Suboptimality Analysis of Receding-Horizon Quadratic Control With Unknown Linear Systems and Its Applications in Learning-Based Control

Shengling Shi, Anastasios Tsiamis, and Bart De Schutter

***Abstract*—This work analyzes how the tradeoff between the modeling error, the terminal value function error, and the prediction horizon affects the performance of a nominal receding-horizon linear quadratic (LQ) controller. By developing a novel perturbation result of the Riccati difference equation, a novel performance upper bound is obtained and suggests that for many cases, the prediction horizon can be either 1 or $+\infty$ to improve the control performance, depending on the relative difference between the modeling error and the terminal value function error. The result also shows that when an infinite horizon is desired, a finite prediction horizon that is larger than the controllability index can be sufficient for achieving a near-optimal performance, revealing a close relation between the prediction horizon and controllability. The obtained suboptimality performance upper bound is applied to provide novel sample complexity and regret guarantees for nominal receding-horizon LQ controllers in a learning-based setting. We show that an adaptive prediction horizon that increases as a logarithmic function of time is beneficial for regret minimization.**

***Index Terms*—Adaptive control, learning-based control, model predictive control, receding-horizon control.**

I. Introduction

Receding-horizon control (RHC), also called model predictive control (MPC) interchangeably, has been applied to various applications, such as process control, building climate control, and robotics control [1], [2]. It has important advantages, such as using optimization based on a model of the system to improve the control performance and the ability to handle constraints and multivariate systems. Due to the importance of RHC in applications, extensive efforts have been devoted to its performance analysis in [3], [4], [5], and [6], and the references therein, such as stability, constraint satisfaction, and suboptimality analysis. These analyses can help understand the behavior of RHC and motivate novel RHC approaches.

While the majority of works focus on stability and constraint satisfaction [3], this work considers the suboptimality analysis of the closed-loop control performance. We briefly illustrate our problem via a generic RHC scheme. Given a discrete-time system $x_{t+1} = f_\star(x_t, u_t)$, where $t \in \mathbb{Z}^+$, and $x_t \in \mathbb{R}^n$ and $u_t \in \mathbb{R}^m$ are the state and input, respectively, we consider the setting where $f_\star$ is unknown, but we have access to an approximate prediction model $\widehat{f}$ instead. At time step t, a generic RHC controller has the following form:

$$\min_{\{u_{k|t}\}_{k=0}^{N-1}} \sum_{k=0}^{N-1} l(x_{k|t}, u_{k|t}) + \widehat{V}(x_{N|t}) \tag{1}$$

$$\text{s.t. } x_{0|t} = x_t, x_{k+1|t} = \widehat{f}(x_{k|t}, u_{k|t}), k \in \{0, \ldots, N-1\}$$

where $N \in \{1, 2, \ldots\}$ is the prediction horizon, l is a stage cost function, and $x_{k|t}$ denotes the predicted state vector at time step $k+t$, obtained by applying the input sequence $\{u_{k|t}\}_{k=0}^{N-1}$ to $\widehat{f}$ starting from x_t. In (1), a terminal value function $\widehat{V}$ is incorporated. At time step t, only the computed $u_{0|t}$ from (1) is applied as the input u_t to the system.

In (1), a modeling error may be present, i.e., $\widehat{f}$ deviates from $f_\star$, which may result from system identification, linearization, or model reduction. While the modeling error can be explicitly considered in the robust controller design [3], the nominal controller (1) is more fundamental and commonly used in practice. The terminal value function $\widehat{V}$ avoids the short-sightedness of the finite-horizon prediction and thus improves the control performance and guarantees stability. It is standard in RHC [4], [6] and advocated in reinforcement learning (RL) to merge RL and RHC for performance improvement [7], where $\widehat{V}$ is trained offline, e.g., by the value iteration [8]. While the optimal value function $V_\star$ is ideal for achieving stability and the optimal infinite-horizon control performance simultaneously, it is typically intractable to

Received 16 December 2024; revised 10 March 2025 and 27 June 2025; accepted 28 August 2025. Date of publication 3 September 2025; date of current version 2 March 2026. This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme under Grant 101018826 - CLariNet. Recommended by Associate Editor T. Yucelen. *(Corresponding author: Shengling Shi.)*

Shengling Shi was with the Department of Chemical Engineering, Massachusetts Institute of Technology, USA. He is now with the Delft Center for Systems and Control, Delft University of Technology, The Netherlands (e-mail: s.shi-3@tudelft.nl).

Anastasios Tsiamis is with Automatic Control Laboratory, ETH Zurich, Switzerland (e-mail: atsiamis@control.ee.ethz.ch).

Bart De Schutter is with the Delft Center for Systems and Control, Delft University of Technology, The Netherlands (e-mail: b.deschutter@tudelft.nl).

Digital Object Identifier 10.1109/TAC.2025.3605539

compute it exactly [9]. Considering an approximation $\widehat{V}$ of $V_\star$ is then a practical choice.

The suboptimality analysis of (1) is an open and challenging problem, due to the interplay between N and the error sources. Particularly, when the prediction model is exact, it is well known that increasing N improves the closed-loop control performance [4], [9]. However, if a modeling error is present, increasing N also propagates the prediction error. On the other hand, having a $\widehat{V}$ closer to the optimal value function $V_\star$ can also improve the control performance. Therefore, the control performance is affected by the complex tradeoff between the error $\|\widehat{V} - V_\star\|$ of the terminal value function, the modeling error $\|\widehat{f} - f_\star\|$, and the prediction horizon N, where the norm will be defined later for the setting of this work. The challenges in performance analysis are understanding how the errors propagate over the prediction horizon, how the length of the prediction horizon affects this propagation, and how the error propagation affects the closed-loop control performance.

This work provides a novel suboptimality analysis of (1) under the joint effect of the modeling error, the prediction horizon, and the terminal value function. We consider the probably most fundamental setting, where the system is linear and the stage cost is a quadratic function, i.e., the nominal receding-horizon linear quadratic (LQ) controller [10]. As shown in our work, the analysis in this setting is already highly challenging.

Moreover, we apply our analysis to the performance analysis of learning-based RHC controllers. In this case, the model is estimated from data, leading to an inevitable modeling error. Our suboptimality analysis can incorporate this modeling error to provide performance guarantees.

Related work: When the model is exact, the suboptimality analysis of RHC controllers, with constraints or economic cost, has been studied in [4], [5], and [6], and the references therein. However, performance analysis in a setting where the system model is uncertain or unknown is rare. The suboptimality analysis of RHC for linear systems with a structured parametric uncertainty is considered in [11]; however, the impact of the approximation in the terminal value function is not investigated. Other relevant works can be found in the performance analysis of learning-based RHC [12], [13], [14], where the controller explores the state space of the unknown system and the model is recursively updated. There, a control performance metric called *regret* is concerned, which measures the cumulative performance difference over a finite time window between the controller and the ideal optimal controller. The modeling error has been investigated in the above analysis; however, the effect of the prediction horizon and the terminal value function is not considered [12], [13], [14].

As we consider the LQ setting, the performance analysis of the LQ regulator (LQR) for unknown systems is relevant, which has received renewed attention from the perspective of learning-based control [15], [16], [17]. The suboptimality analysis of LQR with a modeling error has been considered in [18], [19], and [20]. This analysis is essential for deriving performance guarantees for learning-based LQR. It typically relies on the perturbation analysis of the Riccati equation [19], [20], which characterizes the solution of the Riccati equation under a modeling error. However, since LQR concerns an infinite prediction horizon, the above analysis does not consider the effect of the prediction horizon and the terminal value function.

Contribution: In this work, the main results are developed for controllable systems and then generalized to stabilizable systems. The contributions are summarized here.

First, as an essential step for achieving performance analysis with unknown systems, the performance analysis in the simpler setting, where the model is exact, is considered. A novel performance upper bound for the receding-horizon LQ controller is obtained, which reveals a novel relation between the control performance and controllability. This is achieved by a novel convergence analysis of the Riccati difference equation, generalizing the existing results in [21]. Detailed discussions can be found in Section III-B.

Second, to achieve the suboptimality analysis with a modeling error, a novel convergence rate of the Riccati difference equation is first established when a modeling error is present. It reveals how the errors propagate over the prediction horizon within the Riccati iterations. It generalizes the perturbation results of the discrete Riccati equation [19], [20]. While the analyses in [19] and [20] address the modeling error only, the incorporation of the prediction horizon and the terminal value function error in this work leads to major technical challenges.

Third, based on the above perturbation analysis, a novel performance upper bound for the nominal receding-horizon LQ controller with an inexact model is derived. The derived *upper bound* shows how the control performance may vary with changes in the modeling error, terminal value function error, and prediction horizon. It suggests that for many cases, the prediction horizon can be either 1 or $+\infty$ to improve performance, depending on the relative difference between the modeling error and terminal value function error. When $N \to +\infty$ is desired, the result suggests that for controllable systems, a finite N that is larger than the controllability index is sufficient for achieving a near-optimal performance.

Fourth, the performance upper bound leads to a novel suboptimality bound for the nominal receding-horizon LQ controller, where the unknown system is estimated offline from data. The bound reveals how the control performance depends on the number of data samples. Moreover, a novel regret bound is derived for an adaptive RHC controller. This controller extends the state-of-the-art adaptive LQR controller with ε-greedy exploration [17], [20] from the infinite prediction horizon to an arbitrary finite prediction horizon. We show a novel regret upper bound $\tilde{O}(T\mu_\star^N + \sqrt{T})$ for a fixed N, where $\mu_\star \in (0, 1)$ is a constant and T denotes the total number of time steps for the closed-loop operation, and a regret bound $\tilde{O}(\sqrt{T})$ for an adaptive N being a logarithmic function of time.

Outline: Preliminaries are introduced in Section II. For controllable systems, the suboptimality analysis is developed in Section III for an exact model and is generalized in Section IV to incorporate the modeling error. The above results are extended in Section V to stabilizable systems and applied to learning-based control in Section VI. Finally, Section VII concludes this article. All proofs are presented in the Appendix.

Notations: Let $\mathbb{S}^n \subseteq \mathbb{R}^{n\times n}$ denote the set of symmetric matrices of dimension n, and let $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$ denote the subsets of positive semidefinite and positive-definite symmetric matrices, respectively. Given F_1, $F_2 \in \mathbb{S}^n$, $F_1 \preceq F_2$ denotes $F_2 - F_1 \in \mathbb{S}^n_+$. Moreover, $\mathbb{Z}^+$ denotes the set of nonnegative integers. Let id denote the identity function, and $f \circ g$ denotes the composition of two functions. The symbol $\|\cdot\|$ denotes the spectral norm of a matrix and the l_2 norm of a vector. Given any real sequence $\{a_k\}_{k=0}^{\infty}$, we define $\sum_{k=i}^{j} a_k \triangleq 0$ for the special case $j < i$. Given a real matrix F and a real square matrix A, $\overline{\sigma}(F)$ and $\underline{\sigma}(F)$ denote the maximum and the minimum singular values, respectively, and $\rho(A)$ denotes the spectral radius. Note that $\overline{\sigma}(F) = \|F\|$. The symbol $\mathcal{U}_{[a,b]}$ denotes the uniform distribution over the interval $[a, b]$, $\mathbb{E}$ denotes the expected value, and $\lfloor\cdot\rfloor$ denotes the floor function. Given ε, $t \geqslant 0$ and two nonnegative functions $f(\cdot,\cdot)$ and $g(\cdot,\cdot)$, we write $g(\varepsilon, t) = O(f(\varepsilon, t))$ (as $\varepsilon \to 0$, $t \to \infty$) if $\exists\, c_1, c_2 > 0 : g(\varepsilon, t) \leqslant c_1 f(\varepsilon, t)$ for any $1/\varepsilon \geqslant c_2$ and $t \geqslant c_2$. If g and f depend on t only, then $g(t) = \tilde{O}(f(t))$ (as $t \to \infty$) indicates $g(t) = O(f(t) \log^k(t))$ for some $k \in \mathbb{Z}^+$ [22].

II. Preliminaries and Problem Formulation

A. Preliminaries

We consider the discrete-time linear system

$$x_{t+1} = A_\star x_t + B_\star u_t + w_t \tag{2}$$

where w_t is a white noise signal with $\mathbb{E}(w_t w_t^\top) = \sigma_w^2 I, \sigma_w > 0$, $A_\star \in \mathbb{R}^{n\times n}$, and $B_\star \in \mathbb{R}^{n\times m}$. The initial state x_0 is a zero-mean random vector with $\mathbb{E}(x_0 x_0^\top) \in \mathbb{S}^n_+$, uncorrelated with the noise. In this work, the true system $(A_\star, B_\star)$ is unknown, and we have access to an approximate model $(\hat{A}, \hat{B})$ that satisfies $\|\hat{A} - A_\star\| \leqslant \varepsilon_{\mathrm{m}}$ and $\|\hat{B} - B_\star\| \leqslant \varepsilon_{\mathrm{m}}$ for some $\varepsilon_{\mathrm{m}} \geqslant 0$. This approximate model and its error bound can be obtained, e.g., from linear system identification [23], [24].

In the above setting, we consider the nominal receding-horizon LQ controller [10]. At time step t, x_t is measured, and the RHC controller solves the following optimization problem:

$$\min_{\{u_{k|t}\}_{k=0}^{N-1}} \mathbb{E}_{\{w_{k+t}\}_{k=0}^{N-1}} \left[\sum_{k=0}^{N-1} l(x_{k|t}, u_{k|t}) + x_{N|t}^\top P x_{N|t}\right] \tag{3a}$$

$$\text{s.t. } x_{k+1|t} = \hat{A} x_{k|t} + \hat{B} u_{k|t} + w_{k+t}, x_{0|t} = x_t \tag{3b}$$

where $l(x_{k|t}, u_{k|t}) = x_{k|t}^\top Q x_{k|t} + u_{k|t}^\top R u_{k|t}$, $R \in \mathbb{S}^m_{++}$, and Q, $P \in \mathbb{S}^n_+$. Then, $u_t = u_{0|t}$ is the input at time step t.

We will first consider controllable systems. The extension to stabilizable systems is in Section V. To this end, we define $\mathcal{C}_i(A_\star, B_\star) \triangleq \begin{bmatrix} B_\star & A_\star B_\star & \dots & A_\star^{i-1} B_\star \end{bmatrix}$ for $i \in \{1, 2, \dots\}$ and introduce the following assumptions.

Assumption 1: There exists $i \in \{1, 2, \dots, n\}$ such that $\mathcal{C}_i(A_\star, B_\star)$ has full rank, and the minimum i is denoted by n_{cr}, called the *controllability index*.

Assumption 2: $A_\star \neq 0$, $P \succeq 0$, $R \succeq I$, and $Q \succeq I$ hold.[1]

For some intermediate technical results, we will relax the controllability requirement whenever possible for generality.

Assumption 3: System $(A_\star, B_\star)$ is stabilizable.

[1]Requiring Q to be positive definite leads to a loss of generality. This assumption, however, is commonly imposed to obtain quantitative bounds [20]. Given $Q \succ 0$, $R \succeq I$, and $Q \succeq I$ can be achieved without losing generality by scaling them simultaneously. Letting $A_\star \neq 0$ is not restrictive and ensures that $1 - \beta_\star$ is invertible for technical simplicity.

This RHC controller is further characterized by the Riccati equation. Given any $(A, B) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$, we define

$$S_B \triangleq B R^{-1} B^\top \tag{4}$$

and the Riccati mapping $\mathcal{R}_{A,B} : \mathbb{S}^n_+ \to \mathbb{S}^n_+$ as

$$\begin{aligned}\mathcal{R}_{A,B}(P) &\triangleq A^\top P (I + S_B P)^{-1} A + Q \\ &= A^\top P A - A^\top P B (R + B^\top P B)^{-1} B^\top P A + Q\end{aligned} \tag{5}$$

where the last equality holds by the Woodbury matrix identity. To apply the Riccati mapping recursively, we define the Riccati iteration $\mathcal{R}_{A,B}^{(i)}$ via recursion: for any integer $i \geqslant 0$, $\mathcal{R}_{A,B}^{(i+1)} \triangleq \mathcal{R}_{A,B} \circ \mathcal{R}_{A,B}^{(i)}$, and $\mathcal{R}_{A,B}^{(0)} = \mathrm{id}$. Then, the RHC controller in (3) is equivalent to a linear static controller $u_t = -K_{\mathrm{RHC}} x_t$ [10], where

$$K_{\mathrm{RHC}} = \left[R + \hat{B}^\top \left[\mathcal{R}_{\hat{A},\hat{B}}^{(N-1)}(P)\right] \hat{B}\right]^{-1} \hat{B}^\top \left[\mathcal{R}_{\hat{A},\hat{B}}^{(N-1)}(P)\right] \hat{A}. \tag{6}$$

We consider the expected average infinite-horizon performance of K_{RHC}, applied to the true system (2). Given any controller gain K with $\rho(A_\star - B_\star K) < 1$, we denote its expected average infinite-horizon performance as

$$J_K \triangleq \lim_{T\to\infty} \frac{1}{T} \mathbb{E}_{x_0, \{w_t\}} \left[\sum_{t=0}^{T-1} \left(x_t^\top Q x_t + u_t^\top R u_t\right)\right]$$

$$\text{s.t. } (2), u_t = -K x_t, t \in \mathbb{Z}^+.$$

Under Assumption 2, the optimal controller gain $K_\star$, which achieves the minimum J_K, is the LQR controller gain

$$K_\star = (R + B_\star^\top P_\star B_\star)^{-1} B_\star^\top P_\star A_\star \tag{7}$$

where $P_\star$ is the unique positive-definite solution of the discrete-time Riccati equation [25]

$$P_\star = \mathcal{R}_{A_\star, B_\star}(P_\star). \tag{8}$$

We further define the closed-loop matrix $L_\star$ under $K_\star$ as

$$L_\star \triangleq A_\star - B_\star K_\star = (I + S_{B_\star} P_\star)^{-1} A_\star. \tag{9}$$

It is well known that $L_\star$ is Schur stable, and its decay rate can be characterized by the standard Lyapunov analysis.

Lemma 1: If Assumptions 2 and 3 hold, for any $i \in \mathbb{Z}^+$

$$\|L_\star^i\| \leqslant \sqrt{\beta_\star^{-1}} \left(\sqrt{1 - \beta_\star}\right)^i \tag{10}$$

with $\beta_\star \triangleq \underline{\sigma}(Q)/\overline{\sigma}(P_\star)$, and $\beta_\star \in (0, 1)$ holds.

The proof of this result is presented in the Appendix for completeness. If Assumption 1 holds additionally, then $(A_\star, B_\star)$ is controllable, which implies that $(L_\star, B_\star)$ is controllable, and thus there exists a real number $\nu_\star$ such that

$$\underline{\sigma}\left(\mathcal{C}_{n_{\mathrm{cr}}}(L_\star, B_\star)\right) \geqslant \nu_\star > 0. \tag{11}$$

B. Problem Formulation

In this work, we analyze the performance gap $J_{K_{\mathrm{RHC}}} - J_{K_\star}$ between the RHC controller and the ideal LQR controller.

Problem 1: Given a prediction horizon $N \geqslant 1$, an approximate model $(\hat{A}, \hat{B})$ with $\max\{\|\hat{A} - A_\star\|, \|\hat{B} - B_\star\|\} \leqslant \varepsilon_{\mathrm{m}}$, and a terminal matrix P with $\|P - P_\star\| \leqslant \varepsilon_{\mathrm{p}}$ for some $\varepsilon_{\mathrm{p}} \geqslant 0$, find a nontrivial error bound $g(\varepsilon_{\mathrm{m}}, \varepsilon_{\mathrm{p}}, N)$ such that $J_{K_{\mathrm{RHC}}} - J_{K_\star} \leqslant g(\varepsilon_{\mathrm{m}}, \varepsilon_{\mathrm{p}}, N)$.

In Problem 1, a nonzero g reflects the difference between the performance of the RHC controller, i.e., $J_{K_{\text{RHC}}}$, and the performance of the baseline controller, i.e., $J_{K_\star}$. A nonzero g reveals how the system performance, under the RHC controller, degrades as a function of the errors ε_{m} and ε_{p}, and the prediction horizon N. Here, choosing the infinite-horizon optimal controller as the baseline is standard in studies on learning-based control [19], [20], and MPC [4].

In the special case of a known system and $P = P_\star$, we have $K_{\text{RHC}} = K_\star$ and $J_{K_{\text{RHC}}} - J_{K_\star} = 0$ for any $N \geqslant 1$. However, if $P \neq P_\star$, which happens when the system is unknown or a numerical error is present for the computed $P_\star$ even if the system is known, we can increase N to achieve a smaller $J_{K_{\text{RHC}}} - J_{K_\star}$. In the more complex situation where there is a modeling error, a larger N leads to the propagation of the modeling error and thus may enlarge the performance gap. Our target is to characterize the above complex tradeoff among N, ε_{m}, and ε_{p}.

In this work, the obtained bounds hold regardless of whether ε_{m} and ε_{p} are coupled. The presence of coupling, e.g., $\varepsilon_{\text{p}} = h(\varepsilon_{\text{m}})$ for some h, can be incorporated by plugging function h into g. The error bounds obtained will also depend on the system matrices $A_\star$ and $B_\star$. To simplify the algebraic expressions of the bounds, we upper bound the system matrices as $\max\{\|A_\star\|, \|B_\star\|, \|P_\star\|\} \leqslant \Upsilon_\star$, where $\Upsilon_\star \geqslant \max\{1, \varepsilon_{\text{m}}, \varepsilon_{\text{p}}\}$ is a real constant. Note that there always exists a sufficiently large $\Upsilon_\star$ such that the above holds.

Remark 1: Regarding the bound g in Problem 1, its rate of change as $N, \varepsilon_{\text{m}}$, and ε_{p} change will be the primary interest. The other constants involved in the bound will often be simplified for interpretability. Therefore, the derived bound is more suitable for convergence analysis and qualitative insights instead of being used as a practical error bound.

Remark 2: g is a worst-case bound valid for any $(\hat{A}, \hat{B}) \in \mathcal{S}(\varepsilon_{\text{m}}) \triangleq \{(A, B) \mid \max\{\|A - A_\star\|, \|B - B_\star\|\} \leqslant \varepsilon_{\text{m}}\}$. Therefore, it can be conservative for a particular $(\hat{A}, \hat{B})$ instance.

III. Suboptimality Analysis With Known Model

As the first step, we consider the simpler setting where $(\hat{A}, \hat{B}) = (A_\star, B_\star)$. The goal is to analyze the joint effect of the prediction horizon and the approximation in the terminal value function on the control performance. To obtain the final performance upper bound in Section III-C, we first develop two technical results in Sections III-A and III-B.

Case 1: It holds that $(\hat{A}, \hat{B}) = (A_\star, B_\star)$.

A. Preparatory Performance Analysis

We first characterize the performance gap between the LQR controller and a general linear controller that has a similar structure to the RHC controller (6). Given any $(A, B) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times m}$, we define function $\mathcal{K}_{A,B} : \mathbb{S}^n_+ \to \mathbb{R}^{m\times n}$ as

$$\mathcal{K}_{A,B}(F) \triangleq (R + B^\top F B)^{-1} B^\top F A. \tag{12}$$

Then, given the controller gain $\mathcal{K}_{A_\star,B_\star}(F)$ for any $F \in \mathbb{S}^n_+$, the following result characterizes its performance gap based on the difference $\|F - P_\star\|$.

Lemma 2: Given any $F \in \mathbb{S}^n_+$ with $\|F - P_\star\| \leqslant \varepsilon$ for some $\varepsilon \geqslant 0$, consider the controller gain $K = \mathcal{K}_{A_\star,B_\star}(F)$ and the optimal gain $K_\star$. In Case 1, if Assumptions 2 and 3 hold, and if $\varepsilon \leqslant \min\{\Upsilon_\star, \underline{\sigma}(R)/(40\Upsilon_\star^4\|P_\star\|^{3/2})\}$, then $A_\star - B_\star K$ is Schur stable and $J_K - J_{K_\star} \leqslant c_\star \varepsilon^2$, where

$$c_\star \triangleq \frac{32\min\{n, m\}\sigma_w^2(\overline{\sigma}(R) + \Upsilon_\star^3)\Upsilon_\star^4\|P_\star\|(\sqrt{\|P_\star\|} + 1)^2}{\underline{\sigma}^2(R)}. \tag{13}$$

Lemma 2 directly follows from Lemma 4, which is presented later, as a special case of a known model. These results extend [19, Thm. 1] and [20, Prop. 7] by incorporating the terminal value function error. Lemma 2 shows that if F is sufficiently close to $P_\star$, then the controller gain $\mathcal{K}_{A_\star,B_\star}(F)$ can stabilize the system with a near-optimal performance.

In Case 1, since (6) is a special case of (12) with $F = \mathcal{R}^{(N-1)}_{A_\star,B_\star}(P)$, we will exploit Lemma 2 to characterize the performance gap of K_{RHC}. Particularly, Lemma 2 can be used to bound $J_{K_{\text{RHC}}} - J_{K_\star}$ if we can quantify the difference $\|\mathcal{R}^{(N-1)}_{A_\star,B_\star}(P) - P_\star\|$. Note that this difference captures the convergence of $\mathcal{R}^{(N-1)}_{A_\star,B_\star}(P)$ to $P_\star$ as N increases.

B. Convergence Rate of Riccati Iterations

It is a classical result that $\mathcal{R}^{(i)}_{A_\star,B_\star}(P)$ converges to $P_\star$ exponentially as i increases [26]. Different convergence analyses have been investigated in the literature, leading to different convergence rates. The rates in [27] and [28, Thm. 14.4.1] are complex functions of the system parameters, characterized by the Thompson metric in [27] and a norm weighted by a special matrix in [28, Thm. 14.4.1]. These convergence rates are hard to interpret because of the specific metric and norm. On the other hand, the rates in [21] and [26, Ch. 4.4] utilize the standard matrix 2-norm and thus are easy to interpret; however, [26, Ch. 4.4] does not exploit controllability, leading to a slower convergence rate than [21]. To the best of the authors' knowledge, Moral and Horton [21] provided the most recent state-of-the-art analysis of convergence rates for controllable systems. Therefore, we develop our results based on [21]. While the result in [21] considers only a single convergence rate, we generalize it to incorporate two distinct convergence rates depending on the range of i. We also compute the rates as a function of $\beta_\star$ and thus of $P_\star$. The explicit dependence of the convergence rates on $P_\star$ will facilitate the subsequent perturbation analysis in Section IV, where the effect of the modeling error on $P_\star$ will be exploited.

We define several new notations. Recall S_B defined in (4), and we define the function $\mathcal{L}_{A,B} : \mathbb{S}^n_+ \to \mathbb{R}^{n\times n}$ as

$$\mathcal{L}_{A,B}(P) \triangleq (I + S_B P)^{-1} A = A - B\mathcal{K}_{A,B}(P) \tag{14}$$

which denotes the closed-loop matrix under the controller gain $\mathcal{K}_{A,B}(P)$, and thus, $\mathcal{L}_{A_\star,B_\star}(P_\star) = L_\star$. In addition, for any integers $0 \leqslant j \leqslant i$, define

$$\Phi^{(j:i)}_{A,B}(P) \triangleq \begin{cases} \mathcal{L}_{A,B}\left(P^{(j)}\right) \ldots \mathcal{L}_{A,B}\left(P^{(i-1)}\right), & i > j \geqslant 0 \\ I, & i = j \end{cases} \tag{15}$$

where $P^{(i)}$ is a shorthand notation for $\mathcal{R}^{(i)}_{A,B}(P)$. The definition shows $\Phi^{(0:i)}_{A_\star,B_\star}(P_\star) = L^i_\star$. Note that $\Phi^{(0:i)}_{A,B}(P)$ is a state-transition matrix of a time-varying linear system, whose asymptotic stability is reflected by the convergence of $\Phi^{(0:i)}_{A,B}(P)$ to 0 as $i \to \infty$ [29].

The state-transition matrix is helpful for characterizing $\|\mathcal{R}^{(N-1)}_{A_\star,B_\star}(P) - P_\star\|$ for any $P \in \mathbb{S}^n_+$ [21]

$$\mathcal{R}^{(i)}_{A_\star,B_\star}(P) - P_\star = \left[\Phi^{(0:i)}_{A_\star,B_\star}(P)\right]^\top (P - P_\star)L^i_\star \tag{16}$$

which can be verified by applying Lemma 11(a) recursively. Recall that $L_\star$ is Schur stable, satisfying $\lim_{i\to\infty} L^i_\star = 0$. Therefore, (16) shows that as $i \to \infty$, the convergence of $\mathcal{R}^{(i)}_{A_\star,B_\star}(P)$ to $P_\star$ is governed by the convergence of $L^i_\star$. Furthermore, exploiting the exponential decay of $\Phi^{(0:i)}_{A_\star,B_\star}(P)$ can further contribute to the convergence of $\mathcal{R}^{(i)}_{A_\star,B_\star}(P)$.

The convergence rate of $\mathcal{R}^{(i)}_{A_\star,B_\star}(P)$ to $P_\star$ is as follows.

Lemma 3: Given P with $\|P - P_\star\| \leqslant \varepsilon_{\mathrm{p}}$, if Assumptions 1 and 2 are satisfied, then for $i \in \mathbb{Z}^+$

$$\|\mathcal{R}^{(i)}_{A_\star,B_\star}(P) - P_\star\| \leqslant \begin{cases} \overline{\tau}_\star \left[\sqrt{(1-\beta_\star)\left(1-\overline{\beta}(\varepsilon_{\mathrm{p}})\right)}\right]^i \varepsilon_{\mathrm{p}}, & \text{if } i < n_{\mathrm{cr}} \\ \tau_\star (1-\beta_\star)^i \varepsilon_{\mathrm{p}}, & \text{if } i \geqslant n_{\mathrm{cr}} \end{cases} \tag{17}$$

with constants $\tau_\star$ and $\overline{\tau}_\star$ defined in (41) and (43), respectively, $\beta_\star$ defined in Lemma 1, and

$$\overline{\beta}(\varepsilon_{\mathrm{p}}) \triangleq \underline{\sigma}(Q)/\left(\overline{\sigma}(P_\star) + \varepsilon_{\mathrm{p}}\beta_\star^{-1}\right) \tag{18}$$

where $\overline{\beta}(\varepsilon_{\mathrm{p}}) \leqslant \beta_\star$ holds for any $\varepsilon_{\mathrm{p}} \geqslant 0$.

Since $\overline{\beta} \leqslant \beta_\star$, we have $1 - \beta_\star \leqslant \sqrt{(1-\beta_\star)(1-\overline{\beta})}$. Therefore, as i increases, Lemma 3 shows a slower decay rate[2] $\sqrt{(1-\beta_\star)(1-\overline{\beta})}$ for $i < n_{\mathrm{cr}}$ and a faster rate $1-\beta_\star$ for $i \geqslant n_{\mathrm{cr}}$. Recall $\Phi^{(0:i)}_{A_\star,B_\star}(P)$ in (15) and (16). Intuitively, the case $i < n_{\mathrm{cr}}$ is when the state-transition matrix $\Phi^{(0:i)}_{A_\star,B_\star}(P)$ is affected by the initial matrix P and the consequent transient behaviors, which is evident from the dependence of $\overline{\beta}$ on ε_{p}. Then, if $i \geqslant n_{\mathrm{cr}}$, then the time-varying closed-loop systems in (15) get closer to $L_\star$, leading to a faster convergence rate.

The technical challenge in developing Lemma 3 is the derivation of the two distinct decay rates. Given controllability, the faster rate for $i \geqslant n_{\mathrm{cr}}$ is established from the full rank property of $\mathcal{C}_i(A_\star, B_\star)$ by exploiting Lemma 1 and [21, Lemma 4.1]. Since $\mathcal{C}_i(A_\star, B_\star)$ is not of full rank for $i < n_{\mathrm{cr}}$, the slower rate is established from stabilizability. The analysis is achieved by utilizing the stability analysis of linear time-varying systems [29], introduced in Lemma 13, and further deriving the decay rate as an explicit function of $\beta_\star$. Note that if $\varepsilon_{\mathrm{p}} = 0$, then we obtain the rate $(1-\beta_\star)$ in both cases of (17), recovering the square of the rate $\sqrt{1-\beta_\star}$ in (10).

[2]Given $c_1 a^i$ and $c_2 b^i$ with $0 \leqslant a \leqslant b < 1$ and $c_1, c_2 \geqslant 0$, $i \in \mathbb{Z}^+$, we say a is a faster decay rate than b to indicate a^i decays faster than b^i as i increases.

C. Performance Upper Bound With a Known Model

The following main result is a direct consequence of (6) and Lemmas 2 and 3.

Theorem 1: In Case 1, consider K_{RHC} in (6) with any $P \in \mathbb{S}^n_+$ satisfying $\|P - P_\star\| \leqslant \varepsilon_{\mathrm{p}}$. It holds that $A_\star - B_\star K_{\mathrm{RHC}}$ is Schur stable with $J_{K_{\mathrm{RHC}}} - J_{K_\star} \leqslant v(N, \varepsilon_{\mathrm{p}})$, where

$$v(N,\varepsilon_{\mathrm{p}}) \triangleq \begin{cases} c_\star \overline{\tau}_\star^2 \left[(1-\beta_\star)\left(1-\overline{\beta}(\varepsilon_{\mathrm{p}})\right)\right]^{N-1} \varepsilon_{\mathrm{p}}^2, & \text{if } N \leqslant n_{\mathrm{cr}} \\ c_\star \tau_\star^2 (1-\beta_\star)^{2(N-1)} \varepsilon_{\mathrm{p}}^2, & \text{if } N \geqslant n_{\mathrm{cr}}+1 \end{cases} \tag{19}$$

if Assumptions 1 and 2 hold, and if N and ε_{p} satisfy $\sqrt{v(N,\varepsilon_{\mathrm{p}})/c_\star} \leqslant \min\{\Upsilon_\star, \underline{\sigma}(R)/(40\Upsilon_\star^4\|P_\star\|^{3/2})\}$.

In Theorem 1, the upper bound on $\sqrt{v(N,\varepsilon_{\mathrm{p}})/c_\star}$ means that N should be sufficiently large or ε_{p} should be sufficiently small, such that the RHC controller can stabilize the system with a suboptimal performance. Moreover, the performance gap (19) captures the tradeoff between the error ε_{p} and the prediction horizon N: the performance gap decays quadratically in ε_{p} and exponentially in N. The exponential decay obtained here is analogous to the exponential decay rate in [30] for linear RHC with estimated additive disturbances over the prediction horizon and in [6] for nonlinear RHC, while we have a novel characterization with two distinct decay rates.

The case $N \geqslant n_{\mathrm{cr}} + 1$ in (19) has a faster decay rate than the rate for $N \leqslant n_{\mathrm{cr}}$, which indicates the advantage of choosing a horizon larger than the controllability index. This also suggests that we may choose a smaller horizon for controlling fully actuated systems, with $\mathrm{rank}(B) = n$ and $n_{\mathrm{cr}} = 1$, compared to underactuated systems.

IV. Suboptimality Analysis With Unknown Model

In this section, the results in Section III are generalized to consider the joint performance effect of the modeling error, terminal matrix error, and prediction horizon. To obtain the final performance upper bound in Section IV-C, we first develop several technical results in Sections IV-A and IV-B.

A. Preparatory Performance Analysis

Given any $F \in \mathbb{S}^n_+$ and a general controller gain $\mathcal{K}_{\hat{A},\hat{B}}(F)$ in (12), formulated on the approximate model, the following result characterizes the performance gap between $\mathcal{K}_{\hat{A},\hat{B}}(F)$ and the optimal controller gain $K_\star$.

Lemma 4: Given any $F \in \mathbb{S}^n_+$ with $\|F - P_\star\| \leqslant \varepsilon$ for some $\varepsilon \geqslant 0$, consider the controller gain $K = \mathcal{K}_{\hat{A},\hat{B}}(F)$. Then, $A_\star - B_\star K$ is Schur stable and $J_K - J_{K_\star} \leqslant c_\star(\varepsilon_{\mathrm{m}} + \varepsilon)^2$ with $c_\star$ defined in (13), if Assumptions 2 and 3 hold, and if $\Upsilon_\star \geqslant \varepsilon$ and $8\Upsilon_\star^4(\varepsilon_{\mathrm{m}} + \varepsilon)/\underline{\sigma}(R) \leqslant 1/5\|P_\star\|^{-3/2}$.

The above result is analogous to [19, Thm. 1] and [20, Prop. 7], with extensions to incorporate the additional error term $\|F - P_\star\|$. Lemma 4 shows that if the error $\varepsilon_{\mathrm{m}} + \varepsilon$ is sufficiently small, then the resulting controller gain $K = \mathcal{K}_{\hat{A},\hat{B}}(F)$ can stabilize the real system $(A_\star, B_\star)$ with a suboptimal performance gap.

B. Perturbation Analysis of Riccati Difference Equation

As the RHC controller (6) is a special case of $\mathcal{K}_{\hat{A},\hat{B}}(F)$ with $F = \mathcal{R}^{(N-1)}_{\hat{A},\hat{B}}(P)$, Lemma 4 can be exploited to analyze the suboptimality of the RHC controller. Particularly, Lemma 4 shows that $J_{K_{\text{RHC}}} - J_{K_\star}$ can be upper bounded by a function of the modeling error bound ε_{m} and ε satisfying

$$\left\|\mathcal{R}^{(N-1)}_{\hat{A},\hat{B}}(P) - P_\star\right\| \leqslant \varepsilon. \tag{20}$$

Therefore, the main challenge is to characterize ε by investigating the Riccati iterations using the approximate model.

The above problem has been studied in the classical perturbation analysis of the Riccati difference equation [31]; however, the final bound in [31] does not explicitly show the dependence on N, $P_\star$, and ε_{p}. In this work, we provide a novel upper bound as in (20) that captures this dependence.

To this end, due to the approximate model, we first define a state-transition matrix similar to (15)

$$\bar{\Phi}^{(j:i)}(P) \triangleq \begin{cases} \bar{\mathcal{L}}\left(\mathcal{R}^{(j)}_{A_\star,B_\star}(P)\right)\cdots\bar{\mathcal{L}}\left(\mathcal{R}^{(i-1)}_{A_\star,B_\star}(P)\right), & i > j \geqslant 0 \\ I, & i = j \end{cases} \tag{21}$$

where $\bar{\mathcal{L}}(P) \triangleq \mathcal{W}(P)[\mathcal{H}(P) + \mathcal{L}_{A_\star,B_\star}(P)]$ with the functions $\mathcal{W}$ and $\mathcal{H}$ defined on $\mathbb{S}^n_+$

$$\mathcal{W}(P) \triangleq I + (S_{B_\star} - S_{\hat{B}})P, \mathcal{H}(P) \triangleq (I + S_{B_\star}P)^{-1}(\hat{A} - A_\star).$$

Recall the closed-loop matrix $\mathcal{L}_{A_\star,B_\star}(P)$ in (14), and thus, we can interpret $\bar{\mathcal{L}}(P)$ as a perturbed $\mathcal{L}_{A_\star,B_\star}(P)$: if there is no modeling error, then we have $\bar{\mathcal{L}}(P) = \mathcal{L}_{A_\star,B_\star}(P)$ due to $\mathcal{H}(P) = 0$ and $\mathcal{W}(P) = I$. Therefore, $\bar{\Phi}^{(j:i)}(P)$ can also be interpreted as a perturbed version of the state-transition matrix $\Phi^{(j:i)}_{A_\star,B_\star}(P)$ due to the approximate model.

Then, as an extension of (16) to the approximate model, the following lemma can be obtained.

Lemma 5: For any $P_1, P_2 \in \mathbb{S}^n_+$ and $i \in \mathbb{Z}^+$, define $\hat{P}^{(i)} \triangleq \mathcal{R}^{(i)}_{\hat{A},\hat{B}}(P_1)$ and $P^{(i)}_\star \triangleq \mathcal{R}^{(i)}_{A_\star,B_\star}(P_2)$. It holds that

$$\hat{P}^{(i)} - P^{(i)}_\star = \left[\Phi^{(0:i)}_{\hat{A},\hat{B}}(P_1)\right]^\top (P_1 - P_2)\,\bar{\Phi}^{(0:i)}(P_2) \tag{22a}$$

$$+ \sum_{j=1}^{i}\left[\Phi^{(i-j+1:i)}_{\hat{A},\hat{B}}(P_1)\right]^\top \mathcal{M}\left(\hat{P}^{(i-j)}, P^{(i-j)}_\star\right)\bar{\Phi}^{(i-j+1:i)}(P_2) \tag{22b}$$

where $\mathcal{M}$ is a function on $\mathbb{S}^n_+ \times \mathbb{S}^n_+$: $\mathcal{M}(P_1, P_2) \triangleq \hat{A}^\top P_2(I + S_{B_\star}P_2)^{-1}\hat{A} - A_\star^\top P_2(I + S_{B_\star}P_2)^{-1}A_\star + \hat{A}^\top(I + P_1 S_{\hat{B}})^{-1}P_2(S_{B_\star} - S_{\hat{B}})P_2(I + S_{B_\star}P_2)^{-1}\hat{A}$.

Equation (22) is closely related to (16). While (16) concerns the state-transition matrix $\Phi_{A_\star,B_\star}$ of the true system, (22a) contains the perturbed state-transition matrix $\bar{\Phi}$ and the state-transition matrix $\Phi_{\hat{A},\hat{B}}$ of the approximate model; moreover, (22b) contains an extra term caused by the modeling error.

To upper bound $\|\mathcal{R}^{(N-1)}_{\hat{A},\hat{B}}(P) - P_\star\|$, we exploit (22) with $P_1 = P$ and $P_2 = P_\star$. This motivates us to analyze the spectral norm of the perturbed state-transition matrices in (22). The analysis of these state-transition matrices is analogous to the stability analysis of the corresponding time-varying systems.

Then, we establish the following results for the two perturbed state-transition matrices.

Lemma 6: Given any two integers $i \geqslant j \geqslant 0$ and $\bar{\Phi}^{(j:i)}$ defined in (21), if Assumptions 2 and 3 hold, then

$$\|\bar{\Phi}^{(j:i)}(P_\star)\| \leqslant \sqrt{\beta_\star^{-1}}[\gamma_1(\varepsilon_{\text{m}})]^{i-j} \tag{23}$$

$$\gamma_1(\varepsilon_{\text{m}}) \triangleq \sqrt{\beta_\star^{-1}}\psi(\varepsilon_{\text{m}}) + \sqrt{1 - \beta_\star} \tag{24}$$

where ψ is defined in (50) and satisfies $\psi(\varepsilon_{\text{m}}) = O(\varepsilon_{\text{m}})$.

Comparing (10) and (23), $\bar{\Phi}^{i:j}(P_\star)$ admits a perturbed decay rate γ_1. In γ_1, $\sqrt{1-\beta_\star}$ from (10) is perturbed by the term $\sqrt{\beta_\star^{-1}}\psi(\varepsilon_{\text{m}})$.

To analyze $\Phi^{(0:i)}_{\hat{A},\hat{B}}$ in (22), we limit the modeling error by introducing the following technical assumptions.

Assumption 4: The modeling error ε_{m} is sufficiently small such that the following holds:

a) $\varepsilon_{\text{m}} \leqslant 1/(16\|P_\star\|^2)$;

b) $f_{\mathcal{C}}(\varepsilon_{\text{m}}) \leqslant \nu_\star/2$, where $\nu_\star$ is defined in (11), and $f_{\mathcal{C}}$ is defined in (61) and satisfies $f_{\mathcal{C}}(\varepsilon_{\text{m}}) = O(\varepsilon_{\text{m}})$;

c) $\varepsilon_{\text{m}} \leqslant \alpha_\star$ with the constant $\alpha_\star$ defined in (64).

Assumption 5: The true system $(A_\star, B_\star, Q, R)$ satisfies $\underline{\sigma}(P_{\text{dual}}) \geqslant 1$, where $P_{\text{dual}} \in \mathbb{S}^n_{++}$ is a constant matrix formulated on the true system and is defined in (56).

Assumption 4 limits the modeling error, and Assumption 5 is related to the controllability of the real system; see more details in Remark 7 of Appendix C. Assumption 4(a) ensures the stabilizability of any $(\hat{A}, \hat{B}) \in \mathcal{S}(\varepsilon_{\text{m}})$ and is used to establish a slower decay rate of $\Phi^{(0:i)}_{\hat{A},\hat{B}}$ for $i < n_{\text{cr}}$, analogous to the slower rate in (17). Moreover, Assumption 4(b) ensures the controllability[3] of $(\hat{A}, \hat{B})$. Then, by combining Assumption 4(b) with Assumptions 4(c) and 5, we can exploit controllability to establish a faster decay rate for $i \geqslant n_{\text{cr}}$, analogous to the faster rate in (17). In Assumptions 4(b), 4(c), and 5, the constants $\nu_\star$, $\alpha_\star$, and P_{dual} are independent of the errors and N under study. See more details in Appendix C.

With the above assumptions, we have the following result.

Proposition 1: If Assumptions 1, 2, 4, and 5 hold, then for any $i, j \in \mathbb{Z}^+$, it holds that

$$\|\Phi^{(j:i)}_{\hat{A},\hat{B}}(P)\| \leqslant \begin{cases} \zeta[\gamma_2(\varepsilon_{\text{m}})]^{i-j}, & \text{if } i - j \geqslant n_{\text{cr}} \\ \bar{\zeta}\,[\overline{\gamma}_2(\varepsilon_{\text{m}}, \varepsilon_{\text{p}})]^{i-j}, & \text{if } 0 \leqslant i - j < n_{\text{cr}} \end{cases} \tag{25}$$

where $\gamma_2(\varepsilon_{\text{m}}) \triangleq \sqrt{1 - \alpha_{\varepsilon_{\text{m}}}^{-1}\beta_\star}$, $\overline{\gamma}_2(\varepsilon_{\text{m}}, \varepsilon_{\text{p}}) \triangleq \sqrt{1 - \alpha_{\varepsilon_{\text{m}}}^{-1}\overline{\beta}(\varepsilon_{\text{p}})}$, $\alpha_{\varepsilon_{\text{m}}} \triangleq (1 - 8\|P_\star\|^2\varepsilon_{\text{m}})^{-1/2}$, and the positive constants ζ and $\bar{\zeta}$ are defined in (66) and (73), respectively.

Proposition 1 shows that if the modeling error is sufficiently small, then the approximate state-transition matrix, formulated on any $(\hat{A}, \hat{B}) \in \mathcal{S}(\varepsilon_{\text{m}})$, is guaranteed to converge to zero. The two distinct decay rates in (25) satisfy $\gamma_2 \leqslant \overline{\gamma}_2 < 1$, analogous to the two decay rates in Lemma 3. Proposition 1 can also be of independent interest, e.g., for the stability analysis of finite-horizon LQR or the Kalman filter with modeling errors.

[3]Even if Assumption 4(b) also implies stabilizability, the upper bound for ε_{m} in Assumption 4(a) is still needed for computing the decay rates.

Remark 3: While the constants in the bounds (23) and (25), e.g., ζ and $\overline{\zeta}$, can be conservative, the rates $\gamma_1(\varepsilon_\mathrm{m})$, $\gamma_2(\varepsilon_\mathrm{m})$, and $\overline{\gamma}_2(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p})$ have the desired property that they recover the ideal rate $\sqrt{1-\beta_\star}$ of the true system in (10) if $\varepsilon_\mathrm{m} = \varepsilon_\mathrm{p} = 0$.

In the rest of this work, we will omit the dependence of γ_1, γ_2, and $\overline{\gamma}_2$ on ε_m and ε_p for the simplicity of notation. With Lemmas 5 and 6 and Proposition 1, we can now characterize ε in (20) as a function of ε_p, ε_m, and N in the following result.

Theorem 2: Given P with $\|P - P_\star\| \leqslant \varepsilon_\mathrm{p}$, if Assumptions 1, 2, 4, and 5 hold, then for $i \in \mathbb{Z}^+$, $\|\mathcal{R}^{(i)}_{\hat{A},\hat{B}}(P) - P_\star\| \leqslant \widehat{E}(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, i) \triangleq$

$$\tilde{\zeta}\left[\gamma_1^i \tilde{\gamma}_2^i \varepsilon_\mathrm{p} + \tilde{\psi}(\varepsilon_\mathrm{m}) \left(\sum_{k=0}^{\min\{i, n_\mathrm{cr}\}-1} \overline{\gamma}_2^k \gamma_1^k + \sum_{j=n_\mathrm{cr}}^{i-1} \gamma_2^j \gamma_1^j \right)\right] \quad (26)$$

where $\tilde{\psi}$ is defined in (51) and satisfies $\tilde{\psi} = O(\varepsilon_\mathrm{m})$

$$\tilde{\gamma}_2 = \begin{cases} \gamma_2, & \text{if } i \geqslant n_\mathrm{cr} \\ \overline{\gamma}_2, & \text{otherwise} \end{cases} \quad (27)$$

and $\tilde{\zeta} = \sqrt{\beta_\star^{-1}} \max\{\overline{\zeta}, \zeta\}$, with γ_2, $\overline{\gamma}_2$, ζ, and $\overline{\zeta}$ defined in Proposition 1, and γ_1 defined in (24).

Note that $\lim_{i\to\infty} \widehat{E}(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, i)$ exists iff ε_m is sufficiently small, as in Assumption 4, such that $\gamma_1\gamma_2 < 1$, and thus, the approximate Riccati iterations, formulated on any $(\hat{A}, \hat{B}) \in \mathcal{S}(\varepsilon_\mathrm{m})$, will converge. In this case, the upper bound (26) characterizes the convergence behavior of the Riccati iterations under the modeling error. As the iteration number i increases, the term $\gamma_1^i \tilde{\gamma}_2^i \varepsilon_\mathrm{p}$ in (26) decreases and drives $\mathcal{R}^{(i)}_{\hat{A},\hat{B}}(P)$ closer to $P_\star$; however, the remaining term increases and drives $\mathcal{R}^{(i)}_{\hat{A},\hat{B}}(P)$ away from $P_\star$, which reflects the propagation of the modeling error. The above two terms capture the tradeoff between the prediction horizon and the modeling error in the convergence of the approximate Riccati iterations.

C. Performance Upper Bound With a Modeling Error

Based on (6), Theorem 2, and Lemma 4, the tradeoff between the prediction horizon, the modeling error, and the terminal matrix error can be reflected in the performance of the RHC controller as follows.

Lemma 7: Given the nominal RHC controller (3), $\widehat{E}$ in (26), and $c_\star$ in (13), if Assumptions 1, 2, 4, and 5 hold, and if $\varepsilon_\mathrm{m} + \widehat{E}(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N-1) \leqslant 1/(40\Upsilon_\star^4 \|P_\star\|^2)$, then $\gamma_1\gamma_2 < 1$ holds, and $A_\star - B_\star K_\mathrm{RHC}$ is Schur stable with $J_{K_\mathrm{RHC}} - J_{K_\star}$

$$\leqslant g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N) \triangleq c_\star \left[\varepsilon_\mathrm{m} + \widehat{E}(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N-1)\right]^2. \quad (28)$$

In Lemma 7, $\widehat{E}(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N-1)$ depends on $N-1$ because of $\mathcal{R}^{(N-1)}_{\hat{A},\hat{B}}(P)$ in (6). Moreover, the upper bound for $\varepsilon_\mathrm{m} + \widehat{E}(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N-1)$ limits the joint effect of ε_m, ε_p, and N, such that the resulting nominal RHC controller can stabilize the unknown system with a suboptimality guarantee. When the modeling error is significantly large, the nominal controller may fail to stabilize the unknown system. In this case, a more accurate terminal value function can be considered, or a robust controller can be used to explicitly address the error [3]. In addition, as $c_\star$ in (13) increases as $\min\{n, m\}$ increases, a larger input dimension or state dimension may lead to a larger $c_\star$ and thus a larger g in (28). Deriving the growth rate of g when the state or input dimension increases is a future direction.

In the bound g, due to the tradeoff in the term $\widehat{E}$, increasing N has a complex impact on the control performance due to the propagation of the modeling error. Then, natural questions are whether an optimal prediction horizon $N_\star$ can be found to minimize the performance upper bound g, and how $N_\star$ varies when the errors ε_m and ε_p change.

Example 1: To address the above questions, consider the case $i = N-1 < n_\mathrm{cr}$ as an example. Then, we can rewrite (26) into

$$\widehat{E} = \tilde{\zeta}\left[\left(\varepsilon_\mathrm{p} - \frac{\tilde{\psi}(\varepsilon_\mathrm{m})}{1-\gamma_1\overline{\gamma}_2}\right)(\gamma_1\overline{\gamma}_2)^{N-1} + \frac{\tilde{\psi}(\varepsilon_\mathrm{m})}{1-\gamma_1\overline{\gamma}_2}\right]. \quad (29)$$

Note that in (29), the term

$$\varepsilon_\mathrm{p} - \frac{\tilde{\psi}(\varepsilon_\mathrm{m})}{1-\gamma_1\overline{\gamma}_2} \quad (30)$$

can be interpreted as a measure of the relative difference between ε_p and ε_m, due to $\tilde{\psi} = O(\varepsilon_\mathrm{m})$, and thus, $\tilde{\psi}(\varepsilon_\mathrm{m})/(1-\gamma_1\overline{\gamma}_2) = O(\varepsilon_\mathrm{m})$. Then, (29) suggests that if the prediction horizon N increases, then the overall change of $\widehat{E}$ and g depends on the sign of (30). More specifically, if ε_p is relatively smaller than ε_m, i.e., (30) is negative, then $\inf_N g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N)$ is attained at $N=1$; otherwise, $\inf_N g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N)$ is attained at the largest N possible, i.e., $N = n_\mathrm{cr}$ in this example.

The observation in Example 1 can be generalized beyond the case $i < n_\mathrm{cr}$. To state the formal result, the following terms are relevant, and their relation can be established:

$$\varepsilon_\mathrm{p} - \frac{\tilde{\psi}(\varepsilon_\mathrm{m})}{1 - \frac{\gamma_1\gamma_2}{(\overline{\gamma}_2/\gamma_2)^{n_\mathrm{cr}-1}}} \geqslant \varepsilon_\mathrm{p} - \frac{\tilde{\psi}(\varepsilon_\mathrm{m})}{1-\gamma_1\gamma_2} \geqslant \varepsilon_\mathrm{p} - \frac{\tilde{\psi}(\varepsilon_\mathrm{m})}{1-\gamma_1\overline{\gamma}_2} \quad (31)$$

which follows from $\gamma_2 \leqslant \overline{\gamma}_2$. Similar to (30), the above terms represent three quantitative measures of the difference between the two errors ε_m and ε_p. Their signs decide the optimal prediction horizon $N_\star$, as shown in the following result.

Theorem 3: In the setting of Lemma 7, let $\gamma_\mathrm{ratio} \triangleq (\overline{\gamma}_2/\gamma_2)^{n_\mathrm{cr}-1} \geqslant 1$, and consider $g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N)$ in (28).

a) If $\varepsilon_\mathrm{p} - \tilde{\psi}(\varepsilon_\mathrm{m})/(1-\gamma_1\overline{\gamma}_2) > 0$, then $\inf_N g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N) = \lim_{N\to\infty} g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N)$ and $g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N)$ is strictly decreasing as N increases.

b) If $\varepsilon_\mathrm{p} - \tilde{\psi}(\varepsilon_\mathrm{m})/[1-\gamma_1\gamma_2] < 0$, then $\inf_N g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N) = g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N_\star)$ with $N_\star = 1$, and if $\varepsilon_\mathrm{p} - \tilde{\psi}(\varepsilon_\mathrm{m})/[1-\gamma_1\gamma_2/\gamma_\mathrm{ratio}] < 0$ holds additionally, then $g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N)$ is strictly increasing as N increases.

c) If $\varepsilon_\mathrm{p} - \tilde{\psi}(\varepsilon_\mathrm{m})/(1-\gamma_1\gamma_2) > 0 > \varepsilon_\mathrm{p} - \tilde{\psi}(\varepsilon_\mathrm{m})/(1-\gamma_1\overline{\gamma}_2)$, then $g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N)$ is strictly increasing as $N \in \{1, \ldots, n_\mathrm{cr}\}$ increases and strictly decreasing as $N \geqslant n_\mathrm{cr}+1$ increases. Moreover, the following holds:

   1) $\inf_N g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N) = g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N_\star)$ with $N_\star = 1$ if additionally

$$\varepsilon_\mathrm{p} - \tilde{\psi}(\varepsilon_\mathrm{m})\left[\frac{1-(\overline{\gamma}_2\gamma_1)^{n_\mathrm{cr}}}{1-\gamma_1\overline{\gamma}_2} + \frac{(\gamma_1\gamma_2)^{n_\mathrm{cr}}}{1-\gamma_1\gamma_2}\right] \leqslant 0;$$

   2) $\inf_N g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N) = \lim_{N\to\infty} g(\varepsilon_\mathrm{m}, \varepsilon_\mathrm{p}, N)$ otherwise;

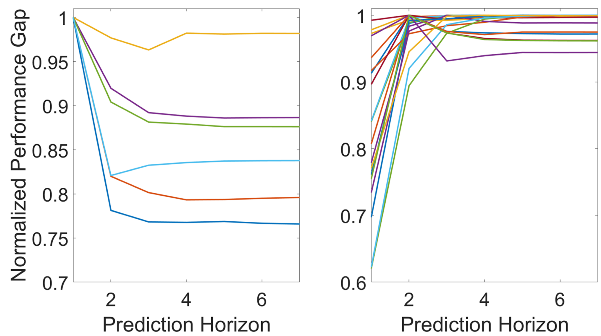


Fig. 1. Given 10-D random real systems, the performance gaps $J_{K_{\text{RHC}}} - J_{K_\star}$ (after normalization) of 25 nominal RHC controllers under varying N are divided into two groups based on the trend of their performance and are shown in these two figures.

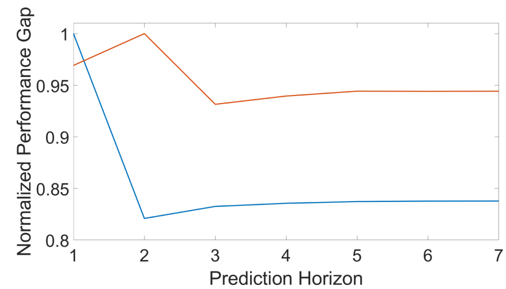


Fig. 2. Two nominal RHC controllers achieve their optimal closed-loop performance at $N = 2$ and $N = 3$, respectively.

d) If $\varepsilon_{\text{p}} - \tilde{\psi}(\varepsilon_{\text{m}})/[1 - \gamma_1\gamma_2] = \varepsilon_{\text{p}} - \tilde{\psi}(\varepsilon_{\text{m}})/(1 - \gamma_1\overline{\gamma}_2) = 0$, then $g(\varepsilon_{\text{m}}, \varepsilon_{\text{p}}, N)$ is a constant as N varies,

where γ_2 and $\overline{\gamma}_2$ are defined in Proposition 1, γ_1 is defined in (24), and $\tilde{\psi}$ is defined in (51) satisfying $\tilde{\psi} = O(\varepsilon_{\text{m}})$.

Theorem 3 shows that in (28) g achieves its infimum either when $N = 1$ or when $N \to \infty$, depending on the relative difference between ε_{m} and ε_{p}, quantified by the terms in (31). More specifically, Theorem 3(a) shows that if ε_{p} is relatively larger than ε_{m} in the sense of (30) being positive, $N \to \infty$ achieves the infimum of g. On the other hand, Theorem 3(b) shows that if ε_{m} is relatively larger than ε_{p}, i.e., when $\varepsilon_{\text{p}} - \tilde{\psi}(\varepsilon_{\text{m}})/[1 - \gamma_1\gamma_2] < 0$, then g is minimized by choosing $N = 1$. Note that g is exponential in N based on (26) and (29). In addition, given the errors, g is a monotonic function of N in Theorem 3(a) and 3(b), but it shows a more complex behavior in (c).

Remark 4: Depending on the relative difference between ε_{m} and ε_{p}, Theorem 3 indicates that choosing $N = 1$ or $N \to \infty$ could achieve a better control performance. However, even if $N \to \infty$ is desired, a finite N can be sufficient to make g close to its infimum, based on its exponential dependence on N. Considering additionally the faster decay rate in (27) when $i \geqslant n_{\text{cr}}$, it can be beneficial to choose a finite N satisfying $N > n_{\text{cr}}$, where we recall that n_{cr} is the controllability index. This also suggests choosing $N = n_{\text{cr}}$ or $n_{\text{cr}} + 1$ could be reasonable initial guesses, which can then be further tuned for performance improvement.[4] If n_{cr} is not known, the state dimension can be an alternative, as a general controllable system has $n_{\text{cr}} = n$. Using $N = n$ is also suggested in [32], and in [33] for controlling nonholonomic vehicles.

The above observation depends on the performance *upper bound* in (28) and thus may not reflect the actual behavior of $J_{K_{\text{RHC}}} - J_{K_\star}$. However, the observation can indeed be seen in extensive simulations in Fig. 1. In this simulation study, five random real systems $(A_\star, B_\star)$, with $n = 10$ and $m = 2$, are generated, where each entry in $A_\star$ and $B_\star$ is sampled from $\mathcal{U}_{[-2,2]}$ and $\mathcal{U}_{[0,1]}$, respectively. All systems have controllability index $n_{\text{cr}} = 5$. Moreover, each $(A_\star, B_\star)$ and the corresponding $P_\star$ are perturbed by random matrices,[5] whose entries are sampled from $\mathcal{U}_{[0,0.001]}$. This is repeated for five times for each real system, leading to 25 approximate $(\hat{A}, \hat{B}, P)$ and 25 nominal RHC controllers. For each controller, we vary its prediction horizon and compute the normalized performance gap as $[J_{K_{\text{RHC}}}(N) - J_{K_\star}]/\max_N(J_{K_{\text{RHC}}}(N) - J_{K_\star})$.

The results are in Fig. 1. The performance gaps have initial transient behaviors, but converge quickly as N increases. Most controllers indeed obtain their optimal control performance at $N = 1$ or the largest N until convergence. All of them converge after $N > n_{\text{cr}}$, showing the potential benefit of a finite N satisfying $N > n_{\text{cr}}$ even if $N \to \infty$ is desired. This is important when the computational aspect is considered, as a larger N leads to a higher computational cost.

However, there are a few cases where the optimal N is finite but greater than 1. Two examples are shown in Fig. 2, where $N = 2$ and $N = 3$ achieve the optimal performance, respectively. However, the performance difference between the optimal performance and the performance upon convergence is not significant. These examples show that the upper bound may not reflect the actual behavior of $J_{K_{\text{RHC}}} - J_{K_\star}$.

Remark 5: Lemma 7 is established under tight conditions, i.e., the bound $1/(40\Upsilon_\star^4\|P_\star\|^2)$ and the bounds on ε_{m} in Assumption 4 can be small and thus may not be satisfied in practice. This is a common limitation of analytical error bounds [19], [20]. Despite this limitation, we can still obtain a nontrivial result in Theorem 3, and this result can be observed in simulations when going beyond the tight conditions, e.g., the modeling errors of the simulation in Fig. 1 are actually greater than $1/(40\Upsilon_\star^4\|P_\star\|^2)$.

We also consider the following practical case of Theorem 3 where $P = 0$ in the RHC controller (3).

Corollary 1: Given the nominal RHC controller (3) with $P = 0$, if Assumptions 1, 2, 4, and 5 hold, and if $\varepsilon_{\text{m}} + \widehat{E}(\varepsilon_{\text{m}}, \|P_\star\|, N-1) \leqslant 1/(40\Upsilon_\star^4\|P_\star\|^2)$, then $A_\star - B_\star K_{\text{RHC}}$ is Schur stable with $J_{K_{\text{RHC}}} - J_{K_\star} \leqslant g(\varepsilon_{\text{m}}, \varepsilon_{\text{p}}, N)$, where $\inf_N g(\varepsilon_{\text{m}}, \varepsilon_{\text{p}}, N) = \lim_{N\to\infty} g(\varepsilon_{\text{m}}, \varepsilon_{\text{p}}, N)$ and $g(\varepsilon_{\text{m}}, \varepsilon_{\text{p}}, N)$ is strictly decreasing as N increases.

The above result shows that for the RHC controller with zero terminal function, using a large prediction horizon, ideally the nominal LQR controller, is beneficial. This is because the other choice $N = 1$ leads to $K_{\text{RHC}} = 0$, which will never stabilize the system if the system is open-loop unstable.

[4] In this work, the control horizon and the prediction horizon are equal.

[5] For $P_\star$, a random matrix is multiplied by its transpose to generate a positive semidefinite matrix perturbation.

V. Extensions to Stabilizable Systems

We generalize the previous results for controllable systems to stabilizable systems. Without controllability to establish the fast decay rates $1-\beta_\star$ in (17) and γ_2 in (25) for $i \geqslant n_{\rm cr}$, we only establish a single decay rate for all $i \geqslant 1$. These results for stabilizable systems are formalized in Appendix D. The single decay rate greatly simplifies the analysis, which leads to a variant of Theorem 2 for stabilizable systems.

Theorem 4: If Assumptions 2, 3, and 4(a) hold, then for $i \in \mathbb{Z}^+$, $\|\mathcal{R}^{(i)}_{\hat{A},\hat{B}}(P) - P_\star\|$

$$\leqslant \widehat{E}_{\rm sta}(\varepsilon_{\rm m}, \varepsilon_{\rm p}, i) \triangleq \overline{\zeta}\left[\gamma_1^i \overline{\gamma}_2^i \varepsilon_{\rm p} + \tilde{\psi}(\varepsilon_{\rm m})\left(\sum_{j=0}^{i-1} \overline{\gamma}_2^j \gamma_1^j\right)\right] \quad (32)$$

where γ_1 is defined in (24), $\tilde{\psi}$ is defined in (51) and satisfies $\tilde{\psi} = O(\varepsilon_{\rm m})$, and $\overline{\zeta}$ and $\overline{\gamma}_2$ are defined in Proposition 1.

Similarly to (28), (32) also leads to an error bound

$$g_{\rm sta}(\varepsilon_{\rm m}, \varepsilon_{\rm p}, N) \triangleq c_\star[\varepsilon_{\rm m} + \widehat{E}_{\rm sta}(\varepsilon_{\rm m}, \varepsilon_{\rm p}, N-1)]^2. \quad (33)$$

Equation (32) can be rewritten as (29). Therefore, the observation from Theorem 3 that the performance upper bound achieves its infimum at $N = 1$ or $N \to \infty$ remains valid for stabilizable systems, now depending on the sign of (30) only.

By observing (16), we note that for stabilizable systems, the faster rate $1-\beta_\star$ in (17) can also be established by letting i be sufficiently large, e.g., larger than a constant i_0, such that the closed-loop matrices in $\Phi^{(0:i)}_{A,B}(P)$ become close to $L_\star$ in (9). However, i_0 can be large and also does not have a clear interpretation, unlike the controllability index $n_{\rm cr}$ in (17). Therefore, this technical extension is not pursued here.

VI. Applications in Learning-Based Control

Besides the obtained theoretical insights, our suboptimality analysis can also be used to analyze the performance of learning-based controllers. Results in this section generalize the results in [18], [19], and [20] for learning-based LQR controllers with an infinite horizon to RHC controllers with an arbitrary prediction horizon. We will compare our performance upper bounds analytically with the above existing results, as is done in similar studies on regret analysis [19], [20].

A. Offline Identification and Control

In this section, we first obtain an estimate $(\hat{A}, \hat{B})$ offline from measured data of the unknown real system (2), and then synthesize a controller (3) with zero terminal matrix $P = 0$. This is the classical receding-horizon LQ controller [10].

There are many recent studies on linear system identification and its finite-sample error bounds [18], [23], [24]. For interpretability, we consider a relatively simple estimator from [18]. Assume that the white noise w_t is Gaussian, and we conduct T independent experiments on the unknown real system (2) by injecting independent Gaussian noise, i.e., $u_t \sim \mathcal{N}(0, \sigma_u^2 I)$ with $\sigma_u > 0$, where each experiment starts from $x_0 = 0$ and lasts for t_h time steps. This leads to the measured data $\{(x_t^{(l)}, u_t^{(l)})\}_{t=0}^{t_h}$, $l = 1, \ldots, T$, which is independent over the experiment index l. To avoid the dependence among the data for establishing the estimation error, we use one data sample from each independent experiment, and then an estimate $(\hat{A}, \hat{B})$ is obtained from the least-squares (LS) estimator [18]

$$\begin{bmatrix}\hat{A} & \hat{B}\end{bmatrix} = \arg\min_{A,B} \sum_{l=1}^{T} \left\| x_{t_h}^{(l)} - \begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} x_{t_h-1}^{(l)} \\ u_{t_h-1}^{(l)} \end{bmatrix} \right\|^2. \quad (34)$$

A bound $\varepsilon_{\rm m}$ can be obtained [18, Prop. 1] so that with high probability and for some constant $c_{\rm ls} > 0$

$$\max\{\|\hat{A} - A_\star\|, \|\hat{B} - B_\star\|\} \leqslant \varepsilon_{\rm m} = c_{\rm ls}/\sqrt{T}. \quad (35)$$

Consider the nominal controller (3) with $P = 0$ and $(\hat{A}, \hat{B})$ from (34). We can provide an end-to-end control performance guarantee by combining the LS estimation error bound (35) and our performance upper bound (33). We consider (33) here for generality as it needs more relaxed assumptions than (28). To this end, we first simplify (33) for interpretability.

Proposition 2: Given the controller (3) with $P = 0$, if Assumptions 2, 3, and 4(a) hold, and if $\varepsilon_{\rm m} + \widehat{E}_{\rm sta}(\varepsilon_{\rm m}, \|P_\star\|, N-1) \leqslant 1/(40\Upsilon_\star^4\|P_\star\|^2)$, then $A_\star - B_\star K_{\rm RHC}$ is Schur stable with $J_{K_{\rm RHC}} - J_{K_\star} \leqslant C\left(\mu_\star^{N-1} + \varepsilon_{\rm m}\right)^2$, for some $C > 0$ that is independent of N and $\varepsilon_{\rm m}$, where $\mu_\star \triangleq \sqrt{1 - \beta_\star/(2)^{3/2}} \in (0, 1)$.

Combining Proposition 2 with the estimation error bound (35) from [18] directly leads to the following end-to-end performance guarantee.

Corollary 2: Consider the estimate $(\hat{A}, \hat{B})$ from the estimator (34) and the controller (3) with $P = 0$. If Assumptions 1 and 2 hold, and if the noise w_t is Gaussian and $c_{\rm ls}/\sqrt{T} + \widehat{E}_{\rm sta}(c_{\rm ls}/\sqrt{T}, \|P_\star\|, N-1) \leqslant 1/(40\Upsilon_\star^4\|P_\star\|^2)$, then for any $\delta \in (0, 1)$, there exists a sufficiently large T such that with probability at least $1 - \delta$, $A_\star - B_\star K_{\rm RHC}$ is Schur stable with

$$J_{K_{\rm RHC}} - J_{K_\star} \leqslant C\left(\mu_\star^{N-1} + \sqrt{\log(1/\delta)/T}\right)^2$$

for some $C > 0$ that is independent of N and T.

Corollary 2 reveals the effect of the prediction horizon and the number of data samples on the control performance. The performance gap is $O(1/T)$ if $N \to \infty$, i.e., when the nominal LQR controller is considered. This growth rate agrees with the recent study [19] of the LQR controller. When N is finite, an additional error $\mu_\star^{N-1}$ exists and decreases as N increases. When $T \to \infty$, the estimated model converges to the real system with high probability, leading to a performance gap $O(\mu_\star^{N-1})$, i.e., the performance of the controller converges exponentially to the optimal one as N increases, which matches our observation in (19). Note that controllability in Assumption 1 is introduced in Corollary 2, required by the estimation error bound [18]. Moreover, Assumption 4(a) is absent, as it is satisfied by a sufficiently large T.

Remark 6: We have used data from multiple independent experiments for estimating the model; however, in practice, maybe only one experiment can be conducted. In this case, an LS estimator with the data from a single experiment can be used, with its error bound studied in [23] and [24].

B. Online Learning Control With Regret Bound

In this section, we consider the adaptive LQR controller algorithm in [20], which has a state-of-the-art regret guarantee with greedy exploration. While there are other exploration strategies [17], the more fundamental greedy exploration is chosen to demonstrate the application of our analysis.

The nominal LQR controller within the adaptive control scheme in [20] is replaced by the receding-horizon LQ controller (3) with $P = 0$. Other cost matrices Q and R are fixed throughout the closed-loop operation. To match the setting in [20], in this section, we assume that w_t is Gaussian with $\sigma_w = 1$ and the initial condition is $x_0 = 0$. Following [20], we assume that a stabilizing but possibly suboptimal controller gain K_0 for the unknown true system is given a priori.

We briefly introduce the main idea of [20, Algorithm 1], and the details can be found in [20]. Starting from $x_0 = 0$ and a random input $u_0 \sim \mathcal{N}(0, I)$, the algorithm utilizes a fixed controller gain K_k within each time step period $[t_k, t_{k+1})$, where $t_k = 2^{k-1}$ and $k \in \{1, 2, \dots\}$ is the period index. In the first few periods, the gain K_0 with a Gaussian perturbation is used, i.e., $u_t = -K_k x_t + g_t$ with $K_k = K_0$ and $g_t \sim \mathcal{N}(0, I)$, to stabilize the unknown system. The perturbation g_t ensures the informativity of the data for estimating the system later.

After collecting sufficient data at time step t_{k_0} for some period k_0, the controller conducts the following steps for every period $k \geqslant k_0$. A new model $(\hat{A}_k, \hat{B}_k)$ is re-estimated at the initial time step t_k of period k. It is obtained from the LS estimator using the data $\{z_t\}_{t=t_{k-1}}^{t_k - 1}$ from the last period $k-1$, where $z_t \triangleq \begin{bmatrix} x_t^\top & u_t^\top \end{bmatrix}^\top$, as

$$\begin{bmatrix} \hat{A}_k & \hat{B}_k \end{bmatrix} = \arg\min_{A,B} \sum_{t=t_{k-1}}^{t_k - 1} \left\| x_{t+1} - \begin{bmatrix} A & B \end{bmatrix} z_t \right\|^2. \quad (36)$$

Then, for $t \in [t_k, t_{k+1})$, the algorithm employs a new RHC controller gain $K_{\text{RHC},k}$, formulated on[6] $(\hat{A}_k, \hat{B}_k)$ via either the implicit form (3) or the explicit form (6), with a random perturbation, i.e., $u_t = -K_{\text{RHC},k} x_t + \sigma_k g_t$, where $\sigma_k > 0$ determines the perturbation size and decays as k increases.

In this case, let T denote the total number of time steps that have passed, and let k_T denote the final period index. Let $\mathcal{C} : \mathbb{R}^n \to \mathbb{R}^m$ denote the above adaptive controller, and its performance is typically characterized by *regret* [23], [34]

$$\text{Regret}(T) \triangleq \sum_{t=0}^{T} \left[x_t^\top Q x_t + \mathcal{C}^\top(x_t) R \mathcal{C}(x_t) - J_{K_\star} \right]$$

which measures the cumulative error of a particular realization under the controller. It is ideal to have a sublinear regret such that as $T \to \infty$, $\text{Regret}(T)/T$ converges to zero with high probability, i.e., the average performance of the adaptive controller is optimal.

As shown in [20, Sect. 5.1] and [20, Appendix G.2], the following holds with high probability for the controller:

$$\text{Regret}(T) = \tilde{O}\left(\sum_{k=k_0}^{k_T} t_k (J_{K_{\text{RHC},k}} - J_{K_\star}) + \sqrt{T} \right).$$

Combining the above equation with Proposition 2 leads to

$$\text{Regret}(T) = \tilde{O}\left(\sum_{k=k_0}^{k_T} t_k \left[\mu_\star^{N-1} + 1/(t_k^{1/4}) \right]^2 + \sqrt{T} \right) \quad (37)$$

where we use the fact that the estimated model at time step t_k has error $\varepsilon_{\text{m}} = O(1/t_k^{1/4})$ with high probability [20, Lemma 5.4].

The regret upper bound (37) provides some interesting information. First, note that $\sum_{k=k_0}^{k_T} \sqrt{t_k} \leqslant \bar{C}\sqrt{T}$ for some constant $\bar{C} > 0$. Therefore, if we have an infinite prediction horizon $N \to \infty$, i.e., when the nominal LQR is considered, then $\text{Regret}(T) = \tilde{O}(\sqrt{T})$, which matches the rate of the adaptive LQR controllers in [19] and [20].

If we have a finite N, then (37) shows with high probability

$$\text{Regret}(T) = \tilde{O}\left(T\mu_\star^N + \sqrt{T} \right) \quad (38)$$

where the regret is linear in T. This observation matches the result in [35], where the regret of a linear unconstrained RHC controller, with a fixed prediction horizon and an exact system model, is linear in T. This linear regret is caused by the fact that even if the model is perfectly identified, the RHC controller still deviates from the optimal LQR controller due to its finite prediction horizon.

To achieve a sublinear regret, (38) suggests that an adaptive prediction horizon is preferred. As also suggested in [35] but for the case of a known system, we can update N via $N = O(\log(t_k))$, e.g., updating $N = \lfloor -\log(t_k)/(4\log(\mu_\star)) \rfloor$, when the model is re-estimated, which again leads to a sublinear regret $\tilde{O}(\sqrt{T})$ according to (37). Note that this is the optimal rate for the regret [20]. The intuition of this choice is that, with a more refined model due to re-estimation, the prediction horizon can be increased adaptively to improve the control performance.

C. Simulation of Adaptive RHC

We use simulations to demonstrate the main theoretical insights from Section VI-B. For the adaptive RHC algorithm, while a fixed prediction horizon leads to a linear regret growth rate $\tilde{O}(T)$, an adaptive horizon, being a logarithmic function of time, leads to a more desired sublinear rate $\tilde{O}(\sqrt{T})$.

While the algorithm based on [20] in Section VI-B has a regret guarantee, a conservative modeling error bound ε_{m}, analogous to (35), is utilized in [20]. This conservatism causes practical issues, e.g., despite the actual modeling error being small, the condition $\varepsilon_{\text{m}} < \bar{\delta}(\hat{A}, \hat{B})$ for some $\bar{\delta}$ in the algorithm is hardly met as ε_{m} is too large. Therefore, by introducing tuning parameters, we slightly modify the algorithm in Section VI-B to avoid analytical error bounds. The resulting algorithm is more practical and suffices to demonstrate the theoretical insights. We summarize the modifications as follows.

i) The key period index k_0 in Section VI-B is chosen to be the first period k such that $\sum_{t=0}^{t_k - 1} z_t z_t^\top \succeq \delta_1 I$ holds, where $\delta_1 \in [1, \infty)$ is a tuning parameter.

ii) The LS estimator (36) uses data $\{z_t\}_{t=0}^{t_k - 1}$ from all the previous periods, instead of only the last period.

iii) We let $\sigma_k = \min\{1, (\delta_2/\sqrt{t_k})^{1/2}\}$, where $\delta_2 \in (0, \infty)$ is a tuning parameter instead of being computed via an analytical formula as in [20].

In the above, (i) ensures that the collected data are informative to obtain an accurate LS estimate of the model. Point (ii) ensures the updated estimate is not worse than the previous one, as more

[6] In [20], the estimated model, if not accurate, is modified via projection. These details are presented in [20]. This projection step is redundant if the estimated model is sufficiently accurate. Moreover, it relies on a conservative error bound, leading to practical issues as discussed in Section VI-C. Therefore, the projection step is ignored later in the simulations of Section VI-C.

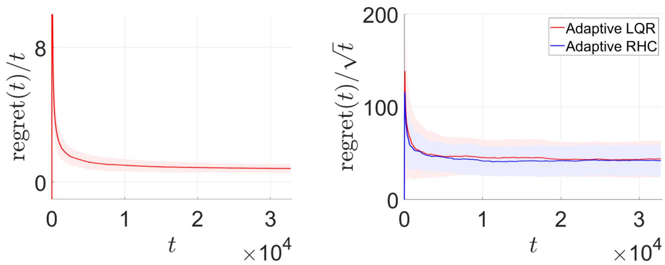


Fig. 3. For the RHC with a fixed N (left) and the one with an adaptive N (right), the mean and standard deviation (shaded region) of the normalized regrets, regret$(t)/t$ for the fixed N and regret$(t)/\sqrt{t}$ for the adaptive N, over 100 simulations are shown. After initial transient behavior, the normalized regrets converge to positive constants. The normalized regret of the adaptive LQR (right) is also shown.

data are used for estimation. With point (iii), the exploration effort decays as the period index increases.

We use the modified algorithm to control the unknown system with the following true system matrices:

$$A_\star = \begin{bmatrix} 0.5 & -2 & 0.9 \\ 0.6 & 0.1 & 1.8 \\ 1.3 & -0.3 & 1.6 \end{bmatrix}, \quad B_\star = \begin{bmatrix} 0.3 & 1 \\ 0.9 & 1 \\ 0 & 0.9 \end{bmatrix}.$$

Starting from $x_0 = 0$, we compare the regret growth of the adaptive RHC controller with a fixed prediction horizon $N = 3$ and the controller with an adaptive prediction horizon $N = \max\{3, \lfloor \log(t_k) \rfloor\}$, where N is updated when the controller is updated. Other parameters are $\sigma_w = 1$, $Q = I$, $R = I$, $P = 0$, $\delta_1 = 4$, $\delta_2 = 0.5$, and K_0 is the LQR controller gain for stabilization. Note that although the true system and its LQR controller are unknown, we choose the LQR controller to be the initial controller for simplicity. The choice of K_0, if stabilizing, does not affect our illustration of the regret growth.

As the regret is a random variable due to disturbance w_t, we conduct 100 simulations for each controller. The regret growth over time, normalized by $\sqrt{t}$ or t, is shown in Fig. 3. The regret grows linearly in T under the fixed N but in the order of $\sqrt{T}$ under the adaptive N. This illustrates the advantage of the adaptive N for regret minimization. When the nominal LQR is used within the adaptive algorithm, the regret remains sublinear with a larger standard deviation in the initial period.

While the modified algorithm does not preserve the regret guarantee in Section VI-B, it remains useful to demonstrate the impact of the adaptive prediction horizon on regret growth. Moreover, this algorithm has a structure similar to that of existing algorithms with regret guarantees, e.g., the adaptive LQ Gaussian controller in [36]. Our future work will investigate the possibility of extending existing analyses to the modified algorithm. The computational increase with a larger N is minor in the current setting and case study, as no online optimization is needed. A larger N only requires more Riccati iterations for computing the controller. For example, in Matlab 2024a with an Intel Core i9-13950HX CPU, the Riccati iteration takes on average around $3 \cdot 10^{-6}$ s for $N = 3$, $1.8 \cdot 10^{-5}$ s for $N = 15$, and solving the Riccati equation for LQR takes $2.4 \cdot 10^{-4}$ s. Therefore, the computational cost is a more important problem for more general settings with constraints or nonlinear systems, requiring online optimization.

Changing the horizon adaptively is also investigated in [37], [38], and the references therein, which are either heuristic or not analyzed in the context of regret. This idea also has implications for, e.g., robot control in a partially unknown environment [39]. There, as more data are collected, the RHC/MPC controller of the robot can increase its prediction horizon to improve performance. The potential challenges are the increase in computation time and handling data that have poor quality. In the latter case, the model may not improve over time, and it is important to decide when to increase the prediction horizon. These challenges will be investigated in future work.

VII. Conclusion

This work analyzes the suboptimality of RHC under the joint effect of the modeling error, the terminal value function error, and the prediction horizon in the LQ setting. By deriving a novel perturbation analysis of the Riccati difference equation, we have obtained a novel performance upper bound of the controller. The bound suggests that letting the prediction horizon be 1 or $+\infty$ can potentially be beneficial, depending on the relative difference between the modeling error and the terminal matrix error. Moreover, when an infinite horizon is desired, a prediction horizon larger than the controllability index can be sufficient for achieving a near-optimal performance. Besides the above insight, this obtained performance upper bound has also been shown to be useful for analyzing the performance of learning-based receding-horizon LQ controllers.

Extending the results to more general settings with constraints and nonlinear systems is an important direction. The overall steps of our analysis can be applied to nonlinear systems; however, the technical derivations differ significantly. For example, instead of the Riccati iteration, the value iteration needs to be considered. This generalization can be found in the follow-up work [40]. Future work also includes the derivation of tighter performance upper bounds to capture the nontrivial optimal horizon, which is finite but larger than 1, and enhancing the adaptive RHC algorithm to achieve both practicality and theoretical guarantees. It is also important to benchmark the algorithm against other state-of-the-art methods.

Appendix A
Technical Tools

Technical tools from the literature are collected here.

Lemma 8: Given $0 < a \leqslant b < 1$ and any positive integer i, we have $(1 - b^{i+1})/(1 - a^{i+1}) - (1 - b^i)/(1 - a^i) \geqslant 0$.

Proof: This result is obtained by applying formula $a^i - b^i = (a - b)(a^{i-1} + a^{i-2}b + \cdots + ab^{i-2} + b^{i-1})$. ■

Lemma 9 (See [19, Lemma 7]): Given $M \in \mathbb{S}_+^n$ and $N \in \mathbb{S}_+^n$, it holds that $\|N(I + MN)^{-1}\| \leqslant \|N\|$.

Lemma 10 (See [19, Lemma 5]): Consider $M \in \mathbb{R}^{n \times n}$, $c \geqslant 1$, and $\rho \geqslant \rho(M)$ such that $\|M^i\| \leqslant c\rho^i$ for any $i \in \mathbb{Z}^+$. For any $i \in \mathbb{Z}^+$ and $\Delta \in \mathbb{R}^{n \times n}$, $\|(M + \Delta)^i\| \leqslant c(c\|\Delta\| + \rho)^i$ holds.

Important identities from [10] and [21, eq. (2.3) and (4.2)] are collected in the following result.

Lemma 11: If $R \succ 0$, for any P, P_1, $P_2 \in \mathbb{S}_+^n$ and $i \in \mathbb{Z}^+$, then the following holds:

a) $\mathcal{R}_{A,B}(P_1) - \mathcal{R}_{A,B}(P_2) = [A - B\mathcal{K}_{A,B}(P_1)]^\top(P_1 - P_2)[A - B\mathcal{K}_{A,B}(P_2)]$;

b) $\mathcal{R}_{A,B}(P) = [A - B\mathcal{K}_{A,B}(P)]^\top P[A - B\mathcal{K}_{A,B}(P)] + \mathcal{K}_{A,B}^\top(P)R\mathcal{K}_{A,B}(P) + Q$;

c) $\mathcal{F}(P) \triangleq B(R + B^\top PB)^{-1}B^\top = S_B(I + PS_B)^{-1}$;

d) $\Phi_{A,B}^{(0:i)}(P_1) = [I + \mathcal{O}_{A,B}^{(i)}(P_1)(P_2 - P_1)]\Phi_{A,B}^{(0:i)}(P_2)$ with $\mathcal{O}_{A,B}^{(i)}$ defined in (40).

The following perturbation analysis of the Riccati equation from [20, Prop. 6] and [20, Lem. B.5, B.8] is utilized.

Lemma 12 (See [20]): Given Q, R, and $(A_\star, B_\star)$, consider any alternative system $(\hat{A}, \hat{B})$ satisfying $\max\{\|\hat{A} - A_\star\|, \|\hat{B} - B_\star\|\} \leqslant \varepsilon_{\rm m}$. If Assumptions 2 and 3 hold, then

1) $P_\star \succeq I$, $\|P_\star\| \geqslant \|L_\star\|^2$, and $\|P_\star\| \geqslant \|K_\star\|^2$;

2) if $\varepsilon_{\rm m} < 1/(8\|P_\star\|^2)$, then $(\hat{A}, \hat{B})$ is stabilizable and

$$\|\hat{P}\| \leqslant \alpha_{\varepsilon_{\rm m}}\|P_\star\|, \quad \|\hat{K} - K_\star\| \leqslant 7\alpha_{\varepsilon_{\rm m}}^{7/2}\|P_\star\|^{7/2}\varepsilon_{\rm m} \tag{39}$$

where $\alpha_{\varepsilon_{\rm m}} = (1 - 8\|P_\star\|^2\varepsilon_{\rm m})^{-1/2}$, $\hat{P}$ is the fixed point of the Riccati equation, and $\hat{K}$ is the LQR control gain for $(\hat{A}, \hat{B})$.

Another tool is the stability analysis of time-varying systems, directly implied by the proof of [29, Thm. 23.3].

Lemma 13: Consider a system $\bar{x}_{k+1} = A_k\bar{x}_k$ with $A_k \in \mathbb{R}^{n\times n}$ and $\phi^{(j:i)} \triangleq A_{i-1}\ldots A_j$ for integers $i \geqslant j \geqslant 0$. If there exists a matrix sequence $W_k \in \mathbb{S}^n$ satisfying for $k \in [j, i]$

$$b_{\rm l}I \preceq W_k \preceq b_{\rm u}I, \text{ and } A_k^\top W_{k+1}A_k - W_k \preceq -aI$$

for some positive constants $b_{\rm l}$, $b_{\rm u}$, and a, then $\|\phi^{(j:i)}\| \leqslant \sqrt{b_{\rm u}/b_{\rm l}}(\sqrt{1 - a/b_{\rm u}})^{i-j}$ with $a/b_{\rm u} \in (0, 1]$.

Appendix B
Main Proofs

Proofs of Lemmas 4, 7, and Corollary 1 are in the extended version of this article [41]

A. Proof of Lemma 1

Lemma 11(b) shows $-\underline{\sigma}(Q)I \succeq L_\star^\top P_\star L_\star - P_\star$, and in addition, we have $\underline{\sigma}(Q)I \preceq \underline{\sigma}(P_\star)I \preceq P_\star \preceq \overline{\sigma}(P_\star)I$. Then, the direct application of Lemma 13 proves the result. Note that $\beta_\star < 1$ holds; otherwise, $P_\star = Q$ and thus $A = 0$ based on Lemma 11(b), which contradicts with Assumption 2.

B. Proof of Lemma 3

We first define the following Gramian matrix: for $i \geqslant 1$

$$\mathcal{O}_{A,B}^{(i)}(P) \triangleq \sum_{k=0}^{i-1}\mathcal{L}_{A,B}^k(P^{(k)})\mathcal{F}\left(P^{(k)}\right)[\mathcal{L}_{A,B}^k(P^{(k)})]^\top \tag{40}$$

where $\mathcal{F}(P) \triangleq B(R + B^\top PB)^{-1}B^\top$, $\mathcal{O}_{A,B}^{(0)}(P) \triangleq 0$, and recall that $P^{(i)}$ is a shorthand notation for $\mathcal{R}_{A,B}^{(i)}(P)$.

The proof for the case $i \geqslant n_{\rm cr}$ is achieved based on Lemma 3 and [21, Lemma 4.1]. We highlight the key steps for completeness. If $i \geqslant n_{\rm cr}$, then matrix $\mathcal{O}_{A_\star,B_\star}^{(i)}(P_\star)$ has full rank due to Assumption 1. Lemma 11(d) shows $[I + \mathcal{O}_{A_\star,B_\star}^{(i)}(P_\star)(P - P_\star)]\Phi_{A_\star,B_\star}^{(0:i)}(P) = L_\star^i$ for $i \in \mathbb{Z}^+$. Then, based on Assumption 2 and [21, Lemma 4.1], $I + \mathcal{O}_{A_\star,B_\star}^{(i)}(P_\star)(P - P_\star)$ is nonsingular for $i \geqslant n_{\rm cr}$, and moreover, its inverse is uniformly upper bounded for any $i \geqslant n_{\rm cr}$ and any $P \succeq 0$. Combining (16) and the above shows for $i \geqslant n_{\rm cr}$, $\mathcal{R}_{A_\star,B_\star}^{(i)}(P) - P_\star = (L_\star^i)^\top[I + \mathcal{O}_{A_\star,B_\star}^{(i)}(P_\star)(P - P_\star)]^{-\top}(P - P_\star)L_\star^i$, which, together with (10), concludes the proof for $i \geqslant n_{\rm cr}$, with

$$\tau_\star \triangleq \beta_\star^{-1}\sup_{P\in\mathbb{S}_+^n}\sup_{i\geqslant n_{\rm cr}}\left\|\left[I + \mathcal{O}_{A_\star,B_\star}^{(i)}(P_\star)(P - P_\star)\right]^{-1}\right\|. \tag{41}$$

When $1 \leqslant i < n_{\rm cr}$, the rank of the controllability matrix is less than n, and thus, $\mathcal{O}_{A_\star,B_\star}^{(i)}(P_\star)$ is not guaranteed to be of full rank. Then, directly exploiting (16) leads to

$$\|\mathcal{R}_{A_\star,B_\star}^{(i)}(P) - P_\star\| \leqslant \|\Phi_{A_\star,B_\star}^{(0:i)}(P)\|\|L_\star^i\|\|P - P_\star\|. \tag{42}$$

One can upper bound $\|\Phi_{A_\star,B_\star}^{(0:i)}(P)\|$ by a constant. However, due to $Q \succeq I$, $\|\Phi_{A_\star,B_\star}^{(0:i)}(P)\|$ decays exponentially as i increases. This result in (68) is exploited here and proved in Appendix D. Combining (10), (42), and (68) proves this case, with

$$\overline{\tau}_\star \triangleq \Upsilon_\star^3(1 + \Upsilon_\star + \|P_\star\|)\sqrt{\beta_\star^{-1}(\|P_\star\| + \Upsilon_\star\beta_\star^{-1})(1 - \beta_\star)^{-1}} \tag{43}$$

derived from (68) using $\Upsilon_\star \geqslant \varepsilon_{\rm p}$. The case $i = 0$ holds trivially due to $\overline{\tau}_\star \geqslant 1$.

C. Proof of Lemma 5

Define the shorthand notations $\hat{S} \triangleq S_{\hat{B}}$ and $S_\star \triangleq S_{B_\star}$. We prove this result by induction. The equality holds trivially when $i = 0$. Assume that it holds for $i = k$, then

$$\begin{aligned}
&\hat{P}^{(k+1)} - P_\star^{(k+1)} = \mathcal{R}_{\hat{A},\hat{B}}^{(k)}\left(\hat{P}^{(1)}\right) - \mathcal{R}_{A_\star,B_\star}^{(k)}\left(P_\star^{(1)}\right)\\
&= \left[\Phi_{\hat{A},\hat{B}}^{(0:k)}(\hat{P}^{(1)})\right]^\top(\hat{P}^{(1)} - P_\star^{(1)})\bar{\Phi}^{(0:k)}(P_\star^{(1)})\\
&\quad + \sum_{j=1}^{k}\left[\Phi_{\hat{A},\hat{B}}^{(k-j+1:k)}(\hat{P}^{(1)})\right]^\top\mathcal{M}\left(\hat{P}^{(k-j+1)}, P_\star^{(k-j+1)}\right)\\
&\quad \times\bar{\Phi}^{(k-j+1:k)}(P_\star^{(1)})\\
&= \left[\Phi_{\hat{A},\hat{B}}^{(1:k+1)}(P_1)\right]^\top(\hat{P}^{(1)} - P_\star^{(1)})\bar{\Phi}^{(1:k+1)}(P_2)\\
&\quad + \sum_{j=1}^{k}\left[\Phi_{\hat{A},\hat{B}}^{(k-j+2:k+1)}(P_1)\right]^\top\mathcal{M}\left(\hat{P}^{(k-j+1)}, P_\star^{(k-j+1)}\right)\\
&\quad \times\bar{\Phi}^{(k-j+2:k+1)}(P_2).
\end{aligned} \tag{44}$$

In (44), it holds that $\hat{P}^{(1)} - P_\star^{(1)}$

$$\begin{aligned}
&= \hat{A}^\top P_1(I + \hat{S}P_1)^{-1}\hat{A} - A_\star^\top P_2(I + S_\star P_2)^{-1}A_\star\\
&= \underbrace{\hat{A}^\top[P_1(I + \hat{S}P_1)^{-1} - P_2(I + S_\star P_2)^{-1}]\hat{A}}_{\text{Term I}}\\
&\quad + \underbrace{\hat{A}^\top P_2(I + S_\star P_2)^{-1}\hat{A} - A_\star^\top P_2(I + S_\star P_2)^{-1}A_\star}_{\text{Term II}}
\end{aligned} \tag{45}$$

where Term I further leads to $\hat{A}^\top[P_1(I+\hat{S}P_1)^{-1} - P_2(I+S_\star P_2)^{-1}]\hat{A} = \hat{A}^\top[(I+P_1\hat{S})^{-1}P_1 - P_2(I+S_\star P_2)^{-1}]\hat{A} = \hat{A}^\top(I+P_1\hat{S})^{-1}[P_1(I+S_\star P_2) - (I+P_1\hat{S})P_2](I+S_\star P_2)^{-1}\hat{A} = \hat{A}^\top(I+P_1\hat{S})^{-1}[(P_1-P_2)+(P_1S_\star P_2 - P_1\hat{S}P_2)](I+S_\star P_2)^{-1}\hat{A} = \hat{A}^\top(I+P_1\hat{S})^{-1}(P_1-P_2)[I+(S_\star-\hat{S})P_2](I+S_\star P_2)^{-1}\hat{A} + \hat{A}^\top(I+P_1\hat{S})^{-1}P_2(S_\star-\hat{S})P_2(I+S_\star P_2)^{-1}\hat{A}$

$$= [\mathcal{L}_{\hat{A},\hat{B}}(P_1)]^\top (P_1-P_2)\bar{\mathcal{L}}(P_2) + \hat{A}^\top(I+P_1\hat{S})^{-1}P_2(S_\star-\hat{S})P_2(I+S_\star P_2)^{-1}\hat{A}. \quad (46)$$

Combining (45) and (46) leads to

$$\hat{P}^{(1)} - P_\star^{(1)} = [\mathcal{L}_{\hat{A},\hat{B}}(P_1)]^\top(P_1-P_2)\bar{\mathcal{L}}(P_2) + \mathcal{M}(P_1,P_2). \quad (47)$$

Then, plugging (47) into (44) shows $\hat{P}^{(k+1)} - P_\star^{(k+1)}$ satisfies the equality in this lemma, concluding the proof by induction.

D. Proof of Lemma 6

The matrix $\bar{\mathcal{L}}(P_\star)$ in (21) is reformulated as

$$\bar{\mathcal{L}}(P_\star) = L_\star + \underbrace{\mathcal{H}(P_\star) + S_\Delta P_\star[\mathcal{H}(P_\star) + L_\star]}_{\Delta\triangleq} \quad (48)$$

where $S_\Delta = S_{B_\star} - S_{\hat{B}}$. To upper bound the norm of Δ, we first upper bound $\mathcal{H}(P_\star)$: $\|\mathcal{H}(P_\star)\| = \|[I - B_\star(R + B_\star^\top P_\star B_\star)^{-1}B_\star^\top P_\star](\hat{A}-A_\star)\| \leqslant (1+\|B_\star\|^2\|P_\star\|\|(R+B_\star^\top P_\star B_\star)^{-1}\|)\varepsilon_{\mathrm{m}} \leqslant (1+\Upsilon_\star^2\|P_\star\|)\varepsilon_{\mathrm{m}}$, where the first identity follows from the matrix inverse lemma, and the final step follows from $R + B_\star^\top P_\star B_\star \succeq R \succeq I$ due to Assumption 2. Due to $R \succeq I$, letting $\varepsilon_s = \|S_\Delta\|$ leads to

$$\varepsilon_s = \|S_{B_\star} - S_{\hat{B}}\| \leqslant \varepsilon_{\mathrm{m}}^2 + 2\Upsilon_\star\varepsilon_{\mathrm{m}}, \text{ and} \quad (49)$$

$$\begin{aligned}\|\Delta\| &\leqslant (1+\Upsilon_\star^2\|P_\star\|)\varepsilon_{\mathrm{m}} + \varepsilon_s\|P_\star\|(\|\mathcal{H}(P_\star)\| + \|L_\star\|)\\ &\leqslant (1+\Upsilon_\star^2\|P_\star\|)\varepsilon_{\mathrm{m}} + \varepsilon_s\|P_\star\|\left[(1+\Upsilon_\star^2\|P_\star\|)\varepsilon_{\mathrm{m}} + \|P_\star\|\right]\\ &\leqslant \psi(\varepsilon_{\mathrm{m}}) \triangleq \varepsilon_{\mathrm{m}}\left\{1 + \Upsilon_\star^2\|P_\star\| + (\varepsilon_{\mathrm{m}} + 2\Upsilon_\star)\right.\\ &\qquad \left.\times\left[(\|P_\star\| + \Upsilon_\star^2\|P_\star\|^2)\varepsilon_{\mathrm{m}} + \|P_\star\|^2\right]\right\} \quad (50)\end{aligned}$$

where we have used $\|L_\star\| \leqslant \|P_\star\|$ from Lemma 12 in the second inequality, and $\psi(\varepsilon_{\mathrm{m}}) = O(\varepsilon_{\mathrm{m}})$ holds. Based on (10) and $\bar{\Phi}^{(j:i)}(P_\star) = \bar{\mathcal{L}}(P_\star)^{i-j} = (L_\star + \Delta)^{i-j}$, applying Lemma 10 concludes the proof.

E. Proof of Theorem 2

Based on (22), we first upper bound $\mathcal{M}$ and then combine it with Lemma 6 and Proposition 1. Recall (49), and for any $P \in \mathbb{S}_+^n$, $\|\mathcal{M}(P,P_\star)\| \leqslant \|(\hat{A}-A_\star)^\top P_\star(I+S_{B_\star}P_\star)^{-1}(\hat{A}-A_\star) + A_\star^\top P_\star(I+S_{B_\star}P_\star)^{-1}(\hat{A}-A_\star) + (\hat{A}-A_\star)^\top P_\star(I+S_{B_\star}P_\star)^{-1}A_\star\| + \|\hat{A}\|\|P_\star\|^2\|\mathcal{L}_{\hat{A},\hat{B}}(P)\|\varepsilon_s$

$$\begin{aligned}&\leqslant \|P_\star\|(\varepsilon_{\mathrm{m}}^2 + 2\Upsilon_\star\varepsilon_{\mathrm{m}}) + (\varepsilon_{\mathrm{m}} + \Upsilon_\star)^4\|P_\star\|^2(1+\|P\|)\varepsilon_s\\ &\leqslant \tilde{\psi}(\varepsilon_{\mathrm{m}}) \triangleq (\varepsilon_{\mathrm{m}}^2 + 2\Upsilon_\star\varepsilon_{\mathrm{m}})\\ &\qquad\times\left[\|P_\star\| + (\varepsilon_{\mathrm{m}}+\Upsilon_\star)^4\|P_\star\|^2(1+\|P_\star\|+\Upsilon_\star)\right] \quad (51)\end{aligned}$$

where (72) and Lemma 9 are used, and $\tilde{\psi}(\varepsilon_{\mathrm{m}}) = O(\varepsilon_{\mathrm{m}})$.

Then, if $i \geqslant n_{\mathrm{cr}}$, combining (22), (23), (25), and (51) shows

$$\|\hat{P}^{(i)} - P_\star\| \leqslant \tilde{\zeta}\varepsilon_{\mathrm{p}}\gamma_2^i\gamma_1^i + \tilde{\zeta}\tilde{\psi}\left(\sum_{k=0}^{n_{\mathrm{cr}}-1}\overline{\gamma}_2^k\gamma_1^k + \sum_{k=n_{\mathrm{cr}}}^{i-1}\gamma_2^k\gamma_1^k\right).$$

Similarly, when $i < n_{\mathrm{cr}}$, we have $\|\hat{P}^{(i)} - P_\star\| \leqslant \tilde{\zeta}\varepsilon_{\mathrm{p}}\overline{\gamma}_2^i\gamma_1^i + \tilde{\zeta}\tilde{\psi}\sum_{k=0}^{i-1}\overline{\gamma}_2^k\gamma_1^k$. Combining these two cases proves the result.

F. Proof of Theorem 3

Equation (26) can be written more compactly as

$$\widehat{E} = \tilde{\zeta}\left\{\left(\varepsilon_{\mathrm{p}} - \frac{\tilde{\psi}}{1-\gamma_1\gamma_2}\right)\gamma_2^{N-1}\gamma_1^{N-1} + \frac{\tilde{\psi}\gamma_1^{n_{\mathrm{cr}}}\gamma_2^{n_{\mathrm{cr}}}}{1-\gamma_1\gamma_2} + \tilde{\psi}\sum_{k=0}^{n_{\mathrm{cr}}-1}\overline{\gamma}_2^k\gamma_1^k\right\} \text{ if } N-1 \geqslant n_{\mathrm{cr}}. \quad (52)$$

We can quantify the change from $N-1 < n_{\mathrm{cr}}$ to $N-1 \geqslant n_{\mathrm{cr}}$ by computing $\Delta\widehat{E} \triangleq \widehat{E}(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},n_{\mathrm{cr}}) - \widehat{E}(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},n_{\mathrm{cr}}-1)$

$$\begin{aligned}&= \tilde{\zeta}\varepsilon_{\mathrm{p}}\left(\gamma_1^{n_{\mathrm{cr}}}\gamma_2^{n_{\mathrm{cr}}} - \gamma_1^{n_{\mathrm{cr}}-1}\overline{\gamma}_2^{n_{\mathrm{cr}}-1}\right) + \tilde{\psi}\tilde{\zeta}\gamma_1^{n_{\mathrm{cr}}-1}\overline{\gamma}_2^{n_{\mathrm{cr}}-1}\\ &= \tilde{\zeta}(\gamma_1\gamma_2)^{n_{\mathrm{cr}}-1}\left\{\varepsilon_{\mathrm{p}}[\gamma_2\gamma_1 - (\overline{\gamma}_2/\gamma_2)^{n_{\mathrm{cr}}-1}] + \tilde{\psi}(\overline{\gamma}_2/\gamma_2)^{n_{\mathrm{cr}}-1}\right\}.\end{aligned}$$

The above equation and $\gamma_2 \leqslant \overline{\gamma}_2$ from (25) show that

$$\Delta\widehat{E} \geqslant 0 \text{ iff } \varepsilon_{\mathrm{p}} - \frac{\tilde{\psi}}{1 - \frac{\gamma_1\gamma_2}{(\overline{\gamma}_2/\gamma_2)^{n_{\mathrm{cr}}-1}}} \leqslant 0. \quad (53)$$

The statement (a) follows from (29), (31), (52), and (53). For statement (b), if $0 > \varepsilon_{\mathrm{p}} - \tilde{\psi}/(1-\gamma_1\gamma_2)$, then (29), (31), and (52) show that g and $\widehat{E}$ are strictly increasing as $N \in \{1,\ldots,n_{\mathrm{cr}}\}$ increases, and they are also strictly increasing as $N \geqslant n_{\mathrm{cr}}+1$ increases. For the transition of the two regions of N, consider $\widehat{E}(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},0) - \widehat{E}(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},n_{\mathrm{cr}})$

$$\begin{aligned}&= \varepsilon_{\mathrm{p}}\tilde{\zeta} - \varepsilon_{\mathrm{p}}\tilde{\zeta}(\gamma_1\gamma_2)^{n_{\mathrm{cr}}} - \tilde{\zeta}\tilde{\psi}\frac{1-(\gamma_1\overline{\gamma}_2)^{n_{\mathrm{cr}}}}{1-\gamma_1\overline{\gamma}_2} = \tilde{\zeta}[1-(\gamma_1\gamma_2)^{n_{\mathrm{cr}}}]\\ &\quad\times\left\{\varepsilon_{\mathrm{p}} - \frac{\tilde{\psi}}{1-\gamma_1\gamma_2}\frac{(1-\gamma_1\gamma_2)[1-(\gamma_1\overline{\gamma}_2)^{n_{\mathrm{cr}}}]}{(1-\gamma_1\overline{\gamma}_2)[1-(\gamma_1\gamma_2)^{n_{\mathrm{cr}}}]}\right\}\\ &\leqslant \tilde{\zeta}[1-(\gamma_1\gamma_2)^{n_{\mathrm{cr}}}]\left(\varepsilon_{\mathrm{p}} - \frac{\tilde{\psi}}{1-\gamma_1\gamma_2}\right) < 0\end{aligned}$$

where the first inequality follows from Lemma 8. The above shows $\arg\min_N \widehat{E}(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},N-1) = \arg\min_N g(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},N) = 1$. Moreover, under the additional condition $\varepsilon_{\mathrm{p}} - \tilde{\psi}/[1-\gamma_1\gamma_2/\gamma_{\mathrm{ratio}}] < 0$, the last part of statement (b) is proved by observing (53).

For statement (c), if $\varepsilon_{\mathrm{p}} - \tilde{\psi}/(1-\gamma_1\gamma_2) > 0 > \varepsilon_{\mathrm{p}} - \tilde{\psi}/(1-\gamma_1\overline{\gamma}_2)$, (29) and (52) show that g is increasing as $N \in \{1,\ldots,n_{\mathrm{cr}}\}$ increases and decreasing as $N \geqslant n_{\mathrm{cr}}+1$ increases. This means that $\inf_N g(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},N)$ equals either $g(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},1)$ or $\lim_{N\to\infty} g(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},N)$ depending on the sign of $\widehat{E}(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},0) - \lim_{N\to\infty}\widehat{E}(\varepsilon_{\mathrm{m}},\varepsilon_{\mathrm{p}},N) =$

$$\tilde{\zeta}\left\{\varepsilon_{\mathrm{p}} - \frac{\tilde{\psi}}{1-\gamma_1\gamma_2}\left[\underbrace{\frac{(1-\gamma_1\gamma_2)[1-(\overline{\gamma}_2\gamma_1)^{n_{\mathrm{cr}}}]}{1-\gamma_1\overline{\gamma}_2} + (\gamma_1\gamma_2)^{n_{\mathrm{cr}}}}_{\mathcal{V}(n_{\mathrm{cr}})}\right]\right\}$$

which can be proved to be in the interval $[\tilde{\zeta}(\varepsilon_{\rm p}-\tilde{\psi}/(1-\gamma_1\overline{\gamma}_2)), \tilde{\zeta}(\varepsilon_{\rm p}-\tilde{\psi}/(1-\gamma_1\gamma_2))]$ as follows. The upper bound of this interval is proved by showing $\mathcal{V}(i+1)-\mathcal{V}(i) \geqslant 0$ and thus $\mathcal{V}(n_{\rm cr}) \geqslant \mathcal{V}(1) = 1$, and the lower bound can be proved similarly. The above proves statement (c). Statement (d) follows trivially from (29) and (52).

G. Proof of Proposition 2

Since $P = 0$, we let $\varepsilon_{\rm p} = \|P_\star\|$. In addition, $\varepsilon_{\rm m} \leqslant 1/(16\|P_\star\|^2)$ holds under the assumption, which leads to

$$\alpha_{\varepsilon_{\rm m}} \in [1, \sqrt{2}] \tag{54}$$

and thus $\overline{\gamma}_2 \leqslant \mu_\star < 1$. Therefore, we have

$$\widehat{E}_{\rm sta} \leqslant \tilde{\zeta}\left[\|P_\star\|\mu_\star^{N-1} + \tilde{\psi}/(1-\mu_\star)\right]. \tag{55}$$

Given the upper bound on $\varepsilon_{\rm m}$, $\tilde{\psi}$ in (51) satisfies $\tilde{\psi} \leqslant C_1\varepsilon_{\rm m}$ for some constant C_1 independent of N and $\varepsilon_{\rm m}$, and $\tilde{\zeta}$ can also be upper bounded by some constant C_2. This shows $\widehat{E}_{\rm sta} \leqslant C_2 \max\{\|P_\star\|, C_1/(1-\gamma_1\mu_\star)\}[\mu_\star^{N-1}+\varepsilon_{\rm m}]$. Combining the above inequality with (28) concludes the proof.

Appendix C
Proof of Proposition 1

To prove this result, the dual Riccati mapping is relevant

$$\mathcal{R}_{\rm dual}(P) \triangleq A_\star P(I+QP)^{-1}A_\star^\top + S_{B_\star} \tag{56}$$

with its fixed point $P_{\rm dual}$. Let $\hat{P}_{\rm dual}$ be a fixed point of (56) with $(A_\star, B_\star)$ replaced by $(\hat{A}, \hat{B})$, and $\hat{L}$ be the closed-loop matrix for $(\hat{A}, \hat{B})$ under the corresponding LQR controller.

The cases $i-j \geqslant n_{\rm cr}$ and $i-j < n_{\rm cr}$ are considered separately. If $i-j \geqslant n_{\rm cr}$, then the proof strategy is as follows. Due to $\Phi_{\hat{A},\hat{B}}^{(j:i)}(P) = \Phi_{\hat{A},\hat{B}}^{(0:i-j)}(\hat{P}^{(j)})$, Lemma 11(d) shows $[I + \mathcal{O}_{\hat{A},\hat{B}}^{(i-j)}(\hat{P})(\hat{P}^{(j)}-\hat{P})]\Phi_{\hat{A},\hat{B}}^{(j:i)}(P) = \hat{L}^{i-j}$ for $i \in \mathbb{Z}^+$. Moreover, [21, Lemma 4.1] shows that if $\mathcal{O}_{\hat{A},\hat{B}}^{(i-j)}(\hat{P})$ is positive definite, then $\|\Phi_{\hat{A},\hat{B}}^{(j:i)}(P)\| = \|[I + \mathcal{O}_{\hat{A},\hat{B}}^{(i-j)}(\hat{P})(\hat{P}^{(j)}-\hat{P})]^{-1}\hat{L}^{i-j}\|$

$$\leqslant \|[\mathcal{O}_{\hat{A},\hat{B}}^{(i-j)}(\hat{P})]^{-1}\|\|\hat{P}_{\rm dual}\|\|\hat{L}^{i-j}\|. \tag{57}$$

Then, the proof consists of the following analysis: 1) perturbation analysis of the closed-loop system $\hat{L}$, 2) perturbation analysis of the Gramian matrix: it shows that $\mathcal{O}_{\hat{A},\hat{B}}^{(i-j)}(\hat{P})$ is invertible for $i-j \geqslant n_{\rm cr}$ and provides an upper bound on $\|[\mathcal{O}_{\hat{A},\hat{B}}^{(i-j)}(\hat{P})]^{-1}\|$, and 3) perturbation analysis of $\|\hat{P}_{\rm dual}\|$. All perturbation bounds will be functions of $\varepsilon_{\rm m}$ and $\varepsilon_{\rm p}$.

The decay rate of the perturbed closed-loop system is obtained here. Given $Q \succeq I$, (10) and Lemma 12 lead to

$$\|\hat{L}^{i-j}\| \leqslant \sqrt{\overline{\sigma}(\hat{P})/\underline{\sigma}(Q)}\left(\sqrt{1-\underline{\sigma}(Q)/\overline{\sigma}(\hat{P})}\right)^{i-j}$$
$$\leqslant \sqrt{\alpha_{\varepsilon_{\rm m}}\beta_\star^{-1}}\left(\sqrt{1-\alpha_{\varepsilon_{\rm m}}^{-1}\beta_\star}\right)^{i-j}. \tag{58}$$

The remaining analysis is presented as follows.

A. Perturbation Analysis of the Gramian Matrix

If $i-j \geqslant n_{\rm cr}$, then $\mathcal{O}_{\hat{A},\hat{B}}^{(i-j)}(\hat{P}) \succeq \mathcal{O}_{\hat{A},\hat{B}}^{(n_{\rm cr})}(\hat{P})$ and

$$\mathcal{O}_{\hat{A},\hat{B}}^{(n_{\rm cr})}(\hat{P}) \succeq (\|R + \hat{B}^\top\hat{P}\hat{B}\|)^{-1}\mathcal{C}_{n_{\rm cr}}(\hat{L},\hat{B})\left[\mathcal{C}_{n_{\rm cr}}(\hat{L},\hat{B})\right]^\top.$$

The Gramian matrix $\mathcal{C}_{n_{\rm cr}}(\hat{L},\hat{B})[\mathcal{C}_{n_{\rm cr}}(\hat{L},\hat{B})]^\top$ is a perturbed version of $\mathcal{C}_{n_{\rm cr}}(L_\star,B_\star)[\mathcal{C}_{n_{\rm cr}}(L_\star,B_\star)]^\top$. Based on Lemma 12, $\|\hat{L}-L_\star\| \leqslant \|A_\star-\hat{A}\| + \|(\hat{B}-B_\star)(\hat{K}-K_\star)\| + \|(\hat{B}-B_\star)K_\star\| + \|B_\star(\hat{K}-K_\star)\|$

$$\leqslant \varepsilon_L \triangleq \varepsilon_{\rm m}\left[1 + 7\alpha_{\varepsilon_{\rm m}}^{7/2}\|P_\star\|^{7/2}(\varepsilon_{\rm m}+\Upsilon_\star) + \|P_\star\|^{1/2}\right] \tag{59}$$

which satisfies

$$\varepsilon_L = O(\varepsilon_{\rm m}). \tag{60}$$

Under Assumption 1, recall $\underline{\sigma}(\mathcal{C}_{n_{\rm cr}}(L_\star,B_\star)) \geqslant \nu_\star$ from (11). This, together with (10), (59), and [19, Lemma 6], shows $\underline{\sigma}(\mathcal{C}_{n_{\rm cr}}(\hat{L},\hat{B})) \geqslant \nu_\star - f_{\mathcal{C}}(\varepsilon_{\rm m})$, with $f_{\mathcal{C}}(\varepsilon_{\rm m}) \triangleq 3\varepsilon_L n_{\rm cr}^{3/2}\beta_\star^{-1}$

$$\times \max\left\{1, \sqrt{\beta_\star^{-1}}\varepsilon_L + \sqrt{1-\beta_\star}\right\}^{n_{\rm cr}-1}(\Upsilon_\star+1)$$
$$= O(\varepsilon_{\rm m}) \tag{61}$$

where ε_L is defined in (59), and the second equality follows from (60). Given $\nu_\star \geqslant 2f_{\mathcal{C}}(\varepsilon_{\rm m})$ in Assumption 4, $\mathcal{C}_{n_{\rm cr}}(\hat{L},\hat{B})$ is of full rank, and we have $\lambda_{\min}(\mathcal{C}_i(\hat{L},\hat{B})\mathcal{C}_i(\hat{L},\hat{B})^\top) \geqslant (\nu_\star - f_{\mathcal{C}}(\varepsilon_{\rm m}))^2$. Algebraic manipulations show $\|\hat{B}^\top\hat{P}\hat{B}\| \leqslant \|\hat{P}\|(\varepsilon_{\rm m}+\Upsilon_\star)^2$, which leads to $[\mathcal{O}_{\hat{A},\hat{B}}^{(n_{\rm cr})}(\hat{P})]^{-1} \preceq (\|R + \hat{B}^\top\hat{P}\hat{B}\|)(\nu_\star - f_{\mathcal{C}}(\varepsilon_{\rm m}))^{-2}I$

$$\preceq \left[\overline{\sigma}(R) + \alpha_{\varepsilon_{\rm m}}\|P_\star\|(\varepsilon_{\rm m}+\Upsilon_\star)^2\right](\nu_\star - f_{\mathcal{C}}(\varepsilon_{\rm m}))^{-2}I \tag{62a}$$
$$= O(1)I \tag{62b}$$

where (62b) follows from the fact that the right-hand side of (62a) approaches a constant as $\varepsilon_{\rm m} \to 0$.

B. Perturbation Analysis of the Dual Riccati Equation

Given $Q \succeq I$ under Assumption 2, let $Q^{1/2}$ denote the symmetric positive-definite square root of Q. We further define

$$K_{\rm dual} \triangleq (I + Q^{\frac{1}{2}}P_{\rm dual}Q^{\frac{1}{2}})^{-1}Q^{\frac{1}{2}}P_{\rm dual}A_\star^\top$$
$$L_{\rm dual} \triangleq A_\star^\top - Q^{\frac{1}{2}}K_{\rm dual}. \tag{63}$$

As $L_{\rm dual}$ is Schur stable, let $\rho_{\rm dual} \in (0,1)$ represent its decay rate such that $\|L_{\rm dual}^i\| \leqslant c\rho_{\rm dual}^i$ for some $c \geqslant 1$ and for any $i \in \mathbb{Z}^+$. Then, applying [19, Prop. 2] leads to the following.

Lemma 14: It holds that

$$\|\hat{P}_{\rm dual} - P_{\rm dual}\| \leqslant O(1)\frac{\rho_{\rm dual}^2}{1-\rho_{\rm dual}^2}\Upsilon_\star(\Upsilon_\star+1)^3\|P_{\rm dual}\|^2\varepsilon_{\rm m}$$

if Assumptions 2, 3, and $\underline{\sigma}(P_{\rm dual}) \geqslant 1$ hold, and if

$$\varepsilon_{\rm m} \leqslant \alpha_\star \triangleq O(1)\frac{(1-\rho_{\rm dual}^2)^2}{\rho_{\rm dual}^4(\Upsilon_\star+1)^5\|P_{\rm dual}\|^2}$$
$$\times \min\{(\|L_{\rm dual}\|+1)^{-2}, \|P_{\rm dual}\|^{-1}\}. \tag{64}$$

Proof: Based on Assumptions 2 and 3, $(A_\star^\top, Q^{1/2})$ is stabilizable, and $(A_\star^\top, R^{-1/2}B_\star^\top)$ is observable, showing $\mathcal{R}_{\rm dual}$ has a unique positive-definite fixed point $P_{\rm dual}$. Moreover, (49) and the assumption $\Upsilon_\star \geqslant \varepsilon_{\rm m}$ imply $\max\{\|\hat{A}^\top - A_\star^\top\|, \|Q^{1/2} -$

$Q^{1/2}\|, \|S_{\hat{B}} - S_{B_\star}\|\} \leqslant 3\Upsilon_\star \varepsilon_{\text{m}}$. Then, applying [19, Prop. 2] proves the current lemma. ■

In Lemma 14, $O(1)$ is a constant and is independent of ε_{m}. Also, the Riccati perturbation result from [20] does not apply to the dual Riccati equation (56) as it requires $S_{B_\star} = B_\star R^{-1} B_\star^\top \succeq I$, which is not met by many practical situations.

Remark 7: Assumption $\underline{\sigma}(P_{\text{dual}}) \geqslant 1$ relates to controllability of $(L_{\text{dual}}^\top, B_\star R^{1/2})$ with L_{dual} defined in (63). Comparing Lemma 11(b) and (56), $P_{\text{dual}} \succeq \sum_{i=0}^{\infty} (L_{\text{dual}}^\top)^i S_{B_\star} L_{\text{dual}}^i$ holds. Therefore, $\underline{\sigma}(P_{\text{dual}}) \geqslant 1$ can be guaranteed by lower bounding the controllability Gramian of $(L_{\text{dual}}^\top, B_\star R^{1/2})$.

C. Final Step When $i - j \geqslant n_{cr}$

When $i - j \geqslant n_{\text{cr}}$, the perturbation analysis can be concluded by combining (57), (58), (62a), and Lemma 14

$$\|\Phi_{\hat{A},\hat{B}}^{(j:i)}(P)\| \leqslant \zeta \left(\sqrt{1 - \alpha_{\varepsilon_{\text{m}}}^{-1} \beta_\star}\right)^{i-j}, \text{ with} \tag{65}$$

$$\zeta \triangleq \left[\overline{\sigma}(R) + \sqrt{2}\|P_\star\|(2\Upsilon_\star)^2\right] [1/2\nu_\star]^{-2} \sqrt{\sqrt{2}\beta_\star^{-1}} \times \left[\|P_{\text{dual}}\| + O(1)\frac{\rho_{\text{dual}}^2}{1 - \rho_{\text{dual}}^2} \Upsilon_\star (\Upsilon_\star + 1)^3 \|P_{\text{dual}}\|^2 \Upsilon_\star\right] \tag{66}$$

where we have used (54) and $\Upsilon_\star \geqslant \max\{\varepsilon_{\text{m}}, \varepsilon_{\text{p}}\}$.

D. Situation Where $i - j < n_{cr}$

If $i - j < n_{\text{cr}}$, $\mathcal{O}_{\hat{A},\hat{B}}^{(i-j)}(\hat{P})$ may not be invertible, and thus, instead of exploiting controllability and (57), we consider the alternative bound in (69), which holds for stabilizable systems.

Appendix D
Proof of Theorem 4

Based on Lemma 5, an essential step to prove Theorem 4 is establishing the convergence rate of the state-transition matrix for stabilizable systems. The stability analysis of time-varying systems in Lemma 13 is a fundamental tool. To exploit it, we first provide a uniform bound on the Riccati iterations as an explicit function of $P_\star$. The proof of Lemma 15 is in [41].

Lemma 15: Let $P \in \mathbb{S}_+^n$ satisfy $\|P - P_\star\| \leqslant \varepsilon_{\text{p}}$, and let $P_\star^{(k)}$ denote $\mathcal{R}_{A_\star,B_\star}^{(k)}(P)$. If Assumptions 2 and 3 hold, then

$$P_\star^{(k)} \preceq \left[\|P_\star\| + \varepsilon_{\text{p}} \beta_\star^{-1}\right] I \quad \forall k \in \mathbb{Z}^+. \tag{67}$$

Lemma 16: For any $i, j \in \mathbb{Z}^+$ with $i \geqslant j$, if Assumptions 2 and 3 hold, then $\|\Phi_{A_\star,B_\star}^{(j:i)}(P)\| \leqslant \Upsilon_\star^3 (1 + \overline{\sigma}(P_\star) + \varepsilon_{\text{p}})$

$$\times \sqrt{\left(\overline{\beta}(\varepsilon_{\text{p}})\right)^{-1} \left(1 - \overline{\beta}(\varepsilon_{\text{p}})\right)^{-1}} \left(\sqrt{1 - \overline{\beta}(\varepsilon_{\text{p}})}\right)^{i-j}. \tag{68}$$

If Assumption 4(a) holds additionally, then

$$\|\Phi_{\hat{A},\hat{B}}^{(j:i)}(P)\| \leqslant \overline{\zeta} \left(\sqrt{1 - \alpha_{\varepsilon_{\text{m}}}^{-1} \overline{\beta}(\varepsilon_{\text{p}})}\right)^{i-j} \tag{69}$$

with $\overline{\zeta}$ defined in (73).

Proof: As $\{P_\star^{(k)}\}_{k=j}^{i-1}$ is a positive-definite matrix sequence when $j \geqslant 1$ and satisfies $P_\star^{(k)} \succeq Q \succeq I$, we first consider $j \geqslant 1$. For $k \in [j, i]$ and from Lemma 11(b), $V(k, \bar{x}) = \bar{x}^\top P_\star^{(i+j-k)} \bar{x}$ is a time-varying Lyapunov function, satisfying $V(k+1, \bar{x}_{k+1}) - V(k, \bar{x}_k) \leqslant -\bar{x}_k^\top Q \bar{x}_k$, for the time-varying nominal system $\bar{x}_{k+1} = \mathcal{L}_{A_\star,B_\star}(P_\star^{(i+j-1-k)})\bar{x}_k$. This nominal system admits the state-transition matrix of interest, i.e., $\bar{x}_i = \Phi_{A_\star,B_\star}^{(j:i)}(P)\bar{x}_j$. Then, when $j \geqslant 1$, we have $\underline{\sigma}(Q)\|x\|^2 \leqslant V(k, x) \leqslant [\|P_\star\| + \varepsilon_{\text{p}}\beta_\star^{-1}]\|x\|^2$ based on (67). Applying Lemma 13 shows

$$\|\Phi_{A_\star,B_\star}^{(j:i)}(P)\| \leqslant \sqrt{\overline{\beta}^{-1}} \left(\sqrt{1 - \overline{\beta}}\right)^{i-j}, \text{ for } j \geqslant 1 \tag{70}$$

where we recall $\overline{\beta}(\varepsilon_{\text{p}}) = \underline{\sigma}(Q)/[\|P_\star\| + \varepsilon_{\text{p}}\beta_\star^{-1}]$. Equation (70) proves (68) for $j \geqslant 1$ due to $\Upsilon_\star^3(1 + \overline{\sigma}(P_\star) + \varepsilon_{\text{p}})\sqrt{(1 - \overline{\beta}(\varepsilon_{\text{p}}))^{-1}} \geqslant 1$.

However, when $j = 0$, recall that $P = P_\star^{(0)}$ is positive semidefinite by Assumption 2; thus, it cannot be lower bounded by a positive-definite matrix. Then, we exploit

$$\|\Phi_{A_\star,B_\star}^{(0:i)}(P)\| \leqslant \|\mathcal{L}_{A_\star,B_\star}(P)\| \|\Phi_{A_\star,B_\star}^{(1:i)}(P)\| \tag{71}$$

for $i \geqslant 1$ and

$$\|\mathcal{L}_{A_\star,B_\star}(P)\| \leqslant \|A_\star\| + \|B_\star\| \|\mathcal{K}_{A_\star,B_\star}(P_\star^{(0)})\| \leqslant \|A_\star\| + \|B_\star\|^2 \|A_\star\| \|P\| \leqslant \Upsilon_\star^3 (1 + \|P\|). \tag{72}$$

Combining the above bound, (70), (71), and $\|P\| = \|P - P_\star + P_\star\| \leqslant \varepsilon_{\text{p}} + \|P_\star\|$ proves (68) for $j = 0$. The case $i = 0$ holds trivially.

Then, if $(\hat{A}, \hat{B})$ is considered instead, $\|\Phi_{\hat{A},\hat{B}}^{(j:i)}(P)\|$ can be upper bounded as in (68), with $P_\star$ replaced by $\hat{P}$. Then, combining this bound with $\|\hat{P}\| \leqslant \alpha_{\varepsilon_{\text{m}}}\|P_\star\|$ from Lemma 12 and $\alpha_{\varepsilon_{\text{m}}} \in [1, \sqrt{2}]$ from (54) implies (69) with

$$\overline{\zeta} \triangleq (2\Upsilon_\star)^3 (1 + \sqrt{2}\|P_\star\| + \Upsilon_\star) \times \sqrt{\sqrt{2}(\|P_\star\| + \Upsilon_\star \beta_\star^{-1})(1 - \beta_\star)^{-1}} \tag{73}$$

where we have used (54), the fact $(\sqrt{1 - \overline{\beta}})^{i-j-1} \leqslant (\sqrt{1 - \alpha_{\varepsilon_{\text{m}}}^{-1}\overline{\beta}})^{i-j-1}$, and $\Upsilon_\star \geqslant \max\{\varepsilon_{\text{m}}, \varepsilon_{\text{p}}\}$. ■

Remark 8: The stability of $\Phi_{A_\star,B_\star}^{(0:i)}(P)$ is a classical result, e.g., see [10] and [42, Lemma 2.9]. The main challenge in Lemma 16 is to derive the decay rate explicitly as a function of $P_\star$, such that the perturbation result in (69) is obtained.

Combining Lemmas 5, 6, and 16 proves Theorem 4.

References

[1] M. Schwenzer, M. Ay, T. Bergs, and D. Abel, "Review on model predictive control: An engineering perspective," *Int. J. Adv. Manuf. Technol.*, vol. 117, no. 5–6, pp. 1327–1349, 2021.

[2] S. Katayama, M. Murooka, and Y. Tazaki, "Model predictive control of legged and humanoid robots: Models and algorithms," *Adv. Robot.*, vol. 37, no. 5, pp. 298–315, 2023.

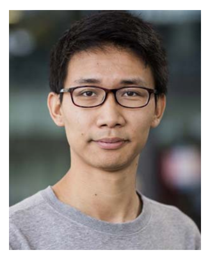
[3] D. Q. Mayne, "Model predictive control: Recent developments and future promise," *Automatica*, vol. 50, no. 12, pp. 2967–2986, 2014.

[4] L. Grüne and J. Pannek, *Nonlinear Model Predictive Control: Theory and Algorithms*. Berlin, Germany: Springer, 2011.

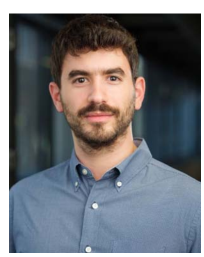
[5] L. Grüne and S. Pirkelmann, "Economic model predictive control for time-varying system: Performance and stability results," *Optim. Control Appl. Methods*, vol. 41, no. 1, pp. 42–64, 2020.

[6] J. Köhler, M. Zeilinger, and L. Grüne, "Stability and performance analysis of NMPC: Detectable stage costs and general terminal costs," *IEEE Trans. Autom. Control*, vol. 68, no. 10, pp. 6114–6129, Oct. 2023.

[7] D. Bertsekas, "Newton's method for reinforcement learning and model predictive control," *Results Control Optim.*, vol. 7, 2022, Art. no. 100121.
[8] F. Moreno-Mora, L. Beckenbach, and S. Streif, "Predictive control with learning-based terminal costs using approximate value iteration," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 3874–3879, 2023.
[9] D. Bertsekas, "Reinforcement Learning and Optimal Control," Belmont, MA, USA: *Athena Sci.*, 2019.
[10] R. R. Bitmead and M. Gevers, "Riccati difference and differential equations: Convergence, monotonicity and stability," in *The Riccati Equation*. Berlin, Germany: Springer, 1991, pp. 263–291.
[11] F. Moreno-Mora, L. Beckenbach, and S. Streif, "Performance bounds of adaptive MPC with bounded parameter uncertainties," *Eur. J. Control*, vol. 68, 2022, Art. no. 100688.
[12] D. Muthirayan, J. Yuan, D. Kalathil, and P. P. Khargonekar, "Online learning for predictive control with provable regret guarantees," in *Proc. 61st IEEE Conf. Decis. Control*, 2022, pp. 6666–6671.
[13] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, "Model learning predictive control in nonlinear dynamical systems," in *Proc. 60th IEEE Conf. Decis. Control*, 2021, pp. 757–762.
[14] I. Dogan, Z.-J. M. Shen, and A. Aswani, "Regret analysis of learning-based MPC with partially-unknown cost function," *IEEE Trans. Autom. Control*, vol. 69, no. 5, pp. 3246–3253, May 2024.
[15] B. Recht, "A tour of reinforcement learning: The view from continuous control," *Annu. Rev. Control Robot. Auton. Syst.*, vol. 2, pp. 253–279, 2019.
[16] N. Matni, A. Proutiere, A. Rantzer, and S. Tu, "From self-tuning regulators to reinforcement learning and back again," in *Proc. 58th IEEE Conf. Decis. Control*, 2019, pp. 3724–3740.
[17] A. Tsiamis, I. Ziemann, N. Matni, and G. J. Pappas, "Statistical learning theory for control: A finite-sample perspective," *IEEE Control Syst*, vol. 43, no. 6, pp. 67–97, Jun. 2023.
[18] S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu, "On the sample complexity of the linear quadratic regulator," *Found. Comput. Math.*, vol. 20, no. 4, pp. 633–679, 2020.
[19] H. Mania, S. Tu, and B. Recht, "Certainty equivalence is efficient for linear quadratic control," in *Proc. 33rd Int. Conf. Neural Inf. Process. Syst.*, 2019, pp. 10154–10164.
[20] M. Simchowitz and D. Foster, "Naive exploration is optimal for online LQR," in *Proc. 37th Int. Conf. Mach. Learn.*, 2020, pp. 8937–8948.
[21] P. Del Moral and E. Horton, "A note on Riccati matrix difference equations," *SIAM J. Control Optim.*, vol. 60, no. 3, pp. 1393–1409, 2022.
[22] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. Hoboken, NJ, USA: MIT Press, 2022.
[23] M. Simchowitz, H. Mania, S. Tu, M. I. Jordan, and B. Recht, "Learning without mixing: Towards a sharp analysis of linear system identification," in *Prof. 31st Conf. Learn. Theory*, 2018, pp. 439–473.
[24] T. Sarkar and A. Rakhlin, "Near optimal finite time identification of arbitrary linear dynamical systems," in *Proc. 36th Int. Conf. Mach. Learn.*, 2019, pp. 5610–5618.
[25] D. Bertsekas, *Dyn. Program. and Optimal Control*, vol. I, Nashua, NH, USA: Athena Sci., 2012.
[26] B. Anderson and J. B. Moore, *Optimal Filtering*. Hoboken, NJ, USA: Prentice-Hall, 1979.
[27] J. Lawson and Y. Lim, "A Birkhoff contraction formula with applications to Riccati equations," *SIAM J. Control Optim.*, vol. 46, no. 3, pp. 930–951, 2007.
[28] B. Hassibi, A. H. Sayed, and T. Kailath, *Indefinite-Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories*, Philadelphia, PA, USA: SIAM, 1999.
[29] W. J. Rugh, *Linear System Theory*. Hoboken, NJ, USA: Prentice-Hall, 1996.
[30] R. Zhang, Y. Li, and N. Li, "On the regret analysis of online LQR control with predictions," in *Proc. Am Control Conf.*, 2021, pp. 697–703.
[31] M. Konstantinov, I. Poptech, and V. Angelova, "Conditioning and sensitivity of the difference matrix Riccati equation," in *Proc. Am Control Conf.*, 1995, pp. 466–466.
[32] B. De Schutter and T. Van Den Boom, "Model predictive control for max-plus-linear discrete event systems," *Automatica*, vol. 37, no. 7, pp. 1049–1056, 2001.
[33] M. Fruchard, G. Allibert, and E. Courtial, "Choice of the control horizon in an NMPC strategy for the full-state control of nonholonomic systems," in *Proc. Am Control Conf.*, 2012, pp. 4149–4154.
[34] S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar, "Reinforcement learning with fast stabilization in linear dynamical systems," in *Proc. 25th Int. Conf. Artif. Intell. Stat.*, 2022, pp. 5354–5390.
[35] C. Yu, G. Shi, S. J. Chung, Y. Yue, and A. Wierman, "The power of predictions in online control," in *Proc. 34th Int. Conf. Adv. Neural Inf. Process. Syst.*, 2020, pp. 1994–2004.
[36] A. Athrey, O. Mazhar, M. Guo, B. De Schutter, and S. Shi, "Regret analysis of learning-based linear quadratic Gaussian control with additive exploration," in *Proc. 22nd Eur. Control Conf.*, 2024, pp. 1795–1801.
[37] A. J. Krener, "Adaptive horizon model predictive control," *IFAC-PapersOnLine*, vol. 51, no. 13, pp. 31–36, 2018.
[38] E. Bøhn, S. Gros, S. Moe, and T. A. Johansen, "Optimization of the model predictive control meta-parameters through reinforcement learning," *Eng. Appl. Artif. Intell.*, vol. 123, 2023, Art. no. 106211.
[39] S. Chen, K. Werling, A. Wu, and C. K. Liu, "Real-time model predictive control and system identification using differentiable simulation," *IEEE Robot. Autom. Lett.*, vol. 8, no. 1, pp. 312–319, Jan. 2022.
[40] C. Liu, S. Shi, and B. De Schutter, "Certainty-equivalence model predictive control: Stability, performance, and beyond," 2024, *arXiv:2412.10625*.
[41] S. Shi, A. Tsiamis, and B. De Schutter, "Suboptimality analysis of receding horizon quadratic control with unknown linear systems and its applications in learning-based control," 2025, *arXiv:2301.07876*.
[42] B. D. O. Anderson et al., *Stability of Adaptive Syst.: Passivity and Averaging Anal.*. Cambridge, MA, USA: MIT Press, 1986.



Shengling Shi is an Assistant Professor with the Delft Center for Systems and Control (DCSC), Delft University of Technology (TU Delft), Delft, The Netherlands. He was a Postdoctoral Researcher with the Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA, and with DCSC, TU Delft. His research interests include system identification, model predictive control, and their applications.



Anastasios Tsiamis received the Ph.D. degree with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA. He is currently a Postdoctoral Researcher with the Automatic Control Laboratory, ETH Zürich, Zurich, Switzerland. His research focuses on the intersection of control theory and machine learning. Dr. Tsiamis was a finalist for the IFAC Young Author Prize in IFAC 2017 World Congress and a finalist for the Best Student Paper Award in ACC 2019.

Bart De Schutter is currently a Full Professor with the Delft Center for Systems and Control, Delft University of Technology, Delft, The Netherlands. His research interests include control of discrete-event and hybrid systems, multilevel and distributed control, intelligent transportation, and infrastructure systems. Dr. De Schutter is a Senior Editor of IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS.