Competitive Control with Delayed Imperfect Information

Chenkai Yu[1] Guanya Shi[2] Soon-Jo Chung[2] Yisong Yue[2] Adam Wierman[2]

[1]Columbia University [2]California Institute of Technology

Linear Quadratic Control with Adversarial Disturbances

- Dynamics: $x_{t+1} = Ax_t + Bu_t + w_t$
- Minimize the total cost: $J = \Sigma_{t=0}^{T-1}(x_t^\top Q x_t + u_t^\top R u_t) + x_T^\top Q_f x_T$
- State $x_t \in \mathbb{R}^n$. Action $u_t \in \mathbb{R}^m$ is selected by the controller. Disturbance $w_t \in \mathbb{R}^n$ is *adaptively* selected by an adversary.
- A, B, Q, R, Q_f are known matrices. $Q, Q_f \succeq 0$, $R \succ 0$.

Modeling Delays and Imperfect Predictions

- In the classic setting (without delays or predictions), at each time step t, the controller observes the current state x_t and then decides the action u_t:
$$u_t = \pi_t(x_0, x_1, \dots, x_t, u_0, \dots, u_{t-1}).$$
Equivalently, the controller is given the initial state x_0 and, at each time t, decides the action u_t before knowing the disturbance w_t:
$$u_t = \pi_t(x_0, u_0, \dots, u_{t-1}, w_0, \dots, w_{t-1}).$$
- **Delays:** Instead of knowing previous disturbances up to time $t-1$, the controller only knows up to time $t-d-1$.
- **Imperfect predictions:** The controller estimates the unknown disturbance of time s to be $\hat{w}_{s|t}$.
- Thus, at each time t, the controller has the following information:
$$x_0, u_0, \dots, u_{t-1}, w_0, \dots, w_{t-d-1}, \hat{w}_{t-d|t}, \dots, \hat{w}_{T-1|t}$$

A Myopic Policy Inspired by Model Predictive Control

Suppose there are d steps of delay and the controller uses k predictions.

At each time t, it minimizes the total future cost with an infinite horizon assuming there is no disturbances after time $t-d+k$:
$$(u_t, u_{t+1}, \dots) = \arg\min_{u} \sum_{i=t-d}^{\infty} (x_i^\top Q x_i + u_i^\top R u_i)$$
where $x_{i+1} = \begin{cases} Ax_i + Bu_i + \hat{w}_{i|t} & \text{if } t-d \le i < t-d+k \\ Ax_i + Bu_i & \text{if } i \ge t-d+k. \end{cases}$

Then, it applies the optimal u_t to the system and discards the rest $u_{t+1}, u_{t+2}, \dots$.

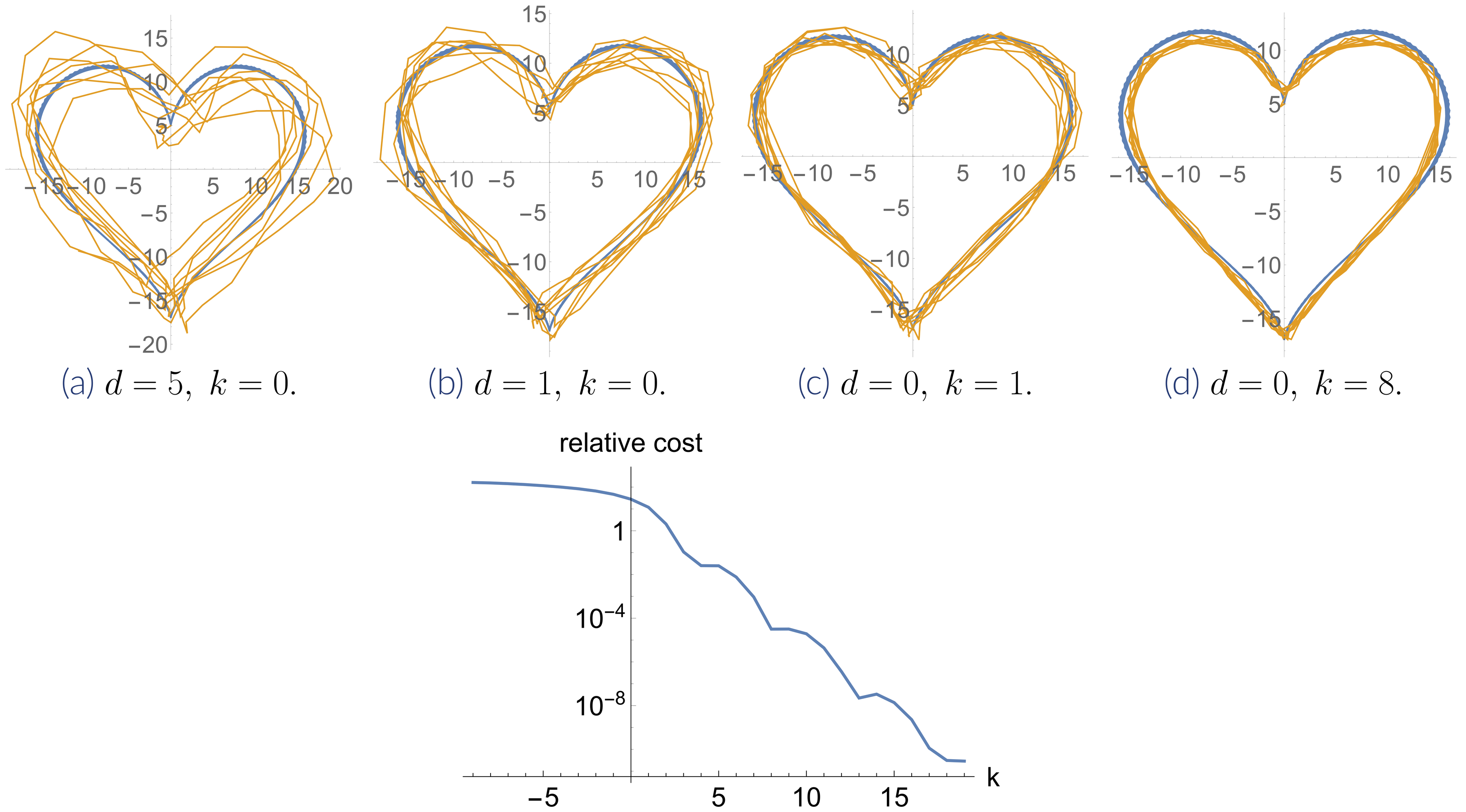
Note that $u_{t-d}, \dots, u_{t-1}$ are fixed and only $u_t, u_{t+1}, \dots$ are optimized. x_{t-d} is the state observed from the system, but $x_{t-d+1}, x_{t-d+2}, \dots$ are based on estimations.

Competitive Ratio Bound (Main Result)

We want to bound the worst-case ratio of the cost of an online policy (Alg) to the cost of the optimal offline policy (Opt) that knows all future information.

Let ϵ_i be an upper bound of the relative estimation errors for disturbances that are i steps into the future, i.e., for all t, $\|w_{t-d+i} - \hat{w}_{t-d+i|t}\| \le \epsilon_i \|w_{t-d+i}\|$. The performance bound of an control policy will depend on $(\epsilon_i)_{i \ge 0}$.

Let $F = A - B(R + B^\top PB)^{-1}B^\top PA = A - BK$, $H = B(R + B^\top PB)^{-1}B^\top$, and $c = \|P\|\|P^{-1}\|(1 + \|F\|)$. Suppose there are d steps of delays and the controller uses k predictions. When $k \ge d$,
$$\text{Alg} \le \left[\frac{\left(c\Sigma_{i=0}^{d-1}\epsilon_i\|A^{d-i}\| + c\Sigma_{i=d}^{k-1}\epsilon_i\|F^{i-d}\| + \|F^{k-d}\|\right)^2}{\|H\|^{-1}\lambda_{\min}(P^{-1} - FP^{-1}F^\top - H)} + 1\right]\text{Opt} + O(1).$$
When $k \le d$,
$$\text{Alg} \le \left[\frac{\left(c\Sigma_{i=0}^{k-1}\epsilon_i\|A^{d-i}\| + c\Sigma_{i=k}^{d-1}\|A^{d-i}\| + 1\right)^2}{\|H\|^{-1}\lambda_{\min}(P^{-1} - FP^{-1}F^\top - H)} + 1\right]\text{Opt} + O(1).$$
The $O(1)$ term does not grow with T, though it may depend on the system parameters A, B, Q, R, Q_f. When $Q_f = P$, this $O(1)$ is zero.



(a) $d = 5,\ k = 0$. (b) $d = 1,\ k = 0$. (c) $d = 0,\ k = 1$. (d) $d = 0,\ k = 8$.

(e) Relative cost: Alg/Opt − 1.

Figure 1. Tracking results with k exact predictions and no delay, or with d steps of delay with no predictions. (a-d) show the desired trajectory (blue) and actual (orange) trajectories. (e) shows both delay and prediction on the x-axis, with the negative part corresponding to delay. The y-axis is in log-scale.

Special Cases

Exact Predictions Without Delay

The competitive ratio goes to 1 exponentially fast in the number k of predictions.
$$\text{Alg} \le \left[1 + \frac{\|F^k\|^2\|H\|}{\lambda_{\min}(P^{-1} - FP^{-1}F^\top - H)}\right]\text{Opt} + O(1).$$

Tight in the sense that there exist systems where the competitive ratio of the *optimal* online algorithm is $1 + \Theta(\|F^k\|^2)$.

Inexact Predictions Without Delay

The prediction quality in the near future is (exponentially) more important than the one in far future. This is consistent with robust MPC literature.
$$\text{Alg} \le \left[\frac{\|H\|\left(c\Sigma_{i=0}^{k-1}\epsilon_i\|F^i\| + \|F^k\|\right)^2}{\lambda_{\min}(P^{-1} - FP^{-1}F^\top - H)} + 1\right]\text{Opt} + O(1).$$

Delay Without Predictions

If $\rho(A) < 1$, the competitive ratio is bounded by a constant irrelevant to the length of delay:
$$\text{Alg} \le \left[\frac{\|H\|\left(c\kappa\frac{a}{1-a} + 1\right)^2}{\lambda_{\min}(P^{-1} - FP^{-1}F^\top - H)} + 1\right]\text{Opt} + O(1).$$

If $\rho(A) > 1$, the competitive ratio bound grows exponentially fast with the of number of delay steps:
$$\text{Alg} \le \left[\frac{\|H\|\left(c\frac{\|A\|^{d+1} - \|A\|}{\|A\| - 1} + 1\right)^2}{\lambda_{\min}(P^{-1} - FP^{-1}F^\top - H)} + 1\right]\text{Opt} + O(1).$$
Tight in the sense that there exist systems such that the competitive ratio of the *optimal* online algorithm is at least $1 + \Theta(\|A^d\|^2)$.

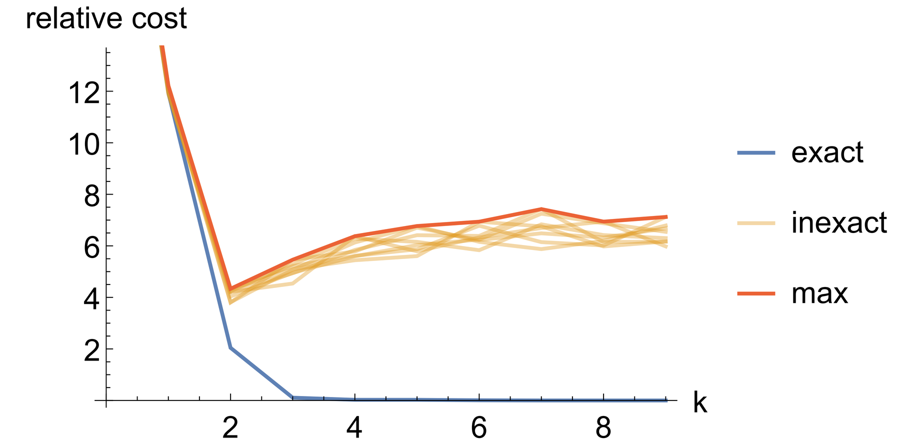


Figure 2. The impact of inexact predictions: the relative cost (Alg/Opt − 1) of our MPC-like policy using k exact (blue) or inexact (orange) predictions.