Some Statistical Fundamental Limits of Learning for Dynamics and Control

Ph.D. candidate: Xiong Zeng

Advisor: Prof. Necmiye Ozay

Electrical and Computer Engineering
University of Michigan

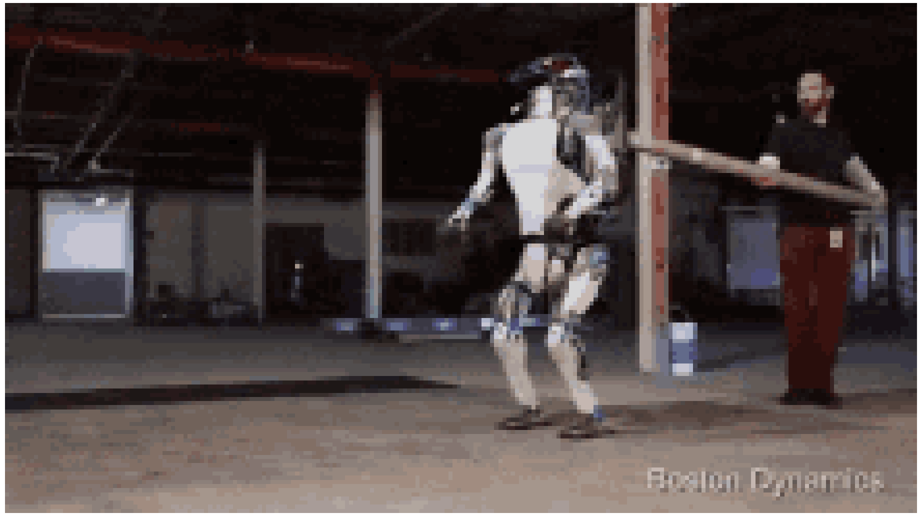
Some Statistical Fundamental Limits of Learning for Dynamics and Control
with Insights for Future Algorithms

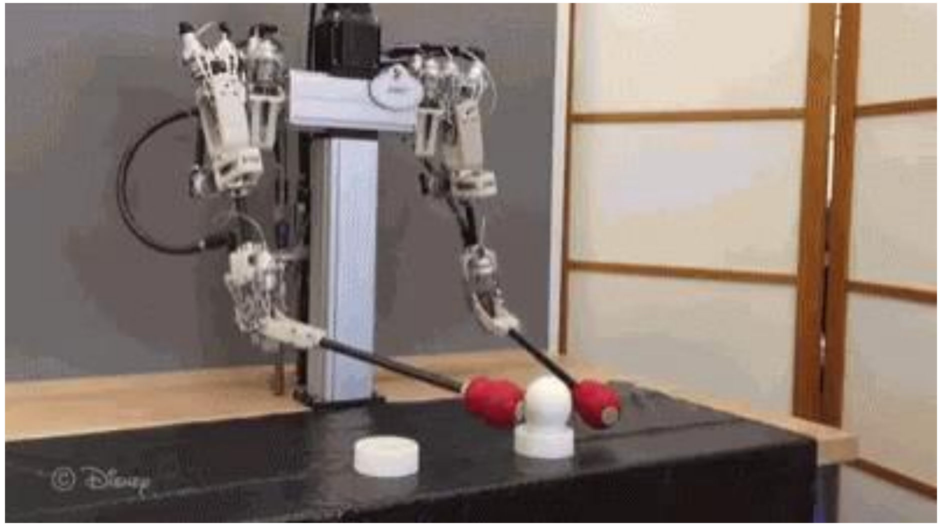
Ph.D. candidate: Xiong Zeng

Advisor: Prof. Necmiye Ozay

Electrical and Computer Engineering
University of Michigan

Unlike ChatGPT, the physical agents are sensitive and dangerous





https://tenor.com/search/robot-fail-gifs

Why?

- Unlike cyberspace with discrete states, the physical world is a continuous state space.
- In continuous state space, infinitesimal error can lead to catastrophic failures in stability and safety.
- Learning for infinitesimal error might be arbitrarily hard.

Therefore, we focus on the statistical fundamental limits of agent learning in continuous state space, like system identification and learning-based control, **etc**.

Summary of Statistical Fundamental Limits

- **Statistical Consistency:** Does the algorithm converge to the ground-truth solution with respect to sample size?

- **Statistical Optimality:** If the algorithm is statistically consistent, does the algorithm achieve the minimax sample complexity lower bound?

- **Statistical Hardness:** If the algorithm is statistically optimal, does the optimal sample complexity increase moderately with the system complexity?

Summary of Our Contributions

- Chapter 2 **for Statistical Consistency ---** Noise Sensitivity of the Semidefinite Programs for Direct Data-Driven LQR (*ACC 2025 and the extension is submitted to TAC*)
- Chapter 3 **for Statistical Optimality ---** System Identification Under Bounded Noise: Optimal Rates Beyond Least Squares *(Submitted to CDC 2025 and L-CSS)*
- Chapter 4 **for Statistical Hardness ---** On the Hardness of Learning to Stabilize Linear Systems *(CDC 2023)*

Chapter 2

Noise Sensitivity of the Semidefinite Programs for Direct Data-Driven LQR

Given a linear time-invariant (LTI) system:

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

where $x_t, w_t \in \mathbb{R}^n, u_t \in \mathbb{R}^m, w_t = 0 \; or \; w_t \sim \mathcal{N}(0, \sigma_w^2 I_n)$, and (A, B) is controllable.

Linear quadratic regulator (LQR) problem:

$$\min_{u_0, u_1, \ldots} \lim_{T \to \infty} \mathbb{E}[\frac{1}{T} \sum_{t=0}^{T} (x_t^T Q x_t + u_t^T R u_t)]$$

s.t. the previous LTI system,

where $Q \succ 0, R \succ 0$. When (A, B) is known, the above optimal solution is $u_t = K_{lqr} x_t$, where

$$K_{lqr} = -\left(R + B^T P B\right)^{-1} B^T P A,$$

and P is the PSD solution of discrete-time algebraic Riccati equation.

When (A, B) is unknown

Offline Data matrices:

$$X_0 = [x_0\ x_1\ \ldots x_{T-1}],$$
$$U_0 = [u_0\ u_1\ \ldots u_{T-1}],$$
$$X_1 = [x_1\ x_2\ \ldots x_T],$$

where $u_t \sim N(0, \sigma_u^2 \mathbf{I}_m)$.

Consider direct data-driven (DDD) control.

Certainty Equivalence (CE) DDD LQR
(De Persis & Tesi, TAC 2019):

$$\min_{X,Y}\ \operatorname{trace}(QX_0Y) + \operatorname{trace}(X)$$
$$\text{s.t.} \begin{bmatrix} X_0Y - I_n & X_1Y \\ Y^TX_1^T & X_0Y \end{bmatrix} \succcurlyeq 0$$
$$\begin{bmatrix} X & \sqrt{R}U_0Y \\ \left(\sqrt{R}U_0Y\right)^T & X_0Y \end{bmatrix} \succcurlyeq 0.$$

Let Y_{ce}^* be an optimal solution, an estimate of K_{lqr}

$$K_{ce}(T) := -U_0Y_{ce}^*(X_0Y_{ce}^*)^{-1}.$$

Theorem 1 (De Persis & Tesi, 2019). Let $\operatorname{rank}\left(\begin{bmatrix} X_0 \\ U_0 \end{bmatrix}\right) = m + n$. When $w_t = 0$ for all t,

$$K_{ce}(T) = K_{lqr}.$$

(CE DDD LQR is perfect for noiseless case)

Theorem 2 (our work). Assume $\sigma_w^2 > 0$. When $T \geq (m+n)(n+1)+n$,

$$P(K_{ce}(T) = 0_{m \times n}) = 1.$$

(CE DDD LQR is trivial for almost all noise)

Key observation: The following equalities hold for all Y_{ce}^* when $\sigma_w^2 > 0$

$$\begin{cases} U_0 Y_{ce}^* = 0_{m \times n} \\ X_0 Y_{ce}^* = I_n \\ X_1 Y_{ce}^* = 0_{n \times n} \end{cases},$$

for any $T \geq (m+n)(n+1)+n$. Recall $K_{ce} = -U_0 Y_{ce}^* (X_0 Y_{ce}^*)^{-1}$.

$$\min_{X,Y} \ \mathrm{trace}(QX_0Y) + \mathrm{trace}(X)$$

$$\text{s.t.} \begin{bmatrix} X_0Y - I_n & X_1Y \\ Y^TX_1^T & X_0Y \end{bmatrix} \succcurlyeq 0$$

$$\begin{bmatrix} X & \sqrt{R}U_0Y \\ \left(\sqrt{R}U_0Y\right)^T & X_0Y \end{bmatrix} \succcurlyeq 0.$$

Explanation for key observations:

- $X_0Y - I_n \succcurlyeq 0$ and $X \succcurlyeq 0 \implies \mathrm{trace}(QX_0Y) + \mathrm{trace}(X) \geq \mathrm{trace}(Q)$.

- All Y *with* $\begin{cases} U_0Y = \mathbf{0}_{m\times n} \\ X_0Y = \mathbf{I}_n \\ X_1Y = \mathbf{0}_{n\times n} \end{cases}$ are feasible solutions, for which $\mathrm{trace}(QX_0Y) + \mathrm{trace}(X) = \mathrm{trace}(Q)$.

Robustness-Promoting (RP) DDD LQR (De Persis & Tesi, Automatica 2021):

$$\min_{X,Y,S} \operatorname{trace}(QX_0Y) + \operatorname{trace}(X) + \operatorname{trace}(S)$$

$$\text{s.t.} \begin{bmatrix} X_0Y - I_n & X_1Y \\ Y^TX_1^T & X_0Y \end{bmatrix} \succcurlyeq 0$$

$$\begin{bmatrix} X & \sqrt{R}U_0Y \\ \left(\sqrt{R}U_0Y\right)^T & X_0Y \end{bmatrix} \succcurlyeq 0$$

$$\begin{bmatrix} S & Y \\ Y^T & X_0Y \end{bmatrix} \succcurlyeq 0.$$

Let Y_{rp}^* denote its optimal solution, $K_{rp}(T) \coloneqq -U_0Y_{rp}^*\left(X_0Y_{rp}^*\right)^{-1}$.

Theorem 3 (our work). Assume $w_t = 0$,

$$\lim_{T\to\infty} P(K_{rp}(T) = K_{lqr}) = 1.$$

(RP DDD LQR is statistically consistent for noiseless case)

Theorem 4 (our work). Assume $\sigma_w^2 > 0$,

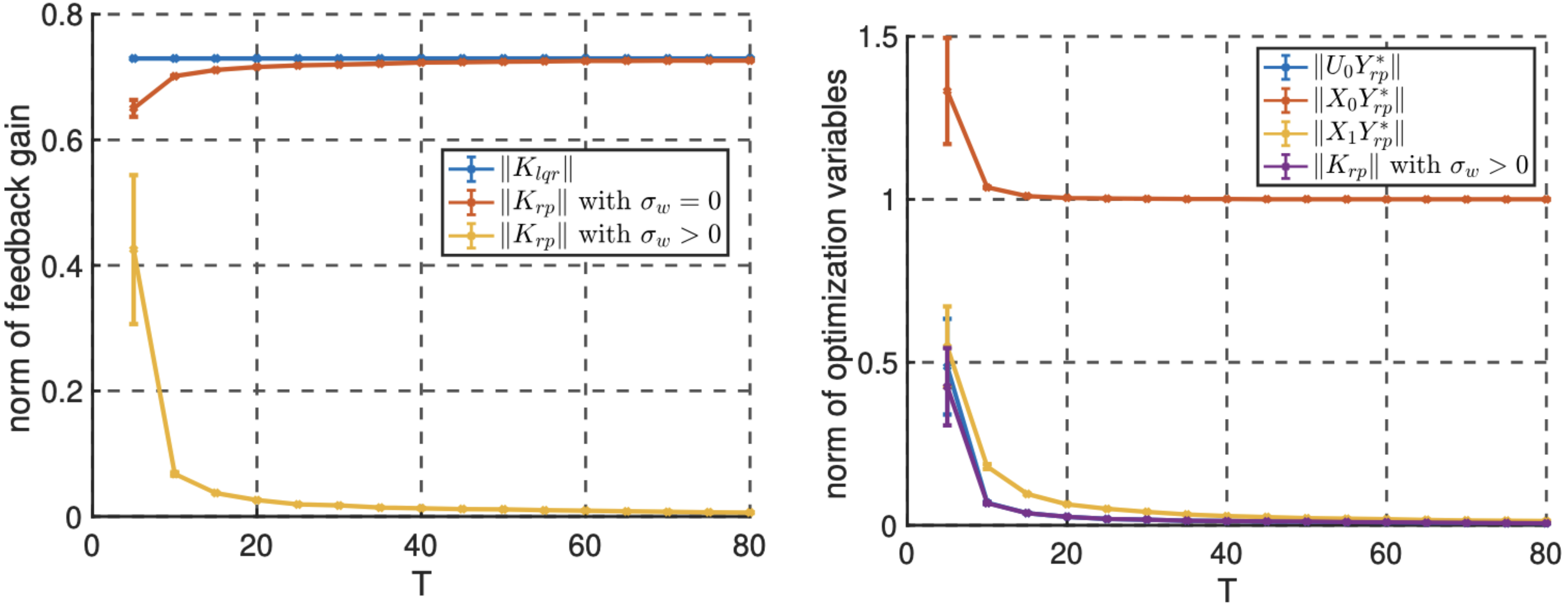
$$\lim_{T\to\infty} \mathrm{P}(K_{rp}(T) = \mathbf{0}_{\boldsymbol{m}\times\boldsymbol{n}}) = 1\,.$$

(RP DDD LQR is not statistically consistent for noisy case)

Key Observation: The following equalities always hold when $\sigma_w^2 > 0$:

$$\begin{cases} \lim_{T\to\infty} \mathrm{P}(U_0 Y_{rp}^* = \mathbf{0}_{\boldsymbol{m}\times\boldsymbol{n}}) = 1 \\ \lim_{T\to\infty} \mathrm{P}\left(X_0 Y_{rp}^* = \boldsymbol{I}_{\boldsymbol{n}}\right) = 1 \\ \lim_{T\to\infty} \mathrm{P}(X_1 Y_{rp}^* = \mathbf{0}_{\boldsymbol{n}\times\boldsymbol{n}}) = 1 \end{cases},$$

and recall $\mathrm{K}_{\mathrm{rp}}(T) = -\mathrm{U}_0 \mathrm{Y}_{\mathrm{rp}}^* \left(\mathrm{X}_0 \mathrm{Y}_{\mathrm{rp}}^*\right)^{-1}$.



Experiments for RP DDD LQR: Consider an order-2 single-input unstable system, for which $K_{lqr} = [-0.7112\ -0.2046]$. $\sigma_w^2 = 1$ and $\sigma_u^2 = 1$.

	CE DDD LQR	RP DDD LQR
noiseless ($\mathbf{w}_t = 0$)	$\mathbb{P}(\mathbf{K}_{ce} = \mathbf{K}_{\text{lqr}}) = 1$ (De Persis & Tesi, 2019)	$\mathbf{K}_{rp}(T) \xrightarrow{p} \mathbf{K}_{\text{lqr}}$ (Theorem 3)
noisy ($\sigma_w > 0$)	$\mathbb{P}(\mathbf{K}_{ce} = \mathbf{0}_{m\times n}) = 1$ (Theorem 2)	$\mathbf{K}_{rp}(T) \xrightarrow{p} \mathbf{0}_{m\times n}$ (Theorem 4)

Summary and Future Work

1. Some SDPs for DDD LQR are sensitive to noise.
2. Check the statistical fundamental limits when designing new DDD control algorithms.
3. The fundamental limits for DDD Robust Control by matrix S-lemma (Waarde et al. 2020, Waarde et al. 2023) are unclear now.

Chapter 3

System Identification Under Bounded Noise: Optimal Rates Beyond Least Squares

Consider an unknown LTI system:

$$x_{t+1} = Ax_t + w_t,$$

where $x_t, w_t \in \mathbb{R}^n$. Assume

- $||w_t||_\infty < \bar{w}$ and w_t are i.i.d. for all t,
- $\rho(A) < 1$,
- and for any $\epsilon \in [0, \bar{w}]$, there exists $C > 0$, such that $\forall j \in [n]$,

$$\max\left(P\left(w_t^{(j)} < -\bar{w} + \epsilon\right), P\left(w_t^{(j)} > \bar{w} - \epsilon\right)\right) < C\epsilon.$$

Theorem 5 (our work). Given a single trajectory $\{x_t\}_{t\in[T]}$. $\mathcal{F}_T$ denotes the σ-algebra generated by $\{x_t\}_{t\in[T]}$ and $\hat{A}_T$ denotes a $\mathcal{F}_T$ -measurable estimator. Then, $\forall\delta \in (0,1)$ and small $\epsilon > 0$,

$$\sup_{\hat{A}_T} \inf_{A\in R^{n\times n}} P(||\hat{A}_T - A||_2 < \epsilon) \geq 1 - \delta \quad \text{only if} \quad T > \frac{1}{4\bar{w}C\epsilon}\left(1 - \frac{2\delta}{n}\right).$$

(The best estimator can achieve $\Omega(\frac{1}{\epsilon})$)

The ordinary least squares (OLS) for scalar case:

$$\hat{a}_T^{OLS} = \mathrm{argmin}_{\mathrm{a}} \sum_{t=1}^{T-1} \left\| x_{t+1} - a x_t \right\|^2$$

Theorem 6 (our work). Assume $|a| < 1$. Then, $\forall \delta \in (0,1)$ and small $\epsilon > 0$,

$$P(||\hat{a}_T^{OLS} - a||_2 < \epsilon) \geq 1 - \delta \quad \text{only if} \quad T > \Omega(\frac{1}{\epsilon^2}).$$

(OLS only achieves $\Omega(\frac{1}{\epsilon^2})$)

The set membership estimator (SME) based on $\{x_t\}_{t\in[T]}$

$$\mathcal{S}_{\mathrm{T}} = \left\{ A \in \mathbb{R}^{n\times n} : \left\| x_{t+1} - A x_t \right\|_\infty \leq \bar{w}, \forall t \in [T-1] \right\}.$$

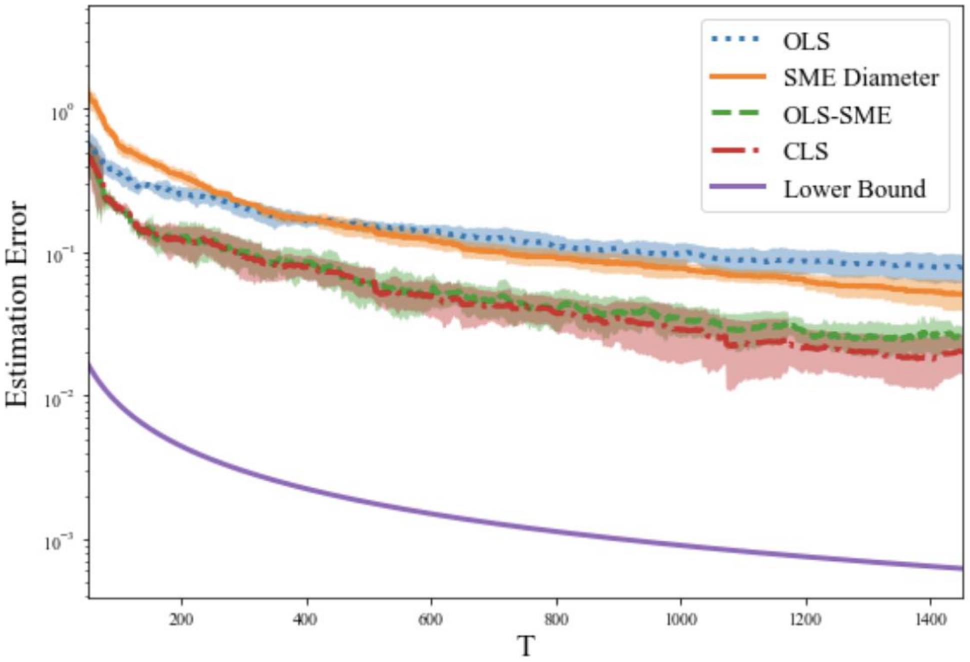
Theorem 7 (Li & Yu et al., ICML 2024). $\forall \delta \in (0,1)$ and small $\epsilon > 0$, with or without knowing $\bar{w}$,

$$\forall \hat{A}_T \in \mathcal{S}_{\mathrm{T}}, \; P(||\hat{A}_T - A||_2 < \epsilon) \geq 1 - \delta \quad \text{if} \quad T > \Omega(\frac{1}{\epsilon}).$$

(SME achieves the optimal $\Omega(\frac{1}{\epsilon})$)

Summary:

		Minimax Lower Bound	Lower Bound for OLS
Regression	Gaussian	$\Omega(1/\sqrt{T})$ (Wainwright, 2019)	$\Omega(1/\sqrt{T})$ (Mourtada, 2022)
	Bounded	$\Omega(1/T)$ (Yi & Neykov, 2024)	$\Omega(1/\sqrt{T})$ (Rudelson & Vershynin, 2008)
LTI Sys Id	Gaussian	$\Omega(1/\sqrt{T})$ (Jedra & Proutiere, 2019)	$\Omega(1/\sqrt{T})$ (Tu et al., 2024)
	Bounded	$\Omega(1/T)$ (Theorem 5)	$\Omega(1/\sqrt{T})$ (Theorem 6)



OLS: $\min_{A} \sum_{t=1}^{T-1} \left|\left| x_{t+1} - Ax_t \right|\right|^2$

SME: $\mathcal{S}_{\mathrm{T}} = \left\{ A \in \mathbb{R}^{n\times n} : \left|\left| x_{t+1} - Ax_t \right|\right|_\infty \leq \bar{w}, \forall t \in [T] \right\}$

OLS-SME: $\min_{A \in \mathcal{S}_{\mathrm{T}}} \left|\left| A - \hat{A}_T^{OLS} \right|\right|$

CLS: $\min_{A \in \mathcal{S}_{\mathrm{T}}} \sum_{t=1}^{T-1} \left|\left| x_{t+1} - Ax_t \right|\right|^2$

OLS achieves $O(\frac{1}{\sqrt{T}})$ ☹

SME, OLS-SME, and CLS achieves $O(\frac{1}{T})$☺

Chapter 4

On the Hardness of Learning to Stabilize Linear Systems

Given an unknown LTI system

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

where $x_t, w_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, and $w_t \sim \mathcal{N}(0, \sigma_w^2 I_n)$. Assume (A, B) is controllable the energy bounded input $\mathbb{E}[\|u_t\|^2] \leq \sigma_u^2$.

Consider a learning to stabilize algorithm π that

- interacts with the above system for T units of time, and
- outputs a **linear static state feedback** controller $\widehat{K}_T$ at time T.

We want $\rho\left(A + B\widehat{K}_T\right) < 1$.

Theorem 7 (our work). Then $\forall \delta \in [0,0.5]$, we have

$$\sup_{\pi} \inf_{(A,B)} P((A + B\,\widehat{K}_T) \text{ is stable}) \geq 1 - \delta,$$

only if

$$\mathrm{T} \geq \Omega(exp(n)).$$

($exp(n)$ hardness of learning to stabilize with easy SysId) 🙁

Comparison with the previous work

Tsiamis et al., COLT 2022	Hard to identify, then hard to learn to stabilize
Our Work	Hard to distinguish and hard to co-stabilize, then hard to learn to stabilize

Summary

1. System identification is easy, but learning to stabilize by linear static state feedback is still exponentially hard with the state dimension

2. We can try other state feedback, like **online switching** (**Proposed Work 1**) or **historical state feedback** (**Proposed Work 2**)

Chapter 5

- **Proposed Work 1 ---** Minimax Optimal Online LQR with A Finite Set of Models or Controllers
- **Proposed Work 2 ---** Co-Stabilization and Learning to Stabilize by Historical State Feedback
- **Proposed Work 3 ---** On the Maximal Stochastic Koopman Invariant Subspace

Proposed Work 1 --- Minimax Optimal Online LQR with A Finite Set of Models or Controllers

Given an unknown LTI system:

$$\mathrm{x}_{t+1} = \mathrm{A}\mathrm{x}_t + \mathrm{B}\mathrm{u}_t + \mathrm{w}_t,$$

where $\mathrm{w}_t \sim \mathcal{N}(0, \sigma_w^2 I_n)$ and (A, B) is controllable.

No offline input-state data, only have some candidate models or controllers:

$$(A, B) \in \mathcal{M}_N := \left\{(A^{(i)}, B^{(i)})\right\}_{i\in[N]}, \quad \text{(Models)}$$

$$\mathcal{P}_N := \left\{K^{(i)}, P^{(i)}\right\}_{i\in[N]}, \quad \text{(Controllers and Quadratic Lyapunov)}$$

where $K^{(i)}$ is from $LQR(A^{(i)}, B^{(i)}, Q, R)$ and $P^{(i)}$ satisfies

$$(A^{(i)}+B^{(i)}K^{(i)})P^{(i)}(A^{(i)}+B^{(i)}K^{(i)})^T + \sigma_w^2 P^{(i)} - P^{(i)} + \Sigma^{(i)} = 0,$$

with $\Sigma^{(i)} \succ 0$.

Online LQR for (A, B): Design $\{u_t\}_{t\in[T]}$ to minimize

$$Regret(T) = \sum_{t=0}^{T} (x_t^T Q x_t + u_t^T R u_t) - T\ Cost_{LQR}.$$

Conjectures:

	Minimax Optimal Regret
$\mathbf{A}$ and $\mathbf{B}$ are unknown	$O(\sqrt{T})$ (Simchowitz & Foster, 2020)
$\mathbf{A}$ is known but $\mathbf{B}$ is unknown	$O(\log(T))$ (Jedra & Proutiere, 2022)
$\mathbf{A}$ is unknown but $\mathbf{B}$ is known	$O(\log(T))$ (Jedra & Proutiere, 2022)
$\mathcal{M}_N$ is known $\mathcal{P}_N$ is known	Is it $O(\log(T))$?
$\mathcal{M}_N$ is unknown $\mathcal{P}_N$ is known	Is it $O(\log(T))$?

Further extensions:

Stabilize unknown **nonlinear** systems with a pool of candidate **Control Lyapunov Functions**

Proposed Work 2 --- Co-Stabilization and Learning to Stabilize by Historical State Feedback

Consider the previous LTI systems:

- For linear static state feedback $\mathrm{u_t} = K\mathrm{x_t}$, check if $\rho(\mathrm{A} + \mathrm{B}K) < 1$;
- For horizon-2 historical state feedback $\mathrm{u_t} = K_0\mathrm{x_t} + K_1\mathrm{x_{t-1}}$, consider $\widetilde{K} = [K_0\ K_1]$ and

$$\begin{bmatrix} \mathbf{x}_{t+1} \\ \mathbf{x}_t \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{A} & 0 \\ \mathbf{I}_n & 0 \end{bmatrix}}_{\tilde{\mathbf{A}}} \begin{bmatrix} \mathbf{x}_t \\ \mathbf{x}_{t-1} \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{B} \\ 0 \end{bmatrix}}_{\tilde{\mathbf{B}}} \mathbf{u}_t + \begin{bmatrix} \mathbf{w}_t \\ 0 \end{bmatrix}$$

check if $\rho\left(\tilde{A} + \tilde{B}\widetilde{K}\right) < 1$.

To do:

We observe $\widetilde{K}$ has a larger co-stabilization gap than K fo the linear scalar case, and we try to observe this result o the multivariate the case empirically and theoretically, and extend them to learning-to-stabilize.

Proposed Work 3 --- On the Maximal Stochastic Koopman Invariant Subspace

Consider a stochastic dynamical system

$$x_{t+1} = \mathrm{f}(\mathrm{x_t}) + \mathrm{w_t},$$

where $\mathcal{X} \in \mathbb{R}^n$ is the state space and $\mathrm{f}: \mathcal{X} \to \mathcal{X}$ is a measurable function and the noise $\mathrm{w_t} \sim \mathcal{N}(0, \sigma_w^2 I_n)$. Let $\mathcal{G}$ denote a linear space of observable functions from $\mathcal{X}$ to $\mathbb{R}$, with some regularity conditions.

Key observation:
Extended Dynamic Mode Decomposition (EDMD) with deterministic Koopman invariant subspace with the deterministic Koopman Operator ($\mathcal{K}_d(g)(x_t) = g(x_{t+1})$, for $\mathrm{g} \in \mathcal{G}$) is statistically inconsistent for stochastic systems, even for linear case.

Stochastic Koopman Operator: $\mathcal{K}_s(g)(x) \coloneqq \mathbb{E}[g(x_{t+1}) | x_t = x]$.

(Maximal) Stochastic Koopman Invariant Subspace: An observable function subspace $\mathcal{S} \subseteq \mathcal{G}$ is **stochastic Koopman-invariant** if for every $g \in \mathcal{S}, \mathcal{K}_s(g) \in \mathcal{S}$. Furthermore, $\mathcal{S}$ is **maximal** stochastic Koopman-invariant in $\mathcal{L} \subseteq \mathcal{G}$ if $\mathcal{S}$ contains **every** Koopman-invariant subspace in $\mathcal{L}$.

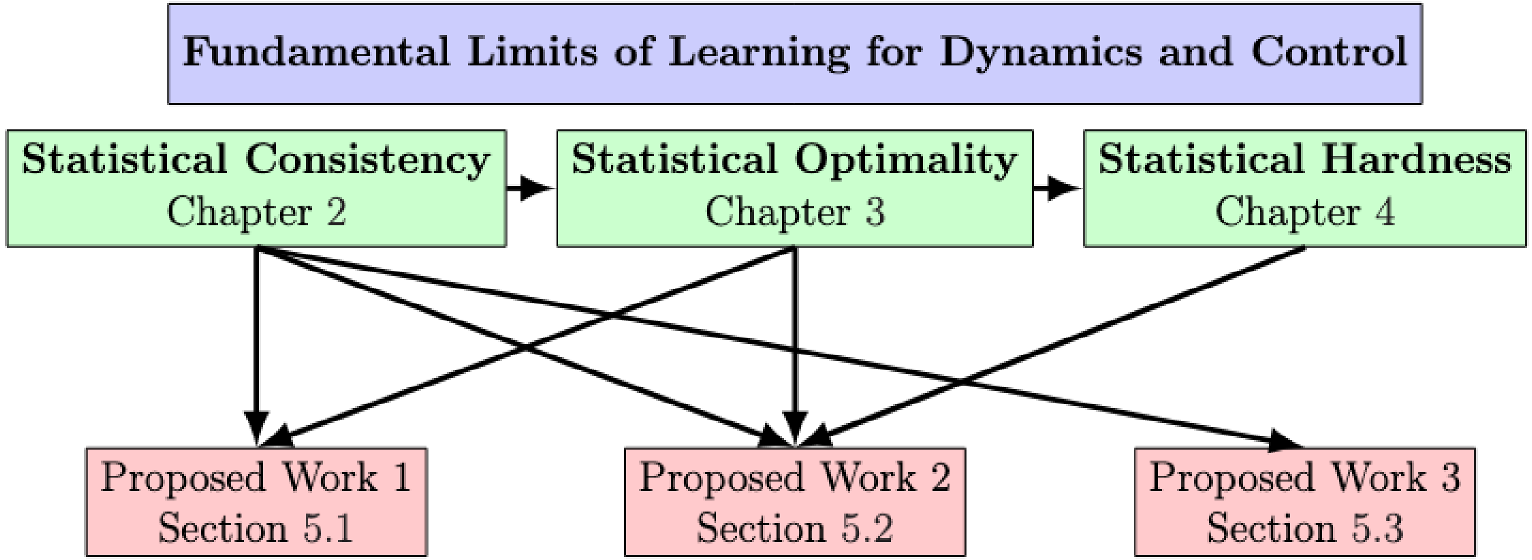
Question:

Given a dictionary of candidate observable functions

$$D(x) = \{g_1(x), g_2(x), \dots, g_d(x)\}.$$

Could we design a provable algorithm to find a basis for the maximal stochastic Koopman-invariant subspace in $\mathrm{span}\{D(x)\}$ for the known or unknown stochastic systems?

Summary and Timeline



References

[3] De Persis & Tesi, “Formulas for data-driven control: Stabilization, optimality, and robustness,” TAC, 2019.
[4] De Persis & Tesi, “Low-complexity learning of linear quadratic regulators from noisy data,” Automatica, 2021.
[5] X. Zeng, L. Bako, and N. Ozay, “Noise sensitivity of direct data-driven linear quadratic regulator by semidefinite programming,” ACC 2025 and long version is submitted to TAC.
[6]. Waarde et al., “From noisy data to feedback controllers: Nonconservative design via a matrix s-lemma,” IEEE Transactions on Automatic Control, 2020.
[7]. Waarde et al., “Quadratic matrix inequalities with applications to data-based control,” SIAM Journal on Control and Optimization, 2023.
[8]. A. Tsiamis, I. M. Ziemann, M. Morari, N. Matni, and G. J. Pappas, “Learning to control linear systems can be hard,” COLT 2022.
[9]. Y. Li, J. Yu, L. Conger, T. Kargin, and A. Wierman, “Learning the uncertainty sets of linear control systems via set membership: A non-asymptotic analysis,” ICML 2024.
[10]. Zeng, X., Liu, Z., Du, Z., Ozay, N., & Sznaier, M. On the hardness of learning to Stabilize linear systems. CDC 2023.
[11]. Zeng, X., Yu, J., & Ozay, N. (2025). System Identification Under Bounded Noise: Optimal Rates Beyond Least Squares. ArXiv 2025.