MARKOV CHAIN STRUCTURES WITH APPLICATIONS TO SYSTEMIC RISK

BY

YU-SIN CHANG

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Applied Mathematics
in the Graduate College of the
Illinois Institute of Technology

Approved _____
Advisor

Approved _____
Co-Advisor

Chicago, Illinois
December 2017

ACKNOWLEDGMENT

It has been an honor and a pleasure to be a student of Professor Tomasz R. Bielecki and Professor Igor Cialenco. I have been privileged to work with them and have their supports throughout my doctoral studies. They were generous in sharing their extensive knowledge, and always inspired deep and helpful discussions. Without their thoughtful guidance, patience, and generosity this thesis would have never happened. I am sincerely grateful for everything that they have done for me.

I also want to thank many professors through my studies at Illinois Institute of Technology who have taught me and promoted my intellectual curiosity. I would also like to thank Professor Shuwang Li for his encouragement and discussion about his philosophy.

My doctoral studies would have not been possible without the generous financial support of the Applied Mathematics Department and my family. I am beholden to everyone who has helped me financially throughout my education.

I am indebted to Professor Matthew Dixon and Professor Ruoting Gong for serving in my Ph.D. thesis committee. I appreciate their time and efforts towards this commitment. I also want to thank Ms. Gladys Collins and Ms. Myrna Walker for all the administrative help that they have provided me during the last four years.

My family's unconditional support and love has been instrumental in this achievement. They have always championed my endeavor to pursue happiness and a life of the mind. I thank my parents for giving me love, encouragement, and support throughout my studies. I also thank my sister Wen-Chi and my brother Chia-Fu for their friendships, love, and always being there for me. I could not ask for a better family. Without my family, none of this would have been possible.

I dedicate this thesis to my parents.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

Symbol	Description
$(\Omega, \mathcal{F}, \mathbb{P})$	Complete underlying probability space
X	Multivariate Markov chain
X^i	The ith component of a multivariate Markov chain X
E_i	Finite state space of X^i
E	Finite state space of X; Cartesian product of finitely many E_i's
$\mathbb{F}^Y$	Natural filtration of any process Y defined on $(\Omega, \mathcal{F}, \mathbb{P})$
$\mathbb{E}^{\mathbb{P}}$	Expectation with respect to probability measure $\mathbb{P}$
$\mathcal{D}$	Dependence structure
$\mathcal{I}$	Independence structure
$X^{\mathcal{D}}$	Multivariate Markov chain endowed with dependence structure $\mathcal{D}$ between its components
$X^{\mathcal{D},i}$	The ith component of $X^{\mathcal{D}}$
$\sigma(\mathcal{G})$	The smallest σ-algebra generated by $\mathcal{G}$
$\mathcal{T}$	Time interval $[0, T]$
$(E, \mathcal{E})$	Measurable space
$E^{\mathcal{T}}$	Collection of all functions from $\mathcal{T}$ into E
π_t	Set of all finite partitions of a closed interval $[0, t]$
Π_t	Collection of all finite partitions of interval $[0, t]$
m	Positive integer
$\mathcal{M}$	Set $\{1, 2, \ldots, m\}$
μ^X	Initial distribution of process X

$\mathbf{P}_{t,s}$	Transition matrix function of X
Λ_u	Infinitesimal generator function of X
$\mathbf{P}^i_{t,s}$	Transition matrix function of X^i
Λ^i_u	Infinitesimal generator function of process X^i
$\mathbf{I}$	Identity matrix
Θ^i_t	Operator acting on a function $g : E \to \mathbb{R}$
Φ^i	Extension operator acting on a function $f : E_i \to \mathbb{R}$
Υ	Projection operator acting on a measure
$\mathcal{K}$	Finite state space $\{0, 1, \ldots K\}$
Ψ^a_n	Collection of vectors in $\mathcal{K}^{n+1}$ such that the last element of the vector is a
$\mathbf{0}$	Zero matrix
$\overset{\mathcal{L}}{=}$	Equality in law
ρ_t	Correlation between random variables at time t

LIST OF SYMBOLS FOR CHAPTER 4

Symbol	Description
I	Identity matrix
$\mathcal{X}$	Collection of Markov chains
$\Lambda_u^{\mathcal{D}}$	Infinitesimal generator of the multivariate Markov chain $X^{\mathcal{D}}$
ν^z	Risk measure associated with vector z
ρ^z	Dependence measure associate with vector z

ABSTRACT

This thesis studies three topics: *Markovian consistency theory*, *Markov structures theory*, and *modeling systemic risk*.

The first part investigates the necessary and sufficient conditions for a component of a multivariate Markov chain to be Markov in its own filtration and coincide with the prescribed marginal law. We first derive a necessary and sufficient condition, called condition (Ci), for a component to be Markovian in its own filtration. We provide a number of sufficient conditions, that are formulated in terms of the transition characteristic of the chain, for condition (Ci) with respect to strong Markovian consistency and weak Markovian consistency. We prove that one sufficient condition, condition (P), for condition (Ci) is equivalent to an existing result, i.e. condition (M). Additionally, we give conditions under which strong Markovian consistency implies condition (M). Next, we provide necessary and sufficient conditions for weak-only Markovian consistency. We finish this chapter by presenting several examples that satisfy the sufficient conditions derived in this work.

The second part studies the constructions of Markov structures. We propose and analyze two approaches to construct Markov structures: the top-down approach and the bottom-up approach. We give a sufficient condition for the one-dimensional distributions obtained by these two approaches to coincide forward in time. In addition, we show that the Markov structures can be constructed forward in time and satisfy the semigroup property if certain sufficient conditions are satisfied. We conclude this chapter by discussing and comparing our results regarding Markovian consistency and Markov structures to existing findings.

The third part works on systemic risk. We propose a novel approach to study systemic risk in financial networks, based on the analysis of the dependence structures between the financial institutions of the network. We construct a dynamic systemic

dependence measure and a dynamic systemic risk measure, and give financial interpretations of these two measures. We propose to use these two measures as a pair to monitor systemic risk. We present several numerical examples to explain how the dependence structures contribute to systemic risk. In particular, certain dependence structure may lead to systemic benefit.

CHAPTER 1
INTRODUCTION

The interconnection or dependence structure between objects within a stochastic dynamical system plays a critical role in modeling various phenomena from a wide range of applied fields. Various applications are related to, among others, the linked financial institutions within an economy, the biological cells of living organisms, sociology, criminology and terrorism, social networks, ruin problems in insurance, credit risk, and seismology. For finite dimensional multivariate random variables the dependence structure between their components is fully characterized in terms of the so-called copulae and the Sklar's Theorem [Skl59]. On the other hand, the analysis or the characterization of the dependence structure between components of multivariate stochastic processes is a relatively new area. In part, this thesis contributes to this area. The obtained theoretical results are applied to the analysis of systemic risk in a financial network.

In this thesis we focus on a class of time-inhomogeneous multivariate Markov chains, with finite state spaces and in a continuous time setup. The main objective of this work is to derive algebraic conditions in terms of the transition characteristics of a multivariate Markov chain, under which the corresponding coordinate processes are Markovian in their own filtrations and coincide in law with the prescribed marginal laws. The detailed research problem is stated in Section 2.3.

The concept of *Markovian consistency* and *Markov structures* for multivariate Markov processes was originated in [BVV08]. In [BJVV08], [BJN10], and [BJN12], the authors enrich the theoretical structures of Markovian consistency and Markov structures, and have successfully applied Markov structures to financial markets, in

particular to the basket-type products in credit risk. Our research builds upon these works as well as upon [BJN13]. We continue the analysis of the theoretical framework of [BJN13]. Furthermore, we apply the framework of Markov structures to study systemic risk. In particular, we construct a dynamic *systemic dependence measure* that takes into account the dependence between the components of a financial system expressed in terms of Markov structures.

1.1 Markovian Consistency and Markov Structures

In this section, we briefly recall some important results from the existing literature that are relevant to our study.

1.1.1 Markovian Consistency. As stated, we deal with the problem regarding the sufficient and necessary conditions under which a coordinate process of a multivariate Markov chain is Markovian in its own filtration with a predetermined law. We are also concerned with the problem of constructing a multivariate Markov chain such that each of its components is Markovian in its own filtration with a predetermined law; such a multivariate Markov chain is called a Markov structure.

To study such problems we first need to investigate conditions under which a given component of a multivariate Markov chain is also a Markov chain. The Markov property of the coordinate processes is studied within the so-called strong and weak Markovian consistency theory. The probabilistic necessary and sufficient conditions for Markovian consistency of multivariate Markov chains have been given in [BJN13, Theorem 1.6] in terms of the semimartingale characteristics. We are mostly interested in the weak Markovian consistency. However, the sufficient and necessary conditions for the weak Markovian consistency given in [BJN13, Theorem 1.6] are extremely difficult to verify. One of the goals of this thesis is to establish easily verifiable 'sufficient conditions' for weak Markovian consistency of a multivariate Markov chain

in terms of its transition matrix function.

It should be noted that our research problem is linked to the so-called *Markov function problem* or *lumpability problem* (cf. [RS93], [GL05], and the reference therein). A Markov function is a measurable function such that when applied to a Markov process the output is a Markov process itself. One important mathematical question is to find sufficient conditions under which a measurable function is a Markov function (cf. [BR58], [Kel82], [RP81], [KO88], and the reference therein). The Markov function problem can be traced back to Burke and Rosenblatt [BR58] and Dynkin [Dyn65a]. Dynkin considered time-homogeneous Markov processes, and the problem is called the transformation of the state space. Dynkin gave sufficient conditions on the infinitesimal operators (and the semigroups) of the original and transformed process Markov process. These conditions are known in the current literature as *Dynkin's criterion.*

Clearly, a coordinate process of the multivariate Markov chain can be obtained by the coordinate projection of the chain, and thus it is a special case of a Markov function. Thus, the study of Markov consistency is related, but not entirely nested, in the study of Markov functions. In particular, the study of Markov functions has been limited to the case of time-homogeneous processes, whereas the study of Markov consistency is not. We also note that the classical study of Markov functions is not concerned with the question of constructing Markov structures.

Since we work on a multivariate Markov chain and its coordinate processes, the state of the coordinate process can be obtained by 'lumping over' the corresponding state of the chain. In this context, our research problem is related to the lumpability problem. The coordinate projection corresponds to a lumping map, and the coordinate process corresponds to a lumped process. A Markov chain is said to be *weakly lumpable* if there exists an initial distribution of the original chain such

that the lumped process is a homogeneous Markov chain. In the weak lumpability problem, the Markov property is defined with respect to the natural filtration of the lumped process. However, as discussed below we consider two notions of Markovian consistency. One notion is related to the natural filtration of the multivariate Markov chain, and the other one is related to the natural filtration of the coordinate process. In general, these two filtrations are not the same. Thus, our research problem is more general than the weak lumpability problem.

The two notions of Markovian consistency (cf. Definitions 2.4.1 and 2.4.2 on page 26 for details) play a fundamental role in our study. We say that a multivariate Markov chain is *strongly Markovian consistent* if each coordinate process is a Markov chain in the filtration of the multivariate Markov chain. A multivariate Markov chain is said to be *weakly Markovian consistent* if each coordinate process is a Markov chain in its own filtration. Indeed, one can show that the latter is a weaker property than the former. As we will see later, an important and significant feature distinguishes these two classes of Markov chains. This unique feature is linked to a desired property in the context of credit risk management that essentially can not be captured by strong Markovian consistency.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete underlying probability space. We consider a continuous time multivariate Markov chain $X = (X^1, \ldots, X^m)$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in finite state space E. For any process Y defined on $(\Omega, \mathcal{F}, \mathbb{P})$, we denote by $\mathbb{F}^Y = (\mathcal{F}_t^Y, t \geq 0)$ the natural filtration of Y, $\mathcal{F}_t^Y = \sigma\{Y_u,\ 0 \leq u \leq t\}$.

Motivated by results in [BJN13, Theorem 1.6 and Theorem 1.11], we naturally propose that one sufficient condition for the weak Markovian consistency is

$$\mathbb{E}_{\mathbb{P}}\left(\mathbb{1}_{\{X_t=x\}} \mid \mathcal{F}_t^{X^i}\right) = \mathbb{E}_{\mathbb{P}}\left(\mathbb{1}_{\{X_t=x\}} \mid X_t^i\right), \quad x \in E,\ i = 1, 2, \ldots, m. \tag{1.1}$$

This condition implies that the future state of X and the past state of X^i are condi-

tionally independent given the present state of X^i.

In Lemma 2.5.3 on page 47, without using [BJN13, Theorem 1.6 and Theorem 1.11], we show that Equation (1.1) is equivalent to condition (Ci-II) which is a sufficient condition for a component to be Markov in its own filtration.

Similar conditions to Equation (1.1) have been studied in [Kel82], [RP81], and [KO88] as well. Kelly [Kel82] shows that Equation (1.1) has an equivalent form with $t \in \mathbb{Z}$ and proves that the equivalent condition is sufficient for the derived process to be Markov. Rogers and Pitman [RP81] formulate conditions (cf. Section 3.4) for Equality (1.1) to hold true for a time-homogeneous Markov process. They further show that under these conditions, a function of a time-homogeneous Markov process is again a time-homogeneous Markov process. Kurtz and Ocone [KO88] derive Equality (1.1) as a result of a filtered Martingale problem, which recovers Rogers and Pitman's results. Whereas, Glover [Glo91] argues that there are more general conditions for a function of a Markov process to be Markovian by providing a counterexample. He further works on the intertwining relation between the semigroups of two Markov processes by a positive kernel, without assuming that this positive kernel is a Markov kernel.

It should be noted that the aforementioned results are all derived under the assumption that the given Markov process is time-homogeneous. We want to emphasize that we are dealing with time-inhomogeneous Markov chains. Of course, a time-inhomogeneous Markov chain X can be homogenized. However, the resulting time-homogeneous process (t, X) is a Markov process with the (non-finite) state space $\mathbb{R} \times E$, and not a Markov chain anymore. Therefore, the transition semigroup of (t, X) becomes an operator on an infinite dimensional space, rather than a matrix. Not only do certain properties or advantages from a transition matrix disappear, but also the computations of the transition semigroup get more complicated and in most of the

cases intractable.

1.1.2 Markov Structures. A Markov structure is a multivariate Markov chain, say $X = (X^1, \ldots, X^m)$, whose components are Markov chains with the prescribed laws under $\mathbb{P}$. Typically, the components of a Markov structure are not independent, but exhibit some stochastic dependence, which is encoded in the initial distribution of the process and in its infinitesimal generator[1]. Let $\mathcal{D}$ denote a generic *dependence structure* of X. The symbol $\mathcal{I}$ will be reserved for the independence structure (all the components $X^1, \ldots, X^m$ are independent under $\mathbb{P}$).

In analogy to Markovian consistency, there are strong Markov structures and weak Markov structures. See Definition 3.2.1 on page 87 and Definition 3.2.2 on page 88. Markov structures theory facilitates the study of the stochastic dependence between the components of a multivariate Markov chain. The main task of this theory is to construct Markov structures for a given collection of marginal laws. As stated above, the theory of Markovian consistency plays an fundamental role in the theory of Markov structures; consequently, it will play an important role in the constructions of Markov structures that we give below. It will be seen though that developing an efficient algorithm for the construction of Markov structures is, in general, quite a complex task.

It should be noted that Markov structures theory is different from Sklar's seminal work [Skl59] on copula functions. The classical concept of copulae used in probability is to construct multivariate finite dimensional random variables with given marginal distributions. A copula captures the dependence structure of multivariate distributions without studying marginal distributions. Bouyé et al. [BDN+00] provide

[1]For simplicity, throughout the thesis it will be assumed that a process start from a deterministic initial position. Thus, the dependence between the components of X is entirely given in terms of its infinitesimal generator.

a detailed survey on applications of copulae in finance. Markov structures theory analyzes the dependence structure between the coordinate processes of a multivariate Markov chain. We refer to [BJN13] for a comprehensive discussion and comparison of Markov structures theory and the classical copulae.

In Section 3.3, we introduce two methods to construct Markov structures: the *top-down approach* and the *bottom-up approach*. The major difference between these two approaches is how to compute the one-dimensional distribution of X. In the top-down approach we obtain the one-dimensional distribution of X in terms of the transition semigroup of X. In the bottom-up approach we construct the one-dimensional distribution of X by the *appropriate* correlations between the components and the one-dimensional distributions of the components at each time. In addition, we discuss the pros and cons of these two approaches, and give conditions for the correlations under which the one-dimensional distributions of X obtained by the top-down approach and the bottom-up approach coincide. Additionally, the constructed Markov structures preserve semigroup property.

1.2 Systemic Risk

The most recent global financial crisis has highlighted interconnectedness in the financial system as a crucial source of systemic risk. In the 1990s, the concept of modeling systemic risk focused on the *too-big-to-fail* issue, whereas the recent financial crisis addresses the *too-connected-to-fail* problem. Since financial institutions are linked by financial activities, the default of one systemically important financial institution may jeopardize the stability of other financial counterparties through financial activities. This may lead to successive rounds of failures, where the default of one financial institution escalates the financial distress of other engaged institutions. Under adverse circumstances the distress may propagate through the whole financial system and cause the collapse of the financial system. We believe this financial

phenomenon is highly related to the dependence structure within a financial system, namely, the dependence between financial institutions.

On the other hand, a default of a "bad player" in a financial system may lead to improved health of the surviving part of the system. Such a phenomenon can be justifiably termed as "systemic benefit".

There is a vast literature devoted to modeling systemic risk. We refer to the *Handbook of Systemic Risk* [FL13] for recent developments on systemic risk and the references therein. Among others, the network models are the mainstream and were originated in the seminal work of Eisenberg and Noe [EN01], that has great influence in the literature on financial contagion and systemic risk. Eisenberg and Noe have developed a framework that combines two issues in traditional finance literature: the uncertainty over the future values of assets and the interdependence of financial claim values. A limitation of the Eisenberg-Noe model is that the valuation of claims is finalized *ex-post*, namely, at the maturity of the financial contracts. Many works have extended the Eisenberg-Noe framework and tried to move toward ex-ante valuations, see for instance [GY15] and [RV13]. For a recent survey of network models, see [Sum13]. Although network models address issues of systemic stability and contagion risk, the default propagation is typically through balance-sheet mechanics.

We briefly mention other approaches that have been developed to study systemic risk. One such approach is to use the theory of large deviations for diffusions interacting through their mean field. Garnier et al. [GPY13] consider a system of diffusion processes that represent interacting agents. In their model, the intrinsic stabilization mechanism keeps the agents near the normal state, and the external destabilizing forces push the agents to a failed state. Garnier et al. study the strength of the intrinsic stabilization, the strength of the external random perturbations, and the degree of cooperation or interaction between them. Another mean field approach

is proposed by Carmona et al. [CFS15]. They model the log-monetary reserves of a bank with finitely many diffusion processes coupled through their drifts. A game feature is then introduced to study the financial stability between banks.

Unlike the aforementioned works, another approach is to construct systemic risk measures and study the properties of the measures. Bisias et al. [BFLV12] surveys 31 quantitative measures of systemic risk in the economics and finance literature. Chen et al. [CIM13] adopts the axiomatic approach to single-firm risk measures introduced in Artzner et al. [ADEH99]. This is the first work that proposes an axiomatic framework for defining the systemic risk measure as a summary statistic that quantifies the level of *risk* associated with an economy from a regulatory point of view. The approach in [CIM13] is general in the sense that it does not focus on any specific systemic risk measure. Biagini et al. [BFF15] provides a general methodological framework that goes beyond designing systemic risk measures through multi-dimensional acceptance sets and aggregation functions. Kromer et al. [KOZ16] extends [CIM13] to a general probability space and relax the assumption on positive homogeneity of systemic risk measures. Feinstein et al. [FRW15] proposes a monetary systemic risk that is set-valued. It measures the set of allocations of additional capital that lead to acceptable outcomes. This concept is similar to the minimal total amount injected into the institutions to secure the financial system in [BFF15]. Hoffmann et al. [HMBS16] axiomatically characterizes systemic risk measures in a conditional framework, where conditional aggregation functions and conditional base risk measures are considered.

1.2.1 Stochastic Dependence. We believe that the dependence structure between financial institutions is at the core of studying the systemic risk. A novel approach is proposed to model and measure systemic dependence.

Suppose that a multivariate Markov chain X taking values in E is endowed

with dependence structure $\mathcal{D}$ between its components ; to emphasize this, we will use the notation $X^{\mathcal{D}} = (X^{\mathcal{D},1} \ldots, X^{\mathcal{D},m})$. Let f be a real valued integrable function. We are interested in the conditional probabilities of of the form

$$\mathbb{P}\left(f(X_T^{\mathcal{D}}) \in B \mid X_t^{\mathcal{D}}\right), \quad 0 \leq t \leq T, \tag{1.2}$$

where B is a Borel set.

One way to normalize (1.2) is to consider the difference

$$\mathbb{P}\left(f(X_T^{\mathcal{D}}) \in B \mid X_t^{\mathcal{D}} = x\right) - \mathbb{P}\left(f(X_T^{\mathcal{I}}) \in B \mid X_t^{\mathcal{I}} = x\right), \quad 0 \leq t \leq T,\ x \in E. \tag{1.3}$$

Clearly, there are various approaches to compute the probability (1.3). For instance, one may use a statistical measure in (1.3), and then compute and analyze (1.3) from historical data.

We propose to study (1.3) in terms of Markov structures theory. Under this framework, we can examine systemic dependence directly through dependence structures between financial institutions. A thorough discussion about the construction of *systemic dependence measure* can be found in Section 4.4.

1.2.2 Systemic Risk Measure. As mentioned earlier, there are different perspectives regarding what a systemic risk measure should be, or what general properties a systemic risk measure needs to preserve. In this thesis, we first define a function. See Definition 4.4.1 on page 120. We want to compute the conditional probability of a fixed number of financial institutions that will be in certain credit ratings at some future time. In particular, if these credit ratings are the default rating. Then we call this function as the systemic risk measure.

A similar perspective about quantifying systemic risk has appeared in [Gig11]. Giglio computes the joint default probability of finite large financial institutions by using the information in bond and credit default swap prices. This probability is called

as systemic default risk. Without making any assumptions on the joint distribution function, Giglio proposes an optimal aggregation method to construct the bounds on the probability of systemic default events.

We argue that, only studying the probability of systemic default events without considering the dependence between financial institutions is not enough. As we will see in Section 4.5, it may happen that the probabilities of the systemic event are the same for two financial systems, but these two financial systems are endowed with distinct dependence structures. Therefore, we propose to use a combined measure, the pair of systemic risk measure and systemic dependence measure, to model systemic risk. A detailed discussion can be found in Section 4.4.

1.3 Original Contributions of the Thesis

In summary, the contribution of the thesis on each major research topic (that corresponds to a chapter) is as follows.

Markovian Consistency; Chapter 2.

- We derive condition (Ci), that is a necessary and sufficient condition for a component of a multivariate Markov chain to be Markovian in its own filtration. We provide several sufficient conditions to condition (Ci) with respect to strong Markovian consistency and weak Markovian consistency.

- One of the sufficient conditions for condition (Ci) is condition (P). We prove that condition (P), formulated in terms of the transition matrix of X, is equivalent to condition (M), derived in [BJN13] and given in terms of the infinitesimal generator of X. Moreover, we provide a condition under which condition (M) is necessary for strong Markovian consistency.

- We give conditions for weak-only Markovian consistency.

- We extend the results by Rogers and Pitman [RP81] to the case of time-inhomogeneous multivariate Markov chain with coordinate projections; see Theorem 2.5.2 on page 56.

- We give a number of examples that satisfy the sufficient conditions derived in this work; see Section 2.6 on page 68.

Markov Structures; Chapter 3.

- We propose and analyze two approaches to construct Markov structures: the top-down approach and the bottom-up approach. We present an example to elucidate the bottom-up approach. In addition, we analyze the conditions under which the Markov structures obtained by the bottom-up approach preserve semigroup property.

- We compare our results on Markovian consistency and Markov structures to the results from the existing literature. In particular, we show our results are more general.

Study of Systemic Risk; Chapter 4.

- We propose a novel approach to study systemic risk in financial networks, based on the analysis of the dependence structures between the components of the network.

- We construct a dynamical systemic risk measure and a dynamical systemic dependence measure. We give financial interpretations of these proposed measures and link them to the concept of monitoring the systemic risk in terms of the pair of systemic risk measure and systemic dependence measure.

- We provide a comprehensive numerical study of the proposed methodology, in particular, by considering various algebraic structures of infinitesimal generators, as well as by varying some model parameters.

This thesis is organized as follows. In Chapter 2, we focus on Markovian consistency and study the sufficient conditions for the component of a multivariate Markov chain to be Markov in its own filtration. Chapter 3 introduces two approaches to construct Markov structures. We apply the framework of Markov structures to model systemic risk in Chapter 4. Two detailed computations for the examples in Chapter 2 are given in Appendix A and Appendix B.

CHAPTER 2
MARKOVIAN CONSISTENCY

2.1 Introduction

In many applications, the objects of interest are usually assumed to have the Markov property. The main advantage of a Markovian assumption is that it simplifies modeling. For instance, as a feature of the Markov property, whatever happens in the past does not affect what will happen in the future. It saves time to deal with the past information. Here, we confine our study to the case of time-inhomogeneous multivariate Markov chains. As mentioned earlier, the components of a multivariate Markov chain is not necessarily Markovian. We give a simple example. Let Z^1 be a non-Markovian process and take $Z^2 = -Z^1$. Clearly, a process $Z = Z^1 + Z^2$ is a constant process, and therefore, a Markov process. In this chapter, we aim at investigating the conditions under which the components of a multivariate Markov chain are Markovian.

Since we are interested in multivariate Markov chains, immediately we have at least two types of filtrations. One is the natural filtration of the multivariate Markov chain, and the others are the natural filtrations of its individual components. If we are in the filtration of the multivariate Markov chain, we know everything about the chain. Typically, one has limited information about the whole chain if one is in the filtration of the component. Therefore, we introduce the concepts of *strong Markovian consistency* and *weak Markovian consistency*. The main difference between these two concepts is based on the type of filtrations.

From here on, we sharpen our focus to our research problem about Markovian consistency: the sufficient conditions under which the components of a multivariate

Markov chain are Markovian in their own filtrations.

This chapter is organized as follows. In Section 2.2, we begin by proving two equivalent forms of the continuous time Markov property for general stochastic processes. Starting from Section 2.3 we restrict ourselves to continuous time multivariate Markov chains and explain our main research problems. In Section 2.4 we give definitions of Markovian consistency and collect relevant results to lay the groundwork to our study. In Section 2.5, we provide necessary and sufficient conditions for weak Markovian consistency, strong Markovian consistency, and weak-only Markovian consistency. In Section 2.6 we give examples for the theoretical results. We conclude this chapter and discuss future work in Section 2.7.

2.2 Two Equivalent Definitions for the Continuous Time Markov Property

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete underlying probability space. Throughout the thesis, for any process Y defined on $(\Omega, \mathcal{F}, \mathbb{P})$, we denote by $\mathbb{F}^Y = (\mathcal{F}_t^Y, t \geq 0)$ the natural filtration of Y, $\mathcal{F}_t^Y = \sigma\{Y_u,\ 0 \leq u \leq t\}$.

The notion of Markovianity can be given in several equivalent forms (cf. [Wen81, Chapter 8.2]). Usually, the definitions of Markovian consistency are given in terms of filtrations. For continuous time Markov chains with finite or countable states, the Markov property can be stated without using the conditional probability with respect to the σ-algebra generated by the process (see Proposition 2.2.1 on page 19). Our results on the sufficient conditions for Markovian consistency are built upon the latter form of the Markov property. To seamlessly connect these two expressions of the Markov property, in this section, we devote ourselves to establishing the equivalence of these two forms of the Markov property.

We begin by recalling some relevent definitions and known results (see [Dyn65b]).

Definition 2.2.1. *A nonempty family $\mathcal{G}$ of subsets of a nonempty set G is*

(i) a π-system, if for any $A, B \in \mathcal{G}$, we have $A \cap B \in \mathcal{G}$;

(ii) a λ-system, if

(1) $G \in \mathcal{G}$;

(2) for any $A, B \in \mathcal{G}$, and $B \subseteq A$, we have $A \setminus B \in \mathcal{G}$;

(3) for any increasing sequence of sets $A_1 \subseteq A_2 \subseteq \ldots A_n \subseteq \ldots \in \mathcal{G}$, it implies $\cup_{n=1}^{\infty} A_n \in \mathcal{G}$.

(iii) a σ-algebra, if

(1) for any $A \in \mathcal{G}$, we have $G \setminus A \in \mathcal{G}$;

(2) for every sequence of sets $A_1, \ldots, A_n \ldots \in \mathcal{G}$, it implies $\cup_{n=1}^{\infty} A_n \in \mathcal{G}$ and $\cap_{n=1}^{\infty} A_n \in \mathcal{G}$.

Recall that $\sigma(\mathcal{G})$ is the smallest σ-algebra generated by $\mathcal{G}$.

Theorem 2.2.1 (Dynkin's π-λ Theorem)**.** *If a λ-system $\mathfrak{D}$ contains a π-system $\mathcal{G}$, then $\mathfrak{D}$ contains the σ-algebra $\sigma(\mathcal{G})$, $\sigma(\mathcal{G}) \subseteq \mathfrak{D}$.*

Let $\mathcal{T} = [0, T]$, and assume that $(E, \mathcal{E})$ is a measurable space. We denote by $E^{\mathcal{T}}$ the collection of all functions from $[0, T]$ into E, namely,

$$E^{\mathcal{T}} = \{y \mid y : [0, T] \to E\}.$$

An element y of $E^{\mathcal{T}}$ at $t \in \mathcal{T}$ will be denoted by y_t. For every $t \in \mathcal{T}$ we define a coordinate function, $Y_t : E^{\mathcal{T}} \to E$,

$$Y_t(y) = y_t. \tag{2.1}$$

If t is fixed, $Y_t(\cdot)$ is a function on $E^{\mathcal{T}}$. If y is fixed, $Y_\cdot(y)$ is a function on $\mathcal{T}$.

Let $\mathcal{E}^{\mathcal{T}}$ be the σ-algebra generated by all the coordinate functions $Y_t,\ t \in \mathcal{T}$. If we have a probability measure on $\left(E^{\mathcal{T}}, \mathcal{E}^{\mathcal{T}}\right)$, $Y_t(y)$ will be the coordinate random variables.

Next, we give the definition of *cylinder sets*, or *finite-dimensional sets.*

Definition 2.2.2 (Cylinder sets)**.** *A subset $C^n \subseteq E^{\mathcal{T}}$ is called an n-dimensional cylinder set, if there exist an n-tuple of distinct points $\{t_1, t_2, \ldots, t_n\} \subset \mathcal{T}$, and a measurable set $\Gamma \in \mathcal{E}^n$ such that*

$$C^n = \{y \in E^{\mathcal{T}} \mid (y_{t_1}, y_{t_2}, \ldots, y_{t_n}) \in \Gamma\}.$$

If the set Γ in Definition 2.2.2 is a singleton, namely, $\Gamma = (y_{t_1}, y_{t_2}, \ldots, y_{t_n})$ with $y_{t_i} \in E,\ i = 1, \ldots, n$, then it is called a *thin cylinder.* We denote by C the family consisting of all n-dimensional cylinder sets $C^n,\ n \in \mathbb{N}$. It is known that C forms an algebra, but not a σ-algebra in the case of an infinite T. The smallest σ-algebra contains all finite-dimensional cylinder sets is called *cylinder σ-algebra*, $\sigma(C)$.

If we have a probability measure $\mathbb{P}$ on $\mathcal{E}^{\mathcal{T}}$, the distribution of a function $Y_t,\ t \in \mathcal{T}$, on $(t_1, \ldots, t_n) \subset \mathcal{T}$ for every $\Gamma \in \mathcal{E}^n$ is defined by

$$\begin{aligned} p_{t_1,\ldots,t_n}(\Gamma) &:= \mathbb{P}\left(y \in E^{\mathcal{T}} \mid (Y_{t_1}(y), Y_{t_2}(y), \ldots, Y_{t_n}(y)) \in \Gamma\right) \\ &= \mathbb{P}\left(y \in E^{\mathcal{T}} \mid (y_{t_1}, y_{t_2}, \ldots, y_{t_n}) \in \Gamma\right), \end{aligned} \tag{2.2}$$

where the second equality follows from Equation (2.1).

Given the probability measure $\mathbb{P}$, the collection of all measurable functions $Y = (Y_t, t \in \mathcal{T})$ on $\left(E^{\mathcal{T}}, \mathcal{E}^{\mathcal{T}}, \mathbb{P}\right)$ is a stochastic process. Accordingly, these probability measures $p_{t_1,\ldots,t_n}$'s defined on the space $\left(E^{\mathcal{T}}, \mathcal{E}^{\mathcal{T}}\right)$ are called the finite-dimensional distributions of Y. It follows from (2.2) that this system of finite-dimensional distributions satisfies two consistency properties:

(i) *Permutation invariance.* Let $(t_{i_1}, \ldots, t_{i_n})$ be a permutation of $(t_1, \ldots, t_n)$. Then, for any $\Gamma_i \in \mathcal{E},\ i = 1, \ldots, n$, we have

$$p_{t_1,\ldots,t_n}(\Gamma_{t_1} \times \Gamma_{t_2} \times \cdots \times \Gamma_{t_n}) = p_{t_{i_1},\ldots,t_{i_n}}(\Gamma_{t_{i_1}} \times \Gamma_{t_{i_2}} \times \cdots \times \Gamma_{t_{i_n}}).$$

(ii) *Projection invariance.* Let $n \geq 2$. Then,

$$p_{t_1,\ldots,t_{n-1},t_n}(\Gamma_{t_1} \times \Gamma_{t_2} \times \cdots \times \Gamma_{t_{n-1}} \times E) = p_{t_1,\ldots,t_{n-1}}(\Gamma_{t_1} \times \Gamma_{t_2} \times \cdots \times \Gamma_{t_{n-1}}).$$

We can associate the process Y with a system of finite-dimensional distributions, which inherits the properties of Y through consistency properties. *Kolmogorov's existence theorem* answers the converse question and provides one way to construct a stochastic process through its finite-dimensional distributions.

Theorem 2.2.2 (Kolmogorov's Existence Theorem)**.** *Let $\mathcal{T}$ be some interval. If a system of distributions $p_{t_1,\ldots,t_n}$ satisfies the consistency properties, then on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ there exists a stochastic process $Y = (Y_t, t \in \mathcal{T})$ whose finite-dimensional distributions is given by $p_{t_1,\ldots,t_n}$.*

Next lemma says that the information contained in the natural filtration of the stochastic process is the same as the information included in the cylinder σ-algebra generated by the finite-dimensional distributions of the process. A detailed proof can be found in [Bil12, Section 36].

Lemma 2.2.1. *The cylinder σ-algebra generated by the finite-dimensional distributions of process Y coincides with the σ-algebra generated by Y, $\sigma(C) = \mathcal{F}_t^Y$.*

In the following proposition, we establish the equivalent forms of the Markov property. One is given in terms of the σ-algebra generated by the process, and another one is defined in terms of its trajectories.

Proposition 2.2.1. *Consider a measurable space* $(E, \mathcal{E})$, $|E| < \infty$. *Let* $X = (X_t, t \geq 0)$ *be a stochastic process taking values in* E. *Then, the following two statements are equivalent:*

(i) For every $s, t \geq 0$ *and* $\Gamma \in \mathcal{E}$,

$$\mathbb{P}\left(X_{t+s} \in \Gamma \mid \mathcal{F}_t^X\right) = \mathbb{P}\left(X_{t+s} \in \Gamma \mid X_t\right). \tag{2.3}$$

(ii) For any $n \in \mathbb{N}$, $0 \leq t_1 \leq \cdots \leq t_n \leq t_{n+1}$, $x_i \in E$, $i = 0, 1, \ldots, n+1$, *the following holds true,*

$$\mathbb{P}\left(X_{t_{n+1}} = x_{n+1} \mid X_{t_n} = x_n,\ X_{t_{n-1}} = x_{n-1}, \ldots,\ X_0 = x_0\right) = \mathbb{P}\left(X_{t_{n+1}} = x_{n+1} \mid X_{t_n} = x_n\right), \tag{2.4}$$

whenever $\mathbb{P}\left(X_{t_n} = x_n,\ X_{t_{n-1}} = x_{n-1}, \ldots,\ X_0 = x_0\right) > 0$.

Proof. Without loss of generality, we set $t_n = t$ and $t_{n+1} = t + s$, $s, t \geq 0$.

We first show that (i) implies (ii). Assume that (i) holds. We will prove that (2.3) implies (2.4) by the tower property of conditional expectations. For any $x_{n+1} \in E$, we know the set $\{X_{t_{n+1}} = x_{n+1}\} \in \mathcal{E}$. Since $|\Gamma| < \infty$, we have only countably many events $\{X_{t_{n+1}} \in \Gamma\} = \cup_{x_{n+1} \in \Gamma}\{X_{t_{n+1}} = x_{n+1}\}$. Due to the countable additivity of probability, it suffices to consider (2.3) in the following form,

$$\mathbb{P}\left(X_{t_{n+1}} = x_{n+1} \mid \mathcal{F}_{t_n}^X\right) = \mathbb{P}\left(X_{t_{n+1}} = x_{n+1} \mid X_{t_n}\right), \quad x_{n+1} \in E. \tag{2.5}$$

We denote by $\pi = \{t_0, t_1, \ldots, t_{n-1}, t_n\}$ a set of all finite partitions of a closed interval $[t_0, t_n]$ such that $t_0 < t_1 < \cdots < t_n$. The notation Π_t is a collection of all partitions,

$$\Pi_t = \{\pi = \{t_0, t_1, \ldots, t_{n-1}, t_n\} \mid t_0 < t_1 < \cdots < t_n,\ n \in \mathbb{N}\}.$$

We denote by $\sigma\left(X_{t_n}, X_{t_{n-1}}, \ldots, X_0; \pi\right)$ the smallest σ-algebra generated by $\{X_{t_n} = x_n,\ X_{t_{n-1}} = x_{n-1}, \ldots,\ X_0 = x_0\}$, $x_i \in E$, $i = 0, 1, \ldots, n \in \mathbb{N}$, with respect to the partition $\pi \in \Pi_t$.

Since

$$\sigma\left(X_{t_n}\right) \subseteq \sigma\left(X_{t_n}, X_{t_{n-1}}, \ldots, X_0; \pi\right) \subseteq \mathcal{F}_{t_n}^X,$$

we have that, for any $\pi \in \Pi_t$,

$$\begin{aligned}
&\mathbb{P}\left(X_{t_{n+1}} = x_{n+1} \mid \sigma\left(X_{t_n}, X_{t_{n-1}}, \ldots, X_0; \pi\right)\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\{X_{t_{n+1}} = x_{n+1}\}} \mid \sigma\left(X_{t_n}, X_{t_{n-1}}, \ldots, X_0; \pi\right)\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t_{n+1}} = x_{n+1}\}} \mid \mathcal{F}_{t_n}^X\right) \mid \sigma\left(X_{t_n}, X_{t_{n-1}}, \ldots, X_0; \pi\right)\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t_{n+1}} = x_{n+1}\}} \mid X_{t_n}\right) \mid \sigma\left(X_{t_n}, X_{t_{n-1}}, \ldots, X_0; \pi\right)\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\{X_{t_{n+1}} = x_{n+1}\}} \mid \sigma\left(X_{t_n}\right)\right) \\
&= \mathbb{P}\left(X_{t_{n+1}} = x_{n+1} \mid \sigma\left(X_{t_n}\right)\right),
\end{aligned}$$

where the second equality follows from the tower property of conditional expectations, the third equality results from (2.5), and the fourth equality is again due to the tower property of conditional expectations. Therefore, we establish (2.4).

Next we prove that (ii) implies (i). By uniqueness of the conditional expectation, there exists $\mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mid \mathcal{F}_t^X\right) \in L^1\left(\Omega, \mathcal{F}_t^X, \mathbb{P}\right)$ such that for every $A \in \mathcal{F}_t^X$,

$$\mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mathbb{1}_A\right) = \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mid \mathcal{F}_t^X\right) \mathbb{1}_A\right).$$

In order to show (2.3), we need to establish

$$\mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mathbb{1}_A\right) = \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mid \sigma\left(X_t\right)\right) \mathbb{1}_A\right), \tag{2.6}$$

for every $A \in \mathcal{F}_t^X$.

Let

$$\mathfrak{D} := \left\{A \in \mathcal{F}_t^X \mid \mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mathbb{1}_A\right) = \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mid \sigma\left(X_t\right)\right) \mathbb{1}_A\right),\ s \geq 0,\ \Gamma \in \mathcal{E}\right\}.$$

Clearly, $\mathfrak{D} \subseteq \mathcal{F}_t^X$. In the remaining part, we will use Lemma 2.2.1 to prove that $\mathfrak{D} \supseteq \mathcal{F}_t^X$. First, we show that $\mathfrak{D}$ is a λ-system. Note that $\Omega \in \mathfrak{D}$,

$$\mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mathbb{1}_\Omega\right) = \mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}}\right) = \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mid X_t\right)\right) = \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s} \in \Gamma\}} \mid X_t\right) \mathbb{1}_\Omega\right).$$

Assume that $A, B \in \mathfrak{D}$ and $B \subseteq A$, then $A \setminus B \in \mathfrak{D}$,

$$\begin{aligned}
\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}}\mathbb{1}_{A\setminus B}\right) &= \mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}}\left(\mathbb{1}_A - \mathbb{1}_B\right)\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}}\mathbb{1}_A\right) - \mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}}\mathbb{1}_B\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}} \mid X_t\right)\mathbb{1}_A\right) - \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}} \mid X_t\right)\mathbb{1}_B\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}} \mid X_t\right)\left(\mathbb{1}_A - \mathbb{1}_B\right)\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}} \mid X_t\right)\mathbb{1}_{A\setminus B}\right).
\end{aligned}$$

Suppose that $A_1, A_2 \ldots$ is a sequence of sets in $\mathfrak{D}$, and $A_n \subseteq A_{n+1}$ for all $n \geq 1$. We construct a new pairwise disjoint sequence as follows,

$$\begin{aligned}
B_1 &= A_1, \\
B_2 &= A_2 \setminus A_1, \\
&\vdots \\
B_n &= A_n \setminus A_{n-1}.
\end{aligned}$$

Then $(B_n)_{n\geq 1} \in \mathfrak{D}$ and $\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} B_n$.

$$\begin{aligned}
\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}}\mathbb{1}_{\cup_{n=1}^{\infty}A_n}\right) &= \mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}}\mathbb{1}_{\cup_{n=1}^{\infty}B_n}\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}}\sum_{n=1}^{\infty}\mathbb{1}_{B_n}\right) \\
&= \sum_{n=1}^{\infty}\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}}\mathbb{1}_{B_n}\right) \\
&= \sum_{n=1}^{\infty}\mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}} \mid X_t\right)\mathbb{1}_{B_n}\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}} \mid X_t\right)\sum_{n=1}^{\infty}\mathbb{1}_{B_n}\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}} \mid X_t\right)\mathbb{1}_{\cup_{n=1}^{\infty}B_n}\right) \\
&= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t+s}\in\Gamma\}} \mid X_t\right)\mathbb{1}_{\cup_{n=1}^{\infty}A_n}\right).
\end{aligned}$$

Thus $\cup_{n=1}^{\infty} A_n \in \mathfrak{D}$. $\mathfrak{D}$ is a λ-system.

We define

$$\mathcal{G} = \{A = \{X_{t_n} = x_n,\ X_{t_{n-1}} = x_{n-1}, \ldots, X_0 = x_0; \pi\} \mid x_i \in E,\ i = 0, 1, \ldots, n \in \mathbb{N}\}.$$

For any $A, B \in \mathcal{G}$, $A \cap B \subseteq A \in \mathcal{G}$. The set $\mathcal{G}$ is closed under finite intersections. Thus $\mathcal{G}$ is a π-system. Since the statement (ii) holds true, it follows that for every $A \in \sigma\left(X_{t_n}, X_{t_{n-1}}, \ldots, X_0; \pi\right)$, $\pi \in \Pi_t$,

$$\begin{aligned}\mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t_{n+1}} \in \Gamma\}} \mid \sigma\left(X_{t_n}\right)\right) \mathbb{1}_A\right) &= \mathbb{E}\left(\mathbb{1}_{\{X_{t_{n+1}} \in \Gamma\}} \mathbb{1}_A\right) \\ &= \mathbb{E}\left(\mathbb{E}\left(\mathbb{1}_{\{X_{t_{n+1}} \in \Gamma\}} \mid \sigma\left(X_{t_n}, X_{t_{n-1}}, \ldots, X_0; \pi\right)\right) \mathbb{1}_A\right).\end{aligned}$$

Note that any element in $\mathcal{G}$ is also in $\mathcal{F}_t^X$. Thus, $\mathcal{G} \subseteq \mathfrak{D}$. By Lemma 2.2.1, $\sigma(\mathcal{G}) \subseteq \mathfrak{D}$. Moreover, $\sigma(\mathcal{G}) = \mathcal{F}_t^X$. Hence, $\mathcal{F}_t^X = \sigma(\mathcal{G}) \subseteq \mathfrak{D}$. Together with $\mathfrak{D} \subseteq \mathcal{F}_t^X$, we conclude that $D = \mathcal{F}_t^X$. Equation (2.6) holds for any $A \in \mathcal{F}_t^X$. Thus (ii) holds true, and this concludes the proof. □

2.3 Main Problems

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete underlying probability space. Throughout the thesis, we fix a positive integer m, and a finite time horizon $T > 0$. Let $\mathcal{M} = \{1, 2, \ldots, m\}$. We consider a collection of finite sets E_i, $i = 1, 2, \ldots, m$. Next, let $E = E_1 \times E_2 \times \cdots \times E_m$ be the Cartesian product. The elements of E are denoted by $z = (z^1, \ldots, z^m)$. We define a set

$$\mathcal{H}\left(z^i\right) = \left\{\left(z^1, \ldots, z^{i-1}, z^i, z^{i+1}, \ldots, z^m\right) \mid z^j \in E_j,\ j = 1, 2, \ldots, m,\ j \neq i\right\}.$$

We consider a continuous time multivariate Markov chain $X = (X^1, \ldots, X^m)$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in E. Thus, we have that for any $t, u \geq 0$ and for every $\Gamma \subset E$ it holds

$$\mathbb{P}\left(X_{t+u} \in \Gamma \mid \mathcal{F}_t^X\right) = \mathbb{P}\left(X_{t+u} \in \Gamma \mid X_t\right). \tag{2.7}$$

Equivalently, as shown in Proposition 2.2.1, the Markov property (2.7) can be written as

$$\mathbb{P}\left(X_{t_{n+1}} = x_{n+1} \mid X_{t_n} = x_n,\ X_{t_{n-1}} = x_{n-1}, \ldots,\ X_0 = x_0\right) \\ = \mathbb{P}\left(X_{t_{n+1}} = x_{n+1} \mid X_{t_n} = x_n\right),$$

for every $n \in \mathbb{N}$, $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n \leq t_{n+1}$ and $x_j \in E$, $j = 0, 1, \ldots, n+1$, and whenever the conditional probabilities on the both sides are well-defined.

The initial distribution μ^X is the probability distribution of X at time 0. Formally, μ^X is a function mapping a subset of E into an interval $[0, 1]$,

$$\mu^X(\Gamma) := \mathbb{P}\left(X_0 \in \Gamma\right) \geq 0, \quad \Gamma \subset E,$$

Alternatively, we can think of μ^X as a vector whose entry is $\mu^X(z) = \mathbb{P}\left(X_0 = z\right),\ z \in E$, such that

$$\sum_{z \in E} \mu^X(z) = 1.$$

We denote by $\mathbf{P}_{t,s} = \left[\mathbf{P}_{t,s}^{x,x'}\right]_{x,x' \in E}$, $0 \leq t \leq s \leq T$, the transition matrix function of X, whose entry is

$$\mathbf{P}_{t,s}^{x,x'} = \mathbb{P}(X_s = x' \mid X_t = x).$$

We assume that for each $t \in [0, T)$ the following limit exists outside of a set of Lebesgue zero,

$$\Lambda_t := \lim_{h \downarrow 0} \frac{\mathbf{P}_{t,t+h} - \mathbf{I}_{|E| \times |E|}}{h}, \tag{2.8}$$

where $\mathbf{I}_{|E| \times |E|}$ is the identity matrix of dimension $|E|$-by-$|E|$, with $|E|$ the cardinality of set E. On the set of Lebesgue zero we take $\Lambda_t = 0$. Followed by (2.8), for all $x, x' \in E$, the entry $\lambda_t^{x,x'} \geq 0$ for $x \neq x'$, $\lambda_t^{x,x} \leq 0$, and

$$\sum_{x' \in E} \lambda_t^{x,x'} = 0, \quad x \in E.$$

The matrix function $\Lambda_t := \left[\lambda_t^{x,x'}\right]$, $t \in [0,T)$, defined in (2.8) is called the infinitesimal generator function of X. We will also postulate measurability and some mild integrable properties about the function Λ_t, $t \in [0,T)$. Namely, each entry $\lambda_t^{x,x'}$ is integrable on $t \in [0,T)$. Then the first derivatives of $\mathbf{P}_{t,s}$ with respect to t and with respect to s exist. Moreover, the Kolmogorov forward equations and the Kolmogorov backward equations are satisfied, particularly, in matrix notation,

$$\frac{\partial}{\partial s}\mathbf{P}_{t,s} = \mathbf{P}_{t,s}\Lambda_s, \quad \mathbf{P}_{t,t} = \mathbf{I}_{|E|\times|E|},$$

and

$$\frac{\partial}{\partial t}\mathbf{P}_{t,s} = -\Lambda_t\mathbf{P}_{t,s}, \quad \mathbf{P}_{s,s} = \mathbf{I}_{|E|\times|E|}.$$

Or, equivalently, in the integral form,

$$\mathbf{P}_{t,s} = \mathbf{I} + \int_t^s \mathbf{P}_{t,u}\Lambda_u du \tag{2.9}$$

$$= \mathbf{I} + \sum_{n=1}^{\infty}\int_t^s\int_t^{u_1}\cdots\int_t^{u_{n-1}} \Lambda_{u_1}\cdots\Lambda_{u_n}\,\mathrm{d}u_n\cdots\mathrm{d}u_2\,\mathrm{d}u_1, \tag{2.10}$$

and

$$\mathbf{P}_{t,s} = \mathbf{I} + \int_t^s \Lambda_u\mathbf{P}_{u,s} du \tag{2.11}$$

$$= \mathbf{I} + \sum_{n=1}^{\infty}\int_t^s\int_{u_1}^{s}\cdots\int_{u_{n-1}}^{s} \Lambda_{u_1}\cdots\Lambda_{u_n}\,\mathrm{d}u_n\cdots\mathrm{d}u_2\,\mathrm{d}u_1, \tag{2.12}$$

Equations (2.10) and (2.12) can be obtained by [RSST98, Lemma 8.4.1 and Theorem 8.4.4] or by Peano-Baker series. It follows from Equations (2.9) and (2.11) there is one-to-one relation between the matrix function $(\Lambda_u, u \geq 0)$ and the transition matrix function $(\mathbf{P}_{t,s}, 0 \leq t \leq s)$ of X. It is well known (cf. [RSST98, Section 8.4.2 and Theorem 8.4.5]) that one can construct an inhomogeneous Markov process with the state space E from the matrix function $(\Lambda_t, t \geq 0)$ and the given initial distribution μ^X. That is, the process $X = (X_t, t \geq 0)$ is characterized by the initial distribution μ^X and the infinitesimal generator function $(\Lambda_t, t \geq 0)$. The proof of existence can be

found in [Ios80, Chapter 2.1 and Chapter 7.1], and the uniqueness follows from the Picard-Lindelöf theorem.

Let γ be a function, $\gamma : E \to E'$, where E' is some set. One important mathematical question is to find under which conditions imposed on function γ such that the process $(\gamma(X_t), t \geq 0)$ is a Markov process. Such function γ is called a Markov function.

Remark 2.3.1. *The above question is more generally posed with respect to a Markov process X with values and a measurable space $(\mathcal{Z}, \boldsymbol{\mathcal{Z}})$ and a measurable function $\gamma : (\mathcal{Z}, \boldsymbol{\mathcal{Z}}) \to (\mathcal{Z}', \boldsymbol{\mathcal{Z}}')$.*

This thesis is devoted to the case when γ is a coordinate projection. In particular, $\gamma(x) = \phi^i(x) := x^i$ for some $i \in \{1, 2, \ldots, m\}$. The major goal of our research is to establish structural conditions such that the laws of its coordinates X^i's agree with prescribed marginal laws, and X^i is a Markov chain with respect to its own natural filtration $\mathbb{F}^{X^i}$. Formally, this can be stated as follows.

We want to find algebraic conditions on $(\Lambda_u, u \geq 0)$ (or, equivalently, conditions on $(\mathbf{P}_{t,s},\ 0 \leq t \leq s)$), and conditions on μ^X, such that each component X^i of X is a Markov chain in its own filtration $\mathbb{F}^{X^i}$, but not a Markov chain in the filtration $\mathbb{F}^X$. Moreover, we want to apply it to developing a new framework to study systemic risk.

2.4 Markovian Consistency

If the components of a multivariate Markov chain are Markovian, then we can analyze both the process X and its components X^i by the well-developed Markov machinery.

As mentioned earlier, we are interested in the case when each component

X^i of X is Markovian in its own filtration, but not the filtration $\mathbb{F}^X$. In [BJN13, Example 3.1] they show that both components of $X = (X^1, X^2)$ are Markovian in their own filtrations, as well as in the filtration $\mathbb{F}^X$. Whereas in [BJN13, Example 3.2] its components are Markovian in their own filtrations, but *not* in the filtration $\mathbb{F}^X$. [BJN13, Example 3.3] gives a case when X^2 is not Markovian neither in the filtration $\mathbb{F}^{X^2}$ or in the filtration $\mathbb{F}^X$.

The analysis of the situations described in these examples quoted above leads to two notions of Markovian consistency (cf. Bielecki et al. [BJN12] and Bielecki et al. [BJN13]). We state the following two definitions in the context of our Markov chain X.

Definition 2.4.1 (Strong Markovian consistency)**.** *A multivariate Markov chain X is strongly Markovian consistent relative to the component X^i if*

$$\mathbb{P}\left(X_s^i \in \Gamma^i \mid \mathcal{F}_t^X\right) = \mathbb{P}\left(X_s^i \in \Gamma^i \mid X_t^i\right), \tag{2.13}$$

for every $\Gamma^i \subset E_i$, $0 \leq t \leq s < \infty$. Equivalently,

$$\mathbb{P}\left(X_s^i \in \Gamma^i \mid X_t\right) = \mathbb{P}\left(X_s^i \in \Gamma^i \mid X_t^i\right), \quad \Gamma^i \subset E_i,\ 0 \leq t \leq s < \infty.$$

If (2.13) *holds for every $i \in \{1, \ldots, m\}$, then we say that X satisfies strong Markovian consistency property, or that X is strongly Markovian consistent.*

Although next concept is weaker than strong Markovian consistency, it arises naturally and is essential.

Definition 2.4.2 (Weak Markovian consistency)**.** *A multivariate Markov chain X is weakly Markovian consistent relative to the component X^i if*

$$\mathbb{P}\left(X_s^i \in \Gamma^i \mid \mathcal{F}_t^{X^i}\right) = \mathbb{P}\left(X_s^i \in \Gamma^i \mid X_t^i\right), \tag{2.14}$$

for every $\Gamma^i \subset E_i$, $0 \leq t \leq s < \infty$. *If* (2.14) *holds for every* $i \in \{1, \ldots, m\}$, *then we say that* X *satisfies weak Markovian consistency property, or that* X *is weakly Markovian consistent.*

Clearly, strong Markovian consistency implies weak Markovian consistency. However, the converse implication is not true in general. If a multivariate Markov chain is weakly Markovian consistent, but not strongly Markovian consistent, we say that this multivariate Markov chain satisfies *weak-only Markovian consistency* property.

Given the above definitions and terminology, we can now state that our primary goal is to study Markov chains which are weakly Markovian consistent but not strongly Markovian consistent. Accordingly, the first reasonable research problem is to find verifiable necessary and sufficient conditions in terms of $(\Lambda_t, t \geq 0)$ and μ^X for the component X^i of X to be a Markov process in its own filtration, and then to classify the situations when weak Markovian consistency holds and the strong one does not hold.

In the rest of this section we collect some relevant concepts from [BJN13] to lay the foundation for our work.

Definition 2.4.3. *For any* $i = 1, 2, \ldots, m$, *and* $t \geq 0$ *the operator* Θ_t^i *acting on any function* $g : E \to \mathbb{R}$ *is defined as*

$$\Theta_t^i g\left(x^i\right) = \mathbb{E}_{\mathbb{P}}\left(g\left(X_t\right) \mid X_t^i = x^i\right), \quad x^i \in E_i. \tag{2.15}$$

For any $i = 1, 2, \ldots, m$, *the extension operator* Φ^i *acting on any function* $f : E_i \to \mathbb{R}$ *is defined by*

$$\Phi^i f\left(x\right) = f\left(x^i\right), \quad x = \left(x^1, x^2, \ldots, x^m\right) \in E. \tag{2.16}$$

Definition 2.4.4. *For any* $i = 1, 2, \ldots, m$, *the projection operator* Υ^i *acting on the*

initial distribution μ^X *is defined as*

$$\Upsilon^i \mu^X(\Gamma) := \mathbb{P}\left(X_0^i \in \Gamma^i\right), \quad \Gamma \subset E,\ \Gamma^i \subset E_i. \tag{2.17}$$

Remark 2.4.1. *We have two remarks.*

(i) The purpose of operator Φ^i *is to extend a subset of* E_i *to the corresponding subsets of* E*. If we order the state spaces* E_i *and* E *in the same ordering, there is a unique representation for each* Φ^i*,* $i \in \{1, 2, \ldots, m\}$*.*

(ii) If we fix the same ordering of the state spaces E_i *and* E *for each* Φ^i *and* Θ_t^i*, then in our setup we have* $\Theta_t^i \Phi^i = \mathbf{I}_{|E_i| \times |E_i|}$*,* $i \in \{1, 2, \ldots, m\}$*,* $t \geq 0$*.*

The operator in (2.15) can be represented by an $|E_i|$-by-$|E|$ matrix:

$$\Theta_t^i = \left[\mathbb{P}\left(X_t = x \mid X_t^i = x^i\right)\right]_{x \in E,\ x^i \in E_i}.$$

Note that, in the above representation, for $t > 0$ and each fixed x^i, the sum of each row is 1.

Likewise, the operator in (2.16) can be represented by an $|E|$-by-$|E_i|$ matrix:

$$\Phi^i = \left[\mathbb{1}_{\Gamma^i}\left(\phi^i(x)\right)\right]_{x \in E,\ \Gamma^i \subset E_i},$$

where

$$\mathbb{1}_{\Gamma^i}\left(\phi^i(x)\right) = \mathbb{1}_{\Gamma^i}\left(x^i\right) = \begin{cases} 1, & \text{if } x^i \in \Gamma^i, \\ 0, & \text{if } x^i \notin \Gamma^i. \end{cases}$$

Remark 2.4.2. *Clearly, the operators* Θ_t^i *and* Φ^i *define probability kernels. Indeed, for any* $x^i \in E_i$ *and* $\Gamma \subset E$*, we have*

$$\Theta_t^i \mathbb{1}_\Gamma\left(x^i\right) = \mathbb{P}(X_t \in \Gamma \mid X_t^i = x^i), \tag{2.18}$$

where $\mathbb{1}_\Gamma$ *is the indicator function of the set* Γ*. Similarly, for any* $x \in E$ *and* $\Gamma^i \subset E_i$ *we know*

$$\Phi^i \mathbb{1}_{\Gamma^i}(x) = \mathbb{1}_{\Gamma^i}\left(x^i\right) = \begin{cases} 1, & \text{if } x^i \in \Gamma^i, \\ 0, & \text{if } x^i \notin \Gamma^i. \end{cases}$$

Note that $\Phi^i \mathbb{1}_\Omega(x) = 1$*. Consider* $\Gamma^i_n \subset E_i$ *and* $\Gamma^i_n \cap \Gamma^i_k = \emptyset$ *for* $n \neq k$*. Since* $|E_i| < \infty$*, we can put* $\Gamma^i_n = \emptyset$ *for* $n > |E_i|$*. For a sequence of subsets* $(\Gamma^i_n)_{n\geq 1}$ *of pairwise disjoint sets in* E_i*, countable additivity holds,*

$$\begin{aligned}\left(\Phi^i \mathbb{1}_{\bigcup_{n=1}^\infty \Gamma^i_n}\right)(x) &= \left(\Phi^i \mathbb{1}_{\bigcup_{n=1}^{|E_i|} \Gamma^i_n}\right)(x) = \mathbb{1}_{\bigcup_{n=1}^{|E_i|} \Gamma^i_n}\left(x^i\right) = \sum_{n=1}^{|E_i|} \mathbb{1}_{\Gamma^i_n}\left(x^i\right) \\ &= \sum_{n=1}^{\infty} \mathbb{1}_{\Gamma^i_n}\left(x^i\right) = \sum_{n=1}^{\infty} \left(\Phi^i \mathbb{1}_{\Gamma^i_n}\right)(x),\end{aligned}$$

where we use the convention that an empty sum is 0*.*

With the introduction of two operators Θ^i_t and Φ^i, we connect our research problem to these two operators. Let us fix i. Recall from [BJN13, Theorem 1.11] that if the component X^i of X is a Markov chain in its own filtration, the property known as weak Markovian consistency of X relative to X^i, then the generator matrix function of X^i, say Λ^i_t, is given as

$$\Lambda^i_t = \Theta^i_t \Lambda_t \Phi^i, \ t \geq 0. \tag{2.19}$$

In other words, Equation (2.19) is a necessary condition for the component X^i of X to be a Markov process in its own filtration, or, equivalently, for X to be weakly Markovian consistent with respect to X^i. We are seeking verifiable sufficient conditions in terms of $(\Lambda_t, t \geq 0)$ for the component X^i of X to be a Markov process in its own filtration.

The tower property of conditional expectations implies that if X is strongly Markovian consistent with respect to X^i, that is, if the component X^i is a Markov

chain in the filtration of X, then X is weakly Markovian consistent with respect to X^i. So, a sufficient condition for X to be strongly Markovian consistent with respect to X^i is also a sufficient condition for X to be weakly Markovian consistent with respect to X^i. Sufficient conditions for strong Markovian consistency have been formulated in an algebraic structure, and are given by so-called condition (M), which can be readily verified:

Condition (M). *We say the generator matrix function Λ_t satisfies condition (M), if Λ_t satisfies for every $t \geq 0$, all $i = 1, 2, \ldots, m$,*

$$(\mathbf{M}i) \qquad \sum_{y \in \mathcal{H}(y^i)} \lambda_t^{x,y} = \sum_{y \in \mathcal{H}(y^i)} \lambda_t^{\bar{x},y},$$

$$x^i, y^i \in E_i,\ x^i \neq y^i,\ x^j, \bar{x}^j \in E_j,\ x^j \neq \bar{x}^j,\ j = 1, 2, \ldots, m,\ j \neq i, \quad (2.20)$$

where

$$x = \left(x^1, \ldots, x^i, \ldots, x^m\right),\ \bar{x} = \left(\bar{x}^1, \ldots, x^i, \ldots, \bar{x}^m\right),\ y = \left(y^1, \ldots, y^i, \ldots, y^m\right).$$

We now formulate our main problem regarding Markovian consistency in terms of the following two questions:

(i) Under which conditions imposed on $(\Lambda_t, t \geq 0)$, is the component X^i of X a Markov chain[2]?

(ii) For which $(\Lambda_t, t \geq 0)$, is the component X^i of X not a Markov chain in the filtration of X?

In other words, we look for conditions under which X is weakly Markovian consistent with respect to X^i, but X is not strongly Markovian consistent with respect

[2]If this happens, then necessarily the generator function of X^i is given as $\Theta_t^i \Lambda_t \Phi^i,\ t \geq 0$.

to X^i. Probabilistic sufficient and necessary conditions are known for both strong and weak Markovian consistency of X with respect to X^i. So, in principle, we can answer our questions using these conditions. However, these conditions are typically impossible to verify. Thus, we look for verifiable conditions, ideally in an algebraic structure.

As said, Equation (2.19) provides a necessary condition for the weak Markovian consistency of X with respect to X^i. In this work, we want to give conditions for $(\Lambda_t, t \geq 0)$ so that the component X^i of X is a Markov chain in its own filtration, but not in the filtration of X. Thus, in particular, we want to give structural conditions for $(\Lambda_t, t \geq 0)$ so that Equation (2.20) is not satisfied, but X^i is a Markov process in its own filtration. Additionally, the matrix function $\Theta_t^i \Lambda_t \Phi^i,\ t \geq 0$, gives the infinitesimal generator of X^i. Needless to say, we aim at providing structural conditions for $(\Lambda_t, t \geq 0)$ that can be verified.

2.5 Conditions for Markovian Consistency

In this section, we study the conditions for Markovian consistency. We want to emphasize that any conditions for strong Markovian consistency certainly imply weak Markovian consistency. The conditions for weak Markovian consistency may or may not imply strong Markovian consistency.

2.5.1 Technical preliminaries. In view of the equation,

$$\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i, \quad i \in \{1, 2, \ldots, m\},$$

if a matrix Λ_t is given, $\Theta_t^i \Lambda_t \Phi^i$ defines a unique matrix Λ_t^i. In the following proposition, we give one sufficient condition for $\Theta_t^i \Lambda_t \Phi^i$ to be a valid generator.

Proposition 2.5.1. *If $(\Lambda_t, t \geq 0)$ is a valid infinitesimal generator matrix function, then the matrix $(\Theta_t^i \Lambda_t \Phi^i, t \geq 0)$, $i \in \{1, 2, \ldots, m\}$, defines a valid generator.*

Proof. Let $(\Lambda_t, t \geq 0)$ be a valid infinitesimal generator matrix with dimension $|E|$, $\Lambda_t = [\lambda_t(j,k)]_{t\in[0,\infty),\ j,k=1,2,\ldots,|E|}$. We fix $t \geq 0$ and $i \in \{1,2,\ldots,m\}$. We know $\lambda_t(j,k) \geq 0$ for $j \neq k$, and

$$\sum_{k=1}^{|E|} \lambda_t(j,k) = 0,$$

$$\lambda_t(j,j) = -\sum_{k\neq j} \lambda_t(j,k), \quad j,k = 1,2,\ldots,|E|\,.$$

Recall from Remark 2.4.2, the extension operator Φ^i acting on $\mathbb{1}_{\Gamma^i}$ for $\Gamma^i \subset E_i$ and $x \in E$ can be represented by

$$\Phi^i \mathbb{1}_{\Gamma^i}(x) = \mathbb{1}_{\Gamma^i}\left(x^i\right) = \begin{cases} 1, & \text{if } x^i \in \Gamma^i, \\ 0, & \text{if } x^i \notin \Gamma^i. \end{cases}$$

which is an $|E| \times |E_i|$ matrix. Thus,

$$\left(\Lambda_t \Phi^i\right)(j,l) = \sum_{k=1}^{|E|} \lambda_t(j,k)\Phi^i(k,l), \quad j,k = 1,2,\ldots,|E|\,,\ l = 1,2,\ldots,|E_i|\,.$$

The matrix multiplication is the left matrix Λ_t multiplying its rows into the columns of the right matrix Φ^i. Because the entries of Φ^i only take values in 0 or 1, the operation $\Lambda_t\Phi^i$ acts by summing up the corresponding entries of each row of Λ_t. The sum of each row in $\Lambda_t\Phi^i$ is still zero,

$$\begin{aligned}\sum_{l=1}^{|E_i|} \left(\Lambda_t\Phi^i\right)(j,l) &= \sum_{l=1}^{|E_i|}\sum_{k=1}^{|E|} \lambda_t(j,k)\Phi^i(k,l) = \sum_{k=1}^{|E|} \lambda_t(j,k) \sum_{l=1}^{|E_i|} \Phi^i(k,l) \\ &= \sum_{k=1}^{|E|} \lambda_t(j,k) \times \mathbf{1}_{|E|\times 1} = \mathbf{0}_{|E|\times 1},\end{aligned}$$

where $\mathbf{1}$ is a matrix with dimension $|E|$-by-1 whose all entries are 1, and $\mathbf{0}$ is a matrix with dimension $|E|$-by-1 whose all entries are 0.

In the sequel, we show that $\left(\Lambda_t\Phi^i\right)(j,l) \leq 0$ if and only if $\Phi^i(j,l) = 1$.

$$\left(\Lambda_t\Phi^i\right)(j,l) = \sum_{k=1}^{|E|} \lambda_t(j,k)\Phi^i(k,l)$$

$$
\begin{aligned}
&= \sum_{k=1}^{|E|} \lambda_t(j,k)\Phi^i(k,l)\left(\mathbb{1}_{\{k\neq j\}}(k) + \mathbb{1}_{\{k=j\}}(k)\right) \\
&= \sum_{k\neq j}^{|E|} \lambda_t(j,k)\Phi^i(k,l) + \lambda_t(j,j)\Phi^i(j,l) \\
&= \sum_{k\neq j}^{|E|} \lambda_t(j,k)\Phi^i(k,l) + \left(-\sum_{k\neq j}^{|E|} \lambda_t(j,k)\right)\Phi^i(j,l) \\
&= \sum_{k\neq j}^{|E|} \lambda_t(j,k)\left(\Phi^i(k,l) - \Phi^i(j,l)\right) \\
&\begin{cases} \leq 0, & \Phi^i(j,l) = 1, \\ \geq 0, & \Phi^i(j,l) = 0. \end{cases}
\end{aligned}
$$

If $\Phi^i(j,l) = 1$, then there exists at least one $k \neq j$ such that $\Phi^i(k,l) = 1$ and $(\Phi^i(k,l) - \Phi^i(j,l)) = 0$. For the rest $\Phi^i(k,l) = 0,\ k \neq j$, we have $(\Phi^i(k,l) - \Phi^i(j,l)) < 0$. Then $(\Lambda_t\Phi^i)(j,l) \leq 0$ because of $\lambda_t(j,k) \geq 0$. Analogically, if $\Phi^i(j,l) = 0$, there exists some $k \neq j$ such that $\Phi^i(k,l) = 1$ and $(\Phi^i(k,l) - \Phi^i(j,l)) = 1$, then $(\Lambda_t\Phi^i)(j,l) \geq 0$. Note that since Φ^i has only one entry with value 1 in each row, for each row of $\Lambda_t\Phi^i$, there is only one entry being non-positive.

Because Θ_t^i is a kernel, the operation $\Theta_t^i(\Lambda_t\Phi^i)$ performs convex combinations of corresponding columns of $\Lambda_t\Phi^i$,

$$
\left(\Theta_t^i\Lambda_t\Phi^i\right)(k,l) = \sum_{j=1}^{|E|} \Theta_t^i(k,j)\left(\Lambda_t\Phi^i\right)(j,l), \quad k,l = 1,2,\ldots,|E_i|.
$$

Next, we show that the sum of each row of $(\Theta_t^i\Lambda_t\Phi^i)$ is zero,

$$
\begin{aligned}
\sum_{l=1}^{|E_i|}\left(\Theta_t^i\Lambda_t\Phi^i\right)(k,l) &= \sum_{l=1}^{|E_i|}\sum_{j=1}^{|E|} \Theta_t^i(k,j)\left(\Lambda_t\Phi^i\right)(j,l) \\
&= \sum_{j=1}^{|E|} \Theta_t^i(k,j)\sum_{l=1}^{|E_i|}\left(\Lambda_t\Phi^i\right)(j,l) \\
&= \sum_{j=1}^{|E|} \Theta_t^i(k,j) \times \mathbf{0}_{|E|\times 1} = \mathbf{0}_{|E_i|\times 1}.
\end{aligned}
$$

It further implies that the diagonal term of $(\Theta_t^i \Lambda_t \Phi^i)$ is given by

$$\left(\Theta_t^i \Lambda_t \Phi^i\right)(k,k) = -\sum_{l \neq k}^{|E_i|} \left(\Theta_t^i \Lambda_t \Phi^i\right)(k,l).$$

It remains to show that $(\Theta_t^i \Lambda_t \Phi^i)(k,l) \geq 0$ for each $k \neq l$. Recall that $\Theta_t^i(k,j) \geq 0$ for all k, j, and the fact that

$$\left(\Lambda_t \Phi^i\right)(j,l) \begin{cases} \leq 0, & \Phi^i(j,l) = 1, \\ \geq 0, & \Phi^i(j,l) = 0. \end{cases}$$

If $(\Lambda_t \Phi^i)(j,l) \leq 0$, the corresponding $\Theta_t^i(k,j) = 0$. Thus, we have for $k \neq j$,

$$\left(\Theta_t^i \Lambda_t \Phi^i\right)(k,l) = \sum_{j=1}^{|E|} \Theta_t^i(k,j) \left(\Lambda_t \Phi^i\right)(j,l) \geq 0.$$

Since $t \geq 0$ and $i \in \{1, 2, \ldots, m\}$ are arbitrary, we conclude that $(\Theta_t^i \Lambda_t \Phi^i, t \geq 0)$, $i \in \{1, 2, \ldots, m\}$, is a valid infinitesimal generator. □

Example 2.5.1. *Assume $E_i = \{0,1\}$, $i = 1,2$, and $E = E_1 \times E_2$. Consider a bivariate Markov chian $X = (X^1, X^2)$ taking values in E with the infinitesimal generator $(\Lambda_t, t \geq 0)$ of the form,*

$$\Lambda_t = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -\sum_{k=2}^{4} \lambda_t(1,k) & \lambda_t(1,2) & \lambda_t(1,3) & \lambda_t(1,4) \\ \lambda_t(2,1) & -\sum_{k=1,3,4} \lambda_t(2,k) & \lambda_t(2,3) & \lambda_t(2,4) \\ \lambda_t(3,1) & \lambda_t(3,2) & -\sum_{k=1,2,4} \lambda_t(3,k) & \lambda_t(3,4) \\ \lambda_t(4,1) & \lambda_t(4,2) & \lambda_t(4,3) & -\sum_{k=1}^{3} \lambda_t(4,k) \end{pmatrix} \end{array},$$

where $\lambda_t(j,k) \geq 0$, $j \neq k$, $j,k = 1,2,3,4$.

We fix $i = 1$ and $t \geq 0$. The extension operator Φ^1 can be represented by

$$\Phi^1 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \end{array}.$$

Then we have

$$\Lambda_t \Phi^1 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -\lambda_t(1,3)-\lambda_t(1,4) & \lambda_t(1,3)+\lambda_t(1,4) \\ -\lambda_t(2,3)-\lambda_t(2,4) & \lambda_t(2,3)+\lambda_t(2,4) \\ \lambda_t(3,1)+\lambda_t(3,2) & -\lambda_t(3,1)-\lambda_t(3,2) \\ \lambda_t(4,1)+\lambda_t(4,2) & -\lambda_t(4,1)-\lambda_t(4,2) \end{pmatrix} \end{array}.$$

Note that Θ_t^1 is of the form,

$$\Theta_t^1 = \left[\mathbb{P}\left(X_t = x_t \mid X_t^1 = x_t^1\right)\right]_{x_t \in E, x_t^1 \in E_1} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \theta_t(1,1) & \theta_t(1,2) & 0 & 0 \\ 0 & 0 & \theta_t(2,3) & \theta_t(2,4) \end{pmatrix} \end{array},$$

where $\theta_t(j,k) \geq 0$, and

$$\theta_t(1,1) + \theta_t(1,2) = 1, \qquad \theta_t(2,3) + \theta_t(2,4) = 1$$

Then we get

$$\Theta_t^1 \Lambda_t \Phi^1 =$$

$$\begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -\theta_t(1,1)(\lambda_t(1,3)+\lambda_t(1,4))-\theta_t(1,2)(\lambda_t(2,3)+\lambda_t(2,4)) & \theta_t(1,1)(\lambda_t(1,3)+\lambda_t(1,4))+\theta_t(1,2)(\lambda_t(2,3)+\lambda_t(2,4)) \\ \theta_t(2,3)(\lambda_t(3,1)+\lambda_t(3,2))+\theta_t(2,4)(\lambda_t(4,1)+\lambda_t(4,2)) & -\theta_t(2,3)(\lambda_t(3,1)+\lambda_t(3,2))-\theta_t(2,4)(\lambda_t(4,1)+\lambda_t(4,2)) \end{pmatrix} \end{array}.$$

Since $\lambda_t(j,k) \geq 0$, $j \neq k$, and $\theta_t(j,k) \geq 0$, the matrix $(\Theta_t^1 \Lambda_t \Phi^1, t \geq 0)$ is a valid infinitesimal generator. Likewise, we can show that $(\Theta_t^2 \Lambda_t \Phi^2, t \geq 0)$ is also a valid generator.

In view of Proposition 2.5.1, because the state space is finite, given a valid infinitesimal generator $(\Lambda_t, t \geq 0)$ of X, the matrix $\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i,\ t \geq 0$, is unique and defines one Markov chain in some filtration. By the Kolmogorov Existence Theorem, on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, there exists a unique in law Markov chain, say Y^i, generated by $(\Lambda_t^i, t \geq 0)$ with some initial distribution μ^{Y^i}. However, we cannot that X^i is generated by $(\Lambda_t^i, t \geq 0)$.

Since Y^i is a Markov chain, the first derivative of the transition matrix function of Y^i satisfies the Kolmogorov forward equation,

$$\frac{\partial}{\partial s}\mathbf{P}_{t,s}^{Y^i} = \mathbf{P}_{t,s}^{Y^i}\Lambda_s^i, \quad \mathbf{P}_{t,t}^{Y^i} = \mathbf{I}_{|E_i|\times|E_i|}.$$

Then we can solve for $\mathbf{P}_{t,s}^{Y^i}$. On the other hand, we know that the transition probability of the component X^i of X for $0 \leq t \leq s$, $x^i, \hat{x}^i \in E_i$, is given by

$$\begin{aligned}\mathbb{P}\left(X_s^i = \hat{x}^i \mid X_t^i = x^i\right) &= \frac{\mathbb{P}\left(X_s^i = \hat{x}^i,\ X_t^i = x^i\right)}{\mathbb{P}\left(X_t^i = x^i\right)} \\ &= \frac{\sum_{\hat{x}\in\mathcal{H}(\hat{x}^i)}\sum_{x\in\mathcal{H}(x^i)}\mathbb{P}\left(X_s = \hat{x},\ X_t = x\right)}{\mathbb{P}\left(X_t^i = x^i\right)},\end{aligned}$$

where $x = (x^1, \ldots, x^i, \ldots, x^m), \hat{x} = (\hat{x}^1, \ldots, \hat{x}^i, \ldots, \hat{x}^m) \in E$. The first derivative of transition matrix function of a Markov process satisfies the Chapman-Kolmogorov equation. But the fulfillment of the Chapman-Kolomogorov equation is not necessary for a process to be Markovian. An example can be found in [Ros71, Chapter III.1]. Even the transition semigroups of Y^i and X^i are the same, it does not conclude that X^i is a Markov chain. However, if the finite-dimensional distributions of Y^i and X^i are identical, then we conclude that the component X^i is indeed a Markov chain generated by $(\Theta_t^i \Lambda_t \Phi^i, t \geq 0)$ with initial distribution μ^{X^i},

$$\mu^{X^i}\left(x_0^i\right) := \mathbb{P}\left(X_0^i = x_0^i\right) = \sum_{x_0\in\mathcal{H}\left(x_0^i\right)}\mathbb{P}\left(X_0 = x_0\right) = \Upsilon^i\mu^X\left(x_0\right),$$

where $x_0 = \left(x_0^1, \ldots, x_0^{i-1}, x_0^i, x_0^{i+1}, \ldots, x_0^m\right)$.

The above arguments can be substantiated by the following lemmas. The first result is a known result. The proof is to apply Dynkin's π-λ theorem (cf. Theorem 2.2.1 on page 16) and use the similar techniques in the proof of Proposition 2.2.1 on page 19.

Lemma 2.5.1. *Let U and V be two $\mathcal{Z}$-valued processes on $[0,T]$. Then U and V have the same distributions if and only if all of their finite-dimensional distributions agree.*

Lemma 2.5.2. *We fix $i \in \{1,2,\ldots,m\}$. Let X^i and Y^i be two E_i-valued processes on $[0,T]$. Assume that Y^i is a Markov chain generated by $(\Lambda^i_u, u \geq 0)$. If $Y^i \overset{\mathcal{L}}{=} X^i$*[3]*, then X^i is a Markov chain generated by $(\Lambda^i_u,\ u \geq 0)$.*

Proof. In view of Lemma 2.5.1, we know that all the finite-dimensional distributions of X^i and Y^i are the same. For any $n \in \mathbb{N}$, any $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq t_{n+1}$, and $x^i_j \in E_i,\ j = 1,2,\ldots,n$, we have

$$
\begin{aligned}
&\mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1} \mid X^i_{t_n} = x^i_n\right) \\
&= \frac{\mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1},\ X^i_{t_n} = x^i_n\right)}{\mathbb{P}\left(X^i_{t_n} = x^i_n\right)} \\
&= \frac{\mathbb{P}\left(Y^i_{t_{n+1}} = x^i_{n+1},\ Y^i_{t_n} = x^i_n\right)}{\mathbb{P}\left(Y^i_{t_n} = x^i_n\right)} \\
&= \mathbb{P}\left(Y^i_{t_{n+1}} = x^i_{n+1} \mid Y^i_{t_n} = x^i_n\right) \\
&= \mathbb{P}\left(Y^i_{t_{n+1}} = x^i_{n+1} \mid Y^i_{t_n} = x^i_n,\ Y^i_{t_{n-1}} = x^i_{n-1}, \ldots, Y^i_{t_0} = x^i_0\right) \\
&= \frac{\mathbb{P}\left(Y^i_{t_{n+1}} = x^i_{n+1},\ Y^i_{t_n} = x^i_n,\ Y^i_{t_{n-1}} = x^i_{n-1}, \ldots, Y^i_{t_0} = x^i_0\right)}{\mathbb{P}\left(Y^i_{t_n} = x^i_n,\ Y^i_{t_{n-1}} = x^i_{n-1}, \ldots, Y^i_{t_0} = x^i_0\right)} \\
&= \frac{\mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1},\ X^i_{t_n} = x^i_n,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots, X^i_{t_0} = x^i_0\right)}{\mathbb{P}\left(X^i_{t_n} = x^i_n,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots, X^i_{t_0} = x^i_0\right)} \\
&= \mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1} \mid X^i_{t_n} = x^i_n,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots, X^i_{t_0} = x^i_0\right),
\end{aligned}
$$

[3]Note that the symbol $\overset{\mathcal{L}}{=}$ means equality in law.

where the fourth equality comes from the fact that Y^i is Markovian. We conclude that X^i is Markov,

$$\mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1} \mid X^i_{t_n} = x^i_n,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots, X^i_{t_0} = x^i_0\right) \\ = \mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1} \mid X^i_{t_n} = x^i_n\right).$$

□

Another way to verify whether X^i is generated by $(\Theta^i_t \Lambda_t \Phi^i, t \geq 0)$ is by [BJN13, Theorem 1.11], which states that if the component X^i of X is a Markov chain, then the infinitesimal generator of X^i is given by $(\Theta^i_t \Lambda_t \Phi^i, t \geq 0)$. In what follows, we study under which conditions the component X^i of X is a Markov chain.

In view of Proposition 2.2.1 on page 19, the Markov property of X^i (with respect to its own filtration) is equivalent to the following statement: for any $n \in \mathbb{N}$, $0 \leq t_1 \leq \cdots \leq t_n \leq t_{n+1}$, $x^i_j \in E_i,\ j = 0, 1, \ldots, n+1$, we have

$$\mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1} \mid X^i_{t_n} = x^i_n,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots,\ X^i_0 = x^i_0\right) \\ = \mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1} \mid X^i_{t_n} = x^i_n\right), \quad (2.21)$$

whenever $\mathbb{P}\left(X^i_{t_n} = x^i_n,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots,\ X^i_0 = x^i_0\right) > 0$.

From now on, in order to simplify the presentation, but without any loss of generality, we assume $E_i = \mathcal{K} := \{0, 1, 2, \ldots, K\},\ i = 1, 2, \ldots, m$. We denote by Π_t a collection of all finite partitions of interval $[0, t]$, namely,

$$\Pi_t = \{\pi_t = \{0, t_1, \ldots, t_{n-1}, t\} \mid 0 = t_0 < t_1 < \cdots < t_n = t,\ n \in \mathbb{N}\}.$$

For $\pi_t \in \Pi_t$ with $|\pi_t| = n+1$, we define

$$X^i|_{\pi_t} := \left(X^i_{t_0}, X^i_{t_1}, \ldots, X^i_{t_{n-1}}, X^i_{t_n}\right)$$

and

$$X|_{\pi_t}^{(i)} := \left(X_{t_0}^i, X_{t_1}^i, \ldots, X_{t_{n-1}}^i, X_{t_n}^1, X_{t_n}^2, \ldots X_{t_n}^{i-1}, X_{t_n}^i, X_{t_n}^{i+1}, \ldots X_{t_n}^m\right).$$

Let $\psi = (\psi_0, \psi_1, \ldots, \psi_n) \in \mathcal{K}^{n+1}$ and $\hat{a} = (\hat{a}_1, \hat{a}_2, \ldots, \hat{a}_{m-1}) \in \mathcal{K}^{m-1}$. Then, we use the notation $X|_{\pi_t}^{(i)} \bowtie (\psi, \hat{a})$ for

$$\begin{aligned}\big(&X_{t_0}^i = \psi_0, \ldots, X_{t_{n-1}}^i = \psi_{n-1},\\ &\qquad X_{t_n}^1 = \hat{a}_1, \ldots,\ X_{t_n}^{i-1} = \hat{a}_{i-1},\ X_{t_n}^i = \psi_n,\ X_{t_n}^{i+1} = \hat{a}_{i+1}, \ldots, X_{t_n}^m = \hat{a}_{m-1}\big).\end{aligned}$$

Saying differently, the vector ψ is the realization of trajectory of X^i sampled according to the partition π_t, and $\hat{a}$ is the realization of X_{t_n} except for the component $X_{t_n}^i$. Note that by countable additivity of probability measures, we have

$$\mathbb{P}\left(X^i|_{\pi_t} = \psi\right) = \sum_{\hat{a} \in \mathcal{K}^{m-1}} \mathbb{P}\left(X|_{\pi_t}^{(i)} \bowtie (\psi, \hat{a})\right).$$

Next, for $n \in \mathbb{N}$ and $a \in \mathcal{K}$ we define

$$\Psi_n^a := \left\{\psi = (\psi_0, \psi_1, \ldots, \psi_n) \in \mathcal{K}^{n+1} \mid \psi_n = a\right\}.$$

Remark 2.5.1. *In view of* (2.21)*, in order to have well-defined conditional probabilities, we are solely interested in the vector* $\psi \in \Psi_{|\pi_t|}^a$ *such that*

$$\mathbb{P}\left(X^i|_{\pi_t} = \psi\right) > 0,$$

which implies $\mathbb{P}\left(X_{t_{|\pi_t|}}^i = a\right) > 0$.

2.5.2 Necessary and Sufficient Conditions for Weak Markovian Consistency. Recall that for any $0 \le t \le s$, $\mathbf{P}_{t,s}^{x,y}$ is the transition probability of X,

$$\mathbf{P}_{t,s}^{x,y} = \mathbb{P}\left(X_s = y \mid X_t = x\right), \quad x, y \in \mathcal{K}^m.$$

We fix $i \in \{1, 2, \ldots, m\}$. For any $0 \le t \le s$, any $x \in \mathcal{K}^m$ and any $y^i \in \mathcal{K}$, we define

$$\mathbf{P}_{t,s}^{x,y^i} := \sum_{y \in \mathcal{H}(y^i)} \mathbb{P}\left(X_s = \left(y^1, \ldots, y^{i-1}, y^i, y^{i+1}, \ldots, y^m\right) \mid X_t = x\right)$$

$$= \mathbb{P}\left(X_s^i = y^i \mid X_t = x\right), \tag{2.22}$$

where

$$\mathcal{H}\left(y^i\right) = \left\{\left(y^1, \ldots, y^{i-1}, y^i, y^{i+1}, \ldots, y^m\right) \mid y^j \in \mathcal{K},\ j = 1, 2, \ldots, m,\ j \neq i\right\}.$$

Remark 2.5.2. *Note that we can represent* $\mathbf{P}_{t,s}^{x,y^i}$ *in the equivalent form,*

$$\begin{aligned}
\mathbf{P}_{t,s}^{x,y^i} &= \sum_{y \in \mathcal{H}(y^i)} \mathbb{P}\left(X_s = \left(y^1, \ldots, y^{i-1}, y^i, y^{i+1}, \ldots, y^m\right) \mid X_t = x\right) \\
&= \mathbb{P}\left(\cup_{y \in \mathcal{H}(y^i)} \{X_s = \left(y^1, \ldots, y^{i-1}, y^i, y^{i+1}, \ldots, y^m\right)\} \mid X_t = x\right) \\
&= \mathbb{P}\left(X_s^i = y^i \mid X_t = x\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\{y^i\}}\left(X_s^i\right) \mid X_t = x\right) \\
&= \mathbb{E}\left(\Phi^i \mathbb{1}_{\{y^i\}}\left(X_s\right) \mid X_t = x\right) \\
&= \mathbf{P}_{t,s} \Phi^i \mathbb{1}_{\{y^i\}}\left(x\right).
\end{aligned}$$

We formally state *condition (Pi)* as when the conditional probability $\mathbf{P}_{t,s}^{x,y^i}$ is independent of x: for any $0 \leq t \leq s$, for any[4] $x, \bar{x} \in E,\ x \neq \bar{x}$ such that

$$\mathbf{P}_{t,s}^{x,y^i} = \mathbf{P}_{t,s}^{\bar{x},y^i}, \quad y^i \in E_i.$$

Condition (P). *The transition probability matrix function* $\mathbf{P}_{t,s}$ *satisfies for any* $0 \leq t \leq s < \infty$, *all* $i = 1, 2, \ldots, m$,

$$(\mathbf{P}i) \qquad \sum_{y \in \mathcal{H}(y^i)} \mathbf{P}_{t,s}^{x,y} = \sum_{y \in \mathcal{H}(y^i)} \mathbf{P}_{t,s}^{\bar{x},y},$$

$$x^i, y^i \in E_i,\ x^i \neq y^i,\ x^j, \bar{x}^j \in E_j,\ x^j \neq \bar{x}^j,\ j = 1, 2, \ldots, m,\ j \neq i, \tag{2.23}$$

where

$$x = \left(x^1, \ldots, x^i, \ldots, x^m\right),\ \bar{x} = \left(\bar{x}^1, \ldots, x^i, \ldots, \bar{x}^m\right),\ y = \left(y^1, \ldots, y^i, \ldots, y^m\right).$$

---

[4]Note that we do not consider the state $x \in E$, if for all $t \geq 0$, we have $\mathbb{P}\left(X_t = x\right) = 0$.

Throughout the thesis, the results are formulated for the component $X^i,\ i \in \{1, 2, \ldots, m\}$. The results for the other components $X^j,\ j \neq i$, can be derived similarly.

In the next theorem, we provide a necessary and sufficient condition for the component X^i of X to be a Markov chain. The conditions are formulated in terms of the transition probability matrix function of X. In this way, we can study weak Markovian consistency through the transition probability matrix $\mathbf{P}_{t,s}$. In particular, we can work on the structure of $\mathbf{P}_{t,s}$.

We state a condition,

Condition (Ci). *For any $0 \leq t \leq s$, any $x_t^i, x_s^i \in \mathcal{K}$, any $\pi_t \in \Pi_t$, and any $\psi \in \Psi_{|\pi_t|}^{x_t^i}$, the following equality holds,*

$$\sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \mathbf{P}_{t,s}^{x_t, x_s^i} \Xi^i \left(t, x_t, \pi_t, \psi, \mu^X\right) = 0, \tag{2.24}$$

where $x_t = \left(x_t^1, \ldots, x_t^{i-1}, x_t^i, x_t^{i+1}, \ldots, x_t^m\right)$,

$$\Xi^i \left(t, x_t, \pi_t, \psi, \mu^X\right) := \mathbb{P}\left(X|_{\pi_t}^{(i)} \bowtie (\psi, \hat{x}_t)\right) \mathbb{P}\left(X_t^i = x_t^i\right)$$

$$- \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \times \left(\sum_{\pi_t' \in \Pi_t, \varphi \in \Psi_{|\pi_t'|}^{x_t^i}} \mathbb{P}\left(X|_{\pi_t'}^{(i)} \bowtie (\varphi, \hat{x}_t)\right) \right),$$

and $\hat{x}_t = \left(x_t^1, \ldots, x_t^{i-1}, x_t^{i+1}, \ldots, x_t^m\right)$.

Theorem 2.5.1. *We fix $i \in \{1, 2, \ldots, m\}$. The component X^i is a Markov chain if and only if condition (Ci) holds.*

Proof. In view of Proposition 2.2.1, X^i is Markovian if and only if for any $n \in \mathbb{N}$, $0 \leq t_1 \leq \cdots \leq t_n \leq t_{n+1}$, $x_j^i \in \mathcal{K},\ j = 0, 1, \ldots, n+1$, we have

$$\mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1} \mid X^i_{t_n} = x^i_n,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots,\ X^i_0 = x^i_0\right)$$
$$= \mathbb{P}\left(X^i_{t_{n+1}} = x^i_{n+1} \mid X^i_{t_n} = x^i_n\right), \quad (2.25)$$

whenever $\mathbb{P}\left(X^i_{t_n} = x^i_n,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots,\ X^i_0 = x^i_0\right) > 0$.

Without loss of generality, we set $t = t_n$ and $s = t_{n+1}$. Then for any $x = (x^1, \ldots x^i, x^{i+1}, \ldots, x^m)$, we have

$$\begin{aligned}
&\mathbb{P}\left(X^i_s = x^i_s,\ X^i_t = x^i_t,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots,\ X^i_0 = x^i_0\right) \\
&= \sum_{x_s \in \mathcal{H}(x^i_s)} \mathbb{P}\big(X_s = \left(x^1_s, \ldots x^i_s, x^j_s, \ldots, x^m_s\right),\ X^i_t = x^i_t,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots,\ X^i_0 = x^i_0\big) \\
&= \sum_{x_s \in \mathcal{H}(x^i_s)} \sum_{x_t \in \mathcal{H}\left(x^i_t\right)} \mathbb{P}\big(X_s = x_s,\ X_t = x_t,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots,\ X^i_0 = x^i_0\big) \\
&= \sum_{x_s \in \mathcal{H}(x^i_s)} \sum_{x_t \in \mathcal{H}\left(x^i_t\right)} \cdots \sum_{x_0 \in \mathcal{H}\left(x^i_0\right)} \mathbb{P}\big(X_s = x_s,\ X_t = x_t,\ X_{t_{n-1}} = x_{n-1}, \ldots,\ X_0 = x_0\big) \\
&= \sum_{x_s \in \mathcal{H}(x^i_s)} \sum_{x_t \in \mathcal{H}\left(x^i_t\right)} \cdots \sum_{x_0 \in \mathcal{H}\left(x^i_0\right)} \mathbb{P}\big(X_s = x_s \mid X_t = x_t,\ X_{t_{n-1}} = x_{n-1}, \ldots,\ X_0 = x_0\big) \\
&\qquad\qquad \times \mathbb{P}\big(X_t = x_t,\ X_{t_{n-1}} = x_{n-1}, \ldots,\ X_0 = x_0\big) \\
&= \sum_{x_s \in \mathcal{H}(x^i_s)} \sum_{x_t \in \mathcal{H}\left(x^i_t\right)} \cdots \sum_{x_0 \in \mathcal{H}\left(x^i_0\right)} \mathbb{P}\big(X_s = x_s \mid X_t = x_t\big) \\
&\qquad\qquad \times \mathbb{P}\big(X_t = x_t,\ X_{t_{n-1}} = x_{n-1}, \ldots,\ X_0 = x_0\big) \\
&= \sum_{x_s \in \mathcal{H}(x^i_s)} \sum_{x_t \in \mathcal{H}\left(x^i_t\right)} \mathbb{P}\big(X_s = x_s \mid X_t = x_t\big)\mathbb{P}\big(X_t = x_t,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots,\ X^i_0 = x^i_0\big) \\
&= \sum_{x_s \in \mathcal{H}(x^i_s)} \sum_{x_t \in \mathcal{H}\left(x^i_t\right)} \mathbf{P}^{x_t, x_s}_{t,s} \mathbb{P}\big(X_t = x_t,\ X^i_{t_{n-1}} = x^i_{n-1}, \ldots,\ X^i_0 = x^i_0\big) \\
&= \sum_{x_s \in \mathcal{H}(x^i_s)} \sum_{x_t \in \mathcal{H}\left(x^i_t\right)} \mathbf{P}^{x_t, x_s}_{t,s} \mathbb{P}\left(X|^{(i)}_{\pi_t} \bowtie (\psi, \hat{x}_t)\right) \\
&= \sum_{x_t \in \mathcal{H}\left(x^i_t\right)} \mathbf{P}^{x_t, x^i_s}_{t,s} \mathbb{P}\left(X|^{(i)}_{\pi_t} \bowtie (\psi, \hat{x}_t)\right),
\end{aligned}$$

where $\pi_t \in \Pi_t$, $\psi = (x^i_0, x^i_1, \ldots, x^i_t) \in \Psi^{x^i_t}_{|\pi_t|}$, $\hat{x}_t = \left(x^1_t, \ldots, x^{i-1}_t, x^{i+1}_t, \ldots, x^m_t\right)$; and the fifth equality from the bottom is because X is a multivariate Markov chain.

It follows that for any $x_t = \left(x_t^1, \dots x_t^i, x_t^{i+1}, \dots, x_t^m\right)$ we get

$$\begin{aligned}\mathbb{P}\left(X_s^i = x_s^i,\ X_t^i = x_t^i\right) &= \sum_{\pi_t' \in \Pi_t, \varphi \in \Psi^{x_t^i}_{|\pi_t'|}} \mathbb{P}\left(X_s^i = x_s^i,\ X^i|_{\pi_t'} = \varphi\right) \\ &= \sum_{\pi_t' \in \Pi_t, \varphi \in \Psi^{x_t^i}_{|\pi_t'|}} \sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \mathbf{P}_{t,s}^{x_t, x_s^i} \mathbb{P}\left(X|^{(i)}_{\pi_t'} \bowtie (\varphi, \hat{x}_t)\right),\end{aligned}$$

where $\hat{x}_t = \left(x_t^1, \dots, x_t^{i-1}, x_t^{i+1}, \dots, x_t^m\right)$.

Since the conditional probabilities in both sides of (2.25) are well-defined, by reducing fractions we have

$$\begin{aligned}&\left(\sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \mathbf{P}_{t,s}^{x_t, x_s^i} \mathbb{P}\left(X|^{(i)}_{\pi_t} \bowtie (\psi, \hat{x}_t)\right)\right) \mathbb{P}\left(X_t^i = x_t^i\right) \\ &\qquad = \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \dots,\ X_0^i = x_0^i\right) \\ &\qquad\qquad \left(\sum_{\pi_t' \in \Pi_t, \varphi \in \Psi^{x_t^i}_{|\pi_t'|}} \sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \mathbf{P}_{t,s}^{x_t, x_s^i} \mathbb{P}\left(X|^{(i)}_{\pi_t'} \bowtie (\varphi, \hat{x}_t)\right)\right). \qquad (2.26)\end{aligned}$$

Note that Equation (2.26) needs to hold for all $x^i \in \mathcal{K}$ and all $\pi_t \in \Pi_t$. Equation (2.26) is equivalent to

$$\begin{aligned}\sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \mathbf{P}_{t,s}^{x_t, x_s^i} \Bigg[&\mathbb{P}\left(X|^{(i)}_{\pi_t} \bowtie (\psi, \hat{x}_t)\right) \mathbb{P}\left(X_t^i = x_t^i\right) \\ &- \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \dots,\ X_0^i = x_0^i\right) \\ &\times \left(\sum_{\pi_t' \in \Pi_t, \varphi \in \Psi^{x_t^i}_{|\pi_t'|}} \mathbb{P}\left(X|^{(i)}_{\pi_t'} \bowtie (\varphi, \hat{x}_t)\right)\right)\Bigg] = 0.\end{aligned}$$

The last equality holds if and only if condition (Ci) holds. □

Sufficient Conditions for Condition (Ci): Part I

We fix $i \in \{1, 2, \ldots, m\}$ throughout the section. In view of Theorem 2.5.1, condition (Ci) is the necessary and sufficient condition for X^i to be a Markov chain.

We will now provide a sufficient condition for (Ci) to hold.

Condition (Ci-I). *For any $0 \le t \le s$, fixed $x_t^i, x_s^i \in \mathcal{K}$, all $x_t = \left(x_t^1, \ldots, x_t^{i-1}, x_t^i, x_t^{i+1}, \ldots, x_t^m\right) \in \mathcal{K}^m$, $\widetilde{x}_t = \left(\widetilde{x}_t^1, \ldots, \widetilde{x}_t^{i-1}, x_t^i, \widetilde{x}_t^{i+1}, \ldots, \widetilde{x}_t^m\right) \in \mathcal{K}^m$ with $x_t \neq \widetilde{x}_t$, we have*

$$\mathbf{P}_{t,s}^{x_t, x_s^i} = \mathbf{P}_{t,s}^{\widetilde{x}_t, x_s^i}.$$

Proposition 2.5.2. *If (Ci-I) is satisfied, then for any $\pi_t \in \Pi_t$ and any $\psi \in \Psi_{|\pi_t|}^{x_t^i}$, Equation* (2.24) *holds true.*

Proof. In view of our assumption and Equation (2.24), we have

$$\begin{aligned}
&\sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \mathbf{P}_{t,s}^{x_t, x_s^i} \Xi^i \left(t, x_t, \pi_t, \psi, \mu^X\right) \\
&= \mathbf{P}_{t,s}^{x_t, x_s^i} \Bigg[\sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \left(\mathbb{P}\left(X|_{\pi_t}^{(i)} \bowtie (\psi, \hat{x}_t)\right) \mathbb{P}\left(X_t^i = x_t^i\right) \right) \\
&\qquad - \sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \Bigg(\mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \\
&\qquad\qquad \left(\sum_{\pi_t' \in \Pi_t, \varphi \in \Psi_{|\pi_t'|}^{x_t^i}} \mathbb{P}\left(X|_{\pi_t'}^{(i)} \bowtie (\varphi, \hat{x}_t)\right) \right) \Bigg) \Bigg] \\
&=: \mathbf{P}_{t,s}^{x_t, x_s^i} \left(I_1 - I_2\right).
\end{aligned}$$

Next, we show that $I_1 = I_2$.

$$
\begin{aligned}
I_1 &= \sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \left(\mathbb{P}\left(X_t = x_t,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \mathbb{P}\left(X_t^i = x_t^i\right) \right) \\
&= \mathbb{P}\left(X_t^i = x_t^i\right) \sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \mathbb{P}\left(X_t = x_t,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \\
&= \mathbb{P}\left(X_t^i = x_t^i\right) \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right).
\end{aligned}
$$

Likewise,

$$
\begin{aligned}
I_2 &= \sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \left(\mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \mathbb{P}\left(X_t = x_t\right) \right) \\
&= \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \mathbb{P}\left(X_t = x_t\right) \\
&= \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \mathbb{P}\left(X_t^i = x_t^i\right).
\end{aligned}
$$

This concludes the proof. □

Corollary 2.5.1. *If condition (Ci-I) is satisfied for all $x_t^i, x_s^i \in \mathcal{K}$, then condition (Ci) holds true.*

Remark 2.5.3. *We have several remarks.*

(i) Under the assumptions of Corollary 2.5.1 the component X^i is a Markov chain in its own filtration. Indeed, note that condition (Ci) is the necessary and sufficient condition for X^i to be a Markov chain in its own filtration. The assumptions in Corollary 2.5.1 are sufficient to condition (Ci).

(ii) The assumptions in Corollary 2.5.1 is a purely algebraic condition, which is related to the structure of the transition probability matrix $(\mathbf{P}_{t,s}, 0 \leq t \leq s)$.

(iii) Condition (P) is the same as condition (Ci-I) holding true for all $x_t^i, x_s^i \in \mathcal{K}$ and all $i \in \{1, 2, \ldots, m\}$. As we will show in Proposition 2.5.5 on page 60, condition (P) is equivalent to condition (M).

Note that condition (Ci) needs to hold for all $x_t^i, x_s^i \in \mathcal{K}$. It turns out though, as the following result shows, that if the component X^i admits absorbing states, say z, then one only needs to verify condition (Ci) for $x_t^i \neq z$.

Proposition 2.5.3. *If $z \in \mathcal{K}$ is an absorbing state for X^i, then for any $0 \leq t \leq s$, any $x_s^i \in \mathcal{K}$, any $\pi_t \in \Pi_t$, and any $\psi \in \Psi^z_{|\pi_t|}$,*

$$\sum_{x_t \in \mathcal{H}\left(x_t^i\right)} \mathbf{P}_{t,s}^{x_t, x_s^i} \Xi^i \left(t, x_t, \pi_t, \psi, \mu^X\right) = 0, \tag{2.27}$$

where $x_t = \left(x_t^1, \ldots, x_t^{i-1}, z, x_t^{i+1}, \ldots, x_t^m\right) \in \mathcal{K}^m$.

Proof. We consider any vector of the form $x_t = \left(x_t^1, \ldots, x_t^{i-1}, z, x_t^{i+1}, \ldots, x_t^m\right)$. In view of the conditional probability defined in Equation (2.22), we know

$$\mathbf{P}_{t,s}^{x_t, x_s^i} = \begin{cases} 0, & \text{if } x_s^i \neq z, \\ 1, & \text{if } x_s^i = z. \end{cases}$$

Then condition (Ci-I) holds for fixed $x_t^i = z$ and all x_s^i. As a result of Proposition 2.5.2, we have Equation (2.27). □

Sufficient Conditions for Condition (Ci): Part II

In view of Equation (2.24), if we have $\Xi^i \left(t, x_t, \pi_t, \psi, \mu^X\right) = 0$ for any input parameters, Equation (2.24) will hold true. In this section, we work on this condition.

Next, we state

Condition (Ci-II). Given μ^X, for any $t \geq 0$, any $x_t^i \in \mathcal{K}$, any $\pi_t \in \Pi_t$, any $\psi \in \Psi^{x_t^i}_{|\pi_t|}$, and for all $x_t = (x_t^1, \ldots, x_t^i, \ldots, x_t^m) \in \mathcal{K}^m$, we have $\Xi^i \left(t, x_t, \pi_t, \psi, \mu^X\right) = 0$.

Note that condition (Ci-II) demands to hold not only for any $t \geq 0$, but also any partitions π_t and $\psi \in \Psi^{x_t^i}_{|\pi_t|}$. Condition (C$i$-II) is a very strong condition.

Clearly, condition (Ci-II) implies condition (Ci). In this section, we provide sufficient conditions for condition (Ci-II) to hold.

We start by presenting equivalent statements of condition (Ci-II), that will be used later.

Lemma 2.5.3. *Condition (Ci-II) is equivalent to each of the following conditions:*

(i) For any $n \in \mathbb{N}$, $0 \leq t_1 \leq \cdots \leq t_n = t$, any $x_j^i \in \mathcal{K}$, $j = 0, 1, \ldots, n-1$, and any $x_t \in \mathcal{K}^m$, the identity holds

$$\mathbb{P}\left(X_t = x_t,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \mathbb{P}\left(X_t^i = x_t^i\right) - \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \mathbb{P}\left(X_t = x_t\right) = 0. \quad (2.28)$$

(ii) For any $x_t = (x_t^1, \ldots, x_t^i, \ldots, x_t^m) \in \mathcal{K}^m$, we have the identity

$$\mathbb{P}\left(X_t = x_t \mid \mathcal{F}_t^{X^i}\right) = \mathbb{P}\left(X_t = x_t \mid X_t^i\right), \quad t \geq 0. \quad (2.29)$$

(iii) For any $n \in \mathbb{N}$, $0 \leq t_1 \leq \cdots \leq t_n = t$, any $x_j^i \in \mathcal{K}$, $j = 0, 1, \ldots, n-1$, and any $x_0, x_t \in \mathcal{K}^m$, the identity holds true

$$\sum_{x_{n-1} \in \mathcal{H}\left(x_{n-1}^i\right)} \sum_{x_0 \in \mathcal{H}\left(x_0^i\right)} \mathbb{P}\left(X_0 = x_0\right) \mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X_{t_{n-2}}^i = x_{n-2}^i, \ldots, X_{t_1}^i = x_1^i \mid X_0 = x_0\right) \left[\mathbb{P}\left(X_t^i = x_t^i\right) \mathbf{P}_{t_{n-1},t_n}^{x_{n-1},x_t} - \mathbb{P}\left(X_t = x_t\right) \mathbf{P}_{t_{n-1},t_n}^{x_{n-1},x_t^i}\right] = 0. \quad (2.30)$$

Proof. Recall that

$$\Xi^i\left(t, x_t, \pi_t, \psi, \mu^X\right) = \mathbb{P}\left(X|_{\pi_t}^{(i)} \bowtie (\psi, \hat{x}_t)\right) \mathbb{P}\left(X_t^i = x_t^i\right) - \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \left(\sum_{\pi_t' \in \Pi_t, \varphi \in \Psi_{|\pi_t'|}^{x_t^i}} \mathbb{P}\left(X|_{\pi_t'}^{(i)} \bowtie (\varphi, \hat{x}_t)\right)\right),$$

where $\hat{x}_t = \left(x_t^1, \ldots, x_t^{i-1}, x_t^{i+1}, \ldots, x_t^m\right)$.

We establish statement (i) first. We fix $n \in \mathbb{N}$, $t \geq 0, x_t \in \mathcal{K}^m, \pi_t \in \Pi_t, \psi \in \Psi^{x_t^i}_{|\pi_t|}$. The equivalence follows from the definition of $X|^{(i)}_{\pi_t} \bowtie (\psi, \hat{x}_t)$ and by finite additivity. Namely, for any $n \in \mathbb{N}$, we have

$$\sum_{\pi'_t \in \Pi_t, \varphi \in \Psi^{x_t^i}_{|\pi'_t|}} \mathbb{P}\left(X|^{(i)}_{\pi'_t} \bowtie (\varphi, \hat{x}_t)\right) = \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right).$$

Thus, $\Xi^i\left(t, x_t, \pi_t, \psi, \mu^X\right) = 0$ is equivalent to

$$\begin{aligned}\mathbb{P}\left(X_t = x_t,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \mathbb{P}\left(X_t^i = x_t^i\right) \\ - \mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) \mathbb{P}\left(X_t = x_t\right) = 0.\end{aligned}$$

Next, we show that statement (i) and (ii) are equivalent. In view of statement (i), each probability is non-zero, then by reducing fractions we have

$$\frac{\mathbb{P}\left(X_t = x_t,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right)}{\mathbb{P}\left(X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right)} = \frac{\mathbb{P}\left(X_t = x_t\right)}{\mathbb{P}\left(X_t^i = x_t^i\right)},$$

which is the same as

$$\mathbb{P}\left(X_t = x_t \mid X_t^i = x_t^i,\ X_{t_{n-1}}^i = x_{n-1}^i, \ldots,\ X_0^i = x_0^i\right) = \mathbb{P}\left(X_t = x_t \mid X_t^i = x_t^i\right).$$

Since the above holds for any $n \in \mathbb{N}$, $0 \leq t_1 \leq \cdots \leq t_n = t$, any $x_j^i \in \mathcal{K}$, $j = 1, 2, \ldots, n-1$, and $x_t \in \mathcal{K}^m$ by Proposition 2.2.1 on page 19, we have statement (ii).

In the succeeding part, we show that statement (iii) is equivalent to statement (i). In the proof, we will use the fact that X is a Markov chain. We derive the formula,

$$\begin{aligned}&\mathbb{P}\left(X_0^i = x_0^i, \ldots, X_{t_{n-1}}^i = x_{n-1}^i,\ X_t = x_t\right) \\ &= \sum_{x_{n-1} \in \mathcal{H}\left(x_{n-1}^i\right)} \cdots \sum_{x_0 \in \mathcal{H}\left(x_0^i\right)} \mathbb{P}\left(X_0 = x_0, \ldots,\ X_{t_{n-1}} = x_{n-1},\ X_t = x_t\right) \\ &= \sum_{x_{n-1} \in \mathcal{H}\left(x_{n-1}^i\right)} \cdots \sum_{x_0 \in \mathcal{H}\left(x_0^i\right)} \mathbb{P}\left(X_t = x_t \mid X_{t_{n-1}} = x_{n-1}, \ldots, X_0 = x_0\right)\end{aligned}$$

$$\times \mathbb{P}\left(X_{t_{n-1}} = x_{n-1}, \ldots, X_0 = x_0\right)$$

$$= \sum_{x_{n-1} \in \mathcal{H}\left(x^i_{n-1}\right)} \cdots \sum_{x_0 \in \mathcal{H}\left(x^i_0\right)} \mathbb{P}\left(X_t = x_t \mid X_{t_{n-1}} = x_{n-1}\right)$$

$$\times \mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X_{t_{n-2}} = x_{n-2}, \ldots, X_0 = x_0\right)$$

$$= \sum_{x_{n-1} \in \mathcal{H}\left(x^i_{n-1}\right)} \mathbb{P}\left(X_t = x_t \mid X_{t_{n-1}} = x_{n-1}\right)$$

$$\times \mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X^i_{t_{n-2}} = x^i_{n-2}, \ldots, X^i_0 = x^i_0\right),$$

where the third equality follows from X being Markovian. Note that we can always fix $x_t = (x^1_t, \ldots, x^i_t, \ldots, x^m_t)$ in the following substitutions,

$$\mathbb{P}\left(X^i_t = x^i_t\right) = \mathbb{P}\left(X_t = x_t\right) + \sum_{\substack{z^j_t \in \mathcal{K}, z^j_t \neq x^j_t \\ j=1,\ldots,m, j \neq i}} \mathbb{P}\left(X_t = \left(z^1_t, \ldots, x^i_t, z^{i+1}_t, \ldots, z^m_t\right)\right),$$

and

$$\begin{aligned}
&\mathbb{P}\left(X^i_0 = x^i_0, \ldots, X^i_{t_{n-1}} = x^i_{n-1},\ X^i_t = x^i_t\right) \\
&= \mathbb{P}\left(X^i_0 = x^i_0, \ldots, X^i_{t_{n-1}} = x^i_{n-1},\ X_t = x_t\right) \\
&\quad + \sum_{\substack{z^j_t \in \mathcal{K}, z^j_t \neq x^j_t \\ j=1,\ldots,m, j \neq i}} \mathbb{P}\left(X^i_0 = x^i_0, \ldots, X^i_{t_{n-1}} = x^i_{n-1},\ X_t = \left(z^1_t, \ldots, x^i_t, z^{i+1}_t, \ldots, z^m_t\right)\right).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\sum_{\substack{z^j_t \in \mathcal{K}, z^j_t \neq x^j_t \\ j=1,\ldots,m, j \neq i}} \mathbb{P}\left(X^1_0 = x^i_0, \ldots, X^i_{t_{n-1}} = x^i_{n-1},\ X_t = \left(z^1_t, \ldots, x^i_t, \ldots, z^m_t\right)\right) \\
&\quad = \sum_{\substack{z^j_t \in \mathcal{K}, z^j_t \neq x^j_t \\ j=1,\ldots,m, j \neq i}} \sum_{x_{n-1} \in \mathcal{H}\left(x^i_{n-1}\right)} \mathbb{P}\left(X_t = \left(z^1_t, \ldots, x^i_t, \ldots, z^m_t\right) \mid X_{t_{n-1}} = x_{n-1}\right) \\
&\qquad \times \mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X^i_{t_{n-2}} = x^i_{n-2}, \ldots, X^i_0 = x^i_0\right).
\end{aligned}$$

Therefore, (2.28) is equivalent to

$$\mathbb{P}\left(X^i_0 = x^i_0,\ X^i_{t_1} = x^i_1, \ldots, X^i_{t_{n-1}} = x^i_{n-1},\ X_t = x_t\right) \mathbb{P}\left(X^i_t = x^i_t\right)$$

$$
\begin{aligned}
&- \mathbb{P}\left(X_t = x_t\right) \mathbb{P}\left(X_0^i = x_0^i,\ X_{t_1}^i = x_1^i, \ldots, X_{t_{n-1}}^i = x_{n-1}^i,\ X_t^i = x_t^i\right) \\
&= \mathbb{P}\left(X_0^i = x_0^i,\ X_{t_1}^i = x_1^i, \ldots, X_{t_{n-1}}^i = x_{n-1}^i,\ X_t = x_t\right) \\
&\qquad \left(\mathbb{P}\left(X_t = x_t\right) + \sum_{\substack{z_t^j \in \mathcal{K}, z_t^j \neq x_t^j \\ j=1,\ldots,m, j\neq i}} \mathbb{P}\left(X_t = \left(z_t^1, \ldots, x_t^i, \ldots, z_t^m\right)\right) \right) \\
&\quad - \mathbb{P}\left(X_t = x_t\right) \Bigg(\mathbb{P}\left(X_0^i = x_0^i, \ldots, X_{t_{n-1}}^i = x_{n-1}^i,\ X_t = x_t\right) \\
&\qquad + \sum_{\substack{z_t^j \in \mathcal{K}, z_t^j \neq x_t^j \\ j=1,\ldots,m, j\neq i}} \mathbb{P}\left(X_0^i = x_0^i, \ldots, X_{t_{n-1}}^i = x_{n-1}^i,\ X_t = \left(z_t^1, \ldots, x_t^i, \ldots, z_t^m\right)\right) \Bigg) \\
&= \Bigg(\sum_{x_{n-1} \in \mathcal{H}\left(x_{n-1}^i\right)} \mathbb{P}\left(X_t = x_t \mid X_{t_{n-1}} = x_{n-1}\right) \\
&\qquad\qquad \times \mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X_{t_{n-2}}^i = x_{n-2}^i, \ldots, X_0^i = x_0^i\right) \Bigg) \\
&\qquad\qquad\qquad \left(\sum_{\substack{z_t^j \in \mathcal{K}, z_t^j \neq x_t^j \\ j=1,\ldots,m, j\neq i}} \mathbb{P}\left(X_t = \left(z_t^1, \ldots, x_t^i, \ldots, z_t^m\right)\right) \right) \\
&\quad - \mathbb{P}\left(X_t = x_t\right) \\
&\qquad \Bigg(\sum_{\substack{z_t^j \in \mathcal{K}, z_t^j \neq x_t^j \\ j=1,\ldots,m, j\neq i}} \sum_{x_{n-1} \in \mathcal{H}\left(x_{n-1}^i\right)} \mathbb{P}\left(X_t = \left(z_t^1, \ldots, x_t^i, \ldots, z_t^m\right) \mid X_{t_{n-1}} = x_{n-1}\right) \\
&\qquad\qquad \times \mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X_{t_{n-2}}^i = x_{n-2}^i, \ldots, X_0^i = x_0^i\right) \Bigg) \\
&= \sum_{x_{n-1} \in \mathcal{H}\left(x_{n-1}^i\right)} \mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X_{t_{n-2}}^i = x_{n-2}^i, \ldots, X_0^i = x_0^i\right) \\
&\quad \Bigg[\left(\sum_{\substack{z_t^j \in \mathcal{K}, z_t^j \neq x_t^j \\ j=1,\ldots,m, j\neq i}} \mathbb{P}\left(X_t = \left(z_t^1, \ldots, x_t^i, \ldots, z_t^m\right)\right) \right) \times \mathbb{P}\left(X_t = x_t \mid X_{t_{n-1}} = x_{n-1}\right)
\end{aligned}
$$

$$
\left. - \mathbb{P}\left(X_t = x_t\right) \sum_{\substack{z_t^j \in \mathcal{K}, z_t^j \neq x_t^j \\ j=1,\ldots,m, j\neq i}} \mathbb{P}\left(X_t = \left(z_t^1, \ldots, x_t^i, \ldots, z_t^m\right) \mid X_{t_{n-1}} = x_{n-1}\right) \right]
$$

$$
= \sum_{x_{n-1} \in \mathcal{H}\left(x_{n-1}^i\right)} \sum_{x_0 \in \mathcal{H}\left(x_0^i\right)} \mathbb{P}\left(X_0 = x_0\right)
$$

$$
\mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X_{t_{n-2}}^i = x_{n-2}^i, \ldots, X_{t_1}^i = x_1^i \mid X_0 = x_0\right)
$$

$$
\left[\left(\sum_{\substack{z_t^j \in \mathcal{K}, z_t^j \neq x_t^j \\ j=1,\ldots,m, j\neq i}} \mathbb{P}\left(X_t = \left(z_t^1, \ldots, x_t^i, \ldots, z_t^m\right)\right) \right) \times \mathbb{P}\left(X_t = x_t \mid X_{t_{n-1}} = x_{n-1}\right) \right.
$$

$$
\left. - \mathbb{P}\left(X_t = x_t\right) \sum_{\substack{z_t^j \in \mathcal{K}, z_t^j \neq x_t^j \\ j=1,\ldots,m, j\neq i}} \mathbb{P}\left(X_t = \left(z_t^1, \ldots, x_t^i, \ldots, z_t^m\right) \mid X_{t_{n-1}} = x_{n-1}\right) \right]
$$

$$
= \sum_{x_{n-1} \in \mathcal{H}\left(x_{n-1}^i\right)} \sum_{x_0 \in \mathcal{H}\left(x_0^i\right)} \mathbb{P}\left(X_0 = x_0\right)
$$

$$
\mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X_{t_{n-2}}^i = x_{n-2}^i, \ldots, X_{t_1}^i = x_1^i \mid X_0 = x_0\right)
$$

$$
\left[\left(\mathbb{P}\left(X_t^i = x_t^i\right) - \mathbb{P}\left(X_t = x_t\right)\right) \mathbf{P}_{t_{n-1},t_n}^{x_{n-1},x_t} - \mathbb{P}\left(X_t = x_t\right) \left(\mathbf{P}_{t_{n-1},t_n}^{x_{n-1},x_t^i} - \mathbf{P}_{t_{n-1},t_n}^{x_{n-1},x_t}\right) \right]
$$

$$
= \sum_{x_{n-1} \in \mathcal{H}\left(x_{n-1}^i\right)} \sum_{x_0 \in \mathcal{H}\left(x_0^i\right)} \mathbb{P}\left(X_0 = x_0\right)
$$

$$
\mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X_{t_{n-2}}^i = x_{n-2}^i, \ldots, X_{t_1}^i = x_1^i \mid X_0 = x_0\right)
$$

$$
\left[\mathbb{P}\left(X_t^i = x_t^i\right) \mathbf{P}_{t_{n-1},t_n}^{x_{n-1},x_t} - \mathbb{P}\left(X_t = x_t\right) \mathbf{P}_{t_{n-1},t_n}^{x_{n-1},x_t^i} \right].
$$

We have statement (iii). □

Remark 2.5.4. *Identity* (2.29) *is exactly the same as our assumption* (1.1) *motivated by [BJN13, Theorem 1.6 and Theorem 1.11], and the same assumption appeared in [RP81]. In Section 3.4, we will discuss identity* (2.29) *in more details.*

Next, we give one sufficient condition for condition (Ci-II) with any initial distribution μ^X.

Proposition 2.5.4. *We fix $i \in \{1, 2, \ldots, m\}$. If the component X^i is independent of*

the component X^j, $j \neq i$, $j = 1, 2, \ldots, m$, then for any $t \geq 0$ the identity holds,

$$\mathbb{P}\left(X_t = x \mid \mathcal{F}_t^{X^i}\right) = \mathbb{P}\left(X_t = x \mid X_t^i\right), \tag{2.31}$$

where $x = (x^1, \ldots, x^i, \ldots, x^m) \in \mathcal{K}^m$.

Proof. Recall $\mathcal{M} = \{1, 2, \ldots, m\}$. In view of the assumption, we have that

$$\begin{aligned}
&\mathbb{P}\left(X_t = \left(x^1, \ldots, x^i, \ldots, x^m\right) \mid X_t^i\right) \\
&= \mathbb{P}\left(X_t^1 = x^1, \ldots, X_t^{i-1} = x^{i-1},\ X_t^{i+1} = x^{i+1}, \ldots, X_t^m = x^m\right) \mathbb{P}\left(X_t^i = x^i \mid X_t^i\right) \\
&= \mathbb{P}\left(X_t^1 = x^1, \ldots, X_t^{i-1} = x^{i-1},\ X_t^{i+1} = x^{i+1}, \ldots, X_t^m = x^m\right) \mathbb{E}_{\mathbb{P}}\left(\mathbb{1}_{\{x^i\}}\left(X_t^i\right) \mid X_t^i\right) \\
&= \mathbb{P}\left(X_t^1 = x^1, \ldots, X_t^{i-1} = x^{i-1},\ X_t^{i+1} = x^{i+1}, \ldots, X_t^m = x^m\right) \mathbb{1}_{\{x^i\}}\left(X_t^i\right) \\
&= \mathbb{P}\left(X_t^1 = x^1, \ldots, X_t^{i-1} = x^{i-1},\ X_t^{i+1} = x^{i+1}, \ldots, X_t^m = x^m\right) \mathbb{E}_{\mathbb{P}}\left(\mathbb{1}_{\{x^i\}}\left(X_t^i\right) \mid \mathcal{F}_t^{X^i}\right) \\
&= \mathbb{E}_{\mathbb{P}}\left(\mathbb{1}_{\{x^1\}}\left(X_t^1\right) \cdots \mathbb{1}_{\{x^{i-1}\}}\left(X_t^{i-1}\right) \mathbb{1}_{\{x^i\}}\left(X_t^i\right) \mathbb{1}_{\{x^{i+1}\}}\left(X_t^{i+1}\right) \cdots \mathbb{1}_{\{x^m\}}\left(X_t^m\right) \mid \mathcal{F}_t^{X^i}\right) \\
&= \mathbb{P}\left(X_t = \left(x^1, \ldots, x^i, \ldots, x^m\right) \mid \mathcal{F}_t^{X^i}\right),
\end{aligned}$$

where the first equality results from the independence of components. □

In what follows we will establish another sufficient condition to condition (Ci-II) in terms of the existence of the initial distribution μ^X and the algebraic structure of the semigroup $(\mathbf{P}_{t,s}, 0 \leq t \leq s)$ of X.

Lemma 2.5.4. *For any $0 \leq t \leq s$, the transition probability of the component X^i is given by*

$$\mathbf{P}_{t,s}^i\left(x^i, \hat{x}^i\right) := \Theta_t^i \mathbf{P}_{t,s} \Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x^i\right), \quad x^i, \hat{x}^i \in \mathcal{K}.$$

Proof. Note that

$$\begin{aligned}
\sum_{\hat{x} \in \mathcal{H}(\hat{x}^i)} \mathbf{P}_{t,s}^{x,\hat{x}} &= \sum_{\hat{x} \in \mathcal{H}(\hat{x}^i)} \mathbb{P}\left(X_s = \hat{x} \mid X_t = x\right) \\
&= \mathbb{P}\left(\cup_{\hat{x} \in \mathcal{H}(\hat{x}^i)} \{X_s = \left(\hat{x}^1, \ldots, \hat{x}^{i-1}, \hat{x}^i, \hat{x}^{i+1} \ldots \hat{x}^m\right)\} \mid X_t = x\right)
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{P}\left(X_s^i = \hat{x}^i \mid X_t = x\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\{\hat{x}^i\}}\left(X_s^i\right) \mid X_t = x\right) \\
&= \mathbb{E}\left(\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(X_s\right) \mid X_t = x\right) \\
&= \mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x\right),
\end{aligned}
$$

where $\hat{x} = (\hat{x}^1, \ldots \hat{x}^i, \ldots \hat{x}^m)$. Let $x = (x^1, \ldots, x^i, \ldots, x^m)$. For any $0 \leq t \leq s$ and any $x^i, \hat{x}^i \in \mathcal{K}$, the transition probability of X^i is given by

$$
\begin{aligned}
\mathbb{P}\left(X_s^i = \hat{x}^i \mid X_t^i = x^i\right) &= \frac{\mathbb{P}\left(X_s^i = \hat{x}^i,\ X_t^i = x^i\right)}{\mathbb{P}\left(X_t^i = x^i\right)} \\
&= \frac{\sum_{x \in \mathcal{H}(x^i)} \sum_{\hat{x} \in \mathcal{H}(\hat{x}^i)} \mathbb{P}\left(X_s = \hat{x},\ X_t = x\right)}{\mathbb{P}\left(X_t^i = x^i\right)} \\
&= \frac{\sum_{x \in \mathcal{H}(x^i)} \left(\sum_{\hat{x} \in \mathcal{H}(\hat{x}^i)} \mathbb{P}\left(X_t = x\right) \mathbf{P}_{t,s}^{x,\hat{x}}\right)}{\mathbb{P}\left(X_t^i = x\right)} \\
&= \frac{\sum_{x \in \mathcal{H}(x^i)} \left(\mathbb{P}\left(X_t = x\right) \sum_{\hat{x} \in \mathcal{H}(\hat{x}^i)} \mathbf{P}_{t,s}^{x,\hat{x}}\right)}{\mathbb{P}\left(X_t^i = x^i\right)} \\
&= \frac{\sum_{x \in \mathcal{H}(x^i)} \left(\mathbb{P}\left(X_t = x\right) \mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x\right)\right)}{\mathbb{P}\left(X_t^i = x^i\right)} \\
&= \sum_{x \in \mathcal{H}(x^i)} \frac{\mathbb{P}\left(X_t = x\right) \mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x\right)}{\mathbb{P}\left(X_t^i = x^i\right)} \\
&= \sum_{x \in \mathcal{H}(x^i)} \mathbb{E}\left(\mathbb{1}_{\{X_t = x\}} \mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x\right) \mid X_t^i = x^i\right) \\
&= \mathbb{E}\left(\sum_{x \in \mathcal{H}(x^i)} \left(\mathbb{1}_{\{X_t = x\}} \mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x\right)\right) \,\middle|\, X_t^i = x^i\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\{X_t^i = x^i\}} \mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x\right) \mid X_t^i = x^i\right) \\
&= \mathbb{1}_{\{X_t^i = x^i\}} \mathbb{E}\left(\mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x\right) \mid X_t^i = x^i\right) \\
&= \mathbb{E}\left(\mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x\right) \mid X_t^i = x^i\right) \\
&= \Theta_t^i \mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x^i\right).
\end{aligned}
$$

Since $i \in \{1, 2, \ldots, m\}$ and $x^i, \hat{x}^i \in \mathcal{K}$ are arbitrary, we conclude that for any $0 \leq t \leq s$

and any $x^i, \hat{x}^i \in \mathcal{K}$,

$$\mathbb{P}\left(X_s^i = \hat{x}^i \mid X_t^i = x^i\right) = \Theta_t^i \mathbf{P}_{t,s} \Phi^i \mathbb{1}_{\{\hat{x}^i\}}\left(x^i\right).$$

□

Note that we do *not* know if $\left(\mathbf{P}_{t,s}^i, 0 \le t \le s\right)$ is a semigroup or not. Next, we establish conditions for $\left(\mathbf{P}_{t,s}^i, 0 \le t \le s\right)$ to be a semigroup.

We define $\widehat{\Lambda}_u^i := \Theta_u^i \Lambda_u \Phi^i,\ u \ge 0$. It follows from Proposition 2.5.1 on page 31 that $\left(\widehat{\Lambda}_u^i, u \ge 0\right)$ is a valid generator, which generates some Markov chain Y^i. We denote by $\widehat{\mathbf{P}}_{t,s}^i,\ 0 \le t \le s$, the transition semigroup of Y^i. We know that $\widehat{\mathbf{P}}_{t,s}^i$ satisfies the Kolmogorov forward equation,

$$\begin{aligned} \frac{\partial}{\partial s}\widehat{\mathbf{P}}_{t,s}^i &= \widehat{\mathbf{P}}_{t,s}^i \widehat{\Lambda}_s^i = \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i \Lambda_s \Phi^i, \\ \widehat{\mathbf{P}}_{t,t}^i &= \mathbf{I}_{|\mathcal{K}|\times|\mathcal{K}|}. \end{aligned} \tag{2.32}$$

Lemma 2.5.5. *Let* $\left(\widehat{\mathbf{P}}_{t,s}^i, 0 \le t \le s\right)$ *be the transition semigroup of* $\widehat{\Lambda}_u^i := \Theta_u^i \Lambda_u \Phi^i$, $u \ge 0$. *If for any* $0 \le t \le s$ *it holds that*

$$\Theta_t^i \mathbf{P}_{t,s} = \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i, \tag{2.33}$$

then we have

(i) $\mathbf{P}_{t,s}^i = \widehat{\mathbf{P}}_{t,s}^i$,

(ii) $\mathbf{P}_{t,s}^i$ *is a semigroup.*

Proof. We first prove (i). In view of Lemma 2.5.4, since differentiation and expectation are linear, for any $x^i, \hat{x}^i \in \mathcal{K}$ the first derivative of $\mathbf{P}_{t,s}^i$ with respect to s is given by

$$\frac{\partial}{\partial s}\mathbf{P}_{t,s}^i\left(x^i, \hat{x}^i\right) = \frac{\partial}{\partial s}\left(\Theta_t^i \mathbf{P}_{t,s} \Phi^i \mathbb{1}_{\{\hat{x}^i\}}\right)\left(x^i\right)$$

$$= \Theta_t^i \frac{\partial}{\partial s} \mathbf{P}_{t,s} \Phi^i \mathbb{1}_{\{\hat{x}^i\}} \left(x^i\right)$$
$$= \Theta_t^i \mathbf{P}_{t,s} \Lambda_s \Phi^i \mathbb{1}_{\{\hat{x}^i\}} \left(x^i\right),$$

where the last equality comes from the fact that the transition semigroup $\mathbf{P}_{t,s}$ satisfies the Kolmogorov forward equation,

$$\frac{\partial}{\partial s} \mathbf{P}_{t,s} = \mathbf{P}_{t,s} \Lambda_s, \qquad \mathbf{P}_{t,t} = \mathbf{I}_{|\mathcal{K}^2| \times |\mathcal{K}^2|}.$$

Moreover, the initial value of $\mathbf{P}_{t,t}^i$ follows from the definition,

$$\mathbf{P}_{t,t}^i \left(x^i, \hat{x}^i\right) := \Theta_t^i \mathbf{P}_{t,t} \Phi^i \mathbb{1}_{\{\hat{x}^i\}} \left(x^i\right) = \begin{cases} 1, & x^i = \hat{x}^i, \\ 0, & x^i \neq \hat{x}^i. \end{cases}$$

Since the transition probability $\mathbf{P}_{t,s}^i \left(x^i, \hat{x}^i\right)$ and the the first derivative of $\mathbf{P}_{t,s}^i \left(x^i, \hat{x}^i\right)$ with respect to s are defined for all $x^i, \hat{x}^i \in \mathcal{K}$, we represent the transition probability by

$$\mathbf{P}_{t,s}^i = \Theta_t^i \mathbf{P}_{t,s} \Phi^i,$$

and the first derivative of $\mathbf{P}_{t,s}^i$ with respect to s by

$$\frac{\partial}{\partial s} \mathbf{P}_{t,s}^i = \Theta_t^i \mathbf{P}_{t,s} \Lambda_s \Phi^i, \tag{2.34}$$

with

$$\mathbf{P}_{t,t}^i = \Theta_t^i \mathbf{P}_{t,t} \Phi^i = \mathbf{I}_{|\mathcal{K}| \times |\mathcal{K}|}.$$

If (2.33) holds true, then together with Equation (2.32) we have

$$\frac{\partial}{\partial s} \mathbf{P}_{t,s}^i = \Theta_t^i \mathbf{P}_{t,s} \Lambda_s \Phi^i = \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i \Lambda_s \Phi^i = \frac{\partial}{\partial s} \widehat{\mathbf{P}}_{t,s}^i,$$

and

$$\mathbf{P}_{t,t}^i = \mathbf{I}_{|\mathcal{K}| \times |\mathcal{K}|} = \widehat{\mathbf{P}}_{t,t}^i.$$

Thus, for any $0 \leq t \leq s$ the first order of $\mathbf{P}_{t,s}^i$ and $\widehat{\mathbf{P}}_{t,s}^i$ with respect to s are the same, and $\mathbf{P}_{t,s}^i$ and $\widehat{\mathbf{P}}_{t,s}^i$ have the same initial condition. We conclude that $\mathbf{P}_{t,s}^i = \widehat{\mathbf{P}}_{t,s}^i$.

Next, we prove (ii). Recall that we have for any $t \geq 0$, $\Theta_t^i \Phi^i = \mathbf{I}$, where $\mathbf{I}$ is the identity kernel on $\mathcal{K}$. Let $0 \leq u \leq t \leq s$.

$$\begin{aligned}
\mathbf{P}_{u,t}^i \mathbf{P}_{t,s}^i &= \left(\Theta_u^i \mathbf{P}_{u,t} \Phi^i\right) \left(\Theta_t^i \mathbf{P}_{t,s} \Phi^i\right) \\
&= \widehat{\mathbf{P}}_{u,t}^i \Theta_t^i \Phi^i \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i \Phi^i \\
&= \widehat{\mathbf{P}}_{u,t}^i \mathbf{I}_{|\mathcal{K}| \times |\mathcal{K}|} \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i \Phi^i \\
&= \widehat{\mathbf{P}}_{u,t}^i \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i \Phi^i \\
&= \widehat{\mathbf{P}}_{u,s}^i \Theta_s^i \Phi^i \\
&= \Theta_u^i \mathbf{P}_{u,s} \Phi^i \\
&=: \mathbf{P}_{u,s}^i,
\end{aligned}$$

where the second equality follows from assumption (ii); the third equality is from assumption (i); the fifth equality is because $\widehat{\mathbf{P}}_{t,s}^i$ is a semigroup; the sixth equality comes from assumption (ii). We conclude that $\mathbf{P}_{t,s}^i$ is a semigroup. □

Remark 2.5.5. *In fact, we have* (2.33) *if and only if* $\mathbf{P}_{t,s}^i = \widehat{\mathbf{P}}_{t,s}^i$.

In the next theorem, we provide sufficient conditions to condition (Ci-II).

Theorem 2.5.2. *Assume that X is generated by infinitesimal generator matrix function $(\Lambda_u, u \geq 0)$ with the corresponding semigroup $(\mathbf{P}_{t,s}, 0 \leq t \leq s)$. If there exists an initial distribution μ^X such that the semigroup $\widehat{\mathbf{P}}_{t,s}^i$ of $(\Theta_u^i \Lambda_u \Phi^i, u \geq 0)$ satisfies the identity*

$$\Theta_t^i \mathbf{P}_{t,s} = \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i, \quad 0 \leq t \leq s, \tag{2.35}$$

then condition (Ci-II) holds.

Proof. By Proposition 2.5.1, $(\Theta_u^i \Lambda_u \Phi^i, u \geq 0)$ is a valid generator, thus $\widehat{\mathbf{P}}_{t,s}^i$ is a semigroup. In view of Lemma 2.5.5, we know that

$$\mathbf{P}_{t,s}^i \left(x^i, \hat{x}^i\right) = \widehat{\mathbf{P}}_{t,s}^i \left(x^i, \hat{x}^i\right), \quad x^i, \hat{x}^i \in \mathcal{K},$$

and $\mathbf{P}^i_{t,s}$ is a semigroup.

Recall the element of $\mathcal{K}^m$ is denoted by $x = (x^1, \dots x^i, \dots x^m)$. Note that for any $n \in \mathbb{N}$, $0 = t_0 \le t_1 \le t_2 \le \dots \le t_n < \infty$, and $x^i \in \mathcal{K}$, the finite-dimensional distribution of X^i is given by

$$\begin{aligned}
&\mathbb{P}\left(X^i_{t_1} = x^i_1,\ X^i_{t_2} = x^i_2, \dots, X^i_{t_n} = x^i_n\right) \\
&= \sum_{x_n \in \mathcal{H}(x^i_n)} \sum_{x_{n-1} \in \mathcal{H}\left(x^i_{n-1}\right)} \cdots \sum_{x_1 \in \mathcal{H}\left(x^i_1\right)} \mathbb{P}\left(X_{t_1} = x_1,\ X_{t_2} = x_2, \dots, X_{t_n} = x_n\right) \\
&= \sum_{x_n \in \mathcal{H}(x^i_n)} \sum_{x_{n-1} \in \mathcal{H}\left(x^i_{n-1}\right)} \cdots \sum_{x_0 \in \mathcal{H}\left(x^i_0\right)} \mathbb{P}\left(X_{t_0} = x_0\right) \mathbf{P}^{x_0,x_1}_{t_0,t_1} \mathbf{P}^{x_1,x_2}_{t_1,t_2} \cdots \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n}, \qquad (2.36)
\end{aligned}$$

and

$$\sum_{x_n \in \mathcal{H}(x^i_n)} \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n} = \mathbf{P}_{t,s} \Phi^i \mathbb{1}_{\{x^i_n\}}\left(x_{n-1}\right).$$

In the following, we verify the finite-dimensional distribution of X^i given by (2.36) is identical to the finite-dimensional distribution of Y^i. Let $f_n = \mathbb{1}_{\{x^i_n\}},\ x^i_n \in \mathcal{K}$, and $g = \mathbb{1}_{\{x_n\}},\ x_n \in \mathcal{K}^m$. We take the entrywise product $(\Phi^i f_n \cdot g)$. The identity $\Theta^i_{t_n} \Phi^i = \mathbf{I}$ implies that

$$\begin{aligned}
\Theta^i_{t_n}\left(\Phi^i f_n \cdot g\right) &= \Theta^i_{t_n}\left(\Phi^i \mathbb{1}_{\{x^i_n\}} \cdot \mathbb{1}_{\{x_n\}}\right) \\
&= \mathbb{E}\left(\Phi^i \mathbb{1}_{\{x^i_n\}}\left(X_{t_n}\right) \cdot \mathbb{1}_{\{x_n\}}\left(X_{t_n}\right) \mid X^i_{t_n} = x^i_n\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\{x^i_n\}}\left(X^i_{t_n}\right) \cdot \mathbb{1}_{\{x_n\}}\left(X_{t_n}\right) \mid X^i_{t_n} = x^i_n\right) \\
&= \mathbb{1}_{\{x^i_n\}} \mathbb{E}\left(\mathbb{1}_{\{x_n\}}\left(X_{t_n}\right) \mid X^i_{t_n} = x^i_n\right) \\
&= \left(f_n \cdot \Theta^i_{t_n} g\right).
\end{aligned}$$

Then we have

$$\Theta^i_{t_n}\left(\Phi^i f_n \cdot g\right) = \left(f_n \cdot \Theta^i_{t_n} g\right). \qquad (2.37)$$

We project (2.37) on the set $\{Y^i_{t_{n-1}} = x^i_{n-1}\}$,

$$\mathbb{E}\left(\mathbb{E}\left(\Phi^i \mathbb{1}_{\{x^i_n\}}\left(X_{t_n}\right) \cdot \mathbb{1}_{\{x_n\}}\left(X_{t_n}\right) \mid X^i_{t_n} = x^i_n\right) \mid Y^i_{t_{n-1}} = x^i_{n-1}\right)$$

$$= \mathbb{E}\left(\mathbb{1}_{\{x_n^i\}}\left(Y_{t_n}^i\right) \cdot \mathbb{E}\left(\mathbb{1}_{\{x_n\}}\left(X_{t_n}\right) \mid X_{t_n}^i = x_n^i\right) \mid Y_{t_{n-1}}^i = x_{n-1}^i\right),$$

which is equivalent to operate $\widehat{\mathbf{P}}_{t_{n-1},t_n}^i$ to (2.37),

$$\widehat{\mathbf{P}}_{t_{n-1},t_n}^i \Theta_{t_n}^i \left(\Phi^i f_n \cdot g\right) = \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i g\right).$$

Identity (2.35) implies that

$$\Theta_{t_{n-1}}^i \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot g\right) = \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i g\right). \tag{2.38}$$

We do entrywise product of f_{n-1} to both sides of (2.38),

$$\left(f_{n-1} \cdot \Theta_{t_{n-1}}^i \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot g\right)\right) = \left(f_{n-1} \cdot \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i g\right)\right).$$

By the same operation in (2.37), we get

$$\left(f_{n-1} \cdot \Theta_{t_{n-1}}^i \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot g\right)\right) = \Theta_{t_{n-1}}^i \left(\Phi^i f_{n-1} \cdot \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot g\right)\right),$$

It follows that

$$\Theta_{t_{n-1}}^i \left(\Phi^i f_{n-1} \cdot \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot g\right)\right) = \left(f_{n-1} \cdot \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i g\right)\right). \tag{2.39}$$

Next, we do backward induction to (2.39). We operate $\widehat{\mathbf{P}}_{t_{n-2},t_{n-1}}^i$ to (2.39),

$$\widehat{\mathbf{P}}_{t_{n-2},t_{n-1}}^i \Theta_{t_{n-1}}^i \left(\Phi^i f_{n-1} \cdot \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot g\right)\right) = \widehat{\mathbf{P}}_{t_{n-2},t_{n-1}}^i \left(f_{n-1} \cdot \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i g\right)\right).$$

Identity (2.35) implies that

$$\Theta_{t_{n-2}}^i \mathbf{P}_{t_{n-2},t_{n-1}} \left(\Phi^i f_{n-1} \cdot \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot g\right)\right) = \widehat{\mathbf{P}}_{t_{n-2},t_{n-1}}^i \left(f_{n-1} \cdot \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i g\right)\right).$$

Backward induction to time t_0, we have

$$\begin{aligned}
&\Theta_{t_0}^i \mathbf{P}_{t_0,t_1} \left(\Phi^i f_1 \cdot \mathbf{P}_{t_1,t_2} \left(\Phi^i f_2 \cdot \cdots \cdot \mathbf{P}_{t_{n-2},t_{n-1}} \left(\Phi^i f_{n-1} \cdot \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot g\right)\right) \cdots\right)\right) \\
&= \widehat{\mathbf{P}}_{t_0,t_1}^i \left(f_1 \cdot \widehat{\mathbf{P}}_{t_1,t_2}^i \left(f_2 \cdot \cdots \cdot \widehat{\mathbf{P}}_{t_{n-2},t_{n-1}}^i \left(f_{n-1} \cdot \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i g\right)\right) \cdots\right)\right).
\end{aligned} \tag{2.40}$$

If there exists an initial distribution $\mu^X > 0$, we can obtain the initial distribution of X^i by

$$\Upsilon^i \mu^X (\Gamma) = \mathbb{P}\left(X_0^i \in \Gamma^i\right), \quad \Gamma \subset \mathcal{K}^m,\ \Gamma^i \subset \mathcal{K}.$$

Now, we do entrywise product $f_0 \cdot \Upsilon^i \mu^X$ to (2.40) to make the conditional probability unconditional,

$$\begin{aligned}
&\left(f_0 \cdot \left(\Upsilon^i \mu^X \cdot \Theta_{t_0}^i \mathbf{P}_{t_0,t_1} \left(\Phi^i f_1 \cdot \mathbf{P}_{t_1,t_2} \left(\Phi^i f_2 \cdot \cdots \cdot \mathbf{P}_{t_{n-2},t_{n-1}} \left(\Phi^i f_{n-1} \cdot \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot g\right)\right) \cdots\right)\right)\right)\right) \\
&= \left(f_0 \cdot \left(\Upsilon^i \mu^X \cdot \widehat{\mathbf{P}}_{t_0,t_1}^i \left(f_1 \cdot \widehat{\mathbf{P}}_{t_1,t_2}^i \left(f_2 \cdot \cdots \cdot \widehat{\mathbf{P}}_{t_{n-2},t_{n-1}}^i \left(f_{n-1} \cdot \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i g\right)\right) \cdots\right)\right)\right)\right).
\end{aligned} \tag{2.41}$$

With $g = \mathbb{1}_{\{x_n\}}$ and $f_j = \mathbb{1}_{\{x_j^i\}},\ j = 0, 1, \ldots, n$, we have

$$f_n \cdot \Theta_{t_n}^i g = \mathbb{1}_{\{x_n^i\}} \cdot \Theta_{t_n}^i \mathbb{1}_{x_n}\left(x_n^i\right) = \mathbb{1}_{\{x_n^i\}} \cdot \mathbb{E}\left(\mathbb{1}_{\{x_n\}}\left(X_{t_n}\right) \mid X_{t_n}^i = x_n^i\right),$$

and

$$\begin{aligned}
\widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i g\right) &= \mathbb{E}\left(\mathbb{1}_{\{x_n^i\}}\left(Y_{t_n}^i\right) \cdot \mathbb{E}\left(\mathbb{1}_{\{x_n\}}\left(X_{t_n}\right) \mid X_{t_n}^i = x_n^i\right) \mid Y_{t_{n-1}}^i = x_{n-1}^i\right) \\
&= \mathbb{E}\left(\mathbb{1}_{\{x_n\}}\left(X_{t_n}\right) \mid X_{t_n}^i = x_n^i\right) \mathbb{E}\left(\mathbb{1}_{\{x_n^i\}}\left(Y_{t_n}^i\right) \mid Y_{t_{n-1}}^i = x_{n-1}^i\right) \\
&= \left(\Theta_{t_n}^i \mathbb{1}_{\{x_n\}}\left(x_n^i\right)\right) \mathbb{E}\left(\mathbb{1}_{\{x_n^i\}}\left(Y_{t_n}^i\right) \mid Y_{t_{n-1}}^i = x_{n-1}^i\right),
\end{aligned}$$

where the second equality is by the independence of conditional expectation. Then (2.41) is equivalent to

$$\begin{aligned}
&\mathbb{P}\left(X_{t_0}^i = x_0^i,\ X_{t_1}^i = x_1^i, \ldots, X_{t_{n-1}}^i = x_{n-1}^i,\ X_{t_n} = x_n\right) \\
&\qquad = \mathbb{P}\left(Y_{t_0}^i = x_0^i,\ Y_{t_1}^i = x_1^i, \ldots, Y_{t_{n-1}}^i = x_{n-1}^i,\ Y_{t_n}^i = x_n^i\right) \Theta_{t_n}^i \mathbb{1}_{\{x_n\}}\left(x_n^i\right).
\end{aligned} \tag{2.42}$$

In particular, (2.41) remains true if we replace g by $\Phi^i f_n$, then we get

$$\begin{aligned}
&\left(f_0 \cdot \left(\Upsilon^i \mu^X \cdot \Theta_{t_0}^i \mathbf{P}_{t_0,t_1} \left(\Phi^i f_1 \cdot \mathbf{P}_{t_1,t_2} \left(\Phi^i f_2 \cdot \cdots \cdot \mathbf{P}_{t_{n-2},t_{n-1}} \left(\Phi^i f_{n-1} \cdot \mathbf{P}_{t_{n-1},t_n} \left(\Phi^i f_n \cdot \Phi^i f_n\right)\right) \cdots\right)\right)\right)\right) \\
&= \left(f_0 \cdot \left(\Upsilon^i \mu^X \cdot \widehat{\mathbf{P}}_{t_0,t_1}^i \left(f_1 \cdot \widehat{\mathbf{P}}_{t_1,t_2}^i \left(f_2 \cdot \cdots \cdot \widehat{\mathbf{P}}_{t_{n-2},t_{n-1}}^i \left(f_{n-1} \cdot \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot \Theta_{t_n}^i \Phi^i f_n\right)\right) \cdots\right)\right)\right)\right) \\
&= \left(f_0 \cdot \left(\Upsilon^i \mu^X \cdot \widehat{\mathbf{P}}_{t_0,t_1}^i \left(f_1 \cdot \widehat{\mathbf{P}}_{t_1,t_2}^i \left(f_2 \cdot \cdots \cdot \widehat{\mathbf{P}}_{t_{n-2},t_{n-1}}^i \left(f_{n-1} \cdot \widehat{\mathbf{P}}_{t_{n-1},t_n}^i \left(f_n \cdot f_n\right)\right) \cdots\right)\right)\right)\right),
\end{aligned} \tag{2.43}$$

where the second equality comes from $\Theta_{t_n}^i \Phi^i = \mathbf{I}$. Then (2.43) is equivalent to

$$\mathbb{P}\left(X^i_{t_0} = x^i_0,\ X^i_{t_1} = x^i_1, \ldots, X^i_{t_{n-1}} = x^i_{n-1},\ X^i_{t_n} = x^i_n\right)$$
$$= \mathbb{P}\left(Y^i_{t_0} = x^i_0,\ Y^i_{t_1} = x^i_1, \ldots, Y^i_{t_{n-1}} = x^i_{n-1},\ Y^i_{t_n} = x^i_n\right). \quad (2.44)$$

In view of (2.42) and (2.44), we get

$$\mathbb{P}\left(X_{t_n} = x_n \mid X^i_{t_0} = x^i_0,\ X^i_{t_1} = x^i_1, \ldots, X^i_{t_{n-1}} = x^i_{n-1},\ X^i_{t_n} = x^i_n\right)$$
$$= \Theta^i_{t_n} \mathbb{1}_{\{x_n\}}\left(x^i_n\right) = \mathbb{P}\left(X_{t_n} = x_n \mid X^i_{t_n} = x^i_n\right).$$

By Lemma 2.5.3, we conclude that condition (Ci-II) holds. □

Remark 2.5.6. *In view of* (2.44)*, Lemma 2.5.1, and Lemma 2.5.2, we show that* X^i *and* Y^i *have identical finite-dimensional distributions. Since* Y^i *is a Markov chain generated by* $(\Theta^i_u \Lambda_u \Phi^i, u \geq 0)$*, it is enough to conclude that* X^i *is a Markov chain generated by* $(\Theta^i_u \Lambda_u \Phi^i, u \geq 0)$ *with initial distribution* $\Upsilon^i \mu^X$*.*

Remark 2.5.7. *Theorem 2.5.2 extends [RP81, Theorem 2] to the class of time-inhomogeneous multivariate Markov chains with coordinate projections.*

2.5.3 Necessary and Sufficient Conditions for Strong Markovian Consistency. We want to emphasize that strong Markovian consistency implies weak Markovian consistency. Therefore, any sufficient condition for strong Markovian consistency remains a sufficient condition for weak Markovian consistency.

In this section, we will show that Proposition 2.5.2 on page 44 is a parallel result to [BJVV08][Proposition 5.1], which states that if the condition (M) holds, then X is strongly Markovian consistent, i.e., all components of X are Markovian in the filtration $\mathbb{F}^X$. Whereas, condition (M) is represented in terms of the infinitesimal generator matrix function Λ_t of X.

Proposition 2.5.5. *Assume that* X *is generated by infinitesimal generator matrix function* $(\Lambda_u, u \geq 0)$*, and* $(\mathbf{P}_{t,s}, 0 \leq t \leq s)$ *is the corresponding transition prob-*

ability matrix function. Then, condition (Pi) is equivalent to condition (Mi), $i \in \{1, 2, \ldots, m\}$.

Proof. Since $\mathbf{P}_{t,s}$ is the transition probability matrix function of X, the Kolmogorov forward equation holds for every $0 \leq t \leq s < \infty$,

$$\frac{\partial}{\partial s}\mathbf{P}_{t,s} = \mathbf{P}_{t,s}\Lambda_s, \quad \mathbf{P}_{t,t} = \mathbf{I}_{|\mathcal{K}^m|\times|\mathcal{K}^m|}.$$

Let us fix $i \in \{1, 2, \ldots, m\}$. First, we prove that condition (Mi) implies condition (Pi). Assume that for any $u \geq 0$, every $x^i \in \mathcal{K}$, every $y^i \in \mathcal{K}$, all $x, \widetilde{x} \in \mathcal{K}^m$, $x \neq \widetilde{x}$, we have $\Lambda_u^{x,y^i} = \Lambda_u^{\widetilde{x},y^i}$, where $x = (x^1, \ldots, x^i, x^{i+1}, \ldots, x^m)$ and $\widetilde{x} = (\widetilde{x}^1, \ldots, x^i, \widetilde{x}^{i+1}, \ldots, \widetilde{x}^m)$. For any $x, y \in \mathcal{K}^m$, the first derivative $\frac{\partial}{\partial s}\mathbf{P}_{t,s}$ is given by

$$\frac{\partial}{\partial s}\mathbf{P}_{t,s}^{x,y} = \sum_{u\in\mathcal{K}^m} \mathbf{P}_{t,s}^{x,u}\Lambda_s^{u,y}, \tag{2.45}$$

where $u = (u^1, \ldots u^i, \ldots u^m)$. We fix $y^i \in \mathcal{K}$ and take $y = (y^1, \ldots, y^i, y^{i+1}, \ldots, y^m)$. By summing over $y^j \in \mathcal{K},\ j \neq i$, in both sides of (2.45), we know

$$\sum_{y\in\mathcal{H}(y^i)} \frac{\partial}{\partial s}\mathbf{P}_{t,s}^{x,y} = \sum_{y\in\mathcal{H}(y^i)} \sum_{u\in\mathcal{K}^m} \mathbf{P}_{t,s}^{x,u}\Lambda_s^{u,y}. \tag{2.46}$$

In particular, we can take $u = (u^1, \ldots x^i, \ldots u^m)$, then (2.46) is equivalent to

$$\begin{aligned}
\frac{\partial}{\partial s}\mathbf{P}_{t,s}^{x,y^i} &= \sum_{u\in\mathcal{K}^m} \mathbf{P}_{t,s}^{x,u}\Lambda_s^{u,y^i} \\
&= \sum_{u^1\in\mathcal{K}} \cdots \sum_{x^i\in\mathcal{K}} \cdots \sum_{u^m\in\mathcal{K}} \mathbf{P}_{t,s}^{x,\left(u^1,\ldots,x^i,\ldots,u^m\right)}\Lambda_s^{\left(u^1,\ldots,x^i,\ldots,u^m\right),y^i} \\
&= \sum_{x^i\in\mathcal{K}} \left(\Lambda_s^{\hat{v},y^i} \sum_{u\in\mathcal{H}(x^i)} \mathbf{P}_{t,s}^{x,u}\right) \\
&= \sum_{x^i\in\mathcal{K}} \Lambda_s^{\hat{v},y^i}\mathbf{P}_{t,s}^{x,u^i} \\
&= \sum_{x^i\in\mathcal{K}} \Lambda_s^{x,y^i}\mathbf{P}_{t,s}^{x,u^i},
\end{aligned} \tag{2.47}$$

where $\hat{v} = (\hat{v}^1, \dots x^i, \dots \hat{v}^m)$. Note that the third equality follows from the assumption, $\Lambda_s^{u,y^i} = \Lambda_s^{\hat{v},y^i}$ for all $u, \hat{v} \in \mathcal{K}^m,\ u \neq \hat{v}$, so that we can move $\Lambda_s^{\hat{v},y^i}$ outside of the summation over $u \in \mathcal{H}(x^i)$. For the last equality, it remains true for $\Lambda_s^{\hat{v},y^i} = \Lambda_s^{x,y^i}$.

Since the transition probability matrix function $\mathbf{P}_{t,s}$ is a kernel, the sum of each row is 1, i.e., $\sum_{y^i \in \mathcal{K}} \mathbf{P}_{t,s}^{x,y^i} = 1$. For every $x \in \mathcal{K}^m$, we have a coupled system of $(|\mathcal{K}| - 1)$ differential equations with $(|\mathcal{K}| - 1)$ dependent variables $\mathbf{P}_{t,s}^{x,y^i},\ y^i \in \mathcal{K}$, and one independent variable s.

Analogically, for every $y^i \in \mathcal{K}$, all $x, \widetilde{x} \in \mathcal{K}^m$, $x \neq \widetilde{x}$, we have

$$\begin{aligned}
\frac{\partial}{\partial s}\mathbf{P}_{t,s}^{\widetilde{x},y^i} &= \sum_{u \in \mathcal{K}^m} \mathbf{P}_{t,s}^{\widetilde{x},u} \Lambda_s^{u,y^i} \\
&= \sum_{u^1 \in \mathcal{K}} \cdots \sum_{x^i \in \mathcal{K}} \cdots \sum_{u^m \in \mathcal{K}} \mathbf{P}_{t,s}^{\widetilde{x},\left(u^1,\dots,x^i,\dots,u^m\right)} \Lambda_s^{\left(u^1,\dots,x^i,\dots,u^m\right),y^i} \\
&= \sum_{x^i \in \mathcal{K}} \left(\Lambda_s^{\hat{v},y^i} \sum_{u \in \mathcal{H}(x^i)} \mathbf{P}_{t,s}^{\widetilde{x},u} \right), \\
&= \sum_{x^i \in \mathcal{K}} \Lambda_s^{\hat{v},y^i} \mathbf{P}_{t,s}^{\widetilde{x},u^i} \\
&= \sum_{x^i \in \mathcal{K}} \Lambda_s^{x,y^i} \mathbf{P}_{t,s}^{\widetilde{x},u^i}. \qquad (2.48)
\end{aligned}$$

Observe that the coupled systems (2.47) and (2.48) have the same coefficients Λ_s^{x,y^i}, $x \in \mathcal{K}^m, y^i \in \mathcal{K}$, in each corresponding differential equation, respectively. Moreover, the initial values are identical in the corresponding differential equations. Therefore, these two coupled systems (2.47) and (2.48) have the same solution,

$$\mathbf{P}_{t,s}^{x,y^i} = \mathbf{P}_{t,s}^{\widetilde{x},y^i}, \quad u^i \in \mathcal{K}.$$

Since $x^i \in \mathcal{K}$ in arbitrary, we have for all $x^i \in \mathcal{K}$, every $y^i \in \mathcal{K}$, all $x, \widetilde{x} \in \mathcal{K}^m$, $x \neq \widetilde{x}$, the identity holds

$$\mathbf{P}_{t,s}^{x,y^i} = \mathbf{P}_{t,s}^{\widetilde{x},y^i}, \quad 0 \leq t \leq s.$$

Next, we show that condition (Pi) implies condition (Mi). Assume that for

all $x^i \in \mathcal{K}$, every $y^i \in \mathcal{K}$, all $x, \widetilde{x} \in \mathcal{K}^m$, $x \neq \widetilde{x}$, we have $\mathbf{P}_{t,s}^{x,y^i} = \mathbf{P}_{t,s}^{\widetilde{x},y^i}$, where $x = (x^1, \ldots, x^i, x^{i+1}, \ldots, x^m)$ and $\widetilde{x} = (\widetilde{x}^1, \ldots, x^i, \widetilde{x}^{i+1}, \ldots, \widetilde{x}^m)$. Since the infinitesimal generator is uniquely determined by

$$\Lambda_s = \lim_{h \downarrow 0} \frac{\mathbf{P}_{t,s+h} - \mathbf{I}_{|\mathcal{K}^m| \times |\mathcal{K}^m|}}{h}, \quad 0 \leq t \leq s. \tag{2.49}$$

We multiply the matrix Φ^i from the right to both sides of (2.49),

$$\Lambda_s \Phi^i = \lim_{h \downarrow 0} \frac{\mathbf{P}_{t,s+h} \Phi^i - \Phi^i}{h}, \quad 0 \leq t \leq s. \tag{2.50}$$

In view of Remark 2.5.2 on page 40, for any $y^i \in \mathcal{K}$ and $z \in \mathcal{K}^m$, the entry of $\mathbf{P}_{t,s}\Phi^i$ is exactly given by

$$\mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{y^i\}}(z) = \sum_{y \in \mathcal{H}(y^i)} \mathbb{P}(X_s = y \mid X_t = z) = \mathbf{P}_{t,s}^{z,y^i}. \tag{2.51}$$

In particular, (2.51) remains true for $z = x$,

$$\mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{y^i\}}(x) = \mathbf{P}_{t,s}^{x,y^i},$$

and $z = \widetilde{x}$, $\widetilde{x} \neq x$,

$$\mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{y^i\}}(\widetilde{x}) = \mathbf{P}_{t,s}^{\widetilde{x},y^i}.$$

By assumption $\mathbf{P}_{t,s}^{x,y^i} = \mathbf{P}_{t,s}^{\widetilde{x},y^i}$, we know for any $0 \leq t \leq s$, every $x^i, y^i \in \mathcal{K}$, and all $x, \widetilde{x} \in \mathcal{K}^m$, $x \neq \widetilde{x}$,

$$\mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{y^i\}}(x) = \mathbf{P}_{t,s}\Phi^i \mathbb{1}_{\{y^i\}}(\widetilde{x}). \tag{2.52}$$

In view of Equations (2.50) and (2.52), for any $0 \leq t \leq s$, every $x^i, y^i \in \mathcal{K}$, all $x, \widetilde{x} \in \mathcal{K}^m$, $x \neq \widetilde{x}$,

$$\begin{aligned}
\Lambda_s^{x,y^i} = \Lambda_s \Phi^i \mathbb{1}_{\{y^i\}}(x) &= \lim_{h \downarrow 0} \frac{\mathbf{P}_{t,s+h}\Phi^i \mathbb{1}_{\{y^i\}}(x) - \Phi^i \mathbb{1}_{\{y^i\}}(x)}{h} \\
&= \lim_{h \downarrow 0} \frac{\mathbf{P}_{t,s+h}\Phi^i \mathbb{1}_{\{y^i\}}(\widetilde{x}) - \Phi^i \mathbb{1}_{\{y^i\}}(\widetilde{x})}{h} \\
&= \Lambda_s \Phi^i \mathbb{1}_{\{y^i\}}(\widetilde{x}) \\
&= \Lambda_s^{\widetilde{x},y^i}.
\end{aligned}$$

This concludes the proof. □

If condition (M) holds, then X is strongly Markovian consistent (cf. [BJVV08, Proposition 5.1]). Next, we present the strong Markovian consistency with respect to the component X^i in terms of condition (Pi).

Corollary 2.5.2. *If condition (Pi) holds, then the component X^i is a Markov chain in the filtration of X.*

Proof. Suppose that condition (Pi) holds. It follows from Proposition 2.5.5, condition (Mi) holds true. In view of [BJVV08, Proposition 5.1], it states if condition (Mi) holds, then the component X^i is Markovian in the filtration of X. The proof is complete. □

In the next proposition we show that the condition in Theorem 2.5.2 is not stronger than condition (Pi).

Proposition 2.5.6. *Assume that X is generated by infinitesimal generator $(\Lambda_u, u \geq 0)$ and the corresponding semigroup is $(\mathbf{P}_{t,s}, 0 \leq t \leq s)$. Condition (P$i$) (or, equivalently, condition (Mi)) implies*

$$\Theta_t^i \mathbf{P}_{t,s} = \mathbf{P}_{t,s}^i \Theta_s^i, \quad 0 \leq t \leq s, \tag{2.53}$$

where $\mathbf{P}_{t,s}^i$ is the transition function of X^i.

Proof. Assume that condition (Pi) holds true. Then X^i is a Markov chain generated by infinitesimal generator $(\Theta_u^i \Lambda_u \Phi^i, u \geq 0)$. We denote by $\mathbf{P}_{t,s}^i$ the transition matrix function of X^i. Since X and X^i are Markov chains, $\mathbf{P}_{t,s}$ and $\mathbf{P}_{t,s}^i$ satisfy the Kolmogorov forward equation,

$$\frac{\partial}{\partial s} \mathbf{P}_{t,s} = \mathbf{P}_{t,s} \Lambda_s, \quad \mathbf{P}_{t,t} = \mathbf{I}_{|\mathcal{K}^m| \times |\mathcal{K}^m|},$$

and

$$\frac{\partial}{\partial s} \mathbf{P}_{t,s}^i = \mathbf{P}_{t,s}^i \Theta_s^i \Lambda_s \Phi^i, \quad \mathbf{P}_{t,t}^i = \mathbf{I}_{|\mathcal{K}| \times |\mathcal{K}|}, \tag{2.54}$$

In view of Lemma 2.5.4, we know $\mathbf{P}^i_{t,s} = \Theta^i_t \mathbf{P}_{t,s} \Phi^i$. The first derivative of $\mathbf{P}^i_{t,s}$ with respect to s is given by

$$\frac{\partial}{\partial s}\mathbf{P}^i_{t,s} = \Theta^i_t \frac{\partial}{\partial s}\mathbf{P}_{t,s}\Phi^i = \Theta^i_t \mathbf{P}_{t,s}\Lambda_s \Phi^i. \tag{2.55}$$

Comparing Equations (2.54) and (2.55), we conclude that for $0 \le t \le s$,

$$\Theta^i_t \mathbf{P}_{t,s} = \mathbf{P}^i_{t,s}\Theta^i_s.$$

□

Remark 2.5.8. *Note that in Proposition 2.5.4, the algebraic structure of the transition probability matrix function* $(\mathbf{P}_{t,s}, 0 \le t \le s)$ *of* X *is not explicitly stated. In Section 4.2 on page 116, we will define the independence of the components in terms of infinitesimal generator of* X*. Then the independence of the components also implies condition (P).*

In [BJVV08, Proposition 5.1], it was showed that condition (M), which is equivalent to condition (P) (cf. Proposition 2.5.5), implies strong Markovian consistency. In the next proposition, we give conditions under which condition (M) is necessary for strong Markovian consistency.

Theorem 2.5.3. *Assume that*

$$\mathbb{P}\left(X_t = z\right) > 0, \ dt\text{-a. e.}, \ z \in \mathcal{K}^m. \tag{2.56}$$

Then strong Markovian consistency implies condition (M).

Proof. As a result of [BJN13, Theorem 1.8], the component X^i of X is Markovian in the filtration $\mathbb{F}^X$ if and only if

$$\mathbb{1}_{\{X^i_t = x^i\}} \sum_{y \in \mathcal{H}(y^i)} \lambda^{x,y}_t = \mathbb{1}_{\{X^i_t = x^i\}} \lambda^{i;x^i,y^i}_t, \quad dt \otimes d\mathbb{P}\text{-a. e.}, \ x^i, y^i \in E_i, \ x^i \neq y^i, \tag{2.57}$$

where $x = \left(X_t^1, \dots x^i, X_t^{i+1}, \dots X_t^m\right)$ and $y = (y^1, \dots y^i, y^{i+1}, \dots y^m)$ for some locally integrable function $\lambda_t^{i,x^i y^i}$.

From (2.57), we have

$$\begin{aligned}
\mathbb{1}_{\{X_t^i = x^i\}} \lambda_t^{i;x^i,y^i} &= \mathbb{1}_{\{X_t^i = x^i\}} \sum_{y \in \mathcal{H}(y^i)} \lambda_t^{x,y} \\
&= \mathbb{1}_{\{X_t^i = x^i,\ \bigcap_{j=1,\dots,m, j \neq i} \{X_t^j \in \mathcal{K}\}\}} \sum_{y \in \mathcal{H}(y^i)} \lambda_t^{x,y} \\
&= \mathbb{1}_{\{X_t^i = x^i,\ \bigcap_{j=1,\dots,m, j \neq i} \bigcup_{x^j \in \mathcal{K}} \{X_t^j = x^j\}\}} \sum_{y \in \mathcal{H}(y^i)} \lambda_t^{z,y} \\
&= \sum_{z \in \mathcal{H}(x^i)} \mathbb{1}_{\{X_t^i = x^i,\ \bigcap_{j=1,\dots,m, j \neq i} \{X_t^j = x^j\}\}} \sum_{y \in \mathcal{H}(y^i)} \lambda_t^{z,y} \\
&= \sum_{z \in \mathcal{H}(x^i)} \mathbb{1}_{\{X_t^1 = x^1, \dots,\ X_t^i = x^i, \dots,\ X_t^m = x^m\}} \sum_{y \in \mathcal{H}(y^i)} \lambda_t^{z,y},
\end{aligned}$$

where $z = (x^1, \dots, x^i, \dots, x^m)$.

Assumption (2.56) implies that for every $\omega \in \{X_t^1 = x^1, \dots, X_t^i = x^i,\ X_t^{i+1} = x^{i+1}, \dots, X_t^m = x^m\} \subseteq \{X_t^i = x^i\}$. Then we have

$$\lambda_t^{i;x^i,y^i} = \sum_{y \in \mathcal{H}(y^i)} \lambda_t^{z,y}, \quad z = \left(x^1, \dots, x^i, \dots, x^m\right),$$

which is exactly condition (Mi). Since i is arbitrary, condition (M) holds. □

Remark 2.5.9. *Note that the assumption* (2.56) *is closely related to the initial distribution* μ^X. *There are two ways in the presence of initial distribution* μ^X. *First, for any* $x_t = (x_t^1, \dots, x_t^i, \dots, x_t^m) \in \mathcal{K}^m$ *, any* $\pi_t \in \Pi_t$, *and any* $\varphi \in \Psi_{|\pi_t|}^{x_t^i}$, *we have*

$$\mathbb{P}\left(X_t = x_t\right) = \sum_{\pi_t \in \Pi_t, \varphi \in \Psi_{|\pi_t|}^{x_t^i}} \mathbb{P}\left(X|_{\pi_t}^{(i)} \bowtie (\varphi, \hat{a})\right),$$

where $\hat{a} = \left(x_t^1, \dots, x_t^{i-1}, x_t^{i+1}, \dots, x_t^m\right)$. *Thus,*

$$\forall x_t \in \mathcal{K}^m,\ \mathbb{P}\left(X_t = x_t\right) > 0$$

$$\iff \forall x_t \in \mathcal{K}^m,\ \exists \psi \in \Psi^{x_t^i}_{|\pi_t|} \text{ such that } \mathbb{P}\left(X|^{(i)}_{\pi_t} \bowtie (\psi, \hat{a})\right) > 0$$

$$\iff \forall x_t \in \mathcal{K}^m,\ \exists \psi \in \Psi^{x_t^i}_{|\pi_t|} \text{ such that } \mathbb{P}\left(X_0^i = \psi_0\right) > 0$$

$$\iff \exists \mu^X > 0.$$

Second, for $x_t \in \mathcal{K}^m$, there exists initial distribution $\mu^X > 0$ such that

$$\mathbb{P}\left(X_t = x_t\right) = \sum_{x_0 \in \mathcal{K}^m} \mathbb{P}\left(X_0 = x_0\right) \mathbf{P}^{x_0, x_t}_{0,t} > 0.$$

2.5.4 Conditions for Weak-Only Markovian Consistency. In view of Theorem 2.5.3 and [BJVV08, Proposition 5.1], if for every $t \geq 0$ all states are accessible, then condition (M) (or, equivalently, condition (P)) is necessary and sufficient for strong Markovian consistency. As mentioned earlier, we are interested in when the component X^i, $i \in \{1, 2, \ldots, m\}$, is Markovian in its own filtration, but not the filtration of X. That is, X satisfies the *weak-only Markovian consistency* property. We close this section with one important theorem which provides criterion for weak-only Markovian consistency.

Theorem 2.5.4. *Let $i \in \{1, 2, \ldots, m\}$. Suppose that X is generated by $(\Lambda_u, u \geq 0)$ with initial distribution μ^X. If the following conditions are satisfied,*

(i) condition (Ci) holds true;

(ii) the inequality holds,

$$\mathbb{P}\left(X_t = x_t\right) > 0,\ dt\text{-a. e.},\ x_t \in \mathcal{K}^m;$$

(iii) condition (Pi) (or, equivalently, condition (Mi)) does not hold,

then X^i is Markovian in its own natural filtration, but not Markovian in the filtration of X.

Proof. Condition (Ci) gives the necessary and sufficient conditions for weak Markovian consistency with respect to the component X^i. According to Theorem 2.5.3, if (ii) holds, then condition (M) is necessary and sufficient for strong Markovian consistency with respect to the component X^i. Therefore, if condition (Mi) (or, equivalently, condition (Pi)) does not hold, then X^i is a Markov chain in its own natural filtration, but not in the filtration $\mathbb{F}^X$. □

We close this section by presenting the relations between the necessary and sufficient conditions for Markovian consistency.

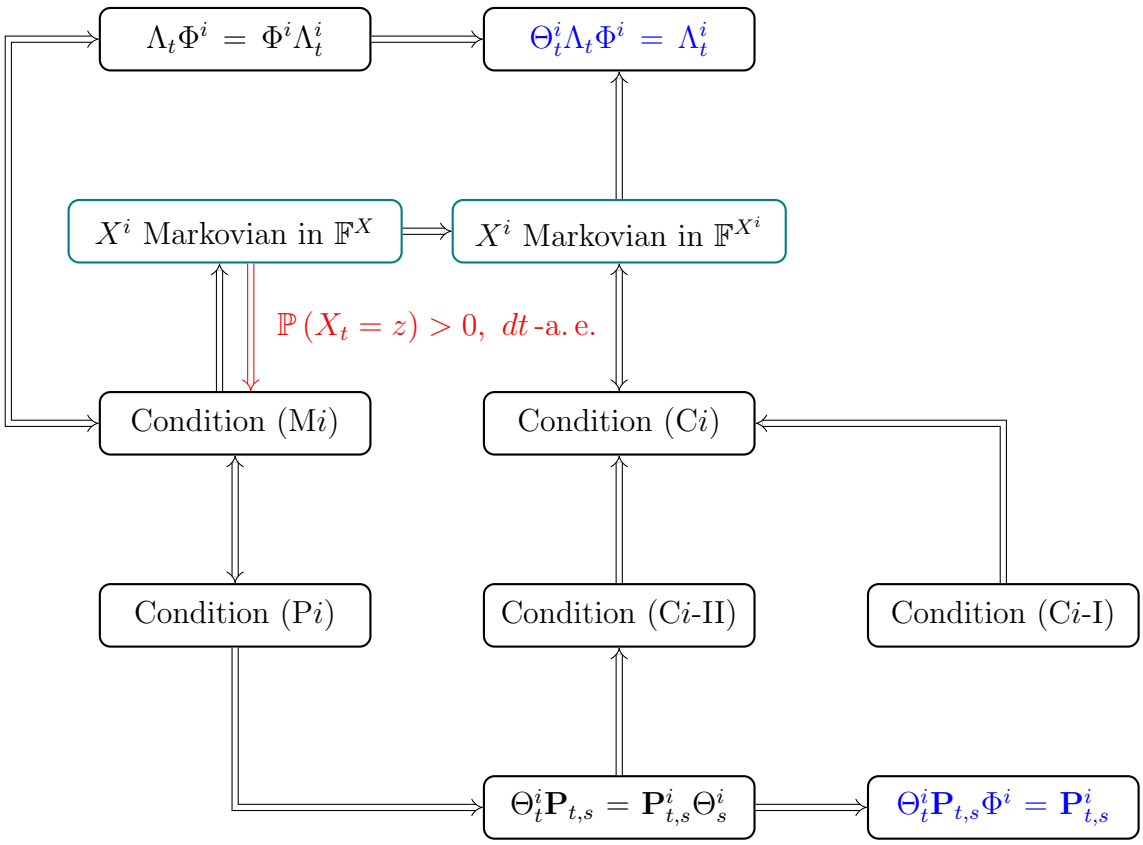

Figure 2.1. The Necessary and Sufficient Conditions for Markovian Consistency

2.6 Examples

We start by giving an example to Theorem 2.5.2 on page 56.

Example 2.6.1. *Consider two processes X^1, X^2 taking values in $\mathcal{K} = \{0,1\}$. Let 1 be the absorbing state. Suppose that $X = (X^1, X^2)$ is a bivariate Markov chain. Assume that $\mathbb{P}(X_0 = (0,0)) = 1$, and the infinitesimal generator of X is given by*

$$\Lambda_u = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a_u + c_u) & a_u & 0 & c_u \\ 0 & -d_u & 0 & d_u \\ 0 & 0 & -f_u & f_u \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \tag{2.58}$$

where for any $u \geq 0$, the parameters $a_u, c_u > 0$, $d_u, f_u \geq 0$, and a_u, c_u, d_u, f_u are locally integrable on $[0,T)$. For any $0 \leq t \leq s$, the transition probability matrix function of X is given by

$$\begin{aligned} \mathbf{P}_{t,s} &= [p_{t,s}(i,j)]_{i,j=1,2,3,4} \\ &= \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \mathrm{e}^{-\int_t^s (a_v + c_v)\,\mathrm{d}v} & p_{t,s}(1,2) & 0 & p_{t,s}(1,4) \\ 0 & \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} & 0 & 1 - \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} \\ 0 & 0 & \mathrm{e}^{-\int_t^s f_v\,\mathrm{d}v} & 1 - \mathrm{e}^{-\int_t^s f_v\,\mathrm{d}v} \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}, \end{aligned}$$

where

$$\begin{aligned} p_{t,s}(1,2) &= \frac{a_s}{a_s + c_s - d_s}\left(\mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} - \mathrm{e}^{-\int_t^s (a_v + c_v)\,\mathrm{d}v}\right) \\ p_{t,s}(1,4) &= 1 - \frac{c_s - d_s}{a_s + c_s - d_s}\mathrm{e}^{-\int_t^s (a_v + c_v)\,\mathrm{d}v} - \frac{a_s}{a_s + c_s - d_s}\mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v}. \end{aligned}$$

For any $t \geq 0$, the one-dimensional probabilities of X are given by

$$\begin{aligned} &\mathbb{P}(X_t = (0,0)) = \sum_{z \in \mathcal{K}^2} \mathbb{P}(X_0 = z)\,\mathbf{P}_{0,t}^{z,(0,0)} = p_{0,t}(1,1) = \mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v} \\ &\mathbb{P}(X_t = (0,1)) = p_{0,t}(1,2) = \frac{a_t}{a_t + c_t - d_t}\left(\mathrm{e}^{-\int_0^t d_v\,\mathrm{d}v} - \mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v}\right) \\ &\mathbb{P}(X_t = (1,0)) = p_{0,t}(1,3) = 0 \end{aligned}$$

$$\mathbb{P}\left(X_t = (1,1)\right) = p_{0,t}(1,4) = 1 - \frac{c_t - d_t}{a_t + c_t - d_t}\mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v} - \frac{a_t}{a_t + c_t - d_t}\mathrm{e}^{-\int_0^t d_v\,\mathrm{d}v},$$

and the one-dimensional probabilities of X^1 *are given by*

$$\begin{aligned}
\mathbb{P}\left(X_t^1 = 0\right) &= \mathbb{P}\left(X_t = (0,0)\right) + \mathbb{P}\left(X_t = (0,1)\right) \\
&= \frac{c_t - d_t}{a_t + c_t - d_t}\mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v} + \frac{a_t}{a_t + c_t - d_t}\mathrm{e}^{-\int_0^t d_v\,\mathrm{d}v} \\
\mathbb{P}\left(X_t^1 = 1\right) &= \mathbb{P}\left(X_t = (1,0)\right) + \mathbb{P}\left(X_t = (1,1)\right) \\
&= 1 - \frac{c_t - d_t}{a_t + c_t - d_t}\mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v} - \frac{a_t}{a_t + c_t - d_t}\mathrm{e}^{-\int_0^t d_v\,\mathrm{d}v}.
\end{aligned}$$

Thus, the operator $\Theta_t^1,\ t \geq 0$, *can be represented by*

$$\begin{aligned}
\Theta_t^1 &= \left[\mathbb{P}\left(X_t = x_t \mid X_t^1 = x_t^1\right)\right]_{x_t \in \mathcal{K}^2, x_t^1 \in \mathcal{K}} \\
&= \begin{array}{c c} & \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} \theta_t & 1-\theta_t & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array},
\end{aligned}$$

where

$$\theta_t = \frac{p_{0,t}(1,1)}{p_{0,t}(1,1) + p_{0,t}(1,2)} = \frac{\mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v}}{\frac{c_t - d_t}{a_t + c_t - d_t}\mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v} + \frac{a_t}{a_t + c_t - d_t}\mathrm{e}^{-\int_0^t d_v\,\mathrm{d}v}}.$$

The extension operator Φ^1 *can be represented by*

$$\Phi^1 = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} & \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \end{array}.$$

Note that $\Theta_t^1 \Phi^1 = \mathbf{I}_{2\times 2},\ t \geq 0$. *Next, the condition in Theorem 2.5.2 holds*

$$\Theta_t^1 \mathbf{P}_{t,s} = \widehat{\mathbf{P}}_{t,s}^1 \Theta_s^1, \quad 0 \leq t \leq s, \tag{2.59}$$

where $\widehat{\mathbf{P}}^1_{t,s} = \Theta^1_t \mathbf{P}_{t,s} \Phi^1$. *The detailed verification of* (2.59) *can be found in Appendix A. In view of Theorem 2.5.2, we conclude that* X^1 *is a Markov chain in its own filtration.*

Now, since $\mathbb{P}(X_t = (1,0)) = 0$, *we cannot apply Theorem 2.5.4 to claim, that the structures of* $(\Lambda_u, u \geq 0)$ *without condition (M) being satisfied are weak-only Markovian consistent with respect to* X^1. *Instead, we verify by the definition of strong Markovian consistency. We compute for any* $0 \leq t \leq s$,

$$\mathbb{P}\left(X^1_s = 1 \mid X_t = (0,0)\right) = p_{t,s}(1,4)$$
$$\mathbb{P}\left(X^1_s = 1 \mid X_t = (0,1)\right) = p_{t,s}(2,4)$$
$$\mathbb{P}\left(X^1_s = 1 \mid X^1_t = 0\right) = \Theta^1_t \mathbf{P}_{t,s} \Phi^1 \mathbb{1}_{\{1\}}(0) = \theta_t p_{t,s}(1,4) + (1-\theta_t)\, p_{t,s}(2,4),$$

where the computation of the last probability can be found in Appendix A. Furthermore, we have

$$\mathbb{P}\left(X^1_s = 1 \mid X_t = z\right) = \mathbb{P}\left(X^1_s = 1 \mid X^1_t = 1\right) = 1, \quad z \in \{(1,0),(1,1)\}.$$

Then

$$\mathbb{P}\left(X^1_s = 1 \mid X_t = z\right) = \mathbb{P}\left(X^1_s = 1 \mid X^1_t = 0\right), \quad z \in \{(0,0),(0,1)\}, \tag{2.60}$$

if and only if for any $0 \leq t \leq s$, $p_{t,s}(1,4) = p_{t,s}(2,4)$. *If Equation* (2.60) *holds, by Definition 2.4.1,* X^1 *is a Markov chain in the filtration of* X. *Saying differently, if* $p_{t,s}(1,4) \neq p_{t,s}(2,4)$, *then* X^1 *is a Markov chain in its own filtration, but not the filtration of* X.

Remark 2.6.1. *Note that in* (2.58) *if* $c_u = d_u$, $u \geq 0$, *then we have for any* $0 \leq t \leq s$, $p_{t,s}(1,4) = p_{t,s}(2,4)$. *However, the converse implication is not true.*

Note that Θ^1_u *depends on* u. *Thus, the infinitesimal generator of* X^1 *is time-*

inhomogeneous and given by

$$\Lambda_u^1 = \Theta_u^1 \Lambda_u \Phi^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -\theta_u (c_u - d_u) - d_u & \theta_u (c_u - d_u) + d_u \\ 0 & 0 \end{pmatrix} \end{array}, \quad u \geq 0,$$

where

$$\theta_u = \frac{\mathrm{e}^{-\int_0^u (a_v + c_v)\,\mathrm{d}v}}{\frac{c_u - d_u}{a_u + c_u - d_u} \mathrm{e}^{-\int_0^u (a_v + c_v)\,\mathrm{d}v} + \frac{a_u}{a_u + c_u - d_u} \mathrm{e}^{-\int_0^u d_v\,\mathrm{d}v}}.$$

Example 2.6.2. *Consider two processes X^1, X^2 taking values in $\{0, 1\}$. Let 1 be the absorbing state. Suppose that $X = (X^1, X^2)$ is a bivariate Markov chain and its generator is given by $(\Lambda_u, u \geq 0)$,*

$$\Lambda_u = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a_u + b_u + c_u) & a_u & b_u & c_u \\ 0 & -d_u & 0 & d_u \\ 0 & 0 & -f_u & f_u \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array},$$

where for any $u \geq 0$, $a_u, b_u, c_u, d_u, f_u \geq 0$, and locally integrable on $[0, T)$. For any $0 \leq t \leq s$, the corresponding transition probability matrix function is of the form,

$$\mathbf{P}_{t,s} = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} 1 - \sum_{j=2}^4 p_{t,s}(1,j) & p_{t,s}(1,2) & p_{t,s}(1,3) & p_{t,s}(1,4) \\ 0 & 1 - p_{t,s}(2,4) & 0 & p_{t,s}(2,4) \\ 0 & 0 & 1 - p_{t,s}(3,4) & p_{t,s}(3,4) \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array},$$

where $0 \leq p_{t,s}(i,j) \leq 1,\ i,j = 1,2,3,4$. We assume that the initial distribution is given by $\mathbb{P}(X_0 = (0,1)) = 1$.

Following the similar computations in Example 2.6.1, the operator Θ_t^1 is rep-

resented by

$$\Theta_t^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}, \quad t \geq 0.$$

Then we have for any $t \geq 0$, $\Theta_t^1 \Phi^1 = \mathbf{I}_{2\times 2}$,

$$\Theta_t^1 \mathbf{P}_{t,s} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} 0 & 1 - p_{t,s}(2,4) & 0 & p_{t,s}(2,4) \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}, \quad 0 \leq t \leq s,$$

and

$$\widehat{\mathbf{P}}_{t,s}^1 \Theta_s^1 = \Theta_t^1 \mathbf{P}_{t,s} \Phi^1 \Theta_s^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} 0 & 1 - p_{t,s}(2,4) & 0 & p_{t,s}(2,4) \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}, \quad 0 \leq t \leq s.$$

In view of Theorem 2.5.2 we know that X^1 *is a Markov chain in its own filtration with time-inhomogeneous infinitesimal generator*

$$\Lambda_u^1 = \Theta_u^1 \Lambda_u \Phi^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -d_u & d_u \\ 0 & 0 \end{pmatrix} \end{array}, \quad u \geq 0.$$

Note that for any $t \geq 0$, *we have* $\mathbb{P}(X_t = (0,0)) = \mathbb{P}(X_t = (1,0)) = 0$. *Condition (M) is not necessary for strong Markovian consistency. We need to check by definition of strong Markovian consistency if* X^1 *is a Markov chain in the filtration of* X. *The transition matrix function of* X^1 *is given by*

$$\mathbf{P}_{t,s}^1 = \widehat{\mathbf{P}}_{t,s}^1 = \Theta_t^1 \mathbf{P}_{t,s} \Phi^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} p_{t,s}(2,2) & p_{t,s}(2,4) \\ 0 & 1 \end{pmatrix} \end{array}, \quad 0 \leq t \leq s.$$

It follows that for any $x = (x^1, x^2) \in \{(0,1),(1,0),(1,1)\}$[5]*, we have*

$$\mathbb{P}\left(X_s^1 = x_s^1 \mid X_t = x_t\right) = \mathbb{P}\left(X_s^1 = x_s^1 \mid X_t^1 = x_t^1\right), \quad 0 \le t \le s.$$

We conclude that X^1 *is a Markov chain in the filtration of* X *as well.*

On page 47 in Lemma 2.5.3, condition (Ci-II) has equivalent statements. Each statement requires to verify all possible trajectories for $t \ge 0$. In Example 2.6.3 we will show that the structure of $(\Lambda_u, u \ge 0)$ can make the verification effortless if X starts from a *right* initial distribution μ^X. Moreover, given a class of $(\Lambda_u, u \ge 0)$, we discuss how the algebraic conditions on the entries of $(\Lambda_u, u \ge 0)$ will make X be weak Markovian consistancy, strong Markovian consistency, or weak-only Markovian consistency.

Example 2.6.3. *Consider a bivariate Markov chain* $X = (X^1, X^2)$ *generated by the same infinitesimal generator* $(\Lambda_u, u \ge 0)$ *in Example 2.6.2. But the initial distribution is* $\mathbb{P}(X_0 = (0,0)) = 1$.

In this example, we will show that

(i) X *is weakly Markovian consistent.*

(ii) Furthermore, if for any $u \ge 0$ *we have* $a_u, b_u, c_u > 0$, $d_u, f_u \ge 0$, *and*

$$a_u + c_u \ne f_u, \qquad\qquad b_u + c_u \ne d_u,$$

then X *is weak-only Markovian consistent.*

Let $0 = t_0 \le t_1 \le \cdots \le t_n = t < \infty$. *We show that* X^1 *and* X^2 *are Markovian in their own filtrations by verifying condition (Ci-II). In view of Lemma*

[5]Since $\mathbb{P}(X_0 = (0,1)) = 1$ and 1 is an absorbing state, for any $t \ge 0$ the one-dimensional probability $\mathbb{P}(X_t = (0,0)) = 0$.

2.5.3, statement (iii), we need to check for $i = 1, 2$, any $x^i \in \mathcal{K}$, $x \in \mathcal{K}^2$, $n \in \mathbb{N}$, we have

$$\sum_{x_{n-1} \in \mathcal{H}\left(x^i_{n-1}\right)} \sum_{x_0 \in \mathcal{H}\left(x^i_0\right)} \mathbb{P}\left(X_{t_{n-1}} = x_{n-1},\ X^i_{t_{n-2}} = x^i_{n-2}, \ldots, X^i_{t_1} = x^i_1,\ X_0 = x_0\right)$$
$$\left[\ \mathbb{P}\left(X^i_{t_n} = x^i_n\right) \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n} - \mathbb{P}\left(X_{t_n} = x_n\right) \mathbf{P}^{x_{n-1},x^i_n}_{t_{n-1},t_n}\ \right] = 0. \quad (2.61)$$

First we reformulate

$$\begin{aligned}\mathbb{P}\left(X^i_{t_n} = x^i_n\right) &= \sum_{x_n \in \mathcal{H}(x^i_n)} \mathbb{P}\left(X_{t_n} = x_n\right) \\ &= \sum_{x_n \in \mathcal{H}(x^i_n)} \sum_{z \in \mathcal{K}^2} \mathbb{P}\left(X_{t_{n-1}} = z\right) \mathbf{P}^{z,x_n}_{t_{n-1},t_n},\end{aligned}$$

and

$$\mathbb{P}\left(X_{t_n} = x_n\right) = \sum_{z \in \mathcal{K}^2} \mathbb{P}\left(X_{t_{n-1}} = z\right) \mathbf{P}^{z,x_n}_{t_{n-1},t_n}.$$

Then the square bracket of Equation (2.61) is equivalent to

$$\begin{aligned}&\mathbb{P}\left(X^i_{t_n} = x^i_n\right) \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n} - \mathbb{P}\left(X_{t_n} = x_n\right) \mathbf{P}^{x_{n-1},x^i_n}_{t_{n-1},t_n} \\ &= \sum_{z \in \mathcal{K}^2} \mathbb{P}\left(X_{t_{n-1}} = z\right) \left[\left(\sum_{x_n \in \mathcal{H}(x^i_n)} \mathbf{P}^{z,x_n}_{t_{n-1},t_n} \right) \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n} - \mathbf{P}^{z,x_n}_{t_{n-1},t_n} \left(\sum_{x_n \in \mathcal{H}(x^i_n)} \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n} \right) \right].\end{aligned} \quad (2.62)$$

Next, we represent $\Xi^i\left(t_n, x_n, \pi_{t_n}, \psi, \mu^X\right) = 0$ as a homogeneous system by Lemma 2.5.3, statement (iii) for any $x^i \in \mathcal{K}$, any $x \in \mathcal{K}^2$, any $n \in \mathbb{N}$,

$$\kappa^i_{t_{n-1},t_n} f\left(t_n, x_n, \pi_{t_n}, \psi, \mu^X\right) = 0,$$

where f is a column vector, and $\kappa^i_{t_{n-1},t_n}$ is a matrix,

$$f = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix} = \begin{pmatrix} \mathbb{P}\left(X_{t_{n-1}} = (0,0),\ X^i_{t_{n-2}} = x^i_{n-2}, \ldots,\ X^i_0 = x^i_0\right) \\ \mathbb{P}\left(X_{t_{n-1}} = (0,1),\ X^i_{t_{n-2}} = x^i_{n-2}, \ldots,\ X^i_0 = x^i_0\right) \\ \mathbb{P}\left(X_{t_{n-1}} = (1,0),\ X^i_{t_{n-2}} = x^i_{n-2}, \ldots,\ X^i_0 = x^i_0\right) \\ \mathbb{P}\left(X_{t_{n-1}} = (1,1),\ X^i_{t_{n-2}} = x^i_{n-2}, \ldots,\ X^i_0 = x^i_0\right) \end{pmatrix}, \quad (2.63)$$

and

$$\kappa^i_{t_{n-1},t_n} = \left[\kappa^{i;x^i_t,x_{n-1}}_{t_{n-1},t_n}\right] = \left[\kappa^i_{t_{n-1},t_n}(j,k)\right]_{j=1,2,\ k=1,2,3,4}$$

$$= \begin{array}{c} \\ x^i_n=0 \\ x^i_n=1 \end{array} \begin{array}{c} \begin{array}{cccc} x_{n-1}=(0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \kappa^i_{t_{n-1},t_n}(1,1) & \kappa^i_{t_{n-1},t_n}(1,2) & \kappa^i_{t_{n-1},t_n}(1,3) & \kappa^i_{t_{n-1},t_n}(1,4) \\ \kappa^i_{t_{n-1},t_n}(2,1) & \kappa^i_{t_{n-1},t_n}(2,2) & \kappa^i_{t_{n-1},t_n}(2,3) & \kappa^i_{t_{n-1},t_n}(2,4) \end{pmatrix} \end{array},$$

the jth row of $\kappa^i_{t_{n-1},t_n}$ corresponding to the jth ordered state space of $X^i_{t_n}$ [6]*, and the kth column of $\kappa^i_{t_{n-1},t_n}$ corresponding to the kth ordered state space of $X_{t_{n-1}}$.*

For example, for $\mathcal{K} = \{0,1\}$, it follows from Equation (2.62)

$$\kappa^i_{t_{n-1},t_n}(2,3) = \sum_{x_n \in \mathcal{H}(x^i_n)} \sum_{z \in \mathcal{K}^2} \mathbb{P}\left(X_{t_{n-1}} = z\right) \times \left[\left(\sum_{x_n \in \mathcal{H}(x^i_n)} \mathbf{P}^{z,x_n}_{t_{n-1},t_n}\right) \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n} - \mathbf{P}^{z,x_n}_{t_{n-1},t_n}\left(\sum_{x_n \in \mathcal{H}(x^i_n)} \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n}\right)\right],$$

with $x^i_n = 1$, $x_{n-1} = (1,0)$.

We fix $i = 1$. Computations of $\kappa^1_{t_{n-1},t_n}(j,k)$ can be found in Appendix B.

Remark 2.6.2. *In Lemma 2.5.3, statement* (*iii*)*, it requires to verify whether $\kappa^1_{t_{n-1},t_n} f = 0$ at each time $t \geq 0$, for $x_t \in \{(0,0),(0,1),(1,0),(1,1)\}$. Because $|\mathcal{K}| = 2$, for $x_n \in \{(0,0),(0,1)\}$ and $x_n \in \{(1,0),(1,1)\}$, the corresponding $\kappa^1_{t_{n-1},t_n}(j,k)$'s are different up to a plus-minus sign. Thus, it suffices to check only one x_n in each set $\{(0,0),(0,1)\}$ and $\{(1,0),(1,1)\}$. Saying differently, we verify for $x_n \in \{(0,0),(1,0)\}$, or $x_n \in \{(0,1),(1,1)\}$, or other combinations.*

Now, by assumption the initial distribution $\mathbb{P}\left(X^1_0 = 0\right) = 1$, it suffice to check $x^1_0 = 0$. In view of (2.63)*, after replacing x^1_0 with value* 0*, each f_i still depends on the*

[6]Note that in the statement (*iii*) of Lemma 2.5.3 we check for x^i_n, but not x_n.

trajectory prior to time t_{n-1}*. However, the structure of* Λ_u *results in many entries of* $\kappa^1_{t_{n-1},t_n}$ *with value* 0*. Next, we compute* $\kappa^1_{t_{n-1},t_n} f$ *for* $x_n = (0,0)$*,*

$$\begin{aligned}
&\kappa^1_{t_{n-1},t_n}(1,1)\cdot f_1 + \kappa^1_{t_{n-1},t_n}(1,2)\cdot f_2 + \kappa^1_{t_{n-1},t_n}(1,3)\cdot f_3 + \kappa^1_{t_{n-1},t_n}(1,4)\cdot f_4 \\
&\qquad = \left(\mathbb{P}\left(X_{t_{n-1}} = (0,1)\right)\mathbf{P}^{(0,1),(0,1)}_{t_{n-1},t_n}\right)\mathbf{P}^{(0,0),(0,0)}_{t_{n-1},t_n}\cdot f_1 \\
&\qquad\qquad - \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right)\mathbf{P}^{(0,0),(0,0)}_{t_{n-1},t_n}\right)\mathbf{P}^{(0,1),(0,1)}_{t_{n-1},t_n}\cdot f_2 \\
&\qquad = 0,
\end{aligned}$$

where the first equality is due to $\kappa^1_{t_{n-1},t_n}(1,3) = \kappa^1_{t_{n-1},t_n}(1,4) = 0$*; the second equality comes from the fact that if* $\mathbb{1}_{\{(0,0)\}}\left(X_{t_n}\right) = 1$*, then* $\mathbb{1}_{\{(0,1)\}}\left(X_{t_{n-1}}\right) = 0$*.*

For $\mathbb{P}\left(X_0 = (0,0)\right) = 1$ *and* $x_n = (0,0)$*, we know*

$$\Xi^1\left(t_n, x_n, \pi_{t_n}, \psi, \mu^X\right) = 0.$$

In view of Remark 2.6.2, for $\mathbb{P}\left(X_0 = (0,0)\right) = 1$ *and* $x_n = (0,1)$*, we have*

$$\Xi^1\left(t_n, (0,1), \pi_{t_n}, \psi, \mu^X\right) = -\Xi^1\left(t_n, (0,0), \pi_{t_n}, \psi, \mu^X\right) = 0.$$

Since $0 \le t_n = t$ *is arbitrary, we have for* $x^1_t = 0$ *and* $0 \le t \le s$*, any* $\pi_t \in \Pi_t$*, and any* $\psi \in \Psi^{x^1_t}_{|\pi_t|}$*,*

$$\sum_{x^2_t \in \mathcal{K}} \mathbf{P}^{x_t, x^1_s}_{t,s}\Xi^1\left(t, x_t, \pi_t, \psi, \mu^X\right) = \sum_{x^2_t \in \mathcal{K}} \mathbf{P}^{\left(0,x^2_t\right), x^1_s}_{t,s}\Xi^1\left(t, \left(0, x^2_t\right), \pi_t, \psi, \mu^X\right) = 0.$$

When $x^1_t = 1$*, by Proposition 2.5.3 on page 46, we have for* $0 \le t \le s$*, any* $\pi_t \in \Pi_t$*, and any* $\psi \in \Psi^{x^1_t}_{|\pi_t|}$*,*

$$\sum_{x^2_t \in \mathcal{K}} \mathbf{P}^{x_t, x^1_s}_{t,s}\Xi^1\left(t, x_t, \pi_t, \psi, \mu^X\right) = \sum_{x^2_t \in \mathcal{K}} \mathbf{P}^{\left(1,x^2_t\right), x^1_s}_{t,s}\Xi^1\left(t, \left(1, x^2_t\right), \pi_t, \psi, \mu^X\right) = 0.$$

We conclude that for every $x^1_t \in \mathcal{K} = \{0,1\}$ *it holds that*

$$\sum_{x^2_t \in \mathcal{K}} \mathbf{P}^{x_t, x^1_s}_{t}\Xi^1\left(t, x_t, \pi_t, \psi, \mu^X\right) = 0.$$

Thus, X^1 is Markovian in its own filtration. By analogy, we can show that X^2 is Markovian in its own filtration as well. Therefore, we conclude that the bivariate Markov chain X is weakly Markovian consistent. By [BJN13, Theorem 1.11], the generator of X^i is given by

$$\Lambda_u^i = \Theta_u^i \Lambda_u \Phi^i, \quad u \geq 0, \ i = 1, 2.$$

Now, in view of Theorem 2.5.3 and Remark 2.5.9, if the parameters a_u, b_u, c_u in Λ_u are such that

$$a_u > 0, \qquad b_u > 0, \qquad c_u > 0,$$

then for $t \geq 0$ the one-dimensional distribution is positive,

$$\mathbb{P}\left(X_t = x_t\right) = \sum_{x_0 \in \mathcal{K}^2} \mathbb{P}\left(X_0 = x_0\right) \mathbf{P}_{0,t}^{x_0, x_t} > 0, \quad x_t \in \mathcal{K}^2.$$

Specifically, the one-dimensional probabilities are given by

$$\mathbb{P}\left(X_t = (0,0)\right) = 1 - \sum_{j=2}^{4} p_{0,t}(1,j) > 0$$

$$\mathbb{P}\left(X_t = (0,1)\right) = p_{0,t}(1,2) > 0$$

$$\mathbb{P}\left(X_t = (1,0)\right) = p_{0,t}(1,3) > 0$$

$$\mathbb{P}\left(X_t = (1,1)\right) = p_{0,t}(1,4) > 0.$$

Therefore, we conclude that X is strongly Markovian consistent if and only if condition (M) holds. Finally, in view of Theorem 2.5.4 on page 67, if, additionally, for any $u \geq 0$,

$$a_u + c_u \neq f_u, \qquad b_u + c_u \neq d_u,$$

then X is weak-only Markovian consistent.

In Example 2.6.3, the generator $(\Lambda_u, u \geq 0)$ of X has a special structure, i.e. the lower triangular entries being all zero. Given $\mathbb{P}\left(X_0 = (0,0)\right) = 1$, we verify that

condition (Ci-II) holds. The approach of verification is independent of functions a, b, c, d, f. In view of Theorem 2.5.2 on page 56, the sufficient conditions are purely algebraic conditions. That is, the sufficient conditions depend on the exact values of parameters $(a_u, b_u, c_u, d_u, f_u, u \geq 0)$, rather than whether some state is absorbing or not. Therefore, given the initial distribution $\mathbb{P}(X_0 = (0,0)) = 1$ and the general form of $(\Lambda_u, u \geq 0)$ in Example 2.6.3, it is inconclusive whether Equation (2.35) on page 56 holds or not.

In view of Example 2.6.3, it might be tempting to conclude that a special structure of $(\Lambda_u, u \geq 0)$ guarantees Markovianity of the components. In the next example, we provide a counterexample showing that such structure does not promise Markovianity of X^i.

Example 2.6.4. *Consider a bivariate Markov chain $X = (X^1, X^2)$ generated by the infinitesimal generator $(\Lambda_u, u \geq 0)$. Let the state space of X^i, $i = 1, 2$, be $\mathcal{K} = \{0, 1, 2\}$, and 2 be the absorbing state. Assume that each component X^i transits solely in one way direction. That is, for $0 \leq t \leq s$,*

$$X_t^i \leq X_s^i, \quad i = 1, 2.$$

Then, all the lower-triangular entries of the generator $(\Lambda_u, u \geq 0)$ are zeros,

$$\Lambda_u = [\lambda_u^{x,y}]_{x,y \in \mathcal{K}^2}$$

$$
= \begin{array}{c} \\ (0,0) \\ (0,1) \\ (0,2) \\ (1,0) \\ (1,1) \\ (1,2) \\ (2,0) \\ (2,1) \\ (2,2) \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} (0,0) & (0,1) & (0,2) & (1,0) & (1,1) & (1,2) & (2,0) & (2,1) & (2,2) \end{array} \\
\begin{pmatrix}
\lambda_u(1,1) & \lambda_u(1,2) & \lambda_u(1,3) & \lambda_u(1,4) & \lambda_u(1,5) & \lambda_u(1,6) & \lambda_u(1,7) & \lambda_u(1,8) & \lambda_u(1,9) \\
0 & \lambda_u(2,2) & \lambda_u(2,3) & 0 & \lambda_u(2,5) & \lambda_u(2,6) & 0 & \lambda_u(2,8) & \lambda_u(2,9) \\
0 & 0 & \lambda_u(3,3) & 0 & 0 & \lambda_u(3,6) & 0 & 0 & \lambda_u(3,9) \\
0 & 0 & 0 & \lambda_u(4,4) & \lambda_u(4,5) & \lambda_u(4,6) & \lambda_u(4,7) & \lambda_u(4,8) & \lambda_u(4,9) \\
0 & 0 & 0 & 0 & \lambda_u(5,5) & \lambda_u(5,6) & 0 & \lambda_u(5,8) & \lambda_u(5,9) \\
0 & 0 & 0 & 0 & 0 & \lambda_u(6,6) & 0 & 0 & \lambda_u(6,9) \\
0 & 0 & 0 & 0 & 0 & 0 & \lambda_u(7,7) & \lambda_u(7,8) & \lambda_u(7,9) \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_u(8,8) & \lambda_u(8,9) \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\end{array}.
$$

We will show that X^1 is not necessarily Markovian in its own filtration. For simplicity, we assume two periods $0 = t_0 < t_1 < t_2$ with equal time interval $\Delta t = t_n - t_{n-1},\ n = 1, 2$. The component X^1 is Markovian in its own filtration if the equality holds

$$
\mathbb{P}\left(X^1_{t_2} = x^1_2 \mid X^1_{t_1} = x^1_1,\ X^1_0 = x^1_0\right) = \mathbb{P}\left(X^1_{t_2} = x^1_2 \mid X^1_{t_1} = x^1_1\right), \quad x^1_0, x^1_1, x^1_2 \in \mathcal{K}. \tag{2.64}
$$

We verify (2.64) *for $x^1_1 = x^1_2 = 1$. There are two possible realizations of X^1 according to the partition $0 = t_0 < t_1 < t_2$ to reach $x^1_2 = 1$,*

$$
\varphi = \{x^1_0, x^1_1, x^1_2\} = \{0, 1, 1\},\ \textit{or}\ \{1, 1, 1\}.
$$

We have the left-hand-side of (2.64)

$$
\mathbb{P}\left(X^1_2 = 1 \mid X^1_1 = 1,\ X^1_0 = x^1_0\right) = \frac{\mathbb{P}\left(X^1_2 = 1,\ X^1_1 = 1,\ X^1_0 = x^1_0\right)}{\mathbb{P}\left(X^1_1 = 1,\ X^1_0 = x^1_0\right)} \tag{2.65}
$$

and the right-hand-side of (2.64)

$$
\mathbb{P}\left(X^1_2 = 1 \mid X^1_1 = 1\right) = \frac{\mathbb{P}\left(X^1_2 = 1,\ X^1_1 = 1\right)}{\mathbb{P}\left(X^1_1 = 1\right)}
$$

$$= \frac{\mathbb{P}\left(X_2^1 = 1,\ X_1^1 = 1,\ X_0 = 0\right) + \mathbb{P}\left(X_2^1 = 1,\ X_1^1 = 1,\ X_0 = 1\right)}{\mathbb{P}\left(X_1^1 = 1,\ X_0 = 0\right) + \mathbb{P}\left(X_1^1 = 1,\ X_0 = 1\right)}. \tag{2.66}$$

If the initial distribution of X^1 *is* $\mathbb{P}\left(X_0^1 = x_0^1\right) = 1,\ x_0^1 \in \{0,1\}$*, then* (2.65) *and* (2.66) *are equivalent. Thus* (2.64) *holds. However, the initial distribution* $\mathbb{P}\left(X_0 = (0,0)\right) = 1$ *is not sufficient for the component* X^1 *to be a Markov chain. Because the initial distribution* $\mathbb{P}\left(X_0 = (0,0)\right) = 1$ *does not guarantee* $\kappa^1_{t_{n-1},t_n} f\left(t_n, x_n, \pi_{t_n}, \psi, \mu^X\right) = 0$*. If the exact formulation of* $(\Lambda_u, u \geq 0)$ *is given, we can use Theorem 2.5.2 to verify if* X^i *is a Markov chain.*

2.7 Conclusions and Future Work

We summarize our research results about Markovian consistency in the following:

Weak Markovian consistency

We derive condition (Ci), which is the necessary and sufficient condition for X to be weakly Markovian consistent with respect to the component X^i (cf. Theorem 2.5.1 on page 41). Thereinafter, we give several sufficient conditions to condition (Ci). One of the important sufficient conditions is condition (Ci-II). The most remarkable result is Theorem 2.5.2 on page 56, which gives sufficient conditions for condition (Ci-II) to hold in terms of the existence of initial distribution and the purely algebraic relations between the semigroups of X and the component X^i.

Strong Markovian consistency

Since strong Markovian consistency implies weak Markovian consistency, any sufficient conditions for strong Markovian consistency are also sufficient conditions for weak Markovian consistency. We give sufficient conditions for strong Markovian consistency in terms of condition (P) and establish the equivalence of condition (P)

and condition (M). Besides, we show that condition (P) implies the condition in Theorem 2.5.2. In Theorem 2.5.3 on page 65, we show that if all states are accessible, then condition (M) (or, equivalently, condition (P)) is necessary for strong Markovian consistency. That is, we can verify whether the component X^i is Markovian in the filtration of X by the algebraic structure of the transition matrix function of X.

Weak-only Markovian consistency

On page 67 in Theorem 2.5.4 we give conditions for weak-only Markovian consistency. Namely, given an initial distribution μ^X and an infinitesimal generator $(\Lambda_u, u \geq 0)$, we verify whether μ^X and $(\Lambda_u, u \geq 0)$ satisfy the condition (Ci). This will determine if X^i is Markovian in its own filtration. If, additionally, all states are accessible, and condition (P) does not hold, then X^i is not Markovian in the filtration of X. If the criterion hold true for all components, we conclude that X is weak-only Markovian consistent.

Future work

We list the future lines of research in the following:

- In view of Lemma 2.5.3, with proper quantifiers the equality

$$\Xi^i\left(t, x_t, \pi_t, \psi, \mu^X\right) = 0, \tag{2.67}$$

  is equivalent to an equality appeared in the existing literature,

$$\mathbb{P}\left(X_t = x_t \mid \mathcal{F}_t^{X^i}\right) = \mathbb{E}\left(X_t = x_t \mid X_t^i\right). \tag{2.68}$$

  We have one sufficient condition for Equation (2.67) to hold. This sufficient condition is related to the existence of initial distribution, and the algebraic relations of the semigroup of X and the transition matrix function of its component. However, this condition is not necessary. There might exist other sufficient conditions to Equation (2.67).

A future work regarding the sufficient condition we gave is to characterize this initial distribution for a given family of generator structures. Or, one may work on finding other sufficient conditions to Equation (2.67).

- To construct an example with all non-absorbing states

  In our examples, we focus on the state space consisting of one absorbing state and one non-absorbing state. When one checks those sufficient conditions for Markovian consistency, the verification involves the transition matrix function, that is a functional of the generators. If all states are non-absorbing, every element of the transition matrix function is non-zero. The analytical form of the transition matrix function is unavailable, thus to find a non-trivial example with all non-absorbing states is not an easy task.

- The study of general time-inhomogeneous Markov processes

  Our results are subject to the class of multivariate Markov chains. An interesting topic is to extend the results to general time-inhomogeneous Markov processes.

CHAPTER 3
MARKOV STRUCTURES

3.1 Introduction

We know from Chapter 2 that if $(\Lambda_t, t \geq 0)$ is a valid generator, the matrix function $(\Theta_t^i \Lambda_t \Phi^i, t \geq 0)$ generates some Markov chain, say Y^i. We study the conditions imposed on $(\Lambda_t, t \geq 0)$ under which the component X^i of X is a Markov chain with generator function $(\Theta_t^i \Lambda_t \Phi^i, t \geq 0)$, which is equivalent to

$$Y^i \stackrel{\mathcal{L}}{=} X^i, \quad \text{if} \quad Y_0^i \stackrel{\mathcal{L}}{=} X_0^i$$

where $\stackrel{\mathcal{L}}{=}$ means equality in law. Preserving the prescribed marginal laws is the core of our interests. Those sufficient conditions allow us to tell if the components of a multivariate Markov chain have the prescribed marginal laws.

In this chapter, we focus on the construction of the infinitesimal generator $(\Lambda_t, t \geq 0)$ such that $(\Lambda_t, t \geq 0)$ satisfies some of the aforementioned sufficient conditions. We call the corresponding multivariate Markov chain a Markov structure for a collection of Markov chains $\{Y^1, \ldots, Y^m\}$. Recall that the sufficient conditions are given in terms of the transition semigroup of X. However, the transition semigroup is a functional of $(\Lambda_t, t \geq 0)$, and the analytical form of the time-inhomogeneous transition semigroup is unavailable in general. The construction is not a straightforward task.

The direct approach is to start with a preliminary $(\Lambda_t, t \geq 0)$. Then one computes the transition semigroup corresponding to $(\Lambda_t, t \geq 0)$ and the one-dimensional distributions of X and X^i's. Next, one verifies if the sufficient condition is satisfied. If the sufficient condition is not satisfied, one changes $(\Lambda_t, t \geq 0)$ and run the same

procedure again. Basically, we need to calibrate $(\Lambda_t, t \geq 0)$ such that it meets the sufficient condition. We refer this method to the *top-down approach*. Although this approach requires intensive computations, this method is comprehensive in the sense that the evolution of X is encoded in the transition semigroup.

On the other hand, since the sufficient condition is related to the one-dimensional distributions of of X and X^i's, it motivates us to study the *bottom-up approach*. That is, we obtain the one-dimensional distributions of X^i's first, then we construct the one-dimensional distribution of X through the joint distributions of random variables X_t^i's. Next, we solve for $(\Lambda_t, t \geq 0)$ such that the sufficient condition is satisfied. Saying differently, we bond all the joint distributions of X_t^i's in such a way that the process X has the desired properties. However, as we will see later, this approach does not always produce satisfactory results as the top-down approach. We need to impose conditions such that those constructed one-dimensional distributions satisfy semigroup property.

This chapter is organized as follows. We begin by introducing Markov structures in Section 3.2. We give definitions of strong Markov structures and weak Markov structures. Thereinafter, we show that, in general, Markov structures are not empty sets. In Section 3.3, we discuss how to construct Markov structures. In particular, we focus on the bottom-up construction and provide sufficient conditions for the construction to be consistent with the top-down approach. Besides, we discuss the limitations of the bottom-up construction. In Section 3.4, we connect our results regarding Markovian consistency and Markov structures to the existing results in the literature. We conclude this cheaper in Section 3.5.

3.2 Markov Structures

The purpose of introducing Markov structures is to study stochastic dependence between the components of a multivariate Markov chain, subject to marginal constraints: we want to construct a multivariate Markov chain $X = (X^1, X^2, \ldots, X^m)$ such that its components are Markov chains with the prescribed laws.

More specifically, given a collection of Markov chains $\{Y^1, Y^2, \ldots, Y^m\}$, we say that a process X is a Markov structure for $\{Y^1, Y^2, \ldots, Y^m\}$, if X is a Markov chain, and each component X^i is a Markov chain with identical law to Y^i. As it will be seen later one can typically construct numerous multivariate chains $X = (X^1, X^2, \ldots, X^m)$ such that their components are Markov and that they have prescribed laws.

In what follows, based on [BJN13] we embark on defining *Markov structures* using infinitesimal generators. In analogy to strong and weak Markovian consistency, we have strong Markov structures and weak Markov structures.

3.2.1 Strong Markov Structures. In view of [BJN13, Theorem 1.8], suppose that for each $i \in \mathcal{M}$, the generator function $\Lambda_t^i = \left[\lambda_t^{i;x^i,y^i}\right],\ t \geq 0$ is given by

$$\lambda_t^{i;x^i,y^i} = \sum_{\substack{y^j \in E_j \\ j=1,\ldots,m,\ j\neq i}} \lambda_t^{x,y}, \quad x^i, y^i \in E_i,\ x^i \neq y^i, \tag{3.1}$$

and

$$\lambda_t^{i;x^i,x^i} = -\sum_{y^i \in E_i,\ x^i \neq y^i} \lambda_t^{i;x^i,y^i}, \quad x^i \in E_i, \tag{3.2}$$

where $x = (x^1, \ldots, x^i, \ldots, x^m)$ and $y = (y^1, \ldots, y^i, \ldots, y^m)$. We want to solve for a valid generator matrix function $\Lambda_t = \left[\lambda_t^{x,x'}\right],\ t \geq 0$ such that (3.1) and (3.2) hold. If there exists such a generator function Λ_t, then by construction we can verify that condition (M) is satisfied. Since condition (M) implies strong Markovian consistency, every component X^i of X is Markov in the filtration $\mathbb{F}^X$. Then each component X^i is generated by $\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i,\ t \geq 0$.

Definition 3.2.1 (Strong Markov structure)**.** *Let* $\{Y^1, Y^2, \ldots, Y^m\}$ *be a family of Markov chains. A multivariate process* $X = (X^1, X^2, \ldots, X^m)$ *is a strong Markov structure for* $\{Y^1, Y^2, \ldots, Y^m\}$*, if*

(i) $X = (X^1, X^2, \ldots, X^m)$ *is a Markov chain, and* X *satisfies the strong Markovian consistency property;*

(ii) each component X^i *of* X *has the same law as* Y^i*,* $X^i \overset{\mathcal{L}}{=} Y^i$*,* $i = 1, 2, \ldots, m$*.*

From the above discussions, we can construct strong Markov structures by solving the system consisting of (3.1) and (3.2). In general, the system admits multiple solutions, and some of the solutions may not be valid generators. Now, we confine ourselves to the valid generators, and fix $t \geq 0$. Since Λ_t is a valid generator with dimension $|E|$-by-$|E|$, Λ_t can be fully determined by appropriately assigning $\left(|E|^2 - |E|\right)$ entries. Note that $|E| = |E_1| \times \cdots \times |E_m|$. For fixed $i \in \mathcal{M}$, the matrix Λ_t^i is a valid generator with dimension $|E_i|$-by-$|E_i|$. In order to satisfy (3.1) and (3.2), it produces $\left(|E_i|^2 - |E_i|\right)$ equations. If (3.1) and (3.2) need to hold for all i's, there are totally $\sum_{i=1}^m \left(|E_i|^2 - |E_i|\right)$ equations. Clearly, if every $|E_i| \geq 2$, we have

$$\begin{aligned}
|E|^2 - |E| = |E|\left(|E| - 1\right) &= |E_1| \times \cdots \times |E_m| \cdot \left(|E_1| \times \cdots \times |E_m| - 1\right) \\
&\geq m \cdot \max\{|E_1|, \ldots, |E_m|\} \cdot \left(\max\{|E_1|, \ldots, |E_m|\} - 1\right) \\
&\geq \sum_{i=1}^m |E_i|\left(|E_i| - 1\right).
\end{aligned}$$

Then the number of unknown entries in Λ_t are greater than the number of equations.

Remark 3.2.1. *Note that there is at least one solution which is a valid generator. This solution corresponds to the tensor independence of* Λ_t^i*'s. For example, let us*

consider

$$\Lambda_t^1 = \begin{array}{c} \\ 0 \\ 1 \end{array}\!\!\begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -a_t & a_t \\ b_t & -b_t \end{pmatrix} \end{array}, \qquad \Lambda_t^2 = \begin{array}{c} \\ 0 \\ 1 \end{array}\!\!\begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -d_t & d_t \\ f_t & -f_t \end{pmatrix} \end{array},$$

where $t \geq 0,\ a_t, b_t, d_t, f_t \geq 0$. *We can verify that* $(\Lambda_t, t \geq 0)$ *given by*

$$\Lambda_t = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array}\!\!\begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a_t+d_t) & d_t & a_t & 0 \\ f_t & -(a_t+f_t) & 0 & a_t \\ b_t & 0 & -(b_t+d_t) & d_t \\ 0 & b_t & f_t & -(b_t+f_t) \end{pmatrix} \end{array},$$

indeed satisfies (3.1) *and* (3.2). *As we will see in Section 4.2, we call this type of structure as independence form. This implies that* X^i*'s are independent.*

3.2.2 Weak Markov Structures. The counterpart of the strong Markov structure is the weak Markov structure.

Definition 3.2.2 (Weak Markov structure)**.** *Let* $\{Y^1, Y^2, \ldots, Y^m\}$ *be a family of Markov chains. A multivariate process* $X = (X^1, X^2, \ldots, X^m)$ *is a weak Markov structure for* $\{Y^1, Y^2, \ldots, Y^m\}$*, if*

(i) $X = (X^1, X^2, \ldots, X^m)$ *is a Markov chain, and* X *satisfies the weak Markovian consistency property;*

(ii) each component X^i *of* X *has the same law as* Y^i*,* $X^i \overset{\mathcal{L}}{=} Y^i,\ i = 1, 2, \ldots, m$.

We know from [BJN13, Theorem 1.11] if the component X^i is Markov in its own filtration, then the generator matrix function of X^i is given by $\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i$.

However, $\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i$ is not sufficient for X^i to be Markov in its own filtration. If Λ_t^i is an input data, for those generator matrix functions Λ_t satisfying the condition

$$\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i,$$

it does not guarantee that the component X^i of X is Markovian in its own filtration. We will use the results from Chapter 2 to verify if X^i of X is a Markov chain.

Since modelling default contagion is excluded in the sense of strong Markov consistency (cf. [BCCH14] and [BJN13]), in practice, an important class is the weak-only Markov structure. We say that a multivariate process X is a *weak-only Markov structure* for $\{Y^1, Y^2, \ldots, Y^m\}$, if X is a weak Markov structure, but it is not a strong Markov structure.

In next section, we tackle one approach to constructing Markov structures.

3.3 Bottom-Up Construction of Markov Structures

Suppose that we are given a family of Markov chains $\{Y^1, Y^2, \ldots, Y^m\}$. Each Y^i is generated by the infinitesimal generator matrix $(\Lambda_t^i, t \geq 0)$, $i \in \{1, 2, \ldots, m\}$. We want to construct a multivariate Markov chain X with the infinitesimal generator matrix function $(\Lambda_t, t \geq 0)$ such that

$$\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i, \quad t \geq 0, \ i = 1, 2, \ldots, m, \tag{3.3}$$

and each component X^i of X has the same law as Y^i. Moreover, X satisfies strong, weak, or weak-only Markovian consistency.

Algebraically, if the infinitesimal generator matrix functions Λ_t^i's are given, the solution matrix Λ_t to Equation (3.3) may not be unique. In the following, we discuss two approaches to constructing a multivariate Markov chain which satisfies the above properties.

Recall that the operator Θ_t^i in Equation (3.3) can be represented by a matrix whose entries are conditional probabilities,

$$\Theta_t^i = \left[\mathbb{P}\left(X_t = x_t \mid X_t^i = x_t^i\right)\right]_{x_t \in \mathcal{K}^m, x_t^i \in \mathcal{K}},$$

where $x_t = (x_t^1, \ldots, x_t^i, \ldots, x_t^m)$. Note that the numerator in the above conditional probability

$$\mathbb{P}\left(X_t = x_t,\ X_t^i = x_t^i\right) = \mathbb{P}\left(X_t = x_t\right)$$

is the one-dimensional distribution of X at time $t \geq 0$, i.e., the joint distribution of random variables $X_t^i,\ i = 1, 2, \ldots, m$. In our setup, there are two methods to compute the one-dimensional distribution of X. A *top-down approach* is working through the transition semigroup of X,

$$\mathbb{P}\left(X_t = x_t\right) = \sum_{z \in \mathcal{K}^m} \mathbb{P}\left(X_0 = z\right) \mathbf{P}_{0,t}^{z,x_t}.$$

A *bottom-up approach* is to construct the joint distribution by the correlations between the components and the marginal distributions at every time $t \geq 0$.

Apparently, the top-down approach is more comprehensive, because the information about the dependence between the components is embedded in the transition semigroup of X. However, the semigroup $(\mathbf{P}_{t,s}, 0 \leq t \leq s)$ is a functional of $(\Lambda_u, u \geq 0)$, and in general, the analytical form of the the semigroup $(\mathbf{P}_{t,s}, 0 \leq t \leq s)$ is unavailable. It is extremely difficult to solve for $(\Lambda_u, u \geq 0)$ in Equation (3.3) analytically, and that makes the top-down approach almost impracticable. This motivates us the succeeding construction: the bottom-up approach.

The Bottom-Up Approach

We illustrate the bottom-up approach through an example. Let $Y^i,\ i = 1, 2$, be a Markov chain with values in $\mathcal{K} = \{0, 1\}$. Assume that Y^i is generated by

$(\Lambda_u^i, u \geq 0)$, $i = 1, 2$, respectively,

$$\Lambda_u^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -a_u & a_u \\ b_u & -b_u \end{pmatrix} \end{array}, \qquad \Lambda_u^2 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -d_u & d_u \\ f_u & -f_u \end{pmatrix} \end{array},$$

where $a_u, b_u, d_u, f_u \geq 0$. The corresponding transition probability matrix functions $\widehat{\mathbf{P}}_{t,s}^i$, $i = 1, 2$, are obtained by the Kolmogorov forward equation, for any $0 \leq t \leq s$,

$$\widehat{\mathbf{P}}_{t,s}^1 = \left[\widehat{\mathbf{P}}_{t,s}^{1;k,r}\right]_{k,r\in\mathcal{K}} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} \frac{b_s + a_s e^{-\int_t^s (a_v+b_v)\,\mathrm{d}v}}{a_s+b_s} & \frac{a_s\left(1-e^{-\int_t^s (a_v+b_v)\,\mathrm{d}v}\right)}{a_s+b_s} \\ \frac{b_s\left(1-e^{-\int_t^s (a_v+b_v)\,\mathrm{d}v}\right)}{a_s+b_s} & \frac{a_s + b_s e^{-\int_t^s (a_v+b_v)\,\mathrm{d}v}}{a_s+b_s} \end{pmatrix} \end{array},$$

and

$$\widehat{\mathbf{P}}_{t,s}^2 = \left[\widehat{\mathbf{P}}_{t,s}^{2;k,r}\right]_{k,r\in\mathcal{K}} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} \frac{f_s + d_s e^{-\int_t^s (d_v+f_v)\,\mathrm{d}v}}{d_s+f_s} & \frac{d_s\left(1-e^{-\int_t^s (d_v+f_v)\,\mathrm{d}v}\right)}{d_s+f_s} \\ \frac{f_s\left(1-e^{-\int_t^s (d_v+f_v)\,\mathrm{d}v}\right)}{d_s+f_s} & \frac{d_s + f_s e^{-\int_t^s (d_v+f_v)\,\mathrm{d}v}}{d_s+f_s} \end{pmatrix} \end{array}.$$

Next, we consider a bivariate Markov chain $X = (X^1, X^2)$ generated by the infinitesimal generator matrix function $(\Lambda_u, u \geq 0)$ with initial distribution μ^X. Each component of X takes value in $\mathcal{K} = \{0, 1\}$. We want to solve for $(\Lambda_u, u \geq 0)$ such that $\Lambda_u^i = \Theta_u^i \Lambda_u \Phi^i$, $u \geq 0$, $i = 1, 2$, and X satisfies Markovian consistency.

Note that at each time $t \geq 0$ the random variables X_t^1 and X_t^2 follow Bernoulli distributions with parameters p_t and q_t respectively. That is,

$$\mathbb{P}\left(X_t^1 = x_t^1\right) = \begin{cases} p_t, & \text{if } x_t^1 = 1, \\ 1 - p_t, & \text{if } x_t^1 = 0, \end{cases}$$

and

$$\mathbb{P}\left(X_t^2 = x_t^2\right) = \begin{cases} q_t, & \text{if } x_t^2 = 1, \\ 1 - q_t, & \text{if } x_t^2 = 0. \end{cases}$$

Since the initial distribution of X is given, the initial distributions of X^1 and X^2 are given by

$$\mu^{X^1}(x_0^1) := \mathbb{P}\left(X_0^1 = x_0^1\right) = \sum_{z \in \mathcal{K}} \mathbb{P}\left(X_0 = \left(x_0^1, z\right)\right) = \begin{cases} p_0, & \text{if } x_0^1 = 1, \\ 1 - p_0, & \text{if } x_0^1 = 0, \end{cases}$$

and

$$\mu^{X^2}(x_0^2) := \mathbb{P}\left(X_0^2 = x_0^2\right) = \sum_{z \in \mathcal{K}} \mathbb{P}\left(X_0 = \left(z, x_0^2\right)\right) = \begin{cases} q_0, & \text{if } x_0^2 = 1, \\ 1 - q_0, & \text{if } x_0^2 = 0, \end{cases}$$

In view of definitions of Markov structure, X^i has the same law as Y^i. That is, for any $t \geq 0$ the one-dimensional distribution of X^i is given by,

$$\mathbb{P}\left(X_t^i = x_t^i\right) = \sum_{k \in \mathcal{K}} \mathbb{P}\left(X_0^i = k\right) \cdot \widehat{\mathbf{P}}_{0,t}^{i;k,x_t^i}, \quad k, x_t^i \in \mathcal{K},\ i = 1, 2.$$

More precisely, for $t \geq 0$ and $\mathcal{K} = \{0, 1\}$, we have

$$\mathbb{P}\left(X_t^1 = x_t^1\right) = \begin{cases} (1 - p_0) \cdot \widehat{\mathbf{P}}_{0,t}^{1;0,1} + p_0 \cdot \widehat{\mathbf{P}}_{0,t}^{1;1,1}, & \text{if } x_t^1 = 1, \\ (1 - p_0) \cdot \widehat{\mathbf{P}}_{0,t}^{1;0,0} + p_0 \cdot \widehat{\mathbf{P}}_{0,t}^{1;1,0}, & \text{if } x_t^1 = 0, \end{cases}$$

and

$$\mathbb{P}\left(X_t^2 = x_t^2\right) = \begin{cases} (1 - q_0) \cdot \widehat{\mathbf{P}}_{0,t}^{2;0,1} + q_0 \cdot \widehat{\mathbf{P}}_{0,t}^{2;1,1}, & \text{if } x_t^2 = 1, \\ (1 - q_0) \cdot \widehat{\mathbf{P}}_{0,t}^{2;0,0} + q_0 \cdot \widehat{\mathbf{P}}_{0,t}^{2;1,0}, & \text{if } x_t^2 = 0. \end{cases}$$

Next, we denote by ρ_t the correlation between the random variables X_t^1 and X_t^2. It follows that the joint distributions of X_t^1 and X_t^2 are given by

$$\begin{aligned} c_t := \mathbb{P}\left(X_t^1 = 1,\ X_t^2 = 1; \rho_t\right) &= p_t q_t + \rho_t \sqrt{p_t(1 - p_t) q_t (1 - q_t)} \\ \mathbb{P}\left(X_t^1 = 1,\ X_t^2 = 0; \rho_t\right) &= p_t - c_t \\ \mathbb{P}\left(X_t^1 = 0,\ X_t^2 = 1; \rho_t\right) &= q_t - c_t \\ \mathbb{P}\left(X_t^1 = 0,\ X_t^2 = 0; \rho_t\right) &= 1 - p_t - q_t + c_t. \end{aligned}$$

Remark 3.3.1. *We require the correlation ρ_t within certain bounds so that for $t \geq 0$ the joint distribution, $\mathbb{P}\left(X_t^1 = x_t^1,\ X_t^2 = x_t^2; \rho_t\right),\ x_t^1, x_t^2 \in \{0,1\}$, is well-defined,*

$$
\begin{aligned}
-1 \leq \max\left(\frac{-p_t q_t}{\sqrt{p_t(1-p_t)q_t(1-q_t)}}, \frac{-(1-p_t)(1-q_t)}{\sqrt{p_t(1-p_t)q_t(1-q_t)}}\right) &\leq \rho_t \\
\leq \min\left(\frac{p_t(1-q_t)}{\sqrt{p_t(1-p_t)q_t(1-q_t)}}, \frac{(1-p_t)q_t}{\sqrt{p_t(1-p_t)q_t(1-q_t)}}\right) &\leq 1. \quad (3.4)
\end{aligned}
$$

Therefore, for $t \geq 0$ the operators $\Theta_t^i,\ i = 1,2$, are represented by

$$
\Theta_t^1 = \left[\mathbb{P}\left(X_t = x_t \mid X_t^1 = x_t^1\right)\right]_{x_t \in \mathcal{K}^2, x_t^1 \in \mathcal{K}} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \frac{1-p_t-q_t+c_t}{1-p_t} & \frac{q_t-c_t}{1-p_t} & 0 & 0 \\ 0 & 0 & \frac{p_t-c_t}{p_t} & \frac{c_t}{p_t} \end{pmatrix} \end{array},
$$

and

$$
\Theta_t^2 = \left[\mathbb{P}\left(X_t = x_t \mid X_t^2 = x_t^2\right)\right]_{x_t \in \mathcal{K}^2, x_t^2 \in \mathcal{K}} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \frac{1-p_t-q_t+c_t}{1-q_t} & 0 & \frac{p_t-c_t}{1-q_t} & 0 \\ 0 & \frac{q_t-c_t}{q_t} & 0 & \frac{c_t}{q_t} \end{pmatrix} \end{array}.
$$

As mentioned earlier, if the matrix function Λ_t^i is given by

$$
\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i, \quad t \geq 0,\ i = 1,2,
$$

Λ_t is not necessarily unique, or a valid generator. Now, we restrict ourselves to the case when Λ_t is indeed a valid infinitesimal generator matrix function,

$$
\Lambda_t = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -\sum_{k=2}^4 \lambda_t(1,k) & \lambda_t(1,2) & \lambda_t(1,3) & \lambda_t(1,4) \\ \lambda_t(2,1) & -\sum_{k=1,3,4} \lambda_t(2,k) & \lambda_t(2,3) & \lambda_t(2,4) \\ \lambda_t(3,1) & \lambda_t(3,2) & -\sum_{k=1,2,4} \lambda_t(3,k) & \lambda_t(3,4) \\ \lambda_t(4,1) & \lambda_t(4,2) & \lambda_t(4,3) & -\sum_{k=1}^3 \lambda_t(4,k) \end{pmatrix} \end{array},
$$

where $\lambda_t(j,k) \geq 0,\ j \neq k,\ j,k = 1,2,3,4$. The extension operators $\Phi^i,\ i = 1,2$, can

be represented by

$$\Phi^1 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \end{array}, \qquad \Phi^2 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}.$$

Since we want to solve for Λ_t with valid generators Λ_t^i as input data such that

$$\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i, \quad t \geq 0, \ i = 1, 2.$$

By comparing entry-wise of the both sides in the above identities, we have for $t \geq 0$ one linear system of four equalities,

$$\begin{aligned}
a_t &= \frac{1}{1-p_t}\big((\lambda_t(1,3) + \lambda_t(1,4))\,(1 - p_t - q_t + c_t) + (\lambda_t(2,3) + \lambda_t(2,4))\,(q_t - c_t)\,\big) \\
b_t &= \frac{1}{p_t}\big((\lambda_t(3,1) + \lambda_t(3,2))\,(p_t - c_t) + (\lambda_t(4,1) + \lambda_t(4,2))\,c_t\big) \\
d_t &= \frac{1}{1-q_t}\big((\lambda_t(1,2) + \lambda_t(1,4))\,(1 - p_t - q_t + c_t) + (\lambda_t(3,2) + \lambda_t(3,4))\,(p_t - c_t)\,\big) \\
f_t &= \frac{1}{q_t}\big((\lambda_t(2,1) + \lambda_t(2,3))\,(q_t - c_t) + (\lambda_t(4,1) + \lambda_t(4,3))\,c_t\big),
\end{aligned} \tag{3.5}$$

where $a_t, b_t, d_t, f_t \geq 0$. Note that in the system (3.5) a_t, b_t, d_t, f_t are input data. Once we choose the initial distribution μ^X and the correlation ρ_t, the probabilities p_t, q_t of the one-dimensional distributions are fully determined. Then we solve for $\lambda_t(j,k)$, $j \neq k$, $j,k = 1,2,3,4$, such that Λ_t is a valid infinitesimal generator. In (3.5), there are 12 unknowns $\lambda_t(j,k)$ and only 4 equalities. In general, there should be infinitely many combinations of $\lambda_t(j,k)$.

Algebraically, if we choose *at least* 8 entries of $\lambda_t(j,k)$ *properly*, the rest entries of $\lambda_t(j,k)$ will be uniquely determined by the system (3.5). Otherwise, the linear system (3.5) is not solvable.

Remark 3.3.2. *We can write down the system* (3.5) *in the matrix form. The matrix has dimension* 4*-by-*12*. If we transform this matrix to the reduced row echelon form, we can interchange the leading terms without changing the solutions, for example, exchanging* $\lambda_t(1,2)$ *with* $\lambda_t(1,4)$*, or swapping* $\lambda_t(1,3)$ *and* $\lambda_t(1,4)$*. As a result, there are many possibilities for these free variables, but not in an arbitrary way.*

Assume that we fix $\lambda_t(1,2), \lambda_t(1,3), \lambda_t(1,4), \lambda_t(2,1), \lambda_t(2,3), \lambda_t(3,1), \lambda_t(3,2)$, $\lambda_t(4,1)$. Because $a_t, b_t, d_t, f_t \geq 0,\ t \geq 0$, in view of the system (3.5), we have constraints on the individual jumps,

$$\begin{aligned}
0 \leq \lambda_t(1,3) + \lambda_t(1,4) &\leq \frac{1}{1-p_t-q_t+c_t}\big(a_t\,(1-p_t) - \lambda_t(2,3)\,(q_t-c_t)\,\big) \\
0 \leq \lambda_t(3,1) + \lambda_t(3,2) &\leq \frac{1}{p_t-c_t}\,(b_t p_t - \lambda_t(4,1)c_t) \\
0 \leq \lambda_t(1,2) + \lambda_t(1,4) &\leq \frac{1}{1-p_t-q_t+c_t}\big(d_t\,(1-q_t) - \lambda_t(3,2)\,(p_t-c_t)\,\big) \\
0 \leq \lambda_t(2,1) + \lambda_t(2,3) &\leq \frac{1}{q_t-c_t}\,(f_t q_t - \lambda_t(4,1)c_t)\,,
\end{aligned} \tag{3.6}$$

and constraints on the common jumps,

$$\begin{aligned}
0 \leq \lambda_t(2,3) &\leq \frac{1}{q_t-c_t}\big(a_t\,(1-p_t) - (\lambda_t(1,3)+\lambda_t(1,4))\,(1-p_t-q_t+c_t)\,\big) \\
0 \leq \lambda_t(3,2) &\leq \frac{1}{p_t-c_t}\big(d_t\,(1-q_t) - (\lambda_t(1,2)+\lambda_t(1,4))\,(1-p_t-q_t+c_t)\,\big) \\
0 \leq \lambda_t(4,1) &\leq \min\Big(\frac{1}{c_t}\big(b_t p_t - (\lambda_t(3,1)+\lambda_t(3,2))\,(p_t-c_t)\,\big) \\
&\qquad\qquad , \frac{1}{c_t}\big(f_t q_t - (\lambda_t(2,1)+\lambda_t(2,3))\,(q_t-c_t)\,\big)\Big)\,.
\end{aligned} \tag{3.7}$$

Once $\lambda_t(1,2), \lambda_t(1,3), \lambda_t(1,4), \lambda_t(2,1), \lambda_t(2,3), \lambda_t(3,1), \lambda_t(3,2), \lambda_t(4,1)$ are chosen such that (3.6) and (3.7) are satisfied, the remaining four entries $\lambda_t(2,4), \lambda_t(3,4), \lambda_t(4,2)$, $\lambda_t(4,3)$ are fully decided. Clearly, we have infinitely many combinations of $\lambda_t(j,k)$, $j \neq k$. Thus, given valid generators $\Lambda_t^i,\ i = 1,2$, the initial distribution $\mu^X > 0$, and the correlation ρ_t satisfying (3.4), we have a family of valid generators Λ_t for any $t \geq 0$,

$$\big\{\Lambda_t \mid \Lambda_t \text{ is a valid generator},\ \Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i,\ i = 1,2,\ t \in [t_0, T]\big\}. \tag{3.8}$$

Algebraically, we can solve for a family $(\Lambda_t, t \geq 0)$ to satisfy (3.8). Now, suppose that we have the family $(\Lambda_t, 0 \leq t \leq T)$. We need to do reality check. Namely, the joint distributions between the components from the bottom-up approach is consistent with the one retrieved from the top-down approach, i.e., for $t \in [0, T]$, $x_t^1, x_t^2 \in \{0, 1\}$, and ρ_t satisfies (3.4),

$$\underbrace{\mathbb{P}\left(X_t^1 = x_t^1,\ X_t^2 = x_t^2; \rho_t\right)}_{\text{by the bottom-up approach}} \overset{?}{=} \underbrace{\sum_{\left(x_0^1, x_0^2\right) \in \mathcal{K}^2} \mathbb{P}\left(X_0 = \left(x_0^1, x_0^2\right)\right) \mathbf{P}_{0,t}^{\left(x_0^1, x_0^2\right), \left(x_t^1, x_t^2\right)}}_{\text{by the top-down approach}}.$$

By construction, the joint distributions obtained from the bottom-up approach strongly depend on the correlation ρ_t. The conditions in Remark 3.3.1 are verified for each time $t \geq 0$ solely. Whereas, Remark 3.3.1 says nothing about the successive correlations $\rho_u, \rho_t, \rho_s,\ u \leq t \leq s$. On the other hand, in the top-down approach the correlation ρ_t between X_t^1 and X_t^2 is embedded in the process $X = (X^1, X^2)$. It is reasonable to ask which correlation series $\{\rho_t\}_{t \geq 0}$ will preserve the consistent property.

In what follows, we provide conditions on the correlation series $\{\rho_t\}_{t \geq 0}$ which give the consistent joint distributions between the top-down approach and the bottom-up approach.

Lemma 3.3.1. *Let $0 \leq u \leq t \leq s$. Assume that $\mathbf{P}_{u,t} = [P_{u,t}(i,j)]_{i,j \in \mathbb{N},\ i,j \leq |\mathcal{K}^m|}$ is a transition probability matrix. Then $\lim_{s \to t} \|\mathbf{P}_{u,s} - \mathbf{P}_{u,t}\| = 0$ if and only if $\lim_{s \to t} |P_{u,s}(i,j) - P_{u,t}(i,j)| = 0$.*

Proof. Notice that all norms are equivalent in a finite-dimensional vector space. It suffices to take 2-norm,

$$\|\mathbf{P}_{u,s} - \mathbf{P}_{u,t}\|_2^m = \sum_{i=1}^{|\mathcal{K}^m|} \sum_{j=1}^{|\mathcal{K}^m|} |P_{u,s}(i,j) - P_{u,t}(i,j)|^m \to 0,$$

which is equivalent to $|P_{u,s}(i,j) - P_{u,t}(i,j)|^m \to 0$. It is the same as

$$\lim_{s \to t} |P_{u,s}(i,j) - P_{u,t}(i,j)| = 0.$$

□

Although the bottom-up approach does not require that the correlation ρ_t must be continuous in time, it is implicitly assumed in the top-down approach. Intuitively, when time proceeds continuously, the transition probability matrix $\mathbf{P}_{t,s}$ also changes continuously in time as $\mathbf{P}_{t,s}$ is a solution of differential equations. The correlations between the components cannot adjust substantially as time changes at the infinitesimal level. If we require the joint distributions obtained from the bottom-up approach and the top-down approach coincide, the correlation ρ_t as a function of time must behave smoothly.

Proposition 3.3.1. *Suppose that* $(\Lambda_t, t \in [0,T])$ *is obtained by the bottom-up approach. If for all* $t \in [0,T]$*, any* $x_t^1, x_t^2 \in \{0,1\}$*, the one-dimensional distribution of* X *is consistent between the bottom-up approach and the top-down approach approach,*

$$\mathbb{P}\left(X_t^1 = x_t^1,\ X_t^2 = x_t^2; \rho_t\right) = \sum_{\left(x_0^1, x_0^2\right) \in \mathcal{K}^2} \mathbb{P}\left(X_0 = \left(x_0^1, x_0^2\right)\right) \mathbf{P}_{0,t}^{\left(x_0^1, x_0^2\right), \left(x_t^1, x_t^2\right)}, \tag{3.9}$$

then $\lim_{s \to t} |\rho_s - \rho_t| = 0$ *for any* $0 \le t \le s$.

Proof. Since ρ_t satisfies (3.4), the probability $\mathbb{P}\left(X_t^1 = x_t^1,\ X_t^2 = x_t^2; \rho_t\right)$ is well-defined. Assume that (3.9) holds for any $t \in [0,T]$. It suffices to consider for $x_t^1 = 1$ and $x_t^2 = 1$,

$$\mathbb{P}\left(X_t^1 = 1,\ X_t^2 = 1; \rho_t\right) = p_t q_t + \rho_t \sqrt{p_t(1-p_t) q_t (1-q_t)},$$

which is equivalent to

$$\rho_t = \frac{\sum_{\left(x_0^1, x_0^2\right) \in \mathcal{K}^2} \mathbb{P}\left(X_0 = (x_0^1, x_0^2)\right) \mathbf{P}_{0,t}^{\left(x_0^1, x_0^2\right), (1,1)} - p_t q_t}{\sqrt{p_t(1-p_t) q_t (1-q_t)}}.$$

For simplicity, we define $h_t := \sqrt{p_t(1-p_t) q_t (1-q_t)}$. Let $x_0 = (x_0^1, x_0^2)$. Thus,

$$|\rho_s - \rho_t|$$

$$
\begin{aligned}
&= \left| \frac{\sum_{x_0\in\mathcal{K}^2} \mathbb{P}\left(X_0 = x_0\right) \mathbf{P}_{0,s}^{x_0,(1,1)} - p_s q_s}{\sqrt{p_s(1-p_s)q_s(1-q_s)}} - \frac{\sum_{x_0\in\mathcal{K}^2} \mathbb{P}\left(X_0 = x_0\right) \mathbf{P}_{0,t}^{x_0,(1,1)} - p_t q_t}{\sqrt{p_t(1-p_t)q_t(1-q_t)}} \right| \\
&\leq \frac{1}{h_t h_s} \left| \sum_{x_0\in\mathcal{K}^2} \mathbb{P}\left(X_0 = x_0\right) \left(h_t \mathbf{P}_{0,s}^{x_0,(1,1)} - h_s \mathbf{P}_{0,t}^{x_0,(1,1)} \right) \right| + \frac{1}{h_t h_s} \left| h_s p_t q_t - h_t p_s q_s \right| \\
&= \frac{\sum_{x_0\in\mathcal{K}^2} \mathbb{P}\left(X_0 = x_0\right)}{h_t h_s} \left| h_t \mathbf{P}_{0,s}^{x_0,(1,1)} - h_t \mathbf{P}_{0,t}^{x_0,(1,1)} + h_t \mathbf{P}_{0,t}^{x_0,(1,1)} - h_s \mathbf{P}_{0,t}^{x_0,(1,1)} \right| \\
&\quad + \frac{1}{h_t h_s} \left| h_s p_t q_t - h_s p_s q_s + h_s p_s q_s - h_t p_s q_s \right| \\
&\leq \frac{\sum_{x_0\in\mathcal{K}^2} \mathbb{P}\left(X_0 = x_0\right)}{h_t h_s} \left(h_t \left| \mathbf{P}_{0,s}^{x_0,(1,1)} - \mathbf{P}_{0,t}^{x_0,(1,1)} \right| + \mathbf{P}_{0,t}^{x_0,(1,1)} \left| h_t - h_s \right| \right) \\
&\quad + \frac{1}{h_t h_s} \left(h_s \left| p_t q_t - p_s q_s \right| + p_s q_s \left| h_s - h_t \right| \right).
\end{aligned}
$$

By Lemma 3.3.1, $\mathbf{P}_{0,t}^{x_0,(1,1)}$ is continuous in the second time-argument. Since p_t, q_t and h_t are continuous in time, we conclude $\lim_{s\to t} |\rho_s - \rho_t| = 0$. □

However, the converse implication of Proposition 3.3.1 is not true in general. In the following, we give one sufficient condition under which the one-dimensional distribution of X between the top-down approach and the bottom-up approach will be consistent forward in time.

Proposition 3.3.2. *Suppose that $(\Lambda_t, t \in [0,T])$ is obtained by the bottom-up approach. If for any $t \in [0,T]$, $x_t^1, x_t^2 \in \{0,1\}$, we have*

$$
\mathbb{P}\left(X_t^1 = x_t^1,\ X_t^2 = x_t^2; \rho_t\right) = \sum_{\left(x_0^1, x_0^2\right)\in\mathcal{K}^2} \mathbb{P}\left(X_0 = \left(x_0^1, x_0^2\right)\right) \mathbf{P}_{0,t}^{\left(x_0^1,x_0^2\right),\left(x_t^1,x_t^2\right)}, \tag{3.10}
$$

and for any $0 \leq t \leq s$ it holds that

$$
\Theta_t^i \mathbf{P}_{t,s} = \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i, \quad i = 1,2, \tag{3.11}
$$

then we have for any $0 \leq t \leq s$, $x_s^1, x_s^2 \in \{0,1\}$,

$$
\mathbb{P}\left(X_s^1 = x_s^1,\ X_s^2 = x_s^2; \rho_s\right) = \sum_{\left(x_0^1, x_0^2\right)\in\mathcal{K}^2} \mathbb{P}\left(X_0 = \left(x_0^1, x_0^2\right)\right) \mathbf{P}_{0,s}^{\left(x_0^1,x_0^2\right),\left(x_s^1,x_s^2\right)}.
$$

Proof. We fix $i \in \{1,2\}$. For any $x_t^i \in \mathcal{K}$ and $x_s \in \mathcal{K}^2$, it holds that

$$
\begin{aligned}
&\Theta_t^i \mathbf{P}_{t,s} \mathbb{1}_{\{(x_s^1,x_s^2)\}}\left(x_t^i\right) \\
&= \sum_{\left(x_t^1,x_t^2\right)\in\mathcal{K}^2} \frac{\mathbb{P}\left(X_t = \left(x_t^1,x_t^2\right);\rho_t\right)}{\mathbb{P}\left(X_t^i = x_t^i\right)} \mathbf{P}_{t,s}^{\left(x_t^1,x_t^2\right),\left(x_s^1,x_s^2\right)} \\
&= \sum_{\left(x_t^1,x_t^2\right)\in\mathcal{K}^2} \frac{\sum_{\left(x_0^1,x_0^2\right)\in\mathcal{K}^2} \mathbb{P}\left(X_0 = \left(x_0^1,x_0^2\right)\right) \mathbf{P}_{0,t}^{\left(x_0^1,x_0^2\right),\left(x_t^1,x_t^2\right)}}{\mathbb{P}\left(X_t^i = x_t^i\right)} \mathbf{P}_{t,s}^{\left(x_t^1,x_t^2\right),\left(x_s^1,x_s^2\right)} \\
&= \frac{1}{\mathbb{P}\left(X_t^i = x_t^i\right)} \sum_{\left(x_0^1,x_0^2\right)\in\mathcal{K}^2} \mathbb{P}\left(X_0 = \left(x_0^1,x_0^2\right)\right) \left(\sum_{\left(x_t^1,x_t^2\right)\in\mathcal{K}^2} \mathbf{P}_{0,t}^{\left(x_0^1,x_0^2\right),\left(x_t^1,x_t^2\right)} \mathbf{P}_{t,s}^{\left(x_t^1,x_t^2\right),\left(x_s^1,x_s^2\right)} \right) \\
&= \frac{1}{\mathbb{P}\left(X_t^i = x_t^i\right)} \sum_{x_0\in\mathcal{K}^2} \mathbb{P}\left(X_0 = \left(x_0^1,x_0^2\right)\right) \mathbf{P}_{0,t}^{\left(x_0^1,x_0^2\right),\left(x_s^1,x_s^2\right)},
\end{aligned}
$$

where the second equality comes from (3.10). For any $x_t^i \in \mathcal{K}$ and $x_s = (x_s^1, x_s^2) \in \mathcal{K}^2$, we have

$$
\begin{aligned}
\widehat{\mathbf{P}}_{t,s}^i \Theta_s^i \mathbb{1}_{\{x_s\}}\left(x_t^i\right) &= \Theta_t^i \mathbf{P}_{t,s} \Phi^i \Theta_s^i \mathbb{1}_{\{x_s\}}\left(x_t^i\right) \\
&= \sum_{x_s^i\in\mathcal{K}} \left(\sum_{x_t\in\mathcal{K}^2} \frac{\sum_{x_0\in\mathcal{K}^2} \mathbb{P}\left(X_0 = x_0\right) \mathbf{P}_{0,t}^{x_0,x_t}}{\mathbb{P}\left(X_t^i = x_t^i\right)} \mathbf{P}_{t,s}^{x_t,x_s} \Phi^i \frac{\mathbb{P}\left(X_s = x_s;\rho_s\right)}{\mathbb{P}\left(X_s^i = x_s^i\right)} \right) \\
&= \frac{1}{\mathbb{P}\left(X_t^i = x_t^i\right)} \mathbb{P}\left(X_s = x_s;\rho_s\right).
\end{aligned}
$$

By (3.11) and $\mathbb{P}\left(X_t^i = x_t^i\right) > 0$, we have for any $x_s = (x_s^1, x_s^2) \in \mathcal{K}^2$,

$$
\mathbb{P}\left(X_s = \left(x_s^1,x_s^2\right);\rho_s\right) = \sum_{\left(x_0^1,x_0^2\right)\in\mathcal{K}^2} \mathbb{P}\left(X_0 = \left(x_0^1,x_0^2\right)\right) \mathbf{P}_{0,s}^{\left(x_0^1,x_0^2\right),\left(x_s^1,x_s^2\right)}.
$$

□

In light of next proposition, we can construct a Markov structure forward in time with a given initial distribution.

Proposition 3.3.3. *Let $i \in \{1,2,\ldots,m\}$. If for any $0 \le t \le u \le s$, it holds that*

$$
\Theta_t^i \mathbf{P}_{t,u} = \widehat{\mathbf{P}}_{t,u}^i \Theta_u^i, \qquad \text{and} \quad \Theta_u^i \mathbf{P}_{u,s} = \widehat{\mathbf{P}}_{u,s}^i \Theta_s^i,
$$

then

$$\Theta_t^i \mathbf{P}_{t,s} = \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i.$$

Proof. Since $\mathbf{P}_{t,s}$ and $\widehat{\mathbf{P}}_{t,s}^i$, $0 \le t \le s$ are semigroups, we know

$$\Theta_t^i \mathbf{P}_{t,s} = \Theta_t^i \mathbf{P}_{t,u} \mathbf{P}_{u,s} = \widehat{\mathbf{P}}_{t,u}^i \Theta_u^i \mathbf{P}_{u,s} = \widehat{\mathbf{P}}_{t,u}^i \widehat{\mathbf{P}}_{u,s}^i \Theta_s^i = \widehat{\mathbf{P}}_{t,s}^i \Theta_s^i.$$

□

In the bottom-up approach, we start with a given initial distribution and want to construct a Markov structure forward in time. In view of Proposition 3.3.2 and Proposition 3.3.3, if for all $0 \le t \le s$, (3.11) is satisfied *continuously* and *forward*, then for any $t \in [0, T]$ the one-dimensional distribution of X obtained by the bottom-up approach will agree with the one obtained by the top-down approach. On the other hands, if we do not impose (3.11), then we have to find $\rho_s,\ s \ge t$ such that

(i) $\lim_{s \to t} |\rho_s - \rho_t| = 0$;

(ii) $\Lambda_s^i = \Theta_s^i \Lambda_s \Phi^i,\ s \ge 0$;

(iii) the consistency property of one-dimensional distributions of X at time s holds.

However, under this circumstance it needs more computations in the bottom-up approach than the top-down approach.

3.4 Relation to Rogers and Pitman [RP81]

In this section, we connect our results to [RP81] and [BJN13, Example 3.2]. In [RP81], Rogers and Pitman provide one sufficient condition [RP81, Theorem 2] for (3.12) (see page 101), such that the function of a time-homogeneous Markov process is again a time-homogeneous Markov process. In [BJN13, Example 3.2], Bielecki et

al. provide an example showing that the coordinate process of a bivariate Markov chain is a time-inhomogeneous Markov chain.

As we will show, when we apply Rogers and Pitman's conditions to the setup of [BJN13, Example 3.2], i.e., a bivariate Markov chain with coordinate projection, the resulting coordinate process is time-homogeneous. The objectives of this section are to discuss the inconsistency of homogeneity of the component processes, as well as the sufficient conditions provided by Rogers and Pitman, and to compare their results to [BJN13, Example 3.2], and to our Example 2.6.1, Example 2.6.2, and Example 2.6.3.

We start by summarizing the relevant contents in [RP81].

3.4.1 Relevant Results in Rogers and Pitman [RP81, Theorem 2]. Assume that $\mathcal{E}$ is the σ-algebra of a state space E, and $(E, \mathcal{E})$ is a measurable state space. Let $(E', \mathcal{E}')$ be another measurable space. Suppose that $X = (X_t, t \geq 0)$ is an E-valued continuous time Markov process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with the initial distribution μ^X and the transition semigroup $\left(\widetilde{P}_t, t \geq 0\right)$. Let ϕ be a measurable transformation, $\phi : E \to E'$. Rogers and Pitman studied the question of when $\phi \circ X$ is a Markov process.

Let $\widetilde{\Lambda}$ be a Markov kernel from E' to E,

$$\widetilde{\Lambda} : (y, A) \to \widetilde{\Lambda}\,(y, A)\,, \quad y \in E',\ A \in \mathcal{E}.$$

such that for every $y \in E'$, $\widetilde{\Lambda}\,(y, \cdot)$ is a probability on E, and for each $A \in \mathcal{E}$, $\widetilde{\Lambda}\,(\cdot, A) \in b\mathcal{E}'$, where $b\mathcal{E}'$ is the space of bounded measurable functions on $\mathcal{E}'$. One can view the kernel $\widetilde{\Lambda}$ as an operator mapping $f \in b\mathcal{E}$ to $\widetilde{\Lambda} f \in b\mathcal{E}'$, where $\widetilde{\Lambda} f(y)$ is the $\widetilde{\Lambda}(y, \cdot)$ integral of f.

Rogers and Pitman gave sufficient conditions for each $t \geq 0$, $A \in \mathcal{E}$,

$$\mathbb{P}\left(X_t \in A \mid \phi \circ X_u,\ 0 \leq u \leq t\right) = \widetilde{\Lambda}\left(\phi \circ X_t, A\right) \quad \text{a.s.,} \tag{3.12}$$

then $\phi \circ X_t$ is Markov with transition kernels $\widetilde{Q}_t$ defined by

$$\widetilde{Q}_t f = \widetilde{\Lambda}\widetilde{P}_t \left(f \circ \phi\right), \quad f \in b\mathcal{E}'.$$

That is, using composition of kernels,

$$\widetilde{Q}_t = \widetilde{\Lambda}\widetilde{P}_t\Phi,$$

where Φ is the Markov kernel from $\mathcal{E}$ to $\mathcal{E}'$ which is induced by ϕ according to the formula

$$\Phi f = f \circ \phi, \quad f \in b\mathcal{E}'.$$

Next theorem gives conditions under which $\phi \circ X$ is a time-homogeneous Markov process.

Theorem 3.4.1 ([RP81, Theorem 2])**.** *Suppose there is a Markov kernel $\widetilde{\Lambda}$ from $\mathcal{E}'$ to $\mathcal{E}$ such that*

(i) $\widetilde{\Lambda}\Phi = \mathrm{I}$*, the identity kernel on* $\mathcal{E}'$*,*

(ii) for each $t \geq 0$ the Markov kernel $\widetilde{Q}_t = \widetilde{\Lambda}\widetilde{P}_t\Phi$ from $\mathcal{E}'$ to $\mathcal{E}'$ satisfies the identity $\widetilde{\Lambda}\widetilde{P}_t = \widetilde{Q}_t\widetilde{\Lambda}$.

Let X be Markov with semigroup $\left(\widetilde{P}_t, t \geq 0\right)$ and initial distribution $\mu = \widetilde{\Lambda}(y, \cdot)$, where $y \in \mathcal{E}'$. Then (3.12) *holds, and $\phi \circ X$ is Markov with starting state y and transition semigroup $\left(\widetilde{Q}_t, t \geq 0\right)$.*

Next, we apply Theorem 3.4.1 to the bivariate Markov chain $X = (X^1, X^2)$, which takes values in $E = E' \times E'$, where $E' = \{0, 1\}$. Because the goal is to make both components X^1 and X^2 be Markov in their own filtrations respectively, Theorem 3.4.1 is required to hold true for both components X^i, $i = 1, 2$. We will

add superscript i to the notations in Theorem 3.4.1 to distinguish conditions for each component.

Now, we consider ϕ^i as the coordinate projection, $\phi^i(X_t) = X_t^i,\ i = 1, 2$. The corresponding $\Phi^i,\ i = 1, 2$, is of the form,

$$\Phi^1 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \end{array}, \qquad \Phi^2 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}.$$

We take the Markov kernels $\widetilde{\Lambda}^i,\ i = 1, 2$, of the form,

$$\widetilde{\Lambda}^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \widetilde{\Lambda}^{1;0,(0,0)} & \widetilde{\Lambda}^{1;0,(0,1)} & 0 & 0 \\ 0 & 0 & \widetilde{\Lambda}^{1;1,(1,0)} & \widetilde{\Lambda}^{1;1,(1,1)} \end{pmatrix} \end{array},$$

and

$$\widetilde{\Lambda}^2 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \widetilde{\Lambda}^{2;0,(0,0)} & 0 & \widetilde{\Lambda}^{2;0,(0,1)} & 0 \\ 0 & \widetilde{\Lambda}^{2;1,(1,0)} & 0 & \widetilde{\Lambda}^{2;1,(1,1)} \end{pmatrix} \end{array},$$

where

$$0 \le \widetilde{\Lambda}^{i;x^i,y} \le 1,$$
$$\sum_{y \in E} \widetilde{\Lambda}^{i;x^i,y} = 1, \quad x^i \in E',\ y \in E,\ i = 1, 2.$$

Then condition (i) is satisfied for both components, $\widetilde{\Lambda}^i \Phi^i = \mathbf{I},\ i = 1, 2$, where $\mathbf{I}$ is an identity matrix.

Remark 3.4.1. *If we order the state space in the same order and ϕ^i is the coordinate projection, there is no alternative form for $\Phi^i,\ i = 1, 2$. Once Φ^i is chosen, the*

form of the corresponding $\widetilde{\Lambda}^i,\ i=1,2,$ *is unique to make condition* (i) *satisfied, i.e.,* $\widetilde{\Lambda}^i\Phi^i=\mathbf{I}$.

Remark 3.4.2. *In view of Theorem 3.4.1 and the form of* $\widetilde{\Lambda}^i,\ i=1,2,$ *if we ask that both components are Markov in their filtrations, the initial distribution of* X *must be either* $\mathbb{P}(X_0=(0,0))=1$ *or* $\mathbb{P}(X_0=(1,1))=1$.

In order to have condition (ii), we need to solve for the transition semigroup $\widetilde{P}_t=\left[\widetilde{P}_t^{x,y}\right]_{x,y\in E}=\left[\widetilde{P}_t(j,k)\right]_{j,k=1,2,\ldots,|E|}$, which satisfies the following linear systems simultaneously,

$$\widetilde{\Lambda}^i\widetilde{P}_t=\widetilde{Q}_t^i\widetilde{\Lambda}^i,\quad i=1,2, \tag{3.13}$$

where

$$\widetilde{Q}_t^i=\widetilde{\Lambda}^i\widetilde{P}_t\Phi^i. \tag{3.14}$$

More precisely, by replacing $\widetilde{Q}_t^i$ from (3.14) to (3.13) the subsequent systems must hold true,

$$\widetilde{\Lambda}^i\widetilde{P}_t=\widetilde{\Lambda}^i\widetilde{P}_t\Phi^i\widetilde{\Lambda}^i,\quad i=1,2. \tag{3.15}$$

In view of (3.15), both sides are $|E'|$-by-$|E|$ dimensional matrices. Comparing entry-wise of the matrices in both sides, there are at most $|E'|\times(|E|-|E'|)\,(=4)$ equalities for each i. However, there are total 20 $(=4+4+12)$ unknowns, namely, $|E'|\times|E'|\,(=4)$ from $\widetilde{\Lambda}^i,\ i=1,2,$ and $|E|^2-|E|\,(=12)$ from the common term $\widetilde{P}_t$. Because the number of unknowns are greater than the number of equalities, generally the solution set of the linear systems is non-empty.

In view of (3.15), if $i = 1$ we have the following,

$$\begin{aligned}
&\left(\widetilde{P}_t(1,1) + \widetilde{P}_t(1,2) - \widetilde{P}_t(2,1) - \widetilde{P}_t(2,2)\right)\alpha_1^2 \\
&\qquad\qquad - \left(\widetilde{P}_t(1,1) - 2\widetilde{P}_t(2,1) - \widetilde{P}_t(2,2)\right)\alpha_1 - \widetilde{P}_t(2,1) = 0 \\
&\left(\widetilde{P}_t(3,3) + \widetilde{P}_t(3,4) - \widetilde{P}_t(4,3) - \widetilde{P}_t(4,4)\right)\alpha_2^2 \\
&\qquad\qquad - \left(\widetilde{P}_t(3,3) - 2\widetilde{P}_t(4,3) - \widetilde{P}_t(4,4)\right)\alpha_2 - \widetilde{P}_t(4,3) = 0 \\
&(1-\alpha_1)\left(\alpha_2\widetilde{P}_t(3,1) + (1-\alpha_2)\,\widetilde{P}_t(4,1)\right) - \alpha_1\left(\alpha_2\widetilde{P}_t(3,2) + (1-\alpha_2)\,\widetilde{P}_t(4,2)\right) = 0 \\
&(1-\alpha_2)\left(\alpha_1\widetilde{P}_t(1,3) + (1-\alpha_1)\,\widetilde{P}_t(2,3)\right) - \alpha_2\left(\alpha_1\widetilde{P}_t(1,4) + (1-\alpha_1)\,\widetilde{P}_t(2,4)\right) = 0,
\end{aligned} \tag{3.16}$$

where

$$\alpha_1 := \widetilde{\Lambda}^{1;0,(0,0)}, \qquad\qquad \alpha_2 := \widetilde{\Lambda}^{1;1,(1,0)}.$$

If $i = 2$, the system is given by

$$\begin{aligned}
&\left(\widetilde{P}_t(1,1) + \widetilde{P}_t(1,3) - \widetilde{P}_t(3,1) - \widetilde{P}_t(3,3)\right)\beta_1^2 \\
&\qquad\qquad - \left(\widetilde{P}_t(1,1) - 2\widetilde{P}_t(3,1) - \widetilde{P}_t(3,3)\right)\beta_1 - \widetilde{P}_t(3,1) = 0 \\
&\left(\widetilde{P}_t(2,2) + \widetilde{P}_t(2,4) - \widetilde{P}_t(4,2) - \widetilde{P}_t(4,4)\right)\beta_2^2 \\
&\qquad\qquad - \left(\widetilde{P}_t(2,2) - 2\widetilde{P}_t(4,2) - \widetilde{P}_t(4,4)\right)\beta_2 - \widetilde{P}_t(4,2) = 0 \\
&(1-\beta_1)\left(\beta_2\widetilde{P}_t(2,1) + (1-\beta_2)\,\widetilde{P}_t(4,1)\right) - \beta_1\left(\beta_2\widetilde{P}_t(2,3) + (1-\beta_2)\,\widetilde{P}_t(4,3)\right) = 0 \\
&(1-\beta_2)\left(\beta_1\widetilde{P}_t(1,2) + (1-\beta_1)\,\widetilde{P}_t(3,2)\right) - \beta_2\left(\beta_1\widetilde{P}_t(1,4) + (1-\beta_1)\,\widetilde{P}_t(3,4)\right) = 0,
\end{aligned} \tag{3.17}$$

where

$$\beta_1 := \widetilde{\Lambda}^{2;0,(0,0)}, \qquad\qquad \beta_2 := \widetilde{\Lambda}^{2;1,(0,1)}.$$

Since we ask both components of X to be Markovian in their filtrations, the semigroup $\widetilde{P}_t$ needs to satisfy the system (3.16) and the system (3.17) simultaneously

with appropriate variables $\alpha_1,\ \alpha_2,\ \beta_1,\ \beta_2$. Note that all these quadratic equations in the systems (3.16) and (3.17) have the properties that when the variable is 0, the function has non-positive value; when the variable is 1, the function has non-negative value. For example,

$$\begin{aligned} \alpha_1 = 0: &\quad -\widetilde{P}_t(2,1) \le 0 \\ \alpha_1 = 1: &\quad \widetilde{P}_t(1,2) \ge 0, \end{aligned}$$

which is equivalent to $-\widetilde{P}_t(2,1) \cdot \widetilde{P}_t(1,2) \le 0$. By intermediate value theorem each quadratic equation has only root in the interval $(0,1)$.

In the sequel, we restrict ourselves to the generator of X of the form,

$$\Lambda = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a+b+c) & a & b & c \\ 0 & -d & 0 & d \\ 0 & 0 & -f & f \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \tag{3.18}$$

where $a, b, c, d, f \ge 0$. The analytic solution of $\widetilde{P}_t$ can be obtained by solving the Kolmogorov forward equation,

$$\frac{\partial}{\partial t}\widetilde{P}_t = \widetilde{P}_t\Lambda, \quad \widetilde{P}_t = \mathbf{I}_{|E|\times|E|},$$

$$\widetilde{P}_t = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \mathrm{e}^{-(a+b+c)t} & \widetilde{P}_t(1,2) & \widetilde{P}_t(1,3) & \widetilde{P}_t(1,4) \\ 0 & \mathrm{e}^{-dt} & 0 & 1-\mathrm{e}^{-dt} \\ 0 & 0 & \mathrm{e}^{-ft} & 1-\mathrm{e}^{-ft} \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array},$$

where

$$\widetilde{P}_t(1,2) = \frac{a}{a+b+c-d}\left(\mathrm{e}^{-dt} - \mathrm{e}^{-(a+b+c)t}\right)$$

$$\widetilde{P}_t(1,3) = \frac{b}{a+b+c-d}\left(\mathrm{e}^{-ft} - \mathrm{e}^{-(a+b+c)t}\right)$$
$$\widetilde{P}_t(1,4) = 1 - \frac{1}{a+b+c-d}\left(a\mathrm{e}^{-dt} + b\mathrm{e}^{-ft} + (c-d)\mathrm{e}^{-(a+b+c)t}\right)$$

To satisfy (3.15), the *only* Markov kernels $\widetilde{\Lambda}^i,\ i = 1, 2$, are given by

$$\widetilde{\Lambda}^1 = \begin{array}{c c} & \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}, \tag{3.19}$$

and

$$\widetilde{\Lambda}^2 = \begin{array}{c c} & \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

Then, we have the semigroup of each component $X^i,\ i = 1, 2$,

$$\widetilde{Q}^1_t = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} \mathrm{e}^{-dt} & 1-\mathrm{e}^{-dt} \\ 0 & 1 \end{pmatrix} \end{array}, \qquad \widetilde{Q}^2_t = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} \mathrm{e}^{-ft} & 1-\mathrm{e}^{-ft} \\ 0 & 1 \end{pmatrix} \end{array},$$

and recover the generator of each component,

$$\Lambda^1 = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} -d & d \\ 0 & 0 \end{pmatrix} \end{array}, \qquad \Lambda^2 = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} -f & f \\ 0 & 0 \end{pmatrix} \end{array},$$

which are time-homogeneous. Furthermore, in view of $\widetilde{\Lambda}^1$ and $\widetilde{\Lambda}^2$, we know the initial distribution of X is $\mathbb{P}(X_0 = (1,1)) = 1$. That is, both components start at the absorbing state 1 with probability 1. Clearly, this bivariate Markov chain will stay in state $(1,1)$ forever, and it behaves like a constant process. Accordingly, its component stays in state 1 and will not change. Thus, the components are surely Markov in their own filtrations, respectively.

Theorem 3.4.1 provides sufficient conditions to verify if $\phi^i(X)$ is Markov in its filtration. If we confine ourselves to a class of generators like (3.18), the bivariate Markov chains solved by Rogers and Pitman's conditions are not helpful in our situations, where we want the bivariate Markov chain to start at non-absorbing state.

3.4.2 Discussions. Let us start by recalling the following example from [BJN13].

Example 3.4.1 (Example 3.2 in [BJN13])**.** *Let the infinitesimal generators of X^1 and X^2 be given by*

$$\Lambda_t^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -(b+c)+\alpha_t & (b+c)-\alpha_t \\ 0 & 0 \end{pmatrix} \end{array}, \tag{3.20}$$

$$\Lambda_t^2 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -(a+c)+\beta_t & (a+c)-\beta_t \\ 0 & 0 \end{pmatrix} \end{array}, \tag{3.21}$$

where with $a, b, c > 0$,

$$\alpha_t = \frac{\frac{ac}{a+c}\mathrm{e}^{-bt}\left(1-\mathrm{e}^{-(a+c)t}\right)}{\mathrm{e}^{-(a+b+c)t} + \frac{a}{a+c}\mathrm{e}^{-bt}\left(1-\mathrm{e}^{-(a+c)t}\right)},$$
$$\beta_t = \frac{\frac{bc}{b+c}\mathrm{e}^{-at}\left(1-\mathrm{e}^{-(b+c)t}\right)}{\mathrm{e}^{-(a+b+c)t} + \frac{b}{b+c}\mathrm{e}^{-at}\left(1-\mathrm{e}^{-(b+c)t}\right)}.$$

It is shown that a valid generator Λ of X given by

$$\Lambda = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a+b+c) & a & b & c \\ 0 & -b & 0 & b \\ 0 & 0 & -a & a \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \tag{3.22}$$

is a solution to the analytic necessary condition of the weak Markovian consistency,

$$\Lambda_t^i = Q_t^i \Lambda \Phi^i, \quad i = 1, 2.$$

Since Λ_t^i depends on time t, the component X^i is time-inhomogeneous. It is implied in this example that the process X starts with the initial distribution $\mathbb{P}(X_0 = (0,0)) = 1$.

The generator Λ in (3.18) is a generalization of [BJN13, Example 3.2], and a special case of Example 2.6.3 on page 74. In Example 2.6.3 we have shown that as long as the bivariate Markov chain starts with the initial distribution $\mathbb{P}(X_0 = (0,0)) = 1$, both components of X are Markovian in their filtrations. Whereas in Example 2.6.3, we do not compute the generator of each component explicitly. We state the generator is given by

$$\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i, \quad t \geq 0, \ i = 1, 2.$$

Even Λ_t does not depend on t, Λ_t^i still depends on t due to Θ_t^i.

[BJN13, Example 3.2] assumes $\mathbb{P}(X_0 = (0,0)) = 1$. Bielecki et al. give analytical formulations (3.20) and (3.21) for Λ_t^i, $i = 1, 2$, and verify that each component is Markovian in its own filtration with the given infinitesimal generator. In view of (3.20) and (3.21), Λ_t^i, $i = 1, 2$, clearly depends on t. Thus both components of X are time-inhomogeneous.

Given generator Λ by (3.18), the bivariate Markov chain obtained by [RP81, Theorem 2], however, has initial distribution $\mathbb{P}(X_0 = (1,1)) = 1$. Besides, both components are time-homogeneous.

There are two layers of information from these examples. First, given the time-homogeneous generator Λ with the structure in (3.22), we have *at least* two bivariate Markov chains whose every component is a Markov chain. One bivariate Markov chain starts at state $(1,1)$ with probability 1, and another bivariate Markov chain begins with the initial distribution $\mathbb{P}(X_0 = (0,0)) = 1$. Because the initial distributions are not identical, it is likely the resulting components are different. Thus, the generators Λ_t^i, $i = 1, 2$, for these two Markov chains will not be the same. We use these examples

to emphasize how the initial distributions of X lead to distinct components of X.

Second, given the same initial distribution $\mathbb{P}(X_0 = (0,0)) = 1$ and two different generators in (3.18) and in (3.22)[7], we have two different bivariate Markov chains. Both Markov chains have time-inhomogeneous Markovian components whose generators are given by $\Theta_t^i \Lambda_t \Phi^i,\ t \geq 0,\ i = 1, 2$, respectively. These generators Λ_t^i's are not necessarily the same.

In view of Theorem 2.5.2 on page 56, our condition is formulated in terms of the operator $\Theta_u^i,\ u \geq 0$. Whereas in [RP81, Theorem 2], their condition is given by the initial distribution $\widetilde{\Lambda}^i$. Now, if we restrict our Example 2.6.2 to a time-homogeneous generator Λ, then we recover the results (3.19) and thereinafter.

3.5 Conclusions and Future Work

How to construct Markov structures efficiently is one of the main subjects in Markov structures theory. We present an example to show an alternative method to construct Markov structures. Unlike the top-down approach, the bottom-up approach establishes the one-dimensional distributions of X by the joint distributions of it components at every time $t \geq 0$. As shown, the continuity of the correlation series is necessary for the consistency in the one-dimensional distributions. By the bottom-up approach we need sufficient conditions to have the the consistent one-dimensional distributions. We show that the condition in Theorem 2.5.2 regarding Markovian consistency theory is also a sufficient condition to preserve consistent one-dimensional distributions.

Although the bottom-up approach is not as intuitive as the top-down approach and does not produce promising results, we use this example to draw attentions to the difference between the Markov structures theory and the classical copulae and

[7]We take (3.22) as a special case of Example 2.6.3, then the results follow.

Sklar Theorem. In the classical copulae, it concerns about the finite dimensional multivariate random variables preserving the marginal distributions. Whereas, in the Markov structures theory we study the components with the prescribed laws as stochastic processes. Markov structures theory certainly incorporates the dependence structures that classical copulae do not include.

Future work

The core issue of Markov structures is to find the solutions for Markov structures effectively. The future work is summarized below.

- To design an efficient algorithm to solve for the solutions of Markov structures

  Given prescribed marginal laws, in order to solver for the solutions of Markov structures, one needs to solve a system of linear equations, that is related to the functional of the generator. In general, the analytical solution of the semigroup is unavailable. Solving for Markov structures is a non-trivial research problem. An efficient method to identify all feasible solutions is therefore a necessity and a topic to work on.

- To study the sufficient conditions for the bottom-up approach

  In the bottom-up approach, we showed that the continuity of the correlation between the one-dimensional distributions in time is necessary for that, the consistency property in one-dimensional distributions will be satisfied in the future. We need to impose additional conditions to have the consistency property. We believe that the bottom-up approach is more practical than the top-down approach. The future work is to investigate the necessary and sufficient conditions in terms of correlations. If this question can be answered, one may extend the binomial case to the multinomial distribution.

- To develop other approaches to solve for the solutions of Markov structures

We proposed the bottom-up approach to solver for the Markov structures. In this approach, we start from a given initial distribution and want to solve the linear equations forward in time. If we already know that some initial distribution guarantees a solution of Markov structures, this approach will be useful to identify the solutions. However, if one has limited knowledge about the initial distribution, this approach may not be useful at all. Therefore, it remains an open question to develop a method to solve for the solutions of Markov structures.

CHAPTER 4

APPLICATIONS TO SYSTEMIC RISK

4.1 Introduction

In this chapter, we apply the Markov structures framework to study systemic risk. In particular, we propose a novel approach to model and to measure systemic risk.

We consider a financial system consisting of a finite number of financial institutions. In general, each financial institution has a corporate credit rating to reflect its current financial stability. For example, the Moody's global long-term rating scale is $\{Aaa, Aa, A, Baa, Ba, B, Caa, Ca, C\}$. A financial institution rated C is in default. Next, we introduce a joint credit migration process whose components are the credit ratings of financial institutions in this financial system. This joint credit migration process contains full information about the interconnections of its components. In our setup, we model the joint migration process as a Markov structure. Each individual rating migration process is therefore a Markov chain itself. The laws of these individual migration processes are identical to the prescribed laws which are obtained from the data provided by the rating agencies. Thus in our approach, the joint migration process encodes the evolution of interdependencies between financial stability of the individual institutions.

Since financial institutions are connected and receiving the same financial services from common markets or intermediaries, the migration of financial stability for one financial institution has a direct or an indirect impact on the other financial institutions. Generally speaking, the phenomenon in which a shock from one financial institution is transmitted to another financial institution is called *domino effect* or

contagion[8]. In our framework, this transmission mechanism can be introduced by imposing conditions on the joint credit migration process. Saying differently, in order to present the contagious phenomenon, we consider weak-only Markov structures to model the joint credit migration process. Thus, the credit migration of one financial institution is allowed to affect the credit migrations of other financial institutions. Accordingly, we construct two measures to quantify systemic risk and the stochastic dependence between the financial institutions within this financial system.

The outline of this chapter is as follows. In Section 4.2, preliminaries are set up, and the concepts of independence and dependence in terms of infinitesimal generator of the joint credit migration process are introduced. In Section 4.3, we provide an example to explain why we consider weak-only Markov structure, and how this structure relates to the contagion of financial shocks. In Section 4.4, we define the systemic risk measure and systemic dependence measure. We explain the financial meanings behind these measures. Section 4.5 is the numerical analysis to our systemic dependence and systemic risk measures. We conclude this chapter and discuss future work in Section 4.6.

4.2 Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete underlying probability space. We consider a financial system consisting of m financial institutions. Without loss of generality, we categorize the corporate credit rating scales to a finite state space $\mathcal{K} = \{0, 1, \ldots, K\}$. By convention, the state K is the default state[9]. Suppose that $Y^i,\ i = 1, 2, \ldots, m$, are Markov chains taking values in $\mathcal{K}$. Each Y^i represents the evolution of credit

[8]In literature, there are serious debates on the definitions of *contagion* and *spillovers*. These definitions are model-dependent. Here, instead of giving precise definitions, we want to introduce the concept of financial chain reactions.

[9]We do not postulate that the default state is necessarily an absorbing state.

ratings of the ith financial institution. Recall that $\mathcal{M} = \{1, 2, \ldots, m\}$. Assume that for every $i \in \mathcal{M}$, Y^i is generated by the infinitesimal generator matrix function $\Lambda_t^i = [\lambda_t^{i;x^i,y^i}]_{x^i,y^i \in \mathcal{K}}$, $t \geq 0$. The function $\lambda_t^{i;x^i,y^i}$ measures at time t how quickly the transition from state x^i to state y^i happens.

We model the joint credit migration process as a weak-only Markov structure. Now, let a continuous time multivariate Markov chain $X = (X^1, X^2, \ldots, X^m)$ take values in $\mathcal{K}^m$ with initial distribution μ^X. We want to construct the infinitesimal generator matrix function $\Lambda_t = [\lambda_t^{x,y}]_{x,y \in \mathcal{K}^m}$, $t \geq 0$ of X, such that X is weak-only Markovian consistent. That is, each component X^i of X is Markovian. Necessarily X^i has the same law as Y^i,

$$\Lambda_t^i = \Theta_t^i \Lambda_t \Phi^i, \quad t \geq 0, \ i \in \mathcal{M}.$$

Note that Θ_t^i depends on the initial distribution of X. Then, we solve for $(\Lambda_t, t \geq 0)$ such that the above conditions are satisfied.

As mentioned in Chapter 3, there always exists one $(\Lambda_t, t \geq 0)$ satisfying the conditions of Markov structures. In general, such a solution $(\Lambda_t, t \geq 0)$ is not unique. Each $(\Lambda_t, t \geq 0)$ has a distinct structure and features its own interdependencies between the components of X. In particular, if we have *sufficient* market information or financial products, we may calibrate one solution $(\Lambda_t, t \geq 0)$ that resembles the current interdependencies between credit ratings of financial institutions. When more than one solution $(\Lambda_t, t \geq 0)$ exists, there is a potential for nonzero systemic risk caused by the dependence structures. We create a relevant systemic risk measure to assess this risk.

In what follows, we explain independence and dependence between the components of X in terms of the infinitesimal generator matrix function of X.

Independence and Dependence Between the Components of X

Before we define independence in terms of the generator structures, we would like to introduce two notations. Let I be the identity of dimension $|\mathcal{K}|$. Recall that m is a finite natural number. The mth tensor power (or the Kronecker product) of I is the m-fold tensor product of I,

$$\mathrm{I}^{\otimes m} := \underbrace{\mathrm{I} \otimes \cdots \otimes \mathrm{I}}_{m}.$$

The notation $\mathrm{I}_A^{\otimes m,j}$ is reserved for replacing the jth matrix I by some matrix A whose dimension is the same as I,

$$\mathrm{I}_A^{\otimes m,j} := \underbrace{\mathrm{I} \otimes \cdots \otimes \overbrace{A}^{j\text{th}} \otimes \cdots \otimes \mathrm{I}}_{m}.$$

The important results in [BJN, Lemma 3.1 and Theorem 3.2] show that, if the semigroup of X is the tensor product of the semigroups from the components, then the components of X are conditional independent given $\mathcal{F}_T$. We apply these results to our setup, and by the fact that condition (P) is equivalent to condition (M). We can show that if the infinitesimal generator of X is the tensor product of the infinitesimal generators from the components, then the components of X are independent. Thus, we now formally give the definition of independence between the components of X by means of infinitesimal generator of X.

Definition 4.2.1 (Independence)**.** *Let $(\Lambda_t, t \geq 0)$ be a valid infinitesimal generator matrix function of X. We say that the components of X are independent if Λ_t admits the representation,*

$$\Lambda_t = \sum_{j=1}^{m} \mathrm{I}_{A_t^j}^{\otimes m,j} = \sum_{j=1}^{m} \underbrace{\mathrm{I} \otimes \cdots \otimes \overbrace{A_t^j}^{jth} \otimes \cdots \otimes \mathrm{I}}_{m}, \quad t \geq 0, \tag{4.1}$$

where every A_t^j, $j = 1, 2, \ldots, m$ is a valid generator matrix function with the same dimension as I.

Note that, an infinitesimal generator given by (4.1) satisfies condition (M). Indeed, we can use

$$\Lambda_t \Phi^i, \quad t \geq 0, \ i = 1, 2, \ldots, m,$$

to examine this property. Thus, the corresponding multivariate Markov chain is strongly Markovian consistent.

4.3 Contagion

In this section, we give an example to explain why we consider a joint credit migration process as a *weak-only* Markov structure. Meanwhile, we use this example to give intuitions about the meaning of contagion in our framework, and to elucidate how the weak-only Markovian consistency relates to contagion.

Example 4.3.1. *Consider a joint credit migration process $X = (X^1, X^2)$ generated by $\Lambda_t, \ t \geq 0$*

$$\Lambda_t = \left[\lambda_t^{x,x'}\right]_{x,x' \in \{0,1\}^2} = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a_t + c_t + d_t) & d_t & a_t & \boldsymbol{c_t} \\ f_t & -(a_t + f_t) & 0 & a_t \\ b_t & 0 & -(b_t + d_t) & d_t \\ 0 & b_t & f_t & -(b_t + f_t) \end{pmatrix} \end{array}, \tag{4.2}$$

where $a_t, b_t, c_t, d_t, f_t \geq 0$.

It should be noted that in general this process X is not Markovian consistent. However, under some additional assumptions on functions a, b, c, d, f, the process X is strongly Markovian consistent, weakly Markovian consistent, or weak-only Markovian consistent. For such examples see for instance Chapter 2.

Next we will argue that the strongly Markovian consistency excludes contagion between the components of X.

At time t, the infinitesimal rate of transitions for the component X^1 from state 0 to state 1 is

$$\begin{aligned} \lambda_t^{(0,0),(1,0)} + \lambda_t^{(0,0),(1,1)} &= a_t + c_t, \qquad \text{if } X_t^2 = 0, \\ \lambda_t^{(0,1),(1,0)} + \lambda_t^{(0,1),(1,1)} &= a_t, \qquad \text{if } X_t^2 = 1. \end{aligned} \tag{4.3}$$

If $c_t = 0$, then both results are the same. Thus, the state of X_t^2 is irrelevant. Similarly, at time t the rates of jumps for X^1 from state 1 to state 0 are identical,

$$\begin{aligned} \lambda_t^{(1,0),(0,0)} + \lambda_t^{(1,0),(0,1)} &= b_t, \qquad \text{if } X_t^2 = 0, \\ \lambda_t^{(1,1),(0,0)} + \lambda_t^{(1,1),(0,1)} &= b_t, \qquad \text{if } X_t^2 = 1. \end{aligned}$$

Thus, the infinitesimal rates of jumps for X^1 from any state to the other states do not depend on the state of X_t^2.

Likewise, if $c_t = 0$, the state of X_t^1 does not influence the transition rate of X^2 from state 0 to state 1,

$$\begin{aligned} \lambda_t^{(0,0),(0,1)} + \lambda_t^{(0,0),(1,1)} &= d_t + c_t, \qquad \text{if } X_t^1 = 0, \\ \lambda_t^{(1,0),(0,1)} + \lambda_t^{(1,0),(1,1)} &= d_t, \qquad \text{if } X_t^1 = 1, \end{aligned} \tag{4.4}$$

and the transition rate of X^2 from state 1 to state 0,

$$\begin{aligned} \lambda_t^{(0,1),(0,0)} + \lambda_t^{(0,0),(1,0)} &= f_t, \qquad \text{if } X_t^1 = 0, \\ \lambda_t^{(1,1),(0,0)} + \lambda_t^{(1,0),(1,0)} &= f_t, \qquad \text{if } X_t^1 = 1. \end{aligned}$$

We say that there is no contagion between X^1 and X^2. In fact, condition (M) is satisfied, thus the process X is strongly Markovian consistent. This is the reason we rule out the structures with strong Markovian consistency.

Instead, if $c_t > 0$, given the state of X_t^2 the rates of jumps for X^1 from state 0 to state 1 in (4.3) are different. The state of X_t^2 will have an impact on how X^1 jumps from state 0 to state 1. Similarly, the state of X_t^1 affects the rate of transitions for X^2 from state 0 to state 1 in (4.4). Then we say there exists contagion between X^1 and X^2.

Remark 4.3.1. *Note that in general the common jump factor with zero value, for example $c_t = 0$, does not imply that the components cannot jump simultaneously, or the probability of the common jump is zero. The probability of the common jump depends on the structure of Λ_t. We need to check the corresponding transition matrix. Actually, given functions $a, b, d, f > 0$ and $c = 0$ in (4.2) the probabilities of all common jumps are nonzero.*

4.4 Systemic Risk Measure and Systemic Dependence Measure

Recall that $\{Y^1, \ldots, Y^m\}$ is a collection of Markov chains with generators $\Lambda_u^i,\ u \geq 0$. We construct a valid infinitesimal generator $\big(\Lambda_u^{\mathcal{I}}, u \geq 0\big)$ by

$$\Lambda_u^{\mathcal{I}} = \sum_{j=1}^{m} \mathrm{I}_{\Lambda_u^j}^{\otimes m,j}, \quad u \geq 0.$$

In view of the Kolmogorov Existence Theorem, on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ there exists a unique multivariate Markov chain $X = (X^1, \ldots, X^m)$ generated by $\big(\Lambda_u^{\mathcal{I}}, u \geq 0\big)$. If we take the initial distribution μ^{X^i} identical to μ^{Y^i}, then each X^i has the same law as Y^i. Therefore, we know from Section 4.2.1 that a joint credit migration process modeled by $\big(\Lambda_u^{\mathcal{I}}, u \geq 0\big)$ has independent components. We say this joint credit migration process has independent structure. We denote by $\mathcal{I}$ the independence structure and call this financial system with fully independent financial institutions as the *neutral* financial system.

Now, we denote by $z \in \mathcal{K}$ the state corresponding to some credit rating. Let z, $0 \leq t \leq T < \infty$, and $h \in \mathcal{M}$ be fixed. We want to compute the conditional probability that, at future time T, at least h financial institutions being in credit rating z conditional on the information available at time t. That is,

$$\mathbb{P}\left(\sum_{j=1}^{m} \mathbb{1}_{\{X_T^{\mathcal{I},j}=z\}} \geq h \;\middle|\; X_t^{\mathcal{I}}\right), \quad z \in \mathcal{K},\ h \in \mathcal{M},\ 0 \leq t \leq T. \tag{4.5}$$

Note that the constructed generator $\big(\Lambda_u^{\mathcal{I}}, u \geq 0\big)$ satisfies condition (M), thus X is

strongly Markovian consistent. We want to emphasize that there is no contagion for this joint credit migration process with independent individual credit evolutions.

Since we model the joint credit migration process X as a weak-only Markov structure for $\{Y^1, \ldots, Y^m\}$, the components of X are Markovian and respectively generated by

$$\Lambda_u^i = \Theta_u^i \Lambda_u \Phi^i, \quad u \geq 0, \ i \in \mathcal{M}.$$

If a nontrivial solution $\left(\Lambda_u^{\mathcal{D}}, u \geq 0\right)$ exists, that is a solution different from $\left(\Lambda_u^{\mathcal{I}}, u \geq 0\right)$, it defines a dependence structure, say $\mathcal{D}$. Accordingly, we compute the conditional probability

$$\mathbb{P}\left(\sum_{j=1}^{m} \mathbb{1}_{\left\{X_T^{\mathcal{D},j}=z\right\}} \geq h \;\middle|\; X_t^{\mathcal{D}}\right), \quad z \in \mathcal{K}, \ h \in \mathcal{M}, \ 0 \leq t \leq T, \tag{4.6}$$

where $X^{\mathcal{D}}$ is generated by $\left(\Lambda_u^{\mathcal{D}}, u \geq 0\right)$. Since $X^{\mathcal{D}}$ is weak-only Markovian consistent, the structure $\mathcal{D}$ is different from structure $\mathcal{I}$. In particular, this joint credit migration process models contagion between the individual credit evolutions.

Before we introduce the following definitions, let $\mathcal{X}$ denote the collection of Markov chains with values in $\mathcal{K}^m$. We give a definition for systemic risk measure.

Definition 4.4.1 (Systemic Risk Measure)**.** *Let $z = (z_1, z_2, \ldots, z_m) \in \mathcal{K}^m$ be fixed. We define a function $\nu^z : [0, \infty) \times [0, T] \times \mathcal{M} \times \mathcal{X} \times \mathcal{K}^m \to [0, 1]$,*

$$\nu^z\left(T, t, h, X^{\mathcal{D}}, x\right) := \mathbb{P}\left(\sum_{j=1}^{m} \mathbb{1}_{\left\{X_T^{\mathcal{D},j}=z_j\right\}} \geq h \;\middle|\; X_t^{\mathcal{D}} = x\right),$$

In particular, if every z_i is the corresponding default rating, we call this function ***the systemic risk measure****.*

When each z_i is the corresponding default rating, we define the systemic dependence measure as the difference between (4.5) and (4.6). Saying differently, we normalize the dependence structure with respect to the independence structure.

Definition 4.4.2 (Systemic Dependence Measure)**.** *Suppose that* $\{Y^1, Y^2, \ldots, Y^m\}$ *is a collection of Markov chains, and* $X^{\mathcal{D}}$ *is a Markov structure for* $\{Y^1, Y^2, \ldots, Y^m\}$*. Assume that* $(\Lambda_u, u \geq 0)$ *is the generator of* $X^{\mathcal{D}}$*. Let* $z = (z_1, z_2, \ldots, z_m) \in \mathcal{K}^m$ *be fixed. We define a function* $\rho^z : [0, \infty) \times [0, T] \times \mathcal{M} \times \mathcal{X} \times \mathcal{K}^m \to [-1, 1]$*,*

$$\rho^z \left(T, t, h, X^{\mathcal{D}}, x\right) := \nu^z \left(T, t, h, X^{\mathcal{D}}, x\right) - \nu^z \left(T, t, h, X^{\mathcal{I}}, x\right). \tag{4.7}$$

In particular, if every z_i *is the corresponding default rating, we call this function* ***the systemic dependence measure****.*

We argue that using only systemic risk measure to monitor systemic risk is not enough. As we will see, different dependence structures might contribute to the same level of systemic risk measure. Consequently, we propose to use the pair $\left(\nu^z \left(T, t, h, X^{\mathcal{D}}, x\right), \rho^z \left(T, t, h, X^{\mathcal{D}}, x\right)\right)$, that takes into account the dependence structure, as a measure to monitor systemic risk.

From now we take K to denote the default rating and we will compute (ν^z, ρ^z) for $z = (K, \ldots, K)$.

The systemic dependence measure depends on T, t, h and $\mathcal{D}$. Given T, t, h and $\mathcal{D}$, we measure the difference of two conditional probabilities of a certain event, that is related to the total number of components X^i's in the default state at some future time T, based on the information available at current time t. It should be noted that the systemic dependence measure is evaluated based on different dependence structures. We compare the conditional probability of this event under a joint credit migration process *with* contagious mechanism to the conditional probability of the same event under a joint credit migration process *without* contagion.

Financially speaking, when we fix $h = m$ and T, the event of interest is all financial institutions in default at some future time T. This financial system will be in *Armageddon* at some future time T. We call this circumstance *systemic risk*. It

follows that the systemic dependence measure depends on t, X_t. In view of (4.7), at time t and $X_t^{\mathcal{D}} = x$, if the systemic dependence measure $\rho^z\left(T, t, h, X^{\mathcal{D}}, x\right) > 0$, it means that the probability of all financial institutions in default rating at future time T under a joint credit migration process allowing contagion is strictly larger than the probability of the same event under the process with independent components. In other words, a financial system with transmission mechanism has greater exposure to the risk in Armageddon. Under this dependence structure, the financial system is no better than the neutral financial system. We say for the pair (t, x), this financial system with dependence structure $\mathcal{D}$ is of *unfavorable systemic dependence.*

At time t and $X_t^{\mathcal{D}} = x$, if the systemic dependence measure $\rho^z\left(T, t, h, X^{\mathcal{D}}, x\right) = 0$, it means that the financial system with contagion mechanism behaves like the neutral financial system. Then we call this financial system with dependence structure $\mathcal{D}$ a *neutral systemic dependence* for the pair (t, x). By analogy, if at time t and $X_t^{\mathcal{D}} = x$ the systemic dependence measure $\rho^z\left(T, t, h, X^{\mathcal{D}}, x\right) < 0$, it suggests that the financial system with transmission mechanism has less exposure to the risk of simultaneously default at future time T. This financial system benefits from the current dependence structure. We say that for the pair (t, x) this financial system with dependence structure $\mathcal{D}$ has *favorable systemic dependence.*

For simplicity we take $t = 0$ and $X_0^{\mathcal{D}} = x$. That is, we evaluate the potential of the *initial systemic dependence* for the financial system. Now, the systemic dependence measure of interest is a function of T, m and X. The systemic dependence measure is given by

$$\rho^z\left(T, 0, m, X^{\mathcal{D}}, x\right) = \mathbb{P}\left(\sum_{j=1}^{m} \mathbb{1}_{\left\{X_T^{\mathcal{D},j}=K\right\}} \geq m \;\middle|\; X_0^{\mathcal{D}} = x\right) - \mathbb{P}\left(\sum_{j=1}^{m} \mathbb{1}_{\left\{X_T^{\mathcal{I},j}=K\right\}} \geq m \;\middle|\; X_0^{\mathcal{I}} = x\right).$$

According to the sign of this systemic dependence measure, we categorize the depen-

dence status of the financial system to the following:

$$\rho^z\left(T,0,m,X^{\mathcal{D}},x\right)\begin{cases}>0, & \text{unfavorable systemic dependence;}\\ =0, & \text{neutral systemic dependence;}\\ <0, & \text{favorable systemic dependence.}\end{cases}$$

Remark 4.4.1. *In general, the weak-only Markov structure for* $\{Y^1,\ldots,Y^m\}$ *is not unique. Namely, in general, there may be more than one solution* $(\Lambda_u, u\geq 0)$ *of*

$$\Lambda_u^i=\Theta_u^i\Lambda_u\Phi^i,\quad u\geq 0,\ i\in\mathcal{M},$$

which produces a weak-only Markov structure. Any such solution $(\Lambda_u, u\geq 0)$ *features unique dependence between the components of the joint credit migration process. These solutions may be found numerically and are used to simulate the evolution of* X *that illustrates the dependence between its components. A practical goal is to calibrate* $(\Lambda_u, u\geq 0)$ *with the data coming from the financial system. One approach is using the prices of the exchange-traded basket products to calibrate, for instance, the Markit CMBX index references a basket of 25 commercial mortgage-backed securities.*

We end this section with providing an important property of our systemic dependence measure.

Theorem 4.4.1 (Law invariance)**.** *Assume that* $X^{\mathcal{D}}, X^{\mathcal{D}'}$ *are Markov structures for* $\{Y^1,Y^2,\ldots,Y^m\}$*. If* $X^{\mathcal{D}}, X^{\mathcal{D}'}$ *have the same law with respect to* $\mathbb{P}$*, then we have*

$$\rho^z\left(T,t,h,X^{\mathcal{D}},x\right)=\rho^z\left(T,t,h,X^{\mathcal{D}'},x\right),\quad z,x\in\mathcal{K}^m,\ 0\leq t\leq T,\ h\in\mathcal{M}.$$

Proof. Note that $X^{\mathcal{D}}, X^{\mathcal{D}'}$ are multivariate Markov chains. We denote by $\Lambda_u^{\mathcal{D}}$ and $\Lambda_u^{\mathcal{D}'},\ u\geq 0$ the infinitesimal generators of $X^{\mathcal{D}}$ and $X^{\mathcal{D}'}$, respectively. Markov chains $X^{\mathcal{D}}, X^{\mathcal{D}'}$ have the same finite-dimensional distribution if and only if $\Lambda_u^{\mathcal{D}}=\Lambda_u^{\mathcal{D}'},\ u\geq 0$. Besides, $\Lambda_u^{\mathcal{D}}=\Lambda_u^{\mathcal{D}'},\ u\geq 0$ if and only if $\mathbf{P}_{t,s}^{X^{\mathcal{D}}}=\mathbf{P}_{t,s}^{X^{\mathcal{D}'}},\ 0\leq t\leq s$. Then we have

$$\Lambda_u^i=\Theta_u^i\Lambda_u^{\mathcal{D}}\Phi^i=\Theta_u^i\Lambda_u^{\mathcal{D}'}\Phi^i,\quad u\geq 0,\ i\in\mathcal{M}.$$

It follows that the corresponding generator of independent structure is identical,

$$\sum_{j=1}^{m} \mathrm{I}_{\Lambda_u^j}^{\otimes m,j} = \sum_{j=1}^{m} \underbrace{\mathrm{I} \otimes \cdots \otimes \overbrace{\Lambda_u^j}^{j\text{th}} \otimes \cdots \otimes \mathrm{I}}_{m}, \quad u \geq 0.$$

For any $z \in \mathcal{K},\ 0 \leq t \leq T,\ h \in \mathcal{M}$, we have

$$\begin{aligned}
&\rho^z\left(T, t, h, X^{\mathcal{D}}, x\right) \\
&= \mathbb{P}\left(\sum_{j=1}^{m} \mathbb{1}_{\left\{X_T^{\mathcal{D},j}=z_j\right\}} \geq h \;\middle|\; X_t^{\mathcal{D}} = x\right) - \mathbb{P}\left(\sum_{j=1}^{m} \mathbb{1}_{\left\{X_T^{\mathcal{I},j}=z_j\right\}} \geq h \;\middle|\; X_t^{\mathcal{I}} = x\right) \\
&= \mathbb{P}\left(\sum_{j=1}^{m} \mathbb{1}_{\left\{X_T^{\mathcal{D}',j}=z_j\right\}} \geq h \;\middle|\; X_t^{\mathcal{D}'} = x\right) - \mathbb{P}\left(\sum_{j=1}^{m} \mathbb{1}_{\left\{X_T^{\mathcal{I},j}=z_j\right\}} \geq h \;\middle|\; X_t^{\mathcal{I}} = x\right) \\
&= \rho^z\left(T, t, h, X^{\mathcal{D}'}, x\right),
\end{aligned}$$

where the second equality comes from the fact that the semigroups of $X^{\mathcal{D}}, X^{\mathcal{D}'}$ are the same. □

4.5 Numerical Results

In this section, we study the numerical analysis on the systemic risk measure and systemic dependence measure.

In the case of Example 2.6.1 on page 68 and Example 2.6.3 on page 74, we have full knowledge about how the initial distribution of a bivariate Markov chain and the algebraic structures of infinitesimal generator $(\Lambda_u, u \geq 0)$ of the chain determine different types of Markovian consistency. Therefore, we will follow up on these examples here, and thus we consider a financial system with two financial institutions. We denote by Y^i the credit rating process of the ith financial institution, $i = 1, 2$. Instead of taking credit ratings as the state space of Y^i, we assume that each financial institution is in either non-default or default state. State 0 represents the non-default state, and state 1 stands for the default state. Moreover, for simplicity, we suppose that once the financial institution defaults, it cannot return to the

non-default state. That means that state 1 is the absorbing state for every Y^i. We further assume that at time 0 each Y^i starts from non-default state with probability 1, $\mathbb{P}\left(Y_0^i = 0\right) = 1,\ i = 1, 2$.

Next, let $0 \leq v_1 \leq v_2 \leq v_3 < \infty$. We consider time intervals: $[0, v_1)$, $[v_1, v_2)$, $[v_2, v_3)$, and $[v_3, \infty)$. Assume that $X^{\mathcal{D}}$ is a bivariate Markov chain with initial distribution $\mathbb{P}\left(X_0 = (0,0)\right) = 1$ and piecewise constant infinitesimal generator $\left(\Lambda_u^{\mathcal{D}}, u \geq 0\right)$,

$$\Lambda_u^{\mathcal{D}} = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a_u + b_u + c_u) & a_u & b_u & c_u \\ 0 & -d_u & 0 & d_u \\ 0 & 0 & -f_u & f_u \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \tag{4.8}$$

where for any $0 \leq u < \infty$, $a_u, b_u, c_u, d_u, f_u \geq 0$, and a_u, b_u, c_u, d_u, f_u are piecewise constant.

Note that, the procedure to construct Markov structures is to take the generator $(\Lambda_u^i, u \geq 0)$ of Y^i, $i = 1, 2$, as initial inputs. Then one constructs Markov structures corresponding to $\{Y^1, Y^2\}$. In view of (4.8), Example 2.6.1, and Example 2.6.3, if, additionally, we have

$$a_u + c_u \neq f_u \qquad\qquad b_u + c_u \neq d_u,$$

then $X^{\mathcal{D}}$ is weak-only Markovian consistent. Each $X^{\mathcal{D},i}$ is therefore generated by

$$\Lambda_u^i = \Theta_u^i \Lambda_u^{\mathcal{D}} \Phi^i, \quad u \geq 0,\ i = 1, 2. \tag{4.9}$$

Subsequently, we construct the independent infinitesimal generator $\left(\Lambda_u^{\mathcal{I}}, u \geq 0\right)$ of X by

$$\Lambda_u^{\mathcal{I}} = \sum_{j=1}^{2} \mathrm{I}_{\Lambda_u^j}^{\otimes 2, j} = \Lambda_u^1 \otimes \mathrm{I} + \mathrm{I} \otimes \Lambda_u^2.$$

Specifications of Model Parameters

In what follows, we investigate the systemic risk measure and systemic dependence measure for different algebraic structures of $\left(\Lambda_u^{\mathcal{D}}, u \geq 0\right)$. Throughout these examples, we will consider the systemic dependence measure corresponding to $z = (1,1)$ as the default state 1 for both components. Also, we will take $h = 2$, and study at time $t = 0$ the systemic dependence measure as a function of the future time T and $X^{\mathcal{D}}$. Thus, the systemic dependence measure is given by

$$\begin{aligned}
&\rho^{(1,1)}\left(T, 0, 2, X^{\mathcal{D}}, (0,0)\right) \\
&= \mathbb{P}\left(\sum_{j=1}^{2} \mathbb{1}_{\left\{X_T^{\mathcal{D},j}=1\right\}} \geq 2 \;\middle|\; X_0^{\mathcal{D}} = (0,0)\right) - \mathbb{P}\left(\sum_{j=1}^{2} \mathbb{1}_{\left\{X_T^{\mathcal{I},j}=1\right\}} \geq 2 \;\middle|\; X_0^{\mathcal{I}} = (0,0)\right).
\end{aligned} \tag{4.10}$$

We will take T from 0 to 30 with step size $\Delta T = 0.2$, and compute a sequence of values of the systemic dependence measure with T changing. The reason we take $t = 0$ and let T vary is that, we can analyze the impacts resulting from changing in dependence structure in the near future and in the far future. We specify the model parameters in Table 4.1.

Table 4.1. Specifications of Model Parameters

Parameters	h	m	v_1	v_2	v_3
Value	2	2	3	10	30

In the following examples, we will analyze the properties of the corresponding systemic dependence measures given by different structures $\Lambda_u^{\mathcal{D}}$.

Example 4.5.1. *Let Y^1, Y^2 be Markov chains with generators*

$$\Lambda_u^i = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} -\lambda_u^i & \lambda_u^i \\ 0 & 0 \end{pmatrix} \end{array}, \quad u \geq 0, \ i = 1, 2,$$

respectively, where

$$\begin{aligned} \lambda_u^1 &= \frac{c_u \left(a + b + c_u\right) \mathrm{e}^{-\int_0^u (a+b+c_v)\,\mathrm{d}v} + ab\mathrm{e}^{-bu}}{a\mathrm{e}^{-bu} + c_u \mathrm{e}^{-\int_0^u (a+b+c_v)\,\mathrm{d}v}} \\ \lambda_u^2 &= \frac{c_u \left(a + b + c_u\right) \mathrm{e}^{-\int_0^u (a+b+c_v)\,\mathrm{d}v} + ab\mathrm{e}^{-au}}{b\mathrm{e}^{-au} + c_u \mathrm{e}^{-\int_0^u (a+b+c_v)\,\mathrm{d}v}}, \end{aligned} \tag{4.11}$$

and $a, b, c_u > 0$, and c_u is piecewise constant.

We construct an independence generator by λ_u^1 and λ_u^2,

$$\Lambda_u^{\mathcal{I}} = \Lambda_u^1 \otimes \mathrm{I} + \mathrm{I} \otimes \Lambda_u^2 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(\lambda_u^1 + \lambda_u^2) & \lambda_u^2 & \lambda_u^1 & 0 \\ 0 & -\lambda_u^1 & 0 & \lambda_u^1 \\ 0 & 0 & -\lambda_u^2 & \lambda_u^2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}. \tag{4.12}$$

The generator $\Lambda_u^{\mathcal{I}}$ satisfies condition (M). Now, we solve for a solution $\left(\Lambda_u^{\mathcal{D}}, u \geq 0\right)$ of

$$\Lambda_u^i = \Theta_u^i \Lambda_u^{\mathcal{D}} \Phi^i, \quad u \geq 0, \ i = 1, 2,$$

where $\left(\Lambda_u^{\mathcal{D}}, u \geq 0\right)$ is of the form

$$\Lambda_u^{\mathcal{D}} = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -(a + b + c_u) & a & b & c_u \\ 0 & -b & 0 & b \\ 0 & 0 & -a & a \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}. \tag{4.13}$$

It should be noted that since $c_u > 0$*, a bivariate Markov chain* $X^{\mathcal{D}} = \left(X^{\mathcal{D},1}, X^{\mathcal{D},2}\right)$ *generated by* (4.13) *is a weak-only Markov structure for* $\{Y^1, Y^2\}$*. In view of Remark 3.2.1 on page 87, if we take* $c_u = 0$ *for all* $u \geq 0$ *in* (4.13)*, then we have the independence structure. However, in such a case, the marginal laws are not given by* (4.11)*. The difference between* (4.13) *and the independence structure is the common jump parameter* c_u*. In our model, the parameter* c_u *is the key element that captures the simultaneous jumps between the components. In this example, we maninly study how the parameter* c_u *affects the systemic dependence measure.*

Next, we consider another generator $\Lambda_u^{\mathcal{S}}$ *of the form,*

$$\Lambda_u^{\mathcal{S}} = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -\left(\lambda_u^1 + \lambda_u^2 - \gamma_u\right) & \lambda_u^2 - \gamma_u & \lambda_u^1 - \gamma_u & \gamma_u \\ 0 & -\lambda_u^1 & 0 & \lambda_u^1 \\ 0 & 0 & -\lambda_u^2 & \lambda_u^2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \quad u \geq 0, \tag{4.14}$$

where $0 \leq \gamma_u \leq \min\left(\lambda_u^1, \lambda_u^2\right)$*. The generator* $\Lambda_u^{\mathcal{S}}$ *satisfies condition (M) as well, and a Markov chain* $X^{\mathcal{S}}$ *generated by* $\left(\Lambda_u^{\mathcal{S}}, u \geq 0\right)$ *is a strong Markov structure for* $\{Y^1, Y^2\}$*. Moreover, if* $\gamma_u = 0$*,* (4.14) *reduces to independence structure* (4.12)*.*

Note that by construction the i*th component of a bivariate chain generated respectively by either* $\Lambda_u, \Lambda_u^{\mathcal{I}}$ *or* $\Lambda_u^{\mathcal{S}}$ *has identical prescribed law* $\Lambda_u^i,\ u \geq 0$*.*

In the first part of the numerical analysis, we restrict ourselves to the structure (4.13) *and investigate how parameter* c_u *affects the systemic dependence measure. The parameters are given in Table 4.2. Two scenarios are considered:* c_u *is piecewise increasing;* c_u *is piecewise decreasing.*

In Figure 4.1, both curves represent the systemic dependence measures for different time horizon T*. We observe that when* c_u *changes between time intervals, the*

Table 4.2. Parameters for Example 4.5.1

Parameters	Time Periods			
	$[0,3)$	$[3,10)$	$[10,30)$	$[30,\infty)$
a	0.01	0.01	0.01	0.01
b	0.02	0.02	0.02	0.02
(1) c_u increasing	0.08	0.19	0.33	0.33
(2) c_u decreasing	0.08	0.03	0.003	0.003

slopes of the curves change as well. Moreover, with c_u larger, the curve approaches to 0 *faster. In view of* (4.11)*, when a, b, u are fixed, the transition rate λ_u^i is approximately linear in c_u. Thus the transition rate λ_u^i with increasing c_u will be larger than the transition rate λ_u^i with decreasing c_u. That means that the probability for the component jumping from state* 0 *to state* 1 *will be higher if c_u is increasing. Both components are tend to default. When c_u is piecewise decreasing, the general trend of the measure is decreasing. Whereas, the level of the measure does not drop immediately. There exists period-lagging before the level of the measure decreases.*

In the second part of the numerical analysis, we examine the systemic dependence measures for two different dependence structures (4.13) *and* (4.14)*. In addition to the parameters in Table 4.2, we take parameter $\gamma_u = \eta \min(\lambda_u^1, \lambda_u^2)$ with $\eta = 0, 0.1, 0.5, 1$. If $\eta = 0$,* (4.13) *becomes independence structure* (4.12)*. Figure 4.2 on page 131 shows the comparisons between c_u and λ_u^1, λ_u^2 for c_u piecewise increasing or decreasing. No matter c_u is piecewise increasing or piecewise decreasing, functions λ^1, λ^2 are piecewise decreasing.*

In Figure 4.3 on page 132 and Figure 4.5 on page 140, we compare the systemic dependence measure and systemic risk measure. Only the red curve with circle marker is the measure for weak-only Markov structure. The rest blue curves are all from

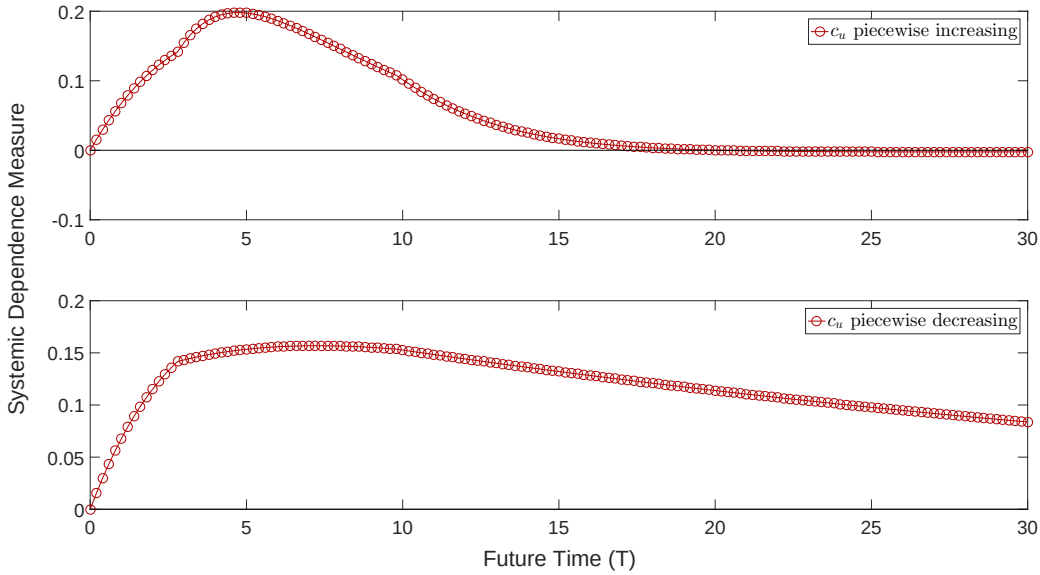

Figure 4.1. The Effect of Common Jump Parameter c_u for Example 4.5.1

strong Markov structures. It should be noted that when $\eta = 1$*, one of the components cannot jump individually. For instance, if* $\min(\lambda_u^1, \lambda_u^2) = \lambda_u^2$*, then the probability for* $X^{\mathcal{S},2}$ *jumping to state* 1 *individually is zero. In such a case, the level of systemic dependence measure is highest. We argue that, since condition (M) is satisfied, the state of* $X^{\mathcal{S},1}$ *will not have an influence on how* $X^{\mathcal{S},2}$ *changes its state. However, if* $X^{\mathcal{S},2}$ *changes to state* 1*,* $X^{\mathcal{S},1}$ *must jump to state* 1 *as well. Financially speaking, the corresponding financial institution* $X^{\mathcal{S},2}$ *will not default individually. However, if* $X^{\mathcal{S},2}$ *defaults, then* $X^{\mathcal{S},1}$ *defaults as well. The component* $X^{\mathcal{S},2}$ *contributes significantly to systemic risk. Our systemic dependence measure labels this circumstance with high systemic dependence. When we look at the pair* $\left(\nu^z\left(T,t,h,X^{\mathcal{D}},x\right), \rho^z\left(T,t,h,X^{\mathcal{D}},x\right)\right)$ *and* $\left(\nu^z\left(T,t,h,X^{\mathcal{S}},x\right), \rho^z\left(T,t,h,X^{\mathcal{S}},x\right)\right)$*, the systemic risk measure is changing the systemic dependence measure.*

In addition, we zoom in on the systemic dependence measure for the case of parameter c_u *piecewise increasing, see Figure 4.4 on page 133. When* T *closes to* 20*, the systemic dependence measure for the weak-only Markov structure changes from*

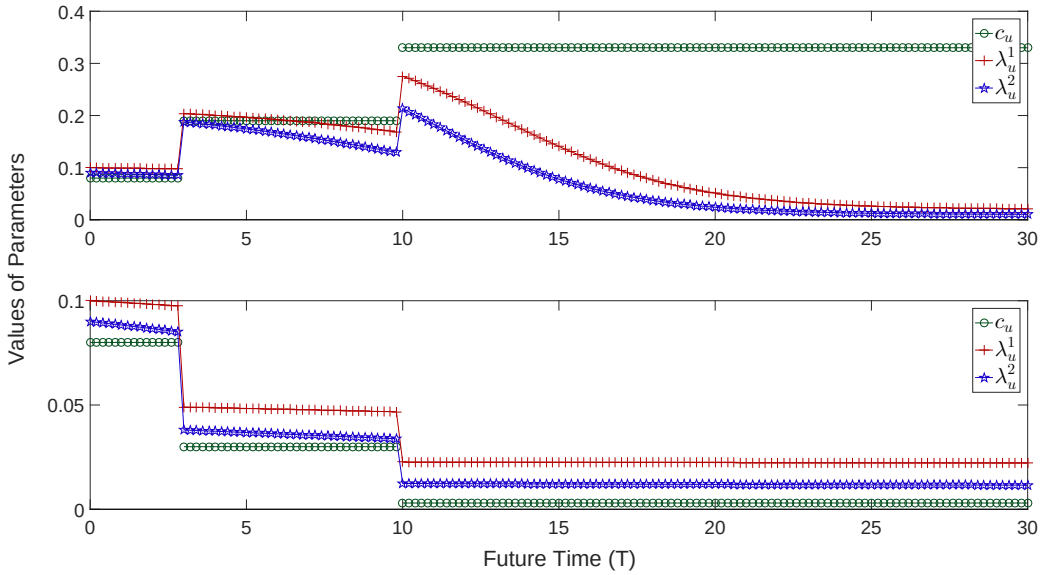

Figure 4.2. Results for Example 4.5.1: Comparisons Between $c_u, \lambda_u^1, \lambda_u^2$

unfavorable systemic dependence to favorable systemic dependence. It implies that in the long run the dependence structure allowing for contagion has advantages over the other dependence structures.

It should be noted that we assume the default state as an absorbing state. As T gets larger, the systemic risk measures approach 1 *regardless of the dependence structures. Moreover, the difference between the systemic risk measures are not significant. However, the corresponding systemic dependence measures behave differently between dependence structures because we normalize the systemic dependence measure. In particular, when c_u is piecewise increasing, the systemic dependence measures are more sensitive to the change of c_u in the near future. This result means that the dependence structure is more important in the near future than in the long run. It implies that if the policy maker decides to intervene in the event of systemic risk, the policy maker needs to step in fast. Otherwise, in the long run there is no significant difference between dependence structures.*

Also, we observe that the level of systemic dependence measure is higher when the parameter η gets larger. This is consistent with the discussion in the first part of numerical analysis.

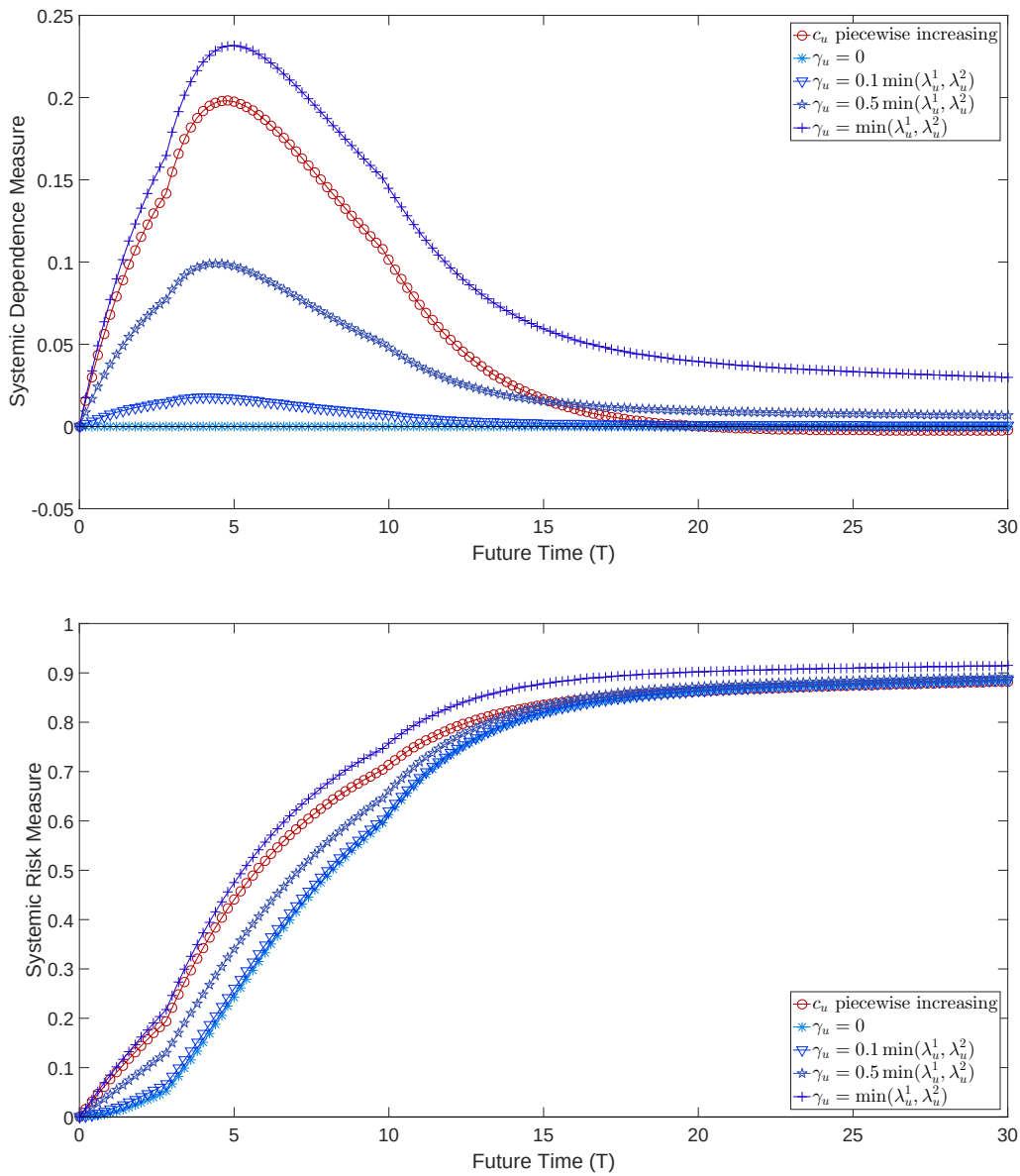

Figure 4.3. Results for Example 4.5.1: Parameter c_u Piecewise Increasing

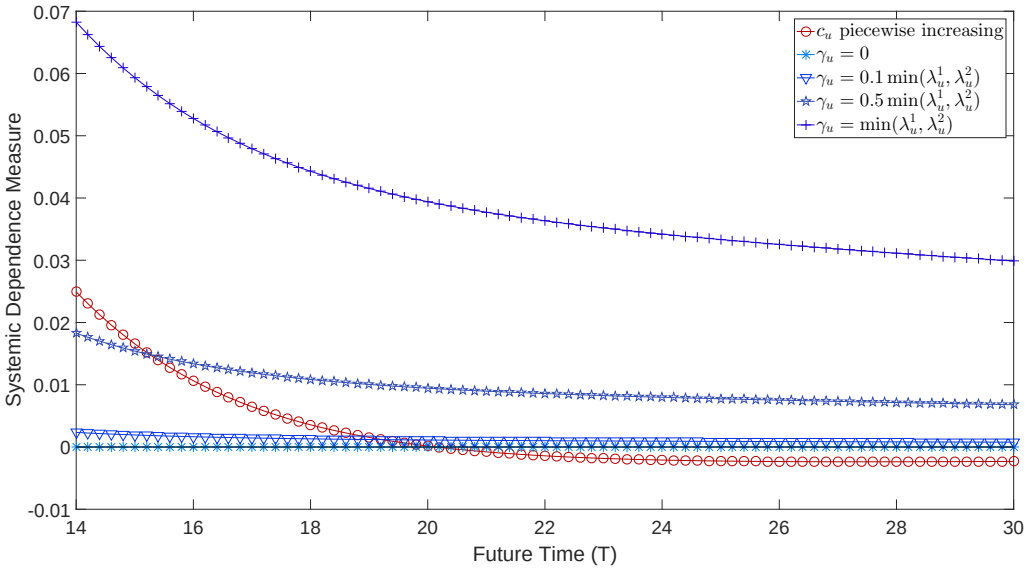

Figure 4.4. Zoom-in of the Systemic Dependence Measure in Figure 4.3

Example 4.5.2 (extreme contagion)**.** *Assume that Y^i is generated by Λ_u^i,*

$$\Lambda_u^i = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} -c_u & c_u \\ 0 & 0 \end{pmatrix} \end{array}, \quad u \geq 0, \ i = 1, 2.$$

We want to solve for the solution $\left(\Lambda_u^{\mathcal{D}}, u \geq 0\right)$ of the form,

$$\Lambda_u^{\mathcal{D}} = \begin{array}{c c} & \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{array}{c} (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} & \begin{pmatrix} -c_u & 0 & 0 & c_u \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \tag{4.15}$$

where $c_u > 0$ and piecewise constant. In view of (4.15), if the components jump, they must jump simultaneously. Note that in this structure, the transition probability for individual jump is zero. Hence, we call this structure as ***extreme contagion****.*

The parameters of this example are summarized in Table 4.3. We will investigate the general properties of the systemic risk measure and systemic dependence

measure for the structure (4.15)*, and then compare the results to the structure* (4.14)*.*

Table 4.3. Parameters for Example 4.5.2

Parameters	Time Periods			
	$[0,3)$	$[3,10)$	$[10,30)$	$[30,\infty)$
(1) c_u increasing	0.08	0.15	0.25	0.25
(2) c_u constant	0.08	0.08	0.08	0.08
(3) c_u decreasing	0.08	0.03	0.005	0.005

In Figure 4.6 on page 141, the curves represent the systemic dependence measure for time horizon T ranging from 0 *to* 30*. We study the properties of the systemic dependence measure within the class* (4.15)*. All the curves are hump-shaped and non-negative for three time periods. The curve with piecewise increasing c_u approaches* 0 *faster than the other curves.*

Next, we compare the systemic dependence measures between the dependence structures (4.14) *and* (4.15) *in Figure 4.7 on page 142 and Figure 4.8 on page 143. The curves with circle marker are obtained by structure* (4.15)*. The other curves are from structure* (4.14) *with $\gamma_u = \eta \min(\lambda_u^1, \lambda_u^2)$, $\eta = 0, 0.1, 0.5, 1$. An interesting observation is that the systemic dependence measure of extreme contagion coincides with the measure obtained by $\eta = 1$. The mathematical explanations are given below. First, Y^1 and Y^2 have the same law. For the structure* (4.15)*, the chain does not move if it is not in state* $(0,0)$*. For the structure* (4.14)*, condition (P) is satisfied and the transition probability of individual jumps from state* $(0,0)$ *is* 0*. Thus both structures produce the same measures.*

Example 4.5.3 (Extreme anti-contagion)**.** *Consider Markov chains Y^1, Y^2 with gen-*

erators

$$\Lambda_u^i = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} -\lambda_u^i & \lambda_u^i \\ 0 & 0 \end{pmatrix} \end{array}, \quad u \geq 0,\ i = 1, 2,$$

respectively, where

$$\lambda_u^1 = \frac{b_u\left(a_u + b_u\right)\mathrm{e}^{-\int_0^u (a_v+b_v)\,\mathrm{d}v}}{a_u + b_u \mathrm{e}^{-\int_0^u (a_v+b_v)\,\mathrm{d}v}}$$
$$\lambda_u^2 = \frac{a_u\left(a_u + b_u\right)\mathrm{e}^{-\int_0^u (a_v+b_v)\,\mathrm{d}v}}{b_u + a_u \mathrm{e}^{-\int_0^u (a_v+b_v)\,\mathrm{d}v}},$$

and $a_u, b_u > 0$ *and piecewise constant.*

We want to solve for a Markov structure $X^{\mathcal{D}}$ *for* $\{Y^1, Y^2\}$ *with generator* $\left(\Lambda_u^{\mathcal{D}}, u \geq 0\right)$ *of the form,*

$$\Lambda_u^{\mathcal{D}} = \begin{array}{c c} & \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{array}{c} (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} & \begin{pmatrix} -(a_u + b_u) & a_u & b_u & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}, \tag{4.16}$$

where $a_u, b_u > 0$ *and locally integrable on* $[0, \infty)$.

Unlike the structure (4.15) *of extreme contagion, in view of the structure* (4.16)*, each component only jumps individually. It is impossible for both components to jump simultaneously,*

$$\mathbb{P}\left(\sum_{j=1}^{2} \mathbb{1}_{\left\{X_T^{\mathcal{D},j}=1\right\}} \geq 2 \;\middle|\; X_0^{\mathcal{D}} = (0,0)\right) = 0, \quad T \in [0, \infty). \tag{4.17}$$

We call this structure as ***extreme anti-contagion****. The theoretical value of the systemic dependence measure is non-positive for any* $0 \leq T < \infty$*. The parameters for this example are given in Table 4.4.*

Table 4.4. Parameters for Example 4.5.3

Parameters	Time Periods			
	$[0,3)$	$[3,10)$	$[10,30)$	$[30,\infty)$
(1) (a_u, b_u) increasing	$(0.05, 0.05)$	$(0.08, 0.08)$	$(0.15, 0.15)$	$(0.15, 0.15)$
(2) (a_u, b_u) constant	$(0.05, 0.05)$	$(0.05, 0.05)$	$(0.05, 0.05)$	$(0.05, 0.05)$
(3) (a_u, b_u) decreasing	$(0.05, 0.05)$	$(0.03, 0.03)$	$(0.005, 0.005)$	$(0.005, 0.005)$

The parameters λ_u^1 and λ_u^2 have piecewise decreasing trends for different time horizon T, see Figure 4.9 on page 144.

In Figure 4.10 on page 144, every curve stands for the systemic dependence measure as a function of T. We observe that the systemic dependence measures are non-positive, and thus the financial system has initially favorable systemic dependence. Besides, the financial system shares more favorable systemic dependence when the values of a_u, b_u get larger.

Next, we are interested in the systemic dependence measure of the joint credit migration process with or without contagion. The comparisons of the systemic dependence measures and systemic risk measure for different dependence structures, see the structures (4.14) *on page 128 and* (4.16) *on page 135, are given in Figure 4.11 on page 145 and Figure 4.12 on page 146. In view of the structure* (4.14)*, the parameter γ_u needs to be bounded within the interval $[0, \min(\lambda_u^1, \lambda_u^2)]$ for all $u \geq 0$. Thus, the systemic dependence measure for structure* (4.14) *is non-negative and the measure attains* 0 *at $\gamma_u = 0$, i.e., the independence structure.*

4.6 Conclusions and Future Work

In this chapter, we proposed an original approach to study systemic risk. As we believe, the dependence between financial institutions is in the core of systemic

risk. We apply the framework of Markov structures to present the issue of dependence structures. In particular, we construct a dynamic systemic dependence measure that takes into account the dependence between the components of the financial system expressed in terms of Markov structures. We also define a dynamic systemic risk measure. We study the numerical analysis on different algebraic structures of infinitesimal generators, and compare these two measures between the weak-only Markov structure and the strong Markov structure. Also, we give financial explanations to the numerical results.

In the numerical analysis, we concentrate on the systemic risk measure and systemic dependence measure for several classes of algebraic structures of the generators. These structures have special meanings in the context of systemic risk. The initial investigations indeed show that different dependence structures contribute to various levels of systemic risk. Furthermore, our proposed systemic dependence measure is dynamic and flexible to deal with more general situations. For example, from a risk management perspective one may have interests in the case when the components are in different credit ratings, not just default state.

Another contribution of this chapter is that we construct a measure to investigate dependence structure between the coordinate processes of a multivariate Markov chain. Our measure assesses the degree to which stochastic processes strays from the independence structure. Although the properties of this measure are not yet fully established, we point out an interesting research direction: how to measure the dependence between stochastic processes with *prescribed* laws.

Future work

Our proposed systemic dependence measure is a function, that depends on many parameters. It is still an interesting topic to analytically study the mathematical

properties of this measure. We summarize the future work in the following.

- To investigate mathematical properties of the systemic dependence measure

  It remains an open question to investigate the properties of the systemic dependence measures in the context of coherent risk measures. The other directions might be: the analysis of the systemic dependence measures for the weak-only Markov structures, that preserve marginals, between different classes of algebraic structures; or the properties for a general state space, for instance, without absorbing state or more than two non-absorbing states.

- To design an efficient algorithm

  We deal with the transition matrices of time-inhomogeneous Markov chains. Since it requires intensive computations to solve for the solutions of Markov structures, the development of productive algorithms is another research direction. The algorithms are expected to identify feasible solutions in an effective fashion.

- Calibrations from the market data

  An empirical question is to calibrate one solution of Markov structures that is consistent with the market data. Although the Markit CMBX index is formed by a basket of 25 commercial mortgage-backed securities, this index is not directly linked to the credit ratings of major banks, for instant, a defaultable bond issued by the major bank. Instead, the CMBX index is indirectly connected to the major banks by holding those commercial mortgage-backed securities. These investments are evaluated and reported in the financial statement. In general, a sound financial statement represents a stable financial status, that is a factor to higher credit rating. However, the methodology to calibrate the dependence between major banks through this type of financial products is not

yet fully established. Thus, one suggesting research problem is to develop a method to calibrate the dependence structure between financial institutions.

- Stochastic dependence measure

  Besides risk management, our proposed systemic dependence measure evaluates the level to which a multivariate Markov chain deviates from the independence structure, that marginal laws are preserved. What we tried to achieve is similar to an existing research problem: comparing Markov chains. In literature, the concepts of comparing Markov chains focus on comparing their transition matrices of time-homogeneous Markov chains. These time-homogeneous Markov chains are not limited to multivariate Markov chains, and thus the concepts about preserving prescribed laws are not clearly mentioned or introduced. In the contexts of Markov structures, we impose additional conditions on the transition matrix functions so that the prescribed laws are preserved. The existing results about comparing Markov chains do not meet our purposes. An interesting problem is to study this dependence measure for time-inhomogeneous multivariate Markov chains whose components satisfy the prescribed laws.

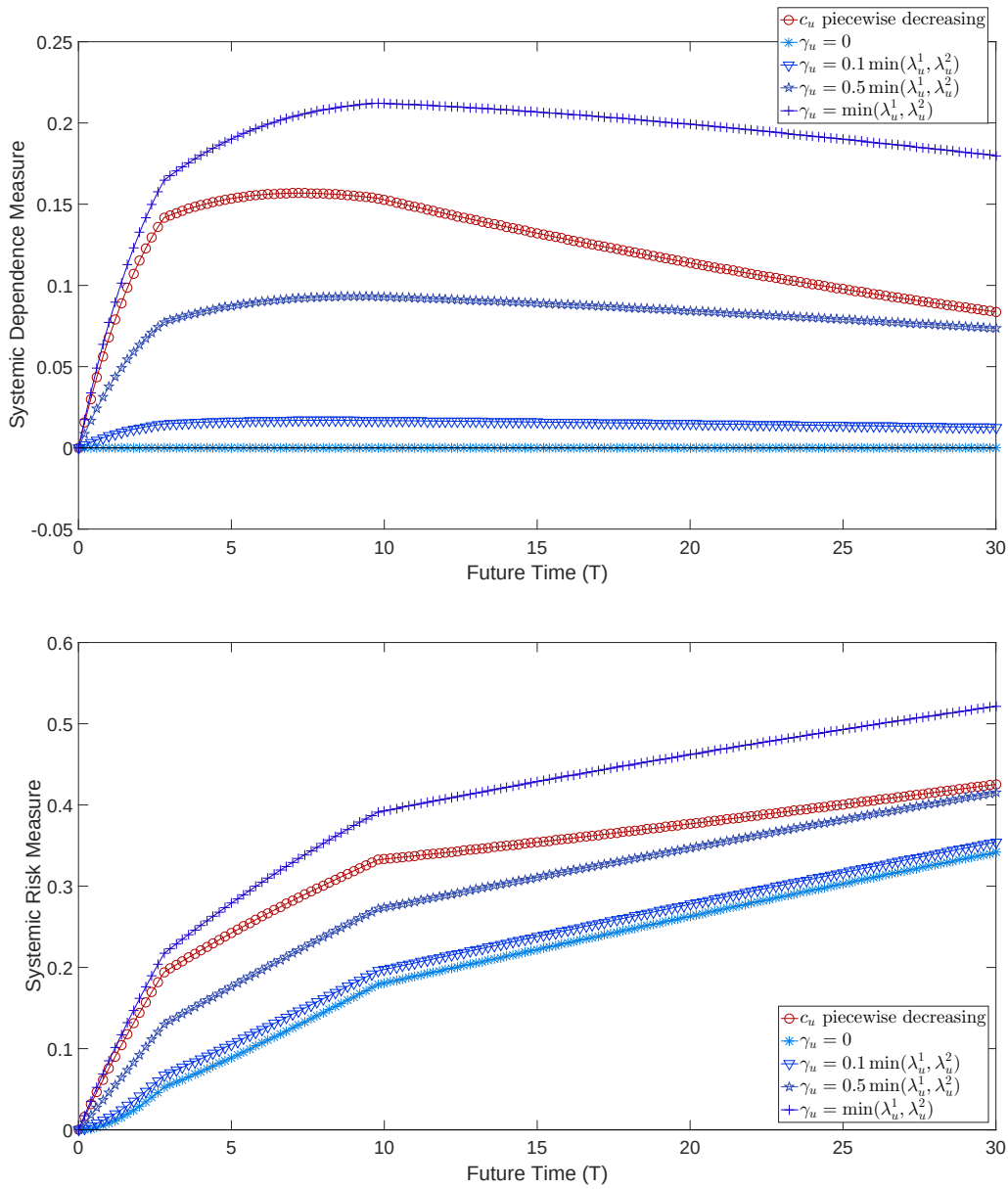

Figure 4.5. Results for Example 4.5.1: Parameter c_u Piecewise Decreasing

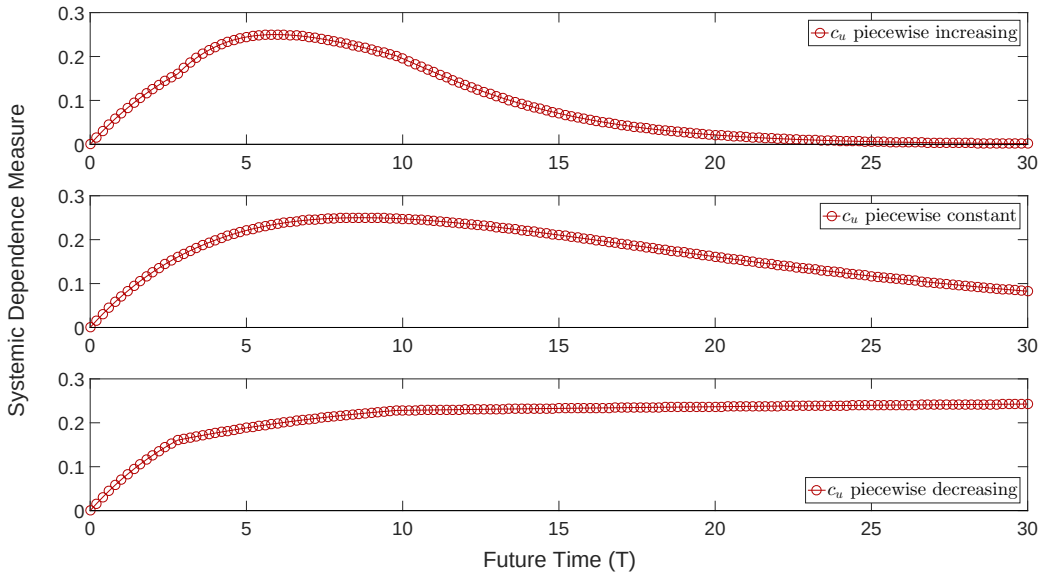

Figure 4.6. Results for Example 4.5.2: Extreme Contagion

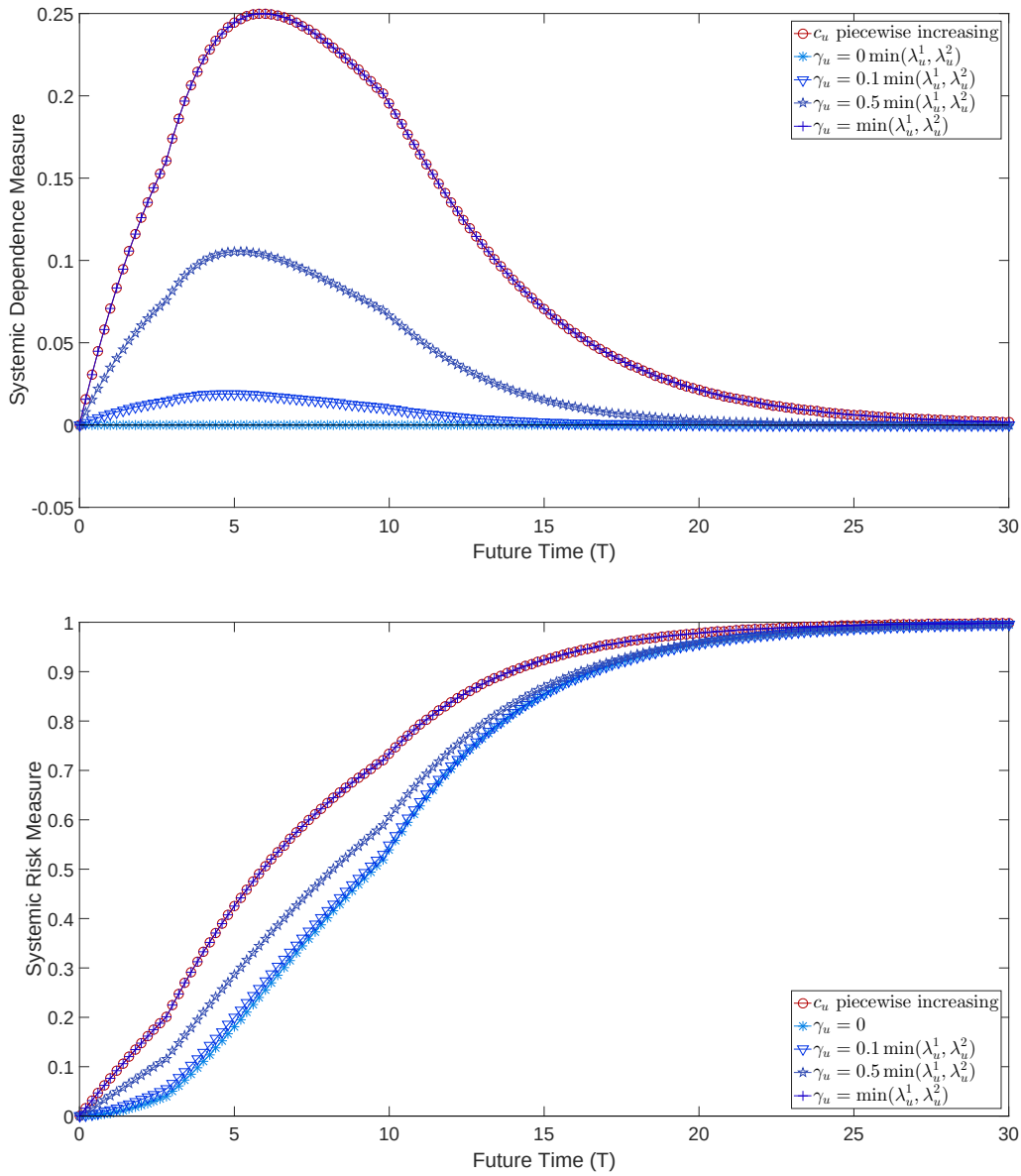

Figure 4.7. Results for Example 4.5.2: Parameter c_u Piecewise Increasing

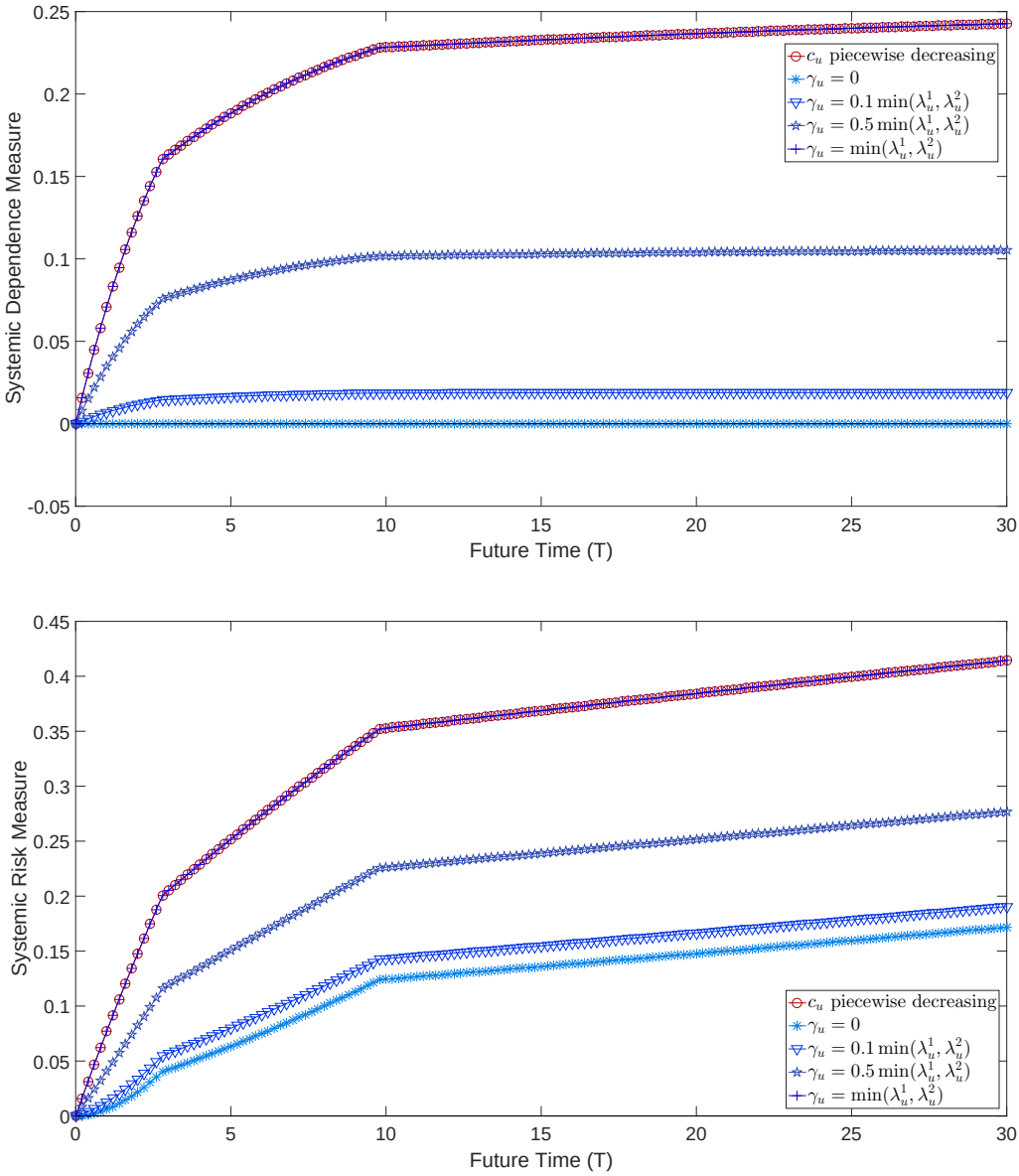

Figure 4.8. Results for Example 4.5.2: Parameter c_u Piecewise Decreasing

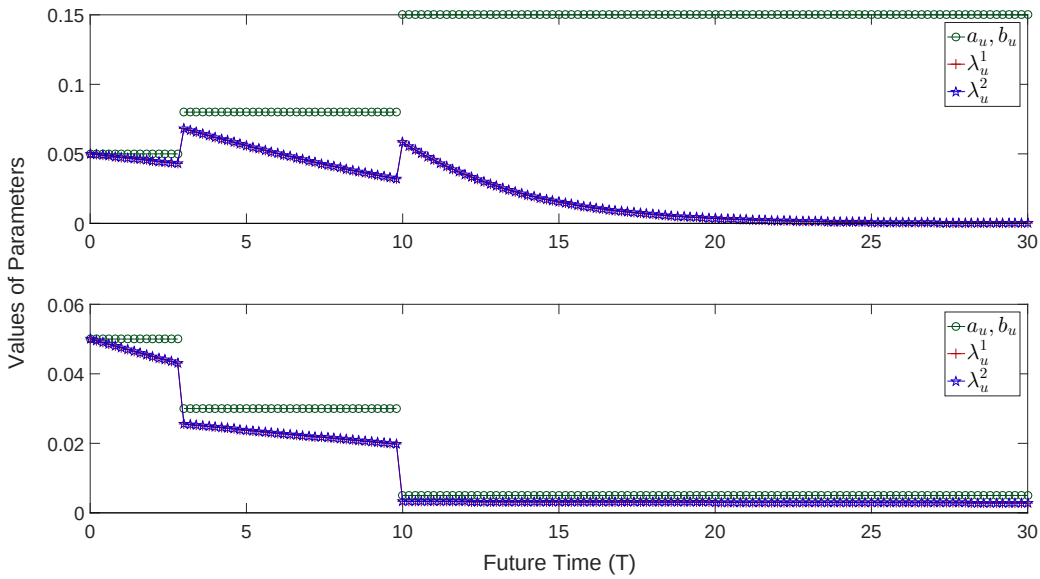

Figure 4.9. Results for Example 4.5.3: Comparision Between $a_u, b_u, \lambda_u^1, \lambda_u^2$

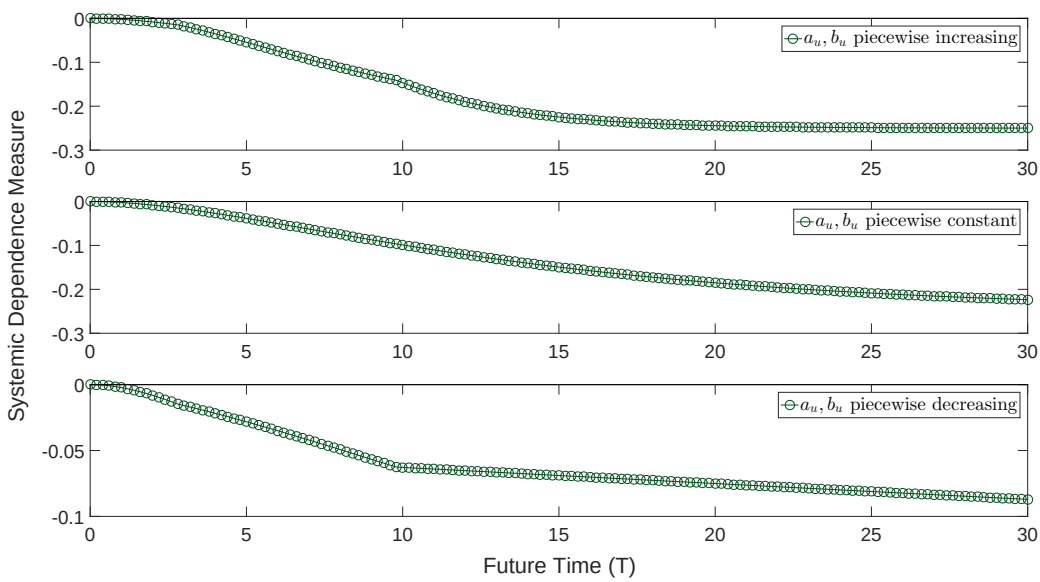

Figure 4.10. Results for Example 4.5.3: Extreme Anti-contagion

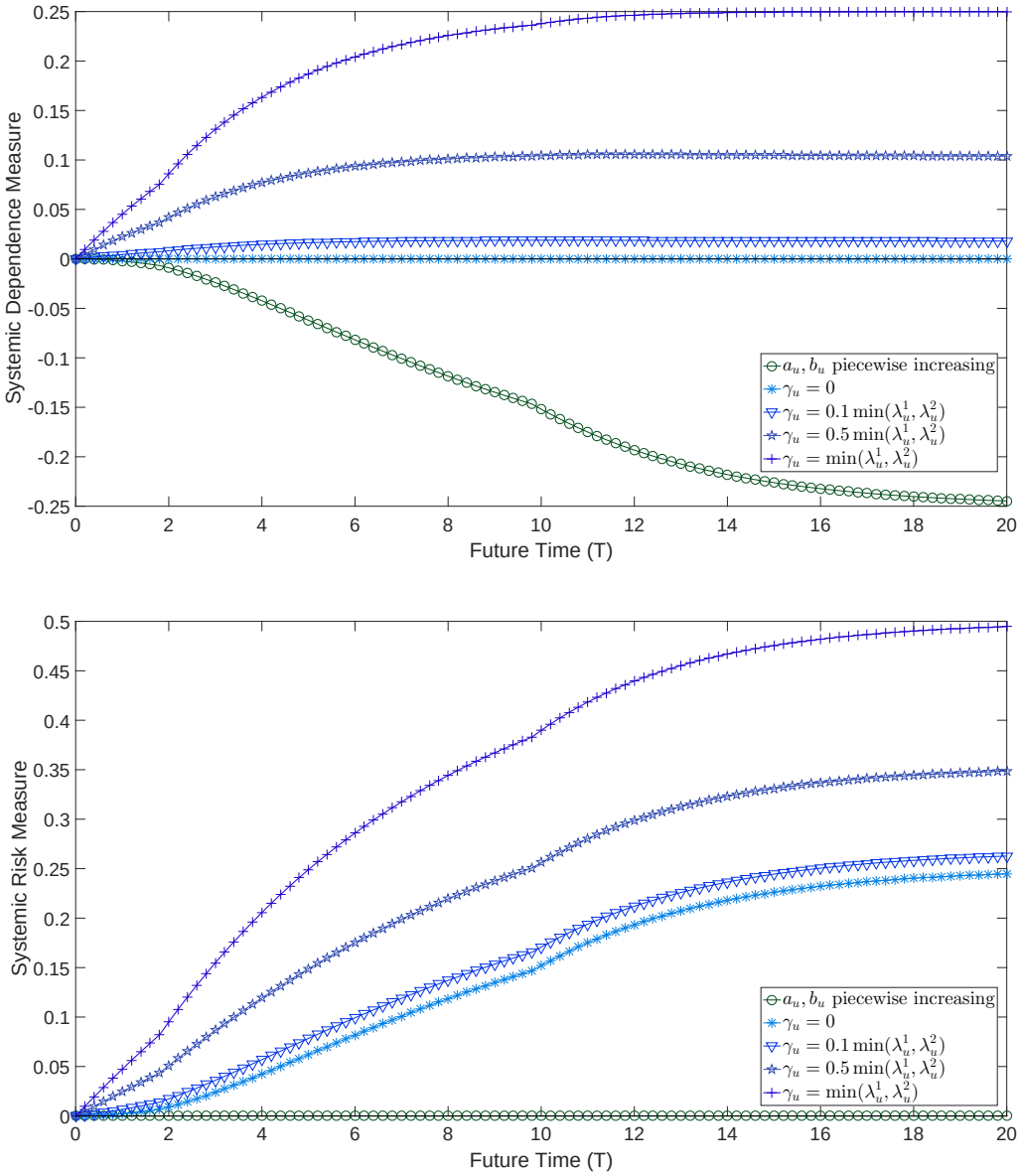

Figure 4.11. Results for Example 4.5.3: Parameters a_u, b_u Piecewise Increasing

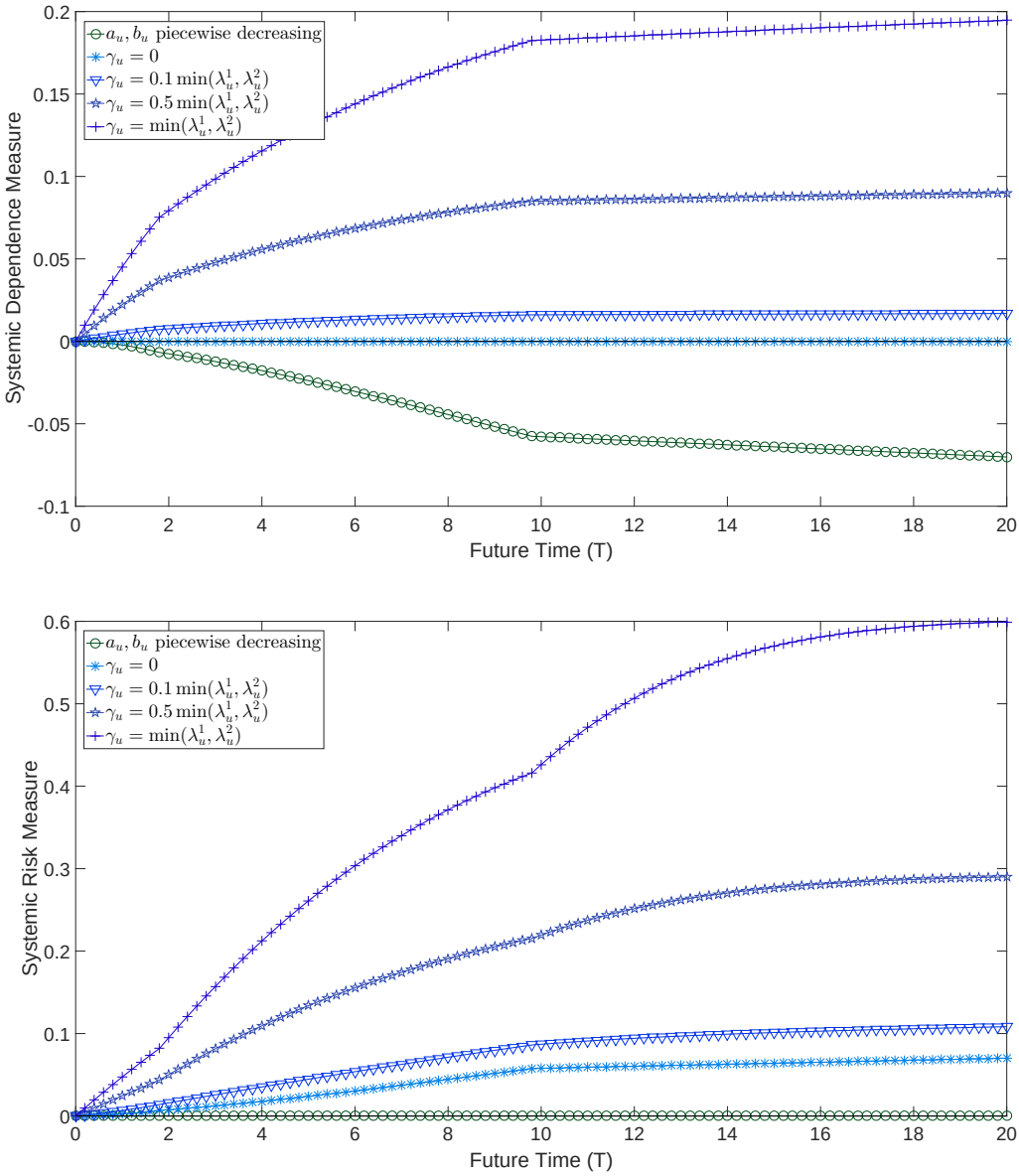

Figure 4.12. Results for Example 4.5.3: Parameters a_u, b_u Piecewise Decreasing

APPENDIX A

COMPUTATIONS FOR EXAMPLE 2.6.1

Recall that for any $0 \leq t \leq s$, the transition probability matrix function of X is given by,

$$\mathbf{P}_{t,s} = \left[p_{t,s}(i,j)\right]_{i,j=1,2,3,4}$$

$$= \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \mathrm{e}^{-\int_t^s (a_v+c_v)\,\mathrm{d}v} & p_{t,s}(1,2) & 0 & p_{t,s}(1,4) \\ 0 & \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} & 0 & 1-\mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} \\ 0 & 0 & \mathrm{e}^{-\int_t^s f_v\,\mathrm{d}v} & 1-\mathrm{e}^{-\int_t^s f_v\,\mathrm{d}v} \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array},$$

where

$$p_{t,s}(1,2) = \frac{a_s}{a_s+c_s-d_s}\left(\mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} - \mathrm{e}^{-\int_t^s (a_v+c_v)\,\mathrm{d}v}\right)$$
$$p_{t,s}(1,4) = 1 - \frac{c_s-d_s}{a_s+c_s-d_s}\mathrm{e}^{-\int_t^s (a_v+c_v)\,\mathrm{d}v} - \frac{a_s}{a_s+c_s-d_s}\mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v}.$$

and for any $t \geq 0$ the operator Θ_t^1 can be represented by

$$\Theta_t^1 = \left[\mathbb{P}\left(X_t = x_t \mid X_t^1 = x_t^1\right)\right]_{x_t \in \mathcal{K}^2, x_t^1 \in \mathcal{K}} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \theta_t & 1-\theta_t & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array},$$

where

$$\theta_t = \frac{p_{0,t}(1,1)}{p_{0,t}(1,1)+p_{0,t}(1,2)} = \frac{\mathrm{e}^{-\int_0^t (a_v+c_v)\,\mathrm{d}v}}{\frac{c_t-d_t}{a_t+c_t-d_t}\mathrm{e}^{-\int_0^t (a_v+c_v)\,\mathrm{d}v} + \frac{a_t}{a_t+c_t-d_t}\mathrm{e}^{-\int_0^t d_v\,\mathrm{d}v}}.$$

The extension operator Φ^1 can be represented by

$$\Phi^1 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \end{array}.$$

We have

$$\Theta_t^1 \mathbf{P}_{t,s} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \alpha_{t,s}(1,1) & \alpha_{t,s}(1,2) & 0 & \alpha_{t,s}(1,4) \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}, \tag{A.1}$$

where

$$\begin{aligned}
\alpha_{t,s}(1,1) &= \theta_t p_{t,s}(1,1) = \theta_t \mathrm{e}^{-\int_t^s (a_v+c_v)\,\mathrm{d}v} \\
\alpha_{t,s}(1,2) &= \theta_t p_{t,s}(1,2) + (1-\theta_t)\, p_{t,s}(2,2) \\
&= \theta_t \left(\frac{a_s}{a_s+c_s-d_s} \left(\mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} - \mathrm{e}^{-\int_t^s (a_v+c_v)\,\mathrm{d}v} \right) \right) + (1-\theta_t)\, \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} \\
\alpha_{t,s}(1,4) &= \theta_t p_{t,s}(1,4) + (1-\theta_t)\, p_{t,s}(2,4) \\
&= \theta_t \left(1 - \frac{c_s+d_s}{a_s+c_s-d_s} \mathrm{e}^{-\int_t^s (a_v+c_v)\,\mathrm{d}v} - \frac{a_s}{a_s+c_s-d_s} \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} \right) \\
&\quad + (1-\theta_t) \left(1 - \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} \right).
\end{aligned}$$

We have the matrix function

$$\mathbf{P}_{t,s}^1 = \Theta_t^1 \mathbf{P}_{t,s} \Phi^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} \alpha_{t,s}(1,1) + \alpha_{t,s}(1,2) & \alpha_{t,s}(1,4) \\ 0 & 1 \end{pmatrix} \end{array},$$

where

$$\begin{aligned}
\alpha_{t,s}(1,1) + \alpha_{t,s}(1,2) &= \theta_t \left(p_{t,s}(1,1) + p_{t,s}(1,2) \right) + (1-\theta_t)\, p_{t,s}(2,2) \\
&= \theta_t \left(\frac{c_s-d_s}{a_s+c_s-d_s} \mathrm{e}^{-\int_t^s (a_v+c_v)\,\mathrm{d}v} + \frac{a_s}{a_s+c_s-d_s} \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} \right) \\
&\quad + (1-\theta_t)\, \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v}.
\end{aligned}$$

Since the infinitesimal rate of $\left(\mathbf{P}_{t,s}^1, 0 \le t \le s\right)$ is the same as $(\Theta_u^1 \Lambda_u \Phi^1, u \ge 0)$, we conclude that $\mathbf{P}_{t,s}^1 = \widehat{\mathbf{P}}_{t,s}^1$. Note that for any $0 \le t \le s$ the one-dimensional probability of X at time s can be computed from the the one-dimensional probability of X at

time t, then we have Θ_s^1

$$\Theta_s^1 = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \theta_s & 1-\theta_s & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array},$$

where

$$\begin{aligned}
\theta_s &= \frac{\mathbb{P}\left(X_s = (0,0)\right)}{\mathbb{P}\left(X_s^1 = 0\right)} \\
\mathbb{P}\left(X_s = (0,0)\right) &= \sum_{z \in \mathcal{K}^2} \mathbb{P}\left(X_t = z\right) \mathbf{P}_{t,s}^{z,(0,0)} = p_{0,t}(1,1) p_{t,s}(1,1) \\
&= \mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v} \mathrm{e}^{-\int_t^s (a_v + c_v)\,\mathrm{d}v} \\
\mathbb{P}\left(X_s^1 = 0\right) &= p_{0,t}(1,1)\left(p_{t,s}(1,1) + p_{t,s}(1,2)\right) + p_{0,t}(1,2) p_{t,s}(2,2) \\
&= \mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v} \left(\frac{c_s - d_s}{a_s + c_s - d_s} \mathrm{e}^{-\int_t^s (a_v + c_v)\,\mathrm{d}v} + \frac{a_s}{a_s + c_s - d_s} \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v} \right) \\
&\quad + \frac{a_t}{a_t + c_t - d_t} \left(\mathrm{e}^{-\int_0^t d_v\,\mathrm{d}v} - \mathrm{e}^{-\int_0^t (a_v + c_v)\,\mathrm{d}v} \right) \mathrm{e}^{-\int_t^s d_v\,\mathrm{d}v}.
\end{aligned}$$

Then we have

$$\begin{aligned}
&\widehat{\mathbf{P}}_{t,s}^1 \Theta_s^1 = \Theta_t^1 \mathbf{P}_{t,s} \Phi^1 \Theta_s^1 \\
&= \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \theta_s\left(\alpha_{t,s}(1,1) + \alpha_{t,s}(1,2)\right) & (1-\theta_s)\left(\alpha_{t,s}(1,1) + \alpha_{t,s}(1,2)\right) & 0 & \alpha_{t,s}(1,4) \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.
\end{aligned} \tag{A.2}$$

Now, we compare each entry of (A.1) and (A.2). We can show that

$$\alpha_{t,s}(1,1) - \theta_s\left(\alpha_{t,s}(1,1) + \alpha_{t,s}(1,2)\right) = 0. \tag{A.3}$$

Moreover, (A.3) is equivalent to

$$\alpha_{t,s}(1,2) - (1-\theta_s)\left(\alpha_{t,s}(1,1) + \alpha_{t,s}(1,2)\right) = 0.$$

Thus, we conclude that for any $0 \le t \le s$, the condition $\Theta_t^1 \mathbf{P}_{t,s} = \widehat{\mathbf{P}}_{t,s}^1 \Theta_s^1$ holds.

APPENDIX B

COMPUTATIONS FOR $\kappa^1_{T_{N-1},T_N}(J,K)$

Recall the matrix $\kappa^i_{t_{n-1},t_n}$, $i \in \{1, 2, \ldots, m\}$, is given by

$$\kappa^i_{t_{n-1},t_n} = \left[\kappa^{i;x^i_t,x_{n-1}}_{t_{n-1},t_n}\right] = \left[\kappa^i_{t_{n-1},t_n}(j,k)\right]$$

$$= \begin{array}{c} \\ x^i_t=0 \\ x^i_t=1 \end{array} \begin{array}{c} \begin{array}{cccc} x_{n-1}=(0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} \kappa^i_{t_{n-1},t_n}(1,1) & \kappa^i_{t_{n-1},t_n}(1,2) & \kappa^i_{t_{n-1},t_n}(1,3) & \kappa^i_{t_{n-1},t_n}(1,4) \\ \kappa^i_{t_{n-1},t_n}(2,1) & \kappa^i_{t_{n-1},t_n}(2,2) & \kappa^i_{t_{n-1},t_n}(2,3) & \kappa^i_{t_{n-1},t_n}(2,4) \end{pmatrix} \end{array},$$

where the jth row of $\kappa^i_{t_{n-1},t_n}$ corresponds to the jth ordered state space of $X^i_{t_n}$, and the kth column of $\kappa^i_{t_{n-1},t_n}$ corresponds to the kth ordered state space of $X_{t_{n-1}}$, and

$$\begin{aligned} \kappa^{i;x^i_t,x_{n-1}}_{t_{n-1},t_n} = & \sum_{x_n \in \mathcal{D}(x^i_n)} \sum_{u \in \mathcal{K}^2} \mathbb{P}\left(X_{t_{n-1}} = z\right) \\ & \times \left[\left(\sum_{x_n \in \mathcal{D}(x^i_n)} \mathbf{P}^{z,x_n}_{t_{n-1},t_n} \right) \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n} - \mathbf{P}^{z,x_n}_{t_{n-1},t_n} \left(\sum_{x_n \in \mathcal{D}(x^i_n)} \mathbf{P}^{x_{n-1},x_n}_{t_{n-1},t_n} \right) \right]. \end{aligned}$$

In the following, we compute $\kappa^1_{t_{n-1},t_n}$ for $\mathcal{K} = \{0, 1\}$ and $m = 2$.

$$\begin{aligned} \kappa^1_{t_{n-1},t_n}(1,1) = & \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(0,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (0,1)\right) \mathbf{P}^{(0,1),(0,1)}_{t_{n-1},t_n} \right) \mathbf{P}^{(0,0),(0,0)}_{t_{n-1},t_n} \\ & - \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(0,0)}_{t_{n-1},t_n} \right) \mathbf{P}^{(0,0),(0,1)}_{t_{n-1},t_n} \\ = & \left(\mathbb{P}\left(X_{t_{n-1}} = (0,1)\right) \mathbf{P}^{(0,1),(0,1)}_{t_{n-1},t_n} \right) \mathbf{P}^{(0,0),(0,0)}_{t_{n-1},t_n} \\ \kappa^1_{t_{n-1},t_n}(1,2) = & \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(0,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (0,1)\right) \mathbf{P}^{(0,1),(0,1)}_{t_{n-1},t_n} \right) \mathbf{P}^{(0,1),(0,0)}_{t_{n-1},t_n} \\ & - \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(0,0)}_{t_{n-1},t_n} \right) \mathbf{P}^{(0,1),(0,1)}_{t_{n-1},t_n} \\ = & - \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(0,0)}_{t_{n-1},t_n} \right) \mathbf{P}^{(0,1),(0,1)}_{t_{n-1},t_n} \\ \kappa^1_{t_{n-1},t_n}(1,3) = & \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(0,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (0,1)\right) \mathbf{P}^{(0,1),(0,1)}_{t_{n-1},t_n} \right) \mathbf{P}^{(1,0),(0,0)}_{t_{n-1},t_n} \\ & - \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(0,0)}_{t_{n-1},t_n} \right) \mathbf{P}^{(1,0),(0,1)}_{t_{n-1},t_n} \\ = & \; 0 \\ \kappa^1_{t_{n-1},t_n}(1,4) = & \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(0,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (0,1)\right) \mathbf{P}^{(0,1),(0,1)}_{t_{n-1},t_n} \right) \mathbf{P}^{(1,1),(0,0)}_{t_{n-1},t_n} \\ & - \left(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(0,0)}_{t_{n-1},t_n} \right) \mathbf{P}^{(1,1),(0,1)}_{t_{n-1},t_n} \end{aligned}$$

$$
\begin{aligned}
&= 0 \\
\kappa^1_{t_{n-1},t_n}(2,1) &= \Big(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(1,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (0,1)\right) \mathbf{P}^{(0,1),(1,1)}_{t_{n-1},t_n} \\
&\quad + \mathbb{P}\left(X_{t_{n-1}} = (1,0)\right) \mathbf{P}^{(1,0),(1,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (1,1)\right) \mathbf{P}^{(1,1),(1,1)}_{t_{n-1},t_n} \Big) \mathbf{P}^{(0,0),(1,0)}_{t_{n-1},t_n} \\
&\quad - \Big(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(1,0)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (1,0)\right) \mathbf{P}^{(1,0),(1,0)}_{t_{n-1},t_n} \Big) \mathbf{P}^{(0,0),(1,1)}_{t_{n-1},t_n} \\
\kappa^1_{t_{n-1},t_n}(2,2) &= \Big(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(1,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (0,1)\right) \mathbf{P}^{(0,1),(1,1)}_{t_{n-1},t_n} \\
&\quad + \mathbb{P}\left(X_{t_{n-1}} = (1,0)\right) \mathbf{P}^{(1,0),(1,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (1,1)\right) \mathbf{P}^{(1,1),(1,1)}_{t_{n-1},t_n} \Big) \mathbf{P}^{(0,1),(1,0)}_{t_{n-1},t_n} \\
&\quad - \Big(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(1,0)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (1,0)\right) \mathbf{P}^{(1,0),(1,0)}_{t_{n-1},t_n} \Big) \mathbf{P}^{(0,1),(1,1)}_{t_{n-1},t_n} \\
\kappa^1_{t_{n-1},t_n}(2,3) &= \Big(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(1,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (0,1)\right) \mathbf{P}^{(0,1),(1,1)}_{t_{n-1},t_n} \\
&\quad + \mathbb{P}\left(X_{t_{n-1}} = (1,0)\right) \mathbf{P}^{(1,0),(1,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (1,1)\right) \mathbf{P}^{(1,1),(1,1)}_{t_{n-1},t_n} \Big) \mathbf{P}^{(1,0),(1,0)}_{t_{n-1},t_n} \\
&\quad - \Big(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(1,0)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (1,0)\right) \mathbf{P}^{(1,0),(1,0)}_{t_{n-1},t_n} \Big) \mathbf{P}^{(1,0),(1,1)}_{t_{n-1},t_n} \\
\kappa^1_{t_{n-1},t_n}(2,4) &= \Big(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(1,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (0,1)\right) \mathbf{P}^{(0,1),(1,1)}_{t_{n-1},t_n} \\
&\quad + \mathbb{P}\left(X_{t_{n-1}} = (1,0)\right) \mathbf{P}^{(1,0),(1,1)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (1,1)\right) \mathbf{P}^{(1,1),(1,1)}_{t_{n-1},t_n} \Big) \mathbf{P}^{(1,1),(1,0)}_{t_{n-1},t_n} \\
&\quad - \Big(\mathbb{P}\left(X_{t_{n-1}} = (0,0)\right) \mathbf{P}^{(0,0),(1,0)}_{t_{n-1},t_n} + \mathbb{P}\left(X_{t_{n-1}} = (1,0)\right) \mathbf{P}^{(1,0),(1,0)}_{t_{n-1},t_n} \Big) \mathbf{P}^{(1,1),(1,1)}_{t_{n-1},t_n}.
\end{aligned}
$$

BIBLIOGRAPHY

[ADEH99] Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.

[BCCH14] Tomasz R. Bielecki, Areski Cousin, Stéphane Crépey, and Alexander Herbertsson. Dynamic hedging of portfolio credit risk in a Markov copula model. *J. Optim. Theory Appl.*, 161(1):90–102, 2014.

[BDN+00] Eric Bouyé, Valdo Durrleman, Ashkan Nikeghbali, Gaël Riboulet, and Thierry Roncalli. Copulas for Finance - A Reading Guide and Some Applications. 2000.

[BFF15] Francesca Biagini, Jean-Pierre Fouque, and Marco Frittelli. A unified approach to systemic risk measures via acceptance sets. *Preprint*, pages 1–54, 2015.

[BFLV12] Dimitrios Bisias, Mark Flood, Andrew W. Lo, and Stavros Valavanis. A survey of systemic risk analytics. *Annual Review of Financial Economics*, 4(1):255–296, 2012.

[Bil12] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, 2012.

[BJN] Tomasz R. Bielecki, Jacek Jakubowski, and Mariusz Nieweglowski. A note on independence copula for conditional Markov chains. In Roderick Melnik, Roman Makarov, , and Jacques Belair, editors, *Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science*. Springer-Verlag New York. to appear in 2017.

[BJN10] Tomasz R. Bielecki, Jacek Jakubowski, and Mariusz Nieweglowski. Dynamic modeling of dependence in finance via copulae between stochastic processes. 198:33–76, 2010.

[BJN12] Tomasz R. Bielecki, Jacek Jakubowski, and Mariusz Nieweglowski. Study of dependence for some stochastic processes: Symbolic Markov copulae. *Stochastic Processes and their Applications*, 122:930–951, 2012.

[BJN13] Tomasz R. Bielecki, Jacek Jakubowski, and Mariusz Nieweglowski. Intricacies of dependence between components of multivariate Markov chains: Weak Markov consistency and weak Markov copulae. *Electronic Journal of Probability*, 18(45):1–21, 2013.

[BJVV08] Tomasz R. Bielecki, Jacek Jakubowski, Andrea Vidozzi, and Luca Vidozzi. Study of dependence for some stochastic processes. *Stochastic Analysis and Applications*, 26:903–924, 2008.

[BR58] C. J. Burke and M. Rosenblatt. A Markovian function of a Markov chain. *Ann. Math. Statist.*, 29:1112–1122, 1958.

[BVV08] Tomasz R. Bielecki, Andrea Vidozzi, and Luca Vidozzi. A Markov copulae approach to pricing and hedging of credit index derivatives and ratings triggered step-up bonds. *Journal of Credit Risk*, 4:47–76, 2008.

[CFS15] René Carmona, Jean-Pierre Fouque, and Li-Hsien Sun. Mean field games and systemic risk. *Communications in Mathematical Sciences*, 13(4):911–933, 2015.

[CIM13] Chen Chen, Garud Iyengar, and Ciamac C. Moallemi. An axiomatic approach to systemic risk. *Management Science*, 59(6):1373–1388, 2013.

[Dyn65a] Eugene B. Dynkin. *Markov Processes*, volume I. Springer-Verlag, 1965.

[Dyn65b] Eugene B. Dynkin. *Markov Processes*, volume II. Springer-Verlag, 1965.

[EN01] Larry Eisenberg and Thomas H. Noe. Systemic risk in financial systems. *Management Science*, 47(2):236–249, 2001.

[FL13] Jean-Pierre Fouque and Joseph A. Langsam, editors. *Handbook on Systemic Risk*. Cambridge, 2013.

[FRW15] Zachary Feinstein, Birgit Rudloff, and Stefan Weber. Measures of systemic risk. 2015. arXiv:1502.07961.

[Gig11] Stefano Giglio. Credit default swap spreads and systemic financial risk. *Proceedings, Federal Reserve Bank of Chicago*, 10:104–141, March 2011.

[GL05] Leonid Gurvits and James Ledoux. Markov property for a function of a Markov chain: a linear algebra approach. *Linear Algebra and its Applications*, 404(2149655):85–117, 2005.

[Glo91] Joseph Glover. Markov functions. *Annales de l'I. H. P., section B*, 2:221–238, 1991.

[GPY13] Josselin Garnier, George Papanicolaou, and Tzu-Wei Yang. Large deviations for a mean field model of systemic risk. *SIAM Journal on Financial Mathematics*, 4(1):151–184, August 2013.

[GY15] Paul Glasserman and H. Peyton Young. How likely is contagion in financial networks? *Journal of Banking & Finance*, 50:383–399, 2015.

[HMBS16] Hannes Hoffmann, Thilo Meyer-Brandis, and Gregor Svindland. Risk-consistent conditional systemic risk measures. *Stochastic Processes and their Applications*, 126(7):2017–2037, 2016.

[Ios80] Marius Iosifescu. *Finite Markov Processes and Their Applications*. John Wiley & Sons, revised and expanded edition, 1980.

[Kel82] Frank P. Kelly. Markovian functions of a Markov chain. *Sankhyā: The Indian Journal of Statistics, Series A*, 44:372–379, 1982.

[KO88] Thomas G. Kurtz and Daniel Ocone. Unique characterization of conditional distribution in nonlinear filtering. *The Annals of Probability*, 16:80–107, 1988.

[KOZ16] Eduard Kromer, Ludger Overbeck, and Katrin A. Zilch. Systemic risk measures on general probability spaces. 2016. Available at SSRN: https://ssrn.com/abstract=2268105 or http://dx.doi.org/10.2139/ssrn.2268105.

[Ros71] Murray Rosenblatt. *Markov Processes. Structure and Asymptotic Behavior.* Springer-Verlag Berlin Heidelberg, 1971.

[RP81] L. C. G. Rogers and Jim W. Pitman. Markov functions. *The Annals of Probability*, 9:573–582, 1981.

[RS93] Gerardo Rubino and Bruno Sericola. A finite characterization of weak lumpable Markov processes. Part II. The continuous time case. *Stochastic Processes and their applications*, 45(1):115–125, 1993.

[RSST98] Tomasz Rolski, Hanspeter Schmidli, Volker Schmidt, and Jozef Teugels. *Stochastic Processes for Insurance and Finance.* John Wiley & Sons, 1998.

[RV13] Leonard C. G. Rogers and Luitgard A. M. Veraart. Failure and rescue in an interbank network. *Management Science*, 59(4):882–898, 2013.

[Skl59] Abe Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de lInstitut de Statistique de lUniversité de Paris*, 8:229–231, 1959.

[Sum13] Martin Summer. Financial contagion and network analysis. *Annual Review of Financial Economics*, 5:277–297, 2013.

[Wen81] A. D. Wentzell. *A Course in the Theory of Stochastic Processes.* McGraw-Hill, 1981.