A Proximal Descent Method for Nonconvex Policy Optimization

Yang Zheng

Assistant Professor,
ECE Department, UC San Diego

62nd Allerton Conference on Communication, Control, and Computing

Sep 17, 2026

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Electrical and Computer Engineering

Scalable Optimization and Control (SOC) Lab
https://zhengy09.github.io/soclab.html

Acknowledgments

Yuto Watanabe
UC San Diego

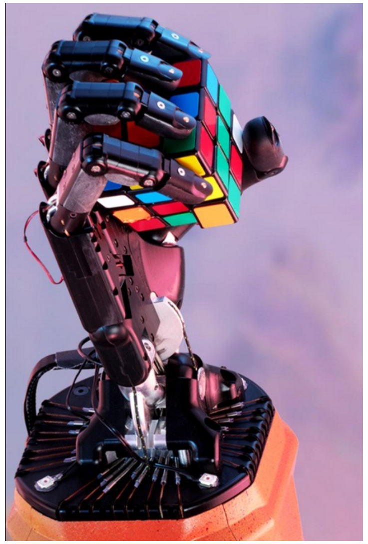
Feng-Yi Liao
UC San Diego

Supported by
NSF ECCS-2154650,
CMMI-2320697,
CAREER-2340713

1. F.Y. Liao, and Y. Zheng. "A proximal descent method for minimizing weakly convex optimization." *arXiv preprint arXiv:2509.02804* (2025).
2. Y. Watanabe, F.Y. Liao, and Y. Zheng. "Weak Convexity and Proximal Bundle Methods for Nonsmooth Policy Optimization in Robust Control." arXiv preprint arXiv:2609.06202 (2026).
3. Y. Watanabe, F.Y. Liao, and Y. Zheng. "Policy optimization in robust control: Weak convexity and subgradient methods." 2026 American Control Conference (ACC). IEEE, 2026.

Data-driven decision making

- Data-driven decision making and Reinforcement learning have achieved striking performance in games, robotics, sequential decision problems, and language models.
- These successes often rely on massive **data**, **computation**, and **flexible policy parameterizations**.
- They shift the design paradigm from "derive a controller" to "optimize a policy".

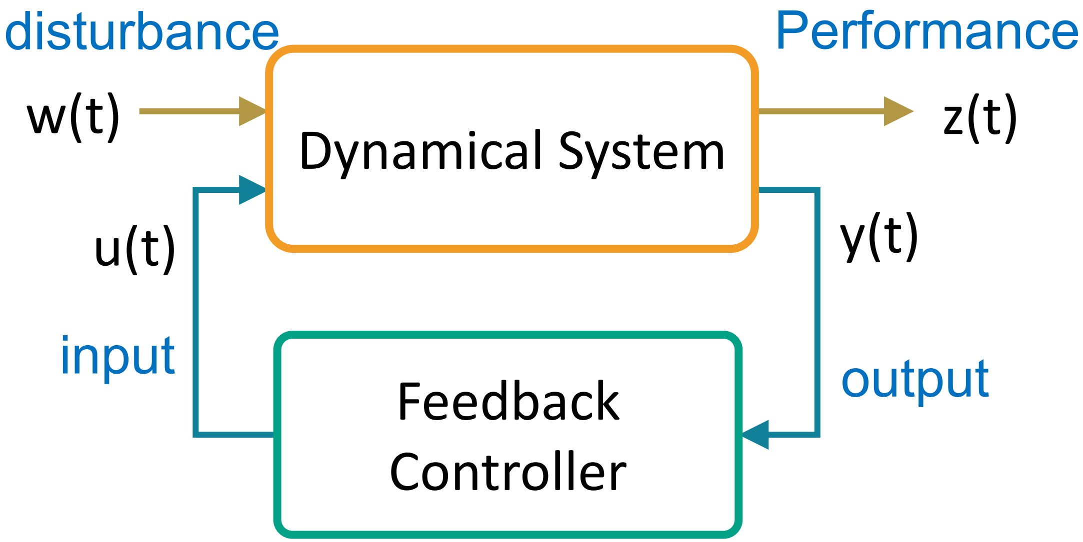


ChatGPT

Silver, et al. 2017; Fazel et al., 2018; Recht, 2019; Hu et al., 2023, etc.

Policy optimization

Feedback control



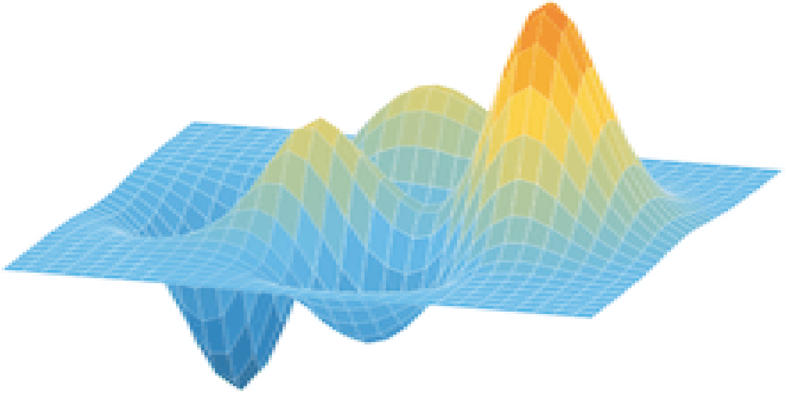
Policy parametrization

Policy Optimization

$$\min_{\mathsf{K}} \; J(\mathsf{K})$$
$$\text{s.t.} \; \mathsf{K} \in \mathcal{C}$$

Use direct policy updates (**model-free**)

$$K_{t+1} = K_t - \alpha_t \nabla J(K_t)$$

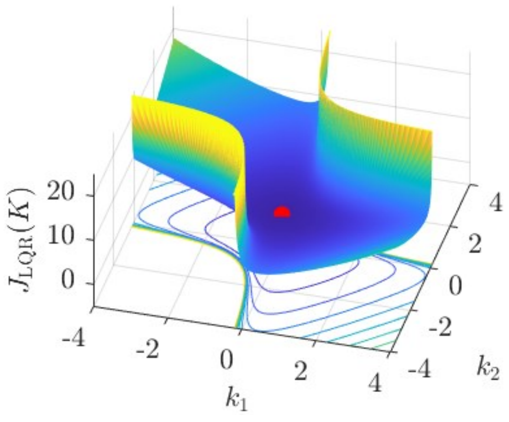


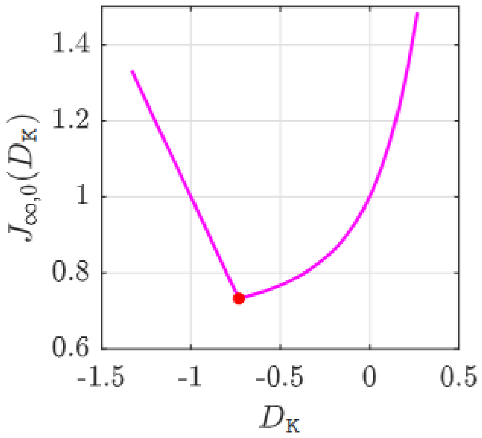
- **Nonconvex optimization**
- Lack of principled algorithms for **optimality**
- Hard to derive **theoretical guarantees**

Today's talk

Benign nonconvex landscapes in benchmark control problems

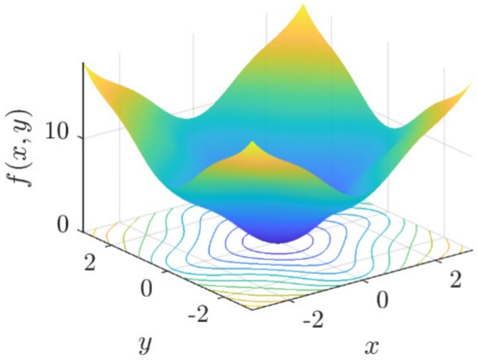
- Informally, any stationary point is **globally optimal** for state-feedback problems.
- Control problems with static policies are often **weakly convex**.

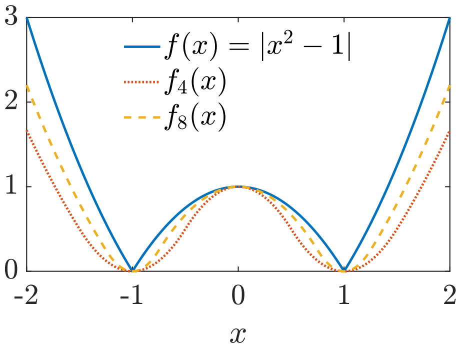




A new proximal descent method for nonconvex policy optimization

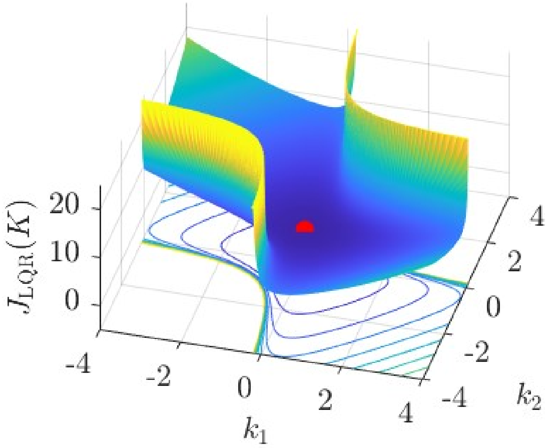
- Thanks to weak convexity, we design a descent local search algorithm based on local convexification.



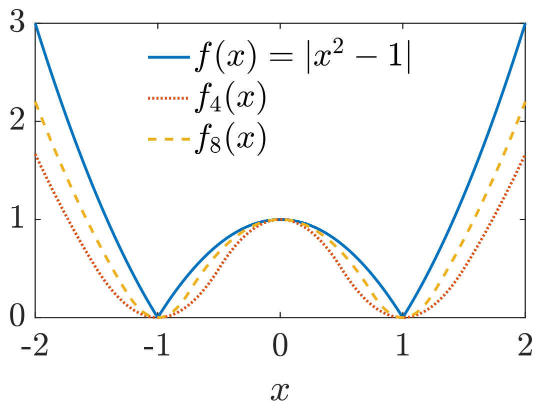


Outline

Geometry:
Benign Nonconvex Landscape



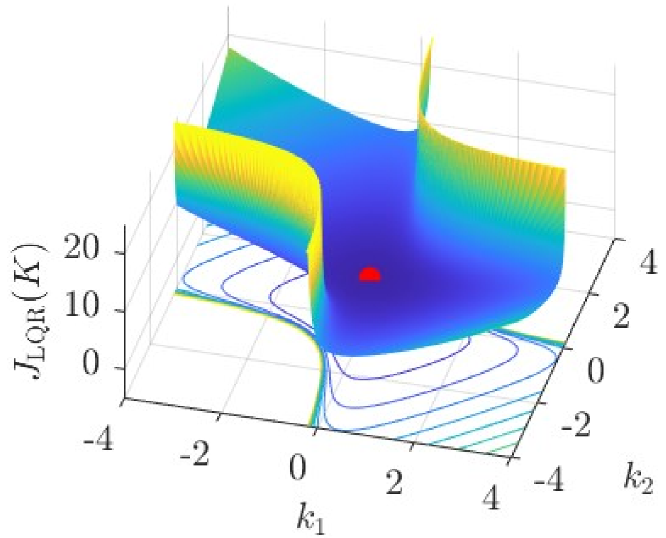
Algorithm:
Proximal Descent Method



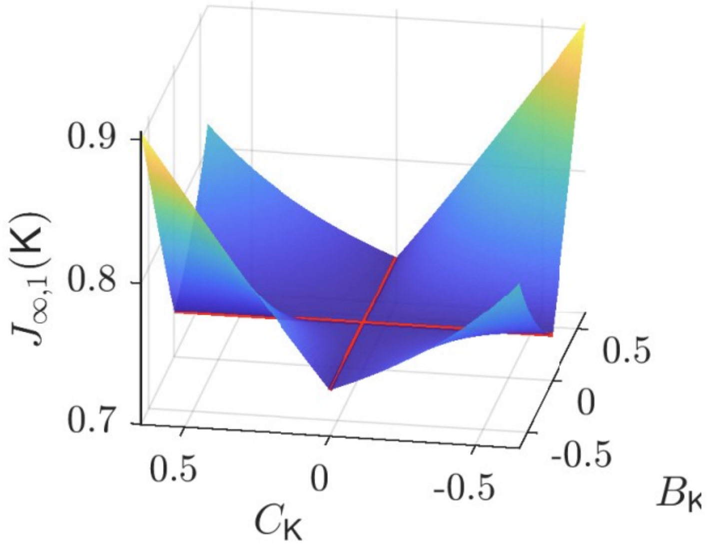
Nonconvexity in control

Policy optimization in control is generally **nonconvex** and may be **non-smooth.**

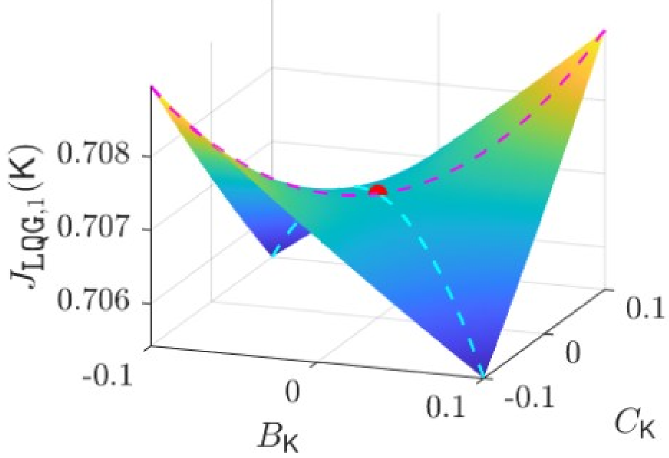
$$\min_{\mathsf{K}} \quad J(\mathsf{K})$$
$$\text{s.t.} \quad \mathsf{K} \in \mathcal{C}$$



An LQR instance
(nonconvex)



An Hinf control instance
(nonsmooth)



An LQG instance
(saddle point)

Optimization landscape in control problems are very rich but not arbitrary.

Global optimality in state-feedback control

(Informal) Any stationary point is **globally optimal** for state-feedback problems.

$$x_{t+1} = Ax_t + Bu_t + w_t,$$
$$u_t = Kx_t$$

	LQR (Stochastic noise)	Hinf control (Adversarial disturbances)
Feasible region	Nonconvex but path-connected	
Stationary point	❖ Globally optimal (**gradient dominance** or PL condition) $J_{\mathrm{LQR}}(K) - J^\star \leq \mu \|\nabla J_{\mathrm{LQR}}(K)\|_F^2$	❖ Globally optimal (**weak PL condition** over a sublevel set) $J_{\inf}(K) - J^\star \leq \mu \|g\|_F, \ \forall g \in \partial J_{\inf}(K)$
Local search	❖ Fast linear convergence for **gradient descent** (like strongly convex problems)	❖ Sublinear convergence for the **subgradient method** (like convex problems)
Refs	Fazel et al., ICML, 2018; Malik et al., 2019; Mohammadi et al., IEEE TAC, 2020; etc.	Gao & Hu, 2022, Watanabe, Liao & **Zheng**, 2026

Weak convexity in output feedback

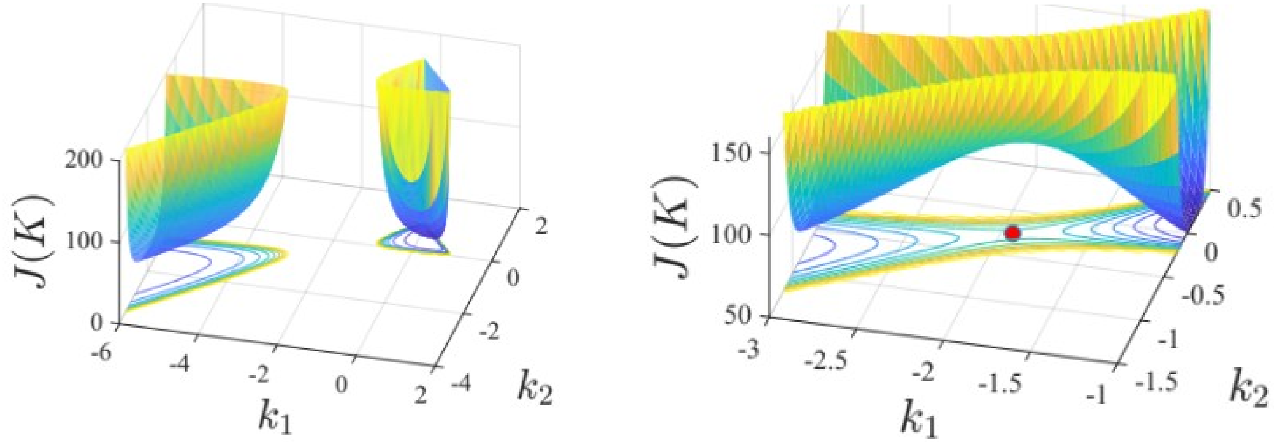
The optimization landscape of static output feedback control is more complicated.

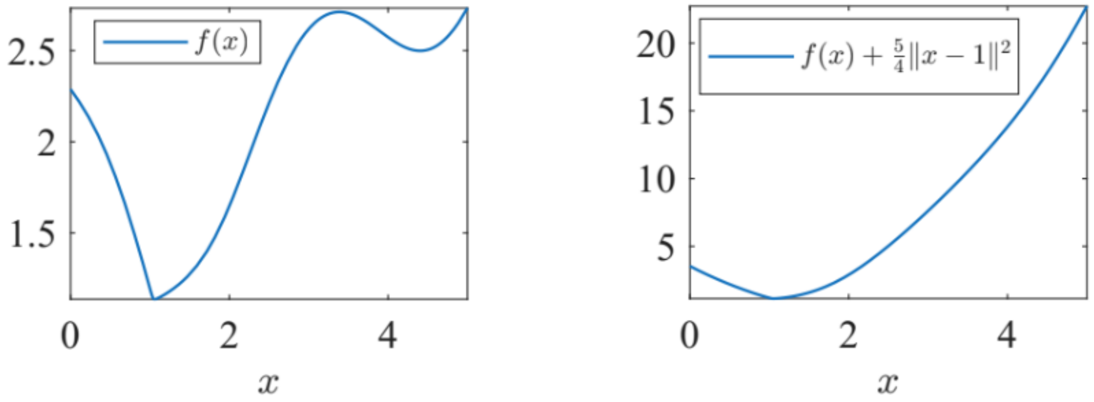
$$x_{t+1} = Ax_t + Bu_t + w_t,$$
$$y_t = Cx_t$$
$$u_t = {\color{red}K} y_t$$

Feasible region	Disconnected
Stationary point	Might be saddle
Cost function (LQR or Hinf)	**Weakly convex** (on any compact sublevel set)

$$J(K) + \frac{m}{2}\|K\|_F^2 \quad \text{is convex}$$

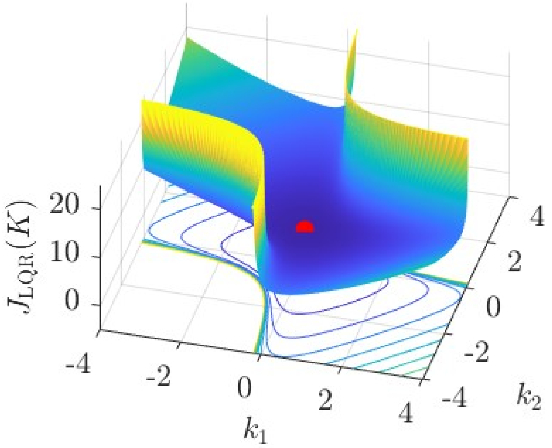
Local convexification with quadratic perturbation



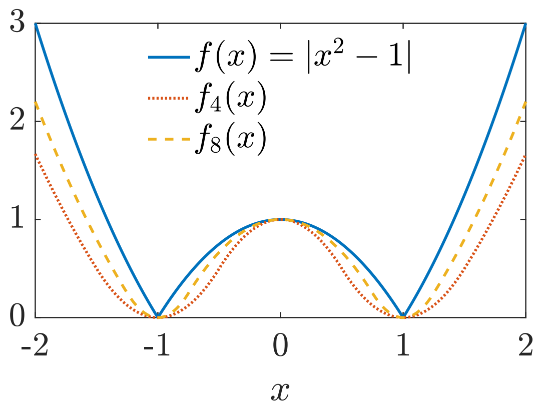


Outline

Geometry:
Benign Nonconvex Landscape

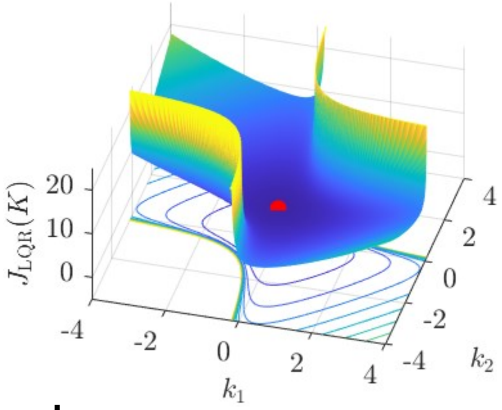


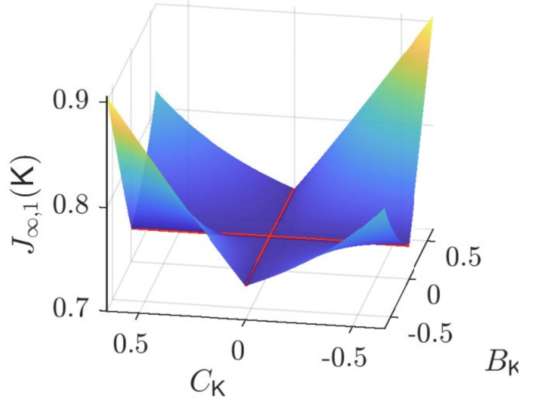
Algorithm:
Proximal Descent Methods



First-order algorithms

$$\min_x f(x)$$





Standard algorithms:

- Gradient descent for L-smooth (non)-convex functions

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k), \qquad 0 < \alpha_k < \frac{2}{L}$$

- Subgradient method for nonsmooth convex functions

$$x_{k+1} = x_k - \alpha_k g_k, \qquad g_k \in \partial f(x_k)$$

Very careful step size tuning

Non-trivial to extend for nonsmooth/nonconvex case

- Their variants (with acceleration) achieve the best first-order complexity for convex functions [Chaps 2 & 3, Nesterov, 2018];

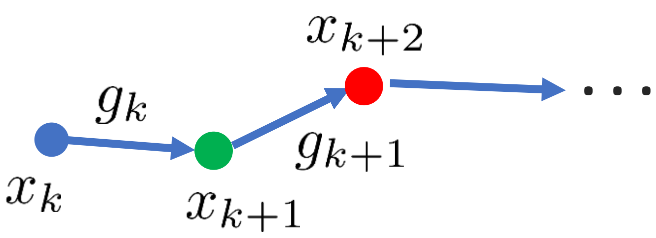
A general descent principle

❑ A conceptual descent algorithm

Search direction

$$x_{k+1} = x_k - \boxed{\alpha} \times \boxed{g_k}$$

Step size



Suppose it satisfies a **cost value drop**

$$f(x_{k+1}) \le f(x_k) - \frac{\alpha}{2}\|g_k\|^2$$

Proof via telescope sum:

$$f(x_T) \le f(x_0) - \frac{\alpha}{2}\sum_{k=0}^{T-1}\|g_k\|^2$$

$$\frac{\alpha}{2}\sum_{k=0}^{T-1}\|g_k\|^2 \le f(x_0) - f(x_T)$$

$$\le f(x_0) - f^\star$$

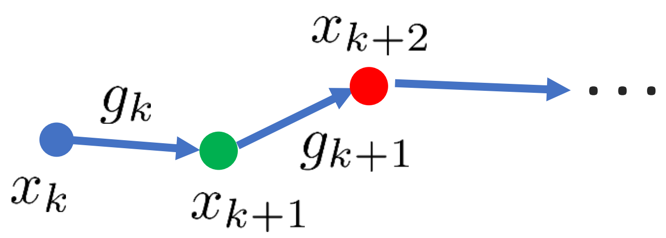
Complexity: the conceptual descent algorithm guarantees

$$\min_{k \le T}\|g_k\|^2 \le \frac{f(x_0) - f^\star}{T} \times \frac{2}{\alpha}$$

Key question: Design g_k so that it both 1) measures stationarity and 2) guarantees sufficient decrease.

Smooth objectives: gradient descent

❑ Gradient descent



$$g_k = \nabla f(x_k)$$

$$x_{k+1} = x_k - \alpha \times \nabla f(x_k)$$

For an L-smooth objective, the **descent lemma** gives [Chap 2, Nesterov, 2018]

$$f(x_{k+1}) \leq f(x_k) - \alpha \left(1 - \frac{L\alpha}{2}\right) \|\nabla f(x_k)\|^2$$

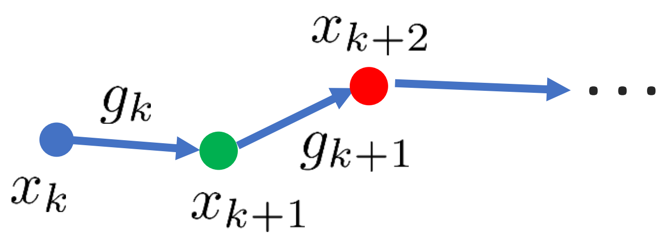
$$\leq f(x_k) - \tfrac{\alpha}{2} \|\nabla f(x_k)\|^2, \qquad 0 < \alpha \leq \tfrac{1}{L}.$$

- $g_k = \nabla f(x_k)$ measures **stationarity**.
- A small stepsize guarantees the **required decrease**.

For nonsmooth objectives, however, a **subgradient step** may not decrease the cost value.

Nonsmooth weakly convex objectives

❑ Proximal point method [Rockafellar, 1976]



- Let f be m-weakly convex, and choose $\alpha < 1/m$

$$x_{k+1} = \arg\min_{x} \left\{ f(x) + \frac{1}{2\alpha}\|x - x_k\|^2 \right\}$$

- The optimality condition gives $g_k = \dfrac{x_k - x_{k+1}}{\alpha} \in \partial f(x_{k+1})$

- **Implicit subgradient** update $x_{k+1} = x_k - \alpha \times \boxed{g_k}$ → a subgradient at the next point

- Comparing the proximal objective at x_{k+1} and x_k gives

$$f(x_{k+1}) \leq f(x_k) - \frac{\alpha}{2}\|g_k\|^2$$

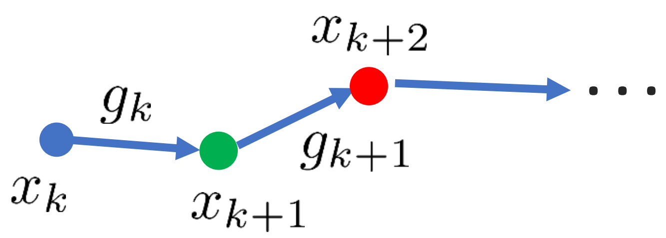
The desired cost value drop

- This works for smooth, nonsmooth, convex and weakly convex objectives.

Descent principle

❑ Gradient descent vs Proximal point update

$$x_{k+1} = x_k - \alpha \times g_k$$



Gradient descent	Proximal point
$g_k = \nabla f(x_k)$	$g_k \in \partial f(x_{k+1})$
Explicit update	Implicit update
L-smooth cost	Weakly convex cost
Both achieve the descent requirement	

Proximal bundle procedure

Use **bundle iterations** to approximate this implicit direction efficiently ([Hare & Sagastizábal, 2009).

$$\min \; f(x) + \frac{1}{2\alpha}\|x - x_k\|^2$$

$$\frac{1}{\alpha} = m + \rho$$

$$\min \; \underbrace{f(x) + \frac{m}{2}\|x - x_k\|^2}_{\text{Convex function}} + \frac{\rho}{2}\|x - x_k\|^2$$

Bundle iterations

❑ Approximate the proximal direction

$$\min \underbrace{f(x)+\frac{m}{2}\|x-x_k\|^2}_{\text{Convex function}}+\frac{\rho}{2}\|x-x_k\|^2$$

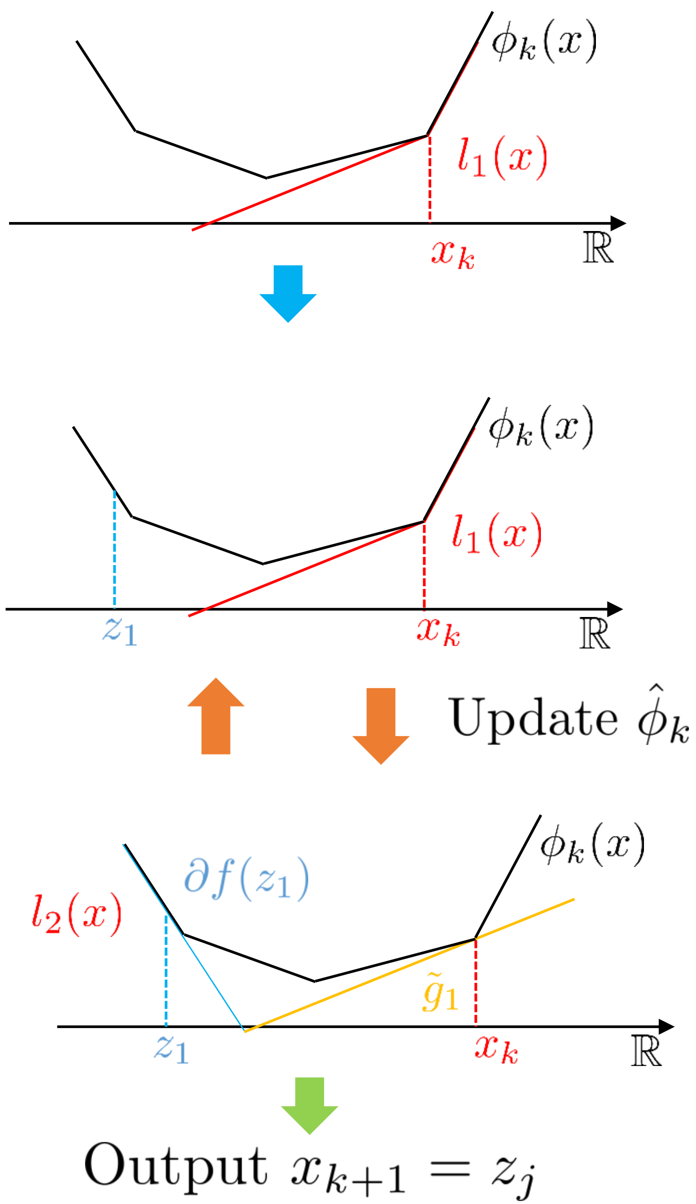
Local convexified function $\phi_k(x) := f(x)+\frac{m}{2}\|x-x_k\|^2$

Every function and subgradient evaluation produces a global affine lower bound of ϕ_k

Intuitively, we can build a cutting-plane approximation

$$\hat{\phi}_k(x) := \max\nolimits_{1\le i\le j} l_i(x) \le \phi_k(x)$$

Testing point $z_{j+1} := \arg\min\ \hat{\phi}_k(x)+\frac{\rho}{2}\|x-x_k\|^2$



Proximal Descent Method

- **Find a search direction efficiently, satisfying** (Liao & **Zheng**, 2025)

$$f(x_{k+1}) \leq f(x_k) - \frac{\tilde{\alpha}}{2}\|g_k\|^2$$

Algorithm 1 Proximal descent method

Require: $x_1 \in \mathbb{R}^n$, $T > 0$,
  for $k = 1, 2, \ldots, T$ **do**
    $x_{k+1} = \texttt{ProxDescent}(x_k)$;
  end for

- Any **convex (nonsmooth)** function
- Any **L-smooth (nonconvex)** function
- **LQR** and **state-feedback robust control** under mild assumptions

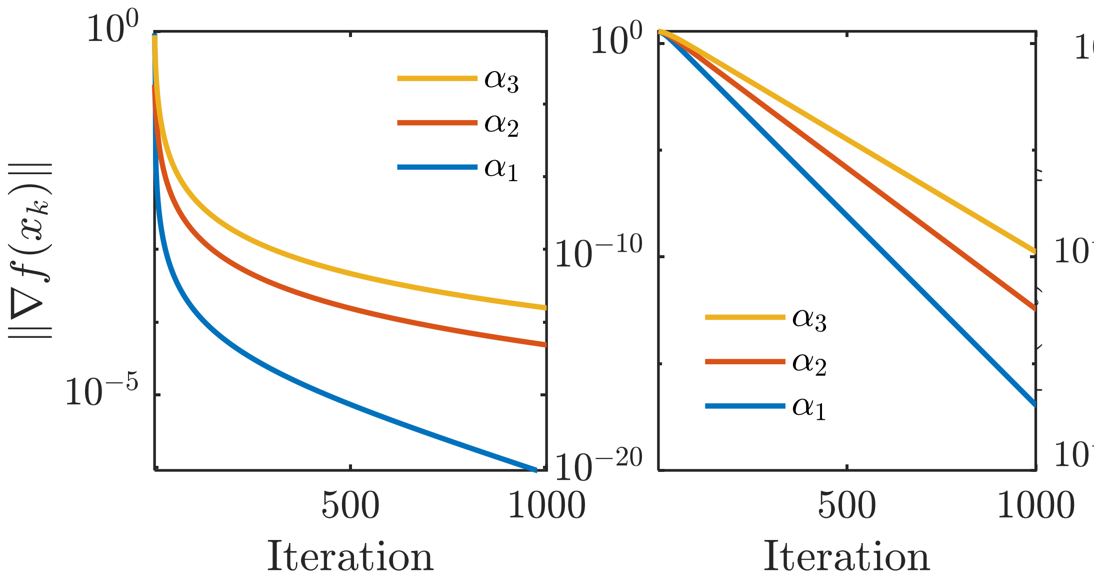
Convergence guarantees

Automatic adaptivity

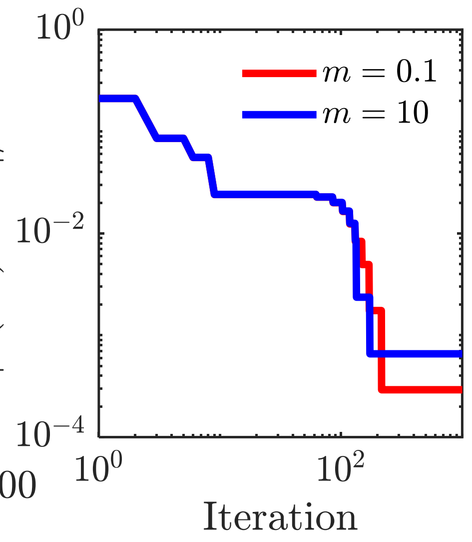
function	Growth	Measure	Complexity
L-Lipchitz	No growth	$\|\nabla f_\alpha(x_k)\| \leq \epsilon$	$\mathcal{O}\left(1/\epsilon^4\right)$
	μ_q-QG and $\mu_q > m$	$\|\nabla f_\alpha(x_k)\| \leq \epsilon$	$\mathcal{O}\left(1/\epsilon^2\right)$
M-Smooth	No growth	$\|\nabla f(x_k)\| \leq \epsilon$	$\mathcal{O}\left(1/\epsilon^2\right)$
	μ_q-QG and $\mu_q > m$	$\|\nabla f(x_k)\| \leq \epsilon$	$\mathcal{O}\left(\log\left(1/\epsilon\right)\right)$

Numerical results

With no parameter tuning, **the algorithm accelerates** in the presence of structural properties

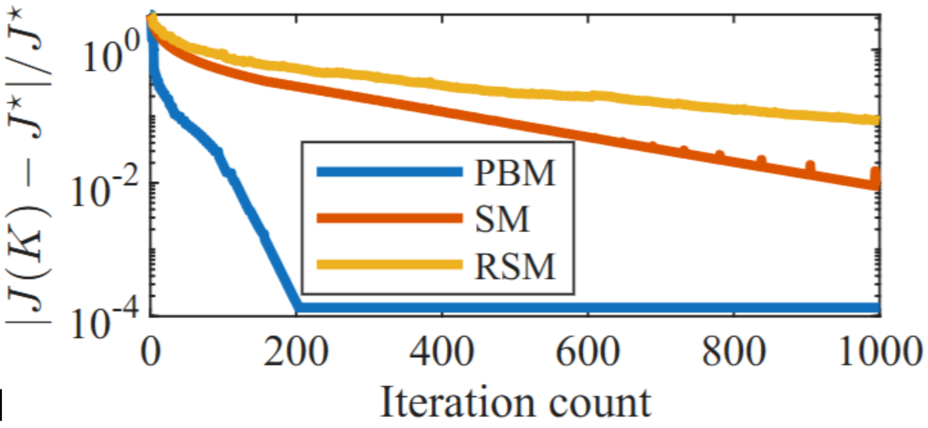


Smooth nonconvex examples

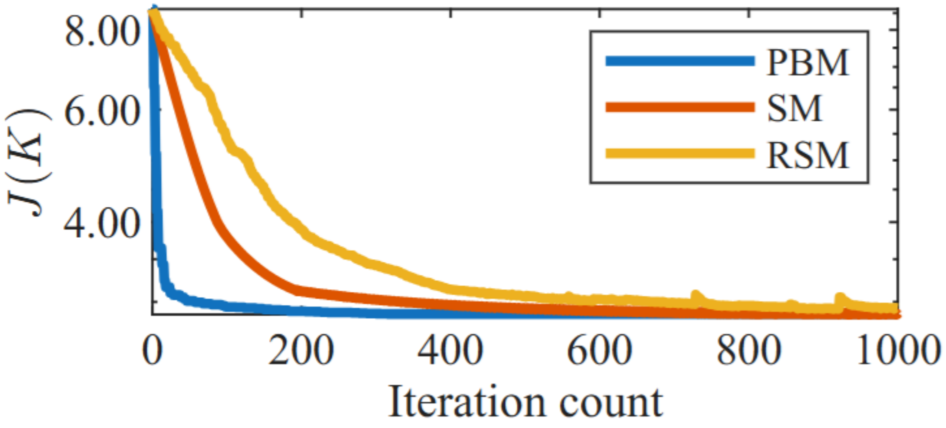


Nonsmooth example

Application to robust control



State-feedback



Output-feedback

PBM: our method

SM: Subgradient method

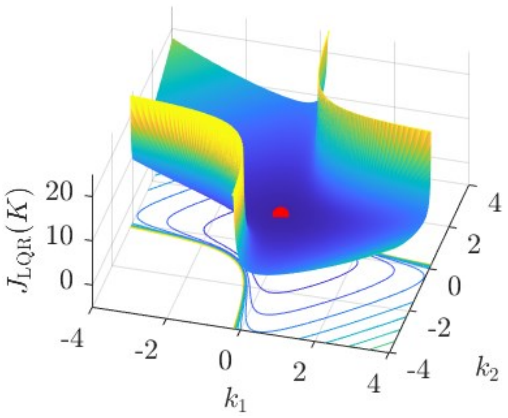
RSM: randomized smoothing

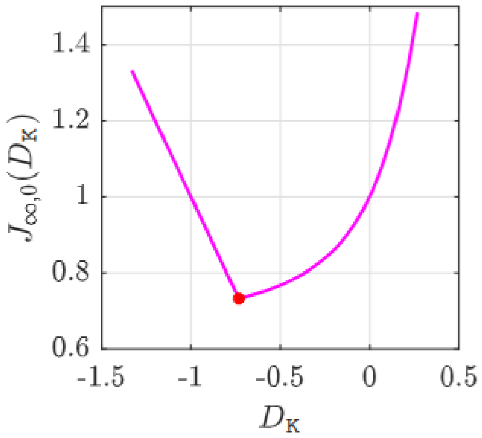
Conclusion

Take-home messages

Nonconvex landscapes in benchmark control problems are benign

- Informally, any stationary point is **globally optimal** for state-feedback problems.
- Control problems with static policies are often **weakly convex**.





A proximal descent method for nonconvex policy optimization

- Weak convexity enables a descent local search algorithm based on local convexification.

Search direction

$$x_{k+1} = x_k - \boxed{\alpha} \times \boxed{g_k}$$

Step size

$$f(x_{k+1}) \leq f(x_k) - \frac{\alpha}{2}\|g_k\|^2$$

Geometry of policy optimization → **Principled & Practical Algorithms**

A Proximal Descent Method for Nonconvex Policy Optimization

Thank you!

1. F.Y. Liao, and Y. Zheng. "A proximal descent method for minimizing weakly convex optimization." *arXiv preprint arXiv:2509.02804* (2025).
2. Y. Watanabe, F.Y. Liao, and Y. Zheng. "Weak Convexity and Proximal Bundle Methods for Nonsmooth Policy Optimization in Robust Control." arXiv preprint arXiv:2609.06202 (2026).
3. Y. Watanabe, F.Y. Liao, and Y. Zheng. "Policy optimization in robust control: Weak convexity and subgradient methods." 2026 American Control Conference (ACC). IEEE, 2026.