Stable Online Control of Linear Time-Varying Systems

Guannan Qu[*]   GQU@CALTECH.EDU
Yuanyuan Shi[*]   YSHI7@CALTECH.EDU
Sahin Lale[*]   ALALE@CALTECH.EDU
Anima Anandkumar   ANIMA@CALTECH.EDU
Adam Wierman   ADAMW@CALTECH.EDU
California Institute of Technology, Pasadena, CA
[*]*Equal contribution*

Editors: A. Jadbabaie, J. Lygeros, G. J. Pappas, P. A. Parrilo, B. Recht, C. J. Tomlin, M. N. Zeilinger

Abstract

Linear time-varying (LTV) systems are widely used for modeling real-world dynamical systems due to their generality and simplicity. Providing stability guarantees for LTV systems is one of the central problems in control theory. However, existing approaches that guarantee stability typically lead to significantly sub-optimal cumulative control cost in online settings where only current or short-term system information is available. In this work, we propose an efficient online control algorithm, COvariance Constrained Online Linear Quadratic (COCO-LQ) control, that guarantees input-to-state stability for a large class of LTV systems while also minimizing the control cost. The proposed method incorporates a state covariance constraint into the semi-definite programming (SDP) formulation of the LQ optimal controller. We empirically demonstrate the performance of COCO-LQ in both synthetic experiments and a power system frequency control example.

Keywords: Time-varying systems, online linear quadratic control, stability guarantee

1. Introduction

Time-invariant systems have traditionally been the main focus of the study for the linear dynamical systems community. However, real-world systems are often *time-varying.* For example, consider a power system that includes renewable generation (e.g. solar/wind). Due to the intermittency of renewable energy, the system dynamics for frequency regulation in the power system are time-varying. Applying a time-invariant controller in this setting may lead to frequency instability and line failures (Ulbig et al., 2014). Time-varying systems are also crucial for many other applications, such as autonomous vehicles and aircraft control (Falcone et al., 2008). While not all time-varying systems have linear dynamics, many applications can be approximated by linear time-varying (LTV) systems via a local linear approximation at each time step (Todorov and Li, 2005), e.g., the frequency control example described above. As a result, LTV systems are widely-used and there is a large literature focused on designing controllers for LTV systems (Amato et al., 2010; Ouyang et al., 2017).

Perhaps the most fundamental challenge in dynamical systems is stability. While the design of stable linear time-invariant (LTI) systems is well understood, the same cannot be said for LTV systems. To this point, several notions of stability have recieved attention, e.g., input-to-state stability (ISS), mean-square stability and Lyapunov stability. ISS is the most widely adopted notion and aims to guarantee the boundedness of the state given bounded initial conditions (Hong et al., 2010).

In most applications of LTV systems, it is crucial to guarantee ISS both in order to avoid saturation, maintain the robustness and validity of linearization (Tarbouriech et al., 2006; Khalil, 2002).

While there is considerable prior work focused on stability in LTV systems, most prior work studies stability in the offline setting where either the sequence of system parameters are known, e.g., (Amato et al., 2010; Li et al., 2019a), or the system parameters have a particular variation pattern, e.g., (Garcia et al., 2009). Maintaining ISS guarantees becomes significantly harder in the online setting where the system parameters are observed in real-time and may have arbitrary variations. This online setting is the most relevant to many applications, e.g., frequency regulation.

Though stability is crucial, it is not enough for a controller to be stable. A controller must also have low cost. For instance, in order to stabilize the dynamics, a controller may use arbitrarily big control inputs, which may result in sub-optimal cost. In classical optimal control problems, e.g. the time-varying linear quadratic (LQ) control setting, the goal is to design a stabilizing controller that minimizes the cost for a particular finite horizon while assuming access to the whole trajectory for that duration. It is possible to characterize the optimal policy in such settings (Bertsekas et al., 1995); however, in the online setting when only current or short-termed system information is available, these methods may not guarantee stability, e.g., see Section 3. There have been recent efforts to provide sub-optimality guarantees on the acquired cost in the online LTV setting, e.g., (Gradu et al., 2020), but it is unclear if the proposed controllers maintain stability for all time-steps since the main focus is on minimizing the cumulative cost.

Thus, despite considerable recent work, much remains to be understood about the design of online LTV controllers. In particular, this paper is motivated by the following question:

Is it possible for an online controller to guarantee stability and maintain low cost in LTV systems?

Contributions. In this work, we answer question above affirmatively. Specifically, we propose **Co**variance **C**onstrained **O**nline **L**inear **Q**uadratic (COCO-LQ) control, a novel online control algorithm that aims to minimize the control cost while ensuring provable stability guarantees in LTV systems without restricting how slow or fast the underlying system changes. Further, we demonstrate the performance of the proposed method in various synthetic LTV systems and in the power system frequency control example that motivated our study.

The main technical contribution of the paper is a stability guarantee for COCO-LQ in LTV systems. Specifically, we show that COCO-LQ guarantees ISS in online time-varying systems. The key technique that underpins the proposed algorithm is the addition of a novel semi-definiteness constraint on the state covariance matrix into the standard online semi-definite programming (SDP) formulation of linear quadratic optimal control. We show that this constraint promotes the sequential strong stability of the controllers (Cohen et al., 2018), which in turn guarantees ISS with a proper choice of an algorithm hyperparameter. Adding this additional constraint is simple and does not result in a significant increase of computational complexity compared to the standard LQ formulation. Moreover, we prove that if the proposed SDP is not directly feasible, short-term predictions on the future system parameters are necessary and can be used in COCO-LQ in order to ensure ISS.

Related work. The work in this paper builds on the design of linear time-invariant (LTI) controllers to provide a new approach for the design of stable controllers for linear-time-varying (LTV) systems. As such, we describe related work on both LTI and LTV systems below.

LTI Systems. In study of control of LTI systems, linear quadratic regulator (LQR) has been considered in detail. In the classical setting where the underlying system is known, the optimal control

law is given by a linear feedback controller obtained by solving Riccati equations (Bertsekas et al., 1995). Alternatively, the optimal control problem can also be posed via semi-definite programming (SDP) (Vandenberghe and Boyd, 1996), which is the approach we build on in the current paper.

Recently, there has been growing interest in online control of these linear systems when the underlying dynamics are unknown. Most of these works study the problem with a regret minimization perspective, e.g., (Abbasi-Yadkori and Szepesvári, 2011; Dean et al., 2018; Lale et al., 2020a,b). However, these methods have so far only been applied in LTI systems with time-varying costs and disturbances. Extensions to LTV dynamics, which are the focus of this paper, are not known.

LTV Systems. As in the case of LTI systems, optimal control of LTV systems where the sequence of system parameters can be obtained by solving backwards Riccati equations (Bertsekas et al., 1995). However, in the online case when the sequence of systems is unknown, the design of controllers is challenging. There are several lines of work in adaptive control and model-predictive control (MPC) that have been studied to this point. In adaptive control of LTV systems, the underlying systems are unknown and the results generally assume slow and bounded or fixed systematic variation of dynamics with bounded disturbances (Middleton and Goodwin, 1988; Marino and Tomei, 2000; Ouyang et al., 2017). In MPC of LTV systems, a finite horizon of sequence of systems (predictions) is known and the system is again assumed to be slowly varying or open-loop stable, e.g., (Zheng and Morari, 1994; Falcone et al., 2007). Different from prior works, in the current work we consider the online problem and make no assumptions about how the system varies over time.

As in the LTI setting, the study of regret minimization in LTV systems has recently received attention. Goel and Hassibi (2020); Gradu et al. (2020) are most related to the current paper. Goel and Hassibi (2020) considers the setting where the sequence of systems is known and provides regret-optimal controller framework. Gradu et al. (2020) studies the adaptive regret of online control in LTV systems with bounded cost. Note that when the cost is bounded, a finite regret need not guarantee stability. In contrast, we use a quadratic (unbounded) cost and we can guarantee stability.

Notation. We denote the Euclidean norm of a vector x as $\|x\|$. For a matrix A, $\|A\|$ is its spectral norm, $A^\top$ is its transpose, and $\mathrm{Tr}(A)$ is its trace. $\mathcal{N}(\mu, \Sigma)$ denotes normal distribution with mean μ and covariance Σ. $A \succ B$ and $A \succeq B$ denote that $A - B$ is positive definite and positive semi-definite respectively. $A \bullet B$ denotes the element-wise inner product of A and B, *i.e.*, $\mathrm{Tr}(A^\top B)$.

2. Model & Background

We consider the following linear time-varying (LTV) system,

$$x_{t+1} = A_t x_t + B_t u_t + w_t, \tag{1}$$

where $x_t \in \mathbb{R}^d$ is the system state, $u_t \in \mathbb{R}^p$ is the control input and $w_t \in \mathbb{R}^d$ is the disturbance at time t. The system is stochastic, *i.e.*, $w_t \sim \mathcal{N}(0, W)$ for $W \succ 0$. The cost at each time-step is a quadratic function of the state and control, $x_t^\top Q x_t + u_t^\top R u_t$, where $Q, R \succ 0$.

The decision maker operates in an online setting. That is, at each time-step t, the learner observes the state x_t and system matrix (A_t, B_t) before choosing action u_t and suffering cost $x_t^\top Q x_t + u_t^\top R u_t$. We assume that the cost matrices (Q, R) are time-invariant and known to the learner. However, future system matrices $(A_{t+1}, ..., A_T)$ and $(B_{t+1}, ..., B_T)$ are unknown to the learner and are chosen by the environment, potentially stochastically or adversarially.

Stability. One of the most central goals for controller design is to ensure stability. In this work, we focus on the notion of input to state stability (ISS) and strive to design controllers that provide ISS. ISS has been the main notion of stability considered in designing stabilizing controllers both in linear and nonlinear systems (Hong et al., 2010; Sontag, 2008; Jiang and Wang, 2001). To formally define ISS, let $\mathcal{K}_\infty$ be the set of functions from nonnegative reals to nonnegative reals that are continuous, strictly increasing, and bijective. Then, ISS is defined as follows.

Definition 1 (ISS) *A LTV system with deterministic policy $\mathcal{A}$ is said to be input to state stable if there exists functions $\beta_1 : [0, \infty) \times \mathbb{N} \to [0, \infty)$ and $\beta_2 \in \mathcal{K}_\infty$ that satisfy $\beta_1(\cdot, t) \in \mathcal{K}_\infty$ for any $t \in \mathbb{N}$, $\lim_{t \to \infty} \beta_1(a, t) = 0$ for any $a \geq 0$ such that, for any disturbance sequence $\{w_t\}_{t=0}^\infty$, any initial time t_0, any initial state x_{t_0}, and any $t \geq t_0$, we have $\|x_t\| \leq \beta_1(\|x_{t_0}\|, t - t_0) + \beta_2(\sup_{t' \in \mathbb{N}} \|w_{t'}\|)$.*

Cost. In addition to stability, another important objective for controller design is maintaining a small, near-optimal control cost. Here we adopt the standard linear quadratic (LQ) cost model, *i.e.*,

$$J_T(\mathcal{A}) = \lim_{T \to \infty} \frac{1}{T} \mathbb{E}\left[\sum_{t=1}^T x_t^\top Q x_t + u_t^\top R u_t\right], \tag{2}$$

where $u_1, \ldots, u_t$ are chosen according to policy $\mathcal{A}$, and the expectation is taken with respect to the randomness of noise sequence w_t.

In this work, our goal is to ensure both stability and near-optimal cost. It should be noted that there is a trade-off between these two goals. On the one hand, a stabilizing controller without cost-awareness may produce arbitrarily large control inputs and induce high cost, which is impractical to implement. On the other hand, a greedy approach that merely focuses on cost minimization may lead to instability, as we highlight in the Section 3 below.

Though our focus is on LTV systems, our approach builds on the SDP formulation of the optimal controller for LTI systems in (Vandenberghe and Boyd, 1996).

Proposition 2 *(Vandenberghe and Boyd, 1996) When $A_t = A, B_t = B$ and (A, B) is controllable, the optimal $K^* = LQR(A, B, Q, R)$ where $u_t = K^* x_t$, can be obtained by the following SDP*

$$\min_{\Sigma \succeq 0} \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \bullet \Sigma \qquad s.t. \qquad \Sigma_{xx} = \begin{bmatrix} A_t & B_t \end{bmatrix} \Sigma \begin{bmatrix} A_t & B_t \end{bmatrix}^\top + W,$$

which has a unique symmetric solution Σ^* that decomposes to the following blocks $\Sigma^* = \begin{bmatrix} \Sigma_{xx}^* & \Sigma_{xu}^* \\ \Sigma_{xu}^{*\top} & \Sigma_{uu}^* \end{bmatrix}$, where $\Sigma_{xx}^* \in \mathbb{R}^{d \times d}$, $\Sigma_{xu}^* \in \mathbb{R}^{d \times p}$ and $\Sigma_{uu}^* \in \mathbb{R}^{p \times p}$. Then, the optimal controller is $K^* = \Sigma_{xu}^{*\top} (\Sigma_{xx}^*)^{-1}$.

The optimal LQR controller described above both stabilizes the system and achieves the minimum cost. The current paper makes a step toward understanding if it is possible to extend this formulation to the case of LTV systems.

3. A Naive Approach

How to achieve stable, cost-optimal control of LTI systems is well-known; however this is not the case in LTV systems. To illustrate the challenge of online control of LTV systems, we start by studying the performance of a naive "plug in" approach where upon receiving (A_t, B_t) an optimal controller for A_t, B_t is computed under the assumption that the system is time-invariant. Due to its

simplicity, this approach has been employed in many contexts, e.g. Li et al. (2019b) for a Markov decision process setting. In this section we provide an example which shows that such a myopic approach based on optimal LTI control described above fails to stabilize the system even in simple settings where A_t can only switch between two possible choices and B_t is fixed. This highlights that one cannot naively apply LTI design approaches in LTV systems and expect to maintain stability.

Example 1 *Consider a system with $Q = \epsilon I$, $R = I$, $w_t = 0$, and $A = \begin{bmatrix} \rho & 0 \\ a & \rho \end{bmatrix}$, $A' = \begin{bmatrix} \rho & a \\ 0 & \rho \end{bmatrix}$, where $0 < \rho < 1$, and $a > \sqrt{2}$. Suppose A_t alternates between A and A' and $B_t = B = I$. Define the optimal LTI controllers for A and A' as $K := LQR(A, B, Q, R)$ and $K' := LQR(A', B, Q, R)$.*

To show that the optimal LTI controllers will not stabilize the system, we consider a case where $\epsilon \to 0$. In this case, one can check that $K, K' \to 0$. Since A_t alternates between A, A', K_t also alternates between K and K' under the myopic design we are considering. Thus, the system state follows $x_{t+2} = (A + K)(A' + K')x_t$. Notice that as $\epsilon \to 0$, $(A + K)(A' + K') \to AA' = \begin{bmatrix} \rho^2 & a\rho \\ a\rho & a^2 + \rho^2 \end{bmatrix}$. Here, AA' is unstable since its largest eigenvalue is greater than $\frac{1}{2}\operatorname{Tr}(AA') = \rho^2 + \frac{a^2}{2} > 1$. Thus, for small enough ϵ, the naive strategy that uses the LTI controller at each time-step leads to instability.

4. Main Result

The previous section highlights that a naive application of LTI control cannot guarantee stability for LTV systems. We now propose a new approach, COvariance Constrained Online LQ (COCO-LQ) control (Section 4.1). Our main technical result shows that COCO-LQ provably guarantees stability in LTV systems (Section 4.2) when the SDP is feasible. In Section 4.3, we discuss how to handle the situation when the SDP is infeasible and Section 4.4 discuss the effect of model estimation error. Detailed proofs could be found in the Appendix of our online report Qu et al. (2021).

4.1. COvariance Constrained Online LQ (COCO-LQ)

The naive approach discussed in Section 3 seeks to solve the LTI problem at every time step, which is equivalent to solving the SDP in Proposition 2 for every (A_t, B_t). The reason this method fails is that it only considers cost minimization without explicitly considering stability. The main idea of COCO-LQ is to enforce stability via a state covariance constraint embedded into the SDP framework. The proposed algorithm is stated formally in Algorithm 1. COCO-LQ solves an SDP (3) at each time step that is similar to that in Proposition 2. The crucial difference is the new constraint (3d), which involves parameter α. Plugging (3b) into constraint (3d) yields the following:

$$\Sigma_{xx} \preceq \frac{1}{1 - \alpha} W.$$

This highlights that constraint (3d) can be interpreted as an upper bound on the state covariance matrix Σ_{xx}. When $\alpha = 0$, the controller essentially cancels out the dynamics, without taking into account the cost of doing so. This ensures stability but can lead to large cost. At another extreme, when $\alpha \to 1$, the SDP solved at each time step is the same as for the LTI setting, and so COCO-LQ matches the naive approach in Section 3. Thus, α trades off between stability and cost. In the following section, we show that this novel state covariance constraint promotes sequential strong stability (Cohen et al., 2018), which in turn guarantees ISS with a proper choice of α.

Algorithm 1: COCO-LQ: COvariance Constrained Online LQ

Parameters: $\alpha \in [0, 1)$
Input: $Q, R, W \succ 0$
1 for $t = 1, 2, ...$ **do**
2   **Receive** state x_t, and system parameter A_t, B_t
3   **Compute policy:** Let $\Sigma_t \in \mathbb{R}^{n \times n}$ be an optimal solution to the SDP program:

$$\text{minimize} \quad \begin{bmatrix} Q & 0 \\ 0 & R \end{bmatrix} \bullet \Sigma \tag{3a}$$

$$\text{subject to } \Sigma_{xx} = \begin{bmatrix} A_t & B_t \end{bmatrix} \Sigma \begin{bmatrix} A_t & B_t \end{bmatrix}^\top + W \tag{3b}$$

$$\Sigma \succeq 0 \tag{3c}$$

$$\begin{bmatrix} A_t & B_t \end{bmatrix} \Sigma \begin{bmatrix} A_t & B_t \end{bmatrix}^\top \preceq \alpha \Sigma_{xx} \tag{3d}$$

   and $K_t = \Sigma_{xu}^\top \Sigma_{xx}^{-1}$
4   **Play** $u_t = K_t x_t$
5   **Update** $x_{t+1} = A_t x_t + B_t u_t + w_t, w_t \sim N(0, W)$
6 end

4.2. Stability

We now state our main technical result, which provides a formal stability guarantee for COCO-LQ.

Theorem 3 *Let $0 \leq \alpha < 1/2$, and suppose (3) is feasible for all t, then the resulting dynamical system satisfies ISS in the sense that for any disturbance sequence $\{w_t\}_{t=0}^\infty$ and for any $t \geq t_0$,*

$$\|x_t\| \leq \rho^{t-t_0} \|x_{t_0}\| + \frac{\kappa \rho}{1-\rho} \sup_{t_0 \leq k < t} \|w_k\|$$

for $\rho = \sqrt{\frac{\alpha}{1-\alpha}} \in [0, 1)$ and $\kappa = \frac{\kappa_W}{\sqrt{1-\alpha}}$, where $\kappa_W = \|W\| \|W^{-1}\|$ is the condition number of W.

The key intuition underlying this result is that the additional state covariance constraint (3d) implicitly enforces sequential strong stability (Cohen et al., 2018), which in turn ensures ISS. More formally, sequential strong stability is defined as follows,

Definition 4 *(Sequential Strong Stability) A sequence of policies $K_1, K_2, ...,$ such that $u_t = K_t x_t$ is (κ, γ, ρ)-sequential strongly stable (for $\kappa > 0$, $0 < \gamma \leq 1$ and $0 \leq \rho < 1$) if there exist matrices $H_1, H_2, ...,$ and $L_1, L_2...,$ such that $A_t + B_t K_t = H_t L_t H_t^{-1}$ for all t, with the following properties: (a) $\|L_t\| \leq 1 - \gamma$; (b) $\|H_t\| \leq \beta_1$ and $\|H_t^{-1}\| \leq 1/\beta_2$ with $\kappa = \beta_1/\beta_2$; (c) $\|H_{t+1}^{-1} H_t\| \leq \frac{\rho}{1-\gamma}$.*

The following lemma formalizes the connection between (3d) and sequential strong stability.

Lemma 5 *Under the conditions in Theorem 3, the policies designed by COCO-LQ are (κ, γ, ρ)-sequential strongly stability for $\kappa = \frac{\kappa_W}{\sqrt{1-\alpha}}, \gamma = 1 - \sqrt{\alpha}, \rho = \sqrt{\frac{\alpha}{1-\alpha}}$ where $\kappa_W = \|W\| \|W^{-1}\|$.*

With Lemma 5, proving the result in Theorem 3 only requires showing that sequential strong stability implies ISS. The complete proof of Lemma 5 and Theorem 3 are given in Appendix A of our online report Qu et al. (2021). A critical assumption in Theorem 3 is that the SDP given in (3) is feasible for $0 \leq \alpha < 1/2$. The following result shows that when B_t is full row rank, the problem is always feasible. The proof of Lemma 6 is postponed to Appendix B in Qu et al. (2021).

Algorithm 2: COCO-LQ-Prediction: COVariance Constrained Online LQ with Predictions

---

Parameters: $\alpha \in [0, 1)$
Input: $Q, R, W \succ 0$

1 **for** $t = 1, 2, ...$ **do**
2   **if** $t \equiv 1 \pmod H$ **then**
3   **Receive** state x_t, and system parameters $(A_t, B_t), \ldots, (A_{t+H-1}, B_{t+H-1})$
4   **Compute policy:** Let $\Sigma_t \in \mathbb{R}^{n \times n}$ be a solution to the constrained SDP in (3) with

$$(R, A_t, B_t) \text{ replaced by } (\tilde{R}, \tilde{A}_t, \tilde{B}_t), \text{ where } \tilde{R} = \begin{bmatrix} R & & \\ & \ddots & \\ & & R \end{bmatrix} (H \text{ repeating blocks}),$$

$$\tilde{A}_t := A_{t+H-1} \cdots A_t, \tilde{B}_t := [B_{t+H-1}, A_{t+H-1}B_{t+H-2}, \cdots, A_{t+H-1}...A_{t+1}B_t]$$

5   Set $K_t = \Sigma_{xu}^\top \Sigma_{xx}^{-1}$ and $[u_{t+H-1}^\top, ..., u_t^\top]^\top = K_t x_t$
6   **end**
7   **Play** Implement the planned control action u_t
8   **Update** $x_{t+1} = A_t x_t + B_t u_t + w_t, w_t \sim \mathcal{N}(0, W)$
9 **end**

Lemma 6 *When B_t is full row rank, then the SDP* (3) *is always feasible.*

Note that having B_t full row rank is a sufficient but not necessary condition for feasibility of (3) of COCO-LQ. When B_t is not full row rank, the feasibility assumption may still hold, and therefore our assumption is weaker than the invertibility assumption used in the literature, e.g. Lai (1986). More broadly, in Theorem 3, $\alpha < 0.5$ is a sufficient condition for stability. For $\alpha \geq 0.5$, stability may still hold for some problem instances (A_t, B_t) as will be shown in the simulations in Section 5. How to provide a more refined instance-dependent threshold on α is an interesting future direction.

4.3. Infeasibility and the Role of Predictions

We now turn our attention to the case when the SDP given in (3) is infeasible. In this case it is necessary for the controller to use additional information in order to stabilize the system. In particular, we provide an example that shows the necessity of predictions when B_t is not full row rank in Appendix C of our online report Qu et al. (2021). This example shows that when B_t is not full row rank, for any (deterministic) online control algorithm that has causal access to system matrices, there exists a future sequence of (A_t, B_t) in which the algorithm cannot stabilize the system. In this section, we show that using (A_t, B_t) together with short-term predictions of future system matrices is enough to stabilize the system under standard controllability assumptions. Specifically, we extend COCO-LQ to include future H steps of predictions in Algorithm 2. The key idea is to rewrite the dynamics as

$$x_{t+H} = \tilde{A}_t x_t + \tilde{B}_t \bar{u}_t + [I, A_{t+H-1}, \cdots, A_{t+H-1}...A_{t+1}]\bar{w}_t \tag{4}$$

where we define $\tilde{A}_t := A_{t+H-1} \cdots A_t, \tilde{B}_t := [B_{t+H-1}, A_{t+H-1}B_{t+H-2}, \cdots, A_{t+H-1}...A_{t+1}B_t]$, $\bar{u}_t := [u_{t+H-1}^\top, u_{t+H-2}^\top, \ldots, u_t^\top]^\top$ and $\bar{w}_t := [w_{t+H-1}^\top, w_{t+H-2}^\top, \ldots, w_t^\top]^\top$. When H is long enough such that $\tilde{B}_t$ is full row rank, we can use Algorithm 1 on $\tilde{A}_t$ and $\tilde{B}_t$ and avoid the infeasibility issue, and our stability guarantee is provided below. The proof of Theorem 7 can be found in Appendix D of our online report Qu et al. (2021).

Theorem 7 *Suppose for each t, matrix $\tilde{B}_t = [B_{t+H-1}, A_{t+H-1}B_{t+H-2}, \cdots, A_{t+H-1}...A_{t+1}B_t]$ satisfies $\tilde{B}_t\tilde{B}_t^\top \succeq \sigma I$ for some $\sigma > 0$, and $\|A_t\| \leq a, \|B_t\| \leq b$ for some $a, b > 0$. Then, the SDP in Algorithm 2 is always feasible. Further, when $\alpha < 1/2$, the closed-loop system is ISS for any t,*

$$\|x_t\| \leq \kappa'_A \rho^{\frac{t}{H}-1}\|x_1\| + \kappa'_A \kappa_A \kappa \max(1, \frac{\rho}{1-\rho}) \sup_{1 \leq s < t} \|w_s\|,$$

where the same as Theorem 3, $\rho = \sqrt{\frac{\alpha}{1-\alpha}} \in [0, 1)$ and $\kappa = \frac{\kappa_W}{\sqrt{1-\alpha}}$ with $\kappa_W = \|W\|\|W^{-1}\|$ being the condition number of W; further, $\kappa_A = 1 + a + \ldots + a^{H-1}$, and $\kappa'_A = a^{H-1} + b^2(1 + a + \cdots + a^{H-1})^2 \kappa_R \frac{\kappa+a^H}{\sigma}$ with κ_R being the condition number of R.

4.4. Estimation Error

In both Algorithm 1 and Algorithm 2, the exact knowledge of state-transition matrices (A_t, B_t) or the extended state-transition matrices $(\tilde{A}_t, \tilde{B}_t)$ are needed when deriving the control actions. In this section, we show that COCO-LQ can still obtain a stabilizing controller in the case where only approximations are known, if the estimation error is controlled. Our main result is the following.

Theorem 8 *Let $(\hat{A}_t, \hat{B}_t)$ be an estimate of (A_t, B_t). Given $\alpha \in [0, \frac{1}{2})$, let $\rho = \sqrt{\frac{\alpha}{1-\alpha}}$, $\kappa = \frac{\|W\|\|W^{-1}\|}{\sqrt{1-\alpha}}$ and $\gamma = 1 - \sqrt{\alpha}$. Let $K_1, K_2, ...$ be the policies designed by COCO-LQ for $(\hat{A}_t, \hat{B}_t)$ with parameter α. When the estimation error satisfies,*

$$\max\{\|\hat{A}_t - A_t\|_2, \|\hat{B}_t - B_t\|_2\} \leq \delta \frac{\gamma}{\kappa(1 + K_{max})} \tag{5}$$

where δ can be any number in $(0, \frac{\sqrt{1-\alpha}-\sqrt{\alpha}}{1-\sqrt{\alpha}})$, and K_{max} is any uniform upper bound on $\|K_t\|$. Then, the policies K_t are ISS when applied to the system (A_t, B_t),

$$\|x_t\| \leq (\rho')^{t-t_0} \|x_{t_0}\| + \frac{\kappa\rho'}{1-\rho'} \sup_{t_0 \leq k < t} \|w_k\|,$$

where $\rho' = \frac{1-(1-\delta)\gamma}{1-\gamma}\rho \in (0, 1)$. Finally, when $\|\hat{A}_t\| \leq \bar{\sigma}_A$, $\|\hat{B}_t\| \leq \bar{\sigma}_B$ and $\hat{B}_t\hat{B}_t^\top \succeq \underline{\sigma}_B^2 I$, one uniform upper bound for $\|K_t\|$ is $K_{max} = \kappa_R \frac{\bar{\sigma}_B}{\underline{\sigma}_B^2}(\kappa(1-\gamma) + \bar{\sigma}_A)$ with $\kappa_R = \|R\|\|R^{-1}\|$.

A proof of Theorem 8 is provided in Appendix E in our online report Qu et al. (2021). This result highlights the tradeoff between the estimation error and the algorithm performance. If we choose a small α, the algorithm can tolerant a larger estimation error (i.e. larger right hand side of (5) can be obtained) but may lead to high control cost due to the tight state co-variance constraint. If we choose a larger α, the algorithm tolerates smaller estimation error while its performance improves due to the less strict state co-variance constraint.

5. Experiments

The results in the previous section focus on stability of COCO-LQ approach. Here, we use experimental results to highlight that COCO-LQ also performs near-optimally in terms of cost while also stabilizing systems that the naive approach based on LTI control cannot. In Section 5.1, we test our method on random, synthetic linear time varying systems, and in Section 5.2 we demonstrate the algorithm performance in real-world power system frequency control settings. Due to space limit, more experiment results on nonlinear systems via local linear approximation can be found in Appendix F of our online report Qu et al. (2021).

5.1. Synthetic Time-Varying Systems

We first consider the control of switching and time-variant systems. The cost function is set as $Q = 0.2I, R = I$, and system is subject to Gaussian disturbance $w_t \sim N(0, 0.1^2)$. We average the simulation results over 5 runs and visualize the mean performance and standard deviation.

a. *Switching systems.* we consider a switching system following Example 1 in Section 3, where A_t alternates between $A = [[0.99, 1.5], [0, 0.99]]$ and $A' = [[0.99, 0], [1.5, 0.99]]$, and $B_t = I$.

b. *Time-variant systems.* We consider a system $A_t = [[0.99, \sin(\frac{\pi t}{2})|e^{t/60}], [|\cos(\frac{\pi t}{2})|e^{t/60}, 0.99]]$ that is continually changing over time, and $B_t = I$.
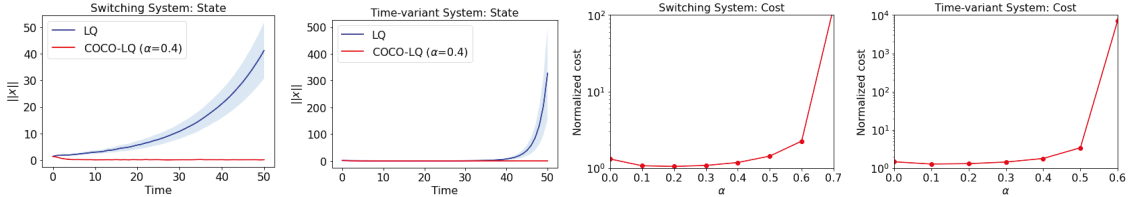

Figure 1: Performance comparison of COCO-LQ and LQ on synthetic time-varying systems. The left two figures show the state evolution, and right two figures show the normalized cost (cost of COCO-LQ divided by cost of the offline optima) under different α.

As we can see in Figure 1, COCO-LQ is able to quickly and effectively stabilize the system under various time-varying scenarios, which validates our theoretical findings. As α increases, the acquired cost of COCO-LQ first decreases and then increases (explosion of state), highlighting that α can explicitly control the tradeoff between cost and stability. With proper selection of α, COCO-LQ achieves near-optimal cost (within 30% of the offline optimal for both system a and b).

5.2. Frequency Control with Renewable Generation

We now consider a power system frequency control problem on standard IEEE WECC 3-machine 9-bus system (Figure 2(a)), which is a widely adopted system used in frequency stability studies. The state space model of power system frequency dynamics follows Hidalgo-Gonzalez et al. (2019),

$$\underbrace{\begin{bmatrix} \dot{\theta} \\ \dot{\omega} \end{bmatrix}}_{\dot{\mathbf{x}}} = \underbrace{\begin{bmatrix} 0 & I \\ -M_t^{-1}L & -M_t^{-1}D \end{bmatrix}}_{A_t} \begin{bmatrix} \theta \\ \omega \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ M_t^{-1} \end{bmatrix}}_{B_t} \underbrace{p_{in}}_{\mathbf{u_t}} \tag{6}$$

where the state variable is defined as the stacked vector of the voltage angle θ and frequency ω. $M_t = diag(m_{t,i})$ is the inertia matrix, where $m_{t,i}$ represents the equivalent rotational inertia at bus i and time t. M_t is time-varying and depends on the mix of online generators, since only thermal generators provide rotational inertia and renewable generation does not Ulbig et al. (2014). $D = diag(d_i)$ is the damping matrix, where d_i is the generator damping coefficient. L is the network susceptance matrix. The control variable p_{in} corresponds to the electric power generation.

We assume the system is changing between two states: a high renewable generation scenario where $m_{t,i} = 2$ (i.e., 80 percent renewable with zero inertia and 20 percent of thermal generation
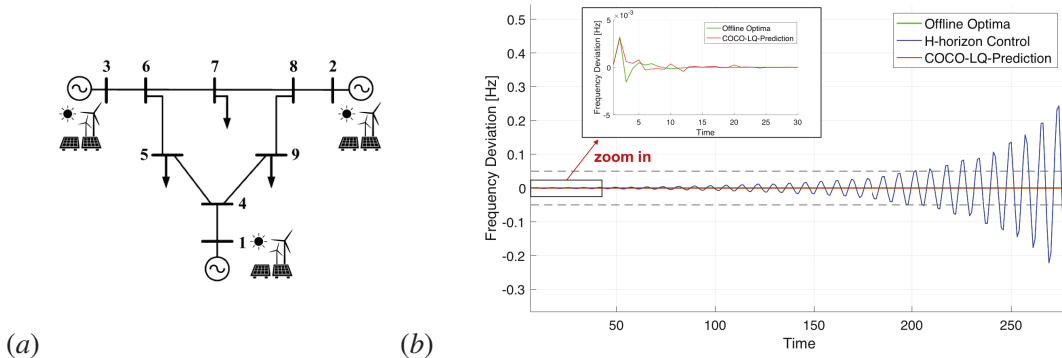
9

Figure 2: (a) IEEE WECC 3-machine 9-bus system schematic with generators at bus 1, 5, 9 are mixture of thermal generation and renewable. (b) Frequency dynamics under offline optima, baseline H-horizon control, and COCO-LQ. The dotted grey lines (± 0.05Hz) are the safety margin of power system frequency variation.

with 10s inertia), and a low renewable generation scenario where $m_{t,i} = 8$ (i.e., 20 percent renewable and 80 percent thermal generation), with additional random fluctuations between $[0, 0.2]$. This setup represents the real-world situation where we have high solar output during the daytime, and low output in the morning/evening, with intra-day variations due to clouds and weather changes. Notice that B_t is not full rank, thus we need to leverage predictions, i.e., A_{t+1} and B_{t+1}. For fair comparison, we compete against the H-horizon optimal control in Bertsekas et al. (1995), which is the extension of naive LTI controller to use H-step predictions. In both cases, we assume the prediction is accurate and use the exact value of A_{t+1} and B_{t+1} for computing control actions.

Figure 2(b) visualizes the power system frequency dynamics under three controllers: the offline optimal control, the baseline H-horizon optimal controller and the proposed COCO-LQ-Prediction method. We ideally desire a controller that is able to maintain the frequency variation within ± 0.05Hz and eventually stabilize the system. It can be observed that our algorithm succeeds at maintaining the frequency stability under random, time-varying renewable generations. Furthermore, the performance of COCO-LQ is very close to the offline optimal, while the system frequency diverges under the baseline H-horizon optimal control.

6. Conclusion

In this paper, we study the stability of LTV systems. Our results demonstrate the challenge of ensuring stability for LTV systems compared to LTI systems. Motivated by this challenge, we propose a COCO-LQ/COCO-LQ-Prediction policy that can guarantee stability for LTV systems under certain assumptions. There are many interesting open questions that remain. For example, the bound $\alpha < 1/2$ in Theorem 3 is a sufficient condition, and studying how to relax the bound and how to derive instance-dependent bounds is an interesting future question. Another important direction is to analyze the performance (e.g. the regret) of the proposed approach in order to quantify the tradeoff between stability and performance.

References

Yasin Abbasi-Yadkori and Csaba Szepesvári. Regret bounds for the adaptive control of linear quadratic systems. In *Proceedings of the 24th Annual Conference on Learning Theory*, pages 1–26, 2011.

Francesco Amato, Marco Ariola, and Carlo Cosentino. Finite-time control of discrete-time linear systems: analysis and design conditions. *Automatica*, 46(5):919–924, 2010.

Dimitri P Bertsekas, Dimitri P Bertsekas, Dimitri P Bertsekas, and Dimitri P Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena scientific Belmont, MA, 1995.

Alon Cohen, Avinatan Hassidim, Tomer Koren, Nevena Lazic, Yishay Mansour, and Kunal Talwar. Online linear quadratic control. *arXiv preprint arXiv:1806.07104*, 2018.

Sarah Dean, Horia Mania, Nikolai Matni, Benjamin Recht, and Stephen Tu. Regret bounds for robust adaptive control of the linear quadratic regulator. In *Advances in Neural Information Processing Systems*, pages 4188–4197, 2018.

Paolo Falcone, Manuela Tufo, Francesco Borrelli, Jahan Asgari, and H Eric Tseng. A linear time varying model predictive control approach to the integrated vehicle dynamics control problem in autonomous systems. In *2007 46th IEEE Conference on Decision and Control*, pages 2980–2985. IEEE, 2007.

Paolo Falcone, Francesco Borrelli, H Eric Tseng, Jahan Asgari, and Davor Hrovat. Linear time-varying model predictive control and its application to active steering systems: Stability analysis and experimental validation. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 18(8):862–875, 2008.

Germain Garcia, Sophie Tarbouriech, and Jacques Bernussou. Finite-time stabilization of linear time-varying continuous systems. *IEEE Transactions on Automatic Control*, 54(2):364–369, 2009.

Gautam Goel and Babak Hassibi. Regret-optimal control in dynamic environments. *arXiv preprint arXiv:2010.10473*, 2020.

Paula Gradu, Elad Hazan, and Edgar Minasyan. Adaptive regret for control of time-varying dynamics. *arXiv preprint arXiv:2007.04393*, 2020.

Patricia Hidalgo-Gonzalez, Rodrigo Henriquez-Auba, Duncan S Callaway, and Claire J Tomlin. Frequency regulation using data-driven controllers in power grids with variable inertia due to renewable energy. In *2019 IEEE Power & Energy Society General Meeting (PESGM)*, pages 1–5. IEEE, 2019.

Yiguang Hong, Zhong-Ping Jiang, and Gang Feng. Finite-time input-to-state stability and applications to finite-time control design. *SIAM Journal on Control and Optimization*, 48(7):4395–4418, 2010.

Zhong-Ping Jiang and Yuan Wang. Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 37(6):857–869, 2001.

Hassan K Khalil. *Nonlinear systems*, volume 3. 2002.

T.L Lai. Asymptotically efficient adaptive control in stochastic regression models. *Advances in Applied Mathematics*, 7(1):23 – 45, 1986.

Sahin Lale, Kamyar Azizzadenesheli, Babak Hassibi, and Anima Anandkumar. Adaptive control and regret minimization in linear quadratic gaussian (lqg) setting. *arXiv preprint arXiv:2003.05999*, 2020a.

Sahin Lale, Kamyar Azizzadenesheli, Babak Hassibi, and Anima Anandkumar. Explore more and improve regret in linear quadratic regulators. *arXiv preprint arXiv:2007.12291*, 2020b.

Xiaodi Li, Xueyan Yang, and Shiji Song. Lyapunov conditions for finite-time stability of time-varying time-delay systems. *Automatica*, 103:135–140, 2019a.

Yingying Li, Aoxiao Zhong, Guannan Qu, and Na Li. Online markov decision processes with time-varying transition probabilities and rewards. In *ICML Real-world Sequential Decision Making workshop*, 2019b.

Riccardo Marino and Patrizio Tomei. Robust adaptive regulation of linear time-varying systems. *IEEE Transactions on Automatic Control*, 45(7):1301–1311, 2000.

Richard H Middleton and Graham C Goodwin. Adaptive control of time-varying linear systems. *IEEE Transactions on Automatic Control*, 33(2):150–155, 1988.

Yi Ouyang, Mukul Gagrani, and Rahul Jain. Learning-based control of unknown linear systems with thompson sampling. *arXiv preprint arXiv:1709.04047*, 2017.

Guannan Qu, Yuanyuan Shi, Sahin Lale, Anima Anandkumar, and Adam Wierman. Stable online control of linear time-varying systems. *arXiv preprint arXiv:2104.14134*, 2021.

Eduardo D Sontag. Input to state stability: Basic concepts and results. In *Nonlinear and optimal control theory*, pages 163–220. Springer, 2008.

Sophie Tarbouriech, Germain Garcia, and Adolf H Glattfelder. *Advanced Strategies in Control Systems with Input and Output Constraints*, volume 346. Springer Science & Business Media, 2006.

Emanuel Todorov and Weiwei Li. A generalized iterative lqg method for locally-optimal feedback control of constrained nonlinear stochastic systems. In *Proceedings of the 2005, American Control Conference, 2005.*, pages 300–306. IEEE, 2005.

Andreas Ulbig, Theodor S Borsche, and Göran Andersson. Impact of low rotational inertia on power system stability and operation. *IFAC Proceedings Volumes*, 47(3):7290–7297, 2014.

Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM review*, 38(1):49–95, 1996.

Alex Zheng and Manfred Morari. Robust control of linear time-varying systems with constraints. In *Proceedings of 1994 American Control Conference-ACC'94*, volume 3, pages 2416–2420. IEEE, 1994.