Perturbation-based Regret Analysis of Predictive Control in Linear Time Varying Systems

Yiheng Lin[1], Yang Hu[2], Guanya Shi[1], Haoyuan Sun[1], Guannan Qu[1], Adam Wierman[1]

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

35th Conference on Neural Information Processing Systems

Problem Setting

We consider a finite-horizon discrete-time online control problem with linear time-varying (LTV) dynamics:

$$\min_{x_{0:T}, u_{0:T-1}} \sum_{t=1}^{T} \left(f_t(x_t) + c_t(u_{t-1})\right)$$
$$\text{s.t. } x_t = A_{t-1}x_{t-1} + B_{t-1}u_{t-1} + w_{t-1}, t = 1, \ldots, T, \quad (1)$$
$$x_0 = x(0),$$

where $x_t \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, and $w_t \in \mathbb{R}^n$ denote the state, the control action, and the disturbance of the system. Define the info tuple at time t as $\vartheta_t := (A_t, B_t, w_t, f_{t+1}, c_{t+1})$. The prediction model is

$$x_0, \vartheta_0, \vartheta_1, \ldots, \vartheta_{k-1}, u_0, x_1, \vartheta_k, u_1, x_1, \vartheta_{k+1}, \ldots.$$

Definition 1. *The transition matrix* $\Phi(t_2, t_1) \in \mathbb{R}^{n \times n}$ *is defined as*

$$\Phi(t_2, t_1) := \begin{cases} A_{t_2-1}A_{t_2-2}\cdots A_{t_1} & \text{if } t_2 > t_1 \\ I & \text{if } t_2 \le t_1 \end{cases},$$

and the controllability matrix $M(t, p) \in \mathbb{R}^{n \times (mp)}$ *is defined as*

$$M(t, p) := [\Phi(t + p, t + 1)B_t, \ldots, \Phi(t + p, t + p)B_{t+p}].$$

Assumption 1. *We assume the costs and dynamics satisfy that*
1. *The state costs f_t and control costs c_t are well-conditioned;*
2. $\arg\min_x f_t(x) = 0$ *and* $\arg\min_u c_t(u) = 0$ *without loss of generality;*
3. $\|A_t\|, \|B_t\|, \|B_t^\dagger\|$ *are bounded, and* $\sigma_{min}(M(t, d)) \ge \sigma$.

Predictive Control PC_k

We define $\tilde{\psi}_t^p(x, \zeta; F)$ as the optimal solution to

$$\arg\min_{y_{0:p}, v_{0:p-1}} \sum_{\tau=1}^{p} f_{t+\tau}(y_\tau) + \sum_{\tau=1}^{p} c_{t+\tau}(v_{\tau-1}) + F(y_k)$$
$$\text{s.t. } y_{\tau+1} = A_{t+\tau}y_\tau + B_{t+\tau}v_\tau + \zeta_\tau, \tau = 0, \ldots, p-1, \quad (2)$$
$$y_0 = x,$$

where the terminal cost F is convex, nonnegative, and satisfies $F(0) = 0$. $\psi_t^p(x, \zeta, z)$ is the optimal solution to

$$\arg\min_{y_{0:p}, v_{0:p-1}} \sum_{\tau=1}^{p} f_{t+\tau}(y_\tau) + \sum_{\tau=1}^{p} c_{t+\tau}(v_{\tau-1})$$
$$\text{s.t. } y_{\tau+1} = A_{t+\tau}y_\tau + B_{t+\tau}v_\tau + \zeta_\tau, \tau = 0, \ldots, p-1, \quad (3)$$
$$y_0 = x, y_p = z.$$

We study predictive control with prediction length k:

Algorithm 1 Predictive Control (PC_k)

```
for t = 0, 1, ..., T − k − 1 do
    Observe current state x_t and receive predictions ϑ_{t:t+k−1}.
    Solve and commit control actions u_t := ψ̃_t^k(x_t, w_{t:t+k−1}; F)_{v_0}.
At time step t = T − k, observe current state x_t and receive predictions ϑ_{t:T−1}.
Solve and commit control actions u_{t:T−1} := ψ̃_t^k(x_t, w_{t:T−1}; 0)_{v_{0:k−1}}.
```

We show the dynamic regret and the competitive ratio for predictive control improve exponentially w.r.t. prediction length in a linear time varying system via a new perturbation approach.

Take a picture to download the full paper

Perturbation Bounds

We first show a perturbation bound for the unconstrained Smoothed Online Convex Optimization (SOCO) problem:

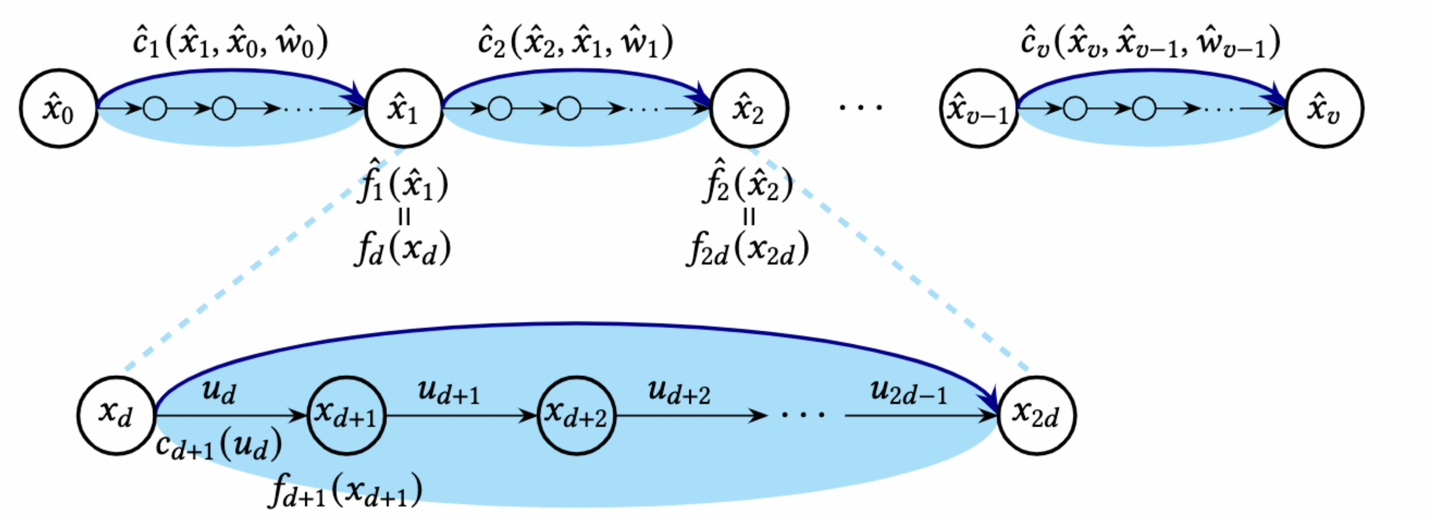
Theorem 1. *Consider the optimal solution of the SOCO problem*

$$\hat{\psi}(\hat{x}_0, \hat{w}, \hat{x}_p) := \arg\min_{\hat{x}_{1:p-1}} \sum_{\tau=1}^{p-1} \hat{f}_\tau(\hat{x}_\tau) + \sum_{\tau=1}^{p} \hat{c}_\tau(\hat{x}_\tau, \hat{x}_{\tau-1}, \hat{w}_{\tau-1})$$

indexed by $1, \ldots, p-1$. *Assume* $\hat{f}_\tau : \mathbb{R}^n \to \mathbb{R}$ *is* μ*-strongly convex,* $\hat{c}_\tau : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^r \to \mathbb{R}$ *is convex and* ℓ*-strongly smooth, then the impact* $\|\hat{\psi}(\hat{x}_0, \hat{w}, \hat{x}_p)_h - \hat{\psi}(\hat{x}_0', \hat{w}', \hat{x}_p')_h\|$ *can be upper bounded by*

$$C_0\left(\lambda_0^{h-1}\|\hat{x}_0 - \hat{x}_0'\| + \sum_{\tau=0}^{p-1} \lambda_0^{|h-\tau|-1}\|\hat{w}_\tau - \hat{w}_\tau'\| + \lambda_0^{p-h-1}\|\hat{x}_p - \hat{x}_p'\|\right),$$

where $C_0 = (2\ell)/\mu$ *and* $\lambda_0 = 1 - 2 \cdot \left(\sqrt{1 + (2\ell/\mu)} + 1\right)^{-1}$.



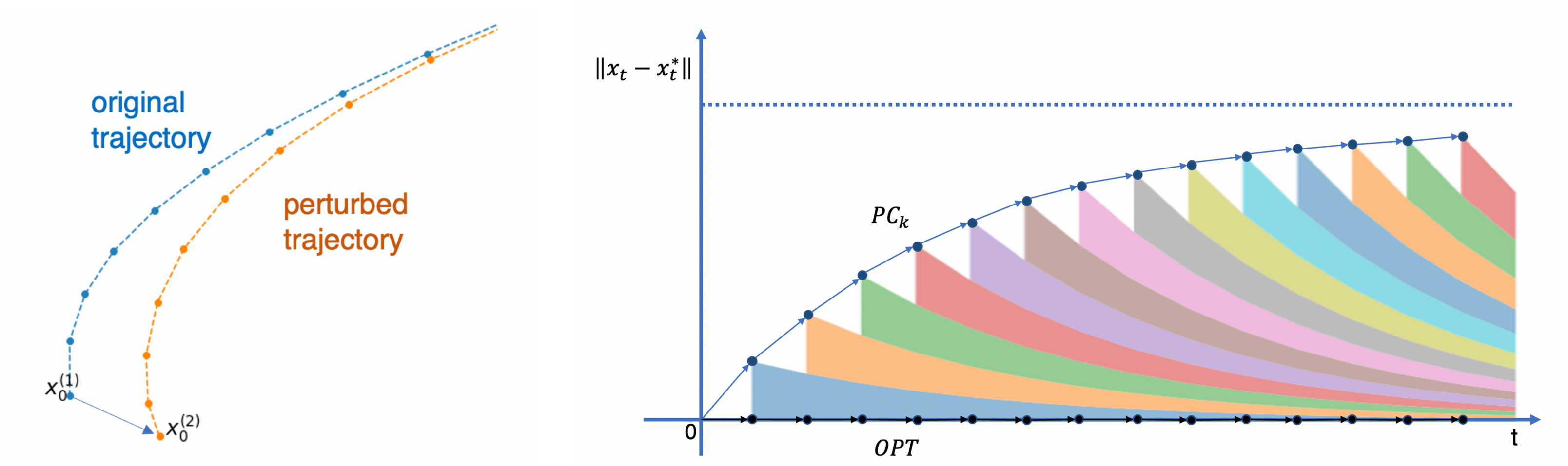
By the controllability assumption and the principle of optimality, we can reduce the LTV problem to a SOCO problem.

Theorem 2. *Under Assumption 1,* $\tilde{\psi}$ *satisfies*

$$\|\tilde{\psi}_t^p(x, \zeta; F)_{y_h} - \tilde{\psi}_t^p(x', \zeta'; F)_{y_h}\| \le C\left(\lambda^h\|x - x'\| + \sum_{\tau=0}^{p-1} \lambda^{|h-\tau|}\|\zeta_\tau - \zeta_\tau'\|\right),$$

and $\|\psi_t^p(x, \zeta, z)_{y_h} - \psi_t^p(x', \zeta', z')_{y_h}\|$ *is upper bounded by*

$$C\left(\lambda^h\|x - x'\| + \sum_{\tau=0}^{p-1} \lambda^{|h-\tau|}\|\zeta_\tau - \zeta_\tau'\| + \lambda^{p-h}\|z - z'\|\right).$$



Performance Guarantees

By Theorem 2, we can show the per-step error injection is $O(\lambda^k)$, and the accumulative error has the same magnitude up to a constant factor.

Theorem 3. *When the prediction window k is large enough,*
1. *The closed-loop dynamics of PC_k is input-to-state stable;*
2. *PC_k achieves an $O(\lambda^k T)$ dynamic regret if $\|w_t\| \le D$;*
3. *PC_k achieves a $1 + O(\lambda^k)$ competitive ratio if F is the indicator of 0.*