SAPIENZA
UNIVERSITÀ DI ROMA

Scuola di Dottorato in Economia

Dottorato di Ricerca in Matematica per le Applicazioni Economico-Finanziarie XXV ciclo

On Semi-Markov Processes and Financial Applications

Thesis submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy in

Mathematics for Economic-Financial Applications

by

Giovanni Salvi

Program Coordinator
Prof. Dr. Maria B. Chiarolla

Thesis Advisors
Prof. Dr. Raimondo Manca
Dr. Guglielmo D'Amico

La probabilità non è nient'altro che il grado di fiducia (speranza, timore, ..)

nel fatto che qualcosa di atteso (temuto, o sperato, o indifferente)

si verifichi e risulti vero

Bruno de Finetti

Ai miei genitori

Ringraziamenti

Giunto alla conclusione di questo lungo cammino è per me un privilegio poter ringraziare alcune delle persone che mi hanno sostenuto lungo la strada.

Prima di tutto vorrei ringraziare i miei supervisors, il Prof. Raimondo Manca e il Dott. Guglielmo D'Amico, per tutte le conoscenze che in questi anni mi hanno trasmesso, per avermi guidato nel mondo della ricerca lasciandomi tuttavia sempre la massima libertà ed autonomia, per avermi incoraggiato ed aver creduto in me nei momenti difficili e per aver guardato con entusiasmo ai risultati che stavo maturando. Per tutto quello che avete fatto per me, grazie.

Vorrei ringraziare la coordinatrice della scuola, la Prof.ssa Maria Chiarolla, per aver creduto in me ed avermi dato la possibilità di intraprendere questo percorso.

Vorrei anche ringraziare il Prof. Anatoliy V. Swishchuk che mi ha ospitato durante il mio visiting all'Università di Calgary, sotto la sua guida ho vissuto un periodo molto produttivo e stimolante.

Ringrazio tutti i membri del dipartimento per la cordiale e calorosa atmosfera che condividono a creare. Ringrazio inoltre il Prof. Blasi per avermi dato l'opportunità di svolgere attività didattica e in particolare la Prof.ssa Novaro, la Dott.ssa Ventura e il Dott. Stabile con cui ho collaborato direttamente.

Un pensiero speciale va a tutti i miei colleghi dottorandi e assegnisti del dipartimento con cui ho condiviso oltre ad all'accogliente stanza 144, una lunga esperienza di vita che porterò sempre con me, grazie a voi tutti. Tra di loro voglio ricordare in particolare Giorgio e Tiziano, miei colleghi fin dai tempi di fisica, con cui si è creato un rapporto sempre più stretto in questi anni. Ilaria e Alessandro con cui ho percorso questo cammino che hanno contribuito a rendere sempre vivaci le nostre giornate. Antonio e Stefano grandi protagonisti delle cene di dottorato. Isabella e Marco che con la loro esperienza hanno guidato noi giovani. Mario, Sofia, Elisa e Naeem a cui lascio il testimone di dottorandi anziani e tutte le giovani leve a cui faccio il mio più sentito in bocca al lupo.

Per quanto riguarda la mia vita extra-lavorativa, questi anni hanno segnato la transizione dalla fase di studente a quella, spero, di lavoratore. Sono stati anni intensi, a tratti duri ma con molti momenti felici. Voglio qui ringraziare tutte le persone che hanno condiviso con me questo periodo.

Prima di tutto voglio ringraziare la mia famiglia che sempre mi ha sostenuto, incoraggiato e sopportato lungo il cammino della mia vita, se ho raggiunto questo traguardo è grazie a voi tutti.

Voglio anche ringraziare i miei amici per tutto quello che consapevolmente o no mi hanno dato durante questi anni, vi abbraccio tutti.

Infine, il mio pensiero va alla persona con cui ho condiviso questo cammino e sto condividendo il percorso della mia vita, Camilla. Un ringraziamento speciale per essere sempre stata al mio fianco, per aver sempre creduto in me ed avermi sostenuto fin dal primo giorno. Grazie.

Contents

Introduction

The complexity arising from financial markets is a challenge to develop new and more sophisticated models. Since the pioneering work of Black and Scholes [10], the stochastic finance literature has grown as well as the interest on the topic. Markov processes have been largely used in an attempt to give a stochastic description of financial markets. However the Markovian memoryless hypothesis imposes strict conditions on the distribution of the waiting times in the states. Semi-Markov processes are a generalization of Markov processes allowing every kind of waiting times distribution and preserving the memoryless hypothesis in a more adaptable way.

Semi-Markov processes can be view as a generalization of both Markov processes and renewal processes. The first works on the topic have been produced independently by Lévy [64], Smith [79] and Takacs [83]. An important theoretical contribution was given by Çinlar [27, 28]. Since the introduction of this concept the interest on the topic has been growing. Nowadays there are applications of semi-Markov processes in various fields as a mark of their flexibility.

In this thesis we will focus our attention on applications in finance, in particular on interest rate, volatility derivatives and credit risk.

A relevant class of models for application in finance is that of (Markov or semi-Markov) modulated models. Indeed, they can give an explanation of the external random elements that influence the phenomena. For example, the classical interest rate diffusion models can be modulated by a switching process that allows to take into account for macroeconomic changing; modulated diffusion model of price returns can give an explanation of low or high volatility periods.

There is a wide literature on interest rate models since the fundamental paper by Vasicek [84]: among the diffusion models of particularly relevance are those of Hull and White [54] and Cox, Ingersoll and Ross (CIR) [30]. Many works have been presented in order to extend and improve diffusion models. Duffie and Kan [47] proposed a model consisting of short rate process function of a state process. Successively, developments have been proposed by Mamon [66] who characterized the term structure of a Markov interest rate model when the interest rate process is assumed to be a function of a continuous time non-homogeneous Markov chain. Then using forward measures Mamon [67] has shown how to price term structure derivative products. In this thesis we assume that the short rate process is a diffusive process modulated by a continuous time semi-Markov process. In this way we provide a general model of evolution that is able to reproduce a great variety of evolutions of the interest rate. The results include renewal type equations for the higher order moments of the zero coupon bond process and for the covariance function of the force of interest.

The market for variance and volatility swaps has been growing, and many investment banks and other financial institutions are now actively quoting volatility swaps on various assets: stock indices, currencies, as well as commodities. Among recent and new financial products there are covariance and correlation swaps, which are useful for volatility hedging and speculation using two different financial underlying assets. Markets with Markov modulated volatility have been analyzed by Elliott and Swishchuk [50], in this work the incompleteness of the market is shown and a minimal martingale measure to price option is found. They give an expression for the price of European option as well as variance swap. In this thesis, within a stochastic Markov modulated volatility market, we obtain an expression for the price of volatility swap. Considering a two risky asset market with Markov modulated volatility, the price of a covariance swap and an expression for the price of a correlation swap are obtained. We apply these results to stochastic volatility driven by a two-state Markov chain. Numerical examples are presented for VIX and VXN volatility indices (S&P 500 and NASDAQ-100, respectively, since January 2004 to June 2012). We also used VIX (January 2004 to June 2012) to price variance and volatility swaps for the two-state Markov-modulated volatility and to present a numerical result in this case.

Using the martingale representation of semi-Markov processes, we were able to generalize this model. Variance swap price has been derived in a semi-Markov stochastic volatility market by Swishchuk [81]. In this work we derive an expression for volatility swap price. In a two risky asset market with semi-Markov volatility we price covariance swap and we obtain an expression for the correlation swap price.

The study of loss probability of any financial subjects is the main challenge of credit risk. There are many approaches to this topic, among them we pay particular attention to credit migration models. The main advantage of these models is their ability to describe not just the probability of default, but what happens to the credit reliability of a debtor during the life of a contract. Semi-Markov processes were proposed for the first time as applied to credit ratings by D'Amico et al. [32], and more recently D'Amico et al. [33] applied this method to credit default swap evaluation.

The recent financial crisis has stressed the importance of considering all the sources of risk associated with financial products. As well known financial markets create a network connecting institutions, banks and companies. In any financial contract the risk of default of the counterpart is then crucial. In order to approach this problem we need to build models able to capture the default correlation between financial subjects.

The first challenge was to build a bivariate model whose components could be credit migration processes. In this thesis we define multivariate semi-Markov chains, i.e. multivariate chains whose components are semi-Markov chain. The evolution of a bivariate chain in time is studied. A bivariate reliability model for the study of credit rating evolution of two debtors is defined. This model has been applied to the study of counterparty risk in credit default swap (CDS).

The evolution of the yield spread as a function of rating evolution has been studied by D'Amico et al. [37]. They used a semi-Markov rewards process for the yield spread, a complete discussion about semi-Markov rewards process is found in Stenberg et al. [80]. Particular relevant, nowadays is the credit spread between two debtors. In this thesis, a bivariate reward model is defined and the evaluation of first and second moments is discussed. In this model we are able to evaluate the credit spread between two debtors.

The thesis is organized in two parts as follows.

The first part is composed of 4 chapters and concerns Markov and semi-Markov modulated models, in particular their applications to interest rate and stochastic volatility. Discrete and continuos time semi-Markov processes are defined in the first chapter. This chapter does not contain new developments, however it recalls the main definitions and fixes the notation for what follows.

A semi-Markov modulated model for interest rate is discussed in the second chapter. In this chapter a continuos time semi-Markov process modulates a diffusion model to describe the force of interest. In this framework we are able to evaluate the Zero Coupon Bond (ZCB) and its higher order moments. We apply this model to some well known diffusion models such as Vasicek, Hull and White and CIR. This chapter is based on a paper [41] currently under review.

Financial markets with Markov modulated volatility are studied in Chapter 3. In this chapter we first describe Markov processes by their martingale representation, then we consider two market models: a single and a double risky assets. In the single risky asset market model with Markov stochastic volatility we show how to price variance and volatility swap. In the two risky assets market model with Markov stochastic volatility we obtain a closed form solution for the price of covariance swaps and we derive an approximation formula for the price of correlation swaps. We discuss an application to a simple two state Markov chain model. We apply our model to S&P 500 and NASDAQ 100 indices. This chapter is based on a paper [75] currently under review. This and the next chapter have been developed during my visiting PhD program at the University of Calgary, Calgary (AB) Canada, under the supervision of Prof. A. V. Swishchuk.

Financial markets with semi-Markov modulated volatility are studied in Chapter 4, that is a generalization of Chapter 3 to semi-Markov environment. A martingale characterization of semi-Markov processes is given at the beginning of the chapter, then we discuss variance and volatility swaps in a single risky asset market with semi-Markov volatility. Covariance and correlation swaps are studied in a two risky assets market with semi-Markov volatilities.

The second part concerns the study of multivariate semi-Markov processes, in particular definition, properties and their application to credit risk.

Bivariate Markov chain are presented in Chapter 5. In the first section we study multidimensional matrices. We use multidimensional matrices as a general framework for bivariate Markov chain. To study this topic we will follow the works of Manca [68, 69] for notation and we refer to them for details and proofs. Bivariate Markov chain are discussed in the second section. We first consider the system as a whole and then we discuss some particular dependence structures between the components.

Discrete time multivariate semi-Markov processes on a finite state space are defined in Chapter 6. First we discuss the assumptions and all the main definitions, then we study the evolution equation of this multivariate process. This chapter and the next are based on a paper [40] currently under review.

A discrete time bivariate semi-Markov reliability model and its applications to counterpart credit risk are presented in Chapter 7. A bivariate semi-Markov reliability model for credit ratings evolution is defined in the first part of this chapter; here we define the system reliability and the marginal reliabilities of the components and we show how to express them in terms of the semi-Markov kernel. The counterparty credit risk in a credit default swap (CDS) contract is analyzed by the bivariate reliability model. The price of a risky CDS and the credit value adjustment due to the counterpart credit risk are obtained. A numerical example is discussed in the final part of the chapter.

A bivariate semi-Markov rewards model and its application to credit spread evaluation is discussed in Chapter 8. The bivariate semi-Markov rewards model and the evaluation of its first and second order moments are discussed on the second section of the chapter. The computation of the credit spread between two debtors is described in the last section. This chapter is based on a working paper [42].

Part I

Markov and Semi-Markov Modulated Models

Chapter 1

Semi-Markov Process

This chapter will recall some standard notations about finite semi-Markov process.

Many books have been written on this topic and we refer to them for more details, eg. Barbu and Limnios [3], Janssen and Manca [59] and Limnios and Oprişan [65].

The chapter is organized as follows: the first section is devoted to discrete time semi-Markov process while in the second section we discuss continuos time processes.

1.1 Discrete Time Semi-Markov Process

In this section we will describe discrete time processes. Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ be a complete filtered probability space and let $E = \{1, \dots, d\}$ be a given finite set.

Definition 1.1.1. ***Markov Chain***

A sequence of random variables $J = (J_n)_{n\in\mathbb{N}^*}$ in $E = \{1, \dots, d\}$ is called time homogenous Markov chain if for every $n \in \mathbb{N}$ we have

$$\mathbb{P}(J_{n+1} = j \mid J_0 = i_0, \dots, J_n = i) = \mathbb{P}(J_{n+1} = j \mid J_n = i) := p_{ij} \qquad \forall\ i, j \in E\ , \tag{1.1}$$

the stochastic matrix $P = (p_{ij})_{i,j\in E}$ is called Markov transition probability.

Let $X = (X_n)_{n\in\mathbb{N}^*}$ be a sequence of positive random variables with values in $\mathbb{N}$, i.e. $X_n > 0$ for any $n \geq 0$. Let $T = (T_n)_{n\in\mathbb{N}^*}$ be the increasing sequence of partial sums

$$T_n = X_0 + X_1 + \dots + X_{n-1} = T_{n-1} + X_{n-1} \qquad \text{for } n \geq 1\ , \tag{1.2}$$

where $T_0 \in \mathbb{Z}$ is the initial time. The $(X_n)_{n\in\mathbb{N}}$ are called lifetimes or waiting times; intuitively they represent the times between the occurrence of two successive events. The $(T_n)_{n\in\mathbb{N}^*}$ are called arrival times and describe the successive instants when a specific event occurs.

Definition 1.1.2. ***Renewal Chain***

A random sequence $T = (T_n)_{n\in\mathbb{N}^*}$ such that the waiting times $X = (X_n)_{n\in\mathbb{N}^*}$ form an i.i.d. sequence and

- $T_0 = 0$ *a.s.;*
- $T_n = X_0 + X_1 + \cdots + X_{n-1}$*;*

is called a renewal chain and the r.v. $T_n,\ n \geq 0$ are called renewal times.

Remark 1.1.3. *If we consider a sequence X with values in $\mathbb{R}_+$ we can define the increasing sequence of partial sums T in exactly the same way.*

Definition 1.1.4. ***Counting Process***

The increasing random sequence $N = (N(t))_{t\in\mathbb{N}^*}$ defined by

$$N(t) = \max\{n \in \mathbb{N} \mid T_n \leq t\} \ , \tag{1.3}$$

is the counting process associated to the renewal chain T. It gives at any time the number of events occurred.

It is interesting to study a process whose sequence of states is determined by a Markov chain and the permanence in any state is triggered by a renewal chain. To this end we first define a semi-Markov kernel and then we introduce the Markov renewal chain.

Definition 1.1.5. ***Discrete-Time Cumulated Semi-Markov Kernel***

A matrix valued function $\mathbf{Q} = (Q_{ij}(t);\ i,j \in E,\ t \in \mathbb{N}^*)$ is a discrete-time cumulated semi-Markov kernel if

- $Q_{ij}(t) \geq 0$ *for every $i,j \in E$ and $t \in \mathbb{N}^*$;*
- $Q_{ij}(0) = 0$ *for every $i,j \in E$;*
- $\left(\lim_{t\to\infty} Q_{ij}(t)\right)_{i,j\in E}$ *is a Markov transition probability.*

Definition 1.1.6. ***Discrete-Time Semi-Markov Kernel***

A matrix valued function $\mathbf{q} = (q_{ij}(t);\ i,j \in E,\ t \in \mathbb{N}^*)$ *is a discrete-time semi-Markov kernel if*

- $q_{ij}(t) \geq 0$ *for every* $i,j \in E$ *and* $t \in \mathbb{N}^*$*;*
- $q_{ij}(0) = 0$ *for every* $i,j \in E$*;*
- $\left(\sum_{t=1}^{\infty} q_{ij}(t)\right)_{i,j\in E}$ *is a Markov transition probability.*

Definition 1.1.7. ***Markov Renewal Chain***

A random sequence $(J,T) = (J_n,T_n)_{n\in\mathbb{N}^*}$ *is a Markov renewal chain if for all* $n \in \mathbb{N}$, $i,j \in E$ *and* $t \in \mathbb{N}$ *it satisfies*

$$\mathbb{P}\{J_{n+1} = j, T_{n+1} - T_n \leq t | \sigma(J_a, T_a), 0 \leq a \leq n\} = \mathbb{P}\{J_{n+1} = j, T_{n+1} - T_n \leq t | J_n = i\} \ . \tag{1.4}$$

If the probability in Eq. (1.4) does not depend on n, (J,T) *is time homogenous and its associated semi-Markov kernel* $\mathbf{q}$ *is defined by*

$$q_{ij}(t) := \mathbb{P}\{J_{n+1} = j, T_{n+1} - T_n = t \mid J_n = i\} \ . \tag{1.5}$$

The cumulated semi-Markov kernel associated to the Markov renewal chain (J,T) is defined by

$$Q_{ij}(t) := \mathbb{P}\{J_{n+1} = j, X_n \leq t \mid J_n = i\} = \sum_{k=0}^{t} q_{ij}(k) \ . \tag{1.6}$$

Remark 1.1.8. *Through this thesis we will only consider time homogenous processes, then, in what follow, we may omit to specify it.*

If (J,T) is a Markov renewal chain then $(J_n)_{n\in\mathbb{N}^*}$ is a Markov chain, called the embedded Markov chain associated to the Markov renewal chain (J,T). The Markov transition probability of J is defined by

$$p_{ij} := \mathbb{P}\{J_{n+1} = j \mid J_n = i\} = \sum_{t=0}^{\infty} q_{ij}(t) \ . \tag{1.7}$$

Definition 1.1.9. ***Distributions of the Waiting Times***

- *The cumulative conditional distribution of* X_n *is defined by*

$$G_{ij}(t) := \mathbb{P}\{X_n \leq t \mid J_n = i, J_{n+1} = j\} \ , \tag{1.8}$$

for every $n \in \mathbb{N}$ *and* $i,j \in E$.

- *The cumulative unconditional distribution of* X_n *is defined by*

$$H_i(t) := \mathbb{P}\{X_n \le t \mid J_n = i\} = \sum_{j\in E} Q_{ij}(t) \ , \tag{1.9}$$

for every $n \in \mathbb{N}$ *and* $i \in E$.

Remark 1.1.10. *We can express the conditional distribution in terms of the semi-Markov kernel as*

$$G_{ij}(t) = \frac{Q_{ij}(t)}{p_{ij}} \qquad \text{for } p_{ij} \neq 0 \ , \tag{1.10}$$

and we set $G_{ij}(t) = 1$ *for* $p_{ij} = 0$ *for every* $t \in \mathbb{N}$.

Definition 1.1.11. ***Semi-Markov Chain***

Let (J, T) *be a Markov renewal chain and* N *its associated counting process. The process* $Z = (Z(t))_{tin\mathbb{N}^*}$ *defined by*

$$Z(t) := J_{N(t)}, \tag{1.11}$$

is a semi-Markov chain associated to the Markov renewal chain (J, T)*. In other words* $Z(t)$ *gives the position of the embedded Markov chain at time* t.

An example of trajectory is shown in Figure 1.1. The evolution of a semi-Markov chain from an initial state can be studied by the associated transition probability.

Definition 1.1.12. ***Transition Probability***

The transition probability of the semi-Markov chain Z *is the matrix valued function* $\phi = (\phi_{ij}(t);\ i, j \in E,\ t \in \mathbb{N}^*)$ *defined by*

$$\phi_{ij}(t) := \mathbb{P}\{Z(t) = j \mid Z(0) = i, T_{N(0)} = 0\} \ . \tag{1.12}$$

The following result allows us to express the transition probability in a recursive way as a function of the semi-Markov kernel.

Proposition 1.1.13. ***Evolution Equation*** *(see for example Howard [52])*

For all $i, j \in E$ *and* $t \in \mathbb{N}$*, we have*

$$\Phi_{ij}(t) = \delta_{ij}[1 - H_i(t)] + \sum_{l\in E}\sum_{\tau=1}^{t} q_{il}(\tau)\Phi_{lj}(t-\tau) \ , \tag{1.13}$$

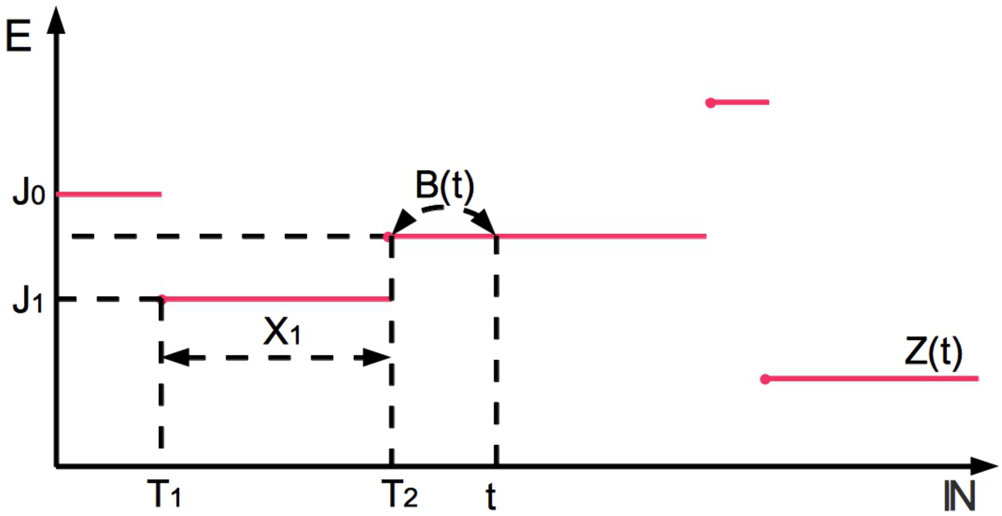


Figure 1.1: A semi-Markov trajectory is shown as a function of time. In the picture sojourn times, transition times and backward recurrence time are shown.

where δ_{ij} represents the Kronecker symbol.

The transition probability is expressed as a sum of two terms. The first is the probability to have no transition up to k while the second gives the probability to have at least one transition. This result completely defines the evolution of the semi-Markov chain and it allows us to solve numerically the process once the semi-Markov kernel is known (see Janssen and Manca [59] for details).

Semi-Markov processes are a very convenient way to describe phenomena which display a duration effect. The duration effect states that the time system spent in a state influence its transition probabilities. One way to detect and quantify it with semi-Markov processes is by using backward and forward recurrence time processes. Recurrence processes were analyzed in Janssen and Manca [60] and more recently in D'Amico et al. [34, 39].

Definition 1.1.14. ***Backward Recurrence Time***

The backward recurrence time process $B = (B(t))_{t\in\mathbb{N}}$ associated to the semi-Markov chain Z is defined by

$$B(t) := t - T_{N(t)}. \tag{1.14}$$

Intuitively, it gives the lifetime of the present state, in other words it is the time since the last transition.

In the semi-Markov evolution the Markovian memoryless property is preserved only at the transition times, i.e. the renewal moments. This feature makes the age of the state particularly important. As a consequence, the transition probabilities of a semi-Markov process change as a function of the values of the backward time. Indeed, the conditional waiting times distribution functions (1.8) can be of any kind and thus, also no memoryless distributions can be used. In this case the time length spent in the starting state (initial backward value) changes the transition probabilities.

The transition probability with initial backward is

$$ {}^b\phi_{ij}(u;t) := \mathbb{P}\{Z(t) = j \mid Z(0) = i, B(0) = u\} \ , $$

and denotes the probability of being in state j after t periods given that at present the process is in state i and it got into this state with the last transition u periods before.
Furthermore one might be interested to know the age of state at time t, i.e. the backward value at time t (final backward). The transition probability with initial and final backward is defined by

$$ {}^b\phi_{ij}(u;v,t) := \mathbb{P}\{Z(t) = j, B(t) = v \mid Z(0) = i, B(0) = u\} \ . $$

It is clear that

$$ {}^b\phi_{ij}(u;t) = \sum_{v\in\mathbb{N}} {}^b\phi_{ij}(u;v,t) \ . \tag{1.15} $$

We now find a recursive formula for calculating backward transition probability

Proposition 1.1.15. *(D'Amico* et al. *[35])*
For all $i, j \in E$, $u, v \in \mathbb{N}$ and $k \in \mathbb{N}$, we have

$$ {}^b\phi_{ij}(u;v,k) = \delta_{ij}\delta_{u+k,v}\frac{1-H_i(u+k)}{1-H_i(u)} + \sum_{l\in E}\sum_{\tau=1}^{k}\frac{q_{ik}(u+\tau)}{1-H_i(u)}{}^b\phi_{lj}(0;v,k-\tau). \tag{1.16} $$

If we sum on the value of the final backward v on both member of the above result we obtain a recursive formula for the initial backward transition probability.

Remark 1.1.16. *If we consider the joint process (Z, B), we record at any step the time already spent by the semi-Markov process in the present state, then it result that (Z, B) is a Markov process (*cf. *Anselone*

[1] and Limnios and Oprişan [65]). In other words, a Markov process is obtained when to the semi-Markov process is added the information regarding the permanence in the states.

1.2 Continuos Time Semi-Markov Process

In this section we summarize the main properties of continuos time processes, recalling the main definition of continuos time semi-Markov process with particular emphasis on the difference with the discrete time case. Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ be a complete filtered probability space, $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ with the usual conditions of completeness and right continuity, and let $E = \{1, \ldots, d\}$ be a given finite set.

Definition 1.2.1. ***Semi-Markov Kernel***

A matrix valued function $\mathbf{Q} = (Q_{ij}(t);\ i, j \in E,\ t \in \mathbb{R}_+)$ is a semi-Markov kernel if

- *for fixed $i, j \in E$, $t \to Q_{ij}(t)$ is a nondecreasing right continuos function and $Q_{ij}(0) = 0$;*
- *$t \to Q_{ij}(t)$ is a probability distribution function for every $i, j \in E$;*
- *$(\lim_{t\to\infty} Q_{ij}(t))_{i,j\in E}$ is a Markov transition probability.*

We define now Markov renewal chain in continuos time, i.e. Markov renewal chain with waiting times taking values in $\mathbb{R}_+$.

Definition 1.2.2. ***Continuos Time Markov Renewal Chain***

A random sequence $(J, T) = (J_n, T_n)_{n\in\mathbb{N}^*}$ is a continuos time Markov renewal chain if for all $n \in \mathbb{N}$, $i, j \in E$ and $t \in \mathbb{R}_+$ it satisfies

$$\mathbb{P}\{J_{n+1} = j, T_{n+1} - T_n \leq t | \sigma(J_a, T_a), 0 \leq a \leq n\} = \mathbb{P}\{J_{n+1} = j, T_{n+1} - T_n \leq t | J_n = i\} \ . \tag{1.17}$$

If the probability in (1.17) does not depend on n, (J, T) is time homogenous and its associated semi-Markov kernel $\mathbf{Q}$ is defined by

$$Q_{ij}(t) := \mathbb{P}\{J_{n+1} = j, T_{n+1} - T_n \leq t \mid J_n = i\} \ . \tag{1.18}$$

Definition 1.2.3. ***Counting Process***

The stochastic process $N = (N(t))_{t\in\mathbb{R}_+}$ defined by

$$N(t) = \sup\{n \in \mathbb{N} \mid T_n \leq t\} \tag{1.19}$$

is the counting process associated to the renewal process T. In other words, it gives at any time the number of events occurred.

Definition 1.2.4. ***Continuos Time Semi-Markov Process***

Let (J, T) be a Markov renewal chain and N its associated counting process. The process $Z = (Z(t))_{tin\mathbb{R}_+}$ defined by

$$Z(t) := J_{N(t)}, \tag{1.20}$$

is a semi-Markov process associated to the Markov renewal chain (J, T). In other words, $Z(t)$ gives the position of the embedded Markov chain at time t.

Definition 1.2.5. ***Transition Function***

The transition function of the semi-Markov process Z is the matrix valued function $\phi = (\phi_{ij}(t);\ i, j \in E,\ t \in \mathbb{R}_+)$ defined by

$$\phi_{ij}(t) := \mathbb{P}\{Z(t) = j \mid Z(0) = i, T_{N(0)} = 0\}\ . \tag{1.21}$$

Proposition 1.2.6. *(see for example Limnios and Oprişan [65])*

The transition function can be expressed as

$$\phi_{ij}(t) = \delta_{ij}(1 - H_i(t)) + \sum_{k\in I} \int_0^t \phi_{kj}(t-\theta) Q_{ik}(d\theta)\ . \tag{1.22}$$

The first term on the right hand side (r.h.s), $\delta_{ij}(1-H_i(t))$ gives the probability that the system does not have transitions up to time t given that it starts in state i at time 0; the second $\sum_{k\in I} \int_0^t \phi_{kj}(t-\theta)Q_{ik}(d\theta)$, takes into account the permanence of the system in state i up to the time θ where a transition in state k occurs. After the transition, the system will move to state j following one of all the possible trajectories going from state k to state j in the remaining time $t - \theta$. All possible states k and times θ are considered by the summation and the integration.

The backward recurrence time process is defined by

$$B(t) := t - T_{N(t)}. \tag{1.23}$$

As we already pointed out in the discrete time case, the transition probabilities of a semi-Markov process change as a function of the values of the backward time.

The transition function with initial backward is defined by

$$ {}^{b}\phi_{ij}(u;t) := \mathbb{P}\{Z(t) = j \mid Z(0) = i, B(0) = u\} \ . $$

Proposition 1.2.7. *(see for example Limnios and Oprişan [65])*

The backward transition function can be expressed as

$$ {}^{b}\phi_{ij}(u;t) = \delta_{ij}\frac{1 - H_i(u+t)}{1 - H_i(u)} + \sum_{k\in I}\int_0^t {}^{b}\phi_{ij}(0;t-\theta)\frac{Q_{ik}(u+d\theta)}{1 - H_i(u)} \ . \tag{1.24} $$

For our purposes we do not need to introduce the transition function with final backward, we refer to the book of Limnios and Oprişan [65] for more details.

Chapter 2

A Semi-Markov Modulated Interest Rate Model

In this Chapter we propose a semi-Markov modulated model of interest rates. We assume that the switching process is a semi-Markov process with finite state space E and the modulated process is a diffusive process.

Under these assumptions we derive recursive equations for the higher order moments of the discount factor and we describe a Monte Carlo algorithm to execute simulations. The results are specialized to classical models as those by Vasicek [84], Hull and White [54] and Cox, Ingersoll and Ross (CIR) [30] with a semi-Markov modulation.

The chapter is organized as follows. After a short introduction to the problem, we present the stochastic models of the short interest rates and we derive the main results concerning the equations for the higher order moments of the discount factor and for the covariance function of the force of interest. In the last section we apply our model to three well known diffusion models for interest rate and we present a Monte Carlo algorithm able to generate the synthetic data of the model.

This chapter is based on a paper (G. D'Amico, R. Manca and G. Salvi [41]) currently under review.

2.1 Introduction

The literature on interest rate models is ample and mainly concerns models based on short rate dynamics and models of forward rate (see e.g. Björk [9]). The advantages and the drawbacks of short rate models are well known in literature and even nowadays they continue to receive attention by researchers and practitioners.

The fundamental paper by Vasicek [84] is the pioneering contribution to short rate models. Different extensions and alternatives to the Vasicek model have been presented. One interesting approach in term-structure modeling is that proposed by Duffie and Kan [47], they assumed the short rate to be a function of a state process. Interesting developments were subsequently obtained by Mamon [66] who characterized the term structure of a Markov interest rate model when the interest rate process is assumed to be a function of a continuous time non-homogeneous Markov chain. By using the forward measure it was shown how to price term structure derivative products, see Mamon [67].

This paper assumes that the short rate is a diffusive process modulated by a continuous time semi-Markov process. The force of interest rates is defined by using the theory of semi-Markov reward processes with initial backward times, as developed by [80], and here opportunely extended to consider the case of stochastic permanence rewards. The resulting model is sufficiently general to be able to reproduce a great variety of interest rate evolutions. This is possible because the parameters of the diffusion are considered to be dependent on the state of the regime (semi-Markov process) and also on the time elapsed in the current regime (backward recurrence time process). Therefore the model is characterized by a rich parameter space, which consequently allows for model flexibility. It is worth noting that, since the Markov modulated model is an instance of the semi-Markov one, the latter will work at least as well as the former.

The results include renewal type equations for the higher order moments of the discount factor (DF) and for the covariance function of the force of interest. Notice that, in the paper by Hunt and Devolder [55], a discrete time regime switching binomial-like model of the term structure, where the regime switches are governed by a discrete time semi-Markov process, is presented. As reported by Hunt and Devolder [55], the semi-Markov modulated model offers a solution to some of the drawbacks of the Markovian switching models. The most important inadequacies are represented by the memoryless property of Markov processes and by the rather unrealistic hypothesis of constant transition intensities for the Markovian switching process for interest rate data (see Dahlquist and Gray [44]). Additional advantages of the semi-Markov approach can be found in Hunt and Devolder [55].

We propose a finite state space continuous time semi-Markov modulated model as opposed to the binomial-like discrete time regime switching model of Hunt and Devolder [55]. There are different reasons motivating our choice. First of all, the continuous time model avoids necessitating the selection of a discrete time scale that, in general, depends on the observation frequency. This, in turn, avoids the problem of temporal aggregation when estimating model parameters, which may cause inconsistent estimations of the short-term interest rate (see Broze *et al.* [18]). The second reason is that very often regimes are considered as states of the economy and although economic measurements occur in discrete periods (e.g. monthly or quarterly) the economy operates continuously in time.

It should be highlighted that our paper doesn't discuss the problem of transformation between the real-world (physical) measure and the pricing measure and therefore the proposed model could not be adopted for pricing interest derivative products, but it is suitable for all applications that require real-world evolution of rates (see Rebonato *et al.* [74] for a complete list). It should be mentioned that determination of real-world evolution of rates should not to be considered as an alternative to sampling the risk-adjusted measure, but rather as a complement (see Rebonato *et al.* [74]).

Particular cases of our model proved appropriate to solve problems with applications to insurance (see Norberg [71]). In this respect, the semi-Markov modulated interest rate model can be useful in studying problems of actuarial mathematics such as, for example, the evolution of mathematical reserves linked to life insurance contracts, where the adoption of simplified and stylized interest rates models can determine misleading results. The mis-modeling of interest rates is a serious problem, which impacts on assets, liabilities and surplus levels (see Wang and Huang [86]).

2.2 The Model

In this section we define a semi-Markov modulated model of interest rates and we assess its probabilistic behavior.

In the following we assume that the force of interest, at any time t, is a stochastic process of diffusive type whose parameters depend on the state of the semi-Markov process, on the backward recurrence time process and on the initial value of the force of interest r_0. To be more precise we assume that within two transition times T_{n-1} and T_n of the semi-Markov process the evolution of the force of interest follows the dynamic of a diffusive process whose parameters depends on the state J_{n-1} of the semi-Markov process.

The dynamic of the force of interest, between two consecutive renewal moments, will be indicated as

$$dr(t) = b_i(r(t), t)dt + \sigma_i(r(t), t)dW(t), \;\; r(0) = r_0 \;, \tag{2.1}$$

where it is supposed that $T_0 = 0$, $T_1 > t$ and $J_0 = i$. The process $W(t)$ is a Brownian motion with respect to its own filtration $\mathcal{F}_t^W$, and we denote the drift and diffusion coefficients with b and σ, respectively. The solution of the stochastic differential equation (2.1) will be denoted by $r_{i,r_0}(s)$, for $s \in [T_0, T_1)$. In T_1 the semi-Markov process transits to another state, say j, the force of interest in the time interval since T_1 to the next transition will evolve according to

$$dr(t) = b_j(r(t), t)dt + \sigma_j(r(t), t)dW(t), \;\; r(T_1) = r_{i,r_0}(T_1) \;, \tag{2.2}$$

for $t \in [T_1, T_2)$, and so on. Therefore, the resulting force of interest will be a continuous process.
We would like to stress that for any $s \in [T_0, T_1)$, and in general between any two transition times, the solution $r_{i,r_0}(s)$ is obtained with standard stochastic calculus methods.

Due to the fact that the force of interest depends on the modulating process, we need to describe the force of interest process at any time s given the information available at the present time, which is time zero as long as we are working with a homogenous time model, and this is represented by the triplet of values $\{Z(0) = i, B(0) = u, r(0) = r_0\}$.

Definition 2.2.1. *The force of interest at the generic time s is defined as:*

$$\delta_{i,u,r_0}(s) \stackrel{d}{=} \chi(T_1 > s | J_0 = i, T_0 = -u, T_1 > 0) r_{i,r_0}(s) + \chi(T_1 \le s | J_0 = i, T_0 = -u, T_1 > 0) \delta_{J_1, 0, r_{i,r_0}(T_1)}(s - T_1). \tag{2.3}$$

Here $\chi(A \mid B)$ is the indicator function of set A given the information B and the symbol $\stackrel{d}{=}$ stands for the equality in distribution.

Remark 2.2.2. *The processes $\chi(T_1 > s | J_0 = i, T_0 = -u, T_1 > 0)$ and $r_{i,r_0}(s)$ are independent for any $s \in \mathbb{R}_+$ and $i \in I$.*

Remark 2.2.3. *The process* $\chi(s < T_1 < s+h|J_0 = i, T_0 = -u, T_1 > 0)\delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1)$ *conditioning to the value of* $r_{i,r_0}(T_1)$ *is independent of* $r_{i,r_0}(s)$.

We are interested in the DF process $v_{i,u,r_t}(t,T) := exp(-\int_t^T \delta_{i,u,r_t}(s)ds)$ expressing the value of 1 Euro at time T, given that at current time t the semi-Markov process is in the state i and it has entered this state u periods before and the force of interest at that time is r_t. To this end, we would like to consider only diffusion processes in the following that allow an explicit representation of the Laplace transform of $\int_0^t r_{i,r_0}(s)ds$ for example, Vasicek [84], Hull and White [54] and Cox, Ingersoll and Ross (1985) (CIR) model.

Theorem 2.2.4. *Let*

$$V_{i,u,r_t}^{(n)}(t,T) = \mathbb{E}[(v_{i,u,r_t}(t,T))^n] \ , \tag{2.4}$$

be the nth order moment of the DF process, then it results that

$$\begin{aligned} V_{i,u,r_t}^{(n)}(t,T) &= \frac{1-H_i(T-t+u)}{1-H_i(u)}\left(B_{i,r_t}^{(n)}(t,T)\right) \\ &+ \sum_{k\in I}\int_t^T \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\left(B_{i,r_t}^{(n)}(t,\tau)\int_{-\infty}^{+\infty} V_{k,0,x}^{(n)}(\tau,T)F_{r_{i,r_t}(\tau)}(dx)\right)d\tau \ , \end{aligned} \tag{2.5}$$

where $B_{i,r_t}^{(n)}$ *is defined by*

$$B_{i,r_t}^{(n)}(t,T) \doteq \mathbb{E}\left\{exp\left(-n\int_t^T r_{i,r_t}(s)ds\right)\right\} \ . \tag{2.6}$$

Proof. Let us consider that at the time t, the process $v_{i,u,r_t}(t,T) := exp(-\int_t^T \delta_{i,u,r_t}(s)ds)$ and the condition $T_{N(t)+1}$ is the time of the next switching process transition. We can partition the state space into two possible events: $\{T_{N(t)+1} > T\}$ or $\{T_{N(t)+1} \leq T\}$.

The first event $\{T_{N(t)+1} > T\}$ corresponds to the possibility of having no transition up to the time T and it has a probability $\frac{1-H_i(T-t+u)}{1-H_i(u)}$. Under this event, the force of interest is given by

$$\delta_{i,u,r_t}(s) = r_{i,r_t}(s) \qquad\qquad \text{for } s \in [t,T] \ . \tag{2.7}$$

In this case, the DF process assumes the value

$$v_{i,u,r_t}(t,T)\Big|_{T_{N(t)+1}>T} = \exp\left(-\int_t^T r_{i,r_t}(s)ds\right), \tag{2.8}$$

taking now the expectation that

$$\mathbb{E}\{(v_{i,u,r_t}(t,T))^n | T_{N(t)+1} > T\} = B_{i,r_t}^{(n)}(t,T) \ . \tag{2.9}$$

Therefore, we have

$$\begin{aligned}&\mathbb{E}\{(v_{i,u,r_t}(t,T))^n\chi(T_{N(t)+1}>T|J_{N(t)}=i,T_{N(t)}=t-u,T_{N(t)+1}>t,r(t)=r_t)\}\\=&\ \frac{1-H_i(T-t+u)}{1-H_i(u)}B^{(n)}_{i,r_t}(t,T)\ .\end{aligned} \tag{2.10}$$

The other event $\{T_{N(t)+1}\in(t,T],J_{N(t)+1}\in I\}$ corresponds to the possibility of having at least one transition in the considered time interval. The probability that the semi-Markov process has the first transition in the time interval $(\tau,\tau+d\tau)$, for $\tau\in(t,T)$, into the state $k\in I$, is given by

$$P(J_{N(t)+1}=k,T_{N(t)+1}\in(\tau,\tau+d\tau)|\mathcal{F}^{SM}_t,Z(t)=i,B(t)=u)=\frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}d\tau\ ,$$

where $(\mathcal{F}^{SM}_t)_t$ is the filtration generated by the process (Z,B), that is the semi-Markov and the backward recurrence time together. To properly evaluate the expectation under this event, we have to consider all the possible values that the force of interest can assume on transition time τ. Denoting $F_{r_{i,r_t}(\tau)}$ as the cumulative distribution function associated to $r_{i,r_t}(\tau)$, and using the continuity property of the force of interest, we have

$$P(J_{N(t)+1}=k,T_{N(t)+1}\in(\tau,\tau+d\tau),\delta_{i,u,r_t}(\tau)\in(x,x+dx)|\mathcal{F}_t)=\frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}F_{r_{i,r_t}(\tau)}(dx)d\tau\ ,$$

where $\mathcal{F}_t$ is the filtration generated by (Z,B,W). Under the event $A=\{J_{N(t)+1}=k,T_{N(t)+1}\in(\tau,\tau+d\tau),\delta_{i,u,r_t}(\tau)\in(x,x+dx),\mathcal{F}_t\}$ the DF process can be expressed as

$$v_{i,u,r_t}(t,T)\Big|_A=exp\left(-\int_t^T\delta_{i,u,r_t}(s)ds\right)\Big|_A=exp\left(-\int_t^\tau r_{i,r_t}(s)ds\right)exp\left(-\int_\tau^T\delta_{k,0,x}(s)ds\right)\Big|_A\ .$$

Moreover from Remark 2.2.3, we have

$$\begin{aligned}&\mathbb{E}\{(v_{i,u,r_t}(t,T))^n|A\}=\mathbb{E}\left\{exp\left(-n\int_t^\tau r_{i,r_t}(s)ds\right)exp\left(-n\int_\tau^T\delta_{k,0,x}(s)ds\right)\Big|A\right\}\\=&\ \mathbb{E}\left\{exp\left(-n\int_t^\tau r_{i,r_t}(s)ds\right)\Big|A\right\}\mathbb{E}\left\{exp\left(-n\int_\tau^T\delta_{k,0,x}(s)ds\right)\Big|A\right\}=B^{(n)}_{i,r_t}(t,\tau)V^{(n)}_{k,0,x}(\tau,T)\ .\end{aligned}$$

and consequently

$$\begin{aligned}&\mathbb{E}\{(v_{i,u,r_t}(t,T))^n\chi(t<T_{N(t)+1}\le T|J_{N(t)}=i,T_{N(t)}=t-u,T_{N(t)+1}>t,r(t)=r_t)\}\\&=\sum_{k\in I}\int_t^T\frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\left(B^{(n)}_{i,r_t}(t,\tau)\int_{-\infty}^{+\infty}V^{(n)}_{k,0,x}(\tau,T)F_{r_{i,r_t}(\tau)}(dx)\right)d\tau\ .\end{aligned} \tag{2.11}$$

The value of the discount factor is given by

$$\mathbb{E}\{(v_{i,u,r_t}(t,T))^n\} = \mathbb{E}\{(v_{i,u,r_t}(t,T))^n \chi(T_{N(t)+1} > T | J_{N(t)} = i, T_{N(t)} = t-u, T_{N(t)+1} > t, r(t) = r_t)\}$$
$$+ \mathbb{E}\{(v_{i,u,r_t}(t,T))^n \chi(t < T_{N(t)+1} \leq T | J_{N(t)} = i, T_{N(t)} = t-u, T_{N(t)+1} > t, r(t) = r_t)\}, \tag{2.12}$$

then by substitution of expressions (2.10) and (2.11) into (2.12) the proof is complete. □

Corollary 2.2.5. *Lets represent* $R_{i,u,r_t}(s) = \mathbb{E}[\delta_{i,u,r_t}(s)]$, *then it results that*

$$R_{i,u,r_t}(s) = \frac{1-H_i(s+u)}{1-H_i(u)} m_{i,r_t}(t,s) + \sum_{k\in I} \int_0^s \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)} \left(\int_{-\infty}^{+\infty} R_{k,0,x}(s-\tau) F_{r_i,r_t(\tau)}(dx) \right) d\tau \ , \tag{2.13}$$

here, $m_{i,r_t}(s) := \mathbb{E}[r_{i,r_t}(s)]$.

We can derive the following result that helps us obtain the covariance function of the force of interest.

Theorem 2.2.6. *Let* $s > 0$ *and* $h > 0$, *and*

$$\Xi_{i,u,r_0}(s,h) = \mathbb{E}[\delta(s)\delta(s+h) | Z(0) = i, B(0) = u, \delta(0) = r_0] \ . \tag{2.14}$$

Then it results that

$$\begin{aligned} \Xi_{i,u,r_0}(s,h) &= \frac{1-H_i(s+h+u)}{1-H_i(u)} \rho_{i,r_0}(s,s+h) \\ &+ \sum_{k\in I} \int_s^{s+h} \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)} m_{i,r_0}(s) \int_{-\infty}^{+\infty} R_{k,0,x}(s+h-\tau) F_{i,r_0(\tau)}(dx) d\tau \\ &+ \sum_{k\in I} \int_0^{s} \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)} \int_{-\infty}^{+\infty} \Xi_{k,0,x}(s-\tau, s+h-\tau) F_{i,r_0(\tau)}(dx) d\tau \ , \end{aligned} \tag{2.15}$$

here, $\rho_{i,r_0}(s,s+h) := \mathbb{E}[r_{i,r_0}(s) r_{i,r_0}(s+h)]$.

Proof. Using the definition of the force of interest process δ, cf. formula (2.3), we have that

$$\begin{aligned} \delta_{i,u,r_0}(s)\delta_{i,u,r_0}(s+h) &\stackrel{d}{=} \{\chi(T_1 > s | J_0 = i, T_0 = -u, T_1 > 0) r_{i,r_0}(s) \\ &+ \chi(T_1 \leq s | J_0 = i, T_0 = -u, T_1 > 0)[\delta_{J_1,0,r_{i,r_0}(T_1)}(s - T_1)]\} \\ &\times \left\{\chi(T_1 > s+h | J_0 = i, T_0 = -u, T_1 > 0) r_{i,r_0}(s+h) \right. \\ &+ \left. \chi(T_1 \leq s+h | J_0 = i, T_0 = -u, T_1 > 0)[\delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1)]\right\} \ , \end{aligned} \tag{2.16}$$

which can be simplified to

$$\begin{aligned} \delta_{i,u,r_0}(s)\delta_{i,u,r_0}(s+h) &= \chi(T_1 > s+h | J_0 = i, T_0 = -u, T_1 > 0) r_{i,r_0}(s) r_{i,r_0}(s+h) \\ &+ \chi(s < T_1 \leq s+h | J_0 = i, T_0 = -u, T_1 > 0) r_{i,r_0}(s) \delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1) \\ &+ \chi(T_1 \leq s | J_0 = i, T_0 = -u, T_1 > 0) \delta_{J_1,0,r_{i,r_0}(T_1)}(s-T_1) \delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1) \ , \end{aligned} \tag{2.17}$$

The first term on the right hand side of (2.17) corresponds to the possibility that the semi-Markov process has no transition up to time $s+h$, the second term considers the case when the first transition occurs during the time interval $(s, s+h)$ and the third takes into account the possibility of the first transition occurring before s.

Taking expectations of the first term on the right hand side (r.h.s.) of (2.17) and noting that (cf. Remark 2.2.2) the random variable $\chi(T_1 > s+h|J_0 = i, T_0 = -u, T_1 > 0)$ is independent from, both, $r_{i,r_0}(s)$ and $r_{i,r_0}(s+h)$ yields,

$$\begin{aligned}
&\mathbb{E}[\chi(T_1 > s+h|J_0 = i, T_0 = -u, T_1 > 0)r_{i,r_0}(s)r_{i,r_0}(s+h)] \\
&= \mathbb{E}[\chi(T_1 > s+h|J_0 = i, T_0 = -u, T_1 > 0)]\mathbb{E}[r_{i,r_0}(s)r_{i,r_0}(s+h)] \\
&= \frac{1 - H_i(s+h+u)}{1 - H_i(u)}\rho_{i,r_0}(s, s+h) \ .
\end{aligned} \tag{2.18}$$

Let us consider the second term on the r.h.s. of (2.17). By taking the expectation and by conditioning on the value of $r_{i,r_0}(T_1)$, we can use the independence between $r_{i,r_0}(s)$ and $\chi(s < T_1 \leq s+h|J_0 = i, T_0 = -u, T_1 > 0)\delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1)$ (cf. Remark 2.2.3) to obtain

$$\begin{aligned}
&\mathbb{E}\{\mathbb{E}[\chi(s < T_1 \leq s+h|J_0 = i, T_0 = -u, T_1 > 0)r_{i,r_0}(s)\delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1)|r_{i,r_0}(T_1) = x]\} \quad (2.19) \\
&= \mathbb{E}\{\mathbb{E}[\chi(T_1 > s+h|J_0 = i, T_0 = -u, T_1 > 0)\delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1)|r_{i,r_0}(T_1) = x]\}\mathbb{E}[r_{i,r_0}(s)] \\
&= \textstyle\sum_{k\in I}\int_s^{s+h} \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\mathbb{E}[r_{i,r_0}(s)]\int_{-\infty}^{+\infty}\mathbb{E}[\delta_{k,0,x}(s+h-\tau)]F_{i,r_0(\tau)}(dx)d\tau \ ,
\end{aligned}$$

since

$$E[r_{i,r_0}(s)] = m_{i,r_0}(s) \ , \tag{2.20}$$

and

$$\mathbb{E}[\delta_{k,0,x}(s+h-\tau)] = R_{k,0,x}(s+h-\tau) \ , \tag{2.21}$$

then, by substitution, we get

$$\begin{aligned}
&E[\chi(s < T_1 \leq s+h|J_0 = i, T_0 = -u, T_1 > 0)r_{i,r_0}(s)\delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1)] \\
&= \textstyle\sum_{k\in I}\int_s^{s+h} \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}m_{i,r_0}(s)\int_{-\infty}^{+\infty}R_{k,0,x}(s+h-\tau)F_{i,r_0(\tau)}(dx)d\tau \ .
\end{aligned} \tag{2.22}$$

Finally, taking the expectation of the third term on the r.h.s. of (2.17) we obtain

$$E[\chi(T_1 \leq s | J_0 = i, T_0 = -u, T_1 > 0)\delta_{J_1,0,r_{i,r_0}(T_1)}(s-T_1)\delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1)] \tag{2.23}$$
$$= \sum_{k\in I}\int_0^s \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\mathbb{E}[\delta_{k,0,r_{i,r_0}(\tau)}(s-\tau)\delta_{k,0,r_{i,r_0}(\tau)}(s+h-\tau)]d\tau \ ,$$

and, by considering all possible values of $r_{i,r_0}(\tau)$, we get

$$E[\chi(T_1 \leq s | J_0 = i, T_0 = -u, T_1 > 0)\delta_{J_1,0,r_{i,r_0}(T_1)}(s-T_1)\delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1)] \tag{2.24}$$
$$= \sum_{k\in I}\int_0^s \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\Big(\int_{-\infty}^{+\infty}\mathbb{E}[\delta_{k,0,x}(s-\tau)\delta_{k,0,x}(s+h-\tau)]F_{i,r_0(\tau)}(dx)\Big)d\tau \ ,$$

but

$$\mathbb{E}[\delta_{k,0,x}(s-\tau)\delta_{k,0,x}(s+h-\tau)] = \Xi_{k,0,x}(s-\tau, s+h-\tau) \ , \tag{2.25}$$

then

$$E[\chi(T_1 \leq s | J_0 = i, T_0 = -u, T_1 > 0)\delta_{J_1,0,r_{i,r_0}(T_1)}(s-T_1)\delta_{J_1,0,r_{i,r_0}(T_1)}(s+h-T_1)] \tag{2.26}$$
$$= \sum_{k\in I}\int_0^s \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\Big(\int_{-\infty}^{+\infty}\Xi_{k,0,x}(s-\tau, s+h-\tau)F_{i,r_0(\tau)}(dx)\Big)d\tau \ .$$

Therefore, in order to determine the expectation of (2.17), it is sufficient to sum (2.18), (2.22) and (2.26). □

2.3 Applications to Some Known Diffusion Models

So far, we have not assumed any specific dynamic driving the force of interest between two consecutive renewal moments. In this section we would like to consider some extensively applied diffusion models for the force of interest such as those proposed by Vasicek [84], Hull and White [54] and CIR [30]. For these models, as it is well known, the Laplace transformation of the integral of the force of interest, r, has an explicit representation. As we will see, our general results specialized for these particular cases.

Without loss of generality, since we are working with an homogeneous time model, we can suppose that the present time is $t = 0$, with $J_0 = i$, and study the dynamic of the force of interest from 0 up to a generic instant $t < T_1$.

2.3.1 Vasicek Modulated Model

We assume that the process $r_{i,r_0}(t)$, satisfies

$$\begin{cases} dr_{i,r_0}(t) = a_i(b_i - r_{i,r_0}(t))dt + \sigma_i dW_t \\ r_{i,r_0}(0) = r_0 \end{cases} ,$$

where a_i, b_i and σ_i, for fixed $i \in I$, are non-negative constants. The solution of the previous stochastic differential equation is

$$r_{i,r_0}(t) = b_i + (r_0 - b_i)e^{-a_i t} + \sigma_i e^{-a_i t} \int_0^t e^{a_i s} dW_s \ . \tag{2.27}$$

For any fixed t, $r_{i,r_0}(t)$ is normally distributed, $r_{i,r_0}(t) \sim \mathcal{N}(m_{i,r_0}(t), \sigma^2_{i,r_0}(t))$ with mean

$$m_{i,r_0}(t) = \mathbb{E}[r_{i,r_0}(t)] = b_i + (r_0 - b_i)e^{-a_i t} \ , \tag{2.28}$$

and variance

$$\sigma^2_{i,r_0}(t) = \frac{\sigma_i^2}{2a_i}(1 - e^{-2a_i t}) \ . \tag{2.29}$$

It follows that (see Lamberton and Lapeyre [63]) $\int_0^t r(s)ds$ is a normal random variable, since the integral can be written as the limit of Riemann sums, which are Gaussians. Then we can completely characterize the distribution of $\int_0^t r_{i,r_0}(s)ds$ with its mean and variance. The mean, using Fubini's Theorem, is given by

$$\begin{aligned} \mathbb{E}\left[\int_0^t r_{i,r_0}(s)ds\right] &= \int_0^t \mathbb{E}[r_{i,r_0}(s)]ds = \int_0^t m_{i,r_0}(s)ds \\ &= \int_0^t [b_i + (r_0 - b_i)e^{-a_i s}]ds = b_i t + \frac{r_0 - b_i}{a_i}(1 - e^{-a_i t}) \ . \end{aligned} \tag{2.30}$$

The variance can be expressed as

$$\begin{aligned} Var\left[\int_0^t r_{i,r_0}(s)ds\right] &= Var\left[\int_0^t \left(b_i + (r_0 - b_i)e^{-a_i s} + \sigma_i e^{-a_i s}\int_0^s e^{a_i u} dW_u\right) ds\right] \\ &= Var\left[\int_0^t \left(\sigma_i e^{-a_i s}\int_0^s e^{a_i u} dW_u\right) ds\right], \end{aligned} \tag{2.31}$$

which can be written as

$$Var\left[\int_0^t r_{i,r_0}(s)ds\right] = Var\left[\int_0^t e^{a_i u}\left(\sigma_i \int_u^t e^{-a_i s} ds\right) dW_u\right] . \tag{2.32}$$

Using Ito's isometry we get

$$Var\left[\int_0^t r_{i,r_0}(s)ds\right] = \sigma_i^2 \int_0^t e^{2a_i u}\left(\int_u^t e^{-a_i s}ds\right)^2 du = \sigma_i^2 \int_0^t e^{2a_i u}\left(\frac{e^{-a_i u} - e^{-a_i t}}{a_i}\right)^2 du \ .$$

This is a standard integral, and we can simply solve it using the ordinary techniques of calculus obtaining

$$Var\left[\int_0^t r_{i,r_0}(s)ds\right] = \frac{\sigma_i^2 t}{a_i^2} - \frac{\sigma_i^2}{a_i^3}(1-e^{-a_i t}) - \frac{\sigma_i^2}{2a_i^3}(1-e^{-a_i t})^2 \ . \tag{2.33}$$

Now, we can give an analytical expression of the nth moment of a DF between two consecutive renewal moments, $B^{(n)}_{i,r_0}(0,t)$ (cf. Theorem 2.2.4), defined by

$$B^{(n)}_{i,r_0}(0,t) = \mathbb{E}\left[exp\{-n\int_0^t r_{i,r_0}(t)dt\}\right] \ , \tag{2.34}$$

which can be seen as the Laplace transformation of the normal random variable $\int_0^T r_{i,r_0}(t)dt$ and then it can be expressed as

$$B^{(n)}_{i,r_0}(0,t) = exp\left\{-n\mathbb{E}\left[\int_0^t r_{i,r_0}(t)dt\right] + \frac{n^2}{2}Var\left[\int_0^t r_{i,r_0}(t)dt\right]\right\}. \tag{2.35}$$

Substituting in the mean and variance expressions we get

$$B^{(n)}_{i,r_0}(0,t) = exp\left\{\left(\frac{\sigma_i^2 n^2}{a_i^2} - nb_i\right)t - \left(\frac{\sigma_i^2 n^2}{a_i^3} + \frac{n(r_0-b_i)}{a_i}\right)\left(1-e^{-a_i t}\right) - \frac{\sigma_i^2 n^2}{2a_i^3}(1-e^{-a_i t})^2\right\}. \tag{2.36}$$

This way, we are able to express all the moments of the DF in any time interval.

Corollary 2.3.1. *The nth order moment of the discount factor in the Vasicek model modulated by the semi-Markov process of kernel* $\mathbf{Q}$ *is*

$$\begin{aligned} V^{(n)}_{i,u,r_t}(t,T) &= \mathbb{E}\left[exp(-n\int_t^T \delta_{i,u,r_t}(s)ds)\right] = \frac{1-H_i(T-t+u)}{1-H_i(u)}\big(B^{(n)}_{i,r_t}(t,T)\big) \\ &+ \sum_{k\in I}\int_t^T \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\left(B^{(n)}_{i,r_t}(t,\tau)\int_{-\infty}^{+\infty} f_{\mathcal{N}(m_{i,r_t}(\tau-t),\sigma^2_{i,r_t}(\tau-t))}(x)V^{(n)}_{k,0,x}(\tau,T)dx\right)d\tau. \end{aligned} \tag{2.37}$$

where $B^{(n)}_{i,r_t}(t,T)$ *is given in equation (??) and* $f_{\mathcal{N}(m_{i,r_t}(\tau-t),\sigma^2_{i,r_t}(\tau-t))}$ *is the probability distribution function of a normal random variable with mean* $m_{i,r_t}(\tau-t)$ *and variance* $\sigma^2_{i,r_t}(\tau-t)$ *(cf. eq. (2.28) and (2.29), respectively).*

Corollary 2.3.2. *The first moment of the force of interest of the Vasicek model modulated by the semi-Markov process of kernel* $\mathbf{Q}$ *is*

$$\begin{aligned} R_{i,u,r_t}(s) &= \mathbb{E}[\delta_{i,u,r_t}(s)] = \frac{1-H_i(s+u)}{1-H_i(u)}\big(b_i + (r_t-b_i)e^{-a_i s}\big) \\ &+ \sum_{k\in I}\int_t^s \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\int_{-\infty}^{+\infty} f_{\mathcal{N}(m_{i,r_t}(\tau),\sigma^2_{i,r_t}(\tau))}(x)R_{k,0,x}(s-\tau)dxd\tau. \end{aligned} \tag{2.38}$$

where $f_{\mathcal{N}(m_{i,r_t}(\tau),\sigma^2_{i,r_t}(\tau))}$ *is the probability distribution function of a normal random variable with mean* $m_{i,r_t}(\tau)$ *and variance* $\sigma^2_{i,r_t}(\tau)$ *(cf. eq. (2.28) and (2.29), respectively) .*

Corollary 2.3.3. *The product moment of the force of interest of the Vasicek model modulated by the semi-Markov process of kernel* $\mathbf{Q}$ *is*

$$\begin{aligned}\Xi_{i,u,r_0}(s,h) &= \mathbb{E}[\delta(s)\delta(s+h)|Z(0)=i,B(0)=u,\delta(0)=r_0] = \frac{1-H_i(s+h+u)}{1-H_i(u)}\rho_{i,r_0}(s,s+h)\\ &+\sum_{k\in I}\int_s^{s+h}\frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\big(b_i+(r_0-b_i)e^{-a_is}\big)\int_{-\infty}^{+\infty}f_{\mathcal{N}(m_{i,r_0}(\tau),\sigma^2_{i,r_0}(\tau))}(x)R_{k,0,x}(s+h-\tau)dxd\tau\\ &+\sum_{k\in I}\int_0^{s}\frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\int_{-\infty}^{+\infty}f_{\mathcal{N}(m_{i,r_0}(\tau),\sigma^2_{i,r_0}(\tau))}(x)\Xi_{k,0,x}(s-\tau,s+h-\tau)dxd\tau\end{aligned} \tag{2.39}$$

where $\rho_{i,r_0}(s,s+h)$ *is given by*

$$\rho_{i,r_0}(s,s+h)=\mathbb{E}[r_{i,r_0}(s)r_{i,r_0}(s+h)] = m_{i,r_t}(s)m_{i,r_t}(s+h)+\frac{\sigma_i^2}{2a_i}e^{-a_ih}(1-e^{-2a_is}), \tag{2.40}$$

and $f_{\mathcal{N}(m_{i,r_t}(\tau),\sigma^2_{i,r_t}(\tau))}$ *is the distribution function of a normal random variable with mean* $m_{i,r_t}(\tau)$ *and variance* $\sigma^2_{i,r_t}(\tau)$ *(cf. eq. (2.28) and (2.29), respectively) .*

Monte Carlo Algorithm

We conclude the discussion with a Monte Carlo algorithm able to generate the trajectories of the Vasicek model modulated by a semi-Markov process of kernel $\mathbf{Q}$ in the time interval $[0,T]$. The algorithm consists in repeated random sampling to compute successive visited states of the random variables $\{J_0, J_1, ...\}$, the jump times $\{T_0, T_1, ...\}$ and the force of interest process $r(t)$ up to the time T.

The algorithm consists of 5 steps:

1. Set $n=0$, $J_0=i$, $T_0=0$, $r(0)=r_0$, horizon time= T; discretization step= h
2. Sample J from $p_{J_n,\cdot}$ and set $J_{n+1}=J(\omega)$;
3. Sample W from $G_{J_n,J_{n+1}}(\cdot)$ and set $T_{n+1}=T_n+W(\omega)$;
4. For each $i=T_n:h:\max(T,T_{k+1}-1)$

   Sample N from $\mathcal{N}(0,h)$

   Set $r(i+1)=b_{J_n}+e^{-a_{J_n}}(r(i)-b_{J_n})+e^{-a_{J_n}}\sigma_{J_n}N(i+1)$

5. if $T_{n+1} \geq T$ stop

   else Set $n = n + 1$ and go to 2).

2.3.2 Hull and White modulated model

Let's now assume that the process $r_{i,r_0}(t)$ satisfies

$$\begin{cases} dr_{i,r_0}(t) = (\alpha_i(t) - \beta_i(t) r_{i,r_0}(t))dt + \sigma_i(t)dW_t \\ r_{i,r_0}(0) = r_0 \end{cases}$$

where α_i, β_i and σ_i are deterministic functions of time. The solution of this stochastic differential equation can be expressed (see Shreve [78]) as

$$r_{i,r_0}(t) = e^{-k_i(t)} \left[r_0 + \int_0^t e^{k_i(u)} \alpha_i(u)du + \int_0^t e^{k_i(u)} \sigma_i(u)dW(u) \right], \tag{2.41}$$

where

$$k_i(t) = \int_0^t \beta_i(u)du. \tag{2.42}$$

Note that $(r_{i,r_0}(t))_{t \in [0,T]}$ is a Gaussian process whose mean is given by

$$m_{i,r_0}(t) = e^{-k_i(t)} \left[r_0 + \int_0^t e^{k_i(u)} \alpha_i(u)du \right], \tag{2.43}$$

and its variance is given by

$$\sigma^2_{i,r_0}(t) = e^{-2k_i(t)} \int_0^t e^{2k_i(u)} \sigma_i(u)^2 du. \tag{2.44}$$

Moreover the process $(\int_0^t r_{i,r_0}(s)ds)_{t \in [0,T]}$ is Gaussian as well, with mean

$$\mathbb{E}\left[\int_0^t r_{i,r_0}(s)ds \right] = \int_0^t e^{-k_i(t)} \left[r_0 + \int_0^s e^{k_i(u)} \alpha_i(u)du \right] ds, \tag{2.45}$$

and variance

$$Var\left(\int_0^t r_{i,r_0}(s)ds \right) = \int_0^t e^{2k_i(u)} \sigma_i^2(u) \left(\int_0^t e^{-k_i(s)} ds \right)^2 du. \tag{2.46}$$

Then, we are able to express the n^{th} moment of a DF value between two consecutive renewal moment in an analytical form

$$B^{(n)}_{i,r_0}(0,t) = \mathbb{E}\left[exp\{-n \int_0^T r_{i,r_0}(t)dt\} \right] = exp\left\{ -\mathbb{E}\left[\int_0^T r_{i,r_0}(t)dt \right] + \frac{1}{2} Var\left[\int_0^T r_{i,r_0}(t)dt \right] \right\}. \tag{2.47}$$

We can now express all the moments of the DF in any time interval.

Corollary 2.3.4. *The nth order moment of the discount factor in the Hull and White model modulated by the semi-Markov process of kernel* $\mathbf{Q}$ *is*

$$\begin{aligned} V^{(n)}_{i,u,r_t}(t,T) = \mathbb{E}\left\{ exp(-n\int_t^T \delta_{i,u,r_t}(s)ds)\right\} = \frac{1-H_i(T-t+u)}{1-H_i(u)}\big(B^{(n)}_{i,r_t}(t,T)\big) \\ +\sum_{k\in I}\int_t^T \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\left(B^{(n)}_{i,r_t}(t,\tau)\int_{-\infty}^{+\infty} f_{\mathcal{N}(m_{i,r_t}(\tau-t),\sigma^2_{i,r_t}(\tau-t))}(x)V^{(n)}_{k,0,x}(\tau,T)dx\right)d\tau \ , \end{aligned} \tag{2.48}$$

where $B^{(n)}_{i,r_t}(t,T)$ *is given in equation (2.47) and* $f_{\mathcal{N}(m_{i,r_t}(\tau-t),\sigma^2_{i,r_t}(\tau-t))}$ *is the probability distribution function of a Normal distribution with mean* $m_{i,r_t}(\tau-t)$ *and variance* $\sigma^2_{i,r_t}(\tau-t)$ *(cf. eq. (2.43) and (2.44) respectively).*

Corollary 2.3.5. *The first moment of the force of interest of the Hull and White model modulated by the semi-Markov process of kernel* $\mathbf{Q}$ *is*

$$\begin{aligned} R_{i,u,r_t}(s) = \mathbb{E}[\delta_{i,u,r_t}(s)] = \frac{1-H_i(s+u)}{1-H_i(u)}e^{-k_i(s-t)}\left[r_t+\int_t^s e^{k_i(u)}\alpha_i(u)du\right] \\ +\sum_{k\in I}\int_t^s \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\int_{-\infty}^{+\infty} f_{\mathcal{N}(m_{i,r_t}(\tau),\sigma^2_{i,r_t}(\tau))}(x)R_{k,0,x}(s-\tau)dxd\tau. \end{aligned} \tag{2.49}$$

where $f_{\mathcal{N}(m_{i,r_t}(\tau),\sigma^2_{i,r_t}(\tau))}$ *is the probability distribution function of a Normal distribution with mean* $m_{i,r_t}(\tau)$ *and variance* $\sigma^2_{i,r_t}(\tau)$ *(cf. eq. (2.43) and (2.44) respectively) .*

Corollary 2.3.6. *The product moment of the force of interest of the Hull and White model modulated by the semi-Markov process of kernel* $\mathbf{Q}$ *is*

$$\begin{aligned} \Xi_{i,u,r_0}(s,h) = \mathbb{E}[\delta(s)\delta(s+h)|Z(0)=i,B(0)=u,\delta(0)=r_0] = \frac{1-H_i(s+h+u)}{1-H_i(u)}\rho_{i,r_0}(s,s+h) \\ +\sum_{k\in I}\int_s^{s+h}\frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}e^{-k_i(s)}\left[r_0+\int_0^s e^{k_i(u)}\alpha_i(u)du\right]\int_{-\infty}^{+\infty} f_{\mathcal{N}(m_{i,r_0}(\tau),\sigma^2_{i,r_0}(\tau))}(x)R_{k,0,x}(s+h-\tau)dxd\tau \\ +\sum_{k\in I}\int_0^s \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\int_{-\infty}^{+\infty} f_{\mathcal{N}(m_{i,r_0}(\tau),\sigma^2_{i,r_0}(\tau))}(x)\Xi_{k,0,x}(s-\tau,s+h-\tau)dxd\tau \ , \end{aligned} \tag{2.50}$$

where $\rho_{i,r_0}(s,s+h)$ *is given by*

$$\rho_{i,r_0}(s,s+h) = \mathbb{E}[r_{i,r_0}(s)r_{i,r_0}(s+h)] = m_{i,r_t}(s)m_{i,r_t}(s+h) + e^{-k_i(s)-k_i(s+h)}\int_0^t e^{2k_i(u)}\sigma_i(u)^2du \ , \tag{2.51}$$

and $f_{\mathcal{N}(m_{i,r_t}(\tau),\sigma^2_{i,r_t}(\tau))}$ *is the probability distribution function of a Normal distribution with mean* $m_{i,r_t}(\tau)$ *and variance* $\sigma^2_{i,r_t}(\tau)$ *(cf. eq. (2.43) and (2.44) respectively) .*

2.3.3 Cox-Ingersoll-Ross (CIR) Modulated Model

In this model we assume that the dynamics of the force of interest between two consecutive renewal moments, is described by the following stochastic differential equation

$$\begin{cases} dr_{i,r_0}(t) = (a_i - b_i r_{i,r_0}(t))dt + \sigma_i \sqrt{r_{i,r_0}(t)}dW_t \\ r_{i,r_0}(0) = r_0 \end{cases},$$

here, a_i and σ_i are non-negative constants, while $b_i \in \mathbb{R}$. This stochastic differential equation, for any $r_0 \in \mathbb{R}$, has a unique solution (see *e.g.* Ikeda and Watanabe [57]). We will not elaborate on the explicit solution of this equation, but following the approach of Lamberton and Lapeyre [63], we will study the property of the distribution of $r_{i,r_0}(t)$. To this aim, let us first study the distribution of $(r_{i,r_0}(t), \int_0^t r_{i,r_0}(s)ds)$. It is possible to show (see Lamberton and Lapeyre [63]) that

$$\mathbb{E}\left(e^{-\lambda r_{i,r_0}(t)} e^{-\mu \int_0^t r_{i,r_0}(s)ds}\right) = \exp(-a_i \phi_{i,\lambda,\mu}(t)) \exp(-r_0 \psi_{i,\lambda,\mu}(t)) \ , \tag{2.52}$$

where the functions $\phi_{i,\lambda,\mu}(t)$ and $\psi_{i,\lambda,\mu}(t)$ are given by

$$\phi_{i,\lambda,\mu}(t) = -\frac{2}{\sigma_i^2} \log\left(\frac{2\gamma_i e^{\frac{t(\gamma_i+b_i)}{2}}}{\sigma_i^2 \lambda(e^{\gamma_i t} - 1) + \gamma_i - b_i + e^{\gamma_i t}(\gamma_i + b_i)}\right) \ , \tag{2.53}$$

and

$$\psi_{i,\lambda,\mu}(t) = \frac{\lambda(\gamma_i + b_i + e^{\gamma_i t}(\gamma_i - b_i)) + 2\mu(e^{\gamma_i t} - 1)}{\sigma_i^2 \lambda(e^{\gamma_i t} - 1) + \gamma_i - b_i + e^{\gamma_i t}(\gamma_i + b_i)} \ , \tag{2.54}$$

with $\gamma_i = \sqrt{b_i^2 + 2\sigma_i^2 \mu}$. Using this result we can explicitly evaluate the Laplace transform of $r_{i,r_0}(t)$, indeed by putting $\mu = 0$ in the above expression we get

$$\mathbb{E}\left(e^{-\lambda r_{i,r_0}(t)}\right) = \left(\frac{2b_i}{\sigma_i^2 \lambda(1 - e^{-b_i t}) + 2b_i}\right)^{\frac{2a_i}{\sigma_i^2}} \exp\left(-r_0 \frac{2\lambda b_i e^{-b_i t}}{\sigma_i^2 \lambda(1 - e^{-b_i t}) + 2b_i}\right). \tag{2.55}$$

Moreover, we can obtain an analytic representation of the *nth* moment of a DF between two renewal moments, in fact setting $\lambda = 0$ and $\mu = n$, we have

$$B_{i,r_0}^{(n)}(0,t) = \mathbb{E}\left(e^{-n \int_0^t r_{i,r_0}(s)ds}\right) = \left(\frac{2\gamma_i e^{\frac{t(\gamma_i+b_i)}{2}}}{\gamma_i - b_i + e^{\gamma_i t}(\gamma_i + b_i)}\right)^{\frac{2a_i}{\sigma_i^2}} \exp\left(-r_0 \frac{2(e^{\gamma_i t} - 1)}{\gamma_i - b_i + e^{\gamma_i t}(\gamma_i + b_i)}\right), \tag{2.56}$$

where $\gamma_i = \sqrt{b_i^2 + 2\sigma_i^2 n}$. Starting from here we can obtain the *nth* moment of a DF in an arbitrary time interval and the first moment of the force of interest.

Corollary 2.3.7. *The nth order moment of the discount factor in the CIR model modulated by the semi-Markov process of kernel* $\mathbf{Q}$ *is*

$$
\begin{aligned}
V^{(n)}_{i,u,r_t}(t,T) &= \mathbb{E}\left\{exp(-n\int_t^T \delta_{i,u,r_t}(s)ds)\right\} = \frac{1-H_i(T-t+u)}{1-H_i(u)}\big(B^{(n)}_{i,r_t}(t,T)\big) \\
&+ \sum_{k\in I}\int_t^T \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\left(B^{(n)}_{i,r_t}(t,\tau)\int_0^{+\infty} f_{\chi^2(2c_i(\tau,t)r_t;2q_i+2,2u_i(\tau,t))}(x)V^{(n)}_{k,0,x}(\tau,T)dx\right)d\tau.
\end{aligned}
\tag{2.57}
$$

where $B^{(n)}_{i,r_t}(t,T)$ *is given in equation (2.56) and* $f_{\chi^2(2c_i(\tau,t)r_t;2q_i+2,2u_i(\tau,t))}$ *is the non-central chi-square distribution function with*

$$
c_i(\tau,t) = \frac{2b_i}{\sigma_i^2(1-e^{-b_i(\tau-t)})}\ ;\ q_i = \frac{2a_i}{\sigma_i^2}-1\ ;\ u_i(\tau,t) = c_i(\tau,t)r_te^{-b_i(\tau-t)}\ . \tag{2.58}
$$

Corollary 2.3.8. *The first moment of the force of interest of the CIR model modulated by a semi-Markov process of kernel* $\mathbf{Q}$ *is*

$$
\begin{aligned}
R_{i,u,r_t}(s) &= \mathbb{E}[\delta_{i,u,r_t}(s)] = \frac{1-H_i(s+u)}{1-H_i(u)}\left[r_te^{-b_i(s-t)}+\frac{a_i}{b_i}\left(1-e^{-b_i(s-t)}\right)\right] \\
&+ \sum_{k\in I}\int_t^s \frac{\dot{Q}_{ik}(\tau+u)}{1-H_i(u)}\int_0^{+\infty} f_{\chi^2(2c_i(\tau,t)r_t;2q_i+2,2u_i(\tau,t))}(x)R_{k,0,x}(s-\tau)dxd\tau.
\end{aligned}
\tag{2.59}
$$

where $f_{\chi^2(2c_i(\tau,t)r_t;2q_i+2,2u_i(\tau,t))}$ *is the non-central chi-square distribution function.*

Chapter 3

Financial Markets with Markov Modulated Stochastic Volatilities

In this chapter we price covariance and correlation swaps for financial markets with Markov-modulated volatilities. As an example, we consider stochastic volatility driven by a two-state continuous time Markov chain. In this case, numerical examples are presented for VIX and VXN volatility indexes (S&P 500 and NASDAQ-100, respectively, since January 2004 to June 2012). We also use VIX to price variance and volatility swaps for the two-state Markov-modulated volatility and to present a numerical result in this case.

The chapter is organized as follows. First, martingale representation of Markov processes is considered. Then, variance and volatility swaps for financial markets with Markov-modulated stochastic volatilities are studied. Furthermore, pricing of covariance and correlation swaps for a two risky assets for financial markets with Markov-modulated stochastic volatilities is presented. Finally, we consider an example for variance, volatility, covariance and correlation swaps for Markov-modulated volatility driven by a two-state continuous time Markov chain. Here, a numerical example for variance and volatility swaps pricing using S&P500 VIX index is given. Moreover, a numerical example for covariance and correlation swaps using S&P500 VIX and NASDAQ-100 VXN indexes is presented.

This chapter is based on the work done while visiting University of Calgary, Calgary (AB) Canada, under the supervision of Prof. A. V. Swishchuk.

This chapter is based on a paper (G. Salvi and A. V. Swishchuk [75]) currently under review.

3.1 Introduction

Among the recent and new financial products there are covariance and correlation swaps, which are useful for volatility hedging and speculation using two different financial underlying assets.

For example, option dependent on exchange rate movements, such as those paying in a currency different from the underlying currency, have an exposure to movements of the correlation between the asset and the exchange rate. This risk may be eliminated by using covariance swap.

A *covariance swap* is a covariance forward contract of the underlying rates S^1 and S^2. Its payoff at expiration is equal to

$$N(Cov_R(S^1, S^2) - K_{cov}) \ ,$$

where K_{cov} is the strike price, N is the notional amount, $Cov_R(S^1, S^2)$ is the covariance between the two assets S^1 and S^2.

A *correlation swap* is a correlation forward contract of two underlying rates S^1 and S^2 which payoff at expiration is equal to:

$$N(Corr_R(S^1, S^2) - K_{corr}) \ ,$$

where $Corr(S^1, S^2)$ is the realized correlation of two underlying assets S^1 and S^2, K_{corr} is the strike price, and N is the notional amount.

Pricing covariance swap, from a theoretical point of view, is similar to pricing variance swaps, since

$$Cov_R(S^1, S^2) = 1/4\{\sigma_R^2(S^1 S^2) - \sigma_R^2(S^1/S^2)\} \ ,$$

where $\sigma_R^2(S)$ is a variance swap for underlying assets.

Thus, we need to know variances for $S^1 S^2$ and for S^1/S^2. Correlation $Corr_R(S^1, S^2)$ is defined as follows:

$$Corr_R(S^1, S^2) = \frac{Cov_R(S^1, S^2)}{\sqrt{\sigma_R^2(S^1)}\sqrt{\sigma_R^2(S^2)}} \ .$$

Given two assets S_t^1 and S_t^2 with $t \in [0, T]$, sampled on days $t_0 = 0 < t_1 < t_2 < ... < t_n = T$ between today and maturity T, the log-return for each asset is: $R_i^j := \log(\frac{S_{t_i}^j}{S_{t_{i-1}}^j}), \quad i = 1, 2, ..., n, \quad j = 1, 2.$

Covariance and correlation can be approximated by

$$Cov_n(S^1, S^2) = \frac{n}{(n-1)T} \sum_{i=1}^{n} R_i^1 R_i^2 \ ,$$

and

$$Corr_n(S^1, S^2) = \frac{Cov_n(S^1, S^2)}{\sqrt{Var_n(S^1)}\sqrt{Var_n(S^2)}} ,$$

respectively.

The literature devoted to the volatility derivatives is growing. We give here a short overview of the latest development in this area. The Non-Gaussian Ornstein-Uhlenbeck stochastic volatility model was used by Benth *et al.* [4] to study volatility and variance swaps. M. Broadie and A. Jain [15] evaluated price and hedging strategy for volatility derivatives in the Heston square root stochastic volatility model and in [16] they compare result from various model in order to investigate the effect of jumps and discrete sampling on variance and volatility swaps. Pure jump process with independent increments return were used by Carr *et al.* [19] to price derivatives written on realized variance, and subsequent development by Carr and Lee [20]. We also refer to Carr and Lee [21] for a good survey on volatility derivatives. Da Fonseca *et al.* [43] analyzed the influence of variance and covariance swap in a market by solving a portfolio optimization problem in a market with risky assets and volatility derivatives. Correlation swap price has been investigated by Bossu [12, 13] for component of an equity index using statistical method. Drissien *et al.* [46] discusses the price of correlation risk for equity options. Pricing volatility swaps under Heston's model with regime-switching and pricing options under a generalized Markov-modulated jump-diffusion model are discussed by Elliott *et al.* [48, 49], respectively. Howison *et al.* [53] considers the pricing of a range of volatility derivatives, including volatility and variance swaps and swaptions. The pricing options on realized variance in the Heston model with jumps in returns and volatility is studied by Sepp [77]. An analytical closed-form pricing of pseudo-variance, pseudo-volatility, pseudo-covariance and pseudo-correlation swaps is studied by Swishchuk *et al.* [82]. Windcliff *et al.* [87] investigates the behaviour and hedging of discretely observed volatility derivatives.

3.2 Martingale Representation of Markov Processes

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ be a filtered probability space, with right-continuous filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ and probability $\mathbb{P}$. Let $(X, \mathscr{X})$ be a measurable space and $(x_t)_{t\in\mathbb{R}_+}$ be a $(X, \mathscr{X})$-valued Markov process with generator Q. The following two results allow us to associate a martingale to this process and to obtain its quadratic variation (we refer to Elliott and Swishchuk [50] for the proofs).

Proposition 3.2.1. *(Elliott and Swishchuk [50])*

Let $(x_t)_{t\in\mathbb{R}_+}$ be a Markov process with generator Q and $f \in$ Domain (Q), then

$$m_t^f := f(x_t) - f(x_0) - \int_0^t Qf(x_s)ds \ , \tag{3.1}$$

is a zero-mean martingale with respect to $\mathscr{F}_t := \sigma\{y(s); 0 \le s \le t\}$.

Let us evaluate the quadratic variation of this martingale.

Proposition 3.2.2. *(Elliott and Swishchuk [50])*

Let $(x_t)_{t\in\mathbb{R}_+}$ be a Markov process with generator Q, $f \in$ Domain (Q) and $(m_t^f)_{t\in\mathbb{R}_+}$ its associated martingale, then

$$\langle m^f \rangle_t := \int_0^t [Qf^2(x_s) - 2f(x_s)Qf(x_s)]ds \ , \tag{3.2}$$

is the quadratic variation of m^f.

In the following it will be useful to consider the quadratic covariation of two martingales associated to a generic couple of functions of a Markov process.

Proposition 3.2.3. *Let $(x_t)_{t\in\mathbb{R}_+}$ be a Markov process with generator Q, $f, g \in$ Domain (Q) such that $fg \in$ Domain (Q). Denote by $(m_t^f)_{t\in\mathbb{R}_+}, (m_t^g)_{t\in\mathbb{R}_+}$ their associated martingale. Then*

$$\langle f(x.), g(x.)\rangle_t := \int_0^t \{Q(f(x_s)g(x_s)) - [g(x_s)Qf(x_s) + f(x_s)Qg(x_s)]\}ds \ , \tag{3.3}$$

is the quadratic covariation of f and g.

Proof. First of all, we note that

$$\begin{aligned} m_t^f m_t^g &= f(x_t)g(x_t) - f(x_t)\int_0^t Qg(x_s)ds - g(x_t)\int_0^t Qf(x_s)ds + \int_0^t Qf(x_s)ds \int_0^t Qg(x_s)ds \\ &= f(x_t)g(x_t) - m_t^f \int_0^t Qg(x_s)ds - m_t^g \int_0^t Qf(x_s)ds - \int_0^t Qf(x_s)ds \int_0^t Qg(x_u)du \ , \end{aligned} \tag{3.4}$$

Moreover,

$$\begin{aligned} & d\left[m_t^f \int_0^t Qg(x_s)ds + m_t^g \int_0^t Qf(x_s)ds + \int_0^t Qf(x_s)ds \int_0^t Qg(x_u)du\right] \\ &= \left(\int_0^t Qg(x_s)ds\right) dm_t^f + \left(\int_0^t Qf(x_s)ds\right) dm_t^g + m_t^f Qg(x_t)dt + m_t^g Qf(x_t)dt \\ &+ \left(\int_0^t Qg(x_s)ds\right) Qf(x_t)dt + \left(\int_0^t Qf(x_s)ds\right) Qg(x_t)dt \ . \end{aligned}$$

Using the expression for m^f and m^g (*cf.* Proposition 3.2.1) we obtain

$$
\begin{aligned}
&d\left[m_t^f \int_0^t Qg(x_s)ds + m_t^g \int_0^t Qf(x_s)ds + \int_0^t Qf(x_s)ds \int_0^t Qg(x_u)du\right] \\
&= \left(\int_0^t Qg(x_s)ds\right) dm_t^f + \left(\int_0^t Qf(x_s)ds\right) dm_t^g + f(x_t)Qg(x_t)dt + g(x_t)Qf(x_t)dt \ .
\end{aligned}
\tag{3.5}
$$

and using Eq. (3.5), Eq. (3.4) becomes

$$
\begin{aligned}
m_t^f m_t^g = f(x_t)g(x_t) &- \left[\int_0^t \left(\int_0^s Qg(x_u)du\right) dm_s^f + \int_0^t \left(\int_0^s Qf(x_u)du\right) dm_s^g\right] \\
&- \int_0^t \left[f(x_s)Qg(x_s) + g(x_s)Qf(x_s)\right] ds \ .
\end{aligned}
$$

Adding and subtracting $\int_0^t Q(f(x_s)g(x_s))ds$ on the right hand side of the previous equation, we have

$$
\begin{aligned}
m_t^f m_t^g &= f(x_t)g(x_t) - \int_0^t Q(f(x_s)g(x_s))ds \\
&- \left[\int_0^t \left(\int_0^s Qg(x_u)du\right) dm_s^f + \int_0^t \left(\int_0^s Qf(x_u)du\right) dm_s^g\right] \\
&+ \int_0^t \left[Q(f(x_s)g(x_s)) - f(x_s)Qg(x_s) - g(x_s)Qf(x_s)\right] ds \ .
\end{aligned}
\tag{3.6}
$$

Since $fg \in$ Domain (Q), then (*cf.* Proposition 3.2.1)

$$
f(x_t)g(x_t) - \int_0^t Q(f(x_s)g(x_s))ds \qquad t \in \mathbb{R}_+ \ , \tag{3.7}
$$

is a martingale. The term in square bracket on the r.h.s of Eq. (3.6) is a martingale too. Therefore

$$
m_t^f m_t^g - \int_0^t \left[Q(f(x_s)g(x_s)) - f(x_s)Qg(x_s) - g(x_s)Qf(x_s)\right] ds \qquad t \in \mathbb{R}_+ \ , \tag{3.8}
$$

is a martingale and we have that

$$
\langle f(x.), g(x.)\rangle_t = \int_0^t \left[Q(f(x_s)g(x_s)) - f(x_s)Qg(x_s) - g(x_s)Qf(x_s)\right] ds \ . \tag{3.9}
$$

□

Now, we are able to evaluate the expectation of a generic function of a Markov process.

Proposition 3.2.4. *(Elliott and Swishchuk [50])*

Let $(x_t)_{t\in\mathbb{R}_+}$ be a Markov process with generator Q and $f \in$ Domain (Q), then

$$
\mathbb{E}\{f(x_t)\} = e^{tQ} f(x_0) \ . \tag{3.10}
$$

Remark 3.2.5. *Let $(x_t)_{t\in\mathbb{R}_+}$ be a Markov process with generator Q, $f,g \in$ Domain (Q) such that $fg \in$ Domain (Q), then*

$$\mathbb{E}\{f(x_t)g(x_t)\} = e^{tQ} f(x_0)g(x_0) \ . \tag{3.11}$$

3.3 Variance and Volatility Swaps for Financial Markets with Markov-Modulated Stochastic Volatilities

Let us consider a financial market with only two securities, the risk free bond $(B_t)_{t\in\mathbb{R}_+}$ and the stock $(S_t)_{t\in\mathbb{R}_+}$. Let us suppose that the stock price and the bond satisfy the following stochastic differential equation

$$\begin{cases} dB_t = B_t r(x_t)dt \\ dS_t = S_t(\mu(x_t)dt + \sigma(x_t)dw_t) \end{cases} , \tag{3.12}$$

where w is a standard Wiener process independent of the Markov process $(x_t)_{t\in\mathbb{R}_+}$. In this model the volatility is stochastic, then it is interesting to study the property of σ and in particular how to price future contracts on realized variance and volatility.

The following results concern the expectation of variance and are simple application of Propositions 3.2.1 and 3.2.4.

Corollary 3.3.1. *Suppose that $\sigma \in$ Domain (Q). Then*

$$\mathbb{E}\{\sigma^2(x_t)|\mathscr{F}_u\} = \sigma^2(x_u) + \int_u^t Q\mathbb{E}\{\sigma^2(x_s)|\mathscr{F}_u\}ds \ , \tag{3.13}$$

for any $0 \le u \le t$.

Corollary 3.3.2. *The conditional expectation of the variance can be expressed as*

$$\mathbb{E}\{\sigma^2(x_t)|\mathscr{F}_u\} = e^{(t-u)Q}\sigma^2(x_u) \ , \tag{3.14}$$

for any $0 \le u \le t$.

This Markov-modulated financial market is incomplete (see Elliott and Swishchuk [50]). In order to price the swaps we will use the minimal martingale measure. We briefly recall here the main definition.

Definition 3.3.3. ***Martingale measure***

The measure $\mathbb{P}^*$ is called a martingale measure, if it is equivalent to $\mathbb{P}$ and such that the discounted capital

$$M_t := \frac{\beta_t B_t + \gamma_t S_t}{B_t} \qquad \text{here } (\beta, \gamma) \text{ is a portfolio} ,$$

is a $\mathbb{P}^*$ martingale.

Definition 3.3.4. ***Strongly orthogonal martingales***

Two martingales are said to be strongly orthogonal if their product is a martingale.

Definition 3.3.5. ***Minimal martingale measure***

A martingale measure $\mathbb{P}^*$ for a discounted capital is called a minimal martingale measure associated with $\mathbb{P}$ if any local $\mathbb{P}$-martingale strongly orthogonal (under $\mathbb{P}$) to each local martingale M remains a local martingale under $\mathbb{P}^*$.

Lemma 3.3.6. *(Elliott and Swishchuk [50])*

The measure defined by

$$\left.\frac{d\hat{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathscr{F}_t} = \eta_t \ , \tag{3.15}$$

where

$$\eta_t := e^{\int_0^t [(r(x_s)-\mu(x_s))/\sigma(x_s)]dw_s - \frac{1}{2}\int_0^t [(r(x_s)-\mu(x_s))/\sigma(x_s)]^2 ds} \ , \tag{3.16}$$

is the minimal martingale measure associated with $\mathbb{P}$.

In the following, for simplicity we will denote the minimal martingale measure with $\mathbb{P}$ as well.

3.3.1 Pricing Variance Swaps

Let us start from the more straightforward variance swap. Variance swaps are forward contract on future realized level of variance. The payoff of a variance swap with expiration date T is given by

$$N(\sigma_R^2(x) - K_{var}) \ . \tag{3.17}$$

Here $\sigma_R^2(x)$ is the realized stock variance over the life of the contract defined by

$$\sigma_R^2(x) := \frac{1}{T}\int_0^T \sigma^2(x_s)ds \ , \tag{3.18}$$

while K_{var} is the strike price for variance and N is the notional amounts of dollars per annualized variance point. Without loss of generality, we can assume $N = 1$. The price of the variance swap is the expected present value of the payoff in the risk-neutral world

$$P_{var}(x) = \mathbb{E}\{e^{-rT}(\sigma_R^2(x) - K_{var})\} \ . \tag{3.19}$$

The following result concern the evaluation of the variance swap. We refer to Elliott and Swishchuk [50] for a complete discussion and proof.

Theorem 3.3.7. *(Elliott and Swishchuk [50])*

The present value of a variance swap for Markov stochastic volatility is

$$P_{var}(x) = e^{-rT}\left\{\frac{1}{T}\int_0^T (e^{tQ}\sigma^2(x) - K_{var})dt\right\} . \tag{3.20}$$

3.3.2 Pricing Volatility Swaps

Volatility swaps are forward contracts on future realized level of volatility. The payoff of a volatility swap with maturity date T is given by

$$N(\sigma_R(x) - K_{vol}) \ , \tag{3.21}$$

where $\sigma_R(x)$ is the realized stock volatility over the life of the contract defined by

$$\sigma_R(x) := \sqrt{\frac{1}{T}\int_0^T \sigma^2(x_s)ds} \ , \tag{3.22}$$

where K_{vol} is the strike price for volatility and N is the notional amounts of dollars per annualized volatility point. We will assume, as before, that $N = 1$ for sake of simplicity. The price of the volatility swap is the expected present value of the payoff in the risk-neutral world

$$P_{vol}(x) = \mathbb{E}\{e^{-rT}(\sigma_R(x) - K_{vol})\} \ . \tag{3.23}$$

In order to evaluate the volatility swaps we need to know the expected value of the square root of the variance, but unfortunately we are not able to evaluate analytically this expected value. Then in order to obtain a closed formula for the price of volatility swaps we have to make an approximation. Following the approach of Brockhaus and Long [17] (see also Javaheri *et al.* [61]), from the second order Taylor expansion we have

$$\mathbb{E}\{\sqrt{\sigma_R^2(x)}\} \approx \sqrt{\mathbb{E}\{\sigma_R^2(x)\}} - \frac{Var\{\sigma_R^2(x)\}}{8\mathbb{E}\{\sigma_R^2(x)\}^{3/2}} \ . \tag{3.24}$$

Then, in order to evaluate the volatility swap price we have to know both expectation and variance of $\sigma_R^2(x)$.

Theorem 3.3.8. *The value of a volatility swap for Markov-modulated stochastic volatility is*

$$P_{vol}(x) \approx \quad e^{-rT}\left\{\sqrt{\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt} - \frac{Var\{\sigma_R^2(x)\}}{8\left(\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt\right)^{3/2}} - K_{vol}\right\} \ ,$$

where the variance is given by

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T\int_0^t e^{sQ}\left[\sigma^2(x)e^{(t-s)Q}\sigma^2(x)\right] dsdt - \left[\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt\right]^2 .$$

Proof. We have already obtained the expectation of the realized variance,

$$\mathbb{E}\{\sigma_R^2(x)\} = \frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt \ , \tag{3.25}$$

then it remains to prove that

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T\int_0^t e^{sQ}\left[\sigma^2(x)e^{tQ}\sigma^2(x)\right] dsdt - \left[\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt\right]^2 .$$

The variance is, from the definition, given by

$$Var\{\sigma_R^2(x)\} = \mathbb{E}\{[\sigma_R^2(x) - \mathbb{E}\{\sigma_R^2(x)\}]^2\} \ . \tag{3.26}$$

Using the definition of realized variance, and Fubini theorem, we have

$$\begin{aligned} Var\{\sigma_R^2(x)\} =& \mathbb{E}\left\{\left[\frac{1}{T}\int_0^T \sigma^2(x_t)dt - \frac{1}{T}\int_0^T \mathbb{E}\{\sigma^2(x_t)\}dt\right]^2\right\} \\ =& \mathbb{E}\left\{\left[\frac{1}{T}\int_0^T \left(\sigma^2(x_t) - \mathbb{E}\{\sigma^2(x_t)\}\right) dt\right]^2\right\} \\ =& \frac{1}{T^2}\int_0^T\int_0^T \mathbb{E}\left\{[\sigma^2(x_t) - \mathbb{E}\{\sigma^2(x_t)\}][\sigma^2(x_s) - \mathbb{E}\{\sigma^2(x_s)\}]\right\} dsdt \ , \end{aligned} \tag{3.27}$$

and then solving the product

$$Var\{\sigma_R^2(x)\} = \frac{1}{T^2}\int_0^T\int_0^T \mathbb{E}\left\{\sigma^2(x_t)\sigma^2(x_s)\right\} dsdt - \left[\frac{1}{T}\int_0^T \mathbb{E}\{\sigma^2(x_t)\}dt\right]^2 . \tag{3.28}$$

The second term on the r.h.s of Eq. (3.28) is known: it follows directly from Proposition 3.3.1. Moreover, we observe that the integrand on the first term is invariant in the exchange of s and t. Then, we can divide the integration set in two zone above and below the line $t = s$. Thanks to symmetry the contribution on the two parts is the same. We can rewrite the variance as

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T \int_0^t \mathbb{E}\left\{\sigma^2(x_t)\sigma^2(x_s)\right\} dsdt - \left[\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt\right]^2 . \tag{3.29}$$

We stress that, in this form, the integration set of the first term is such that the inequality $s \leq t$ holds true. Using the properties of conditional expectation we have

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T \int_0^t \mathbb{E}\{\sigma^2(x_s)\mathbb{E}\{\sigma^2(x_t)|\mathscr{F}_s\}\}dsdt - \left[\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt\right]^2 .$$

Using the Markov property and Corollary 3.3.2, the conditional expected value in the integrand can be viewed as a function of the process at time s, that is

$$\mathbb{E}\{\sigma^2(x_t)|\mathscr{F}_s\} = e^{(t-s)Q}\sigma(x_s) =: g(x_s) . \tag{3.30}$$

Thus, the variance can be expressed as

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T \int_0^t \mathbb{E}\{\sigma^2(x_s)g(x_s)\}dsdt - \left[\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt\right]^2 .$$

Now, using Proposition 3.2.4 we can solve the expectation on the integrand and we finally obtain

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T \int_0^t e^{sQ}\left(\sigma^2(x)g(x)\right) dsdt - \left[\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt\right]^2 .$$

Moreover we observe that function g, evaluated in x, becomes

$$g(x) = e^{(t-s)Q}\sigma^2(x) . \tag{3.31}$$

Then, by substituting in the previous formula we can expressed the variance as

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T \int_0^t e^{sQ}\left[\sigma^2(x)e^{(t-s)Q}\sigma^2(x)\right] dsdt - \left[\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x)dt\right]^2 .$$

□

3.4 Covariance and Correlation Swaps for a Two Risky Assets Financial Markets with Markov-Modulated Stochastic Volatilities

Let's consider a market model with two risky assets and a risk free bond. Let's assume that the risky assets satisfy the following stochastic differential equations

$$\begin{cases} dS_t^{(1)} = S_t^{(1)}(\mu_t^{(1)}dt + \sigma^{(1)}(x_t)dw_t^{(1)}) \\ dS_t^{(2)} = S_t^{(2)}(\mu_t^{(2)}dt + \sigma^{(2)}(x_t)dw_t^{(2)}) \end{cases} , \tag{3.32}$$

where $\mu^{(1)}, \mu^{(2)}$ are deterministic functions of time and $(w_t^{(1)})_{t\in\mathbb{R}_+}$ and $(w_t^{(2)})_{t\in\mathbb{R}_+}$ are standard Wiener processes with quadratic covariance given by

$$d[w_t^{(1)}, w_t^{(2)}] = \rho_t dt \ . \tag{3.33}$$

Here, ρ_t is a deterministic function of time and $(w_t^{(1)})_{t\in\mathbb{R}_+}, (w_t^{(2)})_{t\in\mathbb{R}_+}$ are supposed to be independent of the Markov process $(x_t)_{t\in\mathbb{R}_+}$. This model allows us to study the covariance and correlation structure of two risky assets and how it is possible to price future contract on them.

3.4.1 Pricing Covariance Swaps

A covariance swap is a covariance forward contract on the realized covariance between two risky assets which payoff at maturity is equal to

$$N(Cov_R(S^{(1)}, S^{(2)}) - K_{cov}) \ , \tag{3.34}$$

where K_{cov} is a strike reference value, N is the notional amount and $Cov_R(S^{(1)}, S^{(2)})$ is the realized covariance of the two assets $S^{(1)}$ and $S^{(2)}$ defined by

$$Cov_R(S^{(1)}, S^{(2)}) = \frac{1}{T}[\ln S_T^{(1)}, \ln S_T^{(2)}] = \frac{1}{T}\int_0^T \rho_t \sigma^{(1)}(x_t)\sigma^{(2)}(x_t)dt \ . \tag{3.35}$$

The price of the covariance swap is the expected present value of the payoff in the risk neutral world

$$P_{cov}(x) = \mathbb{E}\{e^{-rT}(Cov_R(S^{(1)}, S^{(2)}) - K_{cov})\} \ , \tag{3.36}$$

where we assumed that $N = 1$.

Theorem 3.4.1. *The value of a covariance swap for Markov-modulated stochastic volatility is*

$$P_{cov}(x) = e^{-rT}\left\{\frac{1}{T}\int_0^T \rho_t e^{tQ}[\sigma^{(1)}(x)\sigma^{(2)}(x)]dt - K_{cov}\right\}. \tag{3.37}$$

Proof. To evaluate the price of covariance swap we only need to know

$$\mathbb{E}\{Cov_R(S^{(1)}, S^{(2)})\} = \frac{1}{T}\int_0^T \rho_t \mathbb{E}\{\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)\}dt\ . \tag{3.38}$$

Then, it remains to show that

$$\mathbb{E}\{\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)\} = e^{tQ}[\sigma^{(1)}(x)\sigma^{(2)}(x)]\ . \tag{3.39}$$

By applying Ito's lemma we have

$$d(\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)) = \sigma^{(1)}(x_t)d\sigma^{(2)}(x_t) + \sigma^{(2)}(x_t)d\sigma^{(1)}(x_t) + d\langle\sigma^{(1)}(x.), \sigma^{(2)}(x.)\rangle_t\ . \tag{3.40}$$

Using Proposition 3.2.3, we can express the covariation as

$$d\langle\sigma^{(1)}(x.), \sigma^{(2)}(x.)\rangle_t = Q[\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)]dt - [\sigma^{(1)}(x_t)Q\sigma^{(2)}(x_t) + \sigma^{(2)}(x_t)Q\sigma^{(1)}(x_t)]dt\ . \tag{3.41}$$

Furthermore, from Proposition 3.2.1, we have

$$d\sigma^{(i)}(x_t) = Q\sigma^{(i)}(x_t)dt + dm^{\sigma^{(i)}} \qquad i = 1, 2\ . \tag{3.42}$$

Substituting Eqs. (3.41), (3.42) in Eq. (3.40) we get

$$d(\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)) = Q[\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)]dt + \sigma^{(1)}(x_t)dm^{\sigma^{(2)}} + \sigma^{(2)}(x_t)dm^{\sigma^{(1)}}\ . \tag{3.43}$$

Taking the expectation on both side of the above equation we obtain

$$\mathbb{E}\{\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)\} = \sigma^{(1)}(x)\sigma^{(2)}(x) + \int_0^t Q\mathbb{E}\{\sigma^{(1)}(x_s)\sigma^{(2)}(x_s)\}dt\ . \tag{3.44}$$

Now, we can solve this differential equation and we get

$$\mathbb{E}\{\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)\} = e^{tQ}[\sigma^{(1)}(x)\sigma^{(2)}(x)]\ . \tag{3.45}$$

This conclude the proof. □

3.4.2 Pricing Correlation Swaps

A correlation swap is a forward contract on the correlation between the underlying assets S^1 and S^2 which payoff at maturity is equal to

$$N(Corr_R(S^1, S^2) - K_{corr}) \ , \tag{3.46}$$

where K_{corr} is a strike reference level, N is the notional amount and $Corr_R(S^1, S^2)$ is the realized correlation defined by

$$Corr_R(S^1, S^2) = \frac{Cov_R(S^1, S^2)}{\sqrt{\sigma_R^{(1)^2}(x)}\sqrt{\sigma_R^{(2)^2}(x)}} \ , \tag{3.47}$$

where the realized variance is given by

$$\sigma_R^{(i)^2}(x) = \frac{1}{T}\int_0^T (\sigma^{(i)}(x_t))^2 dt \qquad\qquad i = 1, 2 \ . \tag{3.48}$$

The price of the correlation swap is the expected present value of the payoff in the risk neutral world

$$P_{corr}(x) = \mathbb{E}\{e^{-rT}(Corr_R(S^1, S^2) - K_{corr})\} \ , \tag{3.49}$$

where we set $N = 1$ for simplicity. Unfortunately the expected value of $Corr_R(S^1, S^2)$ is not known analytically. In order to obtain an explicit formula for the correlation swap price we have to introduce some approximation.

3.4.3 Correlation Swaps Made Simple

First of all, let us introduce the following notations

$$\begin{aligned} X =& Cov_R(S^1, S^2) \\ Y =& \sigma_R^{(1)^2}(x) \\ Z =& \sigma_R^{(2)^2}(x) \ . \end{aligned} \tag{3.50}$$

In what follows we will denote with the subscript 0 the expected value of the above random variables. Starting from the approach we have used for the volatility swap, we would like to approximate the square root of Y and Z at the first order as follows

$$\begin{aligned} \sqrt{Y} \approx& \sqrt{Y_0} + \frac{Y - Y_0}{2\sqrt{Y_0}} \\ \sqrt{Z} \approx& \sqrt{Z_0} + \frac{Z - Z_0}{2\sqrt{Z_0}} \ . \end{aligned} \tag{3.51}$$

The realized correlation can now be approximated by

$$Corr_R(S^1,S^2) \approx \frac{X}{\left(\sqrt{Y_0}+\frac{Y-Y_0}{2\sqrt{Y_0}}\right)\left(\sqrt{Z_0}+\frac{Z-Z_0}{2\sqrt{Z_0}}\right)} = \frac{\frac{X}{\sqrt{Y_0}\sqrt{Z_0}}}{\left(1+\frac{Y-Y_0}{2Y_0}\right)\left(1+\frac{Z-Z_0}{2Z_0}\right)} .$$

Solving the product in the denominator on the r.h.s of last term and keeping only the terms up to the first order in the increment, we have

$$Corr_R(S^1,S^2) \approx \frac{\frac{X}{\sqrt{Y_0}\sqrt{Z_0}}}{1+\left(\frac{Y-Y_0}{2Y_0}+\frac{Z-Z_0}{2Z_0}\right)} \approx \frac{X}{\sqrt{Y_0}\sqrt{Z_0}}\left[1-\left(\frac{Y-Y_0}{2Y_0}+\frac{Z-Z_0}{2Z_0}\right)\right].$$

Finally we obtain the following approximation for the correlation

$$Corr_R(S^1,S^2) \approx \frac{X}{\sqrt{Y_0}\sqrt{Z_0}} - \frac{X}{\sqrt{Y_0}\sqrt{Z_0}}\left(\frac{Y-Y_0}{2Y_0}+\frac{Z-Z_0}{2Z_0}\right) . \tag{3.52}$$

We are going to evaluate the expectation only on the first term on the right hand side, which represent the zero order of approximation and the most intuitive part. We will discuss the first order correction in the next section. In what follows we are going to approximate the realized correlation as

$$Corr_R(S^1,S^2) \approx \frac{X}{\sqrt{Y_0}\sqrt{Z_0}} . \tag{3.53}$$

Substituting X, Y and Z we obtain

$$Corr_R(S^1,S^2) \approx \frac{1}{\sqrt{\mathbb{E}\{\sigma_R^{(1)^2}(x)\}}\sqrt{\mathbb{E}\{\sigma_{2R}^{(2)^2}(x)\}}}\frac{1}{T}\int_0^T \rho_t\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)dt , \tag{3.54}$$

where (*cf.* Theorem 3.3.7), we have

$$\mathbb{E}\left\{\sigma^2_{(i)R}(x)\right\} = \mathbb{E}\left\{\frac{1}{T}\int_0^T(\sigma^{(i)}(x_t))^2dt\right\} = \frac{1}{T}\int_0^T e^{tQ}(\sigma^{(i)}(x))^2dt , \tag{3.55}$$

for $i=1,2$. In order to price a correlation swap we need to be able to evaluate the expectation of both side of Eq. (3.54). From Proposition 3.2.4 the expectation of the integrand on the r.h.s of Eq. (3.54) is given by

$$\mathbb{E}\{\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)\} = e^{tQ}\sigma^{(1)}(x)\sigma^{(2)}(x) . \tag{3.56}$$

We can summarize the previous result in the following statement.

Theorem 3.4.2. *The value of a correlation swap for a Markov-modulated stochastic volatility is*

$$P_{corr}(x) \approx e^{-rT}\left[\frac{\int_0^T \rho_t e^{tQ}\sigma^{(1)}(x)\sigma^{(2)}(x)dt}{\sqrt{\int_0^T e^{tQ}(\sigma^{(1)}(x))^2dt}\sqrt{\int_0^T e^{tQ}(\sigma^{(2)}(x))^2dt}} - K_{corr}\right]. \tag{3.57}$$

3.4.4 Correlation Swaps: First Order Correction

We would like to obtain an approximation for the realized correlation between two risky assets

$$Corr_R(S^1,S^2) = \frac{Cov_R(S^1,S^2)}{\sqrt{\sigma_R^{(1)^2}(x)}\sqrt{\sigma_R^{(2)^2}(x)}} \ . \tag{3.58}$$

In Section 3.4.3 we have already obtained the following approximated expression

$$Corr_R(S^1,S^2) \approx \frac{\frac{X}{\sqrt{Y_0}\sqrt{Z_0}}}{1+\left(\frac{Y-Y_0}{2Y_0}+\frac{Z-Z_0}{2Z_0}\right)} \approx \frac{X}{\sqrt{Y_0}\sqrt{Z_0}}\left[1-\left(\frac{Y-Y_0}{2Y_0}+\frac{Z-Z_0}{2Z_0}\right)\right]. \tag{3.59}$$

where

$$\begin{aligned} X =& Cov_R(S^1,S^2) \\ Y =& \sigma_R^{(1)^2}(x) \\ Z =& \sigma_R^{(2)^2}(x) \ , \end{aligned} \tag{3.60}$$

and with the pedix 0 we have denoted the expected values. We have already evaluated the expectation of the zero order approximation, now we would like to evaluate the first order.

Substituting X, Y and Z in Eq. (3.59) we obtain

$$\begin{aligned} Corr_R(S^1,S^2) \approx& \frac{1}{\sqrt{\mathbb{E}\{\sigma_R^{(1)^2}(x)\}}\sqrt{\mathbb{E}\{\sigma_{2R}^{(2)^2}(x)\}}}\frac{1}{T}\int_0^T \rho_t\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)dt \\ &-\frac{1}{2T^2(\mathbb{E}\{\sigma_R^{(1)^2}(x)\})^{3/2}(\mathbb{E}\{\sigma_R^{(2)^2}(x)\})^{3/2}}\int_0^T \rho_t\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)dt \\ &\times\left\{\mathbb{E}\{\sigma_R^{(2)^2}(x)\}\int_0^T[(\sigma^{(1)}(x_s))^2-\mathbb{E}\{(\sigma^{(1)}(x_s))^2\}]ds\right. \\ &\left.+\,\mathbb{E}\{\sigma_R^{(1)^2}(x)\}\int_0^T[(\sigma^{(2)}(x_u))^2-\mathbb{E}\{(\sigma^{(2)}(x_u))^2\}]du\right\}, \end{aligned} \tag{3.61}$$

where (*cf.* Theorem 3.3.7) we have

$$\mathbb{E}\left\{\sigma^2_{(i)R}(x)\right\} = \mathbb{E}\left\{\frac{1}{T}\int_0^T(\sigma^{(i)}(x_t))^2dt\right\} = \frac{1}{T}\int_0^T e^{tQ}(\sigma^{(i)}(x))^2dt \ , \tag{3.62}$$

for $i=1,2$. We have to evaluate the expectation of the right hand side of Eq. (3.61). In Section 3.4.3 we computed the expectation of the first term. Then we will focus now on the other terms. First of all, let us rewrite them as follows

$$\begin{aligned} \int_0^T\int_0^T \rho_t\sigma^{(1)}(x_t)\sigma^{(2)}(x_t)&\Big(\mathbb{E}\{\sigma_R^{(2)^2}(x)\}[(\sigma^{(1)}(x_s))^2-\mathbb{E}\{(\sigma^{(1)}(x_s))^2\}] \\ &+\mathbb{E}\{\sigma_R^{(1)^2}(x)\}[(\sigma^{(2)}(x_s))^2-\mathbb{E}\{(\sigma^{(2)}(x_s))^2\}]\Big)\,dsdt \ . \end{aligned} \tag{3.63}$$

We have four different contributions to the integrals: the expectation values of the terms

$$\int_0^T \int_0^T \rho_t \sigma^{(1)}(x_t, \gamma(t)) \sigma^{(2)}(x_t) \mathbb{E}\{\sigma^{(i)^2}(x_s)\} \mathbb{E}\{\sigma^{(-i)^2}(x_s)\} dsdt \;, \tag{3.64}$$

for $i = 1, 2$, can be evaluate using Theorem 3.4.1. Then, in order to evaluate the expectation of the approximated realized correlation, it only remains to calculate

$$\mathbb{E}\left\{\int_0^T \int_0^T \rho_t \sigma^{(1)}(x_t) \sigma^{(2)}(x_t) \sigma^{(i)^2}(x_s) dsdt\right\} \qquad\qquad i = 1, 2 \;. \tag{3.65}$$

To this end, let's first divide the range of integration in two intervals as follows

$$\mathbb{E}\left\{\int_0^T \int_0^t \rho_t \sigma^{(1)}(x_t) \sigma^{(2)}(x_t) \sigma^{(i)^2}(x_s) dsdt + \int_0^T \int_t^T \rho_t \sigma^{(1)}(x_t) \sigma^{(2)}(x_t) \sigma^{(i)^2}(x_s) dsdt\right\},$$

for $i = 1, 2$. We notice that the first integral set is such that $t > s$ while the second has $t < s$. We can now use the property of conditional expectation to obtain

$$\begin{aligned} &\mathbb{E}\left\{\int_0^T \int_0^t \rho_t \mathbb{E}\{\sigma^{(1)}(x_t) \sigma^{(2)}(x_t) | \mathscr{F}_s\} \sigma^{(i)^2}(x_s) dsdt\right. \\ &\left. + \int_0^T \int_t^T \rho_t \sigma^{(1)}(x_t) \sigma^{(2)}(x_t) \mathbb{E}\{\sigma^{(i)^2}(x_s) | \mathscr{F}_t\} dsdt\right\}. \end{aligned} \tag{3.66}$$

Using the Markov property, we can express the conditional expectations as

$$\mathbb{E}\{\sigma^{(1)}(x_t) \sigma^{(2)}(x_t) | \mathscr{F}_s\} = e^{(t-s)Q} \sigma^{(1)}(x_s) \sigma^{(2)}(x_s) =: h(x_s) \;, \tag{3.67}$$

for $t > s$ and

$$\mathbb{E}\{\sigma^{(i)^2}(x_s) | \mathscr{F}_t\} = e^{(s-t)Q} \sigma^{(i)^2}(x_t) =: g^{(i)}(x_t) \;, \tag{3.68}$$

for $s > t$. Therefore, the first term of Eq. (3.66) can be expressed (*cf.* Proposition 3.2.4) as

$$\mathbb{E}\left\{\int_0^T \int_0^t \rho_t h(x_s) \sigma^{(i)^2}(x_s) dsdt\right\} = \int_0^T \int_0^t \rho_t e^{sQ} [h(x) \sigma^{(i)^2}(x)] dsdt \;, \tag{3.69}$$

while the second as

$$\begin{aligned} &\mathbb{E}\left\{\int_0^T \int_t^T \rho_t \sigma^{(1)}(x_t) \sigma^{(2)}(x_t) g^{(i)}(x_t) dsdt\right\} \\ &= \int_0^T \int_t^T \rho_t e^{tQ} [\sigma^{(1)}(x) \sigma^{(2)}(x) g^{(i)}(x)] dsdt \;. \end{aligned} \tag{3.70}$$

Now, we evaluate the functions h and g at x obtaining

$$h(x) = e^{(t-s)Q}[\sigma^{(1)}(x)\sigma^{(2)}(x)] \ , \tag{3.71}$$

and

$$g^{(i)}(x) = e^{(s-t)Q}[\sigma^{(i)^2}(x)] \ . \tag{3.72}$$

We can summarize the previous result in the following statement that gives the correlation swap price up to the first order of approximation.

Theorem 3.4.3. *The value of the correlation swap for a Markov-modulated volatility is*

$$P_{corr}(x) = e^{-rT}\left(\mathbb{E}\{Corr_R(S^1,S^2)\} - K_{corr}\right) \ , \tag{3.73}$$

where the realized correlation can be approximated by

$$\begin{aligned}\mathbb{E}\{Corr_R(S^1,S^2)\} \approx \ & \frac{2\int_0^T \rho_t e^{tQ}\sigma^{(1)}(x)\sigma^{(2)}(x)dt}{\sqrt{\int_0^T e^{tQ}(\sigma^{(1)}(x))^2dt}\sqrt{\int_0^T e^{tQ}(\sigma^{(2)}(x))^2dt}} \\ & - \frac{\int_0^T \rho_t(\int_0^t e^{sQ}\{e^{(t-s)Q}[\sigma^{(1)}(x)\sigma^{(2)}(x)]\sigma^{(1)^2}(x)\}ds + \int_t^T e^{tQ}\{\sigma^{(1)}(x)\sigma^{(2)}(x)e^{(u-t)Q}[\sigma^{(1)^2}(x)]\}du)dt}{2\left(\int_0^T e^{tQ}(\sigma^{(1)}(x))^2dt\right)^{3/2}\left(\int_0^T e^{tQ}(\sigma^{(2)}(x))^2dt\right)^{1/2}} \\ & - \frac{\int_0^T \rho_t(\int_0^t e^{sQ}\{e^{(t-s)Q}[\sigma^{(1)}(x)\sigma^{(2)}(x)]\sigma^{(2)^2}(x)\}ds + \int_t^T e^{tQ}\{\sigma^{(1)}(x)\sigma^{(2)}(x)e^{(u-t)Q}[\sigma^{(2)^2}(x)]\}du)dt}{2\left(\int_0^T e^{tQ}(\sigma^{(1)}(x))^2dt\right)^{1/2}\left(\int_0^T e^{tQ}(\sigma^{(2)}(x))^2dt\right)^{3/2}} \ .\end{aligned}$$

3.5 Example: Variance, Volatility, Covariance and Correlation Swaps for Stochastic Volatility Driven by Two State Continuos Markov Chain

Let $(x_t)_{t\in\mathbb{R}_+}$ be a two state continuous time Markov chain, let us denote the states as 'Up' (u) and 'Down' (d). Let Q be the generator of this Markov chain

$$Q = \begin{pmatrix} q_{uu} & q_{ud} \\ q_{du} & q_{dd} \end{pmatrix} \ , \tag{3.74}$$

and let

$$P(t) = \begin{pmatrix} p_{uu}(t) & p_{ud}(t) \\ p_{du}(t) & p_{dd}(t) \end{pmatrix} \ , \tag{3.75}$$

be its transition function, such that

$$P(t) = e^{tQ} \ . \tag{3.76}$$

In this simple model the volatility takes only two values: σ_u and σ_d, thus we can easily express the swap prices of the futures contract so far studied.

The variance swap price in this model is given by

$$P_{var}(i) = e^{-rT} \left\{ \frac{1}{T} \int_0^T (p_{iu}(t)\sigma_u^2 + p_{id}(t)\sigma_d^2) dt - K_{var} \right\} \ , \tag{3.77}$$

where $i = u, d$ is the initial state of the Markov chain. If we are uncertain about the initial state and we have only a probability distribution, say (p_u, p_d) such that $p_u + p_d$, then the price is going be

$$P_{var} = p_u P_{var}(u) + p_d P_{var}(d) \ . \tag{3.78}$$

If we assume that initial distribution is actually the stationary distribution of Markov chain the price simply becomes

$$P_{var} = e^{-rT} \left\{ \pi_u \sigma_u^2 + \pi_d \sigma_d^2 - K_{var} \right\} , \tag{3.79}$$

where (π_u, π_d) is the stationary distribution.

The volatility swap price in this model can be approximateted by

$$P_{vol}(i) \approx \quad e^{-rT} \left\{ \sqrt{\tfrac{1}{T} \int_0^T [p_{iu}(t)\sigma_u^2 + p_{id}(t)\sigma_d^2] dt} - \frac{Var\{\sigma_R^2(i)\}}{8\left(\frac{1}{T}\int_0^T [p_{iu}(t)\sigma_u^2 + p_{id}(t)\sigma_d^2] dt\right)^{3/2}} - K_{vol} \right\} ,$$

for $i = u, d$ and where

$$\begin{aligned} Var\{\sigma_R^2(i)\} = & \frac{2}{T^2} \int_0^T \int_0^t \{p_{iu}(s)p_{uu}(t-s)\sigma_u^4 + [p_{iu}(s)p_{ud}(t-s) + p_{id}(s)p_{du}(t-s)]\sigma_d^2\sigma_u^2 \\ & + p_{id}(s)p_{dd}(t-s)\sigma_d^4\} ds dt - \left[\frac{1}{T} \int_0^T [p_{iu}(t)\sigma_u^2 + p_{id}(t)\sigma_d^2] dt \right]^2 . \end{aligned}$$

Let us now consider a two risky asset market model with volatility modulated by this two state Markov chain. In this setting the covariance swap price is

$$P_{cov}(i) = e^{-rT} \left\{ \frac{1}{T} \int_0^T \rho_t [p_{iu}(t)\sigma_u^{(1)}\sigma_u^{(2)} + p_{id}(t)\sigma_d^{(1)}\sigma_d^{(2)}] dt - K_{cov} \right\} , \tag{3.80}$$

for $i = u, d$ representing the initial state of the chain.

The correlation swap in this model, at the zero order, can be approximated by

$$P_{corr}(i) \approx e^{-rT}\left[\frac{\int_0^T \rho_t[p_{iu}(t)\sigma_u^{(1)}\sigma_u^{(2)} + p_{id}(t)\sigma_d^{(1)}\sigma_d^{(2)}]dt}{\sqrt{\int_0^T [p_{iu}(t)(\sigma_u^{(1)})^2 + p_{id}(t)(\sigma_d^{(1)})^2]dt}\sqrt{\int_0^T [p_{iu}(t)(\sigma_u^{(2)})^2 + p_{id}(t)(\sigma_d^{(2)})^2]dt}} - K_{corr}\right],$$

for $i = u, d$ being the initial state of the chain.

3.6 Numerical Example

3.6.1 S&P 500: Variance and Volatility Swaps

In this section, we give an example of the two states Markov chain model for stochastic volatility modulating a single risky asset market. We use S&P 500 index as risky asset.

In order to estimate the Markov chain transition matrix and the parameters (σ_u, σ_d) we use the CBOE Volatility Index (VIX), daily data from January 2004 to June 2012.

For every day t, a high data v_t^h and a low data v_t^l are available. We interpolated between them defining $v_t = \frac{1}{2}(v_t^h + v_t^l)$, a reference value for day t. Taking a mean of such values we defined

$$\bar{v} = \frac{1}{n}\sum_{t=1}^{n} v_t \ . \tag{3.81}$$

If $v_t > \bar{v}$ the day t state is *Up* otherwise it is *Down*. The transition matrix is estimated The value σ_u is the mean of v_t evaluated only on the *Up* days and similarly σ_d is the mean of v_t evaluated only on the *Down* days. In table 3.1 we show the one step transition probability matrix.

Transition Matrix		
	Up	Down
Up	0.957	0.043
Down	0.026	0.974

Table 3.1: One step transition probability matrix.

Using this probability matrix, given the initial state of the chain, we are able to evaluate the price of covariance and correlation swap as described in Section 3.5. In Figure 3.1 the variance and volatility swap prices as a function of maturity are shown. The rate of convergence of the prices depend on the rate of convergence of the transition probability matrix to the stationary distribution.

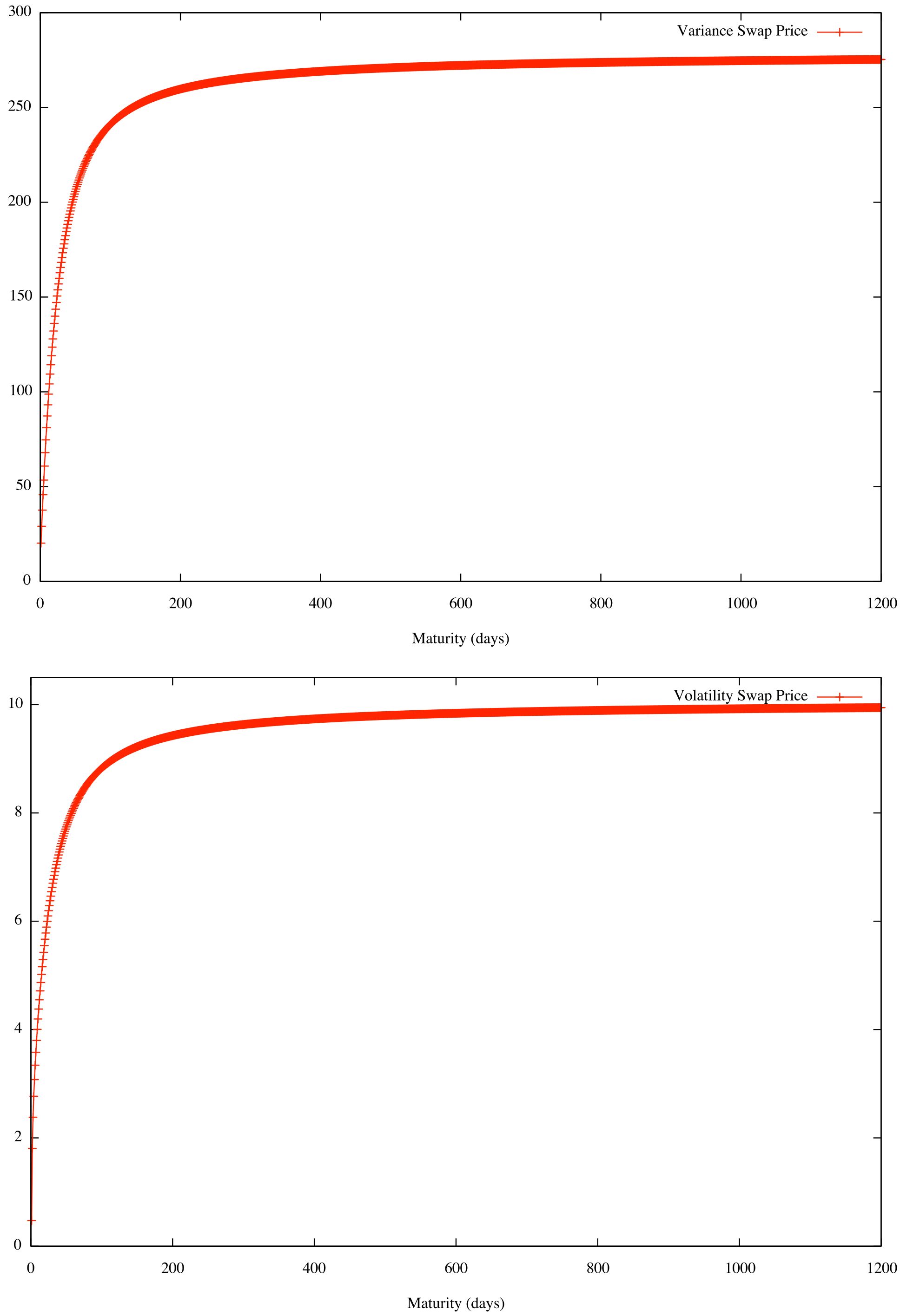


Figure 3.1: Variance and volatility swap prices.

3.6.2 S&P 500 and NASDAQ-100: Covariance and Correlation Swaps

In this section, we give an example of the two states Markov chain model for stochastic volatility modulating a two risky assets market. We use S&P 500 index and NASDAQ-100 index as risky assets.

In order to estimate the Markov chain transition matrix and the parameters $(\sigma_u^{(1)}, \sigma_d^{(1)}, \sigma_u^{(2)}, \sigma_d^{(2)})$ we use the CBOE Volatility Index (VIX) and the CBOE NASDAQ-100 Volatility Index (VXN), from January 2004 to June 2012.

The volatilities of the two assets are modulated by the same Markov chain, then we can consider the two volatility indices as two independent realization of the same process.

For each index and a given day t, a high data (vix_t^h, vxn_t^h) and a low data (vix_t^l, vxn_t^l) are available. We interpolated between them defining a VIX and VXN reference value for day t as

$$vix_t = \frac{1}{2}(vix_t^h + vix_t^l) \ ,$$

and

$$vxn_t = \frac{1}{2}(vxn_t^h + vxn_t^l) \ .$$

We evaluate the mean value $(\overline{vix}, \overline{vxn})$ over the period. If the day t $vix_t > \overline{vix}$ VIX is in the state *Up* otherwise it is in *Down*, similarly if $vxn_t > \overline{vxn}$ VXN is in *Up* otherwise it is in *Down*. This procedure create two (independent) sequences of *Up* and *Down* states of our Markov process: one for VIX and the other for VXN. Using these sequences we can estimate the transition probability for the Markov chain. Regarding the parameters, $\sigma_u^{(1)}$ can be estimated by taking the mean of all the VIX values, vix_t, such that t is an *Up* day and $\sigma_d^{(1)}$ by taking the mean only of vix_t such that t is a *Down* day. Similarly we can estimate $(\sigma_u^{(2)}, \sigma_d^{(2)})$ can be estimated by taking the mean over the *Up* and *Down* days of VXN, respectively. In table **??** we show the one step transition probability matrix built using VIX and VXN indexes.

Using this probability matrix, given the initial state of the chain, we are able to evaluate the price of covariance and correlation swap as described in Section 3.5. In the numerical evaluation we assume that the correlation between the brownian motion ρ (*cf.* Eq. 3.33) is constant.

In Figure 3.2 the covariance and correlation swap prices as a function of maturity are shown. The rate

Transition Matrix		
	Up	Down
Up	0.950	0.050
Down	0.030	0.970

Table 3.2: One step transition probability matrix.

of convergence of the prices depends on the rate of convergence of the transition probability matrix to the stationary distribution. For the correlation swap price we show the result up to the zero order of approximation, as described in Section 3.5, and up to the first order of approximation, as described in Section 3.4.4. The zero order of approximation price is constant, indeed at that order we only take into account the correlation of the assets noise, that is ρ the brownian motions correlation. In the first order the interaction of the volatilities due to the Markov process is taken into account and the price changes as a function of maturity.

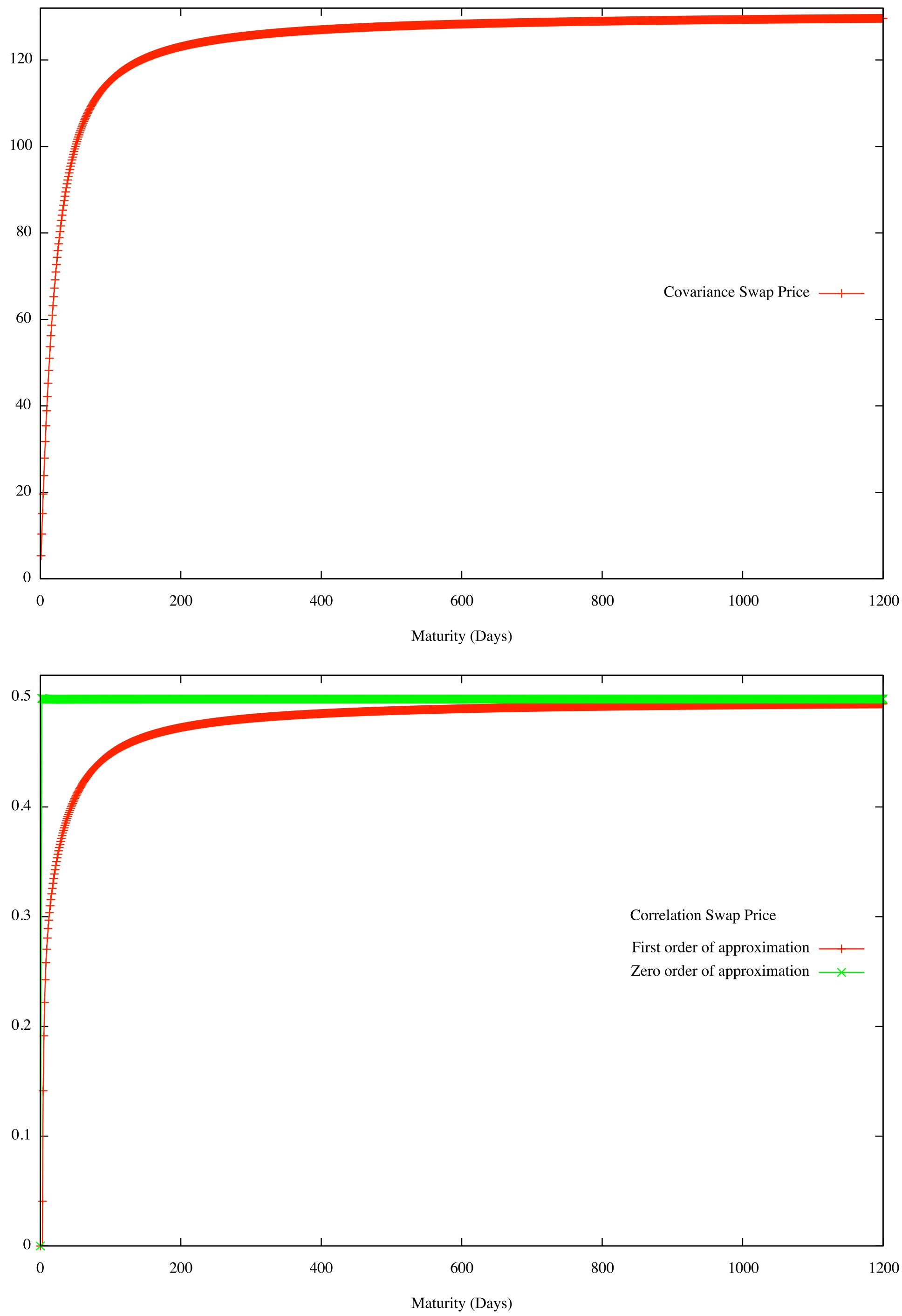


Figure 3.2: Covariance and correlation swap prices.

3.7 Conclusion

In a Markov-modulated stochastic volatility model an expression for the price of variance, volatility, covariance and correlation swap has been obtained. The variance swap price and volatility swap price for a single risky assets market have been studied. The covariance swap price in a two risky assets market has been obtained and an approximated expression for the correlation swap price has been derived. The results have been applied in a two state continuos Markov chain volatility model. A numerical example using data of indices S& P 500 and NASDAQ 100 has been discussed.

Chapter 4

Financial Markets with Semi-Markov Modulated Stochastic Volatilities

In this chapter, we model financial markets with semi-Markov volatilities and price covariance and correlation swaps for this markets. Numerical evaluations of variance, volatility, covariance and correlations swaps with semi-Markov volatility are presented as well. This chapter is a generalization to the more flexible semi-Markov environment of what we discuss in Chapter 3 in Markovian case.

The chapter is organized as follows. First, martingale representation of semi-Markov processes is presented. Then, Variance and volatility swaps for financial markets with semi-Markov modulated stochastic volatilities are studied. Furthermore, pricing of covariance and correlation swaps for a two risky assets for financial markets with semi-Markov modulated stochastic volatilities is presented. The difference between Markov and semi-Markov case will be emphasized.

This chapter is based on the work done while visiting University of Calgary, Calgary (AB) Canada, under the supervision of Prof. A. V. Swishchuk.

This chapter is based on a paper (G. Salvi and A. V. Swishchuk [76]) currently under review.

4.1 Introduction

One of the recent and new financial products are variance and volatility swaps, which are useful for volatility hedging and speculation. The market for variance and volatility swaps has been growing, and many investment banks and other financial institutions are now actively quoting volatility swaps on various assets: stock indexes, currencies, as well as commodities. A stock's volatility is the simplest measure of its riskiness or uncertainty. Formally, the volatility σ_R is the annualized standard deviation of the stock's returns during the period of interest, where the subscript R denotes the observed or 'realized' volatility. Why trade volatility or variance? As mentioned in M. Broadie and A. Jain [16], 'just as stock investors think they know something about the direction of the stock market so we may think we have insight into the level of future volatility. If we think current volatility is low, for the right price we might want to take a position that profits if volatility increases'.

In this chapter, we model financial markets with semi-Markov volatilities and price covariance and correlation swaps for this markets. Numerical evaluations of variance, volatility, covariance and correlations swaps with semi-Markov volatility are presented as well.

Volatility swaps are forward contracts on future realized stock volatility, variance swaps are similar contract on variance, the square of the future volatility, both these instruments provide an easy way for investors to gain exposure to the future level of volatility. A stock's volatility is the simplest measure of its riskness or uncertainty. Formally, the volatility σ_R is the annualized standard deviation of the stock's returns during the period of interest, where the subscript R denotes the observed or "realized" volatility.

The easy way to trade volatility is to use volatility swaps, sometimes called realized volatility forward contracts, because they provide pure exposure to volatility (and only to volatility).

A stock *volatility swap* is a forward contract on the annualized volatility. Its payoff at expiration is equal to

$$N(\sigma_R(S) - K_{vol}) \ ,$$

where $\sigma_R(S)$ is the realized stock volatility (quoted in annual terms) over the life of contract,

$$\sigma_R(S) := \sqrt{\frac{1}{T}\int_0^T \sigma_s^2 ds} \ ,$$

σ_t is a stochastic stock volatility, K_{vol} is the annualized volatility delivery price, and N is the notional amount of the swap in dollar per annualized volatility point. The holder of a volatility swap at expiration receives N dollars for every point by which the stock's realized volatility σ_R has exceeded the volatility delivery price K_{vol}. The holder is swapping a fixed volatility K_{vol} for the actual (floating) future volatility

σ_R. We note that usually $N = \alpha I$, where α is a converting parameter such as 1 per volatility-square, and I is a long-short index (+1 for long and −1 for short).

Although options market participants talk of volatility, it is its variance, or volatility squared, that has more fundamental significance (see Demeterfi *et al* [45]).

A variance swap is a forward contract on annualized variance, the square of the realized volatility. Its payoff at expiration is equal to

$$N(\sigma_R^2(S) - K_{var}) \ ,$$

where $\sigma_R^2(S)$ is the realized stock variance(quoted in annual terms) over the life of the contract,

$$\sigma_R^2(S) := \frac{1}{T}\int_0^T \sigma_s^2 ds \ ,$$

K_{var} is the delivery price for variance, and N is the notional amount of the swap in dollars per annualized volatility point squared. The holder of variance swap at expiration receives N dollars for every point by which the stock's realized variance $\sigma_R^2(S)$ has exceeded the variance delivery price K_{var}.

Therefore, pricing the variance swap reduces to calculating the realized volatility square.

Valuing a variance forward contract or swap is no different from valuing any other derivative security. The value of a forward contract P on future realized variance with strike price K_{var} is the expected present value of the future payoff in the risk-neutral world:

$$P = E\{e^{-rT}(\sigma_R^2(S) - K_{var})\} \ ,$$

where r is the risk-free discount rate corresponding to the expiration date T, and E denotes the expectation.

Thus, for calculating variance swaps we need to know only $E\{\sigma_R^2(S)\}$, namely, the mean value of the underlying variance.

To calculate volatility swaps we need more. From Brockhaus and Long [17] approximation (which is used the second order Taylor expansion for function $\sqrt{x}$) we have (see also Javaheri *et al* [61]):

$$E\{\sqrt{\sigma_R^2(S)}\} \approx \sqrt{E\{V\}} - \frac{Var\{V\}}{8E\{V\}^{3/2}} \ ,$$

where $V := \sigma_R^2(S)$ and $\frac{Var\{V\}}{8E\{V\}^{3/2}}$ is the convexity adjustment.

Thus, to calculate volatility swaps we need both $E\{V\}$ and $Var\{V\}$.

The realized continuously sampled variance is defined in the following way:

$$V := Var(S) := \frac{1}{T}\int_0^T \sigma_t^2 dt \ .$$

Realized continuously sampled volatility is defined as follows:

$$Vol(S) := \sqrt{Var(S)} = \sqrt{V} \ .$$

Options dependent on exchange rate movements, such as those paying in a currency different from the underlying currency, have an exposure to movements of the correlation between the asset and the exchange rate. This risk may be eliminated by using covariance swap. Variance and volatility swaps have been studied by Swishchuk [81]. The novelty of this paper with respect to Swishchuk [81] is that we calculate the volatility swap price explicitly; moreover we price covariance and correlation swap in a two risky assets market model.

A *covariance swap* is a covariance forward contact of the underlying rates S^1 and S^2 which payoff at expiration is equal to

$$N(Cov_R(S^1, S^2) - K_{cov}) \ ,$$

where K_{cov} is a strike price, N is the notional amount, $Cov_R(S^1, S^2)$ is a covariance between two assets S^1 and S^2.

Logically, a *correlation swap* is a correlation forward contract of two underlying rates S^1 and S^2 which payoff at expiration is equal to:

$$N(Corr_R(S^1, S^2) - K_{corr}) \ ,$$

where $Corr(S^1, S^2)$ is a realized correlation of two underlying assets S^1 and S^2, K_{corr} is a strike price, N is the notional amount.

Pricing covariance swap, from a theoretical point of view, is similar to pricing variance swaps, since

$$Cov_R(S^1, S^2) = 1/4\{\sigma_R^2(S^1 S^2) - \sigma_R^2(S^1/S^2)\} \ ,$$

where S^1 and S^2 are two given assets, $\sigma_R^2(S)$ is a variance swap for underlying assets, $Cov_R(S^1, S^2)$ is a realized covariance of the two underlying assets.

Thus, we need to know variances for S^1S^2 and for S^1/S^2. Correlation $Corr_R(S^1, S^2)$ is defined as follows:

$$Corr_R(S^1, S^2) = \frac{Cov_R(S^1, S^2)}{\sqrt{\sigma_R^2(S^1)}\sqrt{\sigma_R^2(S^2)}} \ ,$$

where $Cov_R(S^1, S^2)$ is defined above.

Given two assets S_t^1 and S_t^2 with $t \in [0, T]$, sampled on days $t_0 = 0 < t_1 < t_2 < ... < t_n = T$ between today and maturity T, the log-return of each asset is: $R_i^j := \log(\frac{S_{t_i}^j}{S_{t_{i-1}}^j}), \quad i = 1, 2, ..., n, \quad j = 1, 2.$

Covariance and correlation can be approximated by

$$Cov_n(S^1, S^2) = \frac{n}{(n-1)T} \sum_{i=1}^{n} R_i^1 R_i^2 \ ,$$

and

$$Corr_n(S^1, S^2) = \frac{Cov_n(S^1, S^2)}{\sqrt{Var_n(S^1)}\sqrt{Var_n(S^2)}} \ ,$$

respectively.

The literature devoted to the volatility derivatives is growing. We give here a short overview of the latest development in this area. The Non-Gaussian Ornstein-Uhlenbeck stochastic volatility model was used by Benth *et al.* [4] to study volatility and variance swaps. M. Broadie and A. Jain [15] evaluated price and hedging strategy for volatility derivatives in the Heston square root stochastic volatility model and in M. Broadie and A. Jain [16] they compare result from various model in order to investigate the effect of jumps and discrete sampling on variance and volatility swaps. Pure jump process with independent increments return models were used by Carr *et al.* [19] to price derivatives written on realized variance, and subsequent development by Carr and Lee [20]. We also refer to Carr and Lee [21] for a good survey on volatility derivatives. Da Fonseca *et al.* [43] analyzed the influence of variance and covariance swap in a market by solving a portfolio optimization problem in a market with risky assets and volatility derivatives. Correlation swap price has been investigated by Bossu [12, 13] for component of an equity index using statistical method. Drissien *et al.* [46] discusses the price of correlation risk for equity options. Pricing volatility swaps under Heston's model with regime-switching and pricing options under a generalized Markov-modulated jump-diffusion model are discussed in and Elliott *et al.* [48, 49], respectively. Howison *et al.* [53] considers the pricing of a range of volatility derivatives, including volatility and variance swaps and swaptions. The pricing options on realized variance in the Heston model with jumps in returns and volatility is studied in Sepp [77]. An analytical closed-forms pricing of pseudo-variance, pseudo-volatility, pseudo-covariance and pseudo-correlation swaps is studied in Swishchuk *et al.* [82]. Windcliff *et al.* [87] investigates the behavior and hedging of discretely observed volatility derivatives.

4.2 Martingale Representation of Semi-Markov Processes

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ be a filtered probability space, with a right-continuous filtration $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$ and probability $\mathbb{P}$.

Let $(X, \mathscr{X})$ be a measurable space and

$$Q_{SM}(x, B, t) := P(x, B)G_x(t) \qquad \text{for } x \in X, B \in \mathscr{X}, t \in \mathbb{R}_+ \ , \tag{4.1}$$

be a semi-Markov kernel. Let $(x_n, \tau_n; n \in \mathbb{N})$ be a $(X \times \mathbb{R}_+, \mathscr{X} \otimes \mathscr{B}_+)$-valued Markov renewal process with Q_{SM} the associated kernel, that is

$$\mathbb{P}(x_{n+1} \in B, \tau_{n+1} - \tau_n \leq t \mid \mathscr{F}_n) = Q_{SM}(x_n, B, t) \ . \tag{4.2}$$

Let us define the process

$$\nu_t := sup\{n \in \mathbb{N} : \tau_n \leq t\} \ , \tag{4.3}$$

that gives the number of jumps of the Markov renewal process in the time interval $(0, t]$ and

$$\theta_n := \tau_n - \tau_{n-1} \ , \tag{4.4}$$

that gives the sojourn time of the Markov renewal process in the n-th visited state. The semi-Markov process, associated with the Markov renewal process $(x_n, \tau_n)_{n\in\mathbb{N}}$, is defined by

$$x_t := x_{\nu(t)} \qquad \text{for } t \in \mathbb{R}_+ \ . \tag{4.5}$$

Associated with the semi-Markov process, it is possible to define some auxiliaries processes. We are interested in the backward recurrence time (or life-time) process defined by

$$\gamma(t) := t - \tau_{\nu(t)} \qquad \text{for } t \in \mathbb{R}_+ \ . \tag{4.6}$$

The next result characterizes backward recurrence time process (*cf.* Swishchuk [81]), we give the proof for completeness.

Proposition 4.2.1. *The backward recurrence time* $(\gamma(t))_t$ *is a Markov process with generator*

$$Q_\gamma f(t) = f'(t) + \lambda(t)[f(0) - f(t)] \ , \tag{4.7}$$

where $\lambda(t) = -\frac{\overline{G_x}'(t)}{\overline{G_x}(t)}$, $\overline{G_x}(t) = 1 - G_x(t)$ *and* $Domain(Q_\gamma) = C^1(\mathbb{R}_+)$.

Proof. Let t be the present time such that $\gamma(t) = t$, without loss of generality we can assume that $t < \tau_1$, then for $T > t$ we have

$$\mathbb{E}_t\{f(\gamma(T))\} = \mathbb{E}_t\{f(\gamma(T))\mathbb{I}_{\theta_1 > T}\} + \mathbb{E}_t\{f(\gamma(T))\mathbb{I}_{\theta_1 \le T}\} \ . \tag{4.8}$$

Using the properties of conditional expectation we obtain

$$\begin{aligned}\mathbb{E}_t\{f(\gamma(T))\} =& f(T)\frac{\overline{G}_x(T)}{\overline{G}_x(t)} + \frac{1}{\overline{G}_x(t)}\mathbb{E}\{f(\gamma(T))\mathbb{I}_{t<\theta_1 \le T}\} \\ =& f(T)\frac{\overline{G}_x(T)}{\overline{G}_x(t)} + \frac{1}{\overline{G}_x(t)}\int_t^T f(T-u)G'_x(u)du \ .\end{aligned} \tag{4.9}$$

By adding and subtracting $f(t)$ in the integrand we get

$$\mathbb{E}_t\{f(\gamma(T))\} = f(T)\frac{\overline{G}_x(T)}{\overline{G}_x(t)} + \frac{1}{\overline{G}_x(t)}\int_t^T (f(T-u) - f(t))G'_x(u)du + f(t)\frac{\overline{G}_x(t) - \overline{G}_x(T)}{\overline{G}_x(t)} \ ,$$

then

$$\mathbb{E}_t\{f(\gamma(T))\} - f(t) = (f(T) - f(t))\frac{\overline{G}_x(T)}{\overline{G}_x(t)} + \frac{1}{\overline{G}_x(t)}\int_t^T (f(T-u) - f(t))G'_x(u)du \ . \tag{4.10}$$

Recalling the definition of the generator and using the above equation we have

$$Q_\gamma f(t) = \lim_{T \to t}\frac{\mathbb{E}_t\{f(\gamma(T))\} - f(t)}{T - t} = f'(t) - \frac{\overline{G_x}'(t)}{\overline{G_x}(t)}[f(0) - f(t)] \ , \tag{4.11}$$

this concludes the proof. □

Remark 4.2.2. *As well known, semi-Markov processes preserve the lost-memories property only at transition times, then* $(x_t)_{t \in \mathbb{R}_+}$ *is not Markov. However, if we consider the joint process* $(x_t, \gamma(t))_{t \in \mathbb{R}_+}$*, and we record at any instant the time already spent by the semi-Markov process in the present state, then it result that* $(x_t, \gamma(t))_{t \in \mathbb{R}_+}$ *is a Markov process.*

In Section 3.2 we discussed the martingale representation of a Markov process. Here $(x_t)_t$ belongs to a wider class, but if we consider the joint process with the backward recurrence time we can obtain a martingale representation for the semi-Markov process as well.

We would like to study the martingale associated to the Markov process $(x_t, \gamma(t))_{t\in\mathbb{R}_+}$ and its generator. The following statement concerns this task, and is a direct application of Proposition 3.2.1.

Lemma 4.2.3. *(Swishchuk [81])*

Let $(x_t)_{t\in\mathbb{R}_+}$ be a semi-Markov process with kernel Q_{SM} defined in Eq. (4.1). Then, the process

$$m_t^f := f(x_t, \gamma(t)) - \int_0^t Qf(x_s, \gamma(s))ds \ , \tag{4.12}$$

is a martingale with respect to the filtration $\mathscr{F}_t := \sigma\{x_s, \nu_s; 0 \le s \le t\}$, where Q is the generator of the Markov process $(x_t, \gamma(t))_{t\in\mathbb{R}_+}$ given by

$$Qf(x,t) = \frac{df}{dt}(x,t) + \frac{g_x(t)}{\overline{G}_x(t)} \int_X P(x,dy)[f(y,0) - f(x,t)] \ , \tag{4.13}$$

here $g_x(t) = \frac{dG_x(t)}{dt}$.

The following statement follows directly from Proposition 3.2.2 and it allows us to evaluate the quadratic variation of the martingale associated with $(x_t, \gamma(t))_{t\in\mathbb{R}_+}$.

Lemma 4.2.4. *Let $(x_t)_{t\in\mathbb{R}_+}$ be a semi-Markov process with kernel Q_{SM}, $(x_t, \gamma(t))_{t\in\mathbb{R}_+}$ is a Markov process with generator Q, $f \in$ Domain (Q) and $(m_t^f)_{t\in\mathbb{R}_+}$ its associated martingale, then*

$$\langle m^f \rangle_t := \int_0^t [Qf^2(x_s, \gamma(s)) - 2f(x_s, \gamma(s))Qf(x_s, \gamma(s))]ds \ , \tag{4.14}$$

is the quadratic variation of m^f.

4.3 Variance and Volatility Swaps for Financial Markets with Semi-Markov Stochastic Volatilities

Let us consider a Market model with only two securities, the risk free bond and the stock. Let us suppose that the stock price $(S_t)_{t\in\mathbb{R}_+}$ satisfies the following stochastic differential equation

$$\begin{cases} dB_t = B_t r(x_t, \gamma(t))dt \\ dS_t = S_t(\mu(x_t, \gamma(t))dt + \sigma(x_t, \gamma(t))dw_t) \end{cases} , \tag{4.15}$$

where w is a standard Wiener process independent of (x, γ). We are interested in studying the property of the volatility $\sigma(x, \gamma)$. Salvi and Swishchuk [75] have studied properties of volatility modulated by a Markov process. Here we would like to generalize their work to the semi-Markov case. First of all we study the second moment of the volatility.

Proposition 4.3.1. *(Swishchuk [81])*

Suppose that $\sigma \in$ *Domain* (Q). *Then*

$$\mathbb{E}\{\sigma^2(x_t, \gamma(t))|\mathscr{F}_u\} = \sigma^2(x_u, \gamma(u)) + \int_u^t Q\mathbb{E}\{\sigma^2(x_s, \gamma(s))|\mathscr{F}_u\}ds \ , \tag{4.16}$$

for any $0 \leq u \leq t$.

Remark 4.3.2. *From Proposition 4.3.1, we can directly solve the equation for* $\mathbb{E}\{\sigma^2(x_t, \gamma(t))|\mathscr{F}_u\}$ *and we obtain*

$$\mathbb{E}\{\sigma^2(x_t, \gamma(t))|\mathscr{F}_u\} = e^{(t-u)Q}\sigma^2(x_u, \gamma(u)) \ , \tag{4.17}$$

for any $0 \leq u \leq t$.

It is known that the market model with semi-Markov stochastic volatility is incomplete, see Swishchuk [81]. In order to price the future contracts we will use the minimal martingale measure, we refer to Swishchuk [81] for the details.

4.3.1 Pricing of Variance Swaps

Let us start from the more straightforward variance swap. Variance swaps are forward contracts on future realized level of variance. The payoff of a variance swap with expiration date T is given by

$$N(\sigma_R^2(x) - K_{var}) \ , \tag{4.18}$$

where $\sigma_R^2(x)$ is the realized stock variance over the life of the contract

$$\sigma_R^2(x) := \frac{1}{T}\int_0^T \sigma^2(x_s, \gamma(s))ds \ , \tag{4.19}$$

K_{var} is the strike price for variance and N is the notional amounts of dollars per annualized variance point, we will assume that $N = 1$ just for sake of simplicity. The price of the variance swap is the expected present value of the payoff in the risk-neutral world

$$P_{var}(x) = \mathbb{E}\{e^{-rT}(\sigma_R^2(x) - K_{var})\} \ . \tag{4.20}$$

The following result concerns the evaluation of a variance swap in this semi-Markov volatility model. We refer to Swishchuk [81] for details and proof.

Theorem 4.3.3. *(Swishchuk [81])*

The present value of a variance swap for semi-Markov stochastic volatility is

$$P_{var}(x) = e^{-rT}\left\{\frac{1}{T}\int_0^T (e^{tQ}\sigma^2(x,\gamma) - K_{var})dt\right\} \ , \tag{4.21}$$

where Q is the generator of $(x_t, \gamma(t))_t$, that is

$$Qf(x,t) = \frac{df}{dt}(x,t) + \frac{g_x(t)}{\overline{G}_x(t)}\int_X P(x,dy)[f(y,0) - f(x,t)] \ . \tag{4.22}$$

4.3.2 Pricing of Volatility Swaps

Volatility swaps are forward contract on future realized level of volatility. The payoff of a volatility swap with maturity T is given by

$$N(\sigma_R(x) - K_{vol}) \ , \tag{4.23}$$

where $\sigma_R(x)$ is the realized stock volatility over the life of the contract

$$\sigma_R(x) := \sqrt{\frac{1}{T}\int_0^T \sigma^2(x_s, \gamma(s))ds} \ , \tag{4.24}$$

K_{vol} is the strike price for volatility and N is the notional amounts of dollars per annualized volatility point, as before we will assume that $N = 1$. The price of the volatility swap is the expected present value of the payoff in the risk-neutral world

$$P_{vol}(x) = \mathbb{E}\{e^{-rT}(\sigma_R(x) - K_{vol})\} \ . \tag{4.25}$$

In order to evaluate the volatility swaps we need to know the expected value of the square root of the variance, but unfortunately, in general we are not able to evaluate analytically this expected value. Then in order to obtain a close formula for the price of volatility swaps we have to make an approximation. Using the same approach of the Markov case (see also Brockhaus and Long [17] and Javaheri *at al.* [61]), from the second order Taylor expansion we have

$$\mathbb{E}\{\sqrt{\sigma_R^2(x)}\} \approx \sqrt{\mathbb{E}\{\sigma_R^2(x)\}} - \frac{Var\{\sigma_R^2(x)\}}{8\mathbb{E}\{\sigma_R^2(x)\}^{3/2}} \ . \tag{4.26}$$

Then, to evaluate the volatility swap price we have to know both expectation and variance of $\sigma_R^2(x)$. The next result gives an explicit representation of the price of a volatility swap approximated to the second order for this semi-Markov volatility model.

Theorem 4.3.4. *The value of a volatility swap for semi-Markov stochastic volatility is*

$$P_{vol}(x) \approx \quad e^{-rT}\left\{\sqrt{\tfrac{1}{T}\int_0^T e^{tQ}\sigma^2(x,\gamma)dt} - \frac{Var\{\sigma_R^2(x)\}}{8\left(\frac{1}{T}\int_0^T e^{tQ}\sigma^2(x,\gamma)dt\right)^{3/2}} - K_{vol}\right\},$$

where the variance is given by

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T\int_0^t \left\{e^{sQ}\left[\sigma^2(x,\gamma)e^{(t-s)Q}\sigma^2(x,\gamma)\right] - \left[e^{tQ}\sigma^2(x,\gamma)\right]\left[e^{sQ}\sigma^2(x,\gamma)\right]\right\} dsdt \ ,$$

and Q is the generator of $(x_t, \gamma(t))_t$, that is

$$Qf(x,t) = \frac{df}{dt}(x,t) + \frac{g_x(t)}{\bar{G}_x(t)}\int_X P(x,dy)[f(y,0) - f(x,t)] \ . \tag{4.27}$$

Proof. We have already obtained the expectation of the realized variance,

$$\mathbb{E}\{\sigma_R^2(x)\} = \frac{1}{T}\int_0^T e^{tQ}\sigma^2(x,\gamma)dt \ , \tag{4.28}$$

then it remains to prove that

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T\int_0^t \left\{e^{sQ}\left[\sigma^2(x,\gamma)e^{tQ}\sigma^2(x,\gamma)\right] - \left[e^{tQ}\sigma^2(x,\gamma)\right]\left[e^{sQ}\sigma^2(x,\gamma)\right]\right\} dsdt \ .$$

The variance is, from the definition, given by

$$Var\{\sigma_R^2(x)\} = \mathbb{E}\{[\sigma_R^2(x) - \mathbb{E}\{\sigma_R^2(x)\}]^2\} \ , \tag{4.29}$$

Using the definition of realized variance, and Fubini theorem, we have

$$
\begin{aligned}
Var\{\sigma_R^2(x)\} &= \mathbb{E}\left\{\left[\frac{1}{T}\int_0^T \sigma^2(x_t,\gamma(t))dt - \frac{1}{T}\int_0^T \mathbb{E}\{\sigma^2(x_t,\gamma(t))\}dt\right]^2\right\} \\
&= \mathbb{E}\left\{\left[\frac{1}{T}\int_0^T \left(\sigma^2(x_t,\gamma(t)) - \mathbb{E}\{\sigma^2(x_t,\gamma(t))\}\right)dt\right]^2\right\} \qquad (4.30)\\
&= \frac{1}{T^2}\int_0^T\int_0^T \mathbb{E}\left\{[\sigma^2(x_t,\gamma(t)) - \mathbb{E}\{\sigma^2(x_t,\gamma(t))\}][\sigma^2(x_s,\gamma(s)) - \mathbb{E}\{\sigma^2(x_s,\gamma(s))\}]\right\} dsdt\ .
\end{aligned}
$$

We note that the integrand is symmetric in the exchange of s and t. We can divide the integration on the plan in two areas above and below the graph of $t = s$, thanks to the symmetry the contribution on the two parts is the same. Then we obtain

$$
\begin{aligned}
Var\{\sigma_R^2(x)\} =& \frac{2}{T^2}\int_0^T\int_0^t \mathbb{E}\{[\sigma^2(x_t,\gamma(t)) - \mathbb{E}\{\sigma^2(x_t,\gamma(t))\}][\sigma^2(x_s,\gamma(s)) - \mathbb{E}\{\sigma^2(x_s,\gamma(s))\}]\}dsdt \\
=& \frac{2}{T^2}\int_0^T\int_0^t \left[\mathbb{E}\{\sigma^2(x_t,\gamma(t))\sigma^2(x_s,\gamma(s))\} - \mathbb{E}\{\sigma^2(x_t,\gamma(t))\}\mathbb{E}\{\sigma^2(x_s,\gamma(s))\}\right] dsdt\ .
\end{aligned}
$$

We would like to stress that, in this representation, the integration set is such that the inequality $s \leq t$ holds true. Using the property of conditional expectation and Proposition 4.3.1, we have

$$
Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T\int_0^t \left[\mathbb{E}\{\sigma^2(x_s,\gamma(s))\mathbb{E}\{\sigma^2(x_t,\gamma(t))|\mathscr{F}_s\}\} - \left(e^{tQ}\sigma^2(x,\gamma)\right)\left(e^{sQ}\sigma^2(x,\gamma)\right)\right] dsdt\ .
$$

The process $(x_t,\gamma(t))_t$ is Markovian, then using Remark 4.3.2 the conditional expected value in the integrand, can be expressed as

$$
\mathbb{E}\{\sigma^2(x_t,\gamma(t))|\mathscr{F}_s\} = e^{(t-s)Q}\sigma^2(x_s,\gamma(s)) =: g(x_s,\gamma(s))\ . \qquad (4.31)
$$

Thus, the variance becomes

$$
Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T\int_0^t \left[\mathbb{E}\{\sigma^2(x_s,\gamma(s))g(x_s,\gamma(s))\} - \left(e^{tQ}\sigma^2(x,\gamma)\right)\left(e^{sQ}\sigma^2(x,\gamma)\right)\right] dsdt\ .
$$

Solving the expectation on the r.h.s we obtain

$$
Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T\int_0^t \left[e^{sQ}\left(\sigma^2(x,\gamma)g(x,\gamma)\right) - \left(e^{tQ}\sigma^2(x,\gamma)\right)\left(e^{sQ}\sigma^2(x,\gamma)\right)\right] dsdt\ .
$$

We notice that function g evaluated in (x,γ) is simply given by

$$
g(x,\gamma) = e^{(t-s)Q}\sigma^2(x,\gamma)\ , \qquad (4.32)
$$

then substituting in the previous formula the variance finally becomes

$$Var\{\sigma_R^2(x)\} = \frac{2}{T^2}\int_0^T\int_0^t \left\{ e^{sQ}\left[\sigma^2(x,\gamma)e^{(t-s)Q}\sigma^2(x,\gamma)\right] - \left[e^{tQ}\sigma^2(x,\gamma)\right]\left[e^{sQ}\sigma^2(x,\gamma)\right]\right\} dsdt \ .$$

□

4.3.3 Numerical Evaluation of Variance and Volatility Swaps with Semi-Markov Volatility

When we attempt to evaluate the price of a variance or a volatility swaps we have to deal with numerical problems. The family of exponential operators $(e^{tQ})_t$ involved in Theorems 4.3.3 and 4.3.4 for the semi-Markov stochastic volatility model is usually difficult to evaluate from a numerical point of view. To solve this problem, we first look to the following identity

$$e^{tQ}f(\cdot) = \sum_{n=0}^{\infty}\frac{(tQ)^n}{n!}f(\cdot) \ , \tag{4.33}$$

for any function $f \in Domain(Q)$. This identity allows us to obtain the operator $(e^{tQ})_t$ at any order of approximation. For example, for $n = 1$, we obtain

$$e^{tQ}f(\cdot) \approx (I + tQ)f(\cdot) \ , \tag{4.34}$$

where I is an identity operator. At this order of approximation we allow semi-Markov process to make at most one transition during the life time of contract. If we think to semi-Markov process as a macroeconomic factor this can be plausible. However we can always evaluate the error in this approximation using the subsequent orders. Using the first order approximation the variance swap price becomes

$$\begin{aligned} P_{var}(x) \approx & e^{-rT}\left\{\frac{1}{T}\int_0^T (I+tQ)\sigma^2(x,\gamma)dt - K_{var}\right\} \\ = & e^{-rT}\left\{\sigma^2(x,\gamma) + \frac{T}{2}Q\sigma^2(x,\gamma) - K_{var}\right\} . \end{aligned} \tag{4.35}$$

Using the same approximation the volatility swap price can be expressed as

$$\begin{aligned} P_{vol}(x) \approx & e^{-rT}\left\{\sqrt{\frac{1}{T}\int_0^T (I+tQ)\sigma^2(x,\gamma)dt} - \frac{Var\{\sigma_R^2(x)\}}{8\left(\frac{1}{T}\int_0^T (I+tQ)\sigma^2(x,\gamma)dt\right)^{3/2}} - K_{vol}\right\} \\ = & e^{-rT}\left\{\sqrt{\sigma^2(x,\gamma) + \frac{T}{2}Q\sigma^2(x,\gamma)} - \frac{Var\{\sigma_R^2(x)\}}{8\left(\sigma^2(x,\gamma) + \frac{T}{2}Q\sigma^2(x,\gamma)\right)^{3/2}} - K_{vol}\right\} . \end{aligned}$$

Here, the variance of realized volatility is given by

$$\begin{aligned} Var\{\sigma_R^2(x)\} \approx & \frac{2}{T^2}\int_0^T\int_0^t \{(I+sQ)[\sigma^2(x,\gamma)(I+(t-s)Q)\sigma^2(x,\gamma)] \\ & -[(I+tQ)\sigma^2(x,\gamma)][(I+sQ)\sigma^2(x,\gamma)]\}\, dsdt \ . \end{aligned} \tag{4.36}$$

Solving the product and keeping only the terms up to the first order in Q, we obtain

$$\begin{aligned} Var\{\sigma_R^2(x)\} \approx & \frac{2}{T^2}\int_0^T\int_0^t \{sQ\sigma^4(x,\gamma)-2\sigma^2(x,\gamma)sQ\sigma^2(x,\gamma)\}dsdt \\ = & \frac{T}{3}\left\{Q\sigma^4(x,\gamma)-2\sigma^2(x,\gamma)Q\sigma^2(x,\gamma)\right\} \ . \end{aligned} \tag{4.37}$$

Finally, to first order of approximation in Q the volatility swap price becomes

$$P_{vol}(x) \approx e^{-rT}\left\{\sqrt{\sigma^2(x,\gamma)+\frac{T}{2}Q\sigma^2(x,\gamma)} - \frac{T[Q\sigma^4(x,\gamma)-2\sigma^2(x,\gamma)Q\sigma^2(x,\gamma)]}{24\left(\sigma^2(x,\gamma)+\frac{T}{2}Q\sigma^2(x,\gamma)\right)^{3/2}} - K_{vol}\right\}.$$

4.4 Covariance and Correlation Swaps for a Two Risky Assets in Financial Markets with Semi-Markov Stochastic Volatilities

Let's consider now a market model with two risky assets and one risk free bond. Let's assume that the risky assets are satisfying the following stochastic differential equations

$$\begin{cases} dS_t^{(1)} = S_t^{(1)}(\mu_t^{(1)}dt+\sigma^{(1)}(x_t,\gamma(t))dw_t^{(1)}) \\ dS_t^{(2)} = S_t^{(2)}(\mu_t^{(2)}dt+\sigma^{(2)}(x_t,\gamma(t))dw_t^{(2)}) \end{cases} , \tag{4.38}$$

where $\mu^{(1)}, \mu^{(2)}$ are deterministic functions of time, $(w_t^{(1)})_t$ and $(w_t^{(2)})_t$ are standard Wiener processes with quadratic covariance given by

$$d[w_t^{(1)}, w_t^{(2)}] = \rho_t dt \ . \tag{4.39}$$

Here ρ_t is a deterministic function and $(w_t^{(1)})_t, (w_t^{(2)})_t$ are independent of (x,γ).

In this model it is worth to study the covariance and the correlation swaps between the two risky assets.

4.4.1 Pricing of Covariance Swaps

A covariance swap is a covariance forward contract on the underlying assets $S^{(1)}$ and $S^{(2)}$ which payoff at maturity is equal to

$$N(Cov_R(S^{(1)},S^{(2)})-K_{cov}) \ , \tag{4.40}$$

where K_{cov} is a strike reference value, N is the notional amount and $Cov_R(S^{(1)}, S^{(2)})$ is the realized covariance of the two assets $S^{(1)}$ and $S^{(2)}$ given by

$$Cov_R(S^{(1)}, S^{(2)}) = \frac{1}{T}[\ln S_T^{(1)}, \ln S_T^{(2)}] = \frac{1}{T}\int_0^T \rho_t \sigma^{(1)}(x_t, \gamma(t))\sigma^{(2)}(x_t, \gamma(t))dt \ . \tag{4.41}$$

The price of the covariance swap is the expected present value of the payoff in the risk neutral world

$$P_{cov}(x) = \mathbb{E}\{e^{-rT}(Cov_R(S^{(1)}, S^{(2)}) - K_{cov})\} \ , \tag{4.42}$$

here we set $N = 1$. The next result provides an explicit representation of the covariance swap price.

Theorem 4.4.1. *The value of a covariance swap for semi-Markov stochastic volatility is*

$$P_{cov}(x) = e^{-rT}\left\{\frac{1}{T}\int_0^T \rho_t e^{tQ}[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)]dt - K_{cov}\right\}, \tag{4.43}$$

where Q is the generator of $(x_t, \gamma(t))_t$, that is

$$Qf(x,t) = \frac{df}{dt}(x,t) + \frac{g_x(t)}{\overline{G}_x(t)}\int_X P(x,dy)[f(y,0) - f(x,t)] \ . \tag{4.44}$$

Proof. To evaluate the price of covariance swap we need to know

$$\mathbb{E}\{Cov_R(S^{(1)}, S^{(2)})\} = \frac{1}{T}\int_0^T \rho_t \mathbb{E}\{\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\}dt \ . \tag{4.45}$$

It remains to prove that

$$\mathbb{E}\{\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\} = e^{tQ}[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)] \ . \tag{4.46}$$

By applying Ito's lemma we have

$$\begin{aligned} d(\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))) =& \sigma^{(1)}(x_t,\gamma(t))d\sigma^{(2)}(x_t,\gamma(t)) + \sigma^{(2)}(x_t,\gamma(t))d\sigma^{(1)}(x_t,\gamma(t)) \\ &+ d\langle\sigma^{(1)}(x_\cdot,\gamma(\cdot)), \sigma^{(2)}(x_\cdot,\gamma(\cdot))\rangle_t \ . \end{aligned} \tag{4.47}$$

Using Proposition 3.2.3 we obtain

$$\begin{aligned} d\langle\sigma^{(1)}(x_\cdot,\gamma(\cdot)), \sigma^{(2)}(x_\cdot,\gamma(\cdot))\rangle_t =& Q(\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t)))dt \\ &- [\sigma^{(1)}(x_t,\gamma(t))Q\sigma^{(2)}(x_t,\gamma(t)) + \sigma^{(2)}(x_t,\gamma(t))Q\sigma^{(1)}(x_t,\gamma(t))]dt \ . \end{aligned} \tag{4.48}$$

Furthermore we have

$$d\sigma^{(i)}(x_t,\gamma(t)) = Q\sigma^{(i)}(x_t,\gamma(t))dt + dm^{\sigma^{(i)}} \qquad i = 1, 2 \ . \tag{4.49}$$

Substituting (4.48) and (4.49) in equation (4.47) we get

$$\begin{aligned} d(\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))) =& Q(\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t)))dt \\ &+\sigma^{(1)}(x_t,\gamma(t))dm^{\sigma^{(2)}} + \sigma^{(2)}(x_t,\gamma(t))dm^{\sigma^{(1)}} . \end{aligned} \tag{4.50}$$

Taking the expectation on both side we can rewrite the above equation as

$$\begin{aligned} \mathbb{E}\{\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\} =& \sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma) \\ &+ \int_0^t Q\mathbb{E}\{\sigma^{(1)}(x_s,\gamma(s))\sigma^{(2)}(x_s,\gamma(s))\}dt . \end{aligned} \tag{4.51}$$

Solving this differential equation we obtain

$$\mathbb{E}\{\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\} = e^{tQ}[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)] , \tag{4.52}$$

this conclude the proof. □

4.4.2 Pricing of Correlation Swaps

A correlation swap is a forward contract on the correlation between the underlying assets S^1 and S^2 which payoff at maturity is equal to

$$N(Corr_R(S^1,S^2) - K_{corr}) , \tag{4.53}$$

where K_{corr} is a strike reference level, N is the notional amount and $Corr_R(S^1,S^2)$ is the realized correlation defined by

$$Corr_R(S^1,S^2) = \frac{Cov_R(S^1,S^2)}{\sqrt{\sigma_R^{(1)^2}(x)}\sqrt{\sigma_R^{(2)^2}(x)}} , \tag{4.54}$$

here the realized variance is given by

$$\sigma_R^{(i)^2}(x) = \frac{1}{T}\int_0^T (\sigma^{(i)}(x_t,\gamma(t)))^2 dt \qquad i = 1,2 . \tag{4.55}$$

The price of the correlation swap is the expected present value of the payoff in the risk neutral world, that is

$$P_{corr}(x) = \mathbb{E}\{e^{-rT}(Corr_R(S^1,S^2) - K_{corr})\} , \tag{4.56}$$

where we set $N = 1$ for simplicity. Unfortunately the expected value of $Corr_R(S^1,S^2)$ is not known analytically. Thus, in order to obtain an explicit formula for the correlation swap price, we have to make some approximation.

4.4.3 Correlation Swaps Made Simple

First of all, let us introduce the following notations

$$\begin{aligned} X =& Cov_R(S^1, S^2) \\ Y =& \sigma_R^{(1)^2}(x) \\ Z =& \sigma_R^{(2)^2}(x) \ , \end{aligned} \tag{4.57}$$

and with the subscript 0 we will denote the expected value of the above random variables. Following the approach frequently used for the volatility swap, we would like to approximate the square root of Y and Z to the first order as follows

$$\begin{aligned} \sqrt{Y} \approx& \sqrt{Y_0} + \frac{Y - Y_0}{2\sqrt{Y_0}} \\ \sqrt{Z} \approx& \sqrt{Z_0} + \frac{Z - Z_0}{2\sqrt{Z_0}} \ . \end{aligned} \tag{4.58}$$

The realized correlation can now be approximated by

$$Corr_R(S^1, S^2) \approx \frac{X}{\left(\sqrt{Y_0} + \frac{Y-Y_0}{2\sqrt{Y_0}}\right)\left(\sqrt{Z_0} + \frac{Z-Z_0}{2\sqrt{Z_0}}\right)} = \frac{\frac{X}{\sqrt{Y_0}\sqrt{Z_0}}}{\left(1 + \frac{Y-Y_0}{2Y_0}\right)\left(1 + \frac{Z-Z_0}{2Z_0}\right)} \ . \tag{4.59}$$

Solving the product in the denominator of the last term on the r.h.s and keeping only the terms up to the first order in the increment, we have

$$Corr_R(S^1, S^2) \approx \frac{\frac{X}{\sqrt{Y_0}\sqrt{Z_0}}}{1 + \left(\frac{Y-Y_0}{2Y_0} + \frac{Z-Z_0}{2Z_0}\right)} \approx \frac{X}{\sqrt{Y_0}\sqrt{Z_0}}\left[1 - \left(\frac{Y - Y_0}{2Y_0} + \frac{Z - Z_0}{2Z_0}\right)\right]. \tag{4.60}$$

In what follows, we will consider only the zeroth order of approximation. The first order correction will be discuss in the next section.

Here, we are going to approximate the realized correlation as

$$Corr_R(S^1, S^2) \approx \frac{X}{\sqrt{Y_0}\sqrt{Z_0}} \ , \tag{4.61}$$

Substituting X, Y and Z we obtain

$$Corr_R(S^1, S^2) \approx \frac{1}{\sqrt{\mathbb{E}\{\sigma_R^{(1)^2}(x)\}}\sqrt{\mathbb{E}\{\sigma_{2R}^{(2)^2}(x)\}}} \frac{1}{T}\int_0^T \rho_t \sigma^{(1)}(x_t, \gamma(t))\sigma^{(2)}(x_t, \gamma(t))dt \ , \tag{4.62}$$

where (*cf.* Teorem 4.3.3), we have

$$\mathbb{E}\left\{\sigma_R^{(i)^2}(x)\right\} = \mathbb{E}\left\{\frac{1}{T}\int_0^T (\sigma^{(i)}(x_t, \gamma(t)))^2 dt\right\} = \frac{1}{T}\int_0^T e^{tQ}(\sigma^{(i)}(x, \gamma))^2 dt \ , \tag{4.63}$$

for $i = 1, 2$. In order to price a correlation swap we have to be able to evaluate the expectation of both side of Eq. (4.62), the expectation of the r.h.s becomes

$$\mathbb{E}\left\{\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\right\} = e^{tQ}\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)\ . \tag{4.64}$$

We can summarize the previous result in the following statement.

Theorem 4.4.2. *The value of a correlation swap for semi-Markov stochastic volatility is*

$$P_{corr}(x) \approx e^{-rT}\left\{\frac{\int_0^T \rho_t e^{tQ}[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)]dt}{\sqrt{\int_0^T e^{tQ}(\sigma^{(1)}(x,\gamma))^2dt}\sqrt{\int_0^T e^{tQ}(\sigma^{(2)}(x,\gamma))^2dt}} - K_{corr}\right\}, \tag{4.65}$$

where Q is the generator of $(x_t, \gamma(t))_t$, that is

$$Qf(x,t) = \frac{df}{dt}(x,t) + \frac{g_x(t)}{\bar{G}_x(t)}\int_X P(x,dy)[f(y,0) - f(x,t)]. \tag{4.66}$$

4.4.4 Correlation Swaps: First Order Correction

We would like to obtain a first-order approximation for the realized correlation between two risky assets

$$Corr_R(S^1,S^2) = \frac{Cov_R(S^1,S^2)}{\sqrt{\sigma_R^{(1)^2}(x)}\sqrt{\sigma_R^{(2)^2}(x)}}\ . \tag{4.67}$$

In section 4.4.3 we have already obtained the following approximated expression

$$Corr_R(S^1,S^2) \approx \frac{\frac{X}{\sqrt{Y_0}\sqrt{Z_0}}}{1 + \left(\frac{Y-Y_0}{2Y_0} + \frac{Z-Z_0}{2Z_0}\right)} \approx \frac{X}{\sqrt{Y_0}\sqrt{Z_0}}\left[1 - \left(\frac{Y-Y_0}{2Y_0} + \frac{Z-Z_0}{2Z_0}\right)\right]. \tag{4.68}$$

where

$$\begin{aligned} X =& Cov_R(S^1,S^2) \\ Y =& \sigma_R^{(1)^2}(x) \\ Z =& \sigma_R^{(2)^2}(x)\ , \end{aligned} \tag{4.69}$$

and with the pedix 0 we have denoted the expected values. We have already evaluated the expectation of the zeroth order approximation, now we would like to evaluate the first-order.

Substituting X, Y and Z in Eq. (4.68) we obtain

$$\begin{aligned} Corr_R(S^1, S^2) \approx & \frac{1}{\sqrt{\mathbb{E}\{\sigma_R^{(1)^2}(x)\}}\sqrt{\mathbb{E}\{\sigma_{2R}^{(2)^2}(x)\}}} \frac{1}{T}\int_0^T \rho_t \sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))dt \\ & - \frac{1}{2T^2(\mathbb{E}\{\sigma_R^{(1)^2}(x)\})^{3/2}(\mathbb{E}\{\sigma_R^{(2)^2}(x)\})^{3/2}} \int_0^T \rho_t \sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))dt \\ & \times \Bigg\{ \mathbb{E}\{\sigma_R^{(2)^2}(x)\} \int_0^T [(\sigma^{(1)}(x_s,\gamma(s)))^2 - \mathbb{E}\{(\sigma^{(1)}(x_s,\gamma(s)))^2\}]ds \\ & + \mathbb{E}\{\sigma_R^{(1)^2}(x)\} \int_0^T [(\sigma^{(2)}(x_u,\gamma(u)))^2 - \mathbb{E}\{(\sigma^{(2)}(x_u,\gamma(u)))^2\}]du \Bigg\}, \end{aligned} \tag{4.70}$$

where

$$\mathbb{E}\left\{\sigma^2_{(i)R}(x)\right\} = \mathbb{E}\left\{\frac{1}{T}\int_0^T (\sigma^{(i)}(x_t,\gamma(t)))^2 dt\right\} = \frac{1}{T}\int_0^T e^{tQ}(\sigma^{(i)}(x,\gamma))^2 dt \ , \tag{4.71}$$

for $i = 1, 2$. We have to evaluate the expectation of the r.h.s of equation (4.70). We already calculated the expectation of the first term, which is the zero order approximation for the realized correlation. Then we will focus now on the other terms. First of all, let us rewrite them as follows

$$\begin{aligned} \int_0^T \int_0^T \rho_t \sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t)) \Big(& \mathbb{E}\{\sigma_R^{(2)^2}(x)\}[(\sigma^{(1)}(x_s,\gamma(s)))^2 - \mathbb{E}\{(\sigma^{(1)}(x_s,\gamma(s)))^2\}] \\ & + \mathbb{E}\{\sigma_R^{(1)^2}(x)\}[(\sigma^{(2)}(x_s,\gamma(s)))^2 - \mathbb{E}\{(\sigma^{(2)}(x_s,\gamma(s)))^2\}]\Big) dsdt \ , \end{aligned} \tag{4.72}$$

we have four different contributions in the integrals, the expectation of the terms

$$\int_0^T \int_0^T \rho_t \sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\mathbb{E}\{\sigma^{(i)^2}(x_s,\gamma(s))\}\mathbb{E}\{\sigma^{(-i)^2}(x_s,\gamma(s))\}dsdt \ , \tag{4.73}$$

for $i = 1, 2$, can be evaluate using Theorem 4.4.1. Then, in order to evaluate the expectation of the approximated realized correlation, it only remains to calculate

$$\mathbb{E}\left\{\int_0^T \int_0^T \rho_t \sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\sigma^{(i)^2}(x_s,\gamma(s))dsdt\right\} \qquad i = 1, 2 \ , \tag{4.74}$$

To this end, let's first divide the range of integration in two intervals as follows

$$\begin{aligned} \mathbb{E}\Bigg\{ & \int_0^T \int_0^t \rho_t \sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\sigma^{(i)^2}(x_s,\gamma(s))dsdt \\ & + \int_0^T \int_t^T \rho_t \sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\sigma^{(i)^2}(x_s,\gamma(s))dsdt \Bigg\}, \end{aligned} \tag{4.75}$$

for $i = 1, 2$. We notice that the first integral set is such that $t > s$ while the second has $t < s$. We can now use the property of conditional expectation to obtain

$$\begin{aligned}\mathbb{E}\Bigg\{&\int_0^T\int_0^t \rho_t \mathbb{E}\{\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))|\mathscr{F}_s\}\sigma^{(i)^2}(x_s,\gamma(s))dsdt \\ &+\int_0^T\int_t^T \rho_t\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))\mathbb{E}\{\sigma^{(i)^2}(x_s,\gamma(s))|\mathscr{F}_t\}dsdt\Bigg\}.\end{aligned} \tag{4.76}$$

We notice that $(x_t, \gamma(t))_t$ is a Markov process, then using the Markov property, we can express the conditional expectations as

$$\mathbb{E}\{\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))|\mathscr{F}_s\} = e^{(t-s)Q}\sigma^{(1)}(x_s,\gamma(s))\sigma^{(2)}(x_s,\gamma(s)) =: h(x_s,\gamma(s)) \ ,$$

for $t > s$, and

$$\mathbb{E}\{\sigma^{(i)^2}(x_s,\gamma(s))|\mathscr{F}_t\} = e^{(s-t)Q}\sigma^{(i)^2}(x_t,\gamma(t)) =: g^{(i)}(x_t,\gamma(t)) \ ,$$

for $s > t$. Therefore, the first term of Eq. (4.76) can be expressed as

$$\mathbb{E}\left\{\int_0^T\int_0^t \rho_t h(x_s,\gamma(s))\sigma^{(i)^2}(x_s,\gamma(s))dsdt\right\} = \int_0^T\int_0^t \rho_t e^{sQ}[h(x,\gamma)\sigma^{(i)^2}(x,\gamma)]dsdt \ , \tag{4.77}$$

while the second as

$$\begin{aligned}&\mathbb{E}\left\{\int_0^T\int_t^T \rho_t\sigma^{(1)}(x_t,\gamma(t))\sigma^{(2)}(x_t,\gamma(t))g^{(i)}(x_t,\gamma(t))dsdt\right\} \\ &= \int_0^T\int_t^T \rho_t e^{tQ}[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)g^{(i)}(x,\gamma)]dsdt \ .\end{aligned} \tag{4.78}$$

Now, we can evaluate the functions h and g at x obtaining

$$h(x,\gamma) = e^{(t-s)Q}[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)] \tag{4.79}$$

and

$$g^{(i)}(x,\gamma) = e^{(s-t)Q}[\sigma^{(i)^2}(x,\gamma)]. \tag{4.80}$$

We can summarize the previous result in the following statement which gives the correlation swap price up to the first-order of approximation.

Theorem 4.4.3. *The value of the correlation swap for a semi-Markov volatility is*

$$P_{corr}(x) = e^{-rT}\left(\mathbb{E}\{Corr_R(S^1,S^2)\} - K_{corr}\right) \ , \tag{4.81}$$

where the realized correlation can be approximated by

$$\begin{aligned}
\mathbb{E}\{Corr_R(S^1,S^2)\} \approx & \frac{2\int_0^T \rho_t e^{tQ}\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)dt}{\sqrt{\int_0^T e^{tQ}(\sigma^{(1)}(x,\gamma))^2dt}\sqrt{\int_0^T e^{tQ}(\sigma^{(2)}(x,\gamma))^2dt}} \\
& - \frac{\int_0^T \rho_t(\int_0^t e^{sQ}\{e^{tQ}[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)]\sigma^{(1)^2}(x,\gamma)\}ds + \int_t^T e^{tQ}\{\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)e^{uQ}[\sigma^{(1)^2}(x,\gamma)]\}du)dt}{2\left(\int_0^T e^{tQ}(\sigma^{(1)}(x,\gamma))^2dt\right)^{3/2}\left(\int_0^T e^{tQ}(\sigma^{(2)}(x,\gamma))^2dt\right)^{1/2}} \\
& - \frac{\int_0^T \rho_t(\int_0^t e^{sQ}\{e^{tQ}[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)]\sigma^{(2)^2}(x,\gamma)\}ds + \int_t^T e^{tQ}\{\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)e^{uQ}[\sigma^{(2)^2}(x,\gamma)]\}du)dt}{2\left(\int_0^T e^{tQ}(\sigma^{(1)}(x,\gamma))^2dt\right)^{1/2}\left(\int_0^T e^{tQ}(\sigma^{(2)}(x,\gamma))^2dt\right)^{3/2}} \ ,
\end{aligned} \tag{4.82}$$

here Q is the generator of the Markov process $(x_t,\gamma(t))_{t\in\mathbb{R}_+}$ given by

$$Qf(x,t) = \frac{df}{dt}(x,t) + \frac{g_x(t)}{\overline{G}_x(t)}\int_X P(x,dy)[f(y,0) - f(x,t)] \ . \tag{4.83}$$

4.5 Numerical Evaluation of Covariance and Correlation Swaps with Semi-Markov Stochastic Volatility

In order to obtain a more handy expression for the price of covariance and correlation swaps to use in the application, we will introduce here an approximation for the family of operators $(e^{tQ})_{t\in\mathbb{R}_+}$. Following the approach used for the variance and volatility case, we are going to approximate the operators at the first order in Q as

$$e^{tQ}f(\cdot) \approx (I + tQ)f(\cdot) \ . \tag{4.84}$$

Using this approximation the covariance swap price becomes

$$\begin{aligned}
P_{cov}(x) \approx & e^{-rT}\left\{\frac{1}{T}\int_0^T \rho_t(I+tQ)[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)]dt - K_{cov}\right\} \\
= & e^{-rT}\left\{\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)\int_0^T \rho_t dt + Q[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)]\int_0^T t\rho_t dt - K_{cov}\right\} .
\end{aligned} \tag{4.85}$$

The same approximation allow us to express the zeroth order approximation of correlation swap price as

$$\begin{aligned} P_{corr}(x) \approx & e^{-rT}\left\{\frac{\int_0^T \rho_t(I+tQ)[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)]dt}{\sqrt{\int_0^T (I+tQ)(\sigma^{(1)}(x,\gamma))^2 dt}\sqrt{\int_0^T (I+tQ)(\sigma^{(2)}(x,\gamma))^2 dt}} - K_{corr}\right\} \\ = & e^{-rT}\left\{\frac{\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)\int_0^T \rho_t dt + Q[\sigma^{(1)}(x,\gamma)\sigma^{(2)}(x,\gamma)]\int_0^T t\rho_t dt}{\sqrt{(\sigma^{(1)}(x,\gamma))^2 + \frac{T}{2}Q(\sigma^{(1)}(x,\gamma))^2}\sqrt{(\sigma^{(2)}(x,\gamma))^2 + \frac{T}{2}Q(\sigma^{(2)}(x,\gamma))^2}} - K_{corr}\right\}. \end{aligned} \tag{4.86}$$

4.6 Conclusion

A semi-Markov modulated stochastic volatility model has been defined, in this model variance, volatility, covariance and correlation swap have been studied. In particular, second order approximation for volatility swap price have been explicitly evaluated and a numerical evaluation of both variance and volatility swap has been discussed. The covariance swap price in a two risky assets market with semi-Markov volatility has been obtained and an approximated expression for the correlation swap price has been derived, a numerical evaluation of them has been discussed.

Part II

Multivariate Semi-Markov Models

Chapter 5

Bivariate Markov Chains

In this chapter we will introduce the bivariate Markov chains.

Multivariate Markov chain for stock markets has been studied by Maskawa [70], in this work a two state space is considered and the system is studied as a whole. Multivariate Markov chain has been already studied by Ching *et al.* [24, 25] using a mixing distribution approach; we refer to the book of Ching and Ng [23] for a complete review on their approach. Multivariate Markov process with copula has been investigated by Bielecki *et al.* [7]. Bielecki *et al.* [6] analyzed the dependence structure in a multivariate process whose components are Markov processes.

In the first section we study multidimensional matrices as a general framework, to study this topic we will follow the works of Manca [68, 69] for notation and we refer to them for details and proofs. The second section is devoted to bivariate Markov chains. We begin with the study of the system as a whole and then we discuss some particular dependence structures between the components.

5.1 Multidimensional Matrices: Definition and Properties

In this section we will introduce the multidimensional matrices.

Let us take $\mathbb{R}$ as the field, we define matrices of dimensions 0 and 1 as follows.

Definition 5.1.1. ***Matrices of dimension*** 0 ***and*** 1

- *The elements of* $\mathbb{R}$ *are matrices of dimension* 0*.*
- *A matrix* A *of dimension* 1 *is a matrix whose elements are matrices of dimension* 0*.*

We denote the matrix of order 1 by

$$A_{[r_1][c_1]} \ .$$

Here, r_1 is the number of rows and c_1 is the number of columns. Let $\mathcal{M}^1_{[r_1][c_1]}$ be the set of one dimensional matrices with r_1 rows and c_1 columns.

For example, a matrix A of dimension 1 with 2 rows and 2 columns, $A \in \mathcal{M}^1_{[2][2]}$, can be expressed as

$$A_{[2][2]} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} , \tag{5.1}$$

where $a, b, c, d \in \mathbb{R}$ i.e. they are matrix of dimension 0.

By iteration we can define a matrix of higher dimension.

Definition 5.1.2. ***Matrices of dimension*** n

A matrix of dimension n is a matrix whose elements are matrices of dimension $n-1$. We denote a matrix of dimension n by

$$A_{\underline{r}\ \underline{c}} \qquad \textit{where}\ \ \underline{r} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} \ \textit{and} \ \ \underline{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} . \tag{5.2}$$

Here $\underline{r} = (r_1, \dots, r_n)$ and $\underline{c} = (c_1, \dots, c_n)$ represent the numbers of rows and columns at any dimension level, respectively. We denote by $\mathcal{M}^n_{\underline{r}\ \underline{c}}$ the set of all n-dimensional matrices with order structure of rows $\underline{r}$ and columns $\underline{c}$.

For example, we define a matrix of dimension 2 as a matrix with r_2 rows and c_2 columns, whose elements are matrices of dimension 1, each one of them with r_1 rows and c_1 columns. We denote the two dimensional matrices of this kind as follows

$$A_{\underline{r}\ \underline{c}} \qquad \text{where}\ \ \underline{r} = \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \ \text{and} \ \ \underline{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} . \tag{5.3}$$

Then, a bidimensional matrix with vector of rows $(r_1, 3)$ and vector of columns $(c_1, 3)$ can be represented as

$$A_{\left[\begin{smallmatrix} r_1 \\ 3 \end{smallmatrix}\right]\left[\begin{smallmatrix} c_1 \\ 3 \end{smallmatrix}\right]} = \begin{pmatrix} \alpha_{[r_1][c_1]} & \beta_{[r_1][c_1]} & \gamma_{[r_1][c_1]} \\ \delta_{[r_1][c_1]} & \epsilon_{[r_1][c_1]} & \zeta_{[r_1][c_1]} \\ \eta_{[r_1][c_1]} & \theta_{[r_1][c_1]} & \lambda_{[r_1][c_1]} \end{pmatrix} \qquad \text{where} \ \ \alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta, \theta, \lambda \in \mathcal{M}^1_{r_1 c_1} . \tag{5.4}$$

A generic element (matrix of dimension 0) of $A_{\left[\begin{smallmatrix} r_1 \\ 3 \end{smallmatrix}\right]\left[\begin{smallmatrix} c_1 \\ 3 \end{smallmatrix}\right]}$ can be expressed as

$$A_{\left(\begin{smallmatrix} i_1 \\ i_2 \end{smallmatrix}\right)\left(\begin{smallmatrix} j_1 \\ j_2 \end{smallmatrix}\right)} \qquad \text{where} \ \ i_1 = 1, \dots, r_1 \ ; \ j_1 = 1, \dots, c_1 \ ; \ i_2 = 1, \dots, 3 \ ; \ j_2 = 1, \dots, 3. \tag{5.5}$$

Denote by $\mathcal{M}^2_{\underline{r}\ \underline{c}}$ the set of bidimensional matrices with fixed order structure of rows $\underline{r}$ and columns $\underline{c}$. Therefore, with $\mathcal{M}^2_{\left[\begin{smallmatrix} r_1 \\ 3 \end{smallmatrix}\right]\left[\begin{smallmatrix} c_1 \\ 3 \end{smallmatrix}\right]}$ we will denote the set of all matrices 3×3 whose elements are one dimensional matrices $r_1 \times c_1$.

We define the product between two matrices as a natural generalization of the standard raws by columns product.

Definition 5.1.3. ***Product of Matrices***

- *Let $A, B \in \mathcal{M}^0_{[1][1]}$, we define the product of A and B as the standard multiplication between real number.*

- *Let $A \in \mathcal{M}^1_{[r][h]}$ and $B \in \mathcal{M}^1_{[h][c]}$, we define the product of A and B as the application*

$$PM_1 : \mathcal{M}^1_{[r][h]} \times \mathcal{M}^1_{[h][c]} \longrightarrow \mathcal{M}^1_{[r][c]} \tag{5.6}$$

  given by

$$C_{(i)(j)} = PM_1(A,B)_{(i)(j)} =: (A *_1 B)_{(i)(j)} = \sum_{l=1}^{h} A_{(i)(l)} B_{(l)(j)} \ ,$$

  for $i = 1, \ldots, r$ and $j = 1, \ldots, c$.

- *Let $A \in \mathcal{M}^n_{\underline{r}\ \underline{h}}$ and $B \in \mathcal{M}^n_{\underline{h}\ \underline{c}}$ we define the product of A and B as the application*

$$PM_n : \mathcal{M}^n_{\underline{r}\ \underline{h}} \times \mathcal{M}^n_{\underline{h}\ \underline{c}} \longrightarrow \mathcal{M}^n_{\underline{r}\ \underline{c}} \tag{5.7}$$

  given for $i = 1, \ldots, r_n$ and $j = 1, \ldots, c_n$ by

$$C_{\binom{\underline{r}^{n-1}}{i}\binom{\underline{c}^{n-1}}{j}} = PM_n(A,B)_{\binom{\underline{r}^{n-1}}{i}\binom{\underline{c}^{n-1}}{j}} =: (A *_n B)_{\binom{\underline{r}^{n-1}}{i}\binom{\underline{c}^{n-1}}{j}} = \sum_{l=1}^{h_n} A_{\binom{\underline{r}^{n-1}}{i}\binom{\underline{h}^{n-1}}{l}} *_{n-1} B_{\binom{\underline{h}^{n-1}}{l}\binom{\underline{c}^{n-1}}{j}} \ ,$$

  here $\underline{r}^{n-1}$, $\underline{h}^{n-1}$ and $\underline{c}^{n-1}$ denote the $n-1$ dimensional vectors whose components are the first n-1 components of vectors $\underline{r}$, $\underline{h}$ and $\underline{c}$, respectively.

Remark 5.1.4. *The product of multidimensional matrices is defined recursively from the standard rows by columns product of one dimensional matrices. The properties of the standard raws by columns product can be directly generalized to this multidimensional case.*

5.2 Bivariate Markov Chains

In this section we will introduce the Multivariate Markov Chains.

Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ be a complete filtered probability space and $E = \{1, \ldots, d\}$ be a given finite set.

Let us consider 2 sequences of random variables with values in E, we denote the generic random sequence by $X^\alpha = (X^\alpha_n)_{n\in\mathbb{N}}$, for $\alpha = 1, 2$, and by

$$\mathbf{X} = (\mathbf{X}_n)_{n\in\mathbb{N}} = (X^1_n, X^2_n)_{n\in\mathbb{N}} \ ,$$

the two-dimensional random vector collecting the sequences.

Let $\mathbf{v} = \binom{v^1}{v^2}$ denote a vector in E^2.

Definition 5.2.1. ***Bivariate Markov Chain***

The sequence $(\mathbf{X})_{n\in\mathbb{N}}$ is a bivariate Markov chain if

$$\mathbb{P}\{\mathbf{X}_{n+1} = \mathbf{j} \mid \mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_n = \mathbf{i}_n\} = \mathbb{P}\{\mathbf{X}_{n+1} = \mathbf{j} \mid \mathbf{X}_n = \mathbf{i}_n\} \tag{5.8}$$

for every $n \in \mathbb{N}$, $\mathbf{j} \in E^2$, $\alpha = 1, 2$ and every sequence $\mathbf{i}_0, \ldots, \mathbf{i}_n$ in E^2 such that

$$\mathbb{P}\{\mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_n = \mathbf{i}_n\} > 0 \ .$$

Remark 5.2.2. *The Markovian property is preserved by the system, and each component, with respect to the filtration of the whole system.*

Definition 5.2.3. ***Homogeneous Chain***

The bivariate Markov chain is homogenous if

$$\mathbb{P}\{\mathbf{X}_{n+1} = \mathbf{j} \mid \mathbf{X}_n = \mathbf{i}\} = \mathbb{P}\{\mathbf{X}_1 = \mathbf{j} \mid \mathbf{X}_0 = \mathbf{i}\} \ , \tag{5.9}$$

for all $n \in \mathbb{N}$ and $\mathbf{i}, \mathbf{j} \in E^2$.

Definition 5.2.4. ***Transition Matrix***

The transition matrix $\mathbf{P} = (P_{\binom{i^1}{i^2}\binom{j^1}{j^2}})_{\mathbf{i},\mathbf{j}\in E^2}$ is the two-dimensional matrix in $\mathcal{M}^2_{\left[\substack{d\\d}\right]\left[\substack{d\\d}\right]}$ defined by

$$P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} := \mathbb{P}\{\mathbf{X}_{n+1} = \mathbf{j} \mid \mathbf{X}_n = \mathbf{i}\} \ , \tag{5.10}$$

for all $n \in \mathbb{N}$ and $\mathbf{i}, \mathbf{j} \in E^2$.

Example 5.2.5. *Let us consider a two state space system, i.e.* $E = \{a, b\}$*. Let* $(X^1_n, X^2_n)_{n\in\mathbb{N}}$ *be a couple of sequences like*

$$\begin{array}{rcl} X^1 &=& a\ b\ b\ a\ a\ a\ b\ b\ a\ b\ a\ a\ b\ a\ \ldots \\ X^2 &=& b\ a\ b\ b\ b\ a\ a\ a\ a\ b\ b\ b\ a\ b\ \ldots \end{array} \tag{5.11}$$

The transition matrix $\mathbf{P}$ *has the following structures*

$$\mathbf{P} = \begin{pmatrix} \begin{pmatrix} P_{\binom{a}{a}\binom{a}{a}} & P_{\binom{a}{a}\binom{b}{a}} \\ P_{\binom{b}{a}\binom{a}{a}} & P_{\binom{b}{a}\binom{b}{a}} \end{pmatrix} & \begin{pmatrix} P_{\binom{a}{a}\binom{a}{b}} & P_{\binom{a}{a}\binom{b}{b}} \\ P_{\binom{b}{a}\binom{a}{b}} & P_{\binom{b}{a}\binom{b}{b}} \end{pmatrix} \\ \\ \begin{pmatrix} P_{\binom{a}{b}\binom{a}{a}} & P_{\binom{a}{b}\binom{b}{a}} \\ P_{\binom{b}{b}\binom{a}{a}} & P_{\binom{b}{b}\binom{b}{a}} \end{pmatrix} & \begin{pmatrix} P_{\binom{a}{b}\binom{a}{b}} & P_{\binom{a}{b}\binom{b}{b}} \\ P_{\binom{b}{b}\binom{a}{b}} & P_{\binom{b}{b}\binom{b}{b}} \end{pmatrix} \end{pmatrix} . \tag{5.12}$$

The matrix elements can be estimated counting the number of times that a given transition occurs over the total transitions number.

Lemma 5.2.6. $\mathbf{P}$ *is a stochastic matrix, i.e.*

- $\mathbf{P}$ *has non negative entries,* $P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} \geq 0$ *for all* $\mathbf{i}, \mathbf{j} \in E^2$*;*
- $\mathbf{P}$ *has row sums equal to one,* $\sum_{\mathbf{j}\in E^2} P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = 1$ *for all* $\mathbf{i} \in E^2$*.*

Let $\mu^{(n)}$ be the probability distribution of $\mathbf{X}_n$, for every $n \in \mathbb{N}$, defined by

$$\mu^{(n)}_{\mathbf{i}} := \mathbb{P}\{\mathbf{X}_n = \mathbf{i}\} \ ,$$

for all $\mathbf{i} \in E^2$. For $n = 0$, μ^0 is called the initial distribution of the chain, in the following we will denote it by μ omitting the apex.

Remark 5.2.7. *A probability distribution* μ *on* E^2 *is a two-dimensional matrix in* $\mathcal{M}^2_{\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]\left[\begin{smallmatrix}d\\d\end{smallmatrix}\right]}$*.*

The probability of a given sequence can be expressed as

$$\begin{aligned} &\mathbb{P}\{\mathbf{X}_0 = \mathbf{i}_0, \mathbf{X}_1 = \mathbf{i}_1, \ldots, \mathbf{X}_n = \mathbf{i}_n\} = \mathbb{P}\{\mathbf{X}_0 = \mathbf{i}_0, \mathbf{X}_1 = \mathbf{i}_1, \ldots, \mathbf{X}_{n-1} = \mathbf{i}_{n-1}\}\mathbb{P}\{\mathbf{X}_n = \mathbf{i}_n \mid \mathbf{X}_{n-1} = \mathbf{i}_{n-1}\} \\ &= \cdots = \mathbb{P}\{\mathbf{X}_0 = \mathbf{i}_0\}\mathbb{P}\{\mathbf{X}_1 = \mathbf{i}_1 \mid \mathbf{X}_0 = \mathbf{i}_0\} \cdots \mathbb{P}\{\mathbf{X}_n = \mathbf{i}_n \mid \mathbf{X}_{n-1} = \mathbf{i}_{n-1}\} = \mu^{(0)}_{\mathbf{i}_0} P_{\mathbf{i}_0\mathbf{i}_1} \cdots P_{\mathbf{i}_{n-1}\mathbf{i}_n} \end{aligned} \tag{5.13}$$

for all $n \in \mathbb{N}$ and $\mathbf{i}_0, \ldots, \mathbf{i}_n \in E^2$.

Definition 5.2.8. ***Higher-Order Transitions***

The n-th order transition matrix $\mathbf{P}^{(n)}$ *is defined by*

$$P^{(n)}_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = \mathbb{P}\{\mathbf{X}_{m+n} = \mathbf{j} \mid \mathbf{X}_m = \mathbf{i}\} , \tag{5.14}$$

for all $n, m \in \mathbb{N}$, $n > 0$ *and* $\mathbf{i}, \mathbf{j} \in E^2$. *If* $n = 0$ *we have*

$$P^{(0)}_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = \delta_{i^1,j^1}\delta_{i^2,j^2} = \begin{cases} 1 & \text{if} \quad \mathbf{i} = \mathbf{j} \\ 0 & \text{if} \quad \mathbf{i} \neq \mathbf{j} \end{cases} , \tag{5.15}$$

where δ *represents the Kronecker's delta.*

Lemma 5.2.9. *The transition matrix satisfies*

- $P^{(n)}_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = (P *_2 P^{(n-1)})_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = \sum_{\mathbf{l}\in E^2} P_{\binom{i^1}{i^2}\binom{l^1}{l^2}} P^{(n-1)}_{\binom{l^1}{l^2}\binom{j^1}{j^2}}$;
- $P^{(m+n)}_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = \sum_{\mathbf{l}\in E^2} P^{(m)}_{\binom{i^1}{i^2}\binom{l^1}{l^2}} P^{(n)}_{\binom{l^1}{l^2}\binom{j^1}{j^2}}$;

for all $n, m \in \mathbb{N}$ *and* $\mathbf{i}, \mathbf{j} \in E^2$.

The next result is a generalization of the standard case (*cf.* Billingsley [8]).

Theorem 5.2.10. ***Existence***

Suppose that $\mathbf{P} = (P_{\binom{i^1}{i^2}\binom{j^1}{j^2}})_{\mathbf{i},\mathbf{j}\in E^2}$ *is a stochastic matrix and that* μ *is a probability distribution on* E^2. *There exists on some* $(\Omega, \mathscr{F}, \mathbb{P})$ *a bivariate Markov chain* $\mathbf{X}_0, \mathbf{X}_1, \mathbf{X}_2, \ldots$ *with initial distribution* μ *and transition probability* $\mathbf{P}$.

Proof. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be the unit square, i.e. the Cartesian product of two unit intervals, equipped with the Borel σ-algebra and the Lebesgue measure on the plane.

First construct a partition $Q^{(0)}_{(1,1)}, Q^{(0)}_{(1,2)}, \ldots, Q^{(0)}_{(d,d)}$ of $(0,1] \times (0,1]$ into a finite number of rectangles of area $\mathbb{P}(Q^{(0)}_{(i_1,i_2)}) = \mu_{\mathbf{i}}$. Next decompose each rectangle $Q^{(0)}_{(i_1,i_2)}$ into subsets $Q^{(1)}_{(i_1,i_2),(j_1,j_2)}$ of area $\mathbb{P}(Q^{(1)}_{(i_1,i_2),(j_1,j_2)}) = \mu_{\mathbf{i}} P_{\mathbf{ij}}$. Iterating we obtain a sequence of partitions

$$(Q^{(n)}_{(\mathbf{i}_0),\ldots,(\mathbf{i}_n)} : \mathbf{i}_0, \ldots, \mathbf{i}_n \in E^2) \qquad \text{such that} \qquad \mathbb{P}\{Q^{(n)}_{(\mathbf{i}_0),\ldots,(\mathbf{i}_n)}\} = \mu_{\mathbf{i}_0} P_{\mathbf{i}_0\mathbf{i}_1} \cdots P_{\mathbf{i}_{n-1}\mathbf{i}_n} .$$

We define the bivariate chain as

$$\mathbf{X}(\omega) = \mathbf{i} \qquad \text{if} \qquad \omega \in \bigcup_{\mathbf{i}_0,\ldots,\mathbf{i}_{n-1}} Q^{(n)}_{(\mathbf{i}_0),\ldots,(\mathbf{i}_{n-1}),(\mathbf{i})} .$$

By construction it is clear that the set $\{\mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_n = \mathbf{i}_n\}$ coincides with the rectangle $Q^{(n)}_{(\mathbf{i}_0),\ldots,(\mathbf{i}_n)}$ and thus

$$\mathbb{P}\{\mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_n = \mathbf{i}_n\} = \mu_{\mathbf{i}_0} P_{\mathbf{i}_0\mathbf{i}_1} \cdots P_{\mathbf{i}_{n-1}\mathbf{i}_n} \ .$$

It remains to prove that

$$\mathbb{P}\{\mathbf{X}_{n+1} = \mathbf{j} \mid \mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_n = \mathbf{i}_n\} = \mathbb{P}\{\mathbf{X}_{n+1} = \mathbf{j} \mid \mathbf{X}_n = \mathbf{i}_n\} \ . \tag{5.16}$$

The right hand side of Eq. (5.16) can be expressed as

$$\mathbb{P}\{\mathbf{X}_{n+1} = \mathbf{j} \mid \mathbf{X}_n = \mathbf{i}_n\} = \frac{\mathbb{P}\{\mathbf{X}_{n+1} = \mathbf{j}, \mathbf{X}_n = \mathbf{i}_n\}}{\mathbb{P}\{\mathbf{X}_n = \mathbf{i}_n\}} = \frac{\sum_{\mathbf{i}_0,\ldots,\mathbf{i}_{n-1}} \mu_{\mathbf{i}_0} P_{\mathbf{i}_0\mathbf{i}_1} \cdots P_{\mathbf{i}_{n-1}\mathbf{i}_n} P_{\mathbf{i}_n\mathbf{j}}}{\sum_{\mathbf{i}_0,\ldots,\mathbf{i}_{n-1}} \mu_{\mathbf{i}_0} P_{\mathbf{i}_0\mathbf{i}_1} \cdots P_{\mathbf{i}_{n-1}\mathbf{i}_n}} = P_{\mathbf{i}_n\mathbf{j}} \ , \tag{5.17}$$

on the other hand

$$\begin{aligned} \mathbb{P}\{\mathbf{X}_{n+1} = \mathbf{j} \mid \mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_n = \mathbf{i}_n\} &= \frac{\mathbb{P}\{\mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_n = \mathbf{i}_n, \mathbf{X}_{n+1} = \mathbf{j}\}}{\mathbb{P}\{\mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_n = \mathbf{i}_n\}} \\ = \frac{\mu_{\mathbf{i}_0} P_{\mathbf{i}_0\mathbf{i}_1} \cdots P_{\mathbf{i}_{n-1}\mathbf{i}_n} P_{\mathbf{i}_n\mathbf{j}}}{\mu_{\mathbf{i}_0} P_{\mathbf{i}_0\mathbf{i}_1} \cdots P_{\mathbf{i}_{n-1}\mathbf{i}_n}} = P_{\mathbf{i}_n\mathbf{j}} \ . \end{aligned} \tag{5.18}$$

This concludes the proof. □

Remark 5.2.11. *The bivariate Markov chain, when the two components are studied as a whole system, is equivalent to a standard Markov chain with enlarged state space. In particular all the results regarding the classification of the states and the stationary distribution hold in the bivariate case as well. However, in attempt to study the dependence structure and the evolution of the components the standard approach is not helpful.*

Let us consider a chain starting in $\mathbf{i} \in E^2$, we define the probability of a first visit to $\mathbf{j} \in E^2$ after n steps as

$$f^{(n)}_{\mathbf{ij}} := \mathbb{P}\{\mathbf{X}_1 \neq \mathbf{j}, \ldots, \mathbf{X}_{n-1} \neq \mathbf{j}, \mathbf{X}_n = \mathbf{j} \mid \mathbf{X}_0 = \mathbf{i}\} \ . \tag{5.19}$$

Summing over all possible n, we obtain the probability of an eventual visit in $\mathbf{j}$ as

$$f_{\mathbf{ij}} = \sum_{n=1}^{\infty} f^{(n)}_{\mathbf{ij}} \ . \tag{5.20}$$

In terms of f we can classify the state as follows.

Definition 5.2.12. ***Persistent and Transient States***

- *A state* $\mathbf{i} \in E^2$ *is persistent if eventually the chain return to it, that is* $f_{\mathbf{ii}} = 1$.
- *A state* $\mathbf{i} \in E^2$ *is transient in the opposite case* $f_{\mathbf{ii}} < 1$.

Definition 5.2.13. ***Stationary Distributions***

The distribution π *on* E^2 *is a stationary distribution for the bivariate Markov chain* $(\mathbf{P}, \mu)$ *if*

$$\sum_{i^1, i^2 \in E^2} \pi_{\binom{i^1}{i^2}} P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = \pi_{\binom{j^1}{j^2}} \ . \tag{5.21}$$

Definition 5.2.14. ***Ergodic Markov Chain***

A transition matrix $\mathbf{P}$ *is said to be ergodic if, for all* $\mathbf{i}, \mathbf{j} \in E^2$*, there exists* $r_0 \in \mathbb{N}$ *such that for all* $r > r_0$ *the r-th order transition matrix* $\mathbf{P}^{(r)}$ *has only positive entries, that is* $P^{(r)}_{\binom{i^1}{i^2}\binom{j^1}{j^2}} > 0$ *for all* $\mathbf{i}, \mathbf{j} \in E^2$*. A bivariate Markov chain is said to be ergodic if it can be generated by some initial distribution and an ergodic transition matrix. In other words a bivariate Markov chain is ergodic if every state can be reached, in a finite number of steps, from any initial state.*

The next result is a generalization of the standard case (*cf.* Kolarov and Sinai [62]).

Theorem 5.2.15. *Let* P *be an ergodic transition matrix for a bivariate Markov chain, then there exists a unique stationary distribution* π *on* E^2*. The n-th order transition matrix* $\mathbf{P}^{(n)}$ *converges to the stationary distribution* π*, that is*

$$\lim_{n \to \infty} \mathbf{P}^{(n)}_{\mathbf{ij}} = \pi_{\mathbf{j}} \ . \tag{5.22}$$

The stationary distribution π *is such that* $\pi_{\mathbf{j}} > 0$ *for all* $\mathbf{j} \in E^2$.

Proof. In order to show the result we first define a metric on the space of probability distributions on E^2.

Let μ and ξ be two probability distributions on E^2, we define

$$d(\mu, \xi) := \frac{1}{2} \sum_{\mathbf{i} \in E^2} \left| \mu_{\mathbf{i}} - \xi_{\mathbf{i}} \right|. \tag{5.23}$$

It is to verify that d is a distance on the space of probability distributions on E^2. Moreover, we observe that

$$0 = \sum_{\mathbf{i} \in E^2} \mu_{\binom{i^1}{i^2}} - \sum_{\mathbf{i} \in E^2} \xi_{\binom{i^1}{i^2}} = \sum_{\mathbf{i} \in E^2} (\mu_{\binom{i^1}{i^2}} - \xi_{\binom{i^1}{i^2}}) = \sum_{\mathbf{i} \in E^2} (\mu_{\binom{i^1}{i^2}} - \xi_{\binom{i^1}{i^2}})^+ - \sum_{\mathbf{i} \in E^2} (\xi_{\binom{i^1}{i^2}} - \mu_{\binom{i^1}{i^2}})^+ \ , \tag{5.24}$$

where the apex + denotes the positive part. Then we have

$$\sum_{\mathbf{i}\in E^2}(\mu_{\binom{i^1}{i^2}}-\xi_{\binom{i^1}{i^2}})^+=\sum_{\mathbf{i}\in E^2}(\xi_{\binom{i^1}{i^2}}-\mu_{\binom{i^1}{i^2}})^+ \ .$$

Hence, we can express the distance as

$$d(\mu,\xi)=\frac{1}{2}\sum_{\mathbf{i}\in E^2}\left|\mu_{\mathbf{i}}-\xi_{\mathbf{i}}\right|=\frac{1}{2}\sum_{\mathbf{i}\in E^2}(\mu_{\binom{i^1}{i^2}}-\xi_{\binom{i^1}{i^2}})^+ +\frac{1}{2}\sum_{\mathbf{i}\in E^2}(\xi_{\binom{i^1}{i^2}}-\mu_{\binom{i^1}{i^2}})^+=\sum_{\mathbf{i}\in E^2}(\mu_{\binom{i^1}{i^2}}-\xi_{\binom{i^1}{i^2}})^+ \tag{5.25}$$

We notice that

$$(\mu *_2 \mathbf{P}^{(n)})_{\mathbf{j}}=\sum_{\mathbf{i}\in E^2}\mu_{\binom{i^1}{i^2}}P^{(n)}_{\binom{i^1}{i^2}\binom{j^1}{j^2}} \qquad n\in\mathbb{N} \ ,$$

is a probability distribution on E^2 and in the following will be denoted simply by $\mu\mathbf{P}^{(n)}$.

Let us first show that

$$d(\mu\mathbf{P}^{(r)},\xi\mathbf{P}^{(r)})\leq(1-\epsilon_1)d(\mu,\xi) \ , \tag{5.26}$$

for all $r>r_0$ and some $\epsilon_1>0$. We have

$$d(\mu\mathbf{P}^{(n)},\xi\mathbf{P}^{(n)})=\sum_{\mathbf{i}\in E^2}[(\mu\mathbf{P}^{(n)})_{\binom{i^1}{i^2}}-(\xi\mathbf{P}^{(r)})_{\binom{i^1}{i^2}}]^+=\sum_{\mathbf{i}\in E^2}\left[\sum_{\mathbf{h}\in E^2}(\mu_{\binom{h^1}{h^2}}-\xi_{\binom{h^1}{h^2}})P^{(n)}_{\binom{h^1}{h^2}\binom{i^1}{i^2}}\right]^+ . \tag{5.27}$$

Let us denote by I_+ the subset of E^2 such that

$$\sum_{\mathbf{h}\in E^2}(\mu_{\binom{h^1}{h^2}}-\xi_{\binom{h^1}{h^2}})P^{(n)}_{\binom{h^1}{h^2}\binom{i^1}{i^2}}\geq 0 \ , \tag{5.28}$$

for $(i_1,i_2)\in I^+$. If $\mu\neq\xi$ we have that $I^+\subset E^2$, proper subset, therefore

$$\sum_{\mathbf{i}\in I^+}P^{(n)}_{\binom{h^1}{h^2}\binom{i^1}{i^2}}\leq 1 \ . \tag{5.29}$$

Moreover, being the chain ergodic, there exists $r_0\in\mathbb{N}$ such that

$$\sum_{\mathbf{i}\in I^+}P^{(r)}_{\binom{h^1}{h^2}\binom{i^1}{i^2}}\leq 1-\epsilon_1 \ , \tag{5.30}$$

for all $r>r_0$ and for some $\epsilon_1>0$. Hence, we have

$$d(\mu\mathbf{P},\xi\mathbf{P})\leq\sum_{\mathbf{i}\in I^+}\sum_{\mathbf{h}\in E^2}(\mu_{\binom{h^1}{h^2}}-\xi_{\binom{h^1}{h^2}})^+P_{\binom{h^1}{h^2}\binom{i^1}{i^2}}\leq(1-\epsilon_1)\sum_{\mathbf{i}\in E^2}(\mu_{\binom{h^1}{h^2}}-\xi_{\binom{h^1}{h^2}})^+=(1-\epsilon_1)d(\mu,\xi) \ . \tag{5.31}$$

Let μ^0 be a distribution on E^2 and $\mu^n = \mu^0 \mathbf{P}^{(n)}$ for $n \in \mathbb{N}$. We show that $(\mu^n)_{n\in\mathbb{N}}$ is a Cauchy sequence. That is, for any $\epsilon > 0$ there exists n_0^ϵ such that for any $n > n_0^\epsilon$ and $k \in \mathbb{N}$ we have

$$d(\mu^n, \mu^{n+k}) < \epsilon \; . \tag{5.32}$$

Let us consider $n > r_0$ and $k \in \mathbb{N}$, we have

$$d(\mu^n, \mu^{n+k}) \leq (1 - \epsilon_1) d(\mu^{n-r}, \mu^{n+k-r}), \tag{5.33}$$

for some fixed $r > r_0$. Iterating we get

$$d(\mu^n, \mu^{n+k}) \leq (1 - \epsilon_1)^m d(\mu^{n-mr}, \mu^{n+k-mr}) \leq (1 - \epsilon_1)^m, \tag{5.34}$$

with $m \in \mathbb{N}$ such that $n - r(m-1) > r_0$. For sufficiently large n_0^ϵ there exists m such that $(1-\epsilon_1)^m < \epsilon$ and thus $(\mu^n)_{n\in\mathbb{N}}$ is a Cauchy sequence. Let $\pi = \lim_{n\to\infty} \mu^n$, we have

$$\sum_{i^1,i^2 \in E^2} \pi_{\binom{i^1}{i^2}} P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = \lim_{n\to\infty} \sum_{i^1,i^2\in E^2} (\mu^0 P^{(n)})_{\binom{i^1}{i^2}} P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = \lim_{n\to\infty} (\mu^0 P^{(n+1)})_{\binom{j^1}{j^2}} = \pi_{\binom{j^1}{j^2}}, \tag{5.35}$$

so the limit distribution is stationary. It is easy to show that such stationary distribution is unique. Indeed let π_1 and π_2 two stationary distributions obtained as a limit of the sequences $(\mu^0 \mathbf{P}^{(n)})_{n\in\mathbb{N}}$ and $(\xi^0 \mathbf{P}^{(n)})_{n\in\mathbb{N}}$, respectively. Being the limit distributions stationary, we have

$$d(\pi_1, \pi_2) = d(\pi_1 \mathbf{P}^{(r)}, \pi_2 \mathbf{P}^{(r)}) \leq (1 - \epsilon_1) d(\pi_1, \pi_2) \; , \tag{5.36}$$

for $r > r_0$. It follows that $d(\pi_1, \pi_2) = 0$, that is $\pi_1 = \pi_2$, which concludes the proof. □

The bivariate Markov chain possesses the memoryless property with respect to the filtration of the whole system. The two components are in general not Markovian with respect to their own filtrations. In principle we can distinguish two cases: components Markovian with respect to the filtration of the whole system and to their own filtration. In the following we analyze the dependence structure of the bivariate chain.

Proposition 5.2.16. *The component α satisfies*

$$\mathbb{P}\{X^\alpha_{n+1} = j^\alpha \mid \mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_n = \mathbf{i}_n\} = \mathbb{P}\{X^\alpha_{n+1} = j^\alpha \mid \mathbf{X}_n = \mathbf{i}_n\} \; , \tag{5.37}$$

for $\alpha = 1, 2$.

Proof. This is a direct consequence of the law of total probability and the memoryless property of the bivariate chain. □

Remark 5.2.17. *The next visited state of a component depends on the present state of the whole system.*

Definition 5.2.18. ***Marginal Transition Matrices***

The marginal transition matrices $\mathbf{P}^\alpha = (P^\alpha_{\binom{i^1}{i^2}(j^\alpha)})_{\mathbf{i}\in E^2, j^\alpha\in E}$*, for* $\alpha = 1,2$*, are the two-dimensional matrices in* $\mathcal{M}^2_{\left[\begin{smallmatrix} d \\ d \end{smallmatrix}\right]\left[\begin{smallmatrix} d \\ 1 \end{smallmatrix}\right]}$ *defined by*

$$P^\alpha_{\binom{i^1}{i^2}(j^\alpha)} := \mathbb{P}\{X^\alpha_{n+1} = j^\alpha \mid \mathbf{X}_n = \mathbf{i}\} \ , \tag{5.38}$$

for all $n \in \mathbb{N}$, $\mathbf{i} \in E^2$ *and* $j^\alpha \in E$.

Example 5.2.19. *Let us consider again the two state space system* $E = \{a, b\}$ *and the sequences* $(X^1_n, X^2_n)_{n\in\mathbb{N}}$

$$\begin{array}{rcl} X^1 & = & a\ b\ b\ a\ a\ a\ b\ b\ a\ b\ a\ a\ b\ a\ \ldots \\ X^2 & = & b\ a\ b\ b\ b\ a\ a\ a\ a\ b\ b\ b\ a\ b\ \ldots \end{array} . \tag{5.39}$$

Let us assume that (X^1, X^2) *is a strong Markov chain. Hence, the marginal transition matrix for component 1 has the following structures*

$$\mathbf{P}^1 = \begin{pmatrix} \begin{pmatrix} P_{\binom{a}{a}(a)} & P_{\binom{a}{a}(b)} \\ P_{\binom{b}{a}(a)} & P_{\binom{b}{a}(b)} \end{pmatrix} \\ \\ \begin{pmatrix} P_{\binom{a}{b}(a)} & P_{\binom{a}{b}(b)} \\ P_{\binom{b}{b}(a)} & P_{\binom{b}{b}(b)} \end{pmatrix} \end{pmatrix} . \tag{5.40}$$

and similarly for component 2.

Lemma 5.2.20. *The marginal transition matrices* $\mathbf{P}^1$ *and* $\mathbf{P}^2$ *are stochastic, i.e.*

- $\mathbf{P}^\alpha$ *has non negative entries,* $P^\alpha_{\binom{i^1}{i^2}(j^\alpha)} \geq 0$ *for all* $\mathbf{i} \in E^2$ *and* $j^\alpha \in E$;
- $\mathbf{P}^\alpha$ *has row sums equal to one,* $\sum_{j\in E^2} P_{\binom{i^1}{i^2}(j)} = 1$ *for all* $\mathbf{i} \in E^2$;

for $\alpha = 1, 2$.

Remark 5.2.21. *The marginal transition matrix for component 1 can be expressed in terms of transition matrix as*

$$P^\alpha_{\binom{i^1}{i^2}(j^1)} = \sum_{j^2\in E} P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} \ ,$$

and similarly for component 2.

In general, the bivariate chain transition matrix is not the product of the marginals

$$\mathbb{P}\{X^1_{n+1} = j^1, X^2_{n+1} = j^2 \mid \mathbf{X}_n = \mathbf{i}\} \neq \mathbb{P}\{X^1_{n+1} = j^1 \mid \mathbf{X}_n = \mathbf{i}\}\mathbb{P}\{X^2_{n+1} = j^2 \mid \mathbf{X}_n = \mathbf{i}\} \ . \tag{5.41}$$

However, if the next visited states of the components depends only the present state of the system, i.e. the components are *conditionally independent*, we have

$$\mathbb{P}\{X^1_{n+1} = j^1, X^2_{n+1} = j^2 \mid \mathbf{X}_n = \mathbf{i}\} = \mathbb{P}\{X^1_{n+1} = j^1 \mid \mathbf{X}_n = \mathbf{i}\}\mathbb{P}\{X^2_{n+1} = j^2 \mid \mathbf{X}_n = \mathbf{i}\} \ . \tag{5.42}$$

Definition 5.2.22. ***Marginal Higher-Order Transitions***

The n-th order marginal transition matrix $\mathbf{P}^\alpha(n)$*, for* $\alpha = 1, 2$*, is defined by*

$$P^\alpha_{\binom{i^1}{i^2}(j^\alpha)}(n) = \mathbb{P}\{X^\alpha_{m+n} = j^\alpha \mid \mathbf{X}_m = \mathbf{i}\} \ , \tag{5.43}$$

for all $n, m \in \mathbb{N}$*,* $n > 0$*,* $\mathbf{i} \in E^2$ *and* $j^{alpha} \in E$*. If* $n = 0$ *we have*

$$P^1_{\binom{i^1}{i^2}(j^1)}(0) = \delta_{i^1, j^1} \ , \tag{5.44}$$

and similarly for component 2.

Proposition 5.2.23. *Let* P *be an ergodic transition matrix for a bivariate Markov chain with unique stationary distribution* π *on* E^2*. The n-th order marginal transition matrices* $(\mathbf{P}^1(n), \mathbf{P}^2(n))$ *converge to the marginal stationary distributions* π^α*, for* $\alpha = 1, 2$*, that is*

$$\lim_{n\to\infty} \mathbf{P}^\alpha_{\mathbf{i}j^\alpha}(n) = \pi^\alpha_{j^\alpha} \ . \tag{5.45}$$

The marginal stationary distributions π^α *for* $\alpha = 1, 2$ *are such that*

$$\begin{aligned} \pi^1_{j^1} &= \sum_{j^2 \in E} \pi_{\binom{i^1}{i^2}} \\ \pi^2_{j^2} &= \sum_{j^1 \in E} \pi_{\binom{i^1}{i^2}} \ , \end{aligned} \tag{5.46}$$

for $j^1, j^2 \in E$*.*

As well known, components of a multivariate Markov process are in general not Markovian with respect to their own filtrations. Bielecki *et al.* [7] distinguished between multivariate Markov processes whose components are Markovian with respect to their own filtration (weak Markov consistency) and multivariate Markov processes whose components are Markovian with respect to the filtration of the whole system (strong Markov consistency).

It could be interesting to address the problem of the conditions to impose on a multivariate Markov process in order to obtain components that are Markov process with a specific distribution, Ball and Yeo [2] and Bielecki *et al.* [6] answered to this question for continuous time Markov chain.

In the following, we discuss a consistency condition for bivariate Markov chains (*cf.* Bielecki *et al.* [7]) and we derive the condition a bivariate Markov chain should satisfy in order to obtain given Markov chains as components.

Definition 5.2.24. ***(strong) Markovian Consistency***

Let $\mathbf{X} = (X^1, X^2)$ *be a bivariate Markov chain, we said that* $\mathbf{X}$ *satisfy the Markovian consistency condition if*

$$\begin{aligned}\mathbb{P}\{X^1_{n+1} = j^1 \mid \mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_{n-1} = \mathbf{i}_{n-1}, \mathbf{X}_n = \mathbf{i}\} &= \mathbb{P}\{X^1_{n+1} = j^1 \mid X^1_n = i^1\} \\ \mathbb{P}\{X^2_{n+1} = j^2 \mid \mathbf{X}_0 = \mathbf{i}_0, \ldots, \mathbf{X}_{n-1} = \mathbf{i}_{n-1}, \mathbf{X}_n = \mathbf{i}\} &= \mathbb{P}\{X^2_{n+1} = j^2 \mid X^2_n = i^2\}\ ,\end{aligned} \tag{5.47}$$

for all $i^1, j^1, i^2, j^2 \in E$.

Proposition 5.2.25. *Let* Y^1 *and* Y^2 *be two Markov chains on* E *and let* $\mathbf{p}^1$ *and* $\mathbf{p}^2$ *be the transition matrices of* Y^1 *and* Y^2*, respectively. If the system of linear equations in the unknowns* $P_{\binom{i^1}{i^2}\binom{j^1}{j^2}}$*, for* $\mathbf{i}, \mathbf{j} \in E^2$*:*

$$\begin{aligned}\sum_{j^2\in E} P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} &= p^1_{i^1 j^1} \qquad \text{for all} \qquad i^2 \in E \\ \sum_{j^1\in E} P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} &= p^2_{i^2 j^2} \qquad \text{for all} \qquad i^1 \in E\ ,\end{aligned} \tag{5.48}$$

has a solution, then there exists a Markovian consistent bivariate Markov chain $\mathbf{Y}$ *having* $\mathbf{P} = (P_{\binom{i^1}{i^2}\binom{j^1}{j^2}})_{\mathbf{i},\mathbf{j}\in E^2}$ *as a transition matrix and* (Y^1, Y^2) *as components.*

Proof. If the system of Eqs. (5.48) has a solution, then from Theorem 5.2.10 given an initial distribution on E^2 there exists a bivariate Markov chain, $\mathbf{Y}$, having $\mathbf{P} = (P_{\binom{i^1}{i^2}\binom{j^1}{j^2}})_{\mathbf{i},\mathbf{j}\in E^2}$ as a transition matrix. Moreover if the Eqs. (5.48) are satisfied, $\mathbf{p}^1$ and $\mathbf{p}^2$ are the marginal transition matrices. Thus, Y^1 and Y^2 are the components of $\mathbf{Y}$, that is $\mathbf{Y}$ is Markovian consistent bivariate Markov chain. □

Remark 5.2.26. *The system of Eqs. (5.48) has at least one solution. Indeed, in the independent case, it is easy to verify that*

$$P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} = p^1_{i^1j^1}p^2_{i^2j^2} \tag{5.49}$$

for $\mathbf{i}, \mathbf{j} \in E^2$*, is a solution.*

Corollary 5.2.27. *Let* $\mathbf{X} = (X^1, X^2)$ *be a Markovian consistent bivariate Markov chain with transition matrix* $\mathbf{P} = (P_{\binom{i^1}{i^2}\binom{j^1}{j^2}})_{\mathbf{i},\mathbf{j}\in E^2}$*, then* X^1 *and* X^2 *are Markov chain with transition matrices* $\mathbf{p}^1$ *and* $\mathbf{p}^2$ *if*

$$\begin{aligned} p^1_{i^1j^1} &= \sum_{j^2\in E} P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} \qquad \text{for all} \qquad i^2 \in E \\ p^2_{i^2j^2} &= \sum_{j^1\in E} P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} \qquad \text{for all} \qquad i^1 \in E\ . \end{aligned} \tag{5.50}$$

Remark 5.2.28. *Proposition 5.2.25 and Corollary 5.2.25 give a necessary and sufficient condition for a Markovian consistent bivariate Markov chain to have given Markov chains as marginals.*

In Bielecki *et al.* [7] an example of Markovian consistent continuous-time bivariate Markov chain is discussed. We give an example of Markovian consistent bivariate Markov chain with given Markov chains as components.

Example 5.2.29. *Let us consider again the two state space system* $E = \{a, b\}$*. Let* Y^1 *and* Y^2 *be two Markov chains, with values in* E*, with transition matrices*

$$\mathbf{p}^1 = \begin{pmatrix} p^1_{aa} & p^1_{ab} \\ p^1_{ba} & p^1_{bb} \end{pmatrix} = \begin{pmatrix} 1-(\alpha+\gamma) & \alpha+\gamma \\ 0 & 1 \end{pmatrix} \tag{5.51}$$

and

$$\mathbf{p}^2 = \begin{pmatrix} p^2_{aa} & p^2_{ab} \\ p^2_{ba} & p^2_{bb} \end{pmatrix} = \begin{pmatrix} 1-(\beta+\gamma) & \beta+\gamma \\ 0 & 1 \end{pmatrix}\ . \tag{5.52}$$

It is easy to verify that

$$\mathbf{P} = \begin{pmatrix} \begin{pmatrix} P_{\binom{a}{a}\binom{a}{a}} & P_{\binom{a}{a}\binom{b}{a}} \\ P_{\binom{b}{a}\binom{a}{a}} & P_{\binom{b}{a}\binom{b}{a}} \end{pmatrix} & \begin{pmatrix} P_{\binom{a}{a}\binom{a}{b}} & P_{\binom{a}{a}\binom{b}{b}} \\ P_{\binom{b}{a}\binom{a}{b}} & P_{\binom{b}{a}\binom{b}{b}} \end{pmatrix} \\ & \\ \begin{pmatrix} P_{\binom{a}{b}\binom{a}{a}} & P_{\binom{a}{b}\binom{b}{a}} \\ P_{\binom{b}{b}\binom{a}{a}} & P_{\binom{b}{b}\binom{b}{a}} \end{pmatrix} & \begin{pmatrix} P_{\binom{a}{b}\binom{a}{b}} & P_{\binom{a}{b}\binom{b}{b}} \\ P_{\binom{b}{b}\binom{a}{b}} & P_{\binom{b}{b}\binom{b}{b}} \end{pmatrix} \end{pmatrix} \tag{5.53}$$

$$= \begin{pmatrix} 1-(\alpha+\beta+\gamma) & \alpha & \beta & \gamma \\ 0 & 1-(\alpha+\gamma) & 0 & \alpha+\gamma \\ 0 & 0 & 1-(\beta+\gamma) & \beta+\gamma \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

is a solution of the system of Eqs. (5.48), then $\mathbf{P}$ *is a transition matrix of a bivariate Markov chain* $\mathbf{Y}$ *with values in* E^2*.*

Chapter 6

Bivariate Semi-Markov Chains

In this chapter we would like to define a bivariate process able to capture the semi-Markov environment features. We will define a multivariate discrete time process whose components are semi-Markov with respect to the information of the whole process.

The Markovian memoryless property will be preserved for any component at its transition time, i.e. each component will preserve its own renewal time process. The renewal time processes of the system will be independent, so the waiting times in the states for each component will only depend on its present state. However the next visited state of a component depends on the present state of the whole system, that is the embedded chain is Markovian with respect to the filtration of the whole system.

Thus, in this bivariatete case the one step transition probabilities of each component will depend on the present state of the whole system. The evolution in time and the main properties of this bivariate system will be studied.

6.1 Bivariate Semi-Markov Chain: Main Definitions

Let $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ be a complete filtered probability space and let $E = \{1, \ldots, d\}$ be a given finite set.

Let us consider a system consisting of two parts, each part has values in $E = \{1, \ldots, d\}$. Let us denote by $J^\alpha = (J_n^\alpha)_{n\in\mathbb{N}}$, for $\alpha = 1, 2$, the sequence of states visited by the α part, each J^α takes values in E.

Let $\mathbf{v} = \begin{pmatrix} v^1 \\ v^2 \end{pmatrix}$ denote a vector in E^2.

Assumption 6.1.1. *Let us assume that*

A1 $\mathbf{J} = (J^1, J^2)$ *is a bivariate Markov chain, i.e. the process and its components have the memoryless Markov property with respect to the filtration of the whole system.*

The one step transition probability for the bivariate Markov chain $\mathbf{J}$ *is defined for* $\mathbf{i}, \mathbf{j} \in E^2$ *as*

$$\mathbb{P}\{\mathbf{J}_{n+1} = \mathbf{j} \mid \mathbf{J}_0 = \mathbf{i}_0, \ldots, \mathbf{J}_n = \mathbf{i}_n\} = \mathbb{P}\{\mathbf{J}_{n+1} = \mathbf{j} \mid \mathbf{J}_n = \mathbf{i}_n\} =: P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} \tag{6.1}$$

The process is time homogeneous and then the transition probabilities do not depend on present time.

Let us denote with $(T_n)_{n\in\mathbb{N}}$ the sequence of system's transition times, the state space of transition times will be $\mathbb{N}$ since we are considering a discrete time system.

Let us also introduce the sequence of random variables $X_n = T_{n+1} - T_n$, for every $n \in \mathbb{N}$. X_n is the sojourn time of the system in state $\mathbf{J}_n$.

We define the counting process

$$N(t) = \max\{n \in \mathbb{N} \mid T_n \leq t\} \qquad \forall\, t \in \mathbb{N},$$

which gives the number of transitions up to time t of the system.

Remark 6.1.2. *The bivariate semi-Markov chain, when the two components are studied as a whole system with a univariate renewal time process, is equivalent to a standard semi-Markov chain with enlarged state space.*

Definition 6.1.3. ***Bivariate Cumulated Semi-Markov Kernel***

A two-dimension matrix valued function $\mathbf{Q} = (Q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t);\ \mathbf{i}, \mathbf{j} \in E^2,\ t \in \mathbb{N}^*)$ *is a discrete-time bivariate cumulated semi-Markov kernel if*

- $Q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t) \geq 0$ *for every* $\mathbf{i}, \mathbf{j} \in E^2$ *and* $t \in \mathbb{N}^*$*;*
- $Q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(0) = 0$ *for every* $\mathbf{i}, \mathbf{j} \in E^2$*;*
- $\left(\lim_{t\to\infty} Q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t)\right)_{\mathbf{i},\mathbf{j}\in E^2}$ *is a bivariate Markov chain transition probability.*

Definition 6.1.4. ***Bivariate Semi-Markov Kernel***

A two-dimension matrix valued function $\mathbf{q} = (q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t);\ \mathbf{i},\mathbf{j} \in E^2,\ t \in \mathbb{N}^*)$ *is a discrete-time bivariate semi-Markov kernel if*

- $q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t) \geq 0$ *for every* $\mathbf{i},\mathbf{j} \in E^2$ *and* $t \in \mathbb{N}^*$*;*
- $q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(0) = 0$ *for every* $\mathbf{i},\mathbf{j} \in E^2$*;*
- $\left(\sum_{t=1}^{\infty} q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t)\right)_{\mathbf{i},\mathbf{j}\in E^2}$ *is a bivariate Markov chain transition probability.*

Definition 6.1.5. ***Bivariate Markov Renewal Chain***

A random sequence $(\mathbf{J}, T) = (\mathbf{J}_n, T_n)_{n\in\mathbb{N}^*}$ *is a bivariate Markov renewal chain if for all* $n \in \mathbb{N}$*,* $\mathbf{i},\mathbf{j} \in E^2$ *and* $t \in \mathbb{N}$ *it satisfies*

$$\mathbb{P}\{\mathbf{J}_{n+1} = \mathbf{j}, T_{n+1} - T_n \leq t|\sigma(\mathbf{J}_a, T_a), 0 \leq a \leq n\} = \mathbb{P}\{\mathbf{J}_{n+1} = \mathbf{j}, T_{n+1} - T_n \leq t|\mathbf{J}_n = \mathbf{i}\}\ . \tag{6.2}$$

If the probability in Eq. (6.2) does not depend on n*,* $(\mathbf{J}, T)$ *is time homogenous and its associated bivariate semi-Markov kernel* $\mathbf{q}$ *is defined by*

$$q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t) := \mathbb{P}\{\mathbf{J}_{n+1} = \mathbf{j}, T_{n+1} - T_n = t \mid \mathbf{J}_n = \mathbf{i}\}\ . \tag{6.3}$$

If $(\mathbf{J}, T)$ is a bivariate Markov renewal chain then $(J_n^1, J_n^2)_{n\in\mathbb{N}^*}$ is a bivariate Markov chain, called the embedded bivariate Markov chain associated to the bivariate Markov renewal chain $(\mathbf{J}, T)$. The bivariate Markov transition probability of $\mathbf{J}$ is defined by

$$p_{\binom{i^1}{i^2}\binom{j^1}{j^2}} := \mathbb{P}\{\mathbf{J}_{n+1} = \mathbf{j} \mid \mathbf{J}_n = \mathbf{i}\} = \sum_{t=0}^{\infty} q_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t)\ . \tag{6.4}$$

Let us define the cumulative unconditional distribution of X_n as

$$H_{\binom{i^1}{i^2}}(t) := \mathbb{P}\{X_n \leq t \mid \mathbf{J}_n = \mathbf{i}\}\ , \tag{6.5}$$

for every $n \in \mathbb{N}$ and $\mathbf{i} \in E$.

Assumption 6.1.6. *The distribution of* X_n *does not depend on the next visited state, that is*

$$\begin{aligned} G_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t) &:= \mathbb{P}\{X_n \leq t \mid \mathbf{J}_n = \mathbf{i}, \mathbf{J}_{n+1} = \mathbf{j}\} \\ &= \mathbb{P}\{X_n \leq t \mid \mathbf{J}_n = \mathbf{i}\} = H_{\binom{i^1}{i^2}}(t)\ , \end{aligned} \tag{6.6}$$

for every $n \in \mathbb{N}$ *and* $\mathbf{i}, \mathbf{j} \in E^2$.

Definition 6.1.7. ***Bivariate Semi-Markov Chain***

Let $(\mathbf{J}, T)$ *be a bivariate Markov renewal chain and* N *its associated counting process. The process* $\mathbf{Z} = (Z^1(t), Z^2(t))_{tin\mathbb{N}^*}$ *defined by*

$$\mathbf{Z}(t) := (J^1_{N(t)}, J^2_{N(t)}), \tag{6.7}$$

is a bivariate semi-Markov chain associated to the bivariate Markov renewal chain $(\mathbf{J}, T)$*. In other words* $\mathbf{Z}(t)$ *gives the position of the embedded bivariate Markov chain at time* t.

The evolution of a bivariate semi-Markov chain from an initial state can be studied by the associated transition probability.

Definition 6.1.8. ***Bivariate Transition Probability***

The bivariate transition probability of the bivariate semi-Markov chain $\mathbf{Z}$ *is the two-dimension matrix valued function* $\phi = (\phi_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t);\ \mathbf{i}, \mathbf{j} \in E^2,\ t \in \mathbb{N}^*)$ *defined by*

$$\phi_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t) := \mathbb{P}\{\mathbf{Z}(t) = \mathbf{j} \mid \mathbf{Z}(0) = \mathbf{i}, T_{N(0)} = 0\}\ . \tag{6.8}$$

The following result allows us to express the transition probability in a recursive way as a function of the bivariate semi-Markov kernel.

Proposition 6.1.9. ***Evolution Equation***

For all $\mathbf{i}, \mathbf{j} \in E^2$ *and* $t \in \mathbb{N}$*, we have*

$$\Phi_{\binom{i^1}{i^2}\binom{j^1}{j^2}}(t) = \delta_{\binom{i^1}{i^2}\binom{j^1}{j^2}}[1 - H_{\binom{i^1}{i^2}}(t)] + \sum_{l^1,l^2 \in E}\sum_{\tau=1}^{t} q_{\binom{i^1}{i^2}\binom{l^1}{l^2}}(\tau)\Phi_{\binom{l^1}{l^2}\binom{j^1}{j^2}}(t-\tau)\ , \tag{6.9}$$

where δ *represents the Kronecker symbol.*

Proof. In order to evaluate the transition probability in t let us first distinguish between the trajectories with at least one transition in t-steps and those without, that is

$$\mathbb{P}\{\mathbf{Z}(t) = \mathbf{j} \mid \mathbf{Z}(0) = \mathbf{i}\} = \mathbb{P}\{\mathbf{Z}(t) = \mathbf{j}, T_1 > t \mid \mathbf{Z}(0) = \mathbf{i}\} + \mathbb{P}\{\mathbf{Z}(t) = \mathbf{j}, T_1 \le t \mid \mathbf{Z}(0) = \mathbf{i}\}\ . \tag{6.10}$$

The first term on the r.h.s of Eq. (6.10) is given by

$$\mathbb{P}\{\mathbf{Z}(t)=\mathbf{j}, T_1>t \mid \mathbf{Z}(0)=\mathbf{i}\} = \delta_{\binom{i^1}{i^2}\binom{j^1}{j^2}}[1-H_{\binom{i^1}{i^2}}(t)] \ , \tag{6.11}$$

by applying the law of total probability to the second term on the r.h.s of Eq. (6.10) we get

$$\begin{aligned}\mathbb{P}\{\mathbf{Z}(t)=\mathbf{j}, T_1\le t \mid \mathbf{Z}(0)=\mathbf{i}\} &= \sum_{l^1,l^2\in E}\sum_{\tau=1}^{t}\mathbb{P}\{\mathbf{Z}(t)=\mathbf{j}, T_1=\tau, Z^1(T_1)=l^1, Z^2(T_1)=l^2 \mid \mathbf{Z}(0)=\mathbf{i}\} \\ &= \sum_{l^1,l^2\in E}\sum_{\tau=1}^{t}\mathbb{P}\{\mathbf{Z}(t)=\mathbf{j} \mid T_1=\tau, \mathbf{Z}(\tau)=\mathbf{l}, \mathbf{Z}(0)=\mathbf{i}\}\mathbb{P}\{T_1=\tau, \mathbf{Z}(\tau)=\mathbf{l} \mid \mathbf{Z}(0)=\mathbf{i}\},\end{aligned} \tag{6.12}$$

using the time homogeneity and the memoryless property of the bivariate chain at the transition time, we obtain

$$\mathbb{P}\{\mathbf{Z}(t)=\mathbf{j}, T_1\le t \mid \mathbf{Z}(0)=\mathbf{i}\} = \sum_{l^1,l^2\in E}\sum_{\tau=1}^{t} q_{\binom{i^1}{i^2}\binom{l^1}{l^2}}(\tau)\Phi_{\binom{l^1}{l^2}\binom{j^1}{j^2}}(t-\tau) \ , \tag{6.13}$$

this concludes the proof. □

In many applications, it might be more convenient to introduce a sequence of transition times for each component of the system. In other words, we can define $(T_n^{(\alpha)})_{n\in\mathbb{N}}$ the sequence of transition times of α-th component, for $\alpha=1,2$. We can introduce the sequence of random variables $X_n^\alpha = T_{n+1}^\alpha - T_n^\alpha$, for every $n\in\mathbb{N}$. X_n^α is the sojourn time of component α in state J_n^α. For every component of our system we can define the counting process

$$N^\alpha(t) = \max\{n\in\mathbb{N} \mid T_n^\alpha \le t\} \qquad \forall\ t\in\mathbb{N},$$

which gives the number of transitions up to time t of part α.

Remark 6.1.10. *We notice that, for every $t\in\mathbb{N}$, we have*

$$N(t) = N^1(t) + N^2(t) - \#\{T^1_{N(k)} = T^2_{N(k)}, k\le t\} \ ,$$

here the number operator (#) allows not to count twice a simultaneous transition of the components. We can now define

$$T_n = \min\{t\in\mathbb{N} \mid N(t)=n\} \ ,$$

and

$$X_n = T_{n+1} - T_n \ ,$$

thus, in order to study the evolution of the system the two approaches are equivalent. However, in attempt to study the dependence structure and the evolution of the components the standard approach is not helpful.

Let us define the marginal transition probability for the embedded bivariate Markov chain $\mathbf{J}$ for $\mathbf{i} \in E^2$ and $j \in E$ as

$$\begin{aligned} \mathbb{P}(J^1_{n+1} = j_1 \mid J^1_n = i_i, J^2_{N^2(T^1_n)} = i_2) &=: p^1_{\mathbf{i},j_1} \\ \mathbb{P}(J^2_{n+1} = j_2 \mid J^1_{N^1(T^2_n)} = i_1, J^2_n = i_2) &=: p^2_{\mathbf{i},j_2} \ . \end{aligned} \tag{6.14}$$

Definition 6.1.11. ***Marginal cumulated semi-Markov kernel***

The marginal cumulated semi-Markov kernel is, for each component $\alpha = 1, 2$, the two-dimension matrix valued function $\mathbf{Q}^\alpha = (Q^\alpha_{\mathbf{i},j}(t),\ \text{for } t \in \mathbb{N})_{\mathbf{i}\in E^2, j\in E}$ defined by

$$Q^\alpha_{\mathbf{i},j}(t) := \mathbb{P}(J^\alpha_{N^\alpha(s)+1} = j, T^\alpha_{N^\alpha(s)+1} - T^\alpha_{N^\alpha(s)} \leq t \mid J^1_{N^1(s)} = i_1, J^2_{N^2(s)} = i_2), \tag{6.15}$$

for all $s \in \mathbb{N}$.

In the following, we will refer to the probability to have a transition exactly at a certain time, thus we introduce here the marginal semi-Markov kernel.

Definition 6.1.12. ***Marginal semi-Markov kernel***

The marginal semi-Markov kernel is, for each component $\alpha = 1, \ldots, \gamma$, the matrix valued function $\mathbf{q}^\alpha = (q^\alpha_{\mathbf{i},j}(t),\ \text{for } t \in \mathbb{N})_{\mathbf{i}\in E^\gamma, j\in E}$ defined by

$$q^\alpha_{\mathbf{i},j}(t) := \mathbb{P}(J^\alpha_{N^\alpha(s)+1} = j, T^\alpha_{N^\alpha(s)+1} - T^\alpha_{N^\alpha(s)} = t \mid \mathbf{J}_{\mathbf{N}(s)}) = \begin{cases} Q^\alpha_{\mathbf{i},j}(t+1) - Q^\alpha_{\mathbf{i},j}(t) & \text{for } t > 0 \\ 0 & \text{for } t = 0 \end{cases}, \tag{6.16}$$

for all $s \in \mathbb{N}$, where we used the time homogeneity of the process and we used the notation

$$\mathbf{J}_{\mathbf{N}}(s) = \begin{pmatrix} J^1_{N^1(s)} \\ J^2_{N^2(s)} \end{pmatrix} \ . \tag{6.17}$$

Lemma 6.1.13. *The marginal transition probabilities of the embedded bivariate Markov chain, Eq. (6.14), can be expressed in terms of the marginal semi-Markov kernel as*

$$p^\alpha_{\mathbf{i},j} = \sum_{t=0}^{\infty} q^\alpha_{\mathbf{i},j}(t) \ , \tag{6.18}$$

for every $\alpha = 1, \ldots, \gamma$, $\mathbf{i} \in E^\gamma$ and $j \in E$.

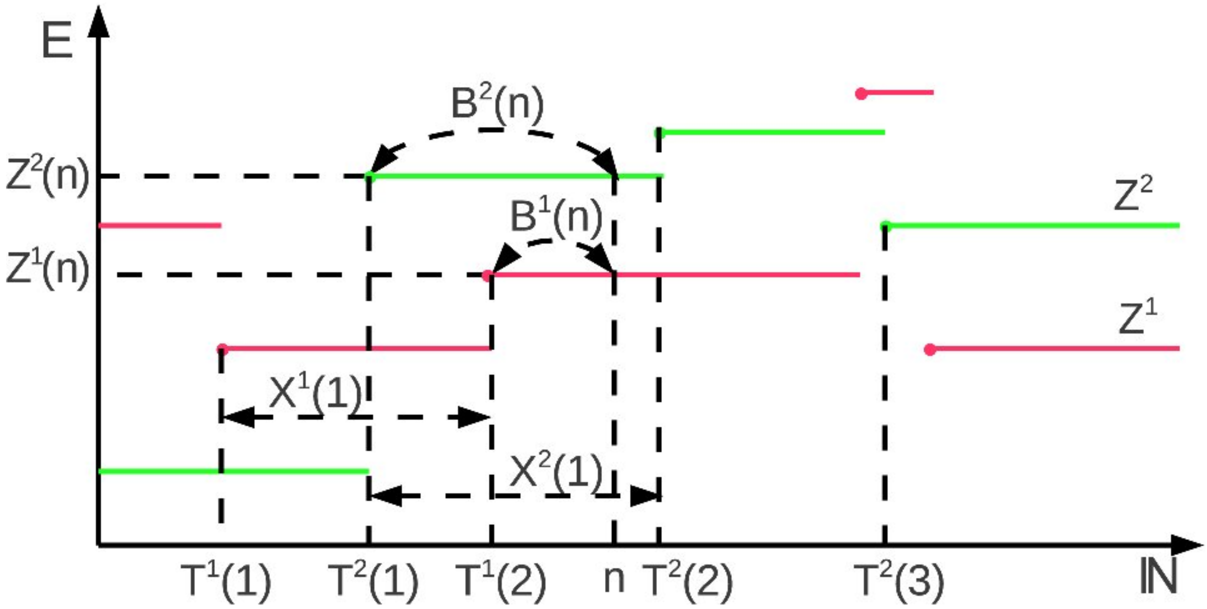


Figure 6.1: The trajectory of the double-component system is shown as a function of time. In the picture sojourn times, transition times and backward recurrence times are shown.

The evolution of the multivariate $(\mathbf{J}, \mathbf{X})$ process can be described as follows, given an initial state the next state occupied by the system is determined according to the evolution of the multivariate Markov chain while the sojourn time in the present state, of every part of the system, is determined according to the joint distribution of $\mathbf{X}$. In the bivariate case an example of trajectory is shown in Figure (6.1).

The marginal transition function of part α of the multivariate semi-Markov chain is defined by

$$\Phi^{\alpha}_{\mathbf{i},j}(t) := \mathbb{P}(Z^{\alpha}(t) = j \mid \mathbf{Z}(0) = \mathbf{i}, T^1_{N^1(0)} = 0, T^2_{N^2(0)} = 0), \tag{6.19}$$

for all $\mathbf{i} \in E^{\gamma}$, $j \in E$ and $t \in \mathbb{N}$. It gives the probability that component α at time t is in a state j given the state of the system at the present time 0.

In the semi-Markov evolution the Markovian memoryless property is preserved only at the transition times, i.e. the renewal moments. This feature makes the age of the state particularly important. As a consequence, the transition probabilities of a semi-Markov process change as a function of the values of the backward time, see for example D'Amico et al. [36].

Definition 6.1.14. ***Multivariate backward recurrence time process***

The multivariate backward recurrence time process associated to the multivariate semi-Markov process $\mathbf{Z}$*, denoted by* $(\mathbf{B}(\mathbf{t}))_{t\in\mathbb{N}} = (B^1(t), B^2(t))_{t\in\mathbb{N}})$*, is defined component by component as*

$$B^\alpha(t) := t - T^\alpha_{N^\alpha(t)} \qquad \text{for } \alpha = 1, 2,$$

its α *component specifies at any time* t *the age of the* α *component's state. In other words,* $B^\alpha(t)$ *gives the time since the last transition of* α*-th component.*

Definition 6.1.15. ***Marginal Backward Semi-Markov Kernel***

The backward marginal semi-Markov kernel is, for each component $\alpha = 1, 2$*, the matrix valued function* ${}^b\mathbf{q}^\alpha = ({}^b q^\alpha_{\mathbf{i},j}(v_\alpha, t),\ \text{for } t \in \mathbb{N}, v_\alpha \in \mathbb{N})_{\mathbf{i}\in E^\gamma, j\in E}$ *defined by*

$${}^b q^\alpha_{\mathbf{i},j}(v_\alpha, t) := \mathbb{P}(J^\alpha_{N^\alpha(s)+1} = j, X^\alpha_{N^\alpha(s)} = t + v_\alpha \mid \mathbf{J}_{\mathbf{N}(s)} = \mathbf{i}\ ,\ \mathbf{T}_{\mathbf{N}(s)} = s\mathbf{1} - \mathbf{v}\ ,\ \mathbf{T}_{\mathbf{N}(s)+\mathbf{1}} > s\mathbf{1})$$

where v_α *stands for the backward of component* α *at time* s*.*

Remark 6.1.16. *If we add to the semi-Markov process the information regarding the permanence into the states we obtain a Markov process. In other words,* $(\mathbf{Z}, \mathbf{B})$ *is a Markov process.*

Definition 6.1.17. ***Transition probability for*** $(\mathbf{Z}, \mathbf{B})$

The transition probability for $(\mathbf{Z}, \mathbf{B})$ *is defined by*

$$\tilde{P}_{\mathbf{i},\mathbf{j}}(\mathbf{v}, \mathbf{w}) := \mathbb{P}(\mathbf{Z}(t+1) = \mathbf{j}, \mathbf{B}(t+1) = \mathbf{w} \mid \mathbf{Z}(t) = \mathbf{i}, \mathbf{B}(t) = \mathbf{v})\ . \tag{6.20}$$

The transition probability does not depend on t *due to the homogeneity of the process.*

Let us denote the marginal transition function for the α-th component by

$${}^b\Phi^\alpha_{\mathbf{i};j_\alpha}(\mathbf{v}; u_\alpha, t) := \mathbb{P}(Z^\alpha(t) = j_\alpha, B^\alpha(t) = u_\alpha \mid \mathbf{Z}(0) = \mathbf{i}\ ,\ \mathbf{B}(0) = \mathbf{v}), \tag{6.21}$$

it gives the probability of the event $\{Z^1(t) = j_1, B^1(t) = u_1\}$, for all $j_1 \in E$ and $u_1 \in \mathbb{N}$, given that $\mathbf{Z}(0) = \mathbf{i}$ be the initial state of the system with backward recurrence time $\mathbf{B}(0) = \mathbf{v}$.

In order to study the evolution equation for the marginal transition function we make some assumption on the waiting times distribution.

6.1.1 Bivariate Semi-Markov Chain with Independent Waiting Times

Assumption 6.1.18. *A2 The sequences of random variables $(X_n^\alpha)_{n\in\mathbb{N}}$, for $\alpha = 1, 2$, are independent. To be more precise the σ algebras generated by $(X^\alpha)_{\alpha=1,2}$ are independent. In other words the sojourn times of the components do not influence each other.*

The sequence of sojourn times $(X_n^{(\alpha)})_{n\in\mathbb{N}}$ depends only on the visited state of the α-component J_n^α. Moreover, the sojourn time of component α depends only on its present state. Thus, the marginal unconditional cumulative distribution function of the waiting times for component α can be defined as

$$\begin{aligned}&\mathbb{P}(X^\alpha_{N^\alpha(s)} \le t \mid \sigma(J^\beta_{N^\beta(h)}), h \le s, \beta = 1, 2) = \mathbb{P}(X^\alpha_{N^\alpha(s)} \le t \mid \sigma(J^\alpha_{N^\alpha(h)}), h \le s)\\ &= \mathbb{P}(X^\alpha_{N^\alpha(s)} \le t \mid J^\alpha_{N^\alpha(s)} = i_\alpha) = \mathbb{P}(X^\alpha_n \le k \mid J^\alpha_n = i_\alpha) =: H^\alpha_{i_\alpha}(t) \qquad \forall\ s \in \mathbb{N},\end{aligned} \tag{6.22}$$

where $\sigma(J^\beta_{N^\beta(h)}, h \le s, \beta = 1, 2)$ is the natural filtration of the multivariate Markov chain $(\mathbf{J})$.

Remark 6.1.19. *A simplifying assumption could be that of equally distributed sojourn times for each component, i.e.*

$$H_i^1(k) = H_i^2(k) \qquad \forall\ k \in \mathbb{N}, i \in E\ . \tag{6.23}$$

This simplification could be adopted if data are not sufficient to estimate different cdf.

Remark 6.1.20. *Using assumptions A1 and A2, we have*

$$\mathbb{P}(J^\alpha_{N^\alpha(s)+1} = j, T^\alpha_{N^\alpha(s)+1} - T^\alpha_{N^\alpha(s)} \le t \mid (\mathbf{J}_{\mathbf{N}(h)}, \mathbf{T}_{\mathbf{N}(h)}, h \le s)) = \mathbb{P}(J^\alpha_{N^\alpha(s)+1} = j, X^\alpha_{N^\alpha(s)} \le t \mid \mathbf{J}_{\mathbf{N}(s)} = \mathbf{i}),$$

that is T^α, for $\alpha = 1, 2$, can be interpreted as the renewal times of component α.

Lemma 6.1.21. *The marginal cumulated semi-Markov kernel can be expressed as*

$$Q^\alpha_{\mathbf{i},j}(t) = p^\alpha_{\mathbf{i},j} F^\alpha_{i_\alpha}(t)\ . \tag{6.24}$$

Proposition 6.1.22. *Let $\mathbf{Z}$ be a semi-Markov chain satisfying assumptions A1 and A2. Then, the marginal backward semi-Markov kernel for component α can be expressed as*

$${}^b q^\alpha_{\mathbf{i},j}(v_\alpha, t) = \frac{H^\alpha_{i_\alpha}(t + v_\alpha) - H^\alpha_{i_\alpha}(t + v_\alpha - 1)}{1 - H^\alpha_{i_\alpha}(v_\alpha)} \cdot p^\alpha_{\mathbf{i},j}\ . \tag{6.25}$$

Proof. By applying Bayes rules we get

$$\begin{aligned} & \mathbb{P}(J^{\alpha}_{N^{\alpha}(s)+1}=j, X^{\alpha}_{N^{\alpha}(s)}=t+v_{\alpha} \mid \mathbf{J}_{\mathbf{N}(s)}=\mathbf{i}\ ,\ \mathbf{T}_{\mathbf{N}(s)}=s-\mathbf{v}\ ,\ \mathbf{T}_{\mathbf{N}(s)+\mathbf{1}}>s) \\ = & \ \mathbb{P}(X^{\alpha}_{N^{\alpha}(s)}=t+v_{\alpha} \mid J^{\alpha}_{N^{\alpha}(s)+1}=j\ ,\ \mathbf{J}_{\mathbf{N}(s)}=\mathbf{i}\ ,\ \mathbf{T}_{\mathbf{N}(s)}=s-\mathbf{v}\ ,\ \mathbf{T}_{\mathbf{N}(s)+\mathbf{1}}>s) \\ \times & \ \mathbb{P}(J^{\alpha}_{N^{\alpha}(s)+1}=j \mid \mathbf{J}_{\mathbf{N}(s)}=\mathbf{i}\ ,\ \mathbf{T}_{\mathbf{N}(s)}=s-\mathbf{v}\ ,\ \mathbf{T}_{\mathbf{N}(s)+\mathbf{1}}>s). \end{aligned} \tag{6.26}$$

Then, by using assumptions A1 and A2 we obtain

$$\begin{aligned} & \mathbb{P}(J^{\alpha}_{N^{\alpha}(s)+1}=j, X^{\alpha}_{N^{\alpha}(s)}=t+v_{\alpha} \mid \mathbf{J}_{\mathbf{N}(s)}=\mathbf{i}\ ,\ \mathbf{T}_{\mathbf{N}(s)}=s-\mathbf{v}\ ,\ \mathbf{T}_{\mathbf{N}(s)+\mathbf{1}}>s) \\ = & \ \mathbb{P}(X^{\alpha}_{N^{\alpha}(s)}=t+v_{\alpha} \mid J^{\alpha}_{N^{\alpha}(s)}=i_{\alpha}\ ,\ T^{\alpha}_{N^{\alpha}(s)}=s-v_{\alpha}\ ,\ T^{\alpha}_{\mathbf{N}(s)+\mathbf{1}}>s) \\ \times & \ \mathbb{P}(J^{\alpha}_{N^{\alpha}(s)+1}=j \mid \mathbf{J}_{\mathbf{N}(s)}=\mathbf{i}), \end{aligned} \tag{6.27}$$

and by using the definitions we get

$$\begin{aligned} & \mathbb{P}(J^{\alpha}_{N^{\alpha}(s)+1}=j, X^{\alpha}_{N^{\alpha}(s)}=t+v_{\alpha} \mid \mathbf{J}_{\mathbf{N}(s)}=\mathbf{i}\ ,\ \mathbf{T}_{\mathbf{N}(s)}=s-\mathbf{v}\ ,\ \mathbf{T}_{\mathbf{N}(s)+\mathbf{1}}>s) \\ = & \ \frac{H^{\alpha}_{i_{\alpha}}(t+v_{\alpha})-H^{\alpha}_{i_{\alpha}}(t+v_{\alpha}-1)}{1-H^{\alpha}_{i_{\alpha}}(v_{\alpha})}\cdot p^{\alpha}_{\mathbf{i},j} = {}^{b}q^{\alpha}_{\mathbf{i},j}(v_{\alpha},t)\ . \end{aligned} \tag{6.28}$$

□

We notice that, if the backward process is zero $B^{\alpha}(s)=0$ $(v_{\alpha}=0)$, as we expect the backward kernel coincides with the kernel, that is

$${}^{b}q^{\alpha}_{\mathbf{i},j}(0,t)=q^{\alpha}_{\mathbf{i},j}(t).$$

Remark 6.1.23. *Formula (6.25) reveals that assumptions A1 and A2 imply that backward semi-Markov kernels* $({}^{b}\mathbf{q}^{\alpha})_{\alpha=1,2}$ *are affected by the permanence in the state only of their own component. That is, the backward values of the other parts of the system does not affect the kernel. However, we will see that transition functions will depend on the backward values of each part of the system.*

Proposition 6.1.24. *Let* $\mathbf{Z}$ *be a bivariate semi-Markov process satisfying assumptions A1 and A2. Then,* *the one-step transition probability for* $(\mathbf{Z},\mathbf{B})$ *can be expressed as*

$$\tilde{P}_{\mathbf{i},\mathbf{j}}(\mathbf{v},\mathbf{w}) = \begin{cases} \left(\frac{1-H^1_{i_1}(1+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(1+v_2)}{1-H^2_{i_2}(v_2)}\right) & \text{if } \mathbf{l}=\mathbf{i}, \mathbf{w}=\mathbf{v}+\mathbf{1} \\ {}^b q^1_{\mathbf{i},j_1}(v_1,1)\left(\frac{1-H^2_{i_2}(1+v_2)}{1-H^2_{i_2}(v_2)}\right) & \text{if } j_1\neq i_1, j_2=i_2, \mathbf{w}=\binom{0}{1+v_2} \\ {}^b q^2_{\mathbf{i},j_2}(v_2,1)\left(\frac{1-H^1_{i_1}(1+v_1)}{1-H^1_{i_1}(v_1)}\right) & \text{if } j_1=i_1, j_2\neq i_2, \mathbf{w}=\binom{1+v_1}{0} \\ \left(\frac{H^1_{i_1}(1+v_1)-H^1_{i_1}(v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{H^2_{i_2}(1+v_2)-H^2_{i_2}(v_2)}{1-H^2_{i_2}(v_2)}\right)P_{\binom{i^1}{i^2}\binom{j^1}{j^2}} & \text{if } j_1\neq i_1, j_2\neq i_2, \mathbf{w}=\binom{0}{0} \\ 0 & \text{otherwise} \end{cases} \tag{6.29}$$

Proof. The backward time process, for each component, takes value in $\mathbb{N}$ so $\tilde{P}$ has in principle infinite entries, however, most of them are null. In fact, the next step backward time of component α can assume only two values depending on whether there is a transition on the next step, backward time null, or there is no transition, backward time increasing of one. These two possibilities partition the trajectories of our bivariate system in 4 parts. These parts can be obtained directly writing the one step transition probability as

$$\tilde{P}_{\mathbf{i},\mathbf{j}}(\mathbf{v},\mathbf{w}) = \begin{cases} \mathbb{P}(\mathbf{Z}(1)=\mathbf{i}, \mathbf{B}(1)=\mathbf{v}+\mathbf{1} \mid \mathbf{Z}(0)=\mathbf{i}, \mathbf{B}(0)=\mathbf{v}) \\ \mathbb{P}(\mathbf{Z}(1)=\binom{j_1}{i_2}, \mathbf{B}(1)=\binom{0}{1+v_2} \mid \mathbf{Z}(0)=\mathbf{i}, \mathbf{B}(0)=\mathbf{v}) \\ \mathbb{P}(\mathbf{Z}(1)=\binom{i_1}{j_2}, \mathbf{B}(1)=\binom{1+v_1}{0} \mid \mathbf{Z}(0)=\mathbf{i}, \mathbf{B}(0)=\mathbf{v}) \\ \mathbb{P}(\mathbf{Z}(1)=\binom{j_1}{j_2}, \mathbf{B}(1)=\binom{0}{0} \mid \mathbf{Z}(0)=\mathbf{i}, \mathbf{B}(0)=\mathbf{v}) \end{cases}, \tag{6.30}$$

here we suppose $i_1 \neq j_1$ and $i_2 \neq j_2$. The first probability on the r.h.s of Eq. (6.30) can be expressed as

$$\mathbb{P}(\mathbf{Z}(1)=\mathbf{i}, \mathbf{B}(1)=\mathbf{v}+\mathbf{1} \mid \mathbf{Z}(0)=\mathbf{i}, \mathbf{B}(0)=\mathbf{v}) = \left(\frac{1-H^1_{i_1}(1+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(1+v_2)}{1-H^2_{i_2}(v_2)}\right), \tag{6.31}$$

it is the probability that no transition occurs.

The second probability on the r.h.s of Eq. (6.30) can be written as

$$\mathbb{P}(\mathbf{Z}(1)=\binom{j_1}{i_2}, \mathbf{B}(1)=\binom{0}{1+v_2} \mid \mathbf{Z}(0)=\mathbf{i}, \mathbf{B}(0)=\mathbf{v}) = {}^b q^1_{\mathbf{i},j_1}(v_1,1)\left(\frac{1-H^2_{i_2}(1+v_2)}{1-H^2_{i_2}(v_2)}\right), \tag{6.32}$$

it is the probability that only component 1 makes a transition. Similarly the third, i.e. the probability that only component 2 makes a transition, can be represented as

$$\mathbb{P}(\mathbf{Z}(1)=\binom{i_1}{j_2}, \mathbf{B}(1)=\binom{1+v_1}{0} \mid \mathbf{Z}(0)=\mathbf{i}, \mathbf{B}(0)=\mathbf{v}) = {}^b q^2_{\mathbf{i},j_2}(v_2,1)\left(\frac{1-H^1_{i_1}(1+v_1)}{1-H^1_{i_1}(v_1)}\right). \tag{6.33}$$

Finally the fourth part gives the probability of a simultaneous transition of both components and it can be expressed as

$$\mathbb{P}(\mathbf{Z}(1)=\binom{j_1}{j_2}, \mathbf{B}(1)=\binom{0}{0} \mid \mathbf{Z}(0)=\mathbf{i}, \mathbf{B}(0)=\mathbf{v}) = \frac{H^1_{i_1}(1+v_1)-H^1_{i_1}(v_1)}{1-H^1_{i_1}(v_1)}\,\frac{H^2_{i_2}(1+v_2)-H^2_{i_2}(v_2)}{1-H^2_{i_2}(v_2)}P_{\binom{i^1}{i^2}\binom{j^1}{j^2}},$$

here **P** is the transition matrix of the embedded bivariate Markov chain. □

Assumption 6.1.25. *A3 The bivariate process* $\mathbf{J} = (J_n^1, J_n^2)_{n\in\mathbb{N}}$ *is such that the next visited state of each component depends only on the present state of the system. The marginal one step transition probabilities satisfy*

$$\mathbb{P}(J_{n+1}^1 = j_1 \mid \sigma(J_n^1, h < n+1), \sigma(J_h^2, h \leq N^2(T_{n+1}^1))) = \mathbb{P}(J_{n+1}^1 = j_1 \mid J_n^1 = i_i, J_{N^2(T_n^1)}^2 = i_2) =: p_{\mathbf{i},j_1}^1$$

$$\mathbb{P}(J_{n+1}^2 = j_2 \mid \sigma(J_h^1, h \leq N^1(T_{n+1}^2)), \sigma(J_h^2, h < n+1)) = \mathbb{P}(J_{n+1}^2 = j_2 \mid J_{N^1(T_n^2)}^1 = i_1, J_n^2 = i_2) =: p_{\mathbf{i},j_2}^2$$

where $\sigma(J_h^\alpha, h < n+1)$ *denotes the natural filtration of* J^α.

Remark 6.1.26. *The bivariate semi-Markov chain satisfying assumptions A1, A2 and A3 is not a multivariate process with independent components. Indeed, the waiting times are independent (see assumption A2) but the sequences of visited states are dependent. This is outlined by assumption A3 stating that the next visited state of component* α *depends on its present state and on the present states of all the other components.*

Theorem 6.1.27. *Let* $\mathbf{Z}$ *be a bivariate semi-Markov process satisfying assumptions* A1, A2 *and* A3. *Let* $\mathbf{B}$ *be the associated backward recurrence time process.*
The transition probability function for component 1 *can be expressed, for all* $\mathbf{i}, \mathbf{j} \in E^2$, $\mathbf{v}, \mathbf{u} \in \mathbb{N}^2$ *and* $t \in \mathbb{N}$, *as follows*

$$\begin{aligned}
{}^b\Phi_{\mathbf{i};j_1}^1(\mathbf{v}; u_1, t) &= \left(\frac{1 - H_{i_1}^1(t+v_1)}{1 - H_{i_1}^1(v_1)}\right)\left(\frac{1 - H_{i_2}^2(t+v_2)}{1 - H_{i_2}^2(v_2)}\right)\delta_{i_1,j_1}\delta_{u_1,t+v_1} \\
&+ \sum_{l_1\in E}\sum_{s=1}^{t} q_{\mathbf{i},l_1}^1(v_1, s)\left(\frac{1 - H_{i_2}^2(s+v_2)}{1 - H_{i_2}^2(v_2)}\right){}^b\Phi_{\binom{l_1}{i_2};j_1}^1\left(\binom{0}{v_2+s}; u_1, t-s\right) \\
&+ \sum_{l_2\in E}\sum_{s=1}^{t} q_{\mathbf{i},l_2}^2(v_2, s)\left(\frac{1 - H_{i_1}^1(s+v_1)}{1 - H_{i_1}^1(v_1)}\right){}^b\Phi_{\binom{i_1}{l_2};j_1}^1\left(\binom{v_1+s}{0}; u_1, t-s\right) \\
&+ \sum_{l_1,l_2\in E}\sum_{s=1}^{t} {}^b q_{\mathbf{i},l_1}^1(v_1, s)\,{}^b q_{\mathbf{i},l_2}^2(v_2, s)\,{}^b\Phi_{\binom{l_1}{l_2};j_1}^1\left(\binom{0}{0}; u_1, t-s\right),
\end{aligned} \tag{6.34}$$

and similarly for component 2 we have

$$
\begin{aligned}
{}^{b}\Phi^{2}_{\mathbf{i};j_2}(\mathbf{v};u_2,t) &= \left(\frac{1-H^1_{i_1}(t+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(t+v_2)}{1-H^2_{i_2}(v_2)}\right)\delta_{i_2,j_2}\delta_{u_2,t+v_2} \\
&+ \sum_{l_1\in E}\sum_{s=1}^{t} q^1_{\mathbf{i},l_1}(v_1,s)\left(\frac{1-H^2_{i_2}(s+v_2)}{1-H^2_{i_2}(v_2)}\right){}^{b}\Phi^{2}_{\binom{l_1}{i_2};j_2}\left(\tbinom{0}{v_2+s};u_2,t-s\right) \\
&+ \sum_{l_2\in E}\sum_{s=1}^{t} q^2_{\mathbf{i},l_2}(v_2,s)\left(\frac{1-H^1_{i_1}(s+v_1)}{1-H^1_{i_1}(v_1)}\right){}^{b}\Phi^{2}_{\binom{i_1}{l_2};j_2}\left(\tbinom{v_1+s}{0};u_2,t-s\right) \\
&+ \sum_{l_1,l_2\in E}\sum_{s=1}^{t} {}^{b}q^1_{\mathbf{i},l_1}(v_1,s)\,{}^{b}q^2_{\mathbf{i},l_2}(v_2,s)\,{}^{b}\Phi^{2}_{\binom{l_1}{l_2};j_2}\left(\tbinom{0}{0};u_2,t-s\right) .
\end{aligned}
\tag{6.35}
$$

Proof. In order to show the result we first notice that, for all $j_1 \in E$ and $u_1 \in \mathbb{N}$, the event $\{Z^1(t) = j_1, B^1(t) = u_1\}$, given that the initial state is $\mathbf{Z}(0) = \mathbf{i}$ with backward recurrence time $\mathbf{B}(0) = \mathbf{v}$, can be obtained by different possible trajectories of the system, these can be divided in the following way

$$
\begin{aligned}
\mathbb{I}_{\{Z^1(t)=j_1,B^1(t)=u_1|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}} &\overset{d}{=} \mathbb{I}_{\{T^1_1>t,T^2_1>t|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\delta_{i_1,j_1}\delta_{u_1,t+v_1} \\
&+ \sum_{l_1\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T^1_1=s,J^1_1=l_1,T^2_1>s|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^1(t)=j_1,B^1(t)=u_1\middle|\mathbf{Z}(s)=\binom{l_1}{i_2},\mathbf{B}(s)=\binom{0}{v_2+s}\right\}} \\
&+ \sum_{l_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T^1_1>s,T^2_1=s,J^2_1=l_2|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^1(t)=j_1,B^1(t)=u_1\middle|\mathbf{Z}(s)=\binom{i_1}{l_2},\mathbf{B}(s)=\binom{v_1+s}{0}\right\}} \\
&+ \sum_{l_1,l_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T^1_1=s,J^1_1=l_1,T^2_1=s,J^2_1=l_2|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^1(t)=j_1,B^1(t)=u_1\middle|\mathbf{Z}(s)=\binom{l_1}{l_2},\mathbf{B}(s)=\binom{0}{0}\right\}},
\end{aligned}
\tag{6.36}
$$

where δ denotes the Kronecker delta function, the upper script d on the equal sign denotes the equality in distribution, and on the second term of the r.h.s we used the following relation

$$
\begin{aligned}
&\mathbb{P}\{Z^1(t)=j_1,B^1(t)=u_1|T^1_1=s,J^1_1=l_1,T^2_1>s,\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\} \\
&= \mathbb{P}\left\{Z^1(t)=j_1,B^1(t)=u_1|\mathbf{Z}(s)=\tbinom{l_1}{i_2},\mathbf{B}(s)=\tbinom{0}{v_2+s},\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\right\} \\
&= \mathbb{P}\left\{Z^1(t)=j_1,B^1(t)=u_1|\mathbf{Z}(s)=\tbinom{l_1}{i_2},\mathbf{B}(s)=\tbinom{0}{v_2+s}\right\},
\end{aligned}
\tag{6.37}
$$

here we applied $(\mathbf{Z}, \mathbf{B})$ as a Markov process. A similar relation has been used for the third and fourth term on the r.h.s of Eq. (6.36).

The relation (6.36) divides the trajectories into four parts, the first takes into account the events with no transition in t steps, the second and third consider the trajectories where the first transition is made by the first and second component respectively, finally the fourth possibility considers the first transition of the first and second component simultaneously. A similar relation can be obtained for the event $\{Z^2(t) = j_2, B^2(t) =$

$u_2\}$, for all $j_2 \in E$ and $u_2 \in \mathbb{N}$, at t time steps given that $(\mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v})$ is the initial state, that is

$$\begin{aligned}
&\mathbb{I}_{\{Z^2(t)=j_2,B^2(t)=u_2|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}} \stackrel{d}{=} \mathbb{I}_{\{T_1^1>t,T_1^2>t|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\delta_{i_2,j_2}\delta_{u_2,t+v_2} \\
&+\sum_{l_1\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=l_1,T_1^2>s|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^2(t)=j_2,B^2(t)=u_2\middle|\mathbf{Z}(s)=\binom{l_1}{i_2},\mathbf{B}(s)=\binom{0}{v_2+s},\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\right\}} \\
&+\sum_{l_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1>s,T_1^2=s,J_1^2=l_2|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^2(t)=j_2,B^2(t)=u_2\middle|\mathbf{Z}(s)=\binom{i_1}{l_2},\mathbf{B}(s)=\binom{v_1+s}{0},\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\right\}} \\
&+\sum_{l_1,l_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=l_1,T_1^2=s,J_1^2=l_2|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^2(t)=j_2,B^2(t)=u_2\middle|\mathbf{Z}(s)=\binom{l_1}{l_2},\mathbf{B}(s)=\binom{0}{0},\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\right\}}.
\end{aligned} \tag{6.38}$$

The transition probabilities for component 1 and 2 can be obtained by taking the conditional expectation on both sides of Eqs. (6.36) and (6.38), respectively. The structures of these two equations is similar, therefore we will explicitly evaluate only the expected value of Eq. (6.36) for component 1. Similar calculations can be done for the expression of the transition probability function for component 2.

We have to evaluate the expectation of four terms on the r.h.s of Eq. (6.36). In the first term the expectation of the indicator function is given by

$$\begin{aligned}
&\mathbb{E}\{\mathbb{I}_{\{T_1^1>t,T_1^2>t|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\} = \mathbb{P}\{T_1^1 > t, T_1^2 > t|\mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}\} \\
&= \mathbb{P}\{X_1^1 > t + v_1|X_1^1 > v_1, J_0^1 = i_1\}\mathbb{P}\{X_1^2 > t + v_2|X_1^2 > v_2, J_0^2 = i_2\} \\
&= \left(\frac{1 - H_{i_1}^1(t + v_1)}{1 - H_{i_1}^1(v_1)}\right)\left(\frac{1 - H_{i_2}^2(t + v_2)}{1 - H_{i_2}^2(v_2)}\right),
\end{aligned} \tag{6.39}$$

where we used the independence between the waiting times of the components.

The second term on the r.h.s of Eq. (6.36) is a sum of terms that are products of indicator functions. By linearity property of the expectation we can directly evaluate the expectation of the single terms of the sum. Using the properties of conditional expectation, the expected value of the generic summand is

$$\begin{aligned}
&\mathbb{E}\left\{\mathbb{I}_{\{T_1^1=s,J_1^1=l_1,T_1^2>s|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^1(t)=j_1,B^1(t)=u_1\middle|\mathbf{Z}(s)=\binom{l_1}{i_2},\mathbf{B}(s)=\binom{0}{v_2+s}\right\}}\middle|\mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}\right\} \\
&= \mathbb{P}\{T_1^1 = s, J_1^1 = l_1, T_1^2 > s|\mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}\}\mathbb{P}\left\{Z^1(t) = j_1, B^1(t) = u_1\middle|\mathbf{Z}(s) = \binom{l_1}{i_2}, \mathbf{B}(s) = \binom{0}{v_2+s}\right\},
\end{aligned} \tag{6.40}$$

where we used the Markov property of the process $(\mathbf{Z}, \mathbf{B})$. Now, by time homogeneity of the process

$$\begin{aligned}
&\mathbb{P}\left\{Z^1(t) = j_1, B^1(t) = u_1\middle|\mathbf{Z}(s) = \binom{l_1}{i_2}, \mathbf{B}(s) = \binom{0}{v_2+s}\right\} \\
&= \mathbb{P}\left\{Z^1(t-s) = j_1, B^1(t-s) = u_1\middle|\mathbf{Z}(0) = \binom{l_1}{i_2}, \mathbf{B}(0) = \binom{0}{v_2+s}\right\} = {}^{b}\Phi^1_{\binom{l_1}{i_2};j_1}(\binom{0}{v_2+s}; u_1, t-s),
\end{aligned} \tag{6.41}$$

and using the hypothesis $A1$ and $A2$

$$\begin{aligned}
&\mathbb{P}\{T_1^1 = s, J_1^1 = l_1, T_1^2 > s | \mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}\} \\
&= \mathbb{P}\{J_1^1 = l_1 | \mathbf{J}_0 = \mathbf{i}\}\mathbb{P}\{T_1^1 = s | X_1^1 > v_1, J_0^1 = i_1\}\mathbb{P}\{X_1^2 > s + v_2 | X_1^2 > v_2, J_0^2 = i_2\} \\
&= p_{\mathbf{i},l_1}^1 \left(\frac{H_{i_1}^1(s+v_1) - H_{i_1}^1(s+v_1-1)}{1-H_{i_1}^1(v_1)}\right)\left(\frac{1-H_{i_2}^2(s+v_2)}{1-H_{i_2}^2(v_2)}\right) = {}^b q_{\mathbf{i},l_1}^1(v_1,s)\left(\frac{1-H_{i_2}^2(s+v_2)}{1-H_{i_2}^2(v_2)}\right) .
\end{aligned} \tag{6.42}$$

Here the last equality is obtained using Eq. (6.25). Finally, by substituting we obtain

$$\begin{aligned}
&\mathbb{E}\left\{\mathbb{I}_{\{T_1^1=s, J_1^1=l_1, T_1^2>s|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^1(t)=j_1, B^1(t)=u_1 \middle| \mathbf{Z}(s)=\binom{l_1}{i_2}, \mathbf{B}(s)=\binom{0}{v_2+s}\right\}}\middle| \mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\right\} \\
&= {}^b q_{\mathbf{i},l_1}^1(v_1,s)\left(\frac{1-H_{i_2}^2(s+v_2)}{1-H_{i_2}^2(v_2)}\right){}^b\Phi^1_{\binom{l_1}{i_2};j_1}(\tbinom{0}{v_2+s}; u_1, t-s).
\end{aligned} \tag{6.43}$$

The expectation of the third term on the r.h.s of Eq. (6.36) can be evaluated in the same way. The fourth term on the r.h.s of Eq. (6.36) is a sum as well, but now the generic summand takes into account for simultaneous transitions of the components and its conditional expectation is

$$\begin{aligned}
&\mathbb{E}\left\{\mathbb{I}_{\{T_1^1=s, J_1^1=l_1, T_1^2=s, J_1^2=l_2|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^1(t)=j_1, B^1(t)=u_1 \middle| \mathbf{Z}(s)=\binom{l_1}{l_2}, \mathbf{B}(s)=\binom{0}{0}\right\}}\middle| \mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\right\} \\
&= \mathbb{P}\{T_1^1 = s, J_1^1 = l_1, T_1^2 = s, J_1^2 = l_2 | \mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}\}\mathbb{P}\left\{Z^1(t) = j_1, B^1(t) = u_1 \middle| \mathbf{Z}(s) = \tbinom{l_1}{l_2}, \mathbf{B}(s) = \tbinom{0}{0}\right\} .
\end{aligned} \tag{6.44}$$

Here by time homogeneity we have

$$\mathbb{P}\left\{Z^1(t) = j_1, B^1(t) = u_1 \middle| \mathbf{Z}(s) = \tbinom{l_1}{l_2}, \mathbf{B}(s) = \tbinom{0}{0}\right\} = {}^b\Phi^1_{\binom{l_1}{l_2};j_1}(\tbinom{0}{0}; u_1, t-s), \tag{6.45}$$

and using hypothesis $A1$ and $A2$ we can express this probability as

$$\begin{aligned}
&\mathbb{P}\{T_1^1 = s, J_1^1 = l_1, T_1^2 = s, J_1^2 = l_2 | \mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}\} \\
&= \mathbb{P}\{J_1^1 = l_1 | \mathbf{J}_0 = \mathbf{i}\}\mathbb{P}\{J_1^2 = l_2 | \mathbf{J}_0 = \mathbf{i}\}\mathbb{P}\{T_1^1 = s | X_1^1 > v_1, J_0^1 = i_1\}\mathbb{P}\{T_1^2 = s | X_1^2 > v_2, J_0^2 = i_2\} \\
&= p_{\mathbf{i},l_1}^1 p_{\mathbf{i},l_2}^2 \left(\frac{H_{i_1}^1(s+v_1) - H_{i_1}^1(s+v_1-1)}{1-H_{i_1}^1(v_1)}\right)\left(\frac{H_{i_2}^2(s+v_2) - H_{i_2}^2(s+v_2-1)}{1-H_{i_2}^2(v_2)}\right) \\
&= {}^b q_{\mathbf{i},l_1}^1(v_1,s){}^b q_{\mathbf{i},l_2}^2(v_2,s) .
\end{aligned} \tag{6.46}$$

By substituting we get

$$\begin{aligned}
&\mathbb{E}\left\{\mathbb{I}_{\{T_1^1=s, J_1^1=l_1, T_1^2=s, J_1^2=l_2|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\}}\mathbb{I}_{\left\{Z^1(t)=j_1, B^1(t)=u_1 \middle| \mathbf{Z}(s)=\binom{l_1}{l_2}, \mathbf{B}(s)=\binom{0}{0}\right\}}\middle| \mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\right\} \\
&= {}^b q_{\mathbf{i},l_1}^1(v_1,s){}^b q_{\mathbf{i},l_2}^2(v_2,s){}^b\Phi^1_{\binom{l_1}{l_2};j_1}(\tbinom{0}{0}; u_1, t-s).
\end{aligned} \tag{6.47}$$

The result is obtained by substituting all the terms in Eq. (6.36). □

We can express the evolution equations for the transition probability functions of component 1 and 2 in a more compact form. To this end, let us define the vector

$$ {}^b\Phi_{\mathbf{i};\mathbf{j}}(\mathbf{v};\mathbf{u},t) = \begin{pmatrix} {}^b\Phi^1_{\mathbf{i};j_1}(\mathbf{v};u_1,t) \\ {}^b\Phi^2_{\mathbf{i};j_2}(\mathbf{v};u_2,t) \end{pmatrix} \tag{6.48} $$

whose components are the transition probability for the two components, and let us introduce the following notation for the Kronecker symbol on the states

$$ \delta_{\mathbf{i},\mathbf{j}} = \begin{pmatrix} \delta_{i_1,j_1} \\ \delta_{i_2,j_2} \end{pmatrix} \tag{6.49} $$

and similarly for the one on backward values. The vector Φ defined in Eq. (6.48), represents the transition probabilities for all the system. In terms of Φ the recursive formulas (6.34) and (6.35) for the two components can be represented as

$$ \begin{aligned} {}^b\Phi_{\mathbf{i};\mathbf{j}}(\mathbf{v};\mathbf{u},t) &= \left(\frac{1-H^1_{i_1}(t+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(t+v_2)}{1-H^2_{i_2}(v_2)}\right)\delta_{\mathbf{i},\mathbf{j}}\delta_{\mathbf{u},\mathbf{v}+\mathbf{1}t} \\ &+ \sum_{l_1\in E}\sum_{s=1}^{t} {}^bq^1_{\mathbf{i},l_1}(v_1,s)\left(\frac{1-H^2_{i_2}(s+v_2)}{1-H^2_{i_2}(v_2)}\right){}^b\Phi_{\binom{l_1}{i_2};\mathbf{j}}\left(\binom{0}{v_2+s};\mathbf{u},t-s\right) \\ &+ \sum_{l_2\in E}\sum_{s=1}^{t} {}^bq^2_{\mathbf{i},l_2}(v_2,s)\left(\frac{1-H^1_{i_1}(s+v_1)}{1-H^1_{i_1}(v_1)}\right){}^b\Phi_{\binom{i_1}{l_2};\mathbf{j}}\left(\binom{v_1+s}{0};\mathbf{u},t-s\right) \\ &+ \sum_{l_1,l_2\in E}\sum_{s=1}^{t} {}^bq^1_{\mathbf{i},l_1}(v_1,s)\,{}^bq^2_{\mathbf{i},l_2}(v_2,s)\,{}^b\Phi_{\binom{l_1}{l_2};\mathbf{j}}\left(\binom{0}{0};\mathbf{u},t-s\right), \end{aligned} \tag{6.50} $$

where $\mathbf{1}$ is the bidimensional vector with the two components equal to one.

Lemma 6.1.28. *Let* $\mathbf{Z}$ *be a bivariate semi-Markov process satisfying assumptions A1, A2 and A3. Then, the one-step transition probability for* $(\mathbf{Z},\mathbf{B})$ *can be expressed as*

$$ \tilde{P}_{\mathbf{i},\mathbf{j}}(\mathbf{v},\mathbf{w}) = \begin{cases} \left(\frac{1-H^1_{i_1}(1+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(1+v_2)}{1-H^2_{i_2}(v_2)}\right) & \textit{if } \mathbf{l}=\mathbf{i}, \mathbf{w}=\mathbf{v}+\mathbf{1} \\ {}^bq^1_{\mathbf{i},j_1}(v_1,1)\left(\frac{1-H^2_{i_2}(1+v_2)}{1-H^2_{i_2}(v_2)}\right) & \textit{if } j_1\neq i_1, j_2=i_2, \mathbf{w}=\binom{0}{1+v_2} \\ {}^bq^2_{\mathbf{i},j_2}(v_2,1)\left(\frac{1-H^1_{i_1}(1+v_1)}{1-H^1_{i_1}(v_1)}\right) & \textit{if } j_1 = i_1, j_2\neq i_2, \mathbf{w}=\binom{1+v_1}{0} \\ {}^bq^1_{\mathbf{i},j_1}\,{}^bq^2_{\mathbf{i},j_2} & \textit{if } j_1\neq i_1, j_2\neq i_2, \mathbf{w}=\binom{0}{0} \\ 0 & \textit{otherwise} \end{cases} . \tag{6.51} $$

Proof. The result can be obtained by Proposition 6.1.24 using assumption A3. □

6.2 Conclusions

A discrete time bivariate semi-Markov chain has been defined and a recursive system for the evaluation of its transition probabilities has been derived. The transition probabilities define the evolution of the chain over time.

Assuming the independence of the two components waiting times, an evolution equation for the marginal transition function has been obtained.

Chapter 7

Bivariate Reliability Model and Application to Counterparty Credit Risk

The current financial crisis has necessitated further study of correlations in the financial market. In this regard, the study of the risk of counterparty default, in any financial contract, has become crucial in determining credit risk. For a complete treatment about credit risk we refer to the classical book by Bielecki and Rutkowski [5]. Many works have been done trying to describe the counterparty risk in a Credit Default Swap (CDS) contract, but all these works are based on the Markovian approach to the credit risk, see for example Crepey *et al.* [31] or Ching and Ng [23].

Markov chain based models are too restrictive for the description of accurate rating dynamics. Indeed, they require that the distribution functions of the sojourn times in a rating class, before transition into a different rating class, should be exponentially or geometrically distributed for continuous and discrete time models, respectively. In the early 90s, Carty and Fons [22] demonstrated that a Weibull distribution most closely models the sojourn times in a given rating class.

In an attempt to produce more efficient credit rating models, homogeneous semi-Markov processes were proposed for the first time as applied to credit ratings in the paper by D'Amico *et al.* [32]; more recent results were given in D'Amico *et al.* [38, 37].

It is important to employ efficient migration models because reliable rating predictions are of interest for pricing rating sensitive derivatives Vasileiou and Vassiliou [85], D'Amico *et al.* [33], for the valuation of portfolios of defaulting bonds, for credit risk management and capital allocation.

No results are available for counterparty credit risk for semi-Markov credit rating models. Such an extension is not straightforward, as randomness in the sojourn times and memory effects have to be appropriately managed.

In this chapter, therefore, we present a novel multivariate semi-Markov model to approach the counterparty risk in a CDS contract. In Section 7.1 we present a 2-component reliability model and results of strict relevance to the financial problems. In Section 7.2 we analyze the counterparty credit risk in a CDS contract. In Section 7.3 we present a numerical example involving real data on credit ratings.

This chapter is based on a submitted paper (G. D'Amico, R. Manca and G. Salvi [40]) currently under review.

7.1 Bivariate Semi-Markov Reliability Model

Let us consider a system with two components supposed to be credit rating of companies or financial institutions. The state space E of the system is a collection of the ratings, e.g. it could be the rating classes of S&P or Moody's.

Let us partition the state space E in two parts respectively U and D, that is

$$E = U \bigcup D, \qquad \text{such that} \qquad U \bigcap D = \emptyset \ . \tag{7.1}$$

The subset U contains all good or alive states, when the components of the system are in this state they are regarded as well working. Conversely, the subset D contains all bad or dead states, when the components are in these state they are not well performing.

Assumption 7.1.1. *Let us assume that*

B1 All states in D are absorbing for the components.

It is a common assumption to consider the subset D as an absorbent class, in what follows we will make this assumption as well.

The assumption B1 is quite common in survival analysis but it can be relaxed easily and it is possible to execute the following computation in the general case.

The main issue of reliability models is the study of survival probability, that is the probability that the system will not default in a given time interval. This probability as a function of time is called reliability function, conditioning on the present state it is defined by

$$R^{sys}_{i_1 i_2}((v_1, v_1); t_1, t_2) := \mathbb{P}(Z^1(t_1) \in U, Z^2(t_2) \in U \mid \mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}). \tag{7.2}$$

The conditional reliability defined above gives the probability that after t_1 and t_2 periods the first and second component, respectively, have not defaulted given the present state.

The following result gives a recursive formula for the conditional reliability of the system.

Theorem 7.1.2. *Suppose that the system is composed of two components, i.e. $\gamma = 2$, and such that hypotheses A1, A2, A3 and B1 hold true. Then, for all $\mathbf{i} \in U^2$, $t, s \in \mathbb{N}$, and $\mathbf{v} \in \mathbb{N}^2$ we have*

$$R^{sys}_{i_1 i_2}((v_1, v_2); s, t) = \sum_{l_1, l_2 \in U} \sum_{w_1, w_2 \in \mathbb{N}} R^{sys}_{l_1 l_2}((w_1, w_2); s-1, t-1) \tilde{P}_{\mathbf{i},\mathbf{l}}(\mathbf{v}, \mathbf{w}) \ , \tag{7.3}$$

for $s, t > 1$ and where $\tilde{P}$ is given by

$$\tilde{P}_{\mathbf{i},\mathbf{l}}(\mathbf{v}, \mathbf{w}) = \begin{cases} \left(\frac{1-H^1_{i_1}(1+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(1+v_2)}{1-H^2_{i_2}(v_2)}\right) & \text{if} \quad \mathbf{l} = \mathbf{i}, \mathbf{w} = \mathbf{v} + \mathbf{1} \\ {}^b q^1_{\mathbf{i},l_1}(v_1, 1) \left(\frac{1-H^2_{i_2}(1+v_2)}{1-H^2_{i_2}(v_2)}\right) & \text{if} \quad l_1 \neq i_1, l_2 = i_2, \mathbf{w} = \binom{0}{1+v_2} \\ {}^b q^2_{\mathbf{i},l_2}(v_2, 1) \left(\frac{1-H^1_{i_1}(1+v_1)}{1-H^1_{i_1}(v_1)}\right) & \text{if} \quad l_1 = i_1, l_2 \neq i_2, \mathbf{w} = \binom{1+v_1}{0} \\ {}^b q^1_{\mathbf{i},l_1}(v_1, 1) {}^b q^2_{\mathbf{i},l_2}(v_2, 1) & \text{if} \quad l_1 \neq i_1, l_2 \neq i_2, \mathbf{w} = \binom{0}{0} \\ 0 & \text{otherwise} \end{cases} . \tag{7.4}$$

Proof. The reliability function is the probability of the event $\{Z^1(s) \in U, Z^2(t) \in U\}$, varying $s, t \in \mathbb{N}$, given the initial state $\mathbf{Z}(0) = \mathbf{i} \in U^2$ with backward recurrence time $\mathbf{B}(0) = \mathbf{v}$. Conditioning on the first step this probability can be written as

$$\begin{aligned} R^{sys}_{i_1 i_2}((v_1, v_2); s, t) &= \mathbb{P}(Z^1(s) \in U, Z^2(t) \in U \mid \mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}) \\ &= \sum_{l_1, l_2 \in U} \sum_{w_1, w_2 \in \mathbb{N}} \mathbb{P}(Z^1(s) \in U, Z^2(t) \in U, \mathbf{Z}(1) = \mathbf{l}, \mathbf{B}(1) = \mathbf{w} \mid \mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}) \ , \end{aligned} \tag{7.5}$$

using the Markov property of (Z, B) we have

$$\begin{aligned} R^{sys}_{i_1 i_2}((v_1, v_2); s, t) = \sum_{l_1, l_2 \in U} \sum_{w_1, w_2 \in \mathbb{N}} & \mathbb{P}(Z^1(s) \in U, Z^2(t) \in U \mid \mathbf{Z}(1) = \mathbf{l}, \mathbf{B}(1) = \mathbf{w}) \\ & \times \mathbb{P}(\mathbf{Z}(1) = \mathbf{l}, \mathbf{B}(1) = \mathbf{w} \mid \mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}) \ . \end{aligned} \tag{7.6}$$

Using the time homogeneity, we obtain

$$R^{sys}_{i_1 i_2}((v_1, v_2); s, t) = \sum_{l_1, l_2 \in U} \sum_{w_1, w_2 \in \mathbb{N}} R^{sys}_{l_1 l_2}((w_1, w_2); s-1, t-1) \tilde{P}_{\mathbf{i},\mathbf{l}}(\mathbf{v}, \mathbf{w}) \ , \tag{7.7}$$

where $\tilde{P}$ is the one-step transition probability of $(\mathbf{Z}, \mathbf{B})$ obtained in Lemma 6.1.28. The proof is complete. □

Remark 7.1.3. *If we consider a model where the two components are supposed to be independent, then, we are able to evaluate the reliability of the single component credit rating. The credit rating, in the independent case, is described by a standard univariate semi-Markov chain (see D'Amico et al. [33]) and the reliability functions of the component is*

$$\begin{aligned} {}^{ind}R^1_{i_1}(v_1, t) &:= \mathbb{P}(Z^1(t) \in U \mid Z^1(0) = i_1, B^1(0) = v_1) \\ {}^{ind}R^2_{i_2}(v_2, t) &:= \mathbb{P}(Z^2(t) \in U \mid Z^2(0) = i_2, B^2(0) = v_2) \ , \end{aligned} \tag{7.8}$$

where the left apexes ind stand for independent case, to stress the difference with the bivariate case. If the two components are independent, the product of these two reliabilities should be equal to the reliability for the system evaluated in Theorem 7.1.2. Indeed in the independent case the joint probability simple factorize in the product of the probabilities. Then, any deviation of the ratio

$$\frac{R^{sys}_{i_1 i_2}((v_1, v_2); s, t)}{{}^{ind}R^1_{i_1}(v_1, s)\,{}^{ind}R^2_{i_2}(v_2, t)} \tag{7.9}$$

by one is an indication of the correlation between the two components.

In many applications at the present time we could only know that the system is in an Up state and it has a given starting state distribution on the state space U^2. In this case it is more natural to define the reliability of the system as

$$R^{sys}(s, t) := \sum_{\mathbf{i} \in U^2} \sum_{v_1, v_2 \in \mathbb{N}} \xi_{\mathbf{i}}(v_1, v_2) R^{sys}_{i_1 i_2}((v_1, v_2); s, t) \qquad \text{for } t \in \mathbb{N} \ , \tag{7.10}$$

where $\xi = (\xi_{\mathbf{i}}(v_1, v_2))_{\mathbf{i} \in U^2; v_1, v_2 \in \mathbb{N}}$ is a starting distribution on possible states and backward values.

In what follows, it might be useful to have the probability that a given component is still alive after a given time interval. Then, let us define the conditional marginal reliability functions for a single component as

$$\begin{aligned} R^1_{i_1, i_2}((v_1, v_2); t) &:= \mathbb{P}(Z^1(t) \in U \mid \mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}) \\ R^2_{i_1, i_2}((v_1, v_2); t) &:= \mathbb{P}(Z^2(t) \in U \mid \mathbf{Z}(0) = \mathbf{i}, \mathbf{B}(0) = \mathbf{v}). \end{aligned} \tag{7.11}$$

The next result expresses the link between the marginal reliabilities and the transition function of the bivariate semi-Markov chain.

Proposition 7.1.4. *The marginal reliabilities can be expressed, for all* $\mathbf{i} \in U^2$, $\mathbf{v} \in \mathbb{N}^2$ *and* $t \in \mathbb{N}$, *as*

$$R^{\alpha}_{i_1,i_2}((v_1, v_2); t) = \sum_{u_\alpha \geq 0} \sum_{j_\alpha \in U} {}^b\Phi^{\alpha}_{\mathbf{i};j_\alpha}(\mathbf{v}, u_\alpha; t) = \sum_{j_\alpha \in U} {}^b\Phi^{\alpha}_{\mathbf{i};j_\alpha}(\mathbf{v}, \cdot; t) \ , \tag{7.12}$$

for $\alpha \in \{1, 2\}$.

Proof. The result follows directly from the definitions of marginal reliability and transition function. □

We can define, even for the marginal case, a reliability function not conditioned to any particular initial state that is

$$R^{\alpha}(t) := \sum_{\mathbf{i} \in U^2} \sum_{v_1, v_2 \in \mathbb{N}} \xi_{\mathbf{i}}(v_1, v_2) R^{\alpha}_{i_1 i_2}((v_1, v_2); t) \qquad \text{for } t \in \mathbb{N} \ , \tag{7.13}$$

for $\alpha = 1, 2$ and where ξ is a distribution on the state space and backward values.

7.2 Counterparty Credit Risk in a CDS Contract

In the financial market all subjects are exposed to the default risk. Therefore, in any financial contract we have to take the risk of default of our counterpart into account. Counterparty credit risk is, in general, 'the risk that a counterpart of a financial contract will default prior to the expiration of the contract and will not make all the payments required by the contract' (*cf.* Pykhtin and Zhu [73]).

We would like to study the counterparty credit risk in a Credit Default Swap (CDS) contract. In this work we would like to emphasize the difference between the CDS contract with and without consideration of the counterparty risk: we will call these two cases *risky CDS* and *risk free CDS*, respectively (Crépey *et al.* [31]). Let us consider a firm C, supposed to be defaultable, emitting an obligation (or bond) on one money unit at time 0 with maturity time T. Let us also consider a bondholder A (or protection buyer), who is supposed to be risk free in all what follows. The possible financial scenarios are

- If C has not been defaulted until T, then it is able to pay the money due to bondholder A.
- In case of C's default before, or at, the maturity date T, it will be able only to pay a fraction (recovery rate ρ_C) of the face value of the obligation to A.

For these reasons, bondholder A is looking for protection against the loss that would occur in the case of C's default. Let us consider a third financial subject that we will generically call protection seller B. A risk free CDS is a contract obligates A (protection buyer) to pay a fee to B (protection seller who is supposed to be risk free) in exchange for protection against the default of the reference credit firm C. The cash flows of a risk free CDS are

- A pays B a stream of premia with spread K, from the initial date until the occurrence of the default event or the maturity date T.
- In case of default of C, B has to cover the loss of A. Then B has to pay $1 - \rho_C$ unit of money to A.

The value of the spread is evaluated in order to guarantee that the contract has a value of zero at the inception time t_0. We assume that the payment of B to A is made at the same time as the default event. The cash amount should be discounted in order to be comparable, for this reason we introduce a structure of discount factors $(\beta_t)_{t\in\mathbb{N}}$. To define β we introduce the deterministic interest rates $(r_t)_{t\in\mathbb{N}}$ and then we can define

$$\beta_t := \begin{cases} 1 & \text{if} \quad t = 0 \\ \prod_{h=1}^{t}(1+r_h)^{-1} & \text{if} \quad t > 0 \end{cases} \tag{7.14}$$

Let τ_C be the time of default for the credit reference firm C. From the above discussion, we can directly write an expression for the cash flows and price process of the risk free CDS contract. The In Cash Flows process from the perspective of the bondholder A in the risk free CDS is given by

$$\beta_{\tau_C}(1-\rho_C)\mathbb{I}_{\{t_0<\tau_C\leq T\}}, \tag{7.15}$$

where β is a discount factor. The Out Cash Flows process is given by

$$-K\sum_{s=t_0}^{T}\beta_s\mathbb{I}_{\{s<\tau_C\}}. \tag{7.16}$$

Then, the discounted value of the risk free CDS with maturity T at time $t > t_0$ is

$$\beta_t p_T(t) = -K\sum_{s=t}^{T}\beta_s\mathbb{I}_{\{s<\tau_C\}} + \beta_{\tau_C}(1-\rho_C)\mathbb{I}_{\{t<\tau_C\leq T\}} \tag{7.17}$$

and its price process is given by $P_T(t) = \mathbb{E}_t[p_T(t)]$. The subscript t, hereafter indicates the information that, at time t, the process is still in one of the Up states.

Remark 7.2.1. *The price of a risk free CDS can be evaluated with a single component reliability model, where the rating of the only defaultable subject is modeled via a standard univariate semi-Markov process (see for example D'Amico et al. [33]).*

A risky CDS is a contract which obligates A (protection buyer) to pay a fee to B (defaultable protection seller) in exchange for protection against the default of the reference credit firm C. The cash flows of a risky CDS are

- A pays to B a stream of premia with spread K, from the initial date until the occurrence of the default event or the maturity date T.

- In case of default of C, if B has not defaulted, B has to cover the loss of A. Then B has to pay $1 - \rho_C$ unit of money to A.

- In case of default of B, if C has not defaulted, the contract is stopped with a *Close-Out Cash Flow* (*cf.* Crépey *et al.* [31]). In this work we assume that the two parties agreed on a termination of the contract with a terminal cash flow paid to A, positive or negative, depending on the value of the risk free CDS computed at the time of default (*cf.* Brigo *et al.* [14]).

- If B defaults at the same time as the firm C, B will be only able to pay to A a fraction (recovery rate ρ_B) of the loss of A, namely $\rho_B(1 - \rho_C)$ unit of money.

The value of the spread is evaluated in order to guarantee, that the contract has zero value at the inception time t_0. We assume that the payment of B to A is made at the same time as the default event(s). The possible loss of A for the joint default event is an effect due to the counterparty risk.

Let us introduce τ_B, the time of default for the protection seller B. The In Cash Flows process for the risky CDS is given by

$$\beta_{\tau_C}(1-\rho_C)\mathbb{I}_{\{t_0<\tau_C\leq T\}}[\mathbb{I}_{\{\tau_C<\tau_B\}} + \rho_B\mathbb{I}_{\{\tau_C=\tau_B\leq T\}}] + \beta_{\tau_B}\mathbb{I}_{\{t_0<\tau_B\leq(T\wedge\tau_C)\}}\rho_B P_T^+(\tau_B) \tag{7.18}$$

here β is a discount factor and with $P_{\cdot}^{+}$ we denote the positive part of the price process for the risk free CDS. The Out Cash Flows process is given by

$$-K\sum_{s=t_0}^{T}\beta_s\mathbb{I}_{\{s<(\tau_C\wedge\tau_B)\}}-\beta_{\tau_B}\mathbb{I}_{\{t_0<\tau_B\leq(T\wedge\tau_C)\}}P_T^{-}(\tau_B) \tag{7.19}$$

where $P_{\cdot}^{-}$ stands for the negative part of the price process for the risk free CDS. Then, the discounted value of the risky CDS with maturity T at time $t>t_0$ is

$$\begin{aligned}\beta_t\pi_T(t) &= -K\sum_{s=t}^{T}\beta_s\mathbb{I}_{\{s<(\tau_C\wedge\tau_B)\}}+\beta_{\tau_C}(1-\rho_C)\mathbb{I}_{\{t<\tau_C\leq T\}}[\mathbb{I}_{\{\tau_C<\tau_B\}}+\\ &+ \rho_B\mathbb{I}_{\{\tau_C=\tau_B\leq T\}}]+\beta_{\tau_B}\mathbb{I}_{\{t<\tau_B\leq(T\wedge\tau_C)\}}(\rho_B P_T^{+}(\tau_B)-P_T^{-}(\tau_B)),\end{aligned} \tag{7.20}$$

and the price process for the risky CDS is $\Pi_T(t)=\mathbb{E}_t[\pi_T(t)]$. Here and in the following the subscript t denotes the information that the process is still in an Up state at time t.

7.2.1 Pricing Risky CDS and CVA Evaluation

In this section we apply the 2-component reliability model to the study of the counterpart risk in a CDS contract. In particular our goal is to price a risky CDS contract and to derive an expression for the credit value adjustment (CVA) which can be seen as a measure of the counterparty credit risk.

In order to price a risky CDS we need to evaluate the expected value of the indicator functions in (7.20) and then we first have to evaluate the price of a risk free CDS.

To evaluate a risk free CDS we consider a reliability model where the credit rating evolution of the firm C is supposed to be independent from B. We will denote the reliability in this independent case by ${}^{ind}R^C$, *cf.* Remark 7.1.3.

Proposition 7.2.2. *The price of a risk free CDS under the real world probability measure is*

$$\beta_t P_T(t)=-K\sum_{h=t+1}^{+\infty}\mathbb{P}_t(\tau_C=h)\sum_{s=t}^{T\wedge h}\beta_s+\sum_{h=t+1}^{T}\beta_h(1-\rho_C)\mathbb{P}_t(\tau_C=h)\ , \tag{7.21}$$

and for any $h>t$

$$\mathbb{P}_t(\tau_C=h)={}^{ind}R^C(h-1)-{}^{ind}R^C(h)\ , \tag{7.22}$$

where the subscript t denotes the information that the process is still in an Up state.

Proof. The result is a direct consequence of formula (7.17) and

$$\mathbb{P}_t(\tau_C = h) = \mathbb{P}_t(\tau_C > h-1) - \mathbb{P}_t(\tau_C > h) = {}^{ind}R^C(h-1) - {}^{ind}R^C(h) \ . \tag{7.23}$$

□

The following two results concern the evaluation of the risky CDS price.

Proposition 7.2.3. *For any $h, s \in \mathbb{N}$ the joint distribution of the stopping times has the following characterization in terms of reliabilities*

$$\mathbb{P}_t(\tau_C = h, \tau_B = h) = R^{sys}(h-1,h-1) - R^{sys}(h,h-1) - R^{sys}(h-1,h) + R^{sys}(h,h)$$

$$\mathbb{P}_t(\tau_C = h, \tau_B = h+s) = R^{sys}(h-1,h+s-1) - R^{sys}(h,h+s-1) - R^{sys}(h-1,h+s) + R^{sys}(h,h+s)$$

$$\mathbb{P}_t(\tau_C = h+s, \tau_B = h) = R^{sys}(h+s-1,h-1) - R^{sys}(h+s,h-1) - R^{sys}(h+s-1,h) + R^{sys}(h+s,h) \ .$$

Proof. The result is composed of three relations, the former can be obtained by

$$\begin{aligned}\mathbb{P}_t(\tau_C = h, \tau_B = h) &= \mathbb{P}_t(\tau_C > h-1, \tau_B > h-1) - \mathbb{P}_t(\tau_C > h, \tau_B > h-1) \\ &\quad - \mathbb{P}_t(\tau_C > h-1, \tau_B > h) + \mathbb{P}_t(\tau_C > h, \tau_B > h) \\ &= R^{sys}(h-1,h-1) - R^{sys}(h,h-1) - R^{sys}(h-1,h) + R^{sys}(h,h) \ .\end{aligned}$$

The second and third relations in the statement are similar, we will show only the second one. The third can be obtained in the same way. To this end, we note that the second can be expressed as

$$\begin{aligned}\mathbb{P}_t(\tau_C = h, \tau_B = h+s) &= \mathbb{P}_t(\tau_C > h-1, \tau_B > h+s-1) - \mathbb{P}_t(\tau_C > h, \tau_B > h+s-1) \\ &\quad - \mathbb{P}_t(\tau_C > h-1, \tau_B > h+s) + \mathbb{P}_t(\tau_C > h, \tau_B > h+s) \\ &= R^{sys}(h-1,h+s-1) - R^{sys}(h,h+s-1) - R^{sys}(h-1,h+s) + R^{sys}(h,h+s) \ .\end{aligned}$$

This concludes the proof. □

Remark 7.2.4. *The joint distribution of the stopping times has been fully characterized by reliability functions.*

Proposition 7.2.5. *The price of a risky CDS under the real world probability measure is*

$$\begin{aligned}\beta_t \Pi_T(t) = -K \sum_{h=t+1}^{+\infty} \mathbb{P}_t(\tau = h) \sum_{s=t}^{T\wedge h} \beta_s + (1-\rho_C) \sum_{h_C=t+1}^{T} \beta_{h_C} \Bigg[\sum_{h_B=1}^{+\infty} \mathbb{P}_t(\tau_C = h_C, \tau_B = h_C + h_B) \\ + \rho_B \mathbb{P}_t(\tau_C = h_C, \tau_B = h_C) \Bigg] + \sum_{h_B=t+1}^{T} \sum_{h_C=1}^{+\infty} \mathbb{P}_t(\tau_C = h_B + h_C, \tau_B = h_B) \beta_{h_B} \left[\rho_B P^+_{h_B} - P^-_{h_B} \right]\end{aligned} \tag{7.24}$$

where $\tau = \tau_C \wedge \tau_B$ is the minimum of the stopping times and then

$$\mathbb{P}_t(\tau = h) = \sum_{h_B=1}^{\infty} \mathbb{P}_t(\tau_C = h, \tau_B = h + h_B) + \sum_{h_C=1}^{\infty} \mathbb{P}_t(\tau_C = h + h_C, \tau_B = h) \ . \tag{7.25}$$

Proof. The result is a direct consequence of formula (7.20) and Proposition 7.2.3 □

Remark 7.2.6. *The difference between the price of a risky CDS and the price of a risk-free CDS has a particular importance: it is a measure of the loss of value a CDS contract undergoes due to the counterpart credit risk. This difference is called credit value adjustment (CVA). The credit value adjustment process (CVA_t) is defined by*

$$CVA_t = P_T(t) - \Pi_T(t) \qquad \text{for } t < \tau_B \ . \tag{7.26}$$

It measures the loss of value of the CDS contract. We notice that, in our model, the CVA process can be completely expressed in term of the reliabilities.

7.3 A numerical example

In this section, a numerical example able to illustrate the previous results is presented. The model is applied to a sample from the historical Standard & Poor's database, which has been managed in order to construct the input for our model directly from real data.

Data refer to entity ratings history, instruments ratings history and issue/maturity ratings history for a sample of Standard & Poor's rated entity, instruments stock or bonds sold by an entity at a particular time and issue/maturity for the Global Issuers and Structured Finance instruments that Standard & Poor's has rated since 1982 to 2007, respectively.

The rating evaluation, done by the rating agency, indicates the degree of reliability of a bond issued by a financial subject.

In the case of the rating agency Standard & Poors, there are eight different classes of rating expressing the creditworthiness of the rated firm. The ratings are listed to form the following set of states:

$$E = \{AAA, AA, A, BBB, BB, B, CCC, D\}.$$

The creditworthiness is highest for the rating AAA, assigned to firm extremely reliable with regard to financial obligations, and decreases towards the rating D, which expresses the occurrence of payment default on some financial obligations. A table showing the financial meaning of the Standard & Poor's rating categories is reported in the book by Bluhm *et al.* [11].

In this paper we apply the model in a simplified form: we consider a three state model. The first state represents the investment grade (INV), it includes the more reliable rating classes $\{AAA, AA, A, BBB\}$. The second state represents the speculative grade (SPE), it includes the less reliable rating classes $\{BB, B, CCC\}$. The third is the default state (D), it includes the rating class $\{D\}$. The state space of the system is $E = \{INV, SPE, D\}$.

Moreover, following the classification of Standard & Poor's, we divide the rating series into two sectors: finance and corporate. Thus, we consider a system of two components, each one corresponds to a given sector: the first component for the finance sector (B) and the second component for the corporate sector (C).

In order to implement the model, we estimate the following distributions from the data:

- The transition matrices of the embedded Markov chains $\mathbf{P}^{\alpha}$, $\alpha \in \{B, C\}$.
- The unconditional waiting times cumulative distribution function $\mathbf{H}^{\alpha}(t)$, $\alpha \in \{B, C\}$.

The matrix $\mathbf{P}^{\alpha}$ gives the transition probability of going from a system rating class $\mathbf{i} \in E^2$ to the rating class $j \in E$ of component α with next transition. We report in Table 1(a) and Table 1(b) the transition matrices of the embedded Markov chain for finance and corporate, respectively. We notice that transition probabilities of finance change as a function of the corporate state and conversely. This is an indication of the dependence between the sectors rating.

In Figure 7.1(a) and Figure 7.1(b), the unconditional waiting times cumulative distribution function for finance and corporate are shown as a function of time (years). In both cases, the waiting times CDF for the speculative rating class is higher than the corresponding for investment rating class. Intuitively this reflects the fact that investment rating classes are more stable than speculative ones.

From knowledge of the transition probabilities of the embedded Markov chain and of the unconditional

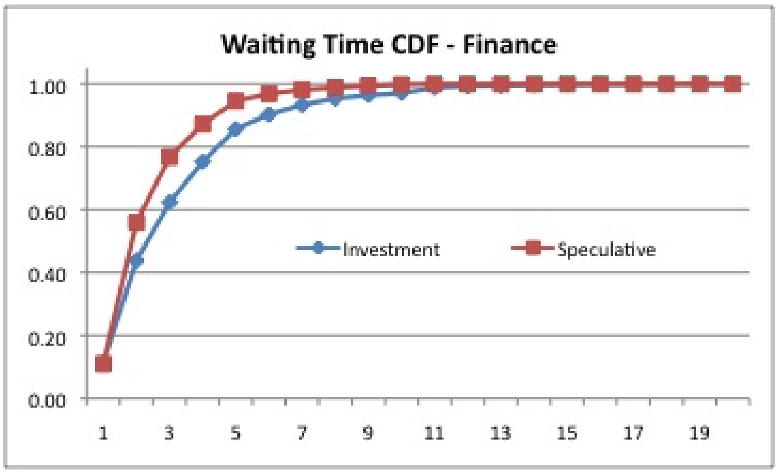
(a) Transition Matrix for Finance (B)

		States of B		
C	B	**INV**	**SPE**	**D**
INV	**INV**	0.885	0.113	0.002
INV	**SPE**	0.235	0.397	0.368
INV	**D**	0.000	0.000	1.000
SPE	**INV**	0.899	0.100	0.001
SPE	**SPE**	0.218	0.410	0.372
SPE	**D**	0.000	0.000	1.000
D	**INV**	0.913	0.086	0.001
D	**SPE**	0.172	0.424	0.404
D	**D**	0.000	0.000	1.000

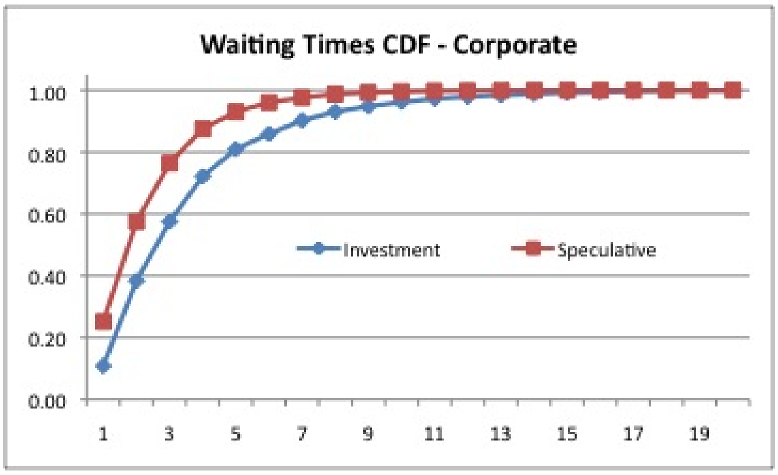
(b) Transition Matrix for Corporate (C)

		States of C		
B	C	**INV**	**SPE**	**D**
INV	**INV**	0.880	0.118	0.002
INV	**SPE**	0.062	0.687	0.251
INV	**D**	0.000	0.000	1.000
SPE	**INV**	0.882	0.116	0.002
SPE	**SPE**	0.068	0.692	0.240
SPE	**D**	0.000	0.000	1.000
D	**INV**	0.881	0.117	0.002
D	**SPE**	0.071	0.678	0.251
D	**D**	0.000	0.000	1.000

Table 7.1: Marginal Transition Matrices



(a) CDF Corporate



(b) CDF Insurance

Figure 7.1: Unconditional Waiting Times Cumulative Distribution Functions

waiting times cumulative distribution function, we can recover the semi-Markov kernel and, through Theorem 6.1.27 and Theorem 7.1.2, we can evaluate the transition probabilities of the system and the most important indicator, which is the reliability function of the system.

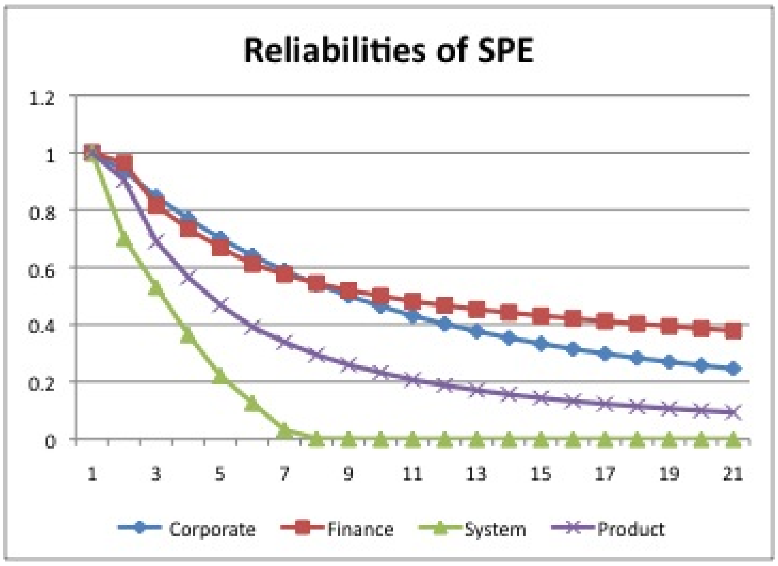


Figure 7.2: Reliabilities for speculative (SPE) state as a function of time.

In Figure 7.2 the system reliability for state $(Z^C, Z^B) = (SPE, SPE)$ and backwards null is shown as a function of time (years). In the same figure, we show the reliabilities of Corporate $({}^{ind}R^C_{SPE}(0,t))$ and Finance $({}^{ind}R^B_{SPE}(0,t))$ in the independent case (cf. Remark 7.1.3). Moreover in order to underline the dependence between the components, the product ${}^{ind}R^C_{SPE}(0,t){}^{ind}R^B_{SPE}(0,t)$ (Product in the figure) is reported as a comparison with ${}^{sys}R_{SPE,SPE}((0,0);t,t)$ (System in the figure) (cf. Eq. 7.9). As it is possible to see from the figure, the system's reliability with non-independent components evaluated with our bivariate semi-Markov chain model differs markedly form the independent case. This result suggests that adoption of the bivariate model enables capturing the dependence between the counter parties and, as a consequence, the risk of a counterparty default.

7.4 Conclusions

This paper proposes a multivariate semi-Markov chain model in discrete time. The multicomponent system is analysed in the transient case by implementing methods that compute the transition probabilities and reliability functions.

The results are applied to the evaluation of the risky credit default swap contracts and they allow the attainment of an explicit formula for the price of a risky CDS and for the credit value adjustment process.

Possible avenues for future developments of our model include:

a) application to real data on credit rating dynamics;

b) asymptotic properties of the multivariate semi-Markov model;

c) construction of a multivariate reward model for the credit spread computation.

Chapter 8

Bivariate Rewards Model and Application to Credit Spread Evaluation

In this chapter we define a bivariate rewards model and we apply it to the term structures and credit spread evaluation.

The chapter is organized as follows. The first section is devoted to the definition of bivariate reward process and evaluation of its moments. In the second section we derive an expression for the credit spread of two debtors.

8.1 Bivariate Reward Model

The main goal of this section is to define a bivariate semi-Markov reward model.

As well known, the rewards can be classified in two main classes: the permanence rewards where the payments are associated with occupancy of a given state and the instant rewards where the payments are associated to the transitions between two states.

The rewards are paid in different moments so we have to consider the discount factor β defined in Eq. (7.14).

Without loss of generality, being our system homogeneous in time by assumption, we can take $t = 0$ as the initial time. Let $\xi_{\mathbf{i}}^{\alpha}(\mathbf{v}, 0, t)$ be the accumulated discounted reward for the component $\alpha = 1, 2$ up to the time t given that at the initial time, the instant 0, the system is in the state $\mathbf{i} = (i_1, i_2)$ with backward recurrence time $\mathbf{v} = (v_1, v_2)$. We assume that $\xi_{\mathbf{i}}^{\alpha}(\mathbf{v}, 0) = 0$ so no reward is paid at the initial time and every payment is done at the end of a period. Let $\psi^{\alpha}_{\mathbf{i}}(\tau,t)$ be the permanence reward of the component α in the state $\mathbf{i}=(i_1,i_2)$, paid at time t, given that τ is the present life time of α. Let $\gamma^{\alpha}_{\mathbf{i},j_\alpha}(t)$ be the instant reward collected by the component α when it makes a transition at time t to the state j_α given that the previous state of the system was $\mathbf{i}=(i_1,i_2)$.

The accumulated reward process for the component 1 up to time t can be expressed as

$$\begin{aligned}
&\xi^1_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},0,t\right) \stackrel{d}{=} \mathbb{I}_{\{T_1^1>t,T_1^2>t\}}\sum_{s=1}^{t}\beta_s\psi^1_{\binom{i_1}{i_2}}(v_1+s,s)\\
&+\sum_{j_1\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2>s\}}\left[\beta_s\gamma^1_{\binom{i_1}{i_2},j_1}(s)+\sum_{h=1}^{s}\beta_h\psi^1_{\binom{i_1}{i_2}}(v_1+h,h)+\beta_s\xi^1_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\right]\\
&+\sum_{j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1>s,T_1^2=s,J_1^2=j_2\}}\left[\sum_{h=1}^{s}\beta_h\psi^1_{\binom{i_1}{i_2}}(v_1+h,h)+\beta_s\xi^1_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},s,t\right)\right]\\
&+\sum_{j_1,j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2=s,J_1^2=j_2\}}\left[\beta_s\gamma^1_{\binom{i_1}{i_2},j_1}(s)+\sum_{h=1}^{s}\beta_h\psi^1_{\binom{i_1}{i_2}}(v_1+h,h)+\beta_s\xi^1_{\binom{j_1}{j_2}}\left(\binom{0}{0},s,t\right)\right],
\end{aligned} \tag{8.1}$$

similarly for the component 2 we have

$$\begin{aligned}
&\xi^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},0,t\right) \stackrel{d}{=} \mathbb{I}_{\{T_1^1>t,T_1^2>t\}}\sum_{s=1}^{t}\beta_s\psi^2_{\binom{i_1}{i_2}}(v_2+s,s)\\
&+\sum_{j_1\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2>s\}}\left[\sum_{h=1}^{s}\beta_h\psi^2_{\binom{i_1}{i_2}}(v_2+h,h)+\beta_s\xi^2_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\right]\\
&+\sum_{j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1>s,T_1^2=s,J_1^2=j_2\}}\left[\beta_s\gamma^2_{\binom{i_1}{i_2},j_2}(s)+\sum_{h=1}^{s}\beta_h\psi^2_{\binom{i_1}{i_2}}(v_2+h,h)+\beta_s\xi^2_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},s,t\right)\right]\\
&+\sum_{j_1,j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2=s,J_1^2=j_2\}}\left[\beta_s\gamma^2_{\binom{i_1}{i_2},j_2}(s)+\sum_{h=1}^{s}\beta_h\psi^2_{\binom{i_1}{i_2}}(v_2+h,h)+\beta_s\xi^2_{\binom{j_1}{j_2}}\left(\binom{0}{0},s,t\right)\right],
\end{aligned} \tag{8.2}$$

Formulas (8.1) and (8.2) give a recursive representation of the bidimensional accumulated reward process $(\xi^1(t),\xi^2(t))$. The expression of the rewards is divided in four parts. The first part corresponds to the rewards accumulated when no jump occurs up to the time t. The second term is the contributions due to the trajectories where component 1 makes a transition before 2. The third conversely consists of the trajectories where the component 2 makes a transition before 1. The fourth term take into account for the rewards accumulated when the two components make a transition in the same time.

Let us define the discounted accumulated permanence reward for component α as

$$\Psi^{\alpha}_{\mathbf{i}}(\mathbf{v},t)=\sum_{s=1}^{t}\beta_s\psi^{\alpha}_{\binom{i_1}{i_2}}(v_\alpha+s,s) \tag{8.3}$$

and let us define

$$\Gamma^{\alpha}_{\mathbf{i},j_\alpha}(\mathbf{v},t) = \Psi^{\alpha}_{\mathbf{i}}(\mathbf{v},t) + \beta_t \gamma^{\alpha}_{\binom{i_1}{i_2},j_\alpha}(t) \tag{8.4}$$

that is the discounted accumulated reward of component α for permanence of the system up to time t in state $\mathbf{i}$ plus the contribution for the transition in t of component α in the state j_α. Using these functions the accumulated reward up to time t for component 1 is given by

$$\begin{aligned}
\xi^1_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},0,t\right) &\stackrel{d}{=} \mathbb{I}_{\{T^1_1>t,T^2_1>t\}}\Psi^1_{\mathbf{i}}(\mathbf{v},t) + \sum_{j_1\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T^1_1=s,J^1_1=j_1,T^2_1>s\}}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+\beta_s\xi^1_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\right]\\
&+\sum_{j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T^1_1>s,T^2_1=s,J^2_1=j_2\}}\left[\Psi^1_{\mathbf{i}}(\mathbf{v},s)+\beta_s\xi^1_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},s,t\right)\right]\\
&+\sum_{j_1,j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T^1_1=s,J^1_1=j_1,T^2_1=s,J^2_1=j_2\}}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+\beta_s\xi^1_{\binom{j_1}{j_2}}\left(\binom{0}{0},s,t\right)\right],
\end{aligned} \tag{8.5}$$

and similarly for component 2. Let us denote by V the expected value of the accumulated reward that is

$$\begin{aligned}
{}^1V_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t-s\right) &:= \mathbb{E}\left\{\xi^1_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},s,t\right)\middle|\mathbf{Z}(s)=\mathbf{i}\ ,\ \mathbf{B}(s)=\mathbf{v}\right\}\\
{}^2V_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t-s\right) &= \mathbb{E}\left\{\xi^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},s,t\right)\middle|\mathbf{Z}(s)=\mathbf{i}\ ,\ \mathbf{B}(s)=\mathbf{v}\right\}
\end{aligned} \tag{8.6}$$

In order to evaluate the expected value of the accumulated reward we have to evaluate the expectations of all the terms in the above representations of ξ. The next result gives us a recursive formula for the evaluation of these expected values.

Proposition 8.1.1. *The expected discount accumulated reward for component 1,* ${}^1V_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right)$*, satisfies the following recursive equations, for all* $\mathbf{v}\in\mathbb{N}^2$*,* $t\in\mathbb{N}$ *and* $\mathbf{i}\in E^2$*,*

$$\begin{aligned}
{}^1V_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right) &= \Psi^1_{\mathbf{i}}(\mathbf{v},t)\left(\frac{1-H^1_{i_1}(t+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(t+v_2)}{1-H^2_{i_2}(v_2)}\right)\\
&+\sum_{j_1\in E}\sum_{s=1}^{t}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+\beta_s{}^1V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right]\left(\frac{1-H^2_{i_2}(s+v_2)}{1-H^2_{i_2}(v_2)}\right){}^bq^1_{\mathbf{i},j_1}(v_1,s)\\
&+\sum_{j_2\in E}\sum_{s=1}^{t}\left[\Psi^1_{\mathbf{i}}(\mathbf{v},s)+\beta_s{}^1V_{\binom{i_1}{j_2}}\left(\binom{v_2+s}{0},t-s\right)\right]\left(\frac{1-H^1_{i_1}(s+v_1)}{1-H^1_{i_1}(v_1)}\right){}^bq^2_{\mathbf{i},j_2}(v_2,s)\\
&+\sum_{j_1,j_2\in E}\sum_{s=1}^{t}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+\beta_s{}^1V_{\binom{j_1}{j_2}}\left(\binom{0}{0},t-s\right)\right]{}^bq^1_{\mathbf{i},j_1}(v_1,s){}^bq^2_{\mathbf{i},j_2}(v_2,s),
\end{aligned} \tag{8.7}$$

and similarly for component 2,

$$\begin{aligned}
&{}^2V_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right)=\Psi^2_{\mathbf{i}}(\mathbf{v},t)\left(\frac{1-H^1_{i_1}(t+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(t+v_2)}{1-H^2_{i_2}(v_2)}\right)\\
&+\sum_{j_1\in E}\sum_{s=1}^{t}\left[\Psi^2_{\mathbf{i}}(\mathbf{v},s)+\beta_s{}^2V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right]\left(\frac{1-H^2_{i_2}(s+v_2)}{1-H^2_{i_2}(v_2)}\right){}^bq^1_{\mathbf{i},j_1}(v_1,s)\\
&+\sum_{j_2\in E}\sum_{s=1}^{t}\left[\Gamma^2_{\mathbf{i},j_2}(\mathbf{v},s)+\beta_s{}^2V_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},t-s\right)\right]\left(\frac{1-H^1_{i_1}(s+v_1)}{1-H^1_{i_1}(v_1)}\right){}^bq^2_{\mathbf{i},j_2}(v_2,s)\\
&+\sum_{j_2,j_2\in E}\sum_{s=1}^{t}\left[\Gamma^2_{\mathbf{i},j_2}(\mathbf{v},s)+\beta_s{}^2V_{\binom{j_1}{j_2}}\left(\binom{0}{0},t-s\right)\right]{}^bq^1_{\mathbf{i},j_1}(v_1,s){}^bq^2_{\mathbf{i},j_2}(v_2,s).
\end{aligned}\tag{8.8}$$

Proof. In order to obtain the result we have to evaluate the expected value of the discounted reward process for both components. We will evaluate only the expectation of ξ^1, with similar calculations we can obtain the expected value of ξ^2.

If we consider the representation of ξ^1 in (8.5) we have to evaluate the expected value of the four terms on the r.h.s., to this end we notice that the expectations of the indicator functions have been already calculated in the proof of Theorem 6.1.27. Then, the expectation of the first term given the present state is

$$\begin{aligned}
&\mathbb{E}\{\mathbb{I}_{\{T^1_1>t,T^2_1>t\}}\Psi^1_{\mathbf{i}}(\mathbf{v},t)|\mathbf{Z}(0)=\mathbf{i}\ ,\ \mathbf{B}(0)=\mathbf{v}\}\\
&=\Psi^1_{\mathbf{i}}(\mathbf{v},t)\mathbb{E}\{\mathbb{I}_{\{T^1_1>t,T^2_1>t\}}|\mathbf{Z}(0)=\mathbf{i}\ ,\ \mathbf{B}(0)=\mathbf{v}\}\\
&=\Psi^1_{\mathbf{i}}(\mathbf{v},t)\left(\frac{1-H^1_{i_1}(t+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(t+v_2)}{1-H^2_{i_2}(v_2)}\right)\ .
\end{aligned}\tag{8.9}$$

The second and third terms on the r.h.s of Eq. (8.5) have similar structure, therefore we consider only one of them. For instance, the expectation of the second term is

$$\begin{aligned}
&\mathbb{E}\left\{\sum_{j_1\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T^1_1=s,J^1_1=j_1,T^2_1>s\}}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+\beta_s\xi^1_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\right]\middle|\mathbf{Z}(0)=\mathbf{i}\ ,\ \mathbf{B}(0)=\mathbf{v}\right\}\\
&=\sum_{j_1\in E}\sum_{s=1}^{t}\mathbb{E}\left\{\mathbb{I}_{\{T^1_1=s,J^1_1=j_1,T^2_1>s\}}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+\beta_s\xi^1_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\right]\middle|\mathbf{Z}(0)=\mathbf{i}\ ,\ \mathbf{B}(0)=\mathbf{v}\right\}\\
&=\sum_{j_1\in E}\sum_{s=1}^{t}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+\beta_s{}^1V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right]\mathbb{E}\left\{\mathbb{I}_{\{T^1_1=s,J^1_1=j_1,T^2_1>s\}}|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\right\},
\end{aligned}\tag{8.10}$$

where in the last equality we use the properties of conditional expectation and the time homogeneity of the process

$${}^1V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)=\mathbb{E}\left\{\xi^1_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\middle|\mathbf{Z}(s)=\binom{j_1}{i_2},\mathbf{B}(s)=\binom{0}{v_2+s}\right\}.\tag{8.11}$$

The expectation of the indicator function on the last term of Eq. (8.10) from Theorem 6.1.27 is given by

$$\mathbb{E}\left\{\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2>s\}}|\mathbf{Z}(0)=\mathbf{i},\mathbf{B}(0)=\mathbf{v}\right\}=\left(\frac{1-H_{i_2}^2(s+v_2)}{1-H_{i_2}^2(v_2)}\right){}^b q_{\mathbf{i},j_1}^1(v_1,s)\ . \tag{8.12}$$

Then the expectation of the second term on the r.h.s of Eq. (8.5) is

$$\begin{aligned}&\mathbb{E}\left\{\sum_{j_1\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2>s\}}\left[\Gamma_{\mathbf{i},j_1}^1(\mathbf{v},s)+\beta_s\xi^1_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\right]\middle|\mathbf{Z}(0)=\mathbf{i}\ ,\ \mathbf{B}(0)=\mathbf{v}\right\}\\&=\sum_{j_1\in E}\sum_{s=1}^{t}\left[\Gamma_{\mathbf{i},j_1}^1(\mathbf{v},s)+\beta_s{}^1V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right]\left(\frac{1-H_{i_2}^2(s+v_2)}{1-H_{i_2}^2(v_2)}\right){}^b q_{\mathbf{i},j_1}^1(v_1,s).\end{aligned} \tag{8.13}$$

Finally for the fourth term on the r.h.s of Eq. (8.5) can be obtained again from Theorem 6.1.27 and we get

$$\mathbb{E}\left\{\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2=s,J_1^2=j_2\}}\middle|\mathbf{Z}(0)=\mathbf{i}\ ,\ \mathbf{B}(0)=\mathbf{v}\right\}={}^b q_{\mathbf{i},j_1}^1(v_1,s){}^b q_{\mathbf{i},j_2}^2(v_2,s)\ . \tag{8.14}$$

The result is obtained by substituting all this terms. □

Remark 8.1.2. *In order to evaluate* ${}^\alpha V$*, for* $\alpha=1,2$*, at time* t *we only need to know the value of* ${}^\alpha V$ *for all the times before* t*, then the previous equations can be solved by iteration given that at time* 0 *the rewards are null.*

Let us consider now the second moments. First of all we evaluate the product of the rewards of the two components, i.e. the second mixed moments. Using the same representation we obtain

$$\begin{aligned}&\xi^1_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},0,t\right)\xi^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},0,t\right)\stackrel{d}{=}\mathbb{I}_{\{T_1^1>t,T_1^2>t\}}\Psi_{\mathbf{i}}^1(\mathbf{v},t)\Psi_{\mathbf{i}}^2(\mathbf{v},t)\\&+\sum_{j_1\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2>s\}}\left[\Gamma_{\mathbf{i},j_1}^1(\mathbf{v},s)+\beta_s\xi^1_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\right]\left[\Psi_{\mathbf{i}}^2(\mathbf{v},s)+\beta_s\xi^2_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\right]\\&+\sum_{j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1>s,T_1^2=s,J_1^2=j_2\}}\left[\Psi_{\mathbf{i}}^1(\mathbf{v},s)+\beta_s\xi^1_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},s,t\right)\right]\left[\Gamma_{\mathbf{i},j_2}^2(\mathbf{v},s)+\beta_s\xi^2_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},s,t\right)\right]\\&+\sum_{j_1,j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2=s,J_1^2=j_2\}}\left[\Gamma_{\mathbf{i},j_1}^1(\mathbf{v},s)+\beta_s\xi^1_{\binom{j_1}{j_2}}\left(\binom{0}{0},s,t\right)\right]\left[\Gamma_{\mathbf{i},j_2}^2(\mathbf{v},s)+\beta_s\xi^2_{\binom{j_1}{j_2}}\left(\binom{0}{0},s,t\right)\right].\end{aligned} \tag{8.15}$$

Let us denote by

$${}^{12}V^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t-s\right):=\mathbb{E}\left\{\xi^1_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},s,t\right)\xi^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},s,t\right)\middle|\mathbf{Z}(s)=\mathbf{i}\ ,\ \mathbf{B}(s)=\mathbf{v}\right\}, \tag{8.16}$$

here the right upper script 2 indicates the second moment and the left upper script 12 denotes the mixed moments.

Using the same techniques of the first moment, we can obtain the present expected value which is formalized in the next result.

Proposition 8.1.3. *The second mixed moment for the discount accumulated reward of the system,* ${}^{12}V^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right)$, *satisfies the following recursive equations, for all* $\mathbf{v}\in\mathbb{N}^2$, $t\in\mathbb{N}$ *and* $\mathbf{i}\in E^2$,

$$
\begin{aligned}
&{}^{12}V^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right)=\left(\frac{1-H^1_{i_1}(t+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(t+v_2)}{1-H^2_{i_2}(v_2)}\right)\Psi^1_{\mathbf{i}}(\mathbf{v},t)\Psi^2_{\mathbf{i}}(\mathbf{v},t)\\
&+\sum_{j_1\in E}\sum_{s=1}^{t}\left(\frac{1-H^2_{i_2}(s+v_2)}{1-H^2_{i_2}(v_2)}\right){}^b q^1_{\mathbf{i},j_1}(v_1,s)\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)\Psi^2_{\mathbf{i}}(\mathbf{v},s)+\beta_s{}^1V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\Psi^2_{\mathbf{i}}(\mathbf{v},s)\right.\\
&\left.+\beta_s{}^2V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+(\beta_s^2){}^{12}V^2_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right]\\
&+\sum_{j_2\in E}\sum_{s=1}^{t}\left(\frac{1-H^1_{i_1}(s+v_1)}{1-H^1_{i_1}(v_1)}\right){}^b q^2_{\mathbf{i},j_2}(v_2,s)\left[\Psi^1_{\mathbf{i}}(\mathbf{v},s)\Gamma^2_{\mathbf{i},j_2}(\mathbf{v},s)+\beta_s{}^1V_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},t-s\right)\Gamma^2_{\mathbf{i},j_2}(\mathbf{v},s)\right.\\
&\left.+\beta_s{}^2V_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},t-s\right)\Psi^1_{\mathbf{i}}(\mathbf{v},s)+(\beta_s^2){}^{12}V^2_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},t-s\right)\right]\\
&+\sum_{j_1,j_2\in E}\sum_{s=1}^{t}{}^b q^1_{\mathbf{i},j_1}(v_1,s){}^b q^2_{\mathbf{i},j_2}(v_2,s)\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)\Gamma^2_{\mathbf{i},j_2}(\mathbf{v},s)+\beta_s{}^1V_{\binom{j_1}{j_2}}\left(\binom{0}{0},t-s\right)\Gamma^2_{\mathbf{i},j_2}(\mathbf{v},s)\right.\\
&\left.+\beta_s{}^2V_{\binom{j_1}{j_2}}\left(\binom{0}{0},t-s\right)\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+(\beta_s^2){}^{12}V^2_{\binom{j_1}{j_2}}\left(\binom{0}{0},t-s\right)\right].
\end{aligned}
\tag{8.17}
$$

Proof. The evaluation can be done following the same steps of the first moments. □

The second moment of component 1 and 2 can be obtained directly by products of the first moments, explicitly we have

$$
\begin{aligned}
&\left(\xi^1_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},0,t\right)\right)^2\overset{d}{=}\mathbb{I}_{\{T_1^1>t,T_1^2>t\}}\left(\Psi^1_{\mathbf{i}}(\mathbf{v},t)\right)^2+\sum_{j_1\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2>s\}}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+\right.\\
&\left.\beta_s\xi^1_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},s,t\right)\right]^2+\sum_{j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1>s,T_1^2=s,J_1^2=j_2\}}\left[\Psi^1_{\mathbf{i}}(\mathbf{v},s)+\beta_s\xi^1_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},s,t\right)\right]^2\\
&+\sum_{j_1,j_2\in E}\sum_{s=1}^{t}\mathbb{I}_{\{T_1^1=s,J_1^1=j_1,T_1^2=s,J_1^2=j_2\}}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)+\beta_s\xi^1_{\binom{j_1}{j_2}}\left(\binom{0}{0},s,t\right)\right]^2,
\end{aligned}
\tag{8.18}
$$

for the square reward of 1 and similar for 2. Let us denote by

$$
\begin{aligned}
{}^1V^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t-s\right)&:=\mathbb{E}\left\{\left(\xi^1_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},s,t\right)\right)^2\middle|\mathbf{Z}(s)=\mathbf{i}\ ,\ \mathbf{B}(s)=\mathbf{v}\right\}\\
{}^2V^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t-s\right)&:=\mathbb{E}\left\{\left(\xi^2_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},s,t\right)\right)^2\middle|\mathbf{Z}(s)=\mathbf{i}\ ,\ \mathbf{B}(s)=\mathbf{v}\right\},
\end{aligned}
\tag{8.19}
$$

here the right upper scripts indicate the second moment and the left upper scripts denotes the component.

Proposition 8.1.4. *The second moments for the discount accumulated reward of the system,* ${}^{1}V^{2}_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right)$ *and* ${}^{2}V^{2}_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right)$*, satisfy for all* $\mathbf{v}\in\mathbb{N}^2$*,* $t\in\mathbb{N}$ *and* $\mathbf{i}\in E^2$*, the following recurrent equations*

$$
\begin{aligned}
&{}^{1}V^{2}_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right)=\left(\frac{1-H^1_{i_1}(t+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(t+v_2)}{1-H^2_{i_2}(v_2)}\right)\left(\Psi^1_{\mathbf{i}}(\mathbf{v},t)\right)^2\\
&+\sum_{j_1\in E}\sum_{s=1}^{t}\left(\frac{1-H^2_{i_2}(s+v_2)}{1-H^2_{i_2}(v_2)}\right){}^{b}q^1_{\mathbf{i},j_1}(v_1,s)\left[\left(\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)\right)^2+2\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)\beta_s{}^{1}V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right.\\
&\left.+(\beta_s^2){}^{1}V^2_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right]+\sum_{j_2\in E}\sum_{s=1}^{t}\left(\frac{1-H^1_{i_1}(s+v_1)}{1-H^1_{i_1}(v_1)}\right){}^{b}q^2_{\mathbf{i},j_2}(v_2,s)\left[\left(\Psi^1_{\mathbf{i}}(\mathbf{v},s)\right)^2\right.\\
&\left.+2\Psi^1_{\mathbf{i}}(\mathbf{v},s)\beta_s{}^{1}V_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},t-s\right)+(\beta_s^2){}^{1}V^2_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},t-s\right)\right]+\sum_{j_1,j_2\in E}\sum_{s=1}^{t}{}^{b}q^1_{\mathbf{i},j_1}(v_1,s){}^{b}q^2_{\mathbf{i},j_2}(v_2,s)\\
&\times\left[\left(\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)\right)^2+2\Gamma^1_{\mathbf{i},j_1}(\mathbf{v},s)\beta_s{}^{1}V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)+(\beta_s^2){}^{1}V^2_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right],
\end{aligned}
\tag{8.20}
$$

and for component 2

$$
\begin{aligned}
&{}^{2}V^{2}_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right)=\left(\frac{1-H^1_{i_1}(t+v_1)}{1-H^1_{i_1}(v_1)}\right)\left(\frac{1-H^2_{i_2}(t+v_2)}{1-H^2_{i_2}(v_2)}\right)\left(\Psi^2_{\mathbf{i}}(\mathbf{v},t)\right)^2\\
&+\sum_{j_1\in E}\sum_{s=1}^{t}\left(\frac{1-H^2_{i_2}(s+v_2)}{1-H^2_{i_2}(v_2)}\right){}^{b}q^1_{\mathbf{i},j_1}(v_1,s)\left[\left(\Gamma^2_{\mathbf{i},j_1}(\mathbf{v},s)\right)^2+2\Gamma^2_{\mathbf{i},j_1}(\mathbf{v},s)\beta_s{}^{2}V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right.\\
&\left.+(\beta_s^2){}^{2}V^2_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right]+\sum_{j_2\in E}\sum_{s=1}^{t}\left(\frac{1-H^1_{i_1}(s+v_1)}{1-H^1_{i_1}(v_1)}\right){}^{b}q^2_{\mathbf{i},j_2}(v_2,s)\left[\left(\Psi^2_{\mathbf{i}}(\mathbf{v},s)\right)^2\right.\\
&\left.+2\Psi^2_{\mathbf{i}}(\mathbf{v},s)\beta_s{}^{2}V_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},t-s\right)+(\beta_s^2){}^{2}V^2_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},t-s\right)\right]+\sum_{j_1,j_2\in E}\sum_{s=1}^{t}{}^{b}q^1_{\mathbf{i},j_1}(v_1,s){}^{b}q^2_{\mathbf{i},j_2}(v_2,s)\\
&\times\left[\left(\Gamma^2_{\mathbf{i},j_1}(\mathbf{v},s)\right)^2+2\Gamma^2_{\mathbf{i},j_1}(\mathbf{v},s)\beta_s{}^{2}V_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)+(\beta_s^2){}^{2}V^2_{\binom{j_1}{i_2}}\left(\binom{0}{v_2+s},t-s\right)\right].
\end{aligned}
\tag{8.21}
$$

Remark 8.1.5. *The same technique can be applied in order to obtain moment of higher order, we omit here details for simplicity.*

8.2 Rating Migration Model for Term Structures and Credit Spread

In this section we develop a model for the credit spread value between two debtors, i.e. two credit rating endowed financial subjects or companies or national governments.

The rating evaluation, done by the rating agency, gives a reliability's degree of a bond issued by a financial subject.

In the case of the rating agency Standard & Poors there are 8 different classes of rating expressing the creditworthiness of the rated firm. The ratings are listed to form the following set of states:

$$E = \{AAA, AA, A, BBB, BB, B, CCC, D\}.$$

The evaluation of the risk structure of interest rates is one of the most important problems in mathematical finance. Fundamentally it consists in trying to explain yield spread between the risk free interest rate and the interest rates of corporate bonds. The rating evaluation of the firm issuing the bond is one of the main reasons for the existence of credit spread. In the paper by D'Amico *et al.* [37] it was presented a model that can follow the mean evolution of the yield spread in the future by considering rating evaluation as the determinant of credit spreads. That paper assumed that the rating evolution of the credit rating was described by a semi-Markov process and the credit spreads were considered as permanence rewards attached to the rating class. In this section we propose a more complete and general approach to this problem. As we know from the last financial crisis, the credit spread of one financial subject depends also on the financial health (rating class) of its counterpart, then it is very important to consider the spread as a reward structure for a bivariate rating model.

Let us consider two debtors whose credit ratings are supposed to migrate via a bivariate reliability semi-Markov model. The expected total credit spreads the debtor 1 will accumulate in a time t, under the assumption that both of them will not make a default, are

$$\begin{aligned}
{}^1V_{\binom{i_1}{i_2}}\left(\tbinom{v_1}{v_2}, t\right) &= \Psi^1_{\mathbf{i}}(\mathbf{v}, t)\left(\frac{1 - H^1_{i_1}(t + v_1)}{1 - H^1_{i_1}(v_1)}\right)\left(\frac{1 - H^2_{i_2}(t + v_2)}{1 - H^2_{i_2}(v_2)}\right) \\
&+ \sum_{j_1 \in U}\sum_{s=1}^{t}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v}, s) + {}^1V_{\binom{j_1}{i_2}}\left(\tbinom{0}{v_2 + s}, t - s\right)\right]\left(\frac{1 - H^2_{i_2}(s + v_2)}{1 - H^2_{i_2}(v_2)}\right){}^b q^1_{\mathbf{i},j_1}(v_1, s) \\
&+ \sum_{j_2 \in U}\sum_{s=1}^{t}\left[\Psi^1_{\mathbf{i}}(\mathbf{v}, s) + {}^1V_{\binom{i_1}{j_2}}\left(\tbinom{v_2 + s}{0}, t - s\right)\right]\left(\frac{1 - H^1_{i_1}(s + v_1)}{1 - H^1_{i_1}(v_1)}\right){}^b q^2_{\mathbf{i},j_2}(v_2, s) \\
&+ \sum_{j_1, j_2 \in U}\sum_{s=1}^{t}\left[\Gamma^1_{\mathbf{i},j_1}(\mathbf{v}, s) + {}^1V_{\binom{j_1}{j_2}}\left(\tbinom{0}{0}, t - s\right)\right]{}^b q^1_{\mathbf{i},j_1}(v_1, s)\,{}^b q^2_{\mathbf{i},j_2}(v_2, s),
\end{aligned} \tag{8.22}$$

where U is the subset of all the good states. Similarly for the debtor 2 we have

$$\begin{aligned}
{}^2V_{\binom{i_1}{i_2}}\left(\tbinom{v_1}{v_2}, t\right) &= \Psi^2_{\mathbf{i}}(\mathbf{v}, t)\left(\frac{1 - H^1_{i_1}(t + v_1)}{1 - H^1_{i_1}(v_1)}\right)\left(\frac{1 - H^2_{i_2}(t + v_2)}{1 - H^2_{i_2}(v_2)}\right) \\
&+ \sum_{j_1 \in U}\sum_{s=1}^{t}\left[\Psi^2_{\mathbf{i}}(\mathbf{v}, s) + {}^2V_{\binom{j_1}{i_2}}\left(\tbinom{0}{v_2 + s}, t - s\right)\right]\left(\frac{1 - H^2_{i_2}(s + v_2)}{1 - H^2_{i_2}(v_2)}\right){}^b q^1_{\mathbf{i},j_1}(v_1, s)
\end{aligned}$$

$$+\sum_{j_2\in U}\sum_{s=1}^{t}\left[\Gamma^2_{\mathbf{i},j_2}(\mathbf{v},s)+{}^2V_{\binom{i_1}{j_2}}\left(\binom{v_1+s}{0},t-s\right)\right]\left(\frac{1-H^1_{i_1}(s+v_1)}{1-H^1_{i_1}(v_1)}\right){}^bq^2_{\mathbf{i},j_2}(v_2,s)$$
$$+\sum_{j_2,j_2\in U}\sum_{s=1}^{t}\left[\Gamma^2_{\mathbf{i},j_2}(\mathbf{v},s)+{}^2V_{\binom{j_1}{j_2}}\left(\binom{0}{0},t-s\right)\right]{}^bq^1_{\mathbf{i},j_1}(v_1,s){}^bq^2_{\mathbf{i},j_2}(v_2,s). \quad (8.23)$$

Remark 8.2.1. *${}^1V$ and ${}^2V$ represent the credit spread accumulated by debtor 1 and 2, respectively. Thus, they do not contain the discount factor.*

The information given by the reward processes allows for the construction of the term structures of forward and spot interest rates and discount factors.

Let ${}^\alpha cs_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right)$ be the expected basis points the debtor α should pay at time t given the present state of the system are defined for $t>1$ by

$${}^\alpha cs_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right):={}^\alpha V_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t\right)-{}^\alpha V_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},t-1\right)$$

and for $t=1$ by

$${}^\alpha cs_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};0,1\right):={}^\alpha V_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2},1\right)-\Psi^1_{\mathbf{i}}(\mathbf{v},0)\ .$$

The expected interest rate i for the debtor α is composed by two parts: the risk free rate r and the contribution due to the rating class occupancy ${}^\alpha cs$

$${}^\alpha i_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right)=r+{}^\alpha cs_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right)\ . \quad (8.24)$$

In other words ${}^\alpha i_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right)$ represents the expected interest rate the debtor α will pay between $t-1$ and t given the present rating class occupancy of the system.

The forward discount factors for the debtor α are obtained by the interest rate using the standard relations

$${}^\alpha v_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right)=\left[1+{}^\alpha i_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right)\right]^{-1}\ . \quad (8.25)$$

Using the forward discount factors we can obtain the spot discount factors for debtor α as

$${}^\alpha v_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t\right)=\prod_{j=0}^{t-1}{}^\alpha v_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};j,j+1\right)\ . \quad (8.26)$$

The credit default spread between the debtors is defined as the difference between their expected interest rate

$$^{12}CS_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right) = {}^{1}cs_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right) - {}^{2}cs_{\binom{i_1}{i_2}}\left(\binom{v_1}{v_2};t-1,t\right) \ . \tag{8.27}$$

^{12}CS represents the expected spread the debtor 1 will pay with respect to debtor 2 due to the credit risk, given the present rating class occupancy.

Bibliography

[1] P. M. Anselone, *Ergodic theory for discrete semi-Markov chains*. Duke Math. J., **27** (1), (1960).

[2] F. Ball and G. F. Yeo, *Lumpability and marginalisability for continuos-time Markov chains*. J. Appl. Prob. **30**, 518-528 (1993).

[3] V. Barbu and N. Limnios, *Semi-Markov Chains and Hidden Semi-Markov Models Toward Applications: Their Use in Reliability and DNA Analysis*, Ed. Springer (2008).

[4] F.E. Benth, M. Groth and R. Kufakunesu. *Valuing Volatility and Variance Swaps for a Non-Gaussian Ornstein-Uhlenbeck Stochastic Volatility Model.* Applied Mathematical Finance, Vol. 14, No. 4, 347-363, (September 2007).

[5] T. R. Bielecki and M. Rutkowski, *Credit Risk: Modeling, Valuation and Hedging*, Ed. Springer (2004).

[6] T.R. Bielecki, J. Jakubowski, A. Vidozzi, L. Vidozzi, *Study of Dependence for Some Stochastic Processes*, Stochastic Analysis and Applications, **26** (2008).

[7] T.R. Bielecki, J. Jakubowski, and M. Niewęgłowski, *Intricacies of dependence between components of multivariate Markov chains: weak Markov consistency and Markov copulae*, electronic paper available at arXiv:1105.2679.

[8] P. Billingsley, *Probability and Measure*, Third Edition, Wiley Series in Probability and Mathematical Statistics (1995).

[9] Björk, T., *Arbitrage Theory in Continuous Time*, Oxford University Press Inc, New York (2004).

[10] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy, **81**(3), (1973).

[11] C.Bluhm, L.Overbeck and C.Wagner *An introduction to Credit Risk Modeling.* Chapman & Hall/CRC Financial Mathematics Series, 2002.

[12] S. Bossu. *Arbitrage Pricing of Equity Correlation Waps*. JPMorgan Equity Derivatives, Working paper (2005).

[13] S. Bossu. *A New Approach For Modelling and Pricing Correlation Swaps*. Equity Structuring - ECD London, Working paper (2007).

[14] D. Brigo and A. Capponi, *Bilateral counterparty risk valuation with stochastic dynamical models and application to Credit Default Swaps*, arXiv 0812:3705v4 (2009).

[15] M. Broadie and A. Jain. *Pricing and Hedging Volatility Derivatives*. The Journal of Derivatives, Vol. 15, No. 3, pp. 7-24 (2008).

[16] M. Broadie and A. Jain. *The Effect of Jumps and Discrete Sampling on Volatility and Variance Swaps*. International Journal of Theoretical and Applied Finance, Vol.11, No.8. pp. 761-797 (2008).

[17] O. Brockhaus and D. Long. *Volatility swap made simple*. Risk, January, 92-96 (2000).

[18] Broze, L., Scaillet, O., Zakoian, J.M., Testing for continuous-time models of the short-term interest rate. *Journal of Empirical Finance*, 2, 199-223 (1995).

[19] P. Carr, H. Geman, D. B. Madan and M. Yor. *Pricing options on realized variance*. Finance Stochast. 9, 453-475 (2005).

[20] P. Carr and R. Lee. *Realized volatility and variance: Options via swaps*. Bloomberg LP and University of Chicago, (2007). Available at: http://math.uchicago.edu/˜rl/OVSwithAppendices.pdf.

[21] P. Carr and R. Lee. *Volatility Derivatives*. Annu. Rev. Financ. Econ., Vol.1, pag. 319-39 (2009).

[22] L. Carty and J. Fons, *Measuring changes in corporate credit quality*, The Journal of Fixed Income, **4**(1), 27–41 (1994).

[23] W. Ching, M.K. Ng, *Markov Chains: Models, Algorithms and Applications*, Ed. Springer (2009).

[24] W. Ching, E.S. Fung and M.K. Ng, *A multivariate Markov chain model for categorical data sequences and its applications in demand predictions*, IMA Journal of Management Mathematics (2003) **13** 187-199.

[25] W. Ching, E.S. Fung and M.K. Ng, *Higher-order multivariate Markov chains and their applications*, Linear Algebra and its Applications (2008) **428** 492-507.

[26] O. Chryssaphinou, M. Karaliopoulou and N. Limnios, *On Discrete Time Semi-Markov Chains and Applications in Words Occurrences*, Communications in Statistics-Theory and Methods (2008) **37** 1306-1322.

[27] Çinlar, E., *Markov Renewal Theory*, Adv. in Appl. Probab. **1** (1969).

[28] Çinlar, E., *Introduction to stochastic processes.*, Prentice Hall, N.Y. (1975).

[29] Corradi, G., Janssen, J., and Manca, R., "Numerical Treatment of homogeneous semi-Markov processes in transient case". *Methodology and Computing in Applied Probability.* v. 6, pp. 233-246 (2004).

[30] Cox, J.C., Ingersoll, J.E., and Ross, S.A., A theory of term structure of interest rates. *Econometrica.* 53, pp. 385-407 (1985).

[31] S. Crépey, M. Jeanblanc and B. Zagari, *Counterparty Risk on a CDS in a Markov Chain Copula Model with Joint Defaults and Stochastic Spreads*, Recent Advances in Financial Engineering World Scientfic (2009).

[32] G. D'Amico, J. Janssen and R. Manca, *Homogeneous discrete time semi-Markov reliability models for credit risk management*, Decisions in Economics and Finance, **28**, 79-93 (2005).

[33] G. D'Amico, J. Janssen and R. Manca, *Valuing Credit Default Swap in a Non-Homogeneous Semi-Markovian Rating Based Model*, Comput. Econ., **29**, pag. 119-138 (2007).

[34] G. D'Amico, J. Janssen, and R. Manca, Semi-Markov Reliability Models with Recurrence Times and Credit Rating Applications, *Journal of Applied Mathematics and Decision Sciences*, Article ID 625712, 17 pages (2009).

[35] G. D'Amico, J. Janssen, and R. Manca, European and American Options: The semi-Markov case, *Physica A*, vol. 388, pp.3181-1194 (2009).

[36] G. D'Amico, J. Janssen and R. Manca, *Initial and Final Backward and Forward Discrete Time Non-Homogeneous Semi-Markov Credit Risk Models*, Methodol. Comput. Appl. Prob., **12**, pag. 215-225 (2010).

[37] G. D'Amico, J. Janssen and R. Manca, *A Non-Homogeneous Semi-Markov Reward Model for the Credit Spread Computation*, Int. Journal of Theo. and Appl. Finance, Vol. 14, No. 2, pag. 1-18 (2011).

[38] G. D'Amico, J. Janssen and R. Manca, *Discrete Time Non-Homogeneous Semi-Markov Reliability Transition Credit Risk Models and the Default Distribution Functions*, Comput. Econ., **38**, 465-481 (2011).

[39] G. D'Amico, J. Janssen, and R. Manca, Duration Dependent Semi-Markov Models, *Applied Mathematical Sciences*, Vol. 5, no. 42, 2097 - 2108 (2011).

[40] G. D'Amico, R. Manca and G. Salvi, *Bivariate semi-Markov Process for Counterparty Credit Risk*, Submitted (2012).
Available at arXiv : 1112.0226v2.

[41] G. D'Amico, R. Manca, G. Salvi : *A semi-Markov Modulated Interest Rate Model*, Submitted (2012).
Available at arXiv : 1210.3164.

[42] G. D'Amico, R. Manca, G. Salvi : *Bivariate semi-Markov rewards model and application to credit spread computation*, Preprint (2012).

[43] J. Da Fonseca, F. Ielpo and M. Grasselli. *Hedging (Co)Variance Risk with Variance Swaps*. Available at SSRN: http://ssrn.com/abstract=1341811 (2009).

[44] M. Dahlquist, S.F. Gray, Regime switching and interest rates in the European monetary system, *Journal of International Economics*, 50, 399,Äì419 (2000).

[45] K. Demeterfi, E. Derman, M. Kamal and J. Zou. *A Guide to Volatility and Variance Swaps*. The Journal of Derivatives (1999).

[46] J. Drissien, P.J. Maenhout and G. Vilkov. *The price of Correlation Risk: Evidence from Equity Options*. The Journal of Finance, Vol. 64, Issue 3, pages 1377-1406 (2009).

[47] Duffie, D., Kan, R., A yield factor model of interest rates. *Math. Finance*, 5, 379-406 (1996).

[48] R.J. Elliott, T.K. Siu and L. Chan. *Pricing Volatility Swaps Under Heston's Stochastic Volatility Model with Regime Switching*. Applied Mathematical Finance, Volume 14, Issue 1, (2007).

[49] R.J. Elliott, T.K. Siu and L. Chan. *Pricing Options Under a Generalized Markov-Modulated Jump-Diffusion Model*. Stochastic Analysis and Applications, 25: 821-843, (2007).

[50] R.J. Elliott and A.V. Swishchuk. *Pricing Options and Variance Swaps in Markov-Modulated Brownian Markets*. In: 'Hidden Markov Model in Finance', Eds. R. Mamon and R. Elliott, Springer (2007).

[51] Elliot, R.J., Hunter, W.C., and Jamieson, B.M., Financial signal processing: a self-calibrating model. *International Journal of Theoretical and Applied Finance*, 4, pp. 567-584 (2001)

[52] R. A. Howard, "Dynamic Probabilistic Systems, Volume II", Ed. Wiley (1971).

[53] S. Howison, A. Rafailidis and H. Rasmussen. *On the pricing and hedging of volatility derivatives*. Applied Mathematical Finance 11, 317-346 (December, 2004).

[54] Hull, J., White, A., Pricing Interest Rate Derivative Securities. *The Review of Financial Studies*, vol. 3, n. 4, pp. 573-592 (1990).

[55] Hunt, J., Devolder, P., Semi-Markov regime switching interest rate models and minimal entropy measure. Physica A, 390(21-22), 3767-3781 (2011).

[56] Hunt, P., Kennedy, J., and Pelsser, A., "Markov-functional interest rate models." *Finance Stochastics,* 4, pp. 391-408 (2000).

[57] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Process*. North-Holland (1988).

[58] Janssen, J., and R. Manca, Numerical solution of non-homogeneous semi-Markov processes in transient case. *Methodology and Computing in Applied Probability,* 3, pp. 271-293 (2001).

[59] J. Janssen and R. Manca, *Applied Semi-Markov Processes*, Ed. Springer (2006).

[60] J. Janssen and R. Manca, *Semi-Markov risk models for Finance, Insurance and Reliability*, Springer, New York, (2007).

[61] A. Javaheri, P. Wilmott and E. Haug. *GARCH and volatility swaps*. Wilmott Technical Article, January, 17p (2002).

[62] L. B. Kolarov and Y. G. Sinai, *Theory of Probability and Random Processes*, Second Edition, Springer-Verlag (2007).

[63] D. Lamberton and B. Lapeyre. *Introduction to Stochastic Calculus Applied to Finance*. Chapman & Hall/CRC (2008).

[64] P. Lévy, *Processus semi-Markoviens*. Proc. of International Congress of Mathematics, Amsterdam, (1954).

[65] N. Limnios and G. Oprişan, *Semi-Markov Process and Reliability*, Ed. Birkäuser (2001).

[66] Mamon, R.S., "A time-varying Markov Chain model of term structure" *Statistics and Probability Letters*, 60 pp. 309-312 (2002).

[67] Mamon, R.S., "Analytic pricing solutions to term structure derivatives in a Markov chain market" *IMA Journal of Management Mathematics*, 15, pp. 243-252 (2004).

[68] R. Manca, *Matrici a più dimensioni*, Quaderno n. 23 dell'Istituto di Matematica della Facoltà di Economia e Commercio di Napoli (Italian).

[69] R. Manca, *Un nuovo tipo di Moltiplicazione tra matrici*, La Ricerca, Anno XXX, n.3 (1979). (Italian).

[70] J. Maskawa, *Multivariate Markov chain modeling for stock markets*, Physica A, **324** (2003).

[71] Norberg, R., "A Time-Continuos Markov Chain Interest Model with Applications to Insurance." *Applied Stochastic Models and Data Analysis*, 11, pp. 245-256 (1995).

[72] Pye, G., "A Markov model of term structure." *Quarterly Journal of Economics,* 25, pp. 60-72 (1966).

[73] M. Pykhtin and S. Zhu, *Global Association of Risk Professionals*, Issue **37**, July/August (2007).

[74] Rebonato, R., Mahal, S., Joshi, M., Buchholz, L.D., Nyholm, K., Evolving Yield Curves in the real-World Measures: A Semi-Parametric Approach. J. Risk 7(3), 29-62 (2005).

[75] G. Salvi and A.V. Swishchuk. *Covariance and Correlation Swaps for Financial Markets with Markov-Modulated Volatilities*. Submitted (2012).

[76] G. Salvi and A.V. Swishchuk. *Modeling and Pricing of Covariance and Correlation Swaps for Financial Markets with Semi-Markov Volatilities*. Submitted (2012).

[77] A. Sepp. *Pricing Options on Realized Variance in the Heston Model with Jumps in Returns and Volatility*. Journal of Computational Finance, 2008, 11(4), 33-70.

[78] S. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer (2004).

[79] W. L. Smith, *Regenerative stochastic processes*, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. (1955).

[80] Stenberg, F., Manca, R., Silvestrov D., An algorithmic approach to discrete time non-homogeneous backward semi-Markov reward processes with an application to disability insurance. Methodol. Comput. Appl. Probab. 9, 497-519 (2007).

[81] A.V. Swishchuk. *Pricing of Variance and Volatility Swaps with semi-Markov volatilities*. Canadian Applied Mathematics Quarterly. Vol. 18, Nu. 4, Winter (2010).

[82] A.V. Swishchuk, R. Cheng, S. Lawi, A. Badescu, H. B. Mekki, A.F. Gashaw, Y. Hua, M. Molyboga, T. Neocleous and Y. Petrachenko. *Price Pseudo-Variance, Pseudo-Covariance, Pseudo-Volatility, and Pseudo-Correlation Swaps-In Analytical Closed-Forms*. Proceedings of the Sixth PIMS Industrial Problems Solving Workshop, PIMS IPSW 6, University of British Columbia, Vancouver, Canada, May 27-31, 2002, pp. 45-55 (2002).

[83] L. Takacs, *Some investigations concerning recurrent stochastic processes of a certain type*. Magyar Tud. Akad. Mat. Kutato Int. Kzl. (1954).

[84] Vasicek, O., "An Equilibrium Characterization of the Term structure." *Journal of Financial Economics*, 5, pp. 177-188 (1977).

[85] A. Vasileiou, and P.-C. G. Vassiliou, *An inhomogeneous semi-Markov model for the term structure of credit risk spreads*, Advances in Applied Probability, **38**, 171-198 (2006).

[86] Wang, J.L., Huang R.J., Model Risks and Surplus Management Under a Stochastic Interest Rate Process. *Journal of Actuarial Practice*, 10, pp. 155-174 (2002).

[87] H. Windcliff, P.A. Forsyth and K.R. Vetzal. *Pricing methods and hedging strategies for volatility derivatives*. Journal of Banking & Finance 30 (2006) 409-431.