Benign Nonconvex Landscapes in Optimal and Robust Control

Yang Zheng

Assistant Professor,
ECE Department, UC San Diego

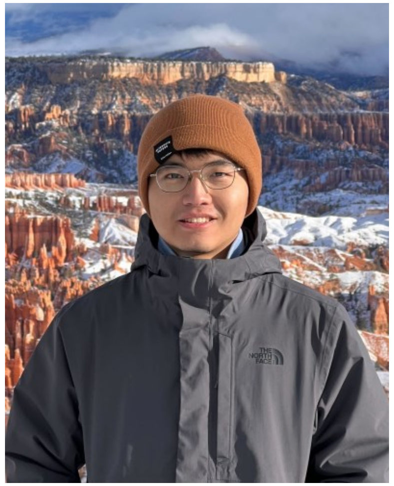
2024 Information Theory and Applications Workshop

Feb 23, 2024

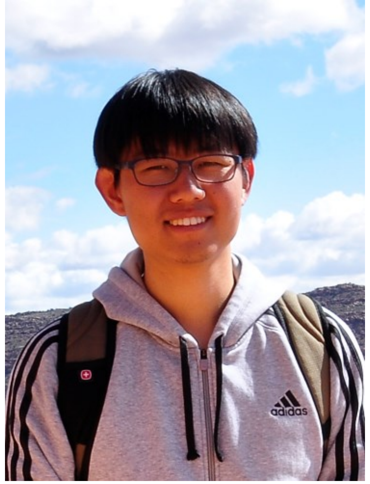
UC San Diego
JACOBS SCHOOL OF ENGINEERING
Electrical and Computer Engineering

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

Acknowledgements



Chih-Fan (Rich) Pai
University of California San Diego



Yujie Tang
Peking University

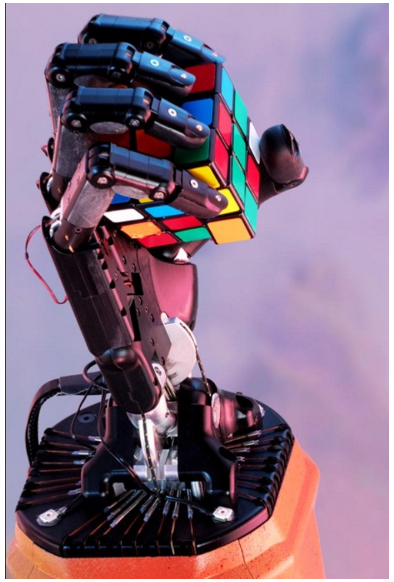
- Zheng, Yang, Chih-Fan Pai, and Yujie Tang. "**Benign Nonconvex Landscapes in Optimal and Robust Control, Part I: Global Optimality**." preprint arXiv:2312.15332 (2023): https://arxiv.org/abs/2312.15332.

- **Part II: Extended Convex Lifting** will be out soon

Supported by
NSF ECCS-2154650
NSF CMMI-2320697

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.

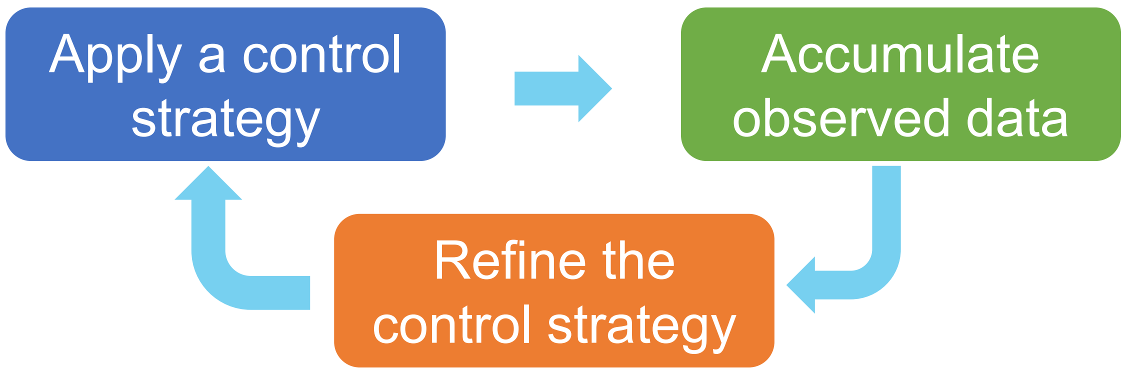


ChatGPT

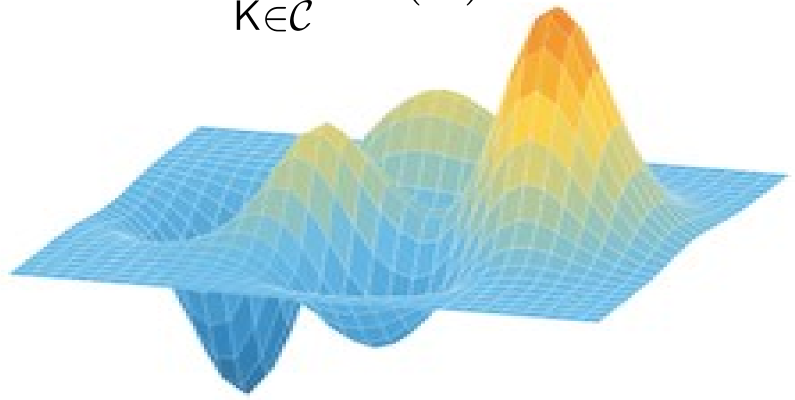
Duan et al. 2016; Silver et al., 2017; Dean et al., 2019; Tu and Recht, 2019; Mania et al., 2019; Fazel et al., 2018; Recht, 2019; https://chat.openai.com/

Policy Optimization for Control

❑ Why policy optimization is so popular



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



Opportunities

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

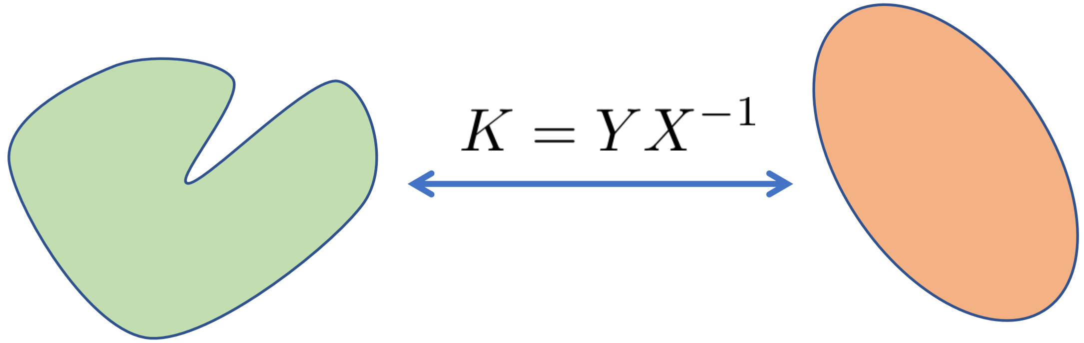
Challenges

- **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)

Convex LMIs vs Policy optimization

❑ Historical background

- Since 1980s, **convex LMIs** become dominant due to **global guarantees** and efficient **interior point methods**
- Rely on **re-parameterizations** (does not optimize controller/policy parameters directly)



- Examples: **State-feedback** or **full-order output-feedback** H2/H∞ control, and many others

❑ Recent progress

- **Favorable properties** have been revealed for policy optimization in a range of 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] [Duan et al., 2023]

✓ H∞ state-feedback/output-feedback [Guo & Hu, 2022] [Hu & Zheng, 2022]

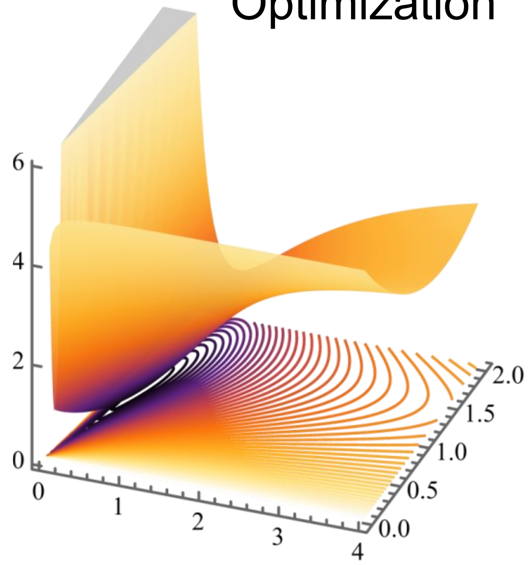
✓ A recent survey paper:

Hu, B., Zhang, K., Li, N., Mesbahi, M., Fazel, M., & Başar, T. (2023). Toward a Theoretical Foundation of Policy Optimization for Learning Control Policies. Annual Review of Control, Robotics, and Autonomous Systems, 6, 123-158.

Our Focus

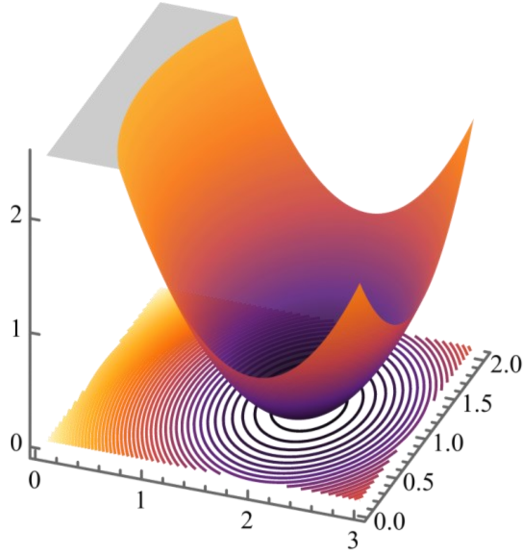
This talk: Benign Nonconvexity in Control via Extended Convex Lifting (ECL)

Nonconvex Policy Optimization



Invertible Map

Convex LMIs



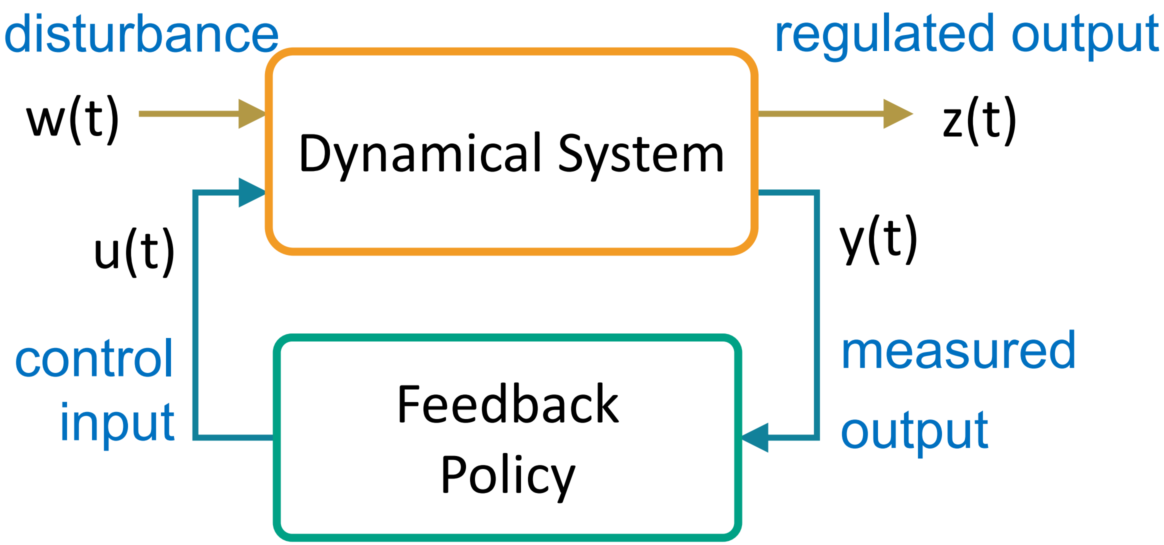
- Reconciles the gap between **nonconvex policy optimization** and **convex reformulations**.
- For a class of non-degenerate policies, **all Clarke stationary points are globally optimal** and there is **no spurious local minimum** in policy optimization.

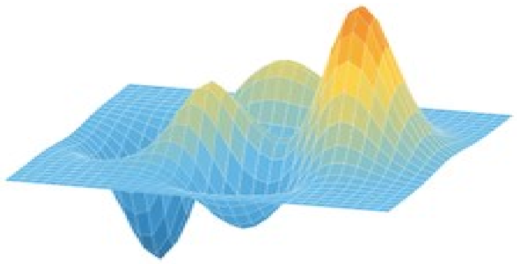
Outline

- ❑ **Problem Setup and Simple Examples**
- ❑ **Benign Nonconvexity via Extended Convex Lifting (ECL)**
- ❑ **ECLs for Optimal and Robust Control**
- ❑ **Conclusions**

Policy Optimization in Control

❑ Optimal and Robust Control





Policy parametrization

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

Non-convex Optimization problem

- Consider the class of **linear dynamic feedback policies** of the form

$$\frac{d\xi(t)}{dt} = A_K\xi(t) + B_K y(t)$$
$$u(t) = C_K\xi(t) + D_K y(t)$$

- Parametrize by K = (A_K, B_K, C_K, D_K)
- LQR, LQG, H2, Hinf robust control

System Dynamics

$$\frac{dx(t)}{dt} = Ax(t) + B_1 w(t) + B_2 u(t)$$
$$z(t) = C_1 x(t) + D_{11} w(t) + D_{12} u(t)$$
$$y(t) = C_2 x(t) + D_{21} w(t)$$

Nonconvexity in Policy Optimization

Policy parametrization

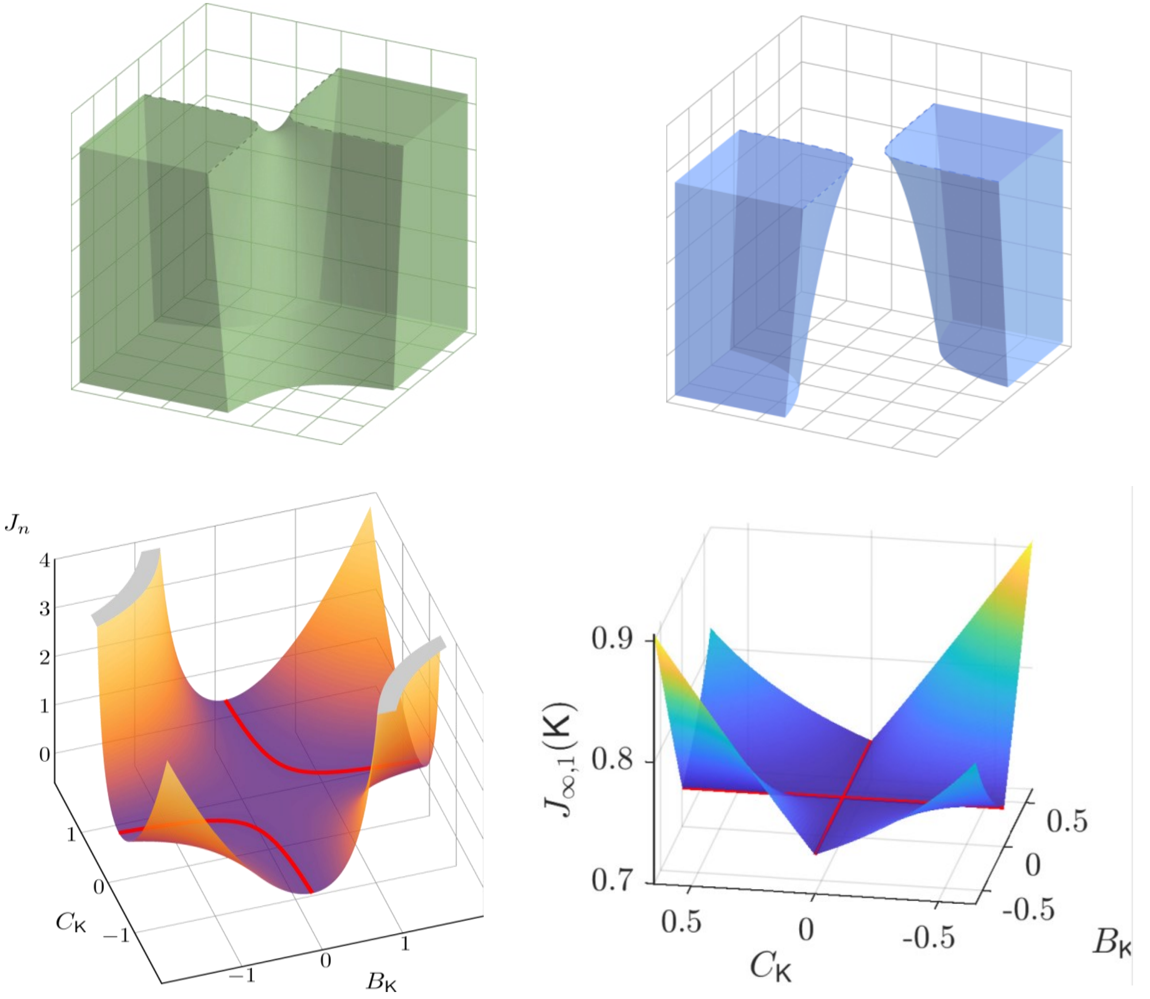
$$\min_{\mathsf{K}} \quad J(\mathsf{K})$$
$$\text{s.t.} \quad \mathsf{K} \in \mathcal{C}$$

Non-convex Optimization problem

- The set of (dynamic) stabilizing policies is **nonconvex** and even might be **not connected.** [Tang, **Zheng**, Li, 2023]
- LQR/LQG costs are **smooth but nonconvex**; Hinf cost are **non-smooth and nonconvex**

Local Stationarity + **Structural Information**

Global Optimality Certificate



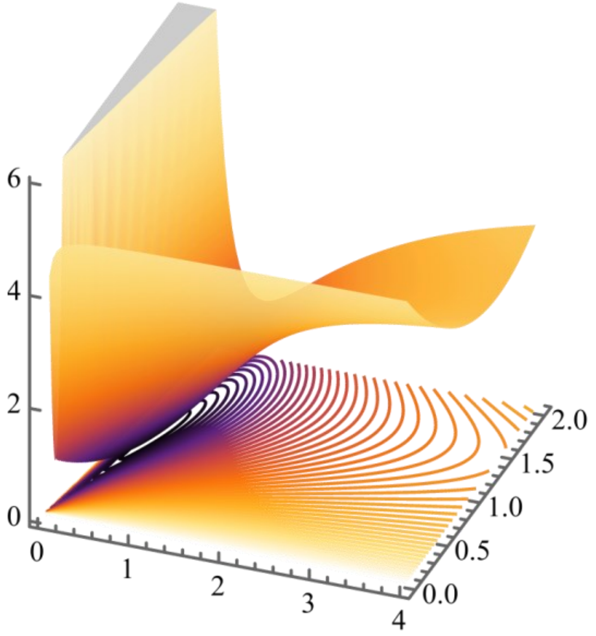
Any (non-degenerate) **Clarke stationary points are globally optimal!**

Example 1

❑ Nonconvex and Smooth function

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

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

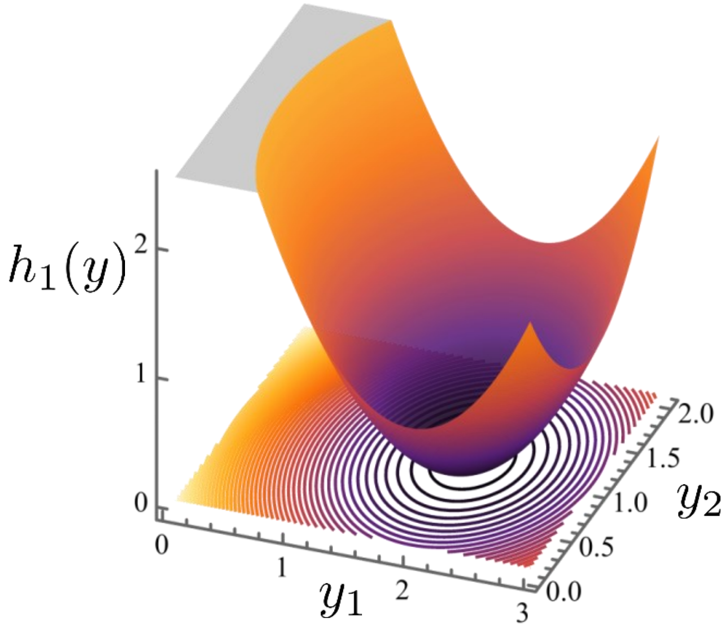


Its global minimizer is

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

Define an invertible map

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



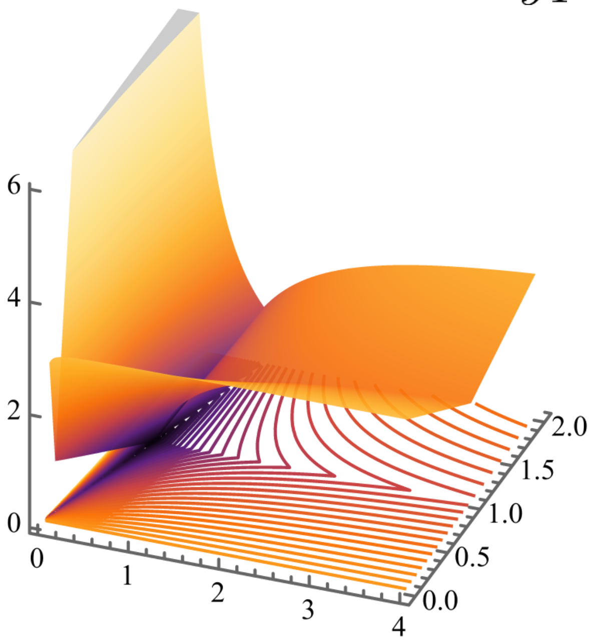
$$h_1(y) := f_1(g^{-1}(y)) = (y_1 - 2)^2 + (y_2 - 1)^2, \qquad \forall y_1 > 0, y_2 > 0.$$

Example 2

❑ Nonconvex and Non-smooth function

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

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

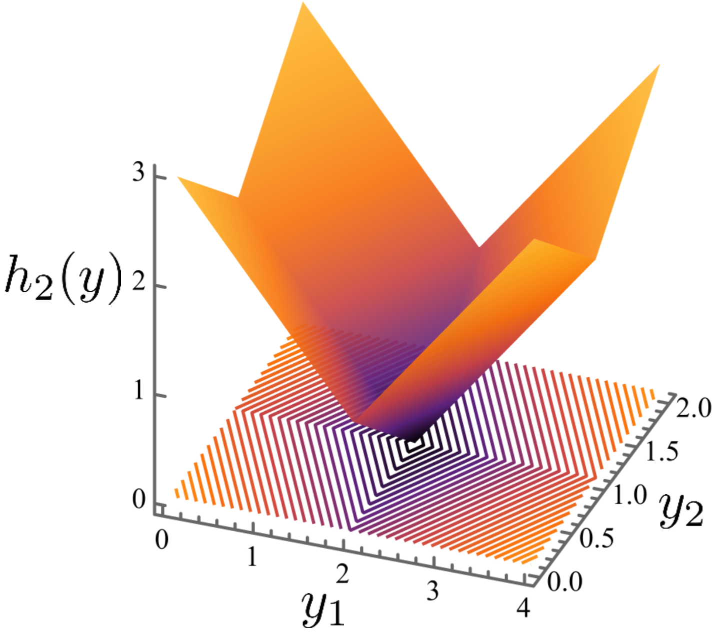


Its global minimizer is

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

Define an invertible map

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



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

Example 3

❑ Linear Quadratic Regulator (LQR)

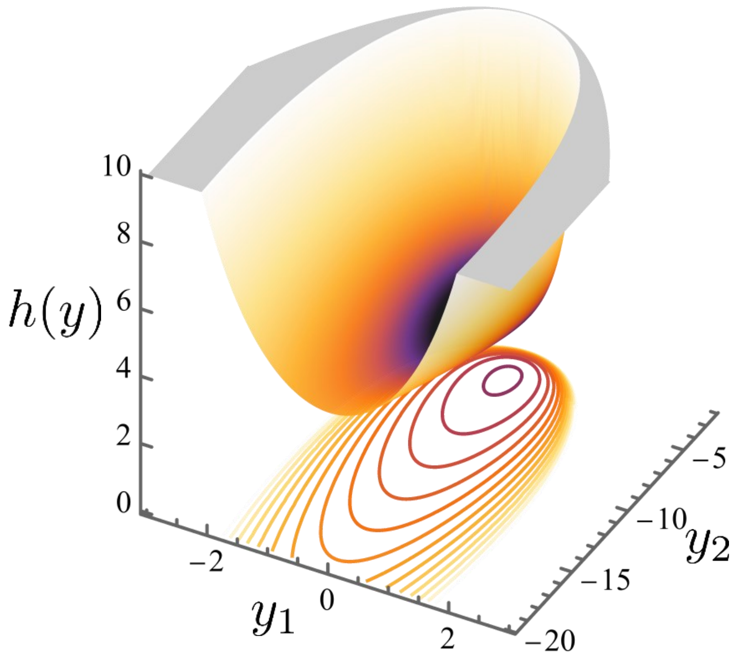
$$J(k_1, k_2) = \frac{1 - 2k_2 + 3k_2^2 - 2k_2^3 - 2k_1^2 k_2}{k_2^2 - 1}, \qquad \forall k_1 \in \mathbb{R}, k_2 < -1.$$

- 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

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

A Useful Fact

❑ Global Optimality (Informal)

- Consider a continuous function $f(x) : \mathcal{D} \to \mathbb{R}$. Denote its epigraph as

$$\mathrm{epi}_{\geq}(f) := \{(x, \gamma) \in \mathcal{D} \times \mathbb{R} \mid \gamma \geq f(x)\}.$$

- Suppose there exists a smooth and invertible map Φ between

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

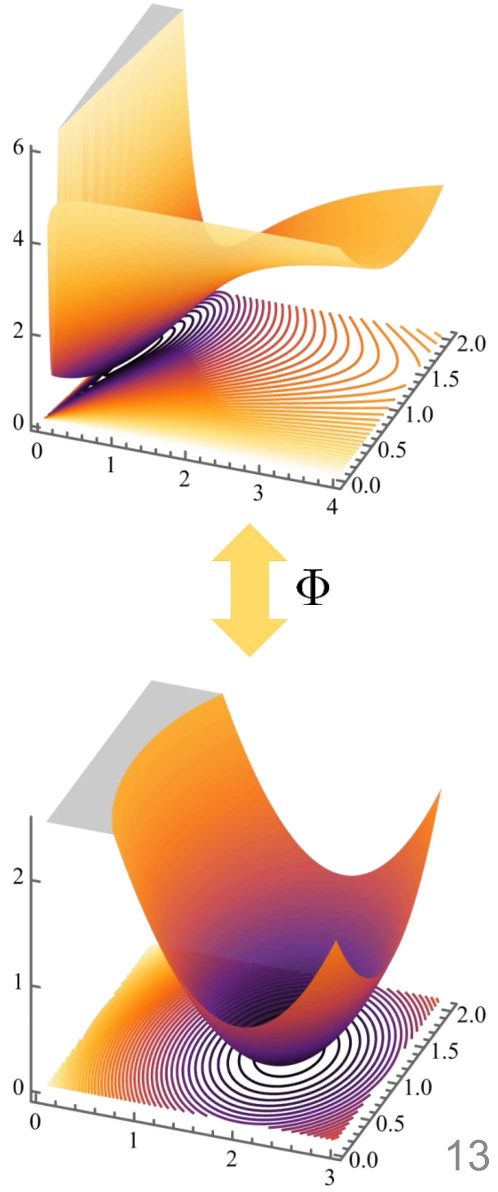
- and we further have $(y, \gamma) = \Phi(x, \gamma), \ \forall (x, \gamma) \in \mathrm{epi}_{\geq}(f)$

Guarantee 1: Optimization over f(x) is equivalent to a convex problem

$$\inf_{x \in \mathcal{D}} f(x) = \inf_{(y,\gamma) \in \mathcal{F}_{\mathrm{cvx}}} \gamma.$$

Guarantee 2: Any stationary point to f(x) is globally optimal

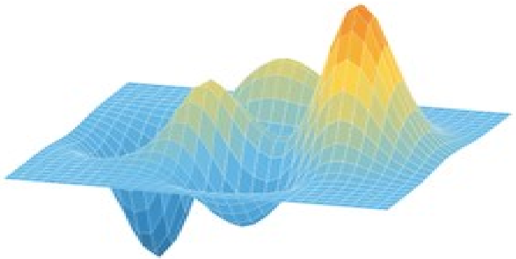
$0 \in \partial f(x^*)$ **implies globally optimality**



Outline

- ❑ Problem Setup and Simple Examples
- ❑ **Benign Nonconvexity via Extended Convex Lifting (ECL)**
- ❑ ECLs for Optimal and Robust Control
- ❑ Conclusions

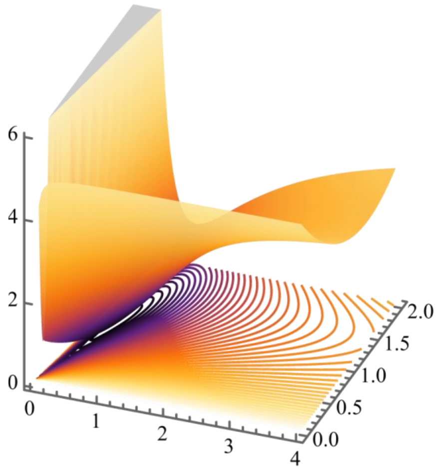
Lifting for Convexity



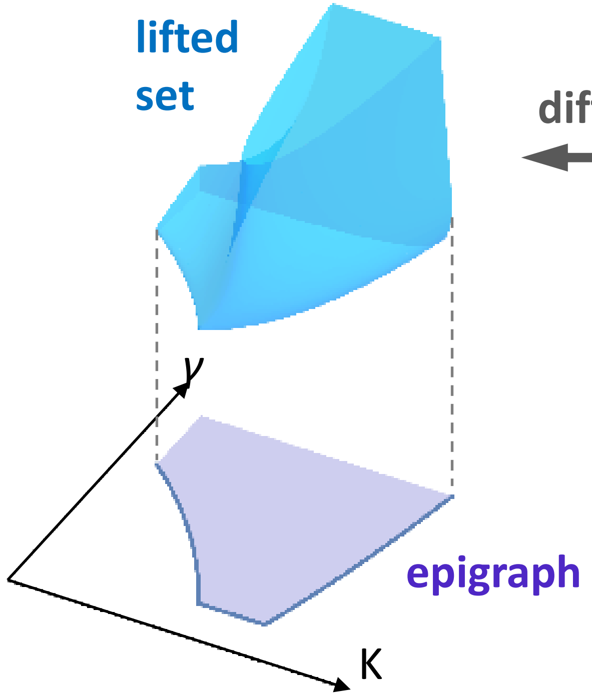
$$\min_{\mathsf{K}} \quad J(\mathsf{K})$$

$$\text{s.t.} \quad \mathsf{K} \in \mathcal{C}$$

Non-convex Optimization problem



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



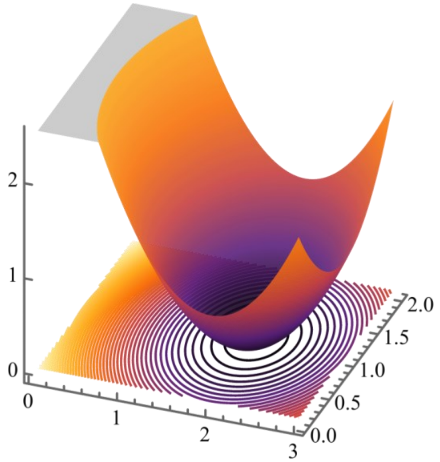
diffeomorphism ⟷ $\mathcal{F} \times \mathrm{GL}_n$

convex ⇓ **Global optimality**

set of invertible matrices

Φ

Does Not Exist!



Extended Convex Lifting (ECL)

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

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

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 f(x) if

- **Condition 1**: $\mathcal{L}_{\mathrm{lft}} \subseteq \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{d_\xi}$ is a lifted set such that its canonical projection satisfies

$$\mathrm{epi}_>(f) \subseteq \pi_{x,\gamma}(\mathcal{L}_{\mathrm{lft}}) \subseteq \mathrm{cl}\,\mathrm{epi}_>(f).$$

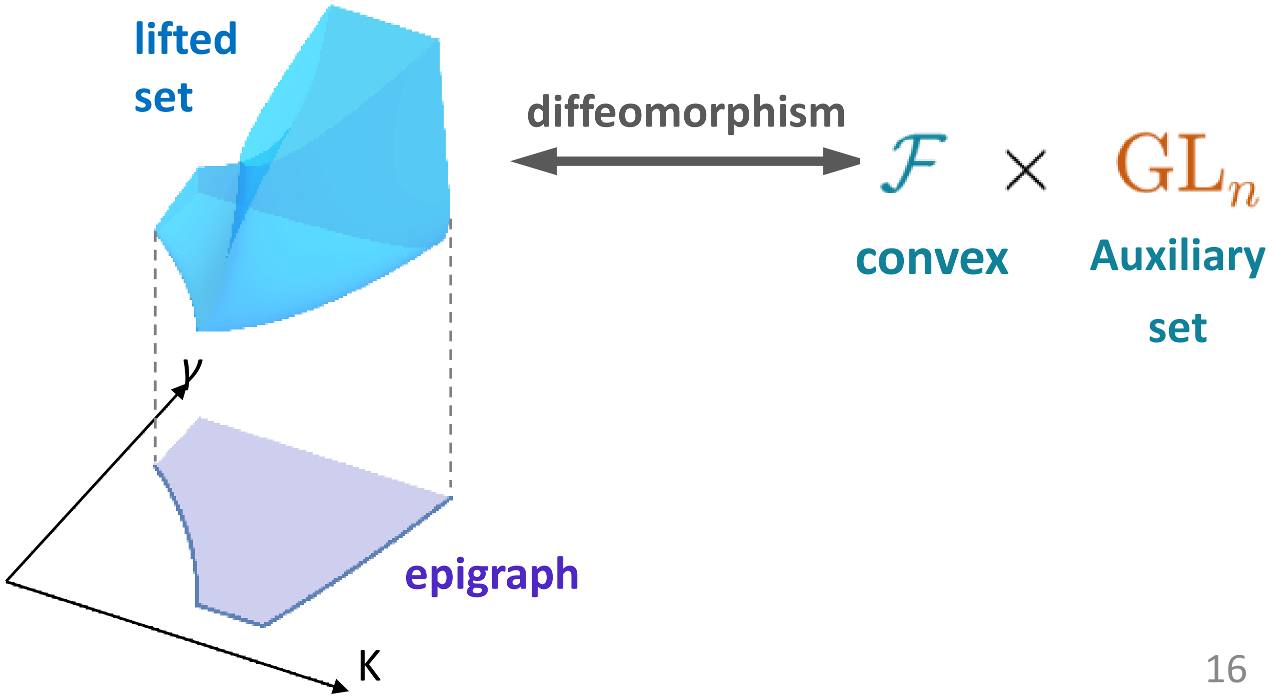
- **Condition 2**: Φ is a diffeomorphism between

$\mathcal{L}_{\mathrm{lft}}$ and $\mathcal{F}_{\mathrm{cvx}} \times \mathcal{G}_{\mathrm{aux}}$

convex Auxiliary set

- **Condition 3**: Φ does not change γ

$$\Phi(x, \gamma, \xi) = (\gamma, \zeta_1, \zeta_2), \quad \forall (x, \gamma, \xi) \in \mathcal{L}_{\mathrm{lft}}$$



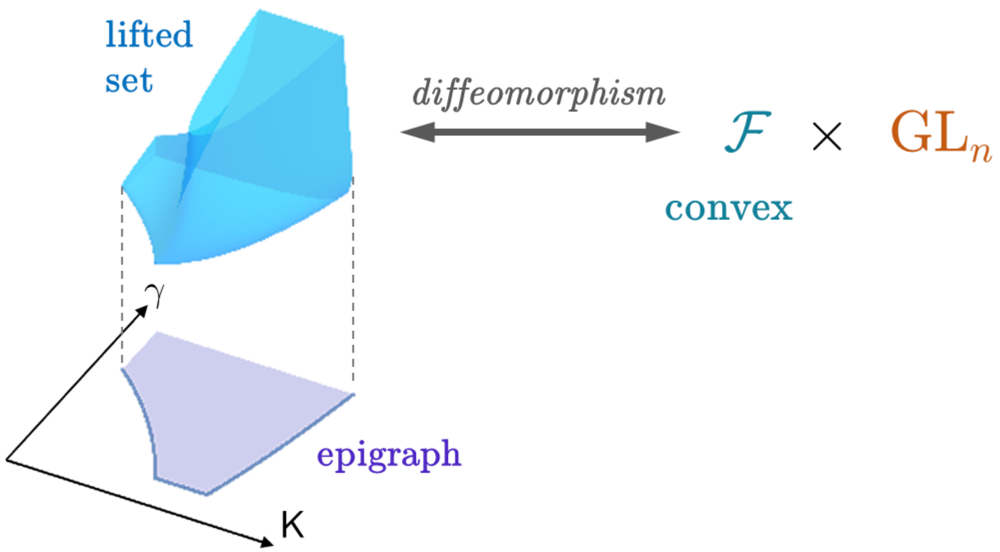
Extended Convex Lifting (ECL)

- If a continuous function $f(x) : \mathcal{D} \to \mathbb{R}$ where $\mathcal{D} \subseteq \mathbb{R}^d$ admits an ECL $(\mathcal{L}_{\text{lft}}, \mathcal{F}_{\text{cvx}}, \mathcal{G}_{\text{aux}}, \Phi)$

Guarantee 1: Optimization over f(x) is equivalent to a convex problem

$$\inf_{x \in \mathcal{D}} f(x) = \inf_{(y,\gamma) \in \mathcal{F}_{\text{cvx}}} \gamma.$$

Guarantee 2: If f(x) is subdifferentially regular, any non-degenerate Clarke stationary point is globally optimal



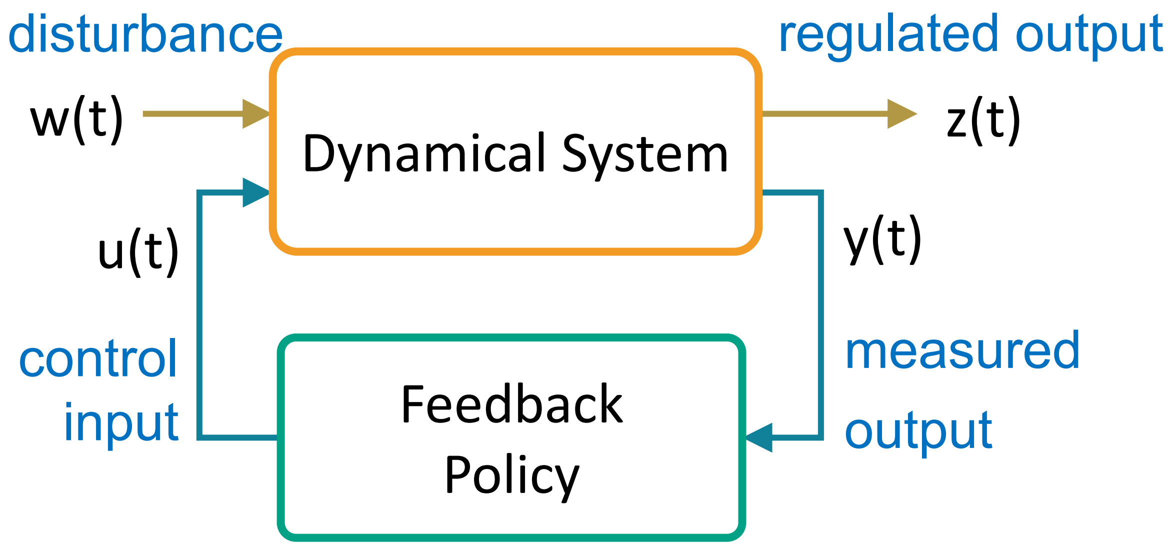
- Subdifferentially regular functions are a very class of functions, including all cost functions in **LQR**, **LQG**, **Hinf robust control**, etc.
- **Non-degenerate points:** covered by the lifted set. $x \in \mathcal{D}$ such that $(x, f(x)) \in \pi_{x,\gamma}\mathcal{L}_{\text{lft}})$.
  - ✓ If we have $\pi_{x,\gamma}(\mathcal{L}_{\text{lft}}) = \text{epi}_{\geq}(f)$, then all feasible points are non-degenerate

Outline

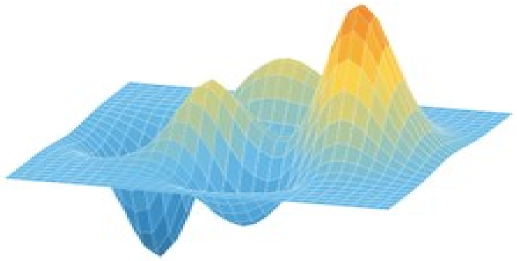
- ❑ Simple Examples
- ❑ Benign Nonconvexity via Extended Convex Lifting (ECL)
- ❑ **ECLs for Optimal and Robust Control**
- ❑ Conclusions

Global Optimality in Control

❑ Optimal and Robust Control



Policy parametrization



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

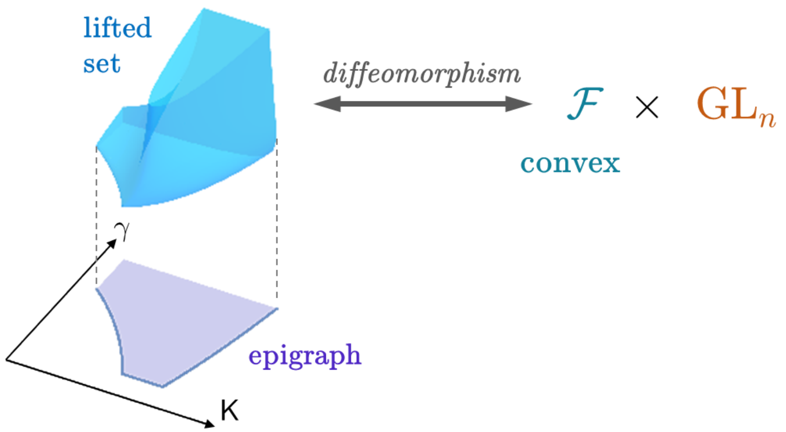
Non-convex Optimization problem

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.

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

❑ Building an ECL

Step 1: Lifting

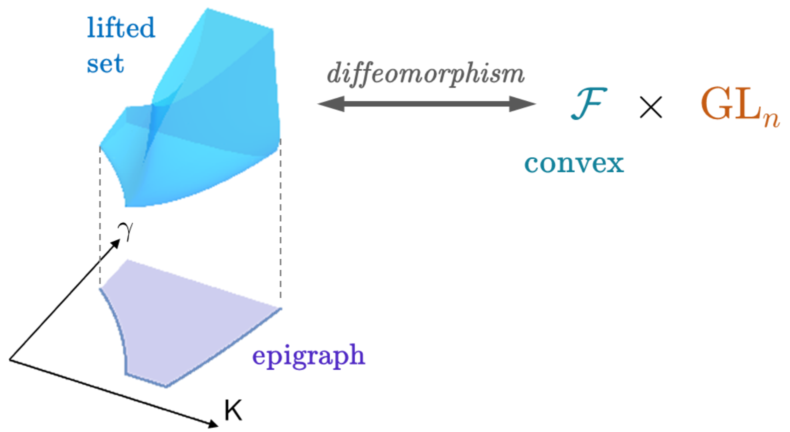
$$\mathcal{L}_{\mathrm{LQR}} := \left\{(K, \gamma, X) : X \succ 0, (A+BK)X + X(A+BK)^\mathsf{T} + W = 0, \gamma \geq \operatorname{Tr}\left[(Q + K^\mathsf{T}RK)X\right]\right\}.$$

Step 2: Convex set

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

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

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

❑ Building an ECL

Step 1: Lifting

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

From the KYP lemma, we can show that $\pi_{K,\gamma}(\mathcal{L}_\infty) = \mathrm{epi}_\geq(J_\infty)$

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

Any Clarke stationary points $0 \in \partial J_\infty(K)$ are globally optimal!

Nonconvexity in Policy Optimization

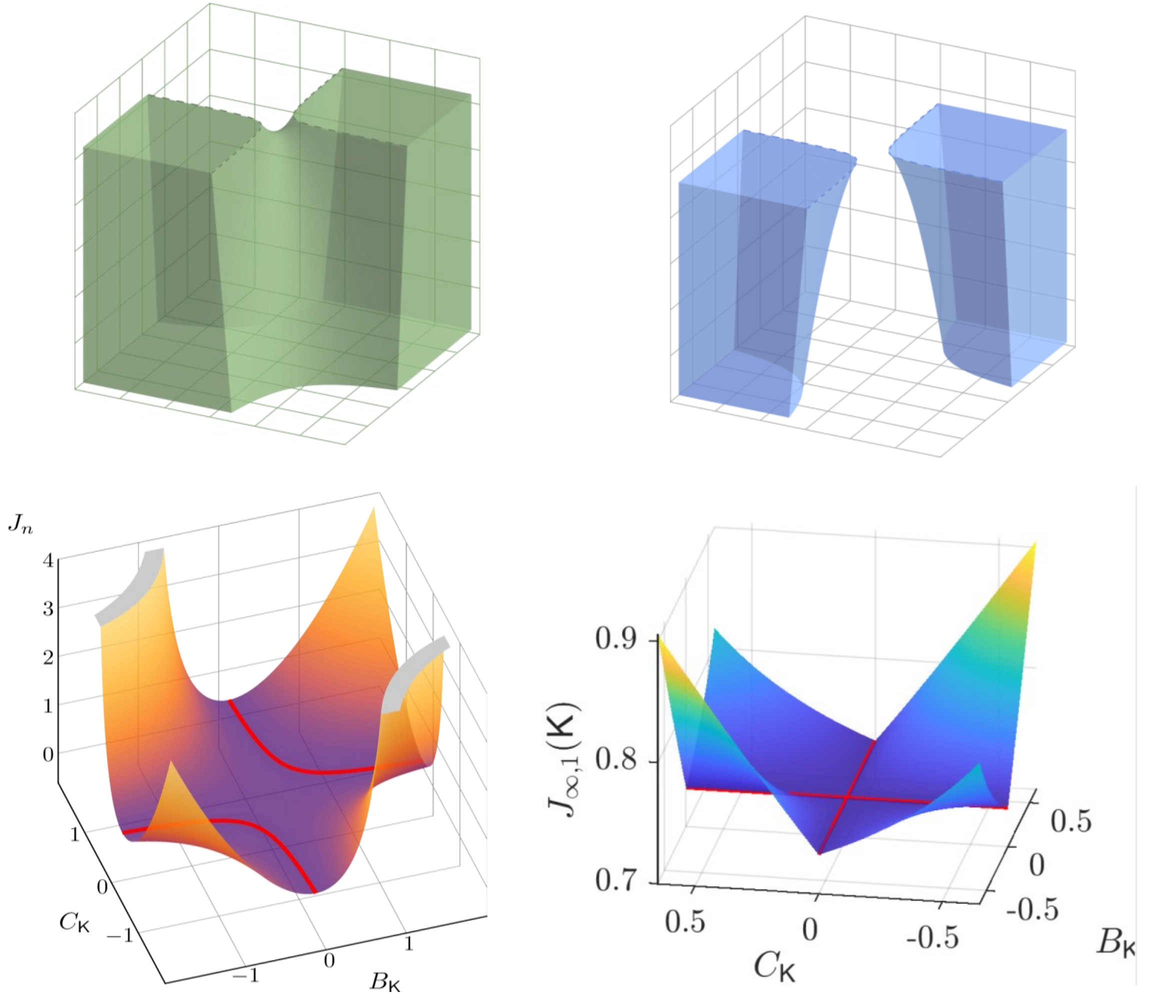
- The set of (dynamic) stabilizing policies is **nonconvex** and even might be **not connected**.

- LQR/LQG costs are **smooth but nonconvex**; Hinf cost are **non-smooth and nonconvex**

- An ECL confirms that any (non-degenerate) **Clarke stationary points are globally optimal**

Local Stationarity + **Structural Information**

Global Optimality Certificate

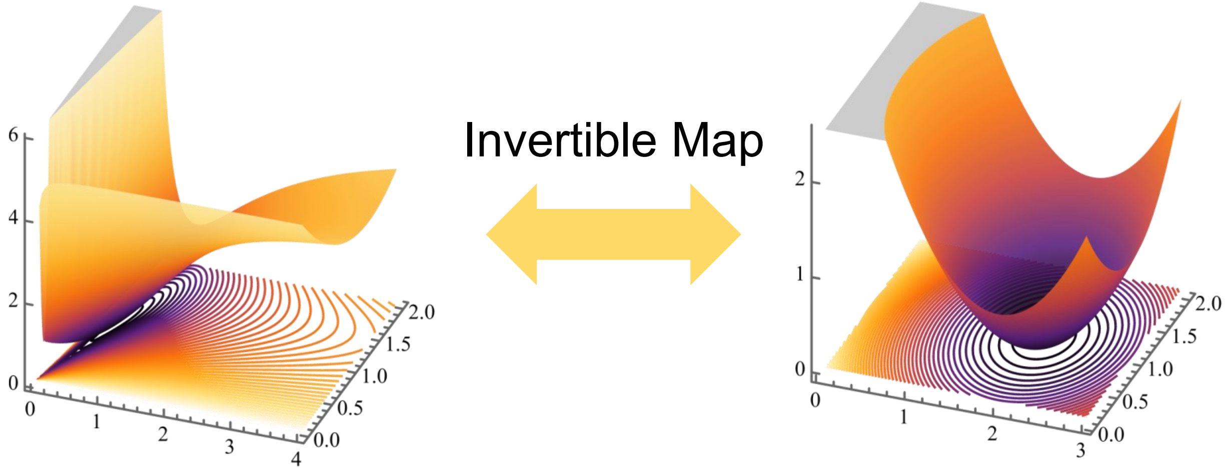


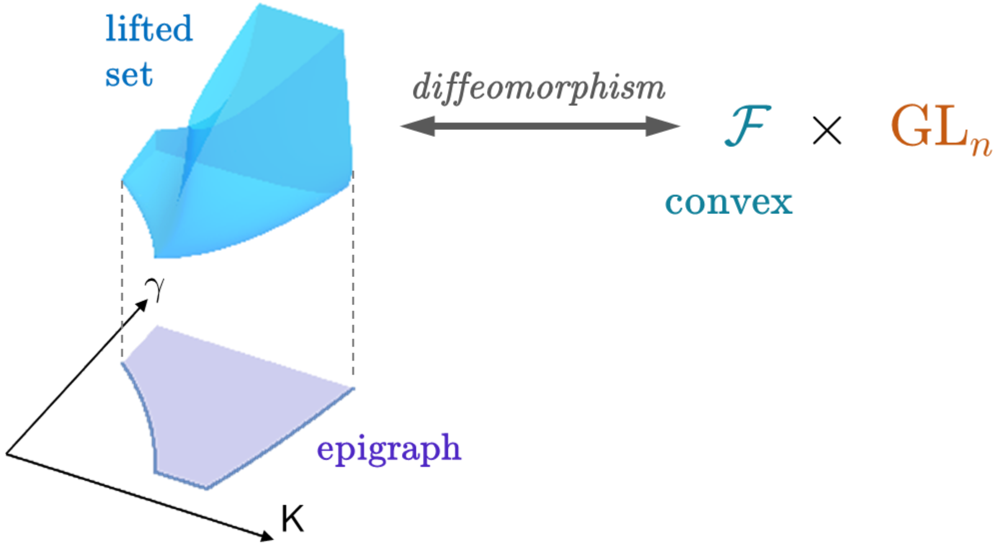
Outline

- ❑ Simple Examples
- ❑ Benign Nonconvexity via Extended Convex Lifting (ECL)
- ❑ ECLs for Optimal and Robust Control
- ❑ **Conclusions**

Nonconvex Policy Optimization for control

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





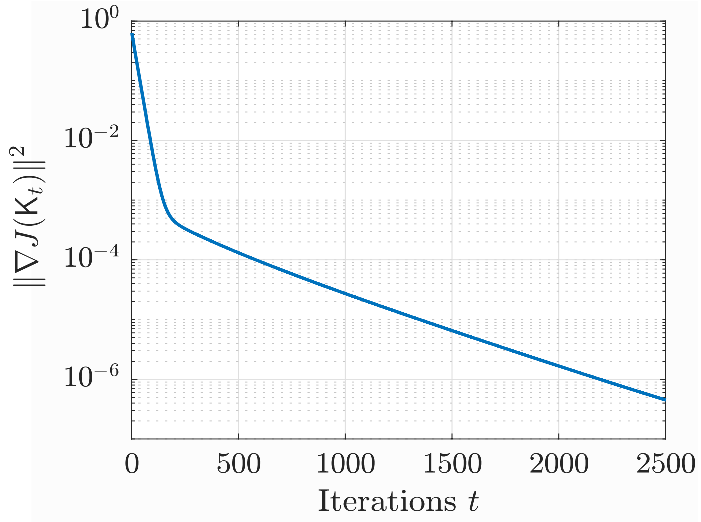
- **Global Optimality**

Local Stationarity + **Structural Information**

Global Optimality Certificate

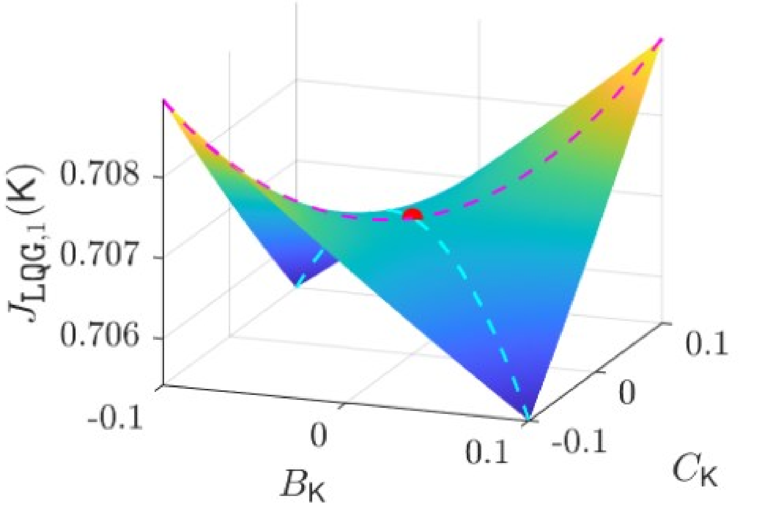
Ongoing and Future work

- How to design efficient local search algorithms?
- How to establish convergence conditions and speeds?
- How to deal with degenerate points in local policy search? Avoiding saddle points?



Control Theory ⟷ (Non)convex Optimization

Reinforcement Learning



Benign Nonconvex Landscapes in Optimal and Robust Control

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**." preprint arXiv:2312.15332 (2023): https://arxiv.org/abs/2312.15332.

- **Part II: Extended Convex Lifting** will be out soon