Sub-optimality of the Separation Principle for Quadratic Control from Bilinear Observations

Yahya Sattar [1] Sunmook Choi [1] Yassir Jedra [2] Maryam Fazel [3] Sarah Dean [1]

[1]Cornell University [2]MIT [3]University of Washington

Cornell University

Problem Setting

Given linear dynamics and bilinear observations with Gaussian noise

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}\boldsymbol{u}_t + \boldsymbol{w}_t \\ \boldsymbol{y}_t &= (\boldsymbol{C}_0 + \sum_{k=1}^{p} (\boldsymbol{u}_t)_k \boldsymbol{C}_k)\boldsymbol{x}_t + \boldsymbol{z}_t \end{aligned} \quad (1)$$

we want to analyze the finite-horizon optimal control problem:

$$\inf_{\boldsymbol{u}_0,\ldots,\boldsymbol{u}_{T-1}} \mathbb{E}\left[\boldsymbol{x}_T^\top \boldsymbol{Q}_T \boldsymbol{x}_T + \sum_{t=0}^{T-1} \left(\boldsymbol{x}_t^\top \boldsymbol{Q}\boldsymbol{x}_t + \boldsymbol{u}_t^\top \boldsymbol{R}\boldsymbol{u}_t\right)\right] \quad (2)$$

Separation Principle

The optimal control from partial observations can be generated by

1. using Kalman filtering to find the optimal state estimate; and
2. using LQR to find the optimal state feedback controller.

State estimation & Separation Principle Policy

Information available at t: $\mathcal{I}_t = \{\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_{t-1}, \boldsymbol{y}_0, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_{t-1}\}$

- $\hat{\boldsymbol{x}}_{t|t-1} = \mathbb{E}[\boldsymbol{x}_t \mid \mathcal{I}_t]$
- $\boldsymbol{\Sigma}_{t|t-1} = \mathbb{E}\left[(\boldsymbol{x}_t - \mathbb{E}[\boldsymbol{x}_t \mid \mathcal{I}_t])(\boldsymbol{x}_t - \mathbb{E}[\boldsymbol{x}_t \mid \mathcal{I}_t])^\top\right]$

Starting from $\hat{\boldsymbol{x}}_{0|-1} = \hat{\boldsymbol{x}}_0$ and $\boldsymbol{\Sigma}_{0|-1} = \boldsymbol{\Sigma}_0$, ($\because \boldsymbol{x}_0 \sim \mathcal{N}(\hat{\boldsymbol{x}}_0, \boldsymbol{\Sigma}_0)$)

$$\begin{aligned} \hat{\boldsymbol{x}}_{t+1|t} &= \boldsymbol{A}\hat{\boldsymbol{x}}_{t|t-1} + \boldsymbol{B}\boldsymbol{u}_t - \boldsymbol{L}(\boldsymbol{u}_t)\big(\boldsymbol{y}_t - \boldsymbol{C}(\boldsymbol{u}_t)\hat{\boldsymbol{x}}_{t|t-1}\big) \\ \boldsymbol{\Sigma}_{t+1|t} &= \boldsymbol{A}\boldsymbol{\Sigma}_{t|t-1}\boldsymbol{A}^\top + \boldsymbol{L}(\boldsymbol{u}_t)\boldsymbol{C}(\boldsymbol{u}_t)\boldsymbol{\Sigma}_{t|t-1}\boldsymbol{A}^\top + \boldsymbol{\Sigma}_w \end{aligned}$$

where $\boldsymbol{L}(\boldsymbol{u}_t) = -\boldsymbol{A}\boldsymbol{\Sigma}_{t|t-1}\boldsymbol{C}(\boldsymbol{u}_t)^\top\left(\boldsymbol{C}(\boldsymbol{u}_t)\boldsymbol{\Sigma}_{t|t-1}\boldsymbol{C}(\boldsymbol{u}_t)^\top + \boldsymbol{\Sigma}_z\right)^{-1}$.

Optimality of KF: It gives the posterior $\boldsymbol{x}_t \mid \mathcal{I}_t \sim \mathcal{N}(\hat{\boldsymbol{x}}_{t|t-1}, \boldsymbol{\Sigma}_{t|t-1})$.

Separation Principle Policy: $\boldsymbol{K}_T = \boldsymbol{Q}_T$, for $t = T-1, \ldots, 1$

$$\begin{aligned} \boldsymbol{u}_t^* &= \boldsymbol{L}_t\hat{\boldsymbol{x}}_{t|t-1} \\ \boldsymbol{L}_t &= -(\boldsymbol{B}^\top\boldsymbol{K}_{t+1}\boldsymbol{B} + \boldsymbol{R})^{-1}\boldsymbol{B}^\top\boldsymbol{K}_{t+1}\boldsymbol{A} \\ \boldsymbol{P}_t &= \boldsymbol{A}^\top\boldsymbol{K}_{t+1}\boldsymbol{B}(\boldsymbol{B}^\top\boldsymbol{K}_{t+1}\boldsymbol{B} + \boldsymbol{R})^{-1}\boldsymbol{B}^\top\boldsymbol{K}_{t+1}\boldsymbol{A} \\ \boldsymbol{K}_t &= \boldsymbol{A}^\top\boldsymbol{K}_{t+1}\boldsymbol{A} - \boldsymbol{P}_t + \boldsymbol{Q} \end{aligned}$$

Sub-optimality of the Separation Principle

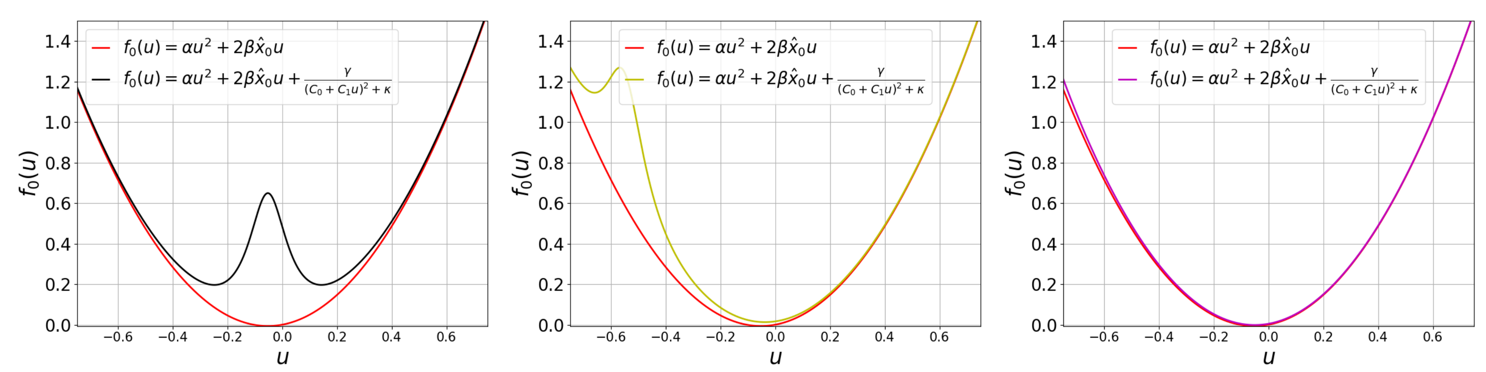
Theorem 1 If $T \geq 2$, the optimal control policy is not affine in the estimated state.

Proof idea: Solution to the scalar case using dynamic programming gives a counterexample.

Theorem 2 There exist instances in which the policy obtained from the separation principle locally maximizes the cost in (2).

Numerical Scalar Example with $T = 2$ for Theorem 2

$$V_{T-2}(\mathcal{I}_{T-2}) = \min_u \underbrace{\mathbb{E}[c(x_{T-2}, u) + V_1(Ax_{T-2} + Bu + w_{T-2}) \mid \mathcal{I}_{T-2}]}_{f_{T-2}(u) = f_{T-2}^{\mathrm{LQG}}(u) + g_{T-2}(u)}$$



(a) $-\frac{C_0}{C_1} = -\frac{\beta\hat{x}_0}{\alpha}$ (b) $-\frac{C_0}{C_1} = -\frac{\beta\hat{x}_0}{\alpha} - 0.5$ (c) $-\frac{C_0}{C_1} = -\frac{\beta\hat{x}_0}{\alpha} - 1.0$

Sufficient Conditions for Uniformly Bounded Cost

Definition The system (1) is **uniformly completely observable** for input sequence $\{\boldsymbol{u}_t\}_{t\geq 0}$ if there exists a $\delta > 0$ such that, for all $\ell \geq 0$,

$$\boldsymbol{O}_\ell = \sum_{k=0}^{n-1} (\boldsymbol{A}^k)^\top \boldsymbol{C}(\boldsymbol{u}_{\ell+k})^\top \boldsymbol{C}(\boldsymbol{u}_{\ell+k})\boldsymbol{A}^k \succ \delta \boldsymbol{I}_n$$

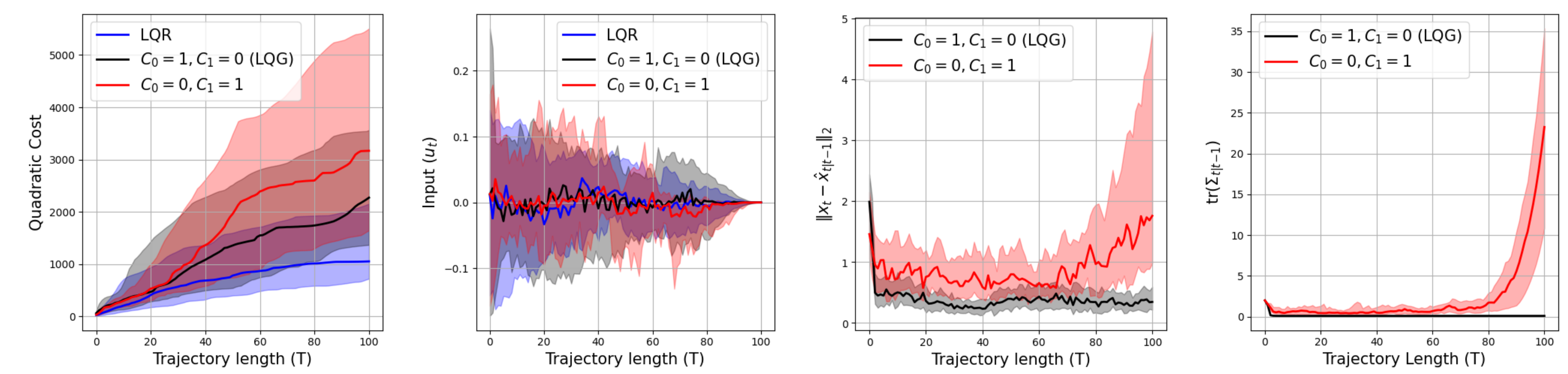
Lemma Suppose (1) is uniformly completely observable for input sequence $\{\boldsymbol{u}_t\}_{t\geq 0}$. Then **the Kalman filtering error covariance matrix remains bounded uniformly in time**.

Proposition Let $\boldsymbol{C}_0^\perp$ be the projection of $\boldsymbol{C}_0$ onto the orthogonal complement of the span $\boldsymbol{C}_1, \ldots, \boldsymbol{C}_p$. Then if $(\boldsymbol{A}, \boldsymbol{C}_0^\perp)$ is observable, for any choice of control inputs $\{\boldsymbol{u}_t\}_{t\geq 0}$, the Kalman filtering error covariance matrix remains bounded for all $t \geq 0$.

Examples

Double Integrator Style Dynamics

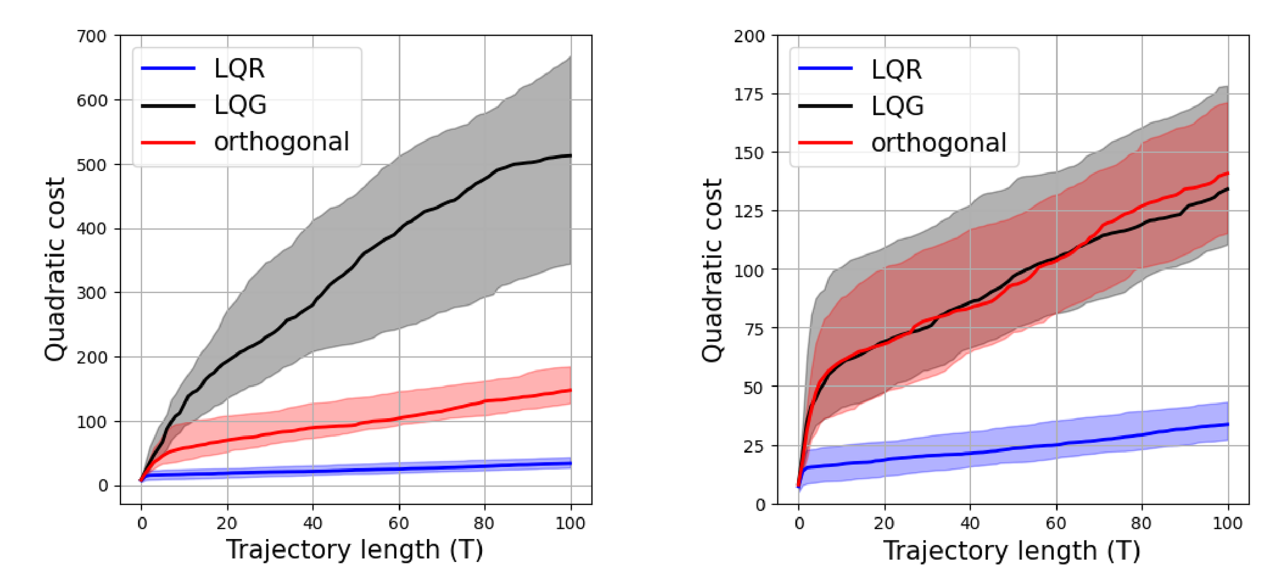
$$\begin{aligned} \boldsymbol{x}_{t+1} &= \begin{bmatrix} 1 & 0.3 \\ 0 & 1 \end{bmatrix} \boldsymbol{x}_t + \begin{bmatrix} 0 \\ 0.3 \end{bmatrix} u_t + \boldsymbol{w}_t \\ y_t &= (C_0 + C_1 u_t) \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}_t + z_t \end{aligned} \qquad \text{with} \quad Q = I,\ R = 1000.$$



- Bilinear observations model incurs higher cost.
- The input magnitude drops towards zero near the end of the trajectory.
- Small inputs lead to a loss of observability, disrupting the state estimate accuracy.

Orthogonal Observations

- Fix $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}_1, \boldsymbol{C}_2, \boldsymbol{C}_3$.
- States $\boldsymbol{x}_t \in \mathbb{R}^6$, inputs $\boldsymbol{u}_t \in \mathbb{R}^3$, outputs $\boldsymbol{y}_t \in \mathbb{R}^3$
- Choose different $\boldsymbol{C}_0$ from the orthogonal complement of $\mathrm{span}(\{\boldsymbol{C}_1, \boldsymbol{C}_2, \boldsymbol{C}_3\})$



The choice of $\boldsymbol{C}_0$ can show different behavior.

- The bilinear observations can improve the quadratic cost compared with LQG.
- The performance can be similar for linear and bilinear observations.