathbb{E}[\xi] = (1, 1)^\top, \mathbb{V}ar(\xi) = I_2$$

Non-convexity of Chance Constraints

- A classical result
  [Kataoka 1963]
  - Take $g : (x, \xi) \mapsto x^\top \xi$
  - Take $\xi \sim \mathcal{N}(\mu, \Sigma)$

Eventual Convexity on 2D Example



$$\mu = \mathbb{E}[\xi] = (1, 1)^\top, \mathbb{V}ar(\xi) = I_2$$

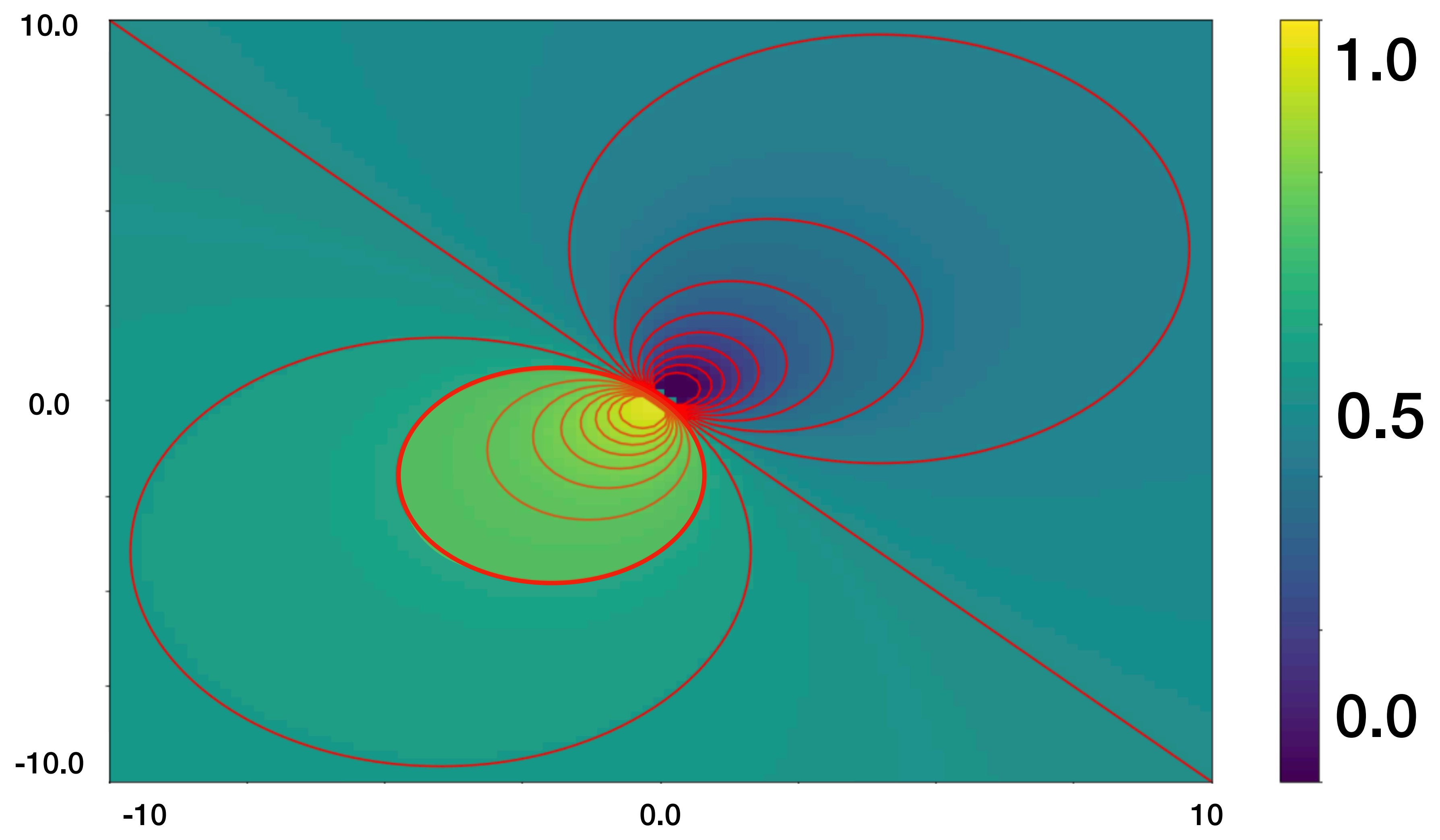
Non-convexity of Chance Constraints

- A classical result

  [Kataoka 1963]

  - Take $g : (x, \xi) \mapsto x^\top \xi$
  - Take $\xi \sim \mathcal{N}(\mu, \Sigma)$

Eventual Convexity on 2D Example



$$\mu = \mathbb{E}[\xi] = (1,1)^\top, \mathbb{V}ar(\xi) = I_2$$

Non-convexity of Chance Constraints

- A classical result

  [Kataoka 1963]

  - Take $g : (x, \xi) \mapsto x^\top \xi$
  - Take $\xi \sim \mathcal{N}(\mu, \Sigma)$

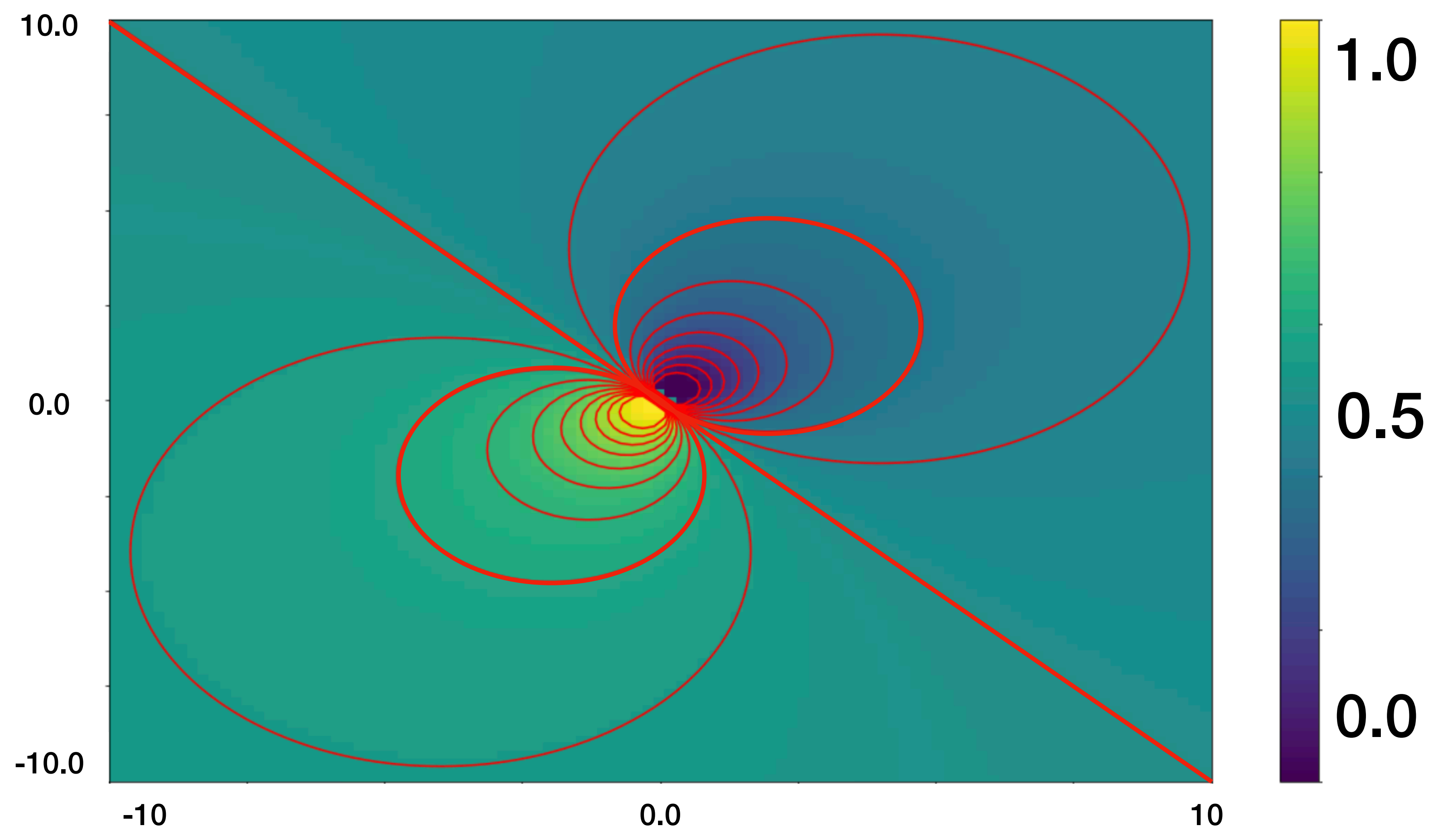
- Proof is elementary

$z \sim \mathcal{N}(0,1)$

$$\mathbb{P}[x^\top \xi \leq 0] = \mathbb{P}[x^\top \mu + \sqrt{x^\top \Sigma x}\; z \leq 0] = \phi\left(\frac{-x^\top \mu}{\sqrt{x^\top \Sigma x}}\right)$$

$$\mathbb{P}[x^\top \xi \leq 0] \geq p \Leftrightarrow \phi\left(\frac{-x^\top \mu}{\sqrt{x^\top \Sigma x}}\right) \geq p \Leftrightarrow x^\top \mu + \sqrt{x^\top \Sigma x}\; \phi^{-1}(p) \leq 0$$

Eventual Convexity on 2D Example



$$\mu = \mathbb{E}[\xi] = (1,1)^\top, \mathbb{V}ar(\xi) = I_2$$

Non-convexity of Chance Constraints

- A classical result

  [Kataoka 1963]

  - Take $g : (x, \xi) \mapsto x^\top \xi$
  - Take $\xi \sim \mathcal{N}(\mu, \Sigma)$

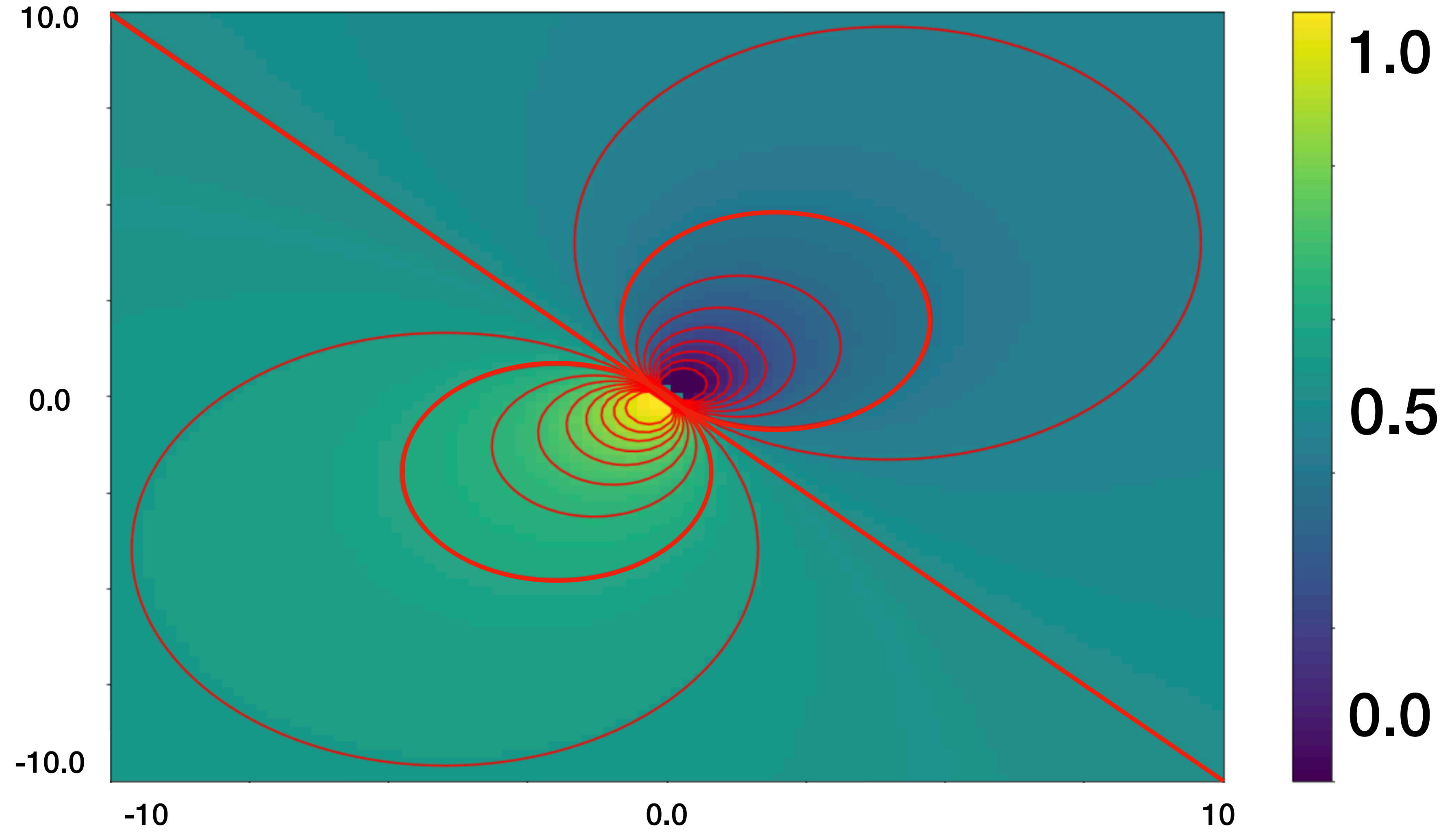
- Proof is elementary

$z \sim \mathcal{N}(0,1)$

$$\mathbb{P}[x^\top \xi \leq 0] = \mathbb{P}[x^\top \mu + \sqrt{x^\top \Sigma x}\, z \leq 0] = \phi\left(\frac{-x^\top \mu}{\sqrt{x^\top \Sigma x}}\right)$$

$$\mathbb{P}[x^\top \xi \leq 0] \geq p \Leftrightarrow \phi\left(\frac{-x^\top \mu}{\sqrt{x^\top \Sigma x}}\right) \geq p \Leftrightarrow x^\top \mu + \sqrt{x^\top \Sigma x}\, \phi^{-1}(p) \leq 0$$

Eventual Convexity on 2D Example



$$\mu = \mathbb{E}[\xi] = (1,1)^\top, \mathbb{V}ar(\xi) = I_2$$

- Studying under which conditions chance constraints are convex

Henrion, Strugarek 06' | Van Ackooij '15 | Van Ackooij, Malick '19

Non-smoothness of Chance Constraints

- Consider the discrete case : $\xi \in \{\xi_1, \dots, \xi_n\}$

$$\mathbb{P}[g(x,\xi) \leq 0] = \frac{1}{n}\sum_{i=1}^{n} \underbrace{\mathbb{1}_{g(x,\xi_i)\leq 0}}$$

Not even continuous !

Non-smoothness of Chance Constraints

- Consider the discrete case : $\xi \in \{\xi_1, \dots, \xi_n\}$

$$\mathbb{P}[g(x,\xi) \leq 0] = \frac{1}{n}\sum_{i=1}^{n} \underbrace{\mathbb{1}_{g(x,\xi_i)\leq 0}}$$

Not even continuous !

- Recent works study the generalized differentiability properties of chance constraints

Van Ackooij, Henrion, '17 **Geletu, Hoffmann, '19** **Heitsch, '19**

PLANNING

I – Eventual Convexity in Chance Constrained Programming

II – A Convex Bilevel Approach to Solve Chance Constrained Problems

I – Eventual Convexity in Chance Constrained Programming

1 From Concavity to Transconcavity

2 Leveraging Structure in a Class of Chance constraints

3 Application examples

Concavity, Quasi-concavity & Transconcavity

- Searching for (Quasi-)concavity
  - The probabilistic function:

$$\varphi : x \mapsto \varphi(x) = \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

Concavity, Quasi-concavity & Transconcavity

- Searching for (Quasi-)concavity
  - The probabilistic function:

$$\varphi : x \mapsto \varphi(x) = \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

  - $\boldsymbol{\varphi}$ is said to be quasi-concave if:

$$\varphi(\lambda x + (1-\lambda)y) \geq \min(\varphi(x), \varphi(y)) \quad \forall x, y \in C, \lambda \in [0,1]$$

Concavity, Quasi-concavity & Transconcavity

- Searching for (Quasi-)concavity
  - The probabilistic function:

$$\varphi : x \mapsto \varphi(x) = \mathbb{P}[g(x, \xi) \leq 0] \geq p$$

  - $\boldsymbol{\varphi}$ is said to be quasi-concave if:

$$\varphi(\lambda x + (1-\lambda)y) \geq \min(\varphi(x), \varphi(y)) \quad , \forall x, y \in X, \lambda \in [0,1]$$

- Transconcavity

Definition Let C be a convex subset of $\mathbb{R}^d$. We say that a function $f : C \to \mathbb{R}$ is G-concave if there exists a strictly monotonic function $G : f(C) \to \mathbb{R}$ such that

$$f(\lambda x + (1-\lambda)y) \geq G^{-1}(\lambda\, G \circ f(x) + (1-\lambda)\, G \circ f(y))$$

holds for any $x, y \in C, \lambda \in [0,1]$.

Concavity, Quasi-concavity & Transconcavity

- Searching for (Quasi-)concavity
  - The probabilistic function:

$$\varphi : x \mapsto \varphi(x) = \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

  - $\boldsymbol{\varphi}$ is said to be quasi-concave if:

$$\varphi(\lambda x + (1-\lambda)y) \geq \min(\varphi(x), \varphi(y)) \quad , \forall x, y \in X, \lambda \in [0,1]$$

- Transconcavity

Definition Let C be a convex subset of $\mathbb{R}^d$. We say that a function $f : C \to \mathbb{R}$ is G-concave if there exists a strictly monotonic function $G : f(C) \to \mathbb{R}$ such that

$$f(\lambda x + (1-\lambda)y) \geq G^{-1}(\lambda\, G \circ f(x) + (1-\lambda)\, G \circ f(y))$$

holds for any $x, y \in C, \lambda \in [0,1]$.

- The family G_α
  - A particularly well studied choice got G is the family:

$$G_\alpha : t \mapsto t^\alpha \, , \alpha \in (-\infty, 1] \setminus \{0\}$$

Concavity, Quasi-concavity & Transconcavity

- Searching for (Quasi-)concavity
  - The probabilistic function:

$$\varphi : x \mapsto \varphi(x) = \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

  - φ is said to be quasi-concave if:

$$\varphi(\lambda x + (1-\lambda)y) \geq \min(\varphi(x), \varphi(y)) \quad , \forall x, y \in X, \lambda \in [0,1]$$

- Transconcavity

Definition Let C be a convex subset of $\mathbb{R}^d$. We say that a function $f : C \to \mathbb{R}$ is G-concave if there exists a strictly monotonic function $G : f(C) \to \mathbb{R}$ such that

$$f(\lambda x + (1-\lambda)y) \geq G^{-1}\left(\lambda\, G \circ f(x) + (1-\lambda)\, G \circ f(y)\right)$$

holds for any $x, y \in C, \lambda \in [0,1]$.

- The family G_α
  - A particularly well studied choice got G is the family:

$$G_\alpha : t \mapsto t^\alpha \, , \alpha \in (-\infty, 1] \setminus \{0\}$$

  - This family may be extended by continuity. Consider $m_\alpha : \mathbb{R}_+ \times \mathbb{R}_+ \times [0,1] \to \mathbb{R}$ defined as:

If $ab = 0$ and $\alpha \leq 0$, $m_\alpha(a,b,\lambda) = 0$

$$\text{otherwise, } m_\alpha(a,b,\lambda) = \begin{cases} a^\lambda b^{1-\lambda} & \text{if } \alpha = 0 \\ \min a, b & \text{if } \alpha = -\infty \\ (\lambda a^\alpha + (1-\lambda) b^\alpha)^{\frac{1}{\alpha}} & \text{otherwise.} \end{cases}$$

Functions satisfying G_α-concavity are called α-concave.

Concavity, Quasi-concavity & Transconcavity

Searching for (Quasi-)concavity

- The probabilistic function:

$$\varphi : x \mapsto \varphi(x) = \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

- φ is said to be quasi-concave if:

$$\varphi(\lambda x + (1-\lambda)y) \geq \min(\varphi(x), \varphi(y)) \quad , \forall x, y \in X, \lambda \in [0,1]$$

Transconcavity

Definition Let C be a convex subset of $\mathbb{R}^d$. We say that a function $f : C \to \mathbb{R}$ is G-concave if there exists a strictly monotonic function $G : f(C) \to \mathbb{R}$ such that

$$f(\lambda x + (1-\lambda)y) \geq G^{-1}\left(\lambda\, G \circ f(x) + (1-\lambda)\, G \circ f(y)\right)$$

holds for any $x, y \in C, \lambda \in [0,1]$.

The family G_α

- A particularly well studied choice got G is the family:

$$G_\alpha : t \mapsto t^\alpha \, , \alpha \in (-\infty, 1] \setminus \{0\}$$

- This family may be extended by continuity. Consider $m_\alpha : \mathbb{R}_+ \times \mathbb{R}_+ \times [0,1] \to \mathbb{R}$ defined as:

If $ab = 0$ and $\alpha \leq 0$, $m_\alpha(a,b,\lambda) = 0$

$$\text{otherwise, } m_\alpha(a,b,\lambda) = \begin{cases} a^\lambda b^{1-\lambda} & \text{if } \alpha = 0 \\ \min a, b & \text{if } \alpha = -\infty \\ (\lambda a^\alpha + (1-\lambda) b^\alpha)^{\frac{1}{\alpha}} & \text{otherwise.} \end{cases}$$

Functions satisfying G_α-concavity are called α-concave.

- For any $f : C \to \mathbb{R}_+$ and $\alpha, \beta \in [-\infty, 1)$ be given, if f is α-concave, it is also β-concave when $\alpha \geq \beta$. In particular, f is quasi-concave.

Transporting Concavity

- Propagation of generalized concavity
  - The propagation lemma

For $f : C \to \mathbb{R}$:

$$\begin{cases} f & \text{is } G_1\text{-concave} \\ G^{-1} & \text{is } G_2\text{-concave} \end{cases} \implies f \text{ is } G_2\text{-concave}$$

Transporting Concavity

- Propagation of generalized concavity
  - The propagation lemma

For $f : C \to \mathbb{R}$:

$$\begin{cases} f & \text{is } G_1\text{-concave} \\ G^{-1} & \text{is } G_2\text{-concave} \end{cases} \implies f \text{ is } G_2\text{-concave}$$

- Inverse transconcavity
  - Concave-G^{-1}functions

Definition Let $F : \mathbb{R} \to \mathbb{R}$ be a non-decreasing function. Let $G : \mathbb{R} \to \mathbb{R}$ be a continuous and strictly monotonic function. We say that F is concave-G^{-1} if $F \circ G^{-1}$ is concave.

Transporting Concavity

- Propagation of generalized concavity
  - The propagation lemma

For $f : C \to \mathbb{R}$:

$$\begin{cases} f & \text{is } G_1\text{-concave} \\ G^{-1} & \text{is } G_2\text{-concave} \end{cases} \implies f \text{ is } G_2\text{-concave}$$

- Inverse transconcavity
  - Concave-G^{-1}functions

Definition Let $F : \mathbb{R} \to \mathbb{R}$ be a non-decreasing function. Let $G : \mathbb{R} \to \mathbb{R}$ be a continuous and strictly monotonic function. We say that F is concave-G^{-1} if $F \circ G^{-1}$ is concave.

- The G-decreasing property

A mapping f is said to be G-decreasing if there exists $G : \mathbb{R} \to \mathbb{R}$, $\mathcal{C}^1$ and strictly monotonic and $t_G^\star$ such that:

$$\begin{cases} G'(t) \neq 0 \,, \forall t \geq t_G^\star \\ r(t) := \frac{f(t)}{G'(t)} \text{ is } \left| \begin{array}{l} \text{decreasing if } G \text{ is increasing} \\ \text{increasing if } G \text{ is decreasing} \end{array} \right. \end{cases}$$

Transporting Concavity

- Propagation of generalized concavity
  - The propagation lemma

For $f : C \to \mathbb{R}$:

$$\begin{cases} f & \text{is } G_1\text{-concave} \\ G^{-1} & \text{is } G_2\text{-concave} \end{cases} \implies f \text{ is } G_2\text{-concave}$$

- Inverse transconcavity
  - Concave-G^{-1}functions

Definition Let $F : \mathbb{R} \to \mathbb{R}$ be a non-decreasing function. Let $G : \mathbb{R} \to \mathbb{R}$ be a continuous and strictly monotonic function. We say that F is concave-G^{-1} if $F \circ G^{-1}$ is concave.

- The G-decreasing property

A mapping f is said to be G-decreasing if there exists $G : \mathbb{R} \to \mathbb{R}$, $\mathcal{C}^1$ and strictly monotonic and $t^\star_G$ such that:

$$\begin{cases} G'(t) \neq 0 \,, \forall t \geq t^\star_G \\ r(t) := \frac{f(t)}{G'(t)} \text{ is } \left| \begin{array}{l} \text{decreasing if } G \text{ is increasing} \\ \text{increasing if } G \text{ is decreasing} \end{array} \right. \end{cases}$$

- Equivalence with inverse transconcavity

For a c.d.f. F with associated continuously differentiable density function f, we have the equivalence between:

$$f \text{ is } G\text{-decreasing} \iff F \text{ is concave-}G^{-1}$$

Transporting Concavity

Propagation of generalized concavity

The propagation lemma

For $f : C \to \mathbb{R}$:

$$\begin{cases} f & \text{is } G_1\text{-concave} \\ G^{-1} & \text{is } G_2\text{-concave} \end{cases} \implies f \text{ is } G_2\text{-concave}$$

Inverse transconcavity

Concave-G^{-1} functions

Definition Let $F : \mathbb{R} \to \mathbb{R}$ be a non-decreasing function. Let $G : \mathbb{R} \to \mathbb{R}$ be a continuous and strictly monotonic function. We say that F is concave-G^{-1} if $F \circ G^{-1}$ is concave.

The G-decreasing property

A mapping f is said to be G-decreasing if there exists $G : \mathbb{R} \to \mathbb{R}$, $\mathcal{C}^1$ and strictly monotonic and $t_G^\star$ such that:

$$\begin{cases} G'(t) \neq 0 \,, \forall t \geq t_G^\star \\ r(t) := \frac{f(t)}{G'(t)} \text{ is } \left| \begin{array}{l} \text{decreasing if } G \text{ is increasing} \\ \text{increasing if } G \text{ is decreasing} \end{array} \right. \end{cases}$$

Equivalence with inverse transconcavity

For a c.d.f. F with associated continuously differentiable density function f, we have the equivalence between:

$$f \text{ is } G\text{-decreasing} \iff F \text{ is concave-}G^{-1}$$

Application with the G_α family.

A good number of distributions have proven to be concave-G-1 including:

Gaussian	Chi-squared	Exponential	Log-Normal
Chi	Fisher-Snedecor	Gamma	Maxwell

Leveraging Structure in a class of Chance constraints

- Separable probabilistic constraints
  - Consider a separable constraint of the form:

$$\mathbb{P}[\xi \leq h(x)] \geq p$$

$\xi : \Omega \to \mathbb{R}^m$ $h : \mathbb{R}^d \to \mathbb{R}^m$

Leveraging Structure in a class of Chance constraints

- Separable probabilistic constraints
  - Consider a separable constraint of the form:

$$\mathbb{P}[\xi \leq h(x)] \geq p$$

$\xi : \Omega \to \mathbb{R}^m$ $\quad$ $h : \mathbb{R}^d \to \mathbb{R}^m$

  - By Sklar's theorem, we may write:

$$\mathbb{P}[\xi \leq h(x)] = \mathcal{C}\left(F_1(h_1(x)), \ldots, F_m(h_m(x))\right)$$

Leveraging Structure in a class of Chance constraints

- Separable probabilistic constraints
  - Consider a separable constraint of the form:

$$\mathbb{P}[\xi \leq h(x)] \geq p$$

$\xi : \Omega \to \mathbb{R}^m$ $\quad$ $h : \mathbb{R}^d \to \mathbb{R}^m$

  - By Sklar's theorem, we may write:

$$\mathbb{P}[\xi \leq h(x)] = \mathcal{C}\left(F_1(h_1(x)), \ldots, F_m(h_m(x))\right)$$

  - Concave-G^{-1} copulae

For $\mathcal{C}$ a given copulae and G such that the G_i are all continuous and strictly monotonic, $\mathcal{C}$ is said to be concave-G^{-1} on $I := \Pi_{i=1}^m I_i$ if:

$$z \mapsto \mathcal{C}(G_1^{-1}(z_1), \ldots, G_m^{-1}(z_m))$$

is quasi-concave.

Leveraging Structure in a class of Chance constraints

- Separable probabilistic constraints
  - Consider a separable constraint of the form:

$$\mathbb{P}[\xi \leq h(x)] \geq p$$

$\xi : \Omega \to \mathbb{R}^m$ $h : \mathbb{R}^d \to \mathbb{R}^m$

  - By Sklar's theorem, we may write:

$$\mathbb{P}[\xi \leq h(x)] = \mathcal{C}\left(F_1(h_1(x)), \ldots, F_m(h_m(x))\right)$$

  - Concave-G^{-1} copulae

For $\mathcal{C}$ a given copulae and G such that the G_i are all continuous and strictly monotonic, $\mathcal{C}$ is said to be concave-G^{-1} on $I := \Pi_{i=1}^m I_i$ if:

$$z \mapsto \mathcal{C}(G_1^{-1}(z_1), \ldots, G_m^{-1}(z_m))$$

is quasi-concave.

- The convexity result
  - Consider a constraint of the form:

$$\mathbb{P}[\xi \leq h(x)] = \mathcal{C}\left(F_1(h_1(x)), \ldots, F_m(h_m(x))\right)$$

Leveraging Structure in a class of Chance constraints

- Separable probabilistic constraints
  - Consider a separable constraint of the form:

$$\mathbb{P}[\xi \leq h(x)] \geq p$$

$\xi : \Omega \to \mathbb{R}^m$ $h : \mathbb{R}^d \to \mathbb{R}^m$

  - By Sklar's theorem, we may write:

$$\mathbb{P}[\xi \leq h(x)] = \mathcal{C}\left(F_1(h_1(x)), \ldots, F_m(h_m(x))\right)$$

  - Concave-G^{-1} copulae

For $\mathcal{C}$ a given copulae and G such that the G_i are all continuous and strictly monotonic, $\mathcal{C}$ is said to be concave-G^{-1} on $I := \Pi_{i=1}^m I_i$ if:

$$z \mapsto \mathcal{C}(G_1^{-1}(z_1), \ldots, G_m^{-1}(z_m))$$

is quasi-concave.

- The convexity result
  - Consider a constraint of the form:

$$\mathbb{P}[\xi \leq h(x)] = \mathcal{C}\left(F_1(h_1(x)), \ldots, F_m(h_m(x))\right)$$

  - Assume there exists strictly monotonous mappings $G_1, \ldots, G_m$ and $\hat{G}^1, \ldots, \hat{G}^m$ such that:

- The components h_i are G_i-concave on $\{x, h_i(x) \geq b_i\}$
- The c.d.f. F_i are concave-G_i^{-1} and $\hat{G}_i$-concave on $[G_i(b_i), \infty)$
- The copulae $\mathcal{C}$ is concave-$\hat{G}^{-1}$on the product of the intervals $(-\infty, \hat{G}_i(b_i)]$.

Leveraging Structure in a class of Chance constraints

- Separable probabilistic constraints
  - Consider a separable constraint of the form:

$$\mathbb{P}[\xi \leq h(x)] \geq p$$

$\xi : \Omega \to \mathbb{R}^m$ $h : \mathbb{R}^d \to \mathbb{R}^m$

  - By Sklar's theorem, we may write:

$$\mathbb{P}[\xi \leq h(x)] = \mathcal{C}\left(F_1(h_1(x)), \ldots, F_m(h_m(x))\right)$$

  - Concave-G^{-1} copulae

For $\mathcal{C}$ a given copulae and G such that the G_i are all continuous and strictly monotonic, $\mathcal{C}$ is said to be concave-G^{-1} on $I := \Pi_{i=1}^m I_i$ if:

$$z \mapsto \mathcal{C}(G_1^{-1}(z_1), \ldots, G_m^{-1}(z_m))$$

is quasi-concave.

- The convexity result
  - Consider a constraint of the form:

$$\mathbb{P}[\xi \leq h(x)] = \mathcal{C}\left(F_1(h_1(x)), \ldots, F_m(h_m(x))\right)$$

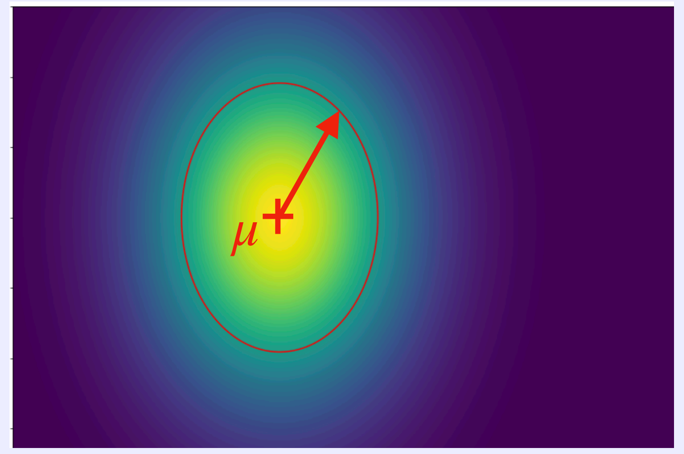
  - Assume there exists strictly monotonous mappings $G_1, \ldots, G_m$ and $\hat{G}^1, \ldots, \hat{G}^m$ such that:

- The components h_i are G_i-concave on $\{x, h_i(x) \geq b_i\}$
- The c.d.f. F_i are concave-G_i^{-1} and $\hat{G}_i$-concave on $[G_i(b_i), \infty)$
- The copulae $\mathcal{C}$ is concave-$\hat{G}^{-1}$on the product of the intervals $(-\infty, \hat{G}_i(b_i)]$.

  - Then, the set $M_p := \{x, \mathbb{P}[\xi \leq h(x)] \geq p\}$ is convex for all $p \geq p^\star := \max_{1 \leq i \leq m} F_i(b_i)$

Leveraging Structure in a class of Chance constraints

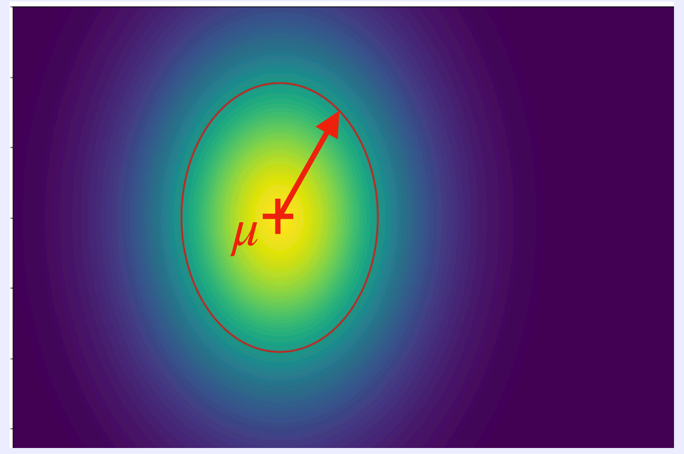
- Non-linear couplings with elliptical distributions
  - The Gaussian case



- $g : C \times \mathbb{R}^m \to \mathbb{R}^k$ is continuous
- $g(., z)$ is convex for all $z \in \mathbb{R}^m$
- $g(x, .)$ is convex for all $x \in \mathbb{R}^d$
- ξ is a gaussian: $\xi = \mu + \mathcal{R} L \zeta$
- $g(x, \mu) \leq 0$

Leveraging Structure in a class of Chance constraints

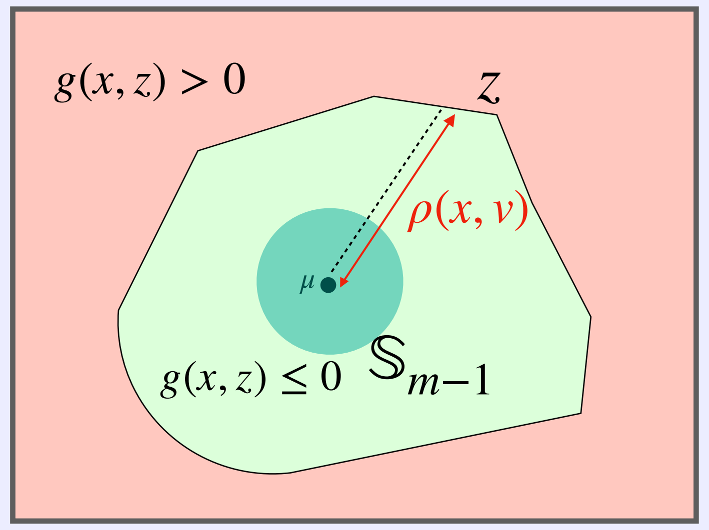
- Non-linear couplings with elliptical distributions
  - The Gaussian case



- $g : C \times \mathbb{R}^m \to \mathbb{R}^k$ is continuous
- $g(., z)$ is convex for all $z \in \mathbb{R}^m$
- $g(x, .)$ is convex for all $x \in \mathbb{R}^d$
- ξ is a gaussian: $\xi = \mu + \mathcal{R}L\zeta$
- $g(x, \mu) \leq 0$

  - The spherical integration formula

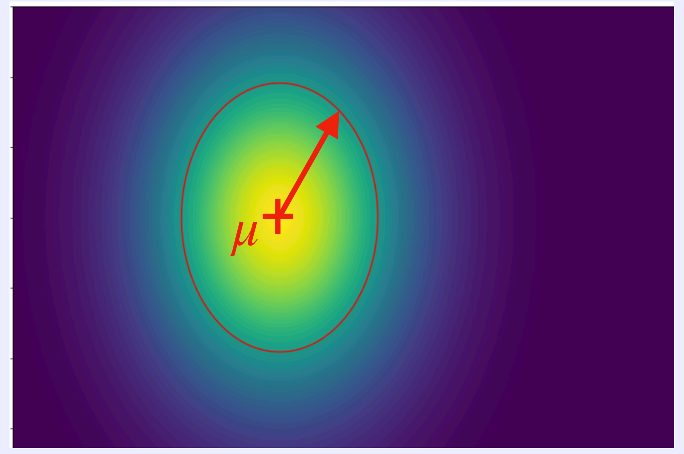
We define: $\rho(x, v) = \sup_{t>0} \{g(x, \mu + tLv) \leq 0\}$



$$\mathbb{P}[g(x, \xi) \leq 0] = \int_{v \in \mathbb{S}_{m-1}} F_{\mathcal{R}}(\rho(x, v)) dv$$

Leveraging Structure in a class of Chance constraints

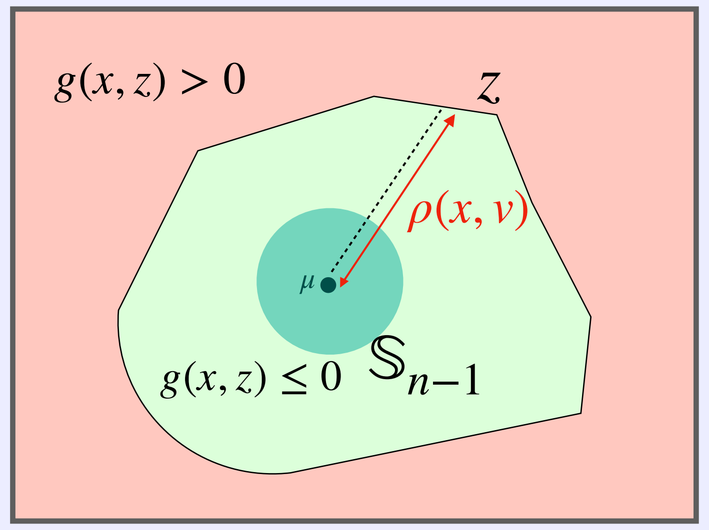
- Non-linear couplings with elliptical distributions
  - The Gaussian case



- $g : C \times \mathbb{R}^m \to \mathbb{R}^k$ is continuous
- $g(., z)$ is convex for all $z \in \mathbb{R}^m$
- $g(x, .)$ is convex for all $x \in \mathbb{R}^d$
- ξ is a gaussian: $\xi = \mu + \mathcal{R}L\zeta$
- $g(x, \mu) \leq 0$

  - The spherical integration formula

We define: $\rho(x, v) = \sup_{t>0} \{g(x, \mu + tLv) \leq 0\}$

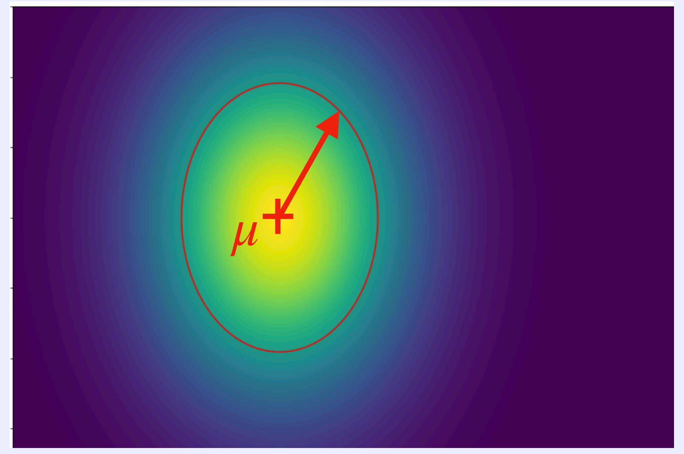


$$\mathbb{P}[g(x, \xi) \leq 0] = \int_{v \in \mathbb{S}_{m-1}} F_{\mathcal{R}}(\rho(x, v)) dv$$

- The convexity result
  - Assume that for all $v \in \mathbb{S}_{m-1}$
    - $\rho(., v)$ is continuous
    - There exists $G_v : \mathbb{R}_+ \to \mathbb{R}$ such that:
      - G_v is strictly monotonic.
      - $\rho(., v)$ is G_v-concave
      - $F_{\mathcal{R}}$ is concave-G_v^{-1}.

Leveraging Structure in a class of Chance constraints

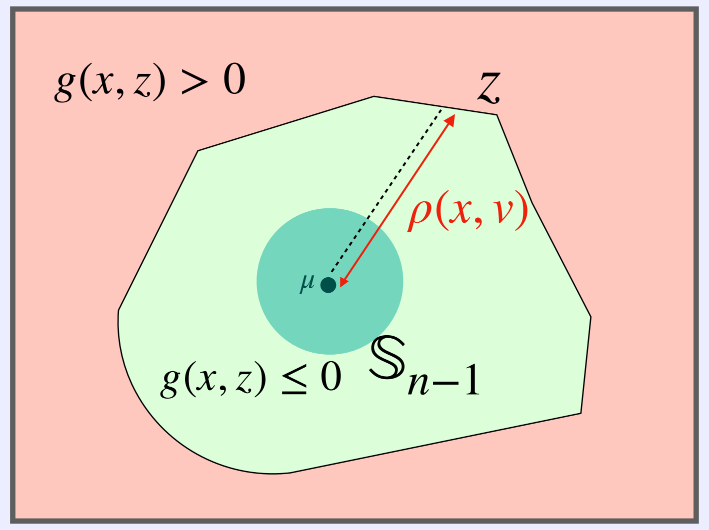
- Non-linear couplings with elliptical distributions
  - The Gaussian case



- $g : C \times \mathbb{R}^m \to \mathbb{R}^k$ is continuous
- $g(., z)$ is convex for all $z \in \mathbb{R}^m$
- $g(x, .)$ is convex for all $x \in \mathbb{R}^d$
- ξ is a gaussian: $\xi = \mu + \mathcal{R}L\zeta$
- $g(x, \mu) \leq 0$

  - The spherical integration formula

We define: $\rho(x, v) = \sup_{t>0} \{g(x, \mu + tLv) \leq 0\}$



$$\mathbb{P}[g(x, \xi) \leq 0] = \int_{v \in \mathbb{S}_{m-1}} F_{\mathcal{R}}(\rho(x, v)dv$$

- The convexity result
  - Assume that for all $v \in \mathbb{S}_{m-1}$
    - $\rho(., v)$ is continuous
    - There exists $G_v : \mathbb{R}_+ \to \mathbb{R}$ such that:
      - G_v is strictly monotonic.
      - $\rho(., v)$ is G_v-concave
      - $F_{\mathcal{R}}$ is concave-G_v^{-1}.
    - There exists $t^\star$ such that: $\{x \in \mathbb{R}^d : \rho(x, v) \geq t^\star\} \subseteq C$
  - Then, the set $M_p = \{x, \mathbb{P}[g(x, \xi) \leq 0] \geq p\}$

is convex for all $p \geq p^\star = \max\left(\frac{1}{2}, p'\right)$ with

$$p' = \inf q \in [0, \frac{1}{2}] \left(\frac{1}{2} - q\right) F_{\mathcal{R}}\left(\frac{t^\star}{\delta(q)}\right) + \frac{1}{2} + q$$

and $\delta(q)$ the unique solution of:

$$\mathcal{B}_i\left(\frac{m-1}{2}, \frac{1}{2}, \sin^2(\arccos(\delta))\right) = (1-2q)\mathcal{B}_c\left(\frac{m-1}{2}, \frac{1}{2}\right)$$

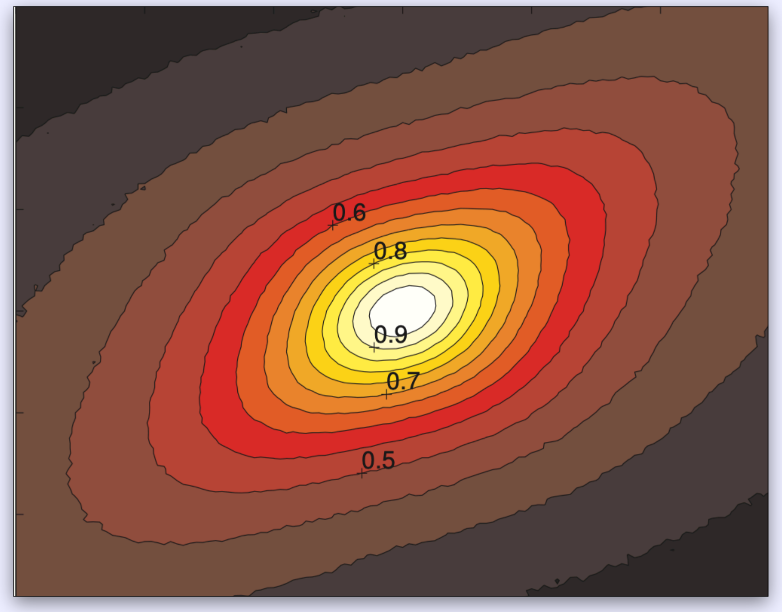
Application Examples

- The linear case
  - Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

$\in (\mathbb{R}_+^*)^m$

$m \times d$ centered multi-variate gaussian random matrix



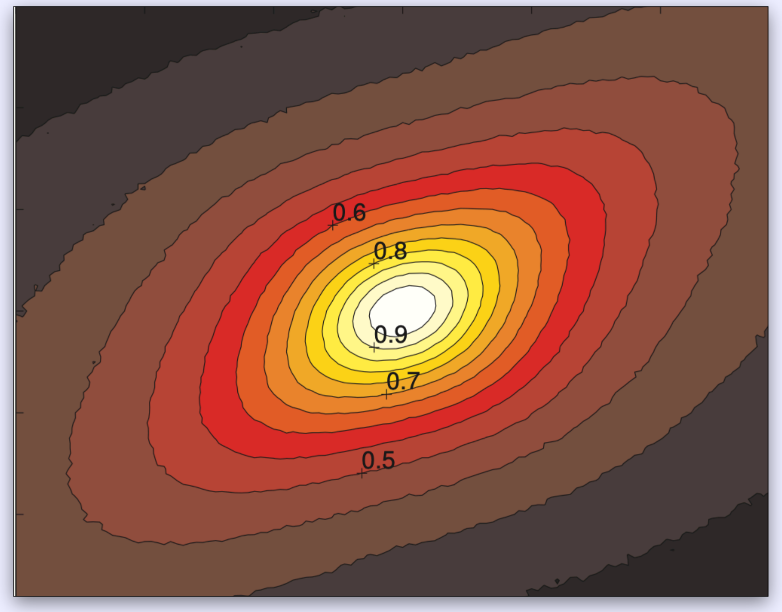
Application Examples

- The linear case
  - Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

$\beta \in (\mathbb{R}_+^*)^m$

Ξ: $m \times d$ centered multi-variate gaussian random matrix



Then $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq p^\star := \Phi(\sqrt{mn+1})$

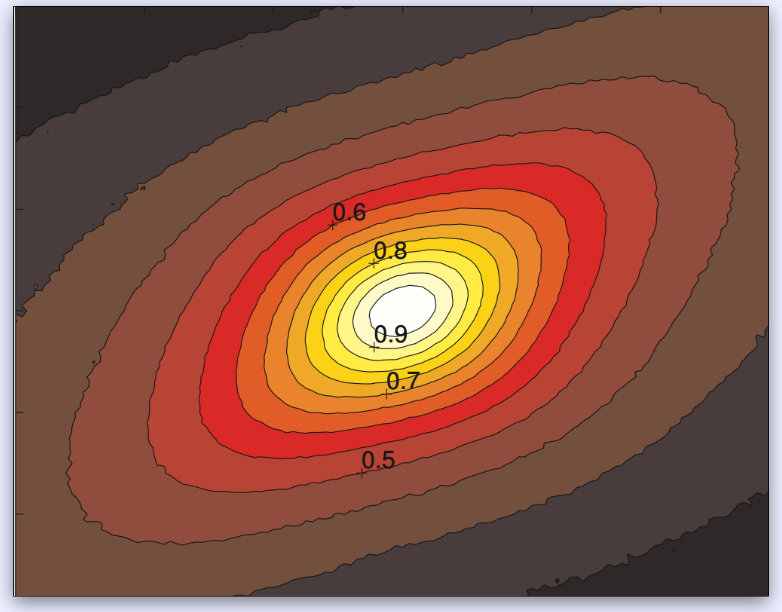
Application Examples

- The linear case
  - Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

Ξ: $m \times d$ centered multi-variate gaussian random matrix

$\beta \in (\mathbb{R}_+^*)^m$



Then $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq p^\star := \Phi(\sqrt{mn+1})$

- Proof leverages α-concavity $(\alpha = -1)$ of the radial function ρ and and the concave-α property of 1D-gaussian random variables.

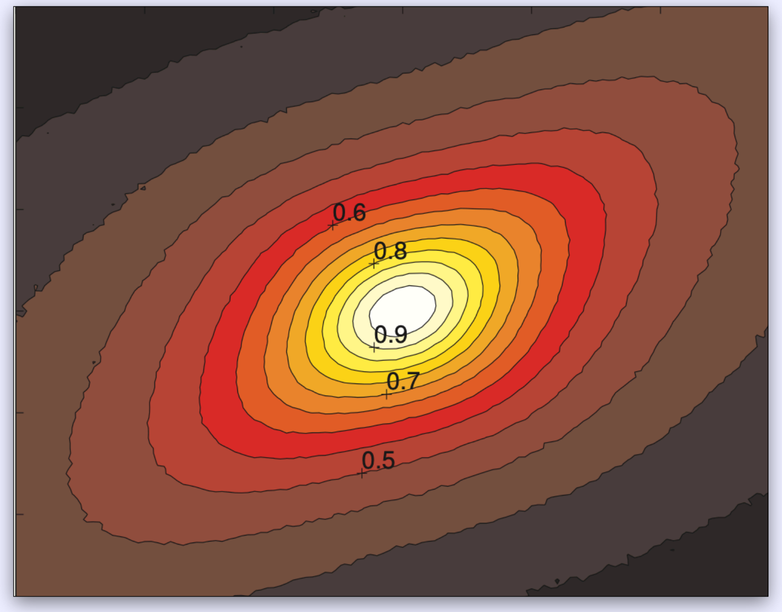
Application Examples

- The linear case
  - Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

Ξ: $m \times d$ centered multi-variate gaussian random matrix

$\beta \in (\mathbb{R}_+^*)^m$



Then $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq p^\star := \Phi(\sqrt{mn+1})$

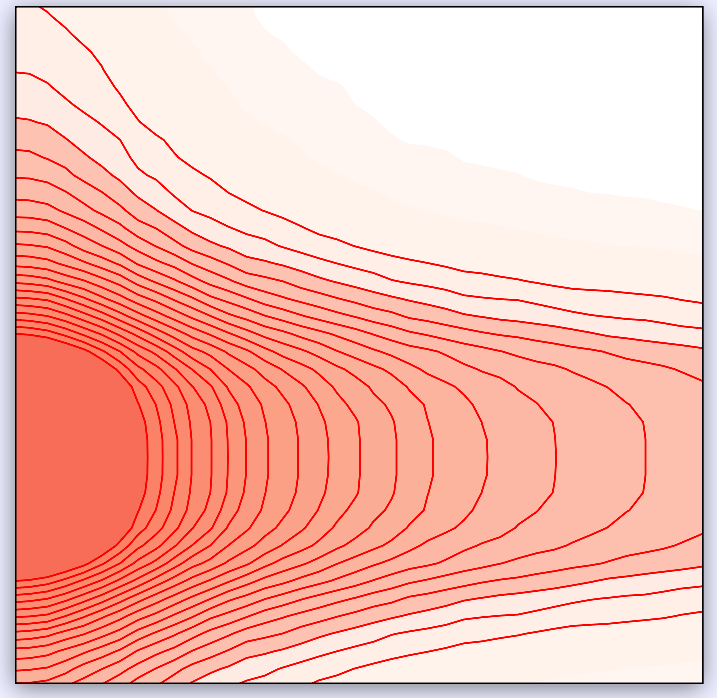
  - Proof leverages α-concavity $(\alpha = -1)$ of the radial function ρ and and the concave-α property of 1D-gaussian random variables.

- The Quadratic case
  - Consider:

$$g(x,z) = z^\top W(x) z + 2\sum_{i=1}^{n} a_i w_i^\top z_i + b$$

$W(x)$: Symmetric definite positive with convex eigenvalues in x

$b \leq 0$



and ξ a multivariate Gaussian random vector.

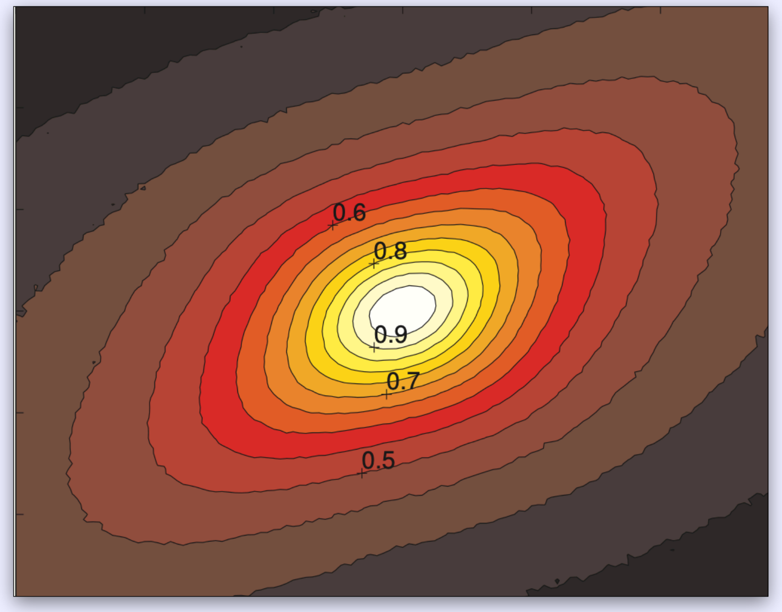
Application Examples

- The linear case
  - Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

Ξ: $m \times d$ centered multi-variate gaussian random matrix

$\beta \in (\mathbb{R}_+^*)^m$



Then $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq p^\star := \Phi(\sqrt{mn+1})$

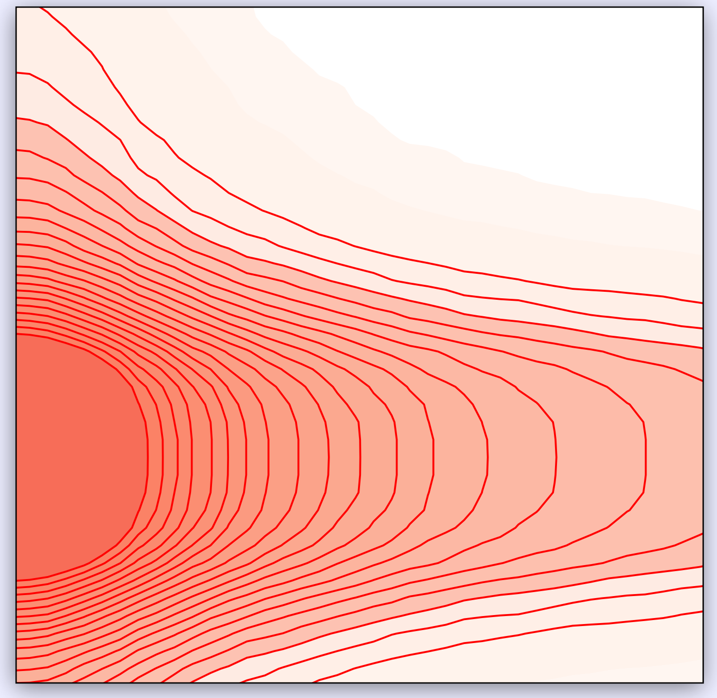
  - Proof leverages α-concavity $(\alpha = -1)$ of the radial function ρ and and the concave-α property of 1D-gaussian random variables.

- The Quadratic case
  - Consider:

$$g(x,z) = z^\top W(x) z + 2\sum_{i=1}^{n} a_i w_i^\top z_i + b$$

$W(x)$: Symmetric definite positive with convex eigenvalues in x

b: ≤ 0



and ξ a multivariate Gaussian random vector.

Then, $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq \Phi(\sqrt{m+3})$

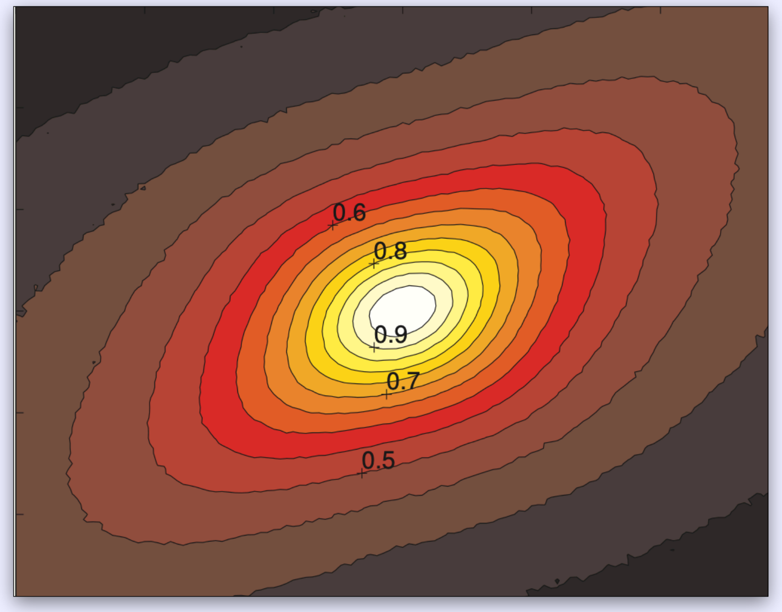
Application Examples

- The linear case
  - Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

$m \times d$ centered multi-variate gaussian random matrix

$\in (\mathbb{R}_+^*)^m$



Then $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq p^\star := \Phi(\sqrt{mn+1})$

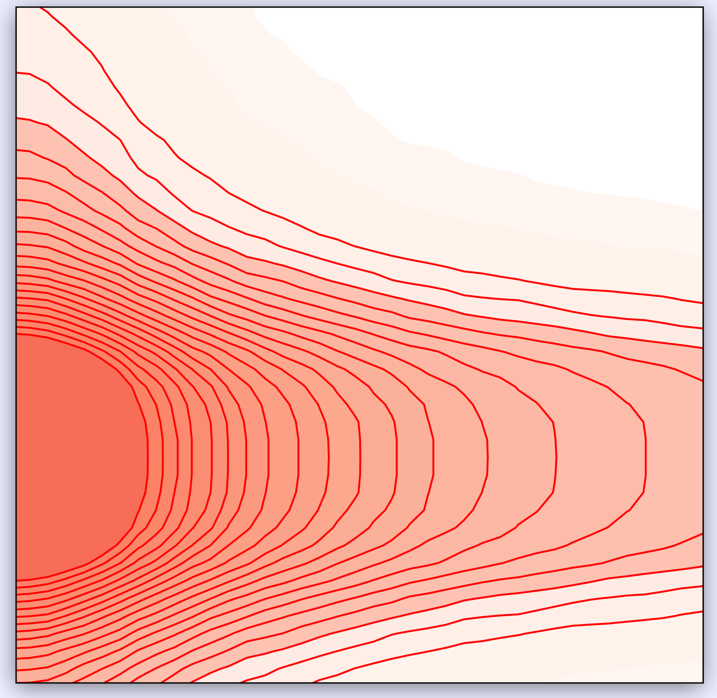
  - Proof leverages α-concavity $(\alpha = -1)$ of the radial function ρ and and the concave-α property of 1D-gaussian random variables.

- The Quadratic case
  - Consider:

$$g(x,z) = z^\top W(x) z + 2\sum_{i=1}^{n} a_i w_i^\top z_i + b$$

Symmetric definite positive with convex eigenvalues in x

≤ 0



and ξ a multivariate Gaussian random vector.

Then, $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq \Phi(\sqrt{m+3})$

  - Proof is more involved
    - $\rho(x,v) = \sup_{t \geq 0}\{t, g(x, \mu + tLv) \leq 0\} = \dfrac{-\beta(v) + \sqrt{\beta^2(v) - bh(x,v)}}{h(x,v)}$

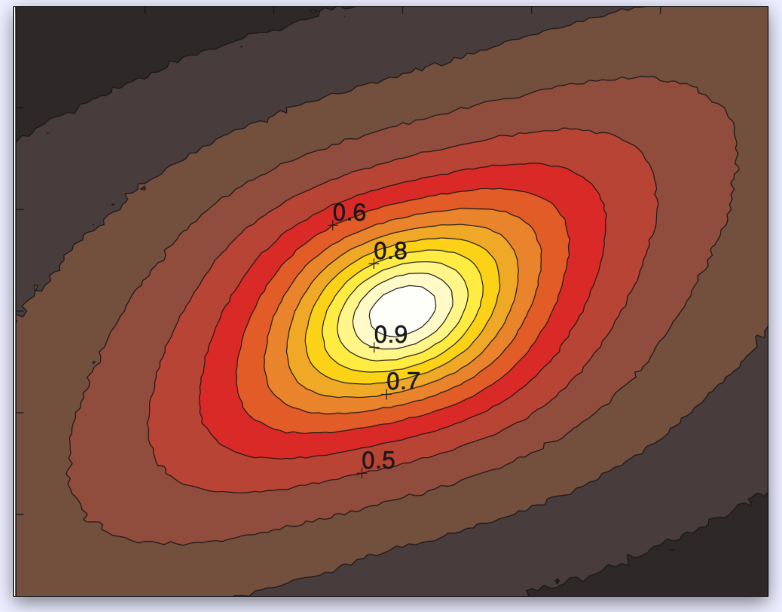
Application Examples

- The linear case
  - Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

Ξ: $m \times d$ centered multi-variate gaussian random matrix

$\beta \in (\mathbb{R}_+^*)^m$



Then $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq p^\star := \Phi(\sqrt{mn+1})$

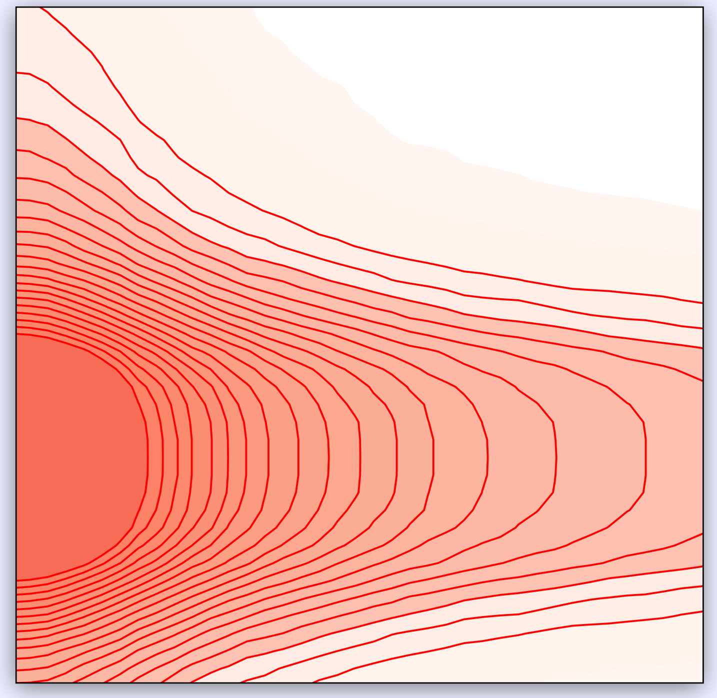
  - Proof leverages α-concavity $(\alpha = -1)$ of the radial function ρ and and the concave-α property of 1D-gaussian random variables.

- The Quadratic case
  - Consider:

$$g(x,z) = z^\top W(x) z + 2\sum_{i=1}^n a_i w_i^\top z_i + b$$

$W(x)$: Symmetric definite positive with convex eigenvalues in x

b: ≤ 0



and ξ a multivariate Gaussian random vector.

Then, $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq \Phi(\sqrt{m+3})$

  - Proof is more involved
    - $\rho(x,v) = \sup_{t \geq 0}\{t, g(x,\mu + tLv) \leq 0\} = \dfrac{-\beta(v) + \sqrt{\beta^2(v) - bh(x,v)}}{h(x,v)}$

$\beta(v) := \sum_{i=1}^d a_i w_i^\top Lv$

$h(x,v) := (Lv)^\top W(x) Lv$

Application Examples

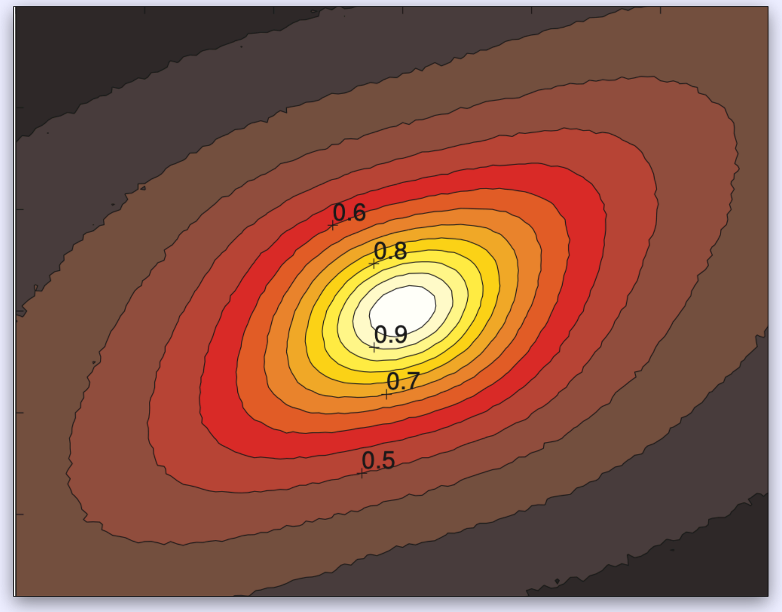
The linear case

- Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

Ξ: $m \times d$ centered multi-variate gaussian random matrix

$\beta \in (\mathbb{R}_+^*)^m$



Then $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq p^\star := \Phi(\sqrt{mn+1})$

- Proof leverages α-concavity $(\alpha = -1)$ of the radial function ρ and and the concave-α property of 1D-gaussian random variables.

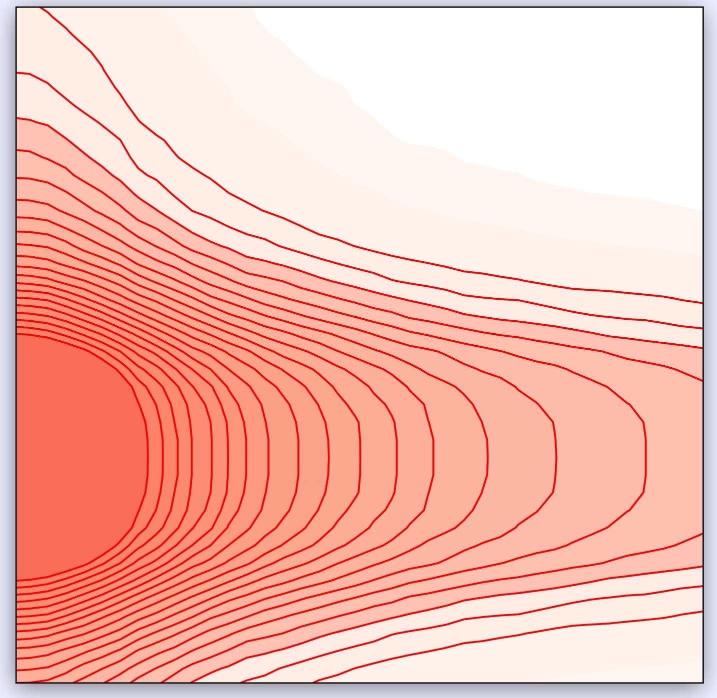
The Quadratic case

- Consider:

$$g(x,z) = z^\top W(x) z + 2\sum_{i=1}^{n} a_i w_i^\top z_i + b$$

$W(x)$: Symmetric definite positive with convex eigenvalues in x

$b \leq 0$



and ξ a multivariate Gaussian random vector.

Then, $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq \Phi(\sqrt{m+3})$

- Proof is more involved
  - $\rho(x,v) = \sup_{t \geq 0}\{t, g(x, \mu + tLv) \leq 0\} = \dfrac{-\beta(v) + \sqrt{\beta^2(v) - bh(x,v)}}{h(x,v)}$, where $\beta(v) := \sum_{i=1}^{d} a_i w_i^\top Lv$ and $h(x,v) := (Lv)^\top W(x) Lv$
  - ρ is G_v-concave with $G_v(t) = -\left(\dfrac{b}{t} + \beta(v)\right)^2$

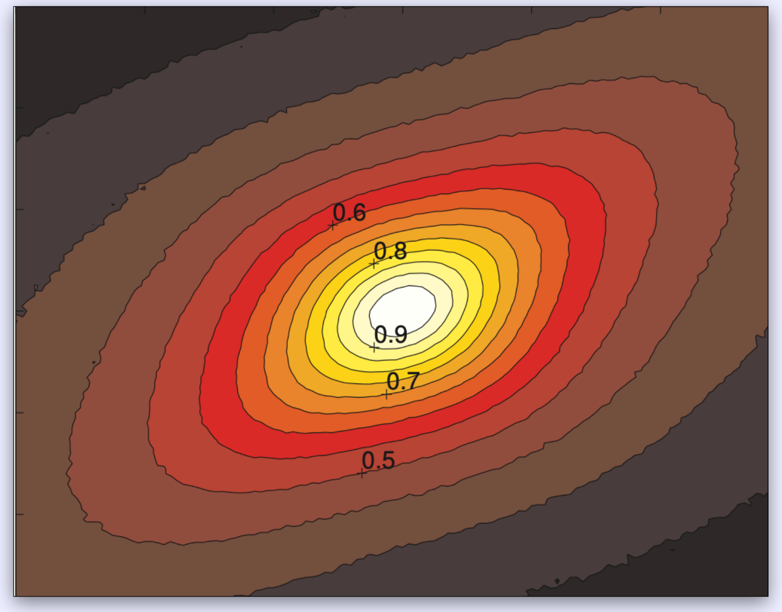
Application Examples

- The linear case
  - Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

Ξ: $m \times d$ centered multi-variate gaussian random matrix

$\beta \in (\mathbb{R}_+^*)^m$



Then $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq p^\star := \Phi(\sqrt{mn+1})$

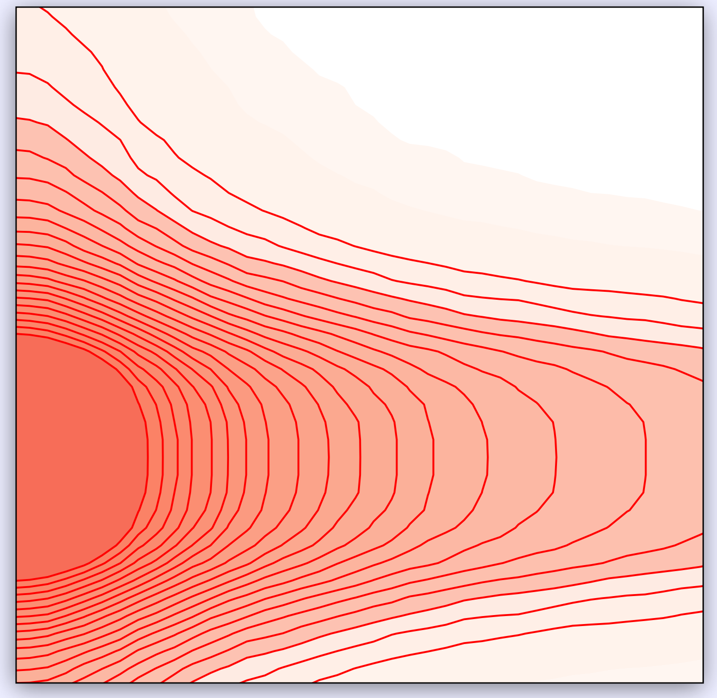
  - Proof leverages α-concavity $(\alpha = -1)$ of the radial function ρ and and the concave-α property of 1D-gaussian random variables.

- The Quadratic case
  - Consider:

$$g(x,z) = z^\top W(x) z + 2\sum_{i=1}^{n} a_i w_i^\top z_i + b$$

$W(x)$: Symmetric definite positive with convex eigenvalues in x

$b \leq 0$



and ξ a multivariate Gaussian random vector.

Then, $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq \Phi(\sqrt{m+3})$

  - Proof is more involved
    - $\rho(x,v) = \sup_{t \geq 0}\{t, g(x,\mu + tLv) \leq 0\} = \dfrac{-\beta(v) + \sqrt{\beta^2(v) - bh(x,v)}}{h(x,v)}$, where $\beta(v) := \sum_{i=1}^{d} a_i w_i^\top Lv$ and $h(x,v) := (Lv)^\top W(x) Lv$
    - ρ is G_v-concave with $G_v(t) = -\left(\frac{b}{t} + \beta(v)\right)^2$
    - How about $F_\mathcal{R} \circ G_v^{-1}$?

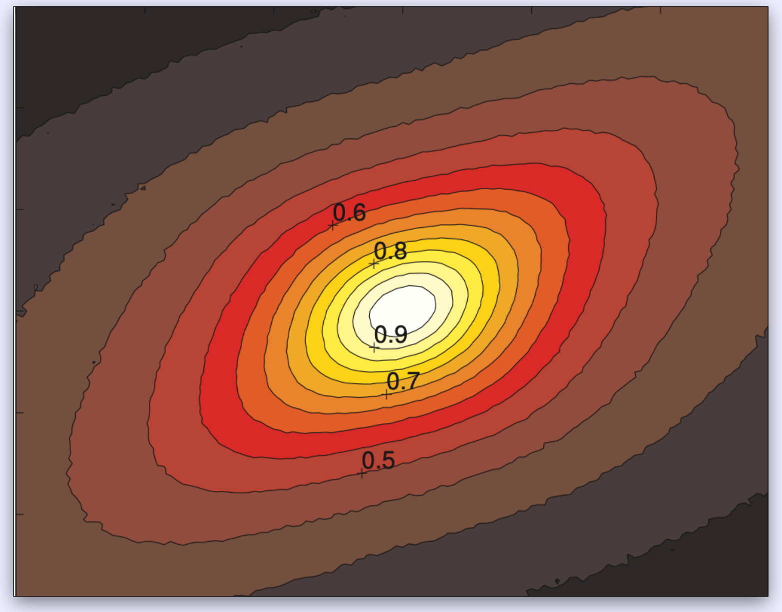
Application Examples

- The linear case
  - Consider a constraint of the form:

$$\mathbb{P}[\Xi x \leq \beta]$$

Ξ: $m \times d$ centered multi-variate gaussian random matrix

$\beta \in (\mathbb{R}_+^*)^m$



Then $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq p^\star := \Phi(\sqrt{mn+1})$

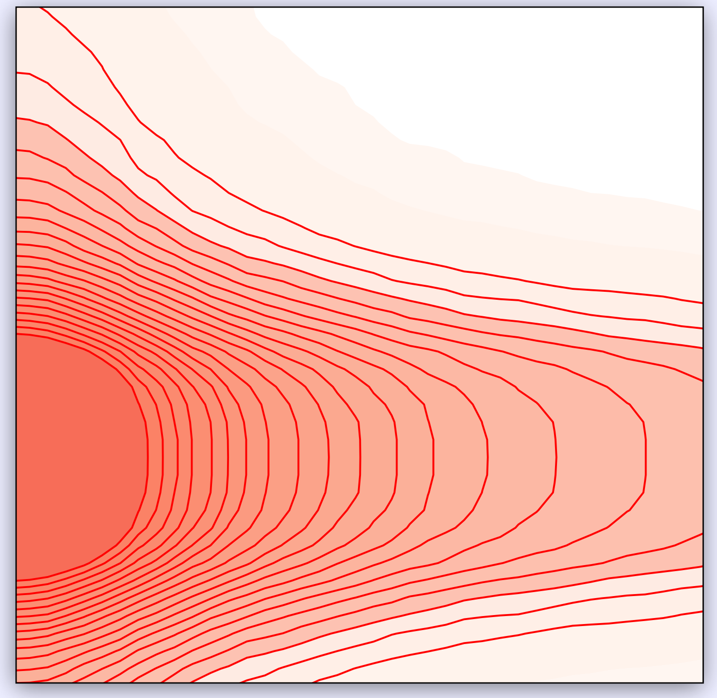
  - Proof leverages α-concavity $(\alpha = -1)$ of the radial function ρ and and the concave-α property of 1D-gaussian random variables.

- The Quadratic case
  - Consider:

$$g(x,z) = z^\top W(x) z + 2\sum_{i=1}^{n} a_i w_i^\top z_i + b$$

$W(x)$: Symmetric definite positive with convex eigenvalues in x

$b \leq 0$



and ξ a multivariate Gaussian random vector.

Then, $M_p = \{x,\ \mathbb{P}[g(x,\xi) \leq 0] \geq p\}$ is convex for all $p \geq \Phi(\sqrt{m+3})$

  - Proof is more involved
    - $\rho(x,v) = \sup_{t \geq 0}\{t, g(x, \mu + tLv) \leq 0\} = \dfrac{-\beta(v) + \sqrt{\beta^2(v) - bh(x,v)}}{h(x,v)}$, where $\beta(v) := \sum_{i=1}^{d} a_i w_i^\top Lv$ and $h(x,v) := (Lv)^\top W(x) Lv$
    - ρ is G_v-concave with $G_v(t) = -\left(\frac{b}{t} + \beta(v)\right)^2$
    - How about $F_{\mathcal{R}} \circ G_v^{-1}$?
      - G_v^{-1} is (-3)-concave!
      - $F_{\mathcal{R}}$ is concave-(-3)!

II – A Convex Bilevel Approach to Solve Chance Constrained Problems

Chance constrained Problems

- We consider now chance constrained problems

$$\min_{x \in \mathbb{R}^d} f(x)$$
$$\text{s.t. } \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

Chance constrained Problems

- We consider now chance constrained problems

$$\min_{x \in \mathbb{R}^d} f(x)$$
$$\text{s.t. } \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

- In this talk,
  - f is convex.
  - $g(., z)$ is convex for all $z \in \mathbb{R}^m$
  - ξ is discrete : $\xi \in \{\xi_1, \ldots, \xi_n\} \subset \mathbb{R}^m$

    For simplicity, we assume: $\mathbb{P}[\xi = \xi_i] = \frac{1}{n}, \; \forall\, 1 \leq i \leq n$

Chance constrained Problems

- We consider now chance constrained problems

$$\min_{x \in \mathbb{R}^d} f(x)$$
$$\text{s.t. } \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

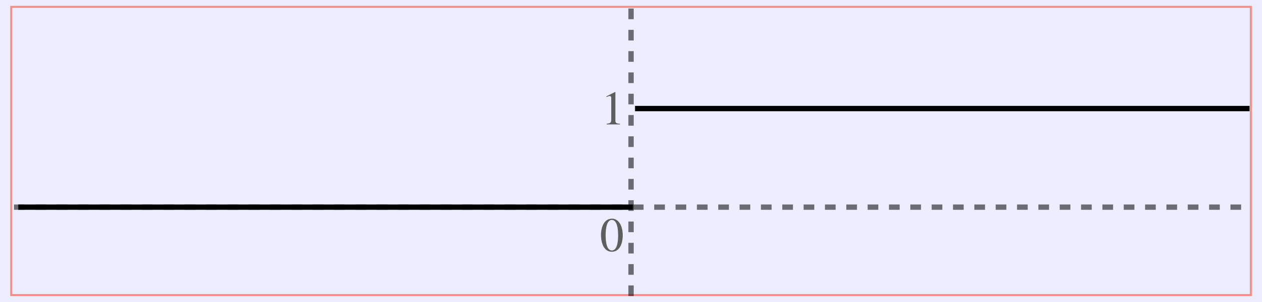
- In this talk,
  - f is convex.
  - $g(., z)$ is convex for all $z \in \mathbb{R}^m$
  - ξ is discrete : $\xi \in \{\xi_1, \ldots, \xi_n\} \subset \mathbb{R}^m$

    For simplicity, we assume: $\mathbb{P}[\xi = \xi_i] = \frac{1}{n}, \; \forall\, 1 \leq i \leq n$

- Many existing approaches including
  - MINLP approaches

    Pagnoncelli, Ahmed, Shapiro(2009)
  - DoC approaches

    Hong, Yang, Zhang (2009)



II - A Convex Bilevel Approach to Solve Chance Constrained Problems

1 Chance Constraints are Bilevel Programs

2 Penalization Method

3 TACO

4 Numerical Illustrations

1 Revealing the bilevel structure of Chance Constraints

1 Chance Constraints are Bilevel Programs

2 Penalization Method

3 TACO

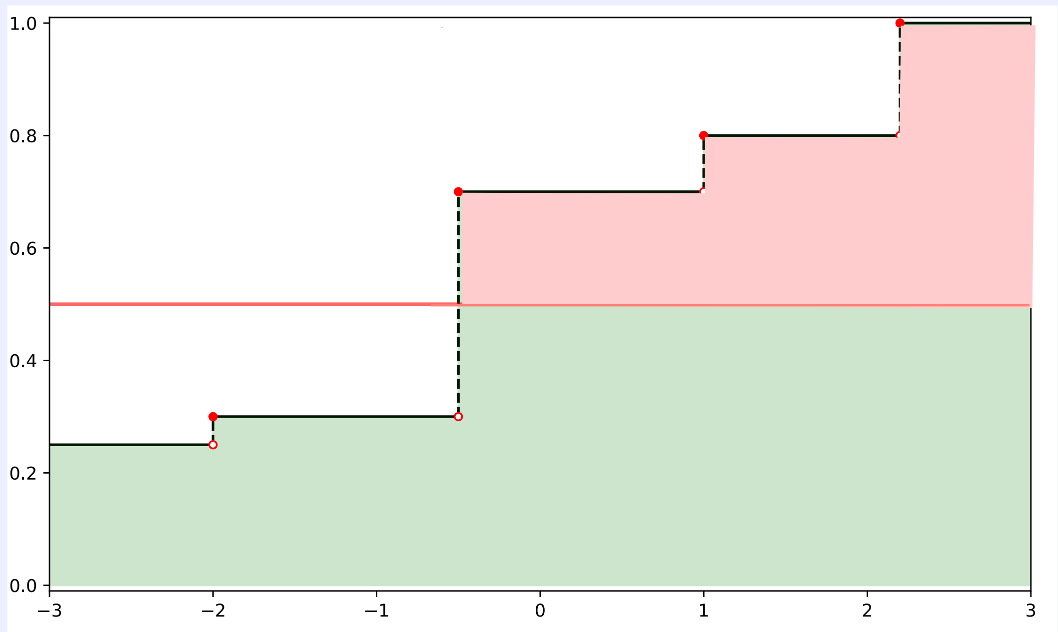
4 Numerical Illustrations

CDFs, Quantiles & Superquantiles

- Recall that for any real random variable U,
  - its cumulative distribution function, $F_U : \mathbb{R} \to [0, 1]$:

$$F_U(t) = \mathbb{P}[U \leq t]$$

Cumulative distribution function



CDFs, Quantiles & Superquantiles

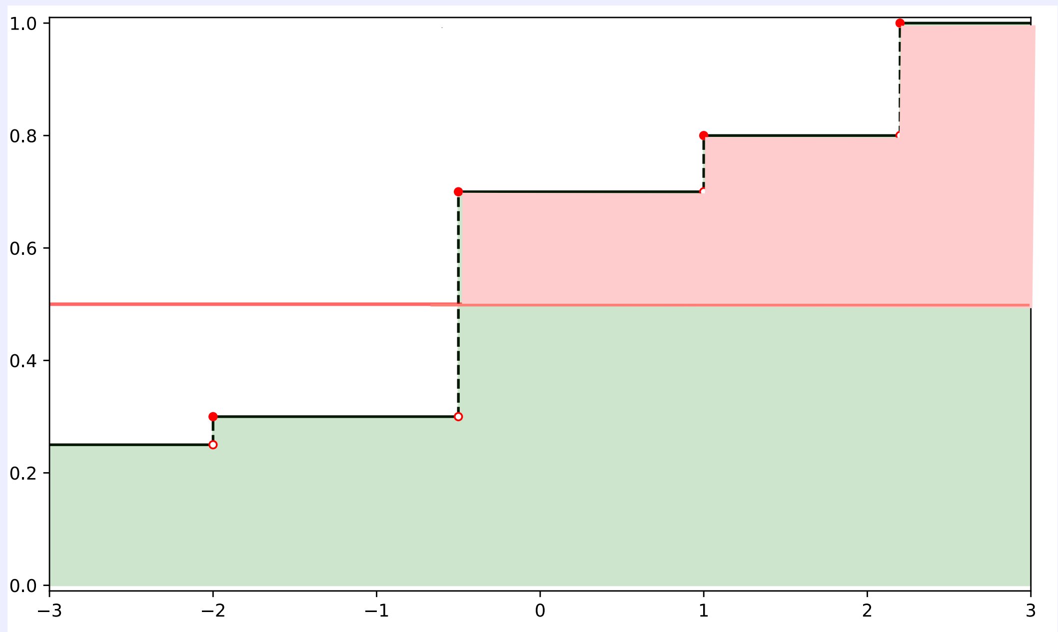
- Recall that for any real random variable U,
  - its cumulative distribution function, $F_U : \mathbb{R} \to [0, 1]$:

$$F_U(t) = \mathbb{P}[U \leq t]$$

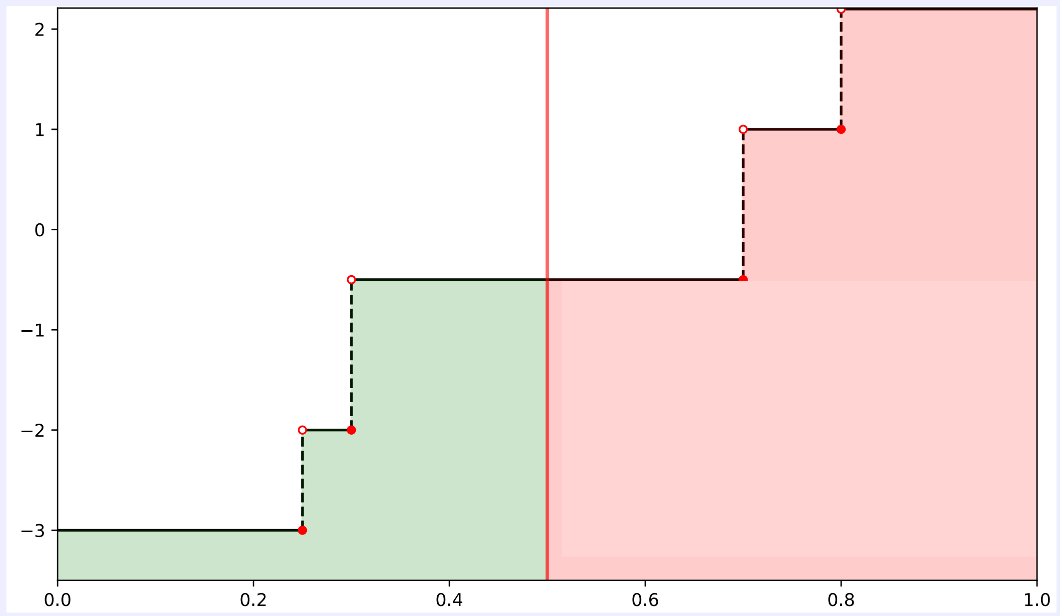
  - for any $p \in [0, 1)$, its p-quantile $Q_p(U)$:

$$Q_p(U) = \inf \{t \in \mathbb{R}, \mathbb{P}[U \leq t] \geq p\}$$

Cumulative distribution function



Quantile function



CDFs, Quantiles & Superquantiles

- Recall that for any real random variable U,
  - its cumulative distribution function, $F_U : \mathbb{R} \to [0,1]$:

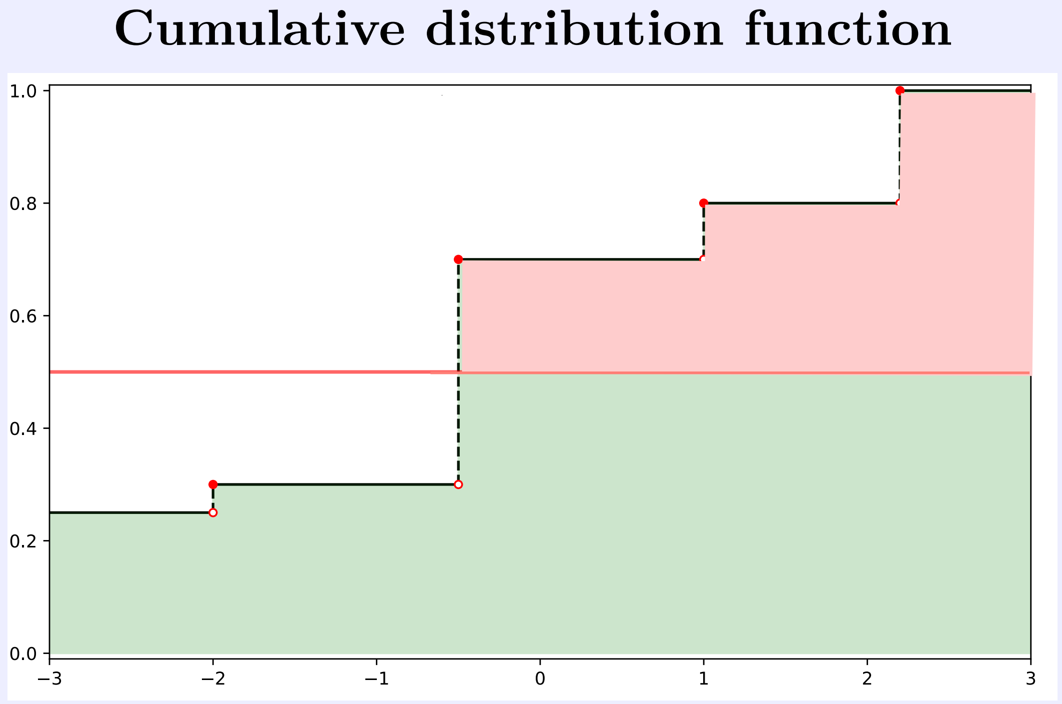
$$F_U(t) = \mathbb{P}[U \leq t]$$

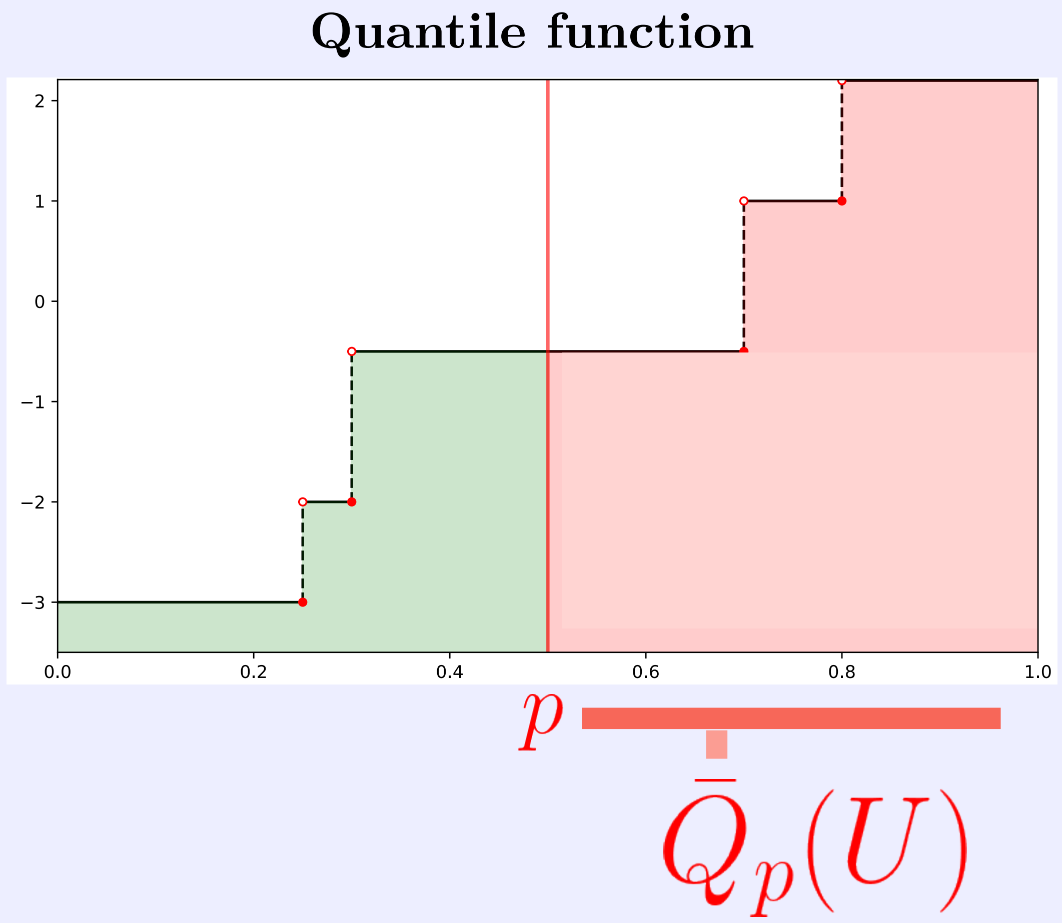
  - for any $p \in [0,1)$, its p-quantile $Q_p(U)$:

$$Q_p(U) = \inf \{t \in \mathbb{R}, \mathbb{P}[U \leq t] \geq p\}$$

  - for any $p \in [0,1)$, its p-superquantile $\bar{Q}_p(U)$:

$$\bar{Q}_p(U) = \frac{1}{1-p} \int_{p'=p}^{1} Q_{p'}(U) dp'$$





CDFs, Quantiles & Superquantiles

- Recall that for any real random variable U,
  - its cumulative distribution function, $F_U : \mathbb{R} \to [0, 1]$:

Invert

$$F_U(t) = \mathbb{P}[U \leq t]$$

  - for any $p \in [0, 1)$, its p-quantile $Q_p(U)$:

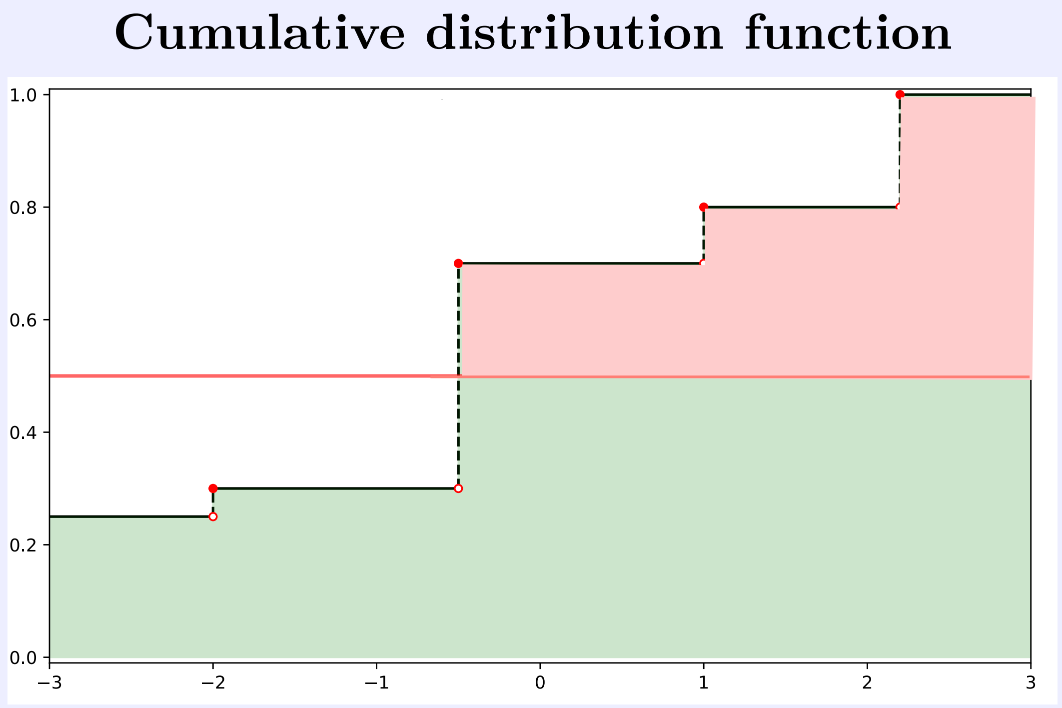
Integrate

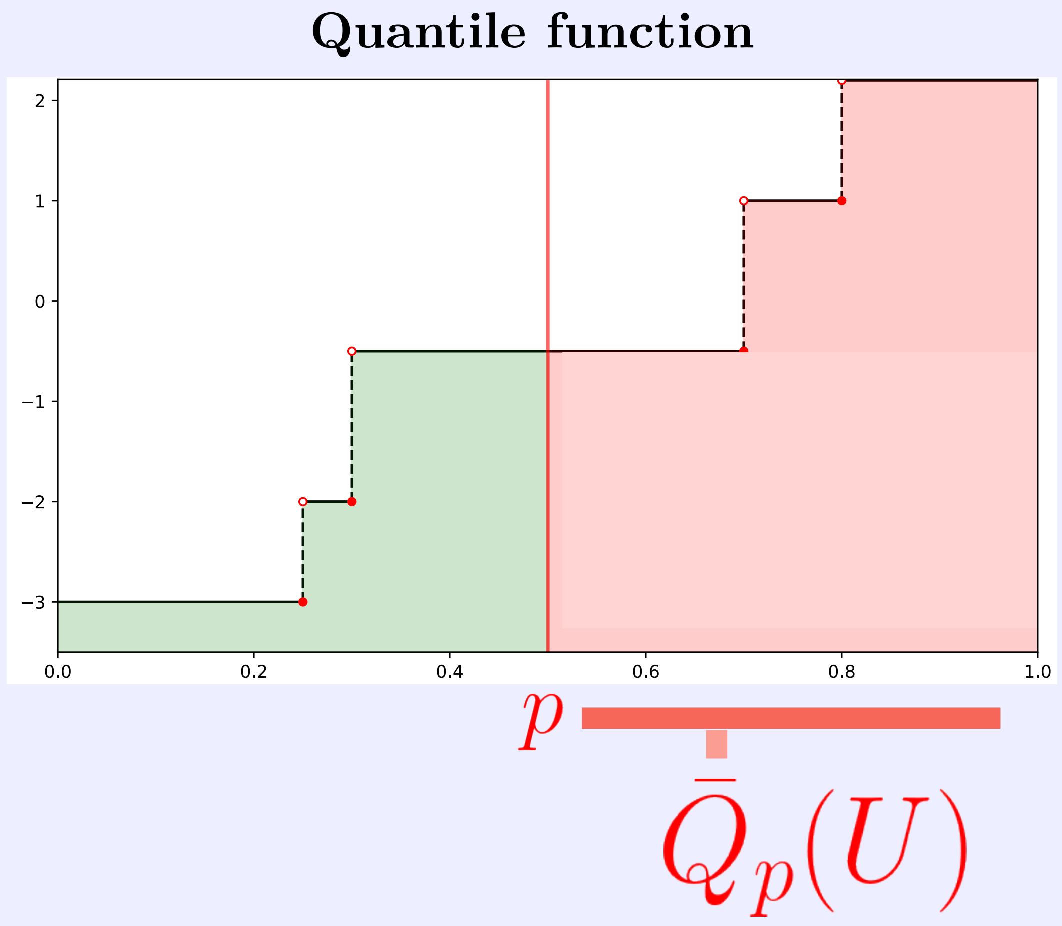
Differentiate

$$Q_p(U) = \inf \{t \in \mathbb{R}, \mathbb{P}[U \leq t] \geq p\}$$

  - for any $p \in [0, 1)$, its p-superquantile $\bar{Q}_p(U)$:

$$\bar{Q}_p(U) = \frac{1}{1-p} \int_{p'=p}^{1} Q_{p'}(U) dp'$$





Rockafellar's Duality Result

- **Rockafellar, Uryasev (2000)**: Superquantile and quantiles are optimal value and optimal solutions resp. of a same one-dimensional convex optimization problem.

$$\bar{Q}_p(U) = \min_{\eta \in \mathbb{R}} \eta + \frac{1}{1-p}\mathbb{E}[\max(U-\eta,0)]$$

$$Q_p(U) = \operatorname*{argmin}_{\eta \in \mathbb{R}} \eta + \frac{1}{1-p}\mathbb{E}[\max(U-\eta,0)]$$

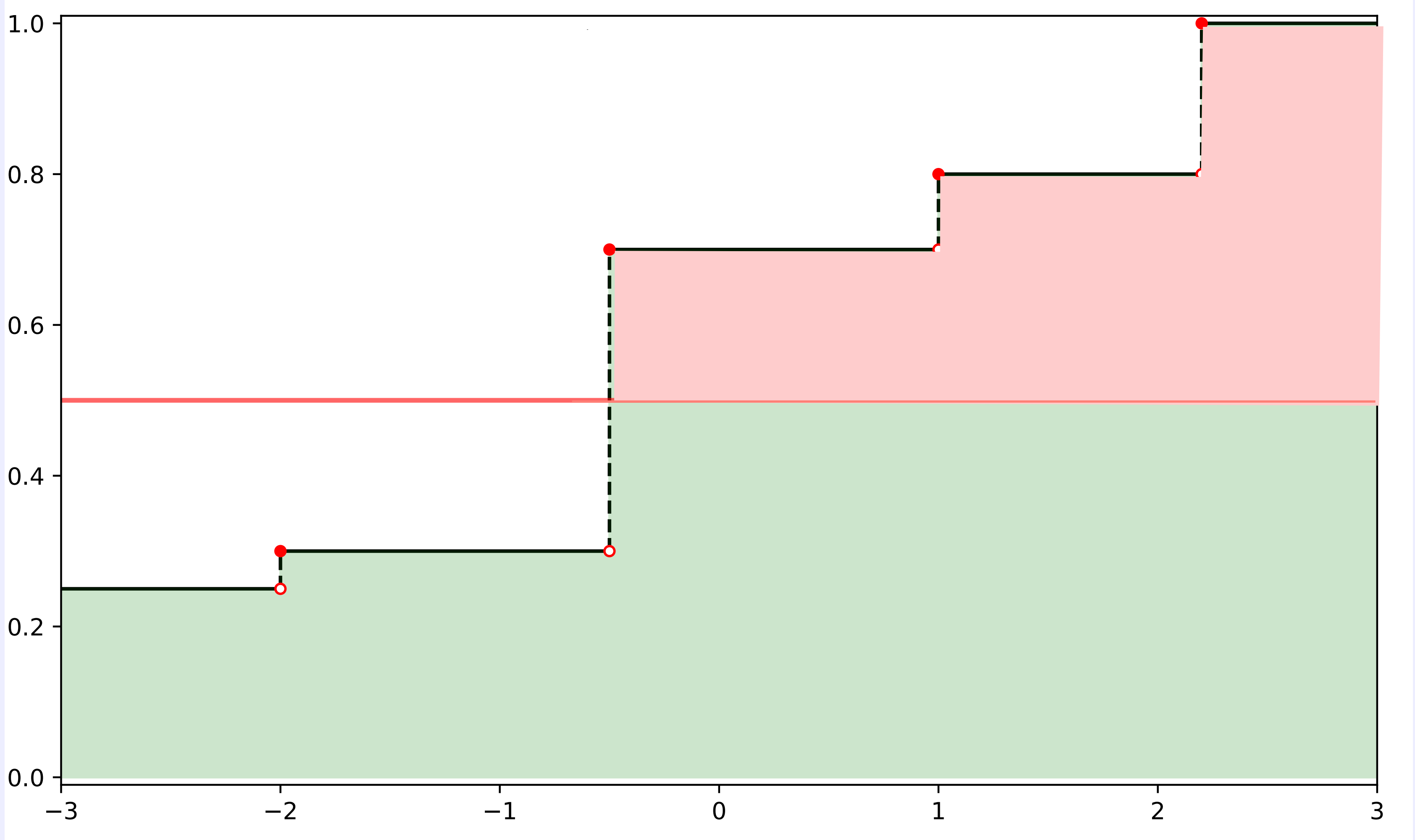
Rockafellar's Duality Result

- **Rockafellar, Uryasev (2000)**: Superquantile and quantiles are optimal value and optimal solutions resp. of a same one-dimensional convex optimization problem.

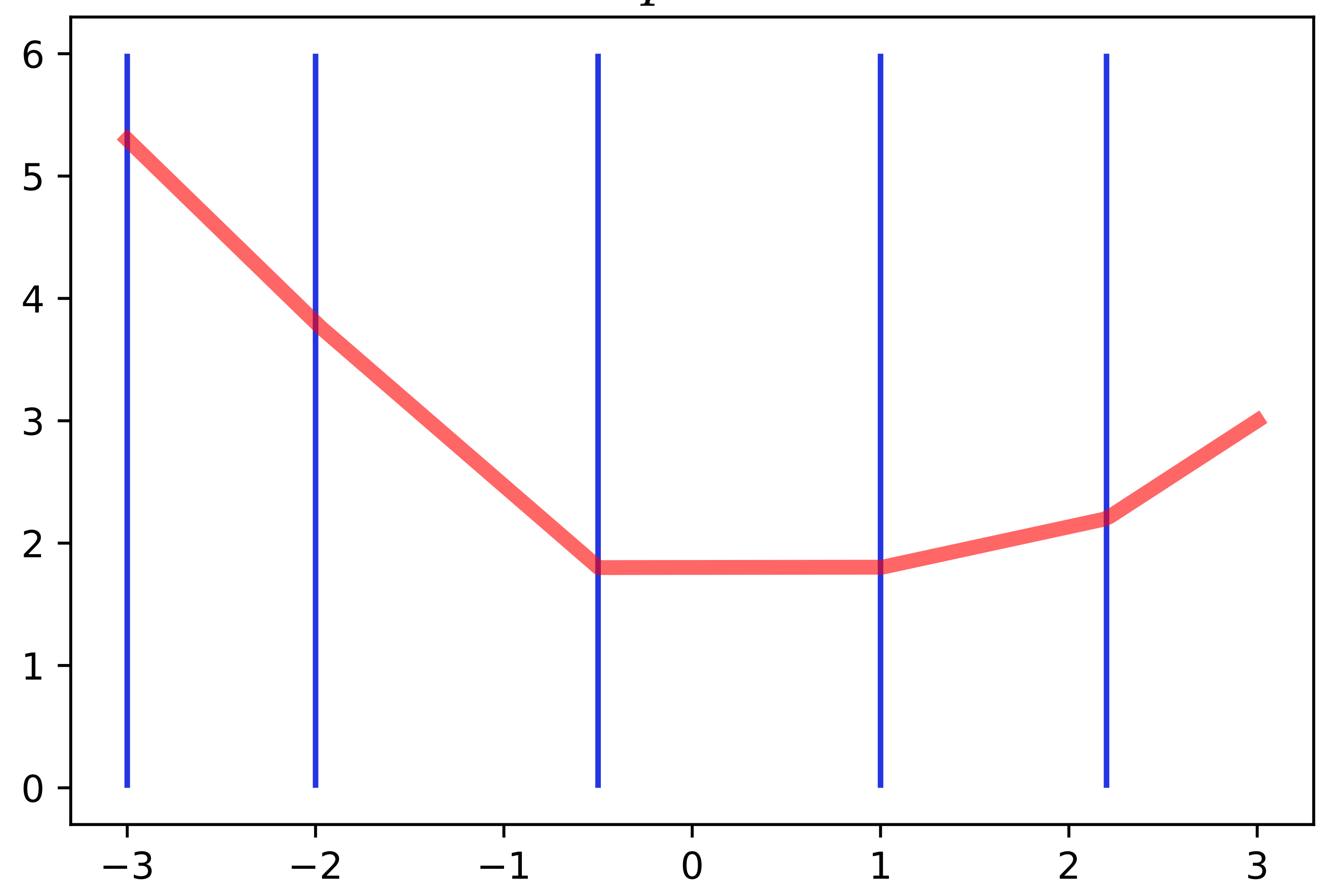
$$\bar{Q}_p(U) = \min_{\eta \in \mathbb{R}} \eta + \frac{1}{1-p}\mathbb{E}[\max(U-\eta, 0)]$$

$$Q_p(U) = \operatorname{argmin}_{\eta \in \mathbb{R}} \eta + \frac{1}{1-p}\mathbb{E}[\max(U-\eta, 0)]$$

Cumulative distribution function



$\eta \mapsto \eta + \frac{1}{1-p}\mathbb{E}[\max(U-\eta, 0)]$



From Chance Constraints to Bilevel Programs

- **Our approach**: rewrite chance constraints as

$$\mathbb{P}[g(x,\xi) \leq 0] \geq p \;\Leftrightarrow\; Q_p(g(x,\xi)) \leq 0$$

From Chance Constraints to Bilevel Programs

- **Our approach**: rewrite chance constraints as

$$\mathbb{P}[g(x,\xi) \leq 0] \geq p \iff Q_p(g(x,\xi)) \leq 0$$

$$\iff \eta \leq 0$$

$$\eta \in \operatorname*{argmin}_{s \in \mathbb{R}} s + \frac{1}{1-p}\mathbb{E}\left[\max(g(x,\xi) - s, 0)\right]$$

- We obtain the following bilevel program:

Upper Level

$$\min_{x \in \mathbb{R}^d, \eta \in \mathbb{R}} \quad f(x)$$
$$\text{s.t.} \quad \eta \leq 0$$

Lower Level

$$\eta \in \operatorname*{argmin}_{s \in \mathbb{R}} s + \frac{1}{1-p}\mathbb{E}\left[\max(g(x,\xi) - s, 0)\right]$$

2 A double penalization method for Chance Constraints

Recall Penalization on a Picture

- The penalization procedure

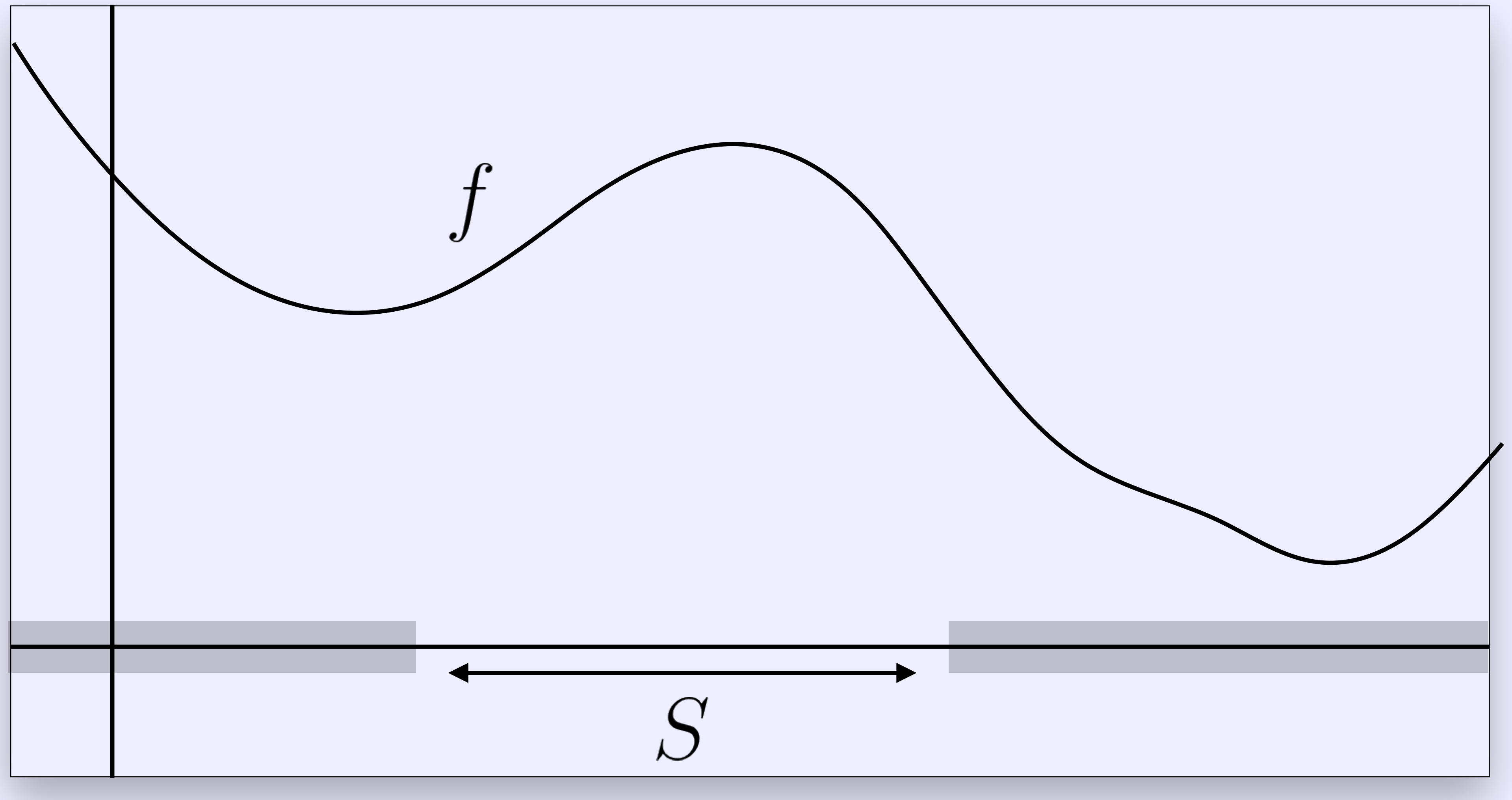
$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

Penalty function

$$P(x) \begin{cases} = 0 \text{ if } x \in S \\ > 0 \text{ if } x \notin S \end{cases}$$

Penalized Problem

$$\min_{x \in \mathbb{R}^d} f(x) + \mu P(x)$$



Recall Penalization on a Picture

- The penalization procedure

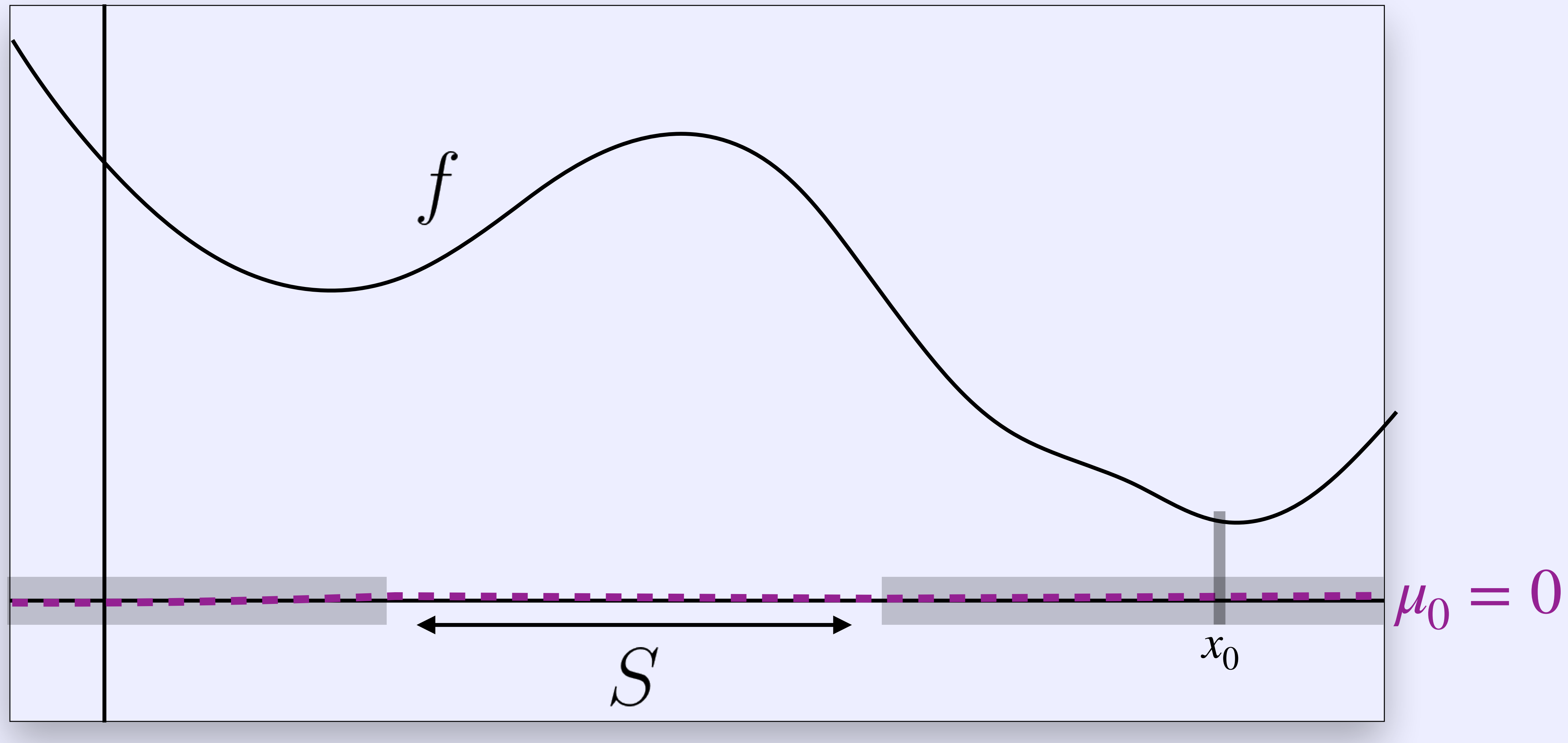
$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

Penalty function

$$P(x) \begin{cases} = 0 \text{ if } x \in S \\ > 0 \text{ if } x \notin S \end{cases}$$

Penalized Problem

$$\min_{x \in \mathbb{R}^d} f(x) + \mu P(x)$$



Recall Penalization on a Picture

- The penalization procedure

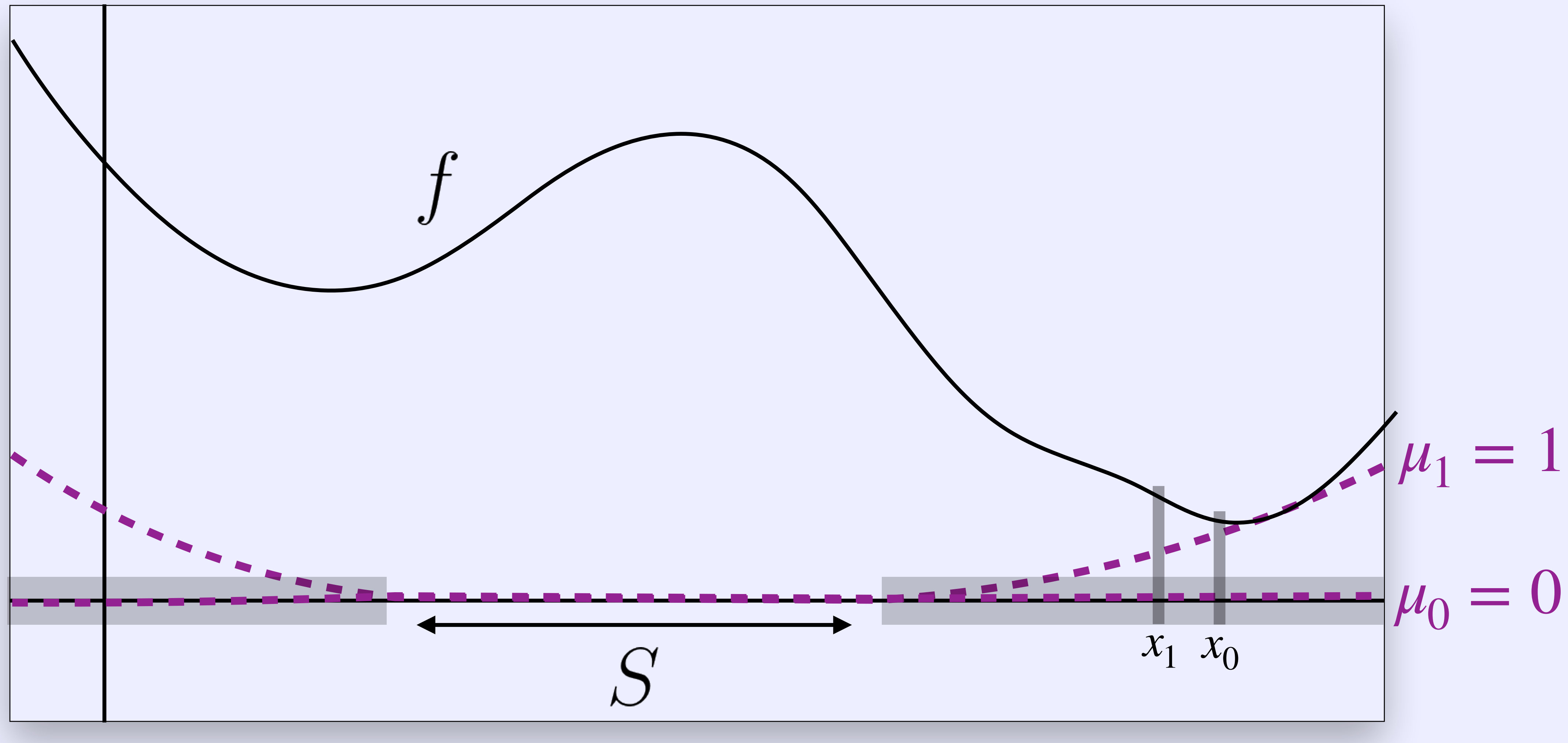
$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

Penalty function

$$P(x) \begin{cases} = 0 \text{ if } x \in S \\ > 0 \text{ if } x \notin S \end{cases}$$

Penalized Problem

$$\min_{x \in \mathbb{R}^d} f(x) + \mu P(x)$$



Recall Penalization on a Picture

- The penalization procedure

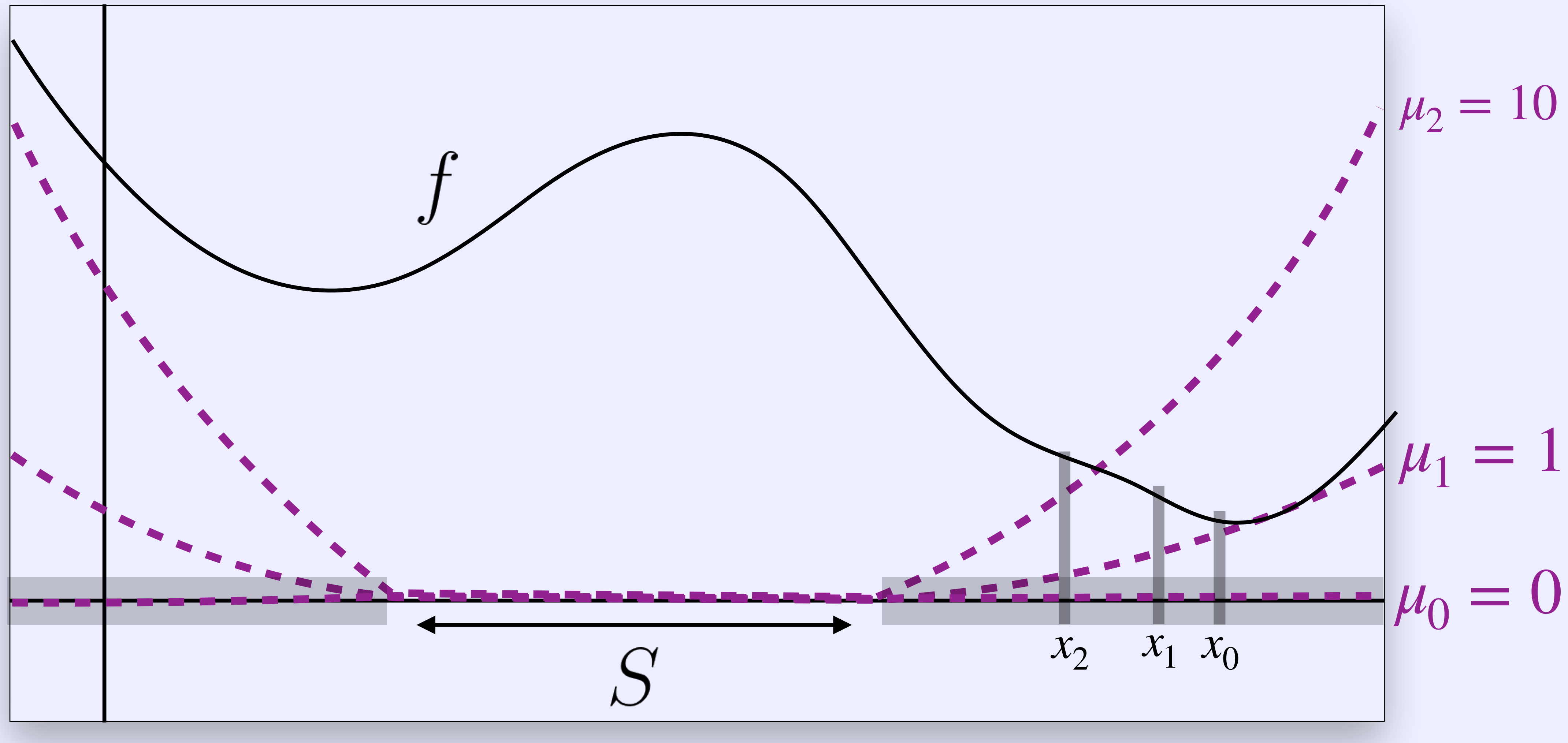
$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

Penalty function

$$P(x) \begin{cases} = 0 \text{ if } x \in S \\ > 0 \text{ if } x \notin S \end{cases}$$

Penalized Problem

$$\min_{x \in \mathbb{R}^d} f(x) + \mu P(x)$$



Recall Penalization on a Picture

- The penalization procedure

$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

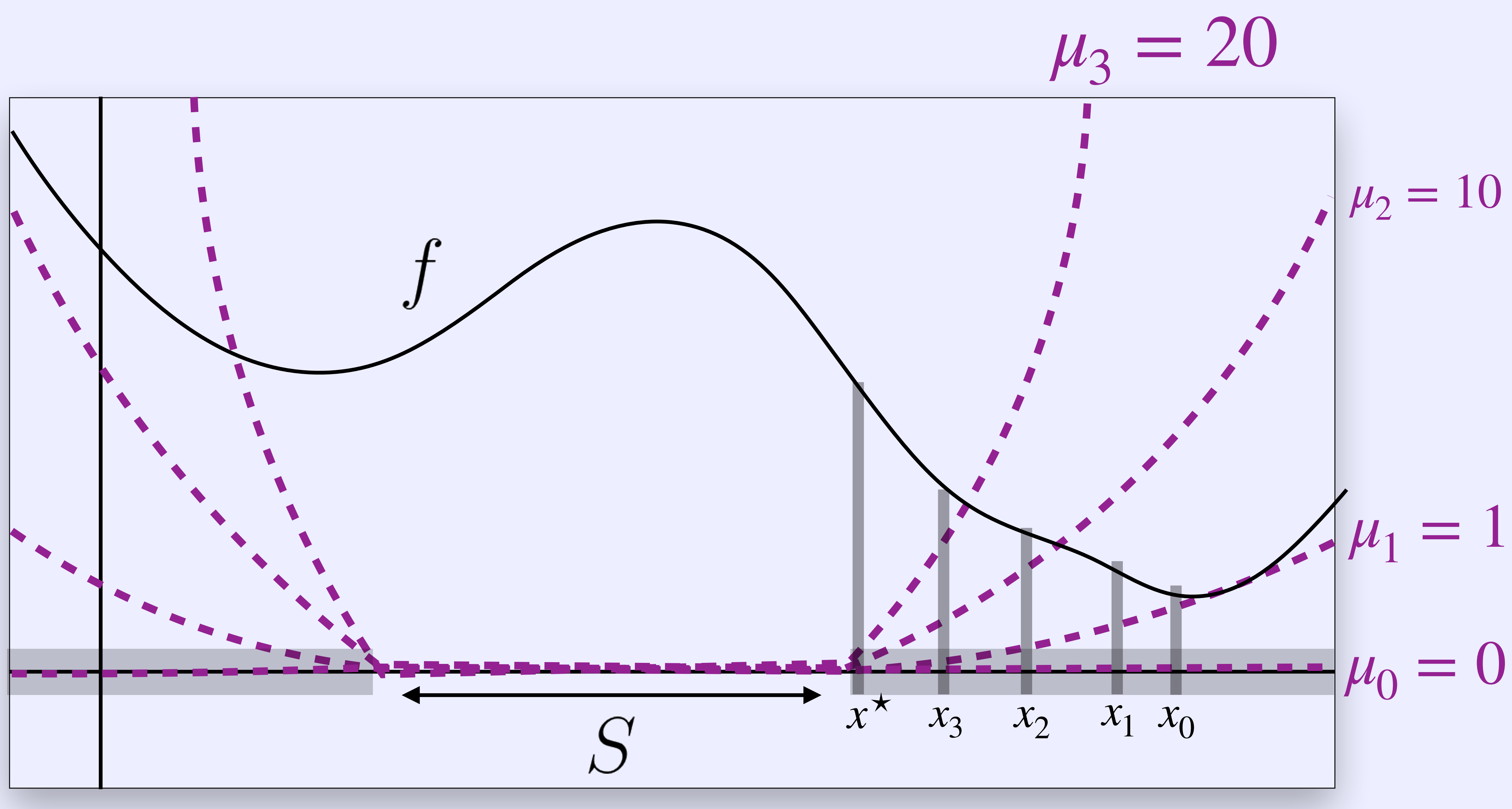
Penalty function

$$P(x) \begin{cases} = 0 \text{ if } x \in S \\ > 0 \text{ if } x \notin S \end{cases}$$

Penalized Problem

$$\min_{x \in \mathbb{R}^d} f(x) + \mu P(x)$$

Any cluster point of the sequence of solutions $(x_k)_{k \geq 0}$ is a solution of the constrained problem.



Recall Penalization on a Picture

- The penalization procedure

$$\min_{x\in\mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

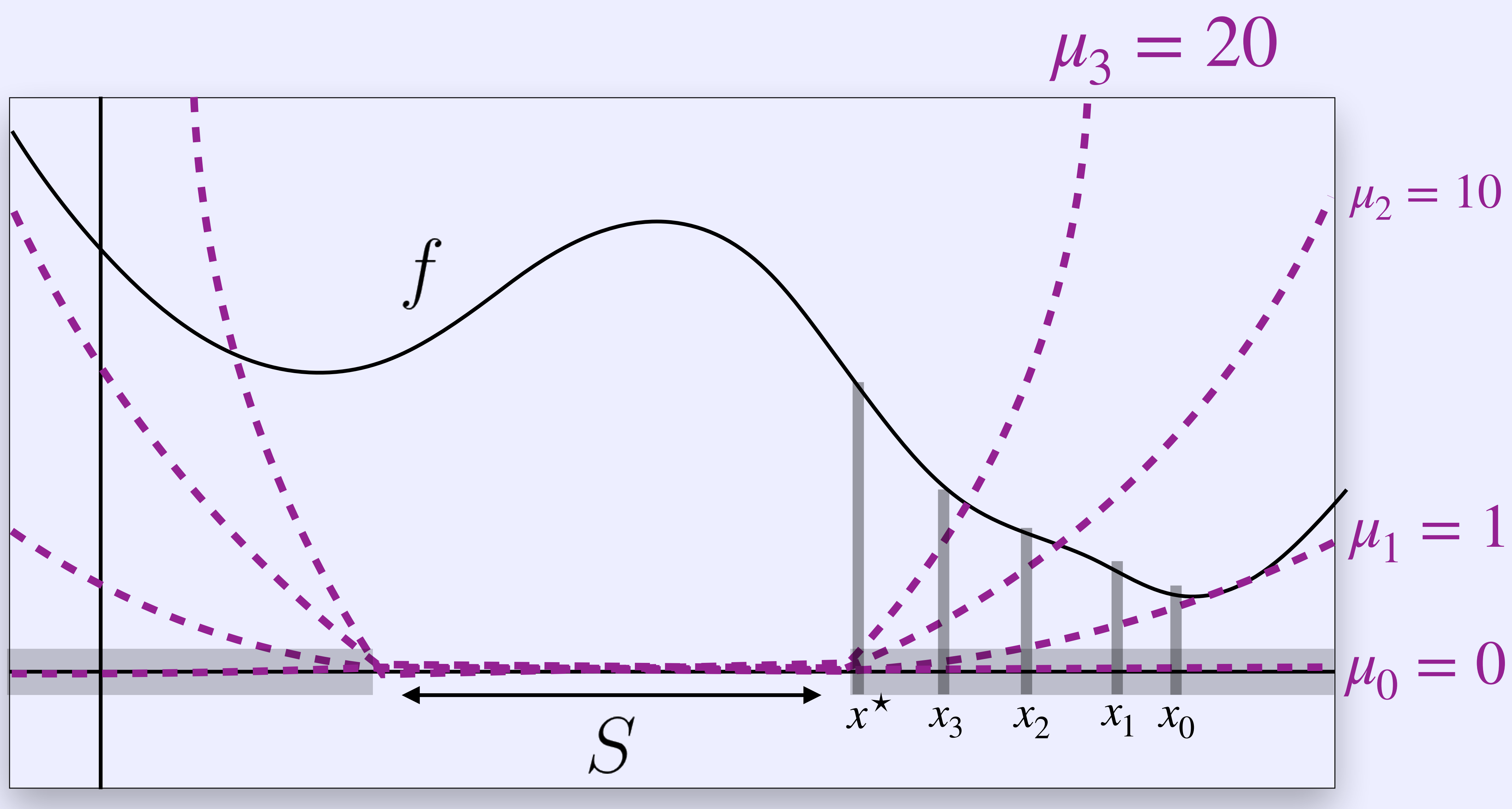
Penalty function

$$P(x) \begin{cases} = 0 \text{ if } x \in S \\ > 0 \text{ if } x \notin S \end{cases}$$

Penalized Problem

$$\min_{x\in\mathbb{R}^d} f(x) + \mu P(x)$$

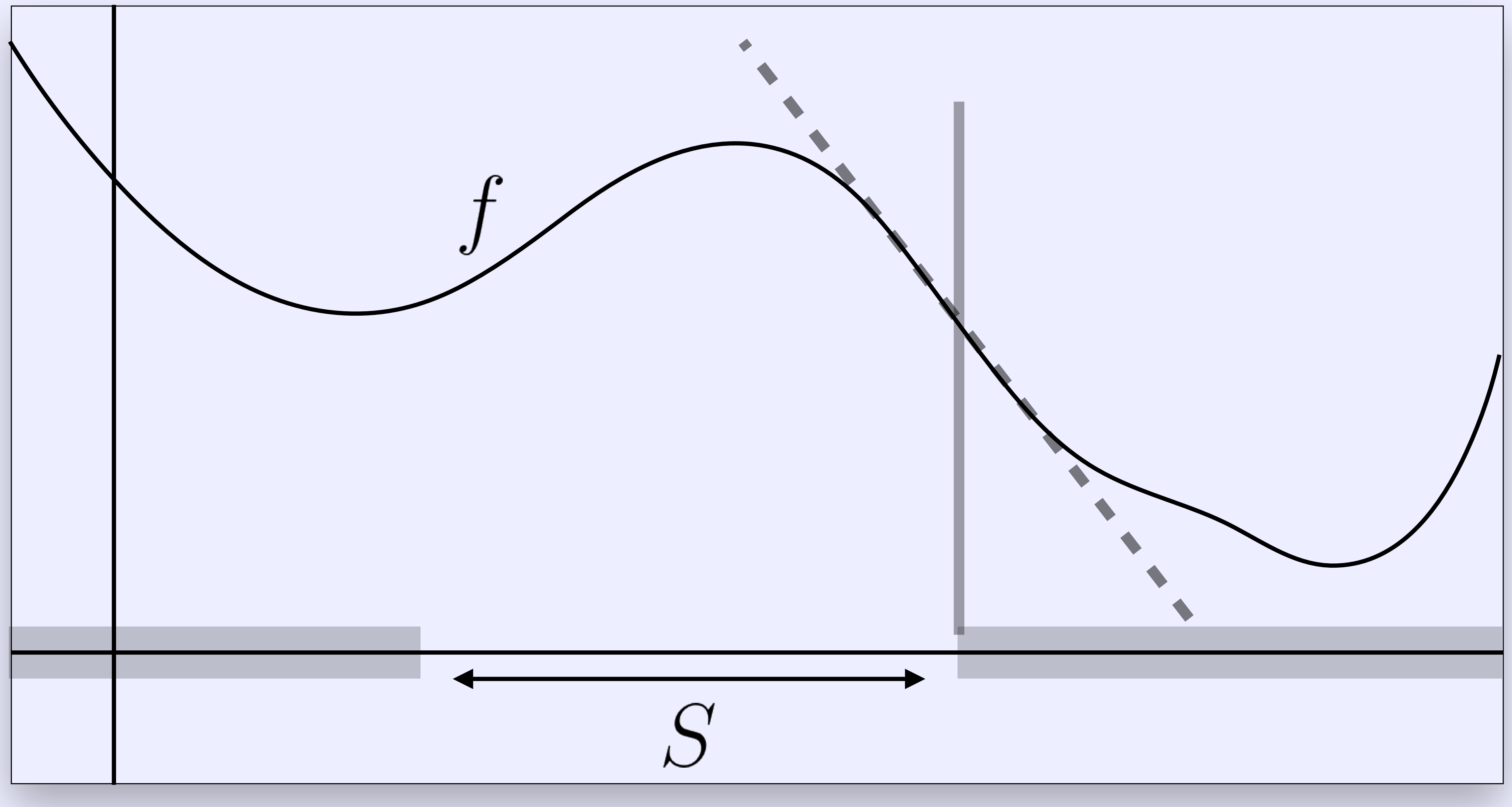
Any cluster point of the sequence of solutions $(x_k)_{k\geq 0}$ is a solution of the constrained problem.



- Exact penalization

$$\min_{x\in\mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

Assume f to be K-lipschitz on $\mathbb{R}^d$.



Recall Penalization on a Picture

- The penalization procedure

$$\min_{x\in\mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

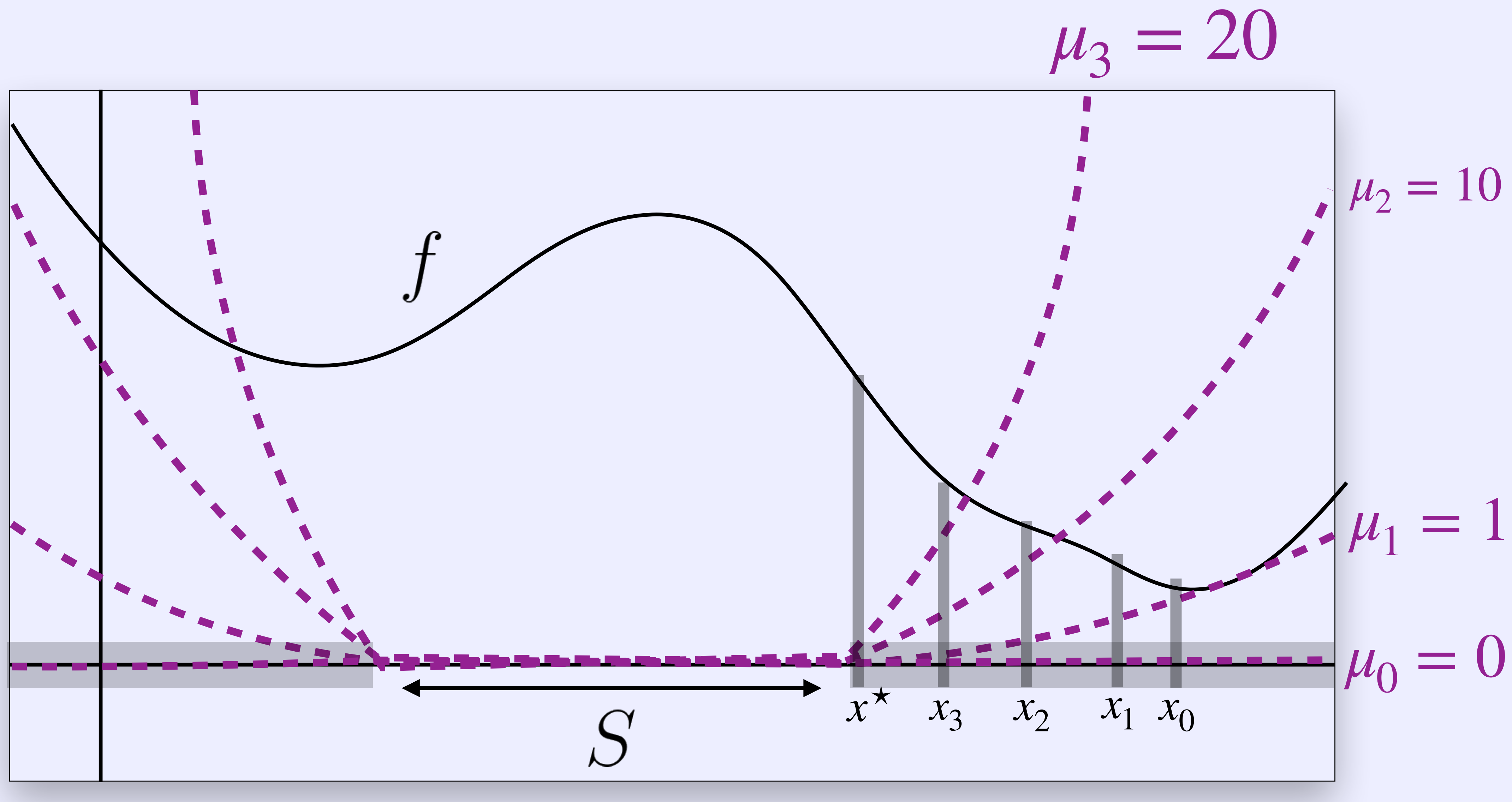
Penalty function

$$P(x) \begin{cases} = 0 \text{ if } x \in S \\ > 0 \text{ if } x \notin S \end{cases}$$

Penalized Problem

$$\min_{x\in\mathbb{R}^d} f(x) + \mu P(x)$$

Any cluster point of the sequence of solutions $(x_k)_{k\geq 0}$ is a solution of the constrained problem.

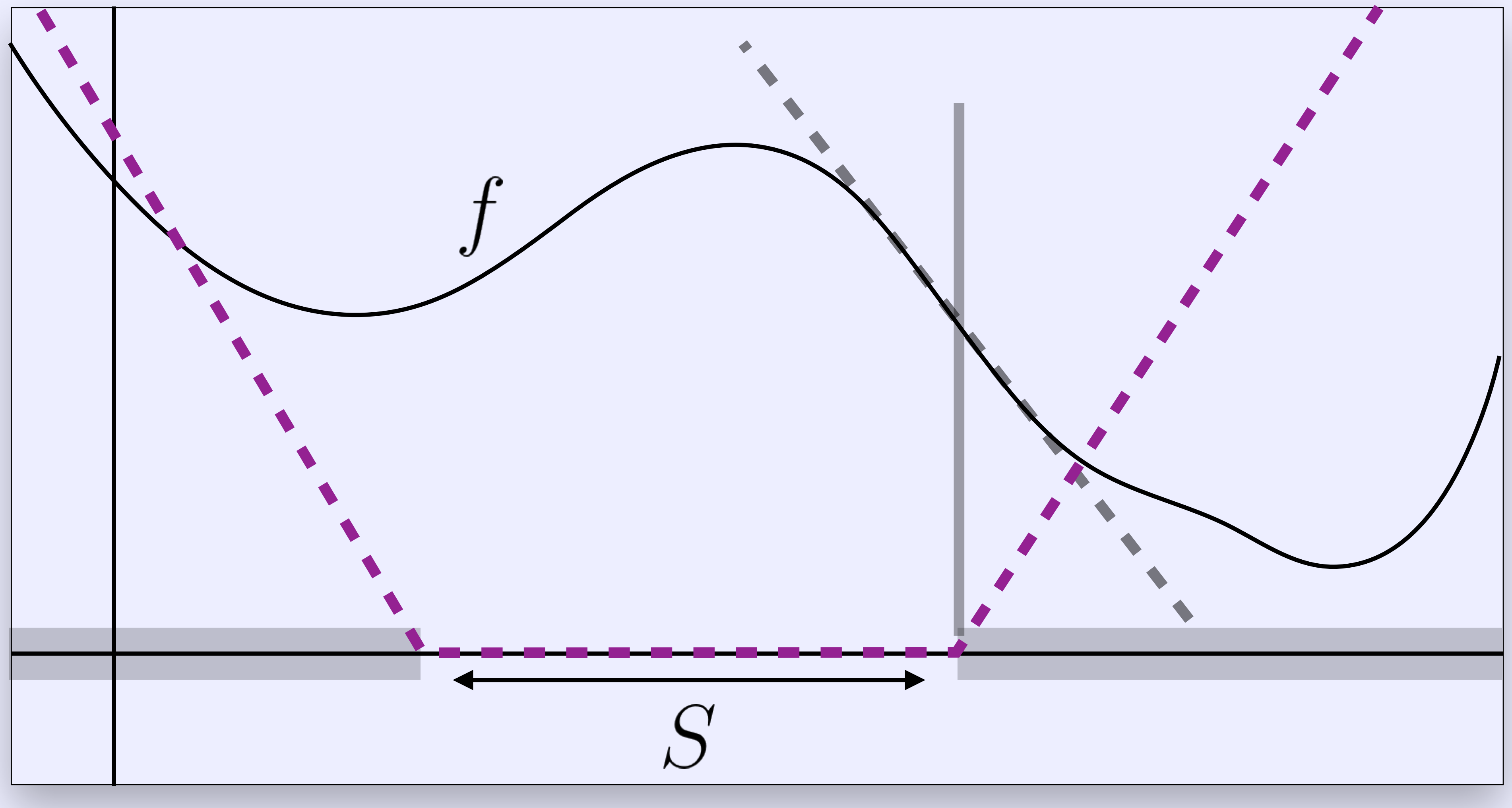


- Exact penalization

$$\min_{x\in\mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

Assume f to be K-lipschitz on $\mathbb{R}^d$.

Then, for any $K' > K$, this problem has the same set of minimisers as $\min_{x\in\mathbb{R}^d} f(x) + K' d_S(x)$



Recall Penalization on a Picture

The penalization procedure

$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

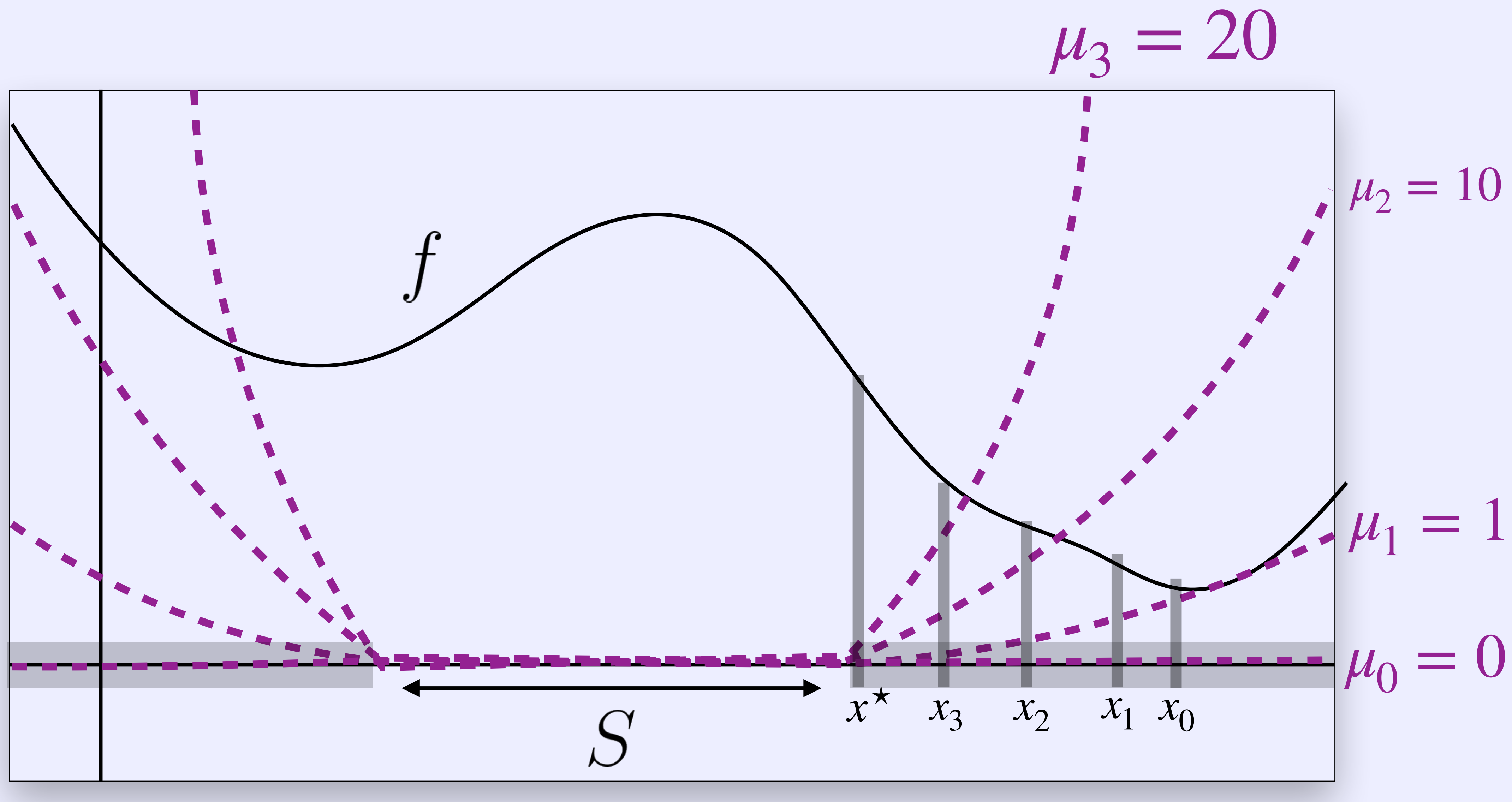
Penalty function

$$P(x) \begin{cases} = 0 \text{ if } x \in S \\ > 0 \text{ if } x \notin S \end{cases}$$

Penalized Problem

$$\min_{x \in \mathbb{R}^d} f(x) + \mu P(x)$$

Any cluster point of the sequence of solutions $(x_k)_{k \geq 0}$ is a solution of the constrained problem.

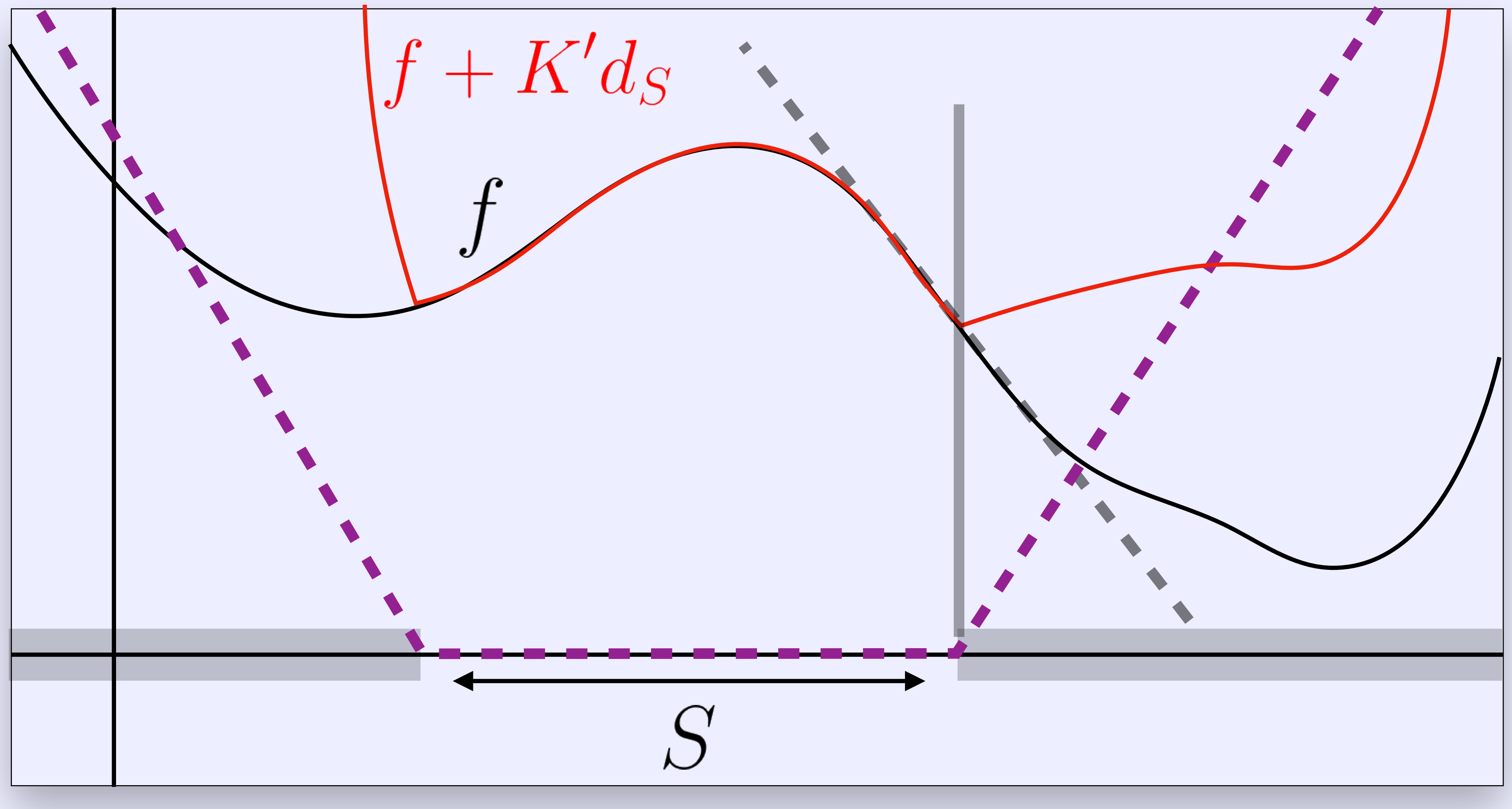


Exact penalization

$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

Assume f to be K-lipschitz on $\mathbb{R}^d$.

Then, for any $K' > K$, this problem has the same set of minimisers as $\min_{x \in \mathbb{R}^d} f(x) + K' d_S(x)$



Recall Penalization on a Picture

The penalization procedure

$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

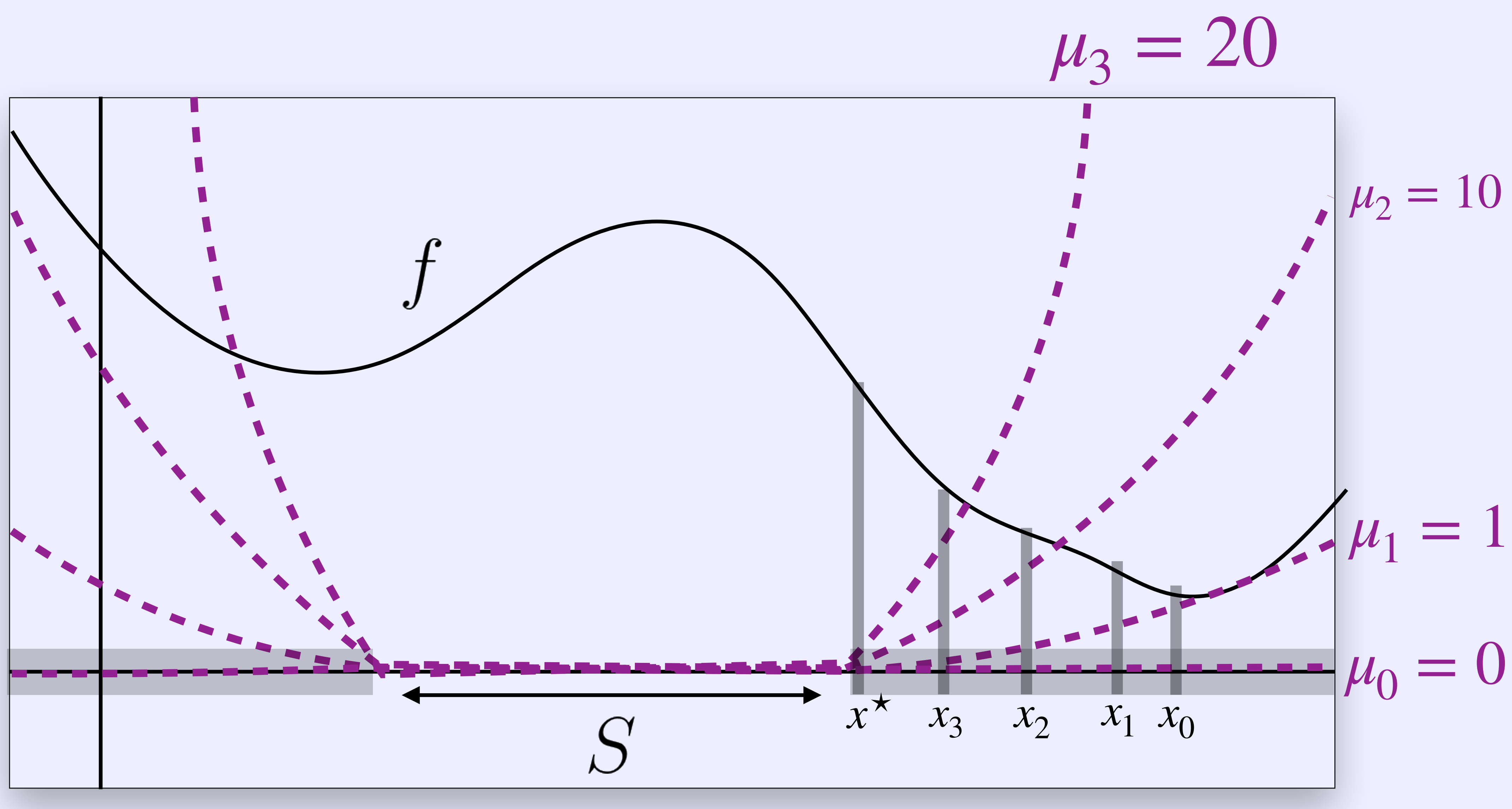
Penalty function

$$P(x) \begin{cases} = 0 \text{ if } x \in S \\ > 0 \text{ if } x \notin S \end{cases}$$

Penalized Problem

$$\min_{x \in \mathbb{R}^d} f(x) + \mu P(x)$$

Any cluster point of the sequence of solutions $(x_k)_{k \geq 0}$ is a solution of the constrained problem.



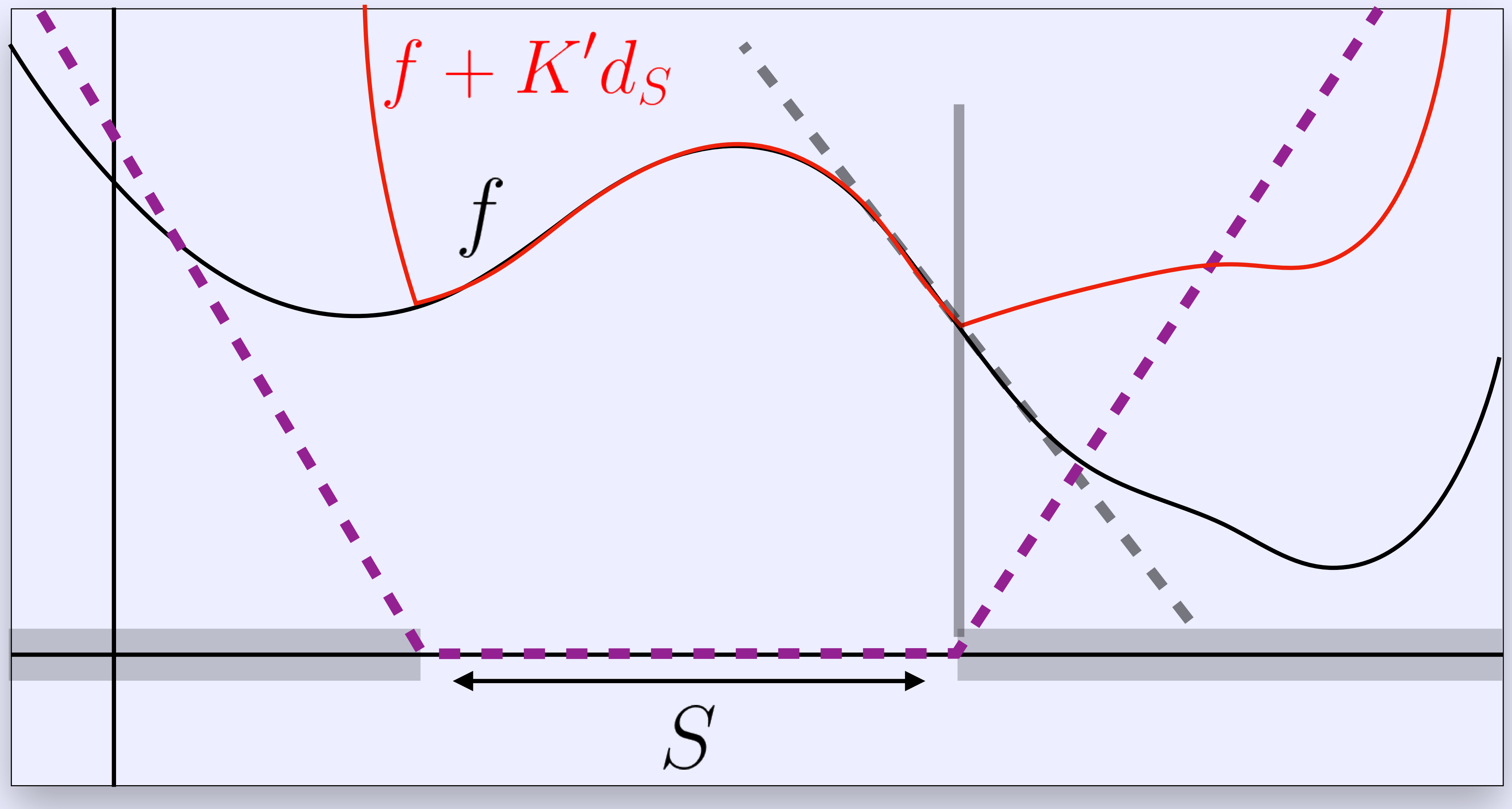
Exact penalization

$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } x \in S$$

Assume f to be K-lipschitz on $\mathbb{R}^d$.

Then, for any $K' > K$, this problem has the same set of minimisers as $\min_{x \in \mathbb{R}^d} f(x) + K' d_S(x)$

Uniform Parametric Error Bound $h(x) \geq \delta\, d_S(x)\ \forall x \in \mathbb{R}^d$, $h(x) = 0 \Leftrightarrow x \in S$



We propose a Double Penalization Procedure

- First penalization

$$(\mathcal{P}) \quad \begin{aligned} \min_{x \in \mathbb{R}^d, \eta \in \mathbb{R}} \quad & f(x) \\ \text{s.t.} \quad & \eta \leq 0 \\ & \eta \in \underset{s \in \mathbb{R}}{\operatorname{argmin}} \; s + \frac{1}{1-p} \mathbb{E}\left[\max(g(x,\xi) - s, 0)\right] \end{aligned}$$

We propose a Double Penalization Procedure

- First penalization

$(\mathcal{P})$

$$\begin{aligned} \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \quad & f(x) \\ \text{s.t.} \quad & \eta \leq 0 \\ & \eta \in \underset{s\,\in\,\mathbb{R}}{\operatorname{argmin}}\; s + \frac{1}{1-p}\mathbb{E}\left[\max(g(x,\xi) - s, 0)\right] \end{aligned}$$

$\Rightarrow$

$(\mathcal{P}_\mu)$

$$\begin{aligned} \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \quad & f(x) + \mu\max(\eta, 0) \\ \text{s.t.} \quad & \eta \in \underset{s\,\in\,\mathbb{R}}{\operatorname{argmin}}\; s + \frac{1}{1-p}\mathbb{E}\left[\max(g(x,\xi) - s, 0)\right] \end{aligned}$$

We propose a Double Penalization Procedure

- First penalization

$(\mathcal{P})$

$$\begin{aligned} \min_{x \in \mathbb{R}^d, \eta \in \mathbb{R}} \quad & f(x) \\ \text{s.t.} \quad & \eta \leq 0 \\ & \eta \in \underset{s \in \mathbb{R}}{\operatorname{argmin}} \; s + \frac{1}{1-p} \mathbb{E}\left[\max(g(x,\xi) - s, 0)\right] \end{aligned}$$

$\Rightarrow$

$(\mathcal{P}_\mu)$

$$\begin{aligned} \min_{x \in \mathbb{R}^d, \eta \in \mathbb{R}} \quad & f(x) + \mu \max(\eta, 0) \\ \text{s.t.} \quad & \eta \in \underset{s \in \mathbb{R}}{\operatorname{argmin}} \; s + \frac{1}{1-p} \mathbb{E}\left[\max(g(x,\xi) - s, 0)\right] \end{aligned}$$

- In practice, the constant μ is a hyperparameter to tune.

We propose a Double Penalization Procedure

- First penalization

$(\mathcal{P})$

$$\begin{aligned} \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \quad & f(x) \\ \text{s.t.} \quad & \eta \leq 0 \\ & \eta \in \underset{s\,\in\,\mathbb{R}}{\operatorname{argmin}}\ s + \frac{1}{1-p}\mathbb{E}\left[\max(g(x,\xi)-s,0)\right] \end{aligned}$$

$\Rightarrow$

$(\mathcal{P}_\mu)$

$$\begin{aligned} \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \quad & f(x) + \mu\max(\eta,0) \\ \text{s.t.} \quad & \eta \in \underset{s\,\in\,\mathbb{R}}{\operatorname{argmin}}\ \underbrace{s + \frac{1}{1-p}\mathbb{E}\left[\max(g(x,\xi)-s,0)\right]}_{=G(x,s)} \end{aligned}$$

$$\geq \bar{Q}_p(g(x,\xi))$$

- In practice, the constant μ is a hyperparameter to tune.

- Using Rockafellar property

$$\begin{aligned} \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \quad & f(x) + \mu\max(\eta,0) \\ \text{s.t.} \quad & G(x,\eta) - \bar{Q}_p(g(x,\xi)) \leq 0 \end{aligned}$$

We propose a Double Penalization Procedure

- Second penalization

$$(\mathcal{P}_\mu) \quad \begin{aligned} \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \quad & f(x) + \mu \max(\eta, 0) \\ \text{s.t.} \quad & G(x,\eta) - \bar{Q}_p(g(x,\xi)) \leq 0 \end{aligned}$$

$$\Rightarrow$$

$$(\mathcal{P}_{\lambda,\mu}) \quad \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} f(x) + \mu \max(\eta, 0) + \lambda \left(G(x,\eta) - \bar{Q}_p(g(x,\xi))\right)$$

We propose a Double Penalization Procedure

- Second penalization

$$(\mathcal{P}_\mu) \quad \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \; f(x) + \mu\max(\eta,0) \quad \text{s.t.} \quad G(x,\eta) - \bar{Q}_p(g(x,\xi)) \le 0$$

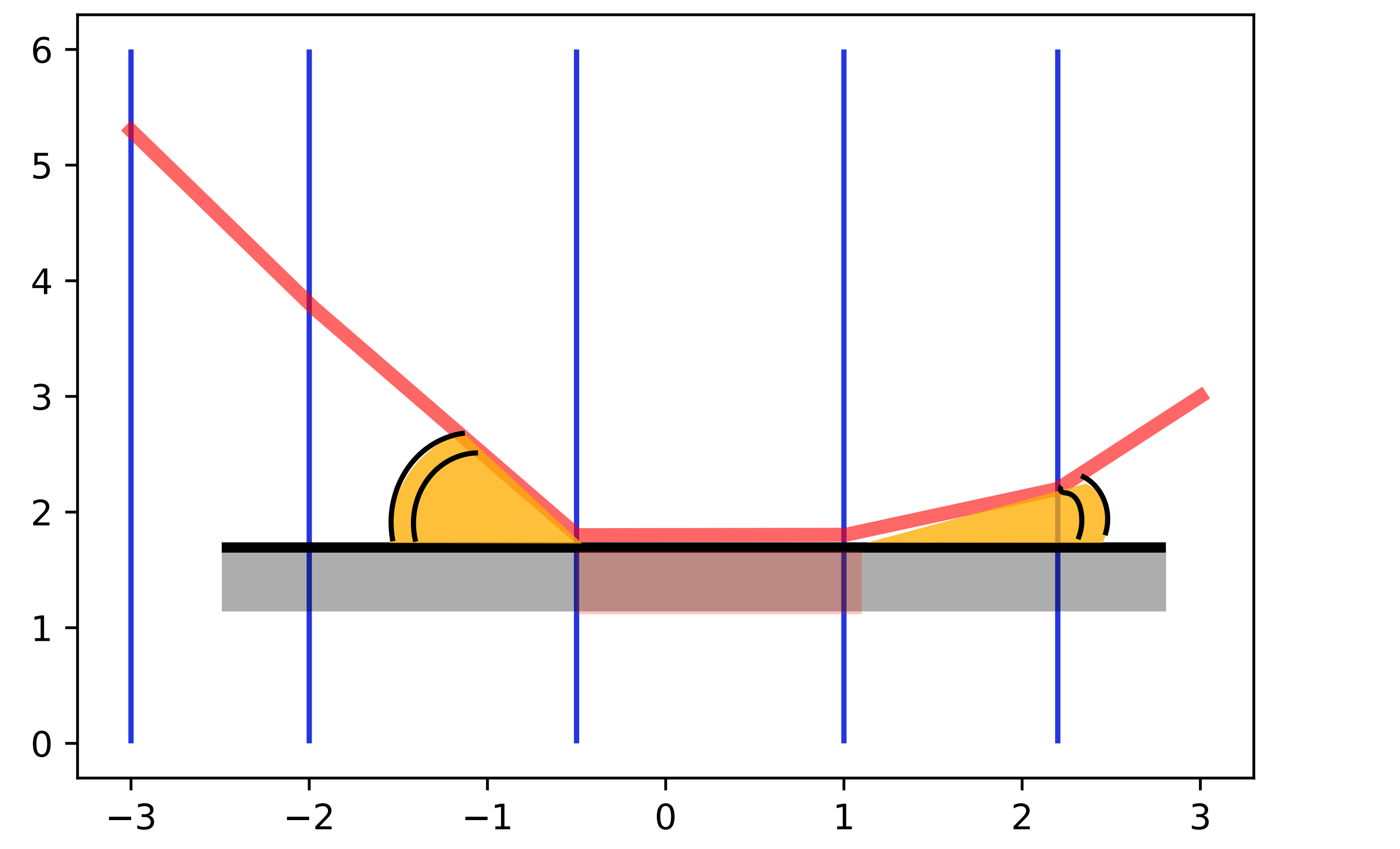
$$\Rightarrow \quad (\mathcal{P}_{\lambda,\mu}) \quad \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \; f(x) + \mu\max(\eta,0) + \lambda\left(G(x,\eta) - \bar{Q}_p(g(x,\xi))\right)$$

- This penalization is exact.

> **Theorem** Let $\mu > 0$ be given and fixed and assume that the solution set of problem $(\mathcal{P}_\mu)$ is not empty. Then for any $\lambda > \lambda_\mu = \frac{\mu}{\delta}$ where:
>
> $$\delta = \begin{cases} \frac{1}{n(1-p)} & \text{if } p \in \mathcal{I} \\ \frac{d_{\mathcal{I}}(p)}{1-p} & \text{otherwise.} \end{cases}$$
>
> the solution set of $(\mathcal{P}_\mu)$ coincides with the solution set of $(\mathcal{P}_{\lambda,\mu})$

Note: $\mathcal{I} := \left\{\frac{i}{n}, 1 \le i \le n\right\}, \quad d_{\mathcal{I}}(p) := \text{distance}(p, \mathcal{I})$

$$\eta \mapsto G(x,\eta) = \eta + \frac{1}{1-p}\mathbb{E}[\max(g(x,\xi)-\eta,0)]$$



Solving of the doubly-penalized problem

- Second penalization

$(\mathcal{P}_\mu)$

$$\min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \quad f(x) + \mu \max(\eta, 0)$$
$$\text{s.t.} \quad G(x,\eta) - \bar{Q}_p(g(x,\xi)) \leq 0$$

$\Rightarrow$

$(\mathcal{P}_{\lambda,\mu})$

$$\min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \underbrace{f(x) + \mu \max(\eta, 0)}_{\text{Convex}} + \underbrace{\lambda \left(G(x,\eta) - \bar{Q}_p(g(x,\xi))\right)}_{\text{Non convex but...}}$$

Convex

Non convex but...

Solving of the doubly-penalized problem

- Second penalization

$(\mathcal{P}_\mu)$

$$\begin{aligned} \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \quad & f(x) + \mu \max(\eta, 0) \\ \text{s.t.} \quad & G(x,\eta) - \bar{Q}_p(g(x,\xi)) \leq 0 \end{aligned}$$

$(\mathcal{P}_{\lambda,\mu})$

$$\min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \underbrace{f(x) + \mu \max(\eta, 0)}_{\text{Convex}} + \lambda \underbrace{\left(G(x,\eta) - \bar{Q}_p(g(x,\xi))\right)}_{\text{Non convex}}$$

Convex

Non convex

but...

$$G(x,\eta) = \eta + \frac{1}{1-p}\mathbb{E}[\max(g(x,\xi) - \eta, 0)]$$

Solving of the doubly-penalized problem

- Second penalization

$$(\mathcal{P}_\mu) \quad \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \ f(x) + \mu\max(\eta, 0) \quad \text{s.t.} \quad G(x,\eta) - \bar{Q}_p(g(x,\xi)) \leq 0$$

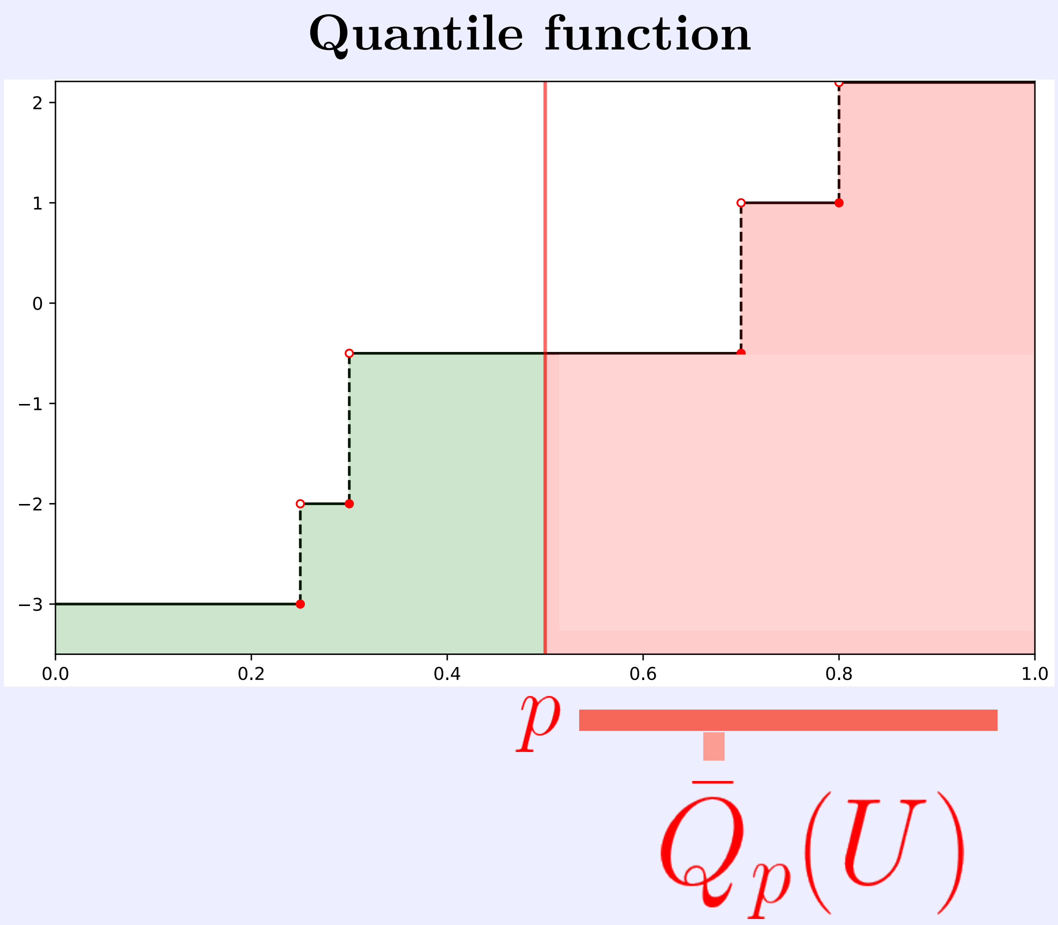
$$(\mathcal{P}_{\lambda,\mu}) \quad \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \ f(x) + \mu\max(\eta,0) + \lambda\left(G(x,\eta) - \bar{Q}_p(g(x,\xi))\right)$$

Convex **Convex**

$$G(x,\eta) = \eta + \frac{1}{1-p}\mathbb{E}[\max(g(x,\xi)-\eta, 0)]$$

- $\bar{Q}_p$ is convex ! $(\mathcal{P}_{\lambda,\mu})$ is a Difference of Convex problem.

$$\bar{Q}_p(U) = \frac{1}{1-p}\int_{p'=p}^{1} Q_{p'}(U)dp'$$



Solving of the doubly-penalized problem

- Second penalization

$(\mathcal{P}_\mu)$

$$\begin{aligned} \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \quad & f(x) + \mu \max(\eta, 0) \\ \text{s.t.} \quad & G(x,\eta) - \bar{Q}_p(g(x,\xi)) \leq 0 \end{aligned}$$

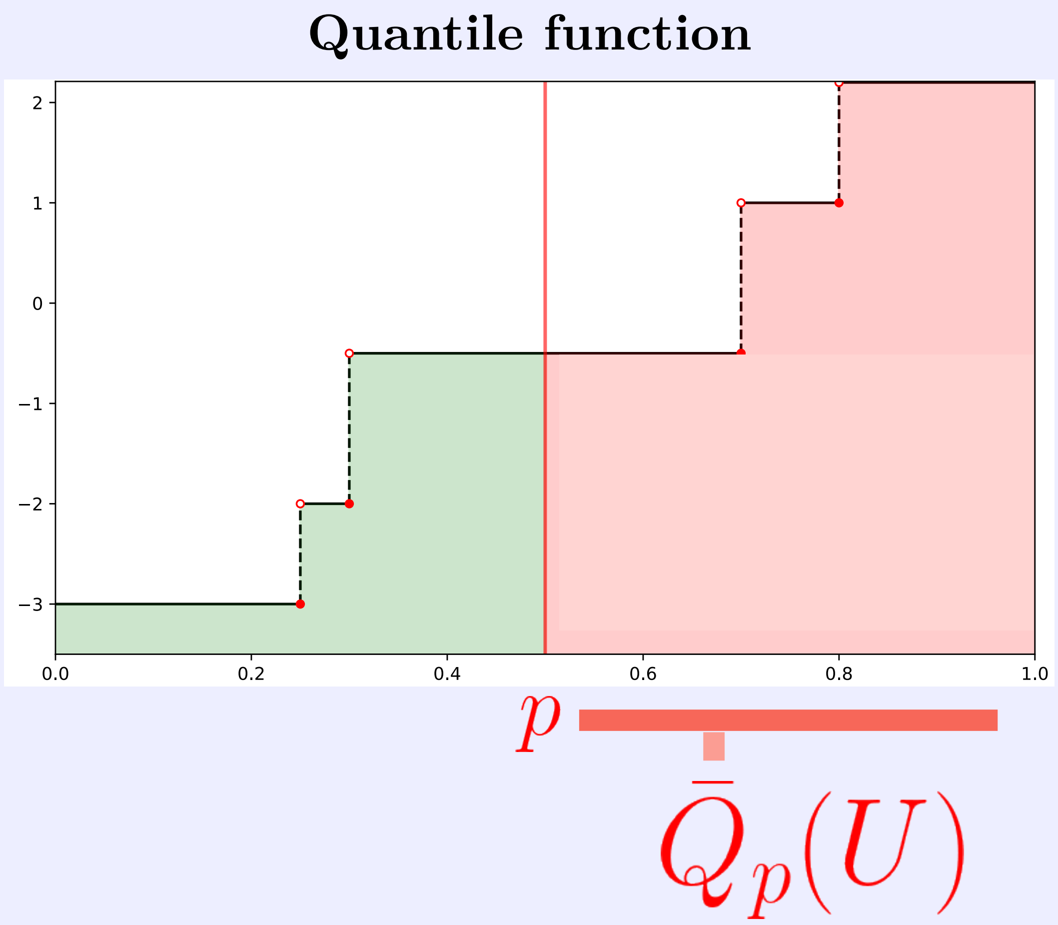
$(\mathcal{P}_{\lambda,\mu})$

$$\min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \underbrace{f(x) + \mu \max(\eta, 0) + \lambda \left(G(x,\eta)\right.}_{\text{Convex}} - \underbrace{\left.\bar{Q}_p(g(x,\xi))\right)}_{\text{Convex}}$$

$$G(x,\eta) = \eta + \frac{1}{1-p}\mathbb{E}[\max(g(x,\xi) - \eta, 0)]$$

- $\bar{Q}_p$ is convex ! $(\mathcal{P}_{\lambda,\mu})$ is a Difference of Convex problem.

$$\bar{Q}_p(U) = \frac{1}{1-p}\int_{p'=p}^{1} Q_{p'}(U)dp' = \sup_{\substack{q\in\mathbb{R}^n \\ \sum_{i=1}^n q_i = 1 \\ 0\leq q_i \leq \frac{1}{n(1-p)}}} \sum_{i=1}^{n} q_i U_i$$

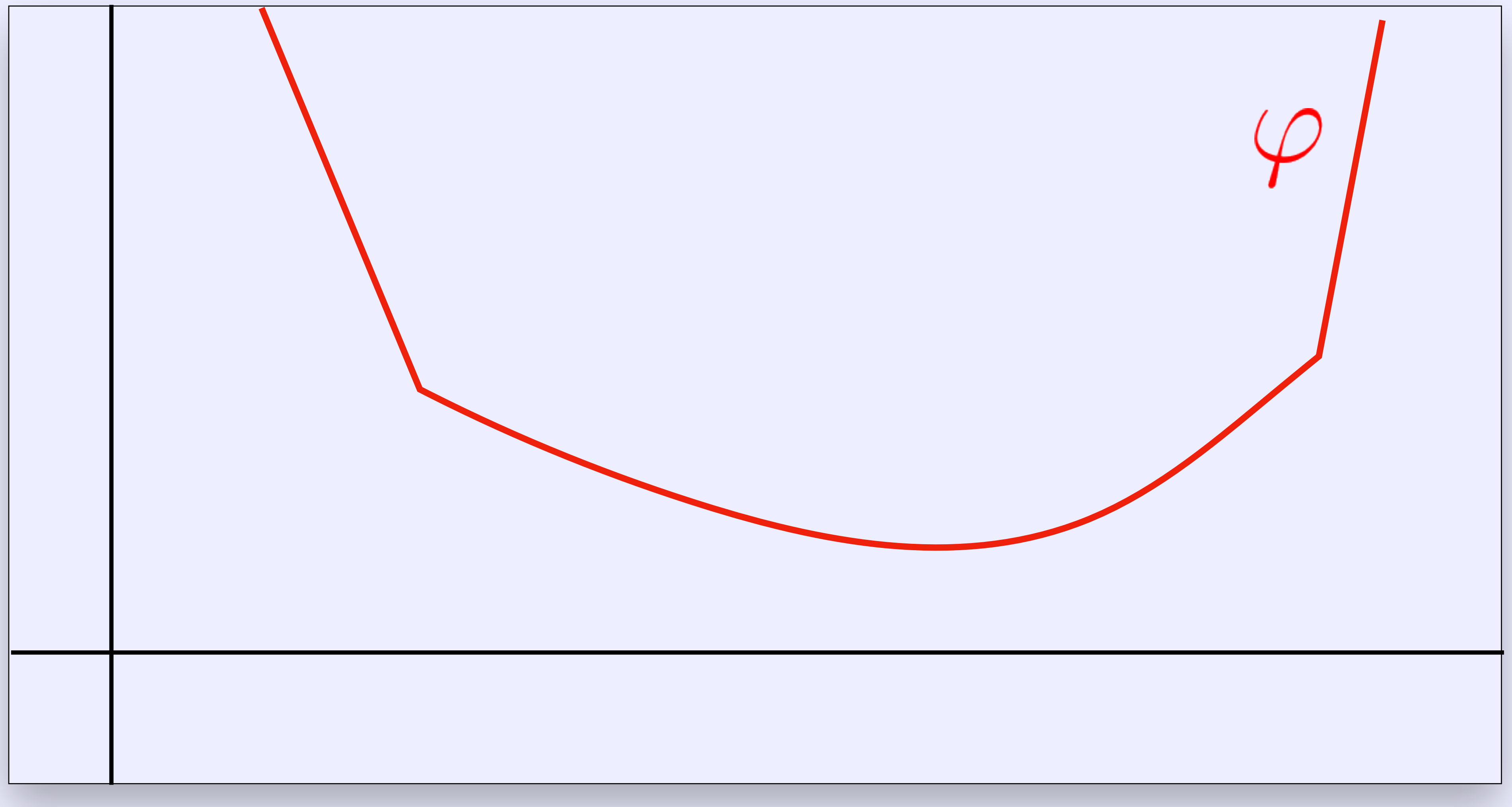


3 TACO

A Python Toolbox for Chance Constrained Problems

Recall: Solving DC programs by Bundle

- Bundle methods in a nutshell
  - Minimization of non-smooth problems



Recall: Solving DC programs by Bundle

- Bundle methods in a nutshell
  - Minimization of non-smooth problems
  - Maintains:
    - the **Bundle Information.**
    - the **Polyhedral Approximation.**

$$x \mapsto \varphi(x_i) + g_\varphi^\top (x - x_i)$$

$$\check{\varphi}(x) = \max_{i \in \text{Bundle}} \varphi(x_i) + g_\varphi^\top (x - x_i)$$

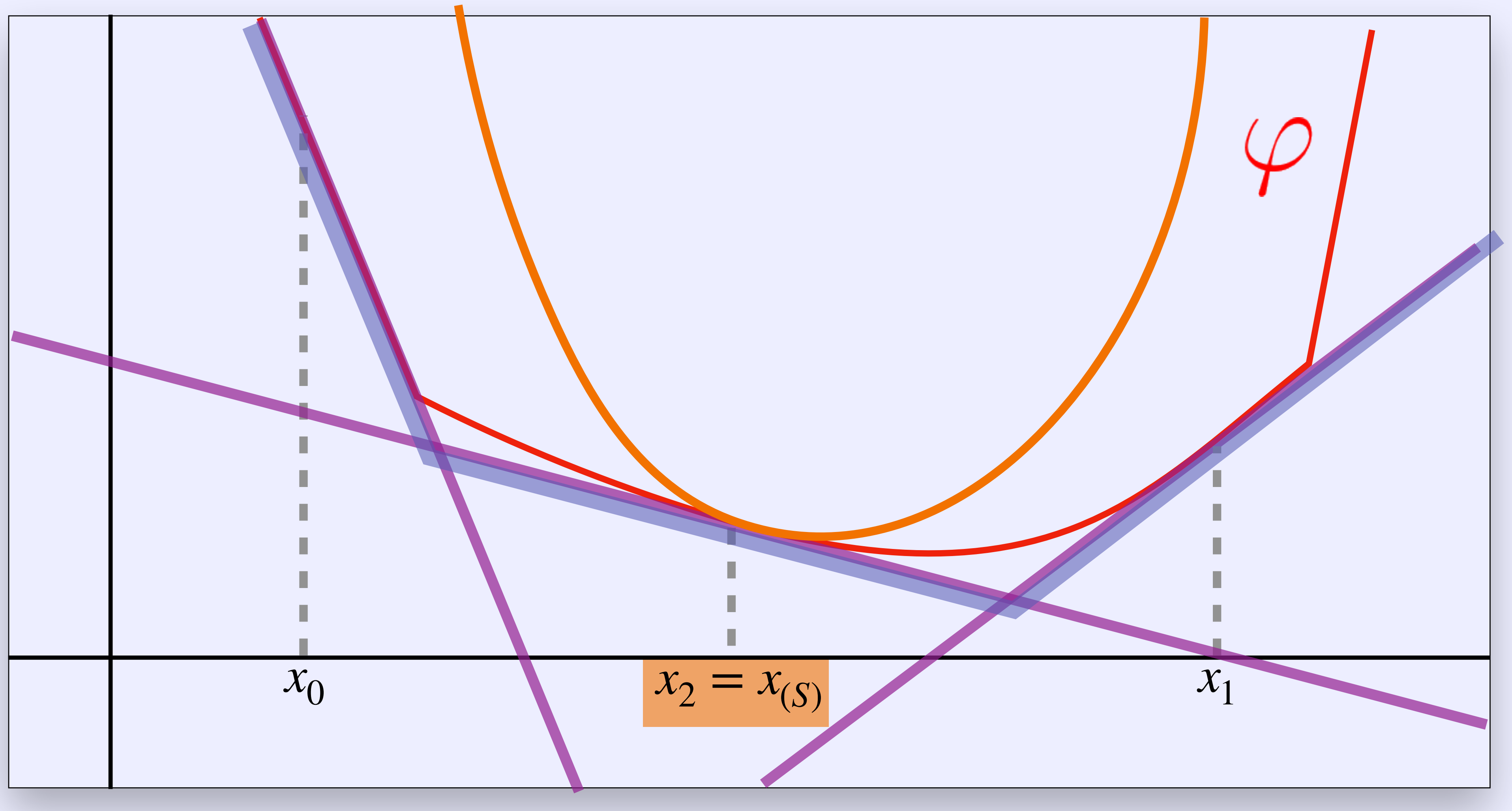
Recall: Solving DC programs by Bundle

- Bundle methods in a nutshell
  - Minimization of non-smooth problems
  - Maintains:
    - the **Bundle Information.**
    - the **Polyhedral Approximation.**
    - the **Stability Center** in the bundle.

$$x \mapsto \varphi(x_i) + g_\varphi^\top (x - x_i)$$

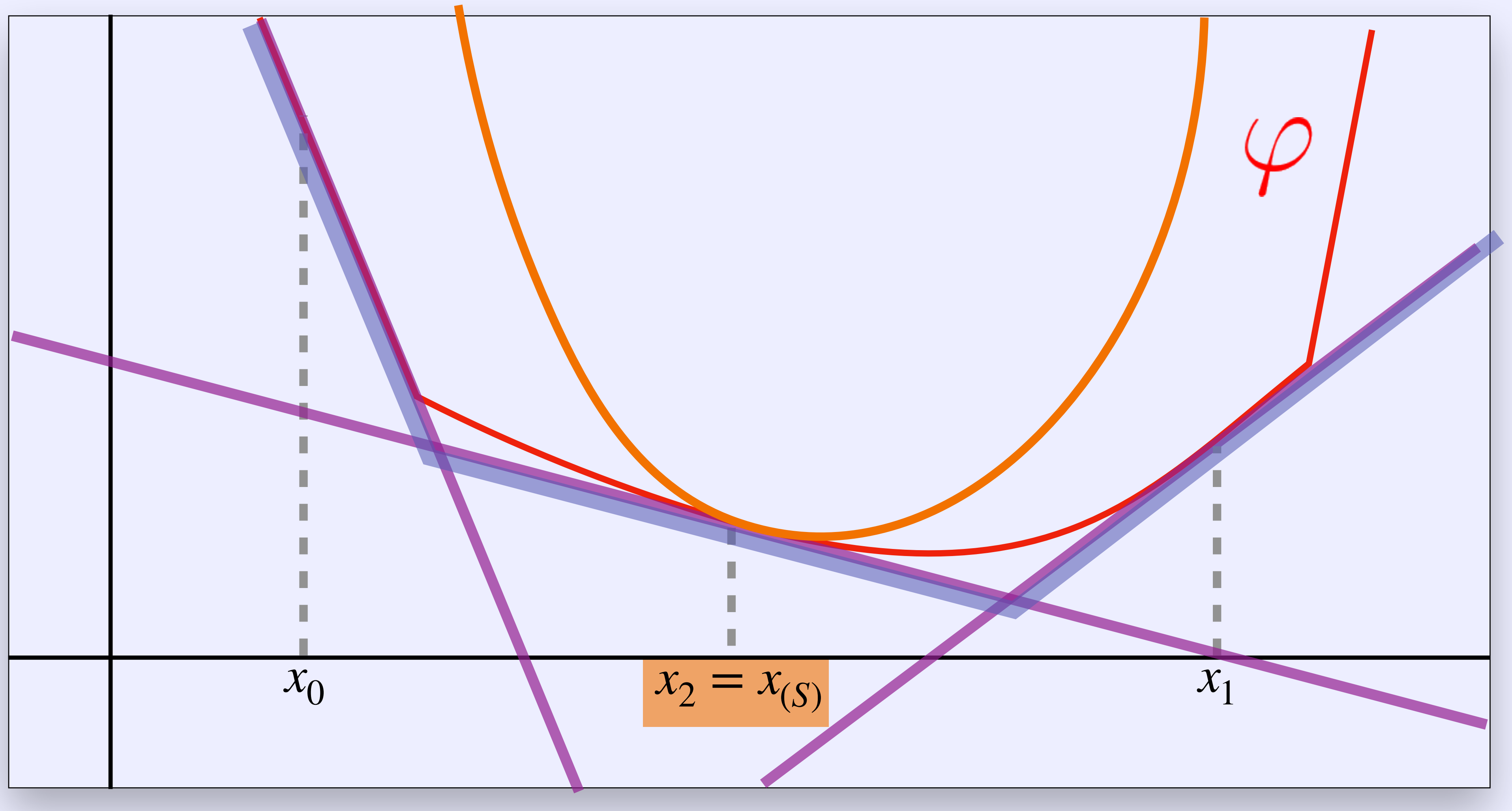
$$\check{\varphi}(x) = \max_{i \in \text{Bundle}} \varphi(x_i) + g_\varphi^\top (x - x_i)$$

$$\min_{x \in \mathbb{R}^d} \check{\varphi}(x) + \alpha \|x - x_{(S)}\|_2^2$$



Recall: Solving DC programs by Bundle

- Bundle methods in a nutshell
  - Minimization of non-smooth problems
  - Maintains:
    - the **Bundle Information.**
    - the **Polyhedral Approximation.**
    - the **Stability Center** in the bundle.



- State-of-the-art methods for DC problems

[De Oliveira 19']

  - Function of the form $\varphi(x) = \varphi_1(x) - \varphi_2(x)$
  - We now solve at each iteration :

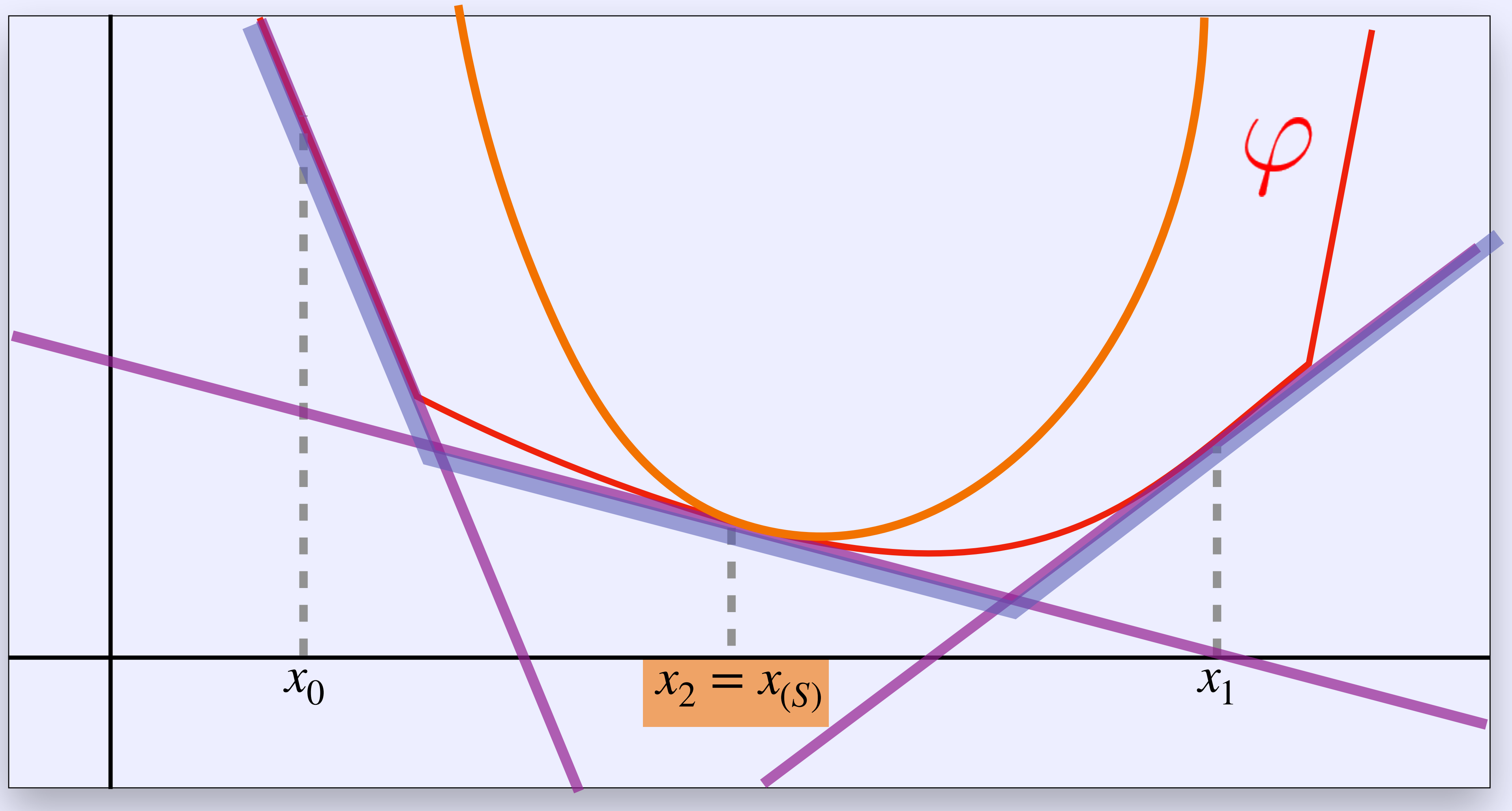
$$\min_{x \in \mathbb{R}^d} \check{\varphi}_1(x) - g_{\varphi_2}^\top (x - x_{(S)}) + \alpha \|x - x_{(S)}\|_2^2$$

  - Update rule for the stability center:

$$\varphi(x_{k+1}) \leq \varphi(x_{(s)}) - \beta \|x_{k+1} - x_{(S)}\|^2$$

Recall: Solving DC programs by Bundle

- Bundle methods in a nutshell
  - Minimization of non-smooth problems
  - Maintains:
    - the **Bundle Information.**
    - the **Polyhedral Approximation.**
    - the **Stability Center** in the bundle.



- State-of-the-art methods for DC problems

[De Oliveira 19']

  - Function of the form $\varphi(x) = \varphi_1(x) - \varphi_2(x)$
  - We now solve at each iteration :

$$\min_{x \in \mathbb{R}^d} \check{\varphi}_1(x) - g_{\varphi_2}^\top (x - x_{(S)}) + \alpha \|x - x_{(S)}\|_2^2$$

  - Update rule for the stability center:

$$\varphi(x_{k+1}) \leq \varphi(x_{(s)}) - \beta \|x_{k+1} - x_{(S)}\|^2$$

- Convergence property

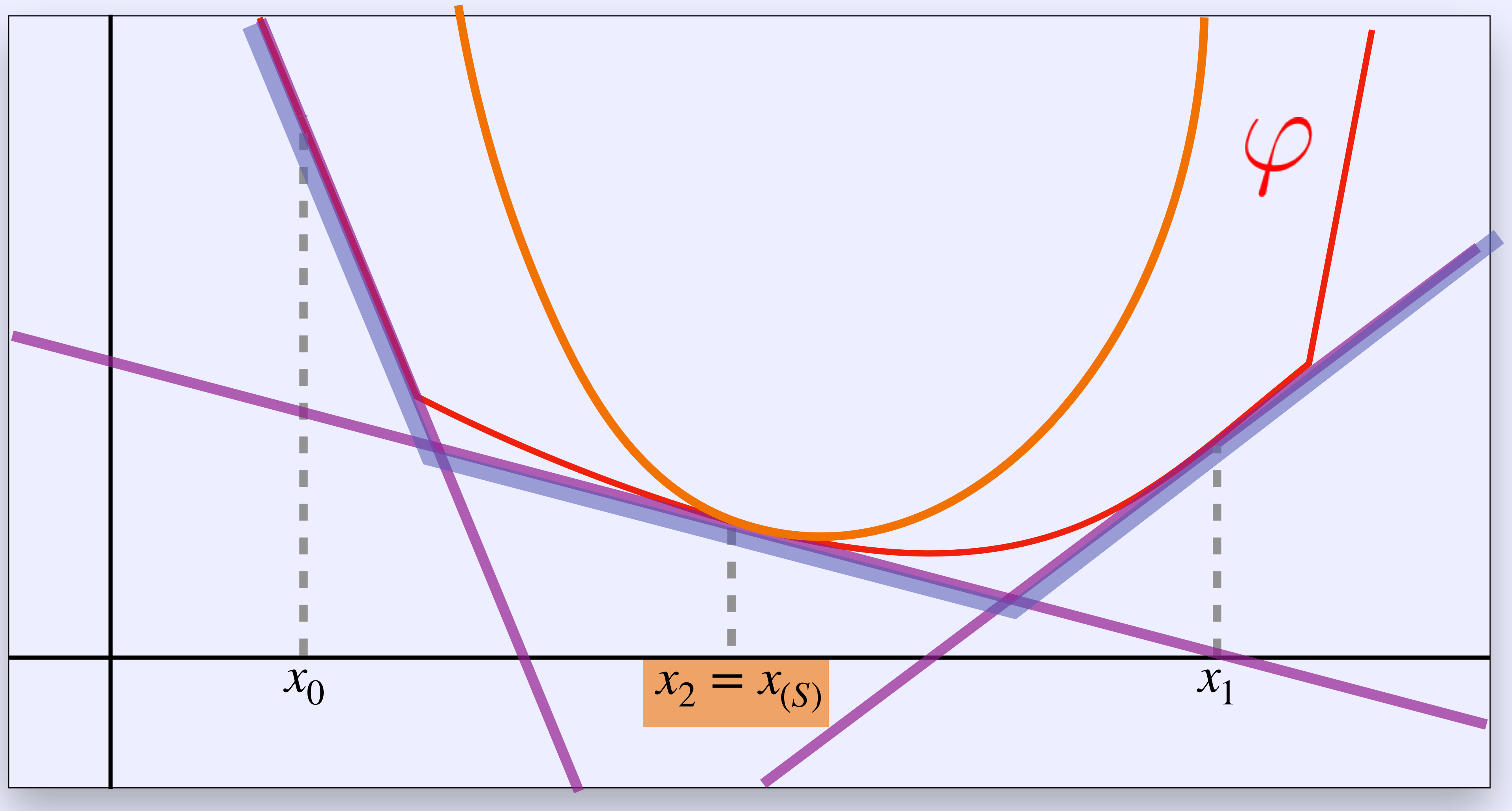
Global → Local → d-stationary → Critical

d-stationary: $\partial\varphi_2(x) \subset \partial\varphi_1(x)$

Critical: $\partial\varphi_2(x) \cap \partial\varphi_1(x) \neq \emptyset$

Recall: Solving DC programs by Bundle

- Bundle methods in a nutshell
  - Minimization of non-smooth problems
  - Maintains:
    - the **Bundle Information.**
    - the **Polyhedral Approximation.**
    - the **Stability Center** in the bundle.



- State-of-the-art methods for DC problems

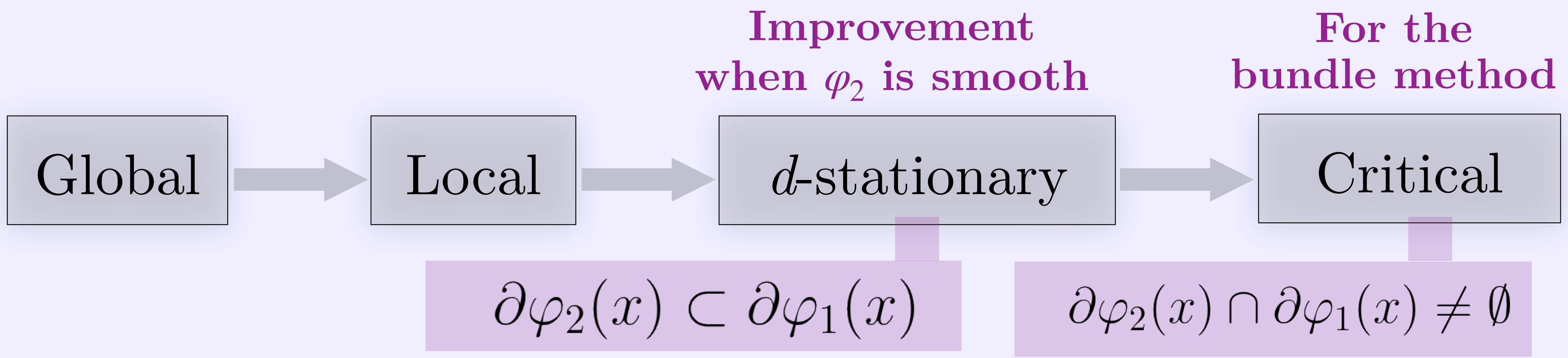
  [De Oliveira 19']
  - Function of the form $\varphi(x) = \varphi_1(x) - \varphi_2(x)$
  - We now solve at each iteration :

$$\min_{x \in \mathbb{R}^d} \check{\varphi}_1(x) - g_{\varphi_2}^\top (x - x_{(S)}) + \alpha \|x - x_{(S)}\|_2^2$$

  - Update rule for the stability center:

$$\varphi(x_{k+1}) \leq \varphi(x_{(s)}) - \beta \|x_{k+1} - x_{(S)}\|^2$$

- Convergence property



For Our DC Problem

- The DC problem

$$(\mathcal{P}_{\lambda,\mu}) \quad \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \overbrace{f(x) + \mu \max(\eta, 0) + \lambda \left(G(x,\eta)\right.}^{\varphi_1} - \left.\overbrace{\bar{Q}_p(g(x,\xi))}^{\varphi_2}\right)$$

For Our DC Problem

- The DC problem

$$(\mathcal{P}_{\lambda,\mu}) \quad \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \overbrace{f(x) + \mu\max(\eta,0) + \lambda\big(G(x,\eta)}^{\varphi_1} - \overbrace{\bar{Q}_p(g(x,\xi))}^{\varphi_2}\big)$$

- Smoothing of the superquantile [L., Malick, Harchaoui 20']

  - Smoothing of f_2 based on Nesterov's technique.

$$\bar{Q}_p(U) = \sup_{\substack{q\in\mathbb{R}^n \\ \sum_{i=1}^n q_i = 1 \\ 0\le q_i\le \frac{1}{n(1-p)}}} \sum_{i=1}^n q_i U_i$$

For Our DC Problem

- The DC problem

$$(\mathcal{P}_{\lambda,\mu}) \quad \min_{x\in\mathbb{R}^d,\eta\in\mathbb{R}} \overbrace{f(x) + \mu\max(\eta,0) + \lambda\big(G(x,\eta)}^{\varphi_1} - \overbrace{\bar{Q}_p(g(x,\xi))}^{\varphi_2}\big)$$

- Smoothing of the superquantile [L., Malick, Harchaoui 20']

  - Smoothing of f_2 based on Nesterov's technique.

$$\bar{Q}_p(U) \simeq \sup_{\substack{q\in\mathbb{R}^n \\ \sum_{i=1}^n q_i = 1 \\ 0\le q_i \le \frac{1}{n(1-p)}}} \sum_{i=1}^n q_i U_i - \frac{\alpha}{2}\|q - \frac{1}{n}(1,\ldots,1)^\top\|^2$$

What a long process !

- To sum up

CHANCE CONSTRAINED PROBLEM → BILEVEL PROGRAM → DOUBLE PENALIZATION → BUNDLE METHOD

- What about the implementation ?

What a long process !

- To sum up

CHANCE CONSTRAINED PROBLEM → BILEVEL PROGRAM → DOUBLE PENALIZATION → BUNDLE METHOD

- What about the implementation ?

TACO : a Toolbox for chAnce Constrained Optimization

- Goal : solve a problem of the form
$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

- Input : the class `Problem`
  - First-order oracles for f and g.
  - A sampled dataset for the values of ξ.
  - A python dictionary of parameters.

Example : Kataoka's Example

In[1]:

```
import numpy as np

class Kataoka:

    def __init__(self, nb_samples=10000, nb_features=2, seed=42):

        np.random.seed(seed)
        mean = np.array([1.0, 1.0])
        cov =  np.eye(2)
        self.data = np.random.multivariate_normal(mean, cov,
size=self.nb_samples)

    def objective_func(self, x):
         return 0.5*np.dot(x,x)

    def objective_grad(self,x):
        return x

    def constraint_func(self, x, z):
        return np.dot(x,z)

    def constraint_grad(self, x, z):
        return z
```

TACO : a Toolbox for chAnce Constrained Optimization

- Goal : solve a problem of the form

$$\min_{x \in \mathbb{R}^d} f(x)$$
$$\text{s.t. } \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

- Input : the class `Problem`
  - First-order oracles for f and g.
  - A sampled dataset for the values of ξ.
  - A python dictionary of parameters.
- The class `Optimizer`
  - Instantiate with the inputs.
  - Optimization launched with the method `run`.

Example : Kataoka's Example

In[2]:
```
optimizer = Optimizer(problem, params=params)
optimiser.run()
```

TACO : a Toolbox for chAnce Constrained Optimization

- Goal : solve a problem of the form
$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } \mathbb{P}[g(x,\xi) \leq 0] \geq p$$
- Input : the class `Problem`
  - First-order oracles for f and g.
  - A sampled dataset for the values of ξ.
  - A python dictionary of parameters.
- The class `Optimizer`
  - Instantiate with the inputs.
  - Optimization launched with the method `run`.
- Output
  - Retrieved from the `Optimizer` class.

Example : Kataoka's Example

In[2]:
```
optimizer = Optimizer(problem, params=params)
optimiser.run()
```

In[3]:
```
sol = optimizer.solution
```

TACO : a Toolbox for chAnce Constrained Optimization

- Goal : solve a problem of the form $\min_{x \in \mathbb{R}^d} f(x)$ s.t. $\mathbb{P}[g(x,\xi) \leq 0] \geq p$
- Input : the class `Problem`
  - First-order oracles for f and g.
  - A sampled dataset for the values of ξ.
  - A python dictionary of parameters.
- The class `Optimizer`
  - Instantiate with the inputs.
  - Optimization launched with the method `run`.
- Output
  - Retrieved from the `Optimizer` class.

Example : Kataoka's Example

In[2]:
```
optimizer = Optimizer(problem, params=params)
optimiser.run()
```

In[3]:
```
sol = optimizer.solution
```

- Hyperparameters
  - Probability threshold p
  - Penalization parameters μ, λ
  - Number of iterations, starting point, target precision, etc.

4 Numerical Illustrations

Proof of concept on a quadratic Chance constraint Problem

- 2D quadratic problem

$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t. } \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

$$f(x) = (x-c)^\top A(x-c)$$
$$g(x,z) = z^\top W(x)^\top z + p^\top z + b$$
$$\xi \sim \mathcal{N}(\mu, \Sigma)$$

Proof of concept on a quadratic Chance constraint Problem

- 2D quadratic problem

$$\min_{x\in\mathbb{R}^d} f(x) \quad \text{s.t. } \mathbb{P}[g(x,\xi)\le 0] \ge p$$

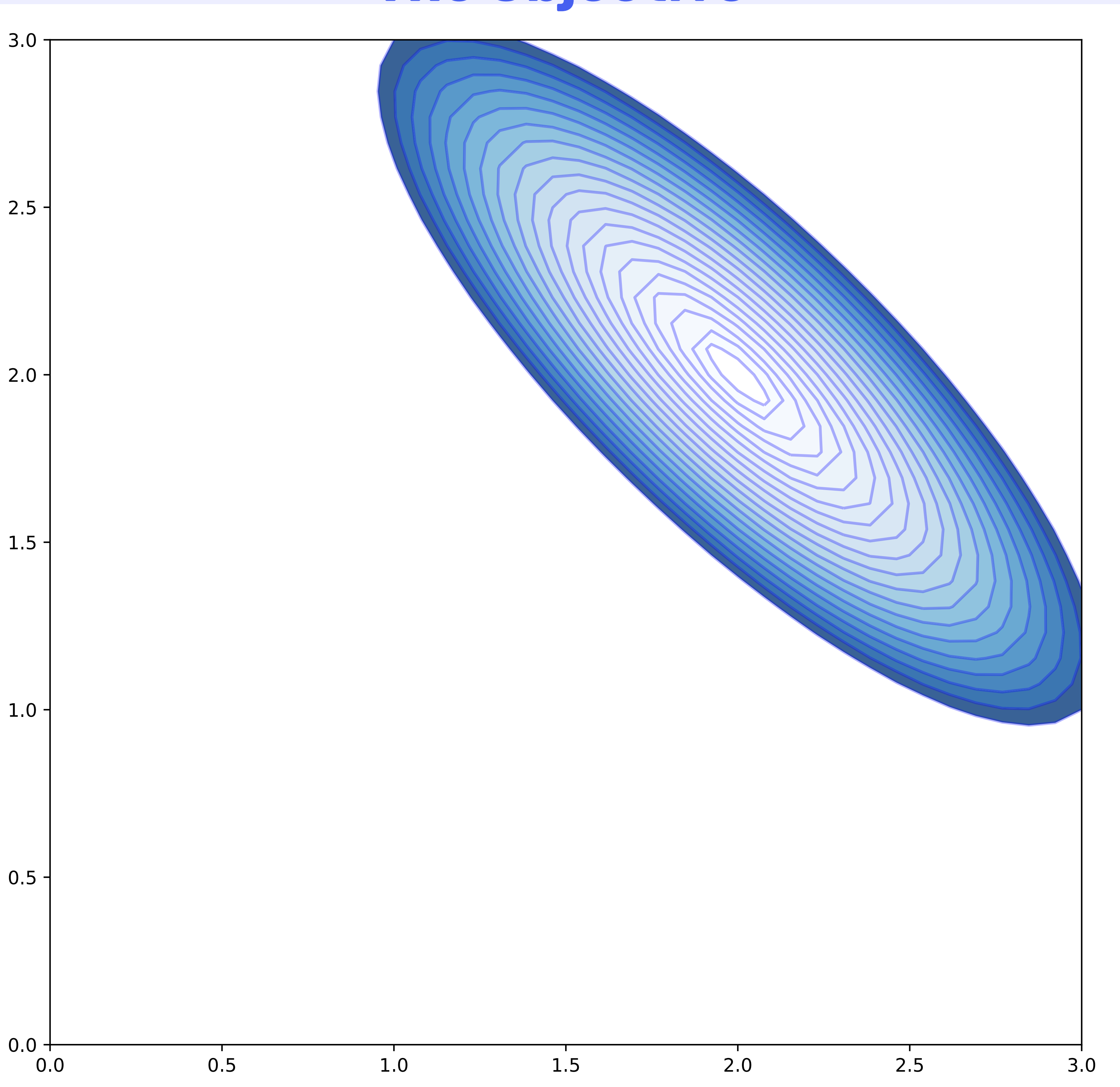
$$f(x) = (x-c)^\top A(x-c)$$

$$g(x,z) = z^\top W(x)^\top z + p^\top z + b$$

$$\xi \sim \mathcal{N}(\mu,\Sigma)$$

$$c = \begin{pmatrix} 2. \\ 2. \end{pmatrix} \quad A = \begin{pmatrix} 5.5 & 4.5 \\ 4.5 & 5.5 \end{pmatrix}$$

The Objective



Proof of concept on a quadratic Chance constraint Problem

- 2D quadratic problem

$$\min_{x\in\mathbb{R}^d} f(x) \qquad f(x) = (x-c)^\top A(x-c)$$

$$\text{s.t. } \mathbb{P}[g(x,\xi)\leq 0]\geq p \qquad g(x,z) = z^\top W(x)^\top z + q^\top z + r$$

$$\xi \sim \mathcal{N}(\mu,\Sigma)$$

$$c = \begin{pmatrix} 2. \\ 2. \end{pmatrix} \quad A = \begin{pmatrix} 5.5 & 4.5 \\ 4.5 & 5.5 \end{pmatrix}$$

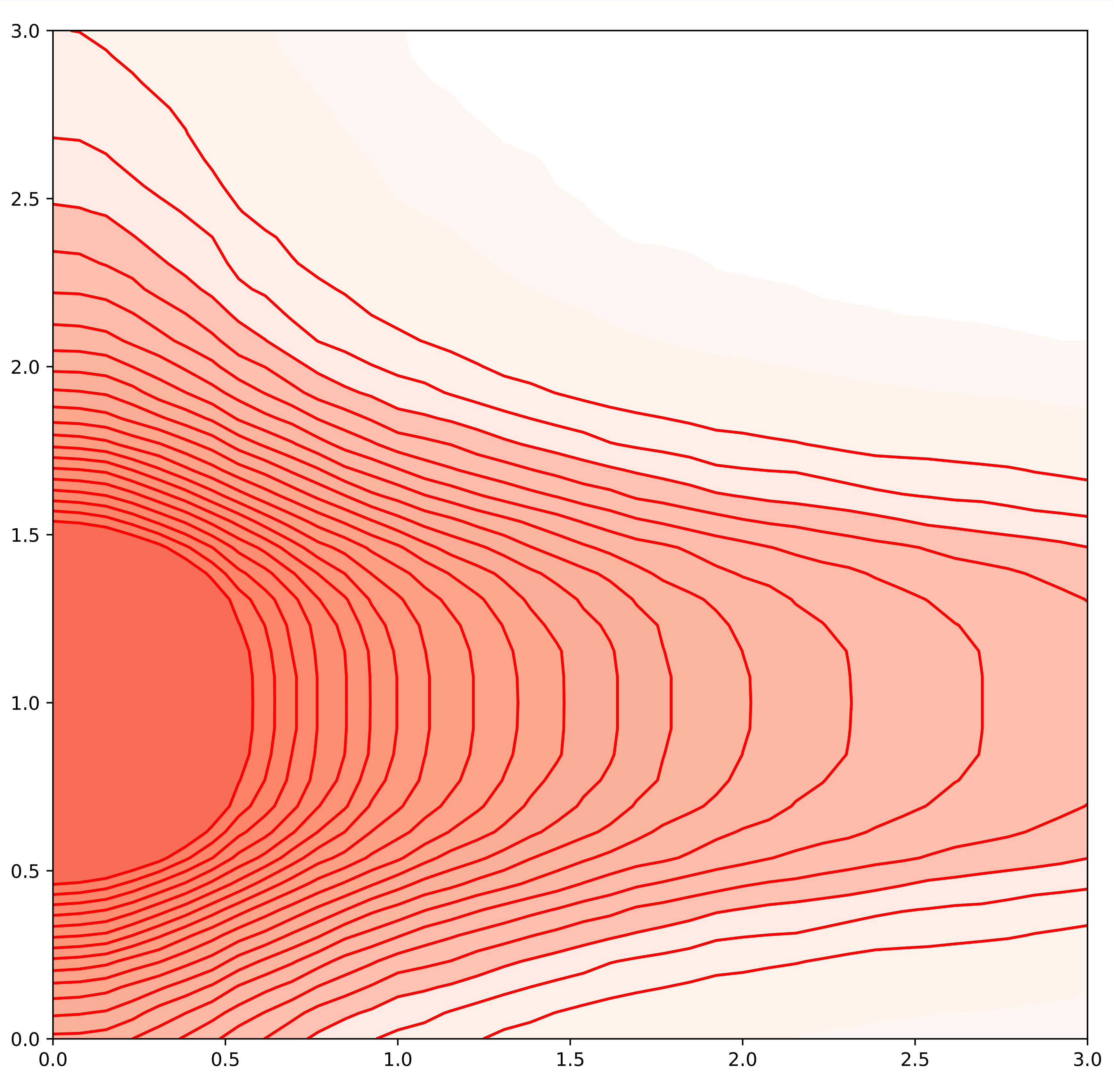
$$W : x = (x_1, x_2)^\top \mapsto \begin{pmatrix} x_1^2 + 0.5 & 0.0 \\ 0.0 & |x_2 - 1|^3 + 1. \end{pmatrix}$$

$$q = \begin{pmatrix} 1. \\ 1. \end{pmatrix}, \; r = -1$$

ξ is sampled 10000 times with parameters $\mu = \begin{pmatrix} 1. \\ 1. \end{pmatrix}$

$$\Sigma = \begin{pmatrix} 20. & 0. \\ 0. & 20. \end{pmatrix}$$

The Chance Constraint



Proof of concept on a quadratic Chance constraint Problem

- 2D quadratic problem

$$\min_{x \in \mathbb{R}^d} f(x) \qquad f(x) = (x-c)^\top A(x-c)$$

$$\text{s.t. } \mathbb{P}[g(x,\xi) \le 0] \ge p \qquad g(x,z) = z^\top W(x)^\top z + q^\top z + r$$

$$\xi \sim \mathcal{N}(\mu, \Sigma)$$

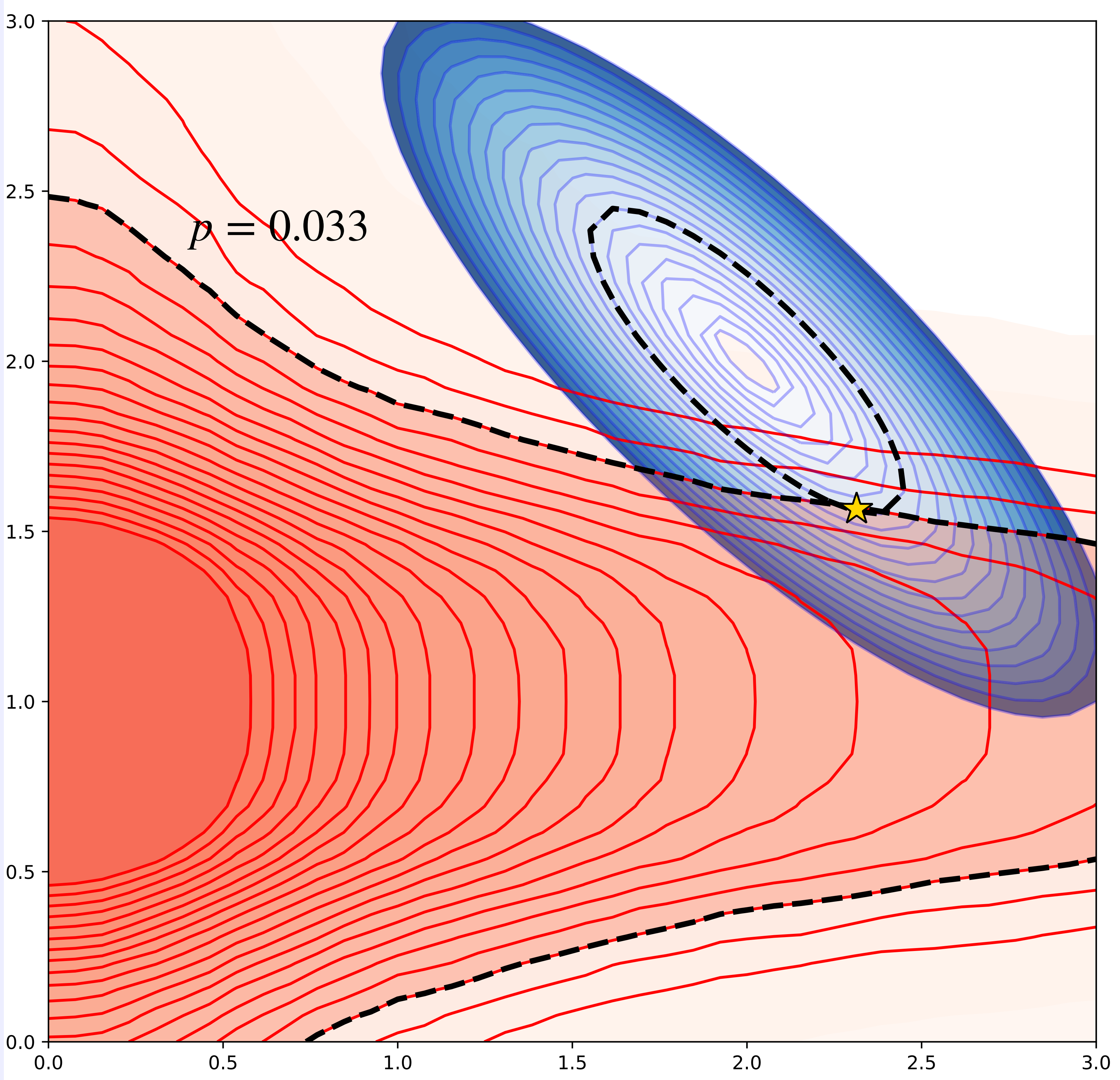
$$c = \begin{pmatrix} 2. \\ 2. \end{pmatrix} \quad A = \begin{pmatrix} 5.5 & 4.5 \\ 4.5 & 5.5 \end{pmatrix}$$

$$W : x = (x_1, x_2)^\top \mapsto \begin{pmatrix} x_1^2 + 0.5 & 0.0 \\ 0.0 & |x_2 - 1|^3 + 1. \end{pmatrix}$$

$$q = \begin{pmatrix} 1. \\ 1. \end{pmatrix}, \; r = -1$$

ξ is sampled 10000 times with parameters $\mu = \begin{pmatrix} 1. \\ 1. \end{pmatrix}$

$$\Sigma = \begin{pmatrix} 20. & 0. \\ 0. & 20. \end{pmatrix}$$



Proof of concept on a quadratic Chance constraint Problem

- 2D quadratic problem

$$\min_{x \in \mathbb{R}^d} f(x) \qquad f(x) = (x-c)^\top A(x-c)$$

$$\text{s.t. } \mathbb{P}[g(x,\xi) \leq 0] \geq p \qquad g(x,z) = z^\top W(x)^\top z + q^\top z + r$$

$$\xi \sim \mathcal{N}(\mu, \Sigma)$$

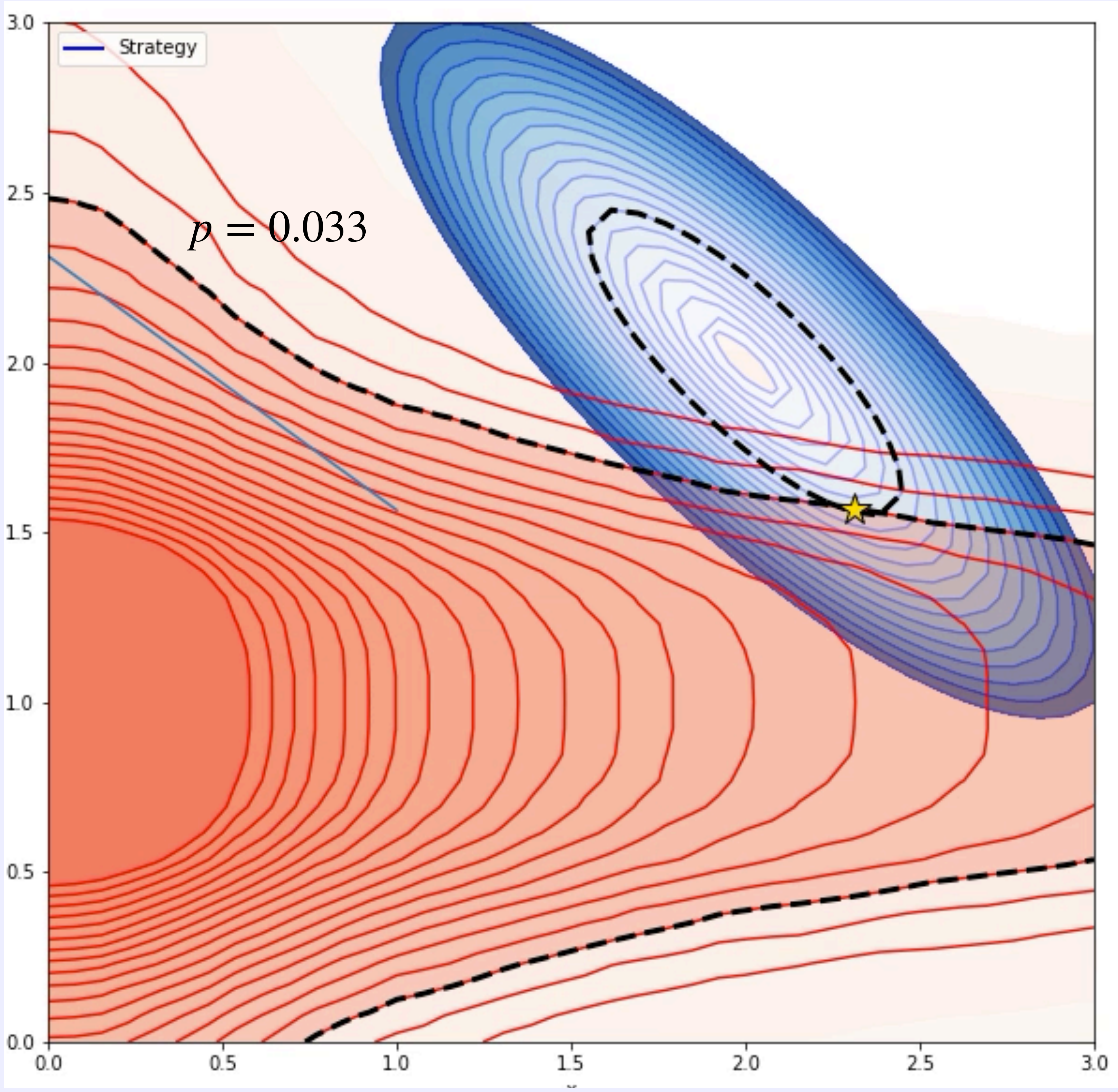
$$c = \begin{pmatrix} 2. \\ 2. \end{pmatrix} \quad A = \begin{pmatrix} 5.5 & 4.5 \\ 4.5 & 5.5 \end{pmatrix}$$

$$W : x = (x_1, x_2)^\top \mapsto \begin{pmatrix} x_1^2 + 0.5 & 0.0 \\ 0.0 & |x_2 - 1|^3 + 1. \end{pmatrix}$$

$$q = \begin{pmatrix} 1. \\ 1. \end{pmatrix}, \; r = -1$$

ξ is sampled 10000 times with parameters $\mu = \begin{pmatrix} 1. \\ 1. \end{pmatrix}$

$$\Sigma = \begin{pmatrix} 20. & 0. \\ 0. & 20. \end{pmatrix}$$



Proof of concept on a quadratic Chance constraint Problem

- 2D quadratic problem

$$\min_{x\in\mathbb{R}^d} f(x) \quad \text{s.t.} \quad \mathbb{P}[g(x,\xi)\leq 0] \geq p$$

$$f(x) = (x-c)^\top A(x-c)$$

$$g(x,z) = z^\top W(x)^\top z + q^\top z + r$$

$$\xi \sim \mathcal{N}(\mu, \Sigma)$$

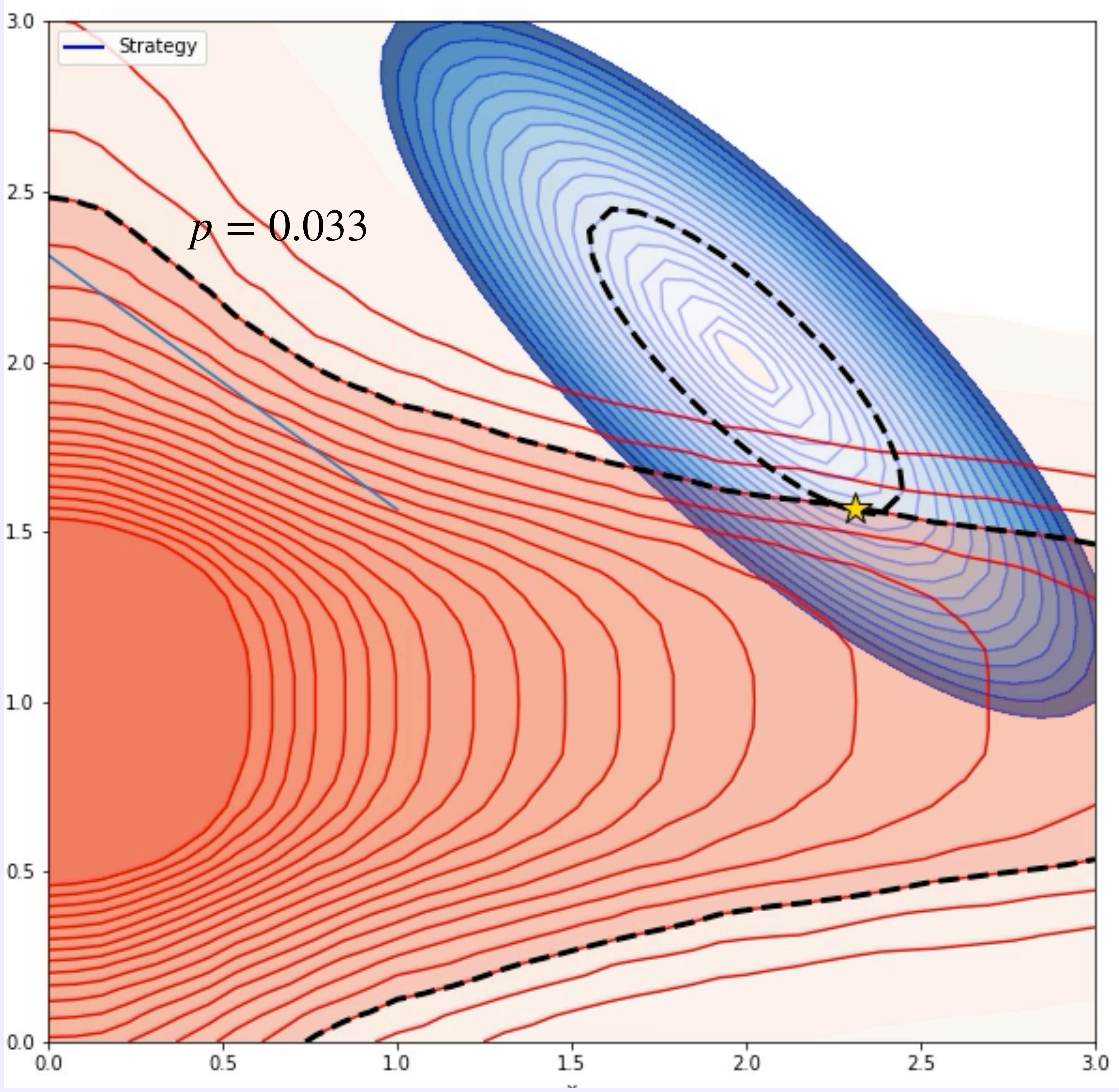
$$c = \begin{pmatrix} 2. \\ 2. \end{pmatrix} \quad A = \begin{pmatrix} 5.5 & 4.5 \\ 4.5 & 5.5 \end{pmatrix}$$

$$W : x = (x_1, x_2)^\top \mapsto \begin{pmatrix} x_1^2 + 0.5 & 0.0 \\ 0.0 & |x_2 - 1|^3 + 1. \end{pmatrix}$$

$$q = \begin{pmatrix} 1. \\ 1. \end{pmatrix}, \; r = -1$$

ξ is sampled 10000 times with parameters $\mu = \begin{pmatrix} 1. \\ 1. \end{pmatrix}$

$$\Sigma = \begin{pmatrix} 20. & 0. \\ 0. & 20. \end{pmatrix}$$



Numerical Experiments on Second Toy Problem

- A norm optimization problem

$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t.} \quad \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

$$f(x) = -\|x\|_1$$

$$g : \mathbb{R}^d \times \mathcal{M}_{n,d} \to \mathbb{R}$$

$$x, Z \mapsto \max_{i \in [n]} \sum_{j=1}^{d} Z_{i,j}^2 x_j^2$$

$$\xi_{i,j} \sim \mathcal{N}(0,1)$$

$$p = 0.8$$

Numerical Experiments on Second Toy Problem

- A norm optimization problem

$$\min_{x \in \mathbb{R}^d} f(x) \quad \text{s.t.} \quad \mathbb{P}[g(x,\xi) \leq 0] \geq p$$

$$f(x) = -\|x\|_1$$

$$g : \mathbb{R}^d \times \mathcal{M}_{n,d} \to \mathbb{R}$$

$$x, Z \mapsto \max_{i \in [n]} \sum_{j=1}^{d} Z_{i,j}^2 x_j^2$$

$$\xi_{i,j} \sim \mathcal{N}(0,1)$$

$$p = 0.8$$

- Optimal value and solution

$$f^\star = \frac{10d}{\sqrt{\underbrace{F^{-1}_{\chi^2_d}}(p^{\frac{1}{10}})}} \qquad x_i^\star = \frac{10}{\sqrt{F^{-1}_{\chi^2_d}(p^{\frac{1}{10}})}}, i \in \{1, \dots, d\}$$

Quantile function of a χ^2 distribution with d degrees of freedom

Numerical Experiments on Second Toy Problem

- A family of norm optimization problems
  Hong, Yang, Zhang (2009)

$$\min_{x\in\mathbb{R}^d} f(x) \quad f(x) = -\|x\|_1$$

$$\text{s.t. } \mathbb{P}[g(x,\xi)\leq 0] \geq p$$

$$g : \mathbb{R}^d \times \mathcal{M}_{n,d} \to \mathbb{R}$$

$$x, Z \mapsto \max_{i\in[n]} \sum_{j=1}^{d} Z_{i,j}^2 x_j^2$$

$$\xi_{i,j} \sim \mathcal{N}(0,1)$$

$$p = 0.8$$

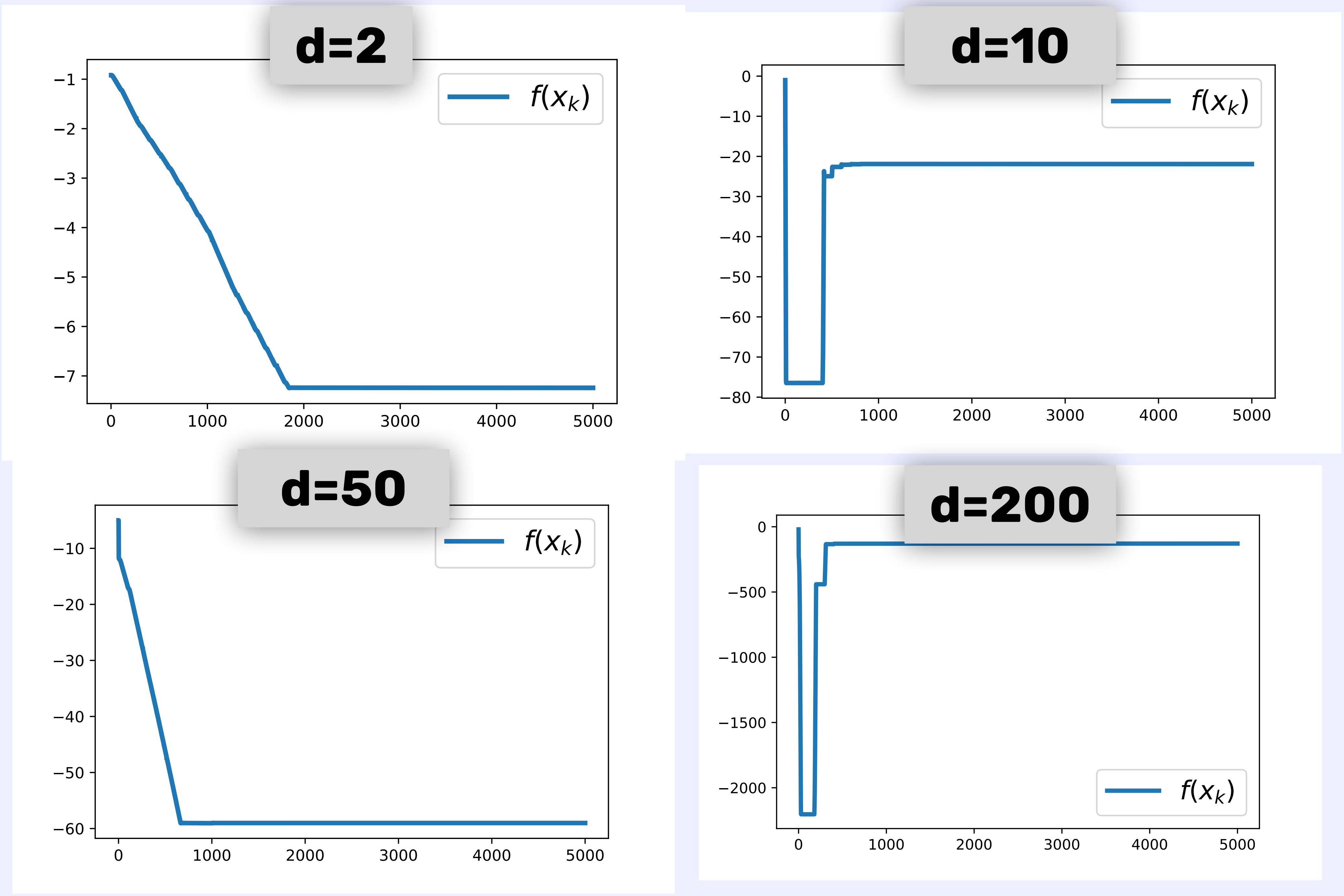
Dimension	Final Sub-optimality	$\mathbb{P}[g(x,\xi)\leq 0]$	μ	λ
$d=2$	5.1×10^{-4}	0.7992	0.01	10.0
$d=10$	2.4×10^{-2}	0.8	1.0	0.01
$d=50$	1.2×10^{-1}	0.7999	1.0	10.0
$d=200$	2.8×10^{-1}	0.7997	1.0	0.01

- Optimal value and solution

$$f^\star = \frac{10d}{\sqrt{F^{-1}_{\chi^2_d}(p^{\frac{1}{10}})}} \qquad x_i^\star = \frac{10}{\sqrt{F^{-1}_{\chi^2_d}(p^{\frac{1}{10}})}}, i \in \{1,\dots,d\}$$

Quantile function of a χ^2 distribution with d degrees of freedom

- Numerical Results



Conclusion

- We proposed a new approach to establish eventual convexity of a broad class of chance constraints, with a computable treshold.
- We propose a new approach to chance constraints via Bilevel Programming.
- We derive a double penalization method for this approach, with an exact penalty for the hard constraint.
- We propose a python toolbox to test out your problems.

- Getting sharper threshold for eventual convexity
- Derive more methods from the bilevel approach
- More numerical experiments

yassine.laguel@univ-grenoble-alpes.fr

References

- Y. Laguel, W. Van Ackooij, J Malick, G Ramalho: On the Convexity of Level-sets of Probability Functions (2021)
- Y. Laguel, J Malick, W. Van Ackooij: Chance constrained problems: a bilevel convex optimization perspective (2021)
- Y. Laguel, J Malick, Z. Harchaoui: First-order optimization for Superquantile-based learning. MLSP (2020)
- S. Kataoka.: A stochastic programming model. Econometrica **31**, (1963)
- R. Henrion, C. Strugarek: Convexity of chance constraints with independent random variables. Computational Optimization and Applications (2008)
- W. Van Ackooij: Eventual convexity of chance constrained feasible sets. Optimization, (2015)
- W. Van Ackooij, J. Malick: Eventual convexity of probability constraints with elliptical distributions. Mathematical Programming, (2019)
- W. Van Ackooij, R. Henrion: (Sub-)Gradient formulae for probability functions of random inequality systems under Gaussian distributionOptimization Letters (2017)
- A. Geletu, A. Hoffmann: Analytic approximation and differentiability of joint chance constraints. Optimization (2019)
- W. Van Ackooij, P. Pérez-Aros: Generalized differentiation of probability functions acting on an infinite system of constraints. SIAM Journal on Optimization (2019)
- RT Rockafellar, S Uryasev: Optimization of conditional value-at-risk. Journal of risk (2000)
- W de Oliveira: Proximal bundle methods for nonsmooth DC programming. Journal of Global Optimization (2019)