Non-Asymptotic Adaptive Control of Linear-Quadratic Systems

by

Mohamad Kazem Shirani Faradonbeh

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Statistics)
in The University of Michigan
2017

Doctoral Committee:

Associate Professor Ambuj Tewari, Co-Chair
Professor George Michailidis, Co-Chair
Professor Robert W. Keener
Professor Demosthenis Teneketzis

THE UNIVERSITY OF MICHIGAN

ARTES · SCIENTIA · VERITAS

1817

shirany@umich.edu

ORCID iD: 0000-0002-3807-5919

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

LQ Linear-Quadratic

LQG Linear-Quadratic-Gaussian

CE Certainty Equivalence

VAR Vector Autoregressive

OFU Optimism in the Face of Uncertainty

BOB Bet On the Best

PD Positive Definite

PSD Positive Semidefinite

pdf Probability Density Function

CDF Cumulative Distribution Function

ABSTRACT

Optimal control for the canonical model of systems with linear dynamics and quadratic operating costs (known as LQ systems) is a well-studied problem in the stochastic control literature. When the true system dynamics are unknown, an adaptive policy is required for learning the model parameters and planning a control policy simultaneously. Addressing this trade-off between accurate estimation and good control represents the main challenge in area of adaptive control. Another important issue is to prevent the system becoming destabilized (in the sense that its state grows in an uncontrolled fashion) due to lack of knowledge of the system dynamics. Asymptotically optimal approaches have been thoroughly investigated in the literature, but non-asymptotic results are few and rather incomplete. To derive such results, new concepts and technical tools need to be developed for the estimation during the stabilization period of the system.

In adaptive control, the system performance is measured by the regret, which is the difference between the cost of the adaptive policy and that of the optimal control designed according to the known dynamics. In this work, we establish non-asymptotic high probability regret bounds, which are modulo a logarithmic factor, optimal, for different LQ systems with and without identifiability assumptions. We also provide high probability guarantees for a stabilization algorithm based on random linear feedbacks. The results obtained are fairly general, since the assumptions needed are those of: (i) stabilizability of the matrices encoding the system's dynamical, and (ii) on the heaviness of the distribution for the noise vectors.

The study provides also novel results regarding the estimation of the parameters for

presumably unstable Vector Autoregressive (VAR) models. In the classical literature, there are hardly any results for the unstable case, especially regarding finite sample bounds, that is the subject of this work. Our results relate the sample size required as a function of the problem dimension and key characteristics of the true underlying transition matrix and the innovation distribution. To obtain them, appropriate concentration inequalities for random matrices and for sequences of martingale differences are leveraged.

CHAPTER 1

Introduction

In this work, we consider adaptive control of the following LQ system. Given the initial state $x(0) \in \mathbb{R}^p$, for $t = 0, 1, \cdots$ we have

$$x(t+1) = A_0 x(t) + B_0 u(t) + w(t+1), \tag{1.1}$$

$$c_t = x(t)' Q x(t) + u(t)' R u(t). \tag{1.2}$$

Above, at time t, the vector $x(t) \in \mathbb{R}^p$ is the state (and output) of the system, $u(t) \in \mathbb{R}^r$ is the control action, and $\{w(t)\}_{t=1}^{\infty}$ is the sequence of noise (disturbance) vectors. Further, c_t is the quadratic instantaneous cost function (the transpose of the vector v is denoted by v'). The dynamics of the system, i.e. both the transition matrix $A_0 \in \mathbb{R}^{p \times p}$, as well as the input matrix $B_0 \in \mathbb{R}^{p \times r}$, are fixed and *unknown*. The positive definite matrices of the cost, $Q \in \mathbb{R}^{p \times p}, R \in \mathbb{R}^{r \times r}$, are however known.

The goal is, roughly speaking, designing a control policy $\{u(t)\}_{t=0}^{\infty}$ in order to minimize the expected average cost. Conceptually, in order to preserve causality, the design of the control action at every time needs to be according to the observations so far. This objective will be formally expressed later in Section 2.1. Designing an efficient adaptive policy is challenging, since it requires to both estimate the unknown true matrices A_0, B_0, as well as design a control policy accordingly. In fact, the exact knowledge of the true parameters A_0, B_0 is required in order to design an optimal control policy, while on the other hand, the user needs to deal with the system by applying some control action, to collect the observations for the estimation of the unknown parameters.

1.1 Asymptotic Literature

The asymptotic analysis of adaptive control for systems evolving according to linear dynamics is a canonical problem in the classical literature. Since the dynamics of the system is unknown, a natural way to design the control action is the certainty equivalence (CE) approach [1]. Intuitively, its prescription is to apply a control policy as if the estimated parameter is the true one the system is evolving according to. It was soon realized that the least-squares estimation does not need to be generally consistent, even if the open-loop system is stable [2]. Later, it was shown for stochastic approximation algorithm, that convergence to incorrect parameter occurs with positive probability [3].

Bypassing the consistency, an extensive amount of the classical literature is devoted to address the problem of adaptive tracking where the objective is to adaptively steer the system to track a reference trajectory. For open-loop stable systems, assuming the reference signal is bounded, a sharp regret bound is provided under the uniform boundedness of the noise terms [4]. Namely, a conservatively defined regret is shown to be of logarithmic order, which is optimal [5]. Later, convergence rates were established for the more general case that the noise is not necessarily bounded [6]. Ensuing works addressed the problem for tracking both a reference trajectory as well as a reference model [7]. The stability assumption of open-loop system can be removed by using a stochastic gradient algorithm parallel to the least-squares, in order to slow down the possible explosion of the system [7], [8].

The asymptotic results for optimality of tracking performance are general, although, the results with regards the consistency issue are fairly restrictive. More precisely, in order to ensure the consistency,

(i) either an additional identifiability assumption [4], [7],

(ii) or a minor deviation from the straightforward CE approach,

is required. The latter case consists of adding a random (independent diminishing) perturbation to the control signal [2], [9]. Even if the noise is assumed to be a white normal process, consistency is still a persistent issue, as the convergence is not necessarily to the true parameter [10]. Thus, despite the non-adaptive version of the problem in stochastic control where output-observation and control-design completely decouple in Linear-Quadratic-Gaussian (LQG) systems [11], it is not the case for consistency the adaptive version. Importantly, what really prevents the guarantee of consistency, is the reduction of open-loop identification to *"closed-loop identification"* [10].

In adaptive tracker type of approaches discussed above, the operating cost is not directly a function of the control signal. Unlikely, when it is the case, the cost of an adaptive policy designed based on certainty equivalence can be strictly non-optimal [12]. The intuition behind both this non-optimaliy, as well as the reduction mentioned above, is intrinsically of the type of the usual exploration-exploitation dilemma. The aforementioned methods (i), (ii) are basically examples of the attempts to address this dilemma. Technically, applying any control policy, the observations lead to accurate information about the unknown parameters of the dynamics, only in some specific directions [12]. Hence, to gather comprehensive information about the open-loop parameter, different control actions are required to be applied. On the other hand, to avoid major deviations from the optimal value in the cost, control actions need to be in some sense similar.

Therefore in general, neither cost optimality, nor consistency will automatically be provided by certainty equivalence. Still, asymptotic cost optimality results in the literature hold, mainly because sufficient assumptions are imposed to enable the planning of an optimal adaptive policy regardless of a consistent learning of the parameter. In addition to *control-free costs*, these assumptions are for instance non-singularity of the true input matrix [9], and lack of common factors [7]. For example, there exist situations where the estimation is almost surely inconsistent, while because it is asymptotically in-line with the true parameter [3], optimality is automatically gifted, needless to consistency. Further, if the instantaneous cost is control-free, closed-loop identification suffices for asymptotic optimality of the average cost [10]. But, for general quadratic costs, availability of an approximation which is in-line with the true parameter, as well as full identification of the closed-loop, are insufficient to design an optimal policy.

While direct application of the least-squares estimation fails to provide an optimal adaptive policy, a modification resolves the issue [12]. In fact, it suffices to use the following idea of *Optimism in the Face of Uncertainty* (OFU), which was originally invented for efficient allocation rules [13]. After constructing a confidence set, the prescription is to "Bet On the Best (BOB)" [14], i.e. to design a control action as if the most optimistic parameter in the confidence set is the true one.

1.2 Non-Asymptotic Literature

Recently, the non-asymptotic approach to adaptive control of LQ systems has been taken first in the work of Abbasi-Yadkori and Szepesvári [15]. The authors provide an adaptive control algorithm, for which the regret is shown to be optimal, apart from a logarithmic factor. In the regret bounds presented in the above paper, there exist constants scaling exponentially with respect to the dimension. This, motivated the second paper due to Ibrahimi et al. [16], which attacks the problem in the high dimensional regime, assuming the true dynamics matrices are sparse. The latter paper also shows that the presented reinforcement learning algorithm leads to an optimal regret bound, apart from a logarithmic factor.

The aforestated works consider a fairly restricted setting, which requires strong assumptions. A concrete example to demonstrate these restrictions will be discussed in the next chapter. There are two assumptions both recent papers share, and the analysis of the presented algorithms fails without. First, controllability and observability are assumed for the true dynamics matrices of the system. Second, the closed-loop transition matrix when the optimal linear feedback is applied to the system (see Proposition 2.1), is assumed to have the operator norm (denoted here by $|||\cdot|||_2$) less than one. Note that the first assumption does not imply the second one because the closed-loop dynamics matrix is only known to be stable, assuming the system is controllable and observable [17].

We relax both of these assumptions to *stabilizability* (see Definition 2.1), which is minimal. More precisely, the problem becomes trivial if one does not assume stabilizability of the system. There are straightforward situations where one can simply see that the strong assumptions of the above papers fail. For example, the first restriction above will be violated when the true dynamics matrices are too sparse that controllability fails, while the system is still stabilizable. The sparsity of the true dynamics matrices is specially common in large systems. Note that, it is exactly the case where the high dimensional setting of the second paper makes sense. The other restriction is even more serious, since a randomly chosen stable matrix does not need to satisfy the operator norm condition mentioned above. As a matter of fact, the class of systems for which the closed-loop transition matrix has operator norm less than one is significantly smaller than the family of stabilizable systems.

Furthermore, the important issue of stabilization is fully overlooked in the second paper. Technically, the constants scaling exponentially with the dimension appear in the regret bounds of the first paper, mainly because of the transient period the system needs to spend, in order to gradually stabilize itself. The second paper bypasses the stabilization period by assuming that a linear feedback which stabilizes the system is automatically provided to the user at the beginning of the interaction with the system. Of course, the way the first paper is addressing the self stabilization of the system, is mainly based on the operator norm restriction we discussed before.

1.3 Contributions

In addition to resolving all aforementioned issues, we generalize the high probability regret bounds to an extensively larger class of noise distributions. Technically, in the papers discussed above, the noise vectors are assumed to be sub-Gaussian or Gaussian, respectively, and the coordinates of the noise vectors are assumed to be uncorrelated. In this work, we assume a sub-Weibull distribution for the noise vectors, and that coordinates can be correlated.

Based on OFU, we provide non-asymptotic regret bounds for a class of adaptive control policies, for a remarkably extensive family of Linear-Quadratic systems. Namely, we prove that the reinforcement learning algorithms presented in Chapter 3 are with high probability near optimal, under the minimal assumption of stabilizability, and heavy-tailed distribution for the noise process. Note that unlike a rich literature providing asymptotic results for the problem at hand, non-asymptotic analysis is sparse with few results. In order to study the finite time behavior of adaptive policies, new conceptual and technical approaches need to be developed.

First, from an exploitation (i.e. cost optimality) viewpoint, whenever non-asymptotically studying a policy, all involved quantities need to be carefully examined. More precisely, one needs to provide a scalable specification of the decay-rate, for the terms (such as the gap between the average and the expected value of a sequence of ransom variables) which vanish in asymptotic. In asymptotic regime however, it is only required to verify that the expressions vanish, if normalized by the leading term. For example, the objective in the aforementioned asymptotic literature is to show that the average cost of the adaptive policy under study converges to the optimal expected average cost of the system. These type of results, lead to sub-linearity (with respect to the time horizon) for the accumulative deviation of the adaptive cost from the optimal value. In the finite time analyses presented in this work, we establish fairly stronger statements, and show that this accumulative deviation approximately scales as the square-root of the time horizon.

Second, from the exploration (i.e. consistency) viewpoint, it is well known that compared to an infinite sample setting, estimation results are essentially harder to achieve when the sample size is finite. For instance, one can immediately apply the Law of Large Numbers to ensure the asymptotic convergence of random matrices. In spite of that, to ensure the high probability convergence of finitely many random matrices, more advanced tools such as concentration inequalities are required. Besides, the theory of concentration inequalities is mainly based on (moment) generating functions which do not necessarily exist for heavy-tailed noise process of this work. So, an extra layer of technicality is necessary to achieve the useful estimation results. Finally, it is needless to mention that from practical viewpoint, the actual horizon is always finite in real world problems.

We also address the important issue of stabilizing a system with unknown dynamics, establishing high probability guarantees for the presented finite time stabilization algorithm. Key estimation results concerning the finite sample learning of the unknown dynamics in both stable and *unstable* dynamical systems are being leveraged to analyze the stabilization algorithm, as well as the adaptive control algorithms. Finite sample analysis of estimation in a more general setting is comprehensively discussed in Chapter 4.

This work is structured as follows. First, in Section 2.1 we rigorously formulate the adaptive control problem for LQ systems. As mentioned before, planning issues, as well as those of learning, need to be addressed. Therefore, the problem from a pure control viewpoint is discussed in Section 2.2, where we investigate the properties of the optimal policies. The estimation approach is the content of Section 2.3, where we establish the key estimation results for vector autoregressive processes as the cornerstone of the algorithms presented in this work. On one hand, in order to stabilize the system, a stabilization algorithm is proposed and analyzed in Section 2.4. On the other hand, once the system is stabilized, an adaptive policy is required to minimize the expected average cost. For this purpose, we present different reinforcement learning algorithms in Chapter 3.

Subsequently, we provide the detailed structures of the contributed results. In order to

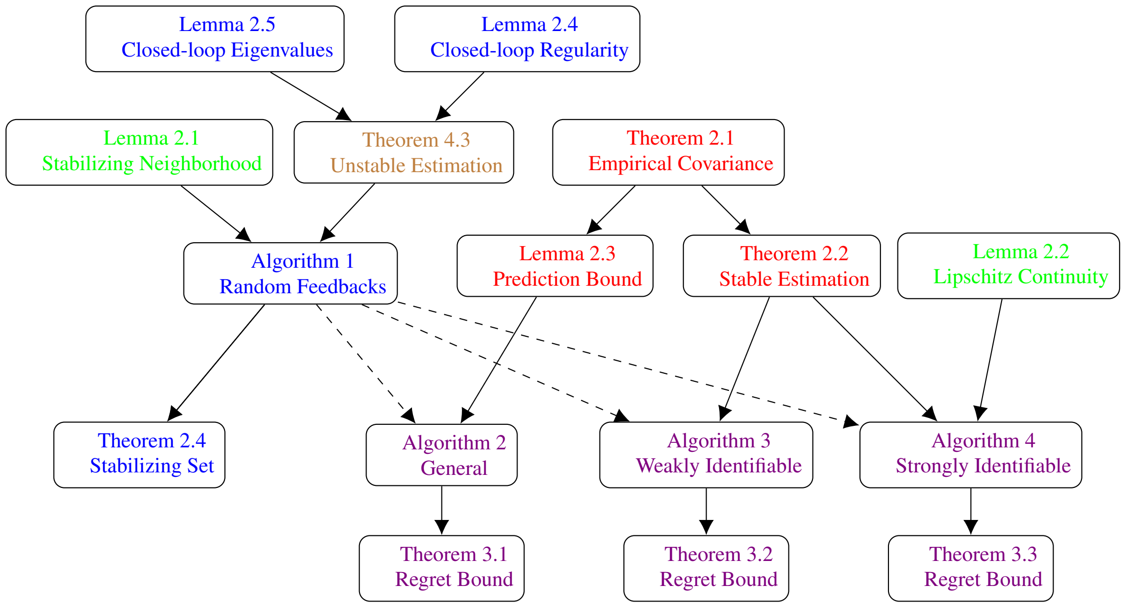

Figure 1.1: Diagram of the contributed results. Sections 2.2, 2.3, and 2.4, are shown respectively by green, red, and blue. Chapters 3 and 4 are shown by violet and brown, respectively.

stabilize the system in finite time, Lemma 2.1 states that the system can be stabilized if one can find a stabilizing neighborhood of the unknown parameters A_0, B_0. Then, according to the estimation results of Section 2.3, the high probability accurate estimation of the closed-loop transition matrix is possible. In fact, by Theorem 4.3, it suffices for the unstable closed-loop transition matrix

(i) to be regular, and

(ii) having no eigenvalue of unit size.

Regularity basically is related to the explosive eigenvalues of the closed-loop transition matrix (i.e. eigenvalues of magnitude larger than one). In fact, as formally defined in Definition 2.3, the geometric multiplicity of every explosive eigenvalue of a regular matrix is one.

Then, applying a random linear feedback, Lemma 2.4 shows that condition (i) is satisfied, while condition (ii) is implied by Lemma 2.5. Putting all together, Section 2.4 provides Theorem 2.4, which states that **Algorithm 1** returns a high probability stabilization set. The diagram of this logical structure is illustrated on the left-side of Figure 1.1.

The next objective, namely design of an adaptive policy to minimize the regret (formally defined in (2.1)), is addressed in Chapter 3. For this purpose, once the system is stabilized (e.g. using Algorithm 1), different reinforcement learning algorithms can be applied. These algorithms, are designed according to the useful estimation and prediction results regarding the stable Vector Autoregressive (VAR) process developed in Section 2.3. More precisely, using the analysis of the empirical covariance matrix of the VAR processes provided by Theorem 2.1, we prove Lemma 2.3. The former, presents a high probability prediction bound, and is used to design **Algorithm 2**. Theorem 3.1 states that the adaptive policy designed by the reinforcement learning scheme of Algorithm 2 leads to a high probability bound for the regret, which is optimal, apart from a logarithmic factor.

The other estimation result, Theorem 2.2, shows that the transition matrix of a stable VAR process can be accurately estimated with high probability. This is what **Algorithm 3** for weakly identifiable systems is leaning on. The weak identifiability condition, formally defined in Definition 3.1, holds, when for an approximation of A_0, B_0, different linear feedbacks yield closed-loop matrices of comparable precision. Theorem 3.2 states the high probability near optimal regret bound of Algorithm 3 under the weak identifiability condition. Both Algorithm 2 and Algorithm 3, are designed using the idea of Optimism in the Face of Uncertainty (OFU) principle, aka Bet On the Best (BOB) [14]. Essentially, BOB prescribes an adaptive control action which is designed according to an optimistic approximation of the true parameters.

Finally, when side information about the true parameter leads to strong identifiability, one can apply **Algorithm 4**. When a system is strongly identifiable, as rigorously presented in Definition 3.2, an accurate approximation of the closed-loop matrix leads to that of the open-loop parameters. In this case, the step based on the OFU principle can be removed from the reinforcement learning algorithm, and the regret is again near optimal, as presented in Theorem 3.3. The analysis is using Theorem 2.2, as well as Lemma 2.2 which is regarding the Lipschitz continuity of the optimal expected average cost with respect to the dynamics parameters. Figure 1.1 describes the structure of the pieces of this work.

1.4 Notations

The following notations will be used throughout this thesis. For matrix $A \in \mathbb{C}^{p\times q}$, A' is the transpose of A. When $p = q$, the smallest (respectively largest) eigenvalue of A (in magnitude) is shown by $\lambda_{\min}(A)$ (respectively $\lambda_{\max}(A)$) and the trace of A is shown by $\mathrm{tr}\,(A)$. For

$$\gamma \in \mathbb{R}, \gamma > 0, x \in \mathbb{C}^q,$$

γ-norm of vector x is

$$\|x\|_\gamma = \left(\sum_{i=1}^{q} |x_i|^\gamma\right)^{1/\gamma}.$$

Further, when $\gamma = \infty$, the norm is defined according to $\|x\|_\infty = \max_{1\le i\le q} |x_i|$.

We also use the following notation for operator norm of matrices. For $\beta, \gamma \in (0, \infty]$, and $A \in \mathbb{C}^{p\times q}$, define

$$|||A|||_{\gamma\to\beta} = \sup_{v\in\mathbb{C}^q\setminus\{0\}} \frac{\|Av\|_\beta}{\|v\|_\gamma}.$$

Whenever $\gamma = \beta$, we simply write $|||A|||_\beta$. To show the dimension of manifold $\mathcal{M}$ over the field F, we use $\dim_F(\mathcal{M})$. Finally, the sigma-field generated by random vectors $X_1, \cdots, X_n$ is denoted by $\sigma(X_1, \cdots, X_n)$.

CHAPTER 2

Optimality, Estimation, and Stabilization

2.1 Introduction

Now, we formally discuss the adaptive control problem this work is addressing. The evolution of the system is governed by the linear dynamics (1.1), and the instantaneous quadratic cost c_t is defined according to (1.2). The true dynamics is assumed to be stabilizable, as defined below.

Definition 2.1 (Stabilizability)**.** $[A_0, B_0]$ is called stabilizable if there is $L \in \mathbb{R}^{r \times p}$ such that

$$|\lambda_{\max} (A_0 + B_0 L)| < 1.$$

The linear feedback matrix L is called a stabilizer.

Remark 2.1. For convenience, henceforth for $A \in \mathbb{R}^{p \times p}$, $B \in \mathbb{R}^{p \times r}$, we use $[A, B]$ and $\Theta \in \mathbb{R}^{p \times q}$ interchangeably, where $q = p + r$.

Next, the following example provides a situation where the system is not controllable, and the operator norm of the closed-loop dynamics can not be less than one, while the stabilizability assumption still holds.

Example 2.1. Let the dynamics matrices be

$$A_0 = \begin{bmatrix} 0.5 & 2 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 3 \end{bmatrix}, B_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & -1 \end{bmatrix}.$$

Computing Kalman's controllability matrix, we get

$$\left[B_0, A_0 B_0, A_0^2 B_0\right] = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 3 & -3 & 9 & -9 \end{bmatrix},$$

which clearly is not full rank, i.e. the system is not controllable [17].

In addition, for an arbitrary linear feedback $L = [L_{ij}] \in \mathbb{R}^{2\times 3}$, the closed-loop transition matrix is

$$A_0 + B_0 L = \begin{bmatrix} 0.5 & 2 & 0 \\ 0 & 0.5 & 0 \\ L_{11} - L_{12} & L_{12} - L_{22} & 3 + L_{13} - L_{23} \end{bmatrix},$$

which is stable if and only if $|3 + L_{13} - L_{23}| < 1$, while the operator norm condition cannot be satisfied because

$$|||A_0 + B_0 L|||_2 \geq 2,$$

for all L.

In the stochastic dynamics of the system presented in (1.1), $\{w(t)\}_{t=1}^{\infty}$ are independent mean-zero homoscedastic noise vectors with full rank covariance matrix C:

$$\mathbb{E}\left[w(t)\right] = 0, \;\; \mathbb{E}\left[w(t)w(t)'\right] = C, \quad \left|\lambda_{\min}\left(C\right)\right| > 0.$$

The tail behavior of every coordinate of the noise vector satisfies the condition below.

Assumption 2.1 (Sub-Weibull noise distribution)**.** There exist positive constants b_1, b_2, and α, such that

$$\mathbb{P}\left(|w_i(t)| > y\right) \leq b_1 \exp\left(-\frac{y^\alpha}{b_2}\right),$$

for all $t = 1, 2, \cdots; i = 1, \cdots, p; y > 0$.

Note that assuming a sub-Weibull distribution for the noise coordinates is more general than the sub-Gaussian assumption routinely made in the literature, where $\alpha \geq 2$, as well as sub-Exponential, where $\alpha \geq 1$. In fact, when $\alpha < 1$, the noise coordinates $w_i(t)$ do not need to have a moment generating function. Therefore, as mentioned before, concentration inequalities of random matrices, which are the foundation of non-asymptotic statistical analyses, cannot be applied directly. Furthermore, the noise coordinates can be either discrete or continuous random variables, and are not assumed to have a probability density function (pdf). Henceforth, the special case of bounded noise can be obtained from the presented results, by simply letting $\alpha \to \infty$.

Remark 2.2. The results established also hold if the noise vectors are martingale difference sequences.

The rigorous formulation of the objective is as follows. For an arbitrary control policy $\{u(t)\}_{t=0}^{\infty}$, let $\overline{\mathcal{J}}_{A_0,B_0}\left(\{u(t)\}_{t=0}^{\infty}\right)$ be the expected average cost of the system:

$$\overline{\mathcal{J}}_{A_0,B_0}\left(\{u(t)\}_{t=0}^{\infty}\right) = \limsup_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}c_t,$$

where c_t is the resulting instantaneous cost, when the policy $\{u(t)\}_{t=0}^{\infty}$ is applied. Above, the dependence of $\overline{\mathcal{J}}_{\Theta_0}(\cdot)$ to the known cost matrices Q, R is suppressed. Moreover, stabilizability of Θ_0 is a minimal assumption, since otherwise, the instantaneous cost c_t will explode, leading to the trivial situation $\overline{\mathcal{J}}_{\Theta_0}(\cdot) = \infty$. Then, among all control policies, an optimal policy is one which minimizes the expected average cost. Note that due to the independence of the noise vectors, an optimal policy can be assumed to be causal (i.e. the control action can not depend on the future observations). Besides, due to Markovity of the state process $\{x(t)\}_{t=0}^{\infty}$, the control can assumed to be memoryless (i.e. for all $t = 0, 1, \cdots$, $u(t)$ is a function of only the current state $x(t)$).

The optimal expected average cost is defined by

$$\mathcal{J}^{\star}(A_0, B_0) = \min_{\{u(t)\}_{t=0}^{\infty}} \overline{\mathcal{J}}_{A_0,B_0}\left(\{u(t)\}_{t=0}^{\infty}\right),$$

where the minimum is taken over *all* control policies $\{u(t)\}_{t=0}^{\infty}$, including non-adaptive policies which are based on the known parameter Θ_0. Since the evolution of the system is stochastic, an open-loop control policy can not achieve the optimal average cost. A linear feedback however, can provide the optimal policy (see Proposition 2.1). As one can expect, the dynamics matrices A_0, B_0 need to be exactly known, in order to find the optimal linear feedback.

For an adaptive control policy though, dynamics matrices A_0, B_0 are unknown. Therefore, the user requires to simultaneously learn the dynamics, and plan a control policy. Here we assume **perfect observation**, i.e. the output of the system provides the state vector accurately. In other words, the sequence of the states $\{x(t)\}_{t=0}^{n}$ is fully observed during the period the user is dealing with the system. So, in a period of length n, the user applies an adaptive policy $\{u(t)\}_{t=0}^{n-1}$ to the system, and observes the sample $\{x(t)\}_{t=0}^{n}$ to estimate A_0, B_0 accordingly.

In order to measure the quality of an adaptive policy, the resulting cost will be compared by the optimal expected average cost defined above. More precisely, for adaptive policy $\{u(t)\}_{t=0}^{\infty}$, letting c_t be as (1.2), the comparison between adaptive control policies is made by *regret*, which at time T is defined formally as

$$\mathcal{R}(T) = \sum_{t=1}^{T} \left[c_t - \mathcal{J}^{\star}(A_0, B_0)\right]. \tag{2.1}$$

The regret is basically nothing but the accumulative deviation of the instantaneous cost of the corresponding adaptive policy from the optimal expected average cost (which is based on a non-adaptive policy).

2.2 Optimal Policies

In this section, we investigate the properties of optimal policies. For general $\Theta \in \mathbb{R}^{p \times q}$, it is well known from classical literature that to achieve $\mathcal{J}^{\star}(\Theta)$, one has to solve *Riccati* equation,

$$
\begin{aligned}
K(\Theta) &= Q + A'K(\Theta)A - A'K(\Theta)B\left(B'K(\Theta)B + R\right)^{-1}B'K(\Theta)A, && (2.2)\\
L(\Theta) &= -\left(B'K(\Theta)B + R\right)^{-1}B'K(\Theta)A. && (2.3)
\end{aligned}
$$

A solution, is a positive semidefinite matrix $K(\Theta)$ satisfying (2.2).

For the sake of completeness, we present and prove the following proposition, which provides an optimal linear feedback to minimize the expected average cost. Moreover, it establishes the existence and uniqueness of the solution of Riccati equation supposing stabilizability.

Proposition 2.1 (Optimal policy)**.** If $[A_0, B_0]$ is stabilizable, (2.2) has a unique solution and $u(t) = L(\Theta_0)\,x(t)$ is an optimal control policy leading to

$$\mathcal{J}^{\star}(\Theta_0) = \operatorname{tr}\left(CK(\Theta_0)\right).$$

Conversely, if $K(\Theta_0)$ is a solution of (2.2), $L(\Theta_0)$ defined by (2.3) is a stabilizer.

Note that in the latter case of Proposition 2.1, the solution $K(\Theta_0)$ is unique and $u(t) = L(\Theta_0)\,x(t)$ is an optimal policy which yields $\mathcal{J}^{\star}(\Theta_0) = \operatorname{tr}(CK(\Theta_0))$.

The next result describes the asymptotic distribution of the regret. In general, since the state of the system $x(t)$, and so the instantaneous cost c_t, are random, $\mathcal{R}(T)$ can not be bounded as T grows. Proposition 2.2, which is basically a Central Limit Theorem for $\mathcal{R}(T)$, states that even if the control action $\{u(t)\}_{t=0}^{\infty}$ is the optimal policy $u(t) = L(\Theta_0)\, x(t)$ described above according to the known dynamics matrices A_0, B_0, the regret $\mathcal{R}(T)$ scales as $O\left(T^{1/2}\right)$.

Proposition 2.2 (Regret lower-bound)**.** Applying optimal control action $u(t) = L(\Theta_0)\, x(t)$, the distribution of $\lim\limits_{T\to\infty} \frac{\mathcal{R}(T)}{T^{1/2}}$ is a mean-zero normal.

An immediate consequence is that the followings hold when applying the optimal linear feedback $u(t) = L(\Theta_0)\, x(t)$:

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[c_t\right] = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} c_t = \mathcal{J}^{\star}(\Theta_0).$$

Moreover, Proposition 2.2 provides a lower bound for the regret of the adaptive policies. Namely, a high probability regret bound to hold with probability at least $1-\delta$, needs to satisfy

$$\mathcal{R}(T) \geq O\left(T^{1/2}\left(-\log\delta\right)^{1/2}\right).$$

Note that the definition of regret in (2.1) covers the accumulative deviation from the optimal expected average cost due to both the stochastic evolution of the system (randomness of $\{w(t)\}_{t=1}^{\infty}$), as well as the uncertainty about the dynamics (unknownness of Θ_0). For example, if one defines the regret as the difference between the instantaneous cost (c_t) of the *adaptive* policy and that of the *non-adaptive optimal* one $u(t) = L(\Theta_0)\, x(t)$, then the regret vanishes by applying the optimal linear feedback $L(\Theta_0)$. The latter definition for regret, takes into account only the fraction due to lack of full knowledge about the dynamics. But, the take home message of Proposition 2.2 is that from pure control point of view, the convergence of accumulative cost is with the rate $O\left(T^{1/2}\right)$. So, trying to push the second fraction of the regret (which is due to learning of the unknown dynamics) to have a rate less than $O\left(T^{1/2}\right)$ is actually unnecessary.

To proceed, we define a notations, helpful to simplify the expressions throughout this chapter. For arbitrary stabilizable $\Theta_1, \Theta_2 \in \mathbb{R}^{p \times q}$, let

$$\tilde{L}(\Theta_1) = \begin{bmatrix} I_p \\ L(\Theta_1) \end{bmatrix} \in \mathbb{R}^{q \times p}.$$

Therefore,

$$\Theta_2 \tilde{L}(\Theta_1) = A_2 + B_2 L(\Theta_1).$$

The adaptive control policy $\{u(t)\}_{t=0}^{\infty}$ is to be defined without knowing Θ_0. Because according to Proposition 2.1 an *optimal* policy is to apply the linear feedback $L(\Theta_0)$, a candidate *adaptive* policy is a linear feedback of the form $L\left(\tilde{\Theta}^{(1)}\right)$, where $\tilde{\Theta}^{(1)}$ is an approximation of the true parameter Θ_0, which is learned when the system evolves.

In the reminder of this section, results concerning properties of these type of policies will be provided, which will be helpful later to analyze the performance of the algorithms of Chapter 3. The first issue is stability of the system (which evolves according to Θ_0), when a linear feedback of the form $L\left(\tilde{\Theta}^{(1)}\right)$ is applied. To address that, existence of a stabilizing neighborhood is established in the following lemma.

Lemma 2.1 (Stabilizing neighborhood)**.** There is $\epsilon_0 > 0$, such that for every stabilizable Θ, if

$$|||\Theta - \Theta_0|||_2 < \epsilon_0,$$

then $\Theta_0 \tilde{L}(\Theta)$ is stable.

Next, the following lemma shows the Lipschitz continuity of matrix $K(\Theta)$ defined in (2.2), with respect to Θ. Note that a direct consequence of Lemma 2.2 is Lipschitz continuity of $L(\Theta)$ and $\mathcal{J}^\star(\Theta)$, respectively using (2.3) and Proposition 2.1.

Lemma 2.2 (Lipschitz continuity)**.** Assume $\Theta_1, \Theta_2 \in \mathbb{R}^{p \times q}$ are stabilizable. There is a constant $\Gamma_K < \infty$, such that

$$|||K(\Theta_1) - K(\Theta_2)|||_2 \leq \Gamma_K |||\Theta_1 - \Theta_2|||_2.$$

Furthermore, there are $\Gamma_L, \Gamma_{\mathcal{J}} < \infty$, such that

$$\begin{aligned} |||L(\Theta_2) - L(\Theta_1)|||_2 &\leq \Gamma_L |||\Theta_2 - \Theta_1|||_2, \\ |\mathcal{J}^\star(\Theta_2) - \mathcal{J}^\star(\Theta_1)| &\leq \Gamma_{\mathcal{J}} |||\Theta_2 - \Theta_1|||_2. \end{aligned}$$

2.3 Estimation

When applying a linear feedback, denoted by $L \in \mathbb{R}^{r \times p}$, to the system, the closed-loop dynamics will be a Vector Autoregressive (VAR) process of the form

$$x(t+1) = Dx(t) + w(t+1),$$

where $D = A_0 + B_0 L$. Moreover, since designing a stabilizer L may be impossible without knowing a neighborhood of Θ_0 (Lemma 2.2), matrix D does not need to be stable. Therefore, studying the performance of estimation for such a process is of interest. The non-asymptotic analysis of unstable VARs under a more general setting, is the subject of Chapter 4.

We present bounds on the number of observations to have arbitrarily small estimation error, with high probability. Results of this section will be used later in Section 2.4, and Chapter 3, to construct a stabilizing neighborhood, as well as design of adaptive control algorithms. First, we define row-wise least-squares estimation, for matrix D, as follows. Observing samples $\{x(t)\}_{t=0}^{n}$, define the sum-of-squares loss function

$$\mathcal{L}_n^{(i)}(\theta) = \sum_{t=0}^{n-1} \left(x_i(t+1) - \theta' x(t)\right)^2.$$

Then, the true transition matrix D is estimated by

$$\hat{D}_n = \left[\hat{d}_1, \cdots, \hat{d}_p\right]',$$

where for $i = 1, \cdots, p$, the vector $\hat{d}_i$ is a minimizer of the above sum-of-squares, i.e.

$$\mathcal{L}_n^{(i)}\left(\hat{d}_i\right) = \min_{\theta \in \mathbb{R}^p} \mathcal{L}_n^{(i)}(\theta).$$

In the sequel, after introducing some notations, we study performance of the estimation $\hat{D}_n$, first when the true transition matrix D is stable, and then, when it is unstable.

Next, for $\lambda \in \mathbb{C}$, the following matrix is called the size m *Jordan matrix of* λ.

$$\begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{m \times m}.$$

Then, for matrix $D \in \mathbb{R}^{p \times p}$, we define $\eta(D)$ as follows.

Definition 2.2 (Constant $\eta(D)$)**.** Let the Jordan decomposition of D be $D = P^{-1}\Lambda P$, i.e. Λ is block diagonal,

$$\Lambda = \operatorname{diag}(\Lambda_1, \cdots, \Lambda_k),$$

where for $i = 1, \cdots, k$, $\Lambda_i \in \mathbb{C}^{m_i \times m_i}$ is a Jordan matrix of λ_i. For $t = 1, 2, \cdots$, letting

$$\eta_t(\Lambda_i) = \inf_{\rho \geq |\lambda_i|} t^{m_i - 1} \rho^t \sum_{j=0}^{m_i - 1} \frac{\rho^{-j}}{j!},$$

define $\eta_t(\Lambda) = \max_{1 \leq i \leq k} \eta_t(\Lambda_i)$. Then, let

$$\eta(D) = \left|\!\left|\!\left| P^{-1} \right|\!\right|\!\right|_{\infty \to 2} |\!|\!| P |\!|\!|_\infty \sum_{t=0}^{\infty} \eta_t(\Lambda),$$

where $\eta_0(\Lambda) = 1$.

Clearly, $\eta(D) < \infty$ if and only if $|\lambda_{\max}(D)| < 1$. Also note that

$$\eta_t(\Lambda) \leq t^{\mu(D)-1} |\lambda_{\max}(D)|^t \sum_{j=0}^{\mu(D)-1} \frac{|\lambda_{\max}(D)|^{-j}}{j!},$$

where $\mu(D) = \max_{1 \leq i \leq k} m_i$. For example, if the stable matrix D is *diagonalizable*, we have $\mu(D) = 1$, i.e.

$$\eta(D) \leq \frac{|||P^{-1}|||_{\infty \to 2} |||P|||_{\infty}}{1 - |\lambda_{\max}(D)|}.$$

2.3.1 Stable Case

Define the following notations which will be used throughout this subsection to establish high probability estimation results.

$$\begin{aligned} \nu_n(\delta) &= b_2^{1/\alpha} \log^{1/\alpha}\left(\frac{b_1 n p}{\delta}\right), \\ \pi_n(\delta) &= \eta(D)\left(\|x(0)\|_{\infty} + \nu_n(\delta)\right). \end{aligned}$$

As the proofs reveal, one can see that $\nu_n(\delta), \pi_n(\delta)$ are respectively the high probability uniform bounds for the size of the noise vectors $\{\|w(t)\|_{\infty}\}_{t=1}^{n}$, and state vectors $\{\|x(t)\|_2\}_{t=0}^{n}$. As a matter of fact, if the noise process is uniformly bounded, $\nu_n(\delta)$, and so $\pi_n(\delta)$, are fixed constants, not depending on n, δ.

Then, let $N_{2.1}(\epsilon, \delta)$ be large enough, such that the followings hold for all $n \geq N_{2.1}(\epsilon, \delta)$:

$$\frac{n}{\nu_n(\delta)^2} \geq \frac{18|\lambda_{\max}(C)| + 2\epsilon}{\epsilon^2} p \log\left(\frac{4p}{\delta}\right), \tag{2.4}$$

$$\frac{n}{\pi_n(\delta)^2 \nu_n(\delta)^2} \geq \frac{288}{\epsilon^2} p |||D|||_2^2 \log\left(\frac{4p}{\delta}\right), \tag{2.5}$$

$$\frac{n}{\pi_n(\delta)^2} \geq \frac{6}{\epsilon}\left(|||D|||_2^2 + 1\right). \tag{2.6}$$

Defining the empirical covariance matrix

$$V_n = \sum_{t=0}^{n-1} x(t)x(t)', \tag{2.7}$$

the following theorem provides a high probability lower bound for the smallest eigenvalue of V_{n+1}.

Theorem 2.1 (empirical covariance matrix)**.** Suppose that D is stable. If $n \geq N_{2.1}(\epsilon, \delta)$, then

$$\mathbb{P}\left(\left|\lambda_{\min}\left(V_{n+1}\right)\right| < n\left(\left|\lambda_{\min}(C)\right| - \epsilon\right)\right) < 2\delta.$$

Moreover, $\lim_{n\to\infty} \frac{1}{n} V_n = \sum_{i=0}^{\infty} D^i C D'^i$.

Using Theorem 2.1, we present the following bound on the prediction error of the least-squares estimator.

Lemma 2.3 (Prediction bound)**.** When D is stable, if $n-1 \geq N_{2.1}\left(\frac{|\lambda_{\min}(C)|}{2}, \delta\right)$, then the following holds, with probability at least $1-3\delta$:

$$\left|\left|\left|\left(\hat{D}_n - D\right) V_n \left(\hat{D}_n - D\right)'\right|\right|\right|_2 \leq \beta_n(\delta),$$

where

$$\beta_n(\delta) = \frac{16np}{(n-1)\left|\lambda_{\min}(C)\right|} \pi_n(\delta)^2 \nu_n(\delta)^2 \log\left(\frac{2p}{\delta}\right).$$

We also present bounds for the number of observations sufficient to guarantee high probability accurate estimation of closed-loop transition matrix D. Indeed, there is $N_{2.2}(\epsilon, \delta)$ as the least number of observations, to estimate D with error at most ϵ, with probability at least $1-2\delta$. For this purpose, assume the followings hold for all $n \geq N_{2.2}(\epsilon, \delta)$:

$$n \geq N_{2.1}\left(\frac{\left|\lambda_{\min}(C)\right|}{2}, \frac{\delta}{2}\right) + 1, \tag{2.8}$$

$$\frac{n-2}{\pi_n(\delta)^2 \nu_n(\delta)^2} \geq \frac{32p}{\left|\lambda_{\min}(C)\right|^2 \epsilon^2} \log\left(\frac{4p}{\delta}\right). \tag{2.9}$$

Theorem 2.2 (Stable estimation)**.** Assume D is stable. If $n \geq N_{2.2}(\epsilon, \delta)$, then

$$\mathbb{P}\left(\left|\left|\left| \hat{D}_n - D \right|\right|\right|_2 > \epsilon \right) < 2\delta.$$

2.3.2 Unstable Case

Results of the previous subsection concern about the estimation when the closed-loop transition matrix D is stable. But, for stabilization of the system, one needs to have accurate estimation of not necessarily stable matrix D. In the sequel, we provide results when D is unstable. As expected, D still needs to meet some requirements, which we will prove later can be satisfied, if a random linear feedback will be applied to the system. In fact, the transition matrix D needs to be regular, according to the following definition, in order to have an accurate estimation.

Definition 2.3 (Regularity)**.** $D \in \mathbb{R}^{p \times p}$ is called regular if for any explosive eigenvalue of D, denoted by λ, the geometric multiplicity of λ is one.

Regularity implies that the eigenspace corresponding to λ is one dimensional, and vice versa. There are other equivalent formulations for regularity. Indeed, D is regular if and only if for any explosive eigenvalue λ, in the Jordan decomposition of D there is only one block corresponding to λ. In other words, no matter how large the algebraic multiplicity of λ is, the geometric multiplicity is one. Another equivalent formulation is the following one. D is regular if and only if

$$\text{rank}\left(D - \lambda I_p\right) \geq p - 1,$$

for all $\lambda \in \mathbb{C}$ such that $|\lambda| > 1$. For example, let $P_1, P_2 \in \mathbb{C}^{2 \times 2}$ be arbitrary invertible matrices, and assume

$$D_1 = P_1^{-1} \begin{bmatrix} \rho & 1 \\ 0 & \rho \end{bmatrix} P_1, D_2 = P_2^{-1} \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix} P_2,$$

are real 2×2 matrices, where $\rho \in \mathbb{C}$ satisfies $|\rho| > 1$. Then, D_1 *is* regular, but D_2 *is not*.

To sum up this section, we present the accuracy of the least-squares estimation, in general (i.e. unstable) VAR processes. The following theorem states the accuracy of the estimation, when the matrix D is regular, and has no eigenvalue exactly on the unit circle of the complex plane. As we will see in Section 2.4, these assumptions are not restrictive when a random linear feedback is applied to a stabilizable system with unknown dynamics.

The sample size for unstable case is based on the constant Δ_{unstable}, as well as the function $\psi(D_0, \delta)$. Technically, Δ_{unstable} is a constant depending on the matrices C, D, as well as the parameters b_1, b_2, α specified in Assumption 2.1. Further, the function $\psi(D_0, \delta)$ depends on both D and the distribution of the noise sequence $\{w(t)\}_{t=1}^{\infty}$, and we have $\psi(D_0, \delta) > 0$, for all $\delta > 0$. If in addition there exists $t_0 \geq 1$ such that the noise $w(t_0)$ is a continuous random vector with a bounded probability density function on $\mathbb{R}^p$, then, for all $\delta > 0$ we have

$$\psi(D_0, \delta) \geq \psi_0 \delta,$$

where ψ_0 is a positive constant, and does not depend on δ. More details are provided in Section 4.4.

Let $N_{2.3}(\epsilon, \delta)$ be large enough, such that for all $n \geq N_{2.3}(\epsilon, \delta)$,

$$\frac{n}{(\log n)^{4/\alpha}} \geq \frac{\Delta_{\text{unstable}}}{\epsilon^2} \left((-\log \delta)^{1+4/\alpha} - \log \psi(D_0, \delta) \right).$$

Theorem 2.3 (Unstable estimation)**.** Suppose that D is regular, and has no eigenvalue of the unit size. As long as $n \geq N_{2.3}(\epsilon, \delta)$, we have

$$\mathbb{P}\left(\left\|\left|\hat{D}_n - D\right|\right\|_2 \leq \epsilon\right) \geq 1 - \delta.$$

2.4 Stabilizing the System

In this section, we show how the system can be stabilized. Even though the true parameter Θ_0 is unknown, according to Lemma 2.1, a stabilizing linear feedback $L(\Theta)$ can be designed, if one can find a stabilizing neighborhood $\Omega^{(0)}$, such that

$$\Omega^{(0)} \subset \left\{\Theta \in \mathbb{R}^{p \times q} : \left\|\left|\Theta - \Theta_0\right|\right\|_2 \leq \epsilon_0\right\}. \tag{2.10}$$

Algorithm 1 : Stabilization
Input: Stabilization Radius ϵ_0, Failure Probability δ
Output: Stabilizing Set $\Omega^{(0)}$

```
Let k = 1 + ⌈r/p⌉, τ_0 = 0
for i = 1, ··· , k do
    for j = 1, ··· , p do
        Draw column j of L_i from N(0, I_r), independently
    end for
end for
Define M, ε̃ according to (2.11), (2.12), respectively
for i = 1, ··· , k do
    Define τ_i by (2.13)
    while t < τ_i do
        Apply control action u(t) = L_i x(t)
    end while
    Estimate D̂^(i) by (2.14), (2.15)
    Construct Ω^(i) by (2.16)
end for
Let Ω^(0) = ⋂_{i=1}^{k} Ω^(i)
return Ω^(0)
```

Once a random linear feedback is applied to the system, using Theorem 2.3 we show that with high probability, $\Omega^{(0)}$ can be learned, if one observes the state vectors $\{x(t)\}_{t=0}^{n}$.

Since in Theorem 2.3 the closed-loop transition matrix needs to be regular with no eigenvalue of the unit size, first we need to show that these conditions can be satisfied.

Lemma 2.4, and Lemma 2.5 do this, with no knowledge beyond stabilizability of $[A_0, B_0]$. Based on the properties of the distribution of a random linear feedback matrix L, the above lemmas provide general statements, which hold almost surely. Then, we present an applicable stabilizing algorithm, and prove that it will provide us the desired stabilizing neighborhood. To proceed, we define the following classes of probability distributions over real valued vectors and matrices.

Definition 2.4 (Full rank distributions)**.** Let X be a random vector in $\mathbb{R}^m$. We say X has a linearly full rank distribution if for any arbitrary hyperplane in $\mathbb{R}^m$ such as $\mathcal{P}$, it holds that

$$\mathbb{P}\left(X \in \mathcal{P}\right) = 0.$$

Further, X has a general full rank distribution, if for every manifold $\mathcal{M} \subset \mathbb{R}^m$ such that $\dim_{\mathbb{R}}\left(\mathcal{M}\right) \leq m - 1$, it holds that

$$\mathbb{P}\left(X \in \mathcal{M}\right) = 0.$$

Example 2.2. Let $Z \in \mathbb{R}^p$ be normally distributed, $Z \sim \mathcal{N}\left(\mu, \Sigma\right)$, with arbitrary mean $\mu \in \mathbb{R}^p$, and positive definite covariance matrix $\Sigma \in \mathbb{R}^{p \times p}$. Then, Z has a general full rank distribution. Letting

$$Y = \frac{Z}{\left\|Z\right\|_2} 1_{\{Z \neq 0\}},$$

the random vector Y has a linearly full rank distribution, but since Y lives on the unit sphere, Y does not have a general full rank distribution.

Random linear feedbacks with full rank distributions, induce the desired properties to the closed-loop transition matrix $A_0 + B_0 L$. In the next lemmas, we rigorously present that the desired properties hold almost surely.

Lemma 2.4 (Closed-loop regularity)**.** Assume $[A_0, B_0]$ is stabilizable. Let the columns of $L \in \mathbb{R}^{r \times p}$ be independent (but not necessarily identically distributed), with linearly full rank distributions. The matrix $A_0 + B_0 L$ is regular, with probability one.

In addition to regularity, $A_0 + B_0 L$ is unit-root free, if the distribution of the linear feedback L is not only linearly full rank, but also generally full rank.

Lemma 2.5 (Eigenvalues of closed-loop)**.** Assume $[A_0, B_0]$ is stabilizable. Let $L \in \mathbb{R}^{r \times p}$ have a general full rank distribution over $\mathbb{R}^{r \times p}$. With probability one, $A_0 + B_0 L$ has no eigenvalue of the unit size.

Subsequently, an algorithmic procedure to find a stabilizing neighborhood will be presented based on random linear feedbacks discussed above. Let the columns of random linear feedbacks $L_1, \cdots, L_k \in \mathbb{R}^{r \times p}$ be drawn from arbitrary general full rank distributions independently, where $k = 1 + \lceil \frac{r}{p} \rceil$. Note that because of independence, for all $i = 1, \cdots, k$, the random feedback L_i has a general full rank distribution, and so, according to Lemma 2.4 and Lemma 2.5, Theorem 2.3 can be applied.

According to Theorem 2.3, every closed-loop transition matrix $D^{(i)} = A_0 + B_0 L_i$ can be estimated with high probability arbitrarily accurate, if the number of observations is large enough. We show how to find a high probability confidence set for Θ_0, using those for $D^{(1)}, \cdots, D^{(k)}$. More precisely, letting $\tau_0 = 0$,

$$M = \begin{bmatrix} I_p & \cdots & I_p \\ L_1 & \cdots & L_k \end{bmatrix} \in \mathbb{R}^{q \times kp}, \tag{2.11}$$

$$\tilde{\epsilon} = \frac{\epsilon_0}{2k} \inf \left\{ \frac{|||\Theta M|||_2}{|||\Theta|||_2} : \Theta \in \mathbb{R}^{p \times q} \right\}, \tag{2.12}$$

note that since $L_1, \cdots, L_k$ are independent, we have $\operatorname{rank}(M) = q$, almost surely, i.e. $\mathbb{P}(\tilde{\epsilon} > 0) = 1$. Even if $\tilde{\epsilon}$ became too small, one can repeat the drawing of the columns of the linear feedbacks, in order to avoid pathologically small values of $\tilde{\epsilon}$.

Using the sample size $N_{2.3}(\cdot,\cdot)$ used in Theorem 2.3, for $i = 1, \cdots, k$ define

$$\tau_i = \tau_{i-1} + N_{2.3}\left(\tilde{\epsilon}, \frac{\delta}{k}\right), \tag{2.13}$$

$$\hat{d}_j^{(i)} = \arg\min_{\theta \in \mathbb{R}^p} \sum_{t=\tau_{i-1}}^{\tau_i - 1} \left(x_j(t+1) - \theta' x(t)\right)^2, \tag{2.14}$$

$$\hat{D}^{(i)} = \left[\hat{d}_1^{(i)}, \cdots, \hat{d}_p^{(i)}\right], \tag{2.15}$$

$$\Omega^{(i)} = \left\{\Theta \in \mathbb{R}^{p \times q} : \left|\left|\left|\Theta \begin{bmatrix} I_p \\ L_i \end{bmatrix} - \hat{D}^{(i)}\right|\right|\right|_2 \leq \tilde{\epsilon}\right\}. \tag{2.16}$$

In fact, τ_i shows the time points where the control action (i.e. the linear feedback) changes, $\hat{D}^{(i)}$ is nothing but the least-squares estimate of $D^{(i)}$, and $\Omega^{(i)}$ is a high probability confidence set for Θ_0, based on L_i. Design of τ_i implies by Theorem 2.3 that with probability at least $1 - \frac{\delta}{k}$, we have $\left|\left|\left|\hat{D}^{(i)} - D^{(i)}\right|\right|\right|_2 \leq \tilde{\epsilon}$. Since $D^{(i)} = \Theta_0 \begin{bmatrix} I_p \\ L_i \end{bmatrix}$, letting

$$\Omega^{(0)} = \bigcap_{i=1}^{k} \Omega^{(i)},$$

clearly

$$\mathbb{P}\left(\Theta_0 \notin \Omega^{(0)}\right) \leq \sum_{i=1}^{k} \mathbb{P}\left(\Theta_0 \notin \Omega^{(i)}\right) \leq \delta.$$

Further, for arbitrary $\Theta_1 \in \Omega^{(0)}$, as long as $\Theta_0 \in \Omega^{(0)}$, we have

$$\begin{aligned} & \left|\left|\left|(\Theta_1 - \Theta_0) \begin{bmatrix} I_p \\ L_i \end{bmatrix}\right|\right|\right|_2 \\ \leq & \left|\left|\left|\Theta_1 \begin{bmatrix} I_p \\ L_i \end{bmatrix} - \hat{D}^{(i)}\right|\right|\right|_2 + \left|\left|\left|\Theta_0 \begin{bmatrix} I_p \\ L_i \end{bmatrix} - \hat{D}^{(i)}\right|\right|\right|_2 \\ \leq & 2\tilde{\epsilon}, \end{aligned}$$

for all $i = 1, \cdots, k$. Thus, $\left\|\left\|\left(\Theta_1 - \Theta_0\right) M\right\|\right\|_2 \leq 2k\tilde{\epsilon}$, which according to (2.12) implies

$$\frac{2k\tilde{\epsilon}}{\left\|\left\|\Theta_1 - \Theta_0\right\|\right\|_2} \geq \frac{\left\|\left\|\left(\Theta_1 - \Theta_0\right) M\right\|\right\|_2}{\left\|\left\|\Theta_1 - \Theta_0\right\|\right\|_2} \geq \frac{2k\tilde{\epsilon}}{\epsilon_0},$$

or equivalently

$$\left\|\left\|\Theta_1 - \Theta_0\right\|\right\|_2 \leq \epsilon_0,$$

i.e. (2.10) holds, with probability at least $1 - \delta$. Algorithm 1 returns $\Omega^{(0)}$, taking stabilization radius ϵ_0 and failure probability δ as inputs. Obviously, the normal distribution $\mathcal{N}\left(0, I_r\right)$ used in Algorithm 1 is not unique, and can be substituted by any general full rank distribution over $\mathbb{R}^r$. Putting all together, we get the following theorem.

Theorem 2.4 (High probability stabilization)**.** Let $\Omega^{(0)}$ be the stabilizing set provided by Algorithm 1. For arbitrary $\Theta \in \Omega^{(0)}$, we have

$$\mathbb{P}\left(\left|\lambda_{\max}\left(\Theta_0 \tilde{L}\left(\Theta\right)\right)\right| < 1\right) \geq 1 - \delta.$$

2.5 Technical Proofs

2.5.1 Proofs of Section 2.2

Proof of Proposition 2.1. For convenience, let $K_0 = K(\Theta_0)$, and $L_0 = L(\Theta_0)$. First, assume $[A_0, B_0]$ is stabilizable, L is a stabilizer, $D = A_0 + B_0 L$, and $|\lambda_{\max}(D)| < 1$. For arbitrary fixed PSD matrix P_0, define $P_t(P_0), t = 1, \cdots, T$ recursively,

$$P_t(P_0) = Q + A_0' P_{t-1}(P_0) A_0 - A_0' P_{t-1}(P_0) B_0 (B_0' P_{t-1}(P_0) B_0 + R)^{-1} B_0' P_{t-1}(P_0) A_0. \tag{2.17}$$

Letting c_t be as defined in (1.2), according to classical literature [17], the optimal control policy for minimizing the finite horizon accumulative cost

$$\mathcal{J}_T = \sum_{t=0}^{T-1} \mathbb{E}[c_t] + \mathbb{E}[x(T)' P_0 x(T)],$$

is $u(t) = L_t x(t), t = 0, \cdots, T-1$, where

$$L_t = -(B_0' P_{T-t-1}(P_0) B_0 + R)^{-1} B_0' P_{T-t-1}(P_0) A_0. \tag{2.18}$$

Moreover, this optimal policy yields the optimal cost

$$\min \mathcal{J}_T = x(0)' P_T(P_0) x(0) + \sum_{t=0}^{T-1} \operatorname{tr}(C P_t(P_0)). \tag{2.19}$$

On the other hand, applying the control policy $u(t) = Lx(t)$, for $t = 0, \cdots, T-2$ we have

$$\begin{aligned}
\mathbb{E}[c_{t+1}|x(t)] &= \mathbb{E}[x(t+1)'(Q + L'RL) x(t+1)|x(t)] \\
&= \mathbb{E}\left[(Dx(t) + w(t+1))'(Q + L'RL)(Dx(t) + w(t+1))|x(t)\right] \\
&= x(t)' D' (Q + L'RL) Dx(t) + \operatorname{tr}(C(Q + L'RL)),
\end{aligned}$$

and

$$\mathbb{E}\left[x(T)'P_0x(T)|x(T-1)\right] \;\; = \;\; x(T-1)'D'P_0Dx(T-1) + \operatorname{tr}\left(CP_0\right).$$

Hence, the finite horizon cost becomes

$$\mathcal{J}_T = x(0)'\tilde{P}_T\left(P_0\right)x(0) + \sum_{t=0}^{T-1}\operatorname{tr}\left(C\tilde{P}_t\left(P_0\right)\right), \tag{2.20}$$

where $\tilde{P}_t\left(P_0\right), t = 1, \cdots, T$ are defined recursively as

$$\tilde{P}_0\left(P_0\right) \;\; = \;\; P_0, \tag{2.21}$$

$$\tilde{P}_t\left(P_0\right) \;\; = \;\; Q + L'RL + D'\tilde{P}_{t-1}\left(P_0\right)D. \tag{2.22}$$

Since $|\lambda_{\max}\left(D\right)| < 1$, $\lim_{T\to\infty}\tilde{P}_T\left(P_0\right) = P_\infty$ for a PSD matrix P_∞. Letting $C \to 0$, by (2.19), (2.20) we have

$$x(0)'P_T\left(P_0\right)x(0) \leq x(0)'\tilde{P}_T\left(P_0\right)x(0),$$

i.e. $x(0)'P_T\left(P_0\right)x(0), T = 1, 2, \cdots$ is bounded. If $P_0 = 0$, this sequence is nondecreasing, because minimizing both sides of

$$\sum_{t=0}^{T-1} c_t \leq \sum_{t=0}^{T} c_t$$

subject to

$$x(t+1) = A_0x(t) + B_0u(t),$$

we get

$$x(0)'P_T\left(0\right)x(0) \leq x(0)'P_{T+1}\left(0\right)x(0).$$

Therefore, the nondecreasing bounded sequence $x(0)'P_T\,(0)\,x(0), T = 1, 2, \cdots$ converges. Since $x(0)$ is arbitrary, $P_T(0), T = 1, 2, \cdots$ itself converges:

$$\lim_{T\to\infty} P_T(0) = P_\infty(0).$$

According to the recursive definition of $P_t(0)$ in (2.17), $P_\infty(0)$ is a solution of (2.2). This shows the existence of solution, and uniqueness will be shown later.

Next, since $\lim\limits_{T\to\infty} P_T(0) = P_\infty(0)$, (2.19) implies $\lim\limits_{t\to\infty} \operatorname{tr}(CP_t(0)) = \operatorname{tr}(CP_\infty(0))$. So, the Cesaro mean also converges to this limit, i.e.

$$\mathcal{J}^\star(\Theta_0) = \operatorname{tr}(CP_\infty(0)).$$

Optimality of the linear feedback $u(t) = L_0 x(t)$, is concluded by (2.18). Now, we are ready to show that L_0 is a stabilizer. Letting

$$D = A_0 + B_0 L_0, C \to 0, K_0 = P_\infty(0),$$

we show that for arbitrary $x(0)$, $x(t) = D^t x(0)$ vanishes as t grows. First, note that by (2.2), (2.3),

$$\begin{aligned} \left(B_0' K_0 B_0 + R\right) L_0 &= -B_0' K_0 A, \\ L_0' \left(B_0' K_0 B_0 + R\right) L_0 &= A_0' K_0 B_0 \left(B_0' K_0 B_0 + R\right)^{-1} B_0' K_0 A_0. \end{aligned}$$

Therefore, doing some algebra we get

$$\begin{aligned}
& Q + L_0' R L_0 + D' K_0 D \\
= \;& Q + A_0' K_0 A_0 + L_0' \left(B_0' K_0 A_0 + R \right) L_0 + A_0' K_0 B_0 L_0 + L_0' B_0' K_0 A_0 \\
= \;& Q + A_0' K_0 A_0 - A_0' K_0 B_0 \left(B_0' K_0 B_0 + R \right)^{-1} B_0' K_0 A_0 \\
+ \;& \left[L_0' \left(B_0' K_0 B_0 + R \right) + A_0' K_0 B_0 \right] L_0 + L_0' \left[\left(B_0' K_0 B_0 + R \right) L_0 + B_0' K_0 A_0 \right] \\
= \;& K_0,
\end{aligned}$$

i.e.

$$K_0 - D' K_0 D = Q + L_0' R L_0. \tag{2.23}$$

So,

$$x(t{+}1)' K_0 x(t{+}1) - x(t)' K_0 x(t) = x(t)' \left(D' K_0 D - K_0 \right) x(t) = -x(t)' \left(Q + L_0' R L_0 \right) x(t). \tag{2.24}$$

Adding up both sides of (2.24), because K_0 is PSD we have

$$-x(0)' K_0 x(0) \le x(t+1)' K_0 x(t+1) - x(0)' K_0 x(0) = -\sum_{i=0}^{t} x(i)' \left(Q + L_0' R L_0 \right) x(i), \tag{2.25}$$

in other words,

$$\lim_{t \to \infty} x(t)' \left(Q + L_0' R L_0 \right) x(t) = 0.$$

Thus, since Q is positive definite, $\lim\limits_{t \to \infty} x(t) = 0$, i.e. L_0 is a stabilizer. Back to the proof of the existence of a solution K_0, we show that for arbitrary PSD P_0, it holds that $\lim\limits_{T \to \infty} P_T(P_0) = P_\infty(0)$. To do so, minimize both sides of

$$\sum_{t=0}^{T-1} c_t \le \sum_{t=0}^{T-1} c_t + x(T)' P_0 x(T),$$

subject to

$$x(t+1) = A_0 x(t) + B_0 u(t),$$

to get

$$x(0)' P_T (0) x(0) \le x(0)' P_T (P_0) x(0). \tag{2.26}$$

On the other hand, applying controller $u(t) = L_0 x(t)$, the cost $\sum_{t=0}^{T-1} c_t + x(T)' P_0 x(T)$ becomes

$$\sum_{t=0}^{T-1} x(0)' D'^t (Q + L_0' R L_0) D^t x(0) + x(0)' D'^T P_0 D^T x(0). \tag{2.27}$$

Note that because of stability $|\lambda_{\max}(D)| < 1$, we have

$$\lim_{T\to\infty} x(0)' D'^T P_0 D^T x(0) = 0.$$

Therefore, by (2.26), (2.27), and (2.23),

$$\begin{aligned}
x(0)' P_\infty (0) x(0) &= \lim_{T\to\infty} x(0)' P_T (0) x(0) \le \lim_{T\to\infty} x(0)' P_T (P_0) x(0) \\
&\le \lim_{T\to\infty} \sum_{t=0}^{T-1} x(0)' D'^t (Q + L_0' R L_0) D^t x(0) + x(0)' D'^T P_0 D^T x(T) \\
&= \lim_{T\to\infty} \sum_{t=0}^{T-1} x(0)' D'^t (K_0 - D' K_0 D) D^t x(0) \\
&= x(0)' K_0 x(0),
\end{aligned}$$

i.e. for arbitrary P_0,

$$\lim_{T\to\infty} P_T (P_0) = P_\infty(0).$$

Using this, we show that K_0 is the unique solution of (2.2). If P_* is another solution, let $P_0 = P_*$, which plugging in (2.17) implies $P_t (P_*) = P_*$, for $t = 1, 2, \cdots$. Hence $P_* = \lim_{T\to\infty} P_T (P_*) = P_\infty(0)$, i.e. the solution K_0 of (2.2) exists, and is unique.

Conversely, if K_0 is a solution of (2.2), define L_0 as (2.3) and $D = A_0 + B_0 L_0$. Note that K_0 is positive semidefinite, and let $P_0 = K_0$. Define P_t by (2.17), which yield $P_t = K_0$, for $t = 0, 1, \cdots$. As before, (2.2), (2.3) imply (2.23). Similarly, (2.24), (2.25) hold, i.e. $\lim_{t\to\infty} D^t x(0) = 0$ for arbitrary $x(0)$, which means L_0 defined in (2.3) is a stabilizer. $\square$

Proof of Proposition 2.2. When applying the linear feedback $L(\Theta_0)$, the closed-loop transition matrix will be $D = \Theta_0 \tilde{L}(\Theta_0) = A_0 + B_0 L(\Theta_0)$. Letting $P = Q + L(\Theta_0)' R L(\Theta_0)$, we have the followings. First,

$$\mathcal{R}(T-1) + x(0)'Px(0) - \mathcal{J}^\star(\Theta_0) = \sum_{t=0}^{T-1} x(t)'Px(t) - T\mathcal{J}^\star(\Theta_0) = \operatorname{tr}(PV_T) - T\mathcal{J}^\star(\Theta_0),$$

where V_T is defined in (2.7).

Second, $x(t+1) = Dx(t) + w(t+1)$ implies $V_T = DV_TD' + E_T$, where

$$\begin{aligned}
E_T &= U_T + C_T + D\left(x(0)x(0)' - x(T-1)x(T-1)'\right)D' + x(0)x(0)', \\
U_T &= \sum_{t=0}^{T-2} \left[Dx(t)w(t+1)' + w(t+1)x(t)'D'\right], \\
C_T &= \sum_{t=1}^{T-1} w(t)w(t)'.
\end{aligned}$$

Third, by Proposition 2.1, $\mathcal{J}^\star(\Theta_0) = \operatorname{tr}(K(\Theta_0)C)$. Finally, stability of the system yields $\lim_{T\to\infty} \frac{1}{T^{1/2}} \|x(T-1)\|_2^2 = 0$, almost surely. Putting all above together, and using (2.23), we get

$$\begin{aligned}
\lim_{T\to\infty} \frac{1}{T^{1/2}} \mathcal{R}(T) &= \lim_{T\to\infty} \frac{1}{T^{1/2}} \operatorname{tr}\left(P \sum_{n=0}^{\infty} D^n E_T D'^n - TC \sum_{n=0}^{\infty} D'^n P D^n\right) \\
&= \sum_{n=0}^{\infty} \operatorname{tr}\left(D'^n P D^n \lim_{T\to\infty} \frac{U_T + C_T - TC}{T^{1/2}}\right).
\end{aligned}$$

According to Lindeberg's Central Limit Theorem [18], the asymptotic distribution of all $2p^2$ entries of the matrices $\frac{C_T - TC}{T^{1/2}}, \frac{U_T}{T^{1/2}}$ is a multivariate mean-zero normal. By $|\lambda_{\max}(D)| < 1$, the matrix $\sum_{n=0}^{\infty} D'^n P D^n$ is bounded. So, the desired result holds, since a linear combination $(\operatorname{tr}(\cdot))$ of jointly normal random variables is normal as well. $\square$

Proof of Lemma 2.1. Since Θ is stabilizable, according to Proposition 2.1, $\Theta\tilde{L}(\Theta)$ is stable,

$$\left|\lambda_{\max}\left(\Theta\tilde{L}(\Theta)\right)\right| \leq 1-2\rho,$$

for some $\rho > 0$. For arbitrary $1 \leq i \leq p, 1 \leq j \leq q$, let all entries of matrix $X_{ij} \in \mathbb{R}^{p\times q}$ be zero, except the ij-th entry, which is one. Then, for $\theta \in \mathbb{R}$, consider the polynomial

$$f_\theta(\lambda) = \det\left((\Theta + \theta X_{ij})\tilde{L}(\Theta) - \lambda I_p\right).$$

All coefficients of $f_\theta(\lambda)$ are linear functions of θ. Further, the magnitudes of the roots of $f_\theta(\lambda)$ are continuous with respect to the coefficients, and so, are also continuous with respect to θ. Since all roots of $f_0(\lambda)$ are in magnitude at most $1-2\rho$, there exists $\epsilon_{ij} > 0$, such that $|\theta| < \epsilon_{ij}$ implies that all roots of $f_\theta(\lambda)$ are in magnitude at most

$$1 - \left(2 - \frac{1}{pq}\right)\rho.$$

Taking $\epsilon_0 = \min_{i,j}\epsilon_{ij}$, by $|||\Theta - \Theta_0|||_2 < \epsilon_0$, Θ_0 can be written in the form of $\Theta_0 = \Theta + \sum_{i=1}^{p}\sum_{j=1}^{q}\theta_{ij}X_{ij}$, where $|\theta_{ij}| < \epsilon_{ij}$, for all i, j. Therefore, all roots of

$$f(\lambda) = \det\left(\Theta_0\tilde{L}(\Theta) - \lambda I_p\right)$$

are in magnitude at most $1-\rho$, which is the desired result. □

Proof of Lemma 2.2. First, Let $D_1, D_2 \in \mathbb{R}^{p\times p}$ be stable, and $P \in \mathbb{R}^{p\times p}$ be a positive semidefinite matrix. For $i = 1, 2$, define $F_i = \sum_{n=0}^{\infty} D_i'^n P D_i^n$. We show that

$$|||F_1 - F_2|||_2 \leq \Gamma_F |||D_1 - D_2|||_2, \tag{2.28}$$

for some $\Gamma_F < \infty$. For $n = 1, 2, \cdots$, we have

$$\begin{aligned}
|||D_2^n - D_1^n|||_2 &= |||(D_1 + D_2 - D_1)^n - D_1^n|||_2 \\
&\leq \sum_{a_0+\sum_{j=1}^m (a_j+b_i)=n,\ \sum_{j=0}^m a_j<n} \left|\left|\left| D_1^{a_0} \prod_{j=1}^m (D_2 - D_1)^{b_j} D_1^{a_j} \right|\right|\right|_2 \\
&\leq \sum_{a_0+\sum_{j=1}^m (a_j+b_i)=n,\ \sum_{j=0}^m a_j<n} |||D_1|||_2^{\sum_{j=0}^m a_j} |||D_2 - D_1|||_2^{\sum_{j=1}^m b_j} \\
&= \sum_{\ell=1}^n \binom{n}{\ell} |||D_1|||_2^{n-\ell} |||D_1 - D_2|||_2^\ell \\
&\leq \frac{(|||D_1|||_2 + |||D_1 - D_2|||_2)^n}{|||D_1|||_2} n |||D_1 - D_2|||_2.
\end{aligned}$$

Then, there is $k < \infty$, such that

$$\max\left\{ \left|\left|\left|D_1'^k\right|\right|\right|_2, \left|\left|\left|D_1^k\right|\right|\right|_2, \left|\left|\left|D_2'^k\right|\right|\right|_2, \left|\left|\left|D_2^k\right|\right|\right|_2 \right\} \leq 1 - 2\rho,$$

for some $\rho > 0$. Define

$$E_i = D_i^k, P_i = \sum_{n=0}^{k-1} D_i'^n P D_i^n.$$

Noting that

$$|||D_2' - D_1'|||_2 \leq \Gamma_0 |||D_2 - D_1|||_2,$$

we have

$$
\begin{aligned}
|||E_2 - E_1|||_2 &\le \frac{\left(|||D_1|||_2 + |||D_2 - D_1|||_2\right)^k}{|||D_1|||_2} k |||D_2 - D_1|||_2 = \Gamma_E |||D_2 - D_1|||_2, \\
|||E_2' - E_1'|||_2 &\le \frac{\left(|||D_1'|||_2 + |||D_2' - D_1'|||_2\right)^k}{|||D_1'|||_2} k |||D_2' - D_1'|||_2 = \Gamma_{E'} |||D_2 - D_1|||_2, \\
|||P_2 - P_1|||_2 &\le \sum_{n=1}^{k-1} \left[|||D_2'^n P \left(D_2^n - D_1^n\right)|||_2 + |||\left(D_2'^n - D_1'^n\right) P D_1^n|||_2 \right] \\
&\le \sum_{n=1}^{k-1} \left[|||D_2'^n|||_2 |||D_2^n - D_1^n|||_2 + |||D_1^n|||_2 |||D_2'^n - D_1'^n|||_2 \right] |||P|||_2 n \\
&\le \Gamma_P |||D_2 - D_1|||_2.
\end{aligned}
$$

Suppose that $|||D_2 - D_1|||_2$ is small enough to satisfy

$$
\max\left\{|||E_2 - E_1|||_2, |||E_2' - E_1'|||_2\right\} \le \rho.
$$

Since $|||E_1|||_2 + |||E_1 - E_2|||_2 \le 1 - \rho$, $|||E_1'|||_2 + |||E_1' - E_2'|||_2 \le 1 - \rho$, and $F_i = \sum_{n=0}^{\infty} E_i'^n P_i E_i^n$, similar to the upper bound above for $|||D_2^n - D_1^n|||_2$ we have

$$
\begin{aligned}
|||E_2^n - E_1^n|||_2 &\le \frac{\left(|||E_1|||_2 + |||E_2 - E_1|||_2\right)^n}{|||E_1|||_2} n |||E_2 - E_1|||_2 \\
&\le \frac{\Gamma_E}{|||E_1|||_2} |||D_2 - D_1|||_2 n \left(1 - \rho\right)^n, \\
|||E_2'^n - E_1'^n|||_2 &\le \frac{\left(|||E_1'|||_2 + |||E_2' - E_1'|||_2\right)^n}{|||E_1'|||_2} n |||E_2' - E_1'|||_2 \\
&\le \frac{\Gamma_{E'}}{|||E_1'|||_2} |||D_2 - D_1|||_2 n \left(1 - \rho\right)^n.
\end{aligned}
$$

So,

$$
\begin{aligned}
|||F_2 - F_1|||_2 &\le \sum_{n=0}^{\infty} \left[|||E_2'^n P_2 \left(E_2^n - E_1^n\right)|||_2 + |||\left(E_2'^n - E_1'^n\right) P_2 E_1^n|||_2 + |||E_1'^n \left(P_2 - P_1\right) E_1^n|||_2 \right] \\
&\le \sum_{n=0}^{\infty} \left[\frac{|||P_2|||_2 \Gamma_E}{|||E_1|||_2} n + \frac{|||P_2|||_2 \Gamma_{E'}}{|||E_1'|||_2} n + \Gamma_P \right] \left(1 - \rho\right)^{2n} |||D_2 - D_1|||_2 \\
&= \Gamma_F |||D_2 - D_1|||_2,
\end{aligned}
$$

i.e. (2.28) holds. Now, we prove the desired inequality. Consider two systems $(1), (2)$, with accumulative costs $\mathcal{J} = \sum_{t=0}^{\infty} c_t$, where for $i = 1, 2$, System (i) evolves according to $x(t+1) = A_i x(t) + B_i u(t)$, and both systems share the initial state $x(0) = x_0$, for $\|x_0\|_2 = 1$. Denoting the optimal accumulative cost of System (i) by $\mathcal{J}^{(i)}$, according to the proof of Proposition 2.1, we have $\mathcal{J}^{(i)} = x_0' K(\Theta_i) x_0$. Without loss of generality, assume $\mathcal{J}^{(1)} \geq \mathcal{J}^{(2)}$.

Using Lemma 2.1, assume $|||\Theta_1 - \Theta_2|||_2$ is sufficiently small, such that both matrices $\Theta_1 \tilde{L}(\Theta_2), \Theta_2 \tilde{L}(\Theta_1)$ are stable. Then, apply control policy $u(t) = L(\Theta_2) x(t)$ to both systems. The closed-loop matrices $D_i = \Theta_i \tilde{L}(\Theta_2)$ are stable, and

$$|||D_1 - D_2|||_2 \leq \left|\left|\left|\tilde{L}(\Theta_2)\right|\right|\right|_2 |||\Theta_1 - \Theta_2|||_2. \tag{2.29}$$

Letting $P = Q + L(\Theta_2)' R L(\Theta_2)$, as we saw in the proof of Proposition 2.1, the accumulative cost of System (i) is $x_0' F_i x_0$, where $F_i = \sum_{n=0}^{\infty} D_i'^n P D_i^n$. The linear feedback $L(\Theta_2)$ is an optimal policy for System (2), i.e. $x_0' K(\Theta_2) x_0 = x_0' F_2 x_0$, and $\mathcal{J}^{(1)}$ is the minimum accumulative cost for System (1), i.e. $x_0' K(\Theta_1) x_0 \leq x_0' F_1 x_0$. Therefore,

$$0 \leq \mathcal{J}^{(1)} - \mathcal{J}^{(2)} = x_0' K(\Theta_1) x_0 - x_0' K(\Theta_2) x_0 \leq x_0' (F_1 - F_2) x_0 \leq |||F_1 - F_2|||_2.$$

Since x_0 is an arbitrary unit vector, by (2.28), (2.29) we have

$$|||K(\Theta_1) - K(\Theta_2)|||_2 \leq \Gamma_F |||D_1 - D_2|||_2 \leq \Gamma_K |||\Theta_1 - \Theta_2|||_2,$$

which is the desired result. □

2.5.2 Proofs of Section 2.3

Lemma 2.6 (Noise upper bound)**.** For $n = 1, 2, \cdots$, and $0 < \delta < 1$, define the following event.

$$\mathcal{W} \;=\; \left\{ \max_{1 \le t \le n} \|w(t)\|_\infty \le \nu_n(\delta) \right\}.$$

We have

$$\mathbb{P}(\mathcal{W}) \;\ge\; 1 - \delta.$$

Proof of Lemma 2.6. First, note that for all $y > 0; i = 1, \cdots, p; t = 1, \cdots, n$, by Assumption 2.1 we have

$$\mathbb{P}(|w_i(t)| > \nu_n(\delta)) \;\le\; b_1 \exp\left(-\frac{\nu_n(\delta)^\alpha}{b_2}\right) = b_1 \exp\left(-\frac{b_2 \log \frac{b_1 np}{\delta}}{b_2}\right) = \frac{\delta}{np}.$$

So, using union bound we get

$$\mathbb{P}(\mathcal{W}^c) \le \sum_{t=1}^{n} \sum_{i=1}^{p} \mathbb{P}(|w_i(t)| > \nu_n(\delta)) \le \delta.$$

□

Lemma 2.7 (State upper bound)**.** Letting $D = P^{-1} \Lambda P$ be the Jordan decomposition of stable matrix D, on the event $\mathcal{W}$ we have

$$\|x(t)\|_2 \le \pi_n(\delta),$$

for all $t = 1, 2, \cdots, n$.

Proof of Lemma 2.7. First, the behavior of $|||\Lambda|||_\infty$ is determined by the blocks of Λ. In fact, letting $\Lambda = \operatorname{diag}(\Lambda_1, \cdots, \Lambda_k)$, simply the definition of operator norm $|||\cdot|||_\infty$ implies

$$|||\Lambda|||_\infty \leq \max_{1 \leq i \leq k} |||\Lambda_i|||_\infty.$$

Then, to find an upper bound for the operator norm of an exponent of an arbitrary block, such as

$$\Lambda_i = \begin{bmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \lambda \end{bmatrix} \in \mathbb{C}^{m \times m},$$

we show that

$$|||\Lambda_i^t|||_\infty \leq t^{m-1} |\lambda|^t \sum_{j=0}^{m-1} \frac{|\lambda|^{-j}}{j!}. \tag{2.30}$$

For this purpose, noting that for $k = 0, 1, \cdots$,

$$\Lambda_i^k = \begin{bmatrix} \lambda^k & \binom{k}{1}\lambda^{k-1} & \cdots & \binom{k}{m-1}\lambda^{k-m+1} \\ 0 & \lambda^k & \cdots & \binom{k}{m-2}\lambda^{k-m+2} \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda^k \end{bmatrix},$$

let $v \in \mathbb{C}^m$ be such that $\|v\|_\infty = 1$. For $\ell = 1, \cdots, m$, the ℓ-th coordinate of $\Lambda_i^t v$ is $\sum_{j=0}^{m-\ell} \binom{t}{j} \lambda^{t-j} v_{j+\ell+1}$, which, because of $\binom{t}{j} \leq \frac{t^j}{j!}$, is at most the right hand side of (2.30). Therefore, because of $\Lambda^t = \operatorname{diag}(\Lambda_1^t, \cdots, \Lambda_k^t)$, we have $|||\Lambda^t|||_\infty \leq \eta_t(\Lambda)$. Now, by $x(t) = D^t x(0) + \sum_{i=1}^{t} D^{t-i} w(i)$, by Lemma 2.6, on the event $\mathcal{W}$ we have

$$\begin{aligned} \|x(t)\|_2 &= \left\| P^{-1}\Lambda^t P x(0) + \sum_{i=1}^{t} P^{-1}\Lambda^{t-i} P w(i) \right\|_2 \\ &\leq |||P^{-1}|||_{\infty \to 2} \left(|||\Lambda^t|||_\infty \|Px(0)\|_\infty + \sum_{i=1}^{t} \left\|\Lambda^{t-i} P w(i)\right\|_\infty \right) \\ &\leq \pi_n(\delta). \end{aligned}$$

□

Lemma 2.8 (Noise empirical covariance)**.** Define $C_n = \frac{1}{n}\sum_{i=1}^{n} w(i)w(i)'$, and assume

$$\frac{n}{\nu_n(\delta)^2} \geq \frac{6\left|\lambda_{\max}(C)\right| + 2\epsilon}{3\epsilon^2} p \log\left(\frac{2p}{\delta}\right). \tag{2.31}$$

On the event $\mathcal{W}$, we have

$$\mathbb{P}\left(\left|\lambda_{\max}(C_n - C)\right| > \epsilon\right) \leq \delta.$$

Proof of Lemma 2.8. In this proof, we use the following Matrix Bernstein inequality [30]:

Lemma 2.9 (Matrix Bernstein)**.** Let $X_i \in \mathbb{R}^{p\times p}, i = 1, \cdots, n$ be a sequence of independent symmetric random matrices. Assume for all $i = 1, \cdots, n$, we have $\mathbb{E}[X_i] = 0$ and $\left|\lambda_{\max}(X_i)\right| \leq \Delta$. Then, for all $y \geq 0$ we have

$$\mathbb{P}\left(\left|\lambda_{\max}\left(\sum_{i=1}^{n} X_i\right)\right| \geq y\right) \leq 2p\text{exp}\left(-\frac{3y^2}{6\sigma^2 + 2\Delta y}\right),$$

where $\sigma^2 = \left|\lambda_{\max}\left(\sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]\right)\right|$.

Letting $X_i = w(i)w(i)' - C$, and $\Delta = \max_{1\leq i\leq n}\|w(i)\|_2^2$, clearly $\mathbb{E}[X_i] = 0$, and

$$\begin{aligned}\sigma^2 &= \left|\lambda_{\max}\left(\sum_{i=1}^{n}\mathbb{E}\left[X_i^2\right]\right)\right| \\ &\leq \sum_{i=1}^{n}\left|\lambda_{\max}\left(\mathbb{E}\left[\|w(i)\|_2^2 w(i)w(i)'\right] - C^2\right)\right| \leq n\Delta\left|\lambda_{\max}(C)\right|.\end{aligned}$$

On $\mathcal{W}$, we have $\Delta \leq p\nu_n(\delta)^2$. Therefore, by (2.31) we have

$$\begin{aligned}\mathbb{P}\left(\left|\lambda_{\max}\left(C_n - C\right)\right| > \epsilon\right) &= \mathbb{P}\left(\left|\lambda_{\max}\left(\sum_{i=1}^{n} X_i\right)\right| > n\epsilon\right) \\ &\leq 2p\exp\left(-\frac{3n\epsilon^2}{6\Delta\left|\lambda_{\max}\left(C\right)\right| + 2\Delta\epsilon}\right) \leq \delta.\end{aligned}$$

□

Lemma 2.10 (Average cross product matrix)**.** Assume $|\lambda_{\max}(D)| < 1$, and define

$$U_n = \frac{1}{n}\sum_{i=0}^{n-1}\left[Dx(i)w(i+1)' + w(i+1)x(i)'D'\right].$$

Assume in addition

$$\frac{n}{\|\!|D|\!\|_2^2 \nu_n(\delta)^2 \pi_n(\delta)^2} \geq \frac{32p}{\epsilon^2}\log\left(\frac{2p}{\delta}\right). \tag{2.32}$$

On the event $\mathcal{W}$, we have

$$\mathbb{P}\left(\left|\lambda_{\max}\left(U_n\right)\right| > \epsilon\right) \leq \delta.$$

Proof of Lemma 2.10. In this proof, we use the following Matrix Azuma inequality [30]:

Lemma 2.11 (Matrix Azuma)**.** Let $X_i \in \mathbb{R}^{p\times p}, i = 1, \cdots, n$ be a martingale difference sequence of symmetric matrices, i.e. for some filter $\{\mathcal{F}_i\}_{i=0}^n$, X_i is $\mathcal{F}_i$-measurable and $\mathbb{E}\left[X_{i+1}|\mathcal{F}_i\right] = 0$. Assume for fixed symmetric matrices $M_i, i = 1, \cdots, n$, all matrices $M_i^2 - X_i^2$ are positive semidefinite. Then, for all $y \geq 0$ we have

$$\mathbb{P}\left(\left|\lambda_{\max}\left(\sum_{i=1}^{n} X_i\right)\right| \geq y\right) \leq 2p\exp\left(-\frac{y^2}{8\sigma^2}\right),$$

where $\sigma^2 = \left|\lambda_{\max}\left(\sum_{i=1}^{n} M_i^2\right)\right|$.

Letting

$$\begin{aligned} X_i &= Dx(i-1)w(i)' + w(i)x(i-1)'D', \\ \mathcal{F}_i &= \sigma\left(w(1), \cdots, w(i)\right), \\ M_i &= 2p^{1/2}\nu_n(\delta)\,\pi_n(\delta)\,\|\!|D|\!\|_2 I_p, \end{aligned}$$

clearly, $\mathbb{E}\left[X_{i+1}|\mathcal{F}_i\right] = 0$. Further, $M_i^2 - X_i^2$ is positive semidefinite, since by Lemma 2.6, and Lemma 2.7, on $\mathcal{W}$ we have

$$\begin{aligned} \max_{1\le i\le n} \|w(i)\|_2 &\le p^{1/2}\nu_n(\delta), \\ \max_{0\le i\le n-1} \|x(i)\|_2 &\le \pi_n(\delta). \end{aligned}$$

Therefore, $\sigma^2 = 4np\|\!|D|\!\|_2^2 \nu_n(\delta)^2\,\pi_n(\delta)^2$, and by (2.32) we have

$$\begin{aligned} \mathbb{P}\left(|\lambda_{\max}(U_n)| > \epsilon\right) &= \mathbb{P}\left(\left|\lambda_{\max}\left(\sum_{i=1}^{n} X_i\right)\right| > n\epsilon\right) \\ &\le 2p\exp\left(-\frac{n\epsilon^2}{32p\|\!|D|\!\|_2^2 \nu_n(\delta)^2\,\pi_n(\delta)^2}\right) \le \delta. \end{aligned}$$

□

Proof of Theorem 2.1. First, by the dynamics equation $x(t+1) = Dx(t) + w(t+1)$, we have

$$\begin{aligned} V_{n+1} &= x(0)x(0)' + \sum_{i=0}^{n-1}\left(Dx(i) + w(i+1)\right)\left(Dx(i) + w(i+1)\right)' \\ &= x(0)x(0)' + D\sum_{i=0}^{n-1} x(i)x(i)'D' \\ &+ \sum_{i=0}^{n-1}\left[Dx(i)w(i+1)' + w(i+1)x(i)'D'\right] + \sum_{i=1}^{n} w(i)w(i)' \\ &= DV_{n+1}D' + nU_n + nC_n + D\left(x(0)x(0)' - x(n)x(n)'\right)D' + x(0)x(0)', \end{aligned}$$

where C_n, U_n are defined in Lemma 2.8, and Lemma 2.10, respectively. Letting

$$E_n = U_n + C_n + \frac{1}{n} D \left(x(0)x(0)' - x(n)x(n)' \right) D' + \frac{1}{n} x(0)x(0)',$$

since $|\lambda_{\max}(D)| < 1$, the Lyapunov equation $V_{n+1} = DV_{n+1}D' + nE_n$ has the solution

$$V_{n+1} = n \sum_{i=0}^{\infty} D^i E_n D'^i, \tag{2.33}$$

i.e. $|\lambda_{\min}(V_{n+1})| \geq n \, |\lambda_{\min}(E_n)|$. Henceforth in the proof, we assume the event $\mathcal{W}$ holds. According to Lemma 2.8, (2.4) implies that

$$\mathbb{P}\left(|\lambda_{\max}(C_n - C)| > \frac{\epsilon}{3} \right) \leq \frac{\delta}{2}. \tag{2.34}$$

In addition, by Lemma 2.10, (2.5) implies that

$$\mathbb{P}\left(|\lambda_{\max}(U_n)| > \frac{\epsilon}{3} \right) \leq \frac{\delta}{2}. \tag{2.35}$$

Finally, using Lemma 2.7, by (2.6) we get

$$\frac{1}{n} \left(|||D|||_2^2 + 1 \right) \left(\|x(0)\|_2^2 + \|x(n)\|_2^2 \right) \leq \frac{\epsilon}{3}. \tag{2.36}$$

Putting (2.34), (2.35), and (2.36) together,

$$\begin{aligned} |\lambda_{\max}(E_n - C)| &\leq |\lambda_{\max}(C_n - C)| + |\lambda_{\max}(U_n)| \\ &+ \frac{1}{n} \left(|||D|||_2^2 + 1 \right) \left(\|x(0)\|_2^2 + \|x(n)\|_2^2 \right) \\ &\leq \epsilon, \end{aligned}$$

i.e. on the event $\mathcal{W}$, with probability at least $1-\delta$, it holds that

$$|\lambda_{\min}(E_n)| \geq |\lambda_{\min}(C)| - \epsilon.$$

Substituting in (2.33), we get the desired result. Moreover, since $|\lambda_{\max}(E_n)| \leq |\lambda_{\max}(C)| + \epsilon$, we have

$$\begin{aligned}
\left|\lambda_{\max}\left(\frac{1}{n}V_{n+1}\right)\right| &= \left|\lambda_{\max}\left(\sum_{i=0}^{\infty} D^i E_n D'^i\right)\right| \\
&\leq \sum_{i=0}^{\infty} \left|\lambda_{\max}\left(D^i E_n D'^i\right)\right| \\
&\leq \left(|\lambda_{\max}(C)| + \epsilon\right) \sum_{i=0}^{\infty} \left|\left|\left| D'^i \right|\right|\right|_2^2 \\
&\leq \left(|\lambda_{\max}(C)| + \epsilon\right) \eta\left(D'\right)^2 .
\end{aligned}$$

Thus, when $2\epsilon = |\lambda_{\min}(C)|$, with probability at least $1-2\delta$ we have

$$\left|\lambda_{\max}\left(\frac{1}{n}V_{n+1}\right)\right| \leq \frac{3}{2}|\lambda_{\max}(C)|\, \eta\left(D'\right)^2 . \tag{2.37}$$

When $n \to \infty$, the conditions hold for arbitrary ϵ, δ. So, $|\lambda_{\min}(E_n)| \to |\lambda_{\min}(C)|$, which implies the desired result. □

Proof of Lemma 2.3. First, since $n \geq N_{2.1}\left(\frac{|\lambda_{\min}(C)|}{2}, \delta\right) + 1$, by the proof of Theorem 2.1, on the event $\mathcal{W}$, with probability at least $1-\delta$, we have

$$|\lambda_{\min}(V_n)| \geq \frac{|\lambda_{\min}(C)|}{2}(n-1).$$

Then, as long as V_n is nonsingular, one can write

$$\hat{D}_n - D = \left(\sum_{t=0}^{n-1} w(t+1)x(t)'\right) V_n^{-1},$$

which yields

$$\left(\hat{D}_n - D\right) V_n \left(\hat{D}_n - D\right)' = U_n' V_n^{-1} U_n,$$

where $U_n = \sum_{t=0}^{n-1} x(t)w(t+1)'$. Therefore,

$$\left|\!\left|\!\left| \left(\hat{D}_n - D\right) V_n \left(\hat{D}_n - D\right)' \right|\!\right|\!\right|_2 \leq \frac{|\!|\!|U_n|\!|\!|_2^2}{|\lambda_{\min}(V_n)|}. \tag{2.38}$$

To proceed, for arbitrary matrix $H \in \mathbb{R}^{k \times \ell}$, define the linear transformation

$$\Phi(H) = \begin{bmatrix} 0_{k\times k} & H \\ H' & 0_{\ell \times \ell} \end{bmatrix} \in \mathbb{R}^{(k+\ell)\times(k+\ell)}.$$

As a well known fact, the equality $|\!|\!|H|\!|\!|_2 = |\lambda_{\max}(\Phi(H))|$ holds [30]. Note that $\Phi(H)$ is always symmetric. Next, letting $X_t = x(t)w(t+1)'$, apply Lemma 2.11 to $\Phi(X_t) \in \mathbb{R}^{2p\times 2p}$. Since

$$\Phi(X_t)^2 = \begin{bmatrix} \|w(t+1)\|_2^2 x(t)x(t)' & 0_{p\times p} \\ 0_{p\times p} & \|x(t)\|_2^2 w(t+1)w(t+1)' \end{bmatrix},$$

by Lemma 2.6, and Lemma 2.7, all matrices $\Phi(M_t)^2 - \Phi(X_t)^2$ are positive semidefinite on the event $\mathcal{W}$, where

$$M_t = p^{1/2}\nu_n(\delta)\,\pi_n(\delta)\, I_p.$$

By

$$\sigma^2 = \left|\lambda_{\max}\left(\sum_{t=0}^{n-1} \Phi(M_t)^2\right)\right| = np\nu_n(\delta)^2\,\pi_n(\delta)^2,$$

letting $y = 8^{1/2}\sigma \log^{1/2}\left(\frac{2p}{\delta}\right)$, Lemma 2.11 implies

$$\mathbb{P}\left(|\!|\!|U_n|\!|\!|_2 > y\right) = \mathbb{P}\left(|\lambda_{\max}(\Phi(U_n))| > y\right) \leq 2p\exp\left(-\frac{y^2}{8\sigma^2}\right) = \delta.$$

Plugging in (2.38), we get the desired result. □

Proof of Theorem 2.2. Similar to the proof of Lemma 2.3, condition (2.8) implies that on the event $\mathcal{W}$, $|\lambda_{\min}(V_n)| \geq \frac{|\lambda_{\min}(C)|}{2}(n-1)$, with probability at least $1-\delta/2$. Further, by $\hat{D}_n - D = U_n V_n^{-1}$, we get

$$\left|\left|\left|\hat{D}_n - D\right|\right|\right|_2 \leq \frac{|||U_n|||_2}{|\lambda_{\min}(V_n)|}. \tag{2.39}$$

Then, $\sigma^2 = np\nu_n(\delta)^2 \pi_n(\delta)^2$, and letting $y = \frac{|\lambda_{\min}(C)|}{2}(n-1)\epsilon$, according to Lemma 2.11, condition (2.9) implies

$$\mathbb{P}\left(|||U_n|||_2 > y\right) \leq 2p\exp\left(-\frac{y^2}{8\sigma^2}\right) \leq \frac{\delta}{2}.$$

Plugging in (2.39), we get the desired result. □

2.5.3 Proofs of Section 2.4

Proof of Lemma 2.4. Let the event $\mathcal{G}$ be irregularity of $D = A_0 + B_0 L$. The statement we prove is slightly stronger than regularity. Indeed, we prove that for all $\lambda \in \mathbb{C}$, $|\lambda| \geq 1$, with probability one,

$$\operatorname{rank}(D - \lambda I_p) \geq p - 1.$$

Note that the recent result implies that $\mathbb{P}(\mathcal{G}) = 0$.

First, let $Y_i \in \mathbb{R}^m, i = 1, \cdots, m$ have linearly full rank distributions. Define $Y = [Y_1, \cdots, Y_m]$, and let $M(\lambda)$ be a $m \times m$ matrix, with all coordinates being real polynomials of λ. Let $f(\lambda)$ be a real polynomial of λ as well. We show that

$$\mathbb{P}\left(\exists \lambda \in \mathbb{C}, f(\lambda) \neq 0 : \operatorname{rank}\left(Y - \frac{1}{f(\lambda)}M(\lambda)\right) < m-1\right) = 0. \tag{2.40}$$

If

$$\operatorname{rank}\left(Y - \frac{1}{f(\lambda_0)} M(\lambda_0)\right) < m - 1,$$

letting $\frac{1}{f(\lambda_0)} M(\lambda_0) = [M_1, \cdots, M_m]$, two of the vectors $Y_i - M_i, i = 1, \cdots, m$, such as $Y_{m-1} - M_{m-1}, Y_m - M_m$, can be written as linear combinations of the others. There are finitely many values of λ_0 for which $Y_{m-1} - M_{m-1}$ is a linear combination of $Y_1 - M_1, \cdots, Y_{m-2} - M_{m-2}$, since for every such a λ_0, $\det\left(\tilde{Y}\right) = 0$, where $\tilde{Y}$ is the square matrix whose columns are $Y_1 - M_1, \cdots, Y_{m-1} - M_{m-1}$, removing an arbitrary row. Note that $\det\left(\tilde{Y}\right)$ is a polynomial of λ_0, divided by $f(\lambda_0)$, and $f(\lambda_0) \neq 0$.

Note that λ_0 is a deterministic function of $Y_1, \cdots, Y_m$. For every such a λ_0, the dimension of the subspace $\mathcal{P}$ spanned by $Y_1 - M_1, \cdots, Y_{m-2} - M_{m-2}, M_m$ is at most $m - 1$. Because Y_m is independent of $Y_1, \cdots, Y_{m-1}$, and Y_m has a linearly full rank distribution, $\mathbb{P}(Y_m \in \mathcal{P}) = 0$, i.e. (2.40) holds.

Now, let $m = \operatorname{rank}(B_0)$. If $m = p$, applying the above argument to

$$Y = D, M(\lambda) = \lambda I_p, f(\lambda) = 1,$$

we have $\mathbb{P}(\mathcal{G}) = 0$, since full rankness of B_0 implies linearly full rank distributions for all columns of $B_0 L$. If $m < p$, there is a $p \times p$ permutation matrix J, and $K \in \mathbb{R}^{(p-m)\times m}$, such that

$$JB_0 = \begin{bmatrix} \tilde{B} \\ K\tilde{B} \end{bmatrix} = \begin{bmatrix} I_m \\ K \end{bmatrix} \tilde{B},$$

where $\tilde{B} \in \mathbb{R}^{m\times r}$ is full rank. Let L_0 be a stabilizer, $D_0 = A_0 + B_0 L_0$, and

$$JD_0 = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}, D_1 \in \mathbb{R}^{m\times p}, D_2 \in \mathbb{R}^{(p-m)\times p},$$

to get

$$J\left(A_0 + B_0 L\right) = JD_0 + JB_0\left(L - L_0\right) = \begin{bmatrix} D_1 + \tilde{B}\left(L - L_0\right) \\ D_2 + K\tilde{B}\left(L - L_0\right) \end{bmatrix}.$$

Writing $J = \begin{bmatrix} J_1 \\ J_2 \end{bmatrix}$ similarly, we have

$$\begin{aligned}
\operatorname{rank}\left(A_0 + B_0 L - \lambda I_p\right) &= \operatorname{rank}\left(J\left(A_0 + B_0 L\right) - \lambda J\right) \\
&= \operatorname{rank}\left(\begin{bmatrix} I_m & 0_{m\times(p-m)} \\ -K & I_{p-m} \end{bmatrix}\left(J\left(A_0 + B_0 L\right) - \lambda J\right)\right) \\
&= \operatorname{rank}\left(\begin{bmatrix} D_1 + \tilde{B}\left(L - L_0\right) - \lambda J_1 \\ \left[-K, I_{p-m}\right] J\left(D_0 - \lambda I_p\right) \end{bmatrix}\right).
\end{aligned}$$

Call the last matrix above $\tilde{X}$. Since $\left|\lambda_{\max}\left(D_0\right)\right| < 1$, for $|\lambda| \geq 1$ the matrix $D_0 - \lambda I_p$ is full rank. Therefore, because of

$$\operatorname{rank}\left(\left[-K, I_{p-m}\right]\right) = p - m,$$

we have

$$\operatorname{rank}\left(\left[-K, I_{p-m}\right] J\left(D_0 - \lambda I_p\right)\right) = p - m.$$

Rearrange columns of matrix $\tilde{X}$ to get

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}, X_{11} \in \mathbb{C}^{m\times m}, X_{22} \in \mathbb{C}^{(p-m)\times(p-m)}, \operatorname{rank}\left(X_{22}\right) = p - m.$$

In other words, $p-m$ linearly independent columns of $\left[-K, I_{p-m}\right] J\left(D_0 - \lambda I_p\right)$ have been put together to form X_{22}. If D is not regular,

$$\begin{aligned}
p-2 &\geq \operatorname{rank}\left(\tilde{X}\right) \\
&= \operatorname{rank}\left(X\right) \\
&= \operatorname{rank}\left(X \begin{bmatrix} I_m & 0_{m\times(p-m)} \\ -X_{22}^{-1}X_{21} & I_{p-m} \end{bmatrix}\right) \\
&= \operatorname{rank}\left(\begin{bmatrix} X_{11}-X_{12}X_{22}^{-1}X_{21} & X_{12} \\ 0_{(p-m)\times m} & X_{22} \end{bmatrix}\right).
\end{aligned}$$

Hence,

$$\operatorname{rank}\left(X_{11}-X_{12}X_{22}^{-1}X_{21}\right) \leq m-2.$$

Remember that columns of $[X_{11}, X_{12}]$ are exactly the same as $D_1 + \tilde{B}\left(L-L_0\right) - \lambda J_1$, and all coordinates of $\det\left(X_{22}\right)X_{12}X_{22}^{-1}X_{21}$ are polynomials of λ (since all coordinates of $\det\left(X_{22}\right)X_{22}^{-1}$ are polynomials of the coordinates of X_{22}). Taking

$$f\left(\lambda\right) = \det\left(X_{22}\right),$$

by (2.40), since full rankness of $\tilde{B}$ implies linearly full rank distributions for all columns of $\tilde{B}\left(L-L_0\right)$, we have

$$\mathbb{P}\left(\operatorname{rank}\left(X_{11}-X_{12}X_{22}^{-1}X_{21}\right) \leq m-2\right) = 0,$$

which is the desired result since $\operatorname{rank}\left(X_{22}\right) = p-m$. □

Proof of Lemma 2.5. Assume $D = A_0 + B_0 L$ has a unit-root eigenvalue, denoted by $\lambda \in \mathbb{C}, |\lambda| = 1$. Assuming $m = \operatorname{rank}\left(B_0\right)$, let permutation matrix J and $K \in \mathbb{R}^{(p-m)\times m}$

be such that

$$JB_0 = \begin{bmatrix} \tilde{B} \\ K\tilde{B} \end{bmatrix} = \begin{bmatrix} I_m \\ K \end{bmatrix} \tilde{B},$$

where $\tilde{B} \in \mathbb{R}^{m\times r}$ is full rank. Letting L_0 be a stabilizer, $D_0 = A_0 + B_0 L_0$, and $X = \tilde{B}(L - L_0) \in \mathbb{R}^{m\times p}$, note that X has a general full rank distribution, thanks to full rankness of $\tilde{B}$. Since D_0 is stable, $\det(D_0 - \lambda I_p) \neq 0$, and

$$\begin{aligned} 0 &= \det(A_0 + B_0 L - \lambda I_p) \\ &= \det\left(JD_0 + \begin{bmatrix} I_m \\ K \end{bmatrix} X - \lambda J \right) \\ &= \det\left((D_0 - \lambda I_p)^{-1} J^{-1} \begin{bmatrix} I_m \\ K \end{bmatrix} X + I_p \right) \\ &= \det\left(X (D_0 - \lambda I_p)^{-1} J^{-1} \begin{bmatrix} I_m \\ K \end{bmatrix} + I_m \right), \end{aligned}$$

where the last equality above is implied by Sylvester's determinant identity. Denote the complex conjugate of λ by $\bar{\lambda}$, and define the real matrix

$$M(\lambda) = M(\bar{\lambda}) = \left(D_0 - \bar{\lambda} I_p\right)^{-1} (D_0 - \lambda I_p)^{-1} J^{-1} \begin{bmatrix} I_m \\ K \end{bmatrix}.$$

Further, define the space of eigenvectors in $\mathbb{C}^m$ as follows. First, consider the equivalence relation $\sim$ on $\mathbb{C}^m$, defined as

$$x \sim y, \text{ if } x = cy \text{ for some } c \in \mathbb{C}, c \neq 0.$$

Letting $S = \frac{\mathbb{C}^m}{\sim}$ be the direction space in $\mathbb{C}^m$, we have $\dim_{\mathbb{C}}(S) = m - 1$, i.e.

$$\dim_{\mathbb{R}}(S) = 2m - 2.$$

Note that for every matrix $Y \in \mathbb{C}^{m \times m}$ and every vector $v \in \mathbb{C}^m$, $Yv = 0$ if and only if $Y\tilde{v} = 0$ for every $\tilde{v} \sim v$. Thus,

$$\det\left(X\left(D_0 - \lambda I_p\right)^{-1} J^{-1} \begin{bmatrix} I_m \\ K \end{bmatrix} + I_m \right) = 0$$

implies that there is $v \in S, v \neq 0$, such that

$$\left(X\left(D_0 - \bar{\lambda} I_p\right) M(\lambda) + I_m\right) v = 0 \tag{2.41}$$

Denote the set of all matrices X satisfying (2.41) by $\mathcal{X}(\lambda, v) \subset \mathbb{R}^{m \times p}$. Separating real ($\Re$) and imaginary ($\Im$) parts, we get

$$\begin{aligned} Xa(v) &= \Re(v), \\ Xb(v) &= \Im(v), \end{aligned}$$

where for $v \in S$, vectors $a(v), b(v) \in \mathbb{R}^p$ are defined as

$$\begin{aligned} a(v) &= M(\lambda)\Re\left(\bar{\lambda} v\right) - D_0 M(\lambda)\Re(v), \\ b(v) &= M(\lambda)\Im\left(\bar{\lambda} v\right) - D_0 M(\lambda)\Im(v). \end{aligned}$$

Now, we partition S to

$$S = S_1 \bigcup S_2, S_1 \bigcap S_2 = \emptyset,$$

where

$$S_1 = \{v \in S : a(v), b(v) \text{ are in-line }\},$$
$$S_2 = \{v \in S : a(v), b(v) \text{ are not in-line }\}.$$

Whenever $v \in S_2$, for $j = 1, \cdots, m$, the j-th row of X needs to be in the intersection of two nonparallel hyperplanes $\mathcal{P}_1, \mathcal{P}_2 \subset \mathbb{R}^p$, where

$$\mathcal{P}_1 = \{y \in \mathbb{R}^p : y'a(v) = \Re(v_j)\},$$
$$\mathcal{P}_2 = \{y \in \mathbb{R}^p : y'b(v) = \Im(v_j)\}.$$

Since $\dim_{\mathbb{R}}(\mathcal{P}_1) \le p-1$, $\dim_{\mathbb{R}}(\mathcal{P}_2) \le p-1$, and $v \in S_2$ we have

$$\dim_{\mathbb{R}}(\mathcal{P}_1 \cap \mathcal{P}_2) \le p-2.$$

Therefore, for $v \in S_2$, we have $\dim_{\mathbb{R}}(\mathcal{X}(\lambda, v)) \le m(p-2)$. Since $\dim_{\mathbb{R}}(|\lambda| = 1) = 1$, using $\dim_{\mathbb{R}}(S_2) \le 2m-2$ we have

$$\dim_{\mathbb{R}}\left(\bigcup_{|\lambda|=1, v\in S_2} \mathcal{X}(\lambda, v)\right) \le 1 + 2m - 2 + m(p-2) = mp - 1. \tag{2.42}$$

On the other hand, for $v \in S_1$, there is a real number, say $\alpha(v)$, such that $b(v) = \alpha(v)a(v)$. Then,

$$\Im(v) = Xb(v) = \alpha(v)Xa(v) = \alpha(v)\Re(v), \tag{2.43}$$

i.e. whenever $v \in S_1$, the vectors $\Re(v), \Im(v)$ are in-line. So, $\dim_{\mathbb{R}}(S_1) = m-1$, and for $v \in S_1$, we have $\mathcal{P}_1 = \mathcal{P}_2$, i.e.

$$\dim_{\mathbb{R}}(\mathcal{X}(\lambda, v)) \le m(p-1).$$

Doing some algebra, we get

$$\begin{aligned} 0 &= \alpha(v)a(v) - b(v) \\ &= \alpha(v)\left(\Re(\lambda) I_p + \alpha(v)\Im(\lambda) I_p - D_0\right) M(\lambda) \Re(v) \\ &- \left(\alpha(v)\Re(\lambda) I_p - \Im(\lambda) I_p - \alpha(v) D_0\right) M(\lambda) \Re(v) \\ &= \left(1 + \alpha(v)^2\right) \Im(\lambda) M(\lambda) \Re(v), \end{aligned}$$

i.e. either $\Im(\lambda) = 0$, or $M(\lambda)\Re(v) = 0$. According to the definition of $M(\lambda)$, the latter case implies $\Re(v) = 0$, which because of (2.43) leads to $v = 0$, and is impossible. So, by $\dim_{\mathbb{R}}(|\lambda| = 1, \Im(\lambda) = 0) = 0$, we have

$$\dim_{\mathbb{R}}\left(\bigcup_{|\lambda|=1,\Im(\lambda)=0} \mathcal{X}(\lambda, v)\right) \leq m - 1 + m(p-1) = mp - 1. \tag{2.44}$$

Writing

$$\mathcal{X} = \bigcup_{|\lambda|=1, v\in S} \mathcal{X}(\lambda, v) \subset \left(\bigcup_{|\lambda|=1, v\in S_2} \mathcal{X}(\lambda, v)\right) \cup \left(\bigcup_{|\lambda|=1,\Im(\lambda)=0} \mathcal{X}(\lambda, v)\right),$$

according to (2.42), (2.44) we have $\dim_{\mathbb{R}}(\mathcal{X}) \leq mp - 1$, and by general full rankness of the distribution of X, $\mathbb{P}(\mathcal{X}) = 0$. □

CHAPTER 3

Reinforcement Learning Algorithms

3.1 Introduction

Once the system is stabilized, an adaptive policy is required to minimize the regret. In this chapter, we present some reinforcement learning algorithms for adaptive control of the LQ systems. Different situations, such as those with and without identifiability assumptions, will be considered. Note that in advance to apply any of the following algorithms, we assume that there is a stabilizing set, provided as an input to the corresponding reinforcement learning algorithm. According Theorem 2.4, a stabilizing set with arbitrary high probability guarantee can be constructed, although, it is not the unique way to stabilize the system. In fact, depending on the application, stabilization can be provided to the user by some additional side information.

If one uses Algorithm 1 in order to find a high probability stabilizing set, the state vector can end up with a large value, since the closed-loop transition matrix is not necessarily stable during Algorithm 1. This will not cause any problem, because the system is fully stabilized now. Indeed, one can let the stabilized system proceed for a while, in order to push down the state vector to a reasonable magnitude.

In the episodic algorithms below, the estimation will be reinforced at the end of every episode. Indeed, the algorithms are based on construction of a sequence of confidence sets, which are constructed according to the estimation results established in Section 2.3. This sequence, will be tightened at the end of every episode, i.e. the provided confidence sets become more and more accurate. According to these sequences, the adaptive control policy will be updated after every episode. We present and prove high probability regret bounds, applying the corresponding algorithm.

First, we provide a high level explanation of the algorithms. Detailed expression of every algorithm will come separately later. Starting with the stabilizing set $\Omega^{(0)}$, we select a parameter denoted by $\tilde{\Theta}^{(1)} \in \Omega^{(0)}$. Depending on the identifiability assumptions, the selection of $\tilde{\Theta}^{(1)}$ can be either arbitrary (see (3.15)), or based on OFU principle, i.e. $\tilde{\Theta}^{(1)}$ is a minimizer of the optimal expected average cost over the corresponding confidence set (see (3.1)).

Then, assuming $\tilde{\Theta}^{(1)}$ is the true parameter the system evolves according to, the algorithm applies the optimal linear feedback $L\left(\tilde{\Theta}^{(1)}\right)$, during the first episode. Once the observations during the first episode are collected, they are used to improve the accuracy of the high probability confidence set. Therefore, $\Omega^{(0)}$ is tightened to get $\Omega^{(1)}$, and the second episode starts by iterating the above procedure, and so on. The lengths of the episodes will be as an increasing fashion, to make every confidence set significantly more accurate than all previous ones.

The intuition behind the proficiency of the OFU principle is as follows. As shown in Section 2.3, applying a linear feedback L, observations of the state vectors will lead only to the accurate estimation of the closed-loop matrix. Letting $\tilde{L} = \begin{bmatrix} I_p \\ L \end{bmatrix}$, the closed-loop transition matrix is $A_0 + B_0 L = \Theta_0 \tilde{L}$. Note that in general, an accurate estimation of $\Theta_0 \tilde{L}$ does not lead to that of Θ_0. In fact, since for arbitrary $\Theta \in \mathbb{R}^{p \times q}$, we have $\Theta \tilde{L} \in \mathbb{R}^{p \times p}$, there is a linear subspace $\mathcal{P} \subset \mathbb{R}^{p \times q}$, such that

$$\left(\Theta_0 + \Theta_1\right) \tilde{L} = \Theta_0 \tilde{L},$$

for all $\Theta_1 \in \mathcal{P}$ (the subspace $\mathcal{P}$ is not trivial because $\dim_{\mathbb{R}} \left(\mathcal{P}\right) \geq pq - p^2 = pr$). Therefore, without an additional side information, approximating Θ_0 is impossible, regardless of the accuracy in the approximation of $\Theta_0 \tilde{L}$.

But, in order to design an adaptive policy to minimize the expected average cost, it is an effective approximation of $L(\Theta_0)$ which is required. More precisely, as long as Θ_1 is available satisfying $L(\Theta_1) = L(\Theta_0)$, one can apply an optimal linear feedback $L(\Theta_1)$, no matter how large $|||\Theta_1 - \Theta_0|||_2$ is. In general, estimation of such a Θ_1 is not possible. Yet, a confidence set in addition to an exact knowledge of the closed-loop dynamics lead to an optimal linear feedback, thanks to the OFU principle.

Lemma 3.1. If $\mathcal{J}^\star(\Theta_1) \leq \mathcal{J}^\star(\Theta_0)$ and $\Theta_1 \tilde{L}(\Theta_1) = \Theta_0 \tilde{L}(\Theta_1)$, then $L(\Theta_1)$ is an optimal linear feedback for the system evolving according to Θ_0.

In other words, applying a linear feedback designed according to optimistically selected parameter Θ_1, as long as the closed-loop matrix is learned absolutely accurate, an optimal control action is automatically provided. Remember that the lengths of the episodes are growing, such that the estimation of the closed-loop matrix becomes more precise at the end of every episode. Thus, the approximation $\Theta_1 \tilde{L}(\Theta_1) \approx \Theta_0 \tilde{L}(\Theta_1)$ is becoming more and more accurate. Rigorous analysis of the discussion above, leads to the high probability near optimal regret bounds.

3.2 General Systems

Now, we explain Algorithm 2, which is designed for general settings, where no identifiability condition is assumed about the true parameter Θ_0. The algorithm takes the inputs

$$\Omega^{(0)} \subset \mathbb{R}^{p \times q}, \quad 6\delta > 0, \quad \gamma > 1,$$

explained below. $\Omega^{(0)}$ is a *bounded* stabilizing set: for every $\Theta \in \Omega^{(0)}$, the system will be stable if the optimal linear feedback of Θ is applied, i.e.

$$\left|\lambda_{\max}\left(\Theta_0 \tilde{L}\left(\Theta\right)\right)\right| < 1.$$

Further, $6\delta > 0$ is the highest probability that the near optimal regret bound fails (see Theorem 3.1). The reinforcement rate γ determines the growth rate of the lengths of the time intervals (episodes) an adaptive policy is applied until being updated (see (3.2)).

Algorithm 2 : Adaptive Control of General Systems

Input: Stabilizing Set $\Omega^{(0)}$, Failure Probability 6δ, Reinforcement Rate $\gamma > 1$.

Let $\tau_0 = 0$
for $i = 1, 2, \cdots$ **do**
  Define $\tilde{\Theta}^{(i)}$ and τ_i according to (3.1) and (3.2), respectively
  while $t < \tau_i$ **do**
    Apply control action $u(t) = L\left(\tilde{\Theta}^{(i)}\right) x(t)$
  end while
  Estimate $\hat{D}^{(i)}$ by (3.3), (3.4)
  Defining $V^{(i)}$ by (3.5), Construct $\Xi^{(i)}$ according to (3.6)
  Update $\Omega^{(i)}$ by (3.7)
end for

The algorithm provides an adaptive policy as follows. For $i = 1, 2, \cdots$, at the beginning of the i-the episode, we apply linear feedback $u(t) = L\left(\tilde{\Theta}^{(i)}\right) x(t)$, where

$$\tilde{\Theta}^{(i)} \in \arg \min_{\Theta \in \Omega^{(i-1)}} \mathcal{J}^{\star}\left(\Theta\right). \tag{3.1}$$

Indeed, based on OFU principle, at the beginning of every episode, the best parameter among the all we are uncertain about is being selected. Note that "the best" parameter is a minimizer of the optimum average cost $\mathcal{J}^{\star}\left(\cdot\right)$.

The length of episode i, which is the time period we apply the adaptive control policy $u(t) = L\left(\tilde{\Theta}^{(i)}\right) x(t)$, is designed according to the following equation. Letting $\tau_0 = 0$, we update the control policy at the end of episode i, i.e. at the time $t = \tau_i$, defined according to

$$\tau_i = \tau_{i-1} + \gamma^{i/q} \left(N_{2.1} \left(\frac{|\lambda_{\min}(C)|}{2}, \frac{\delta}{i^2} \right) + 1 \right), \tag{3.2}$$

where $N_{2.1}(\cdot, \cdot)$ is defined in Section 2.3 by (2.4), (2.5), and (2.6). After the i-th episode, we estimate the closed-loop transition matrix $\Theta_0 \tilde{L}\left(\tilde{\Theta}^{(i)}\right)$ by the following row-wise least-squares estimator:

$$\hat{d}_j^{(i)} = \arg\min_{\theta \in \mathbb{R}^p} \sum_{t=\tau_{i-1}}^{\tau_i - 1} \left(x_j(t+1) - \theta' x(t)\right)^2, \tag{3.3}$$

$$\hat{D}^{(i)} = \left[\hat{d}_1^{(i)}, \cdots, \hat{d}_p^{(i)}\right]'. \tag{3.4}$$

Letting

$$V^{(i)} = \sum_{t=\lceil \tau_{i-1} \rceil}^{\lceil \tau_i \rceil - 1} x(t)x(t)', \tag{3.5}$$

be the empirical covariance matrix of episode i, define the high probability confidence set $\Xi^{(i)}$ by

$$\Xi^{(i)} = \left\{ \Theta \in \mathbb{R}^{p \times q} : \left|\left|\left| V^{(i)\frac{1}{2}} \left(\Theta \tilde{L}\left(\tilde{\Theta}^{(i)}\right) - \hat{D}^{(i)} \right)' \right|\right|\right|_2^2 \le \beta_{\tau_i - \tau_{i-1}} \left(\frac{\delta}{i^2} \right) \right\}, \tag{3.6}$$

where $\beta_n(\delta)$ is defined in Lemma 2.3. Note that according to Lemma 2.3,

$$\mathbb{P}\left(\Theta_0 \in \Xi^{(i)}\right) \ge 1 - \frac{3\delta}{i^2}.$$

Then, at the end of episode i, the confidence set $\Omega^{(i-1)}$ will be updated to

$$\Omega^{(i)} = \Omega^{(i-1)} \bigcap \Xi^{(i)}, \tag{3.7}$$

and episode $i+1$ starts, finding $\tilde{\Theta}^{(i+1)}$ by (3.1), and then iterating all steps described above.

Remark 3.1. The choice of $\tilde{\Theta}^{(i)}$ does not need to be as extreme as (3.1). In fact, it suffices to satisfy

$$\mathcal{J}^{\star}\left(\tilde{\Theta}^{(i)}\right) \leq \left(\tau_i - \tau_{i-1}\right)^{-1/2} + \inf_{\Theta \in \Omega^{(i-1)}} \mathcal{J}^{\star}\left(\Theta\right).$$

The following theorem states that performance of the above adaptive control algorithm is optimal, apart from a logarithmic factor. Compared to $O(\cdot)$, the notation $\tilde{O}(\cdot)$ used below, hides the logarithmic factor.

Theorem 3.1 (Regret bound for general systems)**.** Assuming $\Omega^{(0)}$ is bounded, the following upper bound for the regret of Algorithm 2 holds with probability at least $1 - 6\delta$:

$$\mathcal{R}(T) \leq \tilde{O}\left(T^{1/2}\right)\left(-\log\delta\right)^{\frac{1}{2}+\frac{3}{\alpha}}.$$

A direct consequence of Theorem 3.1 is a generalization of the work of Abbasi-Yadkori and Szepesvári [15], where the uncertainty about the true parameter Θ_0 is limited to a bounded subset of $\mathbb{R}^{p \times q}$, i.e. $\Omega^{(0)}$ is bounded. Note that as mentioned before, Theorem 3.1 is fairly more general than the result of the above paper, because
(i) the controllability assumption is removed,
(ii) the operator norm assumption is relaxed to stabilizability,
(iii) the sub-Gaussian distribution of noise vectors is extended to the sub-Weibull one.

3.3 Weakly Identifiable Systems

In this section, we show that the above reinforcement learning algorithm can be improved under a condition. First, in Theorem 3.1, the stabilizing set $\Omega^{(0)}$ needs to be bounded. Although Algorithm 1 provides a high probability bounded stabilizing set satisfying (2.10), in general, boundedness of $\Omega^{(0)}$ does not require to hold. For instance, considering the system of Example 2.1, sufficient and necessary condition for a linear feedback to be a stabilizer is $2 < L_{23} - L_{13} < 4$. So, in general the user can be provided with an unbounded $\Omega^{(0)}$.

Technically, the situation where the provided stabilizing set can be unbounded is remarkably extensive. Subsequently, we discuss the general case intuitively. First, the domain of function $L(\cdot)$ (which maps the stabilizable Θ to $L(\Theta)$) is $\mathbb{R}^{p\times q}$, while its range is $\mathbb{R}^{r\times p}$. Level sets of $L(\Theta)$ are manifolds of dimension $pq - pr = p^2$. These level sets can be unbounded, i.e. the stabilizing set

$$L^{-1}\left\{L(\Theta) : \left|\!\left|\!\left|\Theta - \Theta_0\right|\!\right|\!\right|_2 \leq \epsilon_0\right\} \subset \mathbb{R}^{p\times q}$$

can be unbounded.

The second issue is regarding the laziness of Algorithm 2. Technically, according to (3.2), the length of every episode of Algorithm 2 is approximately $\gamma^{1/q}$ times larger than the previous one. By Theorem 2.2, accuracy of the estimation scales with square root of the episode length, apart from a logarithmic factor. So, at the end of every episode of Algorithm 2, accuracy of the estimation of the closed-loop matrix improves approximately with rate $\gamma^{1/2q}$. This improvement rate can be small, specifically if q, the number of columns of the parameter matrix Θ_0, is not small. If one substitutes the factor $\gamma^{i/q}$ in (3.2) with γ^i, then a constant of the form γ^q appears in the high probability regret bound of Theorem 3.1.

This slowness of the accuracy improvement, as well as possibly unbounded $\Omega^{(0)}$ discussed above, motivate the modification of Algorithm 2. Particularly, when in the real world adaptive control problem the emphasis is on the situations where the time period $0 \leq t \leq T$ during which the user is interacting with the system is not lengthy. For this purpose, we present Algorithm 3, which provides an adaptive policy, near optimal under weak identifiability condition stated below. Overall, Algorithm 3 outperforms Algorithm 2 in the following ways:

(i) the stabilizing set $\Omega^{(0)}$ can be unbounded,

(ii) improvement rate of the estimation is faster, i.e. the high probability guarantee for regret is better (specially for decent T),

(iii) the number of policy changes (which might affect the performance in real applications) is approximately $2q$ times less.

To proceed, we define weak identifiability. Essentially, it is stating that if the approximation Θ_2 of Θ_0 is accurate in terms of the closed-loop matrix when the linear feedback $L(\Theta_1)$ is applied, it is also accurate if one applies the linear feedback $L(\Theta_2)$.

Definition 3.1 (Weak Identifiability)**.** We say Θ_0 is weakly identifiable, if there is $\Xi^{(0)} \subset \mathbb{R}^{p\times q}$ such that $\Theta_0 \in \Xi^{(0)}$, and for all stabilizable $\Theta_1, \Theta_2 \in \Xi^{(0)}$,

$$\left|\left|\left|(\Theta_2 - \Theta_0)\,\tilde{L}(\Theta_2)\right|\right|\right|_2 \leq \Gamma_{\Theta_0}\left|\left|\left|(\Theta_2 - \Theta_0)\,\tilde{L}(\Theta_1)\right|\right|\right|_2, \tag{3.8}$$

for some constant $\Gamma_{\Theta_0} < \infty$.

Weak identifiability can be implied by some other conditions discussed subsequently. First, consider linear transformations

$$\tilde{L}(\Theta_i) : \mathbb{R}^{p\times q} \to \mathbb{R}^{p\times p},\;\; i = 0, 1, 2$$

mapping $\Theta \in \mathbb{R}^{p\times q}$ to $\Theta\tilde{L}(\Theta_i) \in \mathbb{R}^{p\times p}$. Theoretically, letting $\mathcal{P}_i$ be the null space of $\tilde{L}(\Theta_i)$, (3.8) holds if

$$(\Theta_0 + \mathcal{P}_1) \cap \overline{\Xi^{(0)}} \subset (\Theta_0 + \mathcal{P}_2)\,,$$

where $\overline{\Xi^{(0)}}$ is the closure of $\Xi^{(0)}$. In the sequel, we discuss more intuitive conditions which ensure weak identifiability. The following example illustrates the situations where weak identifiability holds, if a neighborhood and a subspace are available as the side information.

Example 3.1. Suppose that $\mathcal{M}_1$ is a subset, and $\mathcal{M}_2$ is a subspace of $\mathbb{R}^{p\times q}$, such that

$$\sup_{\Theta\in\mathcal{M}_1} |||\Theta - \Theta_0|||_2 \;\; \leq \;\; \epsilon_0, \tag{3.9}$$

$$\mathcal{P}_0 \cap \mathcal{M}_2 \;\; = \;\; \{0\}\,. \tag{3.10}$$

There is $\epsilon_0 > 0$ such that $\Xi^{(0)} = \mathcal{M}_1 \cap \mathcal{M}_2$ satisfies (3.8).

To see that, because $\mathcal{M}_2$ is a subspace, by (3.10), for arbitrary $\Theta_2 \in \mathcal{M}_2$ we have

$$\left|\left|\left|(\Theta_2 - \Theta_0)\, \tilde{L}\,(\Theta_0)\right|\right|\right|_2 \geq \Gamma_0 |||\Theta_2 - \Theta_0|||_2,$$

for some constant $\Gamma_0 > 0$. Then, using Lemma 2.2, by (3.9) we have

$$\left|\left|\left|\tilde{L}\,(\Theta_i) - \tilde{L}\,(\Theta_0)\right|\right|\right|_2 \leq \Gamma_L \epsilon_0,$$

for $i = 1, 2$. So,

$$\begin{aligned} \left|\left|\left|(\Theta_2 - \Theta_0)\, \tilde{L}\,(\Theta_1)\right|\right|\right|_2 &\geq (\Gamma_0 - \Gamma_L \epsilon_0)\, |||\Theta_2 - \Theta_0|||_2, \\ \left|\left|\left|(\Theta_2 - \Theta_0)\, \tilde{L}\,(\Theta_2)\right|\right|\right|_2 &\leq (\Gamma_0 + \Gamma_L \epsilon_0)\, |||\Theta_2 - \Theta_0|||_2, \end{aligned}$$

i.e. if $\Gamma_L \epsilon_0 < \Gamma_0$, the condition (3.8) holds for $\Gamma_{\Theta_0} = \frac{\Gamma_0 + \Gamma_L \epsilon_0}{\Gamma_0 - \Gamma_L \epsilon_0}$.

Therefore, as long as we know that the true parameter Θ_0 is living in a subspace $\mathcal{M}_2$ satisfying (3.10), weak identifiability holds. Note that if one uses Algorithm 1 to stabilize the system, (3.9) can be satisfied.

An application of Example 3.1 is when the true parameter Θ_0 is known to have at most s nonzero entries, for some $s \leq p^2$. Entries of Θ_0 are mostly zeros, for example if the actual dynamics of the system is of smaller dimension with longer memory [9]. Furthermore, assuming sparsity for the dynamics matrices is realistic for example if the system is describing the behavior of a network [19].

To study the case above, for arbitrary $\Theta \in \mathbb{R}^{p \times q}$, let $[\Theta]_{\min}$ be the smallest magnitude of the nonzero entries of $\Theta = [\Theta_{ij}]$, formally defined as

$$[\Theta]_{\min} = \min\{|\Theta_{ij}| : \;\; 1 \leq i \leq p; \;\; 1 \leq j \leq q; \;\; \Theta_{ij} \neq 0\}.$$

First, the set of all $p \times q$ matrices with at most s nonzero entries is a finite union of s dimensional subspaces of $\mathbb{R}^{p\times q}$. Let $\mathcal{M}_2$ be one of these subspaces, which contains Θ_0. If $\epsilon_0 < [\Theta_0]_{\min}$, (3.9) leads to exact identification of $\mathcal{M}_2$ among all subspaces. Namely, any subspace $\mathcal{M}_2$ which has a nonempty intersection with neighborhood $\mathcal{M}_1$, contains Θ_0. Note that if the number of nonzero entries of Θ_0 is exactly s, the subspace $\mathcal{M}_2$ is unique. Otherwise, one can search for $\mathcal{M}_2$ among the subspaces of lower dimensions.

Then, assuming (3.10) holds, Θ_0 is weakly identifiable. The linear transformation $\tilde{L}(\Theta_0)$ is surjective, i.e. $\dim_{\mathbb{R}}(\mathcal{P}_0) = pr$. So, $\dim_{\mathbb{R}}(\mathcal{M}_2) \leq p^2$ is required to satisfy (3.10), i.e. $s \leq p^2$. In the work of Ibrahimi et al. [16] a much stronger identifiability condition is defined and assumed, which implies that both a condition similar to (3.10), as well as the sparsity of at most p, need to hold for every row of the $p \times q$ parameter matrices.

Algorithm 3 : Adaptive Control of Weakly Identifiable Systems
Input: Stabilizing Set $\Omega^{(0)}$, Identifiability Set $\Xi^{(0)}$, Failure Probability 6δ, Precision ϵ, Reinforcement Rate $\gamma > 1$.

Let $\tau_0 = 0$
for $i = 1, 2, \cdots$ **do**
  Define $\tilde{\Theta}^{(i)}$ and τ_i according to (3.11) and (3.12), respectively
  while $t < \tau_i$ **do**
    Apply control action $u(t) = L\left(\tilde{\Theta}^{(i)}\right)x(t)$
  end while
  Estimate $\hat{D}^{(i)}$ by (3.3), (3.4)
  Update $\Omega^{(i)}$ according to (3.13)
end for

When Θ_0 is weakly identifiable, Algorithm 3 provides an adaptive policy of near optimal regret. In addition to

$$\Omega^{(0)} \subset \mathbb{R}^{p\times q}, \quad 6\delta > 0, \quad \gamma > 1,$$

explained before, the inputs include precision $\epsilon > 0$, and identifiability set $\Xi^{(0)}$. Note that despite Algorithm 2, here $\Omega^{(0)}$ can be unbounded. Further, the precision $\epsilon > 0$ is arbitrary and determines the lengths of the episodes, while the growth rate of the episode size is determined by γ (see (3.12)).

Below, we explain Algorithm 3 for weakly identifiable systems, comparing with Algorithm 2 for the general case.

- The linear feedback $u(t) = L\left(\tilde{\Theta}^{(i)}\right) x(t)$, applied during the i-the episode, where $\tilde{\Theta}^{(i)}$ is defined by

$$\tilde{\Theta}^{(i)} \in \arg \min_{\Theta \in \Omega^{(i-1)} \cap \Xi^{(0)}} \mathcal{J}^{\star}(\Theta). \tag{3.11}$$

  Again, similar to Remark 3.1, the choice of $\tilde{\Theta}^{(i)}$ can be relaxed. Note that the set of parameters $\Omega^{(i-1)} \cap \Xi^{(0)}$ over which the minimum is being taken, is different than (3.1).

- The lengths of the episodes, is now determined by

$$\tau_i = \tau_{i-1} + N_{2.2}\left(\frac{\epsilon}{\gamma^i}, \frac{\delta}{i^2}\right), \tag{3.12}$$

  where the function $N_{2.2}(\cdot, \cdot)$ is defined in Section 2.3 according to (2.8), (2.9).

- Estimation of the closed-loop transition matrix remains the same as (3.3), (3.4).

- At the end of episode i, the confidence set $\Omega^{(i-1)}$ will be updated to

$$\Omega^{(i)} = \left\{\Theta \in \mathbb{R}^{p \times q} : \left|\left|\left|\Theta \tilde{L}\left(\tilde{\Theta}^{(i)}\right) - \hat{D}^{(i)}\right|\right|\right|_2 \leq \frac{\epsilon}{\gamma^i}\right\}. \tag{3.13}$$

- Similarly, episode $i+1$ starts by finding $\tilde{\Theta}^{(i+1)}$ according to (3.11), and then iterating all steps described above.

According to Theorem 2.2, for confidence set $\Omega^{(i)}$ defined above we have

$$\mathbb{P}\left(\Theta_0 \in \Omega^{(i)}\right) \geq 1 - \frac{2\delta}{i^2}.$$

The following theorem provides the regret bound of Algorithm 3.

Theorem 3.2 (Regret bound under weak identifiability)**.** If Θ_0 is weakly identifiable, the regret of Algorithm 3 is with probability at least $1-6\delta$ bounded by

$$\mathcal{R}(T) \leq \tilde{O}\left(T^{1/2}\right)(-\log \delta)^{\frac{3}{\alpha}+\frac{1}{2}}.$$

3.4 Strongly Identifiable Systems

Finally, supposing a stronger identifiability condition, certainty equivalence principle holds, and we can remove some steps of the above two algorithms. Namely, the OFU based computation of $\tilde{\Theta}^{(i)}$ in (3.1) and (3.11), as well as the construction of the confidence sets in (3.6), (3.7), and (3.13) can be omitted, if the system is strongly identifiable. This condition, roughly speaking, holds, whenever the dynamics parameter Θ_0 can be estimated, with an accuracy comparable to that of the estimation of the closed-loop transition matrix $\Theta_0 \tilde{L}\left(\tilde{\Theta}^{(i)}\right)$.

Definition 3.2 (Strong Identifiability)**.** Θ_0 is called to be strongly identifiable, if there is $\Xi^{(0)} \subset \mathbb{R}^{p \times q}$ such that $\Theta_0 \in \Xi^{(0)}$, and for all stabilizable $\Theta_1, \Theta_2 \in \Xi^{(0)}$,

$$|||\Theta_2 - \Theta_0|||_2 \leq \Gamma_{\Theta_0} \left|\left|\left|(\Theta_2 - \Theta_0)\,\tilde{L}(\Theta_1)\right|\right|\right|_2, \tag{3.14}$$

for some constant $\Gamma_{\Theta_0} < \infty$.

As we will see later, strongly identifiable systems can be adaptively controlled with near optimal regret, if an *arbitrary* estimation of the dynamics parameter Θ_0 is used to design the adaptive linear feedback.

First, similar to the previous section, there are more intuitive conditions to ensure strong identifiability. Consider linear transformations $\tilde{L}(\Theta_i)$ and the null spaces $\mathcal{P}_i$ defined in the previous section. Theoretically, (3.14) holds if

$$(\Theta_0 + \mathcal{P}_1) \cap \overline{\Xi^{(0)}} = \{\Theta_0\}.$$

Obviously, the strong identifiability implies the weak one according to Lemma 2.2, but the opposite is not true. For example, if

$$\begin{aligned} L(\Theta_1) &= L(\Theta_2), \\ \Theta_2 - \Theta_0 &\in \mathcal{P}_1, \end{aligned}$$

then both sides of (3.8) are zero, while the left-hand side of (3.14) is nonzero if $\Theta_2 \neq \Theta_0$.

Further, the situation of Example 3.1 where a neighborhood and a subspace are available as the side information, implies strong identifiability. To verify that, note that as we saw before,

$$\left|\left|\left| (\Theta_2 - \Theta_0) \tilde{L}(\Theta_1) \right|\right|\right|_2 \geq (\Gamma_0 - \Gamma_L \epsilon_0) \left|\left|\left| \Theta_2 - \Theta_0 \right|\right|\right|_2.$$

So, the condition (3.14) holds for $\Gamma_{\Theta_0} = \Gamma_0 - \Gamma_L \epsilon_0$. Similarly, the situation where Θ_0 has at most s nonzero entries, implies strong identifiability as well.

Algorithm 4 : Adaptive Control of Strongly Identifiable Systems
Input: Stabilizing Set $\Omega^{(0)}$, Identifiability Set $\Xi^{(0)}$, Failure Probability 6δ, Precision ϵ, Reinforcement Rate $\gamma > 1$.

Let $\tau_0 = 0$, and choose $\hat{\Theta}^{(1)} \in \Omega^{(0)} \bigcap \Xi^{(0)}$ arbitrarily
for $i = 1, 2, \cdots$ **do**
  Define τ_i by (3.12)
  while $t < \tau_i$ **do**
    Apply control action $u(t) = L\left(\hat{\Theta}^{(i)}\right) x(t)$
  end while
  Estimating $\hat{D}^{(i)}$ by (3.3), (3.4), update $\hat{\Theta}^{(i+1)}$ by (3.15)
end for

Algorithm 4 is a reinforcement learning one for strongly identifiable systems. Excluding the selection of $\tilde{\Theta}^{(i)}$, other steps are similar to Algorithm 3. Here, during episode i, the linear feedback $u(t) = L\left(\hat{\Theta}^{(i)}\right) x(t)$ is applied, where

$$\hat{\Theta}^{(i+1)} \in \left\{ \Theta \in \Xi^{(0)} : \Theta \tilde{L}\left(\hat{\Theta}^{(i)}\right) = \hat{D}^{(i)} \right\} \tag{3.15}$$

is arbitrarily selected. The following theorem states the non-asymptotic optimality of the CE principle for strongly identifiable systems.

Theorem 3.3 (Regret bound under strong identifiability)**.** If Θ_0 is strongly identifiable, the regret of Algorithm 4 is bounded by

$$\mathcal{R}(T) \leq \tilde{O}\left(T^{1/2}\right)(-\log\delta)^{\frac{3}{\alpha}+\frac{1}{2}},$$

with probability at least $1-6\delta$.

A consequence of Theorem 3.3 is that in the work Ibrahimi et al. [16], the step of the proposed algorithm which is based on OFU principle can be removed. Indeed, as mentioned before, in the above paper it is proven that in high-dimensional setting, under some restrictive identifiability assumptions, the dynamics parameter Θ_0 can be estimated accurately, with high probability. This open-loop identification, which is similar to (3.14), leads to optimality of CE principle.

We finish this chapter with a short explanation about the situation of uniformly bounded noise vectors, as well as the asymptotic behavior of the presented reinforcement learning algorithm.

Remark 3.2. The logarithmic factors becomes double logarithmic for bounded noise. More precisely, in Theorem 3.2 and Theorem 3.3, if the noise sequence is uniformly bounded, we get

$$\mathcal{R}(T) = O\left(T^{1/2}\log\log T\right)(-\log\delta)^{\frac{1}{2}}.$$

Moreover, asymptotic analysis of algorithms 2, 3, and 4 shows that

$$\mathbb{P}\left(\lim_{T\to\infty}\left|\frac{\mathcal{R}(T)}{T^{1/2}}\right| < \infty\right) = 1,$$

which is according to Lemma 2.2, optimal.

3.5 Technical Proofs

Proof of Lemma 3.1. Suppose that $\mathcal{J}^\star(\Theta_1) \leq \mathcal{J}^\star(\Theta_0)$ and $D = \Theta_1 \tilde{L}(\Theta_1) = \Theta_0 \tilde{L}(\Theta_1)$. Applying the linear feedback $u(t) = L(\Theta_1)\, x(t)$ to a system evolving according to the dynamics parameter Θ_0, the closed-loop matrix will be $x(t+1) = Dx(t) + w(t+1)$. Letting $P = Q + L(\Theta_1)' RL(\Theta_1)$, we have

$$\begin{aligned}
\mathbb{E}[c_t] &= \mathbb{E}[x(t)'Px(t)] \\
&= \mathbb{E}[x(t-1)'D'PDx(t-1)] + \mathbb{E}[w(t)'Pw(t)] \\
&= \cdots = x(0)'D'^t PD^t x(0) + \sum_{i=1}^{t} \mathbb{E}\left[w(i)'D'^{t-i} PD^{t-i} w(i)\right].
\end{aligned}$$

Note that by stabilizability of Θ_0, the inequality $\mathcal{J}^\star(\Theta_1) \leq \mathcal{J}^\star(\Theta_0)$ implies that Θ_1 is stabilizable, i.e. by Proposition 2.1, $|\lambda_{\max}(D)| < 1$. Thus,

$$\lim_{t\to\infty} x(0)'D'^t PD^t x(0) = 0. \tag{3.16}$$

Furthermore, by $\mathbb{E}[w(i)w(i)'] = C$, the second term is $\operatorname{tr}\left(C \sum_{i=0}^{t-1} D'^i PD^i\right)$. Therefore, using (3.16) we get

$$\lim_{t\to\infty} \mathbb{E}[c_t] = \operatorname{tr}\left(C \sum_{i=0}^{\infty} D'^i PD^i\right).$$

The above convergence holds for the Cesaro mean of the sequence $\{\mathbb{E}[c_t]\}_{t=1}^{\infty}$ as well, i.e. the expected average cost is

$$\overline{\mathcal{J}}_{\Theta_0}\left(\{u(t)\}_{t=0}^{\infty}\right) = \operatorname{tr}\left(C \sum_{i=0}^{\infty} D'^i PD^i\right).$$

Similarly, since $u(t) = L(\Theta_1) x(t)$ is optimal for a system of open-loop parameter Θ_1,

$$\begin{aligned}
\mathcal{J}^\star(\Theta_0) \geq \mathcal{J}^\star(\Theta_1) &= \operatorname{tr}\left(C \sum_{i=0}^{\infty} D'^{i} P D^{i}\right) \\
&= \overline{\mathcal{J}}_{\Theta_0}\left(\{u(t)\}_{t=0}^{\infty}\right) \geq \mathcal{J}^\star(\Theta_0),
\end{aligned}$$

i.e. the linear feedback $L(\Theta_1)$ is an optimal policy for a system of dynamics parameter Θ_0, which is the desired result. $\square$

Proof of Theorem 3.1. The stabilizing set $\Omega^{(0)}$ is assumed to be bounded, so let

$$\Delta_0 = \sup_{\Theta \in \Omega^{(0)}} |||\Theta'|||_2 < \infty. \tag{3.17}$$

Suppose that for $t = 1, 2, \cdots$, the parameter Θ_t is being used to design the adaptive linear feedback $u(t) = L(\Theta_t) x(t)$. So, during every episode, Θ_t does not change, and for $\tau_{i-1} \leq t < \tau_i$ we have $\Theta_t = \tilde{\Theta}^{(i)}$.

Letting $\mathcal{F}_t = \sigma(w(1), \cdots, w(t))$, the infinite horizon dynamic programming equations [17] are

$$\begin{aligned}
\mathcal{J}^\star(\Theta_t) + x(t)' K(\Theta_t) x(t) &= x(t)' Q x(t) \\
&+ u(t)' R u(t) + \mathbb{E}\left[y(t+1)' K(\Theta_t) y(t+1) \middle| \mathcal{F}_t\right],
\end{aligned}$$

where $u(t) = L(\Theta_t) x(t)$, and

$$y(t+1) = A_t x(t) + B_t u(t) + w(t+1) = \Theta_t \tilde{L}(\Theta_t) x(t) + w(t+1) \tag{3.18}$$

is the desired dynamics of the system. Note that since the true evolution of the system is governed by Θ_0, the next state is in fact

$$x(t+1) = A_0 x(t) + B_0 u(t) + w(t+1) = \Theta_0 \tilde{L}(\Theta_t) x(t) + w(t+1). \tag{3.19}$$

Substituting (3.18), and (3.19) in the dynamic programming equation, and using (1.2) for the instantaneous cost c_t, we have

$$
\begin{aligned}
& \mathcal{J}^\star(\Theta_t) + x(t)'K(\Theta_t)\,x(t) \\
= \;& c_t + \mathbb{E}\left[y(t+1)'K(\Theta_t)\,y(t+1)\middle|\mathcal{F}_t\right] \\
= \;& c_t + \mathbb{E}\left[w(t+1)'K(\Theta_t)\,w(t+1)\middle|\mathcal{F}_t\right] \\
+ \;& x(t)\tilde{L}(\Theta_t)'\,\Theta_t'K(\Theta_t)\,\Theta_t\tilde{L}(\Theta_t)\,x(t) \\
= \;& c_t + \mathbb{E}\left[w(t+1)'K(\Theta_t)\,w(t+1)\middle|\mathcal{F}_t\right] \\
+ \;& x(t)\tilde{L}(\Theta_t)'\,\Theta_0'K(\Theta_t)\,\Theta_0\tilde{L}(\Theta_t)\,x(t) \\
+ \;& x(t)\tilde{L}(\Theta_t)'\,\Theta_t'K(\Theta_t)\,\Theta_t\tilde{L}(\Theta_t)\,x(t) \\
- \;& x(t)\tilde{L}(\Theta_t)'\,\Theta_0'K(\Theta_t)\,\Theta_0\tilde{L}(\Theta_t)\,x(t) \\
= \;& c_t + \mathbb{E}\left[x(t+1)'K(\Theta_t)\,x(t+1)\middle|\mathcal{F}_t\right] \\
+ \;& x(t)\tilde{L}(\Theta_t)'\left[\Theta_t'K(\Theta_t)\,\Theta_t - \Theta_0'K(\Theta_t)\,\Theta_0\right]\tilde{L}(\Theta_t)\,x(t).
\end{aligned}
$$

Adding up for $t = 1, \cdots, T$, we get

$$
\mathcal{R}(T) = \sum_{t=1}^{T}\left[c_t - \mathcal{J}^\star(\Theta_0)\right] = \mathbb{T}_1 + \mathbb{T}_2 + \mathbb{T}_3 + \mathbb{T}_4, \tag{3.20}
$$

where

$$
\begin{aligned}
\mathbb{T}_1 &= \sum_{t=1}^{T}\left[\mathcal{J}^\star(\Theta_t) - \mathcal{J}^\star(\Theta_0)\right], \\
\mathbb{T}_2 &= \sum_{t=1}^{T}\left(x(t)'K(\Theta_t)\,x(t) - \mathbb{E}\left[x(t+1)'K(\Theta_{t+1})\,x(t+1)\middle|\mathcal{F}_t\right]\right), \\
\mathbb{T}_3 &= \sum_{t=1}^{T}\mathbb{E}\left[x(t+1)'\left(K(\Theta_{t+1}) - K(\Theta_t)\right)x(t+1)\middle|\mathcal{F}_t\right], \\
\mathbb{T}_4 &= \sum_{t=1}^{T}x(t)'\tilde{L}(\Theta_t)'\left[\Theta_0'K(\Theta_t)\,\Theta_0 - \Theta_t'K(\Theta_t)\,\Theta_t\right]\tilde{L}(\Theta_t)\,x(t).
\end{aligned}
$$

Let $m(T)$ be the number of episodes started until time T. So,

$$\tau_{m(T)} \leq T < \tau_{m(T)+1}.$$

Now, letting $n_i = \lfloor \tau_i - \tau_{i-1} \rfloor$ be the length of episode i, define the following events

$$\begin{aligned}
\mathcal{G} &= \bigcap_{i=1}^{\infty} \left\{ \max_{\tau_{i-1} \leq t < \tau_i} \|w(t)\|_{\infty} \leq \nu_{n_i} \left(\frac{\delta}{i^2} \right) \right\}, \\
\mathcal{H} &= \bigcap_{i=1}^{\infty} \left\{ \Theta_0 \in \Omega^{(i)} \right\}.
\end{aligned}$$

Similar to the proof of Lemma 2.3, one can simply see that

$$\mathbb{P}(\mathcal{G} \cap \mathcal{H}) \geq 1 - \sum_{i=1}^{\infty} \frac{3\delta}{i^2} \geq 1 - 5\delta. \tag{3.21}$$

Henceforth in the proof, we assume that $\mathcal{G} \cap \mathcal{H}$ holds.

Lemma 3.2 (Bounding $\mathbb{T}_1$)**.** On $\mathcal{G} \cap \mathcal{H}$, we have $\mathbb{T}_1 \leq 0$.

Proof. For all $i = 1, 2, \cdots$, as long as $\Theta_0 \in \Omega^{(i-1)}$, according to (3.1) we have $\mathcal{J}^\star\left(\tilde{\Theta}^{(i)}\right) \leq \mathcal{J}^\star(\Theta_0)$, i.e. $\mathcal{J}^\star(\Theta_t) - \mathcal{J}^\star(\Theta_0) \leq 0$, which implies the desired result. $\square$

Lemma 3.3 (Bounding $\mathbb{T}_2$)**.** On $\mathcal{G} \cap \mathcal{H}$, we have

$$\mathbb{P}\left(\mathbb{T}_2 \geq \Delta_2 + (8T)^{1/2} \Delta_3 \left(\log m(T)\right)^{2/\alpha} \left(-\log \delta\right)^{1/2+2/\alpha}\right) \leq \delta,$$

for some constants $\Delta_2, \Delta_3 < \infty$.

Proof. First, write

$$\begin{aligned}
\mathbb{T}_2 &= \sum_{t=1}^{T} x(t)' K\left(\Theta_t\right) x(t) - \sum_{t=2}^{T+1} \mathbb{E}\left[x(t)' K\left(\Theta_t\right) x(t) \middle| \mathcal{F}_{t-1}\right] \\
&= \mathbb{E}\left[x(1)' K\left(\Theta_1\right) x(1)\right] - \mathbb{E}\left[x(T+1)' K\left(\Theta_{T+1}\right) x(T+1) \middle| \mathcal{F}_T\right] \\
&+ \sum_{t=1}^{T} \left(x(t)' K\left(\Theta_t\right) x(t) - \mathbb{E}\left[x(t)' K\left(\Theta_t\right) x(t) \middle| \mathcal{F}_{t-1}\right] \right).
\end{aligned}$$

Then, letting

$$\Delta_1 = \sup_{1 \le i \le \infty} \left\| K\left(\tilde{\Theta}^{(i)}\right) \right\|_2, \tag{3.22}$$

note that the above sequence we are taking supremum on, is bounded because for positive definite matrix C, on $\mathcal{H}$ the OFU principle of (3.1) implies

$$\mathcal{J}^{\star}\left(\Theta_0\right) \ge \mathcal{J}^{\star}\left(\tilde{\Theta}^{(i)}\right) = \operatorname{tr}\left(K\left(\tilde{\Theta}^{(i)}\right) C\right),$$

hence,

$$\begin{aligned}
\mathcal{J}^{\star}\left(\Theta_0\right) &\ge \operatorname{tr}\left(C^{1/2} K\left(\tilde{\Theta}^{(i)}\right) C^{1/2}\right) \\
&\ge \left|\lambda_{\max}\left(C^{1/2} K\left(\tilde{\Theta}^{(i)}\right) C^{1/2}\right)\right| \\
&= \sup_{v \neq 0} \frac{v' K\left(\tilde{\Theta}^{(i)}\right) v}{\|v\|_2^2} \frac{\|v\|_2^2}{v' C^{-1} v} \\
&\ge \left|\lambda_{\min}\left(C\right)\right| \left|\lambda_{\max}\left(K\left(\tilde{\Theta}^{(i)}\right)\right)\right|,
\end{aligned}$$

i.e.

$$\Delta_1 \le \frac{\mathcal{J}^{\star}\left(\Theta_0\right)}{\left|\lambda_{\min}\left(C\right)\right|} < \infty.$$

To proceed, using boundedness of $\Omega^{(0)}$,

$$\mathbb{E}\left[x(1)' K\left(\Theta_1\right) x(1)\right] = x(0) \tilde{L}\left(\Theta_1\right)' \Theta_1' K\left(\Theta_1\right) \Theta_1 \tilde{L}\left(\Theta_1\right) x(0) + \operatorname{tr}\left(K\left(\Theta_1\right) C\right) \le \Delta_2, \tag{3.23}$$

for some $\Delta_2 < \infty$. Defining the stable closed-loop matrices $D_i = \Theta_0 \tilde{L}\left(\tilde{\Theta}^{(i)}\right), i = 1, \cdots, m(T)$, similar to Lemma 2.7, one can simply show that on the event $\mathcal{G}$, for constant $\eta\left(D_1, \cdots, D_{m(T)}\right) < \infty$, it holds that

$$\max_{1 \le t \le T} \|x(t)\|_2 \le \eta\left(D_1, \cdots, D_{m(T)}\right) \nu_T \left(\frac{\delta}{m(T)^2}\right), \tag{3.24}$$

where the fact

$$\max_{1 \le i \le m(T)} \nu_{n_i}\left(\frac{\delta}{i^2}\right) \le \nu_T\left(\frac{\delta}{m(T)^2}\right)$$

is used above. Therefore, for martingale difference sequence

$$\{X_t\}_{t=1}^T = \left\{x(t)' K(\Theta_t) x(t) - \mathbb{E}\left[x(t)' K(\Theta_t) x(t) \middle| \mathcal{F}_{t-1}\right]\right\}_{t=1}^T,$$

on $\mathcal{G}$ we have

$$\begin{aligned} |X_t| &\le 2 \| K(\Theta_t) \|_2 \|x(t)\|_2^2 \le \Delta_1 \eta\left(D_1, \cdots, D_{m(T)}\right)^2 \nu_T\left(\frac{\delta}{m(T)^2}\right)^2 \\ &\le \Delta_3 \left(\log m(T)\right)^{2/\alpha} \left(-\log \delta\right)^{2/\alpha}, \end{aligned}$$

for some $\Delta_3 < \infty$. letting

$$\begin{aligned} \sigma^2 &= \Delta_3^2 T \left(\log m(T)\right)^{4/\alpha} \left(-\log \delta\right)^{4/\alpha} \ge \sum_{t=1}^T X_t^2, \\ y &= (8T)^{1/2} \Delta_3 \left(\log m(T)\right)^{2/\alpha} \left(-\log \delta\right)^{1/2+2/\alpha}, \end{aligned}$$

apply Lemma 2.11 to get

$$\mathbb{P}\left(\sum_{t=1}^T X_t > y\right) \le \exp\left(-\frac{y^2}{8\sigma^2}\right) \le \delta,$$

which in addition to (3.23) implies the desired result, because of

$$\mathbb{E}\left[x(T+1)'K\left(\Theta_{T+1}\right)x(T+1)\middle|\mathcal{F}_T\right] \geq 0.$$

□

Lemma 3.4 (Bounding $\mathbb{T}_3$)**.** On $\mathcal{G} \cap \mathcal{H}$, we have

$$\mathbb{T}_3 \leq \Delta_3 \left(\log m\left(T\right)\right)^{2/\alpha} \left(-\log \delta\right)^{2/\alpha} m\left(T\right),$$

where Δ_3 is the same as Lemma 3.3.

Proof. Note that as long as both of $t, t+1$ are in episode i, we have

$$\Theta_t = \Theta_{t+1} = \tilde{\Theta}^{(i)}.$$

So, using (3.22), and (3.24), on $\mathcal{G} \cap \mathcal{H}$ we have

$$\begin{aligned}
\mathbb{T}_3 &= \sum_{i=1}^{m(T)-1} \mathbb{E}\left[x(\lceil\tau_i\rceil)'\left(K\left(\tilde{\Theta}^{(i+1)}\right) - K\left(\tilde{\Theta}^{(i)}\right)\right)x(\lceil\tau_i\rceil)\middle|\mathcal{F}_t\right], \\
&\leq m\left(T\right) \max_{1\leq i\leq m(T)-1} \left|\left|\left|K\left(\tilde{\Theta}^{(i+1)}\right)\right|\right|\right|_2 \|x(\lceil\tau_i\rceil)\|_2^2 \\
&\leq \Delta_1 m\left(T\right)\eta\left(D_1, \cdots, D_{m(T)}\right)^2 \nu_T\left(\frac{\delta}{m\left(T\right)^2}\right)^2 \\
&\leq \Delta_3\left(\log m\left(T\right)\right)^{2/\alpha}\left(-\log\delta\right)^{2/\alpha} m\left(T\right).
\end{aligned}$$

□

Lemma 3.5. Letting $U_0 = I_q$, for $i = 1, 2, \cdots$ define the symmetric $q \times q$ matrix U_i as

$$U_i = \tilde{L}\left(\tilde{\Theta}^{(i)}\right)V^{(i)}\tilde{L}\left(\tilde{\Theta}^{(i)}\right)' = \tilde{L}\left(\tilde{\Theta}^{(i)}\right)\sum_{t=\tau_{i-1}}^{\tau_i} x(t)x(t)'\tilde{L}\left(\tilde{\Theta}^{(i)}\right)',$$

and for arbitrary nonzero $\Theta \in \mathbb{R}^{p \times q}$, let real-valued sequence $\{s_j(\Theta)\}_{j=1}^{\infty}$ be

$$s_j(\Theta) = \frac{\left\|\Theta \sum_{i=0}^{j} U_i \Theta'\right\|_2}{\left\|\Theta \sum_{i=0}^{j-1} U_i \Theta'\right\|_2}.$$

Note that $s_j(\Theta)$ does not depend on the magnitude of Θ. On the event $\mathcal{G} \cap \mathcal{H}$, the Cesaro mean of the sequence $\{s_j(\Theta)\}_{j=1}^{\infty}$ is bounded, i.e. for some constant Δ_4, on $\mathcal{G} \cap \mathcal{H}$ we have

$$\sup_{n \geq 1} \frac{1}{n} \sum_{j=1}^{n} s_j(\Theta) \leq \Delta_4.$$

Proof. First, applying the second part of Theorem 2.1, we have

$$\lim_{i \to \infty} \frac{1}{n_i} V^{(i)} = \lim_{i \to \infty} \sum_{\ell=0}^{\infty} {D_i}^{\ell} C {D_i'}^{\ell}, \tag{3.25}$$

where $D_i = \Theta_0 \tilde{L}\left(\tilde{\Theta}^{(i)}\right)$ is the stable closed-loop transition matrix during episode i.

Then, the sequence $\left\{L\left(\tilde{\Theta}^{(i)}\right)\right\}_{i=1}^{\infty}$ converges as follows. According to (3.17), it is bounded. So, divergence of this bounded sequence implies convergence of two subsequences to distinct limits. Let L_∞ be the limit point of a subsequence. According to (3.7), $\left\{\Omega^{(i)}\right\}_{i=0}^{\infty}$ is strictly decreasing: $\Omega^{(i+1)} \subsetneq \Omega^{(i)}$. Further, by Theorem 2.2,

$$0 = \lim_{i \to \infty} \left(\tilde{\Theta}^{(i)} - \Theta_0\right) \tilde{L}\left(\tilde{\Theta}^{(i)}\right) = \lim_{i \to \infty} \left(\tilde{\Theta}^{(i)} - \Theta_0\right) \begin{bmatrix} I_p \\ L_\infty. \end{bmatrix}$$

So, L_∞ is a stationary point in the sense that for some $\Theta_\infty \in \bigcap_{i=0}^{\infty} \Omega^{(i)}$, we have

$$A_\infty + B_\infty L_\infty = A_0 + B_0 L_\infty. \tag{3.26}$$

Since $\mathcal{H}$ holds, and at the end of every episode we are using OFU to pick $\tilde{\Theta}^{(i)}$, we have

$\mathcal{J}^\star(\Theta_\infty) \leq \mathcal{J}^\star(\Theta_0)$. Hence, by Lemma 3.1, (3.26) implies that L_∞ is an optimal linear feedback for the true system Θ_0. But, according to Proposition 2.1, $L(\Theta_0)$ is unique, i.e. $L_\infty = L(\Theta_0)$. Therefore, the limit is unique, which contradicts the divergence. Moreover, the convergence is to $L(\Theta_0)$, i.e.

$$\lim_{i\to\infty} D_i = \Theta_0 \tilde{L}(\Theta_0) = D_0. \tag{3.27}$$

Next, as shown in the proof of Lemma 2.2, $\sum\limits_{\ell=0}^{\infty} {D_i}^\ell C {D_i'}^\ell$ is a Lipschitz function of D_i. Thus, plugging (3.27) in (3.25) we get

$$\lim_{i\to\infty} \det\left(\frac{1}{n_i} V^{(i)}\right) = \det\left(\sum_{\ell=0}^{\infty} {D_0}^\ell C {D_0'}^\ell\right),$$

which yields

$$\lim_{i\to\infty} \det\left(\frac{1}{n_i} U^{(i)}\right) = \det\left(\tilde{L}(\Theta_0) \sum_{\ell=0}^{\infty} {D_0}^\ell C {D_0'}^\ell \tilde{L}(\Theta_0)'\right).$$

Therefore, defining

$$\tilde{s}_j = \frac{\det\left(\sum\limits_{i=0}^{j} U_i\right)}{\det\left(\sum\limits_{i=0}^{j-1} U_i\right)},$$

we have

$$\lim_{j\to\infty} \tilde{s}_j = \lim_{j\to\infty} \left(\frac{n_j}{n_{j-1}}\right)^q \frac{\det\left(\frac{1}{n_j}\sum\limits_{i=0}^{j} U_i\right)}{\det\left(\frac{1}{n_{j-1}}\sum\limits_{i=0}^{j-1} U_i\right)} = \lim_{j\to\infty} \left(\frac{n_j}{n_{j-1}}\right)^q.$$

Note that according to (2.37), on $\mathcal{G} \cap \mathcal{H}$ the matrix $\frac{1}{n_i} U_i$, and so the matrix $\frac{1}{n_j}\sum\limits_{i=0}^{j} U_i$, are bounded. Using the definition of episode size in (3.2), we get

$$\lim_{j\to\infty} \tilde{s}_j = \gamma. \tag{3.28}$$

Finally, according to Lemma 11 in the work of Abbasi-Yadkori and Szepesvári [15],

$$\sup_{\Theta \neq 0} s_j(\Theta) \leq \tilde{s}_j.$$

So, (3.28) implies the desired result. □

Lemma 3.6 (Bounding $\mathbb{T}_4$)**.** On the event $\mathcal{G} \cap \mathcal{H}$, it holds that

$$\mathbb{T}_4 \leq \Delta_6 m(T) \beta_T \left(\frac{\delta}{m(T)^2} \right)^{1/2} (\log m(T))^{1/\alpha} (-\log \delta)^{1/\alpha} T^{1/2},$$

for some constant $\Delta_6 < \infty$.

Proof. Assuming $\mathcal{G} \cap \mathcal{H}$ holds, consider the following expression:

$$\mathbb{T}_5 = \sum_{t=1}^{T} \left\| (\Theta_t - \Theta_0) \tilde{L}(\Theta_t) x(t) \right\|_2^2.$$

Since Θ_t does not change during every episode, we can write

$$\mathbb{T}_5 \leq \sum_{j=1}^{m(T)} \sum_{t=\lceil \tau_{j-1} \rceil}^{\lceil \tau_j \rceil - 1} \left\| \left(\tilde{\Theta}^{(j)} - \Theta_0 \right) \tilde{L}\left(\tilde{\Theta}^{(j)} \right) x(t) \right\|_2^2. \tag{3.29}$$

Letting $\{U_i\}_{i=0}^{\infty}$ be as defined in Lemma 3.5, $\sum_{i=0}^{j} U_i$ is invertible and

$$\begin{aligned}
&\sum_{t=\lceil \tau_{j-1} \rceil}^{\lceil \tau_j \rceil -1} \left\| \left(\sum_{i=0}^{j} U_i \right)^{-1/2} \tilde{L}\left(\tilde{\Theta}^{(i)}\right) x(t) \right\|_2^2 \\
= &\sum_{t=\lceil \tau_{j-1} \rceil}^{\lceil \tau_j \rceil -1} x(t)' \tilde{L}\left(\tilde{\Theta}^{(i)}\right)' \left(\sum_{i=0}^{j} U_i \right)^{-1} \tilde{L}\left(\tilde{\Theta}^{(i)}\right) x(t) \\
= &\sum_{t=\lceil \tau_{j-1} \rceil}^{\lceil \tau_j \rceil -1} \operatorname{tr}\left(\left(\sum_{i=0}^{j} U_i \right)^{-1} \tilde{L}\left(\tilde{\Theta}^{(i)}\right) x(t) x(t)' \tilde{L}\left(\tilde{\Theta}^{(i)}\right)' \right) \\
= &\operatorname{tr}\left(\left(\sum_{i=0}^{j} U_i \right)^{-1} U_j \right) \leq \operatorname{tr}\left(I_q\right) = q.
\end{aligned}$$

Further, using definition of $\{s_j\left(\Theta\right)\}_{j=1}^{\infty}$ in Lemma 3.5 we have

$$\begin{aligned}
&\left|\left|\left| \left(\tilde{\Theta}^{(j)} - \Theta_0\right) \left(\sum_{i=0}^{j} U_i \right)^{1/2} \right|\right|\right|_2^2 \\
= &\left|\left|\left| \left(\sum_{i=0}^{j} U_i \right)^{1/2} \left(\tilde{\Theta}^{(j)} - \Theta_0\right)' \left(\tilde{\Theta}^{(j)} - \Theta_0\right) \left(\sum_{i=0}^{j} U_i \right)^{1/2} \right|\right|\right|_2 \\
\leq &\operatorname{tr}\left(\left(\tilde{\Theta}^{(j)} - \Theta_0\right) \sum_{i=0}^{j} U_i \left(\tilde{\Theta}^{(j)} - \Theta_0\right)' \right) \\
\leq &p \left|\left|\left| \left(\tilde{\Theta}^{(j)} - \Theta_0\right) \sum_{i=0}^{j} U_i \left(\tilde{\Theta}^{(j)} - \Theta_0\right)' \right|\right|\right|_2 \\
\leq &p \left|\left|\left| \left(\tilde{\Theta}^{(j)} - \Theta_0\right) \sum_{i=0}^{j-1} U_i \left(\tilde{\Theta}^{(j)} - \Theta_0\right)' \right|\right|\right|_2 s_j\left(\tilde{\Theta}^{(j)} - \Theta_0\right),
\end{aligned}$$

but according to definition of $\Omega^{(j)}$ in (3.7), both $\tilde{\Theta}^{(j)}$, and Θ_0 belong to $\bigcap_{i=1}^{j-1} \Xi^{(i)}$, i.e. (3.6) implies

$$\begin{aligned}
& \left\|\left\|\left(\tilde{\Theta}^{(j)} - \Theta_0\right) \sum_{i=0}^{j-1} U_i \left(\tilde{\Theta}^{(j)} - \Theta_0\right)'\right\|\right\|_2 \\
\leq\ & \sum_{i=0}^{j-1} \left\|\left\|\left(\tilde{\Theta}^{(j)} - \Theta_0\right) U_i \left(\tilde{\Theta}^{(j)} - \Theta_0\right)'\right\|\right\|_2 \\
\leq\ & \left\|\left\|\left(\tilde{\Theta}^{(j)} - \Theta_0\right)'\right\|\right\|_2^2 + 4 \sum_{i=1}^{j-1} \beta_{n_i}\left(\frac{\delta}{i^2}\right) \\
\leq\ & 4\Delta_0^2 + 4m(T)\,\beta_T\left(\frac{\delta}{m(T)^2}\right),
\end{aligned}$$

where in the last inequality above (3.17), $n_i \leq T$, and $i \leq m(T)$ are used. Now, putting together we have

$$\begin{aligned}
& \sum_{t=\lceil \tau_{j-1} \rceil}^{\lceil \tau_j \rceil - 1} \left\|\left(\tilde{\Theta}^{(j)} - \Theta_0\right) \tilde{L}\left(\tilde{\Theta}^{(j)}\right) x(t)\right\|_2^2 \\
\leq\ & \left\|\left\|\left(\tilde{\Theta}^{(j)} - \Theta_0\right)\left(\sum_{i=0}^{j} U_i\right)^{1/2}\right\|\right\|_2^2 \sum_{t=\lceil \tau_{j-1} \rceil}^{\lceil \tau_j \rceil - 1} \left\|\left(\sum_{i=0}^{j} U_i\right)^{-1/2} \tilde{L}\left(\tilde{\Theta}^{(i)}\right) x(t)\right\|_2^2 \\
\leq\ & 4pqs_j\left(\tilde{\Theta}^{(j)} - \Theta_0\right)\left(\Delta_0^2 + m(T)\,\beta_T\left(\frac{\delta}{m(T)^2}\right)\right)
\end{aligned}$$

Plugging in (3.29), and using Lemma 3.5, leads to

$$\mathbb{T}_5 \leq 4pq\Delta_4 m(T)\left(\Delta_0^2 + m(T)\,\beta_T\left(\frac{\delta}{m(T)^2}\right)\right). \tag{3.30}$$

Going back to $\mathbb{T}_4$, write it as

$$\mathbb{T}_4 = \sum_{t=1}^{T} x(t)'\tilde{L}(\Theta_t)'(\Theta_0 + \Theta_t)' K(\Theta_t)(\Theta_0 - \Theta_t)\tilde{L}(\Theta_t)\, x(t).$$

Cauchy-Schwartz inequality leads to

$$\begin{aligned}\mathbb{T}_4 &\leq \sum_{t=1}^{T} \left\| K(\Theta_t)(\Theta_0+\Theta_t)\tilde{L}(\Theta_t)x(t) \right\|_2 \left\| (\Theta_0-\Theta_t)\tilde{L}(\Theta_t)x(t) \right\|_2 \\ &\leq \mathbb{T}_5^{1/2} \left[\sum_{t=1}^{T} \left\| K(\Theta_t)(\Theta_0+\Theta_t)\tilde{L}(\Theta_t)x(t) \right\|_2^2 \right]^{1/2}.\end{aligned}$$

By (2.23), for all stabilizable $\Theta \in \mathbb{R}^{p\times q}$, the equality

$$\tilde{L}(\Theta)' \Theta' K(\Theta) \Theta \tilde{L}(\Theta) = K(\Theta) - Q - L(\Theta)' RL(\Theta)$$

holds. So, since $Q + L(\Theta)' RL(\Theta)$ is PSD and R is PD, we respectively get

$$\left| \lambda_{\max}\left(\tilde{L}(\Theta)' \Theta' K(\Theta) \Theta \tilde{L}(\Theta) \right) \right| \leq |\lambda_{\max}(K(\Theta))| \leq \Delta_1, \tag{3.31}$$

$$\left\| \left| \tilde{L}(\Theta_t) \right| \right\|_2 \leq \Delta_5 < \infty. \tag{3.32}$$

Now, (3.31) in addition to (3.22) and (3.24) lead to

$$\begin{aligned}\left\| K(\Theta_t)\Theta_t\tilde{L}(\Theta_t)x(t) \right\|_2 &\leq \left\|\left| K(\Theta_t)^{1/2} \right|\right\|_2 \left\|\left| K(\Theta_t)^{1/2}\Theta_t\tilde{L}(\Theta_t) \right|\right\|_2 \|x(t)\|_2 \\ &\leq \Delta_1 \eta\left(D_1,\cdots,D_{m(T)}\right)\nu_T\left(\frac{\delta}{m(T)^2}\right),\end{aligned}$$

while (3.32) similarly implies

$$\left\| K(\Theta_t)\Theta_0\tilde{L}(\Theta_t)x(t) \right\|_2 \leq \Delta_1\Delta_5 |||\Theta_0|||_2 \eta\left(D_1,\cdots,D_{m(T)}\right)\nu_T\left(\frac{\delta}{m(T)^2}\right).$$

Putting all together and using (3.30), on $\mathcal{G}\cap\mathcal{H}$ we have

$$\begin{aligned}\mathbb{T}_4 &\leq \mathbb{T}_5^{1/2}T^{1/2}\Delta_1\left(1+\Delta_5|||\Theta_0|||_2\right)\eta\left(D_1,\cdots,D_{m(T)}\right)\nu_T\left(\frac{\delta}{m(T)^2}\right) \\ &\leq \Delta_6 m(T)\beta_T\left(\frac{\delta}{m(T)^2}\right)^{1/2}(\log m(T))^{1/\alpha}(-\log\delta)^{1/\alpha}T^{1/2},\end{aligned}$$

for some $\Delta_6 < \infty$, which is the desired result. $\square$

Lemma 3.7 (Bounding $m(T)$)**.** On the event $\mathcal{G} \cap \mathcal{H}$, it holds that

$$m(T) \leq \frac{q}{\log \gamma} \log \left(\frac{T\left(\gamma^{1/q} - 1\right)}{\tau_1} + 1 \right).$$

Proof. According to definition of episode size in (3.2), we have

$$\begin{aligned} \tau_i - \tau_{i-1} &= \gamma^{i/q} \left(N_{2.1} \left(\frac{|\lambda_{\min}(C)|}{2}, \frac{\delta}{i^2} \right) + 1 \right) \\ &\geq \gamma^{i/q} \left(N_{2.1} \left(\frac{|\lambda_{\min}(C)|}{2}, \delta \right) + 1 \right) \\ &= \gamma^{\frac{i-1}{q}} \tau_1. \end{aligned}$$

Since $\tau_{m(T)} \leq T$, we have

$$T \geq \sum_{i=1}^{m(T)} (\tau_i - \tau_{i-1}) \geq \frac{\gamma^{\frac{m(T)}{q}} - 1}{\gamma^{\frac{1}{q}} - 1} \tau_1,$$

which yields

$$m(T) \leq \frac{q}{\log \gamma} \log \left(\frac{T\left(\gamma^{1/q} - 1\right)}{\tau_1} + 1 \right).$$

$\square$

Finally, note that definition of $\beta_n(\delta)$ in Theorem 2.3 implies

$$\beta_n(\delta) = O\left((\log n)^{4/\alpha} (-\log \delta)^{1+4/\alpha} \right).$$

Therefore, plugging Lemmas 3.2, 3.3, 3.4, 3.6, and 3.7 into (3.20), we get

$$\mathcal{R}(T) = O\left(T^{1/2} (\log T)^{1+2/\alpha} (\log \log T)^{1/2+3/\alpha} (-\log \delta)^{1/2+3/\alpha} \right),$$

with probability at least $1 - \delta$ on $\mathcal{G} \cap \mathcal{H}$. Hence, according to (3.21), the failure probability is at most 6δ, which completes the proof of Theorem 3.1. □

Proof of Theorem 3.2. Let

$$\mathbb{T}_1, \mathbb{T}_2, \mathbb{T}_3, \mathbb{T}_4, m\left(T\right), \mathcal{G}, \mathcal{H}$$

be as defined in the proof of Theorem 3.1, and let $n_i = \tau_i - \tau_{i-1}$ be the length of episode i. Similarly, the regret can be written as

$$\mathcal{R}\left(T\right) = \sum_{t=1}^{T} \left[c_t - \mathcal{J}^{\star}\left(\Theta_0\right)\right] = \mathbb{T}_1 + \mathbb{T}_2 + \mathbb{T}_3 + \mathbb{T}_4. \tag{3.33}$$

Simply, Lemma 2.6 and Theorem 2.2 imply that

$$\mathbb{P}\left(\mathcal{G} \cap \mathcal{H}\right) \geq 1 - \sum_{i=1}^{\infty} \frac{3\delta}{i^2} \geq 1 - 5\delta. \tag{3.34}$$

Henceforth in the proof, we assume that $\mathcal{G} \cap \mathcal{H}$ holds. Bounding of $\mathbb{T}_1, \mathbb{T}_2, \mathbb{T}_3$ is exactly similar to Lemmas 3.2, 3.3, 3.4, respectively. We state and prove the following lemma, which will be used in order to upper bound $\mathbb{T}_4$.

Lemma 3.8. Let $\left\{M^{(i)}\right\}_{i=1}^{\infty}$ be a sequence of $p \times p$ matrices. Whenever $\tau_{i-1} \leq t < \tau_i$, let $M_t = M^{(i)}$. Define

$$\mathbb{T}_6 = \sum_{t=1}^{T} \|M_t x(t)\|_2^2.$$

On $\mathcal{G} \cap \mathcal{H}$, it holds that

$$\mathbb{T}_6 \leq \Delta_7 \sum_{i=1}^{m(T)} n_i \left|\left|\left|M^{(i)}\right|\right|\right|_2^2,$$

for some constant $\Delta_7 < \infty$.

Proof. Letting $D_i = \Theta_0 \tilde{L}\left(\tilde{\Theta}^{(i)}\right)$ be the stable closed-loop matrix during episode i, and

$$V^{(i)} = \sum_{t=\tau_{i-1}}^{\tau_i - 1} x(t)x(t)',$$

be the empirical covariance matrix of episode i, according to (2.37), on $\mathcal{G} \cap \mathcal{H}$ we have $\left|\lambda_{\max}\left(V^{(i)}\right)\right| \leq \frac{\Delta_7}{p^2} n_i$, where

$$\Delta_7 = \frac{3}{2} p^2 \left|\lambda_{\max}\left(C\right)\right| \eta \left(D_i'\right)^2 < \infty.$$

Therefore,

$$\begin{aligned}
\mathbb{T}_6 &= \sum_{t=1}^{T} x(t)' M_t' M_t x(t) \\
&= \sum_{i=1}^{m(T)} \operatorname{tr}\left(M^{(i)} V^{(i)} M^{(i)\prime}\right) \\
&\leq \sum_{i=1}^{m(T)} p \left|\lambda_{\max}\left(V^{(i)}\right)\right| \left\|\left| M^{(i)\prime} \right\|\right|_2^2 \\
&\leq \frac{\Delta_7}{p} \sum_{i=1}^{m(T)} n_i \left|\lambda_{\max}\left(M^{(i)} M^{(i)\prime}\right)\right| \\
&\leq \frac{\Delta_7}{p} \sum_{i=1}^{m(T)} n_i \operatorname{tr}\left(M^{(i)\prime} M^{(i)}\right) \\
&\leq \Delta_7 \sum_{i=1}^{m(T)} n_i \left\|\left| M^{(i)} \right\|\right|_2^2.
\end{aligned}$$

□

Lemma 3.9 (Bounding $\mathbb{T}_4$)**.** On the event $\mathcal{G} \cap \mathcal{H}$, it holds that

$$\mathbb{T}_4 \leq \Delta_{10} \left(\log T\right)^{2/\alpha} \left(\log m\left(T\right)\right)^{1/2+3/\alpha} \left(-\log \delta\right)^{1/2+3/\alpha} T^{1/2},$$

for some constant $\Delta_{10} < \infty$.

Proof. Defining $M_t = (\Theta_0 - \Theta_t) \tilde{L}(\Theta_t)$, Cauchy-Schwartz inequality leads to

$$\begin{aligned}
\mathbb{T}_4 &\leq \sum_{t=1}^{T} \left\| K(\Theta_t)(\Theta_0 + \Theta_t) \tilde{L}(\Theta_t) x(t) \right\|_2 \left\| (\Theta_0 - \Theta_t) \tilde{L}(\Theta_t) x(t) \right\|_2 \\
&\leq \mathbb{T}_6^{1/2} \left[\sum_{t=1}^{T} \left\| K(\Theta_t)(\Theta_0 + \Theta_t) \tilde{L}(\Theta_t) x(t) \right\|_2^2 \right]^{1/2},
\end{aligned}$$

where $\mathbb{T}_6$ is the same as Lemma 3.8. Then, since $\Theta_0, \tilde{\Theta}^{(i)} \in \Omega^{(i-1)}$, (3.13) implies

$$\left|\!\left|\!\left| \left(\tilde{\Theta}^{(i)} - \Theta_0 \right) \tilde{L}\left(\tilde{\Theta}^{(i-1)} \right) \right|\!\right|\!\right|_2 \leq 2\epsilon\gamma^{-i+1},$$

which by weak identifiability condition of (3.8) leads to

$$\left|\!\left|\!\left| \left(\tilde{\Theta}^{(i)} - \Theta_0 \right) \tilde{L}\left(\tilde{\Theta}^{(i)} \right) \right|\!\right|\!\right|_2 \leq 2\Gamma_{\Theta_0}\epsilon\gamma^{-i+1}.$$

Therefore, according to (3.31), (3.32),

$$\left|\!\left|\!\left| M^{(i)} \right|\!\right|\!\right|_2 \leq 2\left(1 + \Delta_5 |\!|\!| \Theta_0 |\!|\!|_2\right) \Delta_1 \Gamma_{\Theta_0} \epsilon \gamma^{-i+1}. \tag{3.35}$$

On the other hand, by (2.8), (2.9), for some constant $0 < \Delta_8 < \infty$, we have

$$\Delta_8 \gamma^{2i} \leq n_i \leq \Delta_8 (\log n_i)^{4/\alpha} \epsilon^{-2} \gamma^{2i} \left(-\log \frac{\delta}{i^2} \right)^{1+4/\alpha}. \tag{3.36}$$

Thus, by $n_i \leq T, i \leq m(T)$, putting (3.35), (3.36) together we get

$$\mathbb{T}_6 \leq \Delta_7 \sum_{i=1}^{m(T)} n_i \left|\!\left|\!\left| M^{(i)} \right|\!\right|\!\right|_2^2 \leq \Delta_9 \left(\log \frac{m(T)^2}{\delta} \right)^{1+4/\alpha} (\log T)^{4/\alpha}.$$

for some $\Delta_9 < \infty$. The reminder of proof is similar to Lemma 3.6, i.e.

$$\begin{aligned}
\mathbb{T}_4 &\leq \mathbb{T}_6^{1/2} T^{1/2} \Delta_1 \left(1 + \Delta_5 \|\Theta_0\|_2\right) \eta\left(D_1, \cdots, D_{m(T)}\right) \nu_T \left(\frac{\delta}{m\left(T\right)^2}\right) \\
&\leq \Delta_{10} \left(\log T\right)^{2/\alpha} \left(\log m\left(T\right)\right)^{1/2+3/\alpha} \left(-\log \delta\right)^{1/2+3/\alpha} T^{1/2},
\end{aligned}$$

which is the desired result. □

To bound $m\left(T\right)$, note that according to (3.36), we have $\Delta_8 \gamma^{2m(T)-2} \leq T$, i.e.

$$m\left(T\right) \leq \frac{\log T - \log \Delta_8}{2 \log \gamma} + 1. \tag{3.37}$$

Finally, since $\mathbb{T}_1, \mathbb{T}_2, \mathbb{T}_3$ were bounded before exactly similar to the proof of Theorem 3.1, plugging Lemma 3.9, and (3.37) into (3.33), we get the following regret bound which holds with probability at least $1 - 6\delta$;

$$\mathcal{R}\left(T\right) = O\left(T^{1/2} \left(\log T\right)^{2/\alpha} \left(\log \log T\right)^{1/2+3/\alpha} \left(-\log \delta\right)^{1/2+3/\alpha}\right).$$

□

Proof of Theorem 3.3. Letting

$$\mathbb{T}_1, \mathbb{T}_2, \mathbb{T}_3, \mathbb{T}_4, m\left(T\right), \mathcal{G}, \mathcal{H}, n_i$$

be as defined in the proof of Theorem 3.2, the regret can be written as

$$\mathcal{R}\left(T\right) = \sum_{t=1}^{T} \left[c_t - \mathcal{J}^\star\left(\Theta_0\right)\right] = \mathbb{T}_1 + \mathbb{T}_2 + \mathbb{T}_3 + \mathbb{T}_4. \tag{3.38}$$

Further, (3.34) holds similarly. Bounding of $\mathbb{T}_2, \mathbb{T}_3, \mathbb{T}_4$ is exactly similar to Theorem 3.2, with the following upper bound for Δ_1:

$$\begin{aligned}\forall i \geq 2 : \left|\lambda_{\min}\left(C\right)\right| \left|\lambda_{\max}\left(K\left(\hat{\Theta}^{(i)}\right)\right)\right| &\leq \mathcal{J}^{\star}\left(\hat{\Theta}^{(i)}\right) \\ &\leq \mathcal{J}^{\star}\left(\Theta_0\right) + \Gamma_{\mathcal{J}} \left|\left|\left|\hat{\Theta}^{(i)} - \Theta_0\right|\right|\right|_2 \\ &\leq \mathcal{J}^{\star}\left(\Theta_0\right) + \Gamma_{\mathcal{J}} \Gamma_{\Theta_0} \epsilon,\end{aligned}$$

i.e.

$$\Delta_1 \leq \max\left\{\frac{\mathcal{J}^{\star}\left(\Theta_0\right) + \Gamma_{\mathcal{J}}\Gamma_{\Theta_0}\epsilon}{\left|\lambda_{\min}\left(C\right)\right|}, \left|\left|\left|K\left(\hat{\Theta}^{(1)}\right)\right|\right|\right|_2\right\} < \infty.$$

But since Algorithm 4 is not using OFU, the term $\mathbb{T}_1$ is not necessarily non-positive.

Lemma 3.10 (Bounding $\mathbb{T}_1$)**.** On the event $\mathcal{G} \cap \mathcal{H}$, we have

$$\mathbb{T}_1 \leq \Delta_{13} \left(\log m\left(T\right)\right)^{1+4/\alpha} \left(\log T\right)^{4/\alpha} T^{1/2} \left(-\log \delta\right)^{1/2+2/\alpha},$$

for some $\Delta_{13} < \infty$.

Proof. First, the design of episode size in (3.12) implies that

$$\left|\left|\left|\left(\hat{\Theta}^{(i)} - \Theta_0\right) \tilde{L}\left(\hat{\Theta}^{(i-1)}\right)\right|\right|\right|_2 \leq \epsilon \gamma^{-i+1}.$$

Therefore, by strong identifiability condition (3.14),

$$\left|\left|\left|\hat{\Theta}^{(i)} - \Theta_0\right|\right|\right|_2 \leq \Gamma_{\Theta_0} \epsilon \gamma^{-i+1},$$

which according to Lemma 2.2 implies that

$$\left|\mathcal{J}^{\star}\left(\hat{\Theta}^{(i)}\right) - \mathcal{J}^{\star}\left(\Theta_0\right)\right| \leq \Gamma_{\mathcal{J}} \Gamma_{\Theta_0} \epsilon \gamma^{-i+1}.$$

So, using (3.36),

$$\begin{aligned}
\mathbb{T}_1 &\le \sum_{i=1}^{m(T)} n_i \left| \mathcal{J}^\star\left(\hat{\Theta}^{(i)}\right) - \mathcal{J}^\star\left(\Theta_0\right)\right| \\
&\le \sum_{i=1}^{m(T)} \Delta_8 \left(\log n_i\right)^{4/\alpha} \epsilon^{-2}\gamma^{2i}\left(-\log\frac{\delta}{i^2}\right)^{1+4/\alpha} \Gamma_{\mathcal{J}}\Gamma_{\Theta_0}\epsilon\gamma^{-i+1} \\
&\le \Delta_{11}\left(\log T\right)^{4/\alpha}\left(-\log\frac{\delta}{m\left(T\right)^2}\right)^{1+4/\alpha}\gamma^{m(T)},
\end{aligned}$$

for some constant $\Delta_{11} < \infty$. On the other hand,

$$T \ge n_{m(T)-1} \ge \Delta_{12}\left(-\log\delta\right)^{1+4/\alpha}\gamma^{2m(T)},$$

for some $\Delta_{12} > 0$. Thus, for some $\Delta_{13} < \infty$,

$$\mathbb{T}_1 \le \Delta_{13}\left(\log m\left(T\right)\right)^{1+4/\alpha}\left(\log T\right)^{4/\alpha} T^{1/2}\left(-\log\delta\right)^{1/2+2/\alpha},$$

which is the desired result □

This finishes the proof, implying that with probability at least $1 - 6\delta$;

$$\mathcal{R}\left(T\right) = O\left(T^{1/2}\left(\log T\right)^{4/\alpha}\left(\log\log T\right)^{1+4/\alpha}\left(-\log\delta\right)^{1/2+3/\alpha}\right).$$

□

CHAPTER 4

Estimation in General VAR models

4.1 Introduction

Estimation of the transition matrix in VAR models has been extensively studied in the statistics and econometrics literature for the stable case [20]. Further, new work has also addressed this topic under high-dimensional scaling and extra assumptions on sparsity imposed on it [21]. However, in settings where the underlying process is not stable, the topic has *not* been adequately examined. A key issue that arises in this case is that the magnitude of the vector process explodes exponentially over time, with high probability [22]. Nevertheless, in addition to adaptive control of the systems evolving according to linear dynamics, estimation of the transition matrix in the non-stable case is of interest due to a number of applications that give rise to such instances, such as asset bubbles and (Yugoslav) hyperinflation [23], [24].

Some of the work on the topic provides asymptotic results [22]. Early work investigated the asymptotic distribution of the VAR model under a set of restrictive assumptions on the transition matrix [25]. Ensuing work dealt with asymptotic consistency of the estimates for a class of structured transition matrices [26]. Further extensions to more general classes were established by Nielsen [27], [28]. Finally, additional asymptotic results together with the important concept of *irregularity* of the transition matrix (see Definition 2.3) which leads to inconsistency, are presented in the literature [29]. However, finite sample results are not currently available.

In this chapter, we consider a VAR process $x(t) \in \mathbb{R}^p, t = 0, 1, \cdots$ that evolves according to

$$x(t+1) = D_0 x(t) + w(t+1), \tag{4.1}$$

starting from an arbitrary initial state $x(0)$. We examine the general case where the process is not necessarily stable (as technically defined in the next section). The key contributions of this chapter are:
(i) establishing finite sample bounds for the ℓ_2 error of the least-squares estimates of the transition matrix D_0,
(ii) under a fairly general heavy tailed noise process $\{w(t)\}_{t=1}^{\infty}$. In addition, the results due to the presence of a heavy-tailed noise term are of independent interest for the stable case as well (see Corollary 4.1). The novel results provided identify how the sample size required scales both with the dimension of the VAR model, as well as with the characteristics of the transition matrix and the noise process.

In order to establish results for finite sample estimation of D_0, one needs to overcome a set of issues. First, as long as D_0 has eigenvalues outside of the unit circle in the complex plane, the behavior of the Gram matrix is governed by a random matrix. The second issue arises when D_0 has eigenvalues both inside and outside of the unit circle. In this case, the smallest eigenvalue of the Gram matrix scales linearly with respect to sample size, while its largest eigenvalue grows exponentially. This leads to the failure of the classical approaches to establish consistency. The above issues are addressed in Section 4.4 and Section 4.5, respectively. In the proofs, we leverage selected concentration inequalities for random matrices [30], as well as an anti-concentration property of martingale difference sequences (see Lemma 4.2) [31].

Recently, the problem of forecasting non-stationary mixing [32], [33], and non-mixing [34] time series has received attention, assuming the loss function is bounded. Unstable VAR models are a special, yet interesting, case of non-stationary time series. However, the problem of estimation is not still addressed in the existing literature. Moreover, the results on forecasting are not applicable to estimation, since the sum-of-squares loss function employed in this study is not bounded. However, note that our results on estimation imply those on forecasting.

4.2 Technical Details

The VAR process $\{x(t)\}_{t=0}^{\infty}$ evolves according to (4.1), while the unknown transition matrix $D_0 \in \mathbb{R}^{p \times p}$ is not assumed to be stable, i.e. the eigenvalues of D_0 do not necessarily lie inside the unit circle. Further, $\{w(t)\}_{t=1}^{\infty}$ is the sequence of independent mean-zero noise vectors with covariance matrix C, i.e.

$$\mathbb{E}\left[w(t)\right] = 0, \quad \mathbb{E}\left[w(t)w(t)'\right] = C.$$

Remark 4.1. The results established also hold if the noise vectors are martingale difference sequences. Further, the generalization for heteroscedastic noise, where the covariance matrix C is time varying, is rather straightforward.

The objective is to estimate D_0, using the row-wise least-squares estimator. Technically, observing samples $\{x(t)\}_{t=0}^{n}$, define the sum-of-squares loss function

$$\mathcal{L}_n^{(i)}\left(\theta\right) = \sum_{t=0}^{n-1} \left(x_i(t+1) - \theta' x(t)\right)^2.$$

Then, the true transition matrix D_0 is estimated by

$$\hat{D}_n = \left[\hat{d}_1, \cdots, \hat{d}_p\right]',$$

where for $i = 1, \cdots, p$, the vector $\hat{d}_i$ is a minimizer of the above sum-of-squares, i.e.

$$\mathcal{L}_n^{(i)}\left(\hat{d}_i\right) = \min_{\theta \in \mathbb{R}^p} \mathcal{L}_n^{(i)}(\theta).$$

The main contribution of this chapter is showing that with high probability, accurate estimation of the true transition matrix is achieved, excluding a pathological case. Formally, $\hat{D}_n$ is with probability at least $1-\delta$ within an ϵ-neighborhood of D_0, where apart from a logarithmic factor, the sample size n scales quadratically with $\frac{1}{\epsilon}$, and logarithmically with $\frac{1}{\delta}$ (see Theorem 4.3).

To analyze the finite sample behavior of the above estimation procedure, Assumption 2.1 is assumed for the tail-behavior of every coordinate of the noise vector. To proceed, we define a property of the population covariance matrix of a VAR process. It can be seen in the proofs of the presented results, that the following property is necessary and sufficient for accurate estimation of the VAR parameters. The motivation behind Definition 4.1 becomes clear by the example presented later on.

Definition 4.1 (reachability)**.** The pair $[D_0, C]$ is called reachable if

$$\operatorname{rank}\left(\left[C^{1/2}, D_0 C^{1/2}, \cdots, D_0^{p-1} C^{1/2}\right]\right) = p.$$

Clearly, reachability is equivalent to $|\lambda_{\min}(K(C))| > 0$, where

$$K(C) = \sum_{i=0}^{p-1} D_0^i C D_0'^{\,i}.$$

Specifically, if C is positive definite, then $[D_0, C]$ is reachable for all $D_0 \in \mathbb{R}^{p \times p}$, and

$$|\lambda_{\min}(K(C))| > |\lambda_{\min}(C)| > 0.$$

Note that reachability is conceptually equivalent to the population covariance matrix of the process being positive definite. More precisely, the evolution of the process over time implies

$$x(t) = D_0^t x(0) + \sum_{i=1}^{t} D_0^{t-i} w(i).$$

Therefore, since the noise vectors are independent, the covariance matrix of $x(t)$ is given by

$$\sum_{i=0}^{t-1} D_0^i C {D'_0}^i;$$

i.e. reachability is in fact stating that for $t \geq p$, every coordinate of $x(t)$ has a non-degenerate randomness. In the following example, we describe a situation to demonstrate the usefulness of reachability.

Back to Definition 4.1, a natural question arising concerns the motivation behind considering reachable pairs rather than simply assuming positive definiteness of C. There is an extensive family of settings, including the following example, where the latter stronger condition does not essentially hold.

Example 4.1 (Time series with longer memory)**.** Consider a VAR(k) model, where the evolution of the process to the next time step is determined by the k previous lags, for some $k > 1$. For $t \geq k$, the process $\tilde{x}(t) \in \mathbb{R}^m$ evolves according to

$$\tilde{x}(t) = \sum_{j=1}^{k} D_j \tilde{x}(t-j) + \tilde{w}(t),$$

for some initial vectors $\tilde{x}(0), \cdots, \tilde{x}(k-1) \in \mathbb{R}^m$, and transition matrices $D_1, \cdots, D_k \in \mathbb{R}^{m \times m}$, assuming $D_k \neq 0$.

Suppose that the covariance matrix of $\tilde{w}(t)$, $\tilde{C} \in \mathbb{R}^{m \times m}$, is positive definite. Arranging blocks of $\tilde{x}(t)$ accordinlgly, the process can be written in the form of a VAR(1) model, as follows. Letting

$$x(t) = \begin{bmatrix} \tilde{x}(t+k-1) \\ \vdots \\ \tilde{x}(t) \end{bmatrix} \in \mathbb{R}^{km}, w(t) = \begin{bmatrix} \tilde{w}(t+k-1) \\ 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^{km},$$

we get $x(t+1) = D_0 x(t) + w(t+1)$, where

$$D_0 = \begin{bmatrix} D_1 \cdots D_{k-1} & D_k \\ I_{(k-1)m} & 0 \end{bmatrix} \in \mathbb{R}^{km \times km}.$$

Obviously, the covariance matrix of $w(t)$, denoted by C, is not full rank, although, one can show that $[D_0, C]$ is reachable. Hence, for processes exhibiting longer range temporal dependence, reachability constitutes a very natural and critical assumption.

Next, we establish results regarding the sample size needed so that with high probability the least squares estimate of D_0 is accurate within a certain degree. First, we study the stable case where all eigenvalues of D_0 are inside the unit circle, i.e. $|\lambda_{\max}(D_0)| < 1$. Subsequently, the explosive case where all eigenvalues lie outside of the unit circle, i.e. $|\lambda_{\min}(D_0)| > 1$, is examined. Finally, finite sample estimation results are presented for the general case which is the combination of these two regimes.

Some straightforward algebra shows that the least-squares estimator of the transition matrix can be written as

$$\hat{D}_n = \sum_{t=0}^{n-1} x(t+1)x(t)' V_n^{-1},$$

where $V_n = \sum_{t=0}^{n-1} x(t)x(t)'$ denotes the empirical covariance matrix of the VAR process, which is assumed to be non-singular.

The latter result implies that the behavior of V_n needs to be carefully studied and this constitutes a major part of the following two subsections. The proofs of all the results established, as well as the required preliminaries, are provided in Section 4.6. Further, all constants being referred to in this chapter, can be explicitly recovered from the detailed restatements of the results in the corresponding proofs.

4.3 Stable Case

The stable case has been extensively studied before, customarily under the stronger assumption of sub-Gaussian noise [35]. Next, we generalize the results to sub-Weibull noise vectors previously defined in Assumption 2.1. Further, these results will be used for the general case in Section 4.5, where we address the general non-stationary case.

In the stable regime, the process has a stationary limit distribution. In this case, the empirical covariance matrix has an approximately deterministic behavior, which is described by population covariance matrix of the asymptotic distribution. Specifically, if the sample size is large enough, V_n, once normalized properly, can be approximated by $\kappa\left(C\right)$, where

$$\kappa\left(C\right) = \sum_{i=0}^{\infty} D_0^i C D_0'^{\,i}$$

is the asymptotic covariance matrix. The following lemma provides a finite sample bound determined by the estimation error ϵ, the failure probability δ and the tail exponent α. Henceforth, one can let $\alpha \to \infty$ in all inequalities presented for sample size, if the noise vectors $w(1), w(2), \cdots$ are bounded.

Theorem 4.1 (Stable covariance)**.** Assuming $\left|\lambda_{\max}\left(D_0\right)\right| < 1$, there is a constant $\Delta_1 < \infty$, such that for arbitrary $\epsilon, \delta > 0$ if

$$\frac{n}{\left(\log n\right)^{4/\alpha}} \geq \frac{\Delta_1}{\epsilon^2}\left(-\log\delta\right)^{1+4/\alpha},$$

Then

$$\mathbb{P}\left(\left|\lambda_{\max}\left(\frac{1}{n}V_{n+1} - \kappa\left(C\right)\right)\right| > \epsilon\right) \leq \delta.$$

A direct consequence of Theorem 4.1 is the following corollary, which shows that high probability accurate estimation can be ensured, if reachability, as defined in Definition 4.1, is assumed. Note that reachability implies that $\kappa(C)$ is positive definite.

Corollary 4.1 (Stable estimation)**.** Suppose that $|\lambda_{\max}(D_0)| < 1$, and $[D_0, C]$ is reachable. Then, there is $\Delta_2 < \infty$, such that

$$\frac{n}{(\log n)^{4/\alpha}} \geq \frac{\Delta_2}{\epsilon^2} (-\log \delta)^{1+4/\alpha},$$

implies

$$\mathbb{P}\left(\left|\left|\left| \hat{D}_n - D_0 \right|\right|\right|_2 > \epsilon \right) < \delta.$$

Remark 4.2. The scaling of the constants Δ_1, Δ_2 used above, is polynomial with respect to the dimension p. The degree of this polynomial depends on the size of the largest block in the Jordan form of D_0. For example, if D_0 is diagonalizable, Δ_1, Δ_2 scale linearly with respect to p.

The explicit dependence of Δ_1, Δ_2 on D_0 (in fact through $|\lambda_{\max}(D_0)|$), the noise covariance matrix C (namely, $|\lambda_{\max}(C)|$ for Δ_1 and $|\lambda_{\min}(K(C))|$ for Δ_2), as well as the parameters b_1, b_2, α specified in Assumption 2.1, can be found in the corresponding proofs.

4.4 Explosive Case

In the explosive case, the empirical covariance matrix V_n grows exponentially with respect to n. In addition, unlike the stable case, V_n, once normalized properly, can be approximated by a random matrix. Therefore, the eigenvalues of the normalized empirical covariance matrix are stochastic as well. In order to find deterministic bounds for the eigenvalues of V_n, new quantities, denoted by $\phi(D_0), \psi(D_0, \delta)$, need to be defined.

Subsequently, after providing the formal definition of the above quantities, we present in Theorem 4.2 bounds for the eigenvalues. Then, a sufficient and necessary property of D_0

for the accurate estimation will be introduced, followed by Lemmma 4.1, and Lemma 4.2, which establish the positiveness of $\phi\left(\cdot\right), \psi\left(\cdot,\cdot\right)$. This section concludes with Corollary 4.2 that deals with estimation in the explosive case.

First, for explosive D_0, we define the nonnegative functions $\phi\left(D_0\right), \psi\left(D_0, \delta\right)$ as follows. Assuming $\left|\lambda_{\min}\left(D_0\right)\right| > 1$, let $D_0 = P^{-1}\Lambda P$ be the Jordan decomposition (detailed definition, as well as some properties, can be found in Section 4.6). Letting

$$\begin{aligned} z(\infty) &= x(0) + \sum_{i=1}^{\infty} D_0^{-i} w(i), \\ P &= \left[P_1, \cdots, P_p\right]', \end{aligned}$$

for $\delta > 0$ define

$$\psi\left(D_0, \delta\right) = \sup\left\{ y \in \mathbb{R} : \mathbb{P}\left(\min_{1 \leq i \leq p} \left|P_i' z(\infty)\right| < y \right) \leq \delta \right\}.$$

Note that according to this definition, all coordinates of the vector $Pz(\infty)$ are in magnitude at least $\psi\left(D_0, \delta\right)$, with probability at least $1 - \delta$. Next, define

$$\phi\left(D_0\right) = \left|\!\left|\!\left| P \right|\!\right|\!\right|_{2\to\infty}^{-1} \inf_{a \in \mathbb{R}^p \setminus \{0\}} \frac{1}{\left\|a\right\|_1} \left[\sum_{i=0}^{p-1} a_{i+1} \Lambda^{-i} \right]_{\min},$$

where for an arbitrary matrix $M \in \mathbb{C}^{m \times k}$, $\left[M\right]_{\min}$ is the smallest magnitude of the nonzero entries of M:

$$\left[M\right]_{\min} = \min\{|M_{ij}| : 1 \leq i \leq m; 1 \leq j \leq k; M_{ij} \neq 0\}.$$

In fact, as the proofs show, $\phi\left(D_0\right)$ represents the deterministic portion of the smallest eigenvalue of the random matrix F_∞ which approximates the normalized V_n. It only depends on D_0, while $\psi\left(D_0, \delta\right)$ represents the stochastic portion which depends on both D_0 and the distribution of the noise sequence $\{w(t)\}_{t=1}^{\infty}$. Intuitively, $\phi\left(D_0\right)$ denotes the minimum nontrivial distance between the polynomials of D_0^{-1} and the origin, and $\psi\left(D_0, \delta\right)$ denotes the high probability minimum distance of the vector $Pz(\infty)$ from the origin. These minimum distances show up, because for $v \in \mathbb{R}^p$, $v' F_\infty v$ is determined by the product of a polynomial of D_0^{-1} (with coefficients determined by v), and $Pz(\infty)$. More details can be found in the proof of Theorem 4.2.

Now, the behavior of the normalized empirical covariance matrix can be approximated as follows:

Theorem 4.2 (Explosive covariance)**.** Suppose that $\left|\lambda_{\min}\left(D_0\right)\right| > 1$; then, there is a constant $\xi\left(D_0\right) < \infty$ such that

$$\mathbb{P}\left(\left|\lambda_{\max}\left(D_0^{-n} V_{n+1} D_0'^{-n}\right)\right| > \xi\left(D_0\right)\left(-\log \delta\right)^{2/\alpha}\right) \leq \delta.$$

Further, there is $\Delta_3 < \infty$, such that $n \geq \Delta_3 \log\left(\frac{-\log \delta}{\epsilon}\right)$ implies

$$\mathbb{P}\left(\left|\lambda_{\min}\left(D_0^{-n} V_{n+1} D_0'^{-n}\right)\right| < \phi\left(D_0\right)^2 \psi\left(D_0, \delta\right)^2 - \epsilon\right) \leq \delta. \tag{4.2}$$

Remark 4.3. The sample size $n \geq \Delta_3 \log\left(\frac{-\log \delta}{\epsilon}\right)$ in Theorem 4.2 is interesting in following two ways. First, the accuracy ϵ decays exponentially fast when n grows. Second, the failure probability δ decays double exponentially with respect to n.

This surprising strong behavior is intuitively caused by the exponential growth of $x(t)$. Roughly speaking, the growing signal (i.e. $x(t)$) to noise (i.e. $w(t)$) ratio leads to the super fast decay of ϵ and δ.

If $\phi\left(D_0\right) \psi\left(D_0, \delta\right) = 0$, obviously (4.2) holds, and Theorem 4.2 becomes mute. Thus, the main interest is in the case where $\phi\left(D_0\right) \psi\left(D_0, \delta\right) \neq 0$, which we will show that holds, under certain conditions, and is necessary to ensure accurate estimation. In fact, the first case is of no interest, since it can be shown that V_n will be singular, and thus accurate estimation of D_0 fails, even if the sample size becomes infinitely large [29]. For the second case, the transition matrix D_0 needs to be regular, according to Definition 2.3. Regularity (of course in addition to reachability), leads to accurate estimation, as will be established in Corollary 4.2.

Lemma 4.1. Assuming $|\lambda_{\min}(D_0)| > 1$, regularity of D_0 is equivalent to $\phi(D_0) > 0$.

The next lemma shows that positiveness of $\psi(D_0, \delta)$ is implied by reachability. Lemma 4.2 also reveals a linear scaling of $\psi(D_0, \delta)$ with respect to δ, when the noise is a continuous random vector.

Lemma 4.2. Assume $|\lambda_{\min}(D_0)| > 1$, and $[D_0, C]$ is reachable. We then have

$$\psi(D_0, \delta) > 0.$$

Moreover, if there is $i \geq p$, such that $w(i-p+1), \cdots, w(i)$ have bounded pdfs over certain subspaces of $\mathbb{R}^p$, then, $\psi(D_0, \delta) \geq \psi_0 \delta$, for some constant $\psi_0 > 0$. If the bounded pdfs mentioned above correspond to the normal distribution, then

$$\psi_0 \geq \left(\frac{\pi \left|\lambda_{\min}(K(C))\right|}{2\left|\lambda_{\max}\left({D_0}^i {D_0'}^i\right)\right|} \right)^{1/2} p^{-1} \left(\min_{1 \leq i \leq p} \|P_i\|_2 \right).$$

Now, we are ready to state the key result for the sample size required to achieve accurate estimation for an explosive transition matrix.

Corollary 4.2 (Explosive estimation)**.** Suppose that $|\lambda_{\min}(D_0)| > 1$, D_0 is regular, and $[D_0, C]$ is reachable. There exists $\Delta_4 < \infty$, such that

$$n \geq \Delta_4 \log \left(\frac{-\log \delta}{\epsilon \psi(D_0, \delta)} \right) \tag{4.3}$$

implies

$$\mathbb{P}\left(\left|\left|\left| \hat{D}_n - D_0 \right|\right|\right|_2 > \epsilon \right) < \delta.$$

Considering (4.3), while similar to Remark 4.3 the super fast decay of ϵ still holds, the behavior of δ is of the common exponential order of n now (assuming the linear scaling of $\psi(D_0, \delta)$ with respect to δ).

Remark 4.4. Another interesting property of the explosive case is that Δ_3, Δ_4 scale logarithmically with respect to the dimension p.

Further, Δ_3, Δ_4 depend on D_0 (indeed through $|\lambda_{\min}(D_0)|$), as well as b_1, b_2, α specified in Assumption 2.1. The constant Δ_4 also depends on $\phi(D_0)$. These dependencies, are spelled out explicitly in the corresponding proofs.

4.5 General Case

The preliminary results previously stated set the stage for the main result of this chapter. Theorem 4.3 establishes the accuracy of the estimation, when the regular matrix D_0 has no eigenvalue on the unit circle. As the following result shows, this assumption includes almost all matrices, with respect to Lebesgue measure on set of all square matrices.

Lemma 4.3. The set of all $p \times p$ real matrices with at least one eigenvalue on the unit circle has Lebesgue measure zero. Moreover, almost all matrices are regular.

Excluding two pathological cases of square matrices with at least one eigenvalue on the unit circle, and irregular matrices, the estimation of the transition matrix for a general VAR process is with high probability arbitrarily accurate, assuming that the sample size is large enough, as given by the following theorem. Before stating this result, we extend the domain of the non-negative function $\psi(D_0, \delta)$ to arbitrary matrices. Technically, when D_0 is not explosive, (i.e. has some nonexplosive eigenvalues), let D be a real matrix (of smaller size) formed by the explosive eigenvalues of D_0, with exactly the same algebraic and geometric multiplicities. We define $\psi(D_0, \delta) = \psi(D, \delta)$. The relationship between D_0, D is formally discussed in the proof of Theorem 4.3.

Theorem 4.3 (General estimation)**.** Suppose that D_0 is regular, has no eigenvalue on the unit circle, and $[D_0, C]$ is reachable. Then, there exists a constant $\Delta_5 < \infty$, such that for

$$\frac{n}{(\log n)^{4/\alpha}} \geq \frac{\Delta_5}{\epsilon^2} \left((-\log \delta)^{1+4/\alpha} - \log \psi\left(D_0, \delta\right) \right), \tag{4.4}$$

we have

$$\mathbb{P}\left(\left|\left|\left| \hat{D}_n - D_0 \right|\right|\right|_2 > \epsilon \right) < \delta.$$

Remark 4.5. Note that Lemma 4.2 implies that $-\log \psi\left(D_0, \delta\right) < \infty$, and apart from a constant, it is less than $-\log \delta$, if the noise vectors have bounded pdfs. In addition, the behavior of Δ_5 is fully determined by that of the constants Δ_2, Δ_4 (used in Corollary 4.1, and Corollary 4.2). For example, if D_0 is diagonalizable, Δ_5 scales linearly with the number of stable eigenvalues of D_0, and logarithmically with the number of explosive eigenvalues of D_0. Other dependencies are also similar to those of Δ_2, Δ_4, and are explicitly discussed in the proof of Theorem 4.3.

4.6 Technical Proofs

4.6.1 Proofs of Section 4.3

Proof of Theorem 4.1. Indeed, we prove the following. Let $N_{4.1}(\epsilon, \delta)$ be large enough, such that the followings hold for all $n \geq N_{4.1}(\epsilon, \delta)$.

$$\frac{n}{\nu_n(\delta)^2} \geq \frac{18|\lambda_{\max}(C)| + 2\epsilon}{\epsilon^2} p\eta(D_0')^4 \log\left(\frac{4p}{\delta}\right), \tag{4.5}$$

$$\frac{n}{\nu_n(\delta)^2 \pi_n(\delta)^2} \geq \frac{288p}{\epsilon^2}\eta(D_0')^4 |||D_0|||_2^2 \log\left(\frac{4p}{\delta}\right), \tag{4.6}$$

$$\frac{n}{\left(\|x(0)\|_\infty + \nu_n(\delta)\right)^2} \geq \frac{6}{\epsilon}\left(|||D_0|||_2^2 + 1\right)\eta(D_0')^2 \eta(D_0)^2. \tag{4.7}$$

We prove that on the event $\mathcal{W}$ defined in Lemma 2.6, for all $n \geq N_{4.1}(\epsilon, \delta)$ we have

$$\mathbb{P}\left(\left|\lambda_{\max}\left(\frac{1}{n}V_{n+1} - \kappa(C)\right)\right| > \epsilon\right) < \delta.$$

First, according to (2.33), letting

$$E_n = U_n + C_n + \frac{1}{n}D_0\left(x(0)x(0)' - x(n)x(n)'\right)D_0' + \frac{1}{n}x(0)x(0)',$$

since $|\lambda_{\max}(D_0)| < 1$, the Lyapunov equation $V_{n+1} = D_0 V_{n+1} D_0' + nE_n$ has the solution

$$\frac{1}{n}V_{n+1} = \sum_{i=0}^{\infty} D_0^i E_n {D_0'}^i = \kappa(E_n).$$

Henceforth in the proof, we assume the event $\mathcal{W}$ holds. According to Lemma 2.8, (4.5) implies that

$$\mathbb{P}\left(|\lambda_{\max}(C_n - C)| > \frac{\epsilon}{3\eta(D_0')^2}\right) \leq \frac{\delta}{2}. \tag{4.8}$$

In addition, by Lemma 2.10, (4.6) implies that

$$\mathbb{P}\left(|\lambda_{\max}(U_n)| > \frac{\epsilon}{3\eta(D_0')^2}\right) \leq \frac{\delta}{2}. \tag{4.9}$$

Finally, using Lemma 2.7, by (4.7) we get

$$\frac{1}{n}\left(|||D_0|||_2^2 + 1\right)\left(\|x(0)\|_2^2 + \|x(n)\|_2^2\right) \leq \frac{\epsilon}{3\eta(D_0')^2}. \tag{4.10}$$

Now, since

$$\sum_{t=0}^{\infty} \left|\left|\left|{D_0'}^t\right|\right|\right|_{\infty\to 2} \leq \eta(D_0'),$$

and $\left|\left|\left|{D_0'}^t\right|\right|\right|_2 \leq \left|\left|\left|{D_0'}^t\right|\right|\right|_{\infty\to 2}$, we have

$$\sum_{t=0}^{\infty} \left|\left|\left|{D_0'}^t\right|\right|\right|_2^2 \leq \left(\sum_{t=0}^{\infty}\left|\left|\left|{D_0'}^t\right|\right|\right|_2\right)^2 \leq \eta(D_0')^2. \tag{4.11}$$

Putting (4.8), (4.9), (4.10), and (4.11) together, on the event $\mathcal{W}$ we have

$$|\lambda_{\max}(\kappa(E_n - C))| \leq \sum_{t=0}^{\infty}\left|\lambda_{\max}\left(D_0^t(E_n - C){D_0'}^t\right)\right| \leq \eta(D_0')^2 |\lambda_{\max}(E_n - C)| \leq \epsilon,$$

with probability at least $1-\delta$, which is the desired result. □

Proof of Corollary 4.1. We prove that if the followings hold, then, on the event $\mathcal{W}$ we have $\left|\left|\left|\hat{D}_n - D_0\right|\right|\right|_2 \leq \epsilon$, with probability at least $1-\delta$. Letting $N_{4.1}(\cdot,\cdot)$ be as defined in the proof of Theorem 4.1, suppose that

$$n \geq N_{4.1}\left(\frac{|\lambda_{\min}(K(C))|}{2}, \frac{\delta}{2}\right) + 1, \tag{4.12}$$

$$\frac{n-2}{\nu_n(\delta)^2 \pi_n(\delta)^2} \geq \frac{32p}{|\lambda_{\min}(K(C))|^2 \epsilon^2}\log\left(\frac{4p}{\delta}\right). \tag{4.13}$$

First, by Theorem 4.1, (4.12) implies that on the event $\mathcal{W}$,

$$\frac{1}{n-1}\left|\lambda_{\min}\left(V_n\right)\right| \geq \left|\lambda_{\min}\left(\kappa\left(C\right)\right)\right| - \left|\lambda_{\max}\left(\frac{1}{n-1}V_n - \kappa\left(C\right)\right)\right| \geq \frac{\left|\lambda_{\min}\left(K(C)\right)\right|}{2}, \tag{4.14}$$

with probability at least $1-\delta/2$. Since $[D_0, C]$ is reachable, $\left|\lambda_{\min}\left(K(C)\right)\right| > 0$. Thus,

$$\hat{D}_n = \sum_{t=0}^{n-1} x(t+1)x(t)'V_n^{-1} = D_0 + U_n V_n^{-1},$$

where $U_n = \sum\limits_{t=0}^{n-1} w(t+1)x(t)'$, which leads to

$$\left|\left|\left|\hat{D}_n - D_0\right|\right|\right|_2 \leq \frac{|||U_n|||_2}{\left|\lambda_{\min}\left(V_n\right)\right|}. \tag{4.15}$$

To proceed, for arbitrary matrix $H \in \mathbb{R}^{k\times\ell}$, define the linear transformation

$$\Phi\left(H\right) = \begin{bmatrix} 0_{k\times k} & H \\ H' & 0_{\ell\times\ell} \end{bmatrix} \in \mathbb{R}^{(k+\ell)\times(k+\ell)}.$$

As a well known fact, the equality $|||H|||_2 = \left|\lambda_{\max}\left(\Phi\left(H\right)\right)\right|$ holds [30]. Note that $\Phi\left(H\right)$ is always symmetric. Next, letting $X_t = w(t+1)x(t)'$, apply Lemma 2.11 to $\Phi\left(X_t\right) \in \mathbb{R}^{2p\times 2p}$. Since

$$\Phi\left(X_t\right)^2 = \begin{bmatrix} \|x(t)\|_2^2 w(t+1)w(t+1)' & 0_{p\times p} \\ 0_{p\times p} & \|w(t+1)\|_2^2 x(t)x(t)' \end{bmatrix},$$

by Lemma 2.6, and Lemma 2.7, all matrices $\Phi\left(M_t\right)^2 - \Phi\left(X_t\right)^2$ are positive semidefinite on the event $\mathcal{W}$, where

$$M_t = p^{1/2}\nu_n\left(\delta\right)\pi_n\left(\delta\right)I_p.$$

By

$$\sigma^2 = \left|\lambda_{\max}\left(\sum_{t=0}^{n-1}\Phi\left(M_t\right)^2\right)\right| = np\nu_n\left(\delta\right)\pi_n\left(\delta\right)^2,$$

letting $y = \frac{|\lambda_{\min}(K(C))|}{2}\left(n-1\right)\epsilon$, according to Lemma 2.11, (4.13) implies

$$\mathbb{P}\left(|||U_n|||_2 > y\right) = \mathbb{P}\left(\left|\lambda_{\max}\left(\Phi\left(U_n\right)\right)\right| > y\right) \leq 2p\exp\left(-\frac{y^2}{8\sigma^2}\right) \leq \frac{\delta}{2},$$

which in addition to (4.14) gives the desired result, once plugged in (4.15). □

4.6.2 Proofs of Section 4.4

Lemma 4.4. Let $z(n) = x(0) + \sum_{t=1}^{n} D_0^{-t}w(t)$, where D_0 is an explosive matrix with Jordan decomposition $D_0 = P^{-1}\Lambda P$. Define the event

$$\mathcal{V} = \left\{\sup_{1\leq n\leq\infty}\|z(n)\|_2 \leq \xi\left(D_0,\delta\right)\right\},$$

where

$$\xi\left(D_0,\delta\right) = \|x(0)\|_2 + \left|\left|\left|P^{-1}\right|\right|\right|_{\infty\to 2}|||P|||_\infty\sum_{t=1}^{\infty}\eta_t\left(\Lambda^{-1}\right)\left(b_2\log\frac{2b_1pt^2}{\delta}\right)^{1/\alpha} < \infty.$$

We have $\mathbb{P}\left(\mathcal{V}\right) \geq 1-\delta$. Note that apart from a constant, $\xi\left(D_0,\delta\right)$ is less than or equal to $\left(-\log\delta\right)^{1/\alpha}$.

Proof of Lemma 4.4. First, according to Lemma 2.6,

$$\begin{aligned}
\mathbb{P}\left(\|w(t)\|_\infty \leq \nu_1\left(\frac{\delta}{2t^2}\right), \forall t=1,2,\cdots\right) &= 1-\mathbb{P}\left(\|w(t)\|_\infty > \nu_1\left(\frac{\delta}{2t^2}\right), \exists t=1,2,\cdots\right)\\
&\geq 1-\sum_{t=1}^{\infty}\mathbb{P}\left(\|w(t)\|_\infty > \nu_1\left(\frac{\delta}{2t^2}\right)\right)\\
&\geq 1-\sum_{t=1}^{\infty}\frac{\delta}{2t^2}\\
&> 1-\delta.
\end{aligned}$$

Then, similar to the proof of Lemma 2.7, we have

$$\left|\left|\left|\Lambda^{-t}\right|\right|\right|_{\infty} \leq \eta_t\left(\Lambda^{-1}\right).$$

So, on the above event, for all $n = 1, 2, \cdots$, we have

$$\begin{aligned}
\|z(n)\|_2 &\leq \sum_{t=1}^{\infty} \left|\left|\left|D_0^{-t}\right|\right|\right|_{\infty\to 2} \|w(t)\|_{\infty} \\
&\leq \|x(0)\|_2 + \left|\left|\left|P^{-1}\right|\right|\right|_{\infty\to 2} |||P|||_{\infty} \sum_{t=1}^{\infty} \eta_t\left(\Lambda^{-1}\right)\nu_1\left(\frac{\delta}{2t^2}\right) = \xi\left(D_0, \delta\right),
\end{aligned}$$

with probability at least $1-\delta$. □

Proof of Theorem 4.2. Letting $D_0 = P^{-1}\Lambda P$ be the Jordan decomposition of D_0, and $z(0) = x(0)$, for $n = 1, 2, \cdots, \infty$, define

$$\begin{aligned}
z(n) &= x(0) + \sum_{t=1}^{n} D_0^{-t} w(t), \\
U_n &= D_0^{-n} V_{n+1} D_0'^{-n}, \\
F_n &= \sum_{t=0}^{n} D_0^{-t} z(n) z(n)' D_0'^{-t}.
\end{aligned}$$

First, using $x(t) = D_0^t z(t)$, since

$$U_n = \sum_{t=0}^{n} D_0^{-n} x(t) x(t)' D_0'^{-n} = \sum_{t=0}^{n} D_0^{-n+t} z(t) z(t)' D_0'^{-n+t},$$

by Lemma 4.4 and (4.11), on the event $\mathcal{V}$ we have

$$\left|\lambda_{\max}\left(U_n\right)\right| \leq \sum_{t=0}^{\infty} \left\|D_0^{-t} z(n-t)\right\|_2^2 \leq \sum_{t=0}^{\infty} \left|\left|\left|D_0^{-t}\right|\right|\right|_2^2 \|z(n-t)\|_2^2 \leq \eta\left(D_0^{-1}\right)^2 \xi\left(D_0, \delta\right)^2,$$

which is the desired result, because the right hand side above is at most $\xi\left(D_0\right)\left(-\log\delta\right)^{2/\alpha}$, for some constant $\xi\left(D_0\right) < \infty$. In the sequel, we prove the desired result about the smallest eigenvalue. Letting

$$\rho_1 = 2\left(\left|\!\left|\!\left|P^{-1}\right|\!\right|\!\right|_{\infty\to 2}|\!|\!|P|\!|\!|_{\infty}\eta\left(D_0'^{-1}\right)^2 + \eta\left(D_0^{-1}\right)|\!|\!|P'|\!|\!|_{\infty\to 2}^2\left|\!\left|\!\left|P'^{-1}\right|\!\right|\!\right|_{\infty}^2\right)e^{2|\lambda_{\min}(D_0)|},$$
$$\rho_2 = 2\eta\left(D_0'^{-1}\right)^2\left(2+\eta\left(D_0^{-1}\right)\right)\left|\!\left|\!\left|P^{-1}\right|\!\right|\!\right|_{\infty\to 2}|\!|\!|P|\!|\!|_{\infty}e^{|\lambda_{\min}(D_0)|},$$

assume the followings hold for all $n \geq N_{4.2}(\epsilon, \delta)$:

$$\nu_n(\delta)\, n^{2\mu(D_0)}\left|\lambda_{\min}(D_0)\right|^{-2n/3} \leq \frac{\epsilon}{\rho_1\xi(D_0,\delta)}, \tag{4.16}$$
$$n^{\mu(D_0)-1}\left|\lambda_{\min}(D_0)\right|^{-n} \leq \frac{\epsilon}{\rho_2\xi(D_0,\delta)^2}, \tag{4.17}$$

where $\mu(D_0)$ is the largest size of blocks in the Jordan decomposition, as defined in the discussion after Definition 2.2. For all $n \geq N_{4.2}(\epsilon, \delta)$, we show that

$$\mathbb{P}\left(\left|\lambda_{\min}\left(D_0^{-n}V_{n+1}D_0'^{-n}\right)\right| < \phi(D_0)^2\,\psi(D_0,\delta)^2 - \epsilon\right) \leq 4\delta.$$

On the event $\mathcal{W}$, similar to the proof of Lemma 2.7, for all $t = 1, \cdots, n$ we have

$$\begin{aligned}\|z(n) - z(n-t)\|_2 &\leq \sum_{i=n-t+1}^{n}\left\|D_0^{-i}w(i)\right\|_2 \\ &\leq \left|\!\left|\!\left|P^{-1}\right|\!\right|\!\right|_{\infty\to 2}|\!|\!|P|\!|\!|_{\infty}\nu_n(\delta)\sum_{i=n-t+1}^{n}\eta_i\left(\Lambda^{-1}\right).\end{aligned}$$

Similarly, letting $\eta_0(\Lambda^{-1}) = 1$, note that $\eta_t(\Lambda'^{-1}) = \eta_t(\Lambda^{-1})$, for $t = 0, 1, 2, \cdots$. So,

$$\left|\!\left|\!\left|D_0'^{-t}\right|\!\right|\!\right|_2 \leq \left|\!\left|\!\left|D_0'^{-t}\right|\!\right|\!\right|_{\infty\to 2} \leq |\!|\!|P'|\!|\!|_{\infty\to 2}\left|\!\left|\!\left|P'^{-1}\right|\!\right|\!\right|_{\infty}\eta_t\left(\Lambda^{-1}\right). \tag{4.18}$$

According to the discussion after Definition 2.2, we have

$$\eta_t\left(\Lambda^{-1}\right) \leq t^{\mu(D)-1}\left|\lambda_{\min}(D)\right|^{-t}e^{|\lambda_{\min}(D)|}. \tag{4.19}$$

Thus,

$$
\begin{aligned}
& \sum_{t=0}^{n} \|z(n-t)-z(n)\|_2 \left|\left|\left| {D_0'}^{-t} \right|\right|\right|_2^2 \\
\leq\ & \sum_{t=0}^{n/3} \|z(n-t)-z(n)\|_2 \left|\left|\left| {D_0'}^{-t} \right|\right|\right|_2^2 + \sum_{t=n/3}^{n} \|z(n-t)-z(n)\|_2 \left|\left|\left| {D_0'}^{-t} \right|\right|\right|_2^2 \\
\leq\ & \left|\left|\left| P^{-1} \right|\right|\right|_{\infty\to 2} \||P|\|_{\infty} \nu_n(\delta) \left(\sum_{i=2n/3}^{n} \eta_i \left(\Lambda^{-1}\right) \right) \sum_{t=0}^{n/3} \left|\left|\left| {D_0'}^{-t} \right|\right|\right|_2^2 \\
+\ & \sum_{t=n/3}^{n} \|z(n-t)-z(n)\|_2 \left(\||P'|\|_{\infty\to 2} \left|\left|\left| {P'}^{-1} \right|\right|\right|_{\infty} \eta_t \left(\Lambda^{-1}\right) \right)^2 \\
\leq\ & \left|\left|\left| P^{-1} \right|\right|\right|_{\infty\to 2} \||P|\|_{\infty} \nu_n(\delta)\, n^{\mu(D_0)} \left|\lambda_{\min}(D_0)\right|^{-2n/3} e^{|\lambda_{\min}(D_0)|} \eta \left({D_0'}^{-1}\right)^2 \\
+\ & \eta\left(D_0^{-1}\right) \nu_n(\delta) \||P'|\|_{\infty\to 2}^2 \left|\left|\left| {P'}^{-1} \right|\right|\right|_{\infty}^2 n^{2\mu(D_0)} \left|\lambda_{\min}(D_0)\right|^{-2n/3} e^{2|\lambda_{\min}(D_0)|} \\
\leq\ & \frac{1}{2} \rho_1 \nu_n(\delta)\, n^{2\mu(D_0)} \left|\lambda_{\min}(D_0)\right|^{-2n/3},
\end{aligned}
$$

which by (4.16) implies

$$
\sum_{t=0}^{n} \|z(n-t)-z(n)\|_2 \left|\left|\left| {D_0'}^{-t} \right|\right|\right|_2^2 \leq \frac{\epsilon}{2\xi\left(D_0,\delta\right)}. \tag{4.20}
$$

By $x(t) = D_0^t z(t)$, since

$$
\begin{aligned}
U_n - F_n &= \sum_{t=0}^{n} D_0^{-n} x(t) x(t)' D_0'^{-n} - D_0^{-t} z(n) z(n)' D_0'^{-t} \\
&= \sum_{t=0}^{n} D_0^{-n+t} z(t) z(t)' D_0'^{-n+t} - D_0^{-t} z(n) z(n)' D_0'^{-t} \\
&= \sum_{t=0}^{n} D_0^{-t} z(n-t) z(n-t)' D_0'^{-t} - D_0^{-t} z(n) z(n)' D_0'^{-t} \\
&= \sum_{t=0}^{n} D_0^{-t} \left(z(n-t) z(n-t)' - z(n) z(n)' \right) D_0'^{-t}, \\
\left| \lambda_{\max} \left(U_n - F_n \right) \right| &\leq \sum_{t=0}^{n} \| z(n-t) - z(n) \|_2 \| z(n-t) + z(n) \|_2 \left|\left|\left| D_0'^{-t} \right|\right|\right|_2^2 \\
&\leq 2 \left(\sup_{1 \leq n \leq \infty} \| z(n) \|_2 \right) \sum_{t=0}^{n} \| z(n-t) - z(n) \|_2 \left|\left|\left| D_0'^{-t} \right|\right|\right|_2^2,
\end{aligned}
$$

using (4.20), and Lemma 4.4, on the event $\mathcal{W} \bigcap \mathcal{V}$, we get

$$
\left| \lambda_{\max} \left(U_n - F_n \right) \right| \leq \frac{\epsilon}{2}. \tag{4.21}
$$

On the other hand, one can use the same argument used in the proof of Lemma 4.4, to show that the following holds with probability at least $1 - \delta$.

$$
\| z(\infty) - z(n) \|_2 \leq \left|\left|\left| D_0^{-n} \right|\right|\right|_2 \left\| \sum_{t=1}^{\infty} D_0^{-t} w(n+t) \right\|_2 \leq \left|\left|\left| D_0^{-n} \right|\right|\right|_2 \xi \left(D_0, \delta \right).
$$

Therefore, using (4.18), on the event $\mathcal{V}$, with probability at least $1-\delta$ we have

$$\begin{aligned}
|\lambda_{\max}(F_\infty - F_n)| &\leq \sum_{t=0}^{n} \|z(\infty)-z(n)\|_2 \|z(\infty)+z(n)\|_2 \left\|\left| D_0'^{-t} \right|\right\|_2^2 \\
&+ \sum_{t=n+1}^{\infty} \left\| D_0^{-t} z(\infty) \right\|_2^2 \\
&\leq 2\xi(D_0,\delta)\|z(\infty)-z(n)\|_2 \sum_{t=0}^{n} \left\|\left| D_0'^{-t} \right|\right\|_2^2 \\
&+ \xi(D_0,\delta)^2 \left\|\left| D_0^{-n} \right|\right\|_2^2 \sum_{t=1}^{\infty} \left\|\left| D_0'^{-t} \right|\right\|_2^2 \\
&\leq \eta\left(D_0'^{-1}\right)^2 \xi(D_0,\delta)^2 \left(2 + \left\|\left| D_0^{-n} \right|\right\|_2\right) \left\|\left| D_0^{-n} \right|\right\|_2 \\
&\leq \eta\left(D_0'^{-1}\right)^2 \xi(D_0,\delta)^2 \left(2 + \eta\left(D_0^{-1}\right)\right) \left\|\left| P^{-1} \right|\right\|_{\infty\to 2} \| | P | \|_\infty \eta_n\left(\Lambda^{-1}\right).
\end{aligned}$$

By (4.19), (4.17) implies that on $\mathcal{V}$, with probability at least $1-\delta$,

$$|\lambda_{\max}(F_\infty - F_n)| \leq \frac{\epsilon}{2}. \tag{4.22}$$

Next, we show that with probability at least $1-\delta$,

$$|\lambda_{\min}(F_\infty)| \geq \left(\phi(D_0)\,\psi(P,\delta)\right)^2 = \lambda_0. \tag{4.23}$$

For this purpose, we need the following lemmas.

Lemma 4.5. Letting $f(x) = \sum_{i=0}^{p-1} a_{i+1} x^i$ be a real polynomial, we have

$$\mathbb{P}\left(\left\| f\left(D_0^{-1}\right) z(\infty) \right\|_2 < \|a\|_1 \phi(D_0)\,\psi(D_0,\delta)\right) \leq \delta.$$

Lemma 4.6. If $\phi(D_0)\,\psi(D_0,\delta) \neq 0$, then,

$$\mathbb{P}\left(\operatorname{rank}\left(\left[z(\infty), D_0 z(\infty), \cdots, D_0^{-p+1} z(\infty)\right]\right) < p\right) = 0.$$

If $\lambda_0 = 0$, (4.23) is trivial, because F_∞ is always a PSD matrix. Otherwise, assume $|\lambda_{\min}(F_\infty)| < \lambda_0$, and let $v \in \mathbb{R}^p$ be such that $\|v\|_2 = 1$, and $v' F_\infty v < \lambda_0$. Then,

$$\begin{aligned}
\lambda_0 &> \sum_{t=0}^{p-1} v' D_0^{-t} z(\infty) z(\infty)' D_0'^{-t} v \\
&= \sum_{t=0}^{p-1} \left(v' D_0^{-t} z(\infty)\right)^2 \\
&= \left\| v' \left[z(\infty), \cdots, D_0^{-p+1} z(\infty)\right] \right\|_2^2 \\
&\geq \left\| v' \left[z(\infty), \cdots, D_0^{-p+1} z(\infty)\right] \right\|_\infty^2 \\
&= \max_{0 \leq i \leq p-1} \left| v' D_0^{-i} z(\infty) \right|^2 .
\end{aligned}$$

By Lemma 4.6, almost surely, there is $a \in \mathbb{R}^p$, such that $v = \sum_{i=0}^{p-1} a_{i+1} D_0^{-i} z(\infty)$. So,

$$\begin{aligned}
\|v\|_2 &= \left| v' \left(\sum_{i=0}^{p-1} a_{i+1} D_0^{i} \right) z(\infty) \right| \\
&\leq \sum_{i=0}^{p-1} |a_{i+1}| \left| v' D_0^{i} z(\infty) \right| \\
&< \sum_{i=0}^{p-1} |a_{i+1}| \, {\lambda_0}^{1/2} \\
&= {\lambda_0}^{1/2} \|a\|_1,
\end{aligned}$$

which, by Lemma 4.5, holds with probability at most δ, i.e. (4.23) holds. Putting (4.21), (4.22), and (4.23) together, on the event $\mathcal{W} \bigcap \mathcal{V}$, we get the following, which holds with probability at least $1 - 2\delta$.

$$\begin{aligned}
|\lambda_{\min}(U_n)| &\geq |\lambda_{\min}(F_\infty)| - |\lambda_{\max}(F_\infty - F_n)| - |\lambda_{\max}(U_n - F_n)| \\
&\geq \lambda_0 - \epsilon,
\end{aligned}$$

which is the desired result. □

Proof of Lemma 4.5. If $\phi(D_0) = 0$, obviously the statement holds. So, assume $\phi(D_0) > 0$. Letting $D_0 = P^{-1}\Lambda P$ be the Jordan decomposition, we have

$$f\left(D_0^{-1}\right) = P^{-1} f\left(\Lambda^{-1}\right) P.$$

The matrix Λ is block diagonal, thus, $f(\Lambda^{-1})$ is block diagonal as well. Further, every block of Λ^{-i}, as well as every block of $f(\Lambda^{-1})$, is upper triangular (see proof of Lemma 2.7). Therefore, since $\phi(D_0) > 0$, the matrix $f(\Lambda^{-1})$ needs to have at least one nonzero entry. So, there is at least one row of the block-wise upper triangular matrix $f(\Lambda^{-1})$, which has exactly one nonzero entry.

This nonzero entry, by definition of $\phi(D_0)$, is in magnitude at least

$$\|a\|_1 \|\!|P|\!\|_{2\to\infty} \phi(D_0).$$

On the other hand, by definition of $\psi(D_0, \delta)$, all coordinates of the vector $Pz(\infty)$ are in magnitude at least $\psi(D_0, \delta)$, with probability at least $1 - \delta$.

So, with probability at least $1 - \delta$, the vector $u = f(\Lambda^{-1}) Pz(\infty)$ has a coordinate, which is in magnitude at least $\|a\|_1 \|\!|P|\!\|_{2\to\infty} \phi(D_0) \psi(P, \delta)$. This implies the desired inequality, because

$$\begin{aligned}
\|a\|_1 \|\!|P|\!\|_{2\to\infty} \phi(D_0) \psi(P, \delta) &\leq \left\| f\left(\Lambda^{-1}\right) Pz(\infty) \right\|_\infty \\
&= \left\| Pf\left(D_0^{-1}\right) z(\infty) \right\|_\infty \\
&\leq \|\!|P|\!\|_{2\to\infty} \left\| f\left(D_0^{-1}\right) z(\infty) \right\|_2.
\end{aligned}$$

□

Proof of Lemma 4.6. Let $D_0 = P^{-1}\Lambda P$ be the Jordan decomposition of D_0. Whenever

$$\operatorname{rank}\left(\left[z(\infty), \cdots, D_0^{-p+1} z(\infty)\right]\right) < p,$$

there is a nontrivial real polynomial f of degree at most $p-1$, such that

$$f\left(D_0^{-1}\right)z(\infty) = P^{-1}f\left(\Lambda^{-1}\right)Pz(\infty) = 0.$$

Since $\phi\left(D_0\right) > 0$, similar to the proof of Lemma 4.5, there is at least one row of $f\left(\Lambda^{-1}\right)$, say the i-th row, which has exactly one nonzero entry, say the ij-th entry.

Therefore, since the i-th coordinate of the vector $f\left(\Lambda^{-1}\right)Pz(\infty) = 0$, is zero, j-th coordinate of $Pz(\infty) = 0$ must be zero, i.e. $P_j'z(\infty) = 0$, where $P = [P_1, \cdots, P_p]'$. So, the desired result holds because

$$\mathbb{P}\left(\text{rank}\left(\left[z(\infty), \cdots, D^{-p+1}z(\infty)\right]\right) < p\right) = \mathbb{P}\left(\exists j : P_j'z(\infty) = 0\right) = 0.$$

To verify the last equality above, note that as we will see in the proof of Lemma 4.2, for all $j = 1, \cdots, p$, $\left|P_j'z(\infty)\right|$ has a continuous distribution, which yields

$$\mathbb{P}\left(\left|P_j'z(\infty)\right| = 0\right) = 0.$$

□

Proof of Lemma 4.1. Assume D_0 is regular. Clearly, the infimum in the definition of $\phi\left(D_0\right)$, can be taken over $\|a\|_1 = 1$. We will show that there is no polynomial f of degree at most $p-1$, such that $f\left(D_0^{-1}\right) = 0$. Note that this finishes the proof as follows. Let

$$S_1^p = \{a \in \mathbb{R}^p : \|a\|_1 = 1\}.$$

The function $\mathcal{G} : \mathbb{R}^p \to \mathbb{R}$, defined as

$$\mathcal{G}(a) = \left[\sum_{i=0}^{p-1} a_{i+1}\Lambda^{-i}\right]_{\min},$$

is continuous. Since S_1^p is a closed subset of $\mathbb{R}^p$, $\mathcal{G}\left(S_1^p\right) \subset \mathbb{R}$ is closed as well. Therefore, if for all $a \in S_1^p$, we have $\mathcal{G}(a) > 0$, then

$$\inf \mathcal{G}\left(S_1^p\right) > 0,$$

which means $\phi\left(D_0\right) > 0$.

If there is a polynomial f, such that $f\left(D_0^{-1}\right) = 0$, let $D_0^{-1} = P^{-1}\Gamma P$ be the Jordan decomposition of D_0^{-1}, where

$$\Gamma = \operatorname{diag}\left(\Gamma_1, \cdots, \Gamma_k\right),$$

and Γ_i is a size m_i Jordan matrix of γ_i, as defined before Definition 2.2. Now, $f\left(D_0^{-1}\right) = 0$ implies $f\left(\Gamma\right) = 0$, which in turn yields $f\left(\Gamma_i\right) = 0$, for all $i = 1, \cdots, k$. As shown in the proof of Lemma 2.7, diagonal coordinates of $f\left(\Gamma_i\right)$ are all $f\left(\gamma_i\right)$, i.e. $f\left(\gamma_i\right) = 0$.

Let $f(x) = g(x)\left(x - \gamma_1\right)^{n_1} \cdots \left(x - \gamma_k\right)^{n_k}$, where none of $\gamma_1, \cdots, \gamma_k$ is a root of $g(x)$. We show that for all i, $n_i \geq m_i$, so,

$$\deg f \geq \sum_{i=1}^{k} n_i \geq \sum_{i=1}^{k} m_i = p,$$

which is a contradiction. Note that by regularity of D_0, $\gamma_1, \cdots, \gamma_k$ are distinct, i.e. for $i \neq j$, $\Gamma_i - \gamma_j I_{m_i}$ is invertible (since it is a Jordan matrix of $\gamma_i - \gamma_j \neq 0$). Hence, $f\left(\Gamma_i\right) = 0$ implies $\left(\Gamma_i - \gamma_i I_{m_i}\right)^{n_i} = 0$. But, as shown in the proof of Lemma **??**, an exponent of size m Jordan matrix of 0 is zero matrix, only if the exponent is not smaller than m, i.e. $n_i \geq m_i$, which is the desired result.

Conversely, assume D_0 is not regular, i.e. there are $1 \leq i, j \leq k$, such that $\gamma_i = \gamma_j$, and $m_i \geq m_j \geq 1$. Letting $g(x) = \det\left(D_0 - xI_p\right)$, define

$$f(x) = \frac{1}{x - \gamma_i} g(x),$$

if γ_i is real, and

$$f(x) = \frac{1}{(x-\gamma_i)(x-\bar{\gamma}_i)} g(x),$$

otherwise, where $\bar{\gamma}_i$ is the complex conjugate of γ_i.

Clearly, $\deg f \leq p-1$, but we will show that $f\left(D_0^{-1}\right) = 0$, which leads to $\phi\left(D_0\right) = 0$. Note that since D_0 is regular, the polynomial $f(x)$ can not be a trivial one. As seen in the first part of the proof, it suffices to show that $f\left(\Gamma_\ell\right) = 0$, for all $\ell = 1, \cdots, k$. If $\ell \neq i, j$, we have $g\left(\Gamma_\ell\right) = 0$, so, $f\left(\Gamma_\ell\right) = 0$. Since the multiplicity of the root γ_i in $g(x)$ is $m_i + m_j$, its multiplicity in $f(x)$ is at least $m_i + m_j - 1 \geq m_i$, which is greater than or equal to the dimension of Γ_i and Γ_j. Therefore, $f\left(\Gamma_\ell\right) = 0$, for $\ell = i, j$, which finishes the proof. □

Proof of Lemma 4.2. We use the following Lemma [31].

Lemma 4.7. Let $\{\zeta_n\}_{n=1}^{\infty}$ be a martingale difference sequence of random variables with respect to the filter $\{\mathcal{F}_n\}_{n=1}^{\infty}$, such that

$$\liminf_{n\to\infty} \mathbb{E}\left[\zeta_n^2 | \mathcal{F}_{n-1}\right] > 0.$$

If the real sequence $\{a_n\}_{n=1}^{\infty}$, satisfies

$$\sum_{n=1}^{\infty} a_n^2 \leq \infty$$
$$a_n \neq 0, \text{ infinitely often,}$$

then $\sum_{n=1}^{\infty} a_n \zeta_n$ has a continuous distribution.

For an arbitrary row P_i' of P, let v be one of the real vectors $\Re\left(P_i\right)$ or $\Im\left(P_i\right)$. Note that since P is invertible, $P_i \neq 0$, and we can assume $v \neq 0$. Taking

$$\begin{aligned} a_n &= \left\| D_0'^{\,-np} v \right\|_2, \\ \zeta_n &= \frac{1}{a_n} \sum_{i=np-p+1}^{np} v' D_0^{-i} w(i), \end{aligned}$$

we have $a_n \neq 0$ infinitely often, and by $|\lambda_{\min}(D_0)| > 1$ we have $\sum_{n=1}^{\infty} a_n^2 < \infty$. Furthermore, by reachability we have

$$\begin{aligned}
\mathbb{E}\left[\zeta_n^2\right] &= \frac{1}{a_n^2} \sum_{i=np-p+1}^{np} v' D_0^{-i} C D_0'^{-i} v \\
&= \frac{1}{\left\|D_0'^{-np} v\right\|_2^2} v' D_0^{-np} \left(\sum_{i=0}^{p-1} D_0^i C D_0'^{i} \right) D_0'^{-np} v \\
&\geq |\lambda_{\min}(K(C))| > 0.
\end{aligned}$$

So, $v'z(\infty) = v'x(0) + \sum_{n=1}^{\infty} a_n \zeta_n$ has a continuous distribution. Letting $\mathbb{F}_i$ be the Cumulative Distribution Function (CDF) of $|P_i' z(\infty)|$, $\mathbb{F}_i$ is continuous, and because of $|P_i' z(\infty)| \geq |v' z(\infty)|$, one has $\mathbb{F}_i^{-1}\left(\frac{\delta}{p}\right) > 0$. Since,

$$\mathbb{P}\left(|P_i' z(\infty)| < \mathbb{F}_i^{-1}\left(\frac{\delta}{p}\right) \right) = \frac{\delta}{p}, \tag{4.24}$$

we have

$$\mathbb{P}\left(\min_{1\leq i\leq p} |P_i' z(\infty)| < \min_{1\leq i\leq p} \mathbb{F}_i^{-1}\left(\frac{\delta}{p}\right) \right) \leq \sum_{i=1}^{p} \mathbb{P}\left(|P_i' z(\infty)| < \mathbb{F}_i^{-1}\left(\frac{\delta}{p}\right) \right) = \delta,$$

i.e.

$$\psi(D_0, \delta) \geq \min_{1\leq i\leq p} \mathbb{F}_i^{-1}\left(\frac{\delta}{p}\right) > 0,$$

which is the desired result.

To proceed, we use the following fact. For two independent random variables X, Y, if X has bounded pdf f_X, then $X+Y$ has bounded pdf f_{X+Y}, and

$$\sup_{y\in\mathbb{R}} f_{X+Y}(y) \leq \sup_{y\in\mathbb{R}} f_X(y).$$

To see that, note that for all $y \in \mathbb{R}$,

$$f_{X+Y}(y) = \int\limits_{\mathbb{R}} f_X(y-\tau) d\mathbb{P}_Y(\tau) \le \left(\sup_{\tau \in \mathbb{R}} f_X(\tau) \right) \int\limits_{\mathbb{R}} d\mathbb{P}_Y(\tau) = \sup_{\tau \in \mathbb{R}} f_X(\tau).$$

Now, suppose that the supports of $w(i-p+1), \cdots, w(i)$ are certain subspaces of $\mathbb{R}^p$, and they have bounded pdfs. Then, all of the random variables $v' D_0^{-i+p-1} w(i-p+1), \cdots, v' D_0^{-i} w(i)$ cannot be degenerate. Since otherwise,

$$\operatorname{Var}\left(v' D_0^{-i+j} w(i-j) \right) = 0,$$

for all $j = 0, \cdots, p-1$, i.e.

$$\begin{aligned} 0 &= \operatorname{Var}\left(v' \sum_{j=0}^{p-1} D_0^{-i+j} w(i-j) \right) \\ &= v' D_0^{-i} K(C) D_0'^{-i} v \\ &\ge \left| \lambda_{\min}\left(K(C) \right) \right| \left\| D_0'^{-i} v \right\|_2^2 > 0, \end{aligned}$$

which is a contradiction. Therefore, there exists j, such that $D_0^{-i+j} w(i-j)$ lives in a subspace not orthogonal to v, i.e. $v' D_0^{-i+j} w(i-j)$ is a continuous random variable, with a bounded pdf (since pdf of $w(i-j)$ is bounded).

Using the fact mentioned above, pdf of $v' z(\infty)$, as well as pdf of $|P_i z(\infty)|$, denoted by f_i are bounded. Letting $\psi_0^{-1} = p \max_{1 \le i \le p} \sup_{y \in \mathbb{R}} f_i(y) < \infty$,

$$\mathbb{F}_i\left(\psi_0 \delta \right) = \int\limits_0^{\psi_0 \delta} f_i(y) dy \le \psi_0 \delta \sup_{y \in \mathbb{R}} f_i(y) \le \frac{\delta}{p},$$

i.e. $\mathbb{F}_i^{-1}\left(\frac{\delta}{p} \right) \ge \psi_0 \delta$, which is the desired result.

For normal case, $v' \sum_{j=0}^{p-1} D^{-i+j} w(i-j)$ is normal with pdf $\tilde{f}$, and

$$\begin{aligned} \operatorname{Var}\left(v' \sum_{j=0}^{p-1} D_0^{-i+j} w(i-j)\right) &= v' D_0^{-i} K(C) D_0'^{-i} v \\ &\geq |\lambda_{\min}(K(C))| \left\| D_0'^{-i} v \right\|_2^2 > 0, \end{aligned}$$

i.e.

$$\begin{aligned} \sup_{y \in \mathbb{R}} \tilde{f}(y) &\leq \left(2\pi |\lambda_{\min}(K(C))|\right)^{-1/2} \left\| D_0'^{-i} v \right\|_2^{-1} \\ &\leq \left(\frac{\left|\lambda_{\max}\left(D_0{}^i D_0'^i\right)\right|}{2\pi |\lambda_{\min}(K(C))|} \right)^{1/2} \|v\|_2^{-1}. \end{aligned}$$

Denote the last expression above by $\frac{1}{2bp}$. By the fact above, $v'z(\infty)$ has a pdf such as f, bounded by $\frac{1}{2bp}$. Letting $\mathbb{F}$ be CDF of $|v'z(\infty)|$, we have

$$\mathbb{F}(b\delta) = \int_{-b\delta}^{b\delta} f(y) dy \leq 2b\delta \sup_{-b\delta \leq y \leq b\delta} f(y) \leq 2b\delta \sup_{y \in \mathbb{R}} f_1(y) \leq \frac{\delta}{p},$$

which by $|P_i' z(\infty)| \geq |v' z(\infty)|$, implies $\mathbb{F}_i^{-1}\left(\frac{\delta}{p}\right) \geq b\delta$. Plugging in (4.24), we get the desired result. □

Proof of Corollary 4.2. Indeed, we prove that if the followings hold, then, we have

$$\left|\left|\left| \hat{D}_n - D_0 \right|\right|\right|_2 \leq \epsilon,$$

with probability at least $1 - 4\delta$. Letting $N_{4.2}(\cdot, \cdot)$ be as defined by (4.16), and (4.17) in the proof of Theorem 4.2, suppose that

$$\begin{aligned} n &\geq N_{4.2}\left(\lambda_0,\delta\right)+1, && (4.25)\\ \lambda_0\epsilon &\geq \rho\nu_n\left(\delta\right)n^{\mu(D_0)-1}\left|\lambda_{\min}\left(D_0\right)\right|^{-n+1}, && (4.26)\end{aligned}$$

where

$$\begin{aligned} \lambda_0 &= \frac{1}{2}\phi\left(D_0\right)^2\psi\left(D_0,\delta\right)^2,\\ \rho &= p^{1/2}\xi\left(D_0,\delta\right)\eta\left(D_0^{-1}\right)\left|\!\left|\!\left|P^{-1}\right|\!\right|\!\right|_{\infty\to 2}|\!|\!|P|\!|\!|_\infty e^{|\lambda_{\min}(D_0)|}.\end{aligned}$$

First, by Theorem 4.2, (4.25) implies that on the event $\mathcal{W}\bigcap\mathcal{V}$,

$$\left|\lambda_{\min}\left(D_0^{-n+1}V_nD_0'^{-n+1}\right)\right|\geq\lambda_0, \tag{4.27}$$

with probability at least $1-2\delta$. According to Lemma 4.1 and Lemma 4.2, regularity, in addition to reachability, imply $\lambda_0>0$. Thus,

$$\hat{D}_n=\sum_{t=0}^{n-1}x(t+1)x(t)'V_n^{-1}=D_0+U_nD_0'^{n-1}V_n^{-1},$$

where $U_n=\sum\limits_{t=0}^{n-1}w(t+1)x(t)'D_0'^{-n+1}$, which leads to

$$\left|\!\left|\!\left|\hat{D}_n-D_0\right|\!\right|\!\right|_2\leq\frac{|\!|\!|U_n|\!|\!|_2\left|\!\left|\!\left|D_0{}^{-n+1}\right|\!\right|\!\right|_2}{\left|\lambda_{\min}\left(D_0{}^{-n+1}V_nD_0{}^{-n+1}\right)\right|}. \tag{4.28}$$

Since $x(t)=D_0^tz(t)$, Lemma 2.6 and Lemma 4.4 imply that on the event $\mathcal{W}\bigcap\mathcal{V}$,

$$|\!|\!|U_n|\!|\!|_2\leq\sum_{t=0}^{n-1}\|w(t+1)\|_2\left\|D_0^{-n+t+1}z(t)\right\|_2\leq p^{1/2}\nu_n\left(\delta\right)\xi\left(D_0,\delta\right)\eta\left(D_0^{-1}\right) \tag{4.29}$$

Plugging (4.27) and (4.29) in (4.28), and using (4.19), we get

$$\left|\left|\left|\hat{D}_n - D_0\right|\right|\right|_2 \leq \frac{\rho}{\lambda_0} \nu_n(\delta)\, n^{\mu(D_0)-1} \left|\lambda_{\min}(D_0)\right|^{n-1},$$

which by (4.26) is at most ϵ, holding with probability at least $1-2\delta$ on $\mathcal{W} \bigcap \mathcal{V}$. □

4.6.3 Proofs of Section 4.5

Proof of Lemma 4.3. Assume $X \in \mathbb{R}^{p\times p}$ has an eigenvalue of unit size, denoted by $\lambda \in \mathbb{C}, |\lambda| = 1$. Further, define the space of eigenvectors in $\mathbb{C}^p$ as follows. First, consider the equivalence relation $\sim$ on $\mathbb{C}^p$, defined as

$$x \sim y, \text{ if } x = cy \text{ for some } c \in \mathbb{C}, c \neq 0.$$

Letting $S = \frac{\mathbb{C}^p}{\sim}$ be the direction space in $\mathbb{C}^p$, we have $\dim_{\mathbb{C}}(S) = p-1$, i.e.

$$\dim_{\mathbb{R}}(S) = 2p-2.$$

Note that for every matrix $Y \in \mathbb{C}^{p\times p}$ and every vector $v \in \mathbb{C}^p$, $Yv = 0$ if and only if $Y\tilde{v} = 0$ for every $\tilde{v} \sim v$. Thus, $\det(X - \lambda I_p) = 0$ implies that there is $v \in S, v \neq 0$, such that

$$Xv = \lambda v \tag{4.30}$$

Denote the set of all matrices X satisfying (4.30) by $\mathcal{X}(\lambda, v) \subset \mathbb{R}^{p\times p}$. Separating real and imaginary parts, we get

$$\begin{aligned} X\Re(v) &= \Re(\lambda v), \\ X\Im(v) &= \Im(\lambda v). \end{aligned}$$

Then, we partition S to

$$S = S_1 \bigcup S_2, S_1 \bigcap S_2 = \emptyset,$$

where

$$\begin{aligned} S_1 &= \{v \in S : \Re(v), \Im(v) \text{ are in-line } \}, \\ S_2 &= \{v \in S : \Re(v), \Im(v) \text{ are not in-line } \}. \end{aligned}$$

Whenever $v \in S_2$, for $j = 1, \cdots, p$, the j-th row of X needs to be in the intersection of two nonparallel hyperplanes $\mathcal{P}_1, \mathcal{P}_2 \subset \mathbb{R}^p$, where

$$\begin{aligned} \mathcal{P}_1 &= \{y \in \mathbb{R}^p : y'\Re(v) = \Re(\lambda v)_j\}, \\ \mathcal{P}_2 &= \{y \in \mathbb{R}^p : y'\Im(v) = \Im(\lambda v)_j\}. \end{aligned}$$

Since $\dim_{\mathbb{R}}(\mathcal{P}_1) \leq p-1$, $\dim_{\mathbb{R}}(\mathcal{P}_2) \leq p-1$, and $v \in S_2$ we have

$$\dim_{\mathbb{R}}(\mathcal{P}_1 \cap \mathcal{P}_2) \leq p-2.$$

Therefore, for $v \in S_2$, we have $\dim_{\mathbb{R}}(\mathcal{X}(\lambda, v)) \leq p(p-2)$. Because of $\dim_{\mathbb{R}}(|\lambda| = 1) = 1$, and $\dim_{\mathbb{R}}(S_2) \leq 2p-2$, we have

$$\dim_{\mathbb{R}}\left(\bigcup_{|\lambda|=1, v \in S_2} \mathcal{X}(\lambda, v)\right) \leq \dim_{\mathbb{R}}(|\lambda| = 1) + \dim_{\mathbb{R}}(S_2) + \dim_{\mathbb{R}}(\mathcal{X}(\lambda, v)) \leq p^2 - 1. \tag{4.31}$$

On the other hand, for $v \in S_1$, there is a real number, say $\alpha(v)$, such that $\Im(v) = \alpha(v)\Re(v)$. So, $\dim_{\mathbb{R}}(S_1) = p-1$, and for $v \in S_1$, we have $\mathcal{P}_1 = \mathcal{P}_2$, i.e.

$$\dim_{\mathbb{R}}(\mathcal{X}(\lambda, v)) \leq p(p-1),$$

and

$$0 = \alpha(v) X \Re(v) - X \Im(v) = \alpha(v) \Re(\lambda v) - \Im(\lambda v) = \left(1 + \alpha(v)^2\right) \Im(\lambda) \Re(v),$$

i.e. either $\Im(\lambda) = 0$, or $\Re(v) = 0$. Note that the latter case is impossible because it implies $v = 0$. So, since $\{|\lambda| = 1, \Im(\lambda) = 0\} = \{1, -1\}$ is of dimension zero,

$$\dim_{\mathbb{R}} \left(\bigcup_{\lambda=-1,1} \mathcal{X}(\lambda, v) \right) \leq \dim_{\mathbb{R}}(S_1) + \dim_{\mathbb{R}}(\mathcal{X}(\lambda, v)) \leq p^2 - 1. \tag{4.32}$$

Therefore, letting $\mathcal{X} = \bigcup\limits_{\lambda,v} \mathcal{X}(\lambda, v)$, (4.31) and (4.32) imply $\dim_{\mathbb{R}}(\mathcal{X}) \leq p^2 - 1$, i.e. $\mathcal{X}$ is of zero Lebesgue measure in $\mathbb{R}^{p \times p}$.

To prove that irregular matrices are of zero Lebesgue measure, for $|\lambda| > 1$ define

$$\mathcal{Y}(\lambda) = \left\{ Y \in \mathbb{R}^{p \times p} : \operatorname{rank}(Y - \lambda I_p) < p - 1 \right\}.$$

First we show that for a fixed matrix $Y = [Y_1, \cdots, Y_p]$, there are at most $p-1$ values of λ such that $Y \in \mathcal{Y}(\lambda)$. Let $e_1, \cdots, e_p$ be the standard basis of $\mathbb{R}^p$. If $Y \in \mathcal{Y}(\lambda_0)$, two of $Y_i - \lambda_0 e_i, i = 1, \cdots, p$, such as $Y_{p-1} - \lambda_0 e_{p-1}, Y_p - \lambda_0 e_p$, can be written as a linear combinations of the others. There are at most $p-1$ values of λ_0 for which $Y_{p-1} - \lambda_0 e_{p-1}$ is a linear combination of $Y_1 - \lambda_0 e_1, \cdots, Y_{p-2} - \lambda_0 e_{p-2}$, since for every such a λ_0, $\det\left(\tilde{Y}\right) = 0$, where $\tilde{Y}$ is the square matrix whose columns are

$$Y_1 - \lambda_0 e_1, \cdots, Y_{p-1} - \lambda_0 e_{p-1},$$

removing an arbitrary row. Note that $\det\left(\tilde{Y}\right)$ is a polynomial of degree $p-1$.

Now, denote those values of λ by $\lambda_1(Y), \cdots, \lambda_m(Y)$, where $m \leq p-1$. For every

$i = 1, \cdots, m$, the dimension of subspace $\mathcal{P}_i$ spanned by

$$Y_1 - \lambda_i(Y)\, e_1, \cdots, Y_{p-1} - \lambda_i(Y)\, e_{p-1}, e_p$$

is at most $p-1$, which leads to

$$\dim_{\mathbb{R}} \left(\bigcup_{i=1}^{m} \mathcal{P}_i \right) \leq p-1.$$

Because $\lambda_i(Y)$ is uniquely determined by $Y_1, \cdots, Y_{p-1}$, so is $\mathcal{P}_i$. Therefore,

$$\dim_{\mathbb{R}} \left(\bigcup_{\lambda} \mathcal{Y}(\lambda) \right) \leq \dim_{\mathbb{R}} ([Y_1, \cdots, Y_{p-1}]) + \dim_{\mathbb{R}} \left(\bigcup_{i=1}^{m} \mathcal{P}_i \right) \leq p(p-1) + p - 1 = p^2 - 1,$$

which is the desired result. □

Proof of Theorem 4.3. As a well known fact, there is an invertible matrix $M \in \mathbb{R}^{p\times p}$, such that

$$\tilde{A} = M D_0 M^{-1} \in \mathbb{R}^{p\times p}$$

is a block diagonal matrix,

$$\tilde{A} = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix},$$

where for $i = 1, 2$, we have $D_i \in \mathbb{R}^{p_i \times p_i}$, $p_1 + p_2 = p$, and

$$|\lambda_{\max}(D_1)| < 1 < |\lambda_{\min}(D_2)|.$$

We split the original VAR process to two, which are evolving according to transition matrices D_1, and D_2. First, let

$$\tilde{C} = MCM' = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix},$$

where $C_{ij} \in \mathbb{R}^{p_i \times p_j}$ for $i = 1, 2$. Then, for $t = 0, 1, \cdots$, defining

$$\begin{aligned} \tilde{x}(t) &= Mx(t), \\ \tilde{w}(t+1) &= Mw(t+1), \end{aligned}$$

we have

$$\begin{aligned} \tilde{x}(t+1) &= M\left(D_0 x(t) + w(t+1)\right) \\ &= \tilde{A} M x(t) + M w(t+1) \\ &= \tilde{A} \tilde{x}(t) + \tilde{w}(t+1). \end{aligned}$$

Note that letting

$$\nu_{n+1}(\delta) = \max\left\{|||M|||_\infty, 1\right\} \left(b_2 \log \frac{b_1 p\,(n+1)}{\delta}\right)^{1/\alpha},$$

similar to Lemma 2.6, we have $\mathbb{P}(\mathcal{W}) \geq 1 - \delta$, for

$$\mathcal{W} = \left\{ \max_{1 \leq t \leq n+1} \max\left\{\|w(t)\|_\infty, \|\tilde{w}(t)\|_\infty\right\} \leq \nu_{n+1}(\delta) \right\}.$$

Let

$$\begin{aligned} \tilde{x}(t) &= \left[x^{(1)}(t)', x^{(2)}(t)'\right]', \\ \tilde{w}(t+1) &= \left[w^{(1)}(t+1)', w^{(2)}(t+1)'\right]', \end{aligned}$$

where for $i = 1, 2$,

$$x^{(i)}(t), w^{(i)}(t+1) \in \mathbb{R}^{p_i}.$$

Since $\tilde{A}$ is block diagonal, the processes $x^{(1)}(t), x^{(2)}(t)$ are separated:

$$\begin{aligned} x^{(i)}(t+1) &= D_i x^{(i)}(t) + w^{(i)}(t+1), \\ C_{ii} &= \mathbb{E}\left[w^{(i)}(t+1)w^{(i)}(t+1)'\right]. \end{aligned}$$

Both new processes inherit reachability from the original one.

Lemma 4.8. If $[D_0, C]$ is reachable, then for $i = 1, 2$, $[D_i, C_{ii}]$ is reachable as well.

Now, we define the following parameters, which will be used in the proof. Letting $D_2 = P^{-1}\Lambda_2 P$ be the Jordan decomposition of the explosive matrix D_2, and $K_1 = \sum_{t=0}^{\infty} D_1^t C_{11} {D_1'}^t$, define

$$\begin{aligned}
\rho_0 &= \frac{1}{2} - \frac{1}{2}\left(1 - \frac{|\lambda_{\min}(K_1)|}{9\,|\lambda_{\max}(K_1)|}\right)^{1/2}, \\
\rho_1 &= \frac{2^{1/2} p\,|\lambda_{\min}(D_2)|\,\xi(D_2,\delta)\,|||P'|||_{\infty\to 2}|||P'^{-1}|||_{\infty} e^{|\lambda_{\min}(D_2)|}}{\phi(D_2)\,\psi(D_2,\delta)}, \\
\rho_2 &= \frac{8\eta\left(D_2'^{-1}\right)^2 \xi(D_2,\delta)\,|||P^{-1}|||_{\infty\to 2}|||P|||_{\infty} e^{|\lambda_{\min}(D_2)|}}{\phi(D_2)^2\,\psi(D_2,\delta)^2}, \\
\rho_3 &= \frac{4\left(4\,|\lambda_{\min}(K_1)|^{-1}+3\right)^{1/2}|||M|||_2}{|\lambda_{\min}(K_1)|^{1/2}\rho_0}, \\
\rho_4 &= \frac{2p^{1/2}\xi(D_2,\delta)\,|||P^{-1}|||_{\infty\to 2}|||P|||_{\infty} e^{|\lambda_{\min}(D_2)|}}{\phi(D_2)\,\psi(D_2,\delta)}, \\
\rho_5 &= \frac{2|||P^{-1}|||_{\infty\to 2}|||P|||_{\infty} e^{|\lambda_{\min}(D_2)|}}{\phi(D_2)\,\psi(D_2,\delta)}, \\
\rho_6 &= \frac{|||P'|||_{\infty\to 2}|||P'^{-1}|||_{\infty} e^{|\lambda_{\min}(D_2)|}\,|\lambda_{\min}(K_1)|^{1/2}}{\phi(D_2)\,\psi(D_2,\delta)}.
\end{aligned}$$

Note that the constants ρ_0, ρ_3 do not depend on δ, and all other parameters depend on δ, only through $\xi(D_0,\delta)$ and $\psi(D_0,\delta)$. Using $N_{4.1}(\cdot,\cdot)$, and $N_{4.2}(\cdot,\cdot)$ defined in Theorem

4.1 and Theorem 4.2, respectively, suppose that the followings hold.

$$
\begin{aligned}
n &\geq N_{4.2}\left(\frac{\phi\left(D_2\right)^2 \psi\left(D_2, \delta\right)^2}{2}, \delta\right), && (4.33)\\
n &\geq 3N_{4.1}\left(\frac{\left|\lambda_{\min}\left(K_1\right)\right|}{2}, \delta\right), && (4.34)\\
\frac{\rho_0}{\rho_1} &\geq n^{\mu(D_2)-1/2}\left|\lambda_{\min}\left(D_2\right)\right|^{-2n/3}, && (4.35)\\
\frac{1}{\rho_2} &\geq \nu_{n+1}(\delta)\, n^{\mu(D_2)}\left|\lambda_{\min}\left(D_2\right)\right|^{-n/3}, && (4.36)\\
\frac{\epsilon}{3\rho_3\rho_4} &\geq \nu_{n+1}(\delta)\, n^{\mu(D_2)-1/2}\left|\lambda_{\min}\left(D_2\right)\right|^{-2n/3}, && (4.37)\\
\frac{\epsilon}{3\rho_3\rho_5} &\geq \nu_{n+1}(\delta)^2\, n^{\mu(D_2)+1/2}\left|\lambda_{\min}\left(D_2\right)\right|^{-n/3}, && (4.38)\\
\frac{1}{\rho_6} &\geq n^{\mu(D_2)-1/2}\left|\lambda_{\min}\left(D_2\right)\right|^{-n}. && (4.39)
\end{aligned}
$$

In addition, assume the followings.

$$
\begin{aligned}
\frac{n^2(n+1)^{-1}}{\left(\|x^{(1)}(0)\|_\infty + \nu_{n+1}(\delta)\right)^2 \nu_{n+1}(\delta)^2} &\geq \frac{8p\rho_3^2\eta\left(D_1\right)^2}{\epsilon^2}\log\left(\frac{4\left(p+p_1\right)}{\delta}\right), && (4.40)\\
\frac{n}{\nu_{n+1}(\delta)^2} &\geq \frac{72p^2\rho_3^2}{\epsilon^2}\log\left(\frac{4\left(p+p_2\right)}{\delta}\right),. && (4.41)
\end{aligned}
$$

We show that with probability at least $1-6\delta$, it holds that

$$\left|\!\left|\!\left|\hat{D}_{n+1} - D_0\right|\!\right|\!\right|_2 \leq \epsilon.$$

First,

$$MV_{n+1}M' = \sum_{t=0}^{n} \tilde{x}(t)\tilde{x}(t)' = \sum_{t=0}^{n} \begin{bmatrix} x^{(1)}(t) \\ x^{(2)}(t) \end{bmatrix} \left[x^{(1)}(t)', x^{(2)}(t)'\right] = \begin{bmatrix} V_{n+1}^{(1)} & Y_{n+1}' \\ Y_{n+1} & V_{n+1}^{(2)} \end{bmatrix},$$

where for $i = 1, 2$,

$$\begin{aligned} V_n^{(i)} &= \sum_{t=0}^{n-1} x^{(i)}(t) x^{(i)}(t)', \\ Y_n &= \sum_{t=0}^{n-1} x^{(2)}(t) x^{(1)}(t)'. \end{aligned}$$

Let the event $\mathcal{E} \subset \mathcal{W} \bigcap \mathcal{V}$ be the following:

$$\begin{aligned} \left| \lambda_{\min} \left(\frac{1}{n} V_{n+1}^{(1)} \right) \right| &\geq \frac{1}{2} \left| \lambda_{\min} \left(K_1 \right) \right|, \\ \left| \lambda_{\min} \left(D_2^{-n} V_{n+1}^{(2)} D_2'^{-n} \right) \right| &\geq \frac{1}{2} \phi \left(D_2 \right)^2 \psi \left(D_2, \delta \right)^2. \end{aligned}$$

According to Theorem 4.1, and Theorem 4.2, (4.33), (4.34) imply $\mathbb{P}(\mathcal{E}) > 1 - 5\delta$. Henceforth in the proof, we assume the event $\mathcal{E}$ holds. Define the invertible symmetric matrix

$$U_n = \begin{bmatrix} V_{n+1}^{(1)} & 0_{p_1 \times p_2} \\ 0_{p_2 \times p_1} & V_{n+1}^{(2)} \end{bmatrix}^{-1/2} \in \mathbb{R}^{p \times p},$$

and let

$$E_n = U_n M V_{n+1} M' U_n = \begin{bmatrix} I_{p_1} & V_{n+1}^{(1)\,-1/2} Y_{n+1}' V_{n+1}^{(2)\,-1/2} \\ V_{n+1}^{(2)\,-1/2} Y_{n+1} V_{n+1}^{(1)\,-1/2} & I_{p_2} \end{bmatrix}.$$

We show that

$$\left| \lambda_{\min} \left(E_n \right) \right| \geq \rho_0. \tag{4.42}$$

Let $m = \left\lceil \frac{n}{3} \right\rceil$, and $v_i \in \mathbb{R}^{p_i}$ for $i = 1, 2$, where

$$v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \in \mathbb{R}^p, \ \ \|v\|_2 = 1.$$

Then,

$$v'E_n v = 1 + 2v_2' {V_{n+1}^{(2)}}^{-1/2} Y_{n+1} {V_{n+1}^{(1)}}^{-1/2} v_1 = 1 + 2\mathbb{T}_1 + 2\mathbb{T}_2,$$

where

$$\begin{aligned}
\mathbb{T}_1 &= \sum_{t=0}^{m} v_2' {V_{n+1}^{(2)}}^{-1/2} x^{(2)}(t) v_1' {V_{n+1}^{(1)}}^{-1/2} x^{(1)}(t), \\
\mathbb{T}_2 &= \sum_{t=m+1}^{n} v_2' {V_{n+1}^{(2)}}^{-1/2} x^{(2)}(t) v_1' {V_{n+1}^{(1)}}^{-1/2} x^{(1)}(t).
\end{aligned}$$

By Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathbb{T}_1^2 &\leq \left(v_1' {V_{n+1}^{(1)}}^{-1/2} V_{m+1}^{(1)} {V_{n+1}^{(1)}}^{-1/2} v_1 \right) \left(v_2' {V_{n+1}^{(2)}}^{-1/2} V_{m+1}^{(2)} {V_{n+1}^{(2)}}^{-1/2} v_2 \right) \\
&\leq \|v_1\|_2^2 \|v_2\|_2^2 \left| \lambda_{\max} \left({V_{n+1}^{(2)}}^{-1/2} D_2^n D_2^{-n} V_{m+1}^{(2)} {D_2'}^{-n} {D_2'}^{n} {V_{n+1}^{(2)}}^{-1/2} \right) \right| \\
&\leq \|v_1\|_2^2 \|v_2\|_2^2 \left| \lambda_{\max} \left(D_2^{-n} V_{m+1}^{(2)} {D_2'}^{-n} \right) \right| \left| \lambda_{\max} \left({V_{n+1}^{(2)}}^{-1/2} D_2^n {D_2'}^{n} {V_{n+1}^{(2)}}^{-1/2} \right) \right|.
\end{aligned}$$

Letting $z(t) = D_2^{-t} x^{(2)}(t)$, by Lemma 4.4 we have

$$\begin{aligned}
\left| \lambda_{\max} \left(D_2^{-n} V_{m+1}^{(2)} {D_2'}^{-n} \right) \right| &\leq \sum_{t=0}^{m} \|z(t)\|_2^2 \left\| {D_2'}^{-n+t} \right\|_2^2 \\
&\leq \xi (D_2, \delta)^2 \, \|P'\|_{\infty \to 2}^2 \left\| {P'}^{-1} \right\|_\infty^2 \sum_{t=0}^{m} \eta_{n-t} \left(\Lambda_2^{-1} \right)^2, \\
\left| \lambda_{\max} \left({V_{n+1}^{(2)}}^{-1/2} D_2^n {D_2'}^{n} {V_{n+1}^{(2)}}^{-1/2} \right) \right| &\leq \operatorname{tr} \left({V_{n+1}^{(2)}}^{-1/2} D_2^n {D_2'}^{n} {V_{n+1}^{(2)}}^{-1/2} \right) \\
&= \operatorname{tr} \left({D_2'}^{n} {V_{n+1}^{(2)}}^{-1} D_2^n \right) \\
&\leq p \left| \lambda_{\min} \left({D_2}^{-n} V_{n+1}^{(2)} {D_2'}^{-n} \right) \right|^{-1} \\
&\leq 2p \phi (D_2)^{-2} \, \psi (D_2, \delta)^{-2}.
\end{aligned}$$

According to (4.19),

$$\sum_{t=0}^{m} \eta_t \left(\Lambda_2^{-1} \right)^2 \leq e^{2|\lambda_{\min}(D_2)|} n^{2\mu(D_2)-1} \left| \lambda_{\min} (D_2) \right|^{2m-2n}.$$

So, by (4.35), we have

$$\mathbb{T}_1 \leq \|v_1\|_2 \|v_2\|_2 \rho_1 n^{\mu(D_2)-1/2} \left|\lambda_{\min}\left(D_2\right)\right|^{-2n/3} \leq \rho_0 \|v_1\|_2 \|v_2\|_2. \tag{4.43}$$

Similarly, Cauchy-Schwarz inequality implies

$$\mathbb{T}_2 \leq \|v_2\|_2 \left(\sum_{t=m+1}^{n} \left(v_1' {V_{n+1}^{(1)}}^{-1/2} x^{(1)}(t) \right)^2 \right)^{1/2} \leq \|v_1\|_2 \|v_2\|_2 \left(1 - 2\rho_0\right), \tag{4.44}$$

because according to Theorem 4.1, (4.34) implies

$$\begin{aligned}
\sum_{t=m+1}^{n} \left(v_1' {V_{n+1}^{(1)}}^{-1/2} x^{(1)}(t) \right)^2 &= v_1' {V_{n+1}^{(1)}}^{-1/2} \left(V_{n+1}^{(1)} - V_{m+1}^{(1)} \right) {V_{n+1}^{(1)}}^{-1/2} v_1 \\
&\leq \|v_1\|_2^2 \left(1 - \left| \lambda_{\min}\left({V_{n+1}^{(1)}}^{-1/2} V_{m+1}^{(1)} {V_{n+1}^{(1)}}^{-1/2} \right) \right| \right) \\
&\leq \|v_1\|_2^2 \left(1 - \frac{\left| \lambda_{\min}\left(V_{m+1}^{(1)} \right) \right|}{\left| \lambda_{\max}\left(V_{n+1}^{(1)} \right) \right|} \right) \\
&\leq \|v_1\|_2^2 \left(1 - \frac{m \left|\lambda_{\min}\left(K_1\right)\right|}{3n \left|\lambda_{\max}\left(K_1\right)\right|} \right) \\
&\leq \|v_1\|_2^2 \left(1 - \frac{\left|\lambda_{\min}\left(K_1\right)\right|}{9 \left|\lambda_{\max}\left(K_1\right)\right|} \right) \\
&= \|v_1\|_2^2 \left(1 - 2\rho_0\right)^2.
\end{aligned}$$

Thus, by (4.43) and (4.44), for arbitrary unit vector v we have

$$\begin{aligned}
v' E_n v &\geq \|v_1\|_2^2 + \|v_2\|_2^2 - 2\|v_1\|_2 \|v_2\|_2 \left(\rho_0 + 1 - 2\rho_0\right) \\
&= \rho_0 \left(\|v_1\|_2^2 + \|v_2\|_2^2 \right) + \left(1 - \rho_0\right) \left(\|v_1\|_2 - \|v_2\|_2 \right)^2,
\end{aligned}$$

i.e. (4.42) holds.

Then, define

$$\Sigma_n = V_{m+1}^{(2)}(m) + \sum_{t=m+1}^{n} D_2^{t-m} x^{(2)}(m) x^{(2)}(m)' {D_2'}^{t-m}.$$

Since

$$V_{n+1}^{(2)} - \Sigma_n = \sum_{t=m+1}^{n} D_2^t \left(z(t)z(t)' - z(m)z(m)' \right) D_2'^{\,t},$$

and for $m+1 \le t \le n$, according to (4.19),

$$\begin{aligned}
\|z(t) - z(m)\|_2 &\le \sum_{i=m+1}^{t} \left\| D_2^{-i} w^{(2)}(i) \right\|_2 \\
&\le \left\|\left| P^{-1} \right|\right\|_{\infty\to 2} \||P\||_\infty \nu_{n+1}(\delta) \sum_{i=m+1}^{t} \eta_i \left(\Lambda_2^{-1} \right) \\
&\le \left\|\left| P^{-1} \right|\right\|_{\infty\to 2} \||P\||_\infty \nu_{n+1}(\delta)\, e^{|\lambda_{\min}(D_2)|} t^{\mu(D_2)} \left| \lambda_{\min}(D_2) \right|^{-m-1},
\end{aligned}$$

using Lemma 4.4, we have

$$\begin{aligned}
\left| \lambda_{\max} \left(D_2^{-n} \left(V_{n+1}^{(2)} - \Sigma_n \right) D_2'^{\,-n} \right) \right| &\le 2\xi(D_2, \delta) \sum_{t=m+1}^{n} \|z(t) - z(m)\|_2 \left\|\left| D_2'^{\,-n+t} \right|\right\|_2^2 \\
&\le 2\xi(D_2, \delta)\, \eta \left(D_2'^{\,-1} \right)^2 \max_{m+1 \le t \le n} \|z(t) - z(m)\|_2 \\
&\le \frac{1}{4} \rho_2 \phi(D_2)^2 \psi(D_2, \delta)^2 \nu_{n+1}(\delta)\, n^{\mu(D_2)} \left| \lambda_{\min}(D_2) \right|^{-n/3} \\
&\le \frac{1}{2} \left| \lambda_{\min} \left(D_2^{-n} V_{n+1}^{(2)} D_2'^{\,-n} \right) \right|.
\end{aligned}$$

In the last inequality above, (4.36) is used. Hence,

$$\left| \lambda_{\max} \left(V_{n+1}^{(2)} - \Sigma_n \right) \right| \le \frac{1}{2} \left| \lambda_{\min} \left(V_{n+1}^{(2)} \right) \right|,$$

which implies

$$\left\|\left| \Sigma_n^{1/2} V_{n+1}^{(2)\,-1/2} \right|\right\|_2^2 = \left| \lambda_{\max} \left(I_{p_2} + V_{n+1}^{(2)\,-1/2} \left(\Sigma_n - V_{n+1}^{(2)} \right) V_{n+1}^{(2)\,-1/2} \right) \right| \le \frac{3}{2}.$$

Therefore, letting

$$\tilde{U}_n = \begin{bmatrix} n^{-1/2} I_{p_1} & 0_{p_1 \times p_2} \\ 0_{p_2 \times p_1} & \Sigma_n^{-1/2} \end{bmatrix},$$

we have

$$\left\|\left|\tilde{U}_n^{-1}U_n\right|\right\|_2^2 \leq \left\|\left|n^{1/2}V_{n+1}^{(1)\ -1/2}\right|\right\|_2^2 + \left\|\left|\Sigma_n^{1/2}V_{n+1}^{(2)\ -1/2}\right|\right\|_2^2 \leq \frac{2}{|\lambda_{\min}(K_1)|} + \frac{3}{2}. \qquad (4.45)$$

To proceed, define the following matrices:

$$\begin{aligned} G_n &= n^{-1}\sum_{t=0}^{n} w(t+1)x^{(1)}(t)', \\ H_n &= n^{-1/2}\sum_{t=0}^{n} w(t+1)x^{(2)}(t)'\Sigma_n^{-1/2}. \end{aligned}$$

Further, for $t = 0, \cdots, n+1$, define the sigma-fields $\mathcal{F}_t = \sigma(w(1), \cdots, w(t))$. Letting $\Phi(\cdot)$ be as defined in the proof of Theorem 4.1, and

$$X_t = \Phi\left(w(t+1)x^{(1)}(t)'\right)$$

be a martingale difference sequence of symmetric matrices with respect to $\{\mathcal{F}_t\}_{t=0}^n$, all matrices

$$p\left(\eta(D_1)\left(\left\|x^{(1)}(0)\right\|_\infty + \nu_{n+1}(\delta)\right)\nu_{n+1}(\delta)\right)^2 I_{p+p_1} - X_t^2$$

are by Lemma 2.7 positive semidefinite. Letting

$$\sigma^2 = p\eta(D_1)^2\left(\left\|x^{(1)}(0)\right\|_\infty + \nu_{n+1}(\delta)\right)^2 \nu_{n+1}(\delta)^2(n+1),$$

according to Lemma 2.11, by (4.40) we have

$$\mathbb{P}\left(\|G_n\|_2 > \frac{\epsilon}{\rho_3}\right) = \mathbb{P}\left(\left|\lambda_{\max}\left(\sum_{t=0}^{n} X_t\right)\right| > n\frac{\epsilon}{\rho_3}\right) \leq 2(p+p_1)\exp\left(-\frac{n^2\epsilon^2}{8\sigma^2\rho_3^2}\right) \leq \frac{\delta}{2}. \qquad (4.46)$$

On the other hand, $\|H_n\|_2$ can be upper bounded as well. Indeed, using (4.19) and Lemma

4.4, we have

$$\begin{aligned}
& \left|\left|\left| n^{-1/2} \sum_{t=0}^{m-1} w(t+1) x^{(2)}(t)' \Sigma_n^{-1/2} \right|\right|\right|_2 \\
\leq\ & n^{-1/2} \sum_{t=0}^{m-1} \|w(t+1)\|_2 \|z(t)\|_2 \left|\left|\left| \Sigma_n^{-1/2} D_2^t \right|\right|\right|_2 \\
\leq\ & p^{1/2} n^{-1/2} \nu_{n+1}(\delta)\, \xi(D_2, \delta) \left|\left|\left| \Sigma_n^{-1/2} D_2^n \right|\right|\right|_2 \sum_{t=n-m+1}^{n} \left|\left|\left| D_2^{-t} \right|\right|\right|_2 \\
\leq\ & \rho_4 \nu_{n+1}(\delta)\, n^{\mu(D_2)-1/2} \left|\lambda_{\min}(D_2)\right|^{-2n/3}.
\end{aligned}$$

Thus, by (4.37) we have

$$\left|\left|\left| n^{-1/2} \sum_{t=0}^{m-1} w(t+1) x^{(2)}(t)' \Sigma_n^{-1/2} \right|\right|\right|_2 \leq \frac{\epsilon}{3\rho_3}. \tag{4.47}$$

Moreover, for $t = m, \cdots, n$, letting

$$\tilde{X}_t = \Phi\left(w(t+1) x^{(2)}(m)' {D_2'}^{t-m} \Sigma_n^{-1/2} \right)$$

be a martingale difference sequence with respect to $\{\mathcal{F}_t\}_{t=m}^n$ (note that both Σ_n and $x^{(2)}(m)$ are $\mathcal{F}_m$ measurable), all matrices

$$\left(p^{1/2} \nu_{n+1}(\delta) \left\| \Sigma_n^{-1/2} D_2^{t-m} x^{(2)}(m) \right\|_2 \right)^2 I_{p+p_2} - \tilde{X}_t^2$$

are positive semidefinite, so, according to Lemma 2.11, we have

$$\mathbb{P}\left(\left| \lambda_{\max}\left(\sum_{t=m}^{n} \tilde{X}_t \right) \right| > \frac{n^{1/2}\epsilon}{3\rho_3} \middle| \mathcal{F}_m \right) \quad \leq \quad 2\left(p+p_2\right) \exp\left(-\frac{n\epsilon^2}{72\sigma^2\rho_3^2} \right),$$

where

$$\begin{aligned}
\sigma^2 &= \sum_{t=m}^{n} \left(p^{1/2}\nu_{n+1}(\delta) \left\|\Sigma_n^{-1/2} D_2^{t-m} x^{(2)}(m)\right\|_2\right)^2 \\
&= p\nu_{n+1}(\delta)^2 \sum_{t=m}^{n} \left(D_2^{t-m} x^{(2)}(m)\right)' \Sigma_n^{-1} D_2^{t-m} x^{(2)}(m) \\
&= p\nu_{n+1}(\delta)^2 \operatorname{tr}\left(\Sigma_n^{-1} \sum_{t=m}^{n} D_2^{t-m} x^{(2)}(m) x^{(2)}(m)' {D_2'}^{t-m}\right) \le p^2 \nu_{n+1}(\delta)^2 .
\end{aligned}$$

The last inequality above is simply implied by the definition of Σ_n. Now, applying (4.41), we get

$$\mathbb{P}\left(\left|\left|\left|\sum_{t=m}^{n} w(t+1) x^{(2)}(m)' {D_2'}^{t-m} \Sigma_n^{-1/2}\right|\right|\right|_2 > \frac{n^{1/2}\epsilon}{3\rho_3}\right) \le \frac{\delta}{2}. \tag{4.48}$$

Since for $t = m, \cdots, n$,

$$\begin{aligned}
& \left\|\Sigma_n^{-1/2} x^{(2)}(t) - \Sigma_n^{-1/2} D_2^{t-m} x^{(2)}(m)\right\|_2 \\
= \; & \left\|\Sigma_n^{-1/2} D_2^{n} D_2^{-n+t} \left(z(t) - z(m)\right)\right\|_2 \\
\le \; & \left|\lambda_{\min}\left(D_2^{-n} \Sigma_n {D_2'}^{-n}\right)\right|^{-1/2} \left\|\sum_{i=m+1}^{t} D_2^{-n+t-i} w^{(2)}(i)\right\|_2 \\
\le \; & \frac{|||P^{-1}|||_{\infty\to 2} |||P|||_{\infty} e^{|\lambda_{\min}(D_2)|} \nu_{n+1}(\delta)\, n^{\mu(D_2)} \left|\lambda_{\min}(D_2)\right|^{-m-1}}{\left|\lambda_{\min}\left(D_2^{-n} \Sigma_n {D_2'}^{-n}\right)\right|^{1/2}} \\
\le \; & \rho_5 \nu_{n+1}(\delta)\, n^{\mu(D_2)} \left|\lambda_{\min}(D_2)\right|^{-n/3},
\end{aligned}$$

by (4.38),

$$\left|\left|\left|\sum_{t=m}^{n} w(t+1) \left(x^{(2)}(t)' - x^{(2)}(m)' {D_2'}^{t-m}\right) \Sigma_n^{-1/2}\right|\right|\right|_2 \le \frac{n^{1/2}\epsilon}{3\rho_3}.$$

So, (4.48) implies that the following holds, with probability at least $1 - \frac{\delta}{2}$.

$$\left|\left|\left|n^{-1/2} \sum_{t=m}^{n} w(t+1) x^{(2)}(t)' \Sigma_n^{-1/2}\right|\right|\right|_2 \le \frac{2\epsilon}{3\rho_3},$$

which, in addition to (4.47), yields

$$\mathbb{P}\left(|||H_n|||_2 > \frac{\epsilon}{\rho_3}\right) \leq \frac{\delta}{2}. \tag{4.49}$$

Finally, since the event $\mathcal{E}$ holds, (4.39) implies

$$\left|\left|\left|n^{1/2}U_n\right|\right|\right|_2 \leq \left|\left|\left|n^{1/2}V_{n+1}^{(1)\ -1/2}\right|\right|\right|_2 + \left|\left|\left|n^{1/2}V_{n+1}^{(2)\ -1/2}\right|\right|\right|_2 \leq 2^{3/2}\left|\lambda_{\min}\left(K_1\right)\right|^{-1/2}. \tag{4.50}$$

This will finish the proof as follows. Writing

$$\begin{aligned}
\hat{D}_{n+1} - D_0 &= \sum_{t=0}^{n} w(t+1)x(t)'V_{n+1}^{-1} \\
&= \left(n^{-1/2}\sum_{t=0}^{n-1} w(t+1)x(t)'M'\tilde{U}_n\right)\left(\tilde{U}_n^{-1}U_n\right)\left(U_nMV_{n+1}M'U_n\right)^{-1} n^{1/2}U_nM \\
&= [G_n, H_n]\left(\tilde{U}_n^{-1}U_n\right)E_n^{-1}n^{1/2}U_nM,
\end{aligned}$$

according to inequalities (4.42), (4.45), (4.46), (4.49), and (4.50), on the event $\mathcal{E}$, with probability at least $1-\delta$,

$$\begin{aligned}
\left|\left|\left|\hat{D}_{n+1} - D_0\right|\right|\right|_2 &\leq \left(|||G_n|||_2 + |||H_n|||_2\right)\left|\left|\left|\tilde{U}_n^{-1}U_n\right|\right|\right|_2 \left|\left|\left|E_n^{-1}\right|\right|\right|_2 \left|\left|\left|n^{1/2}U_n\right|\right|\right|_2 |||M|||_2 \\
&\leq \frac{2\epsilon}{\rho_3}\left(\frac{2}{\left|\lambda_{\min}\left(K_1\right)\right|} + \frac{3}{2}\right)^{1/2} \rho_0^{-1}2^{3/2}\left|\lambda_{\min}\left(K_1\right)\right|^{-1/2} |||M|||_2 = \epsilon,
\end{aligned}$$

which is the desired result. □

Proof of Lemma 4.8. Assume $\tilde{v} \in \mathbb{R}^{p_1}, \tilde{v} \neq 0$. We show that $[D_1, C_{11}]$ is reachable. Defining $v = [\tilde{v}', 0_{1\times p_2}]' \in \mathbb{R}^p$,

$$0 < \|M'v\|_2^2\left|\lambda_{\min}\left(K(C)\right)\right| \leq v'MK(C)M'v = v'\left(\sum_{j=0}^{p-1}\tilde{A}^j\tilde{C}\tilde{A}'^j\right)v = \tilde{v}'\left(\sum_{j=0}^{p-1}D_1^jC_{11}D_1'^{\,j}\right)\tilde{v},$$

so, the matrix $\sum_{j=0}^{p-1} D_1^j C_{11} D_1'^{\,j}$ is positive definite, or equivalently,

$$\text{rank}\left(\left[C_{11}^{1/2}, D_1 C_{11}^{1/2}, \cdots, D_1^{p-1} C_{11}^{1/2}\right]\right) = p_1. \tag{4.51}$$

But, by Cayley-Hamilton theorem, (4.51) is equivalent to

$$\text{rank}\left(\left[C_{11}^{1/2}, \cdots, D_1^{p_1-1} C_{11}^{1/2}\right]\right) = p_1,$$

which is nothing but the reachability of $[D_1, C_{11}]$. The proof for $[D_2, C_{22}]$ is similar. $\square$

CHAPTER 5

Future Works

We studied the adaptive control schemes for linear dynamical systems with quadratic costs. While the classical literatures mainly focus on the asymptotic analyses, the approach in this work is non-asymptotic. Based on Optimism in the Face of Uncertainty, we established non-asymptotic optimality results for a fairly general settings that certainty equivalence fails to address. Different scenarios with and without identifiability assumptions are considered, under the minimal assumption of stabilizability. Further, the probabilistic properties of the noise process is assumed to be sufficiently general, as it covers a class of heavy-tailed distributions.

Studying the non-asymptotic performance of the analogous reinforcement learning algorithms in different regimes is of potential interests. From a planning viewpoint, extensions to non-asymptotic analysis of optimality under imperfect observations can be a topic of future investigation. Another interesting aspect to scrutinize is trying to formulate sufficient and necessary conditions for the true dynamics, to ensure optimality of the traditional certainty equivalence procedure. On the other hand, approaches leaning to learning challenges such as consistency toward the true dynamics parameter, as well as those considering the problem in a high-dimensional settings (assuming sparsity), can be listed as interesting subjects to be addressed in the future.

In the reminder of this thesis, we studied the problem of providing finite sample bounds for the least-squares estimates of general VAR processes where the transition matrix does not need to be stable. The relationships between different parameters involved, such as sample size, accuracy of the estimation, failure probability, transition matrix, noise covariance matrix, and dimension are discussed. We prove that apart from a pathological case of zero Lebesgue measure, the least-squares estimator is with high probability accurate, if the sample size scales similar to standard results in learning theory, i.e. quadratic scaling with the inverse estimation error and logarithmic scaling with failure probability.

Such finite sample results for such a widely used model can be helpful for analogous results for more complicated models exhibiting temporal dependence, such as nonlinear time series. Further, the techniques used in this work to analyze the estimation accuracy when the transition matrix is not necessarily stable, provide insight for settings where more knowledge about the structure of the transition matrix is available. In particular, potential extensions to a high-dimensional setting (assuming that the transition matrix is sparse), or other structured classes such as low-rank matrices, is a topic of interest and for future investigation.

Bibliography

[1] Y. Bar-Shalom and E. Tse, "Dual effect, certainty equivalence, and separation in stochastic control," *IEEE Transactions on Automatic Control*, vol. 19, no. 5, pp. 494–500, 1974.

[2] T. L. Lai and C. Z. Wei, "Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems," *The Annals of Statistics*, pp. 154–166, 1982.

[3] A. Becker, P. Kumar, and C.-Z. Wei, "Adaptive control with the stochastic approximation algorithm: Geometry and convergence," *IEEE Transactions on Automatic Control*, vol. 30, no. 4, pp. 330–338, 1985.

[4] T. Lai and C.-Z. Wei, "Extended least squares and their applications to adaptive control and prediction in linear systems," *IEEE Transactions on Automatic Control*, vol. 31, no. 10, pp. 898–906, 1986.

[5] T. Lai, "Asymptotically efficient adaptive control in stochastic regression models," *Advances in Applied Mathematics*, vol. 7, no. 1, pp. 23–45, 1986.

[6] L. Guo and H. Chen, "Convergence rate of els based adaptive tracker," *Syst. Sci & Math. Sci*, vol. 1, pp. 131–138, 1988.

[7] H.-F. Chen and J.-F. Zhang, "Convergence rates in stochastic adaptive tracking," *International Journal of Control*, vol. 49, no. 6, pp. 1915–1935, 1989.

[8] T. L. Lai and Z. Ying, "Parallel recursive algorithms in asymptotically efficient adaptive control of linear stochastic systems," *SIAM journal on control and optimization*, vol. 29, no. 5, pp. 1091–1127, 1991.

[9] L. Guo and H.-F. Chen, "The åstrom-wittenmark self-tuning regulator revisited and els-based adaptive trackers," *IEEE Transactions on Automatic Control*, vol. 36, no. 7, pp. 802–812, 1991.

[10] P. Kumar, "Convergence of adaptive control schemes using least-squares parameter estimates," *IEEE Transactions on Automatic Control*, vol. 35, no. 4, pp. 416–424, 1990.

[11] P. R. Kumar and P. Varaiya, *Stochastic systems: Estimation, identification, and adaptive control*. SIAM, 2015.

[12] M. C. Campi and P. Kumar, "Adaptive linear quadratic gaussian control: the cost-biased approach revisited," *SIAM Journal on Control and Optimization*, vol. 36, no. 6, pp. 1890–1907, 1998.

[13] T. L. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Advances in applied mathematics*, vol. 6, no. 1, pp. 4–22, 1985.

[14] S. Bittanti and M. C. Campi, "Adaptive control of linear time invariant systems: the bet on the best principle," *Communications in Information & Systems*, vol. 6, no. 4, pp. 299–320, 2006.

[15] Y. Abbasi-Yadkori and C. Szepesvári, "Regret bounds for the adaptive control of linear quadratic systems." in *COLT*, 2011, pp. 1–26.

[16] M. Ibrahimi, A. Javanmard, and B. V. Roy, "Efficient reinforcement learning for high dimensional linear quadratic systems," in *Advances in Neural Information Processing Systems*, 2012, pp. 2636–2644.

[17] D. P. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific Belmont, MA, 1995, vol. 1, no. 2.

[18] Bruce M Brown et al. Martingale central limit theorems. *The Annals of Mathematical Statistics*, 42(1):59–66, 1971.

[19] M. K. Shirani Faradonbeh, A. Tewari, and G. Michailidis, "Optimality of fast matching algorithms for random networks with applications to structural controllability," *IEEE Transactions on Control of Network Systems*, 2016.

[20] H. Lütkepohl, *New introduction to multiple time series analysis*. Springer Science & Business Media, 2005.

[21] S. Basu and G. Michailidis, "Regularized estimation in sparse high-dimensional time series models," *The Annals of Statistics*, vol. 43, no. 4, pp. 1535–1567, 2015.

[22] T. Lai and C. Wei, "Asymptotic properties of multivariate weighted sums with applications to stochastic regression in linear dynamic systems," *Multivariate Analysis VI*, pp. 375–393, 1985.

[23] K. Juselius and Z. Mladenovic, *High inflation, hyperinflation and explosive roots: the case of Yugoslavia*. Institute of Economics, University of Copenhagen, 2002.

[24] B. Nielsen, "Analysis of coexplosive processes," *Econometric Theory*, vol. 26, no. 03, pp. 882–915, 2010.

[25] T. W. Anderson, "On asymptotic distributions of estimates of parameters of stochastic difference equations," *The Annals of Mathematical Statistics*, pp. 676–687, 1959.

[26] T. Lai and C. Wei, "Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters," *Journal of Multivariate Analysis*, vol. 13, no. 1, pp. 1–23, 1983.

[27] B. Nielsen, "Strong consistency results for least squares estimators in general vector autoregressions with deterministic terms," *Econometric Theory*, pp. 534–561, 2005.

[28] ——, "Order determination in general vector autoregressions," in *Time Series and Related Topics*. Institute of Mathematical Statistics, 2006, pp. 93–112.

[29] ——, "Singular vector autoregressions with deterministic terms: Strong consistency and lag order determination," *Discussion paper, Nuffield College*, 2009.

[30] J. A. Tropp, "User-friendly tail bounds for sums of random matrices," *Foundations of computational mathematics*, vol. 12, no. 4, pp. 389–434, 2012.

[31] T. Lai and C. Wei, "A note on martingale difference sequences satisfying the local Marcinkiewicz-Zygmund condition," *Bulletin of the Institute of Mathematics, Academia Sinica*, vol. 11, no. 1, p. 13, 1983.

[32] V. Kuznetsov and M. Mohri, "Generalization bounds for non-stationary mixing processes," *Machine Learning*, vol. 106, no. 1, pp. 93–117, 2017.

[33] ——, "Generalization bounds for time series prediction with non-stationary processes," in *International Conference on Algorithmic Learning Theory*. Springer, 2014, pp. 260–274.

[34] ——, "Learning theory and algorithms for forecasting non-stationary time series," in *Advances in neural information processing systems*, 2015, pp. 541–549.

[35] Y. Abbasi-Yadkori, D. Pál, and C. Szepesvári, "Online least squares estimation with self-normalized processes: An application to bandit problems," *arXiv preprint arXiv:1102.2670*, 2011.