Robustness and Consistency in Linear Quadratic Control with Untrusted Predictions

Tongxin Li
California Institute of Technology
Pasadena, United States
tongxin@caltech.edu

Ruixiao Yang
Tsinghua University
Beijing, China
yangruixiaoyrx@gmail.com

Guannan Qu
Carnegie Mellon University
Pittsburgh, United States
gqu@andrew.cmu.edu

Guanya Shi
California Institute of Technology
Pasadena, United States
gshi@caltech.edu

Chenkai Yu
Columbia University
New York City, United States
cyu26@gsb.columbia.edu

Adam Wierman
California Institute of Technology
Pasadena, United States
adamw@caltech.edu

Steven Low
California Institute of Technology
Pasadena, United States
slow@caltech.edu

ABSTRACT

We study the problem of learning-augmented predictive linear quadratic control. Our goal is to design a controller that balances *"consistency"*, which measures the competitive ratio when predictions are accurate, and *"robustness"*, which bounds the competitive ratio when predictions are inaccurate. We propose a novel λ-confident controller and prove that it maintains a competitive ratio upper bound of $1 + \min\{O(\lambda^2\varepsilon) + O(1-\lambda)^2, O(1) + O(\lambda^2)\}$ where $\lambda \in [0, 1]$ is a trust parameter set based on the confidence in the predictions, and ε is the prediction error. Further, motivated by online learning methods, we design a self-tuning policy that adaptively learns the trust parameter λ with a competitive ratio that depends on ε and the variation of system perturbations and predictions. We show that its competitive ratio is bounded from above by $1 + O(\varepsilon)/(\Theta(1) + \Theta(\varepsilon)) + O(\mu_{\mathsf{Var}})$ where μ_{Var} measures the variation of perturbations and predictions. It implies that by automatically adjusting the trust parameter online, the self-tuning scheme ensures a competitive ratio that does not scale up with the prediction error ε.

ACM Reference Format:
Tongxin Li, Ruixiao Yang, Guannan Qu, Guanya Shi, Chenkai Yu, Adam Wierman, and Steven Low. 2022. Robustness and Consistency in Linear Quadratic Control with Untrusted Predictions. In *Abstract Proceedings of the 2022 ACM SIGMETRICS/IFIP PERFORMANCE Joint International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS/PERFORMANCE '22 Abstracts), June 6–10, 2022, Mumbai, India.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3489048.3522658

This work is supported by the National Science Foundation, under grants ECCS1931662, CCF 1637598, ECCS 1619352, CPS 1739355, AitF-1637598, CNS-1518941, PIMCO and Amazon Web Services. Tongxin Li and Ruixiao Yang contributed equally to the paper.

ACM ISBN 978-1-4503-9141-2/22/06.
https://doi.org/10.1145/3489048.3522658

The full version paper corresponding to this abstract is [1].

1 PROBLEM STATEMENT

We study a classical online linear quadratic control problem where the controller has access to untrusted predictions/advice during each round, potentially from a black-box AI tool.

Denote by $x_t \in \mathbb{R}^n$ and $u_t \in \mathbb{R}^m$ the system state and action at each time t. We consider a linear dynamic system with adversarial perturbations,

$$x_{t+1} = Ax_t + Bu_t + w_t, \text{ for } t = 0, \dots, T-1, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$, and $w_t \in \mathbb{R}^n$ denotes some unknown perturbation chosen adversarially. We make the standard assumption that the pair (A, B) is stabilizable. Without loss of generality, we also assume the system is initialized with some fixed $x_0 \in \mathbb{R}^n$. The goal of control is to minimize the following quadratic costs given matrices A, B, Q, R :

$$J \coloneqq \sum_{t=0}^{T-1} (x_t^\top Q x_t + u_t^\top R u_t) + x_T P x_T,$$

where $Q, R \succ 0$ are positive definite matrices, and P is the solution of the following discrete algebraic Riccati equation (DARE), which exists because (A, B) is stabilizable and $Q, R \succ 0$.

$$P = Q + A^\top PA - A^\top PB(R + B^\top PB)^{-1} B^\top PA.$$

Given P, we can define $K \coloneqq (R + B^\top PB)^{-1} B^\top PA$ as the optimal LQC controller in the case of no disturbance ($w_t = 0$). Further, let $F \coloneqq A - BK$ be the closed-loop system matrix when using $u_t = -Kx_t$ as the controller.

Our focus is on predictive control and we assume that, at the beginning of the control process, a sequence of predictions of the disturbances $(\widehat{w}_0, \dots, \widehat{w}_{T-1})$ is given to the decision maker. At time t, the decision maker observes x_t, w_{t-1} and picks a decision u_t. Then, the environment picks w_t, and the system transitions to the next step according to (1). We emphasize that, at time t, the decision maker has no access to $(w_t, \dots, w_T)$ and their values may be different from the predictions $(\widehat{w}_t, \ldots, \widehat{w}_T)$. Also, note that w_t can be adversarially chosen at each time t, adaptively.

Formally, we define the prediction error as

$$\varepsilon := \sum_{t=0}^{T-1} \left\| \sum_{\tau=t}^{T-1} \left(F^\top\right)^{\tau-t} P \left(w_t - \widehat{w}_t\right) \right\|^2 . \tag{2}$$

We use the competitive ratio to measure the performance of an online control policy and quantify its robustness and consistency. Specifically, let OPT be the offline optimal cost when all the disturbances $(w_0, \ldots, w_{T-1})$ are known in hindsight, and ALG be the cost achieved by an online algorithm.

Definition 1.1. *The* ***competitive ratio*** *for a given prediction error* ε, $\mathrm{CR}(\varepsilon)$, *is defined as the smallest constant* $C \geq 1$ *such that* $\mathrm{ALG} \leq C \cdot \mathrm{OPT}$ *for fixed* A, B, Q, R *and any adversarially and adaptively chosen perturbations* $(w_0, \ldots, w_{T-1})$ *and predictions* $(\widehat{w}_0, \ldots, \widehat{w}_{T-1})$.

2 ALGORITHM AND MAIN RESULTS

2.1 λ-confident control

We introduce a new *trust parameter* λ and consider a policy

$$\pi(x_t) = -(R + B^\top P B)^{-1} B^\top \left(P A x_t + \lambda \sum_{\tau=t}^{T-1} \left(F^\top\right)^{\tau-t} P \widehat{w}_\tau \right). \tag{3}$$

Note that setting $\lambda = 0$ and $\lambda = 1$ respectively recovers the optimal linear policy $\overline{\pi}(x_t) = -K x_t$ for the LQR problem with Gaussian perturbations and the MPC policy $\widehat{\pi}(x_t)$ below that gives an action u_t at each time t:

$$\min_{(u_t, \ldots, u_{T-1})} \left(\sum_{\tau=t}^{T-1} (x_\tau^\top Q x_\tau + u_\tau^\top R u_\tau) + x_T P x_T \right)$$
$$\text{s.t. (1) for all } \tau = t, \ldots, T-1. \tag{4}$$

The theorem below establishes a consistency and robustness trade-off, i.e, the optimal confidence parameter λ depends on the prediction error and a large λ gives a better competitive ratio if the prediction error is small and vice versa.

Theorem 2.1 (Informal). *Under our model assumptions, with a fixed trust parameter* $\lambda > 0$, *the* λ*-confident control in* (3) *has a worst-case competitive ratio of at most* $\mathrm{CR}(\varepsilon) \leq 1 + \min\{O(\lambda^2 \varepsilon) + O(1-\lambda)^2, O(1) + O(\lambda^2 \overline{W})\}$ *where* $\overline{W} := \sum_{t=0}^{T-1} \left\| \sum_{\tau=t}^{T-1} \left(F^\top\right)^{\tau-t} P \widehat{w}_\tau \right\|^2$.

2.2 Self-tuning control

While the λ-confident control finds a balance between consistency and robustness, selecting the optimal λ parameter requires exogenous knowledge of the quality of the predictions ε, which is often not possible. For example, black-box AI tools typically do not allow uncertainty quantification. In this section, we develop a self-tuning λ-confident control approach that learns to tune λ in an online manner, as shown in Algorithm 1.

The key to the algorithm is the update rule for λ_t. Given previously observed perturbations and predictions, the goal of the algorithm is to find a greedy λ_t that minimizes the gap between the algorithmic and optimal costs. This can be equivalently written as Solving the formed minimization yields the choice of λ_t in the self-tuning policy in Algorithm 1.

Algorithm 1: Self-Tuning λ-Confident Control

for $t = 0, \ldots, T-1$ **do**
  if $t \leq 1$ **then**
    Initialize and choose $\lambda_t = \lambda_0$
  end
  else
    Compute a trust parameter λ_t

$$\lambda_t = \frac{\sum_{s=0}^{t-1} \left(\eta(w; s, t-1)\right)^\top H \left(\eta(\widehat{w}; s, t-1)\right)}{\sum_{s=0}^{t-1} \left(\eta(\widehat{w}; s, t-1)\right)^\top H \left(\eta(\widehat{w}; s, t-1)\right)}$$

$$\text{where } \eta(w; s, t) := \sum_{\tau=s}^{t} \left(F^\top\right)^{\tau-s} P w_\tau$$

  end
  Generate an action u_t using λ_t-confident control in (3)
  Update $x_{t+1} = A x_t + B u_t + w_t$
end

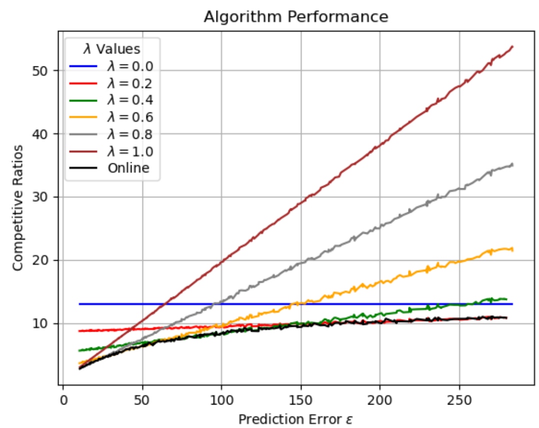

Figure 1: Competitive ratios of Algorithm 1 and λ-confident control with fixed λ's for battery-buffered EV charging.

Theorem 2.2 (Informal). *Under our model assumptions, the self-tuning policy in Algorithm 1 has a competitive ratio* $\mathrm{CR}(\varepsilon) \leq 1 + \frac{O(\varepsilon)}{\Theta(1)+\Theta(\varepsilon)} + O(\mu_{\mathrm{Var}})$ *as a function of the prediction error* ε *where* μ_{Var} *measures the variation of perturbations and predictions.*

Theorem 2.2 implies that when the variations of predictions and perturbations are small, the self-tuning policy is able to achieve a bounded competitive ratio. In Figure 1, we observe a competitive ratio curve (Online) $1 + \Theta(\varepsilon)/(O(1) + \Theta(\varepsilon))$ corresponding to Algorithm 1 that matches the competitive ratio bound given in Theorem 2.2 in order sense (in ε). More case studies are demonstrated in the full paper [1].

In conclusion, we detail online learning-based self-tuning policy that allows the use of untrusted black-box AI tools in a way that ensures worst-case performance bounds for linear quadratic control.

REFERENCES

[1] Tongxin Li, Ruixiao Yang, Guannan Qu, Guanya Shi, Chenkai Yu, Adam Wierman, and Steven Low. Robustness and consistency in linear quadratic control with untrusted predictions. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 6(1):1–35, 2022.