Bias analysis in mode-based Kalman filters for stochastic hybrid systems

by

Wenji Zhang

B.E., University of Electronic Science and Technology of China, 2014

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Abstract

Stochastic hybrid system (SHS) is a class of dynamical systems that experience interaction of both discrete mode and continuous dynamics with uncertainty. State estimation for SHS has attracted research interests for decades with Kalman filter based solutions dominating the area. Mode-based Kalman filter is an extended version of the traditional Kalman filter for SHS. In general, as Kalman filter is unbiased for non-hybrid system estimation, prior research efforts primarily focus on the behavior of error covariance. In SHS state estimate, mode mismatch errors could result in a bias in the mode-based Kalman filter and have impacts on the continuous state estimation quality. The relationship between mode mismatch errors and estimation stability is an open problem that this dissertation attempts to address. Specifically, the probabilistic model of mode mismatch errors can be independent and identically distributed (i.i.d.), correlated across different modes and correlated across time. The proposed approach builds on the idea of modeling the bias evolution as a transformed system. The statistical convergence of the bias dynamics is then mapped to the stability of the transformed system. For each specific model of the mode mismatch error, the system matrix of the transformed system varies which results in challenges for the stability analysis. For the first time, the dissertation derives convergence conditions that provide tolerance regions for the mode mismatch error for three mode mismatch situations. The convergence conditions are derived based on generalized spectral radius theorem, Lyapunov theorem, Schur stability of a matrix polytope and interval matrix method. This research is fundamental in nature and its application is widespread. For example, the spatially and timely correlated mode mismatch errors can effectively capture cyber-attacks and communication link impairments in a cyber-physical system. Therefore, the theory and techniques developed in this dissertation can be used to analyze topology errors in any networked system such as smart grid, smart home, transportation, flight management system etc. The main results provide new

insights on the fidelity in discrete state knowledge needed to maintain the performance of a mode-based Kalman filter and provide guidance on design of estimation strategies for SHS.

Bias analysis in mode-based Kalman filters for stochastic hybrid systems

by

Wenji Zhang

B.E., University of Electronic Science and Technology of China, 2014

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2018

Approved by:

Major Professor
Balasubramaniam Natarajan

Copyright

Abstract

Stochastic hybrid system (SHS) is a class of dynamical systems that experience interaction of both discrete mode and continuous dynamics with uncertainty. State estimation for SHS has attracted research interests for decades with Kalman filter based solutions dominating the area. Mode-based Kalman filter is an extended version of the traditional Kalman filter for SHS. In general, as Kalman filter is unbiased for non-hybrid system estimation, prior research efforts primarily focus on the behavior of error covariance. In SHS state estimate, mode mismatch errors could result in a bias in the mode-based Kalman filter and have impacts on the continuous state estimation quality. The relationship between mode mismatch errors and estimation stability is an open problem that this dissertation attempts to address. Specifically, the probabilistic model of mode mismatch errors can be independent and identically distributed (i.i.d.), correlated across different modes and correlated across time. The proposed approach builds on the idea of modeling the bias evolution as a transformed system. The statistical convergence of the bias dynamics is then mapped to the stability of the transformed system. For each specific model of the mode mismatch error, the system matrix of the transformed system varies which results in challenges for the stability analysis. For the first time, the dissertation derives convergence conditions that provide tolerance regions for the mode mismatch error for three mode mismatch situations. The convergence conditions are derived based on generalized spectral radius theorem, Lyapunov theorem, Schur stability of a matrix polytope and interval matrix method. This research is fundamental in nature and its application is widespread. For example, the spatially and timely correlated mode mismatch errors can effectively capture cyber-attacks and communication link impairments in a cyber-physical system. Therefore, the theory and techniques developed in this dissertation can be used to analyze topology errors in any networked system such as smart grid, smart home, transportation, flight management system etc. The main results provide new

insights on the fidelity in discrete state knowledge needed to maintain the performance of a mode-based Kalman filter and provide guidance on design of estimation strategies for SHS.

Table of Contents

List of Figures

List of Tables

Notations

Symbols

$\mathbf{x}_k$	Continuous state at time k
$\mathbf{y}_k$	Measurements at time k
q_i	Discrete mode, $i = 1, \cdots, d$
$\mathcal{Q}$	Set of discrete state, $\mathcal{Q} = \{q_1, \cdots, q_d\}$
δ_k	Actual discrete state at time k
γ_k	Estimated/Measured discrete state at time k
$\boldsymbol{\delta}_k^s$	Sequence of actual discrete states up to time k, $\boldsymbol{\delta}_k^s = (\delta_1, \cdots, \delta_k)$
$\boldsymbol{\gamma}_k^s$	Sequence of measured discrete states up to time k, $\boldsymbol{\gamma}_k^s = (\gamma_1, \cdots, \gamma_k)$
$\mathbf{y}_k^s$	Sequence of measurements up to time k, $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$
$\mathbb{N}^+$	Space of positive integers
$\mathbb{R}$	Space of real numbers

Operations

$[\cdot]^{-1}$	Inverse of a matrix
$[\cdot]'$	Transpose of a matrix/vector
$\|\cdot\|$	2-norm of a matrix/vector
$\|\cdot\|$	Absolute value of a scalar
$\rho(\cdot)$	Spectral radius of a matrix
$\mathbb{E}(\cdot)$	Expectation of a random variable
$\mathbb{P}(\cdot)$	Probability of a random event
$\mathrm{P}_{\gamma_k}^{q_i}$	Short notation for probability of $\gamma_k = q_i$
$\mathbf{A} \otimes \mathbf{B}$	Kronecker product of two matrices
$\overline{\mathbf{A}}$	Conjugate of matrix $\mathbf{A}$
$\mathbf{a}^{[i]}$	ith element of vector $\mathbf{a}$
$\mathbb{Tr}(\cdot)$	Trace of a matrix
$diag[a_1, \cdots, a_n]$	Diagonal matrix with $a_1, \cdots, a_n$ as diagonal elements
$\mathbf{A} \succ 0$	matrix $\mathbf{A}$ is symmetric positive definite (s.p.d.) matrix
$\mathbf{A} \prec 0$	matrix $\mathbf{A}$ is symmetric negative definite (s.p.d.) matrix

Abbreviations

SHS	Stochastic hybrid system
SDSHS	State-dependant stochastic hybrid system
CPS	Cyber-physical system
MJLS	Markov jump linear system
PV	Photovoltaics
IMM	Interacting multiple model
MMAE	Multiple model adaptive estimation
i.i.d.	Independent and identically distributed
MMSE	Minimum mean square error
MSE	Mean square error
MVUE	Minimum variance unbiased estimator
EPT	Errors per transition
AEPT	Average errors per transition
QLF	Quadratic Lyapunov function
CQLF	Common quadratic Lyapunov function
LMI	Linear matrix inequality
BIBO	Bounded-input bounded-output
NCS	Networked control system
WSS	Wide sense stationary
LQR	Linear quadratic regulator
MSS	Mean squared stability

Acknowledgments

I would like to express my sincere gratitude to my advisor Dr. Bala Natarajan for his guidance, encouragement and support over the years. His passion and enthusiasm kept me constantly engaged with my research, and his personality inspired me to be a lifelong learner. This dissertation would not have been possible without his patient and persistent help. My special appreciation to my doctoral committee members Dr. Pavithra Prabhakar, Dr. Nathan Albin, Dr. Behrooz Mirafzal and my outside chairperson Dr. Wei-Wen Hsu for their constructive advice that helps improve the quality of my research.

I am also grateful to former WICOM lab members Dr. Kan Chen, Dr. Chang Liu, Dr. Kumar Jhala, Dr. S. M. Shafiul Alam and Dr. Mohammed Taj-Eldin for their valuable guidance and mentoring in both research and future career. Additionally, I am thankful to current lab members Alaleh Alivar, Hazhar Sufi Karimi, Solmaz Niknam and Reza Barazideh for their inspiration and assistance. Thanks to my friends at K-State Futing Fan, Xin Li, Haotian Wu, Tianyu Lin, Qihui Yang and Bo Liu for all the fun memories that we shared together.

My sincere appreciation to my parents Zhihua Zhang and Jianying Cao. Thanks for your continuous support over the years and countless sacrifices to raise me up to what I am today. Special thanks to my boyfriend Xiangpeng Li whose kindness, wisdom and endless love encourage me to achieve my dreams over the years. Thanks to my younger self. Thank you for not giving up.

This work was funded in part by the Electrical Power Affiliates Program (EPAP) and NSF-CNS-1544705.

Dedication

To Xiangpeng. You are my light in the darkness.

Chapter 1

Introduction

In this chapter, we introduce the background of this dissertation. Using a motivating example, we introduce the research questions of interest and highlight the contributions of this dissertation.

1.1 Stochastic Hybrid System

Stochastic hybrid systems (SHS) are dynamical systems that involve the interaction of continuous and discrete dynamics (also referred to as mode) with uncertainties. The uncertainties can be part of the continuous dynamics, discrete state transitions, or both. In most cases, the evolution of the continuous state is described via a stochastic differential/difference equation (SDE) whereas the discrete state evolves depending on the application. Discrete state evolution typically follows a random process (such as a Markov chain) or guard conditions (i.e., the discrete state transitions depend on the continuous state). SHS models have been widely used to model cyber-physical systems (CPS) thanks to its capability to capture complex dynamics. The applications of SHS include, but are not limited to, modeling of biochemical processes [1, 2], manufacturing processes [3], communication networks [4], flight management systems [5, 6] and smart grid [7], etc.

Based on different models for the continuous dynamics and discrete transitions, SHS

can be categorized into several subclasses. For example, a state-dependant SHS (SDSHS) denotes an SHS in which the discrete mode transitions are governed by guard conditions. Another important category of SHS is Markov jump linear systems (MJLS). MJLS models are applicable to systems that can be represented by a set of linear systems with modal transitions governed by a Markov chain. MJLS has attracted significant attention in the research community due to its analytical tractability as well as applicability to practical systems, e.g., microgrid [8], networked control systems [9], etc. At a higher level, the generalized SHS can be abstracted as a switched system with arbitrary switchings. This allows researchers to neglect the details of the discrete behavior and instead focus on all possible switching patterns. This represents a significant departure from hybrid systems, especially at the analysis stage [10].

Like many other dynamical systems, a successful and reliable implementation of SHS presents several challenges. Specifically, stability [11–14], reachability [15, 16], situational awareness (state estimation) [17–19] related problems for a general class of SHS are still an ongoing research field and lack a universal solution. The above mentioned challenges are critical for analysis of SHS and need to be well addressed. In this dissertation, we investigate state estimation strategies for SHS and study the impact of inaccurate mode information on continuous state estimation. Before we highlight the contributions of this dissertation, we first introduce a motivating example of a smart grid system that illustrates the mode mismatch problem and the contributions of this work to the community.

1.2 Motivating Example

The conventional power grid is transforming to a "smart grid" with the addition of renewable energy sources (e.g., photovoltaics (PV)), advanced metering and sensing infrastructure, electric vehicles and controllable loads [7]. Integration of these technologies enable it to deliver affordable electric power with improved reliability, security and efficiency. We highlight three characteristic behaviors that may be observed in a smart grid:

- Discrete behaviors: Smart grid contains a variety of operation modes that depend on

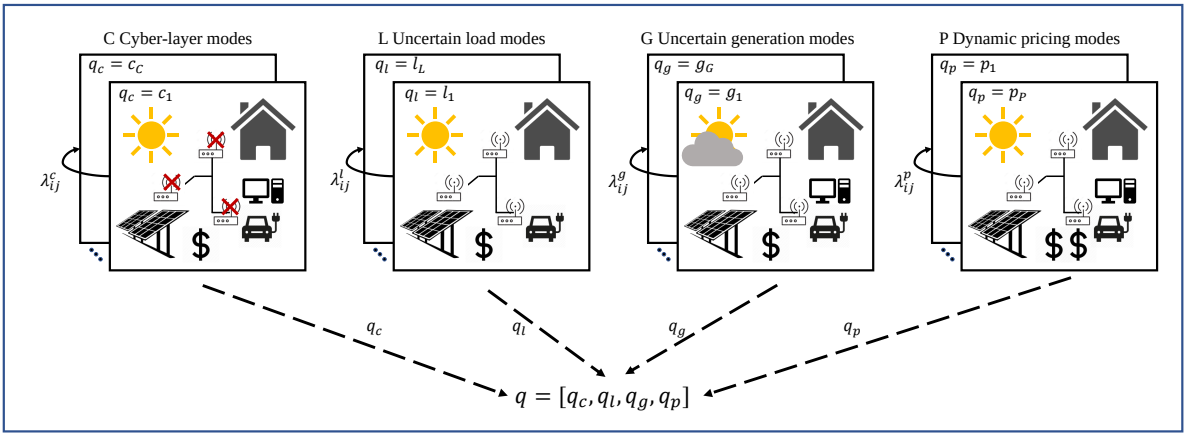


Figure 1.1: *A general SHS model for smart grids*

the network topology. Network topology indicates the connection of main distribution grid, load and storage devices, status of renewable energy generation, etc.

- Continuous behaviors: The physical components in a smart grid include generated power, node voltage, etc. They evolve continuously and the evolution depends on the current system modes (discrete states).

- Stochasticity: There are external and internal uncertainties that have to be taken into consideration. For example, the randomness of human behavior, the failure of the main grid, the influence of weather as well as measurement noise can all impact grid behavior. The above mentioned stochasticity can be modeled as random mode transitions or noise in continuous state knowledge.

Due to the existence of both discrete and continuous behavior as well as the uncertainties within a smart grid, SHS is an appropriate model to capture the above mentioned characteristics. In fact, SHS can not only model the uncertainties in loads, generations and other physical components, but also offer a general framework to model the cyber-infrastructure states and the economic strategies (dynamic pricing) in a smart grid. The general concept of modeling the smart grid as an SHS is illustrated in Figure 1.1.

This dissertation considers network topology errors in a smart grid as a motivating example. A conceptual small-scale smart grid model is shown in Figure 1.2. This toy model

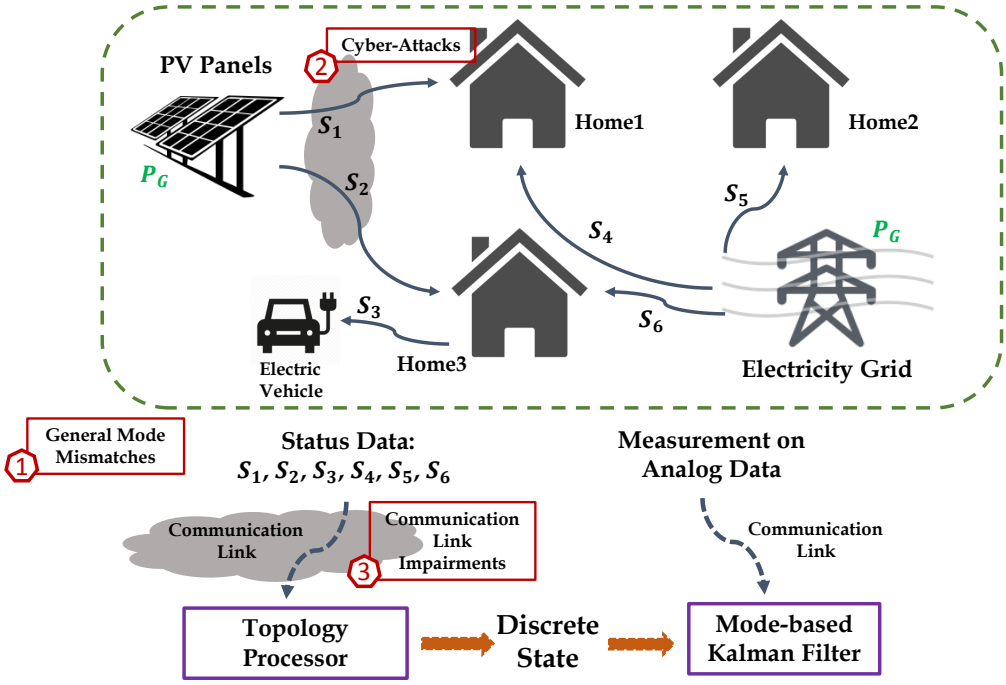


Figure 1.2: *Conceptual smart grid model*

includes a bank of PV panels, electricity grid, home loads and electric vehicles. To aid state estimation, there are two types of data collected in the smart grid:

1. Status data for switches, breakers, and communication links;

2. Analog data such as bus voltage, power flow, reactance, etc.

The continuous state can be estimated based on measurements of analog data and a mode-based Kalman filter is proven to be the optimal estimator if the discrete mode is known [20]. Meanwhile, the status data in a smart grid determines the network topology with each topology representing one discrete mode. The status data is sometimes reported by a human operator, obtained from sensor measurements, or estimated based on analog data. All three approaches are error prone due to human errors, missing data, communication errors or estimation errors. As a consequence, errors in status data result in a network topology (mode) error. Network topology errors in smart grids as discussed in [21] is a critical problem for any networked system. Since continuous state evolves differently for different modes,

errors in the network topology (mode) will impact the performance of mode-based Kalman filter for continuous state estimate. Therefore, in this research, we consider the situation wherein we have information on the discrete states but the information is inaccurate, also referred to as mode mismatch errors. As shown in Figure 1.2, there are different types of mode mismatch errors considered in this dissertation - (i) general mode mismatch; (ii) mode mismatch resulting from cyber-impairment which includes cyber-attacks and communication link failures. Specifically, for each scenario, we investigate the effect of mode mismatch on mode-based Kalman filter and study the fidelity of discrete state knowledge required for maintaining the continuous state estimate quality.

1.3 Motivation and Research Questions

As the preceding discussion suggests, state estimation in SHS is critical for both situational awareness and implementation of control actions. Due to the interaction between continuous states and discrete states, there exists challenges in SHS state estimation. Mode-based Kalman filters have been widely used in continuous dynamics estimation. It has been shown that a mode-based Kalman filter is an optimal estimator if the discrete states (modes) are known [20, 22]. If discrete states are not available, the optimal estimator is composed of a bank of Kalman filters with each filter corresponding to one discrete state. Several widely-used hybrid estimation strategies such as interacting multiple model (IMM) algorithm [9] and multiple model adaptive estimation (MMAE) algorithm [23] are established based on this idea. Related to the problem of state estimation in SHS, we seek to address a few fundamental research questions in this dissertation:

Question 1. *For an SDSHS (SHS with guard conditions), how can one develop a state estimation strategy using mode-based Kalman filter?*

Question 2. *For a general SHS model with mode mismatch errors modeled as independent and identically distributed (i.i.d.) Bernoulli random variables, how does mode mismatch impact the performance of a mode-based Kalman filter?*

Question 3. *For an MJLS with time correlated mode mismatch errors, can we establish*

algebraically solvable conditions under which the bias of mode-based Kalman filter resulting from mode mismatch errors is statistically convergent? Can this result be extended to an MJLS with arbitrary number of modes?

Question 4. *What if the mode mismatch errors are spatially correlated in an MJLS with arbitrary number of modes? Can we still derive algebraically solvable conditions such that the bias of mode-based Kalman filter converges?*

1.4 Contributions of This Dissertation

In this dissertation, we consider the case where we have information on the discrete state but the information is inaccurate. In this case, it is possible to implement a bank of Kalman filters for continuous state estimation. However, this approach suffers from exponentially increasing memory and computational complexity. On the other hand, we can treat the known discrete states as the true state and conduct the estimation via only one Kalman filter (i.e., a mode-based Kalman filter). This Kalman filter is optimal if there is no mode mismatch. A mode mismatch error will introduce a bias to the estimator with the error covariance remains bounded [24–28]. In this regard, the major contributions of this dissertation summarized below captures the foundation of this research.

- Question 1: State estimation design for SHS with guard conditions (SDSHS).
  - Propose a new state estimation strategy for SHS with quadratic guard conditions. Unlike the previous effort [29], only one Kalman filter is needed and the discrete state estimate is derived based on the estimated continuous state.
  - Derive the exact distribution for the guard condition instead of an approximation of the distribution (as in [29]). Based on this distribution, we derive a threshold for deciding whether a discrete transition occurs or not.
  - The proposed approach results in an extremely low error rate for the discrete state estimates even when transitions are frequent.

We discuss these contributions in Chapter 3 and in the following article:

[19] W. Zhang and B. Natarajan, "State estimation in Stochastic Hybrid Systems with Quadratic Guard Conditions," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 752-757, Sept 2016.

- Question 2: Performance analysis for generalized SHS with mode mismatch errors modeled as i.i.d. Bernoulli random variable.

  - Quantify the performance of a mode-based Kalman filter with mode mismatch errors and derive the bias dynamics resulting from mode mismatch errors.
  - Derive a computationally efficient sufficient condition for a special case of SHS with two discrete states. The approach involves solving a straightforward eigenvalue problem to derive the critical region on the mode probability.
  - For a generalized SHS with arbitrary number of modes, we propose the use of a transformed switched system to describe the bias dynamics. The convergence of the bias is then mapped to the stability of the transformed switched system.
  - Derive sufficient and necessary conditions for stability of the corresponding autonomous switched system and investigate the bounded input bounded output stability of the transformed switched system. Acquire a tolerant region on probability of mode mismatch errors that guarantees convergence of the bias dynamics.

  We discuss these contributions in Chapter 4 and in the following articles:

  [24] W. Zhang and B. Natarajan, "Quantifying the Bias Dynamics in a Mode-based Kalman Filter for Stochastic Hybrid Systems," in *2018 Annual American Control Conference (ACC)*, pp. 5849-5856, June 2018.

  [25] W. Zhang and B. Natarajan, "On the Convergence of Bias of a Mode-based Kalman Filter for Stochastic Hybrid Systems," *EURASIP Journal on Advances in Signal Processing (In Press)*, 2018.

- Question 3: Study the impact of time correlated mode mismatch errors on MJLS state estimation.

    - Quantifying the impact of time correlated (Markovian distributed) mode mismatch errors on a mode-based Kalman filter for MJLS state estimation.
    - Begin the analysis with an MJLS with two discrete states and derive sufficient and necessary conditions (based on the results from Schur stability of a matrix polytope) under which the bias dynamics are statistically convergent.
    - Model the mean of bias dynamics as an auxiliary linear system. The system matrix of this linear system is determined by a polytope of matrices with each vertex matrix related to the original MJLS system matrices.
    - For MJLS with arbitrary numbers of modes, by mapping the matrix polytope to an interval matrix, and by leveraging results in Schur stability analysis for an interval matrix, derive sufficient conditions on mode mismatch probabilities under which the bias resulting from mode mismatches is statistically convergent.

  These contributions are discussed in detail in Chapter 5 and in the following articles:

  [26] W. Zhang and B. Natarajan, "Impact of Time Correlated Mode Mismatch on Markov Jump Linear System State Estimation," *IEEE Control Systems Letters*, vol. 2, pp. 489-494, July 2018.

  [27] W. Zhang and B. Natarajan, "On the Performance of Kalman filter for Markov Jump Linear Systems with Mode Mismatch," Manuscript submitted to *IEEE Transaction on Automatic Control*, 2018.

- Question 4: Analyze bias dynamics in mode-based Kalman filter with spatially correlated mode mismatch errors.

    - Consider the case of correlated mode mismatches that can capture spatially correlated cyber-impairments (communication link failures and cyber-attacks) in practical applications.

- Derive sufficient conditions under which the bias resulting from mode mismatches is statistically convergent. The condition is related to mode mismatch probabilities and it provides guidance on the fidelity of discrete state information needed to sustain the quality of the Kalman filter estimate.
- For the first time, we are able to derive an algebraically solvable condition in terms of the mode mismatch probabilities that guarantees the statistical convergence of the bias.

These contributions are discussed in detail in Chapter 6 and in the following articles:

[28] W. Zhang and B. Natarajan, "Bias Analysis in Kalman Filter with Correlated Mode Mismatch Errors," *Signal Processing*, vol. 154, pp. 232-237, 2019.

1.5 Organization of This Dissertation

Chapter 2 presents a literature review on state estimation strategies for SHS and their performance. In Chapter 3, a state estimation strategy based on mode-based Kalman filter for SHS with quadratic guard conditions is proposed. Chapter 4 derives the bias dynamics resulting from mode mismatches modeled as i.i.d. Bernoulli random variables and studies its statistical convergence. Chapter 5 investigates the performance of a mode-based Kalman filter with time correlated mode mismatches. In Chapter 6, spatially correlated mode mismatch errors are considered and a sufficient condition such that bias dynamics remain bounded in an MJLS state estimation is derived. Concluding remarks and future research directions are discussed in Chapter 7.

Chapter 2

Literature Review

In this chapter, we review the prior literature on state estimation for stochastic hybrid systems (SHS). Specifically, we focus on efforts that relate to performance and bias analysis for hybrid estimation strategies.

2.1 State Estimation for SHS

The state space of an SHS is composed of both discrete states and continuous states and the state space reveals the current status of the system. In most practical applications, the continuous state itself may not be directly accessible and instead we observe a noisy measurement that is a function of the states. For this case, state estimation is critical since it benefits both situational awareness and implementation of control actions. The discrete state usually reflects the operational mode of the system and it might be known (directly available) or estimated. For these two different scenarios, the estimation strategies for continuous state are different due to the interaction between continuous states and discrete states. In the following, we present related research works considering the two different scenarios.

2.1.1 Scenario 1: Discrete states are directly available

For the case that discrete states (also referred to as modes) are known and continuous states evolve linearly, mode-based Kalman filter is an optimal estimator for continuous state [20, 22, 30]. Let us take Markov jump linear system (MJLS) [20] as an example. A discrete-time MJLS can be described as

$$\mathbf{x}_{k+1} = \mathbf{A}_{\delta_k}(k)\mathbf{x}_k + \mathbf{B}_{\delta_k}(k)\mathbf{w}_k$$

$$\mathbf{y}_k = \mathbf{C}_{\delta_k}(k)\mathbf{x}_k + \mathbf{v}_k$$

where $\mathbf{x}_k \in \mathbb{R}^n$ is the continuous state, $\delta_k \in \mathcal{Q}$ is the discrete mode which follows a Markov process and $\mathbf{y}_k \in \mathbb{R}^m$ is the measurement. $\mathbf{A}_{\delta_k}(k)$, $\mathbf{B}_{\delta_k}(k)$ and $\mathbf{C}_{\delta_k}(k)$ are matrices corresponding to mode δ_k. The mode-based Kalman filter is updated based on the measurement sequence and mode sequence up to time k, i.e., $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$, $\boldsymbol{\delta}_k^s = (\delta_1, \cdots, \delta_k)$ respectively. The mode-based Kalman filter algorithm is presented in Algorithm 1.

Algorithm 1 Mode-based Kalman filter

function ESTIMATION UPDATE($\boldsymbol{\mu}_0, \mathbf{M}_{0|0}, \mathbf{Q}, \mathbf{R}, \boldsymbol{\delta}_k^s, \mathbf{y}_k^s$)
  $\mathbf{x}_{0|0} = \boldsymbol{\mu}_0, \mathbf{M}_{0|0} = \boldsymbol{\Sigma}_0$
  $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$
  $\boldsymbol{\delta}_k^s = (\delta_1, \cdots, \delta_k)$
  for $i = 1 : k$ **do**
    $\mathbf{x}_{i|i-1} = \mathbf{A}_{\delta_i}\mathbf{x}_{i-1|i-1}$
    $\mathbf{M}_{i|i-1} = \mathbf{A}_{\delta_i}\mathbf{M}_{i-1|i-1}\mathbf{A}'_{\delta_i} + \mathbf{Q}$
    $\mathbf{K}_{\delta_i,i} = \mathbf{M}_{i|i-1}\mathbf{C}'_{\delta_i}(\mathbf{C}_{\delta_i}\mathbf{M}_{i|i-1}\mathbf{C}'_{\delta_i} + \mathbf{R})^{-1}$
    $\mathbf{x}_{i|i} = \mathbf{x}_{i|i-1} + \mathbf{K}_{i,\delta_i}(\mathbf{y}_i - \mathbf{C}_{\delta_i}\mathbf{x}_{i|i-1})$
    $\mathbf{M}_{i|i} = (\mathbf{I} - \mathbf{K}_{\delta_i,i}\mathbf{C}_{\delta_i})\mathbf{M}_{i|i-1}$
  end for
  return $\mathbf{x}_{k|k}$
end function

The estimated state $\mathbf{x}_{k|k}$ is a minimum mean square error (MMSE) estimate of the actual continuous state $\mathbf{x}_k$. MMSE here indicates zero-bias and minimum error covariance. A proof of this can be found in Chapter 5 of [20]. A similar example is for an MJLS with observation of continuous state and delayed measurement of discrete state and a mode-based Kalman

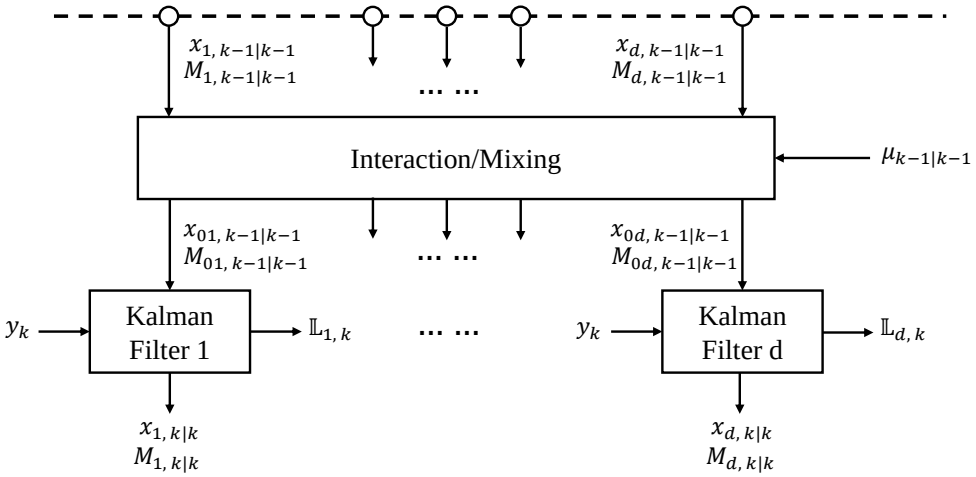
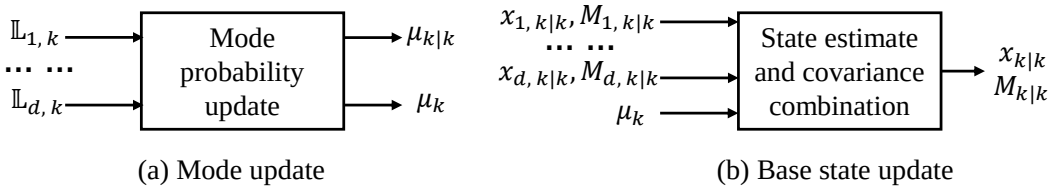


Figure 2.1: *Structure of the IMM algorithm*



(a) Mode update (b) Base state update

Figure 2.2: *Mode update and base state update*

filter can also be applied as a MMSE estimator as demonstrated in [31]. In [32], the authors expand their results to the case of delayed observations of both continuous and discrete states.

2.1.2 Scenario 2: Discrete states are not directly available

In most realistic applications, the continuous states and discrete states may not be directly accessible. For this situation, an optimal estimator is composed of a bank of Kalman filters with each filter corresponding to one discrete state. Several widely-used hybrid estimation strategies such as interacting multiple model (IMM) algorithm [33, 34] and multiple model adaptive estimation (MMAE) algorithm [23] are established based on this idea. The process

of the IMM algorithm is presented in Figure 2.1 and Figure 2.2. Specifically, $\mathbf{x}_{j,k|k}$, $\mathbf{M}_{j,k|k}$ are the state estimate and covariance of Kalman filter corresponding to mode j at time k, respectively. $\mathbf{x}_{0j,k|k}$, $\mathbf{M}_{0j,k|k}$ are the mixed condition of Kalman filter corresponding to mode j at time k. μ_k, $\mu_{k|k}$ are the mode probabilities and mixing probabilities at time k and $\mathbb{L}_{j,k}$ is the likelihood function of Kalman filter corresponding to mode j. This algorithm can be intuitively explained as follows.

1. It contains of a low gain filer (for the nearly uniform motion) and a high gain filter;
2. These filters interact (exchange information) with time-variant weights (the mixing probabilities);
3. The final estimate is a combination (weighted average) of each filter's estimate, with the weights being the mode probabilities;
4. The weights for interaction and combination are based on which model fits better the data (and other factors, such as the expected transition from one mode to another).

More recently, IMM algorithm has been extended to SDSHS and MJLS with time-variant transition rates of the Markov chain [17, 29, 35]. [17, 29] According to Figure 2.1, the IMM algorithm involves d (number of discrete states) Kalman filters. Similarly, the MMAE algorithm also requires a bank of Kalman filters and they therefore suffers from exponentially increasing memory and computational complexity [20, 36].

2.2 Performance Analysis for Kalman Filters

Mean square error (MSE) is one common measure for estimate quality. MSE measures the average squared difference between the estimated values and the estimation. Mathematically, the MSE of $\hat{\theta}$ that is an estimate of parameter θ is defined as

$$MSE(\hat{\theta}) = (bias(\hat{\theta}, \theta))^2 + Var(\hat{\theta}), \tag{2.1}$$

with

$$bias(\hat{\theta}, \theta) = \Big(\mathbb{E}(\hat{\theta}) - \theta\Big),$$
$$Var(\hat{\theta}) = \mathbb{E}\bigg[\Big(\hat{\theta} - \mathbb{E}(\hat{\theta})\Big)^2\bigg].$$

According to (2.1), MSE is defined both on the bias and variance of the estimator. Bias reflects how far off the average estimated value is from the true value and the variance reflects how widely spread the estimates are. Typically, an unbiased estimator with the smallest variance is the best unbiased estimator, also referred to as minimum variance unbiased estimator (MVUE). As the preceding discussion suggests, considering the entire space of estimation algorithms proposed for SHS, Kalman filter based algorithms dominate the area. In the following, we will discuss the performance of Kalman filter and review efforts that have been done to address this problem.

For non-hybrid linear system, the Kalman filter is the optimal filter (in MVUE sense) [37] if the model matches the real system and the system noise is white Gaussian distributed. However, if there exists missing measurements or intermittent observations in a dynamical system, then the error covariance can diverge and become unbounded. This type of system have been studied in [38–40]. [41] extends the analysis for the case of measurement loss in distributed system estimation with Kalman filter. A more sophisticated case for dynamical systems with random delays and packet dropouts (missing measurements) has been considered in [42–45]. All the above mentioned works follow a similar approach that is to derive a bound for the critical probability of missing observation that ensures the statistical convergence of error covariance. The bias is not taken into consideration in the above mentioned papers due to the fact that Kalman filter is unbiased in such situations. Another work [46] considers a different scenario where the model for Kalman filter is mismatched with the true system (mismodeling). They take a unique perspective with focusing on a linear system with possible failures and the failure then can be modeled as a deviation from the true model. The considered metric is the residual of the Kalman filters and it can be shown that mismodeling also introduces a bias to the Kalman filter. [46] derives mean and covariance of the residual

without an algebraic convergence analysis.

As stated in the previous section, several state estimation strategies have been proposed for SHS and Kalman filter is a critical component for most of the estimation strategies. In terms of performance analysis of hybrid estimation strategies, there are only limited prior efforts [47–49] with focus on the stability analysis of MMAE and IMM algorithm. [47] proposes to evaluation the performance of MMAE and IMM using steady-state mean residual for each mode. In [48], the authors propose an algorithm to analyze the performance of IMM algorithm. Specifically, their approach involves several steps to approximate the mean of mixing probability, means and covariance of Kalman filter residuals, mean of likelihood function and mean-squared errors of the state estimation. A more recent work [49] studies the lower and upper bounds on the error covariance of residual from a Kalman filter and derives a sufficient condition for exponential stability of the IMM algorithm. With restricted to IMM and MMAE algorithms, the prior works lack of exploration on the relationship between mode mismatch errors and performance of a general Kalman filter. Let us take an example of IMM algorithm as shown in Figure 2.1. For any hybrid system with d discrete modes, there is only one Kalman filter that corresponds to the true mode is unbiased among the bank of d Kalman filters. How sensitive is the convergence of bias in a mode-based Kalman filter to errors in discrete state knowledge? Is there a critical region within which the error dynamics in a mode-based Kalman filter will converge? These are the fundamental unanswered questions that this dissertation seeks to address.

Chapter 3

State Estimation in SHS with Mode-Based Kalman Filter

In the first research, we design a state estimation strategy for a special class of SHS in which the continuous state evolves linearly and the discrete state is governed by quadratic guard conditions. This work provides us insights on how discrete state and continuous state estimation are interacting and serves as a starting point of this dissertation.

3.1 Introduction

For any SHS, the evolution of continuous states depends on discrete states, whereas the converse does not always hold. State-dependent stochastic hybrid system (SDSHS) is a special subclass of SHS where the discrete state transitions are governed by guard conditions. In this model, a discrete state transition happens only when a certain deterministic guard condition is satisfied. SDSHS with guard condition has been shown to be an appropriate model for many applications such as air traffic control [17, 29] and fuel-transfer system of fighter aircraft [50]. State estimation in SDSHS presents some unique challenges as the continuous state and discrete state are interacting with each other. On the other hand, the discrete state estimation is benefit from the auxiliary information of continuous state

estimates. The objective of this research is to design a Kalman filter based hybrid estimation strategy for SDSHS with quadratic guard conditions.

Research work in state estimation for SHS has been ongoing for decades. One category of well-established SHS model is discrete-time Markov jump linear systems (MJLS) [20] that are applicable for systems that can be represented by a set of discrete-time linear systems with modal transitions given by a Markov chain. [18, 20, 51] consider estimation problems for MJLS. [18] proposes a simulation based algorithms called particle filters that also finds application for non-linear SHS later in [52, 53]. [51] develops interacting-multiple model (IMM) algorithm based on a bank of Kalman filter. Since the above mentioned papers have not considered SDSHS, discrete states estimation is independent with continuous estimation in nature. In the most recent, estimation problem for SDSHS gains research interest as it greatly expands the application area in air traffic control [5, 17, 29, 54] and smart grid [7]. State estimation in such systems have been considered in [17, 29, 55]. In [55], the authors consider a system where the discrete state transitions are described by a Markov chain but the transition rate is dependent on the continuous state. The authors extend the well-known IMM algorithm for MJLS to systems with variable transition rates. In[17], the authors consider a system with guard condition and they propose a bank of Kalman filter-based algorithm called state-dependent-transition hybrid estimation algorithm. However, the algorithm proposed in [17] is only suitable for linear guard conditions. In this case, the problem is simplified as the linear transformation property of Gaussian random variables holds. For SHS with quadratic guard conditions, [29] follows a similar approach as in [17] and approximates the guard condition via the Laurent series and Taylor series expansion that enables a simple form. However, the algorithms in [17, 29] require on-line calculation from a bank of Kalman filter and it suffers from computation complexity. Additionally, the simulation results suggest that discrete state estimation errors occur during every transition. Therefore, the applicability of [29] for many practical systems is limited.

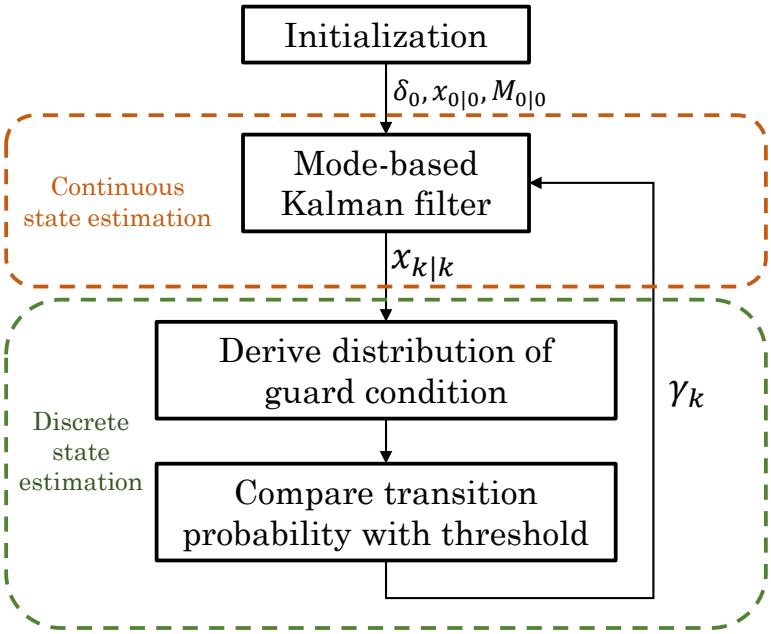


Figure 3.1: *Proposed estimation strategy*

3.2 Proposed Approach

Mathematically, the evolution of continuous dynamics in a linear SHS can be represented by:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}_{\delta_k}\mathbf{x}_{k-1} + \mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{C}_{\delta_k}\mathbf{x}_k + \mathbf{v}_k, \end{aligned} \tag{3.1}$$

where, $\mathbf{x}_k \in \mathbb{R}^n$ is the continuous state and $\mathbf{y}_k \in \mathbb{R}^m$ is the measurement. $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ are both independent Gaussian noise that capture model and measurement uncertainty, respectively. Without lose of generality, we define $\mathcal{Q} = \{q_1, \cdots, q_d\}$ as the discrete states set. The discrete state is $\delta_k = q_i \in \mathcal{Q}$ whose transitions are governed by a set of guard conditions $\mathcal{G} = \{\mathbb{G}(i,j) : i,j \in \mathcal{Q}\}$. $\mathbb{G}(i,j) = \{\mathbf{x} : \mathbb{g}_{i,j}(\mathbf{x}) \leq 0\}$ is a subspace of continuous state space in which the guard condition $\mathbb{g}_{i,j}(\mathbf{x}) \leq 0$ is satisfied. For quadratic guard condition, the function $\mathbb{g}_{i,j}(\mathbf{x}) = \mathbf{x}'\mathbf{\Psi}_{ij}\mathbf{x} - c_{ij}$ where c_{ij} is a scalar constant and $\mathbf{\Psi}_{ij}$

is an n-by-n symmetric positive definite (s.p.d) matrix for each pair of modes (q_i, q_j). If the current mode $\delta_k = q_i$, and continuous dynamics enter the subset $\mathbb{G}(i,j)$, then discrete state makes transition from $\delta_k = q_i$ to $\delta_{k+1} = q_j$. The proposed estimation strategy estimates the system states recursively as shown in Figure 3.1. The continuous state is estimated via a mode-based Kalman filter and the transition probability for discrete states is obtained by derive the distribution of guard condition.

Through the rest of the this dissertation, we assume the initial distribution for continuous state is $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ and there is a unique initial discrete state $\delta_0 = q_1$. We consider the case that discrete states are not available (cannot be directly observed), therefore the mode-based Kalman filter is processed based on estimated modes. Let the measurement sequence and estimated mode sequence up to time k be $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$, $\boldsymbol{\gamma}_k^s = (\gamma_1, \cdots, \gamma_k)$ respectively. The mode-based Kalman filter algorithm is presented in Algorithm 2. Note that $\mathbf{K}_{\gamma_i,i}$ is the Kalman gain related to mode γ_i. $\mathbf{x}_{k|k}$ is the estimate of $\mathbf{x}_k$ and we denote it as $\hat{\mathbf{x}}_k$.

Algorithm 2 Mode-based Kalman filter

function ESTIMATION UPDATE($\boldsymbol{\mu}_0, \mathbf{M}_{0|0}, \mathbf{Q}, \mathbf{R}, \boldsymbol{\gamma}_k^s, \mathbf{y}_k^s$)
  $\mathbf{x}_{0|0} = \boldsymbol{\mu}_0, \mathbf{M}_{0|0} = \boldsymbol{\Sigma}_0$
  $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$
  $\boldsymbol{\gamma}_k^s = (\gamma_1, \cdots, \gamma_k)$
  for $i = 1 : k$ **do**
    $\mathbf{x}_{i|i-1} = \mathbf{A}_{\gamma_i}\mathbf{x}_{i-1|i-1}$
    $\mathbf{M}_{i|i-1} = \mathbf{A}_{\gamma_i}\mathbf{M}_{i-1|i-1}\mathbf{A}'_{\gamma_i} + \mathbf{Q}$
    $\mathbf{K}_{\gamma_i,i} = \mathbf{M}_{i|i-1}\mathbf{C}'_{\gamma_i}(\mathbf{C}_{\gamma_i}\mathbf{M}_{i|i-1}\mathbf{C}'_{\gamma_i} + \mathbf{R})^{-1}$
    $\mathbf{x}_{i|i} = \mathbf{x}_{i|i-1} + \mathbf{K}_{i,\gamma_i}(\mathbf{y}_i - \mathbf{C}_{\gamma_i}\mathbf{x}_{i|i-1})$
    $\mathbf{M}_{i|i} = (\mathbf{I} - \mathbf{K}_{\gamma_i,i}\mathbf{C}_{\gamma_i})\mathbf{M}_{i|i-1}$
  end for
  return $\mathbf{x}_{k|k}$
end function

For discrete state estimation, the transition probability is approximated by:

$$\mathbb{P}(\gamma_{k+1} = q_j | \gamma_k = q_i, \hat{\mathbf{x}}_k) = \mathbb{P}(\mathbb{g}_{i,j}(\hat{\mathbf{x}}_k) \leq 0) = \mathbb{P}(\hat{\mathbf{x}}_k' \boldsymbol{\Psi}_{ij} \hat{\mathbf{x}}_k - c_{ij} \leq 0). \tag{3.2}$$

As the output of Kalman filter $\hat{\mathbf{x}}_k$ is a random variable, the transition probability depends

on the distribution of $\hat{\mathbf{x}}_k' \mathbf{\Psi}_{ij} \hat{\mathbf{x}}_k$ which is also a random variable.

Theorem 3.2.1. *The quadratic form random variable $\hat{\mathbf{x}}_k' \mathbf{\Psi}_{ij} \hat{\mathbf{x}}_k$ has a non-central $\chi^2(\tau, \upsilon_k^2)$ distribution, with degrees of freedom $\tau = n$ and non-centrality parameter $\upsilon_k^2 = \mathbf{b}_k' \mathbf{\Phi}_k \mathbf{b}_k$, where*

$$\mathbf{\Phi}_k = \mathbf{U}' \mathbf{M}_{k|k}^{\frac{1}{2}} \mathbf{\Psi}_{ij} \mathbf{M}_{k|k}^{\frac{1}{2}} \mathbf{U}, \quad \mathbf{b}_k = \mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} \hat{\mathbf{x}}_k \tag{3.3}$$

and $\mathbf{U}$ is a unitary matrix.

Proof. At time k, the continuous state estimation $\hat{\mathbf{x}}_k$ follows distribution of $\mathcal{N}(\hat{\boldsymbol{\mu}_k}, \hat{\mathbf{\Sigma}_k})$, where $\hat{\boldsymbol{\mu}_k} = \mathbf{x}_{k|k}$ and $\hat{\mathbf{\Sigma}_k} = \mathbf{M}_{k|k}$. Since the matrix $\mathbf{\Psi}_{ij}$ is positive definite, we can always find a unitary matrix $\mathbf{U}$ such that

$$\mathbf{U}' \mathbf{M}_{k|k}^{\frac{1}{2}} \mathbf{\Psi}_{ij} \mathbf{M}_{k|k}^{\frac{1}{2}} \mathbf{U} = diag[\phi_{1,k}, \cdots, \phi_{n,k}] = \mathbf{\Phi}_k. \tag{3.4}$$

Therefore,

$$\mathbf{\Psi}_{ij} = \big(\mathbf{M}_{k|k}^{-\frac{1}{2}} \mathbf{U}\big) \mathbf{\Phi}_k \big(\mathbf{M}_{k|k}^{-\frac{1}{2}} \mathbf{U}\big)' \tag{3.5}$$

Let $h := \hat{\mathbf{x}}_k' \mathbf{\Psi}_{ij} \hat{\mathbf{x}}_k$, so we can rewrite h as:

$$\begin{aligned} h &= \hat{\mathbf{x}}_k' \big(\mathbf{M}_{k|k}^{-\frac{1}{2}} \mathbf{U}\big) \mathbf{\Phi}_k \big(\mathbf{M}_{k|k}^{-\frac{1}{2}} \mathbf{U}\big)' \hat{\mathbf{x}}_k \\ &= \big(\mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} \hat{\mathbf{x}}_k\big)' \mathbf{\Phi}_k \big(\mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} \hat{\mathbf{x}}_k\big) \\ &= \mathbb{Tr}\left[\mathbf{\Phi}_k \big(\mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} \hat{\mathbf{x}}_k\big) \big(\mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} \hat{\mathbf{x}}_k\big)'\right] \\ &= \sum_{i=1}^{n} \phi_{i,k} \left[\big(\mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} \hat{\mathbf{x}}_k\big)^{[i]}\right]^2 \end{aligned} \tag{3.6}$$

Specifically,

$$\big(\mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} \hat{\mathbf{x}}_k\big)^{[i]} = \Big(\mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} (\hat{\mathbf{x}}_k - \hat{\boldsymbol{\mu}_k})\Big)^{[i]} + \big(\mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} \hat{\boldsymbol{\mu}_k}\big)^{[i]} \tag{3.7}$$

Let

$$\mathbf{f}_k^{[i]} = \Big(\mathbf{U}' \mathbf{M}_{k|k}^{-\frac{1}{2}} (\hat{\mathbf{x}}_k - \hat{\boldsymbol{\mu}_k})\Big)^{[i]}, \tag{3.8}$$

and

$$\mathbf{b}_k^{[i]} = \left(\mathbf{U}'\mathbf{M}_{k|k}^{\frac{-1}{2}}\hat{\boldsymbol{\mu}}_k\right)^{[i]}. \tag{3.9}$$

It is easy to show that $\mathbf{f}_k^{[i]} \sim \mathcal{N}(0,1)$ and $\mathbf{b}_k^{[i]}$ is a time-variant non-random variable [56]. So $\mathbf{f}_k = \begin{bmatrix}\mathbf{f}_k^{[1]} & \cdots & \mathbf{f}_k^{[n]}\end{bmatrix}'$ is a Gaussian random vector with mean $\mathbf{0}$ and covariance $\mathbf{I}$ while $\mathbf{b}_k = \begin{bmatrix}\mathbf{b}_k^{[1]} & \cdots & \mathbf{b}_k^{[n]}\end{bmatrix}'$ is a time-variant non-random vector. Then, we have

$$h = \hat{\mathbf{x}}_k'\boldsymbol{\Psi}_{ij}\hat{\mathbf{x}}_k = \sum_{i=1}^{b}\lambda_i\left(\mathbf{f}_k^{[i]} + \mathbf{b}_k^{[i]}\right)^2 = \sum_{i=1}^{n}\phi_{i,k}\left(\mathbf{s}_k^{[i]}\right)^2 = \mathbf{s}_k'\boldsymbol{\Phi}_k\mathbf{s}_k, \tag{3.10}$$

where, $\mathbf{s}_k \sim \mathcal{N}(\mathbf{b}_k, \mathbf{I})$. Therefore, h follows a non-central $\chi^2(\tau, \upsilon_k^2)$ distribution [57] with degrees of freedom $\tau = n$ and non-centrality parameter $\upsilon_k^2 = \mathbf{b}_k'\boldsymbol{\Phi}_k\mathbf{b}_k$. □

According to Theorem 3.2.1 and equation (3.2), the probability of state transitions can be approximated as:

$$\mathbb{P}(\gamma_{k+1} = q_j|\gamma_k = q_i, \hat{\mathbf{x}}_k) = \mathbb{P}(\mathbb{g}_{i,j}(\hat{\mathbf{x}}_k) \leq 0) = \mathbb{P}(\hat{\mathbf{x}}_k'\boldsymbol{\Psi}_{ij}\hat{\mathbf{x}}_k - c_{ij} \leq 0) = 1 - \mathfrak{Q}_{\frac{\tau}{2}}(\upsilon_k, \sqrt{c_{ij}}), \tag{3.11}$$

where, $\mathfrak{Q}_{\frac{\tau}{2}}$ is the Marcum-Q-function [58] defined as

$$\mathfrak{Q}_{\frac{\tau}{2}}(\upsilon_k, \sqrt{c_{ij}}) = \frac{1}{\upsilon_k^{\frac{\tau}{2}-1}}\int_{\sqrt{c_{ij}}}^{\infty} t^{\frac{\tau}{2}}e^{-\frac{t^2+\upsilon_k^2}{2}}I_{\frac{\tau}{2}-1}(\upsilon_k t)dt. \tag{3.12}$$

The estimation of discrete states depends on the calculated transition probability. Specifically, if there are more than two discrete states, the transition probability needs to be normalized such that

$$\sum_{j=1}^{d}\mathbb{P}(\gamma_{k+1} = q_j|\gamma_k = q_i, \hat{\mathbf{x}}_k) = 1.$$

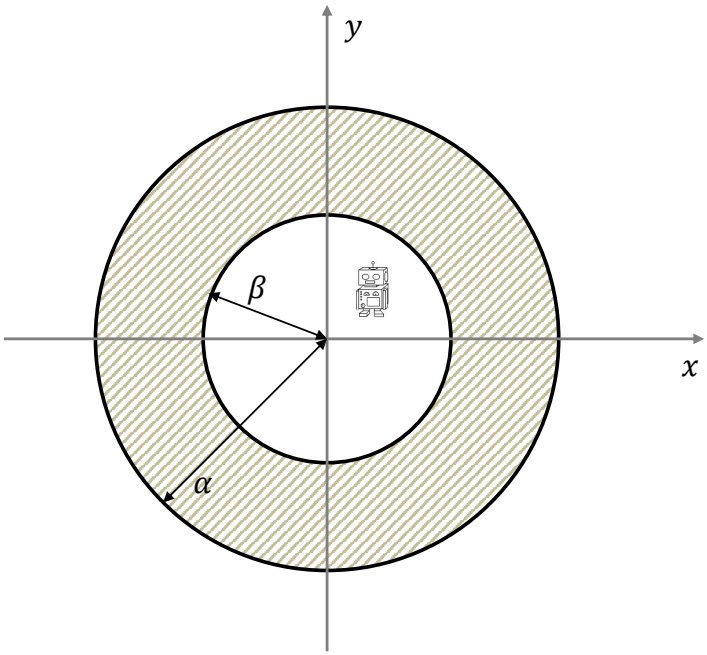


Figure 3.2: *Example of a toy robot moving*

The estimated mode is obtained via

$$\gamma_{k+1} = \underset{q_j}{\text{argmax}}\, \mathbb{P}(\gamma_{k+1} = q_j | \gamma_k = q_i, \hat{\mathbf{x}}_k).$$

3.3 Experimental Results

We consider a simplified autonomous robot motion example as a potential application of our work. As shown in Figure 3.2, the goal is to keep the robot moving in the shaded ring region. We assume a robot is moving in a coordinate plane with two modes of motion - moving away from the origin and moving towards the origin. The continuous state

$$\mathbf{x} = \begin{bmatrix} pos_x & pos_y \end{bmatrix}'$$

is a two-dimensional vector represents x-axis and y-axis of the robot position. The discrete state space is $\mathcal{Q} = \{q_1, q_2\}$ and the guard condition to be $\mathcal{g}_{1,2}(\mathbf{x}_k) = \mathbf{x}_k'\mathbf{I}\mathbf{x}_k - 25^2$ and $\mathcal{g}_{2,1}(\mathbf{x}_k) = \mathbf{x}_k'(-\mathbf{I})\mathbf{x}_k + 15^2$. In Figure 3.3, the small circles present the position of robot

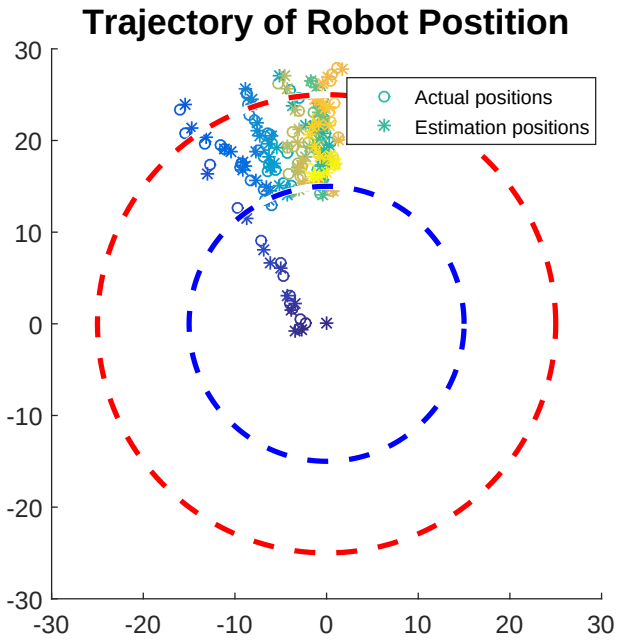


Figure 3.3: *Robot Motion Example: Moving trajectory of the robot*

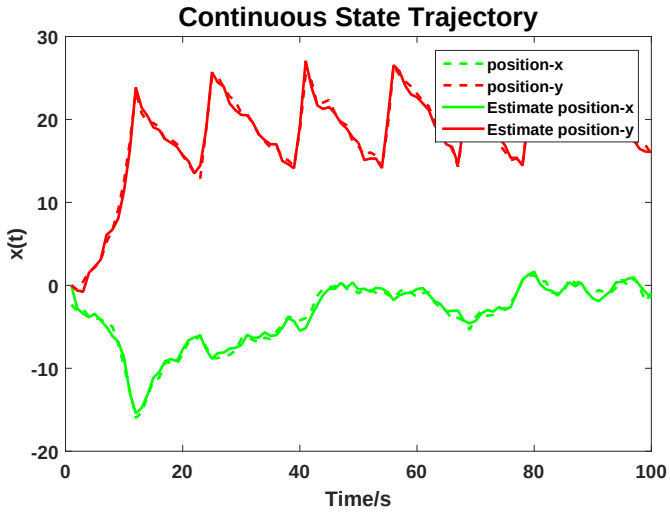


Figure 3.4: *Robot Motion Example: Continuous state and estimation*

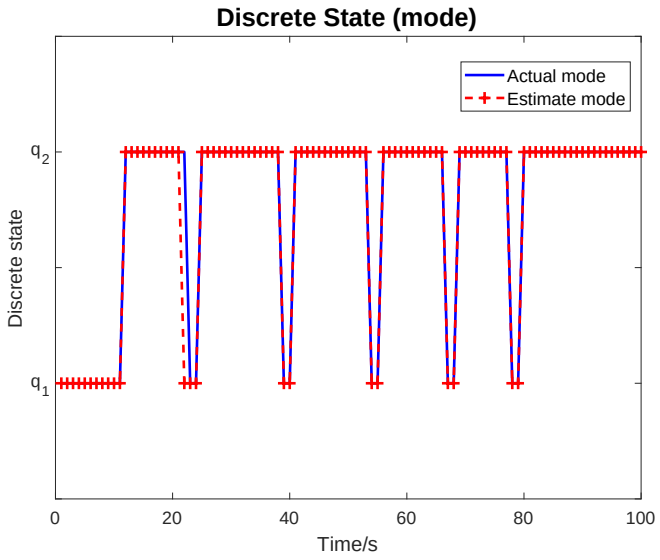


Figure 3.5: *Robot Motion Example: Discrete state and estimation*

during the 100s simulation and the stars are the estimation positions. Figure 3.4 presents the dynamics of $\mathbf{x}_k$ in dashed line and their estimation in solid line. Figure 3.4 indicates that the continuous state estimate tracks the actual state accurately. The solid line in Figure 3.5 shows the actual discrete states while the dashed line with '+' marks shows the estimated states. It should be noted that the number of discrete state transitions has a significant impact on the error performance of estimation strategies. From Figure 3.5, we observe that our algorithm maintains high accuracy even if the state transition happens very frequently. In this simulation, there is only 1 estimation error points over the 100 seconds. We can also define AEPT (average errors per transition) which captures average EPT over multiple runs. We perform 500 Monte-Carlo simulations and compare our results to the results in [29]. In terms of the AEPT, our algorithm has only 0.3691 AEPT while [29] has 1.505. We can conclude that our algorithm offers a significant improvement in AEPT. Additionally, our algorithm only requires one Kalman filter operation at each time instead of d Kalman filters. This advantage broadens applicability in practice.

3.4 Summary

In this chapter, we propose a state estimation strategy for stochastic hybrid systems with quadratic guard conditions. In the proposed approach, we use a Kalman filter for continuous state estimation. We prove that the distribution for guard condition is a non-central χ^2 distribution with parameters related to the dimensional and statistical properties of continuous states. The transition probabilities for discrete states are obtained via the cdf of the guard condition. We also discuss approaches to choose a proper threshold to estimate the discrete states based on the transition probabilities. We find a potential application of our result and simulate this robot motion system with superior error performance.

Chapter 4

Impact of Independent and Identically Distributed Mode Mismatch Errors

In the previous chapter, we proposed a hybrid state estimation strategy for a state-dependant stochastic hybrid system (SDSHS) which is composed of a mode-based Kalman filter and mode switches estimation. In this chapter, we first derive the bias dynamics in a mode-based Kalman filter that results from mode mismatches. By modeling the mode mismatch errors as independent and identically distributed (i.i.d.) Bernoulli random variables, the statistical convergence of the bias dynamics is also investigated in this chapter.

4.1 Introduction

State estimation in stochastic hybrid system (SHS) has attracted research interest for decades. Kalman filter based strategies dominate the area. For one category of SHS where both discrete and continuous states are observable and the discrete state transitions are independent with continuous state, mode-based Kalman filter can be applied as a minimum mean square error (MMSE) estimator [20, 22, 30]. In general SHS applications, discrete state may not be directly observable [17, 19, 20, 23, 29, 33, 59, 60]. In this case, the optimal estimator is obtained from a weighted sum of a bank of Kalman filters with each filter matched to a

possible mode. Therefore, it requires exponentially increasing memory and computing time. A couple of hybrid estimation algorithms have been proposed for Markov jump linear system (MJLS), such as interacting multiple model (IMM) algorithm [33] and multiple model adaptive estimation (MMAE) algorithm [23, 60]. [17, 29] extend the IMM algorithm to state dependent SHS. Note that all the above mentioned algorithms require on-line computation with a bank of Kalman filters and they suffer from high computational complexity. [19, 59] decrease the complexity by formulating the mode estimation as a problem of belief-state update and using only one Kalman filter corresponding to estimated mode for continuous state estimation. As the preceding discussion suggests, the Kalman filter plays an important role in most of the estimation algorithms for SHS. While there has been limited prior work that analyzes the bias (error) of Kalman filter in SHS estimation, multiple efforts have focused on Kalman filter error performance in non-hybrid scenarios [38–43, 45, 46, 61]. Due to the fact that the Kalman filter yields an unbiased estimator in a non-hybrid system framework, when analyzing error performance, only error covariance matrix is taken into consideration [38–43, 45, 61]. Specifically, [38–43, 45, 61] consider dynamical system with missing measures, intermittent observations, random delays and packet dropouts and they follow the similar approach to derive a bound for the critical probability of missing observation that ensures the convergence of error covariance. In [46], the authors consider an estimation problem where the model for the Kalman filter is mismatched with the true system. Unlike the previous mentioned papers, [46] studies the residual of Kalman filter and derives mean and covariance of the Kalman filter residual without analyzing its convergence behavior. In terms of estimation strategies for SHS, there are several prior efforts have been made [47–49, 62]. Their analysis is based on MMAE algorithm and the IMM approach. [47] first considers the problem of quantifying performance of a hybrid estimation algorithm and derive the condition for exponential convergence of the estimator in terms of detection delay and sojourn time [62]. In [48, 49], the authors extend their research on evaluating the stability of IMM algorithm and they focus on the mean and covariance of the Kalman filter residual. However, the existing research efforts have not explored the relationship between mode mismatch error and SHS estimation. It is not known as to how discrete state estimation error influences the

performance of a mode-based Kalman filter. How sensitive is the convergence of bias in a mode-based Kalman filter to errors in discrete state knowledge? Is there a critical region within which the error dynamics in a mode-based Kalman filter will converge? These are the fundamental unanswered questions that this chapter seeks to address.

In this chapter, we study the statistical convergence of the bias dynamics in a mode-based Kalman filter in the presence of i.i.d. mode mismatch errors. We first derive the dynamics of a bias that results from mode mismatch errors for a specific model of SHS with two discrete states. Additionally, the discrete state transitions are modeled via i.i.d. binary Bernoulli random variables. For this specific system, we derive sufficient conditions for statistical convergence of bias. Then, for a generalized SHS with arbitrary numbers of discrete states, we take a fresh perspective and propose to use a transformed switched system to describe the bias dynamics. The convergence of the bias is then mapped to the stability of the transformed switched system. Finally, the theoretical results are verified and validated using numerical examples and simulation of a smart grid with network topology errors. Theoretical and numerical results help us identify the fidelity required in discrete state knowledge in order to meet the performance requirements of continuous state estimates.

4.2 Preliminaries

4.2.1 System Model

We consider a discrete-time autonomous linear SHS. Mathematically, the continuous state $\mathbf{x}_k \in \mathbb{R}^n$ and measurement $\mathbf{y}_k \in \mathbb{R}^m$ are related via the following equations:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}_{\delta_k}\mathbf{x}_{k-1} + \mathbf{B}_{\delta_k}\mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{C}_{\delta_k}\mathbf{x}_k + \mathbf{v}_k, \end{aligned} \tag{4.1}$$

Here, δ_k represents the discrete state at time k, which is sometimes referred to as the mode. Without loss of generality, we define discrete space as $\mathcal{Q} = \{q_1, \cdots, q_d\}$. For each δ_k, the

corresponding $\mathbf{A}_{\delta_k}$ is an n-by-n matrix, $\mathbf{B}_{\delta_k}$ is an n-by-p matrix and $\mathbf{C}_{\delta_k}$ is a m-by-n matrix. Regarding the system model, we have the following assumptions:

1. $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ are mutually independent white Gaussian capturing model and measurement uncertainty, respectively.

2. The initial distribution of the continuous state follows a Gaussian distribution $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$. The discrete state has a unique initial mode $\delta_0 = q_1 \in \mathcal{Q}$.

3. For all $\delta_k \in \mathcal{Q}$, $(\mathbf{A}_{\delta_k}, \mathbf{B}_{\delta_k}\mathbf{Q}\mathbf{B}'_{\delta_k})$ is controllable and $(\mathbf{C}_{\delta_k}, \mathbf{A}_{\delta_k})$ is observable.

4. $\mathbf{A}_{q_i}, \forall q_i \in \mathcal{Q}$ are distinct; $\mathbf{C}_{q_i}\mathbf{A}_{q_i} = \mathbf{C}_{q_j}\mathbf{A}_{q_j}$ if and only if $q_i = q_j$. This assumption guarantees that all modes can be distinguished from each other.

4.2.2 Bias Dynamics in a Mode-based Kalman Filter

Mode-dependent Kalman filter is used to estimate the continuous state in many hybrid estimation algorithms [17, 19, 29, 32]. The mode-based Kalman filter algorithm for system in (4.1) is introduced in Algorithm 3. Note that we denote the estimated/known mode as γ_k and it also takes value in $\mathcal{Q}$.

Algorithm 3 Mode-based Kalman filter

function ESTIMATION UPDATE($\boldsymbol{\mu}_0, \mathbf{M}_{0|0}, \mathbf{Q}, \mathbf{R}, \boldsymbol{\gamma}_k^s, \mathbf{y}_k^s$)
  $\mathbf{x}_{0|0} = \boldsymbol{\mu}_0, \mathbf{M}_{0|0} = \boldsymbol{\Sigma}_0$
  $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$
  $\boldsymbol{\gamma}_k^s = (\gamma_1, \cdots, \gamma_k)$
  for $i = 1 : k$ **do**
    $\mathbf{x}_{i|i-1} = \mathbf{A}_{\gamma_i}\mathbf{x}_{i-1|i-1}$
    $\mathbf{M}_{i|i-1} = \mathbf{A}_{\gamma_i}\mathbf{M}_{i-1|i-1}\mathbf{A}'_{\gamma_i} + \mathbf{B}_{\gamma_i}\mathbf{Q}\mathbf{B}'_{\gamma_i}$
    $\mathbf{K}_{\gamma_i,i} = \mathbf{M}_{i|i-1}\mathbf{C}'_{\gamma_i}(\mathbf{C}_{\gamma_i}\mathbf{M}_{i|i-1}\mathbf{C}'_{\gamma_i} + \mathbf{R})^{-1}$
    $\mathbf{x}_{i|i} = \mathbf{x}_{i|i-1} + \mathbf{K}_{i,\gamma_i}(\mathbf{y}_i - \mathbf{C}_{\gamma_i}\mathbf{x}_{i|i-1})$
    $\mathbf{M}_{i|i} = (\mathbf{I} - \mathbf{K}_{\gamma_i,i}\mathbf{C}_{\gamma_i})\mathbf{M}_{i|i-1}$
  end for
  return $\mathbf{x}_{k|k}$
end function

Here, $\mathbf{K}_{\gamma_i,i}$ is the Kalman gain related to mode γ_i. $\mathbf{x}_{k|k}$ is the estimate of $\mathbf{x}_k$ and we denote it as $\hat{\mathbf{x}}_k$. If the estimator has full knowledge of the actual mode, the mode-based

Kalman filter is a minimum mean square error estimator. This situation is the same as a typical Kalman filter used in a non-hybrid system. However, if $\gamma_k \neq \delta_k$, then $\hat{\mathbf{x}}_k$ is a biased estimate of $\mathbf{x}_k$. We define the estimation error (bias) at time k as

$$\mathbf{e}_k = \mathbb{E}(\hat{\mathbf{x}}_k) - \mathbb{E}(\mathbf{x}_k). \tag{4.2}$$

It should be pointed out that we define the bias to be the difference between mean of estimate and the true state. In general, bias captures the difference between estimate and actual value. However, both $\hat{\mathbf{x}}_k$ and $\mathbf{x}_k$ are random variables that result in the bias being a random variable. Therefore, we capture the difference between $\hat{\mathbf{x}}_k$ and $\mathbf{x}_k$ in a mean sense via $\mathbf{e}_k$. This metric is similar to those considered in [46, 47]. In the following, we first define the unbiased property of Kalman filter for completeness.

Definition 4.2.1. *A mode-based Kalman filter is* unbiased *if and only if* $\forall k$, $\mathbb{E}(\mathbf{x}_k) = \mathbb{E}(\hat{\mathbf{x}}_k)$, *i.e.,* $\forall k$, $\mathbf{e}_k = \mathbf{0}$.

Lemma 4.2.1. *A mode-based Kalman filter is an unbiased estimator if* $\forall k \geq 0$, $\gamma_k = \delta_k$.

With the above definition and lemma, we prove the following corollary that if there exists mode mismatch errors, then the mode-based Kalman filter is biased.

Corollary 4.2.1.1. *A mode-based Kalman filter is biased if* $\exists k > 0$ *such that* $\gamma_k \neq \delta_k$.

Proof. We want to show that if $\exists \bar{k} > 0$ such that $\gamma_{\bar{k}} \neq \delta_{\bar{k}}$, then $\mathbf{e}_k \neq \mathbf{0}$ for some k. Let i denote the estimated mode and t denote the actual mode.

For $k < \bar{k}$ (i.e., time instants before mode mismatch occurs): Since $\mathsf{i} = \mathsf{t}$, as a result of Lemma 4.2.1, $\mathbb{E}(\mathbf{x}_k) = \mathbb{E}(\hat{\mathbf{x}}_k)$.

For $k = \bar{k}$ (i.e., time instant that mode mismatch occurs):

$$\begin{aligned}
\mathbb{E}(\hat{\mathbf{x}}_{\bar{k}}) &= \mathbf{A}_{\mathsf{i}}\mathbb{E}(\hat{\mathbf{x}}_{\bar{k}-1}) + \mathbf{K}_{\mathsf{i},\bar{k}}[\mathbb{E}(\mathbf{y}_{\bar{k}}) - \mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}}\mathbb{E}(\hat{\mathbf{x}}_{\bar{k}-1})] \\
&= (\mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},\bar{k}}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}})\mathbb{E}(\hat{\mathbf{x}}_{\bar{k}-1}) + \mathbf{K}_{\mathsf{i},\bar{k}}\mathbf{C}_{\mathsf{t}}\mathbb{E}(\mathbf{x}_{\bar{k}}) \\
&= (\mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},\bar{k}}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}})\mathbb{E}(\hat{\mathbf{x}}_{\bar{k}-1}) + \mathbf{K}_{\mathsf{i},\bar{k}}\mathbf{C}_{\mathsf{t}}\mathbf{A}_{\mathsf{t}}\mathbb{E}(\mathbf{x}_{\bar{k}-1})
\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbf{e}_{\bar{k}} &= \mathbb{E}(\hat{\mathbf{x}}_{\bar{k}}) - \mathbb{E}(\mathbf{x}_{\bar{k}}) = \mathbb{E}(\hat{\mathbf{x}}_{\bar{k}}) - \mathbf{A}_{\mathsf{t}}\mathbb{E}(\mathbf{x}_{\bar{k}-1}) \\ &= (\mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},\bar{k}}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}})\mathbb{E}(\hat{\mathbf{x}}_{\bar{k}-1}) + (\mathbf{K}_{\mathsf{i},\bar{k}}\mathbf{C}_{\mathsf{t}}\mathbf{A}_{\mathsf{t}} - \mathbf{A}_{\mathsf{t}})\mathbb{E}(\mathbf{x}_{\bar{k}-1}) \\ &\overset{(a)}{=} [\mathbf{A}_{\mathsf{i}} - \mathbf{A}_{\mathsf{t}} + \mathbf{K}_{\mathsf{i},\bar{k}}(\mathbf{C}_{\mathsf{t}}\mathbf{A}_{\mathsf{t}} - \mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}})]\mathbb{E}(\mathbf{x}_{\bar{k}-1})\end{aligned}$$

(a) is due to the fact that $\mathbb{E}(\mathbf{x}_k) = \mathbb{E}(\hat{\mathbf{x}}_k)$ for $k < \bar{k}$. When $k = \bar{k}$, we get $\mathsf{i} \neq \mathsf{t}$, $\mathbf{A}_{\mathsf{i}} - \mathbf{A}_{\mathsf{t}} \neq 0$ and $\mathbf{C}_{\mathsf{t}}\mathbf{A}_{\mathsf{t}} - \mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}} \neq 0$ based on assumptions in system model, which leads to $\mathbf{e}_{\bar{k}} \neq 0$. Therefore, the mode-based Kalman filter is biased. □

Remark 4.2.1. *If $\mathbf{A}_{q_i} = \mathbf{A}_{q_j}$ or $\mathbf{C}_{q_i}\mathbf{A}_{q_i} = \mathbf{C}_{q_j}\mathbf{A}_{q_j}$ for some q_i and q_j, the mode-based Kalman filter cannot discern that a mode mismatch has occurred. The result of Corollary 4.2.1.1 also excludes a trivial case that $\mathbb{E}(\mathbf{x}_{\bar{k}-1})$ lies in the null-space of the matrix $\mathbf{A}_{\mathsf{i}} - \mathbf{A}_{\mathsf{t}} + \mathbf{K}_{\mathsf{i},\bar{k}}(\mathbf{C}_{\mathsf{t}}\mathbf{A}_{\mathsf{t}} - \mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}})$. In this case, the corresponding contribution to bias $\mathbf{e}_k$ is 0.*

For the sake of compactness in notation, we follow the notation in proof of Corollary 4.2.1.1 and introduce t and i to denote actual mode and estimated mode at time k in this chapter. That is, $\mathsf{t} = \delta_k \in \mathcal{Q}$ and $\mathsf{i} = \gamma_k \in \mathcal{Q}$. It needs to be noted that t and i are indeed time-variant random variables. From the discussion in the proof of Corollary 4.2.1.1, we can write

$$\begin{aligned}\mathbf{e}_k =& (\mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}})\mathbb{E}(\hat{\mathbf{x}}_{k-1}) + (\mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{t}}\mathbf{A}_{\mathsf{t}} - \mathbf{A}_{\mathsf{t}})\mathbb{E}(\mathbf{x}_{k-1}) \\ =& (\mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}})\mathbf{e}_{k-1} + (\mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}} + \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{t}}\mathbf{A}_{\mathsf{t}} - \mathbf{A}_{\mathsf{t}})\mathbb{E}(\mathbf{x}_{k-1}) \qquad (4.3)\end{aligned}$$

Lemma 4.2.2. *If $\exists N, s.t. \forall k > N, \mathsf{i} = \mathsf{t}$, then the error of a mode-based Kalman filter will converge, i.e., $\lim_{k\to\infty} \mathbf{e}_k = 0$.*

Proof. $\forall k > N, \mathsf{i} = \mathsf{t}$, so $\mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}} + \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{t}}\mathbf{A}_{\mathsf{t}} - \mathbf{A}_{\mathsf{t}} = 0$. Then $\mathbf{e}_k = (\mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}})\mathbf{e}_{k-1} = \mathbf{\Lambda}_{\mathsf{i},k}\mathbf{e}_{k-1}$ where,

$$\mathbf{\Lambda}_{\mathsf{i},k} = \mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}} = (\mathbf{I} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}})\mathbf{A}_{\mathsf{i}}.$$

In order to prove $\lim_{k\to\infty} \mathbf{e}_k = 0$, we need to show that $\rho(\mathbf{\Lambda}_{\mathsf{i},k}) < 1$.

For any Kalman filter, the observer gain corresponding to mode i is defined as $\mathbf{L}_{\mathsf{i},k} = \mathbf{A}_{\mathsf{i}}\mathbf{M}_{k|k-1}\mathbf{C}_{\mathsf{i}}'(\mathbf{C}_{\mathsf{i}}\mathbf{M}_{k|k-1}\mathbf{C}_{\mathsf{i}}' + \mathbf{R})^{-1}$. Given that $(\mathbf{A}_{\mathsf{i}}, \mathbf{B}_{\mathsf{i}}\mathbf{Q}\mathbf{B}_{\mathsf{i}}')$ is controllable and $(\mathbf{C}_{\mathsf{i}}, \mathbf{A}_{\mathsf{i}})$ is observable for all $\mathsf{i} \in \mathcal{Q}$, the closed-loop dynamics $\mathbf{A}_{\mathsf{i}} - \mathbf{L}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}$ is stable. Then

$$\mathbf{A}_{\mathsf{i}} - \mathbf{L}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}} = \mathbf{A}_{\mathsf{i}} - \mathbf{A}_{\mathsf{i}}\mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}} = \mathbf{A}_{\mathsf{i}}(\mathbf{I} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}).$$

From commutativity property of spectral radius,

$$\rho(\mathbf{A}_{\mathsf{i}} - \mathbf{L}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}) = \rho(\mathbf{\Lambda}_{\mathsf{i},k}) < 1.$$

□

4.2.3 Error Covariance

Mean squared error (MSE) is an important metric to analyze performance of any estimator. MSE of $\hat{\theta}$ that is an estimate of parameter θ is defined as $MSE(\hat{\theta}) = (bias(\hat{\theta}, \theta))^2 + Var(\hat{\theta})$. Here, $Var(\hat{\theta})$ is the error variance. For a typical Kalman filter, since it is unbiased, prior research efforts primarily focus on behavior of the error covariance. For any Kalman filter, estimation error covariance is recursively updated as:

$$\mathbf{M}_{k|k} = \mathbf{M}_{k|k-1} - \mathbf{M}_{k|k-1}\mathbf{C}_{\mathsf{i}}'(\mathbf{C}_{\mathsf{i}}\mathbf{M}_{k|k-1}\mathbf{C}_{\mathsf{i}}' + \mathbf{R})^{-1}\mathbf{C}_{\mathsf{i}}\mathbf{M}_{k|k-1}. \tag{4.4}$$

Equation (4.4) shows that the error covariance matrix is updated depending on the current estimated mode i. The assumptions that $(\mathbf{A}_{\mathsf{i}}, \mathbf{B}_{\mathsf{i}}\mathbf{Q}\mathbf{B}_{\mathsf{i}}')$ is controllable and $(\mathbf{C}_{\mathsf{i}}, \mathbf{A}_{\mathsf{i}})$ is observable guarantee existence of a steady-state error covariance for each mode. As discussed in [63], given $\mathbf{Q}$ and $\mathbf{R}$ are constant, the error covariance $\mathbf{M}_{k|k}$ and Kalman gain will stabilize quickly. Disregarding whether mode i reflects the true state t or not, the updating process (4.4) remains bounded. This situation is different from the analysis of Kalman filter with intermittent measurement problem (see for example [64]) or coupled Riccati equations for

switched system [65] as the error covariance matrix evolves corresponding to a complete Riccati equation defined by mode i. Since the error covariance matrix is bounded, we focus on the bias of the mode-based Kalman filter in the rest of the chapter.

4.3 SHS with Bernoulli Distributed Discrete States

In this section, we focus on the specific class of SHS in which the discrete transitions are modeled via i.i.d. Bernoulli random variables and the discrete space contains only two modes, i.e., $\mathcal{Q} = \{q_1, q_2\}$. We derive conditions under which the bias dynamics is statistically convergent .

4.3.1 Problem Formulation

Irrespective of the methods used to estimate the discrete states, mismatch of discrete states is a typical problem in SHS estimation. Similar to the pioneering work [38–42, 46, 61] in missing observation problem for non-hybrid system, we use i.i.d. binary Bernoulli random variables υ_k to describe the mode mismatch, i.e.,

$$\upsilon_k = \begin{cases} 1 & \text{with probability } \lambda; \\ 0 & \text{with probability } 1-\lambda. \end{cases}$$

Here, $\upsilon_k = 1$ indicates the estimated mode is inconsistent with actual mode while $\upsilon_k = 0$ indicates they are the same. The extension to more general models such as correlated mode mismatch error process will be considered in the following chapters. As discussed, the update of $\mathbf{e}_k$ is described via equation (4.3). Following the notation $\mathbf{\Lambda}_{\mathsf{i},k} = \mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}}$ in the proof of Lemma 4.2.2, and defining $\mathbf{\Gamma}_{\mathsf{i},\mathsf{t},k} = \mathbf{A}_{\mathsf{t}} - \mathbf{K}_{\mathsf{i},k}\mathbf{C}_{\mathsf{t}}\mathbf{A}_{\mathsf{t}}$, when $\mathsf{t} = \mathsf{i}$, $\mathbf{\Gamma}_{\mathsf{i},\mathsf{t},k} = \mathbf{\Lambda}_{\mathsf{i},k}$. We can rewrite (4.3) as:

$$\mathbf{e}_k = \mathbf{\Lambda}_{\mathsf{i},k}\mathbf{e}_{k-1} + (\mathbf{\Lambda}_{\mathsf{i},k} - \mathbf{\Gamma}_{\mathsf{i},\mathsf{t},k})\mathbb{E}(\mathbf{x}_{k-1}). \tag{4.5}$$

After introducing υ_k, we have:

$$\text{for } \upsilon_k = 1,\ \mathbf{e}_k = \mathbf{\Lambda}_{\mathsf{i},k}\mathbf{e}_{k-1} + (\mathbf{\Lambda}_{\mathsf{i},k} - \mathbf{\Gamma}_{\mathsf{i},\mathsf{t},k})\mathbb{E}(\mathbf{x}_{k-1});$$

$$\text{for } \upsilon_k = 0,\ \mathbf{e}_k = \mathbf{\Lambda}_{\mathsf{i},k}\mathbf{e}_{k-1}.$$

So we rewrite equation (4.5) as:

$$\mathbf{e}_k = \mathbf{\Lambda}_{\mathsf{i},k}\mathbf{e}_{k-1} + \upsilon_k(\mathbf{\Lambda}_{\mathsf{i},k} - \mathbf{\Gamma}_{\mathsf{i},\mathsf{t},k})\mathbb{E}(\mathbf{x}_{k-1}). \tag{4.6}$$

Denote the steady Kalman gain $\mathbf{K}_{\mathsf{i},\infty}$ for mode i as $\mathbf{K}_\mathsf{i}$, along with the relative matrices $\mathbf{\Lambda}_\mathsf{i} = \mathbf{A}_\mathsf{i} - \mathbf{K}_\mathsf{i}\mathbf{C}_\mathsf{i}\mathbf{A}_\mathsf{i}$ and $\mathbf{\Gamma}_{\mathsf{i},\mathsf{t}} = \mathbf{A}_\mathsf{t} - \mathbf{K}_\mathsf{i}\mathbf{C}_\mathsf{t}\mathbf{A}_\mathsf{t}$. Since we assume that $\mathbf{Q}$ and $\mathbf{R}$ are same constant for all modes $\mathsf{i} \in \mathcal{Q}$, the Kalman gain $\mathbf{K}_{\mathsf{i},k}$ will converge to the corresponding steady Kalman gain $\mathbf{K}_\mathsf{i}$ quickly [63]. The steady Kalman gain can be computed before any observation is made and it has been used in sub-optimal control problem for decades. As it involves an off-line calculation and can reduce complexity of the estimation process, we use $\mathbf{K}_\mathsf{i}$ to approximate $\mathbf{K}_{\mathsf{i},k}$ and simplify equation (4.6) as:

$$\mathbf{e}_k = \mathbf{\Lambda}_\mathsf{i}\mathbf{e}_{k-1} + \upsilon_k(\mathbf{\Lambda}_\mathsf{i} - \mathbf{\Gamma}_{\mathsf{i},\mathsf{t}})\mathbb{E}(\mathbf{x}_{k-1}). \tag{4.7}$$

It is important to note that even though we remove subscript k from $\mathbf{\Lambda}_\mathsf{i}$ and $\mathbf{\Gamma}_{\mathsf{i},\mathsf{t}}$, they are still time-variant because t and i change with k. Therefore, t and i in (4.7) should be interpreted as the actual mode at time k and estimated mode i at time k, respectively.

4.3.2 Main Results

From equation (4.7), $\{\mathbf{e}_k\}_{k=0}^{\infty}$ is a stochastic process for a given initial value $\mathbf{e}_0$. Note that $\mathbf{e}_k$ is bounded with probability 1 if and only if $\mathbb{E}(\mathbf{e}_k)$ is bounded [39]. Therefore, we consider convergence in mean, i.e., $\lim_{k\to\infty} \mathbb{E}(\mathbf{e}_k) < \infty$. Here, $\mathbb{E}(\mathbf{e}_k)$ is a n-dimensional vector and $\mathbb{E}(\mathbf{e}_k) < \infty$ means each element is finite. From tower rule, we know that $\mathbb{E}(\mathbf{e}_k) = \mathbb{E}(\mathbb{E}(\mathbf{e}_k|\mathbf{e}_{k-1}))$, where

the outer expectation is taken over $\mathbf{e}_{k-1}$ while the inner expectation is taken over the random variable υ_k.

$$\mathbb{E}(\mathbf{e}_k) = \mathbf{\Lambda}_{\mathsf{i}}\mathbb{E}(\mathbf{e}_{k-1}) + \lambda(\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}})\mathbb{E}(\mathbf{x}_{k-1}) \tag{4.8}$$

Our goal is to determine the conditions under which (4.8) is bounded. To help in the analysis, we first define an auxiliary iteration function $\mathbb{f}_\lambda(\mathbf{h}, \mathbf{l}) = \mathbf{S}_1\mathbf{h} + \lambda\mathbf{S}_2\mathbf{l}$. We have:

$$\mathbb{E}(\mathbf{e}_k) = \mathbb{E}\Big[\mathbb{f}_\lambda\big(\mathbf{e}_{k-1}, \mathbb{E}(\mathbf{x}_{k-1})\big)\Big] = \mathbb{f}_\lambda\big(\mathbb{E}(\mathbf{e}_{k-1}), \mathbb{E}(\mathbf{x}_{k-1})\big),$$

with $\mathbf{S}_1 = \mathbf{\Lambda}_{\mathsf{i}}$ and $\mathbf{S}_2 = \mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}}$.

For the vector $\mathbf{e}_k$ with finite dimensions, $\lim_{k\to\infty} \|\mathbb{E}(\mathbf{e}_k)\| < \infty$ gives the sufficient and necessary condition for $\lim_{k\to\infty} \mathbb{E}(\mathbf{e}_k) < \infty$. So if we can prove that $\|\mathbb{f}_\lambda(\mathbb{E}(\mathbf{e}_{k-1}), \mathbb{E}(\mathbf{x}_{k-1}))\|$ is bounded, then $\lim_{k\to\infty} \|\mathbb{E}(\mathbf{e}_k)\| < \infty$ will hold. In the following, we will focus on the convergence property of $\big\|\mathbb{f}_\lambda\big(\mathbb{E}(\mathbf{e}_{k-1}), \mathbb{E}(\mathbf{x}_{k-1})\big)\big\|$.

Theorem 4.3.1. *The iteration function $\|\mathbf{h}\| = \|\mathbb{f}_\lambda(\mathbf{h}, \mathbf{l})\|$ is monotonically increasing w.r.t. λ when the iteration starts from $\mathbf{h} = 0$.*

Proof. Let $0 \le \lambda_1 < \lambda_2 \le 1$. To prove the iteration function is monotonically increasing w.r.t. λ, we need to prove for the same initial condition $\mathbf{h}_0$ and $\mathbf{l}_0$, after $k \in \mathbb{N}^+$ iterations, $\|\mathbf{h}_{\lambda_1,k}\| = \big\|\mathbb{f}^k_{\lambda_1}(\mathbf{h}_0, \mathbf{l}_0)\big\| < \|\mathbf{h}_{\lambda_2,k}\| = \big\|\mathbb{f}^k_{\lambda_2}(\mathbf{h}_0, \mathbf{l}_0)\big\|$. Here,

$$\mathbb{f}^k_\lambda(\mathbf{h}_0, \mathbf{l}_0) = \underbrace{(\mathbb{f}_\lambda \circ \mathbb{f}_\lambda \circ ... \circ \mathbb{f}_\lambda)}_{k \text{ times}}(\mathbf{h}_0, \mathbf{l}_0),$$

and $\circ$ is function composition defined for any function $\mathbb{f}$ as $(\mathbb{f} \circ \mathbb{f})(x) = \mathbb{f}(\mathbb{f}(x))$. The update process is:

$$\mathbf{h}_1 = \mathbf{S}_1\mathbf{h}_0 + \lambda\mathbf{S}_2\mathbf{l}_0;$$

$$\mathbf{h}_2 = \mathbf{S}_1\mathbf{h}_1 + \lambda\mathbf{S}_2\mathbf{l}_1 = \mathbf{S}_1^2\mathbf{h}_0 + \lambda(\mathbf{S}_1\mathbf{S}_2\mathbf{l}_0 + \mathbf{S}_2\mathbf{l}_1);$$

$$\vdots$$

$$\mathbf{h}_k = \mathbf{S}_1^k \mathbf{h}_0 + \lambda \sum_{i=0}^{k-1} \mathbf{S}_1^{k-1-i} \mathbf{S}_2 \mathbf{l}_i.$$

For $\mathbf{h}_0 = 0$, $\mathbf{h}_k = \lambda \sum_{i=0}^{k-1} \mathbf{S}_1^{k-1-i} \mathbf{S}_2 \mathbf{l}_i$. Then $\|\mathbf{h}_k\| = |\lambda| \left\| \sum_{i=0}^{k-1} \mathbf{S}_1^{k-1-i} \mathbf{S}_2 \mathbf{l}_i \right\|$. So

$$\lambda_1 \left\| \sum_{i=0}^{k-1} \mathbf{S}_1^{k-1-i} \mathbf{S}_2 \mathbf{l}_i \right\| < \lambda_2 \left\| \sum_{i=0}^{k-1} \mathbf{S}_1^{k-1-i} \mathbf{S}_2 \mathbf{l}_i \right\|$$

$$\|\mathbf{h}_{\lambda_1,k}\| < \|\mathbf{h}_{\lambda_2,k}\|.$$

□

In our case, $\mathbf{e}_0 = \mathbb{E}(\mathbf{x}_0) - \mathbb{E}(\hat{\mathbf{x}}_0) = 0$ which guarantees that the initial $\mathbb{f}_\lambda(\mathbb{E}(\mathbf{e}_{k-1}), \mathbb{E}(\mathbf{x}_{k-1}))$ is 0. Therefore, Theorem 4.3.1 is applicable to the average error dynamics in (4.8). As defined before, λ represents the probability that estimation mode is inconsistent with actual mode, and Theorem 4.3.1 indicates that as λ increases, $\|\mathbb{E}(\mathbf{e}_k)\|$ will increase at each step update. Additionally, Theorem 4.3.1 also indicates that the increase of $\|\mathbb{E}(\mathbf{e}_k)\|$ is linearly related to λ.

Now, as we have uncovered the relationship between λ and the mean error dynamics, in the following, we will establish conditions for the error dynamics to convergence in mean. We first introduce two definitions of "the largest possible spectral radius" and "Lypunov stability in mean" which will help us derive the convergence conditions.

Definition 4.3.1. *[66] Given $\mathcal{A}$ is a set of matrices, $\bar{\rho}_k(\mathcal{A})$ is the **largest possible spectral radius** of all products of k matrices chosen in the set $\mathcal{A}$, i.e.,*

$$\bar{\rho}_k(\mathcal{A}) := max \left\{ \rho\left(\prod_{i=1}^{k} \mathbf{A}_i \right) : \forall i, \mathbf{A}_i \in \mathcal{A} \right\}.$$

Definition 4.3.2. *A SHS is **Lyapunov stable in mean** if given a $\xi > 0$, there exists*

$\vartheta(\xi, 0)$ *such that* $\mathbb{E}(\|\mathbf{x}_0\|) < \vartheta$ *implies*

$$\mathbb{E}(\sup_{k\geq 0} \|\mathbf{x}_k\|) < \xi.$$

From the proof for theorem 4.3.1:

$$\mathbf{h}_k = \lambda \sum_{i=0}^{k-1} \mathbf{S}_1^{k-1-i} \mathbf{S}_2 \mathbf{l}_i,$$

so we can write

$$\mathbb{E}(\mathbf{e}_k) = \lambda \sum_{i=0}^{k-1} \mathbf{\Lambda}_{\mathsf{i}}^{k-1-i} (\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}}) \mathbb{E}(\mathbf{x}_i) \tag{4.9}$$

One interesting observation is that (4.9) does not depend on $\mathbb{E}(\mathbf{e}_{k-1})$ if $\mathbf{e}_0$ is 0. Moreover, evolution of $\mathbb{E}(\mathbf{e}_k)$ is related to the statistical properties of the actual state $\{\mathbf{x}_i\}_{i=0}^{k-1}$. For a typical non-hybrid system, $\mathbb{E}(\mathbf{x}_k)$ can be obtained based on the system model, where for SHS, since the discrete state transitions are random, $\mathbb{E}(\mathbf{x}_k)$ cannot be obtained without knowing the actual mode. Indeed, $\bar{\rho}_k(\mathcal{A})$ can help to illustrate the convergence property of $\mathbb{E}(\mathbf{x}_k)$. For an SHS with $\mathcal{Q} = \{q_1, \cdots, q_d\}$, let $\mathcal{A} = \{\mathbf{A}_{q_1}, \cdots, \mathbf{A}_{q_d}\}$. Denote $\mathbf{A}_s^k$ as product of matrices on $\mathcal{A}$, i.e.,

$$\mathbf{A}_s^k = \prod_{i=1}^{k} \mathbf{A}_i, \forall i, \mathbf{A}_i \in \mathcal{A}. \tag{4.10}$$

Recall that the initial assumption is $\mathbf{x}_0 \sim \mathcal{N}(\boldsymbol{\mu}_0, \mathbf{\Sigma}_0)$, so we can write $\mathbb{E}(\mathbf{x}_k) = \mathbf{A}_s^k \boldsymbol{\mu}_0$.

Theorem 4.3.2. $\lim_{k\to\infty} \bar{\rho}_k(\mathcal{A}) < 1$ *gives a sufficient condition for* $\lim_{k\to\infty} \|\mathbb{E}(\mathbf{e}_k)\| < \infty$.

Proof. By definition, $\forall k, \rho(\mathbf{A}_s^k) \leq \bar{\rho}_k(\mathcal{A})$, so $\lim_{k\to\infty} \rho(\mathbf{A}_s^k) < 1$, which leads to

$$\lim_{k\to\infty} \|\mathbb{E}(\mathbf{x}_k)\| = \lim_{k\to\infty} \left\|\mathbf{A}_s^k \boldsymbol{\mu}_0\right\| = 0.$$

Therefore, the sequence $\{\|\mathbb{E}(\mathbf{x}_k)\|\}_{k=0}^{\infty}$ is bounded by some constant C. Let assume for $k > k^*$, $\|\mathbb{E}(\mathbf{x}_k)\| = 0$. Then $\mathrm{C} = \max_{0\leq k\leq k^*} \|\mathbb{E}(\mathbf{x}_k)\|$. Since

$$\begin{aligned}\|\mathbb{E}(\mathbf{e}_k)\| &= \lambda \left\| \sum_{i=0}^{k-1} \mathbf{\Lambda}_{\mathsf{i}}^{k-1-i} (\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}}) \mathbb{E}(\mathbf{x}_i) \right\| \\ &\leq \lambda \sum_{i=0}^{k-1} \left\| \mathbf{\Lambda}_{\mathsf{i}}^{k-1-i} \right\| \left\| \mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}} \right\| \left\| \mathbb{E}(\mathbf{x}_i) \right\|,\end{aligned}$$

then,

$$\begin{aligned}\lim_{k\to\infty} \|\mathbb{E}(\mathbf{e}_k)\| &\leq \lim_{k\to\infty} \lambda \sum_{i=0}^{k-1} \left\| \mathbf{\Lambda}_{\mathsf{i}}^{k-1-i} \right\| \left\| \mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}} \right\| \left\| \mathbb{E}(\mathbf{x}_i) \right\| \\ &= \lambda \sum_{i=0}^{k^*} \left\| \mathbf{\Lambda}_{\mathsf{i}}^{k^*-i} \right\| \left\| \mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}} \right\| \left\| \mathbb{E}(\mathbf{x}_i) \right\| \\ &\leq \sum_{i=0}^{k^*} \mathrm{C}_{1,i} \mathrm{C}_2 \mathrm{C} < \infty,\end{aligned}$$

where $\mathrm{C}_{1,i} = \left\| \mathbf{\Lambda}_{\mathsf{i}}^{k^*-i} \right\|$ and $\mathrm{C}_2 = \left\| \mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}} \right\|$, the finite sum of some bounded constants will be bounded. □

Theorem 4.3.3. *If a SHS is Lyapunov stable in mean, then* $\lim_{k\to\infty} \|\mathbb{E}(\mathbf{e}_k)\| < \infty$.

Proof. Since the SHS is Lyapunov stable in mean, $\mathbb{E}(\sup_{k\geq 0} \|\mathbf{x}_k\|) < \xi$, which leads to

$$\lim_{k\to\infty} \mathbb{E}(\|\mathbf{x}_k\|) < \xi \overset{(a)}{\Rightarrow} \lim_{k\to\infty} \|\mathbb{E}(\mathbf{x}_k)\| < \xi.$$

(a) is the result of Jensen's inequality. Therefore, the sequence $\{\|\mathbb{E}(\mathbf{x}_k)\|\}_{k=0}^{\infty}$ is bounded by ξ.

$$\begin{aligned}\lim_{k\to\infty} \|\mathbb{E}(\mathbf{e}_k)\| &\leq \lim_{k\to\infty} \lambda \sum_{i=0}^{k-1} \left\| \mathbf{\Lambda}_{\mathsf{i}}^{k-1-i} \right\| \left\| \mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}} \right\| \left\| \mathbb{E}(\mathbf{x}_i) \right\| \\ &\leq \lim_{k\to\infty} \lambda \xi \sum_{i=0}^{k-1} \left\| \mathbf{\Lambda}_{\mathsf{i}}^{k-1-i} \right\| \left\| \mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i,t}} \right\| < \infty\end{aligned}$$

□

The evolution of $\mathbb{E}(\mathbf{e}_k)$ depends on $\mathbf{\Lambda}_{\mathsf{i}}$, $\mathbf{\Gamma}_{\mathsf{i},\mathsf{t}}$, $\mathbb{E}(\mathbf{x}_{k-1})$ and υ_k. It reflects how close the expectation of continuous state estimates are to the actual states. So far, we have uncovered the relationship between $\mathbf{\Lambda}_{\mathsf{i}}$, $\mathbf{\Gamma}_{\mathsf{i},\mathsf{t}}$, $\mathbb{E}(\mathbf{x}_{k-1})$, υ_k and the convergence of $\mathbb{E}(\mathbf{e}_k)$ as $k \to \infty$. The result in Theorem 4.3.2 is a strong condition but the condition $\lim_{k\to\infty} \bar{\rho}_k(\mathcal{A}) < 1$ is still challenging and it is algorithmically undecidable [67]. In the following, we will consider a specific type of SHS which has two discrete states. With this assumption, all the theorems discussed in this subsection also hold. Additionally, since i and t can only take two values in this scenario, we are able to easily derive a critical region for discrete state probability that will guarantee convergence in a mean sense.

Consider a stochastic hybrid system with two discrete states. Without loss of generality, we denote $\mathcal{Q} = \{q_1, q_2\}$. When $(\mathsf{t} = q_1)\&(\mathsf{i} = q_2)$ or $(\mathsf{t} = q_2)\&(\mathsf{i} = q_1)$, $\upsilon_k = 1$; otherwise, $\upsilon_k = 0$. Assume we have prior knowledge on the discrete states distribution, for example, let $\mathbb{P}(\mathsf{t} = q_1) = \varepsilon$ and $\mathbb{P}(\mathsf{t} = q_2) = 1 - \varepsilon$. Define ζ_k to be a i.i.d Bernoulli random variable that represents the actual mode t as:

$$\zeta_k = \begin{cases} 1(\mathsf{t} = q_1) & \text{with probability } \varepsilon; \\ 0(\mathsf{t} = q_2) & \text{with probability } 1 - \varepsilon. \end{cases} \tag{4.11}$$

$$\begin{aligned} \text{For } \zeta_k = 1,\ \mathbf{e}_k &= \mathbf{\Lambda}_{\mathsf{i}}\mathbf{e}_{k-1} + \upsilon_k(\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},1})\mathbb{E}(\mathbf{x}_{k-1}) \\ &\overset{(a)}{=} \mathbf{\Lambda}_{\mathsf{i}}\mathbf{e}_{k-1} + \gamma_k(\mathbf{\Lambda}_2 - \mathbf{\Gamma}_{2,1})\mathbb{E}(\mathbf{x}_{k-1}). \\ \text{For } \zeta_k = 0,\ \mathbf{e}_k &= \mathbf{\Lambda}_{\mathsf{i}}\mathbf{e}_{k-1} + \upsilon_k(\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},2})\mathbb{E}(\mathbf{x}_{k-1}) \\ &\overset{(b)}{=} \mathbf{\Lambda}_{\mathsf{i}}\mathbf{e}_{k-1} + \gamma_k(\mathbf{\Lambda}_1 - \mathbf{\Gamma}_{1,2})\mathbb{E}(\mathbf{x}_{k-1}). \end{aligned}$$

(a) and (b) result from the fact that when $\mathsf{i} = \mathsf{t}$, $\upsilon_k = 0$, we can replace the term $\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},1}$

with $\mathbf{\Lambda}_2 - \mathbf{\Gamma}_{2,1}$ and $\mathbf{\Lambda}_1 - \mathbf{\Gamma}_{1,2}$ respectively. Equation (4.7) in this case can be rewritten as:

$$\begin{aligned}\mathbf{e}_k =& \zeta_k[\mathbf{\Lambda}_{\text{i}}\mathbf{e}_{k-1} + \upsilon_k(\mathbf{\Lambda}_2 - \mathbf{\Gamma}_{2,1})\mathbb{E}(\mathbf{x}_{k-1})] + (1-\zeta_k)[\mathbf{\Lambda}_{\text{i}}\mathbf{e}_{k-1} + \upsilon_k(\mathbf{\Lambda}_1 - \mathbf{\Gamma}_{1,2})\mathbb{E}(\mathbf{x}_{k-1})] \\ =& \mathbf{\Lambda}_{\text{i}}\mathbf{e}_{k-1} + \zeta_k\upsilon_k(\mathbf{\Lambda}_2 - \mathbf{\Gamma}_{2,1})\mathbb{E}(\mathbf{x}_{k-1}) + (1-\zeta_k)\upsilon_k(\mathbf{\Lambda}_1 - \mathbf{\Gamma}_{1,2})\mathbb{E}(\mathbf{x}_{k-1})\end{aligned} \tag{4.12}$$

Therefore, $\{\mathbf{e}_k\}_{k=0}^{\infty}$ is a stochastic process related to υ_k and ζ_k. As stated, we consider the convergence in mean, i.e., $\lim_{k\to\infty}\mathbb{E}(\mathbf{e}_k) < \infty$. $\mathbb{E}(\mathbf{e}_k) = \mathbb{E}(\mathbb{E}(\mathbf{e}_k|\mathbf{e}_{k-1}))$, the outer expectation is taken over $\mathbf{e}_{k-1}$ while the inner expectation is taken over the random variable ζ_k and υ_k.

$$\mathbb{E}(\mathbf{e}_k|\mathbf{e}_{k-1}) = \mathbf{\Lambda}_{\text{i}}\mathbf{e}_{k-1} + \varepsilon\lambda(\mathbf{\Lambda}_2 - \mathbf{\Gamma}_{2,1})\mathbb{E}(\mathbf{x}_{k-1}) + (1-\varepsilon)\lambda(\mathbf{\Lambda}_1 - \mathbf{\Gamma}_{1,2})\mathbb{E}(\mathbf{x}_{k-1}) \tag{4.13}$$

With iteration function $\mathbf{h} = \mathbb{f}_\lambda(\mathbf{h}, \mathbf{l})$ defined in last section,

$$\mathbb{E}(\mathbf{e}_k) = \mathbb{f}_\lambda(\mathbb{E}(\mathbf{e}_{k-1}), \mathbb{E}(\mathbf{x}_{k-1}))$$

with $\mathbf{S}_1 = \mathbf{\Lambda}_{\text{i}}$ and $\mathbf{S}_2 = \varepsilon(\mathbf{\Lambda}_2 - \mathbf{\Gamma}_{2,1}) + (1-\varepsilon)(\mathbf{\Lambda}_1 - \mathbf{\Gamma}_{1,2})$. Generally, $\mathbb{f}_\lambda(\mathbf{h}, \mathbf{l})$ is not a function of ε. However, if $\mathbf{S}_1$ and $\mathbf{S}_2$ are defined as shown, the 'hidden' parameter ε in $\mathbf{S}_2$ has impact on the convergence property of $\|\mathbb{f}_\lambda(\mathbf{h}, \mathbf{l})\|$. In the following, we will quantify the impact and obtain a critical stable region of ε.

Theorem 4.3.4. *For a system with $\mathcal{Q} = \{q_1, q_2\}$ with ζ_k as defined in (4.11) indicates the actual mode, $\rho(\varepsilon\mathbf{A}_1 + (1-\varepsilon)\mathbf{A}_2) < 1$ gives a sufficient condition for $\lim_{k\to\infty}\|\mathbb{E}(\mathbf{e}_k)\| < \infty$.*

Proof. Since $\mathbf{x}_k = (\zeta_k\mathbf{A}_1 + (1-\zeta_k)\mathbf{A}_2)\mathbf{x}_{k-1} + \mathbf{w}_k$ and $\mathbb{E}(\mathbf{x}_k) = \mathbb{E}[\mathbb{E}(\mathbf{x}_k|\mathbf{x}_{k-1})]$, where the inner expectation is over ζ_k and the outer expectation is over $\mathbf{x}_{k-1}$, we can write

$$\begin{aligned}\mathbb{E}(\mathbf{x}_k) &= (\varepsilon\mathbf{A}_1 + (1-\varepsilon)\mathbf{A}_2)\mathbb{E}(\mathbf{x}_{k-1}) \\ &= (\varepsilon\mathbf{A}_1 + (1-\varepsilon)\mathbf{A}_2)^k\mathbb{E}(\mathbf{x}_0) = (\varepsilon\mathbf{A}_1 + (1-\varepsilon)\mathbf{A}_2)^k\boldsymbol{\mu}_0.\end{aligned}$$

The condition $\rho(\varepsilon\mathbf{A}_1 + (1-\varepsilon)\mathbf{A}_2) < 1$ leads to $\lim_{k\to\infty}\mathbb{E}(\mathbf{x}_k) = 0$. Follow the proof in Theorem 4.3.2, $\lim_{k\to\infty}\|\mathbb{E}(\mathbf{e}_k)\| < \infty$. □

Theorem 4.3.4 shows that for a given SHS with $\mathbf{A}_1$ and $\mathbf{A}_2$, if

$$\varepsilon \in \left\{ \varepsilon^* \;\middle|\; \rho\Big(\varepsilon^* \mathbf{A}_1 + (1-\varepsilon^*)\mathbf{A}_2\Big) < 1 \right\}, \tag{4.14}$$

$\mathbb{E}(\mathbf{e}_k)$ is bounded. Therefore, (4.14) represents the stable region for ε where the error converges in mean. This result is specially useful in practical systems as it helps identify the fidelity needed in estimating the discrete state to ensure that the continuous state estimates converge.

4.4 Generalized SHS with Arbitrary Discrete State Transitions

In the previous section, we derive the formulation of bias dynamics that results from mode mismatch errors. Specifically, for a SHS with two discrete states and the discrete state transitions are modeled via an i.i.d. binary Bernoulli random variables, the previous section presents a sufficient condition such that the bias dynamics is statistically convergent. As an extension, we now relax the constraints on two modes and i.i.d. Bernoulli transitions. The SHS model considered in this section is general and can be applied for many practical systems. The novelty lies in modeling the bias dynamics as a transformed switched system enabling us to exploit techniques developed for stability analysis of switched system to our problem of interest.

4.4.1 Transformed Switched System

Thus far, we have derived the dynamics of the bias in a mode-based Kalman filter. In equation (4.7), the bias evolves based on matrices $\mathbf{\Lambda}_{\mathsf{i}}$ and $\mathbf{\Gamma}_{\mathsf{i},\mathsf{t}}$. As defined in the previous section, i and t are random variables that represent estimated and true mode at time k. In general, for an SHS with discrete state space $\mathcal{Q} = \{q_1, \cdots, q_d\}$, if the actual state is t, there are $d-1$ mode mismatch errors could happen. Intuitively, we want to derive the evolution

of $\mathbf{e}_k$ as a stochastic equation based on the probabilistic event of mode mismatch occurrence. In the following, we will formally model this random process by introducing two sequences of random variables, $\{\mathbf{\Theta}_\mathtt{t}\}_{\mathtt{t}=1}^{\mathtt{t}=d}$ and $\{\mathbf{\Xi}_\mathtt{t}\}_{\mathtt{t}=1}^{\mathtt{t}=d}$ as:

$$\mathbf{\Theta}_\mathtt{t} = \begin{cases} \mathbf{\Lambda}_1 & \text{with probability } \lambda_{1,\mathtt{t}}; \\ \mathbf{\Lambda}_2 & \text{with probability } \lambda_{2,\mathtt{t}}; \\ \vdots & \vdots \\ \mathbf{\Lambda}_d & \text{with probability } \lambda_{d,\mathtt{t}} \end{cases}$$

with $\sum_{\mathtt{i}=1}^{d} \lambda_{\mathtt{i},\mathtt{t}} = 1$. For a given $\mathtt{t}$, $\mathbf{\Theta}_\mathtt{t}$ is a random variable on the outcome space $\{\mathbf{\Lambda}_1, \cdots, \mathbf{\Lambda}_d\}$ and all the events $\mathbf{\Theta}_\mathtt{t} = \mathbf{\Lambda}_1$, $\cdots$, $\mathbf{\Theta}_\mathtt{t} = \mathbf{\Lambda}_d$ are mutually exclusive. The probability $\lambda_{\mathtt{i},\mathtt{t}}$ can be interpreted as the probability that the estimated mode is $\mathtt{i}$ while the true mode is $\mathtt{t}$. It is worth mentioning that in realistic applications, the probability of mode mismatch may not only be a function of $\mathtt{i}$ and $\mathtt{t}$ but can also correlated across time or across mode. Similarly, a random variable $\mathbf{\Xi}_\mathtt{t}$ is defined as:

$$\mathbf{\Xi}_\mathtt{t} = \begin{cases} \mathbf{\Lambda}_1 - \mathbf{\Gamma}_{1,\mathtt{t}} & \text{with probability } \lambda_{1,\mathtt{t}}; \\ \mathbf{\Lambda}_2 - \mathbf{\Gamma}_{2,\mathtt{t}} & \text{with probability } \lambda_{2,\mathtt{t}}; \\ \vdots & \vdots \\ \mathbf{\Lambda}_d - \mathbf{\Gamma}_{d,\mathtt{t}} & \text{with probability } \lambda_{d,\mathtt{t}}. \end{cases}$$

Note that the probabilities are the same as $\mathbf{\Theta}_\mathtt{t}$ for the same $\mathtt{t}$. With $\mathbf{\Theta}_\mathtt{t}$ and $\mathbf{\Xi}_\mathtt{t}$, we can rewrite equation (4.7) as:

$$\mathbf{e}_k = \mathbf{\Theta}_\mathtt{t}\mathbf{e}_{k-1} + \mathbf{\Xi}_\mathtt{t}\mathbb{E}(\mathbf{x}_{k-1}). \tag{4.15}$$

From equation (4.15), $\{\mathbf{e}_k\}_{k=0}^{\infty}$ is a stochastic process for a given initial value $\mathbf{e}_0$. The process $\mathbf{e}_k$ is bounded with probability 1 if and only if $\mathbb{E}(\mathbf{e}_k)$ is bounded [39]. Therefore, we consider convergence in mean, i.e., $\lim_{k\to\infty} \mathbb{E}(\mathbf{e}_k) < \infty$. According to the tower rule, we

have $\mathbb{E}(\mathbf{e}_k) = \mathbb{E}(\mathbb{E}(\mathbf{e}_k|\mathbf{e}_{k-1}))$, where the outer expectation is taken over $\mathbf{e}_{k-1}$ and the inner expectation is taken over the random variables $\mathbf{\Theta}_{\mathsf{t}}$ and $\mathbf{\Xi}_{\mathsf{t}}$. Therefore,

$$\mathbb{E}(\mathbf{e}_k) = \sum_{\mathsf{i}=1}^{d} \lambda_{\mathsf{i},\mathsf{t}} \mathbf{\Lambda}_{\mathsf{i}} \mathbb{E}(\mathbf{e}_{k-1}) + \sum_{\mathsf{i}=1}^{d} \lambda_{\mathsf{i},\mathsf{t}} (\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},\mathsf{t}}) \mathbb{E}(\mathbf{x}_{k-1}) \tag{4.16}$$

Recall that a discrete-time stochastic system is defined on the hybrid space of continuous and discrete state spaces. The dynamics of $\mathbb{E}(\mathbf{e}_k)$ in equation (4.16) follows the structure of the system in (4.1). That is, the evolution of $\mathbb{E}(\mathbf{e}_k)$ is linearly dependant on the previous $\mathbb{E}(\mathbf{e}_{k-1})$ and the current mode t (which by definition is the actual discrete state in the original system). Therefore, we propose to define a transformed stochastic hybrid system to describe (4.16) as:

$$\mathbf{x}_k^* = \mathbf{F}_{q_k} \mathbf{x}_{k-1}^* + \mathbf{G}_{q_k} \mathbf{u}_{k-1}, \tag{4.17}$$

where the continuous state $\mathbf{x}_k^* = \mathbb{E}(\mathbf{e}_k)$ and $\mathbf{u}_k = \mathbb{E}(\mathbf{x}_k)$ can be treated as an external input. We use the same notation q_k to denote the discrete state since it follows the same transitions in both the original system and the transformed switched system. The system matrices are:

$$\mathbf{F}_{q_k} = \sum_{\mathsf{i}=1}^{d} \lambda_{\mathsf{i},q_k} \mathbf{\Lambda}_{\mathsf{i}}, \quad \mathbf{G}_{q_k} = \sum_{\mathsf{i}=1}^{d} \lambda_{\mathsf{i},q_k} (\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q_k})$$

Our goal is to find conditions under which $\mathbb{E}(\mathbf{e}_k)$ converges. It is important to note that the evolution of $\mathbf{x}_k^*$ in (4.17) does not contain uncertainty (i.e., modeling noise) as in (4.1). Thus, (4.17) is effectively a switched system. Since in this section, we consider a generalized SHS model without restricting ourselves to any specific type of discrete state transitions. At a higher level, the generalized SHS can be astracted as a switched system with arbitrary switching. This allows us to neglect specific details of the discrete state behavior and instead incorporate all possible switching patterns [10]. With this connection between switched system and the generalized SHS model in mind, we confine ourselves to the convention of switched systems with arbitrary switching signals throughout the remainder of this section. With the transformed switched system (4.17), this problem is equivalent to analyze the

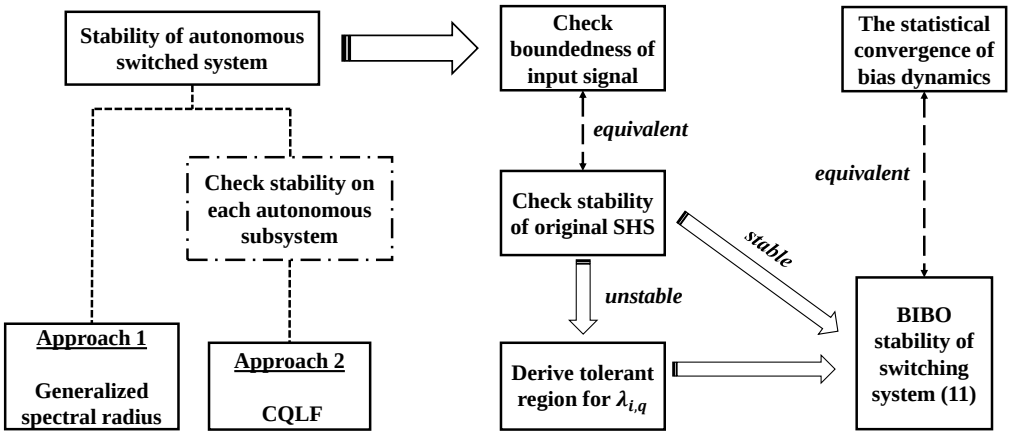


Figure 4.1: *Stability analysis for the transformed switched system*

stability of $\mathbf{x}_k^*$. As stated, we abstract the discrete state transitions in (4.1) as arbitrary switching between each linear subsystem. Therefore, the goal is to find conditions such that the switched system (4.17) with arbitrary switching signal is statistically stable. Additionally, since the system matrices $\mathbf{F}_{q_k}$ and $\mathbf{G}_{q_k}$ depend on the probability of mode mismatch λ_{i,q_k}, the impact of λ_{i,q_k} on the stability of (4.17) also needs to be investigated. In the following, we will first review and summarize the progress that has been made regarding the stability for switched systems and then derive convergence conditions for stability of (4.17).

4.4.2 Main Results

As with general linear systems, numerous concepts of stability have been defined for switched systems. In this section, we use the definition of asymptotic stability for switched systems.

Definition 4.4.1. *The switched system (4.17) is* ***asymptotically stable*** *if there exists some* $\delta > 0$ *such that* $\|\mathbf{x}_0^*\| < \delta$ *implies* $\forall k, \|\mathbf{x}_k^*\| < \epsilon$ *(or* $\lim_{k\to\infty} \|\mathbf{x}_k^*\| = 0$*) for all solutions* $\mathbf{x}_k^*$ *of the system.*

Remark 4.4.1. *A switched system is marginally stable if it is neither asymptotically stable nor unstable.*

Note that asymptotic stability gives a stronger condition for $\lim_{k\to\infty} \|\mathbf{x}_k^*\| < \infty$ since it not

only requires convergence but requires convergence to the origin. The definition of marginal stability implies that the state trajectory is bounded but not necessarily convergent, which is equivalent to $\lim_{k\to\infty} \|\mathbf{x}_k^*\| < \infty$. Therefore, conditions for asymptotic stability is sufficient to guarantee $\lim_{k\to\infty} \|\mathbf{x}_k^*\| < \infty$. Also, because asymptotic stability is closely related to the stability of the corresponding autonomous system, it is typical to consider the stability of the autonomous system first. For the transformed switched system in (4.17), the corresponding autonomous system is:

$$\mathbf{x}_k^* = \mathbf{F}_{q_k} \mathbf{x}_{k-1}^*. \tag{4.18}$$

Among the existing research works, there are primarily two approaches to address the stability of the autonomous system in (4.18). One approach involves solving the generalized spectral radius of bounded semigroups of matrices. This approach leads to the joint spectral radius problem [68]. While its computation is Turing-undecidable in general, the computation and approximation of generalized spectral radius is an active area of research. The other approach is primarily built on the well-known Lyapunov theory. It should be noted that the common quadratic Lyapunov function (CQLF) approach is a well-known approximation of the joint spectral radius and the existence of a CQLF is a sufficient condition for the asymptotic stability. Since generalized spectral radius problem is computationally unsolvable as proved in [67], our main results are built on Lyapunov theory. The complete process of our analysis is summarized in Fig. 4.1.

We use $\mathbb{\Sigma}_{\mathbf{F}_q} : \mathbf{x}_k^* = \mathbf{F}_q \mathbf{x}_{k-1}^*$ to denote the subsystem corresponding to mode q. The autonomous switched system (4.18) switches between $\mathbb{\Sigma}_{\mathbf{F}_q}$ for all q. The following lemma is introduced in [69].

Lemma 4.4.1. *The switched system (4.18) is asymptotically stable under arbitrary switching signal if:*

(i). $\rho(\mathbf{F}_q) < 1, \forall q \in \mathcal{Q}$;

(ii). $\exists \mathbf{P} = \mathbf{P}' \succ 0, \quad \mathbf{F}_q' \mathbf{P} \mathbf{F}_q - \mathbf{P} \prec 0.$

Condition (i) in Lemma 4.4.1 implies asymptotic stability of every subsystem $\mathbb{\Sigma}_{\mathbf{F}_q}$ and condition (ii) is the existence of common Lyapunov quadratic function (CQLF). Also, it is

worth pointing out that the stability for each subsystem does not imply asymptotic stability of the switched system [70]. The converse does not always hold either. As discussed in [71], by choosing the switching signal carefully, the switched system can be made asymptotically stable even though the subsystem is not. In the following, we first study conditions such that

$$\rho(\mathbf{F}_q) < 1, \forall q \in \mathcal{Q} \tag{4.19}$$

holds, i.e., each subsystem is asymptotically stable.

Stability of the Subsystem

By definition, $\mathbf{F}_q$ is composed of convex combination of matrices as:

$$\mathbf{F}_q = \sum_{\mathsf{i}=1}^{d} \lambda_{\mathsf{i},q} \mathbf{\Lambda}_{\mathsf{i}}$$

The task of checking spectral radius of summation of matrices is not trivial in general. If two matrices are commutable, i.e., $\mathbf{AB} = \mathbf{BA}$, then $\rho(\mathbf{A}+\mathbf{B}) \leq \rho(\mathbf{A}) + \rho(\mathbf{B})$ [72]. If all the matrices are non-negative (element-wise), [73] proves that spectral radius is strictly convex. But all the mentioned results cannot be extended to general cases. Therefore, directly checking the spectral radius is not feasible. An alternative approach is built on Lyapunov theory which demonstrates the relationship between a quadratic Lyapunov function (QLF) and the spectral radius of system matrices.

Lemma 4.4.2. *The following statements are equivalent:*
(i) if there exists a positive definite matrix $\mathbf{P}$ *such that* $\mathbf{F}_q'\mathbf{P}\mathbf{F}_q - \mathbf{P} \prec 0$*;*
(ii) $\rho(\mathbf{F}_q) < 1$*;*
(iii) the subsystem $\mathbb{\Sigma}_{\mathbf{F}_q}$ *is asymptotically stable.*

We first illustrate a property related to the spectral radius of $\mathbf{\Lambda}_{\mathsf{i}}$ in the following lemma.

Lemma 4.4.3. *For a SHS defined in (4.1), if* $(\mathbf{A}_{\mathsf{i}}, \mathbf{B}_{\mathsf{i}}\mathbf{Q}\mathbf{B}_{\mathsf{i}}')$ *is controllable and* $(\mathbf{C}_{\mathsf{i}}, \mathbf{A}_{\mathsf{i}})$ *is observable for all* $\mathsf{i} \in \mathcal{Q}$*, then* $\forall \mathsf{i} \in \mathcal{Q}$*,* $\rho(\mathbf{\Lambda}_{\mathsf{i}}) < 1$*.*

Proof. From the definition,

$$\mathbf{\Lambda}_{\mathsf{i}} = \mathbf{A}_{\mathsf{i}} - \mathbf{K}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}}\mathbf{A}_{\mathsf{i}} = (\mathbf{I} - \mathbf{K}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}})\mathbf{A}_{\mathsf{i}}.$$

For any Kalman filter, the observer gain corresponding to mode i is defined as

$$\mathbf{L}_{\mathsf{i}} = \mathbf{A}_{\mathsf{i}}\mathbf{M}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}}'(\mathbf{C}_{\mathsf{i}}\mathbf{M}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}}' + \mathbf{R})^{-1}.$$

Here, $\mathbf{M}_{\mathsf{i}}$ is the steady error covariance related to steady Kalman gain $\mathbf{K}_{\mathsf{i}}$. Given that $(\mathbf{A}_{\mathsf{i}}, \mathbf{B}_{\mathsf{i}}\mathbf{Q}\mathbf{B}_{\mathsf{i}}')$ is controllable and $(\mathbf{C}_{\mathsf{i}}, \mathbf{A}_{\mathsf{i}})$ is observable for all $\mathsf{i} \in \mathcal{Q}$, the closed-loop dynamics $\mathbf{A}_{\mathsf{i}} - \mathbf{L}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}}$ is stable. That is,

$$\rho(\mathbf{A}_{\mathsf{i}} - \mathbf{L}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}}) < 1.$$

Rewrite it as:

$$\mathbf{A}_{\mathsf{i}} - \mathbf{L}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}} = \mathbf{A}_{\mathsf{i}} - \mathbf{A}_{\mathsf{i}}\mathbf{K}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}} = \mathbf{A}_{\mathsf{i}}(\mathbf{I} - \mathbf{K}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}}).$$

From commutativity property of spectral radius,

$$\rho(\mathbf{A}_{\mathsf{i}} - \mathbf{L}_{\mathsf{i}}\mathbf{C}_{\mathsf{i}}) = \rho(\mathbf{\Lambda}_{\mathsf{i}}) < 1.$$

□

With the fact that all the matrices $\mathbf{\Lambda}_{\mathsf{i}}$ are stable, we have the following theorem.

Lemma 4.4.4. *If there is only one $\lambda_{\mathsf{i},q} > 0$ for each $q \in \mathcal{Q}$, then the subsystem $\mathbb{\Sigma}_{\mathbf{F}_q}$ is asymptotically stable for all $q \in \mathcal{Q}$.*

Proof. Let k_q be the index indicating the non-zero $\lambda_{k_q,q}$ for each $q \in \mathcal{Q}$, note that k_q also takes value in $\mathcal{Q}$. Based on the property of random variable $\mathbf{\Xi}_{\mathtt{t}}$, $\lambda_{k_q,q} = 1$. Therefore, we have

$$\mathbf{F}_q = \mathbf{\Lambda}_{k_q}, \forall q.$$

From Lemma 4.4.3, it is straightforward to conclude that $\rho(\mathbf{F}_q) = \rho(\mathbf{\Lambda}_{k_q}) < 1, \forall q \in \mathcal{Q}$.

According to Lemma 4.4.2, all the subsystems $\mathbb{\Sigma}_{\mathbf{F}_q}$ are asymptotically stable. □

Following the notation in proof of Lemma 4.4.4, we use k_q to denote the index indicating the non-zero $\lambda_{k_q,q}$ for each $q \in \mathcal{Q}$. Note that k_q is not necessarily equal to q. As $\rho(\mathbf{\Lambda}_q) < 1$ for all q, even though the probability of mode mismatch between q and mode k_q is 1 (i.e., the mode mismatches always happen), all the subsystems $\mathbb{\Sigma}_{\mathbf{F}_q}$ are still stable. The physical interpretation behind the result seems inconsistent. However, this result is only related to the stability of the autonomous subsystem but not the complete switched system. In fact, if we take a close look at our system in (4.17), the choice of $\lambda_{k_q,q}$ will have impact on the input matrix $\mathbf{G}_q$. We will discuss this result later.

Lemma 4.4.4 gives a non-trivial condition such that the stability of each subsystem $\mathbb{\Sigma}_{\mathbf{F}_q}$ is guaranteed. However, the condition that only one $\lambda_{\mathsf{i},q} > 0$ is not generally realistic since it eliminates the randomness associated with errors. The next theorem is built on the concept of CQLF and it is applicable for broader choices of $\lambda_{\mathsf{i},q}$.

Theorem 4.4.1. *If for all* $\mathsf{i} \in \mathcal{Q}$, $\mathbf{\Lambda}_\mathsf{i}$ *share a common quadratic Lyapunov function. That is, if there exists a s.p.d. matrix* $\mathbf{P} \in \mathbb{R}^{n\times n}$ *such that*

$$\mathbf{\Lambda}_\mathsf{i}'\mathbf{P}\mathbf{\Lambda}_\mathsf{i} - \mathbf{P} \prec 0, \forall \mathsf{i} \in \mathcal{Q}, \tag{4.20}$$

then every subsystem $\mathbb{\Sigma}_{\mathbf{F}_q}$ $\forall q \in \mathcal{Q}$ *is asymptotically stable for all choices of* $\lambda_{\mathsf{i},q}$.

Proof.

$$\mathbf{\Lambda}_\mathsf{i}'\mathbf{P}\mathbf{\Lambda}_\mathsf{i} - \mathbf{P} \prec 0 \overset{(a)}{\Longleftrightarrow} \mathbf{P} - \mathbf{\Lambda}_\mathsf{i}\mathbf{P}\mathbf{\Lambda}_\mathsf{i}' \succ 0 \overset{(b)}{\Longleftrightarrow} \begin{bmatrix} \mathbf{P} & \mathbf{\Lambda}_\mathsf{i} \\ \mathbf{\Lambda}_\mathsf{i}' & \mathbf{P}^{-1} \end{bmatrix} \succ 0.$$

(a) is due to the fact that $\mathbf{P}$ is s.p.d. and (b) is a result of Schur decomposition. According to Lemma 4.4.2, in order to prove $\mathbb{\Sigma}_{\mathbf{F}_q}$ is asymptotically stable for all q, we need to find if there exists some s.p.d. matrix $\mathbf{P}_q$ for each q such that $\mathbf{P}_q - \mathbf{F}_q\mathbf{P}_q\mathbf{F}_q' \succ 0$.

Since $\mathbf{P} - \mathbf{\Lambda}_\mathsf{i}\mathbf{P}\mathbf{\Lambda}_\mathsf{i}' \succ 0$, therefore $\mathbf{P} - \lambda_{\mathsf{i},q}^2\mathbf{\Lambda}_\mathsf{i}\mathbf{P}\mathbf{\Lambda}_\mathsf{i}' \succ 0$ for $0 \leq \lambda_{\mathsf{i},q} \leq 1$. For all $q \in \mathcal{Q}$, we

have:

$$\begin{bmatrix} \mathbf{P} & \lambda_{\mathsf{i},q}\mathbf{\Lambda}_{\mathsf{i}} \\ \lambda_{\mathsf{i},q}\mathbf{\Lambda}_{\mathsf{i}}' & \mathbf{P}^{-1} \end{bmatrix} \succ 0 \Longrightarrow \sum_{\mathsf{i}=1}^{d} \begin{bmatrix} \mathbf{P} & \lambda_{\mathsf{i},q}\mathbf{\Lambda}_{\mathsf{i}} \\ \lambda_{\mathsf{i},q}\mathbf{\Lambda}_{\mathsf{i}}' & \mathbf{P}^{-1} \end{bmatrix} \succ 0$$

$$\Longrightarrow \begin{bmatrix} \mathbf{P} & \sum_{\mathsf{i}=1}^{d} \lambda_{\mathsf{i},q}\mathbf{\Lambda}_{\mathsf{i}} \\ \sum_{\mathsf{i}=1}^{d} \lambda_{\mathsf{i},q}\mathbf{\Lambda}_{\mathsf{i}}' & \mathbf{P}^{-1} \end{bmatrix} \succ 0$$

$$\Longrightarrow \begin{bmatrix} \mathbf{P} & \mathbf{F}_q \\ \mathbf{F}_q' & \mathbf{P}^{-1} \end{bmatrix} \succ 0 \Longrightarrow \mathbf{P} - \mathbf{F}_q\mathbf{P}\mathbf{F}_q' \succ 0.$$

By taking $\mathbf{P}_q = \mathbf{P}$, we proved that there exists positive definite matrix $\mathbf{P}_q$ for each q such that $\mathbf{P}_q - \mathbf{F}_q\mathbf{P}_q\mathbf{F}_q' \succ 0$. Therefore, every subsystem $\mathbb{\Sigma}_{\mathbf{F}_q}$ $\forall q \in \mathcal{Q}$ is asymptotically stable for all choices of $\lambda_{\mathsf{i},q}$. □

As presented in Lemma 4.4.1, there are two conditions that can guarantee the stability of the autonomous switched system. Condition (i) is related to the stability of each subsystem and we have developed Lemma 4.4.4 and Theorem 4.4.1 determine $\rho(\mathbf{F}_q) < 1$ for all $q \in \mathcal{Q}$. To complete the stability analysis for switched autonomous system in (4.18), we will study conditions such that constraint (ii) in Lemma 4.4.1 is satisfied in the following subsection.

Stability of Switched Autonomous Systems

We have introduced the concept of CQLF in Lemma 4.4.1. For stability analysis and CQLF conditions, [14] provides an excellent survey on the progress that have been made in this research area. In general, determining algebraic conditions (on the subsystems' state matrices) for the existence of a CQLF remains an open task. For switched system with only two modes, [74] derives a necessary and sufficient for the existence of a CQLF for a second-order (two dimensional) continuous-time switched system with two modes while a similar approach proposed in [69] by considering discrete-time system. Their approach is based on the stability of the matrix pencil constructed using the state matrices corresponding to the two modes. The matrix pencil presents a different viewpoint on the CQLF existence problem

but it also lacks an analytical solution.

In this section, the switched system in (4.18) contains unknown variable $\lambda_{\mathsf{i},q}$ in the subsystem matrices $\mathbf{F}_q$. Due to the unknown values in $\mathbf{F}_q$ and lack of algebraic solutions, we cannot directly solve the LMI conditions nor derive constraints on $\lambda_{\mathsf{i},q}$ such that the existence of CQLF for $\mathbf{F}_q$ is guaranteed. In the following, we propose to establish a relationship between the existence of CQLF for $\mathbf{\Lambda}_\mathsf{i}$ and $\mathbf{F}_q$ and then obtain conditions for stability of switched system (4.18) regardless of the choice of $\lambda_{\mathsf{i},q}$.

Theorem 4.4.2. *If there exists a CQLF for $\mathbf{\Lambda}_\mathsf{i}, \forall \mathsf{i} \in \mathcal{Q}$, then there exists a CQLF for $\mathbf{F}_q, \forall q \in \mathcal{Q}$. As a consequence, the switched system (4.18) is asymptotically stable under arbitrary switching signal.*

Proof. We will use the similar approach as shown in the proof of Theorem 4.4.1. If there exists a CQLF for $\mathbf{\Lambda}_\mathsf{i}$, we know that there exists a positive definite matrix $\mathbf{P} \in \mathbb{R}^{n\times n}$ such that

$$\mathbf{\Lambda}_\mathsf{i}'\mathbf{P}\mathbf{\Lambda}_\mathsf{i} - \mathbf{P} \prec 0, \forall \mathsf{i} \in \mathcal{Q}.$$

As a result of Theorem 4.4.1, for all $q \in \mathcal{Q}$, we have

$$\sum_{\mathsf{i}=1}^{d} \begin{bmatrix} \mathbf{P} & \lambda_{\mathsf{i},q}\mathbf{\Lambda}_\mathsf{i} \\ \lambda_{\mathsf{i},q}\mathbf{\Lambda}_\mathsf{i}' & \mathbf{P}^{-1} \end{bmatrix} \succ 0 \Longrightarrow \begin{bmatrix} \mathbf{P} & \mathbf{F}_q \\ \mathbf{F}_q & \mathbf{P}^{-1} \end{bmatrix} \succ 0 \Longrightarrow \mathbf{F}_q'\mathbf{P}\mathbf{F}_q - \mathbf{P} \prec 0.$$

Therefore, there exists a CQLF for $\mathbf{F}_q, \forall q \in \mathcal{Q}$. From Lemma 4.4.1, the switched system (4.18) is asymptotically stable under arbitrary switching signal. □

The condition derived in Theorem 4.4.2 is only based on all the matrices $\mathbf{\Lambda}_\mathsf{i}$ which can be determined given the system matrix. The LMI condition can be easily checked in practice via an LMI solver alleviating the lack of an analytical solution. As illustrated in Figure 4.1, we have completed the discussion for the stability of autonomous switched system (4.18) thus far. In the following, we will consider stability of the complete transformed switched system (4.17) including the input term.

Bounded-Input Bounded-Output (BIBO) Stability

For the transformed switched system in (4.17), we introduce the notion of BIBO stability that has been defined in [12].

Definition 4.4.2. *The system in (4.17) is* ***BIBO stable*** *if there exists a positive constant* η *such that for any essentially bounded input signal* $\mathbf{u}$*, the continuous state* $\mathbf{x}^*$ *satisfies*

$$\sup_{k\geq 0} \|\mathbf{x}_k^*\| \leq \eta \sup_{k\geq 0} \|\mathbf{u}_k\| .$$

According to this definition, an input signal cannot be amplified by a factor greater than some finite constant η after passing through the system if the system is BIBO stable. It has been proven that if the corresponding autonomous switched system (4.18) is asymptotically stable, then the input-output system (4.17) is BIBO stable provided the input matrix $\mathbf{G}_q$ is uniformly bounded in time for all q [75]. This in fact is the case when the system switches between a finite family of matrices. In our transformed switched system, the input signal $\mathbf{u}_k = \mathbf{x}_k$, where $\mathbf{x}_k$ is the continuous state of original system (4.1). Therefore, depending on the stability of (4.1), $\mathbf{u}_k$ can be either bounded or unbounded. Therefore, we should consider two different scenarios based on the boundedness of $\mathbf{u}_k$ in the following discussions.

Scenario 1: Original system in (4.1) is not asymptotically stable

If the original system in (4.1) is unstable, then $\sup_{k\geq 0} \|\mathbf{u}_k\| = \sup_{k\geq 0} \|\mathbf{x}_k\| = \infty$. Since $\mathbf{u}_k$ is an n-dimensional vector, when $\mathbf{u}_k$ is unbounded, at least one of the elements in the vector is unbounded. We refer to those elements as unstable components and these components are collected in the set $\mathcal{I}$:

$$\mathcal{I} = \left\{ i : \sup_{k\geq 0} \mathbf{u}_k^{[i]} = \infty \right\}.$$

For this situation, if the columns of $\mathbf{G}_q$ corresponding to those unstable components of $\mathbf{u}_k$ are 0, then the boundedness of $\sup_{j\geq 0,q} \|\mathbf{G}_q \mathbf{u}_j\|$ is guaranteed. The process of finding the stable region for each probability of mode mismatch error is summarized in Algorithm 4:

Generally, $\lambda_{\mathsf{i},q} = 1$ for $\mathsf{i} = q$ should always be a solution of Algorithm 4 because of $\mathbf{\Lambda}_{\mathsf{i}} = \mathbf{\Gamma}_{\mathsf{i},q}$

Algorithm 4 Find stable region of $\lambda_{\mathsf{i},q}$

```
Analyze stability of the original SHS
Find unstable components → 𝓘
for all q ∈ 𝓠 do
    for all i ∈ 𝓠 do
        Let T = Λ_i − Γ_{i,q}
        if ∃j ∈ 𝓘 s.t. j^th column of T is not 0 then
            λ_{i,q} = 0
        else
            0 ≤ λ_{i,q} ≤ 1
        end if
    end for
    Solve Σ_i λ_{i,q} = 1 for all non-zero λ_{i,q}
end for
```

for $\mathsf{i} = q$. Furthermore, this condition along with the result of Lemma 4.4.4 indicate that $\lambda_{\mathsf{i},q} = 1$ for $\mathsf{i} = q$ not only guarantees stability of subsystem but also BIBO stability of the switched system in (4.17). By definition, $\lambda_{\mathsf{i},q}$ represents the probability that true mode is q while estimated mode is i. $\lambda_{\mathsf{i},q} = 1$ for $\mathsf{i} = q$ meaning that there is no mode mismatch error. Therefore, the convergence of $\mathbf{x}_k^*$ (i.e., the bias generated from mode-based Kalman filter) is reasonable. Besides the trivial solution, Algorithm 4 also gives a less conservative result. For those unstable components in the original SHS, if the difference of $\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q}$ at the column corresponding to the unstable components are all 0, the mode-based Kalman filter is still tolerant of the mode mismatch between i and q.

Scenario 2: Original system in (4.1) is asymptotically stable

If the original system in (4.1) is asymptotically stable, then the continuous state $\mathbf{x}_k$ (i.e., $\mathbf{u}_k$ in the transformed switched system) is bounded. Since linear transformations of a vector is a bounded operator in Euclidean space, for a bounded vector $\mathbf{u}$, $\mathbf{Gu}$ is bounded. For this situation, we are interested in minimizing the upper bound of $\|\mathbf{x}_k^*\|$. From the definition of BIBO stability, we can write

$$\|\mathbf{x}_k^*\| \leq \eta \sup_{k\geq 0,q} \|\mathbf{G}_q \mathbf{u}_k\| \overset{(a)}{\leq} \eta \max_q \|\mathbf{G}_q\| \sup_{j\geq 0} \|\mathbf{u}_j\|, \tag{4.21}$$

where η and $\sup_{j\geq 0}\|\mathbf{u}_j\|$ are fixed constant for a given system and $\mathbf{G}_q$ is related to the unknown variable $\lambda_{\mathsf{i},q}$. The equality in (a) holds if and only if each row of $\mathbf{G}_q$ is linearly dependent of $\mathbf{u}_k$ for all q, k. In this framework, we seek to address the following questions:

(1) Given the probability of mode mismatch is $\mathcal{P}$, i.e., $\sum_{\substack{\mathsf{i}=1\\ \mathsf{i}\neq q}}^{d}\lambda_{\mathsf{i},q} = \mathcal{P}, \forall q$, what is the lowest upper bound of $\|\mathbf{x}_k^*\|$?

(2) Given a certain upper bound $\mathcal{B}$ of $\|\mathbf{x}_k^*\|$, what is the largest tolerant region for mode mismatch probability $\mathcal{P}$ that will guarantee that $\mathcal{B}$ is achievable?
The following theorem is developed to answer the first question.

Theorem 4.4.3. *Given the probability of mode mismatch $\mathcal{P} \neq 0$ and the original system in (4.1) is asymptotically stable, the lowest upper bound of $\|\mathbf{x}_k^*\|$ that can be achieved is:*

$$\|\mathbf{x}_k^*\| \leq \eta \cdot \mathcal{P} \cdot \sup_{j\geq 0}\|\mathbf{u}_j\| \cdot \max_{q}\min_{\mathsf{i},\mathsf{i}\neq q}\|\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q}\|\,.$$

Proof. From the definition of $\mathbf{G}_q$,

$$\begin{aligned}\|\mathbf{G}_q\| &= \left\|\sum_{\mathsf{i}=1}^{d}\lambda_{\mathsf{i},q}(\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q})\right\| = \left\|\sum_{\substack{\mathsf{i}=1\\ \mathsf{i}\neq q}}^{d}\lambda_{\mathsf{i},q}(\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q})\right\| \\ &\leq \sum_{\substack{\mathsf{i}=1\\ \mathsf{i}\neq q}}^{d}\lambda_{\mathsf{i},q}\|\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q}\|\,. \end{aligned} \tag{4.22}$$

With the constraint that $\sum_{\substack{\mathsf{i}=1\\ \mathsf{i}\neq q}}^{d}\lambda_{\mathsf{i},q} = \mathcal{P}$, we have:

$$\min_{\lambda_{\mathsf{i},q}}\sum_{\substack{\mathsf{i}=1\\ \mathsf{i}\neq q}}^{d}\lambda_{\mathsf{i},q}\|\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q}\| = \mathcal{P}\min_{\mathsf{i},\mathsf{i}\neq q}\|\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q}\|\,. \tag{4.23}$$

From equation (4.21), we have the lowest bound of $\|\mathbf{x}_k^*\|$ as a function of $\|\mathbf{G}_q\|$. Given the constraint on mode mismatch probability and results of (4.22) and (4.23), we get the lowest

upper bound of $\|\mathbf{x}_k^*\|$ that can be reached is:

$$\|\mathbf{x}_k^*\| \leq \eta \cdot \mathcal{P} \cdot \sup_{j\geq 0} \|\mathbf{u}_j\| \cdot \max_q \min_{\mathrm{i},\mathrm{i}\neq q} \|\mathbf{\Lambda}_\mathrm{i} - \mathbf{\Gamma}_{\mathrm{i},q}\| .$$

□

To assist in the analysis for the second question, we first define an auxiliary function $\phi : \mathbb{R}^{d-1} \to \mathbb{R}$ as:

$$\phi(\boldsymbol{v}) = \max_q \left\| \sum_{i=1}^{d-1} \boldsymbol{v}^{[i]} \mathbf{S}_{i,q} \right\|, \boldsymbol{v} \in \mathbb{R}^{d-1}$$

where $\mathbf{S}_{i,q} \in \mathbb{R}^{n\times n}$ is a series of known matrices for a given q. The following lemma illustrates the convexity of this function.

Lemma 4.4.5. *$\phi(\boldsymbol{v})$ is a convex function respect to $\boldsymbol{v}$.*

Proof. In order prove that $\phi(\boldsymbol{v})$ is a convex function respect to $\boldsymbol{v}$, we want to show that for all $\boldsymbol{v}, \boldsymbol{\nu} \in \mathbb{R}^{d-1}$, and θ with $0 \leq \theta \leq 1$, $\phi(\theta\boldsymbol{v} + (1-\theta)\boldsymbol{\nu}) \leq \theta\phi(\boldsymbol{v}) + (1-\theta)\phi(\boldsymbol{\nu})$. We have

$$\begin{aligned}
\phi(\theta\boldsymbol{v} + (1-\theta)\boldsymbol{\nu}) &= \max_q \left\| \sum_{i=1}^{d-1} (\theta\boldsymbol{v} + (1-\theta)\boldsymbol{\nu})^{[i]} \mathbf{S}_{i,q} \right\| \\
&= \max_q \left\| \theta \sum_{i=1}^{d-1} \boldsymbol{v}^{[i]} \mathbf{S}_{i,q} + (1-\theta) \sum_{i=1}^{d-1} \boldsymbol{\nu}^{[i]} \mathbf{S}_{i,q} \right\| \\
&\leq \max_q \left\| \theta \sum_{i=1}^{d-1} \boldsymbol{v}^{[i]} \mathbf{S}_{i,q} + (1-\theta) \sum_{i=1}^{d-1} \boldsymbol{\nu}^{[i]} \mathbf{S}_{i,q} \right\| \\
&\leq \theta \max_q \left\| \sum_{i=1}^{d-1} \boldsymbol{v}^{[i]} \mathbf{S}_{i,q} \right\| + (1-\theta) \max_q \left\| \sum_{i=1}^{d-1} \boldsymbol{\nu}^{[i]} \mathbf{S}_{i,q} \right\| \\
&= \theta\phi(\boldsymbol{v}) + (1-\theta)\phi(\boldsymbol{\nu}).
\end{aligned}$$

Therefore $\phi(\boldsymbol{v})$ is a convex function on $\boldsymbol{v}$. □

Recall that the second question is to derive the largest tolerant region for mode mismatch probability $\mathcal{P}$ such that an upper bound $\mathcal{B}$ of $\|\mathbf{x}_k^*\|$ is achievable. In other words, we need

to solve for $\lambda_{\mathsf{i},q}$ such that $\sum_{\substack{\mathsf{i}=1\\ \mathsf{i}\neq q}}^{d} \lambda_{\mathsf{i},q} = \mathcal{P}$ and $\|\mathbf{x}_k^*\| \leq \mathcal{B}$ holds. Based on equation (4.21), we have

$$\begin{aligned} &\|\mathbf{x}_k^*\| \leq \eta \max_q \|\mathbf{G}_q\| \sup_{j\geq 0} \|\mathbf{u}_j\| \leq \mathcal{B} \\ \Longrightarrow &\max_q \|\mathbf{G}_q\| \leq \frac{\mathcal{B}}{\eta \cdot \sup_{j\geq 0} \|\mathbf{u}_j\|} \\ \Longrightarrow &\max_q \left\| \sum_{\substack{\mathsf{i}=1\\ \mathsf{i}\neq q}}^{d} \lambda_{\mathsf{i},q} (\mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q}) \right\| \leq \frac{\mathcal{B}}{\eta \cdot \sup_{j\geq 0} \|\mathbf{u}_j\|}. \end{aligned} \tag{4.24}$$

Use the auxiliary function and define $\boldsymbol{\lambda} \in \mathbb{R}^{d-1}$ and $\mathbf{S}_{i,q} = \mathbf{\Lambda}_{\mathsf{i}} - \mathbf{\Gamma}_{\mathsf{i},q}$. We can write the left-hand side of (4.24) as:

$$\phi(\boldsymbol{\lambda}) = \max_q \left\| \sum_{i=1}^{d-1} \boldsymbol{\lambda}^{[i]} \mathbf{S}_{i,q} \right\|.$$

Since $\phi(\boldsymbol{\lambda})$ is convex in $\boldsymbol{\lambda}$, a non-negative bound $\mathcal{B}$ is achievable by taking $\boldsymbol{\lambda}^{[i]} = 0$ for all i. To seek a $\boldsymbol{\lambda}$ such that

$$\phi(\boldsymbol{\lambda}) \leq \frac{\mathcal{B}}{\eta \cdot \sup_{j\geq 0} \|\mathbf{u}_j\|},$$

we will use triangle inequality to approximate $\phi(\boldsymbol{\lambda})$ and get a more conservative condition. Since

$$\phi(\boldsymbol{\lambda}) \leq \max_q \sum_{i=1}^{d-1} \boldsymbol{\lambda}^{[i]} \|\mathbf{S}_{i,q}\|,$$

with $\|\mathbf{S}_{i,q}\|$ is known for all i and q. The condition

$$\max_q \sum_{i=1}^{d-1} \boldsymbol{\lambda}^{[i]} \|\mathbf{S}_{i,q}\| \leq \frac{\mathcal{B}}{\eta \cdot \sup_{j\geq 0} \|\mathbf{u}_j\|} \tag{4.25}$$

is a 1^{st} degree polynomial inequality with $d-1$ variables, and this can provide a feasible region for each $\lambda_{\mathrm{i},q}$ on the $d-1$ dimensions space.

The discussion of BIBO stability completes the convergent analysis of bias dynamics in a mode-based Kalman filter. Both stable and unstable original SHS have been taken into consideration. For an unstable system, we can still stabilize the bias dynamics by specifically choosing the probability $\lambda_{\mathrm{i},q}$. For an asymptotically stable system, we addressed two important questions regarding the minimization of the upper bound for the bias.

4.5 Experimental Results

In this section, we conduct three experiments to verify our main results that presented in Chapter 4.3 and Chapter 4.4. We first consider an SHS with two discrete states and the discrete state transitions are modeled via i.i.d. Bernoulli random variable. Then we simulate a generalized SHS (switched system with arbitrary switching signal). Finally, we illustrate the value of the theoretical results on a small scale smart grid application.

4.5.1 Numerical Example I

We implement an experiment on MATLAB to verify our first main result on an SHS with two discrete states $\mathcal{Q} = \{q_1, q_2\}$. Define matrices $\mathbf{A}_q$ and $\mathbf{C}_q$, $\forall q \in \mathcal{Q}$ as:

$$\mathbf{A}_1 = \begin{pmatrix} 0.9 & 0 \\ 2 & 0.8 \end{pmatrix}, \mathbf{A}_2 = \begin{pmatrix} 0.5 & 0.2 \\ 0.2 & 0.4 \end{pmatrix};$$

$$\mathbf{C}_1 = \mathbf{I}, \mathbf{C}_2 = 5\mathbf{I}.$$

Let the system noise be $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and measurement noise be $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, where $\mathbf{Q} = 0.5\mathbf{I}$ and $\mathbf{R} = 0.3\mathbf{I}$. In this system setting, $(\mathbf{A}_1, \mathbf{Q})$ and $(\mathbf{A}_2, \mathbf{Q})$ are both controllable and $(\mathbf{C}_1, \mathbf{A}_1)$ and $(\mathbf{C}_2, \mathbf{A}_2)$ are observable. From Theorem 4.3.4, we can compute the stable region for ε before executing the system. By solving $\rho(\varepsilon\mathbf{A}_1 + (1-\varepsilon)\mathbf{A}_2) < 1$, we obtain that

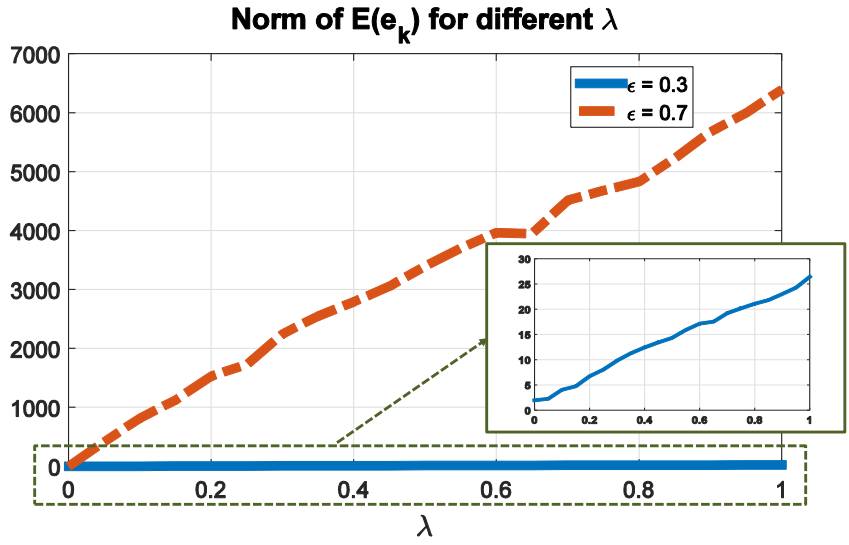


Figure 4.2: *Maximum of* $\|\mathbb{E}(\mathbf{e}_k)\|$ *versus different values of* λ

the stable region of ε as $\varepsilon^* = [0, 0.55] \cup [0.91, 1]$.

To validate the relationship between λ and $\|\mathbb{E}(\mathbf{e}_k)\|$ stated in Theorem 4.3.1, we run $N = 1000$ Monte-Carlo simulations to simulate an SHS over time range $[0, 500]$. Figure 4.2 shows the results for different ε. It can be concluded that as λ increases, $\|\mathbb{E}(\mathbf{e}_k)\|$ increases approximately linearly as postulated by Theorem 4.3.1. We also notice that for different ε, the slope changes. For $\varepsilon = 0.7 \notin \varepsilon^*$, $\|\mathbb{E}(\mathbf{e}_k)\|$ increases rapidly with λ. The slope for ε within the stable region is much smaller than for ε outside of the stable region.

Figure 4.3 reveals the relationship between ε and $\|\mathbb{E}(\mathbf{e}_k)\|$. It is important to remark that all the conditions we have established are sufficient conditions, which means outside the stable region, the evolution of $\mathbb{E}(\mathbf{e}_k)$ can be either stable or unstable. The unshaded regions are the stable region for ε calculated from Theorem 4.3.4. This plot is obtained by executing $N = 1000$ Monte-Carlo simulation over time $[0, 400]$ for two different λ values. It should be noted that the y-axis of Figure 4.3 is $10\log(\|\mathbb{E}(\mathbf{e}_k)\|)$, because outside the guaranteed stable region, $\|\mathbb{E}(\mathbf{e}_k)\|$ might explode to extremely large values. Therefore, the actual difference between stable region and unstable region is quite significant (around 10^5). Additionally, in Fig 4.2 and Figure 4.3, for each λ (or ε), we plot $\max_k \|\mathbb{E}(\mathbf{e}_k)\|$ where $\|\mathbb{E}(\mathbf{e}_k)\|$

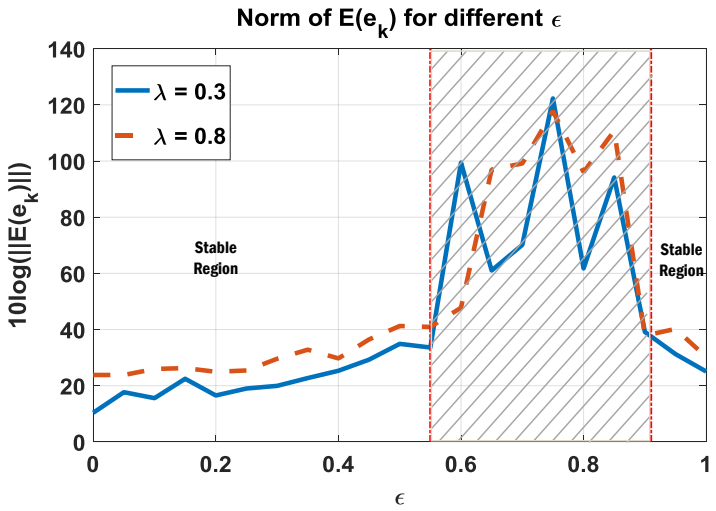


Figure 4.3: *Maximum of* $\|\mathbb{E}(\mathbf{e}_k)\|$ *versus different values of* ε

is average among Monte-Carlo simulations. $\max_k \|\mathbb{E}(\mathbf{e}_k)\|$ can reveal if $\|\mathbb{E}(\mathbf{e}_k)\|$ diverges or not. However, the actual trend of the evolution $\|\mathbb{E}(\mathbf{e}_k)\|$ is not revealed in Figure 4.2 and 4.3. To illustrate how $\|\mathbb{E}(\mathbf{e}_k)\|$ evolves with time k, we choose four sets of value for ε and λ, they are $\{\varepsilon = 0.3, \lambda = 0.2\}$, $\{\varepsilon = 0.3, \lambda = 0.7\}$, $\{\varepsilon = 0.6, \lambda = 0.2\}$ and $\{\varepsilon = 0.6, \lambda = 0.7\}$. For the first two sets of values, $\varepsilon \in \varepsilon^*$ while the other two sets $\varepsilon \notin \varepsilon^*$. We run $N = 6000$ Monte-Carlo simulations over the time range $[0, 300]$ and the results are shown in Figure 4.4. We have the following observations:

1). ε determines the behavior of $\|\mathbb{E}(\mathbf{e}_k)\|$. In this experiment, when ε is within stable region (Figure 4.4 (a)), $\|\mathbb{E}(\mathbf{e}_k)\|$ converges, i.e., $\lim_{k\to\infty} \|\mathbb{E}(\mathbf{e}_k)\| < \infty$. On the other hand, if $\varepsilon \notin \varepsilon^*$ (Figure 4.4 (b)), $\|\mathbb{E}(\mathbf{e}_k)\|$ increases significantly with time k;

2). As expected, increasing of λ increases $\|\mathbb{E}(\mathbf{e}_k)\|$. In Figure 4.4 (a) and (b), while dashed-line ($\lambda = 0.7$) and solid-line ($\lambda = 0.2$) follow the same trend, $\|\mathbb{E}(\mathbf{e}_k)\|$ is larger when $\lambda = 0.7$.

The two observations are consistent with Theorems 4.3.1,4.3.2 and 4.3.4. It suggests that for SHS with $\lim_{k\to\infty} \bar{\rho}_k(\mathcal{A}) < 1$, even the estimated mode is completely incorrect, the error in mode-based Kalman filter can still be bounded. Furthermore, because the conditions in Theorem 4.3.2 and 4.3.4 are only related to system matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, we can verify the convergence conditions apriori.

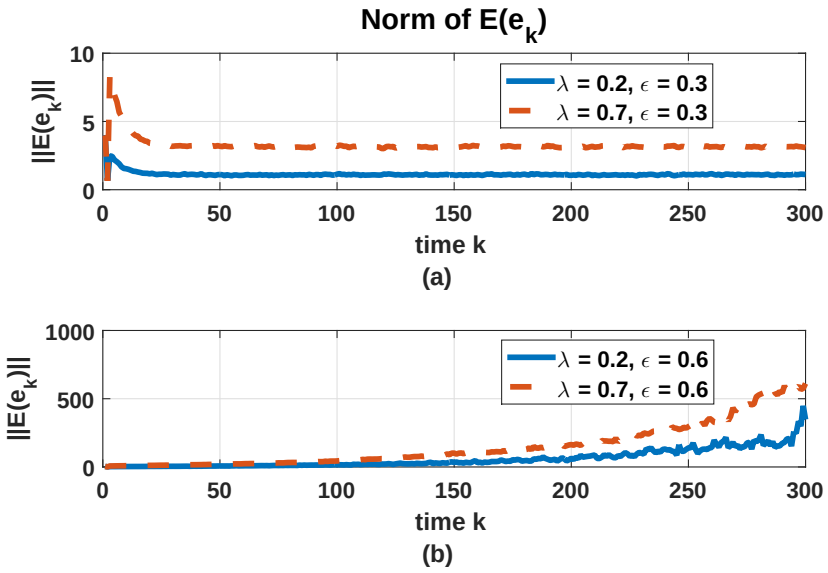


Figure 4.4: $\|\mathbb{E}(\mathbf{e}_k)\|$ *using M-C simulation for different* ε *and* λ

4.5.2 Numerical Example II

Consider a switched system with two discrete states $\mathcal{Q} = \{q_1, q_2\}$. The continuous state is a 2-dimensional vector. Define matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ as:

$$\mathbf{A}_1 = \begin{bmatrix} 0.9 & 0 \\ 0.2 & 0.8 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 0.4 \end{bmatrix}; \mathbf{B}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0.8 \end{bmatrix}, \mathbf{B}_2 = \begin{bmatrix} 1.3 & 0.4 \\ 0.2 & 0.7 \end{bmatrix};$$

$$\mathbf{C}_1 = \mathbf{I}, \mathbf{C}_2 = 5 \times \mathbf{I}.$$

Let the system noise be $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and measurement noise be $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, where $\mathbf{Q} = 0.5 \times \mathbf{I}$ and $\mathbf{R} = 0.3 \times \mathbf{I}$. In this system setting, $(\mathbf{A}_1, \mathbf{B}_1\mathbf{Q}\mathbf{B}_1')$ and $(\mathbf{A}_2, \mathbf{B}_2\mathbf{Q}\mathbf{B}_2')$ are both controllable and $(\mathbf{C}_1, \mathbf{A}_1)$ and $(\mathbf{C}_2, \mathbf{A}_2)$ are observable. The corresponding $\mathbf{\Lambda}_i$ and $\mathbf{\Gamma}_{i,t}$ are calculated as follows:

$$\mathbf{\Lambda}_1 = \begin{bmatrix} 0.2763 & -0.0137 \\ 0.0654 & 0.3232 \end{bmatrix}, \qquad \mathbf{\Lambda}_2 = \begin{bmatrix} 0.0054 & -0.0034 \\ 0.0033 & 0.0205 \end{bmatrix};$$
$$\mathbf{\Gamma}_{1,2} = \begin{bmatrix} -1.2401 & -0.5234 \\ -0.4387 & -0.8091 \end{bmatrix}, \qquad \mathbf{\Gamma}_{2,1} = \begin{bmatrix} 0.7225 & -0.0028 \\ 0.1593 & 0.6496 \end{bmatrix}.$$

For this setup, we get $\|\mathbf{\Lambda}_1\| = 0.3373 < 1$, $\|\mathbf{\Lambda}_2\| = 0.0209 < 1$. Therefore, for any choice of $\lambda_{\mathtt{i},\mathtt{t}}$, we have

$$\begin{aligned} \rho(\mathbf{F}_1) = \rho(\lambda_{1,1}\mathbf{\Lambda}_1 + \lambda_{2,1}\mathbf{\Lambda}_2) \leq \lambda_{1,1}\|\mathbf{\Lambda}_1\| + \lambda_{2,1}\|\mathbf{\Lambda}_2\| < 1, \\ \rho(\mathbf{F}_2) = \rho(\lambda_{1,2}\mathbf{\Lambda}_1 + \lambda_{2,2}\mathbf{\Lambda}_2) \leq \lambda_{1,2}\|\mathbf{\Lambda}_1\| + \lambda_{2,2}\|\mathbf{\Lambda}_2\| < 1. \end{aligned} \tag{4.26}$$

By solving the feasibility of two LMIs that defined in (4.20), the result shows that $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$ share a CQLF. Based on Theorem 4.4.2, there exists a CQLF for $\mathbf{F}_1$ and $\mathbf{F}_2$ with any choice of $\lambda_{1,1}, \lambda_{1,2}, \lambda_{2,1}, \lambda_{2,2}$. Therefore, the switched system composed with $\mathbf{\Sigma}_{\mathbf{F}_1}$ and $\mathbf{\Sigma}_{\mathbf{F}_2}$ is asymptotically stable under arbitrary switching signal.

The next step is to study the boundedness of $\mathbf{u}_k$ (i.e., $\mathbf{x}_k$ of the original system). The boundedness of $\mathbf{x}_k$ can be checked by the existence of CQLF between $\mathbf{A}_1$ and $\mathbf{A}_2$. With a similar LMI condition, it shows that the original system is asymptotically stable. Therefore, the bias dynamics in the mode-based Kalman filter should be BIBO stable with upper bounds derived in (4.21).

Figure 4.5 and figure 4.6 are the experiment results over $N = 5000$ Monte-Carlo Simulation for two different switching signals. For each switching signal, two different probabilities of mode-mismatch error $\lambda_{1,2}$ and $\lambda_{2,1}$ were considered. In both figure 4.5 and figure 4.6, we plot the theoretical bias performance in line with squares. The theoretical bias is obtained via equation (4.16). The actual bias dynamics (difference of $\mathbb{E}(\hat{\mathbf{x}}_k)$ and $\mathbb{E}(\mathbf{x}_k)$) from Monte-Carlo simulation is presented using dashed line with triangles. Since we have verified that the bias evolution should always converge with any switching signal, all the above experiments

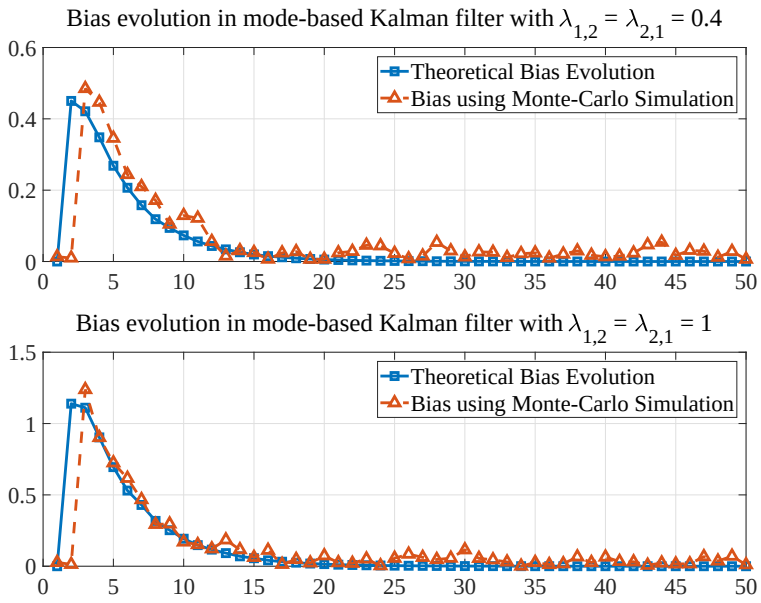


Figure 4.5: *Bias in mode-based Kalman filter using Monte-Carlo simulation and theoretical bias evolution for switching signal 1*

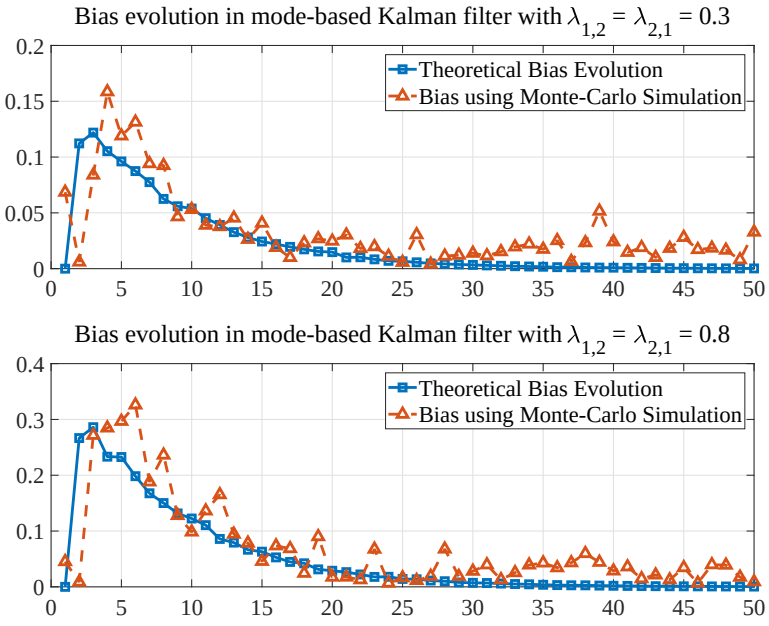


Figure 4.6: *Bias in mode-based Kalman filter using Monte-Carlo simulation and theoretical bias evolution for switching signal 2*

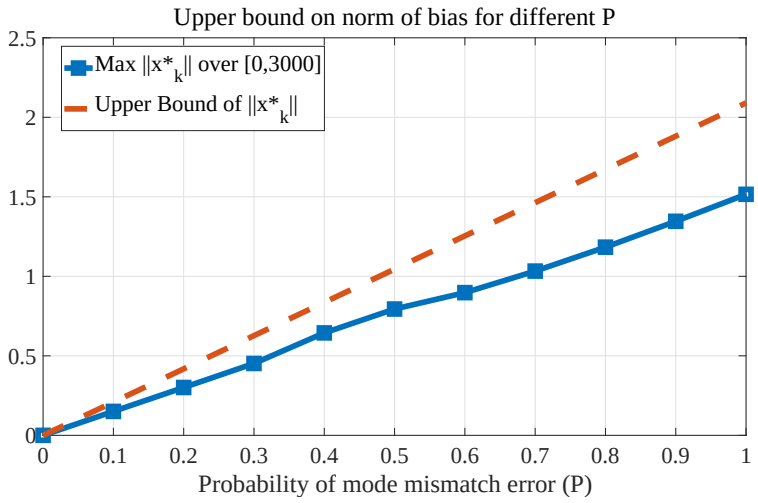


Figure 4.7: *Upper bound of* $\|\mathbf{x}_k^*\|$ *given mode mismatch probability* $\mathcal{P}$

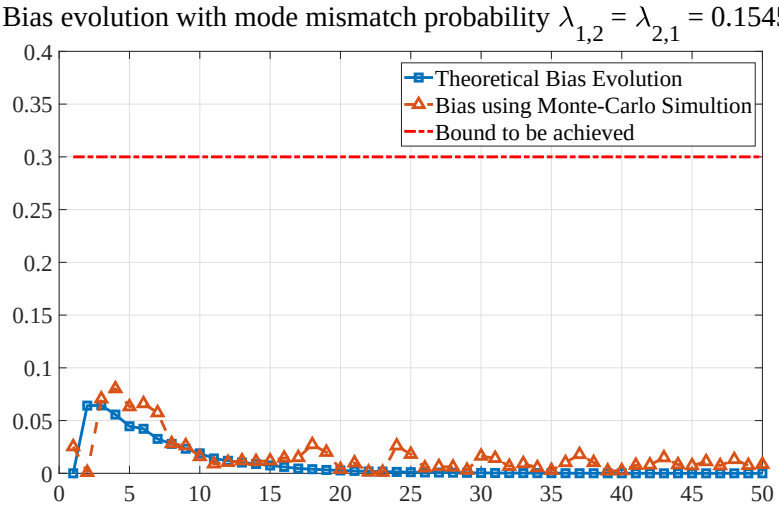


Figure 4.8: *To achieve an upper bound of bias* $\mathcal{B} = 0.3$*, the bias evolution with probability of mode mismatch as* $\lambda_{1,2} = \lambda_{2,1} = 0.154554$

also validate this result.

In Figure 4.7, the line with squares shows the maximum value for norm of bias over Monte-Carlo simulation given that probability of mode mismatch is $\mathcal{P}$. The dashed line is the upper bound calculated using Theorem 4.4.3. In Figure 4.8, we seek to address question (2) proposed in the last section. That is, we want to achieve a certain upper bound $\mathcal{B} = 0.3$ for the bias dynamics. By solving equation (4.25), the maximum probability of mode mismatch is $\lambda_{1,2} = \lambda_{2,1} = 0.154554$. Figure 4.8 shows the actual and theoretical bias evolution with

mode mismatch error $\lambda_{1,2} = \lambda_{2,1} = 0.154554$. We can conclude that the target bound has been achieved.

4.5.3 Case Study: Smart Grid

A classic example of a cyber-physical system that can be modeled in the SHS framework is a smart grid. We consider a toy smart grid set up inspired by [7, 76]. The system consists of three components - main distribution grid, local power network, and electrical loads. The discrete status for each component is:

- **Local power network** (L) - On: 1, Failure mode: 0;
- **Distribution grid** (G) - Connected: 1, Disconnected: 0;
- **Electrical loads** (D) - Connected: 1, Disconnected: 0.

The corresponding power generation and power consumption dynamics are given below:

- **Grid power**: If the micro grid is connected to the main electricity grid ($G = 1$), the grid power P_G has the following dynamics: $\dot{P}_G = k_G P_G + \sigma_G dW$, where k_G is a proportional coefficient and σ_G is a variation parameter [7]. If $G = 0$, both k_G and σ_G are close to 0. dW denotes Wiener process.
- **Electrical loads**: Electrical loads can be modeled via stochastic differential equation. We use Uhlenbeck-Ornstein model to describe electricity loads [76]. Let $\dot{P}_D = \alpha(m - P_D)dt + \sigma_D dW$. Here, we assume $m = 0$. α represents a tracking coefficient. σ_D is a variation coefficient, and dW denotes Wiener process.

Therefore, the continuous state in this smart grid system can be defined as $\mathbf{x} = [P_G, P_D]'$ with corresponding state equation as:

$$\begin{bmatrix} \dot{P}_G \\ \dot{P}_D \end{bmatrix} = \begin{bmatrix} k_G & 0 \\ 0 & -\alpha \end{bmatrix} \begin{bmatrix} P_G \\ P_D \end{bmatrix} + \begin{bmatrix} \sigma_G & 0 \\ 0 & \sigma_D \end{bmatrix} \frac{dW}{dt}.$$

Table 4.1: *Discrete Status and Continuous Dynamics Parameters*

Status	L	G	D	q	k_G	α	σ_G	σ_D
Failure Mode	0	0	0	1	0.1	0.1	0.1	0.1
Grid Connected	1	1	0	2	3	0.5	0.8	0.8
	1	1	1	3	3	0.49	1.5	1

By discretizing the state space with a sampling period of τ, we get a discrete-time SHS:

$$\mathbf{x}_k = \mathbf{A}_{q_k}\mathbf{x}_{k-1} + \mathbf{B}_{q_k}\mathbf{w}_k, \tag{4.27}$$

where

$$\mathbf{A}_{q_k} = \begin{bmatrix} e^{k_G\tau} & 0 \\ 0 & e^{-\alpha\tau} \end{bmatrix}, \tag{4.28}$$

and

$$\mathbf{B}_{q_k} = \begin{bmatrix} k_G & 0 \\ 0 & -\alpha \end{bmatrix}^{-1} (\mathbf{A}_{q_k} - \mathbf{I}) \begin{bmatrix} \sigma_G & 0 \\ 0 & \sigma_D \end{bmatrix}. \tag{4.29}$$

Here, the index k corresponds to the time instant $k\tau$. The discrete state space is defined by combination of different status of L, G, D. Consequently, the value of parameters $k_G, \alpha, \sigma_G, \sigma_D$ are determined by different discrete states. The measurement equation corresponds to

$$\mathbf{y}_k = \mathbf{C}_{q_k}\mathbf{x}_k + \mathbf{v}_k. \tag{4.30}$$

For this case study, the status of components L, G and D and the grid parameters are defined in Table 4.1. Based on system settings, k_G, α, σ_G and σ_D completely determine the system matrices $\mathbf{A}_q$ and $\mathbf{B}_q$. Let $\mathbf{C}_q = \mathbf{I}$ for all modes. Define the noise as $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ with $\mathbf{Q} = 2 \times \mathbf{I}$ and $\mathbf{R} = \mathbf{I}$. With this system setting, we get $\|\mathbf{\Lambda}_1\| = 0.9817$, $\|\mathbf{\Lambda}_2\| = 0.8837$ and $\|\mathbf{\Lambda}_3\| = 0.8611$. Therefore, similar as (4.26), we have $\rho(\mathbf{F}_1), \rho(\mathbf{F}_2), \rho(\mathbf{F}_3) < 1$ for all choices of $\lambda_{\mathrm{i,t}}$. The next step is to solve the LMI conditions on $\mathbf{\Lambda}_1$, $\mathbf{\Lambda}_2$ and $\mathbf{\Lambda}_3$ and the results shows that $\mathbf{\Lambda}_1$, $\mathbf{\Lambda}_2$ and $\mathbf{\Lambda}_3$ share a CQLF. Based on Theorem

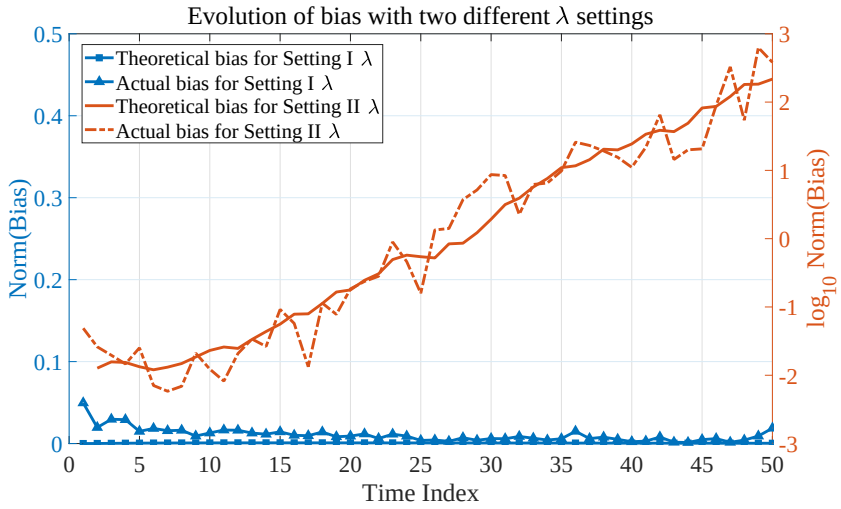


Figure 4.9: *Monte-Carlo simulation for the smart grid system with λ in stable region and unstable region*

4.4.2, the switched system composed with $\mathbf{\Sigma}_{\mathbf{F}_1}$, $\mathbf{\Sigma}_{\mathbf{F}_2}$ and $\mathbf{\Sigma}_{\mathbf{F}_3}$ is asymptotically stable under arbitrary switching signal. In order to check the boundedness of input $\mathbf{u}_k$, we solve for the CQLF for $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$. In this case, the result reveals that the original SHS is not stable (falls into scenario 1). Therefore, we are able to use Algorithm 4 to derive the stable region of each $\lambda_{\mathrm{i},q}$. In this system, the unstable component is: $\mathcal{I} = \{1\}$, i.e., only the first element is unstable. Based on Algorithm 4, we need to calculate $\mathbf{T}_{\mathrm{i},q}$ and find out the corresponding elements on column 1 of each matrix. We get:

$$\mathbf{T}_{1,2} = \begin{bmatrix} -0.3303 & 0 \\ 0 & 0.0385 \end{bmatrix}, \mathbf{T}_{1,3} = \begin{bmatrix} -0.3303 & 0 \\ 0 & 0.0375 \end{bmatrix}, \mathbf{T}_{2,1} = \begin{bmatrix} 0.1827 & 0 \\ 0 & -0.0361 \end{bmatrix},$$
$$\mathbf{T}_{2,3} = \begin{bmatrix} 0 & 0 \\ 0 & -0.0088 \end{bmatrix}, \mathbf{T}_{3,1} = \begin{bmatrix} 0.1745 & 0 \\ 0 & -0.0342 \end{bmatrix}, \mathbf{T}_{3,2} = \begin{bmatrix} 0 & 0 \\ 0 & 0.0086 \end{bmatrix}.$$

It can be observed that the first column in $\mathbf{T}_{2,3}$ and $\mathbf{T}_{3,2}$ are 0. Therefore, the mode-based Kalman filter can be tolerant on mode mismatch error between mode 2 and mode 3. The

stable region for each λ is:

$$\lambda_{1,2} = \lambda_{1,3} = \lambda_{2,1} = \lambda_{3,1} = 0$$

$$0 \leq \lambda_{2,3}, \lambda_{3,2}, \lambda_{1,1}, \lambda_{2,2}, \lambda_{3,3} \leq 1.$$

Note that the condition that $\sum_{i=1}^{3} \lambda_{i,q} = 1$ should also hold for every q. Figure 4.9 shows a Monte-Carlo simulation for two different λ settings. Setting I we use $\lambda_{2,1} = \lambda_{3,1} = \lambda_{1,2} = \lambda_{1,3} = 0, \lambda_{3,2} = 0.4, \lambda_{2,3} = 0.7$ where all the λs are within the stable region. The simulation results for Setting I are shown in lines with squares and triangles with left y-axis. Specifically, the line with squares is the theoretical bias derived using the bias evolution equation (4.17) while the line with triangles shows the bias in a mode-based Kalman filter via Monte-Carlo simulation. We can conclude that when all the λs are in stable region, the bias of the mode-based Kalman filter is convergent and bounded. For Setting II we use $\lambda_{2,1} = \lambda_{3,1} = \lambda_{1,2} = 0.1, \lambda_{1,3} = 0, \lambda_{3,2} = 0.3, \lambda_{2,3} = 0.2$ in which $\lambda_{2,1}, \lambda_{3,1}, \lambda_{1,2}$ are outside the stable region. The solid line and the dashed line with right y-axis present the results for theoretical bias and actual bias generated in a mode-based Kalman filter via Monte-Carlo simulation. Note that the y-axis on the right is $\log(\|\mathbf{x}_k^*\|)$ since the actual $\|\mathbf{x}_k^*\|$ explodes rapidly. As this system does not have tolerance between mode 1,2 and mode 1,3, even a small probability of error (i.e., 0.1 in this case) will result in rapid explosion in the bias dynamics.

4.6 Summary

In this chapter, we consider the open research problem of quantifying the impact of mode-mismatch errors on the performance of a mode-based Kalman filter. Specifically, the mode mismatch errors are captured by i.i.d. Bernoulli random variables. The problem itself is appropriate to describe network topology errors in a smart grid or other cyber-physical systems. We first study the bias performance for SHS with i.i.d. Bernoullian mode switches and then generalize the analysis to SHS with arbitrary mode modelling. The main technique proposed involves modeling the bias dynamics in the Kalman filter as a transformed switched

system. Abstracting the discrete state transitions as arbitrary switching signals not only broaden the application space but also provides us tools from switched system stability analysis to study the statistical convergence of the bias.

Chapter 5

Time Correlated (Markovian Distributed) Mode Mismatch Errors

In the previous chapter, we derived the bias dynamics in a mode-based Kalman filter for a generalized model of stochastic hybrid system (SHS) and studied conditions such that the bias is statistically convergent. Specifically, the bias results from an independent and identically distributed (i.i.d.) Bernoulli distributed mode mismatches. In this chapter, we consider the situation where mode mismatch errors are correlated across time and we derive, for the first time, a sufficient and necessary condition on the statistical convergence of the bias.

5.1 Introduction

SHS have been proven to be a reliable model for many cyber-physical systems (CPS) that experience interaction between continuous dynamics and discrete modes. Markov jump linear systems (MJLS) is a particular subset of SHS in which the discrete mode switches are modeled as a Markov chain and continuous states evolve linearly. MJLS has attracted significant attention in the research community due to its analytical tractability as well as applicability to practical systems such as power systems [8] and networked control system

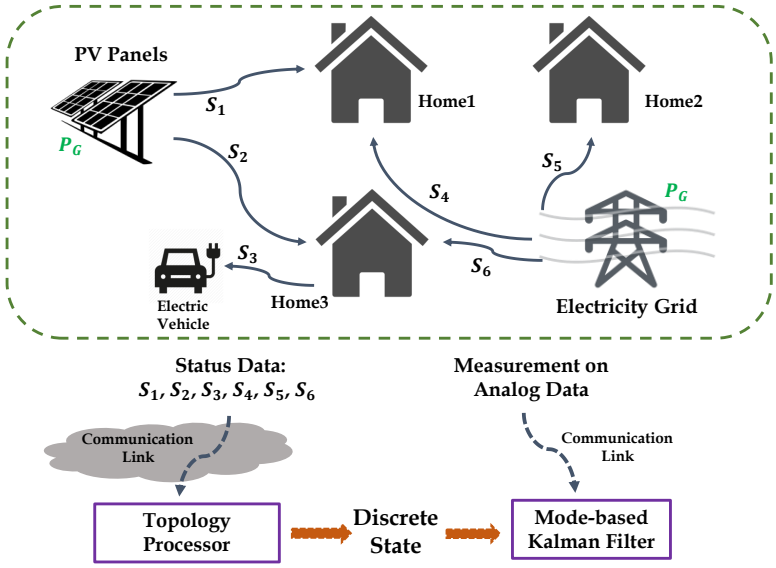


Figure 5.1: *Motivating example: communication link impairments in a smart grid*

(NCS) [9], etc. While the continuous states along with the discrete modes fully describe the MJLS, the continuous states may not be directly accessible in many applications. In such cases, state estimation becomes critical for both situational awareness and implementation of control actions. With many state estimation strategies have been proposed for MJLS, Kalman filter based strategies still dominate the area [18, 20, 22].

Consider a toy example of a smart grid as shown in Figure 5.1. The system consists of a bank of PV panels, three home loads, one electric vehicle and the main grid. There are 'switches' S_1 to S_6 that can be turned ON and OFF depending on the weather, power demands etc. There are sensors that communicate the observation on status of switches S_1 to S_6 to a central topology processor and the status of all switches determines the topology of the network. One topology corresponds to one discrete state (mode) and the evolution of continuous states are based on the mode. Additionally, there are sensors that communicate analog/continuous measurements (such as power consumption, bus voltage, etc.) to a central estimator (mode-based Kalman filter) to estimate the true continuous state. In this chapter, we consider the situation wherein impairments in communication link between switch status sensors and topology processor result in erroneous information about switch status/network

topology. Network topology errors (mode mismatch) in smart grids as discussed in [21] is a critical problem for any networked system. Since continuous state evolves differently for different modes, errors in the network topology (mode) will impact the continuous state estimate. Based on this example, we take a fresh perspective in analyzing the performance of a mode-based Kalman filter for MJLS. Specifically, we consider the errors that due to problems in sensing the discrete states or due to communication network impairments in a CPS setting. The mode mismatches are modeled via a Markov chain as it effectively captures communication link failures and cyber-attacks in many CPS (e.g. topology errors in a smart grid system [21], packet drops/damage in NCS [41, 77]). For this situation, we can treat the (possible) inaccurate discrete mode as the true state and implement an optimal mode-based Kalman filter without suffering from increased complexity. The main purpose of this chapter is to investigate conditions under which the errors in a mode-based Kalman filter are still bounded in the presence of discrete mode mismatch. We first consider an MJLS with two modes derive sufficient and necessary conditions (based on the results from Schur stability of a matrix polytope) under which the bias dynamics are statistically convergent. We then extend the analysis to MJLSs with arbitrary numbers of discrete modes. This extension is not trivial as the results for MJLS with two modes are built on an earlier mathematical result [78] that is applicable to only two modes. For this case, we propose to model the mean of bias dynamics as an auxiliary linear system. The system matrix of this linear system is determined by a polytope of matrices with each vertex matrix related to the original MJLS system matrices. By mapping the matrix polytope to an interval matrix, and by leveraging results in Schur stability analysis for an interval matrix, we derive sufficient conditions on mode mismatch probabilities under which the bias resulting from mode mismatches is statistically convergent. This new fundamental result provides guidance on design of estimation strategies for MJLS and sheds light on their resilience in the presence of mode mismatch errors.

5.2 Preliminaries

We consider a discrete-time MJLS with a finite discrete state space denoted as $\mathcal{Q} = \{q_1, \cdots, q_d\}$. The discrete state transitions are modeled via a time-homogeneous Markov chain. Let $\delta_k : k \to \mathcal{Q}$ denote the discrete state. Then, the transition probability corresponds to,

$$\begin{aligned} &\mathbb{P}(\delta_k = q_j | \delta_{k-1} = q_i, \cdots, \delta_0 = q_1) \\ &\qquad = \mathbb{P}(\delta_k = q_j | \delta_{k-1} = q_i) = \pi_{ij}, \end{aligned} \tag{5.1}$$

and the transition matrix is defined as $\mathbf{P_s} = [\pi_{ij}] \in \mathbb{R}^{d \times d}$. The continuous state $\mathbf{x}_k \in \mathbb{R}^n$ and measurements $\mathbf{y}_k \in \mathbb{R}^m$ evolve as:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}_{\delta_k} \mathbf{x}_{k-1} + \mathbf{B}_{\delta_k} \mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{C}_{\delta_k} \mathbf{x}_k + \mathbf{v}_k. \end{aligned} \tag{5.2}$$

Here, $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ are both independent Gaussian noise capturing model and measurement uncertainties. Given $\delta_k = q_i \in \mathcal{Q}$, $\mathbf{A}_{\delta_k}$, $\mathbf{B}_{\delta_k}$ and $\mathbf{C}_{\delta_k}$ are $n \times n$, $n \times p$ and $m \times n$ matrices, respectively. We assume the initial continuous state is Gaussian with mean $\boldsymbol{\mu}_0$ and covariance $\boldsymbol{\Sigma}_0$ and the initial discrete mode is q_1.

We consider correlated mode mismatch errors modeled by a time-homogeneous Markov chain. The correlation is across time and can result due to multiple practical constraints (e.g., correlated fading of wireless channel). In this case, the probability of mode mismatch at the current instant only depends on the quality of communication channel at previous time instant. Let $\theta_k : k \to \{0, 1\}$ be a variable denoting the occurrence of mode mismatch. That is, $\theta_k = 0$ implies no mode mismatch and $\theta_k = 1$ otherwise. Note that we assume both δ_k and θ_k are irreducible and aperiodic. The transition matrix for the Markov chain θ_k is:

$$\mathbf{P_m} = \begin{bmatrix} \lambda_0 & 1 - \lambda_0 \\ 1 - \lambda_1 & \lambda_1 \end{bmatrix}. \tag{5.3}$$

Algorithm 5 Mode-based Kalman filter

function ESTIMATION UPDATE($\boldsymbol{\mu}_0, \mathbf{M}_{0|0}, \mathbf{Q}, \mathbf{R}, \boldsymbol{\gamma}_k^s, \mathbf{y}_k^s$)
  $\mathbf{x}_{0|0} = \boldsymbol{\mu}_0, \mathbf{M}_{0|0} = \boldsymbol{\Sigma}_0$
  $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$
  $\boldsymbol{\gamma}_k^s = (\gamma_1, \cdots, \gamma_k)$
  for $i = 1 : k$ **do**
    $\mathbf{x}_{i|i-1} = \mathbf{A}_{\gamma_i}\mathbf{x}_{i-1|i-1}$
    $\mathbf{M}_{i|i-1} = \mathbf{A}_{\gamma_i}\mathbf{M}_{i-1|i-1}\mathbf{A}'_{\gamma_i} + \mathbf{B}_{\gamma_i}\mathbf{Q}\mathbf{B}'_{\gamma_i}$
    $\mathbf{K}_{\gamma_i,i} = \mathbf{M}_{i|i-1}\mathbf{C}'_{\gamma_i}(\mathbf{C}_{\gamma_i}\mathbf{M}_{i|i-1}\mathbf{C}'_{\gamma_i} + \mathbf{R})^{-1}$
    $\mathbf{x}_{i|i} = \mathbf{x}_{i|i-1} + \mathbf{K}_{i,\gamma_i}(\mathbf{y}_i - \mathbf{C}_{\gamma_i}\mathbf{x}_{i|i-1})$
    $\mathbf{M}_{i|i} = (\mathbf{I} - \mathbf{K}_{\gamma_i,i}\mathbf{C}_{\gamma_i})\mathbf{M}_{i|i-1}$
  end for
  return $\mathbf{x}_{k|k}$
end function

As the continuous states are not directly observable in system (5.2), a mode-based Kalman filter can be used to estimate the continuous state $\mathbf{x}_k$ based on: (1) the measurements, and (2) discrete states up to time k. If there is a mode mismatch error, the known mode will be different than the actual mode. Let γ_k denote the known (or estimated) mode at time k. Based on definition of θ_k, we have $\gamma_k = \delta_k$ if and only if $\theta_k = 0$. Denote the measurement sequence and mode sequence up to time k as $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$ and $\boldsymbol{\gamma}_k^s = (\gamma_1, \cdots, \gamma_k)$, respectively. The mode-based Kalman filter equations for MJLS (5.2) are given in Algorithm 5. For a hybrid system, the estimate $\hat{\mathbf{x}}_k = \mathbf{x}_{k|k}$ is unbiased only if $\boldsymbol{\gamma}_k^s = \boldsymbol{\delta}_k^s$ where $\boldsymbol{\delta}_k^s = (\delta_1, \cdots, \delta_k)$. It is worth pointing out that even if there is mismatch between $\boldsymbol{\gamma}_k^s$ and $\boldsymbol{\delta}_k^s$, the error covariance matrix still remains bounded as proved in the previous chapter. Therefore, we only focus on the bias term. Since both $\hat{\mathbf{x}}_k$ and $\mathbf{x}_k$ are random variables in an MJLS, we define the bias to be the difference between means of estimator and the true state, i.e.,

$$\mathbf{e}_k = \mathbb{E}(\hat{\mathbf{x}}_k) - \mathbb{E}(\mathbf{x}_k).$$

In other words, we capture the difference between $\hat{\mathbf{x}}_k$ and $\mathbf{x}_k$ in a mean sense via $\mathbf{e}_k$. According to Kalman filter equations and the derivation in the previous chapter, the dynamics

of $\mathbf{e}_k$ corresponds to:

$$\begin{aligned}\mathbf{e}_k = &(\mathbf{A}_{\gamma_k} - \mathbf{K}_{\gamma_k}\mathbf{C}_{\gamma_k}\mathbf{A}_{\gamma_k})\mathbf{e}_{k-1} \\ &+ \theta_k(\mathbf{A}_{\gamma_k} - \mathbf{K}_{\gamma_k}\mathbf{C}_{\gamma_k}\mathbf{A}_{\gamma_k} + \mathbf{K}_{\gamma_k}\mathbf{C}_{\delta_k}\mathbf{A}_{\delta_k} - \mathbf{A}_{\delta_k})\mathbb{E}(\mathbf{x}_{k-1}).\end{aligned}$$

Here, $\mathbf{K}_{\gamma_k}$ is the steady Kalman gain that corresponds to the mode $\gamma_k \in \mathcal{Q}$. Since we assume that $\mathbf{Q}$ and $\mathbf{R}$ are constants for all modes, the Kalman gain will converge to the corresponding steady Kalman gain $\mathbf{K}_{\gamma_k}$ quickly [63]. For a hybrid system, it is reasonable to approximate the time-variant Kalman gain using the steady Kalman gain in practice as the mode switches are infrequent relative to the evolution of the continuous states. This assumption has also been used in sub-optimal control problems for decades [20, 44]. To simplify the presentation of this equation, we define

$$\begin{aligned}\mathbf{\Lambda}_{\gamma_k} &= \mathbf{A}_{\gamma_k} - \mathbf{K}_{\gamma_k}\mathbf{C}_{\gamma_k}\mathbf{A}_{\gamma_k}, \\ \mathbf{\Gamma}_{\gamma_k,\delta_k} &= \mathbf{A}_{\delta_k} - \mathbf{K}_{\gamma_k}\mathbf{C}_{\delta_k}\mathbf{A}_{\delta_k}.\end{aligned}$$

Then, the $\mathbf{e}_k$ dynamics evolves as:

$$\mathbf{e}_k = \mathbf{\Lambda}_{\gamma_k}\mathbf{e}_{k-1} + \theta_k(\mathbf{\Lambda}_{\gamma_k} - \mathbf{\Gamma}_{\gamma_k,\delta_k})\mathbb{E}(\mathbf{x}_{k-1}). \tag{5.4}$$

As the evolution of $\mathbf{e}_k$ depends on γ_k, δ_k and θ_k, the process $\{\mathbf{e}_k\}_{k=0}^{\infty}$ is a stochastic process. We are interested in convergence in mean, i.e., $\lim_{k\to\infty} \mathbb{E}(\mathbf{e}_k) < \infty$ [39, 41]. Analyzing the statistical convergence of $\mathbf{e}_k$ becomes mathematically challenging because of the dependency between γ_k and θ_k, δ_k. In the following, we first consider an MJLS with two modes as a starting point and derive a necessary and sufficient conditions such that the bias is statistically convergent. Finally, we address the same problem for MJLSs with arbitrary numbers of modes.

5.3 MJLS with two discrete states

We first attempt to address this challenge by focusing on an MJLS with two modes as a starting point, i.e., $\mathcal{Q} = \{q_1, q_2\}$. The two-mode assumption guarantees that when mode mismatch happens (i.e., $\theta_k = 1$), there is only one error value that γ_k can take. From the definition, we can rewrite γ_k as a function of δ_k and θ_k:

$$\gamma_k = (1 - \theta_k)\delta_k + \theta_k \bar{\delta}_k, \tag{5.5}$$

where $\bar{\delta}_k$ is the opposite value of δ_k, i.e., $\{\delta_k\} \bigcup \{\bar{\delta}_k\} = \mathcal{Q}$. Assume the initial distributions for δ_0 and θ_0 are:

$$\boldsymbol{\omega}_\delta = \begin{bmatrix} \mathbb{P}(\delta_0 = q_1) \\ \mathbb{P}(\delta_0 = q_2) \end{bmatrix}', \boldsymbol{\omega}_\theta = \begin{bmatrix} \mathbb{P}(\theta_0 = 0) \\ \mathbb{P}(\theta_0 = 1) \end{bmatrix}'.$$

As a property of Markov chain, we get the probability distributions for δ_k and θ_k as:

$$\begin{bmatrix} \mathbb{P}(\delta_k = q_1) \\ \mathbb{P}(\delta_k = q_2) \end{bmatrix}' = \boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k, \quad \begin{bmatrix} \mathbb{P}(\theta_k = 0) \\ \mathbb{P}(\theta_k = 1) \end{bmatrix}' = \boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k.$$

From the definition, γ_k takes value on q_1 and q_2 based on the value of θ_k and δ_k. Specifically,

$$\gamma_k = \begin{cases} q_1, & \text{if } \theta_k = 0 \text{ and } \delta_k = q_1 \text{ or } \theta_k = 1 \text{ and } \delta_k = q_2; \\ q_2, & \text{if } \theta_k = 0 \text{ and } \delta_k = q_2 \text{ or } \theta_k = 1 \text{ and } \delta_k = q_1. \end{cases}$$

Because θ_k and δ_k are independent variables, we can write the probability measure for γ_k as:

$$\mathbb{P}(\gamma_k = q_1) = \begin{bmatrix} \mathbb{P}(\theta_k = 0) \\ \mathbb{P}(\theta_k = 1) \end{bmatrix}' \begin{bmatrix} \mathbb{P}(\delta_k = q_1) \\ \mathbb{P}(\delta_k = q_2) \end{bmatrix} = \boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k (\boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k)', \tag{5.6}$$

$$\mathbb{P}(\gamma_k = q_2) = 1 - \boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k (\boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k)'. \tag{5.7}$$

Based on the values of γ_k, θ_k and δ_k, the bias dynamics in equation (5.4) can be grouped into the following four scenarios:

Scenario 1 ($\theta_k = 0$ and $\delta_k = \boldsymbol{q}_1$):

$$\mathbf{e}_k = \mathbf{\Lambda}_{q_1}\mathbf{e}_{k-1}.$$

Scenario 2 ($\theta_k = 0$ and $\delta_k = \boldsymbol{q}_2$):

$$\mathbf{e}_k = \mathbf{\Lambda}_{q_2}\mathbf{e}_{k-1}.$$

Scenario 3 ($\theta_k = 1$ and $\delta_k = \boldsymbol{q}_2$):

$$\mathbf{e}_k = \mathbf{\Lambda}_{q_1}\mathbf{e}_{k-1} + (\mathbf{\Lambda}_{q_1} - \mathbf{\Gamma}_{q_1,q_2})\mathbb{E}(\mathbf{x}_{k-1}).$$

Scenario 4 ($\theta_k = 1$ and $\delta_k = \boldsymbol{q}_1$):

$$\mathbf{e}_k = \mathbf{\Lambda}_{q_2}\mathbf{e}_{k-1} + (\mathbf{\Lambda}_{q_2} - \mathbf{\Gamma}_{q_2,q_1})\mathbb{E}(\mathbf{x}_{k-1}).$$

Recall that we consider analyzing the convergence of $\mathbf{e}_k$ in a mean sense, i.e., $\lim_{k\to\infty} \mathbb{E}(\mathbf{e}_k) < \infty$. According to the four scenarios and the probability measures derived in (5.6) and (5.7), we have:

$$
\begin{aligned}
&\mathbb{E}(\mathbf{e}_k) = \mathbb{E}(\mathbb{E}(\mathbf{e}_k|\mathbf{e}_{k-1})) \\
&= \Big[\mathbb{P}(\gamma_k = q_1)\mathbf{\Lambda}_{q_1} + \mathbb{P}(\gamma_k = q_2)\mathbf{\Lambda}_{q_2}\Big]\mathbb{E}(\mathbf{e}_{k-1}) \\
&\quad + \mathbb{P}(\theta_k = 1)\mathbb{P}(\delta_k = q_1)(\mathbf{\Lambda}_{q_2} - \mathbf{\Gamma}_{q_2,q_1})\mathbb{E}(\mathbf{x}_{k-1}) \\
&\quad + \mathbb{P}(\theta_k = 1)\mathbb{P}(\delta_k = q_2)(\mathbf{\Lambda}_{q_1} - \mathbf{\Gamma}_{q_1,q_2})\mathbb{E}(\mathbf{x}_{k-1}) \\
&= \Big[\boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k (\boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k)' \mathbf{\Lambda}_{q_1} \\
&\qquad + \Big(1 - \boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k (\boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k)'\Big)\mathbf{\Lambda}_{q_2}\Big]\mathbb{E}(\mathbf{e}_{k-1}) \\
&\quad + \Big[\boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k \mathbf{u}_2 \boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k \mathbf{u}_1 (\mathbf{\Lambda}_{q_2} - \mathbf{\Gamma}_{q_2,q_1}) \\
&\qquad + \boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k \mathbf{u}_2 \boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k \mathbf{u}_2 (\mathbf{\Lambda}_{q_1} - \mathbf{\Gamma}_{q_1,q_2})\Big]\mathbb{E}(\mathbf{x}_{k-1})
\end{aligned}
\tag{5.8}
$$

where the outer expectation is taken over $\mathbf{e}_{k-1}$ and the inner expectation is taken over θ_k and δ_k. In the following, we will focus on deriving conditions under which the dynamics in (5.8) converges. As shown in (5.8), the process of $\mathbb{E}(\mathbf{e}_k)$ evolves based on a dynamic updating equation with a stochastic input signal related to $\mathbb{E}(\mathbf{x}_{k-1})$. To determine the conditions for convergence of (5.8), we first consider the convergence of the first term that is related to $\mathbb{E}(\mathbf{e}_{k-1})$ and then analyze the boundedness on the second (input) term.

5.3.1 Convergence on the Autonomous System

The first term of equation (5.8) is a deterministic update equation on $\mathbb{E}(\mathbf{e}_{k-1})$. Define an auxiliary process $\mathbf{z}_k$ that evolves as:

$$
\mathbf{z}_k = \Big[\boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k (\boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k)' \mathbf{\Lambda}_{q_1} + \Big(1 - \boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k (\boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k)'\Big)\mathbf{\Lambda}_{q_2}\Big]\mathbf{z}_{k-1}. \tag{5.9}
$$

To address convergence of $\mathbb{E}(\mathbf{e}_k)$, we first consider $\lim_{k\to\infty} \mathbf{z}_k$. Since there are stationary distributions for both δ_k and θ_k, we denote them as $\boldsymbol{\omega}_{\delta^*}$ and $\boldsymbol{\omega}_{\theta^*}$ with

$$\boldsymbol{\omega}_{\delta^*} = \lim_{k\to\infty} \boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k, \quad \boldsymbol{\omega}_{\theta^*} = \lim_{k\to\infty} \boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k. \tag{5.10}$$

Let $c_k = \boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k (\boldsymbol{\omega}_\delta {\mathbf{P}_\mathbf{s}}^k)'$, then we have

$$\lim_{k\to\infty} c_k = c^* = \boldsymbol{\omega}_{\theta^*} \boldsymbol{\omega}'_{\delta^*}.$$

It is easy to show that $0 \leq c^* \leq 1$. As the Markov chain δ_k only depends on the system model, the stationary distributions can be computed apriori. Thus, we can treat $\boldsymbol{\omega}_{\delta^*}$ as a fixed value and only focus on the term $\boldsymbol{\omega}_{\theta^*}$. The stationary distribution $\boldsymbol{\omega}_{\theta^*}$ relies on $\mathbf{P}_\mathbf{m}$ that contains the parameters λ_0 and λ_1. Therefore, the main purpose here is to find regions for λ_0 and λ_1 such that $\lim_{k\to\infty} \mathbf{z}_k < \infty$ holds. We first consider the following spectral radius problem as it is closely related to the convergence of $\mathbf{z}_k$. That is, find conditions on $c \in [0, 1]$ such that:

$$\rho(c\boldsymbol{\Lambda}_{q_1} + (1-c)\boldsymbol{\Lambda}_{q_2}) < 1.$$

Before we introduce our main result, the following supporting lemmas that assist in the proof of the main theorem need to be stated. Proofs for Lemmas 5.3.1 and 5.3.2 can be found in [78].

Lemma 5.3.1. *Let* $\mathbf{H}(r) = r^2\mathbf{H}_2 + r\mathbf{H}_1 + \mathbf{H}_0$. *Suppose* $\mathbf{H}_0$ *is nonsingular, then the maximum range* $[r_{min}, r_{max}]$ *for all* $\mathbf{H}(r)$ *to be nonsingular is given by:*

$$r_{min} = \frac{1}{\nu^-_{min}(\mathbf{G})}, r_{max} = \frac{1}{\nu^+_{max}(\mathbf{G})},$$

with

$$\mathbf{G} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{H}_0^{-1}\mathbf{H}_2 & -\mathbf{H}_0^{-1}\mathbf{H}_1 \end{bmatrix}$$

where $\nu_{min}^{-}(\mathbf{G})$ *denotes the minimum negative real eigenvalue of* $\mathbf{G}$ *and* $\nu_{max}^{+}(\mathbf{G})$ *denotes the maximum positive real eigenvalue of* $\mathbf{G}$.

Lemma 5.3.2. *If matrix* $\mathbf{\Lambda}_{q_2}$ *is Schur stable, then*

$$\mathbf{H}_0 = \mathbf{I} \otimes \mathbf{I} - \mathbf{\Lambda}_{q_2} \otimes \overline{\mathbf{\Lambda}_{q_2}}$$

is nonsingular.

Lemma 5.3.3. *The matrix* $\mathbf{\Lambda}(r) = (1-r)\mathbf{\Lambda}_{q_2} + r\mathbf{\Lambda}_{q_1}$ *is Schur stable if and only if* $\mathbf{\Lambda}_{q_2}$ *is strictly Schur and* $\mathbf{I} \otimes \mathbf{I} - \mathbf{\Lambda}(r) \otimes \overline{\mathbf{\Lambda}(r)}$ *is nonsingular.*

We first define the following auxiliary matrices to assist our analysis:

$$\begin{aligned} \mathbf{H}_0 &= \mathbf{I} \otimes \mathbf{I} - \mathbf{\Lambda}_{q_2} \otimes \overline{\mathbf{\Lambda}_{q_2}}, \\ \mathbf{H}_1 &= 2\mathbf{\Lambda}_{q_2} \otimes \overline{\mathbf{\Lambda}_{q_2}} - \mathbf{\Lambda}_{q_2} \otimes \overline{\mathbf{\Lambda}_{q_1}} - \mathbf{\Lambda}_{q_1} \otimes \overline{\mathbf{\Lambda}_{q_2}}, \\ \mathbf{H}_2 &= \mathbf{\Lambda}_{q_2} \otimes \overline{\mathbf{\Lambda}_{q_1}} + \mathbf{\Lambda}_{q_1} \otimes \overline{\mathbf{\Lambda}_{q_2}} - \mathbf{\Lambda}_{q_2} \otimes \overline{\mathbf{\Lambda}_{q_2}} - \mathbf{\Lambda}_{q_1} \otimes \overline{\mathbf{\Lambda}_{q_1}}, \\ \mathbf{G} &= \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{H}_0^{-1}\mathbf{H}_2 & -\mathbf{H}_0^{-1}\mathbf{H}_1 \end{bmatrix}. \end{aligned} \tag{5.11}$$

Note that matrix $\mathbf{G}$ is defined only for the case that $\mathbf{H}_0$ is nonsingular.

Theorem 5.3.1. *The matrix* $c\mathbf{\Lambda}_{q_1} + (1-c)\mathbf{\Lambda}_{q_2}$ *is Schur stable if and only if* $\mathbf{\Lambda}_{q_2}$ *is Schur and* c *is in the range of* $[\frac{1}{\nu_{min}^{-}(\mathbf{G})}, \frac{1}{\nu_{max}^{+}(\mathbf{G})}]$ *with the definition of* $\mathbf{H}_0, \mathbf{H}_1, \mathbf{H}_2$ *and* $\mathbf{G}$ *shown in (5.11).*

Proof. Since $\mathbf{\Lambda}_{q_2}$ is Schur stable, then $\mathbf{H}_0 = \mathbf{I} \otimes \mathbf{I} - \mathbf{\Lambda}_{q_2} \otimes \overline{\mathbf{\Lambda}_{q_2}}$ is nonsingular as a result of Lemma 5.3.2. According to Lemma 5.3.1, if $\mathbf{H}_0$ is nonsingular and c is in the range of

$[\frac{1}{\nu^-_{min}(\mathbf{G})}, \frac{1}{\nu^+_{max}(\mathbf{G})}]$, $\mathbf{H}(c) = c^2\mathbf{H}_2 + c\mathbf{H}_1 + \mathbf{H}_0$ is nonsingular. Given $\mathbf{\Lambda}_{q_2}$ is Schur, based on Lemma 5.3.3, the matrix

$$c\mathbf{\Lambda}_{q_1} + (1-c)\mathbf{\Lambda}_{q_2}$$

is Schur stable if $\mathbf{I}\otimes\mathbf{I} - \mathbf{\Lambda}(c)\otimes\overline{\mathbf{\Lambda}(c)}$ is nonsingular. Since

$$\begin{aligned}
\mathbf{I}\otimes\mathbf{I} - \mathbf{\Lambda}(c)\otimes\overline{\mathbf{\Lambda}(c)} =& \mathbf{I}\otimes\mathbf{I} - \big[(1-c)\mathbf{\Lambda}_{q_2} + c\mathbf{\Lambda}_{q_1}\big]\otimes\overline{\big[(1-c)\mathbf{\Lambda}_{q_2} + c\mathbf{\Lambda}_{q_1}\big]} \\
=& \mathbf{I}\otimes\mathbf{I} - (1-c)^2\mathbf{\Lambda}_{q_2}\otimes\overline{\mathbf{\Lambda}_{q_2}} - (1-c)c\mathbf{\Lambda}_{q_2}\otimes\overline{\mathbf{\Lambda}_{q_1}} \\
& - c(1-c)\mathbf{\Lambda}_{q_1}\otimes\overline{\mathbf{\Lambda}_{q_2}} - c^2\mathbf{\Lambda}_{q_1}\otimes\overline{\mathbf{\Lambda}_{q_1}} \\
=& \mathbf{I}\otimes\mathbf{I} - \mathbf{\Lambda}_{q_2}\otimes\overline{\mathbf{\Lambda}_{q_2}} \\
& + c\big[2\mathbf{\Lambda}_{q_2}\otimes\overline{\mathbf{\Lambda}_{q_2}} - \mathbf{\Lambda}_{q_2}\otimes\overline{\mathbf{\Lambda}_{q_1}} - \mathbf{\Lambda}_{q_1}\otimes\overline{\mathbf{\Lambda}_{q_2}}\big] \\
& + c^2\big(\mathbf{\Lambda}_{q_2}\otimes\overline{\mathbf{\Lambda}_{q_1}} + \mathbf{\Lambda}_{q_1}\otimes\overline{\mathbf{\Lambda}_{q_2}} - \mathbf{\Lambda}_{q_2}\otimes\overline{\mathbf{\Lambda}_{q_2}} - \mathbf{\Lambda}_{q_1}\otimes\overline{\mathbf{\Lambda}_{q_1}}\big) \\
=& c^2\mathbf{H}_2 + c\mathbf{H}_1 + \mathbf{H}_0 = \mathbf{H}(c).
\end{aligned}$$

Therefore, $c\mathbf{\Lambda}_{q_1} + (1-c)\mathbf{\Lambda}_{q_2}$ is Schur stable. □

Recall that we define the scalar $c^* = \boldsymbol{\omega}_{\theta^*}\boldsymbol{\omega}'_{\delta^*}$ and $\boldsymbol{\omega}_{\delta^*}$ is known for a given system. Let us denote $c_{min} = \frac{1}{\nu^-_{min}(\mathbf{G})}$ and $c_{max} = \frac{1}{\nu^+_{max}(\mathbf{G})}$. For c^* in the range of $[c_{min}, c_{max}]$, the corresponding range for $\boldsymbol{\omega}_{\theta^*}$ is obtained as:

$$\boldsymbol{\omega}^{[1]}_{\theta^*} = \frac{c^* - \boldsymbol{\omega}^{[2]}_{\delta^*}}{\boldsymbol{\omega}^{[1]}_{\delta^*} - \boldsymbol{\omega}^{[2]}_{\delta^*}}, \quad \boldsymbol{\omega}^{[2]}_{\theta^*} = 1 - \frac{c^* - \boldsymbol{\omega}^{[2]}_{\delta^*}}{\boldsymbol{\omega}^{[1]}_{\delta^*} - \boldsymbol{\omega}^{[2]}_{\delta^*}}. \tag{5.12}$$

Since $\boldsymbol{\omega}^{[2]}_{\theta^*}$ can be obtained from $\boldsymbol{\omega}^{[1]}_{\theta^*}$, we only look into the range for $\boldsymbol{\omega}^{[1]}_{\theta^*}$. Equation (5.12) illustrates the linear relationship between $\boldsymbol{\omega}^{[1]}_{\theta^*}$ and c^*. Therefore, the valid range for $\boldsymbol{\omega}^{[1]}_{\theta^*}$ is given by:

1) if $\boldsymbol{\omega}^{[1]}_{\delta^*} - \boldsymbol{\omega}^{[2]}_{\delta^*} > 0$,

$$\boldsymbol{\omega}^{[1]}_{\theta^*} \in \left[\frac{c_{min} - \boldsymbol{\omega}^{[2]}_{\delta^*}}{\boldsymbol{\omega}^{[1]}_{\delta^*} - \boldsymbol{\omega}^{[2]}_{\delta^*}}, \frac{c_{max} - \boldsymbol{\omega}^{[2]}_{\delta^*}}{\boldsymbol{\omega}^{[1]}_{\delta^*} - \boldsymbol{\omega}^{[2]}_{\delta^*}}\right] \cap [0,1],$$

2) if $\boldsymbol{\omega}_{\delta^*}^{[1]} - \boldsymbol{\omega}_{\delta^*}^{[2]} < 0$,

$$\boldsymbol{\omega}_{\theta^*}^{[1]} \in \left[\frac{c_{max} - \boldsymbol{\omega}_{\delta^*}^{[2]}}{\boldsymbol{\omega}_{\delta^*}^{[1]} - \boldsymbol{\omega}_{\delta^*}^{[2]}}, \frac{c_{min} - \boldsymbol{\omega}_{\delta^*}^{[2]}}{\boldsymbol{\omega}_{\delta^*}^{[1]} - \boldsymbol{\omega}_{\delta^*}^{[2]}}\right] \cap [0, 1].$$

For both situations, all the values for c_{max}, c_{min} and $\boldsymbol{\omega}_{\delta^*}$ are known so that the range for $\boldsymbol{\omega}_{\theta^*}^{[1]}$ can be computed. To simplify notations, denote $\eta_{min} \geq 0$ be the lower bound and $\eta_{max} \leq 1$ be the upper bound of $\boldsymbol{\omega}_{\theta^*}^{[1]}$, i.e., $\eta_{min} \leq \boldsymbol{\omega}_{\theta^*}^{[1]} \leq \eta_{max}$. Next lemma presents the relationship between $\boldsymbol{\omega}_{\theta^*}^{[1]}$ and λ_0, λ_1.

Lemma 5.3.4. *Given the boundary on the stationary distribution $\eta_{min} \leq \boldsymbol{\omega}_{\theta^*}^{[1]} \leq \eta_{max}$, the corresponding region for λ_0 and λ_1 are:*

Scenario 1: *$\eta_{min} = 0$ and $\eta_{max} < 1$*

$$0 \leq \lambda_0 \leq \Big(\frac{1 - \eta_{max}}{\eta_{max}}\lambda_1 + \frac{2\eta_{max} - 1}{\eta_{max}}\Big). \tag{5.13}$$

Scenario 2: *$\eta_{min} > 0$ and $\eta_{max} = 1$*

$$\Big(\frac{1 - \eta_{min}}{\eta_{min}}\lambda_1 + \frac{2\eta_{min} - 1}{\eta_{min}}\Big) \leq \lambda_0 \leq 1. \tag{5.14}$$

Scenario 3: *$\eta_{min} > 0$ and $\eta_{max} < 1$*

$$\begin{aligned}\Big(\frac{1 - \eta_{min}}{\eta_{min}}\lambda_1 + \frac{2\eta_{min} - 1}{\eta_{min}}\Big) \\ \leq \lambda_0 \leq \Big(\frac{1 - \eta_{max}}{\eta_{max}}\lambda_1 + \frac{2\eta_{max} - 1}{\eta_{max}}\Big).\end{aligned} \tag{5.15}$$

Proof. For transition matrix of a Markov chain defined as in (5.3), we first have:

$$\lim_{k\to\infty} {\mathbf{P}_\mathbf{m}}^k = \frac{1}{2 - \lambda_0 - \lambda_1}\begin{bmatrix} 1 - \lambda_1 & 1 - \lambda_0 \\ 1 - \lambda_1 & 1 - \lambda_0 \end{bmatrix}.$$

Therefore, the stationary distribution is

$$\boldsymbol{\omega}_{\theta^*} = \lim_{k\to\infty} \boldsymbol{\omega}_\theta {\mathbf{P}_\mathbf{m}}^k = \begin{bmatrix} \frac{1-\lambda_1}{2-\lambda_0-\lambda_1} & \frac{1-\lambda_0}{2-\lambda_0-\lambda_1} \end{bmatrix}. \tag{5.16}$$

Given the boundary on $\boldsymbol{\omega}_{\theta^*}^{[1]}$, we have:

$$\eta_{min} \leq \frac{1-\lambda_1}{2-\lambda_0-\lambda_1} \leq \eta_{max}. \tag{5.17}$$

Based on different values of η_{min} and η_{max}, we can get the inequalities in (5.13)-(5.15). □

Equations (5.13)-(5.15) yield an underdetermined problem, i.e., there is one degree of freedom. The relationship in (5.13)-(5.15), along with the constrains $0 \leq \lambda_0, \lambda_1 \leq 1$ fully define the possible ranges for λ_0 and λ_1 that ensures the convergence of the first term in (5.8).

5.3.2 Boundedness on the Continuous States

We have addressed the convergence problem for the system in (5.9). In this section, we deal with the boundedness problem for the second term in equation (5.8):

$$\mathbf{u}_2'\mathbf{P_m}^k\boldsymbol{\omega}_\theta\boldsymbol{\omega}_\delta\mathbf{P_s}^k \times \left[\mathbf{u}_1(\boldsymbol{\Lambda}_{q_2} - \boldsymbol{\Gamma}_{q_2,q_1}) + \mathbf{u}_2(\boldsymbol{\Lambda}_{q_1} - \boldsymbol{\Gamma}_{q_1,q_2})\right]\mathbb{E}(\mathbf{x}_{k-1}). \tag{5.18}$$

For a given system, the matrix

$$\mathbf{u}_2'\mathbf{P_m}^k\boldsymbol{\omega}_\theta\boldsymbol{\omega}_\delta\mathbf{P_s}^k\left[\mathbf{u}_1(\boldsymbol{\Lambda}_{q_2} - \boldsymbol{\Gamma}_{q_2,q_1}) + \mathbf{u}_2(\boldsymbol{\Lambda}_{q_1} - \boldsymbol{\Gamma}_{q_1,q_2})\right]$$

is deterministic and finite. Therefore, the boundedness of (5.18) only depends on $\mathbb{E}(\mathbf{x}_k)$. Recall that $\mathbb{E}(\mathbf{x}_k)$ is the expectation of the continuous state of the MJLS in (5.2). This boundedness problem is equivalent to the stability analysis of MJLS in (5.2). We first investigate the stability problem for the following Markov jump autonomous linear system:

$$\mathbf{x}_k = \mathbf{A}_{\delta_k}\mathbf{x}_{k-1}. \tag{5.19}$$

There are numerous concepts of stability that have been defined for SHS. In this chapter, we use the definition of stochastic second moment stability, also referred to as mean squared

stability (MSS) [79, 80]. The autonomous system (5.19) is MSS if for any initial distribution of $\mathbf{x}_0^*$ and δ_0,

$$\lim_{k\to\infty} \mathbb{E}(\|\mathbf{x}_k^*\|) = 0.$$

Among the existing approaches to study stability, the main approach is primarily built on the well-known Lyapunov theory. A necessary and sufficient condition for stability of system (5.19) with finite Markov chain $\{\delta_k\}$ is presented in [81]. A condition using Kronecker product provides a testable condition and our following analysis is built on this result. The following theorem can be found in [80].

Theorem 5.3.2. *The system in (5.19) is MSS if and only if the matrix* $\mathbf{Z}$ *is Schur stable* ($\rho(\mathbf{Z}) < 1$) *with* $\mathbf{Z}$ *defined as:*

$$\mathbf{Z} = diag[\mathbf{A}_{q_1} \otimes \mathbf{A}_{q_1}, \cdots, \mathbf{A}_{q_d} \otimes \mathbf{A}_{q_d}] \cdot (\mathbf{P_s}' \otimes \mathbf{I}). \tag{5.20}$$

For a given system (5.19), the matrices $\mathbf{A}_{q_1}, \cdots, \mathbf{A}_{q_d}$ and $\mathbf{P_s}$ are known. Therefore, the condition $\rho(\mathbf{Z}) < 1$ is numerically testable. The following theorem exposes the relationship between the stability of (5.2) and (5.19) with a detailed proof in [20].

Theorem 5.3.3. *The following statements are equivalent:*
(i) The MJLS represented in (5.19) is MSS;
(ii) The MJLS represented in (5.2) is MSS if $\mathbf{w}_k$ *is second order independent wide sense stationary (WSS).*

In our system model, $\mathbf{w}_k$ is second order independent WSS since we assume it is Gaussian with mean $\mathbf{0}$ and covariance $\mathbf{Q}$. Therefore, the condition in Theorem 5.3.2 also indicates the stability of (5.2) which results in the boundedness of (5.18). Finally, we can combine all the results introduced thus far and present our main result in the following theorem.

Theorem 5.3.4. *For the MJLS in (5.2), the state estimation bias dynamics introduced by mode mismatch modeled as a Markov chain* θ_k *is statistically convergent as* $\lim_{k\to\infty} \mathbb{E}(\mathbf{e}_k) <$

∞ *if:*

(i) The probabilities λ_0 *and* λ_1 *satisfy (5.13)-(5.15);*

(ii) Matrix $\mathbf{Z}$ *is Schur stable with* $\mathbf{Z}$ *defined as (5.20).*

Proof. The first condition that λ_0 and λ_1 satisfy (5.13)-(5.15) ensures the convergence of (5.9). The second condition guarantees the boundedness of (5.18). Since the mean of bias dynamics evolves as in (5.8), (5.8) is bounded-input bounded-output (BIBO) stable [12]. □

5.4 MJLS with arbitrary numbers of modes

In Chapter 5.2, we have derived the dynamics of bias $\mathbf{e}_k$ in (5.4). In this section, we consider a generalized MJLS and analyze the mean process of $\mathbf{e}_k$. The proposed approach is to model it as an auxiliary linear system and then study the statistical convergence on the mean process.

5.4.1 Mean Process of the Bias Dynamics

Based on equation (5.4), $\{\mathbf{e}_k\}_{k=0}^{\infty}$ is a stochastic process that depends on γ_k, δ_k and θ_k. Therefore, we consider convergence in mean, i.e.,

$$\lim_{k\to\infty} \mathbb{E}(\mathbf{e}_k) < \infty. \tag{5.21}$$

According to the tower rule, we have $\mathbb{E}(\mathbf{e}_k) = \mathbb{E}(\mathbb{E}(\mathbf{e}_k|\mathbf{e}_{k-1}))$ where the outer expectation is taken over $\mathbf{e}_{k-1}$ and the inner expectation is taken over γ_k, δ_k and θ_k. We can write:

$$\mathbb{E}(\mathbf{e}_k) = \mathbb{E}_{\gamma_k}\left[\mathbf{\Lambda}_{\gamma_k}\right]\mathbb{E}(\mathbf{e}_{k-1}) + \mathbb{E}_{\gamma_k,\theta_k,\delta_k}\left[\theta_k(\mathbf{\Lambda}_{\gamma_k} - \mathbf{\Gamma}_{\gamma_k,\delta_k})\right]\mathbb{E}(\mathbf{x}_{k-1}). \tag{5.22}$$

The dynamics of $\mathbb{E}(\mathbf{e}_k)$ in (5.22) can be interpreted as a linear system in the form:

$$\mathbf{z}_k = \mathbf{\Omega}_k\mathbf{z}_{k-1} + \mathbf{\Phi}_k\mathbf{u}_{k-1}, \tag{5.23}$$

where the state and input correspond to:

$$\mathbf{z}_k = \mathbb{E}(\mathbf{e}_k), \quad \mathbf{u}_k = \mathbb{E}(\mathbf{x}_k),$$

and

$$\begin{aligned}\boldsymbol{\Omega}_k &= \mathbb{E}_{\gamma_k}\left[\boldsymbol{\Lambda}_{\gamma_k}\right] \\ \boldsymbol{\Phi}_k &= \mathbb{E}_{\gamma_k,\theta_k,\delta_k}\left[\theta_k(\boldsymbol{\Lambda}_{\gamma_k} - \boldsymbol{\Gamma}_{\gamma_k,\delta_k})\right].\end{aligned}$$

Therefore, the matrices $\boldsymbol{\Omega}_k$ and $\boldsymbol{\Phi}_k$ are closely related to the distributions of γ_k, θ_k and δ_k. Recall that γ_k denotes the observed (or estimated) mode at time k and γ_k is correlated with θ_k and δ_k. In the rest of this chapter, we will use the notation $\mathrm{P}_{\gamma_k}^{q_i}$ to denote that probability of γ_k taking the value q_i, i.e.,

$$\mathrm{P}_{\gamma_k}^{q_i} = \mathbb{P}(\gamma_k = q_i),$$

and $\mathbf{p}_{\gamma_k}$ for the distribution of γ_k as:

$$\mathbf{p}_{\gamma_k} = \left[\mathrm{P}_{\gamma_k}^{q_1}, \quad \cdots, \quad \mathrm{P}_{\gamma_k}^{q_d}\right].$$

We can use a similar notation for the distribution of δ_k and θ_k. Since the distribution on γ_k fully depends on δ_k and θ_k, we will first derive the distribution for δ_k and θ_k and then present the distribution of γ_k. Denote the initial distributions for δ_0 and θ_0 as row vectors as shown below:

$$\begin{aligned}\mathbf{p}_{\delta_0} &= \left[\mathrm{P}_{\delta_0}^{q_1}, \quad \cdots, \quad \mathrm{P}_{\delta_0}^{q_d}\right], \\ \mathbf{p}_{\theta_0} &= \left[\mathrm{P}_{\theta_0}^{0}, \quad \mathrm{P}_{\theta_0}^{1}\right].\end{aligned}$$

Since both δ_k and θ_k are Markovian process, we can write the probability distributions of δ_k and θ_k as:

$$\mathbf{p}_{\delta_k} = \mathbf{p}_{\delta_0}{\mathbf{P}_\mathbf{s}}^k, \quad \mathbf{p}_{\theta_k} = \mathbf{p}_{\theta_0}{\mathbf{P}_\mathbf{m}}^k.$$

We denote the stationary distribution of γ_k and δ_k as:

$$\mathbf{p}_\delta = \lim_{k\to\infty} \mathbf{p}_{\delta_k} = \left[\mathrm{P}_\delta^{q_1}, \quad \cdots, \quad \mathrm{P}_\delta^{q_d}\right],$$
$$\mathbf{p}_\gamma = \lim_{k\to\infty} \mathbf{p}_{\gamma_k} = \left[\mathrm{P}_\gamma^{q_1}, \quad \cdots, \quad \mathrm{P}_\gamma^{q_d}\right].$$

As discussed earlier, we consider a multi-mode discrete state space, i.e., $\mathcal{Q} = \{q_1, \cdots .q_d\}$. Therefore, γ_k follows the same distribution as δ_k if $\theta_k = 0$. However, the distribution for γ_k cannot be fully determined if $\theta_k = 1$. In order to derive the distribution of γ_k, we make the following assumption: if $\theta_k = 1$, γ_k takes values in $\mathcal{Q}\backslash\{\delta_k\}$ with equal probability. Therefore, we have $\forall q_i \in \mathcal{Q}$, the probability of $\gamma_k = q_i$ is:

$$\mathrm{P}_{\gamma_k}^{q_i} = \mathrm{P}_{\theta_k}^0 \mathrm{P}_{\delta_k}^{q_i} + \frac{\mathrm{P}_{\theta_k}^1}{d-1} \sum_{q_j\in\mathcal{Q}}^{q_j\neq q_i} \mathrm{P}_{\delta_k}^{q_j}, \tag{5.24}$$

with

$$\begin{aligned}
\sum_{q_i\in\mathcal{Q}} \mathrm{P}_{\gamma_k}^{q_i} &= \sum_{q_i\in\mathcal{Q}} \left(\mathrm{P}_{\theta_k}^0 \mathrm{P}_{\delta_k}^{q_i}\right) + \sum_{q_i\in\mathcal{Q}} \left(\frac{\mathrm{P}_{\theta_k}^1}{d-1} \sum_{q_j\in\mathcal{Q}}^{q_j\neq q_i} \mathrm{P}_{\delta_k}^{q_j}\right) \\
&= \mathrm{P}_{\theta_k}^0 \sum_{q_i\in\mathcal{Q}} \mathrm{P}_{\delta_k}^{q_i} + \frac{\mathrm{P}_{\theta_k}^1}{d-1} \sum_{q_i\in\mathcal{Q}} \sum_{q_j\in\mathcal{Q}}^{q_j\neq q_i} \mathrm{P}_{\delta_k}^{q_j} \\
&= \mathrm{P}_{\theta_k}^0 + \frac{\mathrm{P}_{\theta_k}^1}{d-1} \sum_{q_i\in\mathcal{Q}} \left(1 - \mathrm{P}_{\delta_k}^{q_i}\right) \\
&= \mathrm{P}_{\theta_k}^0 + \frac{\mathrm{P}_{\theta_k}^1}{d-1}\left(d-1\right) = 1.
\end{aligned}$$

5.4.2 Stability Analysis for Mean Bias Dynamics

Based on preceding discussion, it is easy to see that the convergence of $\mathbb{E}(\mathbf{e}_k)$ as (5.21) is equivalent to the marginal stability of $\mathbf{z}_k$. Definition of marginal stability is based on Lyapunov stability and we include the definition below for the sake of completeness.

Definition 5.4.1. *A linear system in (5.23) is* ***Lyapunov stable*** *if there exists some $\xi > 0$ such that $\|\mathbf{z}_0\| < \xi$ implies $\|\mathbf{z}_k\| < \epsilon$ for all k.*

Remark 5.4.1. *A linear system is* ***marginally stable*** *if it is neither Lyapunov stable nor unstable.*

Note that Lyapunov stability gives a stronger condition than (5.21) since it requires not only convergence but convergence to a region close to 0. In the following, we will develop conditions on Lyapunov stability of (5.23) which can guarantee that (5.21) is satisfied. We first consider stability for the corresponding autonomous system:

$$\mathbf{z}_k = \mathbf{\Omega}_k \mathbf{z}_{k-1}. \tag{5.25}$$

It is known that system (5.25) is Lyapunov stable if and only if $\lim_{k\to\infty} \rho(\mathbf{\Omega}_k) < 1$ with $\mathbf{\Omega}_k$ defined as

$$\mathbf{\Omega}_k = \mathbb{E}_{\gamma_k}\left[\mathbf{\Lambda}_{\gamma_k}\right] = \sum_{i=1}^{d} \mathrm{P}_{\gamma_k}^{q_i} \mathbf{\Lambda}_{q_i}. \tag{5.26}$$

We have:

$$\begin{aligned} \lim_{k\to\infty} \rho(\mathbf{\Omega}_k) &= \rho\left(\lim_{k\to\infty} \mathbf{\Omega}_k \right) \\ &= \rho\left[\lim_{k\to\infty} \left(\sum_{i=1}^{d} \mathrm{P}_{\gamma_k}^{q_i} \mathbf{\Lambda}_{q_i} \right) \right] = \rho\left(\sum_{i=1}^{d} \mathrm{P}_{\gamma}^{q_i} \mathbf{\Lambda}_{q_i} \right). \end{aligned} \tag{5.27}$$

The matrix in (5.27) is a convex combination of d matrices $\mathbf{\Lambda}_{q_1}, \cdots, \mathbf{\Lambda}_{q_d}$ (also referred to as matrix polytope). Derivation of closed form algebraic expression for spectral radius of a

matrix polytope is still an open task. In the previous section (Chapter 5.3), we considered the same research question with discrete state space containing only two modes, i.e., $\mathcal{Q} = \{q_1, q_2\}$. The two mode assumption simplifies the analysis as known δ_k and θ_k can uniquely determine the value of γ_k. Additionally, the two-mode assumption enables us to develop a necessary and sufficient condition that guarantees the statistically convergence on $\mathbf{e}_k$ by leveraging results on Schur stability and nonsingularity of an auxiliary matrix. In this section, due to the fact that γ_k, δ_k and θ_k are correlated and a generalized discrete space $\mathcal{Q} = \{q_1, \cdots . q_d\}$ is considered, we take an different approach than the previous section and derive a sufficient condition such that (5.21) is satisfied. In the following, we will review the prior efforts related to the spectral analysis of a matrix polytope and then present our main results.

A matrix polytope is defined as:

$$\mathcal{V} = \left\{ \mathbf{V} = \sum_{i=1}^{m} \alpha_i \mathbf{V}_i \;\middle|\; \alpha_i \geq 0, \sum_{i=1}^{m} \alpha_i = 1, \mathbf{V_i} \in \mathbb{R}^{n \times n} \right\}. \tag{5.28}$$

Research work on spectral radius of (5.28) spans multiple decades. Approaches can be broadly grouped into three categories: (1) Singularity approach; (2) Interval matrix approach and, (3) Linear matrix inequality (LMI) approach.

Singularity Approach

[78] studies Schur stability for polytope with two matrices and they derive conditions on α_i that assures the stability of (5.28). [82] extends result of [78] to the case of three matrices and they first introduce the equivalence between polytope of matrices and the interval matrix. The approach they proposed for Schur stability of (5.28) is to check a block P-matrix respect to partitions on space $\{1, \cdots, 6n^2\}$. [83] generalizes the results in [82] to all cases of $k \geq 2$.

Interval Matrix Approach

As stated in [82], the convex combination of matrices is equivalent to an interval matrix. [84] first proposes a sufficient and necessary condition for stability of interval matrices but [85] then proves that the condition derived in [84] is not sufficient using a counterexample. [86] proposes a sufficient and necessary condition for a specific class of interval matrices where all diagonal elements are negative and non-diagonal elements are non-negative. [87] presents an algebraically testable sufficient condition for a general class of interval matrices. It also proves a sufficient and necessary condition on subinterval matrices. [88] studies Schur stability of interval matrices based on the definition of M-matrix and derives sufficient element-wise conditions.

LMI-based Approach

Another approach is to derive LMI-based conditions originally used for stability analysis. For polytopic systems, [89] derives sufficient and necessary conditions based on a class of affine parameter-dependent Lyapunov functions. A similar problem has been studied in [90–92]and they provide less conservative necessary and sufficient computationally verifiable conditions for a symmetric positive definite polytope of matrices. Another result is presented in [93] and their approach is based on homogeneous polynomially parameter-dependent quadratic Lyapunov function (HPD-QLF). Other related works can be found in [94, 95].

In this chapter, we intend to derive conditions on probability of mode mismatch such that the bias dynamics is statistical convergent (i.e., (5.21) is satisfied). Due to the fact that both singularity approach and LMI-based approach do not provide an algebraic solution for the coefficient α_i, we will build our results on the interval matrix approach.

To study the statistical convergence of the bias dynamics, we first analyze the spectral radius of the matrix polytope $\mathbf{\Omega}_k$. By mapping the matrix polytope into an interval matrix, we derive sufficient conditions that guarantees the stability of (5.25), Finally, we analyze the stability of the input term in (5.23) and complete the analysis.

5.4.3 Schur Stability of Interval Matrix

Recall that the goal is to find conditions such that $\lim_{k\to\infty} \rho(\mathbf{\Omega}_k) < 1$ with $\lim_{k\to\infty} \rho(\mathbf{\Omega}_k)$ defined in (5.27). A d-dimensional standard (or probability) simplex is defined as:

$$\mathbb{A} = \left\{ \mathbf{a} \;\middle|\; \sum_{i=1}^{d} \mathbf{a}^{[i]} = 1 \right\} \subset [0,1]^d.$$

We denote $\alpha_i = \mathrm{P}_\gamma^{q_i}$ and $\boldsymbol{\alpha} = [\alpha_1, \cdots, \alpha_d] \in \mathbb{A}$ and define

$$\mathbf{\Omega}(\boldsymbol{\alpha}) = \sum_{i=1}^{d} \alpha_i \mathbf{\Lambda}_{q_i}, \quad \boldsymbol{\alpha} \in \mathbb{A}. \tag{5.29}$$

In (5.29), $\mathbf{\Lambda}_{q_i}$ are fixed matrices for a given system. Therefore, the goal is to find a feasible region $\mathbb{A}_\mathsf{f} \subseteq \mathbb{A}$ such that $\forall \boldsymbol{\alpha} \in \mathbb{A}_\mathsf{f}$, $\rho(\mathbf{\Omega}(\boldsymbol{\alpha})) < 1$. The following lemma presents the equivalence between convex set of matrices and interval matrix.

Lemma 5.4.1. *The convex set of matrices $\mathbf{\Lambda}_{q_i}$ is equivalent to the interval matrix $[\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$ with*

$$\begin{aligned} \underline{\mathbf{\Lambda}}^{[k,l]} &= \min_{q_i \in \mathcal{Q}} \mathbf{\Lambda}_{q_i}^{[k,l]}, \quad \forall k,l \in [1,n]^2 \\ \overline{\mathbf{\Lambda}}^{[k,l]} &= \max_{q_i \in \mathcal{Q}} \mathbf{\Lambda}_{q_i}^{[k,l]}, \quad \forall k,l \in [1,n]^2. \end{aligned} \tag{5.30}$$

The following theorem provides a sufficient condition on Schur stability on interval matrix $[\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$ with a detailed proof found in [87].

Theorem 5.4.1. *The interval matrix $[\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$ is Schur stable if:*
(i) $\mathbf{\Lambda}_0 \triangleq \frac{1}{2}(\underline{\mathbf{\Lambda}} + \overline{\mathbf{\Lambda}})$ is Schur stable, and therefore, there exists a s.p.d. matrix $\mathbf{P}$ s.t.

$$\mathbf{\Lambda}_0' \mathbf{P} \mathbf{\Lambda}_0 - \mathbf{P} = -\mathbf{I}$$

(ii) $\frac{1}{2}\beta(\overline{\mathbf{\Lambda}} - \underline{\mathbf{\Lambda}}) < \left[\beta(\mathbf{\Lambda}_0)^2 + \frac{1}{\|\mathbf{P}\|_\infty}\right]^{1/2} - \beta(\mathbf{\Lambda}_0)$,

where,

$$\beta(\mathbf{A}) = \max\left\{\|\mathbf{A}\|_1, \|\mathbf{A}\|_\infty\right\}$$

and

$$\|\mathbf{A}\|_1 = \max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|, \quad \|\mathbf{A}\|_\infty = \max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|.$$

Based on the equivalence relation in Lemma 5.4.1 and the sufficient condition presented in Theorem 5.4.1, the following corollary can be easily obtained.

Corollary 5.4.1.1. *The matrix $\mathbf{\Omega}(\boldsymbol{\alpha})$ is Schur stable for all $\boldsymbol{\alpha} \in \mathbb{A}$ if the interval matrix $[\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$ is Schur stable.*

For a given system described in (5.2), if the calculated $\underline{\mathbf{\Lambda}}$ and $\overline{\mathbf{\Lambda}}$ based on Lemma 5.4.1 satisfy conditions in Theorem 5.4.1, then the bias dynamics in (5.25) is convergent. On the other hand, for the case that $\underline{\mathbf{\Lambda}}$ and $\overline{\mathbf{\Lambda}}$ do not satisfy Theorem 5.4.1, we intend to find constraints on $\boldsymbol{\alpha}$ such that the Schur stability of $\mathbf{\Omega}(\boldsymbol{\alpha})$ can still be guaranteed. Before we present the main theorem, the concept of subinterval matrix of the interval matrix $[\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$ need to be defined.

Definition 5.4.2. *An interval matrix $[\underline{\mathbf{\Delta}}, \overline{\mathbf{\Delta}}]$ is a subinterval matrix of $[\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$ if and only if $[\underline{\mathbf{\Delta}}, \overline{\mathbf{\Delta}}] \subseteq [\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$.*

The following theorem states the non-existence of choice of $\boldsymbol{\alpha}$ such that the matrix $\mathbf{\Omega}(\boldsymbol{\alpha})$ is Schur stable.

Theorem 5.4.2. *For the matrix $\mathbf{\Omega}(\boldsymbol{\alpha})$ defined in (5.29) and the feasible region of $\boldsymbol{\alpha}$ corresponding to $\mathbb{A}_\mathsf{f} \subseteq \mathbb{A}$, $\mathbb{A}_\mathsf{f} \neq \emptyset$ if and only if there exists a subinterval matrix $[\underline{\mathbf{\Delta}}, \overline{\mathbf{\Delta}}]$ of $[\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$ such that $[\underline{\mathbf{\Delta}}, \overline{\mathbf{\Delta}}]$ satisfies Theorem 5.4.1. Specifically, if $[\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$ satisfies Theorem 5.4.1, then $\mathbb{A}_\mathsf{f} = \mathbb{A}$.*

Proof. We first prove sufficiency ($\Longrightarrow$):
Let $[\underline{\mathbf{\Delta}}, \overline{\mathbf{\Delta}}] \subseteq [\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$ and $[\underline{\mathbf{\Delta}}, \overline{\mathbf{\Delta}}]$ satisfies Theorem 5.4.1. By definition, there exists $\boldsymbol{\xi} =$

$[\xi_1, \cdots, \xi_d] \in \mathbb{A}$ such that $\sum_{i=1}^{d} \xi_i \mathbf{\Lambda}_{q_i} = \underline{\mathbf{\Delta}}$ and there exists $\boldsymbol{\zeta} = [\zeta_1, \cdots, \zeta_d] \in \mathbb{A}$ such that $\sum_{i=1}^{d} \zeta_i \mathbf{\Lambda}_{q_i} = \overline{\mathbf{\Delta}}$. We can write:

$$\begin{aligned}\mathbf{\Delta}_0 &\triangleq \frac{1}{2}(\underline{\mathbf{\Delta}} + \overline{\mathbf{\Delta}}) = \frac{1}{2}\Big(\sum_{i=1}^{d} \xi_i \mathbf{\Lambda}_{q_i} + \sum_{i=1}^{d} \zeta_i \mathbf{\Lambda}_{q_i}\Big) \\ &= \sum_{i=1}^{d} \frac{(\xi_i + \zeta_i)}{2} \mathbf{\Lambda}_{q_i}\end{aligned}$$

According to Theorem 5.4.1, we know that $\rho(\mathbf{\Delta}_0) < 1$. Since

$$\begin{aligned}&\sum_{i=1}^{d} \frac{(\xi_i + \zeta_i)}{2} = 1 \\ &0 \leq \frac{(\xi_i + \zeta_i)}{2} \leq 1, \quad \forall i\end{aligned}$$

we have $\frac{1}{2}(\boldsymbol{\xi} + \boldsymbol{\zeta}) \in \mathbb{A}$. The feasible space $\mathbb{A}_\mathsf{f}$ at least contains $\frac{1}{2}(\boldsymbol{\xi} + \boldsymbol{\zeta})$. That is, $\mathbb{A}_\mathsf{f} \neq \emptyset$.

Now we prove necessity ($\Longleftarrow$):

If the feasible region $\mathbb{A}_\mathsf{f} \neq \emptyset$, let $\boldsymbol{\omega} = [\omega_1, \cdots, \omega_d] \in \mathbb{A}_\mathsf{f}$. Since $\rho(\sum_{i=1}^{d} \omega_i \mathbf{\Lambda}_{q_i}) < 1$, then we can construct an interval matrix $[\underline{\mathbf{\Delta}}, \overline{\mathbf{\Delta}}]$ by defining $\underline{\mathbf{\Delta}} = \overline{\mathbf{\Delta}} = \sum_{i=1}^{d} \omega_i \mathbf{\Lambda}_{q_i}$. It is obvious that $[\underline{\mathbf{\Delta}}, \overline{\mathbf{\Delta}}] \subseteq [\underline{\mathbf{\Lambda}}, \overline{\mathbf{\Lambda}}]$. The matrix $\mathbf{\Delta}_0 = \frac{1}{2}(\underline{\mathbf{\Delta}} + \overline{\mathbf{\Delta}}) = \sum_{i=1}^{d} \omega_i \mathbf{\Lambda}_{q_i}$ which is Schur stable, therefore, there exists a s.p.d. matrix $\mathbf{P}$ such that

$$\mathbf{\Lambda}_0' \mathbf{P} \mathbf{\Lambda}_0 - \mathbf{P} = -\mathbf{I}.$$

For the second condition in Theorem 5.4.1, we have the left-hand side $\frac{1}{2}\beta(\overline{\mathbf{\Lambda}} - \underline{\mathbf{\Lambda}}) = 0$. Therefore, we need to show:

$$\begin{aligned}&\Big[\beta(\mathbf{\Lambda}_0)^2 + \frac{1}{\|\mathbf{P}\|_\infty}\Big]^{1/2} - \beta(\mathbf{\Lambda}_0) > 0 \\ \Rightarrow &\Big[\beta(\mathbf{\Lambda}_0)^2 + \frac{1}{\|\mathbf{P}\|_\infty}\Big]^{1/2} > \beta(\mathbf{\Lambda}_0) \\ \Rightarrow &\beta(\mathbf{\Lambda}_0)^2 + \frac{1}{\|\mathbf{P}\|_\infty} > \beta(\mathbf{\Lambda}_0)^2.\end{aligned}$$

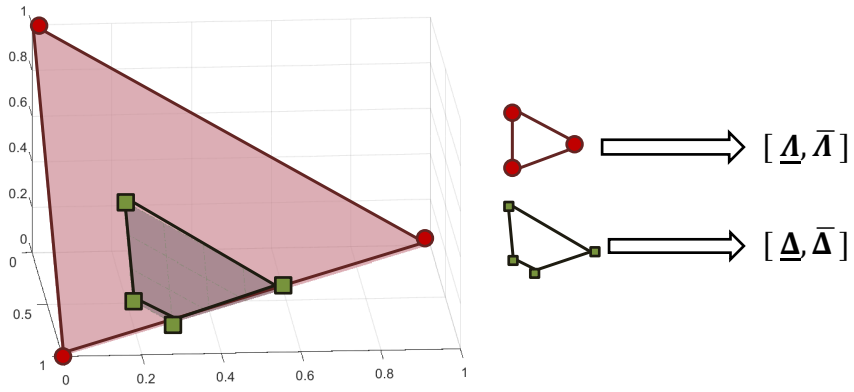


Figure 5.2: *Mapping from space* $\mathbb{A}$, $\mathbb{A}_\mathsf{f}$ *to interval matrix* $[\underline{\boldsymbol{\Lambda}}, \overline{\boldsymbol{\Lambda}}]$ *and* $[\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}]$

This inequality holds obviously. Therefore, we prove that if $\mathbb{A}_\mathsf{f} \neq \emptyset$, then there exists a subinterval matrix $[\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}] \subseteq [\underline{\boldsymbol{\Lambda}}, \overline{\boldsymbol{\Lambda}}]$ such that Theorem 5.4.1 is satisfied. Lastly, it is easy to show that if $[\underline{\boldsymbol{\Lambda}}, \overline{\boldsymbol{\Lambda}}]$ satisfies Theorem 5.4.1, then $\mathbb{A}_\mathsf{f} = \mathbb{A}$ based on the derivation in the proof. We omit the proof for this special case due to page limit constraints. □

Based on the result of Theorem 5.4.2, the following two problems are equivalent:

1. Solve for a feasible region $\mathbb{A}_\mathsf{f}$ such that $\boldsymbol{\Omega}(\boldsymbol{\alpha})$ is Schur stable $\forall \boldsymbol{\alpha} \in \mathbb{A}_\mathsf{f}$;

2. Solve for a feasible region $\mathbb{A}_\mathsf{f}$ such that the corresponding subinterval matrix $[\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}]$ is Schur stable. Here, the subinterval matrix $[\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}]$ corresponding to $\mathbb{A}_\mathsf{f}$ is defined as $\boldsymbol{\Omega}(\boldsymbol{\alpha}) \in [\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}]$ if and only if $\boldsymbol{\alpha} \in \mathbb{A}_\mathsf{f}$.

To assist our analysis, we need to investigate the mapping from space $\mathbb{A}$ to interval matrix $[\underline{\boldsymbol{\Lambda}}, \overline{\boldsymbol{\Lambda}}]$. Figure 5.2 illustrates the space $\mathbb{A}$ and $\mathbb{A}_\mathsf{f}$ to interval matrix $[\underline{\boldsymbol{\Lambda}}, \overline{\boldsymbol{\Lambda}}]$ and $[\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}]$ in 3-dimensional space. The polytope represented by $\mathbb{A}$ is also called the standard simplex. Based on the definition of $\underline{\boldsymbol{\Lambda}}$ and $\overline{\boldsymbol{\Lambda}}$ in (5.30), each element of matrices $\underline{\boldsymbol{\Lambda}}$ and $\overline{\boldsymbol{\Lambda}}$ is obtained at the three boundary points (red circle in Figure 5.2). Therefore, the lower bound $\underline{\boldsymbol{\Delta}}$ and upper bound $\overline{\boldsymbol{\Delta}}$ for the subinterval matrix can also be obtained from the boundary points (green square in Figure 5.2). The following lemma shows the relationship between $\mathbb{A}_\mathsf{f}$ and $[\underline{\boldsymbol{\Lambda}}, \overline{\boldsymbol{\Lambda}}]$.

Algorithm 6 Map From $\mathbb{A}_{\mathsf{f}}$ to Interval Matrix $[\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}]$

function LPS($\{\boldsymbol{\Lambda}_{q_i}\}_{i=1}^{d}, \{\underline{\alpha_{\mathsf{f}_i}}, \overline{\alpha_{\mathsf{f}i}}\}_{i=1}^{d}$)
  for $k, l = (1:n, 1:n)$ **do**
    Minimize: $V = \sum_i \alpha_{\mathsf{f}i} \boldsymbol{\Lambda}_{q_i}^{[k,l]}$
    Subject to: $\underline{\alpha_{\mathsf{f}_i}} \leq \alpha_{\mathsf{f}i} \leq \overline{\alpha_{\mathsf{f}i}}$
      $\sum_{i=1}^{d} \alpha_{\mathsf{f}i} = 1$

    Maximize: $W = \sum_i \alpha_{\mathsf{f}i} \boldsymbol{\Lambda}_{q_i}^{[k,l]}$
    Subject to: $\underline{\alpha_{\mathsf{f}_i}} \leq \alpha_{\mathsf{f}i} \leq \overline{\alpha_{\mathsf{f}i}}$
      $\sum_{i=1}^{d} \alpha_{\mathsf{f}i} = 1$

    $\underline{\boldsymbol{\Delta}}^{[k,l]} = V^*$ where V^* is the optimal value
    $\overline{\boldsymbol{\Delta}}^{[k,l]} = W^*$ where W^* is the optimal value
  end for
  return $\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}$
end function

Lemma 5.4.2. *For a d-dimensional standard simplex $\mathbb{A}$ and a subspace $\mathbb{A}_{\mathsf{f}} \subseteq \mathbb{A}$, let $\boldsymbol{\alpha}_{\mathsf{f}} = [\alpha_{\mathsf{f}1}, \cdots, \alpha_{\mathsf{f}d}] \in \mathbb{A}_{\mathsf{f}}$ with:*

$$0 \leq \underline{\alpha_{\mathsf{f}_i}} \leq \alpha_{\mathsf{f}i} \leq \overline{\alpha_{\mathsf{f}i}} \leq 1, \forall i \tag{5.31}$$

The interval matrix $[\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}]$ corresponding to $\mathbb{A}_{\mathsf{f}}$ is obtained from:

$$\begin{aligned} \underline{\boldsymbol{\Delta}}^{[k,l]} &= \min_{\underline{\alpha_{\mathsf{f}_i}} \leq \alpha_{\mathsf{f}i} \leq \overline{\alpha_{\mathsf{f}i}}} \sum_i \alpha_{\mathsf{f}i} \boldsymbol{\Lambda}_{q_i}^{[k,l]}, \quad \forall k, l \in [1, n]^2 \\ \overline{\boldsymbol{\Delta}}^{[k,l]} &= \max_{\underline{\alpha_{\mathsf{f}_i}} \leq \alpha_{\mathsf{f}i} \leq \overline{\alpha_{\mathsf{f}i}}} \sum_i \alpha_{\mathsf{f}i} \boldsymbol{\Lambda}_{q_i}^{[k,l]}, \quad \forall k, l \in [1, n]^2. \end{aligned} \tag{5.32}$$

Since $\mathbb{A}_{\mathsf{f}}$ is a convex polytope, equation (5.32) forms a linear programming problem with $d-1$ degrees of freedom. Even though the optimal solution of equation (5.32) falls on the extreme points, an analytical solution is difficult to acquire for a general system set up. Algorithm 6 gives a numerical approach to obtain the corresponding interval matrix if given the boundary on $\boldsymbol{\alpha}$. Algorithm 7 is the implementation of Theorem 5.4.1 which illustrates

Algorithm 7 Schur Stable of Interval Matrix

function SCHURSTABLE($\underline{\boldsymbol{\Delta}}, \overline{\boldsymbol{\Delta}}$)
  Let $\boldsymbol{\Delta}_0 = \frac{1}{2}(\underline{\boldsymbol{\Delta}} + \overline{\boldsymbol{\Delta}})$
  if $\rho(\boldsymbol{\Delta}_0) < 1$ **then**
    Find $\mathbf{P}$
    Subject to $\mathbf{P} = \mathbf{P}'$
      $\mathbf{P} \succ 0$
      $\boldsymbol{\Delta}_0' \mathbf{P} \boldsymbol{\Delta}_0 - \mathbf{P} = -\mathbf{I}$
    if $\frac{1}{2}\beta(\overline{\boldsymbol{\Delta}} - \underline{\boldsymbol{\Delta}}) < \left[\beta(\boldsymbol{\Delta}_0)^2 + \frac{1}{\|\mathbf{P}\|_\infty}\right]^{1/2} - \beta(\boldsymbol{\Delta}_0)$ **then**
      return True
    else
      return False
    end if
  else
    return False
  end if
end function

a method to check if an interval matrix is Schur stable (return True) or not (return False). Algorithm 6 and Algorithm 7 can be used to analyze whether a region on $\boldsymbol{\alpha}$ forms a feasible region such that every $\boldsymbol{\alpha}$ in the region implies a Schur stable $\boldsymbol{\Omega}(\boldsymbol{\alpha})$. An exhaustive search can be conducted on the space $[0, 1]^d$ in order to obtain all the stable regions with respect to $\boldsymbol{\alpha}$ for a given system.

5.4.4 Conditions on Transition Matrix

As previously discussed, the Schur stability of a system formed by subspace $\mathbb{A}_\mathsf{f}$ can be analyzed using Algorithm 6 and Algorithm 7. The subspace $\mathbb{A}_\mathsf{f}$ defines the boundaries $\underline{\alpha_\mathsf{f}}_i$ and $\overline{\alpha_\mathsf{f}}_i$ for each $\alpha_{\mathsf{f}i}$. In the following, we will discuss the relationship between the boundary on $\alpha_{\mathsf{f}i}$ and the stationary distribution $\mathbf{p}_\gamma$ and further derive conditions on transition matrix $\mathbf{P_m}$ (λ_0 and λ_1).

Theorem 5.4.3. *For the autonomous system shown in (5.25) with $\boldsymbol{\Omega}_k$ defined in (5.26), the*

system is Lyapunov stable if:

$$\mathrm{P}_\theta^0 \in [\eta_{min}, \eta_{max}] = \bigcap_{i=1}^{d} \left[\underline{\mathrm{P}_\theta^{0,i}}, \overline{\mathrm{P}_\theta^{0,i}}\right],$$

where,

$$\begin{aligned} \underline{\mathrm{P}_\theta^{0,i}} &= \max\left\{\min_{\alpha_{\mathsf{f}i}=\underline{\alpha_{\mathsf{f}_i}},\overline{\alpha_{\mathsf{f}i}}} \mathbb{O}(\alpha_{\mathsf{f}i}), 0\right\} \\ \overline{\mathrm{P}_\theta^{0,i}} &= \min\left\{\max_{\alpha_{\mathsf{f}i}=\underline{\alpha_{\mathsf{f}_i}},\overline{\alpha_{\mathsf{f}i}}} \mathbb{O}(\alpha_{\mathsf{f}i}), 1\right\} \end{aligned} \tag{5.33}$$

and

$$\mathbb{O}(\alpha_{\mathsf{f}i}) = \frac{\alpha_{\mathsf{f}i} - \frac{1}{d-1}\sum\limits_{q_j\in\mathcal{Q}}^{q_j\neq q_i} \mathrm{P}_\delta^{q_j}}{\mathrm{P}_\delta^{q_i} - \frac{1}{d-1}\sum\limits_{q_j\in\mathcal{Q}}^{q_j\neq q_i} \mathrm{P}_\delta^{q_j}}.$$

Proof. By definition,

$$\alpha_{\mathsf{f}i} = \mathrm{P}_\gamma^{q_i} = \mathrm{P}_\theta^0 \mathrm{P}_\delta^{q_i} + \frac{\mathrm{P}_\theta^1}{d-1}\sum_{q_j\in\mathcal{Q}}^{q_j\neq q_i} \mathrm{P}_\delta^{q_j}, \tag{5.34}$$

where the stationary probability $\mathrm{P}_\delta^{q_i}$ for all $q_i \in \mathcal{Q}$ is fixed for a given system. Additionally, because of the constraint $\mathrm{P}_\theta^0 + \mathrm{P}_\theta^1 = 1$, we have:

$$\mathrm{P}_\theta^0 = \frac{\alpha_{\mathsf{f}i} - \frac{1}{d-1}\sum\limits_{q_j\in\mathcal{Q}}^{q_j\neq q_i} \mathrm{P}_\delta^{q_j}}{\mathrm{P}_\delta^{q_i} - \frac{1}{d-1}\sum\limits_{q_j\in\mathcal{Q}}^{q_j\neq q_i} \mathrm{P}_\delta^{q_j}} = \mathbb{O}(\alpha_{\mathsf{f}i}). \tag{5.35}$$

Therefore, given the lower and upper range on $\alpha_{\mathsf{f}i}$, we can solve for the range for P_θ^0. As shown in (5.35), P_θ^0 and $\alpha_{\mathsf{f}i}$ are linearly related, so the boundary for P_θ^0 should be taken on $\underline{\alpha_{\mathsf{f}_i}}$ or $\overline{\alpha_{\mathsf{f}i}}$. Due to the fact that $0 \leq \mathrm{P}_\theta^0 \leq 1$, we get (5.33). For each $\alpha_{\mathsf{f}i}$, there is a

corresponding region $\left[\underline{\mathrm{P}_{\theta}^{0,i}}, \overline{\mathrm{P}_{\theta}^{0,i}}\right]$ and the overall stable region for P_{θ}^{0} should be:

$$[\eta_{min}, \eta_{max}] = \bigcap_{i=1}^{d} \left[\underline{\mathrm{P}_{\theta}^{0,i}}, \overline{\mathrm{P}_{\theta}^{0,i}}\right].$$

□

Given the boundary on the stationary probability as $\eta_{min} \leq \mathrm{P}_{\theta}^{0} \leq \eta_{max}$, the corresponding region for λ_0 and λ_1 (in $\mathbf{P_m}$) can be obtained based on Lemma 5.3.4. The result illustrates conditions on the probabilities λ_0 and λ_1 that are related to the Markovian mode mismatch process. Since both η_{min} and η_{max} are obtained based on the feasible region $\mathbb{A}_{\mathsf{f}}$ and stationary distribution of δ_k, the stable region of λ_0 and λ_1 can be calculated apriori. The previous discussion also implies that if λ_0 and λ_1 satisfies Lemma 5.3.4, the convergence of (5.25) can still be guaranteed for the MJLS even in the presence of mode mismatch.

5.4.5 Boundedness on the Continuous States

Thus far, we have considered the stability of the autonomous system in (5.25). The complete update equation for $\mathbb{E}(\mathbf{e}_k)$ described in (5.23) also contains an input term $\mathbf{\Phi}_k \mathbf{u}_{k-1}$ with $\mathbf{u}_k = \mathbb{E}(\mathbf{x}_k)$. In this subsection, we will consider the boundedness of $\mathbf{\Phi}_k \mathbf{u}_{k-1}$. By definition,

$$\mathbf{\Phi}_k \mathbf{u}_{k-1} = \mathbb{E}_{\gamma_k, \theta_k, \delta_k}\left[\theta_k(\mathbf{\Lambda}_{\gamma_k} - \mathbf{\Gamma}_{\gamma_k, \delta_k})\right]\mathbb{E}(\mathbf{x}_{k-1}). \tag{5.36}$$

For a given system, the matrix $\mathbb{E}_{\gamma_k, \theta_k, \delta_k}\left[\theta_k(\mathbf{\Lambda}_{\gamma_k} - \mathbf{\Gamma}_{\gamma_k, \delta_k})\right]$ is bounded. Therefore, the boundedness of (5.36) fully depends on $\mathbb{E}(\mathbf{x}_k)$. The analysis for stability of $\mathbb{E}(\mathbf{x}_k)$ has been discussed in Chapter 5.3.2 and we will omit the discussion here.

5.5 Experimental Results

In this section, we conduct two experiments base on the previous discussion. One experiment is mode mismatches in an MJLS with two modes and the other is for a generalized MJLS.

5.5.1 Numerical Experiment I

A numerical example is presented to validate our theoretical results. Consider an MJLS in (5.2) with two operational modes $\mathcal{Q} = \{q_1, q_2\}$ and the following system set up:

$$\mathbf{A}_{q_1} = \begin{bmatrix} 0.1571 & 0.1336 & 0.5830 \\ 0.7969 & 0.2882 & 0.0198 \\ 0.1756 & 0.7829 & 0.1454 \end{bmatrix}, \mathbf{B}_{q_1} = \begin{bmatrix} 1 & 0.78 & 0 \\ 0 & 1 & 0.26 \\ 0 & 0.4 & 5 \end{bmatrix},$$

$$\mathbf{A}_{q_2} = \begin{bmatrix} 0.0107 & 0.4652 & 0.5886 \\ 0.7961 & 0.3409 & 0.7606 \\ 0.1415 & 0.0452 & 0.1609 \end{bmatrix}; \mathbf{B}_{q_2} = \begin{bmatrix} 3 & 0 & 0.94 \\ 0 & 4 & 0.11 \\ 1.4 & 0 & 5 \end{bmatrix},$$

$$\mathbf{C}_{q_1} = \begin{bmatrix} 0.0769 & 0.0496 & 0.0340 \\ 0.0006 & 0.0790 & 0.0041 \\ 0.0181 & 0.0493 & 0.0517 \end{bmatrix}, \mathbf{C}_{q_2} = 0.015\mathbf{I}.$$

The transition matrix for discrete state jumps is modeled as follows:

$$\mathbf{P_s} = \begin{bmatrix} 0.2798 & 0.7202 \\ 0.1030 & 0.8970 \end{bmatrix}.$$

According to Theorem 5.3.2 and Theorem 5.3.3, this MJLS is MSS because the matrix $\mathbf{Z}$ as defined in (5.20) satisfies $\rho(\mathbf{Z}) < 1$ and the noise $\mathbf{w}_k$ is WSS. The stationary distribution is $\boldsymbol{\omega}_{\delta^*} = \begin{bmatrix} 0.1251 & 0.8749 \end{bmatrix}$ which falls in the first situation. We can find the stable region of λ_0 and λ_1 using equation (5.13). The obtained region is shown in Figure 5.3.

We choose two sets of values for λ_0 and λ_1: $\{\lambda_0 = 0.9745, \lambda_1 = 0.6\}$ and $\{\lambda_0 = 0.4, \lambda_1 = 0.8\}$ such that the first setting lies in the stable region and the second setting lies outside the region (marked as pentagram and square in Figure 5.3). We then experiment with different mode mismatch probability $\mathbf{P_m}$, and illustrate the evaluation of the actual mean of bias. Based on Monte-Carlo simulations over time $[0, 2000]$, the evolution of bias is plotted in Figure 5.4. We can observe that when λ_0 and λ_1 satisfies the stable condition (5.17), BIBO

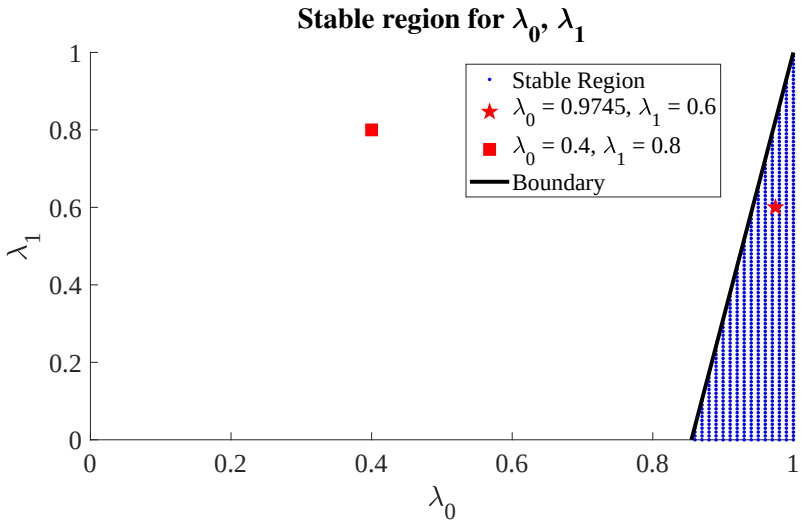


Figure 5.3: *Stable region for* λ_0 *and* λ_1

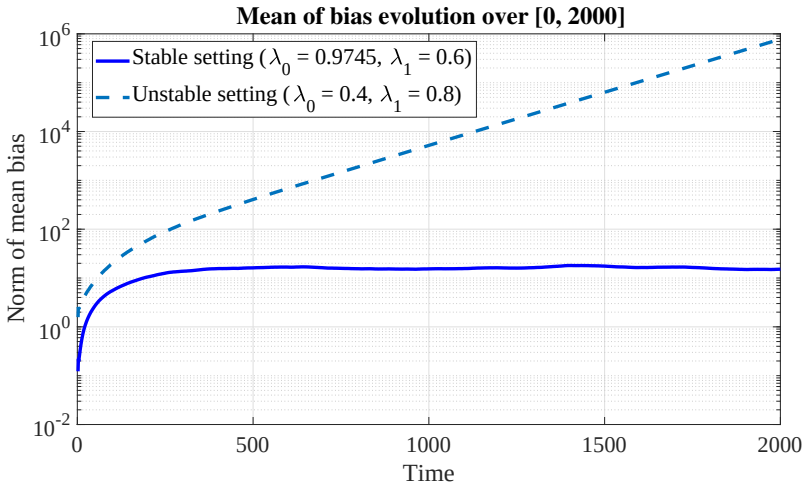


Figure 5.4: *The mean of bias evolution for two different settings on* λ_0 *and* λ_1

stability can be guaranteed. On the other hand, if the choice of λ_0 and λ_1 are outside of the stable region, then the bias can grow exponentially and remain unbounded. These numerical results confirm our analysis and validate the main results introduced in Chapter 5.3.

5.5.2 Numerical Experiment II

In this section, we conduct an experiment to illustrate and verify our main result. We consider an MJLS described as in (5.2) with $\mathbf{x}_k \in \mathbb{R}^2$, $\mathbf{y}_k \in \mathbb{R}^2$ and three modes, e.g.,

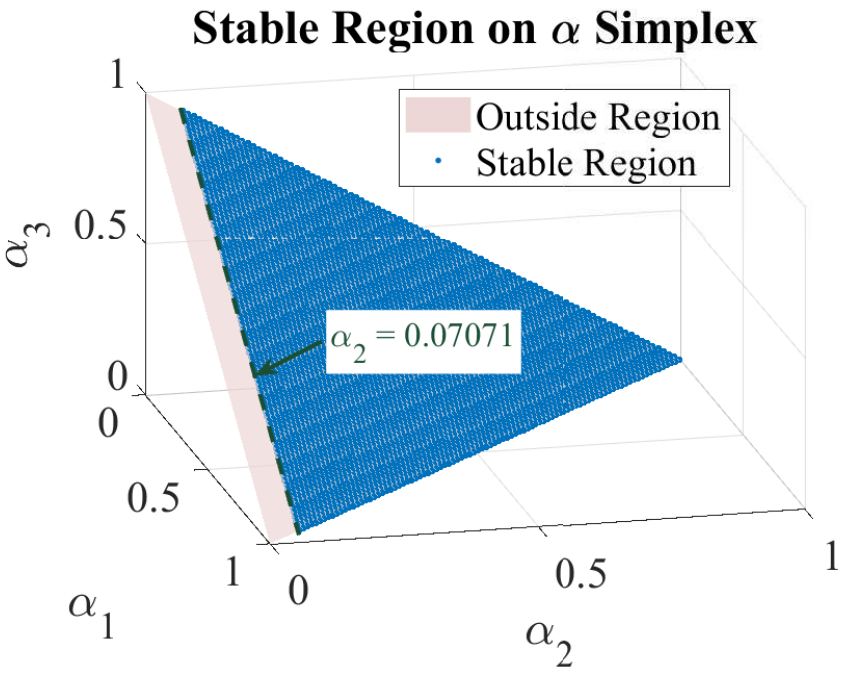


Figure 5.5: *Stable region for* $\boldsymbol{\alpha}$

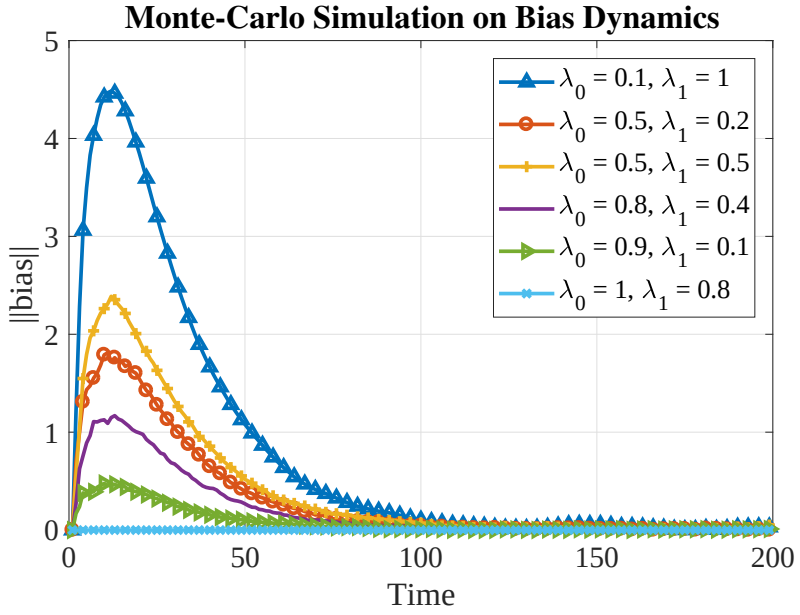


Figure 5.6: *Monte-Carlo simulation for the bias dynamics under different settings of* λ_0 *and* λ_1

$\mathcal{Q} = \{q_1, q_2, q_3\}$. The system matrices are defined as follows.

$$\mathbf{A}_1 = \begin{bmatrix} 0.3395 & 0.4119 \\ 0.3922 & 0.5121 \end{bmatrix}, \mathbf{A}_2 = \begin{bmatrix} 0.7722 & 0.0984 \\ 0.9609 & 0.5098 \end{bmatrix}, \mathbf{A}_3 = \begin{bmatrix} 0.7927 & 0.5183 \\ 0.1458 & 0.5509 \end{bmatrix};$$

$$\mathbf{B}_1 = \begin{bmatrix} 0.3088 & 0.9774 \\ 0.5984 & 0.5340 \end{bmatrix}, \mathbf{B}_2 = \begin{bmatrix} 0.0304 & 0.5629 \\ 0.8077 & 0.4905 \end{bmatrix}, \mathbf{B}_3 = \begin{bmatrix} 0.2991 & 0.3013 \\ 0.2797 & 0.1845 \end{bmatrix};$$

$$\mathbf{C}_1 = \begin{bmatrix} 0.6402 & 0.3258 \\ 0.6477 & 0.7202 \end{bmatrix}, \mathbf{C}_2 = \begin{bmatrix} 0.4140 & 0.0211 \\ 0.3412 & 0.3836 \end{bmatrix}, \mathbf{C}_3 = \begin{bmatrix} 0.8258 & 0.0389 \\ 0.7817 & 0.1139 \end{bmatrix}.$$

The transition matrix for discrete states is

$$\mathbf{P_s} = \begin{bmatrix} 0.3131 & 0.4550 & 0.2319 \\ 0.0939 & 0.1589 & 0.7472 \\ 0.1070 & 0.0077 & 0.8853 \end{bmatrix}.$$

With this transition matrix, we calculate the stationary distribution for this Markov chain is:

$$\mathbf{p}_\delta = \begin{bmatrix} 0.1335 & 0.0794 & 0.7871 \end{bmatrix}.$$

The first step is to obtain matrices $\mathbf{\Lambda}_{q_i}, \forall q_i \in \mathcal{Q}$ and then map the polytope of matrices $\mathbf{\Lambda}_{q_i}$ into an interval matrix $[\overline{\mathbf{\Lambda}}, \underline{\mathbf{\Lambda}}]$. Based on Lemma 5.4.1, we have:

$$\overline{\mathbf{\Lambda}} = \begin{bmatrix} 0.7430 & 0.4844 \\ 0.8739 & 0.5372 \end{bmatrix}, \underline{\mathbf{\Lambda}} = \begin{bmatrix} 0.2979 & 0.0871 \\ 0.1257 & 0.4585 \end{bmatrix}.$$

We then apply algorithm 6 and algorithm 7 to compute the region on $\boldsymbol{\alpha}$ such that the matrix $\mathbf{\Omega}(\boldsymbol{\alpha})$ is Schur stable. As shown in Figure 5.5, the stable region for $\boldsymbol{\alpha}$ is displayed as the

shaded region. The region we calculated is:

$$0 \leq \alpha_{\mathrm{f}1} \leq 0.9192$$

$$0.07071 \leq \alpha_{\mathrm{f}2} \leq 1$$

$$0 \leq \alpha_{\mathrm{f}3} \leq 0.9293$$

It should be noted that for $\boldsymbol{\alpha}$ that falls in the area that is outside the dotted stable region, $\boldsymbol{\Omega}(\boldsymbol{\alpha})$ can be Schur stable or unstable as Theorem 5.4.1 only presents a sufficient condition. For the feasible region on $\boldsymbol{\alpha}$, the corresponding region on P_θ^0 can be obtained via Theorem 5.4.3:

$$0 \leq \mathrm{P}_\theta^{0,1} \leq 1, \quad 0 \leq \mathrm{P}_\theta^{0,2} \leq 1, \quad 0 \leq \mathrm{P}_\theta^{0,2} \leq 1$$

Therefore, the stable range for P_θ^0 is $[0, 1]$. According to Lemma 5.3.4, stationary distribution for Markov chain θ_k is:

$$\lim_{k\to\infty} \mathbf{p}_{\theta_0} {\mathbf{P}_\mathbf{m}}^k = \begin{bmatrix} \frac{1-\lambda_1}{2-\lambda_0-\lambda_1} & \frac{1-\lambda_0}{2-\lambda_0-\lambda_1} \end{bmatrix}.$$

Therefore, we have:

$$0 \leq \frac{1-\lambda_1}{2-\lambda_0-\lambda_1} \leq 1,$$

which results in $0 \leq \lambda_0, \lambda_1 \leq 1$. This result reveals that for this system settings, the bias in a mode-based Kalman filter should always converge for any mode mismatch probabilities λ_0 and λ_1. We then conduct several Monte-Carlo simulations for different combinations of values of λ_0 and λ_1 and track the dynamics of bias. The results shown in Figure 5.6 validates this theoretical analysis and we can see that even for the case that $\lambda_1 = 1$ ($\theta_k = 1$ becomes an absorbing state which means the mode mismatch error always exists), the bias from the Kalman filter still converges.

5.6 Summary

In this chapter, we study the impact of time correlated mode mismatch errors on a mode-based Kalman filter for MJLS state estimation. Specifically, the time correlated mode mismatch error is modeled via a Markov chain and we focus on the resulting estimation bias. Sufficient and necessary conditions regarding the mode mismatch transition probabilities λ_0 and λ_1 are derived for an MJLS with two modes and a sufficient conditions has been presented for a generalized MJLS with arbitrary number of modes. In the next chapter, spatially correlated mode mismatches will be investigated.

Chapter 6

Spatially Correlated Mode Mismatch Errors

In the previous chapters, we have investigated the impact of independent and identically distributed (i.i.d.) Bernoulli distributed and time correlated mode mismatch errors on mode-based Kalman filters. In this chapter, we study the bias dynamics with spatially correlated mode mismatch errors and derive a sufficient condition that guarantees the stability of the bias.

6.1 Introduction

Markov jump linear system (MJLS) [20] is a subclass of stochastic hybrid system (SHS) where the continuous states evolve linearly and discrete modes switch following a Markov chain. MJLS has been used in modeling of cyber-physical systems (CPS) such as microgrid [8], networked control systems [9], etc. State estimation is critical for MJLS if the states are not directly accessible. The quality of state estimation impacts further system analysis and the design of an appropriate control strategy. State estimation in MJLS is dominated by Kalman filter based algorithms. In terms of performance analysis for hybrid system estimation algorithms, due to the complexity of hybrid strategies, there are limited efforts

that have been made in this field. To our best knowledge, only few works have studied the performance of hybrid estimation algorithms [47–49]. Specifically, [47–49] focus on stability of IMM algorithm. [46] studies the effect of mismodeling in Kalman filter for non-hybrid system settings and it derives the mean and covariance matrix for residuals of a mismodeled Kalman filter without analyzing the stability or convergence of the residual. In this chapter, we analyze the performance of a mode-based Kalman filter with mode mismatch errors that are correlated across different modes. The inaccurate mode information will introduce a bias to the estimator as we have discussed in the previous chapters. The focus of this chapter is to derive conditions under which the bias dynamics is statistically convergent. The notion of correlated mode mismatch errors can efficiently capture communication-link failures and spatially correlated cyber-impairments in CPSs. A motivating example is presented as follows.

Network topology error is a typical problem in a smart grid [21]. Consider a conceptual smart grid as shown in Figure 6.1. The system consists of a bank of photovoltaic (PV) panels, three home loads, an electric vehicle and the main electricity grid. As discussed in [7], a smart grid can be well described using a framework of SHS because of the interaction between probabilistic elements and discrete and continuous dynamics. For example, the underlying analog/continuous variables are power consumption, bus voltage, etc. The discrete behaviors are captured by status of switches. The status of switches determines the network topology of the smart grid. As shown in Figure 6.1, there could be communication link failures near switches S_1 and S_2 which in turn result in erroneous status information reported to the topology processor. Consequently, the incorrect status leads to a mode mismatch (which could be spatially correlated as network impairments can impact multiple switch status information) that in turn will introduce a bias in the state estimation for analog variables. For smart grids and many other CPSs that experience similar affects, this work derives conditions for statistical convergence of the bias dynamics of a mode-based Kalman filter estimate.

As stated earlier, mode mismatch is a typical problem in MJLS state estimation. For a general class of SHS, review of the recent literature reveals that there is limited prior

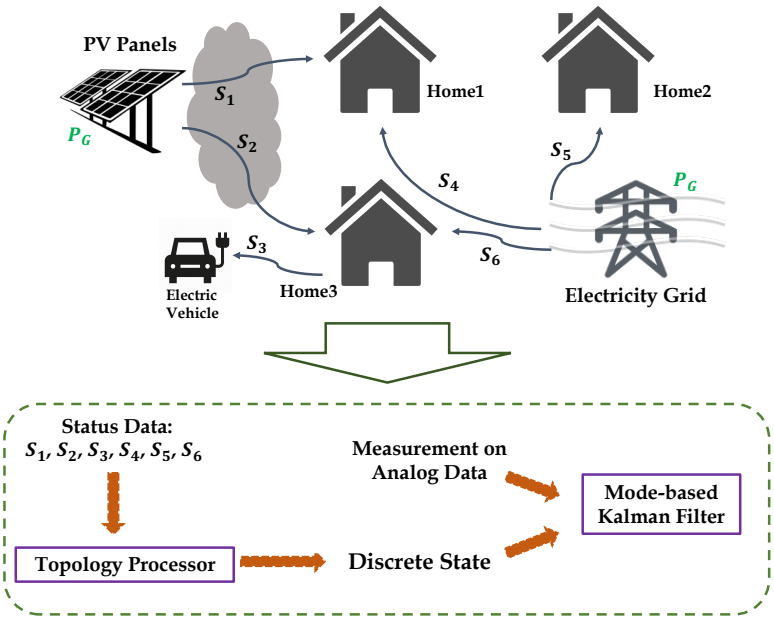


Figure 6.1: *Motivating example: spatial correlated cyber-effects in a smart grid*

work that provides a rigorous performance analysis of SHS estimation techniques [47–49]. Both [47, 48] focus on the convergence of residual for hybrid estimation algorithms IMM. A recent research [49] studies sufficient conditions on error covariance of IMM algorithm for MJLS state estimation. All the above mentioned research works focus on IMM algorithm (composed of multiple mode-based Kalman filters). However, for a general hybrid system estimation setting, the performance of mode-based Kalman filter with presence of mode mismatch errors has not been addressed. This chapter studies the impact of mode mismatch in a mode-based Kalman filter for MJLS estimate. As follow-up of the previous chapters, we consider the case of correlated mode mismatches that can capture spatially correlated cyber-impairments in practical applications. We derive sufficient conditions under which the bias resulting from mode mismatches is statistically convergent. The condition is related to mode mismatch probabilities and it provides guidance on the fidelity of discrete state information needed to sustain the quality of the Kalman filter estimate. The novelty lies in modeling the bias dynamics as a transformed switched system. By leveraging existing results in stability analysis for switched system, we obtain conditions for the convergence of bias.

Furthermore, for the first time, we are able to derive an algebraically solvable condition in terms of the mode mismatch probabilities that guarantees the statistical convergence of the bias.

6.2 System Model

We consider a discrete-time MJLS with a finite discrete state space denoted as $\mathcal{Q} = \{q_1, \cdots, q_d\}$. The discrete state transitions are modeled via a time-homogeneous Markov chain. Let $\delta_k : k \to \mathcal{Q}$ be a switching signal. Then the transition probability corresponds to,

$$\begin{aligned} &\mathbb{P}(\delta_k = q_j | \delta_{k-1} = q_i, \cdots, \delta_0 = q_1) \\ &\qquad = \mathbb{P}(\delta_k = q_j | \delta_{k-1} = q_i) = \pi_{ij}, \end{aligned} \tag{6.1}$$

and the transition matrix is defined as $\mathbf{P_s} = [\pi_{ij}] \in \mathbb{R}^{d \times d}$. The continuous state and measurements evolve as:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}_{\delta_k} \mathbf{x}_{k-1} + \mathbf{B}_{\delta_k} \mathbf{w}_k, \\ \mathbf{y}_k &= \mathbf{C}_{\delta_k} \mathbf{x}_k + \mathbf{v}_k. \end{aligned} \tag{6.2}$$

Here, $\mathbf{x}_k \in \mathbb{R}^n$ is the continuous state and $\mathbf{y}_k \in \mathbb{R}^m$ is the measurement. $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ are both independent Gaussian noise capturing model and measurement uncertainty. Given $\delta_k = q_i \in \mathcal{Q}$, $\mathbf{A}_{\delta_k}$, $\mathbf{B}_{\delta_k}$ and $\mathbf{C}_{\delta_k}$ are $n \times n$, $n \times p$ and $m \times n$ matrices, respectively. We assume the initial continuous state is Gaussian distributed with mean $\boldsymbol{\mu}_0$ and covariance $\boldsymbol{\Sigma}_0$ and the initial discrete mode is q_1.

6.3 Main Results

6.3.1 Correlated Mode Mismatch Errors

In general, state estimation in an SHS is challenging due to the interaction across both discrete and continuous states. In this work, we consider the scenario wherein the continuous state is estimated via a mode-based Kalman filter that is susceptible to correlated mode mismatch errors due to incorrect discrete state information. The correlation considered is between different discrete modes and can result from spatial interaction or cyber-attack. In the following, we will present the generalized error formulation and an illustrative example.

Let the number of correlated-mode groups be $\mathbb{K}$ and the sets of correlated modes be $\mathcal{C} = \{\mathcal{C}_1, \cdots, \mathcal{C}_{\mathbb{K}}\}$. The mode mismatch errors only occur within each set $\mathcal{C}_i$, $\forall \mathcal{C}_i \in \mathcal{C}$. Generally, $\mathcal{C}_i$ and $\mathcal{C}_j$ are mutually exclusive for all distinguished i, j. Define an operator $\mathbb{C} : \mathcal{Q} \to \mathcal{C}$ that returns the correlated group that the state q_i belongs to, i.e., $\mathbb{C}(q_i) = \mathcal{C}_l$ if $q_i \in \mathcal{C}_l$. Let $\gamma_k : k \to \mathcal{Q}$ denote the estimated mode. Note that γ_k is only correlated with δ_k and it is independent with continuous state $\mathbf{x}_k$, $\forall k$. We define the conditional probability of mode mismatch as:

$$\mathbb{P}(\gamma_k = q_j | \delta_k = q_i) = \begin{cases} \lambda_{q_j, q_i} & \exists \mathcal{C}_l \in \mathcal{C} \text{ s.t. } q_i, q_j \in \mathcal{C}_l, \\ 0 & \nexists \mathcal{C}_l \in \mathcal{C} \text{ s.t. } q_i, q_j \in \mathcal{C}_l. \end{cases}$$

Note that for any $q_i \in \mathcal{Q}$, $\sum_{q_j \in \mathcal{Q}} \lambda_{q_j, q_i} = 1$. Here, we assume the probabilities of mode mismatch are the same within each correlated group. That is, $\forall q_i, q_j \in \mathcal{C}_l$, we define

$$\lambda_{q_i, q_j} = \begin{cases} \lambda_{\mathcal{C}_l}, & q_i = q_j \\ \frac{1 - \lambda_{\mathcal{C}_l}}{|\mathcal{C}_l| - 1}, & q_i \neq q_j \end{cases}$$

with $|\mathcal{C}_l|$ represents the number of elements in set $\mathcal{C}_l$. While this assumption will be relaxed in our future work, we believe that this is a reasonable first step as it is a reflection of the correlation that exists within a group. As an illustrative example, Figure 6.2 shows

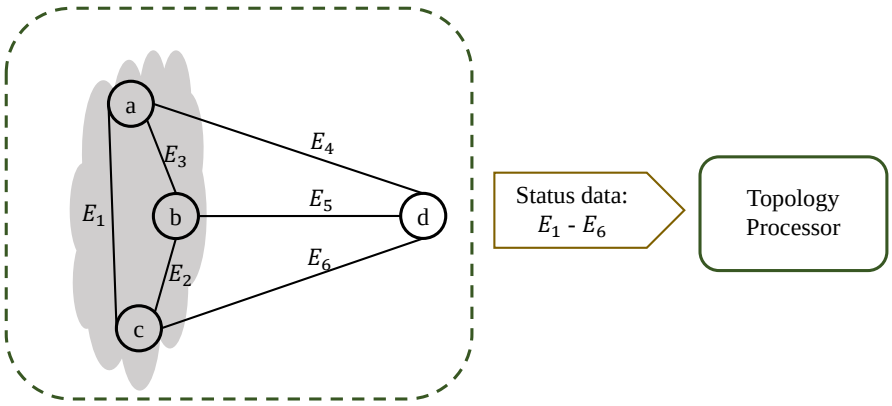


Figure 6.2: *An example of correlated nodes*

a networked system with four nodes (a,b,c and d). Each edge is either connected (1) or disconnected (0). Consequently, the topology of the network changes based on the status of the edges. Let the discrete state represents the topology of the network, i.e., the discrete state space consists of 64 modes corresponding to different status of each edge. Suppose the local network around nodes a,b and c is exposed to attack such that the actual status of E_1, E_2 and E_3 transmitted to the topology processor could be erroneous resulting in correlated mode mismatches. Since there is no error in E_4, E_5 and E_6, and their status contributes eight distinct topologies, i.e., $\mathbb{K} = 8$. Each of these eight groups of states can be mathematically represented as:

$$\mathcal{C}_i = \{q | E_4 = i_1, E_5 = i_2, E_6 = i_3, \forall E_1, E_2, E_3\},$$

with i_1, i_2, $i_3 \in \{0, 1\}$.

6.3.2 Mode-Based Kalman Filter

As the continuous states are not directly observable in system (6.2), a mode-based Kalman filter can be used to estimate the continuous state $\mathbf{x}_k$ based on the measurement and discrete states up to time k. Denote the measurement sequence and mode sequence up to time k as $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$ and $\boldsymbol{\gamma}_k^s = (\gamma_1, \cdots, \gamma_k)$, respectively. The mode-based Kalman filter equations for SHS (6.2) are given in Algorithm 8. For a hybrid system, the mode-based

Algorithm 8 Mode-based Kalman filter

function ESTIMATION UPDATE($\boldsymbol{\mu}_0, \mathbf{M}_{0|0}, \mathbf{Q}, \mathbf{R}, \boldsymbol{\gamma}_k^s, \mathbf{y}_k^s$)
  $\mathbf{x}_{0|0} = \boldsymbol{\mu}_0, \mathbf{M}_{0|0} = \boldsymbol{\Sigma}_0$
  $\mathbf{y}_k^s = (\mathbf{y}_1, \cdots, \mathbf{y}_k)$
  $\boldsymbol{\gamma}_k^s = (\gamma_1, \cdots, \gamma_k)$
  for $i = 1 : k$ **do**
    $\mathbf{x}_{i|i-1} = \mathbf{A}_{\gamma_i}\mathbf{x}_{i-1|i-1}$
    $\mathbf{M}_{i|i-1} = \mathbf{A}_{\gamma_i}\mathbf{M}_{i-1|i-1}\mathbf{A}'_{\gamma_i} + \mathbf{B}_{\gamma_i}\mathbf{Q}\mathbf{B}'_{\gamma_i}$
    $\mathbf{K}_{\gamma_i,i} = \mathbf{M}_{i|i-1}\mathbf{C}'_{\gamma_i}(\mathbf{C}_{\gamma_i}\mathbf{M}_{i|i-1}\mathbf{C}'_{\gamma_i} + \mathbf{R})^{-1}$
    $\mathbf{x}_{i|i} = \mathbf{x}_{i|i-1} + \mathbf{K}_{i,\gamma_i}(\mathbf{y}_i - \mathbf{C}_{\gamma_i}\mathbf{x}_{i|i-1})$
    $\mathbf{M}_{i|i} = (\mathbf{I} - \mathbf{K}_{\gamma_i,i}\mathbf{C}_{\gamma_i})\mathbf{M}_{i|i-1}$
  end for
  return $\mathbf{x}_{k|k}$
end function

Kalman filter is unbiased only if $\boldsymbol{\gamma}_k^s = \boldsymbol{\delta}_k^s$ where $\boldsymbol{\delta}_k^s = (\delta_1, \cdots, \delta_k)$. Similar as previous chapters, we define the bias to be the difference between means of estimator and the true state, i.e., $\mathbf{e}_k = \mathbb{E}(\hat{\mathbf{x}}_k) - \mathbb{E}(\mathbf{x}_k)$. It has been shown in Chapter 4 and 5 that the bias dynamics is:

$$\begin{aligned}\mathbf{e}_k =& (\mathbf{A}_{\gamma_k} - \mathbf{K}_{\gamma_k}\mathbf{C}_{\gamma_k}\mathbf{A}_{\gamma_k})\mathbf{e}_{k-1} \\ &+ (\mathbf{A}_{\gamma_k} - \mathbf{K}_{\gamma_k}\mathbf{C}_{\gamma_k}\mathbf{A}_{\gamma_k} + \mathbf{K}_{\gamma_k}\mathbf{C}_{\delta_k}\mathbf{A}_{\delta_k} - \mathbf{A}_{\delta_k})\mathbb{E}(\mathbf{x}_{k-1}). \end{aligned} \tag{6.3}$$

Here, $\mathbf{K}_{\gamma_k}$ is the steady Kalman gain that corresponds to the mode $\gamma_k \in \mathcal{Q}$. Since we assume that $\mathbf{Q}$ and $\mathbf{R}$ are constants for all modes, the Kalman gain will converge to the corresponding steady Kalman gain $\mathbf{K}_{\gamma_k}$ quickly [63]. For hybrid system, it is reasonable to approximate the time-variant Kalman gain using the steady Kalman gain in practice as the modes switches are much infrequent than the evolution of continuous dynamics. This idea has also been used in sub-optimal control problems for decades. As it only involves an off-line calculation, we use $\mathbf{K}_{\gamma_k}$ for approximation in equation (6.3). It should be mentioned that the bias dynamics analysis is also applicable for a more general class of MJLS where the Markov chain is nonhomogeneous [96]. In fact, the derivation of the dynamics of the bias resulting from mode mismatch errors is not restricted to a specific model for discrete state transitions. We focus on time-homogeneous Markov chain model since the stability of

such systems is well studied and we are able to derive an algebraically solvable condition. As the evolution of $\mathbf{e}_k$ in (6.3) depends on random variable γ_k and δ_k, the process $\{\mathbf{e}_k\}_{k=0}^{\infty}$ is a stochastic process. Analyzing the statistical convergence of $\mathbf{e}_k$ becomes challenging since γ_k is correlated with δ_k. In the following, we will introduce a transformed switched system and analyze the convergence of $\mathbf{e}_k$.

6.3.3 Transformed Switched System

Depending on the actual mode δ_k, the evolution of $\mathbf{e}_k$ varies. Conditioning upon δ_k, the error process is:

$$\begin{aligned}\mathbf{e}_k|\delta_k = (\mathbf{A}_{\gamma_k|\delta_k} - &\mathbf{K}_{\gamma_k|\delta_k}\mathbf{C}_{\gamma_k|\delta_k}\mathbf{A}_{\gamma_k|\delta_k})\mathbf{e}_{k-1} \\ &+(\mathbf{A}_{\gamma_k|\delta_k} - \mathbf{K}_{\gamma_k|\delta_k}\mathbf{C}_{\gamma_k|\delta_k}\mathbf{A}_{\gamma_k|\delta_k} + \mathbf{K}_{\gamma_k|\delta_k}\mathbf{C}_{\delta_k}\mathbf{A}_{\delta_k} - \mathbf{A}_{\delta_k})\mathbb{E}(\mathbf{x}_{k-1}).\end{aligned}$$

Given the actual mode δ_k, the stochastic process $\{\mathbf{e}_k|\delta_k\}_{k=0}^{\infty}$ is a function of γ_k. Here, $\mathbf{A}_{\gamma_k}$, $\mathbf{C}_{\gamma_k}$ and $\mathbf{K}_{\gamma_k}$ are all random matrices that take value in the space of $\left\{\mathbf{A}_{q_i}\right\}_{q_i\in\mathcal{Q}}$, $\left\{\mathbf{C}_{q_i}\right\}_{q_i\in\mathcal{Q}}$ and $\left\{\mathbf{K}_{q_i}\right\}_{q_i\in\mathcal{Q}}$, respectively. The randomness in $\mathbf{A}_{\gamma_k}$, $\mathbf{C}_{\gamma_k}$ and $\mathbf{K}_{\gamma_k}$ results only from γ_k. Define matrices $\mathbf{\Lambda}_{\gamma_k} = \mathbf{A}_{\gamma_k} - \mathbf{K}_{\gamma_k}\mathbf{C}_{\gamma_k}\mathbf{A}_{\gamma_k}$ and $\mathbf{\Gamma}_{\gamma_k,\delta_k} = \mathbf{A}_{\delta_k} - \mathbf{K}_{\gamma_k}\mathbf{C}_{\delta_k}\mathbf{A}_{\delta_k}$, then we can write:

$$\mathbf{e}_k|\delta_k = \mathbf{\Lambda}_{\gamma_k|\delta_k}\mathbf{e}_{k-1} + (\mathbf{\Lambda}_{\gamma_k|\delta_k} - \mathbf{\Gamma}_{\gamma_k,\delta_k|\delta_k})\mathbb{E}(\mathbf{x}_{k-1}). \tag{6.4}$$

It should be noted that $\mathbf{e}_k|\delta_k$ is bounded with probability 1 if and only if $\mathbb{E}(\mathbf{e}_k|\delta_k)$ is bounded [39]. Using the tower rule and taking expectation over $\mathbf{e}_{k-1}$ and γ_k, we get:

$$\mathbb{E}(\mathbf{e}_k|\delta_k) = \sum_{q_j\in\mathbb{C}(\delta_k)} [\lambda_{q_j,\delta_k}\mathbf{\Lambda}_{q_j}\mathbb{E}(\mathbf{e}_{k-1}) + \lambda_{q_j,\delta_k}(\mathbf{\Lambda}_{q_j} - \mathbf{\Gamma}_{q_j,\delta_k})\mathbb{E}(\mathbf{x}_{k-1})]. \tag{6.5}$$

Based on the definition of conditional expectation, $\mathbb{E}(\mathbf{e}_k|\delta_k)$ is a random variable on the same space of δ_k, i.e., $\mathcal{Q}$. As δ_k denotes the true mode which follows a Markov chain, we can interpret the dynamics of $\mathbb{E}(\mathbf{e}_k|\delta_k)$ as a transformed MJLS with discrete state δ_k. To

distinguish the transformed SHS from the original system, we use $\mathbf{x}_k^*$ to denote the continuous states and use $\mathbf{u}_k$ to denote the input. The transformed state evolution is now described as:

$$\mathbf{x}_k^* = \mathbf{F}_{\delta_k}\mathbf{x}_{k-1}^* + \mathbf{G}_{\delta_k}\mathbf{u}_{k-1}. \tag{6.6}$$

Here, $\mathbf{x}_k^* = \mathbb{E}(\mathbf{e}_k|\delta_k)$, $\mathbf{u}_k = \mathbb{E}(\mathbf{x}_k)$, and

$$\mathbf{F}_{\delta_k} = \lambda_{\mathbb{C}(\delta_k)}\mathbf{\Lambda}_{\delta_k} + \frac{1-\lambda_{\mathbb{C}(\delta_k)}}{|\mathbb{C}(\delta_k)|-1}\sum_{\substack{q_j\in\mathbb{C}(\delta_k)\\ q_j\neq\delta_k}}\mathbf{\Lambda}_{q_j},$$

$$\mathbf{G}_{\delta_k} = \frac{1-\lambda_{\mathbb{C}(\delta_k)}}{|\mathbb{C}(\delta_k)|-1}\sum_{\substack{q_j\in\mathbb{C}(\delta_k)\\ q_j\neq\delta_k}}(\mathbf{\Lambda}_{q_j} - \mathbf{\Gamma}_{q_j,\delta_k}).$$

Our goal is to find conditions under which $\mathbf{x}_k^*$ converges. With the transformed MJLS, the problem is equivalent to analyzing the stability of the transformed MJLS. The challenge lies in the uncertainties corresponding to the probability $\lambda_{\mathbb{C}(\delta_k)}$ in system matrices $\mathbf{F}_{\delta_k}$ and $\mathbf{G}_{\delta_k}$. To further address the impact of $\lambda_{\mathbb{C}(\delta_k)}$ on the stability of (6.6), we will leverage the results that have been established regarding the stability of MJLS and derive an algebraically solvable condition on $\lambda_{\mathbb{C}(\delta_k)}$ that guarantees the convergence of (6.6).

6.3.4 Stability for Markov Jump Linear Systems

Numerous concepts of stability have been defined for SHS. Here, we use the notion of stochastic second moment stability (also referred as mean square stability) [79, 80]. An autonomous system

$$\mathbf{x}_k^* = \mathbf{F}_{\delta_k}\mathbf{x}_{k-1}^*, \tag{6.7}$$

is *stochastically second moment stable* if for any initial distribution of $\mathbf{x}_0^*$ and δ_0,

$$\lim_{k\to\infty}\mathbb{E}(\|\mathbf{x}_k^*(\mathbf{x}_0^*,\delta_0)\|^2) = 0.$$

The transformed MJLS (6.6) can be treated as an input-output system by defining the measurement $\mathbf{y}_k^* = \mathbf{x}_k^*$. Since the stability of an input-output system is closely related to the stability of the corresponding autonomous system, it is conventional to consider the stability of the autonomous system (6.7) first. Among the existing approaches to study stability, the main approach is primarily built on the well-known Lyapunov theory. A necessary and sufficient condition for system (6.7) with finite Markov chain $\{\delta_k\}$ is presented in [81]. A condition using Kronecker product provides a testable condition and our main analysis is built on this result. The following theorem can be found in [80].

Theorem 6.3.1. *The system in (6.7) is stochastically second moment stable if and only if the matrix* $\mathbf{Z}$ *is Schur stable, i.e.,* $\rho(\mathbf{Z}) < 1$ *where* $\rho(\cdot)$ *is the spectral radius of a matrix and* $\mathbf{Z}$ *is defined as:*

$$\mathbf{Z} = diag[\mathbf{F}_{q_1} \otimes \mathbf{F}_{q_1}, \cdots, \mathbf{F}_{q_d} \otimes \mathbf{F}_{q_d}] \cdot (\mathbf{P_s}' \otimes \mathbf{I}). \tag{6.8}$$

Before we develop the main result, the following lemma [97] needs to be mentioned since it is essential for proof of the main result.

Lemma 6.3.1. *Let* $\mathbf{Y} = [\mathbf{Y}_{ij}]$ *be a square block matrix, and* $\mathbf{Z} = [\|\mathbf{Y}_{ij}\|]$. *Then* $\rho(\mathbf{Y}) \leq \rho(\mathbf{Z})$.

Theorem 6.3.2. *The matrix* $diag[\mathbf{F}_{q_1} \otimes \mathbf{F}_{q_1}, \cdots, \mathbf{F}_{q_d} \otimes \mathbf{F}_{q_d}]$ *is Schur stable if for all* $i \in [1, d]$, *there exists* $\lambda_{\mathbb{C}(q_i)} \in [0, 1]$ *such that* $\mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] < 1$. *The function* $\mathbb{f}_i(x)$ *is defined as*

$$\mathbb{f}_i(x) = (K_{2i}S_i^2 - K_{1i}S_i + K_{3i})x^2 + (K_{1i}S_i - 2K_{2i}S_i^2)x + K_{2i}S_i^2,$$

where,

$$K_{1i} = \left\| \mathbf{\Lambda}_{q_i} \otimes \left(\sum_{\substack{q_j \in \mathbb{C}(q_i) \\ q_j \neq q_i}} \mathbf{\Lambda}_{q_j} \right) + \left(\sum_{\substack{q_j \in \mathbb{C}(q_i) \\ q_j \neq q_i}} \mathbf{\Lambda}_{q_j} \right) \otimes \mathbf{\Lambda}_{q_i} \right\|,$$

$$K_{2i} = \left\| \left(\sum_{\substack{q_j \in \mathbb{C}(q_i) \\ q_j \neq q_i}} \mathbf{\Lambda}_{q_j} \right) \otimes \left(\sum_{\substack{q_j \in \mathbb{C}(q_i) \\ q_j \neq q_i}} \mathbf{\Lambda}_{q_j} \right) \right\|,$$

$$K_{3i} = \left\| \mathbf{\Lambda}_{q_i} \otimes \mathbf{\Lambda}_{q_i} \right\|, \quad S_i = \frac{1}{|\mathbb{C}(q_i)| - 1}.$$

Proof. Let $\mathbf{H}_i = \mathbf{F}_{q_i} \otimes \mathbf{F}_{q_i}$. From the definition,

$$\begin{aligned} \mathbf{H}_i = & \frac{\lambda_{\mathbb{C}(q_i)}(1 - \lambda_{\mathbb{C}(q_i)})}{|\mathbb{C}(q_i)| - 1} \left[\mathbf{\Lambda}_{q_i} \otimes \left(\sum_{\substack{q_j \in \mathbb{C}(q_i) \\ q_j \neq q_i}} \mathbf{\Lambda}_{q_j} \right) + \left(\sum_{\substack{q_j \in \mathbb{C}(q_i) \\ q_j \neq q_i}} \mathbf{\Lambda}_{q_j} \right) \otimes \mathbf{\Lambda}_{q_i} \right] \\ & + \left(\frac{1 - \lambda_{\mathbb{C}(q_i)}}{|\mathbb{C}(q_i)| - 1} \right)^2 \left(\sum_{\substack{q_j \in \mathbb{C}(q_i) \\ q_j \neq q_i}} \mathbf{\Lambda}_{q_j} \right) \otimes \left(\sum_{\substack{q_j \in \mathbb{C}(q_i) \\ q_j \neq q_i}} \mathbf{\Lambda}_{q_j} \right) + \lambda^2_{\mathbb{C}(q_i)} \mathbf{\Lambda}_{q_i} \otimes \mathbf{\Lambda}_{q_i}. \end{aligned}$$

Then we have:

$$\begin{aligned} \|\mathbf{H}_i\| \leq & K_{1i} S_i \lambda_{\mathbb{C}(q_i)} [1 - \lambda_{\mathbb{C}(q_i)}] \\ & + K_{2i} S_i^2 [1 - \lambda_{\mathbb{C}(q_i)}]^2 + K_{3i} \lambda^2_{\mathbb{C}(q_i)} = \mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] \end{aligned} \tag{6.9}$$

The spectral radius of a diagonal matrix composed of $\|\mathbf{H}_i\|$ is:

$$\rho(diag[\|\mathbf{H}_1\|, \cdots, \|\mathbf{H}_d\|]) = \max_{i \in [1,d]} \|\mathbf{H}_i\|.$$

Therefore, $\rho(diag[\|\mathbf{H}_1\|, \cdots, \|\mathbf{H}_d\|]) < 1$ if and only if $\|\mathbf{H}_i\| < 1$ for all $i \in [1, d]$. According to (6.9), the sufficient condition for each $\|\mathbf{H}_i\| < 1$ is $\mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] < 1$, $\forall i$.

From Lemma 6.3.1, we can write:

$$\rho(diag[\mathbf{H}_1, \cdots, \mathbf{H}_d]) \leq \rho(diag[\|\mathbf{H}_1\|, \cdots, \|\mathbf{H}_d\|]). \tag{6.10}$$

As a result of (6.9) and (6.10), we can conclude that $\mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] < 1$, $\forall i$ is also a sufficient condition for $\rho(diag[\mathbf{H}_1, \cdots, \mathbf{H}_d]) < 1$. Finally, since $\lambda_{\mathbb{C}(q_i)}$ is the probability of no mode mismatch, it needs to satisfy $0 \leq \lambda_{\mathbb{C}(q_i)} \leq 1$. Therefore, if for all $i \in [1, d]$, there exists $\lambda_{\mathbb{C}(q_i)} \in [0, 1]$ such that $\mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] < 1$, then $\rho(diag[\mathbf{H}_1, \cdots, \mathbf{H}_d]) < 1$. □

Based on Theorem 6.3.1 and 6.3.2, a sufficient condition for Schur stability of system in (6.7) can be obtained as the following corollary.

Corollary 6.3.1.1. *The system in (6.7) is stochastically second moment stable if* $\mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] < 1$, $\forall i \in [1, d]$.

Proof. From the property of Kronecker product, we have

$$\rho(\mathbf{P_s}' \otimes \mathbf{I}) = \rho(\mathbf{P_s}')\rho(\mathbf{I}) = \rho(\mathbf{P_s}). \tag{6.11}$$

The following inequality holds:

$$\begin{aligned}\rho(\mathbf{Z}) &\overset{(a)}{\leq} \rho(diag[\mathbf{H}_1, \cdots, \mathbf{H}_d])\rho(\mathbf{P_s}' \otimes \mathbf{I}) \\ &\overset{(b)}{\leq} \max_{i\in[1,d]} \mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}]\rho(\mathbf{P_s}) \overset{(c)}{\leq} \max_{i\in[1,d]} \mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}].\end{aligned}$$

(a) is true because of Gelfand's formula and (b) results from (6.9) and (6.10). (c) is based on the property of a transition matrix, i.e., $\rho(\mathbf{P_s}) = 1$. Therefore, if for all i, $\mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] < 1$, then $\rho(\mathbf{Z}) < 1$. From Theorem 6.3.1, $\rho(\mathbf{Z})$ being Schur stable guarantees stochastic second moment stability of the MJLS represented by (6.7). □

Thus far, we have derived the sufficient conditions for stochastic second moment stability of the autonomous MJLS (6.7). It has been proved that stability of the autonomous system results in bounded-input bounded-output (BIBO) stability, and this result can be extended

to the transformed switched system [12, 75]. It is worth pointing out that in our transformed MJLS (6.6), the input $\mathbf{u}_k$ is not a control signal but a term that depends on the expectation of the continuous states $\mathbf{x}_k$. Therefore, if the original system is second moment stable, then $\mathbf{u}_k$ is bounded which leads to BIBO stability of (6.6). In the following, we present conditions for the stochastic second moment stability of the original SHS in (6.2). This theorem is introduced in [20] with a detailed proof.

Theorem 6.3.3. *The following statements are equivalent:*
(i) The MJLS $\mathbf{x}_k = \mathbf{A}_{\delta_k}\mathbf{x}_{k-1}$ *is stochastically second moment stable;*
(ii) The MJLS $\mathbf{x}_k = \mathbf{A}_{\delta_k}\mathbf{x}_{k-1} + \mathbf{B}_{\delta_k}\mathbf{w}_k$ *is stochastically second moment stable if* $\mathbf{w}_k$ *is second order independent wide sense stationary (WSS).*

As we assume the system noise $\mathbf{w}_k$ is Gaussian with mean $\mathbf{0}$ and covariance $\mathbf{Q}$, analyzing stability of (6.2) is the same as the stability analysis of (6.7). Therefore, Theorem 6.3.1 can be applied to check stochastic second moment stability of (6.2).

6.3.5 Discussion on the Convergence of Bias Dynamics

With all the previously introduced results, we present our main result in the following theorem.

Theorem 6.3.4. *For the state estimate of MJLS in (6.2), the bias dynamics resulting from mode mismatch will converge if the following conditions hold:*
(i) The matrix $diag[\mathbf{A}_{q_1} \otimes \mathbf{A}_{q_1}, \cdots, \mathbf{A}_{q_d} \otimes \mathbf{A}_{q_d}] \cdot (\mathbf{P_s}' \otimes \mathbf{I})$ *is Schur stable;*
(ii) For all $i \in [1, d]$*, the probability* $\lambda_{\mathbb{C}(q_i)}$ *satisfies* $\mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] < 1$.

Proof. Condition (i) implies that the MJLS (6.2) is stochastically second moment stable based on Theorem 6.3.1. The second condition indicates that the autonomous MJLS (6.7) is stochastically second moment stable according to Theorem 6.3.2 and Corollary 6.3.1.1. Both conditions together guarantee the BIBO stability of the bias dynamics in (6.6). □

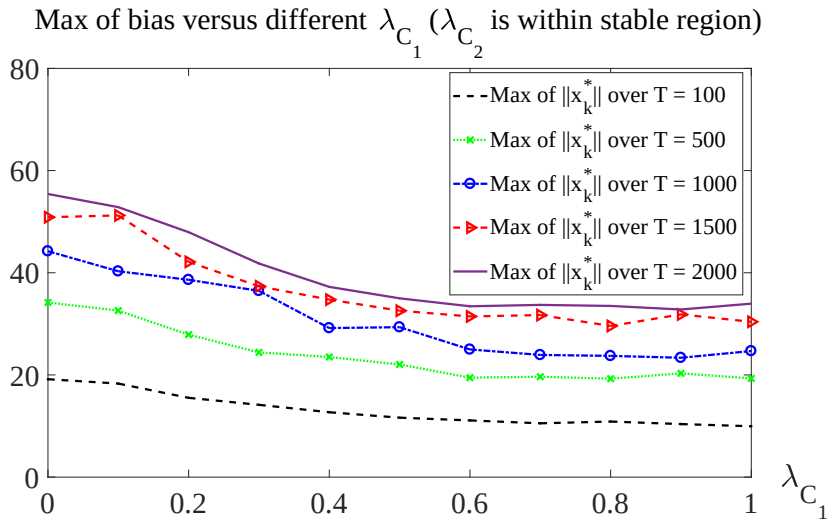


Figure 6.3: *Max of* $\|\mathbf{x}_k^*\|$ *versus* $\lambda_{\mathcal{C}_1}$ *when* $\lambda_{\mathcal{C}_2} = 0.4$ *(stable region)*

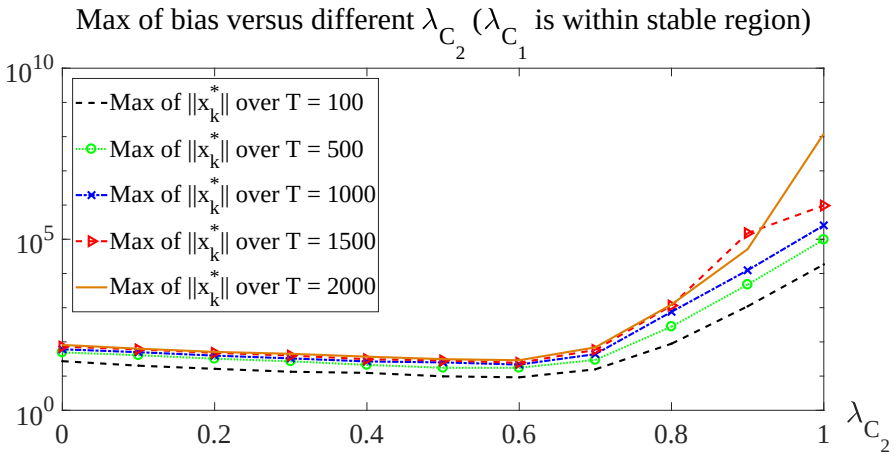


Figure 6.4: *Max of* $\|\mathbf{x}_k^*\|$ *versus* $\lambda_{\mathcal{C}_2}$ *when* $\lambda_{\mathcal{C}_1} = 0.5$ *(stable region)*

Theorem 6.3.2 reveals the region of $\lambda_{\mathcal{C}_l}$ that guarantees convergence of the bias dynamics. Evaluating this region only requires solving a series of quadratic equations. This main result gives us insights on the fidelity of discrete state knowledge required for convergence of bias.

6.4 Experimental Results

In this section, we consider a three dimensional MJLS with five discrete modes. Due to the underlying spatial correlation, the system contains two correlated mode groups:

$\mathcal{C}_1 = \{q_1, q_5\}$, $\mathcal{C}_2 = \{q_2, q_3, q_4\}$. As discussed in the motivating example (Chapter 6.1), this simulation setup can be used to describe a smart grid with continuous states being:

$$\mathbf{x} = \begin{bmatrix} P_G & P_D & \triangle P \end{bmatrix}',$$

where, P_G is the power generation, P_D is the power demand and $\triangle P$ denotes power balance. Depending on the status of switches within the network, there are five network topologies captured by $q_1, \cdots, q_5$. The two topologies q_1 and q_5 in correlated group $\mathcal{C}_1$ correspond to errors in status of one of the switches. Similarly, the three topologies in group $\mathcal{C}_2$ correspond to errors in status of another switch with three settings.

According to Corollary 6.3.1.1, if $\mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] < 1$, $\forall i$, then system in (6.7) is stochastically second moment stable. By solving the quadratic inequality $\mathbb{f}_i[\lambda_{\mathbb{C}(q_i)}] < 1$, we get the values of $\lambda_{\mathcal{C}_1}$ and $\lambda_{\mathcal{C}_2}$ that guarantee convergence of the bias dynamics as $\lambda_{\mathcal{C}_1} \in [0.0111, 0.9890]$ and $\lambda_{\mathcal{C}_2} \in [0, 0.4672]$. Note that the ranges we obtain are sufficient conditions and not necessary. Therefore, while the system will definitely be stable within the region, it may not be unstable outside it.

Since there are two correlated groups and mode mismatch errors $\lambda_{\mathcal{C}_1}$ and $\lambda_{\mathcal{C}_2}$ correspond to each group, respectively, we conduct the simulation by fixing one probability and validate the stability region on another. Figure 6.3 shows the actual $\|\mathbf{x}_k^*\|$ by implementing Monte-Carlo simulation of the system for different values of $\lambda_{\mathcal{C}_1}$ while fixing $\lambda_{\mathcal{C}_2} = 0.4$ which is inside the theoretically derived stable region. We can observe that among $N = 1000$ Monte-Carlo simulations, max of $\|\mathbf{x}_k^*\|$ only slightly increases as simulation time T increases but does not diverge for any choice of $\lambda_{\mathcal{C}_1} \in [0.0111, 0.9890]$. This observation validates the calculated stability region $[0.0111, 0.9890]$ for $\lambda_{\mathcal{C}_1}$. Given $\lambda_{\mathcal{C}_1} = 0.5$ (within the stable region), Figure 6.4 presents the maximum of $\|\mathbf{x}_k^*\|$ over different values of $\lambda_{\mathcal{C}_2}$. In Figure 6.4, we can conclude that the bias dynamics are stable for $\lambda_{\mathcal{C}_2}$ within $[0, 0.4672]$. For $\lambda_{\mathcal{C}_2} \geq 0.7$, the maximum of $\|\mathbf{x}_k^*\|$ diverges rapidly which is consistent with our theoretical result. It is also worth pointing out that ideally, larger the values of $\lambda_{\mathcal{C}_1}$ and $\lambda_{\mathcal{C}_2}$, better the quality of mode information indicating that the bias should be smaller. However, in this particular

system setup, function $\mathbb{f}_4(\lambda)$ dominates the choice of $\lambda_{\mathcal{C}_2}$. As $\mathbb{f}_4(\lambda)$ is a monotonic increasing function in $[0, 1]$, a larger choice of $\lambda_{\mathcal{C}_2}$ leads to $\mathbb{f}_4(\lambda) > 1$.

6.5 Summary

This chapter investigates statistical convergence of bias in a mode-based Kalman filter used for MJLS state estimation in the presence of mode mismatch errors with focus on mode mismatch errors that are correlated across modes. This situation is commonly encountered in many practical systems where mode mismatch errors result due to spatially correlated cyber-impairments. We derive an algebraically solvable condition that ensures the convergence of bias. The theoretical results are also validated via simulations on a MJLS as a numerical example. The main result from the work can serve as a guidance to understand the desired fidelity in discrete state knowledge for maintaining continuous state estimation quality.

Chapter 7

Conclusion and Future Work

This chapter concludes the dissertation with a summary of the research results and possible future directions for this research.

7.1 Summary

In this dissertation, we investigate state estimation strategies for stochastic hybrid systems (SHS) and evaluate the performance of a mode-based Kalman filter in the presence of mode mismatch errors. Due to the fact that the continuous state and discrete state are interacting in an SHS, inaccurate information of discrete state will impact the continuous state estimation quality and introduce a bias in the mode-based Kalman filter. With this background, this dissertation addresses four research questions that represent challenges in modeling and situational awareness in an SHS. Motivated by network topology errors in a smart grid, the four research questions can effectively capture system failure, cyber-attacks or communication link impairments in a cyber physical system (CPS). In this regard, we summary the contributions of this dissertation as follows:

- We propose a new state estimation algorithm for state-dependant SHS (SDSHS). Specifically, we consider an SDSHS with quadratic guard conditions. In the proposed strategy, only one Kalman filter is needed and the discrete state estimate is derived

based on distribution for the guard condition and the distribution is related to the estimated continuous state. Based on this distribution, we propose a threshold framework to decide whether a discrete transition occurs or not. The proposed approach results in an extremely low error rate for the discrete state estimates even when transitions are frequent. We demonstrate the algorithm with simulations on a robot motion system and the result illustrates a superior error performance.

- We study the impact of mode mismatch errors on state estimation for an SHS with mode-based Kalman filter. It is proved that the mode mismatches introduce a bias to the mode-based Kalman filter and we derive the bias dynamics as a function of mode mismatch probabilities. A computationally efficient sufficient condition for a special case of SHS with two discrete states is derived and the approach involves solving a straightforward eigenvalue problem to derive the critical region on the mode probability. Additionally, we propose the use of a transformed switched system to describe the bias dynamics for a generalized SHS with arbitrary number of modes and the convergence of the bias is then mapped to the stability of the transformed switched system. A sufficient and necessary condition such that the bias dynamics is guaranteed to be statistically convergent. Using numerical simulations of a smart grid with network topology errors, we verify and validate the theoretical results and demonstrate the potency of using the analysis in critical infrastructures.

- In a practical CPS, there are situations where the mode information is inaccurate due to cyber-impariments such as communication link failures. This type of failures can be adequately captured by time correlated mode mismatch errors. We address this issue by quantifying the bias dynamics resulting from Markovian distributed mode mismatches for Markov jump linear system (MJLS) state estimation. We begin the analysis with an MJLS with two discrete states and derive sufficient and necessary conditions (based on the results from Schur stability of a matrix polytope) under which the bias dynamics are statistically convergent. Then for a generalized MJLS, we model the mean of dias dynamics as an auxiliary linear system. The system matrix

of this linear system is determined by a polytope of matrices with each vertex matrix related to the original MJLS system matrices. For MJLS with arbitrary numbers of modes, by mapping the matrix polytope to an interval matrix, and by leveraging results in Schur stability analysis for an interval matrix, derive sufficient conditions on mode mismatch probabilities under which the bias resulting from mode mismatches is statistically convergent.

- Another critical situation in any CPS is spatially correlated cyber-effects (e.g., cyber-attacks and communication link impairments, etc.). This cyber-effects could impact the accuracy of the system mode information and then further result spatially correlated mode mismatches in the system. For a generalized MJLS, we derive sufficient conditions under which the bias resulting from mode mismatches is statistically convergent. The condition is related to mode mismatch probabilities and it is algebraically solvable. The result provides new insights on the fidelity of discrete state information needed to sustain the quality of the Kalman filter estimate.

Based on the accomplishment of this dissertation, some future research directions are highlighted in the next section.

7.2 Future Work

In this research, we have thoroughly investigated the impact of mode mismatches on the performance of a mode-based Kalman filter for SHS estimate. Different models of the mode mismatch errors have been considered and those models can widely capture the cyber-effect occurred in a practical system. With the basis developed in this dissertation, several potential research directions are summarized as follows:

- State estimation and control problems are naturally related as the design of controller highly depends on the quality of state estimate. Optimal control for SHS has attracted intense research interests in the last decades. An example of the switched linear quadratic regulator (LQR) has been explored in [98–102]. Another cluster of

research work focuses on LQR problem for MJLS. For the case of complete observations of both continuous state and discrete mode, the related quadratic optimal control problem have been found in [20] (Chapter 4), and [22, 103, 104]. In the situation where only partial information of states is available, the separation theorem [105] is proposed but it still faces enormous difficulty to obtain optimal control. See [20] (Chapter 6) and [22, 106, 107] for reference. Therefore, a subsequent topic of this research is to investigate the stability and robustness of an optimal controller for SHS in the presence of mode mismatch errors.

- Kalman filter based state estimation have found wide acceptance in networked control systems (NCS). Specifically, NCS typically requires a distributed estimation strategies due to (i) states are distributed in the physical space; (ii) sensors are arbitrarily deployed and they jointly sense the complete system; (iii) network infrastructure is used to communicate measurements to the central estimation unit, etc. There are multiple prior works focus on state estimation and performance analysis of Kalman filter for NCS and interested readers can refer to [9, 41, 61, 108, 109], etc. However, a framework of system that involves distributed state with hybrid (discrete and continuous) behavior has not been well established. One future direction of this dissertation is to study a Kalman filter based state estimation strategy for a distributed SHS. Furthermore, the influence of inaccurate mode information on the performance of mode-based Kalman filter for distributed estimate need to be investigated.

- Most practical applications of SHS such as smart grid and air traffic control system involves communication links that connect the sensors to the estimator. In this dissertation, we consider the cyber-impairments that impact the information on discrete state information. However, the communication links that transmit continuous state measurements can be affected resulting in delayed or lost measurements due to packet drops. The stability of Kalman filter in the presence of intermittent observations has been well studied for non-hybrid system [38–40, 42]. In the framework of SHS, intermittent observations (continuous state) and mode mismatches (discrete state) can

both exist at the same time and the quality of a mode-based Kalman filter estimate will be significantly impaired. With this complex but realistic cyber-effect modeling in the system, the stability of an estimator (e.g., mode-based Kalman filter) is certainly worth futher investigation.

Bibliography

[1] J. Hu, W.-C. Wu, and S. Sastry, *Modeling Subtilin Production in Bacillus subtilis Using Stochastic Hybrid Systems*, pp. 417–431. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

[2] A. Singh and J. P. Hespanha, "Stochastic hybrid systems for studying biochemical processes," *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 368, no. 1930, pp. 4995–5011, 2010.

[3] M. K. Ghosh, A. Arapostathis, and S. I. Marcus, "Optimal control of switching diffusions with application to flexible manufacturing systems," *SIAM Journal on Control and Optimization*, vol. 31, no. 5, pp. 1183–1204, 1993.

[4] J. P. Hespanha, *Stochastic Hybrid Systems: Application to Communication Networks*, pp. 387–401. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

[5] W. Glover and J. Lygeros, *A Stochastic Hybrid Model for Air Traffic Control Simulation*, pp. 372–386. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.

[6] M. Prandini and J. Hu, "Application of reachability analysis for stochastic hybrid systems to aircraft conflict prediction," in *47th IEEE Conference on Decision and Control (CDC)*, pp. 4036–4041, Dec 2008.

[7] M. Stelec, K. Macek, and A. Abate, "Modeling and simulation of a microgrid as a stochastic hybrid system," in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, pp. 1–9, Oct 2012.

[8] M. Rasheduzzaman, T. Paul, and J. W. Kimball, "Markov jump linear system analysis of microgrid stability," in *2014 American Control Conference*, pp. 5062–5066, June 2014.

[9] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry, "Foundations of control and estimation over lossy networks," *Proceedings of the IEEE*, vol. 95, pp. 163–187, Jan 2007.

[10] D. Liberzon, *Switching in Systems and Control*. Systems & Control: Foundations & Applications, Birkhäuser Boston, 2003.

[11] A. R. Teel, A. Subbaraman, and A. Sferlazza, "Stability analysis for stochastic hybrid systems: A survey," *Automatica*, vol. 50, no. 10, pp. 2435 – 2456, 2014.

[12] R. Shorten, F. Wirth, O. Mason, K. Wulff, and C. King, "Stability criteria for switched and hybrid systems," *SIAM Review*, vol. 49, no. 4, pp. 545–592, 2007.

[13] L. Zhang and E.-K. Boukas, "Stability and stabilization of markovian jump linear systems with partly unknown transition probabilities," *Automatica*, vol. 45, no. 2, pp. 463 – 468, 2009.

[14] H. Lin and P. J. Antsaklis, "Stability and stabilizability of switched linear systems: A survey of recent results," *IEEE Transactions on Automatic Control*, vol. 54, pp. 308–322, Feb 2009.

[15] A. Abate, M. Prandini, J. Lygeros, and S. Sastry, "Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems," *Automatica*, vol. 44, no. 11, pp. 2724 – 2734, 2008.

[16] W. Zhang, P. Prabhakar, and B. Natarajan, "Abstraction based reachability analysis for finite branching stochastic hybrid systems," in *2017 ACM/IEEE 8th International Conference on Cyber-Physical Systems (ICCPS)*, pp. 121–130, April 2017.

[17] C. E. Seah and I. Hwang, "Stochastic linear hybrid systems: Modeling, estimation, and application in air traffic control," *IEEE Transactions on Control Systems Technology*, vol. 17, pp. 563–575, May 2009.

[18] A. Doucet, N. J. Gordon, and V. Krishnamurthy, "Particle filters for state estimation of jump markov linear systems," *IEEE Transactions on Signal Processing*, vol. 49, pp. 613–624, Mar 2001.

[19] W. Zhang and B. Natarajan, "State estimation in stochastic hybrid systems with quadratic guard conditions," in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 752–757, Sept 2016.

[20] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-time Markov jump linear systems*. Springer Science and Business Media, 2006.

[21] Y. Huang, M. Esmalifalak, Y. Cheng, H. Li, K. A. Campbell, and Z. Han, "Adaptive quickest estimation algorithm for smart grid network topology error," *IEEE Systems Journal*, vol. 8, pp. 430–440, June 2014.

[22] 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*, pp. 1989–1993 vol.3, Dec 1988.

[23] J. Tugnait, "Adaptive estimation and identification for discrete systems with markov jump parameters," *IEEE Transactions on Automatic Control*, vol. 27, pp. 1054–1065, Oct 1982.

[24] W. Zhang and B. Natarajan, "Quantifying the bias dynamics in a mode-based kalman filter for stochastic hybrid systems," in *2018 Annual American Control Conference (ACC)*, pp. 5849–5856, June 2018.

[25] W. Zhang and B. Natarajan, "On the convergence of bias of a mode-based kalman filter for stochastic hybrid systems." EURASIP Journal on Advances in Signal Processing (In Press), 2018.

[26] W. Zhang and B. Natarajan, "Impact of time correlated mode mismatch on markov jump linear system state estimation," *IEEE Control Systems Letters*, vol. 2, pp. 489–494, July 2018.

[27] W. Zhang and B. Natarajan, "On the performance of kalman filter for markov jump linear systems with mode mismatch." Manuscript submitted to IEEE Transactions on Automatic Control, 2018.

[28] W. Zhang and B. Natarajan, "Bias analysis in kalman filter with correlated mode mismatch errors," *Signal Processing*, vol. 154, pp. 232 – 237, 2019.

[29] W. Liu, C. E. Seah, and I. Hwang, "Estimation algorithm for stochastic linear hybrid systems with quadratic guard conditions," in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, pp. 3946–3951, Dec 2009.

[30] M. Davis, *Stochastic modelling and control*. Springer Science and Business Media, 2013.

[31] I. Matei, N. C. Martins, and J. S. Baras, "Optimal state estimation for discrete-time markovian jump linear systems, in the presence of delayed mode observations," in *2008 American Control Conference*, pp. 3560–3565, June 2008.

[32] I. Matei and J. S. Baras, "Optimal state estimation for discrete-time markovian jump linear systems, in the presence of delayed output observations," *IEEE Transactions on Automatic Control*, vol. 56, pp. 2235–2240, Sept 2011.

[33] H. A. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with markovian switching coefficients," *IEEE Transactions on Automatic Control*, vol. 33, pp. 780–783, Aug 1988.

[34] X. R. Li and Y. Bar-Shalom, "Design of an interacting multiple model algorithm for air traffic control tracking," *IEEE Transactions on Control Systems Technology*, vol. 1, pp. 186–194, Sept 1993.

[35] M. Coombes, C. Liu, and W.-H. Chen, "Situation awareness for uav operating in terminal areas using bearing-only observations and circuit flight rules," in *2016 American Control Conference (ACC)*, pp. 479–485, July 2016.

[36] Y. Bar-Shalom and X.-R. Li, "Estimation and tracking- principles, techniques, and software," *Norwood, MA: Artech House, Inc, 1993.*, 1993.

[37] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. New York, NY, USA: Wiley-Interscience, 2006.

[38] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1453–1464, Sept 2004.

[39] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," in *Decision and Control, 2004. CDC. 43rd IEEE Conference on*, vol. 4, pp. 4180–4186 Vol.4, Dec 2004.

[40] E. Rohr, D. Marelli, and M. Fu, "Statistical properties of the error covariance in a kalman filter with random measurement losses," in *49th IEEE Conference on Decision and Control (CDC)*, pp. 5881–5886, Dec 2010.

[41] S. Deshmukh, B. Natarajan, and A. Pahwa, "State estimation over a lossy network in spatially distributed cyber-physical systems," *IEEE Transactions on Signal Processing*, vol. 62, pp. 3911–3923, Aug 2014.

[42] M. Moayedi, Y. C. Soh, and Y. K. Foo, "Optimal kalman filtering with random sensor delays, packet dropouts and missing measurements," in *2009 American Control Conference*, pp. 3405–3410, June 2009.

[43] B. Yan, H. Lev-Ari, and A. M. Stankovic, "Networked state estimation with delayed and irregularly-spaced time-stamped observations," *IEEE Transactions on Control of Network Systems*, vol. PP, no. 99, pp. 1–1, 2017.

[44] M. Nourian, A. S. Leong, S. Dey, and D. E. Quevedo, "An optimal transmission strategy for kalman filtering over packet dropping links with imperfect acknowledgements," *IEEE Transactions on Control of Network Systems*, vol. 1, pp. 259–271, Sept 2014.

[45] S. Dey, A. Chiuso, and L. Schenato, "Remote estimation with noisy measurements subject to packet loss and quantization noise," *IEEE Transactions on Control of Network Systems*, vol. 1, pp. 204–217, Sept 2014.

[46] P. D. Hanlon and P. S. Maybeck, "Characterization of kalman filter residuals in the presence of mismodeling," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 36, pp. 114–131, Jan 2000.

[47] I. Hwang, H. Balakrishnan, and C. Tomlin, "Performance analysis of hybrid estimation algorithms," in *Decision and Control, 2003. Proceedings. 42nd IEEE Conference on*, vol. 5, pp. 5353–5359 Vol.5, Dec 2003.

[48] C. E. Seah and I. Hwang, "Algorithm for performance analysis of the imm algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, pp. 1114–1124, April 2011.

[49] I. Hwang, C. E. Seah, and S. Lee, "A study on stability of the interacting multiple model algorithm," *IEEE Transactions on Automatic Control*, vol. 62, pp. 901–906, Feb 2017.

[50] S. Narasimhan and G. Biswas, "Model-based diagnosis of hybrid systems," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 37, pp. 348–361, May 2007.

[51] Y. Bar-Shalom, X.-R. Li, and T. Kirubarajan, *State Estimation in Discrete-Time Linear Dynamic Systems*, pp. 199–266. John Wiley and Sons, Inc., 2002.

[52] H. A. P. Blom and E. A. Bloem, "Particle filtering for stochastic hybrid systems," in *2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, vol. 3, pp. 3221–3226 Vol.3, Dec 2004.

[53] X. Koutsoukos, J. Kurien, and F. Zhao, *Estimation of Distributed Hybrid Systems Using Particle Filtering Methods*, pp. 298–313. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003.

[54] M. Prandini and J. Hu, "Application of reachability analysis for stochastic hybrid systems to aircraft conflict prediction," *IEEE Transactions on Automatic Control*, vol. 54, pp. 913–917, April 2009.

[55] C. E. Seah and I. Hwang, "State estimation for stochastic linear hybrid systems with continuous-state-dependent transitions: An imm approach," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, pp. 376–392, Jan 2009.

[56] A. A. Mohsenipour, *ON THE DISTRIBUTION OF QUADRATIC EXPRESSIONS IN VARIOUS TYPES OF RANDOM VECTORS*. Ph.d. thesis, The University of Western Ontario, London, Ontario, Canada, 2012.

[57] D. Khoshnevisan, "Quadratic forms." 2014.

[58] J. Marcum, *Table of Q Functions*. Memorandum (Rand Corporation), Rand Corporation, 1950.

[59] M. W. Hofbaur and B. C. Williams, *Mode Estimation of Probabilistic Hybrid Systems*, pp. 253–266. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002.

[60] C. B. Chang and M. Athans, "State estimation for discrete systems with switching parameters," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-14, pp. 418–425, May 1978.

[61] S. M. S. Alam, B. Natarajan, and A. Pahwa, "Agent based optimally weighted kalman consensus filter over a lossy network," in *2015 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, Dec 2015.

[62] I. Hwang, H. Balakrishnan, and C. Tomlin, "Observability criteria and estimator design for stochastic linear hybrid systems," in *European Control Conference (ECC), 2003*, pp. 3317–3322, Sept 2003.

[63] G. Welch and G. Bishop, "An introduction to the kalman filter," tech. rep., Chapel Hill, NC, USA, 1995.

[64] L. Xie and L. Xie, "Stability of a random riccati equation with markovian binary switching," in *2007 46th IEEE Conference on Decision and Control*, pp. 1565–1570, Dec 2007.

[65] O. Costa, "Discrete-time coupled riccati equations for systems with markov switching parameters," *Journal of Mathematical Analysis and Applications*, vol. 194, no. 1, pp. 197 – 216, 1995.

[66] M. R. Sanatkar, W. N. White, B. Natarajan, C. M. Scoglio, and K. A. Garrett, "Epidemic threshold of an sis model in dynamic switching networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 46, pp. 345–355, March 2016.

[67] V. D. Blondel and J. N. Tsitsiklis, "The boundedness of all products of a pair of matrices is undecidable," *Systems and Control Letters*, vol. 41, no. 2, pp. 135 – 140, 2000.

[68] R. Jungers, *The Joint Spectral Radius: Theory and Applications*. Lecture Notes in Control and Information Sciences, Springer Berlin Heidelberg, 2009.

[69] O. Mason and R. Shorten, "On common quadratic lyapunov functions for stable discretetime lti systems," *IMA Journal of Applied Mathematics*, vol. 69, no. 3, p. 271, 2004.

[70] D. Liberzon, J. P. Hespanha, and A. S. Morse, "Stability of switched systems: a lie-algebraic condition," *SYSTEMS CONTROL LETT*, vol. 37, no. 3, pp. 117–122, 1999.

[71] R. A. Decarlo, M. S. Branicky, S. Pettersson, and B. Lennartson, "Perspectives and results on the stability and stabilizability of hybrid systems," *Proceedings of the IEEE*, vol. 88, pp. 1069–1082, July 2000.

[72] F. Kittaneh, "Spectral radius inequalities for hilbert space operators," *Proceedings of the American Mathematical Society*, vol. 134, no. 2, pp. 385–390, 2006.

[73] S. Friedland, "Convex spectral functions," *Linear and Multilinear Algebra*, vol. 9, no. 4, pp. 299–316, 1981.

[74] R. Shorten, K. S. Narendra, and O. Mason, "A result on common quadratic lyapunov functions," *IEEE Transactions on Automatic Control*, vol. 48, pp. 110–113, Jan 2003.

[75] G. Michaletzky and L. Gerencser, "Bibo stability of linear switching systems," *IEEE Transactions on Automatic Control*, vol. 47, pp. 1895–1898, Nov 2002.

[76] R. Weron, B. Kozowska, and J. Nowicka-Zagrajek, "Modeling electricity loads in california: a continuous-time approach," *Physica A: Statistical Mechanics and its Applications*, vol. 299, no. 1, pp. 344 – 350, 2001. Application of Physics in Economic Modelling.

[77] H. Lin and P. J. Antsaklis, "Stability and persistent disturbance attenuation properties for a class of networked control systems: switched system approach," *International Journal of Control*, vol. 78, no. 18, pp. 1447–1458, 2005.

[78] C. Soh, "Schur stability of convex combination of matrices," *Linear Algebra and its Applications*, vol. 128, pp. 159 – 168, 1990.

[79] A. Czornik, "On the discrete timevarying jlqg problem," *Int. J. Appl. Math. Comput. Sci*, pp. 101–105, 2002.

[80] Y. Fang and K. A. Loparo, "Stochastic stability of jump linear systems," *IEEE Transactions on Automatic Control*, vol. 47, pp. 1204–1208, Jul 2002.

[81] T.Morozan, "Optimal stationary control for dynamic systems with markov perturbations," *Stochastic Analysis and Applications*, vol. 1, no. 3, pp. 299–325, 1983.

[82] L. Elsner and T. Szulc, "Convex combinations of matrices — nonsingularity and schur stability," in *Numerical Analysis and Its Applications*, (Berlin, Heidelberg), pp. 170–175, Springer Berlin Heidelberg, 1997.

[83] L. Elsner and T. Szulc, "Convex sets of schur stable and stable matrices," *Linear and Multilinear Algebra*, vol. 48, no. 1, pp. 1–19, 2000.

[84] S. BIALAS, "A necessary and sufficient condition for the stability of interval matrices," *International Journal of Control*, vol. 37, no. 4, pp. 717–722, 1983.

[85] B. R. BARMISH and C. V. HOLLOT, "Counter-example to a recent result on the stability of interval matrices by s. bialas," *International Journal of Control*, vol. 39, no. 5, pp. 1103–1104, 1984.

[86] L. X. XIN, "Necessary and sufficient conditions for stability of a class of interval matrices," *International Journal of Control*, vol. 45, no. 1, pp. 211–214, 1987.

[87] K. Wang, A. N. Michel, and D. Liu, "Necessary and sufficient conditions for the hurwitz and schur stability of interval matrices," *IEEE Transactions on Automatic Control*, vol. 39, pp. 1251–1255, June 1994.

[88] M. E. Sezer and D. D. Siljak, "On stability of interval matrices," *IEEE Transactions on Automatic Control*, vol. 39, pp. 368–371, Feb 1994.

[89] L. Grman, D. RosinovA, V. Vesely, and A. K. KovA, "Robust stability conditions for polytopic systems," *International Journal of Systems Science*, vol. 36, no. 15, pp. 961–973, 2005.

[90] M. de Oliveira, J. Geromel, and L. Hsu, "Lmi characterization of structural and robust stability: the discrete-time case," *Linear Algebra and its Applications*, vol. 296, no. 1, pp. 27 – 38, 1999.

[91] R. C. Oliveira and P. L. Peres, "Stability of polytopes of matrices via affine parameter-dependent lyapunov functions: Asymptotically exact lmi conditions," *Linear Algebra and its Applications*, vol. 405, pp. 209 – 228, 2005.

[92] D. C. Ramos and P. L. Peres, "A less conservative lmi condition for the robust stability

of discrete-time uncertain systems," *Systems & Control Letters*, vol. 43, no. 5, pp. 371 – 378, 2001.

[93] G. Chesi, A. Garulli, A. Tesi, and A. Vicino, "Robust stability of polytopic systems via polynomially parameter-dependent lyapunov functions," in *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)*, vol. 5, pp. 4670–4675 Vol.5, Dec 2003.

[94] P. Bliman, "A convex approach to robust stability for linear systems with uncertain scalar parameters," *SIAM Journal on Control and Optimization*, vol. 42, no. 6, pp. 2016–2042, 2004.

[95] D. Henrion, D. Arzelier, D. Peaucelle, and J. . Lasserre, "On parameter-dependent lyapunov functions for robust stability of linear systems," in *2004 43rd IEEE Conference on Decision and Control (CDC) (IEEE Cat. No.04CH37601)*, vol. 1, pp. 887–892 Vol.1, Dec 2004.

[96] F. Li, P. Shi, C. C. Lim, and L. Wu, "Fault detection filtering for nonhomogeneous markovian jump systems via a fuzzy approach," *IEEE Transactions on Fuzzy Systems*, vol. 26, pp. 131–141, Feb 2018.

[97] M.-Q. Chen and X. Li, "An estimation of the spectral radius of a product of block matrices," *Linear Algebra and its Applications*, vol. 379, no. Supplement C, pp. 267 – 275, 2004. Special Issue on the Tenth ILAS Conference (Auburn, 2002).

[98] W. Zhang, A. Abate, and J. Hu, "Efficient suboptimal solutions of switched lqr problems," in *2009 American Control Conference*, pp. 1084–1091, June 2009.

[99] W. Zhang, J. Hu, and A. Abate, "On the value functions of the discrete-time switched lqr problem," *IEEE Transactions on Automatic Control*, vol. 54, pp. 2669–2674, Nov 2009.

[100] W. Zhang, J. Hu, and A. Abate, "Infinite-horizon switched lqr problems in discrete

time: A suboptimal algorithm with performance analysis," *IEEE Transactions on Automatic Control*, vol. 57, pp. 1815–1821, July 2012.

[101] J. Veneman, "Switched LQR control: Design of a general framework," Master's thesis, Delft Centre for Systems and Control, 2012.

[102] W. Zhang, J. Hu, and J. Lian, "Quadratic optimal control of switched linear stochastic systems," *Systems and Control Letters*, vol. 59, no. 11, pp. 736 – 744, 2010.

[103] H. Abou-Kandil, G. Freiling, and G. Jank, "On the solution of discrete-time markovian jump linear quadratic control problems," *Automatica*, vol. 31, no. 5, pp. 765 – 768, 1995.

[104] A. Czornik and A. Swierniak, "On the discrete jlq and jlqg problems," *Nonlinear Analysis: Theory, Methods and Applications*, vol. 47, no. 1, pp. 423 – 434, 2001. Proceedings of the Third World Congress of Nonlinear Analysts.

[105] W. M. Wonham, "On the separation theorem of stochastic control," *SIAM Journal on Control*, vol. 6, no. 2, pp. 312–326, 1968.

[106] B.-G. Park and W. H. Kwon, "Robust one-step receding horizon control of discrete-time markovian jump uncertain systems," *Automatica*, vol. 38, no. 7, pp. 1229 – 1235, 2002.

[107] O. L. V. Costa and E. F. Tuesta, "Finite horizon quadratic optimal control and a separation principle for markovian jump linear systems," *IEEE Transactions on Automatic Control*, vol. 48, pp. 1836–1842, Oct 2003.

[108] J. P. Hespanha, P. Naghshtabrizi, and Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, pp. 138–162, Jan 2007.

[109] J. Hu, Z. Wang, D. Chen, and F. E. Alsaadi, "Estimation, filtering and fusion for networked systems with network-induced phenomena: New progress and prospects," *Information Fusion*, vol. 31, pp. 65 – 75, 2016.