Logarithmic Regret and Polynomial Scaling in Online Multi-step-ahead Prediction

Jiachen Qian, *Student Member, IEEE,* Yang Zheng, *Senior Member, IEEE*

***Abstract*— This letter studies the problem of online multi-step-ahead prediction for unknown linear stochastic systems. Using conditional distribution theory, we derive an optimal parameterization of the prediction policy as a linear function of future inputs, past inputs, and past outputs. Based on this characterization, we propose an online least-squares algorithm to learn the policy and analyze its regret relative to the optimal model-based predictor. We show that the online algorithm achieves logarithmic regret with respect to the optimal Kalman filter in the multi-step setting. Furthermore, with new proof techniques, we establish an almost-sure regret bound that does not rely on fixed failure probabilities for sufficiently large horizons N. Finally, our analysis also reveals that, while the regret remains logarithmic in N, its constant factor grows polynomially with the prediction horizon H, with the polynomial order set by the largest Jordan block of eigenvalue 1 in the system matrix.**

***Index Terms*— Model-free learning, Multi-step prediction, Logarithmic Regret**

I. Introduction

ONLINE prediction of dynamical system behavior has long been recognized as a fundamental problem in control systems [1], robotics [2], computer vision [3], etc. Classical approaches rely on an accurate model and known noise statistics to propagate system responses and examine the effect of future control inputs [4]. For instance, the celebrated Kalman filter provides *optimal* mean-square-error predictions under correct modeling assumptions [5]. In many applications [6]–[8], however, obtaining explicit models and reliable noise characterizations is impractical, especially when input-output relationships are complex to identify. This has motivated growing interest in learning prediction policies directly from data, without full knowledge of the underlying system.

From model-based prediction to data-driven approaches, a central challenge is how to parameterize the prediction policy. Traditional system identification addresses this by first identifying Markov parameters and then solving a nonconvex problem to extract a system model, upon which a standard Kalman predictor can be used [9]. Recent advances have provided non-asymptotic analysis of the identification process by characterizing convergence rates of the identification error [10], [11] and leveraging multiple trajectories to mitigate potential state divergence [12]. Nonetheless, mapping from Markov parameters to a system model is intrinsically nonlinear and nonconvex [13], which makes it difficult to establish strong theoretical guarantees for online prediction performance.

Instead of identifying an explicit model, some recent studies learn a prediction policy directly from input-output data [14]–[18]. These methods exploit the Kalman filter's structure and parameterize the prediction policy as a linear function of past inputs and outputs, and then estimate the weights via online learning. This approach, also known as *improper learning*, bypasses the intermediate system identification. Notably, [15] shows that by employing a truncated autoregressive (AR) model derived from the Kalman filter, a regret measure (i.e., the cumulative loss of the online predictor relative to the optimal model-based prediction) scales logarithmically with the time horizon. Thus, with sufficiently long trajectories, the average prediction error of the online predictor approaches that of the optimal Kalman filter. Building on this, [16] established a similar logarithmic regret bound using low-rank approximation techniques to address slow convergence under heavy noise. Our recent work [17] has introduced an *exponential forgetting strategy* to address the unbalanced regression model in online prediction while preserving the logarithmic regret.

The aforementioned results [14]–[18] focus primarily on *single-step* prediction. In practical scenarios such as path planning and predictive control [19]–[21], *multi-step* predictions are essential to enforce state constraints and optimize control policies. While empirical studies [22], [23] suggest that recursively applying a single-step predictor can extend the prediction horizon, such autoregressive roll-outs can compound errors and often underperform direct multi-step prediction [24]. Yet, the theory of effective *multi-step* prediction, especially a quantitative characterization of performance degradation with horizon length, remains largely underdeveloped.

In this paper, we focus on model-free online learning of a *multi-step-ahead* predictor for linear stochastic systems. Our contributions are as follows. First, based on conditional distribution theory, we introduce an autoregressive model that parameterizes the H-step-ahead prediction policy as a linear combination of past outputs, past inputs, and future outputs (Theorem 1). Unlike single-step prediction, the innovation in H-step-ahead prediction becomes temporally correlated and *non-orthogonal*. We further establish that the innovation term coincides with the autoregressive roll-out in [15], thereby providing a theoretical justification for the heuristic approach. Second, based on the autoregressive model, we propose an online least-squares-based learning algorithm for H-step-ahead prediction. With a backward horizon chosen proportional to $\log N$, where N denotes the total time horizon,

This work is supported by NSF CMMI 2320697 and NSF CAREER 2340713. Jiachen Qian and Yang Zheng are with the Department of Electrical and Computer Engineering, University of California San Diego, CA 92093 USA (e-mail: jiq012@ucsd.edu; zhengy@ucsd.edu)

we establish that the regret with respect to the optimal model-based Kalman predictor scales logarithmically in N almost surely (Theorem 2). The regret remains logarithmic despite the non-orthogonal innovation process. The prediction horizon H does not change the order in N; instead, it appears only in the constant, which grows polynomially at a rate no larger than $H^{4\kappa+1}$, where κ is the size of the largest Jordan block of eigenvalue 1 in the system matrix. Compared with the prior literature [14]–[18], our results give the first explicit regret scaling in terms of prediction horizon and hold almost surely rather than merely in probability.

Notation: We use $A \succ B$ to denote that $A - B$ is positive definite. We use $\|\cdot\|_2$, $\|\cdot\|_F$, and $\|\cdot\|_1$ to denote the 2-norm, the Frobenius norm, and the 1-norm, respectively. $\mathcal{N}(\mu, V)$ denotes a Gaussian distribution with mean μ and variance V. $\rho(A)$ denotes the spectral radius of A. $\mathrm{poly}(x)$ denotes a polynomial in x. $\mathcal{O}(f(x))$ indicates the function is of the same order as $f(x)$, $o(f(x))$ indicates the function is of a smaller order than $f(x)$.

II. Preliminaries and Problem Statement

A. Linear stochastic system and the Kalman filter

Consider the following linear stochastic system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + w_k, \\ y_k &= Cx_k + v_k, \quad k = 0, 1, 2, \dots \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^{n_u}$ is the input vector, $y_k \in \mathbb{R}^m$ is the output vector, $\omega_k \sim \mathcal{N}(0, Q)$ and $v_k \sim \mathcal{N}(0, R)$ are the process and observation noises, respectively, and we assume $u_k \sim \mathcal{N}(0, I_{n_u})$.

In this paper, we make a standard assumption.

Assumption 1: The matrix A is marginally stable, i.e., $\rho(A) \leq 1$, and the system pair (A, C) is detectable.

If the system parameters (A, B, C, Q, R) are known, we can apply the Kalman filter [1], [5] to predict the future outputs. Let $\mathcal{F}_k \triangleq \sigma(y_0, \dots, y_k)$ be the filtration generated by the observations $y_0, \dots, y_k$. Given $\mathcal{F}_k$, we aim to predict the optimal $\hat{y}_{k+1}$ in the minimum mean-square error sense:

$$\hat{y}_{k+1} \triangleq \arg\min_{z \in \mathcal{F}_k} \mathbb{E}\left[\|y_{k+1} - z\|_2^2 \mid \mathcal{F}_k\right]. \tag{2}$$

It is now well-known that the steady-state optimal predictor takes a recursive form, known as the *Kalman filter* [5],

$$\begin{aligned} \hat{x}_{k+1} &= A\hat{x}_k + Bu_k + L(y_k - \hat{y}_k), \quad \hat{x}_0 = 0 \\ \hat{y}_k &= C\hat{x}_k, \end{aligned} \tag{3}$$

where $L = APC^\mathsf{T}\left(CPC^\mathsf{T} + R\right)^{-1}$ is called the steady-state Kalman gain with P from the algebraic Riccati equation:

$$P = APA^\mathsf{T} + Q - APC^\mathsf{T}\left(CPC^\mathsf{T} + R\right)^{-1} CPA^\mathsf{T}. \tag{4}$$

It is also shown that the optimal prediction $\hat{y}_{k+1}$ is equivalent to the expectation of y_{k+1} conditioned on $\mathcal{F}_k$, i.e., $\hat{y}_{k+1} = \mathbb{E}\{y_{k+1} \mid \mathcal{F}_k\}$ [1, Section 2]. We note that the general Kalman filter takes a time-varying form. However, due to the exponential convergence of the conditioned output process y_k to become steady-state [1], the difference between the time-varying and steady-state filters remains bounded by a constant. Similar to [15]–[17], we focus directly on the steady-state prediction in this paper.

B. Optimal one-step prediction policy

We briefly review here how to utilize the Kalman filter to build an autoregressive model for *single step* prediction [15]. Denote $e_k = y_k - \hat{y}_k$ as the innovation at time step k. By rolling out the Kalman filter (3) backwards for p times, we can reformulate y_{k+1} in terms of the past inputs and outputs as

$$y_{k+1} = \tilde{G}_p Z_{k,p} + C(A - LC)^p \hat{x}_{k-p+1} + e_{k+1}, \tag{5}$$

where $Z_{k,p} \triangleq \left[y_{k-p+1}^\mathsf{T} \cdots y_k^\mathsf{T}, u_{k-p+1}^\mathsf{T} \cdots u_k^\mathsf{T}\right]^\mathsf{T}$ collects the past outputs and inputs, and $\tilde{G}_p$ denotes the optimal weights consisting of $\tilde{G}_p = \left[\tilde{G}_{1,p}, \tilde{G}_{2,p}\right]$, where:

$$\tilde{G}_{1,p} \triangleq \left[C(A - LC)^{p-1}L, \quad \cdots, \quad CL\right] \in \mathbb{R}^{m \times pm} \tag{6a}$$

$$\tilde{G}_{2,p} \triangleq \left[C(A - LC)^{p-1}B, \quad \cdots, \quad CB\right] \in \mathbb{R}^{m \times pn_u}. \tag{6b}$$

This shows that the *optimal* steady-state policy for predicting one-step-ahead output is *a linear function of past outputs and inputs*. We here state another technical assumption [15]–[17]:

Assumption 2: The matrix $A - LC$ is diagonalizable.

This assumption is used only to simplify the regret analysis. It ensures an exponential decay bound $\rho\left((A - LC)^p\right) \leq M\rho(A - LC)^p$ with some constant $M > 0$. There is no fundamental difficulty without Assumption 2. If $A - LC$ contains higher-order Jordan blocks, the convergence speed of $(A - LC)^p$ will be $\rho\left((A - LC)^p\right) = \mathrm{poly}(p)\rho(A - LC)^p$. Hence, the backward horizon p should be further extended to account for the slower convergence rate.

C. Problem statement

With observations up to time k, this paper aims to develop an online *multi-step-ahead* prediction policy of the form

$$\tilde{y}_{k+H} = f_H(y_0, \dots, y_k, u_0, \dots, u_k, \dots, u_{k+H-1}) \tag{7}$$

where the term $\tilde{y}_{k+H}$ means the *H-step ahead prediction* at time step k. If $H = 1$, this is reduced to the linear policy for *one-step* prediction (5), as established in [15]. Following [15]–[17], we quantify the performance of the online prediction in terms of the *regret* measured against the Kalman filter (3) that has full system knowledge. Under this setting, the original benchmark Kalman predictor from (3), i.e.,

$$\hat{y}_{k+H} = \arg\min_{z \in \mathcal{F}_{k+H-1}} \mathbb{E}\left[\|y_{k+H} - z\|_2^2 \mid \mathcal{F}_{k+H-1}\right],$$

will be too strong, since it uses the information up to y_{k+H-1}.

To address this, we consider a modified benchmark, called the *H-step ahead Kalman predictor*, defined as

$$\bar{y}_{k+H} \triangleq \arg\min_{z \in \mathcal{F}_k} \mathbb{E}\left[\|y_{k+H} - z\|_2^2 \mid \mathcal{F}_k\right]. \tag{8}$$

We aim to minimize the following regret

$$\mathcal{R}_N \triangleq \sum_{k=1}^{N} \|y_{k+H} - \tilde{y}_{k+H}\|^2 - \sum_{k=1}^{N} \|y_{k+H} - \bar{y}_{k+H}\|^2, \tag{9}$$

where $\tilde{y}_{k+H}$ is our online model-free prediction (7) and $\bar{y}_{k+H}$ is the optimal model-based Kalman prediction (8).

We address two fundamental questions in designing (7): 1) how to parameterize the optimal multi-step-ahead prediction policy in terms of past outputs and control inputs; and 2) how to effectively learn this policy online and quantify its regret (9) relative to the optimal model-based predictor.

III. Main results

A. Multi-step-ahead regression model

We here derive a closed-form expression of the model-based optimal H-step ahead predictor and then provide a parameterization of the optimal prediction policy.

Lemma 1: Consider the linear stochastic system (1) and the optimal H-step ahead prediction problem (8). The optimal H-step ahead prediction $\bar{y}_{k+H}$ can be obtained recursively as

$$\begin{aligned} \bar{x}_{k+H} &= A^{H-1}\hat{x}_{k+1} + \sum_{i=1}^{H-1} A^{i-1}Bu_{k+i}, \\ \bar{y}_{k+H} &= C\bar{x}_{k+H}, \end{aligned} \tag{10}$$

where $\hat{x}_{k+1} \triangleq \mathbb{E}\{x_{k+1} \mid \mathcal{F}_k\}$ is the standard Kalman state estimation from (3).

The proof is not difficult, and we present some details in Appendix A. Lemma 1 shows that the optimal H-step-ahead predictor can be obtained by first computing the best state estimate $\hat{x}_{k+1}$ via the standard Kalman filter and then propagating this estimate forward $H-1$ steps using the system dynamics together with the planned inputs $\{u_{k+1}, \ldots, u_{k+H-1}\}$. The predicted output $\bar{y}_{k+H}$ is simply the observation of this rolled-out state. Hence, optimal multi-step prediction reduces to applying the Kalman estimator once at the current time, followed by deterministic roll-out under the known dynamics.

Lemma 1 allows us to obtain an optimal prediction policy as a linear function of past outputs, past inputs, and future inputs.

Theorem 1: Let $r_{k+H} = y_{k+H} - \bar{y}_{k+H}$ denote the innovation process for the H-step ahead predictor (10). The following linear regression model holds

$$y_{k+H} = G_p Z_{k,p} + CA^{H-1}(A-LC)^p \hat{x}_{k-p+1} + r_{k+H}, \tag{11}$$

where $Z_{k,p} \triangleq \left[y_{k-p+1}^\mathsf{T} \cdots\ y_k^\mathsf{T}, u_{k-p+1}^\mathsf{T} \cdots u_{k+H-1}^\mathsf{T}\right]^\mathsf{T}$ contains the past outputs, past inputs, and future inputs, and the regressor weights $G_p \triangleq [G_{1,p}, G_{2,p}]$ are of the form:

$$\begin{aligned} G_{1,p} &\triangleq \left[CA^{H-1}(A-LC)^{p-1}L,\ \cdots,\ CA^{H-1}L\right], \\ G_{2,p} &\triangleq \left[CA^{H-1}(A-LC)^{p-1}B,\ \cdots,\ CA^{H-1}B, \cdots, CB\right]. \end{aligned}$$

Furthermore, we have

$$r_{k+H} = e_{k+H} + \textstyle\sum_{i=1}^{H-1} CA^{i-1}Le_{k+H-i}, \tag{12}$$

where $e_k = y_k - \hat{y}_k$ is the innovation in the Kalman filter (3).

We present the proof in Section IV-A. By definition, the H-step innovation r_{k+H} aggregates past process noise $w_{k+1}, \ldots, w_{k+H-1}$. The overlap among these noise components induces correlations between r_{k+H} and r_{k+H-l} for $l < H$. Thus, the usual orthogonality of the one-step innovation process no longer holds in the H-step-ahead setting. As established in (12), this *non-orthogonal* innovation r_{k+H} can be parameterized as a combination of $e_{k+1}, \ldots, e_{k+H}$.

In [15], an H-step autoregressive (AR) model (without control inputs) is obtained by rolling out the one-step AR model H times, and its innovation ϵ_{k+H} satisfies $\epsilon_{k+H} = e_{k+H} + \sum_{i=1}^{H-1} CA^{i-1}Le_{k+H-i}$, which coincides with (12). This observation highlights that including control inputs does not introduce additional uncertainty into the innovation process. Furthermore, (12) provides an equivalent parameterization of r_{k+H} as a summation of temporally uncorrelated one-step innovations e_k across time. These properties are essential in our online learning algorithm and its regret analysis.

B. Online learning and regret guarantee

From the linear regression (11), the H-step-ahead output y_{k+H} is a linear function of the past outputs, past inputs, and future inputs, perturbed by a bias term depending on $\hat{x}_{k-p+1}$ and an innovation process r_{k+H}. Following [14]–[17], we can estimate G_p via least squares by regressing y_{t+H} onto the regressor $Z_{t,p}$, $t \le k$ (past inputs/outputs and future inputs).

Least-squares. We estimate $G_{k,p}$ by ridge regression:

$$\tilde{G}_{k,p} = \arg\min_G \sum_{t=p}^{k-H} \|y_{t+H} - GZ_{t,p}\|_F^2 + \lambda \|G\|_F^2, \tag{13}$$

where $\lambda > 0$ is a regularization parameter. At each time step k, by solving (13), we can obtain a closed form of $\tilde{G}_{k,p}$ as

$$\tilde{G}_{k,p} = \textstyle\sum_{t=p}^{k-H} y_{t+H} Z_{t,p}^\mathsf{T} V_{k-H,p}^{-1}, \tag{14}$$

where $V_{k-H,p} \triangleq \lambda I + \sum_{t=p}^{k-H} Z_{t,p} Z_{t,p}^\mathsf{T}$ is called the Gram matrix, which contains all collected past available samples $Z_{t,p}$. We then predict the future observation by

$$\tilde{y}_{k+H} = \tilde{G}_{k,p} Z_{k,p}. \tag{15}$$

Controlling bias. For nonexplosive systems $\rho(A) \le 1$, the bias term $\hat{x}_{k-p+1}$ in (11) retains the state from previous time steps, potentially growing at a polynomial rate. A persistent bias error $b_{k+H,p} \triangleq CA^{H-1}(A-LC)^p \hat{x}_{k-p+1}$ at each step k could result in linear regret. Fortunately, Lemma 1 and classical theory guarantee that $\left\|A^{H-1}(A-LC)^p\right\|_2 \le c_1 H^{\kappa-1}\rho(A-CL)^p$, with $\rho(A-LC) < 1$. This property allows us to control the accumulation of bias errors $A^{H-1}(A-LC)^p\hat{x}_{k-p+1}$ only with $p = \mathcal{O}(\log H \log k)$. We implement this via the standard doubling trick [25] (as in [15]): partition time into epochs of doubling length, and keep p fixed within each epoch while increasing it between epochs. We note that the prediction policy in [24] chooses $p = 1$, which will induce a significant bias and degrade the prediction performance.

Recursive updates within an epoch. At each time step k, we update the prediction $\tilde{y}_{k+H}$ using (15) and then observe the new observation y_{k+1}. Within each epoch, the predictor can be computed recursively

$$V_{k-H,p} = V_{k-H-1,p} + Z_{k-H,p} Z_{k-H,p}^\mathsf{T}, \tag{16a}$$

$$\tilde{G}_{k,p} = \tilde{G}_{k-1,p} + (y_k - \tilde{y}_k)\, Z_{k-H,p}^\mathsf{T} V_{k-H,p}^{-1}. \tag{16b}$$

A key distinction from the traditional one-step prediction [14]–[18] is the *update delay*: the prediction $\tilde{y}_{k+H}$ is not used immediately to update G at time k; due to the H-step horizon, it contributes to the correction H steps later.

The model-free multi-step-ahead prediction is summarized in Algorithm 1. We have the following regret guarantee.

Theorem 2: Consider the linear stochastic system (1). Suppose Assumptions 1 and 2 hold. For a fixed H-step-ahead

Algorithm 1 H-step-ahead Online Prediction (HOP)

Input: parameter $\beta, \lambda, T_{\text{init}}, N_E$
{*Warm Up*:}
for $k = 1$ **to** $T_{\text{init}} + 1$ **do**
  Observe y_k, Generate u_{k+H-1};
end for
{*Recursive Online Prediction:*}
for $l = 1$ **to** N_E **do**
  Initialize $T_l = 2^{l-1} T_{\text{init}} + 1, p = \beta \log T_l$,
  Compute $V_{T_l - H, p}$ and $\tilde{G}_{T_l, p}$;
  for $k = T_l$ **to** $2T_l - 2$ **do**
    Predict $\tilde{y}_{k+H} = \tilde{G}_{k,p} Z_{k,p}$;
    Observe y_{k+1}, Generate u_{k+H};
    Update $V_{k-H+1,p}$ and $\tilde{G}_{k+1,p}$ as (16).
  end for
end for

prediction, we choose the parameters in Algorithm 1 as

$$\beta = \frac{\mathcal{O}(\kappa + \log H)}{\log(1/\rho(A - LC))}, \tag{17}$$

where κ represents the order of the largest Jordan block of eigenvalue 1 in matrix A. Almost for sure, we have

$$\mathcal{R}_N \leq M H^{4\kappa+1} \beta^3 \mathcal{O}\left(\log^7 N\right), \tag{18}$$

where M is a constant related to the system parameters.

We outline the proof in Section IV and highlight the key technical differences compared with the literature [14]–[17].

Theorem 2 provides the *first* logarithmic regret for online multi-step-ahead prediction with respect to the optimal model-based predictor. When A has Jordan blocks at eigenvalue 1, the regret remains logarithmic in N, but its constant scales polynomially with the prediction horizon H, with degree determined by the size of the largest such block. This logarithmic regret holds true despite the innovation r_{k+H} being non-orthogonal. Compared with [14]–[18], our result gives the first explicit polynomial scaling of the regret in the prediction horizon H.

Moreover, our bound holds almost surely, i.e., it does not depend on a fixed failure probability $\delta \in (0, 1)$. The key idea for this property is that for fixed δ, the original regret in the literature [15]–[17] will scale with $\operatorname{poly}\log \frac{1}{\delta}$. For sufficiently large N, we only need to let $\delta = \frac{1}{N}$, then the above bound can be obtained. This property coincides with the empirical result that the norm of a long Gaussian random process is "very close" to the norm of its expectation [26]. To the best of our knowledge, this is the first almost sure bound for online prediction, although a similar technique has also been utilized in online LQR [27].

IV. Technical Proofs

A. Proof of Theorem 1

First, since $\hat{x}_{k+1} \triangleq \mathbb{E}\{x_{k+1} \mid \mathcal{F}_k\}$ is the standard Kalman's state estimation from (3), we can roll the standard Kalman filter backwards for p times and get $\hat{x}_{k+1} = (A-LC)^p \hat{x}_{k+1-p} + \sum_{l=0}^{p-1} (A-LC)^l (L y_{k-l} + B u_{k-l})$. Then the regression model (11) is a direct consequence of the recursive update (10) combined with the above recursive relation.

In the following, we establish the relationship (12), and only consider the steady-state innovation process. Denote $\tilde{x}_{k+1} \triangleq x_{k+1} - \hat{x}_{k+1}$ as the one-step optimal state prediction error for Kalman filter (3), then we have

$$\mathbb{E}\left\{\tilde{x}_{k+1} \tilde{x}_{k+1}^{\mathsf{T}}\right\} = P, \quad \forall k \in \mathbb{N},$$

where P is from (4), and the following recursion holds

$$\tilde{x}_k = (A - LC)\tilde{x}_{k-1} + w_{k-1} - L v_{k-1}.$$

By comparing y_{k+H} and $\bar{y}_{k+H}$ directly, we obtain

$$r_{k+H} = C A^{H-1} \tilde{x}_{k+1} + v_{k+H} + \sum_{i=1}^{H-1} C A^{i-1} w_{k+H-i}. \tag{19}$$

Consider the structure of $e_k = C\tilde{x}_k + v_k$. We have

$$e_{k+H} = C(A-LC)\tilde{x}_{k+H-1} + C w_{k+H-1} - CL v_{k+H-1} + v_{k+H}.$$

Together with

$$CL e_{k+H-1} = CLC\tilde{x}_{k+H-1} + CL v_{k+H-1},$$

we further have

$$e_{k+H} + CL e_{k+H-1} = CA\tilde{x}_{k+H-1} + v_{k+H} + C w_{k+H-1}.$$

With a similar procedure, we have a recursive deduction below

$$\begin{aligned}
&e_{k+H} + \textstyle\sum_{i=1}^{H-1} CA^{i-1} L e_{k+H-i} \\
=&CA\tilde{x}_{k+H-1} + v_{k+H} + C w_{k+H-1} + \textstyle\sum_{i=2}^{H-1} CA^{i-1} L e_{k+H-i}. \\
=&CA^{H-1}\tilde{x}_{k+1} + v_{k+H} + \textstyle\sum_{i=1}^{H-1} CA^{i-1} w_{k+H-i}
\end{aligned}$$

Combining this with (19) leads to the desired relationship (12).

B. Proof of Theorem 2

We here outline the proof of Theorem 2 in four main steps. More details for each step are postponed to the appendix.

1) Regret decompose: With [15, Theorem 1], the regret $\mathcal{R}_N$ is dominated by $\mathcal{L}_N \triangleq \sum_{k=T_{\text{init}}}^{N} \|\tilde{y}_{k+H} - \bar{y}_{k+H}\|_2^2$., i.e., $\mathcal{R}_N = \mathcal{L}_N + o(\mathcal{L}_N)$. Hence, in the following proof, we mainly analyze the scaling law of $\mathcal{L}_N$ with respect to N and H. We further decompose each $\tilde{y}_{k+H} - \bar{y}_{k+H}$ into three parts,

$$\begin{aligned}
\tilde{y}_{k+H} - \hat{y}_{k+H} = &\underbrace{\sum_{l=p}^{k-H} b_{l+H,p} Z_{l,p}^{\mathsf{T}} V_{k-H,p}^{-1} Z_{k,p} - b_{k+H,p}}_{\text{Bias error}} \\
&+ \underbrace{\sum_{l=p}^{k-H} r_{l+H} Z_{l,p}^{\mathsf{T}} V_{k-H,p}^{-1} Z_{k,p}}_{\text{Regression error}} - \underbrace{\lambda G_p V_{k-H,p}^{-1} Z_{k,p}}_{\text{Regularization error}}.
\end{aligned}$$

We further denote $\mathcal{B}_{k,p} \triangleq \left\| \sum_{l=p}^{k-H} b_{l+H,p} Z_{l,p}^{\mathsf{T}} V_{k-H,p}^{-\frac{1}{2}} \right\|_2^2$ as the bias factor, $\mathcal{E}_{k,p} \triangleq \left\| \sum_{l=p}^{k-H} r_{l+H,p} Z_{l,p}^{\mathsf{T}} V_{k-H,p}^{-\frac{1}{2}} \right\|_2^2$ as the regression factor, $\mathcal{G}_{k,p} \triangleq \left\| \lambda G_p V_{k-H,p}^{-\frac{1}{2}} \right\|_2^2$ as the regularization factor, $\mathcal{V}_{N,p} \triangleq \sum_{k=T_{\text{init}}}^{N} \left\| V_{k-H,p}^{-\frac{1}{2}} Z_{k,p} \right\|_2^2$ as the accumulation

factor, and $\mathfrak{b}_{N+H,p} = \sum_{k=T_{\text{init}}}^{N} \|b_{k+H,p}\|_2^2$ as the accumulation of bias. Then we can further decompose the term $\mathcal{L}_N$ into

$$\mathcal{L}_N \leq 4\big(\max_{k\leq N}(\mathcal{B}_{k,p}+\mathcal{E}_{k,p}+\mathcal{G}_{k,p})\big)\cdot\mathcal{V}_{N,p}+4\mathfrak{b}_{N+H,p}. \quad (20)$$

Note that from the expression of G_p, we can obtain that $\mathcal{G}_{k,p} \leq \frac{MH^{2\kappa}}{1-\rho(A-LC)}$ does not scale with N. Thus, in the following proof, we discuss bias factor $\mathcal{B}_{k,p}$, regression factor $\mathcal{E}_{k,p}$, and accumulation factor $\mathcal{V}_{N,p}$.

2) Suppress $\mathfrak{b}_{k+H,p}$ with large enough β : We state a lemma for bias factor $\mathcal{B}_{k,p}$ and accumulation of bias $\mathfrak{b}_{N+H,p}$.

Lemma 2: Suppose Assumptions 1 and 2 hold. For any fixed β satisfying the condition in Theorem 2, we have

$$\mathfrak{b}_{N+H,p} \leq M\log^2 N, \quad \text{and} \quad \max_{T_{\text{init}}\leq k\leq N}\mathcal{B}_{k,p} \leq M\log N \quad (21)$$

holds almost for sure.

To prove this lemma, we first fix a failure probability δ, and we can guarantee that the system state $\hat{x}_k$ scales polynomially with k with high probability, i.e.,

$$\mathbb{P}\left\{\|\hat{x}_k\|_2^2 \leq Mk^{2\kappa}\log\frac{1}{\delta},\ \forall k\in\mathbb{N}\right\} \geq 1-\delta. \quad (22)$$

Then the bias factor satisfies $\mathcal{B}_{k,p} \leq \sum_{l=p}^{k-H}\|b_{l+H,p}\|_2^2 \leq \left\|CA^{H-1}(A-LC)^p\right\|_2^2 \sum_{l=p}^{k-H}\|\hat{x}_{k-p+1}\|_2^2$. With an exponential decay of $(A-LC)^p$, and further letting $\delta = \frac{1}{N}$, we obtain

$$\mathbb{P}\left\{\max_{k\leq N}\mathcal{B}_{k,p} \leq M\log N\right\} \geq 1-\frac{1}{N},$$

holds for any fixed N. More details are given in Appendix C.

3) Decouple martingale process: We have the following lemma for regression factor $\mathcal{E}_{k,p}$

Lemma 3: Suppose Assumption 1 holds. For any fixed β satisfying the condition in Theorem 2, we have

$$\max_{T_{\text{init}}\leq k\leq N}\mathcal{E}_{k,p} \leq MH^{2\kappa}\beta\log^2 N \quad (23)$$

holds almost for sure.

From (12), we can divide the innovation into the summation of H unrelated innovations e_k, then the regression factor can be further relaxed into

$$\mathcal{E}_{k,p} \leq H\sum_{i=1}^{H-1}\left\|CA^{i-1}L\right\|_2^2\left\|\sum_{l=p}^{k-H}e_{l+H-i}Z_{l,p}^{\mathsf{T}}V_{k-H,p}^{-\frac{1}{2}}\right\|_2^2 + H\left\|\sum_{l=p}^{k-H}e_{l+H}Z_{l,p}^{\mathsf{T}}V_{k-H,p}^{-\frac{1}{2}}\right\|_2^2.$$

For each self-normalized martingale terms, we can apply [15, Theorem 3] to show that for each $0 \leq i \leq H-1$

$$\left\|\sum_{l=p}^{k-H}e_{l+H-i}Z_{l,p}^{\mathsf{T}}V_{k-H,p}^{-\frac{1}{2}}\right\|_2^2 \leq \log\frac{H}{\delta}+\log\frac{\det V_{k-H,p}}{\det\lambda I}$$

holds uniformly for all k with high probability $1-\delta/H$. Similarly, with the almost sure bound for $\det V_{k-H,p}$ as $\max_{k\leq N}\log\det V_{k-H,p} \leq M\beta\log^2 N$, this lemma is proved by choosing $\delta = \frac{1}{N}$. See Appendix D for details.

4) Concentration of accumulation: We have the following lemma for accumulation factor $\mathcal{V}_{N,p}$

Lemma 4: Suppose Assumption 1 holds. For any fixed β satisfying the condition in Theorem 2, we have

$$\mathcal{V}_{N,p} \leq MH^{2\kappa+1}\beta^2\log^5 N \quad (24)$$

holds almost for sure.

Without loss of generality, we directly assume $N = 2^{N_E}T_{\text{init}}$, where N_E is the number of epochs. Then we can decompose the $\mathcal{V}_{N,p}$ into

$$\mathcal{V}_{N,p} \leq \max_{T_{\text{init}}\leq k\leq N}\left\|V_{k-H,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2 \times \sum_{l=1}^{N_E}\sum_{k=T_l}^{2T_l-2}\left\|V_{k,p_l}^{-\frac{1}{2}}Z_{k,p_l}\right\|_2^2,$$

where the subscript of p_l is to highlight the p varies with epoch number l. For the second part, we can directly bound it by the concentration inequality

$$\sum_{k=T_l}^{2T_l-2}\left\|V_{k,p_l}^{-\frac{1}{2}}Z_{k,p_l}\right\|_2^2 \leq \log\left(\det V_{2T_l-2,p_l}/\det V_{T_l-1,p_l}\right).$$

While for the bound $\max_{T_{\text{init}}\leq k\leq N}\left\|V_{k-H,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2$, note that $\left\|V_{k-H,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2 \leq 1+\sum_{l=k-H+1}^{k}Z_{l,p}V_{k-H,p}^{-1}Z_{l,p}^{\mathsf{T}}$, and hence we need the following H-step AR representation of $Z_{k,p}$, which is derived as

$$Z_{k,p} = \sum_{i=1}^{d}a_{i-1}^{(H)}Z_{k-H+i-d,p}+\delta_{k,p}+\sum_{j=1}^{H-1}a_{d-1}^{(j)}\delta_{k-j,p},$$

where each $a_{i-1}^{(j)}$ is obtained from the coefficient of minimal polynomial of A and d is the dimension of minimal polynomial of A. We can further show that $a_{i-1}^{(j)} \leq j^{\kappa-1}$, and hence finally provide a bound $\max_{T_{\text{init}}\leq k\leq N}\left\|V_{k-H,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2 \leq MH^{2\kappa+1}\beta\log^2 N$ holds almost for sure.

Proof of Theorem 2: We can now directly combine (20) and Lemmas 2 to 4 to establish the desired regret bound in (18), where we have eliminated the low-order terms.

V. Numerical Experiments

We here provide numerical experiments to verify the performance of the proposed HOP in Algorithm 1. We consider a modified dynamical system model from [15, Section V] with control inputs. The system parameters are given by

$$A = \begin{bmatrix}1 & 0.5 & 0\\ 0 & 1 & 0.5\\ 0 & 0 & 0.9\end{bmatrix},\quad B = \begin{bmatrix}0\\0\\1\end{bmatrix},\quad C = \begin{bmatrix}1 & 0 & 0\end{bmatrix},$$

and $Q = 0.01 * I_3, R = 0.01$. The control input u_k is randomly generated from i.i.d. standard Gaussian distribution, i.e., $u_k \sim \mathcal{N}(0,1)$. The hyperparameter is chosen to be $\beta = 2$, and $T_{\text{init}} = 400$, the number of epochs is 3.

In Figure 1, we provide the comparison of regret $\mathcal{R}_N$ with different ahead prediction step H, where H are chosen to be $2, 4, 5, 6$ respectively. We can see that the regret remains logarithmic despite different H. For a marginally stable system, the regret scales nonlinearly with the increase of H, which is consistent with the polynomial scaling claim in Theorem 2.

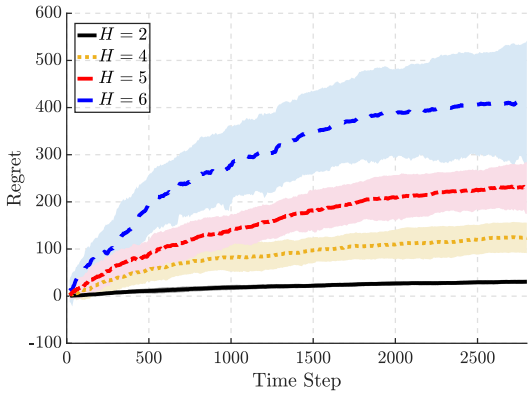

Fig. 1. Comparison of regret $\boldsymbol{\mathcal{R}_N}$ across prediction horizons $\boldsymbol{H} \in \boldsymbol{\{2, 4, 5, 6\}}$. The regret remains logarithmic for all $\boldsymbol{H}$. The multiplicative constant increases nonlinearly with $\boldsymbol{H}$.

TABLE I

COMPARISON OF $\boldsymbol{R_N}$ SCALING WITH $\boldsymbol{H}$ FOR MARGINALLY STABLE SYSTEM AND OPEN-LOOP STABLE SYSTEMS.

Ahead step H	2	4	6	8	10	12
$\mathcal{R}_N$ $(\rho(A) = 1)$	30.7	123.7	410.9	1035	2280	4600
$\mathcal{R}_N$ $(\rho(A) = 0.6)$	2.84	3.49	3.60	4.48	5.08	4.78

To further verify the scaling property of $\mathcal{R}_N$ with respect to H, we also consider an open-loop stable system. The system matrix A is chosen to be $A = \begin{bmatrix} 0.6 & 0.5 & 0 \\ 0 & 0.6 & 0.5 \\ 0 & 0 & 0.6 \end{bmatrix}$, and the parameters B, C, Q, R remain the same as those in the previous experiment. The regrets for the marginally stable system and the open-loop stable system with different H are listed in Table I. We can see that for the open-loop stable system, the regret $\mathcal{R}_N$ scales roughly linearly with H, and even experiences a saturation effect for large enough H. While for marginally stable systems, $\mathcal{R}_N$ scales polynomially with H, with the polynomial order between H^2 and H^3. These results are consistent with our analysis in Theorem 2, though the polynomial exponent there may be conservative.

VI. CONCLUSION

In this letter, we address the problem of multi-step-ahead prediction for unknown linear stochastic systems. We have derived an optimal parameterization of the H-step predictor as a linear combination of future inputs, past inputs, and past outputs. We have proposed an online least–squares-based algorithm to learn this policy. Our algorithm achieves an almost-sure logarithmic regret bound with respect to the optimal model-based H-step predictor. The dependence on the time horizon N is logarithmic, while the multiplicative constant scales polynomially with the prediction horizon H. Future directions include designing online feedback laws to stabilize a linear stochastic system with regret guarantees relative to an optimal LQR/LQG controller.

REFERENCES

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, N.J: Prentice-Hall, 1979.

[2] T. D. Barfoot, *State Estimation for Robotics*. Cambridge University Press, 2024.

[3] H. Coskun, F. Achilles, R. DiPietro, N. Navab, and F. Tombari, "Long short-term memory Kalman filters: Recurrent neural estimators for pose regularization," in *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 5524–5532.

[4] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[5] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 03 1960.

[6] S. E. Li, Y. Zheng, K. Li, Y. Wu, J. K. Hedrick, F. Gao, and H. Zhang, "Dynamical modeling and distributed control of connected and automated vehicles: Challenges and opportunities," *IEEE Intelligent Transportation Systems Magazine*, vol. 9, no. 3, pp. 46–58, 2017.

[7] Z. Wu, A. Tran, D. Rincon, and P. D. Christofides, "Machine learning-based predictive control of nonlinear processes. Part I: theory," *AIChE Journal*, vol. 65, no. 11, p. e16729, 2019.

[8] I. Markovsky, L. Huang, and F. Dörfler, "Data-driven control based on the behavioral approach: From theory to applications in power systems," *IEEE Control Systems Magazine*, vol. 43, no. 5, pp. 28–68, 2023.

[9] L. Ljung, "System identification," in *Signal Analysis and Prediction*. Springer, 1998, pp. 163–173.

[10] I. Ziemann, A. Tsiamis, B. Lee, Y. Jedra, N. Matni, and G. J. Pappas, "A tutorial on the non-asymptotic theory of system identification," in *2023 62nd IEEE Conference on Decision and Control (CDC)*. IEEE, 2023, pp. 8921–8939.

[11] J. He, I. Ziemann, C. R. Rojas, S. J. Qin, and H. Hjalmarsson, "Finite sample analysis of subspace identification methods," *arXiv preprint arXiv:2501.16639*, 2025.

[12] Y. Zheng and N. Li, "Non-asymptotic identification of linear dynamical systems using multiple trajectories," *IEEE Control Systems Letters*, vol. 5, no. 5, pp. 1693–1698, 2020.

[13] M. Hardt, T. Ma, and B. Recht, "Gradient descent learns linear dynamical systems," *Journal of Machine Learning Research*, vol. 19, no. 29, pp. 1–44, 2018.

[14] E. Hazan, K. Singh, and C. Zhang, "Learning linear dynamical systems via spectral filtering," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[15] A. Tsiamis and G. J. Pappas, "Online learning of the Kalman filter with logarithmic regret," *IEEE Transactions on Automatic Control*, vol. 68, no. 5, pp. 2774–2789, 2023.

[16] P. Rashidinejad, J. Jiao, and S. Russell, "SLIP: Learning to predict in unknown dynamical systems with long-term memory," *Advances in Neural Information Processing Systems*, vol. 33, pp. 5716–5728, 2020.

[17] J. Qian and Y. Zheng, "Model-free online learning for the Kalman filter: Forgetting factor and logarithmic regret," *arXiv preprint arXiv:2505.08982*, 2025.

[18] X. Zhang, B. Hu, and T. Bacsar, "Learning the Kalman filter with fine-grained sample complexity," *2023 American Control Conference (ACC)*, pp. 4549–4554, 2023.

[19] A. Tsiamis, A. Karapetyan, Y. Li, E. C. Balta, and J. Lygeros, "Predictive linear online tracking for unknown targets," *arXiv preprint arXiv:2402.10036*, 2024.

[20] J. Wang, Y. Zheng, K. Li, and Q. Xu, "DeeP-LCC: Data-enabled predictive leading cruise control in mixed traffic flow," *IEEE Transactions on Control Systems Technology*, vol. 31, no. 6, pp. 2760–2776, 2023.

[21] C. Liu, S. Lee, S. Varnhagen, and H. E. Tseng, "Path planning for autonomous vehicles using model predictive control," in *2017 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2017, pp. 174–179.

[22] M. Farina and L. Piroddi, "Some convergence properties of multi-step prediction error identification criteria," in *2008 47th IEEE Conference on Decision and Control*. IEEE, 2008, pp. 756–761.

[23] N. Lambert, A. Wilcox, H. Zhang, K. S. Pister, and R. Calandra, "Learning accurate long-term dynamics for model-based reinforcement learning," in *2021 60th IEEE Conference on Decision and Control (CDC)*. IEEE, 2021, pp. 2880–2887.

[24] A. Somalwar, B. D. Lee, G. J. Pappas, and N. Matni, "Learning with imperfect models: When multi-step prediction mitigates compounding error," *arXiv preprint arXiv:2504.01766*, 2025.

[25] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games*. Cambridge university press, 2006.

[26] R. Vershynin, *High-Dimensional Probability: An Introduction with Applications in Data Science*. Cambridge university press, 2018, vol. 47.

[27] Y. Lu and Y. Mo, "Almost surely $\sqrt{T}$ regret for adaptive LQR," *IEEE Transactions on Automatic Control*, vol. 70, no. 8, pp. 5145–5159, 2025.

[28] Y. Abbasi-Yadkori, D. Pál, and C. Szepesvári, "Improved algorithms for

linear stochastic bandits," *Advances in neural Information Processing Systems*, vol. 24, 2011.
[29] T. L. Lai and C. Z. Wei, "Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems," *The Annals of Statistics*, pp. 154–166, 1982.
[30] B. Laurent and P. Massart, "Adaptive estimation of a quadratic functional by model selection," *Annals of Statistics*, pp. 1302–1338, 2000.

APPENDIX

A. Proof of Lemma 1

It is known that the optimal prediction (8) is equivalent to the conditional expectation [1, Section 2], i.e.,

$$\bar{y}_{k+H} = \mathbb{E}\left\{y_{k+H} \mid \mathcal{F}_k\right\}.$$

We thus only need to compute this condition expectation. Let $f(x)$ be the probability distribution function of a random variable x. The Chapman-Kolmogorov (C-K) equation says that

$$f\left(x_{k+2}|\mathcal{F}_k\right) = \int f\left(x_{k+2}|x_{k+1}\right) f\left(x_{k+1}|\mathcal{F}_k\right) \mathrm{d}x_{k+1}. \quad (25)$$

From the system dynamics (1), we have

$$f\left(x_{k+1}|x_k\right) = \mathcal{N}\left(Ax_k + Bu_k, Q\right).$$

Classical Kalman filtering theory guarantees that the steady-state estimation [1] satisfies

$$f\left(x_{k+1} \mid \mathcal{F}_k\right) = \mathcal{N}\left(\hat{x}_{k+1}, P\right),$$

where P is the unique positive semidefinite solution to the ARE (4). Calculating the integral (25), we have

$$f\left(x_{k+2} \mid \mathcal{F}_k\right) = \mathcal{N}\left(A\hat{x}_{k+1} + Bu_{k+1}, APA^{\mathsf{T}} + Q\right).$$

Performing the integral (25) for $H-1$ times recursively leads to

$$f\left(x_{k+H} \mid \mathcal{F}_k\right) = \mathcal{N}\left(\bar{x}_{k+H}, P_H\right),$$

where the mean $\bar{x}_{k+H}$ is defined in (10) and the covariance $P^{(H)}$ takes the form as

$$P_H = A^{H-1}P\left(A^{H-1}\right)^{\mathsf{T}} + \textstyle\sum_{i=1}^{H-1} A^{i-1}Q(A^{i-1})^{\mathsf{T}}. \quad (26)$$

Since we have $\mathbb{E}\left\{y_{k+H} \mid \mathcal{F}_k\right\} = C\mathbb{E}\left\{x_{k+H} \mid \mathcal{F}_k\right\}$, the proof is now completed.

B. Details for Regret decomposition

With classical results in online linear regression techniques [15], [16], we first divide the regret $\mathcal{R}_N$ into two parts, i.e.,

$$\begin{aligned}\mathcal{R}_N &\triangleq \sum_{k=T_{\text{init}}}^{N} \left\|y_{k+H} - \tilde{y}_{k+H}\right\|_2^2 - \sum_{k=T_{\text{init}}}^{N} \left\|y_{k+H} - \bar{y}_{k+H}\right\|_2^2 \\ &= \underbrace{\sum_{k=T_{\text{init}}}^{N} \left\|\bar{y}_{k+H} - \tilde{y}_{k+H}\right\|_2^2}_{\mathcal{L}_N} + 2\underbrace{\sum_{k=T_{\text{init}}}^{N} r_{k+H}^{\mathsf{T}}(\bar{y}_{k+H} - \tilde{y}_{k+H})}_{\text{cross term}}.\end{aligned}$$

The first part is the accumulation of the gap $\bar{y}_k - \tilde{y}_k$, while the second part is a cross term between the innovation r_{k+H} and the gap $\bar{y}_k - \tilde{y}_k$. Since r_{k+H} can be decoupled into the summation of H i.i.d. Gaussian sequences, from the self-normalized martingale theory [28], it is standard to bound

$$\sum_{k=T_{\text{init}}}^{N} r_{k+H}^{\mathsf{T}}\left(\bar{y}_{k+H} - \tilde{y}_{k+H}\right) = \tilde{O}\left(\sqrt{\mathcal{L}_N}\right) = o\left(\mathcal{L}_N\right), \quad (27)$$

i.e., the cross term is dominated by the accumulation term $\mathcal{L}_N$.

Then following standard linear regression techniques [14], [28], we can divide the gap $\bar{y}_{k+H} - \tilde{y}_{k+H}$ at each time step k as:

- 1) the *regularization* error induced by λI,
- 2) the *regression* error induced by Gaussian innovation r_{k+H},
- 3) the *bias* error induced by $b_{k+H,p} = CA^{H-1}(A - KC)^p\hat{x}_{k-p+1}$.

In particular, we rewrite the gap $\tilde{y}_{k+H} - \bar{y}_{k+H}$ as

$$\begin{aligned}\tilde{y}_{k+H} - \hat{y}_{k+H} = &\underbrace{\sum_{l=p}^{k-H} b_{l+H,p} Z_{l,p}^{\mathsf{T}} V_{k-H,p}^{-1} Z_{k,p} - b_{k+H,p}}_{\text{Bias error}} \\ &+ \underbrace{\sum_{l=p}^{k-H} r_{l+H} Z_{l,p}^{\mathsf{T}} V_{k-H,p}^{-1} Z_{k,p}}_{\text{Regression error}} - \underbrace{\lambda G_p V_{k-H,p}^{-1} Z_{k,p}}_{\text{Regularization error}}. \quad (28)\end{aligned}$$

The main difference between (28) and the decomposition in [15], [16] is that the innovation r_{k+H} is inherently correlated, i.e., $\mathbb{E}\left\{r_{k+H} r_{k+H-l}^{\mathsf{T}}\right\} \neq 0, \forall l < H$, which requires some decoupling techniques. Due to the time delay, the bias term will be affected by A^{H-1}, which requires a longer past horizon to suppress. Furthermore, the delay induced asymmetry, i.e., the cross between $V_{k-H,p}^{-\frac{1}{2}} Z_{k,p}$ will further complicate the analysis process. Following a simple argument, we have the following bound.

$$\begin{aligned}\mathcal{L}_N &= \sum_{k=T_{\text{init}}}^{N} \left\|\tilde{y}_{k+H} - \bar{y}_{k+H}\right\|_2^2 \\ &\leq 4\big(\max_{k\leq N}\left(\mathcal{B}_{k,p} + \mathcal{E}_{k,p} + \mathcal{G}_{k,p}\right)\big) \cdot \mathcal{V}_{N,p} + 4\mathfrak{b}_{N,p}, \quad (29)\end{aligned}$$

where the factors are defined as

$$\mathcal{B}_{k,p} \triangleq \left\|B_{k,p}\bar{Z}_{k-H,p}V_{k-H,p}^{-\frac{1}{2}}\right\|_2^2, \mathcal{E}_{k,p} \triangleq \left\|\mathcal{R}_{k,p}\bar{Z}_{k-H,p}^{\mathsf{T}}V_{k-H,p}^{-\frac{1}{2}}\right\|_2^2,$$

$$\mathcal{G}_{k,p} \triangleq \left\|\lambda G_p D_p^{-2} V_{k-H,p}^{-\frac{1}{2}}\right\|_2^2, \quad \mathcal{V}_{N,p} \triangleq \sum_{k=T_{\text{init}}}^{N} \left\|V_{k-H,p}^{-\frac{1}{2}} Z_{k,p}\right\|_2^2,$$

$\mathfrak{b}_{N+H,p} = \sum_{k=T_{\text{init}}}^{N} \left\|b_{k+H,p}\right\|_2^2$. In the terms above, $B_{k,p} \triangleq \left[b_{p+H,p}, \ldots, b_{k,p}\right]$, $\mathcal{R}_{k,p} \triangleq \left[r_{p+H}, \ldots, r_k\right]$, $\bar{Z}_{k-H,p} \triangleq \left[Z_{p,p}, \ldots, Z_{k-H,p}\right]$ are the collections of all past bias $b_{l,p}$, innovations r_l, and samples $Z_{l,p}$, respectively [1]. Without loss of generality, we directly assume $N = 2^{N_E}T_{\text{init}}$, where N_E is the number of epochs.

[1]For a specific k in different epochs, the parameter p will be different due to the doubling trick. In Algorithm 1, we always have $p \leq \beta \log k$.

C. Proof for Lemma 2 and selection of β

We first provide an almost sure bound for the bias factors and derive the requirement for β. Note that at each epoch l, for any $T_l \le k \le 2T_l - 2$ there is

$$\begin{aligned}\mathcal{B}_{k,p} &= \left\| B_{k,p} \bar{Z}_{k-H,p}^{\mathsf{T}} \bar{V}_{k-H,p}^{-1} \bar{Z}_{k-H,p} B_{k,p}^{\mathsf{T}} \right\|_2^2 \le \sum_{i=p+H}^{k} \|b_{i,p}\|_2^2 \\ &\le \sum_{i=p+H}^{k} \left\| CA^{H-1}(A-LC)^p \right\|_2^2 \|\hat{x}_{i-p-H+1}\|_2^2,\end{aligned}$$

where $p = \beta \log(T_l - 1)$. Denote $\Gamma_k \triangleq \mathbb{E}\{\hat{x}_k \hat{x}_k^{\mathsf{T}}\}$, with Lemma 5, we have that for a fixed probability δ and for all $k \in \mathbb{N}$, there is

$$\begin{aligned}\|\hat{x}_k\|_2^2 &\le \left(2n + 3\log\frac{k^2}{\delta}\right) \|\Gamma_k\|_2^2 \\ &\le \left(2n + 3\log\frac{k^2}{\delta}\right) \left(\|Q\|_2 + \|B\|_2^2\right) \sum_{i=0}^{k-1} \left\|A^i\right\|_2^2\end{aligned}$$

uniformly holds for all k with probability at least $1 - \frac{\pi^2\delta}{6}$. Note that $\left\|A^i\right\|_2^2 \le M i^{2\kappa-2}$, where M is a constant only related to system parameters, together with $\left\|CA^{H-1}(A-LC)^p\right\|_2^2 \le MH^{2\kappa-2}\rho(A-LC)^{2p}$, we first have that the inequality

$$\mathcal{B}_{k,p} \le MH^{2\kappa-2}\rho(A-LC)^{2p}k^{2\kappa}\log\frac{1}{\delta},$$

holds for all k uniformly with probability $1 - \frac{\pi^2\delta}{6}$. The above inequality holds due to $\log k \le k$. We then choose $\beta = \frac{M_1}{\log(1/\rho(A-LC))}$, where M_1 is a parameter to be determined. Then we have

$$\begin{aligned}\mathcal{B}_{k,p} \le& MH^{2\kappa-2}k^{2\kappa}\rho(A-LC)^{\frac{2M_1\log\frac{k}{2}}{\log(1/\rho(A-LC))}}\log\frac{1}{\delta} \\ \le& MH^{2\kappa-2}\frac{2^{2M_1}k^{2\kappa}}{k^{2M_1}}\log\frac{1}{\delta},\end{aligned}$$

holds for all k uniformly with probability $1 - \frac{\pi^2\delta}{6}$, where the first inequality is from doubling trick. i.e., $p \ge \beta\log\frac{k}{2}$. Hence we only need to choose $M_1 > \kappa + \log H$, then there is

$$\mathbb{P}\left\{\mathcal{B}_{k,p} \le M\log\frac{1}{\delta},\ \ \forall k \ge T_{\text{init}}\right\} \ge 1 - \frac{\pi^2\delta}{6}.$$

For any fixed N, we choose $\delta = \frac{1}{N}$, then we further have

$$\mathbb{P}\left\{\max_{k\le N}\mathcal{B}_{k,p} \le M\log N\right\} \ge 1 - \frac{\pi^2}{6N}.$$

With the large enough N, we can conclude that $\max_{k\le N}\mathcal{B}_{k,p} \le MH\log N$ holds almost surely. Moreover, for the term $\mathfrak{b}_{N,p}$, note that the value of p varies with the epoch index l. Then we divide $\mathfrak{b}_{N,p}$ apart. Similar to the previous analysis, we have

$$\mathfrak{b}_{N,p} = \sum_{l=1}^{N_E}\sum_{k=T_l}^{2T_l-2}\|b_{k+H,p}\|_2^2 \le \frac{\log(N/T_{\text{init}})}{\log 2}M\log\frac{1}{\delta}$$

holds uniformly for all N with high probability $1 - \frac{\pi^2\delta}{6}$. Therefore by letting $\delta = \frac{1}{N}$, we can conclude that $\mathfrak{b}_{N,p} \le M\log^2 N$ holds almost surely. We have completed the proof that it is sufficient to guarantee the uniform boundedness of bias error only with the parameter β chosen to be proportional to $1/\log\rho(A-LC)$ and $\kappa + \log H$.

D. Proof of Lemma 3

From Theorem 1, we first have

$$r_{k+H} = e_{k+H} + \sum_{i=1}^{H-1} CA^{i-1}Le_{k+H-i}.$$

Then we denote $E_{p:k} \triangleq \left[e_p, \ldots, e_k\right]$, we obtain that

$$\mathcal{R}_{k,p} = E_{p+H:k} + \sum_{i=1}^{H-1} CA^{i-1}LE_{p+H-i:k-i}$$

With the Cauchy-Schwarz inequality, we further have

$$\begin{aligned}\mathcal{E}_{k,p} \le& H\left\|E_{p+H:k}\bar{Z}_{k-H,p}^{\mathsf{T}}\bar{V}_{k-H,p}^{-\frac{1}{2}}\right\|_2^2 + H\sum_{i=1}^{H-1}\left\|CA^{i-1}L\right\|_2^2 \\ &\times\left\|E_{p+H-i:k-i}\bar{Z}_{k-H,p}^{\mathsf{T}}\bar{V}_{k-H,p}^{-\frac{1}{2}}\right\|_2^2.\end{aligned}$$

Due to the conditional independence between e_k and $Z_{k-l,p}$, $\forall l > 0$, with [15, Theorem 3], we have

$$\begin{aligned}&\left\|\bar{R}^{-\frac{1}{2}}E_{p+H-i:k-i}\bar{Z}_{k-H,p}^{\mathsf{T}}V_{k-H,p}^{-\frac{1}{2}}\right\|_2^2 \\ &\qquad\le m\log 5 + \log\frac{H}{\delta} + \log\frac{\det V_{k-H,p}}{\det\lambda I}.\end{aligned}$$

holds for all $k \ge T_{\text{init}}$ and $1 \le i \le H-1$ uniformly with probability at least $1-\delta$, where $\bar{R} \triangleq CPC^{\mathsf{T}} + R$ is the covariance of innovation e_k Then we provide an almost sure bound for the matrix $V_{k-H,p}$. Without loss of generality, we only need to consider the uniform bound of $V_{k,p}$ for all $k \ge T_{\text{init}}$. First note that $V_{k,p} = \lambda I + \sum_{l=p}^{k} Z_{l,p}Z_{l,p}^{\mathsf{T}}$, then denote $\Gamma_{k,p}^{Z} \triangleq \mathbb{E}\left\{Z_{k,p}Z_{k,p}^{\mathsf{T}}\right\}$, we have

$$\begin{aligned}\left\|\Gamma_{k,p}^{Z}\right\|_2^2 \le& \mathbb{E}\left\{Z_{k,p}^{\mathsf{T}}Z_{k,p}\right\} \\ =& \operatorname{tr}\Big(\sum_{i=k-p+1}^{k}\mathbb{E}\left\{y_iy_i^{\mathsf{T}}\right\} + \sum_{i=k-p+1}^{k+H-1}\mathbb{E}\left\{u_iu_i^{\mathsf{T}}\right\}\Big) \\ \le& p\operatorname{tr}(R) + (p+H)n_u \\ &+ \operatorname{tr}\Big(C^{\mathsf{T}}C\sum_{i=k-p+1}^{k}\sum_{l=0}^{i-1}A^l(Q+BB^{\mathsf{T}})(A^l)^{\mathsf{T}}\Big) \\ \le& mp\|R\|_2 + (p+H)n_u + npM\sum_{l=0}^{k-1}l^{2\kappa-2} \\ \le& mp\|R\|_2 + (p+H)n_u + npMk^{2\kappa-1}.\end{aligned}$$

With Lemma 5, we have

$$\left(\Gamma_{k,p}^{Z}\right)^{-\frac{1}{2}}Z_{k,p}Z_{k,p}^{\mathsf{T}}\left(\Gamma_{k,p}^{Z}\right)^{-\frac{1}{2}} \le \left(2(mp+(p+H)n_u) + 3\log\frac{k^2}{\delta}\right)I$$

holds for all $k \ge T_{\text{init}}$ uniformly with probability $1 - \frac{\pi^2\delta}{6}$, together with the condition $p \le \beta\log k$, we can obtain that for each $k \le N$, there is

$$\bar{Z}_{k,p}\bar{Z}_{k,p}^{\mathsf{T}} = \sum_{l=p}^{k}Z_{l,p}Z_{l,p}^{\mathsf{T}} \le \left(M\beta^2k^{2\kappa}\log^2 k\log\frac{1}{\delta}\right)I$$

holds uniformly with probability $1-\frac{\pi^2\delta}{6}$. For any N, we further choose $\delta=\frac{1}{N}$, then we can obtain that

$$\bar{Z}_{k,p}\bar{Z}_{k,p}^{\mathsf{T}} \leq \left(M\beta^2 N^{2\kappa+1}\right) I, \ \ \forall k \leq N$$

holds uniformly with probability at least $1-\frac{\pi^2}{6N}$. Furthermore, we can obtain

$$\begin{aligned}\log\det V_{k-H,p} &\leq \log\det\left(\lambda I+\bar{Z}_{k,p}\bar{Z}_{k,p}^{\mathsf{T}}\right)\\ &\leq (mp+(p+H)n_u)\log(M\beta^2N^{2\kappa+1}),\end{aligned}$$

and for sufficiently large N, the above bound holds almost for sure. Then for the term $\mathcal{E}_{k,p}$, we further have that

$$\begin{aligned}\mathcal{E}_{k,p} \leq & H\left(m\log 5+\log\frac{H}{\delta}+\log\frac{\det V_{k-H,p}}{\det\lambda I}\right)\\ &\times\left\|\bar{R}\right\|_2^2\left(1+\sum_{i=1}^{H-1}\|C\|_2^2\|L\|_2^2\left\|A^{i-1}\right\|_2^2\right)\\ \leq & MH\left(1+\sum_{i=1}^{H-1}i^{2\kappa-2}\right)\\ &\times\left(\log\frac{H}{\delta}+(mp+(p+H)n_u)\log(M\beta N^{2\kappa+1})\right),\end{aligned}$$

uniformly holds for all $k\leq N$ with probability at least $1-\delta$. Similarly, for fixed N, with $p\leq\beta\log N$, we choose $\delta=\frac{1}{N}$, then we have the term $\log\frac{H}{\delta}=\log HN$ dominated by $p\log(M\beta N^{2\kappa+1})\leq\beta\log^2 N$. We finally have that

$$\max_{k\leq N}\mathcal{E}_{k,p}\leq MH^{2\kappa}\beta\log^2 N,$$

holds almost for sure.

E. Proof of Lemma 4

We first relax the term $\mathcal{V}_{N,p}$ as

$$\mathcal{V}_{N,p}\leq\max_{T_{\text{init}}\leq k\leq N}\left\|V_{k-H,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2\times\sum_{l=1}^{N_E}\sum_{k=T_l}^{2T_l-2}\left\|V_{k,p_l}^{-\frac{1}{2}}Z_{k,p_l}\right\|_2^2,$$

and we first consider the uniform boundedness of $\left\|V_{k-H,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2$. Note that

$$V_{k,p}=V_{k-H,p}+\sum_{l=k-H+1}^{k}Z_{l,p}Z_{l,p}^{\mathsf{T}}.$$

Then we have

$$V_{k-H,p}^{-\frac{1}{2}}V_{k,p}V_{k-H,p}^{-\frac{1}{2}}=I+V_{k-H,p}^{-\frac{1}{2}}\Big(\sum_{l=k-H+1}^{k}Z_{l,p}Z_{l,p}^{\mathsf{T}}\Big)V_{k-H,p}^{-\frac{1}{2}}.$$

Therefore we can bound the $\left\|V_{k-H,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2$ by

$$\left\|V_{k-H,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2\leq 1+\sum_{l=k-H+1}^{k}Z_{l,p}^{\mathsf{T}}V_{k-H,p}^{-1}Z_{l,p}$$

To bound $Z_{l,p}^{\mathsf{T}}V_{k-H,p}^{-1}Z_{l,p}$ for $l>k-H$, we need to consider the successive representation of $Z_{l,p}$ with $Z_{s,p}$, $s<l$. Consider the minimal polynomial of A as

$$A^d=a_{d-1}A^{d-1}+\cdots+a_0,$$

and the companion matrix of A can be written as

$$\mathcal{A}=\begin{bmatrix}0&0&\cdots&0&a_0\\1&0&\cdots&0&a_1\\\vdots&\vdots&\vdots&\ddots&\vdots\\0&0&\cdots&1&a_{d-1}\end{bmatrix},$$

where d is the dimension of the minimal polynomial of A. Then similar to [15, Lemma 2], we first derive the innovation representation of output y_k as

$$y_k=CA^d\hat{x}_{k-d}+\sum_{i=1}^{d}CA^{i-1}Bu_{k-i}+e_k+\sum_{i=1}^{d}CA^{i-1}Le_{k-i}.$$

Then, with the minimal polynomial of A, we can derive the successive representation of the y_k as

$$y_k=a_{d-1}y_{k-1}+\ldots+a_0y_{k-d}+\delta_k$$

where

$$\begin{aligned}\delta_k&=\sum_{s=0}^{d}L_se_{k-s}+\sum_{s=1}^{d}K_su_{k-s},\\ L_s&=-a_{d-s}I_m+CA^{s-1}L-\sum_{t=1}^{s-1}a_{d-s+t}CA^{t-1}L,\ L_0=I,\\ K_s&=CA^{s-1}B-\sum_{l=1}^{s-1}a_{d-s+l}CA^{l-1}B.\end{aligned}$$

Then for the augmented form, further denote $\tilde{E}_{k,p}=\left[e_{k-p+1}^{\mathsf{T}},\ldots,e_k^{\mathsf{T}}\right]^{\mathsf{T}}$, $\tilde{U}_{k,p}=\left[u_{k-p+1}^{\mathsf{T}},\ldots,u_k^{\mathsf{T}}\right]^{\mathsf{T}}$, we can rewrite $Z_{k,p}$ as

$$Z_{k,p}=a_{d-1}Z_{k-1,p}+\ldots+a_0Z_{k-d,p}+\delta_{k,p}. \tag{30}$$

The term $\delta_{k,p}$ takes the form as $\delta_{k,p}=\left[\delta_{k,p}^{(1)\mathsf{T}},\delta_{k,p}^{(2)\mathsf{T}}\right]^{\mathsf{T}}$, with

$$\delta_{k,p}^{(1)}=\sum_{s=0}^{d}\text{diag}_p\left(L_s\right)\tilde{E}_{k-s,p}+\sum_{s=1}^{d}\text{diag}_p\left(K_s\right)\tilde{U}_{k-s,p},$$

and $\delta_{k,p}^{(2)}=\tilde{U}_{k+H-1,p+H}-\sum_{s=0}^{d-1}a_s\tilde{U}_{k+H-1+s-d,p+H}$, where $\text{diag}_p\left(L\right)=\text{diag}\left(\underbrace{L,\ldots,L}_{p}\right)$. Furthermore, we denote $a_l^{(s)}$ as the $(l+1,d)$-th element of matrix $\mathcal{A}^s$. Then substitute the successive representation of $Z_{k-1,p}$ into eq. (30), we can obtain

$$Z_{k,p}=a_{d-1}^{(2)}Z_{k-2,p}+\ldots+a_0^{(2)}Z_{k-d-1,p}+\delta_{k,p}+a_{d-1}^{(1)}\delta_{k-1,p}.$$

By performing the above recursion for H times, we can obtain

$$Z_{k,p}=\sum_{i=1}^{d}a_{i-1}^{(H)}Z_{k-H+i-d,p}+\delta_{k,p}+\sum_{j=1}^{H-1}a_{d-1}^{(j)}\delta_{k-j,p}. \tag{31}$$

For each $k-H+1\leq l\leq k$, consider the term $Z_{l,p}^{\mathsf{T}}V_{k-H,p}^{-1}Z_{l,p}$, by substituting the successive representation (31) and applying

Cauchy-Schwarz inequality, we can obtain

$$Z_{l,p}^{\mathsf{T}}V_{k-H,p}^{-1}Z_{l,p} \le (H+1)\Big(\sum_{i=1}^{d} a_{i-1}^{(H)}Z_{l-H+i-d,p}\Big)^{\mathsf{T}}V_{k-H,p}^{-1} \times \Big(\sum_{i=1}^{d} a_{i-1}^{(H)}Z_{l-H+i-d,p}\Big) + (H+1)\delta_{l,p}^{\mathsf{T}}V_{k-H,p}^{-1}\delta_{l,p} + (H+1)\sum_{j=1}^{H-1} a_{d-1}^{(j)}\delta_{l-j,p}^{\mathsf{T}}V_{k-H,p}^{-1}a_{d-1}^{(j)}\delta_{l-j,p}. \quad (32)$$

To provide a uniform bound for the above terms, we need to first provide a uniform bound for each $\delta_{l-j,p}$. We can verify that for fixed d, the norm of L_s is uniformly bounded for all $s = 0, \ldots, d-1$, Then we have the following bound for $\delta_{k,p}$ that

$$\|\delta_{k,p}\|_2^2 \le Mp\max_{k\le N}\|e_k\|_2^2 + M(p+H)\max_{k\le N+H}\|u_k\|_2^2$$

where M is only related to system parameters. With Lemma 5, for fixed N, we have

$$\mathbb{P}\left\{\|e_k\|_2^2 \le M\log\frac{k^2}{\delta},\ \ \forall k \ge T_{\text{init}}\right\} \ge 1 - \frac{\pi^2\delta}{6}.$$

Then choose $\delta = \frac{1}{N}$, we have

$$\max_{k\le N}\|e_k\|_2^2 \le 3M\log N$$

almost for sure for any fixed N. Moreover, due to $u_k \sim \mathcal{N}(0, I_{n_u})$, we also have

$$\max_{k\le N+H}\|u_k\|_2^2 \le 3M\log(N+H)$$

almost for sure for any fixed N and H. Then together with $p \le \beta\log N$ and $N \gg H$, we have

$$\max_{k\le N}\|\delta_{k,p}\|_2^2 \le M\beta\log^2 N$$

holds almost for sure, where H is eliminated as a low-order term.

Then we reconsider the term $Z_{l,p}^{\mathsf{T}}V_{k-H,p}^{-1}Z_{l,p}$, together with $V_{k-H,p} \ge \lambda I$, we have

$$a_{d-1}^{(j)}\delta_{l-j,p}^{\mathsf{T}}V_{k-H,p}^{-1}a_{d-1}^{(j)}\delta_{l-j,p} \le \frac{\left(a_{d-1}^{(j)}\right)^2}{\lambda}M\beta\log^2 N$$

for each $1 \le j \le H-1$. Moreover, from Woodbury Equality, there is also

$$Z_{k,p}^{\mathsf{T}}\left(V_{k-1,p} + Z_{k,p}Z_{k,p}^{\mathsf{T}}\right)^{-1}Z_{k,p} = \frac{Z_{k,p}^{\mathsf{T}}V_{k-1,p}^{-1}Z_{k,p}}{1 + Z_{k,p}^{\mathsf{T}}V_{k-1,p}^{-1}Z_{k,p}} \le 1.$$

Hence we have

$$\Big(\sum_{i=1}^{d} a_{i-1}^{(H)}Z_{l-H+i-d,p}\Big)^{\mathsf{T}}V_{k-H,p}^{-1}\Big(\ \cdot\ \Big) \le d\sum_{i=1}^{d}\left(a_{i-1}^{(H)}\right)^2$$

where $(\cdot) = \sum_{i=1}^{d} a_{i-1}^{(H)}Z_{l-H+i-d,p}$ for brevity and

$$Z_{l,p}^{\mathsf{T}}V_{k-H,p}^{-1}Z_{l,p} \le d(H+1)\sum_{i=1}^{d}\left(a_{i-1}^{(H)}\right)^2 + (H+1)M\beta\log^2 N + (H+1)\sum_{j=1}^{H-1}\frac{\left(a_{d-1}^{(j)}\right)^2}{\lambda}M\beta\log^2 N.$$

Note that the term $a_{d-1}^{(j)}$ is from matrix $\mathcal{A}^j$, and the matrix $\mathcal{A}$ shares the same minimal polynomial with A. Hence we have $\left\|\mathcal{A}^j\right\|_F^2 \le Mj^{2\kappa-2}$, and $\left(a_{i-1}^{(j)}\right)^2 \le Mj^{2\kappa-2}$ uniformly holds for all $i \le d$ and $j \le H$. Finally, we have

$$\max_{T_{\text{init}}\le k\le N}\left\|V_{k-H,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2 \le MH^{2\kappa+1}\beta\log^2 N$$

holds almost for sure, where the constant term $d\sum_{i=1}^{d}\left(a_{i-1}^{(H)}\right)^2$ will also be dominated by $\log^2 N$ for large N.

Then we consider the term $\sum_{l=1}^{N_E}\sum_{k=T_l}^{2T_l-2}\left\|V_{k,p_l}^{-\frac{1}{2}}Z_{k,p_l}\right\|_2^2$. For the accumulation error at the l-th epoch, where $T_l = 2^{l-1}T_{\text{init}} + 1$, from [29, Lemma 2] (also [15, Lemma 1]), we first have the following result for the accumulation error

$$\sum_{k=T_l}^{2T_l-2}\left\|V_{k,p}^{-\frac{1}{2}}Z_{k,p}\right\|_2^2 \le \log\frac{\det(V_{2T_l-2,p})}{\det(V_{T_l-1,p})}.$$

From Section D, we have

$$\log\frac{\det(V_{2T_l-2,p})}{\det(V_{T_l-1,p})} \le \log\frac{\det(V_{2T_l-2,p})}{\det(\lambda I)} \le M\beta\log^2 N$$

almost for sure for each $l = 1, \ldots, N_E$, then we have

$$\sum_{l=1}^{N_E}\sum_{k=T_l}^{2T_l-2}\left\|V_{k,p_l}^{-\frac{1}{2}}Z_{k,p_l}\right\|_2^2 \le M\beta N_E\log^2 N$$

Note that $N_E = \frac{\log N/T_{\text{init}}}{\log 2}$, then we have

$$\sum_{l=1}^{N_E}\sum_{k=T_l}^{2T_l-2}\left\|V_{k,p_l}^{-\frac{1}{2}}Z_{k,p_l}\right\|_2^2 \le M\beta\log^3 N$$

and

$$\mathcal{V}_{N,p} \le \max_{T_{\text{init}}\le k\le N}\left\|V_{k-1,p}^{-\frac{1}{2}}V_{k,p}^{\frac{1}{2}}\right\|_2^2 \times \sum_{l=1}^{N_E}\sum_{k=T_l}^{2T_l-2}\left\|V_{k,p_l}^{-\frac{1}{2}}Z_{k,p_l}\right\|_2^2 \le MH^{2\kappa+1}\beta^2\log^5 N$$

holds almost for sure for any N.

F. Supplementary Lemma

Lemma 5: For any given $\delta \in (0,1)$ and for any Gaussian random vector sequence X_k satisfies $X_k \sim \mathcal{N}(0, I_{p(k)})$, where $p(k)$ is a function of k, we define event $\mathcal{T}_X$ as

$$\mathcal{T}_X \triangleq \left\{\|X_k\|_2^2 \le 2p(k) + 3\log\frac{k^2}{\delta},\ \ \forall k \ge 1\right\},$$

then the event $\mathcal{E}_X$ holds with probability at least $1 - \frac{\pi^2\delta}{6}$.

Proof: From [30, Lemma 1], for any Gaussian random vector $X_k \sim \mathcal{N}(0, I_{p(k)})$. For each $k \ge 1$, we have $\mathbb{P}\left\{\|X_k\|_2^2 \ge p(k) + 2\sqrt{p(k)}\cdot t + 2t^2\right\} \le e^{-t^2}$. For each time step k, take $t = \sqrt{\log\frac{k^2}{\delta}}$ and with inequality $2ab \le a^2 + b^2$, then we have $\mathbb{P}\left\{\|X_k\|_2^2 \le 2n + 3\log\frac{k^2}{\delta}\right\} \ge 1 - \frac{\delta}{k^2}$. Then, take a union bound over all k, we have

$$\mathbb{P}\left\{\|X_k\|_2^2 \le 2p(k) + 3\log\frac{k^2}{\delta}, \forall k \ge 1\right\} \ge 1 - \sum_{k=1}^{\infty}\frac{\delta}{k^2},$$

Then with $\sum_{k=1}^{\infty} 1/k^2 = \pi^2/6$, this lemma is proved. ∎