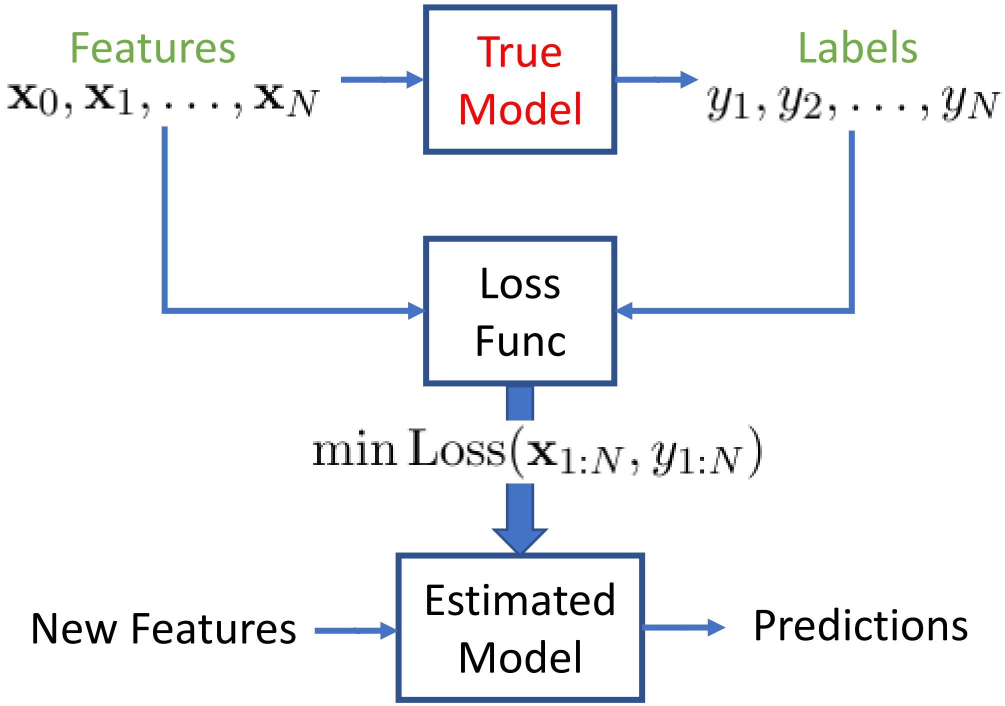
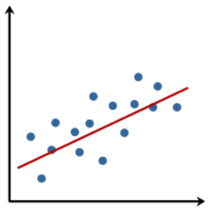
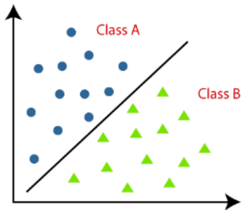
Certainty Equivalent Control for Markov Jump Systems

Zhe Du, Davoud Ataee Tarzanagh, Laura Balzano, Necmiye Ozay
University of Michigan
Yahya Sattar, Samet Oymak
University of California, Riverside

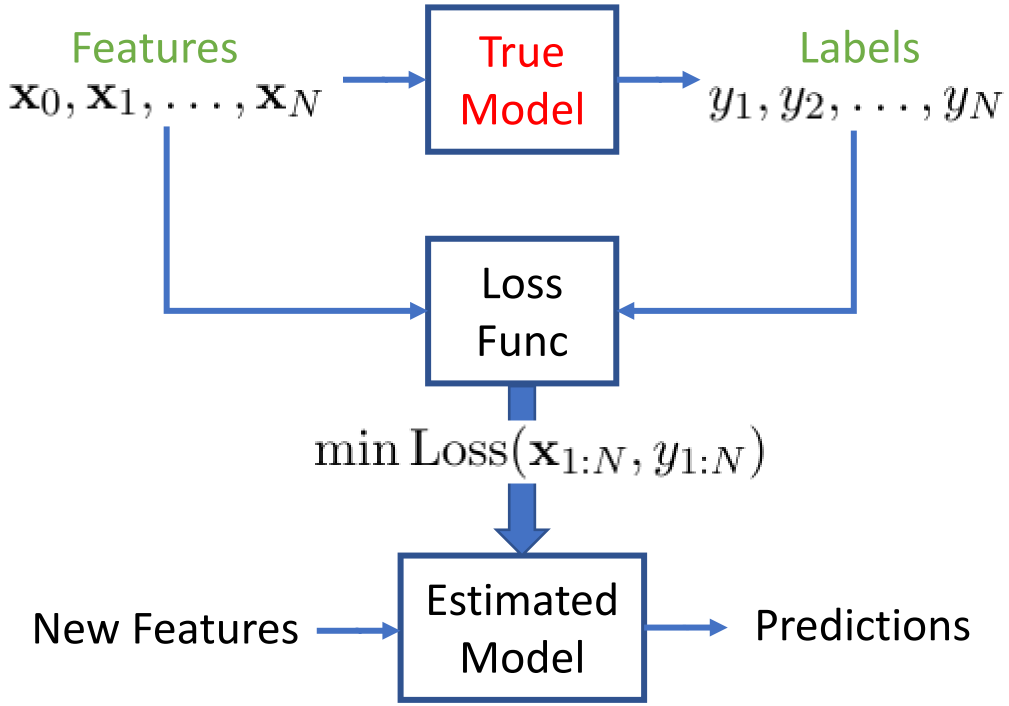
ML vs Control

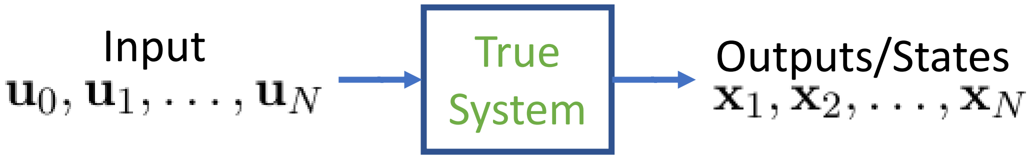


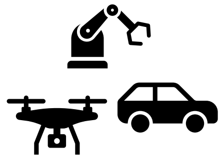


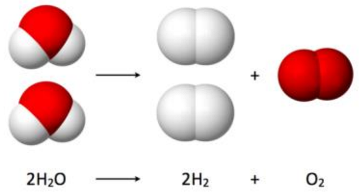


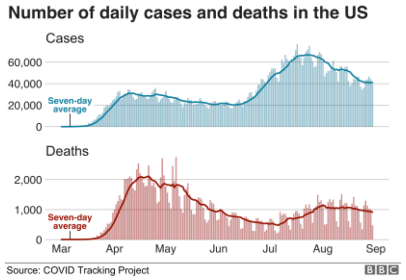
ML vs Control

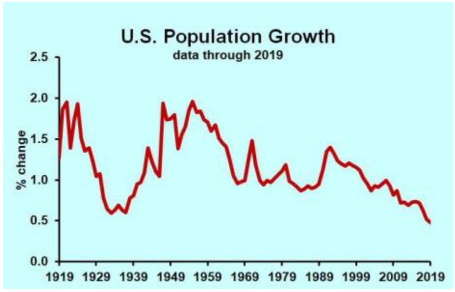


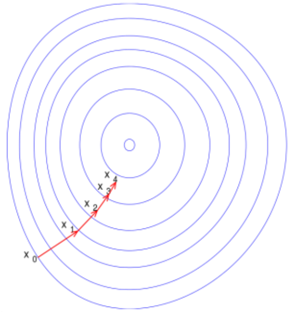




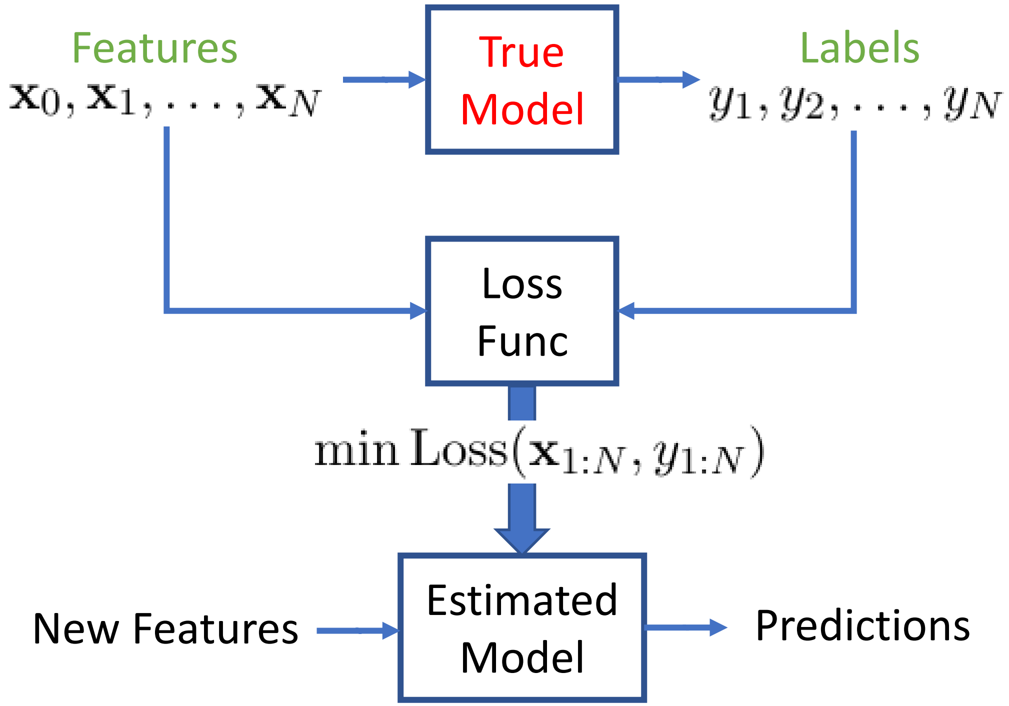


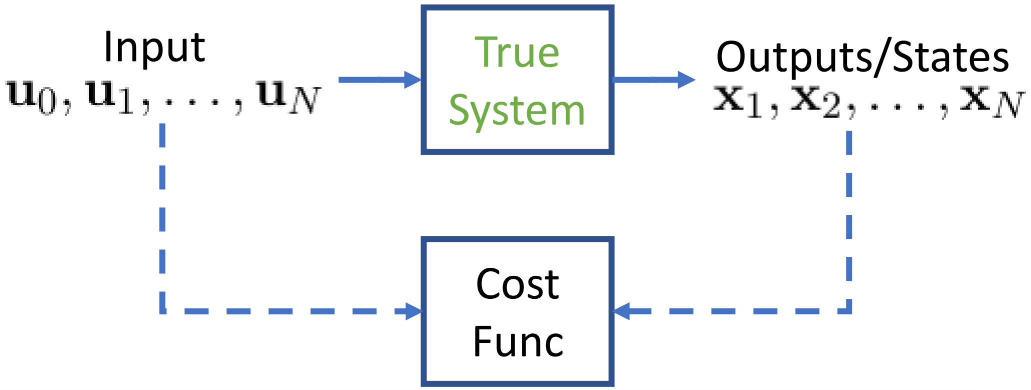




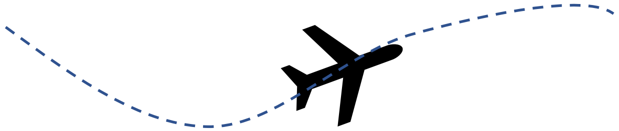


ML vs Control



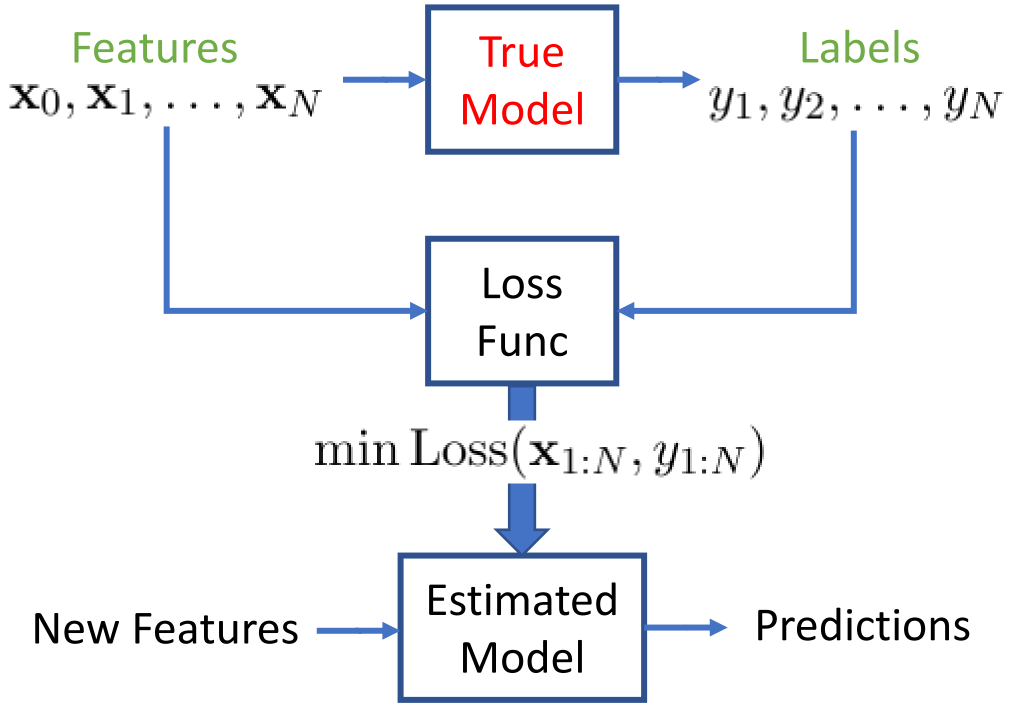


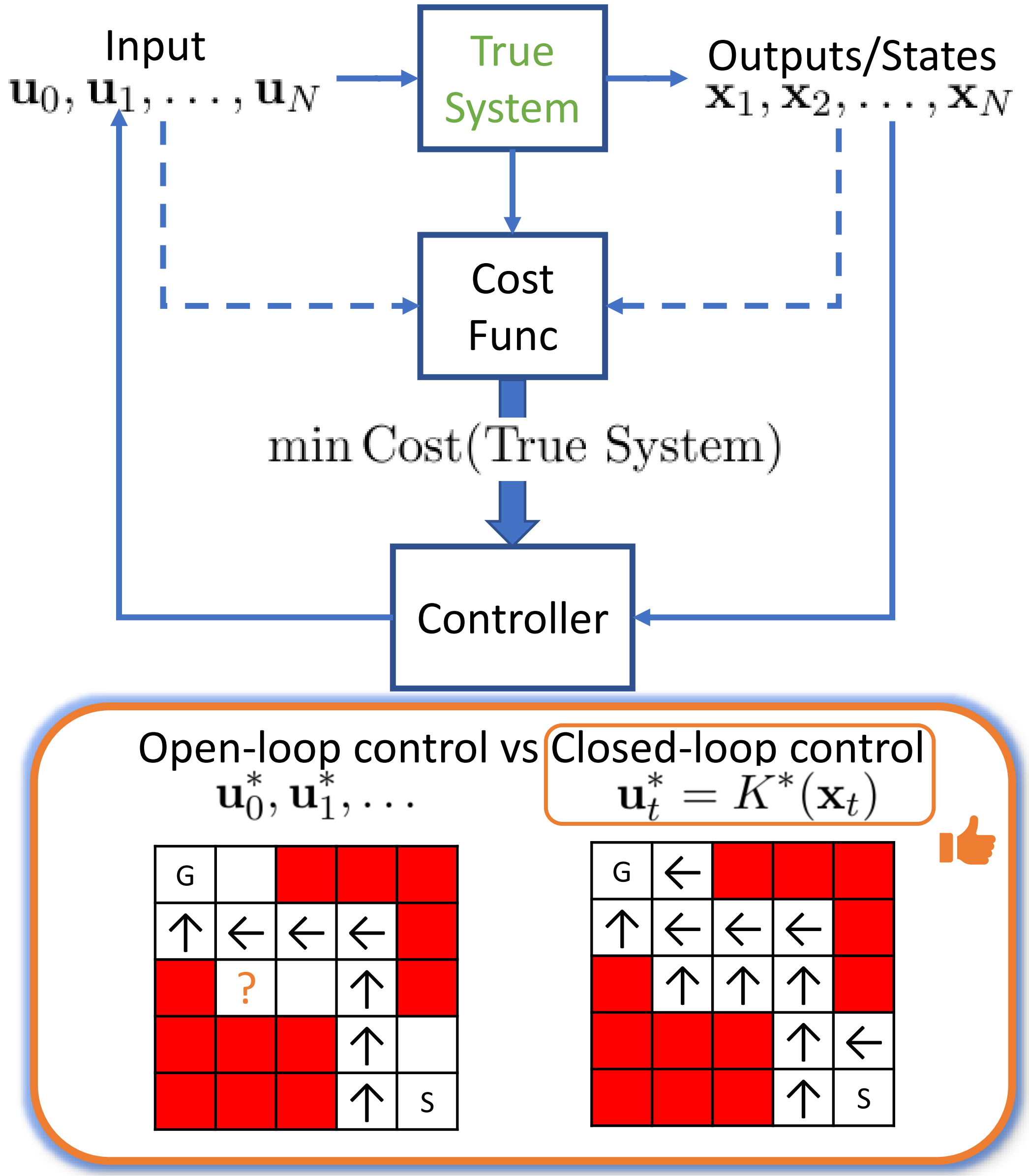
- Deviation from desired states/trajectories



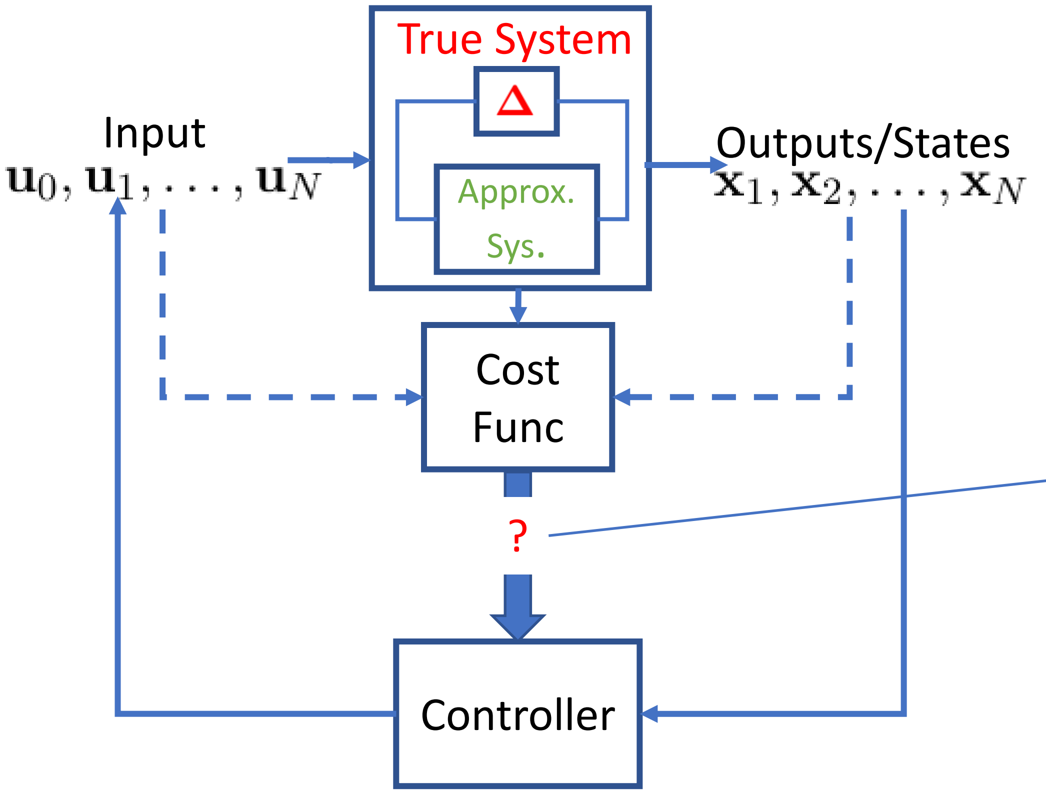
- Control efforts
- Safety guarantee
- Time limitation
- …

ML vs Control



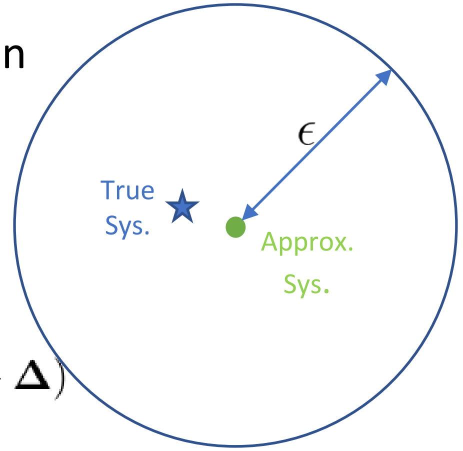


Robust Control vs Certainty Equivalent Control



Sources of $\boldsymbol{\Delta}$

- Estimation error in system identification
- Manufacturing variances
- Environmental changes



Robust Control

$$\min \max_{\|\boldsymbol{\Delta}\| \leq \epsilon} \text{Cost}(\text{Approx. Sys.} + \boldsymbol{\Delta})$$

Pro: robustness, Con: computation, suboptimality

Certainty Equivalent Control

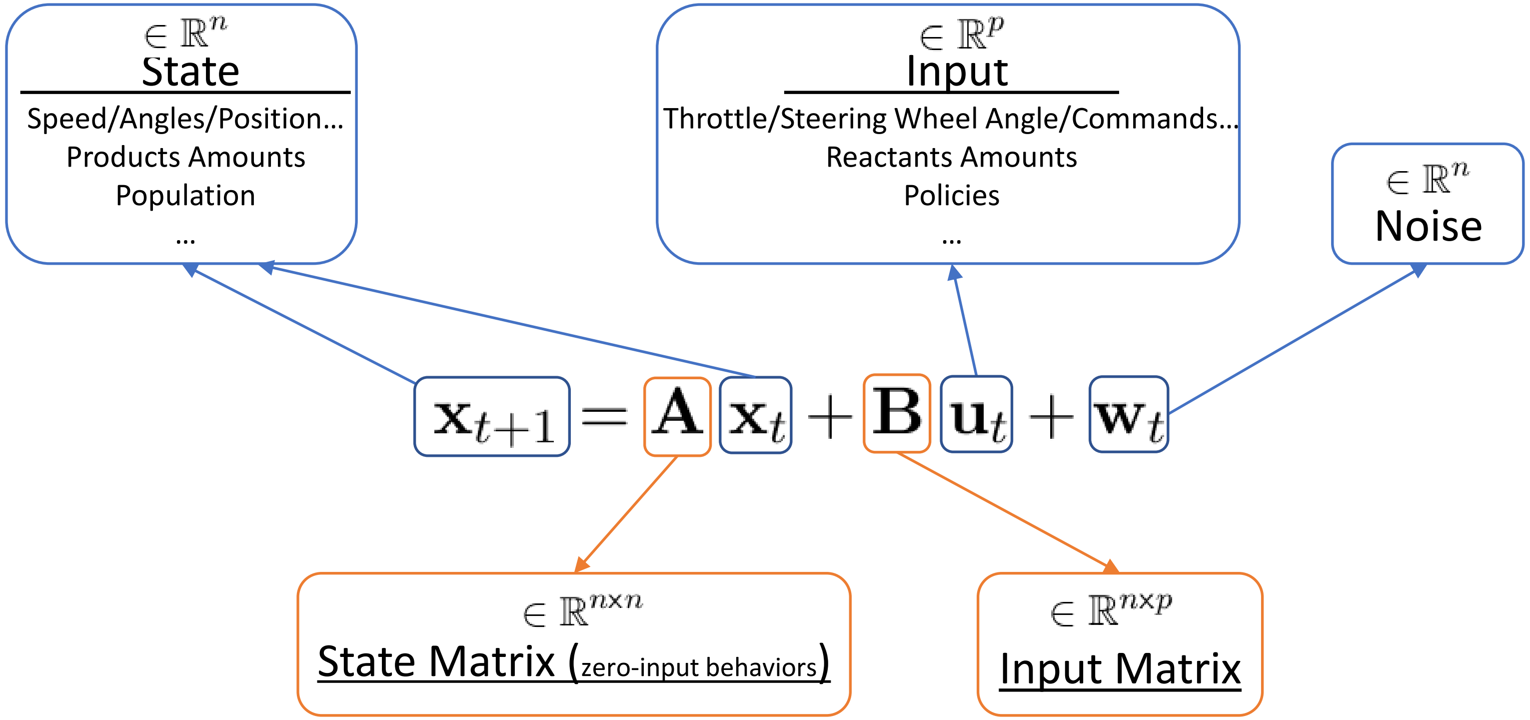
$$\min \text{Cost}(\text{Approx. Sys.})$$

Pro: computation, Con: robustness

Questions:

1. When does certainty equivalent control work?
2. Suboptimality compared w. the optimal controller?
3. Certainty equivalent control vs Robust control?

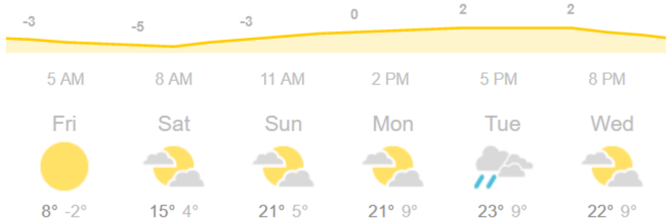
Linear Time-invariant Systems (LTI)

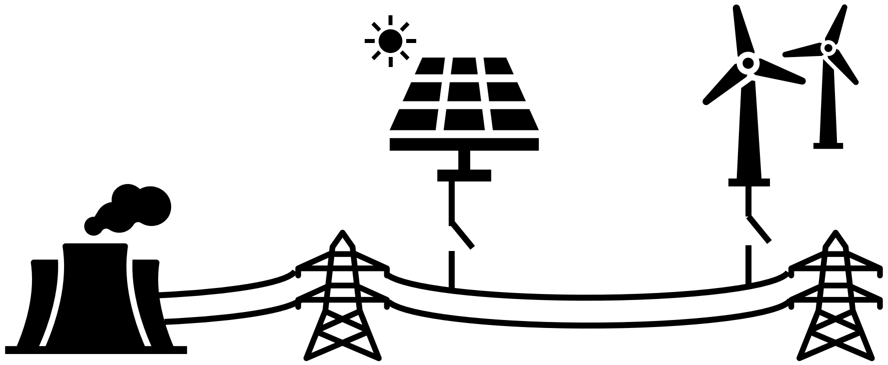


Linear Time-invariant Systems (LTI)

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

Is time-invariant dynamics accurate?





Markov Jump Systems (MJS)

$$\mathbf{x}_{t+1} = \mathbf{A}_{\omega(t)}\mathbf{x}_t + \mathbf{B}_{\omega(t)}\mathbf{u}_t + \mathbf{w}_t$$

$$\{\mathbf{A}_1, \mathbf{B}_1\}, \{\mathbf{A}_2, \mathbf{B}_2\}, \ldots \{\mathbf{A}_s, \mathbf{B}_s\}$$

s modes

Mode switching sequence

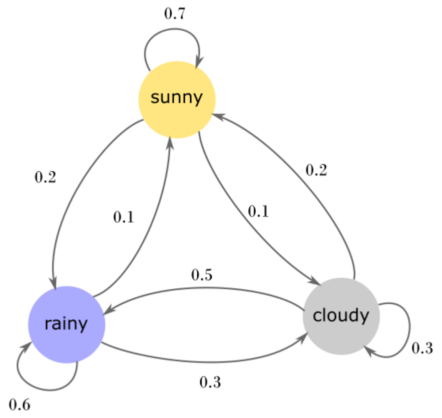
$$\omega(0), \omega(1), \cdots \sim MarkovChain(\mathbf{T})$$

$$\mathrm{P}(\underbrace{\omega(t+1) = j}_{\text{mode } j \text{ is active at time } t+1} \mid \underbrace{\omega(t) = i}_{\text{mode } i \text{ is active at time } t}) = \mathbf{T}_{ij}$$

mode j is active at time t+1

mode i is active at time t

- What is and why Markov chain?



- Given current mode, future modes do not depend on past modes
- Less memory, easier prediction and analysis
- Concepts and control strategies for LTI can be channeled to MJS via Markov chain

Markov Jump Systems (MJS)

$$\mathbf{x}_{t+1} = \mathbf{A}_{\omega(t)}\mathbf{x}_t + \mathbf{B}_{\omega(t)}\mathbf{u}_t + \mathbf{w}_t$$

$$\{\mathbf{A}_1, \mathbf{B}_1\}, \{\mathbf{A}_2, \mathbf{B}_2\}, \ldots \{\mathbf{A}_s, \mathbf{B}_s\}$$

s modes

Mode switching sequence

$$\omega(0), \omega(1), \cdots \sim MarkovChain(\mathbf{T})$$

$$\mathrm{P}(\underbrace{\omega(t+1) = j}_{\text{mode } j \text{ is active at time } t+1} \mid \underbrace{\omega(t) = i}_{\text{mode } i \text{ is active at time } t}) = \mathbf{T}_{ij}$$

mode j is active at time $t+1$ — mode i is active at time t

- Relation to SGD
  - Data: $\{\mathbf{x}_{1:s}, y_{1:s}\}, y_i = \boldsymbol{\theta}_*^\top \mathbf{x}_i$
  - Loss: $\sum_{i=1}^s \frac{1}{2}(\boldsymbol{\theta}^\top \mathbf{x}_i - y_i)^2$
  - At time t,
    1. Pick $\omega(t) \in [s]$ w.p. $\frac{1}{s}$

Linear Quadratic Regulator (LQR) Problem

- Given cost matrices $\mathbf{Q}_{1:s}, \mathbf{R}_{1:s} \succ 0$, define quadratic cost

$$J_t := \mathbb{E}[\mathbf{x}_t^\intercal \mathbf{Q}_{\omega(t)} \mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega(t)} \mathbf{u}_t]$$

Deviation from desired states (annotating $\mathbf{x}_t^\intercal \mathbf{Q}_{\omega(t)} \mathbf{x}_t$)

Control efforts (annotating $\mathbf{u}_t^\intercal \mathbf{R}_{\omega(t)} \mathbf{u}_t$)

and $J_{0:N} := \sum_{t=0}^{N} J_t$

Linear Quadratic Regulator (LQR) Problem

- Given cost matrices $\mathbf{Q}_{1:s}, \mathbf{R}_{1:s} \succ 0$, define quadratic cost

$$J_t := \mathbb{E}[\mathbf{x}_t^\intercal \mathbf{Q}_{\omega(t)} \mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega(t)} \mathbf{u}_t]$$

  and $J_{0:N} := \sum_{t=0}^{N} J_t$

- Finite-time horizon LQR $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$

$$\begin{aligned} \min_{\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_N} \quad & J_{0:N} \\ \text{s.t. } \quad & \mathbf{x}_{t+1} = \mathbf{A}_{\omega(t)} \mathbf{x}_t + \mathbf{B}_{\omega(t)} \mathbf{u}_t + \mathbf{w}_t \\ & \mathrm{P}(\omega(t+1) = j \mid \omega(t) = i) = \mathbf{T}_{ij} \end{aligned}$$

Linear Quadratic Regulator (LQR) Problem

- Given cost matrices $\mathbf{Q}_{1:s}, \mathbf{R}_{1:s} \succ 0$, define quadratic cost

$$J_t := \mathbb{E}[\mathbf{x}_t^\intercal \mathbf{Q}_{\omega(t)} \mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega(t)} \mathbf{u}_t]$$

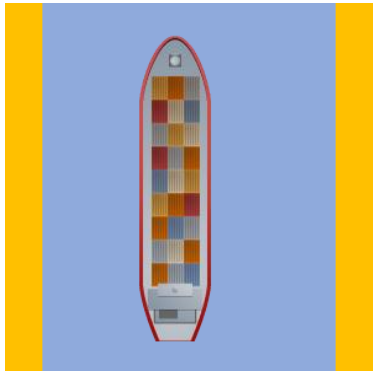
  and $J_{0:N} := \sum_{t=0}^{N} J_t$

- Finite-time horizon LQR $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$

$$\min_{\mathbf{u}_0, \mathbf{u}_1, \ldots, \mathbf{u}_N} \quad J_{0:N}$$
$$\text{s.t. } \mathbf{x}_{t+1} = \mathbf{A}_{\omega(t)} \mathbf{x}_t + \mathbf{B}_{\omega(t)} \mathbf{u}_t + \mathbf{w}_t$$
$$\mathrm{P}(\omega(t+1) = j \mid \omega(t) = i) = \mathbf{T}_{ij}$$

- Infinite-time horizon LQR $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$

$$\min_{\mathbf{u}_0, \mathbf{u}_1, \ldots} \quad \lim_{N \to \infty} \frac{1}{N} J_{0:N}$$
$$\text{s.t. } \mathbf{x}_{t+1} = \mathbf{A}_{\omega(t)} \mathbf{x}_t + \mathbf{B}_{\omega(t)} \mathbf{u}_t + \mathbf{w}_t$$
$$\mathrm{P}(\omega(t+1) = j \mid \omega(t) = i) = \mathbf{T}_{ij}$$



Optimal Solution for MJS-LQR

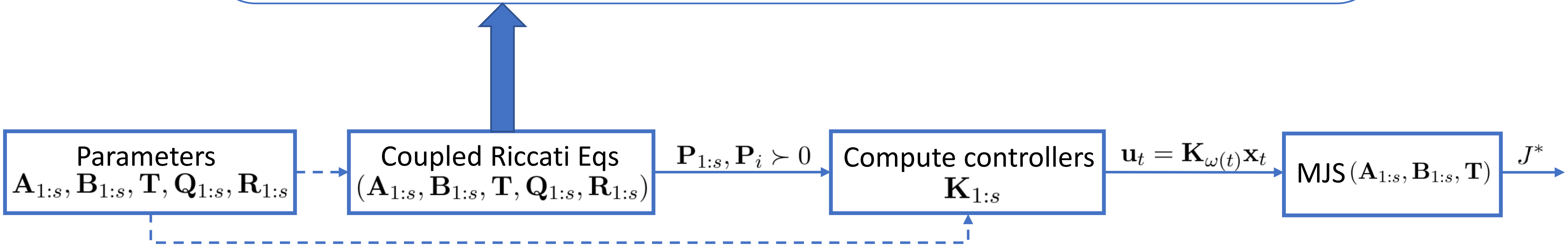
MJS-LQR $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$

$$\min_{\mathbf{u}_0, \mathbf{u}_1, \dots} \lim_{N \to \infty} \frac{1}{T} J_{0:N}$$

$$\text{s.t. } \mathbf{x}_{t+1} = \mathbf{A}_{\omega(t)} \mathbf{x}_t + \mathbf{B}_{\omega(t)} \mathbf{u}_t + \mathbf{w}_t$$

$$\mathrm{P}(\omega(t+1) = j \mid \omega(t) = i) = \mathbf{T}_{ij}$$

$$\begin{cases} \mathbf{P}_1 = \mathbf{A}_1^\intercal \left[\sum_{j=1}^s \mathbf{T}_{1j}\mathbf{P}_j\right] \mathbf{A}_1 + \mathbf{Q}_1 - \mathbf{A}_1^\intercal \left[\sum_{j=1}^s \mathbf{T}_{1j}\mathbf{P}_j\right] \mathbf{B}_1 \left(\mathbf{R}_1 + \mathbf{B}_1^\star \left[\sum_{j=1}^s \mathbf{T}_{1j}\mathbf{P}_j\right] \mathbf{B}_1\right)^{-1} \mathbf{B}_1^\intercal \left[\sum_{j=1}^s \mathbf{T}_{1j}\mathbf{P}_j\right] \mathbf{A}_1 \\ \vdots \\ \mathbf{P}_i = \mathbf{A}_i^\intercal \left[\sum_{j=1}^s \mathbf{T}_{ij}\mathbf{P}_j\right] \mathbf{A}_i + \mathbf{Q}_i - \mathbf{A}_i^\intercal \left[\sum_{j=1}^s \mathbf{T}_{ij}\mathbf{P}_j\right] \mathbf{B}_i \left(\mathbf{R}_i + \mathbf{B}_i^\star \left[\sum_{j=1}^s \mathbf{T}_{ij}\mathbf{P}_j\right] \mathbf{B}_i\right)^{-1} \mathbf{B}_i^\intercal \left[\sum_{j=1}^s \mathbf{T}_{ij}\mathbf{P}_j\right] \mathbf{A}_i \\ \vdots \\ \mathbf{P}_s = \mathbf{A}_s^\intercal \left[\sum_{j=1}^s \mathbf{T}_{sj}\mathbf{P}_j\right] \mathbf{A}_s + \mathbf{Q}_s - \mathbf{A}_s^\intercal \left[\sum_{j=1}^s \mathbf{T}_{sj}\mathbf{P}_j\right] \mathbf{B}_s \left(\mathbf{R}_s + \mathbf{B}_s^\star \left[\sum_{j=1}^s \mathbf{T}_{sj}\mathbf{P}_j\right] \mathbf{B}_s\right)^{-1} \mathbf{B}_s^\intercal \left[\sum_{j=1}^s \mathbf{T}_{sj}\mathbf{P}_j\right] \mathbf{A}_s \end{cases}$$



Optimal Solution for MJS-LQR

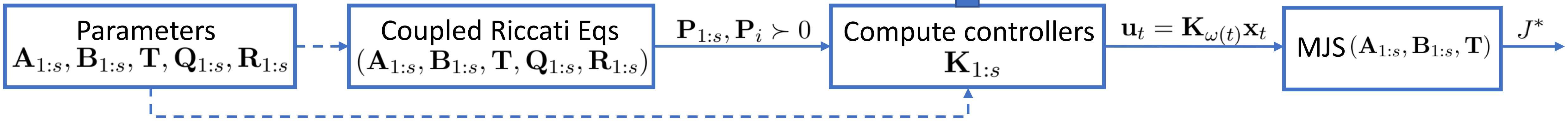
MJS-LQR $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$

$$\min_{\mathbf{u}_0, \mathbf{u}_1, \ldots} \lim_{N \to \infty} \frac{1}{T} J_{0:N}$$

$$\text{s.t. } \mathbf{x}_{t+1} = \mathbf{A}_{\omega(t)} \mathbf{x}_t + \mathbf{B}_{\omega(t)} \mathbf{u}_t + \mathbf{w}_t$$

$$\mathrm{P}(\omega(t+1) = j \mid \omega(t) = i) = \mathbf{T}_{ij}$$

$$\begin{cases} \mathbf{K}_1 = -\left(\mathbf{R}_1 + \mathbf{B}_1^\intercal \left[\sum_{j=1}^s \mathbf{T}_{1j} \mathbf{P}_j\right] \mathbf{B}_1\right)^{-1} \mathbf{B}_1^\intercal \left[\sum_{j=1}^s \mathbf{T}_{1j} \mathbf{P}_j\right] \mathbf{A}_1 \\ \vdots \\ \mathbf{K}_i = -\left(\mathbf{R}_i + \mathbf{B}_i^\intercal \left[\sum_{j=1}^s \mathbf{T}_{ij} \mathbf{P}_j\right] \mathbf{B}_i\right)^{-1} \mathbf{B}_i^\intercal \left[\sum_{j=1}^s \mathbf{T}_{ij} \mathbf{P}_j\right] \mathbf{A}_i \\ \vdots \\ \mathbf{K}_s = -\left(\mathbf{R}_s + \mathbf{B}_s^\intercal \left[\sum_{j=1}^s \mathbf{T}_{sj} \mathbf{P}_j\right] \mathbf{B}_s\right)^{-1} \mathbf{B}_s^\intercal \left[\sum_{j=1}^s \mathbf{T}_{sj} \mathbf{P}_j\right] \mathbf{A}_s \end{cases}$$

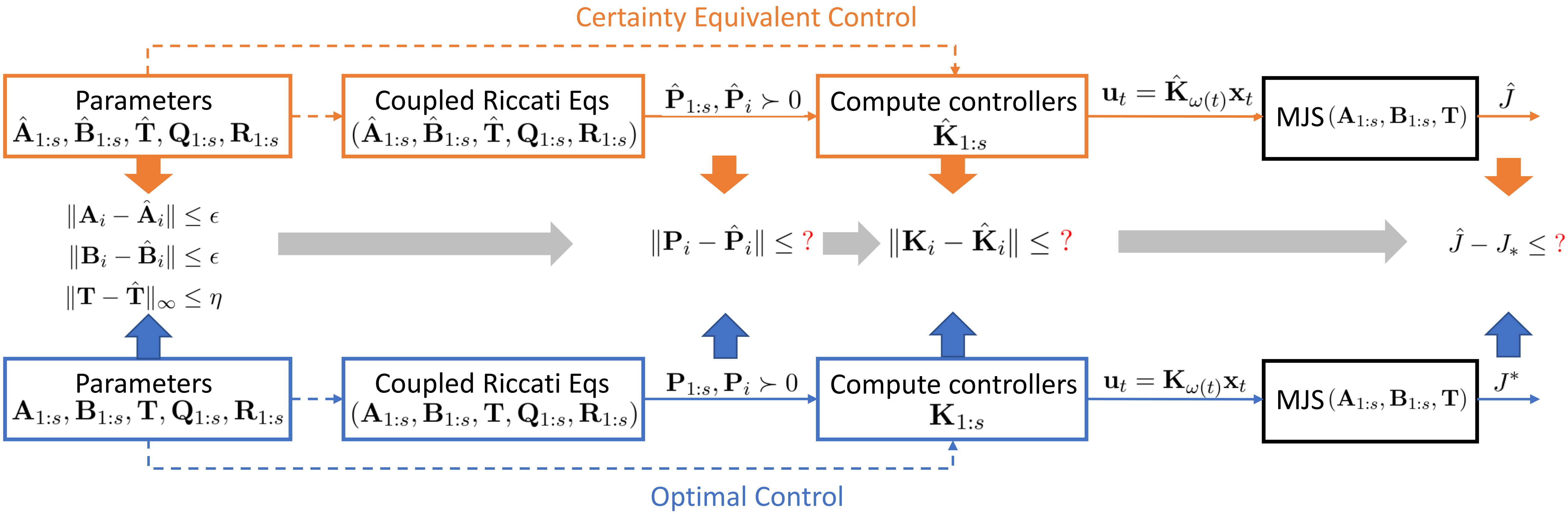


Certainty Equivalent Control for MJS-LQR

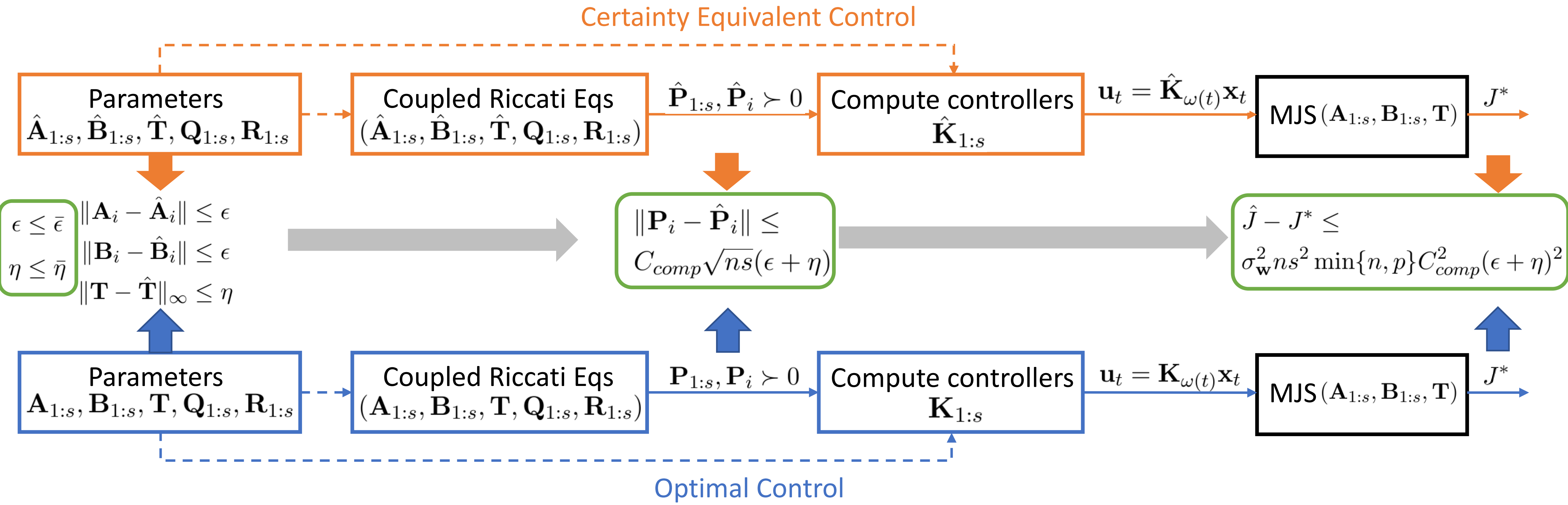
MJS-LQR $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$

$$\min_{\mathbf{u}_0,\mathbf{u}_1,\dots} \lim_{N\to\infty} \frac{1}{T} J_{0:N}$$
$$\text{s.t. } \mathbf{x}_{t+1} = \mathbf{A}_{\omega(t)}\mathbf{x}_t + \mathbf{B}_{\omega(t)}\mathbf{u}_t + \mathbf{w}_t$$
$$\mathrm{P}(\omega(t+1) = j \mid \omega(t) = i) = \mathbf{T}_{ij}$$

However, $\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}$ are unknown, and we only know their *approximate* values $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}$ such that $\|\mathbf{A}_i - \hat{\mathbf{A}}_i\| \leq \epsilon, \|\mathbf{B}_i - \hat{\mathbf{B}}_i\| \leq \epsilon, \|\mathbf{T} - \hat{\mathbf{T}}\|_\infty \leq \eta$

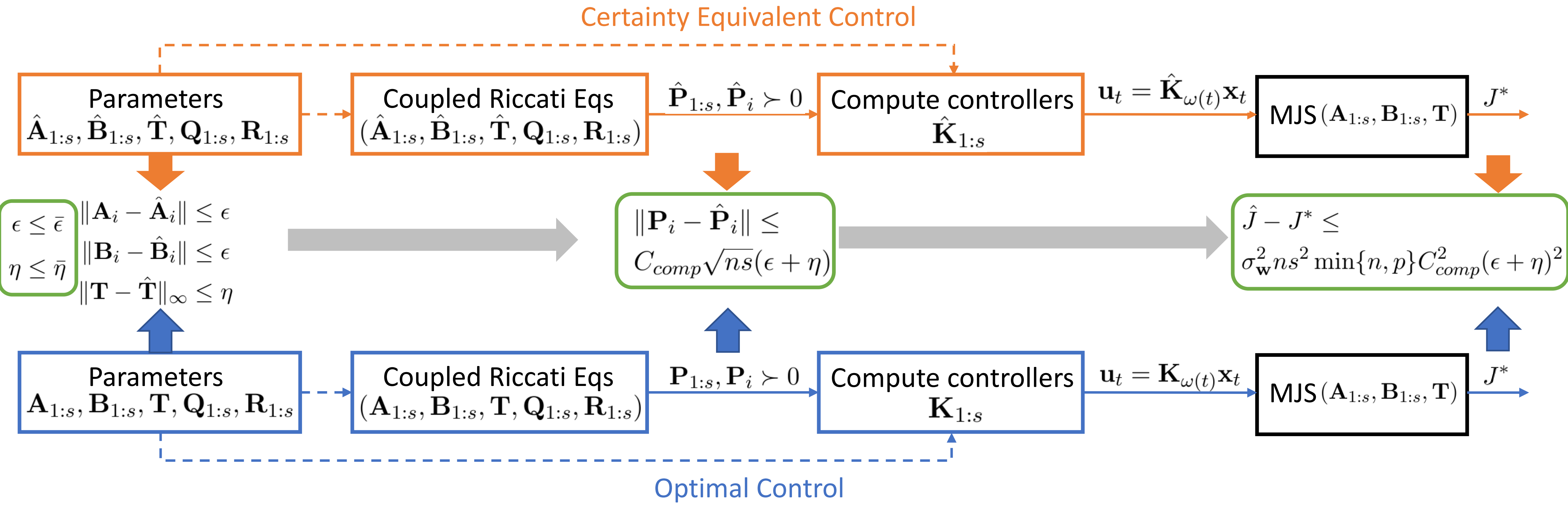


Results



$$C_{comp}, \bar{\epsilon}, \bar{\eta} \downarrow \text{ when stability} \uparrow \text{ and } \mathbf{R}_{1:s} \uparrow$$

Results

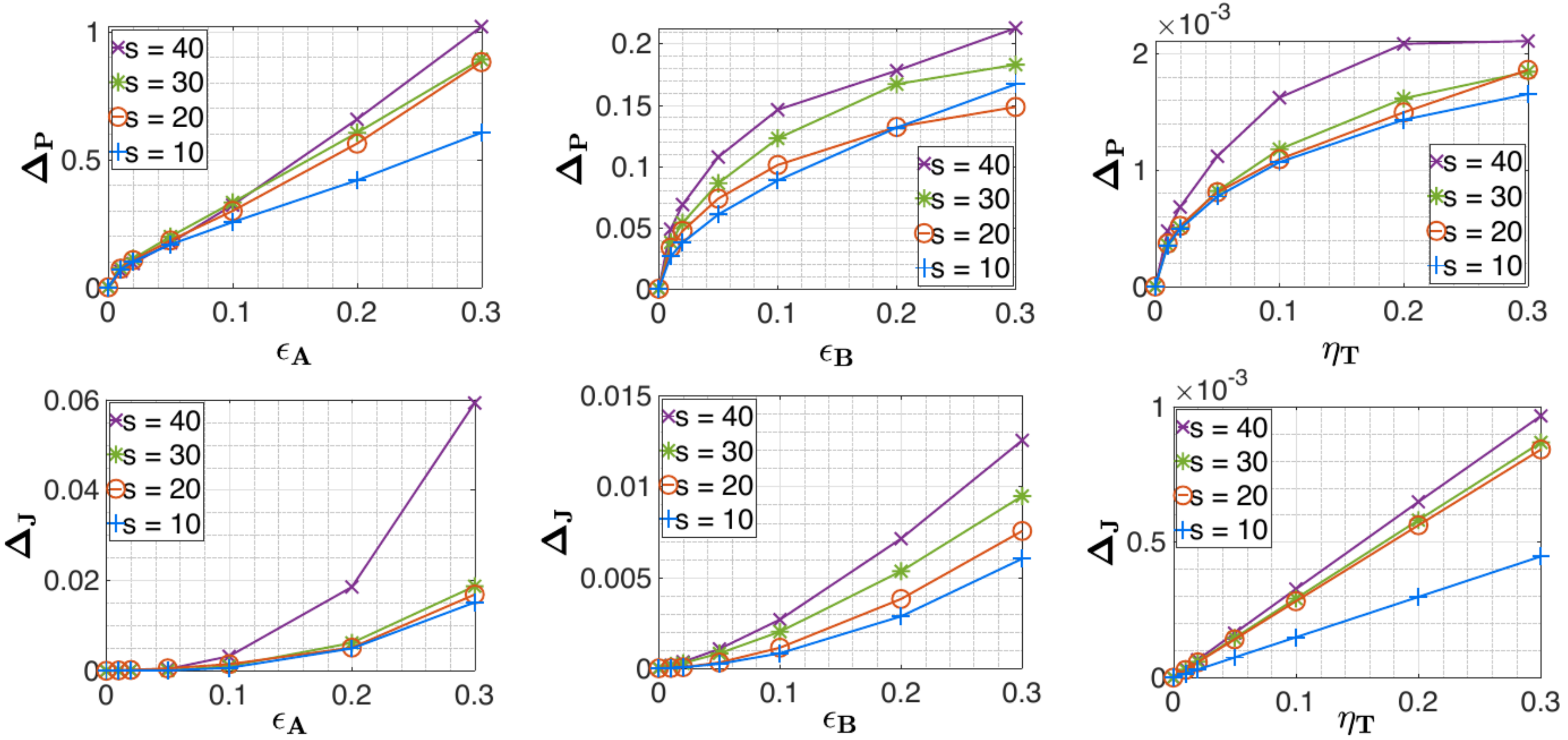


$s = 1$	$\bar{\epsilon}$	$\hat{J} - J^*$
Robust Control $\min \max_{\|\mathbf{\Delta}\| \leq \epsilon} \text{Cost}(\text{Approx. Sys.} + \mathbf{\Delta})$	Moderate	$\mathcal{O}(\epsilon)$
Certainty Equivalent Control $\min \text{Cost}(\text{Approx. Sys.})$	Stringent	$\mathcal{O}(\epsilon^2)$

Experiments

$$\|\mathbf{P}_i - \hat{\mathbf{P}}_i\| \leq C_{comp}\sqrt{ns}(\epsilon + \eta)$$

$$\hat{J} - J^* \leq \sigma_{\mathbf{w}}^2 n s^2 \min\{n, p\} C_{comp}^2 (\epsilon + \eta)^2$$



Future Work

- Offline Data-driven Control

$$\{\mathbf{u}_{0:M}, \mathbf{x}_{0:M}, \omega(0:M)\} \rightarrow \hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}} \rightarrow \hat{J}$$

- Online Adaptive Control

$$\{\mathbf{u}_{0:t}, \mathbf{x}_{0:t}, \omega(0:t)\} \rightarrow \hat{\mathbf{A}}_{1:s,t}, \hat{\mathbf{B}}_{1:s,t}, \hat{\mathbf{T}}_t \rightarrow \mathbf{u}_{t+1}, \mathbf{x}_{t+1}, \omega(t+1)$$

$$\rightarrow \hat{J}$$