Finite Time Logarithmic Regret Bounds for Self-Tuning Regulation

Rahul Singh [1] **Akshay Mete** [2] **Avik Kar** [1] **P. R. Kumar** [2]

Abstract

We establish the first finite-time logarithmic regret bounds for the self-tuning regulation problem. We introduce a modified version of the certainty equivalence algorithm, which we call PIECE, that clips inputs in addition to utilizing probing inputs for exploration. We show that it has a $C \log T$ upper bound on the regret after T time-steps for bounded noise, and $C \log^3 T$ in the case of sub-Gaussian noise, unlike the LQ problem where logarithmic regret is shown to be not possible. The PIECE algorithm is also designed to address the critical challenge of poor initial transient performance of reinforcement learning algorithms for linear systems. Comparative simulation results illustrate the improved performance of PIECE.

1. Introduction

Minimizing the variance of the output of a system with respect to a desired set-point is an important problem in control systems. When the system model is unknown, this results in the extensively studied "self-tuning regulation" problem. It has a large number of engineering applications, including large scale paper production, ore crushing, adaptive autopiloting of supertankers, etc. It differentiates itself from LQ control or self-tuning "controllers" by the fact that it is a singular limit where there is no added penalty on control effort.

Self-tuning regulators have been deeply examined with respect to several asymptotic properties such as stability (Goodwin et al., 1981), whether it asymptotically achieves the minimum variance (Goodwin et al., 1981), whether the controller self-tunes itself asymptotically to the optimal controller for the unknown system (Becker et al., 1985; Kumar & Praly, 1987), and how to achieve asymptotically optimal regret (Lai, 1986; Lai & Wei, 1987).

However, in many applications, in addition to the asymptotics, how to control the system shortly after its initialization so as to control its transient behavior is also important since one wishes to avoid large initial transients. This requires the design of transient control and analysis of the finite time regret of the singular learning system. This problem has remained an open problem. Such an analysis of finite time behavior has been conducted for the so-called Linear Quadratic (LQ) control problems that feature a strictly positive definite weighting on the control inputs in addition to quadratic weighting of the system's outputs (Abbasi-Yadkori & Szepesvári, 2011; Lale et al., 2022; Cohen et al., 2019; Mete et al., 2022; Jedra & Proutiere, 2022; Simchowitz & Foster, 2020; Mania et al., 2019; Faradonbeh et al., 2020; Shirani Faradonbeh et al., 2020; Abeille & Lazaric, 2017). However, these analyses do not apply to the minimum variance control problem since it is a singular problem that does not feature a positive definite quadratic penalty on the control inputs. Moreover, LQ problems generally only admit $C\sqrt{T}$ regret, rather than the logarithmic regret that we establish in the singular minimum variance case.

1.1. Background on Self-Tuning Regulation and Difference from Adaptive LQ

The problem of controlling a paper machine was the seminal application that launched the field of self-tuning regulators (Åström & Wittenmark, 1973), (Fjeld & Wilhelm, 1981), (Borisson & Wittenmark, 1974), (Cegrell & Hedqvist, 1975), (Åström, 1967), (Åström, 1987). Paper manufacture is subject to several uncertainties. These include pulp quality, moisture control, machine speed, pressure distribution, etc. Thus the basis weight of the paper produced, related to its thickness, is "stochastic." At the same time, the sale of the paper produced is related to its minimum basis weight. In order to provide a guarantee on minimum basis weight for the paper produced, the manufacturer therefore has to set a set-point that is larger than the mean by a multiple of the standard deviation of basis weight. If one can reduce the variance of basis weight, then one can accordingly reduce the set-point, which leads to a large savings of paper pulp, and concomitantly, cost. Similarly, in other applications too, reducing the variance of the output improves product

[1]Indian Institute of Science, Bengaluru, Karnataka, India. [2]Texas A&M University, College Station, TX, USA. Correspondence to: Rahul Singh <rahulsingh0188@gmail.com, rahulsingh@iisc.ac.in>.

Proceedings of the 41^{st} International Conference on Machine Learning, Vienna, Austria. PMLR 235, 2024.

quality/control system performance; e.g., supertanker ship steering (Brink & Tiano, 1981) or ore crushers (Borisson & Syding, 1976).

The resulting minimum variance (MV) control problem, called the "regulation" problem, is a singular limit of the linear, quadratic (LQ) control problem. The latter features a positive definite penalty on the quadratic of the control effort, in addition to a quadratic on the system output, while the former has no penalty on control input. As we show in this paper, this allows the attainment of a very different order of regret performance results. The LQ problem admits a $\Omega(\sqrt{T})$ lower bound on regret (Simchowitz & Foster, 2020),(Ziemann & Sandberg, 2022), based upon the positive definiteness of the matrix penalizing control costs. However, in the regulation problem considered here, we show that one can achieve a logarithmic regret upper bounded by $C_1 \log T + C_2$ for each T.

The following is an outline of the highlights of the steps used to establish a $O(\log T)$ regret: The regret during the exploratory steps $\mathcal{I}$ can be bounded proportional to its length. For times $t \notin \mathcal{I}$, we show that "prediction error" associated with predicting the output y_{t+1} is closely connected to instantaneous regret. We then use the recursion in Lemma C.3 to show that regret can be bounded by five terms, each of which can be bounded separately. Central to the analysis is the lower bound on the minimum eigenvalue of the covariance matrix, which is derived in Appendix F. This lower bound yields tight upper bounds on the estimation error, as well as on the norm of Y_t, U_t. The "bottleneck term" which has the highest contribution to the regret turns out to be $\mathcal{T}_{4,2}$, and its upper-bound relies upon the analyses in Appendix F, K.

1.2. Our Contributions

This paper proposes a design of a reinforcement learning (RL) algorithm for the singular minimum variance problem and its finite time analysis. We show how the certainty-equivalence scheme can be modified so as to have provably good transient performance. We propose a RL scheme that employs a "clipping" mechanism to bound the control behavior during learning. We also employ an approach of using increasingly infrequent probing intervals when white noise is injected into the system. This differs from another approach of always using additive noise, but of diminishing excitation, used in (Guo et al., 1991). The resulting algorithm, which we call PIECE, is a simplified and optimized version of the asymptotically regret-optimal policy of (Lai & Wei, 1987). We show the following two results:

Result 1: Bounded noise
If the noise is bounded, then PIECE results in a finite-time regret bound,

$$\text{Regret}(T) \leq C\sigma^2(p+q-1)\log T + C'\sqrt{\log T \cdot \log\log T} + C'',$$

with probability greater than $(1-\delta)$ after T steps, where $C \approx 1$, p and q are the orders of the system model, and σ^2 is the variance of the noise. The dominant first term closely matches the asymptotically optimal regret (1) established by (Lai & Wei, 1987).

Result 2: Sub-Gaussian noise
If the noise is sub-Gaussian with a proxy variance of σ^2, then PIECE results in a finite-time regret bound,

$$\text{Regret}(T) \leq C''' \log T \log^2(T/\delta) + C''''\sigma^2(p+q-1)\log^2 T,$$

with probability greater than $(1-\delta)$ after T steps.

We also present the results of simulations that show how the proposed PIECE controller performs in comparison to the asymptotically regret-optimal policy of (Lai & Wei, 1987) as well as the certainty equivalence scheme.

1.3. Prior Work

Self-tuning regulators were first proposed in (Åström & Wittenmark, 1973). The motivating application was the regulation of the thickness of paper produced in a paper mill, as measured by its variance. The adaptive control law with a stochastic gradient estimator was shown to be stable and asymptotically self-optimal in (Goodwin et al., 1981), and asymptotically self-tuning in (Becker et al., 1985; Kumar & Praly, 1987). For least squares estimates, the stability, self-optimality, and self-tuning under a modification to the adaptive control law featuring an additional diminishing excitation were established in (Guo et al., 1991). The robustness of an appropriately modified minimum variance controller to modeling assumptions, e.g., perturbations with respect to the graph topology in the space of transfer functions (Vidyasagar, 1984), which lead to infinite dimensional systems, was shown in (Praly et al., 1989). A modified certainty equivalence approach was proposed in (Lai, 1986; Lai & Wei, 1987) by introducing "exploration episodes" where the algorithm, which we refer to as LW in the sequel, utilizes probing inputs – in contrast to the diminishing excitation approach of (Guo et al., 1991). Its asymptotic regret was shown to be

$$\limsup_{T\to\infty} \frac{R_T}{\log T} = \sigma^2(p+q-1). \tag{1}$$

It was shown that asymptotic regret of the LW algorithm is optimal in the sense that it matches the asymptotic regret lower bound. There has been no prior work on finite time bounds for the minimum variance problem.

There has been work on finite-time bounds for quadratic costs that include a strictly positive definite quadratic cost on the control inputs. High probability, finite-time analysis of adaptive controllers for linear systems was initiated by (Abbasi-Yadkori & Szepesvári, 2011). However, except in very limited special cases, e.g., where the A or B matrices are known (Jedra & Proutiere, 2022; Cassel et al., 2020), the finite-time regret bounds achievable are shown to be $C\sqrt{T}$ rather than logarithmic. These include the optimism-based designs StabL (Lale et al., 2022), OSLO (Cohen et al., 2019), and ARBMLE (Mete et al., 2022), modified versions of CE algorithms including (Jedra & Proutiere, 2022), CECCE (Simchowitz & Foster, 2020; Mania et al., 2019), RCE (Faradonbeh et al., 2020), and IP (Shirani Faradonbeh et al., 2020), and Thompson Sampling-based algorithms (Abeille & Lazaric, 2017). Notably, (Lale et al., 2020) derives logarithmic regret bound for partially observed LQ systems, in which regret is defined with respect to the best "persistently exciting policy." However, it was noted in (Tsiamis et al., 2023) that the optimal policy might not necessarily be "persistently exciting." In fact, a lower bound of $C\sqrt{T}$ was shown for the fully observable LQ setting in (Simchowitz & Foster, 2020) for any algorithm. Moreover, the minimum variance problem is a singular problem where there is no positive definite weighting on the control cost, and the LQ results are not applicable to it.

1.4. Organization of the Paper

The notations used in the paper are described in Section 1.5. We describe the system model in Section 2. In Section 3, we describe the PIECE self-tuning regulation algorithm. We present the logarithmic bounds on regret and their proofs in Section 4. In Section 5 we present the comparative results obtained in simulation experiments. We conclude with a brief discussion on open problems in Section 6. All the detailed proofs and additional simulation results are provided in the appendices.

1.5. Notation

For a matrix M, we use $\det(M)$,$\mathrm{Tr}(M)$ and $\|M\|$ to denote its determinant, trace, and operator norm induced by the Euclidean norm, respectively. For a vector x, let x' be its transpose, and $\|x\|$ its Euclidean norm. We use $a \wedge b$ to denote the minimum, and $a \vee b$ the maximum of two numbers a, b. We use the abbreviations "w.h.p." to denote "with high probability." Throughout, to keep notation simple, we use $\lesssim$ and $\gtrsim$ to hide problem-dependent constants. We denote the projection of a vector x onto a vector subspace S by $\mathrm{proj}(x, S)$. We use $\mathbb{N}$ to denote the set of natural numbers, $\mathbb{Z}$ the set of integers, and $\mathbb{Z}_+$ the set of positive integers. For $m \in \mathbb{N}$, we use I_m to denote the identity matrix of size m.

2. System Model

Consider an auto-regressive linear system with exogenous inputs (ARX system),

$$\begin{aligned} y_t =& a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + b_1 u_{t-1} \\ &+ b_2 u_{t-2} + \cdots + b_q u_{t-q} + w_t, \ \forall\, t \in \mathbb{N}, \end{aligned} \tag{2}$$

where y_t, u_t are the output and input at time t respectively, and w_t is the system noise at time t. $\{w_t\}$ is assumed to be a martingale difference sequence with respect to filtration $\{\mathcal{F}_t\}$, with initial values $y_0, y_{-1}, \ldots, y_{1-p}, u_0, u_{-1}, \ldots, u_{1-q}$ equal to 0. The optimal control law to minimize the variance of the output, the minimum variance (MV) control law, is (Åström, 2012),

$$\begin{aligned} u_t =& -(1/b_1)(a_1 y_t + a_2 y_{t-1} + \ldots + a_p y_{t-p+1} \\ &+ b_2 u_{t-1} + \cdots + b_q u_{t-q+1}). \end{aligned} \tag{3}$$

The above ARX model (2) can be written as:

$$y_t = \phi'_{t-1}\theta^\star + w_t. \tag{4}$$

where, $\theta^\star := (a_1, a_2, \ldots, a_p, b_1, b_2, \ldots, b_q)'$ and $\phi_{t-1} := (y_{t-1}, y_{t-2}, \ldots, y_{t-p}, u_{t-1}, u_{t-2}, \ldots, u_{t-q})'$.

We make the following assumption regarding the unknown linear system (2).

Assumption 2.1. The parameter $\theta^\star$ associated with the ARX model (2) satisfies the following condition: the polynomials $s^p - a_1 s^{p-1} - a_2 s^{p-2} - \ldots - \alpha_p$ and $b_1 s^{q-1} + b_2 s^{q-2} + \ldots + b_q$ have all zeroes inside the open unit disk. Moreover, $b_1 \neq 0$.

The latter minimum phase assumption is necessary for internal stability of the MV control law (Kumar & Varaiya, 1986). The former assumption can be replaced by the assumption that a stabilizing linear control law is known, and is also used for regret analysis in (Lai, 1986; Lai & Wei, 1987).

Define the vectors $Y_t := (y_t, y_{t-1}, \ldots, y_{t-p+1})'$ and $U_t := (u_t, u_{t-1}, \ldots, u_{t-q+2})'$, and the matrices $A := \begin{pmatrix} a_1 & \cdots & a_{p-1} & a_p \\ I_{p-1} & & & 0 \end{pmatrix}$, and $B := \begin{pmatrix} -b_2/b_1 & \cdots & -b_q/b_1 \\ I_{q-2} & & 0 \end{pmatrix}$, where I_{p-1}, I_{q-2} are identity matrices of sizes $p-1$ and $q-2$ respectively. We have $\|A^n\| \le C_1\rho^n, \|B^n\| \le C_1\rho^n$, where $\rho < 1$ can be taken to be any number greater than the spectral radii of A and B. When we want to depict the dependence of these quantities upon the coefficients in a parameter vector θ, we use $\rho(\theta), A(\theta), B(\theta), C_1(\theta)$. For objects pertaining to $\theta^\star$ (the true parameter), we suppress this dependence and simply write ρ, A, B, C_1, etc.

Assumption 2.2. The noise $\{w_t\}$ is a martingale difference sequence with respect to filtration $\{\mathcal{F}_t\}$, with conditional variance bounded away from 0, i.e.,

$$\inf_t \mathbb{E}\big(w_t^2|\mathcal{F}_{t-1}\big) > c_1 > 0, \text{ a.s.}. \quad (5)$$

Also,

$$\mathbb{E}\big(w_t^2|\mathcal{F}_{t-1}\big) \leq \sigma^2, \text{ a.s., } \forall t. \quad (6)$$

Within this setup we consider two possibilities, either bounded or sub-Gaussian noise:

Assumption 2.3 (Bounded noise)**.** $\{w_t\}$ is uniformly bounded a.s., i.e.,

$$|w_t| \leq B_w, \quad \text{a.s., } \forall t. \quad (7)$$

We present the main result on regret and a proof sketch under Assumption 2.3 in Section 4.1. This is then relaxed in Section 4.2 to allow for unbounded noise:

Assumption 2.4 (Conditionally sub-Gaussian noise)**.** For all $\gamma \in \mathbb{R}$, we have,

$$\mathbb{E}\left\{\exp(\gamma|w_t|)\Big|\mathcal{F}_{t-1}\right\} \leq \exp\big(\gamma^2\sigma^2/2\big) \text{ a.s., } \forall t, \quad (8)$$

where $\sigma > 0$.

We also assume the following prior information about the unknown system, as in (Lai & Wei, 1987):

Assumption 2.5. The learning algorithm has knowledge of a compact set $\Theta \subset \mathbb{R}^{p+q}$ that contains the true parameter value $\theta^\star$.

Regret of Self-Tuning Regulators

Noting that the MV controller results in $y_t \equiv w_t$, we judge the performance of an algorithm $\mathcal{A}$ by its cumulative learning regret,

$$\mathcal{R}_T(\mathcal{A}) := \sum_{t=1}^{T}(y_t - w_t)^2. \quad (9)$$

A discussion on the equivalence of the above definition of regret with the standard definition found in the reinforcement learning literature (Lai & Robbins, 1985; Auer et al., 2002; Auer & Ortner, 2007) can be found in Appendix B.

3. PIECE: An Adaptive Minimum Variance Control Algorithm

The PIECE algorithm is presented in Algorithm 1. It divides the total operation time into two functionalities: (i) Exploration: This consists of a sequence of intervals during which white noise, by which is meant i.i.d., mean 0 and constant variance noise, is used as the control input to ensure sufficient excitation of the system, which in-turn yields consistent estimates, and (ii) Exploitation: The rest of the time, where a standard Certainty Equivalence (CE) controller is applied by generating controls that are optimal under the assumption that the least squares estimates are equal to the true parameter values. Though this structure is inspired by the algorithm of (Lai & Wei, 1987) (henceforth dubbed LW), we will highlight in the sequel some major differences which lead to better transient performance and make possible a finite-time regret analysis.

Exploration: The set of exploration time instants is denoted by $\mathcal{I}$. For $t \in \mathcal{I}$, u_t is an i.i.d. mean 0 sequence that is independent of $\{w_t\}$, and bounded:

$$|u_t| \leq B_w, \ a.s. \ t \in \mathcal{I}. \quad (10)$$

The reason why we clip inputs at B_w is that during this phase, the algorithm is essentially open-loop. Consequently, it behaves "conservatively" and avoids using inputs of large magnitudes.

Let $N^{(\mathcal{I})}(t)$ denote the number of exploratory steps until t. $\mathcal{I}$ is composed of multiple episodes, each comprising of a set of consecutive time steps. For $i = 1, 2, \ldots,$ the i-th such exploratory episode is $[n_i, n_i+m_i]$. The first episode begins at time $t = 1$, i.e., $n_1 = 1$, and lasts until the following stopping-time,

$$n_1 + m_1 := \max\left\{\inf\{t : b_{1,t} \neq 0\}, H_1(\Theta, \epsilon)\right\}, \quad (11)$$

where $b_{1,t}$ denotes the estimate of b_1 generated at time t, $H_1(\Theta, \epsilon)$ depends upon the model parameter set Θ, and ϵ is a parameter choice that decides the length of the first exploratory phase in the algorithm. It is detailed in the Appendix. Its effect on the regret is shown in Theorem 4.1.

The first exploratory phase serves as a special "warm-up" phase, and is of longer duration than the remaining ones. It arises naturally out of the regret analysis, with sufficient exploration in the first few time-steps allowing us to bound the regret as $C \log T$. For the remaining episodes, $i = 2, 3, \ldots,$

$$n_i = \exp(i^2), \ m_i = H := \left\lceil m^\star + \log_\rho\left(\frac{1}{3C_1 q}\right)\right\rceil, \quad (12)$$

$$m^\star = \frac{\log B_w}{\log \rho} - \log \sup_{\theta\in\Theta} C_1(\theta)\Bigg(\|Y_0\|C_1(\theta) + B_u C_1(\theta)\left[1 + \sum_{\ell=1}^{q}|b_\ell(\theta)|\right]\Bigg),$$

where B_u is as in (20).

Estimates: Let $\theta_t^{(\mathcal{I})}$ be the least-squares estimate (LSE) of $\theta^\star$ based upon only the samples in $\mathcal{I}$,

$$\theta_t^{(\mathcal{I})} := \left(\sum_{s\le t, s\in\mathcal{I}} \phi_s\phi_s'\right)^{-1}\left(\sum_{s\le t, s\in\mathcal{I}} \phi_s y_{s+1}\right), \quad (13)$$

where,

$$V_t^{(\mathcal{I})} := I_{p+q} + \sum_{s\le t, s\in\mathcal{I}} \phi_s\phi_s'. \quad (14)$$

Since the estimate of b_1 in $\theta_t^{(\mathcal{I})}$ might be 0, we modify it slightly as follows so that the resulting estimate can be used for estimating λ:

$$\tilde{\theta}_t^{(\mathcal{I})} := \begin{cases}\theta_t^{(\mathcal{I})} \text{ if } b_{1,t}^{(\mathcal{I})} \neq 0,\\ \tilde{\theta}_{t-1}^{(\mathcal{I})} \text{ otherwise}.\end{cases} \quad (15)$$

Let $\lambda := -\frac{1}{b_1}(a_1, a_2, \ldots, a_p, b_2, \ldots, b_q)'$. Since $\tilde{b}_{1,t}^{(\mathcal{I})} \neq 0$, we use it to estimate λ as follows,

$$\tilde{\lambda}_t^{(\mathcal{I})} := -\frac{1}{\tilde{b}_{1,t}^{(\mathcal{I})}}\left(\tilde{a}_{1,t}^{(\mathcal{I})}, \tilde{a}_{2,t}^{(\mathcal{I})}, \ldots, \tilde{b}_{2,t}^{(\mathcal{I})}, \ldots, \tilde{b}_{q,t}^{(\mathcal{I})}\right). \quad (16)$$

Even though we later show $\tilde{\theta}_t^{(\mathcal{I})}, \tilde{\lambda}_t^{(\mathcal{I})}$ to be consistent, they need not be efficient since they use only a small fraction of the total available samples. Hence, while generating $\{u_t\}$, for most of the time we will directly estimate the parameter λ using all the available samples by the following recursive estimator,

$$\lambda_t = \lambda_{t-1} + P_t\psi_t\left(u_t - \tilde{b}_{1,t-1}^{(\mathcal{I})} y_{t+1} - \lambda_{t-1}'\psi_t\right), \quad (17)$$

where P_t is obtained recursively as

$$P_t^{-1} = P_{t-1}^{-1} + \psi_t\psi_t', \quad (18)$$

and ψ_t is as in (26).

Exploitation: We rewrite the system equation (2) as

$$y_{t+1} = b_1(u_t - \lambda'\psi_t) + w_{t+1}, t = 1, 2, \ldots. \quad (19)$$

The inputs are chosen according to the CE rule, i.e., the algorithm assumes that λ_{t-1} is the true value of the optimal gain which yields minimum variance. More specifically, for times $t \notin \mathcal{I}$, we have,[1]

$$u_t = (-B_u) \vee z_t \wedge (B_u), \quad (20)$$

where,

$$z_t := \begin{cases}\lambda_{t-1}'\psi_t \\ \quad \text{if } \left|\lambda_{t-1}'\psi_t - \tilde{\lambda}_{t-1}^{(\mathcal{I})}\psi_t\right| \le B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}}\|\psi_t\|,\\ \tilde{\lambda}_{t-1}^{(\mathcal{I})}\psi_t \text{ otherwise},\end{cases} \quad (21)$$

where the parameters $B_2, B_u > 0$ are user-specified. Note that $\tilde{\lambda}_t^{(\mathcal{I})}$ is used to provide "diagnostic checks" on λ_{t-1}, i.e., in the event that the inputs prescribed by $\tilde{\lambda}_{t-1}^{(\mathcal{I})}$ and λ_{t-1} differ significantly, the algorithm detects that the input prescribed by λ_{t-1} is "bad" and falls back on the estimate $\tilde{\lambda}_{t-1}^{(\mathcal{I})}$.

Clipping Inputs: Let

$$M(\Theta) := 1 + \sup_{\theta\in\Theta}\frac{C_1(\theta)}{1-\rho(\theta)}\left\{1 + \sum_{\ell=1}^{q}|b_\ell(\theta)|\right\}. \quad (22)$$

The clipping threshold B_u in (20) is given by $B_u = \frac{B_w}{\delta_1^2}(1 + M(\Theta))$, where δ_1 is any constant that satisfies the following inequalities,

$$\delta_1 \le \frac{1}{(p+q)\sup_{\theta\in\Theta}(1 + \|\lambda(\theta)\|)}, \quad (23)$$

$$\left(\sup_{\theta\in\Theta}\|\lambda(\theta)\| + 1\right)[M(\Theta) + q]\,\delta_1 \le 1, \text{ and} \quad (24)$$

$$3\delta_1 \le \inf_{\theta\in\Theta}\left[\frac{b_1(\theta)(1-\rho(\theta))}{C_1(\theta)}\right] \times \left[\frac{\delta_1^2}{2B_w M(\Theta)} + \sup_{\theta\in\Theta}\sum_{\ell=1}^{p}|a_\ell(\theta)|\right]^{-1}. \quad (25)$$

To see why a solution exists, we note that the first two inequalities admit a solution set trivially, while in the third case, the l.h.s. is a monotone increasing function with value 0 for $\delta_1 = 0$, while the r.h.s. is decreasing and has a positive value for $\delta_1 = 0$.

3.1. Discussion of PIECE

A key challenge experienced by many adaptive controllers is their poor empirical performance, especially at the initial stages of learning (Lale et al., 2022; Mete et al., 2022). For improving the transient performance, it is important to properly adapt the system in the initial phase. Otherwise, the states of the system can reach arbitrarily high values before settling down to what is predicted by the asymptotic theory. This is even more exacerbated for the MV-CE control law, which can be written as $u_t = \lambda_t'\psi_t$, where

$$\lambda_t := (-1/b_{1,t})(a_{1,t}, \ldots, a_{p,t}, b_{2,t}, \ldots, b_{q,t})',$$
$$\psi_t := (y_t, \ldots, y_{t-p+1}, u_{t-1}, \ldots, u_{t-q+1})'. \quad (26)$$

[1] We note that a similar control law $u_t = k_l \vee \frac{\hat{a}_t}{\hat{b}_t} \wedge k_u$ was proposed in econometrics (Anderson & Taylor, 1976) for the simple model $y_{t+1} = a + bu_t + w_{t+1}$.

Algorithm 1 PIECE: Probing Inputs for Exploration in Certainty Equivalence

Input The exploration set $\mathcal{I}$, $B_2 > 0$.
if $t \in \mathcal{I}$ **then**
Generate an exploratory white noise input u_t such that $|u_t| \leq B_w$ and has mean 0.
else
Compute the estimates $\tilde{\theta}^{(\mathcal{I})}_{t-1}$, $\tilde{\lambda}^{(\mathcal{I})}_{t-1}$ and λ_t as defined in (15).

$$u_t = \begin{cases} (-B_u) \vee \left(\lambda'_{t-1}\psi_t\right) \wedge (B_u) \\ \quad \text{if } \left|\lambda'_{t-1}\psi_t - \left(\tilde{\lambda}^{(\mathcal{I})}_{t-1}\right)'\psi_t\right| \leq \frac{B_2\|\psi_t\| \log N^{(\mathcal{I})}_t}{\sqrt{N^{(\mathcal{I})}_t}}, \\ (-B_u) \vee \left(\left(\tilde{\lambda}^{(\mathcal{I})}_{t-1}\right)'\psi_t\right) \wedge (B_u) \text{ otherwise .} \end{cases}$$

end if

Since λ_t involves a division by $b_{1,t}$ it is susceptible to large errors even for modest values of estimation error of b_t, which in-turn leads to high regret, especially during the initial time steps. To overcome this shortcoming, PIECE clips the inputs suggested by the CE rule to a compact set $[-B_u, B_u]$. The value of threshold B_u is chosen based upon a fine-grained analysis of learning regret and utilizes knowledge of Θ, a compact set in which the true parameter $\theta^\star$ is known to reside.

To establish finite-time regret bounds, as well as improve initial performance, we have also added exploratory episodes where noise is injected to gather information about $\theta^\star$. This ensures that the estimator is well-behaved without greatly affecting the performance. The values of the clipping threshold and the power of the exploratory noise are designed carefully using the information available about the parameter set. The resulting PIECE algorithm is a simplified and optimized version of (Lai & Wei, 1987).

The PIECE algorithm is computationally efficient since it maintains an estimate of the optimal gain and LSEs of the unknown ARX parameters, both of which can be updated recursively and hence require $O(1)$ computation at each time step.

All of these improvements are visible when one compares the empirical performance of LW with PIECE. PIECE is seen to outperform LW and CE by a huge margin, as shown in Section 5.

4. Finite-Time Regret Analysis

We now state our key results which quantify (i) an upper-bound on regret, and (ii) the estimation error $\|\theta^\star - \theta^{(\mathcal{I})}_t\|$ of the PIECE algorithm. We perform analysis under two separate assumptions on $\{w_t\}$. Section 4.1 considers the case when the noise is uniformly bounded, i.e., $|w_s| \leq B_w$, and shows that the regret of the algorithm is upper-bounded by $C \log T$. This is relaxed in Section 4.2, where the noise is allowed to be unbounded, but conditionally sub-Gaussian (Assumption 2.4). Then PIECE suffers a regret that is at most $C' \log^3 T$. Precise values of constants and bounds are given in the Theorems below and in the Appendix.

4.1. Regret for Bounded Noise

Theorem 4.1. *Consider the ARX system (2) that satisfies Assumption 2.1, and $\{w_t\}$ satisfies Assumptions (2.2,2.3). For every $\delta > 0$, there is a set having probability at least $1 - 6\delta$, such that for every $\epsilon > 0$ the cumulative regret until T can be bounded as follows,*

$$\begin{aligned} \mathcal{R}_T \leq & (1 + c(\epsilon))\sigma^2(p + q - 1)\log T \\ & + L_1(\rho)\sqrt{\log T}, \end{aligned} \tag{27}$$

where $c(\epsilon) := \left(1 - \frac{(1+\epsilon)^2}{(1-\epsilon)^2(1-2\epsilon)}\right)^{-1}\left[\frac{1}{(1-\epsilon)^2(1-2\epsilon)}\right] - 1 \to 0$ as $\epsilon \to 0$, while $L_1(\rho) \to \infty$ as $\rho \nearrow 1$.

Outline of Proof. The instantaneous regret at time t, denoted r_t, can be shown to be equal to $b_1^2(u_{t-1} - \lambda'\psi_{t-1})^2$. We analyze this separately for exploratory time steps, i.e. for $t \in \mathcal{I}$, and for $t \notin \mathcal{I}$.

Regret during $t \in \mathcal{I}$: We bound r_t for $t \in \mathcal{I}$ by u^2_{t-1} plus terms $\leq c'\|\psi_{t-1}\|^2$. We derive an upper-bound on $|y_t|$ that holds uniformly for all times after an initial phase. Since during $\mathcal{I}$, the magnitude of input is bounded by B_w, upon combining this with the bound on $|y_t|$ it yields an upper-bound on $\|\psi_t\|$. This shows that the regret incurred during the exploratory episodes is $\leq c'' N^{(\mathcal{I})}_t$.

Regret during $t \notin \mathcal{I}$: To derive an upper-bound, we relate the instantaneous regret r_t with the "prediction error" $e_t := y_{t+1} - b_{1,t}\left(u_t - \lambda'_{t-1}\psi_t\right)$, where $b_{1,t}\left(u_t - \lambda'_{t-1}\psi_t\right)$ is the prediction of the algorithm about the next observation; if $\theta^\star$ were known, this error would have been w_{t+1}. This observation allows us to show that the instantaneous regret can be bounded by the "mismatch" $(e_t - w_t)^2$, but only when this mismatch is "not too large." Following this, the proof for $t \notin \mathcal{I}$ is split into the following two parts:

1. Ensuring that under the proposed algorithm, the mismatch $(e_t - w_t)^2$ does not become too large. While this cannot be ensured at all times and for all sample paths, we show that under the PIECE algorithm, this does hold with a high probability for "most" of the time steps after a sufficiently large duration . We show that a sufficient condition for this to occur is that under

PIECE, the inputs u_t are not clipped too often, i.e., the condition $|z_t| < B_u$ holds. To ensure this, PIECE (i) uses exploratory episodes of sufficiently large duration H (12), and, (ii) explores using white noise of sufficiently small magnitude ($|u_t| \le B_w,\ t \in \mathcal{I}$). Since the roots of the polynomials are strictly inside the unit circle (Assumption 2.1), when the estimation error $\|\theta^\star - \theta_t\|$ is sufficiently small, the magnitude of the output for times $t \notin \mathcal{I}$ lying between two consecutive episodes can be bounded.

2. Central to the ensuing finite-time performance results of the proposed algorithm is the fine analysis of the growth-rate of the minimum eigenvalue of the covariance matrix associated with LSE. We show that it grows as $A\sqrt{N_t^{(\mathcal{I})}}$, where $N_t^{(\mathcal{I})}$ is the number of exploratory steps until t. The pre-factor A is a function of system parameters, and (i) increases with the value of variance of the process noise, (ii) decreases with a_1, (iii) decreases as the stability margin $1-\rho$ (defined in the sequel) approaches 0.

3. Deriving an upper-bound on the cumulative mismatch $\sum_{t\notin\mathcal{I}}(e_t - w_t)^2$. The analysis relies upon a recursion for the quantity $q_t := \mathrm{Tr}\big(b_1^2(\lambda_t - \lambda)P_t^{-1}(\lambda_t-\lambda)'\big)$. Upon summing this recursion, it can be shown that after sufficiently large t, with a high probability, the mismatch can be controlled by deriving upper-bounds on six terms which mostly involve discrete-time martingale transforms. The rest of the analysis relies upon carefully bounding these terms using concentration results for self-normalized martingales (Peña et al., 2009; Abbasi-Yadkori et al., 2011) and the Azuma-Hoeffding inequality for unbounded martingale difference sequences (Tao & Vu, 2015).

□

We note that by letting $\epsilon \searrow 0$, we are able to match the pre-constant as well as the logarithmic growth rate of the asymptotically optimal regret (1 of (Lai & Wei, 1987). Furthermore, our bounds also quantify the transient performance and how it is affected by various parameters such as δ, B_w, B_u, and the operator norm dependent quantity ρ.

In order to minimize the regret, we need to control the estimation error $\|\theta_t^{(\mathcal{I})} - \theta^\star\|$ associated with LSE. Therefore, we provide the following finite-time guarantees on the performance of the LSE operating under PIECE algorithm in Theorem 4.2.

Theorem 4.2. *Consider the ARX system* (2) *that satisfies Assumption 2.1,* $\{w_t\}$ *satisfies Assumptions (2.2,2.3), and LSE is given by* (13). *On a set having probability greater than* $1-4\delta$*, the estimation error can be bounded as follows,*

$$\|\theta^\star - \theta_t^{(\mathcal{I})}\|^2 \le -\frac{\log(\delta)}{2N_t^{(\mathcal{I})}} + \frac{(p+q)}{2N_t^{(\mathcal{I})}}\times \log\left((C_1\|Y_0\| + \frac{C_1 B_u}{1-\rho}\left[1+\sum_{\ell=1}^{q}|b_\ell|\right] + qB_u)^2 N_t^{(\mathcal{I})}\right).$$

Outline of Proof. $V_t^{(\mathcal{I})} = I + \sum_{s\le t, s\in\mathcal{I}} \phi_s\phi_s'$ is the covariance matrix associated with the LSE at t. It can be shown that w.h.p. $\|\theta^\star - \theta_t^{(\mathcal{I})}\|^2 \le c''\left(\frac{\log\lambda_{\max}(V_t^{(\mathcal{I})})}{\lambda_{\min}(V_t^{(\mathcal{I})})}\right)$, and hence it suffices to upper-bound $\lambda_{\max}(V_t)$ and lower-bound $\lambda_{\min}(V_t^{(\mathcal{I})})$. $\lambda_{\max}(V_t^{(\mathcal{I})})$ is bounded by $\sum_s \|\phi_s\|^2$, which in turn is shown to be $\le c''' t B_u^2$. The key challenge in the proof is to derive a lower-bound on $\lambda_{\min}(V_t^{(\mathcal{I})})$ that holds for all times t greater than some finite time w.h.p. This is done in Appendix F.

□

4.2. Unbounded Noise

To deal with unbounded noise, we slightly modify the PIECE algorithm as follows. Firstly, the exploratory episodes are changed to have $n_i = \exp(i)$, and, in the definition of $m^\star$, B_w is replaced with $\sqrt{\log(T/\delta)}$, with this quantity serving as a high-probability upper-bound on $\{w_t\}$. While deciding the threshold B_u for clipping inputs, once again B_w is replaced by $\sqrt{\log(T/\delta)}$. Let $\tilde{H}$ be the resulting episode duration.

Theorem 4.3. *Consider the ARX system* (2) *that satisfies Assumption 2.1, and* $\{w_t\}$ *satisfies Assumptions (2.2,2.4). The regret of PIECE can be bounded as follows: For every* $\delta > 0$*, there is a set having probability at least* $1-7\delta$ *such that for every* $\epsilon > 0$*, the cumulative regret until* T *can be bounded as follows,*

$$\mathcal{R}_T \le C\log T\log^2(T/\delta) + \tilde{H}\tilde{L}_1(\rho)\log T + (1+c(\epsilon))\times \left[\sigma^2(p+q-1)\log T + \log(T/\delta)\right] + \tilde{L}_2(\epsilon,\delta,\rho), \tag{28}$$

where $c(\epsilon)$ *is as in Theorem 4.1, while* $\tilde{L}_1(\rho) \to \infty$ *as* $\rho \nearrow 1$*, and* $\tilde{L}_2 \to \infty$ *as* $\epsilon \searrow 0$*,* $\rho \nearrow 1$*, or* $\delta \searrow 0$*.*

We note that in comparison with Theorem 4.1, there is an additional $\log^2(T/\delta)$ term that arises due to an increase in the high probability upper-bound on the norms of $\|Y_t\|$ and $\|U_t\|$. This, in-turn happens due to an increase in the magnitudes of noise, exploratory inputs, and inputs during the exploitation phase, as compared with the bounded noise case. It is shown in (Lai & Wei, 1987) that the regret of LW is asymptotically $\sigma^2(p+q-1)\log T$ under the assumption that $\sup_t \mathbb{E}\left\{\exp(\gamma|w_t|)\Big|\mathcal{F}_{t-1}\right\} < \infty$, a.s., for some $\gamma > 0$. It remains to be seen if the finite-time regret of our

Example	CE	LW	PIECE
I	313	364	36
II	16224624	5816	66
III	9888	15644	46

Table 1. Cumulative Regret Performance at $T = 1000$.

proposed algorithm can be improved so that it matches this asymptotically as $T \to \infty$.

5. Simulations

In this section, we compare the performance of the PIECE algorithm with the algorithm proposed in (Lai, 1986) (LW), as well as the standard CE controller. Each simulation experiment is performed for 1000 steps. The reported results are the averaged values over 50 runs. The examples of the ARX systems considered in the experiments are the following:

EXAMPLE I (PAPER MACHINE (ÅSTRÖM & WITTENMARK, 1973)): Linear system with $p = 2$ and $q = 2$: $y_t = 1.283y_{t-1} - 0.495y_{t-2} + 2.307u_{t-1} - 2.025u_{t-2} + w_t$.

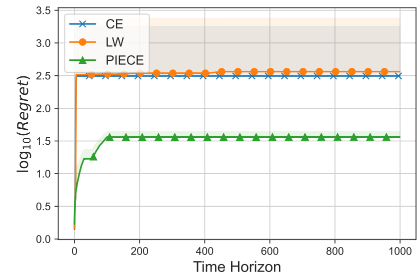
EXAMPLE II: An linear system with $p = 4$ and $q = 4$: $y_t = 1.18y_{t-1} - 0.48y_{t-2} + 0.45y_{t-3} - 0.41y_{t-4} + 0.28u_{t-1} + 0.14u_{t-2} + 0.16u_{t-3} + 0.03u_{t-4} + w_t$,

EXAMPLE III: An linear system with $p = 6$ and $q = 6$: $y_t = -0.66y_{t-1} - 0.79y_{t-2} + 0.2y_{t-3} - 0.03y_{t-4} + 0.09y_{t-6} + 0.32u_{t-1} + 0.06u_{t-2} - 0.2u_{t-3} - 0.01u_{t-4} - 0.03u_{t-5} + 0.001u_{t-6} + w_t$.

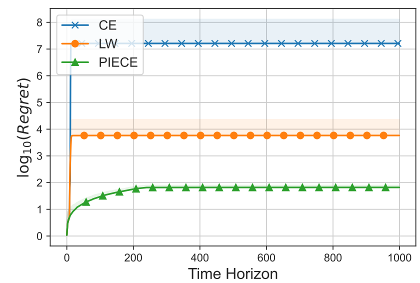
Cumulative Regret Table 1 highlights the cumulative regret at the end of the experiment. In Figure 1, we plot the logarithms of the cumulative regrets, $\log(R_t)$.

One of the key issues with many adaptive controllers is their empirical performance in the initial phase of learning (Lale et al., 2022; Mete et al., 2022). It is evident from the empirical results that CE as well as LW both suffer from this issue. As described in Section 3, the PIECE algorithm differs from LW with regard to the clipping of the input as well as the choice of exploration episodes. The empirical results demonstrate that PIECE does not suffer a large regret at the beginning of the experiments, unlike the LW and the CE controllers. The benefits of the PIECE modifications of clipping, as well as improved exploration strategy, are clearly evident as the resulting algorithm has much lower empirical regret compared to LW or the standard CE controller.

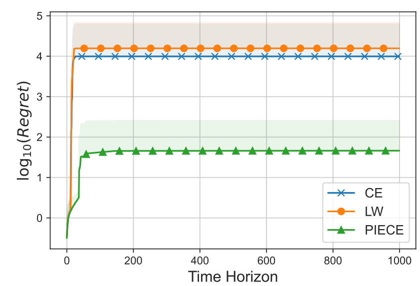
Estimation Error: In Figure 2, we plot the estimation error $||\theta_t - \theta^\star||^2$. It is interesting to note that LW has a better estimation error than PIECE. This reiterates the point that the exploration scheme in PIECE is more efficient in achieving lower regret, which is the primary objective of the controller, at the cost of a higher estimation error.



(a) Example I



(b) Example II



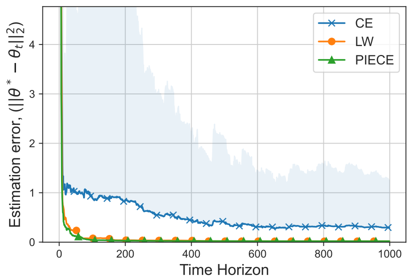
(c) Example III

Figure 1. Log(Cumulative Regret) averaged over 50 runs.

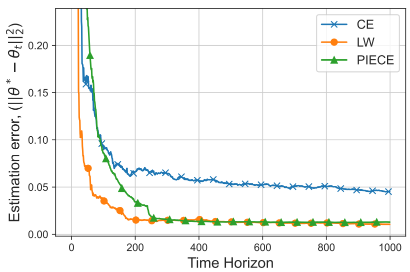
Results for more example, comparison of performance with algorithms LQ learning algorithms like OFU (Abbasi-Yadkori & Szepesvári, 2011), StabL (Lale et al., 2022), hyper-parameter sensitivity analysis, and technical details on implementation are provided in the Appendix L.

6. Concluding Remarks

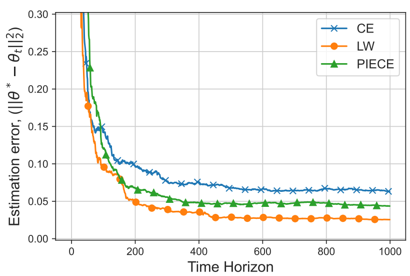
How to obtain a logarithmic, as opposed to polynomial, finite-time regret for the minimum variance controller that is of great value in applications, has been an open problem. To obtain one, it is necessary to re-design the initial learning phase so that it does not have a large transient. We have proposed the PIECE algorithm that employs a modification consisting of clipping the inputs, as well as using



(a) Example I



(b) Example II



(c) Example III

Figure 2. Estimation Error $(||\theta^\star - \theta_t||_2^2)$

probing noise at increasingly infrequent intervals. We have established the first finite-time logarithmic as opposed to polynomial regret bounds for two scenarios: (i) When the system noise is bounded, the regret of the PIECE algorithm is $C \log T$. (ii) When system noise is unbounded, the regret of the PIECE algorithm is $C \log^3 T$. Simulations show the advantage of the PIECE algorithm over LW and the standard certainty equivalence controller.

Whether the regret bound in the unbounded noise case can be improved to $C \log T$ remains an open question. A natural next step is to analyze performance of similar algorithms for an ARMAX system. One can potentially adapt similar algorithms which use probing inputs in other various reinforcement learning settings, including Markov Decision Processes and LQ systems.

Impact Statement

This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

Acknowledgements

This material is based upon work partially supported by the National Science Foundation under Contract Numbers CNS-2328395 and CMMI-2038625, US Army Contracting Command under W911NF-22-1-0151, US ARO under W911NF2120064, US Office of Naval Research under N00014-21-1-2385, and the Science and Engineering Research Board through the grant SRG/2021/002308.

References

Abbasi-Yadkori, Y. and Szepesvári, C. Regret bounds for the adaptive control of linear quadratic systems. In Proceedings of the 24th Annual Conference on Learning Theory, pp. 1–26. JMLR Workshop and Conference Proceedings, 2011.

Abbasi-Yadkori, Y., Pál, D., and Szepesvári, C. Improved algorithms for linear stochastic bandits. Advances in neural information processing systems, 24, 2011.

Abeille, M. and Lazaric, A. Thompson sampling for linear-quadratic control problems. In Artificial Intelligence and Statistics, pp. 1246–1254. PMLR, 2017.

Anderson, T. W. and Taylor, J. B. Some experimental results on the statistical properties of least squares estimates in control problems. Econometrica: Journal of the Econometric Society, pp. 1289–1302, 1976.

Åström, K. J. Computer control of a paper machine—an application of linear stochastic control theory. IBM Journal of research and development, 11(4):389–405, 1967.

Åström, K. J. Papermaking: Adaptive Control. In Systems and Control Encyclopedia: Theory, Technology, Applications, pp. 3590–3591. Pergamon Press Ltd., 1987.

Åström, K. J. Introduction to stochastic control theory. Courier Corporation, 2012.

Åström, K. J. and Wittenmark, B. On self tuning regulators. Automatica, 9(2):185–199, 1973.

Auer, P. and Ortner, R. Logarithmic online regret bounds for undiscounted reinforcement learning. In Advances in neural information processing systems, pp. 49–56, 2007.

Auer, P., Cesa-Bianchi, N., and Fischer, P. Finite-time analysis of the multiarmed bandit problem. Machine learning, 47(2):235–256, 2002.

Bauer, H. Minimalstellen von funktionen und extremalpunkte. Archiv der Mathematik, 9(4):389–393, 1958.

Becker, A., Kumar, P. R., and Wei, C.-Z. Adaptive control with the stochastic approximation algorithm: Geometry and convergence. IEEE Transactions on Automatic Control, 30(4):330–338, 1985.

Borisson, U. and Syding, R. Self-tuning control of an ore crusher. Automatica, 12(1):1–7, 1976.

Borisson, U. and Wittenmark, B. An industrial application of a self-tuning regulator. In 4th IFAC/IFIP International Conference on Digital Computer Applications to Process Control: Part I, pp. 76–87. Springer, Zürich, Switzerland, 1974.

Brink, A. and Tiano, A. Self-tuning adaptive control of large ships in non-stationary conditions. International Shipbuilding Progress, 28(323):162–178, 1981.

Cassel, A., Cohen, A., and Koren, T. Logarithmic regret for learning linear quadratic regulators efficiently. In International Conference on Machine Learning, pp. 1328–1337. PMLR, 2020.

Cegrell, T. and Hedqvist, T. Successful adaptive control of paper machines. Automatica, 11(1):53–59, 1975.

Cohen, A., Koren, T., and Mansour, Y. Learning linear-quadratic regulators efficiently with only $\sqrt{T}$ regret. In International Conference on Machine Learning, pp. 1300–1309. PMLR, 2019.

Faradonbeh, M. K. S., Tewari, A., and Michailidis, G. On adaptive linear-quadratic regulators. Automatica, 117: 108982, 2020.

Fjeld, M. and Wilhelm, R. Self-tuning regulators - The software way. Control Engineering, 28(11):99–102, 1981.

Goodwin, G. C., Ramadge, P. J., and Caines, P. E. Discrete time stochastic adaptive control. SIAM Journal on Control and Optimization, 19(6):829–853, 1981.

Guo, L., Chen, H.-F., et al. The Åström-Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers. IEEE Transactions on Automatic Control, 36(7):802–812, 1991.

Horn, R. A. and Johnson, C. R. Matrix analysis. Cambridge university press, 2012.

Jedra, Y. and Proutiere, A. Minimal expected regret in linear quadratic control. In International Conference on Artificial Intelligence and Statistics, pp. 10234–10321. PMLR, 2022.

Kumar, P. R. and Praly, L. Self-tuning trackers. SIAM journal on control and optimization, 25(4):1053–1071, 1987.

Kumar, P. R. and Varaiya, P. Stochastic systems: Estimation, identification, and adaptive control. Prentice-Hall, Englewood Cliffs, NJ, 1986.

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

Lai, T. L. and Robbins, H. Asymptotically efficient adaptive allocation rules. Advances in applied mathematics, 6(1): 4–22, 1985.

Lai, T. L. and Wei, C. Z. Asymptotic properties of projections with applications to stochastic regression problems. Journal of Multivariate Analysis, 12(3):346–370, 1982a.

Lai, T. L. and Wei, C. Z. Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. The Annals of Statistics, 10(1):154–166, 1982b.

Lai, T. L. and Wei, C.-Z. Asymptotically efficient self-tuning regulators. SIAM Journal on Control and Optimization, 25(2):466–481, 1987.

Lai, T. L., Robbins, H., and Wei, C. Z. Strong consistency of least squares estimates in multiple regression ii. Journal of multivariate analysis, 9(3):343–361, 1979.

Lale, S., Azizzadenesheli, K., Hassibi, B., and Anandkumar, A. Logarithmic regret bound in partially observable linear dynamical systems. Advances in Neural Information Processing Systems, 33:20876–20888, 2020.

Lale, S., Azizzadenesheli, K., Hassibi, B., and Anandkumar, A. Reinforcement learning with fast stabilization in linear dynamical systems. In Camps-Valls, G., Ruiz, F. J. R., and Valera, I. (eds.), Proceedings of The 25th International Conference on Artificial Intelligence and Statistics, volume 151 of Proceedings of Machine Learning Research, pp. 5354–5390. PMLR, 28–30 Mar 2022.

Ljung, L. System identification. In Signal analysis and prediction, pp. 163–173. Springer, 1998.

Mania, H., Tu, S., and Recht, B. Certainty equivalence is efficient for linear quadratic control. Advances in Neural Information Processing Systems, 32, 2019.

Mete, A., Singh, R., and Kumar, P. R. Augmented rbmle-ucb approach for adaptive control of linear quadratic systems. In Advances in Neural Information Processing Systems, 2022.

Peña, V. H., Lai, T. L., and Shao, Q.-M. Self-normalized processes: Limit theory and Statistical Applications. Springer, 2009.

Praly, L., Lin, S.-F., and Kumar, P. R. A robust adaptive minimum variance controller. SIAM journal on control and optimization, 27(2):235–266, 1989.

Shirani Faradonbeh, M. K., Tewari, A., and Michailidis, G. Input perturbations for adaptive control and learning. Automatica, 117:108950, 2020.

Simchowitz, M. and Foster, D. Naive exploration is optimal for online lqr. In International Conference on Machine Learning, pp. 8937–8948. PMLR, 2020.

Tao, T. and Vu, V. Random matrices: Universality of local spectral statistics of non-Hermitian matrices. The Annals of Probability, 43(2):782 – 874, 2015. doi: 10.1214/13-AOP876. URL https://doi.org/10.1214/13-AOP876.

Tsiamis, A., Ziemann, I., Matni, N., and Pappas, G. J. Statistical learning theory for control: A finite-sample perspective. IEEE Control Systems Magazine, 43(6):67–97, 2023.

Vershynin, R. High-dimensional probability: An introduction with applications in data science, volume 47. Cambridge university press, 2018.

Vidyasagar, M. The graph metric for unstable plants and robustness estimates for feedback stability. TAC, AC-29:403–418, 1984.

Ziemann, I. and Sandberg, H. Regret lower bounds for learning linear quadratic gaussian systems. arXiv preprint arXiv:2201.01680, 2022.

Contents

A. Organization of the Appendix

Appendix B provides justification for the regret definition considered by showing its equivalence with the standard regret definition. Appendix C provides the proof of the regret bound presented in Theorem 4.1 for the case of bounded noise. Appendix D relates the estimation error of $\theta^\star$ with the instantaneous regret. Appendix E derives bounds on the estimation error and proves Theorem 4.2. Proof of Theorem 4.2 relies crucially upon the analysis done in Appendix F of $\lambda_{\min}(V_t^{(\mathcal{I})})$, i.e., the minimum eigenvalue of the covariance matrix. The design of PIECE algorithm involves the choice of length of the first episode and the clipping function B_u which is discussed in Appendix G and Appendix H respectively. Appendix I proves Theorem 4.3, i.e., it extends the regret analysis of Theorem 4.1 to the case of sub-Gaussian noise. Finally, the details of simulation setup and additional results are provided in Appendix L.

B. Regret Definition

For the regret analysis for the self-tuning regulators, the regret is defined as follows (Lai, 1986; Lai & Wei, 1987):

$$\mathcal{R}_T(\mathcal{A}) := \sum_{t=1}^{T} (y_t - w_t)^2 \tag{29}$$

Here we show the equivalence between the above definition with the standard definition of regret, widely used in reinforcement learning literature (Lai & Robbins, 1985; Auer et al., 2002; Auer & Ortner, 2007; Abbasi-Yadkori & Szepesvári, 2011).

First, consider the same definition of regret as used in the bandit literature. If μ_a denotes the mean reward of arm a, and a_t denotes the arm pulled at time t by policy π, then regret of π is taken to be Regret$(\pi) := \sum_{t=1}^{T} \max_a \mu_a - \sum_{t=1}^{T} \mathbb{E}^{\pi} \mu_{a_t}$. Similarly, for minimum variance control problem, the cost (negative of reward) incurred at time t is y_t^2. Hence the analogous definition of regret of a policy π is:

$$Regret(\pi) = \sum_{t=1}^{T} \mathbb{E}^{\pi}[y_t^2] - \sum_{t=1}^{T} \min_{\pi'} \mathbb{E}^{\pi'}[y_t^2]. \tag{30}$$

Consider the second term on the right above. We show that when $\{a_\ell\}, \{b_m\}$ are known, then $\min_{\pi'} \mathbb{E}^{\pi'}[y_t^2] = \sigma^2$. To see this, note that

$$y_t = \sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} + w_t,$$

where u_t is measurable w.r.t. $\mathcal{F}_{t-1}$ (sigma-algebra generated by $\{y_s, u_s\}_{s=1}^{t-1}$). Hence we have,

$$\begin{aligned}
\mathbb{E}^{\pi'} y_t^2 &= \mathbb{E}^{\pi'} \left(\sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} + w_t \right)^2 \\
&= \mathbb{E}^{\pi'} \left(\sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} \right)^2 + \mathbb{E}^{\pi'} w_t^2 \\
&+ 2\mathbb{E}^{\pi'} \left[\left\{ \sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} \right\} w_t \right].
\end{aligned} \tag{31}$$

Now, let us focus on the term $\mathbb{E}^{\pi'} [\{\sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m}\} w_t]$. We have,

$$\begin{aligned}
\mathbb{E}^{\pi'} \left[\left\{ \sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} \right\} w_t \right] &= \mathbb{E}^{\pi'} \left(\mathbb{E}^{\pi'} \left[\left\{ \sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} \right\} w_t \middle| \mathcal{F}_{t-1} \right] \right) \\
&= \mathbb{E}^{\pi'} \left[\left\{ \sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} \right\} \mathbb{E}^{\pi'} (w_t | \mathcal{F}_{t-1}) \right] \\
&= \mathbb{E}^{\pi'} \left[\left\{ \sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} \right\} 0 \right] \\
&= 0,
\end{aligned} \tag{32}$$

where in the second-last equality we have used $\mathbb{E}^{\pi'} (w_t | \mathcal{F}_{t-1}) = 0$ a.s. From (31) and (32) we obtain the following,

$$\begin{aligned}
\mathbb{E}^{\pi'} y_t^2 &= \mathbb{E}^{\pi'} \left([\sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m}]^2 \right) + \mathbb{E}^{\pi'} [w_t^2] \\
&\geq \mathbb{E}^{\pi'} w_t^2 \\
&= \sigma^2.
\end{aligned}$$

Note also that the second inequality above becomes an equality only when the policy π' chooses $u_t = -\frac{\sum_{\ell=1}^{p} a_\ell y_{t-\ell}+\sum_{m=2}^{q} b_m u_{t-m}}{b_1}$. Hence we have $\sum_{t=1}^{T} \min_{\pi'} E^{\pi'} y_t^2 = \sigma^2 T$.

Therefore, $\text{Regret}(\pi) = E^{\pi} \sum_{t=1}^{T} [y_t^2] - \sigma^2 T$. Note that this can also be written as $\text{Regret}(\pi) = E^{\pi}[\sum_{t=1}^{T} y_t^2 - \sum_{t=1}^{T} w_t^2]$, since $E \sum_{t=1}^{T} w_t^2 = \sigma^2 T$.

Now, let us begin with the cost criterion used by Lai and Wei (Lai & Wei, 1987) for minimum variance control problem: $\mathbb{E}^{\pi} \sum_{t=1}^{T} (y_t - w_t)^2$. Note that if the system parameters $\{a_\ell\}, \{b_m\}$ were known, then $\min_{\pi'} \mathbb{E}^{\pi'} \left([y_t - w_t]^2\right) = 0$, since one can choose a policy that applies $u_t = -\frac{\sum_{\ell=1}^{p} a_\ell y_{t-\ell}+\sum_{m=2}^{q} b_m u_{t-m}}{b_1}$. Therefore,

$$
\begin{aligned}
\text{Regret}(\pi) &= \mathbb{E}^{\pi} \sum_{t=1}^{T} (y_t - w_t)^2 \\
&= \mathbb{E}^{\pi} \sum_{t=1}^{T} (y_t^2 + w_t^2 - 2 w_t y_t) \\
&= \mathbb{E}^{\pi} \sum_{t=1}^{T} \left(y_t^2 + w_t^2 - 2 w_t \left[\sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} + w_t \right] \right) \\
&= \mathbb{E}^{\pi} \sum_{t=1}^{T} \left(y_t^2 - w_t^2 \right) - 2 \mathbb{E}^{\pi} w_t \left[\sum_{\ell=1}^{p} a_\ell y_{t-\ell} + \sum_{m=1}^{q} b_m u_{t-m} \right] \\
&= \mathbb{E}^{\pi} \sum_{t=1}^{T} \left(y_t^2 - w_t^2 \right) \text{ (since the last term is 0, as shown above)} \\
&= \mathbb{E} \sum_{t=1}^{T} y_t^2 - \sigma^2 T,
\end{aligned}
$$

Therefore, this definition of regret (29) is equivalent to the standard definition of regret (30).

C. Regret Analysis (Proofs)

Let $r_t := (y_t - w_t)^2$ be the instantaneous regret at time t. The regret equation (9) and re-parameterization of ARX model (19) yield,

$$
\begin{aligned}
r_t &= (y_t - w_t)^2 \\
&= b_1^2 (u_{t-1} - \lambda' \psi_{t-1})^2.
\end{aligned} \tag{33}
$$

The behavior of the term $u_t - \lambda' \psi_t$ is different for times $t \notin \mathcal{I}$ and $t \in \mathcal{I}$, and so we study them separately in Section C.1 and Section C.2 respectively. Section C.3 derives results which are used while proving Lemma C.4 of Section C.1. Section C.4 combines the bounds derived in Section C.1 and Section C.2 in order to provide bound on the cumulative regret.

Definition C.1 ($\eta_\delta(\cdot)$)**.** For $\delta > 0$, define the function $\eta_\delta(x)$ as follows,

$$
\eta_\delta(x) := \min \left\{ y \in \mathbb{N} : \log\left(\frac{z}{\delta}\right) \le x \cdot z, \ \forall z \ge y \right\}. \tag{34}
$$

Note that $\lim_{x \to 0} \frac{\eta_\delta(x \cdot \delta)}{\log(1/x)} = 1$.

The relation $\lesssim$ ($\gtrsim$) corresponds to $\le$ ($\ge$) up to a universal multiplicative constant.

C.1. r_t for $t \in \mathcal{I}$

We will discuss only the case of bounded noise, i.e. when $\{w_t\}$ satisfies Assumption 2.3, since the proof under Assumption 2.4 follows through using similar arguments by restricting to the set $\mathcal{G}_w$ defined in (291).

For $t \in \mathcal{I}$, the instantaneous regret (33) can be bounded as follows:

$$\begin{aligned} r_t &= b_1^2 (u_{t-1} - \lambda' \psi_{t-1})^2 \\ &\le 2\, b_1^2 \big(|u_{t-1}|^2 + |\lambda' \psi_{t-1}|^2 \big) \\ &\le 2\, b_1^2 \big(B_w^2 + |\lambda' \psi_{t-1}|^2 \big) \\ &\le 2\, b_1^2 \big(B_w^2 + \|\lambda\|^2 \|\phi_t\|^2 \big) \\ &\le 2\, b_1^2 \big(B_w^2 + \|\lambda\|^2 \big(\|U_t\|^2 + \|Y_t\|^2 \big) \big) \\ &\le 2\, b_1^2 \left(B_w^2 + \|\lambda\|^2 \left(q B_u^2 + \left[C_1 \rho^t \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right]^2 \right) \right) \\ &\le 2\, b_1^2 \left(B_w^2 + \|\lambda\|^2 \left(q B_u^2 + \left[C_1 \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right]^2 \right) \right), \end{aligned} \tag{35}$$

where the second inequality follows since $|u_t| \le B_w$ by the design of the algorithm, the third inequality follows since $\|\psi_t\| \le \|\phi_t\|$, and the fifth inequality follows from Lemma K.2.

(35) yields us the following bound on the cumulative regret incurred during the exploration steps $\mathcal{I}$.

Lemma C.2. *When $\{w_t\}$ satisfies Assumptions (2.2, 2.3), then we have,*

$$\sum_{t \in \mathcal{I}} r_t \le 2\, b_1^2 \left(B_u^2 + \|\lambda\|^2 \left(q B_u^2 + \left[C_1 \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right]^2 \right) \right) N_t^{(\mathcal{I})}, \tag{36}$$

where $N_t^{(\mathcal{I})}$ is the number of exploratory steps until t, and C_1, ρ are as discussed in Section K. The same conclusion holds on the set $\mathcal{G}_w$ where $\{w_t\}$ satisfies Assumption 2.4 instead of Assumption 2.3.

C.2. r_t for $t \notin \mathcal{I}$

On $t \notin \mathcal{I}$, the input u_t is chosen according to the rule (20). This rule can be written equivalently as follows. Define

$$z_t := \begin{cases} \lambda'_{t-1} \psi_t \text{ if } \left| \lambda'_{t-1} \psi_t - \left(\tilde{\lambda}_{t-1}^{(\mathcal{I})} \right)' \psi_t \right| \le B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} \|\psi_t\|, \\ \left(\tilde{\lambda}_{t-1}^{(\mathcal{I})} \right)' \psi_t \text{ otherwise }. \end{cases}, \tag{37}$$

Rule (20) can be written equivalently as follows,

$$u_t = (-B_u) \vee z_t \wedge B_u. \tag{38}$$

The re-parametrization (19) suggests that one can view the quantity $\tilde{b}_{1,t-1}^{(\mathcal{I})} \big(u_t - \lambda'_{t-1} \psi_t \big)$ as the prediction of y_{t+1} based on the information available until t . Hence, define the prediction error at time $t+1$ by

$$\begin{aligned} e_{t+1} :&= y_{t+1} - \tilde{b}_{1,t-1}^{(\mathcal{I})} \big(u_t - \lambda'_{t-1} \psi_t \big) \\ &= b_1 (u_t - \lambda' \psi_t) - \tilde{b}_{1,t-1}^{(\mathcal{I})} \big(u_t - \lambda'_{t-1} \psi_t \big) + w_{t+1}, \end{aligned} \tag{39}$$

where the second equality follows from (19). It is shown in Theorem D.2 that for times $t \ge \max\{t_1^\star(\rho), t_2^\star(\rho), t_3^\star(\rho)\}$, where $t_1^\star(\rho), t_2^\star(\rho), t_3^\star(\rho)$ are as in Definition D.1, the instantaneous regret for $t \notin \mathcal{I}$ can be bounded by the quantity $(e_t - w_t)^2$. Hence, we will now focus on bounding $\sum_{t \ge \max\{t_1^\star(\rho), t_2^\star(\rho), t_3^\star(\rho)\}} (e_t - w_t)^2$. Instead of bounding this summation, we will bound $\sum_{t \ge t^\star} (e_t - w_t)^2$, where $t^\star$ is as in (48), since bounding this expression is simpler.

We begin with some definitions. Define,

$$g_t := (\lambda_t - \lambda) P_t^{-1} (\lambda_t - \lambda)', \tag{40}$$

where P_t^{-1} is obtained recursively as follows,

$$P_t^{-1} = P_{t-1}^{-1} + \psi_t \psi_t'. \tag{41}$$

Also let,

$$q_t := \mathrm{Tr}\big(b_1^2 g_t\big), \tag{42}$$

$$\gamma_t := \frac{b_1}{b_{1,t}}. \tag{43}$$

Let

$$\tau := \inf\left\{ t : \sum_{s=1}^{t} \phi_s \phi_s' \text{ is invertible} \right\}. \tag{44}$$

The following result is essentially (4.12), (4.13) of (Lai, 1986):

Lemma C.3. *For the ARX model (2), we have the following recursion for q_t for times $t \geq \tau$,*

$$\begin{aligned} q_t - q_{t-1} = &- \left[b_1(1-\gamma_t)(u_t - \lambda'\psi_t) - b_1(\lambda_{t-1} - \lambda)'\psi_t\right]^2 \\ &+ \left[b_1(1-\gamma_t)(u_t - \lambda'\psi_t)\right]^2 - 2w_t b_1(\lambda_{t-1}-\lambda)'\psi_t \\ &+ (\psi_t' P_t \psi_t)\left[b_1(1-\gamma_t)(u_t - \lambda'\psi_t) - b_1(\lambda_{t-1}-\lambda)'\psi_t - \gamma_t w_t\right]^2, \; t = 1, 2, \ldots. \end{aligned} \tag{45}$$

Define the times,

$$t_5^\star(\epsilon_1) := \frac{2(C_1\|Y_0\|)^2 + 2\left(\frac{B_u C_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\}\right)^2 + qB_u^2}{\beta_3 \epsilon_1}, \; \epsilon_1 > 0, \tag{46}$$

and,

$$t_6^\star(\epsilon_3, \delta) := \inf\left\{ t \in \mathbb{N} : \mathcal{E}(t;\theta^\star,\delta) \leq b_1\epsilon_3 \right\}, \; \epsilon_3, \delta > 0. \tag{47}$$

where ρ is operator norm of the unknown system as in Section 2. We will occasionally omit dependence of $t_5^\star(\epsilon_1), t_6^\star(\epsilon_3,\delta)$ upon $\epsilon_1, \epsilon_3, \delta$ to ease notation. Let,

$$t^\star := t_1^\star(\rho) \vee t_2^\star(\rho) \vee t_3^\star(\rho) \vee t_5^\star(\epsilon_1) \vee t_6^\star(\epsilon_3,\delta) \vee \tau, \tag{48}$$

where the times $t_1^\star(\rho), t_2^\star(\rho), t_3^\star(\rho)$ are as in Definition D.1. Define,

$$\mathcal{T}_2 := \sum_{t=t^\star}^{T} \left[b_1(1-\gamma_t)(u_t - \lambda'\psi_t)\right]^2, \tag{49}$$

$$\mathcal{T}_3 := \sum_{t=t^\star}^{T} w_t b_1 (\lambda_{t-1} - \lambda)'\psi_t, \tag{50}$$

$$\mathcal{T}_4 := \sum_{s=t^\star}^{T} (\psi_s' P_s \psi_s)\left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s - \gamma_s w_s\right]^2, \tag{51}$$

where β_3 is as in (210). The sets $\mathcal{G}_q, \mathcal{G}_{LSE}, \mathcal{G}_{\text{proj}}, \mathcal{G}_{\mathcal{I}}$ are defined in (60), (144), (250) and (157) respectively. The parameter ϵ_1 here controls the "degree of excitation," i.e. $\lambda_{\min}(V_t)$, during the first exploratory episode. The parameter ϵ_3 controls the estimation error at the end of the first exploratory episode, and δ is the confidence parameter which decides the probability contained in the corresponding event.

Lemma C.4. *Let $\epsilon_1, \epsilon_3 > 0$ be as in (46),(47) respectively. When $\{w_t\}$ satisfies Assumptions (2.2, 2.3), then on the set $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$ we have,*

$$\sum_{s=t^\star}^{t} (w_s - e_s)^2 \leq \frac{1}{(1-\epsilon_3^2)}\left[\frac{(\mathcal{T}_2 + \mathcal{T}_3 + \mathcal{T}_{4,2}) + (1+\epsilon_3)\eta_{\delta/(1+\epsilon_3)^2}(\alpha) + q_{t^\star}}{1 - \epsilon_1 - (1+\epsilon_3)\alpha}\right], \tag{52}$$

for all $\alpha \in (0,1)$, where q_t is as in (42), $\mathcal{T}_{4,2}$ is as in (68), and $\eta_\delta(\cdot)$ is as in Definition C.1. The same conclusion holds on $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_w$ when $\{w_t\}$ satisfies Assumption 2.4 instead of Assumption 2.3.

Proof. The result is obtained by summing the recursions (45), and bounding each of the terms separately. These bounds are derived in Section C.3.

Upon substituting the bounds derived in Section C.3, and summing up the recursions (45) for $t \geq t^\star$, we get the following:

$$\begin{aligned} q_t - q_{t^\star} + \sum_{s=t^\star}^{t} \left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right]^2 \{1 - \psi_s' P_s \psi_s\} \\ \leq \mathcal{T}_2 + \mathcal{T}_3 + \mathcal{T}_{4,2} + \mathcal{T}_{4,3}, \end{aligned} \tag{53}$$

where $\mathcal{T}_{4,2}, \mathcal{T}_{4,3}$ are as in Section C.3 (see (67)). We now derive an upper-bound on $\psi_s' P_s \psi_s$ in order to control the summation in l.h.s. above. Since the duration of the first exploratory episode is greater than $t_5^\star(\epsilon_1)$, we have $\frac{B_u^2}{N_t^{(\mathcal{I})}} \leq \epsilon_1$ for $t > t_5^\star$. For $t > t_5^\star(\epsilon_1)$,

$$\begin{aligned} \psi_t' P_t \psi_t &\leq \frac{\|\psi_t\|^2}{\lambda_{\min}\left(\sum_{s=1}^{t} \psi_s \psi_s'\right)} \\ &\leq \frac{\|\psi_t\|^2}{\lambda_{\min}\left(\sum_{s=1}^{t_5^\star(\epsilon_1)} \psi_s \psi_s'\right)} \\ &\leq \frac{\|\psi_t\|^2}{\beta_3 N_{t_5^\star}^{(\mathcal{I})}} \\ &\leq \frac{\left(C_1\|Y_0\| + \frac{B_u C_1}{1-\rho}\{1 + \sum_{\ell=1}^{q}|b_\ell|\}\right)^2 + qB_u^2}{\beta_3 N_{t_5^\star}^{(\mathcal{I})}} \\ &\leq \frac{2(C_1\|Y_0\|)^2 + 2\left(\frac{B_u C_1}{1-\rho}\{1 + \sum_{\ell=1}^{q}|b_\ell|\}\right)^2 + qB_u^2}{\beta_3 N_{t_5^\star}^{(\mathcal{I})}} \\ &\leq \frac{B_u^2}{N_s^{(\mathcal{I})}} \\ &\leq \epsilon_1, \end{aligned} \tag{54}$$

where the third inequality follows from Theorem F.4, while the fourth follows from Lemma K.2, and β_3 is as in (210). The last inequality follows since the first exploratory episode is of duration greater than $t_5^\star(\epsilon_1)$.

Denote

$$\mathcal{S}_1 := \sum_{s=t^\star}^{t} \left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right]^2. \tag{55}$$

Upon combining (53) with Proposition C.9, we get,

$$\mathcal{S}_1(1-\epsilon_1) \leq \mathcal{T}_2 + \mathcal{T}_3 + \mathcal{T}_{4,2} + (1+\epsilon_3)\big(\alpha \cdot \mathcal{S}_1 + \eta_{\delta/(1+\epsilon_3)^2}(\alpha)\big) + q_{t^\star}, \tag{56}$$

where $\alpha \in (0,1)$. Upon re-arranging we obtain,

$$\mathcal{S}_1 \leq \frac{(\mathcal{T}_2 + \mathcal{T}_3 + \mathcal{T}_{4,2}) + (1+\epsilon_3)\eta_{\delta/(1+\epsilon_3)^2}(\alpha) + q_{t^\star}}{1 - \epsilon_1 - (1+\epsilon_3)\alpha}. \tag{57}$$

We will now relate $\mathcal{S}_1$ to $\sum_{s=t^\star}^{t}(e_s - w_s)^2$.

From Proposition C.12 we have,

$$b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s = \frac{b_1}{b_{1,s}}(w_s - e_s).$$

Let $\epsilon_3 > 0$. For times $s > t_6^\star(\epsilon_3, \delta)$ (47), we have,

$$\big(b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\big)^2 \geq (1-\epsilon_3)^2 (w_s - e_s)^2.$$

Upon performing a summation over s, and substituting the resulting inequality into (57), we obtain,

$$\sum_{s=t^\star}^{t} (w_s - e_s)^2 \leq \frac{1}{(1-\epsilon_3^2)} \left[\frac{(\mathcal{T}_2 + \mathcal{T}_3 + \mathcal{T}_{4,2}) + (1+\epsilon_3)\eta_{\delta/(1+\epsilon_3)^2}(\alpha) + q_{t^\star}}{1 - \epsilon_1 - (1+\epsilon_3)\alpha} \right]. \tag{58}$$

This completes the proof. □

C.3. Bounds on $\mathcal{T}_2, \mathcal{T}_3, \mathcal{T}_4$ used in proof of Lemma C.4

We begin by deriving an upper-bound on the process $\{q_t\}$ (42), on the following set,

$$\mathcal{G}_q := \{\omega : \text{ (1, 2, 3) below hold}\}, \text{ where,} \tag{59}$$

$$\begin{aligned}
&1) \sum_{s=t^\star}^{t} b_1(\lambda_{s-1}-\lambda)'\psi_s w_s \leq \left(\sum_{s=t^\star}^{t} \left\{ b_1(\lambda_{s-1}-\lambda)'\psi_s \right\}^2 \right)^{1/2} \times \log \left[\frac{1}{\delta} \left(\sum_{s=t^\star}^{t} \left\{ b_1(\lambda_{s-1}-\lambda)'\psi_s \right\}^2 \right) \right], \forall t \geq t^\star \\
&2) \sum_{s=\tau+1}^{t} (\psi_s' P_s \psi_s) \left\{ w_s^2 - \mathbb{E}\Big(w_s^2 \Big| \mathcal{F}_{s-1} \Big) \right\} \leq \sqrt{ \sum_{s=1}^{t} (\psi_s' P_s \psi_s)^2 \log \left(\frac{\sum_{s=1}^{t} (\psi_s' P_s \psi_s)^2}{\delta} \right) }, \ \forall t \\
&3) \sum_{s=1}^{t} (\psi_s' P_s \psi_s)(\gamma_s) \left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s \right] w_s \\
&\leq \sqrt{ \sum_{s=1}^{t} (\psi_s' P_s \psi_s)^2 (\gamma_s)^2 \left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s \right]^2 } \\
&\times \sqrt{ \log \left(\frac{\sum_{s=1}^{t} (\psi_s' P_s \psi_s)^2 (\gamma_s)^2 \left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s \right]^2}{\delta} \right) } \Bigg\} \ \forall t.
\end{aligned} \tag{60}$$

Proposition C.5.

$$\mathbb{P}(\mathcal{G}_q \geq 1 - 3\delta). \tag{61}$$

Proof. We will show that the probability with which any of the above conditions 1)-3) is violated can be bounded by δ. For 1), we have $\sum_{s=\tau+1}^{t} \gamma_s b_1 (\lambda_{s-1}-\lambda)'\psi_s w_s = \sum_{s=\tau+1}^{t} \frac{b_1^2}{b_{1,s}} (\lambda_{s-1}-\lambda)'\psi_s w_s$. It then follows from the self-normalized inequality (294), with η_s set equal to w_s and X_s equal to $\frac{b_1^2}{b_{1,s}}(\lambda_{s-1}-\lambda)'\psi_s$, that the probability of violation of 1) is upper-bounded by δ. The probability of violation of 2) is again upper-bounded by δ by using self-normalized martingale concentration (294) since $\{w_s^2 - \mathbb{E}\Big(w_s^2 \Big| \mathcal{F}_{s-1}\Big)\}$ is a martingale difference sequence. Similarly, the probability of violating 3) can also be bounded using (294). □

The sets $\mathcal{G}_{LSE}, \mathcal{G}_{\text{proj}}, \mathcal{G}_{\mathcal{I}}, \mathcal{G}_q$ are defined in (144), (250), (157) and (60) respectively.

Proposition C.6 (Bounding $\mathcal{T}_2$). *On the set $\mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$ (144), $\mathcal{T}_2$ (49) can be bounded as follows,*

$$\mathcal{T}_2 \leq \epsilon_3^2 \sum_{t \geq t_6^\star} r_t. \tag{62}$$

Proof. Recall $\mathcal{T}_2 = \sum_{t=t^\star}^{T} [b_1(1-\gamma_t)(u_t - \lambda'\psi_t)]^2$. Since $\gamma_s = \frac{b_1}{b_{1,s}}$ for $t \geq t_6^\star(\epsilon_3, \delta)$, we have $(b_1(1-\gamma_s)(u_s - \lambda'\psi_s))^2 \leq \epsilon^2 (b_1(u_s - \lambda'\psi_s))^2$. The proof is completed by noting that $r_t = b_1^2 (u_t - \lambda'\psi_t)^2$. □

Proposition C.7 (Bounding $\mathcal{T}_3$). *Under Assumption 2.3, on $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$, we have,*

$$\mathcal{T}_3 \leq \alpha \left(2\mathcal{T}_2 + 2(1+\epsilon_3)^2 \sum_{s=t^\star}^{T} (e_s - \epsilon_s)^2 \right) + \eta_\delta(\alpha), \tag{63}$$

for all $\alpha \in (0,1)$. Same conclusion holds under Assumption 2.4 on $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_w$.

Proof. Recall that $\mathcal{T}_3 = \sum_{t=t^\star}^{T} w_t b_1 (\lambda_{t-1} - \lambda)' \psi_t$. On $\mathcal{G}_q$,

$$\mathcal{T}_3 \leq \left(\sum_{s=1}^{t} \left\{ b_1 (\lambda_{s-1} - \lambda)' \psi_s \right\}^2 \right)^{1/2} \log \left[\frac{1}{\delta} \left(\sum_{s=1}^{t} \left\{ b_1 (\lambda_{s-1} - \lambda)' \psi_s \right\}^2 \right) \right]. \tag{64}$$

Hence, to bound $\mathcal{T}_3$, we will focus on bounding $\sum_{s=1}^{t} \left\{ b_1 (\lambda_{s-1} - \lambda)' \psi_s \right\}^2$. We have,

$$b_1 (\lambda_{s-1} - \lambda)' \psi_s = b_1 (1 - \gamma_s)(u_s - \lambda' \psi_s) + \gamma_s (e_s - \epsilon_s), \tag{65}$$

so that

$$\left\{ b_1 (\lambda_{s-1} - \lambda)' \psi_s \right\}^2 \leq 2 \left[b_1 (1 - \gamma_s)(u_s - \lambda' \psi_s) \right]^2 + 2\gamma_s^2 (e_s - \epsilon_s)^2. \tag{66}$$

This gives,

$$\begin{aligned}
\sum_{s=t^\star}^{T} \left\{ b_1 (\lambda_{s-1} - \lambda)' \psi_s \right\}^2 &\leq 2\mathcal{T}_2 + 2 \sum_{s=t^\star}^{T} \gamma_s^2 (e_s - \epsilon_s)^2 \\
&\leq 2\mathcal{T}_2 + 2(1+\epsilon_3)^2 \sum_{s=t^\star}^{T} (e_s - \epsilon_s)^2,
\end{aligned}$$

where the second inequality follows since we have $\gamma_s \leq 1 + \epsilon_3$ (follows from the definition of $t_6^\star$ (47)). The proof then follows by combining the above inequality with (64), and using the definition of the function $\eta_\delta(\cdot)$. □

We will now derive an upper-bound on $\mathcal{T}_4$. We have

$$\begin{aligned}
\mathcal{T}_4 &= \sum_{s=t^\star}^{t} (\psi_s' P_s \psi_s) \left[b_1 (1 - \gamma_s)(u_s - \lambda' \psi_s) - b_1 (\lambda_{s-1} - \lambda)' \psi_s - \gamma_s w_s \right]^2 \\
&= \sum_{s=t^\star}^{t} (\psi_s' P_s \psi_s) \left[b_1 (1 - \gamma_s)(u_s - \lambda' \psi_s) - b_1 (\lambda_{s-1} - \lambda)' \psi_s \right]^2 \\
&+ \sum_{s=t^\star}^{t} (\psi_s' P_s \psi_s)(\gamma_s w_s)^2 \\
&- 2 \sum_{s=t^\star}^{t} (\psi_s' P_s \psi_s)(\gamma_s w_s) \left[b_1 (1 - \gamma_s)(u_s - \lambda' \psi_s) - b_1 (\lambda_{s-1} - \lambda)' \psi_s \right].
\end{aligned} \tag{67}$$

The term $\mathcal{T}_4$ is therefore composed of three summation terms, which we denote by $\mathcal{T}_{4,1}, \mathcal{T}_{4,2}, \mathcal{T}_{4,3}$ respectively, i.e.,

$$\begin{aligned}
\mathcal{T}_{4,1} &:= \sum_{s=t^\star}^{t} (\psi_s' P_s \psi_s) \left[b_1 (1 - \gamma_s)(u_s - \lambda' \psi_s) - b_1 (\lambda_{s-1} - \lambda)' \psi_s \right]^2, \\
\mathcal{T}_{4,2} &:= \sum_{s=t^\star}^{t} (\psi_s' P_s \psi_s)(\gamma_s w_s)^2,
\end{aligned} \tag{68}$$

and,

$$\mathcal{T}_{4,3} := \sum_{s=t^\star}^{t} (\psi_s' P_s \psi_s)(\gamma_s w_s)\left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right]. \tag{69}$$

Thus,

$$\mathcal{T}_4 = \mathcal{T}_{4,1} + \mathcal{T}_{4,2} - 2\mathcal{T}_{4,3}. \tag{70}$$

Next, we analyze these terms below separately.

Proposition C.8 ($\mathcal{T}_{4,2}$)**.** *On the set $\mathcal{G}_q$ we have,*

$$\begin{aligned}\frac{\mathcal{T}_{4,2}}{(1+\epsilon_3)^2} &\le \sigma^2(p+q-1)\log T + \sigma^2(p+q-1)\log\left(C_1^2\left(\|\psi_0\| + \frac{B_u C_1}{1-\rho}\left\{1+\sum_{\ell=1}^{q}|b_\ell|\right\}\right)^2\right)\\ &+ \sqrt{(p+q-1)\log\left[T\, C_1^2\left(\|\psi_0\| + \frac{B_u C_1}{1-\rho}\left\{1+\sum_{\ell=1}^{q}|b_\ell|\right\}\right)^2\right]}\\ &\times\sqrt{\log\left(\frac{(p+q-1)\log\left[T\, C_1^2\left(\|\psi_0\| + \frac{B_u C_1}{1-\rho}\left\{1+\sum_{\ell=1}^{q}|b_\ell|\right\}\right)^2\right]}{\delta}\right)}.\end{aligned} \tag{71}$$

Proof. Consider

$$\mathcal{T}_{4,2} = \sum_{s=t^\star}^{t}(\psi_s' P_s\psi_s)(\gamma_s w_s)^2 \le \left(\sup_{s\ge t^\star}\gamma_s\right)^2\sum_{s=1}^{t}(\psi_s' P_s\psi_s)\, w_s^2 \le (1+\epsilon_3)^2\sum_{s=1}^{t}(\psi_s' P_s\psi_s)\, w_s^2, \tag{72}$$

where the last inequality follows from the definition of $t_6^\star$. Now,

$$\sum_{s=t^\star}^{t}(\psi_s' P_s\psi_s)\, w_s^2 = \sum_{s=t^\star}^{t}(\psi_s' P_s\psi_s)\,\mathbb{E}\left(w_s^2\middle|\mathcal{F}_{s-1}\right) + \left[\sum_{s=\tau+1}^{t}(\psi_s' P_s\psi_s)\,\left\{w_s^2 - \mathbb{E}\left(w_s^2\middle|\mathcal{F}_{s-1}\right)\right\}\right].$$

The first summation above is bounded as follows,

$$\sum_{s=t^\star}^{t}(\psi_s' P_s\psi_s)\,\mathbb{E}\left(w_s^2\middle|\mathcal{F}_{s-1}\right) \le \sigma^2\sum_{s=\tau+1}^{t}(\psi_s' P_s\psi_s) \le \sigma^2\log\left[\det\left(\sum_{s=1}^{t}\psi_s\psi_s'\right)\right].$$

For second summation,

$$\begin{aligned}\sum_{s=\tau+1}^{t}(\psi_s' P_s\psi_s)\,\left\{w_s^2 - \mathbb{E}\left(w_s^2\middle|\mathcal{F}_{s-1}\right)\right\} &\le \sqrt{\sum_{s=1}^{t}(\psi_s' P_s\psi_s)^2\log\left(\frac{\sum_{s=1}^{t}(\psi_s' P_s\psi_s)^2}{\delta}\right)}\\ &\le \sqrt{\sum_{s=1}^{t}(\psi_s' P_s\psi_s)\log\left(\frac{\sum_{s=1}^{t}(\psi_s' P_s\psi_s)}{\delta}\right)}\\ &\le \sqrt{\log\left[\det\left(\sum_{s=1}^{t}\psi_s\psi_s'\right)\right]\log\left(\frac{\log\left[\det\left(\sum_{s=1}^{t}\psi_s\psi_s'\right)\right]}{\delta}\right)},\end{aligned} \tag{73}$$

where the first inequality follows from the definition of $\mathcal{G}_q$, the second inequality follows since $\psi_s' P_s\psi_s \le 1$, while the third inequality follows since $\sum_{s=1}^{t}\psi_s' P_s\psi_s \le \log\left[\det\left(\sum_{s=1}^{t}\psi_s\psi_s'\right)\right]$.

In summary, we obtain the following,

$$\frac{\mathcal{T}_{4,2}}{(1+\epsilon_3)^2} \leq \sigma^2 \log\left[\det\left(\sum_{s=1}^{t}\psi_s\psi_s'\right)\right] + \sqrt{\log\left[\det\left(\sum_{s=1}^{t}\psi_s\psi_s'\right)\right]\log\left(\frac{\log\left[\det\left(\sum_{s=1}^{t}\psi_s\psi_s'\right)\right]}{\delta}\right)}. \tag{74}$$

We now bound $\log\left[\det\left(\sum_{s=1}^{t}\psi_s\psi_s'\right)\right]$. For a matrix $M \in \mathbb{R}^{(p+q-1)\times(p+q-1)}$, we have $\log\det(M)$ is the sum of logarithm of its eigenvalues, and hence can be upper-bounded by $(p+q-1)$ times the log of the maximum eigenvalue. The eigenvalues of the matrix $\sum_{s=1}^{t}\psi_s\psi_s'$ are bounded by $\sum_{s=1}^{t}\|\psi_s\|^2$, which can be bounded using Lemma K.2 as,

$$\sum_{s=1}^{t}\|\psi_s\|^2 \leq t\left\{\left(\|Y_0\| + \frac{B_u C_1}{1-\rho}\left\{1+\sum_{\ell=1}^{q}|b_\ell|\right\}\right)^2 + (B_u q)^2\right\}. \tag{75}$$

Substituting this into (74), we obtain the following bound on $\mathcal{T}_{4,2}$:

$$\begin{aligned}\frac{\mathcal{T}_{4,2}}{(1+\epsilon_3)^2} &\leq \sigma^2(p+q-1)\log T + \sigma^2(p+q-1)\log\left(C_1^2\left(\|\psi_0\| + \frac{B_u C_1}{1-\rho}\left\{1+\sum_{\ell=1}^{q}|b_\ell|\right\}\right)^2\right)\\ &+ \sqrt{(p+q-1)\log\left[T\, C_1^2\left(\|\psi_0\| + \frac{B_u C_1}{1-\rho}\left\{1+\sum_{\ell=1}^{q}|b_\ell|\right\}\right)^2\right]}\\ &\times\sqrt{\log\left(\frac{(p+q-1)\log\left[T\, C_1^2\left(\|\psi_0\| + \frac{B_u C_1}{1-\rho}\left\{1+\sum_{\ell=1}^{q}|b_\ell|\right\}\right)^2\right]}{\delta}\right)}.\end{aligned} \tag{76}$$

□

We now bound the term $\mathcal{T}_{4,3}$.

Proposition C.9 ($\mathcal{T}_{4,3}$)**.** *On the set $\mathcal{G}_q$ (60) we have,*

$$\begin{aligned}\mathcal{T}_{4,3} &\leq (1+\epsilon_3)\sqrt{\sum_{s=1}^{t}\left(b_1(1-\gamma_s)(u_s-\lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right)^2}\\ &\times\sqrt{\log\left(\frac{(1+\epsilon_3)^2}{\delta}\right)\left(\sum_{s=1}^{t}\left(b_1(1-\gamma_s)(u_s-\lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right)^2\right)}.\end{aligned} \tag{77}$$

Proof. We have,

$$\begin{aligned}
\mathcal{T}_{4,3} &= \sum_{s=t^\star}^{T} (\psi_s' P_s \psi_s)(\gamma_s)\left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right] w_s \\
&\le \sqrt{\sum_{s=t^\star}^{T} (\psi_s' P_s \psi_s)^2 (\gamma_s)^2 \left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right]^2} \\
&\times \sqrt{\log\left(\frac{\sum_{s=1}^{t} (\psi_s' P_s \psi_s)^2 (\gamma_s)^2 \left[b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right]^2}{\delta}\right)} \\
&\le \left(\sup_{s\ge t^\star} \left|\gamma_s\right|\right)\sqrt{\sum_{s=1}^{t}\left(b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right)^2} \\
&\times \sqrt{\log\left[\left(\frac{\sup_{s\ge t^\star}\gamma_s^2}{\delta}\right)\left(\sum_{s=1}^{t}\left(b_1(1-\gamma_s)(u_s - \lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s\right)^2\right)\right]},
\end{aligned} \tag{78}$$

where the first inequality follows from the definition of $\mathcal{G}_q$, while the last inequality follows since $\psi_s' P_s \psi_s \le 1$. Proof is completed by noting that from definition of $t_6^\star$ we have $\sup_{s\ge t^\star} \gamma_s \le 1 + \epsilon_3$. □

We now derive a bound on $q_{t^\star}$.

Lemma C.10. *On $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$ we have,*

$$q_{t^\star} \le b_1^2 \epsilon_3^2 t^\star \left(\left\{C_1\|Y_0\| + \frac{B_w C_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q}|b_\ell|\right\}\right\}^2 + B_w^2\right), \tag{79}$$

where q_t is as in (42), and $t^\star$ as in (48).

Proof. We have $q_t = \mathrm{Tr}\big(b_1^2(\lambda_t - \lambda)P_t^{-1}(\lambda_t-\lambda)'\big)$. Since the trace of a matrix is equal to the sum of its eigenvalues, by using the bound on estimation error derived in Theorem (E.5), we get,

$$\begin{aligned}
q_{t^\star} &\le \mathcal{E}(N_{t^\star}^{(\mathcal{I})};\theta^\star,\delta)^2 \sum_{s\le t}\|\phi_s\|^2 \\
&\le \mathcal{E}(N_{t^\star}^{(\mathcal{I})};\theta^\star,\delta)^2 \cdot t^\star \cdot \left(\left\{C_1\|Y_0\| + \frac{B_w C_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q}|b_\ell|\right\}\right\}^2 + B_w^2\right) && (80)\\
&\le b_1^2 \cdot \epsilon_3^2 \cdot t^\star \left(\left\{C_1\|Y_0\| + \frac{B_w C_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q}|b_\ell|\right\}\right\}^2 + B_w^2\right), && (81)
\end{aligned}$$

where the function $\mathcal{E}(x;\theta,\delta)$ is defined in (152) where the second inequality follows from the definition of $t_6^\star$, and Lemma K.1. □

C.4. Bounding the Cumulative Regret $\mathcal{R}_T$: Proof of Theorem 4.1

The "good sets" $\mathcal{G}_q, \mathcal{G}_{LSE}, \mathcal{G}_{\text{proj}}, \mathcal{G}_{\mathcal{I}}, \mathcal{G}_w, \mathcal{G}_{w_B^2}$ are defined in (60) (144), (250), (157), (291), (270) respectively, and our analysis is performed on the intersection of these. Theorem 4.1 derives bound on the cumulative regret $\mathcal{R}_T$. In order to prove this result, we will instead show the following stronger result. Theorem 4.1 then follows directly from Theorem C.11.

Theorem C.11. *Consider the ARX system (2) in which $\{w_t\}$ satisfies Assumptions (2.2,2.3). On the set $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap$*

$\mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_{w_B^2}$, the cumulative regret of the PIECE algorithm until T is bounded as follows:

$$\begin{aligned}
\mathcal{R}_T \leq\ & D \cdot \sigma^2 (p+q-1) \log T \\
& + 2D' \left(\left(C_1 \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right)^2 + B_w^2 \right) \\
& + 2\, b_1^2 \left(B_u^2 + \|\lambda\|^2 \left[C_1 \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} + B_u q \right]^2 \right) N_T^{(\mathcal{I})} \\
& + D \cdot (1+\epsilon_3) \eta_{\delta/(1+\epsilon_3)^2}(\alpha) \\
& + D \cdot b_1^2 \epsilon_3^2 \cdot D' \left(\left\{ C_1 \|Y_0\| + \frac{B_w C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right\}^2 + B_w^2 \right) \\
& + D \cdot \sigma^2 (p+q-1) \log \left(C_1^2 \left(\|\psi_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right)^2 \right) \\
& + D \cdot \sqrt{ (p+q-1) \log \left[T\, C_1^2 \left(\|\psi_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right)^2 \right] } \\
& \times \sqrt{ \log \left(\frac{ (p+q-1) \log \left[T\, C_1^2 \left(\|\psi_0\| + \frac{B_u C_1}{1-\rho} \{ 1 + \sum_{\ell=1}^{q} |b_\ell| \} \right)^2 \right] }{\delta} \right) },
\end{aligned} \tag{82}$$

where,

$$D = \frac{1}{(1-\epsilon_3^2)} \left[\frac{1}{1-\epsilon_1-(1+\epsilon_3)\alpha} \right] \times \left[1 - \frac{\epsilon_3^2(1+2\alpha) + 2\alpha(1+\epsilon_3)^2}{(1-\epsilon_3^2)(1-\epsilon_1-(1+\epsilon_3)\alpha)} \right]^{-1}, \tag{83}$$

ϵ_1, ϵ_3 are as in (46), (47), $\eta_\delta(\cdot)$ is as in Definition C.1, and D' satisfies $D' \lesssim 1/\epsilon_1, D' \lesssim 1/\epsilon_3$.

Proof. Since the cumulative regret during $\mathcal{I}$ has already been bounded in Lemma C.2, we begin by deriving a bound on the cumulative regret during times $t \notin \mathcal{I}$. From Theorem D.2, for times $t \notin \mathcal{I}$ and greater than $\max\{t_1^\star, t_2^\star, t_3^\star\}$, r_t can be bounded by $(w_t - e_t)^2$. Thus, we decompose the regret $\sum_{t \notin \mathcal{I}} r_t$ into the following two parts,

$$\sum_{t \notin \mathcal{I}} r_t = \sum_{t \notin \mathcal{I}, t \leq t^\star} r_t + \sum_{t \notin \mathcal{I}, t \geq t^\star} r_t, \tag{84}$$

where $t^\star$ is as in (48). The first summation is bounded as follows,

$$\begin{aligned}
\sum_{t \notin \mathcal{I}, t \leq t^\star} r_t &= \sum_{t \notin \mathcal{I}, t \leq t^\star} |y_t - w_t|^2 \\
&\leq t^\star \sup_t |y_t - w_t|^2 \\
&\leq 2t^\star \left(\left(C_1 \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right)^2 + B_w^2 \right),
\end{aligned} \tag{85}$$

where the second inequality follows from Lemma K.2. The second summation $\sum_{t \notin \mathcal{I}, t \geq t^\star} r_t$ in (84) is bounded as follows using Lemma C.4,

$$\sum_{t \notin \mathcal{I}, t \geq t^\star} (w_t - e_t)^2 \leq \frac{1}{(1-\epsilon_3^2)} \left[\frac{(\mathcal{T}_2 + \mathcal{T}_3 + \mathcal{T}_{4,2}) + (1+\epsilon_3)\eta_{\delta/(1+\epsilon_3)^2}(\alpha) + q_{t^\star}}{1-\epsilon_1-(1+\epsilon_3)\alpha} \right]. \tag{86}$$

Consider $\mathcal{T}_2$ in the bound above. The term involved in $\mathcal{T}_2$ at time t is equal to $\left(b_1(1-\gamma_s)(u_s-\lambda'\psi_s)\right)^2$, while $r_t = b_1^2(u_t-\lambda'\psi_t)^2$. Now since $\gamma_s = \frac{b_1}{b_{1,s}}$, for $t \geq t_6^\star(\epsilon_3,\delta)$ (47) we have $\left(b_1(1-\gamma_t)(u_t-\lambda'\psi_t)\right)^2 \leq \epsilon_3^2 r_t$. Thus, we obtain,

$$\mathcal{T}_2 \leq \epsilon_3^2 \sum_{t=t^\star}^{T} r_t. \tag{87}$$

Now, Proposition C.7 yields a bound on $\mathcal{T}_3$ in terms of $\mathcal{T}_2$,

$$\mathcal{T}_3 \leq \alpha\left(2\mathcal{T}_2 + 2(1+\epsilon_3)^2 \sum_{s=t^\star}^{T}(e_s-\epsilon_s)^2\right) + \eta(\alpha). \tag{88}$$

Upon combining the above inequalities, and using Theorem D.2 to relate instantaneous regret with $(e_t-w_t)^2$, we get,

$$\sum_{t\notin\mathcal{I},t\geq t^\star}(w_t-e_t)^2\left[1-\frac{\epsilon_3^2(1+2\alpha)+2\alpha(1+\epsilon_3)^2}{(1-\epsilon_3^2)(1-\epsilon_1-(1+\epsilon_3)\alpha)}\right] \leq \frac{1}{(1-\epsilon_3^2)}\left[\frac{(\mathcal{T}_{4,2})+(1+\epsilon_3)\eta_{\delta/(1+\epsilon_3)^2}(\alpha)+q_{t^\star}}{1-\epsilon_1-(1+\epsilon_3)\alpha}\right]. \tag{89}$$

From Theorem D.2, this also yields a bound on the corresponding regret, i.e.,

$$\sum_{t\notin\mathcal{I},t\geq t^\star} r_t \leq \frac{1}{(1-\epsilon_3^2)}\left[\frac{\mathcal{T}_{4,2}+(1+\epsilon_3)\eta_{\delta/(1+\epsilon_3)^2}(\alpha)+q_{t^\star}}{1-\epsilon_1-(1+\epsilon_3)\alpha}\right]\left[1-\frac{\epsilon_3^2(1+2\alpha)+2\alpha(1+\epsilon_3)^2}{(1-\epsilon_3^2)(1-\epsilon_1-(1+\epsilon_3)\alpha)}\right]^{-1}. \tag{90}$$

Proof is then completed by substituting the bound on $q_{t^\star}$ derived in Lemma C.10, bound on the regret during times $t\in\mathcal{I}$ from Lemma C.2, and bound on $\mathcal{T}_{4,2}$ from Proposition C.8 and also bounding $t^\star$.

□

C.5. Some Auxiliary Results

Proposition C.12.

$$b_1(1-\gamma_s)(u_s-\lambda'\psi_s) - b_1(\lambda_{s-1}-\lambda)'\psi_s = \frac{b_1}{b_{1,t}}(w_t-e_t). \tag{91}$$

Proof. We have,

$$y_t = b_1(u_{t-1}-\lambda'\psi_{t-1}) + w_t, \tag{92}$$

and also,

$$e_t = y_t - b_{1,t}\left(u_{t-1}-\lambda_{t-1}'\psi_{t-1}\right). \tag{93}$$

Hence,

$$\begin{aligned}
-\frac{e_t}{b_{1,t}} &= u_{t-1} - \frac{y_t}{b_{1,t}} - \lambda_{t-1}'\psi_{t-1} \\
&= u_{t-1} - \frac{b_1(u_{t-1}-\lambda'\psi_{t-1})+w_t}{b_{1,t}} - \lambda_{t-1}'\psi_{t-1} \\
&= (1-\gamma_t)(u_{t-1}-\lambda'\psi_{t-1}) - (u_{t-1}-\lambda'\psi_{t-1}) + u_{t-1} - \lambda_{t-1}'\psi_{t-1} - \frac{w_t}{b_{1,t}} \\
&= (1-\gamma_t)(u_{t-1}-\lambda'\psi_{t-1}) - (\lambda_{t-1}-\lambda)'\psi_{t-1} - \frac{w_t}{b_{1,t}},
\end{aligned} \tag{94}$$

where the second equality follows from (92). The proof is completed by re-arranging the terms. □

D. Relation between Prediction Error and Regret for $t \notin \mathcal{I}$

The sets $\mathcal{G}_q, \mathcal{G}_{LSE}, \mathcal{G}_{\text{proj}}, \mathcal{G}_{\mathcal{I}}, \mathcal{G}_w, \mathcal{G}_{w_B^2}$ are defined in (60) (144), (250), (157), (291), (270) respectively.

Definition D.1. For $\rho \in [0,1)$ define,

$$t_1^\star(\rho) := \inf \left\{ t \in \mathbb{N} : B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(t;\theta^\star,\delta) < 1, C_1 \rho^t \|Y_0\| < B_u \right\}, \tag{95}$$

and,

$$\begin{aligned} & t_2^\star(\rho) := \inf \Big\{ t \in \mathbb{N} : \\ & \left[B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(t;\theta^\star,\delta) \right] \cdot \left[b_1 \left(p + 1 + \frac{C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right) \right] \le \frac{\delta_1^2}{2}, \\ & C_1 \rho^t \|Y_0\| < \delta_1 B_u \Big\}, \end{aligned} \tag{96}$$

where the function $\mathcal{E}(x;\theta,\delta)$ is as in (152). For $\delta > 0$ define,

$$t_3^\star(\delta) := \inf \left\{ t \in \mathbb{N} : \mathcal{E}(t;\theta^\star,\delta) \le \frac{b_1}{2}, \text{ and } \frac{B_2 \sqrt{\log N_\ell^{(\mathcal{I})}}}{2} \ge 2 \right\}. \tag{97}$$

Theorem D.2. *Under Assumptions (2.2, 2.3) on $\{w_t\}$, the following holds on the set $\mathcal{G}_q \cap \mathcal{G}_{LSE}$: For times $t \ge \max\{t_1^\star(\rho), t_2^\star(\rho), t_3^\star(\delta)\}$, the instantaneous regret r_t can be bounded by $(e_t - w_t)^2$.*

Proof. From (33), the instantaneous regret is given by

$$b_1^2 (u_{t-1} - \lambda' \psi_{t-1})^2. \tag{98}$$

Recall that from (39), we have

$$e_{t+1} = b_1(u_t - \lambda'\psi_t) - \tilde{b}_{1,t-1}^{(\mathcal{I})} \left(u_t - \lambda_{t-1}' \psi_t \right) + w_{t+1}, \tag{99}$$

so that

$$e_{t+1} - w_{t+1} = b_1(u_t - \lambda'\psi_t) - \tilde{b}_{1,t-1}^{(\mathcal{I})} \left(u_t - \lambda_{t-1}' \psi_t \right). \tag{100}$$

Note that from Lemma D.7 we have that $u_t = z_t$, where z_t is as in (37). Hence, we consider the following two cases, and show that in both these cases, the expression (98) can be bounded by $(e_t - w_t)^2$.

Case-I: $u_t = \lambda_{t-1}' \psi_t$.

In this case,

$$\begin{aligned} e_{t+1} - w_{t+1} &= b_1(u_t - \lambda'\psi_t) - \tilde{b}_{1,t-1}^{(\mathcal{I})} \left(u_t - \lambda_{t-1}' \psi_t \right) \\ &= b_1(u_t - \lambda'\psi_t). \end{aligned} \tag{101}$$

Thus, in this case, $(e_{t+1} - w_{t+1})^2$ is equal to the instantaneous regret $b_1^2 (u_t - \lambda'\psi_t)^2$.

Case-II: $u_t = \left(\tilde{\lambda}_{t-1}^{(\mathcal{I})} \right)' \psi_t$.

We have,

$$\begin{aligned} |u_t - \lambda_{t-1}' \psi_t| &= | \left(\tilde{\lambda}_{t-1}^{(\mathcal{I})} \right)' \psi_t - \lambda_{t-1}' \psi_t | \\ &\ge B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} \|\psi_t\|, \end{aligned} \tag{102}$$

where the inequality follows from (21). Moreover, from guarantees on the estimation error provided in Theorem E.5 on $\mathcal{G}_{LSE}$, we have,

$$\|\lambda - \tilde{\lambda}^{(\mathcal{I})}_{t-1}\| \leq \mathcal{E}(N^{(\mathcal{I})}_t; \theta^\star, \delta). \tag{103}$$

Upon combining these two, we obtain the following,

$$|u_t - \lambda'_{t-1}\psi_t| \geq B_2 \sqrt{\log N^{(\mathcal{I})}_t} \|\lambda - \tilde{\lambda}^{(\mathcal{I})}_{t-1}\| \|\psi_t\|. \tag{104}$$

We have,

$$\begin{aligned} |u_t - \lambda'\psi_t| &= |\left(\tilde{\lambda}^{(\mathcal{I})}_{t-1}\right)'\psi_t - \lambda'\psi_t| \\ &\leq \|\tilde{\lambda}^{(\mathcal{I})}_{t-1} - \lambda'\| \|\psi_t\| \\ &\leq \frac{|u_t - \lambda'_{t-1}\psi_t|}{B_2\sqrt{\log N^{(\mathcal{I})}_t}}, \end{aligned} \tag{105}$$

where the second inequality follows from (104). From (100) we have,

$$\begin{aligned} |e_{t+1} - w_{t+1}| \geq& \left|\tilde{b}^{(\mathcal{I})}_{1,t-1}(u_t - \lambda'_{t-1}\psi_t)\right| - |b_1(u_t - \lambda'\psi_t)| \\ \geq& \left|\tilde{b}^{(\mathcal{I})}_{1,t-1}(u_t - \lambda'\psi_t) B_2 \sqrt{\log N^{(\mathcal{I})}_t}\right| - |b_1(u_t - \lambda'\psi_t)| \\ \geq& \left|\frac{b_1}{2}(u_t - \lambda'\psi_t) B_2 \sqrt{\log N^{(\mathcal{I})}_t}\right| - |b_1(u_t - \lambda'\psi_t)| \\ \geq& \left|b_1(u_t - \lambda'\psi_t) B_2\right|, \end{aligned}$$

where the second inequality follows from (105). This completes the proof. □

D.1. Auxiliary Results

The key result of this section is Lemma D.7, which is used in proof of Theorem D.2. It shows that on a high probability set, after a sufficiently long enough time we have $u_t = z_t$, where z_t is as in (37).

Proposition D.3. *When $\{w_t\}$ satisfies Assumptions (2.2, 2.3), then on the set $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$ we have,*

$$|z_t| \leq \left[\|\lambda\| + B_2 \frac{\log N^{(\mathcal{I})}_t}{\sqrt{N^{(\mathcal{I})}_t}} + \mathcal{E}(N^{(\mathcal{I})}_t; \theta^\star, \delta) \right] \|\psi_t\|, \tag{106}$$

where z_t is as in (21), and B_2 is as in (20). If Assumption 2.4 holds instead of Assumption 2.3, then the same conclusion holds on $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_w$.

Proof. The bound on the estimation error derived in Theorem E.5 yields the following on $\mathcal{G}_{LSE}$:

$$\|\lambda - \tilde{\lambda}^{(\mathcal{I})}_t\| \leq \mathcal{E}(N^{(\mathcal{I})}_t; \theta^\star, \delta). \tag{107}$$

This means that

$$\left|\lambda'\psi_t - \left(\tilde{\lambda}^{(\mathcal{I})}_t\right)'\psi_t\right| \leq \mathcal{E}(N^{(\mathcal{I})}_t; \theta^\star, \delta)\|\psi_t\|. \tag{108}$$

Moreover,

$$\begin{aligned} \left|\lambda'_{t-1}\psi_t - \lambda'\psi_t\right| &\leq \left|\lambda'_{t-1}\psi_t - \left(\tilde{\lambda}^{(\mathcal{I})}_t\right)'\psi_t\right| + |\lambda'\psi_t - \left(\tilde{\lambda}^{(\mathcal{I})}_t\right)'\psi_t| \\ &\leq \left|\lambda'_{t-1}\psi_t - \left(\tilde{\lambda}^{(\mathcal{I})}_t\right)'\psi_t\right| + \mathcal{E}(N^{(\mathcal{I})}_t; \theta^\star, \delta)\|\psi_t\|, \end{aligned} \tag{109}$$

where the second inequality follows from Theorem E.5. It then follows from the definition of z_t in (21), and the bounds (108), (109), that

$$\left|z_t - \lambda'\psi_t\right| \le \left[B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(N_t^{(\mathcal{I})};\theta^\star,\delta)\right]\|\psi_t\|, \tag{110}$$

or

$$|z_t| \le \left[\|\lambda\| + B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(N_t^{(\mathcal{I})};\theta^\star,\delta)\right]\|\psi_t\|. \tag{111}$$

This completes the proof. □

Proposition D.4. *When $\{w_t\}$ satisfies Assumptions (2.2, 2.3), then on the set $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$, we have the following: For times $t \ge t_1^\star$ that satisfy $t \notin \mathcal{I}$, on the event*

$$(\|\lambda\| + 1)\|\psi_t\| \le B_u, \tag{112}$$

we have,

$$|y_{t+1} - w_{t+1}| \le b_1 \left[B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(N_t^{(\mathcal{I})};\theta^\star,\delta)\right]\left[B_u\left(p + 1 + \frac{C_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q}|b_\ell|\right\}\right)\right], \tag{113}$$

where B_2 is as in (21). If $\{w_t\}$ satisfies Assumption 2.4 instead of Assumption 2.3, then the same conclusion holds on $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_w$.

Proof. From Lemma K.2,

$$\|Y_t\| \le C_1\rho^t\|Y_{t_0}\| + \frac{B_uC_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q}|b_\ell|\right\}, \tag{114}$$

where $Y_t = (y_t, y_{t-1}, \ldots, y_{t-p+1})'$. Recall $U_t = (u_t, u_{t-1}, \ldots, u_{t-q+2})'$. Since $\|\psi_t\| \le \|U_t\| + \|Y_t\|$, and $\|U_t\| \le B_u q$, for $t \ge t_1^\star$ we have,

$$\|\psi_t\| \le B_u q + B_u + \frac{B_uC_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q}|b_\ell|\right\}. \tag{115}$$

For $t \ge t_1^\star$,

$$B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(N_t^{(\mathcal{I})};\theta^\star,\delta) < 1, \tag{116}$$

so that from Proposition D.3 we have,

$$|z_t| \le (\|\lambda\| + 1)\|\psi_t\|.$$

Thus, when $(\|\lambda\| + 1)\|\psi_t\| < B_u$, we have $|z_t| < B_u$, so that from (20) we have $u_t = z_t$. This gives,

$$\begin{aligned} |y_{t+1} - w_{t+1}| &= |b_1(u_t - \lambda'\psi_t)| \qquad (117)\\ &= |b_1(z_t - \lambda'\psi_t)| \\ &\le b_1 \left[B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(N_t^{(\mathcal{I})};\theta^\star,\delta)\right]\|\psi_t\| \\ &\le b_1 \left[B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(N_t^{(\mathcal{I})};\theta^\star,\delta)\right]\left[B_u\left(p + 1 + \frac{C_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q}|b_\ell|\right\}\right)\right], \end{aligned} \tag{118}$$

where the second equality follows since $u_t = z_t$, the first inequality follows from (110), and the second inequality from (115). This completes the proof. □

Proposition D.5. *When $\{w_t\}$ satisfies Assumptions (2.2, 2.3), we have the following on $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$: For $t \geq \max\{t_1^\star, t_2^\star\}$ (116),*

$$|u_t| \leq \delta_1 B_u, \tag{119}$$

where $\delta_1 > 0$ satisfies (279)-(281).

Proof. It follows from Proposition D.3 that for $t \geq \max\{t_1^\star, t_2^\star\}$,

$$|z_t| \leq (\|\lambda\| + 1)\|\psi_t\|. \tag{120}$$

Moreover, from Proposition D.4 for $t \geq \max\{t_1^\star, t_2^\star\}$ where $t \notin \mathcal{I}$, on the event

$$(\|\lambda\| + 1)\|\psi_t\| \leq B_u, \tag{121}$$

we have,

$$|y_{t+1} - w_{t+1}| \leq b_1 \left[B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(N_t^{(\mathcal{I})}; \theta^\star, \delta) \right] \left[B_u \left(p + 1 + \frac{C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right) \right]. \tag{122}$$

From definition of $t_2^\star$ we have that for $t \geq t_2^\star$,

$$|y_{t+1} - w_{t+1}| \leq \frac{\delta_1^2}{2} B_u. \tag{123}$$

We have,

$$\begin{aligned} |w_t| &\leq B_w \\ &\leq \frac{\delta_1^2}{2} B_u, \end{aligned} \tag{124}$$

where the second inequality follows from (286). Upon combining this with (123), we obtain that when $(\|\lambda\| + 1)\psi_t \leq B_u$, then we have

$$\begin{aligned} |y_{t+1}| &\leq B_w + \frac{\delta_1^2}{2} B_u \\ &\leq \delta_1^2 \, B_u. \end{aligned} \tag{125}$$

Now, choose a sufficiently large episode i, so that its start time satisfies $n_i \geq \max\{t_1^\star, t_2^\star\}$. We will now show that for times $t \in \{n_i + m^\star, \ldots, n_{i+1}\}$, where $m_i \geq m^\star$[2], the following holds:

$$|u_t| \leq \delta_1 B_u, \tag{126}$$

$$\text{and } |y_{t+1}| \leq \delta_1^2 B_u. \tag{127}$$

We consider the following two cases separately.

Case 1). $t \in \{n_i + m^\star, \ldots, n_i + m_i\}$:

From (305) we have the following bound,

$$\|Y_{n_i+m}\| \leq C_1 \rho^m \|Y_{n_i}\| + B_w C_1 \sum_{s=0}^{m} \rho^s \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\}, \tag{128}$$

[2] $m^\star$ is as in (306).

where $m \le m_i$. When $m \ge m^\star$, so that it satisfies (307), this bound yields,

$$\begin{aligned}\|Y_{n_i+m}\| &\le B_w\left(1 + C_1 \sum_{s=0}^{m} \rho^s \left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\}\right)\\ &= \frac{\delta_1^2}{1+M(\Theta)} B_u \left(1 + C_1 \sum_{s=0}^{m} \rho^s \left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\}\right)\\ &\le \delta_1^2 B_u, \end{aligned} \tag{129}$$

where the equality follows from (286) and the last inequality from (290). This shows that (127) holds. Since $|u_t| \le B_w,\ \forall t \in \mathcal{I}$, (126) also clearly holds.

Case 2). $t \in \{n_i + m_i + 1, n_i + m_i + 2, \dots, n_{i+1}\}$: For $\nu < n_{i+1}$, assume that (126), (127) hold for times $t \in \{n_i + m^\star, \dots, n_i + m_i, \dots, \nu\}$. We will show that (126), (127) also hold for $t = \nu + 1$. Note that we have already shown above that they hold for $t \in \{n_i + m^\star, \dots, n_i + m_i\}$.

We have,

$$\begin{aligned}(\|\lambda\|+1)\|\psi_{\nu+1}\| &\le (\|\lambda\|+1)(|y_{\nu+1}| + |y_\nu| + \ldots + |y_{\nu+1-p}| + |u_\nu| + |u_{\nu-1}| + \ldots + |u_{\nu-q+1}|)\\ &\le (\|\lambda\|+1)\big(p\delta_1^2 + q\delta_1\big)B_u\\ &\le B_u, \end{aligned} \tag{130}$$

where the second inequality follows from the induction hypothesis (126), (127), and the last inequality follows since δ_1 has been chosen to satisfy (279) for $\theta = \theta^\star$. Since $(\|\lambda\|+1)\|\psi_{\nu+1}\| \le B_u$, it follows from (125) that

$$|y_{\nu+2}| \le \delta_1^2\ B_u. \tag{131}$$

This shows that (127) holds for $t = \nu + 1$. It remains to show that we have $|u_{\nu+1}| \le \delta_1 B_u$. Consider the bound on $\|U_t\|$ derived in Lemma K.1 (ii),

$$\|U_{t_1}\| \le C_1 \rho^{t_1 - t_0}\|U_{t_0}\| + \frac{C_1}{b_1} \sum_{s=0}^{t_1 - t_0 - 1} \rho^s \left\{|w_{t_1+1-s}| + \sum_{\ell=1}^{p} |a_\ell||y_{t_1+1-s-\ell}|\right\}, \tag{132}$$

where $t_1 > t_0$. Consider time $t_0 \in \{n_i + m^\star, n_i + m^\star + 1, \dots, n_i + m_i\}$. Then $|w_t| \le B_w \le \frac{B_u \delta^2}{2}$, where the second inequality follows from (286). Our induction hypothesis yields that $|y_t| \le B_u \delta_1^2$ for times $t \in \{t_0, t_0 + 1, \dots, \nu\}$, and also $|u_t| \le \delta_1 B_u$ for $t \in \{t_0 + 1, \dots, \nu - 1\}$, and we have shown above that $|y_{\nu+2}| \le \delta^2$. Consider the vector $U_{\nu+1} = (u_{\nu+1}, u_\nu, \dots, u_{\nu+1-q+2})'$. Note that it follows from our induction hypothesis that $\|U_{t_0}\| \le \delta_1 B_u q$. Upon substituting these bounds into (132), and setting $t_1 = \nu + 1$, we obtain the following,

$$\begin{aligned}\|U_{\nu+1}\| &\le C_1 \rho^{\nu+1-t_0} \delta_1 B_u q + \frac{C_1}{b_1} \sum_{s=0}^{\nu - t_0} \rho^s \left\{\frac{B_u \delta_1^2}{2} + \sum_{\ell=1}^{p} |a_\ell| \delta_1^2 B_u\right\}\\ &\le B_u \left(C_1 \rho^{\nu+1-t_0} \delta_1 q + \frac{C_1}{b_1} \sum_{s=0}^{\nu - t_0} \rho^s \left\{\frac{\delta_1^2}{2} + \sum_{\ell=1}^{p} |a_\ell| \delta_1^2\right\}\right)\\ &= \delta_1\ B_u \left(C_1 \rho^{\nu+1-t_0} q + \delta_1 \frac{C_1}{b_1} \sum_{s=0}^{\nu - t_0} \rho^s \left\{\frac{1}{2} + \sum_{\ell=1}^{p} |a_\ell|\right\}\right)\\ &\le \delta_1 B_u, \end{aligned} \tag{133}$$

where the last inequality holds since ν is sufficiently large and δ_1 sufficiently small, i.e.,

$$\nu + 1 - t_0 \ge \log_\rho \frac{1}{2C_1 q}, \tag{134}$$

$$\delta_1 \le \frac{1}{2}\left[\frac{C_1}{b_1(1-\rho)} \left\{\frac{1}{2} + \sum_{\ell=1}^{p} |a_\ell|\right\}\right]^{-1}. \tag{135}$$

The first condition holds since episode duration H is sufficiently large, i.e.,

$$H + 1 - (n_i + m^\star) \geq \log_\rho\left(\frac{1}{2C_1 q}\right). \tag{136}$$

while the second holds because δ_1 satisfies (279)-(280).

We have thus completed the induction step, and shown that (126), (127) holds for all $t \in \{n_i + m^\star, \ldots, n_{i+1}\}$.

It remains to be shown that $|u_t| \leq \delta_1 B_u$ for times $t \in \mathcal{I}$. But we already have $|u_t| \leq B_w < \delta_1 B_u$, where the last inequality follows from (283). This proves the claim. □

Proposition D.6. *When $\{w_t\}$ satisfies Assumptions (2.2, 2.3), then the following holds on the set $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$ for $t \geq \max\{t_1^\star, t_2^\star\}$:*

$$\|\psi_t\| \leq \left[\left(1 + C_1 \sum_{s=0}^{t} \rho^s \left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\}\right) + q\right] B_u \delta_1. \tag{137}$$

If instead $\{w_t\}$ satisfies Assumption 2.4, then the same conclusion holds on $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_w$.

Proof. We have shown in Proposition D.5 that on $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{\text{proj}} \cap \mathcal{G}_{\mathcal{I}}$, for $t \geq \max\{t_1^\star, t_2^\star\}$ we have,

$$|u_t| \leq \delta_1 B_u,$$

and hence,

$$\|U_t\| \leq q B_u \delta_1. \tag{138}$$

Also,

$$\begin{aligned} |w_t| &\leq B_w \\ &< B_u \delta_1, \end{aligned} \tag{139}$$

where the second inequality follows from (283). Upon substituting these bounds in Lemma K.1-(i), we obtain,

$$\begin{aligned} \|Y_t\| &\leq C_1 \rho^t \|Y_0\| + C_1 \sum_{s=0}^{t} \rho^s \left\{|w_{t-s}| + \sum_{\ell=1}^{q} |b_\ell||u_{t-s-\ell}|\right\} \\ &\leq C_1 \rho^t \|Y_0\| + B_u \delta_1 C_1 \sum_{s=0}^{t} \rho^s \left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\} \\ &\leq B_u \delta_1 \left(1 + C_1 \sum_{s=0}^{t} \rho^s \left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\}\right), \end{aligned} \tag{140}$$

where the first inequality follows since $|u_s| \leq \delta_1 B_u$, and the last inequality follows from the definition of $t_2(\rho)$ (96), and since $t \geq t_2^\star$. Since

$$\psi_t = (y_t, y_{t-1}, \ldots, y_{t-p+1}, u_{t-1}, u_{t-2}, \ldots, u_{t-q+1})',$$

we have $\|\psi_t\| \leq \|Y_t\| + \|U_t\|$. The proof is then completed by substituting the bounds on $\|Y_t\|$ and $\|U_t\|$. □

Lemma D.7. *When $\{w_t\}$ satisfies Assumptions (2.2, 2.3), then on the set $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$, for $t \geq \max\{t_1^\star, t_2^\star\}$, and $t \notin \mathcal{I}$, we have $u_t = z_t$. If instead $\{w_t\}$ satisfies Assumption 2.4, then the same conclusion holds on $\mathcal{G}_q \cap \mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_w$.*

Proof. From Proposition D.3 we have,

$$|z_t| \leq (\|\lambda\| + 1)\|\psi_t\|.$$

From Proposition D.6, we have,

$$\|\psi_t\| \leq \left[\left(1 + C_1 \sum_{s=0}^{t} \rho^s \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right) + q \right] B_u \delta_1. \tag{141}$$

Upon combining these two bounds, we obtain,

$$|z_t| \leq (\|\lambda\| + 1) \left[\left(1 + C_1 \sum_{s=0}^{t} \rho^s \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right) + q \right] B_u \delta_1.$$

Now, it follows from (20) that for $t \notin \mathcal{I}$ whenever $|z_t| < B_u$, the input u_t is set equal to z_t. Thus, when

$$(\|\lambda\| + 1) \left[\left(1 + C_1 \sum_{s=0}^{t} \rho^s \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right) + q \right] \delta_1 < 1,$$

we have $u_t = z_t$ for sufficiently large t. This condition is true since δ_1 satisfies (281). □

E. Estimation Error

PIECE uses multiple estimators while designing inputs $\{u_t\}$. A recursive estimate of λ using all the samples until t is generated as in (17), and denoted by λ_t. Let θ_t be the LS estimate of $\theta^\star$ using all the samples until time t. Let

$$\theta_t^{(\mathcal{I})} = \left(a_{1,t}^{(\mathcal{I})}, a_{2,t}^{(\mathcal{I})}, \ldots, a_{p,t}^{(\mathcal{I})}, b_{1,t}^{(\mathcal{I})}, b_{2,t}^{(\mathcal{I})}, \ldots, b_{q,t}^{(\mathcal{I})} \right),$$

be the least squares estimate of $\theta^\star$ using only the samples collected during the exploratory instants $\mathcal{I}$. We have,

$$V_t^{(\mathcal{I})} = I_{p+q} + \sum_{s \leq t, s \in \mathcal{I}} \phi_s \phi_s'. \tag{142}$$

We have,

$$\theta_t^{(\mathcal{I})} = \left(V_t^{(\mathcal{I})} \right)^{-1} \left(\sum_{s \leq t, s \in \mathcal{I}} \phi_s y_{s+1} \right). \tag{143}$$

Definition E.1. Define,

$$\mathcal{G}_{LSE} := \left\{ \omega : \left(\sum_{s \leq t, s \in \mathcal{I}} \phi_s w_s \right)' \left(\sum_{s \leq t, s \in \mathcal{I}} \phi_s \phi_s' \right)^{-1} \left(\sum_{s \leq t, s \in \mathcal{I}} \phi_s w_s \right) \leq 4\sigma_w^2 \log \left(\frac{\det(V_t^{(\mathcal{I})})^{\frac{1}{2}}}{\delta \lambda_{\min} \left(V_t^{(\mathcal{I})} \right)^{(p+q)/2}} \right), \ \forall t \right\}. \tag{144}$$

Lemma E.2. *We have,*

$$\mathbb{P}(\mathcal{G}_{LSE}) \geq 1 - \delta. \tag{145}$$

Proof. The estimation error at time t satisfies,

$$\begin{aligned}\|\theta^\star - \theta_t^{(\mathcal{I})}\|^2 &= \left\| (V_t^{(\mathcal{I})})^{-1} \left(\sum_{s \le t, s\in\mathcal{I}} \phi(s) w(s) \right) \right\|^2 \\ &\le \|(V_t^{(\mathcal{I})})^{-1/2}\|^2 \left\| (V_t^{(\mathcal{I})})^{-1/2} \left(\sum_{s \le t, s\in\mathcal{I}} \phi_s w_s \right) \right\|^2 \\ &= \|(V_t^{(\mathcal{I})})^{-1/2}\|^2 \left(\sum_{s \le t, s\in\mathcal{I}} \phi_s w_s \right)' \left[\left(V_t^{(\mathcal{I})}\right)^{-1/2} \right]' \left(V_t^{(\mathcal{I})}\right)^{-1/2} \left(\sum_{s \le t, s\in\mathcal{I}} \phi_s w_s \right) \\ &= \| \left(V_t^{(\mathcal{I})}\right)^{-1/2} \|^2 \left(\sum_{s \le t, s\in\mathcal{I}} \phi_s w_s \right)' \left(V_t^{(\mathcal{I})}\right)^{-1} \left(\sum_{s \le t, s\in\mathcal{I}} \phi_s w_s \right) \\ &\le \frac{1}{\lambda_{\min}(V_t^{(\mathcal{I})})} \left(\sum_{s \le t, s\in\mathcal{I}} \phi_s w_s \right)' \left(V_t^{(\mathcal{I})}\right)^{-1} \left(\sum_{s \le t, s\in\mathcal{I}} \phi_s w_s \right) \qquad (146) \\ &\le \frac{1}{\lambda_{\min}(V_t^{(\mathcal{I})})} \left(\sum_{s \le t, s\in\mathcal{I}} \phi_s w_s \right)' \left(V_t^{(\mathcal{I})}\right)^{-1} \left(\sum_{s \le t, s\in\mathcal{I}} \phi_s w_s \right). \qquad (147)\end{aligned}$$

The proof then follows from (294) by letting $\eta_s = w_s$ and $X_s = \phi_s$. ☐

Lemma E.3. *On $\mathcal{G}_{LSE}$,*

$$\|\theta^\star - \theta_t^{(\mathcal{I})}\|^2 \le \sigma_w^2 \left(\frac{(p+q)\log\left(\lambda_{\max}\left(V_t^{(\mathcal{I})}\right)\right) - 2\log(\delta)}{\lambda_{\min}(V_t^{(\mathcal{I})})} \right).$$

Proof. Follows from the bound (146) and the definition of $\mathcal{G}_{LSE}$. ☐

It follows from Lemma E.3 that in order to bound the estimation error, we need to derive an upper-bound on $\lambda_{\max}\left(V_t^{(\mathcal{I})}\right)$. This is done in the following result.

Lemma E.4. *If $\{w_t\}$ satisfies Assumptions (2.2, 2.3), then*

$$\lambda_{\max}(V_t^{(\mathcal{I})}) \le \left(\left(C_1\|Y_0\| + \frac{C_1 B_u}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right)^2 + qB_u^2 \right) N_t^{(\mathcal{I})} + 1, \ \forall t. \qquad (148)$$

Proof. Since $\lambda_{\max}(V_t^{(\mathcal{I})}) \le \sum_{s\le t, s\in\mathcal{I}} \|\phi_s\|^2$, we will derive an upper-bound on $\|\phi_s\|$:

$$\|\phi_s\| \le C_1\|Y_0\| + \frac{C_1 B_u}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} + qB_u, \ \forall s \in \mathcal{I}. \qquad (149)$$

Since $|u_s| \le B_u$ and $|w_s| \le B_w < B_u$, Lemma K.1-(i) yields

$$\begin{aligned}\|Y_t\| &\le C_1\rho^t\|Y_0\| + C_1 B_u \sum_{s=0}^{t-1} \rho^s \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \\ &\le C_1\|Y_0\| + \frac{C_1 B_u}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\}. \qquad (150)\end{aligned}$$

Since $\phi_t = (y_t, y_{t-1}, \ldots, y_{t-p+1}, u_t, u_{t-1}, \ldots, u_{t-q+1})'$, we have,

$$\|\phi_t\|^2 = \|Y_t\|^2 + \|U_t\|^2 \\ \leq \left(C_1\|Y_0\| + \frac{C_1 B_u}{1-\rho}\left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\}\right)^2 + qB_u^2, \tag{151}$$

where $U_t = (u_t, u_{t-1}, \ldots, u_{t-q+2})'$. The proof is completed by noting that $\lambda_{\max}(V_t^{(\mathcal{I})}) \leq \sum_{s\leq t, s\in\mathcal{I}} \|\phi_s\|^2$. □

It follows from Lemma E.3 that in order to obtain an upper-bound on the estimation error, a lower bound on $\lambda_{\min}(V_t^{(\mathcal{I})})$ under the proposed learning algorithm is required. This is derived in Section F. This is then used to prove the main result on estimation error below. Define the function,

$$\mathcal{E}(x;\theta,\delta) := \frac{(p+q)\log(x)}{2\beta_3 x} \\ + 2\sigma_w^2 \left[\frac{(p+q)\log\left[\left(C_1(\theta)\|Y_0\| + \frac{C_1(\theta)B_u}{1-\rho(\theta)}\left\{1+\sum_{\ell=1}^{q}|b_\ell(\theta)|\right\}\right)^2 + qB_u^2 + 1\right] - 2\log(\delta)}{\beta_3 x} \right], \tag{152}$$

for $x > 0$ where β_3 is as in (210), and σ_e^2 is as in (158). The sets $\mathcal{G}_q, \mathcal{G}_{LSE}, \mathcal{G}_{\text{proj}}, \mathcal{G}_{\mathcal{I}}, \mathcal{G}_w, \mathcal{G}_{w_B^2}$ are defined in (60) (144), (250), (157), (291), (270) respectively. We now state the main result of this section, that states a high-probability upper-bound on the error $\|\theta^\star - \theta_t^{(\mathcal{I})}\|$.

Theorem E.5. *If $\{w_t\}$ satisfies Assumptions (2.2, 2.3), then on the set $\mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_{w_B^2}$, the estimation error can be bounded as follows,*

$$\|\theta^\star - \theta_t^{(\mathcal{I})}\|^2 \leq \mathcal{E}(N_t^{(\mathcal{I})};\theta^\star,\delta), \tag{153}$$

where the function $\mathcal{E}(\cdot,;\delta,\theta)$ is as in (152). If instead $\{w_t\}$ satisfies Assumption 2.4, then the same conclusion holds on the set $\mathcal{G}_{LSE} \cap \mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_{w_{UB}^2}$

Proof. Follows from Lemma E.3, Lemma E.4 and Theorem F.4. □

F. Lower Bound on $\lambda_{\min}(V_t^{(\mathcal{I})})$

The required theory of projections is developed in Section F.1. This is utilized in order to obtain a lower bound on the quantity $\lambda_{\min}(V_t^{(\mathcal{I})})$. This lower bound is used in Section E for controlling the estimation error. This is stated in Theorem F.4, which is the main result of this section.

Consider the ARX model (2), repeated here for convenience,

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \ldots + a_p y_{t-p} + b_1 u_{t-1} + b_2 u_{t-2} + \cdots + b_q u_{t-q} + w_t, t = 1, 2, \ldots, \tag{154}$$

where u_s is $\mathcal{F}_{s-1}$-measurable, while y_s is $\mathcal{F}_s$-measurable. Consider the design matrix associated with θ_t, the LS estimate of θ at time t,

$$\Psi_t := \begin{pmatrix} y_{I-1} & \cdots & y_{I-p} & u_{I-1} & \cdots & u_{I-q} \\ \vdots & & & & & \\ y_{t-1} & \cdots & y_{t-p} & u_{t-1} & \cdots & u_{t-q} \end{pmatrix} \tag{155}$$

where $I > \max\{p, q\}$. Also let,

$$U_s = (u_{s-1}, u_{s-2}, \ldots, u_{s-q})'. \tag{156}$$

Lemma F.1. *Define the event,*

$$\mathcal{G}_{\mathcal{I}} := \left\{ \lambda_{\min}\left(\sum_{s=1}^{t} U_s U_s'\right) \geq \sigma_e^2 N_t^{(\mathcal{I})} - qB_u^2 \sqrt{2N_t^{(\mathcal{I})} \log\left(\frac{qN_T^{(\mathcal{I})}}{\delta}\right)} \right\}, \tag{157}$$

where σ_e^2 is the variance of the exploratory noise, i.e.

$$\sigma_e^2 := \mathbb{E}\big(u_t^2\big),\ t \in \mathcal{I}. \tag{158}$$

We have,

$$\mathbb{P}(\mathcal{G}_{\mathcal{I}}) \geq 1 - \delta. \tag{159}$$

For

$$t \geq \frac{2B_u^4}{\sigma_e^4} \log\left(\frac{qN_T^{(\mathcal{I})}}{\delta}\right), \tag{160}$$

on $\mathcal{G}_{\mathcal{I}}$, we have,

$$\lambda_{\min}\left(\sum_{s=1}^{t} U_s U_s'\right) \geq \frac{\sigma_e^2}{2} N_t^{(\mathcal{I})}. \tag{161}$$

Proof. Since,

$$\lambda_{\min}\left(\sum_{s\leq t} U_s U_s'\right) \geq \lambda_{\min}\left(\sum_{s\in\mathcal{I}, s\leq t} U_s U_s'\right), \tag{162}$$

we will instead derive a lower bound on the latter quantity. For $i \neq j$, consider the (i,j)-th element of the matrix $\sum_{s\in\mathcal{I},s\leq t} U_s U_s'$. This is given by $\sum_{s\in\mathcal{I}} u_{s-i}u_{s-j}$, without loss of generality assume $i < j$. Define new random variables $\{\tilde{u}_s\}$ such that $\tilde{u}_s$ is the s-th exploratory input.

This sum $\sum_{s\in\mathcal{I}} u_{s-i}u_{s-j}$ is equivalent to $\sum_s \tilde{u}_s\tilde{u}_{s+j-i}$. Now consider the filtration $\{\tilde{\mathcal{F}}_s\}$ defined as follows: $\tilde{\mathcal{F}}_s$ is the sigma-algebra generated by $\{\tilde{u}_\ell\}_{\ell=1}^{s}$. Now, $\{\tilde{u}_s\tilde{u}_{s+j-i}, \tilde{\mathcal{F}}_{s+j-i-1}\}$ is a martingale difference sequence. By using the Azuma-Hoeffding inequality we deduce,

$$\mathbb{P}\left(\Big|\sum_{s=1}^{t} \tilde{u}_s\tilde{u}_{s+j-i}\Big| > \epsilon\right) \leq 2\exp\left(-\frac{\epsilon^2}{2t(B_u^2)^2}\right),$$

where $t < N_T^{(\mathcal{I})}$. Letting $\epsilon = \sqrt{2t(B_u^2)^2 \log\left(\frac{q^2 N_T^{(\mathcal{I})}}{\delta}\right)}$, we deduce that the event $\left\{\Big|\sum_{s=1}^{t} \tilde{u}_s\tilde{u}_{s+j-i}\Big| > \sqrt{2t(B_u^2)^2 \log\left(\frac{q^2 N_T^{(\mathcal{I})}}{\delta}\right)}\right\}$, has a probability less than $\frac{\delta}{q^2 N_T^{(\mathcal{I})}}$. Upon using a union bound over t, and all possible $i \neq j$, we conclude that the probability of the following event is less than $\delta/2$,

$$\left\{\Big|\sum_{s=1}^{t} \tilde{u}_s\tilde{u}_{s+j-i}\Big| \geq \sqrt{2t(B_u^2)^2 \log\left(\frac{q^2 N_T^{(\mathcal{I})}}{\delta}\right)},\ \forall t = 1, 2, \ldots, N_T^{(\mathcal{I})}, \forall i, j \in \{1, 2, \ldots, q\}, i \neq j\right\}. \tag{163}$$

One may note that (163) is equivalent to the off-diagonal entries of $\sum_{s\in\mathcal{I},s\leq t} U_s U_s'$ being less than $B_u^2\sqrt{2t\log\left(\frac{qN_T^{(\mathcal{I})}}{\delta}\right)}$.

The diagonal terms of $\sum_{s\in\mathcal{I},s\leq t} U_s U_s'$ are $\sum_{s\leq N_t^{(\mathcal{I})}} \tilde{u}_s^2$. Upon using Azuma-Hoeffding and a union-bound on t on the process $\{\tilde{u}_s^2 - \mathbb{E}(\tilde{u}_s^2)\}$, we deduce that the following event has a probability less than $\delta/2$,

$$\left\{ \Big| \sum_{s=1}^{t} \tilde{u}_s^2 - t\mathbb{E}(\tilde{u}_s^2) \Big| \geq \sqrt{2t(B_u^2)^2 \log\left(\frac{4N_T^{(\mathcal{I})}}{\delta}\right)},\ \forall t = 1, 2, \ldots, N_T^{(\mathcal{I})} \right\}. \tag{164}$$

The proof then follows from the Gershgorin circle theorem (Horn & Johnson, 2012). □

Let $I > \max\{p, q\}$. For $\ell \in \mathbb{Z}_+$, define,

$$y^{(t)}(\ell) := (y_{I-\ell}, \ldots, y_{t-\ell}), \tag{165}$$

$$u^{(t)}(\ell) := (u_{I-\ell}, u_{I+1-\ell}, \ldots, u_{t-\ell})',\ \text{and} \tag{166}$$

$$\underline{w}^{(t)}(\ell) := (w_{I-\ell}, w_{I+1-\ell}, \ldots, w_{t-\ell})', \tag{167}$$

Define,

$$D_t := \begin{pmatrix} u_{I-1} & \cdots u_{I-q} \\ \vdots & \\ u_{t-1} & \cdots u_{t-q} \end{pmatrix} = \Big(u^{(t)}(1), \ldots, u^{(t)}(q)\Big). \tag{168}$$

The design matrix Ψ_t (155) can thus be written as

$$\Psi_t = \Big(y^{(t)}(1), \ldots, y^{(t)}(p), D_t\Big). \tag{169}$$

In the sequel, we will omit the superscript t when it is clear from the context. Let d be a column of Ψ_t, and $\hat{d}$ its projection onto the linear space spanned by the remaining columns of Ψ_t. We will derive a lower bound on the quantity $\|d - \hat{d}\|$. This will yield us a lower bound on $\lambda_{\min}(\Psi_t'\Psi_t)$ since from Lemma J.1 we have,

$$(p+q)\ \|d - \hat{d}\| \geq \lambda_{\min}(\Psi_t'\Psi_t) \geq (p+q)^{-1}\|d - \hat{d}\|. \tag{170}$$

Also define

$$\begin{aligned} t^\star_{cov}(\rho, \delta) := \inf \Bigg\{ & t \in \mathbb{N} : \frac{pqB_u^2}{2c_1} \vee \frac{8pqB_u^2}{c_1} \vee \frac{2c^2 \log\big(\frac{1}{\delta}\big)}{c_1^2} \\ & \vee p \frac{B_u^2 + 2\|Y_0\|^2 + \frac{4\|b\|^2}{(1-\rho)^2} B_u^2 q + \frac{4B_w^2}{(1-\rho)^2}}{(c_1/3)\delta} \vee \frac{2c^2 \log\big(\frac{1}{\delta}\big)}{(c_1/3)^2} \vee \frac{2B_w^4}{\sigma_e^4} \log\left(\frac{qN_T^{(\mathcal{I})}}{\delta}\right) \Bigg\}. \end{aligned} \tag{171}$$

To ease notation, we will occasionally omit the dependence of $t^\star_{cov}(\rho, \delta)$ on ρ, δ. The sets $\mathcal{G}_{\text{proj}}$, $\mathcal{G}_{\mathcal{I}}$ are defined in (250) and (157) respectively.

Proposition F.2. *Consider the ARX system (2) and assume that $\{w_s\}$ satisfies Assumptions 2.2 and 2.3. Let d be a column of D_t, and $\hat{d}$ its projection onto the linear space spanned by the remaining columns of Ψ_t. Let $\tilde{c} > 2c_1$, and define,*

$$\beta_1 := \frac{c_1/4}{6\left[\frac{2}{p}\|Y_0\|^2 + \frac{4\|b\|^2}{p(1-\rho)^2} B_u^2 q + \frac{4B_w^2}{p(1-\rho)^2}\right] + 4B_w^2}, \tag{172}$$

where c_1 is as in (5). On the set $\mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$, we have,

$$\|d - \hat{d}\| \geq \frac{\sigma_e^2 N_t^{(\mathcal{I})}}{2q} \left[1 \wedge \min_{\ell \in \{1,2,\ldots,p\}} \beta_1^{\ell/2}\right],\ \text{for } t \geq t^\star_{cov}(\rho, \delta). \tag{173}$$

Proof. Consider a column $d = (d_I, \ldots, d_t)$ of D_t. Let $D^\star$ be the sub-matrix of D_t consisting of all the other columns except d. For $\ell = 1, 2, \ldots$, let $D^\star(\ell)$ be the matrix $\left(y^{(t)}(\ell), y^{(t)}(\ell+1), \ldots, y^{(t)}(p), D^\star\right)$. For $\ell = 1, 2, \ldots, p$, let $\hat{d}_\ell$ denote the projection of d onto $L(y(\ell), y(\ell+1), \ldots, y(p), D^\star)$. Let $\hat{d}_0$ be the projection of d onto $L(D^\star)$.

We begin with deriving a lower-bound on $\|d - \hat{d}_0\|$. We have,

$$\begin{aligned} \|d - \hat{d}_0\| &\geq \frac{\lambda_{\min}(D_t' D_t)}{q} \\ &\geq \frac{\sigma_e^2 N_t^{(\mathcal{I})}}{2q}, \end{aligned} \tag{174}$$

where the first inequality follows from Lemma J.1, while the second follows from Lemma F.1 since $D_t' D_t = \sum_{s=1}^t U_s U_s'$. This shows $\lambda_{\min}(D_t' D_t) \geq \frac{\sigma_e^2}{2} N_t^{(\mathcal{I})}$. Upon substituting this into (174), we obtain

$$\|d - \hat{d}_0\| \geq \frac{\sigma_e^2 N_t^{(\mathcal{I})}}{2q}. \tag{175}$$

Now we will derive lower bounds for $\|d - \hat{d}_\ell\|$ for $\ell = 1, 2, \ldots, p$. We will show that

$$\|d - \hat{d}_\ell\|^2 \geq \|d - \hat{d}_0\|^2 \ \beta_1^{p+1-\ell}, \ \ell = 1, 2, \ldots, p, \tag{176}$$

where $\beta_1 > 0$ is as in (172). We will prove this via induction. We begin with $\ell = p$. Consider the vector $y^{(t)}(p)$. Its $i - I$-th element ($i \geq I$) is y_{i-p}, and is equal to

$$y_{i-p} = a_1 y_{i-p-1} + a_2 y_{i-p-2} + \ldots + a_p y_{i-p-p} + b_1 u_{i-p-1} + \ldots + b_q u_{i-p-q} + w_{i-p}.$$

This can be written in vector form as,

$$\begin{aligned} y^{(t)}(p) &= (y_{I-p}, \ldots, y_{t-p})' \\ &= (v_I, \ldots, v_t)' + (w_{I-p}, \ldots, w_{t-p})' \qquad (177) \\ &= (v_I, \ldots, v_t)' + \underline{w}^{(t)}(p), \qquad (178) \end{aligned}$$

where

$$v_i := a_1 y_{i-p-1} + \ldots + a_p y_{i-p-p} + b_1 u_{i-p} + \ldots + b_q u_{i-p-q}. \tag{179}$$

Let $v := (v_I, \ldots, v_t)'$. Note that v_i is $\mathcal{F}_{i-p-1}$ measurable. Hence u_i, and therefore also h_i are $\mathcal{F}_{i-k-1}$ measurable. From (177), $\hat{d}_p$ is the projection of d onto $L(D^\star, v + \underline{w}(p))$. Let $\hat{\underline{w}}_0(p)$ be projection of $\underline{w}(p)$ onto $D^\star$. Let $v^\star$ be the projection of v onto $L(D, y^{(t)}(p))$. Let $\hat{y}_0(p)$ be the projection of $y^{(t)}(p)$ onto $L(D^\star)$. Let $\hat{v}$ be the projection of v onto $L(D^\star)$. Define,

$$\mathcal{S}_3 : = 1 \vee \sqrt{\log^+\left(\frac{\|d - \hat{d}_p\|}{\delta}\right)} \vee \sqrt{2\log\left(\sum_{s=1}^t \|\phi_s\|^2\right)}, \tag{180}$$

$$\mathcal{S}_4 : = 1 \vee \sqrt{\log^+\left(\frac{\|v - \hat{v}\|}{\delta}\right)} \vee \sqrt{2\log\left(\sum_{s=1}^t \|\phi_s\|^2\right)}. \tag{181}$$

It follows from (310) that the conditions of Theorem F.10 are satisfied, and hence we can use Theorem F.10 and obtain the following after performing some algebraic manipulations,

$$\begin{aligned}\|d-\hat{d}_p\|^2 &\geq \frac{\|d-\hat{d}_0\|^2\left\{\|v-v^\star\|^2+\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2-\mathcal{S}_3\right\}}{\|(v-\hat{v})\|^2+\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2+2\|v-\hat{v}\|\left\{\|v-\hat{v}\|+\mathcal{S}_4\right\}}\\ &\geq \frac{\|d-\hat{d}_0\|^2\left\{\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2-\mathcal{S}_3\right\}}{\|(v-\hat{v})\|^2+\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2+2\|v-\hat{v}\|\left\{\|v-\hat{v}\|+\mathcal{S}_4\right\}}\\ &= \frac{\|d-\hat{d}_0\|^2\left\{\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2-\mathcal{S}_3\right\}}{3\|(v-\hat{v})\|^2+\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2+2\|v-\hat{v}\|\mathcal{S}_4}\\ &\geq \frac{\|d-\hat{d}_0\|^2\left\{\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2-\mathcal{S}_3\right\}}{6\|\left(y^{(t)}(p)-\hat{y}_0(p)\right)\|^2+7\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2+2\|v-\hat{v}\|\mathcal{S}_4} \qquad (182)\\ &\geq \frac{\|d-\hat{d}_0\|^2\left\{\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2-\mathcal{S}_3\right\}}{6\|y^{(t)}(p)\|^2+7\|\underline{w}(p)-\underline{\hat{w}}_0(p)\|^2+2\|v-\hat{v}\|\mathcal{S}_4}, \qquad (183)\end{aligned}$$

where (182) follows since $y(p)=v+\underline{w}$, so that $y(p)-\hat{y}(p)=v-\hat{v}+\underline{w}(p)-\underline{\hat{w}}(p)$, and hence $\|v-\hat{v}\|^2\leq 2\|y(p)-\hat{y}(p)\|^2+2\|\underline{w}(p)-\underline{\hat{w}}(p)\|^2$. We will now derive bounds on various terms in the numerator and denominator of (183), which will allow us to lower-bound this expression.

Recall $Y_t=(y_t,y_{t-1},\ldots,y_{t-p+1})'$. Since $y^{(t)}(p)=(y_{I-p},\ldots,y_{t-p})$, we have,

$$\|y^{(t)}(p)\|^2=\|Y_{t-p}\|^2+\|Y_{t-2p}\|^2+\ldots+\|Y_{t-\lfloor t\rfloor}\|^2. \qquad (184)$$

After bounding $\|Y_{t-p}\|^2, \|Y_{t-2p}\|^2,\ldots,\|Y_{t-\lfloor t\rfloor}\|^2$ using Proposition (K.5), and performing algebraic manipulations, we obtain,

$$\|y^{(t)}(p)\|^2\leq \frac{2t}{p}\|Y_0\|^2+\frac{4\|b\|^2 t}{p(1-\rho)^2}B_u^2 q+\frac{4B_w^2 t}{p(1-\rho)^2}. \qquad (185)$$

Proposition F.7 gives us the following lower bound,

$$\begin{aligned}\|\underline{w}(p)-\underline{\hat{w}}(p)\|^2 &\geq tc_1-\sqrt{2tc^2\log\left(\frac{1}{\delta}\right)}\\ &-p\left\{1\vee\log^+\left(\frac{q\times\sum_{j=1}^t u_j^2}{\delta}\right)\vee 2\log\left(q\times\sum_{j=1}^t u_j^2\right)\right\},\ \forall t, \qquad (186)\end{aligned}$$

where c_1 is as in Assumption 2.2.

We also have,

$$\begin{aligned}\|\underline{w}(p)-\underline{\hat{w}}(p)\|^2 &\leq \|\underline{w}^{(t)}(p)\|^2\\ &\leq B_w^2 t,\ \forall t. \qquad (187)\end{aligned}$$

Since from (179) we have $v_i = a_1 y_{i-p-1} + \ldots + a_p y_{i-p-p} + b_1 u_{i-p-1} + \ldots + b_q u_{i-p-q} = (a,b)\cdot\phi_{i-p}$ we get,

$$\begin{aligned}\|v-\hat{v}\| &\le \|v\| \\ &\le \|(a,b)\|\sqrt{\sum_{s=1}^{t-p}\|\phi_s\|^2} \\ &\le \|(a,b)\|\sqrt{\sum_{s=1}^{t} 2\|A^s\|^2\|Y_0\|^2 + \frac{4\|b\|^2}{1-\rho}\left(\sum_{s=1}^{t}\sum_{\ell=1}^{s}\rho^{s-\ell}\|U_\ell\|^2\right) + \frac{4}{(1-\rho)^2}B_w^2 t + qB_u^2 t} \\ &\le \|(a,b)\|\sqrt{\sum_{s=1}^{t} 2\|A^s\|^2\|Y_0\|^2 + \frac{4\|b\|^2}{(1-\rho)^2}\left(\sum_{s=1}^{t}\|U_s\|^2\right) + \frac{4}{(1-\rho)^2}B_w^2 t + qB_u^2 t} \\ &\le \|(a,b)\|\sqrt{\sum_{s=1}^{t} 2\|A^s\|^2\|Y_0\|^2 + \frac{4\|b\|^2}{(1-\rho)^2}B_u^2 qt + \frac{4}{(1-\rho)^2}B_w^2 t + qB_u^2 t}\quad ,\end{aligned} \tag{188}$$

where the third inequality follows from (314).

Substituting the bounds (185)-(188), and also $|u_s| \le B_u$ into (183), we get,

$$\|d^{(t)} - \hat{d}_p^{(t)}\|^2 \ge \|d^{(t)} - \hat{d}_0^{(t)}\|^2 \cdot \mathcal{T}_5,\ \forall t, \tag{189}$$

where,

$$\mathcal{T}_5 := \frac{tc_1 - \sqrt{2tc^2\log\left(\frac{1}{\delta}\right)} - p\left\{1 \vee \log^+\left(\frac{qB_u^2 t}{\delta}\right) \vee 2\log\left(qB_u^2 t\right)\right\} - \mathcal{S}_3}{6\left[\frac{2t}{p}\|Y_0\|^2 + \frac{4\|b\|^2 t}{p(1-\rho)^2}B_u^2 q + \frac{4B_w^2 t}{p(1-\rho)^2}\right] + 7tB_w^2 + 2\cdot\mathcal{T}_6\cdot\mathcal{S}_4}, \tag{190}$$

with,

$$\mathcal{T}_6 := \|(a,b)\|\sqrt{\sum_{s=1}^{t} 2\|A^s\|^2\|Y_0\|^2 + \frac{4\|b\|^2}{(1-\rho)^2}B_u^2 qt + \frac{4}{(1-\rho)^2}B_w^2 t + qB_u^2 t}. \tag{191}$$

After performing some algebraic manipulations, we get that for $t \ge t^\star_{cov}$ (171), we have,

$$\begin{aligned}\mathcal{T}_5 &\ge \frac{c_1/4}{6\left[\frac{2}{p}\|Y_0\|^2 + \frac{4\|b\|^2}{p(1-\rho)^2}B_u^2 q + \frac{4B_w^2}{p(1-\rho)^2}\right] + 7B_w^2} \\ &= \beta_1,\end{aligned} \tag{192}$$

or equivalently,

$$\|d^{(t)} - \hat{d}_p^{(t)}\|^2 \ge \|d^{(t)} - \hat{d}_0^{(t)}\|^2\beta_1,\ \forall t. \tag{193}$$

Next, suppose that (176) holds for $\ell = m+1, m+2, \ldots, p$. We will show that (176) holds for $\ell = m$. Once again, similar to (177), we have,

$$y(m) = v + \underline{w}(m), \tag{194}$$

where the $i-I$-th element ($i \ge I$) of v is y_{i-m}, and is given by,

$$y_{i-m} = a_1 y_{i-m-1} + a_2 y_{i-m-2} + \ldots + a_p y_{i-m-p} + b_0 u_{i-m} + \ldots + b_q u_{i-m-q}.$$

This shows that $\hat{d}_m$ is the projection of d onto $L(D^\star(m+1), v + \underline{w}(m))$. Let $\hat{\underline{w}}_{\ell_2}(\ell_1)$ be the projection of $\underline{w}(\ell_1)$ onto $D^\star(\ell_2)$, and $v^\star$ the projection of v onto $L(D^\star(m+1), y(m))$. It follows from (310) that the conditions of Theorem F.10 are satisfied, and hence we can use Theorem F.10 and arguments similar to (183) to obtain the following:

$$\|d - \hat{d}_m\|^2 \ge \|d - \hat{d}_{m+1}\|^2\beta_1. \tag{195}$$

This completes the induction step, and hence we have shown (176). The proof of the claim then follows by substituting (174) into (176). □

Proposition F.3. *Consider the ARX system (2) and assume that $\{w_s\}$ satisfies Assumptions (2.2, 2.3). Let c be a column of the matrix $(y_n(1), \dots, y_n(p))$, and $\hat{c}$ be its projection onto the linear space spanned by the remaining columns of*

$$\Psi_t = \left(y^{(t)}(1), \dots, y^{(t)}(p), D_t\right), \text{ where } D_t = \left(u^{(t)}(1), \dots, u^{(t)}(q)\right). \tag{196}$$

If instead $\{w_t\}$ satisfies Assumptions (2.2, 2.4), then on $\mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}}$ we have,

$$\|c - \hat{c}\| \geq \sqrt{\frac{c_1 t}{4}} \left[1 \wedge \min_{\ell \in \{1,2,\dots,p\}} \beta_2^{\ell/2}\right], \ \forall t \geq t^{\star}_{cov}, \tag{197}$$

where $t^{\star}_{cov}$ is as in (171), and

$$\beta_2 := \frac{(c_1/3)}{3a_1^2 \left[\frac{2}{p}\|Y_0\|^2 + \frac{4\|b\|^2}{p(1-\rho)^2} B_u^2 q + \frac{4B_w^2}{p(1-\rho)^2}\right] + B_w^2}. \tag{198}$$

Proof. Let the column c be $y^{(t)}(\ell)$, where $\ell \in \{1, 2, \dots, p\}$. Also let $\Psi^{\star}$ be the sub-matrix of Ψ_t obtained by removing $y^{(t)}(\ell)$. Recalling that $y^{(t)}(\ell) = (y_{I-\ell}, \dots, y_{t-\ell})'$ and

$$u^{(t)}(m) = (u_{I-m}, u_{I+1-m}, \dots, u_{t-m})',$$

define the matrix,

$$\Psi(\ell) := (y(\ell+1), y(\ell+2), \dots, y(\ell+p), u(0), u(1), \dots, u(\ell+q)).$$

Let L_ℓ be the linear space spanned by $y(1), y(2), \dots, y(\ell-1)$ and the columns of $\Psi(\ell)$. Since $L(\Psi^{\star})$ is a subspace of L_ℓ, clearly,

$$\|y(\ell) - \text{proj}(y(\ell), L(\Psi^{\star}))\| \geq \|y(\ell) - \text{proj}(y(\ell), L_\ell)\|.$$

Thus, in order to show the claim, we will derive a lower bound on $\|y(\ell) - \text{proj}(y(\ell), L_\ell)\|$. Let $\hat{\underline{w}}_0(\ell)$ be the projection of $\underline{w}(\ell)$ onto $L(\Psi(\ell))$.

For $i = 1, 2, \dots, \ell-1$, let π_i be the projection of $y(\ell)$ onto $L(y(\ell-i), \dots, y(\ell-1), \Psi(\ell))$. Also let π_0 be the projection of $y(\ell)$ onto $L(\Psi(\ell))$. We will now derive lower bounds on $\|y(\ell) - \pi_i\|$. We begin with $i = 0$. Now, $y(\ell)$ is a linear combination of the columns of $\Psi(\ell)$, and $\underline{w}(\ell) = (w_{I-\ell}, \dots, w_{t-\ell})'$,

$$y(\ell) = \sum_{s=1}^{p} a_s y(\ell+s) + \sum_{s=1}^{q} b_s u(\ell+s) + \underline{w}(\ell). \tag{199}$$

Since the vectors $\{y(\ell+s)\}_{s=1}^{p}$, $\{u(\ell+s)\}_{s=1}^{q}$ belong to $\Psi(\ell)$, we have,

$$y(\ell) - \pi_0 = \underline{w}(\ell) - \hat{\underline{w}}_0(\ell). \tag{200}$$

Hence, we will now derive a lower bound on $\underline{w}(\ell) - \hat{\underline{w}}_0(\ell)$. From Proposition F.7 we have (c is as in 265),

$$\|\underline{w}(\ell) - \hat{\underline{w}}_0(\ell)\|^2 \geq c_1 t - \sqrt{2tc^2 \log\left(\frac{1}{\delta}\right)} - p\left\{1 \vee \log^+\left(\frac{\sum_{s=0}^{\ell+q}\|u(s)\|^2 + \sum_{s=\ell+1}^{\ell+p}\|y(s)\|^2}{\delta}\right)\right.$$
$$\left. \vee \log\left(\sum_{s=0}^{\ell+q}\|u(s)\|^2 + \sum_{s=\ell+1}^{\ell+p}\|y(s)\|^2\right)\right\}. \tag{201}$$

Now $\sum_{s=0}^{\ell+q}\|u(s)\|^2 + \sum_{s=\ell+1}^{\ell+p}\|y(s)\|^2$ can be bounded by $B_u^2 t + 2t\|Y_0\|^2 + \frac{4\|b\|^2 t}{(1-\rho)^2} B_u^2 q + \frac{4B_w^2 t}{(1-\rho)^2}$ using techniques as in (184), (185). Thus, when $t \geq t^{\star}_{cov}$,

$$\|\underline{w}(\ell) - \hat{\underline{w}}_0(\ell)\|^2 \geq \frac{c_1 t}{3}, \tag{202}$$

which when combined with (200) yields,

$$\|y(\ell)-\pi_0\|^2 \geq \frac{c_1 t}{3}. \tag{203}$$

Next, consider $i=1$. We have

$$y(\ell-1) = a_1 y(\ell) + \sum_{s=1}^{p-1} a_{s+1} y(\ell+s) + \underline{w}(\ell-1), \tag{204}$$

where $\underline{w}(\ell-1) := (w_{I-\ell+1},\dots,w_{n-\ell+1})'$. Since the columns $\{y(\ell+s)\}_{s=1}^{p-1}$ belong to $\Psi(\ell)$, we get,

$$L(y(\ell-1),\Psi(\ell)) = L(a_1 y(\ell) + \underline{w}(\ell-1),\Psi(\ell)). \tag{205}$$

Hence, setting $v = a_1 y(\ell)$, we can use Theorem F.10 to obtain

$$\begin{aligned}&\|y(\ell)-\pi_1\|^2 \geq \|y(\ell)-\pi_0\|^2 \times \\ &\frac{\|\underline{w}(\ell-1)-\underline{\hat{w}}_0(\ell-1)\|^2 - \mathcal{S}_3}{3a_1^2\|y(\ell)-\pi_0\|^2 + \|\underline{w}(\ell-1)-\underline{\hat{w}}_0(\ell-1)\|^2 + 2\mathcal{S}_4 a_1\|y(\ell)-\pi_0\|},\end{aligned} \tag{206}$$

where $\underline{\hat{w}}_0(\ell-1)$ is the projection of $\underline{w}(\ell-1)$ onto $L(\Psi(\ell))$, and $\mathcal{S}_3,\mathcal{S}_4$ are as in (180), (181).

Similar to (185) we have,

$$\begin{aligned}\|y(\ell)-\pi_0\|^2 &\leq \|y(\ell)\|^2 \\ &\leq \frac{2t}{p}\|Y_0\|^2 + \frac{4\|b\|^2 t}{p(1-\rho)^2}B_u^2 q + \frac{4B_w^2 t}{p(1-\rho)^2}.\end{aligned} \tag{207}$$

Upon substituting this and the bounds (186)-(188) into (206), and performing algebraic manipulations similar to the proof of Proposition F.2, we obtain,

$$\|y(\ell)-\pi_1\|^2 \geq \|y(\ell)-\pi_0\|^2 \cdot \beta_2. \tag{208}$$

The proof is then completed by induction.

□

The following is the main result of this section and provides a lower bound on the minimum eigenvalue of $\Psi_t'\Psi_t$ that holds w.h.p.

Theorem F.4. *Consider the ARX system* (2) *and let* $\{w_t\}$ *satisfy Assumptions (2.2, 2.3). On* $\mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_{w_B^2}$*, for times* $t \geq t^\star_{cov}$*,*

$$\lambda_{\min}(\Psi_t'\Psi_t) \geq \beta_3 N_t^{(\mathcal{I})}, \tag{209}$$

where,

$$\beta_3 := \left(\frac{\sigma_e^2}{2q}\min_{\ell\in\{0,1,\dots,p\}}\beta_1^{\ell/2}\right) \wedge \left(\frac{c_1}{4}\min_{\ell\in\{0,1,\dots,p\}}\beta_2^{\ell/2}\right), \tag{210}$$

and Ψ_t *is as in* (155)*. Same conclusion holds for* $\{w_t\}$ *satisfying Assumptions (2.2, 2.4) on the set* $\mathcal{G}_{proj} \cap \mathcal{G}_{\mathcal{I}} \cap \mathcal{G}_{w_{UB}^2}$.

Proof. Follows from Propositions F.2, F.3, (170) and noting that $N_t^{(\mathcal{I})} \leq \sqrt{t}$ under the proposed algorithm, after some algebraic manipulations. □

F.1. Properties of Projections

The material in this section contains "finite-time version" of the results in (Lai & Wei, 1982a). More specifically, the proof of Theorem F.4 relies upon finite-time versions of Corollary 2 and Theorem 5 of (Lai & Wei, 1982a). In order to obtain these finite-time results, we will derive non-asymptotic versions of several results from (Lai & Wei, 1982a). The results in this section are of independent interest, and have much wider applications. The main result of this section is Theorem F.10, and it is used in the proof of Propositions F.2 and F.3 while deriving a lower bound on the minimum eigenvalue of the covariance matrix.

Within this section we consider stochastic processes $\{x_s\}_{s=1}^T, \{w_s\}_{s=1}^T, \{z_s\}_{s=1}^T, \{v_s\}_{s=1}^T$, where $z_s = (z_{s,1}, z_{s,2}, \dots, z_{s,p})$ is a vector-valued process. While performing analysis, we will be interested in t-dimensional vectors created from the first t components of these processes, with time index ranging from 1 to t. Hence denote $Z^{(t)} = \{z_{i,j}\}_{1\le i\le t, 1\le j\le p}$, $x^{(t)} = (x_1, x_2, \dots, x_t)'$, $w^{(t)} = (w_1, w_2, \dots, w_t)'$ and $v^{(t)} = (v_1, v_2, \dots, v_t)'$. For a matrix M, we let $L(M)$ be the linear space spanned by its columns. When the time t is clear from the context, we will omit the superscript t, which will be mostly the case in this section since the analysis is performed for a fixed t. So we will write x in lieu of $x^{(t)}$, and so on. Only when we explicitly want to depict the dependence upon t, will we use a super-script. Let $\hat{x}, \hat{w}, \hat{v}$ be the projections of the vectors x, w, v onto $L(Z)$. $\{w_s\}$ is a martingale difference sequence w.r.t. $\{\mathcal{F}_s\}$. For each $s \ge 1$, x_s, v_s, z_s are $\mathcal{F}_{s-1}$ measurable random variables. In this section, we will derive the results for the case when $\{w_s\}$ is either bounded (Assumption 2.3) or sub-Gaussian (Assumption 2.4).

Theorem F.10 is the main result of this section, and allows us to lower-bound the minimum eigenvalue of the covariance matrix. The following is the finite-time version of Theorem 4 of (Lai & Wei, 1982a).

Theorem F.5. *If $\{w_s\}$ satisfies Assumptions (2.2, 2.3), then on $\mathcal{G}_{proj}$ (250) we have,*

$$\begin{aligned}
&(x-\hat{x})\cdot(w-\hat{w}) = (x-\hat{x})\cdot w = x\cdot(w-\hat{w}) \\
&\qquad \le \|x-\hat{x}\| \left\{ 1 \vee \sqrt{\log^+\left(\frac{\|x-\hat{x}\|}{\delta}\right)} \vee \sqrt{2\sigma_w^2 \log\left(\sum_{s=1}^{t}\sum_{j=1}^{p} z_{s,j}^2\right)} \right\}. \qquad (211)
\end{aligned}$$

If instead $\{w_s\}$ satisfies Assumptions (2.2, 2.4), then (211) holds on $\mathcal{G}_{proj} \cap \mathcal{G}_w$.

Proof. Let $\mathcal{J}$ be a non-empty subset of $\{1, 2, \dots, p\}$. We will use $Z_{\mathcal{J}}$ to denote the $t \times |\mathcal{J}|$ matrix formed by the vectors $\{(z_{1,j}, z_{2,j}, \dots, z_{t,j})' : j \in \mathcal{J}\}$. Also, for $s = 1, 2, \dots, t$, let $Z_{s,\mathcal{J}}$ be the $|\mathcal{J}|$-dimensional column vector with components $z_{s,i}, i \in \mathcal{J}$. For each non-empty subset $\mathcal{J}$ of $\{1, 2, \dots, p\}$, define the following stopping-time,

$$\tilde{\tau}_{\mathcal{J}} := \inf\left\{ \ell : \sum_{s=1}^{\ell} Z_{s,\mathcal{J}} Z_{s,\mathcal{J}}' \text{ is non-singular} \right\},$$

where we let $\tilde{\tau}_{\mathcal{J}} = \infty$ if the set on the r.h.s. is empty. For times $\ell \ge \tilde{\tau}_{\mathcal{J}}$ define,

$$V_{\ell,\mathcal{J}} := \left(\sum_{s=1}^{\ell} Z_{s,\mathcal{J}} Z_{s,\mathcal{J}}' \right)^{-1}, \qquad (212)$$

while for $\ell < \tilde{\tau}_{\mathcal{J}}$ we let $V_{\ell,\mathcal{J}}$ be the Moore-Penrose generalized inverse of $\sum_{s=1}^{\ell} Z_{s,\mathcal{J}} Z_{s,\mathcal{J}}'$. Let,

$$X_{\ell,\mathcal{J}} := \sum_{s=1}^{\ell} x_s Z_{s,\mathcal{J}}, \quad \text{where } \ell = 1, 2, \dots, t. \qquad (213)$$

Let $\hat{x}(\mathcal{J})$ be the projection of x onto $L(Z_{\mathcal{J}})$. We have,

$$\text{proj}(x, L(Z_{\mathcal{J}})) = (Z_{1,\mathcal{J}}, Z_{2,\mathcal{J}}, \dots, Z_{t,\mathcal{J}})' V_{t,\mathcal{J}} X_{t,\mathcal{J}}. \qquad (214)$$

Thus,

$$(x - \text{proj}(x, L(Z_{\mathcal{J}}))) \cdot w = \sum_{s=1}^{t} \left\{ x_s - X_{t,\mathcal{J}}' V_{t,\mathcal{J}} Z_{s,\mathcal{J}} \right\} w_s, \qquad (215)$$

and also,

$$\|x - \text{proj}(x, L(Z_{\mathcal{J}}))\|^2 = \sum_{s=1}^{t} \Big(x_s - X_{t,\mathcal{J}}^{'} V_{t,J} Z_{s,\mathcal{J}} \Big)^2. \tag{216}$$

Assume that we have that there exists a $\mathcal{J}$ s.t. we have $\tilde{\tau}_{\mathcal{J}} < \infty$. By using Lemma F.13, we obtain that the following holds on $\mathcal{G}_{\text{proj}}$ (250),

$$\begin{aligned} &|(x - \text{proj}(x, L(Z_{\mathcal{J}}))) \cdot w| \\ &\le \|x - \text{proj}(x, L(Z_{\mathcal{J}}))\| \cdot \max \left\{ \sqrt{\log^+ \Big(\frac{\|x - \text{proj}(x, L(Z_{\mathcal{J}}))\|}{\delta} \Big)}, \sqrt{2\sigma_w^2 \log \left(\sum_{j \in \mathcal{J}} \sum_{s=1}^{t} z_{s,j}^2 \right)} \right\}. \end{aligned} \tag{217}$$

For times $t > \tilde{\tau}_{\mathcal{J}}$, we have $\hat{x}^{(t)} = \text{proj}\Big(x^{(t)}, L(Z_{\mathcal{J}}^{(t)}) \Big)$.

This completes the proof. ☐

Theorem F.6. *If $\{w_s\}$ satisfies Assumptions (2.2, 2.3), then on $\mathcal{G}_{proj}$ (250) the following holds,*

$$\|\hat{w}^{(t)}\|^2 \le p \left\{ 1 \vee \max_{j \in \{1,2,\ldots,p\}} \log^+ \left(\frac{\sum_{s=1}^{t} z_{s,j}^2}{\delta} \right) \vee 2\sigma_w^2 \log \left(\sum_{s=1}^{t} \sum_{\ell=1}^{j} z_{s,\ell}^2 \right) \right\}, \ \forall t. \tag{218}$$

If instead $\{w_s\}$ satisfies Assumption 2.4, then (218) holds on $\mathcal{G}_w \cap \mathcal{G}_{proj}$.

Proof. Let $Z_{\cdot,j}$ be the j-th column of Z, and $\tilde{Z}_{\cdot,j}$ the projection of $Z_{\cdot,j}$ onto the linear space spanned by $Z_{\cdot,1}, Z_{\cdot,2}, \ldots, Z_{\cdot,j-1}$. We let $\tilde{Z}_{\cdot,1} = \mathbf{0}$. Consider the orthogonal vectors $Z_{\cdot,1}, Z_{\cdot,2} - \tilde{Z}_{\cdot,2}, \ldots, Z_{\cdot,p} - \tilde{Z}_{\cdot,p}$. These span the space $L(Z)$. Since $\hat{w}$ is the projection of w onto $L(Z)$, we have,

$$\|\hat{w}\|^2 = \sum_{j=1}^{p} \frac{\left\{ \Big(Z_{\cdot,j} - \tilde{Z}_{\cdot,j} \Big) \cdot w \right\}^2}{\|Z_{\cdot,j} - \tilde{Z}_{\cdot,j}\|^2}. \tag{219}$$

In case the denominator of a summand in the above is 0, we set that term to 0. From Theorem F.5,

$$\begin{aligned} \left| \Big(Z_{\cdot,j} - \tilde{Z}_{\cdot,j} \Big) \cdot w \right| &\le \|Z_{\cdot,j} - \tilde{Z}_{\cdot,j}\| \times \left\{ 1 \vee \sqrt{\log^+ \left(\frac{\|Z_{\cdot,j} - \tilde{Z}_{\cdot,j}\|}{\delta} \right)} \vee \sqrt{2\sigma_w^2 \log \left(\sum_{s=1}^{t} \sum_{m=1}^{j-1} z_{s,m}^2 \right)} \right\} \\ &\le \|Z_{\cdot,j} - \tilde{Z}_{\cdot,j}\| \times \left\{ 1 \vee \sqrt{\log^+ \left(\frac{\sum_{s=1}^{t} z_{s,j}^2}{\delta} \right)} \vee \sqrt{2\sigma_w^2 \log \left(\sum_{s=1}^{t} \sum_{m=1}^{j-1} z_{s,m}^2 \right)} \right\}, \end{aligned} \tag{220}$$

where in the last inequality we have used $\|Z_{\cdot,j} - \tilde{Z}_{\cdot,j}\|^2 \le \|Z_{\cdot,j}\|^2 = \sum_{s=1}^{t} z_{s,j}^2$. The proof is then completed by substituting the above bound into (219). ☐

Proposition F.7. *If $\{w_t\}$ satisfies Assumption 2.2 and Assumption 2.4, then the folowing holds on the set $\mathcal{G}_{proj} \cap \mathcal{G}_{w_{UB}^2}$,*

$$\begin{aligned} &\|w^{(t)} - \hat{w}^{(t)}\|^2 \ge tc_1 - \sqrt{2tc^2 \log \Big(\frac{1}{\delta} \Big)} \\ &- p \left\{ 1 \vee \max_{j \in \{1,2,\ldots,p\}} \log^+ \left(\frac{\sum_{s=1}^{t} z_{s,j}^2}{\delta} \right) \vee 2\sigma_w^2 \log \left(\sum_{s=1}^{t} \sum_{m=1}^{j} z_{s,m}^2 \right) \right\}, \ \forall t, \end{aligned} \tag{221}$$

where $\mathcal{G}_{w_{UB}^2}$ is as in (266), while $\mathcal{G}_{proj}$ is as in (250). If instead $\{w_t\}$ satisfies Assumption 2.2 and Assumption 2.3, then (221) holds on $\mathcal{G}_{proj} \cap \mathcal{G}_{w_B^2}$. The sets $\mathcal{G}_{w_B^2}$ and $\mathcal{G}_{w_{UB}^2}$ are as in Lemma F.15 and Lemma F.14 respectively.

Proof. We have,

$$\begin{aligned}\|w^{(t)} - \hat{w}^{(t)}\|^2 &= \|w^{(t)}\|^2 - \|\hat{w}^{(t)}\|^2 \\ &\geq tc_1 - \sqrt{2t\left(\frac{1}{c}\log\left(\frac{T}{\epsilon'}\right)\right)^2 \log\left(\frac{1}{\delta}\right)} - \|\hat{w}^{(t)}\|^2 \\ &\geq tc_1 - \sqrt{2t\left(\frac{1}{c}\log\left(\frac{T}{\epsilon'}\right)\right)^2 \log\left(\frac{1}{\delta}\right)} \\ &\quad - p\left\{1 \vee \max_{j\in\{1,2,\ldots,p\}} \log^+\left(\frac{\sum_{s=1}^t z_{s,j}^2}{\delta}\right) \vee 2\sigma_w^2 \log\left(\sum_{s=1}^t \sum_{m=1}^j z_{s,m}^2\right)\right\}, \end{aligned} \tag{222}$$

where the first inequality follows from Lemma F.14, while the second follows from Theorem F.6. Both the inequalities hold on high probability sets. □

Proposition F.8. *Define,*

$$r := x - \hat{x}. \tag{223}$$

Let $\hat{\hat{r}}$ be the projection of r onto $L(v+w)$, and $\hat{\hat{x}}$ the projection of x onto $L(Z, v+w)$. Then,

$$\|x - \hat{x}\|^2 = \|\hat{\hat{r}}\|^2 + \|x - \hat{\hat{x}}\|^2. \tag{224}$$

Also,

$$\|\hat{\hat{r}}\|^2 = \frac{\left[r \cdot (v + w - \hat{v} - \hat{w})\right]^2}{\|v + w - \hat{v} - \hat{w}\|^2}. \tag{225}$$

Proof. We clearly have,

$$\hat{\hat{x}} = \hat{x} + \hat{\hat{r}}, \tag{226}$$

or equivalently,

$$r = x - \hat{x} = \hat{\hat{r}} + (x - \hat{\hat{x}}). \tag{227}$$

Since $\hat{\hat{r}}$ and $x - \hat{\hat{x}}$ are orthogonal, we get

$$\|x - \hat{x}\|^2 = \|\hat{\hat{r}}\|^2 + \|x - \hat{\hat{x}}\|^2. \tag{228}$$

This proves (224).

Since $r(= x - \hat{x})$ is orthogonal to $L(Z)$, and $\hat{v} + \hat{w}$ is the projection of $v + w$ onto $L(Z)$, we have that $\hat{\hat{r}}$ is also equal to the projection of r onto $L(v + w - \hat{v} - \hat{w})$. Hence,

$$\|\hat{\hat{r}}\|^2 = \frac{r \cdot (v + w - \hat{v} - \hat{w})}{\|v + w - \hat{v} - \hat{w}\|^2}, \tag{229}$$

which proves (225).

□

Proposition F.9. *Let $v^\star$ denote the projection of v onto $L(Z, x)$. Then,*

$$\begin{aligned} &\|r\|^2 \|v - \hat{v} + w - \hat{w}\|^2 - \left| r \cdot (v + w - \hat{v} - \hat{w}) \right|^2 \\ &= \|r\|^2 \left\{ \|v - v^\star\|^2 + \|w - \hat{w}\|^2 + 2(v - v^\star) \cdot w \right\} - (r \cdot w)^2, \end{aligned} \tag{230}$$

where $r = x - \hat{x}$ is as in (223).

Proof. Since $v-\hat{v}$ is orthogonal to $L(Z)$, its projection on $L(Z,x)$ is the same as its projection onto $L(x-\hat{x})$. Suppose that this projection is equal to $a(x-\hat{x}) = a\,r$. Then we have,

$$v-\hat{v} = a\,r + (v-v^\star), \tag{231}$$

where the second quantity in the r.h.s. above is the component that is orthogonal to $L(Z,x)$. Since the vectors r and $v-v^\star$ are orthogonal,

$$\|v-\hat{v}\|^2 = a^2\,\|r\|^2 + \|v-v^\star\|^2. \tag{232}$$

Upon taking dot product with the vector r on both sides of (231), we get,

$$r\cdot(v-\hat{v}) = a\,\|r\|^2. \tag{233}$$

This gives,

$$r\cdot(v-\hat{v}+w-\hat{w}) = a\|r\|^2 + r\cdot(w-\hat{w}). \tag{234}$$

Note that since $r = x-\hat{x}$, it is orthogonal to Z, hence it is also orthogonal to $\hat{w}$ so that we have $r\cdot\hat{w}=0$. Upon substituting this into the above relation, we get

$$r\cdot(v-\hat{v}+w-\hat{w}) = a\|r\|^2 + r\cdot w. \tag{235}$$

Taking squares on both sides,

$$\left[r\cdot(v-\hat{v}+w-\hat{w})\right]^2 = a^2\|r\|^4 + (r\cdot w)^2 + 2a\|r\|^2(r\cdot w). \tag{236}$$

Now,

$$\begin{aligned}\|(v-\hat{v})+(w-\hat{w})\|^2 &= \|v-\hat{v}\|^2 + \|w-\hat{w}\|^2 + 2(v-\hat{v})\cdot(w-\hat{w})\\ &= \|v-\hat{v}\|^2 + \|w-\hat{w}\|^2 + 2(v-\hat{v})\cdot w\\ &= a^2\,\|r\|^2 + \|v-v^\star\|^2 + \|w-\hat{w}\|^2 + 2(v-\hat{v})\cdot w\\ &= a^2\,\|r\|^2 + \|v-v^\star\|^2 + \|w-\hat{w}\|^2 + 2a\,r\cdot w + 2(v-v^\star)\cdot w,\end{aligned} \tag{237}$$

where the second equality follows since $v-\hat{v}$ is orthogonal to $L(Z)$, and hence $(v-\hat{v})\cdot\hat{w}=0$, the third follows from (232), and the last follows from (231). Upon multiplying (237) by $\|r\|^2$ and subtracting (236) from it, we get,

$$\begin{aligned}&\|r\|^2\|(v-\hat{v})+(w-\hat{w})\|^2 - \left[r\cdot(v-\hat{v}+w-\hat{w})\right]^2\\ &= \|r\|^2\left\{\|v-v^\star\|^2 + \|w-\hat{w}\|^2 + 2(v-v^\star)\cdot w\right\} - (r\cdot w)^2.\end{aligned} \tag{238}$$

This completes the proof. ☐

Theorem F.10. *If $\{w_t\}$ satisfies Assumptions (2.2, 2.3), then on $\mathcal{G}_{proj}$ (250) we have the following for all t,*

$$\|x^{(t)}-\hat{\hat{x}}^{(t)}\|^2 \ge \frac{Nr.}{\|(v-\hat{v})\|^2 + \|w-\hat{w}\|^2 + 2\|v-\hat{v}\|\left\{1\vee\sqrt{\log^+\left(\frac{\|v-\hat{v}\|}{\delta}\right)}\vee\sqrt{2\sigma_w^2\log\left(\sum_{s=1}^t\sum_{j=1}^p z_{s,p}^2\right)}\right\}}, \tag{239}$$

where,

$$\begin{aligned}Nr. = {}&\|r\|^2\left\{\|v-v^\star\|^2 + \|w-\hat{w}\|^2\right\} - \|x-\hat{x}\|^2\left\{1\vee\sqrt{\log^+\left(\frac{\|x-\hat{x}\|}{\delta}\right)}\vee\sqrt{2\sigma_w^2\log\left(\sum_{s=1}^t\sum_{j=1}^p z_{s,j}^2\right)}\right\}\\ &-2\|v-v^\star\|\left\{1\vee\log^+\frac{\|v-v^\star\|}{\delta}\vee 2\sigma_w^2\log\left(\sum_{s=1}^t\sum_{j=1}^p z_{s,j}^2 + \sum_{s=1}^t x_s^2\right)\right\}^{1/2},\end{aligned} \tag{240}$$

and where $r = x^{(t)}-\hat{x}^{(t)}$, $\hat{\hat{x}}^{(t)}$ is the projection of $x^{(t)}$ onto $L(Z^{(t)}, v^{(t)}+w^{(t)})$, and $(v^{(t)})^\star$ is the projection of $v^{(t)}$ onto $L(Z^{(t)}, x^{(t)})$. If the assumption on $\{w_t\}$ is replaced by Assumption 2.4, then the same conclusion holds on $\mathcal{G}_{proj}$ (250).

Proof. We note that,

$$\begin{aligned}\|x - \hat{\hat{x}}\|^2 &= \|r\|^2 - \|\hat{\hat{r}}\|^2 \\ &= \frac{\|r\|^2 \|v + w - \hat{v} - \hat{w}\|^2 - [r \cdot (v + w - \hat{v} - \hat{w})]^2}{\|v + w - \hat{v} - \hat{w}\|^2},\end{aligned} \tag{241}$$

where the first equality follows from (224), while the second one follows from Proposition F.8.

Next, we derive an upper-bound on the denominator of the above expression. We have,

$$\begin{aligned}\|(v - \hat{v}) + (w - \hat{w})\|^2 &= \|(v - \hat{v})\|^2 + \|w - \hat{w}\|^2 + 2(v - \hat{v}) \cdot (w - \hat{w}) \\ &\le \|(v - \hat{v})\|^2 + \|w - \hat{w}\|^2 \\ &\quad + 2\|v - \hat{v}\| \left\{ 1 \vee \log^+ \left(\frac{\|v - \hat{v}\|}{\delta} \right) \vee 2\sigma_w^2 \log \left(\sum_{s=1}^t \sum_{j=1}^p z_{s,j}^2 \right) \right\}^{1/2},\end{aligned} \tag{242}$$

where the inequality follows from Theorem F.5. From Proposition F.9, the numerator in (241) can be bounded as follows,

$$\begin{aligned}&\|r\|^2 \|v + w - \hat{v} - \hat{w}\|^2 - [r \cdot (v + w - \hat{v} - \hat{w})]^2 \\ &= \|r\|^2 \left\{ \|v - v^\star\|^2 + \|w - \hat{w}\|^2 + 2(v - v^\star) \cdot w \right\} - (r \cdot w)^2.\end{aligned} \tag{243}$$

The terms $(r \cdot w)^2$ and $\left|(v - v^\star) \cdot w\right|$ can be bounded using Theorem F.5 as follows,

$$\begin{aligned}(r \cdot w)^2 &= [(x - \hat{x}) \cdot w]^2 \\ &\le \|x - \hat{x}\|^2 \left\{ 1 \vee \log^+ \left(\frac{\|x - \hat{x}\|}{\delta} \right) \vee 2\sigma_w^2 \log \left(\sum_{s=1}^t \sum_{j=1}^p z_{s,j}^2 \right) \right\},\end{aligned} \tag{244}$$

and,

$$\left|(v - v^\star) \cdot w\right| \le \|v - v^\star\| \left\{ 1 \vee \log^+ \frac{\|v - v^\star\|}{\delta} \vee 2\sigma_w^2 \log \left(\sum_{s=1}^t \sum_{j=1}^p z_{s,j}^2 + \sum_{s=1}^t x_s^2 \right) \right\}^{1/2}. \tag{245}$$

Upon substituting these into (243), the numerator in (241) can be lower-bounded as follows,

$$\begin{aligned}&\|r\|^2 \|v + w - \hat{v} - \hat{w}\|^2 - [r \cdot (v + w - \hat{v} - \hat{w})]^2 \\ &\ge \|r\|^2 \left\{ \|v - v^\star\|^2 + \|w - \hat{w}\|^2 \right\} - \|x - \hat{x}\|^2 \left\{ 1 \vee \log^+ \frac{\|x - \hat{x}\|}{\delta} \vee 2\sigma_w^2 \log \left(\sum_{s=1}^t \sum_{j=1}^p z_{s,j}^2 \right) \right\} \\ &- 2\|v - v^\star\| \left\{ 1 \vee \log^+ \frac{\|v - v^\star\|}{\delta} \vee 2\sigma_w^2 \log \left(\sum_{s=1}^t \sum_{j=1}^p z_{s,j}^2 + \sum_{s=1}^t x_s^2 \right) \right\}^{1/2}.\end{aligned} \tag{246}$$

The proof is then completed by substituting the bounds (242) and (246) into (241). □

F.2. Auxiliary Results

Recall that $\{w_s\}$ is a martingale difference sequence with respect to $\{\mathcal{F}_s\}$. If $\mathcal{J}$ is a non-empty subset of $\{1, 2, \ldots, p\}$, then define $\tilde{\tau}_{\mathcal{J}} := \inf \left\{ \ell : \sum_{s=1}^{\ell} Z_{i,\mathcal{J}} Z'_{i,\mathcal{J}} \text{ is non-singular} \right\}$, and for times $s \ge \tilde{\tau}_{\mathcal{J}}$ define,

$$V_{s,\mathcal{J}} = \left(\sum_{k=1}^s Z_{k,\mathcal{J}} Z'_{k,\mathcal{J}} \right)^{-1}. \tag{247}$$

Recall $X_t = \sum_{s=1}^{t} x_s Z_{s,\mathcal{J}}$. Define,

$$s_t := \sum_{s=1}^{t}\Big(x_s - X_t' V_{t,\mathcal{J}} Z_{s,\mathcal{J}}\Big)^2, \tag{248}$$

$$d_t := x_t - X_t' V_{t,\mathcal{J}} Z_{t,\mathcal{J}}, and \tag{249}$$

$$\begin{aligned}
&\mathcal{G}_{\text{proj}} := \{\omega : \text{ (1, 2) below hold}\}, \text{ where}\\
&1)\sum_{k=1}^{s} \frac{Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}}\left(\sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right) w_k}{1 + Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} Z_k} \le \sqrt{\sum_{k=1}^{s} \frac{\left(Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{\left(1 + Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}\right)^2} + 1}\\
&\times \sqrt{\sigma_w^2 \log\left(\frac{1}{\delta}\left(\sum_{k=1}^{s} \frac{\left(Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{\left(1 + Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}\right)^2} + 1\right)\right)}, \forall \mathcal{J},\ s \text{ and}\\
&2)\sum_{s=1}^{t} d_s w_s \le \sqrt{\sigma_w^2\left(1 + \sum_{s=1}^{t} d_s^2\right)\log\left(\frac{1 + \sum_{s=1}^{t} d_s^2}{\delta}\right)}\ \forall \mathcal{J} \subseteq \{1, 2, \dots, p\}, \quad \forall t \Bigg\}\\
&3)\Big|\sum_{s=1}^{t} x_s w_s\Big| \le \sqrt{1 + \sum_s x_s^2}\sqrt{\sigma_w^2 \log\left(\frac{1 + \sum_s x_s^2}{\delta}\right)},\ \forall t.
\end{aligned} \tag{250}$$

Lemma F.11.

$$\mathbb{P}\Big(\mathcal{G}_{proj}^c\Big) \le 3\delta. \tag{251}$$

Proof. Follows from the self-normalization bound. (294) □

Lemma F.12. *If $\{w_s\}$ satisfies either Assumption 2.3 or Assumption 2.4, then on $\mathcal{G}_{proj}$ (250) we have,*

$$\begin{aligned}
\sum_{k=1}^{s} \frac{\left(Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{1 + Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}} &\le 2\sigma_w^2 \log\left(\lambda_{\max}\left(\sum_{k=1}^{s} Z_{k,\mathcal{J}} Z_{k,\mathcal{J}}'\right)\right)\\
&+ 8\log\left(\frac{1}{\delta}\right) + \log 8,\ \forall s.
\end{aligned} \tag{252}$$

Proof. The following is essentially (2.17) of (Lai & Wei, 1982b),

$$\begin{aligned}
&\left(\sum_{i=1}^{s} Z_{i,\mathcal{J}}' w_i\right) V_{s,\mathcal{J}} \left(\sum_{i=1}^{s} Z_{i,\mathcal{J}} w_i\right) + \sum_{k=1}^{s} \frac{\left(Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{1 + Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}}\\
&= \sum_{k=1}^{s} Z_{k,\mathcal{J}}' V_{k,\mathcal{J}} Z_{k,\mathcal{J}}\ w_k^2 + 2\sum_{k=1}^{s} \frac{Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}}\left(\sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right) w_k}{1 + Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}}.
\end{aligned} \tag{253}$$

Since $V_{s,\mathcal{J}}$ is positive semi-definite, this yields,

$$\begin{aligned}
&\sum_{k=1}^{s} \frac{\left(Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{1 + Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}}\\
&\le \sum_{k=1}^{s} Z_{k,\mathcal{J}}' V_{k,\mathcal{J}} Z_{k,\mathcal{J}}\ w_k^2 + 2\sum_{k=1}^{s} \frac{Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}}\left(\sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right) w_k}{1 + Z_{k,\mathcal{J}}' V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}}.
\end{aligned} \tag{254}$$

We will derive upper bounds on both the terms in the r.h.s. above. For the first term we have,

$$\sum_{k=1}^{s} Z^{'}_{k,\mathcal{J}} V_{k,\mathcal{J}} Z_{k,\mathcal{J}}\ w_k^2 = \sum_{k=1}^{s} Z^{'}_{k,\mathcal{J}} V_{k,\mathcal{J}} Z_{k,\mathcal{J}}\ \sigma_w^2 + \sum_{k=1}^{s} Z^{'}_{k,\mathcal{J}} V_{k,\mathcal{J}} Z_{k,\mathcal{J}} \big(w_k^2 - \mathbb{E}\left\{w_k^2 | \mathcal{F}_{k-1}\right\}\big). \tag{255}$$

After performing some algebraic manipulations we obtain the following,

$$\sum_{k=1}^{s} Z^{'}_{k,\mathcal{J}} V_{k,\mathcal{J}} Z_{k,\mathcal{J}}\ \sigma_w^2 \le \sigma_w^2 \log\left(\lambda_{\max}\left(\sum_{k=1}^{s} Z_{k,\mathcal{J}} Z^{\prime}_{k,\mathcal{J}}\right)\right). \tag{256}$$

To bound the second term on the r.h.s. of (254), we note that from the definition of $\mathcal{G}_{\text{proj}}$,

$$\sum_{k=1}^{s} \frac{Z^{'}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} \left(\sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right) w_k}{1 + Z^{'}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} Z_k} \le \sqrt{1 + \sum_{k=1}^{s} \frac{\left(Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{\left(1 + Z^{'}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}\right)^2}}$$

$$\times \sqrt{\sigma_w^2 \log\left(\frac{1}{\delta} \sum_{k=1}^{s} \frac{\left(Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{\left(1 + Z^{'}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}\right)^2} + \frac{1}{\delta}\right)}$$

$$\le \sqrt{1 + \sum_{k=1}^{s} \frac{\left(Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{\left(1 + Z^{'}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}\right)}} \sqrt{\sigma_w^2 \log\left(\frac{1}{\delta} \sum_{k=1}^{s} \frac{\left(Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{\left(1 + Z^{'}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}\right)} + \frac{1}{\delta}\right)}. \tag{257}$$

Upon substituting (256), (257) into (254), we get the following relation,

$$\sum_{k=1}^{s} \frac{\left(Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{1 + Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}} \le \sigma_w^2 \log\left(\lambda_{\max}\left(\sum_{k=1}^{s} Z_{k,\mathcal{J}} Z^{\prime}_{k,\mathcal{J}}\right)\right) \tag{258}$$

$$+ \sqrt{\sum_{k=1}^{s} \frac{\left(Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{1 + Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}} \log\left(\frac{\sum_{k=1}^{s} \frac{\left(Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} \sum_{i=1}^{k-1} Z_{i,\mathcal{J}} w_i\right)^2}{1 + Z^{\prime}_{k,\mathcal{J}} V_{k-1,\mathcal{J}} Z_{k,\mathcal{J}}}}{\delta}\right)}. \tag{259}$$

The proof is then completed by performing algebraic manipulations. □

Lemma F.13. *If $\{w_s\}$ satisfies Assumption 2.2 and either Assumption 2.3 or Assumption 2.4, then on $\mathcal{G}_{proj}$ (250) we have the following bound,*

$$\left|\sum_{s=1}^{t} \left(x_s - X^{'}_t V_t Z_{s,\mathcal{J}}\right) w_s\right| \le (s_t)^{1/2}$$

$$\times \max\left\{\sqrt{\log\left(\frac{s_t}{\delta}\right)}, \left[2B_w^2 \log\left(\lambda_{\max}\left(\sum_{s=1}^{t} Y_s Y^{\prime}_s\right)\right)\right]^{1/2}\right\},\ \forall t. \tag{260}$$

Proof. The following results are essentially Lemma 3 of (Lai et al., 1979),

$$\sum_{s=\tilde{\tau}_{\mathcal{J}}+1}^{t} \left(x_s - X^{'}_t V_t Z_{s,\mathcal{J}}\right) w_s = \sum_{s=\tilde{\tau}_{\mathcal{J}}+1}^{t} d_s \left\{w_s - Z^{'}_{s,\mathcal{J}} V_{s-1}\left(\sum_{j=1}^{s-1} Z_{j,\mathcal{J}} w_j\right)\right\}, \tag{261}$$

$$s_t = s_{\tilde{\tau}_{\mathcal{J}}} + \sum_{s=\tilde{\tau}_{\mathcal{J}}+1}^{t} d_s^2 \left(1 + Z^{'}_{s,\mathcal{J}} V_{s-1} Z_{s,\mathcal{J}}\right). \tag{262}$$

We now bound each term on the r.h.s. of (261) separately. It follows from the definition of $\mathcal{G}_{\text{proj}}$, that we have the following bound on the first term,

$$\Big|\sum_s d_s w_s\Big| \leq \sqrt{\left(\sum_{s=1}^t d_s^2\right)\log\left(\frac{\sum_{s=1}^t d_s^2}{\delta}\right)}$$
$$\leq \sqrt{s_t \log\Big(\frac{s_t}{\delta}\Big)}, \tag{263}$$

where the second inequality follows from (262). Using the Cauchy-Schwartz inequality,

$$\Big|\sum_{s=\tilde{\tau}_{\mathcal{J}}+1}^{t} d_s Z'_{i,\mathcal{J}} V_{s-1}\left(\sum_{j=1}^{s-1} Z_{j,\mathcal{J}} w_j\right)\Big|$$
$$\leq \left\{\sum_{s=\tilde{\tau}_{\mathcal{J}}+1}^{t} d_s^2\Big(1+Z'_{s,\mathcal{J}}V_{s-1}Z_{s,\mathcal{J}}\Big)\right\}^{1/2}\cdot\left\{\sum_{s=\tilde{\tau}_{\mathcal{J}}+1}^{t}\frac{\Big(Z'_{s,\mathcal{J}}V_{s-1}\sum_{j=1}^{s-1}Z_{j,\mathcal{J}}w_j\Big)^2}{\Big(1+Z'_{s,\mathcal{J}}V_{s-1}Z_{s,\mathcal{J}}\Big)}\right\}^{1/2}$$
$$\leq (s_t)^{1/2}\cdot\left\{\sum_{s=\tilde{\tau}_{\mathcal{J}}+1}^{t}\frac{\Big(Z'_{s,\mathcal{J}}V_{s-1}\sum_{j=1}^{s-1}Z_{j,\mathcal{J}}w_j\Big)^2}{\Big(1+Z'_{i,\mathcal{J}}V_{s-1}Z_{i,\mathcal{J}}\Big)}\right\}^{1/2}$$
$$\leq (s_t)^{1/2}\cdot\left[2\sigma_w^2\log\left(\lambda_{\max}\left(\sum_{k=1}^{s}Z_{k,\mathcal{J}}Z'_{k,\mathcal{J}}\right)\right)+8\log\Big(\frac{1}{\delta}\Big)+\log 8\right]^{1/2}, \tag{264}$$

where the last inequality follows from Lemma F.12. The proof is completed by substituting (263), (264) into (261). □

Under Assumption 2.4, the process $\{w_t\}$ is conditionally sub-Gaussian, and hence the process $|w_t^2 - \mathbb{E}\big(w_{t-1}^2|\mathcal{F}_{t-1}\big)|$ is sub-exponential (Vershynin, 2018), i.e. we have

$$\mathbb{P}\big(|w_t^2-\mathbb{E}\big(w_{t-1}^2|\mathcal{F}_{t-1}\big)|>x\big)\leq\exp(-cx),\ \forall x>0,\ \text{ for some } c>0. \tag{265}$$

Lemma F.14. *Define,*

$$\mathcal{G}_{w^2_{UB}} := \left\{\sum_{s=1}^t w_s^2 \geq c_1 t - \sqrt{2t\left(\frac{1}{c}\log\Big(\frac{T}{\epsilon'}\Big)\right)^2\log\Big(\frac{1}{\delta}\Big)},\ \forall t=1,2,\ldots,T\right\}. \tag{266}$$

Let $\{w_t\}$ satisfy Assumption 2.2 and Assumption 2.4. Then,

$$\mathbb{P}\Big(\mathcal{G}_{w^2_{UB}}\Big)\geq 1-(\delta+\epsilon'). \tag{267}$$

Proof. Using the union bound on individual increments and (265), we obtain that the following occurs w.p. less than ϵ',

$$\mathbb{P}\left(\exists s\in\{1,2,\ldots,T\}\ \text{ s.t. }\Big|w_s^2-\mathbb{E}\big(w_s^2|\mathcal{F}_{s-1}\big)\Big|>\frac{1}{c}\log\Big(\frac{T}{\epsilon'}\Big)\right)\leq\epsilon', \tag{268}$$

where $\epsilon'>0$.

Upon letting $B=\frac{1}{c}\log\big(\frac{T}{\epsilon'}\big)$ in Theorem J.2, and applying (268), the union bound over t, and letting $\epsilon'\leftarrow\frac{\epsilon}{T},\ \delta\leftarrow\frac{\delta}{T}$, we get,

$$\mathbb{P}\left(\left\{\sum_{s=1}^t\{w_s^2-\mathbb{E}\big(w_s^2|\mathcal{F}_{s-1}\big)\}<\sqrt{2t\left(\frac{1}{c}\log\Big(\frac{T}{\epsilon'}\Big)\right)^2\log\Big(\frac{1}{\delta}\Big)},\ \forall t=1,2,\ldots,T\right\}\right)\leq\delta+\epsilon'. \tag{269}$$

The proof is then completed by noting that $\mathbb{E}\big(w_t^2|\mathcal{F}_{t-1}\big)>c_1$. □

Lemma F.15. *Let* $\{w_t\}$ *satisfy Assumption 2.2 and Assumption 2.4. Define,*

$$\mathcal{G}_{w_B^2} := \left\{ \sum_{s=1}^{t} w_s^2 \geq c_1 t - \sqrt{2t(2B_w)^2 \log\left(\frac{1}{\delta}\right)} \right\}. \tag{270}$$

Then,

$$\mathbb{P}\left(\mathcal{G}_{w_B^2}\right) \geq 1 - \delta. \tag{271}$$

Proof. Follows from Azuma-Hoeffding J.2 after noting that $|w_t^2 - \mathbb{E}(w_t^2|\mathcal{F}_{t-1})|$ is bounded by $2B_w$. □

G. Duration of the First Exploratory Episode

Recall (11) lasts until the time $n_1 + m_1$ that is given as follows,

$$n_1 + m_1 = \max\left\{\inf\{t : b_{1,t} \neq 0\}, H_1(\Theta, \epsilon)\right\}. \tag{272}$$

The quantity $H_1(\Theta, \epsilon)$ is as follows,

$$H_1(\Theta, \epsilon) = t_1^\star(\rho) \vee t_2^\star(\rho) \vee t_3^\star(\rho) \vee t_5^\star(\epsilon_1) \vee t_6^\star(\epsilon_3, \delta). \tag{273}$$

Though these have been defined earlier, we repeat these for convenience,

$$t_1^\star(\rho) = \inf\left\{ t \in \mathbb{N} : B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(t; \theta^\star, \delta) < 1, C_1 \rho^t \|Y_0\| < B_u \right\}, \tag{274}$$

$$\begin{aligned} &t_2^\star(\rho) = \inf\Big\{t \in \mathbb{N} : \\ &\left[B_2 \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} + \mathcal{E}(t; \theta^\star, \delta)\right] \cdot \left[b_1 \left(p + 1 + \frac{C_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\}\right)\right] \leq \frac{\delta_1^2}{2}, \\ &\quad C_1 \rho^t \|Y_0\| < \delta_1 B_u \Big\}. \end{aligned} \tag{275}$$

$$t_3^\star(\rho) = \inf\left\{ t \in \mathbb{N} : \mathcal{E}(t; \theta^\star, \delta) \leq \frac{b_1}{2}, \text{ and } \frac{B_2 \sqrt{\log N_\ell^{(\mathcal{I})}}}{2} \geq 2 \right\}. \tag{276}$$

$$t_5^\star(\epsilon_1) = \frac{2(C_1\|Y_0\|)^2 + 2\left(\frac{B_u C_1}{1-\rho}\left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\}\right)^2 + qB_u^2}{\min\left\{\frac{\sigma_e^2}{2q}\min_{\ell \in \{1,2,\ldots,p\}} \beta_1^{\ell/2}, \frac{c_1}{4}\min_{\ell \in \{1,2,\ldots,p\}} \beta_2^{\ell/2}\right\}\epsilon_1}, \ \epsilon_1 > 0 \tag{277}$$

$$t_6^\star(\epsilon_3, \delta) = \inf\left\{t \in \mathbb{N} : \mathcal{E}(t; \theta^\star, \delta) \leq b_1 \epsilon_3\right\}, \ \epsilon_3 > 0. \tag{278}$$

H. Choosing B_u, the Threshold for Clipping Inputs

We will now describe how to choose B_u. We will also discuss the sensitivity of the algorithm to the choice of B_u.

We begin with few definitions.

Definition H.1. Let $\theta \in \mathbb{R}^{p+q}$ be a possible parameter associated with ARX (2). Let $(\delta_1(\theta), B_u(\theta))$ be a tuple that satisfies the following set of inequalities,

$$(\|\lambda(\theta)\| + 1)\big(p\delta_1(\theta)^2 + q\delta_1(\theta)\big) < 1, \tag{279}$$

$$\frac{1}{3}\left[\frac{C_1(\theta)}{b_1(\theta)(1-\rho(\theta))}\left\{\frac{1}{2B_u(\theta)} + \sum_{\ell=1}^{p}|a_\ell(\theta)|\right\}\right]^{-1} \geq \delta_1(\theta), \tag{280}$$

$$(\|\lambda(\theta)\| + 1)\left[\left(1 + \frac{C_1(\theta)}{1-\rho(\theta)}\left\{1 + \sum_{\ell=1}^{q}|b_\ell(\theta)|\right\}\right) + q\right]\delta_1(\theta) < 1. \tag{281}$$

$$\frac{B_w}{B_u(\theta)} \leq \frac{\delta_1(\theta)}{2},\ \frac{B_w}{B_u(\theta)} \leq \frac{\delta_1(\theta)}{\left(1 + \frac{C_1(\theta)}{1-\rho(\theta)}\left\{1 + \sum_{\ell=1}^{q}|b_\ell(\theta)|\right\}\right)}, \tag{282}$$

$$\frac{B_w}{B_u(\theta)} \leq \delta_1(\theta), \tag{283}$$

$$B_u(\theta) > 1. \tag{284}$$

Here $C_1(\theta), \rho(\theta)$ are as in Lemma K.1. When (279)-(283) are required to hold for every $\theta \in \Theta$, denote a solution to those inequalities by $(\delta_1(\Theta), B_u(\Theta))$ and denote $\delta_1 = \delta_1(\Theta), B_u = B_u(\Theta)$.

Obtaining δ_1, B_u: Our interest will be in obtaining a $(\delta_1(\Theta), B_u(\Theta))$. A solution to the above set of inequalities can be found using the following set of simplified inequalities. It can be verified that a solution to these inequalities also satisfies the above set of inequalities. Define

$$M(\Theta) := \sup_{\theta\in\Theta}\frac{C_1(\theta)}{1-\rho(\theta)}\left\{1 + \sum_{\ell=1}^{q}|b_\ell(\theta)|\right\}. \tag{285}$$

We use

$$B_u = \frac{B_w}{\delta_1^2}\cdot(1 + M(\Theta)), \tag{286}$$

where δ_1 satisfies the following inequalities,

$$\delta_1 \leq \frac{1}{(p+q)\sup_{\theta\in\Theta}(1 + \|\lambda(\theta)\|)} \tag{287}$$

$$\delta_1 \leq \frac{1}{3}\inf_{\theta\in\Theta}\left[\frac{b_1(\theta)(1-\rho(\theta))}{C_1(\theta)}\right]\left[\frac{\delta_1^2}{2B_w M(\Theta)} + \sup_{\theta\in\Theta}\sum_{\ell=1}^{p}|a_\ell(\theta)|\right]^{-1} \tag{288}$$

$$1 \geq \left(\sup_{\theta\in\Theta}\|\lambda(\theta)\| + 1\right)[M(\Theta) + q]\,\delta_1. \tag{289}$$

We now discuss how to compute B_u. In order to compute B_u efficiently, we will assume the following about the set Θ: (i) Θ is a polytope, (ii) $\sup_{\theta\in\Theta}\rho(\theta) \leq 1 - \epsilon_1$, where $\epsilon_1 > 0$. Recall that $M(\Theta)$ is as follows,

$$M(\Theta) = \sup_{\theta\in\Theta}\frac{C_1(\theta)}{1-\rho(\theta)}\left\{1 + \sum_{\ell=1}^{q}|b_\ell(\theta)|\right\}.$$

From the discussion above, we see that in order to set the value of B_u, it suffices to compute the following quantities:

- Lower bound on $M(\Theta)$: A lower bound is easily computed by evaluating the function $\frac{C_1(\theta)}{1-\rho(\theta)}\left\{1 + \sum_{\ell=1}^{q}|b_\ell(\theta)|\right\}$ at an arbitrary $\theta \in \Theta$.

- An upper-bound on $M(\Theta)$:Under our assumption on Θ, we have
$$M(\Theta) \leq \frac{1}{\epsilon_1} \cdot \left(\sup_{\theta\in\Theta} C_1(\theta) \right) \cdot \left(\sup_{\theta\in\Theta} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell(\theta)| \right\} \right).$$
Since the function $\sum_{\ell=1}^{q} |b_\ell(\theta)|$ is convex, and since from Bauer maximum principle (Bauer, 1958) we have that for a convex function maxima occurs at an extreme point, we can compute the term $(\sup_{\theta\in\Theta} \{1 + \sum_{\ell=1}^{q} |b_\ell(\theta)|\})$ by evaluating the function $\sum_{\ell=1}^{q} |b_\ell(\theta)|$ only at the extreme points of Θ. It remains to compute an upper-bound on $\sup_{\theta\in\Theta} C_1(\theta)$, which is discussed next.

- Computing $\sup_{\theta\in\Theta} C_1(\theta)$: Note that $C_1(\theta)$ is as follows: $\|A(\theta)\| \leq C_1(\theta)\rho^n$, $\|B(\theta)^n\| \leq C_1(\theta)\rho^n$, where $\rho < 1$ can be taken to be any number greater than the spectral radii of $A(\theta)$ and $B(\theta)$. Note that the dimensions of $A(\theta)$ and $B(\theta)$ are equal to p and $q-1$ respectively. In the rest of discussion, in order to keep notation simple, we will carry out derivations using the inequality $\|A(\theta)\| \leq C_1(\theta)\rho^n$ only, and it is straightforward to show that the same arguments hold for that involving $B(\theta)$ too. In order to compute an upper-bound on $C_1(\theta)$, we write $A(\theta)$ in its Jordan form, and after performing algebraic manipulations obtain the following,
$$C_1(\theta) \leq \max_{m\in\{1,2,\ldots,p\}} \max_{k\in\mathbb{N}} k^m \frac{\rho(\theta)^k}{(\rho(\theta)+\epsilon_2)^k},$$
where $\mathbb{N}$ is the set of natural numbers. Since the objective of the above optimization problem can equivalently be written as $\frac{1}{(1+\frac{\epsilon_2}{\rho(\theta)})^k}$, and since under our assumption we have $\rho(\theta) \leq 1-\epsilon_1$ for all $\theta \in \Theta$, we obtain
$$\sup_{\theta\in\Theta} C_1(\theta) \leq \max_{m\in\{1,2,\ldots,p\}} \max_{k\in\mathbb{N}} k^m \beta^k,$$
where,
$$\beta := 1 + \frac{\epsilon_2}{1-\epsilon_1},$$
is a constant that satisfies $\beta \in (0,1)$. Since relaxing the constraint $k \in \mathbb{N}$ to $k \in \mathbb{R}$ (the set of real numbers) only increases the optimal value of the optimization problem, we have
$$\sup_{\theta\in\Theta} C_1(\theta) \leq \max_{m\in\{1,2,\ldots,p\}} \max_{x\in\mathbb{R}} x^m \beta^x.$$
After solving the inner optimization problem, we get
$$\sup_{\theta\in\Theta} C_1(\theta) \leq \max_{m\in\{1,2,\ldots,p\}} \left[\frac{m}{\ln(1/\beta)} \right]^m (\beta)^{m/\ln(1/\beta)}.$$

- an upper-bound on $\sup_{\theta\in\Theta} \|\lambda(\theta)\|$: note that $\|\lambda(\theta)\|$ can be bounded from above by $p \max_{\ell=1,2,\ldots,p} \left|\frac{a_\ell(\theta)}{b_1(\theta)}\right|$, $(q-1)\max_{m=2,\ldots,q} \left|\frac{b_m(\theta)}{b_1(\theta)}\right|$. Since the functions $\left|\frac{a_\ell(\theta)}{b_1(\theta)}\right|$, $\left|\frac{b_m(\theta)}{b_1(\theta)}\right|$ can be optimized over Θ using linear fractional programming, an upper-bound on $\sup_{\theta\in\Theta} \|\lambda(\theta)\|$ is obtained efficiently.

- upper-bound on $\sup_{\theta\in\Theta} \sum_{\ell=1}^{p} |a_\ell(\theta)|$: obtained easily since the L_1 norm is a convex function, and hence maxima occurs at an extreme point. An upper-bound is thus computed by evaluating the function $\sum_{\ell=1}^{p} |a_\ell(\theta)|$ at the extreme points of Θ.

SENSITIVITY OF ALGORITHM TO CLIPPING THRESHOLD B_u:

We make the following observations: (i) B_u decides the duration of the warm-up phase $t^\star$ as $O(B_u^2)$ (Appendix B), and hence the regret during times $t \in \mathcal{I}$ is $O(B_u^3)$ (Appendix B.4). (ii) $\mathcal{T}_3 = \tilde{O}(B_u)$[3] and $\mathcal{T}_{4,2} = O(\log B_u)$, so that the regret during $t \notin \mathcal{I}$ is bounded as $\tilde{O}(B_u)$ (proof of Theorem B.5). In summary, the dependence of regret on B_u is $O(B_u^3)$.

[3] $\tilde{O}$ hides logarithmic factors.

I. Unbounded Noise Case

Recall that we assumed the following holds,

$$\sup_t \mathbb{E}\left\{\exp(\gamma|w_t|)\Big|\mathcal{F}_{t-1}\right\} \le \exp\big(\gamma^2\sigma^2/2\big), a.s.\ \forall t. \tag{290}$$

Since the noise is not bounded, we will restrict our analysis to the following set.

Lemma I.1. *Define,*

$$\mathcal{G}_w := \left\{|w_t| \le \sigma\sqrt{\log\left(\frac{T}{\delta}\right)},\ \forall t = 1, 2, \ldots, T\right\}, \tag{291}$$

where $\sigma > 0$*. Then,*

$$\mathbb{P}(\mathcal{G}_w^c) \le \delta. \tag{292}$$

Proof. It follows from Chernoff bound that $\mathbb{P}(|w_t| > x) \le \exp\big(-x^2/(2\sigma^2)\big)$. The proof then follows by letting $x = \sqrt{\log\left(\frac{T}{\delta}\right)}$, and using union bound for $t = 1, 2, \ldots, T$. □

Define,

$$B_w(T) := \sigma\sqrt{\log\left(\frac{T}{\delta}\right)}. \tag{293}$$

Since unlike the bounded noise case, in which we had $|w_t| \le B_w,\ \forall t$, there is no upper-bound on the noise values, the quantity $B_w(T)$ serves as a high-probability upper-bound. Indeed, most of the results derived so far under Assumption 2.3 continue to hold under Assumption 2.4 upon replacing B_w by $B_w(T)$. Since the analysis of regret for sub-Gaussian noise closely follows that of bounded noise, we will only highlight the differences between the two.

We begin with the lower-bound on $\lambda_{\min}(V_t)$ that was derived in Theorem F.4. Notice that β_3 involves B_w^2 in the denominator, and hence after replacing it by $B_w(T)$, we have that $\beta_3 \propto \frac{1}{\lesssim \log(T/\delta)}$, and hence decays with time-horizon T. In order to compensate for this, the algorithm explores more often, so that we let the episode duration H be equal to $1/\beta_3$, and let number of episodes until t be $\log t$, which yields $N_t^{(\mathcal{I})} \approx \frac{\log t}{\beta_3}$. With this change, the high-probability bound on the estimation error derived in Theorem E.5 is modified, so that after x episodes the error is bounded by

$$\frac{(p+q)\log(x)}{2x} + \frac{(p+q)\log\left[\left(C_1(\theta)\|Y_0\| + \frac{C_1(\theta)B_u}{1-\rho(\theta)}\left\{1 + \sum_{\ell=1}^{q}|b_\ell(\theta)|\right\}\right)^2 + qB_u^2\right] - 2\log(\delta)}{x}.$$

The relation $r_t \le (e_t - w_t)^2$ for $t \notin \mathcal{I}$ and $t \ge t_1^\star \vee t_2^\star \vee t_3^\star$ that was derived in Section D continues to hold, except that now in the definition of $t_1^\star, t_2^\star, t_3^\star$ we replace B_w by $B_w(T)$ and this introduces additional dependency upon T. We now discuss changes made while analyzing regret in Section C.

The cumulative regret $\sum_{t\in\mathcal{I}} r_t$ for times $t \in \mathcal{I}$ was determined by multiplying the cumulative number of exploratory instants $N_T^{(\mathcal{I})}$, with the bound $\log(T/\delta)\log(T)$ on $\|\phi_t\|^2$, where we had $\|\phi_t\|^2 \lesssim B_w^2$. We now have $N_T^{(\mathcal{I})} \gtrsim \log(T/\delta)\log(T)$, while $\|\phi_t\|^2 \lesssim \log(T/\delta)$. In summary, this regret is now bounded $(\log(T/\delta))^2\log(T)$.

We recollect that the analysis of $\sum_{t\notin\mathcal{I}} r_t$ involved summation of (45), and bounding $\mathcal{T}_2, \mathcal{T}_3, \mathcal{T}_4$. Thus, we will have to consider the dependence upon B_w of the bounds derived in Section C.3 on $\mathcal{T}_2, \mathcal{T}_3, \mathcal{T}_4$ therein. From Proposition C.7, the bound on $\mathcal{T}_3$ is $\lesssim B_u$ (after hiding terms that are $\log B_u, \log\log B_u$), or equivalently $\lesssim B_w$, and hence this bound is now $\lesssim B_w(T)$. Bound on $\mathcal{T}_{4,2}$ is $\lesssim \log(B_u)$, which is same as $\lesssim \log(B_w)$, and hence contributes a term that grows as $\log\log\left(\frac{T}{\delta}\right)$. Upon making the changes described above, we obtain the desired result.

J. Useful Results

J.1. Self-Normalized Martingales Concentration Results

Let $\{\mathcal{F}_t, t \in \mathbb{N}\}$ be a filtration and $\{\eta_t, t \in \mathbb{N}\}$ an $\mathcal{F}_t$-adapted process such that $\mathbb{E}\left\{\exp(\lambda\eta(t)\middle|\mathcal{F}_{t-1}\right\} \leq \exp(\lambda^2 R^2/2)$. Let $\{X(t)\}_{t\in\mathbb{N}}$ be a predictable process, i.e., $X(t)$ is $\mathcal{F}_{t-1}$ measurable. Define $\bar{V}(t) := V + \sum_{s=1}^{t} X(s)X(s)'$ and $S(t) := \sum_{s=1}^{t} \eta(s)X(s)$. The following holds w.p. greater than $1-\delta$:

$$\|S(t)\|^2_{\bar{V}(t)^{-1}} \leq R^2 \log\left(\frac{\det(V(t))\det(V)}{\delta}\right), \ \forall t \in \mathbb{N}. \tag{294}$$

The following result is essentially (3.6) of (Lai & Wei, 1982a).

Lemma J.1. *Consider an $n \times m$ matrix $A = \{a_{i,j}\}$, and denote its columns by $A_{\cdot,1}, A_{\cdot,2}, \dots, A_{\cdot,m}$. Let $\hat{A}_{\cdot,j}$ denote the projection of the j-th column on the linear space spanned by the remaining $m-1$ columns. Then*

$$m^{-1} \min_{1\leq j\leq m} \|A_{\cdot,j} - \hat{A}_{\cdot,j}\|^2 \leq \lambda_{\min}(A'A) \leq m \min_{1\leq j\leq m} \|A_{\cdot,j} - \hat{A}_{\cdot,j}\|^2. \tag{295}$$

The following result from (Tao & Vu, 2015) is essentially the Azuma-Hoeffding concentration inequality for unbounded random variables.

Theorem J.2. *Let $\{X_i\}_{i=1}^n$ be a supermartingale such that the differences are bounded w.h.p., i.e.*

$$\mathbb{P}(\ \exists\, i \ \text{ s.t. } |X_i - X_{i-1}| > B) \leq \epsilon, \tag{296}$$

where $\epsilon, B > 0$. Then,

$$\mathbb{P}(X_n > X_0 + x) \leq \exp\left(-\frac{x^2}{2nB^2}\right) + \epsilon. \tag{297}$$

K. Bounds on $\|Y_t\|$, $\|U_t\|$

Consider the following vector-valued processes associated with the ARX model (2): $Y_t = (y_t, y_{t-1}, \dots, y_{t-p+1})'$ and $U_t = (u_t, u_{t-1}, \dots, u_{t-q+2})'$. Consider the matrices

$$A = \begin{pmatrix} a_1 & \cdots & a_{p-1} & a_p \\ I_{p-1} & & & 0 \end{pmatrix}, \tag{298}$$

and,

$$B = \begin{pmatrix} -b_2/b_1 & \cdots & -b_q/b_1 \\ I_{q-2} & & 0 \end{pmatrix}, \tag{299}$$

where I_{p-1}, I_{q-2} are identity matrices of sizes $p-1$ and $q-2$ respectively. We have the following bounds, which are essentially Lemma 2-(i), (ii) of (Lai & Wei, 1987).

Lemma K.1. *Consider times $t_1 > t_0$, and let Assumption 2.1 hold true for the ARX model (2). There exists $0 < \rho < 1$, $C_1 > 0$ such that:*

I

$$\|Y_{t_1}\| \leq C_1 \rho^{t_1-t_0} \|Y_{t_0}\| + C_1 \sum_{s=0}^{t_1-t_0-1} \rho^s \left\{ |w_{t_1-s}| + \sum_{\ell=1}^{q} |b_\ell||u_{t_1-s-\ell}| \right\}. \tag{300}$$

II We have the following bound on $\|U_t\|$:

$$\|U_t\| \leq C_1 \rho^{t_1-t_0} \|U_{t_0}\| + \frac{C_1}{b_1} \sum_{s=0}^{t_1-t_0-1} \rho^s \left\{ |w_{t_1+1-s}| + \sum_{\ell=1}^{p} |a_\ell||y_{t_1+1-s-\ell}| \right\}. \tag{301}$$

We note that ρ above can be taken to be any number greater than the spectral radius of A but less than 1.

Lemma K.2. *For the case when $|w_s| \le B_w$ for all $s = 1, 2, \ldots$, we have,*

$$\|Y_t\| \le C_1 \rho^{t_1 - t_0} \|Y_{t_0}\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\}.$$

Proof. The proof follows from Lemma K.1-(i) after noting that $|w_s| \le B_w \le B_u$, and also $|u_s| \le B_u$ for all s. □

Remark K.3. When we want to indicate the dependence of ρ on the system parameter, we will write $\rho(\theta)$. Similarly for $a_\ell(\theta), b_\ell(\theta), C_1(\theta)$.

We now exhibit a result for $|y_t|,\ t \in \mathcal{I}$ that holds for the ARX process evolving under the PIECE algorithm. Note that the inputs $\{u_t\}$ are chosen so as to satisfy the following bounds,

$$|u_t| \le B_u\ \forall t, \ \ |u_t| \le B_w, \text{ for } t \in \mathcal{I}. \tag{302}$$

Moreover, the noise process $\{w_t\}$ is also bounded as,

$$|w_t| \le B_w, \ \text{ where } B_w \le B_u. \tag{303}$$

The exploratory phase $\mathcal{I}$ is comprised of several episodes, where the i-th episode consists of m_i consecutive steps, starting at time n_i and ending at time-step $n_i + m_i$.

Since $|u_s|, |w_s| \le B_u$, we obtain the following bound from Lemma K.1-(i) by setting $t_1 = n_i$ and $t_0 = 0$:

$$\begin{aligned} \|Y_{n_i}\| &\le C_1 \rho^{n_i} \|Y_0\| + B_u C_1 \sum_{s=0}^{t_1 - 1} \rho^s \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \\ &\le C_1 \rho^{n_i} \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\}. \end{aligned} \tag{304}$$

Now, in Lemma K.1-(i), we let $t_0 = n_i$, and $t_1 = n_i + m$ where $m < m_i$, so that we have $t_1 < n_i + m_i$, and also $|u_t|, |w_t| \le B_3$ for $n_i < t < n_i + m_i$. So,

$$\begin{aligned} \|Y_{n_i+m}\| &\le C_1 \rho^m \|Y_{n_i}\| + C_1 \sum_{s=0}^{m} \rho^s \left\{ |w_{n_i+m-s}| + \sum_{\ell=1}^{q} |b_\ell| |u_{n_i+m-s-\ell}| \right\} \\ &\le C_1 \rho^m \|Y_{n_i}\| + B_w C_1 \sum_{s=0}^{m} \rho^s \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \\ &\le C_1 \rho^m \left[C_1 \rho^{n_i} \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right] + B_w C_1 \sum_{s=0}^{m} \rho^s \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \\ &\le C_1 \rho^m \left[C_1 \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right] + B_w C_1 \sum_{s=0}^{m} \rho^s \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\}. \end{aligned} \tag{305}$$

Define,

$$m^\star := \left\lceil \frac{1}{\log \rho} \log \left(\frac{B_w}{\left[C_1 \|Y_0\| + \frac{B_u C_1}{1-\rho} \left\{ 1 + \sum_{\ell=1}^{q} |b_\ell| \right\} \right] \cdot C_1} \right) \right\rceil. \tag{306}$$

Upon choosing $m > m^\star$, we obtain the following bound on $\|Y_{n_i+m}\|$.

Lemma K.4. *Consider the ARX system (2) evolving under PIECE, in which $\{w_t\}$ satisfies Assumption 2.3, and $\{u_t\}$ satisfies (302). Consider the i-th exploratory episode, and let m satisfy,*

$$m^\star \le m \le m_i. \tag{307}$$

Then,

$$\|Y_{n_i+m}\| \le B_w \left[1 + C_1 \sum_{s=0}^{m} \rho^s \left\{1 + \sum_{\ell=1}^{q} |b_\ell|\right\}\right]. \tag{308}$$

If instead $\{w_t\}$ satisfies Assumption 2.4, the same conclusion holds on $\mathcal{G}_w$.

Proposition K.5. *Consider the ARX model (2). Under Assumption 2.3,*

$$\|Y_s\|^2 \le 2\|A^s\|^2\|Y_0\|^2 + \frac{4\|b\|^2}{1-\rho}\left(\sum_{\ell=1}^{s} \rho^{s-\ell}\|U_\ell\|^2\right) + \frac{4}{(1-\rho)^2}B_w^2, \tag{309}$$

and,

$$\operatorname{Tr}(\Psi_t'\Psi_t) \le \left(4\left[\frac{\|b\|}{1-\rho}\right]^2 + 1\right) q \operatorname{Tr}\left(\sum_{s=p+1}^{t} U_s U_s'\right) + 4\left[\frac{1}{1-\rho}\right]^2 B_w^2 t, \tag{310}$$

where $b = (b_1, b_2, \ldots, b_q)$, and ρ is as in Lemma K.1. Note that $tr(\Psi_t'\Psi_t) = \sum_{s=I}^{t} \|\phi_s\|^2$, where ϕ_s is the regressor during time s.

Proof. We have,

$$|w_t| \le B_w. \tag{311}$$

We will derive an upper-bound on $\sum_{s=1}^{t} y_s^2$. Define,

$$\tilde{U}_s := (b_1 u_{s-1} + b_2 u_{s-2} + \ldots + b_q u_{s-q} + w_s, 0, \ldots, 0)'. \tag{312}$$

We have,

$$Y_s = AY_{s-1} + \tilde{U}_s, \tag{313}$$

where A is as in (2). This yields,

$$\begin{aligned}
\|Y_s\|^2 &= \|A^s Y_0 + \sum_{\ell=1}^{s} A^{s-\ell}\tilde{U}_\ell\|^2 \\
&\le 2\|A^s\|^2\|Y_0\|^2 + 2\|\sum_{\ell=1}^{s} A^{s-\ell}\tilde{U}_\ell\|^2 \\
&\le 2\|A^s\|^2\|Y_0\|^2 + 2\left(\sum_{\ell=1}^{s} \|A^{s-\ell}\|\|\tilde{U}_\ell\|\right)^2 \\
&\le 2\|A^s\|^2\|Y_0\|^2 + 2\left(\sum_{\ell=1}^{s} \|A^{s-\ell}\|\right)\left(\sum_{\ell=1}^{s} \|A^{s-\ell}\|(b_1 u_{\ell-1} + b_2 u_{\ell-2} + \ldots + b_q u_{\ell-q} + w_\ell)^2\right) \\
&\le 2\|A^s\|^2\|Y_0\|^2 + \frac{2}{1-\rho}\left(\sum_{\ell=1}^{s} \|A^{s-\ell}\|(b_1 u_{\ell-1} + b_2 u_{\ell-2} + \ldots + b_q u_{\ell-q} + w_\ell)^2\right) \\
&= 2\|A^s\|^2\|Y_0\|^2 + \frac{2}{1-\rho}\left(\sum_{\ell=1}^{s} \|A^{s-\ell}\|(b \cdot U_\ell + w_\ell)^2\right) \\
&= 2\|A^s\|^2\|Y_0\|^2 + \frac{4}{1-\rho}\left(\sum_{\ell=1}^{s} \|A^{s-\ell}\|\left\{\|b\|^2\|U_\ell\|^2 + w_\ell^2\right\}\right) \\
&= 2\|A^s\|^2\|Y_0\|^2 + \frac{4}{1-\rho}\left(\sum_{\ell=1}^{s} \rho^{s-\ell}\left\{\|b\|^2\|U_\ell\|^2 + w_\ell^2\right\}\right) \\
&= 2\|A^s\|^2\|Y_0\|^2 + \frac{4\|b\|^2}{1-\rho}\left(\sum_{\ell=1}^{s} \rho^{s-\ell}\|U_\ell\|^2\right) + \frac{4}{(1-\rho)^2}B_w^2,
\end{aligned}$$

where the third inequality follows from the Cauchy-Schwartz inequality, and ρ is as in Lemma K.1. Upon using $\|\phi_s\|^2 = \|Y_s\|^2 + \|U_s\|^2$, we get,

$$\|\phi_s\|^2 \le 2\|A^s\|^2\|Y_0\|^2 + \frac{4\|b\|^2}{1-\rho}\left(\sum_{\ell=1}^{s}\rho^{s-\ell}\|U_\ell\|^2\right) + \frac{4}{(1-\rho)^2}B_w^2 + qB_u^2. \tag{314}$$

Summing up the above inequality from $s=1$ to t we get,

$$\sum_{s=1}^{t}\|\phi_s\|^2 \le 2\sum_{s=1}^{t}\|A^s\|^2\|Y_0\|^2 + \frac{4\|b\|^2}{1-\rho}\left(\sum_{s=1}^{t}\sum_{\ell=1}^{s}\rho^{s-\ell}\|U_\ell\|^2\right) + \frac{4}{(1-\rho)^2}B_w^2 t + \sum_{s=1}^{t}\|U_s\|^2 \tag{315}$$

$$\le \frac{2}{(1-\rho)^2}Y_0\|^2 + \frac{4\|b\|^2}{(1-\rho)^2}\left(\sum_{s=1}^{t}\|U_\ell\|^2\right) + \frac{4}{(1-\rho)^2}B_w^2 t + \sum_{s=1}^{t}\|U_s\|^2. \tag{316}$$

□

L. Simulations

L.1. Methodology

We demonstrate the empirical performance of the PIECE algorithm for a variety of auto-regressive linear systems. We compare the PIECE algorithm with the standard certainty equivalence approach (Algorithm 2) and the modified certainty equivalence approach by (Lai & Wei, 1987) (dubbed as LW, Algorithm 3). Table 2 describes the ARX models considered in the simulation experiments. Example I represents a linearized model of a paper machine (Åström & Wittenmark, 1973). Example II is an ARX model for a hairdryer (Ljung, 1998), where the y_t is the temperature of the air, and u_t is the power delivered to the hairdryer. The remaining examples are randomly generated ARX systems. Table 3 summarizes the results.

Ex.	p	q	y_t
I	2	2	$1.283y_{t-1} - 0.495y_{t-2} + 2.307u_{t-1} - 2.025u_{t-2} + w_t$
II	3	3	$1.4898y_{t-1} - 0.7025y_{t-2} + 0.1123y_{t-3} + 0.0039u_{t-1} + 0.0621u_{t-2} - 0.0284u_{t-3} + w_t$
III	4	4	$1.18y_{t-1} - 0.48y_{t-2} + 0.45y_{t-3} - 0.41y_{t-4} + 0.28u_{t-1} + 0.14u_{t-2} + 0.16u_{t-3} + 0.03u_{t-4} + w_t$
IV	2	3	$-0.01y_{t-1} - 0.46y_{t-2} + 0.1u_{t-1} + 0.086u_{t-2} + 0.02u_{t-3} + w_t$
V	6	6	$-0.66y_{t-1} - 0.79y_{t-2} - 0.2y_{t-3} - 0.03y_{t-4} + 0.09y_{t-6} + 0.32u_{t-1} + 0.06u_{t-2} - 0.2u_{t-3}$ $-0.01u_{t-4} - 0.03u_{t-5} + 0.001u_{t-6} + w_t$

Table 2. ARX Models

Algorithm 2 Certainty Equivalence (CE)

Input Horizon, H.
$\tau = \inf \left\{ t > 0 : \sum_{s \leq t} \phi_s \phi_s' \text{ is invertible and } b_{1,t} \neq 0 \right\}$.
if $t \leq \tau$ **then**
  Generate an exploratory input $u_t \sim \mathcal{N}(0, \sigma_u^2)$
else
  Compute the estimates θ_{t-1} and λ_{t-1} as follows:

$$te_{t-1} = (a_{1,t-1}, \dots, a_{p,t-1}, b_{1,t-1}, \dots, b_{q,t-1})' := \left(\sum_{s<t} \phi_s \phi_s' \right)^{-1} \left(\sum_{s<t} \phi_s y_{s+1} \right), \text{ and}$$

$$\lambda_{t-1} := (-1/b_{1,t-1}) \left(a_{1,t-1}, \dots, a_{p,t-1}, b_{2,t-1}, \dots, b_{q,t-1} \right)'$$

  Apply control, $u_t = \lambda_{t-1}' \psi_t$
end if

Hyper-Parameters: PIECE needs two system-dependent parameters and a bound on the absolute value of the noise in order to compute the algorithm's hyper-parameters. ρ is an upper bound on the eigenvalues of matrix A (see (2)) and $\|\lambda\|_2$ is the ℓ_2-norm of vector λ. The duration of the first exploration episode is $\|\lambda\|_2^3$. In the following table, their values are given for the three examples described in (5). Other hyper-parameters depend on the system as well as the noise process. B_w is the upper bound for the absolute value of the noise sequence. B_u, the threshold for the clipping input, and $\delta > 0$ are such that (B_u, δ) satisfies (287, 288 and 289). H, defined as in (12), is the exploration episode duration after the first episode. Among the noise process-dependent hyper-parameters, we observed that δ does not vary significantly across experiments.

System Noise	Regret	Estimation Error	Terminal Regret
IID Gaussian Noise, $\sigma = 0.2$	Figure 3	Figure 4	Table 5
IID Gaussian Noise, $\sigma = 0.6$	Figure 5	Figure 6	Table 7
IID Gaussian Noise, $\sigma = 1$	Figure 7	Figure 8	Table 9
Random walk with IID Gaussian steps, $\sigma = 0.2$	Figure 9	Figure 10	Table 11
Random walk with IID Gaussian steps, $\sigma = 0.5$	Figure 11	Figure 12	Table 13

Table 3. Summary of Results

Algorithm 3 Lai and Wei (LW)

Input Algorithm parameters, $\delta > 0, \rho > 1, B_2 > 0, B_w > 0$; horizon, H.
$\mathcal{I} = \{1, \ldots, \tau\} \cup (\cup_{i:n_i<H} \{n_i + 1, \ldots, n_i + m_i\})$ where
$\tau = \inf \left\{ t > 0 : \sum_{s \le t} \phi_s \phi_s' \textit{ is invertible and } b_{1,t} \neq 0 \right\}$, $n_i = e^{i^\rho (1+o(1))}$ and $m_i = (\log i)^\delta$.
if $t \in \mathcal{I}$ **then**
Generate an exploratory white noise input u_t such that $|u_t| \le B_w \log \log t$ and has mean 0.
else
Compute the estimates $\tilde{\theta}_{t-1}^{(\mathcal{I})}$, $\tilde{\lambda}_{t-1}^{(\mathcal{I})}$ and λ_t as defined in (15).

$$u_t = \begin{cases} (-B_u) \vee \left(\lambda_{t-1}' \psi_t\right) \wedge (B_u) \text{ if } \left|\lambda_{t-1}' \psi_t - \left(\tilde{\lambda}_{t-1}^{(\mathcal{I})}\right)' \psi_t\right| \le B_2 \times \frac{\log N_t^{(\mathcal{I})}}{\sqrt{N_t^{(\mathcal{I})}}} \|\psi_t\|, \\ (-B_u) \vee \left(\left(\tilde{\lambda}_{t-1}^{(\mathcal{I})}\right)' \psi_t\right) \wedge (B_u) \text{ otherwise .} \end{cases}$$

end if

	ρ	$\|\lambda\|_2$	δ
Example I	0.7036	1.06	0.0231
Example II	0.7596	423.68	0.0001
Example III	0.8986	5.1	0.0038
Example IV	0.6782	4.69	0.0227
Example V	0.8282	3.36	0.0104

Table 4. System parameters

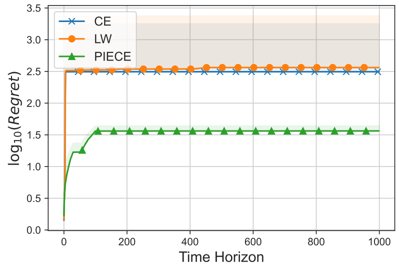
So, we kept it constant for a particular system and added its values in the following table. The experiment-dependent hyper-parameter values are given along with the corresponding experimental results.

System Noise: We perform our experiments for two different noise processes : (i) Gaussian Noise and (ii) Random walk with IID Gaussian Steps.

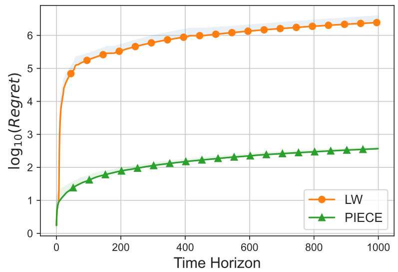
L.2. Regret and Estimation Error

Cumulative Regret: We plot the logarithms of the cumulative regrets, $\log(R_t)$ and highlight the cumulative regret at the end of the experiment. One of the key issues with many adaptive controllers is their empirical performance in the initial phase of learning (Lale et al., 2022; Mete et al., 2022). It is evident from the empirical results that CE as well as LW both suffer from this issue. As described in Section 3, the PIECE algorithm differs from LW with regard to the clipping of the input as well as the choice of exploration episodes. The empirical results demonstrate that PIECE does not suffer a large regret at the beginning of the experiments, unlike the LW and the CE controllers. The benefits of the PIECE modifications of clipping as well as improved exploration strategy are clearly evident as the resulting algorithm has much lower empirical regret compared to LW or the standard CE controller.

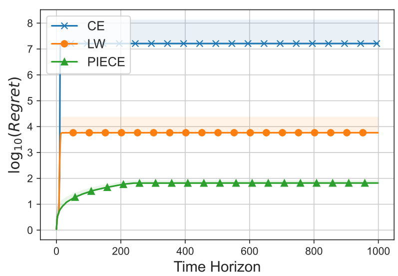
Estimation Error: We also plot the estimation error $||\theta_t - \theta^\star||^2$. It is interesting to note that LW has better estimation error than PIECE. This reiterates the point that the exploration scheme in PIECE is more efficient in achieving lower regret, which is the primary objective of the controller, at the cost of a higher estimation error.



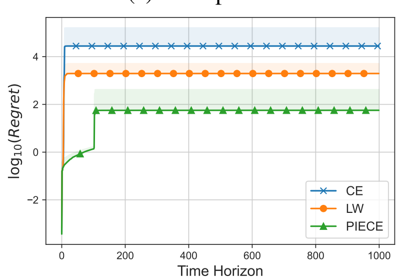
(a) Example I



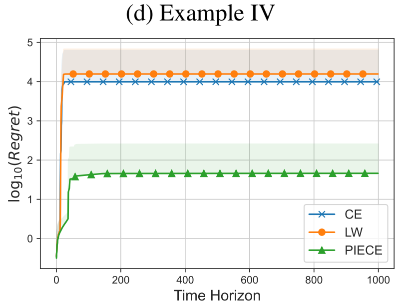
(b) Example II



(c) Example III

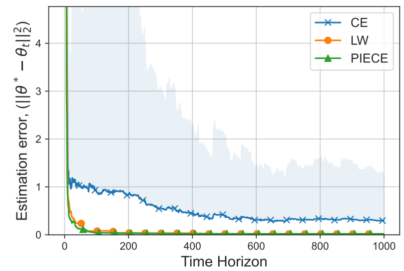


(d) Example IV

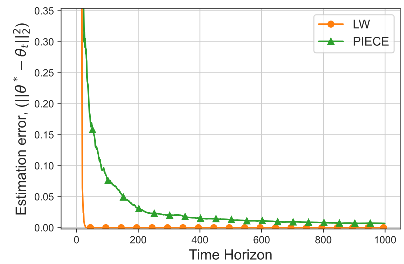


(e) Example V

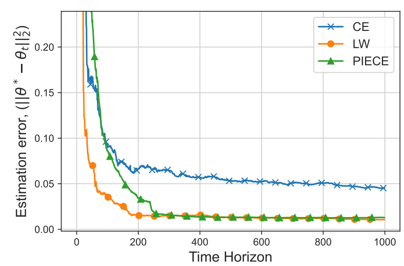
Figure 3. Log(Cumulative Regret) averaged over 50 runs for Gaussian noise with mean 0 and standard deviation 0.2.



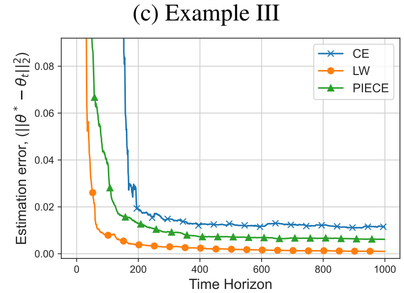
(a) Example I



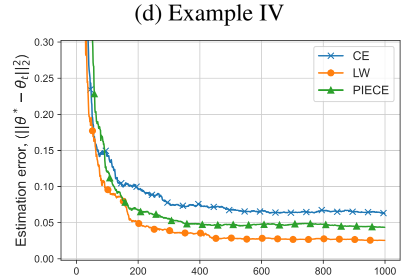
(b) Example II



(c) Example III

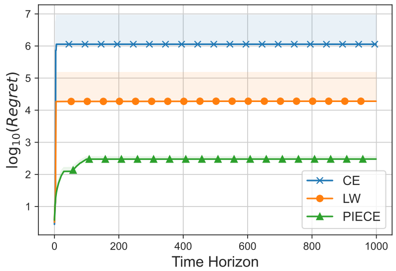


(d) Example IV

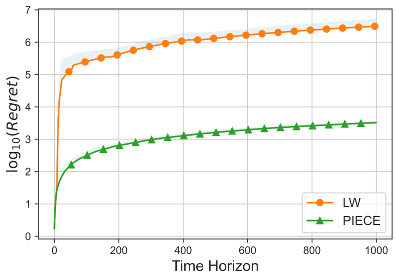


(e) Example V

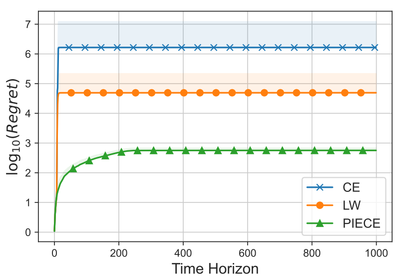
Figure 4. Estimation Error ($||\theta^\star - \theta_t||_2^2$) for Gaussian noise with mean 0 and standard deviation 0.2.



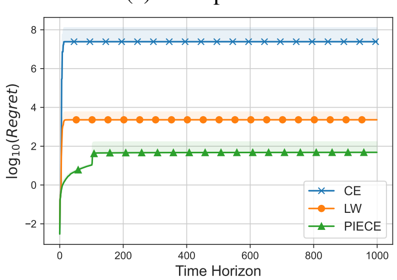
(a) Example I



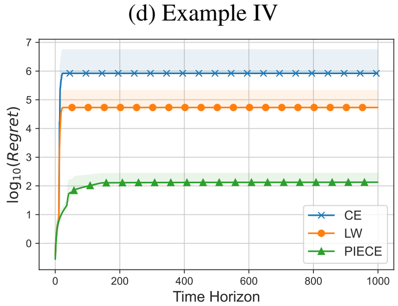
(b) Example II



(c) Example III

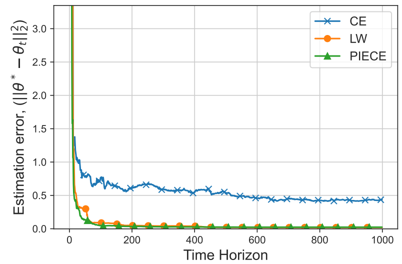


(d) Example IV

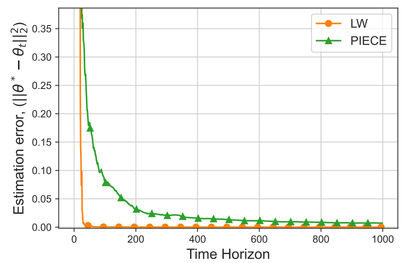


(e) Example V

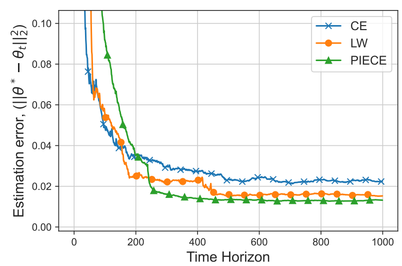
Figure 5. Log(Cumulative Regret) averaged over 50 runs for Gaussian noise with mean 0 and standard deviation 0.6.



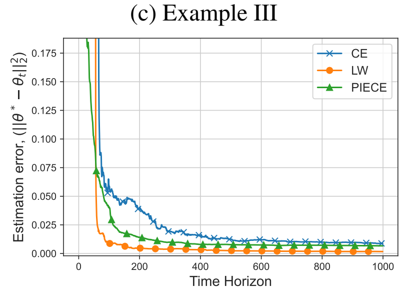
(a) Example I



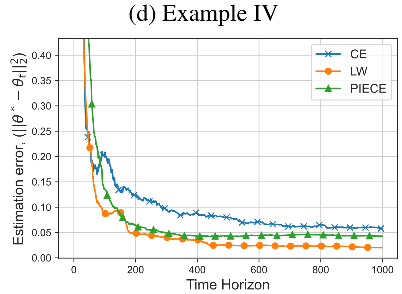
(b) Example II



(c) Example III

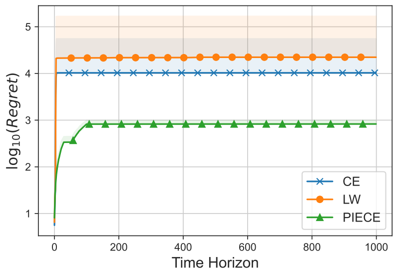


(d) Example IV

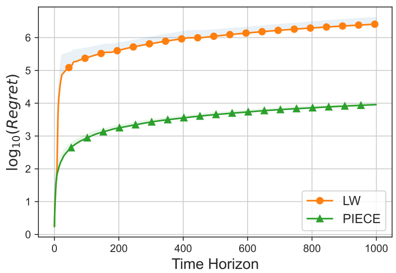


(e) Example V

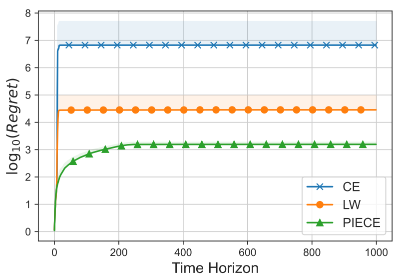
Figure 6. Estimation Error ($||\theta^\star - \theta_t||_2^2$) for Gaussian noise with mean 0 and standard deviation 0.6.



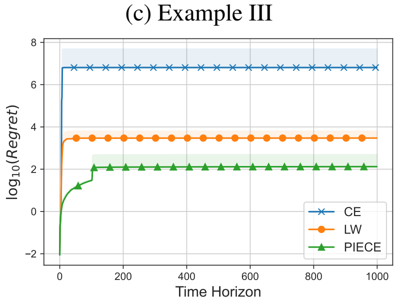
(a) Example I



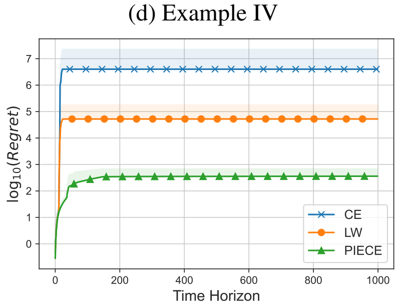
(b) Example II



(c) Example III

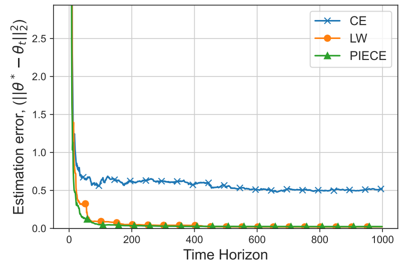


(d) Example IV

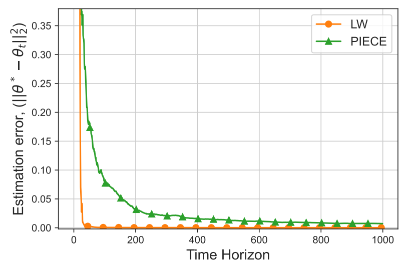


(e) Example V

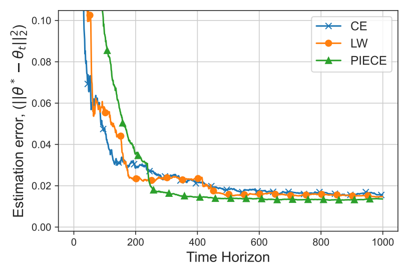
Figure 7. Log(Cumulative Regret) averaged over 50 runs for Gaussian noise with mean 0 and standard deviation 1.0.



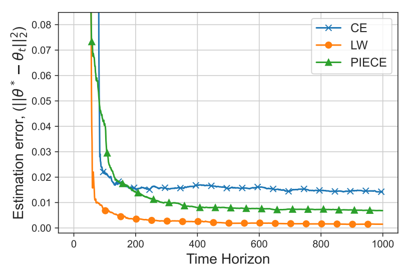
(a) Example I



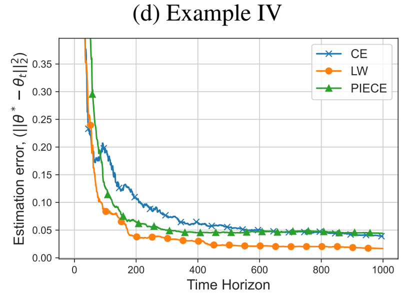
(b) Example II



(c) Example III

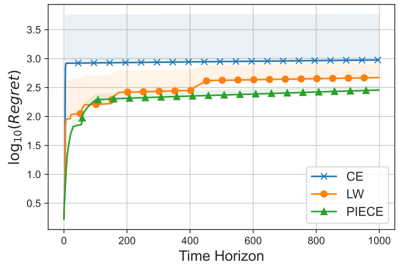


(d) Example IV

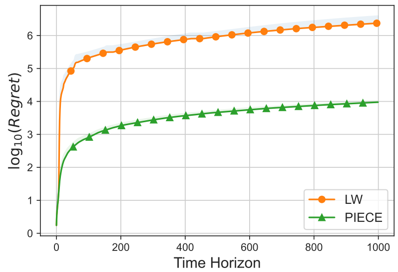


(e) Example V

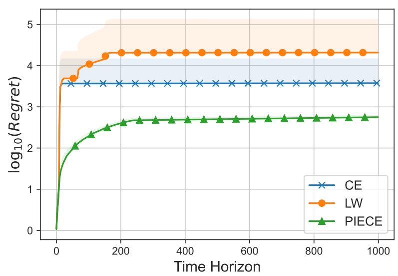
Figure 8. Estimation Error ($||\theta^\star - \theta_t||_2^2$) for Gaussian noise with mean 0 and standard deviation 1.0.



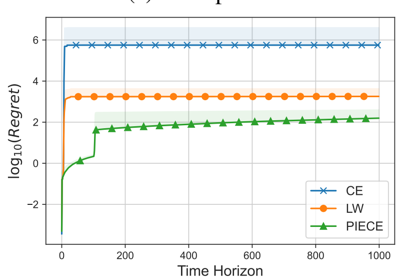
(a) Example I



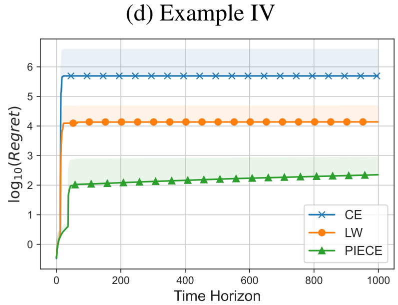
(b) Example II



(c) Example III

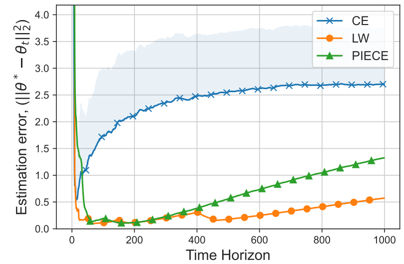


(d) Example IV

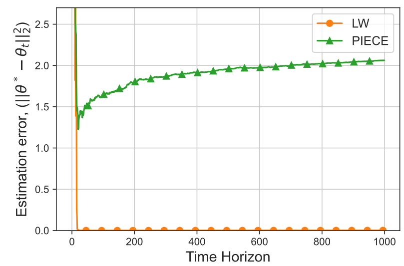


(e) Example V

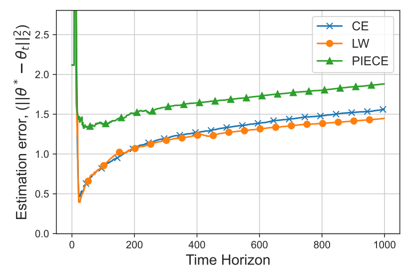
Figure 9. Log(Cumulative Regret) averaged over 50 runs (Noise: Random walk with i.i.d. Gaussian steps, $\sigma = 0.2$).



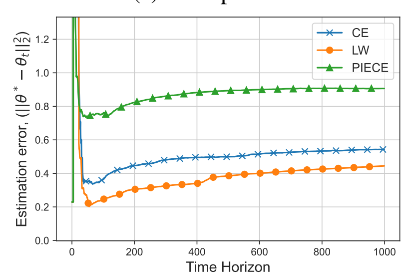
(a) Example I



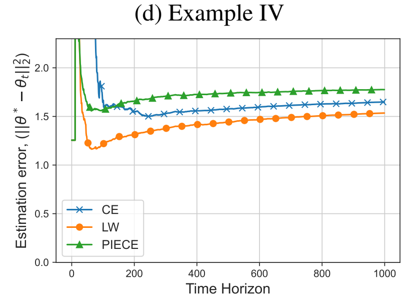
(b) Example II



(c) Example III

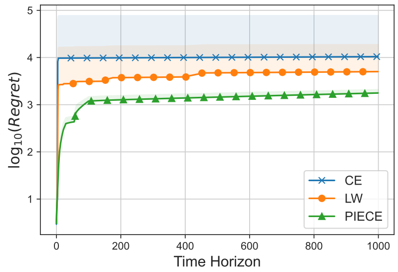


(d) Example IV

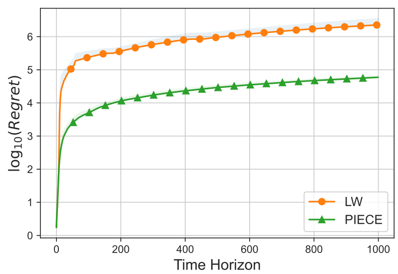


(e) Example V

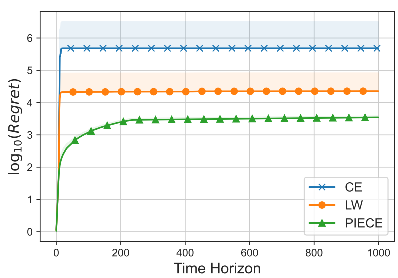
Figure 10. Estimation Error $(||\theta^{\star} - \theta_t||_2^2)$ (Noise: Random walk with i.i.d. Gaussian steps, $\sigma = 0.2$).



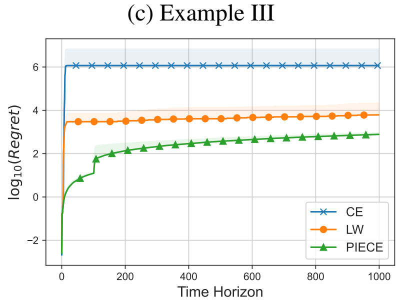
(a) Example I



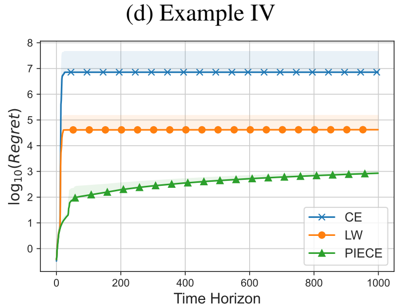
(b) Example II



(c) Example III

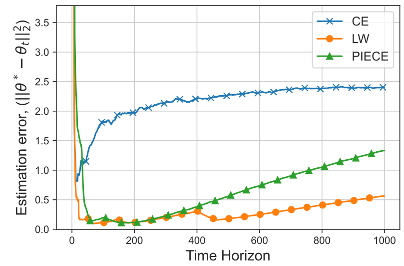


(d) Example IV

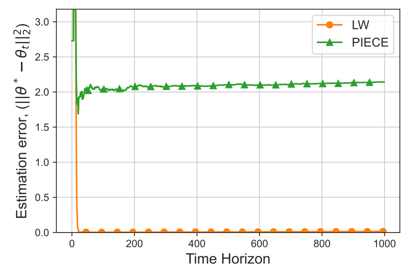


(e) Example V

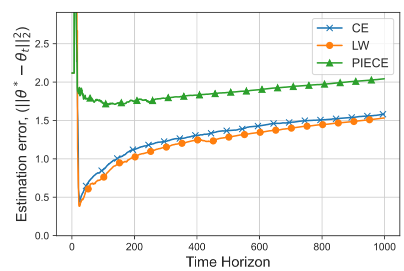
Figure 11. Log(Cumulative Regret) averaged over 50 runs (Noise: Random walk with i.i.d. Gaussian steps, $\sigma = 0.5$).



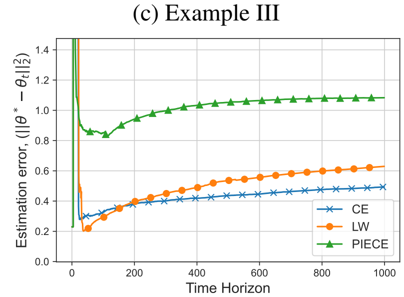
(a) Example I



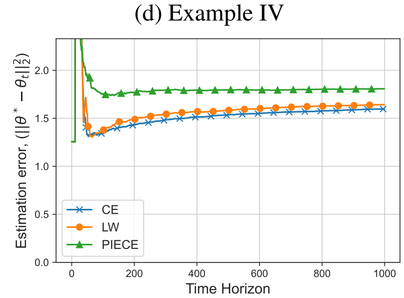
(b) Example II



(c) Example III



(d) Example IV



(e) Example V

Figure 12. Estimation Error ($||\theta^\star - \theta_t||_2^2$) (Noise: Random walk with i.i.d. Gaussian steps, $\sigma = 0.5$).

	CE	LW	PIECE
Example I	313	364	36
Example II	N/A	2460117	368
Example III	16224624	5816	66
Example IV	27935	1969	57
Example V	9888	15644	46

Table 5. Average Regret at $T = 1000$ for Gaussian noise with mean 0 and standard deviation 0.2.

	B_w	B_u	H
Example I	0.6	22876.16	45
Example II	0.6	195743858.13	85
Example III	0.6	777879.64	182
Example IV	0.6	6024.01	34
Example V	0.6	59334.86	89

Table 6. PIECE hyper-parameters for Gaussian noise with mean 0 and standard deviation 0.2.

	CE	LW	PIECE
Example I	1136155	19059	301
Example II	N/A	3115039	3259
Example III	1646089	49267	565
Example IV	24324377	2304	48
Example V	836005	53870	134

Table 7. Average Regret at $T = 1000$ for Gaussian noise with mean 0 and standard deviation 0.6.

	B_w	B_u	H
Example I	1.8	65268.34	45
Example II	1.8	558480071.75	85
Example III	1.8	2219381.4	182
Example IV	1.8	17187.19	34
Example V	1.8	186218.2	90

Table 8. PIECE hyper-parameters for Gaussian noise with mean 0 and standard deviation 0.6.

	CE	LW	PIECE
Example I	10300	22253	827
Example II	N/A	2573490	9041
Example III	6701137	28624	1562
Example IV	6442807	2972	132
Example V	4012054	52550	361

Table 9. Average Regret at $T = 1000$ for Gaussian noise with mean 0 and standard deviation 1.0.

	B_w	B_u	H
Example I	3.0	115626.85	45
Example II	3.0	989381514.38	85
Example III	3.0	3931769.53	182
Example IV	3.0	27680.15	33
Example V	3.0	299906.28	89

Table 10. PIECE hyper-parameters for Gaussian noise with mean 0 and standard deviation 1.0.

	CE	LW	PIECE
Example I	947	469	286
Example II	N/A	2368557	9478
Example III	3760	20722	563
Example IV	561426	1791	157
Example V	492781	13770	225

Table 11. Average Regret at $T = 1000$ (Noise: Random walk with iid Gaussian steps, $\sigma = 0.2$).

	B_w	B_u	H
Example I	0.6	22876.16	45
Example II	0.6	195743858.13	85
Example III	0.6	777879.64	182
Example IV	0.6	6024.01	34
Example V	0.6	59334.86	89

Table 12. PIECE hyper-parameters (Noise: Random walk with iid Gaussian steps, $\sigma = 0.2$).

	CE	LW	PIECE
Example I	10438	5030	1772
Example II	N/A	2258821	59209
Example III	481603	22630	3490
Example IV	1168033	6144	777
Example V	7172206	41842	851

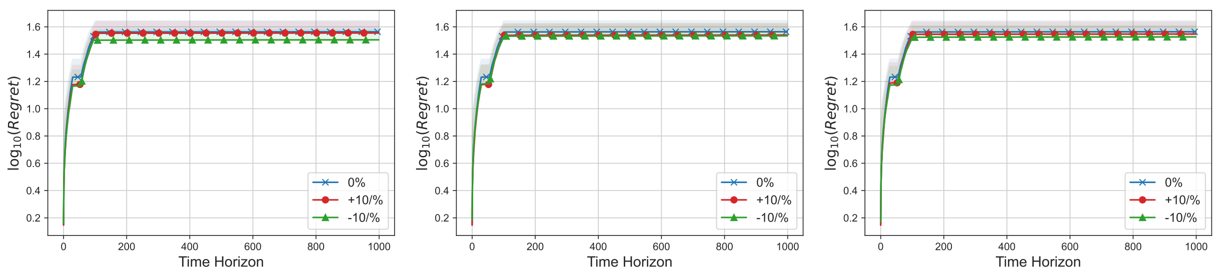
Table 13. Average Regret at $T = 1000$ (Noise: Random walk with iid Gaussian steps, $\sigma = 0.5$).

	B_w	B_u	H
Example I	1.5	53940.78	45
Example II	1.5	461553778.3	85
Example III	1.5	1834199.5	181
Example IV	1.5	14204.29	33
Example V	1.5	153899.34	90

Table 14. PIECE hyper-parameters (Noise: Random walk with iid Gaussian steps, $\sigma = 0.5$).

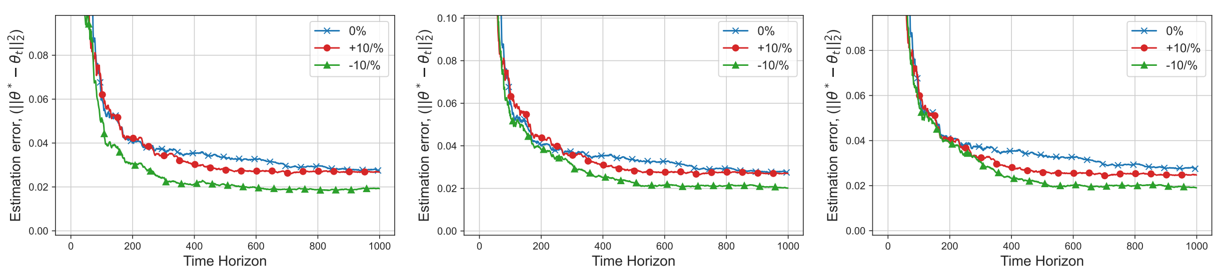
L.3. Sensitivity Analysis

We have examined the sensitivity of three hyperparameters of PIECE, H, B_u and B_w. For Example I, we have plotted the Log(Cumulative Regret) and the Estimation Error in Figs. 13 and 14, respectively, with each parameter varied $\pm 10\%$ from its calculated value keeping the other hyperparameters fixed. We created two systems by perturbing each parameter of the system in Example I with Gaussian noise of mean 0 and standard deviation 0.025. The plots for Log(Cumulative Regret) and the Estimation error of PIECE on three systems (Example I and two perturbed systems) are shown in Fig. 15.



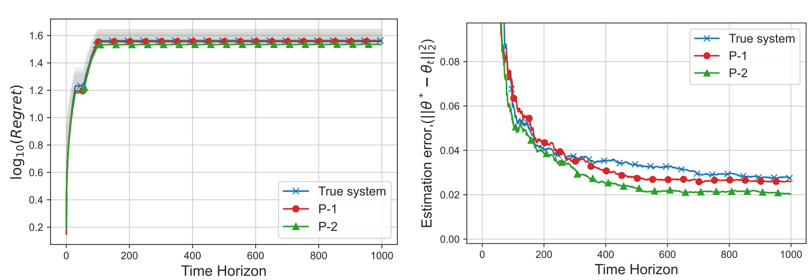
(a) Varying H (b) Varying B_u (c) Varying B_w

Figure 13. Log(Cumulative Regret) of Example I for three different values (standard and $\pm 10\%$) of each hyperparameter while keeping the other hyperparameters fixed. The standard values of H, B_u and B_w are 45, 22876 and 0.6, respectively.



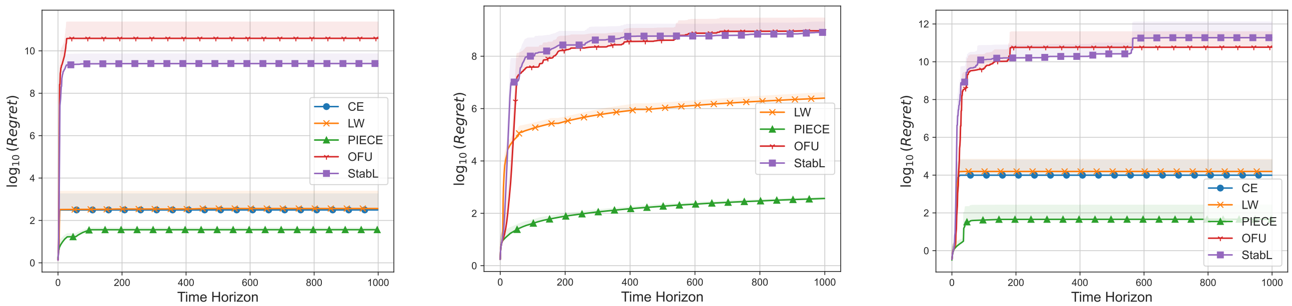
(a) Varying H (b) Varying B_u (c) Varying B_w

Figure 14. Estimation Error ($||\theta^\star - \theta_t||_2^2$) of Example I for three different values (standard and $\pm 10\%$) of each hyperparameter while keeping the other hyper parameters fixed. The standard values of H, B_u and B_w are 45, 22876 and 0.6, respectively.



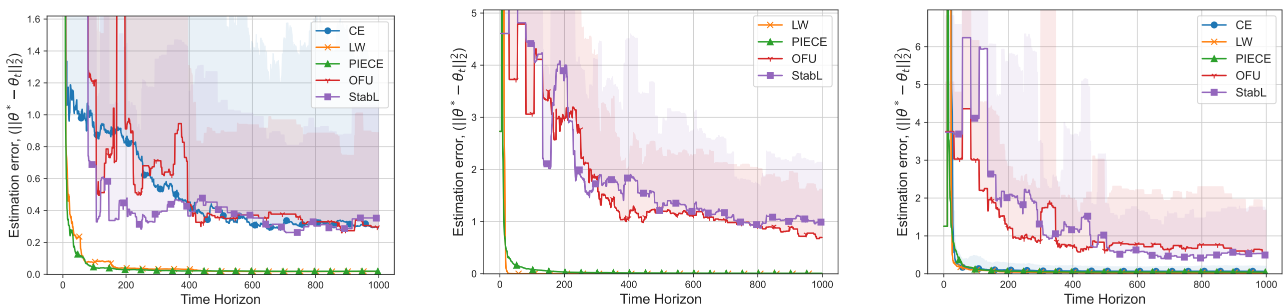
(a) Log(Cumulative Regret) (b) $||\theta^\star - \theta_t||_2^2$

Figure 15. 15a and 15b are, respectively, the Log(Cumulative Regret) plot and the Estimation error plot of PIECE algorithm for the true system of Example I and two perturbed versions of the same system.



(a) Example I (b) Example II (c) Example V

Figure 16. Log(Cumulative Regret) averaged over 50 runs (Noise: IID Gaussian noise with mean 0 and $\sigma = 0.2$).



(a) Example I (b) Example II (c) Example V

Figure 17. Estimation Error ($||\theta^\star - \theta_t||_2^2$) (Noise: IID Gaussian noise with mean 0 and $\sigma = 0.2$).

L.4. Comparison with Algorithms Originally Designed for LQ Systems

One could be inquisitive about the performance of learning algorithms originally developed for adaptive LQ problems on minimum variance control tasks, hence we demonstrate the performance of LQ algorithms for minimum variance tasks. A popular class of algorithms for LQ system are the algorithms based on the optimism principle such StabL (Lale et al., 2022), OFU (Abbasi-Yadkori & Szepesvári, 2011). These adaptive LQ algorithms involve choosing a estimate θ_t:

$$\theta_t \in \arg\min_{\theta \in C_t} J(\theta)$$

where, C_t is a high confidence parameter set and $J(\theta)$ is optimal average quadratic cost for parameter θ.

In the case of minimum variance control, $J(\theta)$ is same for all θ. Hence the optimism principle and algorithms based on it can not be directly adapted for minimum variance control. One approximate way to employ LQ learning algorithms on MV tasks is to utilize a low-norm positive definite matrix in order to weigh the control cost, in the hope that this will approximate the MV task well. We compare the performance of CE, LW, and PIECE with the following two LQ learning algorithms: OFU (Abbasi-Yadkori & Szepesvári, 2011), and StabL (Lale et al., 2022). We set the coefficient (since inputs are scalar) for the control cost to be 10^{-2}, and the noise process is IID Gaussian with $\sigma = 0.2$. The plots for Log(Cumulative Regret) and the Estimation error of all the five algorithms on Example I, Example II, and Example V are shown in Fig. 16, and Fig. 17, respectively. CE, LW, and PIECE outperform OFU and StabL by a large margin.