Benign Nonconvex Landscapes in Optimal and Robust Control

Yang Zheng

Assistant Professor,
ECE Department, UC San Diego

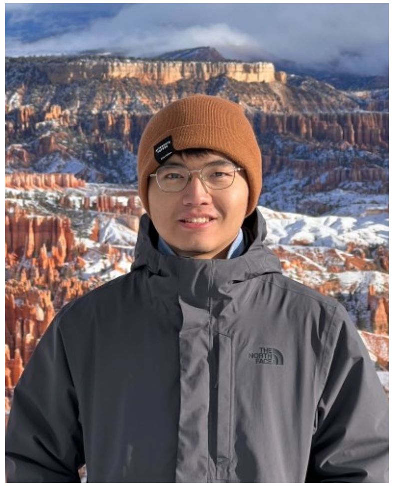
UCSD MAE Dynamic Systems & Controls Seminar

May 23, 2025

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Electrical and Computer Engineering

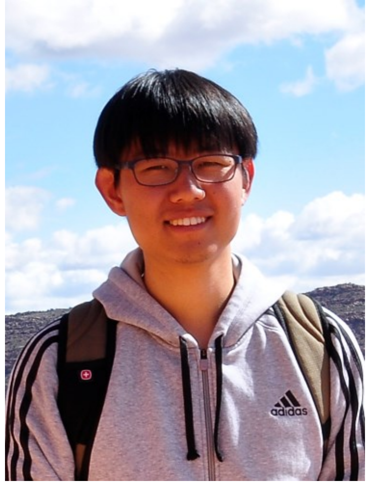
Scalable Optimization
and Control (SOC) Lab
https://zhengy09.github.io/soclab.html

Acknowledgements



Chih-Fan (Rich) Pai
UC San Diego

Yuto Watanabe
UC San Diego



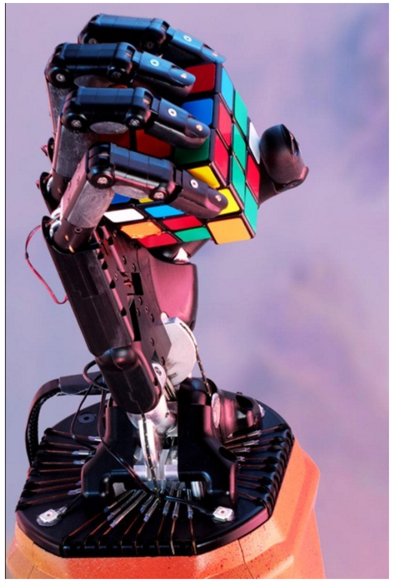
Yujie Tang
Peking University

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

- Zheng, Yang, Chih-Fan Pai, and Yujie Tang. **"Benign Nonconvex Landscapes in Optimal and Robust Control, Part I: Global Optimality."** arXiv preprint arXiv:2312.15332 (2023): https://arxiv.org/abs/2312.15332.
- Zheng, Yang, Chih-Fan Pai, and Yujie Tang. **"Benign Nonconvex Landscapes in Optimal and Robust Control, Part II: Extended Convex Lifting."** arXiv preprint arXiv:2406.04001 (2024): https://arxiv.org/abs/2406.04001
- Watanabe, Yuto, and Yang Zheng. **"Revisiting Strong Duality, Hidden Convexity, and Gradient Dominance in the Linear Quadratic Regulator."** arXiv preprint arXiv:2503.10964 (2025): https://arxiv.org/abs/2503.10964

Success of Data-driven Decision Making

- **Data-driven decision-making** for complex tasks in dynamical systems, e.g., game playing, robotic manipulation/locomotion, networked systems, ChatGPT, etc.
- Reinforcement learning (RL) has served as one backbone of the recent successes of data-driven decision-making.
- Policy optimization as one of the major workhorses of modern RL.

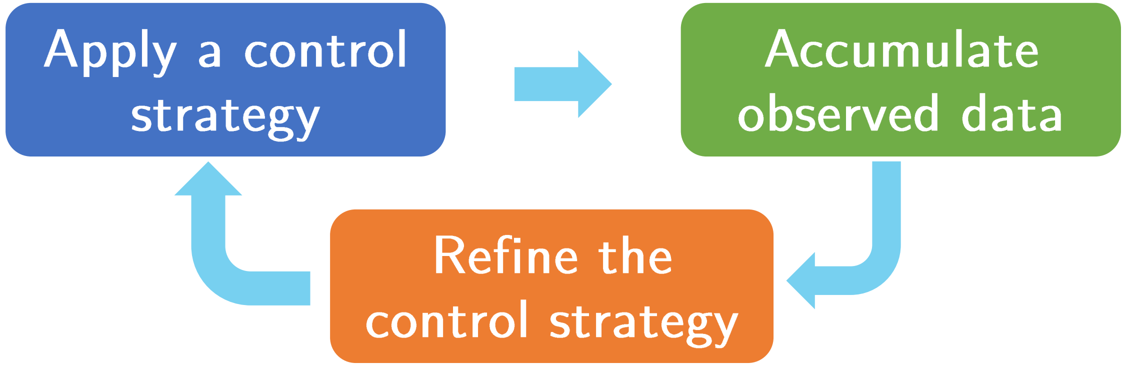


ChatGPT

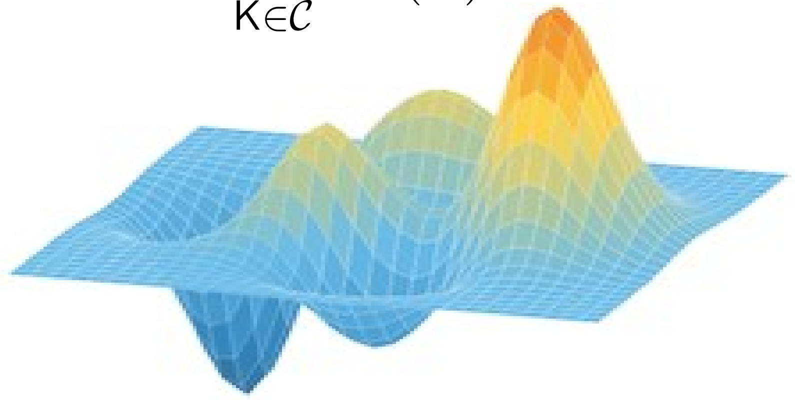
Duan et al. 2016; Silver et al., 2017; Dean et al., 2019; Tu and Recht, 2019; Mania et al., 2019; Fazel et al., 2018; Recht, 2019; https://chat.openai.com/

Policy Optimization for Control

❑ **Why policy optimization is so popular**



$$\min_{\mathsf{K}\in\mathcal{C}} \quad J(\mathsf{K})$$



Opportunities

- Easy-to-implement
- Scalable to high-dimensional problems
- Enable model-free search with rich observations (e.g. images)

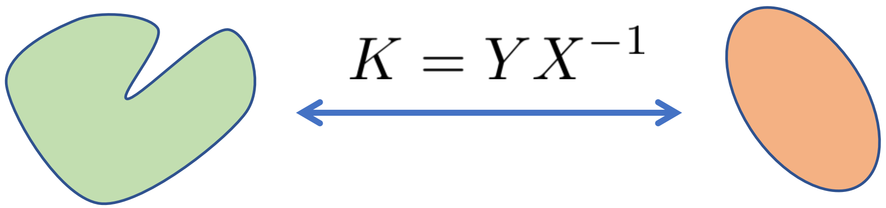
Challenges

- Nonconvex optimization
- Lack of principled algorithms for optimality (e.g., avoiding saddles/local minimizers)
- Hard to obtain theoretical guarantees (e.g., robustness/stability, sample efficiency)

Our Focus: Optimal & Robust Control

Some Historical Background

- **LMI-based convex reformulation**
  - Has became popular since 1980s due to global guarantees and efficient interior point solvers
  - Relies on re-parameterizations (does not optimize over controller/policy directly)



  - Examples: State-feedback or full-order output-feedback $\mathcal{H}_\infty$ / $\mathcal{H}_2$ control, etc.

- **Policy optimization**
  - Has a long history in control theory
    - [Apkarian & Noll, 2006] [Saeki, 2006] [Apkarian et al., 2008] [Gumussoy et al., 2009] [Arzelier et al., 2011], etc.
    - HIFOO, hinfstruct
  - Good empirical performance
    - Scalability, flexibility, …
  - Weak guarantees, unpopular among control theorists

Convex LMIs vs Nonconvex Policy Optimization

❑ Recent progress on non-convex policy optimization

- Favorable properties have been revealed for policy optimization in many benchmark control problems:
  - ✓ LQR [Fazel et al., 2018] [Malik et al., 2020] [Mohammad et al., 2022] [Fatkhullin & Polyak, 2021], etc.
  - ✓ LQG [Zheng, Tang & Li, 2021], [Mohammadi et al., 2021] [Zheng et al., 2022], [Ren et al., 2023]]
  - ✓ $\mathcal{H}_\infty$ state-feedback/output-feedback, [Guo & Hu, 2022] [Hu & Zheng, 2022]

ANNUAL REVIEW OF CONTROL, ROBOTICS, AND AUTONOMOUS SYSTEMS Volume 6, 2023

Review Article | Open Access

**Toward a Theoretical Foundat
Learning Control Policies**

Bin Hu[1], Kaiqing Zhang[2,3], Na Li[4], Mehran Mesbahi[5], Maryam F

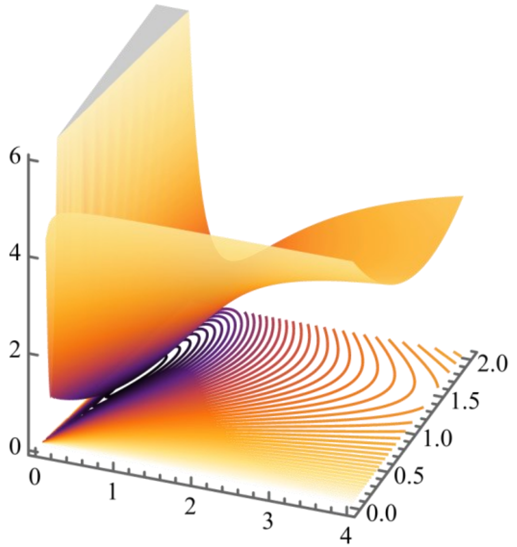
Policy Optimization in Control: Geometry and Algorithmic Implications

Shahriar Talebi[a], Yang Zheng[b], Spencer Kraisler[c], Na Li[a], Mehran Mesbahi[c]

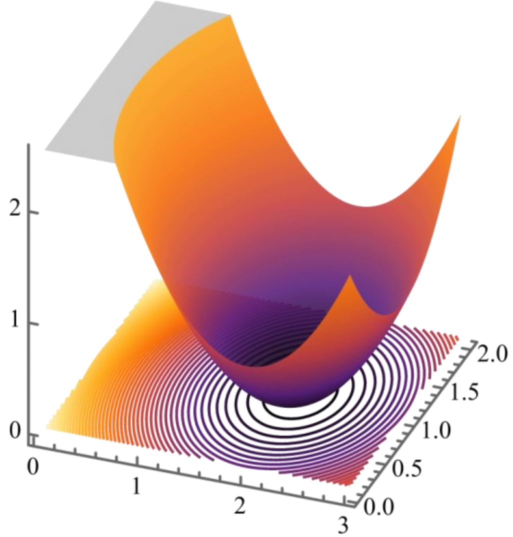
[a] *Harvard University, School of Engineering and Applied Sciences, 150 Western Ave, Boston, 02134, MA, US*
[b] *University of California San Diego, Department of Electrical and Computer Engineering, 9500 Gilman Drive, La Jolla, 92093, CA, US*
[c] *University of Washington, Department of Aeronautics and Astronautics, 3940 Benton Ln NE, Seattle, 98195, WA, US*

This Talk

Benign Nonconvexity in Control via Extended Convex Lifting (ECL)



Nonconvex policy optimization ⟷ **LMI-based convex reformulation**

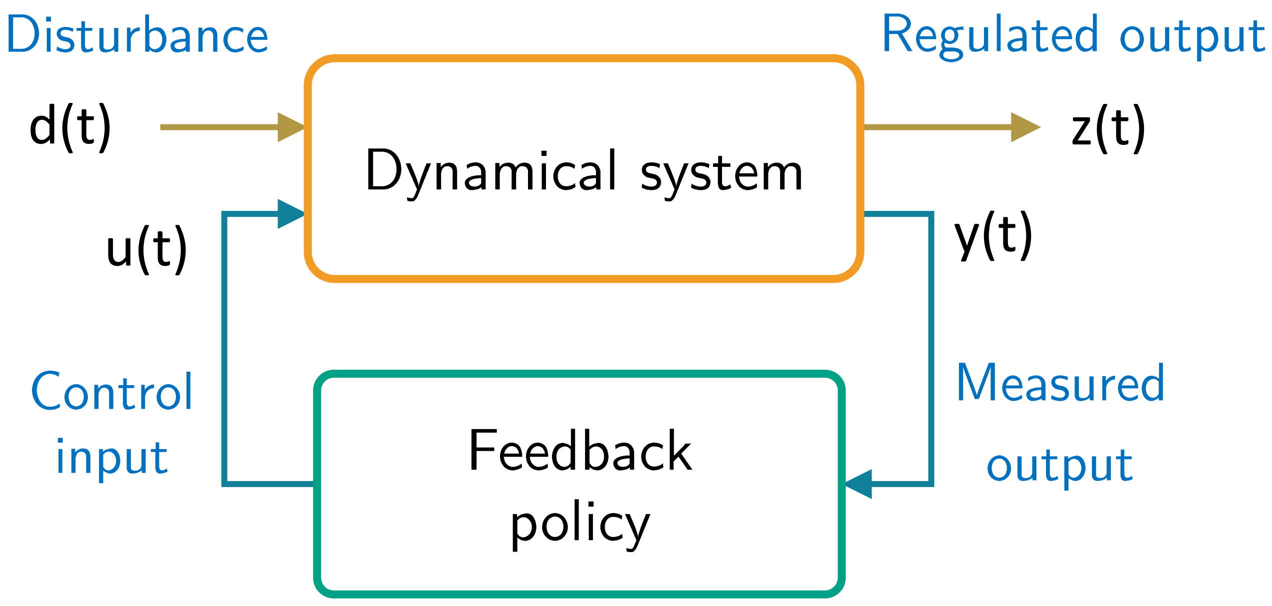


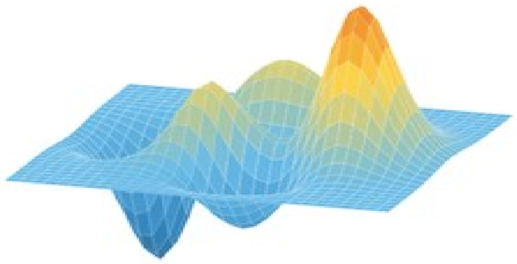
- Reconciles the gap between **nonconvex policy optimization** and **LMI-based convex reformulations**.
- For **non-degenerate** policies, **all Clarke stationary points are globally optimal** and there is **no spurious local minimum** in policy optimization.

Outline

- ❑ Problem Setup and Motivating Examples
- ❑ Extended Convex Lifting (ECL)
- ❑ ECLs for Optimal and Robust Control
- ❑ Escaping Degenerate Saddle Points

Policy Optimization in Control





Policy parametrization

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

Non-convex Optimization problem

System dynamics

$$\dot{x}(t) = Ax(t) + Bu(t) + B_w w(t)$$
$$y(t) = Cx(t) + D_v v(t)$$

Performance signal

$$z(t) = \begin{bmatrix} Q^{1/2}x(t) \\ R^{1/2}u(t) \end{bmatrix}$$

State feedback

$$u(t) = Kx(t)$$

Output feedback

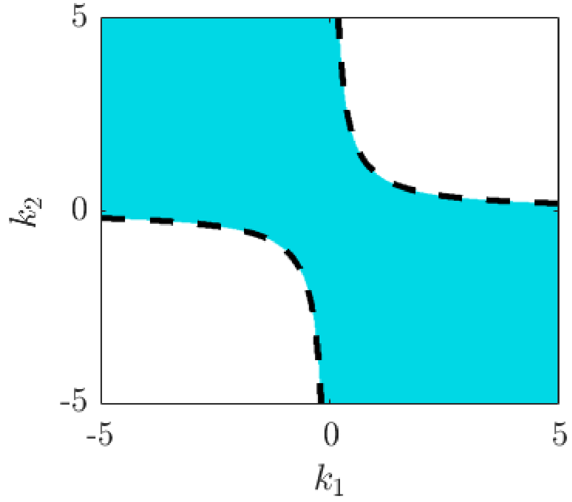
$$\dot{\xi}(t) = A_{\mathsf{K}}\xi(t) + B_{\mathsf{K}}y(t)$$
$$u(t) = C_{\mathsf{K}}\xi(t)$$

$$\mathcal{C} = \{\mathsf{K} : \text{Closed-loop system is stable}\}$$

Nonconvexity in Policy Optimization

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

Policy optimization is generally **nonconvex**!



- The basic problem of **stabilization** is **non-convex**

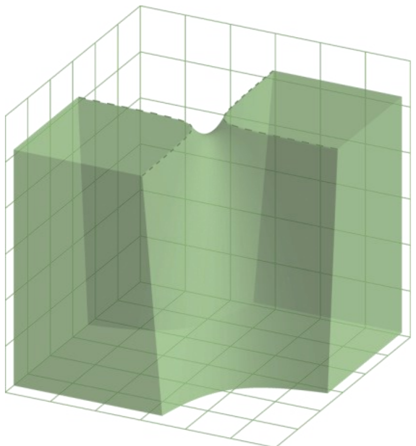
A simple example: $A = 0, \quad B = I_2$

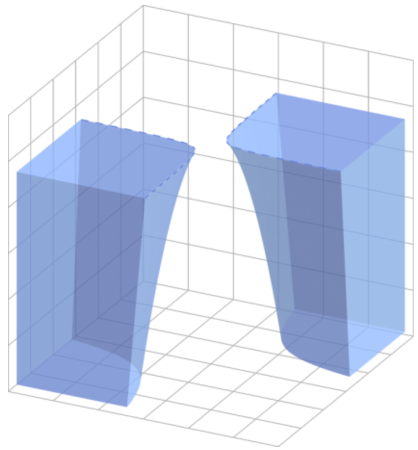
$$\mathcal{C} = \{K \in \mathbb{R}^{2\times 2} \mid A + BK \text{ is stable}\}$$

$$K_1 = \begin{bmatrix} -1 & 2 \\ 0 & -1 \end{bmatrix} \in \mathcal{C}, \quad K_2 = \begin{bmatrix} -1 & 0 \\ 2 & -1 \end{bmatrix} \in \mathcal{C}, \qquad \frac{1}{2}(K_1 + K_2) = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \notin \mathcal{C}$$

- The set of dynamic stabilizing policies is **nonconvex** and may even be **disconnected**.

[Tang, Zheng, Li, 2023]





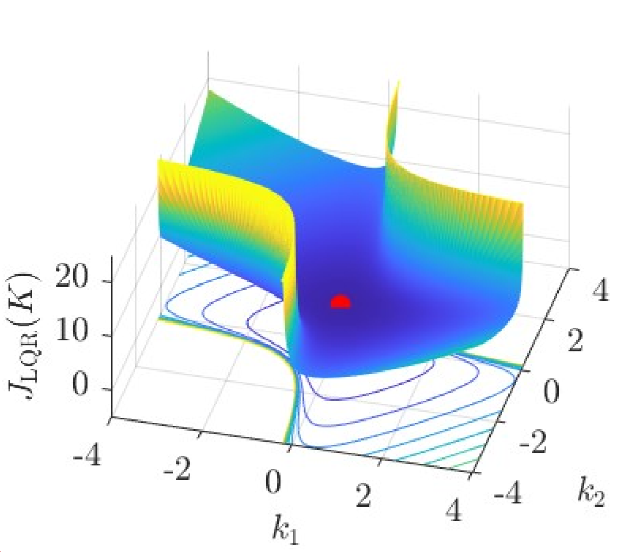
Nonconvexity in Policy Optimization

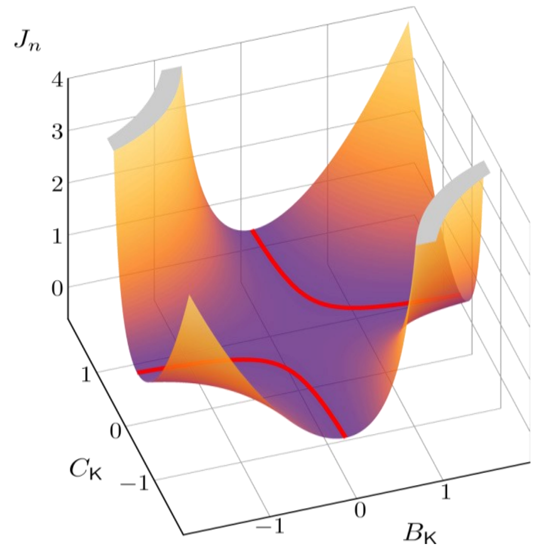
$$\min_{\mathsf{K}} \quad J(\mathsf{K})$$
$$\text{s.t.} \quad \mathsf{K} \in \mathcal{C}$$

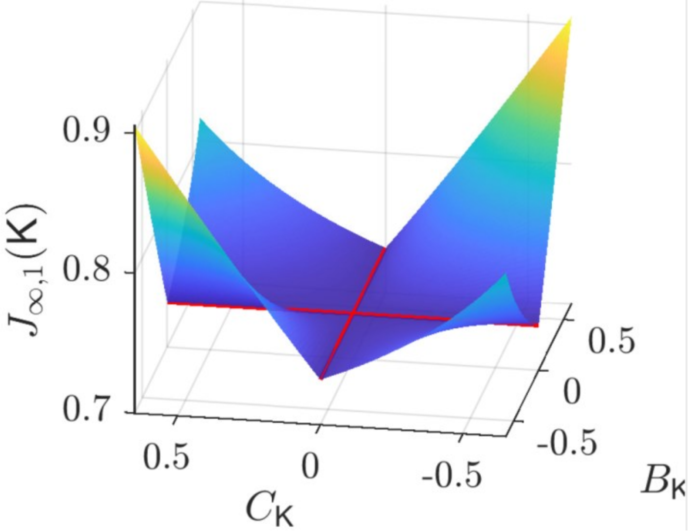
Policy optimization is generally **nonconvex**!

- The costs of Linear Quadratic Regulator (LQR)/LQG costs are **smooth** but **nonconvex**

- The cost of $\mathcal{H}_\infty$ robust control are **non-smooth** and **nonconvex**







Highly non-trivial to establish theoretical guarantees!

A very basic question:

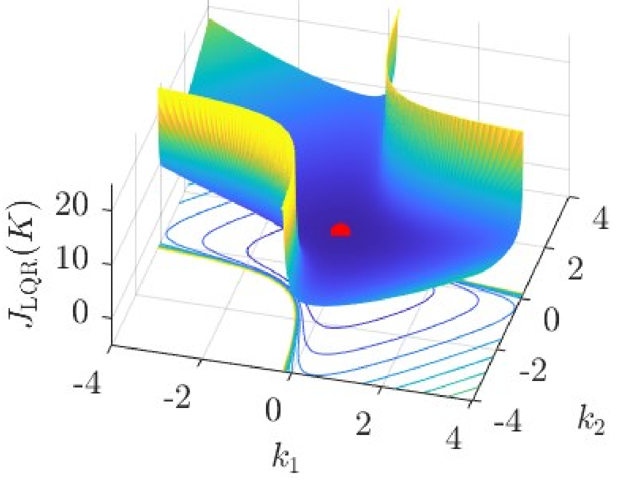
When is a stationary point globally optimal?

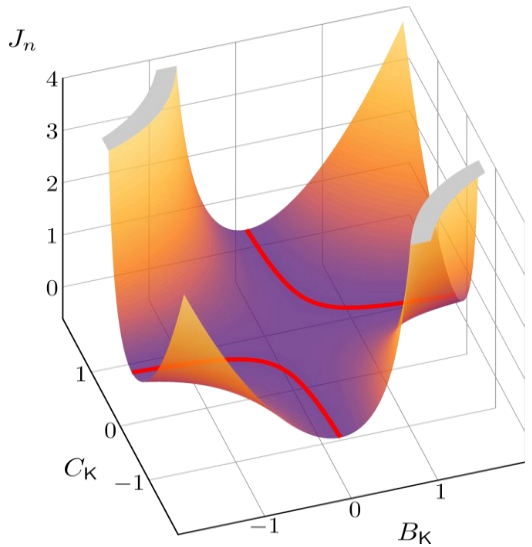
Benign Nonconvex Landscape

Policy parametrization

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

Non-convex Optimization problem





Question: When is a stationary point globally optimal?

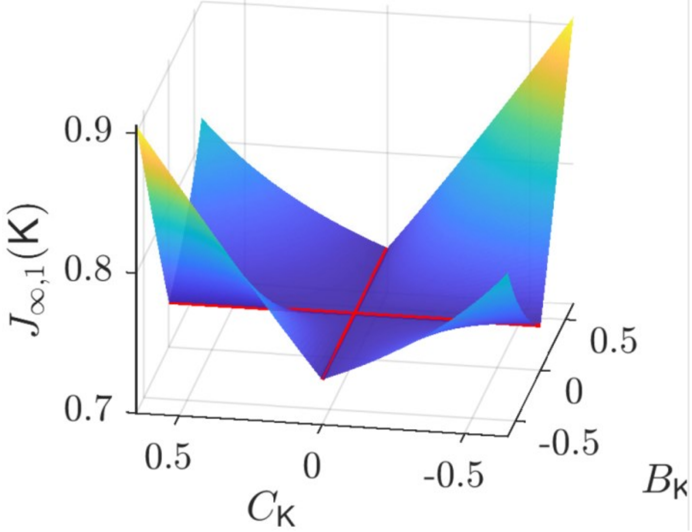
Answer: Any (non-degenerate) Clarke stationary points are globally optimal!

Local Stationarity + Structural Information

Global Optimality Certificate

Our tool:

Extended Convex Lifting



Inspirations of Convex Reformulation

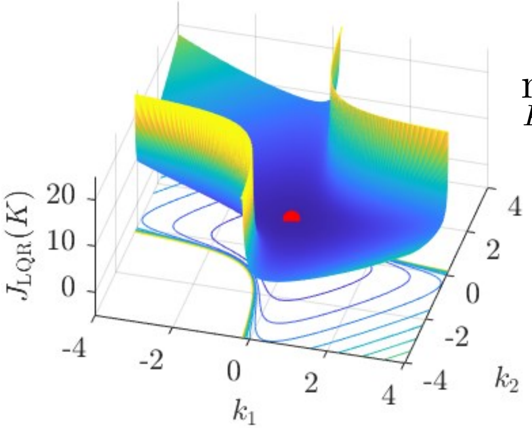
Policy parametrization

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

Non-convex Optimization problem

Our idea: Exploit **LMI-based convex reformulations** of control problems

- They reveal the hidden convexity of policy optimization landscapes

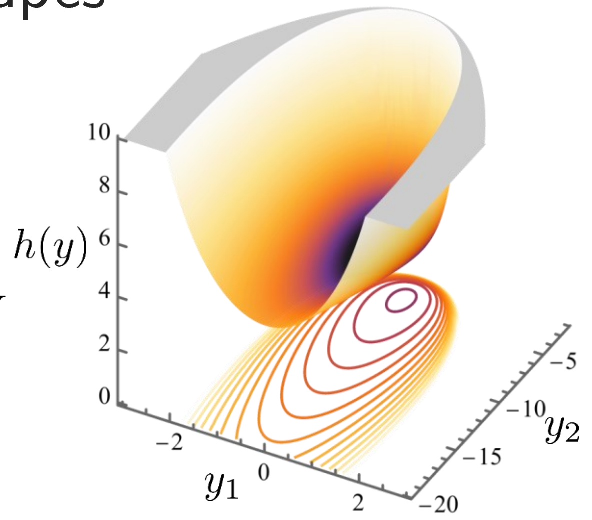


$$\min_{K,X} \quad \operatorname{tr}\left[(Q + K^{\mathsf{T}} R K) X\right]$$
$$\text{s.t.} \quad X = \operatorname{Lyap}(A + BK, W)$$
$$X \succ 0$$

$Y = KX$

Change of variable

$$\min_{X,Y} \quad \operatorname{tr}\left(Q + X^{-1} Y^{\mathsf{T}} R Y\right)$$
$$\text{s.t.} \quad 0 = AX + BY + XA^{\mathsf{T}} + Y^{\mathsf{T}} B^{\mathsf{T}} + W$$
$$X \succ 0$$

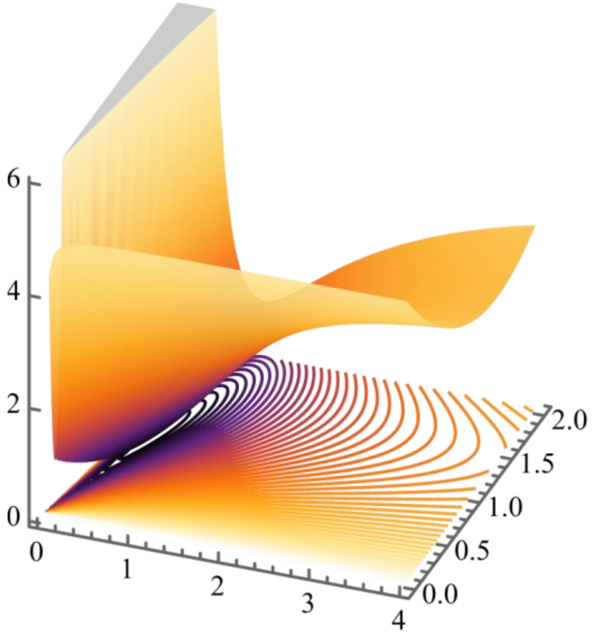


Example 1

❑ Nonconvex and Smooth Function

$$f_1(x_1, x_2) = \left(\frac{x_2}{x_1} - 2\right)^2 + (x_2 - 1)^2, \qquad \mathrm{dom}(f_1) = \{x \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0\}.$$

y_1 y_2

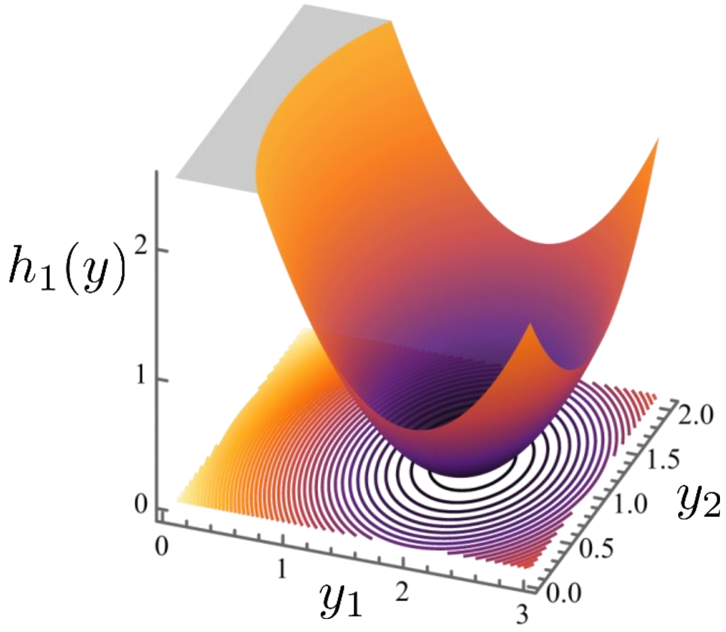


Its global **minimizer** is

$$x^* = (0.5, 1)$$

Define an **invertible map**

$$g(x) := (x_2/x_1, x_2),$$
$$\forall x_1 > 0, x_2 > 0,$$



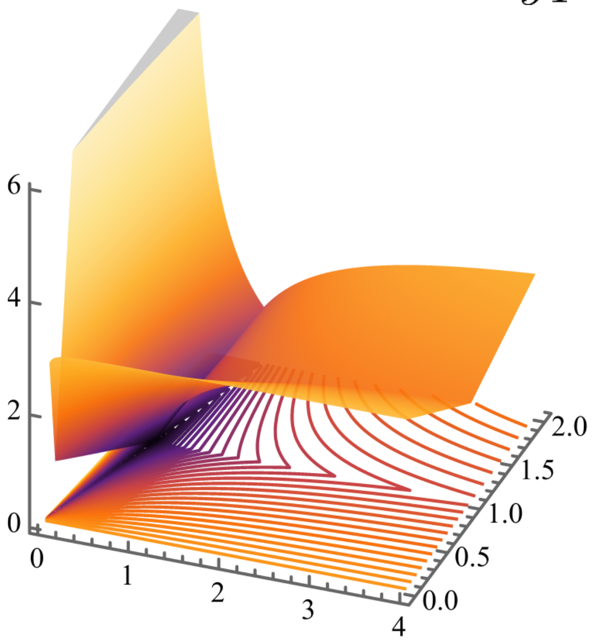
$$h_1(y) := f_1(g^{-1}(y)) = (y_1 - 2)^2 + (y_2 - 1)^2, \qquad \forall y_1 > 0, y_2 > 0.$$

Example 2

❑ Nonconvex and Non-smooth Function

$$f_2(x_1, x_2) = \left|\frac{x_2}{x_1} - 2\right| + |x_2 - 1|, \qquad \mathrm{dom}(f_2) = \{x \in \mathbb{R}^2 \mid x_1 > 0, x_2 > 0\}.$$

$\frac{x_2}{x_1} \to y_1$, $x_2 \to y_2$

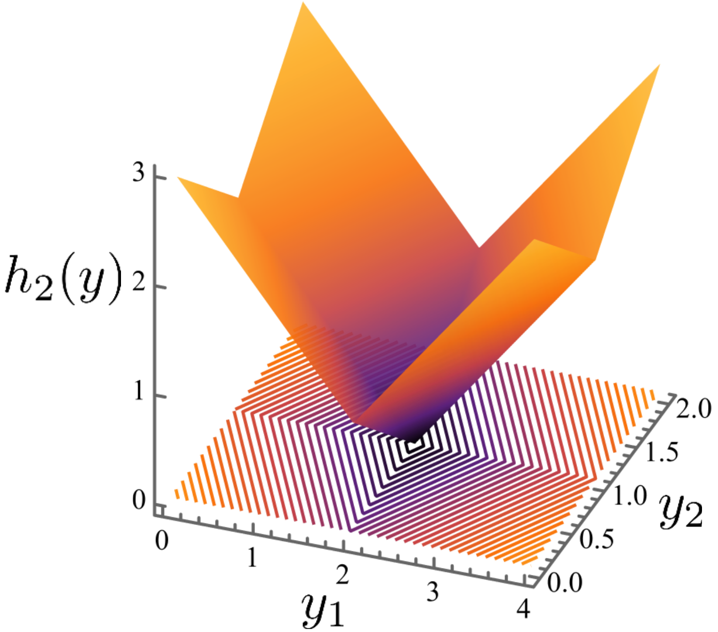


Its global **minimizer** is

$$x^* = (0.5, 1)$$

Define an **invertible map**

$$g(x) := (x_2/x_1, x_2),$$
$$\forall x_1 > 0, x_2 > 0,$$



$$h_2(y) := f_2(g^{-1}(y)) = |y_1 - 2| + |y_2 - 1|, \qquad \forall y_1 > 0, y_2 > 0,$$

Example 3

❑ Linear Quadratic Regulator (LQR)

$$J(k_1, k_2) = \frac{1 - 2k_2 + 3k_2^2 - 2k_2^3 - 2k_1^2 k_2}{k_2^2 - 1}, \qquad \forall k_1 \in \mathbb{R}, k_2 < -1.$$

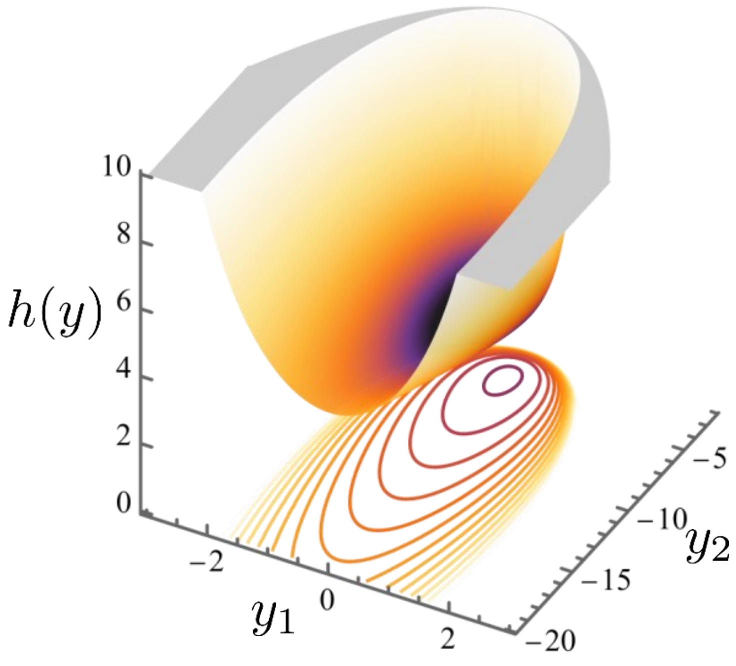
- Not easy to see whether it is convex in the current form
- This cost function comes from an LQR instance

$$A = \begin{bmatrix} -2 & 0 \\ 0 & 1 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; Q = I_2, R = 1$$

- There exists an invertible mapping

$$g(k) := \left(\frac{k_1}{1 - k_2}, \frac{2k_2 - k_1^2 - 2k_2^2}{k_2^2 - 1} \right) \qquad \forall k_1 \in \mathbb{R}, k_2 < -1.$$

- We get a convex function in terms of the new variable y



$$h(y) := J(g^{-1}(y)) = -y_2 - 1 + y^{\mathsf{T}} \begin{bmatrix} 1 & y_1 \\ y_1 & -y_2 - 2 \end{bmatrix}^{-1} y, \qquad \forall \begin{bmatrix} 1 & y_1 \\ y_1 & -y_2 - 2 \end{bmatrix} \succ 0.$$

Direct Convex Reformulation

❑ Direct convex reformulation (the simplest ECL; no lifting)

- Consider a continuous function $J(\mathsf{K}) : \mathcal{D} \to \mathbb{R}$.

  Denote its epigraph as $\operatorname{epi}_{\geq}(f) := \{(\mathsf{K}, \gamma) \in \mathcal{D} \times \mathbb{R} \mid \gamma \geq J(\mathsf{K})\}$.

- Suppose there exists a **smooth and invertible map** Φ between

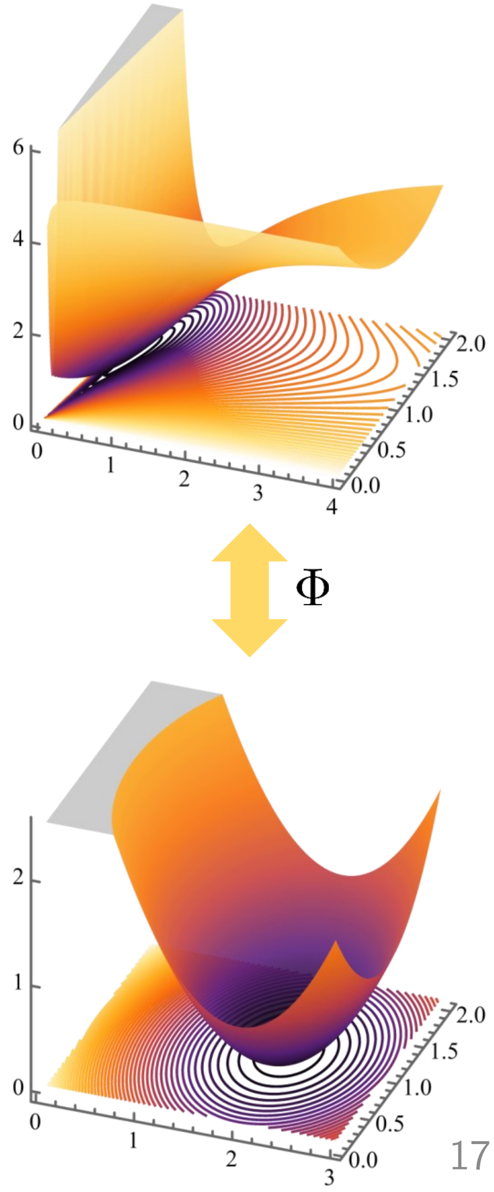
  $\operatorname{epi}_{\geq}(J)$ and a **convex set** $\mathcal{F}_{\mathrm{cvx}}$

- and we further have $(y, \gamma) = \Phi(\mathsf{K}, \gamma), \ \forall (\mathsf{K}, \gamma) \in \operatorname{epi}_{\geq}(J)$

<u>Guarantee 1</u>: Optimization over J(x) is equivalent to a convex problem

$$\inf_{\mathsf{K} \in \mathcal{D}} J(\mathsf{K}) = \inf_{(y,\gamma) \in \mathcal{F}_{\mathrm{cvx}}} \gamma.$$

<u>Guarantee 2</u>: Any stationary point to J(x) is globally optimal; in other words, $0 \in \partial J(\mathsf{K}^*)$ implies globally optimality

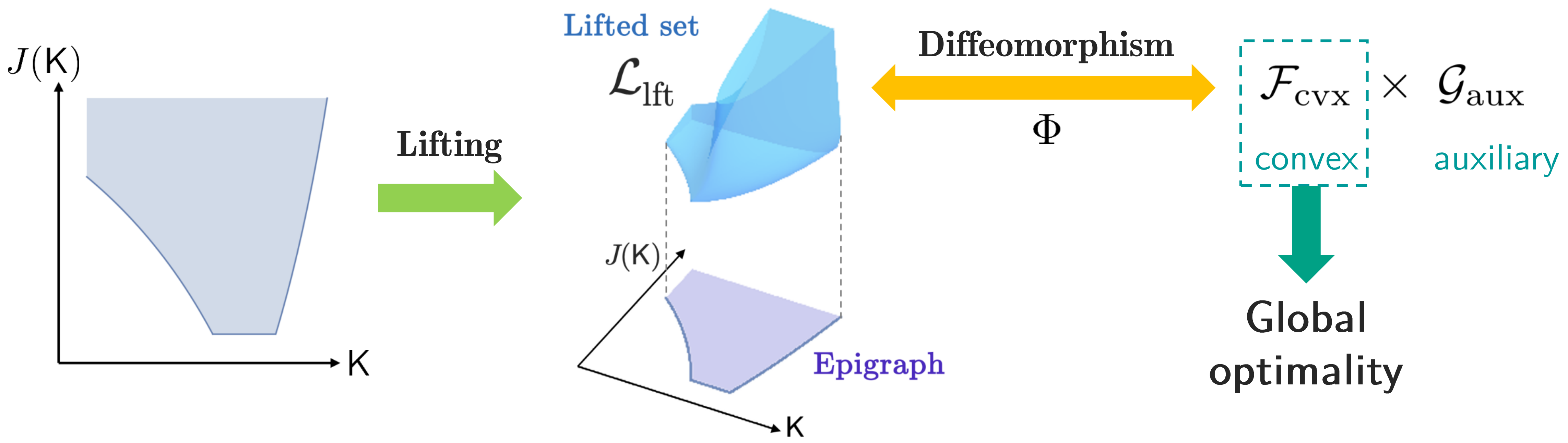


Outline

- ❑ Problem Setup and Motivating Examples
- ❑ Extended Convex Lifting (ECL)
- ❑ ECLs for Optimal and Robust Control
- ❑ Escaping degenerate saddle points

Extended Convex Lifting (ECL)

A schematic illustration of ECL:



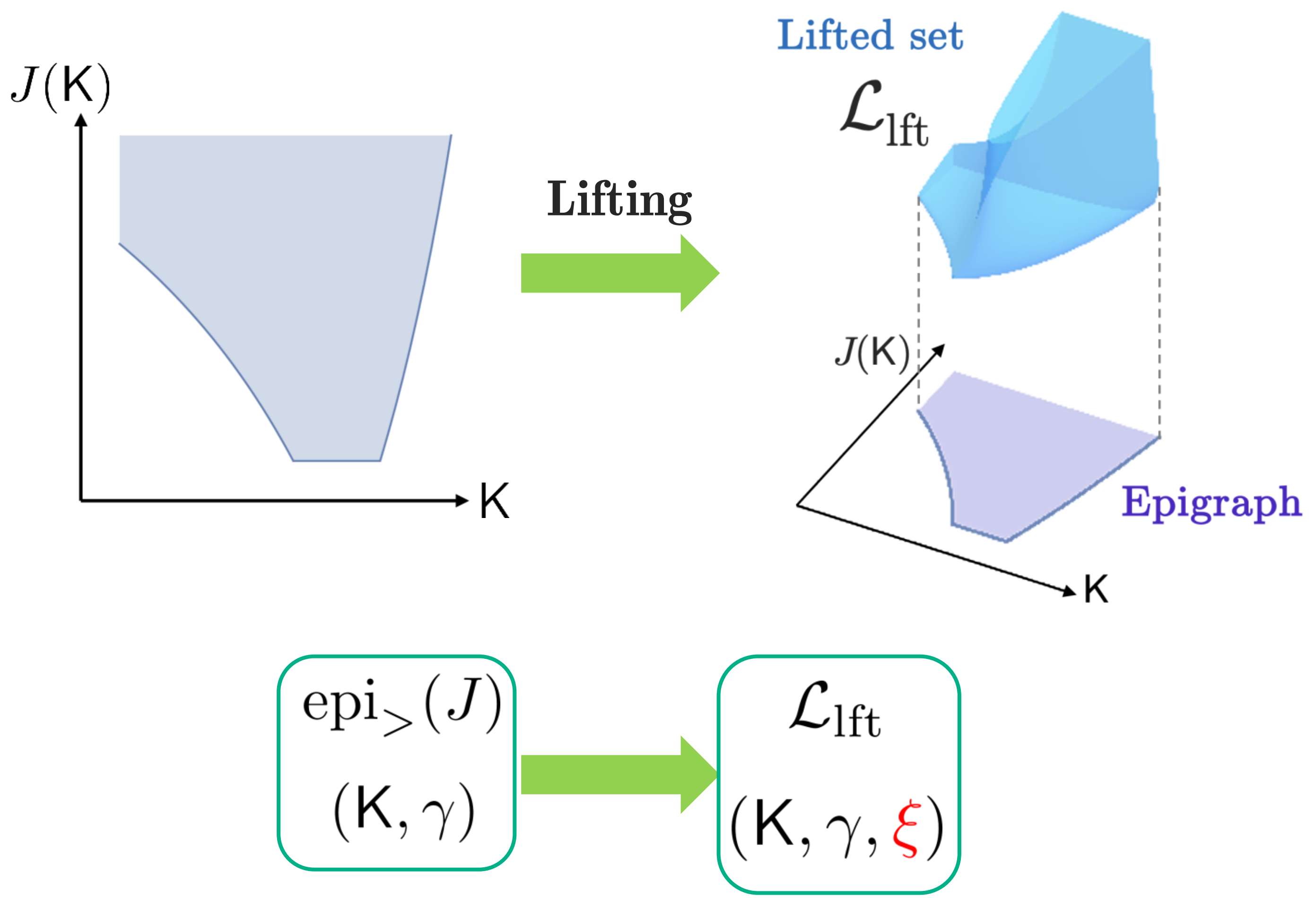
Two key features

Feature 1: a lifting procedure

Feature 2: an auxiliary set

Extended Convex Lifting (ECL)

A schematic illustration of ECL:



Why lifting?

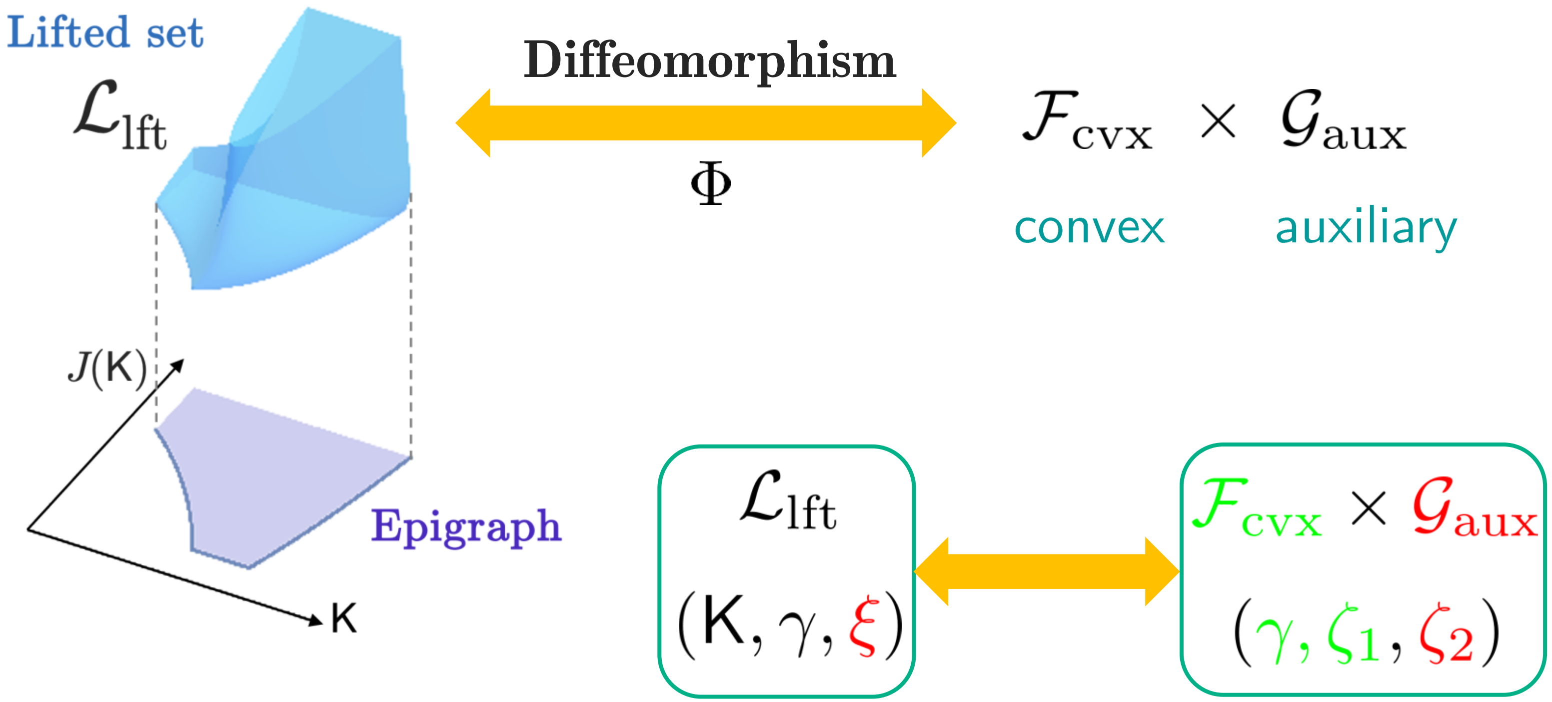
- For many control problems, a direct convexification is not possible
- A lifting procedure corresponding to Lyapunov variables is necessary.

Extended Convex Lifting (ECL)

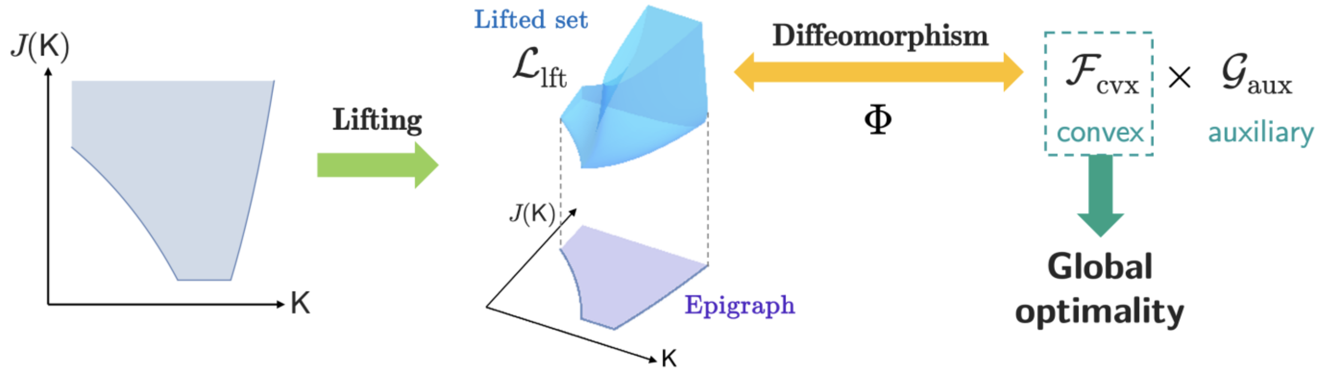
A schematic illustration of ECL:

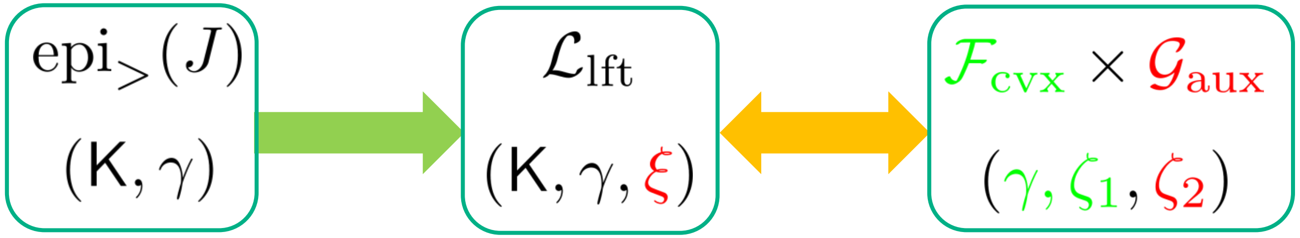
Why auxiliary set?

- Allows us to isolate the **redundancy** or **symmetry** in the original nonconvex domain
- Related to **similarity transformations** of dynamic policies in control
- Needed for **output-feedback control** problems



Formal ECL Definition





- Consider a continuous function $J(\mathsf{K}) : \mathcal{D} \to \mathbb{R}$ where $\mathcal{D} \subseteq \mathbb{R}^d$.
- Denote its strict and non-strict epigraph as

$$\mathrm{epi}_{>}(J) := \{(\mathsf{K}, \gamma) \in \mathcal{D} \times \mathbb{R} \mid \gamma > J(\mathsf{K})\},$$
$$\mathrm{epi}_{\geq}(J) := \{(\mathsf{K}, \gamma) \in \mathcal{D} \times \mathbb{R} \mid \gamma \geq J(\mathsf{K})\}.$$

Extended Convex Lifting (ECL)

We say a tuple $(\mathcal{L}_{\mathrm{lft}}, \mathcal{F}_{\mathrm{cvx}}, \mathcal{G}_{\mathrm{aux}}, \Phi)$ is an ECL of $J(\mathsf{K}) : \mathcal{D} \to \mathbb{R}$ if

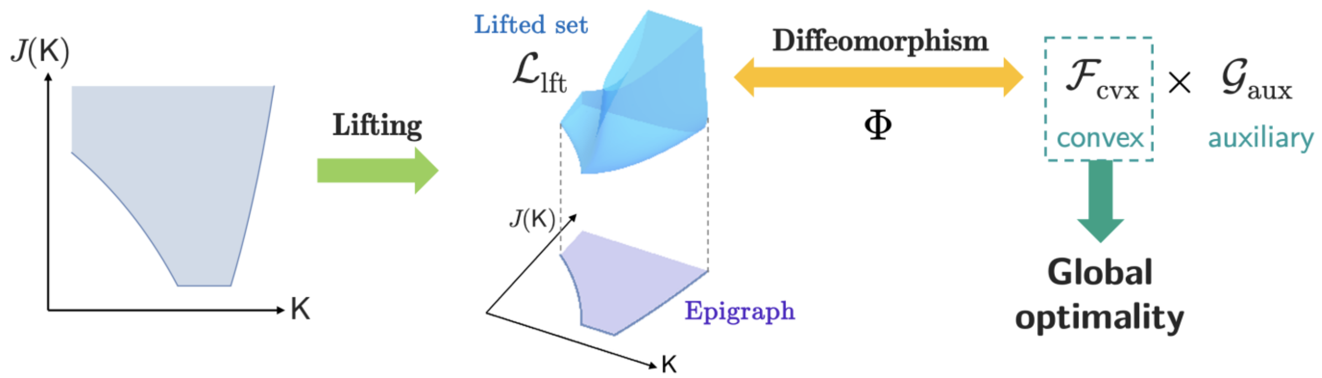
- A lifted set $\mathcal{L}_{\mathrm{lft}}$ satisfying

$$\mathrm{epi}_{>}(J) \subseteq \pi_{\mathsf{K},\gamma}(\mathcal{L}_{\mathrm{lft}}) \subseteq \mathrm{cl}\,\mathrm{epi}_{\geq}(J)$$

- A diffeomorphism $\Phi : \mathcal{L}_{\mathrm{lft}} \to \mathcal{F}_{\mathrm{cvx}} \times \mathcal{G}_{\mathrm{aux}}$ such that

$$\Phi(\mathsf{K}, \gamma, \xi) = (\gamma, \zeta_1, \zeta_2)$$

A special ECL



A more intuitive condition

- A lifted set $\mathcal{L}_{\rm lft}$ satisfying

$$\mathrm{epi}_>(J) \subseteq \pi_{\mathsf{K},\gamma}(\mathcal{L}_{\rm lft}) \subseteq \mathrm{cl}\,\mathrm{epi}_\geq(J)$$

→

- A lifted set $\mathcal{L}_{\rm lft}$ satisfying

$$\pi_{\mathsf{K},\gamma}(\mathcal{L}_{\rm lft}) = \mathrm{epi}_\geq(f).$$

Does this "simpler" lifting condition work?

- Apparently, the condition on the left is less restrictive, and works for more general situations
- The simpler condition on the right is indeed sufficient for state-feedback control problems
- However, it is too restrictive for dynamic output-feedback control problems

Strict vs. Non-strict Epi-graphs

$$\operatorname{epi}_>(J) \subseteq \pi_{\mathsf{K},\gamma}(\mathcal{L}_{\mathrm{lft}}) \subseteq \operatorname{cl}\operatorname{epi}_\geq(J)$$

$$\pi_{\mathsf{K},\gamma}(\mathcal{L}_{\mathrm{lft}}) = \operatorname{epi}_\geq(f).$$

What could the left condition go wrong?

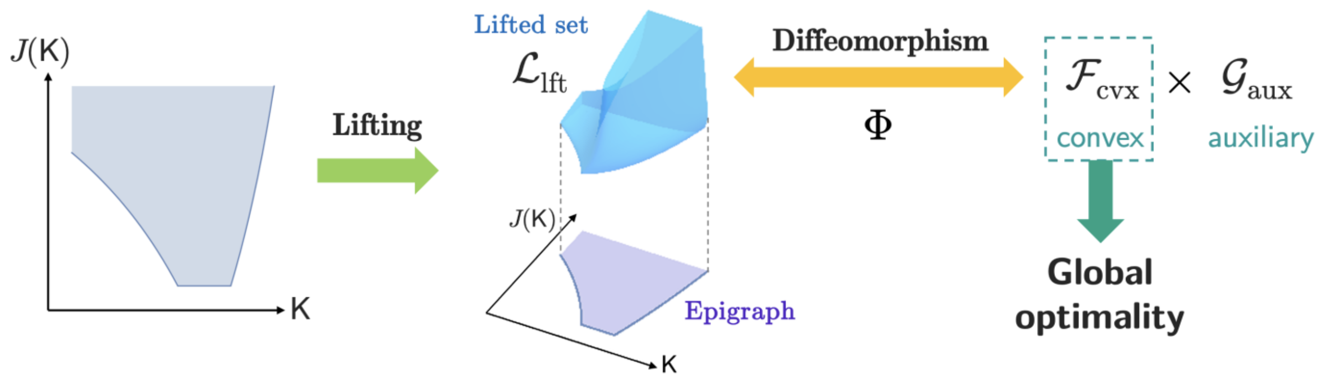
- "Convexifications" of LQG and $\mathcal{H}_\infty$output-feedback control are all based on **strict LMIs**:
  - LQG : $\begin{bmatrix} A^\mathsf{T}P + PA & PB \\ B^\mathsf{T}P & -\gamma I \end{bmatrix} \prec 0, \quad \begin{bmatrix} P & C^\mathsf{T} \\ C & \Gamma \end{bmatrix} \succ 0, \quad \operatorname{trace}(\Gamma) < \gamma$
  - $\mathcal{H}_\infty$: $\begin{bmatrix} A^\mathsf{T}P + PA & PB & C^\mathsf{T} \\ B^\mathsf{T}P & -\gamma I & D^\mathsf{T} \\ C & D & -\gamma I \end{bmatrix} \prec 0$ (bounded real lemma)

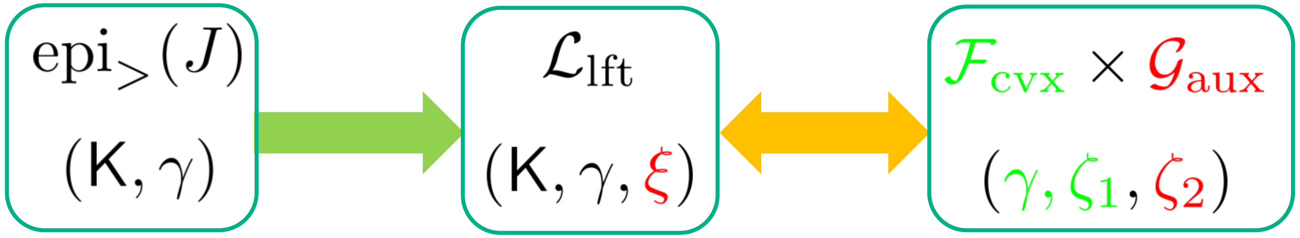
  Used to construct the lifted set $\mathcal{L}_{\mathrm{lft}}$

- **Strict LMIs** only characterize the **strict epigraph** $\operatorname{epi}_>(J) := \{(\mathsf{K}, \gamma) \mid \gamma > J(\mathsf{K})\}$
- They cannot directly characterize the true cost value, i.e., **non-strict epigraphs**

Some classical LMI formulations **are not "equivalent" convex parameterizations** for original control problems, especially in dynamic output feedback cases

Non-degenerate points





Extended Convex Lifting:

- A lifted set $\mathcal{L}_{\text{lft}}$ satisfying

$$\text{epi}_>(J) \subseteq \pi_{\mathsf{K},\gamma}(\mathcal{L}_{\text{lft}}) \subseteq \text{cl}\,\text{epi}_\geq(J)$$

- A diffeomorphism $\Phi : \mathcal{L}_{\text{lft}} \to \mathcal{F}_{\text{cvx}} \times \mathcal{G}_{\text{aux}}$ such that

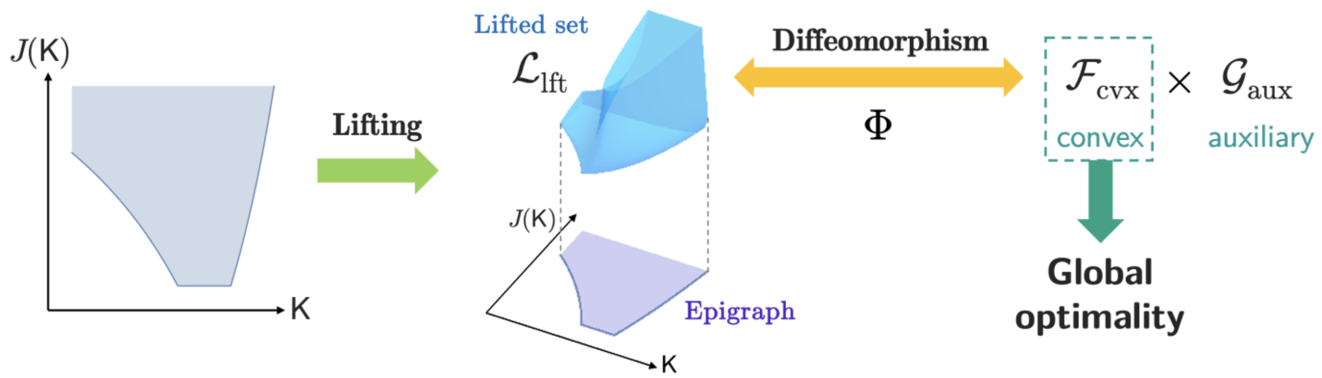
$$\Phi(\mathsf{K}, \gamma, \xi) = (\gamma, \zeta_1, \zeta_2)$$

- By construction, some points in $\text{epi}_\geq(J)$ may **not be covered** in the lifted set

→ Those points will be called **degenerate** → **bad behavior (e.g., saddles) may exist**

Definition. K is called **non-degenerate** if $(\mathsf{K}, J(\mathsf{K})) \in \pi_{\mathsf{K},\gamma}(\mathcal{L}_{\text{lft}})$ → **well-behaved**

ECL Guarantees



Given an Extended Convex Lifting

$$(\mathcal{L}_{\mathrm{lft}}, \mathcal{F}_{\mathrm{cvx}}, \mathcal{G}_{\mathrm{aux}}, \Phi)$$

Guarantee 1: Convex Reformulation

Optimization $J(\mathsf{K})$ is equivalent to a convex problem

$$\inf_{\mathsf{K}\in\mathcal{D}} J(\mathsf{K}) = \inf_{(y,\gamma)\in\mathcal{F}_{\mathrm{cvx}}} \gamma.$$

Guarantee 2: Global Optimality

All non-degenerate Clarke stationary points are globally optimal

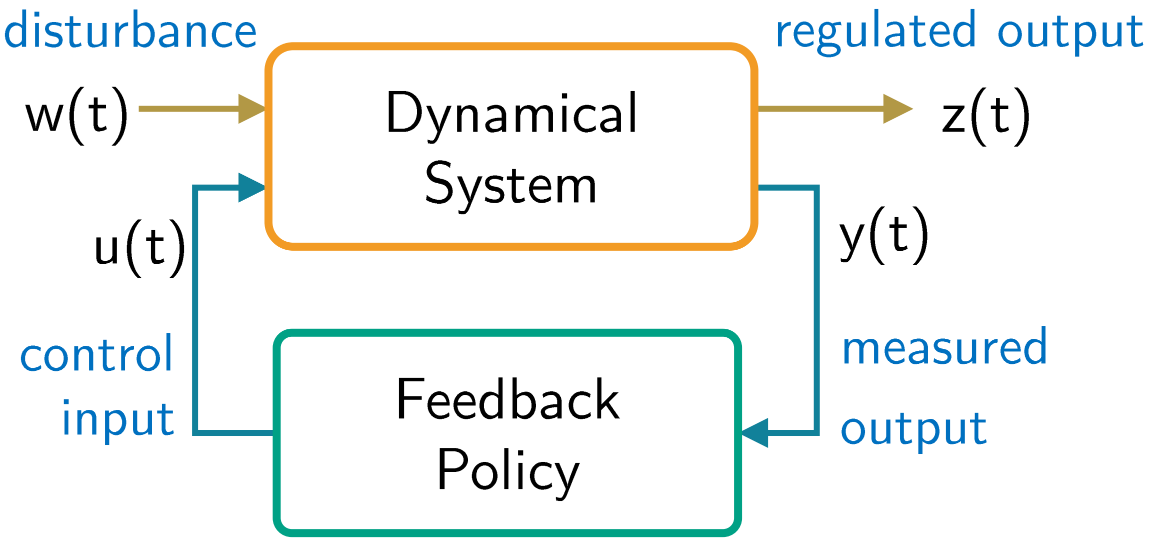
- **Clarke stationary points**: Generalization of stationary points to nonsmooth functions, based on the notion of Clarke subdifferential

Outline

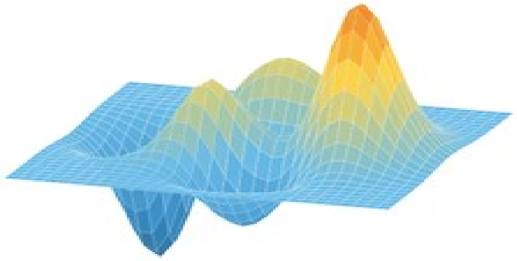
- Problem Setup and Motivating Examples
- Extended Convex Lifting (ECL)
- ECLs for Optimal and Robust Control
- Escaping degenerate saddle points

Global Optimality in Control

❑ Optimal and Robust Control



Policy parametrization



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

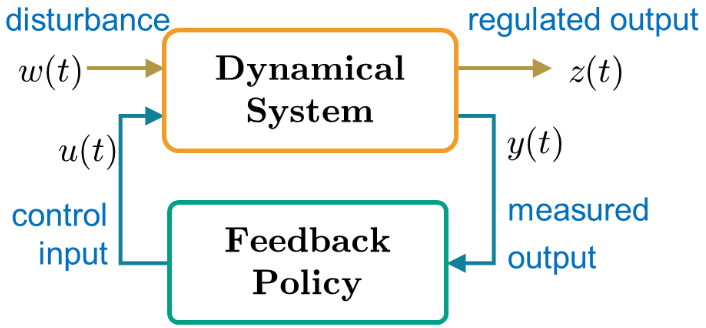
Non-convex Optimization problem

Main Results (informal):

1. **Static state feedback**: Any (Clarke) stationary points in LQR or Hinf control are globally optimal ([Fazel et al., 2018]; [Guo & Hu, 2022]);
2. **Dynamic output feedback**: Any **non-degenerate** (Clarke) stationary points in **LQG** or **Hinf dynamic output control** are globally optimal.

Linear Quadratic Regulator (LQR)

- Problem setup



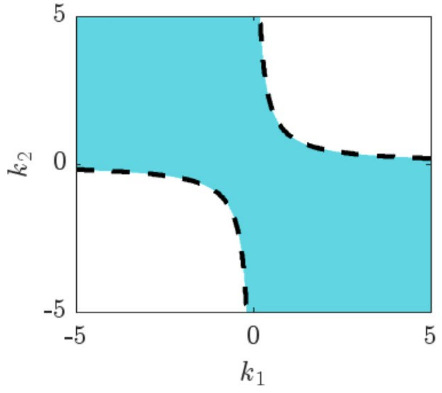
Dynamics: $\dot{x}(t) = Ax(t) + Bu(t) + B_w w(t),$

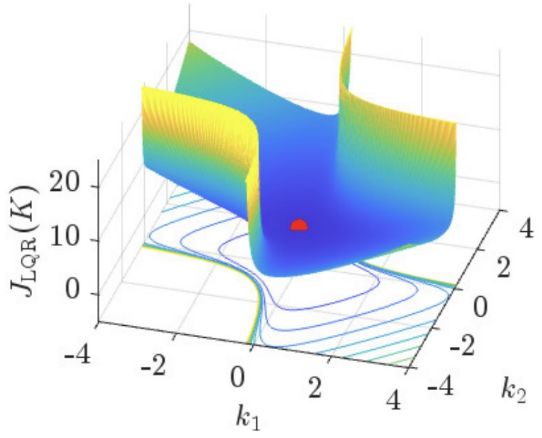
Static policies: $u(t) = Kx(t)$

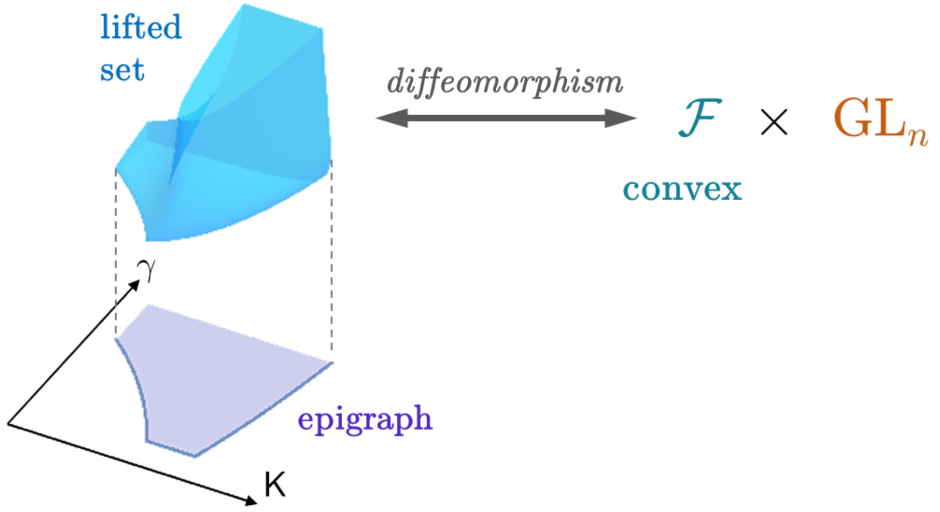
Stability: $\mathcal{C} = \{K \in \mathbb{R}^{m\times n} \mid A + BK \text{ is stable}\}$

Performance: $J_{\mathrm{LQR}}(K) := \lim_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\int_0^T x^{\mathsf{T}}(t)Qx(t) + u^{\mathsf{T}}(t)Ru(t)\,dt\right]$

- Nonconvex and smooth landscape







Linear Quadratic Regulator (LQR)

❑ Construction of ECL

Step 1: Lifting

$$\mathcal{L}_{\mathrm{LQR}} := \left\{(K,\gamma,X) : X \succ 0, (A+BK)X + X(A+BK)^{\mathsf{T}} + W = 0, \gamma \geq \mathrm{Tr}\left[(Q+K^{\mathsf{T}}RK)X\right]\right\}.$$

Step 2: Convex set

$$\mathcal{F}_{\mathtt{LQR}} = \left\{(\gamma,Y,X) : X \succ 0, AX + BY + XA^{\mathsf{T}} + Y^{\mathsf{T}}B^{\mathsf{T}} + W = 0, \gamma \geq \mathrm{tr}\big(QX + X^{-1}Y^{\mathsf{T}}RY\big)\right\}$$

Step 3: Diffeomorphism $\quad \Phi(K,\gamma,X) = (\gamma, KX, X), \quad \forall (K,\gamma,X) \in \mathcal{L}_{\mathtt{LQR}}$

- No auxiliary set
- Lifted set satisfies $\mathrm{epi}_{\geq}(J) = \pi_{K,\gamma}(\mathcal{L}_{\mathtt{LQR}})$

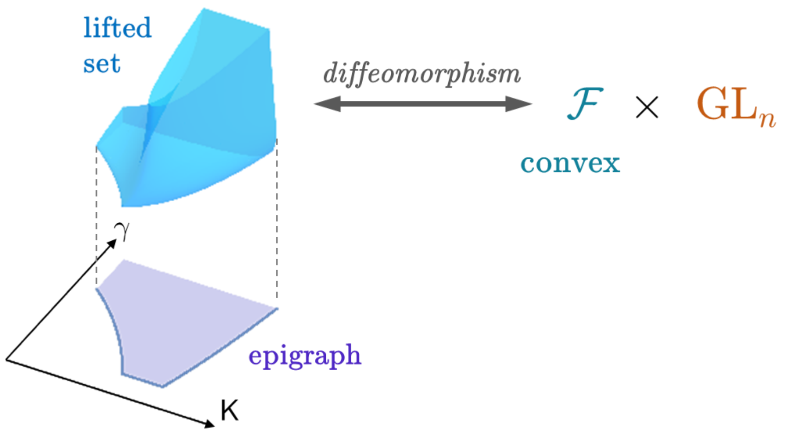
➡ All policies are non-degenerate

> **Theorem.** Any stationary point of the LQR cost function is globally optimal.

Under mild assumptions, LQR behaves like a strongly convex problem,

→ satisfying Gradient Dominance property

State-feedback Robust Control



Problem setup

Dynamics: $$\dot{x}(t) = Ax(t) + Bu(t) + B_w w(t),$$

Static policies: $$u(t) = Kx(t)$$

Stability: $$\mathcal{C} = \{K \in \mathbb{R}^{m \times n} \mid A + BK \text{ is stable}\}$$

Performance: $$J_\infty(K) := \sup_{\|w(t)\|_2 \le 1} \int_0^\infty x^\mathsf{T}(t)Qx(t) + u^\mathsf{T}(t)Ru(t)\, dt$$

Building an ECL

Step 1: Lifting

$$\mathcal{L}_\infty := \left\{ (K, \gamma, P) : P \succ 0, \begin{bmatrix} (A+BK)^\mathsf{T} P + P(A+BK) & PB_w & C^\mathsf{T} \\ B_w^\mathsf{T} P & -\gamma I & 0 \\ C & 0 & -\gamma I \end{bmatrix} \preceq 0 \right\},$$

State-feedback Robust Control

- Building an ECL

Step 1: Lifting

$$\mathcal{L}_\infty := \left\{ (K,\gamma,P) : P \succ 0, \begin{bmatrix} (A+BK)^\top P + P(A+BK) & PB_w & C^\top \\ B_w^\top P & -\gamma I & 0 \\ C & 0 & -\gamma I \end{bmatrix} \preceq 0 \right\},$$

Step 2: Convex set

$$\mathcal{F}_\infty = \left\{ (\gamma, Y, X) \middle| \begin{array}{r} X \succ 0, \\ Y \in \mathbb{R}^{m\times n}, \end{array} \begin{bmatrix} AX + XA^\top + BY + Y^\top B^\top & B_w & XQ^{1/2} & Y^\top R^{1/2} \\ B_w^\top & -\gamma I & 0 & 0 \\ Q^{1/2}X & 0 & -\gamma I & 0 \\ R^{1/2}Y & 0 & 0 & -\gamma I \end{bmatrix} \preceq 0 \right\},$$

Step 3: Diffeomorphism $\Phi(K,\gamma,P) = (\gamma, KP^{-1}, P^{-1}), \qquad \forall (K,\gamma,P) \in \mathcal{L}_\infty.$

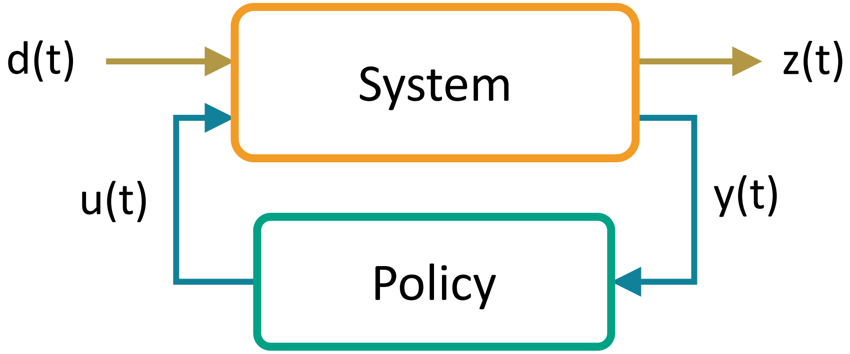
- No auxiliary set
- Lifted set satisfies $\pi_{K,\gamma}(\mathcal{L}_\infty) = \mathrm{epi}_{\geq}(J_\infty)$

➡ **All policies are non-degenerate**

Theorem: Any Clarke stationary points are globally optimal!

Linear Quadratic Gaussian (LQG)

- Problem setup



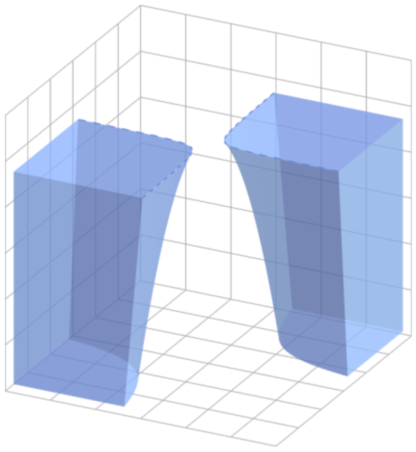
Dynamics: $$\dot{x}(t) = Ax(t) + Bu(t) + B_w w(t)$$
$$y(t) = Cx(t) + D_v v(t)$$

Performance: $$J = \|\mathbf{T}_{zd}\|_{\mathcal{H}_2}$$

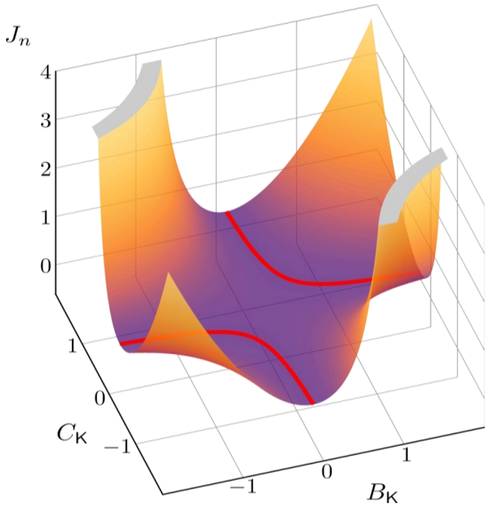
$$z(t) = \begin{bmatrix} Q^{1/2}x(t) \\ R^{1/2}u(t) \end{bmatrix} \quad d(t) = \begin{bmatrix} w(t) \\ v(t) \end{bmatrix}$$

Policy: $$\dot{\xi}(t) = A_{\mathsf{K}}\xi(t) + B_{\mathsf{K}}y(t)$$
$$u(t) = C_{\mathsf{K}}\xi(t)$$

$$\mathsf{K} = (A_{\mathsf{K}}, B_{\mathsf{K}}, C_{\mathsf{K}})$$



disconnected domain



multiple globally optimal points

Linear Quadratic Gaussian (LQG)

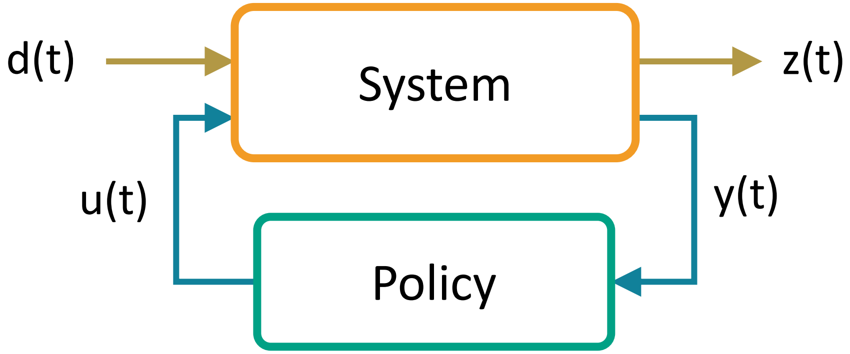
- **Construction of the ECL**: Based on the convexification proposed in [Scherer et al., 1997]

Theorem.
1. An ECL for LQG exists, of which $\mathcal{G}_{\text{aux}}$ is the set of invertible matrices.
2. A policy K is non-degenerate if and only if it is **informative** in the sense that
$$\lim_{t\to\infty} \mathbb{E}\left[x(t)\xi(t)^{\mathsf{T}}\right]$$
has full rank. So **any informative stationary point is globally optimal**.
3. Non-degenerate policies are **generic** in the sense that degenerate policies form a **set of measure zero**.

- Part 2 extends [Umenberger et al., 2022, Theorem 1(ii)] from Kalman filtering to LQG.
- We also show that **minimal stationary policies are non-degenerate**, generalizing our exisiting results in [Tang, Zheng, Li, 2023].

$\mathcal{H}_\infty$ Output-Feedback Control

❑ Problem setup



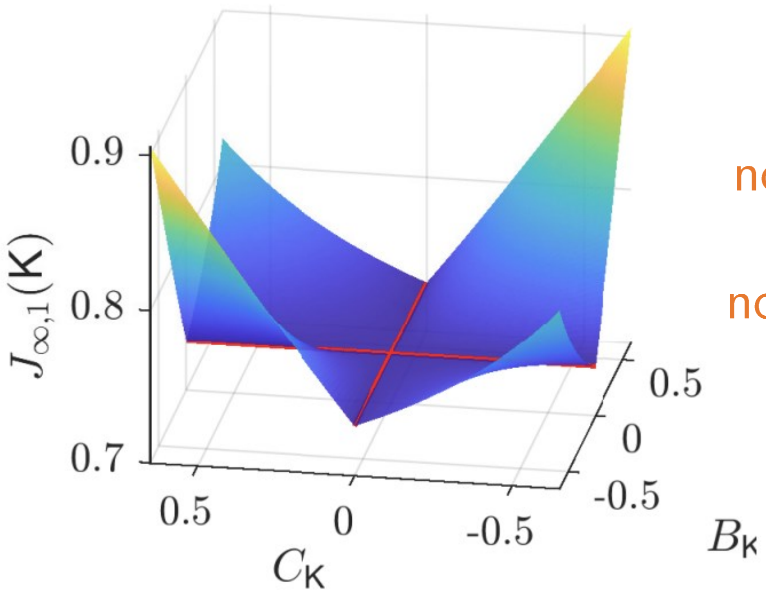
Policy: $$\dot{\xi}(t) = A_{\mathsf{K}}\xi(t) + B_{\mathsf{K}}y(t)$$
$$u(t) = C_{\mathsf{K}}\xi(t) + D_{\mathsf{K}}y(t)$$

$$\mathsf{K} = (A_{\mathsf{K}}, B_{\mathsf{K}}, C_{\mathsf{K}}, D_{\mathsf{K}})$$

Dynamics: $$\dot{x}(t) = Ax(t) + Bu(t) + B_w w(t)$$
$$y(t) = Cx(t) + D_v v(t)$$

Performance: $$J = \|\mathbf{T}_{zd}\|_{\mathcal{H}_\infty}$$

$$z(t) = \begin{bmatrix} Q^{1/2}x(t) \\ R^{1/2}u(t) \end{bmatrix} \quad d(t) = \begin{bmatrix} w(t) \\ v(t) \end{bmatrix}$$



nonconvex
+
nonsmooth

$\mathcal{H}_\infty$ Output-Feedback Control

- Construction of the ECL: Based on the convexification proposed in [Scherer et al., 1997]

Theorem.
1. An ECL for $\mathcal{H}_\infty$ output-feedback control exists.
2. A policy K is non-degenerate if and only if

   a) There exists a non-strict certificate $P \succeq 0$ of the $\mathcal{H}_\infty$ cost.

$$\begin{bmatrix} A_{\mathrm{cl}}^{\mathsf{T}}(\mathsf{K})P + PA_{\mathrm{cl}}(\mathsf{K}) & PB_{\mathrm{cl}}(\mathsf{K}) & C_{\mathrm{cl}}^{\mathsf{T}}(\mathsf{K}) \\ B_{\mathrm{cl}}^{\mathsf{T}}(\mathsf{K})P & -J(\mathsf{K})I & D_{\mathrm{cl}}^{\mathsf{T}}(\mathsf{K}) \\ C_{\mathrm{cl}}(\mathsf{K}) & D_{\mathrm{cl}}(\mathsf{K}) & -J(\mathsf{K})I \end{bmatrix} \preceq 0$$

   b) The block P_{12} is invertible.

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^{\mathsf{T}} & P_{22} \end{bmatrix}$$

So **a Clarke stationary point is globally optimal if these conditions hold**.

- Physical interpretation of non-degeneracy is not as clear as LQG.
- We conjecture that non-degenerate policies for $\mathcal{H}_\infty$ output-feedback control are also **generic**, with some numerical evidence, but a proof is not known yet.

Outline

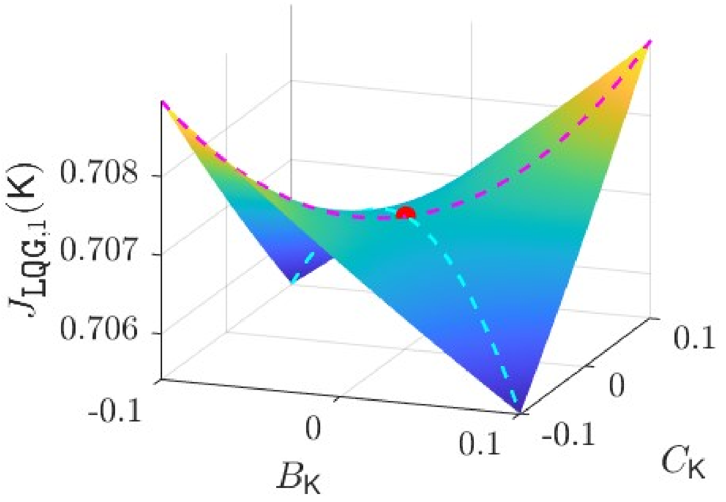
- ❑ Problem Setup and Motivating Examples
- ❑ Extended Convex Lifting (ECL)
- ❑ ECLs for Optimal and Robust Control
- ❑ Escaping degenerate saddle points

Degenerate Saddle points

Policy Optimization for LQG

$$\min_{\mathsf{K}} \quad J(\mathsf{K})$$
$$\text{s.t.} \quad \mathsf{K} = (A_{\mathsf{K}}, B_{\mathsf{K}}, C_{\mathsf{K}}) \in \mathcal{C}_{\text{full}}$$

- **Minimal** (aka controllable and observable) stationary policies are **non-degenerate**;
- They are globally optimal;
- There are also other degenerate points.



❑ Local geometry

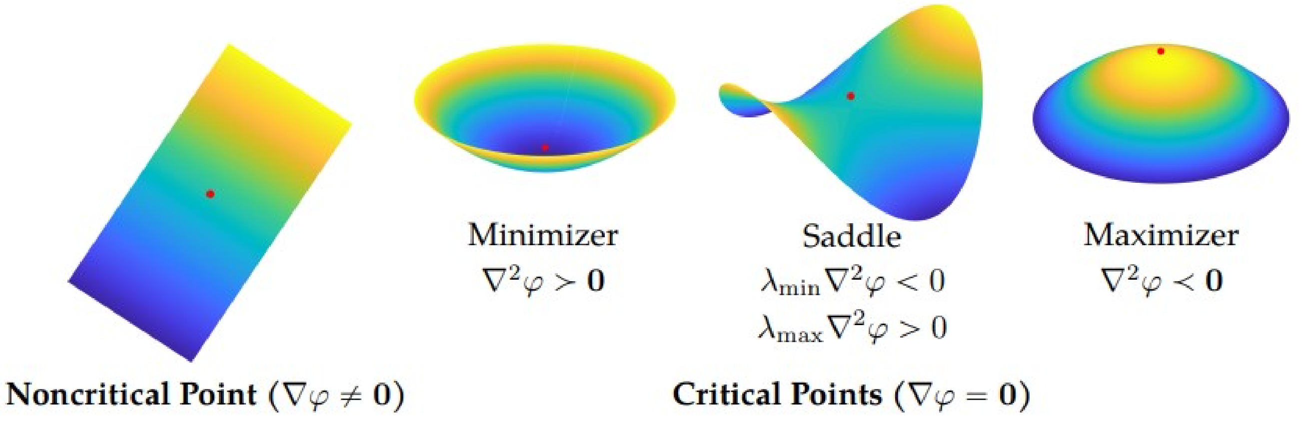


Figure taken from Zhang et al., 2020

- ❑ **Strict saddle points**: the hessian has a strict negative eigenvalue (i.e., escaping direction)
- ❑ **Non-strict (high-order) saddle points**: no such escaping direction, i.e., minimum eigenvalue is zero.
- ❑ **Simple perturbed gradient descent (PGD)** methods can escape strict saddle points efficiently (e.g., Jin et al., 2017)

✓ Zhang, Yuqian, Qing Qu, and John Wright. "From symmetry to geometry: Tractable nonconvex problems." *arXiv preprint arXiv:2007.06753* (2020).

Structure of stationary points

- Theorem (informal): all bad stationary points are in the same form

$$\left\{ \mathsf{K} \in \mathcal{C}_n \;\middle|\; \frac{\partial J(\mathsf{K})}{\partial A_{\mathsf{K}}} = 0, \; \frac{\partial J(\mathsf{K})}{\partial B_{\mathsf{K}}} = 0, \; \frac{\partial J(\mathsf{K})}{\partial C_{\mathsf{K}}} = 0, \right\}$$

a stationary point

$$\mathsf{K} = \begin{bmatrix} 0 & C_{\mathsf{K}} \\ B_{\mathsf{K}} & A_{\mathsf{K}} \end{bmatrix} \in \mathcal{C}_n$$

If it is **minimal**, then it is globally optimal

If it is not minimal, find a minimal realization

$$\hat{\mathsf{K}} = \begin{bmatrix} 0 & \hat{C}_{\mathsf{K}} \\ \hat{B}_{\mathsf{K}} & \hat{A}_{\mathsf{K}} \end{bmatrix} \in \mathcal{C}_q$$

$$\dot{\xi}(t) = A_{\mathsf{K}}\xi(t) + B_{\mathsf{K}}y(t),$$
$$u(t) = C_{\mathsf{K}}\xi(t),$$

$$\dot{\hat{\xi}}(t) = \begin{bmatrix} \hat{A}_{\mathsf{K}} & 0 \\ 0 & \Lambda \end{bmatrix} \hat{\xi}(t) + \begin{bmatrix} \hat{B}_{\mathsf{K}} \\ 0 \end{bmatrix} y(t),$$
$$u(t) = \begin{bmatrix} \hat{C}_{\mathsf{K}} & 0 \end{bmatrix} \hat{\xi}(t),$$

The following full-order controller with any stable Λ is also a stationary point with the same LQG cost

$$\tilde{\mathsf{K}} = \left[\begin{array}{c|cc} 0 & \hat{C}_{\mathsf{K}} & 0 \\ \hline \hat{B}_{\mathsf{K}} & \hat{A}_{\mathsf{K}} & 0 \\ 0 & 0 & \Lambda \end{array}\right] \in \mathcal{C}_n$$

where we isolate the uncontrollable and unobservable part

Strict saddle points

$$\left\{ \mathsf{K} \in \mathcal{C}_n \;\middle|\; \frac{\partial J(\mathsf{K})}{\partial A_\mathsf{K}} = 0, \; \frac{\partial J(\mathsf{K})}{\partial B_\mathsf{K}} = 0, \; \frac{\partial J(\mathsf{K})}{\partial C_\mathsf{K}} = 0, \right\}$$

a stationary point

$$\mathsf{K} = \begin{bmatrix} 0 & C_\mathsf{K} \\ B_\mathsf{K} & A_\mathsf{K} \end{bmatrix} \in \mathcal{C}_n$$

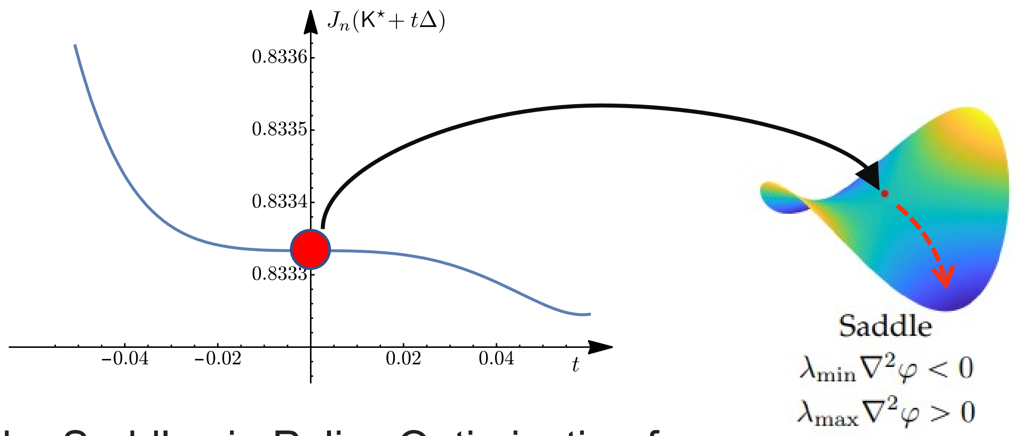
The same form

$$\tilde{\mathsf{K}} = \left[\begin{array}{c|cc} 0 & \hat{C}_\mathsf{K} & 0 \\ \hline \hat{B}_\mathsf{K} & \hat{A}_\mathsf{K} & 0 \\ 0 & 0 & \Lambda \end{array}\right] \in \mathcal{C}_n$$

- **Theorem (informal): Under a mild condition, choosing the diagonal stable block Λ randomly leads to a strict saddle point with probability 1**

Our idea: a structural perturbation

A high-order saddle → A strict saddle point with the same LQG cost



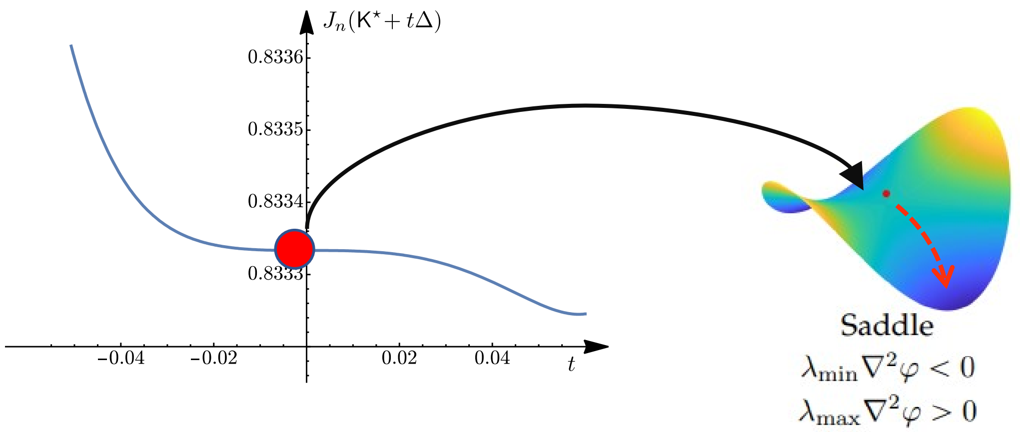
✓ Yang Zheng*, Yue Sun*, Maryam Fazel, and Na Li. "Escaping High-order Saddles in Policy Optimization for Linear Quadratic Gaussian (LQG) Control." arXiv preprint arXiv:2204.00912 (2022). *Equal contribution

Perturbed Gradient Descent

- **Theorem (informal)**: all bad stationary points are in the same form

$$\tilde{\mathsf{K}} = \left[\begin{array}{c|cc} 0 & \hat{C}_{\mathsf{K}} & 0 \\ \hline \hat{B}_{\mathsf{K}} & \hat{A}_{\mathsf{K}} & 0 \\ 0 & 0 & \Lambda \end{array}\right] \in \mathcal{C}_n$$

- **Theorem (informal)**: Choosing the diagonal stable block Λ randomly leads to a strict saddle point with probability almost 1



Our idea: a structural perturbation + standard PGD

A non-optimal stationary point → A strict saddle point with the same LQG cost → Standard PGD algorithm (Jin et al., 2017)

Perturbation on Λ

Perturbation on gradients

Jin, C., Ge, R., Netrapalli, P., Kakade, S. M., & Jordan, M. I. (2017, July). How to escape saddle points efficiently. In *International Conference on Machine Learning* (pp. 1724-1732). PMLR.

Numerical simulations

Three policy gradient algorithms

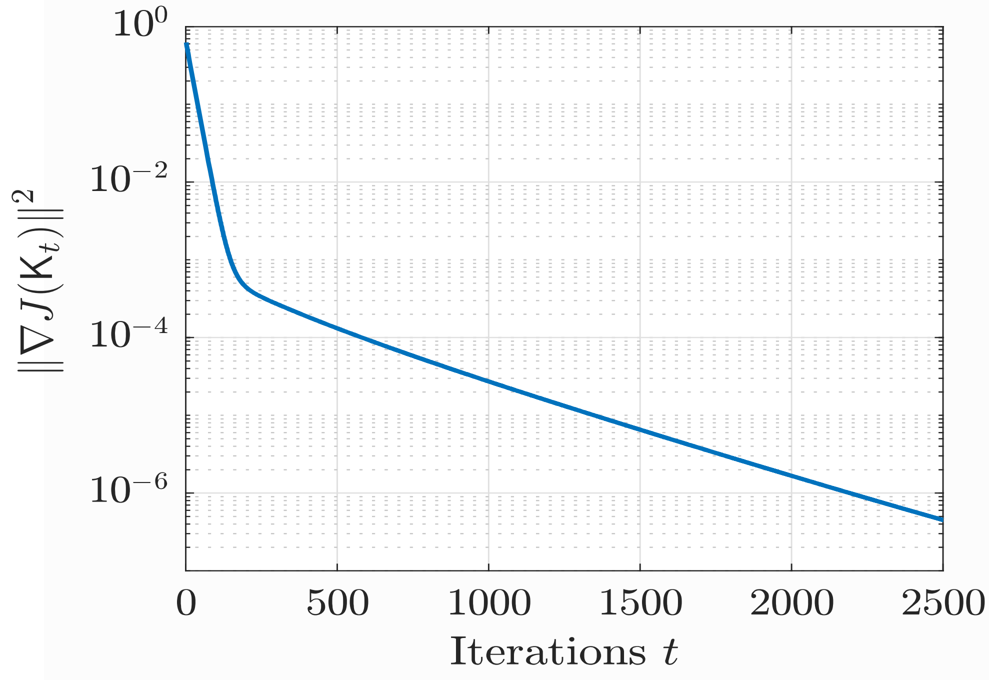
1. Vanilla gradient descent $\mathsf{K}_{i+1} = \mathsf{K}_i - \alpha_i \nabla J(\mathsf{K}_i)$
2. Standard PGD algorithm (adding a small random perturbation on iterates; Jin et al., 2017;)
3. Structural perturbation + standard PGD

Example: System dynamics

$$A = \begin{bmatrix} -0.5 & 0 \\ 0.5 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad C = \begin{bmatrix} -\frac{1}{6} & \frac{11}{12} \end{bmatrix},$$

Performance weights

$$W = Q = I_2, \; V = R = 1$$



Numerical simulations

Three policy gradient algorithms

1. Vanilla gradient descent $\mathsf{K}_{i+1} = \mathsf{K}_i - \alpha_i \nabla J(\mathsf{K}_i)$
2. Standard PGD algorithm (adding a small random perturbation on iterates; Jin et al., 2017;)
3. Structural perturbation + standard PGD

Example: System dynamics

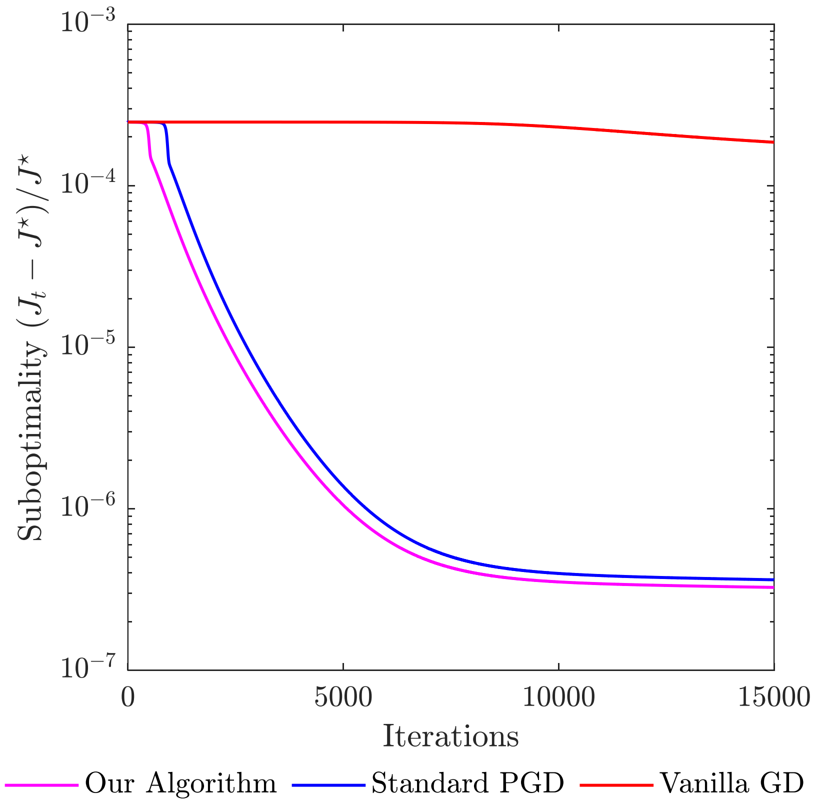
$$A = \begin{bmatrix} -0.5 & 0 \\ 0.5 & -1 \end{bmatrix}, \; B = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \; C = \begin{bmatrix} -\frac{1}{6} & \frac{11}{12} \end{bmatrix},$$

Performance weights

$$W = Q = I_2, \; V = R = 1$$

A point that is close to a high-order saddle with zero hessian

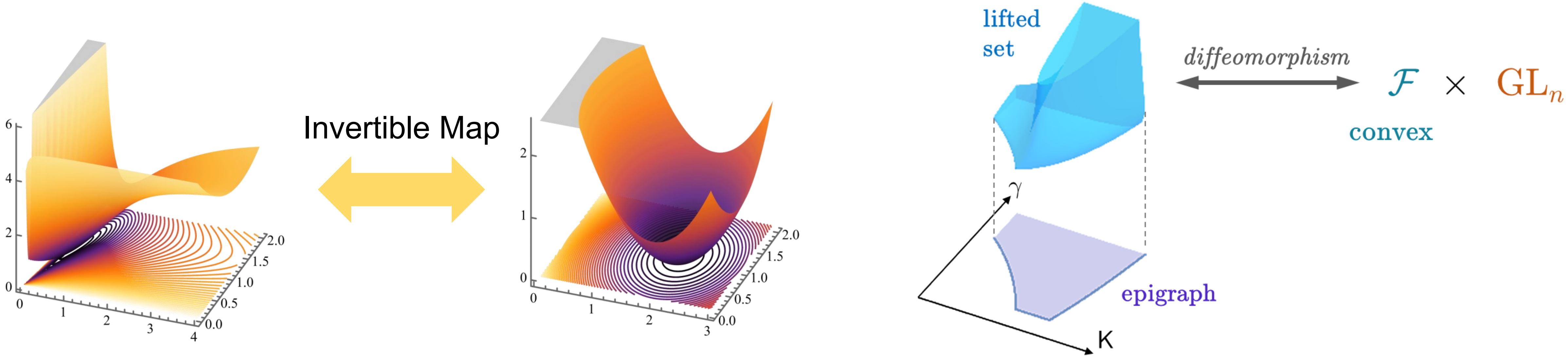
$$A_{\mathsf{K},0} = -0.5 I_2, \; B_{\mathsf{K},0} = \begin{bmatrix} 0 \\ 0.01 \end{bmatrix}, \; C_{\mathsf{K},0} = [0, -0.01]$$



Conclusion

Nonconvex Policy Optimization for control

- ❑ Policy optimization in control can be **nonconvex and non-smooth.**
- ❑ **Extended Convex Lifting (ECL)** reveals benign nonconvexity.



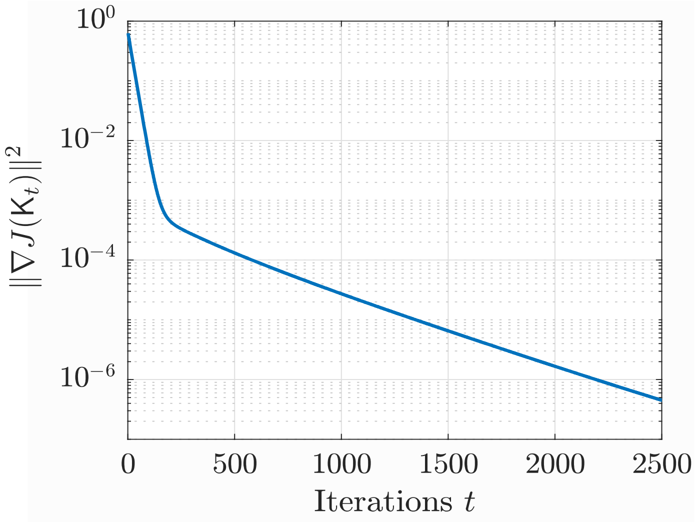
- ❑ **Global Optimality**

Local Stationarity + **Structural Information**

Global Optimality Certificate

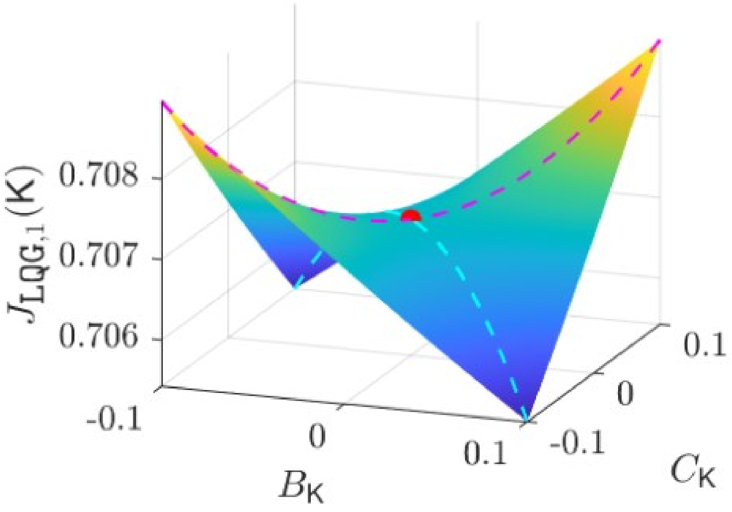
Ongoing and Future work

- How to design principled local search algorithms for **nonconvex** **and** **non-smooth** policy optimization?
- How to establish **convergence conditions and speeds**?
- How to deal with **degenerate points** in local policy search? Avoiding saddle points with guarantees?



Control Theory ↔ **(Non)convex Optimization**

↕ **Reinforcement Learning** ↕



Thank you for your attention!

Q & A

- Zheng, Yang, Chih-Fan Pai, and Yujie Tang. "**Benign Nonconvex Landscapes in Optimal and Robust Control, Part I: Global Optimality.**" arXiv preprint arXiv:2312.15332 (2023): https://arxiv.org/abs/2312.15332.
- Zheng, Yang, Chih-Fan Pai, and Yujie Tang. "**Benign Nonconvex Landscapes in Optimal and Robust Control, Part II: Extended Convex Lifting.**" arXiv preprint arXiv:2406.04001 (2024): https://arxiv.org/abs/2406.04001
- Watanabe, Yuto, and Yang Zheng. "**Revisiting Strong Duality, Hidden Convexity, and Gradient Dominance in the Linear Quadratic Regulator.**" arXiv preprint arXiv:2503.10964 (2025): https://arxiv.org/abs/2503.10964