Learning-based Optimal Control for Linear Systems with Model Uncertainties

Zitong Wang *, **Xuda Ding** *, **Xiaoming Duan** *, **Jianping He** *

* *Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240. E-mail:{wangzitong, dingxuda, xduan, jphe}@sjtu.edu.cn*

Abstract: The problem of optimal control with unknown dynamics is important and challenging. Most existing learning-based methods usually do not utilize the prior knowledge of the model and choose to learn the system model from scratch, leading to high identification costs and unbounded identification time. In practice, prior knowledge of the model may be available. Therefore, in this paper, we investigate the Linear Quadratic Gaussian (LQG) control problem for an unknown system but with prior knowledge of model uncertainties. We develop an effective learning procedure to identify the optimal controller parameters, leading to the least possible cost. Specifically, we analyze the performance of the optimal observer under unknown system dynamics and reveal the relationship between the regression error and the observer performance. We propose a learning procedure to design the optimal controller incorporating prior knowledge of the system. Simulations are conducted to illustrate the effectiveness of our controller design.

Keywords: system identification, linear systems, unknown systems, LQG control

1. INTRODUCTION

Many control methods depend on accurate system models, which are usually difficult to obtain in practice (Oymak and Ozay, 2021; Tsiamis et al., 2020). Control for unknown systems is of more importance and has drawn considerable interest from the community. In typical industrial applications, approximate models for the plants from the same batch may be available, which can be treated as the prior knowledge about the plants. Due to model uncertainties, the approximate system model could be inaccurate, leading to control performance degradation and necessitating data-driven control methods. Related work on data-driven control can be roughly divided into two categories: regressing the control input or learning the system model. Since the latter approach is more prevalent and has broader applications, recent studies focused on identifying the system model and then designing controllers (Lu and Mo, 2021; Jha et al., 2017; Dean et al., 2020). The various inevitable noises usually result in inaccuracy in the system identification process, affect the observer and controller design, and deteriorate the control performance. As a result, how to identify the unknown system and design a controller based on the prior knowledge of model uncertainties become essential research questions.

There are fruitful research results on unknown system identification and optimal observer design (Oymak and Ozay, 2021; Tsiamis et al., 2020). The difficulties in general system model regression problems lie in two parts: data volume and noise (Sabzekar and Hasheminejad, 2021). Various model regression algorithms require massive data samples, leading to a slow identification speed for unknown systems. Noisy data and outliers also limit the generalization capability and the identification accuracy of the models. The authors in (Ding et al., 2021) use support vector regression to identify system parameters and design feedback gain to ensure stability. Another work (Oymak and Ozay, 2019) learns the Markov parameters of an unknown system from a single input/output trajectory using the classical Ho-Kalman algorithm. Researchers proposed various mechanisms to overcome the difficulties in state estimation caused by inaccurate system dynamics. Since the Kalman filter serves as an optimal observer and has good properties (Xie and Soh, 1994), the authors in (Tsiamis et al., 2020) manage to solve the Kalman filter design problem with unknown systems and show the end-to-end sample complexity bounds. In situations when the system dynamics are unknown, combining system identification with optimal control is critical to guarantee the performance of the system.

Recently, the well-studied optimal Linear Quadratic Regulator (LQR) has been re-investigated in the data-driven control and reinforcement learning framework (Dean et al., 2020). Furthermore, Linear Quadratic Gaussian (LQG) control deals with linear dynamical systems driven by additive white Gaussian noises. To help with the controller design, parameterization methods serve to find the parameters of the controller, e.g., Youla Parameterization (Youla et al., 1976), System-Level Parameterization (SLP) (Wang et al., 2019), and Input-Output Parameterization (IOP) (Furieri et al., 2019). In addition, many works study the control problem with system identification under unknown

⋆ This work was supported by the National Natural Science Foundation of China under Grants 61973218, the Open Fund of State Key Laboratory of Ocean Engineering under Grant GKZD010084, and Shanghai Pujiang Program under grant 22PJ1404900.

dynamics (Jha et al., 2017; Dean et al., 2020; Lu and Mo, 2021; Zheng et al., 2021). For example, the authors in (Dean et al., 2020) propose Coarse-ID control based on LQR, which employs a small amount of experimental data and designs a controller using both the system model and uncertainty estimate. The authors identify the system parameters via the least squares method and provide end-to-end bounds on the relative error in the control. However, their approach assumes that the states of the system can be fully observed, which limits its practical applicability. In (Lu and Mo, 2021), the authors focus on the linear quadratic dual control problem. They identify the system parameters and optimize the control objective online with Markov parameters. The authors in (Zheng et al., 2021) present a method to identify a model from a single input-output trajectory of finite length and use IOP to design a robust controller that can deal with model uncertainties.

In practice, some prior knowledge can be useful (Berberich et al., 2022; Sinquin, 2019), such as the approximation of the plant dynamics and a stable controller. The most straightforward method for the control synthesis is directly using the approximate stable controller. The identification procedure could be skipped when the identification takes too long, and the optimality of the controller is not a concern. Motivated by the above observations, this paper studies whether it is worth learning the dynamic parameters for optimal control of an unknown linear system when prior knowledge is available. Overall, we propose a learning procedure that utilizes the least squares (LS) method to regress the system model, the Kalman filter to design the optimal observer, and the LQG method to retrieve the controller based on the former two steps. The challenges lie in i) characterizing the connection between system identification error and control performance, and ii) utilizing the prior information of the system to decide the required identification time. The main contributions of our work are listed as follows:

- We combine the LS method, the observer design, and the LQG controller to design an optimal controller for linear systems with model uncertainties. The performance of the observer under unknown system dynamics is derived. The relationship between the regression error and the observer performance is revealed.
- We propose a learning procedure for the controller design, which incorporates the prior knowledge of model uncertainties and stabilizes the system while improving the performance of the controller. Numerical experiments have been conducted to illustrate the effectiveness of the procedure.

The rest of the paper is organized as follows: Sec. 2 provides some preliminary knowledge and formulates the problem. The proposed design and its performance analysis are presented in Sec. 3. Sec. 4 shows the simulation and comparison results. Finally, Sec. 5 concludes the work.

2. PRELIMINARIES

Notation: We use slim symbols to denote scalars and bold symbols to denote vectors and matrices. We denote $\lambda(\boldsymbol{M})$ as the eigenvalues of the matrix $\boldsymbol{M}$. The $\mathcal{H}_2$ system norm is defined by $\|\boldsymbol{\Phi}\|_{\mathcal{H}_2}^2 \triangleq \sum_{t=0}^{\infty} \|\Phi_t\|_F^2$, where $\boldsymbol{\Phi} = \sum_{t=0}^{\infty} \Phi_t z^{-t}$ is the frequency representation of Φ.

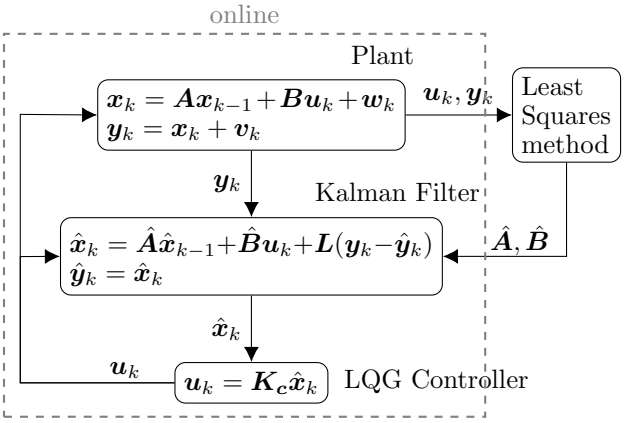


Fig. 1. System block diagram

2.1 System Models

Suppose the plant is described by a linear model:

$$\begin{cases} \boldsymbol{x}_{k+1} = \boldsymbol{A}\boldsymbol{x}_k + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{w}_k, \\ \boldsymbol{y}_k = \boldsymbol{x}_k + \boldsymbol{v}_k, \end{cases} \tag{1}$$

where $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{n\times p}$ are the system matrices, $\boldsymbol{x}_k$ and $\boldsymbol{y}_k$ are the state and the output, and $\boldsymbol{u}_k = \pi(\boldsymbol{y}_k, \cdots, \boldsymbol{y}_0)$ represents the designed input. The process noise $\boldsymbol{w}_k$ and the observation noise $\boldsymbol{v}_k$ are assumed to be i.i.d. zero mean Gaussian noises, with positive definite covariance matrices $\boldsymbol{Q_w} = \sigma_w^2 \boldsymbol{I}_n$ and $\boldsymbol{R_v} = \sigma_v^2 \boldsymbol{I}_n$, respectively.

The cost of the LQG control problem is defined as:

$$J_c = \min_{\boldsymbol{u}_0, \boldsymbol{u}_1, \cdots} \lim_{T\to\infty} \mathbb{E}\left[\frac{1}{T}\sum_{k=0}^{T} \left(\boldsymbol{y}_k^\top \boldsymbol{Q}\boldsymbol{y}_k + \boldsymbol{u}_k^\top \boldsymbol{R}\boldsymbol{u}_k\right)\right], \tag{2}$$

where $\boldsymbol{Q}$ and $\boldsymbol{R}$ are weight matrices of the states and the control inputs in the cost function, respectively. The optimal solution to (2) is $\boldsymbol{u}_k = \boldsymbol{K_c}\hat{\boldsymbol{x}}_k$ with a fixed $p \times n$ matrix $\boldsymbol{K_c}$ and $\hat{\boldsymbol{x}}_k$ is the estimated state from the observation $(\boldsymbol{y}_0, \cdots, \boldsymbol{y}_k)$ using the Kalman filter. The quality of the estimated states generated by the Kalman filter depends on the system parameters and the noise statistics (Anderson and Moore, 2012). The optimal linear feedback control gain $\boldsymbol{K_c^\star}$ of the LQG controller is determined by

$$\boldsymbol{K_c^\star} = -\left(\boldsymbol{R} + \boldsymbol{B}^\top \boldsymbol{P}^\star \boldsymbol{B}\right)^{-1} \boldsymbol{B}^\top \boldsymbol{P}^\star \boldsymbol{A},$$

where $\boldsymbol{P}^\star$ is the solution to the discrete-time algebraic Riccati equation:

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

2.2 Problem Formulation

The system block diagram is shown in Fig. 1. The system consists of the plant, the Kalman filter, the LS-based estimator, and the LQG controller. Specifically, the plant generates the states and provides the outputs to the estimator where the system dynamics and states are estimated. The Kalman filter estimates the states based on the outputs of the plant and the identified system dynamics. Finally, the LQG controller uses the estimated states $\hat{\boldsymbol{x}}_k$ to control the linear plant.

Suppose the system matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are unknown, with their approximations (prior knowledge) being $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{B}}$,

respectively. The initial LQG controller designed based on $(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}})$ is $\tilde{\boldsymbol{K}}_c$.

Assumption 1. $(\boldsymbol{A}, \boldsymbol{B})$ is stabilizable, and the initial controller stabilizes the system, i.e., $\rho(\boldsymbol{A} - \boldsymbol{B}\tilde{\boldsymbol{K}}_c) < 1$.

Assumption 2. Suppose $\|\tilde{\boldsymbol{A}} - \boldsymbol{A}\| \leq \varepsilon_{\boldsymbol{A}}$ and $\|\tilde{\boldsymbol{B}} - \boldsymbol{B}\| \leq \varepsilon_{\boldsymbol{B}}$, where $\varepsilon_{\boldsymbol{A}}$ and $\varepsilon_{\boldsymbol{B}}$ are the upper bounds on the approximation errors.

Mathematically, the system identification error is the deviation between the actual matrices $\boldsymbol{A}, \boldsymbol{B}$ and the estimated matrices $\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}$ based on the LS method. The LS estimators $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ are unbiased, and their variances depend on the amplitude of the noise (Li et al., 2021). Denote $\hat{\boldsymbol{A}} = \boldsymbol{A} + \Delta_{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}} = \boldsymbol{B} + \Delta_{\boldsymbol{B}}$. We further determine the relationship between the system identification error and the number of samples for learning as $\|\Delta_{\boldsymbol{A}}\|, \|\Delta_{\boldsymbol{B}}\| \sim \mathcal{O}(1/\sqrt{N})$.

In situations when system dynamics are unknown, the discrepancy between the system matrices and the estimated matrices results in inaccuracy in the Kalman filter design. These uncertainties are fed as inputs to the filter, and their effects are described in the robustness analysis of the Kalman filter (Lu and Niu, 2014). Suppose that with a high probability, the error bounds are respectively $\|\Delta_{\boldsymbol{A}}\| \leq \epsilon_{\boldsymbol{A}}$ and $\|\Delta_{\boldsymbol{B}}\| \leq \epsilon_{\boldsymbol{B}}$. Here, $\epsilon_{\boldsymbol{A}}$ and $\epsilon_{\boldsymbol{B}}$ refer to the upper bounds of system identification errors, which are different from $\varepsilon_{\boldsymbol{A}}$ and $\varepsilon_{\boldsymbol{B}}$ mentioned in Assumption 2. Similar to (Zheng et al., 2021) and (Dean et al., 2020), we consider a robust variant of the LQG problem that seeks to minimize the worst-case LQG performance of the system given the norm bounds on $\Delta_{\boldsymbol{A}}$ and $\Delta_{\boldsymbol{B}}$:

$$\min_{\boldsymbol{K}_c} \sup_{\substack{\|\Delta_{\boldsymbol{A}}\| \leq \epsilon_{\boldsymbol{A}} \\ \|\Delta_{\boldsymbol{B}}\| \leq \epsilon_{\boldsymbol{B}}}} \lim_{T \to \infty} \mathbb{E}\left[\frac{1}{T}\sum_{k=0}^{T}\left(\boldsymbol{y}_k^\top \boldsymbol{Q}\boldsymbol{y}_k + \boldsymbol{u}_k^\top \boldsymbol{R}\boldsymbol{u}_k\right)\right]. \quad (3)$$

Learning-based methods are generally used to minimize the control cost in the LQG problem (3), which usually have high identification costs. With the prior knowledge of model uncertainties and the proper learning procedure of the controller design, the balance between the learning cost and the control cost can be achieved.

Finally, the performance of the prior knowledge-based controller and the learning-based controller are compared. $J_c^\star$ is the optimal cost when the system dynamics are known. However, as the identified dynamics are with error, the cost J_c is worse than $J_c^\star$, and their difference can be represented by the system identification errors.

In general, the objectives of this paper are as follows:

- Characterizing the system identification performance of the proposed learning procedure with respect to the number of samples.
- Finding the relationship between the optimality of the controller and system identification error;
- Determining the controller optimality based on the initial and learning-based controllers and designing an optimal control strategy to incorporate the prior knowledge and the learning-based knowledge.

3. DESIGN AND ANALYSIS

This section presents the details of the overall design and the performance analysis. In the first part, the process of system identification is given. In the second and third parts, the analysis of the observer design and the discussion about the controller design are given.

3.1 System Identification

Since $\boldsymbol{A}$ and $\boldsymbol{B}$ are unknown, the dynamics of the plant need to be identified before the state estimation and the LQG control. We use the multiple roll-outs procedure and Gaussian inputs to generate data for identification, which is the same as in the previous studies (Dean et al., 2020; Lu and Mo, 2021). Our goal is to estimate $\boldsymbol{A}$ and $\boldsymbol{B}$ using a series of input and output data $\{(\boldsymbol{u}_k^{(l)}, \boldsymbol{y}_k^{(l)}) : 1 \leq l \leq N, 0 \leq k \leq T\}$ where the superscript l suggests the index of the roll-out. According to (1), it is easy to derive the recursive expression of observation data as follows

$$\boldsymbol{y}_{k+1}^{(l)} = \boldsymbol{A}\boldsymbol{y}_k^{(l)} + \boldsymbol{B}\boldsymbol{u}_k^{(l)} + \boldsymbol{w}_k^{(l)} + \boldsymbol{v}_{k+1}^{(l)} - \boldsymbol{A}\boldsymbol{v}_k^{(l)}.$$

For simplicity, let $\boldsymbol{W} = [\boldsymbol{A}, \boldsymbol{B}]^\top$ and $\boldsymbol{Y}_k^{(l)} = [\boldsymbol{y}_k^{(l)}; \boldsymbol{u}_k^{(l)}]^\top$. The Gaussian input $\boldsymbol{u} \sim (0, \sigma_u^2 \boldsymbol{I}_p)$ is used to excite the plant for identification purposes. We then obtain

$$\boldsymbol{y}_{k+1}^{(l)\top} = \boldsymbol{Y}_k^{(l)}\boldsymbol{W} + \boldsymbol{\Theta}_k^{(l)\top},$$

where $\boldsymbol{\Theta}_k^{(l)} = \boldsymbol{w}_k^{(l)} + \boldsymbol{v}_{k+1}^{(l)} - \boldsymbol{A}\boldsymbol{v}_k^{(l)}$ is a variable indicating the noise term. The LS method is adopted to estimate the system matrices. To avoid the interdependencies among the residual term and the input/output data in the LS method, we use the last sample of each trajectory as in (Dean et al., 2020).

Let $\boldsymbol{Y} = [\boldsymbol{Y}_{T-1}^{(1)}; \cdots; \boldsymbol{Y}_{T-1}^{(N)}]$, $\boldsymbol{\Theta} = [\boldsymbol{\Theta}_{T-1}^{(1)}, \cdots, \boldsymbol{\Theta}_{T-1}^{(N)}]^\top$ and $\boldsymbol{Z} = [\boldsymbol{y}_T^{(1)}, \cdots, \boldsymbol{y}_T^{(N)}]^\top$. The identification problem can be formulated as follows:

$$\min_{\hat{\boldsymbol{W}}} \|\boldsymbol{Y}\hat{\boldsymbol{W}} - \boldsymbol{Z}\|_F^2. \quad (4)$$

The main parameter we are going to optimize is $\hat{\boldsymbol{W}} = [\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}]^\top$. When $\boldsymbol{Y}$ has full column rank, the optimal solution of (4) is given by $\hat{\boldsymbol{W}}^\star = (\boldsymbol{Y}^\top\boldsymbol{Y})^{-1}\boldsymbol{Y}^\top\boldsymbol{Z}$. Due to the impact of the Gaussian noises, $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ differ from $\boldsymbol{A}$ and $\boldsymbol{B}$ with the error

$$\boldsymbol{E}_W = \hat{\boldsymbol{W}}^\star - \boldsymbol{W} = (\boldsymbol{Y}^\top\boldsymbol{Y})^{-1}\boldsymbol{Y}^\top\boldsymbol{\Theta}.$$

To obtain the identification error bounds, we first define the following two matrices

$$\boldsymbol{F}_T = \begin{bmatrix}\boldsymbol{A}^{T-1} & \boldsymbol{A}^{T-2} & \dots & \boldsymbol{I}_n\end{bmatrix},$$
$$\boldsymbol{G}_T = \begin{bmatrix}\boldsymbol{A}^{T-1}\boldsymbol{B} & \boldsymbol{A}^{T-2}\boldsymbol{B} & \dots & \boldsymbol{B}\end{bmatrix}.$$

Then, we give a theorem indicating the relationship between the identification error bounds and the number of samples N. Let

$$S = \lambda_{\min}\left(\sigma_u^2 \boldsymbol{G}_T\boldsymbol{G}_T^\top + \sigma_w^2 \boldsymbol{F}_T\boldsymbol{F}_T^\top + \sigma_v^2 \boldsymbol{I}_n\right),$$

and

$$\sigma_a = \sigma_w + \sigma_v\left(1 + \|\boldsymbol{A}\|\right).$$

Then, we have the following theorem.

Theorem 1. (The bounds of $\Delta_{\boldsymbol{A}}$ and $\Delta_{\boldsymbol{B}}$) Suppose that the sample size N satisfies the following inequality related to the state dimension n and the parameter $\delta \in (0, 1)$:

$$N \geq 8n + 16\log(2/\delta). \quad (5)$$

Algorithm 1 System Identification

Input: $N, T, \boldsymbol{x}_0, Q_w, R_v, \sigma_u$;
Output: $\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}$;

```
for all l = 1, ..., N do
    Initialize x_0 = 0;
    for all k = 0, 1, ...,T do
        Generate noises w_k^(l) ~ N(0, Q_w), v_k^(l) ~ N(0, R_v) and Gaussian input u_k^(l) ~ N(0, R_u);
        Update x_{k+1}^(l) and y_k^(l) by (1)
    end for
end for
Estimate Â and B̂ by the LS method (4);
```

Then, the bounds of $\Delta_{\boldsymbol{A}}$ and $\Delta_{\boldsymbol{B}}$ satisfy

$$\|\Delta_{\boldsymbol{A}}\| \leq \frac{16\sigma_a}{\sqrt{S}}\sqrt{\frac{(2n+p)\log(9/\delta)}{N}},$$

$$\|\Delta_{\boldsymbol{B}}\| \leq \frac{16\sigma_a}{\sigma_u}\sqrt{\frac{(2n+p)\log(9/\delta)}{N}},$$

with probability at least $1-\delta$.

Proof. The proof is shown in Appendix A.

Theorem 1 shows that the upper bounds of $\Delta_{\boldsymbol{A}}$ and $\Delta_{\boldsymbol{B}}$ depend on the sample size N. A larger N implies a more accurate identification result. The converge rate of the error is $\mathcal{O}(1/\sqrt{N})$, indicating that infinite data are needed to get the exact identification result. Note that similar theoretical results are given by (Dean et al., 2020), where the system model differs.

3.2 Observer Design and Analysis

The effectiveness of the state estimation is analyzed from an observer-designing point of view in this subsection. The Kalman filter can be viewed as an optimal observer, which can be expressed as:

$$\hat{\boldsymbol{x}}_k = \boldsymbol{A}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{B}\boldsymbol{u}_k + \boldsymbol{K}_k(\boldsymbol{y}_k - (\boldsymbol{A}\hat{\boldsymbol{x}}_{k-1} + \boldsymbol{B}\boldsymbol{u}_k)),$$

where $\boldsymbol{K}_k = \boldsymbol{P}_k^-(\boldsymbol{P}_k^- + \boldsymbol{R}_{\boldsymbol{v}})^{-1}$ is Kalman gain. The variable $\boldsymbol{P}_k^-$ satisfies the algebraic Riccati recursion:

$$\boldsymbol{P}_{k+1}^- = \boldsymbol{A}\boldsymbol{P}_k^-\boldsymbol{A}^\top + \boldsymbol{Q}_{\boldsymbol{w}} - \boldsymbol{A}\boldsymbol{P}_k^-(\boldsymbol{P}_k^- + \boldsymbol{R}_{\boldsymbol{v}})^{-1}\boldsymbol{P}_k^-\boldsymbol{A}^\top.$$

A steady-state filter is a time-invariant observer. Since the actual matrices are unavailable for the designer, $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ are used to solve for the observer, leading to

$$\begin{cases} \hat{\boldsymbol{x}}_{k+1} = (\boldsymbol{A}+\Delta_{\boldsymbol{A}})\hat{\boldsymbol{x}}_k + (\boldsymbol{B}+\Delta_{\boldsymbol{B}})\boldsymbol{u}_k + \boldsymbol{L}(\boldsymbol{x}_k + \boldsymbol{v}_k - \hat{\boldsymbol{x}}_k), \\ \hat{\boldsymbol{y}}_k = \hat{\boldsymbol{x}}_k, \end{cases}$$

where $\boldsymbol{L}$ is the observer gain. Denote the state estimation error by $\hat{\boldsymbol{e}}_k = \hat{\boldsymbol{x}}_k - \boldsymbol{x}_k$. The recursive form of the state estimation error is

$$\hat{\boldsymbol{e}}_{k+1} = (\boldsymbol{A}-\boldsymbol{L})\hat{\boldsymbol{e}}_k + \Delta_{\boldsymbol{A}}\hat{\boldsymbol{x}}_k + \Delta_{\boldsymbol{B}}\boldsymbol{u}_k + \boldsymbol{L}\boldsymbol{v}_k - \boldsymbol{w}_k.$$

Thus, the propagation of the estimation error is driven not only by the noise $\boldsymbol{w}_t - \boldsymbol{L}\boldsymbol{v}_t$, but also by $\Delta_{\boldsymbol{A}}\hat{\boldsymbol{x}}_k + \Delta_{\boldsymbol{B}}\boldsymbol{u}_k$.

In the observer design, the eigenvalues of $\boldsymbol{A}-\boldsymbol{L}$ need to lie within the unit circle. The estimation error propagates according to the linear system with closed-loop dynamics $\boldsymbol{A}-\boldsymbol{L}$, disturbed by the process noise $\boldsymbol{w}_t - \boldsymbol{L}\boldsymbol{v}_t$, which is i.i.d. zero means and has covariance $\boldsymbol{Q}_{\boldsymbol{w}} + \boldsymbol{L}\boldsymbol{R}_{\boldsymbol{v}}\boldsymbol{L}^\top$. If $(\boldsymbol{A}, \boldsymbol{Q}_{\boldsymbol{w}})$ is controllable, then $\boldsymbol{A}-\boldsymbol{L}$ is stable.

The mean square prediction error of the filter can be defined as the following with respect to the optimal Kalman filter:

$$J_e \triangleq \sqrt{\lim_{T\to\infty}\frac{1}{T}\sum_{k=0}^{T}\|\hat{\boldsymbol{x}}_k - \boldsymbol{x}_k\|^2}.$$

Theorem 2. (Error analysis of Kalman filter) Suppose the zero mean Gaussian input $\boldsymbol{u}$, whose covariance matrix is $\boldsymbol{R}_{\boldsymbol{u}}$, is used to excite the system. The mean square prediction error J_e is bounded by

$$J_e \leq \sqrt{2}\epsilon\|\boldsymbol{\Phi}_w\|_{\mathcal{H}_2}\left\{\left\|(\boldsymbol{\mathcal{R}}_{\boldsymbol{A}}\boldsymbol{B}+\boldsymbol{I})\boldsymbol{R}_{\boldsymbol{u}}^{\frac{1}{2}}\right\|_{\mathcal{H}_2} + \left\|\boldsymbol{\mathcal{R}}_{\boldsymbol{A}}\boldsymbol{K}\boldsymbol{R}_{\boldsymbol{v}}^{\frac{1}{2}}\right\|_{\mathcal{H}_2}\right\} + \left\|[\boldsymbol{\Phi}_w \;\; \boldsymbol{\Phi}_v]\begin{bmatrix}\boldsymbol{K}\\ \boldsymbol{I}\end{bmatrix}\boldsymbol{R}_{\boldsymbol{v}}^{\frac{1}{2}}\right\|_{\mathcal{H}_2},$$

where $\epsilon = \max\{\epsilon_{\boldsymbol{A}}, \epsilon_{\boldsymbol{B}}\}$, $\boldsymbol{K} = \sigma_w^2/\sigma_v^2\boldsymbol{I}_n$, $\boldsymbol{\Phi}_w = (z\boldsymbol{I} - \hat{\boldsymbol{A}} + \boldsymbol{L})^{-1}$, $\boldsymbol{\Phi}_v = -\boldsymbol{\Phi}_w\boldsymbol{L}$, and $\boldsymbol{\mathcal{R}}_{\boldsymbol{A}} = (z\boldsymbol{I} - \boldsymbol{A})^{-1}$.

Proof. The proof is shown in Appendix B.

Theorem 2 establishes a connection between the mean square prediction error and the upper bounds of the system identification errors $\Delta_{\boldsymbol{A}}$ and $\Delta_{\boldsymbol{B}}$. To get a good observer performance, the Gaussian input $\boldsymbol{u}$ should have a low amplitude. Furthermore, the error bound of the observer grows linearly with respect to the system identification errors, which implies that the sample size N also affects the upper bound of the prediction error with $J_e \sim \mathcal{O}(1/\sqrt{N})$.

3.3 Controller Design and Analysis

To control the actual system with optimal controller parameters, we first use Algorithm 1 to identify the system matrices and then design the new controller by (3). The core of the overall procedure is to determine when the performance under the controller $\boldsymbol{K}_{\boldsymbol{c}}$ exceeds that of the initial controller $\tilde{\boldsymbol{K}}_c$.

Theorem 3. Under Assumption 2, if the sample size N satisfies

$$N \geq \max\left\{\frac{1}{\varepsilon_{\boldsymbol{A}}^2 S}, \frac{1}{\varepsilon_{\boldsymbol{B}}^2\sigma_u^2}\right\} \times 256(2n+p)\log(9/\delta)\sigma_a^2, \quad (6)$$

then the LQG cost of $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ is bounded by

$$J_c(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}) \leq \max_{\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}} J_c(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}),$$

and the relative LQG error is

$$J_c(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}) - J_c^\star \sim \mathcal{O}\left(\frac{1}{\sqrt{N}}\right),$$

where $J_c^\star$ is the globally optimal LQG cost to (2).

Proof. The proof is shown in Appendix C.

Theorem 3 demonstrates that when (6) holds, the LQG controller based on the identified system matrices is no worse than the controller designed based on the prior knowledge of the system with high probability.

The problem with (6) is that the sample complexity N depends on the actual system matrices, and for the system whose true parameters are unknown, it is impossible to estimate the lower bound of N, i.e., this bound does not provide precise information on how many data samples

should be collected. In some situations, N needs to be extremely large in order to fulfill the control performance requirement. The identification procedure could be omitted if the system performance is bearable with coarse control. We next replace the terms in (6) that depend on the true system parameters with the approximations $\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}$, the upper bound of the approximation errors $\varepsilon_{\boldsymbol{A}}, \varepsilon_{\boldsymbol{B}}$, and the condition numbers $\kappa(\boldsymbol{A})$ and $\kappa(\boldsymbol{B})$ of the actual matrices.

Theorem 4. To bound the LQG cost of $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ by

$$J_c(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}) \leq \max_{\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}} J_c(\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}),$$

the estimated bound of the sample size N is

$$N^{'} \geq \max\left\{\frac{1}{\varepsilon_{\boldsymbol{A}}^2 S^{'}}, \frac{1}{\varepsilon_{\boldsymbol{B}}^2 \sigma_u^2}\right\} \times 256(2n+p)\log(9/\delta)\sigma_b^2. \quad (7)$$

In (7), we have

$$\sigma_b = \sigma_w + \sigma_v\left(1 + \|\tilde{\boldsymbol{A}}\| + \varepsilon_{\boldsymbol{A}}\right),$$

and $S^{'}$ is given by

$$S^{'} = \begin{cases} \sum_{t=0}^{T-1}\left[\frac{\|\tilde{\boldsymbol{A}}\|-\varepsilon_{\boldsymbol{A}}}{\kappa(\boldsymbol{A})}\right]^t \sigma_w^2 + \sigma_v^2, & \text{if } \operatorname{rank}(\boldsymbol{B}) < n, \\ \sum_{t=0}^{T-1}\left[\frac{\|\tilde{\boldsymbol{A}}\|-\varepsilon_{\boldsymbol{A}}}{\kappa(\boldsymbol{A})}\right]^t \left[\sigma_u^2 \frac{\|\tilde{\boldsymbol{B}}\|-\varepsilon_{\boldsymbol{B}}}{\kappa(\boldsymbol{B})} + \sigma_w^2\right] + \sigma_v^2, & \text{otherwise.} \end{cases}$$

Proof. The proof is shown in Appendix D.

Although the bound given by (7) is inevitably looser than the original bound in (6), it uses less information about the actual matrices and could almost be computed using only prior information.

4. SIMULATION

In this section, comparative simulations are conducted to show the effectiveness of the whole learning procedure introduced in Fig. 1 and reveal the relationships among the sample size N, identification errors, and performance of the LQG controller for unknown systems.

4.1 Simulation Setting

In the simulations, system matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ and their approximate matrices $\tilde{\boldsymbol{A}}$ and $\tilde{\boldsymbol{B}}$ are set as

$$\boldsymbol{A} = \begin{bmatrix} 1.1 & 0 \\ 0.5 & 0.1 \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 0.5 \\ 0.8 \end{bmatrix}; \tilde{\boldsymbol{A}} = \begin{bmatrix} 1.2 & 0.2 \\ 0.2 & 0.2 \end{bmatrix}, \tilde{\boldsymbol{B}} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}.$$

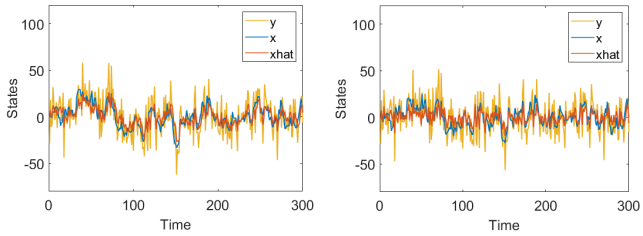
The initial state is set as $\boldsymbol{x}_0 = [0; 0]$ and the upper bounds on the prior error are $\varepsilon_{\boldsymbol{A}} = 0.6$ and $\varepsilon_{\boldsymbol{B}} = 0.6$. To design a new optimal controller, we follow the procedure described in the block diagram in Fig. 1.

4.2 Results and Analysis

In the first simulation, the variances of the process noise, the observation noise, and the Gaussian input are set as $\sigma_w = 1$, $\sigma_v = 1$, and $\sigma_u = 2$, respectively. The sample sizes are chosen in different scales to investigate the relationship among the identification error, observer performance, and controller performance. Each experiment with a different sample size N is repeated 100 times.

Table 1. Performance under different sample size N

	$N = 0$	$N = 2.2 \times 10^3$	$N = 2.2 \times 10^4$	$N = 2.2 \times 10^5$
J_c	6.811	5.715	5.629	5.625
J_e	1.152	1.084	1.082	1.081
ϵ	N/A	0.262	0.209	0.202



(a) N satisfies (7) (b) $N = 0$

Fig. 2. The changes of the states with or without system identification

Under this simulation setting, the smallest sample sizes that satisfy (6) and (7) are 61169 and 219060, respectively. These two numbers give the theoretical bound that leads to better controller performance than that based on the approximations of the system with high probabilities. Although the smallest sample size satisfying (6) is much larger than that satisfying (7), they are of the same order of magnitude. Table 1 shows that sample size N is the fundamental variable in the system design. Smaller system identification errors result in better LQG cost performance. The simulation data trends align with theoretical derivation. Here, $N = 0$ implies that we skip the system identification process and use the approximate stable controller. For scenarios where the demand on control performance is not strict, using the approximate model may be more efficient.

In the second simulation, the variances of the process noise, the observation noise, and the Gaussian input are set as $\sigma_w = 10$, $\sigma_v = 100$, and $\sigma_u = 2$, respectively.

This simulation demonstrates the performance of the controller designed with the framework in Fig. 1. Fig. 2 shows the overall control performance of both N satisfying (7) and $N = 0$. The horizontal axis represents time, and the vertical axis shows changes in the states of the first dimension. It can be seen that the system works well in situations when the noise is large and unavoidable.

5. CONCLUSION

In this work, we study whether learning dynamics for optimal control of an unknown linear system is worthwhile when prior knowledge of model uncertainties is available. We analyzed the relationship between system identification errors and the number of samples. The performance of the controller synthesis concerning optimality is discussed under different system identification error bounds. The strategy for adapting the learning procedure for control or adopting the approximate controller is given. Extensive simulations are conducted to illustrate the effectiveness of the learning procedure.

Appendix A. PROOF OF THEOREM 1

Proof. To analyze $\Delta_{\boldsymbol{A}}$ and $\Delta_{\boldsymbol{B}}$, we first present a few properties of the Gaussian noises.

Corollary 5. (Gaussian matrices, deviation; Corollary 5.35 in (Vershynin, 2010)). Let $\boldsymbol{M}$ be an $N\times n$ matrix whose entries are independent standard normal random variables, and $s_{\min}(\boldsymbol{M})$, $s_{\max}(\boldsymbol{M})$ be the smallest and the largest singular values of $\boldsymbol{M}$, respectively. Then for every $t\geq 0$, with probability at least $1-2\exp(-t^2/2)$, one has

$$\sqrt{N}-\sqrt{n}-t\leq s_{\min}(\boldsymbol{M})\leq s_{\max}(\boldsymbol{M})\leq\sqrt{N}+\sqrt{n}+t.$$

Besides this corollary, Lemma 2.1 in (Dean et al., 2020) shows that if $\boldsymbol{M}_1$ is $m_1\times N$ with i.i.d. $\mathcal{N}(0,1)$ entries and $\boldsymbol{M}_2$ is $N\times m_2$ with i.i.d. $\mathcal{N}(0,1)$ entries, both $\boldsymbol{M}_1$ and $\boldsymbol{M}_2$ are independent, then with probability at least $1-\delta$, we have

$$\|\boldsymbol{M}_1\boldsymbol{M}_2\|\leq 4\sqrt{N(m_1+m_2)\log(9/\delta)}. \tag{A.1}$$

Suppose that the matrix $\boldsymbol{M}$ has i.i.d. $\mathcal{N}(0,1)$ entries and $\boldsymbol{\Sigma}$ represents the covariance matrix of $\boldsymbol{Y}$, then $\boldsymbol{Y}$ can be represented by the product of $\boldsymbol{M}$ and $\boldsymbol{\Sigma}^{\frac{1}{2}}$, i.e., $\boldsymbol{Y}=\boldsymbol{M}\boldsymbol{\Sigma}^{\frac{1}{2}}$. The last observation $\boldsymbol{y}_T$ by each trajectory could be written in the following form

$$\boldsymbol{y}_T=\boldsymbol{G}_T\begin{bmatrix}\boldsymbol{u}_0\\ \boldsymbol{u}_1\\ \vdots\\ \boldsymbol{u}_{T-1}\end{bmatrix}+\boldsymbol{F}_T\begin{bmatrix}\boldsymbol{w}_0\\ \boldsymbol{w}_1\\ \vdots\\ \boldsymbol{w}_{T-1}\end{bmatrix}+\boldsymbol{v}_T.$$

Since the input is Gaussian $u\sim\mathcal{N}(0,\sigma_u^2\boldsymbol{I}_p)$, we have that

$$\begin{bmatrix}\boldsymbol{y}_T\\ \boldsymbol{u}_T\end{bmatrix}\sim\mathcal{N}\left(0,\begin{bmatrix}\sigma_u^2\boldsymbol{G}_T\boldsymbol{G}_T^\top+\sigma_w^2\boldsymbol{F}_T\boldsymbol{F}_T^\top+\sigma_v^2\boldsymbol{I}_n & 0\\ 0 & \sigma_u^2\boldsymbol{I}_p\end{bmatrix}\right).$$

Similar to Lemma 2.3 in (Dean et al., 2020), we calculate the upper bound of $\boldsymbol{V}\boldsymbol{E}_W$, where $\boldsymbol{V}$ can be any fixed matrix,

$$\begin{aligned}\boldsymbol{V}\boldsymbol{E}_W&=\boldsymbol{V}\left(\boldsymbol{Y}^\top\boldsymbol{Y}\right)^{-1}\boldsymbol{Y}^\top\boldsymbol{\Theta}\\&=\boldsymbol{V}\left(\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{M}^\top\boldsymbol{M}\boldsymbol{\Sigma}^{\frac{1}{2}}\right)^{-1}\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{M}^\top\boldsymbol{\Theta}\\&=\boldsymbol{V}\boldsymbol{\Sigma}^{-\frac{1}{2}}\left(\boldsymbol{M}^\top\boldsymbol{M}\right)^{-1}\boldsymbol{M}^\top\boldsymbol{\Theta}.\end{aligned}$$

The 2-norm of $\boldsymbol{V}\boldsymbol{E}_W$ satisfies

$$\|\boldsymbol{V}\boldsymbol{E}_W\|\leq\left\|\boldsymbol{V}\boldsymbol{\Sigma}^{-\frac{1}{2}}\right\|\frac{\|\boldsymbol{M}^\top\boldsymbol{\Theta}\|}{\lambda_{\min}(\boldsymbol{M}^\top\boldsymbol{M})}. \tag{A.2}$$

According to Corollary 5 and (5), the smallest eigenvalue of $\boldsymbol{M}^\top\boldsymbol{M}$ has the property that $\lambda_{\min}(\boldsymbol{M}^\top\boldsymbol{M})=[s_{\min}(\boldsymbol{M})]^2\geq N/4$ with probability at least $1-\delta$. According to Lemma 2.3 in (Dean et al., 2020), we have

$$\begin{aligned}\|\boldsymbol{M}^\top\boldsymbol{\Theta}\|&\leq 4\sigma_w\sqrt{(2n+p)\log(9/\delta)}\\&\quad+4\sigma_v(1+\|\boldsymbol{A}\|)\sqrt{(2n+p)\log(9/\delta)}.\end{aligned} \tag{A.3}$$

Together with (A.2), we have

$$\begin{aligned}\|\boldsymbol{V}\boldsymbol{E}_W\|&\leq 16\left(\sigma_w+\sigma_v(1+\|\boldsymbol{A}\|)\right)\\&\quad\left\|\boldsymbol{V}\boldsymbol{\Sigma}^{-\frac{1}{2}}\right\|\sqrt{\frac{(2n+p)\log(9/\delta)}{N}}\end{aligned} \tag{A.4}$$

with probability at least $1-\delta$. Applying (A.2) with matrix $\boldsymbol{V}_{\boldsymbol{A}}=[\boldsymbol{I}_n\ \boldsymbol{0}]$, $\boldsymbol{V}_{\boldsymbol{A}}\boldsymbol{E}_W$ extracts the estimate for $\boldsymbol{A}$. With probability at least $1-\delta$, we conclude

$$\|\Delta_{\boldsymbol{A}}\|\leq\frac{16\sigma_a}{\sqrt{S}}\sqrt{\frac{(2n+p)\log(9/\delta)}{N}},$$

and with $\boldsymbol{V}_{\boldsymbol{B}}=[\boldsymbol{0}\ \boldsymbol{I}_p]$, we conclude

$$\|\Delta_{\boldsymbol{B}}\|\leq\frac{16\sigma_a}{\sigma_u}\sqrt{\frac{(2n+p)\log(9/\delta)}{N}}.$$

The proof is completed.

Appendix B. PROOF OF THEOREM 2

Proof. (Tsiamis et al., 2020) manages to obtain the estimation guarantees for robust Kalman filtering and shows that imposing additional robustness constraints on the filter leads to sub-optimality guarantees. Similarly, in this paper, we have

$$\begin{aligned}(z\boldsymbol{I}-\boldsymbol{A}+\boldsymbol{L})\boldsymbol{x}&=\boldsymbol{B}\boldsymbol{u}+\boldsymbol{K}\boldsymbol{y}+(\boldsymbol{L}-\boldsymbol{K})\boldsymbol{x},\\(z\boldsymbol{I}-\hat{\boldsymbol{A}}+\boldsymbol{L})\hat{\boldsymbol{x}}&=\hat{\boldsymbol{B}}\boldsymbol{u}+\boldsymbol{L}\boldsymbol{y}.\end{aligned}$$

Subtracting the second equation from the first one, we have

$$(z\boldsymbol{I}-\hat{\boldsymbol{A}}+\boldsymbol{L})(\boldsymbol{x}-\hat{\boldsymbol{x}})=-\Delta_{\boldsymbol{A}}\boldsymbol{x}-\Delta_{\boldsymbol{B}}\boldsymbol{u}+(\boldsymbol{K}-\boldsymbol{L})\boldsymbol{v},$$

or equivalently,

$$\boldsymbol{x}-\hat{\boldsymbol{x}}=-\boldsymbol{\Phi}_w\Delta_{\boldsymbol{A}}\boldsymbol{x}-\boldsymbol{\Phi}_w\Delta_{\boldsymbol{B}}\boldsymbol{u}+\left(\boldsymbol{\Phi}_w\boldsymbol{K}+\boldsymbol{\Phi}_v\right)\boldsymbol{v},$$

where $\boldsymbol{\Phi}_w=(z\boldsymbol{I}-\hat{\boldsymbol{A}}+\boldsymbol{L})^{-1}$ and $\boldsymbol{\Phi}_v=-\boldsymbol{\Phi}_w\boldsymbol{L}$. Then $\boldsymbol{x}-\hat{\boldsymbol{x}}$ equals to

$$\begin{aligned}\boldsymbol{x}-\hat{\boldsymbol{x}}=&-\boldsymbol{\Phi}_w(\Delta_{\boldsymbol{A}}\boldsymbol{\mathcal{R}}_{\boldsymbol{A}}\boldsymbol{B}+\Delta_{\boldsymbol{B}})\boldsymbol{u}\\&+(\boldsymbol{\Phi}_w\boldsymbol{K}+\boldsymbol{\Phi}_v-\boldsymbol{\Phi}_w\Delta_{\boldsymbol{A}}\boldsymbol{\mathcal{R}}_{\boldsymbol{A}}\boldsymbol{K})\boldsymbol{v}.\end{aligned}$$

The mean square prediction error is

$$\begin{aligned}J_e=&\Big\|\boldsymbol{\Phi}_w(\Delta_{\boldsymbol{A}}\boldsymbol{\mathcal{R}}_{\boldsymbol{A}}\boldsymbol{B}+\Delta_{\boldsymbol{B}})\boldsymbol{R}_{\boldsymbol{u}}^{\frac{1}{2}}\\&+\left(\boldsymbol{\Phi}_w\boldsymbol{K}+\boldsymbol{\Phi}_v-\boldsymbol{\Phi}_w\Delta_{\boldsymbol{A}}\boldsymbol{\mathcal{R}}_{\boldsymbol{A}}\boldsymbol{K}\right)\boldsymbol{R}_{\boldsymbol{v}}^{\frac{1}{2}}\Big\|_{\mathcal{H}_2}.\end{aligned}$$

The upper bound of J_e is given by

$$\begin{aligned}J_e\leq&\sqrt{2}\epsilon\left\|\boldsymbol{\Phi}_w\right\|_{\mathcal{H}_2}\left\{\left\|(\boldsymbol{\mathcal{R}}_{\boldsymbol{A}}\boldsymbol{B}+\boldsymbol{I})\boldsymbol{R}_{\boldsymbol{u}}^{\frac{1}{2}}\right\|_{\mathcal{H}_2}+\left\|\boldsymbol{\mathcal{R}}_{\boldsymbol{A}}\boldsymbol{K}\boldsymbol{R}_{\boldsymbol{v}}^{\frac{1}{2}}\right\|_{\mathcal{H}_2}\right\}\\&+\left\|[\boldsymbol{\Phi}_w\ \boldsymbol{\Phi}_v]\begin{bmatrix}\boldsymbol{K}\\ \boldsymbol{I}\end{bmatrix}\boldsymbol{R}_{\boldsymbol{v}}^{\frac{1}{2}}\right\|_{\mathcal{H}_2}.\end{aligned}$$

Appendix C. PROOF OF THEOREM 3

Proof. The inequalities

$$\frac{16\sigma_a}{\sqrt{S}}\sqrt{\frac{(n+p)\log(9/\delta)}{N}}\leq\varepsilon_{\boldsymbol{A}},\ \frac{16\sigma_a}{\sigma_u}\sqrt{\frac{(n+p)\log(9/\delta)}{N}}\leq\varepsilon_{\boldsymbol{B}},$$

are equivalent to (6). If the above inequalities hold, then it is guaranteed that

$$\|\hat{\boldsymbol{A}}-\boldsymbol{A}\|<\max\|\tilde{\boldsymbol{A}}-\boldsymbol{A}\|,\quad\|\hat{\boldsymbol{B}}-\boldsymbol{B}\|<\max\|\tilde{\boldsymbol{B}}-\boldsymbol{B}\|,$$

leading to $J_e(\hat{\boldsymbol{A}},\hat{\boldsymbol{B}})<\max\limits_{\tilde{\boldsymbol{A}},\tilde{\boldsymbol{B}}}J_e(\tilde{\boldsymbol{A}},\tilde{\boldsymbol{B}})$. Moreover,

$$\begin{aligned}J_c(\hat{\boldsymbol{A}},\hat{\boldsymbol{B}})-J_c^\star&=\lim_{T\to\infty}\mathbb{E}\left[\frac{1}{T}\sum_{k=0}^{T}\boldsymbol{u}_k^\top(\boldsymbol{B}^\top\boldsymbol{Q}\boldsymbol{B}+\boldsymbol{R})\boldsymbol{u}_k\right]\\&\sim s_{\max}\left[(\boldsymbol{B}^\top\boldsymbol{Q}\boldsymbol{B}+\boldsymbol{R})\boldsymbol{K}_{\boldsymbol{c}}^\star\right]J_e,\end{aligned}$$

which means that $J_c(\hat{\boldsymbol{A}},\hat{\boldsymbol{B}})$ is of the same order as $J_e(\hat{\boldsymbol{A}},\hat{\boldsymbol{B}})$. We conclude that $J_c(\hat{\boldsymbol{A}},\hat{\boldsymbol{B}})<\max\limits_{\tilde{\boldsymbol{A}},\tilde{\boldsymbol{B}}}J_c(\tilde{\boldsymbol{A}},\tilde{\boldsymbol{B}})$. Since $J_e\sim\mathcal{O}\left(1/\sqrt{N}\right)$, the relative LQG cost satisfies

$$J_c(\hat{\boldsymbol{A}},\hat{\boldsymbol{B}})-J_c^\star\sim\mathcal{O}\left(\frac{1}{\sqrt{N}}\right). \tag{C.1}$$

Appendix D. PROOF OF THEOREM 4

Proof. The matrices $\boldsymbol{G}_T\boldsymbol{G}_T^\top = \sum_{t=1}^{T} \boldsymbol{A}^{T-t}\boldsymbol{B}\boldsymbol{B}^\top {\boldsymbol{A}^{T-t}}^\top$ and $\boldsymbol{F}_T\boldsymbol{F}_T^\top = \sum_{t=1}^{T} \boldsymbol{A}^{T-t}{\boldsymbol{A}^{T-t}}^\top$ are sums of a sequence of real symmetric matrices with non-negative main diagonals. Therefore,

$$\begin{aligned} S &= \lambda_{\min}\left(\sigma_u^2\boldsymbol{G}_T\boldsymbol{G}_T^\top + \sigma_w^2\boldsymbol{F}_T\boldsymbol{F}_T^\top + \sigma_v^2\boldsymbol{I}_n\right) \\ &\geq \sigma_u^2 \sum_{t=1}^{T} \sigma_{\min}\left(\boldsymbol{A}^{T-t}\boldsymbol{B}\right) + \sigma_w^2 \sum_{t=1}^{T} \sigma_{\min}\left(\boldsymbol{A}^{T-t}\right) + \sigma_v^2. \end{aligned}$$

For any matrices $\boldsymbol{M} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{N} \in \mathbb{R}^{n\times p}$, it holds that

$$\begin{aligned} \sigma_{\min}(\boldsymbol{M}\boldsymbol{N}) &= \min_{\boldsymbol{x}\neq 0} \frac{\|\boldsymbol{M}\boldsymbol{N}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \min_{\boldsymbol{x}\neq 0} \frac{\|\boldsymbol{M}\boldsymbol{N}\boldsymbol{x}\|}{\|\boldsymbol{N}\boldsymbol{x}\|}\frac{\|\boldsymbol{N}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \\ &\geq \min_{\boldsymbol{y}\neq 0} \frac{\|\boldsymbol{M}\boldsymbol{y}\|}{\|\boldsymbol{y}\|} \min_{\boldsymbol{x}\neq 0} \frac{\|\boldsymbol{N}\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \sigma_{\min}(\boldsymbol{M})\sigma_{\min}(\boldsymbol{N}). \end{aligned}$$

Note that both sides of the inequality are zero if $\boldsymbol{N}$ has a deficient column rank. Suppose that the condition numbers $\kappa(\boldsymbol{A}) = \frac{\sigma_{\max}(\boldsymbol{A})}{\sigma_{\min}(\boldsymbol{A})}$ and $\kappa(\boldsymbol{B}) = \frac{\sigma_{\max}(\boldsymbol{B})}{\sigma_{\min}(\boldsymbol{B})}$ are the only information available about the actual matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. It is easy to obtain the lower bound of S as

$$S \geq \begin{cases} \sum_{t=0}^{T-1} \left[\frac{\sigma_{\max}(\boldsymbol{A})}{\kappa(\boldsymbol{A})}\right]^t \sigma_w^2 + \sigma_v^2, & \text{if } \operatorname{rank}(\boldsymbol{B}) < n, \\ \sum_{t=0}^{T-1} \left[\frac{\sigma_{\max}(\boldsymbol{A})}{\kappa(\boldsymbol{A})}\right]^t \left[\sigma_u^2 \frac{\sigma_{\max}(\boldsymbol{B})}{\kappa(\boldsymbol{B})} + \sigma_w^2\right] + \sigma_v^2, & \text{otherwise.} \end{cases}$$

Assumption 2 implies that $\|\tilde{\boldsymbol{A}}\| - \varepsilon_{\boldsymbol{A}} \leq \|\boldsymbol{A}\| \leq \|\tilde{\boldsymbol{A}}\| + \varepsilon_{\boldsymbol{A}}$ and $\|\tilde{\boldsymbol{B}}\| - \varepsilon_{\boldsymbol{B}} \leq \|\boldsymbol{B}\|$. Substitute the prior knowledge $\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{B}}, \varepsilon_{\boldsymbol{A}}, \varepsilon_{\boldsymbol{B}}$ for unknown system matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, we obtain $S^{'}$. Together with

$$\sigma_a \leq \sigma_w + \sigma_v\left(1 + \|\tilde{\boldsymbol{A}}\| + \varepsilon_{\boldsymbol{A}}\right) = \sigma_b,$$

the proof is completed.

REFERENCES

Anderson, B.D. and Moore, J.B. (2012). *Optimal filtering*. Courier Corporation.

Berberich, J., Scherer, C.W., and Allgöwer, F. (2022). Combining prior knowledge and data for robust controller design. *IEEE Transactions on Automatic Control*.

Dean, S., Mania, H., Matni, N., Recht, B., and Tu, S. (2020). On the sample complexity of the linear quadratic regulator. *Foundations of Computational Mathematics*, 20(4), 633–679.

Ding, X., He, J., Chen, C., and Guan, X. (2021). System parameter and state estimator over unknown linear systems. In *American Control Conference*, 4007–4012.

Furieri, L., Zheng, Y., Papachristodoulou, A., and Kamgarpour, M. (2019). An input–output parametrization of stabilizing controllers: Amidst Youla and system level synthesis. *IEEE Control Systems Letters*, 3(4), 1014–1019.

Jha, S.K., Roy, S.B., and Bhasin, S. (2017). Data-driven adaptive LQR for completely unknown LTI systems. In *IFAC World Congress*, 4156–4161.

Li, Y., He, J., Chen, C., and Guan, X. (2021). On topology inference for networked dynamical systems: Principles and performances. *arXiv:2106.01031*.

Lu, J. and Niu, R. (2014). False information injection attack on dynamic state estimation in multi-sensor systems. In *International Conference on Information Fusion*, 1–8.

Lu, Y. and Mo, Y. (2021). Safe linear-quadratic dual control with almost sure performance guarantee. *arXiv:2103.13278*.

Oymak, S. and Ozay, N. (2019). Non-asymptotic identification of LTI systems from a single trajectory. In *American Control Conference*, 5655–5661.

Oymak, S. and Ozay, N. (2021). Revisiting Ho-Kalman based system identification: Robustness and finite-sample analysis. *IEEE Transactions on Automatic Control*, 67(4), 1914–1928.

Sabzekar, M. and Hasheminejad, S.M.H. (2021). Robust regression using support vector regressions. *Chaos, Solitons & Fractals*, 144, 110738.

Sinquin, B. (2019). *Structured matrices for predictive control of large and multi-dimensional systems*. Ph.D. thesis.

Tsiamis, A., Matni, N., and Pappas, G. (2020). Sample complexity of Kalman filtering for unknown systems. In *Learning for Dynamics and Control*, 435–444.

Vershynin, R. (2010). Introduction to the non-asymptotic analysis of random matrices. *arXiv:1011.3027*.

Wang, Y.S., Matni, N., and Doyle, J.C. (2019). A system-level approach to controller synthesis. *IEEE Transactions on Automatic Control*, 64(10), 4079–4093.

Xie, L. and Soh, Y.C. (1994). Robust Kalman filtering for uncertain systems. *Systems & Control Letters*, 22(2), 123–129.

Youla, D., Jabr, H., and Bongiorno, J. (1976). Modern Wiener-Hopf design of optimal controllers–Part II: The multivariable case. *IEEE Transactions on Automatic Control*, 21(3), 319–338.

Zheng, Y., Furieri, L., Kamgarpour, M., and Li, N. (2021). Sample complexity of linear quadratic gaussian (LQG) control for output feedback systems. In *Learning for Dynamics and Control*, 559–570.