Learning, Control, and Reduction for Markov Jump Systems

by

Zhe Du

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical and Computer Engineering)
in The University of Michigan
2023

Doctoral Committee:

Associate Professor Laura Balzano, Co-Chair
Associate Professor Necmiye Ozay, Co-Chair
Assistant Professor Salar Fattahi
Associate Professor Peter Seiler
Professor Lei Ying

Zhe Du

zhedu@umich.edu

ORCID: 0000-0002-8245-9215

Acknowledgments

This dissertation would not have been possible without the help from many people. First, I would like to express my deepest gratitude to my academic advisors Prof. Laura Balzano and Prof. Necmiye Ozay. They helped and supported me in every possible way. In addition to providing me with constructive feedback when I encountered technical difficulties in research, they never hesitated to reach out to me when I was in frustration. I am also grateful that I learned from them how to be a nice person, how to be sympathetic and care for others, and how to communicate, which, I believe, would help me even beyond being a researcher. I wish I could be like them someday and help and guide others like they did to me.

I am also grateful to the rest of my committee members: Prof. Salar Fattahi, Prof. Peter Seiler, and Prof. Lei Ying. They helped me in idea formulation, proposal defense, and the final dissertation defense. Through talking to them, I realized how my work can be viewed from different perspectives and have the potential to be improved and generalized.

I would like to extend my thankfulness to all of my amazing collaborators. Their different backgrounds, expertise, and personalities made the collaborations productive and enjoyable. I would not have been able to make the same amount of progress without their help and efforts.

Thanks should also go to all my course instructors as they have laid the foundation for my research. Special thanks to Prof. Demosthenis Teneketzis who taught me random processes and stochastic control. He showed me through his teaching how theory derivations can be so elegant and intriguing, which helped shape my research directions.

Table of Contents

List of Figures

List of Tables

List of Appendices

Abstract

Many real-world phenomena or systems with temporally changing dynamics can be effectively characterized by time-varying models — typically ensembles of simple models (modes) among which the active mode switches over time. Coupling different modes through switching increases the model capacity so that more complex behaviors can be explained, but this meanwhile brings new challenges: one cannot simply study the entire model by looking at individual modes separately, and the coupling structures need to be factored in. Because of this, recent advances for simple time-invariant models such as data-driven methodologies and sharp finite-sample guarantees are yet to be generalized to time-varying models.

An important class of time-varying models is given by the Markov jump linear systems (MJSs), where each MJS is made up of a collection of linear modes and a Markov chain modeling their switching. MJSs strike a good balance between describing complex temporal variations in dynamics and possessing simple solutions to many classical control problems. This dissertation focuses on MJSs and revisits classical problems including the identification, data-driven control, and model reduction for MJSs. With recent advances in machine learning, optimization, and statistics, this work seeks to address challenges incurred by the mode switching in MJSs and bring new perspectives to these problems. For an MJS with unknown dynamics, we propose an identification scheme to develop its model, which requires only a single data trajectory and is guaranteed to have near-optimal sample complexity. Then, by establishing novel perturbation analysis, two classical control problems, certainty equivalent control and adaptive control, are studied with focus on solving for the linear quadratic regulator (LQR) problems. For the latter, the proposed adaptive control method invokes our identification scheme and is guaranteed to achieve performance with sublinear regret. Sometimes, a designed or learned MJS may suffer from complexities incurred by the sheer number of modes. Using clustering techniques from unsupervised learning, we develop a model reduction scheme that constructs a reduced-mode MJS by grouping modes with similar dynamics, which provably and empirically approximates the original system.

Chapter 1
Introduction

Many real-world systems have behaviors that can change over time. For example, some systems have dynamics determined by the environment, such as the dependency of solar panel power output on the sunlight radiation. Certain systems may have components that naturally deteriorate, e.g., rubber aging, metal corrosion, battery drain, etc., which can affect the overall system dynamics. Distributed systems that operate based on information transmitted through communication networks may suffer from occasional packet losses or delays, either of which may drastically and adversely impact system operations. Outside of engineering, time-varying behaviors also show up in economic indicators, pandemic infection rates, chemical reaction rates, etc.

Time-invariant models are well studied and understood theoretically and practically, but they fall short of capturing dynamics that can change over time. One workaround is to reactively build a new temporary model every time changes in the dynamics are detected via expert knowledge. This approach lacks responsiveness and does not exploit the recurring nature of the dynamics. Moreover, due to the short lifespan of each temporary model, performing control would necessarily be "myopic" and yield suboptimal performance in the long run and, even worse, destabilize the system.

On the other hand, using time-varying models that proactively take future variations of dynamics into consideration is a more systematic choice. A time-varying model is typically an ensemble of simple models (modes) with only one mode being active at a time, and the active mode can switch over time. One important class of time-varying models is known as the discrete-time Markov jump linear systems (MJSs). An MJS is composed of a collection of linear modes and a Markov chain that governs the switching of the active mode. MJSs strike a balance between describing complex behaviors with temporal variations and possessing simple solutions to many classical control problems. MJSs have applications in power electronics (Stankovic et al., 1995), power systems (Loparo and Abdel-Malek, 1990; Ugrinovskii and Pota, 2005), manufacturing (Boukas and Liu, 2001), economics and finance (Cajueiro, 2002; Hatjispyros and Yannacopoulos, 2005), solar thermal receiver (Sworder and

Rogers, 1983), networked control with packet losses (Sinopoli et al., 2005; Truong et al., 2021), etc.

Having modes that can switch allows MJSs to characterize a wider class of real-world systems but also poses new challenges to a variety of problems.

- **Identification**: In an MJS, the Markov chain that models the mode switching is a random process by itself. This introduces inherent randomness and makes the MJS a stochastic system. Because of this, the stability of an MJS is typically considered in the expectation sense. Though the probability is small, an MJS that is stable in expectation may still produce unstable trajectory realizations. This calls for new analytical tools to study sample complexities of learning an MJS, since existing tools usually require deterministic stability.

- **Control**: The algebraic Riccati equation (ARE) is the key to solving for optimal control problems such as the linear quadratic regulator (LQR) problems. ARE perturbation analysis helps understand the suboptimality incurred by having imperfect knowledge of the model, which can happen when such knowledge is obtained from estimation using a finite amount of data. In terms of an MJS with switched modes, its ARE is a set of coupled equations, each of which correspond to an individual mode. The coupling structure in the ARE adds to the difficulty of establishing perturbation analysis.

- **Model Reduction**: Compared with non-switched systems, MJSs may suffer from a new type of model complexity: the number of modes. Having more modes increases the computation cost in the analysis and controller design, thus it is imperative to study mode reduction for MJSs. Though there is preliminary work on reducing Markov chains or hidden Markov models, this problem is yet to be systematically studied for MJSs.

1.1 Related Work

Many classical control problems for linear time-invariant (LTI) systems find their counterparts for MJSs. Optimal quadratic control for MJSs, where one seeks to design inputs to minimize a quadratic cost function in terms of both the states and inputs, has been studied in works such as Chizeck and Ji (1988); Abou-Kandil et al. (1995); Costa and do Val (2002), and the constrained case is considered in Costa et al. (1999); Vargas et al. (2013). When the state of the MJS is only partially observed, the state estimation in real time is known as filtering, which has been discussed in Ackerson and Fu (1970); Chang and Athans (1978); Blom

(1984); Allam et al. (2001); Smith and Seiler (2003). When one wants to perform optimal control on a partially observed MJS, the optimal solution usually follows from the separation principle — the optimal control and optimal filtering are conducted separately and then combined together to yield the optimal performance (Caines and Chen, 1985; Caines and Zhang, 1995; De Farias et al., 2000). When there are model uncertainties or system delays, sometimes the worst-case control performance is of interest. This is known as robust control and studied in Costa and do Val (1996); Cao and Lam (1999); Boukas et al. (2001); de Farias et al. (2002).

The aforementioned control problems require exact (or at least coarse) knowledge of the MJS model, which includes the dynamics of all modes and the underlying Markov chain transition probabilities. When the model, however, is unknown, one needs to learn the dynamics from one or more data trajectories generated by the MJS, which is known as the identification problem and studied in Cao and Lam (1999); Cinquemani et al. (2007); Özkan et al. (2014); Barao and Marques (2011); Lale et al. (2021). However, their contributions are mainly methodological and lack theoretical results in terms of finite-sample analysis. Toward this end, recent work (Sarkar et al., 2019) considers identifying the partially observed stochastic switched system with unknown orders, which can be considered as a simplified MJS with independent mode switching. Though finite-sample guarantees are provided, strong assumptions on the system stability and multiple independent trajectories are needed.

As for the MJS adaptive control, where the goal is to control an unknown-dynamics MJS, earlier works (Yang et al., 1990; Caines and Zhang, 1995; Xue and Guo, 2001; Tan et al., 2005; Baltaoglu et al., 2016) are more interested in closed-loop stability guarantees and asymptotic analysis. Lacking non-asymptotic suboptimality analysis makes it difficult to compare the performance of two adaptive control schemes that are both stabilizing. Jansch-Porto et al. (2020) proposes a model-free approach based on the policy gradient and provides suboptimality guarantees, but it requires multiple independent rollout trajectories to obtain an accurate enough estimate of the gradients, which can be impossible in practice.

When the model complexity of the MJS, e.g., the dimension of the states or the number of the modes, is large, it can make the computations intractable for tasks such as analysis, verification, or controller design. To address this issue, one can construct a reduced system with smaller model complexity and use it as a surrogate for the original one in the computations. In terms of the model reduction for MJS, the majority of the work aims at order reduction, i.e., the reduction of state dimension, using methods such as the $\mathcal{H}_\infty$ reduction, $\mathcal{H}_2$ reduction, and balanced truncation (Zhang et al., 2003; Kotsalis and Rantzer, 2010; Sun and Lam, 2016). However, model complexity incurred by the number of modes is rarely studied.

Other than the classical control scenarios, MJSs have been novelly applied to the convergence analysis of stochastic optimization and reinforcement learning methods. In Hu et al. (2017), it is shown that several stochastic optimization methods can be modeled by MJSs with independent mode switching; in Hu and Syed (2019), a large family of temporal difference (TD) methods for reinforcement learning (RL) are shown to be well approximated by MJSs. For both works, using control theoretic ideas developed for MJSs, sufficient conditions for the algorithm convergence and the corresponding convergence rates are provided. This MJS interpretation of optimization algorithms is further exploited in Zhang et al. (2021) to analyze first-order methods for stochastic games.

1.2 Contributions and Organization

This dissertation has made contributions toward the following problems for MJSs.

- **Identification** (Chapter 3): The problem of learning an unknown-dynamics MJS using a single data trajectory is considered. A method is proposed to learn the dynamics of each individual mode of the MJS as well as the Markov chain that governs the mode switching. Finite-sample guarantees are established, which show that the dynamics can be learned with estimation error rate $\mathcal{O}(1/\sqrt{T})$ with T being the trajectory length. Our analysis tackles the heavy-tailed statistical property of the data, which is a new challenge for MJSs compared with LTI systems.

- **Certainty Equivalent Control** (Chapter 4): We consider the scenario of designing controllers for LQR problems using inaccurate knowledge of the MJS dynamics. It is shown that the performance degradation can be upper bounded by $\mathcal{O}(\epsilon^2)$ where ϵ measures the knowledge inaccuracy. As an intermediate result, we also establish perturbation analysis for the coupled algebraic Riccati equations, which can be of independent interest and benefit problems such as filter design and the linear quadratic Gaussian (LQG) problems.

- **Adaptive Control** (Chapter 5): We look into solving for the LQR control problem without any prior knowledge of the MJS model. An epoch-based adaptive control framework is proposed: each epoch is driven by a temporary controller, and at the end of each epoch the model knowledge and the controller are updated. It is shown that the proposed adaptive control scheme is guaranteed to achieve performance with sublinear regret $\mathcal{O}(\sqrt{T})$, where T is the planning time horizon.

- **Model Reduction** (Chapter 6): For an MJS with a large number of modes, a clustering-based model reduction scheme is developed, which constructs a reduced MJS with fewer modes. The reduced MJS is guaranteed to approximate the original MJS under various metrics. Furthermore, both theoretically and empirically, we show how one can use the reduced MJS to analyze stability and design controllers with significant computational cost reduction while achieving guaranteed accuracy.

The organization of this dissertation is as follows: Chapter 2 introduces the preliminaries including notations and basics of Markov chains and MJSs; Chapter 3 - Chapter 6 respectively present the works on identification, certainty equivalent control, adaptive control, and model reduction for MJSs. Chapter 7 lists several future directions that can follow from the established work in this dissertation.

Chapter 2
Preliminaries

In this chapter, we introduce the preliminaries to be covered in this dissertation. The notations are provided in Section 2.1. The basics of Markov chains including convergence results and special properties are discussed in Section 2.2. Section 2.3 introduces the basics of MJSs, notions of stability, and the LQR problem. Note that this chapter seeks to provide supporting concepts and lemmas regarding Markov chains and MJSs to be used in this dissertation and does not intend to cover every fundamental aspect of them. Interested readers can refer to the references herein for more details.

2.1 Notations

2.1.1 Sets

We let $\mathbb{R}$ denote the set of real numbers and $\mathbb{N} := \{0, 1, 2, \ldots\}$ denote the set of natural numbers. Notations $\mathbb{R}_+$ and $\mathbb{N}_+$ are the sets of positive real numbers and positive integers respectively. For some positive integer $n \in \mathbb{N}_+$, we let $[n] := \{1, 2, \ldots, n\}$. When some positive integer $k \in \mathbb{N}_+$ shows up as the superscript of some set S, this denotes the k-ary Cartesian power of S, i.e., $S^k := \underbrace{S \times S \times \cdots \times S}_{k}$. We use $\mathbb{R}^n_\Delta := \{(x_1, \ldots, x_n) \in \mathbb{R}^n : \sum_i x_i = 1, x_j \geq 0, \forall j\}$ to denote the standard $(n-1)$-simplex. Similarly, for the space of $n \times n$ dimensional matrices where each row is in $\mathbb{R}^n_\Delta$, we denote it by $\mathbb{R}^{n \times n}_\Delta$. In this dissertation, the positive (semi-)definiteness of matrices is discussed for symmetric matrices only. Notations $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$ denote the sets of $n \times n$, positive semi-definite and positive definite matrices respectively. We use $a_{1:k}$ as a shorthand notation for a sequence of variables (or sets) $(a_1, a_2, \ldots, a_k)$. Given a set S and a collection of its subsets $\Omega_{1:r}$, where each $\Omega_i \subseteq S$, we say $\Omega_{1:r}$ is an r-cluster partition of S if (i) $\Omega_i \neq \phi$ for all i, (ii) $\bigcup_{i=1}^r \Omega_i = S$, and (iii) $\Omega_i \bigcap \Omega_j = \phi$ for any $i \neq j$. We let $\Omega_{(i)}$ denote the cluster with i-th largest cardinality.

2.1.2 Linear Algebra

We use boldface uppercase letters to denote matrices and boldface lowercase letters to denote vectors, e.g., matrix $\mathbf{A}$ and vector $\mathbf{a}$. For a matrix $\mathbf{A}$, $\mathbf{A}(i,j)$ indexes the (i,j)-th element in $\mathbf{A}$, $\mathbf{A}(i,:)$ denotes the row vector given by the i-th row of $\mathbf{A}$, and $\mathbf{A}(:,j)$ denotes the column vector given by the j-th column. $\mathbf{A}(i,j{:}k)$ denotes the row vector given by the i-th row preserving only the j-th to k-th elements. For any index set $\mathcal{I} \subset \mathbb{N}_+$, $\mathbf{A}(i,\mathcal{I})$ denotes the row vector given by the i-th row preserving elements indexed by $\mathcal{I}$.

The n-dimensional identity matrix is written as $\mathbf{I}_n$, and the n-dimensional all-ones vector is written as $\mathbf{1}_n$, where the dimension subscript n may be omitted when it is clear from the context. For a vector $\mathbf{a}$, $\mathbf{diag}(\mathbf{a})$ stands for the diagonalization operator; for a matrix $\mathbf{A}$ with n columns, we let $\mathbf{vec}(\mathbf{A}) := [\mathbf{A}(:,1)^\intercal, \ldots, \mathbf{A}(:,n)^\intercal]^\intercal$ denote the vectorization operator. The Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$ is denoted as $\mathbf{A} \otimes \mathbf{B}$.

Notation $\|\cdot\|_p$ for $1 \leq p \leq \infty$ denotes the p-norm for vectors or the induced p-norm for matrices. For conciseness, when $p = 2$, we omit the subscript and simply use $\|\cdot\|$. We let $\|\cdot\|_+ := \|\cdot\| + 1$. The matrix Frobenius norm of a matrix is denoted by $\|\cdot\|_\mathrm{F}$. For a matrix $\mathbf{A}$, $\sigma_i(\mathbf{A})$ denotes its i-th largest singular value, and its largest and smallest singular values may also be represented by $\bar{\sigma}(\mathbf{A})$ and $\underline{\sigma}(\mathbf{A})$. If $\mathbf{A}$ is a square matrix, $\lambda_i(\mathbf{A})$ denotes its i-th largest eigenvalue in terms of magnitude, and $\rho(\mathbf{A})$ denotes its spectral radius. When the decomposition $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}$ is referred to as the eigenvalue decomposition of the matrix $\mathbf{A}$, the eigenvalues on the diagonal of $\mathbf{\Lambda}$ are arranged in descending order of magnitude. The same applies to the singular value decomposition $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\intercal$.

A sequence of matrices $(\mathbf{A}_1, \ldots, \mathbf{A}_s)$ sharing the same dimensions is compactly denoted by $\mathbf{A}_{1:s}$. We let $\mathbb{R}_s^{m\times n} := \{\mathbf{A}_{1:s} : \mathbf{A}_i \in \mathbb{R}^{m\times n}, \forall\ i\}$, $\mathbb{S}_{s,+}^n := \{\mathbf{A}_{1:s} : \mathbf{A}_i \in \mathbb{S}_+^n, \forall\ i\}$, and $\mathbb{S}_{s,++}^n := \{\mathbf{A}_{1:s} : \mathbf{A}_i \in \mathbb{S}_{++}^n, \forall\ i\}$. If $\mathbf{A}_{1:s} \in \mathbb{S}_{s,+}^n$ (or $\mathbf{A}_{1:s} \in \mathbb{S}_{s,++}^n$), we also say $\mathbf{A}_{1:s} \succeq 0$ (or $\mathbf{A}_{1:s} \succ 0$). Operator $\mathbf{diag}(\mathbf{A}_{1:s})$ produces a block diagonal matrix whose i-th diagonal block is given by matrix $\mathbf{A}_i$. We define the norm operator $\|\mathbf{A}_{1:s}\| := \max_{i\in[s]} \|\mathbf{A}_i\|$ and $\|\mathbf{A}_{1:s}\|_+ = \|\mathbf{A}_{1:s}\| + 1$. Given two sequences of matrices $\mathbf{A}_{1:s}$ and $\mathbf{B}_{1:s}$, and scalars $\alpha, \beta \in \mathbb{R}$, the notation $\alpha\mathbf{A}_{1:s} + \beta\mathbf{B}_{1:s}$ stands for the sequence $(\alpha\mathbf{A}_1 + \beta\mathbf{B}_1, \ldots, \alpha\mathbf{A}_s + \beta\mathbf{B}_s)$. Given $\mathbf{A}_{1:s} \in \mathbb{R}_s^{m\times n}$, we let $\xi(\mathbf{A}_{1:s}) := \lim_{k\to\infty} \max_{\sigma_{1:k}\in[s]^k} \|\mathbf{A}_{\sigma_1}\cdots\mathbf{A}_{\sigma_k}\|^{\frac{1}{k}}$ denote the joint spectral radius of $\mathbf{A}_{1:s}$.

2.1.3 Probability

The multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{\Sigma}$ is denoted by $\mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$. For a zero-mean random vector $\mathbf{x}$, we let $\mathbf{Cov}(\mathbf{x}) := \mathbb{E}[\mathbf{x}\mathbf{x}^\intercal]$ denote its covariance matrix. Claims such as "for all j, with probability at least $1-\delta$, for all i, event $\mathcal{E}_{i,j}$ occurs" is equivalent to "for all j, $\mathbb{P}(\cap_i \mathcal{E}_{i,j}) \geq 1-\delta$". Given an event $\mathcal{E}$, we let $\mathbf{1}_{\{\mathcal{E}\}}$ denote the indicator function for $\mathcal{E}$. Orders of magnitude notation $\hat{\mathcal{O}}(\cdot)$ hides $\log(1/\delta)$ or $\log^2(1/\delta)$ terms.

2.2 Markov Chains

In this dissertation, we restrict our attention to Markov chains that are discrete-time, time-homogeneous, and have a finite number of states. We say the random process $\{\omega_t\}_{t\in\mathbb{N}}$ follows an s-state Markov chain if the state transition probabilities satisfy (i) $\mathbb{P}(\omega_{t+1}=i_{t+1} \mid \omega_t=i_t, \omega_{t-1}=i_{t-1}, \cdots, \omega_0=i_0) = \mathbb{P}(\omega_{t+1}=i_{t+1} \mid \omega_t=i_t)$ for all $t \in \mathbb{N}$ and $i_{0:t+1} \in [s]^{t+2}$, and (ii) $\mathbb{P}(\omega_{t_1+1}=i \mid \omega_{t_1}=j) = \mathbb{P}(\omega_{t_2+1}=i \mid \omega_{t_2}=j)$ for all $t_1, t_2 \in \mathbb{N}$ and $i, j \in [s]$. We use the Markov transition matrix $\mathbf{T} \in \mathbb{R}_{\Delta}^{s \times s}$ to collect those transition probabilities such that $\mathbf{T}(i,j) := \mathbb{P}(\omega_{t+1}=j \mid \omega_t=i)$, then the process can be described as "$\omega_t \sim \text{MarkovChain}(\mathbf{T})$".

We use the vector $\boldsymbol{\pi}_t \in \mathbb{R}_{\Delta}^{s}$ to denote the transient state distribution of the Markov chain such that $\boldsymbol{\pi}_t(i) := \mathbb{P}(\omega_t = i)$. Particularly, $\boldsymbol{\pi}_0$ is referred to as the initial distribution. The evolution of $\boldsymbol{\pi}_t$ is then given by $\boldsymbol{\pi}_{t+1}^\intercal = \boldsymbol{\pi}_t^\intercal \mathbf{T}$. We say the Markov chain is ergodic if it is irreducible, positive recurrent, and aperiodic (Gallager, 2013). Under ergodicity, it is known that regardless of the initial distribution $\boldsymbol{\pi}_0$, $\lim_{t\to\infty} \boldsymbol{\pi}_t$ uniquely exists, i.e., the Markov chain has a unique stationary distribution. We denote the stationary distribution by $\boldsymbol{\pi} \in \mathbb{R}_{\Delta}^{s}$ such that $\boldsymbol{\pi}(i) := \lim_{t\to\infty} \mathbb{P}(\omega_t = i) = \lim_{t\to\infty} \boldsymbol{\pi}_t(i)$. It is easy to see $\boldsymbol{\pi}^\intercal = \boldsymbol{\pi}^\intercal \mathbf{T}$. Let $\pi_{\max} := \max_i \boldsymbol{\pi}(i)$ and $\pi_{\min} := \min_i \boldsymbol{\pi}(i)$, then it is also known that $\pi_{\min} > 0$ under ergodicity.

2.2.1 Convergence of Markov Chains

To quantify the convergence rate of the Markov chain transient distribution to its stationary distribution, i.e. $\boldsymbol{\pi}_t \to \boldsymbol{\pi}$ as $t \to \infty$, we first introduce the following notion.

Definition 2.1 (Normalized Power Supremum)**.** *Consider an arbitrary square matrix* $\mathbf{A} \in \mathbb{R}^{n \times n}$*. Given a free parameter* ρ *such that* $\rho \geq \rho(\mathbf{A})$*, define*

$$\tau(\mathbf{A}, \rho) := \sup_{k\in\mathbb{N}} \|\mathbf{A}^k / \rho^k\|. \tag{2.1}$$

We call $\tau(\mathbf{A}, \rho)$ *the power supremum of the matrix* $\mathbf{A}$ *normalized by* ρ*.*

This notion is also studied in Mania et al. (2019). In the definition, ρ is a free parameter, and different choices of ρ yield different $\tau(\mathbf{A}, \rho)$. We can also see that $\tau(\mathbf{A}, \rho)$ monotonically decreases with respect to increasing ρ. Note that when $\rho > \rho(\mathbf{A})$, $\lim_{k\to\infty} \|\mathbf{A}^k/\rho^k\| = 0$ and $\|\mathbf{A}^0/\rho^0\| = 1$, hence by definition $\tau(\mathbf{A}, \rho)$ takes the supremum at some finite k. In other words, $\tau(\mathbf{A}, \rho)$ compares the transient behaviors of $\|\mathbf{A}^k\|$ and ρ^k. It is easy to see $\tau(\mathbf{A}, \rho)$ is bounded as long as $\rho > \rho(\mathbf{A})$. When the equal case is considered as well, i.e., $\rho \geq \rho(\mathbf{A})$, the boundedness of $\tau(\mathbf{A}, \rho)$ can be guaranteed when eigenvalues with magnitudes $\rho(\mathbf{A})$ have equal geometric and algebraic multiplicities. In other words, the Jordan blocks associated with the leading eigenvalues are trivial 1×1 matrices, i.e., scalars. When $\mathbf{A}$ is diagonalizable, i.e., all the Jordan blocks are trivial, an explicit upper bound is given by $\tau(\mathbf{A}, \rho) \leq \|\mathbf{V}\|\|\mathbf{V}^{-1}\|$, where $\mathbf{V}$ is a matrix whose columns span the eigenspace of $\mathbf{A}$. Particularly, when $\mathbf{A}$ is normal, $\tau(\mathbf{A}, \rho) = 1$ by the spectral theorem. For general matrix $\mathbf{A}$ that are not necessarily diagonalizable, by the Kreiss matrix theorem (Trefethen and Embree, 2005), $\tau(\mathbf{A}, \rho)$ can be bounded by the Kreiss constant of the normalized matrix $\mathbf{A}/\rho$. From Definition 2.1, the following properties are immediate.

Lemma 2.2 (Properties of the Normalized Power Supremum)**.** *For the pair $\{\rho, \tau(\mathbf{A}, \rho)\}$ in Definition 2.1, we have the following.*

(a) $\tau(\mathbf{A}, \rho) \geq 1$.

(b) For all $k \in \mathbb{N}$, $\|\mathbf{A}^k\| \leq \tau(\mathbf{A}, \rho)\rho^k$.

(c) If $\rho(\mathbf{A}) < 1$ and $\rho \in [\rho(\mathbf{A}), 1)$, $\sum_{k=0}^{\infty} \|\mathbf{A}^k\| \leq \frac{\tau(\mathbf{A},\rho)}{1-\rho}$.

As for the term $\frac{\tau(\mathbf{A},\rho)}{1-\rho}$ shows up in the upper bound of Lemma 2.2, since increasing ρ decreases $\tau(\mathbf{A}, \rho)$, the freedom of selecting ρ makes it possible to attain the tightest upper bound given by $\inf_{\rho \in [\rho(\mathbf{A}),1)} \frac{\tau(\mathbf{A},\rho)}{1-\rho}$. One will find the term $\frac{\tau(\mathbf{A},\rho)}{1-\rho}$ shows up in theoretical results throughout this dissertation.

Using the normalized power supremum in Definition 2.1, we define the following to analyze the convergence of Markov chains.

Definition 2.3 (Quantification of Markov Chain Convergence)**.** *Consider an ergodic Markov chain with Markov matrix $\mathbf{T} \in \mathbb{R}_{\Delta}^{s \times s}$ and stationary distribution $\boldsymbol{\pi} \in \mathbb{R}_{\Delta}^{s}$. Let $\mathbf{R} := \mathbf{T} - \mathbf{1}_s \boldsymbol{\pi}^{\intercal}$. Then for any $\rho_{\mathrm{MC}} \in [\rho(\mathbf{R}), 1)$, let $\tau_{\mathrm{MC}} := \tau(\mathbf{R}, \rho_{\mathrm{MC}})$, where $\tau(\cdot, \cdot)$ is as in Definition 2.1.*

By the Perron-Frobenius theorem and ergodicity, we know the leading eigenvalue of the Markov matrix $\mathbf{T}$ is 1, and only one eigenvalue has magnitude equal to 1. It can be verified that its eigenvalue component is given by $\mathbf{1}_s \boldsymbol{\pi}^{\intercal}$. We then see from the definition that the matrix $\mathbf{R}$ shares the same eigenvalues and eigenvectors as $\mathbf{T}$ except for replacing the eigenvalue 1 of $\mathbf{T}$ by 0. Hence, the spectral radius $\rho(\mathbf{R})$ is nothing but the magnitude of the second largest eigenvalue of $\mathbf{T}$, which is strictly smaller than 1. This guarantees that the range $[\rho(\mathbf{R}), 1)$ for ρ_{MC} in Definition 2.3 is valid. The pair $\{\rho_{\mathrm{MC}}, \tau_{\mathrm{MC}}\}$ can be used to quantify the convergence of Markov chains, which is provided in the following lemma.

Lemma 2.4 (Markov Chain Convergence)**.** *For $\{\rho_{\mathrm{MC}}, \tau_{\mathrm{MC}}\}$ in Definition 2.3, we have, for all $t \in \mathbb{N}$,*

$$\|\mathbf{T}^t - \mathbf{1}_s \boldsymbol{\pi}^\intercal\| \leq \tau_{\mathrm{MC}} \rho_{\mathrm{MC}}^t, \tag{2.2}$$

and for any initial distribution $\boldsymbol{\pi}_0$, the transient distribution $\boldsymbol{\pi}_t$ converges to the stationary distribution $\boldsymbol{\pi}$ with at least the following rate:

$$\|\boldsymbol{\pi}_t - \boldsymbol{\pi}\| \leq \tau_{\mathrm{MC}} \rho_{\mathrm{MC}}^t. \tag{2.3}$$

Proof. From Definition 2.3 and Lemma 2.2, we have $\|(\mathbf{T} - \mathbf{1}_s \boldsymbol{\pi}^\intercal)^t\| \leq \tau_{\mathrm{MC}} \rho_{\mathrm{MC}}^t$. From the discussion above regarding the eigenvalues and eigenvectors of $\mathbf{T}$, we can infer that $(\mathbf{T} - \mathbf{1}_s \boldsymbol{\pi}^\intercal)^t = \mathbf{T}^t - \mathbf{1}_s \boldsymbol{\pi}^\intercal$, which gives (2.2). As for (2.3), by noticing $\boldsymbol{\pi}_t = (\sum_{i \in [s]} \boldsymbol{\pi}_0(i) \mathbf{T}^t(i,:))^\intercal$, we obtain $\|\boldsymbol{\pi}_t - \boldsymbol{\pi}\| \leq \sum_{i \in [s]} \boldsymbol{\pi}_0(i) \|\mathbf{T}^t(i,:)^\intercal - \boldsymbol{\pi}\|$ using the triangle inequality. For each $\|\mathbf{T}^t(i,:)^\intercal - \boldsymbol{\pi}\|$, invoking (2.2) gives that $\|\mathbf{T}^t(i,:)^\intercal - \boldsymbol{\pi}\| \leq \|\mathbf{T}^t - \mathbf{1}_s \boldsymbol{\pi}^\intercal\| \leq \tau_{\mathrm{MC}} \rho_{\mathrm{MC}}^t$. Combining these results, (2.3) can be shown. □

Other than using $\{\rho_{\mathrm{MC}}, \tau_{\mathrm{MC}}\}$ to upper bound the convergence of Markov chains, the convergence can also be studied in terms of the minimum number of time steps to reach certain convergence tolerance, which is known as the mixing time.

Definition 2.5 (Markov Chain Mixing Time (Levin and Peres, 2017))**.** *Consider an ergodic Markov chain with Markov matrix $\mathbf{T} \in \mathbb{R}_\Delta^{s \times s}$ and stationary distribution $\boldsymbol{\pi} \in \mathbb{R}_\Delta^s$. Define*

$$t_{\mathrm{MC}}(\epsilon) := \min \left\{ t \in \mathbb{N} : \max_{i \in [s]} \frac{1}{2} \|(\mathbf{T}^t(i,:)^\intercal - \boldsymbol{\pi}\|_1 \leq \epsilon \right\}. \tag{2.4}$$

When the argument ϵ is omitted, it denotes that $t_{\mathrm{MC}} := t_{\mathrm{MC}}(\frac{1}{4})$.

In the above definition, $t_{\mathrm{MC}}(\frac{1}{4})$ is of particular interest since one can upper bound $t_{\mathrm{MC}}(\epsilon)$ for any $\epsilon < \frac{1}{4}$ using $t_{\mathrm{MC}}(\frac{1}{4})$ (see, for example, Zhang and Wang (2019, Lemma 5)). Using Lemma 2.4, we can derive that the mixing time and the pair $\{\rho_{\mathrm{MC}}, \tau_{\mathrm{MC}}\}$ in Definition 2.3 are related by $t_{\mathrm{MC}}(\epsilon) \leq \max\{0, \frac{\log(2\sqrt{s}\epsilon/\tau_{\mathrm{MC}})}{\log(\rho_{\mathrm{MC}})}\}$.

2.2.2 Some Special Markov Chains

Markov chains with the following special properties will be considered in Chapter 6 for the reduction of MJSs.

Definition 2.6 (Reversibility (Gallager, 2013))**.** *Consider an ergodic Markov chain with Markov matrix $\mathbf{T} \in \mathbb{R}_\Delta^{s \times s}$ and stationary distribution $\boldsymbol{\pi} \in \mathbb{R}_\Delta^s$. We say the Markov chain* *is reversible if the Markov matrix* $\mathbf{T}$ *and stationary distribution* $\boldsymbol{\pi}$ *satisfy* $\boldsymbol{\pi}(i)\mathbf{T}(i,j) = \boldsymbol{\pi}(j)\mathbf{T}(j,i)$ *for all* $i, j \in [s]$.

It can be seen that the condition for reversibility translates to $\mathbf{diag}(\boldsymbol{\pi})\mathbf{T} = \mathbf{T}^\intercal \mathbf{diag}(\boldsymbol{\pi})$. Reversibility tells that under the stationary distribution $\boldsymbol{\pi}$, the random process $\{\omega_t\}_{t\in\mathbb{N}}$ has the same statistical properties as its time-reversed process. Sometimes, certain states in the Markov chain can be lumped into a meta-state such that the process for these meta-states also follows a Markov chain. This makes it possible to reduce a large-scale Markov chain to a small-scale one. Such properties are characterized below.

Definition 2.7 (Lumpability and Aggregatability (Buchholz, 1994))**.** *Consider a Markov chain with Markov matrix* $\mathbf{T} \in \mathbb{R}_\Delta^{s\times s}$*. We say the Markov chain is lumpable with respect to the* r*-cluster partition* $\Omega_{1:r}$ *on the state space* $[s]$ *if states in the same cluster have equal probabilities of visiting any cluster, i.e., for any* $k, l \in [r]$ *and* $i, i' \in \Omega_k$*, we have*

$$\sum_{j\in\Omega_l} \mathbf{T}(i,j) = \sum_{j\in\Omega_l} \mathbf{T}(i',j). \tag{2.5}$$

If they further have equal probabilities of visiting any state, i.e., $\mathbf{T}(i,j) = \mathbf{T}(i',j)$ *for all* $j \in [s]$*, we say the Markov chain is aggregatable with respect to* $\Omega_{1:r}$.

2.3 Markov Jump Systems

In this dissertation, we study the following discrete-time MJS with dynamics given by

$$\Sigma := \begin{cases} \mathbf{x}_{t+1} & = \mathbf{A}_{\omega_t}\mathbf{x}_t + \mathbf{B}_{\omega_t}\mathbf{u}_t + \mathbf{w}_t \\ \quad \omega_t & \sim \text{Markov Chain}(\mathbf{T}) \end{cases} \tag{2.6}$$

where $\mathbf{x}_t \in \mathbb{R}^n$, $\mathbf{u}_t \in \mathbb{R}^p$, and $\mathbf{w}_t \in \mathbb{R}^n$ denote the state, input, and process noise at time t. We assume the noise $\{\mathbf{w}_t\}_{t\in\mathbb{N}} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_\mathbf{w}^2 \mathbf{I}_n)$, i.e., i.i.d. zero mean Gaussian with covariance $\sigma_\mathbf{w}^2 \mathbf{I}_n$. The dynamics is time-varying and can switch among s modes $\{\mathbf{A}_i, \mathbf{B}_i\}_{i=1}^s$ where $\mathbf{A}_i \in \mathbb{R}^{n\times n}$ and $\mathbf{B}_i \in \mathbb{R}^{n\times p}$ are state and input matrices for mode i. At time t, only one mode is active, which is indexed by $\omega_t \in [s]$. The mode switching sequence $\{\omega_t\}_{t\in\mathbb{N}}$ follows a Markov chain with Markov matrix $\mathbf{T} \in \mathbb{R}_\Delta^{s\times s}$, i.e., for any $t \in \mathbb{N}$, any $i, j \in [s]$, $\mathbb{P}(\omega_{t+1} = j \mid \omega_t = i) = \mathbf{T}(i,j)$. The mode switching sequence $\{\omega_t\}_{t\in\mathbb{N}}$ and the noise $\{\mathbf{w}_t\}_{t\in\mathbb{N}}$ are mutually independent. We assume the state $\mathbf{x}_t$ and mode ω_t can be observed at time t.

For the ergodic Markov matrix $\mathbf{T}$ underlying the MJS Σ, as in Section 2.2, we let $\boldsymbol{\pi}_t$, $\boldsymbol{\pi}$, and t_{MC} denote its transient distribution, stationary distribution, and mixing time; let $\pi_{\max} := \max_i \boldsymbol{\pi}(i)$ and $\pi_{\min} := \min_i \boldsymbol{\pi}(i)$.

In the remaining dissertation, we use the shorthand notation $\mathrm{MJS}(\mathbf{A}'_{1:s}, \mathbf{B}'_{1:s}, \mathbf{T}')$ to parameterize an arbitrary MJS in the form of (2.6) with mode dynamics $\{\mathbf{A}'_i, \mathbf{B}'_i\}_{i=1}^s$ and Markov matrix $\mathbf{T}'$. Particularly, we use $\Sigma := \mathrm{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ to denote the target MJS that is under study, e.g., the ground truth MJS that we want to identify, control, or perform model reduction to.

In terms of the controller design for MJSs, a typical choice is to assign each mode i a linear state-feedback controller $\mathbf{K}_i \in \mathbb{R}^{p \times n}$ such that the input is given by $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t$. We refer to the controller ensemble $\mathbf{K}_{1:s}$ as a mode-dependent controller for MJS. We say an MJS is autonomous if $\mathbf{u}_t = 0$ for all t and noise-free if $\mathbf{w}_t = 0$ for all t, or simply $\sigma_{\mathbf{w}} = 0$.

Remark. *As a slight abuse of terminology, we have used "state" to refer to both the Markov chain state in Section 2.2 and the MJS state in this section. In the remaining dissertation, "state" mainly refers to the MJS state $\mathbf{x}_t$. When it occasionally refers to the Markov chain state, it should be clear from the context.*

2.3.1 Stability

In this section, we introduce two types of stability commonly considered for MJSs: mean-square stability and uniform stability.

2.3.1.1 Mean-Square Stability

Even in the autonomous and noise-free case, the state $\{\mathbf{x}_t\}_{t\in\mathbb{N}}$ of an MJS is a random process due to the inherent randomness induced by the Markovian mode switching sequence $\{\omega_t\}_{t\in\mathbb{N}}$. Because of this, when we study the stability of an MJS, such randomness is also factored in.

Definition 2.8 (Mean-Square Stability (Costa et al., 2006, Definition 3.8))**.** *We say the MJS Σ in (2.6) is mean-square stable if, when setting $\mathbf{u}_t = 0$ for all t, there exists $\mathbf{x}_\infty \in \mathbb{R}^n, \mathbf{\Sigma}_\infty \in \mathbb{S}_+^n$ such that for any initial state $\mathbf{x}_0$ and mode ω_0, as $t \to \infty$,*

$$\|\mathbb{E}[\mathbf{x}_t] - \mathbf{x}_\infty\| \to 0, \quad \|\mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal] - \mathbf{\Sigma}_\infty\| \to 0. \tag{2.7}$$

In the above definition, the expectation is taken with respect to the mode switching sequence $\{\omega_t\}_{t\in\mathbb{N}}$ and process noise $\{\mathbf{w}_t\}_{t\in\mathbb{N}}$. In the noise-free case, condition (2.7) is equivalent to $\|\mathbb{E}[\mathbf{x}_t]\| \to 0, \|\mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal]\| \to 0$, and the convergence rate is exponential (Costa et al., 2006, Theorem 3.9). Mean-square stability only requires the convergence of the state $\mathbf{x}_t$ in the expectation sense, thus explosive state trajectory realizations may still occur with certain probability. Furthermore, an MJS being mean-square stable does not imply each individual mode, when treated as an LTI system, is Lyapunov stable. And every mode being Lyapunov stable does not imply the mean-square stability either. Below is an example for the former claim.

Example 2.9. *Consider the following scalar two-mode autonomous MJS.*

$$\textit{Mode 1}: x_{t+1} = 1.1x_t, \quad \textit{Mode 2}: x_{t+1} = 0.7x_t, \quad x_0 = 1, \quad \mathbf{T} = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix}. \tag{2.8}$$

It is obvious that, individually Mode 1 and Mode 2 are unstable and stable respectively. Fig. 2.1 shows the state trajectories under different mode switching sequences. We see even though the state x_t is explosive when Mode 1 is active all the time, but the state average is decaying over time, which implies the mean-square stability.

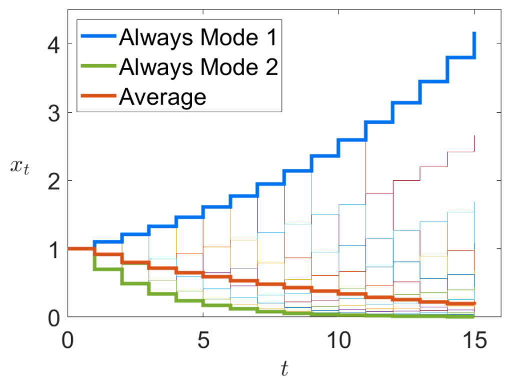


Figure 2.1: Mean-Square stability demonstration. Thick blue and green curves: mode switching sequences $\Omega_1 = \{1, 1, \dots\}$ and $\Omega_2 = \{2, 2, \dots\}$. Thick orange curve: average over all realizations. Thin curves: possible trajectory realizations.

Though the mean-square stability in Definition 2.8 is defined for open-loop autonomous MJSs due to $\mathbf{u}_t = 0$, it also applies to closed-loop MJSs under the mode-dependent state-feedback controller $\mathbf{K}_{1:s}$ since the closed-loop dynamics $\mathbf{x}_{t+1} = (\mathbf{A}_{\omega_t} + \mathbf{B}_{\omega_t}\mathbf{K}_{\omega_t})\mathbf{x}_t + \mathbf{w}_t$ is an autonomous MJS. This brings the notion of stabilizability.

Definition 2.10 (Mean-Square Stabilizability)**.** *Given a mode-dependent controller $\mathbf{K}_{1:s}$ for the MJS Σ in (2.6), let $\mathbf{L}_{1:s}$ denote the closed-loop state matrices such that $\mathbf{L}_i = \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i$. We say Σ is (mean-square) stabilizable if there exists a mode-dependent controller $\mathbf{K}_{1:s}$ such that the closed-loop MJS, i.e.,* $\mathrm{MJS}(\mathbf{L}_{1:s}, 0, \mathbf{T})$*, is mean-square stable. And we say such a controller $\mathbf{K}_{1:s}$ stabilizes Σ or is a stabilizing controller for Σ.*

The stabilizability of an MJS can be verified by solving for linear matrix inequalities (Costa et al., 2006, Proposition 3.42). The verification of mean-square stability, on the other hand, is much easier. It is well-known that the stability of a non-switched LTI system is related to the spectral radius of the state matrix. Similarly, mean-square stability of the MJS in (2.6) relates to the spectral radius of the following block matrix.

Definition 2.11 (Augmented State Matrix for MJSs)**.** *For the the MJS* Σ *in* (2.6)*, we define the augmented state matrix* $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{sn^2 \times sn^2}$ *such that*

$$\boldsymbol{\mathcal{A}} := \begin{bmatrix} \mathbf{T}(1,1)\cdot(\mathbf{A}_1 \otimes \mathbf{A}_1) & \mathbf{T}(2,1)\cdot(\mathbf{A}_2 \otimes \mathbf{A}_2) & \cdots & \mathbf{T}(s,1)\cdot(\mathbf{A}_s \otimes \mathbf{A}_s) \\ \mathbf{T}(1,2)\cdot(\mathbf{A}_1 \otimes \mathbf{A}_1) & \mathbf{T}(2,2)\cdot(\mathbf{A}_2 \otimes \mathbf{A}_2) & \cdots & \mathbf{T}(s,2)\cdot(\mathbf{A}_s \otimes \mathbf{A}_s) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{T}(1,s)\cdot(\mathbf{A}_1 \otimes \mathbf{A}_1) & \mathbf{T}(2,s)\cdot(\mathbf{A}_2 \otimes \mathbf{A}_2) & \cdots & \mathbf{T}(s,s)\cdot(\mathbf{A}_s \otimes \mathbf{A}_s) \end{bmatrix} \tag{2.9}$$

For a mode-dependent controller $\mathbf{K}_{1:s}$ *and its corresponding closed-loop state matrices* $\mathbf{L}_{1:s}$ *such that* $\mathbf{L}_i = \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i$*, we similarly define the augmented closed-loop state matrix* $\boldsymbol{\mathcal{L}} \in \mathbb{R}^{sn^2 \times sn^2}$ *such that*

$$\boldsymbol{\mathcal{L}} := \begin{bmatrix} \mathbf{T}(1,1)\cdot(\mathbf{L}_1 \otimes \mathbf{L}_1) & \mathbf{T}(2,1)\cdot(\mathbf{L}_2 \otimes \mathbf{L}_2) & \cdots & \mathbf{T}(s,1)\cdot(\mathbf{L}_s \otimes \mathbf{L}_s) \\ \mathbf{T}(1,2)\cdot(\mathbf{L}_1 \otimes \mathbf{L}_1) & \mathbf{T}(2,2)\cdot(\mathbf{L}_2 \otimes \mathbf{L}_2) & \cdots & \mathbf{T}(s,2)\cdot(\mathbf{L}_s \otimes \mathbf{L}_s) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{T}(1,s)\cdot(\mathbf{L}_1 \otimes \mathbf{L}_1) & \mathbf{T}(2,s)\cdot(\mathbf{L}_2 \otimes \mathbf{L}_2) & \cdots & \mathbf{T}(s,s)\cdot(\mathbf{L}_s \otimes \mathbf{L}_s) \end{bmatrix} \tag{2.10}$$

Then, we have the following results regarding mean-square stability.

Lemma 2.12 ((Costa et al., 2006, Theorem 3.9))**.** *For the the MJS* Σ *in* (2.6) *and the augmented state matrix* $\boldsymbol{\mathcal{A}}$ *in* (2.9)*, the following are equivalent.*

(a) Σ *is mean-square stable.*

(b) $\rho(\boldsymbol{\mathcal{A}}) < 1$.

(c) There exists $\mathbf{V}_{1:s} \in \mathbb{S}^n_{s,++}$ *such that for all* $i \in [s]$,

$$\mathbf{V}_i - \mathbf{A}_i^\intercal\Big(\sum_{j=1}^{s} \mathbf{T}(i,j)\mathbf{V}_j\Big)\mathbf{A}_i \succ 0. \tag{2.11}$$

It is easy to see that when $s = 1$, these assertions reduce to the Lyapunov stability criteria for LTI systems. From Lemma 2.12, we can trivially obtain the following corollary regarding whether a given controller stabilizes the MJS.

Corollary 2.13. *For the the MJS* Σ *in* (2.6)*, a mode-dependent controller* $\mathbf{K}_{1:s}$*, the corresponding closed-loop state matrices* $\mathbf{L}_{1:s}$*, and the augmented closed-loop state matrix* $\boldsymbol{\mathcal{L}}$ *in* (2.10)*, the following are equivalent.*

(a) $\mathbf{K}_{1:s}$ *is a stabilizing controller for* Σ.

(b) $\rho(\boldsymbol{\mathcal{L}}) < 1$.

(c) There exists $\mathbf{V}_{1:s} \in \mathbb{S}^n_{s,++}$ *such that for all* $i \in [s]$,

$$\mathbf{V}_i - \mathbf{L}_i^\intercal \Big(\sum_{j=1}^{s} \mathbf{T}(i,j)\mathbf{V}_j \Big) \mathbf{L}_i \succ 0. \tag{2.12}$$

The following Lemma 2.14 regarding the evolution of the state covariance $\mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal]$ tells how $\rho(\boldsymbol{\mathcal{A}})$ or $\rho(\boldsymbol{\mathcal{L}})$ would determine the mean-square stability. In Lemma 2.14, we consider a more general case where the input $\mathbf{u}_t$ is given the state-feedback term $\mathbf{K}_{\omega_t}\mathbf{x}_t$ plus an additional random noise term $\mathbf{z}_t$, which is also known as the random excitation. This applies to cases when the input signal is sent to the system through some noisy channel or when one wants to fully excite the system for better identification performance. The latter will be covered in later chapters when the identification and adaptive control problems are studied.

Lemma 2.14 (State Covariance Dynamics)**.** *Consider the MJS* $\Sigma = \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *in* (2.6) *except for more general covariance* $\boldsymbol{\Sigma}_\mathbf{w}$ *for* $\mathbf{w}_t$. *Let* $\boldsymbol{\pi}_t \in \mathbb{R}^s_\Delta$ *denote the transient distribution of the underlying Markov chain for the modes. Given a controller* $\mathbf{K}_{1:s}$, *suppose the input is given by* $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$ *where* $\{\mathbf{z}_t\}_{t\in\mathbb{N}} \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, \boldsymbol{\Sigma}_\mathbf{z})$, *and* $\{\mathbf{z}_t\}_{t\in\mathbb{N}}$ *is independent of* $\{\mathbf{w}_t\}_{t\in\mathbb{N}}$ *and* $\{\omega_t\}_{t\in\mathbb{N}}$. *Let* $\mathbf{L}_{1:s}$ *and* $\boldsymbol{\mathcal{L}}$ *respectively denote the individual and augmented closed-loop state matrices. Define*

$$\begin{gathered}\boldsymbol{\Sigma}_{t,i} := \mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal \mathbf{1}_{\{\omega_t = i\}}], \quad \boldsymbol{\Sigma}_t := \mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal], \quad \boldsymbol{\Pi}_t := \boldsymbol{\pi}_t \otimes \mathbf{I}_{n^2}, \\ \mathbf{s}_t := \begin{bmatrix} \mathbf{vec}(\boldsymbol{\Sigma}_{t,1}) \\ \vdots \\ \mathbf{vec}(\boldsymbol{\Sigma}_{t,s}) \end{bmatrix}, \quad \boldsymbol{\mathcal{B}}_t := \begin{bmatrix} \sum_{j=1}^s \boldsymbol{\pi}_{t-1}(j)\mathbf{T}(j,1)(\mathbf{B}_j \otimes \mathbf{B}_j) \\ \vdots \\ \sum_{j=1}^s \boldsymbol{\pi}_{t-1}(j)\mathbf{T}(j,s)(\mathbf{B}_j \otimes \mathbf{B}_j). \end{bmatrix}\end{gathered} \tag{2.13}$$

Then, $\mathbf{s}_t$ *evolves as follows.*

$$\mathbf{s}_t = \boldsymbol{\mathcal{L}}\mathbf{s}_0 + \boldsymbol{\mathcal{B}}_t \mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{z}) + \boldsymbol{\Pi}_t \mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{w}). \tag{2.14}$$

Propagating this backward gives

$$\begin{aligned}\mathbf{s}_t = \boldsymbol{\mathcal{L}}^t\mathbf{s}_0 &+ (\boldsymbol{\mathcal{B}}_t + \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{B}}_{t-1} + \cdots + \boldsymbol{\mathcal{L}}^{t-1}\boldsymbol{\mathcal{B}}_1)\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{z}) \\ &+ (\boldsymbol{\Pi}_t + \boldsymbol{\mathcal{L}}\boldsymbol{\Pi}_{t-1} + \cdots + \boldsymbol{\mathcal{L}}^{t-1}\boldsymbol{\Pi}_1)\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{w}).\end{aligned} \tag{2.15}$$

Proof. Note that the closed-loop MJS dynamics is given by $\mathbf{x}_{t+1} = \mathbf{L}_{\omega_t}\mathbf{x}_t + \mathbf{B}_{\omega_t}\mathbf{z}_t + \mathbf{w}_t$.

Taking outer-products of both sides followed by expectations gives

$$\begin{aligned}\mathbb{E}[\mathbf{x}_{t+1}\mathbf{x}_{t+1}^\intercal \mathbf{1}_{\{\omega_{t+1}=i\}}] = &\sum_{j=1}^{s} \mathbb{E}[\mathbf{L}_j \mathbf{x}_t \mathbf{x}_t^\intercal \mathbf{L}_j \mathbf{1}_{\{\omega_{t+1}=i,\omega_t=j\}}] \\ &+ \sum_{j=1}^{s} \mathbb{E}[\mathbf{B}_j \mathbf{z}_t \mathbf{z}_t^\intercal \mathbf{B}_j^\intercal \mathbf{1}_{\{\omega_{t+1}=i,\omega_t=j\}}] + \mathbb{E}[\mathbf{w}_t \mathbf{w}_t^\intercal \mathbf{1}_{\{\omega_{t+1}=i\}}],\end{aligned} \tag{2.16}$$

where the mutual independence between $\{\mathbf{z}_t\}_{t\in\mathbb{N}}$, $\{\mathbf{w}_t\}_{t\in\mathbb{N}}$, and $\{\omega_t\}_{t\in\mathbb{N}}$. This equation can be further written as

$$\mathbf{\Sigma}_{t+1,i} = \sum_{j=1}^{s} \mathbf{T}(j,i)\mathbf{L}_j \mathbf{\Sigma}_{t,j} \mathbf{L}_j^\intercal + \sum_{j=1}^{s} \boldsymbol{\pi}_t(j)\mathbf{T}(j,i)\mathbf{B}_j \mathbf{\Sigma}_{\mathbf{z}} \mathbf{B}_j^\intercal + \boldsymbol{\pi}_{t+1}(i)\mathbf{\Sigma}_{\mathbf{w}}. \tag{2.17}$$

Vectorizing both sides of the above equation, we have

$$\begin{aligned}\mathbf{vec}(\mathbf{\Sigma}_{t+1,i}) = &\sum_{j=1}^{s} \mathbf{T}(j,i)(\mathbf{L}_j \otimes \mathbf{L}_j)\mathbf{vec}(\mathbf{\Sigma}_{t,j}) \\ &+ \sum_{j=1}^{s} \boldsymbol{\pi}_t(j)\mathbf{T}(j,i)(\mathbf{B}_j \otimes \mathbf{B}_j)\mathbf{vec}(\mathbf{\Sigma}_{\mathbf{z}}) + \boldsymbol{\pi}_{t+1}(i)\mathbf{vec}(\mathbf{\Sigma}_{\mathbf{w}}).\end{aligned}$$

Stacking this for all $i \in [s]$, we obtain

$$\begin{bmatrix}\mathbf{vec}(\mathbf{\Sigma}_{1,t+1}) \\ \vdots \\ \mathbf{vec}(\mathbf{\Sigma}_{s,t+1}))\end{bmatrix} = \boldsymbol{\mathcal{L}} \begin{bmatrix}\mathbf{vec}(\mathbf{\Sigma}_{1,t}) \\ \vdots \\ \mathbf{vec}(\mathbf{\Sigma}_{s,t})\end{bmatrix} + \boldsymbol{\mathcal{B}}_{t+1}\mathbf{vec}(\mathbf{\Sigma}_{\mathbf{z}}) + \mathbf{\Pi}_{t+1}\mathbf{vec}(\mathbf{\Sigma}_{\mathbf{w}}), \tag{2.18}$$

which concludes the proof. □

Note that one can obtain the state covariance $\mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal]$ from $\mathbf{s}_t$ simply through a linear mapping. Hence, Lemma 2.14 describes the dynamics of the state covariance $\mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal]$. When $\mathbf{\Sigma}_{\mathbf{z}} = 0$ and $\mathbf{\Sigma}_{\mathbf{w}} = 0$, it is the augmented matrix $\boldsymbol{\mathcal{L}}$ only that governs the evolution and thus convergence of $\mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal]$. From this, we can see why the spectral radii $\rho(\boldsymbol{\mathcal{A}})$ and $\rho(\boldsymbol{\mathcal{L}})$ would determine the mean-square stability in Lemma 2.12.

From the definition of mean-square stability in Definition 2.8, we know mean-square stability implies that $\mathbb{E}[\mathbf{x}_t]$ and $\mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal]$ both converges; and from Lemma 2.14, we can speculate their asymptotic convergence rates would be $\sqrt{\rho(\mathbf{L})}$ (or $\sqrt{\rho(\mathbf{A})}$) and $\rho(\mathbf{L})$ (or $\rho(\mathbf{A})$). On the other hand, their non-asymptotic convergence behaviors are of interest to finite-time or finite-sample analysis. This is studied in the next lemma.

Lemma 2.15 (State and Covariance Bounds for MJS)**.** *Consider the setup in Lemma 2.14. Assume the controller* $\mathbf{K}_{1:s}$ *stabilizes* Σ*. For any* $\rho \in [\rho(\boldsymbol{\mathcal{L}}, 1)$*, let* $\tau := \tau(\boldsymbol{\mathcal{L}}, \rho)$ *denote the power supremum of* $\boldsymbol{\mathcal{L}}$ *normalized by* ρ *as in Definition 2.1. Let* $\bar{\sigma} := \sqrt{\|\mathbf{B}_{1:s}\|^2\|\boldsymbol{\Sigma}_{\mathbf{z}}\| + \|\boldsymbol{\Sigma}_{\mathbf{w}}\|}$*. For* $\mathbb{E}[\|\mathbf{x}_t\|^2]$ *and* $\|\boldsymbol{\Sigma}_t\|_{\mathrm{F}}$*, we have*

$$\mathbb{E}[\|\mathbf{x}_t\|^2] \leq \sqrt{ns}\tau\rho^t\mathbb{E}[\|\mathbf{x}_0\|^2] + \frac{n\sqrt{s}\tau\bar{\sigma}^2}{1-\rho}, \tag{2.19}$$

$$\|\boldsymbol{\Sigma}_t\|_{\mathrm{F}} \leq \sqrt{s}\tau\rho^t\mathbb{E}[\|\mathbf{x}_0\|^2] + \frac{\sqrt{ns}\tau\bar{\sigma}^2}{1-\rho}. \tag{2.20}$$

Proof. First we derive the upper bound for $\mathbb{E}[\|\mathbf{x}_t\|^2]$. The upper bound for $\|\boldsymbol{\Sigma}_t\|_{\mathrm{F}}$ follows similarly. For state $\mathbf{x}_t$, we have

$$\begin{aligned}
\mathbb{E}[\|\mathbf{x}_t\|^2] = \sum_{i=1}^{s}\mathbb{E}[\|\mathbf{x}_t\|^2\mathbf{1}_{\{\omega_t=i\}}] &= \sum_{i=1}^{s}\mathbf{tr}\left(\mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal\mathbf{1}_{\{\omega_t=i\}}]\right) = \sum_{i=1}^{s}\mathbf{tr}(\boldsymbol{\Sigma}_{t,i}) \\
&= \sum_{i=1}^{s}\sum_{j=1}^{n}\lambda_j(\boldsymbol{\Sigma}_{t,i}) \leq \sqrt{ns}\sqrt{\sum_{i=1}^{s}\sum_{j=1}^{n}\lambda_j^2(\boldsymbol{\Sigma}_{t,i})} \\
&\leq \sqrt{ns}\sqrt{\sum_{i=1}^{s}\|\boldsymbol{\Sigma}_{t,i}\|_{\mathrm{F}}^2}.
\end{aligned}$$

From the definition of $\mathbf{s}_t$ in (2.13), we see $\|\sqrt{\sum_{i=1}^{s}\|\boldsymbol{\Sigma}_{t,i}\|_{\mathrm{F}}^2}\| = \|\mathbf{s}_t\|$. This gives

$$\mathbb{E}[\|\mathbf{x}_t\|^2] \leq \sqrt{ns}\|\mathbf{s}_t\|. \tag{2.21}$$

Now, plugging in the expression for $\mathbf{s}_t$ in (2.14) of Lemma 2.14 yields

$$\begin{aligned}
\mathbb{E}[\|\mathbf{x}_t\|^2] &\leq \sqrt{ns}\Big(\|\boldsymbol{\mathcal{L}}^t\|\|\mathbf{s}_0\| + \sum_{t'=1}^{t}\|\boldsymbol{\mathcal{L}}^{t-t'}\|\|\boldsymbol{\mathcal{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\| + \sum_{t'=1}^{t}\|\boldsymbol{\mathcal{L}}^{t-t'}\|\|\boldsymbol{\Pi}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\|\Big) \\
&\leq \sqrt{ns}\tau\Big(\rho^t\|\mathbf{s}_0\| + \sum_{t'=1}^{t}\rho^{t-t'}\|\boldsymbol{\mathcal{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\| + \sum_{t'=1}^{t}\rho^{t-t'}\|\boldsymbol{\Pi}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\|\Big),
\end{aligned} \tag{2.22}$$

where the second line follows from $\|\boldsymbol{\mathcal{L}}^t\| \leq \tau\rho^t$ in Lemma 2.2.

Now, we evaluate the terms $\|\mathbf{s}_0\|$, $\|\boldsymbol{\mathcal{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\|$, and $\|\boldsymbol{\Pi}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\|$ in (2.22) separately. For $\|\mathbf{s}_0\|$, following from $\|\mathbf{s}_0\| = \sqrt{\sum_{i=1}^{s}\|\boldsymbol{\Sigma}_{0,i}\|_{\mathrm{F}}^2}$, we have

$$\|\mathbf{s}_0\| = \sqrt{\sum_{i=1}^{s}\boldsymbol{\pi}_i(0)^2\|\mathbb{E}[\mathbf{x}_0\mathbf{x}_0^\intercal]\|_{\mathrm{F}}^2} \leq \|\mathbb{E}[\mathbf{x}_0\mathbf{x}_0^\intercal]\|_{\mathrm{F}} \leq \mathbb{E}[\|\mathbf{x}_0\mathbf{x}_0^\intercal\|_{\mathrm{F}}] = \mathbb{E}[\|\mathbf{x}_0\|^2]. \tag{2.23}$$

Let $[\boldsymbol{\mathcal{B}}_{t'}]_i := \sum_{j=1}^{s} \boldsymbol{\pi}_{t-1}(j)\mathbf{T}(j,i)(\mathbf{B}_j \otimes \mathbf{B}_j)$, i.e., the i-th block of $\boldsymbol{\mathcal{B}}_{t'}$ in (2.13). Then, following from $\|\boldsymbol{\mathcal{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\| = \sqrt{\sum_{i=1}^{s} \|[\boldsymbol{\mathcal{B}}_{t'}]_i\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\|^2}$, we have

$$\begin{aligned}
\|\boldsymbol{\mathcal{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\| &\leq \sum_{i=1}^{s} \|[\boldsymbol{\mathcal{B}}_{t'}]_i\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\| \\
&= \sum_{i=1}^{s} \|\sum_{j=1}^{s} \boldsymbol{\pi}_{t'-1}(j)\mathbf{T}(j,i)(\mathbf{B}_j \otimes \mathbf{B}_j)\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\| \\
&= \sum_{i=1}^{s} \|\sum_{j=1}^{s} \boldsymbol{\pi}_{t'-1}(j)\mathbf{T}(j,i)(\mathbf{B}_j\boldsymbol{\Sigma}_{\mathbf{z}}\mathbf{B}_j^\intercal)\|_{\mathrm{F}} \\
&\leq \|\mathbf{B}_{1:s}\|^2\|\boldsymbol{\Sigma}_{\mathbf{z}}\| \cdot \sum_{i=1}^{s} \|\sum_{j=1}^{s} \boldsymbol{\pi}_{t'-1}(j)\mathbf{T}(j,i)\mathbf{I}_n\|_{\mathrm{F}} \\
&= \|\mathbf{B}_{1:s}\|^2\|\boldsymbol{\Sigma}_{\mathbf{z}}\| \cdot \sum_{i=1}^{s} \|\boldsymbol{\pi}_{t'}(i)\mathbf{I}_n\|_{\mathrm{F}} \\
&\leq \sqrt{n}\|\mathbf{B}_{1:s}\|^2\|\boldsymbol{\Sigma}_{\mathbf{z}}\|.
\end{aligned} \tag{2.24}$$

Lastly, we have

$$\|\tilde{\boldsymbol{\Pi}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\| = \sqrt{\sum_{i=1}^{s} \|\boldsymbol{\pi}_{t'}(i)\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\|^2} \leq \|\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\| = \|\boldsymbol{\Sigma}_{\mathbf{w}}\|_{\mathrm{F}} = \sqrt{n}\|\boldsymbol{\Sigma}_{\mathbf{w}}\|. \tag{2.25}$$

Plugging (2.23)–(2.25) into (2.22), we obtain

$$\begin{aligned}
\mathbb{E}[\|\mathbf{x}_t\|^2] &\leq \sqrt{ns}\tau\left(\rho^t\mathbb{E}[\|\mathbf{x}_0\|^2] + \sqrt{n}\|\mathbf{B}_{1:s}\|^2\|\boldsymbol{\Sigma}_{\mathbf{z}}\|\sum_{t'=1}^{t}\rho^{t-t'} + \sqrt{n}\|\boldsymbol{\Sigma}_{\mathbf{w}}\|\sum_{t'=1}^{t}\rho^{t-t'}\right), \\
&\leq \sqrt{ns}\cdot\tau\rho^t\cdot\mathbb{E}[\|\mathbf{x}_0\|^2] + n\sqrt{s}(\|\mathbf{B}_{1:s}\|^2\|\boldsymbol{\Sigma}_{\mathbf{z}}\| + \|\boldsymbol{\Sigma}_{\mathbf{w}}\|)\frac{\tau}{1-\rho},
\end{aligned} \tag{2.26}$$

which gives the bound for $\mathbb{E}[\|\mathbf{x}_t\|^2]$ in (2.19).

To obtain the bound for $\|\boldsymbol{\Sigma}_t\|_{\mathrm{F}}$ in (2.20), first notice that $\|\boldsymbol{\Sigma}_t\|_{\mathrm{F}} = \|\sum_{i=1}^{s}\boldsymbol{\Sigma}_{t,i}\|_{\mathrm{F}} \leq \sqrt{s}\sqrt{\sum_{i=1}^{s}\|\boldsymbol{\Sigma}_{t,i}\|_{\mathrm{F}}^2} \leq \sqrt{s}\|\mathbf{s}_t\|$, then we can use similar steps to prove the claim. □

2.3.1.2 Uniform Stability

As a stability notion in the weak sense, mean-square stability does not prevent certain mode switching sequence from yielding explosive state realizations. There is a stronger type of stability, namely uniform stability (Liberzon, 2003; Lee and Dullerud, 2006) that guarantees the state convergence under an arbitrary switching sequence.

Definition 2.16 (Uniform Stability)**.** *We say the MJS* Σ *in* (2.6) *is uniformly stable if the joint spectral radius* $\xi(\mathbf{A}_{1:s}) < 1$.

By definition of the spectral radius and the sub-multiplicative property of matrix norms, it is easy to see a sufficient condition for uniform stability is $\|\mathbf{A}_i\| < 1$ for all i. To quantify the convergence rate of a uniformly stable MJS, similar to the normalized power supremum in Definition 2.1, we define the following.

Definition 2.17 (Normalized Joint Power Supremum)**.** *Consider a sequence of matrices* $\mathbf{A}_{1:s} \in \mathbb{R}_s^{n \times n}$. *Given a free parameter* ξ *such that* $\xi \geq \xi(\mathbf{A}_{1:s})$, *define*

$$\kappa(\mathbf{A}_{1:s}, \xi) := \sup_{k \in \mathbb{N}} \max_{\omega_{1:k} \in [s]^k} \|\mathbf{A}_{\omega_1} \cdots \mathbf{A}_{\omega_k}\| / \xi^k. \tag{2.27}$$

We call $\kappa(\mathbf{A}_{1:s}, \xi)$ *the joint power supremum of matrices* $\mathbf{A}_{1:s}$ *normalized by* ξ.

By definition of the joint spectral radius, we see $\kappa(\mathbf{A}_{1:s}, \xi)$ is finite for any $\xi > \xi(\mathbf{A}_{1:s})$. Note that the pair $\{\xi, \kappa(\mathbf{A}_{1:s}, \xi)\}$ for a sequence of matrices $\mathbf{A}_{1:s}$ is just the counterpart of $\{\rho, \tau(\mathbf{A}, \rho)\}$ defined for a single matrix $\mathbf{A}$. Similar to Lemma 2.2, we have the following properties.

Lemma 2.18 (Properties of the Normalized Joint Power Supremum)**.** *For the pair* $\{\xi, \kappa(\mathbf{A}_{1:s}, \xi)\}$ *in Definition 2.1, we have the following.*

(a) $\kappa(\mathbf{A}_{1:s}, \xi) \geq 1$.

(b) For all $k \in \mathbb{N}$ *and all* $\omega_{1:k} \in [s]^k$, $\|\mathbf{A}_{\omega_1} \cdots \mathbf{A}_{\omega_k}\| \leq \kappa(\mathbf{A}_{1:s}, \xi)\xi^k$.

(c) If $\xi(\mathbf{A}_{1:s}) < 1$ *and* $\xi \in [\xi(\mathbf{A}_{1:s}), 1)$, $\sum_{k=0}^{\infty} \max_{\omega_{1:k} \in [s]^k} \|\mathbf{A}_{\omega_1} \cdots \mathbf{A}_{\omega_k}\| \leq \frac{\kappa(\mathbf{A}_{1:s}, \xi)}{1-\xi}$.

2.3.2 Linear Quadratic Regulator

In this dissertation, the control problem we mainly consider for MJS is the linear quadratic regulator (LQR) problem. Given positive semi-definite cost matrices $\mathbf{Q}_{1:s} \in \mathbb{S}_{s,+}^n$ and $\mathbf{R}_{1:s} \in \mathbb{S}_{s,+}^p$, we define the following cost function with respect to the tuple $\{\mathbf{x}_t, \mathbf{u}_t, \omega_t\}$ from the MJS Σ in (2.6).

$$J_T = \mathbb{E}\left[\sum_{t=0}^{T-1} \left(\mathbf{x}_t^\intercal \mathbf{Q}_{\omega_t} \mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega_t} \mathbf{u}_t\right) + \mathbf{x}_T^\intercal \mathbf{Q}_{\omega_t} \mathbf{x}_T\right]. \tag{2.28}$$

Matrices $\mathbf{Q}_{\omega_t}$ and $\mathbf{R}_{\omega_t}$ are mode-dependent cost matrices chosen by users. The expectation is taken over the randomness of the initial state $\mathbf{x}_0$, noise $\{\mathbf{w}_t\}_{t \in \mathbb{N}}$ and Markovian modes $\{\omega_t\}_{t \in \mathbb{N}}$. The quadratic cost $\mathbf{x}_t^\intercal \mathbf{Q}_{\omega_t} \mathbf{x}_t$ usually represents the deviation of states, e.g. velocity, position, angle, from the target value, whereas $\mathbf{u}_t^\intercal \mathbf{R}_{\omega_t} \mathbf{u}_t$ represents the control effort, e.g. battery/fuel consumption. Unlike classical LQR for LTI systems, where cost matrices are usually fixed throughout the time horizon, the mode-dependent cost matrices in MJS-LQR allows us to have different control goals under different modes. The goal is to design the input $\mathbf{u}_t$ such that the cost function is minimized.

$$\begin{aligned} \min_{\mathbf{u}_{0:T-1}} \quad & J_T \\ \text{s.t.} \quad & \mathbf{x}_t, \omega_t \sim \Sigma \\ & \mathbf{x}_t, \omega_t \text{ observed at time } t. \end{aligned} \tag{2.29}$$

The constraint "$\mathbf{x}_t, \omega_t \sim \Sigma$" denote that $\{\mathbf{x}_t, \omega_t\}_{t=0}^T$ is generated by the MJS Σ under the input $\{\mathbf{u}_t\}_{t=0}^\intercal$. This problem is known as the finite-horizon MJS-LQR problem. It has seen many real world applications, including networked control with random packet losses (Vargas et al., 2010) or delays (Chan and Ozguner, 1995), single-link robot arm with time-varying payloads and inertia (Palm and Driankov, 1998; Wu and Cai, 2006; Zhong et al., 2014), optimal control for a solar thermal receiver (Costa et al., 2006), and public expenditure policy-making (Costa et al., 2006).

To ease the discussion of its optimal solution, we define the following operators.

Definition 2.19 (Operators Related to MJS-LQR)**.** *Consider the MJS* $\Sigma = \mathrm{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *and LQR cost matrices* $\mathbf{Q}_{1:s}$ *and* $\mathbf{R}_{1:s}$*. For a sequence of positive semi-definite matrices* $\mathbf{X}_{1:s} \in \mathbb{S}_{s,+}^n$*, define the following operators: for all* $i \in [s]$

$$\varphi_i(\mathbf{X}_{1:s}) := \sum_{j \in [s]} \mathbf{T}(i,j) \mathbf{X}_j, \tag{2.30}$$

$$\begin{aligned} \mathcal{R}_i(\mathbf{X}_{1:s}) := & \, \mathbf{Q}_i + \mathbf{A}_i^\intercal \varphi_i(\mathbf{X}_{1:s}) \mathbf{A}_i \\ & - \mathbf{A}_i^\intercal \varphi_i(\mathbf{X}_{1:s})^\intercal \mathbf{B}_i \left(\mathbf{R}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{X}_{1:s}) \mathbf{B}_i\right)^{-1} \mathbf{B}_i^\intercal \varphi_i(\mathbf{X}_{1:s}) \mathbf{A}_i, \end{aligned} \tag{2.31}$$

$$\mathcal{K}_i(\mathbf{X}_{1:s}) := -\left(\mathbf{R}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{X}_{1:s}) \mathbf{B}_i\right)^{-1} \left(\mathbf{B}_i^\intercal \varphi_i(\mathbf{X}_{1:s}) \mathbf{A}_i\right). \tag{2.32}$$

Then, the optimal solution to the finite-horizon MJS-LQR problem (2.29) is given by the following, which can be obtained by dynamic programming.

Lemma 2.20 (Optimal Solution to Finite-Horizon MJS-LQR (Costa et al., 2006, Theorem 4.2))**.** *Consider the finite-horizon MJS-LQR problem* (2.29) *with MJS* $\Sigma = \mathrm{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *and cost matrices* $\mathbf{Q}_{1:s}$, $\mathbf{R}_{1:s}$*. Its optimal solution is given by* $\mathbf{u}_t = \mathbf{K}_{t,\omega_t}^\star \mathbf{x}_t$ *for all* t *where* $\{\mathbf{K}_{t,1:s}^\star\}_{t=0}^{T-1}$ *is the optimal time-varying mode-dependent state-feedback controller computed* *as follows.*

$$\begin{aligned} \mathbf{P}^\star_{T,i} &= \mathbf{Q}_i, && \forall\ i \in [s]; \\ \mathbf{P}^\star_{t,i} &= \mathcal{R}_i(\mathbf{P}^\star_{t+1,1:s}), && \forall\ t = 0, \dots, T-1,\ \forall\ i \in [s]; \\ \mathbf{K}^\star_{t,i} &= \mathcal{K}_i(\mathbf{P}^\star_{t+1,1:s}), && \forall\ t = 0, \dots, T-1,\ \forall\ i \in [s], \end{aligned} \tag{2.33}$$

where the operators $\mathcal{R}_{1:s}$ and $\mathcal{K}_{1:s}$ are as in Definition 2.19.

Other than picking a finite horizon, one can also formulate the LQR problem using the infinite horizon:

$$\begin{aligned} \inf_{\mathbf{u}_0, \mathbf{u}_1, \dots} \quad & \limsup_{T \to \infty} \frac{1}{T} J_T \\ \text{s.t.} \quad & \mathbf{x}_t, \omega_t \sim \Sigma, \\ & \mathbf{x}_t, \omega_t \text{ observed at time } t. \end{aligned} \tag{2.34}$$

Here, the time-averaged cost is considered since it is possible that the cumulative cost $J_T \to \infty$ as $T \to \infty$ due to the presence of process noise $\mathbf{w}_t$. Let $J^\star$ denote the optimal cost in (2.34). Its optimal solution is related to the solution of the following Riccati equations.

Definition 2.21 (Coupled Discrete-Time Algebraic Riccati Equations (cDARE) for MJS-LQR)**.** *Consider the MJS $\Sigma = \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ and LQR cost matrices $\mathbf{Q}_{1:s}$ and $\mathbf{R}_{1:s}$. We call the following set of equations the coupled discrete-time algebraic Riccati equations (cDARE)*

$$\begin{aligned} \mathbf{X}_1 &= \mathcal{R}_1(\mathbf{X}_{1:s}) \\ \mathbf{X}_2 &= \mathcal{R}_2(\mathbf{X}_{1:s}) \\ &\vdots \\ \mathbf{X}_s &= \mathcal{R}_s(\mathbf{X}_{1:s}) \end{aligned} \tag{2.35}$$

with respect to positive semi-definite matrices $\mathbf{X}_{1:s} \in \mathbb{S}^n_+$, where the operators $\mathcal{R}_{1:s}$ are as in Definition 2.19.

Throughout the dissertation, we use $\text{cDARE}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ to parameterize the cDARE with $\text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ and cost matrices $\mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$. In practice, cDARE can be solved efficiently either with LMIs or via value iteration (Costa et al., 2006). The following lemma discusses its solvability and the optimal solution to the infinite-horizon MJS-LQR problem.

Lemma 2.22 (Optimal Solution to Infinite-Horizon MJS-LQR (Costa et al., 2006, Theorem 4.6 and Corollary A.21))**.** *Consider the infinite-horizon MJS-LQR problem (2.34) and the cDARE (2.35) with MJS $\Sigma = \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ and cost matrices $\mathbf{Q}_{1:s}$, $\mathbf{R}_{1:s}$. Assume that*

(a) the cost matrices $\mathbf{R}_{1:s} \succ 0$;

(b) for all $i \in [s]$*, the pair* $(\mathbf{Q}_i^{\frac{1}{2}}, \mathbf{A}_i)$ *is observable;*

(c) the MJS Σ *is stabilizable.*

Then the cDARE has a unique solution $\mathbf{P}_{1:s}^\star$ *in* $\mathbb{S}_{s,+}^n$*, and the solution is positive definite, i.e.,* $\mathbf{P}_{1:s}^\star \succ 0$*. Consider the mode-dependent state-feedback controller* $\mathbf{K}_{1:s}^\star$ *such that*

$$\mathbf{K}_i^\star = \mathcal{K}_i(\mathbf{P}_{1:s}^\star), \ \forall i \in [s], \tag{2.36}$$

where the operators $\mathcal{K}_{1:s}$ *are as in Definition 2.19. Then* $\mathbf{K}_{1:s}^\star$ *stabilizes the MJS* Σ*. Assume additionally that*

(d) the Markov matrix $\mathbf{T}$ *is ergodic.*

Then, the optimal solution to the infinite-horizon MJS-LQR is given by $\mathbf{u}_t = \mathbf{K}_{\omega_t}^\star \mathbf{x}_t$ *for all* t*. The resulting optimal cost is given by* $J^\star = \sigma_{\mathbf{w}}^2 \mathbf{tr}(\sum_{i\in[s]} \boldsymbol{\pi}(i)\mathbf{P}_i)$ *where* $\boldsymbol{\pi}$ *is the stationary distribution of the Markov matrix* $\mathbf{T}$*.*

Throughout the dissertation, we use MJS-LQR($\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$) to parameterize an MJS-LQR problem with dynamics MJS($\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}$) and cost matrices $\mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$. Whether it is finite-horizon (2.29) or infinite-horizon (2.34) will be explicitly mentioned.

Chapter 3
System Identification

For an unknown system, using its data to obtain a mathematical model that characterizes its dynamics is known as system identification. This typically involves applying estimation techniques to find system parameters that best explain the given data. An accurate enough model is fundamental to downstream tasks such as analyzing system properties and designing controllers. In this chapter, we consider the system identification problem for MJSs. In Section 3.1, we introduce the problem setup and related work. Our identification approach is presented in Section 3.2, and its sample complexity guarantees are in Section 3.3. Section 3.5 provides experimental results that show the efficacy of our approach and support our theory.

3.1 Introduction and Related Work

Given an MJS, we study its identification problem with the following setup.

Problem P3.1 (System Identification for MJSs)**.** *Consider the MJS* $\Sigma = \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *in* (2.6) *with unknown state/input matrices* $\mathbf{A}_{1:s}, \mathbf{B}_{1:s}$ *and Markov matrix* $\mathbf{T}$*. The goal is to design the input* $\mathbf{u}_t$ *and then estimate* $\mathbf{A}_{1:s}, \mathbf{B}_{1:s}$ *and* $\mathbf{T}$ *from a single trajectory of states, inputs, and modes* $\{\mathbf{x}_t, \mathbf{u}_t, \omega_t\}_{t=0}^T$*. Denoting the estimates by* $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}$*, and* $\hat{\mathbf{T}}$*, the second goal is to analyze how the trajectory length* T *affects the estimation errors, i.e.,* $\|\hat{\mathbf{A}}_i - \mathbf{A}_i\|$*,* $\|\hat{\mathbf{B}}_i - \mathbf{B}_i\|$*, and* $\|\hat{\mathbf{T}} - \mathbf{T}\|_\infty$*.*

In this problem, we want to estimate the unknown MJS dynamics from a single state-input-mode trajectory. One has the flexibility to design the inputs so that the collected data has nice statistical properties. Meanwhile, we seek to establish finite-sample analysis for the estimation errors.

*The contents of this chapter are published in Du et al. (2022c) and Sattar et al. (2021) and represent an equal contribution from Zhe Du and Yahya Sattar. The proofs in this chapter are provided in Appendix A.

3.1.1 Contribution

In this chapter, we look into Problem P3.1 and provide an algorithm in Section 3.2 to estimate the MJS dynamics. In Section 3.3, we show that when the MJS is equipped with the mean-square stability, the estimation error is guaranteed to have a rate of $\mathcal{O}((n+p)\log(T)\sqrt{s/T})$, where n and p are the state and input dimensions respectively, and the $\mathcal{O}(1/\sqrt{T})$ dependence on the trajectory length T is optimal.

Our proof strategy introduces multiple innovations. To address Markovian mode switching, we introduce a mixing time argument to jointly track the approximate-dependence across the states and the modes. This in turn helps ensure each mode has sufficient samples and these samples are sufficiently informative. Secondly, due to the mean-square stability which allows for unstable trajectory realizations (see Example 2.9), it becomes non-trivial to determine whether the states have a light-tailed distribution (e.g., sub-Gaussian or sub-exponential). To circumvent this, we develop intricate system identification arguments that allow for heavy-tailed states. Such arguments can potentially benefit other problems involving heavy-tailed data.

3.1.2 Related Work

Learning dynamical models has a long history in the control community, with major theoretical results being either asymptotic (infinite sample) or non-asymptotic (finite sample) but with strong assumptions on persistence of excitation (Ljung, 1999). The main obstacle toward establishing non-asymptotic results using a single trajectory is the statistical dependency between the data.

In terms of learning the non-switched LTI systems, there is a recent surge of interest towards understanding the non-asymptotic performance from a single trajectory under mild assumptions (Oymak and Ozay, 2021), using statistical tools like martingales (Sarkar and Rakhlin, 2019; Simchowitz et al., 2018; Tsiamis and Pappas, 2019) or mixing time arguments (Kuznetsov and Mohri, 2017; Mohri and Rostamizadeh, 2008). Recently, Jedra and Proutiere (2020) provides precise rates for the finite-time identification of LTI systems using a single trajectory.

The problem becomes harder for hybrid and switched systems where the initial focus was on computational complexity as opposed to sample complexity of learning (Lauer and Bloch, 2018; Ozay et al., 2011). MJSs present unique statistical analysis challenges due to Markovian jumps and the weak mean-square stability. Preliminary asymptotic consistency result is established in Hespanhol and Aswani (2020) for the set-membership estimator. For stochastic switched systems, a special case of MJSs where the modes switch in an independently and identically distributed manner, Lale et al. (2021) proposes a novel identification method based on Lyapunov equations but without guarantees. Sarkar et al. (2019) is one of the early works to provide finite-sample result for learning stochastic switched systems, but its strong assumption on the system stability simplifies the statistical analysis and cannot be generalized to the mean-square stability case.

3.2 Identification Approach

The proposed MJS identification procedure is given in Algorithm 1. We assume one has access to an initial stabilizing controller $\mathbf{K}_{1:s}$ to start the identification. This has been a standard assumption in data-driven control (Abeille and Lazaric, 2018; Cohen et al., 2019; Dean et al., 2018; Ibrahimi et al., 2012; Simchowitz and Foster, 2020) for LTI systems. For MJSs, a thorough discussion on the validity of this assumption is provided later in Section 3.4. Note that, if the open-loop MJS is already mean-square stable, then one can simply set $\mathbf{K}_{1:s} = 0$ and carry out the rest of the MJS identification.

Algorithm 1: Identification of MJS

Input: A mean-square stabilizing controller $\mathbf{K}_{1:s}$; process and exploration noise variances $\sigma_{\mathbf{w}}^2$ and $\sigma_{\mathbf{z}}^2$; MJS trajectory $\{\mathbf{x}_t, \mathbf{z}_t, \omega_t\}_{t=0}^T$ generated using input $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$ with $\mathbf{z}_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{z}}^2 \mathbf{I}_p)$; and data clipping thresholds $c_{\mathbf{x}}$, $c_{\mathbf{z}}$.

Estimate $\mathbf{A}_{1:s}, \mathbf{B}_{1:s}$: **for all** modes $i \in [s]$ **do**

  $S_i = \left\{t \;\middle|\; \omega_t = i,\ \|\mathbf{x}_t\| \leq c_{\mathbf{x}}\sigma_{\mathbf{w}}\sqrt{\log(T)},\ \|\mathbf{z}_t\| \leq c_{\mathbf{z}}\sigma_{\mathbf{z}}\right\}$ // sub-sample data

  $\hat{\mathbf{\Theta}}_{1,i}, \hat{\mathbf{\Theta}}_{2,i} = \underset{\mathbf{\Theta}_1, \mathbf{\Theta}_2}{\arg\min} \sum_{t \in S_i} \|\mathbf{x}_{t+1} - \mathbf{\Theta}_1 \mathbf{x}_t/\sigma_{\mathbf{w}} - \mathbf{\Theta}_2 \mathbf{z}_t/\sigma_{\mathbf{z}}\|^2$ // regress with data S_i

  $\hat{\mathbf{B}}_i = \hat{\mathbf{\Theta}}_{2,i}/\sigma_{\mathbf{z}}$, $\hat{\mathbf{A}}_i = \hat{\mathbf{\Theta}}_{1,i}/\sigma_{\mathbf{w}} - \hat{\mathbf{B}}_i\mathbf{K}_i$

Estimate T: $\hat{\mathbf{T}}(i,j) = \frac{\sum_{t=1}^T \mathbf{1}_{\{\omega_t = j, \omega_{t-1} = i\}}}{\sum_{t=1}^T \mathbf{1}_{\{\omega_{t-1} = i\}}}$ //empirical frequency of transitions

Output: $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}$

With the controller $\mathbf{K}_{1:s}$, the input to the MJS is given by $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$, where $\{\mathbf{z}_t\}_{t=0}^{\mathsf{T}} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{z}}^2 \mathbf{I}_p)$ is known as the exploration noise. After collecting a length-T trajectory $\{\mathbf{x}_t, \mathbf{z}_t, \omega_t\}_{t=0}^T$, we sub-sample it and preserve data with bounded states $\mathbf{x}_t$ and excitations $\mathbf{z}_t$. By sub-sampling, the obtained samples will have manageable statistical properties, which is a compromise made to establish the analysis under the mean-square stability. In practice, one is free to use all the data without any sub-sampling.

After appropriate scaling, we regress over these samples to obtain the estimates $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}$. Lastly, we use the empirical frequency of the mode sequence as the estimate $\hat{\mathbf{T}}$.

3.3 Theory

The following theorem gives our main result on learning the dynamics of an unknown MJS from finite samples obtained from a single trajectory. One can refer to Theorems A.6 and A.22 in Appendix A for the detailed theorem statements and proofs. The result for estimating $\mathbf{T}$ in (3.2) mainly comes as a corollary from the work Zhang and Wang (2019).

Theorem 3.1 (Identification of MJS)**.** *Assume the MJS* $\Sigma = \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *to be identified is stabilizable, and the Markov matrix* $\mathbf{T}$ *is ergodic. Let* $\boldsymbol{\mathcal{L}}$ *denote the augmented closed-loop state matrix of the MJS under the initial stabilizing controller* $\mathbf{K}_{1:s}$ *as in* (2.10)*. For any* $\rho \in [\rho(\boldsymbol{\mathcal{L}}), 1)$*, let* $\tau := \tau(\boldsymbol{\mathcal{L}}, \rho)$ *denote the normalized power supremum of* $\boldsymbol{\mathcal{L}}$ *as in* (2.1)*. Suppose we run Algorithm 1 with* $c_{\mathbf{x}} = \mathcal{O}(\sqrt{n})$*,* $c_{\mathbf{z}} = \mathcal{O}(\sqrt{p})$*, and the trajectory length obeys* $T \geq \hat{\mathcal{O}}\big(\frac{\sqrt{s} t_{\text{MC}} \log^2(T)}{\pi_{\min}(1-\rho)}(n+p)\big)$*. Then, with probability at least* $1-\delta$*, for all* $i \in [s]$*, we have*

$$\max\{\|\hat{\mathbf{A}}_i - \mathbf{A}_i\|, \|\hat{\mathbf{B}}_i - \mathbf{B}_i\|\} \leq \hat{\mathcal{O}}\Bigg(\frac{(\sigma_{\mathbf{z}} + \sigma_{\mathbf{w}})}{\sigma_{\mathbf{z}}} \frac{(n+p)\tau \log(T)}{\pi_{\min}(1-\rho)} \sqrt{\frac{s}{T}} \Bigg), \tag{3.1}$$

$$\|\hat{\mathbf{T}} - \mathbf{T}\|_\infty \leq \hat{\mathcal{O}}\Bigg(\frac{t_{\text{MC}}}{\pi_{\min}} \sqrt{\frac{\log(T)}{T}} \Bigg). \tag{3.2}$$

Corollary 3.2 (Identification with known $\mathbf{B}_{1:s}$)**.** *Consider the same setting of Theorem 3.1. Additionally, suppose* $\mathbf{B}_{1:s}$ *is known. Then, setting* $\sigma_{\mathbf{z}} = 0$ *and solving only for the state matrices leads to a stronger upper bound* $\|\hat{\mathbf{A}}_i - \mathbf{A}_i\| \leq \hat{\mathcal{O}}(\frac{(n+p)\tau \log(T)}{\pi_{\min}(1-\rho)} \sqrt{\frac{s}{T}})$ *for all* $i \in [s]$*.*

Proof Sketch. Our proof strategy addresses the key challenges introduced by MJS and mean-square stability. We only emphasize the core technical challenges here. The idea is to think of the set S_i as a union of L subsets $S_i^{(\ell)}$ defined as follows:

$$S_i^{(\ell)} := \{\ell + kL \;\big|\; \omega_{\ell+kL} = i,\; \|\mathbf{x}_{\ell+kL}\| \leq c\sigma_{\mathbf{w}}\sqrt{n\log(T)},\; \|\mathbf{z}_{\ell+kL}\| \leq c\sigma_{\mathbf{z}}\sqrt{p}\}, \tag{3.3}$$

where $0 \leq \ell \leq L-1$ is a fixed offset and $k = 1, 2, \ldots, \lfloor \frac{T-L}{L} \rfloor$. The spacing of samples by $L \geq 1$ in each subset $S_i^{(\ell)}$ aims to reduce the statistical dependence across the samples belonging to that subset, to obtain weakly-dependent sub-trajectories. This weak dependence is due to the Markovian mode switching sequence $\{\omega_t\}_{t\geq 0}$ – unique to the MJS setting – and the system's memory (contributions from the past states). Thus L is primarily a function of the mixing time (t_{MC}) of the Markov chain and the spectral radius ($\rho(\boldsymbol{\mathcal{L}})$) of the MJS. At a high-level, by choosing sufficiently large L (e.g., $\mathcal{O}(\log(T))$), we can upper/lower bound the empirical covariance of the concatenated vector $\mathbf{h}_{\ell_k} := [\mathbf{x}_{\ell_k}^\intercal/\sigma_{\mathbf{w}} \; \mathbf{z}_{\ell_k}^\intercal/\sigma_{\mathbf{z}}]^\intercal$ for all $\ell_k \in S_i^{(\ell)}$.

Unlike related works on system identification and regret analysis (Dean et al., 2018; Lale et al., 2020a,b; Oymak and Ozay, 2021; Simchowitz et al., 2018), mean-square stability does not lead to strong high-probability bounds, as one can only upper bound $\|\mathbf{x}_t\|$ or $\mathbf{x}_t\mathbf{x}_t^\intercal$ in expectation. Therefore, in Algorithm 1, we sample only bounded state-excitation pairs $(\mathbf{x}_t, \mathbf{z}_t)$ on each mode $i \in [s]$. This boundedness enables us to control the covariance matrix of $\mathbf{h}_{\ell_k}$, despite mean-square stability and potentially heavy-tailed states, via non-asymptotic tool-sets (e.g., Vershynin (2012, Theorem 5.41)). When $\{\mathbf{z}_t\}_{t=0}^{T} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{z}}^2 \mathbf{I}_p)$, it can be easily shown that $\|\mathbf{z}_t\| \leq \hat{\mathcal{O}}(\sigma_{\mathbf{z}}\sqrt{p})$ for all $0 \leq t \leq T$ with high probability. We sample bounded $\mathbf{z}_t$ mainly for the simplicity of the analysis. The tightness of the upper bounds and the empirical performance are not affected significantly by sub-sampling $\{\mathbf{z}_t\}_{t=0}^{T}$ using the threshold $\hat{\mathcal{O}}(\sigma_{\mathbf{z}}\sqrt{p})$.

Heavy-tailed empirical covariance lower bound requires independence, and our sub-sampled data are only "approximately independent" (coupled over modes and history). To make matters worse, the fact that we sub-sample only bounded states introduces further dependencies. To resolve this, we introduce a novel strategy to construct (for the purpose of analysis) an independent subset of *processed states* from this larger weakly-dependent set. The independence is ensured by conditioning on the mode-sequence and truncating the contribution of earlier states. We then use perturbation-based techniques (see e.g., Sattar and Oymak (2022)) to deal with actual (non-truncated) states. The final ingredient is showing that, for each mode $i \in [s]$, with high probability, this carefully-crafted subset contains enough samples to ensure a well-conditioned covariance (with excitation provided by $\mathbf{z}_t$ and $\mathbf{w}_t$). With this in place, after stitching together the estimation error from L sub-trajectories $\{\mathbf{x}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$ for $0 \leq \ell \leq L-1$, least-squares will accurately estimate $\mathbf{A}_i$ and $\mathbf{B}_i$ from $\{\mathbf{x}_t, \mathbf{z}_t, \omega_t\}_{t \in S_i}$ for all $i \in [s]$. □

Our system identification result achieves near-optimal ($\hat{\mathcal{O}}(1/\sqrt{T})$) dependence on the trajectory length T. However, the effective sample complexity of our system identification algorithm is $\hat{\mathcal{O}}(s(n+p)^2 \log^2(T)/(\pi_{\min}^2(1-\rho)^2))$, that is, our sample complexity bound grows quadratically in the state dimension n, which can potentially be improved to linear via a more refined analysis of the state-covariance (see e.g., Dean et al. (2020); Simchowitz et al. (2018) for standard LTI systems). It also grows with the inverse of the minimum mode frequency as $1/\pi_{\min}^2$. Note that, $\pi_{\min}$ dictates the trajectory fraction of the least-frequent mode, thus, in the result $1/\pi_{\min}$ multiplier is unavoidable. Moreover, our sample complexity bound degrades as the Markov chain mixing time t_{MC} or the spectral radius $\rho(\boldsymbol{\mathcal{L}})$ of the augmented state matrix of the closed-loop MJS increase. This is because these parameters control the rate of mixing of the underlying process and we are using mixing-time arguments to derive our bounds. It is not desirable to have a sample complexity bound which degrades as t_{MC} or $\rho(\boldsymbol{\mathcal{L}})$ increase. In the case of standard LTI systems, it is well known that this behavior is qualitatively incorrect (Simchowitz et al., 2018). Interestingly, the more unstable LTI systems are easier to estimate. However, a fundamental limitation of mixing-time arguments is that the bounds deteriorate when the underlying process is slower to mix (Foster et al., 2020; Boffi et al., 2021; Ziemann et al., 2022; Sattar and Oymak, 2022).

In Corollary 3.2, we show that, when $\mathbf{B}_{1:s}$ is assumed to be known, $\mathbf{A}_{1:s}$ can be estimated regardless of the exploration strength $\sigma_{\mathbf{z}}$. This is because the excitation for the state matrix arises from noise $\mathbf{w}_t$. As we will see in Chapter 5, the distinct $\sigma_{\mathbf{z}}$ dependencies in Theorem 3.1 and Corollary 3.2 will lead to different regret bounds for MJS-LQR (albeit both bounds will be optimal up to $\mathrm{polylog}(T)$).

3.4 Discussion

Having access to an initial stabilizing controller has become a very common assumption in system identification (see for instance Lee and Lamperski (2020) and references therein) and adaptive control (Abeille and Lazaric, 2018; Cohen et al., 2019; Dean et al., 2018; Ibrahimi et al., 2012; Simchowitz and Foster, 2020) for LTI systems. On the other hand, for works where no initial stabilizing controller is required, there is usually a separate warm-up phase at the beginning, where coarse dynamics is learned, upon which a stabilizing controller is computed. Recent non-asymptotic system identification results (Faradonbeh et al., 2018a; Sarkar and Rakhlin, 2019) on potentially unstable LTI systems can be used to obtain coarse dynamics without a stabilizing controller. One can use random linear feedback to construct a confidence set of the dynamics such that any point in this set can produce a stabilizing controller by solving Riccati equations (Faradonbeh et al., 2018b). In the model-free setting, Lamperski (2020) provides asymptotic results and relies on persistent excitation assumption. Chen and Hazan (2021) designs subtle scaled one-hot vector input and collects the trajectory to estimate the dynamics, then a stabilizing controller can be solved via semi-definite programming. For MJS or general switched systems, to the best of our knowledge, there is no work on stabilizing unknown dynamics using single trajectory with guarantees. One challenge is the individual mode stability and overall mean-square stability does not imply each other due to mode switching. However, as outlined below, we can approach this problem leveraging what is recently done for the LTI case in the aforementioned literature (modulo some additional assumptions).

Similar to the LTI case, suppose we could obtain some coarse dynamics estimate $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}$, $\hat{\mathbf{T}}$, then we can solve for the optimal controller $\hat{\mathbf{K}}_{1:s}$ for the infinite-horizon MJS-LQR($\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}$, $\hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$) via coupled discrete-time algebraic Riccati equations. To investigate when $\hat{\mathbf{K}}_{1:s}$

can stabilize the MJS($\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}$), the key is to obtain sample complexity guarantees for this coarse dynamics, i.e. dependence of estimation error $\|\hat{\mathbf{A}}_i - \mathbf{A}_i\|$, $\|\hat{\mathbf{B}}_i - \mathbf{B}\|$, and $\|\hat{\mathbf{T}} - \mathbf{T}\|$ on sample size. Fortunately Theorem 4.2 provides the required estimation accuracy under which $\hat{\mathbf{K}}_{1:s}$ is guaranteed to be stabilizing. Thus, combining Theorem 4.2 with the estimation error bounds (in terms of sample size), the required accuracy can be translated to the required number of samples. Note that learning $\mathbf{T}$ is the same as learning a Markov chain, thus using the mode transition pair frequencies in an arbitrary single MJS trajectory, we can obtain an estimate $\hat{\mathbf{T}}$ as in Algorithm 1, and its sample complexity is given in Lemma A.6 in Appendix A. The more challenging part is the identification scheme and corresponding sample complexity for $\hat{\mathbf{A}}_{1:s}$ and $\hat{\mathbf{B}}_{1:s}$. Here, we outline two potential schemes.

- Suppose we can generate N i.i.d. MJS rollout trajectories, each with length T (small T, e.g. $T = 1$, is preferred to avoid potential unstable behavior and for the ease of the implementation). We can obtain least squares estimates $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}$ using only $\{\mathbf{x}_T, \mathbf{x}_{T-1}, \mathbf{u}_{T-1}, \omega_{T-1}\}$ from each trajectory, which is similar to the scheme in Dean et al. (2020) for LTI systems. Since only i.i.d. data is used in the computation, one can easily obtain the sample complexity in terms of N.

- If each mode in the MJS can run in isolation (i.e. for any $i \in [s]$, $\omega_t = i$ for all t) so that it acts as an LTI system, we can use recent advances on single-trajectory open-loop LTI system identification (Faradonbeh et al., 2018a; Sarkar and Rakhlin, 2019) to obtain coarse estimates together with sample complexity for $\hat{\mathbf{A}}_i$ and $\hat{\mathbf{B}}_i$ for every mode i.

We also note that while finding an initial stabilizing controller is theoretically very interesting and challenging, most results we know of are limited to simulated or numerical examples (see for instance Lee and Lamperski (2020) and references therein). This is because, from a practical standpoint, an initial stabilizing controller is almost always required in model-based approaches since running experiments with open-loop unstable plants can be very dangerous as the state could explode quickly.

3.5 Experiments

We provide experiments to investigate the efficiency and verify the theory of the proposed algorithms on synthetic datasets. Throughout, we show results from a synthetic experiment where entries of the true system matrices ($\mathbf{A}_{1:s}, \mathbf{B}_{1:s}$) are generated randomly from a standard normal distribution. We further scale each $\mathbf{A}_i$ to have $\|\mathbf{A}_i\| \leq 0.5$. Since this guarantees the MJS itself is mean-square stable, as we discussed in Section 2.3.1.1, we set controller $\mathbf{K}_{1:s} = 0$ in MJS identification Algorithm 1. The Markov matrix $\mathbf{T} \in \mathbb{R}_{\Delta}^{s \times s}$ is sampled from a Dirichlet distribution $\texttt{Dir}((s-1)\cdot\mathbf{I}_s+1)$, where $\mathbf{I}_s$ denotes the identity matrix. We assume that we have equal probability of starting in any initial mode.

Since for system identification, our main contribution is estimating $\mathbf{A}_{1:s}$ and $\mathbf{B}_{1:s}$ of the MJS, we omit the plots for estimating $\mathbf{T}$. Let $\hat{\mathbf{\Psi}}_i = [\hat{\mathbf{A}}_i, \hat{\mathbf{B}}_i]$ and $\mathbf{\Psi}_i = [\mathbf{A}_i, \mathbf{B}_i]$. We use $\|\hat{\mathbf{\Psi}} - \mathbf{\Psi}\|/\|\mathbf{\Psi}\| := \max_{i\in[s]} \|\hat{\mathbf{\Psi}}_i - \mathbf{\Psi}_i\|/\|\mathbf{\Psi}_i\|$ to investigate the convergence behavior of Algorithm 1. The clipping constants in this algorithm, $c_{\mathbf{x}}$, and $c_{\mathbf{z}}$ are chosen based on their lower bounds provided in Theorem 3.1. Fig. 3.1 provides the results that are averaged over 10 independent Monte Carlo runs.

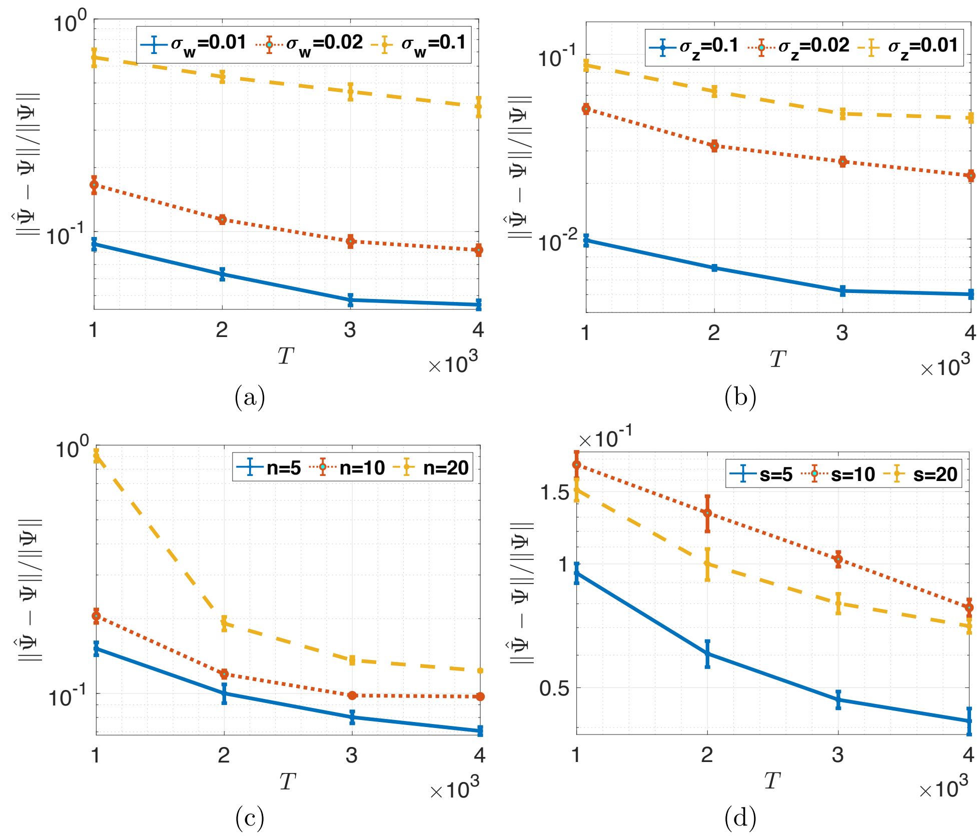


Figure 3.1: Performance of MJS identification algorithm. Influence of: (a) process noise $\sigma_{\mathbf{w}}$, (b) exploration noise $\sigma_{\mathbf{z}}$, (c) state dimension n, and (d) number of modes s.

We first empirically evaluate the effect of the noise variances $\sigma_{\mathbf{w}}$ and $\sigma_{\mathbf{z}}$. In particular, we study how the estimation errors vary with (i) $\sigma_{\mathbf{w}} = 0.01, \sigma_{\mathbf{z}} \in \{0.01, 0.02, 0.1\}$ and (ii) $\sigma_{\mathbf{z}} = 0.01, \sigma_{\mathbf{w}} \in \{0.01, 0.02, 0.1\}$. The number of states, inputs, and modes are set to $n = 5$, $p = 3$, and $s = 5$, respectively. Fig. 3.1 (a) and (b) demonstrate how the relative estimation error $\|\hat{\mathbf{\Psi}} - \mathbf{\Psi}\|/\|\mathbf{\Psi}\|$ changes as T increases. Each curve on the plot represents a fixed $\sigma_{\mathbf{w}}$ and $\sigma_{\mathbf{z}}$. These empirical results are all consistent with the theoretical bound in (3.1). In particular, the estimation errors degrade with increasing $\sigma_{\mathbf{w}}$ and decreasing $\sigma_{\mathbf{z}}$, respectively.

Now, we fix $\sigma_{\mathbf{w}} = \sigma_{\mathbf{z}} = 0.01$ and investigate the identification performance with varying numbers of states, inputs, and modes. Fig. 3.1 (c) and (d) show how the estimation error $\|\hat{\mathbf{\Psi}} - \mathbf{\Psi}\|/\|\mathbf{\Psi}\|$ changes with (left) $s = 5$, $n \in \{5, 10, 20\}$, $p = n - 2$ and (right) $n = 5$, $p = n - 2$, $s \in \{5, 10, 20\}$. As we can see, the performance is better with small n, p and s which is consistent with (3.1).

3.6 Conclusion

This chapter considers the problem of learning an MJS from a single trajectory. An identification method is proposed together with finite-sample guarantees. Numerical experimental results demonstrate the performance of our method and support our theory. Our statistical analysis tackles data distribution with heavy-tail caused by the mean-square stability. This allows us to provide high-probability non-asymptotic guarantees and may have further implications on analyzing heavy-tailed data ariseing from other problems.

Chapter 4
Certainty Equivalent Control

As in Chapter 3, the model of a real-world system is often obtained from learning from its data, i.e. system identification. Since one can only have access to finitely many data, and the data itself is noisy due to system process noise, there is usually an inevitable mismatch between such obtained model and the ground truth one. Using such an approximate model as a surrogate for the ground truth one to design control laws is known as certainty equivalent control. Certainty equivalent control can have decent performance when the approximate model is accurate enough. Perhaps, every model-based control in the real-world can be considered as certainty equivalent control since the exact model knowledge is never possible due to, for example, rounding errors. On the other hand, when the approximate model is way off, the resulting control law may destabilize thus endanger the ground truth system. Therefore, it is of importance to understand when certainty equivalent control works, e.g., ensures stability, and how well it works, e.g., suboptimality. Analysis of this kind is sometimes referred to as sensitivity or perturbation analysis.

Despite enormous studies on MJSs, theoretical understanding of MJS certainty equivalent control is somewhat lacking. This chapter investigates the performance of certainty equivalent control on MJS-LQR problems. Roughly speaking, suppose the mismatch level of the approximate MJS is ϵ, then this chapter establishes perturbation results for (i) the solution to the coupled Riccati equations and (ii) the cost in LQR problems, by providing explicit perturbation bounds with orders $\mathcal{O}(\epsilon)$ and $\mathcal{O}(\epsilon^2)$ respectively.

The outline is as follows: the certainty equivalent control scheme together with problems of interest is presented in Section 4.2; Section 4.3 provides the main perturbation analysis; numerical experimental results are in Section 4.4.

---

*The contents of this chapter are published in Sattar et al. (2022) and represent an equal contribution from Zhe Du and Yahya Sattar. The proofs in this chapter are provided in Appendix B.

4.1 Introduction and Related Work

The LQR problems are both theoretically well understood and commonly used in practice when the system dynamics are known. Its nice properties, e.g., admitting an elegant linear state feedback solution, make it a popular benchmark problem in reinforcement learning and adaptive control (Campi and Kumar, 1998; Abbasi-Yadkori and Szepesvári, 2011; Dean et al., 2020; Mania et al., 2019; Simchowitz and Foster, 2020; Abeille and Lazaric, 2020; Lale et al., 2020a).

A natural generalization of LTI systems are MJSs, which allow the dynamics of the underlying system to switch between multiple linear systems according to an underlying finite Markov chain. Similarly, a natural generalization of the LQR problem to MJS is to use mode-dependent cost matrices, which enables different control goals under different modes. While the optimal control for MJS-LQR is well understood when one has perfect knowledge of the system dynamics (Chizeck et al., 1986; Costa et al., 2006), in practice we do not always know the exact system dynamics and the transition matrix. For instance, one might use system identification techniques to learn an approximate model for the system. Designing optimal controllers for MJS-LQR with this approximate system dynamics and transition matrix in place of the true ones leads to so-called certainty equivalent control which is used extensively in practice. However, a theoretical understanding of the suboptimality of the CE control for MJS-LQR is somewhat lacking. The main challenge here is the hybrid nature of the problem that requires consideration of both the system dynamics uncertainty ϵ, and the underlying Markov transition matrix uncertainty η.

The solution to the infinite-horizon MJS-LQR involves coupled algebraic Riccati equations. Our goal is to understand the sensitivity of the solution of these equations and the corresponding optimal cost to perturbations in the system model. Toward this aim, we first establish an explicit $\mathcal{O}(\epsilon + \eta)$ perturbation bound for the solution to coupled algebraic Riccati equations that arise in the context of MJS-LQR. This in turn is used to establish an explicit $\mathcal{O}((\epsilon + \eta)^2)$ suboptimality bound on the cost. Finally, numerical experiments are provided to support our theoretical claims. Our proof strategy requires nontrivial advances over those of Mania et al. (2019); Konstantinov et al. (1993). Specifically, the coupled nature of these Riccati equations requires novel perturbation arguments, as they lack some of the nice properties of the standard Riccati equations, like uniqueness of solution under certain conditions or being amenable to matrix factorization based approaches.

Related Work: The performance analysis of certainty equivalent control for the classical LQR problem for LTI systems relies on the perturbation/sensitivity analysis of the underlying algebraic Riccati equations (ARE), i.e. how much the ARE solution changes when the parameters in the equation are perturbed. This problem is studied in many works (Konstantinov et al., 2003b). Early results on ARE solution perturbation bound are presented in Kenney and Hewer (1990) (continuous-time) and Konstantinov et al. (1993) (discrete-time). Most literature, however, only discusses perturbed solutions within the vicinity of the ground-truth solution. The uniqueness of such a perturbed solution is not discussed until Sun (1998), which is further refined in Sun (2002) to provide explicit perturbation bounds and generalization to complex equations. Tighter bounds are obtained (Zhou et al., 2009) when the parameters have a special structure like sparsity.

Channelled by ARE perturbation results, the end-to-end certainty equivalent LQR control suboptimality bound in terms of the dynamics perturbation is established in Mania et al. (2019). The field of certainty equivalent MJS-LQR control and the corresponding coupled ARE (cARE) perturbation analysis, however, is not well studied. Two perturbation results (Konstantinov et al., 2002, 2003a) for cARE only consider continuous-time cARE that arises in robust control applications and they are not directly applicable in MJS-LQR setting. Our work is also related to robust control for MJS (see, e.g., Shi et al. (1999); Costa et al. (2006)), where the focus is to numerically compute a controller to achieve a guaranteed cost under a given uncertainty bound. Whereas, we aim to characterize how the degradation in performance depends on perturbations in different parameters when certainty equivalent control is used. Therefore, our work contributes to the body of work in robust control and certainty equivalent control of MJS from a different perspective, and also paves the way to use these ideas in the context of learning-based adaptive control, which will be followed in Chapter 5.

4.2 Problem Formulation

In this chapter, we consider the certainty equivalent control for the infinite-horizon MJS-LQR problem (2.34) with the ground truth MJS $\Sigma = \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ and cost matrices $\mathbf{Q}_{1:s}$, $\mathbf{R}_{1:s}$. Following the notations in Lemma 2.22, let $\mathbf{P}^\star_{1:s}$, $\mathbf{K}^\star_{1:s}$, $J^\star$ respectively denote the solution to the $\text{cDARE}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$, the optimal controller, and the optimal LQR cost.

It is assumed that the model parameters, i.e., $\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}$, of the ground truth MJS Σ is unknown, but some approximate parameters, $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}$, are accessible that satisfy

$$\|\hat{\mathbf{A}}_i - \mathbf{A}_i\| \le \epsilon, \;\; \|\hat{\mathbf{B}}_i - \mathbf{B}_i\| \le \epsilon, \;\; \|\hat{\mathbf{T}} - \mathbf{T}\|_\infty \le \eta, \quad \forall i \in [s]. \tag{4.1}$$

In this chapter, we let $\hat{\Sigma} := \text{MJS}(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}})$ denote the MJS parameterized by the approximate parameters.

To perform certainty equivalent control, we solve for a controller using the approximate MJS $\hat{\Sigma}$ and then apply it to the ground truth MJS Σ. To ease the exposition, similar to Section 2.3.2, we define the following perturbed operators. For a sequence of positive semi-definite matrices $\mathbf{X}_{1:s} \in \mathbb{S}^n_{s,+}$, for all $i \in [s]$, define

$$\hat{\varphi}_i(\mathbf{X}_{1:s}) := \sum_{j\in[s]} \hat{\mathbf{T}}(i,j)\mathbf{X}_j, \tag{4.2}$$

$$\begin{aligned}\hat{\mathcal{R}}_i(\mathbf{X}_{1:s}) := \mathbf{Q}_i + \hat{\mathbf{A}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:s})\hat{\mathbf{A}}_i \\ - \hat{\mathbf{A}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:s})^\intercal \hat{\mathbf{B}}_i \left(\mathbf{R}_i + \hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:s})\hat{\mathbf{B}}_i\right)^{-1} \hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:s})\hat{\mathbf{A}}_i,\end{aligned} \tag{4.3}$$

$$\hat{\mathcal{K}}_i(\mathbf{X}_{1:s}) := -\left(\mathbf{R}_i + \hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:s})\hat{\mathbf{B}}_i\right)^{-1}\left(\hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:s})\hat{\mathbf{A}}_i\right). \tag{4.4}$$

We first solve for the following perturbed cDARE given by cDARE($\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$):

$$\begin{aligned}\mathbf{X}_1 &= \hat{\mathcal{R}}_1(\mathbf{X}_{1:s}) \\ \mathbf{X}_2 &= \hat{\mathcal{R}}_2(\mathbf{X}_{1:s}) \\ &\vdots \\ \mathbf{X}_s &= \hat{\mathcal{R}}_s(\mathbf{X}_{1:s})\end{aligned} \tag{4.5}$$

with respect to positive semi-definite matrices $\mathbf{X}_{1:s} \in \mathbb{S}^n_{s,+}$. Let $\hat{\mathbf{P}}_{1:s} \in \mathbb{S}^{s,n}_+$ denote the solution. Then the certainty equivalent controller $\hat{\mathbf{K}}_{1:s}$ is given by

$$\hat{\mathbf{K}}_i = \hat{\mathcal{K}}_i(\hat{\mathbf{P}}_{1:s}), \quad \forall\, i \in [s]. \tag{4.6}$$

Lastly, we apply the input $\hat{\mathbf{u}}_t = \hat{\mathbf{K}}_{\omega_t}\mathbf{x}_t$ to control the ground truth MJS Σ.

Let $\hat{J}$ denote the cost incurred by playing the certainty equivalent controller $\hat{\mathbf{K}}_{1:s}$. In the next section, we address the following questions.

(a) When is the perturbed cDARE in (4.5) guaranteed to have a unique positive semi-definite solution $\hat{\mathbf{P}}_{1:s}$?

(b) What is a tight upper bound on $\|\hat{\mathbf{P}}_{1:s} - \mathbf{P}^\star_{1:s}\|$?

(c) When does $\hat{\mathbf{K}}_{1:s}$ stabilize the true MJS?

(d) How large is the suboptimality gap $\hat{J} - J^\star$?

4.3 Theory

In this chapter, the main assumption is as follows, which guarantees the existence of solution to cDARE according to Lemma 2.22.

Assumption A4.1. *The* MJS-LQR$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ *problem satisfies the following.*

(a) For all $i \in [s]$, $\mathbf{Q}_i \succ 0$ and $\mathbf{R}_i \succ 0$.

(b) The MJS $\Sigma = \mathrm{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ is stabilizable.

Let $\mathbf{L}^\star_{1:s}$ denote the closed-loop state matrices under the optimal controller $\mathbf{K}^\star_{1:s}$ such that $\mathbf{L}^\star_i := \mathbf{A}_i + \mathbf{B}_i\mathbf{K}^\star_i$. Similar to the augmented state matrix in (2.10), define the augmented closed-loop state matrix $\boldsymbol{\mathcal{L}}^\star$ as

$$\boldsymbol{\mathcal{L}}^\star := \begin{bmatrix} \mathbf{T}(1,1)\cdot(\mathbf{L}^\star_1 \otimes \mathbf{L}^\star_1) & \mathbf{T}(2,1)\cdot(\mathbf{L}^\star_2 \otimes \mathbf{L}^\star_2) & \cdots & \mathbf{T}(s,1)\cdot(\mathbf{L}^\star_s \otimes \mathbf{L}^\star_s) \\ \mathbf{T}(1,2)\cdot(\mathbf{L}^\star_1 \otimes \mathbf{L}^\star_1) & \mathbf{T}(2,2)\cdot(\mathbf{L}^\star_2 \otimes \mathbf{L}^\star_2) & \cdots & \mathbf{T}(s,2)\cdot(\mathbf{L}^\star_s \otimes \mathbf{L}^\star_s) \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{T}(1,s)\cdot(\mathbf{L}^\star_1 \otimes \mathbf{L}^\star_1) & \mathbf{T}(2,s)\cdot(\mathbf{L}^\star_2 \otimes \mathbf{L}^\star_2) & \cdots & \mathbf{T}(s,s)\cdot(\mathbf{L}^\star_s \otimes \mathbf{L}^\star_s) \end{bmatrix}. \tag{4.7}$$

From Lemma 2.22, we know the closed-loop MJS is mean-square stable, thus $\rho(\boldsymbol{\mathcal{L}}^\star) < 1$ by Corollary 2.13. For any $\rho \in [\rho(\boldsymbol{\mathcal{L}}^\star), 1)$, let $\tau := \tau(\boldsymbol{\mathcal{L}}^\star, \rho)$ denote the normalized power supremum of $\boldsymbol{\mathcal{L}}^\star$ as in Definition 2.1. Table 4.1 summarizes some notations that will be used in the statements of the main results.

Table 4.1: Notations — MJS-LQR Certainty Equivalent Control

ξ	$\min\{\lVert\mathbf{B}_{1:s}\rVert_+^{-2}\lVert\mathbf{R}_{1:s}^{-1}\rVert_+^{-1}\lVert\mathbf{L}^\star_{1:s}\rVert_+^{-2}, \underline{\sigma}(\mathbf{P}^\star_{1:s})\}$
C_ϵ	$6\lVert\mathbf{A}_{1:s}\rVert_+^2\lVert\mathbf{B}_{1:s}\rVert_+\lVert\mathbf{P}^\star_{1:s}\rVert_+^2\lVert\mathbf{R}_{1:s}^{-1}\rVert_+$
C_ϵ^u	$6C_\epsilon^{-1}\lVert\mathbf{B}_{1:s}\rVert_+^{-2}\lVert\mathbf{P}^\star_{1:s}\rVert_+^{-1}\lVert\mathbf{R}_{1:s}^{-1}\rVert_+^{-1}$
C_η	$2\lVert\mathbf{A}_{1:s}\rVert_+^2\lVert\mathbf{B}_{1:s}\rVert_+^4\lVert\mathbf{P}^\star_{1:s}\rVert_+^3\lVert\mathbf{R}_{1:s}^{-1}\rVert_+^2$
C_η^u	$6C_\eta^{-1}$
$\Gamma_\star$	$\max\{\lVert\mathbf{A}_{1:s}\rVert_+, \lVert\mathbf{B}_{1:s}\rVert_+, \lVert\mathbf{P}^\star_{1:s}\rVert_+, \lVert\mathbf{K}^\star_{1:s}\rVert_+\}$
$\bar{\epsilon}_{\mathbf{K}}$	$\frac{1-\rho}{2\sqrt{s}\tau(1+2\lVert\mathbf{L}^\star_{1:s}\rVert)\lVert\mathbf{B}_{1:s}\rVert}$

In the following, we will show that despite being coupled, cDARE for MJS-LQR satisfies nice local Lipschitz properties. To be more precise, we show that if the approximate MJS is accurate enough, i.e., ϵ and η are sufficiently small, we can guarantee that, not only the positive definite solution $\hat{\mathbf{P}}_{1:s}$ to the perturbed cDARE uniquely exists, but also $\hat{\mathbf{P}}_{1:s}$ is guaranteed to be close to $\mathbf{P}^\star_{1:s}$.

Theorem 4.1. *Suppose Assumption A4.1 holds and* $\epsilon \leq \min\left\{\frac{C_\epsilon^u \xi(1-\rho)^2}{204 n s \tau^2},\ \|\mathbf{B}_{1:s}\|\right\}$*,* $\eta \leq \frac{C_\eta^u \xi(1-\rho)^2}{48 n s \tau^2}$*. Then, the perturbed cDARE in (4.5) is guaranteed to have a unique solution* $\hat{\mathbf{P}}_{1:s}$ *in* $\mathbb{S}_{s,+}^n$ *such that* $\hat{\mathbf{P}}_i \succ 0$ *for all* i *and*

$$\|\hat{\mathbf{P}}_{1:s} - \mathbf{P}_{1:s}^\star\| \leq \frac{\sqrt{ns}\tau}{1-\rho}(C_\epsilon \epsilon + C_\eta \eta). \tag{4.8}$$

From the constants, we see we would have milder requirements on ϵ and η and a tighter bound on $\|\hat{\mathbf{P}}_{1:s} - \mathbf{P}_{1:s}^\star\|$ when (i) $\|\mathbf{A}_{1:s}\|, \|\mathbf{B}_{1:s}\|$, (ii) $\|\mathbf{L}_{1:s}^\star\|$, τ, and (iii) $\|\mathbf{R}_{1:s}^{-1}\|$ are smaller. These translate to the cases when (i) the true MJS is easier to stabilize; (ii) the closed-loop MJS under the optimal controller is more stable; and (iii) the input dominates more in the cost function. The role of τ in this theorem is closely related to the damping property in ARE perturbation analysis (Kenney and Hewer, 1990). The coefficients for ϵ and η on the RHS of (4.8) are also known as condition numbers in algebraic Riccati equation sensitivity literature (Sun, 2002).

Note that the perturbation upper bound in Theorem 4.1, when setting $s = 1$ and $\eta = 0$, is consistent with (Mania et al., 2019, Proposition 1) developed for the LTI case except that we suffer an additional $\sqrt{n}$ term. This is because, due to the coupled nature of $\hat{\mathbf{P}}_{1:s}$ through cDARE, we proceed by first vectorizing and stacking cDARE into a single equation to evaluate $[\mathbf{vec}(\hat{\mathbf{P}}_1)^\intercal, \cdots, \mathbf{vec}(\hat{\mathbf{P}}_s)^\intercal]^\intercal$, then convert it back to $\hat{\mathbf{P}}_{1:s}$ through reshaping. Certain norm equivalency arguments (Fact B.2) are needed to carry perturbation results through these back-and-forth reshaping steps, which produces this additional $\sqrt{n}$. On the other hand, these steps and thus the $\sqrt{n}$ term are not needed for the LTI case, since only a single Riccati equation is involved.

It is easy to extend this result to the cases when an approximate $\hat{\mathbf{Q}}_{1:s}$ with $\|\hat{\mathbf{Q}}_{1:s} - \mathbf{Q}_{1:s}\| \leq \epsilon$ is used in place of $\mathbf{Q}_{1:s}$ in the computations, which can be useful when the cost includes a term of the form $\|\mathbf{y}_t\|^2$ where $\mathbf{y}_t = \mathbf{C}_{\omega_t}\mathbf{x}_t$ represents the output, and we only have an approximate parameter $\hat{\mathbf{C}}_{1:s}$. In this case, $\mathbf{Q}_i = \mathbf{C}_i^\intercal \mathbf{C}_i$ and $\hat{\mathbf{Q}}_i = \hat{\mathbf{C}}_i^\intercal \hat{\mathbf{C}}_i$.

In the next result, we leverage Theorem 4.1 to show how the controller $\hat{\mathbf{K}}_{1:s}$ computed from a perturbed cDARE solution deviates from the optimal one, i.e., how $\|\hat{\mathbf{K}}_{1:s} - \mathbf{K}_{1:s}^\star\|$ depends on ϵ and η, and when $\hat{\mathbf{K}}_{1:s}$ stabilizes the true MJS (such that $\hat{J}$ will be bounded). Moreover, with the help of (Jansch-Porto et al., 2020, Lemma 3), which provides a relation between suboptimality gap $\hat{J} - J^\star$ and $\|\hat{\mathbf{K}}_{1:s} - \mathbf{K}_{1:s}^\star\|$, we establish an upper bound for $\hat{J} - J^\star$ in terms of ϵ and η.

Theorem 4.2. *Suppose Assumption A4.1 holds, the Markov matrix* $\mathbf{T}$ *is ergodic, and* ϵ, η *satisfy the bounds in Theorem 4.1 and* $C_\epsilon \epsilon + C_\eta \eta \leq \frac{(1-\rho)\min\{\Gamma_\star, \underline{\sigma}(\mathbf{R}_{1:s})^2 \bar{\epsilon}_{\mathbf{K}}\}}{28\sqrt{ns}\tau \Gamma_\star^3(\underline{\sigma}(\mathbf{R}_{1:s}) + \Gamma_\star^3)}$*. Then, the certainty* *equivalent controller $\hat{\mathbf{K}}_{1:s}$ stabilizes the ground truth MJS Σ and*

$$\|\mathbf{K}^{\star}_{1:s} - \hat{\mathbf{K}}_{1:s}\| \leq 28\sqrt{ns}\tau\Gamma_{\star}^{3}\frac{(\underline{\sigma}(\mathbf{R}_{1:s}) + \Gamma_{\star}^{3})}{(1-\rho)\underline{\sigma}(\mathbf{R}_{1:s})^{2}}(C_{\epsilon}\epsilon + C_{\eta}\eta) \tag{4.9}$$

$$\hat{J} - J^{\star} \leq 1600n^{1.5}s^{2.5}\min\{n,p\}\tau^{3}\Gamma_{\star}^{6}\frac{(\|\mathbf{R}_{1:s}\| + \Gamma_{\star}^{3})^{3}}{(1-\rho)^{3}\underline{\sigma}(\mathbf{R}_{1:s})^{4}}(C_{\epsilon}\epsilon + C_{\eta}\eta)^{2}\sigma_{\mathbf{w}}^{2}. \tag{4.10}$$

This result states that the suboptimality has quadratic dependency on the uncertainties ϵ and η, and degrades when the MJS has larger number of modes s, system order n, or noise variance $\sigma_{\mathbf{w}}^2$. Similar to the earlier discussion, Theorem 4.2 is also consistent with its LTI counterpart (Mania et al., 2019, Theorem 1) except the n term.

Our suboptimality result has important implications in data-driven control for MJS. Suppose the uncertainties ϵ and η in the system dynamics and the transition matrix are due to estimation errors induced by a system identification procedure that uses T samples. Then, if the estimation error decays as $\mathcal{O}(1/\sqrt{T})$ (which is typical for ϵ as in learning LTI systems (Oymak and Ozay, 2021; Sarkar and Rakhlin, 2019) and for η in learning Markov chains (Zhang and Wang, 2019)), Theorem 4.2 implies that the suboptimality decays as $\mathcal{O}(1/T)$. Thus, given a desired suboptimality level for the certainty equivalent controller, one can use this relation to infer the required number of samples, which has been employed in Chapter 5 to establish regret analysis for adaptive control.

4.4 Experiments

In this section, we present some numerical results to support our proposed theory. We fix $n = 10$ and $p = 5$. The true system matrices $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s})$ were generated randomly from the standard normal distribution. We scaled each $\mathbf{A}_i$ to have a spectral radius equal to 0.3 to obtain a mean square stable MJS. We set $\mathbf{Q}_i = \underline{\mathbf{Q}}_i\underline{\mathbf{Q}}_i^{\top}, \mathbf{R}_i = \underline{\mathbf{R}}_i\underline{\mathbf{R}}_i^{\top}, \hat{\mathbf{A}}_i = \mathbf{A}_i + \epsilon_{\mathbf{A}}\underline{\mathbf{A}}_i$, and $\hat{\mathbf{B}}_i = \mathbf{B}_i + \epsilon_{\mathbf{B}}\underline{\mathbf{B}}_i$, where $\underline{\mathbf{Q}}_i$, $\underline{\mathbf{R}}_i$, $\underline{\mathbf{A}}_i$, and $\underline{\mathbf{B}}_i$ were generated randomly from the standard normal distribution; and $\epsilon_{\mathbf{A}}$ and $\epsilon_{\mathbf{B}}$ are some fixed scalars. Here we experimentally study the influences of perturbation on $\mathbf{A}_{1:s}$ and $\mathbf{B}_{1:s}$ separately with $\epsilon_{\mathbf{A}}$ and $\epsilon_{\mathbf{B}}$. Note that ϵ defined in (4.1) is equal to $\max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\}$. The true Markov transition matrix $\mathbf{T}$ was sampled from a Dirichlet distribution $\texttt{Dir}((s-1)\cdot \mathbf{I}_s + 1)$, and we let the approximate $\hat{\mathbf{T}} = \mathbf{T} + \mathbf{E}$, where the perturbation $\mathbf{E} = \eta_{\mathbf{T}}(\texttt{Dir}((s-1)\cdot\mathbf{I}_s + 1)) - \hat{\mathbf{T}})$ for $\eta_{\mathbf{T}} \in [0,1]$.

We study how the Riccati equation solution perturbation and suboptimality gap vary with $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \eta_{\mathbf{T}} \in \{0.01, 0.02, 0.05, 0.1, 0.2, 0.3\}$ and the number of modes $s \in \{10, 20, 30, 40\}$. For each choice of $\epsilon_{\mathbf{A}}$, $\epsilon_{\mathbf{B}}$, and $\eta_{\mathbf{T}}$, we run 100 experiments and record the respective maximums of $\Delta_{\mathbf{P}} := \max_i \|\hat{\mathbf{P}}_i - \mathbf{P}_i^{\star}\|/\|\mathbf{P}_i^{\star}\|$ and $\Delta_J := (\hat{J} - J^{\star})/J^{\star}$ over all 100 runs. In Figures 4.1 and

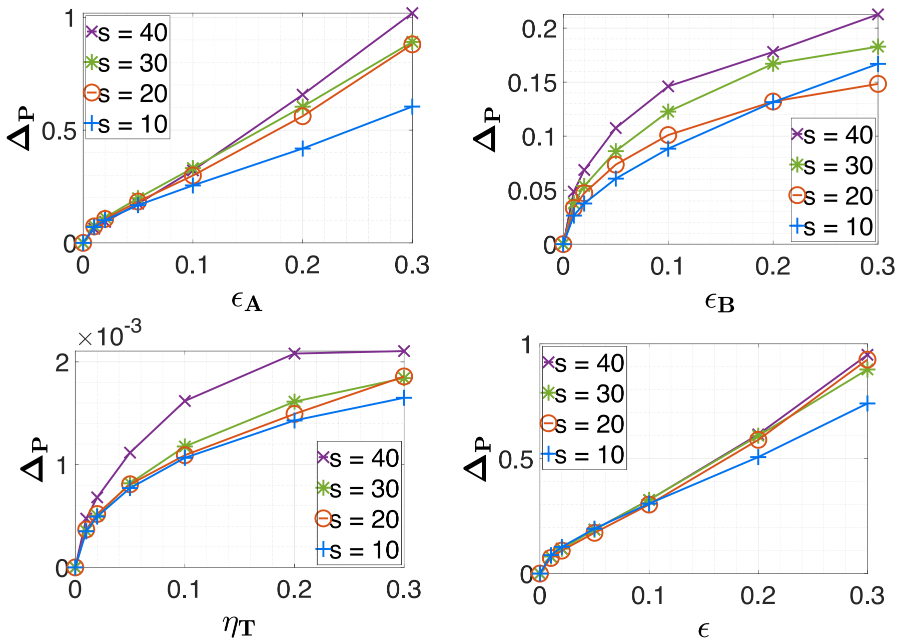


Figure 4.1: Riccati solution perturbation. Left to right: $\epsilon_{\mathbf{B}} = \eta_{\mathbf{T}} = 0$, $\epsilon_{\mathbf{A}} = \eta_{\mathbf{T}} = 0$, $\epsilon_{\mathbf{A}} = \epsilon_{\mathbf{B}} = 0$, and $\epsilon = \epsilon_{\mathbf{A}} = \epsilon_{\mathbf{B}} = \eta_{\mathbf{T}}$.

4.2, we have four plots showing $\Delta_{\mathbf{P}}$ and Δ_J versus uncertainties (i) $\epsilon_{\mathbf{A}}$ ($\epsilon_{\mathbf{B}} = \eta_{\mathbf{T}} = 0$), (ii) $\epsilon_{\mathbf{B}}$ ($\epsilon_{\mathbf{A}} = \eta_{\mathbf{T}} = 0$), (iii) $\eta_{\mathbf{T}}$ ($\epsilon_{\mathbf{A}} = \epsilon_{\mathbf{B}} = 0$), and (iv) $\epsilon = \epsilon_{\mathbf{A}} = \epsilon_{\mathbf{B}} = \eta_{\mathbf{T}}$.

Figure 4.1 presents four plots that demonstrate how $\Delta_{\mathbf{P}}$ changes as $\epsilon_{\mathbf{A}}$, $\epsilon_{\mathbf{B}}$, $\eta_{\mathbf{T}}$, and ϵ increase, respectively. Each curve on the plot represents a fixed number of modes s. These empirical results are all consistent with (4.8). In particular, Figure 4.1 (right) shows that given the uncertainty in the system matrices and in the Markov transition matrix is bounded by ϵ, the perturbation bound to coupled Riccati equations has the rate $\mathcal{O}(\epsilon)$ which degrades linearly as ϵ increases. Further, it can be easily seen that the gaps indeed increase with the number of modes in the system. Figure 4.2 demonstrates the relationship between the relative suboptimality Δ_J and the five parameters $\epsilon_{\mathbf{A}}$, $\epsilon_{\mathbf{B}}$, $\eta_{\mathbf{T}}$, ϵ and s. As can be seen in Figure 4.2 (right), given the uncertainty in the system matrices and in the Markov transition matrix is bounded by ϵ, the perturbation bounds to the optimal cost decay quadratically which is consistent with our theory.

4.5 Conclusion

In this chapter, we provide a perturbation analysis for cDARE, which arises in the solution of MJS-LQR, and an end-to-end suboptimality guarantee for certainty equivalence control for MJS-LQR. Our results show the robustness of the optimal policy to perturbations in system dynamics and establish the validity of the certainty equivalent control in a neighborhood of the original system. This chapter opens up multiple future directions. First,

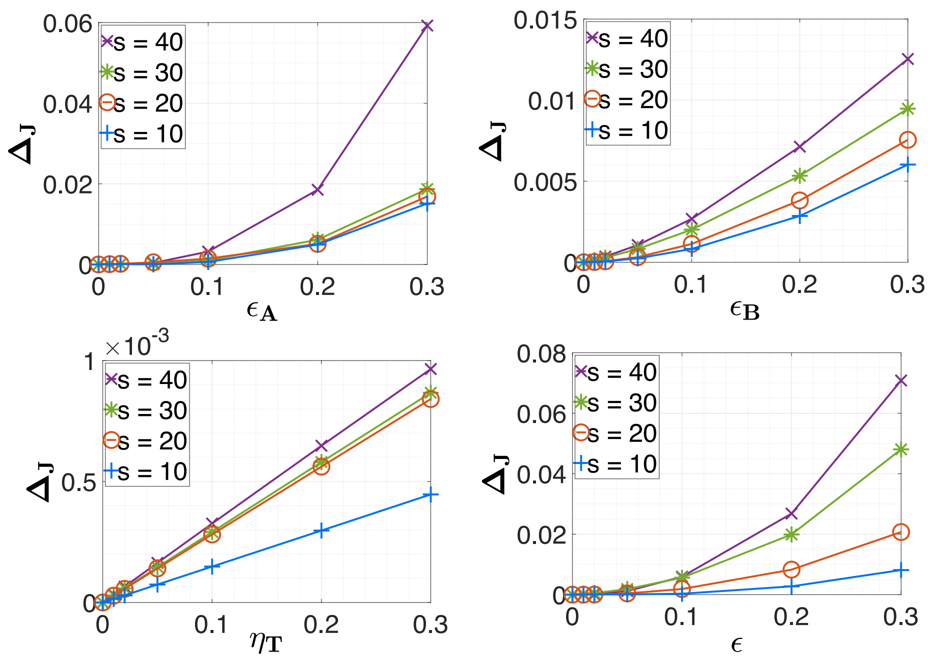


Figure 4.2: Suboptimality gap. Left to right: $\epsilon_{\mathbf{B}} = \eta_{\mathbf{T}} = 0$, $\epsilon_{\mathbf{A}} = \eta_{\mathbf{T}} = 0$, $\epsilon_{\mathbf{A}} = \epsilon_{\mathbf{B}} = 0$, and $\epsilon = \epsilon_{\mathbf{A}} = \epsilon_{\mathbf{B}} = \eta_{\mathbf{T}}$.

with proper system identification algorithms, e.g., Chapter 3, we can analyze model-based online/adaptive algorithms where control policy is updated continuously over a single trajectory. Second, a natural extension would be to study MJS with output measurements where states are only partially observed, i.e., the LQG setting. This will require considering the dual coupled Riccati equations for filtering.

Chapter 5
Adaptive Quadratic Control

Learning how to effectively control unknown dynamical systems from data is crucial for intelligent autonomous systems. This task becomes a significant challenge when the underlying dynamics are changing with time. Motivated by this challenge, this chapter considers solving the MJS-LQR problem with unknown MJS dynamics in an online way. By taking a model-based perspective, we consider identification-based adaptive control for MJS.

We then propose an adaptive control scheme that incorporates the system identification procedure in Chapter 3 together with the certainty equivalent control in Chapter 4 to adapt the controllers in an episodic fashion. Combining our identification sample complexity results with perturbation results for certainty equivalent control, we prove that when the episode lengths are appropriately chosen, the proposed adaptive control scheme achieves $\mathcal{O}(\sqrt{T})$ regret.

This chapter is organized as follows: the problem formulation is in Section 5.2; the adaptive control scheme is proposed in Section 5.3 with performance guarantees in Section 5.4; Section 5.5 presents experimental results that demonstrate the performance of our approach and support our theory.

5.1 Introduction and Related Work

A canonical problem at the intersection of machine learning and control is that of adaptive control of an unknown dynamical system. An intelligent autonomous system is likely to encounter such a task; from an observation of the inputs and outputs, it needs to both learn and effectively control the dynamics. A commonly used control paradigm is the LQR problem, which is theoretically well understood when system dynamics are linear and known. LQR also provides an interesting benchmark, when system dynamics are unknown, for reinforcement learning (RL) with continuous state and action spaces and for adaptive control

*The contents of this chapter are published in Du et al. (2022c) and Sattar et al. (2021). The proofs in this chapter are provided in Appendix C.

(Abbasi-Yadkori and Szepesvári, 2011; Abeille and Lazaric, 2020; Campi and Kumar, 1998; Dean et al., 2020; Lale et al., 2020a; Mania et al., 2019).

While the MJS-LQR problem is also well understood when one has perfect knowledge of the system dynamics (Chizeck et al., 1986; Costa et al., 2006), in practice, it is not always possible to have a perfect knowledge of the system dynamics and the Markov transition matrix. For instance, a Mars rover optimally exploring an unknown heterogeneous terrain, optimal solar power generation on a cloudy day, or controlling investments in financial markets may be modeled as MJS-LQR problems with unknown system dynamics (Blackmore et al., 2005; Cajueiro, 2002; Loparo and Abdel-Malek, 1990; Svensson and Williams, 2008; Ugrinovskii and Pota, 2005).

Earlier works have aimed at analyzing the asymptotic properties (i.e., stability) of adaptive controllers for unknown MJSs both in continuous-time (Caines and Zhang, 1995) and discrete-time (Xue and Guo, 2001) settings, however, despite the practical importance of MJSs, non-asymptotic sample complexity results and regret analysis for MJSs are lacking. In the case of stochastic jump systems, (which are the switched dynamical systems where the switching of modes are i.i.d.), when the only unknown is the mode distribution, recent works study data-driven stability verification (Gatsis and Pappas, 2021) and stabilization (Schuurmans et al., 2019) with non-asymptotic guarantees. However, it is difficult to generalize these approaches to general MJSs as well as MJSs with unknown dynamics.

The high-level challenge here is the hybrid nature of the problem that requires consideration of both the system dynamics and the underlying Markov transition matrix. A related challenge is that, typically, the stability of MJS is understood only in the *mean-square sense*. This is in stark contrast to the deterministic stability (e.g., as in LQR), where the system is guaranteed to converge towards an equilibrium point in the absence of noise. In contrast, the convergence of MJS trajectories towards an equilibrium depends heavily on how the switching between modes occurs. As we have discussed in Example 2.9, an MJS that is mean-square stable may still have an explosive trajectory realization under an unfavorable mode switching sequence. High probability light-tail bounds are, therefore, not applicable without very strong assumptions on the joint spectral radius of different modes (cf. Sarkar et al. (2019)). Perhaps more surprisingly, there are examples of MJS with all modes individually stable, however due to switching, the system exhibits an unstable behavior on average, and the MJS is not mean-square stable (see Example 3.17 of Costa et al. (2006)). Therefore, finding controllers to individually stabilize the mode dynamics does not guarantee that the overall system will be stable when mode switches over time. This more relaxed notion of mean-square stability presents major challenges in learning, controlling, and statistical analysis.

5.1.1 Contribution

In this chapter, by incorporating the system identification method in Chapter 3, we propose an adaptive control scheme to solve MJS-LQR problems with unknown MJS dynamics and provide its performance guarantees in terms of regret analysis, while assuming only mean-square stability. Specifically, our theoretical contributions are as follows[1]:

- $\mathcal{O}(\sqrt{T})$**-regret bound:** When the system dynamics are unknown, we show that the certainty-equivalent adaptive MJS-LQR Algorithm (Algorithm 2) achieves a regret bound of $\mathcal{O}(\sqrt{T})$. Remarkably, this coincides with the optimal regret bound for the standard LQR problem obtained via certainty equivalence (Mania et al., 2019).

- $\mathcal{O}(\textbf{polylog}(T))$**-regret with partial knowledge:** We also consider the practically relevant setting where the state matrices are unknown but the input matrices are known. We show that the regret bound can be significantly improved to $\mathcal{O}(\text{polylog}(T))$. This bound also coincides with the poly-logarithmic regret bound for the standard LQR with the knowledge of the input matrix $\mathbf{B}$ (Cassel et al., 2020).

5.1.2 Related Work

There are mainly two types of approaches for solving the optimal control problem with unknown system dynamics: model-based and model-free ones. Model-based ones rely on system identification to obtain an estimate of the system dynamics, based on which the controllers are designed. Model-free ones solve for the controller directly from the data without relying on model estimates. A comparison with the related works, in the LQR setting, is provided in Table 5.1.

• **Model-Based Approaches:** For LTI systems, there have been a large body of work on model-based adaptive control, and they have sophisticated control performance guarantees from the regret perspective (Abbasi-Yadkori et al., 2019; Abbasi-Yadkori and Szepesvári, 2011; Dean et al., 2020; Faradonbeh et al., 2020b; Hazan et al., 2020; Mania et al., 2019). Specifically, Simchowitz and Foster (2020) achieves $\mathcal{O}(\sqrt{T})$ regret lower bound for adaptive LQR control with unknown system dynamics, while Cassel et al. (2020) and Lale et al. (2020b) achieve logarithmic regret upper bound, with partial knowledge of the system and persistence of excitation assumption, respectively, as no additional exploration noise is needed to guarantee learnability of the system. However, in the MJS setting, due to the lack of well established identification analysis, prior works (Caines and Zhang, 1995; Xue and Guo, 2001) are only able to provide guarantees from the stability aspect. The case of

[1] orders of magnitude here are up to poly-logarithmic factors

input design without system state dynamics is considered in Baltaoglu et al. (2016), which can be thought of as a generalization of linear bandits to have a Markovian structure in the reward function without any continuous dynamic structure. However, only a regret lower bound is provided in Baltaoglu et al. (2016). Finally, we refer the reader to the survey papers (Gaudio et al., 2019; Matni et al., 2019; Recht, 2019) for a broad overview of the recent developments on non-asymptotic system identification, adaptive control and reinforcement learning from the perspective of optimization and control.

• **Model-Free Approaches:** Somehow orthogonal to the above developments, but still highly relevant, are approaches that sidestep system identification and try to learn an optimal controller (policy) directly (among many others, see e.g., Fazel et al. (2018); Mohammadi et al. (2020); Zhang et al. (2020); Zheng et al. (2021)). These works analyze the optimization landscape of LQR and related optimal control problems and provide polynomial-time algorithms that lead to a globally convergent search in the space of controllers. Importantly, these optimization algorithms do not require the knowledge of the system parameters as long as relevant quantities like gradients can be approximated from simulated system trajectories. More recently, this line of work is extended to MJSs in Jansch-Porto et al. (2020), significantly expanding their utility. However, these works require multiple trajectories to estimate the gradients as opposed to a controller that adapts at run-time, therefore, they provide a complementary perspective to the single trajectory adaptation and regret analysis in our work.

Table 5.1: Comparison with prior works in the LQR setting.

Model	**Reference**	**Regret**	**Computational Complexity**	**Cost**	**Stabilizability/ Controllability**
LTI	Abbasi-Yadkori and Szepesvári (2011)	$\sqrt{T}$	Exponential	Strongly Convex	Controllable
	Ibrahimi et al. (2012)	$\sqrt{T}$	Exponential	Convex	Controllable
	Abeille and Lazaric (2018) (one dim. systems)	$\sqrt{T}$	Polynomial	Strongly Convex	Stabilizable
	Dean et al. (2018)	$T^{2/3}$	Polynomial	Convex	Stabilizable
	Mania et al. (2019)	$\sqrt{T}$	Polynomial	Strongly Convex	Controllable
	Cohen et al. (2019)	$\sqrt{T}$	Polynomial	Strongly Convex	Strongly Stabilizable
	Faradonbeh et al. (2020a); Simchowitz and Foster (2020)	$\sqrt{T}$	Polynomial	Strongly Convex	Stabilizable
	Cassel et al. (2020) (known $\mathbf{A}$ or $\mathbf{B}$)	$\text{polylog}(T)$	Polynomial	Strongly Convex	Strongly Stabilizable
MJS	**Ours**	$s^2\sqrt{T}$	Polynomial	Strongly Convex	Mean-square stabilizable
	Ours (known $\mathbf{B}_{1:s}$)	$s^2\text{polylog}(T)$	Polynomial	Strongly Convex	Mean-square stabilizable

5.2 Problem Formulation

In this chapter, we consider the following adaptive control for finite-horizon MJS-LQR problems.

Problem P5.1 (Adaptive Control for MJS-LQR)**.** *Solve the finite-horizon MJS-LQR problem* (2.29) *with unknown MJS* $\Sigma = \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *and known LQR cost matrices* $\mathbf{Q}_{1:s}$

and $\mathbf{R}_{1:s}$.

The solution to the above problem usually involves procedures of learning, either the dynamics or directly the controllers. Adaptive control suffers additional costs as (i) the lack of the exact knowledge of the system and (ii) the exploration-exploitation trade-off – the necessity to sacrifice short-term input optimality to boost learning, so that overall long-term optimality can be improved.

Because of this, to evaluate the performance of an adaptive scheme, one is interested in the notion of regret – how much more cost it will incur if one could have applied the optimal controllers? In our setting, we compare the resulting cost against the optimal cost $T \cdot J^\star$ where $J^\star$ is the optimal infinite-horizon average cost in Lemma 2.22.

Compared to the regret analysis of adaptive LQR problem (Dean et al., 2018) for LTI systems, the cost analysis here requires additional consideration of Markov chain mixing, which is addressed in this chapter.

5.3 Approach

Our adaptive MJS-LQR control scheme is given in Algorithm 2. It is performed on an epoch-by-epoch basis; a fixed controller is used for each epoch, and from epoch to epoch, the controller is updated using the trajectory generated in the most recent epoch. Note that a new epoch is just a continuation of previous epochs instead of restarting the MJS. Similar to the discussion in Section 3.2, we assume, at the beginning of epoch 0, that one has access to a stabilizing controller $\mathbf{K}_{1:s}^{(0)}$. During epoch q, the controller $\mathbf{K}_{1:s}^{(q)}$ is used together with additive exploration noise $\mathbf{z}_t^{(q)} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{z},q}^2 \mathbf{I}_p)$ to boost learning. At the end of epoch q, the trajectory during that epoch is used to obtain a new MJS dynamics estimate $\mathbf{A}_{1:s}^{(q)}, \mathbf{B}_{1:s}^{(q)}, \mathbf{T}^{(q)}$ using the MJS identification Algorithm 1. Then, we set the controller $\mathbf{K}_{1:s}^{(q+1)}$ for epoch $q+1$ to be the optimal controller for the infinite-horizon MJS-LQR$(\mathbf{A}_{1:s}^{(q)}, \mathbf{B}_{1:s}^{(q)}, \mathbf{T}^{(q)}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$, which can be computed as in Lemma 2.22. Note that cDARE may not be solvable for arbitrary MJS models or cost matrices, but our theory guarantees that when epoch lengths are appropriately chosen, cDARE parameterized by $\mathbf{A}_{1:s}^{(q)}, \mathbf{B}_{1:s}^{(q)}, \mathbf{T}^{(q)}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$ is solvable for every epoch q. This control design based on the estimated dynamics is also referred to as certainty equivalent control.

To achieve theoretically guaranteed performance, i.e., sublinear regret, the key is to have a subtle scheduling of epoch lengths T_q and exploration noise variance $\sigma_{\mathbf{z},q}^2$. We choose T_q to increase exponentially with rate $\gamma > 1$, and set $\sigma_{\mathbf{z},q}^2 = \sigma_{\mathbf{w}}^2 / \sqrt{T_q}$, which collectively guarantee $\hat{\mathcal{O}}(\sqrt{T})$ regret when combined with the system identification result from Theorem 3.1. Intuitively, this scheduling can be interpreted as follows: (i) the increase of epoch lengths

Algorithm 2: Adaptive MJS-LQR

Input: Initial epoch length T_0; initial stabilizing controller $\mathbf{K}_{1:s}^{(0)}$; epoch incremental ratio $\gamma > 1$; and data clipping thresholds $c_{\mathbf{x}}$, $c_{\mathbf{z}}$

for $q = 0, 1, 2, \dots$ **do**

  Set epoch length $T_q = \lfloor T_0 \gamma^q \rfloor$.

  Set exploration noise variance $\sigma_{\mathbf{z},q}^2 = \frac{\sigma_{\mathbf{w}}^2}{\sqrt{T_q}}$.

  Evolve the MJS for T_q steps with $\mathbf{u}_t^{(q)} = \mathbf{K}_{\omega^{(q)}(t)}^{(q)} \mathbf{x}_t^{(q)} + \mathbf{z}_t^{(q)}$ with $\mathbf{z}_t^{(q)} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{z},q}^2 \mathbf{I}_p)$ and record the trajectory $\{\mathbf{x}_t^{(q)}, \mathbf{z}_t^{(q)}, \omega^{(q)}(t)\}_{t=0}^{T_q}$.

  $\mathbf{A}_{1:s}^{(q)}, \mathbf{B}_{1:s}^{(q)}, \mathbf{T}^{(q)} =$ MJS Identification Algorithm 1$(\mathbf{K}_{1:s}^{(q)}, \sigma_{\mathbf{w}}^2, \sigma_{\mathbf{z},q}^2, \{\mathbf{x}_t^{(q)}, \mathbf{z}_t^{(q)}, \omega^{(q)}(t)\}_{t=0}^{T_q}, c_{\mathbf{x}}, c_{\mathbf{z}})$.

  Set the controller $\mathbf{K}_{1:s}^{(q+1)}$ for the next epoch to be the optimal controller for the infinite-horizon MJS-LQR$(\mathbf{A}_{1:s}^{(q)}, \mathbf{B}_{1:s}^{(q)}, \mathbf{T}^{(q)}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$.

end

guarantees we have more accurate MJS estimates thus more optimal controllers; (ii) as the controller becomes more optimal we can gradually decrease the exploration noise and deploy (exploit) the controller for a longer time. Note that the scheduling rate γ has a similar role to the discount factor in reinforcement learning: smaller γ aims to reduce short-term cost while larger γ aims to reduce long-term cost.

5.4 Theory

In this chapter, the main assumption is as follows, which guarantees the existence of optimal solution to the MJS-LQR problem according to Lemma 2.22.

Assumption A5.1. *The* MJS-LQR$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ *problem satisfies the following.*

(a) For all $i \in [s]$, $\mathbf{Q}_i \succ 0$ *and* $\mathbf{R}_i \succ 0$*;*

(b) The MJS $\Sigma = \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *is stabilizable.*

(c) The Markov matrix $\mathbf{T}$ *is ergodic.*

We define filtration $\mathcal{F}_{-1}, \mathcal{F}_0, \mathcal{F}_1, \dots$ such that $\mathcal{F}_{-1} := \sigma(\mathbf{x}_0, \omega_0)$ is the sigma-algebra generated by the initial state and initial mode, and $\mathcal{F}_q := \sigma(\mathbf{x}_0, \omega_0, \{\{\omega^{(j)}(t)\}_{t=1}^{T_j}\}_{j=0}^{q}, \mathbf{w}_0, \{\{\mathbf{w}_t^{(j)}\}_{t=1}^{T_j}\}_{j=0}^{q}, \mathbf{z}_0, \{\{\mathbf{z}_t^{(j)}\}_{t=1}^{T_j}\}_{j=0}^{q})$ is the sigma-algebra generated by the randomness up to epoch q. Note that the initial state $\mathbf{x}_0^{(q)}$ of epoch q is also the final state $\mathbf{x}_{T_{q-1}}^{(q-1)}$ of epoch $q-1$, therefore, $\mathbf{x}_0^{(q)}$ is $\mathcal{F}_{q-1}$-measurable, and so is $\omega^{(q)}(0)$. Suppose time step t belongs to epoch q, then we define the following conditional expected cost at time t as,

$$c_t = \mathbb{E}[\mathbf{x}_t^\intercal \mathbf{Q}_{\omega_t} \mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega_t} \mathbf{u}_t \mid \mathcal{F}_{q-1}]. \tag{5.1}$$

The cost for epoch q is defined as $J_{(q)} := \sum_{t \in \text{ epoch } q} c_t$, and the cumulative cost is defined as $J_T := \sum_q J_{(q)}$. We define the total regret and epoch-q regret as,

$$\text{Regret}(T) := J_T - TJ^\star, \quad \text{Regret}^{(q)} := J_{(q)} - T_q J^\star. \tag{5.2}$$

Then, we have, $\text{Regret}(T) = \mathcal{O}(\sum_{q=1}^{\mathcal{O}(\log_\gamma(T/T_0))} \text{Regret}^{(q)})$, where regret of epoch 0 is ignored as it does not scale with time T. In the definition of the regret, we evaluate the expected cost conditioning on the randomness up to the previous epoch. This is the middle ground between the expected cost $\mathbb{E}[\sum_t \mathbf{x}_t^\intercal \mathbf{Q}_{\omega_t} \mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega_t} \mathbf{u}_t]$ (Cassel et al., 2020) and random cost $\sum_t \mathbf{x}_t^\intercal \mathbf{Q}_{\omega_t} \mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega_t} \mathbf{u}_t$ (Lale et al., 2020b) typically considered in previous online learning works. We show in the next subsection that under certain stronger stability of the MJS, regret based on the random cost can also be bounded.

Let $\mathbf{K}_{1:s}^\star$ denote the optimal controller for the infinite-horizon MJS-LQR($\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$) problem. $\boldsymbol{\mathcal{L}}^{(0)}$ and $\boldsymbol{\mathcal{L}}^\star$ denote the augmented closed-loop state matrices (2.10) under the initial controller $\mathbf{K}_{1:s}^{(0)}$ and the optimal controller $\mathbf{K}_{1:s}^\star$ respectively. For a free parameter $\bar{\rho} \in [\max\{\rho(\boldsymbol{\mathcal{L}}^{(0)}), \rho(\boldsymbol{\mathcal{L}}^\star)\}, 1)$, let $\bar{\tau} := \max\{\tau(\boldsymbol{\mathcal{L}}^{(0)}, \bar{\rho}), \tau(\boldsymbol{\mathcal{L}}^\star, \bar{\rho})\}$, where $\tau(\cdot, \cdot)$ is the normalized power supremum in Definition 2.1. With these definitions, we have the following sublinear regret guarantee. Please refer to Theorem C.11 in Appendix C for the complete version and proof.

Theorem 5.1 (sublinear regret)**.** *Assume that the initial state* $\mathbf{x}_0 = 0$*, and Assumption A5.1 holds. In Algorithm 2, suppose hyper-parameters* $c_\mathbf{x} = \mathcal{O}(\sqrt{n})$*,* $c_\mathbf{z} = \mathcal{O}(\sqrt{p})$*, and* $T_0 \geq \hat{\mathcal{O}}\big(\frac{\sqrt{s} t_{\text{MC}} \log^2(T_0)}{\pi_{\min}(1-\bar{\rho})}(n+p)\big)$*. Then, with probability at least* $1-\delta$*, Algorithm 2 achieves*

$$\text{Regret}(T) \leq \hat{\mathcal{O}}\left(\frac{s^2 p (n^2+p^2) \bar{\tau}^2 \sigma_\mathbf{w}^2}{\pi_{\min}^2 (1-\bar{\rho})^2} \log^2(T) \sqrt{T}\right) + \mathcal{O}\left(\frac{\sqrt{ns} \log^3(T)}{\delta}\right). \tag{5.3}$$

Proof Sketch. For simplicity, we only show the dominant $\hat{\mathcal{O}}(\cdot)$ term here and leave the complete proof to appendix. Define the estimation error after epoch q as

$$\epsilon_{\mathbf{A},\mathbf{B}}^{(q)} := \max_{j \in [s]} \max\{\|\mathbf{A}_j^{(q)} - \mathbf{A}_j\|, \|\mathbf{B}_j^{(q)} - \mathbf{B}_j\|\}, \quad \epsilon_\mathbf{T}^{(q)} := \|\mathbf{T}^{(q)} - \mathbf{T}\|_\infty.$$

Using Theorem 4.2, we can bound epoch-q regret as

$$\text{Regret}^{(q)} \leq \mathcal{O}\left(T_q \sigma_{\mathbf{z},q}^2 + T_q \sigma_{\mathbf{w}}^2 \left(\epsilon_{\mathbf{A},\mathbf{B}}^{(q-1)} + \epsilon_{\mathbf{T}}^{(q-1)}\right)^2\right). \tag{5.4}$$

Plugging in the exploration noise variance $\sigma_{\mathbf{z},q}^2 = \frac{\sigma_{\mathbf{w}}^2}{\sqrt{T_q}}$, the upper bounds on the estimation errors $\epsilon_{\mathbf{A},\mathbf{B}}^{(q)} \leq \hat{\mathcal{O}}\left(\frac{\sigma_{\mathbf{z},q}+\sigma_{\mathbf{w}}}{\sigma_{\mathbf{z},q}\pi_{\min}} \frac{\sqrt{s}(n+p)\bar{\tau}\log(T_q)}{(1-\bar{\rho})\sqrt{T_q}}\right)$ and $\epsilon_{\mathbf{T}}^{(q)} \leq \hat{\mathcal{O}}\left(\sqrt{\frac{\log(T_q)}{T_q}}\right)$ from Theorem 3.1, we have $\text{Regret}^{(q)} \leq \hat{\mathcal{O}}\left(\frac{s^2 p(n^2+p^2)\bar{\tau}^2\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\gamma\sqrt{T_q}\log^2(T_q)\right)$. Finally, since $T_q = \mathcal{O}(T_0\gamma^q)$ from Algorithm 2, we have

$$\begin{aligned}
\text{Regret}(T) = & \sum_{q=1}^{\mathcal{O}(\log_\gamma(\frac{T}{T_0}))} \text{Regret}^{(q)}, \\
& \leq \hat{\mathcal{O}}\left(\frac{s^2 p(n^2+p^2)\bar{\tau}^2\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\sqrt{T}\log(\frac{T}{T_0})\left(\frac{\sqrt{\gamma}}{\sqrt{\gamma}-1}\right)^3\left(\gamma\log(\frac{T}{T_0}) - \sqrt{\gamma}\log(\gamma)\right)\right), \\
& \leq \hat{\mathcal{O}}\left(\frac{s^2 p(n^2+p^2)\bar{\tau}^2\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\text{polylog}(T)\sqrt{T}\right).
\end{aligned} \tag{5.5}$$

□

Note that the state dimension n, the input dimension p, number of modes s, and spectral radius $\bar{\rho}$ affect the regret bound in Theorem 5.1 and the identification error bound in Theorem 3.1 in a similar way. The factor that exclusively affects the regret bound is the epoch incremental ratio γ. One can see the interplay between T and γ from the term $\left(\frac{\sqrt{\gamma}}{\sqrt{\gamma}-1}\right)^3\left(\gamma\log(\frac{T}{T_0}) - \sqrt{\gamma}\log(\gamma)\right)$ in the proof sketch. Specifically, when horizon T is smaller, a smaller γ minimizes the upper bound, and vice versa. This further provides a mathematical justification for γ being similar to the discount factor in reinforcement learning in early discussions. In our regret upper bound, there is a heavy-tailed probability term $1/\delta$. In the next subsection, we discuss how this term is unavoidable under mean-square stability, but can be improved to sub-exponential tail term $\log(1/\delta)$ when stronger stability exists.

5.4.1 Two Special Cases

5.4.1.1 Regret under uniform stability

Note that the second term in the regret upper bound (5.3) in Theorem 5.1 depends on the failure probability δ through $1/\delta$. Though this term has a much milder dependency on the time horizon T, when setting δ to be small, it can still easily outweigh the $\hat{\mathcal{O}}(\cdot)$ term in

(5.3), which only has sub-exponential tail $\log(1/\delta)$ dependency, and can result in overly pessimistic regret bounds. The main cause of this $1/\delta$ term is that, in the regret analysis, one needs to factor in the cumulative impact of initial state of every epoch, i.e. $\sum_q \|\mathbf{x}_0^{(q)}\|^2$. Since mean-square stability guarantees the stability and state convergence only in the mean-square sense, we can, at best, only bound $\mathbb{E}[\|\mathbf{x}_0^{(q)}\|^2]$ and then use the Markov inequality: with probability at least $1-\delta$, $\|\mathbf{x}_0^{(q)}\|^2 \leq \mathbb{E}[\|\mathbf{x}_0^{(q)}\|^2]/\delta$. Furthermore, in Appendix C.5.1, we construct an MJS example that is mean-square stable, but dependency no better than $1/\delta$ is possible. Fortunately, there exists an easy workaround to get rid of this $1/\delta$ dependency if the MJS is uniformly stable (Definition 2.16), which enforces stability under arbitrary switching sequences, thus is stronger than mean-square stability. It allows us to bound $\mathbf{x}_0^{(q)}$ using tail inequalities much tighter than the Markov inequality and obtain $\|\mathbf{x}_0^{(q)}\|^2 \leq \mathcal{O}(\log(1/\delta))$ with probability at least $1-\delta$. As a result, the $1/\delta$ dependency in the regret bound can be improved to $\log(1/\delta)$.

One type of uniform stability assumption that can help us in this case is related to the closed-loop MJS under the optimal controllers. We let $\mathbf{K}_{1:s}^\star$ denote the optimal controller for the infinite-horizon MJS-LQR$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ and define closed-loop state matrices $\mathbf{L}_i^\star = \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i^\star$ for all $i \in [s]$. We let $\theta^\star$ denote the joint spectral radius of $\mathbf{L}_{1:s}^\star$, i.e. $\theta^\star := \lim_{l\to\infty} \max_{\omega_{1:l}\in[s]^l} \|\mathbf{L}_{\omega_1}^\star \cdots \mathbf{L}_{\omega_l}^\star\|^{\frac{1}{l}}$, and we say $\mathbf{L}_{1:s}^\star$ is uniformly stable if and only if $\theta^\star < 1$. Let $\bar{\theta} := \frac{1+\theta^\star}{2}$. The resulting regret bound is outlined in the following theorem, with its complete version and proof provided in Theorem C.12 of Appendix C.5.1.

Theorem 5.2 (Regret under uniform stability)**.** *Assume that the initial state* $\mathbf{x}_0 = 0$*, Assumption A5.1 holds, and* $\mathbf{L}_{1:s}^\star$ *is* uniformly stable*. If hyper-parameters* T_0*,* $c_\mathbf{x}$*, and* $c_\mathbf{z}$ *are chosen as sufficiently large, with probability at least* $1-\delta$*, Algorithm 2 achieves*

$$\text{Regret}(T) \leq \hat{\mathcal{O}}\left(\frac{s^2 p(n^2+p^2)\bar{\tau}^2\sigma_\mathbf{w}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\log^2(T)\sqrt{T}\right). \tag{5.6}$$

Another benefit of assuming uniform stability is that we can establish a sublinear bound for the regret defined using the random cost. Denote the random cost at time t as c_t°, the random cost for epoch q as $J_{(q)}^\circ$, and random regret as $\text{Regret}^\circ(T)$, defined as follows:

$$c_t^\circ := \mathbf{x}_t^\intercal \mathbf{Q}_{\omega_t}\mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega_t}\mathbf{u}_t, \quad J_{(q)}^\circ := \sum_{t\,\in\,\text{epoch}\,q} c_t^\circ, \quad \text{Regret}^\circ(T) := \sum_q J_{(q)}^\circ - TJ^\star. \tag{5.7}$$

Since we already have an upper bound for the $\text{Regret}(T) = \sum_q J_{(q)} - TJ^\star$ in Theorem 5.2, it suffices to upper bound $\sum_q J_{(q)}^\circ - J_{(q)}$ to establish an upper bound for the $\text{Regret}^\circ(T)$. In each summand $J_{(q)}^\circ - J_{(q)}$, we see $J_{(q)} = \mathbb{E}[J_{(q)}^\circ \mid \mathcal{F}_{q-1}]$ where $\mathcal{F}_{q-1}$ affects the expectation only through the initial state $\mathbf{x}_0^{(q)}$, initial mode $\omega^{(q)}(0)$, and the controller $\mathbf{K}_{1:s}^{(q)}$. Thus, the summand $J^\circ_{(q)} - J_{(q)}$ measures the deviation of the epoch's random cost $J^\circ_{(q)}$ from its conditional expectation with given initial conditions and controllers. Under the uniform stability assumption, we can show that $J^\circ_{(q)}$ is sub-exponential, which allows us to obtain $J^\circ_{(q)} - J_{(q)} \leq \mathcal{O}(\sqrt{T_q}\log(1/\delta))$ and $\text{Regret}^\circ(T) \leq \mathcal{O}(\sqrt{T}\log(1/\delta))$. On the other hand, in the case of mean-square stability, for similar reasons we discussed above, $J^\circ_{(q)}$ can be heavy-tailed, and the dependency on δ can at best be $1/\delta$. The formal result is provided below and the proof is provided in Appendix C.5.2

Theorem 5.3 (Random regret under uniform stability)**.** *Under the same setup of Theorem 5.2, with probability at least* $1-\delta$, *Algorithm 2 achieves*

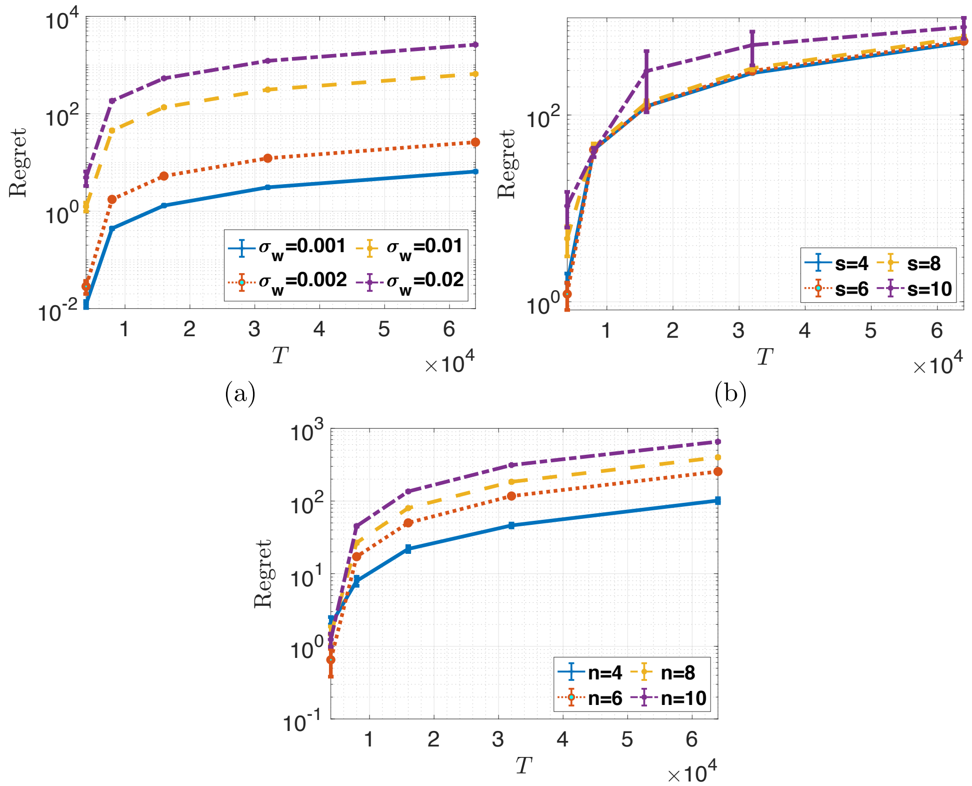
$$\text{Regret}^\circ(T) \leq \hat{\mathcal{O}}\left(\frac{s^2 p(n^2+p^2)\bar{\tau}^2\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\log^2(T)\sqrt{T} + \frac{(np)^{1.5}}{(1-\bar{\theta})^2}\sigma_{\mathbf{w}}^2\sqrt{T}\right). \tag{5.8}$$

5.4.1.2 Partial knowledge of dynamics

In practice, the input matrices $\mathbf{B}_{1:s}$ correspond to the actuators. One may have their knowledge either from the manufacturers or through various estimation techniques designed for non-dynamical models. From Corollary 3.2, we know that when $\mathbf{B}_{1:s}$ is known, no further exploration noise is needed to identify the state matrices $\mathbf{A}_{1:s}$ or Markov transition matrix $\mathbf{T}$. This can also be applied to the adaptive MJS-LQR setting, and the resulting regret bound can improve (from $\hat{\mathcal{O}}(\log^2(T)\sqrt{T})$ to $\hat{\mathcal{O}}(\log^3(T))$), since exploration noise incurs additional costs. The result is given in the following corollary, and we omit the proof due to its similarity to the proofs of Theorems 5.1 and 5.2.

Corollary 5.4 (Poly-logarithmic regret)**.** *When* $\mathbf{B}_{1:s}$ *is known, it suffices to set the exploration noise to be* $\sigma_{\mathbf{z},q} = 0$ *for all* q *in Algorithm 2. Then, the regret bound in Theorem 5.1 becomes,* $\text{Regret}(T) \leq \hat{\mathcal{O}}\left(\frac{s^2p(n^2+p^2)\bar{\tau}^2\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\log^3(T)\right) + \mathcal{O}\left(\frac{\sqrt{ns}\log^3(T)}{\delta}\right)$. *Additionally, the regret bound in Theorem 5.2 becomes,* $\text{Regret}(T) \leq \hat{\mathcal{O}}\left(\frac{s^2p(n^2+p^2)\bar{\tau}^2\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\log^3(T)\right)$.

As for the other special case when $\mathbf{A}_{1:s}$ is known but $\mathbf{B}_{1:s}$ is unknown, the exploration noise is still needed. One can analyze it as a special case of the general case when neither of them is known. For LTI systems, under certain strong assumptions, e.g. controller non-degeneracy, it is shown that poly-logarithmic regret is attainable for this case (Cassel et al., 2020). We speculate similar assumptions can lead to poly-logarithmic regret for MJS as well and leave this to the future work.



(a) (b) (c)

Figure 5.1: Performance of adaptive MJS-LQR. Influence of: (a) process noise $\sigma_{\mathbf{w}}$, (b) number of modes s, (c) state dimension n.

5.5 Experiments

We provide experiments to investigate the efficiency and verify the theory of the proposed algorithms on synthetic datasets. Throughout, we show results from a synthetic experiment where entries of the true system matrices $\mathbf{A}_{1:s}, \mathbf{B}_{1:s}$ are generated randomly from a standard normal distribution. We further scale each $\mathbf{A}_i$ to have $\|\mathbf{A}_i\| \leq 0.5$. Since this guarantees the MJS itself is mean-square stable, as we discussed in Section 2.3.1.1, we set the initial stabilizing controller $\mathbf{K}_{1:s}^{(0)} = 0$ in adaptive MJS-LQR Algorithm 2. For the cost matrices $(\mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$, we set $\mathbf{Q}_i = \underline{\mathbf{Q}}_i \underline{\mathbf{Q}}_i^\intercal$, and $\mathbf{R}_i = \underline{\mathbf{R}}_i \underline{\mathbf{R}}_i^\intercal$ where $\underline{\mathbf{Q}}_i \in \mathbb{R}^{n \times n}$ and $\underline{\mathbf{R}}_i \in \mathbb{R}^{p \times p}$ are generated from a standard normal distribution. The Markov matrix $\mathbf{T} \in \mathbb{R}_{\Delta}^{s \times s}$ is sampled from a Dirichlet distribution $\texttt{Dir}((s-1) \cdot \mathbf{I}_s + 1)$, where $\mathbf{I}_s$ denotes the identity matrix. We assume that we have equal probability of starting in any initial mode. The depicted results are averaged over 10 independent Monte Carlo runs. In our experiments, we explore the sensitivity of the regret bounds to the system parameters. We set the initial epoch length $T_0 = 2000$ and incremental ratio $\gamma = 2$. We select five epochs to run Algorithm 2. As an intermediate step for computing controller $\mathbf{K}_{1:s}^{(q+1)}$ in Algorithm 2, the coupled Riccati equations (2.35) are solved via value iteration, and the iteration stops when the parameter variation between two iterations falls below 10^{-6}, or iteration number reaches 10^4.

Fig. 5.1 demonstrates how regret bounds vary with (a) $\sigma_{\mathbf{w}} \in \{0.001, 0.002, 0.01, 0.02\}$, $n = 10$, $p = s = 5$; (b) $\sigma_{\mathbf{w}} = 0.01$, $n = 10$, $p = 5$, $s \in \{4, 6, 8, 10\}$, and (c) $\sigma_{\mathbf{w}} = 0.01$, $s = 10$, $p = 5$, $n \in \{4, 6, 8, 10\}$. We see that the regret degrades as $\sigma_{\mathbf{w}}, n$, and s increase. We also see that when $\sigma_{\mathbf{w}}$ is large (T is small), the regret becomes worse quickly as n and s grow larger. These results are consistent with the theoretical bounds in Theorem 5.1.

5.6 Conclusion

This chapter considers solving the optimal quadratic control problem for MJSs with unknown dynamics. We propose a model-based adaptive control scheme that alternates between system identification (Chapter 3) and certainty equivalent control (Chapter 4) on an epoch-by-epoch basis. Sublinear regret guarantees are established for the proposed scheme, which can be improved to poly-logarithmic regret when partial knowledge of the MJS is available. As a future work, it would be interesting and of practical importance to investigate the case when mode is not observed, which makes both system identification and adaptive quadratic control problems non-trivial.

Chapter 6
Mode Reduction

Switched systems are capable of modeling processes with underlying dynamics that may change abruptly over time. To achieve accurate modeling in practice, one may need a large number of modes, but this may in turn increase the model complexity drastically. Existing work on reducing system complexity mainly considers state space reduction, whereas reducing the number of modes is less studied. In this chapter, we consider the mode complexity for MJSs. Specifically, inspired by clustering techniques from unsupervised learning, we are able to construct a reduced MJS with fewer modes that approximates the original MJS well under various metrics. Furthermore, both theoretically and empirically, we show how one can use the reduced MJS to analyze stability and design controllers with significant reduction in computational cost while achieving guaranteed accuracy.

This chapter is organized as follows: we present the preliminaries and mode reduction problem setup in Section 6.3; an clustering-based reduction approach is proposed in Section 6.4; in Section 6.5, we discuss how the reduced MJS approximates the original one under setups; Section 6.6 and 6.7 respectively show that one can use the reduced MJS as a surrogate to evaluate stability and design LQR controllers for the original MJS; simulation experiments are presented in Section 6.8.

6.1 Introduction and Related Work

Switched systems generalize time-invariant systems by allowing the dynamics to switch over time. However, this is accompanied by new complexity challenges: the number of modes that is needed to model systems accurately and thoroughly may grow undesirably large. For example, for controlled plants composed of multiple components, if we model each combination of health statuses, e.g. working and faulty, of all components as a mode, then the number of modes grows exponentially with the number of components. Given this rate, there can

*The contents of this chapter are published in Du et al. (2022b) and Du et al. (2022a). The proofs in this chapter are provided in Appendix D.

be a huge amount of modes even with a moderate number of components. Given a system with this many modes, analysis can become computationally intractable. For example, in finite-horizon LQR problems with horizon T, the total number of controllers to be computed is s^T where s denotes the number of modes. This lack of scalability calls for systematic and theoretically guaranteed ways to reduce the number of modes.

Existing work on (switched) system reduction mainly focuses on reducing the state dimension (Zhang et al., 2003) or constructing finite abstractions for the continuous state space (Zamani and Abate, 2014). Reducing the mode complexity, however, is still mainly an uncharted territory.

In this chapter, we study how one can perform mode reduction for MJSs. Our main contributions are the following:

- We propose a clustering-based method that takes mode dynamics as features and use the estimated clusters to construct a reduced-mode MJS.
- The reduced MJS is proved to well approximate the original MJS under several approximation metrics.
- We show the reduced MJS can be used as a surrogate for the original MJS to analyze stability and design controllers with guaranteed performance and significant reduction in computational cost.

Our work adds a new dimension, i.e., reduction of modes, to the research of switched system reduction. This framework can be generalized to other problems such as robust and optimal control, invariance analysis, partially observed systems, etc. Other than constructing and analyzing the reduced MJS, the technical tools we develop in this chapter regarding perturbations can be applied to cases when there are model mismatches.

6.2 Related Work

Depending on the problems of interest and methodologies, the work on reduction for stochastic (switched) systems can be roughly divided into three categories: bisimulation, symbolic abstraction, and order reduction.

Bisimulation: To evaluate the equivalency between two stochastic switched systems, notions of (approximate) probabilistic bisimulation are proposed in Larsen and Skou (1991); Desharnais et al. (2002, 2004). Approximation metrics (Abate, 2013) from different perspectives are developed to compare two systems, e.g. one(multi)-step transition kernels (Abate et al., 2011) and trajectories (Girard and Pappas, 2007; Tkachev and Abate, 2014; Julius and

Table 6.1: Related Work on Reduction for Stochastic Switched Systems

<table>
<tr><th>Reference</th><th>Model</th><th>Reduction Target</th><th>Exact Bisimulation Condition</th><th>Reduction Method</th><th>Approximation Metric</th></tr>
<tr><td>(Desharnais et al., 2004)</td><td>Controlled Markov Process</td><td rowspan="7">State Space Cardinality</td><td rowspan="5">Yes</td><td rowspan="2">N.A.</td><td>Formula Metric</td></tr>
<tr><td>(Tkachev and Abate, 2014)
(Bian and Abate, 2017)</td><td rowspan="2">Labelled Markov Chains (autonomous)</td><td rowspan="2">Trajectory</td></tr>
<tr><td>(Lun et al., 2018)</td><td rowspan="2">Clustering</td></tr>
<tr><td>(Zhang and Wang, 2019)
(Du et al., 2019a)
(Bittracher and Schütte, 2021)</td><td>Markov Chains</td><td>N.A.</td></tr>
<tr><td>(Zamani and Abate, 2014)</td><td>Switching Stochastic Sys.</td><td rowspan="3">State Space Discretization</td><td>Trajectory</td></tr>
<tr><td>(Abate et al., 2011)</td><td rowspan="2">Stochastic Hybrid System (autonomous)</td><td rowspan="6">N.A.</td><td>Transition Kernel</td></tr>
<tr><td>(Soudjani and Abate, 2011)</td><td>Invariance Probability</td></tr>
<tr><td>(Julius et al., 2006)
(Julius and Pappas, 2009)
(Zamani et al., 2016)</td><td>Jump Linear Stochastic System</td><td rowspan="4">State Space Dimension (Order)</td><td>(Bi)simulation Function</td><td>Trajectory</td></tr>
<tr><td>(Zhang et al., 2003)
(Shen et al., 2019)</td><td rowspan="3">MJS</td><td>$\mathcal{H}_\infty$ Reduction</td><td rowspan="2">$\mathcal{H}_\infty$ norm</td></tr>
<tr><td>(Kotsalis and Rantzer, 2010)</td><td>Balanced Truncation</td></tr>
<tr><td>(Sun and Lam, 2016)</td><td>$\mathcal{H}_2$ Reduction</td><td>$\mathcal{H}_2$ norm</td></tr>
<tr><td>This Chapter</td><td>MJS</td><td>Modes</td><td>Yes</td><td>Clustering</td><td>Trajectory</td></tr>
</table>

Pappas, 2009). Based on the approximate bisimulation notion in Bian and Abate (2017), a technique for reducing the state space of labeled Markov chains through state aggregation is proposed in Lun et al. (2018). Unlike existing work that typically defines the notions of (approximate) bisimulation on the state space, we provide an algorithm that constructs a reduced system by aggregating the mode space, which provably approximates the original one. Our work shares the idea of aggregation of Markov chains with Lun et al. (2018), but we also seek to recover the best aggregation partition which is otherwise assumed as prior knowledge in Lun et al. (2018).

Symbolic Abstraction: Given a system with continuous state space, abstraction (Alur et al., 2000) considers discretizing the state space and then constructing a finite state symbolic model, which can be used as a surrogate for model verification (Clarke Jr. et al., 2018; Kurshan, 2014) or controller synthesis (Maler et al., 1995). The work on abstraction for stochastic hybrid systems starts with the autonomous cases. Under uniform discretization, Abate et al. (2010) and Abate et al. (2011) provide approximation guarantees that depend on the discretization width. An adaptive partition scheme is proposed in Soudjani and Abate (2011), which mitigates the curse of dimensionality suffered by uniform sampling. Since the systems under consideration are autonomous, these works mainly serve verification purposes, but fall short toward controller synthesis goals. Zamani and Abate (2014) addressed this by allowing inputs in the systems. The idea of partitioning the continuous state space is similar to our work except that our partition is performed on the mode space, i.e., the discrete state space in hybrid systems, which provides a new yet closely related dimension to the existing abstraction work.

Order Reduction: Another important line of research on system reduction is order reduction (Gugercin and Antoulas, 2004), where one seeks to reduce the dimension of the state space to satisfy certain criteria. With the help of linear matrix inequalities (LMIs), various methods have been applied for MJS, including $\mathcal{H}_\infty$ reduction (Zhang et al., 2003), balanced truncation (Kotsalis and Rantzer, 2010), and $\mathcal{H}_2$ reduction (Sun and Lam, 2016), etc. Order reduction is also applied to more complex models with time-varying transition probabilities (Shen et al., 2019).

The reduction of Markov chains, a class of simplified yet fundamental stochastic switched models, has also attracted the learning and statistics communities. Several notions of lumpability are proposed in Buchholz (1994), which coincide with the notion of bisimulation in Larsen and Skou (1991). Lumpability allows one to reduce the original Markov chain to a smaller scale yet equivalent Markov chain by lumping the Markovian states. Similar research focusing on the equivalence metrics and bounding the difference of transition kernel, can be found in Hoffmann and Salamon (2009); Schulman and Gaveau (2001); Gaveau and Schulman (2005) under the name of coarse graining. Compared with the bisimulation work for general stochastic systems, which is mostly conceptual, the restriction to Markov chains allows for practical "low-rank + clustering" methods (Meilă and Shi, 2001) to uncover the lumpability structure. Zhang and Wang (2019) considers the case when the Markov matrix is estimated from a trajectory, and the approximate lumpability case is studied in Du et al. (2019a); Bittracher and Schütte (2021). Furthermore Du et al. (2019a) studies the reduction of Markov chains that are embedded in switched autoregressive exogenous models, but the overall dynamical models are not reduced. Based on the ideas in Du et al. (2019a), our work further extends the reduction to the overall MJS.

A comprehensive comparison of the related work together with our work is listed in Table 6.1. The entry "exact bisimulation condition" tells whether ideal case sufficient conditions are provided under which a system can be reduced without introducing any model inaccuracy, i.e., they are bisimilar. In practice, when the reduced system is constructed, these principled conditions can help gain more insight into the original system. In practice, these model reduction methods developed under different perspectives can be combined to achieve overall better performance. For example, for the continuous state space, one can apply order reduction followed by finite abstraction (the former can help remove the curse of dimensionality for the latter), and meanwhile our work can further help reduce the discrete mode space.

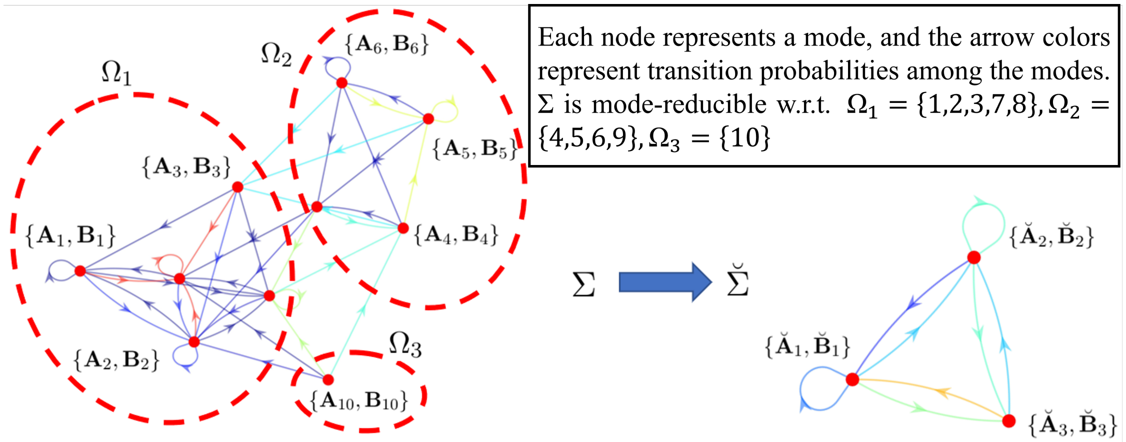


Figure 6.1: Illustration of MJS mode reduction

6.3 Problem Formulation

In this chapter, we consider the MJS Σ in (2.6) but without process noise. In other words, Σ is given by

$$\Sigma := \begin{cases} \mathbf{x}_{t+1} & = \mathbf{A}_{\omega_t}\mathbf{x}_t + \mathbf{B}_{\omega_t}\mathbf{u}_t \\ \omega_t & \sim \text{Markov Chain}(\mathbf{T}) \end{cases} \tag{6.1}$$

We assume the Markov matrix $\mathbf{T}$ is ergodic. In the remainder of this chapter, we use $\Sigma := \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ to denote the groundtruth MJS in (6.1) that we want to study, and similarly use notation $\text{MJS}(\cdot,\cdot,\cdot)$ to parameterize any noiseless MJS with expressions given in (6.1).

This chapter will consider Markov chains for the Markov matrix $\mathbf{T}$ with the following properties: lumpability and aggregatability (Definition 2.7). Lumpability of a Markov chain coincides with the definition of probabilistic bisimulation in Desharnais et al. (2002), which describes an equivalence relation on the Markovian state space $[s]$, i.e., two members are equivalent if they belong to the same cluster. For a Markov chain $\mathbf{T}$ that is lumpable with respect to partition $\Omega_{1:s}$, we use $\zeta_t \in [r]$ to index the active cluster at time t, i.e., $\zeta_t = k$ if and only if $\omega_t \in \Omega_k$, and use $\zeta_{0:t}$ to denote the active cluster sequence. The implication of lumpability is that $\zeta_{0:t}$ also follows a Markov chain.

6.3.1 Problem Setup

With the notions of Markov chain lumpability and aggregatability, in this chapter, we consider reducing the number of modes for Σ under the following two problem settings.

Problem P6.1 (Lumpable Case). *Assume the dynamics of* $\Sigma = \text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *is* *known. Suppose there exists a hidden partition $\Omega_{1:r}$ on $[s]$ and $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}} \geq 0$ such that*

$$\begin{aligned} \sum_{k\in[r]} \sum_{i,i'\in\Omega_k} \|\mathbf{A}_i - \mathbf{A}_{i'}\|_{\mathrm{F}} \leq \epsilon_{\mathbf{A}}, \\ \sum_{k\in[r]} \sum_{i,i'\in\Omega_k} \|\mathbf{B}_i - \mathbf{B}_{i'}\|_{\mathrm{F}} \leq \epsilon_{\mathbf{B}}, \end{aligned} \tag{6.2}$$

$$\sum_{k,l\in[r]} \sum_{i,i'\in\Omega_k} \Big| \sum_{j\in\Omega_l} \mathbf{T}(i,j) - \sum_{j\in\Omega_l} \mathbf{T}(i',j) \Big| \leq \epsilon_{\mathbf{T}}. \tag{6.3}$$

Then, given $\{\mathbf{A}_i, \mathbf{B}_i\}_{i=1}^s, \mathbf{T}$ and r, we seek to estimate the partition $\Omega_{1:r}$ by clustering the modes, construct a reduced MJS for Σ, and provide guarantees on the behavior difference incurred by the reduction.

Throughout this chapter, we refer to $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}$, and $\epsilon_{\mathbf{T}}$ as *perturbations*. When $\epsilon_{\mathbf{T}} = 0$, by Definition 2.7, the Markov matrix $\mathbf{T}$ is lumpable with respect to the partition $\Omega_{1:r}$. Thus, in condition (6.3), one can view $\mathbf{T}$ as *approximately* lumpable. One can think of $\{\omega_t, \mathbf{x}_t\} \in [s]\times\mathbb{R}^n$ as a hybrid state (Abate et al., 2011). Then, condition (6.3) guarantees the existence of an approximate equivalence relation in the discrete domain $[s]$, while condition (6.2) guarantees this in the continuous domain $\mathbb{R}^n$. More discussions for the special case when $\epsilon_{\mathbf{A}} = \epsilon_{\mathbf{B}} = \epsilon_{\mathbf{T}} = 0$ follow in the next subsection.

For the aggregatable case, we separately formulate a similar problem in Problem P6.2.

Problem P6.2 (Aggregatable Case)**.** *In Problem P6.1, replace the condition in (6.3) with $\sum_{k\in[r]} \sum_{i,i'\in\Omega_k} \|\mathbf{T}(i,:)^\intercal - \mathbf{T}(i',:)^\intercal\|_1 \leq \epsilon_{\mathbf{T}}$.*

When $\epsilon_{\mathbf{T}} = 0$, by Definition 2.7, the Markov matrix $\mathbf{T}$ is aggregatable with respect to the partition $\Omega_{1:r}$. P6.2 differs from P6.1 in terms of $\epsilon_{\mathbf{T}}$, which quantifies the violation of lumpability and aggregatability respectively for the matrix $\mathbf{T}$. Consider a 3-state Markov matrix $\mathbf{T}$ with rows $\mathbf{T}(1,:) = [0.2, 0.4, 0.4]$, $\mathbf{T}(2,:) = [0.7, 0.1, 0.2]$, and $\mathbf{T}(3,:) = [0.7, 0, 0.3]$. Then for the partition $\Omega_1 = \{1\}, \Omega_2 = \{2,3\}$, we obtain $\epsilon_{\mathbf{T}} = 0$ in P6.1 but $\epsilon_{\mathbf{T}} = 0.4$ in P6.2. In other words, $\mathbf{T}$ is exactly lumpable but non-aggregatable with a violation level of 0.4. In P6.2, $\epsilon_{\mathbf{T}} = 0$ only when $\mathbf{T}(i,:) = \mathbf{T}(i',:)$ for all $i, i' \in \Omega_k$, i.e. the rows are equal. On the other hand, in P6.1, $\epsilon_{\mathbf{T}} = 0$ is possible even if no row equalities exist. Hence, being $\epsilon_{\mathbf{T}}$-aggregatable in P6.2 is a stronger assumption than being $\epsilon_{\mathbf{T}}$-lumpable in P6.1. As a result, in Section 6.4, the clustering guarantee for P6.2 is stronger and more interpretable than that of P6.1.

6.3.2 Equivalency between MJSs

To compare the original and reduced-mode MJSs as mentioned in P6.1, we need a notion of equivalency between two MJSs with different numbers of modes. This is provided below via a surjection from modes of the larger MJS to the smaller one, which extends the bijection idea in Julius and Pappas (2009); Zhang et al. (2003) that can only compare two MJSs with equal amounts of modes.

Definition 6.1 (Equivalency between MJSs)**.** *Consider two* MJSs Σ_1 *and* Σ_2 *with the same state and input dimensions* n, p, *but different number of modes* s_1 *and* s_2 *respectively. WLOG, assume* $s_1 > s_2$. *Let* $\{\mathbf{x}_t^{(1)}, \mathbf{u}_t^{(1)}, \omega_t^{(1)}\}$ *and* $\{\mathbf{x}_t^{(2)}, \mathbf{u}_t^{(2)}, \omega_t^{(2)}\}$ *denote their respective state, input, and mode index.* Σ_1 *and* Σ_2 *are equivalent if there exists a partition* $\Omega_{1:s_2}$ *on* $[s_1]$ *such that* Σ_1 *and* Σ_2 *have the same transition kernels, i.e. for any time* t, *any mode* $k, k' \in [s_2]$, *any* $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$, *and any* $\mathbf{u} \in \mathbb{R}^p$

$$\begin{aligned}\mathbb{P}\big(\omega_{t+1}^{(1)} \in \Omega_{k'}, \mathbf{x}_{t+1}^{(1)} = \mathbf{x}' \mid \ \omega_t^{(1)} \in \Omega_k, \mathbf{x}_t^{(1)} = \mathbf{x}, \mathbf{u}_t^{(1)} = \mathbf{u}\big) \\ = \mathbb{P}\big(\omega_{t+1}^{(2)} = k', \mathbf{x}_{t+1}^{(2)} = \mathbf{x}' \mid \ \omega_t^{(2)} = k, \mathbf{x}_t^{(2)} = \mathbf{x}, \mathbf{u}_t^{(2)} = \mathbf{u}\big). \quad (6.4)\end{aligned}$$

The trivial perturbation-free case, i.e., $\epsilon_\mathbf{A}, \epsilon_\mathbf{B}, \epsilon_\mathbf{T} = 0$, provides a sufficient condition that guarantees that an MJS can be reduced to a smaller MJS with equivalency between them.

Definition 6.2 (Mode-reducibility Condition)**.** *If in P6.1,* $\epsilon_\mathbf{A}, \epsilon_\mathbf{B}$, $\epsilon_\mathbf{T} = 0$, *we say* Σ *is mode-reducible with respect to* $\Omega_{1:r}$.

If this condition holds for Σ, we can construct a reduced-mode MJS $\breve{\Sigma} := \mathrm{MJS}(\breve{\mathbf{A}}_{1:r}, \breve{\mathbf{B}}_{1:r}, \breve{\mathbf{T}})$ such that for any $k, l \in [r]$, any $i \in \Omega_k$, $\breve{\mathbf{A}}_k = \mathbf{A}_i$, $\breve{\mathbf{B}}_k = \mathbf{B}_i$, and $\breve{\mathbf{T}} \in \mathbb{R}^{r\mathrm{x}r}$ with $\breve{\mathbf{T}}(k,l) = \sum_{j \in \Omega_l} \mathbf{T}(i,j)$, which is illustrated in Fig. 6.1. Let $\{\breve{\mathbf{x}}_t, \breve{\mathbf{u}}_t, \breve{\omega}_t\}$ denote the state, input, and mode index for the reduced $\breve{\Sigma}$. Then, the following fact shows that $\breve{\Sigma}$ and Σ are equivalent according to Definition 6.1.

Fact 6.3. *Suppose* Σ *is mode-reducible and* $\breve{\Sigma}$ *is constructed as above. Consider the case when* Σ *and* $\hat{\Sigma}$ *have (i) initial mode distributions satisfy* $\mathbb{P}(\omega_0 \in \Omega_k) = \mathbb{P}(\breve{\omega}_0 = k)$ *for all* $k \in [r]$, *(ii) the same initial states* ($\mathbf{x}_0 = \breve{\mathbf{x}}_0$), *and (iii) the same input sequences* ($\mathbf{u}_{0:t-1} = \breve{\mathbf{u}}_{0:t-1}$). *Then, these two MJSs have the same mode and state transition kernels, i.e.* $\mathbb{P}(\omega_t \in \Omega_k, \mathbf{x}_t = \mathbf{x}) = \mathbb{P}(\breve{\omega}_t = k, \breve{\mathbf{x}}_t = \mathbf{x})$ *for all* t, *all* $k \in [r]$ *and* $\mathbf{x} \in \mathbb{R}^n$. *Particularly, there exists a special type of reduced* $\breve{\Sigma}$ *such that the modes are synchronized: for all* t, $\breve{\omega}_t = \zeta_t$. *In this case,* $\breve{\mathbf{x}}_t = \mathbf{x}_t$ *for all* t.

Proof. For Σ, using its dynamics in (6.1), we obtain

$$\begin{aligned}\mathbf{x}_t &= \Big(\prod_{h=0}^{t-1} \mathbf{A}_{\omega_h}\Big)\mathbf{x}_0 + \sum_{t'=0}^{t-2}\Big(\prod_{h=t'+1}^{t-1} \mathbf{A}_{\omega_h}\Big)\mathbf{B}_{\omega_{t'}}\mathbf{u}_{t'} + \mathbf{B}_{\omega_{t-1}}\mathbf{u}_{t-1} \\ &= \Big(\prod_{h=0}^{t-1} \breve{\mathbf{A}}_{\zeta_h}\Big)\mathbf{x}_0 + \sum_{t'=0}^{t-2}\Big(\prod_{h=t'+1}^{t-1} \breve{\mathbf{A}}_{\zeta_h}\Big)\breve{\mathbf{B}}_{\zeta_{t'}}\mathbf{u}_{t'} + \breve{\mathbf{B}}_{\zeta_{t-1}}\mathbf{u}_{t-1}.\end{aligned} \tag{6.5}$$

Similarly, for the reduced $\breve{\Sigma}$, we have $\breve{\mathbf{x}}_t = \big(\prod_{h=0}^{t-1} \breve{\mathbf{A}}_{\breve{\omega}_h}\big)\breve{\mathbf{x}}_0 + \sum_{t'=0}^{t-2}\big(\prod_{h=t'+1}^{t-1} \breve{\mathbf{A}}_{\breve{\omega}_h}\big)\breve{\mathbf{B}}_{\breve{\omega}_{t'}}\breve{\mathbf{u}}_{t'} + \breve{\mathbf{B}}_{\breve{\omega}_{t-1}}\breve{\mathbf{u}}_{t-1}$. Note that when initial conditions $\mathbf{x}_0 = \breve{\mathbf{x}}_0$ and input sequence $\mathbf{u}_{0:t-1} = \breve{\mathbf{u}}_{0:t-1}$ are fixed, $\mathbf{x}_t$ depends on $\zeta_{0:t-1}$ in the same way as $\breve{\mathbf{x}}_t$ depends on $\breve{\omega}_{0:t-1}$. Thus to prove the claim $\mathbb{P}(\omega_t \in \Omega_k, \mathbf{x}_t = \mathbf{x}) = \mathbb{P}(\breve{\omega}_t = k, \breve{\mathbf{x}}_t = \mathbf{x})$, it suffices to show sequences $\zeta_{0:t-1}$ and $\breve{\omega}_{0:t-1}$ have the same distributions, i.e.

$$\mathbb{P}(\zeta_{0:t-1} = \sigma_{0:t-1}) = \mathbb{P}(\breve{\omega}_{0:t-1} = \sigma_{0:t-1}) \tag{6.6}$$

for any possible sequence $\sigma_{0:t-1} \in [r]^t$. We will show this through induction. Note that due to the assumption for the same initial conditions, we have $\mathbb{P}(\zeta_0 = \sigma_0) = \mathbb{P}(\breve{\omega}_0 = \sigma_0)$. Now, suppose $\mathbb{P}(\zeta_{0:h-1} = \sigma_{0:h-1}) = \mathbb{P}(\breve{\omega}_{0:h-1} = \sigma_{0:h-1})$, then to complete the induction argument, by Bayes' theorem, it suffices to show $\mathbb{P}(\zeta_h = \sigma_h \mid \zeta_{0:h-1} = \sigma_{0:h-1}) = \mathbb{P}(\breve{\omega}_h = \sigma_h \mid \breve{\omega}_{0:h-1} = \sigma_{0:h-1})$. For the LHS,

$$\begin{aligned}&\mathbb{P}(\zeta_h = \sigma_h \mid \zeta_{0:h-1} = \sigma_{0:h-1}) \\ =&\sum_{\omega \in \Omega_{\sigma_{h-1}}} \mathbb{P}(\omega_h \in \Omega_{\sigma_h}, \omega_{h-1} = \omega \mid \zeta_{0:h-1} = \sigma_{0:h-1}) \\ =&\sum_{\omega \in \Omega_{\sigma_{h-1}}} \mathbb{P}(\omega_h \in \Omega_{\sigma_h} \mid \omega_{h-1} = \omega)\mathbb{P}(\omega_{h-1} = \omega \mid \zeta_{0:h-1} = \sigma_{0:h-1}) \\ =&\breve{\mathbf{T}}(\sigma_{h-1}, \sigma_h) \sum_{\omega \in \Omega_{\sigma_{h-1}}} \mathbb{P}(\omega_{h-1} = \omega \mid \zeta_{0:h-1} = \sigma_{0:h-1}) \\ =&\breve{\mathbf{T}}(\sigma_{h-1}, \sigma_h).\end{aligned} \tag{6.7}$$

For the RHS, by the Markov property, we have $\mathbb{P}(\breve{\omega}_h = \sigma_h \mid \breve{\omega}_{0:h-1} = \sigma_{0:h-1}) = \mathbb{P}(\breve{\omega}_h = \sigma_h \mid \breve{\omega}_{h-1} = \sigma_{h-1}) = \breve{\mathbf{T}}(\sigma_{h-1}, \sigma_h)$. The induction and proof are complete. ☐

Fact 6.3 first shows the equivalency between Σ and $\breve{\Sigma}$ in terms of the transition kernels, which is then extended to trajectory realizations if certain synchrony exists between $\zeta_{0:t}$ and $\breve{\omega}_{0:t}$. The condition $\breve{\omega}_t = \zeta_t$ in Fact 6.3 essentially establishes a coupling between the Markov chains $\omega_{0:t}$ and $\breve{\omega}_{0:t}$ such that $\mathbb{P}(\omega_t \in \Omega_k, \breve{\omega}_t = k) = \mathbb{P}(\omega_t \in \Omega_k) = \mathbb{P}(\breve{\omega}_t = k)$.

Establishing coupling between the stochastic systems usually allows for stronger equivalency and approximation result. Similar coupling scheme is implicitly used in Julius and Pappas (2009); Zhang et al. (2003); an optimal coupling by minimizing Wasserstein distance is discussed in Tkachev and Abate (2014); and a weaker coupling using the idea of HMM is discussed in Shen et al. (2019).

In Definition 6.1 and Fact 6.3, viewing $\{\omega_t, \mathbf{x}_t\} \in [s] \times \mathbb{R}^n$ as a hybrid state, $\mathbf{T}$ being lumpable guarantees the existence of an equivalence relation in the discrete domain $[s]$ as in Definition 2.7, while state/input matrices being the same guarantees this in the continuous domain $\mathbb{R}^n$.

6.4 Clustering-based Mode Reduction

In this section, we first propose Algorithm 3 to estimate the latent partition $\Omega_{1:r}$ and construct the reduced MJS for Problem P6.1 and P6.2, and then provide its theoretical guarantees for partition estimation in Section 6.4.1.

Algorithm 3: Mode Reduction for MJS

Input: $\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \boldsymbol{\pi}$, r, and non-negative tuning weights $\alpha_{\mathbf{A}}, \alpha_{\mathbf{B}}, \alpha_{\mathbf{T}}$ that sum to 1

1 Construct feature matrix $\mathbf{\Phi}$: $\forall i \in [s]$,

2 **case** *Problem P6.2* **do**

3 $\quad \mathbf{\Phi}(i,:) = [\alpha_{\mathbf{A}} \mathbf{vec}(\mathbf{A}_i)^\intercal, \alpha_{\mathbf{B}} \mathbf{vec}(\mathbf{B}_i)^\intercal, \alpha_{\mathbf{T}} \mathbf{T}(i,:)]$.

4 **case** *Problem P6.1* **do**

5 $\quad \mathbf{H} = \mathbf{diag}(\boldsymbol{\pi})^{\frac{1}{2}} \mathbf{T} \mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}}$

6 $\quad \mathbf{W}_r \leftarrow$ top r left singular vectors of $\mathbf{H}$

7 $\quad \mathbf{S}_r = \mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}} \mathbf{W}_r$

8 $\quad \mathbf{\Phi}(i,:) = [\alpha_{\mathbf{A}} \mathbf{vec}(\mathbf{A}_i)^\intercal, \alpha_{\mathbf{B}} \mathbf{vec}(\mathbf{B}_i)^\intercal, \alpha_{\mathbf{T}} \mathbf{S}_r(i,:)]$

9 $\mathbf{U}_r \leftarrow$ top r left singular vectors of $\mathbf{\Phi}$

10 Solve k-means problem: $\hat{\Omega}_{1:r}, \hat{\mathbf{c}}_{1:r} = \arg\min_{\hat{\Omega}_{1:r}, \hat{\mathbf{c}}_{1:r}} \sum_{k \in [r]} \sum_{i \in \hat{\Omega}_k} \|\mathbf{U}_r(i,:) - \hat{\mathbf{c}}_k\|^2$

11 Construct $\hat{\Sigma}$, $\forall k, l \in [r]$ $\hat{\mathbf{A}}_k = \frac{1}{|\hat{\Omega}_k|} \sum_{i \in \hat{\Omega}_k} \mathbf{A}_i, \quad \hat{\mathbf{B}}_k = \frac{1}{|\hat{\Omega}_k|} \sum_{i \in \hat{\Omega}_k} \mathbf{B}_i$,
$\hat{\mathbf{T}}(k,l) = \frac{1}{|\hat{\Omega}_k|} \sum_{i \in \hat{\Omega}_k, j \in \hat{\Omega}_l} \mathbf{T}(i,j)$

Output: $\hat{\Sigma}$: MJS$(\hat{\mathbf{A}}_{1:r}, \hat{\mathbf{B}}_{1:r}, \hat{\mathbf{T}})$

We treat the estimation of partition $\Omega_{1:r}$ essentially as a mode clustering problem with the dynamics matrices $\mathbf{A}_i$, $\mathbf{B}_i$ and transition distribution $\mathbf{T}(i,:)$ serving as features for mode i. In Algorithm 3, we first construct the feature matrix $\mathbf{\Phi}$ from Line 2 to Line 8, with $\mathbf{\Phi}(i,:)$ denoting the features of mode i. For the aggregatable case in Problem P6.2, we simply stack the vectorized $\mathbf{A}_i$, $\mathbf{B}_i$ and $\mathbf{T}(i,:)$, and use $\alpha_{\mathbf{A}}, \alpha_{\mathbf{B}}, \alpha_{\mathbf{T}}$ to denote their weights respectively. One way to choose these weights is as a normalization, e.g. $\alpha_{\mathbf{A}} \propto 1/\max_i \|\mathbf{A}_i\|$, so that these three features would have the same scales. Though in the aggregatable case P6.2, similarities among the rows of $\mathbf{T}$ shed light on the groundtruth partition $\Omega_{1:r}$, this is no longer valid in the lumpable case P6.1 as two modes belonging to the same cluster can still have different transition probabilities $\mathbf{T}(i,:)$, even if $\epsilon_{\mathbf{T}} = 0$. According to (6.3), the groundtruth partition $\Omega_{1:r}$ is only embodied in the mode-to-cluster transition probabilities $\sum_{j \in \Omega_l} \mathbf{T}(i,j)$ constructed using the groundtruth partition itself. This leaves us in a "chicken-and-egg" dilemma. To deal with this, from Line 5 to 8, we compute the first r left singular vectors $\mathbf{W}_r$ of matrix $\mathbf{diag}(\boldsymbol{\pi})^{\frac{1}{2}}\mathbf{T}\mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}}$, and then weight it by $\mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}}$ to obtain matrix $\mathbf{S}_r \in \mathbb{R}^{s \mathrm{x} r}$, which is used to construct features in $\mathbf{\Phi}$ for the lumpable case P6.1. We will later justify using $\mathbf{S}_r$ as features by showing row similarities in $\mathbf{S}_r$ reflect the partition under certain assumptions.

With the feature matrix $\mathbf{\Phi}$, to recover the partition, we resort to k-means: in Line 10, k-means is applied to the first r left singular vector $\mathbf{U}_r$ of $\mathbf{\Phi}$. The typical algorithm for k-means is Lloyd's algorithm, where the cluster centers and partition membership are updated alternately. Based on the solution $\hat{\Omega}_{1:r}$ obtained via k-means, we construct the reduced $\hat{\Sigma}$ by averaging modes within the same estimated cluster. A more subtle averaging scheme is through the weights provided in the stationary distribution $\boldsymbol{\pi}$ which describes the frequency of each mode being active in the long run. $\hat{\Sigma}$ generated by this scheme (or any weighted averaging) would have the same performance guarantees as the uniform averaging, which is provided in Section 6.4.1.

In practice, if we have no good prior knowledge which model of Problem P6.1 and P6.2 would yield the best model reduction performance, we can first obtain the partitions for both cases and then pick the one that yields smaller hindsight perturbations $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}}$ in Problem P6.1 and P6.2. When one picks $\alpha_{\mathbf{A}} = \alpha_{\mathbf{B}} = 0$, i.e only the Markov matrix $\mathbf{T}$ is used to cluster the modes, then our clustering scheme under the aggregatable case P6.2 is equivalent to Zhang and Wang (2019) which studies clustering for Markov matrices that is estimated from a single trajectory. The lumpable case P6.1, on the other hand, is based on preliminary analysis in Meilă and Shi (2001).

We note that several aspects of this algorithm we have guarantees for do not directly consider the metrics important to this problem; for example averaging dynamics matrices within the same cluster may not yield the dynamics that gives an optimal fit for prediction or controller design. That said, even for this straightforward approach, Section 6.5 provides several strong approximation guarantees. We are hopeful that future generalizations will be able to build on this theory and further improve the control performance of our mode-reduction approach.

6.4.1 Theoretical Guarantees for Clustering

In this section, we discuss the clustering performance by comparing the estimated partition $\hat{\Omega}_{1:r}$ and the true $\Omega_{1:r}$. As k-means algorithms are known to have local convergence properties (Bottou and Bengio, 1994), we instead assume for the k-means problem in Algorithm 3, a $(1+\epsilon)$ approximate solution can be obtained, i.e., $\sum_{k\in[r],i\in\hat{\Omega}_k} \|\mathbf{U}_r(i,:) - \hat{\mathbf{c}}_k\|^2 \leq (1+\epsilon)\min_{\Omega'_{1:r},\mathbf{c}'_{1:r}} \sum_{k\in[r],i\in\Omega'_k} \|\mathbf{U}_r(i,:) - \mathbf{c}'_k\|^2$. Many efficient algorithms have been developed that can provide $(1+\epsilon)$ approximate solutions. For $\epsilon = 1$, a linear time (in terms of r and s) algorithm is provided in Gonzalez (1985). For smaller ϵ, Kumar et al. (2004) proposes a linear time algorithm using random sampling; Song and Rajasekaran (2010) gives a polynomial time algorithm with computational complexity independent of s. We later show how ϵ affects the overall clustering performance.

To evaluate the performance of partition estimation, we define misclustering rate (MR) as $\mathrm{MR}(\hat{\Omega}_{1:r}) = \min_{h\in\mathcal{H}} \sum_{k\in[r]} \frac{|\{i: i\in\Omega_k, i\notin\hat{\Omega}_{h(k)}\}|}{|\Omega_k|}$, where $\mathcal{H}$ is the set of all bijections from $[r]$ to $[r]$ so that the comparison finds the best cluster label matching. The error metric MR counts the total misclustered modes normalized by the cluster sizes, which implies clustering errors occurring in smaller clusters would yield larger MR.

We define the following averaged feature matrix $\bar{\mathbf{\Phi}}$ based on the underlying partition $\Omega_{1:r}$: for all $i \in [s]$ (suppose $i \in \Omega_k$ for some $k \in [r]$), $\bar{\mathbf{\Phi}}(i,:) = \frac{1}{|\Omega_k|}\sum_{i'\in\Omega_k} \mathbf{\Phi}(i',:)$. By construction, there are up to r unique rows in $\bar{\mathbf{\Phi}}$, hence $\mathrm{rank}(\bar{\mathbf{\Phi}}) \leq r$. We first present the clustering guarantee for Problem P6.2, i.e., the aggregatable case.

Theorem 6.4. *Consider Problem P6.2 and Algorithm 3. Suppose $\hat{\Omega}_{1:r}$ is a $(1+\epsilon)$ k-means solution. Let $\epsilon_{Agg}^2 := \alpha_{\mathbf{A}}^2\epsilon_{\mathbf{A}}^2 + \alpha_{\mathbf{B}}^2\epsilon_{\mathbf{B}}^2 + \alpha_{\mathbf{T}}^2\epsilon_{\mathbf{T}}^2$. Then, if $\mathrm{rank}(\bar{\mathbf{\Phi}}) = r$ and $\epsilon_{Agg} \leq \frac{\sigma_r(\bar{\mathbf{\Phi}})\sqrt{|\Omega_{(r)}|+|\Omega_{(1)}|}}{8\sqrt{(2+\epsilon)|\Omega_{(1)}|}}$, we have*

$$\mathrm{MR}(\hat{\Omega}_{1:r}) \leq 64(2+\epsilon)\sigma_r(\bar{\mathbf{\Phi}})^{-2}\epsilon_{Agg}^2. \tag{6.8}$$

Additionally, if $\epsilon_{Agg} \leq \frac{\sigma_r(\bar{\mathbf{\Phi}})}{8\sqrt{(2+\epsilon)|\Omega_{(1)}|}}$, then $\mathrm{MR}(\hat{\Omega}_{1:r}) = 0$.

The key term ϵ_{Agg} measures how modes within the same cluster differ from each other, i.e., inner-cluster distance. On the other hand, the singular value $\sigma_r(\bar{\mathbf{\Phi}})$ measures the differences of modes from different clusters, i.e., inter-cluster distance. This is because when modes belonging to different clusters have similar features, their corresponding rows in the averaged feature matrix $\bar{\mathbf{\Phi}}$ will also be similar, which could give small $\sigma_r(\bar{\mathbf{\Phi}})$. Particularly, if two different clusters share the same features, then $\mathrm{rank}(\bar{\mathbf{\Phi}}) = r-1$ and $\sigma_r(\bar{\mathbf{\Phi}}) = 0$. In the theorem, when the inner-cluster distance is small compared to the inter-cluster distance, the misclustering rate can be bounded by their ratio $\epsilon_{Agg}/\sigma_r(\bar{\mathbf{\Phi}})$. By definition of misclustering rate, the smallest nonzero value it can take is given by $\frac{1}{|\Omega_{(1)}|}$. Therefore, whenever the upper bound in (6.8) is smaller than $\frac{1}{|\Omega_{(1)}|}$, one can guarantee $\mathrm{MR}(\hat{\Omega}_{1:r}) = 0$, which yields the final claim in Theorem 6.4.

The clustering guarantee for the lumpable case in Problem P6.1 is more involved than the aggregatable case. We first provide a few more notions and definitions that can help the exposition. We say a Markov matrix $\mathbf{T}$ is *reversible* if there exists a distribution $\boldsymbol{\pi} \in \mathbb{R}^s$ such that $\boldsymbol{\pi}(i)\mathbf{T}(i,j) = \boldsymbol{\pi}(j)\mathbf{T}(j,i)$ for all $i, j \in [s]$. This condition translates to $\mathbf{diag}(\boldsymbol{\pi})\mathbf{T} = \mathbf{T}^\intercal \mathbf{diag}(\boldsymbol{\pi})$ when $\mathbf{T}$ is ergodic with stationary distribution $\boldsymbol{\pi}$. For a reversible Markov matrix that is also lumpable, we have the following property.

Lemma 6.5 ((Meilă and Shi, 2001, Appendix A))**.** *For a reversible Markov matrix $\mathbf{T}$ that is also lumpable with respect to partition $\Omega_{1:r}$, it is diagonalizable with real eigenvalues. Let $\mathbf{S} \in \mathbb{R}^{s \times s}$ denote an arbitrary eigenvector matrix of $\mathbf{T}$. Then, there exists an index set $\mathcal{A} \subseteq [s]$ with $|\mathcal{A}| = r$ such that for all $k \in [r]$, for all $i, i' \in \Omega_k$, we have $\mathbf{S}(i, \mathcal{A}) = \mathbf{S}(i', \mathcal{A})$.*

We say $\mathbf{T}$ in Lemma 6.5 has *informative spectrum* if $\mathcal{A} = [r]$ and $|\lambda_r(\mathbf{T})| > |\lambda_{r+1}(\mathbf{T})|$, which implies that the r eigenvectors that carry partition information in Lemma 6.5 correspond to the r leading eigenvalues. For lumpable Markov matrices, we define the $\epsilon_{\mathbf{T}}$-neighborhood of $\mathbf{T}$:

$$\begin{aligned}\mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_{\mathbf{T}}) := \Big\{ & \mathbf{T}_0 \in \mathbb{R}^{s \times s} : \mathbf{T}_0 \text{ is Markovian}, \\ & \|\mathbf{T}_0 - \mathbf{T}\|_\infty \le \epsilon_{\mathbf{T}}, \|\mathbf{T}_0 - \mathbf{T}\|_{\mathrm{F}} \le \epsilon_{\mathbf{T}}, \\ & \forall k, l \in [r], \forall i \in \Omega_k, \sum_{j \in \Omega_l} \mathbf{T}_0(i,j) = \frac{1}{|\Omega_k|} \sum_{\substack{i' \in \Omega_k \\ j \in \Omega_l}} \mathbf{T}(i', j) \Big\}.\end{aligned} \tag{6.9}$$

Under the approximate lumpability condition in (6.3), one can show this neighborhood set is non-empty. Related discussions are provided in Appendix D.2. To find such a $\mathbf{T}_0 \in \mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_{\mathbf{T}})$, one only needs to solve a feasibility linear programming problem. Then we provide the clustering guarantee for the lumpable case.

Theorem 6.6. *Consider Problem P6.1 and Algorithm 3. Define notations $\gamma_1 := \sum_{i=2}^s \frac{1}{1-\lambda_i(\mathbf{T})}$, $\gamma_2 := \min\{\sigma_r(\mathbf{H}) - \sigma_{r+1}(\mathbf{H}), 1\}$, $\gamma_3 := \frac{16\gamma_1\sqrt{r\pi_{\max}}\|\mathbf{T}\|_{\mathrm{F}}}{\gamma_2 \pi_{\min}^2}$, and $\epsilon_{Lmp}^2 := \alpha_{\mathbf{A}}^2\epsilon_{\mathbf{A}}^2 + \alpha_{\mathbf{B}}^2\epsilon_{\mathbf{B}}^2 + \alpha_{\mathbf{T}}^2\gamma_3^2\epsilon_{\mathbf{T}}^2$. Assume there exists an ergodic and reversible $\mathbf{T}_0 \in \mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_{\mathbf{T}})$ with informative spectrum. Suppose $\hat{\Omega}_{1:r}$ is a $(1+\epsilon)$ k-means solution. Then, if $\mathrm{rank}(\bar{\mathbf{\Phi}}) = r$, $\epsilon_{\mathbf{T}} \le \frac{\pi_{\min}}{\gamma_1}$, $\epsilon_{Lmp} \le \frac{\sigma_r(\bar{\mathbf{\Phi}})\sqrt{|\Omega_{(r)}| + |\Omega_{(1)}|}}{8\sqrt{s(2+\epsilon)|\Omega_{(1)}|}}$, we have*

$$\mathrm{MR}(\hat{\Omega}_{1:r}) \le 64(2+\epsilon)\sigma_r(\bar{\mathbf{\Phi}})^{-2}\epsilon_{Lmp}^2. \tag{6.10}$$

Additionally, if $\epsilon_{Lmp} \leq \frac{\sigma_r(\bar{\mathbf{\Phi}})}{8\sqrt{(2+\epsilon)|\Omega_{(1)}|}}$*, then* $\mathrm{MR}(\hat{\Omega}_{1:r}) = 0$.

Theorem 6.6 for the lumpable case is similar to Theorem 6.4 for the aggregatable case with an additional γ_3 term. This is a result of using $\mathbf{S}_r$ and $\mathbf{T}$ to construct features in Algorithm 3 for these two cases. γ_3 describes how much the lumpability perturbation $\epsilon_{\mathbf{T}}$ on $\mathbf{T}$ affects the row equalities of its spectrum-related matrix $\mathbf{S}_r$ in Lemma 6.5. The assumption on the existence of $\mathbf{T}_0$ with informative spectrum guarantees (i) the partition $\Omega_{1:r}$ information is carried by the leading eigenvectors of $\mathbf{T}_0$ as introduced in Lemma 6.5, and (ii) this information can still be preserved in $\mathbf{S}_r$ as long as $\mathbf{T}$ is close to $\mathbf{T}_0$. Because of this, Theorem 6.6 may not hold for arbitrary lumpable $\mathbf{T}$, but only those close to Markov matrices with informative spectra.

6.5 Approximation Guarantees

With perturbations $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}}$, the reduced $\hat{\Sigma}$ may not be equivalent to the original Σ as in Fact 6.3. In this case, if certain approximation guarantees can be established, they can be used in verification tasks such as safety (Julius and Pappas, 2009) and invariance (Soudjani and Abate, 2011) evaluations. In this section, we show that the reduced system $\hat{\Sigma}$ can be guaranteed to well approximate the original system Σ under metrics such as transition kernels (distributions) and trajectory realizations. Particularly, these metrics reach 0 when $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}} = 0$, i.e., the mode-reducibility condition in Definition 6.2 holds. We have shown in Theorem 6.4 and 6.6 that $\mathrm{MR}(\hat{\Omega}_{1:r}) = 0$ when perturbations $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}}$ are small. Hence, in this section together with Section 6.6 and 6.7, we assume $\Omega_{1:r} = \hat{\Omega}_{1:r}$ for simplicity. In these sections, the theory holds for perturbations $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}}$ introduced in either P6.1 or P6.2.

We study the approximation in a setup where Σ and $\hat{\Sigma}$ start with the same initial condition and are driven by the same input.

Setup S1 (Initialization-Excitation Setup)**.** *Systems* Σ *and* $\hat{\Sigma}$ *have (i) initial mode distributions satisfy* $\mathbb{P}(\omega_0 \in \hat{\Omega}_k) = \mathbb{P}(\hat{\omega}_0 = k)$ *for all* $k \in [r]$*; (ii) the same initial states, i.e.,* $\mathbf{x}_0 = \hat{\mathbf{x}}_0$*, and (iii) the same inputs* $\mathbf{u}_t = \hat{\mathbf{u}}_t$ *for all* t*.*

Note that when Σ and $\hat{\Sigma}$ have fixed and shared initial conditions and inputs as in setup S1, we can at most evaluate the difference between $\mathbf{x}_t$ and $\hat{\mathbf{x}}_t$ in terms of their distributions (or, the transition kernels of Σ and $\hat{\Sigma}$). However, the actual realizations of $\mathbf{x}_t$ and $\hat{\mathbf{x}}_t$, i.e., when we only generate a single sample for each, can be very different. This is because $\mathbf{x}_t$ and $\hat{\mathbf{x}}_t$ are driven not only by the input excitation, but also the mode switching sequences $\omega_{0:t-1}$ and $\hat{\omega}_{0:t-1}$; thus, $\mathbf{x}_t$ and $\hat{\mathbf{x}}_t$ are likely to be far away from each other if the realizations of $\omega_{0:t-1}$ and $\hat{\omega}_{0:t-1}$ are different. On the other hand, if the reduced model $\hat{\Sigma}$ is to be used online to predict the future behavior of Σ, and if the mode ω_t can be observed at run-time, we can assume the following and derive stronger relations on the state realization difference $\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|$.

Setup S2 (Mode Synchrony). *Mode $\hat{\omega}_t$ of $\hat{\Sigma}$ is synchronous to ω_t of Σ, i.e., for all t, if $\omega_t \in \hat{\Omega}_k$ then $\hat{\omega}_t = k$.*

Mode synchrony setup essentially establishes the strongest possible coupling between $\omega_{0:t}$ and $\hat{\omega}_{0:t}$ as discussed in Section 6.3.2. When the mode sequence $\hat{\omega}_{0:t}$ of $\hat{\Sigma}$ is synchronized with that of Σ, this amounts to having $\hat{\Sigma}$ being driven by an external switching signal $\omega_{0:t}$.

In the following, we provide bounds on how close $\hat{\Sigma}$ is to Σ in terms of the following approximation metrics: (i) under the mean-square stability of Σ, the difference $\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|$ in trajectories (Theorem 6.7); (ii) under uniform stability, the difference in trajectories (Theorem 6.8 (T1)) and the difference of transition kernels (Theorem 6.8 (T2)).

6.5.1 Result with Mean-square Stability

From Lemma 2.12, we know the MJS Σ is mean-square stable if and only if the the augmented state matrix $\boldsymbol{\mathcal{A}}$ in (2.9) has spectral radius $\rho(\boldsymbol{\mathcal{A}}) < 1$. For any $\rho \geq \rho(\boldsymbol{\mathcal{A}})$, let us define $\tau := \tau(\boldsymbol{\mathcal{A}}, \rho)$, where $\tau(\cdot,\cdot)$ is the normalized power supremum in Definition 2.1. Keep in mind that τ depends on the choice of the free parameter ρ. We let $\bar{A} := \max_i \|\mathbf{A}_i\|$, $\bar{B} := \max_i \|\mathbf{B}_i\|$. The following theorem provides an upper bound for $\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|$ under mean-square stability.

Theorem 6.7. *Consider setup S1 and S2 where the shared initial state $\mathbf{x}_0$ and inputs $\mathbf{u}_{0:t}$ can be arbitrary as long as for all t, $\mathbf{u}_t$ is bounded, i.e., $\|\mathbf{u}_t\| \leq \bar{u}$. Assume Σ is mean-square stable and $\hat{\Omega}_{1:r} = \Omega_{1:r}$ in Algorithm 3. For any $\rho \in [\rho(\boldsymbol{\mathcal{A}}), 1)$ and its corresponding τ, let $\rho_0 := \frac{1+\rho}{2}$. For perturbation, assume $\epsilon_{\mathbf{A}} \leq \min\{\bar{A}, \frac{1-\rho}{6\tau\bar{A}\|\mathbf{T}\|}\}$ and $\epsilon_{\mathbf{B}} \leq \bar{B}$. Then, $\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|] \leq 4\sqrt{n\sqrt{s}}\tau\epsilon_t^{mss}$ where $\epsilon_t^{mss} := \rho_0^{\frac{t-1}{2}}\sqrt{t\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}}\|\mathbf{x}_0\| + \sqrt{\bar{B}}\bar{u}\big(\frac{\sqrt{\rho_0}}{(1-\sqrt{\rho_0})^2}\sqrt{\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}} + \frac{\sqrt{2}}{1-\sqrt{\rho_0}}\sqrt{\epsilon_{\mathbf{B}}}\big)$.*

In this theorem, ϵ_t^{mss} is the key element in the upper bound. In its definition, the first term describes the effect of $\epsilon_{\mathbf{A}}$ through initial the state $\mathbf{x}_0$. Since $\rho < 1$ due to Σ being mean-square stable, we know $\rho_0 < 1$, which implies exponential decay. The rest of the terms in ϵ_t^{mss} characterize the effects of $\epsilon_{\mathbf{A}}$ and $\epsilon_{\mathbf{B}}$ through the inputs. And if there is no input, the trajectory difference $\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|$ converges to 0 exponentially with t. The condition $\epsilon_{\mathbf{A}} \leq \frac{1-\rho}{6\sqrt{s}\tau\bar{A}}$ is used to guarantee perturbation $\epsilon_{\mathbf{A}}$ is small such that $\hat{\Sigma}$ is also mean-square stable, as otherwise the difference will grow exponentially, and no meaningful results can be established in this case. Conditions $\epsilon_{\mathbf{A}} < \bar{A}$ and $\epsilon_{\mathbf{B}} \leq \bar{B}$ are only used to simplify the expressions, and similar bounds can be established without them.

Fact 6.3 provides a sanity check for Theorem 6.7: when $\epsilon_{\mathbf{A}} = \epsilon_{\mathbf{B}} = 0$, we have $\mathbf{x}_t = \hat{\mathbf{x}}_t$. In the autonomous case, i.e., $\mathbf{u}_t = 0$, as a direct corollary of Theorem 6.7, we can further obtain a probabilistic bound on the difference over an entire trajectory using Markov inequality: with probability at least $1-\delta$, $\sum_{t=0}^{\infty} \|\mathbf{x}_t - \hat{\mathbf{x}}_t\| \leq \frac{4\sqrt{np}\tau\|\mathbf{x}_0\|\sqrt{\bar{A}\epsilon_{\mathbf{A}}}}{\delta(1-\sqrt{\rho_0})^2}$.

6.5.2 Results with Uniform Stability

Mean-square stability in Section 6.5.1 is a weak notion of stability in that it only requires stability in expectation while still allowing a set of mode switching sequences that result in explosive $\mathbf{x}_t$, even a set with nonzero probability. In this section, we consider uniform stability, which guarantees stable $\mathbf{x}_t$ even with an arbitrary switching sequence. Uniform stability allows us to further build approximation results without enforcing mode synchrony as in S2.

From Definition 2.16, we know the MJS Σ is uniformly stable if the matrices $\mathbf{A}_{1:s}$ have joint spectral radius $\xi(\mathbf{A}_{1:s}) < 1$. For any $\xi \geq \xi(\mathbf{A}_{1:s})$, let us define $\kappa := \kappa(\mathbf{A}_{1:s}, \xi)$, where $\kappa(\cdot,\cdot)$ is the normalized joint power supremum in Definition 2.17. Furthermore, we let $\bar{\mathcal{T}} := \max_{i,j} \mathbf{T}(i,j)$.

We formally define the transition kernels for Σ and $\hat{\Sigma}$ and their distance. Under fixed initial state $\mathbf{x}_0$ and input sequence $\mathbf{u}_{0:t-1}$, we let $\mathcal{X}_t := \{\mathbf{x}_t : \mathbf{x}_{0:t} \text{ is a solution to } \Sigma, \forall\, \omega_{0:t-1} \in [s]^t\}$ denote the reachable set of Σ. Then we define the t-step transition kernel as $p_t(\mathbf{x}) := \mathbb{P}(\mathbf{x}_t = \mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}_t$. Note that both $\mathcal{X}_t$ and $p_t(\mathbf{x})$ depend on the choice of the initial state and input sequence as well. We omit this dependency in the notation not only for simplicity but also because the approximation results we provide hold for arbitrary initial state and input sequence. Similarly, for the reduced $\hat{\Sigma}$, we use $\hat{\mathcal{X}}_t$ to denote the reachable set at time t, and for $\hat{\mathbf{x}} \in \hat{\mathcal{X}}_t$, we let $\hat{p}_t(\hat{\mathbf{x}}) := \mathbb{P}(\hat{\mathbf{x}}_t = \hat{\mathbf{x}})$. Then, for $\ell \geq 1$ the ℓ-Wasserstein distance $W_\ell(p_t, \hat{p}_t)$, between distributions p_t and $\hat{p}_t$ is defined as the optimal objective value of the following mass transportation problem:

$$\begin{aligned} \min_{f\geq 0} \quad & \left(\textstyle\sum_{\mathbf{x}\in\mathcal{X}_t, \hat{\mathbf{x}}\in\hat{\mathcal{X}}_t} f(\mathbf{x},\hat{\mathbf{x}}) \|\mathbf{x}-\hat{\mathbf{x}}\|^\ell\right)^{1/\ell} \\ \text{s.t.} \quad & \textstyle\sum_{\mathbf{x}\in\mathcal{X}_t} f(\mathbf{x},\hat{\mathbf{x}}) = \hat{p}_t(\hat{\mathbf{x}}), \forall\, \hat{\mathbf{x}} \\ & \textstyle\sum_{\hat{\mathbf{x}}\in\hat{\mathcal{X}}_t} f(\mathbf{x},\hat{\mathbf{x}}) = p_t(\mathbf{x}), \forall\, \mathbf{x}. \end{aligned} \tag{6.11}$$

The constraints describe the transportation of probability mass distributed according to p_t to the support of $\hat{p}_t$ so that the mass after transportation distributes the same as $\hat{p}_t$. We can view $f(\mathbf{x}, \hat{\mathbf{x}})$ as the mass that is transported from point $\mathbf{x}$ to $\hat{\mathbf{x}}$ and $\|\mathbf{x} - \hat{\mathbf{x}}\|$ as the distance it travels. When $\ell = 1$, the goal is to minimize the total weighted travel distance, and the resulting W_1 is also known as the earth mover's distance. Now we are ready to present our results for the uniform stability assumption.

Theorem 6.8. *Consider setup S1 where the shared initial state $\mathbf{x}_0$ and inputs $\mathbf{u}_{0:t}$ can be arbitrary as long as for all t, $\mathbf{u}_t$ is bounded, i.e., $\|\mathbf{u}_t\| \leq \bar{u}$. Assume Σ is uniformly stable and $\hat{\Omega}_{1:r} = \Omega_{1:r}$ in Algorithm 3. For any $\xi \in [\xi(\mathbf{A}_{1:s}), 1)$ and its corresponding κ, let $\xi_0 := \frac{1+\xi}{2}$. For perturbation, we assume $\epsilon_{\mathbf{A}} \leq \frac{1-\xi}{2\kappa}$ and $\epsilon_{\mathbf{B}} \leq \bar{B}$. Then, we have the following results.*

(T1) *Under S2, $\|\mathbf{x}_t - \hat{\mathbf{x}}_t\| \leq \epsilon_t^{us} := t\xi_0^{t-1}\kappa^2\|\mathbf{x}_0\|\epsilon_{\mathbf{A}} + \frac{2(1+t\xi_0^t)\kappa^2\bar{B}\bar{u}}{1-\xi_0}\epsilon_{\mathbf{A}} + \frac{\kappa\bar{u}}{1-\xi}\epsilon_{\mathbf{B}}$ almost surely.*

(T2) *Consider the autonomous case, i.e., $\mathbf{B}_{1:s} = 0$. (S2 is not mandatory.) Then,*

$$W_\ell(p_t, \hat{p}_t) \leq t\xi_0^{t-1}\kappa^2\|\mathbf{x}_0\|\epsilon_{\mathbf{A}} + 2r^2t\kappa\|\mathbf{x}_0\|r^t(\kappa\epsilon_{\mathbf{A}} + \xi)^t(\bar{\mathcal{T}} + \epsilon_{\mathbf{T}})^{(t-2)/\ell}\epsilon_{\mathbf{T}}^{1/\ell}.$$

In Theorem 6.8, condition $\epsilon_{\mathbf{A}} \leq \frac{1-\xi}{2\kappa}$ guarantees the reduced $\hat{\Sigma}$ is uniformly stable with joint spectral radius upper bounded by ξ_0. The condition $\epsilon_{\mathbf{B}} \leq \bar{B}$ simplifies the expression, and similar results can be obtained when it is relaxed. (T1) upper bounds the realization difference with the mode synchrony setup. We can see the similarity between the upper bounds ϵ_t^{us} and ϵ_t^{mss} of Theorem 6.7 under the mean-square stability assumption. The fact that uniform stability and mean-square stability upper bound $\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|$ deterministically and in expectation respectively is a manifestation of the difference between these two stability notions for MJS.

In (T2), we bound the Wasserstein distance between p_t and $\hat{p}_t$. This bound depends on both perturbations $\epsilon_{\mathbf{A}}$ and $\epsilon_{\mathbf{T}}$. Let $\boldsymbol{\mu}$ and $\mathbf{S}$ denote the mean and covariance for $\mathbf{x}_t$; and similarly define $\hat{\boldsymbol{\mu}}$ and $\hat{\mathbf{S}}$ for $\hat{\mathbf{x}}_t$. From Kuhn et al. (2019, Theorem 4), we obtain $\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|^2 + d(\mathbf{S}, \hat{\mathbf{S}}) \leq W_2(p_t, \hat{p}_t)^2$, where $d(\mathbf{S}, \hat{\mathbf{S}}) := \mathbf{tr}(\mathbf{S} + \hat{\mathbf{S}} - 2(\mathbf{S}^{\frac{1}{2}}\hat{\mathbf{S}}\mathbf{S}^{\frac{1}{2}})^{\frac{1}{2}})$ is a metric between $\mathbf{S}$ and $\hat{\mathbf{S}}$. Hence, by setting $\ell = 2$ in (T2), we also obtain upper bounds for the differences between p_t and $\hat{p}_t$ in terms of their first and second order moments, i.e., $\|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|$ and $d(\mathbf{S}, \hat{\mathbf{S}})$. These metrics can be used to obtain performance bounds in other control problems such as covariance steering (Chen et al., 2015; Goldshtein and Tsiotras, 2017; Okamoto et al., 2018) and ensemble control (Li and Qi, 2015).

6.6 Stability Analysis

In this section, we study whether the stability properties of Σ can be deduced from those of $\hat{\Sigma}$. Recall that the mean-square stability of Σ depends on $\rho(\boldsymbol{\mathcal{A}})$, the spectral radius of its augmented state matrix $\boldsymbol{\mathcal{A}}$, and its uniform stability depends on $\xi(\mathbf{A}_{1:s})$, the joint spectral radius of state matrices $\mathbf{A}_{1:s}$. Similarly, for $\hat{\Sigma}$, we define its augmented state matrix $\hat{\boldsymbol{\mathcal{A}}} \in \mathbb{R}^{rn^2 \times rn^2}$ with its (i,j)-th $n^2 \times n^2$ block given by $[\hat{\boldsymbol{\mathcal{A}}}]_{ij} := \hat{\mathbf{T}}(j,i) \cdot \hat{\mathbf{A}}_j \otimes \hat{\mathbf{A}}_j$ and let $\rho(\hat{\boldsymbol{\mathcal{A}}})$ denote its spectral radius; we let $\xi(\hat{\mathbf{A}}_{1:r})$ denote the joint spectral radius of state matrices $\hat{\mathbf{A}}_{1:r}$. With these notations, we want to analyze when $\rho(\hat{\boldsymbol{\mathcal{A}}})$ (or $\xi(\hat{\mathbf{A}}_{1:r})$) can be taken as an approximation for $\rho(\boldsymbol{\mathcal{A}})$ (or $\xi(\mathbf{A}_{1:s})$) since computing or approximating $\rho(\hat{\boldsymbol{\mathcal{A}}})$ and $\xi(\hat{\mathbf{A}}_{1:r})$ may require much less computation compared with $\rho(\boldsymbol{\mathcal{A}})$ and $\xi(\mathbf{A}_{1:s})$ as $\hat{\Sigma}$ has much fewer number of modes than Σ.

To begin with, we first construct an intermediate MJS by *expanding* the reduced $\hat{\Sigma}$: we let $\bar{\Sigma} := \text{MJS}(\bar{\mathbf{A}}_{1:s}, \bar{\mathbf{B}}_{1:s}, \bar{\mathbf{T}})$ such that $\bar{\mathbf{T}} \in \mathcal{L}(\mathbf{T}, \hat{\Omega}_{1:r}, \epsilon_{\mathbf{T}})$, and for all $i \in [s]$ (suppose $i \in \hat{\Omega}_k$), $\bar{\mathbf{A}}_i = \hat{\mathbf{A}}_k$, $\bar{\mathbf{B}}_i = \hat{\mathbf{B}}_k$. By definition of $\mathcal{L}(\mathbf{T}, \hat{\Omega}_{1:r}, \epsilon_{\mathbf{T}})$, we can solve for $\bar{\mathbf{T}}$ through a linear programming feasibility problem with constraints given by the definition of $\mathcal{L}(\cdot,\cdot,\cdot)$ in (6.9). Particularly, if it is the aggregatable case P6.2, it suffices to let $\bar{\mathbf{T}}(i,:) := |\hat{\Omega}_k|^{-1} \sum_{i \in \hat{\Omega}_k} \mathbf{T}(i,:)$ if $i \in \hat{\Omega}_k$. Note that by construction, $\bar{\Sigma}$ is mode-reducible with respect to $\hat{\Omega}_{1:r}$ and can be reduced to $\hat{\Sigma}$. According to Fact 6.3, $\bar{\Sigma}$ has the same dynamics as $\hat{\Sigma}$. Since $\bar{\Sigma}$ has the same number of modes as Σ, we can use $\bar{\Sigma}$ as a bridge to compare Σ and $\hat{\Sigma}$. We let $\rho(\bar{\boldsymbol{\mathcal{A}}})$ denote the spectral radius of $\bar{\boldsymbol{\mathcal{A}}} \in \mathbb{R}^{sn^2 \times sn^2}$ whose (i,j)-th $n^2 \times n^2$ block is given by $[\bar{\boldsymbol{\mathcal{A}}}]_{ij} := \bar{\mathbf{T}}(j,i) \cdot \bar{\mathbf{A}}_j \otimes \bar{\mathbf{A}}_j$ and let $\xi(\bar{\mathbf{A}}_{1:s})$ denote the joint spectral radius of $\bar{\mathbf{A}}_{1:s}$. The following preliminary result (proof omitted due to its simplicity) says $\bar{\Sigma}$ and $\hat{\Sigma}$ have the same stability properties.

Lemma 6.9. *For $\hat{\Sigma}$ and $\bar{\Sigma}$, we have $\rho(\hat{\boldsymbol{\mathcal{A}}}) = \rho(\bar{\boldsymbol{\mathcal{A}}})$ and $\xi(\hat{\mathbf{A}}_{1:s}) = \xi(\bar{\mathbf{A}}_{1:s})$.*

One implication of Lemma 6.9 is that if an MJS is mode-reducible, the reduced MJS has the same mean-square stability and uniform stability as the original MJS in terms of (joint) spectral radius. When Σ is not exactly mode-reducible, Lemma 6.9 allows us to compare the stability properties of $\hat{\Sigma}$ and Σ via the intermediate expanded $\bar{\Sigma}$ as presented in Theorem 6.10. For analysis purposes, similar to τ and κ defined for Σ, we define $\bar{\tau} := \sup_{k \in \mathbb{N}} \|\bar{\boldsymbol{\mathcal{A}}}^k\| / \hat{\rho}^k$ for any $\hat{\rho} \geq \rho(\hat{\boldsymbol{\mathcal{A}}})$ and $\bar{\kappa} := \sup_{k \in \mathbb{N}} \max_{\omega_{1:k} \in [r]^k} \|\bar{\mathbf{A}}_{\omega_1} \cdots \bar{\mathbf{A}}_{\omega_k}\| / \hat{\xi}^k$ for any $\hat{\xi} \geq \xi(\hat{\mathbf{A}}_{1:s})$

Theorem 6.10 (Stability Analysis)**.** *Assume $\hat{\Omega}_{1:r} = \Omega_{1:r}$ in Algorithm 3, then Σ and $\hat{\Sigma}$ have the following relations.*

(T1) *(Mean-square stability) For any $\rho \geq \rho(\boldsymbol{\mathcal{A}})$ and its corresponding τ, any $\hat{\rho} \geq \rho(\hat{\boldsymbol{\mathcal{A}}})$ and* *its corresponding $\bar{\tau}$, we have*

$$\begin{aligned}\rho(\hat{\boldsymbol{\mathcal{A}}}) - \rho(\boldsymbol{\mathcal{A}}) &\leq \tau\epsilon_\rho + (\rho - \rho(\boldsymbol{\mathcal{A}}))\\ \rho(\boldsymbol{\mathcal{A}}) - \rho(\hat{\boldsymbol{\mathcal{A}}}) &\leq \bar{\tau}\epsilon_\rho + (\hat{\rho} - \rho(\hat{\boldsymbol{\mathcal{A}}}))\end{aligned} \tag{6.12}$$

where $\epsilon_\rho := \sqrt{s}((2\bar{A} + \epsilon_{\mathbf{A}})\epsilon_{\mathbf{A}} + \bar{A}^2\epsilon_{\mathbf{T}})$.

(T2) *(Uniform stability) For any $\xi \geq \xi(\mathbf{A}_{1:s})$ and its corresponding κ, any $\hat{\xi} \geq \xi(\hat{\mathbf{A}}_{1:r})$ and its corresponding $\bar{\tau}$, we have*

$$\begin{aligned}\xi(\hat{\mathbf{A}}_{1:r}) - \xi(\mathbf{A}_{1:s}) &\leq \kappa\epsilon_{\mathbf{A}} + (\xi - \xi(\mathbf{A}_{1:s}))\\ \xi(\mathbf{A}_{1:s}) - \xi(\hat{\mathbf{A}}_{1:r}) &\leq \bar{\kappa}\epsilon_{\mathbf{A}} + (\hat{\xi} - \xi(\hat{\mathbf{A}}_{1:r})).\end{aligned} \tag{6.13}$$

Proof. From Lemma 6.9, it suffices to prove

$$\rho(\bar{\boldsymbol{\mathcal{A}}}) \leq \tau\epsilon_\rho + \rho, \quad \rho(\boldsymbol{\mathcal{A}}) \leq \bar{\tau}\epsilon_\rho + \hat{\rho}. \tag{6.14}$$

$$\xi(\bar{\mathbf{A}}_{1:s}) \leq \kappa\epsilon_{\mathbf{A}} + \xi, \quad \xi(\mathbf{A}_{1:s}) \leq \bar{\kappa}\epsilon_{\mathbf{A}} + \hat{\xi}. \tag{6.15}$$

Since we assume $\hat{\Omega}_{1:r} = \Omega_{1:r}$, then for Σ and $\bar{\Sigma}$, we have $\|\bar{\mathbf{A}}_i - \mathbf{A}_i\| \leq \epsilon_{\mathbf{A}}$, $\|\bar{\mathbf{B}}_i - \mathbf{B}_i\| \leq \epsilon_{\mathbf{B}}$, and $\|\bar{\mathbf{T}} - \mathbf{T}\|_\infty \leq \epsilon_{\mathbf{T}}$. Consider matrix $\bar{\boldsymbol{\mathcal{A}}}$ and $\boldsymbol{\mathcal{A}}$, we have $[\bar{\boldsymbol{\mathcal{A}}}]_{ij} - [\boldsymbol{\mathcal{A}}]_{ij} = \bar{\mathbf{T}}(j,i)\bar{\mathbf{A}}_j \otimes \bar{\mathbf{A}}_j - \mathbf{T}(j,i)\mathbf{A}_j \otimes \mathbf{A}_j = \bar{\mathbf{T}}(j,i)(\bar{\mathbf{A}}_j \otimes \bar{\mathbf{A}}_j - \mathbf{A}_j \otimes \mathbf{A}_j) + (\bar{\mathbf{T}}(j,i) - \mathbf{T}(j,i))\mathbf{A}_j \otimes \mathbf{A}_j$. Note that $\bar{\mathbf{A}}_j \otimes \bar{\mathbf{A}}_j - \mathbf{A}_j \otimes \mathbf{A}_j = (\bar{\mathbf{A}}_j - \mathbf{A}_j) \otimes \mathbf{A}_j + \mathbf{A}_j \otimes (\bar{\mathbf{A}}_j - \mathbf{A}_j) + (\bar{\mathbf{A}}_j - \mathbf{A}_j) \otimes (\bar{\mathbf{A}}_j - \mathbf{A}_j)$, which gives $\|\bar{\mathbf{A}}_j \otimes \bar{\mathbf{A}}_j - \mathbf{A}_j \otimes \mathbf{A}_j\| \leq (2\bar{A} + \epsilon_{\mathbf{A}})\epsilon_{\mathbf{A}}$. Then, we have $\|[\bar{\boldsymbol{\mathcal{A}}}]_{ij} - [\boldsymbol{\mathcal{A}}]_{ij}\| \leq \bar{\mathbf{T}}(j,i)(2\bar{A} + \epsilon_{\mathbf{A}})\epsilon_{\mathbf{A}} + |\bar{\mathbf{T}}(j,i) - \mathbf{T}(j,i)|\bar{A}^2$. To simplify the notation, we let $c_1 := (2\bar{A} + \epsilon_{\mathbf{A}})\epsilon_{\mathbf{A}}$ and $c_2 := \bar{A}^2$. By Cauchy-Schwarz inequality, we have $\sum_i \|[\bar{\boldsymbol{\mathcal{A}}}]_{ij} - [\boldsymbol{\mathcal{A}}]_{ij}\|^2 \leq (c_1\|\bar{\mathbf{T}}(j,:)\| + c_2\|\bar{\mathbf{T}}(j,:) - \mathbf{T}(j,:)\|)^2$. Thus, $\|\bar{\boldsymbol{\mathcal{A}}} - \boldsymbol{\mathcal{A}}\| \leq \sqrt{s}\max_j(\sum_i \|[\bar{\boldsymbol{\mathcal{A}}}]_{ij} - [\boldsymbol{\mathcal{A}}]_{ij}\|)^{0.5} \leq \sqrt{s}(c_1 + c_2\epsilon_{\mathbf{T}}) =: \epsilon_\rho$.

With Corollary D.9 in the appendix, we have $\|\bar{\boldsymbol{\mathcal{A}}}^k\| \leq \tau(\tau\epsilon_\rho + \rho)^k$. By Gelfand's formula, $\rho(\bar{\boldsymbol{\mathcal{A}}}) = \limsup_{k\to\infty} \|\bar{\boldsymbol{\mathcal{A}}}^k\|^{\frac{1}{k}} \leq \tau\epsilon_\rho + \rho$, which shows the left inequality of (6.14). If we use Corollary D.9 the other way, we have $\|\boldsymbol{\mathcal{A}}^k\| \leq \bar{\tau}(\bar{\tau}\epsilon_\rho + \hat{\rho})^k$, which similarly implies $\rho(\boldsymbol{\mathcal{A}}) \leq \bar{\tau}\epsilon_\rho + \hat{\rho}$. With these results, (6.14) is proved. (6.15) can be shown similarly by noticing $\|\bar{\mathbf{A}}_i - \mathbf{A}_i\| \leq \epsilon_{\mathbf{A}}$ and then using Lemma D.8 in the appendix. □

Theorem 6.10 provides upper bounds on $|\rho(\hat{\boldsymbol{\mathcal{A}}}) - \rho(\boldsymbol{\mathcal{A}})|$ and $|\xi(\hat{\mathbf{A}}_{1:r}) - \xi(\mathbf{A}_{1:s})|$. By definition, τ decreases when ρ increases, and the same applies to the pairs $\{\bar{\tau}, \hat{\rho}\}$, $\{\kappa, \xi\}$, and $\{\bar{\kappa}, \hat{\xi}\}$. Hence, for fixed ϵ_ρ and $\epsilon_{\mathbf{A}}$, by tuning the free parameters ρ, $\hat{\rho}$, ξ, and $\hat{\xi}$, one may obtain tighter upper bounds in Theorem 6.10. When $\rho = \rho(\boldsymbol{\mathcal{A}})$, $\hat{\rho} = \rho(\hat{\boldsymbol{\mathcal{A}}})$, the bound in (T1) becomes tight at $\epsilon_\rho = 0$ as the upper and lower bounds meet at 0. Note that these results hold for both stable and unstable Σ, and does not require perturbation $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}}$ to be small, which is in contrast to approximation results in Theorem 6.7 and 6.8.

Now we briefly compare the complexities for computing or approximating $\rho(\boldsymbol{\mathcal{A}}), \rho(\hat{\boldsymbol{\mathcal{A}}}), \xi(\mathbf{A}_{1:s})$, and $\xi(\hat{\mathbf{A}}_{1:r})$. Since $\boldsymbol{\mathcal{A}}$ has dimension $sn^2 \times sn^2$, the complexity to compute its spectral radius $\rho(\boldsymbol{\mathcal{A}})$ is $\mathcal{O}(s^3n^6)$, but it only requires $\mathcal{O}(r^3n^6)$ for $\rho(\hat{\boldsymbol{\mathcal{A}}})$. Computation of the joint spectral radius is in general undecidable (Jungers, 2009). An iterative approach (Parrilo and Jadbabaie, 2008) provides an approximation for $\xi(\mathbf{A}_{1:s})$ with computational complexity $\mathcal{O}(s)$, whereas it only requires $\mathcal{O}(r)$ for $\xi(\hat{\mathbf{A}}_{1:r})$.

6.7 Controller Design with Case Study on LQR

When the mode of an MJS can be observed at run-time, one can use *mode-dependent* controllers. A mode-dependent controller is essentially a collection of individual controllers, one per mode, and the deployed controller switches with corresponding modes. Therefore, if we can reduce the modes, that would also reduce the number of controllers in a mode-dependent control. That is, with the reduced $\hat{\Sigma}$, we can design mode-dependent controller $\hat{\mathbf{K}}_{1:r}$ for $\hat{\Sigma}$ and then associate every mode i in Σ with $\hat{\mathbf{K}}_k$ if $i \in \hat{\Omega}_k$. Since $\hat{\Sigma}$ has a smaller scale than Σ, the computational cost may be reduced but the question is how this simplified controller performs on the original system Σ. In this section, we show how this idea can be used for the infinite-horizon LQR problems for MJS and provide suboptimality guarantees for the reduced controller.

We consider the infinite-horizon MJS-LQR problem (2.34) with mode-independent cost matrices, i.e, $\mathbf{Q}_{1:s} = \mathbf{Q}$ and $\mathbf{R}_{1:s} = \mathbf{R}$. The main assumption in this section is as follows, which guarantees the existence and uniqueness of the optimal solution to the MJS-LQR problem according to Lemma 2.22.

Assumption A6.1. *The problem* MJS-LQR$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ *satisfies the following.*

(a) For all $i \in [s]$, $\mathbf{Q}_i = \mathbf{Q}$ *and* $\mathbf{R}_i = \mathbf{R}$ *for some* $\mathbf{Q}, \mathbf{R} \succ 0$.

(b) The MJS $\Sigma = \mathrm{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *is stabilizable.*

To design controllers with the reduced $\hat{\Sigma} = \mathrm{MJS}(\hat{\mathbf{A}}_{1:r}, \hat{\mathbf{B}}_{1:r}, \hat{\mathbf{T}})$, we can first compute controller $\hat{\mathbf{K}}_{1:r}$ by solving LQR problem with $\hat{\Sigma}$ as the MJS dynamics. To ease the exposition, similar to Section 2.3.2, we define the following operators. For a collection of positive semi-definite matrices $\mathbf{X}_{1:r} \in \mathbb{S}^n_{r,+}$, for all $i \in [r]$, define

$$\hat{\varphi}_i(\mathbf{X}_{1:r}) := \sum_{j \in [r]} \hat{\mathbf{T}}(i,j)\mathbf{X}_j, \tag{6.16}$$

$$\hat{\mathcal{R}}_i(\mathbf{X}_{1:r}) := \mathbf{Q} + \hat{\mathbf{A}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:r})\hat{\mathbf{A}}_i - \hat{\mathbf{A}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:r})^\intercal \hat{\mathbf{B}}_i \left(\mathbf{R} + \hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:r})\hat{\mathbf{B}}_i\right)^{-1} \hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:r})\hat{\mathbf{A}}_i, \tag{6.17}$$

$$\hat{\mathcal{K}}_i(\mathbf{X}_{1:r}) := -\left(\mathbf{R} + \hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:r})\hat{\mathbf{B}}_i\right)^{-1} \left(\hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\mathbf{X}_{1:r})\hat{\mathbf{A}}_i\right). \tag{6.18}$$

We first solve for the following cDARE:

$$\begin{aligned} \mathbf{X}_1 &= \hat{\mathcal{R}}_1(\mathbf{X}_{1:r}) \\ \mathbf{X}_2 &= \hat{\mathcal{R}}_2(\mathbf{X}_{1:r}) \\ &\vdots \\ \mathbf{X}_r &= \hat{\mathcal{R}}_r(\mathbf{X}_{1:r}) \end{aligned} \tag{6.19}$$

with respect to positive semi-definite matrices $\mathbf{X}_{1:r} \in \mathbb{S}^n_{r,+}$. Let $\hat{\mathbf{P}}_{1:r} \in \mathbb{S}^n_{r,+}$ denote the solution. Then we compute controller $\hat{\mathbf{K}}_{1:r}$ as

$$\hat{\mathbf{K}}_i = \hat{\mathcal{K}}_i(\hat{\mathbf{P}}_{1:r}), \quad \forall\, i \in [r]. \tag{6.20}$$

To apply the controller $\hat{\mathbf{K}}_{1:r}$ to the original MJS Σ, we simply let $\mathbf{u}_t = \hat{\mathbf{K}}_k \mathbf{x}_t$ if $\omega_t = k$. Note that in MJS-LQR problems, the number of coupled Riccati equations is the same as the number of modes, the computational cost for Σ is $\mathcal{O}(s)$ while only $\mathcal{O}(r)$ for $\hat{\Sigma}$, thus the saving is prominent when $r \ll s$.

Next, we analyze the suboptimality when applying controllers computed with $\hat{\Sigma}$. We will take the expanded and mode-reducible MJS $\bar{\Sigma}$ constructed with $\hat{\Sigma}$ and $\hat{\Omega}_{1:r}$ in Section 6.6 as a bridge. To begin with, similar to the notations for $\hat{\Sigma}$, we define $\bar{\varphi}_{1:s}, \bar{\mathcal{R}}_{1:s}$, $\bar{\mathcal{K}}_{1:s}$, $\bar{\mathbf{P}}_{1:s}$, and $\bar{\mathbf{K}}_{1:s}$. In terms of LQR solutions, the relation between $\hat{\Sigma}$ and $\bar{\Sigma}$ is given below.

Lemma 6.11. *Assume the Riccati solution $\bar{\mathbf{P}}_{1:s}$ exists and $\bar{\mathbf{P}}_i \succ 0$ for all i. Then, (i) there exists a unique Riccati solution $\hat{\mathbf{P}}_{1:r}$ in $\mathbb{S}^n_{r,+}$; (ii) $\hat{\mathbf{P}}_k = \bar{\mathbf{P}}_i, \hat{\mathbf{K}}_k = \bar{\mathbf{K}}_i$ for any $i \in \hat{\Omega}_k$ for any k.*

Proof. We consider the Riccati operator iteration defined as follows: $\bar{\mathbf{P}}_i^{(0)} = \mathbf{Q}$, $\bar{\mathbf{P}}_i^{(h+1)} = \bar{\mathcal{R}}_i(\bar{\mathbf{P}}_i^{(h)})$ for all $i \in [s], h \in \mathbb{N}$ and $\hat{\mathbf{P}}_k^{(0)} = \mathbf{Q}$, $\hat{\mathbf{P}}_i^{(h+1)} = \hat{\mathcal{R}}_i(\hat{\mathbf{P}}_i^{(h)})$ for all $k \in [r], h \in \mathbb{N}$. Then, note that by construction, for all $i \in \hat{\Omega}_k$ and all $l \in [r]$, we have $\sum_{j \in \hat{\Omega}_l} \bar{\mathbf{T}}(i,j) = \hat{\mathbf{T}}(k,l)$. Through induction and algebra, it is easy to show that for all $h \in \mathbb{N}$, and for any $i, i' \in \hat{\Omega}_k$

for any k, we have $\bar{\mathbf{P}}_i^{(h)} = \bar{\mathbf{P}}_{i'}^{(h)} = \hat{\mathbf{P}}_k^{(h)}$.

Since $\bar{\mathbf{P}}_i \succ 0$, by Fact B.4 in the appendix, we know $\bar{\mathbf{P}}_{1:s}$ is the unique solution among $\mathbb{S}_{s,+}^n$, and $\bar{\mathbf{K}}_{1:s}$ stabilizes $\bar{\Sigma}$. According to Costa et al. (2006, Proposition A.23), the stabilizability of $\bar{\Sigma}$ and the fact $\mathbf{Q}, \mathbf{R} \succ 0$ imply $\lim_{h \to \infty} \bar{\mathbf{P}}_i^{(h)} = \bar{\mathbf{P}}_i$. Combining this convergence result with the Riccati iteration results we just showed, we further have, for any $i, i' \in \hat{\Omega}_k$ and any k, we have $\bar{\mathbf{P}}_i = \bar{\mathbf{P}}_{i'} = \hat{\mathbf{P}}_k$. Then, it is easy to show that $\bar{\mathbf{K}}_i = \bar{\mathbf{K}}_{i'} = \hat{\mathbf{K}}_k$. The uniqueness of $\hat{\mathbf{P}}_{1:s}$ can be shown by contradiction. □

With this lemma, we have the following suboptimality guarantees in terms of applying controller $\hat{\mathbf{K}}_{1:r}$ to Σ.

Theorem 6.12 (LQR Suboptimality)**.** *Suppose A6.1 holds, and Σ has additive Gaussian noise $\mathcal{N}(0, \sigma_{\mathbf{w}}^2 \mathbf{I}_n)$ that is independent of the mode switching. Let $J^\star$ and $\hat{J}$ respectively denote the infinite time average cost incurred by the optimal controller $\mathbf{K}_{1:s}^\star$ and controller $\hat{\mathbf{K}}_{1:r}$ (at time t, $\mathbf{u}_t = \hat{\mathbf{K}}_k \mathbf{x}_t$ if $\omega_t \in \hat{\Omega}_k$). Then, there exists constants $\bar{\epsilon}_{\mathbf{A},\mathbf{B}}$, $\bar{\epsilon}_{\mathbf{T}}$, $C_{\mathbf{A},\mathbf{B}}$, and $C_{\mathbf{T}}$, such that when $\max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B}}$ and $\epsilon_{\mathbf{T}} \leq \bar{\epsilon}_{\mathbf{T}}$,*

$$\hat{J} - J^\star \leq \sigma_{\mathbf{w}}^2 (C_{\mathbf{A},\mathbf{B}} \max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} + C_{\mathbf{T}} \epsilon_{\mathbf{T}})^2 \tag{6.21}$$

Let $J_\infty^\star$ and $\hat{J}_\infty$ denote the infinite time cumulative cost incurred by $\mathbf{K}_{1:s}^\star$ and $\hat{\mathbf{K}}_{1:r}$ respectively. Then, when $\sigma_{\mathbf{w}} = 0$, $\epsilon_{\mathbf{T}} \leq \bar{\epsilon}_{\mathbf{T}}$, and $\max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B}}$,

$$\hat{J}_\infty - J_\infty^\star \leq (C'_{\mathbf{A},\mathbf{B}} \max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} + C'_{\mathbf{T}} \epsilon_{\mathbf{T}}) \|\mathbf{x}_0\|^2, \tag{6.22}$$

for some constants $C'_{\mathbf{A},\mathbf{B}}$ and $C'_{\mathbf{T}}$.

Proof. We will use Lemma 6.11 and $\bar{\Sigma}$ and $\bar{\mathbf{K}}_{1:s}$ as a bridge to compare $\hat{\mathbf{K}}_{1:r}$ and $\mathbf{K}_{1:s}^\star$. First, we prove (6.21). Comparing $\bar{\Sigma}$ and Σ, one can see $\|\bar{\mathbf{A}}_i - \mathbf{A}_i\| \leq \epsilon_{\mathbf{A}}$, $\|\bar{\mathbf{B}}_i - \mathbf{B}_i\| \leq \epsilon_{\mathbf{B}}$, and $\|\bar{\mathbf{T}} - \mathbf{T}\|_\infty \leq \epsilon_{\mathbf{T}}$. Then, from Theorem 4.1 we know when $\max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B}}$ and $\epsilon_{\mathbf{T}} \leq \bar{\epsilon}_{\mathbf{T}}$ for some constants $\bar{\epsilon}_{\mathbf{A},\mathbf{B}}$ and $\bar{\epsilon}_{\mathbf{T}}$, the Riccati solution $\bar{\mathbf{P}}_{1:s}$ uniquely exists among $\mathbb{S}_{s,+}^n$ and are positive definite, and the cost $\bar{J}$ when applying $\bar{\mathbf{K}}_{1:s}$ to Σ has suboptimality $\bar{J} - J^\star \leq \sigma_{\mathbf{w}}^2 (C_{\mathbf{A},\mathbf{B}} \max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} + C_{\mathbf{T}} \epsilon_{\mathbf{T}})$ for some constants $C_{\mathbf{A},\mathbf{B}}$ and $C_{\mathbf{T}}$. Using Lemma 6.11, we know $\hat{\mathbf{P}}_{1:r}$ uniquely exists $\mathbb{S}_{s,+}^n$, and $\hat{\mathbf{K}}_k = \bar{\mathbf{K}}_i$ for any i belonging to any $\hat{\Omega}_k$, which implies applying $\bar{\mathbf{K}}_{1:s}$ is equivalent to applying $\hat{\mathbf{K}}_{1:r}$ as in the theorem statement. Thus $\bar{J} = \hat{J}$, and $\hat{J} - J^\star = \bar{J} - J^\star \leq \sigma_{\mathbf{w}}^2 (C_{\mathbf{A},\mathbf{B}} \max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} + C_{\mathbf{T}} \epsilon_{\mathbf{T}})$.

Next, we prove (6.22). Similar as above, we let $\bar{J}_\infty$ denote the cumulative cost when applying $\bar{\mathbf{K}}_{1:s}$ to Σ, then we have $\hat{J}_\infty = \bar{J}_\infty$. From the proof of Costa et al. (2006, Theorem 4.5), we have $\bar{J}_\infty - J_\infty^\star = \sum_{t=0}^\infty \mathbb{E}[\|\mathbf{M}_{\omega_t} (\bar{\mathbf{K}}_{\omega_t} - \mathbf{K}_{\omega_t}^\star) \mathbf{x}_t\|^2]$ where $\mathbf{M}_i = \mathbf{R} + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}_{1:s}) \mathbf{B}_i$ and

$\mathbf{x}_t$ is driven by controller $\bar{\mathbf{K}}_{1:s}$. From Theorem 4.1, we know when $\max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B}}$ and $\epsilon_{\mathbf{T}} \leq \bar{\epsilon}_{\mathbf{T}}$, then $\bar{\mathbf{K}}_{1:s}$ is a stabilizing controller and $\|\bar{\mathbf{K}}_{1:s} - \mathbf{K}^\star_{1:s}\| \leq C^{\mathbf{K}}_{\mathbf{A},\mathbf{B}} \max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} + C^{\mathbf{K}}_{\mathbf{T}}\epsilon_{\mathbf{T}}$ for some constants $C^{\mathbf{K}}_{\mathbf{A},\mathbf{B}}$ and $C^{\mathbf{K}}_{\mathbf{T}}$. Following from Lemma 2.15, we know $\sum_{t=0}^{\infty} \mathbb{E}[\|\mathbf{x}_t\|^2] \leq C_{\mathbf{x}}\|\mathbf{x}_0\|^2$ for some constant $C_{\mathbf{x}}$. Combining these results, the suboptimality is bounded by $\hat{J}_\infty - J^\star_\infty \leq \|\mathbf{M}_{1:s}\| \|\bar{\mathbf{K}}_{1:s} - \mathbf{K}^\star_{1:s}\| \sum_{t=0}^{\infty} \mathbb{E}[\|\mathbf{x}_t\|^2] \leq \|\mathbf{M}_{1:s}\| C_{\mathbf{x}}(C^{\mathbf{K}}_{\mathbf{A},\mathbf{B}} \max\{\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}\} + C^{\mathbf{K}}_{\mathbf{T}}\epsilon_{\mathbf{T}})\|\mathbf{x}_0\|^2$. □

In Theorem 6.12, constants $\bar{\epsilon}_{\mathbf{A},\mathbf{B}}$, $\bar{\epsilon}_{\mathbf{T}}$, $C_{\mathbf{A},\mathbf{B}}$, $C_{\mathbf{T}}$, $C'_{\mathbf{A},\mathbf{B}}$, and $C'_{\mathbf{T}}$ only depend on the original MJS Σ and cost matrices $\mathbf{Q}$ and $\mathbf{R}$, and their exact expressions can be obtained following the proof and corresponding references. As a sanity check, when there is no perturbation, i.e., mode-reducible case, then we have $\hat{J} = J^\star$ and $\hat{J}_\infty = J^\star_\infty$, which can also be implied from Lemma 6.11. For the reduced MJS $\hat{\Sigma}$, its Riccati solution $\hat{\mathbf{P}}_{1:r}$ and thus controllers $\hat{\mathbf{K}}_{1:r}$ are guaranteed to exist when perturbation $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}$, and $\epsilon_{\mathbf{T}}$ are small enough as required in Theorem 6.12.

In the noisy case, both $\hat{J}_\infty$ and $J^\star_\infty$ are infinite, so the cumulative suboptimality $\hat{J}_\infty - J^\star_\infty$ is only studied for the noise-free case as in (6.22). On the other hand, in the noise-free case, we have not only $J^\star = \hat{J}$ as implied by (6.21), but also $J^\star = \hat{J} = 0$ as long as $\hat{\mathbf{K}}_{1:r}$ is stabilizing.

6.8 Experiments

In this section, we present synthetic experiments to evaluate the main results in this chapter. We evaluate the clustering performance of Algorithm 3 and the LQR controller designed with the reduced MJS $\hat{\Sigma}$ as discussed in Section 6.7. All the experiments are performed using MATLAB R2020a on a laptop with Xeon E3-1505M CPU. We use the kmeans() function from the Statistics and Machine Learning Toolbox in MATLAB for the k-means problem in Algorithm 3.

6.8.1 Clustering Evaluation

We consider the uniform partition $\Omega_{1:r}$, i.e. $|\Omega_i| = \bar{s} := s/r$ for any i. The system Σ is randomly generated according to P6.1 or P6.2 with desired levels of perturbation $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}}$ so that in (6.2) each summand $\|\mathbf{A}_i - \mathbf{A}_{i'}\| \leq \epsilon_{\mathbf{A}}/(r\bar{s}^2)$. The same applies to $\mathbf{B}_{1:s}$ and $\mathbf{T}$. Specifically, we first randomly generate a small scale MJS $\breve{\Sigma} = \text{MJS}(\breve{\mathbf{A}}_{1:r}, \breve{\mathbf{B}}_{1:r}, \breve{\mathbf{T}})$: we sample each matrix element in $\breve{\mathbf{A}}_k$ and $\breve{\mathbf{B}}_k$ from standard Gaussian distributions and then scale the matrices so that each $\|\breve{\mathbf{A}}_k\| = 0.5$ and $\|\breve{\mathbf{B}}_k\| = 1$ unless otherwise mentioned; and each $\breve{\mathbf{T}}(i, :)$ is sampled from the flat Dirichlet distribution. Then, we generate Σ by augmenting $\breve{\Sigma}$. For every mode $i \in \Omega_k$, we let $\mathbf{A}_i = \breve{\mathbf{A}}_k + \mathbf{E}_i$ and $\mathbf{B}_i = \breve{\mathbf{B}}_k + \mathbf{F}_i$ where we sample elements in $\mathbf{E}_i$ and $\mathbf{E}_i$ from standard Gaussian and then scale them so that $\|\mathbf{E}_i\|_{\mathrm{F}} = \frac{\epsilon_{\mathbf{A}}}{2r\bar{s}^2}$ and $\|\mathbf{F}_i\|_{\mathrm{F}} = \frac{\epsilon_{\mathbf{B}}}{2r\bar{s}^2}$. The generation of $\mathbf{T}$ is a bit involved. For the aggregatable case P6.2, we first generate a Markov matrix $\bar{\mathbf{T}} \in \mathbb{R}^{s \times s}$ such that for every $i \in \Omega_k$, $\bar{\mathbf{T}}(i, \Omega_l) = \mathbf{a}_{k,l}\breve{\mathbf{T}}(k,l)$ where $\mathbf{a}_{k,l} \in \mathbb{R}^{1 \times |\Omega_l|}$ is sampled from the flat Dirichlet distribution; then we let $\mathbf{T}(i,:) = (1 - \frac{\epsilon_{\mathbf{T}}}{2r\bar{s}^2})\bar{\mathbf{T}}(i,:) + \frac{\epsilon_{\mathbf{T}}}{2r\bar{s}^2}\mathbf{b}_i$ where $\mathbf{b}_i \in \mathbb{R}^{1 \times s}$ is again sampled from the flat Dirichlet distribution. The same steps are used to generate $\mathbf{T}$ for the lumpable case P6.1 except that $\bar{\mathbf{T}}(i, \Omega_l) = \mathbf{a}_{i,l}\breve{\mathbf{T}}(k,l)$. Following these steps, Σ satisfies the perturbation conditions in P6.1 and P6.2.

To evaluate Algorithm 3, we fix $n = 5$, $p = 3$, $r = 4$ and record the misclustering rate (MR) defined in Section 6.4.1 over 100 runs. Fig. 6.2 presents the clustering performances under different number of modes s and perturbations $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}$ and $\epsilon_{\mathbf{T}}$. In the plots, we normalized the perturbation on the x-axis by s^2 so that the trends under different s can be better visualized. This also follows from the experiment setup: each summand in (6.2) has $\|\mathbf{A}_i - \mathbf{A}_{i'}\| \leq \mathcal{O}(\epsilon_{\mathbf{A}}/s^2)$. It is clear that the clustering performance degrades with increasing s and perturbations. We can also observe that when the perturbation is small, there are no misclustered modes.

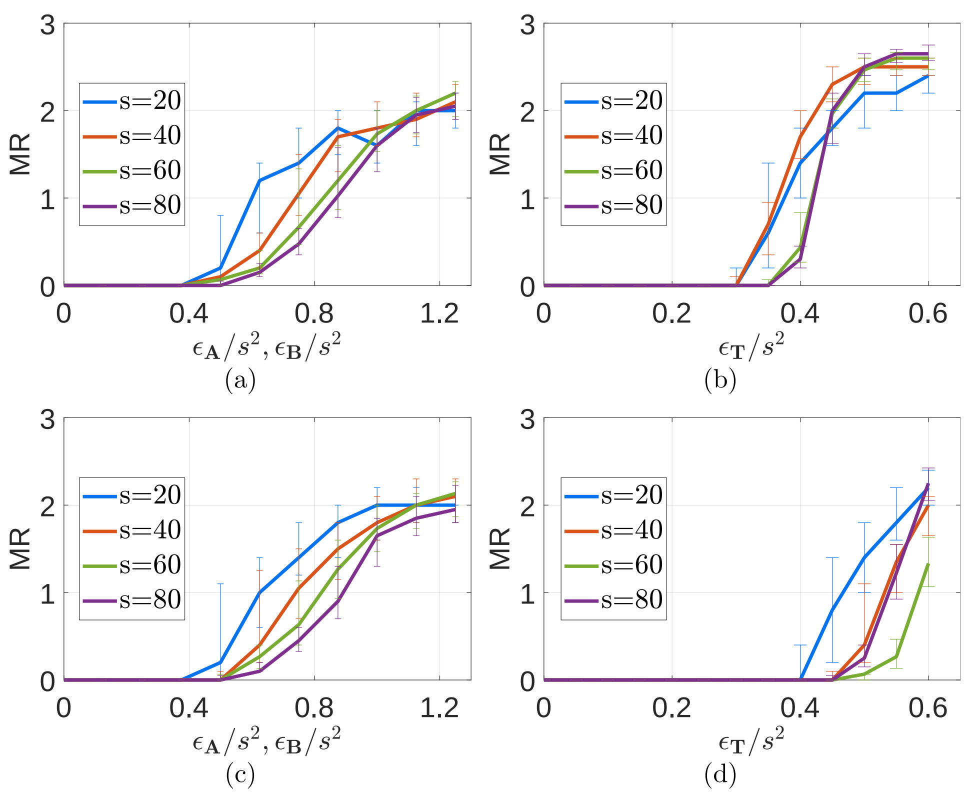


Figure 6.2: MJS mode reduction — clustering performance. MR (median, first and third quartiles) vs number of modes s vs perturbation level. First row: aggregatable case P6.2. Second row: lumpable case P6.1. First column: $\epsilon_{\mathbf{T}} = 0.5s^2$, $\alpha_{\mathbf{A}} \propto 1/\max_i\|\mathbf{A}_i\|$, $\alpha_{\mathbf{B}} \propto 1/\max_i\|\mathbf{B}_i\|$, $\alpha_{\mathbf{T}} \propto 0.01/\|\mathbf{T}\|$. Second column: $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}} = 0.5s^2$, $\alpha_{\mathbf{A}} \propto 0.01/\max_i\|\mathbf{A}_i\|$, $\alpha_{\mathbf{B}} \propto 0.01/\max_i\|\mathbf{B}_i\|$, $\alpha_{\mathbf{T}} \propto 1/\|\mathbf{T}\|$.

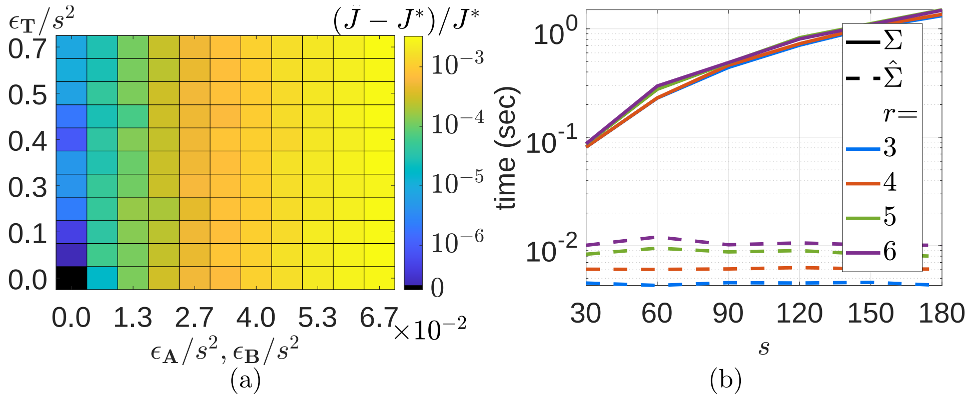


Figure 6.3: MJS reduction — LQR suboptimality. (a) LQR suboptimality (median) vs perturbations; (b) LQR computation time (median) for the original MJS Σ and the reduced MJS $\hat{\Sigma}$ with different number of modes and clusters. (We omit the quartiles as they are very close to the median.)

6.8.2 LQR Controller Design

Then, we implement the idea of designing LQR controllers for the original Σ through the reduced $\hat{\Sigma}$ as discussed in Section 6.7. We let $r = 4$, $n = 10$, $p = 5$, and the system dynamics is generated the same way as the previous section. The process noise variance is $\sigma_{\mathbf{w}}^2 = 0.1$, and the initial state is $\mathbf{x}_0 = \mathbf{1}$. Fig. 6.3a shows the suboptimality against perturbations for $s = 100$. As one would expect, the suboptimality increases with the perturbation levels and is 0 when there is no perturbation. The trend on $\epsilon_{\mathbf{T}}$ is evident when $\epsilon_{\mathbf{A}}$ and $\epsilon_{\mathbf{B}}$ are small but imperceptible for larger values of $\epsilon_{\mathbf{A}}$ and $\epsilon_{\mathbf{B}}$. Fig. 6.3b shows the time to compute controllers via Riccati iterations using Σ and $\hat{\Sigma}$ as a function of s. The computation terminates when the controller difference between two consecutive iterations falls below 10^{-12}. We see when s is large, Σ needs significantly more time than $\hat{\Sigma}$.

Next, we consider a more practical scenario where one has no knowledge of the true number of cluster r, and replace it in Algorithm 3 with a hyper-parameter $\hat{r}$ as the number of modes in $\hat{\Sigma}$. We fix $s = 100$ and $r = 30$, and the rest of the experiment setup is the same as Fig. 6.3. We record the suboptimality and computation time under different choices of $\hat{r}$ in Table 6.2. When increases $\hat{r}$, the suboptimality achieves the minimum when $\hat{r} = r = 30$ and then gradually increases until $\hat{r} = s$, i.e. no reduction is performed at all. This comes as a bit of surprise as one would expect no worse performance when using more clusters than needed. Further investigation suggests that when $\hat{r} > r$, misclustering occurs more frequently than the case of $\hat{r} = r$, which is likely to account for the performance degradation. In practice, to find the best $\hat{r}$, one could try multiple $\hat{r}$ in Algorithm 3, plug in the resulting partitions into P6.1 and P6.2, and select the one that gives the smallest perturbation $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}}$.

Table 6.2: Suboptimality vs computation time vs selected number of modes in $\hat{\Sigma}$

$\hat{r}$	10	20	30	40	50
$\frac{\hat{J}-J^\star}{J^\star}$	$1.2e{-1}$	$3.8e{-2}$	$4.1e{-7}$	$6.9e{-5}$	$8.9e{-4}$
Time (sec)	$3.7e{-2}$	$6.8e{-2}$	$1.4e{-1}$	$2.2e{-1}$	$3.2e{-1}$
$\hat{r}$	60	70	80	90	100
$\frac{\hat{J}-J^\star}{J^\star}$	$7.1e{-3}$	$1.6e{-2}$	$2.1e{-2}$	$1.5e{-2}$	0
Time (sec)	$6.8e{-1}$	$8.4e{-1}$	$1.0e0$	$1.1e0$	$1.2e0$

6.8.3 Trajectory Approximation

In this section, we evaluate the trajectory approximation results from Section 6.5. Let $\theta = \pi/16$, $\breve{\mathbf{A}}_1 = [[\cos(\theta), \sin(\theta)]^\intercal, [-\sin(\theta), \cos(\theta)]^\intercal]^\intercal$, $\breve{\mathbf{A}}_2 = [[0.8, 0]^\intercal, [0, 0.8]^\intercal]^\intercal$, and $\breve{\mathbf{A}}_3 = [[1.2, 0]^\intercal, [0, 1.2]^\intercal]^\intercal$. Then, we construct an autonomous MJS Σ with 6 modes: for $k = \{1, 2, 3\}$, $\mathbf{A}_{2k-1} = \breve{\mathbf{A}}_k + [[0.1, 0]^\intercal, [0, 0.1]^\intercal]^\intercal$ and $\mathbf{A}_{2k} = \breve{\mathbf{A}}_k - [[0.1, 0]^\intercal, [0, 0.1]^\intercal]^\intercal$. The uniform partition $\{\{1, 2\}, \{3, 4\}, \{5, 6\}\}$ gives $\epsilon_{\mathbf{A}} = 0.6\sqrt{2}$ according to P6.1. Define $\mathbf{T}$ such that for all i, $\mathbf{T}(i, j) = 0.2$ if $j \in \{1, 2, 3, 4\}$ and $\mathbf{T}(i, j) = 0.1$ if $j \in \{5, 6\}$. By relevant definitions in Section 6.5, the constructed Σ is mean-square stable but not uniformly stable.

We fix the initial state $\mathbf{x}_0 = [1, 1]^\intercal$, generate 500 independent trajectories for states $\mathbf{x}_t$ and $\hat{\mathbf{x}}_t$, and record the difference $\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|$. In Fig. 6.4, each thin solid line represents the difference, in log-scale, for each trajectory, the yellow dashed line shows their average, and the blue dashed line depicts the upper bound in Theorem 6.7. Throughout the time horizon, though not very tight, the theoretical upper bound stays above the averaged difference. Note that, for a given δ, by Markov inequality, shifting the upper bound in the plot upward by $\log(\delta)$ would give a bound on the individual error trajectories with probability $1 - \delta$. As seen in the figure, even the non-shifted version serves a good bound for individual error trajectories.

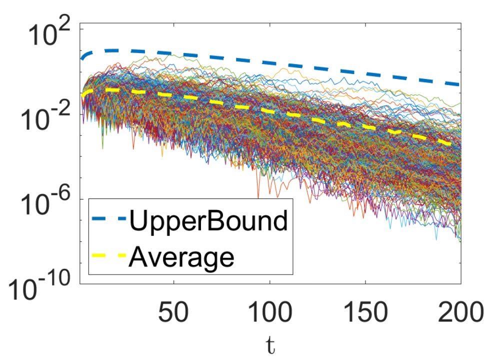


Figure 6.4: MJS reduction — trajectory difference and the upper bound

6.9 Conclusion

In this chapter, we propose a clustering-based method to reduce the number of modes in an MJS. The reduced MJS provably well approximates the original MJS in terms of trajectory, transition kernels, stability, and controller optimality. One future direction could be the generalization of the offline controller design scheme in Section 6.7 to settings where controllers need to be computed in the runtime, such as model predictive control and adaptive control. In these problems, the savings of computation time would be even more prominent. Another potential future direction could be the extension of the fully observed MJS in this chapter to partially observed MJS, i.e., the state $\mathbf{x}_t$ is observed through $\mathbf{y}_t = \mathbf{C}_{\omega_t}\mathbf{x}_t$ for some mode-dependent output matrices $\mathbf{C}_{1:s}$. As a side note, the similarity between MJS and Markov decision processes (MDP) hints that the framework and principles developed for MJS in this chapter may also help the complexity reduction of MDP and reinforcement learning problems.

Chapter 7
Conclusion and Future Work

7.1 Conclusion

Markov jump systems (MJSs) provide a systematic paradigm to model systems with time-varying dynamics. In this dissertation, several problems regarding MJSs are studied, including system identification, certainty equivalent control, adaptive quadratic control, and model reduction. We show how recent advances in statistics, online optimization, and machine learning can bring new solutions, insights, and analyses to these problems. Specifically, finite-sample analysis is developed for the identification of MJS dynamics using a single trajectory. Based on this, we propose a model-based adaptive control scheme to solve for LQR problems with unknown MJS dynamics and establish performance guarantees in terms of the regret. Finally, we show how clustering techniques from unsupervised learning can help reduce the mode complexity of an MJS.

We hope this dissertation could encourage and facilitate more studies in MJSs and the fusions between control and other fields. Though mode switching complicates the formulation and computation of many problems, MJSs have preserved solutions and guarantees as neat as LTI systems. Besides, the computational costs that scale with the number of modes may be brought down by our work on mode complexity reduction. In other words, there is almost no harm in choosing MJSs over LTI systems when it comes to modeling, especially when the underlying dynamics is time-varying. On the other hand, there have been many connections and intersections between control and learning problems, e.g., model order reduction vs. principal component analysis, system stability vs. gradient descent convergence, stochastic control vs. reinforcement learning. This dissertation pushes forward the frontiers by revisiting classical problems for MJSs. We believe this is a direction with many potentials and interesting problems, which would lead to a promised land. Some future directions are listed below.

7.2 Future Work

In this section, we list a few future directions that follow from the established work in this dissertation.

Identification with unobserved mode: When the mode ω_t in the MJS is unobserved, it poses significant challenges to the identification of an MJS. Without the knowledge of mode switching information ω_t, we can no longer collect the data generated by a specific mode and feed it to estimation algorithms as in Chapter 3.. In this case, one may need to estimate both the mode dynamics and mode switching sequence simultaneously. Similar problems have been studied for scalar switched autoregressive exogenous (SARX) systems in offline fashion (Ozay et al., 2011, 2015) using algebraic methods. Unlike the Markovian mode switching in MJS, the mode in SARX is typically assumed to be active for a minimum amount of dwell time before switching to a new mode. Thus, to adapt these approaches developed for SARX to MJS would require additional consideration of the underlying Markov chains. Another approach motivated by Bemporad et al. (2018) and our earlier work Du et al. (2018) is to alternate between dynamics and mode estimations:

- Repeat until convergence
  - given certain estimates of the MJS dynamics $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}$, we can use them to obtain mode sequence estimates $\hat{\omega}_{1:t}$ from the trajectory according to certain criteria, e.g. prediction error, likelihood.
  - given the the obtained mode sequence estimates $\hat{\omega}_{1:t}$, we can use them as the true mode sequence and apply Algorithm 1 to update the dynamics estimates $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}$.

In the ideal case, both the dynamics estimates $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}$ and mode sequence estimates $\hat{\omega}_{1:t}$ will converge to the ground truth ones. The initialization of the dynamics $\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}$ may be important as only local convergence is shown in Du et al. (2018) for SARX systems. Earlier works on time series clustering (Aghabozorgi et al., 2015) and recent advances in MJS mode filtering (Vergés and Fragoso, 2020) might be of help in the development of methodologies and analyses.

Learning a reduced-mode MJS: Given an unknown-dynamics MJS with a large number of modes, one can follow Chapter 6 to learn its dynamics from the data and then Chapter 6 to obtain a reduced-mode MJS. However, this two-step procedure may not be the most data-efficient approach if a reduced-mode MJS is the ultimate goal — for modes that are similar, data generated by them can be used collectively to learn a mode in the reduced

MJS. Hence, it would be interesting to develop a more principled method that can learn a reduced-mode MJS directly from the data. This problem involves estimating mode clustering and dynamics simultaneously, which shares ideas similar to the identification problem with unobserved mode discussed above. Both problems consist of estimation and clustering except that the former problem seeks to cluster the data while this problem aims at clustering the modes. Being able to solve one problem would surely bring insights to and shed light on the other.

Partially observed state in adaptive control: In practice, the state $\mathbf{x}_t$ in the MJS may only be observed through $\mathbf{y}_t = \mathbf{C}_{\omega_t}\mathbf{x}_t + \mathbf{v}_t$ where the observation $\mathbf{y}_t \in \mathbb{R}^m$ is generated by the mode-dependent output matrices $\mathbf{C}_{1:s}$ subject to measurement noise $\mathbf{v}_t$. Since direct feedback using the ground truth state $\mathbf{x}_t$ is no longer possible, the optimal control problems now require online estimation of the state $\mathbf{x}_t$. This is known as the filtering problem and can be solved via the Kalman filters. Extending the adaptive control framework in Chapter 5 to this model would require additional effort to study the perturbation of the state filtering. This part can benefit from the established perturbation results in Chapter 4 since the Kalman filters also involve solving for coupled Riccati equations.

Appendix A
Proofs for Results in Chapter 3

A.1 Supporting Lemmas

In this section, we provide a list of lemmas that will be useful for the subsequent proofs.

Lemma A.1. *Suppose* $\mathbf{z} \sim \mathcal{N}(0, \mathbf{\Sigma_z})$ *with* $\mathbf{\Sigma_z} \in \mathbb{R}^{p\times p}$*. For any* $t \geq (3+2\sqrt{2})p$*, we have*

$$\mathbb{P}(\|\mathbf{z}\|^2 \geq 3\|\mathbf{\Sigma_z}\|t) \leq e^{-t}.$$

Proof. From Hsu et al. (2012, Proposition 1), we have for any $t > 0$,

$$\mathbb{P}(\|\mathbf{z}\|^2 \geq \mathbf{tr}(\mathbf{\Sigma_z}) + 2\sqrt{\mathbf{tr}(\mathbf{\Sigma_z^2})t} + 2\|\mathbf{\Sigma_z}\|t) \leq e^{-t},$$

which implies

$$\mathbb{P}(\|\mathbf{z}\|^2 \geq p\|\mathbf{\Sigma_z}\| + 2\sqrt{p}\|\mathbf{\Sigma_z}\|\sqrt{t} + 2\|\mathbf{\Sigma_z}\|t) \leq e^{-t}.$$

We can see that when $t \geq (3+2\sqrt{2})p$, we have $p + 2\sqrt{p}\sqrt{t} \leq t$, which implies $p\|\mathbf{\Sigma_z}\| + 2\sqrt{p}\|\mathbf{\Sigma_z}\|\sqrt{t} \leq \|\mathbf{\Sigma_z}\|t$. Therefore, we have $\mathbb{P}(\|\mathbf{z}\|^2 \geq 3\|\mathbf{\Sigma_z}\|t) \leq e^{-t}$. □

Lemma A.2. *Let* $\mathbf{x}_t$ *be the MJS state and define the noise-removed state* $\tilde{\mathbf{x}}_t = \mathbf{x}_t - \mathbf{w}_{t-1}$ *which is independent of* $\mathbf{w}_{t-1}$*. Let* $\mathbb{E}[\|\tilde{\mathbf{x}}_t\|] \leq B$ *and* $\mathbf{w}_t$ *has i.i.d. entries with variance* $\sigma_\mathbf{w}^2$ *bounded in absolute value by* $c_\mathbf{w}\sigma_\mathbf{w}$ *for some* $c_\mathbf{w} > 0$*. Consider the conditional random vector*

$$\mathbf{y}_t \sim \{\mathbf{x}_t \mid \|\mathbf{x}_t\| \leq 3B,\ \omega_t = i\}.$$

If $c_\mathbf{w}\sigma_\mathbf{w}\sqrt{n} \leq B$*, then* $\mathbb{E}[\mathbf{y}_t\mathbf{y}_t^\intercal] \succeq \sigma_\mathbf{w}^2\mathbf{I}_n/2$*.*

Proof. Observe that $\|\mathbf{w}_t\| \leq c_\mathbf{w}\sigma_\mathbf{w}\sqrt{n} \leq B$. Define the events

$$E_1 = \{\|\mathbf{x}_t\| \leq 3B\},\ E_2 = \{\|\tilde{\mathbf{x}}_t\| \leq 2B\}. \tag{A.1}$$

Clearly $E_2 \subset E_1$ as $\|\mathbf{w}_{t-1}\| \le B$ and on the event E_2, $\tilde{\mathbf{x}}_t$ and $\mathbf{w}_{t-1}$ are independent. Now, observe that

$$\mathbb{E}[\mathbf{y}_t\mathbf{y}_t^\intercal] \succeq \mathbb{E}[\mathbf{y}_t\mathbf{y}_t^\intercal \mid E_2]\mathbb{P}(E_2). \tag{A.2}$$

Note that $\mathbb{P}(E_2) \ge 1/2$ from Markov bound as $\mathbb{E}[\|\tilde{\mathbf{x}}_t\|] \le B$. Additionally, on the event E_2, $\tilde{\mathbf{x}}_t$ and $\mathbf{w}_{t-1}$ are independent. Similarly, ω_t and $\mathbf{w}_{t-1}$ are also independent. Thus, we further have

$$\begin{aligned}
\mathbb{E}[\mathbf{y}_t\mathbf{y}_t^\intercal] &\succeq \mathbb{E}[\mathbf{y}_t\mathbf{y}_t^\intercal \mid E_2]\mathbb{P}(E_2), \\
&= \mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal \mid \|\tilde{\mathbf{x}}_t\| \le 2B,\ \omega_t = i]\mathbb{P}(\|\tilde{\mathbf{x}}_t\| \le 2B), \\
&\succeq \mathbb{E}[\mathbf{w}_{t-1}\mathbf{w}_{t-1}^\intercal \mid \|\tilde{\mathbf{x}}_t\| \le 2B,\ \omega_t = i]\mathbb{P}(\|\tilde{\mathbf{x}}_t\| \le 2B), \\
&\succeq (1/2)\mathbb{E}[\mathbf{w}_{t-1}\mathbf{w}_{t-1}^\intercal] && \text{(A.3)} \\
&\succeq \sigma_{\mathbf{w}}^2\mathbf{I}_n/2. && \text{(A.4)}
\end{aligned}$$

This completes the proof. ☐

Lemma A.3. *Let* $\mathbf{z} \sim \mathcal{N}(0, \sigma_{\mathbf{z}}^2\mathbf{I}_p)$*. Consider the conditional random vector* $\mathbf{y} \sim \{\mathbf{z} \mid \|\mathbf{z}\| \le c\sigma_{\mathbf{z}}\sqrt{p}\}$*, where* $c \ge 6$ *is a fixed constant. Then* $\mathbb{E}[\mathbf{y}\mathbf{y}^\intercal] \succeq \sigma_{\mathbf{z}}^2\mathbf{I}_p/2$*.*

Proof. This proof gives a lower bound on the covariance of truncated Gaussian vector $\mathbf{z} \mid \|\mathbf{z}\| \le c\sigma_{\mathbf{z}}\sqrt{p}$. Note that, $\mathbf{z}' = \mathbf{z}/\sigma_{\mathbf{z}}$ is $\mathcal{N}(0, \mathbf{I}_p)$. Set variable $X = \|\mathbf{z}'\|^2$. We have the following Lipschitz Gaussian tail bound (we use Lipschitzness of the ℓ_2 norm and use minor calculus and relaxations)

$$\mathbb{P}(\|\mathbf{z}'\|^2 \ge 4tp) \le \begin{cases} 1 \quad \text{if} \quad t \le 1 \\ e^{-tp/2} \quad \text{if} \quad t \ge 1. \end{cases}$$

This implies the following tail bound for X

$$Q(t) = \mathbb{P}(X \ge t) \le \begin{cases} 1 \quad \text{if} \quad t \le 4p \\ e^{-t/8} \quad \text{if} \quad t \ge 4p. \end{cases}$$

Fix $\kappa \ge 4$. Using integration-by-parts, this implies that

$$\begin{aligned}
\mathbb{E}[X \mid X \ge \kappa p]\mathbb{P}(X \ge \kappa p) &= -\int_{\kappa p}^{\infty} x dQ(x) = -[xQ(x)]_{\kappa n}^{\infty} + \int_{\kappa p}^{\infty} Q(x)dx, \\
&\le (\kappa p + 8)e^{-\kappa p/8}. && \text{(A.5)}
\end{aligned}$$

The final line as a function of κp is decreasing when $\kappa p \geq 36$. Specifically it is upper bounded by $1/2$ when $\kappa \geq 36$ (as $p \geq 1$). Now define the event

$$E_z = \{\|\mathbf{z}\| \leq \sigma_{\mathbf{z}}\sqrt{\kappa}\sqrt{p}\}.$$

For $\sqrt{\kappa} \geq 6$, $\sqrt{\kappa}$ will map to the c in the statement of the lemma. Observe that this is also the event $X \leq \kappa p$. Following (A.5), this implies

$$\begin{aligned}\mathbb{E}[\|\mathbf{z}\|^2 \mid E_z^c]\mathbb{P}(E_z^c) &\leq \mathbb{E}[\sigma_{\mathbf{z}}^2 X \mid E_z^c]\mathbb{P}({E_z}^c) \\ &\leq \sigma_{\mathbf{z}}^2\mathbb{E}[X \mid E_z^c]\mathbb{P}(E_z^c) \\ &\leq \sigma_{\mathbf{z}}^2/2.\end{aligned}$$

This also yields the covariance bound of the tail event

$$\mathbb{E}[\mathbf{z}\mathbf{z}^\intercal \mid E_z^c]\mathbb{P}(E_z^c) \preceq \mathbb{E}[\|\mathbf{z}\|^2\mathbf{I}_p \mid E_z^c]\mathbb{P}(E_z^c) \preceq \sigma_{\mathbf{z}}^2\mathbf{I}_p/2.$$

Finally, from the conditional decomposition, observe that

$$\mathbb{E}[\mathbf{z}\mathbf{z}^\intercal] = \mathbb{E}[\mathbf{z}\mathbf{z}^\intercal \mid E_z^c]\mathbb{P}(E_z^c) + \mathbb{E}[\mathbf{z}\mathbf{z}^\intercal \mid E_z]\mathbb{P}(E_z) \implies \mathbb{E}[\mathbf{z}\mathbf{z}^\intercal \mid E_z]\mathbb{P}(E_z) \succeq \sigma_{\mathbf{z}}^2\mathbf{I}_p/2.$$

To conclude, observe that $\mathbb{E}[\mathbf{z}\mathbf{z}^\intercal \mid E_z] = \mathbb{E}[\mathbf{y}\mathbf{y}^\intercal]$, where $\mathbf{y}$ is the conditional vector defined by truncating $\mathbf{z}$. Thus, we found

$$\mathbb{E}[\mathbf{y}\mathbf{y}^\intercal]\mathbb{P}(E_z) \succeq \sigma_{\mathbf{z}}^2\mathbf{I}_p/2 \implies \mathbb{E}[\mathbf{y}\mathbf{y}^\intercal] \succeq \sigma_{\mathbf{z}}^2\mathbf{I}_p/2.$$

□

Theorem A.4. *(Vershynin, 2012, Theorem 5.41 (Isotropic)) Let $\mathbf{X}$ be an $N \times d$ matrix whose rows $\mathbf{x}_i \in \mathbb{R}^d$ are independent isotropic. Let m be such that $\|\mathbf{x}_i\| \leq \sqrt{m}$ almost surely for all $i \in [N]$. Then, for every $t \geq 0$, with probability $1 - 2d \cdot e^{-ct^2}$, we have*

$$\sqrt{N} - t\sqrt{m} \leq \underline{\sigma}(\mathbf{X}) \leq \|\mathbf{X}\| \leq \sqrt{N} + t\sqrt{m}.$$

Corollary A.5 (Non-isotropic)**.** *Let $\mathbf{X}$ be an $N \times d$ matrix whose rows $\mathbf{x}_i \in \mathbb{R}^d$ are independent with covariance $\mathbf{\Sigma}_i$. Suppose each covariance obeys*

$$\sigma_{\min}^2 \leq \underline{\sigma}(\mathbf{\Sigma}_i) \leq \|\mathbf{\Sigma}_i\| \leq \sigma_{\max}^2.$$

Let m be such that $\|\mathbf{x}_i\| \leq \sqrt{m}$ almost surely for all $i \in [N]$. Then, for every $t \geq 0$, with *probability* $1 - 2d \cdot e^{-ct^2}$*, we have*

$$\sigma_{\min}\sqrt{N} - t\sqrt{m} \le \underline{\sigma}(\mathbf{X}) \le \|\mathbf{X}\| \le \sigma_{\max}\sqrt{N} + t\frac{\sigma_{\max}}{\sigma_{\min}}\sqrt{m}.$$

Proof. Let $\mathbf{x}_i' = \mathbf{\Sigma}_i^{-1/2}\mathbf{x}_i$. Observe that $\mathbf{x}_i'$ are independent isotropic. Define the matrix $\mathbf{X}'$ with rows $\mathbf{x}'$. Note that $\|\mathbf{x}_i'\| \le \|\mathbf{x}_i\|/\sigma_{\min} \le \sigma_{\min}^{-1}\sqrt{m}$. Thus, applying Theorem A.4 on $\mathbf{X}'$, for every $t \ge 0$, with probability $1 - 2d \cdot e^{-ct^2}$, we have

$$\sqrt{N} - t\sigma_{\min}^{-1}\sqrt{m} \le \underline{\sigma}(\mathbf{X}') \le \|\mathbf{X}'\| \le \sqrt{N} + t\sigma_{\min}^{-1}\sqrt{m}. \tag{A.6}$$

Next, observing that $\mathbf{X}^\intercal\mathbf{X} = \sum_{i=1}^N \mathbf{x}_i\mathbf{x}_i^\intercal = \sum_{i=1}^N \sqrt{\mathbf{\Sigma}_i}\mathbf{x}_i'\mathbf{x}_i'^\intercal\sqrt{\mathbf{\Sigma}_i}$, we find that

$$\begin{aligned}\sigma_{\min}^2\mathbf{X}'^\intercal\mathbf{X}' = \sigma_{\min}^2\sum_{i=1}^N \mathbf{x}_i'\mathbf{x}_i'^\intercal \preceq \mathbf{X}^\intercal\mathbf{X} &= \sum_{i=1}^N \sqrt{\mathbf{\Sigma}_i}\mathbf{x}_i'\mathbf{x}_i'^\intercal\sqrt{\mathbf{\Sigma}_i} \\ &\preceq \sigma_{\max}^2\sum_{i=1}^N \mathbf{x}_i'\mathbf{x}_i'^\intercal = \sigma_{\max}^2\mathbf{X}'^T\mathbf{X}',\end{aligned}$$

which implies that

$$\sigma_{\min}\underline{\sigma}(\mathbf{X}') \le \underline{\sigma}(\mathbf{X}) \le \|\mathbf{X}\| \le \sigma_{\max}\|\mathbf{X}'\|.$$

Plugging this into (A.6) completes the proof. □

We first list in Table A.1 a few shorthand notations to be used in this appendix. They are mainly used in the fictional sub-trajectories analysis in Appendices A.3 and A.4. Notations on the inside the parentheses are arguments to be replaced with context-depending variables.

A.2 Estimating T

The following theorem adapted from Zhang and Wang (2019, Lemma 7) provides the sample complexity result for estimating Markov matrix $\mathbf{T}$, which is a corresponds to the sample complexity on $\|\hat{\mathbf{T}} - \mathbf{T}\|$ in Theorem 3.1.

Theorem A.6. *Suppose we have an ergodic Markov chain* $\mathbf{T} \in \mathbb{R}^{s\times s}$ *with mixing time* t_{MC} *and stationary distribution* $\boldsymbol{\pi}_\infty \in \mathbb{R}^s$. *Let* $\pi_{\max} := \max_{i\in[s]} \boldsymbol{\pi}_\infty(i)$ *and* $\pi_{\min} := \min_{i\in[s]} \boldsymbol{\pi}_\infty(i)$. *Given a state sequence* $\omega_0, \omega_1, \ldots, \omega_T$ *of the Markov chain, define the empirical estimator* $\hat{\mathbf{T}}$ *of the Markov matrix as follows,*

$$\hat{\mathbf{T}}(i,j) = \frac{\sum_{t=1}^{T-1}\mathbf{1}_{\{\omega_t=i,\omega_{t+1}=j\}}}{\sum_{t=1}^{T-1}\mathbf{1}_{\{\omega_t=i\}}},$$

Table A.1: Notations — Sampling Periods

$\underline{c}_{\mathbf{w}}(T,\delta)$	$\sqrt{2\log(nT)}+\sqrt{2\log(2/\delta)}$
$\beta_+(\rho,\tau,c_{\mathbf{w}})$	$\sqrt{\frac{2\sqrt{s}(c_{\mathbf{w}}^2+C_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|^2)\tau}{1-\rho}}$
$\beta'_+(\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s})$	$c_{\mathbf{w}}+\beta_+(\rho,\tau,c_{\mathbf{w}})(\|\mathbf{A}_{1:s}\|+\|\mathbf{B}_{1:s}\|\|\mathbf{K}_{1:s}\|)+C_{\mathbf{z}}\sqrt{p/n}\|\mathbf{B}_{1:s}\|$
$\underline{C}_{sub,\mathbf{x}}(\bar{x}_0,\delta,T,\rho,\tau)$	$\frac{2}{\log(\rho^{-1})}+\frac{2}{\log(\rho^{-1})\log(T)}\log\left(\frac{24n\sqrt{s}\tau\max\{\bar{x}_0^2,\bar{\sigma}^2\}}{\delta(1-\rho)}\right)\Big\}$
$\underline{C}_{sub,\bar{\mathbf{x}}}(\delta,T,\rho,\tau)$	$\frac{1}{\log(\rho^{-1})}+\frac{1}{\log(\rho^{-1})\log(T)}\log(\frac{72\sqrt{n}s^{1.5}\tau}{\delta})$
$\underline{C}_{sub,N}(\bar{x}_0,\delta,T,\rho,\tau)$	$\max\left\{C_{MC},\underline{C}_{sub,\mathbf{x}}(\bar{x}_0,\frac{\delta}{2},T,\rho,\tau),\underline{C}_{sub,\bar{\mathbf{x}}}(\delta,T,\rho,\tau)\right\}$
$\underline{L}_{id,t_0}(\bar{x}_0,\rho,\tau,c_{\mathbf{w}})$	$\frac{\log\left((1-\rho)\bar{x}_0^2/(c_{\mathbf{w}}^2\bar{\sigma}_{\mathbf{w}}^2+\bar{\sigma}_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|)\right)}{1-\rho}$
$\underline{L}_{id,cov}(\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s})$	$1+\frac{2\log\left(8c^2\beta_+(\rho,\tau,c_{\mathbf{w}})\beta'_+(\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s})n\sqrt{ns}\tau\right)}{1-\rho}$
$\underline{L}_{id,tr1}(\delta,T,\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s})$	$1+\frac{2\log\left(2\sqrt{ns}\tau T\beta'_+(\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s})/(\beta_+(\rho,\tau,c_{\mathbf{w}})\delta)\right)}{1-\rho}$
$\underline{L}_{id,tr2}(\delta,T,\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s})$	$1+\frac{2}{(1-\rho)}\log\left(\frac{192c^2\tau\beta'_+(\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s})(1+\beta_+(\rho,\tau,c_{\mathbf{w}}))n\sqrt{s(n+p)}T}{\pi_{\min}\delta}\right)$
$\underline{L}_{id,tr3}(\delta,T,\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s})$	$1+\frac{2}{(1-\rho)}\log\left(\frac{c_{\mathbf{w}}\sigma_{\mathbf{w}}\beta'_+(\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s})\tau n\sqrt{ns}T^2}{\delta(1+\beta_+(\rho,\tau,c_{\mathbf{w}}))\sqrt{(n+p)}\left(C\sigma_{\mathbf{w}}\sqrt{n+p}+C_0\sqrt{\log(2s/\delta)}\right)}\right)$
$\underline{L}_{id}(\bar{x}_0,\delta,T,\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s},L)$	$\max\left\{\underline{L}_{id,t_0}(\bar{x}_0,\rho,\tau,c_{\mathbf{w}}),\underline{L}_{id,cov}(\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s}),\right.$ $\underline{L}_{id,tr1}(\frac{\delta}{36L},T,\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s}),\underline{L}_{id,tr2}(\frac{\delta}{36L},T,\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s}),$ $\left.\underline{L}_{id,tr3}(\frac{\delta}{36L},T,\rho,\tau,c_{\mathbf{w}},\mathbf{K}_{1:s}),\underline{C}_{sub,N}\left(\bar{x}_0,\frac{\delta}{2L},T,\rho,\tau\right)\log(T)\right\}$

Assume for some $\delta>0$, $T\geq\underline{T}_{MC,1}(C_{MC},\frac{\delta}{4}):=\left(68C_{MC}\pi_{\max}\pi_{\min}^{-2}\log(\frac{4s}{\delta})\right)^2$, *where* C_{MC} *is defined in Table C.1. Then, we have with probability at least* $1-\delta$,

$$\|\hat{\mathbf{T}}-\mathbf{T}\|\leq 4\pi_{\min}^{-1}\|\mathbf{T}\|\sqrt{\frac{17\pi_{\max}C_{MC}\log(T)}{T}\log\left(\frac{4sC_{MC}\log(T)}{\delta}\right)}. \tag{A.7}$$

Proof. We first consider estimators computed using a sub-trajectory of $\omega_0,\omega_1,\ldots,\omega_T$, then combine them together to show the error bound for $\hat{\mathbf{T}}$ in the claim. For C_{MC} defined in Table C.1, let $L=C_{MC}\log(T)$. Then, for $\ell=0,1,\ldots,L-1$, define $\hat{\mathbf{T}}^{(\ell)}\in\mathbb{R}^{s\times s}$ such that $[\hat{\mathbf{T}}^{(\ell)}](i,j)=\frac{\sum_{k=1}^{\lfloor T/L\rfloor}\mathbf{1}_{\{\omega_{kL+\ell}=i,\omega_{kL+1+\ell}=j\}}}{\sum_{k=1}^{\lfloor T/L\rfloor}\mathbf{1}_{\{\omega_{kL+\ell}=i\}}}$. In other words, $\hat{\mathbf{T}}^{(\ell)}$ is the estimator computed using data with sub-sampling period L. Following the proof of (Zhang and Wang, 2019, Lemma 7), we know for any $\epsilon<\pi_{\min}/2$, suppose $L\geq 6t_{\mathrm{MC}}\log(\epsilon^{-1})$.

$$\mathbb{P}\left(\|\hat{\mathbf{T}}^{(\ell)}-\mathbf{T}\|\leq 4\pi_{\min}^{-1}\|\mathbf{T}\|\epsilon\right)\geq 1-4s\exp\left(-\frac{T\epsilon^2}{17\pi_{\max}L}\right). \tag{A.8}$$

By setting $\delta=4s\exp\left(-\frac{T\epsilon^2}{17\pi_{\max}L}\right)$, one can also interpret the above result as: for all $\delta>0$, suppose

$$L \geq 3t_{\mathrm{MC}} \log\left(\frac{T}{17\pi_{\max} L \log(\frac{4s}{\delta})}\right), \tag{A.9}$$

then when

$$T \geq 68L\pi_{\max}\pi_{\min}^{-2}\log(\frac{4s}{\delta}), \tag{A.10}$$

we have with probability at least $1-\delta$

$$\|\hat{\mathbf{T}}^{(\ell)} - \mathbf{T}\| \leq 4\pi_{\min}^{-1}\|\mathbf{T}\|\sqrt{\frac{17\pi_{\max}C_{MC}\log(T)\log(\frac{4s}{\delta})}{T}}. \tag{A.11}$$

One can verify (A.9) holds by plugging in $L = C_{MC}\log(T)$ and using definition $C_{MC} := t_{\mathrm{MC}} \cdot \max\{3, 3-3\log(\pi_{\max}\log(s))\}$; (A.10) holds under the premise condition $T \geq \underline{T}_{MC,1}(C_{MC}, \frac{\delta}{4}) := \left(68C_{MC}\pi_{\max}\pi_{\min}^{-2}\log(\frac{4s}{\delta})\right)^2$.

Note that by definition, $\hat{\mathbf{T}}$ can be viewed as a convex combination of $\hat{\mathbf{T}}^{(\ell)}$ for all $\ell = 0, 1, \ldots, L$, thus by triangle inequality and union bound, we have with probability $1 - L\delta$,

$$\|\hat{\mathbf{T}} - \mathbf{T}\| \leq 4\pi_{\min}^{-1}\|\mathbf{T}\|\sqrt{\frac{17\pi_{\max}C_{MC}\log(T)\log(\frac{4s}{\delta})}{T}}. \tag{A.12}$$

Finally, by replacing $L\delta$ with δ, we could show (A.7) and conclude the proof. □

A.3 Estimation of $\mathbf{A}_{1:s}$ and $\mathbf{B}_{1:s}$ from Single Trajectory (Main SYSID Analysis)

A.3.1 Architecture of the proof

Let $\mathbf{h}_t := [\mathbf{x}_t^\intercal/\sigma_{\mathbf{w}} \ \mathbf{z}_t^\intercal/\sigma_{\mathbf{z}}]^\intercal$ be the (scaled) concatenated state vector and define $\mathbf{\Theta}_{\omega_t}^\star := [\sigma_{\mathbf{w}}\mathbf{L}_{\omega_t} \ \sigma_{\mathbf{z}}\mathbf{B}_{\omega_t}]$. Then, the closed-loop MJS is governed by the following state equation,

$$\mathbf{x}_{t+1} = \mathbf{\Theta}_{\omega_t}^\star \mathbf{h}_t + \mathbf{w}_t \quad \text{s.t.} \quad \omega_t \sim \text{Markov Chain}(\mathbf{T}). \tag{A.13}$$

Hence, to estimate $\mathbf{\Theta}_i^\star$, for each $i \in [s]$, we sub-sample the bounded samples corresponding to $\omega_t = i$ from the given MJS trajectory $\{\mathbf{x}_t, \mathbf{z}_t, \omega_t\}_{t=0}^T$ to obtain s sub-trajectories $\{(\mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \omega_t)\}_{t\in S_i}$, for $i = 1, \ldots, s$ and solve s regression problems given by $\hat{\mathbf{\Theta}}_i = (\arg\min_{\mathbf{\Theta}_i} \|\mathbf{Y}_i - \mathbf{H}_i\mathbf{\Theta}_i^\intercal\|_{\mathrm{F}}^2)^\intercal$, where the rows of $\mathbf{Y}_i$ and $\mathbf{H}_i$ are $\{\mathbf{x}_{t+1}^\intercal\}_{t\in S_i}$ and $\{\mathbf{h}_t^\intercal\}_{t\in S_i}$, respectively. Similarly, the rows of $\mathbf{W}_i$ are $\{\mathbf{w}_t^\intercal\}_{t\in S_i}$. Then, the estimation error of the least-squares estimator can be bounded as $\|\hat{\mathbf{\Theta}}_i - \mathbf{\Theta}_i^\star\| \leq \|\mathbf{H}_i^\intercal\mathbf{W}_i\|/\lambda_{\min}(\mathbf{H}_i^\intercal\mathbf{H}_i)$. Since $\mathbf{H}_i$ has non-i.i.d. rows, it is therefore not straightforward to upper bound $\|\mathbf{H}_i^\intercal \mathbf{W}_i\|$ and lower bound $\lambda_{\min}(\mathbf{H}_i^\intercal \mathbf{H}_i)$ directly. To resolve this issue, we rely on the assumption of mean-square stability and use perturbation-based techniques to indirectly bound these terms. For the ease of analysis, we first derive our estimation error bounds with **Assumption A1.1** which assumes bounded noise, that is, we have $\|\mathbf{w}_t\|_\infty \leq c_\mathbf{w}\sigma_\mathbf{w}$ for some $c_\mathbf{w} > 0$. Later on, we will remove this assumption to extend our results to the Gaussian noise. The main ingredients of our proof are as follows:

- **Definition A.8** splits each sub-trajectory $\{(\mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \omega_t)\}_{t\in S_i}$ into L sub-trajectories $\{(\mathbf{x}_{\ell_k+1}, \mathbf{x}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$ through shifting by $0 \leq \ell \leq L-1$ and sub-sampling by $L \geq 1$, that is, given the set $S_i = \big\{t \mid \omega_t = i,\ \|\mathbf{x}_t\| \leq \mathcal{O}(\sigma_\mathbf{w}\sqrt{n\log(T)}),\ \|\mathbf{z}_t\| \leq \mathcal{O}(\sigma_\mathbf{z}\sqrt{p})\big\}$ from Algorithm 1, we split it into L sub-sets, defined as $S_i^{(\ell)} := \big\{\ell_k := \ell + kL \mid \omega_{\ell_k} = i,\ \|\mathbf{x}_{\ell_k}\| \leq \mathcal{O}(\sigma_\mathbf{w}\sqrt{n\log(T)}),\ \|\mathbf{z}_{\ell_k}\| \leq \mathcal{O}(\sigma_\mathbf{z}\sqrt{p})\big\}$, where $k = 1, 2, \dots, \lfloor \frac{T-L}{L} \rfloor$ and $S_i = \bigcup_{\ell=0}^{L-1} S_i^{(\ell)}$.

- **Lemma A.9** bounds the covariance matrix of the states belonging to the sub-trajectory $\{(\mathbf{x}_{\ell_k+1}, \mathbf{x}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$.

- Using **Definition A.8** and triangle inequality, we have $\|\mathbf{H}_i^\intercal \mathbf{W}_i\| \leq \sum_{\ell=0}^{L-1} \|\mathbf{H}_i^{(\ell)\intercal} \mathbf{W}_i^{(\ell)}\|$. Similarly, using **Definition A.8** along-with Weyl's inequality for Hermitian matrices, we have $\lambda_{\min}(\mathbf{H}_i^\intercal \mathbf{H}_i) \geq \sum_{\ell=0}^{L-1} \lambda_{\min}\big(\mathbf{H}_i^{(\ell)\intercal} \mathbf{H}_i^{(\ell)}\big)$, where the rows of $\mathbf{H}_i^{(\ell)}$ and $\mathbf{W}_i^{(\ell)}$ are $\{\mathbf{h}_{\ell_k}^\intercal\}_{\ell_k \in S_i^{(\ell)}}$ and $\{\mathbf{w}_{\ell_k}^\intercal\}_{\ell_k \in S_i^{(\ell)}}$, respectively. Hence, we can upper bound the estimation error by upper bounding $\|\mathbf{H}_i^{(\ell)\intercal} \mathbf{W}_i^{(\ell)}\|$ and lower bounding $\lambda_{\min}\big(\mathbf{H}_i^{(\ell)\intercal} \mathbf{H}_i^{(\ell)}\big)$ for each $0 \leq \ell \leq L-1$.

- Observe that the rows of $\mathbf{H}_i^{(\ell)}$ are still not independent. However, if we choose L to be large ($\hat{\mathcal{O}}(\log(T))$), then they are weakly dependent (conditioned on the modes) due to mean-square stability. Therefore, for the purposes of analysis, for each state $\mathbf{x}_{\ell_k}$, where $\ell_k \in S_i^{(\ell)}$, **Definition A.10** introduces its fictional proxy $\bar{\mathbf{x}}_{\ell_k}$, called truncated state, by resetting $\mathbf{x}_{\ell_{k'}+1} = 0$ but preserving the mode switching sequence $\omega_{t'}$, the excitation $\mathbf{z}_{t'}$ and the noise $\mathbf{w}_{t'}$ from $\ell_{k'}+1$ to $\ell_k - 1$, where $\ell_{k'}$ denotes the largest time index smaller than ℓ_k in $S_i^{(\ell)}$. One can view $\bar{\mathbf{x}}_{\ell_k}$ as the zero-initial-state response starting from time $\ell_{k'}+1$, and $\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}$ as the zero-input response.

- **Definition A.11** truncates the states in sub-trajectories $\{(\mathbf{x}_{\ell_k+1}, \mathbf{x}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$ to obtain the truncated sub-trajectories $\{(\bar{\mathbf{x}}_{\ell_k+1}, \bar{\mathbf{x}}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$. Let the rows of $\bar{\mathbf{H}}_i^{(\ell)}$ be $\{\bar{\mathbf{h}}_{\ell_k}^\intercal := [\bar{\mathbf{x}}_{\ell_k}^\intercal/\sigma_\mathbf{w}\ \ \mathbf{z}_{\ell_k}^\intercal/\sigma_\mathbf{z}]\}_{\ell_k \in S_i^{(\ell)}}$. Then using triangle inequality, we have $\|\mathbf{H}_i^{(\ell)\intercal} \mathbf{W}_i^{(\ell)}\| \leq \|\bar{\mathbf{H}}_i^{(\ell)\intercal} \mathbf{W}_i^{(\ell)}\| + \|\mathbf{H}_i^{(\ell)\intercal} \mathbf{W}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal} \mathbf{W}_i^{(\ell)}\|$. Let $\mathbf{v}$ be the eigenvector of

$\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}$ with eigenvalue $\lambda_{\min}(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)})$. We have $\mathbf{v}^\intercal\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}\mathbf{v} = \mathbf{v}^\intercal\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\mathbf{v} + \mathbf{v}^\intercal(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)})\mathbf{v} \implies \lambda_{\min}(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)})\|\mathbf{v}\|^2 \geq \lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)})\|\mathbf{v}\|^2 - |\mathbf{v}^\intercal(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)})\mathbf{v}|$. This implies $\lambda_{\min}(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}) \geq \lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}) - \|\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\|$.

- **Theorem A.16** provides an upper bound on **(a)** $\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$, and a lower bound on **(b)** $\lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)})$ as follows:

  (a) Using **Lemma A.12**, when L is large, we have $\|\bar{\mathbf{h}}_{\ell_k}\| \leq \mathcal{O}(\sqrt{(n+p)\log(T)})$. This implies $\|\bar{\mathbf{H}}_i^{(\ell)}\| \leq \|\bar{\mathbf{H}}_i^{(\ell)}\|_{\mathrm{F}} \leq \mathcal{O}(\sqrt{T(n+p)\log(T)})$. Let $\bar{\mathbf{H}}_i^{(\ell)}$ has singular value decomposition $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ with $\|\boldsymbol{\Sigma}\| \leq \mathcal{O}(\sqrt{T(n+p)\log(T)})$. Since $\mathbf{W}_i^{(\ell)}$ has i.i.d. sub-Gaussian entries, $\mathbf{U}^T\mathbf{W}_i^{(\ell)} \in \mathbb{R}^{(n+p)\times n}$ has i.i.d. sub-Gaussian columns. As a result, applying Theorem 5.39 of Vershynin (2012), we have $\|\mathbf{U}^T\mathbf{W}_i^{(\ell)}\| \leq \hat{\mathcal{O}}(\sigma_{\mathbf{w}}\sqrt{n+p})$. Therefore, $\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\| \leq \|\boldsymbol{\Sigma}\|\|\mathbf{U}^T\mathbf{W}_i^{(\ell)}\| \leq \hat{\mathcal{O}}(\sigma_{\mathbf{w}}(n+p)\sqrt{T\log(T)})$.

  (b) By construction, conditioned on the modes, the rows of $\bar{\mathbf{H}}_i^{(\ell)}$ contains a subset of independent rows. **Definition A.13** introduces this subset as $\{\bar{\mathbf{h}}_{\ell_k}^\intercal\}_{\ell_k\in\bar{S}_i^{(\ell)}}$, where $\bar{S}_i^{(\ell)} := \{\ell_k := \ell + kL \mid \omega_{\ell_k} = i,\ \|\mathbf{x}_{\ell_k}\| \leq \mathcal{O}(\sigma_{\mathbf{w}}\sqrt{n\log(T)}),\ \|\bar{\mathbf{x}}_{\ell_k}\| \leq \mathcal{O}((1/2)\sigma_{\mathbf{w}}\sqrt{n\log(T)}),\ \|\mathbf{z}_{\ell_k}\| \leq \mathcal{O}(\sigma_{\mathbf{z}}\sqrt{p})\}$. Let the rows of $\tilde{\mathbf{H}}_i^{(\ell)}$ be $\{\bar{\mathbf{h}}_{\ell_k}^\intercal\}_{\ell_k\in\bar{S}_i^{(\ell)}}$. **Lemma A.14** proves that the rows of $\tilde{\mathbf{H}}_i^{(\ell)}$ are independent, and **Lemma A.15** shows that they have bounded covariance. These enable us to use Theorem 5.41 of Vershynin (2012) (by specializing it to non-isotropic rows), to obtain $\lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}) \geq \lambda_{\min}(\tilde{\mathbf{H}}_i^{(\ell)\intercal}\tilde{\mathbf{H}}_i^{(\ell)}) \geq \sqrt{|\bar{S}_i^{(\ell)}|/2} - \hat{\mathcal{O}}(\sqrt{s(n+p)\log(T)})$.

- **Theorem A.17** upper bounds the perturbation terms $\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$ and $\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\|$. Both of these terms decay exponentially with L. Hence, by picking large L (e.g., $\hat{\mathcal{O}}(\log(T))$), they can be upper bounded up to the scale of $\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$ and $\lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)})$, respectively.

- **Theorem A.18** provides finite time estimation error bounds, assuming enough number of independent rows in $\bar{\mathbf{H}}_i^{(\ell)}$, specifically, $|\bar{S}_i^{(\ell)}| \geq \mathcal{O}(\pi_{\min}T/L)$, and the noise satisfies Assumption A1.1.

- **Lemma A.19** extends the results of Theorem A.18 to the setting of Gaussian noise. Moreover, **Lemma A.21** proves that, with high probability $|\bar{S}_i^{(\ell)}| \geq \mathcal{O}(\pi_{\min}T/L)$, given T is sufficiently large. Putting these results together we state our main result on the finite time identification of MJS in **Theorem A.22**.

We lower bound $|S_i^{(\ell)}|$ and $|\bar{S}_i^{(\ell)}|$ in Appendix A.4. The proofs of all intermediate theorems and lemmas are provided in Appendix A.5.

A.3.2 Preliminaries

Before, we present the proof of Theorem 3.1, we present some preliminary results which capture the properties of the samples used to estimate the MJS dynamics in Algorithm 1. Given a stabilizing controller $\mathbf{K}_{1:s}$, under the input $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$, the MJS state equation (2.6) becomes,

$$\mathbf{x}_{t+1} = (\mathbf{A}_{\omega_t} + \mathbf{B}_{\omega_t}\mathbf{K}_{\omega_t})\mathbf{x}_t + \mathbf{B}_{\omega_t}\mathbf{z}_t + \mathbf{w}_t = \mathbf{L}_{\omega_t}\mathbf{x}_t + \mathbf{B}_{\omega_t}\mathbf{z}_t + \mathbf{w}_t, \tag{A.14}$$

where $\{\mathbf{z}_t\}_{t=0}^{\infty} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{z}}^2 \mathbf{I}_p)$ is the i.i.d. excitation for exploration and we let $\mathbf{L}_{\omega_t} := \mathbf{A}_{\omega_t} + \mathbf{B}_{\omega_t}\mathbf{K}_{\omega_t}$. To estimate the unknown system dynamics $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s})$, we run the closed-loop MJS (A.14) for T time-steps and collect the trajectory $(\mathbf{x}_t, \mathbf{z}_t, \omega_t)_{t=0}^{T}$. Then, we run Algorithm 1 on the collected trajectory to obtain the estimates $(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s})$. For the ease of analysis, we first derive our estimation error bounds with the following assumption on the noise.

Assumption A1.1 (Bounded noise)**.** *Let* $\{\mathbf{w}_t\}_{t=0}^{T-1} \overset{\text{i.i.d.}}{\sim} \mathcal{D}_w$. *There exists* $\sigma_{\mathbf{w}} > 0$ *and* $c_{\mathbf{w}} \geq 1$ *such that, each entry of* $\mathbf{w}_t$ *is i.i.d. zero-mean sub-Gaussian with variance* $\sigma_{\mathbf{w}}^2$ *and we have* $\|\mathbf{w}_t\|_\infty \leq c_{\mathbf{w}}\sigma_{\mathbf{w}}$.

Later on, we will relax this assumption to get the estimation error bounds with the Gaussian noise. To proceed, we first show that the Euclidean norm of the states $\mathbf{x}_t$ in (A.14) can be upper bounded in expectation. The following result, which is a corollary of Lemma 2.15, accomplishes this.

Corollary A.7 (Bounded states)**.** *Let* $\mathbf{x}_t$ *be the state at time* t *of the MJS* (A.14)*, with initial state* $\mathbf{x}_0 \sim \mathcal{D}_x$. *Suppose Assumption A5.1 on the system and the Markov chain and Assumption A1.1 on the process noise hold. Suppose* $\{\mathbf{z}_t\}_{t=0}^{\infty} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{z}}^2 \mathbf{I}_p)$. *Let* $C_{\mathbf{z}} := \sigma_{\mathbf{z}}/\sigma_{\mathbf{w}}$ *be a constant,* $\mathbf{h}_t := [\mathbf{x}_t^\intercal/\sigma_{\mathbf{w}}\ \mathbf{z}_t^\intercal/\sigma_{\mathbf{z}}]^\intercal$ *be the concatenated state and define*

$$t_0 := \frac{\log\left((1-\rho_{\mathcal{L}})\mathbb{E}[\|\mathbf{x}_0\|^2]/(c_{\mathbf{w}}^2\sigma_{\mathbf{w}}^2 + \sigma_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|)\right)}{1-\rho_{\mathcal{L}}}, \tag{A.15}$$

$$\beta_+^2 := \frac{2\sqrt{s}(c_{\mathbf{w}}^2 + C_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|^2)\tau_{\mathcal{L}}}{1-\rho_{\mathcal{L}}}. \tag{A.16}$$

Then, for all $t \geq t_0$*, we have*

$$\mathbb{E}[\|\mathbf{x}_t\|^2] \leq \sigma_{\mathbf{w}}^2\beta_+^2 n \quad \textit{and} \quad \mathbb{E}[\|\mathbf{h}_t\|^2] \leq (1+\beta_+^2)(n+p). \tag{A.17}$$

To use mixing-time arguments to estimate the unknown system dynamics $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s})$ from a single trajectory $(\mathbf{x}_t, \mathbf{z}_t, \omega_t)_{t=0}^{T}$ of (A.14), we split the trajectory $(\mathbf{x}_t, \mathbf{z}_t, \omega_t)_{t=0}^{T}$ into multiple weakly dependent sub-trajectories, defined as follows.

Definition A.8 (Sub-trajectories of bounded states)**.** *Let sampling period* $L \geq 1$ *be an integer. Let* $\ell_k := \ell + kL$ *be the sub-sampling indices, where* $0 \leq \ell \leq L-1$ *is a fixed offset and* $k = 1, 2, \dots, \lfloor \frac{T-L}{L} \rfloor$*. We sub-sample the trajectory* $(\mathbf{x}_t, \mathbf{z}_t, \omega_t)_{t=0}^T$ *at time indices* $\ell_k \in S_i^{(\ell)}$*, where*

$$S_i^{(\ell)} := \{\ell_k \mid \omega_{\ell_k} = i,\ \|\mathbf{x}_{\ell_k}\| \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n},\ \|\mathbf{z}_{\ell_k}\| \leq c\sigma_{\mathbf{z}}\sqrt{p}\}, \tag{A.18}$$

to obtain the ℓ_{th} *sub-trajectory* $\{(\mathbf{x}_{\ell_k+1}, \mathbf{x}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$.

Note that, $S_i = \bigcup_{\ell=0}^{L-1} S_i^{(\ell)}$, where S_i is as defined in Algorithm 1. This shows that a single trajectory with bounded states and excitations $\{(\mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \omega_t)\}_{t \in S_i}$ is composed of L weakly dependent sub-trajectories $\{(\mathbf{x}_{\ell_k+1}, \mathbf{x}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$ for $0 \leq \ell \leq L-1$. To proceed, we first lower and upper bound the covariance of the states belonging to the weakly dependent sub-trajectories $\{(\mathbf{x}_{\ell_k+1}, \mathbf{x}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$.

Lemma A.9 (Covariance of bounded states)**.** *Consider the same setup of Corollary A.7. Let* t_0 *and* β_+ *be as in* (A.15) *and* (A.16) *respectively. Let* $S_i^{(\ell)}$ *be as in Definition A.8 and* $c \geq 6$ *be a fixed constant. Then, for all* $\ell_k \in S_i^{(\ell)}$ *such that* $\ell_k = \ell + kL \geq t_0$*, we have*

$$(\sigma_{\mathbf{w}}^2/2)\mathbf{I}_n \preceq \mathbf{\Sigma}[\mathbf{x}_{\ell_k}] := \mathbb{E}[\mathbf{x}_{\ell_k}\mathbf{x}_{\ell_k}^\intercal] \preceq c^2\sigma_{\mathbf{w}}^2\beta_+^2 n\mathbf{I}_n, \tag{A.19}$$

$$(1/2)\mathbf{I}_{n+p} \preceq \mathbf{\Sigma}[\mathbf{h}_{\ell_k}] := \mathbb{E}[\mathbf{h}_{\ell_k}\mathbf{h}_{\ell_k}^\intercal] \preceq c^2(1+\beta_+^2)(n+p)\mathbf{I}_{n+p}. \tag{A.20}$$

To proceed, let $\mathbf{h}_t$ be as in Corollary A.7 and $\mathbf{\Theta}_i^\star := [\sigma_{\mathbf{w}}\mathbf{L}_i\ \sigma_{\mathbf{z}}\mathbf{B}_i]$ for all $i \in [s]$. Then the output of each sample in $\{(\mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{z}_t, \omega_t)\}_{t \in S_i}$ can be related to the inputs as follows,

$$\mathbf{x}_{t+1} = \mathbf{\Theta}_{\omega_t}^\star \mathbf{h}_t + \mathbf{w}_t \quad \text{for all} \quad t \in S_i. \tag{A.21}$$

Next, to carry out finite sample identification of $\mathbf{\Theta}_i^\star$ using the method of linear least squares, we define the following concatenated matrices,

$$\mathbf{Y}_i = \begin{bmatrix} \mathbf{x}_{t_1+1}^\intercal \\ \mathbf{x}_{t_2+1}^\intercal \\ \vdots \\ \mathbf{x}_{t_{|S_i|}+1}^\intercal \end{bmatrix}, \quad \mathbf{H}_i = \begin{bmatrix} \mathbf{h}_{t_1}^\intercal \\ \mathbf{h}_{t_2}^\intercal \\ \vdots \\ \mathbf{h}_{t_{|S_i|}}^\intercal \end{bmatrix}, \quad \mathbf{W}_i = \begin{bmatrix} \mathbf{w}_{t_1}^\intercal \\ \mathbf{w}_{t_2}^\intercal \\ \vdots \\ \mathbf{w}_{t_{|S_i|}}^\intercal \end{bmatrix}, \tag{A.22}$$

that is, $\mathbf{Y}_i$ has $\{\mathbf{x}_{t+1}^\intercal\}_{t \in S_i}$ on its rows, $\mathbf{H}_i$ has $\{\mathbf{h}_t^\intercal\}_{t \in S_i}$ on its rows and $\mathbf{W}_i$ has $\{\mathbf{w}_t^\intercal\}_{t \in S_i}$ on its rows. Similarly, we also construct $\mathbf{Y}_i^{(\ell)}$, $\mathbf{H}_i^{(\ell)}$ and $\mathbf{W}_i^{(\ell)}$ by (row-wise) stacking $\{\mathbf{x}_{\ell_k+1}^\intercal\}_{\ell_k \in S_i^{(\ell)}}$, $\{\mathbf{h}_{\ell_k}^\intercal\}_{\ell_k \in S_i^{(\ell)}}$ and $\{\mathbf{w}_{\ell_k}^\intercal\}_{\ell_k \in S_i^{(\ell)}}$ respectively. We have $\mathbf{Y}_i = \mathbf{H}_i\mathbf{\Theta}_i^{\star\intercal} + \mathbf{W}_i$ and the regression problem in Algorithm 1 is alternately represented as

$$\hat{\mathbf{\Theta}}_i^\intercal = \arg\min_{\mathbf{\Theta}_i} \frac{1}{2|S_i|}\|\mathbf{Y}_i - \mathbf{H}_i\mathbf{\Theta}_i^\intercal\|_{\mathrm{F}}^2. \tag{A.23}$$

The least squares estimator $\hat{\mathbf{\Theta}}_i^\intercal = \mathbf{H}_i^\dagger \mathbf{Y}_i = \big(\mathbf{H}_i^\intercal\mathbf{H}_i\big)^{-1}\mathbf{H}_i^\intercal\mathbf{Y}_i$ has the following estimation error,

$$\begin{aligned}\|\hat{\mathbf{\Theta}}_i - \mathbf{\Theta}_i^\star\| \leq \|\big(\mathbf{H}_i^\intercal\mathbf{H}_i\big)^{-1}\|\|\mathbf{H}_i^\intercal\mathbf{W}_i\| = \frac{\|\mathbf{H}_i^\intercal\mathbf{W}_i\|}{\lambda_{\min}\big(\mathbf{H}_i^\intercal\mathbf{H}_i\big)} &\overset{\text{(a)}}{=} \frac{\|\sum_{\ell=0}^{L-1}\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|}{\lambda_{\min}\big(\sum_{\ell=0}^{L-1}\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}\big)}, \\ &\overset{\text{(b)}}{=} \frac{\sum_{\ell=0}^{L-1}\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|}{\sum_{\ell=0}^{L-1}\lambda_{\min}\big(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}\big)},\end{aligned} \tag{A.24}$$

where we obtain (a) from the fact that $S_i = \bigcup_{\ell=0}^{L-1} S_i^{(\ell)}$ and (b) follows from using triangular inequality and Weyl's inequality for Hermitian matrices. If we upper bound the terms $\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$ and lower bound the terms $\lambda_{\min}\big(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}\big)$, for all $0 \leq \ell \leq L-1$, we can use (A.24) to upper bound the estimation error $\|\hat{\mathbf{\Theta}}_i - \mathbf{\Theta}_i^\star\|$. However, because $\mathbf{H}_i^{(\ell)}$ has non-i.i.d. rows, it is not straightforward to bound the terms $\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$ and $\lambda_{\min}\big(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}\big)$ directly. To resolve this issue, we rely on the notion of stability and use perturbation based techniques to indirectly bound these terms in the following sub-sections.

A.3.3 Truncated Sub-trajectories

Definition A.10 (Truncated state vector (Oymak, 2019))**.** *Consider state equation (A.14). Given, $t \geq L > 0$, for each state $\mathbf{x}_t$, we define its fictional proxy $\mathbf{x}_{t,L}$ by resetting $\mathbf{x}_{t-L} = 0$ but preserving the excitation $\mathbf{z}_{t'}$, noise $\mathbf{w}_{t'}$, and modes $\omega_{t'}$ for $t' = t-L, \dots, t-1$. Alternately, $\mathbf{x}_{t,L}$ is obtained by driving the system with excitation $\mathbf{z}'_{t'}$ and additive noise $\mathbf{w}'_{t'}$ until time $t-1$, where*

$$\mathbf{z}'_{t'} = \begin{cases} 0 & \text{if} \quad t' < t-L \\ \mathbf{z}_{t'} & \text{else} \end{cases}, \quad \text{and} \quad \mathbf{w}'_{t'} = \begin{cases} 0 & \text{if} \quad t' < t-L \\ \mathbf{w}_{t'} & \text{else} \end{cases}. \tag{A.25}$$

We call the obtained state $\mathbf{x}_{t,L}$ as the L-truncated (or simply truncated) state at time t.

Using Definition A.10, we can obtain independent samples (for the purpose of analysis only) from a single trajectory which will be used to capture the effect of learning from a single trajectory. With high probability over the mode observation, truncated states can be made very close to the original states with sufficiently large truncation length. To show this, we expand the closed-loop MJS state equation (A.14) as follows,

$$\mathbf{x}_t = \begin{cases} \mathbf{L}_{\omega_0}\mathbf{x}_0 + \mathbf{B}_{\omega_0}\mathbf{z}_0 + \mathbf{w}_0 & \text{if} \quad t=1, \\ \prod_{j=0}^{t-1}\mathbf{L}_{\omega_j}\mathbf{x}_0 + \sum_{t'=0}^{t-2}\prod_{j=t'+1}^{t-1}\mathbf{L}_{\omega_j}\mathbf{B}_{\omega_{t'}}\mathbf{z}_{t'} + \mathbf{B}_{\omega_{t-1}}\mathbf{z}_{t-1} & \\ \qquad + \sum_{t'=0}^{t-2}\prod_{j=t'+1}^{t-1}\mathbf{L}_{\omega_j}\mathbf{w}_{t'} + \mathbf{w}_{t-1} & \text{if} \quad t \geq 2, \end{cases} \tag{A.26}$$

where $\mathbf{x}_0$ denotes the state at time $t=0$. Using (A.26), the difference between $\mathbf{x}_t$ and $\mathbf{x}_{t,L}$ is given by

$$\mathbf{x}_t - \mathbf{x}_{t,L} = \prod_{j=t-L}^{t-1}\mathbf{L}_{\omega_j}\mathbf{x}_{t-L}. \tag{A.27}$$

As a corollary of Lemma 2.15, observe that for a closed loop autonomous system $\mathbf{x}_{t+1} = \mathbf{L}_{\omega_t}\mathbf{x}_t$, mean-square stability implies that, for any initial conditions $\mathbf{x}_0$ and ω_0, we have $\mathbb{E}[\|\mathbf{x}_t\|^2] \leq \sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^t\|\mathbf{x}_0\|^2$. Combining this argument with (A.27), we have

$$\begin{aligned} \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t,L}\|^2] = \mathbb{E}[\|\prod_{j=t-L}^{t-1}\mathbf{L}_{\omega_j}^{+}\mathbf{x}_{t-L}\|^2] \leq \sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^L\|\mathbf{x}_{t-L}\|^2, \\ \Longrightarrow \mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t,L}\|] \leq \sqrt{(ns)^{1/2}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^L}\|\mathbf{x}_{t-L}\| \leq \sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{L/2}\|\mathbf{x}_{t-L}\|, \end{aligned} \tag{A.28}$$

where the expectation is over the Markov modes $\{\omega_j\}_{j=t-L}^{t-1}$ and we get the last relation by using Jensen's inequality. Moreover, if we also have $\|\mathbf{x}_{t-L}\| \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n}$, then we can make $\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t,L}\|]$ arbitrarily small by picking a sufficiently large truncation length $L \geq 1$. We have

$$\mathbb{E}[\|\mathbf{x}_t - \mathbf{x}_{t,L}\| \mid \|\mathbf{x}_{t-L}\| \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n}] \leq \tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^L\|\mathbf{x}_{t-L}\| \leq c\sigma_{\mathbf{w}}\beta_+ n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{L/2}. \tag{A.29}$$

To proceed, we carry out the truncation of the sub-trajectories introduced in Definition A.8 to get the truncated sub-trajectories defined as follows.

Definition A.11 (Truncated sub-trajectories)**.** *Let sampling period $L \geq 1$ be an integer. Let $\ell_k := \ell + kL$ be the sub-sampling indices, where $0 \leq \ell \leq L-1$ is a fixed offset and $k = 1, 2, \dots, \lfloor \frac{T-L}{L} \rfloor$. Let $S_i^{(\ell)}$ be as in Definition A.8. For each $\ell_k \in S_i^{(\ell)}$, let $\ell_{k'} \in S_i^{(\ell)}$ denotes the largest time index smaller than ℓ_k. Given the ℓ_{th} sub-trajectory $\{(\mathbf{x}_{\ell_k+1}, \mathbf{x}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$ from Definition A.8, we truncate each state $\mathbf{x}_{\ell_k}$ by $\ell_k - \ell_{k'} - 1$ to get the ℓ_{th} truncated sub-trajectory $\{(\bar{\mathbf{x}}_{\ell_k+1}, \bar{\mathbf{x}}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$, where*

$$\bar{\mathbf{x}}_{\ell_k} := \mathbf{x}_{\ell_k, \ell_k - \ell_{k'} - 1} \quad \textit{and} \quad \bar{\mathbf{x}}_{\ell_k+1} := \mathbf{L}_{\omega_{\ell_k}}\bar{\mathbf{x}}_{\ell_k} + \mathbf{B}_{\omega_{\ell_k}}\mathbf{z}_{\ell_k} + \mathbf{w}_{\ell_k}. \tag{A.30}$$

If ℓ_k is the smallest time index in $S_i^{(\ell)}$, we set $\ell_{k'} = 0$.

Note that $\ell_k - \ell_{k'} \geq L$ by definition. Hence, the truncation lengths used to obtain $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$ are always larger than $L-1$. Next, we show that when L is sufficiently large enough, then the truncated states $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$ as well as the Euclidean distance between the truncated and non-truncated states can be bounded with high probability over the modes.

Lemma A.12 (Bounded states (truncated))**.** *Consider the same setup of Corollary A.7. Let $\{\mathbf{x}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$ be the bounded states and $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$ be the truncated states from Definition A.8 and A.11 respectively. Let*

$$\beta'_+ := c_{\mathbf{w}} + \beta_+ \|\mathbf{L}_{1:s}\| + C_{\mathbf{z}} \sqrt{p/n} \|\mathbf{B}_{1:s}\|, \tag{A.31}$$

$$L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}}, \delta) := 1 + \frac{2 \log \left(2\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}} T \beta'_+ / (\beta_+ \delta)\right)}{1 - \rho_{\boldsymbol{\mathcal{L}}}}, \tag{A.32}$$

$$\text{and} \quad L \geq \max\{t_0, L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}}, \delta)\}. \tag{A.33}$$

Then, with probability at least $1-\delta$ over the modes, for all $\ell_k \in S_i^{(\ell)}$ and all $i \in [s]$, we have

$$\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\| \leq (1/2) c \sigma_{\mathbf{w}} \beta_+ \sqrt{n} \quad \text{and} \quad \|\bar{\mathbf{x}}_{\ell_k}\| \leq (3/2) c \sigma_{\mathbf{w}} \beta_+ \sqrt{n}. \tag{A.34}$$

By construction, conditioned on the modes, $\bar{\mathbf{x}}_{\ell_k} = \mathbf{x}_{\ell_k, \ell_k - \ell_{k'} - 1}$ only depends on the excitation and noise $\{\mathbf{z}_t, \mathbf{w}_t\}_{t=\ell+k'L+1}^{\ell+kL-1}$. Note that the dependence ranges $[\ell + k'L + 1, \ell + kL - 1]$ are disjoint intervals for each (k, k') pairs. Hence, $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$ should all be independent of each other. However, this is not the case because $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$ are obtained by truncating only bounded states $\{\mathbf{x}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$. Therefore, we will look for a subset of independent truncated states within $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$, as follows.

Definition A.13 (Subset of bounded states)**.** *Let sampling period $L \geq 1$ be an integer. Let $\ell_k = \ell + kL$ be the sub-sampling indices, where $0 \leq \ell \leq L-1$ is a fixed offset and $k = 1, 2, \dots, \lfloor \frac{T-L}{L} \rfloor$. We sub-sample the trajectory $(\mathbf{x}_t, \mathbf{z}_t, \omega_t)_{t=0}^T$ at time indices $\ell_k \in \bar{S}_i^{(\ell)}$, where*

$$\bar{S}_i^{(\ell)} := \{\ell_k \bigm| \omega_{\ell_k} = i,\ \|\mathbf{x}_{\ell_k}\| \leq c\sigma_{\mathbf{w}} \beta_+ \sqrt{n},\ \|\bar{\mathbf{x}}_{\ell_k}\| \leq (1/2) c \sigma_{\mathbf{w}} \beta_+ \sqrt{n},\ \|\mathbf{z}_{\ell_k}\| \leq c \sigma_{\mathbf{z}} \sqrt{p}\}, \tag{A.35}$$

to obtain a subset of the ℓ_{th} sub-trajectory, denoted by $\{(\mathbf{x}_{\ell_k+1}, \mathbf{x}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in \bar{S}_i^{(\ell)}}$.

Next, we show that, conditioned on the modes, the samples in $\{(\bar{\mathbf{x}}_{\ell_k+1}, \bar{\mathbf{x}}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ are independent.

Lemma A.14 (Conditional independence)**.** *Consider the MJS in* (A.14)*. Suppose* $\{\mathbf{z}_t\}_{t=0}^{\infty} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{z}}^2 \mathbf{I}_p)$ *and* $\{\mathbf{w}_t\}_{t=0}^{\infty} \overset{\text{i.i.d.}}{\sim} \mathcal{D}_w$ *satisfies Assumption A1.1. Suppose the sampling period* L *satisfies* (A.33)*. Let* $S_i^{(\ell)}$ *and* $\bar{S}_i^{(\ell)}$ *be as in Definition A.8 and A.13 respectively. Then, with probability at least* $1-\delta$ *over the mode, we have,* **(a)** $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ *is a subset of* $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$*,* **(b)** *conditioned on the modes,* $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ *are all independent,* **(c)** *conditioned on the modes,* $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$*,* $\{\mathbf{z}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ *and* $\{\mathbf{w}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ *are all independent of each other.*

Next, we state a lemma similar to Lemma A.9 to show that the truncated states have nice covariance properties.

Lemma A.15 (Covariance of truncated states)**.** *Consider the setup of Corollary A.7. Let* t_0*,* β_+ *and* β'_+ *be as in* (A.15)*,* (A.16) *and* (A.31) *respectively. Let* $c \geq 6$ *be a fixed constant and* $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ *be as in Lemma A.14. Define*

$$L_{cov}(\rho_{\mathcal{L}}) := 1 + \frac{2\log\left(8c^2\beta_+\beta'_+ n\sqrt{ns}\tau_{\mathcal{L}}\right)}{1-\rho_{\mathcal{L}}}, \tag{A.36}$$

and suppose, the sampling period L *obeys,*

$$L \geq \max\{t_0, L_{cov}(\rho_{\mathcal{L}})\}. \tag{A.37}$$

Then, for all $\ell_k \in \bar{S}_i^{(\ell)}$*, we have*

$$(\sigma_{\mathbf{w}}^2/4)\mathbf{I}_n \preceq \mathbf{\Sigma}[\bar{\mathbf{x}}_{\ell_k}] := \mathbb{E}[\bar{\mathbf{x}}_{\ell_k}\bar{\mathbf{x}}_{\ell_k}^\intercal] \preceq 2c^2\sigma_{\mathbf{w}}^2\beta_+^2 n\mathbf{I}_n. \tag{A.38}$$
$$(1/4)\mathbf{I}_{n+p} \preceq \mathbf{\Sigma}[\bar{\mathbf{h}}_{\ell_k}] := \mathbb{E}[\bar{\mathbf{h}}_{\ell_k}\bar{\mathbf{h}}_{\ell_k}^\intercal] \preceq 2c^2(1+\beta_+^2)(n+p)\mathbf{I}_{n+p}. \tag{A.39}$$

To proceed, consider the ℓ_{th} truncated sub-trajectory $\{(\bar{\mathbf{x}}_{\ell_k+1}, \bar{\mathbf{x}}_{\ell_k}, \mathbf{z}_{\ell_k}, \omega_{\ell_k})\}_{\ell_k \in S_i^{(\ell)}}$ given by Definition A.11. Let $\bar{\mathbf{h}}_{\ell_k} := [\bar{\mathbf{x}}_{\ell_k}^\intercal/\sigma_{\mathbf{w}} \; \mathbf{z}_{\ell_k}^\intercal/\sigma_{\mathbf{z}}]^\intercal$. Similar to (A.22), we construct $\bar{\mathbf{Y}}_i^{(\ell)}$, $\bar{\mathbf{H}}_i^{(\ell)}$ and $\mathbf{W}_i^{(\ell)}$ by (row-wise) stacking $\{\bar{\mathbf{x}}_{\ell_k+1}^\intercal\}_{\ell_k \in S_i^{(\ell)}}$, $\{\bar{\mathbf{h}}_{\ell_k}^\intercal\}_{\ell_k \in S_i^{(\ell)}}$ and $\{\mathbf{w}_{\ell_k}^\intercal\}_{\ell_k \in S_i^{(\ell)}}$ respectively. As an intermediate step, in the following, we will lower bound $\lambda_{\min}\big(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\big)$ and upper bound $\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$. This will, in turn, allow us to lower and upper bound the non-truncated terms $\lambda_{\min}\big(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}\big)$ and $\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$ respectively via,

$$\lambda_{\min}\big(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}\big) \geq \lambda_{\min}\big(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\big) - \|\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\|, \tag{A.40}$$
$$\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\| \leq \|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\| + \|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|. \tag{A.41}$$

For this purpose, our next lemma lower bounds the eigenvalues of the matrix $\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}$ and upper bounds the error term $\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$.

Theorem A.16 (Bounding $\lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)})$ and $\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$)**.** *Consider the setup of Corollary A.7. Let t_0, β_+, β_+', $L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}},\delta)$ and $L_{cov}(\rho_{\boldsymbol{\mathcal{L}}})$ be as in (A.15), (A.16), (A.31), (A.32) and (A.36), respectively. Let $C, C_0 > 0$ and $c \geq 6$ be fixed constants. Let $\bar{\mathbf{H}}_i^{(\ell)}$, $\tilde{\mathbf{H}}_i^{(\ell)}$ and $\mathbf{W}_i^{(\ell)}$ be constructed by (row-wise) stacking $\{\bar{\mathbf{h}}_{\ell_k}^\intercal\}_{\ell_k\in S_i^{(\ell)}}$, $\{\bar{\mathbf{h}}_{\ell_k}^\intercal\}_{\ell_k\in \bar{S}_i^{(\ell)}}$ and $\{\mathbf{w}_{\ell_k}^\intercal\}_{\ell_k\in S_i^{(\ell)}}$ respectively. Suppose the sampling period L and the number of independent samples $|\bar{S}_i^{(\ell)}|$ satisfy the following lower bounds,*

$$L \geq \max\{t_0, L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}},\delta), L_{cov}(\rho_{\boldsymbol{\mathcal{L}}})\}, \tag{A.42}$$

$$|\bar{S}_i^{(\ell)}| \geq 16c^2(1+\beta_+^2)\log\big(\frac{2s(n+p)}{\delta}\big)(n+p). \tag{A.43}$$

Then, with probability at least $1-3\delta$, for all $i \in [s]$, we have

$$\lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}) \geq \lambda_{\min}(\tilde{\mathbf{H}}_i^{(\ell)\intercal}\tilde{\mathbf{H}}_i^{(\ell)}) \geq \frac{|\bar{S}_i^{(\ell)}|}{16}, \tag{A.44}$$

$$\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\| \leq 2c(1+\beta_+)\sqrt{|S_i^{(\ell)}|(n+p)}\Big(C\sigma_{\mathbf{w}}\sqrt{n+p} + C_0\sqrt{\log\big(\frac{2s}{\delta}\big)}\Big). \tag{A.45}$$

A.3.4 Bounding the Estimation Error

Coming back to the initial problem of estimating the unknown MJS dynamics from dependent samples, by solving the regression problem (A.23), observe that the estimation error in (A.24) can be upper bounded as follows,

$$\begin{aligned}\|\hat{\boldsymbol{\Theta}}_i - \boldsymbol{\Theta}_i^\star\| &\leq \frac{\sum_{\ell=0}^{L-1}\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|}{\sum_{\ell=0}^{L-1}\lambda_{\min}(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)})}, \\ &\leq \frac{\sum_{\ell=0}^{L-1}\big(\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\| + \|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|\big)}{\sum_{\ell=0}^{L-1}\big(\lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}) - \|\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\|\big)}.\end{aligned} \tag{A.46}$$

To upper bound the estimation error $\|\hat{\boldsymbol{\Theta}}_i - \boldsymbol{\Theta}_i^\star\|$ in (A.46), we need to upper bound the impact of truncation, captured by $\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$ and $\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\|$ for all $i \in [s]$ and all $0 \leq \ell \leq L-1$. This is done by the following theorem.

Theorem A.17 (Small impact of truncation)**.** *Consider the same setup of Corollary A.7. Let t_0, β_+, β_+' and $L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}},\delta)$ be as in (A.15), (A.16), (A.31) and (A.32), respectively. Suppose the sampling period L obeys $L \geq \max\{t_0, L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}},\delta)\}$. Let $\bar{\mathbf{H}}_i^{(\ell)}$ and $\mathbf{W}_i^{(\ell)}$ be constructed by (row-wise) stacking $\{\bar{\mathbf{h}}_{\ell_k}^\intercal\}_{\ell_k\in S_i^{(\ell)}}$ and $\{\mathbf{w}_{\ell_k}^\intercal\}_{\ell_k\in S_i^{(\ell)}}$ respectively. Then, with probability at least*

$1-\delta$ *over the modes, for all* $i \in [s]$, *we have*

$$\|\mathbf{H}_i^{(\ell)\mathsf{T}}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\mathsf{T}}\bar{\mathbf{H}}_i^{(\ell)}\| \leq \frac{3c^2\beta'_+(1+\beta_+)\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2}n\sqrt{s(n+p)}|S_i^{(\ell)}|T}{\delta}, \tag{A.47}$$

$$\|\mathbf{H}_i^{(\ell)\mathsf{T}}\mathbf{W}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\mathsf{T}}\mathbf{W}_i^{(\ell)}\| \leq \frac{cc_{\mathbf{w}}\sigma_{\mathbf{w}}\beta'_+\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2}n\sqrt{ns}|S_i^{(\ell)}|T}{\delta}. \tag{A.48}$$

Combining Theorems A.16 and A.17, we obtain our result on the estimation of MJS in (A.14) from finite samples obtained from a single trajectory.

Theorem A.18 (Learning with bounded noise)**.** *Consider the setup of Corollary A.7. Let* t_0, β_+, β'_+, $L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}}, \delta)$ *and* $L_{cov}(\rho_{\boldsymbol{\mathcal{L}}})$ *be as in* (A.15), (A.16), (A.31), (A.32) *and* (A.36), *respectively. Let* $S_i^{(\ell)}$ *and* $\bar{S}_i^{(\ell)}$ *be as in Definition A.8 and A.13 respectively and assume* $|\bar{S}_i^{(\ell)}| \geq \frac{\pi_{\min}T}{2L}$, *for all* $i \in [s]$, *with probability at least* $1-\delta$. *Suppose* $\|\mathbf{K}_{1:s}\| \leq C_K$ *for some constant* $C_K > 0$. *Let* $C, C_0 > 0$, *and* $c \geq 6$ *be fixed constants. Define*

$$L_{tr_2}(\rho_{\boldsymbol{\mathcal{L}}}, \delta) := 1 + \frac{2}{(1-\rho_{\boldsymbol{\mathcal{L}}})}\log\big(\frac{192c^2\tau_{\boldsymbol{\mathcal{L}}}\beta'_+(1+\beta_+)n\sqrt{s(n+p)}T}{\pi_{\min}\delta}\big), \tag{A.49}$$

$$L_{tr_3}(\rho_{\boldsymbol{\mathcal{L}}}, \delta) := 1 + \frac{2}{(1-\rho_{\boldsymbol{\mathcal{L}}})}\log\big(\frac{c_{\mathbf{w}}\sigma_{\mathbf{w}}\beta'_+\tau_{\boldsymbol{\mathcal{L}}}n\sqrt{nsT}T}{\delta(1+\beta_+)\sqrt{(n+p)}\big(C\sigma_{\mathbf{w}}\sqrt{n+p} + C_0\sqrt{\log(2s/\delta)}\big)}\big). \tag{A.50}$$

Suppose the sampling period L *and the trajectory length* T *satisfy*

$$L \geq \max\{t_0, L_{cov}(\rho_{\boldsymbol{\mathcal{L}}}), L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}}, \frac{\delta}{18L}), L_{tr2}(\rho_{\boldsymbol{\mathcal{L}}}, \frac{\delta}{18L}), L_{tr3}(\rho_{\boldsymbol{\mathcal{L}}}, \frac{\delta}{18L})\} \tag{A.51}$$

$$T \geq \frac{32L}{\pi_{\min}}c^2(1+\beta_+^2)\log\big(\frac{36sL(n+p)}{\delta}\big)(n+p). \tag{A.52}$$

Then, solving the least-squares problem (A.23), *with probability at least* $1-\delta/2$, *for all* $i \in [s]$, *we have*

$$\|\hat{\mathbf{A}}_i - \mathbf{A}_i\| \leq \frac{(\sigma_{\mathbf{z}} + C_K\sigma_{\mathbf{w}})}{\sigma_{\mathbf{z}}}\frac{192c(1+\beta_+)}{\pi_{\min}}\sqrt{\frac{L(n+p)}{T}}\Big(C\sqrt{n+p} + (C_0/\sigma_{\mathbf{w}})\sqrt{\log\big(\frac{36sL}{\delta}\big)}\Big),$$

$$\|\hat{\mathbf{B}}_i - \mathbf{B}_i\| \leq \frac{\sigma_{\mathbf{w}}}{\sigma_{\mathbf{z}}}\frac{192c(1+\beta_+)}{\pi_{\min}}\sqrt{\frac{L(n+p)}{T}}\Big(C\sqrt{n+p} + (C_0/\sigma_{\mathbf{w}})\sqrt{\log\big(\frac{36sL}{\delta}\big)}\Big). \tag{A.53}$$

Next, we use the following lemma to relax the Assumption A1.1 on the noise.

Lemma A.19 (From Bounded to Unbounded Noise)**.** *Let* $\mathbf{g} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{w}}^2\mathbf{I}_{nT})$ *and* $\mathbf{a}$ *be two independent vectors. Let* $\mathbf{g}'$ *be the truncated Gaussian distribution* $\mathbf{g}' \sim \{\mathbf{g} \mid \|\mathbf{g}\|_\infty \leq c_{\mathbf{w}}\sigma_{\mathbf{w}}\}$.

Let $S_{\mathbf{g},\mathbf{a}}$ be the indicator function of an event defined on vectors $\mathbf{g}, \mathbf{a}$ s.t.

$$\mathbb{E}[S_{\mathbf{g}',\mathbf{a}}] \geq 1 - \delta/2.$$

That is, the event holds, on the bounded variable $\mathbf{g}'$, with probability at least $1 - \delta/2$. Then, if the bound above holds for $c_{\mathbf{w}} > C_\delta := \sqrt{2\log(nT)} + \sqrt{2\log(2/\delta)}$, we also have that

$$\mathbb{E}[S_{\mathbf{g},\mathbf{a}}] \geq 1 - \delta.$$

That is, the probability that event holds on the unbounded variable $\mathbf{g}$ is at least $1 - \delta$.

Combining Theorem A.18 and Lemma A.19, we get the following result on learning the MJS dynamics when the process noise is Gaussian.

Corollary A.20 (Learning with un-bounded noise)**.** *Consider the same setup of Theorem A.18 except that Assumption A1.1 is replaced with $\{\mathbf{w}_t\}_{t=0}^{\infty} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{w}}^2 \mathbf{I}_n)$ and the threshold for bounding the noise satisfies,*

$$c_{\mathbf{w}} \geq C_\delta := \sqrt{2\log(nT)} + \sqrt{2\log(2/\delta)}. \tag{A.54}$$

Suppose $\|\mathbf{B}_{1:s}\| \leq C_B$ for some $C_B > 0$ and the trajectory length T satisfies,

$$T \gtrsim \frac{\tau_{\boldsymbol{\mathcal{L}}}\sqrt{s}L}{\pi_{\min}(1-\rho_{\boldsymbol{\mathcal{L}}})}\big(C_\delta^2 + C_{\mathbf{z}}^2 C_B^2\big)\log\big(\frac{36sL(n+p)}{\delta}\big)(n+p). \tag{A.55}$$

Then, solving the least-squares problem (A.23), with probability at least $1 - \delta$, for all $i \in [s]$, we have

$$\begin{aligned}
&\|\hat{\mathbf{A}}_i - \mathbf{A}_i\| \\
&\quad\lesssim \frac{(\sigma_{\mathbf{z}} + C_K\sigma_{\mathbf{w}})}{\sigma_{\mathbf{z}}}\frac{\tau_{\boldsymbol{\mathcal{L}}}\big(C_\delta + C_{\mathbf{z}}C_B\big)}{\pi_{\min}(1-\rho_{\boldsymbol{\mathcal{L}}})}\sqrt{\frac{sL(n+p)}{T}}\Big(C\sqrt{n+p} + (C_0/\sigma_{\mathbf{w}})\sqrt{\log\big(\frac{36sL}{\delta}\big)}\Big), \\
&\|\hat{\mathbf{B}}_i - \mathbf{B}_i\| \\
&\quad\lesssim \frac{\sigma_{\mathbf{w}}}{\sigma_{\mathbf{z}}}\frac{\tau_{\boldsymbol{\mathcal{L}}}\big(C_\delta + C_{\mathbf{z}}C_B\big)}{\pi_{\min}(1-\rho_{\boldsymbol{\mathcal{L}}})}\sqrt{\frac{sL(n+p)}{T}}\Big(C\sqrt{n+p} + (C_0/\sigma_{\mathbf{w}})\sqrt{\log\big(\frac{36sL}{\delta}\big)}\Big).
\end{aligned} \tag{A.56}$$

At this point, we are only left with verifying the assumption that, for all $i \in [s]$, with probability at least $1-\delta$, we have $|\bar{S}_i^{(\ell)}| \geq \frac{\pi_{\min}T}{2L}$ for some choice of L and T. In the following, we will state a lemma to show that the above assumption indeed holds for certain choice of L and T. The detailed analysis for obtaining a lower bound on $|\bar{S}_i^{(\ell)}|$ is given in Section A.4.

Specifically, the following result can be obtained by applying union bound to Lemma A.27 over $\ell = 0, 1, \dots, L-1$.

Lemma A.21. *Let $\bar{S}_i^{(\ell)}$ be as in Definition A.13 and consider the setup of Algorithm 1. Assume $c_{\mathbf{x}} \geq \underline{c}_{\mathbf{x}}(\rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})$, $c_{\mathbf{z}} \geq \underline{c}_{\mathbf{z}}$, $L \geq \underline{C}_{sub,N}\left(\beta_0\sqrt{n}, \frac{\delta}{L}, T, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}}\right)\log(T)$, and $T \geq \underline{T}_N(\frac{L}{\log(T)}, \frac{\delta}{L}, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})$, where $\underline{c}_{\mathbf{x}}(\rho, \tau)$, $\underline{c}_{\mathbf{z}}$, and $\underline{T}_N(\mathcal{C}, \delta, \rho, \tau)$ are defined in Table C.1, and $C_{sub,N}(\bar{x}_0, \delta, T, \rho, \tau)$ is defined in Table A.1. Then with probability at least $1-\delta$, for all $i \in [s]$ and all $\ell = 0, 1, \dots, L-1$ we have*

$$|\bar{S}_i^{(\ell)}| \geq \frac{\pi_{\min} T}{2L}. \tag{A.57}$$

A.3.5 Finalizing the SYSID: Proof of Theorem 3.1

To finalize, we combine Corollary A.20 and Lemma A.21 to get our main result on learning the unknown MJS dynamics. The following theorem is a more refined and precise version of our main system identification result in Theorem 3.1.

Theorem A.22 (Main result)**.** *Consider the MJS (A.14), with initial state $\mathbf{x}_0 \sim \mathcal{D}_x$. Suppose Assumption A5.1 on the system and the Markov chain holds. Suppose $\{\mathbf{z}_t\}_{t=0}^{\infty} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{z}}^2 \mathbf{I}_p)$, $\{\mathbf{w}_t\}_{t=0}^{\infty} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_{\mathbf{w}}^2 \mathbf{I}_n)$ and the threshold for bounding the noise satisfies,*

$$c_{\mathbf{w}} \geq C_\delta := \sqrt{2\log(nT)} + \sqrt{2\log(2/\delta)}. \tag{A.58}$$

Suppose $\|\mathbf{B}_{1:s}\| \leq C_B$ and $\|\mathbf{K}_{1:s}\| \leq C_K$ for some $C_B, C_K > 0$. Let t_0, β_+, β'_+, $L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}}, \delta)$, $L_{cov}(\rho_{\boldsymbol{\mathcal{L}}})$, $L_{tr2}(\rho_{\boldsymbol{\mathcal{L}}}, \delta)$ and $L_{tr3}(\rho_{\boldsymbol{\mathcal{L}}}, \delta)$ be as in (A.15), (A.16), (A.31), (A.32), (A.36), (A.49) and (A.50), respectively. Suppose $c_{\mathbf{x}} \geq \underline{c}_{\mathbf{x}}(\rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})$, $c_{\mathbf{z}} \geq \underline{c}_{\mathbf{z}}$, where $\underline{c}_{\mathbf{x}}(\rho, \tau)$ and $\underline{c}_{\mathbf{z}}$ are defined in Table C.1. Let $C, C_0 > 0$, and $c \geq 6$ be fixed constants. Suppose the trajectory length T satisfies

$$T \gtrsim \max\left\{\frac{\tau_{\boldsymbol{\mathcal{L}}}\sqrt{s}L}{\pi_{\min}(1-\rho_{\boldsymbol{\mathcal{L}}})}\left(C_\delta^2 + C_{\mathbf{z}}^2 C_B^2\right)\log\left(\frac{36sL(n+p)}{\delta}\right)(n+p), \underline{T}_N\left(\frac{L}{\log(T)}, \frac{\delta}{L}, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}}\right)\right\} \tag{A.59}$$

$$\text{where,}\quad L \geq \max\left\{t_0, L_{cov}(\rho_{\boldsymbol{\mathcal{L}}}), L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}}, \frac{\delta}{18L}), L_{tr2}(\rho_{\boldsymbol{\mathcal{L}}}, \frac{\delta}{18L}), L_{tr3}(\rho_{\boldsymbol{\mathcal{L}}}, \frac{\delta}{18L}), \underline{C}_{sub,N}\left(\beta_0\sqrt{n}, \frac{\delta}{L}, T, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}}\right)\log(T)\right\}, \tag{A.60}$$

where $\underline{T}_N(\mathcal{C}, \delta, \rho, \tau)$ and $C_{sub,N}(\bar{x}_0, \delta, T, \rho, \tau)$ are defined in Table C.1 and A.1 respectively. Then, solving the least-squares problem (A.23), with probability at least $1-\delta$, for all $i \in [s]$, *we have*

$$\|\hat{\mathbf{A}}_i - \mathbf{A}_i\| \lesssim \frac{(\sigma_{\mathbf{z}} + C_K\sigma_{\mathbf{w}})}{\sigma_{\mathbf{z}}} \frac{\tau_{\boldsymbol{\mathcal{L}}}\big(C_\delta + C_{\mathbf{z}}C_B\big)}{\pi_{\min}(1-\rho_{\boldsymbol{\mathcal{L}}})} \sqrt{\frac{sL(n+p)}{T}} \Big(C\sqrt{n+p} + (C_0/\sigma_{\mathbf{w}})\sqrt{\log\big(\frac{36sL}{\delta}\big)}\Big),$$

$$\|\hat{\mathbf{B}}_i - \mathbf{B}_i\| \lesssim \frac{\sigma_{\mathbf{w}}}{\sigma_{\mathbf{z}}} \frac{\tau_{\boldsymbol{\mathcal{L}}}\big(C_\delta + C_{\mathbf{z}}C_B\big)}{\pi_{\min}(1-\rho_{\boldsymbol{\mathcal{L}}})} \sqrt{\frac{sL(n+p)}{T}} \Big(C\sqrt{n+p} + (C_0/\sigma_{\mathbf{w}})\sqrt{\log\big(\frac{36sL}{\delta}\big)}\Big). \tag{A.61}$$

Remark A.23. *Note that in Theorem A.22, with the shorthand notations defined in Tables C.1 and A.1, the premise conditions* (A.58)*,* (A.59)*, and* (A.60) *can also be interpreted as the following.*

$$c_{\mathbf{w}} = \underline{c}_{\mathbf{w}}(T,\delta) \tag{A.62}$$

$$T \geq \underline{T}_{id,N}(L,\delta,T,\rho_{\boldsymbol{\mathcal{L}}},\tau_{\boldsymbol{\mathcal{L}}}) \tag{A.63}$$

$$L \geq \underline{L}_{id}\left(\beta_0\sqrt{n},\delta,T,\rho_{\boldsymbol{\mathcal{L}}},\tau_{\boldsymbol{\mathcal{L}}},\underline{c}_{\mathbf{w}}(T,\delta),\mathbf{K}_{1:s},L\right). \tag{A.64}$$

From the definition of $\underline{L}_{id}$, one can see there exists $\underline{L} = \tilde{\mathcal{O}}(\log(T))$ such that (A.64) *holds by choosing $L = \underline{L}$. Define shorthand notation $\underline{T}_{id,N,L}(\delta,T,\rho,\tau) := \underline{T}_{id,N}(\underline{L},\delta,T,\rho_{\boldsymbol{\mathcal{L}}},\tau_{\boldsymbol{\mathcal{L}}})$, then the premise conditions* (A.58)*,* (A.59)*, and* (A.60) *can be implied by by the single condition $T \geq \underline{T}_{id,N,L}(\delta,T,\rho,\tau)$, under which the main results in Theorem A.22 still hold.*

A.3.6 Discussion

• **Sample complexity:** Here, a few remarks are in place. First, the result appears to be convoluted however most of the dependencies are logarithmic (specifically dependency on the failure probability δ and $\log(T)$ terms). Besides these, the dominant term (when estimating $\mathbf{A}$) reduces to

$$\frac{(\sigma_{\mathbf{z}} + C_K\sigma_{\mathbf{w}})}{\sigma_{\mathbf{z}}} \frac{\tau_{\boldsymbol{\mathcal{L}}}(n+p)}{\pi_{\min}(1-\rho_{\boldsymbol{\mathcal{L}}})} \sqrt{\frac{s}{T}}.$$

which is identical to our statement in Theorem 3.1. Note that the overall sample complexity grows as $T \gtrsim s(n+p)^2/\pi_{\min}^2$. We remark that, this quadratic growth is somewhat undesirable. A degrees-of-freedom counting argument would lead to an ideal dependency of $T \gtrsim s(n+p)/\pi_{\min}$. The reason is that, each vector state equation we fit has n scalar equations. The total degrees of freedom for each dynamics pair $(\mathbf{A}_i, \mathbf{B}_i)$ is $n \times (n+p)$. Additionally, for the least-frequent mode, in steady-state, we should observe $\pi_{\min}T$ equations. Putting these together, we would minimally need $n \times \pi_{\min}T \geq n \times (n+p)$, which means we need $T \geq$

$s(n+p)/\pi_{\min}$ samples to estimate s dynamic pairs $(\mathbf{A}_{1:s}, \mathbf{B}_{1:s})$. Our analysis indicates that this suboptimality (at least the quadratic growth in n) can be addressed to achieve optimal dependence by establishing a stronger control on the state covariance (e.g. refining (A.19)) as well as a better control on the degree of independence across sampled states (this issue arises during the proof of Theorem A.16).

• **To what extent sub-sampling is necessary?** We recall that our argument is based on mixing time arguments which are well-studied in the literature. In Algorithm 1, we sub-sample the trajectory for bounded samples and use them to estimate the unknown MJS dynamics. Unfortunately, such a sub-sampling seems unavoidable as long as we don't have a good tail control on the distribution of the state vectors. Specifically, as long as the feature vectors (in our case state vectors) are allowed to be heavy-tailed, existing – to the best of our knowledge – minimum singular value concentration guarantees for the empirical covariance apply under the assumption of boundedness (Vershynin, 2012). More recently, self-normalized martingale arguments are employed to address temporal dependencies (Sarkar and Rakhlin, 2019; Simchowitz et al., 2018). We remark that, using martingale-based arguments, it may be possible to mitigate the spectral radius dependency by shaving a factor of $1/(1-\rho_{\boldsymbol{\mathcal{L}}})$ (e.g. martingale based arguments have milder $\rho_{\boldsymbol{\mathcal{L}}}$ dependence (Simchowitz et al., 2018; Sarkar and Rakhlin, 2019)). We have left this as a possible future work.

A.4 Lower Bounding $|\bar{S}_i^{(\ell)}|$

To begin, we define sub-sampling period $L = C_{sub} \log(T)$, sub-sampling indices $\ell_k = \ell + kL$ for $k = 1, 2, \ldots, \lfloor T/L \rfloor$, and the time index set

$$S_i^{(\ell)} = \left\{ \ell_k \mid \omega_{\ell_k} = i, \quad \|\mathbf{x}_{\ell_k}\| \le c_{\mathbf{x}} \sqrt{\|\mathbf{\Sigma}_{\mathbf{w}}\| \log(T)}, \quad \|\mathbf{z}_{\ell_k}\| \le c_{\mathbf{z}} \sqrt{\|\mathbf{\Sigma}_{\mathbf{z}}\|} \right\} \tag{A.65}$$

by bounding $\|\mathbf{x}_t\|$ and $\|\mathbf{z}_t\|$, which is used to estimate $\mathbf{A}_{1:s}$ and $\mathbf{B}_{1:s}$ through least squares (Here we generalize isotropic noise $\mathbf{w}_t \sim \mathcal{N}(0, \sigma_{\mathbf{w}}^2 \mathbf{I}_n)$ and $\mathbf{z}_t \sim \mathcal{N}(0, \sigma_{\mathbf{z}}^2 \mathbf{I}_p)$ to $\mathcal{N}(0, \mathbf{\Sigma}_{\mathbf{w}})$ and $\mathcal{N}(0, \mathbf{\Sigma}_{\mathbf{z}})$, respectively.). A fundamental question is: Is $|S_i^{(\ell)}|$ big enough such that there will be enough data available when applying least squares? We provide answer to this question in this section. Lemma A.24 acts as a building block for the later result; Lemma A.25 provides the lower bound on $|S_i^{(\ell)}|$; Corollary A.26 gives a more interpretable lower bound on $|S_i^{(\ell)}|$ when $c_{\mathbf{x}}$ and $c_{\mathbf{z}}$ are large enough; and finally, Lemma A.27 shows how many samples in $S_i^{(\ell)}$ are "weakly" independent, which is the quantity that essentially determines the sample complexity of estimating $\mathbf{A}_{1:s}$ and $\mathbf{B}_{1:s}$.

For clarity, we reiterate some definitions and define a few new ones here. We are given an

MJS($\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}$) with process noise $\mathbf{w}_t \sim \mathcal{N}(0, \mathbf{\Sigma_w})$ and ergodic Markov matrix $\mathbf{T}$. With some stabilizing controller $\mathbf{K}_{1:s}$, the input is given by $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$ where $\mathbf{z}_t \sim \mathcal{N}(0, \mathbf{\Sigma_z})$. Let $\bar{\sigma}^2 := \|\mathbf{B}_{1:s}\|^2\|\mathbf{\Sigma_z}\| + \|\mathbf{\Sigma_w}\|$. Let $\mathbf{L}_i := \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i$. Let $\boldsymbol{\mathcal{L}} \in \mathbb{R}^{sn^2 \times sn^2}$ denote the augmented closed-loop state matrix with (i,j)-th $n^2 \times n^2$ block given by $[\boldsymbol{\mathcal{L}}]_{ij} := \mathbf{T}(j,i)\mathbf{L}_j \otimes \mathbf{L}_j$. Let $\rho_{\boldsymbol{\mathcal{L}}} \in [0,1)$ and $\tau_{\boldsymbol{\mathcal{L}}} > 0$ be two constants such that $\|\boldsymbol{\mathcal{L}}^k\| \leq \tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^k$. By definition, one available choice for $\tau_{\boldsymbol{\mathcal{L}}}$ and $\rho_{\boldsymbol{\mathcal{L}}}$ are $\tau_{\boldsymbol{\mathcal{L}}}$ and $\rho(\boldsymbol{\mathcal{L}})$, respectively. Let $t_{\mathrm{MC}}(\cdot)$ and t_{MC} denote the mixing time of $\mathbf{T}$ as in Definition 2.5. Let $\boldsymbol{\pi}_\infty$ denote the stationary distribution of $\mathbf{T}$, $\pi_{\min} = \min_i \boldsymbol{\pi}_\infty(i)$, and $\pi_{\max} = \max_i \boldsymbol{\pi}_\infty(i)$. Assume the initial state $\mathbf{x}_0$ satisfies $\mathbb{E}[\|\mathbf{x}_0\|^2] \leq \bar{x}_0^2$ for some $\bar{x}_0 \geq 0$. Lastly, without loss of generality, we consider the sub-trajectory with zero shift, that is, $\ell = 0$, which is identical to any $0 \leq \ell \leq L-1$.

Lemma A.24. *Suppose the Markov chain trajectory $\{\omega_0, \omega_1, \dots\}$ and a sequence of events $\{A_0, A_1, \dots\}$ are both adapted to filtration $\{\mathcal{F}_0, \mathcal{F}_1, \dots\}$, i.e. ω_t and $\mathbf{1}_{\{A_t\}}$ are both $\mathcal{F}_t$-measurable. We assume $\mathbb{E}[\mathbf{1}_{\{\omega_t = j\}} \mid \mathcal{F}_{t-r}] = \mathbb{P}(\omega_t = j \mid \omega_{t-r})$ for all $j \in [s], t$, and $r < t$. For all $i \in [s]$, let*

$$N_i = \sum_{k=1}^{\lfloor T/L \rfloor} \mathbf{1}_{\{\omega_{kL}=i\}}\mathbf{1}_{\{A_{kL}\}} \tag{A.66}$$

and suppose

$$\mathbb{E}[\mathbf{1}_{\{A_t\}} \mid \mathcal{F}_{t-L}] \geq 1 - p_t \tag{A.67}$$

for some $p_t \in [0,1)$ and $L = C_{sub}\log(T)$. Assume $C_{sub} \geq C_{MC}$, and for some $\delta > 0$, $T \geq \underline{T}_{MC,1}(C_{sub}, \delta)$, where C_{MC} and $\underline{T}_{MC,1}(\mathcal{C}, \delta)$ are defined in Tables C.1 and A.1, respectively. Then we have

$$\mathbb{P}\left(\bigcap_{i=1}^{s}\left\{N_i \geq \frac{T\boldsymbol{\pi}_\infty(i)}{C_{sub}\log(T)}\left(1 - \frac{1}{\boldsymbol{\pi}_\infty(i)}\sqrt{\log(\frac{s}{\delta})\frac{17C_{sub}\pi_{\max}\log(T)}{T}}\right) - \sum_{k=1}^{\lfloor T/L \rfloor} p_{kL}\right\}\right) \geq 1 - \delta.$$

Proof. For some $\epsilon < \pi_{\min}/2$, we temporarily let $L \geq 6t_{\mathrm{MC}}\log(\epsilon^{-1})$. From the proof of Du et al. (2019b, Lemma 13 (47)), we know this guarantees $L \geq t_{\mathrm{MC}}(\epsilon/2)$. By definition of $t_{\mathrm{MC}}(\cdot)$, we know $\max_i \|([\mathbf{T}^L](i,:))^\intercal - \boldsymbol{\pi}_\infty\|_1 \leq \epsilon \leq \pi_{\min}/2$, and since $([\mathbf{T}^L](i,:))\mathbf{1} = \boldsymbol{\pi}_\infty^\intercal \mathbf{1} = 1$, we further have

$$\max_i \|([\mathbf{T}^L](i,:))^\intercal - \boldsymbol{\pi}_\infty\|_\infty \leq \frac{\epsilon}{2} \leq \frac{\pi_{\min}}{4}. \tag{A.68}$$

For simplicity, we assume $\lfloor T/L \rfloor = T/L =: \tilde{T}$. To ease the notation, we let $\tilde{\omega}_k := \omega_{kL}$, $\tilde{A}_k := A_{kL}$, and $\tilde{\mathcal{F}}_k := \mathcal{F}_{kL}$. Then, one can see $\tilde{\omega}_k$ and $\tilde{A}_k$ are both $\tilde{\mathcal{F}}_k$-measurable. Define $\boldsymbol{\delta}_k, \boldsymbol{\Delta}_k \in \mathbb{R}^s$ such that

$$\boldsymbol{\delta}_j(i) := \mathbf{1}_{\{\tilde{\omega}_j = i\}}\mathbf{1}_{\{\tilde{A}_j\}} - \mathbb{E}[\mathbf{1}_{\{\tilde{\omega}_j = i\}}\mathbf{1}_{\{\tilde{A}_j\}} \mid \tilde{\mathcal{F}}_{j-1}], \tag{A.69}$$

$$\boldsymbol{\Delta}_k(i) := \sum_{j=1}^{k} \boldsymbol{\delta}_j(i). \tag{A.70}$$

Note that for all $i \in [s]$, $\{\boldsymbol{\Delta}_k(i), \tilde{\mathcal{F}}_k\}$ forms a martingale as

$$\begin{aligned}
\mathbb{E}[\boldsymbol{\Delta}_{k+1}(i) \mid \tilde{\mathcal{F}}_k] &= \mathbb{E}[\sum_{j=1}^{k+1} \boldsymbol{\delta}_j(i) \mid \tilde{\mathcal{F}}_k] \\
&= \sum_{j=1}^{k} \boldsymbol{\delta}_j(i) + \mathbb{E}[\mathbf{1}_{\{\tilde{\omega}_{k+1}=i\}} \mathbf{1}_{\{\tilde{A}_{k+1}\}} - \mathbb{E}[\mathbf{1}_{\{\tilde{\omega}_{k+1}=i\}} \mathbf{1}_{\{\tilde{A}_{k+1}\}} \mid \tilde{\mathcal{F}}_k] \mid \tilde{\mathcal{F}}_k] \qquad (A.71) \\
&= \sum_{j=1}^{k} \boldsymbol{\delta}_j(i) = \boldsymbol{\Delta}_k(i),
\end{aligned}$$

thus $\boldsymbol{\delta}_k(i) = \boldsymbol{\Delta}_k(i) - \boldsymbol{\Delta}_{k-1}(i)$ can be viewed as the martingale difference sequence. Since $\mathbb{E}[\boldsymbol{\delta}_k(i) \mid \tilde{\mathcal{F}}_{k-1}] = 0$, we have $\mathbb{E}[\boldsymbol{\delta}_k(i)^2 \mid \tilde{\mathcal{F}}_{k-1}] = \mathrm{Var}(\boldsymbol{\delta}_k(i) \mid \tilde{\mathcal{F}}_{k-1}) = \mathrm{Var}(\mathbf{1}_{\{\tilde{\omega}_j=i\}} \mathbf{1}_{\{\tilde{A}_j\}} \mid \tilde{\mathcal{F}}_{k-1}) \le \mathbb{E}[\mathbf{1}^2_{\{\tilde{\omega}_k=i\}} \mathbf{1}^2_{\{\tilde{A}_k\}} \mid \tilde{\mathcal{F}}_{k-1}] \le \mathbb{E}[\mathbf{1}_{\{\tilde{\omega}_k=i\}} \mid \tilde{\mathcal{F}}_{k-1}] = \mathbb{P}(\tilde{\omega}_k = i \mid \tilde{\omega}_{k-1}) = [\mathbf{T}^L](\omega_{(k-1)L}, i)$. By the choice of L, using (A.68), we know $[\mathbf{T}^L](\omega_{(k-1)L}, i) \le \boldsymbol{\pi}_\infty(i) + \max_j \|([\mathbf{T}^L](j,:))^\intercal - \boldsymbol{\pi}_\infty\|_\infty \le 2\pi_{\max}$. Thus,

$$\sum_{k=1}^{\tilde{T}} \mathbb{E}[\boldsymbol{\delta}_k(i)^2 \mid \tilde{\mathcal{F}}_{k-1}] \le 2\pi_{\max}\tilde{T}. \tag{A.72}$$

With this, and the fact that $|\boldsymbol{\delta}_k(i)| < 1$, we have

$$\begin{aligned}
\mathbb{P}\left(N_i - \sum_{k=1}^{\tilde{T}} \mathbb{E}[\mathbf{1}_{\{\tilde{\omega}_k=i\}} \mathbf{1}_{\{\tilde{A}_k\}} \mid \tilde{\mathcal{F}}_{k-1}] \ge \tilde{T}\frac{\epsilon}{2}\right) &\overset{\text{(i)}}{=} \mathbb{P}(\boldsymbol{\Delta}_{\tilde{T}}(i) \ge \tilde{T}\epsilon/2) \\
&\overset{\text{(ii)}}{\le} \exp(-\frac{\tilde{T}\epsilon^2/8}{2\pi_{\max} + \epsilon/6}) \qquad (A.73) \\
&\overset{\text{(iii)}}{\le} \exp(-\frac{T\epsilon^2}{17\pi_{\max}L}),
\end{aligned}$$

where (i) follows from the definition of N_i and $\boldsymbol{\Delta}_{\tilde{T}}(i)$; (ii) follows from Freedman's inequality

(Freedman, 1975, Theorem 1.6), and (iii) follows since $\epsilon \leq \pi_{\min}/2$. Note that

$$\begin{aligned}
&\left|\sum_{k=1}^{\tilde{T}} \mathbb{E}[\mathbf{1}_{\{\tilde{\omega}_k=i\}}\mathbf{1}_{\{\tilde{A}_k\}} \mid \tilde{\mathcal{F}}_{k-1}] - \tilde{T}\boldsymbol{\pi}_\infty(i)\right| \\
\leq &\left|\sum_{k=1}^{\tilde{T}} \mathbb{E}[\mathbf{1}_{\{\tilde{\omega}_k=i\}} \mid \tilde{\mathcal{F}}_{k-1}] - \tilde{T}\boldsymbol{\pi}_\infty(i)\right| + \left|\sum_{k=1}^{\tilde{T}} \mathbb{E}[\mathbf{1}_{\{\tilde{\omega}_k=i\}} \mid \tilde{\mathcal{F}}_{k-1}] - \mathbb{E}[\mathbf{1}_{\{\tilde{\omega}_k=i\}}\mathbf{1}_{\{\tilde{A}_k\}} \mid \tilde{\mathcal{F}}_{k-1}]\right| \\
\leq &\tilde{T}\max_j |[\mathbf{T}^L](j,i) - \boldsymbol{\pi}_\infty(i)| + \left|\sum_{k=1}^{\tilde{T}} \mathbb{E}[\mathbf{1}_{\{\tilde{A}_k^c\}} \mid \tilde{\mathcal{F}}_{k-1}]\right| \\
\leq &\tilde{T}\frac{\epsilon}{2} + \sum_{k=1}^{\tilde{T}} p_{kL}.
\end{aligned}$$

Then, combining this with (A.73) and applying union bound, we have with probability at least $1 - s\exp(-\frac{T\epsilon^2}{17\pi_{\max}L})$,

$$\bigcap_{i=1}^{s}\left\{N_i \geq \frac{T}{L}\boldsymbol{\pi}_\infty(i) - \frac{T}{L}\epsilon - \sum_{k=1}^{T/L} p_{kL}\right\} \tag{A.74}$$

when $\epsilon < \pi_{\min}/2$ and $L \geq 6t_{\mathrm{MC}}\log(\epsilon^{-1})$. Then, similar to the proof of Lemma A.6, we know if we pick $L = C_{sub}\log(T)$ with $C_{sub} \geq t_{\mathrm{MC}} \cdot \max\{3, 3 - 3\log(\pi_{\max}\log(s))\}$, and for some $\delta > 0$, we pick the trajectory length $T \geq \left(68C_{sub}\pi_{\max}\pi_{\min}^{-2}\log(\frac{s}{\delta})\right)^2$, with probability at least $1-\delta$, we have

$$\bigcap_{i=1}^{s}\left\{N_i \geq \frac{T\boldsymbol{\pi}_\infty(i)}{C_{sub}\log(T)}\left(1 - \frac{1}{\boldsymbol{\pi}_\infty(i)}\sqrt{\log(\frac{s}{\delta})\frac{17C_{sub}\pi_{\max}\log(T)}{T}}\right) - \sum_{k=1}^{T/L} p_{kL}\right\}. \tag{A.75}$$

□

Lemma A.25. *For some $\delta > 0$, we assume $C_{sub} \geq \max\{C_{MC}, \underline{C}_{sub,\mathbf{x}}(\bar{x}_0, \delta, T, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})\}$, $c_{\mathbf{z}} \geq (\sqrt{3} + \sqrt{6})\sqrt{p}$, and $T \geq \max\{\underline{T}_{MC,1}(C_{sub}, \frac{\delta}{2}), \underline{T}_{cl,1}(\rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})\}$, where C_{MC}, $\underline{T}_{MC,1}(\mathcal{C}, \delta)$ and $\underline{T}_{cl,1}(\rho, \tau)$ are defined in Table C.1, and $\underline{C}_{sub,\mathbf{x}}(\bar{x}_0, \delta, T, \rho, \tau)$ is defined in Table A.1. Then, with probability at least $1-\delta$, the following intersected events occur*

$$\begin{aligned}
\bigcap_{i=1}^{s}\Bigg\{|S_i^{(\ell)}| \geq \frac{T\boldsymbol{\pi}_\infty(i)}{C_{sub}\log(T)}\Bigg(1 - \frac{1}{\boldsymbol{\pi}_\infty(i)}&\sqrt{\log(\frac{2s}{\delta})\frac{17C_{sub}\pi_{\max}\log(T)}{T}} \\
&- \frac{2n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{\boldsymbol{\pi}_\infty(i)c_{\mathbf{x}}^2\|\boldsymbol{\Sigma}_{\mathbf{w}}\|\log(T)(1-\rho_{\boldsymbol{\mathcal{L}}})} - \frac{1}{\boldsymbol{\pi}_\infty(i)}e^{-\frac{c_{\mathbf{z}}^2}{3}}\Bigg\}.
\end{aligned} \tag{A.76}$$

Proof. For simplicity, we assume $\lfloor T/L \rfloor = T/L$. We let $\mathcal{F}_t$ denote the sigma algebra generated by $\{\{\omega_r\}_{r=0}^t, \mathbf{w}_{0:t}, \mathbf{z}_{0:t}, \mathbf{x}_0\}$, and let $A_t = \{\|\mathbf{x}_t\| \le c_{\mathbf{x}}\sqrt{\|\mathbf{\Sigma_w}\|\log(T)}, \|\mathbf{z}_t\| \le c_{\mathbf{z}}\sqrt{\|\mathbf{\Sigma_z}\|}\}$, then by Lemma A.24, when $C_{sub} \ge C_{MC} = t_{\mathrm{MC}} \cdot \max\{3, 3 - 3\log(\pi_{\max}\log(s))\}$ and $T \ge \underline{T}_{MC,1}(C_{sub}, \frac{\delta}{2}) = \left(68 C_{sub}\pi_{\max}\pi_{\min}^{-2}\log(\frac{2s}{\delta})\right)^2$, with probability at least $1 - \frac{\delta}{2}$, we have

$$\bigcap_{i=1}^{s} \Bigg\{ |S_i^{(\ell)}| \ge \frac{T\boldsymbol{\pi}_\infty(i)}{C_{sub}\log(T)} \cdot \left(1 - \frac{1}{\boldsymbol{\pi}_\infty(i)}\sqrt{\log(\frac{2s}{\delta})\frac{17C_{sub}\pi_{\max}\log(T)}{T}}\right) \\ - \underbrace{\sum_{k=1}^{T/L} \mathbb{P}\left(A_t^c \mid \mathbf{x}_{(k-1)L+1}, \omega_{(k-1)L}\right)}_{=:P} \Bigg\}. \tag{A.77}$$

For term P, we have

$$\begin{aligned} P &= \sum_{k=1}^{T/L} \mathbb{P}\left(\|\mathbf{x}_{kL}\| \ge c_{\mathbf{x}}\sqrt{\|\mathbf{\Sigma_w}\|\log(T)} \bigcup \|\mathbf{z}_{kL}\| \ge c_{\mathbf{z}}\sqrt{\|\mathbf{\Sigma_z}\|} \mid \mathbf{x}_{(k-1)L+1}, \omega_{(k-1)L}\right) \\ &\le \underbrace{\sum_{k=1}^{T/L} \mathbb{P}\left(\|\mathbf{z}_{kL}\| \ge c_{\mathbf{z}}\sqrt{\|\mathbf{\Sigma_z}\|} \mid \mathbf{x}_{(k-1)L+1}, \omega_{(k-1)L}\right)}_{=:P_1} \\ &\quad + \underbrace{\sum_{k=1}^{T/L} \mathbb{P}\left(\|\mathbf{x}_{kL}\| \ge c_{\mathbf{x}}\sqrt{\|\mathbf{\Sigma_w}\|\log(T)} \mid \mathbf{x}_{(k-1)L+1}, \omega_{(k-1)L}\right)}_{=:P_2}. \end{aligned} \tag{A.78}$$

For term P_1, we know from Lemma A.1 that when $c_{\mathbf{z}} \ge (\sqrt{3} + \sqrt{6})\sqrt{p}$, we have $P_1 = \sum_{k=1}^{T/L} \mathbb{P}\left(\|\mathbf{z}_{kL}\| \ge c_{\mathbf{z}}\sqrt{\|\mathbf{\Sigma_z}\|}\right) \le \frac{T}{L}e^{-\frac{c_{\mathbf{z}}^2}{3}}$. Now we consider term P_2. From Lemma 2.15, we know

$$\mathbb{E}[\|\mathbf{x}_{kL}\|^2 \mid \mathbf{x}_{(k-1)L+1}, \omega_{(k-1)L}] \le \sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{L-1}\|\mathbf{x}_{(k-1)L+1}\|^2 + \frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{1 - \rho_{\boldsymbol{\mathcal{L}}}}, \tag{A.79}$$

thus by Markov inequality, we have

$$\begin{aligned} P_2 &\le \sum_{k=1}^{T/L} \frac{1}{c_{\mathbf{x}}^2\|\mathbf{\Sigma_w}\|\log(T)}\left(\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{L-1}\|\mathbf{x}_{(k-1)L+1}\|^2 + \frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{1 - \rho_{\boldsymbol{\mathcal{L}}}}\right) \\ &\le \frac{1}{c_{\mathbf{x}}^2\|\mathbf{\Sigma_w}\|\log(T)}\left(\frac{T}{L}\frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{1 - \rho_{\boldsymbol{\mathcal{L}}}} + \sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{L-1}\sum_{k=1}^{T/L}\|\mathbf{x}_{(k-1)L+1}\|^2\right). \end{aligned} \tag{A.80}$$

Now, we seek to upper bound $\rho_{\boldsymbol{\mathcal{L}}}^{L-1}\sum_{k=1}^{T/L}\|\mathbf{x}_{(k-1)L+1}\|^2$ with high probability. Note that the assumption $C_{sub} \geq \underline{C}_{sub,\mathbf{x}}(\bar{x}_0, \delta, T, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})$ implies the following

$$L = C_{sub}\log(T) \geq \frac{1}{\log(\rho_{\boldsymbol{\mathcal{L}}}^{-1})}\max\left\{\log(2), 2\log(\frac{8\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\bar{x}_0^2}{\delta}), 2\log(4T\frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{\delta(1-\rho_{\boldsymbol{\mathcal{L}}})}) + 2,\right\}. \quad \text{(A.81)}$$

Then, we have

$$\begin{aligned}
\mathbb{P}\left(\rho_{\boldsymbol{\mathcal{L}}}^{L-1}\sum_{k=1}^{T/L}\|\mathbf{x}_{(k-1)L+1}\|^2 \leq \frac{T}{L}\frac{(1-\rho_{\boldsymbol{\mathcal{L}}})}{4Tn\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}\right) &\overset{\text{(i)}}{\geq} \mathbb{P}\left(\rho_{\boldsymbol{\mathcal{L}}}^{L-1}\sum_{k=1}^{T/L}\|\mathbf{x}_{(k-1)L+1}\|^2 \leq \frac{T}{L}\rho_{\boldsymbol{\mathcal{L}}}^{\frac{L}{2}-1}\right)\\
&\geq \mathbb{P}\left(\bigcap_{k=1}^{T/L}\left\{\|\mathbf{x}_{(k-1)L+1}\|^2 \leq \rho_{\boldsymbol{\mathcal{L}}}^{-\frac{L}{2}}\right\}\right)\\
&\geq 1 - \sum_{k=1}^{T/L}\mathbb{P}\left(\|\mathbf{x}_{(k-1)L+1}\|^2 \geq \rho_{\boldsymbol{\mathcal{L}}}^{-\frac{L}{2}}\right)\\
&\overset{\text{(ii)}}{\geq} 1 - \sum_{k=1}^{T/L}\rho_{\boldsymbol{\mathcal{L}}}^{\frac{L}{2}}\left(\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(k-1)L+1}\bar{x}_0^2 + \frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{1-\rho_{\boldsymbol{\mathcal{L}}}}\right)\\
&\geq 1 - \rho_{\boldsymbol{\mathcal{L}}}^{\frac{L}{2}+1}\frac{\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\bar{x}_0^2}{1-\rho_{\boldsymbol{\mathcal{L}}}^L} - \frac{T}{L}\rho_{\boldsymbol{\mathcal{L}}}^{\frac{L}{2}}\frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{1-\rho_{\boldsymbol{\mathcal{L}}}}\\
&\overset{\text{(iii)}}{\geq} 1 - 2\rho_{\boldsymbol{\mathcal{L}}}^{\frac{L}{2}}\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\bar{x}_0^2 - \frac{\delta}{4L}\\
&\overset{\text{(iv)}}{\geq} 1 - \frac{\delta}{4} - \frac{\delta}{4} = 1 - \frac{\delta}{2},
\end{aligned}$$

where (i) follows from (A.81) which gives $\rho_{\boldsymbol{\mathcal{L}}}^{\frac{L}{2}-1} \leq \frac{(1-\rho_{\boldsymbol{\mathcal{L}}})}{4Tn\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}$; (ii) follows from Lemma 2.15 and Markov inequality; (iii) follows from (A.81) which gives $\rho_{\boldsymbol{\mathcal{L}}}^L \leq \frac{1}{2}$ and $\rho_{\boldsymbol{\mathcal{L}}}^{\frac{L}{2}} \leq \frac{\delta(1-\rho_{\boldsymbol{\mathcal{L}}})}{4Tn\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}$ and (iv) follows from (A.81) which gives $\rho_{\boldsymbol{\mathcal{L}}}^{\frac{L}{2}} \leq \frac{\delta}{8\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\bar{x}_0^2}$. Therefore, we have with probability at least $1-\frac{\delta}{2}$

$$P_2 \leq \frac{1}{c_{\mathbf{x}}^2\|\mathbf{\Sigma}_{\mathbf{w}}\|\log(T)}\left(\frac{T}{L}\frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{1-\rho_{\boldsymbol{\mathcal{L}}}} + \frac{1-\rho_{\boldsymbol{\mathcal{L}}}}{4L\sqrt{n}\bar{\sigma}^2}\right), \quad \text{(A.82)}$$

and thus,

$$\begin{aligned}
P \leq P_1 + P_2 &\leq \frac{1}{c_{\mathbf{x}}^2\|\mathbf{\Sigma}_{\mathbf{w}}\|\log(T)}\left(\frac{T}{L}\frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{1-\rho_{\boldsymbol{\mathcal{L}}}} + \frac{1-\rho_{\boldsymbol{\mathcal{L}}}}{4L\sqrt{n}\bar{\sigma}^2}\right) + \frac{T}{L}e^{-\frac{c_{\mathbf{z}}^2}{3}}\\
&\leq \frac{1}{c_{\mathbf{x}}^2\|\mathbf{\Sigma}_{\mathbf{w}}\|\log(T)}\left(\frac{T}{L}\frac{2n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\bar{\sigma}^2}{1-\rho_{\boldsymbol{\mathcal{L}}}}\right) + \frac{T}{L}e^{-\frac{c_{\mathbf{z}}^2}{3}},
\end{aligned} \quad \text{(A.83)}$$

where the second inequality follows from $T \geq \underline{T}_{cl,1}(\rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})$. Plugging this into (A.77), we have with probability at least $1-\delta$,

$$\bigcap_{i=1}^{s}\left\{|S_i^{(\ell)}| \geq \frac{T\boldsymbol{\pi}_\infty(i)}{C_{sub}\log(T)}\cdot\left(1-\frac{1}{\boldsymbol{\pi}_\infty(i)}\sqrt{\log(\frac{2s}{\delta})\frac{17C_{sub}\pi_{\max}\log(T)}{T}}\right.\right.$$
$$\left.\left.-\frac{2n\sqrt{s}\tau_{\mathcal{L}}\bar{\sigma}^2}{\boldsymbol{\pi}_\infty(i)c_{\mathbf{x}}^2\|\boldsymbol{\Sigma}_{\mathbf{w}}\|\log(T)(1-\rho_{\mathcal{L}})}-\frac{1}{\boldsymbol{\pi}_\infty(i)}e^{-\frac{c_{\mathbf{z}}^2}{3}}.\right)\right\}, \quad \text{(A.84)}$$

which concludes the proof. □

Note that when T, $c_{\mathbf{x}}$, and $c_{\mathbf{z}}$ are sufficiently large enough, we could obtain a more interpretable version of Lemma A.25 which is presented as follows.

Corollary A.26. *For some $\delta>0$, assume $C_{sub}\geq\max\{C_{MC},\underline{C}_{sub,\mathbf{x}}(\bar{x}_0,\delta,T,\rho_{\mathcal{L}},\tau_{\mathcal{L}})\}$, $c_{\mathbf{x}}\geq\frac{1}{3}\underline{c}_{\mathbf{x}}(\rho_{\mathcal{L}},\tau_{\mathcal{L}})$, $c_{\mathbf{z}}\geq\underline{c}_{\mathbf{z}}$ and $T\geq\underline{T}_N(C_{sub},2\delta,\rho_{\mathcal{L}},\tau_{\mathcal{L}}):=\max\{\underline{T}_{MC}(C_{sub},\delta),\underline{T}_{cl,1}(\rho_{\mathcal{L}},\tau_{\mathcal{L}})\}$, where $\underline{c}_{\mathbf{x}}(\rho,\tau)$, $\underline{c}_{\mathbf{z}}$, C_{MC}, $\underline{T}_{MC}(\mathcal{C},\delta)$, $\underline{T}_{cl,1}(\rho,\tau)$ are defined in Table C.1 and $\underline{C}_{sub,\mathbf{x}}(\bar{x}_0,\delta,T,\rho,\tau)$ is defined in Table A.1. Then, with probability at least $1-\delta/2$, the following intersected events occur*

$$\bigcap_{i=1}^{s}\left\{|S_i^{(\ell)}|\geq\frac{T\pi_{\min}}{2C_{sub}\log(T)}\right\}. \quad \text{(A.85)}$$

Now we provide a result on how many data in $S_i^{(\ell)}$ are "weakly" independent, which is the quantity that essentially determines the sample complexity of estimating $\mathbf{A}_{1:s}$ and $\mathbf{B}_{1:s}$ in Algorithm 1. We first define a few notations. Let $\ell_{i,1},\ldots,\ell_{i,|S_i^{(\ell)}|}$ denote the elements in $S_i^{(\ell)}$, and let $\ell_{i,0}=0$. Define $\bar{\mathbf{x}}_{\ell_{i,k}}$ such that

$$\bar{\mathbf{x}}_{\ell_{i,k}}=\sum_{j=1}^{\ell_{i,k-1}}\left(\prod_{k=1}^{j-1}\mathbf{L}_{\omega_{t-k}}\right)\left(\mathbf{B}_{\omega_{t-j}}\mathbf{z}_{t-j}+\mathbf{w}_{t-j}\right)+\mathbf{B}_{\omega_{\ell_{i,k}-1}}\mathbf{z}_{\ell_{i,k}-1}+\mathbf{w}_{\ell_{i,k}-1}. \quad \text{(A.86)}$$

One can view $\bar{\mathbf{x}}_{\ell_{i,k}}$ as follows: set $\mathbf{x}_{\ell_{i,k-1}}=0$, then propagate the dynamics to time $\ell_{i,k}$ following the same noise and mode switching sequences, $\mathbf{w}_{\ell_{i,k-1}:\ell_{i,k}-1}$, $\mathbf{z}_{\ell_{i,k-1}:\ell_{i,k}-1}$, $\{\omega_{t'}\}_{t'=\ell_{i,k-1}}^{\ell_{i,k}-1}$. Or, one can also view $\bar{\mathbf{x}}_{\ell_{i,k}}$ as the contribution of noise $\mathbf{x}_t$ and $\mathbf{z}_t$ that propagate $\mathbf{x}_{\ell_{i,k-1}}$ to $\mathbf{x}_{\ell_{i,k}}$. And it is easy to see that

$$\mathbf{x}_{\ell_{i,k}}-\bar{\mathbf{x}}_{\ell_{i,k}}=\left(\prod_{k=1}^{\ell_{i,k-1}}\mathbf{L}_{\omega_{t-k}}\right)\mathbf{x}_{\ell_{i,k-1}}. \quad \text{(A.87)}$$

Define $\bar{S}_i^{(\ell)} \subseteq S_i^{(\ell)}$ such that

$$\bar{S}_i^{(\ell)} := \left\{ \ell_k \middle| \omega_{\ell_k} = i, \|\mathbf{x}_{\ell_k}\| \leq c_{\mathbf{x}} \sqrt{\|\boldsymbol{\Sigma}_{\mathbf{w}}\| \log(T)}, \|\mathbf{z}_{\ell_k}\| \leq c_{\mathbf{z}} \sqrt{\boldsymbol{\Sigma}_{\mathbf{z}}}, \|\bar{\mathbf{x}}_{\ell_k}\| \leq \frac{c_{\mathbf{x}} \sqrt{\|\boldsymbol{\Sigma}_{\mathbf{w}}\| \log(T)}}{2} \right\}.$$

The next lemma provides a lower bound on $|\bar{S}_i^{(\ell)}|$.

Lemma A.27. *Assume $c_{\mathbf{x}} \geq \underline{c}_{\mathbf{x}}(\rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})$, $c_{\mathbf{z}} \geq \underline{c}_{\mathbf{z}}$, $C_{sub} \geq \underline{C}_{sub,N}(\bar{x}_0, \delta, T, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}}) := \max\{C_{MC}, \underline{C}_{sub,\mathbf{x}}(\bar{x}_0, \frac{\delta}{2}, T, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}}), \underline{C}_{sub,\bar{\mathbf{x}}}(\delta, T, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})\}$, and $T \geq \underline{T}_N(C_{sub}, \delta, \rho_{\boldsymbol{\mathcal{L}}}, \tau_{\boldsymbol{\mathcal{L}}})$, where $\underline{c}_{\mathbf{x}}(\rho, \tau)$, $\underline{c}_{\mathbf{z}}$, $\underline{C}_{sub,\mathbf{x}}(\bar{x}_0, \delta, T, \rho, \tau)$, and $\underline{C}_{sub,\bar{\mathbf{x}}}(\delta, T, \rho, \tau)$ are defined in Table A.1, and C_{MS} and $\underline{T}_N(\mathcal{C}, \delta, \rho, \tau)$ are defined in Table C.1. Then with probability at least $1-\delta$, the following intersected events occur*

$$\bigcap_{i=1}^{s} \left\{ |\bar{S}_i^{(\ell)}| \geq \frac{T\pi_{\min}}{2C_{sub}\log(T)} \right\}. \tag{A.88}$$

Proof. We define sets $R_i^{(\ell)} \subseteq S_i^{(\ell)}$ and $\bar{R}_i^{(\ell)} \subseteq \bar{S}_i^{(\ell)}$ such that

$$R_i^{(\ell)} := \left\{ \ell_k \middle| \omega_{\ell_k} = i, \|\mathbf{x}_{\ell_k}\| \leq \frac{c_{\mathbf{x}} \sqrt{\|\boldsymbol{\Sigma}_{\mathbf{w}}\| \log(T)}}{3}, \|\mathbf{z}_{\ell_k}\| \leq c_{\mathbf{z}} \sqrt{\boldsymbol{\Sigma}_{\mathbf{z}}} \right\}.$$
$$\bar{R}_i^{(\ell)} := \left\{ \ell_k \middle| \omega_{\ell_k} = i, \|\mathbf{x}_{\ell_k}\| \leq \frac{c_{\mathbf{x}} \sqrt{\|\boldsymbol{\Sigma}_{\mathbf{w}}\| \log(T)}}{3}, \|\mathbf{z}_{\ell_k}\| \leq c_{\mathbf{z}} \sqrt{\boldsymbol{\Sigma}_{\mathbf{z}}}, \|\bar{\mathbf{x}}_{\ell_k}\| \leq \frac{c_{\mathbf{x}} \sqrt{\|\boldsymbol{\Sigma}_{\mathbf{w}}\| \log(T)}}{2} \right\}.$$

Note that $\bar{R}_i^{(\ell)} \subseteq R_i^{(\ell)}$. We will first (i) lower bound $|R_i^{(\ell)}|$ and (ii) show $|\bar{R}_i^{(\ell)}| = |R_i^{(\ell)}|$, then we could lower bound $|\bar{S}_i^{(\ell)}|$ since $|\bar{S}_i^{(\ell)}| \geq |\bar{R}_i^{(\ell)}|$ and conclude the proof.

Using Corollary A.26, we see under given assumptions, with probability at least $1 - \frac{\delta}{2}$,

$$\bigcap_{i=1}^{s} \left\{ |R_i^{(\ell)}| \geq \frac{T\pi_{\min}}{2C_{sub}\log(T)} \right\}. \tag{A.89}$$

Let $\zeta_{i,1}, \ldots, \zeta_{i,|R_i^{(\ell)}|}$ denote the elements in $R_i^{(\ell)}$. It is easy to see $\{\zeta_{i,1}, \ldots, \zeta_{i,|R_i^{(\ell)}|}\} \subseteq \{\ell_{i,1}, \ldots, \ell_{i,|S_i^{(\ell)}|}\}$. Consider an arbitrary $\zeta_{i,j} \in R_i^{(\ell)}$ and $\ell_{i,j'} \in S_i^{(\ell)}$ denote the counterpart of $\zeta_{i,j}$ such that $\ell_{i,j'} = \zeta_{i,j}$. By definition of $R_i^{(\ell)}$, we have

$$\|\mathbf{x}_{\ell_{i,j'}}\| \leq \frac{c_{\mathbf{x}} \sqrt{\|\boldsymbol{\Sigma}_{\mathbf{w}}\| \log(T)}}{3}. \tag{A.90}$$

From (A.87), together with Lemma 2.15, we have

$$\begin{aligned} \mathbb{E}[\|\mathbf{x}_{\ell_{i,j'}} - \bar{\mathbf{x}}_{\ell_{i,j'}}\|^2] &\leq \sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}} \rho_{\boldsymbol{\mathcal{L}}}^{\ell_{i,j'} - \ell_{i,j'-1}} \mathbb{E}[\|\mathbf{x}_{\ell_{i,j'-1}}\|^2] \\ &\leq \sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}} \rho_{\boldsymbol{\mathcal{L}}}^{L} (c_{\mathbf{x}}^2 \|\boldsymbol{\Sigma}_{\mathbf{w}}\| \log(T)), \end{aligned} \tag{A.91}$$

where the second inequality follows from $\ell_{i,j'} - \ell_{i,j'-1} \geq L$ and $\|\mathbf{x}_{\ell_{i,j'-1}}\| \leq c_{\mathbf{x}}\sqrt{\|\mathbf{\Sigma}_{\mathbf{w}}\|\log(T)}$ by definition of $S_i^{(\ell)}$. Then, by Markov inequality, we have

$$\mathbb{P}\left(\|\mathbf{x}_{\ell_{i,j'}} - \bar{\mathbf{x}}_{\ell_{i,j'}}\| \leq \frac{c_{\mathbf{x}}\sqrt{\|\mathbf{\Sigma}_{\mathbf{w}}\|\log(T)}}{6}\right) \geq 1 - 36\sqrt{ns}\tau_{\mathcal{L}}\rho_{\mathcal{L}}^L. \tag{A.92}$$

Then, using union bound, we have

$$\begin{aligned}
&\mathbb{P}\left(\bigcap_{i\in[s]}\bigcap_{j'}\left\{\|\mathbf{x}_{\ell_{i,j'}} - \bar{\mathbf{x}}_{\ell_{i,j'}}\| \leq \frac{c_{\mathbf{x}}\sqrt{\|\mathbf{\Sigma}_{\mathbf{w}}\|\log(T)}}{6}\right\}\right) \\
\geq& 1 - 36\sqrt{n}s^{1.5}|R_i^{(\ell)}|\tau_{\mathcal{L}}\rho_{\mathcal{L}}^L \\
\geq& 1 - 36\sqrt{n}s^{1.5}T\tau_{\mathcal{L}}\rho_{\mathcal{L}}^L \\
\geq& 1 - \frac{\delta}{2},
\end{aligned} \tag{A.93}$$

where the last line follows from $L = C_{sub}\log(T)$ and $C_{sub} \geq \underline{C}_{sub,\bar{\mathbf{x}}}(\delta, T, \rho_{\mathcal{L}}, \tau_{\mathcal{L}})$ in the assumption. Note that $\|\bar{\mathbf{x}}_{\zeta_{i,j}}\| = \|\bar{\mathbf{x}}_{\ell_{i,j'}}\| \leq \|\mathbf{x}_{\ell_{i,j'}}\| + \|\mathbf{x}_{\ell_{i,j'}} - \bar{\mathbf{x}}_{\ell_{i,j'}}\|$. This together with (A.90) and (A.93) gives, with probability at least $1 - \frac{\delta}{2}$,

$$\bigcap_{i\in[s]}\bigcap_{j\in[|R_i^{(\ell)}|]}\left\{\|\bar{\mathbf{x}}_{\zeta_{i,j}}\| \leq \frac{c_{\mathbf{x}}\sqrt{\|\mathbf{\Sigma}_{\mathbf{w}}\|\log(T)}}{2}\right\}. \tag{A.94}$$

This implies for any i, for any $\ell_k \in R_i^{(\ell)}$, we have $\ell_k \in \bar{R}_i^{(\ell)}$, i.e. $R_i^{(\ell)} \subseteq \bar{R}_i^{(\ell)}$. Thus, we have $R_i^{(\ell)} = \bar{R}_i^{(\ell)}$ and $|R_i^{(\ell)}| = |\bar{R}_i^{(\ell)}|$. Combining this with (A.89), we have with probability at least $1 - \delta$,

$$\bigcap_{i=1}^{s}\left\{|\bar{R}_i^{(\ell)}| \geq \frac{T\pi_{\min}}{2C_{sub}\log(T)}\right\}. \tag{A.95}$$

Finally, we could conclude the proof by noticing $|\bar{S}_i^{(\ell)}| \geq |\bar{R}_i^{(\ell)}|$. □

A.5 Proofs of Intermediate Theorems and Lemmas

A.5.1 Proof of Corollary A.7

Proof. Recall from Lemma 2.15 that the states $\mathbf{x}_t$ can be bounded in expectation as follows,

$$\mathbb{E}[\|\mathbf{x}_t\|^2] \leq \tau_{\mathcal{L}}\sqrt{s}\big(\rho_{\mathcal{L}}^t \mathbb{E}[\|\mathbf{x}_0\|^2]\sqrt{n} + \frac{c_{\mathbf{w}}^2\sigma_{\mathbf{w}}^2 + \sigma_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|^2}{1-\rho_{\mathcal{L}}} n\big), \\ \leq \frac{2\sigma_{\mathbf{w}}^2\sqrt{s}(c_{\mathbf{w}}^2 + C_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|^2)\tau_{\mathcal{L}} n}{1-\rho_{\mathcal{L}}}, \tag{A.96}$$

where we get the last inequality by choosing the timestep t to satisfy the following lower bound,

$$\rho_{\mathcal{L}}^t \leq \frac{(c_{\mathbf{w}}^2\sigma_{\mathbf{w}}^2 + \sigma_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|^2)\sqrt{n}}{(1-\rho_{\mathcal{L}})\mathbb{E}[\|\mathbf{x}_0\|^2]} \impliedby t \geq t_0 := \frac{\log\big((1-\rho_{\mathcal{L}})\mathbb{E}[\|\mathbf{x}_0\|^2]/(c_{\mathbf{w}}^2\sigma_{\mathbf{w}}^2 + \sigma_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|)\big)}{1-\rho_{\mathcal{L}}}. \tag{A.97}$$

This gives the advertised upper bound on $\mathbb{E}[\|\mathbf{x}_t\|^2]$ for $t \geq t_0$. Using Jensen's inequality, this further implies

$$\mathbb{E}[\|\mathbf{x}_t\|] \leq \sigma_{\mathbf{w}}\sqrt{\frac{2s^{1/2}(c_{\mathbf{w}}^2 + C_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|^2)\tau_{\mathcal{L}} n}{1-\rho_{\mathcal{L}}}} \quad \text{for} \quad t \geq t_0. \tag{A.98}$$

Next, using standard results on the distribution of squared Euclidean norm of a Gaussian vector, we have $\mathbb{E}[\|\mathbf{z}_t\|^2] = \sigma_{\mathbf{z}}^2 p$ for all $t \geq 0$. Combining this with (A.96), we get the following upper bound on the expected squared norm of $\mathbf{h}_t := [\mathbf{x}_t^\intercal/\sigma_{\mathbf{w}}\ \mathbf{z}_t^\intercal/\sigma_{\mathbf{z}}]^\intercal$, that is, for all $t \geq t_0$, we have

$$\mathbb{E}[\|\mathbf{h}_t\|^2] = \frac{1}{\sigma_{\mathbf{w}}^2}\mathbb{E}[\|\mathbf{x}_t\|^2] + \frac{1}{\sigma_{\mathbf{z}}^2}\mathbb{E}[\|\mathbf{z}_t\|^2] \leq \frac{2\sqrt{s}(c_{\mathbf{w}}^2 + C_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|^2)\tau_{\mathcal{L}} n}{1-\rho_{\mathcal{L}}} + p, \\ \leq \left(1 + \frac{2\sqrt{s}(c_{\mathbf{w}}^2 + C_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|^2)\tau_{\mathcal{L}}}{1-\rho_{\mathcal{L}}}\right)(n+p). \tag{A.99}$$

This gives the advertised upper bound on $\mathbb{E}[\|\mathbf{h}_t\|^2]$ for $t \geq t_0$. Using Jensen's inequality, this further implies

$$\mathbb{E}[\|\mathbf{h}_t\|] \leq \sqrt{\left(1 + \frac{2s^{1/2}(c_{\mathbf{w}}^2 + C_{\mathbf{z}}^2\|\mathbf{B}_{1:s}\|^2)\tau_{\mathcal{L}}}{1-\rho_{\mathcal{L}}}\right)(n+p)} \quad \text{for} \quad t \geq t_0. \tag{A.100}$$

This completes the proof. □

A.5.2 Proof of Lemma A.9

Proof. Let $\mathbf{x}_t$ be the state at time t of the MJS given by (A.14), with initial state $\mathbf{x}_0 \sim \mathcal{D}_x$ and define the noise-removed state $\tilde{\mathbf{x}}_t = \mathbf{x}_t - \mathbf{w}_{t-1}$ which is independent of $\mathbf{w}_{t-1}$. From Corollary A.7, for all $t \geq t_0$, we have $\mathbb{E}[\|\mathbf{x}_t\|] \leq \sigma_{\mathbf{w}}\beta_+\sqrt{n}$. Similarly, from Assumption A1.1, we have $\|\mathbf{w}_{t-1}\| \leq c_{\mathbf{w}}\sigma_{\mathbf{w}}\sqrt{n}$. This implies

$$\mathbb{E}[\|\tilde{\mathbf{x}}_t\|] \leq \mathbb{E}[\|\mathbf{x}_t\|] + \mathbb{E}[\|\mathbf{w}_{t-1}\|] \leq 2\sigma_{\mathbf{w}}\beta_+\sqrt{n}. \tag{A.101}$$

To proceed, consider the conditional random variable $\mathbf{y}_t \sim \{\mathbf{x}_t \mid \|\mathbf{x}_t\| \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n},\ \|\mathbf{z}_t\| \leq c\sigma_{\mathbf{z}}\sqrt{p},\ \omega_t = i\} \sim \{\mathbf{x}_t \mid \|\mathbf{x}_t\| \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n},\ \omega_t = i\}$. To lower bound the covariance matrix $\boldsymbol{\Sigma}[\mathbf{y}_t] := \mathbb{E}[\mathbf{y}_t\mathbf{y}_t^\intercal]$, observe that $\|\mathbf{w}_{t-1}\| \leq c_{\mathbf{w}}\sigma_{\mathbf{w}}\sqrt{n} \leq \sigma_{\mathbf{w}}\beta_+\sqrt{n}$ and $\mathbb{E}[\|\tilde{\mathbf{x}}_t\|] \leq 2\sigma_{\mathbf{w}}\beta_+\sqrt{n}$, where β_+ is given by (A.16). Therefore, using Lemma A.2 from Section A.1 with $B = 2\sigma_{\mathbf{w}}\beta_+\sqrt{n}$, we can lower bound $\boldsymbol{\Sigma}[\mathbf{y}_t]$ as follows,

$$\boldsymbol{\Sigma}[\mathbf{y}_t] = \mathbb{E}[\mathbf{y}_t\mathbf{y}_t^\intercal] \succeq (\sigma_{\mathbf{w}}^2/2)\mathbf{I}_n, \tag{A.102}$$

where we use $c \geq 6$ to get the last inequality. Next, we upper bound $\boldsymbol{\Sigma}[\mathbf{y}_t]$ as follows,

$$\begin{aligned}
\|\boldsymbol{\Sigma}[\mathbf{y}_t]\| = \|\mathbb{E}[\mathbf{y}_t\mathbf{y}_t^\intercal]\| &\leq \mathbb{E}[\|\mathbf{y}_t\|^2], \\
&= \mathbb{E}[\|\mathbf{x}_t\|^2 \mid \|\mathbf{x}_t\| \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n},\ \|\mathbf{z}_t\| \leq c\sigma_{\mathbf{z}}\sqrt{p},\ \omega_t = i], \\
&\leq c^2\sigma_{\mathbf{w}}^2\beta_+^2 n,
\end{aligned} \tag{A.103}$$

Combining (A.102) and (A.103), we get the first statement of the lemma. To proceed, consider another conditional random variable $\mathbf{z}'_t \sim \{\mathbf{z}_t \mid \|\mathbf{x}_t\| \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n},\ \|\mathbf{z}_t\| \leq c\sigma_{\mathbf{z}}\sqrt{p},\ \omega_t = i\} \sim \{\mathbf{z}_t \mid \|\mathbf{z}_t\| \leq c\sigma_{\mathbf{z}}\sqrt{p}\}$. Note that $\mathbf{z}_t$ is independent of both $\mathbf{x}_t$ and ω_t. Then, using similar arguments as above with Lemma A.2 replaced by Lemma A.3, we can show that, when $c \geq 6$,

$$(\sigma_{\mathbf{z}}^2/2)\mathbf{I}_p \preceq \boldsymbol{\Sigma}[\mathbf{z}'_t] := \mathbb{E}[\mathbf{z}'_t\mathbf{z}_t^{'\intercal}] \preceq (c^2\sigma_{\mathbf{z}}^2 p)\mathbf{I}_p. \tag{A.104}$$

Finally, combing the derived bounds for $\boldsymbol{\Sigma}[\mathbf{y}_t]$ and $\boldsymbol{\Sigma}[\mathbf{z}'_t]$, we get the second statement of the lemma. This completes the proof. □

A.5.3 Proof of Lemma A.12

Proof. To begin, using Assumption A5.1 and (A.28), the impact of truncation can be bounded in expectation over the modes as follows,

$$\mathbb{E}[\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\|] = \mathbb{E}[\|\mathbf{x}_{\ell+kL} - \mathbf{x}_{\ell+kL,(k-k')L-1}\|] \leq \sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{((k-k')L-1)/2}\|\mathbf{x}_{\ell+k'L+1}\|,$$
$$\leq \sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2}\|\mathbf{x}_{\ell_{k'}+1}\|, \tag{A.105}$$

where we get the last inequality from the fact that $k - k' \geq 1$, and the expectation is over the Markov modes at timesteps $\ell + k'L + 1, \ell + k'L + 2, \ldots \ell + kL - 1$. To proceed, observe that, for all $\ell_k \in S_i^{(\ell)}$, we have

$$\|\mathbf{x}_{\ell_k+1}\| = \|\mathbf{L}_{\omega_{\ell_k}}\mathbf{x}_{\ell_k} + \mathbf{B}_{\omega_{\ell_k}}\mathbf{z}_{\ell_k} + \mathbf{w}_{\ell_k}\| \leq \max_{i\in[s]}\|\mathbf{L}_i\|\|\mathbf{x}_{\ell_k}\| + \max_{i\in[s]}\|\mathbf{B}_i\|\|\mathbf{z}_{\ell_k}\| + \|\mathbf{w}_{\ell_k}\|,$$
$$\leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n}\|\mathbf{L}_{1:s}\| + c\sigma_{\mathbf{z}}\sqrt{p}\|\mathbf{B}_{1:s}\| + c_{\mathbf{w}}\sigma_{\mathbf{w}}\sqrt{n},$$
$$\leq c\sigma_{\mathbf{w}}(c_{\mathbf{w}} + \beta_+\|\mathbf{L}_{1:s}\| + C_{\mathbf{z}}\sqrt{p/n}\|\mathbf{B}_{1:s}\|)\sqrt{n},$$
$$= c\sigma_{\mathbf{w}}\beta'_+\sqrt{n}, \tag{A.106}$$

where we set $\beta'_+ := c_{\mathbf{w}} + \beta_+\|\mathbf{L}_{1:s}\| + C_{\mathbf{z}}\sqrt{p/n}\|\mathbf{B}_{1:s}\|$ and $C_{\mathbf{z}} := \sigma_{\mathbf{z}}/\sigma_{\mathbf{w}}$. Combining (A.106) with (A.105), for all $\ell_k \in S_i^{(\ell)}$ and all $i \in [s]$, we have

$$\mathbb{E}[\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\|] \leq c\sigma_{\mathbf{w}}\beta'_+ n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2}, \tag{A.107}$$
$$\implies \mathbb{P}\big(\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\| \leq \frac{c\sigma_{\mathbf{w}}\beta'_+ n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2}T}{\delta}\big) \geq 1-\delta, \tag{A.108}$$

where we get the high probability bound by using Markov inequality and union bounding over all bounded states. This further implies that, with probability at least $1-\delta$ over the modes, for all $\ell_k \in S_i^{(\ell)}$ and all $i \in [s]$, we have

$$\|\bar{\mathbf{x}}_{\ell_k}\| \leq \|\mathbf{x}_{\ell_k}\| + \|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\| \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n} + \frac{c\sigma_{\mathbf{w}}\beta'_+ n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2}T}{\delta} \leq (3/2)c\sigma_{\mathbf{w}}\beta_+\sqrt{n}, \tag{A.109}$$

where we get the last inequality by choosing $L \geq 1$ via

$$\frac{c\sigma_{\mathbf{w}}\beta'_+ n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2}T}{\delta} \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n}/2 \iff \rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2} \leq \frac{\delta\beta_+}{2\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}T\beta'_+},$$
$$\Longleftarrow L \geq 1 + \frac{2\log\left(2\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}T\beta'_+/(\beta_+\delta)\right)}{1-\rho_{\boldsymbol{\mathcal{L}}}}. \tag{A.110}$$

This also implies that, with probability at least $1-\delta$ over the modes, for all $\ell_k \in S_i^{(\ell)}$ and all $i \in [s]$, we have $\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\| \leq (1/2)c\sigma_{\mathbf{w}}\beta_+\sqrt{n}$. This completes the proof. □

A.5.4 Proof of Lemma A.14

Proof. The first statement is a direct implication of Definition A.8 and A.13. To prove the second statement, consider $\{\mathbf{x}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$ which contains states bounded by $c\sigma_{\mathbf{w}}\beta_+\sqrt{n}$. From (A.26), observe that, each state can be decomposed into $\mathbf{x}_{\ell_k} = \bar{\mathbf{x}}_{\ell_k} + \tilde{\mathbf{x}}_{\ell_k}$, where $\bar{\mathbf{x}}_{\ell_k}$ is the truncated state and $\tilde{\mathbf{x}}_{\ell_k}$ captures the impact of the past states at time index $\ell + k'L + 1$. When the sampling period L satisfies (A.33), then from Lemma A.12, with probability at least $1-\delta$ over the modes, for all $\ell_k \in S_i^{(\ell)}$ and all $i \in [s]$, we have $\|\tilde{\mathbf{x}}_{\ell_k}\| \leq (1/2)c\sigma_{\mathbf{w}}\beta_+\sqrt{n}$. Furthermore, from Definition A.13, for all $\ell_k \in \bar{S}_i^{(\ell)}$ and all $i \in [s]$, we have $\|\bar{\mathbf{x}}_{\ell_k}\| \leq (1/2)c\sigma_{\mathbf{w}}\beta_+\sqrt{n}$. Combining these results, with probability at least $1-\delta$ over the modes, for all $\ell_k \in \bar{S}_i^{(\ell)}$ and all $i \in [s]$, we have, $\|\mathbf{x}_{\ell_k}\| \leq \|\bar{\mathbf{x}}_{\ell_k}\| + \|\tilde{\mathbf{x}}_{\ell_k}\| \leq c\sigma_{\mathbf{w}}\beta_+\sqrt{n}$. This implies that, with probability at least $1-\delta$ over the modes, bounding $\mathbf{x}_{\ell_k}$ by $c\sigma_{\mathbf{w}}\beta_+\sqrt{n}$ will not introduce further dependence between $\bar{\mathbf{x}}_{\ell_k}$ and $\tilde{\mathbf{x}}_{\ell_k}$ when $\ell_k \in \bar{S}_i^{(\ell)}$. Secondly, by construction, conditioned on the modes, $\bar{\mathbf{x}}_{\ell_k} = \mathbf{x}_{\ell_k, \ell_k - \ell_{k'} - 1}$ only depends on the excitation and noise $\{\mathbf{z}_t, \mathbf{w}_t\}_{t=\ell+k'L+1}^{\ell+kL-1}$. Note that the dependence ranges $[\ell + k'L + 1, \ell + kL - 1]$ are disjoint intervals for each (k, k') pairs. Hence, conditioned on the modes, the samples in the set $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ are all independent of each other.

To show the independence of $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ and $\{\mathbf{z}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$, observe that $\mathbf{z}_{\ell_k} = \mathbf{z}_{\ell+kL}$ have timestamps $\ell + kL$; which is not covered by $[\ell + k'L + 1, \ell + kL - 1]$ – the dependence ranges of $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$. Identical argument shows the independence of $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ and $\{\mathbf{w}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$. Lastly, $\{\mathbf{z}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ and $\{\mathbf{w}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ are independent of each other by definition. Hence, $\{\bar{\mathbf{x}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$, $\{\mathbf{z}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ and $\{\mathbf{w}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ are all independent of each other. This completes the proof. □

A.5.5 Proof of Lemma A.15

Proof. To begin, for all $\ell_k \in \bar{S}_i^{(\ell)}$, we upper bound the difference between the covariance of truncated and non-truncated states as follows,

$$\begin{aligned}
\|\mathbb{E}[\bar{\mathbf{x}}_{\ell_k}\bar{\mathbf{x}}_{\ell_k}^\intercal - \mathbf{x}_{\ell_k}\mathbf{x}_{\ell_k}^\intercal]\| &= \|\mathbb{E}[\bar{\mathbf{x}}_{\ell_k}\bar{\mathbf{x}}_{\ell_k}^\intercal - \mathbf{x}_{\ell_k}\bar{\mathbf{x}}_{\ell_k}^\intercal + \mathbf{x}_{\ell_k}\bar{\mathbf{x}}_{\ell_k}^\intercal - \mathbf{x}_{\ell_k}\mathbf{x}_{\ell_k}^\intercal]\|, \\
&\leq \mathbb{E}[\|\bar{\mathbf{x}}_{\ell_k}\|\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\|] + \mathbb{E}[\|\mathbf{x}_{\ell_k}\|\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\|], \\
&\leq (1/2)c\sigma_{\mathbf{w}}\beta_+\sqrt{n}\mathbb{E}[\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\|] + c\sigma_{\mathbf{w}}\beta_+\sqrt{n}\mathbb{E}[\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\|], \\
&\leq 2c^2\sigma_{\mathbf{w}}^2\beta_+\beta_+'n\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2}, \qquad (A.111)
\end{aligned}$$

where we use Definition A.13 to obtain the second last inequality and (A.107) to obtain the last inequality. Combining this with Lemma A.9, for all $\ell_k \in \bar{S}_i^{(\ell)}$, assuming $c \geq 6$, we have

$$\begin{aligned}\lambda_{\min}(\mathbf{\Sigma}[\bar{\mathbf{x}}_{\ell_k}]) &\geq \lambda_{\min}(\mathbf{\Sigma}[\mathbf{x}_{\ell_k}]) - \|\mathbb{E}[\bar{\mathbf{x}}_{\ell_k}\bar{\mathbf{x}}_{\ell_k}^\intercal - \mathbf{x}_{\ell_k}\mathbf{x}_{\ell_k}^\intercal]\|, \\ &\geq \sigma_{\mathbf{w}}^2/2 - 2c^2\sigma_{\mathbf{w}}^2\beta_+\beta_+' n\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2} \geq \sigma_{\mathbf{w}}^2/4, \end{aligned} \tag{A.112}$$

where we get the last inequality by choosing L via

$$\begin{aligned}\sigma_{\mathbf{w}}^2/4 \geq 2c^2\sigma_{\mathbf{w}}^2\beta_+\beta_+' n\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2} &\iff \rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2} \leq \frac{1}{8c^2\beta_+\beta_+' n\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}}, \\ &\Longleftarrow L \geq 1 + \frac{2\log\left(8c^2\beta_+\beta_+' n\sqrt{ns}\tau_{\boldsymbol{\mathcal{L}}}\right)}{1-\rho_{\boldsymbol{\mathcal{L}}}}. \end{aligned} \tag{A.113}$$

This also implies that we have the following upper bound on the covariance spectral norm, that is, for all $\ell_k \in \bar{S}_i^{(\ell)}$, assuming $c \geq 6$, we have

$$\|\mathbf{\Sigma}[\bar{\mathbf{x}}_{\ell_k}]\| \leq \|\mathbf{\Sigma}[\mathbf{x}_{\ell_k}]\| + \|\mathbb{E}[\bar{\mathbf{x}}_{\ell_k}\bar{\mathbf{x}}_{\ell_k}^\intercal - \mathbf{x}_{\ell_k}\mathbf{x}_{\ell_k}^\intercal]\| \leq c^2\sigma_{\mathbf{w}}^2\beta_+^2 n + \sigma_{\mathbf{w}}^2/4 \leq 2c^2\sigma_{\mathbf{w}}^2\beta_+^2 n. \tag{A.114}$$

Combining (A.112) and (A.114), we get the first statement of the lemma, which is combined with (A.104) to obtain the second statement of the lemma. This completes the proof. □

A.5.6 Proof of Theorem A.16

Proof. To begin, recall that not all the rows of $\bar{\mathbf{H}}_i^{(\ell)}$ are independent. Therefore, to lower bound $\lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)})$, we first consider the matrix $\tilde{\mathbf{H}}_i^{(\ell)}$ which is constructed by (row-wise) stacking $\{\bar{\mathbf{h}}_{\ell_k}^\intercal\}_{\ell_k \in \bar{S}_i^{(\ell)}}$. Observe that, conditioned on the modes, the matrix $\tilde{\mathbf{H}}_i^{(\ell)}$, which is a sub-matrix of $\bar{\mathbf{H}}_i^{(\ell)}$, has independent rows from Lemma A.14.

• **Lower bounding $\underline{\sigma}(\tilde{\mathbf{H}}_i^{(\ell)})$:** Using Lemma A.14, we observe that, conditioned on the modes, the rows of $\tilde{\mathbf{H}}_i^{(\ell)}$ are all independent. Secondly, by definition, each row of $\tilde{\mathbf{H}}_i^{(\ell)}$ can be deterministically bounded as follows: for all $\ell_k \in \bar{S}_i^{(\ell)}$, we have

$$\|\bar{\mathbf{h}}_{\ell_k}\|^2 \leq \frac{1}{\sigma_{\mathbf{w}}^2}\|\bar{\mathbf{x}}_{\ell_k}\|^2 + \frac{1}{\sigma_{\mathbf{z}}^2}\|\mathbf{z}_{\ell_k}\|^2 \leq (1/4)c^2\beta_+^2 n + c^2 p \leq c^2(1+\beta_+^2)(n+p). \tag{A.115}$$

Thirdly, from Lemma A.15, when $c \geq 6$ and $L \geq \max\{t_0, L_{cov}(\rho_{\boldsymbol{\mathcal{L}}})\}$, then for all $\ell_k \in \bar{S}_i^{(\ell)}$, we have

$$(1/4)\mathbf{I}_{n+p} \preceq \mathbf{\Sigma}[\bar{\mathbf{h}}_{\ell_k}] = \mathbb{E}[\bar{\mathbf{h}}_{\ell_k}\bar{\mathbf{h}}_{\ell_k}^\intercal] \preceq 2c^2(1+\beta_+^2)(n+p)\mathbf{I}_{n+p}. \tag{A.116}$$

Therefore, we can use Corollary A.5 to lower bound $\underline{\sigma}(\tilde{\mathbf{H}}_i^{(\ell)})$. Specifically, using Corollary A.5 with $\sigma_{\min} = 1/2$ and $m = c^2(1+\beta_+^2)(n+p)$, with probability at least $1-\delta$, for all $i \in [s]$, we have

$$\underline{\sigma}(\tilde{\mathbf{H}}_i^{(\ell)}) \geq \frac{\sqrt{|\bar{S}_i^{(\ell)}|}}{2} - c\sqrt{(1+\beta_+^2)(n+p)\log\big(\frac{2s(n+p)}{\delta}\big)} \geq \frac{\sqrt{|\bar{S}_i^{(\ell)}|}}{4}, \tag{A.117}$$

as long as $|\bar{S}_i^{(\ell)}|$ satisfies the following lower bound,

$$\begin{aligned}&\frac{\sqrt{|\bar{S}_i^{(\ell)}|}}{4} \geq c\sqrt{(1+\beta_+^2)(n+p)\log\big(\frac{2s(n+p)}{\delta}\big)} \\ \Longleftarrow\ & |\bar{S}_i^{(\ell)}| \geq 16c^2(1+\beta_+^2)(n+p)\log\big(\frac{2s(n+p)}{\delta}\big).\end{aligned} \tag{A.118}$$

• **Lower bounding $\lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)})$:** Using Lemma A.14, we have, $\{\bar{\mathbf{h}}_{\ell_k}\}_{\ell_k \in \bar{S}_i^{(\ell)}}$ is a subset of $\{\bar{\mathbf{h}}_{\ell_k}\}_{\ell_k \in S_i^{(\ell)}}$. As a result, (A.117) also implies that, with probability at least $1-\delta$, for all $i \in [s]$, we have

$$\underline{\sigma}(\bar{\mathbf{H}}_i^{(\ell)}) \geq \underline{\sigma}(\tilde{\mathbf{H}}_i^{(\ell)}) \geq \frac{\sqrt{|\bar{S}_i^{(\ell)}|}}{4} \implies \lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}) \geq \frac{|\bar{S}_i^{(\ell)}|}{16}, \tag{A.119}$$

as long as $|\bar{S}_i^{(\ell)}|$ satisfies the lower bound in (A.118).

• **Upper bounding $\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|$:** Using Lemma A.12, when $L \geq \max\{t_0, L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}}, \delta)\}$, with probability at least $1-\delta$ over the modes, for all $\ell_k \in S_i^{(\ell)}$ and all $i \in [s]$, we have, $\|\bar{\mathbf{h}}_{\ell_k}\|^2 \leq c^2(1+(9/4)\beta_+^2)(n+p)$. This implies that, with probability at least $1-\delta$ over the modes, for all $i \in [s]$, we have

$$\|\bar{\mathbf{H}}_i^{(\ell)}\| \leq \|\bar{\mathbf{H}}_i^{(\ell)}\|_{\mathrm{F}} \leq c(1+2\beta_+)\sqrt{|S_i^{(\ell)}|(n+p)} \leq 2c(1+\beta_+)\sqrt{|S_i^{(\ell)}|(n+p)}.$$

Let $\bar{\mathbf{H}}_i^{(\ell)}$ have singular value decomposition $\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\intercal$ with $\|\boldsymbol{\Sigma}\| \leq 2c(1+\beta_+)\sqrt{|S_i^{(\ell)}|(n+p)}$. Since $\mathbf{W}_i^{(\ell)}$ has i.i.d. $\sigma_{\mathbf{w}}$-sub-Gaussian entries (Assumption A1.1), $\mathbf{U}^\intercal\mathbf{W}_i^{(\ell)} \in \mathbb{R}^{(n+p)\times n}$ has i.i.d. $\sigma_{\mathbf{w}}$-sub-Gaussian columns. As a result, applying Theorem 5.39 of (Vershynin, 2012), with probability at least $1-\delta$, for all $i \in [s]$, we have

$$\|\mathbf{U}^\intercal\mathbf{W}_i^{(\ell)}\| \leq C\sigma_{\mathbf{w}}\sqrt{n+p} + C_0\sqrt{\log\big(\frac{2s}{\delta}\big)}. \tag{A.120}$$

This implies that, with probability at least $1-2\delta$, for all $i \in [s]$, we have

$$\|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\| \leq \|\mathbf{\Sigma}\|\|\mathbf{U}^\intercal\mathbf{W}_i^{(\ell)}\| \leq 2c(1+\beta_+)\sqrt{|S_i^{(\ell)}|(n+p)}\bigg(C\sigma_{\mathbf{w}}\sqrt{n+p}+C_0\sqrt{\log\big(\frac{2s}{\delta}\big)}\bigg).$$

This completes the proof. □

A.5.7 Proof of Theorem A.17

Proof. We begin by simplifying the the term $\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\|$ as follows,

$$\begin{aligned}
\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\| &= \|\sum_{\ell_k \in S_i^{(\ell)}}\big(\mathbf{h}_{\ell_k}\mathbf{h}_{\ell_k}^\intercal - \bar{\mathbf{h}}_{\ell_k}\bar{\mathbf{h}}_{\ell_k}^\intercal\big)\|, \\
&\leq |S_i^{(\ell)}| \max_{\ell_k \in S_i^{(\ell)}} \|\mathbf{h}_{\ell_k}\mathbf{h}_{\ell_k}^\intercal - \bar{\mathbf{h}}_{\ell_k}\bar{\mathbf{h}}_{\ell_k}^\intercal\|, \\
&= |S_i^{(\ell)}| \max_{\ell_k \in S_i^{(\ell)}} \|\mathbf{h}_{\ell_k}\mathbf{h}_{\ell_k}^\intercal - \mathbf{h}_{\ell_k}\bar{\mathbf{h}}_{\ell_k}^\intercal + \mathbf{h}_{\ell_k}\bar{\mathbf{h}}_{\ell_k}^\intercal - \bar{\mathbf{h}}_{\ell_k}\bar{\mathbf{h}}_{\ell_k}^\intercal\|, \\
&\leq |S_i^{(\ell)}| \max_{\ell_k \in S_i^{(\ell)}} \big(\|\mathbf{h}_{\ell_k}\|\|\mathbf{h}_{\ell_k} - \bar{\mathbf{h}}_{\ell_k}\| + \|\bar{\mathbf{h}}_{\ell_k}\|\|\mathbf{h}_{\ell_k} - \bar{\mathbf{h}}_{\ell_k}\|\big). \qquad \text{(A.121)}
\end{aligned}$$

We will upper bound each of these terms separately and combine them together in (A.121) to get the desired upper bound. First of all, observe that, for all $i \in [s]$, each row of $\mathbf{H}_i^{(\ell)}$ is deterministically bounded as follows,

$$\|\mathbf{h}_{\ell_k}\|^2 \leq \frac{1}{\sigma_{\mathbf{w}}^2}\|\mathbf{x}_{\ell_k}\|^2 + \frac{1}{\sigma_{\mathbf{z}}^2}\|\mathbf{z}_{\ell_k}\|^2 \leq c^2\beta_+^2 n + c^2 p \leq c^2(1+\beta_+^2)(n+p). \qquad \text{(A.122)}$$

Similarly, from Lemma A.12, when $L \geq \max\{t_0, L_{tr1}(\rho_{\boldsymbol{\mathcal{L}}}, \delta)$, with probability at least $1-\delta$ over the modes, for all $i \in [s]$, each row of $\bar{\mathbf{H}}_i^{(\ell)}$ can be bounded as follows,

$$\|\bar{\mathbf{h}}_{\ell_k}\|^2 \leq \frac{1}{\sigma_{\mathbf{w}}^2}\|\bar{\mathbf{x}}_{\ell_k}\|^2 + \frac{1}{\sigma_{\mathbf{z}}^2}\|\mathbf{z}_{\ell_k}\|^2 \leq c^2(9/4)\beta_+^2 n + c^2 p \leq 4c^2(1+\beta_+^2)(n+p). \qquad \text{(A.123)}$$

To proceed, recall from (A.108) that, with probability at least $1-\delta$ over the modes, for all $\ell_k \in S_i^{(\ell)}$ and all $i \in [s]$, we have

$$\|\mathbf{h}_{\ell_k} - \bar{\mathbf{h}}_{\ell_k}\| = \left\| \begin{bmatrix} \frac{1}{\sigma_{\mathbf{w}}}\mathbf{x}_{\ell_k} \\ \frac{1}{\sigma_{\mathbf{z}}}\mathbf{z}_{\ell_k} \end{bmatrix} - \begin{bmatrix} \frac{1}{\sigma_{\mathbf{w}}}\bar{\mathbf{x}}_{\ell_k} \\ \frac{1}{\sigma_{\mathbf{z}}}\mathbf{z}_{\ell_k} \end{bmatrix} \right\| = \frac{1}{\sigma_{\mathbf{w}}}\|\mathbf{x}_{\ell_k} - \bar{\mathbf{x}}_{\ell_k}\| \leq \frac{c\beta_+' n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2}T}{\delta}. \qquad \text{(A.124)}$$

Combining (A.122), (A.123) and (A.124) into (A.121), with probability at least $1-\delta$ over the modes, for all $i \in [s]$, we have

$$\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\| \leq \frac{3c^2\beta_+'(1+\beta_+)\tau_{\mathcal{L}}\rho_{\mathcal{L}}^{(L-1)/2}n\sqrt{s(n+p)}|S_i^{(\ell)}|T}{\delta}. \tag{A.125}$$

Using a similar line of reasoning, with probability at least $1-\delta$ over the modes, for all $i \in [s]$, we also have

$$\begin{aligned}\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\| &= \|\sum_{\ell_k \in S_i^{(\ell)}} \big(\mathbf{h}_{\ell_k}\mathbf{w}_{\ell_k}^\intercal - \bar{\mathbf{h}}_{\ell_k}\mathbf{w}_{\ell_k}^\intercal\|\big), \\ &\leq |S_i^{(\ell)}| \max_{\ell_k \in S_i^{(\ell)}} \|\mathbf{h}_{\ell_k} - \bar{\mathbf{h}}_{\ell_k}\|\|\mathbf{w}_{(j_k)}^\intercal\|, \\ &\leq \frac{cc_{\mathbf{w}}\sigma_{\mathbf{w}}\beta_+'\tau_{\mathcal{L}}\rho_{\mathcal{L}}^{(L-1)/2}n\sqrt{ns}|S_i^{(\ell)}|T}{\delta}. \end{aligned} \tag{A.126}$$

This completes the proof. □

A.5.8 Proof of Theorem A.18

Proof. To begin, using Theorem A.16 along-with the assumption made in the statement of the theorem regarding $|\bar{S}_i^{(\ell)}|$, with probability at least $1-4\delta$, for all $i \in [s]$, we have

$$\lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}) \geq \frac{|\bar{S}_i^{(\ell)}|}{16} \geq \frac{\pi_{\min}T}{32L}, \tag{A.127}$$

as long as the trajectory length T satisfies the following lower bound,

$$T \geq \frac{32L}{\pi_{\min}}c^2(1+\beta_+^2)\log\big(\frac{2s(n+p)}{\delta}\big)(n+p). \tag{A.128}$$

Combining this with Theorem A.17, with probability at least $1-5\delta$, for all $i \in [s]$, we have

$$\begin{aligned}\lambda_{\min}(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}) &\geq \lambda_{\min}(\bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}) - \|\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\bar{\mathbf{H}}_i^{(\ell)}\|, \\ &\geq \frac{\pi_{\min}T}{32L} - \frac{3c^2\beta_+'(1+\beta_+)\tau_{\mathcal{L}}\rho_{\mathcal{L}}^{(L-1)/2}n\sqrt{s(n+p)}T^2}{\delta L}, \\ &\geq \frac{\pi_{\min}T}{64L}, \end{aligned} \tag{A.129}$$

where we get the last inequality by choosing L via,

$$\begin{aligned}
&\frac{\pi_{\min} T}{64L} \geq \frac{3c^2\beta'_+(1+\beta_+)\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2} n\sqrt{s(n+p)}T^2}{\delta L}, \\
\iff\ & \rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2} \leq \frac{\pi_{\min}\delta}{192c^2\tau_{\boldsymbol{\mathcal{L}}}\beta'_+(1+\beta_+)n\sqrt{s(n+p)}T}, \\
\Longleftarrow\ & \qquad L \geq 1 + \frac{2\log\big(192c^2\tau_{\boldsymbol{\mathcal{L}}}\beta'_+(1+\beta_+)n\sqrt{s(n+p)}T/(\pi_{\min}\delta)\big)}{(1-\rho_{\boldsymbol{\mathcal{L}}})}.
\end{aligned} \tag{A.130}$$

Similarly, combing Theorems A.16 and A.17, with probability at least $1-4\delta$, for all $i \in [s]$, we also have

$$\begin{aligned}
\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\| &\leq \|\bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\| + \|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)} - \bar{\mathbf{H}}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|, \\
&\leq 2c(1+\beta_+)\sqrt{\frac{T(n+p)}{L}}\Big(C\sigma_{\mathbf{w}}\sqrt{n+p} + C_0\sqrt{\log\big(\frac{2s}{\delta}\big)}\Big) \\
&\quad + \frac{cc_{\mathbf{w}}\sigma_{\mathbf{w}}\beta'_+\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2} n\sqrt{ns}T^2}{\delta L}, \\
&\leq 3c(1+\beta_+)\sqrt{\frac{T(n+p)}{L}}\Big(C\sigma_{\mathbf{w}}\sqrt{n+p} + C_0\sqrt{\log\big(\frac{2s}{\delta}\big)}\Big),
\end{aligned} \tag{A.131}$$

where we get the last inequality by choosing L via,

$$\begin{aligned}
&\frac{cc_{\mathbf{w}}\sigma_{\mathbf{w}}\beta'_+\tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2} n\sqrt{ns}T^2}{\delta L} \leq c(1+\beta_+)\sqrt{\frac{T(n+p)}{L}}\Big(C\sigma_{\mathbf{w}}\sqrt{n+p} + C_0\sqrt{\log\big(\frac{2s}{\delta}\big)}\Big) \\
\iff\ & \rho_{\boldsymbol{\mathcal{L}}}^{(L-1)/2} \leq \frac{\delta(1+\beta_+)\sqrt{TL(n+p)}\big(C\sigma_{\mathbf{w}}\sqrt{n+p} + C_0\sqrt{\log(2s/\delta)}\big)}{c_{\mathbf{w}}\sigma_{\mathbf{w}}\beta'_+\tau_{\boldsymbol{\mathcal{L}}}n\sqrt{ns}T^2}, \\
\Longleftarrow\ & \qquad L \geq 1 + \frac{2}{(1-\rho_{\boldsymbol{\mathcal{L}}})}\log\Big(\frac{c_{\mathbf{w}}\sigma_{\mathbf{w}}\beta'_+\tau_{\boldsymbol{\mathcal{L}}}n\sqrt{ns}TT}{\delta(1+\beta_+)\sqrt{(n+p)}\big(C\sigma_{\mathbf{w}}\sqrt{n+p} + C_0\sqrt{\log(2s/\delta)}\big)}\Big).
\end{aligned} \tag{A.132}$$

Finally combining (A.129) and (A.131) into (A.24) and union bounding over all $0 \leq \ell \leq L-1$, with probability at least $1-9L\delta$, for all $i \in [s]$, we have

$$\begin{aligned}
\|\hat{\mathbf{\Theta}}_i - \mathbf{\Theta}_i^\star\| &\leq \frac{\sum_{\ell=0}^{L-1}\|\mathbf{H}_i^{(\ell)\intercal}\mathbf{W}_i^{(\ell)}\|}{\sum_{\ell=0}^{L-1}\lambda_{\min}\big(\mathbf{H}_i^{(\ell)\intercal}\mathbf{H}_i^{(\ell)}\big)}, \\
&\leq \frac{192c(1+\beta_+)}{\pi_{\min}}\sqrt{\frac{L(n+p)}{T}}\Big(C\sigma_{\mathbf{w}}\sqrt{n+p} + C_0\sqrt{\log\big(\frac{2s}{\delta}\big)}\Big).
\end{aligned} \tag{A.133}$$

To proceed, using standard result from linear algebra that the spectral norm of a sub-matrix is upper bounded by the norm of the original matrix, with probability at least $1 - 9L\delta$, for all $i \in [s]$, we have

$$\begin{aligned}
\|\hat{\mathbf{L}}_i - \mathbf{L}_i\| &\leq \frac{192c(1+\beta_+)}{\pi_{\min}} \sqrt{\frac{L(n+p)}{T}} \Big(C\sqrt{n+p} + (C_0/\sigma_{\mathbf{w}}) \sqrt{\log\big(\frac{2s}{\delta}\big)} \Big), \\
\|\hat{\mathbf{B}}_i - \mathbf{B}_i\| &\leq \frac{192c(1+\beta_+)}{\pi_{\min}} \frac{\sigma_{\mathbf{w}}}{\sigma_{\mathbf{z}}} \sqrt{\frac{L(n+p)}{T}} \Big(C\sqrt{n+p} + (C_0/\sigma_{\mathbf{w}}) \sqrt{\log\big(\frac{2s}{\delta}\big)} \Big), \\
\implies \|\hat{\mathbf{A}}_i - \mathbf{A}_i\| &\leq \|\hat{\mathbf{L}}_i - \mathbf{L}_i\| + \|\mathbf{K}_i\| \|\hat{\mathbf{B}}_i - \mathbf{B}_i\|, \\
&\leq \frac{192c(1+\beta_+)}{\pi_{\min}} \frac{(\sigma_{\mathbf{z}} + C_K \sigma_{\mathbf{w}})}{\sigma_{\mathbf{z}}} \sqrt{\frac{L(n+p)}{T}} \Big(C\sqrt{n+p} + (C_0/\sigma_{\mathbf{w}}) \sqrt{\log\big(\frac{2s}{\delta}\big)} \Big). \quad \text{(A.134)}
\end{aligned}$$

Finally replacing δ with $\delta/(18L)$ we get the statement of the theorem. This completes the proof. □

A.5.9 Proof of Lemma A.19

Proof. Let E be the event $\{\mathbf{g} \mid \|\mathbf{g}\|_\infty \leq c_{\mathbf{w}}\sigma_{\mathbf{w}}\}$. If $c_{\mathbf{w}} \geq \sqrt{2\log(nT)} + \sqrt{2\log(2/\delta)}$, using Gaussian tail bound and the fact that $\mathbb{E}[\|\mathbf{g}\|_\infty] \leq \sigma_{\mathbf{w}}\sqrt{2\log(nT)}$, observe that $\mathbb{P}(E) \geq 1 - e^{-\frac{(c_{\mathbf{w}}\sigma_{\mathbf{w}} - \mathbb{E}[\|\mathbf{g}\|_\infty])^2}{2}} \geq 1 - \delta/2$. Therefore, we have

$$\mathbb{E}[S_{\mathbf{g},\mathbf{a}}] \geq \mathbb{E}[S_{\mathbf{g},\mathbf{a}}|E]\mathbb{P}(E) = \mathbb{E}[S_{\mathbf{g}',\mathbf{a}}]\mathbb{P}(E) \geq (1-\delta/2)^2 \geq 1 - \delta.$$

This completes the proof. □

Appendix B
Proofs for Results in Chapter 4

B.1 Useful Facts

Fact B.1 (Matrix Facts)**.** *For arbitrary matrices* $\mathbf{M}, \mathbf{N}, \mathbf{X}$ *with appropriate dimensions, we have the following facts.*

1. *If* $\mathbf{M}, \mathbf{N} \succeq 0$*, then*

$$\|\mathbf{N}(\mathbf{I} + \mathbf{M}\mathbf{N})^{-1}\| \leq \|\mathbf{N}\|, \tag{B.1}$$

$$\|(\mathbf{I} + \mathbf{M}\mathbf{N})^{-1}\| \leq 1 + \|\mathbf{N}\|\|\mathbf{M}\|. \tag{B.2}$$

2. *If* $\mathbf{M}$ *and* $\mathbf{M} + \mathbf{N}$ *are invertible, then*

$$\begin{aligned}(\mathbf{M} + \mathbf{N})^{-1} &= \mathbf{M}^{-1} - \mathbf{M}^{-1}\mathbf{N}(\mathbf{M} + \mathbf{N})^{-1} \\ &= \mathbf{M}^{-1} - (\mathbf{M} + \mathbf{N})^{-1}\mathbf{N}\mathbf{M}^{-1}.\end{aligned} \tag{B.3}$$

3. *If* $\mathbf{I} + \mathbf{M}$ *and* $\mathbf{I} + \mathbf{N}$ *are invertible, then*

$$(\mathbf{I} + \mathbf{M})^{-1} - (\mathbf{I} + \mathbf{N})^{-1} = (\mathbf{I} + \mathbf{M})^{-1}(\mathbf{N} - \mathbf{M})(\mathbf{I} + \mathbf{N})^{-1}. \tag{B.4}$$

4.

$$\mathbf{vec}(\mathbf{M}\mathbf{X}\mathbf{N}) = (\mathbf{N}^{\intercal} \otimes \mathbf{M})\mathbf{vec}(\mathbf{X}). \tag{B.5}$$

5. *For a collection of matrices* $\mathbf{M}_{1:s}$*, and for all* $i \in [s]$*,*

$$\|\varphi_i(\mathbf{M}_{1:s})\| = \|\sum\nolimits_{j=1}^{s} \mathbf{T}(i,j)\mathbf{M}_j\| \leq \|\mathbf{M}_{1:s}\|. \tag{B.6}$$

In Fact B.1, (B.1) is due to (Mania et al., 2019, Lemma 7) (in their supplement); to see (B.2), first note that $(\mathbf{I} + \mathbf{M}\mathbf{N})^{-1} = \mathbf{I} - \mathbf{M}\mathbf{N}(\mathbf{I} + \mathbf{M}\mathbf{N})^{-1}$ by matrix inversion lemma, and then apply (B.1). (B.3) and (B.4) also follow from matrix inversion lemma.

Fact B.2. *For nxn matrices* $\mathbf{X}_{1:s}$*, let* $\mathbf{X} = \mathbf{diag}(\mathbf{X}_{1:s})$*. Let* $\tilde{\mathbf{vec}}(\cdot)$ *be the operator that vectorizes all diagonal blocks of* $\mathbf{X}$ *into a vector, i.e.* $\tilde{\mathbf{vec}}(\mathbf{X}) := (\mathbf{vec}(\mathbf{X}_1), \dots, \mathbf{vec}(\mathbf{X}_s))$*. Let* $\tilde{\mathbf{vec}}^{-1}$ *denote the inverse, i.e.* $\tilde{\mathbf{vec}}^{-1}(\tilde{\mathbf{vec}}(\mathbf{X})) = \mathbf{X}$*. Then,*

$$\|\tilde{\mathbf{vec}}\| := \sup_{\mathbf{X}=\mathbf{diag}(\mathbf{X}_{1:s}), \|\mathbf{X}\|=1} \|\tilde{\mathbf{vec}}(\mathbf{X})\| \overset{\text{(i)}}{=} \sqrt{ns} \tag{B.7}$$

$$\|\tilde{\mathbf{vec}}^{-1}\| := \sup_{\|\mathbf{x}\|=1} \|\tilde{\mathbf{vec}}^{-1}(\mathbf{x})\| \overset{\text{(ii)}}{=} 1. \tag{B.8}$$

Fact B.2 follows by noting that (i) achieves the supremum when $\mathbf{X}_i = \mathbf{I}_n$ for all i and (ii) achieves the supremum when $\mathbf{x} = (1, 0, \dots, 0)$. For a matrix $\mathbf{M}$ and perturbation $\mathbf{\Delta}$, we have the following result adapted from (Mania et al., 2019, Lemma 5).

Fact B.3. *Let* $\rho := \rho(\mathbf{M})$ *and* $\tau := \sup_{k \in \mathbb{N}} \|\mathbf{M}^k\|/\rho^k$*. Then,* (a) $\rho(\mathbf{M} + \mathbf{\Delta}) \leq \tau\|\mathbf{\Delta}\| + \rho$*;* (b) $\|(\mathbf{M} + \mathbf{\Delta})^k\| \leq \tau\big(\tau\|\mathbf{\Delta}\| + \rho\big)^k$*.*

Fact B.4. *Consider a generic MJS-LQR problem given by* MJS-LQR$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ *and its corresponding* cDARE$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$*. Assume* $\mathbf{Q}_{1:s}, \mathbf{R}_{1:s} \succ 0$ *for all* $i \in [s]$*. If there exists a positive definite solution* $\mathbf{P}_{1:s} \succ 0$ *to* cDARE$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$*, then it is the unique solution among* $\{\mathbf{P}_{1:s} : \mathbf{P}_i \succeq 0, \forall\, i \in [s]\}$*.*

To see this, first note that cDARE$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ can be written as the Joseph stabilized form (Lewis et al., 2012, (2.2-62)), i.e. $\mathbf{P}_i - \mathbf{L}_i^\intercal \varphi_i(\mathbf{P}_{1:s})\mathbf{L}_i = \mathbf{K}_i^\intercal \mathbf{R}_i \mathbf{K}_i + \mathbf{Q}_i$ where $\mathbf{K}_i = -\big(\mathbf{R}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}_{1:s})\mathbf{B}_i\big)^{-1} \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}_{1:s}^\star)\mathbf{A}_i$ and $\mathbf{L}_i := \mathbf{A}_i + \mathbf{B}_i \mathbf{K}_i$. Since $\mathbf{Q}_i \succ 0$, we know by Corollary 2.13 the closed-loop MJS $\mathbf{x}_{t+1} = \mathbf{L}_{\omega_t}\mathbf{x}_t$ is mean-square stable. Then one can obtain Fact B.4 by invoking (Costa et al., 2006, Lemma A.14) which says cDARE has at most one solution with resulting controller stabilizes the MJS.

B.2 Proof of Theorem 4.1

We first provide the road map of the proof.

(a) We construct an operator $\mathcal{K}(\mathbf{X}_{1:s})$ using the difference between the ground truth cDARE$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ and perturbed cDARE$(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$, whose fixed point $\mathbf{X}_{1:s}^\star$ (if exists) guarantees $\hat{\mathbf{P}}_{1:s} := \mathbf{P}_{1:s}^\star + \mathbf{X}_{1:s}^\star$ to be a solution to the perturbed cDARE$(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$.

(b) We show when ϵ, η are small, $\mathcal{K}(\mathbf{X}_{1:s})$ is a contraction mapping on a closed set $\mathcal{S}_\nu$ whose radius ν is a function of ϵ and η. Thus, there exists a unique fixed point $\mathbf{X}^\star_{1:s} \in \mathcal{S}_\nu$ and $\|\hat{\mathbf{P}}_{1:s} - \mathbf{P}^\star_{1:s}\| = \|\mathbf{X}^\star_{1:s}\| \leq \nu(\epsilon, \eta)$.

(c) Finally, we show $\hat{\mathbf{P}}_{1:s}$ is unique by showing $\hat{\mathbf{P}}_i \succ 0$ and then invoking Fact B.4.

B.2.1 Construct operator $\mathcal{K}$

First we define a few notations for the ease of exposition. For all $i \in [s]$, let $\mathbf{S}_i := \mathbf{B}_i \mathbf{R}_i^{-1} \mathbf{B}_i^\intercal$ and $\hat{\mathbf{S}}_i := \hat{\mathbf{B}}_i \mathbf{R}_i^{-1} \hat{\mathbf{B}}_i^\intercal$. Define block diagonal matrices $\mathbf{A}$, $\hat{\mathbf{A}}$, $\mathbf{B}$, $\hat{\mathbf{B}}$, $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{P}^\star$, $\hat{\mathbf{P}}$, $\mathbf{K}^\star$, $\mathbf{L}^\star$, $\mathbf{S}$, $\hat{\mathbf{S}}$, $\mathbf{P}$, $\mathbf{X}$, $\mathbf{\Phi}(\mathbf{X})$, $\hat{\mathbf{\Phi}}(\mathbf{X})$ such that their i-th diagonal blocks are given by $\mathbf{A}_i$, $\hat{\mathbf{A}}_i$, $\mathbf{B}_i$, $\hat{\mathbf{B}}_i$, $\mathbf{Q}_i$, $\mathbf{R}_i$, $\mathbf{P}^\star_i$, $\hat{\mathbf{P}}_i$, $\mathbf{K}^\star_i$, $\mathbf{L}^\star_i$, $\mathbf{S}_i$, $\hat{\mathbf{S}}_i$, $\mathbf{P}_i$, $\mathbf{X}_i$, $\varphi_i(\mathbf{X}_{1:s})$, $\hat{\varphi}_i(\mathbf{X}_{1:s})$ respectively. Note that $\mathbf{P}_i, \mathbf{X}_i \succeq 0$ are just generic variables to be used in function arguments. We will see many equations that hold for each single block also hold for the diagonally concatenated notations.

We have $\mathbf{K}^\star = -(\mathbf{R} + \mathbf{B}^\intercal \mathbf{\Phi}(\mathbf{P}^\star)\mathbf{B})^{-1}\mathbf{B}^\intercal \mathbf{\Phi}(\mathbf{P}^\star)\mathbf{A}$ from (2.36), then using the matrix inversion lemma, we get

$$\mathbf{L}^\star = \mathbf{A} + \mathbf{B}\mathbf{K}^\star = (\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star))^{-1}\mathbf{A}. \tag{B.9}$$

Furthermore, by diagonally concatenating cDARE (2.35) and then applying the matrix inversion lemma again, we have

$$\mathbf{X} = \mathbf{A}^\intercal \mathbf{\Phi}(\mathbf{X})(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{X}))^{-1}\mathbf{A} + \mathbf{Q}. \tag{B.10}$$

Then, we define the following Riccati difference function using the difference between LHS and RHS of (B.10), with $\mathbf{P}$ as argument and $\mathbf{A}, \mathbf{B}, \mathbf{T}$ as parameters:

$$\mathcal{F}(\mathbf{P}; \mathbf{A}, \mathbf{B}, \mathbf{T}) := \mathbf{P} - \mathbf{A}^\intercal \mathbf{\Phi}(\mathbf{P})(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}))^{-1}\mathbf{A} - \mathbf{Q}. \tag{B.11}$$

Though not explicitly listed, $\mathbf{\Phi}$ and $\mathbf{S}$ on the RHS of (B.11) depend on $\mathbf{T}$ and $\mathbf{B}$ respectively. Since $\mathbf{P}^\star_{1:s}$ is the solution to cDARE$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$, we have $\mathcal{F}(\mathbf{P}^\star; \mathbf{A}, \mathbf{B}, \mathbf{T}) = 0$. Similarly, if there exists solution $\hat{\mathbf{P}}_{1:s}$ to cDARE$(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$, then we also have $\mathcal{F}(\hat{\mathbf{P}}; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{T}}) = 0$.

For $\mathbf{X}$ such that $\mathbf{P}^\star + \mathbf{X} \succeq 0$, we know $\mathbf{I} + \mathbf{S}(\mathbf{P}^\star + \mathbf{X})$ is invertible. Then, for $\mathcal{F}(\mathbf{P}^\star +$

$\mathbf{X}; \mathbf{A}, \mathbf{B}, \mathbf{T})$, we have

$$\begin{aligned}
&\mathcal{F}(\mathbf{P}^\star + \mathbf{X}; \mathbf{A}, \mathbf{B}, \mathbf{T}) \\
\overset{(B.3)}{=}&\mathbf{P}^\star + \mathbf{X} - \mathbf{A}^\intercal \mathbf{\Phi}(\mathbf{P}^\star + \mathbf{X}) \cdot \big[(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star))^{-1} - \\
&\underbrace{(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star + \mathbf{X}))^{-1}}_{=:\mathbf{\Gamma}} \mathbf{S}\mathbf{\Phi}(\mathbf{X})(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star))^{-1}\big] \mathbf{A} - \mathbf{Q} \\
=&\mathbf{P}^\star + \mathbf{X} - \mathbf{A}^\intercal \mathbf{\Phi}(\mathbf{P}^\star + \mathbf{X})(\mathbf{I} - \mathbf{\Gamma}\mathbf{S}\mathbf{\Phi}(\mathbf{X}))(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star))^{-1}\mathbf{A} - \mathbf{Q} \\
\overset{(B.9)}{=}&\mathbf{P}^\star + \mathbf{X} - \mathbf{A}^\intercal \mathbf{\Phi}(\mathbf{P}^\star + \mathbf{X})(\mathbf{I} - \mathbf{\Gamma}\mathbf{S}\mathbf{\Phi}(\mathbf{X}))\mathbf{L}^\star - \mathbf{Q} \\
\overset{(i)}{=}&\mathbf{X} - \mathbf{A}^\intercal \mathbf{\Phi}(\mathbf{P}^\star + \mathbf{X})(\mathbf{I} - \mathbf{\Gamma}\mathbf{S}\mathbf{\Phi}(\mathbf{X}))\mathbf{L}^\star + \mathbf{A}^\intercal\mathbf{\Phi}(\mathbf{P}^\star)\mathbf{L}^\star \\
=&\mathbf{X} - \mathbf{A}^\intercal \left[\mathbf{\Phi}(\mathbf{P}^\star + \mathbf{X})(\mathbf{I} - \mathbf{\Gamma}\mathbf{S}\mathbf{\Phi}(\mathbf{X})) - \mathbf{\Phi}(\mathbf{P}^\star)\right] \mathbf{L}^\star \\
\overset{(B.9)}{=}&\mathbf{X} - \mathbf{L}^{\star\intercal}(\mathbf{I} + \mathbf{\Phi}(\mathbf{P}^\star)\mathbf{S}) \left[\mathbf{\Phi}(\mathbf{P}^\star + \mathbf{X})(\mathbf{I} - \mathbf{\Gamma}\mathbf{S}\mathbf{\Phi}(\mathbf{X})) - \mathbf{\Phi}(\mathbf{P}^\star)\right] \mathbf{L}^\star \\
=&\mathbf{X} - \mathbf{L}^{\star\intercal} \underbrace{(\mathbf{I} + \mathbf{\Phi}(\mathbf{P}^\star)\mathbf{S}) \left[-\mathbf{\Phi}(\mathbf{P}^\star)\mathbf{\Gamma}\mathbf{S} + \mathbf{I} - \mathbf{\Phi}(\mathbf{X})\mathbf{\Gamma}\mathbf{S}\right]}_{=:\mathbf{\Lambda}} \mathbf{\Phi}(\mathbf{X})\mathbf{L}^\star
\end{aligned}$$

where (i) follows from $\mathbf{P}^\star - \mathbf{Q} = \mathbf{A}^\intercal\mathbf{\Phi}(\mathbf{X})\mathbf{L}^\star$ which can be seen from the fact $\mathcal{F}(\mathbf{P}^\star; \mathbf{A}, \mathbf{B}, \mathbf{T}) = 0$. By expanding $\mathbf{\Lambda}$, one can check $\mathbf{\Lambda} = \mathbf{I} - \mathbf{\Phi}(\mathbf{X})\mathbf{\Gamma}\mathbf{S}$. Plugging this back and using the definition of $\mathbf{\Gamma}$, we have

$$\begin{aligned}
\mathcal{F}(\mathbf{P}^\star + \mathbf{X}; \mathbf{A}, \mathbf{B}, \mathbf{T}) = \mathbf{X} - \mathbf{L}^{\star\intercal}\mathbf{\Phi}(\mathbf{X})\mathbf{L}^\star + \\
\mathbf{L}^{\star\intercal}\mathbf{\Phi}(\mathbf{X})(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star) + \mathbf{S}\mathbf{\Phi}(\mathbf{X}))^{-1}\mathbf{S}\mathbf{\Phi}(\mathbf{X})\mathbf{L}^\star. \quad \text{(B.12)}
\end{aligned}$$

If we define

$$\begin{aligned}
\mathcal{T}(\mathbf{X}) &= \mathbf{X} - \mathbf{L}^{\star\intercal}\mathbf{\Phi}(\mathbf{X})\mathbf{L}^\star, \\
\mathcal{H}(\mathbf{X}) &= \mathbf{L}^{\star\intercal}\mathbf{\Phi}(\mathbf{X})(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star) + \mathbf{S}\mathbf{\Phi}(\mathbf{X}))^{-1}\mathbf{S}\mathbf{\Phi}(\mathbf{X})\mathbf{L}^\star,
\end{aligned} \quad \text{(B.13)}$$

we can write $\mathcal{F}(\mathbf{P}^\star + \mathbf{X}; \mathbf{A}, \mathbf{B}, \mathbf{T})$ as

$$\mathcal{F}(\mathbf{P}^\star + \mathbf{X}; \mathbf{A}, \mathbf{B}, \mathbf{T}) = \mathcal{T}(\mathbf{X}) + \mathcal{H}(\mathbf{X}). \quad \text{(B.14)}$$

We now study the invertibility of operator $\mathcal{T}$. Let $\mathbf{Y}_i := \mathbf{X}_i - \mathbf{L}_i^{\star\intercal}\varphi_i(\mathbf{X}_{1:s})\mathbf{L}_i^\star$, and $\mathbf{Y} := \mathbf{diag}(\mathbf{Y}_{1:s})$, then we see $\mathbf{Y} = \mathcal{T}(\mathbf{X}) = \mathbf{X} - \mathbf{L}^{\star\intercal}\mathbf{\Phi}(\mathbf{X})\mathbf{L}^\star$. Apply (B.5) to $\mathbf{Y}_i$, we have $\mathbf{vec}(\mathbf{Y}_i) = (\mathbf{I} - \mathbf{T}(i,i) \cdot \mathbf{L}_i^{\star\intercal} \otimes \mathbf{L}_i^{\star\intercal})\mathbf{vec}(\mathbf{X}_i) - \sum_{j \neq i} \mathbf{T}(i,j)\mathbf{L}_i^{\star\intercal} \otimes \mathbf{L}_i^{\star\intercal}\mathbf{vec}(\mathbf{X}_j)$. Stacking this equation for all i, we have $(\mathbf{I} - \boldsymbol{\mathcal{L}}^{\star\intercal})\tilde{\mathbf{vec}}(\mathbf{X}) = \tilde{\mathbf{vec}}(\mathbf{Y})$, where $\tilde{\mathbf{vec}}(\cdot)$ is defined in Fact B.2. We know $\rho(\boldsymbol{\mathcal{L}}^{\star\intercal}) < 1$, thus $(\mathbf{I} - \boldsymbol{\mathcal{L}}^{\star\intercal})$ is invertible, and inverse operator $\mathcal{T}^{-1}$ exists and is given by

$$\mathbf{X} = \mathcal{T}^{-1}(\mathbf{Y}) = \tilde{\mathbf{vec}}^{-1} \circ (\mathbf{I} - \boldsymbol{\mathcal{L}}^{\star\intercal})^{-1} \circ \tilde{\mathbf{vec}}(\mathbf{Y}), \quad \text{(B.15)}$$

where $\circ$ denotes operator composition, and $\tilde{\mathbf{vec}}(\cdot)^{-1}$ is defined in Fact B.2. With $\mathcal{T}^{-1}$, we define the following operator:

$$\mathcal{K}(\mathbf{X}) := \mathcal{T}^{-1}\big(\mathcal{F}(\mathbf{P}^\star + \mathbf{X}; \mathbf{A}, \mathbf{B}, \mathbf{T}) - \mathcal{F}(\mathbf{P}^\star + \mathbf{X}; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{T}}) - \mathcal{H}(\mathbf{X})\big). \tag{B.16}$$

Suppose there exists a fixed point $\mathbf{X}^\star$ for $\mathcal{K}$, then we see $\mathcal{F}(\mathbf{P}^\star + \mathbf{X}^\star; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{T}}) = \mathcal{F}(\mathbf{P}^\star + \mathbf{X}^\star; \mathbf{A}, \mathbf{B}, \mathbf{T}) - \mathcal{T}(\mathbf{X}^\star) - \mathcal{H}(\mathbf{X}^\star) = 0$, i.e. $\hat{\mathbf{P}}_{1:s} = \mathbf{P}^\star_{1:s} + \mathbf{X}^\star_{1:s}$ is a solution to the perturbed cDARE$(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$.

B.2.2 $\mathcal{K}$ is a Contraction

We will show $\mathcal{K}(\mathbf{X})$ is a contraction mapping on the closed set

$$\mathcal{S}_\nu := \{\mathbf{X} : \|\mathbf{X}\| \le \nu, \mathbf{X} = \mathbf{diag}(\mathbf{X}_{1:s}), \mathbf{P}^\star + \mathbf{X} \succeq 0\} \tag{B.17}$$

so that $\mathcal{K}(\mathbf{X})$ is guaranteed to have a fixed point in $\mathcal{S}_\nu$. To do this, we first present the following lemma (proof in Appendix B.4) regarding properties of $\mathcal{K}(\mathbf{X})$.

Lemma B.5. *Assume* $\epsilon \le \min\{\|\mathbf{B}\|, 1\}$. *Suppose* $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2 \in \mathcal{S}_\nu$ *with* $\nu \le \min\{1, \|\mathbf{S}\|^{-1}\}$, *then*

$$\|\mathcal{K}(\mathbf{X})\| \le \frac{\sqrt{ns}\tau}{1-\rho}\big(\|\mathbf{L}^\star\|^2\|\mathbf{S}\|\nu^2 + \frac{C_\epsilon\epsilon + C_\eta\eta}{2}\big), \tag{B.18}$$

$$\|\mathcal{K}(\mathbf{X}_1) - \mathcal{K}(\mathbf{X}_2)\| \le \frac{\sqrt{ns}\tau}{1-\rho}\|\mathbf{X}_1 - \mathbf{X}_2\| \cdot \big(3\|\mathbf{L}^\star\|^2\|\mathbf{S}\|\nu + \|\mathbf{B}\|_+^2\|\mathbf{R}^{-1}\|_+(51\epsilon/C_\epsilon^u + 2\eta/C_\eta^u)\big). \tag{B.19}$$

To use this lemma, we pick $\nu = \frac{\sqrt{ns}\tau}{1-\rho}\left(C_\epsilon\epsilon + C_\eta\eta\right)$. We first show $\mathcal{K}$ maps $\mathcal{S}_\nu$ into itself and then show it is a contraction mapping. Plugging in the upper bounds for ϵ and η in the premises of Theorem 4.1, we have

$$\nu \le \min\left\{1, \frac{1}{\|\mathbf{S}\|}, \frac{1-\rho}{12\sqrt{ns}\tau\|\mathbf{L}^\star\|^2\|\mathbf{S}\|}, \frac{\underline{\sigma}(\mathbf{P}^\star)}{12}\right\}, \tag{B.20}$$

Following the premise upper bound of ϵ in Theorem 4.1 we have $\epsilon \le \min\{\|\mathbf{B}\|, 1\}$. This together with (B.20) makes Lemma B.5 applicable, and we get $\|\mathcal{K}(\mathbf{X})\| \le \frac{1}{12}\nu + \frac{1}{2}\nu = \frac{7}{12}\nu$ by cancelling off ϵ and η in (B.18) with the definition of ν, and applying the third upper bound for ν in (B.20). We know $\nu \le \underline{\sigma}(\mathbf{P}^\star)/12$ from (B.20), we have $\|\mathcal{K}(\mathbf{X})\| \le \frac{7}{144}\underline{\sigma}(\mathbf{P}^\star)$, thus $\mathbf{P}^\star + \mathcal{K}(\mathbf{X}) \succ 0$. This shows $\mathcal{K}(\mathbf{X}) \in \mathcal{S}_\nu$, i.e. $\mathcal{K}$ maps $\mathcal{S}_\nu$ into itself. Plugging the premise upper bounds for ϵ, η in Theorem 4.1 and the third upper bound for ν in (B.20) into (B.19) gives

$\|\mathcal{K}(\mathbf{X}_1) - \mathcal{K}(\mathbf{X}_2)\| \leq \frac{13}{24}\|\mathbf{X}_1 - \mathbf{X}_2\|$, i.e. $\mathcal{K}(\mathbf{X})$ is a contraction mapping on $\mathcal{S}_\nu$, which means $\mathcal{K}(\mathbf{X})$ has a unique fixed point $\mathbf{X}^\star \in \mathcal{S}_\nu$. From the discussion below (B.16), we know $\hat{\mathbf{P}}_{1:s}$ is a solution to cDARE$(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ and $\|\hat{\mathbf{P}}_{1:s} - \mathbf{P}^\star_{1:s}\| = \|\mathbf{X}^\star_{1:s}\| = \|\mathbf{X}^\star\| \leq \nu$, which shows (4.8).

B.2.3 Uniqueness of $\hat{\mathbf{P}}_{1:s}$

Note that $\mathbf{X}^\star \in \mathcal{S}_\nu$ gives $\|\mathbf{X}^\star\| < \nu$, and using (B.20), we have $\|\mathbf{X}^\star\| < \underline{\sigma}(\mathbf{P}^\star)$, thus $\mathbf{P}^\star + \mathbf{X}^\star \succ 0$. This implies $\hat{\mathbf{P}}_i = \mathbf{P}^\star_i + \mathbf{X}^\star_i \succ 0$ for all i. By Fact B.4, we know $\hat{\mathbf{P}}_{1:s}$ is the only possible solution to cDARE$(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ among $\mathbb{S}^n_{s,+}$. □

B.3 Proof of Theorem 4.2

We first provide the road map of the proof.

(a) We bound the controller difference $\|\mathbf{K}^\star_{1:s} - \hat{\mathbf{K}}_{1:s}\|$ in terms of $\|\hat{\mathbf{P}}_{1:s} - \mathbf{P}^\star_{1:s}\|$ and provide conditions under which $\hat{\mathbf{K}}_{1:s}$ stabilizes the true MJS$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$.

(b) For $\hat{J}$ incurred by the stabilizing $\hat{\mathbf{K}}_{1:s}$, we bound the suboptimality gap $\hat{J} - J^\star$ in terms of $\|\mathbf{K}^\star_{1:s} - \hat{\mathbf{K}}_{1:s}\|$.

(c) We combine steps (a), (b) and Theorem 4.1 to obtain the final result.

B.3.1 Properties of $\hat{\mathbf{K}}_{1:s}$

We show that when $\hat{\mathbf{P}}_{1:s}$ is close to $\mathbf{P}_{1:s}$, then $\hat{\mathbf{K}}_{1:s}$ is also close to $\mathbf{K}_{1:s}$.

Lemma B.6 (Controller mismatch)**.** *Suppose* $\|\hat{\mathbf{P}}_{1:s} - \mathbf{P}^\star_{1:s}\| \leq f(\epsilon, \eta)$ *for some function* $f(\epsilon, \eta)$ *such that* $\max\{\epsilon, \eta\} \leq f(\epsilon, \eta) \leq \Gamma_\star$. *Then, under Assumption A5.1, we have*

$$\|\mathbf{K}^\star_{1:s} - \hat{\mathbf{K}}_{1:s}\| \leq 28\Gamma_\star^3 \frac{(\underline{\sigma}(\mathbf{R}_{1:s}) + \Gamma_\star^3)}{\underline{\sigma}(\mathbf{R}_{1:s})^2} f(\epsilon, \eta) \tag{B.21}$$

Proof. Recall

$$\mathbf{K}^\star_i = -\big(\mathbf{R}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{B}_i\big)^{-1}\mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{A}_i,$$

$$\hat{\mathbf{K}}_i = -\big(\mathbf{R}_i + \hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{B}}_i\big)^{-1}\hat{\mathbf{B}}_i^\intercal \hat{\varphi}_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{A}}_i.$$

As an auxiliary step, we define $\tilde{\mathbf{K}}_i := -\big(\mathbf{R}_i + \hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{B}}_i\big)^{-1} \cdot \hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{A}}_i$. Then, we have

$$\|\mathbf{K}^\star_i - \hat{\mathbf{K}}_i\| \leq \|\mathbf{K}^\star_i - \tilde{\mathbf{K}}_i\| + \|\tilde{\mathbf{K}}_i - \hat{\mathbf{K}}_i\|. \tag{B.22}$$

Note that $\|\mathbf{K}^\star_{1:s} - \hat{\mathbf{K}}_{1:s}\| = \max_i \|\mathbf{K}^\star_i - \hat{\mathbf{K}}_i\|$, thus it suffices to bound $\|\mathbf{K}^\star_i - \tilde{\mathbf{K}}_i\|$ and $\|\tilde{\mathbf{K}}_i - \hat{\mathbf{K}}_i\|$ respectively. For $\|\mathbf{K}^\star_i - \tilde{\mathbf{K}}_i\|$, we can see $\|\mathbf{K}^\star_i - \tilde{\mathbf{K}}_i\| \leq M\delta_N + \delta_M N$ where

$$\begin{aligned}
M &= \|\big(\mathbf{R}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{B}_i\big)^{-1}\|, \quad N = \|\hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{A}}_i\| \\
\delta_M &= \|\big(\mathbf{R}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{B}_i\big)^{-1} - \big(\mathbf{R}_i + \hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{B}}_i\big)^{-1}\| \\
\delta_N &= \|\mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{A}_i - \hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{A}}_i\|
\end{aligned}$$

We next upper bound M, δ_N, δ_M, and N. Since we assume $\mathbb{R}_i \succ 0$, it is easy to see $M = \|\big(\mathbf{R}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{B}_i\big)^{-1}\| \leq \frac{1}{\underline{\sigma}(\mathbf{R}_i)}$. For δ_N, let $\Delta_{\mathbf{A}_i} = \hat{\mathbf{A}}_i - \mathbf{A}_i$, $\Delta_{\mathbf{B}_i} = \hat{\mathbf{B}}_i - \mathbf{B}_i$, $\Delta_{\mathbf{P}_i} = \hat{\mathbf{P}}_i - \mathbf{P}^\star_i$, then we have

$$\begin{aligned}
\delta_N &= \|\mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{A}_i - \hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{A}}_i\| \\
&= \|\mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{A}_i - (\mathbf{B}_i + \Delta_{\mathbf{B}_i})^\intercal \\
&\quad \cdot \big[\varphi_i(\mathbf{P}^\star_{1:s})\mathbf{A}_i + \varphi_i(\Delta_{\mathbf{P}^\star_{1:s}})\mathbf{A}_i + \varphi_i(\mathbf{P}^\star_{1:s})\Delta_{\mathbf{A}_i} + \varphi_i(\Delta_{\mathbf{P}^\star_{1:s}})\Delta_{\mathbf{A}_i}\big]\| \\
&= \|\mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{A}_i - \big[\mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{A}_i + \mathbf{B}_i^\intercal \varphi_i(\Delta_{\mathbf{P}^\star_{1:s}})\mathbf{A}_i \\
&\quad + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\Delta_{\mathbf{A}_i} + \mathbf{B}_i^\intercal \varphi_i(\Delta_{\mathbf{P}^\star_{1:s}})\Delta_{\mathbf{A}_i} + \Delta_{\mathbf{B}_i}^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{A}_i \\
&\quad + \Delta_{\mathbf{B}_i}^\intercal \varphi_i(\Delta_{\mathbf{P}^\star_{1:s}})\mathbf{A}_i + \Delta_{\mathbf{B}_i}^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\Delta_{\mathbf{A}_i} + \Delta_{\mathbf{B}_i}^\intercal \varphi_i(\Delta_{\mathbf{P}^\star_{1:s}})\Delta_{\mathbf{A}_i}\big]\| \\
&\overset{\text{(B.6)}}{\leq} \|\mathbf{A}_i\|\|\mathbf{B}_i\| f(\epsilon,\eta) + \|\mathbf{B}_i\|\|\mathbf{P}^\star_{1:s}\|\epsilon + \|\mathbf{B}_i\| f(\epsilon,\eta)\epsilon \\
&\quad + \|\mathbf{A}_i\|\|\mathbf{P}^\star_{1:s}\|\epsilon + \|\mathbf{A}_i\| f(\epsilon,\eta)\epsilon + \|\mathbf{P}^\star_{1:s}\|\epsilon^2 + f(\epsilon,\eta)\epsilon^2, \\
&\leq 3\Gamma_\star^2 f(\epsilon,\eta),
\end{aligned}$$

where the last line follows from the assumption that $\epsilon < f(\epsilon,\eta)$. For δ_M, we have

$$\begin{aligned}
\delta_M &= \|\big(\mathbf{R}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{B}_i\big)^{-1} - \big(\mathbf{R}_i + \hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{B}}_i\big)^{-1}\| \\
&\overset{\text{(B.3)}}{\leq} \|\big(\mathbf{R}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{B}_i\big)^{-1}\| \cdot \|\big(\mathbf{R}_i + \hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{B}}_i\big)^{-1}\| \\
&\quad \cdot \|\hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{B}}_i - \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}^\star_{1:s})\mathbf{B}_i\| \\
&\leq \frac{3\Gamma_\star^2 f(\epsilon,\eta)}{\underline{\sigma}(\mathbf{R}_i)^2}.
\end{aligned}$$

Similarly, we have the following for N.

$$\begin{aligned}N &= \|\hat{\mathbf{B}}_i^\intercal \varphi_i(\hat{\mathbf{P}}_{1:s})\hat{\mathbf{A}}_i\| \\ &= \|\mathbf{B}_i^\intercal \varphi_i(\mathbf{P}_{1:s}^\star)\mathbf{A}_i + \mathbf{B}_i^\intercal \varphi_i(\Delta_{\mathbf{P}_{1:s}^\star})\mathbf{A}_i + \mathbf{B}_i^\intercal \varphi_i(\mathbf{P}_{1:s}^\star)\Delta_{\mathbf{A}_i} \\ &\quad + \mathbf{B}_i^\intercal \varphi_i(\Delta_{\mathbf{P}_{1:s}^\star})\Delta_{\mathbf{A}_i} + \Delta_{\mathbf{B}_i}^\intercal \varphi_i(\mathbf{P}_{1:s}^\star)\mathbf{A}_i + \Delta_{\mathbf{B}_i}^\intercal \varphi_i(\Delta_{\mathbf{P}_{1:s}^\star})\mathbf{A}_i \\ &\quad + \Delta_{\mathbf{B}_i}^\intercal \varphi_i(\mathbf{P}_{1:s}^\star)\Delta_{\mathbf{A}_i} + \Delta_{\mathbf{B}_i}^\intercal \varphi_i(\Delta_{\mathbf{P}_{1:s}^\star})\Delta_{\mathbf{A}_i}\| \\ &\leq \big(\|\mathbf{A}_i\|\|\mathbf{B}_i\| + \|\mathbf{A}_i\|\|\mathbf{P}_{1:s}^\star\| + \|\mathbf{B}_i\|\|\mathbf{P}_{1:s}^\star\| + \|\mathbf{A}_i\|\epsilon + \|\mathbf{B}_i\|\epsilon \\ &\quad + \|\mathbf{P}_{1:s}^\star\|\epsilon + \epsilon^2) \cdot f(\epsilon,\eta\big) + \|\mathbf{A}_i\|\|\mathbf{B}_i\|\|\mathbf{P}_{1:s}^\star\| \\ &\leq 3\Gamma_\star^2 f(\epsilon,\eta) + \Gamma_\star^3.\end{aligned}$$

Combining the bounds for M, δ_N, δ_M, and N we obtained thus far, we have $\|\mathbf{K}_i^\star - \tilde{\mathbf{K}}_i\| \leq 12\Gamma_\star^2 \frac{(\sigma(\mathbf{R}_i)+\Gamma_\star^3)}{\underline{\sigma}(\mathbf{R}_i)^2} f(\epsilon,\eta)$. Using similar techniques, for the other term on the RHS of (B.22), we can show $\|\tilde{\mathbf{K}}_i - \hat{\mathbf{K}}_i\| \leq 16\Gamma_\star^3 \frac{(\sigma(\mathbf{R}_i)+\Gamma_\star^3)}{\underline{\sigma}(\mathbf{R}_i)^2}\eta$. Recall we assume $\eta \leq f(\epsilon,\eta)$, then by triangle inequality, we have $\|\mathbf{K}_i^\star - \hat{\mathbf{K}}_i\| \leq 28\Gamma_\star^3 \frac{(\sigma(\mathbf{R}_i)+\Gamma_\star^3)}{\underline{\sigma}(\mathbf{R}_i)^2} f(\epsilon,\eta)$. ☐

For $\hat{\mathbf{K}}_{1:s}$, let $\hat{\mathbf{L}}_i := \mathbf{A}_i + \mathbf{B}_i\hat{\mathbf{K}}_i$ and define the augmented closed-loop state matrix $\hat{\boldsymbol{\mathcal{L}}} \in \mathbb{R}^{sn^2 \times sn^2}$ with (i,j)-th $n^2 \times n^2$ block given by $[\hat{\boldsymbol{\mathcal{L}}}]_{ij} := \mathbf{T}(j,i)\hat{\mathbf{L}}_j \otimes \hat{\mathbf{L}}_j$.

Lemma B.7 (Stabilizability of $\hat{\mathbf{K}}$)**.** *Suppose* $\|\hat{\mathbf{K}}_{1:s} - \mathbf{K}_{1:s}^\star\| \leq \frac{1-\rho}{2\sqrt{s}\tau(1+2\|\mathbf{L}_{1:s}^\star\|)\|\mathbf{B}_{1:s}\|} =: \bar{\epsilon}_{\mathbf{K}}$, *then (a)* $\rho(\hat{\boldsymbol{\mathcal{L}}}) < \frac{1+\rho}{2}$, *i.e.* $\hat{\mathbf{K}}_{1:s}$ *is a stabilizing controller; (b)* $\|\hat{\boldsymbol{\mathcal{L}}}^k\| \leq \tau(\frac{1+\rho}{2})^k$*;*

Proof. Let $\boldsymbol{\Delta}_{\mathbf{K}_i} := \hat{\mathbf{K}}_i - \mathbf{K}_i^\star$, then we have $[\hat{\boldsymbol{\mathcal{L}}}]_{ij} - [\boldsymbol{\mathcal{L}}^\star]_{ij} = \mathbf{T}(j,i)[(\mathbf{B}_j\boldsymbol{\Delta}_{\mathbf{K}_j}) \otimes (\mathbf{B}_j\boldsymbol{\Delta}_{\mathbf{K}_j}) + (\mathbf{B}_j\boldsymbol{\Delta}_{\mathbf{K}_j}) \otimes \mathbf{L}_j^\star + \mathbf{L}_j^\star \otimes (\mathbf{B}_j\boldsymbol{\Delta}_{\mathbf{K}_j})]$. Taking the norms gives $\|[\hat{\boldsymbol{\mathcal{L}}}]_{ij} - [\boldsymbol{\mathcal{L}}^\star]_{ij}\| \leq \mathbf{T}(j,i)(\|\mathbf{B}_j\|^2 \cdot \|\boldsymbol{\Delta}_{\mathbf{K}_j}\|^2 + 2\|\mathbf{B}_j\|\|\mathbf{L}_j^\star\|\|\boldsymbol{\Delta}_{\mathbf{K}_j}\|) \leq \mathbf{T}(j,i)(1+2\|\mathbf{L}_j\|)\|\mathbf{B}_j\|\|\boldsymbol{\Delta}_{\mathbf{K}_j}\| \leq \mathbf{T}(j,i)\frac{1-\rho}{2\sqrt{s}\tau}$. Using Cauchy-Schwartz inequality gives $\|\hat{\boldsymbol{\mathcal{L}}} - \boldsymbol{\mathcal{L}}^\star\| \leq \big(\sum_{i,j} \|[\hat{\boldsymbol{\mathcal{L}}}]_{ij} - [\boldsymbol{\mathcal{L}}^\star]_{ij}\|^2\big)^{0.5} \leq \frac{1-\rho}{2\tau}$. Finally, we can conclude the proof by invoking Fact B.3. ☐

B.3.2 $\hat{J} - J^\star$ vs. $\|\mathbf{K}_{1:s}^\star - \hat{\mathbf{K}}_{1:s}\|$

Adapting (Jansch-Porto et al., 2020, Lemma 3-(2)) to noisy MJS and infinite-horizon average cost case, we have the following result.

Lemma B.8. *Suppose the controller* $\hat{\mathbf{K}}_{1:s}$ *is a stabilizing controller. Let* $\boldsymbol{\Pi}_i := \pi_\infty(i)\mathbf{I}_n$ *and* $\boldsymbol{\Pi} := \mathbf{diag}(\boldsymbol{\Pi}_{1:s})$*. Let* $\boldsymbol{\Sigma}_{1:s}$ *be the solution to* $\tilde{\mathbf{vec}}(\boldsymbol{\Sigma}) = \hat{\boldsymbol{\mathcal{L}}}\tilde{\mathbf{vec}}(\boldsymbol{\Sigma}) + \sigma_{\mathbf{w}}^2\tilde{\mathbf{vec}}(\boldsymbol{\Pi})$*, where* $\boldsymbol{\Sigma} := \mathbf{diag}(\boldsymbol{\Sigma}_{1:s})$*. Then,* $\hat{J} - J^\star = \sum_i \mathbf{tr}\big(\boldsymbol{\Sigma}_i(\hat{\mathbf{K}}_i - \mathbf{K}_i^\star)^\intercal(\mathbf{R}_i + \mathbf{B}_i^\intercal\varphi_i(\mathbf{P}_{1:s}^\star)\mathbf{B}_i(\hat{\mathbf{K}}_i - \mathbf{K}_i^\star))\big)$

In Lemma B.8, the equation described by $\boldsymbol{\Sigma}$ is essentially the coupled Lyapunov equation for MJS, and it can be shown $\boldsymbol{\Sigma}_i = \lim_{t\to\infty}\mathbb{E}[\mathbf{x}_t\mathbf{x}_t^\intercal \mathbf{1}_{\{\omega_t=i\}}]$ where $\mathbf{x}_t$ is the state under controller $\hat{\mathbf{K}}_{1:s}$. Combining Lemma B.7 and B.8, we have

Corollary B.9. *Suppose* $\|\hat{\mathbf{K}}_{1:s} - \mathbf{K}^\star_{1:s}\| \le \bar{\epsilon}_{\mathbf{K}}$, *then*

$$\hat{J} - J^\star \le \frac{2\sigma_{\mathbf{w}}^2 s^{1.5}\sqrt{n}\min\{n,p\}\tau}{1-\rho}(\|\mathbf{R}_{1:s}\| + \Gamma_\star^3)\|\hat{\mathbf{K}}_{1:s} - \mathbf{K}^\star_{1:s}\|^2.$$

Proof. We first bound $\|\boldsymbol{\Sigma}_i\|$ in Lemma B.8. Similar to (B.15), we have $\boldsymbol{\Sigma} = \sigma_{\mathbf{w}}^2 \cdot \tilde{\mathbf{vec}}^{-1} \circ (\mathbf{I} - \hat{\boldsymbol{\mathcal{L}}})^{-1} \circ \tilde{\mathbf{vec}}(\boldsymbol{\Pi})$. Using Fact (B.2) and the sub-multiplicative property of operator norms, we have $\|\boldsymbol{\Sigma}_i\| \le \|\boldsymbol{\Sigma}\| \le \sqrt{ns}\|(\mathbf{I} - \hat{\boldsymbol{\mathcal{L}}})^{-1}\|\|\boldsymbol{\Pi}\|$. Note that $\|\boldsymbol{\Pi}\| \le 1$ and $\|(\mathbf{I} - \hat{\boldsymbol{\mathcal{L}}})^{-1}\| = \|\sum_{k=0}^\infty (\hat{\boldsymbol{\mathcal{L}}})^k\| \le \sum_{k=0}^\infty \|(\hat{\boldsymbol{\mathcal{L}}})^k\| \le \frac{2\tau}{1-\rho}$, where the last inequality follows from Lemma B.7 (b). Thus, $\|\boldsymbol{\Sigma}_i\| \le \frac{2\sigma_{\mathbf{w}}^2\sqrt{sn}\tau}{1-\rho}$. Finally, applying Lemma B.8 gives $\hat{J} - J^\star \le s\|\boldsymbol{\Sigma}_i\|(\|\mathbf{R}_{1:s}\| + \|\mathbf{B}_{1:s}\|^2\|\mathbf{P}^\star\|)\|\hat{\mathbf{K}}_{1:s} - \mathbf{K}^\star_{1:s}\|_{\mathrm{F}}^2 \le \frac{2\sigma_{\mathbf{w}}^2 s^{1.5}\sqrt{n}\min\{n,p\}\tau}{1-\rho}(\|\mathbf{R}_{1:s}\| + \Gamma_\star^3)\|\hat{\mathbf{K}}_{1:s} - \mathbf{K}^\star_{1:s}\|^2$. □

B.3.3 Proof of Theorem 4.2

To prove Theorem 4.2, we only need to combine Theorem 4.1, Lemma B.6, and Corollary B.9. By Theorem 4.1, we can choose $f(\epsilon, \eta) := \frac{\sqrt{ns}\tau}{1-\rho}(C_\epsilon \epsilon + C_\eta \eta)$ in Lemma B.6. The, when $C_\epsilon\epsilon + C_\eta\eta \le \frac{(1-\rho)\min\{\Gamma_\star, \underline{\sigma}(\mathbf{R}_{1:s})^2\bar{\epsilon}_{\mathbf{K}}\}}{28\sqrt{ns}\tau\Gamma_\star^3(\underline{\sigma}(\mathbf{R}_{1:s}) + \Gamma_\star^3)}$, the premise conditions $\max\{\epsilon, \eta\} \le f(\epsilon, \eta) \le \Gamma_\star$ in Lemma B.6 and $\|\hat{\mathbf{K}}_{1:s} - \mathbf{K}^\star_{1:s}\| \le \bar{\epsilon}_{\mathbf{K}}$ in Corollary B.9 hold. Theorem 4.1 and Lemma B.6 give $\|\mathbf{K}^\star_{1:s} - \hat{\mathbf{K}}_{1:s}\| \le 28\sqrt{ns}\tau\Gamma_\star^3\frac{(\underline{\sigma}(\mathbf{R}_{1:s}) + \Gamma_\star^3)}{(1-\rho)\underline{\sigma}(\mathbf{R}_{1:s})^2}(C_\epsilon\epsilon + C_\eta\eta)$ which shows (4.9). Combining this with Corollary B.9 shows (4.10). □

B.4 Proof of Lemma B.5

To ease the exposition, let $\mathbf{P}^\star_{\mathbf{X}} := \mathbf{P}^\star + \mathbf{X}$ and define $\mathbf{P}^\star_{\mathbf{X}_1}, \mathbf{P}^\star_{\mathbf{X}_1}$ similarly. Let $\boldsymbol{\Delta}_{\mathbf{A}} := \hat{\mathbf{A}} - \mathbf{A}$, $\boldsymbol{\Delta}_{\mathbf{B}} := \hat{\mathbf{B}} - \mathbf{B}$, and $\boldsymbol{\Delta}_{\mathbf{S}} := \hat{\mathbf{S}} - \mathbf{S}$. We list a few preliminary results (when $\mathbf{X} \in \mathcal{S}_\nu$) to be used later.

- $\|\boldsymbol{\Phi}(\mathbf{X})\| \le \|\mathbf{X}\|, \quad \|\hat{\boldsymbol{\Phi}}(\mathbf{X})\| \le \|\mathbf{X}\|.$ (B.23)
- $\|\boldsymbol{\Phi}(\mathbf{X}) - \hat{\boldsymbol{\Phi}}(\mathbf{X})\| \le \eta\|\mathbf{X}\|.$ (B.24)
- $\max\{\|\mathbf{P}^\star_{\mathbf{X}}\|, \|\boldsymbol{\Phi}(\mathbf{P}^\star_{\mathbf{X}})\|, \|\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}})\|\} \le \|\mathbf{P}^\star\|_+.$ (B.25)
- $\|\mathbf{S}\| \le \|\mathbf{B}\|^2\|\mathbf{R}^{-1}\|, \quad \|\boldsymbol{\Delta}_{\mathbf{S}}\| \le 3\|\mathbf{B}\|\|\mathbf{R}^{-1}\|\epsilon, \quad \|\hat{\mathbf{S}}\| \le 4\|\mathbf{B}\|^2\|\mathbf{R}^{-1}\|$ (B.26)
- $\max\left\{\|(\mathbf{I} + \mathbf{S}\boldsymbol{\Phi}(\mathbf{P}^\star_{\mathbf{X}}))^{-1}\|, \|(\mathbf{I} + \mathbf{S}\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}}))^{-1}\|\right\} \le \|\mathbf{B}\|_+^2\|\mathbf{R}^{-1}\|_+\|\mathbf{P}^\star\|_+$ (B.27)
- $\max\left\{\|(\mathbf{I} + \hat{\mathbf{S}}\boldsymbol{\Phi}(\mathbf{P}^\star_{\mathbf{X}}))^{-1}\|, \|(\mathbf{I} + \hat{\mathbf{S}}\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}}))^{-1}\|\right\} \le 4\|\mathbf{B}\|_+^2\|\mathbf{R}^{-1}\|_+\|\mathbf{P}^\star\|_+.$ (B.28)

(B.25) is due to $\nu \le 1$, and (B.26) uses $\|\boldsymbol{\Delta}_{\mathbf{B}}\| \le \epsilon \le \|\mathbf{B}\|$. (B.27) and (B.28) follows from (B.2), (B.25), (B.26).

Now, we are ready to begin the main proof. We first define

$$\mathcal{G}_1(\mathbf{X}) := \mathcal{F}(\mathbf{P}^\star_{\mathbf{X}}; \mathbf{A}, \mathbf{B}, \hat{\mathbf{T}}) - \mathcal{F}(\mathbf{P}^\star_{\mathbf{X}}; \hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{T}})$$
$$\mathcal{G}_2(\mathbf{X}) := \mathcal{F}(\mathbf{P}^\star_{\mathbf{X}}; \mathbf{A}, \mathbf{B}, \mathbf{T}) - \mathcal{F}(\mathbf{P}^\star_{\mathbf{X}}; \mathbf{A}, \mathbf{B}, \hat{\mathbf{T}}).$$

Then, we have the following decomposition.

$$\mathcal{K}(\mathbf{X}) = \mathcal{T}^{-1}(\mathcal{G}_1(\mathbf{X}) + \mathcal{G}_2(\mathbf{X}) - \mathcal{H}(\mathbf{X})), \tag{B.29}$$
$$\begin{aligned}\mathcal{K}(\mathbf{X}_1) - \mathcal{K}(\mathbf{X}_2) = \mathcal{T}^{-1}(&\mathcal{G}_1(\mathbf{X}_1) - \mathcal{G}_1(\mathbf{X}_2)\\ &+ \mathcal{G}_2(\mathbf{X}_1) - \mathcal{G}_2(\mathbf{X}_2) - \mathcal{H}(\mathbf{X}_1) + \mathcal{H}(\mathbf{X}_2))\end{aligned} \tag{B.30}$$

To bound the $\|\mathcal{K}(\mathbf{X})\|$ and $\|\mathcal{K}(\mathbf{X}_1) - \mathcal{K}(\mathbf{X}_2)\|$, we will bound $\|\mathcal{T}^{-1}\|$, $\|\mathcal{H}(\mathbf{X})\|$, $\|\mathcal{G}_1(\mathbf{X})\|$, $\|\mathcal{G}_2(\mathbf{X})\|$, $\|\mathcal{H}(\mathbf{X}_1) - \mathcal{H}(\mathbf{X}_2)\|$, $\|\mathcal{G}_1(\mathbf{X}_1) - \mathcal{G}_1(\mathbf{X}_2)\|$, $\|\mathcal{G}_2(\mathbf{X}_1) - \mathcal{G}_2(\mathbf{X}_2)\|$ individually, for any $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2 \in \mathcal{S}_\nu$ and then combine them using triangle inequality and operator composition sub-multiplicativity, i.e.

$$\|\mathcal{K}(\mathbf{X})\| \leq \|\mathcal{T}^{-1}\|(\|\mathcal{G}_1(\mathbf{X})\| + \|\mathcal{G}_2(\mathbf{X})\| + \|\mathcal{H}(\mathbf{X})\|) \tag{B.31}$$
$$\begin{aligned}\|\mathcal{K}(\mathbf{X}_1) - \mathcal{K}(\mathbf{X}_2)\| \leq \|\mathcal{T}^{-1}\|(&\|\mathcal{G}_1(\mathbf{X}_1) - \mathcal{G}_1(\mathbf{X}_2)\|\\ &+ \|\mathcal{G}_2(\mathbf{X}_1) - \mathcal{G}_2(\mathbf{X}_2)\| + \|\mathcal{H}(\mathbf{X}_1) - \mathcal{H}(\mathbf{X}_2)\|)\end{aligned} \tag{B.32}$$

B.4.1 Bound $\|\mathcal{K}(\mathbf{X})\|$

By the definition of $\mathcal{T}^{-1}$ in (B.15), we know $\mathcal{T}^{-1}(\mathbf{Y}) = \tilde{\mathbf{vec}}^{-1} \circ (\mathbf{I} - \boldsymbol{\mathcal{L}}^{\star\intercal})^{-1} \circ \tilde{\mathbf{vec}}(\mathbf{Y})$. Then, for $\|\mathcal{T}^{-1}\|$, similar to the proof for Corollary B.9, we have $\|\mathcal{T}^{-1}\| \leq \frac{\sqrt{sn\tau}}{1-\rho}$. By definition of $\mathcal{H}(\mathbf{X})$ in (B.14), we have $\|\mathcal{H}(\mathbf{X})\| \leq \|\mathbf{L}^\star\|^2\|\mathbf{S}\|\|\mathbf{X}\|^2 \leq \|\mathbf{L}^\star\|^2\|\mathbf{S}\|\nu^2$, where (B.1) and (B.23) are used. For term $\mathcal{G}_1(\mathbf{X})$, using (B.4), we can decompose it as

$$\begin{aligned}&\mathcal{G}_1(\mathbf{X}) =\\ &\quad - \mathbf{A}^\intercal\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}})(\mathbf{I} + \mathbf{S}\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}}))^{-1}\boldsymbol{\Delta}_{\mathbf{S}}\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}})(\mathbf{I} + \hat{\mathbf{S}}\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}}))^{-1}\mathbf{A}\\ &\quad + \boldsymbol{\Delta}_{\mathbf{A}}^\intercal\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}})(\mathbf{I} + \hat{\mathbf{S}}\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}}))^{-1}\mathbf{A} + \mathbf{A}^\intercal\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}})(\mathbf{I} + \hat{\mathbf{S}}\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}}))^{-1}\boldsymbol{\Delta}_{\mathbf{A}}\\ &\quad + \boldsymbol{\Delta}_{\mathbf{A}}^\intercal\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}})(\mathbf{I} + \hat{\mathbf{S}}\hat{\boldsymbol{\Phi}}(\mathbf{P}^\star_{\mathbf{X}}))^{-1}\boldsymbol{\Delta}_{\mathbf{A}}.\end{aligned}$$

With properties (B.1), (B.25), (B.26), and the premise assumption $\epsilon \leq \|\mathbf{B}\|$, we can show $\|\mathcal{G}_1(\mathbf{X})\| \leq 3\|\mathbf{A}\|_+^2\|\mathbf{B}\|_+ \cdot \|\mathbf{P}^\star\|_+^2\|\mathbf{R}^{-1}\|_+\epsilon$. Similarly, we can obtain that $\|\mathcal{G}_2(\mathbf{X})\| \leq \|\mathbf{A}\|_+^2 \cdot \|\mathbf{B}\|_+^4\|\mathbf{P}^\star\|_+^3\|\mathbf{R}^{-1}\|_+^2\eta$ by invoking (B.1), (B.4), (B.24), (B.25), (B.26), and (B.27). Finally, using the relation in (B.31), we can show the upper bound for $\|\mathcal{K}(\mathbf{X})\|$ in (B.18).

B.4.2 Bound $\|\mathcal{K}(\mathbf{X}_1) - \mathcal{K}(\mathbf{X}_2)\|$

We first derive bounds for $\|\mathcal{H}(\mathbf{X}_1) - \mathcal{H}(\mathbf{X}_2)\|$, $\|\mathcal{G}_1(\mathbf{X}_1) - \mathcal{G}_1(\mathbf{X}_2)\|$, and $\|\mathcal{G}_2(\mathbf{X}_1) - \mathcal{G}_2(\mathbf{X}_2)\|$. With the help of (B.4), the following can be obtained.

$$\begin{aligned}\mathcal{H}(\mathbf{X}_1) - \mathcal{H}(\mathbf{X}_2) = &\mathbf{L}^{\star\mathsf{T}}\mathbf{\Phi}(\mathbf{X}_1)(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star_{\mathbf{X}_1}))^{-1} \\ &\cdot \mathbf{S}\mathbf{\Phi}(\mathbf{X}_2 - \mathbf{X}_1)(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star_{\mathbf{X}_2}))^{-1}\mathbf{S}\mathbf{\Phi}(\mathbf{X}_1)\mathbf{L}^\star \\ &- \mathbf{L}^{\star\mathsf{T}}\mathbf{\Phi}(\mathbf{X}_2 - \mathbf{X}_1)(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star_{\mathbf{X}_2}))^{-1}\mathbf{S}\mathbf{\Phi}(\mathbf{X}_2)\mathbf{L}^\star \\ &- \mathbf{L}^{\star\mathsf{T}}\mathbf{\Phi}(\mathbf{X}_1)(\mathbf{I} + \mathbf{S}\mathbf{\Phi}(\mathbf{P}^\star_{\mathbf{X}_2}))^{-1}\mathbf{S}\mathbf{\Phi}(\mathbf{X}_2 - \mathbf{X}_1)\mathbf{L}^\star.\end{aligned} \tag{B.33}$$

Using (B.1), (B.23), and $\nu \le \|\mathbf{S}\|^{-1}$, we have $\|\mathcal{H}(\mathbf{X}_1) - \mathcal{H}(\mathbf{X}_2)\| \le 3\|\mathbf{L}^\star\|^2\|\mathbf{S}\|\nu\|\mathbf{X}_2 - \mathbf{X}_1\|$. Similarly, $\|\mathcal{G}_1(\mathbf{X}_1) - \mathcal{G}_1(\mathbf{X}_2)\| \le 51\|\mathbf{A}\|_+^2\|\mathbf{B}\|_+^5\|\mathbf{P}^\star\|_+^3\|\mathbf{R}^{-1}\|_+^3\|\mathbf{X}_2 - \mathbf{X}_1\|\epsilon$ and $\|\mathcal{G}_2(\mathbf{X}_1) - \mathcal{G}_2(\mathbf{X}_2)\| \le 2\|\mathbf{A}\|_+^2\|\mathbf{B}\|_+^6\|\mathbf{P}^\star\|_+^3\|\mathbf{R}^{-1}\|_+^3\|\mathbf{X}_2 - \mathbf{X}_1\|\eta$ can be established. Plugging these results into the relation in (B.32) shows the bound for $\|\mathcal{K}(\mathbf{X}_1) - \mathcal{K}(\mathbf{X}_2)\|$ in (B.19) □

Appendix C
Proofs for Results in Chapter 5

Consider MJS-LQR($\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$) with dynamics noise $\mathbf{w}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}_\mathbf{w})$, some arbitrary initial state $\mathbf{x}_0$ and stabilizing controller $\mathbf{K}_{1:s}$. The input is $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$ where exploration noise $\mathbf{z}_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}_\mathbf{z})$. Let $\mathbf{L}_i := \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i$. Let $\boldsymbol{\mathcal{L}} \in \mathbb{R}^{sn^2 \times sn^2}$ denote the augmented closed-loop state matrix with (i,j)-th $n^2 \times n^2$ block given by $[\boldsymbol{\mathcal{L}}]_{ij} := \mathbf{T}(j,i)\mathbf{L}_j \otimes \mathbf{L}_j$. Let $\tau_{\boldsymbol{\mathcal{L}}} > 0$ and $\rho_{\boldsymbol{\mathcal{L}}} \in [0,1)$ be two constants such that $\|\boldsymbol{\mathcal{L}}^k\| \le \tau_{\boldsymbol{\mathcal{L}}}\rho_{\boldsymbol{\mathcal{L}}}^k$. By definition, one available choice for $\tau_{\boldsymbol{\mathcal{L}}}$ and $\rho_{\boldsymbol{\mathcal{L}}}$ are $\tau(\boldsymbol{\mathcal{L}})$ and $\rho(\boldsymbol{\mathcal{L}})$.

We define the following cumulative cost conditioned on the initial state $\mathbf{x}_0$, initial mode ω_0, and controller $\mathbf{K}_{1:s}$.

$$J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \boldsymbol{\Sigma}_\mathbf{z}\}) := \sum_{t=1}^{T} \mathbb{E}[\mathbf{x}_t^\intercal \mathbf{Q}_{\omega_t}\mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega_t}\mathbf{u}_t \mid \mathbf{x}_0, \omega_0, \mathbf{K}_{1:s}]. \tag{C.1}$$

The definition of this cumulative cost coincides with the cost $\sum_{t=1}^{T_q} c_{T_0+\dots+T_{q-1}+t}$ in the definition of Regret_q in (5.2) with $\mathbf{x}_0, \omega_0, \mathbf{K}_{1:s}$ setting to $\mathbf{x}_0^{(q)}, \omega^{(q)}(0), \mathbf{K}_{1:s}^{(q)}$ since Regret_q depends on randomness in $\mathcal{F}_{q-1}$ only through $\mathbf{x}_0^{(q)}, \omega^{(q)}(0), \mathbf{K}_{1:s}^{(q)}$. In the remainder of this appendix, for simplicity, we will drop the conditions $\mathbf{x}_0, \omega_0, \mathbf{K}_{1:s}$ in the expectation and simply write $\mathbb{E}[\,\cdot \mid \mathbf{x}_0, \omega_0, \mathbf{K}_{1:s}]$ as $\mathbb{E}[\cdot]$. So, for any measurable function f, $\mathbb{E}[f(\mathbf{x}_0, \omega_0, \mathbf{K}_{1:s})] = f(\mathbf{x}_0, \omega_0, \mathbf{K}_{1:s})$. Note that even though the results in this appendix are derived for conditional expectation $\mathbb{E}[\,\cdot \mid \mathbf{x}_0, \omega_0, \mathbf{K}_{1:s}]$, most of them also hold for the total expectation $\mathbb{E}[\cdot]$.

For the infinite-horizon case, we define the following infinite-horizon average cost without exploration noise $\mathbf{z}_t$ and starting from $\mathbf{x}_0 = 0$.

$$J(0, \omega_0, \{\mathbf{K}_{1:s}\}) := \limsup_{T\to\infty} \frac{1}{T} J_T(0, \omega_0, \{\mathbf{K}_{1:s}, 0\}). \tag{C.2}$$

Let $\mathbf{P}_{1:s}^\star$ denote the solution to cDARE($\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$) defined in (2.35). Let $\mathbf{K}_{1:s}^\star$ denote the resulting infinite-horizon optimal controller computed using $\mathbf{P}_{1:s}^\star$ and following (2.36). Note that the optimal infinite-horizon average cost $J^\star$ in Lemma 2.22 is achieved if the optimal controller $\mathbf{K}_{1:s}^\star$ is used, i.e.

$$J^\star = J(0, \omega_0, \{\mathbf{K}_{1:s}^\star\}). \tag{C.3}$$

Note that if the underlying Markov chain $\mathbf{T}$ is ergodic, for any initial state $\mathbf{x}_0$ and mode ω_0, $J^\star = J(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}^\star\})$. Let $\mathbf{L}_i^\star = \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i^\star$ for all $i \in [s]$ denote the closed-loop state matrix when the optimal controller $\mathbf{K}_{1:s}^\star$ is used. Define the augmented state matrix $\boldsymbol{\mathcal{L}}^\star$ such that its (i,j)-th block is given by $[\boldsymbol{\mathcal{L}}^\star]_{ij} := \mathbf{T}(j,i)\mathbf{L}_j^\star \otimes \mathbf{L}_j^\star$. From Costa et al. (2006), we know $\mathbf{K}_{1:s}^\star$ is stabilizes the MJS, thus $\rho^\star := \rho(\boldsymbol{\mathcal{L}}^\star) < 1$.

Since Regret_q defined in (5.2) can be written as

$$\mathrm{Regret}_q = J_T(\mathbf{x}_0^{(q)}, \omega^{(q)}(0), \{\mathbf{K}_{1:s}^{(q)}, \sigma_{\mathbf{z},q}^2\mathbf{I}_p\}) - TJ^\star, \tag{C.4}$$

to evaluate $\mathrm{Regret}(T)$, it suffices to evaluate $J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \boldsymbol{\Sigma}_\mathbf{z}\}) - TJ^\star$ for generic $\mathbf{x}_0$, ω_0, $\mathbf{K}_{1:s}$, and $\boldsymbol{\Sigma}_\mathbf{z}$. The outline of this Appendix C is as follows.

- In Appendix C.1, we list a few shorthand notations to ease later expositions.
- In Appendix C.2, we provide perturbation results $J(0, \omega_0, \{\mathbf{K}_{1:s}\}) - J^\star$ following from Chapter 4.1.
- In Appendix C.3, we evaluate $J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \boldsymbol{\Sigma}_\mathbf{z}\}) - TJ(0, \omega_0, \{\mathbf{K}_{1:s}\})$. Then, applying the results in Appendix C.2, for each epoch, we can bound the single epoch regret $J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \boldsymbol{\Sigma}_\mathbf{z}\}) - TJ^\star$.
- In Appendix C.4, we stitch regrets for all epochs together, and combine them with identification results in Appendix A to bound $\mathrm{Regret}(T)$.

C.1 Preliminaries

In the following, we define a few notations to ease the exposition in the appendix. Note that, for notations under parameterized form, i.e., notations which are functions of (δ, ρ, τ) etc., one can choose these parameters freely to get different deterministic quantities.

Table C.1 introduces notations and constants related to the choice of tuning parameters $c_\mathbf{x}$, $c_\mathbf{z}$, and the shortest trajectory (initial epoch) length such that theoretical performance guarantees can be achieved. Recall that $\mathbf{K}_{1:s}^{(0)}$ is the stabilizing controller for epoch 0 in Algorithm 2. We let $\mathbf{L}_i^{(0)} := \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i^{(0)}$, for all $i \in [s]$, denote the closed-loop state matrix, and $\boldsymbol{\mathcal{L}}^{(0)} \in \mathbb{R}^{sn^2 \times sn^2}$ denotes the augmented closed-loop state matrix with (i,j)-th $n^2 \times n^2$ block given by $[\boldsymbol{\mathcal{L}}^{(0)}]_{ij} = \mathbf{T}(j,i)\mathbf{L}_j^{(0)} \otimes \mathbf{L}_j^{(0)}$. $\tau(\cdot)$ is as in Definition 2.1 and $\rho(\cdot)$ denotes the

Table C.1: Notations — Tuning Parameters and Trajectory Length

$\bar{\sigma}_{\mathbf{z}}$ (*depending on context*)	$\sigma_{\mathbf{z}}$ or $\sigma_{\mathbf{z},0}$ or $\sqrt{\\|\mathbf{\Sigma}_{\mathbf{z}}\\|}$
$\bar{\sigma}_{\mathbf{w}}$ (*depending on context*)	$\sigma_{\mathbf{w}}$ or $\sqrt{\\|\mathbf{\Sigma}_{\mathbf{w}}\\|}$
$C_{\mathbf{z}}$	$\bar{\sigma}_{\mathbf{z}}/\bar{\sigma}_{\mathbf{w}}$
$\bar{\sigma}^2$	$\\|\mathbf{B}_{1:s}\\|^2\bar{\sigma}_{\mathbf{z}}^2 + \bar{\sigma}_{\mathbf{w}}^2$
$\underline{c}_{\mathbf{x}}(\rho,\tau)$	$3\sqrt{\frac{18n\sqrt{s}\tau\bar{\sigma}^2}{\pi_{\min}\bar{\sigma}_{\mathbf{w}}^2(1-\rho)}}$
$\underline{c}_{\mathbf{z}}$	$\max\left\{(\sqrt{3}+\sqrt{6})\sqrt{p}, \sqrt{3\log(\frac{6}{\pi_{\min}})}\right\}$
$\bar{\tau}$	$\max\{\tau(\boldsymbol{\mathcal{L}}^{(0)}), \tau(\boldsymbol{\mathcal{L}}^\star)\}$
$\bar{\rho}$	$\max\{\rho(\boldsymbol{\mathcal{L}}^{(0)}), \frac{1+\rho}{2}\}$
C_{MC}	$t_{\mathrm{MC}} \cdot \max\{3, 3-3\log(\pi_{\max}\log(s))\}$
$\underline{T}_{MC,1}(\mathcal{C},\delta)$	$\left(68\mathcal{C}\pi_{\max}\pi_{\min}^{-2}\log(\frac{s}{\delta})\right)^2$
$\underline{T}_{MC}(\mathcal{C},\delta)$	$\left(612\mathcal{C}\pi_{\max}\pi_{\min}^{-2}\log(\frac{2s}{\delta})\right)^2$
$\underline{T}_{cl,1}(\rho,\tau)$	$\frac{(1-\rho)^2}{4n^{1.5}\sqrt{s}\tau\bar{\sigma}^4}$
$\underline{T}_{N}(\mathcal{C},\delta,\rho,\tau)$	$\max\{\underline{T}_{MC}(\mathcal{C},\frac{\delta}{2}), \underline{T}_{cl,1}(\rho,\tau)\}$
$\underline{T}_{id}(\mathcal{C},\delta,T,\rho,\tau)$	$\frac{\tau\sqrt{s}\mathcal{C}\log(T)}{\pi_{\min}(1-\rho)}\left((\sqrt{2\log(nT)}+\sqrt{2\log(2/\delta)})^2 + C_{\mathbf{z}}^2\\|\mathbf{B}_{1:s}\\|^2\right)\log\left(\frac{36s(n+p)\mathcal{C}\log(T)}{\delta}\right)(n+p)$
$\underline{T}_{id,N}(L,\delta,T,\rho,\tau)$	$\max\left\{\underline{T}_{id}(\frac{L}{\log(T)},\frac{\delta}{2L},T,\rho,\tau), \underline{T}_{N}(\frac{L}{\log(T)},\frac{\delta}{2L},\rho,\tau)\right\}$
$\underline{T}_{rgt,\bar{\epsilon}}(\delta,T)$	$\mathcal{O}(\frac{\sqrt{s}(n+p)}{\pi_{\min}(1-\bar{\rho})}\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}^{-4}\log(\frac{1}{\delta})\log^4(T))$ $\mathcal{O}(\frac{\sqrt{s}(n+p)}{\pi_{\min}(1-\bar{\rho})}\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}^{-2}\log(\frac{1}{\delta})\log^2(T))$ (when $\mathbf{B}_{1:s}$ is known)
$\underline{T}_{\mathbf{x}_0}(\delta)$	$\frac{1}{\gamma\log(1/\bar{\rho})}\max\{\frac{2}{\log(\gamma)}, \log(\frac{\pi^2\sqrt{ns}\bar{\tau}}{3\delta})\}$
$\underline{T}_{rgt}(\delta,T)$	$\max\left\{\underline{T}_{\mathbf{x}_0}(\delta), \underline{T}_{rgt,\bar{\epsilon}}(\delta,T), \underline{T}_{MC,1}(\delta), \underline{T}_{id,N}(\underline{L},\delta,T,\bar{\rho},\bar{\tau})\right\}$

spectral radius. For the infinite-horizon MJS-LQR($\mathbf{A}_{1:s},\mathbf{B}_{1:s},\mathbf{T},\mathbf{Q}_{1:s},\mathbf{R}_{1:s}$) problem, we let $\mathbf{P}_{1:s}^\star$ denote the solution to cDARE given by (2.35) and $\mathbf{K}_{1:s}^\star$ denotes the optimal controller which can be computed via (2.36) with $\mathbf{P}_{1:s}^\star$. Similarly, we define $\mathbf{L}_{1:s}^\star$ and $\boldsymbol{\mathcal{L}}^\star$ to be the corresponding closed-loop state matrix and augmented closed-loop state matrix respectively and $\rho^\star := \rho(\boldsymbol{\mathcal{L}}^\star)$. $\pi_{\max}$ and $\pi_{\min}$ are the largest and smallest elements in the stationary distribution of the ergodic Markov matrix $\mathbf{T}$. For the definition of $\underline{T}_{rgt,\bar{\epsilon}}(\delta,T)$, notation $\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}$ is defined in Table C.2. As a slight abuse of notation, T in $\underline{T}_{rgt,\bar{\epsilon}}(\delta,T)$ (as well as Table A.1) and $\mathcal{C}$ are merely arguments to be replaced with specific quantities depending on the context.

Table C.2 lists the notations related to infinite-horizon MJS perturbation results closely following the notations in Table 4.1. It provides several sensitivity parameters, e.g., how the optimal controller $\mathbf{K}_{1:s}^\star$ varies with perturbations in the MJS parameters $\mathbf{A}_{1:s},\mathbf{B}_{1:s}$, and $\mathbf{T}$ and how the MJS-LQR cost J varies with the controller $\mathbf{K}_{1:s}$. It also provides certain upper

Table C.2: Notations — MJS-LQR Perturbation

ξ	$\min\{\lVert\mathbf{B}_{1:s}\rVert_+^{-2}\lVert\mathbf{R}_{1:s}^{-1}\rVert_+^{-1}\lVert\mathbf{L}_{1:s}^\star\rVert_+^{-2}, \underline{\sigma}(\mathbf{P}_{1:s}^\star)\}$
ξ'	$\lVert\mathbf{A}_{1:s}\rVert_+^2\lVert\mathbf{B}_{1:s}\rVert_+^4\lVert\mathbf{P}_{1:s}^\star\rVert_+^3\lVert\mathbf{R}_{1:s}^{-1}\rVert_+^2$
$\Gamma_\star$	$\max\{\lVert\mathbf{A}_{1:s}\rVert_+, \lVert\mathbf{B}_{1:s}\rVert_+, \lVert\mathbf{P}_{1:s}^\star\rVert_+, \lVert\mathbf{K}_{1:s}^\star\rVert_+\}$
$C_{\mathbf{A},\mathbf{B},\mathbf{T}}^{\mathbf{K}}$	$28\sqrt{ns}\tau(\boldsymbol{\mathcal{L}}^\star)(1-\rho^\star)^{-1}\left(\underline{\sigma}(\mathbf{R}_{1:s})^{-1}+\Gamma_\star^3\underline{\sigma}(\mathbf{R}_{1:s})^{-2}\right)\Gamma_\star^3\xi'$
$C_{\mathbf{K}}^{J}$	$2s^{1.5}\sqrt{n}\min\{n,p\}(\lVert\mathbf{R}_{1:s}\rVert+\Gamma_\star^3)\frac{\tau(\boldsymbol{\mathcal{L}}^\star)}{1-\rho^\star}$
$\bar{\epsilon}_{\mathbf{K}}$	$\min\left\{\lVert\mathbf{K}_{1:s}^\star\rVert, \frac{1-\rho^\star}{2\sqrt{s}\tau(\boldsymbol{\mathcal{L}}^\star)(1+2\lVert\mathbf{L}_{1:s}^\star\rVert)\lVert\mathbf{B}_{1:s}\rVert)}\right\}$
$\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}^{LQR}$	$\frac{(1-\rho^\star)\min\{\Gamma_\star,\underline{\sigma}(\mathbf{R}_{1:s})^2\bar{\epsilon}_{\mathbf{K}}\}}{28\sqrt{ns}\tau(\boldsymbol{\mathcal{L}}^\star)\Gamma_\star^3(\underline{\sigma}(\mathbf{R}_{1:s})+\Gamma_\star^3)}\xi'^{-1}$
$\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}$	$\min\left\{\frac{\xi(1-\rho^\star)^2}{204ns\tau(\boldsymbol{\mathcal{L}}^\star)^2\xi'}, \lVert\mathbf{B}_{1:s}\rVert,\ \underline{\sigma}(\mathbf{Q}_{1:s}),\ \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}^{LQR}\right\}$

bounds on the variations in $\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}$, and $\mathbf{K}_{1:s}$ such that the perturbation theory holds. In this table, $\mathbf{R}_{1:s}^{-1} := \{\mathbf{R}_i^{-1}\}_{i=1}^s$ and recall $\lVert\cdot\rVert_+ := \lVert\cdot\rVert + 1$.

C.2 MJS-LQR Perturbation Results

We first restate Lemma B.7 on the perturbation of augmented closed-loop state matrix if we use a controller $\mathbf{K}_{1:s}$ that is close to the optimal $\mathbf{K}_{1:s}^\star$.

Lemma C.1. *For an arbitrary controller $\mathbf{K}_{1:s}$, let $\mathbf{L}_i = \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i$ for all $i \in [s]$, and let $\boldsymbol{\mathcal{L}}$ be the augmented state matrix such that its (i,j)-th block is given by $[\boldsymbol{\mathcal{L}}]_{ij} := \mathbf{T}(j,i)\mathbf{L}_j \otimes \mathbf{L}_j$. Assume $\lVert\mathbf{K}_{1:s} - \mathbf{K}_{1:s}^\star\rVert \le \bar{\epsilon}_{\mathbf{K}}$, where $\bar{\epsilon}_{\mathbf{K}}$ is defined in Table C.2. Then, we have*

$$\lVert\boldsymbol{\mathcal{L}}^k\rVert \le \tau(\boldsymbol{\mathcal{L}}^\star)(\frac{1+\rho^\star}{2})^k, \quad \forall k \in \mathbb{N}, \tag{C.5}$$

$$\rho(\boldsymbol{\mathcal{L}}) \le \frac{1+\rho^\star}{2}. \tag{C.6}$$

Thus controller $\mathbf{K}_{1:s}$ is stabilizing.

The following perturbation results show how much the infinite-horizon average cost deviates depending on the deviations from the optimal controller and how much the optimal controller deviates depending on the model accuracy for the MJS-LQR problem. The results follow from Corollary B.9 and Theorem 4.2.

Lemma C.2 (Perturbation Results of Infinite-horizon MJS-LQR)**.** *For the infinite-horizon* MJS-LQR$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ *problem, its optimal controller $\mathbf{K}_{1:s}^\star$, and the optimal cost $J^\star$, we have the following perturbation results. Note that notations $\bar{\epsilon}_{\mathbf{K}}, \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}$, and $C_{\mathbf{A},\mathbf{B},\mathbf{T}}^{\mathbf{K}}$ are defined in Table C.2.*

1. *Suppose we have an arbitrary controller* $\mathbf{K}_{1:s}$ *such that* $\|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\| \leq \bar{\epsilon}_{\mathbf{K}}$. *Then, we have*
$$J(0, \omega_0, \{\mathbf{K}_{1:s}\}) - J^\star \leq C^J_{\mathbf{K}} \|\mathbf{\Sigma}_{\mathbf{w}}\| \|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\|^2. \tag{C.7}$$

2. *Suppose there is an arbitrary* $\mathrm{MJS}(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}})$ *such that* $\epsilon_{\mathbf{A},\mathbf{B}} := \max\{\|\hat{\mathbf{A}}_{1:s} - \mathbf{A}_{1:s}\|, \|\hat{\mathbf{B}}_{1:s} - \mathbf{B}_{1:s}\|\} \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}$, *and* $\epsilon_{\mathbf{T}} := \|\hat{\mathbf{T}} - \mathbf{T}\|_\infty \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}$. *Then, there exists an optimal controller* $\mathbf{K}_{1:s}$ *to the infinite-horizon* $\mathrm{MJS\text{-}LQR}(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ *and it can be computed using* (2.36) *and* (2.35), *and we have*
$$\|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\| \leq C^{\mathbf{K}}_{\mathbf{A},\mathbf{B},\mathbf{T}} (\epsilon_{\mathbf{A},\mathbf{B}} + \epsilon_{\mathbf{T}}). \tag{C.8}$$

By definition of $\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}$, *we see* $\|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\| \leq \bar{\epsilon}_{\mathbf{K}}$, *thus Lemma C.1 is applicable.*

C.3 Single Epoch Regret Analysis

Recall the definitions of $\tilde{\mathbf{B}}_t$ and $\tilde{\mathbf{\Pi}}_t$ in (2.13). Furthermore, we define
$$\tilde{\mathbf{\Pi}}_\infty = \boldsymbol{\pi}_\infty \otimes \mathbf{I}_{n^2}, \quad \tilde{\mathbf{R}}_t = \sum_{i=1}^s \boldsymbol{\pi}_t(i) \mathbf{R}_i. \tag{C.9}$$

For a sequence of matrices $\mathbf{V}_{1:s}$, define the following reshaping mapping
$$\mathcal{H}\left(\begin{bmatrix}\mathbf{V}_1 \\ \vdots \\ \mathbf{V}_s\end{bmatrix}\right) = \begin{bmatrix}\mathbf{vec}(\mathbf{V}_1) \\ \vdots \\ \mathbf{vec}(\mathbf{V}_s)\end{bmatrix}, \tag{C.10}$$

and let $\mathcal{H}^{-1}$ denote the inverse mapping of $\mathcal{H}$. Let
$$\mathbf{M}_i := \mathbf{Q}_i + \mathbf{K}_i^\intercal \mathbf{R}_i \mathbf{K}_i, \qquad \mathbf{M} := [\mathbf{M}_1, \dots, \mathbf{M}_s]. \tag{C.11}$$

We define

$$\begin{aligned} N_{0,t} &= \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t \begin{bmatrix} \mathbf{vec}(\boldsymbol{\Sigma}_1(0)) \\ \vdots \\ \mathbf{vec}(\boldsymbol{\Sigma}_s(0)) \end{bmatrix})), \\ N_{\mathbf{z},1,t} &= \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}((\tilde{\mathbf{B}}_t + \boldsymbol{\mathcal{L}}\tilde{\mathbf{B}}_{t-1} + \cdots + \boldsymbol{\mathcal{L}}^{t-1}\tilde{\mathbf{B}}_1)\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}}))), \\ N_{\mathbf{w},t} &= \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}((\tilde{\boldsymbol{\Pi}}_t + \boldsymbol{\mathcal{L}}\tilde{\boldsymbol{\Pi}}_{t-1} + \cdots + \boldsymbol{\mathcal{L}}^{t-1}\tilde{\boldsymbol{\Pi}}_1)\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}}))), \\ N_{\mathbf{z},2,t} &= \mathbf{tr}(\tilde{\mathbf{R}}_t \boldsymbol{\Sigma}_z), \end{aligned} \tag{C.12}$$

and

$$S_{0,T} = \sum_{t=1}^{T} N_{0,t}, \quad S_{\mathbf{z},1,T} = \sum_{t=1}^{T} N_{\mathbf{z},1,t}, \quad S_{\mathbf{w},T} = \sum_{t=1}^{T} N_{\mathbf{w},t}, \quad S_{\mathbf{z},2,T} = \sum_{t=1}^{T} N_{\mathbf{z},2,t}. \tag{C.13}$$

First, we provide the exact expression for the cumulative cost. It will be used later to analyze the regret.

Lemma C.3 (Cumulative Cost Expression)**.** *For the cost* $J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \boldsymbol{\Sigma}_{\mathbf{z}}\})$ *defined in* (C.1)*, we have*

$$J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \boldsymbol{\Sigma}_{\mathbf{z}}\}) = S_{0,T} + S_{\mathbf{z},1,T} + S_{\mathbf{z},2,T} + S_{\mathbf{w},T}. \tag{C.14}$$

Proof. For the expected cost at time t, we have

$$\begin{aligned} \mathbb{E}[\mathbf{x}_t^\intercal \mathbf{Q}_{\omega_t}\mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega_t}\mathbf{u}_t] &= \sum_{i=1}^{s} \mathbf{tr}\left(\mathbb{E}[\mathbf{Q}_{\omega_t}\mathbf{x}_t\mathbf{x}_t^\intercal \mathbf{1}_{\{\omega_t=i\}}] + \mathbb{E}[\mathbf{R}_{\omega_t}\mathbf{u}_t\mathbf{u}_t^\intercal \mathbf{1}_{\{\omega_t=i\}}]\right) \\ &= \sum_{i=1}^{s} \mathbf{tr}\left((\mathbf{Q}_i + \mathbf{K}_i^\intercal \mathbf{R}_i \mathbf{K}_i)\boldsymbol{\Sigma}_i(t) + \boldsymbol{\pi}_t(i)\mathbf{R}_i\boldsymbol{\Sigma}_{\mathbf{z}}\right) \\ &= \sum_{i=1}^{s} \mathbf{tr}\left(\mathbf{M}_i\boldsymbol{\Sigma}_i(t)\right) + N_{\mathbf{z},2,t}, \end{aligned} \tag{C.15}$$

where the second equality follows since $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$. Now plugging in the dynamics of $\boldsymbol{\Sigma}_i(t)$ in Lemma 2.14, we can conclude the proof. $\square$

Next, before proceeding, we provide several properties regarding the operator $\mathbf{tr}(\mathbf{M}\mathcal{H}(\cdot))$ that shows up in (C.12) and (C.13), which will be used later to evaluate $J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \boldsymbol{\Sigma}_{\mathbf{z}}\}) - TJ(0, \omega_0, \{\mathbf{K}_{1:s}\})$.

Lemma C.4 (Properties of Cost Building Bricks)**.** *For any* $t, t' \in \mathbb{N}$*, we have*

(L1) $\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})) \leq \sqrt{ns}\|\mathbf{M}_{1:s}\|\|\boldsymbol{\mathcal{L}}^t\|\|\mathbf{v}\|$*, where* $\mathbf{v} := [\mathbf{vec}(\mathbf{V}_1)^\intercal, \ldots, \mathbf{vec}(\mathbf{V}_s)^\intercal]^\intercal$ *for some* $\mathbf{V}_{1:s}$ *such that* $\mathbf{V}_i \succeq 0$ *for all* $i \in [s]$*;*

*(*L2) $\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\tilde{\mathbf{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{z}))) \leq n\sqrt{s}\|\mathbf{M}_{1:s}\|\|\boldsymbol{\mathcal{L}}^t\|\|\mathbf{B}_{1:s}\|^2\|\boldsymbol{\Sigma}_\mathbf{z}\|$*;*

*(*L3) $\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\tilde{\boldsymbol{\Pi}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{w}))) \leq n\sqrt{s}\|\mathbf{M}_{1:s}\|\|\boldsymbol{\mathcal{L}}^t\|\|\boldsymbol{\Sigma}_\mathbf{w}\|$*;*

*(*L4) $|\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t(\tilde{\boldsymbol{\Pi}}_{t'}-\tilde{\boldsymbol{\Pi}}_\infty)\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{w})))| \leq \tau_{MC}n\sqrt{s}\|\mathbf{M}_{1:s}\|\|\boldsymbol{\mathcal{L}}^t\|\|\boldsymbol{\Sigma}_\mathbf{w}\|\rho_{MC}^{t'}$*, where* τ_{MC} *are* ρ_{MC} *are given in Definition 2.3, and* $\tilde{\boldsymbol{\Pi}}_\infty$ *is given in* (C.9)

Proof. Let $[\cdot]_i$ denote the i-th sub-block of an $s \times 1$ block matrix. Let $\mathbf{vec}^{-1}$ denote the inverse mapping of $\mathbf{vec}$, i.e., $\mathbf{vec}^{-1}([\mathbf{v}_1^\top, \cdots, \mathbf{v}_r^\top]^\top) = [\mathbf{v}_1, \cdots, \mathbf{v}_r]$ for a set of vectors $\{\mathbf{v}_i\}_{i=1}^r$. It can be easily seen that for any set of matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and $\mathbf{X}$, we have $\mathbf{AXB} = \mathbf{C}$ if and only if $(\mathbf{B}^\top \otimes \mathbf{A})\mathbf{vec}(\mathbf{X}) = \mathbf{vec}(\mathbf{C})$. This together with the definitions of $\tilde{\mathbf{B}}_t$, $\tilde{\boldsymbol{\Pi}}_t$ in (2.13), $\tilde{\boldsymbol{\Pi}}_\infty$, $\tilde{\mathbf{R}}_t$ in (C.9), and $\mathcal{H}()$ in (C.10) yields the following preliminary results

$$[\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})]_i \succeq 0, \tag{C.16a}$$

$$\mathbf{vec}^{-1}([\tilde{\mathbf{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{z})]_i) \succeq 0, \tag{C.16b}$$

$$\mathbf{vec}^{-1}([\tilde{\boldsymbol{\Pi}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{w})]_i) \succeq 0, \tag{C.16c}$$

$$\mathbf{vec}^{-1}([|\tilde{\boldsymbol{\Pi}}_{t'} - \tilde{\boldsymbol{\Pi}}_\infty|\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{w})]_i) \succeq 0, \tag{C.16d}$$

$$|\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t(\tilde{\boldsymbol{\Pi}}_{t'} - \tilde{\boldsymbol{\Pi}}_\infty)\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{w})))| \leq \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t|\tilde{\boldsymbol{\Pi}}_{t'} - \tilde{\boldsymbol{\Pi}}_\infty|\mathbf{vec}(\boldsymbol{\Sigma}_\mathbf{w}))), \tag{C.16e}$$

where $|\cdot|$ here denotes the element-wise absolute value of a matrix. Now, let us consider (L1). We observe that

$$\begin{aligned}\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})) = \mathbf{tr}(\sum_{i=1}^s \mathbf{M}_i[\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})]_i) &\leq \|\mathbf{M}_{1:s}\| \cdot \mathbf{tr}(\sum_{i=1}^s[\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})]_i) \\ &\leq \sqrt{n}\|\mathbf{M}_{1:s}\|\|\sum_{i=1}^s[\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})]_i\|_\mathrm{F},\end{aligned} \tag{C.17}$$

where the first inequality uses (C.16a) and the definition that $\|\mathbf{M}_{1:s}\| = \max_{i\in[s]}\|\mathbf{M}_i\|$; and the last inequality follows from Cauchy-Schwarz inequality and the fact that $[\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})]_i \in \mathbb{R}^{n\times n}$. Now, for the last term on the R.H.S. of (C.17), we have

$$\begin{aligned}\|\sum_{i=1}^s[\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})]_i\|_\mathrm{F} \leq \sum_{i=1}^s\|[\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})]_i\|_\mathrm{F} &\leq \sqrt{s}\sqrt{\sum_{i=1}^s\|[\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})]_i\|_\mathrm{F}^2} \\ &= \sqrt{s}\|\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\mathbf{v})\|_\mathrm{F} \\ &= \sqrt{s}\|\boldsymbol{\mathcal{L}}^t\mathbf{v}\| \\ &\leq \sqrt{s}\|\boldsymbol{\mathcal{L}}^t\|\|\mathbf{v}\|,\end{aligned} \tag{C.18}$$

where the second equality holds since $\mathcal{H}^{-1}$ is a reshaping operator, and $\boldsymbol{\mathcal{L}}^t\mathbf{v}$ is a vector. Substituting (C.18) into (C.17) gives (L1).

To show (L2), we combine (C.16b) with (L1) to get

$$\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\tilde{\mathbf{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}}))) \leq \sqrt{ns}\|\mathbf{M}_{1:s}\|\|\boldsymbol{\mathcal{L}}^t\|\|\tilde{\mathbf{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\|.$$

Then, using the upper bound for $\|\tilde{\mathbf{B}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{z}})\|$ derived in (2.24) completes the proof of (L2).

To establish (L3), we combine (C.16c) with (L1) to obtain

$$\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t\tilde{\boldsymbol{\Pi}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}}))) \leq \sqrt{ns}\|\mathbf{M}_{1:s}\|\|\boldsymbol{\mathcal{L}}^t\|\|\tilde{\boldsymbol{\Pi}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\|. \tag{C.19}$$

Then, using the upper bound for $\|\tilde{\boldsymbol{\Pi}}_{t'}\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\|$ derived in (2.25) gives (L2).

Finally, let us consider (L4). It follows from (C.16e) and (C.16d) in conjunction with (L1) that

$$|\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t|\tilde{\boldsymbol{\Pi}}_{t'} - \tilde{\boldsymbol{\Pi}}_{\infty}|\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})))| \leq \sqrt{ns}\|\mathbf{M}_{1:s}\|\|\boldsymbol{\mathcal{L}}^t\|\||\tilde{\boldsymbol{\Pi}}_{t'} - \tilde{\boldsymbol{\Pi}}_{\infty}|\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\|. \tag{C.20}$$

Now, using (C.9), we obtain

$$\begin{aligned}
\||\tilde{\boldsymbol{\Pi}}_{t'} - \tilde{\boldsymbol{\Pi}}_{\infty}|\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\| =& \sqrt{\sum_{i=1}^{s}\||[\tilde{\boldsymbol{\Pi}}_{t'}]_i - [\tilde{\boldsymbol{\Pi}}_{\infty}]_i|\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\|^2} \\
=& \sqrt{\sum_{i=1}^{s}\||\boldsymbol{\pi}_{t'}(i) - \boldsymbol{\pi}_{\infty}(i)|\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\|^2} \\
=& \|\boldsymbol{\pi}_{t'} - \boldsymbol{\pi}_{\infty}\|\|\mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}})\| \\
=& \|\boldsymbol{\pi}_{t'} - \boldsymbol{\pi}_{\infty}\|\|\boldsymbol{\Sigma}_{\mathbf{w}}\|_{\mathrm{F}} \\
\leq& \tau_{MC}\sqrt{n}\|\boldsymbol{\Sigma}_{\mathbf{w}}\|\rho_{MC}^{t'},
\end{aligned}$$

where the last line follows from Lemma 2.4. Substituting the above inequality in (C.20) completes the proof of (L4). □

The following lemma bounds the difference $J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \boldsymbol{\Sigma}_{\mathbf{z}}\}) - TJ(0, \omega_0, \{\mathbf{K}_{1:s}\})$ using an arbitrary stabilizing controller $\mathbf{K}_{1:s}$. Based on this result, we will provide in Proposition C.6 a uniform upper bound for this difference when using any controllers $\mathbf{K}_{1:s}$ that are close to $\mathbf{K}^{\star}_{1:s}$.

Lemma C.5. *For an arbitrary stabilizing controller* $\mathbf{K}_{1:s}$*, we have*

$$\begin{aligned}
&J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \mathbf{\Sigma}_\mathbf{z}\}) - TJ(0, \omega_0, \{\mathbf{K}_{1:s}\}) \\
\leq& \sqrt{ns}\|\mathbf{M}_{1:s}\| \cdot \|\mathbf{x}_0\|^2 + \frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}}{1-\rho_{\boldsymbol{\mathcal{L}}}}\|\mathbf{M}_{1:s}\|\|\mathbf{B}_{1:s}\|^2\|\mathbf{\Sigma}_\mathbf{z}\|T \\
&+ n\|\mathbf{R}_{1:s}\|\|\mathbf{\Sigma}_\mathbf{z}\|T + n\sqrt{s}\tau_{MC}\tau_{\boldsymbol{\mathcal{L}}}\|\mathbf{M}_{1:s}\|\|\mathbf{\Sigma}_\mathbf{w}\|\frac{\rho_{MC}}{\rho_{MC}-\rho_{\boldsymbol{\mathcal{L}}}}\left(\frac{\rho_{MC}}{1-\rho_{MC}} - \frac{\rho_{\boldsymbol{\mathcal{L}}}}{1-\rho_{\boldsymbol{\mathcal{L}}}}\right),
\end{aligned} \tag{C.21}$$

where τ_{MC} *and* ρ_{MC} *are given in Definition 2.3,* $\tau_{\boldsymbol{\mathcal{L}}}$ *and* $\rho_{\boldsymbol{\mathcal{L}}}$ *are constants defined in the beginning of Appendix C, and* $\mathbf{M} = [\mathbf{M}_1, \dots, \mathbf{M}_s]$ *with* $\mathbf{M}_i = \mathbf{Q}_i + \mathbf{K}_i^\intercal \mathbf{R}_i \mathbf{K}_i$*.*

Proof. From Lemma C.3, we know

$$\begin{aligned}
J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \mathbf{\Sigma}_\mathbf{z}\}) &= S_{0,T} + S_{\mathbf{z},1,T} + S_{\mathbf{z},2,T} + S_{\mathbf{w},T}, \\
J(0, \omega_0, \{\mathbf{K}_{1:s}\}) &= \limsup_{T\to\infty} \frac{1}{T}(S_{0,T} + S_{\mathbf{w},T}) =: S_0 + S_\mathbf{w}.
\end{aligned}$$

where $S_0 := \limsup_{T\to\infty} \frac{1}{T} S_{0,T}$ and $S_\mathbf{w} := \limsup_{T\to\infty} \frac{1}{T} S_{\mathbf{w},T}$. Next, we will evaluate each term on the RHSs separately.

For $S_{0,T}$, letting $\mathbf{s}_0 = \begin{bmatrix} \mathbf{vec}(\mathbf{\Sigma}_1(0)) \\ \vdots \\ \mathbf{vec}(\mathbf{\Sigma}_s(0)) \end{bmatrix}$, we have

$$\begin{aligned}
S_{0,T} = \sum_{t=1}^{T} \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^t \mathbf{s}_0)) &\leq \sqrt{ns}\|\mathbf{M}_{1:s}\|\|\boldsymbol{\mathcal{L}}^t\|\|\mathbf{s}_0\| \\
&\leq \sqrt{ns}\|\mathbf{M}_{1:s}\| \cdot \mathbb{E}[\|\mathbf{x}_0\|^2] \\
&= \sqrt{ns}\|\mathbf{M}_{1:s}\| \cdot \|\mathbf{x}_0\|^2,
\end{aligned}$$

where the second line follows from Item (L1) in Lemma C.4; the third line follows from (2.23) in Lemma 2.15. And from the discussion at the beginning of Appendix C, we can get rid of $\mathbb{E}[\cdot]$. Then it is easy to see $S_0 = 0$, as long as $\|\mathbf{x}_0\|^2$ is bounded.

For $S_{\mathbf{z},1,T}$, we have

$$\begin{aligned}
S_{\mathbf{z},1,T} &= \sum_{t=1}^{T}\sum_{t'=0}^{t-1} \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^{t'}\tilde{\mathbf{B}}_{t-t'}\mathbf{vec}(\mathbf{\Sigma}_\mathbf{z}))) \\
&\leq n\sqrt{s}\|\mathbf{M}_{1:s}\|\|\mathbf{B}_{1:s}\|^2\|\mathbf{\Sigma}_\mathbf{z}\|\left(\sum_{t=1}^{T}\sum_{t'=0}^{t-1}\|\boldsymbol{\mathcal{L}}^{t'}\|\right) \\
&\leq \frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}}{1-\rho_{\boldsymbol{\mathcal{L}}}}\|\mathbf{M}_{1:s}\|\|\mathbf{B}_{1:s}\|^2\|\mathbf{\Sigma}_\mathbf{z}\|T,
\end{aligned} \tag{C.22}$$

where the first inequality follows from Item (L2) in Lemma C.4, and the second inequality follows from the fact $\|\boldsymbol{\mathcal{L}}^{t'}\| \leq \tau_{\boldsymbol{\mathcal{L}}} \rho_{\boldsymbol{\mathcal{L}}}^{t'}$.

For $S_{\mathbf{z},2,T}$, we have

$$S_{\mathbf{z},2,T} = \sum_{t=1}^{T} \mathbf{tr}(\sum_{i=1}^{s} \boldsymbol{\pi}_t(i) \mathbf{R}_i \boldsymbol{\Sigma}_{\mathbf{z}}) \leq n \|\mathbf{R}_{1:s}\| \|\boldsymbol{\Sigma}_{\mathbf{z}}\| T. \tag{C.23}$$

For $S_{\mathbf{w},T}$, we have

$$S_{\mathbf{w},T} = \sum_{t=1}^{T} \sum_{t'=0}^{t-1} \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^{t'} \tilde{\boldsymbol{\Pi}}_{t-t'} \mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}}))). \tag{C.24}$$

To evaluate it, we first define the following terms:

$$S_{\mathbf{w},T}^{(\infty)} := \sum_{t=1}^{T} \sum_{t'=0}^{t-1} \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^{t'} \tilde{\boldsymbol{\Pi}}_{\infty} \mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}}))), \tag{C.25}$$

$$S_{\mathbf{w}}^{(\infty)} := \limsup_{T \to \infty} \frac{1}{T} S_{\mathbf{w},T}^{(\infty)}, \tag{C.26}$$

where $\tilde{\boldsymbol{\Pi}}_{\infty}$ is defined in (C.9). Note that $S_{\mathbf{w},T}^{(\infty)}$ and $S_{\mathbf{w}}^{(\infty)}$ are the counterparts of $S_{\mathbf{w},T}$ and $S_{\mathbf{w}}$ except that the initial mode distribution $\boldsymbol{\pi}_0$ is the stationary distribution $\boldsymbol{\pi}_{\infty}$. Then, we have

$$\begin{aligned}
|S_{\mathbf{w},T} - S_{\mathbf{w},T}^{(\infty)}| &= \left| \sum_{t=1}^{T} \sum_{t'=0}^{t-1} \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^{t'} (\tilde{\boldsymbol{\Pi}}_{t-t'} - \tilde{\boldsymbol{\Pi}}_{\infty}) \mathbf{vec}(\boldsymbol{\Sigma}_{\mathbf{w}}))) \right| \\
&\leq \tau_{MC} n \sqrt{s} \|\mathbf{M}_{1:s}\| \|\boldsymbol{\Sigma}_{\mathbf{w}}\| (\sum_{t=1}^{T} \sum_{t'=0}^{t-1} \|\boldsymbol{\mathcal{L}}^{t'}\| \rho_{MC}^{t-t'}) \\
&\leq \tau_{MC} n \sqrt{s} \|\mathbf{M}_{1:s}\| \|\boldsymbol{\Sigma}_{\mathbf{w}}\| (\sum_{t=1}^{\infty} \sum_{t'=0}^{t-1} \tau_{\boldsymbol{\mathcal{L}}} \rho_{\boldsymbol{\mathcal{L}}}^{t'} \rho_{MC}^{t-t'}) \\
&\leq n \sqrt{s} \tau_{MC} \tau_{\boldsymbol{\mathcal{L}}} \|\mathbf{M}_{1:s}\| \|\boldsymbol{\Sigma}_{\mathbf{w}}\| \frac{\rho_{MC}}{\rho_{MC} - \rho_{\boldsymbol{\mathcal{L}}}} (\frac{\rho_{MC}}{1 - \rho_{MC}} - \frac{\rho_{\boldsymbol{\mathcal{L}}}}{1 - \rho_{\boldsymbol{\mathcal{L}}}})
\end{aligned} \tag{C.27}$$

where the first inequality follows from Item (L4) in Lemma C.4. Thus,

$$S_{\mathbf{w}} = \limsup_{T \to \infty} \frac{1}{T} S_{\mathbf{w},T} = \limsup_{T \to \infty} \frac{1}{T} (S_{\mathbf{w},T} - S_{\mathbf{w},T}^{(\infty)}) + \limsup_{T \to \infty} \frac{1}{T} S_{\mathbf{w},T}^{(\infty)} = S_{\mathbf{w}}^{(\infty)}. \tag{C.28}$$

Since $\sum_{t=1}^{T}\sum_{t'=0}^{t-1}\boldsymbol{\mathcal{L}}^{t'} = (\mathbf{I}-\boldsymbol{\mathcal{L}})^{-1}T - (\mathbf{I}-\boldsymbol{\mathcal{L}})^{-2}\boldsymbol{\mathcal{L}}(\mathbf{I}-\boldsymbol{\mathcal{L}}^{T})$ and $\sum_{t'=0}^{\infty}\boldsymbol{\mathcal{L}}^{t'} = (\mathbf{I}-\boldsymbol{\mathcal{L}})^{-1}$ we have

$$\begin{aligned} S_{\mathbf{w}} = S_{\mathbf{w}}^{(\infty)} &= \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\limsup_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sum_{t'=0}^{t-1}\boldsymbol{\mathcal{L}}^{t'}\tilde{\mathbf{\Pi}}_{\infty}\mathbf{vec}(\mathbf{\Sigma_w}))) \\ &= \mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}((\mathbf{I}-\boldsymbol{\mathcal{L}})^{-1}\tilde{\mathbf{\Pi}}_{\infty}\mathbf{vec}(\mathbf{\Sigma_w}))) \\ &= \sum_{t'=0}^{\infty}\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^{t'}\tilde{\mathbf{\Pi}}_{\infty}\mathbf{vec}(\mathbf{\Sigma_w}))). \end{aligned} \tag{C.29}$$

Thus,

$$\begin{aligned} TS_{\mathbf{w}} = TS_{\mathbf{w}}^{(\infty)} &= \sum_{t=1}^{T}\sum_{t'=0}^{\infty}\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^{t'}\tilde{\mathbf{\Pi}}_{\infty}\mathbf{vec}(\mathbf{\Sigma_w}))) \\ &\geq \sum_{t=1}^{T}\sum_{t'=0}^{t-1}\mathbf{tr}(\mathbf{M}\mathcal{H}^{-1}(\boldsymbol{\mathcal{L}}^{t'}\tilde{\mathbf{\Pi}}_{\infty}\mathbf{vec}(\mathbf{\Sigma_w}))) \\ &= S_{\mathbf{w},T}^{(\infty)} \end{aligned} \tag{C.30}$$

where the inequality holds since each trace summand is non-negative. Therefore,

$$\begin{aligned} S_{\mathbf{w},T} &\leq S_{\mathbf{w},T}^{(\infty)} + |S_{\mathbf{w},T} - S_{\mathbf{w},T}^{(\infty)}| \\ &\overset{(C.28)}{\leq} TS_{\mathbf{w}} + |S_{\mathbf{w},T} - S_{\mathbf{w},T}^{(\infty)}| \\ &\overset{(C.27)}{\leq} TS_{\mathbf{w}} + n\sqrt{s}\tau_{MC}\tau_{\boldsymbol{\mathcal{L}}}\|\mathbf{M}_{1:s}\|\|\mathbf{\Sigma_w}\|\frac{\rho_{MC}}{\rho_{MC}-\rho_{\boldsymbol{\mathcal{L}}}}(\frac{\rho_{MC}}{1-\rho_{MC}} - \frac{\rho_{\boldsymbol{\mathcal{L}}}}{1-\rho_{\boldsymbol{\mathcal{L}}}}). \end{aligned} \tag{C.31}$$

Finally, combining all the results we have so far, we have

$$\begin{aligned} & J_T(\mathbf{x}_0,\omega_0,\{\mathbf{K}_{1:s},\mathbf{\Sigma_z}\}) - TJ(0,\omega_0,\{\mathbf{K}_{1:s}\}) \\ =& S_{0,T} + S_{\mathbf{z},1,T} + S_{\mathbf{z},2,T} + S_{\mathbf{w},T} - T(S_0 + S_{\mathbf{w}}) \\ \leq& \sqrt{ns}\|\mathbf{M}_{1:s}\|\cdot\|\mathbf{x}_0\|^2 \\ &+ \frac{n\sqrt{s}\tau_{\boldsymbol{\mathcal{L}}}}{1-\rho_{\boldsymbol{\mathcal{L}}}}\|\mathbf{M}_{1:s}\|\|\mathbf{B}_{1:s}\|^2\|\mathbf{\Sigma_z}\|T \\ &+ n\|\mathbf{R}_{1:s}\|\|\mathbf{\Sigma_z}\|T \\ &+ n\sqrt{s}\tau_{MC}\tau_{\boldsymbol{\mathcal{L}}}\|\mathbf{M}_{1:s}\|\|\mathbf{\Sigma_w}\|\frac{\rho_{MC}}{\rho_{MC}-\rho_{\boldsymbol{\mathcal{L}}}}(\frac{\rho_{MC}}{1-\rho_{MC}} - \frac{\rho_{\boldsymbol{\mathcal{L}}}}{1-\rho_{\boldsymbol{\mathcal{L}}}}) \end{aligned} \tag{C.32}$$

which concludes the proof. □

We now provide a uniform upper bound on the regret $J_T(\mathbf{x}_0,\omega_0,\{\mathbf{K}_{1:s},\mathbf{\Sigma_z}\}) - TJ^\star$ for any stabilizing controller $\mathbf{K}_{1:s}$ that is close enough to the optimal controller $\mathbf{K}_{1:s}^\star$.

Proposition C.6. *For every* $\mathbf{K}_{1:s}$ *such that* $\|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\| \leq \bar{\epsilon}_\mathbf{K}$, *we have*

$$
\begin{aligned}
J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \mathbf{\Sigma_z}\}) - TJ^\star \leq & C^J_\mathbf{K} \|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\|^2 \|\mathbf{\Sigma_w}\| T \\
& + \sqrt{ns} M \|\mathbf{x}_0\|^2 \\
& + n\sqrt{s} \frac{2\tau(\boldsymbol{\mathcal{L}}^\star)\|\mathbf{B}_{1:s}\|^2 M}{1-\rho^\star} \|\mathbf{\Sigma_z}\| T \\
& + n\|\mathbf{R}_{1:s}\|\|\mathbf{\Sigma_z}\| T \\
& + n\sqrt{s} \frac{2\tau(\boldsymbol{\mathcal{L}}^\star)\tau_{MC} M \rho_{MC}}{2\rho_{MC} - 1 - \rho^\star} \left(\frac{\rho_{MC}}{1-\rho_{MC}} - \frac{1+\rho^\star}{1-\rho^\star}\right) \|\mathbf{\Sigma_w}\|,
\end{aligned}
\tag{C.33}
$$

where $M := \|\mathbf{Q}_{1:s}\| + 4\|\mathbf{R}_{1:s}\|\|\mathbf{K}^\star_{1:s}\|^2$, *and* $\bar{\epsilon}_\mathbf{K}$ *and* $C^J_\mathbf{K}$ *are defined in Table C.2.*

Proof. When $\|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\| \leq \bar{\epsilon}_\mathbf{K}$, from Lemma C.1, we know $\|\boldsymbol{\mathcal{L}}^k\| \leq \tau(\boldsymbol{\mathcal{L}}^\star)(\frac{1+\rho^\star}{2})^k$, thus we could set $\tau_{\boldsymbol{\mathcal{L}}}$ and $\rho_{\boldsymbol{\mathcal{L}}}$ to be $\tau(\boldsymbol{\mathcal{L}}^\star)$ and $\frac{1+\rho^\star}{2}$. By definition, we know $\bar{\epsilon}_\mathbf{K} \leq \|\mathbf{K}^\star_{1:s}\|$, thus $\|\mathbf{M}_{1:s}\| \leq \|\mathbf{Q}_{1:s}\| + \|\mathbf{R}_{1:s}\|\|\mathbf{K}_{1:s}\|^2 \leq \|\mathbf{Q}_{1:s}\| + \|\mathbf{R}_{1:s}\|(\|\mathbf{K}^\star_{1:s}\| + \bar{\epsilon}_\mathbf{K})^2 \leq \|\mathbf{Q}_{1:s}\| + 4\|\mathbf{R}_{1:s}\|\|\mathbf{K}^\star_{1:s}\|^2 = M$. Then applying Lemma C.5, we have

$$
\begin{aligned}
& J_T(\mathbf{x}_0, \omega_0, \{\mathbf{K}_{1:s}, \mathbf{\Sigma_z}\}) - TJ(0, \omega_0, \{\mathbf{K}_{1:s}\}) \\
\leq & \sqrt{ns} M \|\mathbf{x}_0\|^2 \\
& + n\sqrt{s} \frac{2\tau(\boldsymbol{\mathcal{L}}^\star)\|\mathbf{B}_{1:s}\|^2 M}{1-\rho^\star} \|\mathbf{\Sigma_z}\| T \\
& + n\|\mathbf{R}_{1:s}\|\|\mathbf{\Sigma_z}\| T \\
& + n\sqrt{s} \frac{2\tau(\boldsymbol{\mathcal{L}}^\star)\tau_{MC} M \rho_{MC}}{2\rho_{MC} - 1 - \rho^\star} \left(\frac{\rho_{MC}}{1-\rho_{MC}} - \frac{1+\rho^\star}{1-\rho^\star}\right) \|\mathbf{\Sigma_w}\|
\end{aligned}
\tag{C.34}
$$

Now note that when $\|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\| \leq \bar{\epsilon}_\mathbf{K}$, we have $J(0, \omega_0, \{\mathbf{K}_{1:s}\}) - J^\star \leq C^J_\mathbf{K}\|\mathbf{\Sigma_w}\|\|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\|^2$ using Lemma C.2. Combining this with (C.34), we could conclude the proof. □

C.4 Stitching Every Epoch

In this section, we stitch the upper bounds on Regret_q for every epoch q and build a bound on the overall regret $\text{Regret}(T)$.

We define the estimation error after epoch q as $\epsilon^{(q)}_{\mathbf{A},\mathbf{B}} = \max\{\|\mathbf{A}^{(q)}_{1:s} - \mathbf{A}_{1:s}\|, \|\mathbf{B}^{(q)}_{1:s} - \mathbf{B}_{1:s}\|\}$, $\epsilon^{(q)}_\mathbf{T} = \|\mathbf{T}^{(q)} - \mathbf{T}\|_\infty$. Furthermore, we also define $\epsilon^{(q)}_\mathbf{K} := \|\mathbf{K}^{(q)}_{1:s} - \mathbf{K}^\star_{1:s}\|$ where $\mathbf{K}^\star_{1:s}$ is the optimal controller for the infinite-horizon MJS-LQR($\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s}$). We define the following events for every epoch q.

$$\begin{aligned}
\mathcal{A}_q &= \left\{ \text{Regret}_q \leq \mathcal{O}\left(sp \left(\epsilon_{\mathbf{A},\mathbf{B}}^{(q-1)} + \epsilon_{\mathbf{T}}^{(q-1)} \right)^2 \sigma_{\mathbf{w}}^2 T_q + \sqrt{ns} \|\mathbf{x}_0^{(q)}\|^2 + \frac{n\sqrt{s}}{1-\rho^\star} \sigma_{\mathbf{z},q}^2 T_q + c_{\mathcal{A}} \right) \right\} \\
\mathcal{B}_q &= \left\{ \epsilon_{\mathbf{A},\mathbf{B}}^{(q)} \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}, \epsilon_{\mathbf{T}}^{(q)} \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}, \epsilon_{\mathbf{K}}^{(q+1)} \leq \bar{\epsilon}_{\mathbf{K}} \right\} \\
\mathcal{C}_q &= \left\{ \epsilon_{\mathbf{A},\mathbf{B}}^{(q)} \leq \mathcal{O}\left(\log(\frac{1}{\delta_{id,q}}) \frac{\sigma_{\mathbf{z},q} + \sigma_{\mathbf{w}}}{\sigma_{\mathbf{z},q} \pi_{\min}} \frac{\sqrt{s}(n+p)\log(T_q)}{(1-\bar{\rho})\sqrt{T_q}} \right), \right. \\
&\qquad \left. \epsilon_{\mathbf{T}}^{(q)} \leq \mathcal{O}\left(\log(\frac{1}{\delta_{id,q}}) \frac{1}{\pi_{\min}} \sqrt{\frac{\log(T_q)}{T_q}} \right) \right\} \\
\mathcal{D}_q &= \left\{ \|\mathbf{x}_0^{(q+1)}\|^2 = \|\mathbf{x}_{T_q}^{(q)}\|^2 \leq \frac{\bar{x}_0^2}{\delta_{\mathbf{x}_0,q}} \right\}.
\end{aligned} \tag{C.35}$$

where $c_{\mathcal{A}}, \bar{x}_0$ are constants, $\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}$ and $\bar{\epsilon}_{\mathbf{K}}$ are defined in Table C.2, and $\delta_{id,q}$ and $\delta_{\mathbf{x}_0,q}$ within $[0,1]$ denotes the failure probability for event $\mathcal{C}_q$ and $\mathcal{D}_q$. Note that $\mathcal{O}(\cdot)$ hides terms that are invariant to epochs such as $\rho^\star$, $\|\mathbf{A}_{1:s}\|$, $\|\mathbf{B}_{1:s}\|$, etc.

Event $\mathcal{A}_q$ describes how epoch q regret depends on initial state $\|\mathbf{x}_0^{(q)}\|^2$, exploration noise variance $\sigma_{\mathbf{z},q}^2$, and the accuracy of the estimated MJS dynamics $\mathbf{A}_{1:s}^{(q-1)}, \mathbf{B}_{1:s}^{(q-1)}, \hat{\mathbf{T}}$ after epoch $q-1$, which is used to computed epoch q controller $\mathbf{K}_{1:s}^{(q)}$. Event $\mathcal{B}_q$ indicates whether the estimated dynamics and resulting controllers are good enough. $\mathcal{C}_q$ describes the dynamics estimation error after epoch q, and when epoch T_q is chosen appropriately, $\mathcal{B}_q$ can be implied. Lastly, event $\mathcal{D}_q$ bounds the initial state of each epoch, as the initial state plays a vital role in regret upper bound $\mathcal{A}_q$. We see events $\mathcal{A}_{q+1}, \mathcal{B}_q, \mathcal{C}_q, \mathcal{D}_q$ are $\mathcal{F}_q$-measurable, i.e. these events can be determined using random variables $\mathbf{x}_0, \mathbf{w}_t, \mathbf{z}_t, \omega_t$ up to epoch q. Let $\mathcal{E}_q := \mathcal{A}_{q+1} \cap \mathcal{B}_q \cap \mathcal{C}_q \cap \mathcal{D}_q$. Note that even though $\mathcal{A}_{q+1}$ is for the conditional expected regret of the epoch $q+1$ with randomness coming from $\mathbf{x}_0^{(q+1)} = \mathbf{x}_{T_q}^{(q)}, \omega^{(q+1)}(0) = \omega^{(q)}(T_q)$, and controller $\mathbf{K}_{1:s}^{(q+1)}$ computed from $\mathbf{A}_{1:s}^{(q)}, \mathbf{B}_{1:s}^{(q)}, \mathbf{T}^{(q)}$, thus $\mathcal{A}_{q+1}$ is $\mathcal{F}_q$-measurable.

Then, we have the following results regarding the conditional probabilities of these events. First, Proposition C.7 says given the event $\mathcal{B}_{q-1}$ (a good controller is applied during epoch q) and event $\mathcal{D}_q$ (the initial state of epoch q, $\mathbf{x}_0^{(q)}$ is bounded), then $\mathcal{D}_q$ could occur, i.e. $\mathbf{x}_{T_q}^{(q)}$, the final state of epoch q, a.k.a. $\mathbf{x}_0^{(q+1)}$ the initial state of epoch $q+1$, is also bounded.

Proposition C.7. *Suppose* $\frac{\sqrt{ns}\bar{\tau}\bar{\rho}^{T_q}}{\delta_{\mathbf{x}_0,q-1}} < 1$ *and* $\bar{x}_0^2 \geq \frac{n\sqrt{s}(\|\mathbf{B}_{1:s}\|^2+1)\sigma_{\mathbf{w}}^2\bar{\tau}}{(1-\bar{\rho})(1-\sqrt{ns}\cdot\bar{\tau}\bar{\rho}^{T_q}/\delta_{\mathbf{x}_0,q-1})}$ for $i \geq 1$. *Then,*

$$\mathbb{P}(\mathcal{D}_q \mid \cap_{j=0}^{q-1} \mathcal{E}_j) = \mathbb{P}(\mathcal{D}_q \mid \mathcal{B}_{q-1}\mathcal{D}_{q-1}) > 1 - \delta_{\mathbf{x}_0,q},$$

and $\mathbb{P}(\mathcal{D}_0) \geq 1 - \delta_{\mathbf{x}_0,0}$.

Proof. For epoch $q = 1, 2, \ldots$, given event $\mathcal{B}_{q-1}$, we know $\epsilon_{\mathbf{K}}^{(q+1)} \leq \bar{\epsilon}_{\mathbf{K}}$. Let $\boldsymbol{\mathcal{L}}^{(q)}$ denote the augmented closed-loop state matrix. By Lemma C.1, we know $\|(\boldsymbol{\mathcal{L}}^{(q)})^k\| \leq \tau(\boldsymbol{\mathcal{L}}^\star)(\frac{1+\rho^\star}{2})^k$. Thus, if we pick $\bar{\tau} := \max\{\tau(\boldsymbol{\mathcal{L}}^{(0)}), \tau(\boldsymbol{\mathcal{L}}^\star)\}, \bar{\rho} := \max\{\rho(\boldsymbol{\mathcal{L}}^{(0)}), \frac{1+\rho^\star}{2}\}$, this can be generalized to $q = 0$ case, i.e. for every epoch $q = 0, 1, 2, \ldots$, we have $\|(\boldsymbol{\mathcal{L}}^{(q)})^k\| \leq \bar{\tau}\bar{\rho}^k$.

For $q = 1, 2, \ldots$, event $\mathcal{D}_{q-1}$ implies $\|\mathbf{x}_0^{(q)}\|^2 \leq \frac{\bar{x}_0^2}{\delta_{\mathbf{x}_0, q-1}}$. Then, according to Lemma 2.15, we know

$$\begin{aligned}
\mathbb{E}[\|\mathbf{x}_{T_q}^{(q)}\|^2 \mid \mathcal{B}_{q-1}, \mathcal{D}_{q-1}] &\leq \sqrt{ns} \cdot \bar{\tau}\bar{\rho}^{T_q} \frac{\bar{x}_0^2}{\delta_{\mathbf{x}_0, q-1}} + n\sqrt{s}(\|\mathbf{B}_{1:s}\|^2 \frac{\sigma_{\mathbf{w}}^2}{\sqrt{T_q}} + \sigma_{\mathbf{w}}^2) \frac{\bar{\tau}}{1 - \bar{\rho}}. \\
&\leq \frac{\sqrt{ns} \cdot \bar{\tau}\bar{\rho}^{T_q}}{\delta_{\mathbf{x}_0, q-1}} \bar{x}_0^2 + (1 - \frac{\sqrt{ns}\bar{\tau}\bar{\rho}^{T_q}}{\delta_{\mathbf{x}_0, q-1}}) \bar{x}_0^2 \\
&\leq \bar{x}_0^2,
\end{aligned} \tag{C.36}$$

where the second line follows from the assumptions in the proposition statement. Using Markov inequality, we have

$$\mathbb{P}(\|\mathbf{x}_{T_q}^{(q)}\|^2 \leq \frac{\bar{x}_0^2}{\delta_{\mathbf{x}_0, q}} \mid \mathcal{B}_{q-1}, \mathcal{D}_{q-1}) \geq 1 - \delta_{\mathbf{x}_0, q},$$

which implies $\mathbb{P}(\mathcal{D}_q \mid \mathcal{B}_{q-1}, \mathcal{D}_{q-1}) \geq 1 - \delta_{\mathbf{x}_0, q}$. For $q = 0$, similarly, we have $\mathbb{E}[\|\mathbf{x}_{T_0}^{(0)}\|^2] \leq n\sqrt{s}(\|\mathbf{B}_{1:s}\|^2 \frac{\sigma_{\mathbf{w}}^2}{\sqrt{T_q}} + \sigma_{\mathbf{w}}^2) \frac{\bar{\tau}}{1-\bar{\rho}} \leq \bar{x}_0^2$, thus $\mathbb{P}(\mathcal{D}_0) \geq 1 - \delta_{\mathbf{x}_0, q}$.

Finally, note that given a good stabilizing controller (event $\mathcal{B}_{q-1}$) and a bounded initial state (event $\mathcal{D}_{q-1}$) for epoch q, the final state of epoch q only depends on randomness in epoch q, thus $\mathbb{P}(\mathcal{D}_q \mid \cap_{j=0}^{q-1} \mathcal{E}_j) = \mathbb{P}(\mathcal{D}_q \mid \mathcal{B}_{q-1} \mathcal{D}_{q-1})$. □

Proposition C.8 describes that given the event $\mathcal{C}_q$ (the estimated MJS dynamics after epoch q has estimation errors decays with T_q), when epoch q has length T_q large enough, then the event $\mathcal{B}_q$ (the estimated dynamics and controllers computed with it will be good enough) occurs.

Proposition C.8. *Suppose every epoch q has length $T_q \geq \underline{T}_{rgt, \bar{\epsilon}}(\delta_{id,q}, T_q)$. Then,*

$$\mathbb{P}(\mathcal{B}_q \mid \mathcal{C}_q, \cap_{j=0}^{q-1} \mathcal{E}_j) = \mathbb{P}(\mathcal{B}_q \mid \mathcal{C}_q) = 1$$

Proof. When $\mathcal{C}_q$ occurs, since $\sigma_{\mathbf{z},q}^2 = \frac{\sigma_{\mathbf{w}}^2}{\sqrt{T_q}}$, we have

$$\epsilon_{\mathbf{A},\mathbf{B}}^{(q)} \leq \mathcal{O}(\log(\frac{1}{\delta_{id,q}}) \frac{\sqrt{s}(n+p)}{\pi_{\min}} \frac{\log(T_q)}{(1 - \bar{\rho}) T_q^{0.25}}), \epsilon_{\mathbf{T}}^{(q)} \leq \mathcal{O}(\log(\frac{1}{\delta_{id,q}}) \frac{1}{\pi_{\min}} \sqrt{\frac{\log(T_q)}{T_q}}).$$

We know when $T_q \geq \mathcal{O}(\frac{\sqrt{s}(n+p)}{\pi_{\min}(1-\bar{\rho})}\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}^{-4}\log(\frac{1}{\delta_{id,q}})\log^4(T_q)) =: \underline{T}_{rgt,\bar{\epsilon}}(\delta_{id,q}, T_q)$, we have $\epsilon_{\mathbf{A},\mathbf{B}}^{(q)} \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}, \epsilon_{\mathbf{T}}^{(q)} \leq \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}$. Then according to Lemma C.2, we have $\epsilon_{\mathbf{K}}^{(q+1)} \leq \bar{\epsilon}_{\mathbf{K}}$. Thus $\mathbb{P}(\mathcal{B}_q \mid \mathcal{C}_q) = 1$. Finally, note that given the estimation error sample complexity in $\mathcal{C}_q$ for epoch q, events happen before epoch q does not influence $\mathcal{B}_q$, so $\mathbb{P}(\mathcal{B}_q \mid \mathcal{C}_q, \cap_{j=0}^{q-1}\mathcal{E}_j) = \mathbb{P}(\mathcal{B}_q \mid \mathcal{C}_q)$=1. □

Next, Proposition C.9 says given the $\mathcal{B}_{q-1}$ (a good controller is used in epoch q), then the event $\mathcal{C}_q$ could occur, i.e. dynamics learned using the trajectory of epoch q, will be accurate enough.

Proposition C.9. *For $c_{\mathbf{x}} \geq \underline{c}_{\mathbf{x}}(\bar{\rho}, \bar{\tau}), c_{\mathbf{z}} \geq \underline{c}_{\mathbf{z}}$, $T_q \geq \max\left\{\underline{T}_{MC,1}(\frac{\delta_{id,q}}{8}), \underline{T}_{id,N}(\frac{\delta_{id,q}}{2}, \bar{\rho}, \bar{\tau})\right\}$, we have for $q = 1, 2, \ldots$,*

$$\mathbb{P}(\mathcal{C}_q \mid \cap_{j=0}^{q-1}\mathcal{E}_j) = \mathbb{P}(\mathcal{C}_q \mid \mathcal{B}_{q-1}) \geq 1 - \delta_{id,q}. \tag{C.37}$$

And $\mathbb{P}(\mathcal{C}_0) \geq 1 - \delta_{id,0}$.

Proof. By Lemma A.6, we know for every epoch $q = 0, 1, \ldots$, when $T_q \geq \underline{T}_{MC,1}(\frac{\delta_{id,q}}{8})$, we have with probability at least $1 - \frac{\delta_{id,q}}{2}$, $\epsilon_{\mathbf{T}}^{(q)} \leq \mathcal{O}\left(\log(\frac{1}{\delta_{id,q}})\frac{1}{\pi_{\min}}\sqrt{\frac{\log(T_q)}{T_q}}\right)$.

For epoch $q = 1, 2, \ldots$, given event $\mathcal{B}_{q-1}$, we know $\epsilon_{\mathbf{K}}^{(q)} \leq \bar{\epsilon}_{\mathbf{K}}$. Let $\boldsymbol{\mathcal{L}}^{(q)}$ denote the augmented closed-loop state matrix. By Lemma C.1, we know $\|(\boldsymbol{\mathcal{L}}^{(q)})^k\| \leq \tau(\boldsymbol{\mathcal{L}}^\star)(\frac{1+\rho^\star}{2})^k$. Thus, if we pick $\bar{\tau} := \max\{\tau(\boldsymbol{\mathcal{L}}^{(0)}), \tau(\boldsymbol{\mathcal{L}}^\star)\}, \bar{\rho} := \max\{\rho(\boldsymbol{\mathcal{L}}^{(0)}), \frac{1+\rho^\star}{2}\}$, this can be generalized to $q = 0$ case, i.e. for every epoch $q = 0, 1, 2, \ldots$, we have $\|(\boldsymbol{\mathcal{L}}^{(q)})^k\| \leq \bar{\tau}\bar{\rho}^k$.

Suppose $c_{\mathbf{x}} \geq \underline{c}_{\mathbf{x}}(\bar{\rho}, \bar{\tau}), c_{\mathbf{z}} \geq \underline{c}_{\mathbf{z}}$, and $T_q \geq \underline{T}_{id,N}(\frac{\delta_{id,q}}{2}, \bar{\rho}, \bar{\tau})$ hold for $q = 0, 1, \ldots$. Then, from Theorem A.22, we know for every $q = 0, 1, \ldots$, with probability at least $1 - \frac{\delta_{id,q}}{2}$, $\epsilon_{\mathbf{A},\mathbf{B}}^{(q)} \leq \mathcal{O}\left(\log(\frac{1}{\delta_{id,q}})\frac{\sigma_{\mathbf{z},q}+\sigma_{\mathbf{w}}}{\sigma_{\mathbf{z},q}\pi_{\min}}\frac{\sqrt{s}(n+p)\log(T_q)}{\sqrt{T_q}}\right)$.

Applying union bound to $\epsilon_{\mathbf{T}}^{(q)}$ and $\epsilon_{\mathbf{A},\mathbf{B}}^{(q)}$, we could show $\mathbb{P}(\mathcal{C}_0) \geq 1 - \delta_{id,q}$ and $\mathbb{P}(\mathcal{C}_q \mid \mathcal{B}_{q-1}, \mathcal{D}_{q-1}) \geq 1 - \delta_{id,q}$. Finally, note that given a good stabilizing controller (event $\mathcal{B}_{q-1}$) and bounded initial state (event $\mathcal{D}_{q-1}$) for epoch q, the estimation error sample complexity (event $\mathcal{C}_q$) does not depend on events happen before epoch q, so $\mathbb{P}(\mathcal{C}_q \mid \cap_{j=0}^{q-1}\mathcal{E}_j)) = \mathbb{P}(\mathcal{C}_q \mid \mathcal{B}_{q-1}, \mathcal{D}_{q-1})$. □

Finally, Proposition C.10 simply describes how the regret of epoch q depends on the accuracy of the estimated dynamics after epoch $q - 1$.

Proposition C.10. *For $\mathcal{A}_q$– $\mathcal{C}_q$ given in (C.35), we have*

$$\mathbb{P}(\mathcal{A}_q \mid \mathcal{B}_{q-1}, \mathcal{C}_{q-1}, \mathcal{D}_{q-1}, \cap_{j=0}^{q-2}\mathcal{E}_j) = \mathbb{P}(\mathcal{A}_q \mid \mathcal{B}_{q-1}) = 1.$$

Proof. From Proposition C.6, we know that for every epoch $q = 1, 2, \dots$, given $\|\mathbf{K}_{1:s}^{(q)} - \mathbf{K}_{1:s}^\star\| \le \bar{\epsilon}_\mathbf{K}$ in $\mathcal{B}_{q-1}$, we have with probability 1

$$
\begin{aligned}
\text{Regret}_q \le & C_\mathbf{K}^J \|\mathbf{K}_{1:s}^{(q)} - \mathbf{K}_{1:s}^\star\|^2 \sigma_\mathbf{w}^2 T_q \\
& + \sqrt{ns} M \|\mathbf{x}_0^{(q)}\|^2 \\
& + n\sqrt{s} \frac{2\tau(\boldsymbol{\mathcal{L}}^\star)\|\mathbf{B}_{1:s}\|^2 M}{1-\rho^\star} \sigma_{\mathbf{z},q}^2 T_q \\
& + n\|\mathbf{R}_{1:s}\| \sigma_{\mathbf{z},q}^2 T_q \\
& + n\sqrt{s} \frac{2\tau(\boldsymbol{\mathcal{L}}^\star)\tau_{MC} M \rho_{MC}}{2\rho_{MC} - 1 - \rho^\star} \left(\frac{\rho_{MC}}{1-\rho_{MC}} - \frac{1+\rho^\star}{1-\rho^\star}\right) \sigma_\mathbf{w}^2.
\end{aligned} \tag{C.38}
$$

Let $c_\mathcal{A}$ denote the last term in (C.38), which is a constant over epochs. Note that from $\epsilon_{\mathbf{A},\mathbf{B}}^{(q-1)} \le \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}, \epsilon_\mathbf{T}^{(q-1)} \le \bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}$ in event $\mathcal{B}_{q-1}$, we know $\|\mathbf{K}_{1:s}^{(q)} - \mathbf{K}_{1:s}^\star\| \le C_{\mathbf{A},\mathbf{B},\mathbf{T}}^\mathbf{K}(\epsilon_{\mathbf{A},\mathbf{B}}^{(q-1)} + \epsilon_\mathbf{T}^{(q-1)})$ by Lemma C.2. Plugging this into (C.38), we have

$$
\text{Regret}_q \le O\left(s \cdot p \left(\epsilon_{\mathbf{A},\mathbf{B}}^{(q-1)} + \epsilon_\mathbf{T}^{(q-1)}\right)^2 \sigma_\mathbf{w}^2 T_q + \sqrt{ns}\|\mathbf{x}_0^{(q)}\|^2 + \frac{n\sqrt{s}}{1-\rho^\star}\sigma_{\mathbf{z},q}^2 T_q + c_\mathcal{A}\right) \tag{C.39}
$$

where term $s \cdot p$ comes from term $s \min\{n, p\}$ in the definition of $C_\mathbf{K}^J$ in Appendix C.2. This shows $\mathbb{P}(\mathcal{A}_q \mid \mathcal{B}_{q-1}) = 1$. Finally, note that given a good controller (event $\mathcal{B}_{q-1}$) for epoch q, the regret for epoch q can be upper bounded (event $\mathcal{A}_q$) without dependence on other events, thus $\mathbb{P}(\mathcal{A}_q \mid \mathcal{B}_{q-1}, \mathcal{C}_{q-1}, \mathcal{D}_{q-1}, \cap_{j=0}^{q-2}\mathcal{E}_j) = \mathbb{P}(\mathcal{A}_q \mid \mathcal{B}_{q-1})$. □

C.4.1 Proof for Theorem 5.1

Theorem C.11 (Complete version of Theorem 5.1)**.** *Assume that the initial state* $\mathbf{x}_0 = 0$*, and Assumption A5.1 holds. Suppose* $c_\mathbf{x} \ge \underline{c}_\mathbf{x}(\bar{\rho}, \bar{\tau}), c_\mathbf{z} \ge \underline{c}_\mathbf{z}$*,* $T_0 \ge \mathcal{O}(\underline{T}_{rgt}(\delta, T_0))$*, and* $\bar{x}_0^2 = \frac{n\sqrt{s}(\|\mathbf{B}_{1:s}\|^2+1)\sigma_\mathbf{w}^2\bar{\tau}}{(1-\bar{\rho})(1-\sqrt{ns}\cdot\bar{\tau}\bar{\rho}^{T_0\gamma}\pi^2/3\delta)}$*. Then, with probability at least* $1-\delta$*, Algorithm 2 achieves*

$$
\text{Regret}(T) \le \mathcal{O}\left(\frac{s^2 p(n^2+p^2)\sigma_\mathbf{w}^2}{\pi_{\min}^2(1-\bar{\rho})^2} \log^2\left(\frac{\log^2(T)}{\delta}\right) \log^2(T)\sqrt{T} + \frac{\sqrt{ns}\log^3(T)}{\delta}\right). \tag{C.40}
$$

Proof. In this proof, we will first show the intersected event $\cap_q \mathcal{E}_q = \cap_q\{\mathcal{A}_{q+1} \cap \mathcal{B}_q \cap \mathcal{C}_q \cap \mathcal{D}_q\}$ implies the desired regret bound, then we evaluate the occurrence probability of $\cap_q \mathcal{E}_q$ using Proposition C.8 to C.10. In the following, we set $\delta_{id,q} = \delta_{\mathbf{x}_0,q} = \frac{3}{\pi^2} \cdot \frac{\delta}{(q+1)^2}$. With the choices $T_q = \gamma T_{q-1}$, $\sigma_{\mathbf{z},q}^2 = \frac{\sigma_\mathbf{w}^2}{\sqrt{T_q}}$, and $\delta_{id,q} = \delta_{\mathbf{x}_0,q} = \frac{3}{\pi^2} \cdot \frac{\delta}{(q+1)^2}$, event $\mathcal{E}_q = \mathcal{A}_{q+1} \cap \mathcal{B}_q \cap \mathcal{C}_q \cap \mathcal{D}_q$ implies the following.

$$\begin{aligned}
\text{Regret}_{q+1} \leq & \mathcal{O}(1)\log^2\left(\frac{(q+1)^2}{\delta}\right) sp\left(\frac{\sigma_{\mathbf{z},q}+\sigma_{\mathbf{w}}}{\sigma_{\mathbf{z},q}\pi_{\min}}\cdot\frac{\sqrt{s}(n+p)\log(T_q)}{(1-\bar{\rho})\sqrt{T_q}}+\frac{\sqrt{\log(T_q)}}{\pi_{\min}\sqrt{T_q}}\right)^2 \sigma_{\mathbf{w}}^2 T_{q+1} \\
& + \mathcal{O}\left(\frac{(q+1)^2}{\delta}\right)\sqrt{ns}\bar{x}_0^2 + \mathcal{O}(\frac{n\sqrt{s}}{1-\rho^\star}\sigma_{\mathbf{z},q+1}^2 T_{q+1}) + \mathcal{O}(1) \\
\leq & \mathcal{O}(1)\log^2\left(\frac{(q+1)^2}{\delta}\right)\frac{s^2p(n^2+p^2)\gamma}{\pi_{\min}^2(1-\bar{\rho})^2}\frac{(\sigma_{\mathbf{z},q}+\sigma_{\mathbf{w}})^2}{\sigma_{\mathbf{z},q}^2}\sigma_{\mathbf{w}}^2\log^2(T_q) \\
& + \mathcal{O}\left(\frac{(q+1)^2}{\delta}\right)\sqrt{ns}\bar{x}_0^2 + \mathcal{O}(\frac{n\sqrt{s}}{1-\rho^\star}\sigma_{\mathbf{z},q+1}^2 T_{q+1}) \\
\leq & \mathcal{O}(1)\log^2\left(\frac{(q+1)^2}{\delta}\right)\frac{s^2p(n^2+p^2)\gamma}{\pi_{\min}^2(1-\bar{\rho})^2}\left(\frac{\sigma_{\mathbf{w}}^4}{\sigma_{\mathbf{z},q}^2}\log^2(T_q)+\sigma_{\mathbf{z},q+1}^2 T_q\right) \\
& + \mathcal{O}\left(\frac{(q+1)^2}{\delta}\right)\sqrt{ns}\bar{x}_0^2 \\
\leq & \mathcal{O}(1)\log^2\left(\frac{(q+1)^2}{\delta}\right)\frac{s^2p(n^2+p^2)\gamma}{\pi_{\min}^2(1-\bar{\rho})^2}\sigma_{\mathbf{w}}^2\sqrt{T_q}\log^2(T_q) + \mathcal{O}\left(\frac{(q+1)^2}{\delta}\right)\sqrt{ns}\bar{x}_0^2
\end{aligned} \tag{C.41}$$

We have $M := \mathcal{O}(\log_\gamma(\frac{T}{T_0}))$ epochs at time T. Using $T_q = \mathcal{O}(T_0\gamma^q)$, event $\cap_{q=0}^{M-1}\mathcal{E}_q$ implies

$$\begin{aligned}
\text{Regret}(T) = & \mathcal{O}(\sum_{q=1}^{M}\text{Regret}_q) \\
\leq & \mathcal{O}(1)\log^2\big(\frac{\log^2(T)}{\delta}\big)\frac{s^2p(n^2+p^2)\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\left(\gamma\sum_{q=1}^{M}\sqrt{T_q}\log^2(T_q)\right) + \mathcal{O}\big(\frac{\sqrt{ns}\log^3(T)}{\delta}\big)
\end{aligned} \tag{C.42}$$

For the term $\gamma\sum_{q=1}^{M}\sqrt{T_q}\log^2(T_q)$, we have

$$\begin{aligned}\gamma\sum_{q=1}^{M}\sqrt{T_q}\log^2(T_q) \le& \mathcal{O}(1)\gamma\sqrt{T_0}\left(\log^2(T_0)\sum_{q=1}^{M}\sqrt{\gamma}^q+\log^2(\gamma)\sum_{q=1}^{M}\sqrt{\gamma}^q i^2\right)\\ \le& \mathcal{O}(1)\gamma\sqrt{T_0}\log^2(\gamma)\sum_{q=1}^{M}\sqrt{\gamma}^q i^2\\ \le& \mathcal{O}(1)\gamma\sqrt{T_0}\log^2(\gamma)M\sqrt{\gamma}^M\left(\frac{\sqrt{\gamma}}{\sqrt{\gamma}-1}\right)^3\left(M-\frac{1}{\sqrt{\gamma}}\right)\\ \le& \mathcal{O}(1)\gamma\sqrt{T}\log^2(\gamma)\frac{\log(\frac{T}{T_0})}{\log(\gamma)}\left(\frac{\sqrt{\gamma}}{\sqrt{\gamma}-1}\right)^3\left(\frac{\log(\frac{T}{T_0})}{\log(\gamma)}-\frac{1}{\sqrt{\gamma}}\right)\\ \le& \mathcal{O}(1)\sqrt{T}\log(\frac{T}{T_0})\left(\frac{\sqrt{\gamma}}{\sqrt{\gamma}-1}\right)^3\left(\gamma\log(\frac{T}{T_0})-\sqrt{\gamma}\log(\gamma)\right)\\ \le& \mathcal{O}(\log^2(T)\sqrt{T}).\end{aligned} \tag{C.43}$$

Plugging this back into (C.42), we have

$$\text{Regret}(T) \le \mathcal{O}\left(\frac{s^2p(n^2+p^2)\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\log^2\big(\frac{\log^2(T)}{\delta}\big)\log^2(T)\sqrt{T}+\frac{\sqrt{ns}\log^3(T)}{\delta}\right) \tag{C.44}$$

which shows the regret bound in (C.40).

Now we are only left to show the occurrence probability of regret bound (C.40) is larger than $1-\delta$. To do this, we will combine Proposition C.7, C.8, C.9, and C.10 over all $q=0,1,\dots,M-1$. Note that for each individual q, these propositions hold only when certain prerequisite conditions on hyper-parameters $c_{\mathbf{x}}$, $c_{\mathbf{z}}$, T_0, and $\bar{x}_0$ are satisfied. We first show that under the choices $T_q=\gamma T_{q-1}$, $\sigma_{\mathbf{z},q}^2=\frac{\sigma_{\mathbf{w}}^2}{\sqrt{T_q}}$, and $\delta_{id,q}=\delta_{\mathbf{x}_0,q}=\frac{3}{\pi^2}\cdot\frac{\delta}{(q+1)^2}$ these hyper-parameter conditions can be satisfied for all $q=0,1,\dots,M-1$.

- Proposition C.7 requires that for $q=1,2,\dots,$ conditions $\frac{\sqrt{ns}\bar{\tau}\bar{\rho}^{T_0\gamma^q}q^2\pi^2}{3\delta}<1$ and $\bar{x}_0^2\ge\frac{n\sqrt{s}(\|\mathbf{B}_{1:s}\|^2+1)\sigma_{\mathbf{w}}^2\bar{\tau}}{(1-\bar{\rho})(1-\sqrt{ns}\cdot\bar{\tau}\bar{\rho}^{T_0\gamma^q}q^2\pi^2/3\delta)}$ need to be satisfied. $T_0\ge\frac{1}{\gamma\log(1/\bar{\rho})}\max\{\frac{2}{\log(\gamma)},\log(\frac{\pi^2\sqrt{ns}\bar{\tau}}{3\delta})\}=:\underline{T}_{\mathbf{x}_0}(\delta)$, and picking $\bar{x}_0^2\ge\frac{n\sqrt{s}(\|\mathbf{B}_{1:s}\|^2+1)\sigma_{\mathbf{w}}^2\bar{\tau}}{(1-\bar{\rho})(1-\sqrt{ns}\cdot\bar{\tau}\bar{\rho}^{T_0\gamma}\pi^2/3\delta)}$ would suffice for this.
- Proposition C.8 requires that for $q=0,1,\dots,$ condition $T_0\gamma^q\ge\underline{T}_{rgt,\bar{\epsilon}}(\frac{3\delta}{\pi^2(q+1)^2},T_0\gamma^q)$ holds, which can be satisfied when one chooses $T_0\ge\mathcal{O}(\underline{T}_{rgt,\bar{\epsilon}}(\delta,T_0))$.
- Proposition C.9 requires the following to hold: $c_{\mathbf{x}}\ge\underline{c}_{\mathbf{x}}(\bar{\rho},\bar{\tau}),c_{\mathbf{z}}\ge\underline{c}_{\mathbf{z}}$, and $T_0\gamma^q\ge\max\left\{\underline{T}_{MC,1}(\frac{3\delta}{8\pi^2q^2}),\underline{T}_{id,N}(\frac{3\delta}{2\pi^2(q+1)^2},\bar{\rho},\bar{\tau})\right\}$. The last one can be satisfied when $T_0\ge\mathcal{O}(\max\{\underline{T}_{MC,1}(\delta),\underline{T}_{id,N}(\delta,\bar{\rho},\bar{\tau})\})$.
- Proposition C.10 requires no conditions on hyper-parameters.

Therefore, when $c_{\mathbf{x}} \geq \underline{c}_{\mathbf{x}}(\bar{\rho}, \bar{\tau}), c_{\mathbf{z}} \geq \underline{c}_{\mathbf{z}}$,

$$T_0 \geq \mathcal{O}(\max\{\underline{T}_{\mathbf{x}_0}(\delta), \underline{T}_{rgt,\bar{\epsilon}}(\delta, T_0), \underline{T}_{MC,1}(\delta), \underline{T}_{id,N}(\delta, \bar{\rho}, \bar{\tau})\}) =: \mathcal{O}(\underline{T}_{rgt}(\delta, T_0)),$$

we can apply Propositions C.7, C.8, C.9, and C.10 to every epoch $q = 0, 1, \ldots, M-1$. First note that Propositions C.7 and C.9 give the following

$$\mathbb{P}(\mathcal{D}_q \mid \cap_{j=0}^{q-1}\mathcal{E}_j) = \mathbb{P}(\mathcal{D}_q \mid \mathcal{B}_{q-1}\mathcal{D}_{q-1}) > 1 - \frac{3\delta}{\pi^2(q+1)^2}, \quad \mathbb{P}(\mathcal{D}_0) \geq 1 - \frac{3\delta}{\pi^2}$$
$$\mathbb{P}(\mathcal{C}_q \mid \cap_{j=0}^{q-1}\mathcal{E}_j) = \mathbb{P}(\mathcal{C}_q \mid \mathcal{B}_{q-1}) \geq 1 - \frac{3\delta}{\pi^2(q+1)^2}, \quad \mathbb{P}(\mathcal{C}_0) \geq 1 - \frac{3\delta}{\pi^2}.$$

Then combining the probability bounds in Propositions C.7, C.8, C.9, and C.10, we have

$$\begin{aligned}
&\mathbb{P}\left(\text{Regret bounds in (C.40) holds}\right)\\
\geq&\mathbb{P}(\cap_{q=0}^{M-1}\mathcal{E}_q)\\
=&\mathbb{P}(\mathcal{A}_M, \mathcal{B}_{M-1}, \mathcal{C}_{M-1}, \mathcal{D}_{M-1} \mid \cap_{q=0}^{M-2}\mathcal{E}_q) \cdot \mathbb{P}(\cap_{q=0}^{M-2}\mathcal{E}_q)\\
=&\mathbb{P}(\mathcal{C}_{M-1}, \mathcal{D}_{M-1} \mid \cap_{q=0}^{M-2}\mathcal{E}_q) \cdot \mathbb{P}(\cap_{q=0}^{M-2}\mathcal{E}_q)\\
\geq& (1 - \delta_{id,M-1} - \delta_{\mathbf{x}_0,M-1}) \cdot \mathbb{P}(\cap_{q=0}^{M-2}\mathcal{E}_q)\\
\geq& \prod_{q=0}^{M-1}(1 - \delta_{id,q} - \delta_{\mathbf{x}_0,q})\\
\geq&1 - \sum_{q=0}^{M-1}(\delta_{id,q} + \delta_{\mathbf{x}_0,q})\\
\geq&1 - \delta.
\end{aligned} \tag{C.45}$$

where the last line holds since $\sum_{q=0}^{M-1} \frac{1}{(q+1)^2} \leq \frac{\pi^2}{6}$. □

C.5 Regret Under Uniform Stability

C.5.1 Proof for Theorem 5.2

As we discussed in Section 5.4.1, under mean-square stability, the regret upper bound in Theorem 5.1 (or the complete version Theorem C.11) involves $\frac{1}{\delta}$ dependency on failure probability δ. By checking the proof for Theorem C.11, we can see the only source for $\frac{1}{\delta}$ is event $\mathcal{D}_q$ in (C.35) and the corresponding Proposition C.7, which provides $1-\delta$ probability bound for event $\mathcal{D}_q$ – the initial state $\mathbf{x}_0^{(q+1)}$ of epoch $q+1$, a.k.a. the final state $\mathbf{x}_{T_q}^{(q)}$ of epoch q, is bounded by $\|\mathbf{x}_0^{(q+1)}\|^2 = \|\mathbf{x}_{T_q}^{(q)}\|^2 \leq \mathcal{O}(\frac{1}{\delta})$. In Proposition C.7, we get this bound using

Markov inequality $\|\mathbf{x}_{T_q}^{(q)}\|^2 \leq \mathbb{E}[\|\mathbf{x}_{T_q}^{(q)}\|^2]/\delta$ and Lemma 2.15 which provides an upper bound on the numerator $\mathbb{E}[\|\mathbf{x}_{T_q}^{(q)}\|^2]$ under mean-square stability. From event $\mathcal{A}_q$ in (C.35) we see the regret of epoch q directly depends on its epoch initial state $\|\mathbf{x}_0^{(q)}\|^2$, thus in the final cumulative regret, the cumulative impact of initial states from all epochs, $\sum_q \|\mathbf{x}_0^{(q)}\|^2$ with order $\frac{1}{\delta}$, will show up, as given in (C.42). Therefore, whether $\frac{1}{\delta}$ terms can be relaxed directly hinges on whether one could refine Proposition C.7 to get a tighter dependency on δ.

This refinement, however, is not possible under the mean-square stability assumption only, and we can easily construct a toy example to show that the $\frac{1}{\delta}$ dependency resulting from the Markov inequality cannot be improved. Consider a two-mode, one-dimensional, autonomous MJS:

$$\begin{cases} x_{t+1} = 2x_t \\ x_{t+1} = 0.5x_t \end{cases} \text{with Markov matrix } \mathbf{T} = \begin{bmatrix} 0.1 & 0.9 \\ 0.1 & 0.9 \end{bmatrix}$$

with $x_0 \sim \mathcal{N}(0,1)$, and $\mathbb{P}(\omega_0 = 1) = 0.1$. It is easy to check this MJS is mean-square stable by the spectral radius criterion discussed below Definition 2.8. Also note that with probability 0.1^t, $\omega_{0:t-1} = 1$ and $x_t = 2^t x_0$. Therefore, for any $a > 0$,

$$\mathbb{P}(x_t \geq a) = \sum_{\omega_{0:t-1}} \mathbb{P}(x_t \geq a \mid \omega_{0:t-1})\mathbb{P}(\omega_{0:t-1}) \tag{C.46}$$

$$\geq \mathbb{P}(x_t \geq a \mid \omega_{0:t-1} = 1)\mathbb{P}(\omega_{0:t-1} = 1) = 0.1^t \cdot \mathbb{P}(x_0 \geq 2^{-t}a) \, . \tag{C.47}$$

where the inequality in (C.47) is extremely loose since we condition only on the most improbable event. For standard Gaussian x_0, $\mathbb{P}(x_0 \geq a) \geq \frac{C}{a}\exp(-\frac{a^2}{2})$ for some absolute constant C. Thus $\mathbb{P}(x_t \geq a) \geq C\frac{0.2^t}{a}\exp(-\frac{2^{-2t}a^2}{2})$. From this, we see that for any $a > 0$, any $t \geq \log(a)/\log(2)$, we have $\mathbb{P}(x_t \geq a) \geq C\frac{0.2^t}{\sqrt{e}a}$. We can observe that though when t grows slower than $\log(a)$, the tail of x_t has exponential decay, the Markov inequality decay, i.e. $\frac{1}{a}$, will eventually show up when t gets larger. Interpretation from failure probability δ perspective is the following: letting $\delta = C\frac{0.2^t}{\sqrt{e}a}$, we have $\mathbb{P}(x_t \leq C\frac{0.2^t}{\sqrt{e}\delta}) \leq 1 - \delta$, which means any δ dependency lighter than $\frac{1}{\delta}$ must have probability less than $1-\delta$. This further implies that in the regret analysis of adaptive control, in order to obtain better probability dependency, the time horizon has to be limited, which greatly impairs its value in practice.

Intuitively, the mean-square stability assumption only provides us with stable behavior of $\|\mathbf{x}_t\|^2$ in the expectation (w.r.t. mode switching) sense, and having only this first-order moment information is of little use compared with the deterministic Lyapunov stability typically used for LTI systems, which allows one to bound $\|\mathbf{x}_t\|^2$ with only $\log(\frac{1}{\delta})$ dependence

((Dean et al., 2018, Lemma C.5)). Then, one may wonder naturally: Does there exist a deterministic version of stability for switched systems? Can this stability (if exists) help build similar dependence for switched systems? The answers to both questions are yes and will be discussed in this appendix. In short, if there exists uniform stability for the MJS, we can adapt Proposition C.7 such that $\|\mathbf{x}_0^{(q)}\|^2$ can instead be bounded much more tightly by $\|\mathbf{x}_0^{(q)}\|^2 \leq \mathcal{O}(\log(\frac{1}{\delta}))$, thus the $\frac{1}{\delta}$ dependency can improve to $\log(\frac{1}{\delta})$ in the regret bound (5.3) (or (C.40)). The final improved regret bound is presented in Theorem 5.2. In order to show it, we will need to adapt Proposition C.7 together with several related results (Lemma 2.15, Lemma C.1, Lemma C.2) to the uniform stability case, and we append suffix "a" in the result label to denote the adapted versions.

To begin with, recall that $\mathbf{K}^\star_{1:s}$ is the optimal controller for the infinite-horizon problem MJS-LQR$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ and define the closed-loop state matrix $\mathbf{L}^\star_i = \mathbf{A}_i + \mathbf{B}_i\mathbf{K}^\star_i$ for all i. Let $\theta^\star$ denote the joint spectral radius of $\mathbf{L}^\star_{1:s}$, i.e. $\theta^\star := \lim_{l\to\infty} \max_{\omega_{1:l}\in[s]^l} \|\mathbf{L}^\star_{\omega_1} \cdots \mathbf{L}^\star_{\omega_l}\|^{\frac{1}{l}}$. We say $\mathbf{L}^\star_{1:s}$ is uniformly stable if and only if $\theta^\star < 1$. Similar to τ in Definition 2.1, define $\kappa^\star := \sup_{l\in\mathbb{N}} \max_{\omega_{1:l}\in[s]^l} \|\mathbf{L}^\star_{\omega_1} \cdots \mathbf{L}^\star_{\omega_l}\|/(\theta^\star)^l$. Note that the pair $\{\theta^\star, \kappa^\star\}$ for uniform stability is just the counterpart of $\{\rho^\star, \tau(\boldsymbol{\mathcal{L}}^\star)\}$ for mean-square stability defined in Appendix C. Similar as before, Table C.3 lists all the shorthand notations to be used in this appendix for quick reference.

Table C.3: Notations — Uniform Stability

$\bar{\sigma}^2$	$\|\mathbf{B}_{1:s}\|^2\|\mathbf{\Sigma_z}\| + \|\mathbf{\Sigma_w}\|$ or $\|\mathbf{B}_{1:s}\|^2\sigma^2_{\mathbf{z},0} + \sigma^2_{\mathbf{w}}$
$\bar{\theta}$	$(1+\theta^\star)/2$
$\bar{\kappa}$	$\kappa^\star$
$\bar{\epsilon}^{us}_{\mathbf{K}}$	$\frac{1-\rho^\star}{2\kappa^\star\|\mathbf{B}_{1:s}\|}$
$\bar{\bar{\epsilon}}_{\mathbf{K}}$	$\min\{\bar{\epsilon}^{us}_{\mathbf{K}}, \bar{\epsilon}_{\mathbf{K}}\}$
$\bar{\bar{\epsilon}}_{\mathbf{A},\mathbf{B},\mathbf{T}}$	$\min\{\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}, \frac{\bar{\bar{\epsilon}}_{\mathbf{K}}}{2C^{\mathbf{K}}_{\mathbf{A},\mathbf{B},\mathbf{T}}}\}$
$\bar{x}^{us}$	$2\bar{\kappa}^2\bar{\sigma}^2(6\max\{\sqrt{n}e^{3n}, \sqrt{p}e^{3p}\} + \frac{5}{(1-\bar{\theta})^2})^2$
$\underline{T}^{us}_{\mathbf{x}_0}(\delta)$	$\max\{\frac{54\bar{\kappa}^4\bar{\sigma}^2}{(1-\bar{\theta})\bar{x}^{us}\log(1/\bar{\theta})\log(\gamma)}, \frac{1}{\gamma\log(1/\bar{\theta})}\log(6\bar{\kappa}^2 + \frac{54n\sqrt{s}\bar{\kappa}^4\bar{\sigma}^2\log(\pi^2/3\delta)}{(1-\bar{\theta})(1-\bar{\rho})\bar{x}^{us}\delta})\}$
$\underline{T}^{us}_{rgt,\bar{\epsilon}}(\delta, T)$	$\mathcal{O}(\frac{\sqrt{s}(n+p)}{\pi_{\min}(1-\bar{\rho})}\bar{\bar{\epsilon}}^{-4}_{\mathbf{A},\mathbf{B},\mathbf{T}}\log(\frac{1}{\delta})\log^4(T))$ $\mathcal{O}(\frac{\sqrt{s}(n+p)}{\pi_{\min}(1-\bar{\rho})^2}\bar{\bar{\epsilon}}^{-2}_{\mathbf{A},\mathbf{B},\mathbf{T}}\log(\frac{1}{\delta})\log^2(T))$ (when $\mathbf{B}_{1:s}$ is known)
$\underline{T}^{us}_{rgt}(\delta, T)$	$\max\{\underline{T}^{us}_{\mathbf{x}_0}(\delta), \underline{T}^{us}_{rgt,\bar{\epsilon}}(\delta, T), \underline{T}_{MC,1}(\delta), \underline{T}_{id,N}(\underline{L}, \delta, T, \bar{\rho}, \bar{\tau})\}$

The following Lemma 2.15a bounds the state $\mathbf{x}_t$ under the designed input. Compared with its counterpart Lemma 2.15 which is only able to bound $\mathbb{E}[\|\mathbf{x}_t\|^2]$, Lemma 2.15a provides high-probability bound for $\|\mathbf{x}_t\|^2$.

Lemma 2.15a. *Consider an* $\text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$ *with noise* $\mathbf{w}_t \sim \mathcal{N}(0, \mathbf{\Sigma_w})$. *Consider controller* $\mathbf{K}_{1:s}$, *and let* $\mathbf{L}_{1:s}$ *denote the closed-loop state matrices with* $\mathbf{L}_i = \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i$. *Assume there exist constants* κ *and* $\theta \in [0, 1)$ *such that, for any sequence* $\omega_{1:l} \in [s]^l$ *with any length* l, $\|\mathbf{L}_{\omega_1} \cdots \mathbf{L}_{\omega_l}\| \leq \kappa\theta^l$. *Let the input be* $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$ *with* $\mathbf{z}_t \sim \mathcal{N}(0, \Sigma_\mathbf{z})$. *Then, for any* $t \geq e^{6\max\{n,p\}}$, *with probability at least* $1 - \delta$, *we have*

$$\|\mathbf{x}_t\|^2 \leq 3\kappa^2\theta^{2t}\|\mathbf{x}_0\|^2 + \frac{18\kappa^2\bar{\sigma}^2}{(1-\theta)^2}\log(\frac{1}{\delta}) + c \tag{C.48}$$

where $\bar{\sigma}^2 := \|\mathbf{\Sigma_w}\| + \|\mathbf{B}_{1:s}\|^2\|\mathbf{\Sigma_z}\|$ *and* $c := 2\kappa^2\bar{\sigma}^2(6\max\{\sqrt{n}e^{3n}, \sqrt{p}e^{3p}\} + \frac{5}{(1-\theta)^2})^2$.

Proof. From the MJS dynamics (2.6) and plugging in the input $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$, we have the following.

$$\begin{aligned}\mathbf{x}_t = \Big(\prod_{h=0}^{t-1}\mathbf{L}_{\omega_h}\Big)\mathbf{x}_0 + &\sum_{i=0}^{t-2}\Big(\prod_{h=i+1}^{t-1}\mathbf{L}_{\omega_h}\Big)\mathbf{B}_{\omega_i}\mathbf{z}_i + \mathbf{B}_{\omega_{t-1}}\mathbf{z}_{t-1} \\ &\sum_{i=0}^{t-2}\Big(\prod_{h=i+1}^{t-1}\mathbf{L}_{\omega_h}\Big)\mathbf{w}_i + \mathbf{w}_{t-1}.\end{aligned} \tag{C.49}$$

Then, by triangle inequality and the assumption that $\|\mathbf{L}_{\omega_1} \cdots \mathbf{L}_{\omega_l}\| \leq \kappa\theta^l$, we have

$$\begin{aligned}\|\mathbf{x}_t\| &\leq \kappa\theta^t\|\mathbf{x}_0\| + \kappa\|\mathbf{B}_{1:s}\|\sum_{i=0}^{t-1}\theta^{t-i-1}\|\mathbf{z}_i\| + \kappa\sum_{i=0}^{t-1}\theta^{t-i-1}\|\mathbf{w}_i\| \\ &= \kappa\theta^t\|\mathbf{x}_0\| + \kappa\|\mathbf{B}_{1:s}\|\sum_{i=0}^{t-1}\theta^{i}\|\mathbf{z}_{t-i-1}\| + \kappa\sum_{i=0}^{t-1}\theta^{i}\|\mathbf{w}_{t-i-1}\|.\end{aligned} \tag{C.50}$$

For each $\mathbf{w}_{t-i-1}$, using Lemma A.1 (replacing e^{-t} with δ_i), we have with probability $1 - \delta_i$,

$$\|\mathbf{w}_{t-i-1}\| \leq \sqrt{3\|\mathbf{\Sigma_w}\|}\log^{0.5}(\frac{1}{\min\{\delta_i, \bar{\delta}_n\}}), \tag{C.51}$$

where $\bar{\delta}_n := e^{-(3+2\sqrt{2})n}$, and n is the dimension of vector $\mathbf{w}_{t-i-1}$. In the following, for all $i = 0, 1, \ldots, t-1$, we set $\delta_i = \frac{3}{\pi^2}\frac{\delta}{(i+1)^2}$. First note that when $i \geq \bar{i} := \sqrt{\frac{3\delta}{\pi^2\bar{\delta}_n}} - 1$, we have $\min\{\delta_i, \bar{\delta}_n\} = \delta_i$, i.e. $\delta_i \leq \bar{\delta}_n$, and $\min\{\delta_i, \bar{\delta}_n\} = \bar{\delta}_n$ otherwise. Then, applying union bound for all i, we know with probability at least $1-\frac{\delta}{2}$,

$$\begin{aligned}\sum_{i=0}^{t-1}\theta^i\|\mathbf{w}_{t-i-1}\| &\leq \sqrt{3\|\boldsymbol{\Sigma}_\mathbf{w}\|}\sum_{i=0}^{t-1}\theta^i\log^{0.5}(\frac{1}{\min\{\delta_i,\bar{\delta}_n\}}) \\ &\leq \sqrt{3\|\boldsymbol{\Sigma}_\mathbf{w}\|}\left(\sum_{i=0}^{t-1}\theta^i\log^{0.5}(\frac{1}{\delta_i})+(\bar{i}+1)\log^{0.5}(\frac{1}{\bar{\delta}_n})\right).\end{aligned} \tag{C.52}$$

For the term $\sum_{i=0}^{t-1}\theta^i\log^{0.5}(\frac{1}{\delta_i})$ above, we have $\sum_{i=0}^{t-1}\theta^i\log^{0.5}(\frac{1}{\delta_i}) = \sum_i\theta^i\log^{0.5}(\frac{\pi^2(i+1)^2}{3\delta}) \leq \sum_i\theta^i(\log^{0.5}(\frac{1}{\delta})+\sqrt{2}\log^{0.5}(\frac{\pi(i+1)}{\sqrt{3}})) \leq \frac{1}{1-\theta}\log^{0.5}(\frac{1}{\delta})+\sqrt{2}\sum_i\theta^i\frac{\pi(i+1)}{\sqrt{3}} \leq \frac{1}{1-\theta}\log^{0.5}(\frac{1}{\delta})+\frac{\sqrt{2}\pi}{\sqrt{3}}\frac{1}{(1-\theta)^2}$. And for the term $(\bar{i}+1)\log^{0.5}(\frac{1}{\bar{\delta}_n})$ in (C.52), by the definitions of $\bar{i}$ and $\bar{\delta}_n$, we have $(\bar{i}+1)\log^{0.5}(\frac{1}{\bar{\delta}_n}) \leq \sqrt{2n}e^{3n}$. Plugging these two results back into (C.52), we have, with probability at least $1-\frac{\delta}{2}$,

$$\sum_{i=0}^{t-1}\theta^i\|\mathbf{w}_{t-i-1}\| \leq \frac{\sqrt{3\|\boldsymbol{\Sigma}_\mathbf{w}\|}}{1-\theta}\log^{0.5}(\frac{1}{\delta})+\frac{5\sqrt{\|\boldsymbol{\Sigma}_\mathbf{w}\|}}{(1-\theta)^2}+3\sqrt{n}e^{3n}\sqrt{\|\boldsymbol{\Sigma}_\mathbf{w}\|}. \tag{C.53}$$

Similarly, with probability at least $1-\frac{\delta}{2}$,

$$\sum_{i=0}^{t-1}\theta^i\|\mathbf{z}_{t-i-1}\| \leq \frac{\sqrt{3\|\boldsymbol{\Sigma}_\mathbf{z}\|}}{1-\theta}\log^{0.5}(\frac{1}{\delta})+\frac{5\sqrt{\|\boldsymbol{\Sigma}_\mathbf{z}\|}}{(1-\theta)^2}+3\sqrt{p}e^{3p}\sqrt{\|\boldsymbol{\Sigma}_\mathbf{z}\|}. \tag{C.54}$$

Plugging (C.53) and (C.54) back into (C.50) and applying union bound, we have, with probability $1-\delta$,

$$\begin{aligned}\|\mathbf{x}_t\| \leq \kappa\theta^t\|\mathbf{x}_0\| &+ \frac{\sqrt{3}\kappa(\sqrt{\|\boldsymbol{\Sigma}_\mathbf{w}\|}+\|\mathbf{B}_{1:s}\|\sqrt{\|\boldsymbol{\Sigma}_\mathbf{z}\|})}{(1-\theta)^2}\log^{0.5}(\frac{1}{\delta}) \\ &+\kappa(\sqrt{\|\boldsymbol{\Sigma}_\mathbf{w}\|}+\|\mathbf{B}_{1:s}\|\sqrt{\|\boldsymbol{\Sigma}_\mathbf{z}\|})\left(3\max\{\sqrt{n}e^{3n},\sqrt{p}e^{3p}\}+\frac{5}{(1-\theta)^2}\right).\end{aligned} \tag{C.55}$$

Taking squares of both sides and using Cauchy-Schwartz inequality, we have

$$\|\mathbf{x}_t\|^2 \leq 3\kappa^2\theta^{2t}\|\mathbf{x}_0\|^2+\frac{18\kappa^2\bar{\sigma}^2}{1-\theta}\log(\frac{1}{\delta})+c \tag{C.56}$$

where $\bar{\sigma}^2 := \|\boldsymbol{\Sigma}_\mathbf{w}\|+\|\mathbf{B}_{1:s}\|^2\|\boldsymbol{\Sigma}_\mathbf{z}\|$ and $c := 6\kappa^2\bar{\sigma}^2(3\max\{\sqrt{n}e^{3n},\sqrt{p}e^{3p}\}+\frac{5}{(1-\theta)^2})^2$. □

The following Lemma C.1a describes that given a set of matrices that have joint spectral radius smaller than 1, i.e. uniformly stable, moderate perturbation can preserve the uniform stability. On the other hand, its counterpart, Lemma C.1, considers perturbation results for mean-square stability.

Lemma C.1a (Joint Spectral Radius Perturbation)**.** *Assume $\theta^\star < 1$. For an arbitrary controller $\mathbf{K}_{1:s}$ and resulting closed-loop state matrices $\mathbf{L}_{1:s}$ with $\mathbf{L}_i = \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i$, let $\theta(\mathbf{L}_{1:s})$ denote the joint spectral radius of $\mathbf{L}_{1:s}$. Assume $\|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\| \leq \bar{\epsilon}^{us}_{\mathbf{K}} := \frac{1-\theta^\star}{2\kappa^\star \|\mathbf{B}_{1:s}\|}$, then for any sequence $\omega_{1:l} \in [s]^l$ with any length l,*

$$\|\prod_{j=1}^{l} \mathbf{L}_{\omega_j}\| \leq \bar{\kappa}\bar{\theta}^l \tag{C.57}$$

$$\theta(\mathbf{L}_{1:s}) \leq \bar{\theta}. \tag{C.58}$$

where $\bar{\kappa} = \kappa^\star$ and $\bar{\theta} = \frac{1+\theta^\star}{2}$.

Proof. Let $\mathbf{E}_i := \mathbf{L}_i - \mathbf{L}^\star_i$, then we see $\|\mathbf{E}_i\| \leq \|\mathbf{B}_{1:s}\|\bar{\epsilon}^{us}_{\mathbf{K}}$ and $\prod_{j=1}^{l} \mathbf{L}_{\omega_j} = \prod_{j=1}^{l}(\mathbf{L}^\star_{\omega_j} + \mathbf{E}_{\omega_j})$. In the expansion of $\prod_{j=1}^{l}(\mathbf{L}^\star_{\omega_j} + \mathbf{E}_{\omega_j})$, for each $h = 0, 1, \ldots, l$, there are $\binom{l}{h}$ terms, each of which is a product where $\mathbf{E}$ has degree h and $\mathbf{L}^\star$ has degree $l-h$. We let $\mathbf{F}_{h,l}$ with $h = 0, 1, \ldots, l$ and $l \in [\binom{l}{h}]$ to index such terms. Note that $\|\mathbf{F}_{h,l}\| \leq (\kappa^\star)^{h+1}(\theta^\star)^{l-h}(\|\mathbf{B}_{1:s}\|\bar{\epsilon}^{us}_{\mathbf{K}})^h$. Then, we have

$$\begin{aligned}
\|\prod_{j=1}^{l} \mathbf{L}_{\omega_j}\| \leq & \sum_{h=0}^{l} \sum_{l \in [\binom{l}{h}]} \|\mathbf{F}_{h,l}\| \\
\leq & \sum_{h=0}^{l} \binom{l}{h} (\kappa^\star)^{h+1}(\theta^\star)^{l-h}(\|\mathbf{B}_{1:s}\|\bar{\epsilon}^{us}_{\mathbf{K}})^h \\
\leq & \kappa^\star(\kappa^\star\|\mathbf{B}_{1:s}\|\bar{\epsilon}^{us}_{\mathbf{K}} + \theta^\star)^l.
\end{aligned} \tag{C.59}$$

Then (C.57) follows from the fact that $\bar{\epsilon}^{us}_{\mathbf{K}} \leq \frac{1-\theta^\star}{2\kappa^\star\|\mathbf{B}_{1:s}\|}$ and $\bar{\theta} := \frac{1+\theta^\star}{2}$. To proceed, noticing that $\theta(\mathbf{L}_{1:s}) = \lim_{l\to\infty} \max_{\omega_{1:l}\in[s]^l} \|\prod_{j=1}^{l} \mathbf{L}_{\omega_j}\|^{\frac{1}{l}}$ and using the result in (C.57), we can show (C.58). □

In the Lemma C.1a, if the controller $\mathbf{K}_{1:s}$ is obtained by solving the infinite-horizon MJS-LQR$(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ for some estimated MJS$(\hat{\mathbf{A}}_{1:s}, \hat{\mathbf{B}}_{1:s}, \hat{\mathbf{T}})$, the following result provides the required estimation accuracy such that the resulting $\mathbf{K}_{1:s}$ is uniformly stabilizing.

Lemma C.2a. *Under the setup of Lemma C.2, if $\max\{\bar{\epsilon}_{\mathbf{A},\mathbf{B}}, \bar{\epsilon}_{\mathbf{T}}\} \leq \bar{\bar{\epsilon}}_{\mathbf{A},\mathbf{B},\mathbf{T}}$, then we have $\|\mathbf{K}_{1:s} - \mathbf{K}^\star_{1:s}\| \leq \bar{\epsilon}_{\mathbf{K}}$, and Lemma C.1a is applicable.*

Recall we defined events $\mathcal{A}_q, \mathcal{B}_q, \mathcal{C}_q, \mathcal{D}_q$ in (C.60) to analyze the events happen in each epoch of the regret. To adapt to the uniform stability assumption, we redefine event $\mathcal{B}_q$ and $\mathcal{D}_q$ while keep $\mathcal{A}_q$ and $\mathcal{C}_q$ as before. For easier reference, We list all of them below.

$$
\begin{aligned}
\mathcal{A}_q &= \left\{ \text{Regret}_q \leq \mathcal{O}\left(sp\left(\epsilon_{\mathbf{A},\mathbf{B}}^{(q-1)} + \epsilon_{\mathbf{T}}^{(q-1)}\right)^2 \sigma_{\mathbf{w}}^2 T_q + \sqrt{ns}\|\mathbf{x}_0^{(q)}\|^2 + \frac{n\sqrt{s}}{1-\rho^\star}\sigma_{\mathbf{z},q}^2 T_q + c_{\mathcal{A}} \right) \right\} \\
\mathcal{B}_q &= \left\{ \epsilon_{\mathbf{A},\mathbf{B}}^{(q)} \leq \bar{\bar{\epsilon}}_{\mathbf{A},\mathbf{B},\mathbf{T}}, \epsilon_{\mathbf{T}}^{(q)} \leq \bar{\bar{\epsilon}}_{\mathbf{A},\mathbf{B},\mathbf{T}}, \epsilon_{\mathbf{K}}^{(q+1)} \leq \bar{\bar{\epsilon}}_{\mathbf{K}} \right\}, \forall q = 0, 1, \dots \\
\mathcal{C}_q &= \left\{ \epsilon_{\mathbf{A},\mathbf{B}}^{(q)} \leq \mathcal{O}\left(\log(\frac{1}{\delta_{id,q}}) \frac{\sigma_{\mathbf{z},q} + \sigma_{\mathbf{w}}}{\sigma_{\mathbf{z},q}\pi_{\min}} \frac{\sqrt{s}(n+p)\log(T_q)}{(1-\bar{\rho})\sqrt{T_q}} \right), \right. \\
&\qquad \left. \epsilon_{\mathbf{T}}^{(q)} \leq \mathcal{O}\left(\log(\frac{1}{\delta_{id,q}}) \frac{1}{\pi_{\min}} \sqrt{\frac{\log(T_q)}{T_q}} \right) \right\} \\
\mathcal{D}_q &= \left\{ \|\mathbf{x}_0^{(q+1)}\|^2 = \|\mathbf{x}_{T_q}^{(q)}\|^2 \leq \frac{18\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2} \log(\frac{1}{\delta_{\mathbf{x}_0,q}}) + 2\bar{x}^{us} \right\}, \forall q = 1, 2, \dots, \\
\mathcal{D}_0 &= \left\{ \|\mathbf{x}_0^{(1)}\|^2 = \|\mathbf{x}_{T_0}^{(0)}\|^2 \leq \frac{n\sqrt{s}\bar{\tau}\bar{\sigma}^2/(1-\bar{\rho})}{\delta_{\mathbf{x}_0,0}} \right\},
\end{aligned}
\tag{C.60}
$$

where we define $\bar{x}^{us} := 2\bar{\kappa}^2\bar{\sigma}^2(6\max\{\sqrt{n}e^{3n}, \sqrt{p}e^{3p}\} + \frac{5}{(1-\bar{\theta})^2})^2$, $\bar{\bar{\epsilon}}_{\mathbf{K}} := \min\{\bar{\epsilon}_{\mathbf{K}}^{us}, \bar{\epsilon}_{\mathbf{K}}\}$, $\bar{\bar{\epsilon}}_{\mathbf{A},\mathbf{B},\mathbf{T}} := \min\{\bar{\epsilon}_{\mathbf{A},\mathbf{B},\mathbf{T}}, \frac{\bar{\bar{\epsilon}}_{\mathbf{K}}}{2C_{\mathbf{A},\mathbf{B},\mathbf{T}}^{\mathbf{K}}}\}$ and $\bar{\sigma}^2 := \|\mathbf{B}_{1:s}\|^2\sigma_{\mathbf{z},0}^2 + \sigma_{\mathbf{w}}^2$. Event $\mathcal{D}_q$ describes the initial state magnitude of epoch $q+1$. Since Algorithm 2 requires initial mean-square stabilizing controller $\mathbf{K}_{1:s}^{(0)}$ for epoch 0, and as in the proof for the following Proposition C.7a, epoch $1, 2, \dots$ have uniformly stabilizing controller, thus we define $\mathcal{D}_0$ and $\mathcal{D}_1, \mathcal{D}_2, \dots$ separately.

Proposition C.7a. *Assuming that* $T_q \geq \frac{1}{2\log(1/\bar{\theta})} \log\left(6\bar{\kappa}^2 + \frac{54\bar{\kappa}^4\bar{\sigma}^2}{(1-\bar{\theta})\bar{x}^{us}} \log(\frac{1}{\delta_{\mathbf{x}_0,q-1}})\right)$ *and* $T_1 \geq \frac{1}{2\log(1/\bar{\theta})} \log\left(\frac{3n\sqrt{s}\bar{\kappa}^2\bar{\tau}\bar{\sigma}^2}{(1-\bar{\rho})\bar{x}^{us}\delta_{\mathbf{x}_0,0}}\right)$, *we have*

$$
\mathbb{P}(\mathcal{D}_q \mid \mathcal{B}_{q-1}, \mathcal{D}_{q-1}) \geq 1 - \delta_{\mathbf{x}_0,q} \tag{C.61}
$$

and $\mathbb{P}(\mathcal{D}_0) \geq 1 - \delta_{\mathbf{x}_0,0}$.

Proof. For the initial epoch 0, i.e. $q = 0$, since we assume in Algorithm 2 that the initial controller $\mathbf{K}_{1:s}^{(0)}$ stabilizes the MJS in the mean-squared sense, similar to the proof for Proposition C.7, we have $\mathbb{E}[\|\mathbf{x}_{T_0}^{(0)}\|^2] \leq n\sqrt{s}(\|\mathbf{B}_{1:s}\|^2\sigma_{\mathbf{z},0}^2 + \sigma_{\mathbf{w}}^2)\frac{\bar{\tau}}{1-\bar{\rho}}$. Then by Markov inequality, with probability $1-\delta_{\mathbf{x}_0,0}$, $\|\mathbf{x}_{T_0}^{(0)}\|^2 \leq \frac{n\sqrt{s}\bar{\tau}\bar{\sigma}^2/(1-\bar{\rho})}{\delta_{\mathbf{x}_0,0}}$ where $\bar{\sigma}^2 := \|\mathbf{B}_{1:s}\|^2\sigma_{\mathbf{z},0}^2 + \sigma_{\mathbf{w}}^2$. This shows $\mathbb{P}(\mathcal{D}_0) \geq 1 - \delta_{\mathbf{x}_0,0}$.

For epoch $q = 1, 2, \dots$, given event $\mathcal{B}_{q-1}$, we know $\epsilon_{\mathbf{K}}^{(q)} \leq \bar{\bar{\epsilon}}_{\mathbf{K}} \leq \bar{\epsilon}_{\mathbf{K}}^{us}$. Let $\mathbf{L}_{1:s}^{(q)}$ denote the closed-loop state matrices for epoch q, then by Lemma C.1a, $\epsilon_{\mathbf{K}}^{(q)} \leq \bar{\epsilon}_{\mathbf{K}}^{us}$ implies that for any l and any sequence $\omega_{1:l} \in [s]^l$, $\|\prod_{j=1}^{l} \mathbf{L}_{\omega_j}^{(q)}\| \leq \bar{\kappa}\bar{\theta}^l$. Then using the bound on $\|\mathbf{x}_t\|$ in

Lemma 2.15a, we have, with probability $1-\delta_{\mathbf{x}_0,q}$,

$$\|\mathbf{x}_{T_q}^{(q)}\|^2 \le \frac{18\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2}\log(\frac{1}{\delta_{\mathbf{x}_0,q}}) + 3\bar{\kappa}^2\bar{\theta}^{2T_q}\|\mathbf{x}_0^{(q)}\|^2 + \bar{x}^{us} \tag{C.62}$$

where $\bar{x}^{us} := 2\bar{\kappa}^2\bar{\sigma}^2(6\max\{\sqrt{n}e^{3n}, \sqrt{p}e^{3p}\} + \frac{5}{(1-\bar{\theta})^2})^2$.

- When $q=1$, given D_0, i.e. $\|\mathbf{x}_0^{(1)}\|^2 \le \frac{n\sqrt{s}\bar{\tau}\bar{\sigma}^2/(1-\bar{\rho})}{\delta_{\mathbf{x}_0,0}}$, the above (C.62) gives $\|\mathbf{x}_{T_1}^{(1)}\|^2 \le \frac{18\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2}\log(\frac{1}{\delta_{\mathbf{x}_0,1}}) + 3\bar{\kappa}^2\bar{\theta}^{2T_1}\frac{n\sqrt{s}\bar{\tau}\bar{\sigma}^2/(1-\bar{\rho})}{\delta_{\mathbf{x}_0,0}} + \bar{x}^{us}$. One can check that when we choose $T_1 \ge \frac{1}{2\log(1/\bar{\theta})}\log\left(\frac{3n\sqrt{s}\bar{\kappa}^2\bar{\tau}\bar{\sigma}^2}{(1-\bar{\rho})\bar{x}^{us}\delta_{\mathbf{x}_0,0}}\right)$, we have that $3\bar{\kappa}^2\bar{\theta}^{2T_1}\frac{n\sqrt{s}\bar{\tau}\bar{\sigma}^2/(1-\bar{\rho})}{\delta_{\mathbf{x}_0,0}} \le \bar{x}^{us}$, which gives

$$\|\mathbf{x}_{T_1}^{(1)}\|^2 \le \frac{18\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2}\log(\frac{1}{\delta_{\mathbf{x}_0,1}}) + 2\bar{x}^{us}. \tag{C.63}$$

- When $q=2,3,\dots$, given event $\mathcal{D}_{q-1}$, i.e. $\|\mathbf{x}_0^{(q)}\|^2 \le \frac{18\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2}\log(\frac{1}{\delta_{\mathbf{x}_0,q-1}}) + 2\bar{x}^{us}$, the above (C.62) gives $\|\mathbf{x}_{T_q}^{(q)}\|^2 \le \frac{18\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2}\log(\frac{1}{\delta_{\mathbf{x}_0,q}}) + 3\bar{\kappa}^2\bar{\theta}^{2T_q}\left(\frac{18\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2}\log(\frac{1}{\delta_{\mathbf{x}_0,q-1}}) + 2\bar{x}^{us}\right) + \bar{x}^{us}$. Similarly, when $T_q \ge \frac{1}{2\log(1/\bar{\theta})}\log\left(6\bar{\kappa}^2 + \frac{54\bar{\kappa}^4\bar{\sigma}^2}{(1-\bar{\theta})\bar{x}^{us}}\log(\frac{1}{\delta_{\mathbf{x}_0,q-1}})\right)$, we further have

$$\|\mathbf{x}_{T_q}^{(q)}\|^2 \le \frac{18\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2}\log(\frac{1}{\delta_{\mathbf{x}_0,q}}) + 2\bar{x}^{us}. \tag{C.64}$$

Combining (C.63) and (C.64), we can claim the following: for epoch $q=1,2,\dots$, when $T_1 \ge \frac{1}{2\log(1/\bar{\theta})}\log\left(\frac{3n\sqrt{s}\bar{\kappa}^2\bar{\tau}\bar{\sigma}^2}{(1-\bar{\rho})\bar{x}^{us}\delta_{\mathbf{x}_0,0}}\right)$ and $T_q \ge \frac{1}{2\log(1/\bar{\theta})}\log\left(6\bar{\kappa}^2 + \frac{54\bar{\kappa}^4\bar{\sigma}^2}{(1-\bar{\theta})\bar{x}^{us}}\log(\frac{1}{\delta_{\mathbf{x}_0,q-1}})\right)$, we have $\mathbb{P}(\mathcal{D}_q \mid \mathcal{B}_{q-1}, \mathcal{D}_{q-1}) \ge 1-\delta_{\mathbf{x}_0,q}$. □

The following Proposition C.8a says that if a good controller is used in epoch q, then the final state $x_{T_q}^{(q)}$ of epoch q (the initial state of epoch $q+1$) can be bounded.

Proposition C.8a. *Suppose every epoch q has length $T_q \ge \underline{T}_{rgt,\bar{\epsilon}}^{us}(\delta_{id,q}, T_q)$. Then,*

$$\mathbb{P}(\mathcal{B}_q \mid \mathcal{C}_q, \cap_{j=0}^{q-1}\mathcal{E}_j) = \mathbb{P}(\mathcal{B}_q \mid \mathcal{C}_q) = 1 \tag{C.65}$$

Now, we are ready to present the main proof of Theorem 5.2.

Theorem C.12 (Complete version of Theorem 5.2)**.** *Assume that the initial state $\mathbf{x}_0 = 0$, Assumption A5.1 holds, and $\mathbf{L}_{1:s}^\star$ is* uniformly stable. *Suppose $c_{\mathbf{x}} \ge \underline{c}_{\mathbf{x}}(\bar{\rho}, \bar{\tau}), c_{\mathbf{z}} \ge \underline{c}_{\mathbf{z}}$, $T_0 \ge \mathcal{O}(\underline{T}_{rgt}^{us}(\delta, T_0))$. Then, with probability at least $1-\delta$, Algorithm 2 achieves*

$$\text{Regret}(T) \le \mathcal{O}\left(\frac{s^2p(n^2+p^2)\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\log\big(\frac{\log^2(T)}{\delta}\big)\log^2(T)\sqrt{T}\right). \tag{C.66}$$

Proof. The proof is almost the same as the proof for the regret upper bound in Theorem C.11 for the mean-square stability case in Appendix C.4.1, thus we only present the key steps and omit certain details of intermediate steps.

In the following, we set $\delta_{id,q} = \delta_{\mathbf{x}_0,q} = \frac{3}{\pi^2} \cdot \frac{\delta}{(q+1)^2}$. Similar to the counterpart (C.41), event $\mathcal{E}_q = \mathcal{A}_{q+1} \cap \mathcal{B}_q \cap \mathcal{C}_q \cap \mathcal{D}_q$ implies the following: for $q = 1, 2, \dots,$

$$
\begin{aligned}
&\text{Regret}_{q+1} \\
\le & \mathcal{O}(1) \log\big(\frac{(q+1)^2}{\delta}\big) sp\big(\frac{\sigma_{\mathbf{z},q} + \sigma_{\mathbf{w}}}{\sigma_{\mathbf{z},q}\pi_{\min}} \cdot \frac{\sqrt{s}(n+p)\log(T_q)}{\sqrt{T_q}} + \frac{\sqrt{\log(T_q)}}{\pi_{\min}(1-\bar{\rho})\sqrt{T_q}}\big)^2 \sigma_{\mathbf{w}}^2 T_{q+1} \\
& + \mathcal{O}(1)\log\big(\frac{q+1}{\delta}\big)\frac{18\sqrt{ns}\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2} + \mathcal{O}(\frac{n\sqrt{s}}{1-\rho^\star}\sigma_{\mathbf{z},q+1}^2 T_{q+1}) + \mathcal{O}(1) \\
\le & \mathcal{O}(1)\log\big(\frac{(q+1)^2}{\delta}\big)\frac{s^2 p(n^2+p^2)\gamma}{\pi_{\min}^2(1-\bar{\rho})^2}\sigma_{\mathbf{w}}^2\sqrt{T_q}\log^2(T_q) + \mathcal{O}(1)\log\big(\frac{(q+1)^2}{\delta}\big)\frac{18\sqrt{ns}\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2};
\end{aligned}
\tag{C.67}
$$

and for $q = 0$,

$$
\text{Regret}_1 \le \mathcal{O}(1)\log\big(\frac{1}{\delta}\big)\frac{s^2p(n^2+p^2)\gamma}{\pi_{\min}^2(1-\bar{\rho})^2}\sigma_{\mathbf{w}}^2\sqrt{T_0}\log^2(T_0) + \mathcal{O}(1)\big(\frac{1}{\delta}\big)\frac{n^{1.5}s\bar{\tau}\bar{\sigma}^2}{1-\bar{\rho}}. \tag{C.68}
$$

Note that the difference between (C.67) ($q = 1, 2, \dots$) and (C.68) ($q = 0$) is due to the difference between the event $\mathcal{D}_q$ for $q = 1, 2, \dots$ and event $\mathcal{D}_0$. Compared with the mean-square stability counterpart (C.41), we see the $\frac{(q+1)^2}{\delta}$ dependence in (C.41) is now replaced with $\log\big(\frac{q+1}{\delta}\big)$. For all $M := \mathcal{O}(\log_\gamma(\frac{T}{T_0}))$ epochs, similar to the counterpart (C.42), event $\cap_{q=0}^{M-1}\mathcal{E}_q$ implies

$$
\begin{aligned}
&\text{Regret}(T) \\
=& \mathcal{O}(\sum_{q=1}^{M}\text{Regret}_q) \\
\le& \mathcal{O}\Bigg(\frac{s^2p(n^2+p^2)\sigma_{\mathbf{w}}^2}{\pi_{\min}^2(1-\bar{\rho})^2}\log\big(\frac{\log^2(T)}{\delta}\big)\sqrt{T}\log^2(T) + \frac{18\sqrt{ns}\bar{\kappa}^2\bar{\sigma}^2}{(1-\bar{\theta})^2}\log\big(\frac{\log^2(T)}{\delta}\big)\log(T)\Bigg)
\end{aligned}
\tag{C.69}
$$

which shows the main result (C.66). Note that in the above summation, we have omit $\frac{1}{\delta}$ term in Regret_1 since it does not scale with time and can be dominated by the rest.

Now we are only left to show the occurrence probability of regret bound (C.66) is larger than $1 - \delta$. To do this, we will combine Proposition C.7a, C.8a, C.9, and C.10 over all $q = 0, 1, \dots, M-1$. Note that for each individual q, these propositions hold only when certain prerequisite conditions on hyper-parameters $c_{\mathbf{x}}$, $c_{\mathbf{z}}$, and T_0 are satisfied. We first show that under the choices $T_q = \gamma T_{q-1}$, $\sigma_{\mathbf{z},q}^2 = \frac{\sigma_{\mathbf{w}}^2}{\sqrt{T_q}}$, and $\delta_{id,q} = \delta_{\mathbf{x}_0,q} = \frac{3}{\pi^2} \cdot \frac{\delta}{(q+1)^2}$ these hyper-parameter conditions can be satisfied for all $q = 0, 1, \ldots, M-1$.

- Proposition C.7a requires these to hold: $T_0\gamma^q \geq \frac{1}{2\log(1/\bar{\theta})} \log\left(6\bar{\kappa}^2 + \frac{54\bar{\kappa}^4\bar{\sigma}^2}{(1-\bar{\theta})\bar{x}^{us}} \log(\frac{i^2\pi^2}{3\delta})\right)$ and $T_0\gamma \geq \frac{1}{2\log(1/\bar{\theta})} \log\left(\frac{\pi^2 n\sqrt{\bar{s}}\bar{\kappa}^2\bar{\tau}\bar{\sigma}^2}{(1-\bar{\rho})\bar{x}^{us}\delta}\right)$. One can check $T_0 \geq \max\{\frac{54\bar{\kappa}^4\bar{\sigma}^2}{(1-\bar{\theta})\bar{x}^{us}\log(1/\bar{\theta})\log(\gamma)}, \frac{1}{\gamma\log(1/\bar{\theta})}\log(6\bar{\kappa}^2 + \frac{54n\sqrt{\bar{s}}\bar{\kappa}^4\bar{\sigma}^2\log(\pi^2/3\delta)}{(1-\bar{\theta})(1-\bar{\rho})\bar{x}^{us}\delta})\} =: \underline{T}^{us}_{\mathbf{x}_0}(\delta)$ would suffice.
- Proposition C.8a requires that for $q = 0, 1, \ldots,$ condition $T_0\gamma^q \geq \underline{T}^{us}_{rgt,\bar{\epsilon}}(\frac{3\delta}{\pi^2(q+1)^2}, T_0\gamma^q)$ holds, which can be satisfied when one chooses $T_0 \geq \mathcal{O}(\underline{T}^{us}_{rgt,\bar{\epsilon}}(\delta, T_0))$.
- Proposition C.9 requires the following to hold: $c_{\mathbf{x}} \geq \underline{c}_{\mathbf{x}}(\bar{\rho}, \bar{\tau}), c_{\mathbf{z}} \geq \underline{c}_{\mathbf{z}}$, and $T_0\gamma^q \geq \max\left\{\underline{T}_{MC,1}(\frac{3\delta}{8\pi^2 q^2}), \underline{T}_{id,N}(\frac{3\delta}{2\pi^2(q+1)^2}, \bar{\rho}, \bar{\tau})\right\}$. The last one can be satisfied when we have $T_0 \geq \mathcal{O}(\max\{\underline{T}_{MC,1}(\delta), \underline{T}_{id,N}(\delta, \bar{\rho}, \bar{\tau})\})$.
- Proposition C.10 requires no conditions on hyper-parameters.

Therefore, when $c_{\mathbf{x}} \geq \underline{c}_{\mathbf{x}}(\bar{\rho}, \bar{\tau}), c_{\mathbf{z}} \geq \underline{c}_{\mathbf{z}}$, $T_0 \geq \mathcal{O}(\max\{\underline{T}^{us}_{\mathbf{x}_0}(\delta), \underline{T}^{us}_{rgt,\bar{\epsilon}}(\delta, T_0), \underline{T}_{MC,1}(\delta), \underline{T}_{id,N}(\delta, \bar{\rho}, \bar{\tau})\}) =: \mathcal{O}(\underline{T}^{us}_{rgt}(\delta, T_0))$, we can apply Proposition C.7a, C.8a, C.9, and C.10 to every epoch $q = 0, 1, \ldots, M-1$. Similar to (C.45), this gives $\mathbb{P}$ (Regret bounds in (C.66) holds) $\geq \mathbb{P}(\cap_{q=0}^{M-1}\mathcal{E}_q) \geq 1-\delta$. □

C.5.2 Proof for Theorem 5.3

Since Theorem 5.2 shows that $\text{Regret}(T) := \sum_q J_{(q)} - TJ^\star \leq \mathcal{O}(\sqrt{T}\log(\frac{1}{\delta}))$, to upper bound $\text{Regret}^\circ(T) := \sum_q J^\circ_{(q)} - TJ^\star$ in Theorem 5.3, it suffices to upper bound each summand $J^\circ_{(q)} - J_{(q)}$. By definition, we further have

$$J^\circ_{(q)} - J_{(q)} = J^\circ_{(q)} - \mathbb{E}[J^\circ_{(q)} \mid \mathcal{F}_{q-1}] = J^\circ_{(q)} - \mathbb{E}[J^\circ_{(q)} \mid \mathbf{x}_0^{(q)}, \omega^{(q)}(0), \mathbf{K}^{(q)}_{1:s}]\,.$$

Hence, we only need to study the deviation of the random cost $J^\circ_{(q)}$ from its conditional mean $\mathbb{E}[J^\circ_{(q)} \mid \mathbf{x}_0^{(q)}, \omega^{(q)}(0), \mathbf{K}^{(q)}_{1:s}]$. Before presenting this result in Lemma C.17, we first provide several supporting results from high-dimensional statistics. In this section, c denotes an absolute constant.

Lemma C.13 (Theorem 1.1 in Rudelson and Vershynin (2013))**.** *Consider a random vector $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x} \sim \mathcal{N}(0, \mathbf{\Sigma}_{\mathbf{x}})$ and an arbitrary matrix $\mathbf{S} \in \mathbb{R}^{n \times n}$. Then, with probability at least $1-\delta$,*

$$|\mathbf{x}^\intercal \mathbf{S}\mathbf{x} - \mathbb{E}[\mathbf{x}^\intercal \mathbf{S}\mathbf{x}]| \leq c\|\mathbf{\Sigma}_{\mathbf{x}}\|\|\mathbf{S}\|_{\text{F}}\log(\frac{3}{\delta}). \tag{C.70}$$

Lemma C.14 (Proposition 5.10 in Vershynin (2012))**.** *Consider a random vector* $\mathbf{x} \in \mathbb{R}^n$ *such that* $\mathbf{x} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\mathbf{x}})$ *and an arbitrary vector* $\mathbf{a} \in \mathbb{R}^n$*. Then, with probability at least* $1-\delta$*,*

$$|\mathbf{a}^\intercal \mathbf{x}| \leq c\sqrt{\|\boldsymbol{\Sigma}_{\mathbf{x}}\|}\|\mathbf{a}\|\sqrt{\log(\frac{3}{\delta})}. \tag{C.71}$$

Lemma C.15. *Consider two independent random vectors* $\mathbf{x} \in \mathbb{R}^{n_{\mathbf{x}}}, \mathbf{y} \in \mathbb{R}^{n_{\mathbf{y}}}$ *such that* $\mathbf{x} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\mathbf{x}})$ *and* $\mathbf{y} \sim \mathcal{N}(0, \boldsymbol{\Sigma}_{\mathbf{y}})$*, and an arbitrary matrix* $\mathbf{S} \in \mathbb{R}^{n_{\mathbf{x}} \times n_{\mathbf{y}}}$*, then with probability at least* $1-\delta$*,*

$$|\mathbf{x}^\intercal \mathbf{S}\mathbf{y}| \leq c\sqrt{\min\{n_{\mathbf{x}}, n_{\mathbf{y}}\}}\sqrt{\|\boldsymbol{\Sigma}_{\mathbf{x}}\|\|\boldsymbol{\Sigma}_{\mathbf{y}}\|}\|\mathbf{S}\| \log(\frac{6}{\delta}). \tag{C.72}$$

Proof. By Lemma C.14, with probability at least $1-\delta/2$, $\mathbf{x}^\intercal \mathbf{S}\mathbf{y} \leq c\sqrt{\|\mathbf{S}\boldsymbol{\Sigma}_{\mathbf{y}}\mathbf{S}^\intercal\|}\|\mathbf{x}\|\sqrt{\log(\frac{6}{\delta})}$. By Lemma C.13, with probability at least $1-\delta/2$, $\|\mathbf{x}\|^2 \leq \mathbf{tr}(\boldsymbol{\Sigma}_{\mathbf{x}}) + c\|\boldsymbol{\Sigma}_{\mathbf{x}}\|\sqrt{n_{\mathbf{x}}}\log(\frac{6}{\delta})$, which further gives $\|\mathbf{x}\| \leq c\sqrt{n_{\mathbf{x}}\|\boldsymbol{\Sigma}_{\mathbf{x}}\|\log(\frac{6}{\delta})}$. Combining these two results shows $|\mathbf{x}^\intercal \mathbf{S}\mathbf{y}| \leq c\sqrt{n_{\mathbf{x}}}\sqrt{\|\boldsymbol{\Sigma}_{\mathbf{x}}\|\|\boldsymbol{\Sigma}_{\mathbf{y}}\|}\,\|\mathbf{S}\|\log(\frac{6}{\delta})$. Similarly, we can show $|\mathbf{x}^\intercal \mathbf{S}\mathbf{y}| \leq c\sqrt{n_{\mathbf{y}}}\sqrt{\|\boldsymbol{\Sigma}_{\mathbf{x}}\|\|\boldsymbol{\Sigma}_{\mathbf{y}}\|}\|\mathbf{S}\|\log(\frac{6}{\delta})$, which completes the proof. □

Lemma C.16. *Consider a vector* $\mathbf{v} := [\mathbf{v}_1^\intercal, \mathbf{v}_2^\intercal, \mathbf{v}_3^\intercal]^\intercal$ *where* $\mathbf{v}_1 \in \mathbb{R}^{n_1}$ *is deterministic with* $\|\mathbf{v}_1\| \leq \bar{v}_1$*, and* $\mathbf{v}_2 \in \mathbb{R}^{n_2}$*,* $\mathbf{v}_3 \in \mathbb{R}^{n_3}$ *are random vectors such that* $\mathbf{v}_2 \sim \mathcal{N}(0, \boldsymbol{\Sigma}_2)$*,* $\mathbf{v}_3 \sim \mathcal{N}(0, \boldsymbol{\Sigma}_3)$*. Consider an arbitrary symmetric matrix* $\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23} \\ \mathbf{S}_{31} & \mathbf{S}_{32} & \mathbf{S}_{33} \end{bmatrix}$ *where* $\mathbf{S}_{11} \in \mathbb{R}^{n_1 \times n_1}, \mathbf{S}_{22} \in \mathbb{R}^{n_2 \times n_2}, \mathbf{S}_{33} \in \mathbb{R}^{n_3 \times n_3}$*. Then, with probability at least* $1-\delta$*,*

$$\begin{aligned} |\mathbf{v}^\intercal \mathbf{S}\mathbf{v} - \mathbb{E}[\mathbf{v}^\intercal \mathbf{S}\mathbf{v}]| \leq c\Big(\|\boldsymbol{\Sigma}_2\|\|\mathbf{S}_{22}\|_{\mathrm{F}} + \|\boldsymbol{\Sigma}_3\|\|\mathbf{S}_{33}\|_{\mathrm{F}} + \sqrt{\min\{n_2, n_3\}}\sqrt{\|\boldsymbol{\Sigma}_2\|\|\boldsymbol{\Sigma}_3\|}\|\mathbf{S}_{23}\|\Big)\log(\frac{18}{\delta}) \\ + c\Big(\sqrt{\|\boldsymbol{\Sigma}_2\|}\|\mathbf{S}_{12}\| + \sqrt{\|\boldsymbol{\Sigma}_3\|}\|\mathbf{S}_{13}\|\Big)\bar{v}_1\sqrt{\log(\frac{18}{\delta})}. \end{aligned} \tag{C.73}$$

Proof. By triangle inequality,

$$|\mathbf{v}^\intercal \mathbf{S}\mathbf{v} - \mathbb{E}[\mathbf{v}^\intercal \mathbf{S}\mathbf{v}]| \leq d_{22} + d_{33} + 2d_{23} + 2d_{12} + 2d_{23}, \tag{C.74}$$

where $d_{ij} = |\mathbf{v}_i^\intercal \mathbf{S}_{ij}\mathbf{v}_j - \mathbb{E}[\mathbf{v}_i^\intercal \mathbf{S}_{ij}\mathbf{v}_j]|$. Then

- By Lemma C.13, with probability at least $1-\delta/6$, $d_{22} \leq c\|\boldsymbol{\Sigma}_2\|\|\mathbf{S}_{22}\|_{\mathrm{F}}\log(\frac{18}{\delta})$.
- By Lemma C.13, with probability at least $1-\delta/6$, $d_{33} \leq c\|\boldsymbol{\Sigma}_3\|\|\mathbf{S}_{33}\|_{\mathrm{F}}\log(\frac{18}{\delta})$.
- By Lemma C.15, with probability at least $1-\delta/3$, $d_{23} \leq c\sqrt{\min\{n_2, n_3\}\|\boldsymbol{\Sigma}_2\|\|\boldsymbol{\Sigma}_3\|}\|\mathbf{S}_{23}\|\log(\frac{18}{\delta})$.
- By Lemma C.14, with probability at least $1-\delta/6$, $d_{12} \leq c\sqrt{\|\boldsymbol{\Sigma}_2\|}\|\mathbf{S}_{12}\|\bar{v}_1\sqrt{\log(\frac{18}{\delta})}$.

- By Lemma C.14, with probability at least $1-\delta/6$, $d_{13} \leq c\sqrt{\|\mathbf{\Sigma}_3\|}\|\mathbf{S}_{13}\|\bar{v}_1\sqrt{\log(\frac{18}{\delta})}$.

Combining these with the union bound concludes the proof. □

With Lemma C.16, we can analyze the concentration of the MJS-LQR cumulative cost around its mean under uniform stability.

Lemma C.17. *Consider* MJS-LQR$(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T}, \mathbf{Q}_{1:s}, \mathbf{R}_{1:s})$ *with process noise* $\mathcal{N}(0, \sigma_{\mathbf{w}}^2 \mathbf{I})$*, given initial mode* ω_0 *and initial state* $\mathbf{x}_0$ *such that* $\|\mathbf{x}_0\| \leq \bar{x}_0$*. For a controller* $\mathbf{K}_{1:s}$*, the input is given by* $\mathbf{u}_t = \mathbf{K}_{\omega_t}\mathbf{x}_t + \mathbf{z}_t$ *where* $\mathbf{z}_t \sim \mathcal{N}(0, \sigma_{\mathbf{z}}^2 \mathbf{I})$*. Let* $\mathbf{L}_i = \mathbf{A}_i + \mathbf{B}_i\mathbf{K}_i$ *for all* i*. Assume there exists* $\kappa \geq 1$ *and* $\theta \in [0,1)$ *such that for any sequence* $\omega_{1:l} \in [s]^l$ *with any* $l \in \mathbb{N}$ *such that* $\|\prod_{j=1}^l \mathbf{L}_{\omega_j}\| \leq \kappa\theta^l$*. Let* $J_T = \sum_{t=0}^T \mathbf{x}_t^\intercal \mathbf{Q}_{\omega_t}\mathbf{x}_t + \mathbf{u}_t^\intercal \mathbf{R}_{\omega_t}\mathbf{u}_t$ *denote the cumulative cost over time horizon* T*. Then, with probability at least* $1-\delta$*,*

$$|J_T - \mathbb{E}[J_T \mid \omega_0, \mathbf{x}_0, \mathbf{K}_{1:s}]| \leq \frac{c(np)^{1.5}\kappa^2}{(1-\theta)^2}\Big[(\gamma_1\sigma_{\mathbf{w}}^2 + \gamma_2\sigma_{\mathbf{z}}^2 + \gamma_3\sigma_{\mathbf{w}}\sigma_{\mathbf{z}})\sqrt{T}\log(\frac{18}{\delta}) + (\gamma_1\sigma_{\mathbf{w}} + \gamma_3\sigma_{\mathbf{z}})\bar{x}_0\sqrt{\log(\frac{18}{\delta})}\Big], \quad \text{(C.75)}$$

where $\gamma_1 := \|\mathbf{M}_{1:s}\|$ *for* $\mathbf{M}_i := \mathbf{Q}_i + \mathbf{K}_i^\intercal \mathbf{R}_i \mathbf{K}_i$*,* $\gamma_2 := \|\mathbf{M}_{1:s}\|\|\mathbf{B}_{1:s}\| + \|\mathbf{R}_{1:s}\|\|\mathbf{K}_{1:s}\|$*, and* $\gamma_3 := \|\mathbf{M}_{1:s}\|\|\mathbf{B}_{1:s}\|^2 + 2\|\mathbf{B}_{1:s}\|\|\mathbf{R}_{1:s}\|\|\mathbf{K}_{1:s}\| + \|\mathbf{R}_{1:s}\|$*.*

Proof. First we define a few notations that can convert J_T into the form of vector-matrix multiplications. Define the block-diagonal matrix $\mathbf{K}$ with $T+1$ diagonal blocks such that the t-th block is given by $\mathbf{K}_{\omega_{t-1}}$ for all t. Similarly, define $\mathbf{Q}$ for $\mathbf{Q}_{\omega_0:\omega_T}$, $\mathbf{R}$ for $\mathbf{R}_{\omega_0:\omega_T}$, and $\mathbf{M}$ for $\mathbf{M}_{\omega_0:\omega_T}$. For all t, define

$$\begin{aligned}
&\mathbf{G}_{0,0}^{(0)} := \mathbf{I}_n, \quad \mathbf{G}_{0,t}^{(0)} := \prod_{h=0}^{t-1}\mathbf{L}_{\omega_h};\\
&\mathbf{G}_{t,t}^{(\mathbf{w})} := 0, \quad \mathbf{G}_{t-1,t}^{(\mathbf{w})} := \mathbf{I}_n, \quad \mathbf{G}_{r,t}^{(\mathbf{w})} := \prod_{h=r+1}^{t-1}\mathbf{L}_{\omega_h}, \forall r \leq t-2;\\
&\mathbf{G}_{t,t}^{(\mathbf{z})} := 0, \quad \mathbf{G}_{t-1,t}^{(\mathbf{z})} := \mathbf{B}_{\omega_{t-1}}, \quad \mathbf{G}_{r,t}^{(\mathbf{z})} := (\prod_{h=r+1}^{t-1}\mathbf{L}_{\omega_h})\mathbf{B}_{\omega_r}, \forall r \leq t-2
\end{aligned} \quad \text{(C.76)}$$

Then, it is easy to derive that

$$\mathbf{x}_t = \mathbf{G}_{0,t}^{(0)}\mathbf{x}_0 + \sum_{r=0}^t \mathbf{G}_{r,t}^{(\mathbf{w})}\mathbf{w}_r + \sum_{r=0}^t \mathbf{G}_{r,t}^{(\mathbf{z})}\mathbf{z}_r \quad \text{(C.77)}$$

and $\|\mathbf{G}_{0,t}\| \leq \kappa\theta^t$, $\|\mathbf{G}_{r,t}^{(\mathbf{w})}\| \leq \kappa\theta^{t-r-1}$, and $\|\mathbf{G}_{r,t}^{(\mathbf{z})}\| \leq \kappa\theta^{t-r-1}\|\mathbf{B}_{1:s}\|$. Define the following vectors by concatenation.

$$\begin{aligned}\mathbf{x} &:= [\mathbf{x}_0^\intercal, \mathbf{x}_1^\intercal, \dots, \mathbf{x}_T^\intercal]^\intercal,\ \mathbf{u} := [\mathbf{u}_0^\intercal, \mathbf{u}_1^\intercal, \dots, \mathbf{u}_T^\intercal]^\intercal,\ \boldsymbol{\phi} := [\mathbf{x}^\intercal, \mathbf{u}^\intercal]^\intercal,\\ \mathbf{w} &:= [\mathbf{w}_0^\intercal, \mathbf{w}_1^\intercal, \dots, \mathbf{w}_T^\intercal]^\intercal,\ \mathbf{z} := [\mathbf{z}_0^\intercal, \mathbf{z}_1^\intercal, \dots, \mathbf{z}_T^\intercal]^\intercal,\ \mathbf{v} := [\mathbf{x}_0^\intercal, \mathbf{w}^\intercal, \mathbf{z}^\intercal]^\intercal,\end{aligned} \tag{C.78}$$

Define the following block matrices.

$$\mathbf{G}^{(0)} := \begin{bmatrix} \mathbf{G}^{(0)}_{0,0} \\ \mathbf{G}^{(0)}_{0,1} \\ \vdots \\ \mathbf{G}^{(0)}_{0,T} \end{bmatrix},\ \mathbf{G}^{(\mathbf{w})} := \begin{bmatrix} \mathbf{G}^{(\mathbf{w})}_{0,0} & & & \\ \mathbf{G}^{(\mathbf{w})}_{0,1} & \mathbf{G}^{(\mathbf{w})}_{1,1} & & \\ \vdots & & \ddots & \\ \mathbf{G}^{(\mathbf{w})}_{0,T} & \mathbf{G}^{(\mathbf{w})}_{1,T} & \dots & \mathbf{G}^{(\mathbf{w})}_{T,T} \end{bmatrix},\ \mathbf{G}^{(\mathbf{z})} := \begin{bmatrix} \mathbf{G}^{(\mathbf{z})}_{0,0} & & & \\ \mathbf{G}^{(\mathbf{z})}_{0,1} & \mathbf{G}^{(\mathbf{z})}_{1,1} & & \\ \vdots & & \ddots & \\ \mathbf{G}^{(\mathbf{z})}_{0,T} & \mathbf{G}^{(\mathbf{z})}_{1,T} & \dots & \mathbf{G}^{(\mathbf{z})}_{T,T} \end{bmatrix},$$
$$\mathbf{G} := [\mathbf{G}^{(0)}, \mathbf{G}^{(\mathbf{w})}, \mathbf{G}^{(\mathbf{z})}], \quad \tilde{\mathbf{I}} := [0_{(T+1)p\times(T+2)n}, \mathbf{I}_{(T+1)p}]. \tag{C.79}$$

One can see $\mathbf{x} = \mathbf{G}\mathbf{v}$, $\mathbf{u} = \mathbf{K}\mathbf{x} + \mathbf{z} = (\mathbf{K}\mathbf{G} + \tilde{\mathbf{I}})\mathbf{v}$, $\boldsymbol{\phi} = \begin{bmatrix} \mathbf{G} \\ \mathbf{K}\mathbf{G} + \tilde{\mathbf{I}} \end{bmatrix}\mathbf{v}$, and $J_T = \boldsymbol{\phi}^\intercal \begin{bmatrix} \mathbf{Q} & \\ & \mathbf{R} \end{bmatrix} \boldsymbol{\phi}$. Let $\mathbf{S} := \mathbf{G}^\intercal\mathbf{Q}\mathbf{G} + (\mathbf{K}\mathbf{G} + \tilde{\mathbf{I}})^\intercal\mathbf{R}(\mathbf{K}\mathbf{G} + \tilde{\mathbf{I}})$, then these relations give

$$J_T = \mathbf{v}^\intercal\mathbf{S}\mathbf{v}. \tag{C.80}$$

Block-partition matrix $\mathbf{S}$ by $\mathbf{S} = \begin{bmatrix} \mathbf{S}^{(0,0)} & \mathbf{S}^{(0,\mathbf{w})} & \mathbf{S}^{(0,\mathbf{z})} \\ \mathbf{S}^{(\mathbf{w},0)} & \mathbf{S}^{(\mathbf{w},\mathbf{w})} & \mathbf{S}^{(\mathbf{w},\mathbf{z})} \\ \mathbf{S}^{(\mathbf{z},0)} & \mathbf{S}^{(\mathbf{z},\mathbf{w})} & \mathbf{S}^{(\mathbf{z},\mathbf{z})} \end{bmatrix}$ such that $\mathbf{S}^{(0,0)} \in \mathbb{R}^{n\times n}$, $\mathbf{S}^{(\mathbf{w},\mathbf{w})} \in \mathbb{R}^{(T+1)n\times(T+1)n}$, $\mathbf{S}^{(\mathbf{z},\mathbf{z})} \in \mathbb{R}^{(T+1)p\times(T+1)p}$. Then, we have

$$\begin{aligned} \mathbf{S}^{(0,0)} &= {\mathbf{G}^{(0)}}^\intercal\mathbf{M}\mathbf{G}^{(0)} \\ \mathbf{S}^{(\mathbf{w},\mathbf{w})} &= {\mathbf{G}^{(\mathbf{w})}}^\intercal\mathbf{M}\mathbf{G}^{(\mathbf{w})},\quad \mathbf{S}^{(\mathbf{z},\mathbf{z})} = {\mathbf{G}^{(\mathbf{z})}}^\intercal\mathbf{M}\mathbf{G}^{(\mathbf{z})} + \mathbf{R}\mathbf{K}\mathbf{G}^{(\mathbf{z})} + {\mathbf{G}^{(\mathbf{z})}}^\intercal\mathbf{K}^\intercal\mathbf{R} + \mathbf{R}, \\ \mathbf{S}^{(\mathbf{w},\mathbf{z})} &= {\mathbf{G}^{(\mathbf{w})}}^\intercal\mathbf{M}\mathbf{G}^{(\mathbf{z})} + {\mathbf{G}^{(\mathbf{w})}}^\intercal\mathbf{K}^\intercal\mathbf{R}, \\ \mathbf{S}^{(0,\mathbf{w})} &= {\mathbf{G}^{(0)}}^\intercal\mathbf{M}\mathbf{G}^{(\mathbf{w})},\quad \mathbf{S}^{(0,\mathbf{z})} = {\mathbf{G}^{(0)}}^\intercal\mathbf{M}\mathbf{G}^{(\mathbf{z})} + {\mathbf{G}^{(0)}}^\intercal\mathbf{K}^\intercal\mathbf{R}. \end{aligned} \tag{C.81}$$

Matrices $\mathbf{G}^{(0)}, \mathbf{G}^{(\mathbf{w})}, \mathbf{G}^{(\mathbf{z})}$ can be bounded as follows.

$$\begin{aligned} &\|\mathbf{G}^{(0)}\| \le \|\mathbf{G}^{(0)}\|_{\mathrm{F}} \le \sqrt{\sum_{i=0}^{T} \kappa^2\theta^{2i}} \le \frac{\kappa}{1-\theta}. \\ &\|\mathbf{G}^{(\mathbf{w})}\| \le \sqrt{\|\mathbf{G}^{(\mathbf{w})}\|_1\|\mathbf{G}^{(\mathbf{w})}\|_\infty} \le \sqrt{\frac{\sqrt{n}\kappa}{1-\theta}\cdot\frac{\sqrt{n}\kappa}{1-\theta}} = \frac{\sqrt{n}\kappa}{1-\theta}. \\ &\|\mathbf{G}^{(\mathbf{z})}\| \le \sqrt{\|\mathbf{G}^{(\mathbf{z})}\|_1\|\mathbf{G}^{(\mathbf{z})}\|_\infty} \le \sqrt{\frac{\sqrt{n}\kappa\|\mathbf{B}_{1:s}\|}{1-\theta}\cdot\frac{\sqrt{p}\kappa\|\mathbf{B}_{1:s}\|}{1-\theta}} = \frac{(np)^{0.25}\kappa\|\mathbf{B}_{1:s}\|}{1-\theta}. \end{aligned} \tag{C.82}$$

These results further give

$$\begin{aligned} &\|\mathbf{S}^{(\mathbf{w},\mathbf{w})}\| \leq \frac{n\kappa^2\gamma_1}{(1-\theta)^2}, \quad \|\mathbf{S}^{(\mathbf{z},\mathbf{z})}\| \leq \frac{(np)^{0.5}\kappa^2\gamma_2}{(1-\theta)^2}, \quad \|\mathbf{S}^{(\mathbf{w},\mathbf{z})}\| \leq \frac{n^{0.75}p^{0.25}\kappa^2\gamma_3}{(1-\theta)^2}, \\ &\|\mathbf{S}^{(0,\mathbf{w})}\| \leq \frac{n^{0.5}\kappa^2\gamma_1}{(1-\theta)^2}, \quad \|\mathbf{S}^{(0,\mathbf{z})}\| \leq \frac{(np)^{0.25}\kappa^2\gamma_3}{(1-\theta)^2}. \end{aligned} \tag{C.83}$$

Finally, we can conclude the proof by invoking Lemma C.16. □

Now, we are ready to present the main proof of Theorem 5.3.

Proof. Following from Lemma C.2a, Proposition C.7a, and the proof of Theorem C.12, we know with probability at least $1-\delta/2$, for all epochs q,

$$\begin{aligned} &\|\mathbf{K}_{1:s}^{(q)}\| \leq 2\|\mathbf{K}_{1:s}^{\star}\|, \\ &\|\prod_{j=1}^{l}\mathbf{L}_{\omega_j}^{(q)}\| \leq \kappa\theta^l, \quad \forall\omega_{1:l}\in[s]^l, \forall l\in\mathbb{N}, \\ &\|\mathbf{x}_0^{(q)}\| \leq \mathcal{O}\left(\sqrt{\frac{\bar{\kappa}\bar{\sigma}^2}{(1-\bar{\boldsymbol{\theta}})^2}\log(\frac{q^2}{\delta})}\right), \quad (q\geq 2). \end{aligned} \tag{C.84}$$

Under these conditions, and applying Lemma C.17, we know for epoch q with probability at least $1-\frac{3}{\pi^2}\cdot\frac{\delta}{q^2}$,

$$\begin{aligned} &\left|J^{\circ}_{(q)} - J_{(q)}\right| = \left|J^{\circ}_{(q)} - \mathbb{E}[J^{\circ}_{(q)} \mid \mathcal{F}_{q-1}]\right| \\ \leq&\mathcal{O}\left(\frac{(np)^{1.5}\bar{\kappa}^2}{(1-\bar{\theta})^2}\left[(\sigma_{\mathbf{w}}^2+\sigma_{\mathbf{z},q}^2)\sqrt{T_q}\log(\frac{q^2}{\delta}) + (\sigma_{\mathbf{w}}+\sigma_{\mathbf{z},q})\|\mathbf{x}_0^{(q)}\|\sqrt{\log(\frac{q^2}{\delta})}\right]\right) \\ \leq&\mathcal{O}\left(\frac{(np)^{1.5}\bar{\kappa}^2}{(1-\bar{\theta})^2}\left[\sigma_{\mathbf{w}}^2\sqrt{\gamma}^q\log(\frac{q^2}{\delta}) + \sigma_{\mathbf{w}}^2\frac{\sqrt{\bar{\kappa}}}{1-\bar{\theta}}\log(\frac{q^2}{\delta})\right]\right) \end{aligned} \tag{C.85}$$

where the second line follows from $\sigma_{\mathbf{z},q}^2 = \frac{\sigma_{\mathbf{w}}^2}{\sqrt{T_q}}$, $T_q = \mathcal{O}(\gamma^q)$, and the bound of $\|\mathbf{x}_0^{(q)}\|$ in (C.84). Taking the summation over all $M=\mathcal{O}(\log(T))$ epochs (for simplicity, epoch 0 and 1 are ignored) and applying the union bound, we obtain with probability $1-\delta$,

$$\left|\sum_q J^{\circ}_{(q)} - J_{(q)}\right| \leq \left(\frac{(np)^{1.5}\bar{\kappa}^2\sigma_{\mathbf{w}}^2}{(1-\bar{\theta})^2}\left[\sqrt{T}\log(\frac{\log^2(T)}{\delta}) + \frac{\sqrt{\bar{\kappa}}}{1-\bar{\theta}}\log(\frac{\log^2(T)}{\delta})\right]\right). \tag{C.86}$$

Combining this with the upper bound on $\text{Regret}(T) := \sum_q J^{\circ}_{(q)} - TJ^{\star}$ provided in Theorem 5.2 completes the proof. □

Appendix D
Proofs for Results in Chapter 6

D.1 Aggregatable Clustering — Proof for Theorem 6.4

We first provide several supporting lemmas. The first one is regarding the perturbation of the left singular vector space.

Lemma D.1 (Singular Vectors Perturbation Bound)**.** *Consider two arbitrary matrices* $\bar{\mathbf{\Phi}}, \mathbf{\Phi} \in \mathbb{R}^{s \times r}$. *Let* $\bar{\mathbf{U}}, \mathbf{U} \in \mathbb{R}^{s \times r}$ *respectively denote the top-r left singular vectors of* $\bar{\mathbf{\Phi}}$ *and* $\mathbf{\Phi}$ *with* $\bar{\mathbf{U}}^\intercal \bar{\mathbf{U}} = \mathbf{U}^\intercal \mathbf{U} = \mathbf{I}_r$. *Then*

$$\min_{\mathbf{O} \in \mathcal{O}(r)} \|\bar{\mathbf{U}}\mathbf{O} - \mathbf{U}\|_{\mathrm{F}} \leq \frac{2\sqrt{2}\|\bar{\mathbf{\Phi}} - \mathbf{\Phi}\|_{\mathrm{F}}}{\sigma_r(\bar{\mathbf{\Phi}}) - \sigma_{r+1}(\bar{\mathbf{\Phi}})}, \tag{D.1}$$

where $\mathcal{O}(r)$ *denotes the set of all* r×r *orthonormal matrices.*

This result can be seen simply by combining Lemma 10 and Lemma 11 in Du et al. (2019b), where Lemma 10 requires a trivial generalization from spectral norm to the Frobenius norm. The next result says if a matrix has certain rows being identical, its singular vectors share the same identity pattern.

Lemma D.2 (Lemma 12 in Du et al. (2019b))**.** *Consider a matrix* $\bar{\mathbf{\Phi}} \in \mathbb{R}^{s \times r}$ *and a partition* $\Omega_{1:r}$ *on* $[s]$ *such that for any* $i, i' \in \Omega_k$, $\bar{\mathbf{\Phi}}(i,:) = \bar{\mathbf{\Phi}}(i',:)$. *Assume* $\mathrm{rank}(\bar{\mathbf{\Phi}}) = r$. *Let* $\bar{\mathbf{U}} \in \mathbb{R}^{s \times r}$ *denote the top-r left singular vectors of* $\bar{\mathbf{\Phi}}$ *with* $\bar{\mathbf{U}}^\intercal \bar{\mathbf{U}} = \mathbf{I}_r$. *Then for any* $i \in \Omega_k$ *and* $j \in \Omega_l$, $\|\bar{\mathbf{U}}(i,:) - \bar{\mathbf{U}}(j,:)\| = (\frac{1}{|\Omega_k|} + \frac{1}{|\Omega_l|})^{0.5}$ *if* $k \neq l$ *and* 0 *if* $k = l$.

The next lemma provides a preliminary result on the performance of k-means when it is applied to a data matrix with feature dimension same as the number of clusters.

Lemma D.3 (Lemma 5.3 in Lei and Rinaldo (2015))**.** *Consider two arbitrary matrices* $\bar{\mathbf{U}}, \mathbf{U} \in \mathbb{R}^{s \times r}$ *with* $\Delta_{\mathbf{U}} := \|\bar{\mathbf{U}} - \mathbf{U}\|_{\mathrm{F}}$. *Suppose there exists a partition* $\Omega_{1:r}$ *on* $[s]$ *such that for any* $i, i' \in \Omega_k$, $\bar{\mathbf{U}}(i,:) = \bar{\mathbf{U}}(i',:)$. *Define the inter-cluster distance for cluster* k *as*

$\delta_k := \min_{l\in[r]\setminus k} \min_{i\in\Omega_k, j\in\Omega_l} \|\bar{\mathbf{U}}(i,:) - \bar{\mathbf{U}}(j,:)\|$. *Let* $\{\hat{\Omega}_{1:r}, \hat{c}_{1:r}\}$ *be a* $(1+\epsilon)$ *solution to the k-means problem on the rows of* $\mathbf{U}$. *Then, when* $\Delta_{\mathbf{U}} \leq \frac{\min_k \sqrt{|\Omega_k|}\delta_k}{\sqrt{8(2+\epsilon)}}$, *we have*

$$\min_{h\in\mathcal{H}} \sum_{k\in[r]} |\{i : i \in \Omega_k, i \notin \hat{\Omega}_{h(k)}\}| \cdot \delta_k^2 \leq 8(2+\epsilon)\Delta_{\mathbf{U}}^2, \tag{D.2}$$

where $\mathcal{H}$ *is the set of all bijections from* $[r]$ *to* $[r]$.

By combining Lemma D.1, D.2, and D.3, we obtain guarantee on the performance of k-means when it is applied to the left singular vectors of the data matrix, which is the key lemma we will use to show Theorem 6.4 and Theorem 6.6.

Lemma D.4 (Approximate k-means error bound)**.** *Consider two arbitrary matrices* $\bar{\mathbf{\Phi}}, \mathbf{\Phi} \in \mathbb{R}^{s\times r}$ *with* $\Delta_{\mathbf{\Phi}} := \|\bar{\mathbf{\Phi}} - \mathbf{\Phi}\|_{\mathrm{F}}$. *Suppose there exists a partition* $\Omega_{1:r}$ *on* $[s]$ *such that for any* $i, i' \in \Omega_k$, $\bar{\mathbf{\Phi}}(i,:) = \bar{\mathbf{\Phi}}(i',:)$. *Assume* $\mathrm{rank}(\bar{\mathbf{\Phi}}) = r$. *Let* $\mathbf{U} \in \mathbb{R}^{s\times r}$ *denote the top-r left singular vectors of* $\mathbf{\Phi}$ *with* $\mathbf{U}^\intercal\mathbf{U} = \mathbf{I}_r$. *Let* $\{\hat{\Omega}_{1:r}, \hat{c}_{1:r}\}$ *be a* $(1+\epsilon)$ *solution to the k-means problem on clustering the rows of* $\mathbf{U}$. *Then, when* $\Delta_{\mathbf{\Phi}} \leq \frac{\sigma_r(\bar{\mathbf{\Phi}})\sqrt{|\Omega_{(r)}|+|\Omega_{(1)}|}}{8\sqrt{(2+\epsilon)|\Omega_{(1)}|}}$, *we have* $\mathrm{MR}(\hat{\Omega}_{1:r}) \leq \frac{64(2+\epsilon)}{\sigma_r(\bar{\mathbf{\Phi}})^2}\Delta_{\mathbf{\Phi}}^2$.

Proof. Let $\bar{\mathbf{U}} \in \mathbb{R}^{s\times r}$ denote the top-r left singular vectors of $\bar{\mathbf{\Phi}}$ with $\bar{\mathbf{U}}^\intercal\bar{\mathbf{U}} = \mathbf{I}_r$. Then, Lemma D.1 implies that exists $\mathbf{O}^\star \in \mathcal{O}(r)$ such that $\|\bar{\mathbf{U}}\mathbf{O}^\star - \mathbf{U}\|_{\mathrm{F}} \leq \frac{2\sqrt{2}\Delta_\Phi}{\sigma_r(\bar{\mathbf{\Phi}})}$. Note that $\|[\bar{\mathbf{U}}\mathbf{O}^\star](i,:) - [\bar{\mathbf{U}}\mathbf{O}^\star](j,:)\| = \|(\bar{\mathbf{U}}(i,:) - \bar{\mathbf{U}}(j,:))\mathbf{O}^\star\| = \|\bar{\mathbf{U}}(i,:) - \bar{\mathbf{U}}(j,:)\|$. By Lemma D.2, we know for any $i \in \Omega_k$, $j \in \Omega_l$, $\|[\bar{\mathbf{U}}\mathbf{O}^\star](i,:) - [\bar{\mathbf{U}}\mathbf{O}^\star](j,:)\| = \sqrt{\frac{1}{|\Omega_k|} + \frac{1}{|\Omega_l|}}$ if $k \neq l$ and 0 if $k = l$. Then, for any $k \in [r]$, let $\delta_k := \min_{l\in[r]\setminus k} \min_{i\in\Omega_k, j\in\Omega_l} \|[\bar{\mathbf{U}}\mathbf{O}^\star](i,:) - [\bar{\mathbf{U}}\mathbf{O}^\star](j,:)\|$, we see $\delta_k \geq \sqrt{\frac{1}{|\Omega_k|} + \frac{1}{|\Omega_{(1)}|}}$.

Note that when $\Delta_{\mathbf{\Phi}} \leq \frac{\sigma_r(\bar{\mathbf{\Phi}})\sqrt{|\Omega_{(r)}|+|\Omega_{(1)}|}}{8\sqrt{(2+\epsilon)|\Omega_{(1)}|}}$, one can check $\|\bar{\mathbf{U}}\mathbf{O}^\star - \mathbf{U}\|_{\mathrm{F}} \leq \frac{\min_k \sqrt{|\Omega_k|}\delta_k}{\sqrt{8(2+\epsilon)}}$. Then, by Lemma D.3, we obtain that $\mathrm{MR}(\hat{\Omega}_{1:r}) = \min_{h\in\mathcal{H}} \sum_{k=1}^r |\{i : i \in \Omega_k, i \notin \hat{\Omega}_{h(k)}\}| \frac{1}{|\Omega_k|} \leq \min_{h\in\mathcal{H}} \sum_{k=1}^r |\{i : i \in \Omega_k, i \notin \hat{\Omega}_{h(k)}\}|\delta_k^2 \leq 64(2+\epsilon)\sigma_r(\bar{\mathbf{\Phi}})^{-2}\Delta_{\mathbf{\Phi}}^2$. ☐

Main Proof for Theorem 6.4. Consider $\mathbf{\Phi}$ in Algorithm 3 Line 3 and its averaged version $\bar{\mathbf{\Phi}}$ defined in Section 6.4.1. By definition, we have $\|\bar{\mathbf{\Phi}} - \mathbf{\Phi}\|_{\mathrm{F}}^2 = \sum_{k\in[r]} \sum_{i\in\Omega_k} \|\bar{\mathbf{\Phi}}(i,:) - \mathbf{\Phi}(i,:)\|_{\mathrm{F}}^2 = \alpha_{\mathbf{T}}^2 \cdot \sum_{k\in[r]} \sum_{i\in\Omega_k} \|\mathbf{T}(i,:) - |\Omega_k|^{-1} \cdot \sum_{i'\in\Omega_k} \mathbf{T}(i',:)\|^2 + \alpha_{\mathbf{A}}^2 \cdot \sum_{k\in[r]} \sum_{i\in\Omega_k} \|\mathbf{A}_i - |\Omega_k|^{-1} \cdot \sum_{i'\in\Omega_k} \mathbf{A}_{i'}\|_{\mathrm{F}}^2 + \alpha_{\mathbf{B}}^2 \cdot \sum_{k\in[r]} \sum_{i\in\Omega_k} \|\mathbf{B}_i - |\Omega_k|^{-1} \sum_{i'\in\Omega_k} \mathbf{B}_{i'}\|_{\mathrm{F}}^2$. By the definitions of $\epsilon_{\mathbf{A}}, \epsilon_{\mathbf{B}}, \epsilon_{\mathbf{T}}$ in Problem P6.2, triangle inequality ,and Cauchy-Schwarz inequality, we have $\|\bar{\mathbf{\Phi}} - \mathbf{\Phi}\|_{\mathrm{F}} \leq \epsilon_{Agg}$ where $\epsilon_{Agg} := \sqrt{\alpha_{\mathbf{A}}^2\epsilon_{\mathbf{A}}^2 + \alpha_{\mathbf{B}}^2\epsilon_{\mathbf{B}}^2 + \alpha_{\mathbf{T}}^2\epsilon_{\mathbf{T}}^2}$. By construction, in matrix $\bar{\mathbf{\Phi}}$, rows that belong to the same cluster are identical, thus we can apply Lemma D.4 to $\{\bar{\mathbf{\Phi}}, \mathbf{\Phi}\}$ and obtain that when $\epsilon_{Agg} \leq \frac{\sigma_r(\bar{\mathbf{\Phi}})\sqrt{|\Omega_{(r)}|+|\Omega_{(1)}|}}{8\sqrt{(2+\epsilon)|\Omega_{(1)}|}}$, we have $\mathrm{MR}(\hat{\Omega}_{1:r}) \leq 64(2+\epsilon)\sigma_r(\bar{\mathbf{\Phi}})^{-2}\epsilon_{Agg}^2$. ☐

D.2 Lumpable Clustering — Proof for Theorem 6.6

We first provide a supporting result regarding the perturbation of stationary distribution of Markov chains.

Lemma D.5 (Section 3.6 in Cho and Meyer (2001))**.** *For two Markov matrices* $\mathbf{T}, \mathbf{T}_0 \in \mathbb{R}^{s \times s}$ *and their stationary distributions* $\boldsymbol{\pi}, \boldsymbol{\pi}_0 \in \mathbb{R}^s$*, we have* $\|\boldsymbol{\pi} - \boldsymbol{\pi}_0\|_1 \leq \gamma_1 \|\mathbf{T} - \mathbf{T}_0\|_\infty$*, where* $\gamma_1 := \sum_{i=2}^{s} \frac{1}{1-\lambda_i(\mathbf{T})}$.

When the difference $\|\mathbf{T} - \mathbf{T}_0\|$ is small, we further have the following corollary.

Corollary D.6. *In Lemma D.5, let* $\pi_{\min} := \min_i \boldsymbol{\pi}(i), \pi_{\max} := \max_i \boldsymbol{\pi}(i)$*. Suppose that* $\|\mathbf{T} - \mathbf{T}_0\|_\infty \leq \frac{\pi_{\min}}{\gamma_1}$*, then we have*

$$\max_i |\boldsymbol{\pi}(i) - \boldsymbol{\pi}_0(i)| \leq \frac{\pi_{\min}}{2}, \tag{D.3}$$

$$\min_i \boldsymbol{\pi}_0(i) \geq \frac{\pi_{\min}}{2}, \quad \max_i \boldsymbol{\pi}_0(i) \leq \pi_{\max} + \frac{\pi_{\min}}{2} \tag{D.4}$$

$$\max_i |\boldsymbol{\pi}(i)^{-\frac{1}{2}} - \boldsymbol{\pi}_0(i)^{-\frac{1}{2}}| \leq (\sqrt{2}-1)\gamma_1 \pi_{\min}^{-\frac{3}{2}} \|\mathbf{T} - \mathbf{T}_0\|_\infty \tag{D.5}$$

$$\max_i |\boldsymbol{\pi}(i)^{\frac{1}{2}} - \boldsymbol{\pi}_0(i)^{\frac{1}{2}}| \leq (1 - \frac{\sqrt{2}}{2})\gamma_1 \pi_{\min}^{-\frac{1}{2}} \|\mathbf{T} - \mathbf{T}_0\|_\infty. \tag{D.6}$$

Proof. Since $\mathbf{1}^\intercal \boldsymbol{\pi} = \mathbf{1}^\intercal \boldsymbol{\pi}_0 = 1$, we have $\max_i |\boldsymbol{\pi}(i) - \boldsymbol{\pi}_0(i)| \leq \frac{1}{2}\|\boldsymbol{\pi} - \boldsymbol{\pi}_0\|_1 \leq \frac{\gamma_1}{2}\|\mathbf{T} - \mathbf{T}_0\|_\infty \leq \frac{\pi_{\min}}{2}$. Then using triangle inequality, we can show (D.3) and (D.4). Note that the LHS of (D.5) is equivalent to $\max_i \frac{|\boldsymbol{\pi}_0(i) - \boldsymbol{\pi}(i)|}{\sqrt{\boldsymbol{\pi}(i)\boldsymbol{\pi}_0(i)}(\sqrt{\boldsymbol{\pi}(i)} + \sqrt{\boldsymbol{\pi}_0(i)})}$, then plugging in (D.4) gives (D.5). And (D.6) follows similarly. □

When the lumpability perturbation $\epsilon_{\mathbf{T}} \neq 0$, matrix $\mathbf{S}_r$ in Algorithm 3 Line 7 no longer has the row identity pattern as discussed in Lemma 6.5. The next result measures this effect.

Lemma D.7. *Consider an ergodic Markov matrix* $\mathbf{T} \in \mathbb{R}^{s \times s}$ *with stationary distribution* $\boldsymbol{\pi}$ *and a partition* $\Omega_{1:r}$ *such that it is approximately lumpable as in (6.3) with perturbation* $\epsilon_{\mathbf{T}}$*. Consider the neighborhood of* $\mathbf{T}$ *given by* $\mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_{\mathbf{T}})$ *defined in (6.9). Assume there exists an ergodic and reversible* $\mathbf{T}_0 \in \mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_{\mathbf{T}})$ *that has informative spectrum. Construct* $\mathbf{S}_r \in \mathbb{R}^{s \times r}$ *with* $\mathbf{T}$ *and* $\boldsymbol{\pi}$ *as in Algorithm 3 Line 7. Construct* $\bar{\mathbf{S}}_r \in \mathbb{R}^{s \times r}$ *such that for any* $i \in [s]$ *(suppose* $i \in \Omega_k$*),* $\bar{\mathbf{S}}_r(i,:) = \frac{1}{|\Omega_k|}\sum_{i' \in \Omega_k} \mathbf{S}_r(i',:)$*. Let* $\pi_{\min} := \max_i \boldsymbol{\pi}(i)$*,* $\pi_{\max} := \min_i \boldsymbol{\pi}(i)$*,* $\gamma_1 := \sum_{i=2}^{s} \frac{1}{1-\lambda_i(\mathbf{T})}$*,* $\gamma_2 := \min\{\sigma_r(\mathbf{H}) - \sigma_{r+1}(\mathbf{H}), 1\}$*, and* $\gamma_3 := \frac{16\gamma_1\sqrt{r\pi_{\max}}\|\mathbf{T}\|_{\mathrm{F}}}{\gamma_2 \pi_{\min}^2}$ *where* $\mathbf{H}$ *is defined in Algorithm 3. Then, when perturbation* $\epsilon_{\mathbf{T}} \leq \frac{\pi_{\min}}{\gamma_1}$*, we have* $\|\mathbf{S}_r - \bar{\mathbf{S}}_r\|_{\mathrm{F}} \leq \gamma_3 \epsilon_{\mathbf{T}}$.

Proof. We will start with analyzing $\mathbf{T}_0$ and use it as a bridge to prove the claim. Let $\boldsymbol{\pi}_0 \in \mathbb{R}^s$ denote the stationary distribution of $\mathbf{T}_0$. Since $\mathbf{T}_0$ is ergodic, we know $\boldsymbol{\pi}_0$ is strictly positive.

By definition of reversibility, we know $\mathbf{diag}(\boldsymbol{\pi}_0)\mathbf{T}_0 = \mathbf{T}_0^\intercal \mathbf{diag}(\boldsymbol{\pi}_0)$, and this further gives $\mathbf{diag}(\boldsymbol{\pi}_0)^{\frac{1}{2}}\mathbf{T}_0\mathbf{diag}(\boldsymbol{\pi}_0)^{-\frac{1}{2}} = \mathbf{diag}(\boldsymbol{\pi}_0)^{-\frac{1}{2}}\mathbf{T}_0^\intercal\mathbf{diag}(\boldsymbol{\pi}_0)^{\frac{1}{2}}$. Let $\mathbf{H}_0 := \mathbf{diag}(\boldsymbol{\pi}_0)^{\frac{1}{2}}\mathbf{T}_0\mathbf{diag}(\boldsymbol{\pi}_0)^{-\frac{1}{2}}$, then we see $\mathbf{H}_0$ is symmetric. Let $\mathbf{W}_{0,r} \in \mathbb{R}^{s \times r}$ denote the top r left singular vectors of $\mathbf{H}_0$, by spectral theorem, we know the columns of $\mathbf{W}_{0,r}$ also serve as the top r eigenvectors of $\mathbf{H}_0$. Let $\mathbf{S}_{0,r} := \mathbf{diag}(\boldsymbol{\pi}_0)^{-\frac{1}{2}}\mathbf{W}_{0,r}$, by definition of $\mathbf{H}_0$, it is easy to see that the columns of $\mathbf{S}_{0,r}$ are also the top r eigenvectors of $\mathbf{T}_0$. Then, by Lemma 6.5 and the definition of informative spectrum, for any $i, i' \in \Omega_k$, we have $\mathbf{S}_{0,r}(i,:) = \mathbf{S}_{0,r}(i',:)$.

Recall in Algorithm 3, the matrix $\mathbf{W}_r$ denotes the top r left singular vectors of $\mathbf{H} := \mathbf{diag}(\boldsymbol{\pi})^{\frac{1}{2}}\mathbf{T}\mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}}$ and let $\mathbf{S}_r = \mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}}\mathbf{W}_r$. Let $\mathbf{O}^\star := \min_{\mathbf{O}\in\mathcal{O}(r)} \|\mathbf{W}_{0,r}\mathbf{O} - \mathbf{W}_r\|_{\mathrm{F}}$, where $\mathcal{O}(r)$ is the set of all $r \times r$ orthonormal matrices. Then, for any $i, i' \in \Omega_k$, we have $[\mathbf{S}_{0,r}\mathbf{O}^\star](i,:) = [\mathbf{S}_{0,r}\mathbf{O}^\star](i',:)$. Using this, for any $i \in [s]$ (suppose $i \in \Omega_k$), we have

$$\begin{aligned} &\mathbf{S}_r(i,:) - \bar{\mathbf{S}}_r(i,:) \\ =&\frac{|\Omega_k|-1}{|\Omega_k|}\mathbf{S}_r(i,:) - \frac{1}{|\Omega_k|}\sum_{i':i'\in\Omega_k, i'\neq i}\mathbf{S}_r(i',:) \\ \leq&\frac{|\Omega_k|-1}{|\Omega_k|}(\mathbf{S}_r(i,:) - [\mathbf{S}_{0,r}\mathbf{O}^\star](i,:)) \\ &+ \frac{1}{|\Omega_k|}\sum_{i':i'\in\Omega_k, i'\neq i}\big([\mathbf{S}_{0,r}\mathbf{O}^\star](i',:) - \mathbf{S}_r(i',:)\big). \end{aligned} \tag{D.7}$$

WLOG, assume $\{1,\ldots,|\Omega_1|\} = \Omega_1$, $\{|\Omega_1|+1,\ldots,|\Omega_1|+|\Omega_2|\} = \Omega_2, \cdots$ and define block diagonal matrices $\mathbf{D}, \mathbf{P} \in \mathbb{R}^{s \times s}$ both with r diagonal blocks such that their k-th diagonal blocks $[\mathbf{D}]_k, [\mathbf{P}]_k \in \mathbb{R}^{|\Omega_k| \times |\Omega_k|}$ are given by

$$[\mathbf{D}]_k = \frac{|\Omega_k|-1}{|\Omega_k|}\mathbf{I}_{|\Omega_k|},\ [\mathbf{P}]_k = \frac{1}{|\Omega_k|}(\mathbf{1}_{|\Omega_k|}\mathbf{1}_{|\Omega_k|}^\intercal - \mathbf{I}_{|\Omega_k|}). \tag{D.8}$$

Then, stacking (D.7) for all i, one can verify that $\mathbf{S}_r - \bar{\mathbf{S}}_r = \mathbf{D}(\mathbf{S}_r - \mathbf{S}_{0,r}\mathbf{O}^\star) + \mathbf{P}(\mathbf{S}_{0,r}\mathbf{O}^\star - \mathbf{S}_r)$. Note that for an arbitrary matrix $\mathbf{E} \in \mathbb{R}^{s \times s}$, we have $\|\mathbf{PE}\|_{\mathrm{F}}^2 = \mathbf{tr}(\mathbf{P}^\intercal\mathbf{PEE}^\intercal) \leq \mathbf{tr}(\mathbf{D}^\intercal \cdot \mathbf{DEE}^\intercal) = \|\mathbf{DE}\|_{\mathrm{F}}^2$ where the inequality holds since for each diagonal block we have $[\mathbf{P}]_k^\intercal[\mathbf{P}]_k \preceq [\mathbf{D}]_k^\intercal[\mathbf{D}]_k$. Therefore, $\|\mathbf{S}_r - \bar{\mathbf{S}}_r\|_{\mathrm{F}} \leq 2\|\mathbf{D}(\mathbf{S}_r - \mathbf{S}_{0,r}\mathbf{O}^\star)\|_{\mathrm{F}} \leq 2\max_k \frac{|\Omega_k|-1}{|\Omega_k|}\|\mathbf{S}_r - \mathbf{S}_{0,r}\mathbf{O}^\star|_{\mathrm{F}} \leq 2\|\mathbf{S}_r - \mathbf{S}_{0,r}\mathbf{O}^\star\|_{\mathrm{F}}$. To complete the proof, it suffices to study $\|\mathbf{S}_r - \mathbf{S}_{0,r}\mathbf{O}^\star\|_{\mathrm{F}}$.

$$\begin{aligned} &\|\mathbf{S}_r - \mathbf{S}_{0,r}\mathbf{O}^\star\|_{\mathrm{F}} \\ =&\|\mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}}(\mathbf{W}_r - \mathbf{W}_{0,r}\mathbf{O}^\star) \\ &+ (\mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}} - \mathbf{diag}(\boldsymbol{\pi}_0)^{-\frac{1}{2}})\mathbf{W}_{0,r}\mathbf{O}^\star\|_{\mathrm{F}} \\ \leq&\pi_{\min}^{-0.5}\|\mathbf{W}_r - \mathbf{W}_{0,r}\mathbf{O}^\star\|_{\mathrm{F}} + \sqrt{r}\max_i|\boldsymbol{\pi}(i)^{-\frac{1}{2}} - \boldsymbol{\pi}_0(i)^{-\frac{1}{2}}|. \end{aligned} \tag{D.9}$$

According to Lemma D.1, we know $\|\mathbf{W}_r - \mathbf{W}_{0,r}\mathbf{O}^\star\|_\text{F} \le \frac{2\sqrt{2}}{\sigma_r(\mathbf{H})-\sigma_{r+1}(\mathbf{H})}\|\mathbf{H}-\mathbf{H}_0\|_\text{F}$. This together with the upper bound for $\max_i |\boldsymbol{\pi}(i)^{-\frac{1}{2}} - \boldsymbol{\pi}_0(i)^{-\frac{1}{2}}|$ in (D.5) gives

$$\|\mathbf{S}_r - \mathbf{S}_{0,r}\mathbf{O}^\star\|_\text{F} \le \frac{2\sqrt{2}}{(\sigma_r(\mathbf{H}) - \sigma_{r+1}(\mathbf{H}))\pi_{\min}^{0.5}}\|\mathbf{H}-\mathbf{H}_0\|_\text{F} + \frac{(\sqrt{2}-1)\gamma_1\sqrt{r}}{\pi_{\min}^{1.5}}\epsilon_\mathbf{T}. \tag{D.10}$$

By the definitions of $\mathbf{H}$ and $\mathbf{H}_0$, we have $\|\mathbf{H}-\mathbf{H}_0\|_\text{F} \le \|(\mathbf{diag}(\boldsymbol{\pi})^{\frac{1}{2}} - \mathbf{diag}(\boldsymbol{\pi}_0)^{\frac{1}{2}}) \cdot \mathbf{T} \cdot \mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}}\|_\text{F} + \|\mathbf{diag}(\boldsymbol{\pi}_0)^{\frac{1}{2}} \cdot \mathbf{T} \cdot (\mathbf{diag}(\boldsymbol{\pi})^{-\frac{1}{2}} - \mathbf{diag}(\boldsymbol{\pi}_0)^{-\frac{1}{2}})\|_\text{F} + \|\mathbf{diag}(\boldsymbol{\pi}_0)^{\frac{1}{2}} \cdot (\mathbf{T}-\mathbf{T}_0) \cdot \mathbf{diag}(\boldsymbol{\pi}_0)^{-\frac{1}{2}}\|_\text{F}$. Applying Corollary D.6 gives $\|\mathbf{H}-\mathbf{H}_0\|_\text{F} \le 2.56\gamma_1\pi_{\max}^{0.5}\pi_{\min}^{-1.5}\|\mathbf{T}\|_\text{F}\epsilon_\mathbf{T}$. Plugging this into (D.10), we have $\|\mathbf{S}_r - \mathbf{S}_{0,r}\mathbf{O}^\star\|_\text{F} \le \frac{8\gamma_1\sqrt{r}\sqrt{\pi_{\max}}\|\mathbf{T}\|_\text{F}}{\gamma_2\pi_{\min}^2}\epsilon_\mathbf{T}$, where $\gamma_2 := \min\{\sigma_r(\mathbf{H}) - \sigma_{r+1}(\mathbf{H}), 1\}$. This concludes the proof as we showed that $\|\mathbf{S}_r - \bar{\mathbf{S}}_r\|_\text{F} \le 2\|\mathbf{S}_r - \mathbf{S}_{0,r}\mathbf{O}^\star\|_\text{F}$. □

Main Proof for Theorem 6.6. Consider $\mathbf{\Phi}$ in Algorithm 3 Line 8 and its averaged version $\bar{\mathbf{\Phi}}$ defined in Section 6.4.1. Then, by definition, we have $\|\bar{\mathbf{\Phi}} - \mathbf{\Phi}\|_\text{F}^2 = \alpha_\mathbf{T}^2 \cdot \|\mathbf{S}_r - \bar{\mathbf{S}}_r\|_\text{F}^2 + \alpha_\mathbf{A}^2 \cdot \sum_{k\in[r]}\sum_{i\in\Omega_k}\|\mathbf{A}_i - |\Omega_k|^{-1}\cdot\sum_{i'\in\Omega_k}\mathbf{A}_{i'}\|_\text{F}^2 + \alpha_\mathbf{B}^2 \cdot \sum_{k\in[r]}\sum_{i\in\Omega_k}\|\mathbf{B}_i - |\Omega_k|^{-1}\sum_{i'\in\Omega_k}\mathbf{B}_{i'}\|_\text{F}^2$. where $\bar{\mathbf{S}}_r$ is defined in Lemma D.7. By Lemma D.7 and the definitions of $\epsilon_\mathbf{A}$ and $\epsilon_\mathbf{B}$ in Problem P6.1, we have $\|\bar{\mathbf{\Phi}} - \mathbf{\Phi}\|_\text{F} \le \epsilon_{Lmp}$ where $\epsilon_{Lmp} := \sqrt{\alpha_\mathbf{A}^2\epsilon_\mathbf{A}^2 + \alpha_\mathbf{B}^2\epsilon_\mathbf{B}^2 + \alpha_\mathbf{T}^2\gamma_3^2\epsilon_\mathbf{T}^2}$. By construction, in $\bar{\mathbf{\Phi}}$, rows that belong to the same cluster have the same rows, thus we can apply Lemma D.4 to $\{\bar{\mathbf{\Phi}}, \mathbf{\Phi}\}$ and obtain that when $\epsilon_{Lmp} \le \frac{\sigma_r(\bar{\mathbf{\Phi}})\sqrt{|\Omega_{(r)}|+|\Omega_{(1)}|}}{8\sqrt{(2+\epsilon)|\Omega_{(1)}|}}$, we have $\text{MR}(\hat{\Omega}_{1:r}) \le 64(2+\epsilon)\sigma_r(\bar{\mathbf{\Phi}})^{-2}\epsilon_{Lmp}^2$. □

D.2.1 Non-emptiness of $\mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_\mathbf{T})$

Note that both Lemma D.7 and Theorem 6.6 require the set $\mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_\mathbf{T})$, a neighborhood of $\mathbf{T}$. Now, we show it is non-empty under the approximate lumpability condition (6.3).

Let $\mathbf{T}_0 := \mathbf{T} + \mathbf{\Delta}$ for $\mathbf{\Delta} \in \mathcal{D}$ where

$$\begin{aligned}\mathcal{D} := \Big\{\mathbf{\Delta} \in \mathbb{R}^{sxs} : \forall k, l \in [r], \forall i \in \Omega_k,& \\ -\mathbf{T}(i,j) \le \mathbf{\Delta}(i,j) \le 1 - \mathbf{T}(i,j) \quad \forall j \in [s],& \qquad \text{(D.11)} \\ \sum_{j\in\Omega_l}\mathbf{\Delta}(i,j) = -\sum_{j\in\Omega_l}\mathbf{T}(i,j) + |\Omega_k|^{-1}\sum_{\substack{i'\in\Omega_k\\ j\in\Omega_l}}\mathbf{T}(i',j),& \qquad \text{(D.12)} \\ \|\mathbf{\Delta}\|_\text{F} \le \epsilon_\mathbf{T}, \quad \|\mathbf{\Delta}\|_\infty \le \epsilon_\mathbf{T}. \Big\}&\end{aligned}$$

Then, we see to show there exists $\mathbf{T}_0 \in \mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_\mathbf{T})$, i.e. $\mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_\mathbf{T})$ is non-empty, it is equivalent to show there exists $\mathbf{\Delta} \in \mathcal{D}$.

Note that (D.11) gives that for all $i \in \Omega_k, l \in [r]$, $-\sum_{j\in\Omega_l}\mathbf{T}(i,j) \le \sum_{j\in\Omega_l}\mathbf{\Delta}(i,j) \le |\Omega_l| - \sum_{j\in\Omega_l}\mathbf{T}(i,j)$. This together with (D.11) and (D.12) imply that there exists $\mathbf{\Delta}$ satisfying both (D.11) and (D.12) such that among its elements $\{\mathbf{\Delta}(i,j)\}_{j\in\Omega_l}$, the nonzero ones have the same signs as the RHS of (D.12). Then, for all $i \in \Omega_k, l \in [r]$, we have $\sum_{j\in\Omega_l}|\mathbf{\Delta}(i,j)| = |\sum_{j\in\Omega_l}\mathbf{\Delta}(i,j)| = |\text{RHS of (D.12)}| \leq |\Omega_k|^{-1}\sum_{i'\in\Omega_k}|\sum_{j\in\Omega_l}\mathbf{T}(i,j) - \sum_{j\in\Omega_l}\mathbf{T}(i',j)|$. This further gives $\|\mathbf{\Delta}\|_\text{F} \leq \sum_{k,l\in[r]}\sum_{i\in\Omega_k,j\in\Omega_l}|\mathbf{\Delta}(i,j)| \leq |\Omega_k|^{-1}\sum_{k,l\in[r]}\sum_{i,i'\in\Omega_k}|\sum_{j\in\Omega_l}\mathbf{T}(i,j) - \sum_{j\in\Omega_l}\mathbf{T}(i',j)| \leq |\Omega_k|^{-1}\epsilon_\mathbf{T}$, where the last inequality follows from (6.3). These steps also show $\|\mathbf{\Delta}\|_\infty \leq \epsilon_\mathbf{T}$. We have shown $\mathbf{\Delta} \in \mathcal{D}$, i.e. $\mathcal{D}$ is non-empty, and so is $\mathcal{L}(\mathbf{T}, \Omega_{1:r}, \epsilon_\mathbf{T})$.

D.3 Approximation with Mean-Square Stability — Proof for Theorem 6.7

We first provide several supporting results regarding the perturbation of matrix product.

Lemma D.8. *Consider two sets of matrices $\mathbf{A}_1, \dots, \mathbf{A}_s$ and $\hat{\mathbf{A}}_1, \dots, \hat{\mathbf{A}}_s$ with $\|\mathbf{A}_i - \hat{\mathbf{A}}_i\| \leq \epsilon_\mathbf{A}$ for all $i \in [s]$. Assume that there exists a pair $\{\xi, \kappa\}$ such that for all $t \in \mathbb{N}$, we have $\max_{\sigma_{1:t}\in[s]^t}\|\mathbf{A}_{\sigma_1}\cdots\mathbf{A}_{\sigma_t}\|^{\frac{1}{t}} \leq \kappa\cdot\xi^t$. Then, for all t and any sequence $\sigma_{1:t} \in [s]^t$, we have (i) $\|\prod_{h=1}^t \hat{\mathbf{A}}_{\sigma_h}\| \leq \kappa(\kappa\epsilon_\mathbf{A} + \xi)^t$; (ii) $\|\prod_{h=1}^t \hat{\mathbf{A}}_{\sigma_h} - \prod_{h=1}^t \mathbf{A}_{\sigma_h}\| \leq \kappa^2 t(\kappa\epsilon_\mathbf{A} + \xi)^{t-1}\epsilon_\mathbf{A}$.*

Proof. Let $\mathbf{E}_i := \hat{\mathbf{A}}_i - \mathbf{A}_i$, then we see $\|\mathbf{E}_i\| \leq \epsilon_\mathbf{A}$ and $\prod_{h=1}^t \hat{\mathbf{A}}_{\sigma_h} = \prod_{h=1}^t(\mathbf{A}_{\sigma_h} + \mathbf{E}_{\sigma_h})$. In the expansion of $\prod_{h=1}^t(\mathbf{A}_{\sigma_h} + \mathbf{E}_{\sigma_h})$, for each $i = 0, 1, \dots, t$, there are $\binom{t}{i}$ terms, each of which is a product where $\mathbf{E}_{\sigma_h}$ has degree i and $\mathbf{A}_{\sigma_h}$ has degree $t - i$. We let $\mathbf{F}_{i,j}$ with $i = 0, 1, \dots, t$ and $j \in [\binom{t}{i}]$ to index these expansion terms. Note that $\|\mathbf{F}_{i,j}\| \leq \kappa^{i+1}\xi^{t-i}\epsilon_\mathbf{A}^i$. Then, we have $\|\prod_{h=1}^t \hat{\mathbf{A}}_{\sigma_h}\| \leq \sum_{i=0}^t \sum_{j\in[\binom{t}{i}]}\|\mathbf{F}_{i,j}\| \leq \sum_{i=0}^t \binom{t}{i}\kappa^{i+1}\xi^{t-i}\epsilon_\mathbf{A}^i \leq \kappa(\kappa\epsilon_\mathbf{A} + \xi)^t$.

Similarly, $\|\prod_{h=1}^t \hat{\mathbf{A}}_{\sigma_h} - \prod_{h=1}^t \mathbf{A}_{\sigma_h}\| = \|\sum_{i=1}^t\sum_{j\in[\binom{t}{i}]}\mathbf{F}_{i,j}\| \leq \sum_{i=0}^t\sum_{j\in[\binom{t}{i}]}\|\mathbf{F}_{i,j}\| - \|\mathbf{F}_{0,1}\| \leq \kappa(\kappa\epsilon_\mathbf{A} + \xi)^t - \kappa\xi^t \leq \kappa^2 t(\kappa\epsilon_\mathbf{A} + \xi)^{t-1}\epsilon_\mathbf{A}$, where the last line follows from the fact that for function $f(x) := x^t$ and $x, a \geq 0$, $f(x) \geq f(x+a) - a\cdot f'(x+a)$. □

Based on Lemma D.8, we have the following corollaries, which will be used in different settings in later derivations.

Corollary D.9. *Consider two matrices $\boldsymbol{\mathcal{A}}$ and $\bar{\boldsymbol{\mathcal{A}}}$ with $\|\boldsymbol{\mathcal{A}} - \bar{\boldsymbol{\mathcal{A}}}\| \leq \epsilon_{\boldsymbol{\mathcal{A}}}$. Suppose there exists a pair $\{\rho, \tau\}$ such that for all $k \in \mathbb{N}$, $\|\boldsymbol{\mathcal{A}}^k\| \leq \tau\rho^k$. Then, we have $\|\bar{\boldsymbol{\mathcal{A}}}^t\| \leq \tau\left(\tau\epsilon_{\boldsymbol{\mathcal{A}}} + \rho\right)^t$ and $\|\bar{\boldsymbol{\mathcal{A}}}^t - \boldsymbol{\mathcal{A}}^t\| \leq \tau^2 t\left(\tau\epsilon_{\boldsymbol{\mathcal{A}}} + \rho\right)^{t-1}\epsilon_{\boldsymbol{\mathcal{A}}}$.*

Corollary D.10. *Consider two sets of scalars $a_1, \dots, a_s$ and $\hat{a}_1, \dots, \hat{a}_s$ with $|a_i - \hat{a}_i| < \epsilon_a$ and $|a_i| < \bar{a}$ for all $i \in [s]$. Then, for all t and any sequence $\sigma_{1:t} \in [s]^t$, we have $|\prod_{h=1}^t \hat{a}_{\sigma_h}| \leq (\epsilon_a + \bar{a})^t$ and $|\prod_{h=1}^t \hat{a}_{\sigma_h} - \prod_{h=1}^t a_{\sigma_h}| \leq t(\epsilon_a + \bar{a})^{t-1}\epsilon_a$.*

The next result considers the evolution of state $\mathbf{x}_t$ in the mean-square sense for autonomous MJSs, which can be obtained from Lemma 2.14.

Lemma D.11. *Consider* $\text{MJS}(\mathbf{A}_{1:s}, 0, \mathbf{T})$ *and define matrix* $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{sn^2 \times sn^2}$ *with its* (i,j)*-th* $n^2 \times n^2$ *block given by* $[\boldsymbol{\mathcal{A}}]_{ij} := \mathbf{T}(j,i) \cdot \mathbf{A}_j \otimes \mathbf{A}_j$. *Let* $\boldsymbol{\Sigma}_{t,i} := \mathbb{E}[\mathbf{x}_t \mathbf{x}_t^\intercal \mathbf{1}_{\{\omega_t = i\}}]$ *and* $\mathbf{s}_t := [\mathbf{vec}(\boldsymbol{\Sigma}_{t,1}^\intercal, \dots, \mathbf{vec}(\boldsymbol{\Sigma}_{t,s})^\intercal]^\intercal$, *then* $\mathbf{s}_t = \boldsymbol{\mathcal{A}}^t \mathbf{s}_0$.

Recall in Section 6.5, for Σ, i.e., $\text{MJS}(\mathbf{A}_{1:s}, \mathbf{B}_{1:s}, \mathbf{T})$, we define the augmented state matrix $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{sn^2 \times sn^2}$ with its (i,j)-th $n^2 \times n^2$ block given by $[\boldsymbol{\mathcal{A}}]_{ij} := \mathbf{T}(j,i) \cdot \mathbf{A}_j \otimes \mathbf{A}_j$; and for any $\rho \geq \rho(\boldsymbol{\mathcal{A}})$ and all $k \in \mathbb{N}$, we have $\|\boldsymbol{\mathcal{A}}^k\| \leq \tau \rho^k$. The next lemma is regarding the augmentation of two MJS with the same $\mathbf{A}$ matrix.

Lemma D.12. *Construct matrix* $\check{\boldsymbol{\mathcal{A}}} \in \mathbb{R}^{4sn^2 \times 4sn^2}$ *with its* (i,j)*-th* $4n^2 \times 4n^2$ *block given by* $[\check{\boldsymbol{\mathcal{A}}}]_{ij} := \mathbf{T}(j,i) \cdot \begin{bmatrix} \mathbf{A}_j & \\ & \mathbf{A}_j \end{bmatrix} \otimes \begin{bmatrix} \mathbf{A}_j & \\ & \mathbf{A}_j \end{bmatrix}$. *Then, for all* $k \in \mathbb{N}$, $\|\check{\boldsymbol{\mathcal{A}}}^k\| \leq \tau \rho^k$.

To see this result, first notice that there exists a permutation matrix $\mathbf{P}$ such that $\mathbf{P}\check{\boldsymbol{\mathcal{A}}}\mathbf{P}^\intercal = \mathbf{I}_4 \otimes \boldsymbol{\mathcal{A}}$, where $\mathbf{I}_4$ denotes the 4×4 identity matrix. This gives $\|\check{\boldsymbol{\mathcal{A}}}^k\| = \|\boldsymbol{\mathcal{A}}^k\|$ and shows the claim.

To prove the result in Theorem 6.7, we first consider the simplified autonomous case but with potentially different initial states $\mathbf{x}_0$ and $\hat{\mathbf{x}}_0$.

Proposition D.13. *Consider the setup in Theorem 6.7 except that* $\mathbf{u}_t = 0$ *for all* t, *and* $\mathbf{x}_0$ *and* $\hat{\mathbf{x}}_0$ *can be different such that* $\|\mathbf{x}_0 - \hat{\mathbf{x}}_0\| \leq \epsilon_0$ *for some* $\epsilon_0 \geq 0$. *For perturbation, assume* $\epsilon_{\mathbf{A}} \leq \min\{\bar{A}, \frac{1-\rho}{6\tau \bar{A} \|\mathbf{T}\|}\}$. *Then,* $\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|] \leq 4\sqrt{n\sqrt{s}}\tau \rho_0^{\frac{t-1}{2}} \left(\|\mathbf{x}_0\| \sqrt{t\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}} + \sqrt{(\|\mathbf{x}_0\| + \epsilon_0)\epsilon_0} \right)$.

Proof. First, we construct two autonomous switched systems:

$$\check{\Pi} := \begin{cases} \check{\mathbf{x}}_{t+1} = \check{\mathbf{A}}_{\check{\omega}_t} \check{\mathbf{x}}_t \\ \check{\omega}_t = \omega_t, \end{cases}, \quad \bar{\Pi} := \begin{cases} \bar{\mathbf{x}}_{t+1} = \bar{\mathbf{A}}_{\bar{\omega}_t} \bar{\mathbf{x}}_t \\ \bar{\omega}_t = \omega_t, \end{cases} \tag{D.13}$$

where for $i \in [s]$ (suppose $i \in \Omega_k$), $\check{\mathbf{A}}_i := \begin{bmatrix} \mathbf{A}_i & \\ & \mathbf{A}_i \end{bmatrix}$, $\bar{\mathbf{A}}_i := \begin{bmatrix} \mathbf{A}_i & \\ & \hat{\mathbf{A}}_k \end{bmatrix}$. Since ω_t of Σ follows Markov chain $\mathbf{T}$, systems $\check{\Pi}$ and $\bar{\Pi}$ can be viewed as $\text{MJS}(\check{\mathbf{A}}_{1:s}, 0, \check{\mathbf{T}})$ and $\text{MJS}(\bar{\mathbf{A}}_{1:s}, 0, \bar{\mathbf{T}})$ respectively with $\check{\mathbf{T}} = \bar{\mathbf{T}} = \mathbf{T}$. We then define observations for $\check{\Pi}$ and $\bar{\Pi}$: $\check{\mathbf{y}}_t = \check{\mathbf{C}}\check{\mathbf{x}}_t$ and $\bar{\mathbf{y}}_t = \bar{\mathbf{C}}\bar{\mathbf{x}}_t$ where $\check{\mathbf{C}} = \bar{\mathbf{C}} = [\mathbf{I}_n, -\mathbf{I}_n]$. We set their initial states as $\check{\mathbf{x}}_0 = [\mathbf{x}_0^\intercal, \mathbf{x}_0^\intercal]^\intercal$, $\bar{\mathbf{x}}_0 = [\mathbf{x}_0^\intercal, \hat{\mathbf{x}}_0^\intercal]^\intercal$ where $\mathbf{x}_0$ and $\hat{\mathbf{x}}_0$ are the initial states of Σ and $\hat{\Sigma}$ respectively.

By construction, we have, for all t, $\check{\mathbf{x}}_t = [\mathbf{x}_t^\intercal, \mathbf{x}_t^\intercal]^\intercal$ and $\bar{\mathbf{x}}_t = [\mathbf{x}_t^\intercal, \hat{\mathbf{x}}_t^\intercal]^\intercal$, thus $\check{\mathbf{y}}_t = 0$ and $\bar{\mathbf{y}}_t = \mathbf{x}_t - \hat{\mathbf{x}}_t$. Define $\check{\boldsymbol{\Sigma}}_t := \mathbb{E}[\check{\mathbf{x}}_t \check{\mathbf{x}}_t^\intercal]$ and $\bar{\boldsymbol{\Sigma}}_t := \mathbb{E}[\bar{\mathbf{x}}_t \bar{\mathbf{x}}_t^\intercal]$, then we have $\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|^2] = \mathbb{E}[\bar{\mathbf{y}}_t \bar{\mathbf{y}}_t^\intercal] = \mathbb{E}[\bar{\mathbf{y}}_t \bar{\mathbf{y}}_t^\intercal] - \mathbb{E}[\check{\mathbf{y}}_t \check{\mathbf{y}}_t^\intercal] = \mathbf{tr}(\bar{\mathbf{C}}^\intercal \bar{\mathbf{C}} \bar{\boldsymbol{\Sigma}}_t) - \mathbf{tr}(\check{\mathbf{C}}^\intercal \check{\mathbf{C}} \check{\boldsymbol{\Sigma}}_t) = \mathbf{tr}(\bar{\mathbf{C}}^\intercal \bar{\mathbf{C}}(\bar{\boldsymbol{\Sigma}}_t - \check{\boldsymbol{\Sigma}}_t))$. Since $\bar{\mathbf{C}}^\intercal \bar{\mathbf{C}} \succeq 0$, we further have

$$\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|^2] \leq \mathbf{tr}(\bar{\mathbf{C}}^\intercal \bar{\mathbf{C}}) \|\bar{\boldsymbol{\Sigma}}_t - \check{\boldsymbol{\Sigma}}_t\| = 2n \|\bar{\boldsymbol{\Sigma}}_t - \check{\boldsymbol{\Sigma}}_t\|. \tag{D.14}$$

Let $\check{\boldsymbol{\Sigma}}_{t,i} := \mathbb{E}[\check{\mathbf{x}}_t\check{\mathbf{x}}_t^\intercal \mathbf{1}_{\{\check{\omega}_t=i\}}]$, $\bar{\boldsymbol{\Sigma}}_{t,i} := \mathbb{E}[\bar{\mathbf{x}}_t\bar{\mathbf{x}}_t^\intercal \mathbf{1}_{\{\bar{\omega}_t=i\}}]$, $\check{\mathbf{s}}_t := [\mathbf{vec}(\check{\boldsymbol{\Sigma}}_{t,1})^\intercal, \dots, \mathbf{vec}(\check{\boldsymbol{\Sigma}}_{t,s})^\intercal]^\intercal$ and $\bar{\mathbf{s}}_t := [\mathbf{vec}(\bar{\boldsymbol{\Sigma}}_{t,1})^\intercal, \dots , \mathbf{vec}(\bar{\boldsymbol{\Sigma}}_{t,s})^\intercal]^\intercal$. Note that $\mathbf{vec}(\check{\boldsymbol{\Sigma}}_t) = [\mathbf{I}_{4n^2}, \dots, \mathbf{I}_{4n^2}]\check{\mathbf{s}}_t$ and $\mathbf{vec}(\bar{\boldsymbol{\Sigma}}_t) = [\mathbf{I}_{4n^2}, \dots, \mathbf{I}_{4n^2}]\bar{\mathbf{s}}_t$, thus we have $\|\check{\boldsymbol{\Sigma}}_t - \bar{\boldsymbol{\Sigma}}_t\| \le \|\check{\boldsymbol{\Sigma}}_t - \bar{\boldsymbol{\Sigma}}_t\|_{\mathrm{F}} = \|\mathbf{vec}(\check{\boldsymbol{\Sigma}}_t - \bar{\boldsymbol{\Sigma}}_t)\| \le \sqrt{s}\|\check{\mathbf{s}}_t - \bar{\mathbf{s}}_t\|$. Plugging this into (D.14), we have

$$\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|^2] \le 2n\sqrt{s}\|\check{\mathbf{s}}_t - \bar{\mathbf{s}}_t\|. \tag{D.15}$$

By Lemma D.11, we have $\check{\mathbf{s}}_t = \check{\boldsymbol{\mathcal{A}}}^t\check{\mathbf{s}}_0$ and $\bar{\mathbf{s}}_t = \bar{\boldsymbol{\mathcal{A}}}^t\bar{\mathbf{s}}_0$, where $\check{\boldsymbol{\mathcal{A}}} \in \mathbb{R}^{4sn^2 \times 4sn^2}$ is constructed such that its (i,j)-th $4n^2 \times 4n^2$ block given by $[\check{\boldsymbol{\mathcal{A}}}]_{ij} = \check{\mathbf{T}}(j,i)\check{\mathbf{A}}_j \otimes \check{\mathbf{A}}_j$, and $\bar{\boldsymbol{\mathcal{A}}}$ is constructed similarly. By triangle inequality, we further have

$$\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|^2] \le 2n\sqrt{s}(\|\check{\boldsymbol{\mathcal{A}}}^t - \bar{\boldsymbol{\mathcal{A}}}^t\|\|\check{\mathbf{s}}_0\| + \|\bar{\boldsymbol{\mathcal{A}}}^t\|\|\check{\mathbf{s}}_0 - \bar{\mathbf{s}}_0\|) \tag{D.16}$$

To bound $\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|^2]$, we seek to bound the terms on the RHS individually. Since $\check{\mathbf{s}}_0 = [\mathbf{vec}(\check{\mathbf{x}}_0\check{\mathbf{x}}_0^\intercal)^\intercal \cdot \mathbb{P}(\omega_t = 1), \dots, \mathbf{vec}(\check{\mathbf{x}}_0\check{\mathbf{x}}_0^\intercal)^\intercal \cdot \mathbb{P}(\omega_t = s)]^\intercal$, we have $\|\check{\mathbf{s}}_0\| = \|\check{\mathbf{x}}_0\check{\mathbf{x}}_0^\intercal\|_{\mathrm{F}} \cdot (\sum_{i\in[s]} \mathbb{P}(\omega_t = i)^2)^{\frac{1}{2}} \le \|\check{\mathbf{x}}_0\check{\mathbf{x}}_0^\intercal\|_{\mathrm{F}} = 2\|\mathbf{x}_0\|^2$. Similarly, we have $\|\check{\mathbf{s}}_0 - \bar{\mathbf{s}}_0\| \le \|\bar{\mathbf{x}}_0\bar{\mathbf{x}}_0^\intercal - \check{\mathbf{x}}_0\check{\mathbf{x}}_0^\intercal\|_{\mathrm{F}} \le \|\bar{\mathbf{x}}_0(\bar{\mathbf{x}}_0 - \check{\mathbf{x}}_0)^\intercal\|_{\mathrm{F}} + \|(\bar{\mathbf{x}}_0 - \check{\mathbf{x}}_0)\check{\mathbf{x}}_0^\intercal\|_{\mathrm{F}} \le \sqrt{2}(\sqrt{3}\|\mathbf{x}_0\| + \epsilon_0)\epsilon_0$.

To bound $\|\bar{\boldsymbol{\mathcal{A}}}^t\|$ and $\|\check{\boldsymbol{\mathcal{A}}}^t - \bar{\boldsymbol{\mathcal{A}}}^t\|$, we first evaluate $\|\check{\boldsymbol{\mathcal{A}}} - \bar{\boldsymbol{\mathcal{A}}}\|$. Define $\boldsymbol{\Delta}_i := \check{\mathbf{A}}_i \otimes \check{\mathbf{A}}_i - \bar{\mathbf{A}}_i \otimes \bar{\mathbf{A}}_i$ for all i, and block diagonal matrix $\boldsymbol{\Delta} \in \mathbb{R}^{sn^2 \times sn^2}$ such that the i-th $n^2 \times n^2$ block is given by $\boldsymbol{\Delta}_i$. Then one can verify that $\check{\boldsymbol{\mathcal{A}}} - \bar{\boldsymbol{\mathcal{A}}} = (\mathbf{T} \otimes \mathbf{I}_{n^2})\boldsymbol{\Delta}$, which gives $\|\check{\boldsymbol{\mathcal{A}}} - \bar{\boldsymbol{\mathcal{A}}}\| \le \|\mathbf{T}\| \max_i \|\boldsymbol{\Delta}_i\|$. For $\|\boldsymbol{\Delta}_i\|$, we have $\|\boldsymbol{\Delta}_i\| \le \|\check{\mathbf{A}}_i\|\|\check{\mathbf{A}}_i - \bar{\mathbf{A}}_i\| + \|\check{\mathbf{A}}_i - \bar{\mathbf{A}}_i\|\|\bar{\mathbf{A}}_i\|$. It is easy to see $\|\check{\mathbf{A}}_i\| \le \bar{A}$, $\|\check{\mathbf{A}}_i - \bar{\mathbf{A}}_i\| \le \epsilon_{\mathbf{A}}$, and $\|\bar{\mathbf{A}}_i\| \le \bar{A} + \epsilon_{\mathbf{A}} \le 2\bar{A}$. These give $\|\check{\boldsymbol{\mathcal{A}}} - \bar{\boldsymbol{\mathcal{A}}}\| \le 3\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}$. From Lemma D.12, we know for all $k \in \mathbb{N}$, $\|\check{\boldsymbol{\mathcal{A}}}^k\| \le \tau\rho^k$. Then, according to Corollary D.9, we have $\|\bar{\boldsymbol{\mathcal{A}}}^t\| \le \tau(3\tau\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}} + \rho)^t \le \tau\rho_0^t$ and $\|\check{\boldsymbol{\mathcal{A}}}^t - \bar{\boldsymbol{\mathcal{A}}}^t\| \le 3t(3\tau\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}} + \rho)^{t-1}\tau^2\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}} \le 3t\rho_0^{t-1}\tau^2\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}$, where the premise $\epsilon_{\mathbf{A}} \le \frac{1-\rho}{6\tau\bar{A}\|\mathbf{T}\|}$ and notation $\rho_0 := \frac{1+\rho}{2}$ are used.

Finally, plugging in the bounds we just derived for each term on the RHS of (D.16) back, we have $\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|] \le \sqrt{\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|^2]} \le 4\sqrt{n\sqrt{s}}\tau\rho_0^{\frac{t-1}{2}}\big(\|\mathbf{x}_0\|\sqrt{t\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}} + \sqrt{(\|\mathbf{x}_0\| + \epsilon_0)\epsilon_0}\big)$, which concludes the proof. □

Main Proof for Theorem 6.7. We first decompose $\mathbf{x}_t$ in terms of the contribution from $\mathbf{x}_0$, $\mathbf{u}_{0:t-1}$: define $\mathbf{x}_t^{(0')} := \big(\prod_{h=0}^{t-1} \mathbf{A}_{\omega_h}\big)\mathbf{x}_0$, for $l = 0, \dots, t-2$, define $\mathbf{x}_t^{(l)} := \big(\prod_{h=l+1}^{t-1} \mathbf{A}_{\hat{\omega}_h}\big)\mathbf{B}_{\omega_l}\mathbf{u}_l$, and $\mathbf{x}_t^{(t-1)} := \mathbf{B}_{\omega_{t-1}}\mathbf{u}_{t-1}$. Then it is easy to see $\mathbf{x}_t = \mathbf{x}_t^{(0')} + \sum_{l=0}^{t-1}\mathbf{x}_t^{(l)}$. Similarly, we define $\hat{\mathbf{x}}_t^{(0')}$ and $\hat{\mathbf{x}}_t^{(l)}$ for $\hat{\mathbf{x}}_t$ such that $\hat{\mathbf{x}}_t = \hat{\mathbf{x}}_t^{(0')} + \sum_{l=0}^{t-1}\hat{\mathbf{x}}_t^{(l)}$. According to Proposition D.13, we have

$$\mathbb{E}[\|\mathbf{x}_t^{(0')} - \hat{\mathbf{x}}_t^{(0')}\|] \le 4\sqrt{n\sqrt{s}}\tau\rho_0^{\frac{t-1}{2}}\sqrt{t\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}}\|\mathbf{x}_0\|. \tag{D.17}$$

Note that $\mathbf{x}_t^{(l)}$ and $\hat{\mathbf{x}}_t^{(l)}$ can be viewed as the states at time $t-l$ with respective initial states

$\mathbf{B}_{\omega_l}\mathbf{u}_l$ and $\hat{\mathbf{B}}_{\omega_l}\mathbf{u}_l$ and zero inputs. Therefore, applying Proposition D.13 again, we have

$$\mathbb{E}[\|\mathbf{x}_t^{(l)} - \hat{\mathbf{x}}_t^{(l)}\|] \leq 4\sqrt{n\sqrt{s}}\tau\rho_0^{\frac{t-l-1}{2}}\sqrt{\bar{B}}\bar{u}\big(\sqrt{(t-l-1)\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}} + \sqrt{2\epsilon_{\mathbf{B}}}\big), \tag{D.18}$$

where the premise $\epsilon_{\mathbf{B}} \leq \bar{B}$ is applied. Finally, using (D.17) and (D.18), we obtain that $\mathbb{E}[\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|] \leq \mathbb{E}[\|\mathbf{x}_t^{(0')} - \hat{\mathbf{x}}_t^{(0')}\|] + \sum_{l=0}^{t-1}\mathbb{E}[\|\mathbf{x}_t^{(l)} - \hat{\mathbf{x}}_t^{(l)}\|] \leq 4\sqrt{n\sqrt{s}}\tau\rho_0^{\frac{t-1}{2}}\sqrt{t\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}}\|\mathbf{x}_0\| + 4\sqrt{n\sqrt{s}\bar{B}}\tau\bar{u}\cdot\,(\frac{\sqrt{\rho_0}}{(1-\sqrt{\rho_0})^2}\sqrt{\bar{A}\|\mathbf{T}\|\epsilon_{\mathbf{A}}} + \frac{\sqrt{2}}{1-\sqrt{\rho_0}}\sqrt{\epsilon_{\mathbf{B}}})$, which concludes the proof. □

D.4 Approximation with Unif. Stability — Proof for Theorem 6.8

Proof for Theorem 6.8 (T1). $\mathbf{x}_t$ and $\hat{\mathbf{x}}_t$ can be decomposed as:

$$\mathbf{x}_t = (\prod_{h=0}^{t-1}\mathbf{A}_{\omega_h})\mathbf{x}_0 + \sum_{t'=0}^{t-2}(\prod_{h=t'+1}^{t-1}\mathbf{A}_{\omega_h})\mathbf{B}_{\omega_{t'}}\mathbf{u}_{t'} + \mathbf{B}_{\omega_{t-1}}\mathbf{u}_{t-1},$$

$$\hat{\mathbf{x}}_t = (\prod_{h=0}^{t-1}\hat{\mathbf{A}}_{\hat{\omega}_h})\hat{\mathbf{x}}_0 + \sum_{t'=0}^{t-2}(\prod_{h=t'+1}^{t-1}\hat{\mathbf{A}}_{\hat{\omega}_h})\hat{\mathbf{B}}_{\hat{\omega}_{t'}}\hat{\mathbf{u}}_{t'} + \hat{\mathbf{B}}_{\hat{\omega}_{t-1}}\hat{\mathbf{u}}_{t-1}.$$

Since in Algorithm 3 we let $\hat{\mathbf{A}}_k = |\hat{\Omega}_k|^{-1}\sum_{i\in\hat{\Omega}_k}\mathbf{A}_i$, and the premise gives $\hat{\Omega}_{1:r} = \Omega_{1:r}$, we have $\|\hat{\mathbf{A}}_k - \mathbf{A}_i\| \leq \epsilon_{\mathbf{A}}$ for all $i \in \Omega_k$. Based on the mode synchrony Setup S2, i.e., $\omega_t \in \Omega_{\hat{\omega}_t}$, we further have $\|\hat{\mathbf{A}}_{\hat{\omega}_t} - \mathbf{A}_{\omega_t}\| \leq \epsilon_{\mathbf{A}}$. Similarly, we obtain $\|\hat{\mathbf{B}}_{\hat{\omega}_t} - \mathbf{B}_{\omega_t}\| \leq \epsilon_{\mathbf{B}}$. Then, by Lemma D.8 : (i) $\|\prod_{h=t'+1}^{t-1}\hat{\mathbf{A}}_{\hat{\omega}_h}\| \leq \kappa(\kappa\epsilon_{\mathbf{A}}+\xi)^{t-t'-1}$ and (ii) $\|\prod_{h=t'+1}^{t-1}\mathbf{A}_{\omega_h} - \prod_{h=t'+1}^{t-1}\hat{\mathbf{A}}_{\hat{\omega}_h}\| \leq \kappa^2(t-t'-1)(\kappa\epsilon_{\mathbf{A}}+\xi)^{t-t'-2}\epsilon_{\mathbf{A}}$.

With (i) and (ii), and the fact that $\prod_{h=t'+1}^{t-1}\mathbf{A}_{\omega_h}\mathbf{B}_{\omega_{t'}} - \prod_{h=t'+1}^{t-1}\hat{\mathbf{A}}_{\hat{\omega}_h}\hat{\mathbf{B}}_{\hat{\omega}_{t'}} = \big(\prod_{h=t'+1}^{t-1}\mathbf{A}_{\omega_h} - \prod_{h=t'+1}^{t-1}\hat{\mathbf{A}}_{\hat{\omega}_h}\big)\hat{\mathbf{B}}_{\hat{\omega}_{t'}} - \big(\prod_{h=t'+1}^{t-1}\mathbf{A}_{\omega_h}\big)\,(\hat{\mathbf{B}}_{\hat{\omega}_{t'}} - \mathbf{B}_{\omega_{t'}})$, we have

$$\begin{aligned}\|\textstyle\prod_{h=t'+1}^{t-1}\mathbf{A}_{\omega_h}\mathbf{B}_{\omega_{t'}} - \prod_{h=t'+1}^{t-1}\hat{\mathbf{A}}_{\hat{\omega}_h}\hat{\mathbf{B}}_{\hat{\omega}_{t'}}\| \leq{}& \kappa\xi^{t-t'-1}\epsilon_{\mathbf{B}} \\ &+ \kappa^2(t-t'-1)(\kappa\epsilon_{\mathbf{A}}+\xi)^{t-t'-2}(\bar{B}+\epsilon_{\mathbf{B}})\epsilon_{\mathbf{A}}.\end{aligned} \tag{D.19}$$

According to Setup S1, Σ and $\hat{\Sigma}$ have the same initial states and inputs. Then, applying triangle inequality to the difference $\|\mathbf{x}_t - \hat{\mathbf{x}}_t\|$, we have

$$\|\mathbf{x}_t - \hat{\mathbf{x}}_t\| \leq \kappa^2 t(\kappa\epsilon_{\mathbf{A}}+\xi)^{t-1}\|\mathbf{x}_0\|\epsilon_{\mathbf{A}} + \kappa^2\frac{1+t(\kappa\epsilon_{\mathbf{A}}+\xi)^t}{1-\kappa\epsilon_{\mathbf{A}}-\xi}(\bar{B}+\epsilon_{\mathbf{B}})\bar{u}\epsilon_{\mathbf{A}} + \frac{\kappa}{1-\xi}\bar{u}\epsilon_{\mathbf{B}}, \tag{D.20}$$

where the following facts are implicitly used: (i) $\kappa \geq 1$ by definition; (ii) $\kappa\epsilon_{\mathbf{A}} + \xi < 1$

according to the premise. Finally, note that we assume perturbation $\epsilon_{\mathbf{A}} \leq \frac{1-\xi}{2\kappa}$ and $\epsilon_{\mathbf{B}} \leq \bar{B}$, we have $\|\mathbf{x}_t - \hat{\mathbf{x}}_t\| \leq t\xi_0^{t-1}\kappa^2\|\mathbf{x}_0\|\epsilon_{\mathbf{A}} + \frac{2(1+t\xi_0^t)\kappa^2\bar{B}\bar{u}}{1-\xi_0}\epsilon_{\mathbf{A}} + \frac{\kappa\bar{u}}{1-\xi}\epsilon_{\mathbf{B}}$, which concludes the proof. □

D.4.1 Proof for Theorem 6.8 (T2)

To ease the proof exposition, we first define a few notations and concepts. For the original system Σ, fixing the initial state $\mathbf{x}_0$ and input sequence $\mathbf{u}_{0:t-1}$, there can be at most s^t possible $\mathbf{x}_t$, each of which correspond to one possible mode switching sequence $\omega_{0:t-1} \in [s]^t$. We use $g \in [s^t]$ to index these states and mode sequences, i.e., mode sequence $\omega_{0:t-1}^{(g)}$ generates state $\mathbf{x}_t^{(g)}$. Then, the reachable set $\mathcal{X}_t$ defined in Section 6.5 satisfies $\mathcal{X}_t = \bigcup_{g\in[s^t]}\{\mathbf{x}_t^{(g)}\}$. Define probability measure $q_t(g) := \mathbb{P}(\omega_{0:t-1} = \omega_{0:t-1}^{(g)})$, then we see $p_t(\mathbf{x})$ defined in Section 6.5 satisfies $p_t(\mathbf{x}) = \sum_{g:\mathbf{x}_t^{(g)}=\mathbf{x}} q_t(g)$. For the reduced $\hat{\Sigma}$ and for all $\hat{g} \in [r^t]$, we similarly define notations $\hat{\omega}_{0:t-1}^{(\hat{g})}$ for the mode sequence, $\hat{\mathbf{x}}_t^{(\hat{g})}$ for the state, and $\hat{q}_t(\hat{g}) := \mathbb{P}(\hat{\omega}_{0:t-1} = \hat{\omega}_{0:t-1}^{(\hat{g})})$ for the measure. Then, the following holds: $\hat{\mathcal{X}}_t = \bigcup_{\hat{g}\in[r^t]}\{\hat{\mathbf{x}}_t^{(\hat{g})}\}$ and $\hat{p}_t(\hat{\mathbf{x}}) = \sum_{\hat{g}:\hat{\mathbf{x}}_t^{(\hat{g})}=\hat{\mathbf{x}}} \hat{q}(\hat{g})$. Next, we introduce the following relation regarding mode sequences between Σ and $\hat{\Sigma}$.

Definition D.14 (Mode Sequence Synchrony)**.** *For any $g \in [s^t], \hat{g} \in [r^t]$, we say $\omega_{0:t-1}^{(g)}$ is synchronous to $\hat{\omega}_{0:t-1}^{(\hat{g})}$ (denoted by $g \triangleright \hat{g}$) if $\omega_h \in \Omega_h$ for all $h = 0, 1, \ldots, t-1$.*

Note that the synchrony definition here coincides with the mode synchrony in Setup S2. With this synchrony relation, we first present a preliminary result.

Lemma D.15. *For any $\hat{g} \in [r]^t$, we have $\left|\hat{q}_t(\hat{g}) - \sum_{g:g\triangleright\hat{g}} q_t(g)\right| \leq (t-1)(\bar{\mathcal{T}} + \epsilon_{\mathbf{T}})^{t-2}\epsilon_{\mathbf{T}}$.*

Proof. Recall ζ_t indexes the active cluster of Σ at time t, i.e., $\zeta_t = k$ if and only if $\omega_t \in \Omega_k$. First observe that $\sum_{g:g\triangleright\hat{g}'} q_t(g) = \sum_{g:g\triangleright\hat{g}} \mathbb{P}(\omega_{0:t-1} = \omega_{0:t-1}^{(g)}) = \mathbb{P}(\omega_{t-1} \in \Omega_{\hat{\omega}_{t-1}^{(\hat{g})}}, \ldots, \omega_0 \in \Omega_{\hat{\omega}_0^{(\hat{g})}}) = \mathbb{P}(\zeta_{0:t-1} = \hat{\omega}_{0:t-1}^{(\hat{g})})$. Also note that $\hat{q}_t(\hat{g}) = \mathbb{P}(\hat{\omega}_{0:t-1} = \hat{\omega}_{0:t-1}^{(\hat{g})})$. So, to show the claim, it suffices to show for any $\sigma_{0:t} \in [r]^t$,

$$|\mathbb{P}(\hat{\omega}_{0:t} = \sigma_{0:t}) - \mathbb{P}(\zeta_{0:t} = \sigma_{0:t})| \leq t(\bar{\mathcal{T}} + \epsilon_{\mathbf{T}})^{t-1}\epsilon_{\mathbf{T}}. \tag{D.21}$$

For the LHS of (D.21), we have

$$\mathbb{P}(\hat{\omega}_{0:t} = \sigma_{0:t}) = \mathbb{P}(\hat{\omega}_0 = \sigma_0) \cdot \textstyle\prod_{h=1}^t \hat{\mathbf{T}}(\sigma_{h-1}, \sigma_h) \tag{D.22}$$

$$\mathbb{P}(\zeta_{0:t} = \sigma_{0:t}) = \mathbb{P}(\omega_0 \in \Omega_{\sigma_0}) \cdot \textstyle\prod_{h=1}^t \tilde{T}_h \tag{D.23}$$

where $\tilde{T}_h := \mathbb{P}(\omega_h \in \Omega_{\sigma_h} \mid \omega_{h-1} \in \Omega_{\sigma_{h-1}}, \ldots, \omega_0 \in \Omega_{\sigma_0})$. Note that $\zeta_{0:t}$ may not be a Markov process when $\epsilon_{\mathbf{T}} \neq 0$, so we cannot drop the past conditional events in (D.23).

Let $\alpha_i := \mathbb{P}(\omega_{h-1} = i \mid \omega_{h-2} \in \Omega_{\sigma_{h-2}}, \ldots, \omega_0 \in \Omega_{\sigma_0})$, then $\tilde{T}_h = \sum_{i \in \Omega_{\sigma_{h-1}}} \big[\mathbb{P}(\omega_h \in \Omega_{\sigma_h} \mid \omega_{h-1} = i) \cdot \mathbb{P}(\omega_{h-1} = i \mid \omega_{h-2} \in \Omega_{\sigma_{h-2}}, \ldots, \omega_0 \in \Omega_{\sigma_0})\big] = \sum_{i \in \Omega_{\sigma_{h-1}}} \big[\big(\sum_{j \in \Omega_{\sigma_h}} \mathbf{T}(i,j)\big)\alpha_i\big]$.

Let $\beta_i := |\Omega_{\sigma_{h-1}}|^{-1}$. For $\hat{\mathbf{T}}(\sigma_{h-1}, \sigma_h)$, by definition in Algorithm 3 and the assumption $\hat{\Omega}_{1:r} = \Omega_{1:r}$, we obtain that $\hat{\mathbf{T}}(\sigma_{h-1}, \sigma_h) = |\Omega_{\sigma_{h-1}}|^{-1} \cdot \sum_{i \in \Omega_{\sigma_{h-1}}} (\sum_{j \in \Omega_{\sigma_h}} \mathbf{T}(i,j)) = \sum_{i \in \Omega_{\sigma_{h-1}}} \big[\big(\sum_{j \in \Omega_{\sigma_h}} \mathbf{T}(i,j)\big)\beta_i\big]$.

Then, it follows that the difference $|\tilde{T}_h - \hat{\mathbf{T}}(\sigma_{h-1}, \sigma_h)| = \big|\sum_{i,i' \in \Omega_{\sigma_{h-1}}} \big[\big(\sum_{j \in \Omega_{\sigma_h}} \mathbf{T}(i,j)\big)\alpha_i\beta_{i'}\big] - \sum_{i,i' \in \Omega_{\sigma_{h-1}}} \big[\big(\sum_{j \in \Omega_{\sigma_h}} \mathbf{T}(i',j)\big)\alpha_i\beta_{i'}\big]\big| \le \sum_{i,i' \in \Omega_{\sigma_{h-1}}} \big[\big|\sum_{j \in \Omega_{\sigma_h}} \mathbf{T}(i,j) - \sum_{j \in \Omega_{\sigma_h}} \mathbf{T}(i',j)\big|\alpha_i\beta_{i'}\big] \le \epsilon_{\mathbf{T}}$, where the first inequality follows from triangle inequality on the absolute values; the second inequality holds since the definition of perturbation $\epsilon_{\mathbf{T}}$ in either Problem P6.1 or P6.2 gives $|\sum_{j \in \Omega_{\sigma_h}} \mathbf{T}(i,j) - \sum_{j \in \Omega_{\sigma_h}} \mathbf{T}(i',j)| \le \epsilon_{\mathbf{T}}$ for any $i, i' \in \Omega_{\sigma_{h-1}}$.

We have established upper bounds for the differences between each multiplier in (D.22) and (D.23), by Corollary D.10, we obtain $|\mathbb{P}(\hat{\omega}_{0:t} = \sigma_{0:t}) - \mathbb{P}(\zeta_{0:t} = \sigma_{0:t})| \le t(\bar{\mathcal{T}} + \epsilon_{\mathbf{T}})^{t-1}\epsilon_{\mathbf{T}}$ which shows (D.21) and concludes the proof. □

Main Proof for Theorem 6.8 (T2). To lower bound the Wasserstein distance $W_\ell(p_t, \hat{p}_t)$ defined in the mass transportation problem (6.11), we consider the objective value given by a constrained mass transportation scheme. Recall with measures q_t and $\hat{q}_t$, we have $p_t(\mathbf{x}) = \sum_{g:\mathbf{x}_t^{(g)}=\mathbf{x}} q_t(g)$ and $\hat{p}_t(\hat{\mathbf{x}}) = \sum_{\hat{g}:\hat{\mathbf{x}}_t^{(\hat{g})}=\hat{\mathbf{x}}} \hat{q}_t(\hat{g})$. With these relations, we consider the following transportation scheme in terms of q_t and $\hat{q}_t$: for all the mass $q_t(g)$ with mode sequence $\omega_t^{(g)}$ synchronous to mode sequence $\hat{\omega}_t^{(\hat{g})}$, it is prioritized to be moved to location $\hat{g}$; if there is surplus, i.e., $\sum_{g:g\triangleright\hat{g}} q_t(g) > \hat{q}_t(\hat{g})$, we move the surplus portion $\sum_{g:g\triangleright\hat{g}} q_t(g) - \hat{q}_t(\hat{g})$ elsewhere.

Under this moving scheme, let $\bar{W}_\ell(q_t, \hat{q}_t)$ denote the optimal objective value of the mass transportation problem (6.11). Let $\hat{\mathcal{G}}_1 = \{\hat{g} : \hat{g} \in [r^t], \sum_{g:g\triangleright\hat{g}} q_t(g) \le \hat{q}_t(\hat{g})\}$, $\hat{\mathcal{G}}_2 = [r^t] \backslash \hat{\mathcal{G}}_1$. Then, $\bar{W}_\ell(q_t, \hat{q}_t)$ can be viewed as the optimal objective of the following problem:

$$\begin{aligned}
\min_{f \ge 0} \quad & \big(\textstyle\sum_{g \in [s^t], \hat{g} \in [r^t]} f(g, \hat{g}) \|\mathbf{x}_t^{(g)} - \hat{\mathbf{x}}_t^{(\hat{g})}\|^\ell\big)^{1/\ell} && \text{(D.24)} \\
\text{s.t.} \quad & \textstyle\sum_{g \in [s^t]} f(g, \hat{g}) = \hat{q}_t(\hat{g}), \forall\ \hat{g} \\
& \textstyle\sum_{\hat{g} \in [r^t]} f(g, \hat{g}) = q_t(g), \forall\ g. \\
& f(g, \hat{g}) = q_t(g), \qquad \forall\ g \triangleright \hat{g}, \forall\ \hat{g} \in \hat{\mathcal{G}}_1 && \text{(D.25)} \\
& \textstyle\sum_{g \triangleright \hat{g}} f(g, \hat{g}) = \hat{q}_t(\hat{g}), \quad \forall\ \hat{g} \in \hat{\mathcal{G}}_2 && \text{(D.26)}
\end{aligned}$$

where constraints (D.25) and (D.26) characterize the moving scheme outlined above. Without them, the problem reduces to (6.11), thus $W_\ell(p_t, \hat{p}_t) \le \bar{W}_\ell(q_t, \hat{q}_t)$. To prove the main claim, it suffices to show

$$\bar{W}_\ell(q_t, \hat{q}_t) \le t\xi_0^{t-1}\kappa^2\|\mathbf{x}_0\|\epsilon_{\mathbf{A}} + 2r^2 t\kappa\|\mathbf{x}_0\|r^t(\kappa\epsilon_{\mathbf{A}}+\xi)^t(\bar{\mathcal{T}}+\epsilon_{\mathbf{T}})^{\frac{t-2}{\ell}}\epsilon_{\mathbf{T}}^{\frac{1}{\ell}}. \tag{D.27}$$

For all $\hat{g} \in [r^t]$, define its synchrony set $\mathcal{S}(\hat{g}) := \{g : g \in [s^t],\, g \triangleright \hat{g}\}$ and the asynchrony set $\mathcal{S}^c(\hat{g}) := [s^t]\backslash\mathcal{S}(\hat{g})$. For the synchrony flow, define total flow $F_s := \sum_{\hat{g}\in[r^t], g\in\mathcal{S}(\hat{g})} f(g, \hat{g})$ and maximum travel distance $D_s := \max_{\hat{g}\in[r^t], g\in\mathcal{S}(\hat{g})} \|\mathbf{x}_t^{(g)} - \hat{\mathbf{x}}_t^{(\hat{g})}\|$. For the asynchrony flow, similarly define $F_a := \sum_{\hat{g}\in[r^t], g\in\mathcal{S}^c(\hat{g})} f(g, \hat{g})$ and $D_a := \max_{\hat{g}\in[r^t], g\in\mathcal{S}^c(\hat{g})} \|\mathbf{x}_t^{(g)} - \hat{\mathbf{x}}_t^{(\hat{g})}\|$. Then, we have $\bar{W}_\ell(q_t, \hat{q}_t) \le (F_s D_s^\ell + F_a D_a^\ell)^{\frac{1}{\ell}} \le F_s^{\frac{1}{\ell}} D_s + F_a^{\frac{1}{\ell}} D_a$. We next bound F_s, D_s, F_a, D_a separately.

For the synchrony maximum travel distance D_s, since $g \triangleright \hat{g}$, by Theorem 6.8 (T1), we know $D_s \le t\xi_0^{t-1}\kappa^2\|\mathbf{x}_0\|\epsilon_{\mathbf{A}}$. For the synchrony total flow F_s, we simply bound it with $F_s \le 1$.

Now we consider the asynchrony maximum travel distance D_a. First note that for any g and $\hat{g}$, we have $\|\mathbf{x}_t^{(g)} - \hat{\mathbf{x}}_t^{(\hat{g})}\| = \|\prod_{h=0}^{t-1}\mathbf{A}_{\omega_h^{(g)}}\mathbf{x}_0 - \prod_{h=0}^{t-1}\hat{\mathbf{A}}_{\hat{\omega}_h^{(\hat{g})}}\hat{\mathbf{x}}_0\| \le \|\prod_{h=0}^{t-1}\mathbf{A}_{\omega_h^{(g)}}\mathbf{x}_0\| + \|\prod_{h=0}^{t-1}\hat{\mathbf{A}}_{\hat{\omega}_h^{(\hat{g})}}\hat{\mathbf{x}}_0\| \le 2\kappa(\kappa\epsilon_{\mathbf{A}}+\xi)^t\|\mathbf{x}_0\|$, where the second inequality follows from Lemma D.8. Then, it follows that $D_a \le 2\kappa(\kappa\epsilon_{\mathbf{A}}+\xi)^t\|\mathbf{x}_0\|$.

For the asynchrony total flow F_a, define $F_{a,\hat{g}} := \sum_{g\in\mathcal{S}^c(\hat{g})} f(g,\hat{g})$, then $F_a = \sum_{\hat{g}\in[r^t]} F_{a,\hat{g}}$. By constraints (D.25) and (D.26), $F_{a,\hat{g}} = \sum_{g\in\mathcal{S}^c(\hat{g})} f(g,\hat{g}) = \hat{q}_t(\hat{g}) - \sum_{g:g\in\mathcal{S}(\hat{g})} f(g,\hat{g}) = \hat{q}_t(\hat{g}) - \sum_{g:g\triangleright\hat{g}} f(g,\hat{g})$. Thus, if $\hat{g} \in \hat{\mathcal{G}}_2$, $F_{a,\hat{g}} = 0$; and if $\hat{g} \in \hat{\mathcal{G}}_1$, $F_{a,\hat{g}} = \hat{q}_t(\hat{g}) - \sum_{g:g\triangleright\hat{g}} q_t(g) > 0$. For the latter case, according to Lemma D.15, we have $F_{a,\hat{g}} = |\hat{q}_t(\hat{g}) - \sum_{g:g\triangleright\hat{g}} q_t(g)| \le (t-1)(\bar{\mathcal{T}}+\epsilon_{\mathbf{T}})^{t-2}\epsilon_{\mathbf{T}}$, which further implies that $F_a = \sum_{\hat{g}\in[r^t]} F_{a,\hat{g}} \le r^2 t(r(\bar{\mathcal{T}}+\epsilon_{\mathbf{T}}))^{t-2}\epsilon_{\mathbf{T}}$.

Finally, (D.27) can be shown by plugging the upper bounds for F_s, D_s, F_a, D_a into the relation that $\bar{W}_\ell(q_t, \hat{q}_t) \le F_s^{\frac{1}{\ell}} D_s + F_a^{\frac{1}{\ell}} D_a$, which concludes the proof. □

Bibliography

A. Abate. Approximation metrics based on probabilistic bisimulations for general state-space Markov processes: a survey. *Electronic Notes in Theoretical Computer Science*, 297:3–25, 2013.

A. Abate, J.P. Katoen, J. Lygeros, and M. Prandini. Approximate model checking of stochastic hybrid systems. *European Journal of Control*, 16(6):624–641, 2010.

A. Abate, A. D'Innocenzo, and M.D. Di Benedetto. Approximate abstractions of stochastic hybrid systems. *IEEE Transactions on Automatic Control*, 56(11):2688–2694, 2011.

Y. Abbasi-Yadkori and C. Szepesvári. Regret bounds for the adaptive control of linear quadratic systems. In *Proc. of COLT*, pages 1–26. JMLR Workshop and Conference Proceedings, 2011.

Y. Abbasi-Yadkori, N. Lazic, and C. Szepesvári. Model-free linear quadratic control via reduction to expert prediction. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 3108–3117. PMLR, 2019.

M. Abeille and A. Lazaric. Improved regret bounds for thompson sampling in linear quadratic control problems. In *International Conference on Machine Learning*, pages 1–9. PMLR, 2018.

M. Abeille and A. Lazaric. Efficient optimistic exploration in linear-quadratic regulators via lagrangian relaxation. In *ICML*, pages 23–31. PMLR, 2020.

H. Abou-Kandil, G. Freiling, and G. Jank. On the solution of discrete-time Markovian jump linear quadratic control problems. *Automatica*, 31(5):765–768, 1995.

G.A. Ackerson and K.S. Fu. On state estimation in switching environments. *IEEE transactions on automatic control*, 15(1):10–17, 1970.

S. Aghabozorgi, A.S. Shirkhorshidi, and T.Y. Wah. Time-series clustering–a decade review. *Information Systems*, 53:16–38, 2015.

S. Allam, F. Dufour, and P. Bertrand. Discrete-time estimation of a Markov chain with marked point process observations. application to Markovian jump filtering. *IEEE Transactions on Automatic Control*, 46(6):903–908, 2001.

R. Alur, T.A. Henzinger, G. Lafferriere, and G.J. Pappas. Discrete abstractions of hybrid systems. *Proceedings of the IEEE*, 88(7):971–984, 2000.

S. Baltaoglu, L. Tong, and Q. Zhao. Online learning and optimization of Markov jump affine models. *arXiv preprint arXiv:1605.02213*, 2016.

M. Barao and J.S. Marques. Offline Bayesian identification of jump Markov nonlinear systems. *IFAC Proceedings Volumes*, 44(1):7761–7766, 2011.

A. Bemporad, V. Breschi, D. Piga, and S.P. Boyd. Fitting jump models. *Automatica*, 96: 11–21, 2018.

G. Bian and A. Abate. On the relationship between bisimulation and trace equivalence in an approximate probabilistic context. In *International Conference on Foundations of Software Science and Computation Structures*, pages 321–337. Springer, 2017.

A. Bittracher and C. Schütte. A probabilistic algorithm for aggregating vastly undersampled large Markov chains. *Physica D: Nonlinear Phenomena*, 416:132799, 2021.

L. Blackmore, S. Funiak, and B.C. Williams. Combining stochastic and greedy search in hybrid estimation. In *AAAI*, pages 282–287, 2005.

H.A. Blom. An efficient filter for abruptly changing systems. In *The 23rd IEEE Conference on Decision and Control*, pages 656–658. IEEE, 1984.

N.M. Boffi, S. Tu, and J.J.E. Slotine. Regret bounds for adaptive nonlinear control. In *Learning for Dynamics and Control*, pages 471–483. PMLR, 2021.

L. Bottou and Y. Bengio. Convergence properties of the k-means algorithms. *Advances in neural information processing systems*, 7, 1994.

E.K. Boukas and Z.K. Liu. Manufacturing systems with random breakdowns and deteriorating items. *Automatica*, 37(3):401–408, 2001.

E.K. Boukas, Z.K. Liu, and G.X. Liu. Delay-dependent robust stability and $\mathcal{H}_\infty$ control of jump linear systems. *International Journal of Control*, 74(4):329–340, 2001.

P. Buchholz. Exact and ordinary lumpability in finite Markov chains. *Journal of applied probability*, 31(1):59–75, 1994.

P. Caines and H. Chen. Optimal adaptive LQG control for systems with finite state process parameters. *IEEE Transactions on Automatic Control*, 30(2):185–189, 1985.

P.E. Caines and J.F. Zhang. On the adaptive control of jump parameter systems via nonlinear filtering. *SIAM journal on control and optimization*, 33(6):1758–1777, 1995.

D.O. Cajueiro. *Stochastic optimal control of jumping Markov parameter processes with applications to finance*. PhD thesis, PhD thesis, 2002, Instituto Tecnológico de Aeronáutica-ITA, Brazil, 2002.

M.C. Campi and P. Kumar. Adaptive linear quadratic Gaussian control: the cost-biased approach revisited. *SIAM J. Control Optim.*, 36(6):1890–1907, 1998.

Y.Y. Cao and J. Lam. Robust $\mathcal{H}_\infty$ control of discrete-time Markovian jump linear systems with mode-dependent time-delays. *Journal of the Franklin Institute*, 336(8):1263–1281, 1999.

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

H. Chan and U. Ozguner. Optimal control of systems over a communication network with queues via a jump system approach. In *Proceedings of International Conference on Control Applications*, pages 1148–1153. IEEE, 1995.

C.B. Chang and M. Athans. State estimation for discrete systems with switching parameters. *IEEE Transactions on Aerospace and Electronic Systems*, AES-14(3):418–425, 1978. doi: 10.1109/TAES.1978.308603.

X. Chen and E. Hazan. Black-box control for linear dynamical systems. In *Conference on Learning Theory*, pages 1114–1143. PMLR, 2021.

Y. Chen, T.T. Georgiou, and M. Pavon. Optimal steering of a linear stochastic system to a final probability distribution, part i. *IEEE Transactions on Automatic Control*, 61(5): 1158–1169, 2015.

H.J. Chizeck and Y. Ji. Optimal quadratic control of jump linear systems with Gaussian noise in discrete-time. In *Proceedings of the 27th IEEE Conference on Decision and Control*, pages 1989–1993. IEEE, 1988.

H.J. Chizeck, A.S. Willsky, and D. Castanon. Discrete-time Markovian-jump linear quadratic optimal control. *International Journal of Control*, 43(1):213–231, 1986.

G.E. Cho and C.D. Meyer. Comparison of perturbation bounds for the stationary distribution of a Markov chain. *Linear Algebra and its Applications*, 335(1-3):137–150, 2001.

E. Cinquemani, R. Porreca, G. Ferrari-Trecate, and J. Lygeros. A general framework for the identification of jump Markov linear systems. In *2007 46th IEEE Conference on Decision and Control*, pages 5737–5742. IEEE, 2007.

E.M. Clarke Jr., O. Grumberg, D. Kroening, D. Peled, and H. Veith. *Model checking*. MIT press, 2018.

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

E.F. Costa and J.B.R. do Val. Weak detectability and the linear-quadratic control problem of discrete-time Markov jump linear systems. *International Journal of Control*, 75(16-17):

1282–1292, 2002.

O.L.V. Costa and J.B.R. do Val. Full information $\mathcal{H}_\infty$-control for discrete-time infinite Markov jump parameter systems. *Journal of Mathematical Analysis and Applications*, 202(2):578–603, 1996.

O.L.V. Costa, E.O. Assumpção Filho, E.K. Boukas, and R.P. Marques. Constrained quadratic state feedback control of discrete-time Markovian jump linear systems. *Automatica*, 35(4):617–626, 1999.

O.L.V. Costa, M.D. Fragoso, and R.P. Marques. *Discrete-time Markov jump linear systems*. Springer, 2006.

D.P. De Farias, J.C. Geromel, J.B.R. do Val, and O.L.V. Costa. Output feedback control of Markov jump linear systems in continuous-time. *IEEE Transactions on Automatic Control*, 45(5):944–949, 2000.

D.P. de Farias, J.C. Geromel, and J.B.R. do Val. A note on the robust control of Markov jump linear uncertain systems. *Optimal Control Applications and Methods*, 23(2):105–112, 2002.

S. Dean, H. Mania, N. Matni, B. Recht, and S. Tu. Regret bounds for robust adaptive control of the linear quadratic regulator. In *Advances in Neural Information Processing Systems*, pages 4188–4197, 2018.

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

J. Desharnais, A. Edalat, and P. Panangaden. Bisimulation for labelled Markov processes. *Information and Computation*, 179(2):163–193, 2002.

J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled Markov processes. *Theoretical computer science*, 318(3):323–354, 2004.

Z. Du, L. Balzano, and N. Ozay. A robust algorithm for online switched system identification. *IFAC-PapersOnLine*, 51(15):293–298, 2018.

Z. Du, N. Ozay, and L. Balzano. Mode clustering for Markov jump systems. In *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 126–130, 2019a. doi: 10.1109/CAMSAP45676.2019.9022650.

Z. Du, N. Ozay, and L. Balzano. Mode clustering for Markov jump systems. *arXiv preprint arXiv:1910.02193*, 2019b.

Z. Du, L. Balzano, and N. Ozay. Mode reduction for Markov jump systems. *IEEE Open Journal of Control Systems*, 1:335–353, 2022a. doi: 10.1109/OJCSYS.2022.3212613.

Z. Du, N. Ozay, and L. Balzano. Clustering-based mode reduction for Markov jump systems. In *Proceedings of The 4th Annual Learning for Dynamics and Control Conference*, volume

168 of *Proceedings of Machine Learning Research*, pages 689–701. PMLR, 23–24 Jun 2022b.

Z. Du, Y. Sattar, D.A. Tarzanagh, L. Balzano, N. Ozay, and S. Oymak. Data-driven control of Markov jump systems: Sample complexity and regret bounds. In *2022 American Control Conference (ACC)*, pages 4901–4908, 2022c. doi: 10.23919/ACC53348.2022.9867863.

M.K.S. Faradonbeh, A. Tewari, and G. Michailidis. Finite time identification in unstable linear systems. *Automatica*, 96:342–353, 2018a.

M.K.S. Faradonbeh, A. Tewari, and G. Michailidis. Finite-time adaptive stabilization of linear systems. *IEEE Transactions on Automatic Control*, 64(8):3498–3505, 2018b.

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

M.K.S. Faradonbeh, A. Tewari, and G. Michailidis. Optimism-based adaptive regulation of linear-quadratic systems. *IEEE Transactions on Automatic Control*, 2020b.

M. Fazel, R. Ge, S. Kakade, and M. Mesbahi. Global convergence of policy gradient methods for the linear quadratic regulator. In *International Conference on Machine Learning*, pages 1467–1476. PMLR, 2018.

D. Foster, T. Sarkar, and A. Rakhlin. Learning nonlinear dynamical systems from a single trajectory. In *Learning for Dynamics and Control*, pages 851–861. PMLR, 2020.

D.A. Freedman. On tail probabilities for martingales. *the Annals of Probability*, pages 100–118, 1975.

R.G. Gallager. *Stochastic processes: theory for applications.* Cambridge University Press, 2013.

K. Gatsis and G.J. Pappas. Statistical learning for analysis of networked control systems over unknown channels. *Automatica*, 125:109386, 2021.

J.E. Gaudio, T.E. Gibson, A.M. Annaswamy, M.A. Bolender, and E. Lavretsky. Connections between adaptive control and optimization in machine learning. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 4563–4568. IEEE, 2019.

B. Gaveau and L.S. Schulman. Dynamical distance: coarse grains, pattern recognition, and network analysis. *Bulletin des sciences mathematiques*, 129(8):631–642, 2005.

A. Girard and G.J. Pappas. Approximation metrics for discrete and continuous systems. *IEEE Transactions on Automatic Control*, 52(5):782–798, 2007.

M. Goldshtein and P. Tsiotras. Finite-horizon covariance control of linear time-varying systems. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 3606–3611. IEEE, 2017.

T.F. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical* *computer science*, 38:293–306, 1985.

S. Gugercin and A.C. Antoulas. A survey of model reduction by balanced truncation and some new results. *International Journal of Control*, 77(8):748–766, 2004.

S.J. Hatjispyros and A.N. Yannacopoulos. A random dynamical system model of a stylized equity market. *Physica A: Statistical Mechanics and its Applications*, 347:583–612, 2005.

E. Hazan, S. Kakade, and K. Singh. The nonstochastic control problem. In *Algorithmic Learning Theory*, pages 408–421. PMLR, 2020.

P. Hespanhol and A. Aswani. Statistical consistency of set-membership estimator for linear systems. *IEEE Control Systems Letters*, 4(3):668–673, 2020.

K.H. Hoffmann and P. Salamon. Bounding the lumping error in Markov chain dynamics. *Applied mathematics letters*, 22(9):1471–1475, 2009.

D. Hsu, S. Kakade, and T. Zhang. A tail inequality for quadratic forms of subgaussian random vectors. *Electronic Communications in Probability*, 17:1–6, 2012.

B. Hu and U.A. Syed. Characterizing the exact behaviors of temporal difference learning algorithms using Markov jump linear system theory. *Advances in Neural Information Processing Systems*, 32, 2019.

B. Hu, P. Seiler, and A. Rantzer. A unified analysis of stochastic optimization methods using jump system theory and quadratic constraints. In *Conference on Learning Theory*, pages 1157–1189. PMLR, 2017.

M. Ibrahimi, A. Javanmard, and B. Van Roy. Efficient reinforcement learning for high dimensional linear quadratic systems. In *NeurIPS*, pages 2645–2653, 2012.

J.P. Jansch-Porto, B. Hu, and G. Dullerud. Policy learning of MDPs with mixed continuous/discrete variables: A case study on model-free control of Markovian jump systems. In *Learning for Dynamics and Control*, pages 947–957. PMLR, 2020.

Y. Jedra and A. Proutiere. Finite-time identification of stable linear systems optimality of the least-squares estimator. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 996–1001. IEEE, 2020.

A.A. Julius and G.J. Pappas. Approximations of stochastic hybrid systems. *IEEE Transactions on Automatic Control*, 54(6):1193–1203, 2009.

A.A. Julius, A. Girard, and G.J. Pappas. Approximate bisimulation for a class of stochastic hybrid systems. In *2006 American Control Conference*, pages 6–pp. IEEE, 2006.

R. Jungers. *The joint spectral radius: theory and applications*, volume 385. Springer Science & Business Media, 2009.

C. Kenney and G. Hewer. The sensitivity of the algebraic and differential Riccati equations.

SIAM J. Control Optim., 28(1):50–69, 1990.

M. Konstantinov, P.H. Petkov, and N.D. Christov. Perturbation analysis of the discrete Riccati equation. *Kybernetika*, 29(1):18–29, 1993.

M. Konstantinov, V. Angelova, P. Petkov, D. Gu, and V. Tsachouridis. Perturbation analysis of coupled matrix Riccati equations. *IFAC Proc. Volumes*, 35(1):307–312, 2002.

M. Konstantinov, V. Angelova, P. Petkov, D. Gu, and V. Tsachouridis. Perturbation bounds for coupled matrix riccati equations. *Linear algebra and its applications*, 359(1-3):197–218, 2003a.

M. Konstantinov, D. Gu, V. Mehrmann, and P. Petkov. *Perturbation theory for matrix equations*. Gulf Professional Publishing, 2003b.

G. Kotsalis and A. Rantzer. Balanced truncation for discrete time Markov jump linear systems. *IEEE Transactions on Automatic Control*, 55(11):2606–2611, 2010.

D. Kuhn, P.M. Esfahani, V.A. Nguyen, and S. Shafieezadeh-Abadeh. Wasserstein distributionally robust optimization: Theory and applications in machine learning. In *Operations research & management science in the age of analytics*, pages 130–166. Informs, 2019.

A. Kumar, Y. Sabharwal, and S. Sen. A simple linear time $(1+\epsilon)$-approximation algorithm for k-means clustering in any dimensions. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 454–462. IEEE, 2004.

R.P. Kurshan. *Computer-aided verification of coordinating processes: the automata-theoretic approach*, volume 302. Princeton university press, 2014.

V. Kuznetsov and M. Mohri. Generalization bounds for non-stationary mixing processes. *Machine Learning*, 106(1):93–117, 2017.

S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar. Explore more and improve regret in linear quadratic regulators. *arXiv preprint arXiv:2007.12291*, 2020a.

S. Lale, K. Azizzadenesheli, B. Hassibi, and A. Anandkumar. Logarithmic regret bound in partially observable linear dynamical systems. In *Advances in Neural Information Processing Systems*, 2020b.

S. Lale, O. Teke, B. Hassibi, and A. Anandkumar. Stability and identification of random asynchronous linear time-invariant systems. In *Learning for Dynamics and Control*, pages 651–663. PMLR, 2021.

A. Lamperski. Computing stabilizing linear controllers via policy iteration. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 1902–1907. IEEE, 2020.

K.G. Larsen and A. Skou. Bisimulation through probabilistic testing. *Information and computation*, 94(1):1–28, 1991.

F. Lauer and G. Bloch. *Hybrid System Identification : Theory and Algorithms for Learning Switching Models*. Springer Cham, 2018.

B. Lee and A. Lamperski. Non-asymptotic closed-loop system identification using autoregressive processes and hankel model reduction. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 3419–3424. IEEE, 2020.

J.W. Lee and G.E. Dullerud. Uniform stabilization of discrete-time switched and Markovian jump linear systems. *Automatica*, 42(2):205–218, 2006. ISSN 0005-1098. doi: https://doi.org/10.1016/j.automatica.2005.08.019.

J. Lei and A. Rinaldo. Consistency of spectral clustering in stochastic block models. *The Annals of Statistics*, 43(1):215–237, 2015.

D.A. Levin and Y. Peres. *Markov chains and mixing times*, volume 107. American Mathematical Soc., 2017.

F.L. Lewis, D. Vrabie, and V.L. Syrmos. *Optimal control*. John Wiley & Sons, 2012.

J.S. Li and J. Qi. Ensemble control of time-invariant linear systems with linear parameter variation. *IEEE Transactions on Automatic Control*, 61(10):2808–2820, 2015.

D. Liberzon. *Switching in systems and control*. Springer Science & Business Media, 2003.

L. Ljung. *System Identification: Theory for the User*. Prentice Hall information and system sciences series. Prentice Hall, 1999. ISBN 9780136566953.

K.A. Loparo and F. Abdel-Malek. A probabilistic approach to dynamic power system security. *IEEE transactions on circuits and systems*, 37(6):787–798, 1990.

Y.Z. Lun, J. Wheatley, A. D'Innocenzo, and A. Abate. Approximate abstractions of Markov chains with interval decision processes. *IFAC-PapersOnLine*, 51(16):91–96, 2018.

O. Maler, A. Pnueli, and J. Sifakis. On the synthesis of discrete controllers for timed systems. In *Annual symposium on theoretical aspects of computer science*, pages 229–242. Springer, 1995.

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

N. Matni, A. Proutiere, A. Rantzer, and S. Tu. From self-tuning regulators to reinforcement learning and back again. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 3724–3740. IEEE, 2019.

M. Meilă and J. Shi. A random walks view of spectral segmentation. In *Proceedings of the Eighth International Workshop on Artificial Intelligence and Statistics*, volume R3 of *Proceedings of Machine Learning Research*, pages 203–208. PMLR, 04–07 Jan 2001.

H. Mohammadi, M. Soltanolkotabi, and M.R. Jovanović. On the linear convergence of random search for discrete-time LQR. *IEEE Control Systems Letters*, 5(3):989–994, 2020.

M. Mohri and A. Rostamizadeh. Stability bounds for non-iid processes. In *Advances in Neural Information Processing Systems*, pages 1025–1032, 2008.

K. Okamoto, M. Goldshtein, and P. Tsiotras. Optimal covariance control for stochastic systems under chance constraints. *IEEE Control Systems Letters*, 2(2):266–271, 2018.

S. Oymak. Stochastic gradient descent learns state equations with nonlinear activations. In *Conference on Learning Theory*, pages 2551–2579, 2019.

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

N. Ozay, M. Sznaier, C.M. Lagoa, and O.I. Camps. A sparsification approach to set membership identification of switched affine systems. *IEEE Transactions on Automatic Control*, 57(3):634–648, 2011.

N. Ozay, C. Lagoa, and M. Sznaier. Set membership identification of switched linear systems with known number of subsystems. *Automatica*, 51:180–191, Jan. 2015. ISSN 0005-1098.

E. Özkan, F. Lindsten, C. Fritsche, and F. Gustafsson. Recursive maximum likelihood identification of jump Markov nonlinear systems. *IEEE Transactions on Signal Processing*, 63(3):754–765, 2014.

R. Palm and D. Driankov. Fuzzy switched hybrid systems-modeling and identification. In *Proceedings of the 1998 IEEE International Symposium on Intelligent Control (ISIC) held jointly with IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA) Intell*, pages 130–135. IEEE, 1998.

P.A. Parrilo and A. Jadbabaie. Approximation of the joint spectral radius using sum of squares. *Linear Algebra and its Applications*, 428(10):2385–2402, 2008.

B. Recht. A tour of reinforcement learning: The view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*, 2:253–279, 2019.

M. Rudelson and R. Vershynin. Hanson-Wright inequality and sub-Gaussian concentration. *Electronic Communications in Probability*, 18:1–9, 2013.

T. Sarkar and A. Rakhlin. Near optimal finite time identification of arbitrary linear dynamical systems. In *International Conference on Machine Learning*, pages 5610–5618. PMLR, 2019.

T. Sarkar, A. Rakhlin, and M. Dahleh. Nonparametric system identification of stochastic switched linear systems. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 3623–3628. IEEE, 2019.

Y. Sattar and S. Oymak. Non-asymptotic and accurate learning of nonlinear dynamical systems. *Journal of Machine Learning Research*, 23(140):1–49, 2022.

Y. Sattar, Z. Du, D.A. Tarzanagh, L. Balzano, N. Ozay, and S. Oymak. Identification and adaptive control of Markov jump systems: Sample complexity and regret bounds. *arXiv preprint arXiv:2111.07018*, 2021. doi: 10.48550/ARXIV.2111.07018.

Y. Sattar, Z. Du, D.A. Tarzanagh, S. Oymak, L. Balzano, and N. Ozay. Certainty equivalent quadratic control for Markov jump systems. In *2022 American Control Conference (ACC)*, pages 2871–2878, 2022. doi: 10.23919/ACC53348.2022.9867208.

L. Schulman and B. Gaveau. Coarse grains: The emergence of space and order. *Foundations of Physics*, 31(4):713–731, 2001. doi: 10.1023/A:1017577211902.

M. Schuurmans, P. Sopasakis, and P. Patrinos. Safe learning-based control of stochastic jump linear systems: a distributionally robust approach. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 6498–6503. IEEE, 2019.

Y. Shen, Z.G. Wu, P. Shi, and C.K. Ahn. Model reduction of Markovian jump systems with uncertain probabilities. *IEEE Transactions on Automatic Control*, 65(1):382–388, 2019.

P. Shi, E.K. Boukas, and R.K. Agarwal. Control of Markovian jump discrete-time systems with norm bounded uncertainty and unknown delay. *IEEE Transactions on automatic control*, 44(11):2139–2144, 1999.

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

M. Simchowitz, H. Mania, S. Tu, M.I. Jordan, and B. Recht. Learning without mixing: Towards a sharp analysis of linear system identification. In *Conference On Learning Theory*, pages 439–473. PMLR, 2018.

B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, and S. Sastry. An LQG optimal linear controller for control systems with packet losses. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 458–463. IEEE, 2005.

S.C. Smith and P. Seiler. Estimation with lossy measurements: jump estimators for jump systems. *IEEE Transactions on Automatic Control*, 48(12):2163–2171, 2003.

M. Song and S. Rajasekaran. Fast algorithms for constant approximation k-means clustering. *Transactions on Machine Learning and Data Mining*, 3(2):67–79, 2010.

S.E.Z. Soudjani and A. Abate. Adaptive gridding for abstraction and verification of stochastic hybrid systems. In *2011 Eighth International Conference on Quantitative Evaluation of SysTems*, pages 59–68. IEEE, 2011.

A.M. Stankovic, G.C. Verghese, and D.J. Perreault. Analysis and synthesis of randomized modulation schemes for power converters. *IEEE Transactions on Power Electronics*, 10 (6):680–693, 1995.

J.G. Sun. Perturbation theory for algebraic Riccati equations. *SIAM Journal on Matrix Analysis and Applications*, 19(1):39–65, 1998.

J.G. Sun. Condition numbers of algebraic Riccati equations in the Frobenius norm. *Linear algebra and its applications*, 350(1-3):237–261, 2002.

M. Sun and J. Lam. Model reduction of discrete Markovian jump systems with time-weighted h2 performance. *International Journal of Robust and Nonlinear Control*, 26(3):401–425, 2016.

L.E.O. Svensson and N. Williams. Optimal monetary policy under uncertainty: a Markov jump-linear-quadratic approach. *Federal Reserve Bank of St. Louis Review*, 90(4):275–293, 2008.

D. Sworder and R. Rogers. An LQ-solution to a control problem associated with a solar thermal central receiver. *IEEE Transactions on Automatic Control*, 28(10):971–978, 1983. doi: 10.1109/TAC.1983.1103151.

S. Tan, J.F. Zhang, and L. Yao. Optimality analysis of adaptive sampled control of hybrid systems with quadratic index. *IEEE transactions on automatic control*, 50(7):1044–1051, 2005.

I. Tkachev and A. Abate. On approximation metrics for linear temporal model-checking of stochastic systems. In *Proceedings of the 17th international conference on Hybrid systems: computation and control*, pages 193–202, 2014.

L.N. Trefethen and M. Embree. *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press, 2005. ISBN 9780691119465.

T.H.T. Truong, P. Seiler, and L.E. Linderman. Analysis of networked structural control with packet loss. *IEEE Transactions on Control Systems Technology*, 2021.

A. Tsiamis and G.J. Pappas. Finite sample analysis of stochastic system identification. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 3648–3654. IEEE, 2019.

V. Ugrinovskii and H.R. Pota. Decentralized control of power systems via robust control of uncertain Markov jump parameter systems. *International Journal of Control*, 78(9): 662–677, 2005. doi: 10.1080/00207170500105384.

A.N. Vargas, J.Y. Ishihara, and J.B.R. do Val. Linear quadratic regulator for a class of Markovian jump systems with control in jumps. In *49th IEEE Conference on Decision and Control (CDC)*, pages 2282–2285. IEEE, 2010.

A.N. Vargas, W. Furloni, and J.B.R. do Val. Second moment constraints and the control problem of Markov jump linear systems. *Numerical Linear Algebra with Applications*, 20 (2):357–368, 2013.

F.V. Vergés and M.D. Fragoso. Optimal linear mean square filter and an associated stationary filter for hidden Markov chain. In *2020 59th IEEE Conference on Decision and Control (CDC)*, pages 4865–4870. IEEE, 2020.

R. Vershynin. *Introduction to the non-asymptotic analysis of random matrices*, page 210–268. Cambridge University Press, 2012. doi: 10.1017/CBO9780511794308.006.

H.N. Wu and K.Y. Cai. Mode-independent robust stabilization for uncertain Markovian jump nonlinear systems via fuzzy control. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 36(3):509–519, 2006.

F. Xue and L. Guo. Necessary and sufficient conditions for adaptive stablizability of jump linear systems. *Communications in Information and Systems*, 1(2):205–224, 2001.

C. Yang, P. Bertrand, and M. Mariton. Adaptive control in the presence of Markovian parameter jumps. *International Journal of Control*, 52(2):473–484, 1990.

M. Zamani and A. Abate. Approximately bisimilar symbolic models for randomly switched stochastic systems. *Systems & Control Letters*, 69:38–46, 2014.

M. Zamani, M. Rungger, and P.M. Esfahani. Approximations of stochastic hybrid systems: A compositional approach. *IEEE Transactions on Automatic Control*, 62(6):2838–2853, 2016.

A. Zhang and M. Wang. Spectral state compression of Markov processes. *IEEE transactions on information theory*, 66(5):3202–3231, 2019.

G. Zhang, X. Bao, L. Lessard, and R. Grosse. A unified analysis of first-order methods for smooth games via integral quadratic constraints. *Journal of Machine Learning Research*, 22(103):1–39, 2021.

K. Zhang, B. Hu, and T. Basar. Policy optimization for $\mathcal{H}_2$ linear control with $\mathcal{H}_\infty$ robustness guarantee: Implicit regularization and global convergence. In *Learning for Dynamics and Control*, pages 179–190. PMLR, 2020.

L. Zhang, B. Huang, and J. Lam. $\mathcal{H}_\infty$ model reduction of Markovian jump linear systems. *Systems & Control Letters*, 50(2):103–118, 2003.

Y. Zheng, Y. Tang, and N. Li. Analysis of the optimization landscape of linear quadratic Gaussian (LQG) control. *arXiv preprint arXiv:2102.04393*, 2021.

X. Zhong, H. He, H. Zhang, and Z. Wang. Optimal control for unknown discrete-time nonlinear Markov jump systems using adaptive dynamic programming. *IEEE Transactions on Neural Networks and Learning Systems*, 25(12):2141–2155, 2014.

L. Zhou, Y. Lin, Y. Wei, and S. Qiao. Perturbation analysis and condition numbers of symmetric algebraic Riccati equations. *Automatica*, 45(4):1005–1011, 2009.

I.M. Ziemann, H. Sandberg, and N. Matni. Single trajectory nonparametric learning of nonlinear dynamics. In *Conference on Learning Theory*, pages 3333–3364. PMLR, 2022.