Policy Optimization for Linear Quadratic Gaussian (LQG) Control

Yang Zheng

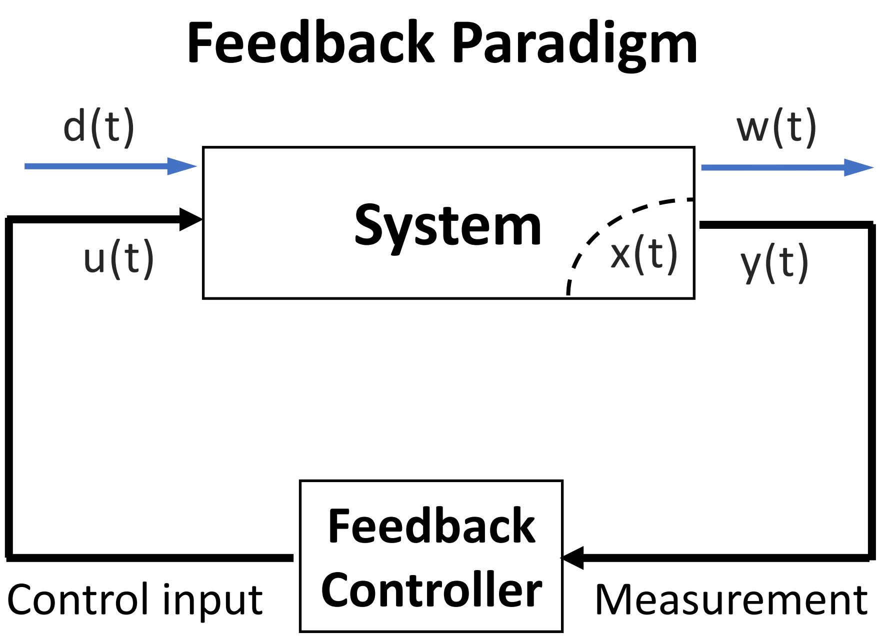
Assistant Professor,
ECE Department, UC San Diego

Joint work with Yujie Tang, Na Li, Maryam Fazel, and Yue Sun

2023 American Control Conference (ACC) Workshop,
San Diego, CA. May 30, 2023

Today's talk

❑ Optimal Control



Control theory: the principled use of feedback loops and algorithms to drive a dynamical system to its desired goal

Linear Quadratic Optimal control

$$\min_{u_1, u_2, \ldots,} \quad \lim_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\left(x_t^{\mathsf{T}} Q x_t + u_t^{\mathsf{T}} R u_t\right)\right]$$

$$\text{subject to} \quad x_{t+1} = Ax_t + Bu_t + w_t$$

$$y_t = Cx_t + v_t$$

- Many practical applications
- **Linear Quadratic Regulator (LQR)** when the state x_t is directly observable
- **Linear Quadratic Gaussian (LQG) control** when only partial output y_t is observed
- Extensive classical results (Dynamic programming, Separation principle, Riccati equations, etc.)

They are all model-based. Are there any guarantees for non-convex policy optimization?

Direct policy optimization

❑ Controller parameterization

- Give a parameterization of control policies; say **neural networks?** ✖
- Control theory already tells us many structural properties
- **Linear feedback is sufficient for LQR**

$$u_t = K x_t, \qquad K \in \mathbb{R}^{m \times n}$$

$$\lim_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\sum_{t=1}^{T}\left(x_t^{\mathsf{T}} Q x_t + u_t^{\mathsf{T}} R u_t\right)\right] := J(K)$$

- Set of stabilizing controllers $K \in \mathcal{K}$
- A fast-growing list of references

Policy optimization for LQR

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

Direct policy iteration

$$K_{i+1} = K_i - \alpha_i \nabla J(K_i)$$

✓ Good optimization landscape properties (Fazel et al., 2018); **Morning session**!

- Connected feasible region
- Unique stationary point
- Gradient dominance

✓ Fast global convergence (linear)

➢ Fazel et al., ICML, 2018; Malik et al., 2019; Mohammadi et al., IEEE TAC, 2020; Li et al., 2019; K. Zhang, B. Hu, and T. Başar, 2021; Furieri et al., 2019; Feiran Zhao & Keyou You, 2021, and many others

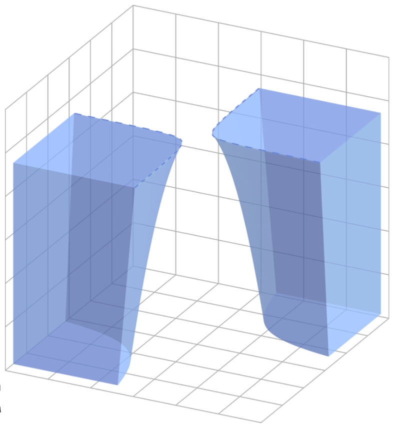
Challenges for partially observed LQG

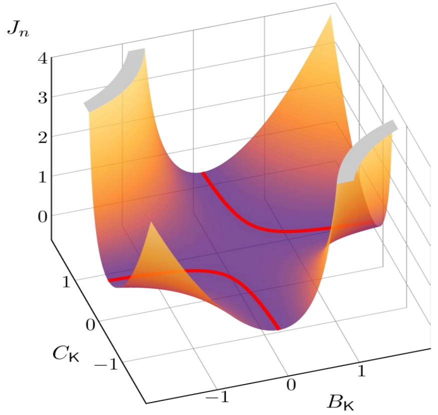
❑ Policy optimization for LQG control

- LQG is more complicated than LQR
- Requires dynamical controllers
- Its non-convex landscape properties are much richer and more complicated than LQR

Our focus: non-convex LQG landscape

- **Q1: Properties of the domain** (set of stabilizing controllers)
  - convexity, connectivity, open/closed?
- **Q2: Properties of the accumulated LQG cost**
  - convexity, differentiability, coercivity?
  - set of stationary points/local minima/global minima?
- **Q3: How to escape saddle points via Perturbed Gradient Descent (PGD)**





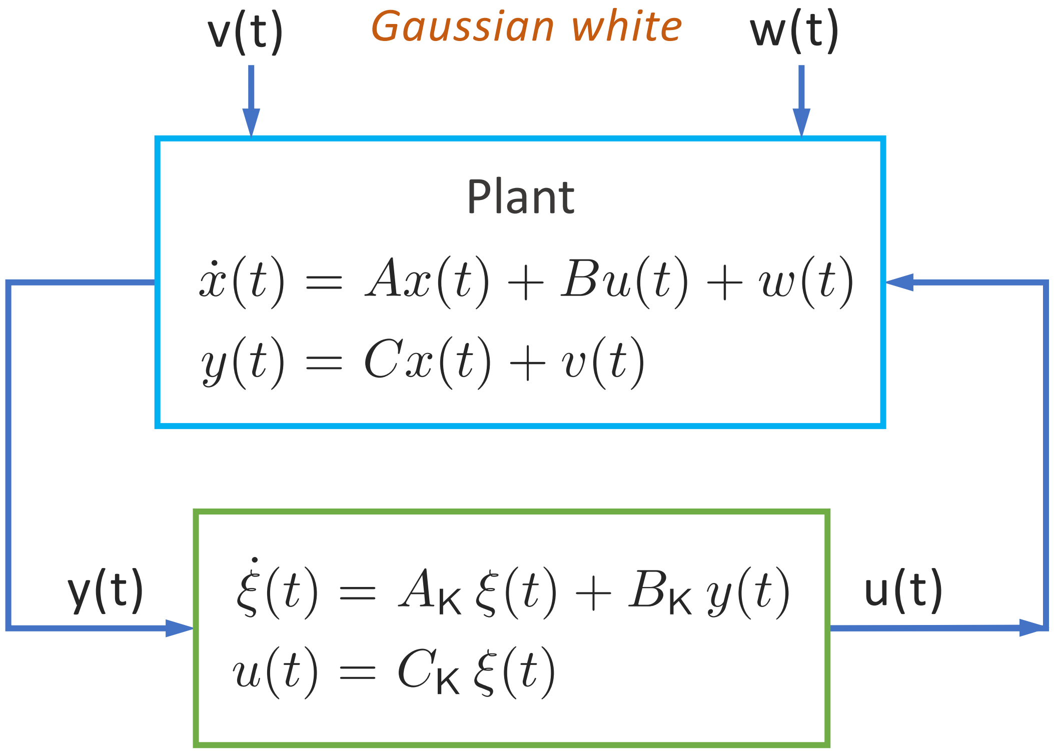
Outline

- LQG problem Setup
- Connectivity of the Set of Stabilizing Controllers
- Structure of Stationary Points of the LQG cost
- Escaping saddle points via PGD

Outline

- ❑ LQG problem Setup
- ❑ Connectivity of the Set of Stabilizing Controllers
- ❑ Structure of Stationary Points of the LQG cost
- ❑ Escaping saddle points via PGD

LQG Problem Setup



dynamical controller

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

Standard Assumption		
	$(A, B),\ (A, W^{1/2})$	Controllable
	$(C, A),\ (Q^{1/2}, A)$	Observable

Objective: The LQG cost

$$\lim_{T \to +\infty} \frac{1}{T} \mathbb{E} \int_0^T (x^\top Q x + u^\top R u)\, dt$$

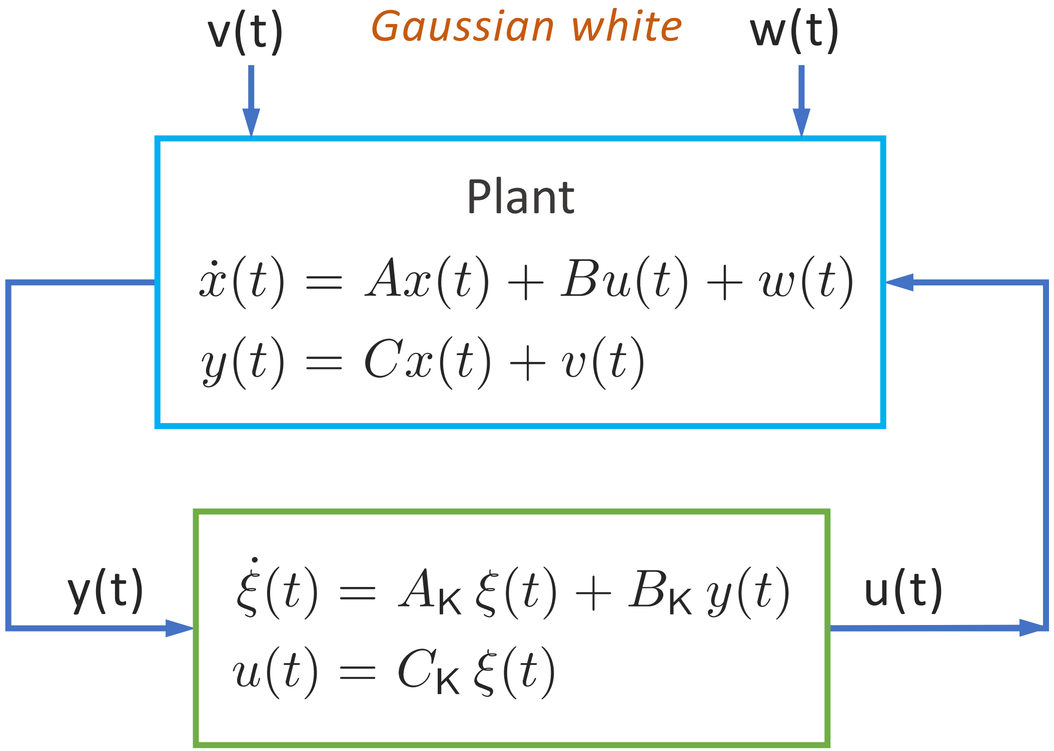
- $\xi(t)$ internal state of the controller
- $\dim \xi(t)$ order of the controller
- $\dim \xi(t) = \dim x(t)$ full-order
- $\dim \xi(t) < \dim x(t)$ reduced-order

Minimal controller

The input-output behavior cannot be replicated by a lower order controller.

* $(A_{\mathsf{K}}, B_{\mathsf{K}}, C_{\mathsf{K}})$ controllable and observable

Separation principle



dynamical controller

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

Explicit dependence on the dynamics

Objective: The LQG cost

$$\lim_{T \to +\infty} \frac{1}{T} \mathbb{E} \int_0^T (x^\top Q x + u^\top R u)\, dt$$

Solution: Kalman filter for state estimation
+ LQR based on the estimated state

$$\dot{\xi} = (A - BK)\xi + L(y - C\xi),$$
$$u = -K\xi.$$

Two Riccati equations

- **Kalman gain** $L = PC^{\mathsf{T}}V^{-1}$

$$AP + PA^{\mathsf{T}} - PC^{\mathsf{T}}V^{-1}CP + W = 0,$$

- **Feedback gain** $K = R^{-1}B^{\mathsf{T}}S$

$$A^{\mathsf{T}}S + SA - SBR^{-1}B^{\mathsf{T}}S + Q = 0$$

Policy Optimization formulation

❑ Closed-loop dynamics

$$\frac{d}{dt}\begin{bmatrix} x \\ \xi \end{bmatrix} = \begin{bmatrix} A & BC_{\mathsf{K}} \\ B_{\mathsf{K}}C & A_{\mathsf{K}} \end{bmatrix}\begin{bmatrix} x \\ \xi \end{bmatrix} + \begin{bmatrix} I & 0 \\ 0 & B_{\mathsf{K}} \end{bmatrix}\begin{bmatrix} w \\ v \end{bmatrix},$$

$$\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} C & 0 \\ 0 & C_{\mathsf{K}} \end{bmatrix}\begin{bmatrix} x \\ \xi \end{bmatrix} + \begin{bmatrix} v \\ 0 \end{bmatrix}.$$

❑ Feasible region of the controller parameters

$$\mathcal{C}_{\text{full}} = \left\{ \mathsf{K} \,|\, \mathsf{K} = (A_{\mathsf{K}}, B_{\mathsf{K}}, C_{\mathsf{K}}) \text{ is full order } \begin{bmatrix} A & BC_{\mathsf{K}} \\ B_{\mathsf{K}}C & A_{\mathsf{K}} \end{bmatrix} \text{ is Hurwitz stable} \right\}$$

❑ Cost function $\lim_{T\to+\infty} \frac{1}{T}\mathbb{E}\int_0^T (x^\top Q x + u^\top R u)\, dt$

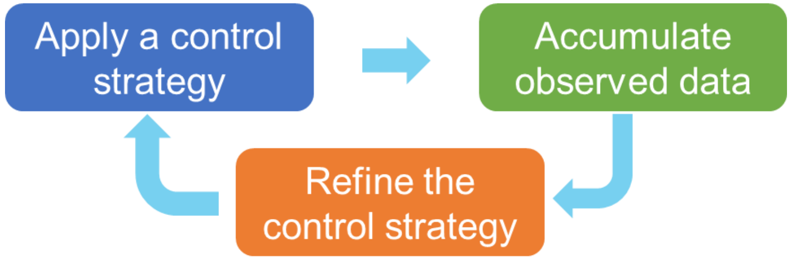
$$J(\mathsf{K}) = \operatorname{tr}\left(\begin{bmatrix} Q & 0 \\ 0 & C_{\mathsf{K}}^{\mathsf{T}} R C_{\mathsf{K}} \end{bmatrix} X_{\mathsf{K}}\right) = \operatorname{tr}\left(\begin{bmatrix} W & 0 \\ 0 & B_{\mathsf{K}} V B_{\mathsf{K}}^{\mathsf{T}} \end{bmatrix} Y_{\mathsf{K}}\right)$$

$X_{\mathsf{K}}, Y_{\mathsf{K}}$ Solution to Lyapunov equations

Policy optimization for LQG

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

Direct policy iteration $\mathsf{K}_{i+1} = \mathsf{K}_i - \alpha_i \nabla J(\mathsf{K}_i)$

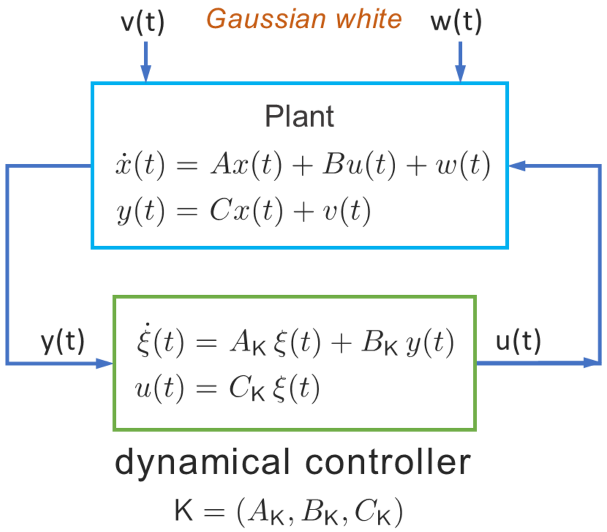


- ✓ **Does it converge at all?**
- ✓ **Converge to which point?**
- ✓ **Convergence speed?**

Optimization Landscape Analysis

Hyland, David, and Dennis Bernstein. "The optimal projection equations for fixed-order dynamic compensation." *IEEE Transactions on Automatic Control* 29.11 (1984): 1034-1037.

Main questions



Policy optimization for LQG

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

Non-convex Landscape Analysis

- **Q1: Connectivity of the feasible region $\mathcal{C}_{\text{full}}$**
  - Is it connected?
  - If not, how many connected components can it have?
- **Q2: Structure of stationary points of $J(\mathsf{K})$**
  - Are there spurious (strictly suboptimal, saddle) stationary points?
  - How to check if a stationary point is globally optimal?
- **Q3: How to escape high-order saddle points via PGD?**

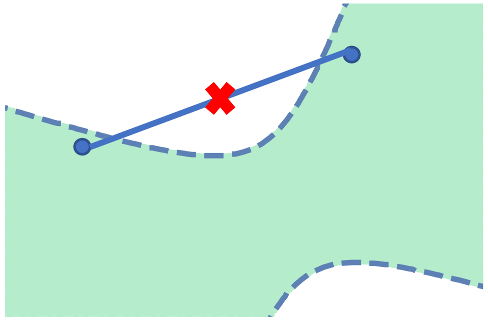
Outline

- ❑ LQG problem Setup
- ❑ **Connectivity of the Set of Stabilizing Controllers**
- ❑ Structure of Stationary Points of the LQG cost
- ❑ Escaping saddle points via PGD

Connectivity of the feasible region

❑ Simple observation: non-convex and unbounded

Lemma 1: the set $\mathcal{C}_{\text{full}}$ is non-empty, unbounded, and can be non-convex.



Example

$$\dot{x}(t) = x(t) + u(t) + w(t)$$
$$y(t) = x(t) + v(t)$$

$$\mathcal{C}_{\text{full}} = \left\{ \mathsf{K} = \begin{bmatrix} 0 & C_{\mathsf{K}} \\ B_{\mathsf{K}} & A_{\mathsf{K}} \end{bmatrix} \in \mathbb{R}^{2\times 2} \middle| \begin{bmatrix} 1 & C_{\mathsf{K}} \\ B_{\mathsf{K}} & A_{\mathsf{K}} \end{bmatrix} \text{ is stable} \right\}.$$

$$\mathsf{K}^{(1)} = \begin{bmatrix} 0 & 2 \\ -2 & -2 \end{bmatrix}, \qquad \mathsf{K}^{(2)} = \begin{bmatrix} 0 & -2 \\ 2 & -2 \end{bmatrix}$$

Stabilize the plant, and thus belong to $\mathcal{C}_{\text{full}}$

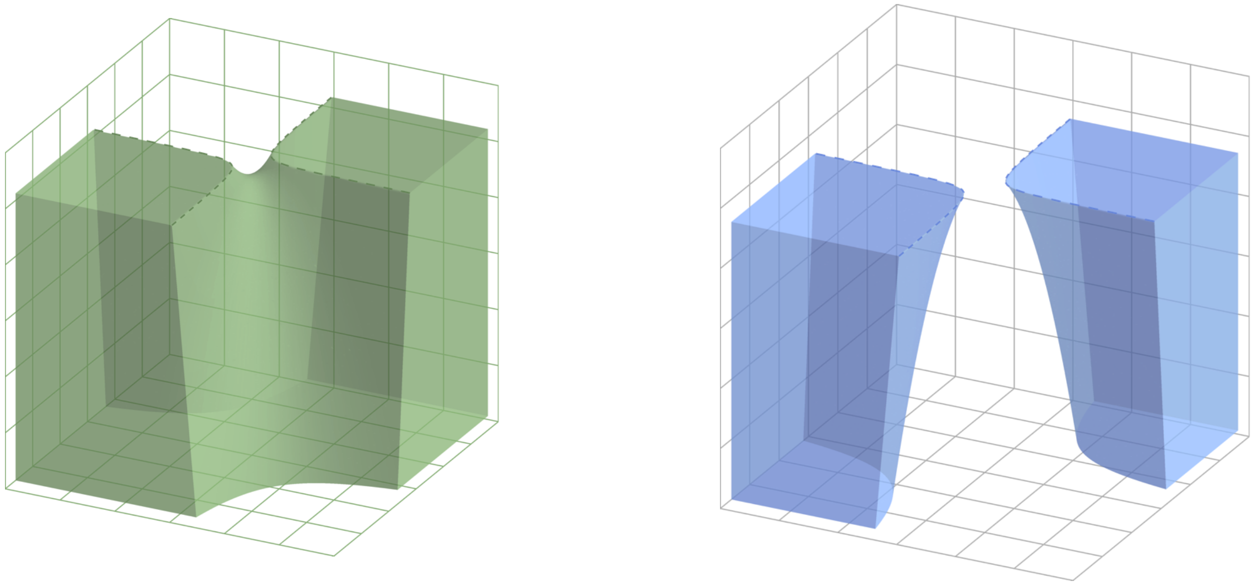
$$\hat{\mathsf{K}} = \frac{1}{2}\left(\mathsf{K}^{(1)} + \mathsf{K}^{(2)}\right) = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix}$$

Fails to stabilize the plant, and thus outside $\mathcal{C}_{\text{full}}$

Connectivity of the feasible region

❑ Main Result 1: dis-connectivity

Theorem 1: The set $\mathcal{C}_{\text{full}}$ can be disconnected but has at most 2 connected components.

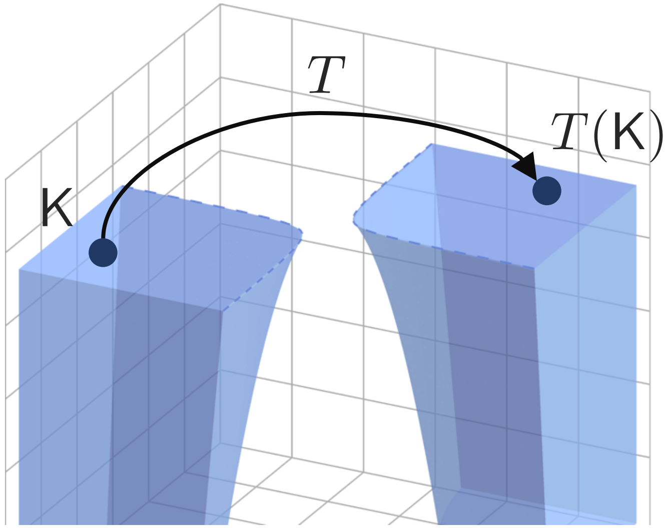


- ✓ Different from the connectivity of static stabilizing state-feedback controllers, which is always connected!
- ✓ **Is this a negative result for gradient-based algorithms?** → **No**

Connectivity of the feasible region

❑ Main Result 2: dis-connectivity

Theorem 2: If $\mathcal{C}_{\text{full}}$ has 2 connected components, then there is a smooth bijection T between the 2 connected components that has the same cost function value.



$$J(\mathsf{K}) = J(T(\mathsf{K}))$$

✓ In fact, the bijection T is defined by a similarity transformation (change of controller state coordinates)

$$\mathscr{T}_T(\mathsf{K}) := \begin{bmatrix} D_{\mathsf{K}} & C_{\mathsf{K}} T^{-1} \\ T B_{\mathsf{K}} & T A_{\mathsf{K}} T^{-1} \end{bmatrix}.$$

Positive news: For gradient-based local search methods, it makes no difference to search over either connected component.

Connectivity of the feasible region

Main Result 3: conditions for connectivity

Theorem 3: 1) $\mathcal{C}_{\text{full}}$ is connected if there exists a reduced-order stabilizing controller.

2) The sufficient condition above becomes necessary if the plant is single-input or single-output.

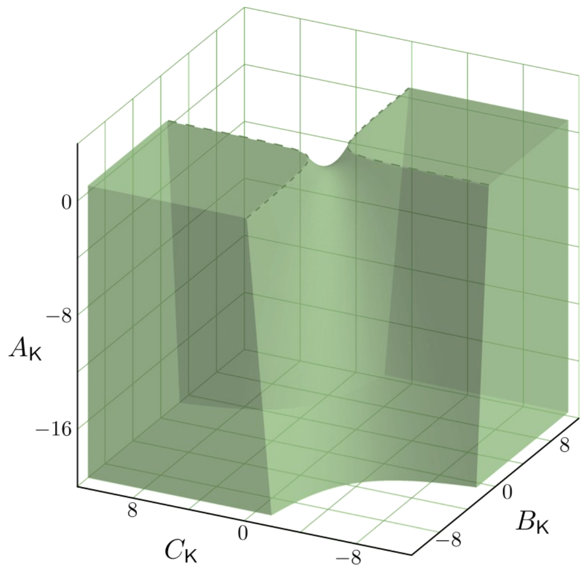
Corollary 1: Given any open-loop stable plant, the set of stabilizing controllers $\mathcal{C}_{\text{full}}$ is connected.

Example: Open-loop stable system

$$\dot{x}(t) = -x(t) + u(t) + w(t)$$
$$y(t) = x(t) + v(t)$$

Routh--Hurwitz stability criterion

$$\mathcal{C}_{\text{full}} = \left\{ \mathsf{K} = \begin{bmatrix} 0 & C_{\mathsf{K}} \\ B_{\mathsf{K}} & A_{\mathsf{K}} \end{bmatrix} \in \mathbb{R}^{2\times 2} \middle| A_{\mathsf{K}} < 1, B_{\mathsf{K}} C_{\mathsf{K}} < -A_{\mathsf{K}} \right\}.$$



Connectivity of the feasible region

❑ Main Result 3: conditions for connectivity

Example: Open-loop unstable system (SISO)

$$\dot{x}(t) = x(t) + u(t) + w(t)$$
$$y(t) = x(t) + v(t)$$

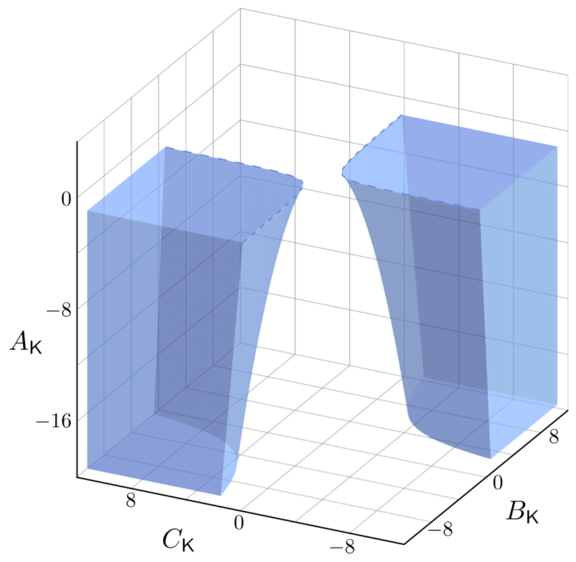
- **Routh--Hurwitz stability criterion**

$$\mathcal{C}_{\text{full}} = \left\{ \mathsf{K} = \begin{bmatrix} 0 & C_{\mathsf{K}} \\ B_{\mathsf{K}} & A_{\mathsf{K}} \end{bmatrix} \in \mathbb{R}^{2\times 2} \middle| \begin{bmatrix} A & BC_{\mathsf{K}} \\ B_{\mathsf{K}}C & A_{\mathsf{K}} \end{bmatrix} \text{ is stable} \right\}$$
$$= \left\{ \mathsf{K} = \begin{bmatrix} 0 & C_{\mathsf{K}} \\ B_{\mathsf{K}} & A_{\mathsf{K}} \end{bmatrix} \in \mathbb{R}^{2\times 2} \middle| A_{\mathsf{K}} < -1,\ B_{\mathsf{K}}C_{\mathsf{K}} < A_{\mathsf{K}} \right\}.$$

- **Two path-connected components**

$$\mathcal{C}_1^+ := \left\{ \mathsf{K} = \begin{bmatrix} 0 & C_{\mathsf{K}} \\ B_{\mathsf{K}} & A_{\mathsf{K}} \end{bmatrix} \in \mathbb{R}^{2\times 2} \middle| A_{\mathsf{K}} < -1,\ B_{\mathsf{K}}C_{\mathsf{K}} < A_{\mathsf{K}},\ B_{\mathsf{K}} > 0 \right\},$$
$$\mathcal{C}_1^- := \left\{ \mathsf{K} = \begin{bmatrix} 0 & C_{\mathsf{K}} \\ B_{\mathsf{K}} & A_{\mathsf{K}} \end{bmatrix} \in \mathbb{R}^{2\times 2} \middle| A_{\mathsf{K}} < -1,\ B_{\mathsf{K}}C_{\mathsf{K}} < A_{\mathsf{K}},\ B_{\mathsf{K}} < 0 \right\}.$$

Disconnected feasible region

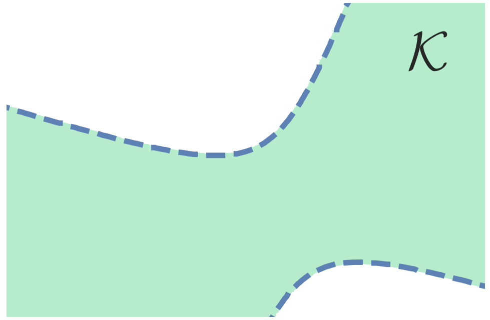


Proof idea: Lifting via Change of Variables

❑ Change of variables in state-space domain: Lyapunov theory

- Connectivity of the static stabilizing state feedback gains

$$\{K \in \mathbb{R}^{m\times n} \mid A - BK \text{ is stable}\}$$
$$\iff \{K \in \mathbb{R}^{m\times n} \mid \exists P \succ 0, P(A-BK)^{\mathsf{T}} + (A-BK)P \prec 0\}$$
$$\iff \{K \in \mathbb{R}^{m\times n} \mid \exists P \succ 0, PA^{\mathsf{T}} - L^{\mathsf{T}}B^{\mathsf{T}} + AP - BL \prec 0, L = KP\}$$
$$\iff \{K = LP^{-1} \in \mathbb{R}^{m\times n} \mid \exists P \succ 0, PA^{\mathsf{T}} - L^{\mathsf{T}}B^{\mathsf{T}} + AP - BL \prec 0\}.$$



Open, connected, possibly nonconvex

- How about the set of stabilizing dynamical controllers

$$\begin{bmatrix} A & BC_{\mathsf{K}} \\ B_{\mathsf{K}}C & A_{\mathsf{K}} \end{bmatrix} \text{ is stable}$$
$$\iff \quad \exists P \succ 0,\ P\begin{bmatrix} A & BC_{\mathsf{K}} \\ B_{\mathsf{K}}C & A_{\mathsf{K}} \end{bmatrix}^{\mathsf{T}} + \begin{bmatrix} A & BC_{\mathsf{K}} \\ B_{\mathsf{K}}C & A_{\mathsf{K}} \end{bmatrix} P \prec 0,$$

Change of variables for output feedback control is highly non-trivial

[Gahinet and Apkarian, 1994]
[Scherer et al., IEEE TAC 1997]

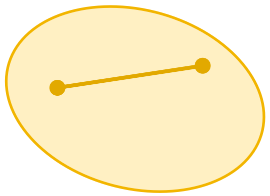
Proof idea: Lifting via Change of Variables

❑ Change of variables in state-space domain: Lyapunov theory

$$\Phi(\mathsf{Z}) = \begin{bmatrix}\Phi_D(\mathsf{Z}) & \Phi_C(\mathsf{Z})\\ \Phi_B(\mathsf{Z}) & \Phi_A(\mathsf{Z})\end{bmatrix} := \begin{bmatrix} I & 0 \\ YB & \Xi\end{bmatrix}^{-1}\begin{bmatrix} G & H \\ F & M - YAX\end{bmatrix}\begin{bmatrix} I & CX \\ 0 & \Pi\end{bmatrix}^{-1}.$$

[Scherer et al., IEEE TAC 1997]
[Gahinet and Apkarian, 1994]

$\mathcal{F} \quad \times \quad \mathrm{GL} \quad \xrightarrow[\text{surjective}]{\Phi} \quad \mathcal{C}_{\mathrm{full}}$



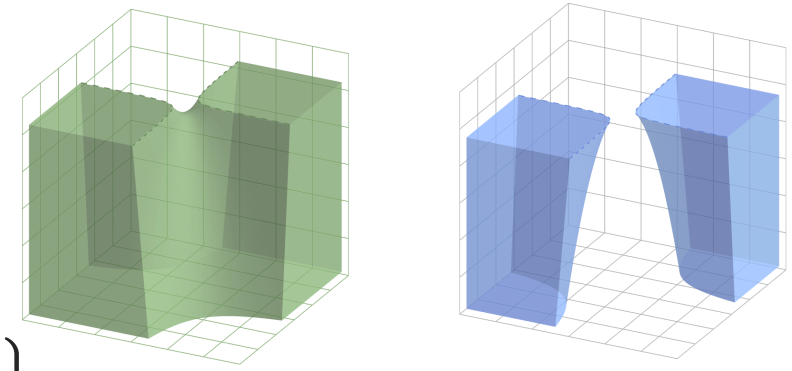
Convex thus connected

General linear group: the set of invertible matrices (similarity transformation)

at most 2 connected components

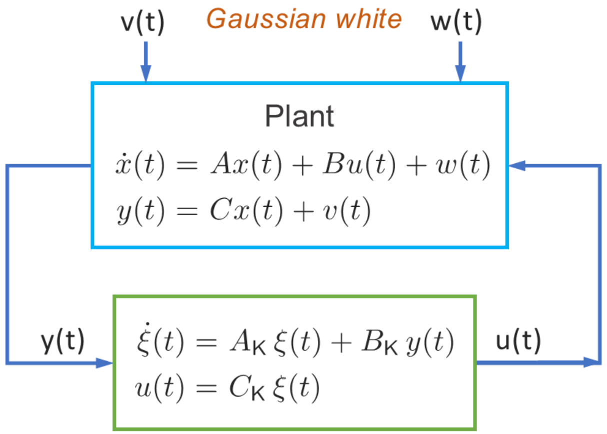
Two connected components

$$\mathrm{GL}_n^+ = \{\Pi \in \mathbb{R}^{n\times n} \mid \det \Pi > 0\},$$
$$\mathrm{GL}_n^- = \{\Pi \in \mathbb{R}^{n\times n} \mid \det \Pi < 0\}.$$



$$\mathcal{F} = \left\{(X,Y,M,H,F) | X, Y \in \mathbb{S}^n, M \in \mathbb{R}^{n\times n}, H \in \mathbb{R}^{n\times p}, F \in \mathbb{R}^{m\times n},\right.$$
$$\left.\begin{bmatrix} X & I \\ I & Y\end{bmatrix} \succ 0, \begin{bmatrix} AX+BF & A \\ M & YA+HC\end{bmatrix} + \begin{bmatrix} AX+BF & A \\ M & YA+HC\end{bmatrix}^{\top} \prec 0\right\}$$

Policy Optimization formulation



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

Non-convex Landscape Analysis

Policy optimization for LQG

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

- **Q1: Connectivity of the feasible region $\mathcal{C}_{\text{full}}$**
  - Is it connected? **No**
  - If not, how many connected components can it have? **Two**
- **Q2: Structure of stationary points of $J(\mathsf{K})$**
  - Are there spurious (strictly suboptimal, saddle) stationary points?
  - How to check if a stationary point is globally optimal?
- **Q3: How to escape high-order saddle points via PGD?**

Outline

- ❑ LQG problem Setup
- ❑ Connectivity of the Set of Stabilizing Controllers
- ❑ **Structure of Stationary Points of the LQG cost**
- ❑ Escaping saddle points via PGD

Structure of Stationary Points

Simple observations

1) $J(\mathsf{K})$ is a real analytic function over its domain (smooth, infinitely differentiable)
2) $J(\mathsf{K})$ has **non-unique** and **non-isolated** global optima

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

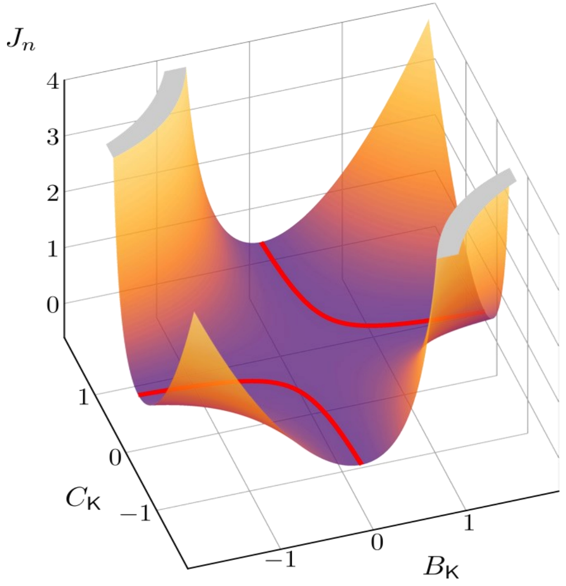
Similarity transformation

$$(A_{\mathsf{K}}, B_{\mathsf{K}}, C_{\mathsf{K}}) \mapsto (TA_{\mathsf{K}}T^{-1}, TB_{\mathsf{K}}, C_{\mathsf{K}}T^{-1})$$

- $J(\mathsf{K})$ is invariant under similarity transformations.
- It has many stationary points, unlike the LQR with a unique stationary point

Policy optimization for LQG

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



Structure of Stationary Points

❑ Gradient computation

Lemma 2: For every $\mathsf{K} = (A_\mathsf{K}, B_\mathsf{K}, C_\mathsf{K}) \in \mathcal{C}_{\text{full}}$, we have

$$\frac{\partial J(\mathsf{K})}{\partial A_\mathsf{K}} = 2\left(Y_{12}^\mathsf{T} X_{12} + Y_{22} X_{22}\right),$$
$$\frac{\partial J(\mathsf{K})}{\partial B_\mathsf{K}} = 2\left(Y_{22} B_\mathsf{K} V + Y_{22} X_{12}^\mathsf{T} C^\mathsf{T} + Y_{12}^\mathsf{T} X_{11} C^\mathsf{T}\right),$$
$$\frac{\partial J(\mathsf{K})}{\partial C_\mathsf{K}} = 2\left(R C_\mathsf{K} X_{22} + B^\mathsf{T} Y_{11} X_{12} + B^\mathsf{T} Y_{12} X_{22}\right),$$

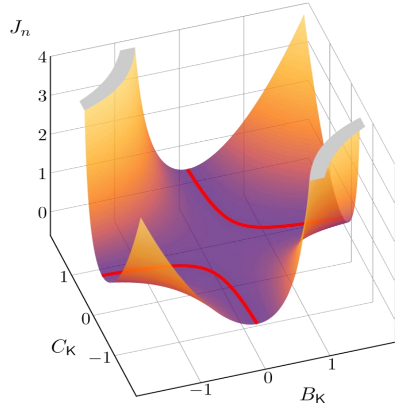
where $X_\mathsf{K} = \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^\mathsf{T} & X_{22} \end{bmatrix}, \qquad Y_\mathsf{K} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{12}^\mathsf{T} & Y_{22} \end{bmatrix}$

are the unique positive semidefinite solutions to two Lyapunov equations.

How does the set of Stationary Points look like?

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

- ❑ Non-unique, non-isolated
- ❑ Local minimum, local maximum, saddle points, or globally minimum?



Structure of Stationary Points

❑ Main Result: existences of strict saddle points

Theorem 4: Consider any open-loop stable plant. The zero controller with any stable A_{K}

$$\mathsf{K} = (A_{\mathsf{K}}, 0, 0) \in \mathcal{C}_{\text{full}}$$

is a stationary point. Furthermore, the corresponding hessian is either indefinite (**strict saddle point**) or equal to zero (**high-order saddle or else**).

Example:

$$\dot{x}(t) = -x(t) + u(t) + w(t)$$
$$y(t) = x(t) + v(t)$$

$$Q = 1, R = 1, V = 1, W = 1$$

Stationary point: $\mathsf{K}^\star = \begin{bmatrix} 0 & 0 \\ 0 & a \end{bmatrix} \in \mathbb{R}^{2\times 2}, \qquad \text{with } a < 0$

- **Cost function:** $J\left(\begin{bmatrix} 0 & C_{\mathsf{K}} \\ B_{\mathsf{K}} & A_{\mathsf{K}} \end{bmatrix}\right) = \dfrac{A_{\mathsf{K}}^2 - A_{\mathsf{K}}(1 + B_{\mathsf{K}}^2 C_{\mathsf{K}}^2) - B_{\mathsf{K}} C_{\mathsf{K}}(1 - 3B_{\mathsf{K}} C_{\mathsf{K}} + B_{\mathsf{K}}^2 C_{\mathsf{K}}^2)}{2(-1 + A_{\mathsf{K}})(A_{\mathsf{K}} + B_{\mathsf{K}} C_{\mathsf{K}})}.$

- **Hessian:** $\begin{bmatrix} \frac{\partial J^2(\mathsf{K})}{\partial A_{\mathsf{K}}^2} & \frac{\partial J^2(\mathsf{K})}{\partial A_{\mathsf{K}} \partial B_{\mathsf{K}}} & \frac{\partial J^2(\mathsf{K})}{\partial A_{\mathsf{K}} \partial C_{\mathsf{K}}} \\ \frac{\partial J^2(\mathsf{K})}{\partial B_{\mathsf{K}} A_{\mathsf{K}}} & \frac{\partial J^2(\mathsf{K})}{\partial B_{\mathsf{K}}^2} & \frac{\partial J^2(\mathsf{K})}{\partial B_{\mathsf{K}} \partial C_{\mathsf{K}}} \\ \frac{\partial J^2(\mathsf{K})}{\partial C_{\mathsf{K}} A_{\mathsf{K}}} & \frac{\partial J^2(\mathsf{K})}{\partial C_{\mathsf{K}} B_{\mathsf{K}}} & \frac{\partial J^2(\mathsf{K})}{\partial \partial C_{\mathsf{K}}^2} \end{bmatrix} \Bigg|_{\mathsf{K}^\star = \begin{bmatrix} 0 & 0 \\ 0 & a \end{bmatrix}} = \dfrac{1}{2(1-a)} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix},$

Indefinite with eigenvalues:

$0 \text{ and } \pm \frac{1}{2(1-a)}$

Structure of Stationary Points

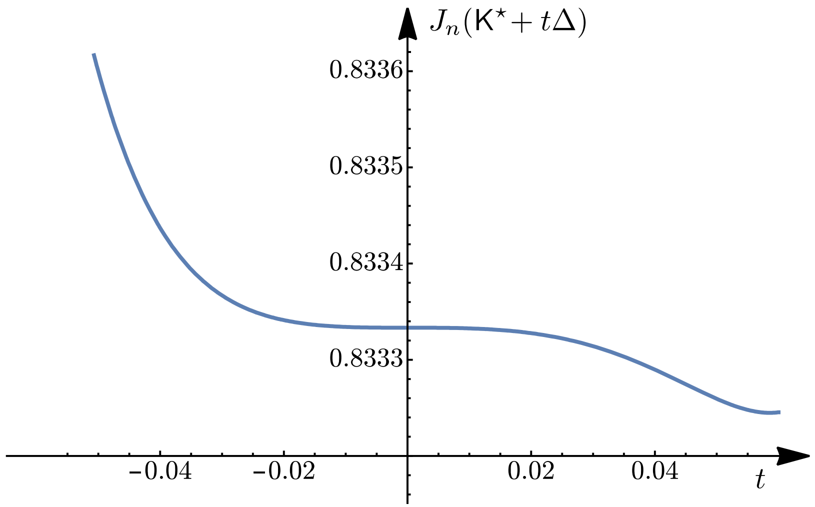
❑ Main Result: existences of strict saddle points

Theorem 4: Consider any open-loop stable plant. The zero controller with any stable A_{K}

$$\mathsf{K} = (A_{\mathsf{K}}, 0, 0) \in \mathcal{C}_{\text{full}}$$

is a stationary point. Furthermore, the corresponding hessian is either indefinite (**strict saddle point**) or equal to zero (**high-order saddle or else**).

Another example with zero Hessian



How does the set of Stationary Points look like?

$$\left\{ \mathsf{K} \in \mathcal{C}_{\text{full}} \;\middle|\; \begin{aligned} \frac{\partial J(\mathsf{K})}{\partial A_{\mathsf{K}}} &= 0, \\ \frac{\partial J(\mathsf{K})}{\partial B_{\mathsf{K}}} &= 0, \\ \frac{\partial J(\mathsf{K})}{\partial C_{\mathsf{K}}} &= 0, \end{aligned} \right\}$$

- ❑ **Non-unique, non-isolated**
- ❑ **Strictly suboptimal points; Strict saddle points**
- ❑ **All bad stationary points correspond to non-minimal controllers**

Structure of Stationary Points

❑ Main Result

Theorem 5: All stationary points corresponding to controllable and observable controllers are globally optimum.

$$\left\{ \mathsf{K} \in \mathcal{C}_{\text{full}} \;\middle|\; \begin{aligned} \frac{\partial J(\mathsf{K})}{\partial A_{\mathsf{K}}} &= 0, \\ \frac{\partial J(\mathsf{K})}{\partial B_{\mathsf{K}}} &= 0, \\ \frac{\partial J(\mathsf{K})}{\partial C_{\mathsf{K}}} &= 0, \end{aligned} \right\}$$

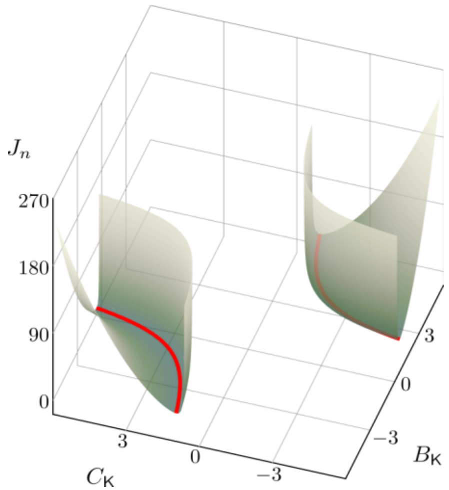
Local Zero Gradient + Structural Information → **Global Optimality Certificate**

Particularly, given a stationary point that is a **minimal** controller

1) It is **globally optimal**, and the set of all global optima forms a manifold with 2 connected components.

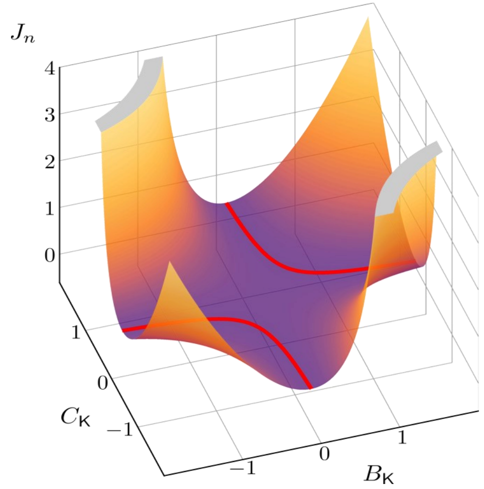
Example: open-loop unstable system

$$\dot{x}(t) = x(t) + u(t) + w(t)$$
$$y(t) = x(t) + v(t)$$



Example: open-loop stable system

$$\dot{x}(t) = -x(t) + u(t) + w(t)$$
$$y(t) = x(t) + v(t)$$



Proof idea

Proof: all minimal stationary points are unique up to a similarity transformation

All minimal stationary points $\mathsf{K} = (A_{\mathsf{K}}, B_{\mathsf{K}}, C_{\mathsf{K}}) \in \mathcal{C}_{\text{full}}$ to the LQG problem are in the form of

$$A_{\mathsf{K}} = T(A - BK - LC)T^{-1}, \qquad B_{\mathsf{K}} = -TL, \qquad C_{\mathsf{K}} = KT^{-1},$$

$$K = R^{-1}B^{\mathsf{T}}S,\ L = PC^{\mathsf{T}}V^{-1},$$

T is an invertible matrix and P, S are the unique positive definite solutions to the Riccati equations

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

Minimal controller ⇒

$$\frac{\partial J_n(\mathsf{K})}{\partial B_{\mathsf{K}}} = 0 \quad \Longrightarrow \quad B_{\mathsf{K}} = -TPC^{\mathsf{T}}V^{-1}$$

$$\frac{\partial J_n(\mathsf{K})}{\partial C_{\mathsf{K}}} = 0 \quad \Longrightarrow \quad C_{\mathsf{K}} = R^{-1}B^{\mathsf{T}}ST^{-1}$$

$$\frac{\partial J(\mathsf{K})}{\partial A_{\mathsf{K}}} = 0 \quad \Longrightarrow \quad A_{\mathsf{K}} = T(A - PC^{\mathsf{T}}V^{-1}C - BR^{-1}B^{\mathsf{T}}S)T^{-1}$$

$$\frac{\partial J(\mathsf{K})}{\partial A_{\mathsf{K}}} = 2\left(Y_{12}^{\mathsf{T}}X_{12} + Y_{22}X_{22}\right),$$

$$\frac{\partial J(\mathsf{K})}{\partial B_{\mathsf{K}}} = 2\left(Y_{22}B_{\mathsf{K}}V + Y_{22}X_{12}^{\mathsf{T}}C^{\mathsf{T}} + Y_{12}^{\mathsf{T}}X_{11}C^{\mathsf{T}}\right),$$

$$\frac{\partial J(\mathsf{K})}{\partial C_{\mathsf{K}}} = 2\left(RC_{\mathsf{K}}X_{22} + B^{\mathsf{T}}Y_{11}X_{12} + B^{\mathsf{T}}Y_{12}X_{22}\right),$$

Special case in Theorem 20.6 of Zhou et al., 1996 and Section II of Hyland, 1984

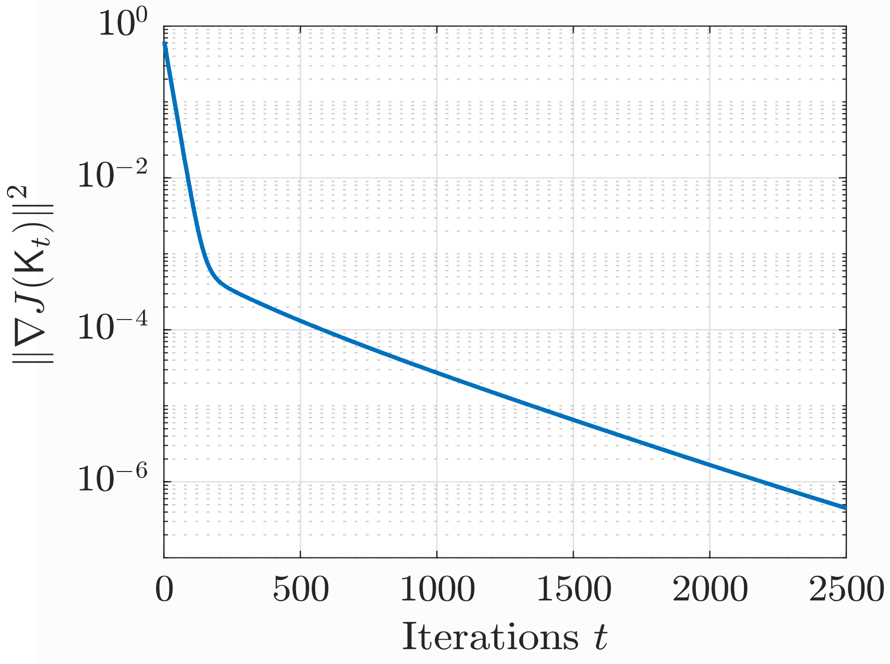
Structure of Stationary Points

❑ Implication

Corollary: Consider gradient descent iterations

$$\mathsf{K}_{t+1} = \mathsf{K}_t - \alpha \nabla J(\mathsf{K}_t)$$

If the iterates converge to a minimal controller, then this minimal controller is a global optima.



More questions:

- ✓ **Escaping saddle points**?
- ✓ Convergence conditions?
- ✓ Convergence speed?
- ✓ Alternative model-free parameterization?

Comparison with LQR

	Policy optimization for LQR $\min_{K} \ J(K)$ s.t. $K \in \mathcal{K}$	**Policy optimization for LQG** $\min_{\mathsf{K}} \ J(\mathsf{K})$ s.t. $\mathsf{K} = (A_{\mathsf{K}}, B_{\mathsf{K}}, C_{\mathsf{K}}) \in \mathcal{C}_{\text{full}}$
Connectivity of feasible region	❖ Always connected	❖ Disconnected, but at most 2 connected comp. ❖ They are almost identical to each other
Stationary points	❖ Unique	❖ Non-unique, non-isolated stationary points ❖ Spurious stationary points (strict saddle, nonminimal controller) ❖ **All mini. stationary points are globally optimal**
Gradient Descent	❖ Gradient dominance ❖ Global fast convergence (like strictly convex)	❖ No gradient dominance ❖ Local convergence/speed (**unknown**) ❖ **Many open questions**
References	Fazel et al., ICML, 2018; Malik et al., 2019; Mohammadi et al., IEEE TAC, 2020; Li et al., 2019; K. Zhang, B. Hu, and T. Başar, 2021; Furieri et al., 2019; Feiran Zhao & Keyou You, 2021, and many others	Zheng*, Tang*, Li. 2021, link (* equal contribution)

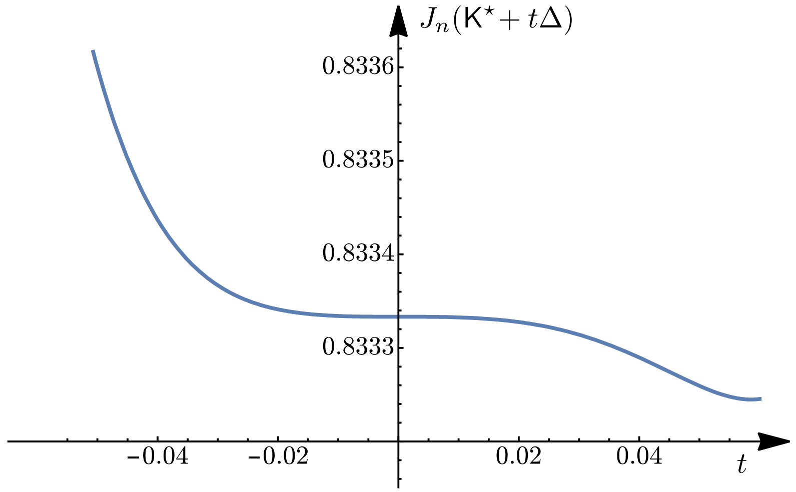
Outline

- ❑ LQG problem Setup
- ❑ Connectivity of the Set of Stabilizing Controllers
- ❑ Structure of Stationary Points of the LQG cost
- ❑ **Escaping saddle points via PGD**

Saddle points

Policy Optimization for LQG

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



High-order saddle point with zero hessian

❑ Local geometry

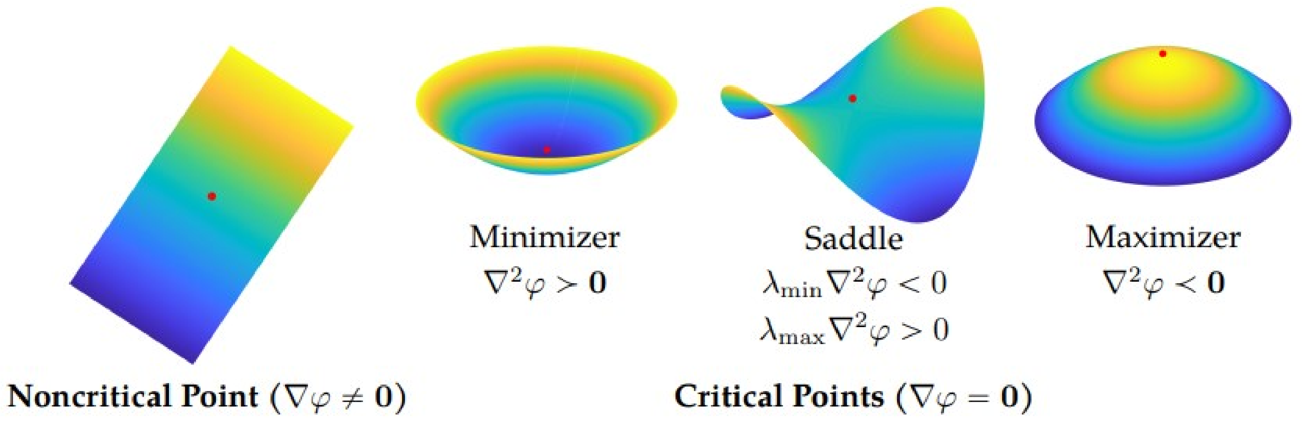


Figure taken from Zhang et al., 2020

- ❑ **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

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

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

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

where we isolate the uncontrollable and unobservable part

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

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

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

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

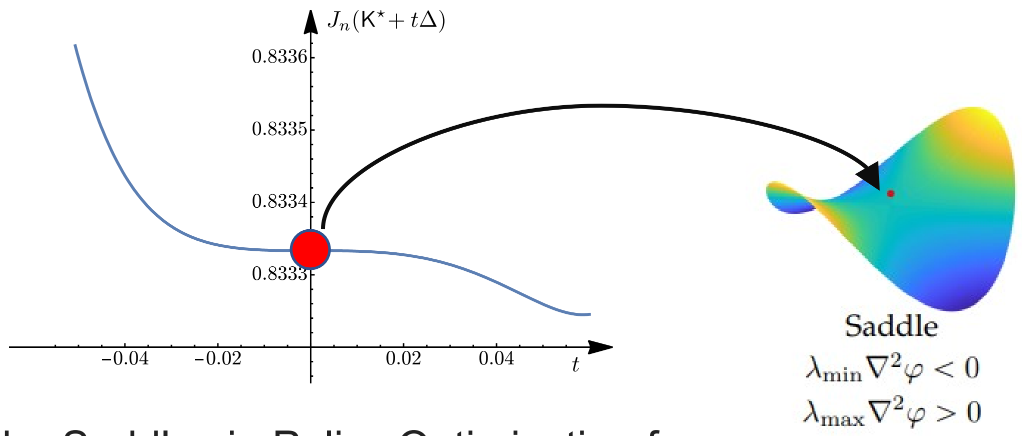
The same form

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

- **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



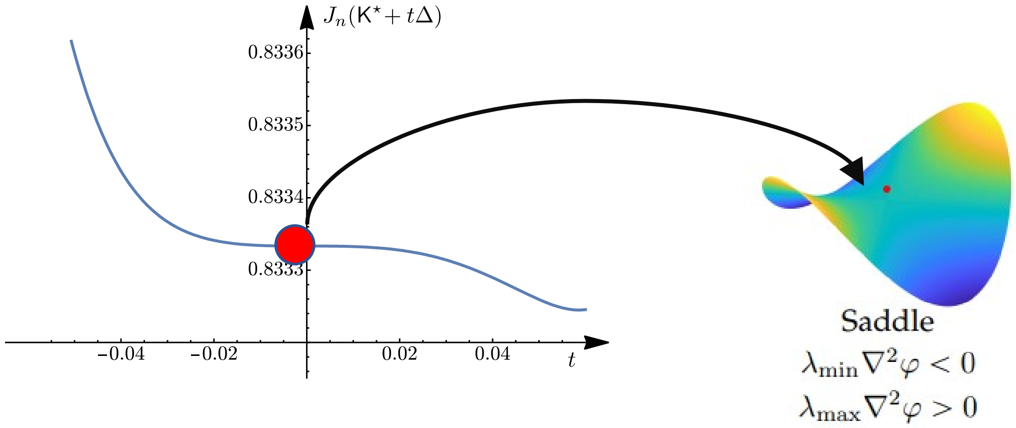
✓ Yang Zheng*, Yue Sun*, Maryam Fazel, and Na Li. "Escaping High-order Saddles in Policy Optimization for Linear Quadratic Gaussian (LQG) Control." arXiv preprint arXiv:2204.00912 (2022). *Equal contribution

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

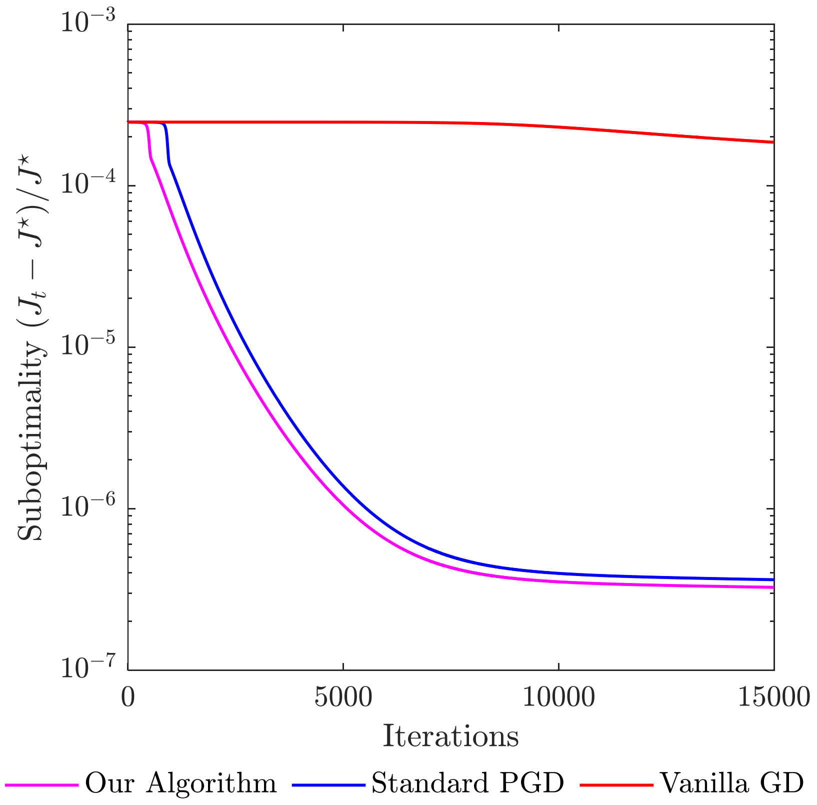
$$A = \begin{bmatrix} -0.5 & 0 \\ 0.5 & -1 \end{bmatrix}, \; B = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \; C = \begin{bmatrix} -\frac{1}{6} & \frac{11}{12} \end{bmatrix},$$

Performance weights

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

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

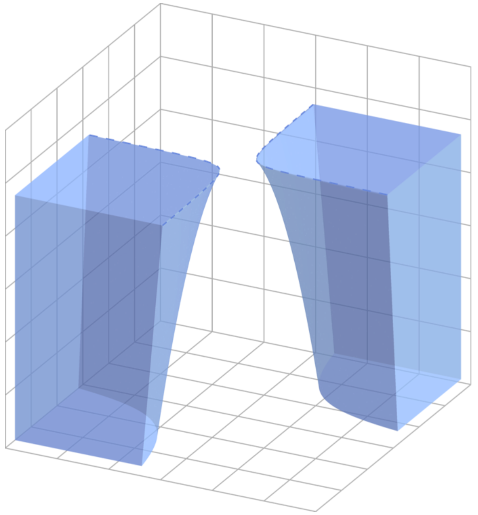
$$A_{\mathsf{K},0} = -0.5 I_2, \; B_{\mathsf{K},0} = \begin{bmatrix} 0 \\ 0.01 \end{bmatrix}, \; C_{\mathsf{K},0} = [0, -0.01]$$

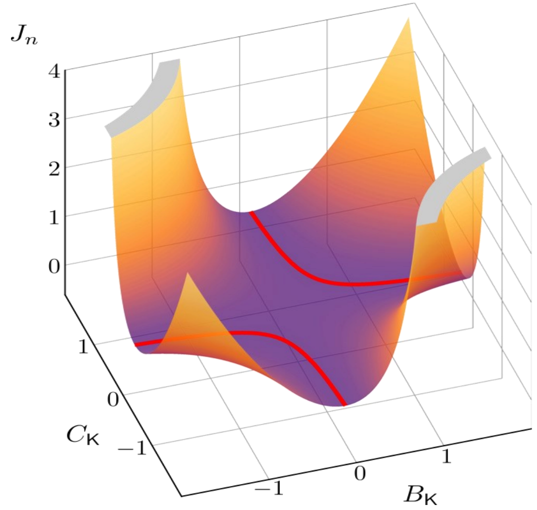


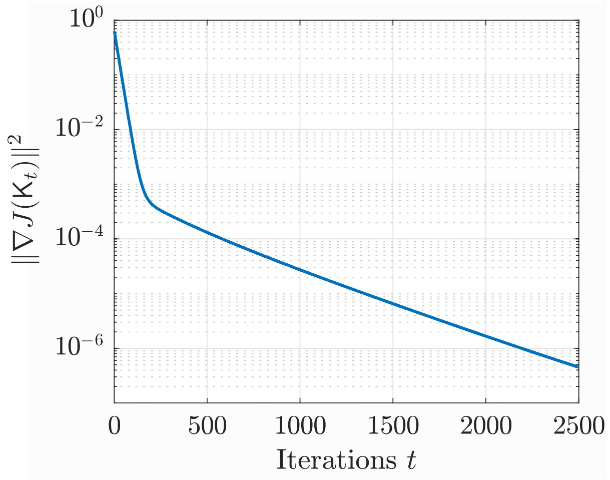
Conclusions

Policy optimization for LQG control

- Much richer and more complicated than LQR
- Disconnected, but at most 2 connected components
- Non-unique, non-isolated stationary points, strict saddle points
- Minimal (controllable and observable) stationary points are globally optimal
- A new perturbed gradient descent algorithm







Ongoing and Future work

- ❑ How to certify the optimality of a non-minimal stationary point
- ❑ Convergence proof of perturbed policy gradient (PGD)
- ❑ More quantitative analysis of PGD algorithms for LQG
- ❑ Alternative model-free parametrization of dynamical controllers (e.g., Makdah & Pasqualetti, 2023; Zhao, Fu &You, 2022.)
  - ✓ Better optimization landscape structures, smaller dimension
- ❑ Nonconvex Landscape of Hinf dynamical output feedback control (Tang & Zheng, 2023 https://arxiv.org/abs/2304.00753;)

Policy Optimization for Linear Quadratic Gaussian (LQG) Control

Thank you for your attention!

Q & A

1. Y. Tang*, Y. Zheng*, and N. Li, “Analysis of the optimization landscape of Linear Quadratic Gaussian (LQG) control,” Mathematical Programming, 2023. Available: https://arxiv.org/abs/2102.04393 *Equal contribution

2. Y. Zheng*, Y. Sun*, M. Fazel, and N. Li. "Escaping High-order Saddles in Policy Optimization for Linear Quadratic Gaussian (LQG) Control." arXiv preprint, 2022 https://arxiv.org/abs/2204.00912. *Equal contribution