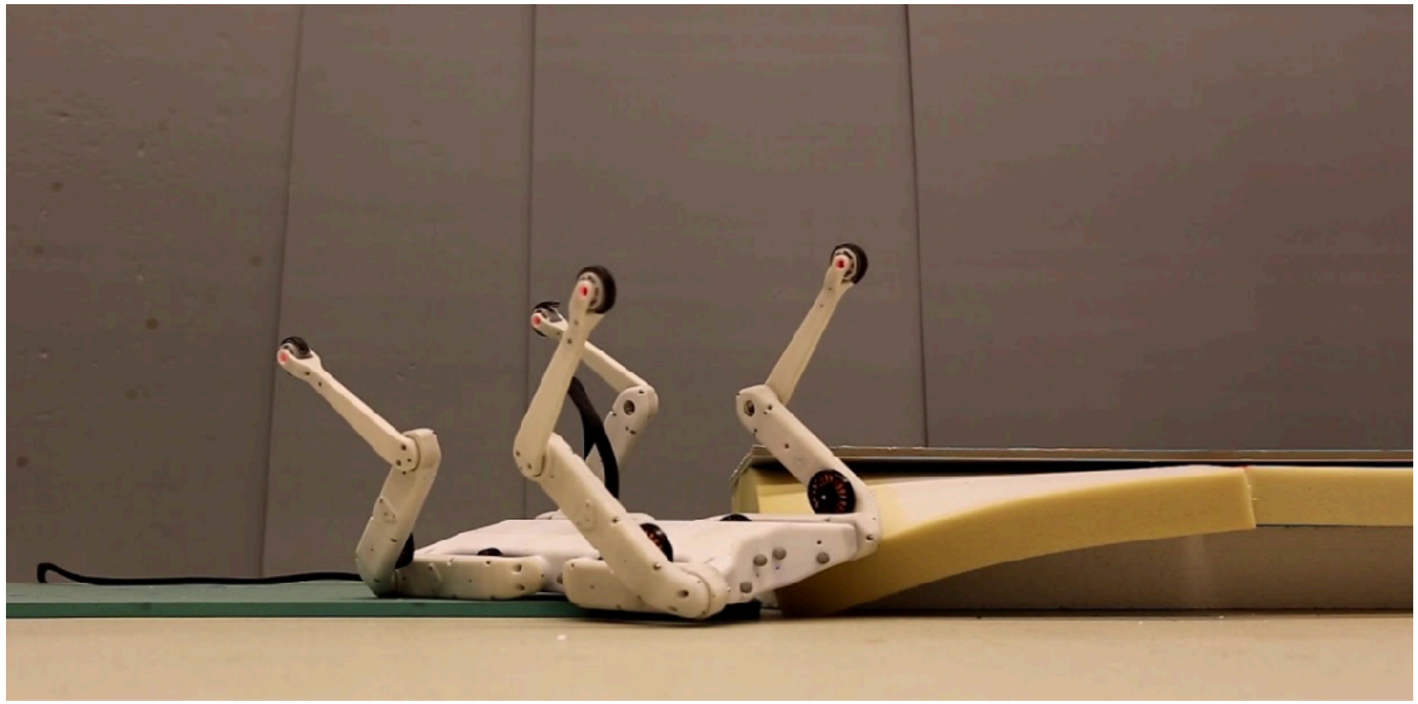
Sampling-Based Methods for Optimal Control

Theory, Algorithms, and Applications

Zeji Yi*, Chaoyi Pan*, Guanya Shi+, Guannan Qu+

*Equal Contribution + Equal Advising

Motivation

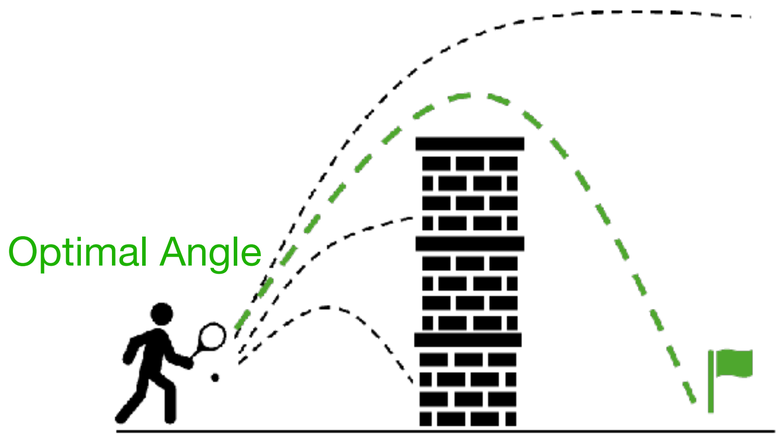


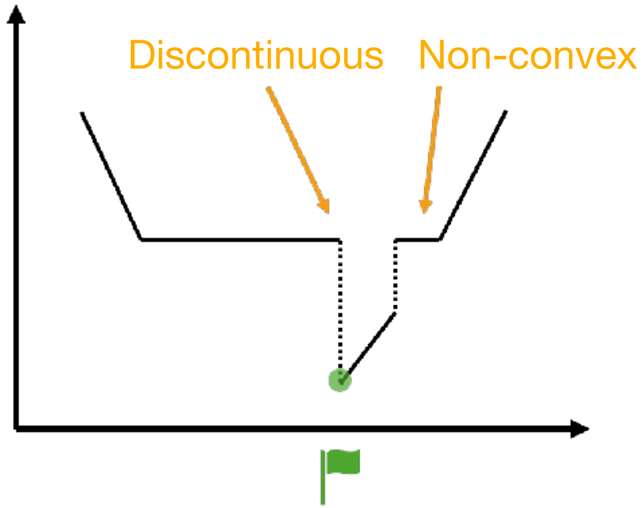
Optimal control for contact-rich robots are hard!

Image credit: Open dynamic robot initiative
https://www.youtube.com/watch?v=snMI95d05xA

Motivation

- Many optimal control problems in robotics are nonconvex

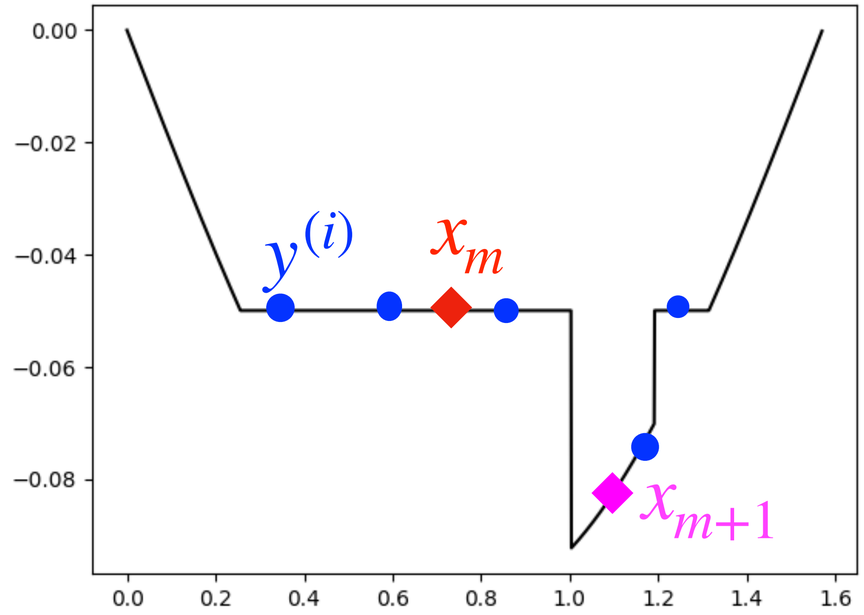




- Practitioners have found “Sampling-based Optimization”, to be surprisingly effective in solving these problems

Sampling-based Optimization

$$\min f(x)$$



For $m = 0,1,2,...$

Sample $y^{(1)}, y^{(2)}, \ldots, y^{(N)} \sim \mathcal{N}(x_m, \sigma^2 I)$

Evaluate $f(y^{(1)}), f(y^{(2)}), \ldots, f(y^{(N)})$

Update $x_{m+1} = \sum_{i=1}^{N} w^{(i)} y^{(i)}$ where $w^{(i)} = \dfrac{e^{-\frac{1}{\lambda} f(y^{(i)})}}{\sum_{j=1}^{N} e^{-\frac{1}{\lambda} f(y^{(j)})}}$

Sampling-based Optimization

For $m = 0,1,2,...$

Variance

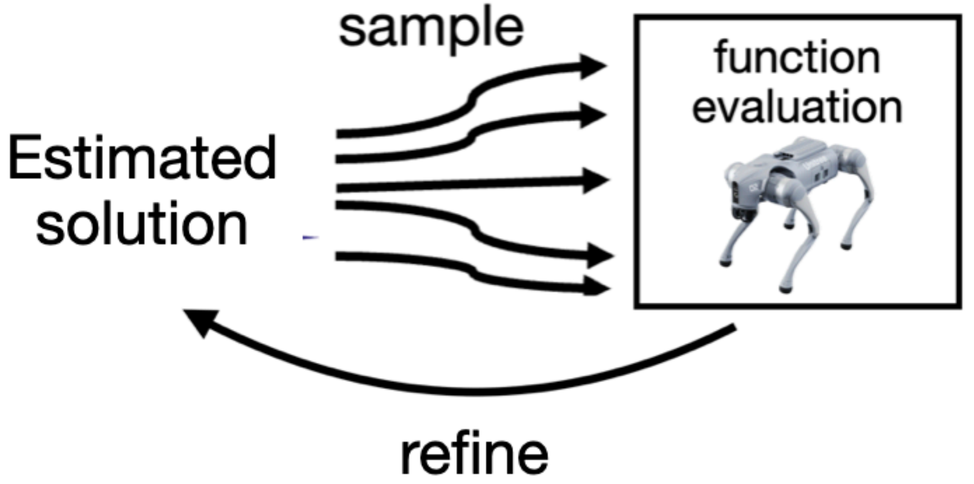
Sample $y^{(1)}, y^{(2)}, \ldots, y^{(N)} \sim \mathcal{N}(x_m, \sigma^2 I)$

Soft-max with temperature

Evaluate $f(y^{(1)}), f(y^{(2)}), \ldots, f(y^{(N)})$

Update $x_{m+1} = \sum_{i=1}^{N} w^{(i)} y^{(i)}$ where $w^{(i)} = \dfrac{e^{-\frac{1}{\lambda} f(y^{(i)})}}{\sum_{j=1}^{N} e^{-\frac{1}{\lambda} f(y^{(j)})}}$

Sampling-based Optimization



For $m = 0,1,2,...$

Variance

Sample $y^{(1)}, y^{(2)}, \ldots, y^{(N)} \sim \mathcal{N}(x_m, \sigma^2 I)$

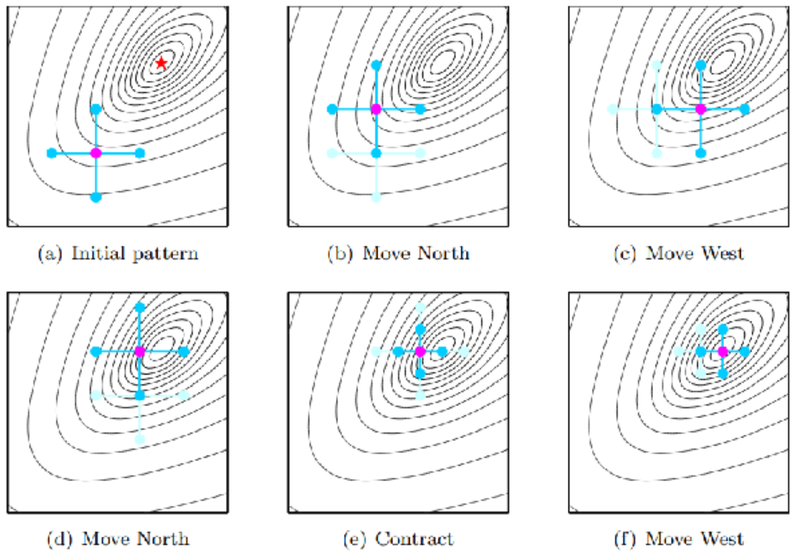
Soft-max with temperature

Evaluate $f(y^{(1)}), f(y^{(2)}), \ldots, f(y^{(N)})$

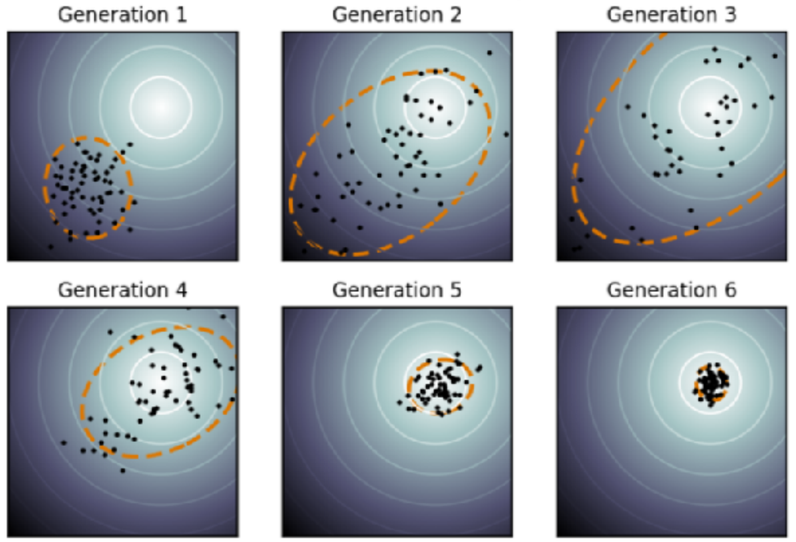
Update $x_{m+1} = \sum_{i=1}^{N} w^{(i)} y^{(i)}$ where $w^{(i)} = \frac{e^{-\frac{1}{\lambda} f(y^{(i)})}}{\sum_{j=1}^{N} e^{-\frac{1}{\lambda} f(y^{(j)})}}$

History of Sampling-based Optimal Control

- 1950s-1980s: Zeroth-order optimization: The root of sampling
  - Sample $u_k \sim \mathcal{N}(0, \sigma_k^2 I), \quad y_k = x_k + u_k,$ accept if improvement
  - Motivation: Optimize functions when gradients are unavailable or unreliable
  - Random search, pattern search, and evolutionary strategy



[Biscani et al., 2020]



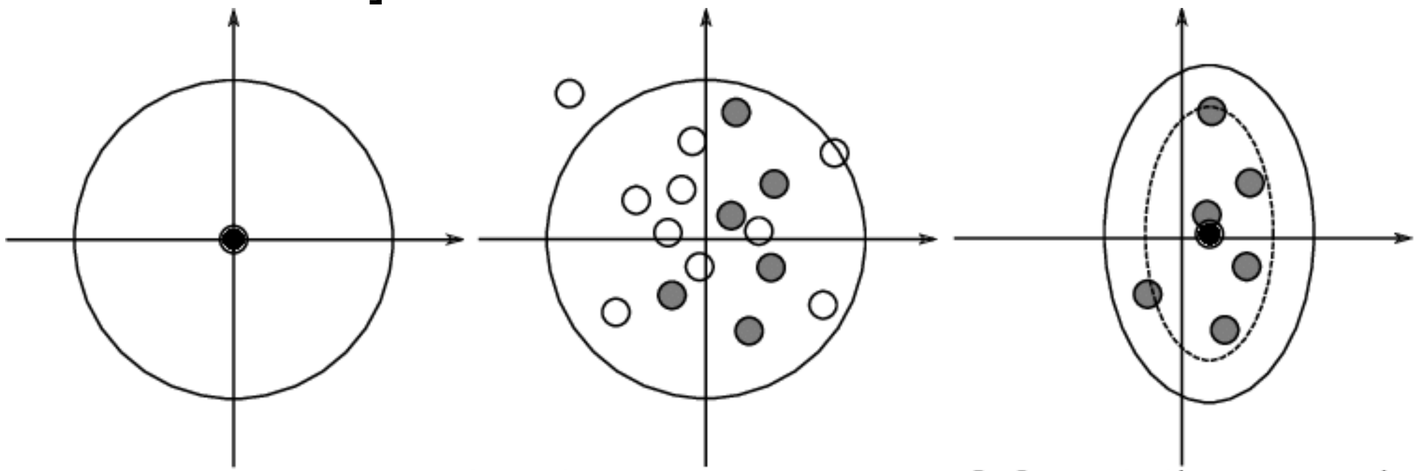
https://en.wikipedia.org/wiki/CMA-ES

History of Sampling-based Optimal Control

- 1990s-2000s: **Path Integrals Methods**
  - Optimal control problems admit a path-integral formulation [Todorov, NeurIPS 2006][Kappen, J. Stat. Mech. 2005][Theodorou et al., JMLR 2010]
    - **Control as weighted averaging** over stochastic trajectories
    - Explains **why softmax cost weighting** emerges naturally.

History of Sampling-based Optimal Control

- 1990s-2000s: **Cross-Entropy Methods**
  - Rubinstein (1997) created Cross-Entropy Methods (CEM) as a rare-event optimization method; quickly adopted for global black-box optimization. [Rubinstein, MCAP 1999]

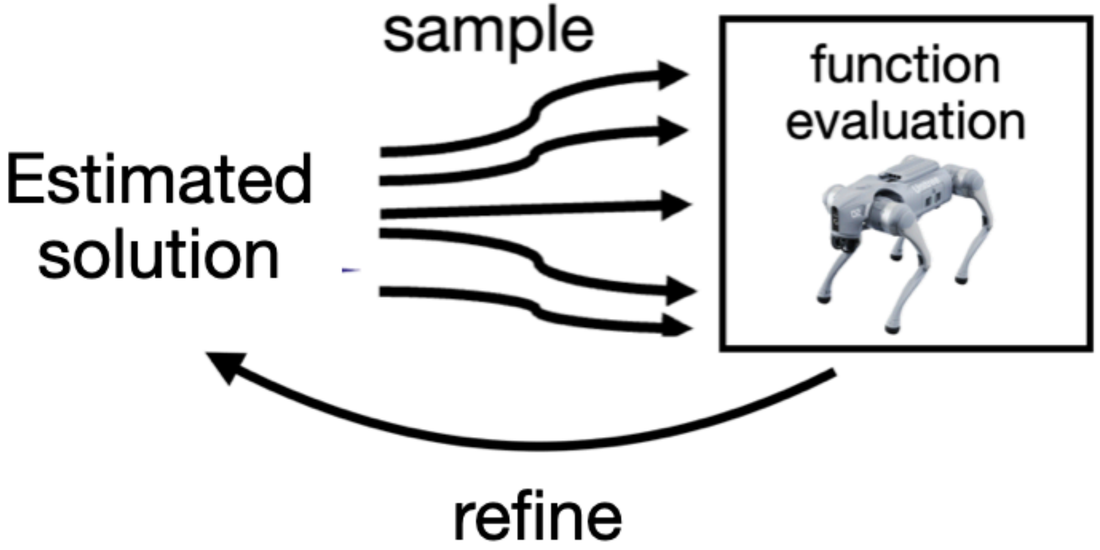


[Didier et al.,GECCO2010]

History of Sampling-based Optimal Control

- **2010s: Practical algorithms for robotics on modern GPUs**
  - GPU can easily run thousands of samples in parallel

Evaluation on the samples can be done in parallel



History of Sampling-based Optimal Control

- **2010s: Practical algorithms for robotics on modern GPUs**
  - GPU can easily run thousands of samples in parallel
  - Practical algorithms that works on real robots. Example: Model Predictive Path Integral Control (MPPI), [Williams et al., ICRA 2015,2016]

Aggressive Driving with Model Predictive Path Integral Control

Grady Williams, Paul Drews, Brian Goldfain, James M. Rehg, and Evangelos A. Theodorou

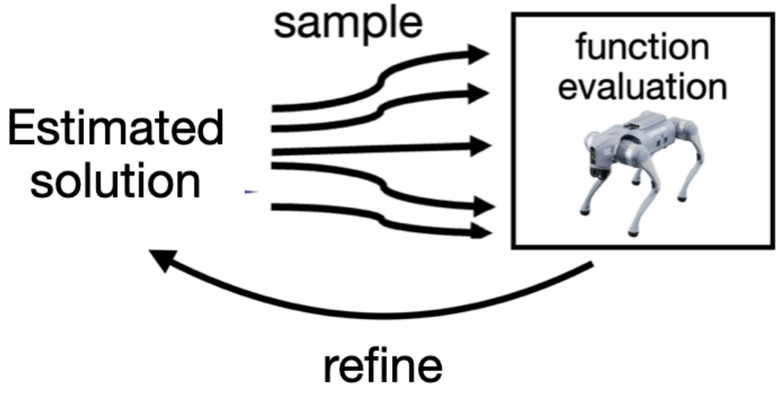
ICRA 2016

  - Early adoption in model-based RL: PETS, 2018.[Chua et al., NeurIPS 2018]

- **2020s: Improved algorithms, deep integration with learning, new theory, etc**

Sampling-based Optimization

This version of algorithm is
Model Predictive Path Integral (MPPI)
[Williams et al., ICRA 2015,2016]



For $m = 0,1,2,...$

Variance

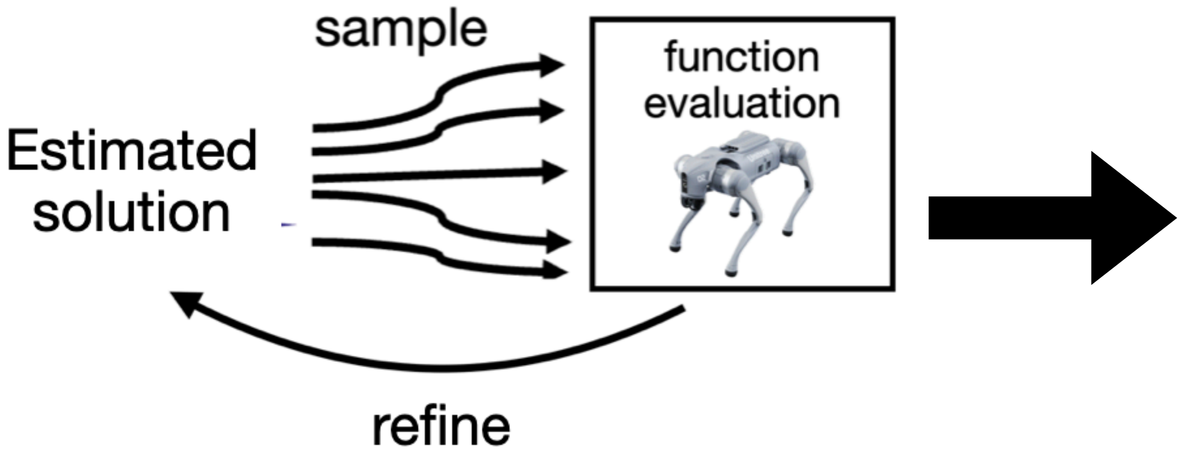
Sample $y^{(1)}, y^{(2)}, \ldots, y^{(N)} \sim \mathcal{N}(x_m, \sigma^2 I)$

Soft-max with temperature

Evaluate $f(y^{(1)}), f(y^{(2)}), \ldots, f(y^{(N)})$

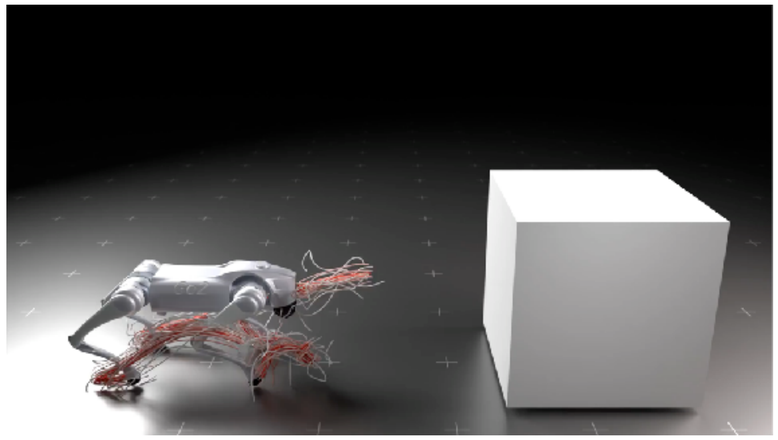
Update $x_{m+1} = \sum_{i=1}^{N} w^{(i)} y^{(i)}$ where $w^{(i)} = \frac{e^{-\frac{1}{\lambda} f(y^{(i)})}}{\sum_{j=1}^{N} e^{-\frac{1}{\lambda} f(y^{(j)})}}$

Simple yet Surprisingly Effective!



Aggressive Driving with Model Predictive Path Integral Control

Grady Williams, Paul Drews, Brian Goldfain, James M. Rehg, and Evangelos A. Theodorou



Our ICRA 2025 paper

Today's Tutorial

Part 1: Theoretical Framework of Sampling-based Optimization (SBO)

SBO often empirically succeeds in non-convex optimization

However, there is very little
theoretical understanding on why

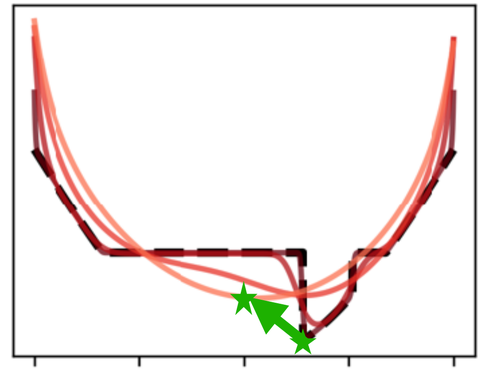
Today's Tutorial

Part 1: Theoretical Framework of Sampling-based Optimization (SBO)

How can SBO effectively navigate the highly non-convex objective?

Can we have global convergence guarantee for SBO for non convex problems?

How to choose hyper parameters optimally?



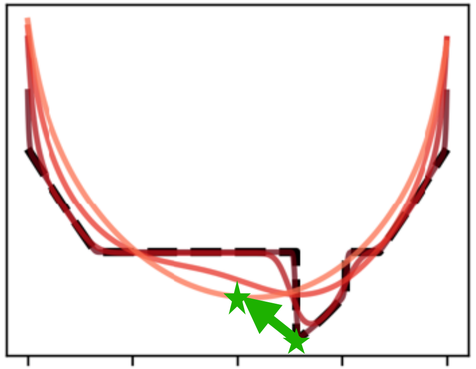
Key insight: smoothing

Focusing on pure optimization setting (not control yet)

Part 1: Theoretical Framework of Sampling-based Optimization (SBO)

Presented by Guannan Qu

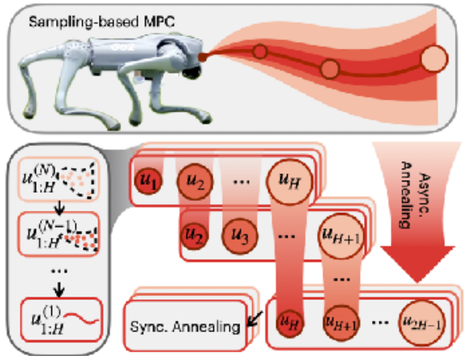
Focusing on pure optimization setting (not control yet)



Part 2: Sampling-based Optimal Control (SBOC) for Robotics

Presented by Chaoyi Pan

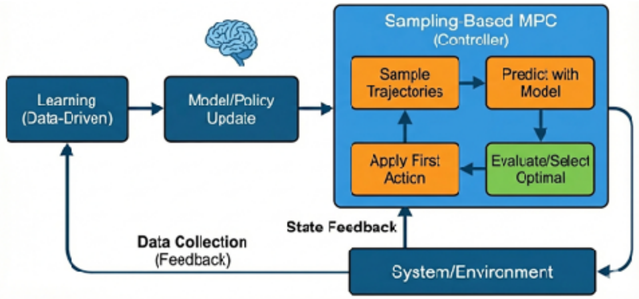
Applying SBO in a MPC manner for real-time control



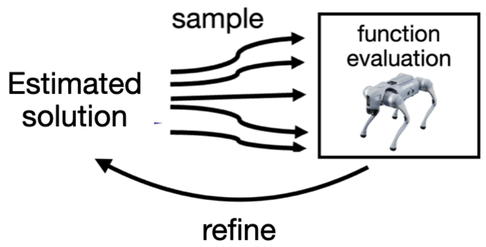
Part 3: SBOC + Learning

Presented by Zeji Yi

SBOC with learned dynamics, values, etc



Part 1: Theoretical Framework of Sampling-based Optimization (SBO)



How can SBO effectively navigate the highly non-convex objective?

Can we have global convergence guarantee for SBO for non convex problems?

How to choose hyper parameters optimally?

SBO is Zeroth Order GD on Smoothed Objective

Sample $y^{(1)}, y^{(2)}, \ldots, y^{(N)} \sim \mathcal{N}(x_m, \sigma^2 I)$

Update $x_{m+1} = \sum_{i=1}^{N} w^{(i)} y^{(i)}$ where $w^{(i)} = \dfrac{e^{-\frac{1}{\lambda} f(y^{(i)})}}{\sum_{j=1}^{N} e^{-\frac{1}{\lambda} f(y^{(j)})}}$

What if we let $N \to \infty$?

Proposition

SBO can be written as $x_{m+1} = x_m - \dfrac{\sigma^2}{\lambda} \widehat{\nabla_x g(x_m; \sigma^2)}$ with

$$\widehat{\nabla_x g(x_m; \sigma^2)} \to \nabla_x g(x_m; \sigma^2) \text{ as } N \to \infty$$

where $\widehat{\nabla_x g(x_m; \sigma^2)} = \dfrac{\lambda}{\sigma^2}(x_m - \sum_{i=1}^{N} w^{(i)} y^{(i)})$

SBO is Zeroth Order GD on Smoothed Objective

$g(x; \sigma^2)$ is a "smoothed objective" defined by:

$$p_0(x) \propto \exp(-\frac{f(x)}{\lambda})$$

Proposition

SBO can be written as $x_{m+1} = x_m - \frac{\sigma^2}{\lambda} \widehat{\nabla_x g(x_m; \sigma^2)}$ with

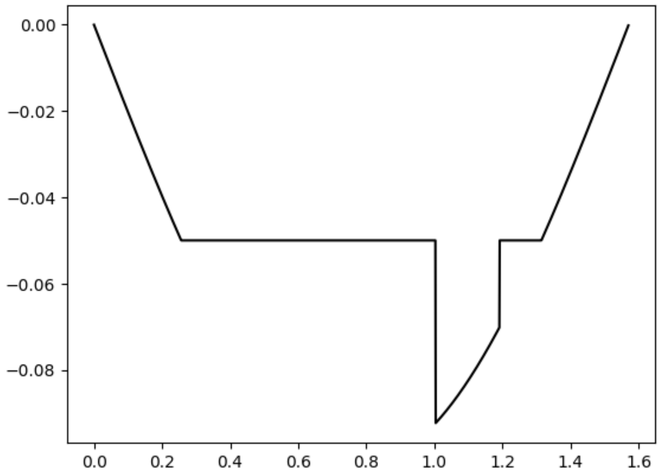
$$\widehat{\nabla_x g(x_m; \sigma^2)} \to \nabla_x g(x_m; \sigma^2) \text{ as } N \to \infty$$

SBO is Zeroth Order GD on Smoothed Objective

$g(x; \sigma^2)$ is a “smoothed objective” defined by:

$$p_0(x) \propto \exp(-\frac{f(x)}{\lambda})$$

Recall this is $f(x)$

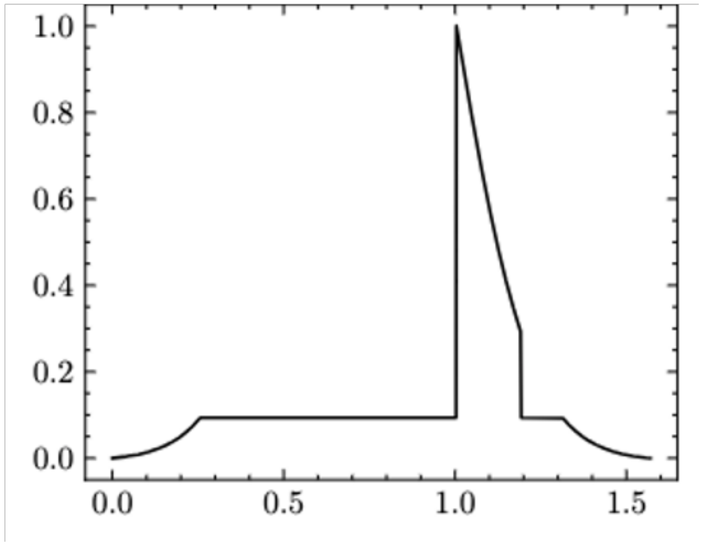


SBO is Zeroth Order GD on Smoothed Objective

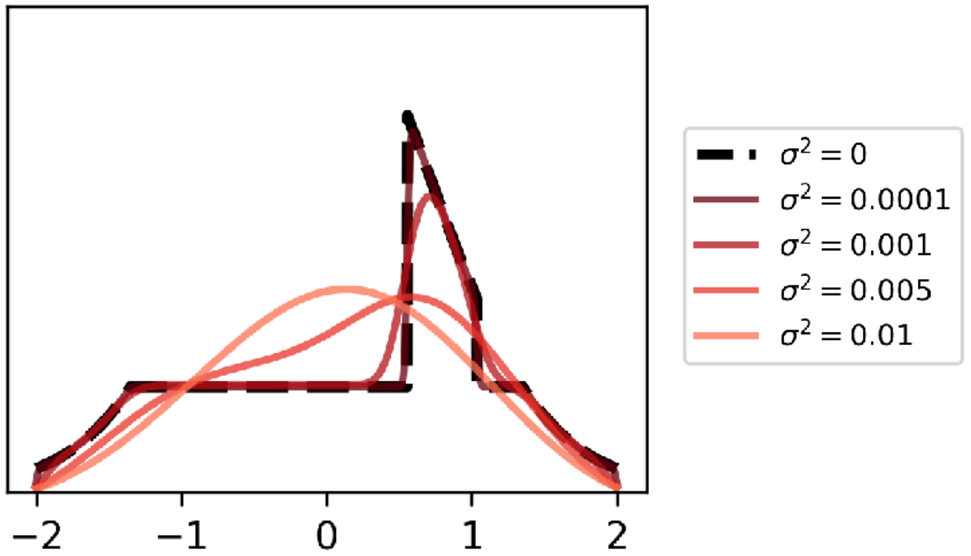
$g(x; \sigma^2)$ is a "smoothed objective" defined by:

$$p_0(x) \propto \exp(-\frac{f(x)}{\lambda}) \Rightarrow p(x; \sigma^2) = p_0 * \mathcal{N}(0, \sigma^2 I) \Rightarrow g(x; \sigma^2) = -\lambda \log p(x; \sigma^2)$$

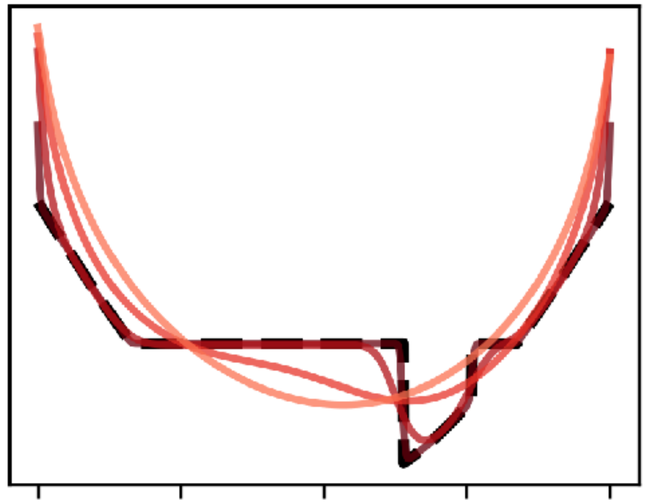
Gibbs distribution for f



Convolute with Gaussian



Taking log



SBO is Zeroth Order GD on Smoothed Objective

Proposition

SBO can be written as $x_{m+1} = x_m - \frac{\sigma^2}{\lambda} \widehat{\nabla_x g(x_m; \sigma^2)}$ with

$$\widehat{\nabla_x g(x_m; \sigma^2)} \to \nabla_x g(x_m; \sigma^2) \text{ as } N \to \infty$$

Proof Idea

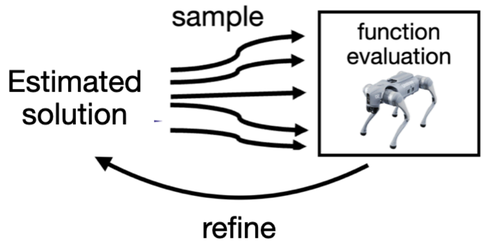
$$x_{m+1} = \frac{\sum_i y^{(i)} \exp(-\frac{1}{\lambda} f(y^{(i)}))}{\sum_i \exp(-\frac{1}{\lambda} f(y^{(i)}))} \overset{\text{def of } p_0(\cdot)}{=} \frac{\sum_i y^{(i)} p_0(y^{(i)})}{\sum_i p_0(y^{(i)})}$$

$X_0 \sim p_0(\cdot)$ $\quad X_{\sigma^2} := X_0 + \mathcal{N}(0; \sigma^2 I)$

$$\longrightarrow \mathbb{E}(X_0 \,|\, X_{\sigma^2} = x_m) \quad \text{(Using } y^{(i)} \sim \mathcal{N}(\cdot \,|\, x_m, \sigma^2 I)\text{)}$$

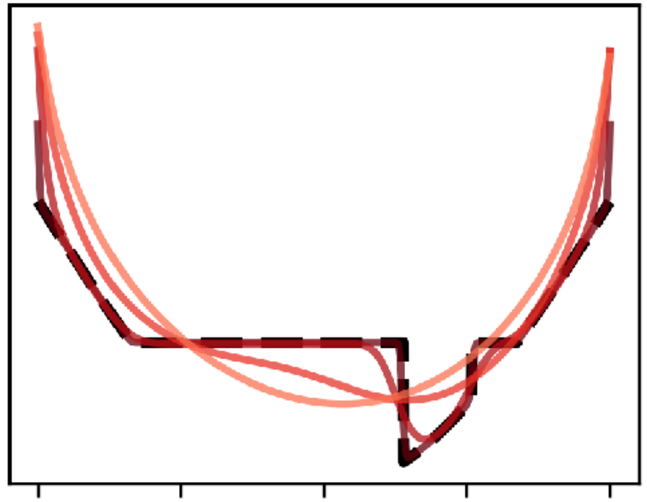
(By Tweedie's Formula) $= x_m + \sigma^2 \nabla \log p(x_m; \sigma^2)$

(Definition of g) $= x_m - \frac{\sigma^2}{\lambda} \nabla g(x_m; \sigma^2)$

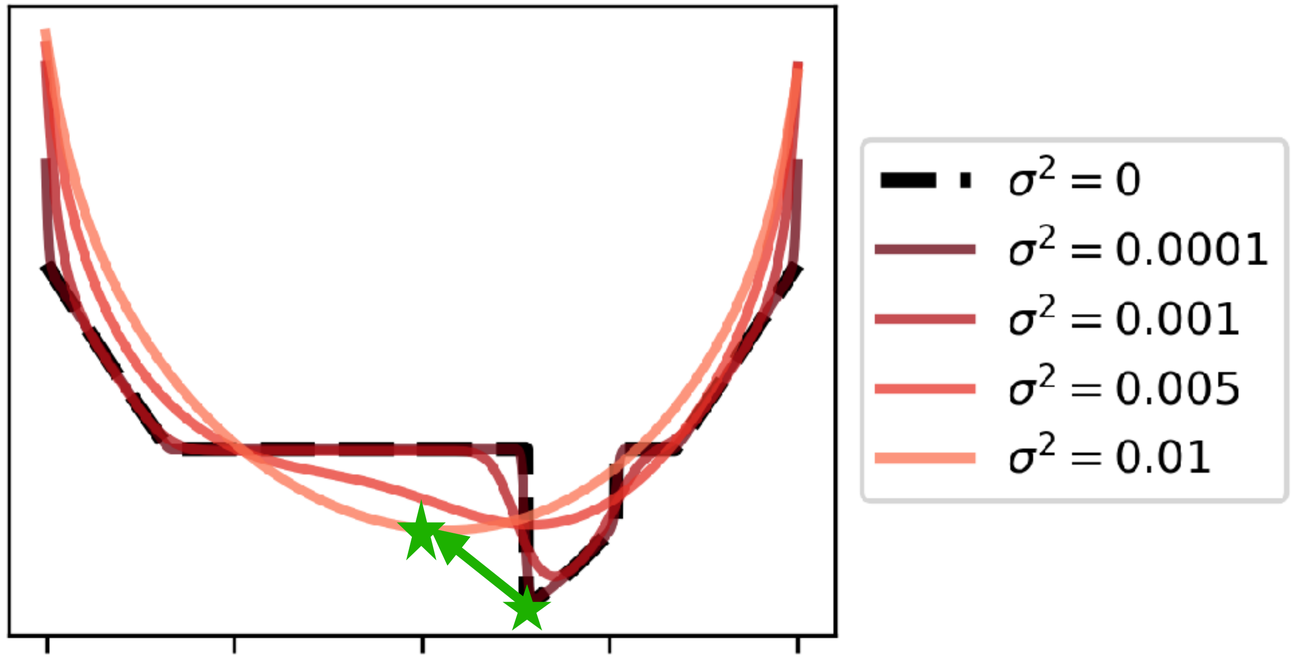


How can SBO effectively navigate the highly non-convex objective?

Boils down to analyzing the landscape of $g(x; \sigma^2)$!



As σ^2 increases, convex basin becomes larger (larger coverage)

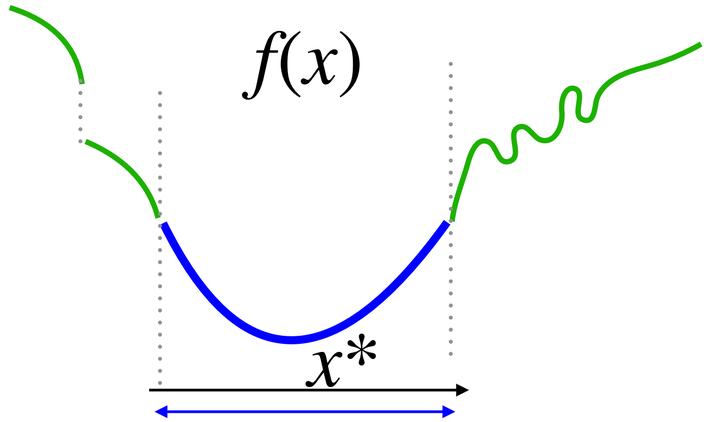


introduces “optimality gap”

Key result: Coverage and Optimality Tradeoff

Assumption

Assump. 1: x^* sits within convex basin D_τ-ball, in which $f(x)$ is α-strongly convex and L-smooth

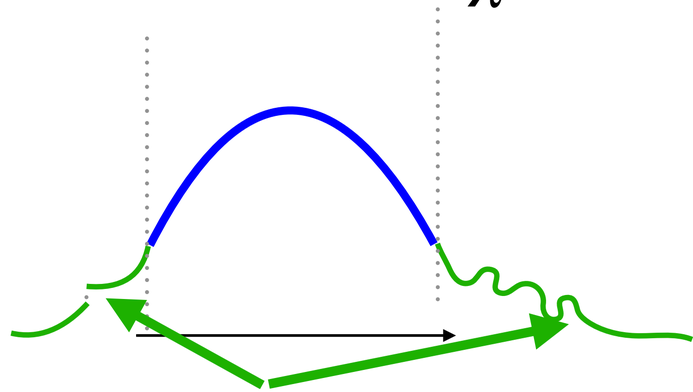


D_τ-ball: α-strongly convex L-smooth

Assump. 2: this convex basin is dominating outside region. Specifically, p_0 is subgaussian outside the D_τ ball

$$P_0(\|x - x^*\| > a) \leq \exp\left(-\frac{a^2}{2\tau^2}\right), \quad \forall a \geq D_\tau,$$

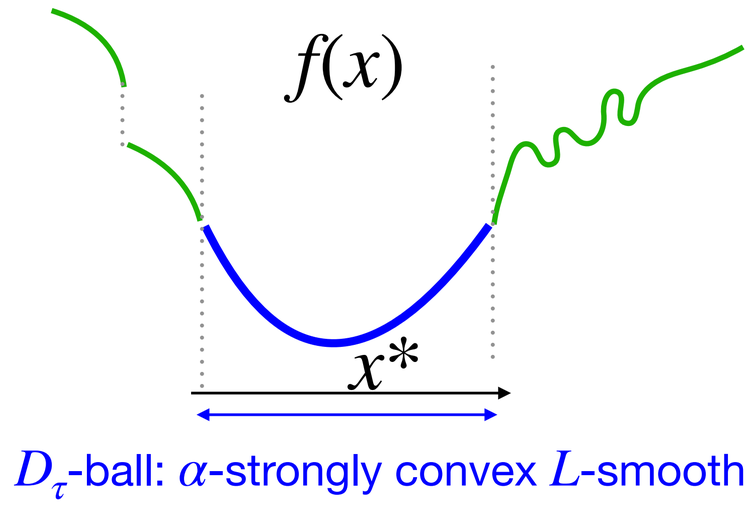
$$p_0(x) \propto \exp(-\frac{1}{\lambda} f(x))$$



Subgaussian outside the ball

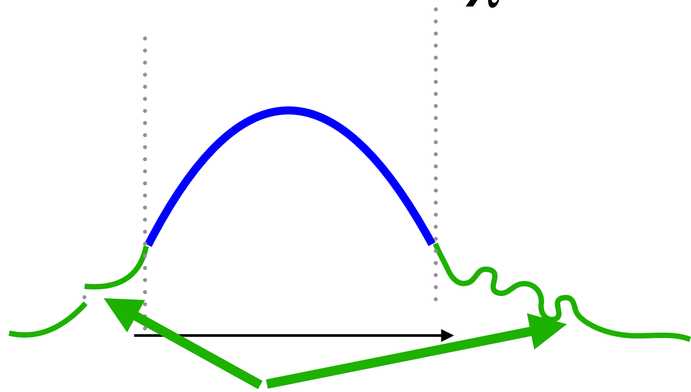
Assumption

- The “dominating basin” setting allows clean analysis for the coverage-optimality tradeoff
- Allow complex non-convexity and discontinuity
- The Subgaussian tail can potentially be relaxed



D_τ-ball: α-strongly convex L-smooth

$$p_0(x) \propto \exp(-\frac{1}{\lambda} f(x))$$



Subgaussian outside the ball

Convex Region Expansion

Theorem. $g(x;\sigma^2)$ is $\frac{\lambda}{2(\sigma^2+\frac{\lambda}{\alpha})}$-strongly convex in the following region

$$\{x \in \mathbb{R}^n \mid \|x - x^*\|^2 \leq \frac{\max(\lambda/L, \tau)+\sigma^2}{4\max(\lambda/L, \tau)} D_\tau\}$$

provided the following condition holds: $\sqrt{P_{\text{out}}} \lesssim \frac{1}{d^4}\frac{\alpha}{\lambda}\frac{D_\tau^4}{\tau^2}$

Convex region diameters increases in σ^2

Convex Region Expansion

Theorem. $g(x;\sigma^2)$ is $\frac{\lambda}{2(\sigma^2+\frac{\lambda}{\alpha})}$-strongly convex in the following region

$$\{x \in \mathbb{R}^n \mid \|x - x^*\|^2 \leq \frac{\max(\lambda/L, \tau)+\sigma^2}{4\max(\lambda/L, \tau)} D_\tau\}$$

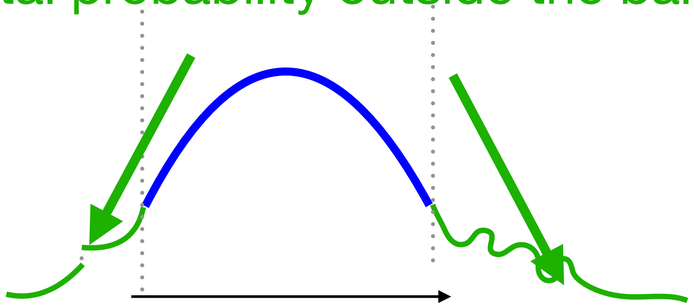
provided the following condition holds: $\sqrt{P_{\text{out}}} \lesssim \frac{1}{d^4}\frac{\alpha}{\lambda}\frac{D_\tau^4}{\tau^2}$

Ball Size

Total probability outside the ball

Dimension

Gaussian tail parameter



Optimality Gap

Theorem. The optimality gap is bounded by

Opt of $g(x;\sigma^2)$ Opt of $f(x)$

$$\left\| x^*_{\sigma^2} - x^* \right\| \leq \min \left\{ \sigma^2 \frac{1}{12 D_\tau}, (D_\tau + \tau) \frac{\sigma^2 + \frac{\lambda}{\alpha}}{0.5\sigma^2} \right\}$$

Linear growth

Upper bounded by constant

🌟Optimality gap grows linearly until certain threshold are reached

Illustration of the Tradeoff

$g(x; \sigma^2)$

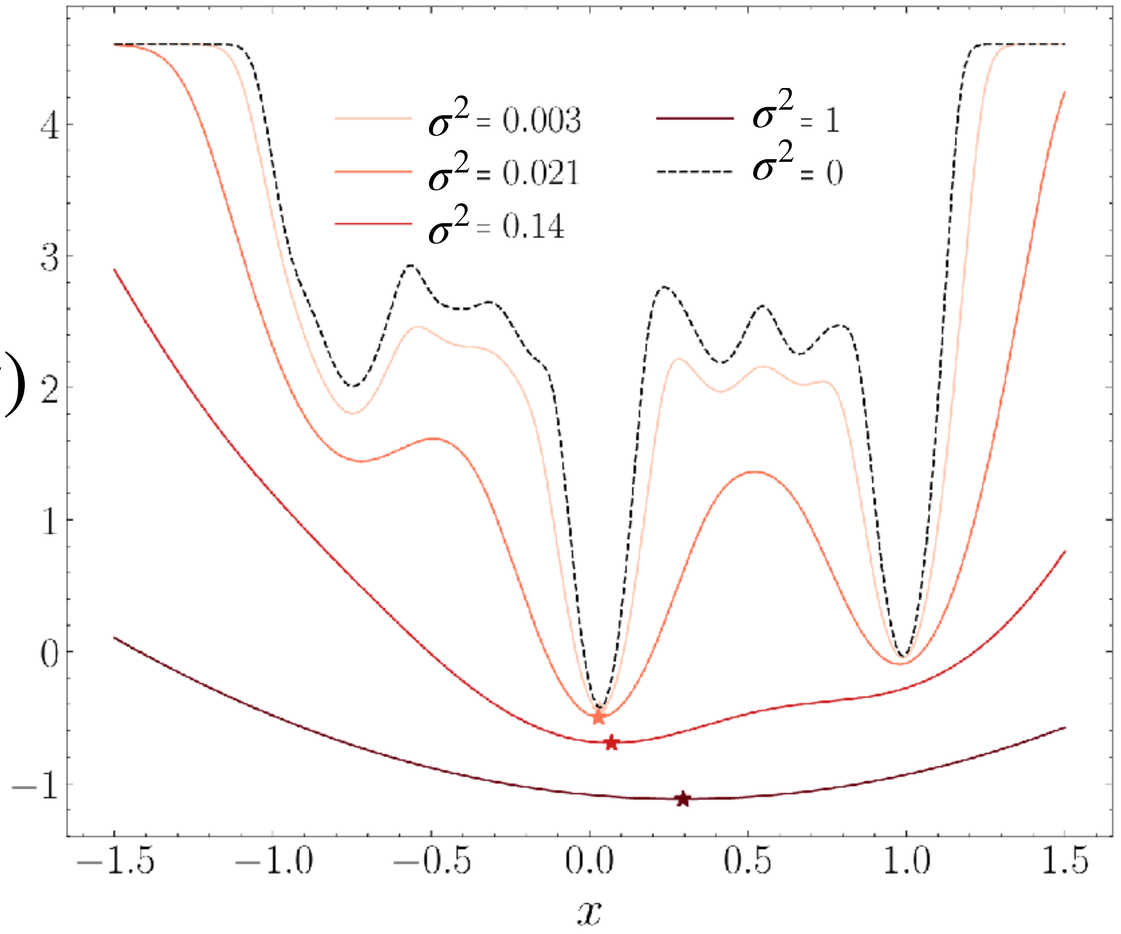
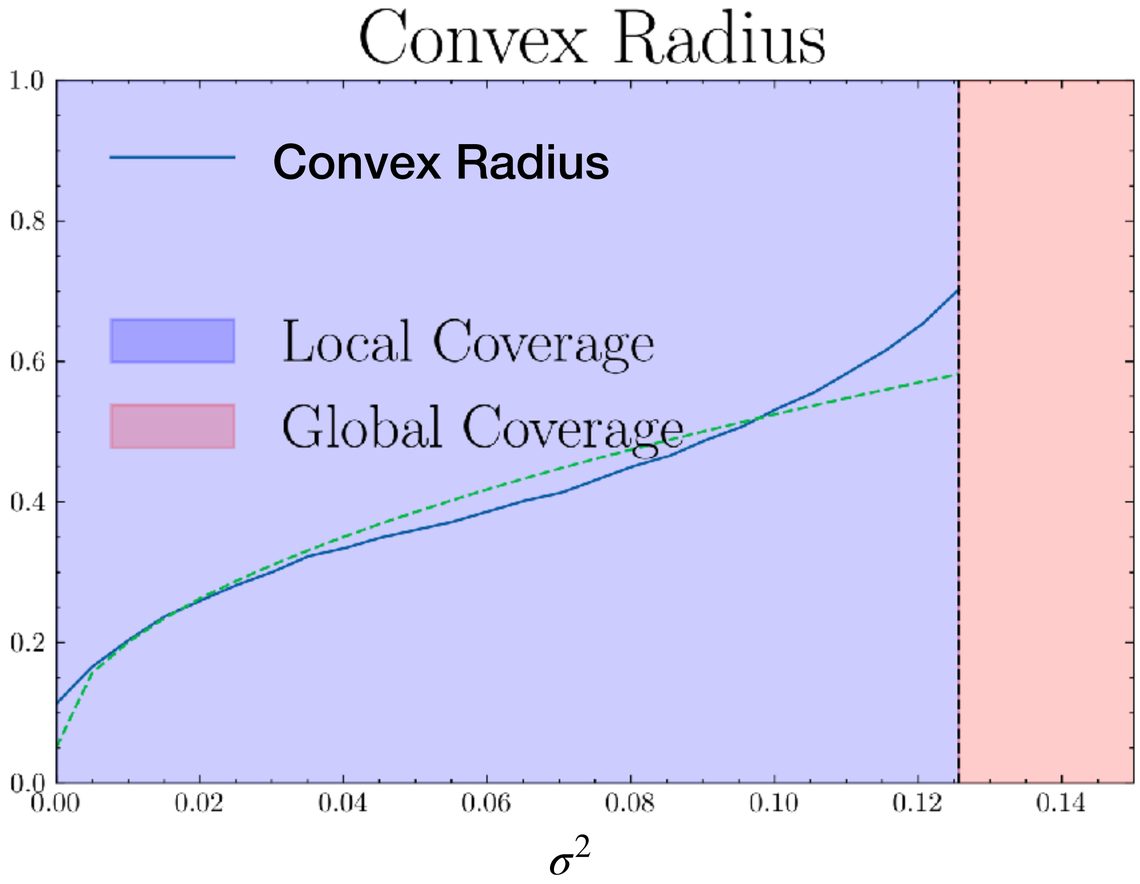
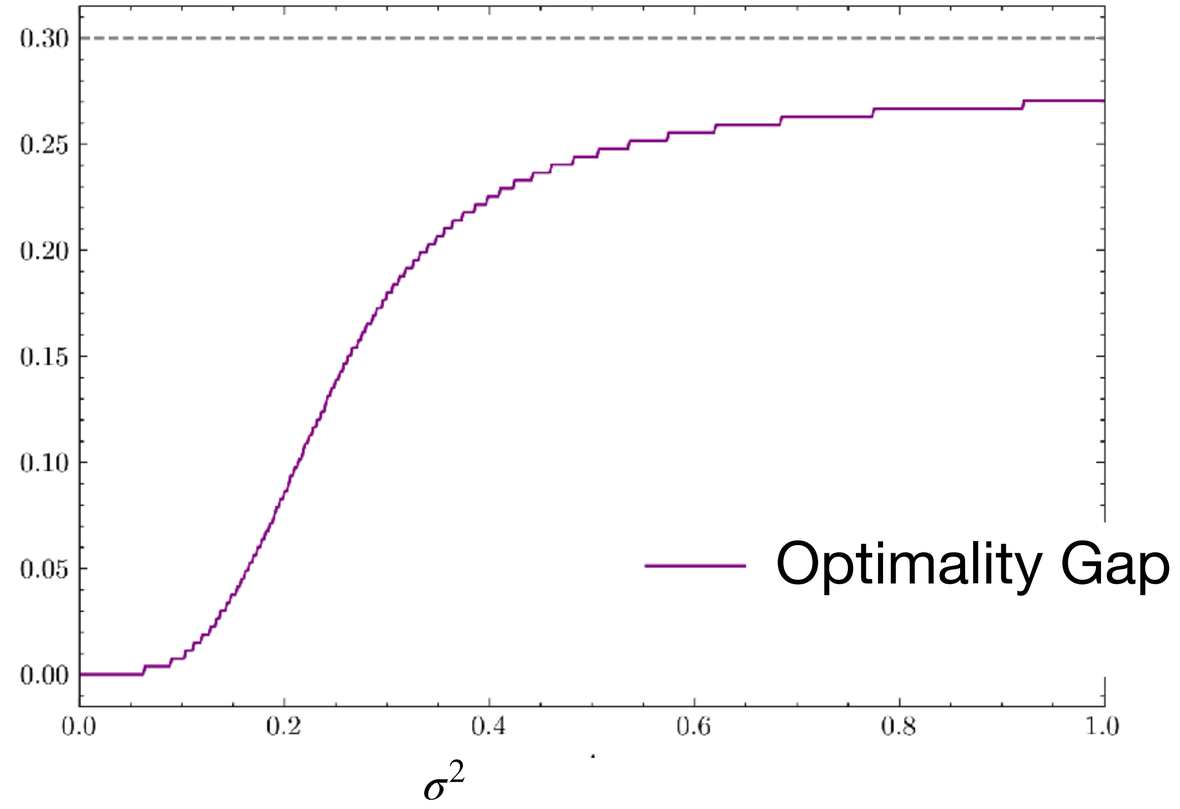


Illustration of the Tradeoff





Proof Idea

Theorem. The $g(x;\sigma^2)$ is $\frac{\lambda}{2(\sigma^2+\frac{\lambda}{\alpha})}$-strongly convex in the following region

$$\{x \in \mathbb{R}^n \mid \|x - x^*\|^2 \leq \frac{\max(\lambda/L, \tau)+\sigma^2}{4\max(\lambda/L, \tau)} D_\tau\}$$

provided the following condition holds: $\sqrt{P_{\text{out}}} \lesssim \frac{1}{d^4}\frac{\alpha}{\lambda}\frac{D_\tau^4}{\tau^2}$

$\nabla_x^2 g(x;\sigma^2) = \frac{\lambda}{\sigma^4}\left[\sigma^2 \mathbf{I} - \text{Cov}[X_0 \,|\, X_{\sigma^2} = x]\right] \leq \frac{\lambda}{\sigma^2}\mathbf{I}$ indicating smoothness

Proof Idea

To show convexity, need to upper bound Cov

$$\nabla_x^2 g(x;\sigma^2) = \frac{\lambda}{\sigma^4}\left[\sigma^2 \mathbf{I} - \mathrm{Cov}[X_0 \,|\, X_{\sigma^2} = x]\right]$$

X_0 sampled from p_0

$$p_0(x) \propto \exp(-\frac{f(x)}{\lambda})$$

$$X_{\sigma^2} = X_0 + \sigma^2 \epsilon \text{ where } \epsilon \sim \mathcal{N}(0,\mathbf{I})$$

$$p(x;\sigma^2) = p_0 * \mathcal{N}(0,\sigma^2\mathbf{I})$$

Proof Idea

To show convexity, need to upper bound Cov

$$\nabla_x^2 g(x;\sigma^2) = \frac{\lambda}{\sigma^4}\left[\sigma^2 \mathbf{I} - \mathrm{Cov}[X_0 \,|\, X_{\sigma^2} = x]\right]$$

X_0 sampled from p_0

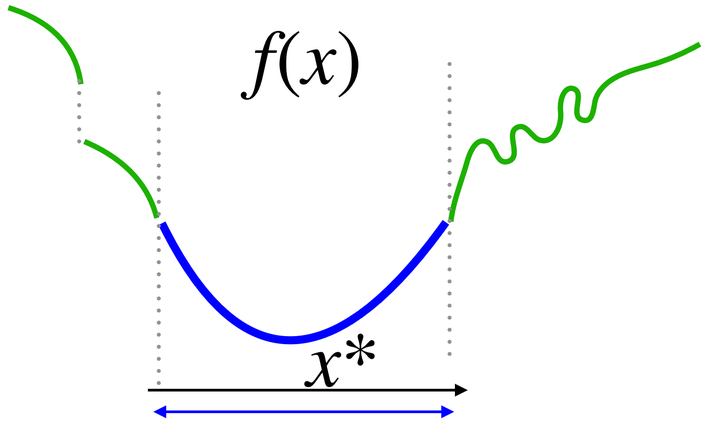
$$p_0(x) \propto \exp(-\frac{f(x)}{\lambda})$$

$$X_{\sigma^2} = X_0 + \sigma^2\epsilon \text{ where } \epsilon \sim \mathcal{N}(0,\mathbf{I})$$

$$p(x;\sigma^2) = p_0 * \mathcal{N}(0,\sigma^2\mathbf{I})$$

Recall this assumption

Assump. 1: x^* sits within convex basin D_τ-ball, in which $f(x)$ is α-strongly convex and L-smooth



D_τ-ball: α-strongly convex L-smooth

Proof Idea

$$\nabla_x^2 g(x;\sigma^2) = \frac{\lambda}{\sigma^4}\Big[\sigma^2\,\mathbf{I} - \mathrm{Cov}[X_0 \,|\, X_{\sigma^2} = x]\Big]$$

X_0 sampled from p_0

$$p_0(x) \propto \exp(-\frac{f(x)}{\lambda})$$

$X_{\sigma^2} = X_0 + \sigma^2\epsilon$ where $\epsilon \sim \mathcal{N}(0,\mathbf{I})$

$$p(x;\sigma^2) = p_0 * \mathcal{N}(0,\sigma^2\mathbf{I})$$

Conditioned on $X_0 \in D_\tau$-Ball

f is α-strongly-convex

X_0 is approx. Gaussian with var $\geq \lambda/\alpha$

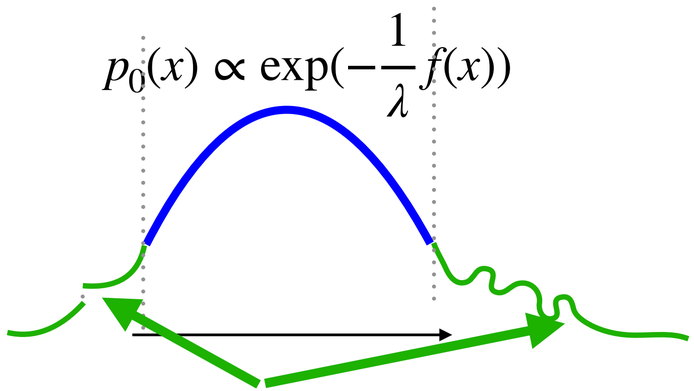
X_{σ^2} is approx. Gaussian with var $\geq \lambda/\alpha + \sigma^2$

Let's upper bound

$$\mathrm{Cov}[X_0 \,|\, X_{\sigma^2} = x, X_0 \in D_\tau\text{-Ball}] \preceq \frac{\lambda/\alpha}{\lambda/\alpha + \sigma^2}\sigma^2\mathbf{I}$$

Proof Idea

$$\nabla_x^2 g(x;\sigma^2) = \frac{\lambda}{\sigma^4}\left[\sigma^2 \mathbf{I} - \mathrm{Cov}[X_0 \,|\, X_{\sigma^2} = x]\right]$$



Assumption 2: Subgaussian outside the ball

We are only proving strong convexity within a radius

$$\{x \in \mathbb{R}^n \mid \|x - x^*\|^2 \leq \frac{\max(\lambda/L, \tau) + \sigma^2}{4\max(\lambda/L, \tau)} D_\tau\}$$

$$\sqrt{P_{\text{out}}} \lesssim \frac{1}{d^4} \frac{\alpha}{\lambda} \frac{D_\tau^4}{\tau^2}$$

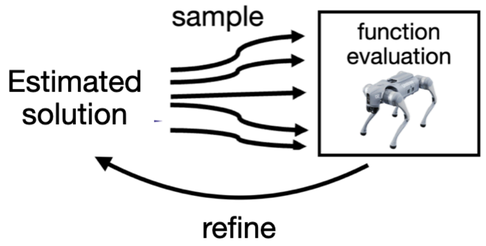
Upper bound $\mathrm{Cov}[X_0 \,|\, X_{\sigma^2} = x, X_0 \notin D_\tau\text{-Ball}]$

Proof Idea

Theorem. The $g(x;\sigma^2)$ is $\frac{\lambda}{2(\sigma^2+\frac{\lambda}{\alpha})}$-strongly convex in the following region

$$\{x \in \mathbb{R}^n \mid \|x - x^*\|^2 \leq \frac{\max(\lambda/L, \tau)+\sigma^2}{4\max(\lambda/L, \tau)} D_\tau\}$$

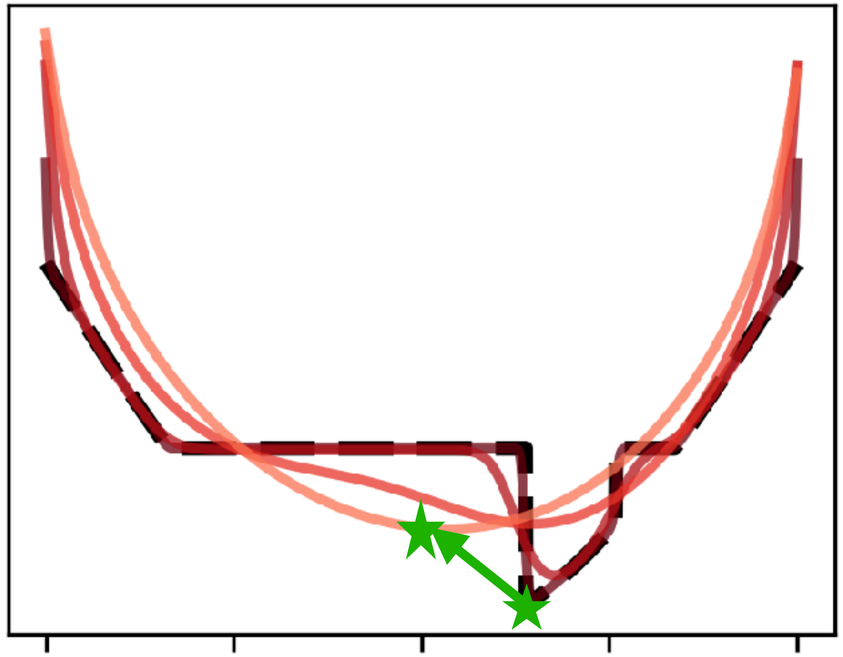
provided the following condition holds: $\sqrt{P_{\text{out}}} \lesssim \frac{1}{d^4}\frac{\alpha}{\lambda}\frac{D_\tau^4}{\tau^2}$



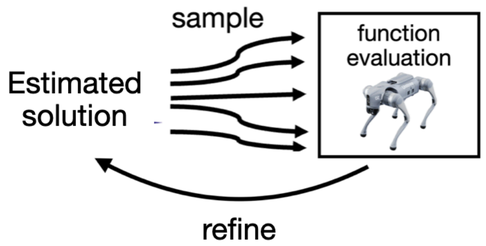
How can SBO effectively navigate the highly non-convex objective?

Key result: coverage and optimality tradeoff

As σ^2 increases, convex basin becomes larger



introduces "optimality gap"



How can SBO effectively navigate the highly non-convex objective?

Key result: coverage and optimality tradeoff

Can we have global convergence guarantee for SBO for non convex problems?

How to choose hyper parameters optimally?

Recall SBO can be written as

$$x_{m+1} = x_m - \frac{\sigma^2}{\lambda} \widehat{\nabla_x g(x_m; \sigma^2)}$$

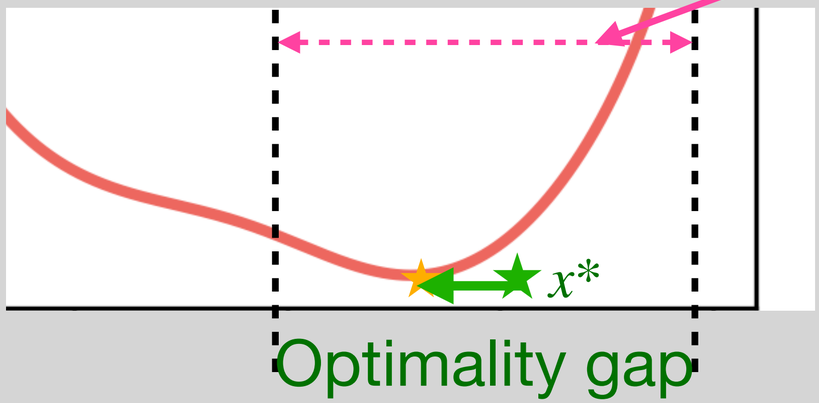
We can also show a bound on

$$\widehat{\nabla_x g(x_m; \sigma^2)} - \nabla_x g(x_m; \sigma^2)$$

Recall we've shown $g(x, \sigma^2)$ is

$\frac{\lambda}{2(\sigma^2+\frac{\lambda}{\alpha})}$-**strongly convex** and $\frac{\lambda}{\sigma^2}$ -**smooth**

within radius $\frac{\max(\lambda/L, \tau)+\sigma^2}{4\max(\lambda/L, \tau)} D_\tau$ of x^*



if x_0 starts in the convex region,

applying standard SGD analysis

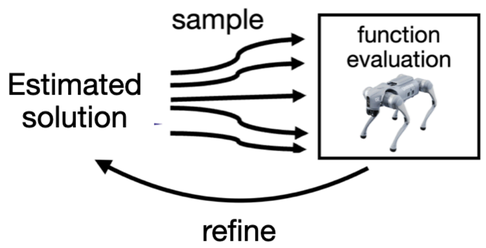
we get convergence to neighborhood of x^*.

Convergence of SBO

Theorem (Informal). Under the assumption 1,2 and the condition in the previous theorems, suppose x_0 is in the enlarged convex region. Then

$$\|x_m - x^*\|^2 \lesssim \rho^m \|x_0 - x_\sigma^*\|^2 + \underbrace{\|x_{\sigma^2}^* - x^*\|^2}_{\text{Optimality Gap}} + \underbrace{\frac{1}{N}\kappa_{\sigma^2}\left(\lambda^2 U_0 + U_1 L^2 \sigma^2 + \frac{1}{\lambda^2} U_1 L^4 \sigma^4\right)}_{\text{Statistical Error}}$$

Contraction due to strong-convexity



How can SBO effectively navigate the highly non-convex objective?

Key result: coverage and optimality tradeoff

Can we have global convergence guarantee for SBO for non convex problems?

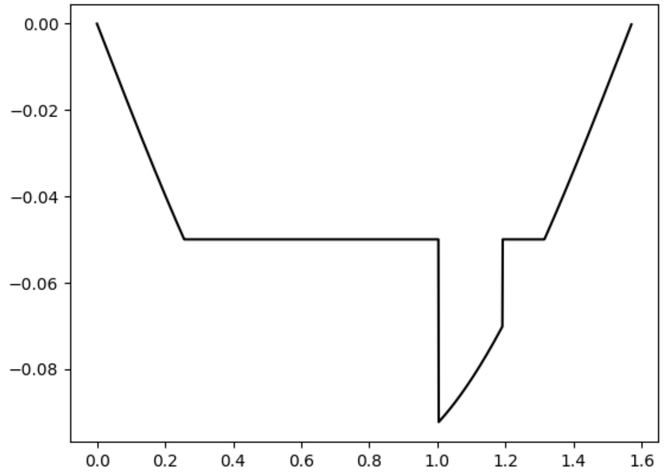
Key result: convergence to neighborhood, when initialized in the enlarged convex region

How to choose hyper parameters optimally?

How to choose hyper-parameters

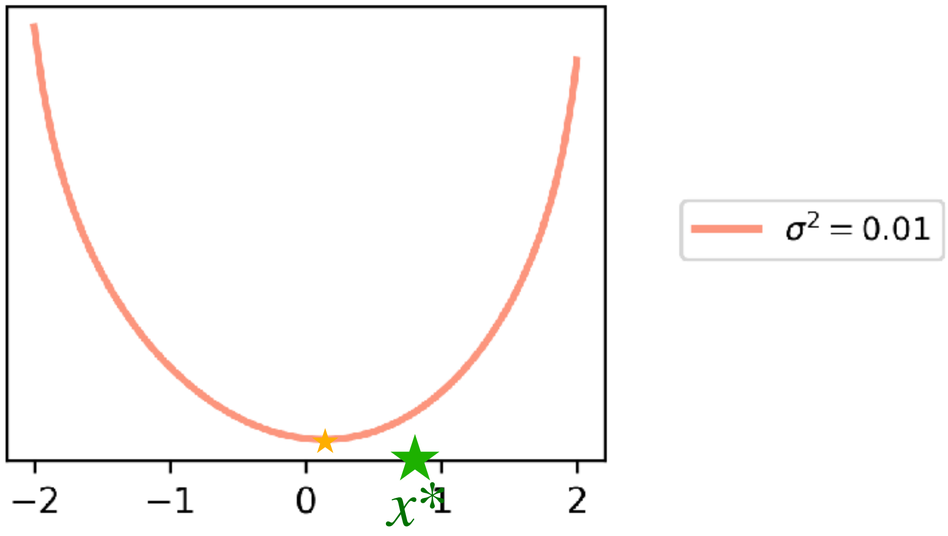
Existing SBO uses fixed parameters

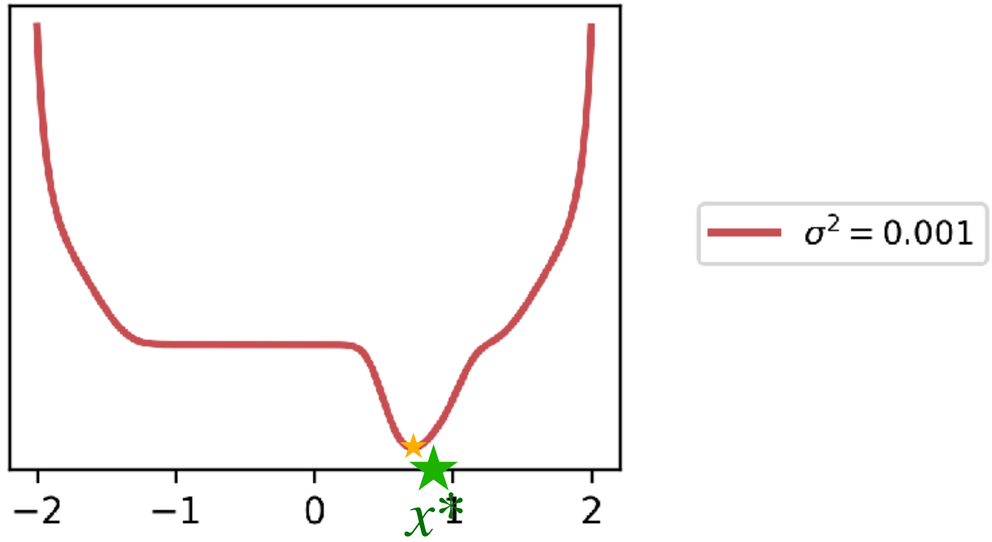
Recall this is $f(x)$



How to choose hyper-parameters

Existing SBO uses fixed parameters

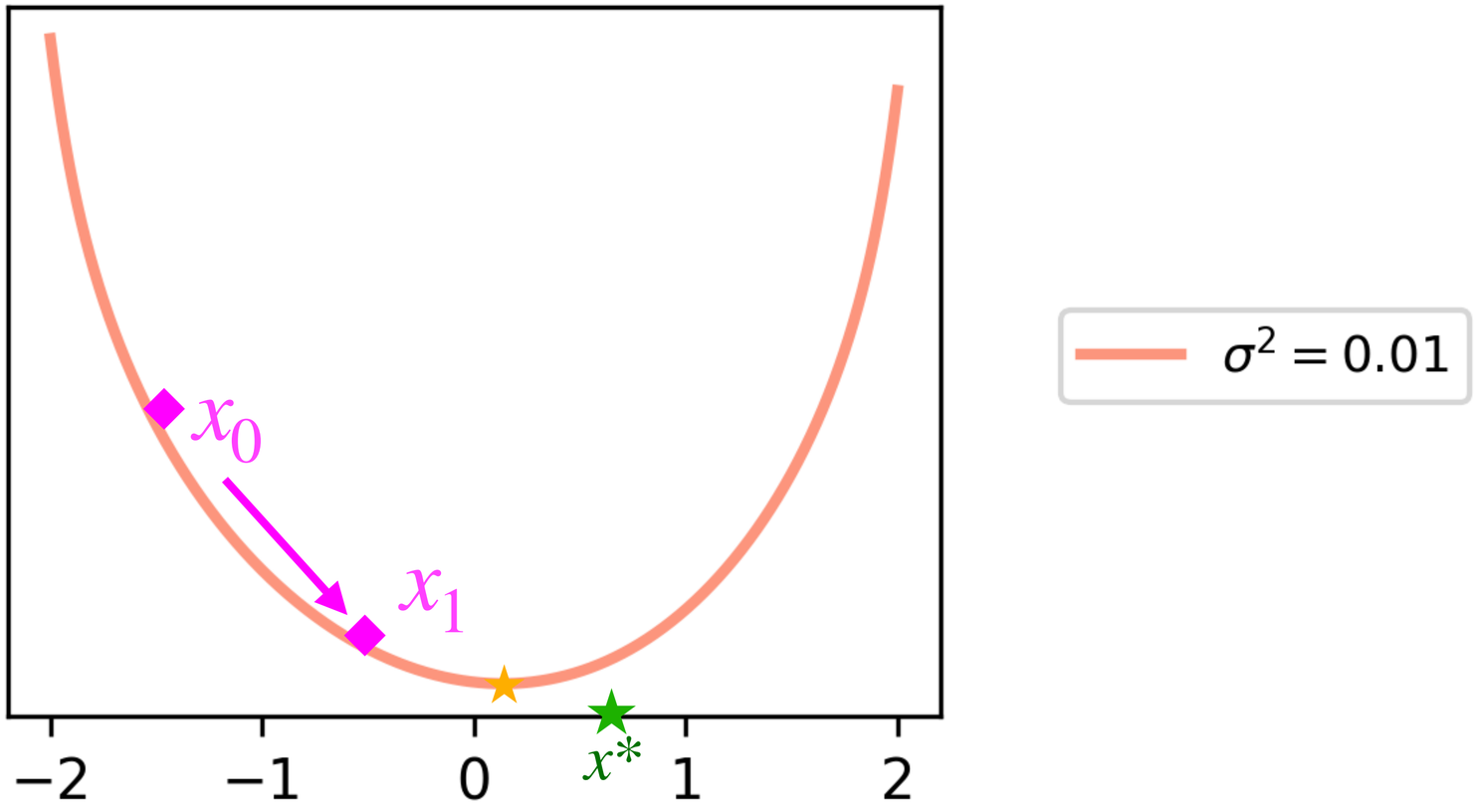




Larger σ^2: convex but large optimality gap

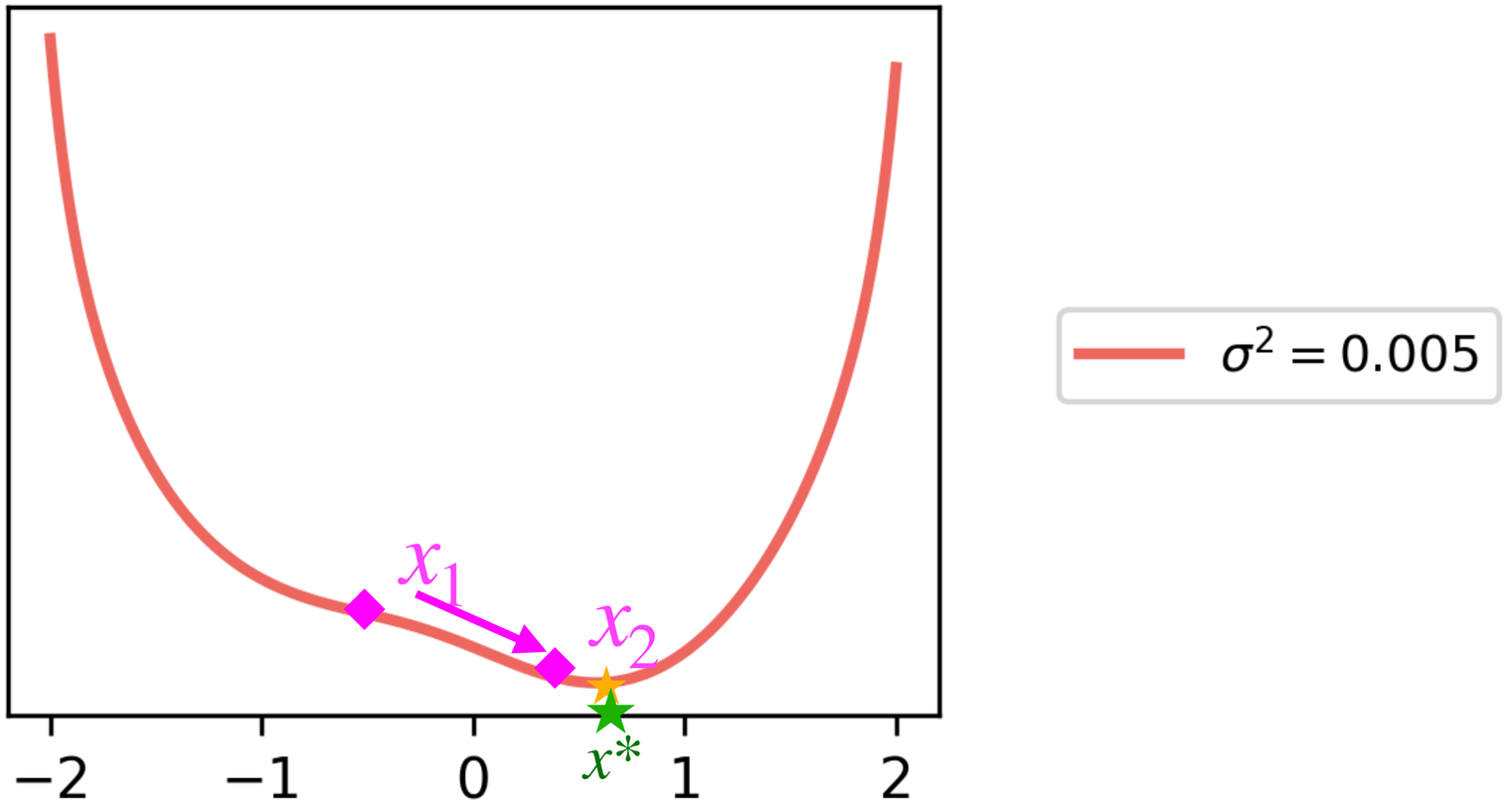
smaller σ^2: very small convex basin

Annealed sampling to ensure global convergence!



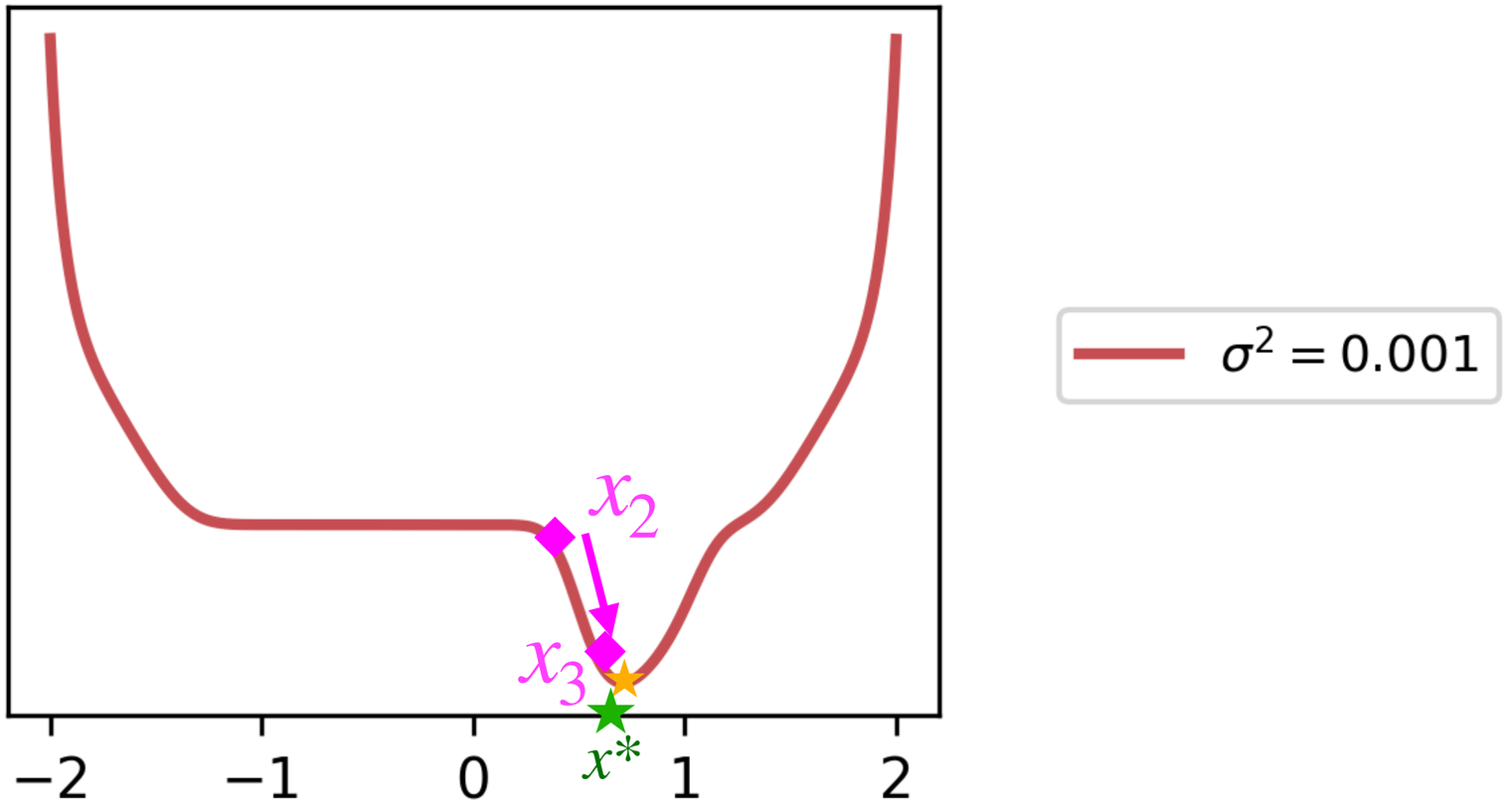
Start from a large σ^2 that has a large convex basin

Annealed sampling to ensure global convergence!



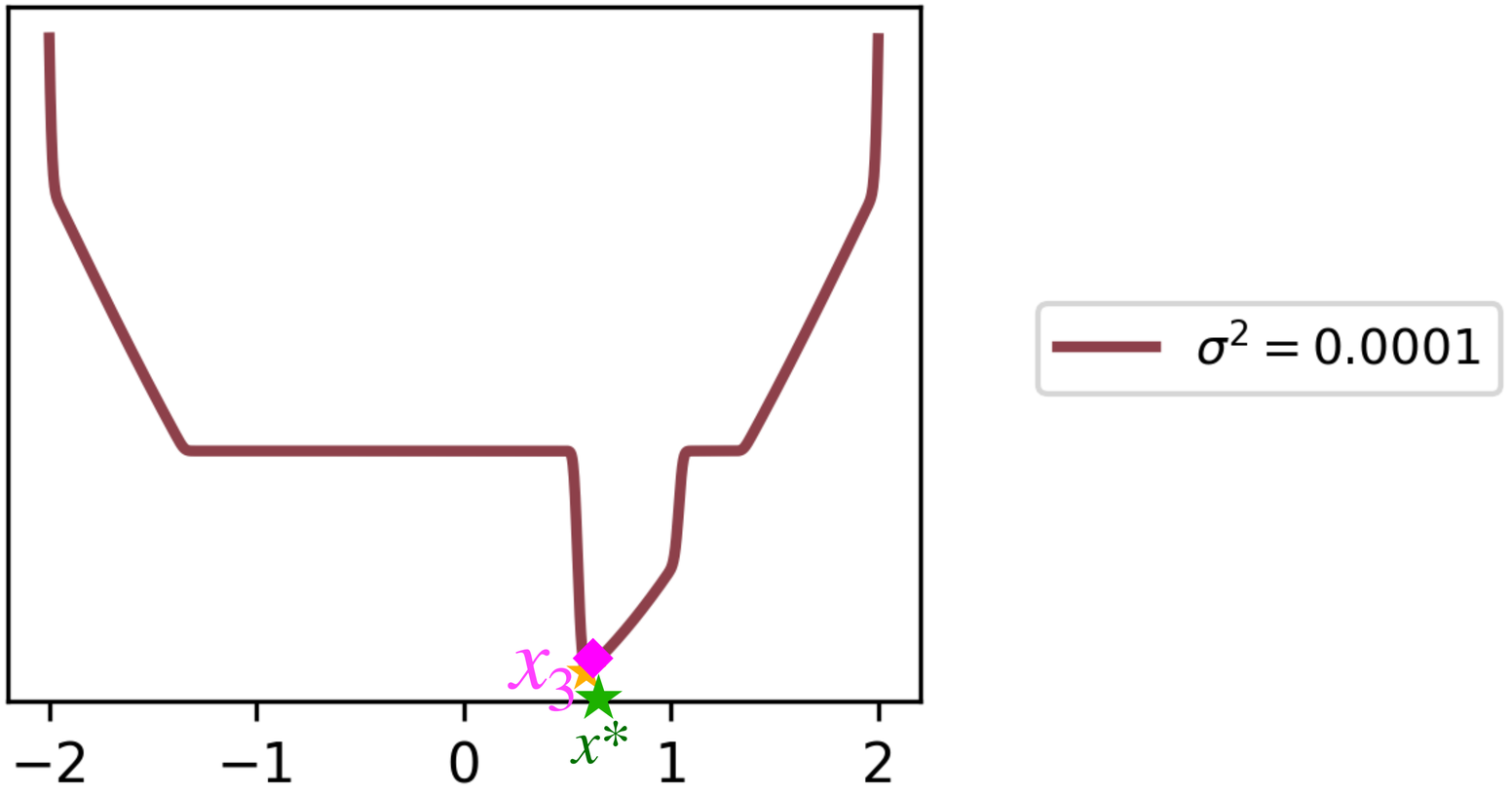
Gradually decrease σ^2, but also make sure x_m stays within convex region

Annealed sampling to ensure global convergence!



Gradually decrease σ^2, but also make sure x_m stays within convex region

Annealed sampling to ensure global convergence!



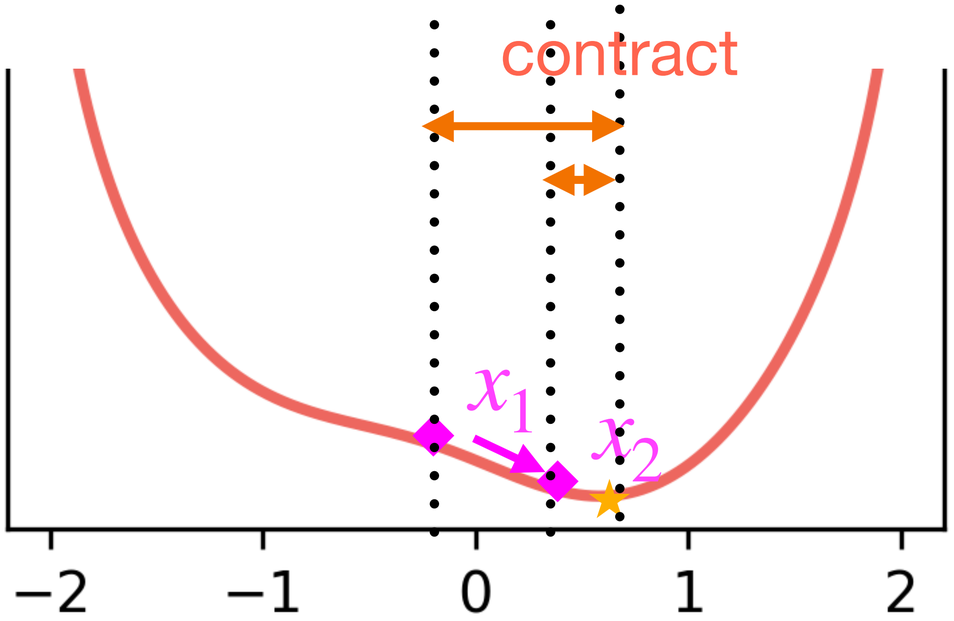
Reaching the global optimizer!

How to select the variance schedule?

Variance should shrink fast to enable faster vanishing of optimality gap…

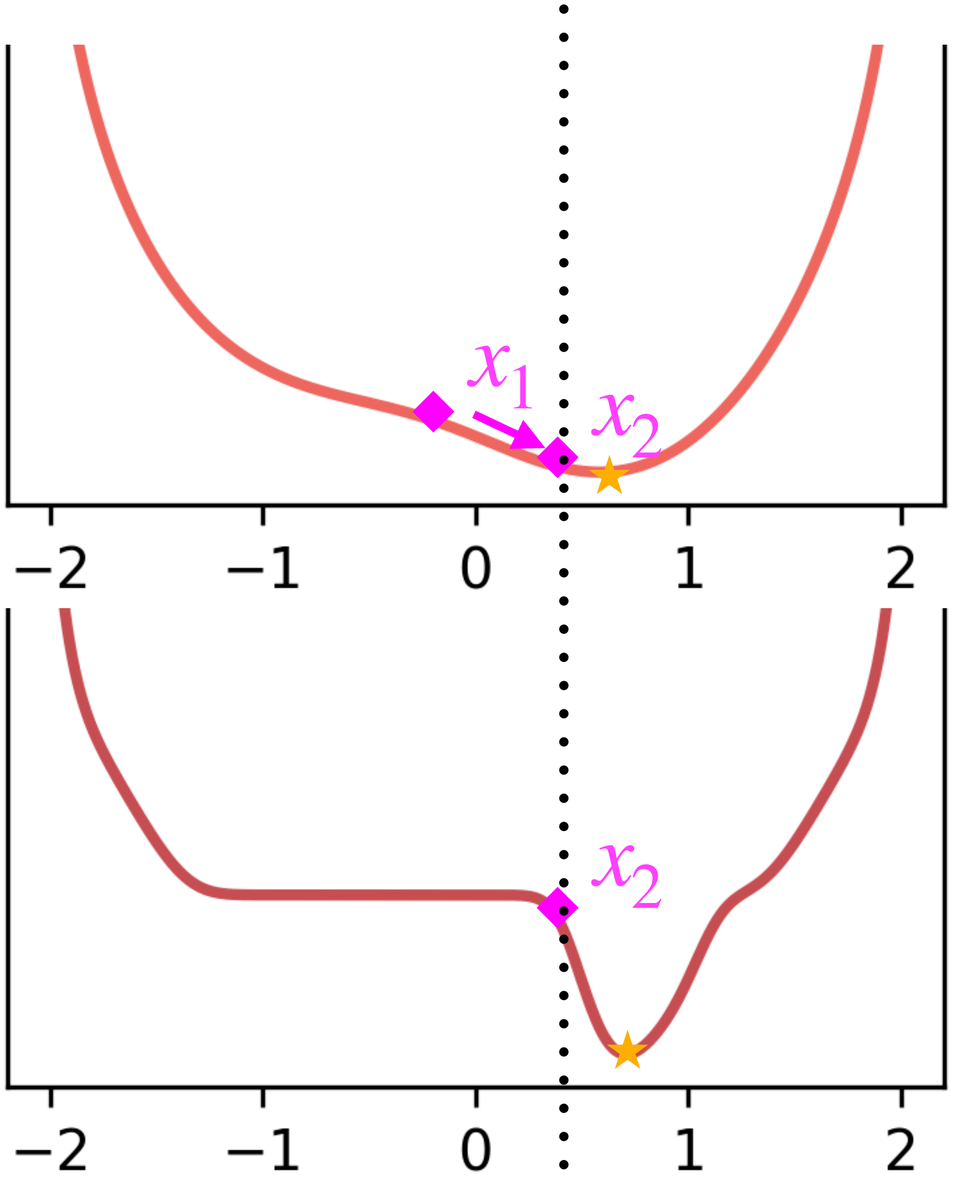
… but not too fast so that iterates **stay inside the convex region**

… but not too fast so that iterates **stay inside the convex region**



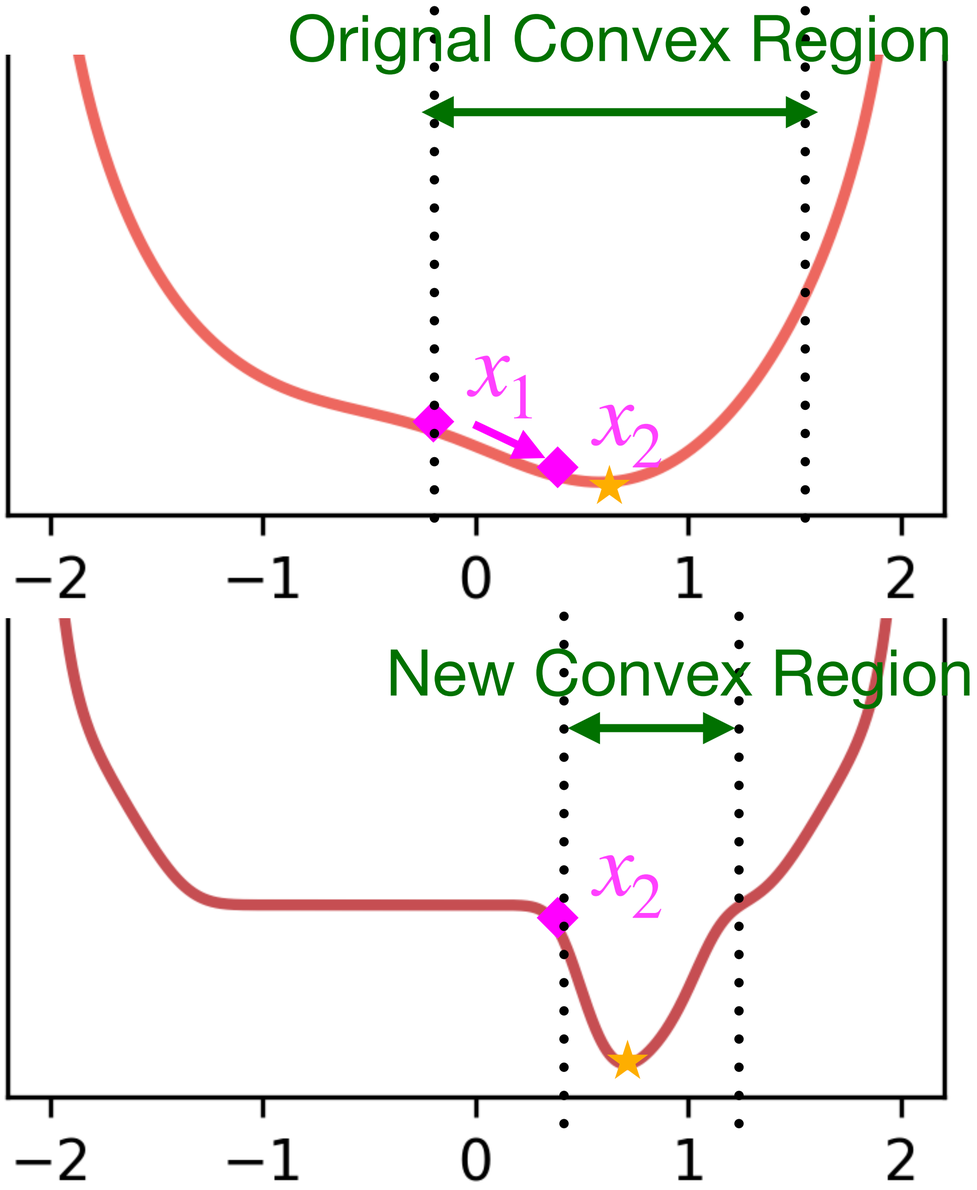
Distance to optimizer approximately “contract” due to strong convexity

… but not too fast so that iterates **stay inside the convex region**



After convex region shrinks, we need x_2 **stay inside the new convex region**

… but not too fast so that iterates **stay inside the convex region**



Convex region should shrink in a similar rate as how x_2 contracts

Convex region radius $\sqrt{\Theta(1) + \Theta(\sigma^2)}$

σ^2 should shrink geometrically

New SBO Algorithms: MBD

Model Based Diffusion

For $m = 0,1,2,...$

Sample $y^{(1)}, y^{(2)}, \ldots, y^{(N)} \sim \mathcal{N}(x_m, \sigma_m^2)$

Annealing $\sigma_m^2 \sim O(\beta^m)$

From coarse to fine sampling

Evaluate $f(y^{(1)}), f(y^{(2)}), \ldots, f(y^{(N)})$

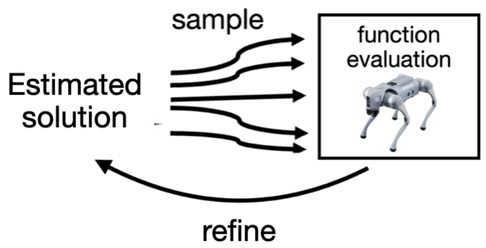
Update $x_{m+1} = \sum_{i=1}^{N} w^{(i)} y^{(i)}$ where $w^{(i)} = \frac{e^{-\frac{1}{\lambda} f(y^{(i)})}}{\sum_{j=1}^{N} e^{-\frac{1}{\lambda} f(y^{(j)})}}$

Empirical Performance for Robot tasks

- Leading on most tasks

Task	CMA-ES	CEM	MPPI	RL	MBD
Hopper	1.12 ± 0.10	0.65 ± 0.12	0.91 ± 0.15	1.40 ± 0.04	$\mathbf{1.53 \pm 0.03}$
Half Cheetah	0.44 ± 0.10	0.22 ± 0.15	0.20 ± 0.14	1.59 ± 0.05	$\mathbf{2.31 \pm 0.19}$
Ant	1.18 ± 0.52	0.85 ± 0.17	0.33 ± 0.45	3.26 ± 1.61	$\mathbf{3.80 \pm 0.35}$
Walker2D	0.83 ± 0.04	1.06 ± 0.04	0.90 ± 0.05	1.09 ± 0.28	$\mathbf{2.63 \pm 0.23}$
Humanoid Standup	0.58 ± 0.01	0.47 ± 0.01	0.53 ± 0.05	0.83 ± 0.02	$\mathbf{0.99 \pm 0.07}$
Humanoid Running	0.60 ± 0.11	0.41 ± 0.16	0.59 ± 0.14	1.80 ± 0.03	$\mathbf{2.92 \pm 0.26}$
Push T	0.39 ± 0.07	0.25 ± 0.09	-0.13 ± 0.09	-0.63 ± 0.16	$\mathbf{0.67 \pm 0.10}$

Table 2: Reward of different methods on non-continuous tasks.



How can SBO effectively navigate the highly non-convex objective?

Key result: coverage and optimality tradeoff

Can we have global convergence guarantee for SBO for non convex problems?

Key result: convergence to neighborhood, when initialized in the enlarged convex region

How to choose hyper parameters optimally?

Key idea: Annealing the variance!

Better algorithm design for robotics

Better Algorithms for Robotics: DIAL-MPC

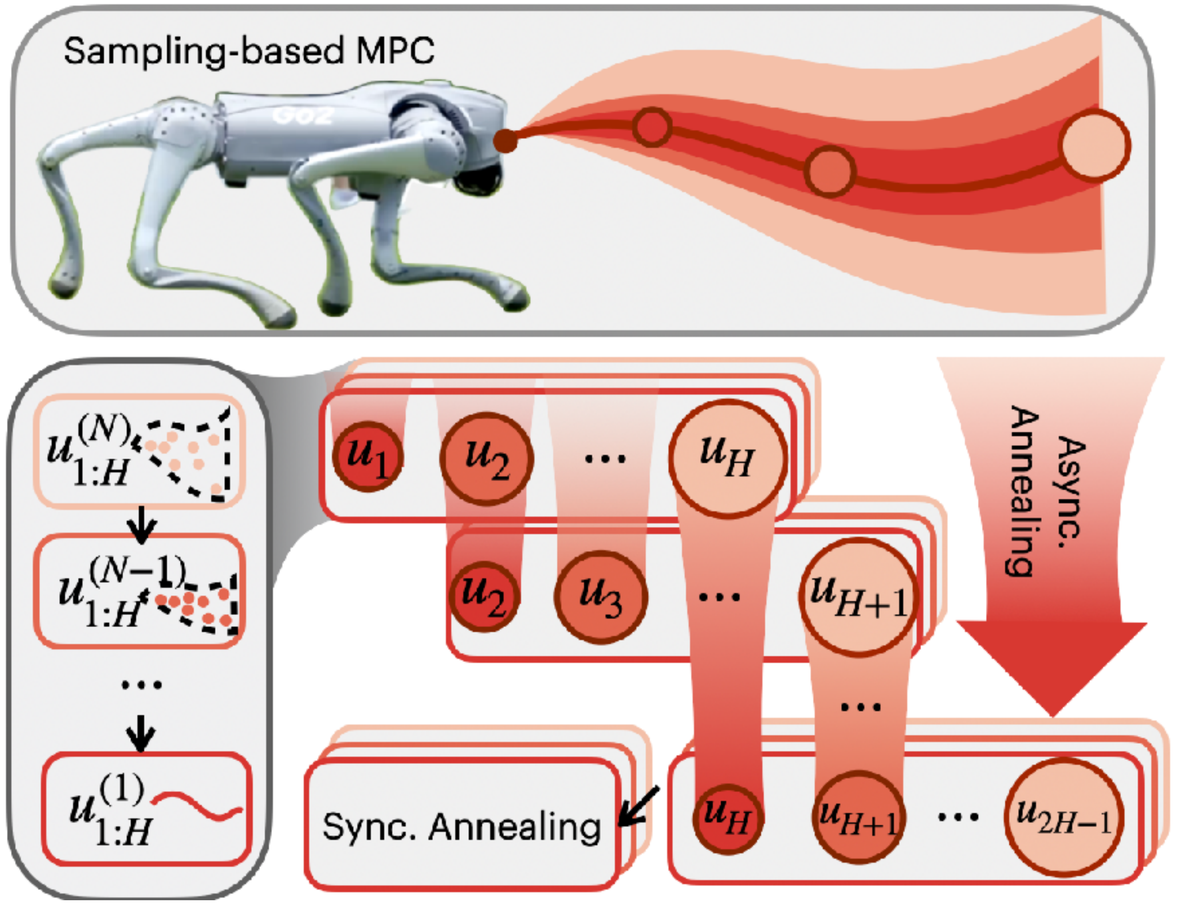
Based on the previous insights, we developed DIAL-MPC

Rough Idea

Receding horizon implementaiton of MBD

Use previous solution (right shift) to initialize current search

Asynchronous annealing



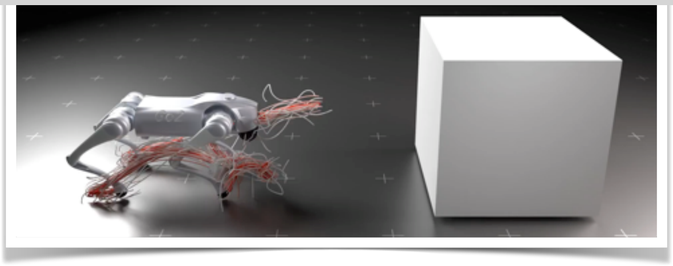
Better Algorithms for Robotics: DIAL-MPC

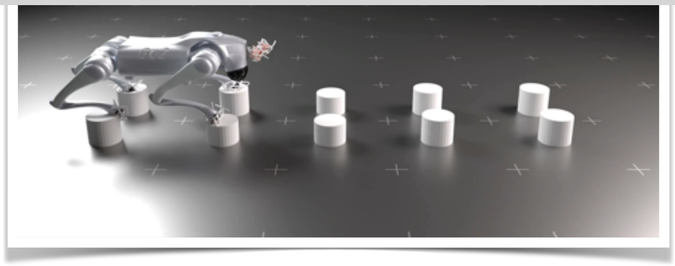
First to enable real-time, training-free control of full-order quadrupeds:

– Precise real-world jumping with payload

– 13.4× lower tracking error than MPPI

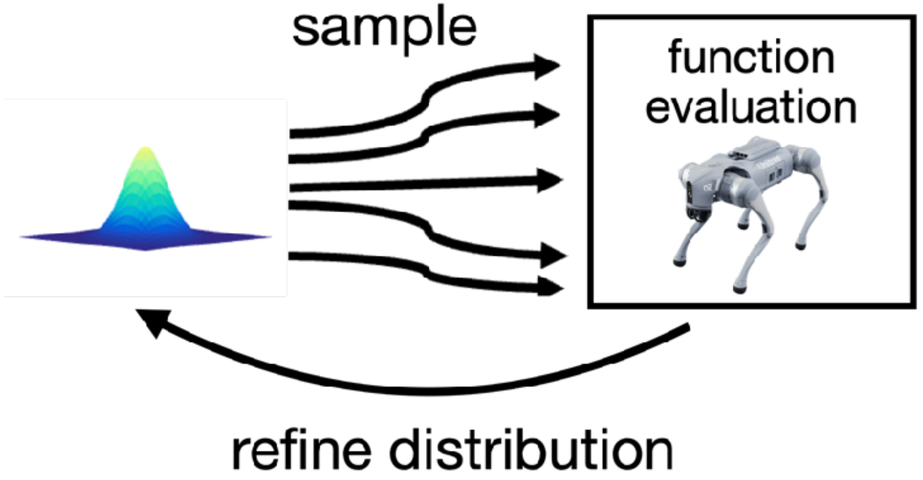
More details of DIAL-MPC will be covered in Part 2 of this tutorial



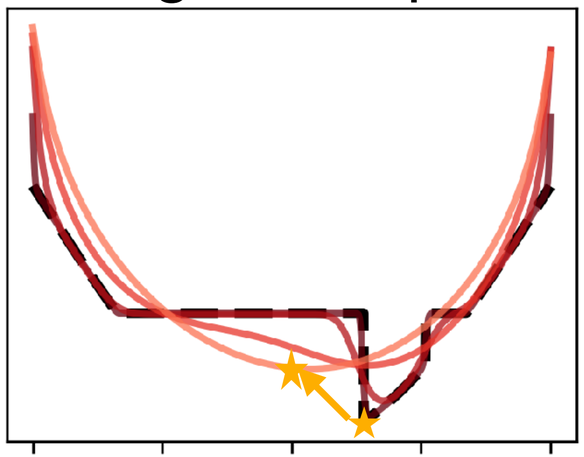


Summary of Part 1

Empirical Success of SBO



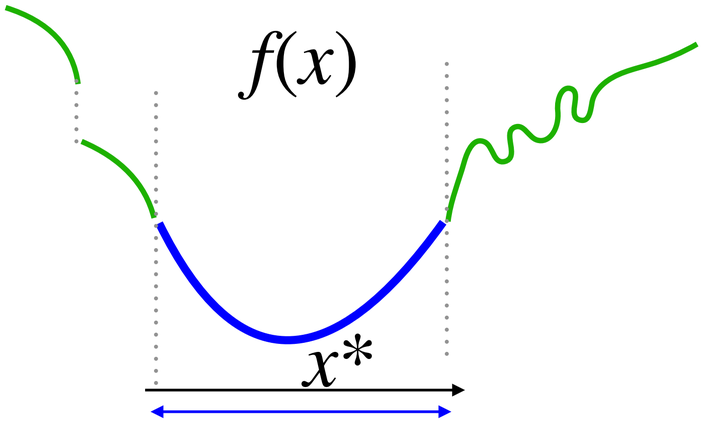
Key theory: coverage and optimality tradeof



Imform better algorithm design for robotics

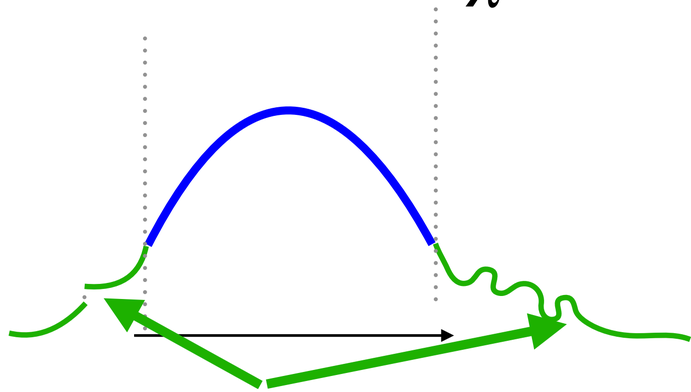
Future Direction

Extending beyond the “dominating basin” regime!



D_τ-ball: α-strongly convex L-smooth

$$p_0(x) \propto \exp(-\frac{1}{\lambda} f(x))$$



Subgaussian outside the ball

Future Direction: Connection to Diffusion

$$p_0(x) \propto \exp(-\frac{f(x)}{\lambda}) \Rightarrow p(x;\sigma^2) = p_0 * \mathcal{N}(0,\sigma^2) \Rightarrow g(x;\sigma^2) = -\lambda \log p(x;\sigma^2)$$

This process is very similar to the diffusion forward process

Proposition

As $N \to \infty$, the update equation converges in probability to:

$$x_{m+1} = x_m - \frac{\sigma^2}{\lambda} \nabla_x g(x_m;\sigma^2)$$

This is (almost) the score function used in diffusion