Université d'Evry Val d'Essonne

Ecole Doctorale: Sciences et Ingénierie

Laboratoire de Mathématiques et Modélisation d'Evry

THÈSE

presentée et soutenue publiquement le 24 Août 2015

pour l'obtention du grade de

Docteur de l'Université d'Evry-Val-d'Essonne

Discipline: Mathématiques Appliquées

par

Tuyet Mai NGUYEN

Calcul Malliavin pour Chaînes de Markov et Risque de Contrepartie
Malliavin Calculus for Markov Chains and Counterparty Risk

Composition du Jury :

M. Rama CONT	Rapporteur
M. Stéphane CREPEY	Directeur de thèse
M. Laurent DENIS	Directeur de thèse
M. Romuald ELIE	Rapporteur
M. Emmanuel GOBET	Examinateur
Mme. Zorana GRBAC	Examinatrice
M. Pierre HENRY-LABORDERE	Rapporteur

Remerciements

Mes premiers remerciements s'adressent à mes directeurs de thèse, M. Stéphane Crépey et M. Laurent Denis, sans qui cette thèse n'aurait jamais pu voir le jour.

Je suis infiniment reconnaissante envers M. Stéphane Crépey pour la qualité exceptionnelle de son encadrement. Sa disponibilité, son enthousiasme, ses conseils avisés ainsi que son exigence m'ont été très précieux. J'apprécie tous le temps qu'il m'a consacré, tous les opportunités qu'il m'a données pour élargir ma connaissance dans le monde de la recherche. Je lui exprime toute mon admiration, tant sur le plan humain que sur le plan professionnel.

Toute ma gratitude s'adresse également à M. Laurent Denis, qui m'a guidé, partagé ses connaissances et ses expériences précieuses, avec une immense générosité dès les premiers jours de ma thèse. Je n'oublierai pas la gentillesse, les soutiens et la grande confiance qu'il m'a accordé tout au long de cette thèse. Qu'il veuille trouver ici l'expression de ma plus vive reconnaissance.

Je voudrais remercier à M. Rama Cont, M. Romuald Elie, et M. Pierre Henry-Labordère qui, malgré leur emplois du temps très chargés, ont accepté d'être les rapporteurs de ma thèse. Je remercie très vivement M. Emmanuel Gobet, et Mme Zorana Grbac qui ont accepté de faire partie de mon jury de thèse.

Je suis très reconnaissante envers M. Huyen Pham, qui m'a inspiré et guidé vers le chemin de cette thèse. Ma gratitude s'adresse également à Mme. Monique Jeanblanc, qui m'a acceuilli au sein de son équipe et conseillé tout au long de ce chemin. J'exprime également toute ma reconnaissance à M. Andrea Macrina, M. Stéphane Menozzi, M. David Skovmand, et M. Shi Qi Song, avec lesquels j'ai eu l'opportunité d'échanger et de collaborer. Ces expériences ont été pour moi très enrichissantes.

Je tiens à remercier tous les membres du laboratoire LaMME, dont les membres de l'équipe Probabilités et Mathématiques Financières, avec une attention particulière pour Mme. Valérie Gontier-Picot et M. Maouloud Ould Baba qui m'ont toujours apporté leur soutien et leur gentillesse. Merci à tous mes collègues, mes amis doctorants, postdoctorants et maîtres de conférence pour les moments sympas qu'on a vécu ensemble, particulèrement à Dongli et Nam pour avoir partagé avec moi leur expérience.

Enfin, je ne peux conclure sans une pensé toute particulière pour ma famille qui, malgré la distance, a toujours été présente, à ma belle famille qui m'a beaucoup soutenu, et à mon mari qui a toujours été à mes côtés. Cette thèse lui est spécialement dédiée.

Résumé

Cette thèse traite de deux domaines d'analyse stochastique et de mathématiques financières: le calcul Malliavin pour chaînes de Markov (Partie I) et le risque de contrepartie (Partie II).

La partie I a pour objectif l'étude du calcul Malliavin pour chaînes de Markov en temps continu. Il y est présenté deux points : démontrer l'existence de la densité pour les solutions d'une équation différentielle stochastique et calculer les sensibilités des produits dérivés. Dans le premier point, on considère une chaîne de Markov en temps continu comme une généralisation d'un processus de Poisson dont l'intensité de saut et l'état suivant dépendent de l'état actuel. Nous perturbons les temps de saut de la chaîne de Markov conditionnellement au nombre de sauts à un horizon de temps fixe. La perturbation est compensée par un changement de mesure via l'utilisation de Girsanov, on obtient la dérivée directionnelle ainsi qu'une formule d'intégration par partie. Une propriété de densité de l'image de l'énergie en termes de la forme de Dirichlet associée et des opérateurs de gradient et de divergence est établie pour déduire les conditions de l'existence d'une densité pour les solutions d'équations différentielles stochastiques. Cette approche permet le calcul des sensiblités des options asiatiques dans des modèles financiers concrets. Dans le deuxième point, on considère une chaîne de Markov en temps continu comme un processus de Poisson ayant subi un changement de mesure, qui permet le calcul Malliavin via les opérateurs de différences finies pour les processus de Lévy. Cette approche est appliquée au calcul de sensibilités de dérivés de credit.

La partie II traite de sujets d'actualité dans le domaine du risque de marché, à savoir les XVA (ajustements de prix) et la modélisation multi-courbe. Dans un premier temps, nous développons une approche multi-courbe. La modélisation s'exprime par des fonctions rationnelles de processus de Markov. On calibre le modèle à partir des données de swaptions sur LIBOR et montre qu'un modèle multi-courbe rationnelle à deux facteurs log-normaux est suffisant pour s'ajuster aux données de marché. On élucide la relation entre les modèles développés et calibrés sous la mesure risque neutre $\mathbb{Q}$ et ceux sous la mesure historique $\mathbb{P}$. Le modèle de valorisation sous $\mathbb{P}$ est utilisé pour calculer les expositions aux risques éventuellement exigées par la réglementation. Afin de calculer des ajustements de valeur de crédit, comme la CVA, nous modélisons les processus d'intensité de défaut sous la forme de fonctions rationnelles. Nous étoffons notre étude en appliquant les résultats à un contrat de swap de base. Dans un second temps on considère les calculs numériques de XVA. Pour les problèmes non linéaires et de très haute-dimensions, le seul schéma numérique réalisable est purement à terme. Nous comparons deux de ces schémas dans le cadre du calcul des ajustements de risques de crédit et des ajustements de

risques de financement pour des dérivés de crédit, à savoir une expansion linéaire de Monte Carlo et un schéma de séparation de particules.

Abstract

This thesis deals with two areas of stochastic analysis and mathematical finance: Malliavin calculus for Markov chains (Part I) and counterparty risk (Part II).

Part I is devoted to the study of Malliavin calculus for continuous-time Markov chains, in two respects: proving the existence of a density for the solution of a stochastic differential equation and computing sensitivities of financial derivatives. In a first approach, we consider a continuous-time Markov chain as a generalization of a Poisson process with jump intensity and next state depending on the current state. We perturb the jump times of the Markov chain conditionally on the number of jumps up to a fixed time horizon. The perturbation of time is compensated by a Girsanov's measure change, from which we deduce the directional derivation together with an integration by parts formula. An energy image density (EID) property is derived in terms of the associated Dirichlet form and gradient and divergence operators. This property is then applied to deduce conditions for the existence of a density for solutions to stochastic differential equations. This approach also permits the computation of the sensitivities of Asian options in concrete financial models. In a second approach, we consider a continuous-time Markov chain as a measure changed Poisson process, which allows using the Malliavin calculus via finite difference operators for Lévy processes. This is applied to the computation of sensitivities of credit derivatives.

Part II addresses topical issues in interest rates and credit, namely XVA (pricing adjustments) and multi-curve modeling. In a first work, we develop a multi-curve term structure setup in which the modelling ingredients are expressed by rational functionals of Markov processes. We calibrate to LIBOR swaptions data and show that a rational two-factor lognormal multi-curve model is sufficient to match market data with accuracy. We elucidate the relationship between the models developed and calibrated under a risk-neutral measure $\mathbb{Q}$ and their consistent equivalence class under the real-world probability measure $\mathbb{P}$. The consistent $\mathbb{P}$-pricing models are applied to compute the risk exposures which may be required to comply with regulatory obligations. In order to compute counterparty-risk valuation adjustments, such as CVA, we show how default intensity processes with rational form can be derived. We flesh out our study by applying the results to a basis swap contract. The second work regards numerical XVA computations. For nonlinear and very high-dimensional problems, the only feasible numerical schemes are purely forward simulation schemes. We compare two such schemes regarding the computation of counterparty risk and funding valuation adjustments on credit derivatives, namely a linear Monte Carlo expansion and a branching particles schemes.

Contents

Chapter 0

Introduction

This thesis deals with two areas of stochastic analysis and mathematical finance: Malliavin calculus for Markov chains (Part I) and counterparty risk (Part II).

0.1 Part I

Malliavin calculus extends the calculus of variations from functions to stochastic processes, which allows one to compute derivatives of random variables. It was initiated in the late 1970s by the mathematician Paul Malliavin to prove the existence and smoothness of a density for the solution of a stochastic differential equation. It was then applied in financial mathematics to compute the sensitivities of financial derivatives, using the integration by parts formula.

This method first involved the Wiener space and processes based on the Brownian motion (see the historical notes of Bouleau and Hirsch (1991), Ma and Röckner (1992), Nualart (1995), Malliavin (1997), Watanabe (1987), Shigekawa (2004)). However extensions of the Malliavin calculus to the case of stochastic differential equations with jumps were rapidly proposed by Bichteler, Gravereaux and Jacod (1987) and gave rise to an extensive literature. There are three main approaches using *local operators acting on the size of the jumps* (Bichteler et al. (1987), Coquio (1993), Ma and Röckner (2000), Bouleau and Denis (2015) etc.) or *acting on instants of the jumps* (Carlen and Pardoux (1990), Denis (2000), etc.), or using *finite difference operators and the Fock space representation* that exploit similarities between the Poisson space and the Wiener space (Dermoune, Kree, and Wu (1988), Nualart and Vives (1990), Picard (1996), Ishikawa and Kunita (2006), Applebaum (2008), Privault (2009), etc.). In other words, unlike Malliavin calculus on the Wiener space, there is no "natural" choice for the gradient operator on the Poisson space and the different approaches yield different operators. In the first and second approaches, the Dirichlet structure is *local*, i.e. the gradient satisfies the chain rule and the *Energy Image Density* property, which is a powerful criterion, introduced in Bouleau and Hirsch (1991), permitting to establish absolute continuity of the law of Poissonian functionals. The third approach yields a structure which is non local but satisfies a Clark-Ocone formula.

Speaking of applications of the stochastic calculus for jump processes in finance, apart from Lévy processes, continuous-time Markov chains are other examples of semimartingales with jumps not only important in theory but also with vast ap-

plications in diverse fields. Continuous-time Markov chains are distinct from the class of Lévy processes since they have memory, i.e. the probability of jump to the next state may depend on the current state. But they are Markov processes, so it is natural to wonder whether the calculus of variations for jump processes, which is well established for Poisson random measures and Lévy processes, can be adapted to Markov chains. Relatively little attention has been paid to this issue. Biane (1990) and Elliott (1991) introduce and study martingales associated with Markov chains. In addition, Biane (1990) establishes an homogeneous chaos expansion, which is a starting point of the above-mentioned third approach regarding Malliavin calculus with jumps. However, it did not go further due to the complexity of dealing with many martingales associated to Markov chains. Siu (2014) uses perturbations of the jump intensities to obtain an integration by parts formula and thereby a martingale representation.

The first part of this thesis is devoted to the study of Malliavin calculus for continuous-time Markov chains, in two respects: on the one hand, proving the existence of a density for functionals of the Markov chain such as the solution of a stochastic differential equation driven, by it and, on the other hand, computing sensitivities of financial derivatives. Since the set of jump sizes of a Markov chain is discrete, the first approach above is not available. In the line of the second approach, **Chapter 2**, which is based on Denis and Nguyen (2015), considers a continuous-time Markov chain as a generalization of a Poisson process with jump intensity and next state depending on the current state. We follow the second approach of the Malliavin calculus and consider a continuous-time Markov chain as a generalization of a Poisson process whose jump intensity is not constant but depends on the current state, then we perturb the jump times of the Markov chain conditionally to the number of jumps up to a fixed time horizon. The perturbation of time is then compensated by a Girsanov's measure change from which we establish the directional derivation together with an integration by part formula. Then by means of a sequence of well-chosen directions in a space playing the role of a *Cameron-Martin space*, we construct a local Dirichlet structure and the gradient, divergence and carré du champ operators. Then, we prove that it satisfies the Energy Image Density (EID) property, which is then applied to deduce conditions of the existence of a density for solutions to stochastic differential equations. in the last part of this chapter, we apply the Malliavin calculus to the computation of greeks in a concrete model. More precisely, we consider asian type derivatives such that the price of the underlying asset satisfies a stochastic differential equations driven by a Markov chain and compute the delta in terms of directional derivations. Finally,we make some simulations and compute the delta by two methods: the first one using the Malliavin calculus and the second one using the finite difference approach.

Chapter 3 is an illustration of the third approach of the Malliavin calculus. More precisely, we still consider a non-homogeneous continuous-time Markov chain with finite state space. The first step consists in transforming it into a homogeneous Markov chain by a change of probability measure. Then, since a homogeneous Markov chain is naturally associated with a random Poisson measure, we apply the standard Malliavin calculus based on the creation operator (see for example Di Nunno, Øksendal, and Proske (2008)). Then, we apply it to the computation of sensitivities of credit derivatives following the homogeneous-group model (see also

Crépey and Nguyen (2014)), where we have the inverse change of measure, so that we can profit the Clark-Ocone formula in the changed measure but do the simulation under the original measure, which is practically simpler due to many absorbing states.

The ***main contributions*** of Part I are the two constructions of the Malliavin calculus for non homogeneous Markov chains and the applications to solutions of SDEs and to concrete models in finance.

0.2 Part II

The world wide credit crisis and the European sovereign debt crisis have highlighted the native form of credit risk, namely counterparty risk. This is the risk that the counterparties might not live up to the fulfillment of their contractual obligations in a financial transaction. In a bilateral perspective, counterparty risk affects both parties in the contract and should be considered when evaluating a contract via the credit valuation adjustment (CVA, which prices the other party's default risk) and the debt valuation adjustment (DVA, which prices own default risk). In this context, the classical assumption of a locally risk-free asset used for both lending and borrowing is no longer sustainable, which raises a companion issue of proper accounting for funding costs, priced by the funding valuation adjustment (FVA). These adjustments (XVA) need be accounted for both in pricing and regulatory capital.

Taking a look inside, CVA (and similarly DVA) is an option on the future value of the contract with random maturity given by the first to default time of the two parties, whereas the funding cost coefficient depends on the future value of the contract. This leads to a non linear valuation and interdependence between the adjustments, so that they must be computed jointly. The works of Crépey (2012a, 2012b) and recently the book Crépey, Bielecki and Brigo (2014) have constructed a backward stochastic differential equation (BSDE) approach to counteparty risk under funding constraints valuation with all these criteria, where the total valuation adjustment (TVA, i.e. the sum of CVA, DVA and FVA) is the solution of a BSDE. We do not have an explicit solution for such a BSDE but only numerical methods to approximate the solution.

Since August 2007, one has also seen the emergence of a systemic counterparty risk, referring to various significant spreads between quantities that were similar before, such as the Overnight Index Swap (OIS) rate versus the London Interbank Offer Rate (LIBOR). Through its relation with the concept of discounting, this systemic counterparty risk has impacted all derivative markets. Hence, the assumption that all interest rates can be modeled as one (single curve modeling) is no longer sustainable.

Part II deals with the above topical issues. First, in **Chapter 4**, which is based on Crépey, Macrina, Nguyen, and Skovmand (2015), we develop multi-curve interest rate models which extend to counterparty risk models in a consistent fashion. The aim is the pricing and risk management of financial instruments with price models capable of discounting at multiple rates (e.g. OIS and LIBOR) and which allow for corrections in the asset's valuation scheme in order to adjust for counterparty-risk inclusive of credit, debt, and liquidity risk. We thus propose factor-models for

the OIS rate, the LIBOR, and the default intensities of two counterparties involved in bilateral OTC derivative transactions. The three ingredients are characterised by a feature they share in common: the rate and intensity models are all rational functions of the underlying factor processes. Since we have in mind the pricing of assets as well as the management of risk exposures, we also need to work within a setup that maintains price consistency under various probability measures. We will for instance want to price derivatives by making use of a risk-neutral measure $\mathbb{Q}$ while analysing the statistics of risk exposures under the real-world measure $\mathbb{P}$. This point is particularly important when we calibrate the interest rate models to derivatives data, such as implied volatilities, and then apply the calibrated models to compute counterparty-risk valuation adjustments to comply with regulatory requirements. The presented rational models allow us to develop a comprehensive framework that begins with an OIS model, evolves to an approach for constructing the LIBOR process, includes the pricing of fixed-income assets and model calibration, analyses risk exposures, and concludes with a credit risk model that leads to the analysis of counterparty-risk valuation adjustments (XVA).

The issue of how to model multi-curve interest rates and incorporate counterparty-risk valuation adjustments in a pricing framework has motivated much research. For instance, research on multi-curve interest rate modelling is presented in Kijima, Tanaka, and Wong (2009), Kenyon (2010), Henrard (2007, 2010, 2014), Bianchetti (2010), Mercurio (2010b, 2010a, 2010c), Fujii (2011, 2010), Moreni and Pallavicini (2014), Bianchetti and Morini (2013), Filipović and Trolle (2013) or Crépey, Grbac, Ngor and Skovmand (2015). Pricing models with rational form have also appeared before. Flesaker and Hughston (1996) pioneered such pricing models and in particular introduced the so-called rational log-normal model for discount bond prices. Further related studies include Rutkowski (1997), Döberlein and Schweizer (2001) and Hunt and Kennedy (2004), Brody and Hughston (2004), Hughston and Rafailidis (2005), Brody, Hughston and Mackie (2012), Akahori, Hishida, Teichmann and Tsuchiya (2014), Filipović, Larsson and Trolle (2014), Macrina and Parbhoo (2014) or Nguyen and Seifried (2014). However, as far as we know, our work is the first to apply rational pricing models in a multi-curve setup, along with Nguyen and Seifried (2014) who develop a rational multi-curve model in the spirit of Rogers (1997) based on a multiplicative spread, and it is the only rational pricing work dealing with XVA computations. We shall see that, despite the simplicity of these models, they perform surprisingly well when comparing to other, in principle more elaborate, proposals such as Crépey et al. (2015) or Moreni and Pallavicini (2013, 2014).

In **Chapter 5**, which is based on Crépey and Nguyen (2015), we endeavour to study counterparty risk on credit derivatives. For nonlinear and very high-dimensional problems, any numerical scheme based, even to some extent, on dynamic programming, such as purely backward deterministic PDE schemes, but also forward/backward simulation/regression BSDE schemes, are ruled out by the curse of dimensionality (see e.g. Crépey (2013)). The only feasible numerical schemes are purely forward simulation schemes, such as the linear Monte Carlo expansion of Fujii and Takahashi (2012a,2012b) or the branching particles scheme of Henry-Labordère (2012). We compare two such schemes regarding the computation of counterparty risk and funding valuation adjustments on credit derivatives. In both

cases, a Markov structure is required to justify the use of the method, but is too heavy for practical use. Instead, fast and exact simulation and forward pricing schemes are available based on the dynamic copula features of credit portfolio models. However, the branching particles scheme requires a stronger dynamic copula property where, conditionally on the past, the future again obeys to some copula structure. In the end, it's only in one of two proposed models that the two schemes can be run and compared numerically. They show similar performance but the branching scheme requires a fine-tuning that can only be achieved by a preliminary knowledge on the solution (which can be provided by linear approximation).

The ***main contributions*** of Part II are: on the one hand, the demonstration of the practical abilities of the linear-rational models of chapter 4 in terms of calibration and post-crisis multicurve and counterparty risk modeling, whereas most previous work on linear-rational or pricing kernels models had stayed more theoretical; on the other hand, the mathematical and algorithmic understanding of the realm of validity of the two numerical schemes of chapter 5, the adaptations required for fitting these schemes to our credit problems and the comparison of their empirical performance.

Part I

Malliavin calculus for Markov chains

Chapter 1

Preliminaries

1.1 Basic facts on Poisson random measures, Lévy processes and Dirichlet forms

1.1.1 Poisson random measures

Here we adopt the definition of Poisson random measures that was introduced in Cont and Tankov (2003).

Definition 1.1.1. *(Poisson random measure)*
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $E \subset \mathbb{R}^d$ and ν a given positive Radon measure on $(E, \mathcal{E})$. A Poisson random measure on E with intensity ν is an integer-valued radom measure

$$\begin{aligned} M : \Omega \times \mathcal{E} &\rightarrow \mathbb{N} \\ (\omega, A) &\mapsto M(\omega, A), \end{aligned}$$

such that

a) For every $\omega \in \Omega$, $M(\omega, .) = M(\omega) : \mathcal{E} \rightarrow \mathbb{N}$ is an integer-valued Radon measure on E. This measure depends on a random source $\omega \in \Omega$, which is the reason why M is called a random measure.

b) For every $A \in \mathcal{E}$ such that $\nu(A) < +\infty$, $M(., A) = M(A) : \Omega \rightarrow \mathbb{N}$ is an integer-valued random variable on Ω following the Poisson law with parameter $\nu(A)$, i.e.

$$\forall k \in \mathbb{N},\ \mathbb{P}(M(A) = k) = e^{-\nu(A)} \frac{(\nu(A))^k}{k!}.$$

As a consequence $\nu(A) = \mathbb{E}[M(A)]$.

c) If $A_1, \ldots, A_n$ are pathwise disjoint sets in $\mathcal{E}$ then the random variables $M(A_1), \ldots, M(A_n)$ are independent.

Since $M(\omega)$ is a Radon measure on E, we can define integrals with respect to this measure. First, for simple functions $f = \sum_{i=1}^n c_i \mathbf{1}_{A_i} : E \rightarrow \mathbb{R}_+$ where $c_i \geq 0$ and $A_i \in \mathcal{E}$ are disjoint, we define $M(f) = \sum_{i=1}^n c_i M(A_i)$, then $M(f)$ is a random variable on Ω with expectation $\mathbb{E}[M(f)] = \sum_{i=1}^n c_i \nu(A_i)$. Next, for

positive measurable function $f : E \to \mathbb{R}_+$, we define $M(f) = \lim_{n\to\infty} M(f_n)$ where $f_n \to f$ is an increasing sequence of simple functions. By the monotone convergence theorem, $M(f)$ is then a random variable on Ω with values in $\mathbb{R}_+ \cup \{\infty\}$ with expectation $\mathbb{E}[M(f)] = \nu(f)$. For a real-valued measurable function $f : E \to \mathbb{R}$, we can decompose f into a positive part and a negative part: $f = f^+ - f^-$ where $f^+ = max(f, 0)$ and $f^- = max(-f, 0)$ are positive measurable functions with values in $\mathbb{R}_+$. If the function f verifies

$$\nu(|f|) = \int_E |f|\nu(dx) < \infty \tag{1.1.1}$$

then the positive random variables $M(f^+)$, $M(f^-)$ have finite expectations. In particular, $M(f^+)$, $M(f^-)$ are almost surely finite, so we can define $M(f) := M(f^+) - M(f^-)$. $M(f)$ is thus a random variable with expectation

$$\mathbb{E}[M(f)] = \nu(f) = \int_E f\nu(dx).$$

We set $\tilde{M} = M - \nu$, *the compensated random measure.* Similarly, under the condition (1.1.1), one can define the integral of f with respect to $\tilde{M}$. The following lemma is classical:

Lemma 1.1.1. *Let $f \in L^1(\nu) \cap L^2(\nu)$, then*

$$\mathbb{E}[(\tilde{M}(f))^2] = \int_E f^2\, d\nu, \quad (*)$$

as a consequence, the mapping $f \in L^1(\nu) \cap L^2(\nu) \mapsto \tilde{M}(f)$ can be extended uniquely to a continuous mapping from $L^2(\nu)$ into $L^2(\mathbb{P})$, it is still denoted by $f \mapsto \tilde{M}(f)$ and relation $(*)$ holds.

In our framework, we consider $E = \mathbb{R}_+ \times \mathcal{Z}$, where $\mathcal{Z} \subset \mathbb{R}^d \backslash \{0\}$, with $\mathcal{E}$ is the family of Borel sets in E, and the random measure M defined on the canonical probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is the configuration space, the space of measures which are countable sums of Dirac measures on E

$$\Omega = \{\omega = \sum_{i\in I} \epsilon_{(t_i,z_i)} : I \subset \mathbb{N};\, \forall i \in I,\, t_i \in \mathbb{R}_+,\, z_i \in \mathcal{Z};\, \forall i \neq j,\, (t_i, z_i) \neq (t_j, z_j)\}.$$

M is the canonical map $M(\omega) = \omega$, $\forall \omega \in \Omega$.

1.1.2 Lévy Processes

See Bertoin (1998) Chapters 0 to III, Cont and Tankov (2003) Chapters 3 and 4 for the details of Lévy processes. We introduce here the definition in Cont and Tankov (2003).

Definition 1.1.2. *A cadlag stochastic process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^d$ such that $X_0 = 0$ is called a Lévy process if it possesses the following properties*

1. *Independent increments: for every increasing sequence of times $t_0, \cdots, t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \cdots, X_{t_n} - X_{t_{n-1}}$ are independent.*

2. *Stationary increments: the law of* $X_{t+h} - X_t$ *does not depend on* t.

3. *Stochastic continuity:* $\forall \epsilon > 0, \lim_{h\to 0} \mathbb{P}(|X_{t+h} - X_t| \geq \epsilon) = 0$.

Example 1.1.1. 1. Brownian motions and Poisson processes are two typical examples of continuous and jump Lévy processes.

2. A stochastic process X_t defined as

$$X_t = \sum_{i=1}^{N_t} Y_i$$

where jump sizes Y_i are i.i.d. with distribution μ and (N_t) is a Poisson process with intensity λ, independent from $(Y_i)_{i\geq 1}$, is called a *compound Poisson process*. Every compound Poisson process is a Lévy process, and every Lévy process with piecewise constant sample paths is a compound Poisson process.

If $(X_t)_{t\geq 0}$ is a compound Poisson process with intensity λ and jump size distribution μ, then its jump measure J_X defined, for every $0 \leq t_1 \leq t_2$, $A \in \mathcal{B}(\mathbb{R}^d \backslash 0)$, by

$$J_X([t_1, t_2] \times A) = \#\{(t, X_t - X_{t-}) : t \in [t_1, t_2], X_t - X_{t-} \in A\},$$

(the number of jump times of X between t_1 and t_2 such that their jump sizes are in A) is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}^d \backslash \{0\}$ with intensity $\nu(dt \times dx) = \lambda dt \mu(dx)$.

1.1.3 Dirichlet forms

We shall only consider local, symmetric Dirichlet forms admitting a *carré du champ operator* and defined on locally compact spaces. Our main reference is Bouleau and Hirsch (1991). The locality assumption implies that the form we consider satisfies some algebraic properties. The fact that the Dirichlet form admits a carré du champ means that, in some sense, it is regular and that the domain of the generator associated to the underlying semigroup contains a dense algebra. The main point is that if a Dirichlet form is local and admits a carré du champ, one can develop a very useful and intuitive functional calculus, as in the case of the historical energy form.

We now briefly recall the main objects we consider, all the details can be found in Bouleau and Hirsch (1991). We consider $(E, \mathcal{E}, \nu)$ a measured space such that ν is σ-finite measure.

A symmetric *closed form* is a non-negative quadratic form e defined on a dense subspace $\mathbf{d} \subset L^2(\nu)$ such that $\mathbf{d}$ equipped with the norm

$$||u||^2_{\mathbf{d}} = ||u||^2_{L^2(\nu)} + e[u]$$

is a Hilbert space. We denote by $e[u, v]$ the bilinear form associated. Then

$$e[u, v] = \frac{1}{4}(e[u + v] - e[u - v]),\ \forall u, v \in \mathbf{d}.$$

If normal contractions operate on e, in the sense that if $\varphi : \mathbb{R} \to \mathbb{R}$ is a contraction ($|\varphi(x) - \varphi(y)| \leq |x - y|$, $\forall x, y \in \mathbb{R}$) such that $\varphi(0) = 0$ then

$$u \in \mathbf{d} \quad \Rightarrow \quad \varphi(u) \in \mathbf{d} \text{ and } e[\varphi(u)] \leq e[u],$$

then we say that $(\mathbf{d}, e)$ is a *Dirichlet form*. To construct a Dirichlet form, normally, we construct a "pre-Dirichlet form" on a set of "smooth" function and then extend it by density and closability properties as in Example 1.1.3. If a Dirichlet form $(\mathbf{d}, e)$ is *local* ($\forall u \in \mathbf{d}\ \ e[|u+1|-1] = e[u]$), we say that it admits a *carré du champ* γ (cf. Bouleau and Hirsch (1991) p17) if and only if there exists a continuous operator γ from $\mathbf{d} \times \mathbf{d}$ into $L^1(\nu)$ such that

$$e[u, v] = \int \gamma[u, v] d\nu, \quad \forall u, v \in \mathbf{d}.$$

We shall simply denote $\gamma[u] = \gamma[u, u]$. For all $u, v \in \mathbf{d}^n$, all F, G of class $\mathcal{C}^1$ on $\mathbb{R}^n$ with bounded derivatives and such that $F(0) = G(0) = 0$, we have the following functional calculus for carré du champ operator

$$\gamma[F(u), G(v)] = \sum_{ij} \partial_i F(u) \partial_j G(v) \gamma[u_i, v_j]. \tag{1.1.2}$$

The space $\mathbf{d}$ equipped with the scalar product $(\|.\|^2_{L^2(\nu)} + e[.,.])^{\frac{1}{2}}$ is a Hilbert space, we assume that it is separable. As a consequence (see Bouleau and Hirsch (1991) ex.5.9 p.242), we can construct a linear operator which has the same properties as a derivation operator that we call the *gradient*. More precisely, there exists a separable Hilbert space H and a continuous linear map, D, from $\mathbf{d}$ into $L^2(E, \nu; H)$ such that

- $\forall u \in \mathbf{d}$, $\|D(u)\|^2_H = \gamma[u]$.
- If $F : \mathbb{R} \to \mathbb{R}$ is Lipschitz $\forall u \in \mathbf{d}$, $D(F \circ u) = (F' \circ u) D(u)$, where F' the derivate of F defined almost everywhere with respect to the Lebesgue measure.
- If F is $\mathcal{C}^1$ and Lipschitz from $\mathbb{R}^d$ into $\mathbb{R}$, then

$$\forall u = (u_1, \cdots, u_d) \in \mathbf{d}^d,\ D(F \circ u) = \sum_{i=1}^{d} (\partial_i F \circ u) D(u_i). \tag{1.1.3}$$

One of the main interest of Dirichlet Structures is that they permit to establish existence of density for random variables without integration by parts formula (cf Bouleau and Hirsch (1986 1991), Denis (2000)). It based on the following property that we call *energy image density property*, or EID for short.

For any integer $d \geq 1$, let $\mathcal{B}(\mathbb{R}^d)$ denote the Borelian σ-field on $\mathbb{R}^d$ an λ^d the Lebesgue measure on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$. For any measurable function u, we denote by $u_*\nu$ the image measure of ν by u.

Definition 1.1.3. *The Dirichlet structure $(E, \mathcal{E}, \nu, \mathbf{d}, \gamma)$ is said to satisfy* EID *if for all $d \in \mathbb{N}^*$ and all $U = (U_1, \cdots, U_d) \in \mathbf{d}^d$,*

$$U_*[(\det \gamma[U, U^t]) \cdot \nu] \ll \lambda^d$$

where det *denotes the determinant and $\gamma[U, U^t]$ is nothing but the matrix*

$$\gamma[U, U^t] = (\gamma[U_i, U_j])_{1 \leq i,j \leq d}.$$

The EID property is always satisfied in the case $d = 1$ if the Dirichlet form is local and admits a carré du champ operator (cf. Bouleau (1984) Thm 5 and Corol 6). In 1986, it was conjectured that EID were always true for local Dirichlet forms with carré du champ (Bouleau and Hirsch (1986) p251). To better illustrate the definitions, we introduce a classical example below.

Example 1.1.2. Let $E = \mathbb{R}^d$, $\mathcal{E} = \mathcal{B}(\mathbb{R}^d)$, the family of Borel sets in $\mathbb{R}^d$, and $\nu = dx$, the Lebesgue measure on $\mathbb{R}^d$. We consider the Sobolev space

$$H^1(\mathbb{R}^d) = \{u \in L^2(\mathbb{R}^d, dx); \frac{\partial u}{\partial x_i} \in L^2(\mathbb{R}^d), i = 1, \cdots, n\}.$$

It is well-known that $H^1(\mathbb{R}^d)$ is dense on $L^2(\mathbb{R}^d)$ and $H^1(\mathbb{R}^d)$ equipped with the following norm

$$||u||_1^2 = ||u||^2_{L^2(\mathbb{R}^d)} + ||u||^2_{H^1(\mathbb{R}^d)}$$

is a Hilbert space, where

$$||u||^2_{H^1(\mathbb{R}^d)} = \int_{\mathbb{R}^d} \sum_{i=1}^d \left(\frac{\partial u}{\partial x_i}\right)^2 dx.$$

Hence, $e = ||.||^2_{H^1(\mathbb{R}^d)} = \sum_{i=1}^d \int_{\mathbb{R}^d} \left(\frac{\partial}{\partial x_i}\right)^2 dx$ is a closed form defined on $\mathbf{d} = H^1(\mathbb{R}^d)$. If $F : \mathbb{R} \to \mathbb{R}$ is a contraction with $F(0) = 0$ then $F(u) \in H^1(\mathbb{R}^d)$ for all $u \in H^1(\mathbb{R}^d)$ and $||F(u)||^2_{H^1(\mathbb{R}^d)} \leq ||u||^2_{H^1(\mathbb{R}^d)}$, which deduces that the closed form ($\mathbf{d}$,e) defined above is a Dirichlet form. This Dirichlet form admit a carré du champ

$$\gamma[u, v] = \sum_{i,j=1}^d \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j}$$

and a gradient taking values in $\mathbb{R}^d$

$$Du := \nabla u = (\frac{\partial u}{\partial x_1}, \cdots, \frac{\partial u}{\partial x_d}), \forall u \in H^1(\mathbb{R}^d).$$

Moreover, for all $u \in H^1(\mathbb{R}^d)$ and $A \subset \mathbb{R}$ with Lebesgue measure 0, we have

$$\int_{\mathbb{R}^d} \mathbf{1}_A(u(x))||\nabla u||^2 dx = 0,$$

which shows that the Dirichlet structure $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), dx, \mathbf{d}, \gamma)$ satisfies EID for dimension 1.

Example 1.1.3. More generally, we consider $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and $\nu = k(x)dx$ where k belongs to $H^1_{loc}(\mathbb{R}^d)$ and k^{-1} is locally bounded. Let $\xi = (\xi_{ij})_{1 \leq i,j \leq d} : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ be a symmetric Borel function that is locally elliptic on an open set of $\mathbb{R}^d$ and

$$\mathbf{d} = \{u \in L^2(kdx) : \sum_{i,j} \xi_{ij} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} \in L^1(kdx)\}.$$

Then the bilinear form e defined by

$$\forall u, v \in \mathbf{d},\ e[u,v] = \int_{\mathbb{R}^d} \sum_{i,j=1}^{d} \xi_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} k(x) dx$$

is a local Dirichlet form on $L^2(kdx)$ which admits a carré du champ operator γ given by

$$\forall u, v \in \mathbf{d},\ \gamma[u,v] = \sum_{i,j=1}^{d} \xi_{ij}(x) \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j}.$$

Moreover, the Dirichlet structure $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), kdx, \mathbf{d}, \gamma)$ satisfies the EID property (see Bouleau and Denis (2009)).

The EID property has been established on the Wiener space, for the Dirichlet structure associated with the Ornstein-Uhlenbeck operator and for some other examples by Bouleau and Hirsch (1991) Chap. II § 5 and Chap. V Example 2.2.4, but the EID conjecture being at present neither proved nor refuted in full generality, it has to be established in each particular framework.

On the Poisson space, it was first proved by Coquio (1993) in the case where the intensity is the Lebesgue on an open domain and then has been proved in a more general case (see Bouleau and Denis (2009) Section 2 Thm 2 and Section 4).

The EID property on the *Wiener space* is a powerful tool widely used now (cf. Nualart and Quer-Sardanyons (2007), Nualart (2010), Chighouby, Djehichez and Mezerdix (2009)). It is also recently developped and adapted for the *Poisson space* in Bouleau and Denis (2015). One of the aim of our work is to study the EID property in the context of *Markov chains*.

1.2 Malliavin calculus via chaos decomposition

1.2.1 Multiple integrals and chaos decomposition

Let us assume that ν is diffused, and that the space $(E, \mathcal{E})$ is separable and Hausdorff, what means (cf. Dellacherie and Meyer (1975) Chap. I §9) that the σ-field $\mathcal{E}$ is generated by a countable family and that the atoms are the points of E i.e. $(\forall A \in \mathcal{E},\ (x \in A \Leftrightarrow y \in A)) \Rightarrow x = y$. This implies that for all $x \in E$, $\{x\}$ belongs to $\mathcal{E}$ and that the diagonal of $E \times E$ is measurable (cf. Dellacherie and Meyer (1975) Chap. I Thm 12). This implies that all the diagonals of the form $\{(x_1, \ldots, x_n) : \exists i \neq j\ x_i = x_j\}$ are measurable and ν-negligible.

We are going to define multiple integral with respect to a Poisson random measure, a generalization of the the integral defined in Section 1.1.1. The construction is mostly based on Bouleau and Denis (2015) who follow the idea of the beautiful paper of Itô (1951), but in the context of a Poisson random measure.

A real-valued function f defined on $(E, \mathcal{E}, \nu)^n$ is said *elementary* if it is a weighted sum of indicator functions of sets of the form $A_1 \times \cdots \times A_n$ where for all i, $\nu(A_i)$ is finite.

If f is elementary, we can re-number the sets $A_1, A_2, \ldots, A_k$ which appear in the expression of f and even assume that they are disjoint (since $\nu^{\times n}$ does not charge

the diagonals) in such a way that

$$f(x_1,\dots,x_n)=\sum a_{i_1,\cdots,i_n}1_{A_{i_1}}(x_1)\cdots 1_{A_{i_n}}(x_n)\ \nu^{\times n}-a.e.$$

where the $A_1,A_2,\dots,A_k$ are disjoint, with finite mesure and $a_{i_1,\cdots,i_n}=0$ if two indexes among $i_1,\dots,i_n$ are equal.

Since for any $n\in\mathbb{N}^*$, $\nu^{\times n}$ does not charge the diagonal in X^n, the set of elementary functions is a dense subvector space in $L^2(\nu^{\times n})$.

Then we put

$$I_n(f)=\sum a_{i_1,\cdots,i_n}\tilde{M}(A_{i_1})\cdots\tilde{M}(A_{i_n})$$

I_n is clearly a linear map on the set of elementary functions, moreover

$$I_n(f)=I_n(\tilde{f}) \tag{1.2.4}$$

where $\tilde{f}=\frac{1}{n!}\sum_{\sigma\in\mathcal{S}}f(x_{\sigma(1)},\dots,x_{\sigma(n)})$ is the symetrized of f, this is due to the fact that the products of the form $\tilde{M}(A_{i_1})\cdots\tilde{M}(A_{i_n})$ are commutative. We also have

$$\langle I_n(f),I_m(g)\rangle_{L^2(\mathbb{P})}=\delta_{m,n}\,n!\langle\tilde{f},\tilde{g}\rangle_{L^2(\nu^{\times n})}. \tag{1.2.5}$$

This equality is not so obvious, we refer to the article of Itô (1951) for a rigorous proof of it. As a consequence of (1.2.5)

$$\|I_n(f)\|^2_{L^2(\mathbb{P})}=n!\|\tilde{f}\|^2_{L^2(\nu^{\times n})}\le n!\|f\|^2_{L^2(\nu^{\times n})}$$

so, by a density argument, I_n can be extended uniquely to $f\in L^2(\nu^{\times n})$ and we have $\forall f\in L^2(\nu^{\times n})$, $\forall g\in L^2(\nu^{\times m})$

$$I_n(f)=I_n(\tilde{f})\qquad\qquad\langle I_n(f),I_m(g)\rangle_{L^2(\mathbb{P})}=\delta_{m,n}\,n!\langle\tilde{f},\tilde{g}\rangle_{L^2(\nu^{\times n})}. \tag{1.2.6}$$

Let $\tilde{L}^2(\nu^n)$ be the subset of all symmetric function in $L^2(\nu^n)$. The following theorem and also the introduction of Malliavin derivative in the next subsection are based on Di Nunno, Øksendal, and Proske (2008).

Theorem 1.2.1. *(Wiener-Itô chaos expansion for Poisson random measure) Let $F\in L^2(\mathbb{P})$ be a $\mathcal{F}_T$-measurable random variable. Then F admits a representation*

$$F=\sum_{n=0}^{\infty}I_n(f_n)$$

via a unique sequence of elements $f_n\in\tilde{L}^2(\nu^n),n=1,2,\cdots$. Here we set $I_0(f_0):=f_0$ for the constant values $f_0\in\mathbb{R}_0$. Moreover, we have that

$$||F||^2_{L^2(\mathbb{P})}=\sum_{n=0}^{\infty}n!||f_n||^2_{L^2(\nu^n)}<\infty. \tag{1.2.7}$$

1.2.2 Malliavin derivative

Definition 1.2.4. *The stochastic Sobolev space $\mathbb{D}_{1,2}$ consists of all $\mathcal{F}_T$-measurable random variables $F \in L^2(\mathbb{P})$ with chaos expansion*

$$F = \sum_{n=0}^{\infty} I_n(f_n),\ f_n \in \tilde{L}^2(\nu^n), \tag{1.2.8}$$

satisfying the convergence criterion

$$||F||^2_{\mathbb{D}_{1,2}} := \sum_{n=0}^{\infty} nn!||f_n||^2_{L^2(\nu^n)} < \infty.$$

In comparing with (1.2.7), we see that $\mathbb{D}_{1,2}$ is strictly contained in the space of all $\mathcal{F}_T$-measurable random variables in $L^2(\mathbb{P})$.

Definition 1.2.5. *(Malliavin derivative) We define the operator $D : F \in \mathbb{D}_{1,2} \subset L^2(\mathbb{P}) \mapsto DF \in L^2(\mathbb{P} \times \nu)$ by*

$$D_{t,z}F = \sum_{n=0}^{\infty} nI_{n-1}(f_n(.,t,z)) \tag{1.2.9}$$

for all $F \in \mathbb{D}_{1,2}$ of the form (1.2.8)*. Here $I_{n-1}(f_n(.,t,z))$ means that the $(n-1)$-fold interated integral of f_n is regarded as a function of its $(n-1)$ first pairs of variables $(t_1,z_1),\cdots,(t_{n-1},z_{n-1})$, while a final pair (t,z) is kept as a parameter. $D_{t,z}F$ is called the Malliavin derivative of F at (t,z).*

The Malliavin derivative defined as above is a closed operator.

1.2.3 Creation operator

The creation operator with its companion, the annihilation operator are well known in quantum mechanics.They are one of the main ingredients for the lent particle method, a new approach to study the Malliavin calculus on Poisson spaces in Bouleau and Denis (2015). We recall here the creation operator ϵ^+ for the purpose of giving an intuitive understanding of the Malliavin derivative defined via the chaos decomposition on the configuration space:

$$\forall x \in E,\ \forall \omega \in \Omega,\ \epsilon_x^+(\omega) = \omega \mathbf{1}_{\{x \in supp\omega\}} + (\omega + \epsilon_x)\mathbf{1}_{\{x \notin supp\omega\}}, \tag{1.2.10}$$

where $supp\omega = \{x_i, i \in I\}$ if $\omega = \sum_{i\in I} \epsilon_{x_i}$, $I \subset \mathbb{N}$, $x_i \in E\, \forall i \in I$, $x_i \neq x_j\, \forall i \neq j$. It is extended to the functionals $H : \Omega \times E \to \mathbb{R}$ by

$$\epsilon^+ H(\omega, x) = H(\epsilon_x^+ \omega, x),\ \forall \omega \in \Omega, x \in E.$$

We recall the following lemma whose proof can be found in Nualart and Vives (1990).

Lemma 1.2.2. *Let $u \in L^2(\nu^n)$, then*

$$\epsilon^+ I_n(u^{\otimes n}) = I_n(u^{\otimes n}) + nuI_{n-1}(u^{\otimes n-1}).$$

Lemma 1.2.2 shows that if u is elementary, to add a mass at point x_0 to $M(dx)$ adds $nu(x_0)I_{n-1}(u^{\otimes n-1})$ to $I_n(u^{\otimes n})$. Now for any function $f_n \in \tilde{L}^2(\nu^n)$, we obtain by a density argument

$$(\epsilon^+ - I)I_n(f_n) = nI_{n-1}(f_n) \quad \mathbb{P} \times \nu - a.s,$$

where I represents the indentity operator, and the mutiple integral on the right hand side acts only on the $n-1$ first arguments of f_n. Thanks to this relation, (1.2.9) can be rewritten as

$$D_{t,z}F = \sum_{n=0}^{\infty}(\epsilon^+_{(t,z)} - I)I_n(f_n) = (\epsilon^+_{(t,z)} - I)\sum_{n=0}^{\infty} I_n(f_n) = (\epsilon^+_{(t,z)} - I)F. \qquad (1.2.11)$$

This equality gives a sense to the Malliavin derivative: taking Malliavin derivative of the random variable F at (t,z) is to measure the sensitivity of F when adding a particle (t,z) (adding a jump of size z at time t).

1.2.4 Chain rule and product rule

Theorem 1.2.2. *(Chain rule) Let $F \in \mathbb{D}_{1,2}$ and let ϕ be a real continuous function on $\mathbb{R}$. Suppose that $\phi(F) \in L^2(\mathbb{P})$ and $\phi(F + D_{t,z}F) \in L^2(\mathbb{P} \times \nu)$. Then $\phi(F) \in \mathbb{D}_{1,2}$ and*

$$D_{t,z}\phi(F) = \phi(F + D_{t,z}F) - \phi(F).$$

Proof. By applying (1.2.11), we obtain

$$D_{t,z}\phi(F) = (\epsilon^+_{(t,z)} - I)\phi(F) = \epsilon^+_{(t,z)}\phi(F) - \phi(F) = \phi(\epsilon^+_{(t,z)}F) - \phi(F) = \phi(F + D_{t,z}F) - \phi(F).$$

□

Similarly, we obtain the product rule for Malliavin derivative:

Theorem 1.2.3. *Let $F, G \in \mathbb{D}_{1,2}$. Then $FG \in \mathbb{D}_{1,2}$ and*

$$D_{t,z}(FG) = FD_{t,z}G + GD_{t,z}F + D_{t,z}FD_{t,z}G.$$

Proof. By applying (1.2.11), we have

$$\begin{aligned} D_{t,z}(FG) &= (\epsilon^+_{(t,z)} - I)(FG) = \epsilon^+_{(t,z)}(FG) - FG = (\epsilon^+_{(t,z)}F)(\epsilon^+_{(t,z)}G) - FG \\ &= (F + D_{t,z}F)(G + D_{t,z}G) - FG = FD_{t,z}G + GD_{t,z}F + D_{t,z}FD_{t,z}G. \end{aligned}$$

□

The rest of this section is a recall from Di Nunno, Øksendal, and Proske (2008).

1.2.5 Malliavin derivative and Skorohod Integral

Definition 1.2.6. *(Skorohod integral) Let $X = X(t,z), 0 \le t \le T, z \in \mathbb{R}_0$, be an stochastic process such that $X(t,z)$ is an $\mathcal{F}_T$-measurable random variable for all $(t,z) \in [0,T] \times \mathbb{R}_0$ and*

$$\mathbb{E}\left[\int_0^T \int_{\mathbb{R}_0} X^2(t,z)\nu(dz)dt\right] < \infty, \forall (t,z) \in [0,T] \times \mathbb{R}_0. \tag{1.2.12}$$

Then for each (t,z), the random variable $X(t,z)$ has an expansion of the form

$$X(t,z) = \sum_{n=0}^{\infty} I_n(f_n(.,t,z)), \text{ where } f_n(.,t,z) \in \tilde{L}^2(\nu^n). \tag{1.2.13}$$

Let $\tilde{f}_n(t_1,z_1,\cdots,t_n,z_n,t_{n+1},z_{n+1})$ be the symmetrization of $f(t_1,z_1,\cdots,t_n,z_n,t,z)$ as a function of the $n+1$ pairs $(t_1,z_1),\cdots,(t_n,z_n),(t,z) = (t_{n+1},z_{n+1})$. Suppose that

$$\sum_{n=0}^{\infty}(n+1)!||\tilde{f}_n||^2_{L^2(\nu^{n+1})} < \infty. \tag{1.2.14}$$

Then we say that X is Skorohod integrable, write $X \in Dom(\delta)$. We define the Skohorod integral $\delta(X)$ of X with respect to $\tilde{M}$ by

$$\delta(X) = \int_0^T \int_{\mathbb{R}_0} X(t,z)\tilde{M}(\delta t, dz) := \sum_{n=0}^{\infty} I_{n+1}(\tilde{f}_n). \tag{1.2.15}$$

The Skorohod integral defined as above is an extension of the Itô integral in the sense that if X is predictable process satisfying (1.2.12), then X is both Itô and Skorohod integrable with respect to $\tilde{M}$ and these two integrals are equal.

Theorem 1.2.4. *(Duality formula) Let $X(t,z)$, $t \in [0,T]$, $z \in \mathbb{R}$ be Skorohod integrable and $F \in \mathbb{D}_{1,2}$. Then*

$$\mathbb{E}\left[\int_0^T \int_{\mathbb{R}_0} X(t,z)D_{t,z}F\nu(dz)dt\right] = \mathbb{E}\left[F\int_0^T \int_{\mathbb{R}_0} X(t,z)\tilde{M}(\delta t, dz)\right].$$

Theorem 1.2.5. *(Integration by parts) Let $X(t,z), t \in [0,T], z \in \mathbb{R}$ be a Skorohod integrable stochastic process and $F \in \mathbb{D}_{1,2}$ such that the product $X(t,z)\epsilon_{t,z}F$, $t \in [0,T]$, $z \in \mathbb{R}$ is Skorohod integrable. Then*

$$F\int_0^T \int_{\mathbb{R}_0} X(t,z)\tilde{M}(\delta t, dz) = \int_0^T \int_{\mathbb{R}_0} X(t,z)\epsilon_{t,z}F\tilde{M}(\delta t, dz) + \int_0^T \int_{\mathbb{R}_0} X(t,z)D_{t,z}F\nu(dz)dt.$$

By using this integration by part formula, the closability of the Skorohod integral is obvious.

1.2.6 The Clark-Ocone formula

Theorem 1.2.6. *Let $F \in \mathbb{D}_{1,2}$. Then*

$$F = \mathbb{E}[F] + \int_0^T \int_{\mathbb{R}_0} \mathbb{E}[D_{t,z}F|\mathcal{F}_t]\tilde{M}(dt,dz) \tag{1.2.16}$$

where we have chosen a predictable version of the conditional expectation process $\mathbb{E}[D_{t,z}F|\mathcal{F}_t]$, $t \ge 0$.

1.3 Malliavin calculus by perturbations of time

In addition to the Malliavin calculus introduced in Section 1.2, which based on a different finite operator, there are other ways to construct a Malliavin-type calculus for a Poisson random measure on $\mathbb{R}^+ \times \mathcal{Z}$, where $\mathcal{Z}$ is a measured space, with compensator $dt \times d\nu$. The idea is to define a "real" derivation operator called the "gradient" which allow to be able to prove the existence of densities for functionals and also to define an analog of the Skorohod integral. In general, there are two main ways to do so.

If $\mathcal{Z}$ is an open subset of $\mathbb{R}^d$ and ν is the Lebesgue measure, the popular way is to "differentiate" with respect to the size of jumps (see Bichteler, Gravereaux and Jacod (1987)). But when the space of jump size is not rich enough to take the derivation, for example, when $\mathcal{Z}$ is discrete like the case of Poisson process, we are obliged to "differentiate" with respect to the jump times as did Carlen and Pardoux (1990) for the Poisson space and then Denis (2000) generalize for Poisson random measure thanks to the construction of a Dirichlet structure related to the Poisson random measure. For the purpose of our work, that is to generalize Malliavin calculus for Markov chains of finite state, we recall here some main ideas in Carlen and Pardoux (1990) and Denis (2000).

Let $(N_t)_{t\geq 0}$ be a Poisson process in its configuration space $(\Omega, \mathcal{F}, \mathbb{P})$ with natural filtration $(\mathcal{F}_t)_{t\geq 0}$. So Ω is the set of all maps $\omega : [0,1] \to \mathbb{N}$ such that $\omega(0) = 0$, ω is increasing and right continuous, and has finitely many jumps of size one. We have the sequence of jump times

$$0 < T_1(\omega) < T_2(\omega) < \cdots < T_n(\omega) < \cdots .$$

Let $\mathcal{H}$ be the subspace of $L^2([0,1])$ orthogonal to the constant functions. For a fixed bounded function $m \in \mathcal{H}$ and $\epsilon > 0$ small enough, we define a reparametrization of $[0,1]$ by $\tau_\epsilon(t) = t + \epsilon \int_0^t m_s ds$. This reparametrization does not change the number and the order of jump times in $[0,1]$. We denote $\mathcal{T}_\epsilon$ is the transformation $(\mathcal{T}_\epsilon(w))(t) = w(\tau_\epsilon(t))$ and $\mathbb{P}^\epsilon$ is the image of $\mathbb{P}$ through this transformation. As shown in Carlen and Pardoux (1990), $\mathbb{P}^\epsilon$ is absolutely continuous with respect to $\mathbb{P}$ with density

$$\frac{d\mathbb{P}^\epsilon}{d\mathbb{P}} = \prod_{i=1}^{N_1}(1 + \epsilon m(T_i)).$$

Next, the directional derivative is defined by the sensitivity of random variables with respect to this reparametrization of time. More precisely, we define

$$\mathbb{D}_m^0 = \{F \in L^2(\Omega) : L^2(\Omega) - \lim_{\epsilon\to 0} \frac{1}{\epsilon}(\mathcal{T}_\epsilon F - F) \text{ exists}\}.$$

Then, for $F \in \mathbb{D}_m^0$, we define $D_m F$ as the above limit. For example, $T_i \in \mathbb{D}_m^0$ and $D_m T_i = -\int_0^{T_i} m(t)dt$. As a consequence, $\mathbb{D}_m^0$ contains the class $\mathcal{S}$ of "smooth" functions on $L^2(\Omega)$. A random variable F belongs to $\mathcal{S}$ if there exists $a \in \mathbb{R}$, $m \in \mathbb{N}^*$ and for any $n \in \{1, \cdots, m\}$, a smooth function with bounded derivatives of any order $f_n : \mathbb{R}^n \to \mathbb{R}$ such that

$$F = a\mathbf{1}_{\{N_1=0\}} + \sum_{n=1}^{m} f_n(T_1, \cdots, T_n)\mathbf{1}_{\{N_1=n\}}.$$

Thanks to the definition of "real derivative" type, the operator D_m maintains properties of a real derivative, such as product rule and chain rule, etc. Moreover, D_m is a closed operator, so we can densely extend it to its closed extension.

The gradient operator $D : L^2(\Omega) \to L^2(\Omega; \mathcal{H})$ is defined compatibly with the directional derivatives defined above. That means

$$\int_0^1 D_t F m(t) dt = D_m F.$$

For $F \in \mathcal{S}$, $F = f(T_1, \cdots, T_n)$, we have $D_t F = -\sum_{j=1}^n \frac{\partial f}{\partial t_j}(T_1, \cdots, T_n)(\mathbf{1}_{[0,T_j]}(t) - T_j)$. Again, D is a closable densely defined operator from $L^2(\Omega)$ to $L^2(\Omega; \mathcal{H})$.

Let δ denote the adjoint of D. Then we have the integration by part formula which comes from the duality between δ and D:

$$\mathbb{E}[\delta(u)F] = \mathbb{E}\int_0^1 u_t D_t F dt, \forall F \in \mathbb{D}^{1,2}.$$

In particularly when u is predictable, the divergence operator δ coincides with the Skohorod integral:

$$\delta(u) = \int_0^1 u_t dN_t.$$

Next, we can infer the existence of a density from the non-degeneracy of the Malliavin covariance matrix: *Let $F \in \mathbb{D}^{1,2}$ and $B \subset \mathbb{N} - \{0\}$ be such that $\int -0^1 |D_t F|^2 dt > 0$ a.s. on $A = \{N_1 \in B\}$ then $(\mathbf{1}_A \mathbb{P})F^{-1}$, the image by F of the restriction of $\mathbb{P}$ to A, is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}$.*

Denis (2000) proposed another criterion for the existence of the density with respect to the Lebesgue measure, by introducing a Dirichlet structure with the "carré du champ" operator and the Dirichlet form firstly defined on $\mathcal{S}$ by

$$\begin{aligned} \gamma(F) &= \sum_{n=1}^m \mathbf{1}_{\{N_1=n\}} \sum_{i=1}^n \left(\frac{\partial}{\partial t_i} f_n(T_1, \cdots, T_n) \right)^2, \\ e(F) &= \mathbb{E}[\gamma(F)], \end{aligned}$$

and then densely extend to its closure to d to obtain a local Dirichlet form with the "carré du champ" γ. By the energy image density property, for all $F \in d$, $F * (\gamma(F).\mathbb{P})$, the image of $\gamma(F).\mathbb{P}$ by F, is absolutely continuous with respect to the Lebesgue measure.

This approach gives also a criterion of the existence of density in multidimensional case and then is extended for Poisson random measure. However, this approach does not base on a real derivative, so one can not develop an extension of Itô's integral and calculus as the case of Carlen and Pardoux (1990) or Bichteler, Gravereaux and Jacod (1987).

1.4 Continuous-time Markov chains

The aim for study Markov chains in our work is to study portfolio credit risk models, so we recall here some basic properties of Markov chains in the view of these

applications. The main reference for this section is Bielecki, Crépey, and Herbertsson (2009). Given $C = (C_t)_{t\geq 0}$ be a Markov chain in an underlying *probability space* $(\Omega, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ with a finite *state space* $\mathcal{K} = \{k_1, k_2, \cdots, k_l\}$, where $(\mathcal{F}_t)_{t\geq 0}$ is the natural filtration of the Markov chain $(C_t)_{t\geq 0}$. Then each path of C is a right-continuous function from $[0, \infty)$ to $\mathcal{K}$.

A two-parameter family $\mathcal{P}(t,s) = (p_{ij}(t,s))_{1\leq i,j\leq l}, t, s \in \mathbb{R}_+, t \leq s$, of stochastic matrices is called the family of *transition probability matrices* of the Markov chain C under $\mathbb{P}$ if, for every $t, s \in \mathbb{R}_+, t \leq s$,

$$\mathbb{P}(C_s = k_j | C_t = k_i) = p_{ij}(t,s), \quad \forall 1 \leq i, j \leq l.$$

In particlular, we have $\mathcal{P}(t,t) = Id$ for every $t \in \mathbb{R}_+$.

Under some regularity conditions for the family $\mathcal{P}(t,s)$, $t, s \in \mathbb{R}_+, t \leq s$, we can define the time-dependent *transition intensity matrix* (or *infinitesimal generator matrix*) $\Lambda(t) = (\lambda_{ij}(t))_{1\leq i,j\leq l}$ through the formula

$$\lambda_{ij}(t) = \lim_{h\downarrow 0} \frac{p_{ij}(t, t+h) - \delta_{ij}}{h}. \tag{1.4.17}$$

It is obvious by the definition that $\lambda_{ij}(t) \geq 0$ for arbitrary $i \neq j$, and

$$\lambda_{ii}(t) = \lim_{h\downarrow 0} \frac{p_{ij}(t, t+h) - 1}{h} = -\lim_{h\downarrow 0} \frac{\sum_{j=1, j\neq i}^{l} p_{ij}(t, t+h)}{h} = -\sum_{j=1, j\neq i}^{l} \lambda_{ij}(t). \tag{1.4.18}$$

For every $1 \leq i \neq j \leq l$, the term $\lambda_{ij}(t)$ defined by (1.4.17) gives the intensity that the Markov chain move from state k_i to state k_j at t. From (1.4.18), the term $-\lambda_{ii}(t)$ is nothing but the intensity of the Markov chain move to another state at time t.

In case of a time-homogeneous Markov chain C, the family of transition probability matrix does not depend on the starting point t, but only depends on the difference $s - t$, so we can define the family of transition probability matrix as the one-parameter family $\mathcal{P}(t) = (p_{ij}(t))_{1\leq i,j\leq l} := \mathcal{P}(0,t)$, $t \in \mathbb{R}_+$. Then we have

$$\mathbb{P}(C_{s+t} = k_j | C_s = k_i) = p_{ij}(t), \, \forall 1 \leq i, j \leq l.$$

This family satisfies the Chapman-Kolmogorov equation, namely,

$$\mathcal{P}(t+s) = \mathcal{P}(t)\mathcal{P}(s) = \mathcal{P}(s)\mathcal{P}(t), \quad \forall s, t \in \mathbb{R}_+.$$

Let $u_0 = (u_0(i))_{1\leq i\leq l} := (\mathbb{P}(C_0 = k_i))_{1\leq i\leq l}$ denote the initial probability distribution for the Markov chain under $\mathbb{P}$. Likewise, let $u_t = (u_t(i))_{1\leq i\leq l} := (\mathbb{P}(C_t = k_i))_{1\leq i\leq l}$ be the probability distribution of C at time $t \in \mathbb{R}_+$. We have

$$u_{t+s} = u_0\mathcal{P}(t+s) = u_t\mathcal{P}(s) = u_s\mathcal{P}(t), \quad \forall s, t \in \mathbb{R}_+.$$

Next, if we impose that the family $\mathcal{P}(.)$ is right-continuous at time $t = 0$, which implies also the differentiability of the family at $t = 0$, then the Markov chain C admits an *infinitesimal generator matrix* (which, in this case, does not depend on time) $\Lambda = (\lambda_{ij})_{1\leq i,j\leq l}$, where each λ_{ij} represents the intensity of transition from the state k_i to the state k_j. The infinitesimal generator matrix Λ is also called the *intensity matrix* of the Markov chain C. This matrix uniquely determines all relevant probabilistic properties of a time-homogeneous Markov chain and moreover, $C_t - \int_0^t \sum_{j=1}^{l} j\lambda_{C_u j} du$ is an $\mathcal{F}_t$-martingale.

Chapter 2

Malliavin calculus for Markov chains using perturbations of time

2.1 Introduction

The aim of this chapter is to construct a Dirichlet structure associated to a time-homogeneous Markov chain with finite state space in the spirit of Carlen and Pardoux (1990) who obtained criteria of density for stochastic differential equations (SDEs) driven by a Poisson process. More precisely, we develop a Malliavin calculus on the canonical space by "derivating" with respect to the jump times of the Markov chain, the main difficulty is that the times of jumps of Markov chain we consider are no more distributed according to an homogeneous Poisson process.
Extensions of Malliavin calculus to the case of SDEs with jumps have been soon proposed and gave rise to an extensive literature. The approach is either dealing with local operators acting on the size of the jumps (cf Bichteler, Gravereaux and Jacod (1987), Coquio (1993), Ma and Röckner (2000) etc.) or acting on the instants of the jumps (cf Carlen and Pardoux (1990), Denis (2000)) or based on the Fock space representation of the Poisson space and finite difference operators (cf Nualart and Vives (1990), Picard (1996), Ishikawa and Kunita (2006) etc.). Let us mention that developping a Malliavin Calculus or a Dirichlet structure on the Poisson space is not the only way to prove the absolutely continuity of the law of the solution of SDE's driven by a Lévy process, see for example Nourdin and Simon (2006), or the recent works of Bally and Clément (2011) who consider a very general case.
In this chapter we consider a Markov chain and we construct "explicitly" a *gradient operator* and a local Dirichlet form. With respect to the Malliavin analysis or the integration by part formula approach , what brings the Dirichlet forms approach is threefold: a) The arguments hold under only Lipschitz hypotheses, e.g. for density of solutions of stochastic differential equations cf Bouleau and Hirsch (1991), this is due to the celebrated property that contractions operate on Dirichlet forms and Émile Picard's iteration scheme may be performed under the Dirichlet norm. b) A general criterion exists, the energy image density property (EID), proved on the Wiener space for the Ornstein-Uhlenbeck form, and in several other cases (but still a conjecture in general since 1986 cf Bouleau and Hirsch (1986)), which provides an

efficient tool for obtaining existence of densities in stochastic calculus. c) Dirichlet forms are easy to construct in the infinite dimensional frameworks encountered in probability theory (cf Bouleau and Hirsch (1991) Chap.V) and this yields a theory of error propagation, especially for finance and physics cf Bouleau (2003).
Moreover, since the gradient operator may be calculated easily, this permits to make numerical simulations, for example in order to compute greeks of an asset in a market with jumps.
The plan of the chapter is as follows. In Section 2, we describe the probabilistic framework and introduce the Markov chain. Then in Section 3, we introduce the directional derivative w.r.t. an element of the Cameron-Martin space and give some basic properties. The next section is devoted to the construction of the local Dirichlet structure and the associated operators namely the gradient and divergence, and also to the establishment of an integration by parts formula. In Section 5, we prove that this Dirichlet form satisfies the *Energy Image Density (EID)* property that we apply in Section 6 to get a criterion of density for solution of SDEs involving the Markov chain. In the last section, we show that this Malliavin calculus may be applied to compute some greeks in finance.

2.2 Probability space

Let $C = (C_t)_{t\geq 0}$ be a time-homogeneous Markov chain, with finite state space $\mathcal{K} = \{k_1, k_2, \cdots, k_l\}$. The one-step transition probability matrix is $\mathcal{P} = (p_{ij})_{1\leq i,j\leq l}$, and the infinitesimal generator matrix $\Lambda = (\lambda_{ij})_{1\leq i,j\leq l}$. We assume that $(C_t)_{t\geq 0}$ has a finite number of jumps over a bounded horizon of time so that we consider that it is defined on the canonical probability space $(\Omega, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ where Ω is the set of $\mathcal{K}$-valued right continuous maps $w : [0, \infty) \to \mathcal{K}$ starting at $w(0) = c_0 \in \mathcal{K}$ such that there exists sequences $c_1, c_2, \cdots \in \mathcal{K}$ and $0 = t_0 < t_1 < \cdots$ with

$$w(t) = \sum_{i=0}^{\infty} c_i 1_{[t_i, t_{i+1}[}(t). \tag{2.2.1}$$

The filtration $(\mathcal{F}_t)_{t\geq 0}$ we consider is the natural filtration of the Markov chain $(C_t)_{t\geq 0}$ satisfying the usual hypotheses.
Let $(T_n)_{n\in\mathbb{N}}$ denote the sequence of successive jump times of C and Z_n denote the position of C at time T_n. More explicitly, for any $n \in \mathbb{N}$, the random variables T_n and Z_n are defined as follow

$$\begin{cases} T_0 = 0, Z_0 = c_0, \\ T_n = inf\{t > T_{n-1} : C_t \neq Z_{n-1}\},\ Z_n = C_{T_n}. \end{cases}$$

We have $0 = T_0 < T_1 < T_2 < \cdots$, and $T_n \to \infty$ when $n \to \infty$ almost surely. These are two well-known properties:

1. $\mathbb{P}(T_n - T_{n-1} > t | Z_{n-1} = k_i) = e^{\lambda_{ii} t}$.

2. $\mathbb{P}(Z_n = j | Z_{n-1} = k_i) = p_{ij} = -\dfrac{\lambda_{ij}}{\lambda_{ii}}$.

Let $u_n = (u_n^i)_{1\leq i\leq l}$ be the distribution of Z_n, i.e., $u_n^i = \mathbb{P}(Z_n = k_i), i = 1, \cdots, l$. From the second property and the law of total probability we have

$$u_n^j = \sum_{i=1}^{l} \mathbb{P}(Z_n = k_j | Z_{n-1} = k_i)\mathbb{P}(Z_{n-1} = k_i) = \sum_{i=1}^{l} p_{ij} u_{n-1}^i.$$

Hence, $u_n = \mathcal{P}u_{n-1} = \mathcal{P}^2 u_{n-2} = \cdots = \mathcal{P}^n u_0$. The first property means that conditionally on the position $Z_{n-1} = k_i$ at the jump time T_{n-1}, the random time that elapses until the next jump has an exponential probability law with parameter $-\lambda_{ii} > 0$. So conditionally on the positions $Z_1 = c_1, \cdots, Z_{n-1} = c_{n-1}$, the increments $\tau_1 = T_1, \tau_2 = T_2 - T_1, \cdots, \tau_n = T_n - T_{n-1}$ are independently exponential distributed with parameters $\lambda_1 = -\lambda_{c_0 c_0}, \cdots, \lambda_n = -\lambda_{c_{n-1}c_{n-1}}$. Moreover, we have the following proposition:

Proposition 2.2.1. *Conditionally to $Z_1 = c_1, \cdots, Z_{n-1} = c_{n-1}$, $(T_1, T_2, \cdots, T_n)$ has a probability density function on $\mathbb{R}^n$ given by*

$$\prod_{i=1}^{n} \lambda_i e^{-\lambda_i (t_i - t_{i-1})} \mathbf{1}_{\{0<t_1<\cdots<t_n\}}(t_1, \cdots, t_n).$$

Proof. For every bounded measurable function $f : \mathbb{R}^n \to \mathbb{R}$, we have

$$\begin{aligned}
& \mathbb{E}[f(T_1, T_2, \cdots, T_n) | Z_1 = c_1, Z_2 = c_2, \cdots, Z_{n-1} = c_{n-1}] \\
= \; & \mathbb{E}[f(\tau_1, \tau_1 + \tau_2, \cdots, \tau_1 + \tau_2 + \cdots + \tau_n) | Z_1 = c_1, Z_2 = c_2, \cdots, Z_{n-1} = c_{n-1}] \\
= \; & \iint_{t_i>0} f(t_1, t_1 + t_2, \cdots, t_1 + t_2 + \cdots + t_n) \prod_{i=1}^{n} \lambda_i e^{-\lambda_i t_i} dt_1 dt_2 \cdots dt_n \\
= \; & \iint_{0<t_1<t_2<\cdots<t_n} f(t_1, t_2, \cdots, t_n) \prod_{i=1}^{n} \lambda_i e^{-\lambda_i (t_i - t_{i-1})} dt_1 dt_2 \cdots dt_n.
\end{aligned}$$

□

We can also compute the law of $\tau_n = T_n - T_{n-1}$ in the same way as Z_n

$$\mathbb{P}(\tau_n > t) = \sum_{i=1}^{l} \mathbb{P}(T_n - T_{n-1} > t | Z_{n-1} = k_i)\mathbb{P}(Z_{n-1} = k_i) = \sum_{i=1}^{l} e^{\lambda_{ii} t} u_{n-1}^i. \quad (2.2.2)$$

Let $(N_t)_{t\geq 0}$ be the process that counts the number of jumps of the Markov chain $(C_t)_{t\geq 0}$ up to t

$$N_t = \sum_{n\geq 1} \mathbf{1}_{\{T_n \leq t\}}. \quad (2.2.3)$$

With some abuse of notation, we define the function $\lambda : \mathcal{K} \to \mathbb{R}$ such that $\lambda(k_i) = -\lambda_{ii}$, $\forall 1 \leq i \leq l$. Then the process $\lambda(C_t) = \sum_{n\geq 0} \lambda(Z_n) \mathbf{1}_{\{T_n \leq t < T_{n+1}\}}$ is the intensity of $(N_t)_{t\geq 0}$. The three processes $(C_t)_{t\geq 0}, (N_t)_{t\geq 0}$ and $(\lambda(C_t))_{t\geq 0}$ have the same jump times. More importantly, the compensated process defined by

$$\tilde{N}_t = N_t - \int_0^t \lambda(C_u) du \quad (2.2.4)$$

is an $\mathcal{F}_t$-martingale. Indeed, conditionally on the positions $Z_1, Z_2, \cdots, Z_{n-1}$, $\int_0^{t\wedge T_n} \lambda(C_u)du$ is the compensator of the process $N_{t\wedge T_n}$. Hence for every $0 \leq s \leq t$,

$$\mathbb{E}[N_{t\wedge T_n} - N_{s\wedge T_n} - \int_{s\wedge T_n}^{t\wedge T_n} \lambda(C_u)du|\mathcal{F}_s, Z_1 = c_1, \cdots, Z_{n-1} = c_{n-1}] = 0,$$

which deduces

$$\mathbb{E}[N_{t\wedge T_n} - N_{s\wedge T_n} - \int_{s\wedge T_n}^{t\wedge T_n} \lambda(u)du|\mathcal{F}_s] = \sum_{c_1,\cdots,c_{n-1}\in\mathcal{K}} \mathbb{P}(Z_1 = c_1, \cdots, Z_{n-1} = c_{n-1})$$

$$\times \quad \mathbb{E}[N_{t\wedge T_n} - N_{s\wedge T_n} - \int_{s\wedge T_n}^{t\wedge T_n} \lambda(u)du|\mathcal{F}_s, Z_1 = c_1, \cdots, Z_{n-1} = c_{n-1}] = 0.$$

Let n tend to $+\infty$ we receive $\mathbb{E}[N_t - N_s - \int_s^t \lambda(u)du|\mathcal{F}_s] = 0$, so the process $(\tilde{N}_t)_{t\geq 0}$ is an $\mathcal{F}_t$-martingale.

2.3 Directional differentiation

In this section, we will consider the Markov chain (C_t) defined on the filtered probability space $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T})$ where $\mathcal{F} = \mathcal{F}_T$ for a fixed time horizon $0 < T < \infty$. We apply the same approach as in Carlen and Pardoux (1990) to define the directional derivative using the reparametrization of time with respect to a function in a Cameron-Martin space.
Let $\mathcal{H}$ be the closed subspace of $L^2([0,T])$ orthogonal to the constant functions, i.e.,

$$\mathcal{H} = \{m \in L^2([0,T]) \text{ such that } \int_0^T m(s)ds = 0\}.$$

We denote $\hat{m} = \int_0^{\cdot} m(s)ds$ for every $m \in \mathcal{H}$, then $\hat{m}(0) = \hat{m}(T) = 0$. In a natural way , $\mathcal{H}$ inherits the Hilbert structure of $L^2([0,T])$ and we denote by $\| \|_{\mathcal{H}}$ and $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ the norm and the scalar product on it. From now on in this section, we fix a function $m \in \mathcal{H}$. The condition $\int_0^T m(s)ds = 0$ ensures that the change of intensity that we are about to define simply shifts the jump times without affecting the total number of jumps. Let define

$$\tilde{m}_\epsilon(s) = \begin{cases} \frac{-1}{3\epsilon} & \text{if } m(s) \leq \frac{-1}{3\epsilon}, \\ m(s) & \text{if } \frac{-1}{3\epsilon} \leq m(s) \leq \frac{1}{3\epsilon}, \\ \frac{1}{3\epsilon} & \text{if } m(s) \geq \frac{1}{3\epsilon}, \end{cases}$$

and $m_\epsilon \in \mathcal{H}$ such that $m_\epsilon(s) = \tilde{m}_\epsilon(s) - \frac{1}{T}\int_0^T \tilde{m}_\epsilon(s)ds$. Then we have again $\int_0^T m_\epsilon(s)ds = 0$, and $\frac{1}{3} \leq 1 + \epsilon m_\epsilon(s) \leq \frac{5}{3}$ (since $-\frac{1}{3\epsilon} \leq \tilde{m}_\epsilon(s) \leq \frac{1}{3\epsilon}$). Moreover, $||m - m_\epsilon||_{\mathcal{H}} \to 0$ as $\epsilon \to 0$. We define the reparametrization of time with respect to m_ϵ as follow

$$\tau_\epsilon(s) = s + \epsilon\hat{m}_\epsilon(s),\ s \geq 0.$$

Notice that $\tau_\epsilon(0) = 0, \tau_\epsilon(T) = T$, and $\frac{\partial\tau_\epsilon}{\partial s}(s) = 1 + \epsilon m_\epsilon(s) > 0$, so the number and the order of jump times between 0 and T remain unchanged. Let $\mathcal{T}_\epsilon : \Omega \to \Omega$ be the map defined by

$$(\mathcal{T}_\epsilon(w))(s) = w(\tau_\epsilon(s)) \text{ for all } w \in \Omega,$$

$\mathcal{T}_\epsilon F = F \circ \mathcal{T}_\epsilon$ for every $F \in L^2(\Omega)$, and $\mathbb{P}^\epsilon$ be the probability measure $\mathbb{P}\mathcal{T}_\epsilon^{-1}$. We denote

$$\mathbb{D}_m^0 = \{F \in L^2(\Omega) : \frac{\partial \mathcal{T}_\epsilon F}{\partial \epsilon}|_{\epsilon=0} = \lim_{\epsilon \to 0} \frac{1}{\epsilon}(\mathcal{T}_\epsilon F - F) \text{ in } L^2(\Omega) \text{ exists}\}.$$

For $F \in \mathbb{D}_m^0$, $D_m F$ is defined as the above limit.

Example 2.3.4. Now we give some examples of random variables whose directional derivatives can be computed directly from the definition.

1. The random variables Z_i do not change under $\mathcal{T}_\epsilon$, so $Z_i \in \mathbb{D}_m^0$ and $D_m Z_i = 0$.

2. Let $w \in \Omega$ have the form (2.2.1). Then

$$(\mathcal{T}_\epsilon(w))(s) = w(\tau_\epsilon(s)) = \sum_{i=0}^{\infty} c_i 1_{[t_i, t_{i+1}[}(\tau_\epsilon(s)) = \sum_{i=0}^{\infty} c_i 1_{[\tau_\epsilon^{-1}(t_i), \tau_\epsilon^{-1}(t_{i+1})[}(s), \tag{2.3.5}$$

which deduces $\mathcal{T}_\epsilon T_i(w) = T_i(\mathcal{T}_\epsilon(w)) = \tau_\epsilon^{-1}(t_i) = \tau_\epsilon^{-1}(T_i(w))$. Hence $\mathcal{T}_\epsilon T_i = \tau_\epsilon^{-1} \circ T_i = U_i$ where U_i is the random variable satisfying $\tau_\epsilon \circ U_i = U_i + \epsilon \hat{m}_\epsilon(U_i) = T_i$. We define the random variables $\bar{T}_i = T_i \wedge T$, $\bar{U}_i = U_i \wedge T$. Then we also have $\mathcal{T}_\epsilon \bar{T}_i = \tau_\epsilon^{-1} \circ \bar{T}_i = \bar{U}_i$ and $\tau_\epsilon \circ \bar{U}_i = \bar{U}_i + \epsilon \hat{m}_\epsilon(\bar{U}_i) = \bar{T}_i$. Hence,

$$\frac{1}{\epsilon}(\mathcal{T}_\epsilon \bar{T}_i - \bar{T}_i) = \frac{1}{\epsilon}(\bar{U}_i - (\bar{U}_i + \epsilon \hat{m}_\epsilon(\bar{U}_i))) = -\hat{m}_\epsilon(\bar{U}_i).$$

Reminding that $m_\epsilon \to m$ in $\mathcal{H}$ as $\epsilon \to 0$, we can expect that $\bar{U}_i \to \bar{T}_i$ and $\hat{m}_\epsilon(\bar{U}_i) \to \hat{m}(\bar{T}_i)$ in $L^2(\Omega)$ as $\epsilon \to 0$. Indeed,

$$\begin{aligned}
&\left|\frac{1}{\epsilon}(\mathcal{T}_\epsilon \bar{T}_i - \bar{T}_i) + \hat{m}(\bar{T}_i)\right| \leq \left|-\hat{m}_\epsilon(\bar{U}_i) + \hat{m}_\epsilon(\bar{T}_i)\right| + \left|\hat{m}(\bar{T}_i) - \hat{m}_\epsilon(\bar{T}_i)\right| \\
&\leq \left|\int_{\bar{U}_i}^{\bar{T}_i} m_\epsilon(s) ds\right| + \int_0^{\bar{T}_i} |m(s) - m_\epsilon(s)| ds \\
&\leq \left(|\bar{T}_i - \bar{U}_i| \int_0^T m_\epsilon^2(s) ds\right)^{\frac{1}{2}} + \left(\bar{T}_i \int_0^T (m(s) - m_\epsilon(s))^2 ds\right)^{\frac{1}{2}} \\
&\leq \left(\epsilon \left|\int_0^{\bar{U}_i} m_\epsilon(s) ds\right|\right)^{\frac{1}{2}} ||m_\epsilon||_{\mathcal{H}} + T^{\frac{1}{2}} ||m_\epsilon - m||_{\mathcal{H}} \\
&\leq \left(\epsilon \left(\bar{U}_i \int_0^T m_\epsilon^2(s) ds\right)^{\frac{1}{2}}\right)^{\frac{1}{2}} ||m_\epsilon||_{\mathcal{H}} + ||m_\epsilon - m||_{\mathcal{H}} \\
&\leq \epsilon^{\frac{1}{2}} T^{\frac{1}{4}} ||m_\epsilon||_{\mathcal{H}}^{\frac{3}{2}} + T^{\frac{1}{2}} ||m_\epsilon - m||_{\mathcal{H}} \to 0 \ as \ \epsilon \to 0.
\end{aligned}$$

Therefore, $\bar{T}_i \in \mathbb{D}_m^0$, and $D_m \bar{T}_i = -\hat{m}(\bar{T}_i)$.

3. Since the number of jumps of $(C_t)_{t \in [0,T]}$ does not change after the reparametrization of time, we have $\mathcal{T}_\epsilon N_T = N_T$. Hence $N_T \in \mathbb{D}_m^0$, and $D_m N_T = 0$.

Remark 2.3.1. If the assumption $\int_0^T m(s)ds = 0$ was relaxed, the number of jumps in $[0,T]$ would change, which does not ensure that $N_T \in \mathbb{D}_m^0$. Indeed, by using (2.3.5), we have

$$\mathcal{T}_\epsilon N_T(w) = N_T(\mathcal{T}_\epsilon(w)) = \sum_{i\geq 1} \mathbf{1}_{\{\tau_\epsilon^{-1}(t_i)\leq T\}} = \sum_{i\geq 1} \mathbf{1}_{\{t_i\leq \tau_\epsilon(T)\}}, \tag{2.3.6}$$

so $\mathcal{T}_\epsilon N_T = \sum_{i\geq 1} \mathbf{1}_{\{T_i\leq \tau_\epsilon(T)\}}$. Put $\bar{\lambda} = max_{1\leq i\leq l}\lambda_{ii}$, then from (2.2.2) and notice that $\bar{\lambda} < 0$, we have

$$\mathbb{P}(T_n - T_{n-1} \leq t) = 1 - \sum_{i=1}^{l} e^{\lambda_{ii}t} u_{n-1}^i \geq 1 - e^{\bar{\lambda}t} > 0 \text{ for every } n \geq 1, t > 0. \tag{2.3.7}$$

Without loss of generality, we can assume that $\int_0^T m(s)ds > 0$. Then $\tau_\epsilon(T) > T$ for $\epsilon > 0$. From (2.3.6) and (2.3.7), we obtain

$$\begin{aligned}
&\mathbb{E}\left[\frac{1}{\epsilon}(\mathcal{T}_\epsilon N_T - N_T)\right]^2 = \frac{1}{\epsilon^2}\mathbb{E}\left[(\mathcal{T}_\epsilon N_T - N_T)^2\right] \\
&\geq \frac{1}{\epsilon^2}\mathbb{E}\left[(\mathcal{T}_\epsilon N_T - N_T)^2 | \mathcal{T}_\epsilon N_T - N_T \geq 1\right]\mathbb{P}(\mathcal{T}_\epsilon N_T - N_T \geq 1) \\
&\geq \frac{1}{\epsilon^2}\mathbb{P}(T_{N_T+1} - T_{N_T} \leq \tau_\epsilon(T) - T) \geq \frac{1}{\epsilon^2}\left(1 - e^{\bar{\lambda}(\tau_\epsilon(T)-T)}\right) \to \infty \text{ as } \epsilon \downarrow 0.
\end{aligned}$$

Thus, in this case we would have $N_T \notin \mathbb{D}_m^0$. This explains why we need the assumption $\int_0^T m(s)ds = 0$ in our construction.

Let define the set $\mathcal{S}$ of "smooth" functions. A map $F : \Omega \to \mathbb{R}$ belongs to $\mathcal{S}$ if and only if there exists $a \in \mathbb{R}, d \in \mathbb{N}^*$ and for any $n \in \{1, \cdots, d\}$, $c_1, \cdots, c_n \in \mathcal{K}$, there exists a function $f_n^{c_1,\cdots,c_n} : \mathbb{R}^n \to \mathbb{R}$ such that

1. $F = a\mathbf{1}_{\{N_T=0\}} + \sum_{n=1}^{d}\sum_{c_1,\cdots,c_n\in\mathcal{K}} f_n^{c_1,\cdots,c_n}(T_1,\cdots,T_n)\mathbf{1}_{\{N_T=n, Z_1=c_1,\cdots,Z_n=c_n\}}$;
2. for any $n \in \{1,\cdots,d\}$, any $c_1,\cdots,c_n \in \mathcal{K}$, $f_n^{c_1,\cdots,c_n}$ is smooth with bounded derivatives of any order.

It is known that $\mathcal{S}$ is dense in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Here are some basic properties of directional derivatives on $\mathcal{S}$.

Proposition 2.3.2. *$\mathcal{S} \subset \mathbb{D}_m^0$ and for any smooth function $f : \mathbb{R}^n \to \mathbb{R}$,*

$$D_m f(\bar{T}_1, \bar{T}_2, \cdots, \bar{T}_n) = -\sum_{j=1}^{n} \frac{\partial f}{\partial t_j}(\bar{T}_1, \bar{T}_2, \cdots, \bar{T}_n)\, \hat{m}(\bar{T}_j).$$

Proof. By the definition of D_m we have

$$\begin{aligned}
&D_m f(\bar{T}_1, \bar{T}_2, \cdots, \bar{T}_n) = \frac{\partial \mathcal{T}_\epsilon f(\bar{T}_1, \bar{T}_2, \cdots, \bar{T}_n)}{\partial \epsilon}|_{\epsilon=0} = \frac{\partial}{\partial \epsilon} f(\mathcal{T}_\epsilon \bar{T}_1, \mathcal{T}_\epsilon \bar{T}_2, \cdots, \mathcal{T}_\epsilon \bar{T}_n)|_{\epsilon=0} \\
&= \sum_{j=1}^{n} \frac{\partial f}{\partial t_j}\frac{\partial}{\partial \epsilon}\mathcal{T}_\epsilon \bar{T}_i|_{\epsilon=0} = \sum_{j=1}^{n} \frac{\partial f}{\partial t_j} D_m \bar{T}_i = -\sum_{j=1}^{n} \frac{\partial f}{\partial t_j}(\bar{T}_1, \bar{T}_2, \cdots, \bar{T}_n)\, \hat{m}(\bar{T}_j).
\end{aligned}$$

□

Proposition 2.3.3. *(Functional calculus properties)*

1. *If $F, G \in \mathcal{S}$ then $FG \in \mathcal{S}$ and $D_m(FG) = (D_mF)G + F(D_mG)$.*
2. *If $F_1, F_2, \cdots, F_n \in \mathcal{S}$ and $\Phi : \mathbb{R}^n \to \mathbb{R}$ is a smooth function then $\Phi(F_1, F_2, \cdots, F_n) \in \mathcal{S}$ and*
$$D_m\Phi(F_1, F_2, \cdots, F_n) = \sum_{j=1}^{n} \frac{\partial \Phi}{\partial x_j}(F_1, F_2, \cdots, F_n) D_m F_j.$$

Now we study the absolute continuity of $\mathbb{P}^\epsilon$ with respect to $\mathbb{P}$. Let $\mathbb{E}^\epsilon$ be the expectation taken under the probability $\mathbb{P}^\epsilon$. For every $n \in \mathbb{N}$, $c_1, \cdots, c_n \in \mathcal{K}$, $\lambda_1, \cdots, \lambda_n$ as in Proposition 2.2.1, and every measurable function $f : \mathbb{R}^n \to \mathbb{R}$, we have

$$\begin{aligned}
& \mathbb{E}^\epsilon[f(T_1, \cdots, T_n)\mathbf{1}_{\{N_T=n\}} | Z_1 = c_1, \cdots, Z_n = c_n] \\
= \; & \mathbb{E}[f(U_1, \cdots, U_n)\mathbf{1}_{\{N_T=n\}} | Z_1 = c_1, \cdots, Z_n = c_n] \\
= \; & \mathbb{E}[(f \circ \Phi^{-1})(T_1, \cdots, T_n)\mathbf{1}_{\{T_n \le T < T_{n+1}\}} | Z_1 = c_1, \cdots, Z_n = c_n] \\
= \; & \iint_{0<t_1<\cdots<t_n\le T<t_{n+1}} (f \circ \Phi^{-1})(t_1, \cdots, t_n) \prod_{i=1}^{n+1} \lambda_i e^{-\lambda_i(t_i - t_{i-1})} dt_1 \cdots dt_{n+1} \\
= \; & \iint_{0<t_1<\cdots<t_n\le T} (f \circ \Phi^{-1}) \prod_{i=1}^{n} \lambda_i e^{-\lambda_i(t_i - t_{i-1})} e^{\lambda_{n+1} t_n} dt_1 \cdots dt_n \\
& \times \int_T^\infty \lambda_{n+1} e^{-\lambda_{n+1} t_{n+1}} dt_{n+1} \\
= \; & \iint_{0<t_1<\cdots<t_n\le T} (f \circ \Phi^{-1})(t_1, \cdots, t_n)\varphi(t_1, \cdots, t_n) dt_1 \cdots dt_n \\
= \; & \iint_{0<u_1<\cdots<u_n\le T} f(u_1, \cdots, u_n)(\varphi \circ \Phi)(u_1, \cdots, u_n) |det J_\Phi| du_1 \cdots du_n \\
= \; & \iint_{0<u_1<\cdots<u_n\le T} f\varphi e^{\epsilon(\lambda_{n+1}\int_0^{u_n} m_\epsilon(s)ds - \sum_{i=1}^n \lambda_i \int_{u_{i-1}}^{u_i} m_\epsilon(s)ds)} \\
& \times \prod_{i=1}^{n} (1 + \epsilon m_\epsilon(u_i)) du_1 \cdots du_n \\
= \; & \mathbb{E}\left[f(T_1, \cdots, T_n)\mathbf{1}_{\{N_T=n\}} p_n | Z_1 = c_1, \cdots, Z_n = c_n\right], \text{ where}
\end{aligned}$$

$$\begin{aligned}
\Phi(u_1, \cdots, u_n) &= (u_1 + \epsilon \hat{m}_\epsilon(u_1), \cdots, u_n + \epsilon \hat{m}_\epsilon(u_n)) \\
\varphi(t_1, \cdots, t_n) &= e^{-\lambda_{n+1}(T - t_n)} \prod_{i=1}^{n} \lambda_i e^{-\lambda_i(t_i - t_{i-1})} \\
p_n &= e^{\epsilon(\lambda_{n+1}\int_0^{T_n} m_\epsilon(s)ds - \sum_{i=1}^n \lambda_i \int_{T_{i-1}}^{T_i} m_\epsilon(s)ds)} \prod_{i=1}^{n} (1 + \epsilon m_\epsilon(T_i)).
\end{aligned}$$

Notice that $\int_0^{T_n} m_\epsilon(s)ds = -\int_{T_n}^{T} m_\epsilon(s)ds$, so conditionally to $N_T = n$,

$$\sum_{i=1}^{n} \lambda_i \int_{T_{i-1}}^{T_i} m_\epsilon ds + \lambda_{n+1} \int_{T_n}^{T} m_\epsilon ds = \int_0^T \lambda(C_s) m_\epsilon(s) ds,$$

and we can rewrite

$$p_n = e^{-\epsilon \int_0^T \lambda(C_s) m_\epsilon(s) ds} \prod_{i=1}^{n} (1 + \epsilon m_\epsilon(T_i))$$

Proposition 2.3.4. *$\mathbb{P}^\epsilon$ is absolutely continuous with respect to $\mathbb{P}$ with density*

$$\frac{d\mathbb{P}^\epsilon}{d\mathbb{P}} = e^{-\epsilon \int_0^T \lambda(C_s) m_\epsilon(s) ds} \prod_{i=1}^{N_T} (1 + \epsilon m_\epsilon(T_i)) := G^\epsilon.$$

In case of Poisson process $\lambda_i = 1$, so $\int_0^T \lambda(C_s) m_\epsilon(s) ds = 0$. Hence we have again the result obtained in Carlen and Pardoux (1990) for standard Poisson processes

$$\frac{d\mathbb{P}^\epsilon}{d\mathbb{P}} = \prod_{i=1}^{N_T} (1 + \epsilon m_\epsilon(T_i)).$$

2.4 Gradient and divergence

For any $h \in \mathcal{H}$, we define $\hat{\delta}(h) = \int_0^T h(s) d\tilde{N}_s$, where the compensated process $\tilde{N}$ is defined in (2.2.4). More precisely, $\hat{\delta}(h)$ is defined as

$$\int_0^T h(s) dN_s - \int_0^T \lambda(C_s) h(s) ds = \sum_{i=1}^{N_T} h(T_i) - \lambda_{N_T+1} \int_{T_{N_T}}^{T} h(s) ds - \sum_{i=1}^{N_T} \lambda_i \int_{T_{i-1}}^{T_i} h(s) ds, \tag{2.4.8}$$

with convention that $\sum_{i=1}^{0} = 0$. If we assume that $h \in \mathcal{H} \cap \mathcal{C}^1([0,T])$, by taking the directional derivative of (2.4.8), we obtain $D_m \hat{\delta}(h)$ is equal to

$$\begin{aligned}
& \sum_{i=1}^{N_T} \dot{h}(T_i) D_m T_i + \lambda_{N_T+1} h(T_{N_T}) D_m T_{N_T} - \sum_{i=1}^{N_T} \lambda_i (h(T_i) D_m T_i - h(T_{i-1}) D_m T_{i-1}) \\
= & -\sum_{i=1}^{N_T} \dot{h}(T_i) \hat{m}(T_i) - \sum_{i=0}^{N_T} \lambda_{i+1} h(T_i) \hat{m}(T_i) + \sum_{i=1}^{N_T} \lambda_i h(T_i) \hat{m}(T_i) \\
= & -\int_0^T \dot{h}(s) \hat{m}(s) dN_s - \int_0^T \lambda(C_s) h(s) \hat{m}(s) dN_s + \int_0^T \lambda(C_{s-}) h(s) \hat{m}(s) dN_s \\
= & -\int_0^T \dot{h}(s) \hat{m}(s) dN_s - \int_0^T (\lambda(C_s) - \lambda(C_{s-})) h(s) \hat{m}(s) dN_s
\end{aligned} \tag{2.4.9}$$

By applying Itô formula for the process $\lambda(C_s) h(s) \hat{m}(s)$ in $[0,T]$, we have

$$\begin{aligned}
0 = & \int_0^T \lambda(C_s) [\dot{h}(s) \hat{m}(s) + h(s) m(s)] ds + \sum_{n \geq 1, T_n \leq T} (\lambda(T_n) - \lambda(T_n -)) h(T_n) \hat{m}(T_n) \\
= & \int_0^T \lambda(C_s) \dot{h}(s) \hat{m}(s) ds + \int_0^T \lambda(C_s) h(s) m(s) ds + \int_0^T (\lambda(C_s) - \lambda(C_{s-})) h(s) \hat{m}(s) dN_s.
\end{aligned} \tag{2.4.10}$$

From (2.4.9) and (2.4.10), we obtain

$$\begin{aligned} D_m\hat{\delta}(h) &= -\int_0^T \dot{h}(s)\hat{m}(s)dN_s + \int_0^T \lambda(C_s)\dot{h}(s)\hat{m}(s)ds + \int_0^T \lambda(C_s)h(s)m(s)ds \\ &= -\int_0^T \dot{h}(s)\hat{m}(s)d\tilde{N}_s + \int_0^T \lambda(C_s)h(s)m(s)ds. \end{aligned} \tag{2.4.11}$$

By taking expectation, we obtain

$$\mathbb{E}[D_m\hat{\delta}(h)] = \mathbb{E}\left[\int_0^T \lambda(C_s)h(s)m(s)ds\right]. \tag{2.4.12}$$

2.4.1 Integration by part formula by Bismut's way

In this section, we will establish an integration by part formula directly from the perturbation of measure as Bismut's way (c.f. Bichteler, Gravereaux and Jacod (1987)).

Proposition 2.4.5. *For all $F \in \mathbb{D}_m^0$ and m in $\mathcal{H}$,*

$$\mathbb{E}[D_m F] = \mathbb{E}[\hat{\delta}(m)F].$$

Proof. By the definitions of $\mathcal{T}_\epsilon$ and $\mathbb{P}^\epsilon$, we have

$$\mathbb{E}[\mathcal{T}_\epsilon F] = \mathbb{E}^\epsilon[F] = \mathbb{E}\left[G^\epsilon F\right]. \tag{2.4.13}$$

Indeed, for every Borel set $A \subset \mathbb{R}$,

$$\mathbb{P}^\epsilon(F \in A) = \mathbb{P}(\mathcal{T}_\epsilon^{-1}(\{w \in \Omega : F(w) \in A\})) = \mathbb{P}(\{w \in \Omega : F(\mathcal{T}_\epsilon(w)) \in A\}) = \mathbb{P}(\mathcal{T}_\epsilon F \in A).$$

By taking the derivation in both sides of (2.4.13) with respect to ϵ at $\epsilon = 0$, we obtain

$$\mathbb{E}[D_m F] = \mathbb{E}\left[\frac{\partial G^\epsilon}{\partial \epsilon}|_{\epsilon=0} F\right].$$

Notice that

$$\frac{\partial G^\epsilon}{\partial \epsilon}|_{\epsilon=0} = -\int_0^T \lambda(C_s)m(s)ds + \sum_{i=1}^{N_T} m(T_i) = \hat{\delta}(m)$$

we have the integration by part formula of the form $\mathbb{E}[D_m F] = \mathbb{E}[\hat{\delta}(m)F]$. □

Remark 2.4.2. We can also retrieve the equality (2.4.12) by using Proposition 2.4.5 for $F = \hat{\delta}(h)$ as follow

$$\mathbb{E}[D_m\hat{\delta}(h)] = \mathbb{E}[\hat{\delta}(m)\hat{\delta}(h)] = \mathbb{E}\left[\int_0^T m(s)d\tilde{N}_s \int_0^T h(s)d\tilde{N}_s\right] = \mathbb{E}\left[\int_0^T \lambda(C_s)h(s)m(s)ds\right].$$

Corollary 2.4.1. *For all $F, G \in \mathbb{D}_m^0$ such that $FG \in \mathbb{D}_m^0$ and m in $\mathcal{H}$,*

$$\mathbb{E}[GD_m F] = \mathbb{E}[F(\hat{\delta}(m)G - D_m G)].$$

Proof. The result is direct by applying Proposition 2.4.5 for FG. □

Corollary 2.4.2. *Let $\psi \in \mathbb{D}_m^0$, $\Phi = (\phi_1, \cdots, \phi_d)$ with $\phi_i \in \mathbb{D}_m^0, i = 1, \cdots, d$, and $f \in C^1(\mathbb{R}^d)$. Then*

$$\mathbb{E}[\psi D_m(f \circ \Phi)] = \mathbb{E}[\psi \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\Phi) D_m \phi_i] = \mathbb{E}[f(\Phi)(\hat{\delta}(m)\psi - D_m\psi)].$$

Proof. By applying Proposition 2.4.5 for $F = \psi f(\Phi)$, we obtain

$$\mathbb{E}[D_m(\psi f(\Phi))] = \mathbb{E}[\hat{\delta}(m)\psi f(\Phi)].$$

But $D_m(\psi f(\Phi)) = D_m\psi f(\Phi) + \psi D_m f(\Phi) = D_m\psi f(\Phi) + \psi \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\Phi) D_m\phi_i$. Hence,

$$\mathbb{E}[\psi \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(\Phi) D_m \phi_i] = \mathbb{E}[\hat{\delta}(m)\psi f(\Phi)] - \mathbb{E}[D_m\psi f(\Phi)] = \mathbb{E}[f(\Phi)(\hat{\delta}(m)\psi - D_m\psi)].$$

□

Proposition 2.4.6. *(Closability of D_m) For any $m \in \mathcal{H}$, D_m is closable.*

Proof. To prove that D_m is closable, we have to prove that if $(F_k)_{k\in\mathbb{N}} \subset \mathcal{S}$ satisfying $F_k \to 0$ and $D_m F_k \to u$ in $L^2(\Omega)$ then $u = 0$ $\mathbb{P}$ a.s.. From Corollary 2.4.1 we have

$$\mathbb{E}[G D_m F_k] = \mathbb{E}[F_k(\hat{\delta}(m)G - D_m G)].$$

Let k tend to infinite we get $\mathbb{E}[Gu] = 0$ for every $G \in \mathcal{S}$. This is also true for $G \in L^2(\Omega)$ since $\mathcal{S}$ is dense in $L^2(\Omega)$, so we can take $G = u$ and receive $u = 0$ $\mathbb{P}$ a.s. □

We shall identify D_m with its closed extension and denote by $\mathbb{D}_m^{1,2}$ its domain, i.e.,

$$\mathbb{D}_m^{1,2} = \{F \in L^2(\Omega) : \exists (F_k)_{k\in\mathbb{N}} \subset \mathcal{S} \text{ s.t } (F_k)_{k\in\mathbb{N}} \to F, (D_m F_k)_{k\in\mathbb{N}} \text{ converges in } L^2(\Omega)\}.$$

Obviously, $\mathcal{S} \subset \mathbb{D}_m^{1,2}$. For every $F \in \mathbb{D}_m^{1,2}$, we define $D_m F = \lim_{k\to\infty} D_m F_k$. Thanks to Proposition 2.4.6, this limit does not depend on the choice of the sequence $(F_k)_{k\in\mathbb{N}}$, so $D_m F$ is well-defined and coincides with the already defined $D_m F$ in case $F \in \mathcal{S}$. For every $F \in \mathbb{D}_m^{1,2}$, we have

$$||F||^2_{\mathbb{D}_m^{1,2}} = ||F||^2_{L^2(\Omega)} + ||D_m F||^2_{L^2(\Omega)} < +\infty.$$

2.4.2 Local Dirichlet form

Next we would like to define an operator D from $L^2(\Omega)$ into $L^2(\Omega; \mathcal{H})$ which play the role of stochastic derivative, i.e., for every F in its domain which should belong to every domain $\mathbb{D}_m^{1,2}$, this equality should be satisfied

$$D_m F = \langle DF, m\rangle_{\mathcal{H}} = \int_0^T D_s F m(s) ds.$$

Let $(m_k)_{k\in\mathbb{N}}$ be an orthonormal basis of the space $\mathcal{H}$. Then every function $m \in \mathcal{H}$ can be expressed as $m = \sum_{k=1}^{+\infty} \langle m, m_k\rangle_{\mathcal{H}} m_k$.
We now follow the construction of Bouleau and Hirsch (1991). We set

$$\mathbb{D}^{1,2} = \{X \in \bigcap_{k=1}^{+\infty} \mathbb{D}^{1,2}_{m_k} : \sum_{k=1}^{+\infty} \| D_{m_k} X \|^2_{L^2(\Omega)} < +\infty\}, \text{ and}$$

$$\forall X, Y \in \mathbb{D}^{1,2},\ \mathcal{E}[X,Y] = \sum_{k=1}^{+\infty} \mathbb{E}[D_{m_k} X D_{m_k} Y].$$

We denote $\mathcal{E}[X] := \mathcal{E}[X,X]$ for convenience. The next proposition defines the *local Dirichlet form.* Concerning the definition of local Dirichlet forms, our reference is Bouleau and Hirsch (1991).

Proposition 2.4.7. *The bilinear form $(\mathbb{D}^{1,2}, \mathcal{E})$ is a local Dirichlet admitting a gradient, D, and a carré du champ, Γ, given respectively by the following formulas:*

$$\begin{aligned} \forall X \in \mathbb{D}^{1,2},\ DX &= \sum_{k=1}^{+\infty} D_{m_k} X\, m_k \in L^2(\Omega; \mathcal{H}), \\ \forall X, Y \in \mathbb{D}^{1,2},\ \Gamma[X,Y] &= \langle DX, DY\rangle_{\mathcal{H}}. \end{aligned}$$

As a consequence $\mathbb{D}^{1,2}$ is a Hilbert space equipped with the norm:

$$\forall X \in \mathbb{D}^{1,2},\ \| X \|^2_{\mathbb{D}^{1,2}} = ||X||^2_{L^2(\Omega)} + \mathcal{E}(X).$$

Proof. The proof is obvious and uses the same arguments as the proof of Proposition 4.2.1 in (Bouleau and Hirsch 1991, Chapter II). Let us remark that the locality property is a direct consequence of the functional calculus, see Proposition 2.3.3 in (Bouleau and Hirsch 1991, Chapter II). □

Example 2.4.5. By using the result in Example 2.3.4, for any i we have

1. $DZ_i = \sum_{k=1}^{+\infty} \langle DZ_i, m_k\rangle_{\mathcal{H}} m_k = \sum_{k=1}^{+\infty} D_{m_k} Z_i\, m_k = 0$. Similarly, $DN_T = 0$.
2. In the same way,

$$\begin{aligned} D\bar{T}_i &= \sum_k \langle D\bar{T}_i, m_k\rangle_{\mathcal{H}} m_k = \sum_k D_{m_k} \bar{T}_i\, m_k = -\sum_k \hat{m}_k(\bar{T}_i) m_k \\ &= -\sum_k m_k \int_0^T m_k(s) 1_{[0,\bar{T}_i]}(s) ds = \sum_k m_k \int_0^T m_k(s) \left(\frac{\bar{T}_i}{T} - 1_{[0,\bar{T}_i]}(s)\right) ds \\ &= \sum_k \langle m_k, \frac{\bar{T}_i}{T} - 1_{[0,\bar{T}_i]}\rangle_{\mathcal{H}} m_k = \frac{\bar{T}_i}{T} - 1_{[0,\bar{T}_i]}. \end{aligned}$$

Moreover, as a consequence of the functional calculus specific to local Dirichlet forms (see Bouleau and Hirsch (1991), Section I.6) the set of smooth function $\mathcal{S}$ is dense in $\mathbb{D}^{1,2}$ and if $F \in \mathcal{S}$, $F = f(\bar{T}_1, \bar{T}_2, \cdots, \bar{T}_n)$ where f is smooth with bounded derivatives of any order, then

$$DF = \sum_{i=1}^{n} \frac{\partial f}{\partial t_i} \left(\frac{\bar{T}_i}{T} - 1_{[0,\bar{T}_i]}\right).$$

Remark 2.4.3. The expression of the gradient operator D above prove that nor the Dirichlet form $(\mathbb{D}^{1,2}, \mathcal{E})$ nor the gradient D depend on the choice of the orthonormal basis $(m_k)_{k\in\mathbb{N}}$ of $\mathcal{H}$.

2.4.3 Divergence operator

Let $\delta : L^2(\Omega;\mathcal{H}) \to L^2(\Omega)$ be the adjoint operator of D. Its domain, $Dom(\delta)$, is the set of $u \in L^2(\Omega;\mathcal{H})$ such that there exists $c > 0$ satisfying

$$\left|\mathbb{E}\left[\int_0^T D_s F u_s ds\right]\right| \leq c||F||_{\mathbb{D}^{1,2}},\ \forall F \in \mathbb{D}^{1,2}.$$

It follows from the properties of D that δ is also a closed densely defined operator. We have the integral by part formula by the duality:

$$\mathbb{E}[\delta(u)F] = \mathbb{E}[\langle u, DF\rangle_{\mathcal{H}}] = \mathbb{E}\left[\int_0^T u_s D_s F ds\right],\ \forall F \in \mathbb{D}^{1,2}, u \in Dom(\delta).$$

Proposition 2.4.8. *For every $u \in \mathbb{D}^{1,2}\otimes\mathcal{H}$, we have $u \in Dom(\delta)$ and*

$$\delta(u) = \int_0^T u_s d\tilde{N}_s - \int_0^T D_s u_s ds.$$

Proof. First of all, if $u = mG$ with $m \in \mathcal{H}, G \in \mathbb{D}^{1,2}$ then

$$\mathbb{E}\left[\int_0^T u_s D_s F ds\right] = \mathbb{E}\left[G\int_0^T m(s) D_s F dt\right] = \mathbb{E}[GD_m F] = \mathbb{E}[F(\hat{\delta}(m)G - D_m G)]$$

for every $F \in \mathbb{D}^{1,2}$. From the uniqueness of δ, we have

$$\delta(u) = \hat{\delta}(m)G - D_m G = G\int_0^T m(s) d\tilde{N}_s - \int_0^T D_s G m(s) ds = \int_0^T u_s d\tilde{N}_s - \int_0^T D_s u_s ds.$$

By linearity, this is true for every function in

$$\{u \in \mathbb{D}^{1,2}\otimes\mathcal{H} : u = \sum_{i=1}^n m_i G_i, m_i \in \mathcal{H}, G_i \in \mathbb{D}^{1,2}\}.$$

The result of the proposition follows since this set is dense in $\mathbb{D}^{1,2}\otimes\mathcal{H}$. □

Remark 2.4.4. If $u \in \mathcal{H}$, then $D_s u_s = 0$ for every $s \in [0,T]$, so the divergence operator δ coincides with the integral w.r.t. the compensated Poisson process $\hat{\delta}$. From the proof of Proposition 2.4.8, we can retain that

1. $\delta(mG) = \delta(m)G - D_m G$ for every $m \in \mathcal{H}, G \in \mathbb{D}^{1,2}$.
2. $\mathbb{E}[GD_m F] = \mathbb{E}[F\delta(mG)]$ for every $m \in \mathcal{H}, F, G \in \mathbb{D}^{1,2}$.

Corollary 2.4.3. *If $u \in \mathbb{D}^{1,2}\otimes\mathcal{H}$ is an adapted process then*

$$\delta(u) = \int_0^T u_s d\tilde{N}_s.$$

Proof. We have

$$D_s u_s = \sum_{T_i \leq T} \frac{\partial u_s}{\partial T_i} D_s T_i = \sum_{T_i \leq T} \frac{\partial u_s}{\partial T_i} \left(\frac{T_i}{T} - 1_{[0,T_i]}(s) \right), \text{ where}$$

$$\int_0^T \frac{\partial u_s}{\partial T_i} \frac{T_i}{T} dt = \frac{T_i}{T} \frac{\partial}{\partial T_i} \int_0^T u_s ds = 0, \text{ so } \int_0^T D_s u_s ds = - \sum_{T_i \leq T} \int_0^{T_i} \frac{\partial u_s}{\partial T_i} ds.$$

If u is adapted, $\dfrac{\partial u_s}{\partial T_i} = 0$ for all $t < T_i$, so $\int_0^T D_s u_s ds = 0$. □

Corollary 2.4.4. *If $u \in L^2(\Omega; \mathcal{H})$ is an adapted process then*

$$\delta(u) = \int_0^T u_s d\tilde{N}_s.$$

Proof. The result follows from Corollary 2.4.3 by approximating u by a sequence of adapted processes in $\mathbb{D}^{1,2} \otimes \mathcal{H}$. □

The next proposition is an important property of the operator δ, which follows from the fact that D is a derivative.

Proposition 2.4.9. *Let $F \in \mathbb{D}^{1,2}$, $X \in Dom(\delta)$ such that $F\delta(X) - \int_0^T D_s F X_s ds \in L^2(\Omega)$, then $FX \in Dom(\delta)$, and*

$$\delta(FX) = F\delta(X) - \int_0^T D_s F X_s ds.$$

Proof. For every $G \in \mathcal{S}$, by the integration by parts formula, we have

$$\begin{aligned} \mathbb{E}[\delta(FX)G] &= \mathbb{E}\left[\int_0^T F X_s D_s G ds \right] = \mathbb{E}\left[\int_0^T X_s (D_s(GF) - G D_s F) ds \right] \\ &= \mathbb{E}\left[G(F\delta(X) - \int_0^T D_s F X_s ds) \right]. \end{aligned}$$

□

Corollary 2.4.5. *Let $m \in \mathcal{H}$ and $F \in \mathbb{D}^{1,2}$, then $mF \in Dom(\delta)$, and*

$$\delta(mF) = F \int_0^T m(s) d\tilde{N}_s - D_m F.$$

Proof. We obtain the result by applying Proposition 2.4.9 for $X = m$, $\delta(m) = \int_0^T m(s) d\tilde{N}_s$ and $\int_0^T D_s F m(s) ds = D_m F$. □

Remark 2.4.5. In this approach, the Clark-Ocone formula (cf. Theorem 1.2.6) does not hold. For example, we have already shown that $D_m N_T = 0, \forall m \in \mathcal{H}$, so $D_s N_T = 0, \forall s \in [0,T]$, but we do not have $N_T = \mathbb{E}[N_T]$.

2.5 An absolute continuity criterion

In this section, we denote

1. $A_0 = \{N_T = 0\}$, $p_0 = \mathbb{P}(A_0)$.
2. $A_n^{c_1,\cdots,c_n} = \{N_T = n, Z_1 = c_1, \cdots, Z_n = c_n\}$, and $p_n^{c_1,\cdots,c_n} = \mathbb{P}(A_n^{c_1,\cdots,c_n})$, for every $n \in \mathbb{N}^*$, $c_1, \cdots, c_n \in \mathcal{K}$.

Lemma 2.5.3. *Let $\Delta_n = \{t = (t_1, \cdots, t_n) : 0 < t_1 < \cdots < t_n < T\}$. The distribution of $(T_1, \cdots, T_n)$ conditionally to $A_n^{c_1,\cdots,c_n}$ has a density*

$$\begin{aligned} k_n^{c_1,\cdots,c_n} : \mathbb{R}^n &\to \mathbb{R}^+ \\ t = (t_1, \cdots, t_n) &\mapsto k_n^{c_1,\cdots,c_n}(t) = n!\, \mathbf{1}_{\Delta_n}(t) \prod_{i=1, \lambda_{i+1} \neq \lambda_i}^{n} \frac{(\lambda_{i+1} - \lambda_i) e^{t_i(\lambda_{i+1} - \lambda_i)}}{e^{(\lambda_{i+1} - \lambda_i)T} - 1} \end{aligned}$$

with respect to the Lebesgue measure ν_n on $\mathbb{R}^n$, where $\lambda_i = -\lambda_{c_i c_i}, i = 1, \cdots, n$.

Proof. Let f be a measureable function on $\mathbb{R}^n$. Then

$$\begin{aligned} &\mathbb{E}[f(T_1, \cdots, T_n) | A_n^{c_1,\cdots,c_n}] = \mathbb{E}[\mathbf{1}_{A_n^{c_1,\cdots,c_n}} f(T_1, \cdots, T_n)] / p_n^{c_1,\cdots,c_n} \\ &= \frac{\mathbb{E}[f(T_1, \cdots, T_n) \mathbf{1}_{\{N_T = n\}} \mathbf{1}_{\{Z_1 = c_1, \cdots, Z_n = c_n\}}]}{\mathbb{E}[\mathbf{1}_{\{N_T = n\}} \mathbf{1}_{\{Z_1 = c_1, \cdots, Z_n = c_n\}}]} \\ &= \frac{\mathbb{E}[f(T_1, \cdots, T_n) \mathbf{1}_{\{N_T = n\}} | \mathbf{1}_{\{Z_1 = c_1, \cdots, Z_n = c_n\}}] \, \mathbb{E}[\mathbf{1}_{\{Z_1 = c_1, \cdots, Z_n = c_n\}}]}{\mathbb{E}[\mathbf{1}_{\{N_T = n\}} | \mathbf{1}_{\{Z_1 = c_1, \cdots, Z_n = c_n\}}] \, \mathbb{E}[\mathbf{1}_{\{Z_1 = c_1, \cdots, Z_n = c_n\}}]} \\ &= \frac{\mathbb{E}[f(T_1, \cdots, T_n) \mathbf{1}_{\{T_n \leq T < T_{n+1}\}} | \mathbf{1}_{\{Z_1 = c_1, \cdots, Z_n = c_n\}}]}{\mathbb{E}[\mathbf{1}_{\{T_n \leq T < T_{n+1}\}} | \mathbf{1}_{\{Z_1 = c_1, \cdots, Z_n = c_n\}}]} \\ &= \frac{\iint_{0 < t_1 < \cdots < t_n \leq T < t_{n+1}} f(t_1, \cdots, t_n) \prod_{i=1}^{n+1} \lambda_i e^{-\lambda_i (t_i - t_{i-1})} dt_1 \cdots dt_{n+1}}{\iint_{0 < t_1 < \cdots < t_n \leq T < t_{n+1}} \prod_{i=1}^{n+1} \lambda_i e^{-\lambda_i (t_i - t_{i-1})} dt_1 \cdots dt_{n+1}} \\ &= \frac{\iint_{\Delta_n} f \prod_{i=1}^{n} \lambda_i e^{-\lambda_i (t_i - t_{i-1})} e^{\lambda_{n+1} t_n} dt_1 \cdots dt_n \int_T^\infty \lambda_{n+1} e^{-\lambda_{n+1} t_{n+1}} dt_{n+1}}{\iint_{\Delta_n} \prod_{i=1}^{n} \lambda_i e^{-\lambda_i (t_i - t_{i-1})} e^{\lambda_{n+1} t_n} dt_1 \cdots dt_n \int_T^\infty \lambda_{n+1} e^{-\lambda_{n+1} t_{n+1}} dt_{n+1}} \\ &= \frac{\iint_{\Delta_n} f\varphi(t_1, \cdots, t_n) dt_1 \cdots dt_n}{\iint_{\Delta_n} \varphi(t_1, \cdots, t_n) dt_1 \cdots dt_n}, \text{ where} \\ &\varphi(t_1, \cdots, t_n) = e^{-\lambda_{n+1}(T - t_n)} \prod_{i=1}^{n} \lambda_i e^{-\lambda_i (t_i - t_{i-1})} = e^{-\lambda_{n+1} T} \prod_{i=1}^{n} \lambda_i e^{t_i (\lambda_{i+1} - \lambda_i)}, \\ &\iint_{\Delta_n} \varphi(t_1, \cdots, t_n) dt_1 \cdots dt_n = \frac{1}{n!} \iint_{[0,T]^n} e^{-\lambda_{n+1} T} \prod_{i=1}^{n} \lambda_i e^{t_i (\lambda_{i+1} - \lambda_i)} dt_1 \cdots dt_n \\ &= \frac{e^{-\lambda_{n+1} T}}{n!} \prod_{i=1}^{n} \lambda_i \prod_{i=1, \lambda_{i+1} \neq \lambda_i}^{n} \frac{e^{(\lambda_{i+1} - \lambda_i)T} - 1}{\lambda_{i+1} - \lambda_i}. \end{aligned}$$

□

Remark 2.5.6. $k_n^{c_1,\cdots,c_n}$ is a positive function of class $\mathcal{C}^\infty$ on Δ_n. Particularly, in case C is a homogeneous Poisson process, $\lambda_1 = \cdots = \lambda_n$, so $k_n^{c_1,\cdots,c_n}(t_1, \cdots, t_n) = n!\, \mathbf{1}_{\Delta_n}(t)$ which means $(T_1, \cdots, T_n)$ has uniform distribution on Δ_n conditionnally to $A_n^{c_1,\cdots,c_n}$.

Now fixe $n, c_1, \cdots, c_n$, we consider $\mathbf{d}_n^{c_1,\cdots,c_n}$ the set of $\mathcal{B}(\mathbb{R}^n)$-measurable functions u in $L^2(k_n^{c_1,\cdots,c_n}dt)$ such that for any $i \in \{1, \cdots, n\}$, and ν_{n-1}-almost all $\bar{t} = (t_1, \cdots, t_{i-1}, t_{i+1}, \cdots, t_n) \in \mathbb{R}^{n-1}$, $u_{\bar{t}}^{(i)} = u(t_1, \cdots, t_{i-1}, ., t_{i+1}, \cdots, t_n)$ has an absolute continuous version $\tilde{u}_{\bar{t}}^{(i)}$ on $\{t_i : (t_1, \cdots, t_n) \in \Delta_n\}$ such that

$$\sum_{i,j=1}^{n} \frac{\partial u}{\partial t_i}\frac{\partial u}{\partial t_j}\left(t_i \wedge t_j - \frac{t_i t_j}{T}\right) \in L^1(k_n^{c_1,\cdots,c_n}dt),$$

where $\dfrac{\partial u}{\partial t_i} = \dfrac{\partial \tilde{u}_{\bar{t}}^{(i)}}{\partial s}$. We consider the following quadratic form on $\mathbf{d}_n^{c_1,\cdots,c_n}$

$$e_n^{c_1,\cdots,c_n}[u,v] = \frac{1}{2}\int_{\mathbb{R}^n} \sum_{i,j=1}^{n} \frac{\partial u}{\partial t_i}\frac{\partial v}{\partial t_j}\left(t_i \wedge t_j - \frac{t_i t_j}{T}\right) k_n^{c_1,\cdots,c_n}dt, \forall u, v \in \mathbf{d}_n^{c_1,\cdots,c_n}.$$

As usual, we denote $e[u,u]$ by $e[u]$.

Proposition 2.5.10. *1. $(\mathbf{d}_n^{c_1,\cdots,c_n}, e_n^{c_1,\cdots,c_n})$ is a local Dirichlet form on $L^2(k_n^{c_1,\cdots,c_n}dt)$ which admits a "carré du champ" operator γ_n and a gradient operator $\tilde{D}^n$ given by*

$$\gamma_n[u,v](t) = \sum_{i,j=1}^{n} \frac{\partial u}{\partial t_i}\frac{\partial v}{\partial t_j}\left(t_i \wedge t_j - \frac{t_i t_j}{T}\right), \quad \tilde{D}_s^n u(t) = \sum_{i=1}^{n} \frac{\partial u}{\partial t_i}\left(\frac{t_i}{T} - \mathbf{1}_{[0,t_i]}(s)\right)$$

for all $u, v \in \mathbf{d}_n^{c_1,\cdots,c_n}$, $t = (t_1, \cdots, t_n) \in \mathbb{R}^n$, $s \in [0,T]$.

2. The structure $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), k_n^{c_1,\cdots,c_n}dt, \mathbf{d}_n^{c_1,\cdots,c_n}, \gamma_n)$ satisfies the EID property, i.e., for every $d \in \mathbb{N}^*, \forall u = (u_1, \cdots, u_d) \in (\mathbf{d}_n^{c_1,\cdots,c_n})^d$, we have

$$u_*[(det\gamma_n[u]).k_n^{c_1,\cdots,c_n}\nu_n] \ll \nu_d,$$

where $\gamma_n[u]$ denotes the matrix $(\gamma_n[u_i,u_j])_{1\le i,j\le d}$.

Proof. The results are obtained by applying Proposition 1 and Theorem 2 in Bouleau and Denis (2009) (see also Example 1.1.3) for $\mathbf{d} = \mathbf{d}_n^{c_1,\cdots,c_n}, k = k_n^{c_1,\cdots,c_n}$ and $\xi_{ij}(t) = t_i \wedge t_j - \frac{t_i t_j}{T}, \forall 1 \le i,j \le n, t = (t_1, \cdots, t_n)$. We only have to prove that ξ is locally elliptic on Δ_n. Indeed, for every $\alpha = (\alpha_1, \cdots, \alpha_n) \in \mathbb{R}^n$, $t = (t_1, \cdots, t_n) \in \Delta_n$, we have

$$\alpha^*\xi(t)\alpha = \sum_{i,j=1}^{n} \alpha_i\alpha_j\left(t_i \wedge t_j - \frac{t_i t_j}{T}\right) = \int_0^T \left(\sum_{i=1}^{n} \alpha_i\left(\mathbf{1}_{[0,t_i]}(s) - \frac{t_i}{T}\right)\right)^2 ds \ge 0.$$

$\alpha^*\xi(t)\alpha = 0$ if and only if $\sum_{i=1}^n \alpha_i\left(\mathbf{1}_{[0,t_i]}(s) - \frac{t_i}{T}\right) = 0, \forall s \in [0,T]$. By taking $t_n < s < T$, we obtain $\sum_{i=1}^n \alpha_i t_i = 0$, so $\sum_{i=1}^n \alpha_i \mathbf{1}_{[0,t_i]}(s) = 0, \forall s \in [0,T]$. By taking $t_{n-1} < s < t_n$, we obtain $\alpha_n = 0$, so $\sum_{i=1}^{n-1} \alpha_i \mathbf{1}_{[0,t_i]}(s) = 0, \forall s \in [0,T]$. Continue this process and finally we obtain $\alpha_i = 0, \forall i = 1, \cdots, n$. □

Remark 2.5.7. Since the density function $k_n^{c_1,\cdots,c_n}$ is bounded in Δ_n both below and above by positive constants, we have $L^p(k_n^{c_1,\cdots,c_n}dt) = L^p(dt)$, so the domain $d_n^{c_1,\cdots,c_n}$ does not depend on $c_1, \cdots, c_n$.

For every measurable random variable $F : \Omega \to \mathbb{R}$, there exists a constant $a \in \mathbb{R}$ and measurable functions $f_n^{c_1,\cdots,c_n} : \mathbb{R}^n \to \mathbb{R}, \forall n \in \mathbb{N}^*, c_1, \cdots, c_n \in \mathcal{K}$ such that

$$F = a\mathbf{1}_{A_0} + \sum_{n=1}^{\infty} \sum_{c_1,\cdots,c_n \in \mathcal{K}} \mathbf{1}_{A_n^{c_1,\cdots,c_n}} f_n^{c_1,\cdots,c_n}(T_1, \cdots, T_n) \quad \mathbb{P} - a.s.. \tag{2.5.14}$$

From now on, we will write $\sum_{n,c}$ instead of $\sum_{n=1}^{\infty} \sum_{c_1,\cdots,c_n \in \mathcal{K}}$ for simplicity.

Proposition 2.5.11. *For every measurable random variable F of the form (2.5.14), we have $F \in \mathbb{D}^{1,2}$ if and only if $f_n^{c_1,\cdots,c_n} \in \mathbf{d}_n^{c_1,\cdots,c_n}, \forall n \in \mathbb{N}^*, \forall c_1, \cdots, c_n \in \mathcal{K}$ and*

$$\sum_{n,c} p_n^{c_1,\cdots,c_n} ||f_n^{c_1,\cdots,c_n}||^2_{\mathbf{d}_n^{c_1,\cdots,c_n}} < \infty.$$

Moreover, if $F \in \mathbb{D}^{1,2}$ then

$$||F||^2_{\mathbb{D}^{1,2}} = a^2 p_0 + \sum_{n,c} p_n^{c_1,\cdots,c_n} ||f_n^{c_1,\cdots,c_n}||^2_{\mathbf{d}_n^{c_1,\cdots,c_n}}. \tag{2.5.15}$$

Proof. For every F of the form (2.5.14), we have

$$\begin{aligned}
||F||^2_{L^2(\Omega)} &= \mathbb{E}[F^2] = a^2 p_0 + \sum_{n,c} \mathbb{E}\left[\mathbf{1}_{A_n^{c_1,\cdots,c_n}} \left(f_n^{c_1,\cdots,c_n}(T_1, \cdots, T_n)\right)^2\right] \\
&= a^2 p_0 + \sum_{n,c} p_n^{c_1,\cdots,c_n} \mathbb{E}\left[\left(f_n^{c_1,\cdots,c_n}(T_1, \cdots, T_n)\right)^2 | A_n^{c_1,\cdots,c_n}\right] \\
&= a^2 p_0 + \sum_{n,c} p_n^{c_1,\cdots,c_n} \int_{\mathbb{R}^n} \left(f_n^{c_1,\cdots,c_n}(t)\right)^2 k_n^{c_1,\cdots,c_n}(t) dt \\
&= a^2 p_0 + \sum_{n,c} p_n^{c_1,\cdots,c_n} ||f_n^{c_1,\cdots,c_n}||^2_{L^2(k_n^{c_1,\cdots,c_n} dt)}, \text{ and}
\end{aligned}$$

$$\begin{aligned}
D_s F &= \sum_{n,c} \mathbf{1}_{A_n^{c_1,\cdots,c_n}} \left(\sum_{i=1}^{n} \frac{\partial f_n^{c_1,\cdots,c_n}}{\partial t_i}(T_1, \cdots, T_n) \left(\frac{T_i}{T} - \mathbf{1}_{[0,T_i]}(s)\right)\right) \\
&= \sum_{n,c} \mathbf{1}_{A_n^{c_1,\cdots,c_n}} \tilde{D}_s^n f_n^{c_1,\cdots,c_n}(T_1, \cdots, T_n). \text{ Hence,}
\end{aligned}$$

$$\begin{aligned}
||DF||^2_{L^2(\Omega;\mathcal{H})} &= \int_0^T \mathbb{E}[(D_s F)^2] ds \\
&= \sum_{n,c} \int_0^T \mathbb{E}\left[\mathbf{1}_{A_n^{c_1,\cdots,c_n}} \left(\tilde{D}_s^n f_n^{c_1,\cdots,c_n}(T_1, \cdots, T_n)\right)^2\right] ds \\
&= \sum_{n,c} \int_0^T p_n^{c_1,\cdots,c_n} \mathbb{E}\left[\left(\tilde{D}_s^n f_n^{c_1,\cdots,c_n}(T_1, \cdots, T_n)\right)^2 | A_n^{c_1,\cdots,c_n}\right] ds \\
&= \sum_{n,c} p_n^{c_1,\cdots,c_n} \int_0^T \int_{\mathbb{R}^n} \left(\tilde{D}_s^n f_n^{c_1,\cdots,c_n}(t)\right)^2 k_n^{c_1,\cdots,c_n}(t) dt ds \\
&= \sum_{n,c} p_n^{c_1,\cdots,c_n} ||\tilde{D}^n f_n^{c_1,\cdots,c_n}||^2_{L^2((\mathbb{R}^n, k_n^{c_1,\cdots,c_n} dt);\mathcal{H})} \\
&= \sum_{n=1}^{\infty} \sum_{c_1,\cdots,c_n \in \mathcal{K}} p_n^{c_1,\cdots,c_n} e_n^{c_1,\cdots,c_n} [f_n^{c_1,\cdots,c_n}].
\end{aligned}$$

Therefore,

$$||F||^2_{L^2(\Omega)} + ||DF||^2_{L^2(\Omega;\mathcal{H})} = \sum_{n,c} p_n^{c_1,\cdots,c_n} ||f_n^{c_1,\cdots,c_n}||^2_{\mathbf{d}_n^{c_1,\cdots,c_n}}.$$

From here we obtain the condition for $F \in \mathbb{D}^{1,2}$. The equation (2.5.15) is obvious since the fact that $||F||^2_{\mathbb{D}^{1,2}} = ||F||^2_{L^2(\Omega)} + \mathcal{E}[F] = ||F||^2_{L^2(\Omega)} + ||DF||^2_{L^2(\Omega;\mathcal{H})}$. □

Remark 2.5.8. In summary, we have the following relations between the Dirichlet structure $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n), k_n^{c_1,\cdots,c_n} dt, \mathbf{d}_n^{c_1,\cdots,c_n}, \gamma_n)$ and the Dirichlet structures $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{D}^{1,2}, \Gamma)$: for every $F \in \mathbb{D}^{1,2}$ having the form (2.5.14),

1. $||F||^2_{L^2(\Omega)} = a^2 p_0 + \sum_{n,c} p_n^{c_1,\cdots,c_n} ||f_n^{c_1,\cdots,c_n}||^2_{L^2(k_n^{c_1,\cdots,c_n} dt)}$.
2. $D_s F = \sum_{n,c} \mathbf{1}_{A_n^{c_1,\cdots,c_n}} \tilde{D}_s^n f_n^{c_1,\cdots,c_n}(T_1,\cdots,T_n)$, $\forall s \in [0,T]$.
3. $\Gamma[F] = \sum_{n,c} \mathbf{1}_{A_n^{c_1,\cdots,c_n}} \gamma_n[f_n^{c_1,\cdots,c_n}](T_1,\cdots,T_n)$
4. $\mathcal{E}(F) = \sum_{n,c} p_n^{c_1,\cdots,c_n} e_n^{c_1,\cdots,c_n}[f_n^{c_1,\cdots,c_n}]$.
5. $||F||^2_{\mathbb{D}^{1,2}} = a^2 p_0 + \sum_{n,c} p_n^{c_1,\cdots,c_n} ||f_n^{c_1,\cdots,c_n}||^2_{\mathbf{d}_n^{c_1,\cdots,c_n}}$.

Now let $d \in \mathbb{N}^*$ and $F = (F_1,\cdots,F_d) \in (\mathbb{D}^{1,2})^d$, we denote by $\Gamma[F]$ the $d \times d$ symmetric matrix $(\Gamma[F_i, F_j])_{1 \le i,j \le d}$.

Theorem 2.5.7. *(Energy image density property)*
If $F \in (\mathbb{D}^{1,2})^d$, then $F_*[(det\Gamma[F]).\mathbb{P}]$ is absolutely continuous with respect to the Lebesgue measure ν_d on $\mathbb{R}^d$.

Proof. Let $B \subset \mathbb{R}^d$ such that $\nu_d(B) = 0$. We have to prove that

$$F_*[(det\Gamma[F]).\mathbb{P}](B) = 0 \text{ or equivalently, } \mathbb{E}[\mathbf{1}_B(F) det\Gamma[F]] = 0.$$

Thanks to Proposition 2.5.11, we can write

$$F = a\mathbf{1}_{A_0} + \sum_{n=1}^{\infty} \sum_{c_1,\cdots,c_n \in \mathcal{K}} \mathbf{1}_{A_n^{c_1,\cdots,c_n}} f_n^{c_1,\cdots,c_n}(T_1,\cdots,T_n),$$

with $f_n^{c_1,\cdots,c_n} \in (\mathbf{d}_n^{c_1,\cdots,c_n})^d, \forall n \in \mathbb{N}^*, \forall c_1,\cdots,c_n \in \mathcal{K}$. Then

$$\begin{aligned} \Gamma[F,F] &= \sum_{n,c} \mathbf{1}_{A_n^{c_1,\cdots,c_n}} \gamma_n[f_n^{c_1,\cdots,c_n}](T_1,\cdots,T_n), \text{ and} \\ \mathbb{E}[\mathbf{1}_B(F) det\Gamma[F]] &= \sum_{n,c} p_n^{c_1,\cdots,c_n} \int_{\mathbb{R}^n} \mathbf{1}_B det\gamma_n[f_n^{c_1,\cdots,c_n}] k_n^{c_1,\cdots,c_n} dt. \end{aligned}$$

By applying the part 2 of Proposition 2.5.10 for $u = f_n^{c_1,\cdots,c_n}$, we obtain

$$\int_{\mathbb{R}^n} \mathbf{1}_B det\gamma_n[f_n^{c_1,\cdots,c_n}] k_n^{c_1,\cdots,c_n} dt = 0, \forall n \in \mathbb{N}^*.$$

Hence $\mathbb{E}[\mathbf{1}_B(F) det\Gamma[F]] = 0$. □

2.6 Application to SDEs involving the Markov chain

Reminding that the original works on Malliavin calculus aimed to study the existence and the smoothness of densities of solutions to stochastic differential equations, this section is contributed to infer the existence in some sense of a density from the non-degeneracy of the Malliavin covariance matrix and deduce a simple result for Markov chain driven stochastic differential equations.

First notice that every random variable F in Ω can not have a density nor the Malliavin variance $\int_0^T |D_sF|^2 ds$ be a.s strictly positive. Indeed, from the decomposition

$$F = a\mathbf{1}_{A_0} + \sum_{n=1}^{\infty} \sum_{c_1,\cdots,c_n \in \mathcal{K}} \mathbf{1}_{A_n^{c_1,\cdots,c_n}} f_n^{c_1,\cdots,c_n}(T_1,\cdots,T_n)$$

we deduce that the law of F has a point mass at a (since $\mathbb{P}(F=a) \geq \mathbb{P}(N_T=0) = \mathbb{P}(T_1 > T) = e^{-\lambda_1 T} > 0$), and $\int_0^T |D_sF|^2 ds = 0$ on $\{N_T = 0\}$. Therefore, we shall rather give conditions under which $(\mathbf{1}_{\{N_T \geq 1\}}\mathbb{P})F^{-1}$ has a density.
Now we will study the regularity of the solution of a stochastic differential equation driven by the Markov chain (C_t). Let $d \in \mathbb{N}^*$ and consider the SDE

$$X_t = x_0 + \int_0^t f(s, X_s, C_s)ds + \int_0^t g(s, X_{s-}, C_s)dN_s, \tag{2.6.16}$$

or in the differential form

$$dX_t = f(t, X_t, C_t)dt + g(t, X_{t-}, C_t)dN_t, \quad X_0 = x_0.$$

where $x_0 \in \mathbb{R}^d$ fixed, (N_t) is the process of culmulative of jumps of (C_t) defined by (2.2.3) and the functions $f, g : \mathbb{R}^+ \times \mathbb{R}^d \times \mathcal{K} \to \mathbb{R}^d$ are measurable and satisfy

1. $\forall t \in \mathbb{R}^+$, $\forall c \in \mathcal{K}$, the maps $x \mapsto f(t,x,c)$, $x \mapsto g(t,x,c)$ are $\mathcal{C}^1$.
2. $\sup_{(t,x,c)} |\nabla_x f(t,x,c)| + |\nabla_x g(t,x,c)| < +\infty$.

Remark 2.6.9. Here, the term $g(s, X_{s^-}, C_s)$ is not predictable but it is not a real problem since N is of finite variation so that that this equation may be solved pathwise. By adapting the proofs in Bichteler, Gravereaux and Jacod (1987), Jacod (1979) or using the explicit expression of the solution given below, it is clear that the equation (2.6.16) admits a unique solution, X, such that

$$\forall T > 0, \sup_{t\in[0,T]} |X_t| \in \bigcap_{p \geq 1} L^p(\Omega)$$

We would like to apply Theorem 2.5.7 to $F = X_T$. For each $k \in \mathcal{K}$, let $\{\Phi_{s,t}(x,k), t \geq s\}$ denote the deterministic flow defined by

$$\Phi_{s,t}(x,k) = x + \int_s^t f(u, \Phi_{s,u}(x,k), k)du, \text{ or,}$$

$$\partial_t \Phi_{s,t}(x,k) = f(t, \Phi_{s,t}(x,k), k), \quad \Phi_{s,s}(x,k) = x. \tag{2.6.17}$$

We can see that $\Phi_{s,t}(x,k)$ exists and unique for every $x \in \mathbb{R}^d, k \in \mathcal{K}$. On the set $\{N_t = 0\}$, $X_s, 0 \leq s \leq t$ is the solution of

$$dX_s = f(s, X_s, Z_0)ds, \quad X_0 = x_0.$$

Hence we have $X_s = \Phi_{0,s}(x_0, Z_0), 0 \leq s \leq t$, and particularly $X_t = \Phi_{0,t}(x_0, Z_0)$. In case $N_t \geq 1$, we have the following result:

Proposition 2.6.12. *Let Ψ is the map: $(t, x, k) \in \mathbb{R}^+ \times \mathbb{R}^d \times \mathcal{K} \mapsto \Psi(t, x, k) = x + g(t, x, k)$. Then for any $t \geq 0$ and $i \in \mathbb{N}^*$,*

$$X_t = \Phi_{T_i,t}(., Z_i) \circ \Psi(T_i, ., Z_i) \circ \cdots \circ \Phi_{T_1,T_2}(., Z_1) \circ \Psi(T_1, ., Z_1) \circ \Phi_{0,T_1}(x_0, Z_0) \quad (2.6.18)$$

$\mathbb{P}$ a.e. on the set $\{N_t = i\}$.

Proof. We will prove (2.6.18) by induction. On the set $\{N_t = 1\}$, the process (N_t) has a jump at T_1. The equation remain unchanged for $0 \leq s < T_1$, so $X_s = \Phi_{0,s}(x_0, Z_0), 0 \leq s < T_1$. For $T_1 \leq s \leq t$, X_s satisfies the following equation

$$dX_s = f(s, X_s, Z_1)ds, \quad X_{T_1} = \Psi(T_1, X_{T_1-}, Z_1).$$

We obtain $X_s = \Phi_{T_1,s}(X_{T_1}, Z_1) = \Phi_{T_1,s}(\Psi(T_1, \Phi_{0,T_1}(x_0, Z_0), Z_1), Z_1), T_1 \leq s \leq t$. Particularly,

$$X_t = \Phi_{T_1,t}(X_{T_1}, Z_1) = \Phi_{T_1,t}(., Z_1) \circ \Psi(T_1, ., Z_1) \circ \Phi_{0,T_1}(x_0, Z_0).$$

Now suppose that (2.6.18) holds until i. Let consider the set $\{N_t = i + 1\}$, i.e., the set of trajectories of the Markov chain having $i+1$ jumps up to t. For $T_{i+1} \leq s \leq t$, X_s satisfies the following equation

$$dX_s = f(s, X_s, Z_{i+1})ds, \quad X_{T_{i+1}} = \Psi(T_{i+1}, X_{T_{i+1}-}, Z_{i+1}).$$

Hence $X_s = \Phi_{T_{i+1},s}(\Psi(T_{i+1}, X_{T_{i+1}-}, Z_{i+1}), Z_{i+1}))$. We obtain the formula for $\{N_t = i + 1\}$ by using the induction assumption

$$X_{T_{i+1}-} = \Phi_{T_i,T_{i+1}}(., Z_i) \circ \Psi(T_i, ., Z_i) \circ \cdots \circ \Phi_{T_1,T_2}(., Z_1) \circ \Psi(T_1, ., Z_1) \circ \Phi_{0,T_1}(x_0, Z_0).$$

□

From Equation (2.6.17), the process $\partial_t \Phi_{s,t}(x, k)$ satisfies

$$\partial_t \left(\partial_t \Phi_{s,t}(x, k)\right) = \nabla_x f(t, \Phi_{s,t}(x, k), k) \partial_t \Phi_{s,t}(x, k), \quad \partial_t \Phi_{s,s}(x, k) = f(s, x, k),$$

which deduces

$$\partial_t \Phi_{s,t}(x, k) = f(s, x, k) exp\left(\int_s^t \nabla_x f(u, \Phi_{s,u}(x, k), k) du\right).$$

Similarly, the processes $\nabla_x \Phi_{s,t}(x, k)$ satisfies

$$\partial_t (\nabla_x \Phi_{s,t}(x, k)) = \nabla_x f(t, \Phi_{s,t}(x, k), k) \nabla_x \Phi_{s,t}(x, k), \quad \nabla_x \Phi_{s,s}(x, k) = I_d.$$

Hence,

$$\nabla_x \Phi_{s,t}(x, k) = exp\left(\int_s^t \nabla_x f(u, \Phi_{s,u}(x, k), k) du\right).$$

We define

$$K_t(x) = \nabla_x X_t(x), \quad K_t^s(x) = \left(\nabla_x X_t^s\right)(X_s(x)), \forall t \geq s \geq 0,$$

where $(X_t^s(x))_{t\geq s}$ is the solution of

$$X_t^s = x + \int_s^t f(u, X_u^s, C_u)du + \int_s^t g(u, X_{u-}^s, C_u)dN_u. \tag{2.6.19}$$

(K_t) satisfies the following SDE

$$K_t = I_d + \int_0^t \nabla_x f(s, X_s, C_s)K_s ds + \int_0^t \nabla_x g(s, X_{s-}, C_s)K_{s-}dN_s. \tag{2.6.20}$$

Suppose that $det[I_d + \nabla g(., x, .)] \neq 0$ and let $(\bar{K}_t)$ be the solution of the following SDE

$$\bar{K}_t = I_d - \int_0^t \nabla_x f(s, X_s, C_s)\bar{K}_s ds - \int_0^t (I_d + \nabla_x g(s, X_{s-}, C_s))^{-1}\nabla_x g(s, X_{s-}, C_s)\bar{K}_{s-}dN_s. \tag{2.6.21}$$

Proposition 2.6.13. $K_t\bar{K}_t = 1, \forall t \geq 0.$

Proof. Indeed, the process $Y_t = K_t\bar{K}_t$ satisfies $Y_0 = I_d$, and

$$\begin{aligned} dY_t &= \bar{K}_{t-}dK_t + K_{t-}d\bar{K}_t + d[K, \bar{K}]_t \\ &= Y_{t-}(\nabla_x f(t, X_t, C_t)dt + \nabla_x g(t, X_{t-}, C_t)dN_t) \\ &\quad + Y_{t-}(-\nabla_x f(t, X_t, C_t)dt - (I_d + \nabla_x g(t, X_{t-}, C_t))^{-1}\nabla_x g(t, X_{t-}, C_t)dN_t) \\ &\quad - Y_{t-}(\nabla_x g(t, X_{t-}, C_t))^2(I_d + \nabla_x g(t, X_{t-}, C_t))^{-1}dN_t = 0. \end{aligned}$$

□

From Equations (2.6.20) and (2.6.21), we obtain the recurrence property of K_t and $\bar{K}_t$

$$\begin{aligned} K_{T_i} &= (I_d + \nabla_x g(T_i, X_{T_i-}, Z_i))K_{T_i-} \\ \bar{K}_{T_i} &= (I_d + \nabla_x g(T_i, X_{T_i-}, Z_i))^{-1}\bar{K}_{T_i-}. \end{aligned}$$

Proposition 2.6.14. $K_t^s(x) = K_t(x)\bar{K}_s(x), \forall t \geq s \geq 0.$

Proof. We fix $s \geq 0$. From Equation (2.6.19), we deduce the SDE satisfied by K_t^s

$$K_t^s = I_d + \int_s^t \nabla_x f(u, X_u^s(X_s(x)), C_u)K_u^s du + \int_s^t \nabla_x g(u, X_{u-}^s(X_s(x)), C_u)K_{u-}^s dN_u, \text{ or,}$$

$$K_t^s = I_d + \int_s^t \nabla_x f(u, X_u, C_u)K_u^s du + \int_s^t \nabla_x g(u, X_{u-}, C_u)K_{u-}^s dN_u, \forall t \geq s.$$

In addition, K_t satisfies

$$K_t = K_s + \int_s^t \nabla_x f(u, X_u, C_u)K_u du + \int_s^t \nabla_x g(u, X_{u-}, C_u)K_{u-}dN_u, \forall t \geq s. \tag{2.6.22}$$

Therefore, $K_t = K_t^s K_s$. □

Now we study the relationship between the Malliavin derivative and the derivative of flow of (X_t). In the Brownian case, we have (cf. Bouleau and Denis (2015))

$$D_s X_t = f(s, X_s, C_s) K_t^s, \ \forall 0 \leq s \leq t.$$

In our case, the result is follow

Proposition 2.6.15. *Let $\varphi : \mathbb{R}^+ \times \mathbb{R}^d \times \mathcal{K} \to \mathbb{R}^d$ defined by: $\forall (t,x,k) \in \mathbb{R}^+ \times \mathbb{R}^d \times \mathcal{K}$,*

$$\varphi(t,x,k) = f(t, x + g(t,x,k), k) - (I_d + \nabla_x g(t,x,k)) f(t,x,k) - \partial_t g(t,x,k).$$

Then we have

1. $D_s X_T = -\int_0^T K_T^t \varphi(t, X_{t-}, C_t) \left(\frac{t}{T} - \mathbf{1}_{[0,t]}(s)\right) dN_t$.
2. $\Gamma[X_T] = \int_0^T \int_0^T K_T^t \varphi(t, X_{t-}, C_t) \varphi^*(u, X_{u-}, C_u) (K_T^u)^* \left(u \wedge t - \frac{ut}{T}\right) dN_t dN_u$.

Proof. On the set $\{N_T = i\}$, we have, for every $0 \leq j \leq i$,

$$\begin{aligned} X_T^{T_j}(x) &= \Phi_{T_i,T}(., Z_i) \circ \Psi(T_i, ., Z_i) \circ \cdots \circ \Psi(T_{j+1}, ., Z_{j+1}) \circ \Phi_{T_j,T_{j+1}}(x, Z_j) \\ &= F(\Phi_{T_j,T_{j+1}}(x, Z_j)), \end{aligned}$$

where $F = \Phi_{T_i,T}(., Z_i) \circ \Psi(T_i, ., Z_i) \circ \cdots \circ \Psi(T_{j+1}, ., Z_{j+1})$. We will prove that

$$\partial_{T_j} X_T^{T_j}(x) = -\nabla_x X_T^{T_j}(x) f(T_j, x, Z_j). \tag{2.6.23}$$

Indeed, the left hand side equals to

$$-F'(\Phi_{T_j,T_{j+1}}(x, Z_j)) \partial_t \Phi_{T_j,T_{j+1}}(x, Z_j),$$

and the right hand side equals to

$$-F'(\Phi_{T_j,T_{j+1}}(x, Z_j)) \nabla_x \Phi_{T_j,T_{j+1}}(x, Z_j) f(T_j, x, Z_j).$$

The equation (2.6.23) is obtained by using the result

$$\partial_t \Phi_{T_j,T_{j+1}}(x, Z_j) = \nabla_x \Phi_{T_j,T_{j+1}}(x, Z_j) f(T_j, x, Z_j).$$

We have

$$X_{T_j-}(x) = \Phi_{T_{j-1},T_j}(., Z_{j-1}) \circ \Psi(T_{j-1}, ., Z_{j-1}) \circ \cdots \circ \Psi(T_1, ., Z_1) \circ \Phi_{0,T_1}(x, Z_0),$$

which deduces

$$\begin{aligned} \partial_{T_j} X_{T_j-}(x) &= \partial_t \Phi_{T_{j-1},T_j}(., Z_{j-1}) \left(\Psi(T_{j-1}, ., Z_{j-1}) \circ \cdots \circ \Psi(T_1, ., Z_1) \circ \Phi_{0,T_1}(x, Z_0)\right) \\ &= f(T_j, X_{T_j-}(x), Z_j). \end{aligned}$$

Moreover, $X_{T_j}(x) = \Psi(T_j, X_{T_j-}(x), Z_j)$, so

$$\begin{aligned} \partial_{T_j} X_{T_j}(x) &= \partial_t \Psi(T_j, X_{T_j-}(x), Z_j) + \nabla_x \Psi(T_j, X_{T_j-}(x), Z_j) \partial_{T_j} X_{T_j-}(x) \\ &= \partial_t \Psi(T_j, X_{T_j-}(x), Z_j) + \nabla_x \Psi(T_j, X_{T_j-}(x), Z_j) f(T_j, X_{T_j-}(x), Z_j). \end{aligned}$$

As $X_T(x) = X_T^{T_j}(X_{T_j}(x))$, by using Equation (2.6.23) we deduce

$$\begin{aligned}\partial_{T_j} X_T(x) &= \partial_{T_j} X_T^{T_j}(X_{T_j}(x)) + \nabla_x X_T^{T_j}(X_{T_j}(x))\partial_{T_j} X_{T_j}(x) \\ &= -K_T^{T_j}(x)\varphi(T_j, X_{T_j-}(x), Z_j).\end{aligned}$$

Hence,

$$\begin{aligned}D_s X_T &= \sum_{j=1}^{i} \frac{\partial X_T}{\partial T_j} D_s T_j = -\sum_{j=1}^{i} K_T^{T_j}\varphi(T_j, X_{T_j-}, Z_j) D_s T_j \\ &= -\sum_{j=1}^{i} K_T^{T_j}\varphi(T_j, X_{T_j-}, Z_j)\left(\frac{T_j}{T} - \mathbf{1}_{[0,T_j]}(s)\right) \\ &= -\int_0^T K_T^t \varphi(t, X_{t-}, C_t)\left(\frac{t}{T} - \mathbf{1}_{[0,t]}(s)\right) dN_t.\end{aligned}$$

And the second statement follows as

$$\begin{aligned}&\Gamma[X_T] = \int_0^T D_s X_T (D_s X_T)^* ds \\ &= \int_0^T ds\Big(\int_0^T K_T^t \varphi(t, X_{t-}, C_t)\left(\frac{t}{T} - \mathbf{1}_{[0,t]}(s)\right) dN_t \\ &\quad \times \int_0^T \varphi^*(u, X_{u-}, C_u)(K_T^u)^*\left(\frac{u}{T} - \mathbf{1}_{[0,u]}(s)\right) dN_u\Big) \\ &= \int_0^T \int_0^T K_T^t \varphi(t, X_{t-}, C_t)\varphi^*(u, X_{u-}, C_u)(K_T^u)^* \\ &\quad \times \left(\int_0^T \left(\frac{t}{T} - \mathbf{1}_{[0,t]}(s)\right)\left(\frac{u}{T} - \mathbf{1}_{[0,u]}(s)\right) ds\right) dN_t dN_u \\ &= \int_0^T \int_0^T K_T^t \varphi(t, X_{t-}, C_t)\varphi^*(u, X_{u-}, C_u)(K_T^u)^*\left(u \wedge t - \frac{ut}{T}\right) dN_t dN_u.\end{aligned}$$

□

Now we can use the criterion of density in $\mathbb{D}^{1,2}$. We define

$$\mathcal{C} = \{det\left(\int_0^T \int_0^T K_T^t \varphi(t, X_{t-}, C_t)\varphi^*(u, X_{u-}, C_u)(K_T^u)^*\left(u \wedge t - \frac{ut}{T}\right) dN_t dN_u\right) > 0\}.$$

As a consequence of Theorem 2.5.7, we have

Proposition 2.6.16. *(Existence of density of X_T) If $\mathbb{P}(\mathcal{C}) > 0$ then the conditional law of $X_T(x)$ given $\mathcal{C}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$.*

Now we consider the case $d = 1$. For each subset A of $\mathcal{K}$, we denote by N_t^A the number of times that the Markov chain (C_t) passes through A up to t.

Proposition 2.6.17. *Assume that there exists a subset A of $\mathcal{K}$ such that for every $k \in A$ and every $t \in \mathbb{R}^+$, $x \in \mathbb{R}^d$, $\varphi(t, x, k) \neq 0$. Then the conditional law of $X_T(x)$ given $\{N_T^A \geq 1\}$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}$.*

Proof. The results is obtained by using Proposition 2.6.16 and

$$\Gamma[X_T] = \int_0^T \left(\int_0^T K_T^t \varphi(t, X_{t-}, C_t) \left(\frac{t}{T} - \mathbf{1}_{[0,t]}(s) \right) dN_t \right)^2 ds \geq 0.$$

By using the same argument as Proposition 2.5.10, $\Gamma[X_T] = 0$ deduces that

$$K_T^t \varphi(t, X_{t-}, C_t) = 0, \text{ a.s. } \forall t \in [0, T].$$

This can not happen since $K_T^t \neq 0\, a.s. \forall t \in [0,T]$ and $\varphi(t, X_{t-}, C_t) \neq 0\, a.s.$ when $C_t \in A$. □

In case the functions f and g do not depend on t, we can see that the function φ vanishes when $g(x,k) = -x$ and $f(0,k) = 0$. In this case, the solution jumps to 0 at the first jump and then stays there. In the next proposition, we will give a sufficient condition for which the condition in Proposition 2.6.17 will be satisfied.

Theorem 2.6.8. *A sufficient condition for the measure $(\mathbf{1}_{\{N_T \geq 1\}}\mathbb{P})X_T^{-1}$ to be absolutely continuous with respect to one dimensional Lebesgue measure is that:*

$$|W(g,f)(x,k)| > \frac{1}{2} ||f''(.,k)||_\infty ||g(.,k)||_\infty^2, \, \forall x \in \mathbb{R}, k \in \mathcal{K},$$

where $W(g,f)(x,k) = g'(x,k)f(x,k) - f'(x,k)g(x,k)$ is the Wronskian of f and g, and all the derivatives are with respect to x.

Proof. By applying Taylor expansion for $f(x,k)$ on x, we have

$$\begin{aligned} \varphi(x,k) &= f(x + g(x,k), k) - f(x,k) - g'(x,k)f(x,k) \\ &= g(x,k)f'(x,k) + \frac{1}{2} f''(\xi_x, k) g^2(x,k) - g'(x,k)f(x,k) \\ &= \frac{1}{2} f''(\xi_x, k) g^2(x,k) - W(g,f)(x,k), \end{aligned}$$

for some ξ_x between x and $x + g(x,k)$. □

2.7 Computation of Greeks

In this section, we would like to use the same technique as the classical Malliavin calculus for the case of Markov chain to compute greeks by using the integration by part formula. We start with a technical lemma.

Lemma 2.7.4. *Let (a,b) be an open interval of $\mathbb{R}$. Let $(F^x)_{x\in(a,b)}$ and $(G^x)_{x\in(a,b)}$ be two families of random variables, continuously differentiable in $Dom(D)$ depending on the parameter $x \in (a,b)$. Let $m \in \mathcal{H}$ satisfy*

$$D_m F^x \neq 0, \, a.s \, on \, \{\partial_x F^x \neq 0\}, \, x \in (a,b),$$

and such that $mG^x \partial_x F^x / D_m F^x$ is continuous in x in $Dom(\delta)$. We have

$$\frac{\partial}{\partial x} \mathbb{E}[G^x f(F^x)] = \mathbb{E}\left[f(F^x) \delta \left(G^x m \frac{\partial_x F^x}{D_m F^x} \right) \right] + \mathbb{E}[\partial_x G^x f(F^x)]$$

where f is a function such that $f(F^x) \in L^2(\Omega), x \in (a,b)$.

Proof. We first prove Lemma with $f \in \mathcal{C}_b^\infty(\mathbb{R})$.

$$\begin{aligned}
\frac{\partial}{\partial x}\mathbb{E}[G^x f(F^x)] &= \mathbb{E}[G^x f'(F^x)\partial_x F^x] + \mathbb{E}[\partial_x G^x f(F^x)]. \\
&= \mathbb{E}\left[G^x \frac{D_m f(F^x)}{D_m F^x}\partial_x F^x\right] + \mathbb{E}[\partial_x G^x f(F^x)] \\
&= \mathbb{E}\left[G^x \frac{\partial_x F^x}{D_m F^x} D_m f(F^x)\right] + \mathbb{E}[\partial_x G^x f(F^x)] \\
&= \mathbb{E}\left[f(F^x)\delta\left(G^x m \frac{\partial_x F^x}{D_m F^x}\right)\right] + \mathbb{E}[\partial_x G^x f(F^x)].
\end{aligned}$$

The last equation follows from Remark 2.4.4. For an arbitrary function f such that $f(F^x) \in L^2(\Omega), x \in (a,b)$, we can aprroximate f by a sequence $(f_n)_{n\in\mathbb{N}}$ of smooth functions. We have

$$\frac{\partial}{\partial x}\mathbb{E}[G^x f_n(F^x)] = \mathbb{E}\left[f_n(F^x)\delta\left(G^x m \frac{\partial_x F^x}{D_m F^x}\right)\right] + \mathbb{E}[\partial_x G^x f_n(F^x)],$$

so

$$\begin{aligned}
& \left|\mathbb{E}\left[f(F^x)\delta\left(G^x m \frac{\partial_x F^x}{D_m F^x}\right)\right] + \mathbb{E}[\partial_x G^x f(F^x)] - \frac{\partial}{\partial x}\mathbb{E}[G^x f_n(F^x)]\right| \\
= & \left|\mathbb{E}\left[(f(F^x) - f_n(F^x))\left(\delta\left(G^x m \frac{\partial_x F^x}{D_m F^x}\right) + \partial_x G^x\right)\right]\right| \\
\leq & ||f(F^x) - f_n(F^x)||_{L^2(\Omega)} \left|\left|\delta\left(G^x m \frac{\partial_x F^x}{D_m F^x}\right) + \partial_x G^x\right|\right|_{L^2(\Omega)} \\
\to & \; 0 \, as \, n \to \infty.
\end{aligned}$$

We obtain the result for f by using the closability and the fact that

$$\mathbb{E}[G^x f_n(F^x)] \to \mathbb{E}[G^x f(F^x)] \, as \, n \to \infty.$$

□

Again, by Remark 2.4.4, the weight $\delta\left(G^x m \frac{\partial_x F^x}{D_m F^x}\right)$ can be computed as

$$\begin{aligned}
& \delta\left(G^x m \frac{\partial_x F^x}{D_m F^x}\right) = G^x \frac{\partial_x F^x}{D_m F^x}\delta(m) - D_m\left(G^x \frac{\partial_x F^x}{D_m F^x}\right) \\
= & \; G^x \frac{\partial_x F^x}{D_m F^x}\int_0^T m(t) d\tilde{N}_t - G^x \frac{D_m \partial_x F^x}{D_m F^x} + G^x \frac{\partial_x F^x}{(D_m F^x)^2} D^2_{m^2} F^x - D_m G^x \frac{\partial_x F^x}{D_m F^x}.
\end{aligned}$$

2.7.1 First derivatives

For first derivatives, we have $G^x = 1$, so

$$\frac{\partial}{\partial x}\mathbb{E}[f(F^x)] = \mathbb{E}\left[f(F^x)\delta\left(m \frac{\partial_x F^x}{D_m F^x}\right)\right]$$

with

$$\delta^1 = \delta\left(m \frac{\partial_x F^x}{D_m F^x}\right) = \frac{\partial_x F^x}{D_m F^x}\delta(m) - \frac{D_m \partial_x F^x}{D_m F^x} + \frac{\partial_x F^x}{(D_m F^x)^2} D^2_{m^2} F^x.$$

Linear case

In the linear case, we have $F^x = xF$. Hence the weight for first derivatives is

$$\delta^1 = \delta\left(m\frac{\partial_x F^x}{D_m F^x}\right) = \frac{1}{x}\left(\frac{F}{D_m F}\delta(m) - 1 + \frac{F}{(D_m F)^2}D^2_{m^2}F\right). \tag{2.7.24}$$

2.7.2 Second derivatives

For second derivatives, by assuming that $m \in \mathcal{C}^2_c([0,T])$ and applying the technical lemma for $G^x = \delta^1$ we have

$$\frac{\partial^2}{\partial x^2}\mathbb{E}[f(F^x)] = \frac{\partial}{\partial x}\mathbb{E}\left[f(F^x)\delta^1\right] = \mathbb{E}\left[f(F^x)\delta\left(\delta^1 m\frac{\partial_x F^x}{D_m F^x}\right)\right] + \mathbb{E}\left[\partial_x\delta^1 f(F^x)\right],$$

where

$$\begin{aligned}
&\delta\left(\delta^1 m\frac{\partial_x F^x}{D_m F^x}\right) = \delta^1\frac{\partial_x F^x}{D_m F^x}\delta(m) - D_m\left(\delta^1\frac{\partial_x F^x}{D_m F^x}\right)\\
= \;& \delta^1\frac{\partial_x F^x}{D_m F^x}\delta(m) - D_m\delta^1\frac{\partial_x F^x}{D_m F^x} - \delta^1\left(\frac{D_m\partial_x F^x}{D_m F^x} - \frac{\partial_x F^x}{(D_m F^x)^2}D^2_{m^2}F^x\right).
\end{aligned}$$

Linear case

In the linear case, the weight for second derivatives is then

$$\begin{aligned}
&\delta^2 = \partial_x\delta^1 + \delta\left(\delta^1 m\frac{\partial_x F^x}{D_m F^x}\right)\\
= \;& \frac{-\delta_1}{x} + \frac{1}{x}\left(\delta^1\frac{F}{D_m F}\delta(m) - D_m\delta^1\frac{F}{D_m F} - \delta^1\left(1 - \frac{F}{(D_m F)^2}D^2_{m^2}F\right)\right).
\end{aligned}$$

Now we compute $D_m\delta^1$. From Equation (2.7.24), we have

$$\begin{aligned}
xD_m\delta^1 &= D_m\frac{F}{D_m F}\delta(m) + \frac{F}{D_m F}D_m\delta(m) + D_m\frac{F}{(D_m F)^2}D^2_{m^2}F + \frac{F}{(D_m F)^2}D^3_{m^3}F\\
&= \left(1 - \frac{F}{(D_m F)^2}D^2_{m^2}F\right)\delta(m) - \frac{F}{D_m F}\left(\int_0^T \dot{m}(s)\hat{m}(s)d\tilde{N}_s - \int_0^T \lambda(C_s)m^2(s)ds\right)\\
&\quad + \left(1 - \frac{2F}{(D_m F)^2}D^2_{m^2}F\right)\frac{D^2_{m^2}F}{D_m F} + \frac{F}{(D_m F)^2}D^3_{m^3}F.
\end{aligned}$$

where $D_m\delta(m)$ is defined by (2.4.11).

2.7.3 Fundamental computations

Let $F : \mathbb{R}_+ \times \mathbb{N} \times \Omega \to \mathbb{R}$ such that for every $t \in \mathbb{R}_+$, $k \in \mathbb{N}$, $F(t,k)$ is a random variable in $L^2(\Omega)$. The partial finite difference operator ∇_k is defined as

$$\nabla_k F(t,k) = F(t,k) - F(t,k-1).$$

We are now interested in the derivation of the quantities $\int_0^T F(t,N_t)dt$ and $\int_0^T F(t,N_t)dN_t$, which usually appear in the solutions of stochastic differential equations.

Proposition 2.7.18. *Let $m \in \mathcal{H}$ and assume that $F(t,k) \in \mathbb{D}_m^{1,2}$ for every $t \in \mathbb{R}_+, k \in \mathbb{N}$. We have*

$$D_m \int_0^T F(t,N_t)dt = \int_0^T \nabla_k F(t,N_t)\hat{m}(t)dN_t + \int_0^T [D_m F](t,N_t)dt.$$

Proof.

$$\begin{aligned}
D_m \int_0^T F(t,N_t)dt &= D_m \sum_{k=0}^{N_T-1} \int_{T_k}^{T_{k+1}} F(t,k)dt + D_m \int_{T_{N_T}}^T F(t,N_T)dt \\
&= \sum_{k=0}^{N_T-1} \Big(F(T_{k+1},k) D_m T_{k+1} - F(T_k,k) D_m T_k + \int_{T_k}^{T_{k+1}} [D_m F](t,k)dt \Big) \\
&\quad - F(T_{N_T},N_T) D_m T_{N_T} + \int_{T_{N_T}}^T [D_m F](t,N_T)dt \\
&= \sum_{k=1}^{N_T} D_m T_k (F(T_k,k-1) - F(T_k,k)) + \int_0^T [D_m F](t,N_t)dt \\
&= -\sum_{k=1}^{N_T} D_m T_k \nabla_k F(T_k,k) + \int_0^T [D_m F](t,N_t)dt \\
&= \int_0^T \nabla_k F(t,N_t)\hat{m}(t)dN_t + \int_0^T [D_m F](t,N_t)dt.
\end{aligned}$$

□

Proposition 2.7.19. *Let $m \in \mathcal{H}$ and assume that $F(t,k) \in \mathbb{D}_m^{1,2}$ for every $t \in \mathbb{R}_+, k \in \mathbb{N}$, $F(.,k) \in \mathcal{C}_c^1([0,T])$ for every $k \in \mathbb{N}$. We have*

$$D_m \int_0^T F(t,N_t)dN_t = -\int_0^T \partial_t F(t,N_t)\hat{m}(t)dN_t + \int_0^T [D_m F](t,N_t)dN_t$$

Proof. We have

$$\begin{aligned}
D_m \int_0^T F(t,N_t)dN_t &= D_m \sum_{k=0}^{N_T} F(T_k,k) = \sum_{k=0}^{N_T} \partial_t F(T_k,k) D_m T_k + [D_m F](T_k,k) \\
&= -\int_0^T \partial_t F(t,N_t)\hat{m}(t)dN_t + \int_0^T [D_m F](t,N_t)dN_t.
\end{aligned}$$

□

Corollary 2.7.6. *Let $m, u \in \mathcal{H}$ and assume that $F(t,k) \in \mathbb{D}_m^{1,2}$ for every $t \in \mathbb{R}_+, k \in \mathbb{N}$, $F(.,k) \in \mathcal{C}_c^1([0,T])$ a.s. for every $k \in \mathbb{N}$. We have*

$$\begin{aligned}
D_{um}^2 \int_0^T F(t,N_t)dt &= -\int_0^T \Big(m(t)\nabla_k F(t,N_t) + [\partial_t \nabla_k F](t,N_t)\hat{m}(t) \Big)\hat{u}(t)dN_t \\
&+ \int_0^T \Big([D_u \nabla_k F](t,N_t)\hat{m}(t) + \nabla_k [D_m F](t,N_t)\hat{u}(t) \Big)dN_t + \int_0^T [D_{um}^2 F](t,N_t)dt.
\end{aligned}$$

Proof. By applying Proposition 2.7.18, we have

$$D^2_{um}\int_0^T F(t,N_t)dt = D_u\int_0^T \nabla_k F(t,N_t)\hat{m}(t)dN_t + D_u\int_0^T [D_m F](t,N_t)dt.$$

Again, by Proposition 2.7.18, the second term on the right hand side is equal to

$$\int_0^T \nabla_k[D_m F](t,N_t)\hat{u}(t)dN_t + \int_0^T [D^2_{um}F](t,N_t)dt.$$

The first term on the right hand side is equal to, by Proposition 2.7.19,

$$\begin{aligned}
&-\int_0^T \partial_t\Big(\nabla_k F(t,N_t)\hat{m}(t)\Big)\hat{u}(t)dN_t + \int_0^T D_u\Big(\nabla_k F(t,N_t)\hat{m}(t)\Big)dN_t\\
&= -\int_0^T \Big(m(t)\nabla_k F(t,N_t) + [\partial_t\nabla_k F](t,N_t)\hat{m}(t)\Big)\hat{u}(t)dN_t\\
&+ \int_0^T [D_u\nabla_k F](t,N_t)\hat{m}(t)dN_t.
\end{aligned}$$

□

Corollary 2.7.7. *Assume that $F(t,k)$ does not depend on the jump times of (N_t), i.e., $[D_m F](t,k) = 0$, $\forall t\in[0,T]$, $k\in\mathbb{N}$, $m\in\mathcal{H}$, and that $F(.,k)\in\mathcal{C}^2_c([0,T])$ a.s. for every $k\in\mathbb{N}$. Then for every $u,v\in\mathcal{H}$ and $m\in\mathcal{C}^1_c([0,T])$, we have*

$$D^3_{vum}\int_0^T F(t,N_t)dt = \int_0^T \Big(\left(mu+\dot{m}\hat{u}\right)\nabla_k F + (2m\hat{u}+u\hat{m})\,\partial_t\nabla_k F + \hat{m}\hat{u}\partial^2_t\nabla_k F\Big)\hat{v}\,dN_t.$$

Proof. As $F(t,k)$ does not depend on the jump times of (N_t), so these terms $[D_u\nabla_k F](t,N_t)$, $[\nabla_k D_u F](t,N_t)$, $\nabla_k[D_m F](t,N_t)$, $[D^2_{um}F](t,N_t)$ are all equal to 0. By using Corollary 2.7.6, we obtain

$$D^2_{um}\int_0^T F(t,N_t)dt = -\int_0^T \left(m\nabla_k F(t,N_t) + [\partial_t\nabla_k F](t,N_t)\hat{m}\right)\hat{u}dN_t.$$

By taking the directional derivative with respect to v in both sides and using Proposition 2.7.19, $D^3_{vum}\int_0^T F(t,N_t)dt$ is equal to

$$\int_0^T \partial_t\left[\left(m\nabla_k F + \partial_t\nabla_k F\hat{m}\right)\hat{u}\right]\hat{v}dN_t - \int_0^T D_v\left[\left(m\nabla_k F + \partial_t\nabla_k F\hat{m}\right)\hat{u}\right]dN_t,$$

where the second integral is equal to 0 by assumption, and the first integral is equal to

$$\begin{aligned}
&\int_0^T \Big(\left(\dot{m}\nabla_k F + m\partial_t\nabla_k F + m\partial_t\nabla_k F + \partial^2_t\nabla_k F\hat{m}\right)\hat{u}\Big)\hat{v}dN_t\\
&\quad + \int_0^T \Big(\left(m\nabla_k F + \partial_t\nabla_k F\hat{m}\right)u\Big)\hat{v}dN_t\\
&= \int_0^T \Big(\left(mu+\dot{m}\hat{u}\right)\nabla_k F + (2m\hat{u}+u\hat{m})\,\partial_t\nabla_k F + \hat{m}\hat{u}\,\partial^2_t\nabla_k F\Big)\hat{v}\,dN_t.
\end{aligned}$$

□

2.7.4 Market model

In this section we assume that the underlying asset price under the risk-neutral probability follows the following SDE

$$dS_t^x = r(t, C_t)S_t^x dt + \sigma(t, C_t)S_{t-}^x d\tilde{N}_t,\ S_0^x = x, \tag{2.7.25}$$

where r, σ are deterministic function and $\tilde{N}_t = N_t - \int_0^t \lambda(C_s)ds$. Put

$$\alpha(t,k) = r(t,k) - \lambda(k)\sigma(t,k), k \in \mathcal{K},$$

Equation (2.7.25) can be rewritten as

$$dS_t^x = \alpha(t, C_t)S_t^x dt + \sigma(t, C_t)S_{t-}^x dN_t,\ S_0^x = x.$$

Now we are in the framework of Equation (2.6.16) with $f(t,x,k) = \alpha(t,k)x$, $g(t,x,k) = \sigma(t,k)x$. Let $\{\Phi_{s,t}(x,k), t \geq s\}$ be the solution of

$$\partial_t \Phi_{s,t}(x,k) = \alpha(t,k)\Phi_{s,t}(x,k), \quad \Phi_{s,s}(x,k) = x.$$

Therefore,

$$\Phi_{s,t}(x,k) = x\, exp\left(\int_s^t \alpha(u,k)du\right).$$

By applying Proposition 2.6.12, we obtain

$$S_t^x = \Phi_{T_i,t}(.,Z_i) \circ \Psi(T_i,.,Z_i) \circ \cdots \circ \Phi_{T_1,T_2}(.,Z_1) \circ \Psi(T_1,.,Z_1) \circ \Phi_{0,T_1}(x,Z_0),$$

on the set $\{N_t = i\}$ for any $t \geq 0$ and $i \in \mathbb{N}^*$, where $\Psi(t,x,k) = x(1+\sigma(t,k))$, $n \geq 1$. Hence,

$$S_t^x = xS(t,N_t), \text{ where } S(t,k) = exp\left(\int_0^t \alpha(u,C_u)du\right)\prod_{j=1}^k (1+\sigma(T_j,Z_j)).$$

For every $k \in \mathcal{K}, t \geq 0$, we have

$$\begin{aligned}
\nabla_k S(t,k) &= e^{\int_0^t \alpha(u,C_u)du}\prod_{j=1}^k (1+\sigma(T_j,Z_j)) - e^{\int_0^t \alpha(u,C_u)du}\prod_{j=1}^{k-1}(1+\sigma(T_j,Z_j)) \\
&= S(t,k-1)\sigma(T_k,Z_k). \\
[D_m S](t,k) &= S(t,k)\left(D_m\int_0^t \alpha(u,C_u)du + \sum_{j=1}^k \frac{\partial_t\sigma(T_j,Z_j)}{1+\sigma(T_j,Z_j)}D_m T_j\right) \\
&= S(t,k)\sum_{j=1}^k \left(\alpha(T_j,Z_{j-1}) - \alpha(T_j,Z_j) + \frac{\partial_t\sigma(T_j,Z_j)}{1+\sigma(T_j,Z_j)}\right)D_m T_j.
\end{aligned}$$

In our numerical implementation, we choose $r(t,k) = r \in \mathbb{R}^+$, $\sigma(t,k) = k$, then $\alpha(t,k) = r - \lambda(k)k$ does not depend on t, and

$$\begin{aligned}
\nabla_k S(t,k) &= S(t,k-1)Z_k, \\
[D_m S](t,k) &= S(t,k)D_m\hat{\alpha}(t) = S(t,k)\sum_{j=1}^k (\alpha_{j-1} - \alpha_j)\, D_m T_j,
\end{aligned}$$

where $\hat{\alpha}(t) = \int_0^t \alpha(u, C_u)du$, $\alpha_i = r - Z_i\lambda(Z_i)$. By applying Proposition 2.7.18, Proposition 2.7.19, Corollaries 2.7.6 and 2.7.7, we have

$$
\begin{aligned}
D_m \int_0^T S(t, N_t)dt &= \int_0^T S(t, N_t - 1)C_t\hat{m}(t)dN_t + \int_0^T S(t, N_t)D_m\hat{\alpha}(t)dt \\
D^2_{m^2} \int_0^T S(t, N_t)dt &= -\int_0^T \Big(m(t)\nabla_k S(t, N_t) + [\partial_t\nabla_k S](t, N_t)\hat{m}(t)\Big)\hat{m}(t)dN_t \\
&+ \int_0^T \Big([D_m\nabla_k S](t, N_t)\hat{m}(t) + \nabla_k[D_m S](t, N_t)\hat{m}(t)\Big)dN_t + \int_0^T [D^2_{m^2}S](t, N_t)dt \\
&= -\int_0^T \Big(m(t)S(t, N_t - 1)C_t + \partial_t S(t, N_t - 1)C_t\hat{m}(t)\Big)\hat{m}(t)dN_t \\
&+ \int_0^T \Big([D_m S](t, N_t - 1)C_t\hat{m}(t) + \nabla_k S(t, N_t)D_m\hat{\alpha}(t)\hat{m}(t)\Big)dN_t + \int_0^T [D^2_{m^2}S](t, N_t)dt \\
&= -\int_0^T \Big(m(t) + \alpha(t, C_t)\hat{m}(t)\Big)S(t, N_t - 1)C_t\hat{m}(t)dN_t, \\
&+ 2\int_0^T \Big(S(t, N_t - 1)C_t D_m\hat{\alpha}(t)\hat{m}(t)\Big)dN_t + \int_0^T [D^2_{m^2}S](t, N_t)dt.
\end{aligned}
$$

for every $m \in \mathcal{C}^1_c([0, T])$. The last term $[D^2_{m^2}S](t, N_t)$ is equal to

$$
D_m S(t, N_t)D_m\hat{\alpha}(t) + S(t, N_t)D^2_{m^2}\hat{\alpha}(t) = S(t, N_t)\Big(\,(D_m\hat{\alpha}(t))^2 + D^2_{m^2}\hat{\alpha}(t)\Big),
$$

where $D^2_{m^2}\hat{\alpha}(t)$ is equal to

$$
\sum_{j=1}^{N_t} (\alpha_{j-1} - \alpha_j)\, D^2_{m^2}T_j = \sum_{j=1}^{N_t} (\alpha_{j-1} - \alpha_j)\, m(T_j)D_m T_j.
$$

2.7.5 Simulation

In this section, we will compute the delta by two methods, Malliavin calculus and the finite difference method, for a binary (resp. standard) asian call with payoff $1_{[K,\infty[}(\frac{1}{T}\int_0^T S_t^x dt)$ (resp. $(\frac{1}{T}\int_0^T S_t^x dt - K)^+$), where K is the strike price. The call price is computed as $P(x) = \mathbb{E}[f(F^x)]$, where $F^x = xF$, $F = \int_0^T S(t, N_t)dt$ with $f(u) = e^{-rT}1_{[K,\infty[}(\frac{u}{T})$ (resp. $f(u) = e^{-rT}(\frac{u}{T} - K)^+$). With the convention that $T_0 = 0$, $T_{N_T+1} = T$ and $\prod_{i=1}^0(1 + Z_i) = 1$, we have

$$
\begin{aligned}
F &= \int_0^T S(t, N_t)dt = \int_0^T e^{\int_0^t \alpha(u,C_u)du} \prod_{i=1}^{N_t}(1 + Z_i)dt \\
&= \sum_{j=0}^{N_T}\prod_{i=1}^{j}(1 + Z_i)\int_{T_j}^{T_{j+1}} e^{\int_0^t \alpha(u,C_u)du}dt \\
&= \sum_{j=0}^{N_T}\prod_{i=1}^{j}(1 + Z_i)e^{\int_0^{T_j} \alpha(u,C_u)du}\frac{e^{\alpha_j(T_{j+1}-T_j)} - 1}{\alpha_j}.
\end{aligned}
$$

By using the formula of $D_m F, D^2_{m^2} F, D^3_{m^3} F$ computed in the previous section, we receive the first and second derivatives. More precisely,

$$D_m F = \sum_{j=1}^{N_T} S_{T_j-} Z_j \hat{m}(T_j) + \sum_{j=0}^{N_T} D_m \hat{\alpha}(T_j) \prod_{i=1}^{j} (1+Z_i) e^{\int_0^{T_j} \alpha(u,C_u)du} \frac{e^{\alpha_j (T_{j+1}-T_j)} - 1}{\alpha_j},$$

$$D^2_{m^2} F = -\sum_{j=1}^{N_T} \left(m(T_j) + \alpha_j \hat{m}(T_j)\right) S_{T_j-} Z_j \hat{m}(T_j) + 2 \sum_{j=1}^{N_T} S_{T_j-} Z_j D_m \hat{\alpha}(T_j) \hat{m}(T_j)$$
$$+ \sum_{j=0}^{N_T} \left((D_m \hat{\alpha}(T_j))^2 + D^2_{m^2} \hat{\alpha}(T_j) \right) \prod_{i=1}^{j} (1+Z_i) e^{\int_0^{T_j} \alpha(u,C_u)du} \frac{e^{\alpha_j (T_{j+1}-T_j)} - 1}{\alpha_j}.$$

where $S_{T_j^-} := S(T_j, j-1) = e^{\int_0^{T_j} \alpha(u,C_u)dt} \prod_{i=1}^{j-1}(1+Z_i) = e^{\sum_{i=0}^{j-1} \alpha_i (T_{i+1}-T_i)} \prod_{i=1}^{j-1}(1+Z_i)$. For numerical simulation, we consider a Markov chain of two states $\mathcal{K} = \{0.1, 0.2\}$ with corresponding intensities $\lambda(1) = 2.5$, and $\lambda(2) = 1.5$. We take $T = 2$ (years), $x = 10$, $r = 0$, $\hat{m}(t) = sin(\pi t/T)$, $m(t) = \pi/T cos(\pi t/T)$. Hence,

$$\begin{aligned}
\delta(m) &= \int_0^T m(t) d\tilde{N}_t = \int_0^T m(t) dN_t - \int_0^T m(t)\lambda(C_t)dt \\
&= \sum_{i=1}^{N_T} m(T_i) - \lambda(Z_{i-1}) \int_{T_{i-1}}^{T_i} m(t)dt - \lambda(Z_{N_T}) \int_{T_{N_T}}^{T} m(t)dt \\
&= \sum_{i=1}^{N_T} m(T_i) + (\lambda(Z_i) - \lambda(Z_{i-1}))\hat{m}(T_i).
\end{aligned}$$

Remind that the finite difference method gives delta as

$$delta = \frac{P(x+\epsilon) - P(x-\epsilon)}{2\epsilon},$$

where ϵ is taken small enough. We choose $\epsilon = 0,001$.

Figure 2.1 compares the convergence of the delta computed by the finite difference method and the Malliavin method for at the money binary Asian call and standard Asian call. The Mallavin method is more stable than the finite difference method for binary Asian calls and vice versa for the standard Asian calls. Table 2.1 shows more details for different strikes. We observe that the finite difference method converges rapidly for standard Asian calls but explodes for binary Asian calls (the relative standard errors are extremely high). Contrarily, the Malliavin method converges more slowly for standard Asian calls but it works stably for two kinds of call since the Malliavin method uses same weights for every payoff.

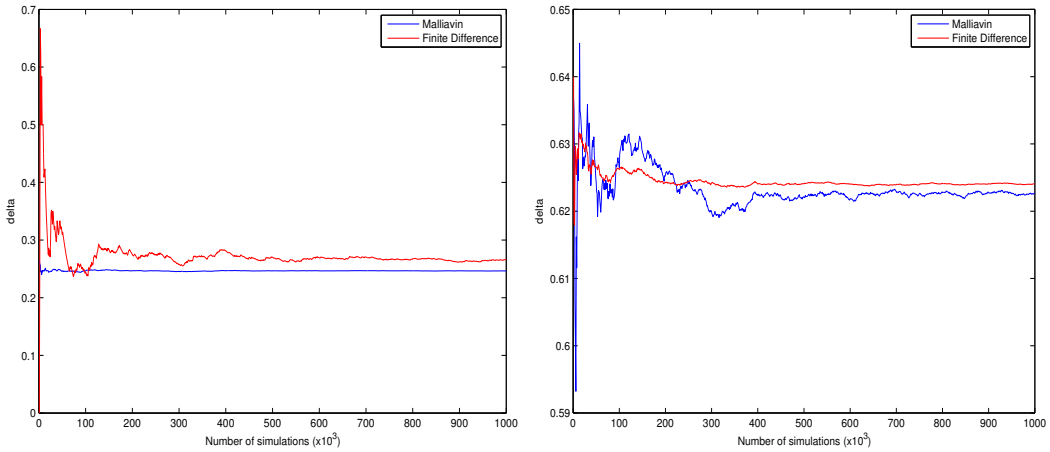

Figure 2.1: Comparison the convergence of the delta computed by the finite difference method and the Malliavin method in case $K = 10$ for binary Asian call *(left)* and standard Asian call *(right)* with 10^6 simulations.

Method	delta	95% CI	Rel. SE
FD	0.454	$[-0.50, 1.41]$	104.8 %
M	0.238	$[0.13, 0.34]$	21.9%

Method	delta	95% CI	Rel. SE
FD	0.875	$[0.85, 0.90]$	1.7%
M	0.869	$[0.68, 1.06]$	10.8%

Method	delta	95% CI	Rel. SE
FD	0.266	$[-0.46, 0.99]$	137.1 %
M	0.246	$[0.19, 0.30]$	11.5%

Method	delta	95% CI	Rel. SE
FD	0.624	$[0.59, 0.66]$	3.0%
M	0.623	$[0.48, 0.76]$	11.1%

Method	delta	95% CI	Rel. SE
FD	0.212	$[-0.44, 0.86]$	153.5%
M	0.209	$[0.16, 0.26]$	11.2%

Method	delta	95% CI	Rel. SE
FD	0.390	$[0.35, 0.43]$	4.8%
M	0.386	$[0.29, 0.49]$	12.8%

Table 2.1: Comparison of deltas (value, 95% confidence interval, and relative standard error) computed by the finite difference (FD) method and the Malliavin (M) method for binary Asian call *(left)* and standard Asian call *(right)* with 10^6 simulations for different strikes: *(top)* $K = 9$, *(middle)* $K = 10$, *(bottom)* $K = 11$.

Chapter 3

Malliavin calculus for Markov chains using chaos decomposition

3.1 Fully homogeneous Markov chain

Let consider a time-continuous homogeneous Markov chain $\hat{C} = (\hat{C}_t)_{t\in[0,T]}$ on a filtered probability space $(\hat{\Omega}, \hat{\mathcal{F}}, (\hat{\mathcal{F}}_t)_{t\in[0,T]}, \hat{\mathbb{P}})$ where $(\hat{\mathcal{F}}_t)_{t\in[0,T]}$ is the natural filtration of $\hat{C}$, with a finite state space $\mathcal{K}$ of l elements, which we shall, for algebraic purpose, identify with $\mathbb{Z}/l\mathbb{Z} = \{\bar{0}, \bar{1}, \cdots, \overline{l-1}\}$. Moreover, we assume that $\hat{C}$ admits the associated infinitesimal generator matrix

$$\Lambda = \begin{pmatrix} -(l-1) & 1 & \cdots & 1 \\ 1 & -(l-1) & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & -(l-1) \end{pmatrix}. \tag{3.1.1}$$

That means at any time $0 \leq t \leq T$, the intensity of jump of $\hat{C}_t$ is $l-1$, and the probability that the Markov chain jump to any other state is equal to $\frac{1}{l-1}$. We call a Markov chain with such a generator matrix a "fully homogeneous Markov chain". We refer to Section 1.4 and Bielecki, Crépey, and Herbertsson (2009) for more properties of $\hat{C}$ and Λ.

3.1.1 Fundamental martingales

For any $k \in \mathcal{K}^* = \mathcal{K}\backslash\{\bar{0}\}$, let the counting process N_t^k denote the number of times that the Markov chain $\hat{C}$ jumps with size k up to time t, i.e., $N_t^k = \sum_{s\leq t} \mathbf{1}_{\{\hat{C}_s = \hat{C}_{s-}+k\}}$ where the sum takes over only a finite number of terms. Then N_t^k is a homogeneous Poisson process with intensity 1, and for $k \neq k'$, the processes (N_t^k) and $(N_t^{k'})$ are independent since they never jump together. The compensated martingales $(\hat{N}^k)_{k\in\mathcal{K}^*}$ are defined by

$$d\hat{N}_t^k = dN_t^k - dt, \hat{N}_0^k = 0.$$

The family of martingales $(\hat{N}^k)_{k\in\mathcal{K}^*}$ together with the initial state of $\hat{C}$ (normally starts at 0) give us the total information about the Markov chain $\hat{C}$, which is the reason why we call $(\hat{N}^k)_{k\in\mathcal{K}^*}$ the fundamental martingales of the model.

Let also N_t be the number of jumps of $\hat{C}$ up to time t

$$N_t = \sum_{k\in\mathcal{K}^*} N_t^k = \sum_{s\leq t} \mathbf{1}_{\{\hat{C}_s \neq \hat{C}_{s-}\}}. \tag{3.1.2}$$

N_t is the sum of $l-1$ homogeneous Poisson processes with intensity 1 which are mutually independent (since they never jump together), so it is also a homogeneous Poisson process with intensity $l-1$ with compensated martingale is defined by

$$d\hat{N}_t = dN_t - (l-1)dt, \hat{N}_0 = 0.$$

3.1.2 Poisson random measure interpretation

Let $(Z_i)_{i\geq 1}$ denote the sequence of successive jump sizes of the Markov chain $\hat{C}$. For every $i \geq 1$, Z_i is a random variable with value in $\mathcal{K}^*$ and the Markov chain $(\hat{C}_t)$ can be represented as follow

$$\hat{C}_t = \sum_{i=1}^{N_t} Z_i, \tag{3.1.3}$$

where N_t is the Poisson process defined by (3.1.2). Since the probability that the Markov chain jump to any other state is equal, the jump sizes $(Z_i)_{i\geq 0}$ are i.i.d. with uniform distribution σ in $\mathcal{K}^*$ and hence, $\hat{C}_t$ is a compound Poisson process.

Let M be the jump measure of $\hat{C}$ then M is a Poisson random measure on $[0,T]\times\mathcal{K}^*$ with intensity measure $\mu(dt,dz) = (l-1)dt\sigma(dz)$ and compensated measure $\tilde{M}(dt,dz) = M(dt,dz) - \mu(dt,dz)$ (see Cont and Tankov (2003)). All results in Section 1.2 follow for this $\tilde{M}$ measure with $\mathbb{R}_0 = \mathcal{K}^*$.

Remark 3.1.10. In our case $\mathbb{R}_0 = \mathcal{K}^*$, the integrals in $L^2(\nu)$ can be simplified as

$$\int_0^T \int_{\mathcal{K}^*} f(t,z)\mu(dt,dz) = \int_0^T \int_{\mathcal{K}^*} \sum_{k\in\mathcal{K}^*} f(t,z)\mathbf{1}_{z=k}(l-1)dt\sigma(dz) = \int_0^T \sum_{k\in\mathcal{K}^*} f(t,k)dt$$

for every real function $f \in L^2(\nu)$. Similarly,

$$\begin{aligned}&\int_0^T \int_{\mathcal{K}^*} \cdots \int_0^T \int_{\mathcal{K}^*} f_n(t_1,z_1,\cdots,t_n,z_n)\mu(dt_1,dz_1)\cdots\mu(dt_n,dz_n)\\ &= \int_0^T \cdots \int_0^T \sum_{k_i\in\mathcal{K}^*} f(t_1,k_1,\cdots,t_n,k_n)dt_1\cdots dt_n\end{aligned}$$

for every real function $f_n \in L^2(\nu^n)$.

3.2 Non-homogeneous Markov chain

Let

$$\Lambda(t) = \begin{pmatrix} \lambda_{0,0}(t) & \lambda_{0,1}(t) & \cdots & \lambda_{0,K-1}(t) \\ \lambda_{1,0}(t) & \lambda_{1,1}(t) & \cdots & \lambda_{1,K-1}(t) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{K-1,0}(t) & \lambda_{K-1,1}(t) & \cdots & \lambda_{K-1,K-1}(t) \end{pmatrix} \tag{3.2.4}$$

which will play the role of an infinitesimal generator matrix associated with a Markov chain, where $\lambda_{i,i}(t) = -\sum_{j\neq i}\lambda_{i,j}(t)$, for every $0 \leq i \leq K-1$, $0 \leq t \leq T$.

3.2.1 Change of measure

In this section, we will show that we can obtain a non-homogeneous Markov chain with the infinitesimal generator matrix $\Lambda(t)$ defined by (3.2.4) from a fully homogeneous Markov chain $\hat{C}_t$ considered in Section 3.1 having infinitesimal generator matrix Λ defined by (3.1.1), by a change of measure.

Let Γ be a $\hat{\mathbb{P}}$-martingale defined by $\Gamma_0 = 1$, and for $0 \leq t \leq T$,

$$\frac{d\Gamma_t}{\Gamma_{t-}} = \sum_{k\in\mathcal{K}^*}\left(\lambda^k(t,\hat{C}_{t-}) - 1\right)d\hat{N}_t^k, \tag{3.2.5}$$

where $\lambda^k(t,\hat{C}_{t-}) = \lambda_{\hat{C}_{t-},\hat{C}_{t-}+k}(t)$. We deduce

$$\Gamma_t = \prod_{k\in\mathcal{K}^*}\mathcal{E}\left(\int_0^t\left(\lambda^k(s,\hat{C}_{s-}) - 1\right)d\hat{N}_s^k\right).$$

where $\mathcal{E}$ signifies the Doléans-Dade exponential.By applying the formula

$$\mathcal{E}(Y_t) = e^{Y_t - Y_0}\prod_{0\leq s\leq t}(1+\Delta Y_s)e^{-\Delta Y_s}$$

for $Y_t = \int_0^t\left(\lambda^k(s,\hat{C}_{s-}) - 1\right)d\hat{N}_s^k$, we have

$$\begin{aligned}
\Gamma_t &= \prod_{k\in\mathcal{K}^*}\left(e^{\int_0^t\left(1-\lambda^k(s,\hat{C}_{s-})\right)ds}\prod_{\tau_k\leq t,\hat{C}_{\tau_k}=\hat{C}_{\tau_k-}+k}\lambda^k(\tau_k,\hat{C}_{\tau_k-})\right)\\
&= e^{\sum_{k\in\mathcal{K}^*}\int_0^t\left(1-\lambda^k(s,\hat{C}_{s-})\right)ds}\prod_{k\in\mathcal{K}^*}\prod_{\tau_k\leq t,\hat{C}_{\tau_k}=\hat{C}_{\tau_k-}+k}\lambda^k(\tau_k,\hat{C}_{\tau_k-})\\
&= e^{\int_0^t\left(\lambda(s,\hat{C}_{s-})+K-1\right)ds}\prod_{\tau\leq t,\hat{C}_\tau\neq\hat{C}_{\tau-}}\lambda(\tau,\hat{C}_{\tau-},\hat{C}_\tau)
\end{aligned} \tag{3.2.6}$$

We use Γ_t as the Radon- Nikodym density to define a measure $\mathbb{P}$ satisfying

$$\frac{d\mathbb{P}}{d\hat{\mathbb{P}}}\Big|_{\mathcal{F}_t} = \Gamma_t,\ 0 \leq t \leq T. \tag{3.2.7}$$

We will prove that under the new measure $\mathbb{P}$, the Markov chain $\hat{C}_{t\in[0,T]}$ has the infinitesimal generator matrix $\Lambda(t)$, i.e., we have to prove that the processes $(N_t^k)_{t\in[0,T]}, k \in \mathcal{K}^*$ defined by

$$d\tilde{N}_t^k = dN_t^k - \lambda^k(t,\hat{C}_{t-})dt, \tilde{N}_0^k = 0$$

are $\mathbb{P}$-martingales. Indeed, for every $k \in \mathcal{K}^*$,

$$\begin{aligned}
d(\tilde{N}_t^k\Gamma_t) &= \tilde{N}_{t-}^k d\Gamma_t + \Gamma_{t-}d\tilde{N}_t^k + d[\Gamma,\tilde{N}^k]_t\\
&= \tilde{N}_{t-}^k d\Gamma_t + \Gamma_{t-}(dN_t^k - \lambda^k(t,\hat{C}_{t-})dt) + \Gamma_{t-}\left(\lambda^k(t,\hat{C}_{t-}) - 1\right)dN_t^k\\
&= \tilde{N}_{t-}^k d\Gamma_t + \Gamma_{t-}\lambda^k(t,\hat{C}_{t-})(dN_t^k - dt)\\
&= \tilde{N}_{t-}^k d\Gamma_t + \Gamma_{t-}\lambda^k(t,\hat{C}_{t-})d\hat{N}_t^k.
\end{aligned}$$

Hence $(\tilde{N}_t^k\Gamma_t)_{t\in[0,T]}$is a $\hat{\mathbb{P}}$-martingale. By the definition of martingale, we have

$$\hat{\mathbb{E}}[\tilde{N}_t^k\Gamma_t|\mathcal{F}_s] = \tilde{N}_s^k\Gamma_s,\ \forall 0 \le s \le t \le T. \tag{3.2.8}$$

From the conditional version of the Bayes formula, we obtain

$$\hat{\mathbb{E}}[\tilde{N}_t^k\Gamma_t|\mathcal{F}_s] = \Gamma_s\mathbb{E}[\tilde{N}_t^k|\mathcal{F}_s],\ \forall 0 \le s \le t \le T. \tag{3.2.9}$$

Moreover, we have $\Gamma_s > 0, \mathbb{P}$ almost surely since

$$P(\Gamma_s = 0) = \mathbb{E}[\mathbf{1}_{\{\Gamma_s=0\}}] = \hat{\mathbb{E}}[\Gamma_T\mathbf{1}_{\{\Gamma_s=0\}}] = 0,$$

combining with (3.2.8) and (3.2.9), we have

$$\mathbb{E}[\tilde{N}_t^k|\mathcal{F}_s] = \tilde{N}_s^k\ \mathbb{P}\,a.s.\,\forall 0 \le s \le t \le T,$$

i.e., $(\tilde{N}_t^k)_{t\in[0,T]}$ is a $\mathbb{P}$-martingale.

3.3 Application in greeking CDO

CDO tranches are credit derivatives that ensure protection to the buyer against losses due to the defaults of the names of a pool of reference entities. Though CDO issuances have become quite rare since the crisis, there is still a huge amount of outstanding CDO contracts which need to be marked to market and hedged up to their maturity dates. Therefore, the task of pricing and greeking CDOs and the credit portfolios in general is still relevant.

We consider a risk neutral pricing model $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ where $T \ge 0$ is a fixed time horizon and $(\mathcal{F}_t)_{t\in[0,T]}$ is the natural filtration of the Markov chain. We denote $\mathcal{N}_n$ the set of all subsets of $\{1, 2, \cdots, n\}$ where n is the number of obligors in the underlying credit portfolio.

3.3.1 Homogeneous-group model

In the homogeneous-group model, the n names of the pool are shared into k groups of $\nu - 1 = \dfrac{n}{k}$ homogeneous obligors (taking n as a multiple of k). The cumulative default processes $N^l, l = 1, \cdots, k$ in the different groups are jointly modeled as a continuous-time d-variate Markov chain $\mathcal{N} = (N^1, \cdots, N^k)$ with the components in $\mathbb{N}_\nu = \{0, 1, \cdots, \nu - 1\}$. So $\mathcal{N}$ lives in the state space $\mathcal{I} = \mathbb{N}_\nu^k$. Moreover, we assume no simultaneous default, so the cumulative default processes N^l never jump together.

The homogeneity in the name of the model comes from the fact that at every time t, knowing the number of defaults in each group, the default intensities of all survivors in the group are the same and given by some pre-default individual intensity functions $\tilde{\lambda}^l : [0, T] \times \mathcal{I} \to \mathbb{R}^+$. Hence, the intensity of jump in the group l is given by

$$\lambda^l(t, \imath) = (\nu - 1 - i_l)\tilde{\lambda}^l(t, \imath),$$

where $\imath = (i_1, \cdots, i_k) \in \mathcal{I}$ is the current state of $\mathcal{N}$, and i_l is the number of defaults in the l^{th} group. The compensated process

$$M_t^l = N_t^l - \int_0^t \lambda^l(s, \mathcal{N}_s)ds$$

is an $\mathcal{F}$-martingale under $\mathbb{P}$.

The cumulative default process N_t on a credit risk portfolio is modeled as

$$N_t = \phi(\mathcal{N}_t) = \sum_{l=1}^{k} N^l. \tag{3.3.10}$$

Given a credit derivative payoff $\xi = \pi(N_T) = \pi(\phi(\mathcal{N}_T)) = \Phi(\mathcal{N}_T)$, where $\Phi = \pi \circ \phi$. In case of CDO tranche with attachement point a and detachement point b,

$$\pi(N_t) = (L_t - a)^+ \wedge (b - a),$$

where $L_t = (1 - R)N_t/n$ represent the relative cumulative porfolio loss process, R is the recovery rate. We have the corresponding price process, by the Markov property of $\mathcal{N}$ assuming zero risk-free rate for simplicity:

$$\Pi_t = \mathbb{E}[\xi|\mathcal{F}_t] = \mathbb{E}[\Phi(\mathcal{N}_T)|\mathcal{F}_t] = \mathbb{E}[\Phi(\mathcal{N}_T)|\mathcal{N}_t] = u(t, \mathcal{N}_t), \text{ for } t \in [0, T], \tag{3.3.11}$$

where $u(t, \imath)$, with $t \in [0, T]$ and $\imath \in \mathcal{I}$ defines the pricing function of the credit derivative. We are interested in the sensitivity of a credit derivative price when there is one more default in some group k, which can be represented as

$$\delta u^l(t, \imath) = u(t, \imath^l) - u(t, \imath),$$

where $\imath^l$ is the state obtained from the state $\imath$ if there is one more default in the group l. Here

$$u(t, \imath) = \mathbb{E}[\Phi(\mathcal{N}_T)|\mathcal{N}_t = \imath], \text{ and } u(t, \imath^l) = \mathbb{E}[\Phi(\mathcal{N}_T)|\mathcal{N}_t = \imath^l].$$

To compute $\delta u^l(t, \imath)$, normally we have to simulate the Markov chain conditionally to $\mathcal{N}_t = \imath$, and then resimulate it conditionally to $\mathcal{N}_t = \imath^l$ since the law of $(\mathcal{N}_T|\mathcal{N}_t = \imath)$ is different from the law of of $(\mathcal{N}_T|\mathcal{N}_t = \imath^l)$. Our aim is looking for some kind of spatial homogeneity of the Markov chain, so that the law of $(\mathcal{N}_T|\mathcal{N}_t = \imath^l)$ can be deduced directly from the law of $(\mathcal{N}_T|\mathcal{N}_t = \imath)$, which permits to avoid the resimulation.

3.3.2 Compound Poisson form

Since there is no simultaneous default, our Markov chain $\mathcal{N}$ can be represented as

$$\mathcal{N}_t = \sum_{i=1}^{N_t} Z_i, \tag{3.3.12}$$

where N_t is the cumulative default process defined by (3.3.10), and $(Z_i)_{i\geq 0}$ is the sequence of jump sizes of $\mathcal{N}$ living in the set

$$\mathcal{Z} = \{0^1 := (1, 0, \cdots, 0), 0^2 := (0, 1, 0, \cdots, 0), \cdots, 0^k := (0, \cdots, 0, 1)\} \subset \mathbb{N}^k.$$

The process $\mathcal{N}_t$ defined by (3.3.12) will become a compound Poisson process if N_t is a Poisson process and $(Z_i)_{i\geq 1}$ are i.i.d. random variables.

3.3.3 Homogeneity in homogeneous-group model

Similarly to Section 3.2, we can construct a Markov chain with intensities $\lambda^l(t,\mathcal{N}_t)$ under $\mathbb{P}$ from a homogeneous Markov chain by a change of measure. More precisely, let consider a Markov chain $\mathcal{N}$ under a probability measure $\hat{\mathbb{P}}$ where all the counting processes N^l have intensity 1. We define the process $(\Gamma_t)_{t\in[0,T]}$ such that

$$\frac{d\Gamma_t}{\Gamma_{t-}} = \sum_{l=1}^{k}(\lambda^l(t,\mathcal{N}_{t-})-1)d\hat{M}_t^l,\ \Gamma_0 = 1, \tag{3.3.13}$$

where $\hat{M}_t^l = N_t^l - t$ is the compensated martingale of the process N_t^l under $\hat{\mathbb{P}}$. Hence Γ_t is a $\hat{\mathbb{P}}$-martingale playing the role of a Radon-Nikodym density in the change of measure and can be computed explicitly by Doléan-Dade exponentials:

$$\begin{aligned}
\Gamma_t &= \prod_{l=1}^{k}\mathcal{E}\left(\int_0^t(\lambda^l(s,\mathcal{N}_{s-})-1)d\hat{M}_s^l\right)\\
&= \prod_{l=1}^{k}e^{\int_0^t\left(1-\lambda^l(s,\mathcal{N}_s)\right)ds}\prod_{\tau^l\le t, N^l_{\tau^l}\neq N^l_{\tau^l-}}\lambda^l(\tau^l,\mathcal{N}_{\tau^l-}) &(3.3.14)\\
&= e^{\int_0^t(k-\lambda(s,\mathcal{N}_s))ds}\prod_{l=1}^{k}\prod_{\tau^l\le t, N^l_{\tau^l}\neq N^l_{\tau^l-}}\lambda^l(\tau^l,\mathcal{N}_{\tau^l-}) &(3.3.15)
\end{aligned}$$

where $\lambda(s,\mathcal{N}_s)=\sum_{l=1}^k\lambda^l(s,\mathcal{N}_s)$ is the intensity of jump of $\mathcal{N}$ at s. In (3.3.14), for each l, the second product takes over all jump times of the process N^l up to t. And in (3.3.15), the double product takes over all jump times of the process N up to t.

Remark 3.3.11. Γ is a function of the trajectory of $\mathcal{N}$, so we would have to write $\Gamma(\mathcal{N})_t$ instead of Γ_t, but we keep using Γ_t for simplicity and will point out the trajectory whenever there may have some risk of confusion.

By defining a change of measure

$$\frac{d\mathbb{P}}{d\hat{\mathbb{P}}} = \Gamma_T,$$

we obtain the process N_t^l having intensity $\lambda^l(t,\mathcal{N}_t)$ under $\mathbb{P}$, or equivalently, M_t^l is a $\mathbb{P}$-martingale for every $l=1,\cdots,k$. Indeed,

$$\begin{aligned}
d(M_t^l\Gamma_t) &= M_{t-}^l d\Gamma_t + \Gamma_{t-}dM_t^l + d[M^l,\Gamma]_t\\
&= M_{t-}^l d\Gamma_t + \Gamma_{t-}(dN_t^l - \lambda^l(t,\mathcal{N}_t)dt) + \Gamma_{t-}(\lambda^l(t,\mathcal{N}_{t-})-1)dN_t^l\\
&= M_{t-}^l d\Gamma_t + \Gamma_{t-}\lambda^l(t,\mathcal{N}_{t-})d\hat{M}_t^l.
\end{aligned}$$

Both M^l and Γ are bounded, so $M^l\Gamma$ is a $\hat{\mathbb{P}}$-martingale which deduces that M^l is a $\mathbb{P}$-martingale.

3.3.4 Greeking problem via Malliavin calculus

In this model, the martingale representation has the form

$$\Pi_t = \Pi_0 + \sum_{l=1}^{k} \int_0^t \delta u^l(s, \mathcal{N}_{s-}) dM_s^l, \tag{3.3.16}$$

where $\delta u^l(t, \mathcal{N}_t)$ is the sensitivity of the pricing function with respect to the jump of the process N_t^l, which we would like to evaluate thanks to the homogeneity under $\hat{\mathbb{P}}$.

Proposition 3.3.20. *For every $t \in [0,T]$ such that $\Gamma_{t-} \neq 0$ and $\lambda^l(t, \mathcal{N}_{t-}) \neq 0$,*

$$\delta u^l(t, \mathcal{N}_{t-}) = \left(\frac{1}{\lambda^l(t, \mathcal{N}_{t-})} \mathbb{E}\left[\frac{\epsilon^+_{t,0^l}(\Gamma_T \xi)}{\Gamma_T} \Big| \mathcal{F}_t \right] - \mathbb{E}[\xi | \mathcal{F}_t] \right), \tag{3.3.17}$$

where $\epsilon^+_{t,z}$ is the creation operator defined in (1.2.10)*. In particular,*

$$\delta u^l(0, \mathcal{N}_0) = \frac{1}{\lambda^l(0, \mathcal{N}_0)} \mathbb{E}\left[\frac{\epsilon^+_{0,0^l}(\Gamma_T \xi)}{\Gamma_T} \right] - \mathbb{E}[\xi]. \tag{3.3.18}$$

Proof. Under the probability $\hat{\mathbb{P}}$, $\mathcal{N}$ has the form (3.3.12), where N_t is a homogeneous Poisson process of intensity k and $(Z_i)_{i \geq 0}$ are i.i.d. with uniform distribution σ on $\mathcal{Z}$. Hence, $\mathcal{N}$ is a compound Poisson process under $\hat{\mathbb{P}}$. Let M be the jump measure of $\mathcal{N}$ then M is a Poisson random measure on $\mathbb{R}_+ \times \mathcal{Z}$ with intensity measure $\mu(dt, dz) = k dt \sigma(dz)$ and the compensated random measure $\tilde{M}(dt, dz) = M(dt, dz) - \mu(dt, dz)$ (see Cont and Tankov (2003)). By applying the Clark-Ocone formula for the random variable $\Gamma_T \xi$ (cf. Theorem 1.2.6) under $\hat{\mathbb{P}}$, we have

$$\Gamma_T \xi = \hat{\mathbb{E}}[\Gamma_T \xi] + \int_0^T \hat{\mathbb{E}}[D_{s,z}(\Gamma_T \xi) | \mathcal{F}_s] \tilde{M}(ds, dz),$$

where $D_{s,z}(\Gamma_T \xi)$ is the Malliavin derivative of $\Gamma_T \xi$ at (s,z) (cf. Definition 1.2.5), and we have chosen a predictable version of the conditional expectation process $\hat{\mathbb{E}}[D_{s,z}(\Gamma_T \xi) | \mathcal{F}_s], s \geq 0$. Hence,

$$\Gamma_t \Pi_t = \Gamma_t \mathbb{E}[\xi | \mathcal{F}_t] = \hat{\mathbb{E}}[\Gamma_T \xi | \mathcal{F}_t] = \hat{\mathbb{E}}[\Gamma_T \xi] + \int_0^t \hat{\mathbb{E}}[D_{s,z}(\Gamma_T \xi) | \mathcal{F}_s] \tilde{M}(ds, dz),$$

where the first equality is based on the definition of Π_t, the second one is from Bayes' formula and the third one is from martingale property of $\tilde{M}$. We deduce

$$d(\Gamma_t \Pi_t) = \hat{\mathbb{E}}[D_{t,z}(\Gamma_T \xi) | \mathcal{F}_t] \tilde{M}(dt, dz) = \sum_{l=1}^{k} \hat{\mathbb{E}}[D_{t,0^l}(\Gamma_T \xi) | \mathcal{F}_t] d\hat{M}_t^l. \tag{3.3.19}$$

Moreover, from (3.3.13) and (3.3.16), we obtain

$$\begin{aligned} d(\Gamma_t\Pi_t) &= \Gamma_{t-}d\Pi_t + \Pi_{t-}d\Gamma_t + d[\Pi,\Gamma]_t \\ &= \Gamma_{t-}\sum_{l=1}^{k}\delta u^l(t,\mathcal{N}_{t-})dM_t^l + \Pi_{t-}\Gamma_{t-}\sum_{l=1}^{k}(\lambda^l(t,\mathcal{N}_{t-})-1)d\hat{M}_t^l \\ &\quad +\Gamma_{t-}\sum_{l=1}^{k}\delta u^l(t,\mathcal{N}_{t-})(\lambda^l(t,\mathcal{N}_{t-})-1)dN_t^l \\ &= \Gamma_{t-}\sum_{l=1}^{k}[\delta u^l(t,\mathcal{N}_{t-})\lambda^l(t,\mathcal{N}_{t-}) + \Pi_{t-}(\lambda^l(t,\mathcal{N}_{t-})-1)]d\hat{M}_t^l \quad (3.3.20) \end{aligned}$$

By identifying (3.3.19) and (3.3.20) we get

$$\Gamma_{t-}[\delta u^l(t,\mathcal{N}_{t-})\lambda^l(t,\mathcal{N}_{t-}) + \Pi_{t-}(\lambda^l(t,\mathcal{N}_{t-})-1)] = \hat{\mathbb{E}}[D_{t,0^l}(\Gamma_T\xi)|\mathcal{F}_t].$$

But applying (1.2.11), we have $D_{t,0^l}(\Gamma_T\xi) = \epsilon^+_{t,0^l}(\Gamma_T\xi) - \Gamma_T\xi$, and

$$\hat{\mathbb{E}}[D_{t,0^l}(\Gamma_T\xi)|\mathcal{F}_t] = \hat{\mathbb{E}}[\epsilon^+_{t,0^l}(\Gamma_T\xi) - \Gamma_T\xi|\mathcal{F}_t] = \hat{\mathbb{E}}[\epsilon^+_{t,0^l}(\Gamma_T\xi)|\mathcal{F}_t] - \Gamma_{t-}\Pi_{t-},$$

where a predictable version of the conditional expectation $\hat{\mathbb{E}}[\epsilon^+_{t,0^l}(\Gamma_T\xi)|\mathcal{F}_t]$ is used. Therefore,

$$\Gamma_{t-}\lambda^l(t,\mathcal{N}_{t-})[\delta u^l(t,\mathcal{N}_{t-}) + \Pi_{t-}] = \hat{\mathbb{E}}[\epsilon^+_{t,0^l}(\Gamma_T\xi)|\mathcal{F}_t] = \Gamma_{t-}\mathbb{E}\left[\frac{\epsilon^+_{t,0^l}(\Gamma_T\xi)}{\Gamma_T}\Big|\mathcal{F}_t\right], \quad (3.3.21)$$

(with the convention that the ratio $\frac{\epsilon^+_{t,0^l}(\Gamma_T\xi)}{\Gamma_T}$ equals to 0 when $\Gamma_T = 0$, which deduce also $\epsilon^+_{t,0^l}(\Gamma_T\xi) = 0$). In the case $\Gamma_{t-} \neq 0$ and $\lambda^l(t,\mathcal{N}_{t-}) \neq 0$, we deduce

$$\delta u^l(t,\mathcal{N}_{t-}) + \Pi_{t-} = \frac{1}{\lambda^l(t,\mathcal{N}_{t-})}\mathbb{E}\left[\frac{\epsilon^+_{t,0^l}(\Gamma_T\xi)}{\Gamma_T}|\mathcal{F}_t\right].$$

□

Remark 3.3.12. In case the payoff depends only on the final state of the portfolio, we can obtain (3.3.21) by the Bayes formula and the Markov property. Indeed, by using the Bayes' formula and (3.3.11), we have

$$\Gamma_t u(t,\mathcal{N}_t) = \hat{\mathbb{E}}[\Gamma_T\xi|\mathcal{F}_t],$$

or more precisely,

$$\Gamma(\mathcal{N})_t u(t,\mathcal{N}_t) = \hat{\mathbb{E}}[\Gamma(\mathcal{N})_T\xi(\mathcal{N}_T)|\mathcal{F}_t].$$

By applying this formula for $\epsilon^+_{t,0^l}\mathcal{N}$, noticing that $(\Gamma(\epsilon^+_{t,0^l}\mathcal{N}))_t = \Gamma_t\lambda^l(t,\mathcal{N}_t)$ (by (3.3.15)), we obtain

$$\Gamma_t\lambda^l(t,\mathcal{N}_t)u(t,\mathcal{N}_t + 0^l) = \hat{\mathbb{E}}[\epsilon^+_{t,0^l}(\Gamma_T\xi)|\mathcal{F}_t],$$

and (3.3.21) follows directly since

$$u(t,\mathcal{N}_t + 0^l) = u(t,\mathcal{N}_t + 0^l) - u(t,\mathcal{N}_t) + u(t,\mathcal{N}_t) = \delta u^l(t,\mathcal{N}_t) + \Pi_t.$$

3.3.5 Min-variance hedging CDO tranches

Now we consider the problem of min-variance hedging an equity or senior tranche by the index, i.e. the tranche with attachement point $a = 0$ and detachement point $b = 0$. Let Π and P (resp. u and v) denote the cumulative price processes (resp. pricing functions) of the tranche and of the index. By application of the formula (67) in Crépey (2013), we can min-variance hedge a tranche by the index and the riskless constant asset by using the strategy ζ in the index defined by

$$\zeta_t = \frac{\sum_{l=1}^d \lambda^l (\delta u^l)(\delta v^l)}{\sum_{l=1}^d \lambda^l (\delta v^l)^2}(t, \mathcal{N}_{t-}) = \sum_{l=1}^d w^l \left(\frac{\delta u^l}{\delta v^l}\right) \text{ with } w^l = \frac{(\delta v^l)^2}{\sum_{j=1}^d \lambda^j (\delta v^j)^2}, \tag{3.3.22}$$

for $t \in [0, T]$, where δu^l and δv^l can be computed by (3.3.17) and (3.3.18). In case of a local intensity model ($k = 1$), the martingale representation (3.3.16) reduces to

$$d\Pi_t = \delta u(t, \mathcal{N}_{t-}) dM_t, \; dP_t = \delta v(t, \mathcal{N}_{t-}) dM_t.$$

Therefore,

$$d\Pi_t = \delta_t dP_t, \text{ where } \delta_t = \delta(t, \mathcal{N}_{t-}) = \frac{u(t, \mathcal{N}_t) - u(t, \mathcal{N}_{t-})}{v(t, \mathcal{N}_t) - v(t, \mathcal{N}_{t-})}. \tag{3.3.23}$$

In this case, it is thus possible to replicate the tranche by the index using the strategy δ_t defined by (3.3.23), which coincides with the min-variance hedging strategy ζ_t in (3.3.22).

3.3.6 Numerical results

We compute the delta and the replication strategy of the equity tranche and the senior tranche with the same strike $k = 20\%$ or $k = 45\%$ (this is the detachment point of the equity tranche and attachment point of the senior tranche, the attachment point of the equity tranche is 0, the detachment point of the senior tranche is 100%) by the index with maturity $T = 5$, the recovery rate $R = 0.4$ and then compare with the explicit results computed by matrix exponentiation method. Without loss of generality, the nominal is set to 1. For the simulation/regression method, we use $m = 4 \times 10^4$ simulations, and for Monte-Carlo simulation based on spatial homogeneity, we use $m = 10^4$ or $m = 10^6$ simulations.

- **Fully-homogeneous model**
  This is the special case where $k = 1$. The number of obligors $n = 8$. This may be not realistic since a CDO contract usually contains over 100 names. In general, our method will work similarly but there will be difficulties in the matrix exponentiation method when we increase the number of groups. Hence, we restrict in a small number names to better illustrate the results. The pre-default individual intensity function is given by

  $$\tilde{\lambda}(i) = \frac{1 + i}{n}.$$

  Tables 3.1 and 3.2 show the numerical results of deltas by replicating the equity tranche and the senior tranche by the index.

$k = 45\%$	val δ	err $\hat{\delta}_1^1$	err $\hat{\delta}_s^1$	err $\hat{\delta}_s^2$
Eq	0.415	0.292	-7.972	-0.090
Sen	0.585	-0.207	5.659	0.064

Table 3.1: Exact values (column 2) and percentage relative errors for various estimators $\delta = \frac{\delta u_0^{\pm}(0)}{\delta v_0(0)}$ by simulation/regression method (column 3) and by simulation based on spatial homogeneity with $m = 10^4$ simulations (column 4) or $m = 10^6$ simulations (column 5) in the fully-homogeneous model ($u^{\pm}$ = equity or senior pricing function, v =index pricing function).

$k = 20\%$	val δ	err $\hat{\delta}_1^1$	err $\hat{\delta}_s^1$	err $\hat{\delta}_s^2$
Eq	0.065	0.587	-25.071	-0.611
Sen	0.935	-0.041	1.746	0.043

Table 3.2: Like Table 3.1, but for $k = 20\%$.

- **Semi-homogeneous group model**
  We also consider $n = 8$ names divided into $k = 2$ groups. The pre-default individual intensity function of each group is given by

$$\tilde{\lambda}^l(\imath) = l\frac{1 + i_l}{n}.$$

  We keep the other parameters as in the local intensity model. The results of deltas by hedging the equity tranche and the senior tranche by the index are shown in tables 3.3 and 3.4.

$k = 45\%$	val δ	err $\hat{\delta}_1^1$	err $\hat{\delta}_s^1$	err $\hat{\delta}_s^2$
Eq1	0.395	0.052	3.223	0.454
Eq2	0.532	-6.330	14.883	0.342
Sen1	0.605	-0.034	-2.100	-0.296
Sen2	0.468	7.188	-16.899	-0.388

Table 3.3: Exact values (column 2) and percentages relative errors for estimators of $\delta = \frac{\delta^1 u^{\pm}}{\delta^1 v}(0,0,0)$ or $\frac{\delta^2 u^{\pm}}{\delta^2 v}(0,0,0)$ by simulation/regression method (column 3) and by simulation based on spatial homogeneity with $m = 10^4$ simulations (column 4) or $m = 10^6$ simulations (column 5) in the semi-homogeneous model ($u^{\pm}$ = equity or senior pricing function, v =index pricing function).

In tables 3.1 to 3.4, we show the exact deltas and the errors of estimates. $\hat{\delta}_1^1$ is the best simulation/regression estimate obtained in Crépey and Rahal (2013), where the indices mean that the regression is affine in time and restricted to the scenarios where the first default takes place before $T_1 = 1$ year. The error of this estimator, and also of estimators by the simulation/regression method in general, varies a lot

$k = 20\%$	val δ	err $\hat{\delta}_1^1$	err $\hat{\delta}_s^1$	err $\hat{\delta}_s^2$
Eq1	0.009	4.901	116.482	18.608
Eq2	0.043	1.878	230.300	1.036
Sen1	0.991	-0.044	-1.038	-0.166
Sen2	0.957	-0.084	-10.327	-0.046

Table 3.4: Like Table 3.3, but for $k = 20\%$.

with respect to the parameters of the problem, whereas our estimators $\hat{\delta}_s$ are quite robust in the sense that their errors are stable with respect to the parameters of the problem and, unlike $\hat{\delta}_1^1$, they are unbiased. Therefore, more precise results can be obtained by increasing the number of simulations as we did with $\hat{\delta}_s^1$ and $\hat{\delta}_s^2$.

Part II

Counterparty risk and multi-curve modeling

Chapter 4

Rational multi-curve models with counterparty risk valuation adjustments

.

4.1 Introduction

In this work we endeavour to develop multi-curve interest rate models which extend to counterparty risk models in a consistent fashion. The aim is the pricing and risk management of financial instruments with price models capable of discounting at multiple rates (e.g. OIS and LIBOR) and which allow for corrections in the asset's valuation scheme so to adjust for counterparty-risk inclusive of credit, debt, and liquidity risk. We thus propose factor-models for (i) the Overnight Index Swap (OIS) rate, (ii) the London Interbank Offer Rate (LIBOR), and (iii) the default intensities of two counterparties involved in bilateral OTC derivative transactions. The three ingredients are characterised by a feature they share in common: the rate and intensity models are all rational functions of the underlying factor processes. Since we have in mind the pricing of assets as well as the management of risk exposures, we also need to work within a setup that maintains price consistency under various probability measures. We will for instance want to price derivatives by making use of a risk-neutral measure $\mathbb{Q}$ while analysing the statistics of risk exposures under the real-world measure $\mathbb{P}$. This point is particularly important when we calibrate the interest rate models to derivatives data, such as implied volatilities, and then apply the calibrated models to compute counterparty-risk valuation adjustments to comply with regulatory requirements. The presented rational models allow us to develop a comprehensive framework that begins with an OIS model, evolves to an approach for constructing the LIBOR process, includes the pricing of fixed-income assets and model calibration, analyses risk exposures, and concludes with a credit risk model that leads to the analysis of counterparty-risk valuation adjustments (XVA).

The issue of how to model multi-curve interest rates and incorporate counterparty-risk valuation adjustments in a pricing framework has motivated much research. For instance, research on multi-curve interest rate modelling is presented in Kijima,

Tanaka, and Wong (2009), Kenyon (2010), Henrard (2007, 2010, 2014), Bianchetti (2010), Mercurio (2010b, 2010a, 2010c), Fujii, Shimada, and Takahashi (2011, 2010), Moreni and Pallavicini (2014), Bianchetti and Morini (2013), Filipović and Trolle (2013) or Crépey, Grbac, Ngor and Skovmand (2015). On counterparty-risk valuation adjustment, we mention two recent books by Brigo, Morini, and Pallavicini (2013) and Crépey, Bielecki and Brigo (2014); more references are given as we go along. Pricing models with rational form have also appeared before. Flesaker and Hughston (1996) pioneered such pricing models and in particular introduced the so-called rational log-normal model for discount bond prices. Further related studies include Rutkowski (1997), Döberlein and Schweizer (2001) and Hunt and Kennedy (2004), Brody and Hughston (2004), Hughston and Rafailidis (2005), Brody, Hughston and Mackie (2012), Akahori, Hishida, Teichmann and Tsuchiya (2014), Filipović, Larsson and Trolle (2014), Macrina and Parbhoo (2014) or Nguyen and Seifried (2014). However, as far as we know, the present work is the first to apply rational pricing models in a multi-curve setup, along with Nguyen and Seifried (2014) who develop a rational multi-curve model in the spirit of Rogers (1997) based on a multiplicative spread, and it is the only rational pricing work dealing with XVA computations. We shall see that, despite the simplicity of these models, they perform surprisingly well when comparing to other, in principle more elaborate, proposals such as Crépey et al. (2015) or Moreni and Pallavicini (2013, 2014). Other recent related research includes Filipović, Larsson and Trolle (2014), for the study of unspanned volatility and its regulatory implications, Cuchiero, Keller-Ressel and Teichmann (2012), for moment computations in financial applications, and Cheng and Tehranchi (2014), motivated by stochastic volatility modelling.

We give a brief overview of this chapter. In Section 4.2, we introduce the rational models for multi-curve term structures whereby we derive the forward LIBOR process by pricing a forward rate agreement under the real-world probability measure. In doing so we apply a pricing kernel model. The short rate model arising from the pricing kernel process is taken as a proxy model for the OIS rate. In view of derivative pricing in subsequent sections, we also derive the multi-curve interest rate models by starting with the risk-neutral measure. We call this method "bottom-up risk-neutral approach". In Section 4.3, we perform the so-called "clean valuation" of swaptions written on LIBOR, and analyse three different specifications for the OIS-LIBOR dynamics. We explain the advantages one gains from the chosen "codebook" for the LIBOR process, which we model as a rational function where the denominator is the stochastic discount factor associated with the utilised probability measure. In Section 4.4, we calibrate the three specified multi-curve models and assess them for the quality of fit and on positivity of rates and spread. We conclude by singling out a two-factor lognormal OIS-LIBOR model for its good tractability and calibration properties. In Section 4.5, we price a basis swap in closed form without taking into account counterparty-risk, that is we again perform a "clean valuation". In this section we take the opportunity to show the explicit relationship in our setup between pricing under an equivalent measure and the real-world measure. We compute the risk exposure associated with holding a basis swap and plot the quantiles under both probability measures for comparison. As an example, we apply Lévy random bridges to describe the dynamics of the factor processes under $\mathbb{P}$. This enables us to interpret the re-weighting of the risk exposure under $\mathbb{P}$ as an

effect that could be related to, e.g., "forward guidance" provided by a central bank. In the last section, we present default intensity processes with rational form and compute XVA, that is, the valuation adjustments due to credit, debt, and liquidity risk.

4.2 Rational multi-curve term structures

We model a financial market by a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\}_{0\leq t})$, where $\mathbb{P}$ denotes the real probability measure and $\{\mathcal{F}_t\}_{0\leq t}$ is the market filtration. The no-arbitrage pricing formula for a generic (non-dividend-paying) financial asset with price process $\{S_{tT}\}_{0\leq t\leq T}$, which is characterised by a cash flow S_{TT} at the fixed date T, is given by

$$S_{tT} = \frac{1}{\pi_t}\mathbb{E}^{\mathbb{P}}[\pi_T S_{TT} \,|\, \mathcal{F}_t], \tag{4.2.1}$$

where $\{\pi_t\}_{0\leq t\leq U}$ is the pricing kernel embodying the inter-temporal discounting and risk-adjustments, see e.g. Hunt and Kennedy (2004). Once the model for the pricing kernel is specified, the OIS discount bond price process $\{P_{tT}\}_{0\leq t\leq T\leq U}$ is determined as a special case of formula (4.2.1) by

$$P_{tT} = \frac{1}{\pi_t}\mathbb{E}^{\mathbb{P}}[\pi_T \,|\, \mathcal{F}_t]. \tag{4.2.2}$$

The associated OIS short rate of interest is obtained by

$$r_t = -\left(\partial_T \ln P_{tT}\right)|_{T=t}, \tag{4.2.3}$$

where it is assumed that the discount bond system is differentiable in its maturity parameter T. The rate $\{r_t\}$ is non-negative if the pricing kernel $\{\pi_t\}$ is a super-martingale and vice versa. We next go on to infer a pricing formula for financial derivatives written on LIBOR. In doing so, we also derive a price process (4.2.6) that we identify as determining the dynamics of the forward LIBOR or, as we shall call it for brevity, the LIBOR process. It is this formula for the LIBOR process that reveals the nature of the so-called multi-curve term structure whereby the OIS rate and the LIBOR rates of different tenors are treated as distinct discount rates.

4.2.1 Generic multi-curve interest rate models

We derive multi-curve pricing models for securities written on the LIBOR by starting with the valuation of a forward rate agreement (FRA). We consider $0 \leq t \leq T_0 \leq T_2 \leq \cdots \leq T_i \leq \cdots \leq T_n$, where $T_0, T_i, \ldots, T_n$ are fixed dates, and let N be a notional, K a strike rate and $\delta_i = T_i - T_{i-1}$. The fixed leg of the FRA contract is given by $NK\delta_i$ and the floating leg payable in arrear at time T_i is modelled by $N\delta_i L(T_i; T_{i-1}, T_i)$, where the random rate $L(T_i; T_{i-1}, T_i)$ is $\mathcal{F}_{T_{i-1}}$-measurable. The net cash flow at the maturity date T_i of the FRA contract reads

$$H_{T_i} = N\delta_i \left[K - L(T_i; T_{i-1}, T_i)\right]. \tag{4.2.4}$$

The FRA price process is then given by an application of (4.2.1), that is, for $0 \leq t \leq T_{i-1}$, by

$$\begin{aligned} H_{tT_i} &= \frac{1}{\pi_t}\mathbb{E}^{\mathbb{P}}\left[\pi_{T_i} H_{T_i} \,\middle|\, \mathcal{F}_t\right] \\ &= N\delta_i \left[K P_{tT_i} - L(t, T_{i-1}, T_i)\right], \end{aligned} \tag{4.2.5}$$

where we define the (forward) LIBOR process by

$$L(t; T_{i-1}, T_i) := \frac{1}{\pi_t}\mathbb{E}^{\mathbb{P}}\left[\pi_{T_i} L(T_i; T_{i-1}, T_i) \,\middle|\, \mathcal{F}_t\right]. \tag{4.2.6}$$

The fair spread of the FRA at time t (the value of K at time t such that $H_{tT_i} = 0$) is then expressed in terms of $L(t; T_{i-1}, T_i)$ by

$$K_t = \frac{L(t; T_{i-1}, T_i)}{P_{tT_i}}. \tag{4.2.7}$$

For times up to and including T_{i-1}, our LIBOR process can be written in terms of a conditional expectation of an $\mathcal{F}_{T_{i-1}}$-measurable random variable. In fact, for $t \leq T_{i-1}$,

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}\left[\pi_{T_i} L(T_i; T_{i-1}, T_i) \,\middle|\, \mathcal{F}_t\right] &= \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\pi_{T_i} L(T_i; T_{i-1}, T_i) \,\middle|\, \mathcal{F}_{T_{i-1}}\right] \,\middle|\, \mathcal{F}_t\right] & (4.2.8) \\ &= \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\pi_{T_i} \,\middle|\, \mathcal{F}_{T_{i-1}}\right] L(T_i; T_{i-1}, T_i) \,\middle|\, \mathcal{F}_t\right], & (4.2.9) \end{aligned}$$

and thus

$$L(t, T_{i-1}, T_i) = \frac{1}{\pi_t}\mathbb{E}^{\mathbb{P}}\left[\mathbb{E}^{\mathbb{P}}\left[\pi_{T_i} \,\middle|\, \mathcal{F}_{T_{i-1}}\right] L(T_i; T_{i-1}, T_i) \,\middle|\, \mathcal{F}_t\right]. \tag{4.2.10}$$

The (pre-crisis) classical approach to LIBOR modelling defines the price process $\{H_{tT_i}\}$ of a FRA by

$$H_{tT_i} = N\left[(1 + \delta_i K) P_{tT_i} - P_{tT_{i-1}}\right], \tag{4.2.11}$$

see, e.g., Hunt and Kennedy (2004). By equating with (4.2.5), we see that the classical single-curve LIBOR model is obtained in the special case where

$$L(t; T_{i-1}, T_i) = \frac{1}{\delta_i}\left(P_{tT_{i-1}} - P_{tT_i}\right). \tag{4.2.12}$$

Remark 4.2.1. In normal market conditions, one expects the positive-spread relation $L(t; T, T+\delta_i) < L(t; T, T+\delta_j)$, for tenors $\delta_j > \delta_i$, to hold. We will return to this relationship in Section 4.4 where various model specifications are calibrated and the positivity of the spread is checked. LIBOR tenor spreads play a role in the pricing of basis swaps, which are contracts that exchange two LIBORs with different tenor (see Section 4.5). For recent work on multi-curve modelling with focus on spread modelling, we refer to Cuchiero, Fontana and Gnoatto (2014).

4.2.2 Multi-curve models with rational form

In order to construct explicit LIBOR processes, the pricing kernel $\{\pi_t\}$ and the random variable $L(T_i; T_{i-1}, T_i)$ need to be specified in the definition (4.2.6). For reasons that will become apparent as we move forward in this chapter, we opt to apply the rational pricing models proposed in Macrina (2014). These models bestow a rational form on the price processes, here intended as a "quotient of summands" (slightly abusing the terminology that usually refers to a "quotient of polynomials"). The basic pricing model with rational form for a generic financial asset (for short "rational pricing model") that we consider is given by

$$S_{tT} = \frac{S_{0T} + b_2(T)A_t^{(2)} + b_3(T)A_t^{(3)}}{P_{0t} + b_1(t)A_t^{(1)}}, \tag{4.2.13}$$

where S_{0T} is the value of the asset at $t = 0$. There may be more bA-terms in the numerator, but two (at most) will be enough for all our purposes in this work. For $0 \leq t \leq T$ and $i = 1, 2, 3$, $b_i(t)$ are deterministic functions and $A_t^{(i)} = A_i(t, X_t^{(i)})$ are martingale processes, not necessarily under $\mathbb{P}$ but under an equivalent martingale measure $\mathbb{M}$, which are driven by $\mathbb{M}$-Markov processes $\{X_t^{(i)}\}$. The details of how the expression (4.2.13) is derived from the formula (4.2.1), and in particular how explicit examples for $\{A_t^{(i)}\}$ can be constructed, are shown in Macrina (2014). Here we only give the pricing kernel model associated with the price process (4.2.13), that is

$$\pi_t = \frac{\pi_0}{M_0}\left[P_{0t} + b_1(t)A_t^{(1)}\right] M_t, \tag{4.2.14}$$

where $\{M_t\}$ is the $\mathbb{P}$-martingale that induces the change of measure from $\mathbb{P}$ to an auxiliary measure $\mathbb{M}$ under which the $\{A_t^{(i)}\}$ are martingales. The deterministic functions P_{0t} and $b_1(t)$ are defined such that $P_{0t} + b_1(t)A_t^{(1)}$ is a non-negative $\mathbb{M}$-supermartingale (see e.g. Example 4.2.1), and thus in such a way that $\{\pi_t\}$ is a non-negative $\mathbb{P}$-supermartingale. By the equations (4.2.2) and (4.2.3), it is straightforward to see that

$$P_{tT} = \frac{P_{0T} + b_1(T)A_t^{(1)}}{P_{0t} + b_1(t)A_t^{(1)}}, \quad r_t = -\frac{\dot{P}_{0t} + \dot{b_1}(t)A_t^{(1)}}{P_{0t} + b_1(t)A_t^{(1)}}, \tag{4.2.15}$$

where the "dot-notation" means differentiation with respect to time t.

Let us return to the modelling of rational multi-curve term structures and in particular to the definition of the (forward) LIBOR process. Putting equations (4.2.6) and (4.2.1) in relation, we see that the model (4.2.13) naturally offers itself as a model for the LIBOR process (4.2.6) in the considered setup. Since (4.2.13) satisfies (4.2.1) by construction, so does the LIBOR model

$$L(t; T_{i-1}, T_i) = \frac{L(0; T_{i-1}, T_i) + b_2(T_{i-1}, T_i)A_t^{(2)} + b_3(T_{i-1}, T_i)A_t^{(3)}}{P_{0t} + b_1(t)A_t^{(1)}} \tag{4.2.16}$$

satisfy the martingale equation (4.2.6) and in particular (4.2.10) for $t \leq T_{i-1}$. In Macrina (2014) a method based on the use of weighted heat kernels is provided for the explicit construction of the $\mathbb{M}$-martingales $\{A_t^{(i)}\}_{i=1,2}$ and thus in turn for explicit

LIBOR processes. The method allows for the development of LIBOR processes, which, if circumstances in financial markets require it, by construction take positive values at all times.

4.2.3 Bottom-up risk-neutral approach

Since we also deal with counterparty-risk valuation adjustments, we present another scheme for the construction of the LIBOR models, which we call "bottom-up risk-neutral approach". As the name suggests, we model the multi-curve term structure by making use of the risk-neutral measure (via the auxiliary measure $\mathbb{M}$) while the connection to the $\mathbb{P}$-dynamics of prices can be reintroduced at a later stage, which is important for the calculation of risk exposures and their management. "Bottom-up" refers to the fact that the short interest rate will be modelled first, then followed by the discount bond price and LIBOR processes. Similarly, in Section 4.6.1, the default intensity processes will be modelled first, and thereafter the price processes of counterparty-risky assets will be derived thereof. We utilise the notation $\mathbb{E}[\ldots|\mathcal{F}_t] = \mathbb{E}_t[\ldots]$. In the bottom-up setting, we directly model the short risk-free rate $\{r_t\}$ in the manner of the right-hand side in (4.2.15), i.e.

$$r_t = -\frac{\dot{c_1}(t) + \dot{b_1}(t)A_t^{(1)}}{c_1(t) + b_1(t)A_t^{(1)}}, \tag{4.2.17}$$

by postulating (i) non-increasing deterministic functions $b_1(t)$ and $c_1(t)$ with $c_1(0) = 1$ (later $c_1(t)$ will be seen to coincide with P_{0t}), and (ii) an $(\{\mathcal{F}_t\}, \mathbb{M})$-martingale $\{A_t^{(1)}\}$ with $A_0^{(1)} = 0$ such that

$$h_t = c_1(t) + b_1(t)A_t^{(1)} \tag{4.2.18}$$

is a positive $(\{\mathcal{F}_t\}, \mathbb{M})$-supermartingale for all $t > 0$.

Example 4.2.1. Let $A_t^{(1)} = S_t^{(1)} - 1$, where $\{S_t^{(1)}\}$ is a positive $\mathbb{M}$-martingale with $S_0^{(1)} = 1$, for example a unit-initialised exponential Lévy martingale. Then the supermartingale (4.2.18) is positive for any given t if $0 < b_1(t) \leq c_1(t)$.

Associated with the supermartingale (4.2.18), we characterise the (risk-neutral) pricing measure $\mathbb{Q}$ by the $\mathbb{M}$-density process $\{\mu_t\}_{0\leq t\leq T}$, given by

$$\mu_t = \frac{d\mathbb{Q}}{d\mathbb{M}}\Big|_{\mathcal{F}_t} = \mathcal{E}\left(\int_0^{\cdot} \frac{b_1(t)dA_t^{(1)}}{c_1(t) + b_1(t)A_{t-}^{(1)}}\right), \tag{4.2.19}$$

which is taken to be a positive $(\{\mathcal{F}_t\}, \mathbb{M})$-martingale. Furthermore, we denote by $D_t = \exp\left(-\int_0^t r_s\, ds\right)$ the discount factor associated with the risk-neutral measure $\mathbb{Q}$.

Lemma 4.2.1. $h_t = D_t\, \mu_t$.

Proof. The Ito semimartingale formula applied to $\varphi(t, A_t^{(1)}) = \ln(c_1(t)+b_1(t)A_t^{(1)}) = \ln(h_t)$ and to $\ln(D_t\mu_t)$ gives the following relations:

$$\begin{aligned} d\ln\left(c_1(t)+b_1(t)A_t^{(1)}\right) &= -r_t dt + \frac{b_1(t)dA_t^{(1)}}{c_1(t)+b_1(t)A_{t-}^{(1)}} - \frac{b_1^2(t)d[A^{(1)},A^{(1)}]_t^c}{2(c_1(t)+b_1(t)A_{t-}^{(1)})^2} \\ &+ d\sum_{s\le t}\left(\Delta\ln\left(c_1(t)+b_1(t)A_t^{(1)}\right) - \frac{b_1(t)\Delta A_t^{(1)}}{c_1(t)+b_1(t)A_{t-}^{(1)}}\right), \end{aligned} \tag{4.2.20}$$

where (4.2.17) was used in the first line, and

$$\begin{aligned} d\ln(D_t\mu_t) &= d\ln D_t + d\ln\mu_t \\ &= -r_t dt + \frac{d\mu_t}{\mu_{t-}} - \frac{d[\mu,\mu]_t^c}{2(\mu_{t-})^2} + d\sum_{s\le t}\left(\Delta\ln(\mu_t) - \frac{\Delta\mu_t}{\mu_{t-}}\right) \\ &= -r_t dt + \frac{b_1(t)dA_t^{(1)}}{c_1(t)+b_1(t)A_{t-}^{(1)}} - \frac{b_1^2(t)d[A^{(1)},A^{(1)}]_t^c}{2(c_1(t)+b_1(t)A_{t-}^{(1)})^2} \\ &\quad + d\sum_{s\le t}\left(\Delta\ln(\mu_t) - \frac{b_1(t)\Delta A_t^{(1)}}{c_1(t)+b_1(t)A_{t-}^{(1)}}\right) \end{aligned} \tag{4.2.21}$$

where

$$\begin{aligned} \Delta\ln(\mu_t) &= \ln\left(\frac{\mu_t}{\mu_{t-}}\right) = \ln\left(1+\frac{b_1(t)\Delta A_t^{(1)}}{c_1(t)+b_1(t)A_{t-}^{(1)}}\right) = \ln\left(\frac{c_1(t)+b_1(t)A_t^{(1)}}{c_1(t)+b_1(t)A_{t-}^{(1)}}\right) \\ &= \Delta\ln\left(c_1(t)+b_1(t)A_t^{(1)}\right). \end{aligned}$$

Therefore, $d\ln(h_t) = d\ln(D_t\mu_t)$. Moreover, $h_0 = D_0\mu_0 = 1$. Hence $h_t = D_t\mu_t$. ∎

It then follows that the price process of the OIS discount bond with maturity T can be expressed, for $0 \le t \le T$, by

$$P_{tT} = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{D_T}{D_t}\right] = \frac{1}{D_t\,\mu_t}\mathbb{E}^{\mathbb{M}}\left[D_T\,\mu_T\,|\,\mathcal{F}_t\right] = \mathbb{E}_t^{\mathbb{M}}\left[\frac{h_T}{h_t}\right] = \frac{c_1(T)+b_1(T)A_t^{(1)}}{c_1(t)+b_1(t)A_t^{(1)}}. \tag{4.2.22}$$

Thus, the process $\{h_t\}$ plays the role of the pricing kernel associated with the OIS market under the measure $\mathbb{M}$. In particular, we note that $c_1(t) = P_{0t}$ for $t \in [0,T]$ and $r_t = -\left(\partial_T \ln P_{tT}\right)_{|T=t}$. A construction inspired by the above formula for the OIS bond leads to the rational model for the LIBOR prevailing over the interval $[T_{i-1}, T_i)$. The $\mathcal{F}_{T_{i-1}}$-measurable spot LIBOR rate $L(T_i; T_{i-1}, T_i)$ is modelled in terms of $\{A_t^{(1)}\}$ and, in this chapter, at most two other $\mathbb{M}$-martingales $\{A_t^{(2)}\}$ and $\{A_t^{(3)}\}$ evaluated at T_{i-1}:

$$L(T_i; T_{i-1}, T_i) = \frac{L(0; T_{i-1}, T_i) + b_2(T_{i-1}, T_i)A_{T_{i-1}}^{(2)} + b_3(T_{i-1}, T_i)A_{T_{i-1}}^{(3)}}{P_{0T_i} + b_1(T_i)A_{T_{i-1}}^{(1)}}. \tag{4.2.23}$$

The (forward) LIBOR process is then defined by an application of the risk-neutral valuation formula (which is equivalent to the pricing formula (4.2.1) under $\mathbb{P}$) as follows. For $t \leq T_{i-1}$ we let

$$L(t;T_{i-1},T_i) = \frac{1}{D_t}\mathbb{E}_t^{\mathbb{Q}}\left[D_{T_i}\,L(T_i;T_{i-1},T_i)\right] = \mathbb{E}_t^{\mathbb{M}}\left[\frac{D_{T_i}\,\mu_{T_i}}{D_t\,\mu_t}\,L(T_i;T_{i-1},T_i)\right] \tag{4.2.24}$$

$$= \mathbb{E}_t^{\mathbb{M}}\left[\frac{\mathbb{E}_{T_{i-1}}^{\mathbb{M}}[h_{T_i}]L(T_i;T_{i-1},T_i)}{h_t}\right], \tag{4.2.25}$$

and thus, by applying (4.2.18) and (4.2.23),

$$L(t;T_{i-1},T_i) = \frac{L(0;T_{i-1},T_i) + b_2(T_{i-1},T_i)A_t^{(2)} + b_3(T_{i-1},T_i)A_t^{(3)}}{P_{0t} + b_1(t)A_t^{(1)}}. \tag{4.2.26}$$

Hence, we recover the same model and expression as in (4.2.16). The LIBOR models (4.2.26) (or (4.2.16)) are compatible with an HJM multi-curve setup where, in the spirit of Heath, Jarrow and Morton (1992), the initial term structures P_{0T_i} and $L(0;T_{i-1},T_i)$ are fitted by construction.

Example 4.2.2. Let $A_t^{(i)} = S_t^{(i)} - 1$, where $S_t^{(i)}$ is a positive $\mathbb{M}$-martingale with $S_0^{(i)} = 1$. For example, one could consider a unit-initialised exponential Lévy martingale defined in terms of a function of an $\mathbb{M}$-Lévy process $\{X_t^{(i)}\}$, for $i = 2, 3$. Such a construction produces non-negative LIBOR rates if

$$0 \leq b_2(T_{i-1},T_i) + b_3(T_{i-1},T_i) \leq L(0;T_{i-1},T_i). \tag{4.2.27}$$

If this condition is not satisfied, then the LIBOR model may be viewed as a shifted model, in which the LIBOR rates may become negative with positive probability. For different kinds of shifts used in the multi-curve term structure literature we refer to, e.g., Mercurio (2010a) or Moreni and Pallavicini (2014).

4.3 Clean valuation

The next questions we address are centred around the pricing of LIBOR derivatives and their calibration to market data, especially LIBOR swaptions, which are the most liquidly traded (nonlinear) interest rate derivatives. Since market data typically reflect prices of fully collaterallised transactions, which are funded at a remuneration rate of the collateral that is best proxied by the OIS rate, we consider in this section, from the perspective of model calibration, clean valuation ignoring counterparty-risk and assume funding at the rate r_t.

An interest rate swap (see, e.g., Brigo and Mercurio (2006)) is an agreement between two counterparties, where one stream of future interest payments is exchanged for another based on a specified nominal amount N. A popular interest rate swap is the exchange of a fixed rate (contractual swap spread) against the LIBOR at the end of successive time intervals $[T_{i-1}, T_i]$ of length δ. Such a swap can also be viewed

as a collection of n forward rate agreements. The swap price Sw_t at time $t \leq T_0$ is given by the following model-independent formula:

$$Sw_t = N\delta \sum_{i=1}^{n} [L(t; T_{i-1}, T_i) - KP_{tT_i}].$$

A swaption is an option between two parties to enter a swap at the expiry date T_k (the maturity date of the option). Its price at time $t \leq T_k$ is given by the following $\mathbb{M}$-pricing formula:

$$\begin{aligned} Swn_{tT_k} &= \frac{N\delta}{h_t}\mathbb{E}^{\mathbb{M}}[h_{T_k}(Sw_{T_k})^+|\mathcal{F}_t] \\ &= \frac{N\delta}{h_t}\mathbb{E}^{\mathbb{M}}\left[h_{T_k}\left(\sum_{i=k+1}^{n}[L(T_k; T_{i-1}, T_i) - KP_{T_kT_i}]\right)^+ \Big|\mathcal{F}_t\right] \\ &= \frac{N\delta}{P_{0t} + b_1 A_t^{(1)}}\mathbb{E}^{\mathbb{M}}\Bigg[\Big(\sum_{i=k+1}^{m}\big[L(0; T_{i-1}, T_i) + b_2(T_{i-1}, T_i)A_{T_k}^{(2)} + b_3(T_{i-1}, T_i)A_{T_k}^{(3)} \\ &\qquad - K(P_{0T_i} + b_1(T_i)A_{T_k}^{(1)})\big]\Big)^+ \Big|\mathcal{F}_t\Bigg], \end{aligned} \tag{4.3.28}$$

using the formulae (4.2.22) and (4.2.26) for $P_{T_kT_i}$ and $L(T_k; T_{i-1}, T_i)$. In particular, the swaption prices at time $t = 0$ can be rewritten by use of $A_t^{(i)} = S_t^{(i)} - 1$ so that

$$\begin{aligned} Swn_{0T_k} &= N\delta\,\mathbb{E}^{\mathbb{M}}\left[\left(c_2 A_{T_k}^{(2)} + c_3 A_{T_k}^{(3)} - c_1 A_{T_k}^{(1)} + c_0\right)^+\right] \\ &= N\delta\,\mathbb{E}^{\mathbb{M}}\left[\left(c_2 S_{T_k}^{(2)} + c_3 S_{T_k}^{(3)} - c_1 S_{T_k}^{(1)} + \tilde{c}_0\right)^+\right], \end{aligned} \tag{4.3.29}$$

where

$$\begin{aligned} c_2 &= \sum_{i=k+1}^{m} b_2(T_{i-1}, T_i), \quad c_3 = \sum_{i=k+1}^{m} b_3(T_{i-1}, T_i), \quad c_1 = K\sum_{i=k+1}^{m} b_1(T_i), \\ c_0 &= \sum_{i=k+1}^{m}[L(0; T_{i-1}, T_i) - KP_{0T_i}], \quad \tilde{c}_0 = c_0 + c_1 - c_2 - c_3. \end{aligned}$$

As we will see in several instance of interest, these expectations can be computed efficiently with high accuracy by various numerical schemes.

Remark 4.3.2. The advantages of modelling the LIBOR process $\{L(t; T_{i-1}, T_i)\}$ by a rational function of which denominator is the discount factor (pricing kernel) associated with the employed pricing measure (in this case $\mathbb{M}$) are: (i) The rational form of $\{L(t; T_{i-1}, T_i)\}$ and also of $\{P_{tT_i}\}$ produces, when multiplied with the discount factor $\{h_t\}$, a linear expression in the $\mathbb{M}$-martingale drivers $\{A_t^{(i)}\}$. This is in contrast to other akin pricing formulae in which the factors appear as sums of exponentials, see e.g. Crépey et al. (2015), Equation (33). (ii) The dependence structure between the LIBOR process and the OIS discount factor $\{h_t\}$—or the pricing kernel

$\{\pi_t\}$ under the $\mathbb{P}$-measure—is clear-cut. The numerator of $\{L(t;T_{i-1},T_i)\}$ is driven only by idiosyncratic stochastic factors that influence the dynamics of the LIBOR process. We may call such drivers the "LIBOR risk factors". Dependence on the "OIS risk factors", in our model example $\{A_t^{(1)}\}$, is produced solely by the denominator of the LIBOR process. (iii) Usually, the FRA process $K_t = L(t;T_{i-1},T_i)/P_{tT_i}$ is modelled directly and more commonly applied to develop multi-curve frameworks. With such models, however, it is not guaranteed that simple pricing formulae like (4.3.28) can be derived. We think that the "codebook" (4.2.6), and (4.2.26) in the considered example, is more suitable for the development of consistent, flexible and tractable multi-curve models.

4.3.1 Univariate Fourier pricing

Since in current markets there are no liquidly-traded OIS derivatives and hence no useful data is available, a pragmatic simplification is to assume deterministic OIS rates r_t. That is to say $A_t^{(1)} = 0$, and hence $b_1(t)$ plays no role either, so that it can be assumed equal to zero. Furthermore, for a start, we assume $A_t^{(3)} = 0$ and $b_3(t) = 0$, and (4.3.29) simplifies to

$$Swn_{0T_k} = N\delta\,\mathbb{E}^{\mathbb{M}}\left[\left(c_2A_{T_k}^{(2)} + c_0\right)^+\right] = N\delta\,\mathbb{E}^{\mathbb{M}}\left[\left(c_2S_{T_k}^{(2)} + \tilde{c}_0\right)^+\right],$$

where here $\tilde{c}_0 = c_0 - c_2$. For $\tilde{c}_0 > 0$ the price is simply $Swn_{0T_k} = N\delta c_0$. For $\tilde{c}_0 < 0$, and in the case of an exponential-Lévy martingale model with

$$S_t^{(2)} = \mathrm{e}^{X_t^{(2)} - t\,\psi_2(1)},$$

where $\{X_t^{(2)}\}$ is a Lévy process with cumulant ψ_2 such that

$$\mathbb{E}\left[\mathrm{e}^{zX_t^{(2)}}\right] = \exp\left[t\psi_2(z)\right], \tag{4.3.30}$$

we have

$$Swn_{0T_k} = \frac{N\delta}{2\pi}\int_{\mathbb{R}} \frac{\tilde{c}_0^{\,1-iv-R}\; M_{T_k}^{(2)}(R+iv)}{(R+iv)(R+iv-1)}dv, \tag{4.3.31}$$

where

$$M_{T_k}^{(2)}(z) = \mathrm{e}^{T_k\psi_2(z) + z\left(\ln(c_2) - \psi_2(1)\right)}$$

and R is an arbitrary constant ensuring finiteness of $M_{T_k}^{(2)}(R+iv)$ for $v \in \mathbb{R}$. For details concerning (4.3.31), we refer to Eberlein, Glau and Papapantoleon (2010).

4.3.2 One-factor lognormal model

In the event that $\{A_t^{(1)}\} = \{A_t^{(3)}\} = 0$ and $\{A_t^{(2)}\}$ is of the form

$$A_t^{(2)} = \exp\left(a_2X_t^{(2)} - \frac{1}{2}a_2^2t\right) - 1, \tag{4.3.32}$$

where $\{X_t^{(2)}\}$ is a standard Brownian motion and a_2 is a real constant, it follows from simple calculations that the swaption price is given, for $\tilde{c}_0 = c_0 - c_2$, by

$$Swn_{0T_k} = N\delta\, \mathbb{E}^{\mathbb{M}}\left[\left(c_2 A_{T_k}^{(2)} + c_0\right)^+\right] \tag{4.3.33}$$

$$= N\delta\left(c_2\Phi\left(\frac{\frac{1}{2}a_2^2 T - \ln(\tilde{c}_0/c_2)}{a_2\sqrt{T}}\right) + \tilde{c}_0\Phi\left(\frac{-\frac{1}{2}a_2^2 T - \ln(\tilde{c}_0/c_2)}{a_2\sqrt{T}}\right)\right), \tag{4.3.34}$$

where $\Phi(x)$ is the standard normal distribution function.

4.3.3 Two-factor lognormal model

We return to the price formula (4.3.29) and consider the case where the martingales $\{A_t^{(i)}\}$ are given, for $i = 1, 2, 3$, by

$$A_t^{(i)} = \exp\left(a_i X_t^{(i)} - \frac{1}{2}a_i^2 t\right) - 1, \tag{4.3.35}$$

for real constants a_i and standard Brownian motions $\{X_t^{(1)}\} = \{X_t^{(3)}\}$ and $\{X_t^{(2)}\}$ with correlation ρ. Then it follows that

$$Swn_{0T_k} = \mathbb{E}^{\mathbb{M}}\left[\left(c_2 \mathrm{e}^{X\sqrt{T_k}a_2 - \frac{1}{2}a_2^2 T_k} + c_3 \mathrm{e}^{Y\sqrt{T_k}a_3 - \frac{1}{2}a_3^2 T_k} - c_1 \mathrm{e}^{Y\sqrt{T_k}a_1 - \frac{1}{2}a_1^2 T_k} + \tilde{c}_0\right)^+\right], \tag{4.3.36}$$

where $X \sim \mathcal{N}(0,1)$, $Y \sim \mathcal{N}(0,1)$, $(X|Y) = y \sim \mathcal{N}(\rho y, (1-\rho^2))$. Hence,

$$\begin{aligned} Swn_{0T_k} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(c_2 \mathrm{e}^{x\sqrt{T_k}a_2 - \frac{1}{2}a_2^2 T_k} - K(y))^+ f(x|y) f(y)\,dx\,dy \\ &= \int_{K(y)>0}\left(\int_{-\infty}^{\infty}(c_2 \mathrm{e}^{x\sqrt{T_k}a_2 - \frac{1}{2}a_2^2 T_k} - K(y))^+ f(x|y)\,dx\right) f(y)\,dy \\ &\qquad + \int_{K(y)<0}\left(\int_{-\infty}^{\infty}(c_2 \mathrm{e}^{x\sqrt{T_k}a_2 - \frac{1}{2}a_2^2 T_k} - K(y))^+ f(x|y)\,dx\right) f(y)\,dy, \end{aligned}$$

where

$$\begin{aligned} K(y) &= c_1(\mathrm{e}^{a_1\sqrt{T_k}y - \frac{1}{2}a_1^2 T_k} - 1) - c_3(\mathrm{e}^{a_3\sqrt{T_k}y - \frac{1}{2}a_3^2 T_k} - 1) - c_0, \\ f(y) &= \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{y^2}{2}}, \\ f(x|y) &= \frac{1}{\sqrt{2\pi(1-\rho^2)}}\mathrm{e}^{\frac{-(x-\rho y)^2}{2(1-\rho^2)}}. \end{aligned}$$

This expression can be simplified further to obtain

$$\begin{aligned} &Swn_{0T_k} \\ &= \int_{K(y)>0}\left[c_2 \mathrm{e}^{a_2\sqrt{T_k}\,\rho y + \frac{1}{2}a_2^2 T_k(1-\rho^2)}\Phi\left(\frac{\rho y + a_2\sqrt{T_k}(1-\rho^2) + \ln(c_2) - \frac{1}{2}a_2^2 T_k - K(y)}{\sqrt{1-\rho^2}}\right)\right. \\ &\qquad \left. - K(y)\Phi\left(\frac{\rho y + \ln(c_2) - \frac{1}{2}a_2^2 T_k - K(y)}{\sqrt{1-\rho^2}}\right)\right] f(y)\,dy \\ &+ \int_{K(y)<0}\left(c_2 \mathrm{e}^{a_2\sqrt{T_k}\rho(y - \frac{1}{2}a_2\sqrt{T_k}\rho} - K(y)\right) f(y)\,dy. \end{aligned}$$

The calculation of the swaption price is then reduced to calculating two one-dimensional integrals. Since the regions of integration are not explicitly known, one has to numerically solve for the roots of $K(y)$, which may have up to two roots. Nevertheless a full swaption smile can be calculated in a small fraction of a second by means of this formula.

4.4 Calibration

The counterparty-risk valuation adjustments, abbreviated by XVAs (CVA, DVA, etc.), can be viewed as long-term options on the underlying contracts. For their computation, the effects by the volatility smile and term structure matter. Furthermore, for the planned XVA computations regarding the multi-curve product (basis swap) considered in later sections, it is necessary to calibrate the proposed pricing model to financial instruments with underlying tenors of $\delta = 3m$ and $\delta = 6m$. Similar to Crépey et al. (2015), we make use of the following EUR market Bloomberg data of January 4, 2011 to calibrate our model: EONIA, three-month EURIBOR and six-month EURIBOR initial term structures on the one hand, and three-month and six-month tenor swaptions on the other. As in the HJM framework of Crépey et al. (2015), to which the reader is referred for more details in this regard, the initial term structures are fitted by construction in our setup. Regarding swaption calibration, at first, we calibrate the non-maturity/tenor-dependent parameters to the swaption smile for the 9×1 years swaption with a three-month tenor underlying. The market smile corresponds to a vector of strikes $[-200, -100, -50, -25, 0, 25, 50, 100, 200]$ bp around the underlying swap spread. Then, we make use of at-the-money swaptions on three and six-month tenor swaps all terminating at exactly ten years, but with maturities from one to nine years. This co-terminal procedure is chosen with a view towards the XVA application in Section 4.6, where a basis swap with a ten-year terminal date is considered.

In particular, in a single factor $\{A_t^{(2)}\}$ setting:

1. We calibrate the parameters of the driving martingale $\{A_t^{(2)}\}$ to the smile of the 9×1 years swaption with tenor $\delta = 3m$. This part of the calibration procedure gives us also the values of $b_2(9, 9.25)$, $b_2(9.25, 9.5)$, $b_2(9.5, 9.75)$ and $b_2(9.75, 10)$, which we assume to be equal.

2. Next, we consider the co-terminal, $\Delta \times (10 - \Delta)$, ATM swaptions with $\Delta =$ 1, 2,..., 9 years. These are available written on the three and six-month rates. We calibrate the remaining values of b_2 one maturity at a time, going backwards and starting with the 8×2 years for the three-month tenor and with the 9×1 years for the six-month tenor. This is done assuming that the parameters are piecewise constant such that $b_2(T, T+0.25) = b_2(T+0.25, T+0.5) = b_2(T+0.5, T+0.75) = b_2(T+0.75, T+1)$ for each $T = 0, 1, \ldots, 8$ and that $b_2(T, T+0.5) = b_3(T+0.5, T+1)$ hold for each $T = 0, 1, \ldots, 9$.

4.4.1 Calibration of the one-factor lognormal model

In the one-factor lognormal specification of Section 4.3.2, we calibrate the parameter a_2 and $b = b_2(9, 9.25) = b_2(9.25, 9.5) = b_2(9.5, 9.75) = b_2(9.75, 10)$ with Matlab utilising the procedure "lsqnonlin" based on the pricing formula (4.3.33) (if $\tilde{c}_0 < 0$, otherwise $Swn_{0T_k} = N\delta c_0$). This calibration yields:

$$a_2 = 0.0537, b = 0.1107.$$

Forcing positivity of the underlying LIBOR rates means, in this particular case, restricting $b \leq L(0; 9.75, 10) = 0.0328$ (cf. (4.2.27)). The constrained calibration yields:

$$a_2 = 0.1864, b = 0.0328.$$

The two resulting smiles can be found in Figure 4.1, where we can see that the unconstrained model achieves a reasonably good calibration. However, enforcing positivity is highly restrictive since the Gaussian model, in this setting, cannot produce a downward sloping smile.

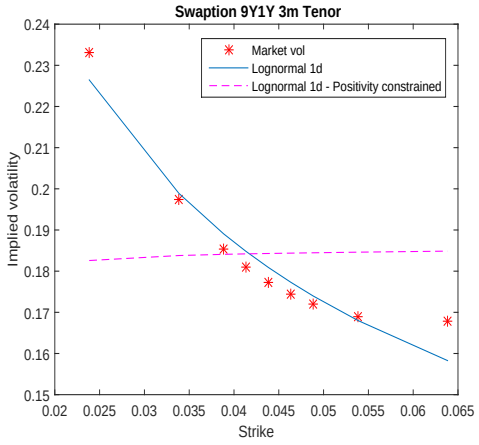

Figure 4.1: Lognormal one-factor calibration

Next we calibrate the b_2 parameters to the ATM swaption term structures of 3 months and 6 months tenors. The results are shown in Figure 4.2. When positivity is not enforced the model can be calibrated with no error to the market quotes of the ATM co-terminal swaptions. However, one can see from the figure that the positivity constraint does not allow the b_2 function to take the necessary values, and thus a very poor fit to the data is obtained, in particular for shorter maturities.

With this in mind the natural question is whether the positivity constraint is too restrictive. Informal discussions with market participants reveal that positive probability for negative rates is not such a critical issue for a model. As long as the probability mass for negative values is not substantial, it is a feature that can be lived with. Indeed assigning a small probability to this event may even be realistic. A broad panel of money market rates have been negative in the last months, including

DKK (CIBOR), short term EURIBOR and CHF LIBOR. Multiplicatively positive LIBOR model then pose practical problems for valuing options with nonpositive strike. In order to investigate the significance of the negative rates and spreads issue mentioned in Remark 4.2.1, we calculate lower quantiles for spot rates as well as the spot spread for the model calibrated without the positivity constraint. As Figure 4.3 shows, the lower quantiles for the rates are of no concern. Indeed it can hardly be considered pathological that rates will be below -14 basis points with 1% probability on a three year time horizon. Similarly, with regard to the spot spread, the lower quantile is in fact positive for all time horizons. Further calculations reveal that the probabilities of the eight year spot spread being negative is 1.1×10^{-5} and the nine year is 0.008 – which again can hardly be deemed pathologically high.

We find that the model performs surprisingly well despite the parsimony of a one-factor lognormal setup. While positivity of rates and spreads are not achieved, the model assigns only small probabilities to negative values. However, the ability of fitting the smile with such a parsimonious model is not satisfactory (cf. Fig. 4.1), which is our motivation for the next specification.

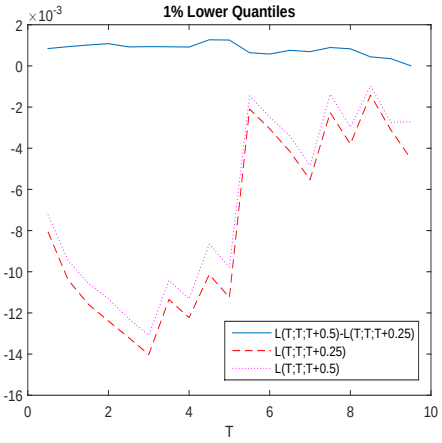

Figure 4.3: One-Factor Lognormal calibration. 1% lower quantiles

4.4.2 Calibration of exponential normal inverse Gaussian model

The one-factor model, which is driven by a Gaussian factor $\{A_t^{(2)}\}$, is able to capture the level of the volatility smile. Nevertheless, the model implied skew is slightly different from the market skew. To overcome this issue, we now consider a one-factor model driven by a richer family of Lévy processes. The process $\{A_t^{(2)}\}$ is now assumed to be the exponential normal inverse Gaussian (NIG) $\mathbb{M}$-martingale

$$A_t^{(2)} = \exp\left(X_t^{(2)} - t\psi(1)\right) - 1, \tag{4.4.37}$$

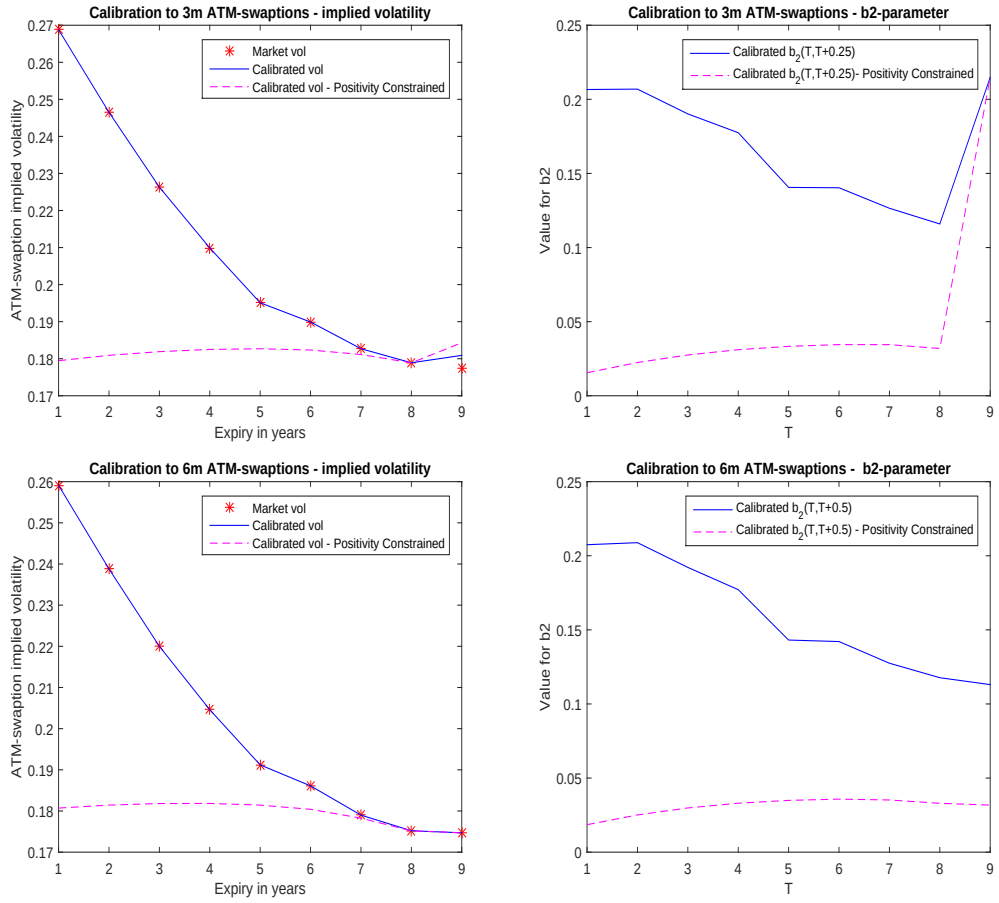

Figure 4.2: One-Factor Lognormal calibration. *(Left)* Fit to ATM swaption implied volatility term structures. *(Right)* Calibrated values of the b_2 parameters. *(Top)* $\delta = 3m$. *(Bottom)* $\delta = 6m$.

where $\{X_t^{(2)}\}$ is an $\mathbb{M}$-NIG-process with cumulant $\psi(z)$, see (4.3.30), expressed in terms of the parametrisation[1] (ν, θ, σ) from Cont and Tankov (2003) as

$$\psi(z) = -\nu \left(\sqrt{\nu^2 - 2z\theta - z^2\sigma^2} - \nu \right), \tag{4.4.38}$$

where $\nu, \sigma > 0$ and $\theta \in \mathbb{R}$. The parameters that need to be calibrated at first are ν, θ, σ and $b = b_2(9, 9.25) = b_2(9.25, 9.5) = b_2(9.5, 9.75) = b_2(9.75, 10)$. After the calibration, we obtain

$$b = 0.0431, \nu = 0.2498, \theta = -0.0242, \sigma = 0.1584.$$

Imposing $b \leq L(0; 9.75, 10) = 0.0328$ to get positive rates we obtain instead

$$b = 0.0291, \nu = 0.1354, \theta = -0.0802, \sigma = 0.3048.$$

The two fits are plotted in Figure 4.4. Here, imposing positivity comes at a much smaller cost when compared to the one factor Gaussian case. The NIG process has a richer structure (more parametric freedom) and therefore is able to compensate for an imposed smaller level of the parameter b_2.

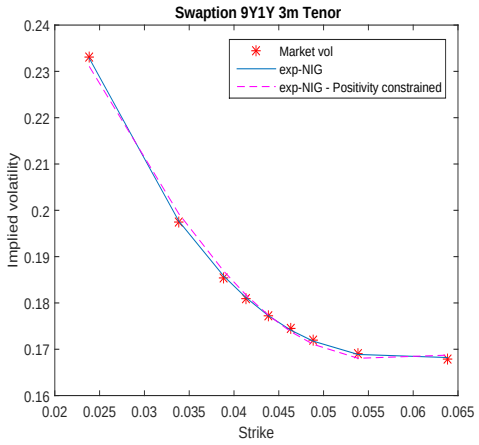

Figure 4.4: Exponential-NIG calibration

We continue with the second part of the calibration of which results are found in Figure 4.5. Here we see that enforcing positivity may have a small effect on the smile but it means that the volatility structure cannot be made to match swaptions with maturity smaller than 7 years. Thus, enforcing positivity in this model produces limitations which we wish to avoid. In Figure 4.6, we plot lower quantiles for the rates and spreads as for the one-factor lognormal model. While spot spreads remain positive, the levels do not, and, as shown, the model assigns an unrealistically high probability mass to negative values. In fact the model assigns a 1% probability

[1] The Barndorff-Nielsen (1997) parametrisation is recovered by setting $\mu = 0$, $\alpha = \frac{1}{\sigma}\sqrt{\frac{\theta_i^2}{\sigma_i^2} + \nu_i^2}$, $\beta = \frac{\theta_i}{\sigma_i^2}$ and $\delta = \sigma\nu$.

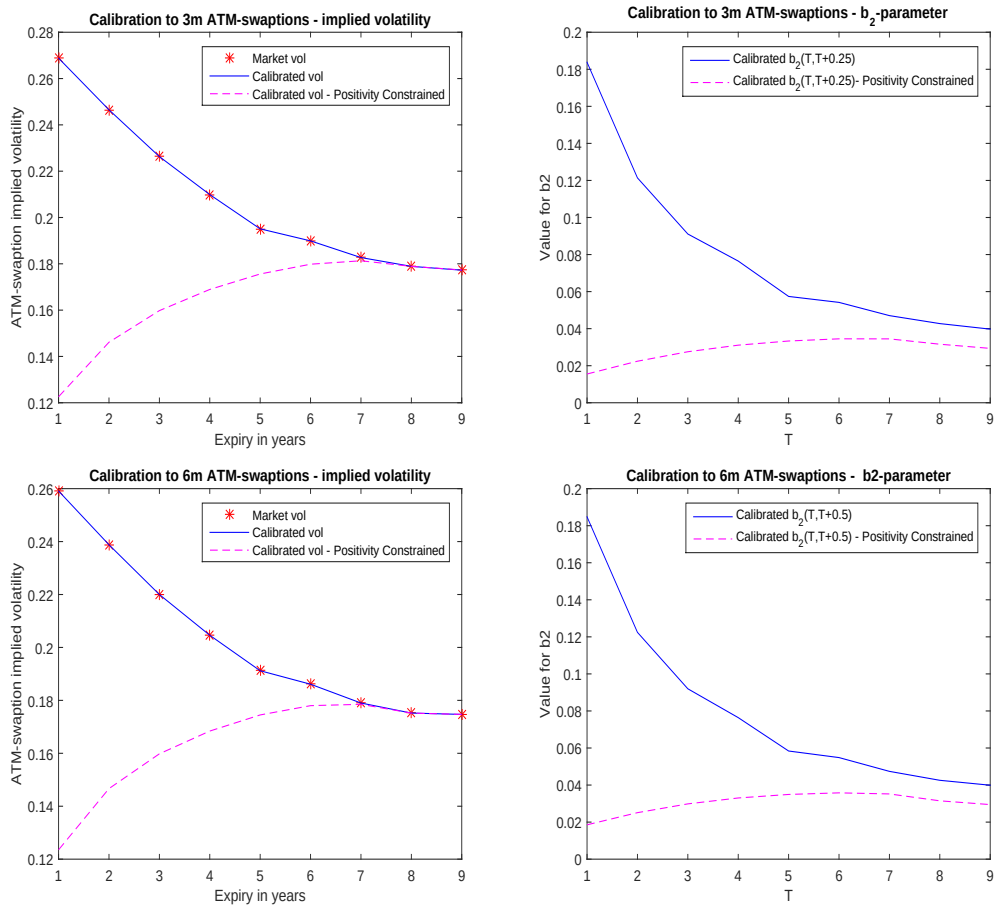

Figure 4.5: Exponential-NIG calibration. *(Left)* Fit to ATM swaption implied volatility term structures. *(Right)* Calibrated values of the b_2 parameters. *(Top)* $\delta = 3m$. *(Bottom)* $\delta = 6m$.

to rates falling below -12% within 2 years! Thus, the one-factor exponential-NIG model loses much of its appeal for it cannot, in a realistic manner regarding signs of interest rates, be made to fit long-term smiles and shorter-term ATM volatilities.

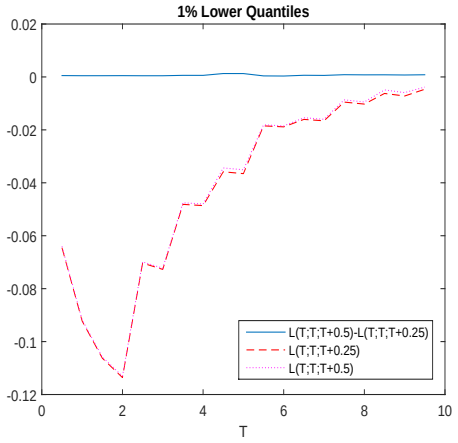

Figure 4.6: Exponential-NIG calibration calibration. 1% lower quantiles.

4.4.3 Calibration of a two-factor lognormal model

The necessity to produce a better fit to the smile than what can be achieved with the one-factor Gaussian model, while maintaining realistically positive rates and spreads, leads us to proposing the two-factor specification presented in Section 4.3.3. This model is heavily parametrised and the parameters at hand are not all identified by the considered data. We therefore fix the following parameters:

$$a_1 = 1, \quad a_3 = 1.6, \tag{4.4.39}$$

$$b_3(T, T+0.25) = 0.15L(0;T;T+0.25), \quad T \in [9, 9.75], \tag{4.4.40}$$

$$b_2(T, T+0.25) = 0.55L(0,T;T+0.25), \quad T \in [0, 8.75]. \tag{4.4.41}$$

We assume that b_1 is constant, i.e. $b_1 = b_1(T)$ for $T \in [0, 10]$, and that b_3, outside of the region defined above, is piecewise constant such that $b_3(T, T+0.25) = b_3(T+0.25, T+0.5) = b_3(T+0.5, T+0.75) = b_3(T+0.75, T+1)$ for each $T = 0, 1 \ldots, 8$ and $b_3(T, T+0.5) = b_3(T+0.5, T+1)$ holds for each $T = 0, 1 \ldots, 9$. We furthermore assume that $b_2(T, T+0.5) = b_2(T, T+0.25), \quad T \in [0, 9.5]$. These somewhat *ad hoc* choices are made with a view towards b_2 and b_3 being fairly smooth functions of time. We herewith apply a slightly altered procedure to calibrate the remaining parameters if compared to the scheme utilised for the one-factor models.

1. We first calibrate to the smile of the 9×1 years swaption which gives us the parameters a_2, ρ, the assumed constant value of b_1, and $b_2(9, 9.25)$ to $b_2(9.75, 10)$ which are assumed equal to a constant b. Similar to the exponential-NIG model, we make use of four parameters in total to fit the smile.

2. The remaining b_2 parameters are determined *a priori*, so what remains is to calibrate the values of b_3. The three-month tenor values $b_3(T, T + 0.25)$ for $T \in [0, 8.75]$ are calibrated to ATM, co-terminal swaptions starting from the 8×2 years and then continuing backwards to the 1×9 years instruments. For the six-month tenor products, we calibrate $b_3(T, T + 0.5)$ for $T \in [0, 9.5]$ starting with 9×1 years and proceed backwards.

These are the values we obtain from the first calibration phase: $b_1 = 0.2434, \quad b = 0.02, \quad a_2 = 0.1888, \quad \rho = 0.9530$. The corresponding fit is plotted in the upper left quadrant of Figure 4.7. In order to check the robustness of the calibrated fit through

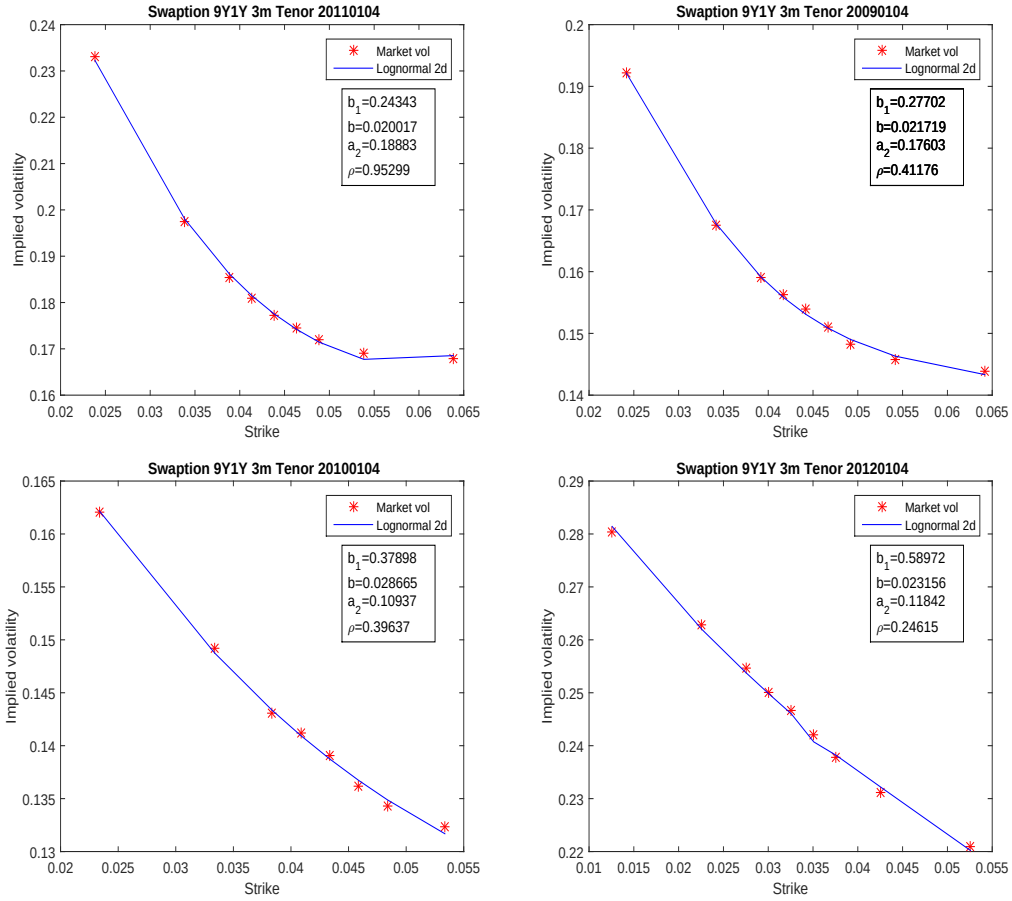

Figure 4.7: Lognormal two-factor calibration.

time, we also calibrate to three alternative dates. The quality of the fit appears quite satisfactory and comparable to the exponential-NIG model. For all four dates the calibration is done enforcing the positivity condition $b_2(T, T + 0.25) + b_3(T, T + 0.25) \leq L(0; T, T + 0.25)$. However, the procedure yields the exact same parameters even if the constraint is relaxed. We thus conclude that a better calibration appears not to be possible for these datasets by allowing negative rates. Note that it is only for our first data set that the calibrated correlation ρ is as high as 0.9530. In the other three cases we have $\rho = 0.4118$, $\rho = 0.3964$, and $\rho = 0.2461$. Figure 4.8 shows

the parameters b_2 and b_3 obtained at the second phase of the calibration to the data of 4 January 2011. As with the previous model (cf. the left graphs of Figures 4.2 and 4.5), the volatilities are matched to market data without any error. We add

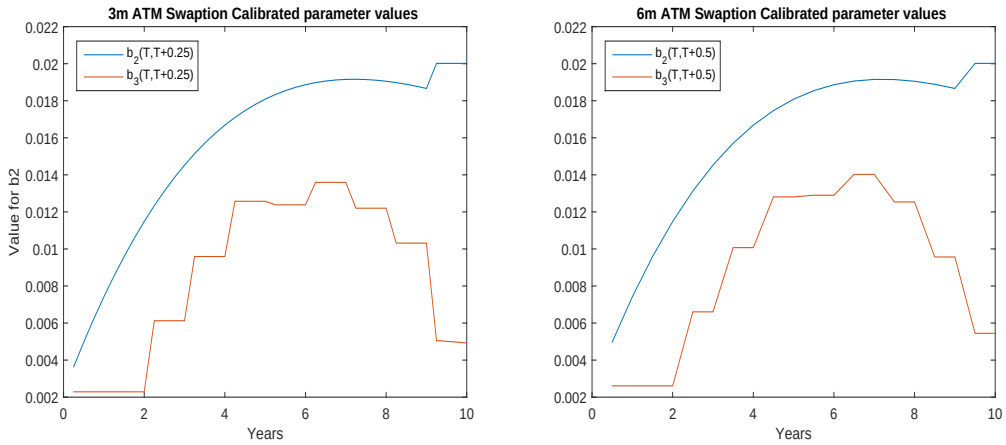

Figure 4.8: Two-factor lognormal calibration. *(Left)* Parameter values fitted to three-month ATM swaption implied volatility term structures. *(Right)* Parameter values fitted to six-month ATM swaption implied volatility term structures.

here that, although not visible from the graphs, the calibrated parameters satisfy the LIBOR spread positivity discussed in Remark 4.2.1.

In conclusion, we find that the two-factor log-normal has the ability to fit the swaption smile very well, it can be controlled to generate positive rates and positive spreads, and it is tractable with numerically-efficient closed-form expressions for the swaption prices. Given these desirable properties, we discard the one-factor models and retain the two-factor log-normal model for all the analyses in the remaining part of the chapter.

4.5 Basis swap

In this section, we prepare the ground for counterparty-risk analysis, which we shall treat in detail in Section 4.6. A typical multi-curve financial product, i.e. one that significantly manifests the difference between single-curve and a multi-curve discounting, is the so-called basis swap. Such an instrument consists of exchanging two streams of floating payments based on a nominal cash amount N or, more generally, a floating leg against another floating leg plus a fixed leg. In the classical single-curve setup, the value of a basis swap (without fixed leg) is zero throughout its life. Since the onset of the financial crisis in 2007, markets quote positive basis swap spreads that have to be added to the smaller tenor leg, which is clear evidence that LIBOR is no longer accepted as an interest rate free of credit and liquidity risk. We consider a basis swap with a duration of ten years where payments based on LIBOR of six-month tenor are exchanged against payments based on LIBOR of

three-month tenor plus a fixed spread. The two payment streams start and end at the same times $T_0 = T_0^1 = T_0^2$, $T = T_{n_1}^1 = T_{n_2}^2$. The value at time t of the basis swap with spread K is given, for $t \leq T_0$, by

$$BS_t = N\left(\sum_{i=1}^{n_1} \delta_i^{6m} L(t; T_{i-1}^1, T_i^1) - \sum_{j=1}^{n_2} \delta_j^{3m} (L(t; T_{j-1}^2, T_j^2) + KP_{tT_j^2}\right).$$

After the swap has begun, i.e. for $T_0 \leq t < T$, the value is given by

$$\begin{aligned} BS_t &= N\Bigg(\delta_{i_t}^{6m} L(T_{i_t-1}^1; T_{i_t-1}^1, T_{i_t}^1) + \sum_{i=i_t+1}^{n_1} \delta_i^{6m} L(t; T_{i-1}^1, T_i^1) \\ &\quad -\delta_{j_t}^{3m} \left(L(T_{j_t-1}^2; T_{j_t-1}^2, T_{j_t}^2) + KP_{tT_{j_t}^2}\right) - \sum_{j=j_t+1}^{n_2} \delta_j^{3m} \left(L(t; T_{j-1}^2, T_j^2) + KP_{tT_j^2}\right)\Bigg), \end{aligned}$$

where $T_{i_t}^1$ (respectively $T_{j_t}^2$) denotes the smallest T_i^1 (respectively T_i^2) that is strictly greater than t. The spread K is chosen to be the fair basis swap spread at T_0 so that the basis swap has value zero at inception. We have

$$K = \frac{\sum_{i=1}^{n_1} \delta_i^{6m} L(T_0; T_{i-1}^1, T_i^1) - \sum_{j=1}^{n_2} \delta_j^{3m} L(T_0; T_{j-1}^2, T_j^2)}{\sum_{j=1}^{n_2} \delta_j^{3m} P_{T_0 T_j^2}}.$$

The price processes on which the numerical illustration in Figure 4.9 have been obtained was simulated by applying the calibrated two-factor lognormal model developed in Section 4.4.3. The basis swap is assumed to have a notional cash amount $N = 100$ and maturity $T = 10$ years. In the two-factor lognormal setup, the basis swap spread at time $t = 0$ is $K = 12$ basis points, which is added to the three-month leg so that the basis swap is incepted at par. The $t = 0$ value of both legs is then equal to EUR 27.96. The resulting risk exposure, in the sense of the expectation and quantiles of the corresponding price process at each point in time, is shown in the left graphs of Figure 4.9, where the right plots correspond to the $\mathbb{P}$ exposure discussed in Section 4.5.1. Due to the discrete coupon payments, there are two distinct patterns of the price process exposure, most clearly visible at times preceding payments of the six-month tenor coupons for the first one and at times preceding payments of the three-month tenor coupons without the payments of the six-month tenor coupons for the second one. We show the exposures at such respective dates on the upper and lower plots in Figure 4.9.

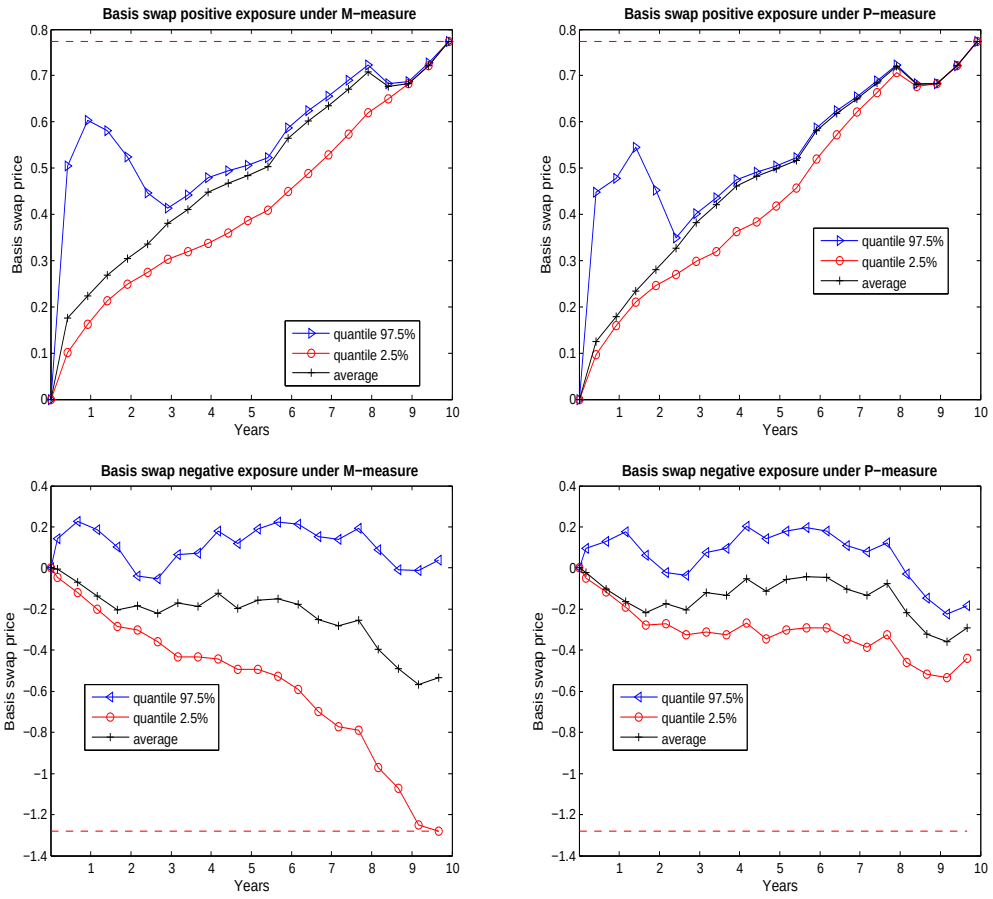

Figure 4.9: Exposures of a basis swap (price process with mean and quantiles) in the calibrated two-factor Gaussian model. *(Top)* Exposure of the basis swap at $t = 5m, 11m, etc.$ *(Bottom)* Exposure of the basis swap price at $t = 2m, 8m, 14m, etc.$ *(Left)* Exposure under the $\mathbb{M}$-measure. *(Right)* Exposure under the $\mathbb{P}$-measure with the prediction that the LIBOR rate $L(10.75y; 10.75y, 11y)$ will be either $2\,\%$ with probability $p = 0.7$ or $5\,\%$ with probability $1 - p = 0.3$.

4.5.1 Lévy random bridges

The basis swap exposures in Figure 4.9 are computed under the auxiliary $\mathbb{M}$-measure. The XVAs that are computed in later sections are derived from these $\mathbb{M}$-exposures. However, exposures are also needed for risk management and as such need to be evaluated under the real-world measure $\mathbb{P}$. This means that a measure change from $\mathbb{M}$ to $\mathbb{P}$ needs to be defined, which requires some thoughts as to what features of a price dynamics under $\mathbb{P}$ one might like to capture through a specific type of measure change and hence by the induced $\mathbb{P}$-model. In other words, we design a measure change so as to induce a particular stochastic behaviour of the $\{A_t\}$ processes under $\mathbb{P}$, and in particular of the underlying Markov processes $\{X_t\}$ driving them.

A special case we consider in what follows is where $\{X_t\}$ is a Lévy process under $\mathbb{M}$, while it adopts the law of a corresponding (possibly multivariate, componentwise)

Lévy random bridge (LRB) under $\mathbb{P}$. Several explicit asset price models driven by LRBs have been developed in Macrina (2014). The LRB-driven rational pricing models have a finite time horizon. The LRB is characterised, apart from the type of underlying Lévy process, by the terminal $\mathbb{P}$-marginal distribution to which it is pinned at a fixed time horizon U. The terminal distribution can be arbitrarily chosen, but its specification influences the behaviour of the LRB as time approaches U. In turn, the properties of a specified LRB influence the behaviour of $\{A_t\}$ and hence the dynamics of the considered price process. We see an advantage in having the freedom of specifying the $\mathbb{P}$-distribution of the factor process at some fixed future date. This way, we can implement experts' opinions (e.g. personal beliefs based on some expert analysis) in the $\mathbb{P}$-dynamics of the price process as to what level, say, an interest rate (e.g. OIS, LIBOR) is likely to be centred around at a fixed future date.

The recipe for the construction of an LRB can be found in Hoyle, Hughston and Macrina (2011), Definition 3.1, which is extended for the development of a multivariate LRB in Macrina (2014). LRBs have the property, as shown in Proposition 3.7 of Hoyle, Hughston and Macrina (2011), that there exists a measure change to an auxiliary measure with respect to which the LRB has the law of the constituting Lévy process. That is, we suppose the auxiliary measure is $\mathbb{M}$ and we have an LRB $\{X_t\}_{0\leq t\leq U}$ defined on the finite time interval $[0, U]$ where U is fixed. Under $\mathbb{M}$ and on $[0, U)$, $\{X_t\}$ has the law of the underlying Lévy process. To illustrate further, let us assume a univariate LRB; the analogous measure change for multivariate LRBs is given in Macrina (2014). Under $\mathbb{P}$, which stands in relation with $\mathbb{M}$ via the measure change

$$\eta_t = \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{M}}\Big|_{\mathcal{F}_t} = \int_{\mathbb{R}} \frac{f_{U-t}(z - X_t)}{f_U(z)}\, \nu(\mathrm{d}z),\ t < U, \tag{4.5.42}$$

where $f_t(x)$ is the density function of the underlying Lévy process for all $t \in (0, U]$ and ν is the $\mathbb{P}$-marginal law of the LRB at the terminal date U, the process $\{X_t\}$ is an LRB (note that the change of measure is singular at U).

Now, returning to the calibrated two-factor lognormal model of Section 4.4.3, but similarly also to the other models in Section 4.4, we may model the drivers $\{X_t^{(1)}\} = \{X_t^{(3)}\}$ and $\{X_t^{(2)}\}$ by two dependent Brownian random bridges under $\mathbb{P}$. The computed $\mathbb{M}$ exposures in Figure 4.9 thus need to be re-weighted by the corresponding amount η_t in order to obtain the $\mathbb{P}$-exposures of the basis swap. Since here we employ LRBs, we have the opportunity to include an expert opinion through the LRB marginals ν as to what level one believes the interest rates will tend to by time U. The re-weighted $\mathbb{P}$-exposures of the basis swap are plotted in the graphs of the right-hand side of Figure 4.9. The maximum of the upper quantile curves shown in the graphs is known as the potential future exposure (PFE) at the level 97.5%[2].

Hence, we now have the means to propose a risk-neutral model that can be calibrated to option data, and which after an explicit measure change can be applied

[2]In fact, people rather consider the expected positive exposure (expectation of the positive part of the price rather than the price) in the PFE computation, but the methodology is the same.

for risk management purposes while offering a way to incorporate economic views in the dynamic of asset prices. Recalling (4.2.19) and (4.5.42), the $\mathbb{Q}$-to-$\mathbb{P}$ measure change is obtained by

$$\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}\mathbb{Q}}\Big|_{\mathcal{F}_t} = \frac{\eta_t}{\mu_t}, \tag{4.5.43}$$

and the pricing formula for financial assets (4.2.1) may be utilised under the various measures as follows:

$$\begin{aligned} S_{tT} &= \frac{1}{D_t}\mathbb{E}^{\mathbb{Q}}[D_T\, S_{TT}\,|\,\mathcal{F}_t] = \frac{1}{D_t\,\mu_t}\mathbb{E}^{\mathbb{M}}[D_T\,\mu_T\, S_{TT}\,|\,\mathcal{F}_t] = \frac{1}{h_t}\mathbb{E}^{\mathbb{M}}[h_T\, S_{TT}\,|\,\mathcal{F}_t] \\ &= \frac{\eta_t}{D_t\,\mu_t}\mathbb{E}^{\mathbb{P}}\left[\frac{D_T\,\mu_T}{\eta_T}\, S_{TT}\,\Big|\,\mathcal{F}_t\right] = \frac{1}{\pi_t}\mathbb{E}^{\mathbb{P}}[\pi_T\, S_{TT}\,|\,\mathcal{F}_t], \end{aligned} \tag{4.5.44}$$

for $0 \le t \le T < U$ (since we consider price models driven by LRBs). It follows that the pricing kernel is given by $\pi_t = D_t\,\mu_t\,\eta_t^{-1} = \eta_t^{-1}\,h_t$. Measure changes from a risk-neutral to the real-world probability measure are discussed for similar applications also elsewhere. For a recent study in this area of research, we refer to, e.g., Hull, Sokol and White (2014).

4.6 Adjustments

So far we have focused on so-called "clean computations", i.e. ignoring counterparty-risk and assuming that funding is obtained at the risk-free OIS rate. In reality, contractually specified counterparties at the end of a financial agreement may default, and funding to enter or honour a financial agreement may come at a higher cost than at OIS rate. Thus, various valuation adjustments need to be included in the pricing of a financial position. The price of a counterparty-risky financial contract is computed as the difference between the clean price, as in previous sections, and an adjustment accounting for counterparty-risk and funding costs.

4.6.1 Rational credit model

As we shall see below, in addition to their use for the computation of PFE, the exposures in Section 4.5 can be used to compute various adjustments: CVA (credit valuation adjustment), DVA (debt valuation adjustment) and LVA (liquidity-funding valuation adjustment). With this goal in mind, we equip the bottom-up construction in Section 4.2.3, the notation of which is used henceforth, with a credit component in the following manner.

We consider $\{X_t^{(i)}\}_{t\ge 0}^{i=1,2,\ldots,n}$, which are assumed to be $(\{\mathcal{F}_t\},\mathbb{M})$-Markov processes. For any multi-index $(i_1,\ldots,i_d)$, we write $\mathcal{F}_t^{(i_1,\ldots,i_d)} = \bigvee_{l=1,\ldots,d}\mathcal{F}_t^{X^{(i_l)}}$. The (market) filtration $\{\mathcal{F}_t\}$ is given by $\{\mathcal{F}_t^{(1,\ldots,n)}\}$. For the application in the present section, we fix $n = 6$. Markov processes $\{X_t^{(1)}\} = \{X^{(3)}\}$ and $\{X_t^{(2)}\}$ are utilised to drive the OIS and LIBOR models as described in Section 4.2.3, in particular the zero-initialised $(\{\mathcal{F}_t\},\mathbb{M})$-martingales $\{A_t^{(i)}\}^{i=1,2,3}$. The Markov processes $\{X_t^{(i)}\}$, $i = 4,5,6$, which are assumed to be $\mathbb{M}$-independent between them and of the Markov

processes $i = 1, 2, 3$, are applied to model $\{\mathcal{F}_t\}$-adapted processes $\{\gamma_t^{(i)}\}^{i=4,5,6}$ defined by

$$\gamma_t^{(i)} = -\frac{\dot{c}_i(t) + \dot{b}_i(t)A_t^{(i)}}{c_i(t) + b_i(t)A_t^{(i)}}, \tag{4.6.45}$$

where $b_i(t)$ and $c_i(t)$, with $c_i(0) = 1$, are non-increasing deterministic functions, and where $\{A_t^{(i)}\}^{i=4,5,6}$ are zero-initialised $(\{\mathcal{F}_t\}, \mathbb{M})$-martingales of the form $A(t, X_t^{(i)})$. Comparing with (4.2.17), we see that (4.6.45) is modelled in the same way as the OIS rate (4.2.17), non-negative in particular, as an intensity should be (see Remark 4.6.3).

In line with the "bottom-up" construction in Section (4.2.3), we now introduce a density $(\{\mathcal{F}_t\}, \mathbb{M})$-martingale $\{\mu_t\nu_t\}_{0\le t\le T}$ that induces a measure change from $\mathbb{M}$ to the risk-neutral measure $\mathbb{Q}$:

$$\frac{d\mathbb{Q}}{d\mathbb{M}}\Big|_{\mathcal{F}_t} = \mu_t\nu_t,\ 0 \le t \le T,$$

where $\{\mu_t\}$ is defined as in Section 4.2.3. Here, we furthermore define $\nu_t = \prod_{i\ge 4}\nu_t^{(i)}$ where the processes

$$\nu_t^{(i)} = \mathcal{E}\left(\int_0^{\cdot} \frac{\dot{b}_i(t)dA_t^{(i)}}{\dot{c}_i(t) + \dot{b}_i(t)A_{t-}^{(i)}}\right)$$

are assumed to be positive true $(\{\mathcal{F}_t\}, \mathbb{M})$-martingales.

Lemma 4.6.2. *Let ξ denote any non-negative $\mathcal{F}_T^{(1,2,3)}$-measurable random variable and let $\chi = \prod_{j\ge 4}\chi_i$ where, for $j = 4, 5, 6$, χ_j is $\mathcal{F}_T^{(j)}$-measurable. Then*

$$\mathbb{E}_t^{\mathbb{R}}[\xi\,\chi] = \mathbb{E}_t^{\mathbb{R}}[\xi]\prod_{j\ge 4}\mathbb{E}_t^{\mathbb{R}}[\chi_i], \tag{4.6.46}$$

for $\mathbb{R} = \mathbb{M}$ or $\mathbb{Q}$ and for $0 \le t \le T$.

Proof. Since $\mathcal{F}_T^{(4,5,6)}$ is independent of $\mathcal{F}_t^{(1,2,3)}$ and of ξ,

$$\mathbb{E}^{\mathbb{M}}\left[\xi \,\middle|\, \mathcal{F}_t^{(1,2,3)} \vee \mathcal{F}_T^{(4,5,6)}\right] = \mathbb{E}^{\mathbb{M}}\left[\xi \,\middle|\, \mathcal{F}_t^{(1,2,3)}\right].$$

Therefore,

$$\begin{aligned}
\mathbb{E}_t^{\mathbb{M}}[\xi\,\chi] &= \mathbb{E}^{\mathbb{M}}\left[\mathbb{E}^{\mathbb{M}}\left[\xi\,\chi \,\middle|\, \mathcal{F}_t^{(1,2,3)} \vee \mathcal{F}_T^{(4,5,6)}\right] \,\middle|\, \mathcal{F}_t^{(1,2,3)} \vee \mathcal{F}_t^{(4,5,6)}\right] \\
&= \mathbb{E}^{\mathbb{M}}\left[\mathbb{E}^{\mathbb{M}}\left[\xi \,\middle|\, \mathcal{F}_t^{(1,2,3)} \vee \mathcal{F}_T^{(4,5,6)}\right]\chi \,\middle|\, \mathcal{F}_t^{(1,2,3)} \vee \mathcal{F}_t^{(4,5,6)}\right] \\
&= \mathbb{E}^{\mathbb{M}}\left[\mathbb{E}^{\mathbb{M}}\left[\xi \,\middle|\, \mathcal{F}_t^{(1,2,3)}\right]\chi \,\middle|\, \mathcal{F}_t^{(1,2,3)} \vee \mathcal{F}_t^{(4,5,6)}\right] \\
&= \mathbb{E}^{\mathbb{M}}\left[\xi \,\middle|\, \mathcal{F}_t^{(1,2,3)}\right]\mathbb{E}^{\mathbb{M}}\left[\chi \,\middle|\, \mathcal{F}_t^{(1,2,3)} \vee \mathcal{F}_t^{(4,5,6)}\right] = \mathbb{E}_t^{\mathbb{M}}[\xi]\,\mathbb{E}_t^{\mathbb{M}}[\chi].
\end{aligned}$$

Next, the Girsanov formula in combination with the result for $\mathbb{M}$-conditional expectation yields:

$$\begin{aligned}
\mathbb{E}_t^{\mathbb{Q}}[\xi\chi] &= \mathbb{E}_t^{\mathbb{M}}\left[\frac{\mu_T\nu_T\xi\chi}{\mu_t\nu_t}\right] = \mathbb{E}_t^{\mathbb{M}}\left[\frac{\mu_T\xi}{\mu_t}\right]\mathbb{E}_t^{\mathbb{M}}\left[\frac{\nu_T\chi}{\nu_t}\right] \\
&= \mathbb{E}_t^{\mathbb{M}}\left[\frac{\nu_T\mu_T\xi}{\nu_t\mu_t}\right]\mathbb{E}_t^{\mathbb{M}}\left[\frac{\mu_T\nu_T\chi}{\mu_t\nu_t}\right] = \mathbb{E}_t^{\mathbb{Q}}[\xi]\,\mathbb{E}_t^{\mathbb{Q}}[\chi].
\end{aligned}$$

The result remains to be proven for the case $\xi = 1$, which is done similarly. ∎

For the XVA computations, we shall use a reduced-form counterparty-risk approach where the default times of a bank "b" (we adopt its point of view) and of its counterparty "c" are modeled in terms of three Cox times τ_i defined by

$$\tau_i = \inf\left\{ t > 0 \,\middle|\, \int_0^t \gamma_s^{(i)}\, ds \geq E_i \right\}. \tag{4.6.47}$$

Under $\mathbb{Q}$, the random variables E_i $(i = 4, 5, 6)$ are independent and exponentially distributed. Furthermore, $\tau_c = \tau_4 \wedge \tau_6$, $\tau_b = \tau_5 \wedge \tau_6$, hence $\tau = \tau_b \wedge \tau_c = \tau_4 \wedge \tau_5 \wedge \tau_6$.

We write

$$\gamma_t^c = \gamma_t^{(4)} + \gamma_t^{(6)},\ \gamma_t^b = \gamma_t^{(5)} + \gamma_t^{(6)},\ \gamma_t = \gamma_t^{(4)} + \gamma_t^{(5)} + \gamma_t^{(6)},$$

which are the so called $(\{\mathcal{F}_t\}, \mathbb{Q})$-hazard intensity processes of the $\{\mathcal{G}_t\}$ stopping times τ_c, τ_b and τ, where the full model filtration $\{\mathcal{G}_t\}$ is given as the market filtration $\{\mathcal{F}_t\}$-progressively enlarged by τ_c and τ_b (see, e.g., Bielecki, Jeanblanc, and Rutkowski (2009), Chapter 5). Writing as before $D_t = \exp(-\int_0^t r_s\, ds)$, we note that Lemma 4.2.1 still holds in the present setup. That is,

$$h = c_1 + b_1 A^{(1)} = D\,\mu,$$

an $(\{\mathcal{F}_t\}, \mathbb{M})$-supermartingale, assumed to be positive (e.g. under an exponential Lévy martingale specification for $A^{(1)}$ as of Example 4.2.2). Further, we introduce $Z_t^{(i)} = \exp(-\int_0^t \gamma_s^{(i)}\, ds)$, for $i = 4, 5, 6$, and obtain analogously that

$$k^{(i)} := c_i + b_i A^{(i)} = Z^i\, \nu^{(i)}. \tag{4.6.48}$$

With these observations at hand, the following results follow from Lemma 4.6.2. We write $k_t = \prod_{i \geq 4} k^{(i)}$ and $Z_t = \prod_{i \geq 4} Z_t^{(i)}$.

Proposition 4.6.1. *The identities* (4.2.22) *and* (4.2.26) *still hold in the present setup, that is*

$$P_{tT} = \mathbb{E}_t^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T r_s\, ds}\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{D_T}{D_t}\right] = \mathbb{E}_t^{\mathbb{M}}\left[\frac{h_T}{h_t}\right] = \frac{c_1(T) + b_1(T) A_t^{(1)}}{c_1(t) + b_1(t) A_t^{(1)}} \tag{4.6.49}$$

and, for $t \leq T_{i-1}$,

$$L(t; T_{i-1}, T_i) = \frac{L(0; T_{i-1}, T_i) + b_2(T_{i-1}, T_i) A_t^{(2)} + b_3(T_{i-1}, T_i) A_t^{(3)}}{P_{0t} + b_1(t) A_t^{(1)}}. \tag{4.6.50}$$

Likewise,

$$\mathbb{E}_t^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T \gamma_s\, ds}\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{Z_T}{Z_t}\right] = \mathbb{E}_t^{\mathbb{M}}\left[\frac{k_T}{k_t}\right] = \prod_{i=4,5,6} \frac{c_i(T) + b_i(T) A_t^{(i)}}{c_i(t) + b_i(t) A_t^{(i)}}, \tag{4.6.51}$$

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T \gamma_s\, ds}\,\gamma_T^c\right] &= -\mathbb{E}_t^{\mathbb{Q}}\left[\frac{Z_T^{(5)}}{Z_t^{(5)}}\right]\partial_T\,\mathbb{E}_t^{\mathbb{Q}}\left[\frac{Z_T^{(4)}Z_T^{(6)}}{Z_t^{(4)}Z_t^{(6)}}\right]\\ &= -\mathbb{E}_t^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T \gamma_s\, ds}\right]\sum_{i=4,6}\frac{\dot{c}_i(T)+\dot{b}_i(T)A_t^{(i)}}{c_i(T)+b_i(T)A_t^{(i)}}, \quad (4.6.52)\end{aligned}$$

$$\mathbb{E}_t^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T (r_s+\gamma_s^c)ds}\right] = \mathbb{E}_t^{\mathbb{Q}}\left[\frac{D_T Z_T^{(4)}Z_T^{(6)}}{D_t Z_t^{(4)}Z_t^{(6)}}\right] = \prod_{i=1,4,6}\frac{c_i(T)+b_i(T)A_t^{(i)}}{c_i(t)+b_i(t)A_t^{(i)}}. \quad (4.6.53)$$

Proof. Using Lemma 4.6.2, we compute

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}\left[\mathrm{e}^{-\int_t^T r_s ds}\right] &= \mathbb{E}_t^{\mathbb{Q}}\left[\frac{D_T}{D_t}\right] = \mathbb{E}_t^{\mathbb{M}}\left[\frac{h_T\nu_T}{h_t\nu_t}\right] = \mathbb{E}_t^{\mathbb{M}}\left[\frac{h_T}{h_t}\right]\mathbb{E}_t^{\mathbb{M}}\left[\frac{\nu_T}{\nu_t}\right]\\ &= \mathbb{E}_t^{\mathbb{M}}\left[\frac{h_T}{h_t}\right] = \frac{c_1(T)+b_1(T)A_t^{(1)}}{c_1(t)+b_1(t)A_t^{(1)}}, \quad (4.6.54)\end{aligned}$$

where the last equality holds by Lemma 4.2.1. This proves (4.6.49). The other identities are proven similarly. ∎

Remark 4.6.3. Equations (4.6.49) and (4.6.51) are similar in nature and appearance. As it is the case for the resulting OIS rate $\{r_t\}$ (4.2.17), the fact that (4.6.48) is designed to be a supermartingale has as a consequence that the associated intensity (4.6.45) is a non-negative process. This is readily seen by observing that $\{\nu_t^{(i)}\}$ is a martingale and thus the drift of the supermartingale (4.6.48) is given by the necessarily non-negative process $\{\gamma_t^{(i)}\}$ that drives $\{Z_t^{(i)}\}$.

At time $t = 0$, all the $A_0^{(i)} = 0$, hence only the terms $c_i(T)$ remain in these formulas. Since the formulas (4.6.49) and (4.6.50) are not affected by the inclusion of the credit component in this approach, the valuation of the basis swap of Section 4.5 remains unchanged. By making use of the so-called "Key Lemma" of credit risk, see for instance Bielecki, Jeanblanc, and Rutkowski (2009), the identity (4.6.53) is the main building block for the pre-default price process of a "clean" CDS on the counterparty (respectively the bank, substituting τ_b for τ_c in (4.6.53)). In particular, the identities at $t = 0$

$$\mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-\int_0^T (r_s+\gamma_s^c)ds}\right] = c_1(T)c_4(T)c_6(T),\ \mathbb{E}^{\mathbb{Q}}\left[\mathrm{e}^{-\int_0^T (r_s+\gamma_s^b)ds}\right] = c_1(T)c_5(T)c_6(T), \quad (4.6.55)$$

for $T \geq 0$, can be applied to calibrate the functions $c_i(T)$, $i = 4, 5, 6$, to CDS curves of the counterparty and the bank, once the dependence on the respective credit risk factors has been specified. The calibration of the "noisy" credit model components $b_i(T)A_t^{(i)}$, $i = 4, 5, 6$, would require CDS option data or views on CDS option volatilities. If the entire model is judged underdetermined, more parsimonious specifications may be obtained by removing the common default component τ_6 (just

letting $\tau_c = \tau_4, \tau_b = \tau_5$) and/or restricting oneself to deterministic default intensities by settting some of the stochastic terms equal to zero, i.e. $b_i(T)A_t^{(i)} = 0$, $i = 4, 5$ and/or 6 (as is the case for the one-factor interest rate models in Section 4.3). The core building blocks of our multi-curve LIBOR model with counterparty-risk are the couterparty-risk kernels $\{k_t^{(i)}\}$, $i = 4, 5, 6$, the OIS kernel $\{h_t\}$, and the LIBOR kernel given by the numerator of the LIBOR process (4.2.26). We may view all kernels as defined under the $\mathbb{M}$-measure, *a priori*. The respective kernels under the $\mathbb{P}$-measure, e.g. the pricing kernel $\{\pi_t\}$, are obtained as explained at the end of Section 4.5.

4.6.2 XVA analysis

In the above reduced-form counterparty-risk setup, following Bielecki and Crépey (2014, Part III), given a contract (or portfolio of contracts) with "clean" price process $\{P_t\}$ and a time horizon T, the total valuation adjustment (TVA) process $\{\Theta_t\}$ accounting for counterparty-risk and funding cost, can be modelled as a solution to an equation of the form

$$\Theta_t = \mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T \exp\left(-\int_t^s (r_u + \gamma_u)du\right) f_s(\Theta_s)ds\right], \; t \in [0, T], \tag{4.6.56}$$

for some coefficient $\{f_t(\vartheta)\}$. We note that (4.6.56) is a backward stochastic differential equation (BSDE) for the TVA process $\{\Theta_t\}$. For accounts on BSDEs and their use in mathematical finance in general and counterparty-risk in particular, we refer to, e.g., El Karoui, Peng, and Quenez (1997), Brigo et al. (2013) and Crépey, Bielecki and Brigo (2014) or (Bielecki and Crépey 2014, Part III). An analysis in line with Crépey, Bielecki and Brigo (2014) yields a coefficient of the BSDE (4.6.56) given, for $\vartheta \in \mathbb{R}$, by:

$$\begin{aligned} f_t(\vartheta) = & \underbrace{\gamma_t^c \; (1 - R_c)(P_t - \Gamma_t)^+}_{\textit{CVA coefficient } (cva_t)} - \underbrace{\gamma_t^b (1 - R_b)(P_t - \Gamma_t)^-}_{\textit{DVA coefficient } (dva_t)} \\ & + \underbrace{\bar{b}_t \Gamma_t^+ - b_t \Gamma_t^- + \tilde{\lambda}_t \big(P_t - \vartheta - \Gamma_t\big)^+ - \lambda_t \big(P_t - \vartheta - \Gamma_t\big)^-}_{\textit{LVA coefficient } (lva_t(\vartheta))}, \end{aligned} \tag{4.6.57}$$

where:

- R_b and R_c are the recovery rates of the bank towards the counterparty and vice versa.
- $\Gamma_t = \Gamma_t^+ - \Gamma_t^-$, where $\{\Gamma_t^+\}$ (resp. $\{\Gamma_t^-\}$) denotes the value process of the collateral posted by the counterparty to the bank (resp. by the bank to the counterparty), for instance $\Gamma_t = 0$ (used henceforth unless otherwise stated) or $\Gamma_t = P_t$.
- The processes $\{\bar{b}_t\}$ and $\{b_t\}$ are the spreads with respect to the OIS short rate $\{r_t\}$ for the remuneration of the collateral $\{\Gamma_t^+\}$ and $\{\Gamma_t^-\}$ posted by the counterparty and the bank to each other.

– The process $\{\lambda_t\}$ (resp. $\{\tilde{\lambda}_t\}$) is the liquidity funding (resp. investment) spread of the bank with respect to $\{r_t\}$. By liquidity funding spreads we mean that these are free of credit risk. In particular,

$$\tilde{\lambda}_t = \bar{\lambda}_t - \gamma_t^b(1 - \bar{R}_b), \tag{4.6.58}$$

where $\{\bar{\lambda}_t\}$ is the all-inclusive funding borrowing spread of the bank and where $\bar{R}_b$ stands for a recovery rate of the bank to its unsecured lender (which is assumed risk-free, for simplicity, so that in the case of $\{\lambda_t\}$ there is no credit risk involved in any case).

The data $\{\Gamma_t\}$, $\{b_t\}$ and $\bar{b}_t$ are specified in a credit support annex (CSA) contracted between the two parties. We note that

$$\begin{aligned}\mathbb{E}_t^{\mathbb{Q}}\left[\int_t^T \exp\left(-\int_t^s (r_u+\gamma_u)du\right) f_s(\Theta_s)ds\right] &= \mathbb{E}_t^{\mathbb{M}}\left[\int_t^T \frac{\mu_s\nu_s D_s Z_s}{\mu_s\nu_t D_t Z_t} f_s(\Theta_s)ds\right] \\ &= \mathbb{E}_t^{\mathbb{M}}\left[\int_t^T \frac{h_s k_s}{h_t k_t} f_s(\Theta_s)ds\right]\end{aligned} \tag{4.6.59}$$

Hence, by setting $\widetilde{\Theta}_t = h_t\, k_t\, \Theta_t$, one obtains the following equivalent formulation of (4.6.56) and (4.6.57) under $\mathbb{M}$:

$$\widetilde{\Theta}_t = \mathbb{E}_t^{\mathbb{M}}\left[\int_t^T \tilde{f}_s(\widetilde{\Theta}_s)ds\right],\ t \in [0,T], \tag{4.6.60}$$

where

$$\begin{aligned}\frac{\tilde{f}_t(\tilde{\vartheta})}{h_t k_t} = f_t\big(\frac{\tilde{\vartheta}}{h_t k_t}\big) = \gamma_t^c\ (1-R_c)(P_t-\Gamma_t)^+ - \gamma_t^b(1-R_b)(P_t-\Gamma_t)^- \\ +\ \bar{b}_t\Gamma_t^+ - b_t\Gamma_t^- + \tilde{\lambda}_t\left(P_t - \frac{\tilde{\vartheta}}{h_t k_t} - \Gamma_t\right)^+ - \lambda_t\left(P_t - \frac{\tilde{\vartheta}}{h_t k_t} - \Gamma_t\right)^-.\end{aligned} \tag{4.6.61}$$

For the numerical implementations presented in the following section, unless stated otherwise, we set:

$$\begin{aligned}&\gamma^b = 5\%,\ \gamma^c = 7\%,\ \gamma = 10\%, \\ &R_b = R_c = 40\%, \\ &b = \bar{b} = \lambda = \tilde{\lambda} = 1.5\%.\end{aligned} \tag{4.6.62}$$

In the simulation grid one time-step corresponds to one month and $m = 10^4$ or 10^5 scenarios are produced. We recall the comments made after (4.6.55) and note that (i) the counterparty and the bank may default jointly, which is reflected by the fact that $\gamma_t < \gamma_t^b + \gamma_t^c$, and (ii) we consider a case where default intensities are assumed deterministic, that is $b_i A^{(i)} = 0$ $(i = 4, 5, 6)$. In fact, any stochasticity of the default intensities $\{\gamma_t^{(i)}\}$ would be averaged out in all the $t = 0$ pricing formulas that are derived below (but it would appear in more general t pricing formulas or in the XVA Greeks even for $t = 0$).

BSDE-based computations

The BSDE (4.6.60)-(4.6.61) can be solved numerically by simulation/regression schemes similar to those used for the pricing of American-style options, see Crépey, Gerboud, Grbac, and Ngor (2013), and Crépey et al. (2015). Since in (4.6.62) we have $\lambda_t = \tilde{\lambda}_t$, the coefficients of the terms $(P_t - \frac{\tilde{\vartheta}}{h_t k_t} - \Gamma_t)^{\pm}$ coincide in (4.6.61). This is the case of a "linear TVA" where the coefficient $\tilde{f}_t$ depends linearly on $\tilde{\vartheta}$. The results emerging from the numerical BSDE scheme for (4.6.61) can thus be verified by a standard Monte Carlo computation. Table 4.1 displays the value of the TVA and its CVA, DVA and LVA components at time zero, where the components are obtained by substituting for ϑ, in the respective term of (4.6.61), the TVA process $\tilde{\Theta}_t$ computed by simulation/regression in the first place (see Section 5.2 in Crépey et al. (2013) for the details of this procedure). The sum of the CVA, DVA and LVA, which in theory equals the TVA, is shown in the sixth column. Therefore, columns two, six and seven yield three different estimates for $\tilde{\Theta}_0 = \Theta_0$. Table 4.2 displays the relative differences between these estimates, as well as the Monte Carlo confidence interval in a comparable scale, which is shown in the last column. The TVA repriced by the sum of its components is more accurate than the regressed TVA. This observation is consistent with the better performance of Longstaff and Schwartz (2001) when compared with Tsitsiklis and Van Roy (2001) in the case of American-style option pricing by Monte Carlo methods (see, e.g., Chapter 10 in Crépey (2013)).

m	Regr TVA	CVA	DVA	LVA	Sum	MC TVA
10^4	0.0447	0.0614	-0.0243	0.0067	0.0438	0.0438
10^5	0.0443	0.0602	-0.0234	0.0067	0.0435	0.0435

Table 4.1: TVA at time zero and its decomposition (all quoted in EUR) computed by regression for $m = 10^4$ or 10^5 against $X_t^{(1)}$ and $X_t^{(2)}$. *Column 2*: TVA Θ_0. *Columns 3 to 5*: CVA, DVA, LVA at time zero repriced individually by plugging $\tilde{\Theta}_t$ for $\tilde{\vartheta}$ in the respective term of (4.6.61). *Column 6*: Sum of the three components. *Column 7*: TVA computed by a standard Monte Carlo scheme.

m	Sum/TVA	TVA/MC	Sum/MC	CI//|MC|
10^4	-2.0114%	2.0637%	0.0108 %	9.7471%
10^5	-1.7344 %	1.7386 %	-0.0259%	2.9380%

Table 4.2: Relative errors of the TVA at time zero corresponding to the results of Table 4.1. "A/B" represents the relative difference $(A - B)/B$. "CI//|MC|", in the last column, refers to the half-size of the 95%-Monte Carlo confidence interval divided by the absolute value of the standard Monte Carlo estimate of the TVA at time zero.

In Table 4.3, in order to compare alternative CSA specifications, we repeat the

above numerical implementation in each of the following four cases, with $\bar{\lambda}_t$ set equal to the constant 4.5% everywhere and all other parameters as in (4.6.62):

$$\begin{array}{lllll} 1. & (\bar{R}_b, R_b, R_c) = (100, 40, 40)\%, & Q = P, & \Gamma = 0, \\ 2. & (\bar{R}_b, R_b, R_c) = (100, 40, 40)\%, & Q = P, & \Gamma = Q = P, \\ 3. & (\bar{R}_b, R_b, R_c) = (40, 40, 40)\%, & Q = P, & \Gamma = 0, \\ 4. & (\bar{R}_b, R_b, R_c) = (100, 100, 40)\%, & Q = P, & \Gamma = 0. \end{array} \tag{4.6.63}$$

Remembering that the $t = 0$ value of both legs of the basis swap is equal to EUR

Case	Regr TVA	CVA	DVA	LVA	Sum	Sum/TVA
1	0.0776	0.0602	-0.0234	0.0408	0.0776	-0.0464 %
2	0.0095	0.0000	0.0000	0.0092	0.0092	-3.6499%
3	0.0443	0.0602	-0.0234	0.0067	0.0435	-1.7344 %
4	0.0964	0.0602	0.0000	0.0376	0.0978	1.4472%

Table 4.3: TVA at time zero and its decomposition (all quoted in EUR) computed by regression for $m = 10^5$ against $X_t^{(1)}$ and $X_t^{(2)}$. *Column 2*: TVA Θ_0. *Columns 3 to 5*: CVA, DVA and LVA at time zero, repriced individually by plugging $\tilde{\Theta}_t$ for $\tilde{\vartheta}$ in the respective term of (4.6.60). *Column 6*: Sum of the three components. *Column 7*: Relative difference between the second and the sixth columns.

27.96, the numbers in Table 4.3 may seem quite small, but one must also bear in mind that the toy model that is used here doesn't account for any wrong-way risk effect (see Crépey and Song (2015a)). In fact, the most informative conclusion of the table is the impact of the choice of the parameters on the relative weight of the different XVA components.

Exposure-based computations

Let's restrict attention to the case of interest rate derivatives with $\{P_t\}$ adapted with respect to $\{\mathcal{F}_t^{(1,2,3)}\}$. We introduce $c(s) = \prod_{i\geq 4} c_i(s)$ and the function of time

$$EPE(s) := \mathbb{E}^{\mathbb{M}}\left[h_s P_s^+\right] = \mathbb{E}^{\mathbb{Q}}\left[D_s P_s^+\right], \text{ resp. } ENE(s) := \mathbb{E}^{\mathbb{M}}\left[h_s P_s^-\right] = \mathbb{E}^{\mathbb{Q}}\left[D_s P_s^-\right],$$

called the expected positive exposure, resp. expected negative exposure. For an interest-rate swap, the EPE and ENE correspond to the mark-to-market of swaptions with maturity s written on the swap, which can be recovered analytically if available in a suitable model specification. In general, the EPE/ENE can be retrieved numerically by simulating the exposure.

In view of (4.6.60)-(4.6.61), by the time $t = 0$ forms of (4.6.51) and (4.6.52), the noncollateralised CVA at $t = 0$ satisfies (for $R_c \neq 1$, otherwise CVA_0=0):

$$\begin{aligned} \frac{1}{(1-R_c)} CVA_0 &= \mathbb{E}^{\mathbb{M}}\left[\int_0^T h_s k_s \gamma_s^c P_s^+ ds\right] = \int_0^T \mathbb{E}^{\mathbb{M}}\left[h_s P_s^+\right] \mathbb{E}^{\mathbb{M}}\left[k_s \gamma_s^c\right] ds \\ &= \int_0^T \mathbb{E}^{\mathbb{M}}\left[h_s P_s^+\right] \mathbb{E}^{\mathbb{Q}}\left[Z_s \gamma_s^c\right] ds = -\int_0^T EPE(s) \left(\frac{\dot{c}_6(s)}{c_6(s)} + \frac{\dot{c}_4(s)}{c_4(s)}\right) c(s) ds. \end{aligned}$$

Similarly, for the DVA (for $R_b \neq 1$, otherwise $DVA_0 = 0$) we have:

$$\frac{1}{(1-R_b)} DVA_0 = -\int_0^T ENE(s) \left(\frac{\dot{c}_6(s)}{c_6(s)} + \frac{\dot{c}_4(s)}{c_4(s)} \right) c(s) ds.$$

For the basis swap of Section 4.5 and the counterparty-risk data (4.6.62), we obtain by this manner $CVA_0 = 0.0600$ and $DVA_0 = -0.0234$, quite consistent with the corresponding entries of the second row (i.e. for $m = 10^5$) in Table 4.1. As for the LVA, to simplify its computation, one may be tempted to neglect the nonlinearity that is inherent to $lva_t(\vartheta)$ (unless $\tilde{\lambda}_t = \lambda_t$), replacing ϑ by 0 in $lva_t(\vartheta)$. Then, assuming $lva_t(0) \in \mathcal{X}_t^{(1,2,3)}$, by (4.6.56)-(4.6.57), one can compute a linearised LVA at time zero given by

$$\begin{aligned} \widehat{LVA}_0 &= \mathbb{E}^{\mathbb{M}} \left[\int_0^T h_s k_s \, lva_s(0) ds \right] \\ &= \int_0^T \mathbb{E}^{\mathbb{M}} \left[h_s \, lva_s(0) \right] \mathbb{E}^{\mathbb{M}} \left[k_s \right] ds = \int_0^T \mathbb{E}^{\mathbb{M}} \left[h_s \, lva_s(0) \right] c(s) ds, \end{aligned}$$

by (4.6.51) for $t = 0$. This is based on the expected (linearised) liquidity exposure

$$\mathbb{E}^{\mathbb{M}} \left[h_s lva_s(0) \right] = \mathbb{E}^{\mathbb{Q}} \left[D_s lva_s(0) \right].$$

In case of no collateralisation ($\Gamma_t = 0$) and of deterministic $\tilde{\lambda}_t$ and λ_t, we have

$$lva_s(0) = \tilde{\lambda}_s P_s^+ - \lambda_s P_s^-, \; \widehat{LVA}_0 = \int_0^T \Big(\tilde{\lambda}_s EPE(s) - \lambda_s ENE(s) \Big) c(s) ds.$$

In case of continuous collateralisation ($\Gamma_t = P_t$) and of deterministic $\bar{b}_t$ and b_t, the formulas read

$$lva_s(0) = \bar{b}_s P_s^+ - b_s P_s^-, \; \widehat{LVA}_0 = \int_0^T \big(\bar{b}_s EPE(s) - b_s ENE(s) \big) c(s) ds.$$

As for CVA/DVA, the LVA exposure is controlled by the EPE/ENE functions, but for different "weighting functions", depending on the CSA. For instance, for the data (4.6.62), the LVA on the basis swap of Section 4.5 (collateralised or not, since in this case $\bar{b}_t = b_t = \tilde{\lambda}_t = \lambda_t = 1.5\%$), we obtain $\widehat{LVA}_0 = 0.0098$, quite different in relative terms (but these are small numbers) from the exact (as opposed to linearised) value of 0.0067 in Table 4.1.

Chapter 5

Nonlinear Monte Carlo schemes for counterparty risk on credit derivatives

5.1 Introduction

Counterparty risk is a major issue since the global credit crisis and the ongoing European sovereign debt crisis. In a bilateral counterparty risk setup, counterparty risk is valued as the so-called credit valuation adjustment (CVA), for the risk of default of the counterparty, and debt valuation adjustment (DVA), for own default risk. In such a setup, the classical assumption of a locally risk-free funding asset used for both investing and unsecured borrowing is no longer sustainable. The proper accounting of the funding costs of a position leads to the funding valuation adjustment (FVA). Moreover, these adjustments are interdependent and must be computed jointly through a global correction dubbed total valuation adjustment (TVA). The pricing equation for the TVA is nonlinear due to the funding costs. It is posed over a random time interval determined by the first default time of the two counterparties. To deal with the corresponding backward stochastic differential equation (BSDE), a first reduced-form modeling approach has been proposed in Crépey (2012b), under a rather standard immersion hypothesis between a reference (or market) filtration and the full model filtration progressively enlarged by the default times of the counterparties. This basic immersion setup is fine for standard applications, such as counterparty risk on interest rate derivatives. But it is too restrictive for situations of strong dependence between the underlying exposure and the default risk of the two counterparties, such as counterparty risk on credit derivatives (which involves strong adverse dependence, called wrong-way risk). For this reason, an extended reduced-form modeling approach has been recently developed in Crépey and Song (2014a, 2014b, 2015a, 2015b). With credit derivatives, the problem is also very high-dimensional. From a numerical point of view, for high-dimensional nonlinear problems, only purely forward simulation schemes can

be used. In Crépey and Song (2015a), the problem is addressed by the linear Monte Carlo expansion with randomization of Fujii and Takahashi (2012a,2012b). In the present work, we assess another scheme, namely the marked branching diffusion approach of Henry-Labordère (2012), which we compare with the previous one in terms of applicability and numerical behavior. This is done in two dynamic copula models of portfolio credit risk: the dynamic Gaussian copula model and where default dependence stems from joint defaults.

The chapter is organized as follows. Sect. 5.2 and 5.3 provide a summary of the main pricing and TVA BSDEs that are derived in Crépey and Song (2014a, 2014b, 2015a). Sect. 5.4 exposes two nonlinear Monte Carlo schemes that can be considered for solving these in high-dimensional models, such as the portfolio credit models of Sect. 5.5. Comparative numerics in these models are presented in Sect. 5.6. Sect. 5.7 concludes.

5.2 Prices

5.2.1 Setup

We consider a netted portfolio of OTC derivatives between two defaultable counterparties, generally referred to as the contract between a bank, the perspective of which is taken, and its counterparty. After having bought the contract from its counterparty at time 0, the bank sets-up a hedging, collateralization (or margining) and funding portfolio. We call the funder of the bank a third party, possibly composed in practice of several entities or devices, insuring funding of the bank's strategy. The funder, assumed default-free for simplicity, plays the role of lender/borrower of last resort after the exhaustion of the internal sources of funding provided to the bank through its hedge and collateral.

For notational simplicity we assume no collateralization. All the numerical considerations, our main focus in this work, can be readily extended to the case of collateralized portfolios using the corresponding developments in Crépey and Song (2015a). Likewise, we assume hedging in the simplest sense of replication by the bank and we consider the case of a fully securely funded hedge, so that the the cost of the hedge of the bank is exactly reflected by the wealth of its hedging and funding portfolio.

We consider a stochastic basis $(\Omega, \mathcal{G}_T, \mathcal{G}, \mathbb{Q})$, where $\mathcal{G} = (\mathcal{G}_t)_{t\in[0,T]}$ is interpreted as a risk-neutral pricing model on the primary market of the instruments that are used by the bank for hedging its TVA. The reference filtration $\mathcal{F}$ is a subfiltration of $\mathcal{G}$ representing the counterparty risk free filtration, not carrying any direct information about the defaults of the two counterparties. The relation between these two filtrations will be pointed out in the condition (C) to be introduced later. We denote by:

- $\mathbb{E}_t$, the conditional expectation under $\mathbb{Q}$ given $\mathcal{G}_t$,

- r, the OIS (risk-free) short rate process, with related discount factor $\beta_t = e^{-\int_0^t r_s ds}$,
- T, the maturity of the contract,
- τ_b and τ_c, the default time of the bank and of the counterparty, modeled as $\mathcal{G}$ stopping times with $(\mathcal{G}, \mathbb{Q})$ intensities γ^b and γ^c,
- $\tau = \tau_b \wedge \tau_c$, the first-to-default time of the two counterparties, also a $\mathcal{G}$ stopping time, with intensity γ such that $\max(\gamma^b, \gamma^c) \leq \gamma \leq \gamma^b + \gamma^c$,
- $\bar{\tau} = \tau \wedge T$, the effective time horizon of our problem (there is no cashflow after $\bar{\tau}$),
- D, the contractual dividend process,
- $\Delta = D - D_-$, the jump process of D.

5.2.2 Clean price

We denote by P be the reference (or clean) price of the contract ignoring counterparty risk and assuming the position of the bank financed at the OIS rate r, i.e. the $\mathcal{G}$ conditional expectation of the future contractual cash-flows discounted at the OIS risk-free rate r. In particular,

$$\beta_t P_t = \mathbb{E}_t \left[\int_t^{\bar{\tau}} \beta_s dD_s + \beta_{\bar{\tau}} P_{\bar{\tau}} \right], \; t \in [0, \bar{\tau}]. \tag{5.2.1}$$

We also define $Q_t = P_t + \mathbf{1}_{\{t=\tau<T\}} \Delta_\tau$, so that Q_τ represents the clean value of the contract inclusive of the promised dividend at default (if any) Δ_τ, which also belongs to the "debt" of the counterparty to the bank (or vice versa depending on the sign of Q_τ) in case of default of a party. Accordingly, at time τ (if $< T$), the close-out cash-flow of the counterparty to the bank is modeled as

$$\mathcal{R} = \mathbf{1}_{\{\tau=\tau_c\}} \big(R_c Q_\tau^+ - Q_\tau^- \big) - \mathbf{1}_{\{\tau=\tau_b\}} \big(R_b Q_\tau^- - Q_\tau^+ \big) - \mathbf{1}_{\{\tau_b=\tau_c\}} Q_\tau, \tag{5.2.2}$$

where R_b and R_c are the recovery rates of the bank and of the counterparty to each other.

5.2.3 All-inclusive price

Let Π be the all-inclusive price of the contract for the bank, including the cost of counterparty risk and funding costs. Since we assume a securely funded hedge (in the sense of replication) and no collateralization, the amounts invested and funded by the bank at time t are respectively given by Π_{t-}^- and Π_{t-}^+. The all-inclusive price Π is the discounted conditional expectation of all effective future cash flows

including the contractual dividends before τ, the cost of funding the position prior to time τ and the terminal cash flow at time τ. Hence,

$$\beta_t \Pi_t = \mathbb{E}_t\Big[\int_t^{\bar{\tau}} \beta_s \mathbf{1}_{s<\tau} dD_s - \int_t^{\bar{\tau}} \beta_s \bar{\lambda}_s \Pi_s^+ ds + \beta_{\bar{\tau}} \mathbf{1}_{\tau<T} \mathcal{R}\Big], \tag{5.2.3}$$

where $\bar{\lambda}$ is the funding spread over r of the bank toward the external funder, i.e. the bank borrows cash from its funder at rate $r + \bar{\lambda}$ (and invests cash at the risk-free rate r). Since the right hand side in (5.2.3) depends also on Π, (5.2.3) is in fact a backward stochastic differential equation (BSDE). Consistent with the no arbitrage principle, the gain process on the hedge is a $\mathbb{Q}$ martingale, which explains why it does not appear in (5.2.3).

5.3 TVA BSDEs

The total valuation adjustment (TVA) process Θ is defined as

$$\Theta = Q - \Pi. \tag{5.3.4}$$

5.3.1 Full TVA BSDE

By taking the difference between (5.2.1) and (5.2.3), we obtain

$$\beta_t \Theta_t = \mathbb{E}_t\left[\int_t^{\bar{\tau}} \beta_s g_s(\Theta_s) ds + \beta_{\bar{\tau}} \mathbf{1}_{\tau<T} \xi\right], \; t \in [0, \bar{\tau}], \tag{5.3.5}$$

where $g_t(\vartheta) = \bar{\lambda}_t (P_t - \vartheta)^+$ is the funding coefficient and where

$$\xi = Q_\tau - \mathcal{R} = \mathbf{1}_{\{\tau=\tau_c\}}(1 - R_c)(P_\tau + \Delta_\tau)^+ - \mathbf{1}_{\{\tau=\tau_b\}}(1 - R_b)(P_\tau + \Delta_\tau)^- \tag{5.3.6}$$

is the exposure at default of the bank. Equivalent to (5.3.5), the "full TVA BSDE" is written as

$$\Theta_t = \mathbb{E}_t\left[\int_t^{\bar{\tau}} f_s(\Theta_s) ds + \mathbf{1}_{\tau<T} \xi\right], \; 0 \le t \le \bar{\tau}, \tag{I}$$

for the coefficient $f_t(\vartheta) = g_t(\vartheta) - r_t \vartheta$.

5.3.2 Partially reduced TVA BSDE

Let $\hat{\xi}$ be a $\mathcal{G}$ predictable process, which exists by Corollary 3.23 2) in He, Wang, and Yan (1992), such that $\hat{\xi}_\tau = \mathbb{E}[\xi | \mathcal{G}_{\tau-}]$ on $\tau < \infty$ and let $\bar{f}$ be the modified coefficient such that

$$\bar{f}_t(\vartheta) + r_t \vartheta = g_t(\vartheta) + (r_t + \gamma_t)\hat{\xi}_t. \tag{5.3.7}$$

As easily shown (cf. Crépey and Song (2014a, Lemma 2.2)), the full TVA BSDE (I) can be simplified into the "partially reduced BSDE"

$$\bar{\Theta}_t = \mathbb{E}_t\left[\int_t^{\bar{\tau}} \bar{f}_s(\bar{\Theta}_s) ds\right], \; 0 \le t \le \bar{\tau}, \tag{II}$$

in the sense that if Θ solves (I), then $\bar{\Theta} = \Theta \mathbf{1}_{[0,\tau)}$ solves (II), whilst if $\bar{\Theta}$ solves (II), then $\Theta = \bar{\Theta}\mathbf{1}_{[0,\tau)} + \mathbf{1}_{[\tau]}\mathbf{1}_{\tau<T}\xi$ solves (I). Note that both BSDEs (I) and (II) are $(\mathcal{G}, \mathbb{Q})$ BSDEs posed over the random time interval $[0, \bar{\tau}]$, but with the terminal condition ξ for (I) as opposed to a null terminal condition (and a modified coefficient) for (II).

5.3.3 Fully reduced TVA BSDE

Let

$$\hat{f}_t(\vartheta) = \bar{f}_t(\vartheta) - \gamma_t \vartheta = cdva_t + fva_t(\vartheta) - (r_t + \gamma_t)\vartheta.$$

Assume the following conditions, which are studied in Crépey and Song (2014a, 2014b, 2015a, 2015b):

Condition (C). There exist:

(C.1) a subfiltration $\mathcal{F}$ of $\mathcal{G}$ satisfying the usual conditions and such that $\mathcal{F}$ semimartingales stopped at τ are $\mathcal{G}$ semimartingales,

(C.2) a probability measure $\mathbb{P}$ equivalent to $\mathbb{Q}$ on $\mathcal{F}_T$ such that any $(\mathcal{F}, \mathbb{P})$ local martingale stopped at $(\tau-)$ is a $(\mathcal{G}, \mathbb{Q})$ local martingale on $[0, T]$,

(C.3) an $\mathcal{F}$ progressive "reduction" $\tilde{f}_t(\vartheta)$ of $\hat{f}_t(\vartheta)$ such that $\int_0^\cdot \hat{f}_t(\vartheta)dt = \int_0^\cdot \tilde{f}_t(\vartheta)dt$ on $[0, \bar{\tau}]$.

Let $\tilde{\mathbb{E}}_t$ denote the conditional expectation under $\mathbb{P}$ given $\mathcal{F}_t$. It is shown in Crépey and Song (2014a, 2014b, 2015a) that the full TVA BSDE (I) is equivalent to the following "fully reduced BSDE":

$$\tilde{\Theta}_t = \tilde{\mathbb{E}}_t \left[\int_t^T \tilde{f}_s(\tilde{\Theta}_s) ds \right], \quad t \in [0, T], \tag{III}$$

equivalent in the sense that if Θ solves (I), then the "$\mathcal{F}$ optional reduction" $\tilde{\Theta}$ of Θ ($\mathcal{F}$ optional process that coincides with Θ before τ) solves (III), whilst if $\tilde{\Theta}$ solves (III), then $\Theta = \tilde{\Theta}\mathbf{1}_{[0,\tau)} + \mathbf{1}_{[\tau]}\mathbf{1}_{\tau<T}\xi$ solves (I).

Moreover, under mild assumptions (see e.g. Crépey and Song (2015a, Theorem 4.1)), one can easily check that $\bar{f}_t(\vartheta)$ in (5.3.7) (resp. $\tilde{f}_t(\vartheta)$) satisfies the classical BSDE monotonicity assumption

$$\big(\bar{f}_t(\vartheta) - \bar{f}_t(\vartheta')\big)(\vartheta - \vartheta') \leq C(\vartheta - \vartheta')^2$$

(and likewise for $\tilde{f}$), for some constant C. Hence, by classical BSDE results nicely surveyed in Kruse and Popier (2014, Section 2 (resp. 3)), the partially reduced TVA BSDE (II), hence the equivalent full TVA BSDE (I) (resp. the fully reduced BSDE (III)), is well-posed in the space of $(\mathcal{G}, \mathbb{Q})$ (resp. $(\mathcal{F}, \mathbb{P})$) square integrable solutions, where well-posedness includes existence, uniqueness, comparison and BSDE standard estimates.

5.3.4 Marked default time setup

In order to be able to compute $\gamma\hat{\xi}$ in $\bar{f}$, we assume that τ is endowed with a mark e in a finite set E, in the sense that

$$\tau = \min_{e\in E} \tau_e, \tag{5.3.8}$$

where each τ_e is a stopping time with intensity γ_t^e such that $\mathbb{Q}(\tau_e \neq \tau_{e'}) = 1$, $e \neq e'$, and

$$\mathcal{G}_\tau = \mathcal{G}_{\tau^-} \vee \sigma(\epsilon),$$

where $\epsilon = \text{argmin}_{e\in E}\tau_e$ yields the "identity" of the mark. Then, by Lemma 5.1 in Crépey and Song (2015a), there exists $\mathcal{G}$-predictable processes $\tilde{P}_t^e$ and $\tilde{\Delta}_t^e$ such that

$$P_\tau = \tilde{P}_\tau^e \text{ and } \Delta_\tau = \tilde{\Delta}_\tau^e \text{ on the event } \{\tau = \tau_e\}.$$

Assuming further that $\tau_b = \min_{e\in E_b} \tau_e$ and $\tau_c = \min_{e\in E_c} \tau_e$, where $E = E_b \cup E_c$ (not necessarily a disjoint union), one can then take on $[0, \bar{\tau}]$:

$$\gamma_t\hat{\xi}_t = (1-R_c)\sum_{e\in E_c} \gamma_t^e \left(\tilde{P}_t^e + \tilde{\Delta}_t^e\right)^+ - (1-R_b)\sum_{e\in E_b} \gamma_t^e \left(\tilde{P}_t^e + \tilde{\Delta}_t^e\right)^-,$$

where the two terms have clear respective CVA and DVA interpretation. Hence, (5.3.7) is rewritten, on $[0, \bar{\tau}]$, as

$$\begin{aligned}\bar{f}_t(\vartheta) + r_t\vartheta = &\underbrace{(1-R_c)\sum_{e\in E_c} \gamma_t^e \left(\tilde{P}_t^e + \tilde{\Delta}_t^e\right)^+}_{\text{CVA coefficient } (cva_t)} - \underbrace{(1-R_b)\sum_{e\in E_b} \gamma_t^e \left(\tilde{P}_t^e + \tilde{\Delta}_t^e\right)^-}_{\text{DVA coefficient } (dva_t)} \\ &+ \underbrace{\bar{\lambda}_t(P_t - \vartheta)^+}_{\text{FVA coefficient } (fva_t(\vartheta)=g_t(\vartheta))}.\end{aligned} \tag{5.3.9}$$

If the functions $\tilde{P}_t^e$ and $\tilde{\Delta}_t^e$ above not only exist, but can be computed explicitly (as will be the case in the concrete models of 5.5.1 and 5.5.2), once stated in a Markov setup where

$$\bar{f}_t(\vartheta) = \bar{f}(t, X_t, \vartheta),\ t \in [0,T], \tag{5.3.10}$$

for some $(\mathcal{G}, \mathbb{Q})$ jump diffusion X, then the partially reduced TVA BSDE (II) can be tackled numerically. Similarly, once stated in a Markov setup where

$$\tilde{f}_t(\vartheta) = \tilde{f}(t, \tilde{X}_t, \vartheta),\ t \in [0,T], \tag{5.3.11}$$

for some $(\mathcal{F}, \mathbb{P})$ jump diffusion $\tilde{X}$, then the fully reduced TVA BSDE (III) can be tackled numerically.

5.4 TVA numerical schemes

5.4.1 Linear approximation

Our first TVA approximation is obtained replacing Θ_s by 0 in the right hand side of (I), i.e.

$$\Theta_0 \approx \mathbb{E}\left[\int_0^{\bar{\tau}} f_s(0)ds + \mathbf{1}_{\tau<T}\xi\right] = \mathbb{E}\left[\int_0^{\bar{\tau}} \bar{\lambda}_s P_s^+ ds + \mathbf{1}_{\tau<T}\xi\right]. \tag{5.4.12}$$

One we then approximate the TVA by standard Monte-Carlo, with randomization of the integral to reduce the computation time (at the cost of a small increase in the variance). Hence, introducing an exponential time ζ of parameter μ, i.e. a random variable with density $\phi(s) = \mathbf{1}_{s\geq 0}\,\mu\, e^{-\mu s}$, we have

$$\mathbb{E}\left[\int_0^{\bar{\tau}} f_s(0)ds\right] = \mathbb{E}\left[\int_0^{\bar{\tau}} \phi(s)\frac{1}{\mu}e^{\mu s} f_s(0)ds\right] = \mathbb{E}\left[\mathbf{1}_{\zeta<\bar{\tau}}\frac{e^{\mu\zeta}}{\mu} f_\zeta(0)\right]. \tag{5.4.13}$$

We can use the same technic for (II) and (III), which yields:

$$\Theta_0 = \bar{\Theta}_0 \approx \mathbb{E}\left[\int_0^{\bar{\tau}} \bar{f}_s(0)ds\right] = \mathbb{E}\left[\mathbf{1}_{\zeta<\bar{\tau}}\frac{e^{\mu\zeta}}{\mu}\bar{f}_\zeta(0)\right], \tag{5.4.14}$$

$$\Theta_0 = \tilde{\Theta}_0 \approx \tilde{\mathbb{E}}\left[\int_0^{T} \tilde{f}_s(0)ds\right] = \tilde{\mathbb{E}}\left[\mathbf{1}_{\zeta<T}\frac{e^{\mu\zeta}}{\mu}\tilde{f}_\zeta(0)\right]. \tag{5.4.15}$$

5.4.2 Linear Expansion and interacting particle implementation

Following Fujii and Takahashi (2012a,2012b), we can introduce a perturbation parameter ϵ and the following perturbed form of the fully reduced BSDE (III):

$$\tilde{\Theta}_t^\epsilon = \tilde{\mathbb{E}}_t\left[\int_t^T \epsilon\tilde{f}_s(\tilde{\Theta}_s^\epsilon)ds\right], \quad t\in[0,T], \tag{5.4.16}$$

where $\epsilon = 1$ corresponds to the original BSDE (III). Suppose that the solution of (5.4.16) can be expanded in a power series of ϵ:

$$\tilde{\Theta}_t^\epsilon = \tilde{\Theta}_t^{(0)} + \epsilon\tilde{\Theta}_t^{(1)} + \epsilon^2\tilde{\Theta}_t^{(2)} + \epsilon^3\tilde{\Theta}_t^{(3)} + \cdots. \tag{5.4.17}$$

The Taylor expansion of f at $\tilde{\Theta}^{(0)}$ reads

$$\tilde{f}_t(\tilde{\Theta}_t^\epsilon) = \tilde{f}_t(\tilde{\Theta}_t^{(0)}) + (\epsilon\tilde{\Theta}_t^{(1)} + \epsilon^2\tilde{\Theta}_t^{(2)} + \cdots)\partial_\vartheta\tilde{f}_t(\tilde{\Theta}_t^{(0)}) + \frac{1}{2}(\epsilon\tilde{\Theta}_t^{(1)} + \epsilon^2\tilde{\Theta}_t^{(2)} + \cdots)^2\partial^2_{\vartheta^2}\tilde{f}_t(\tilde{\Theta}_t^{(0)}) + \cdots$$

Collecting the terms of the same order with respect to ϵ in (5.4.16), we obtain $\tilde{\Theta}_t^{(0)} = 0$, due to the null terminal condition of the fully reduced BSDE (III), and

$$\begin{aligned}
\tilde{\Theta}_t^{(1)} &= \tilde{\mathbb{E}}_t\left[\int_t^T \tilde{f}_s(\tilde{\Theta}_s^{(0)})ds\right],\\
\tilde{\Theta}_t^{(2)} &= \tilde{\mathbb{E}}_t\left[\int_t^T \tilde{\Theta}_s^{(1)}\partial_\vartheta\tilde{f}_s(\tilde{\Theta}_s^{(0)})ds\right],\\
\tilde{\Theta}_t^{(3)} &= \tilde{\mathbb{E}}_t\left[\int_t^T \tilde{\Theta}_s^{(2)}\partial_\vartheta\tilde{f}_s(\tilde{\Theta}_s^{(0)})ds\right],
\end{aligned} \tag{5.4.18}$$

where the complete third order term comprises another component based on $\partial^2_{\vartheta^2}\tilde{f}$. In our case, $\partial^2_{\vartheta^2}\tilde{f}$ involves a Dirac measure via the terms $(P_t-\vartheta)^+$ in $fva_t(\vartheta)$, so that we truncate the expansion to the term $\tilde{\Theta}^{(3)}_t$ as above. If the non-linearity in (III) is sub-dominant, one can expect to obtain a reasonable approximation of the original equation by setting $\epsilon = 1$ at the end of the calculation, i.e.

$$\tilde{\Theta}_0 \approx \tilde{\Theta}^{(1)}_0 + \tilde{\Theta}^{(2)}_0 + \tilde{\Theta}^{(3)}_0.$$

Carrying out a Monte Carlo simulation by an Euler scheme for every time s in a time grid and integrating to obtain $\tilde{\Theta}^{(1)}_0$ would be quite heavy. Moreover, this would become completely unpractical for the higher order terms that involve iterated (multivariate) time integrals. For these reasons, Fujii and Takahashi (2012b) have introduced a particle interpretation to randomize and compute numerically the integrals in (5.4.18), which we call the FT scheme. Let η_1 be the interaction time of a particle drawn independently as the first jump time of a Poisson process with an arbitrary intensity $\mu > 0$ starting from time $t \geq 0$, i.e., η_1 is a random variable with density

$$\phi(t,s) = \mathbf{1}_{s\geq t}\,\mu\, e^{-\mu(s-t)}. \tag{5.4.19}$$

From the first line in (5.4.18), we have

$$\tilde{\Theta}^{(1)}_t = \tilde{\mathbb{E}}_t\left[\int_t^T \phi(t,s)\frac{e^{\mu(s-t)}}{\mu}\tilde{f}_s(\tilde{\Theta}^{(0)}_s)ds\right] = \tilde{\mathbb{E}}_t\left[\mathbf{1}_{\eta_1<T}\frac{e^{\mu(\eta_1-t)}}{\mu}\tilde{f}_{\eta_1}(\tilde{\Theta}^{(0)}_{\eta_1})\right]. \tag{5.4.20}$$

Similarly, the particle representation is available for the higher order. By applying the same procedure as above, we obtain

$$\tilde{\Theta}^{(2)}_t = \tilde{\mathbb{E}}_t\left[\mathbf{1}_{\eta_1<T}\tilde{\Theta}^{(1)}_{\eta_1}\frac{e^{\mu(\eta_1-t)}}{\mu}\partial_\vartheta\tilde{f}_{\eta_1}(\tilde{\Theta}^{(0)}_{\eta_1})\right],$$

where $\tilde{\Theta}^{(1)}_{\eta_1}$ can be computed by (5.4.20). Therefore, by using the tower property of conditional expectations, we obtain

$$\tilde{\Theta}^{(2)}_t = \tilde{\mathbb{E}}_t\left[\mathbf{1}_{\eta_2<T}\frac{e^{\mu(\eta_2-\eta_1)}}{\mu}\tilde{f}_{\eta_2}(\tilde{\Theta}^{(0)}_{\eta_2})\frac{e^{\mu(\eta_1-t)}}{\mu}\partial_\vartheta\tilde{f}_{\eta_1}(\tilde{\Theta}^{(0)}_{\eta_1})\right], \tag{5.4.21}$$

where η_1, η_2 are the two consecutive interaction times of a particle randomly drawn with intensity μ starting from t. Similarly, for the third order, we get

$$\tilde{\Theta}^{(3)}_t = \tilde{\mathbb{E}}_t\left[\mathbf{1}_{\eta_3<T}\frac{e^{\mu(\eta_3-\eta_2)}}{\mu}\tilde{f}_{\eta_3}(\tilde{\Theta}^{(0)}_{\eta_3})\frac{e^{\mu(\eta_2-\eta_1)}}{\mu}\partial_\vartheta\tilde{f}_{\eta_2}(\tilde{\Theta}^{(0)}_{\eta_2})\frac{e^{\mu(\eta_1-t)}}{\mu}\partial_\vartheta\tilde{f}_{\eta_1}(\tilde{\Theta}^{(0)}_{\eta_1})\right], \tag{5.4.22}$$

where η_1, η_2, η_3 are consecutive interaction times of a particle randomly drawn with intensity μ starting from t. In case $t = 0$, (5.4.20), (5.4.21) and (5.4.22) can be

simplified as

$$\begin{aligned}
\tilde{\Theta}_0^{(1)} &= \tilde{\mathbb{E}}\left[\mathbf{1}_{\zeta_1<T}\frac{e^{\mu\zeta_1}}{\mu}\tilde{f}_{\zeta_1}(\tilde{\Theta}^{(0)}_{\zeta_1})\right]\\
\tilde{\Theta}_0^{(2)} &= \tilde{\mathbb{E}}\left[\mathbf{1}_{\zeta_1+\zeta_2<T}\frac{e^{\mu\zeta_1}}{\mu}\partial_\vartheta\tilde{f}_{\zeta_1}(\tilde{\Theta}^{(0)}_{\zeta_1})\frac{e^{\mu\zeta_2}}{\mu}\tilde{f}_{\zeta_1+\zeta_2}(\tilde{\Theta}^{(0)}_{\zeta_1+\zeta_2})\right]\\
\tilde{\Theta}_0^{(3)} &= \tilde{\mathbb{E}}\left[\mathbf{1}_{\zeta_1+\zeta_2+\zeta_3<T}\frac{e^{\mu\zeta_1}}{\mu}\partial_\vartheta\tilde{f}_{\zeta_1}(\tilde{\Theta}^{(0)}_{\zeta_1})\frac{e^{\mu\zeta_2}}{\mu}\partial_\vartheta\tilde{f}_{\zeta_1+\zeta_2}(\tilde{\Theta}^{(0)}_{\zeta_1+\zeta_2})\frac{e^{\mu\zeta_3}}{\mu}\tilde{f}_{\zeta_1+\zeta_2+\zeta_3}(\tilde{\Theta}^{(0)}_{\zeta_1+\zeta_2+\zeta_3})\right]
\end{aligned} \tag{5.4.23}$$

where ζ_1, ζ_2, ζ_3 are the elapsed time from the last interaction until the next interaction, which are independent exponential random variables with parameter μ.

Note that the pricing model is originally defined with respect to the full stochastic basic $(\mathcal{G},\mathbb{Q})$. Even in the case where there exists a stochastic basis $(\mathcal{F},\mathbb{P})$ satisfying the condition (C), $(\mathcal{F},\mathbb{P})$ simulation may be nontrivial. Lemma 8.1 in Crépey and Song (2015a) allows us to reformulate the $\mathbb{P}$ expectations in (5.4.23) as the following $\mathbb{Q}$ expectations, with $\bar{\Theta}^{(0)}=0$:

$$\begin{aligned}
\tilde{\Theta}_0^{(1)} &= \bar{\Theta}_0^{(1)} = \mathbb{E}\left[\mathbf{1}_{\zeta_1<\bar{\tau}}\frac{e^{\mu\zeta_1}}{\mu}\bar{f}_{\zeta_1}(\bar{\Theta}^{(0)}_{\zeta_1})\right]\\
\tilde{\Theta}_0^{(2)} &= \bar{\Theta}_0^{(2)} = \mathbb{E}\left[\mathbf{1}_{\zeta_1+\zeta_2<\bar{\tau}}\frac{e^{\mu\zeta_1}}{\mu}\partial_\vartheta\bar{f}_{\zeta_1}(\bar{\Theta}^{(0)}_{\zeta_1})\frac{e^{\mu\zeta_2}}{\mu}\bar{f}_{\zeta_1+\zeta_2}(\bar{\Theta}^{(0)}_{\zeta_1+\zeta_2})\right]\\
\tilde{\Theta}_0^{(3)} &= \bar{\Theta}_0^{(3)} = \mathbb{E}\Big[\mathbf{1}_{\zeta_1+\zeta_2+\zeta_3<\bar{\tau}}\frac{e^{\mu\zeta_1}}{\mu}\partial_\vartheta\bar{f}_{\zeta_1}(\bar{\Theta}^{(0)}_{\zeta_1})\frac{e^{\mu\zeta_2}}{\mu}\partial_\vartheta\bar{f}_{\zeta_1+\zeta_2}(\bar{\Theta}^{(0)}_{\zeta_1+\zeta_2})\\
&\qquad\times\frac{e^{\mu\zeta_3}}{\mu}\bar{f}_{\zeta_1+\zeta_2+\zeta_3}(\bar{\Theta}^{(0)}_{\zeta_1+\zeta_2+\zeta_3})\Big],
\end{aligned} \tag{5.4.24}$$

which is nothing but the FT scheme applied to the partially reduced BSDE (II). The tractability of the FT schemes (5.4.23) and (5.4.24) relies on the nullity of the terminal condition of the related BSDEs (III) and (II), which implies that $\bar{\Theta}^{(0)} = \tilde{\Theta}^{(0)} = 0$. By contrast, an FT scheme would not be practical for the full TVA BSDE (5.3.5) with terminal condition $\xi \neq 0$. Also note that the first order in the FT scheme (5.4.23) (resp (5.4.24)) is nothing but the linear approximation (5.4.15) (resp. (5.4.14)).

5.4.3 Marked branching diffusion approach

Based on an old idea of McKean (1975), the solution $u(t_0,x_0)$ to a PDE

$$\partial_t u + \mathcal{L}u + \mu(F(u)-u) = 0, \quad u(T,x) = \Psi(x), \tag{5.4.25}$$

where $\mathcal{L}$ is the infinitesimal generator of a strong Markov process X and $F(y) = \sum_{k=0}^{d} a_k y^k$ is a polynomial of order d, admits a probabilistic representation in terms of a random tree $\mathcal{T}$ (branching diffusion). The tree starts from a single particle ("trunk") born from (t_0,x_0). Subsequently, every particle born from a node (t,x) evolves independently according to the generator $\mathcal{L}$ of X until it dies at time $t' = (t+\zeta)$ in a state x', where ζ is an independent μ-exponential time (one for each

particle). Moreover, in dying, a particle gives birth to an independent number of k' new particles starting from the node (t', x'), where k' is drawn in the finite set $\{0, 1, \cdots, d\}$ with some fixed probabilities $p_0, p_1, \cdots, p_d$. The marked branching diffusion probabilistic representation reads

$$\begin{aligned} u(t_0, x_0) &= \mathbb{E}_{t_0,x_0}\left[\prod_{\{\text{inner nodes } (t,x,k) \text{ of } \mathcal{T}\}} \frac{a_k}{p_k} \prod_{\{\text{states } x \text{ of particles alive at } T\}} \Psi(x)\right] \\ &= \mathbb{E}_{t_0,x_0}\left[\prod_{k=0}^{d}\left(\frac{a_k}{p_k}\right)^{n_k}\prod_{l=1}^{\nu}\Psi(x_l)\right], \end{aligned} \tag{5.4.26}$$

where n_k is the number of branching with k descendants up on $(0, T)$ and ν is the number of particles alive at T, with corresponding locations $x_1, \ldots, x_\nu$.

The marked branching diffusion method of Henry-Labordère (2012) for CVA computations, dubbed PHL scheme henceforth, is based on the idea that, by approximating y^+ by a well-chosen polynomial $F(y)$, the solution to the PDE

$$\partial_t u + \mathcal{L}u + \mu(u^+ - u) = 0, \quad u(T, x) = \Psi(x), \tag{5.4.27}$$

can be approximated by the solution to the PDE (5.4.25), hence by (5.4.26). We want to apply this approach to solve the TVA BSDEs (I), (II) or (III) for which, instead of fixing the approximating polynomial $F(y)$ once for all in the simulations, we need a state dependent polynomial approximation to $g_t(y) = (P_t - y)^+$ (cf. (5.3.7)) in a suitable range for y. Moreover, (I) and (II) are BSDEs with random terminal time $\bar{\tau}$, equivalently written in a Markov setup as Cauchy-Dirichlet PDE problems, as opposed to the pure Cauchy problem (5.4.27). Hence, some adaptation of the method is required. We show how to do it for (II), after which we directly give the algorithm in the similar case of (I) and in the more classical (pure Cauchy) case of (III). Assuming τ given in terms of a $(\mathcal{G}, \mathbb{Q})$ Markov factor process X as $\tau = \inf\{t > 0 : X_t \notin \mathcal{D}\}$ for some domain $\mathcal{D}$, the Cauchy-Dirichlet PDE used for approximating the partially reduced BSDE (II) reads:

$$(\partial_t + \mathcal{A})\bar{u} + \mu\left(\bar{F}(\bar{u}) - \bar{u}\right) = 0 \text{ on } [0, T] \times \mathcal{D}, \quad \bar{u}(t, x) = 0 \text{ for } t = T \text{ or } x \notin \mathcal{D}, \tag{5.4.28}$$

where $\mathcal{A}$ is the generator of X and $\bar{F}_{t,x}(y) = \sum_{k=0}^{d} \bar{a}_k(t, x)y^k$ is such that

$$\mu(\bar{F}_{t,x}(y) - y) \approx \bar{f}(t, x, y), \text{ i.e. } \bar{F}_{t,x}(y) \approx \frac{\bar{f}(t, x, y)}{\mu} + y. \tag{5.4.29}$$

Specifically, in view of (5.3.9), one can set

$$\bar{F}_{t,x}(y) = \frac{1}{\mu}\left(cdva(t, x) + \bar{\lambda} pol\big(P(t, x) - y\big) - ry\right) + y = \sum_{k=0}^{d} \bar{a}_k(t, x)y^k, \tag{5.4.30}$$

where $pol(r)$ is a d-order polynomial approximation of r^+ in a suitable range for r. The marked branching diffusion probabilistic representation of $\bar{u}(t_0, x_0) \in \mathcal{D}$

involves a random tree $\overline{\mathcal{T}}$ made of nodes and "particles" between consecutive nodes as follows. The tree starts from a single particle (trunk) born from the root (t_0, x_0). Subsequently, every particle born from a node (t, x) evolves independently according to the generator $\mathcal{A}$ of X until it dies at time $t' = (t + \zeta)$ in a state x', where ζ is an independent μ-exponential time. Moreover, in dying, if its position x' at time t' lies in $\mathcal{D}$, the particle gives birth to an independent number of k' new particles starting from the note (t', x'), where k' is drawn in the finite set $\{0, 1, \cdots, d\}$ with some fixed probabilities $p_0, p_1, \cdots, p_d$. Figure 5.1 describes such a random tree in case $d = 2$. The first particle starts from the root (t_0, x_0) and dies at time t_1, generating two new particles. The first one dies at time t_{11} and generates a new particle, who dies at time $t_{111} > T$ without descendant. The second one dies at time t_{12} and generates two new particles, where the first one dies at time t_{121} without descendant and the second one dies at time t_{122} outside the domain $\mathcal{D}$, hence also without descendant. The blue points represent the inner nodes, the red points the outer nodes and the green points the exit points of the tree out of the time-space domain $[0, T] \times \mathcal{D}$. The

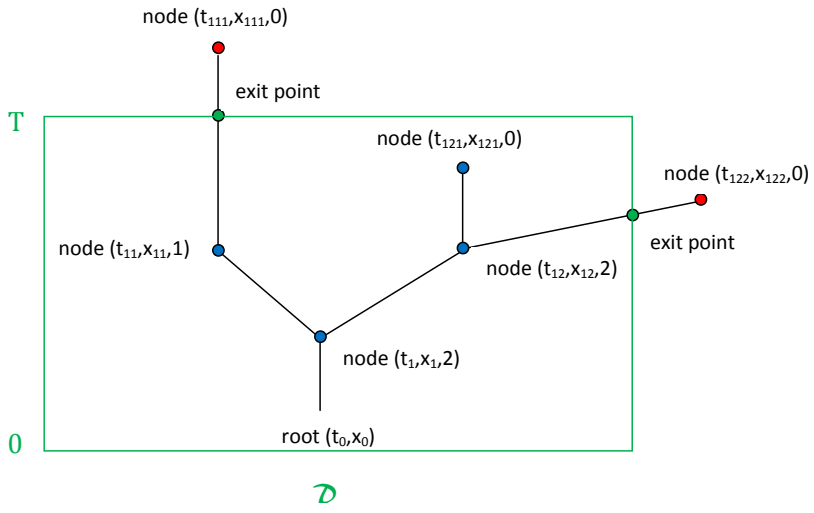

Figure 5.1: PHL random tree

marked branching diffusion probabilistic representation of $\bar{u}$ is written as

$$\bar{u}(t_0, x_0) = \mathbb{E}_{t_0,x_0}\left[\mathbf{1}_{\overline{\mathcal{T}}\subset[0,T]\times\mathcal{D}} \prod_{\{\text{inner nodes } (t,x,k) \text{ of } \overline{\mathcal{T}}\}} \frac{\bar{a}_k(t,x)}{p_k}\right], \ (t_0, x_0) \in [0, T] \times \mathcal{D}. \tag{5.4.31}$$

Specifically:

Proposition 5.4.2. *Denoting by $\bar{u}$ the function defined by the right hand side in* (5.4.31) *(assuming integrability of the integrand on the domain $[0, T]\times\mathcal{D}$), the process $Y_t = \bar{u}(t, X_t), 0 \le t \le \bar{\tau}$, solves the BSDE associated with the Cauchy-Dirichlet PDE*

(5.4.28), *namely*

$$Y_t = \mathbb{E}_t\Big[\int_t^{\bar\tau} \mu\Big(\bar F_{s,X_s}(Y_s) - Y_s\Big)ds\Big], \quad t \in [0,\bar\tau] \tag{5.4.32}$$

(which, in view of (5.4.29)*, approximates the partially reduced BSDE* (II)*, so that* $Y \approx \bar\Theta$ *provided* Y *is square integrable).*

Proof. Let (t_1, x_1, k_1) be the first branching point in the tree rooted at $(0, X_0)$ and let $\overline{\mathcal{T}}_j$ denote k_1 independent trees of the same kind rooted at (t_1, x_1). By using the independence and the strong Markov property postulated for X, we obtain

$$\begin{aligned}
\bar u(t,X_t) &= \sum_{k_1=0}^{d} \mathbb{E}_{t,X_t}\left[\mathbf{1}_{t_1<T} p_{k_1} \frac{a_{k_1}(t_1,x_1)}{p_{k_1}} \times \right. \\
&\qquad \left. \prod_{j=1}^{k_1} \mathbb{E}_{t_1,x_1}\left[\mathbf{1}_{\overline{\mathcal{T}}_j \subset [0,T]\times\mathcal{D}\}} \prod_{\{\text{inner node } (s,x,k) \text{ of } \overline{\mathcal{T}}_j\}} \frac{a_k(s,x)}{p_k}\right]\right] \\
&= \mathbb{E}_{t,X_t}\left[\mathbf{1}_{t_1<T} \sum_{k_1=0}^{d} a_{k_1}(t_1,x_1) \prod_{j=1}^{k_1} \mathbb{E}_{t_1,x_1}\left[\mathbf{1}_{\overline{\mathcal{T}}_j \subset [0,T]\times\mathcal{D}} \prod_{\{\text{inner node } (s,x,k) \text{ of } \overline{\mathcal{T}}_j\}} \frac{a_k(s,x)}{p_k}\right]\right] \\
&= \mathbb{E}_{t,X_t}\left[\mathbf{1}_{t_1<T} \sum_{k_1=0}^{d} a_{k_1}(t_1,x_1) \prod_{j=1}^{k_1} \bar u(t_1,x_1)\right] \\
&= \mathbb{E}_{t,X_t}\left[\mathbf{1}_{t_1<T} \bar F_{t_1,x_1}(\bar u(t, X_t^{t_1,x_1}))\right] \\
&= \mathbb{E}_{t,X_t}\left[\int_t^{\bar\tau} \mu(s) e^{-\int_t^s \mu(u)du} \bar F_{s,X_s^{t,x}}(\bar u(s, X_s^{t,x}))ds\right],\ 0 \le t \le \bar\tau,
\end{aligned}$$

i.e. $Y_t = \bar u(t, X_t)$ solves (5.4.32). □

If $\mathbf{1}_{\tau<T}\xi$ is given as a deterministic function $\Psi(\tau, X_\tau)$, then a similar approach (using the same tree $\overline{\mathcal{T}}$) can be applied to the full BSDE (I) in terms of the Cauchy-Dirichlet PDE

$$(\partial_t + \mathcal{A})u + \mu\,(F(u) - u) = 0 \text{ on } [0,T]\times\mathcal{D}, \quad u(t,x) = \Psi(t,x) \text{ for } t = T \text{ or } x \notin \mathcal{D}, \tag{5.4.33}$$

where $F_{t,x}(y) = \sum_{k=0}^{d}\ a_k(t,x)y^k$ is such that

$$\mu(F_{t,x}(y) - y) \approx f(t,x,y), \text{ i.e. } F_{t,x}(y) \approx \frac{f(t,x,y)}{\mu} + y.$$

This yields the approximation formula alternative to (5.4.31):

$$\Theta_0 \approx \mathbb{E}\left[\prod_{\{\text{inner node } (t,x,k) \text{ of } \overline{\mathcal{T}}\}} \frac{a_k(t,x)}{p_k} \prod_{\{\text{exit point } (t,x) \text{ of } \overline{\mathcal{T}}\}} \Psi(t,x)\right], \tag{5.4.34}$$

where an exit point of $\overline{\mathcal{T}}$ means a point where a branch of the tree leaves for the first time the time-space domain $[0,T]\times\mathcal{D}$. Last, regarding the $(\mathcal{F},\mathbb{P})$ reduced BSDE

(III), assuming an $(\mathcal{F}, \mathbb{P})$ Markov factor process $\tilde{X}$ with generator $\tilde{\mathcal{A}}$ and domain $\mathcal{D}$, we can apply a similar approach in terms of the Cauchy PDE

$$(\partial_t + \tilde{\mathcal{A}})\tilde{u} + \mu\left(\tilde{F}_{t,x}(\tilde{u}) - \tilde{u}\right) = 0 \text{ on } [0,T]\times\mathcal{D}, \quad \tilde{u}(t,x) = 0 \text{ for } t = T \text{ or } x \notin \mathcal{D}, \tag{5.4.35}$$

where $\tilde{F}_{t,x}(y) = \sum_{k=0}^{d} \tilde{a}_k(t,x)y^k$ is such that

$$\mu(\tilde{F}_{t,x}(y) - y) \approx \tilde{f}(t,x,y), \text{ i.e. } \tilde{F}_{t,x}(y) \approx \frac{\tilde{f}(t,x,y)}{\mu} + y.$$

We obtain

$$\Theta_0 = \tilde{\Theta}_0 \approx \tilde{\mathbb{E}}\left[\mathbf{1}_{\widetilde{\mathcal{T}}\subset[0,T]\times\mathcal{D}} \prod_{\text{inner node } (t,x,k) \text{ of } \widetilde{\mathcal{T}}} \frac{\tilde{a}_k(t,x)}{p_k}\right], \tag{5.4.36}$$

where $\widetilde{\mathcal{T}}$ is the branching tree associated with the Cauchy PDE (5.4.35) (similar to $\widetilde{\mathcal{T}}$ but for the generator $\widetilde{\mathcal{A}}$.

5.5 TVA models for credit derivative

Our goal is to apply the above approaches to TVA computations on credit derivatives referencing the names in $N^\star = \{1,\dots,n\}$, for some positive integer n, traded between the bank and the counterparty respectively labeled as -1 and 0. In this section we briefly survey two models of the default times τ_i, $i \in N = \{-1,0,1,\dots,n\}$, that will be used for that purpose with $\tau_b = \tau_{-1}$ and $\tau_c = \tau_0$, namely the dynamic Gaussian copula (DGC) model and the dynamic Marshall-Olkin copula (DMO) model. For more details the reader is referred to Crépey, Bielecki and Brigo (2014, Chapters 7 and 8) and Crépey and Song (2015a, Sections 6 and 7).

5.5.1 Dynamic Gaussian copula TVA model

Model of Default Times

Let there be given a function $\varsigma(\cdot)$ with unit L^2 norm on $\mathbb{R}_+$ and a multivariate Brownian motion $\mathbf{B} = (B^i)_{i\in N}$ with pairwise constant correlation $\varrho \geq 0$ in its own completed filtration $\mathcal{B} = (\mathcal{B}_t)_{t\geq 0}$. For each $i \in N$, let h_i be a continuously differentiable increasing function from $\mathbb{R}_+^*$ to $\mathbb{R}$, with $\lim_0 h_i(s) = -\infty$ and $\lim_{+\infty} h_i(s) = +\infty$, and let

$$\tau_i = h_i^{-1}(\varepsilon_i), \text{ where } \varepsilon_i = \int_0^{+\infty} \varsigma(u)dB_u^i. \tag{5.5.37}$$

Thus the $(\tau_i)_{l\in N}$ follow the standard Gaussian copula model of Li (2000), with correlation parameter ϱ and with marginal survival function $\Phi \circ h_i$ of τ_i, where Φ is the standard normal survival function. In particular, these τ_i don't intersect each other. In order to make the model dynamic as required by counterparty risk

applications, the model filtration $\mathcal{G}$ is given as the Brownian filtration $\mathcal{B}$ progressively enlarged by the τ_i, i.e.

$$\mathcal{G}_t = \mathcal{B}_t \vee \bigvee_{i\in N} \big(\sigma(\tau_i \wedge t) \vee \sigma(\{\tau_i > t\})\big),\ t \geq 0, \tag{5.5.38}$$

and the reference filtration $\mathcal{F}$ is given as $\mathcal{B}$ progressively enlarged by the default times of the reference names, i.e.

$$\mathcal{F}_t = \mathcal{B}_t \vee \bigvee_{i\in N^\star} \big(\sigma(\tau_i \wedge t) \vee \sigma(\{\tau_i > t\})\big),\ t \geq 0. \tag{5.5.39}$$

As shown in Section 6.2 of Crépey and Song (2015a), for the filtrations $\mathcal{G}$ and $\mathcal{F}$ as above, there exists a (unique) probability measure $\mathbb{P}$ equivalent to $\mathbb{Q}$ such that the condition (C) holds. For every $i \in N$, let

$$m_t^i = \int_0^t \varsigma(u) dB_u^i,\ k_t^i = \tau_i \mathbf{1}_{\{\tau_i \leq t\}},$$

and let $\mathbf{m}_t = (m_t^i)_{i\in N}$, $\mathbf{k}_t = (k_t^i)_{i\in N}$, $\tilde{\mathbf{k}}_t = (\mathbf{1}_{i\in N^\star} k_t^i)_{i\in N}$. The couple $X_t = (\mathbf{m}_t, \mathbf{k}_t)$ (resp. $\tilde{X}_t = (\mathbf{m}_t, \tilde{\mathbf{k}}_t)$) plays the role of a $(\mathcal{G}, \mathbb{Q})$ (resp. $(\mathcal{F}, \mathbb{P})$) Markov factor process in the dynamic Gaussian copula (DGC) model.

TVA Model

A DGC setup can be used as a TVA model for credit derivatives, with mark $i - 1, 0$ and $E_b = \{-1\}$, $E_c = \{0\}$. Since there are no joint defaults in this model, it is harmless to assume that the contract promises no cash-flow at τ, i.e., $\Delta_\tau = 0$, so that $Q_\tau = P_\tau$. By Crépey, Bielecki and Brigo (2014, Propositions 7.3.1 page 178 and 7.3.3 page 181), in the case of vanilla credit derivatives on the reference names, namely CDS contracts and CDO tranches (cf. (5.6.47)), there exists a continuous, explicit function $\tilde{P}^i$ such that

$$P_\tau = \tilde{P}^i(\tau, \mathbf{m}_\tau, \mathbf{k}_{\tau-}), \tag{5.5.40}$$

or $\tilde{P}_\tau^i$ in a shorthand notation, on the event $\{\tau = \tau_i\}$. Hence, (5.3.9) yields

$$\bar{f}_t(\vartheta) + r_t\vartheta = (1 - R_c)\gamma_t^0(\tilde{P}_t^0)^+ - (1 - R_b)\gamma_t^{-1}(\tilde{P}_t^{-1})^- + \bar{\lambda}_t(P_t - \vartheta)^+, \quad t \in [0, \bar{\tau}].$$

Assume that the processes r and $\bar{\lambda}$ are given before τ as continuous functions of (t, X_t), which also holds for P in the case of vanilla credit derivatives on names in N. Then the coefficients $\bar{f}$ and in turn $\tilde{f}$ are deterministically given in terms of the corresponding factor processes as

$$\bar{f}_t(\vartheta) = \bar{f}(t, X_t, \vartheta),\ \tilde{f}_t(\vartheta) = \tilde{f}(t, \tilde{X}_t, \vartheta),$$

so that we are in the Markovian setup where the FT and the PHL schemes are valid and, in principle, applicable.

5.5.2 Dynamic Marshall-Olkin copula TVA model

The above dynamic Gaussian copula model allows dealing with TVA on CDS contracts. But a Gaussian copula dependence structure is not rich enough for ensuring a proper calibration to CDS and CDO quotes at the same time. If CDO tranches are also present in a portfolio, a possible alternative is the following dynamic Marshall-Olkin (DMO) copula model, also known as the "common shock" model.

Model of default times

We define a family $\mathcal{Y}$ of "shocks", i.e. subsets $Y \subseteq N$ of obligors, usually consisting of the singletons $\{-1\}, \{0\}, \{1\}, \ldots, \{n\}$, and a few "common shocks" $I_1, I_2, \cdots, I_m$ representing simultaneous defaults. For $Y \in \mathcal{Y}$, the shock time' η_Y is defined as an i.i.d. exponential random variable with parameter γ_Y. The default time of obligor i in the common shock model is then defined as

$$\tau_i = \min_{Y \in \mathcal{Y}, i \in Y} \eta_Y. \tag{5.5.41}$$

Example 5.5.3. Figure 5.2 shows one possible default path in a common-shock model with $n = 3$ and $\mathcal{Y} = \{\{-1\}, \{0\}, \{1\}, \{2\}, \{3\}, \{2, 3\}, \{0, 1, 2\}, \{-1, 0\}\}$. The inner oval shows which shocks happened and caused the observed default scenarios at successive default times.

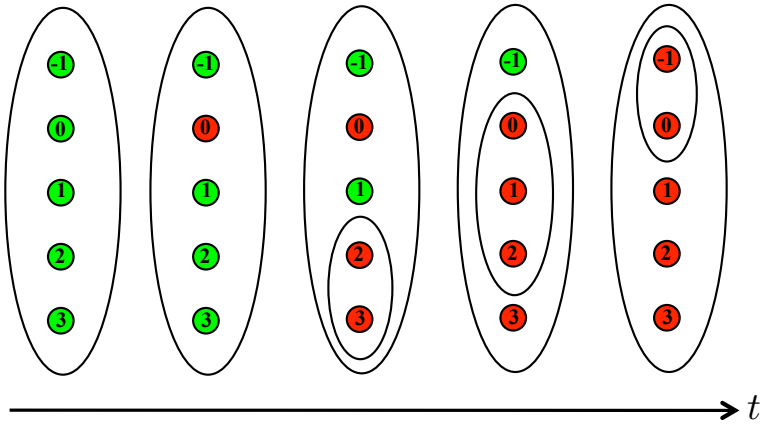

Figure 5.2: One possible default path in the common-shock model with $n = 3$ and $\mathcal{Y} = \{\{-1\}, \{0\}, \{1\}, \{2\}, \{3\}, \{2, 3\}, \{0, 1, 2\}, \{-1, 0\}\}$.

The full model filtration $\mathcal{G}$ is defined as

$$\mathcal{G}_t = \bigvee_{Y \in \mathcal{Y}} \big(\sigma(\eta_Y \wedge t) \vee \sigma(\{\eta_Y > t\})\big),\ t \geq 0.$$

Letting $\mathcal{Y}_\circ = \{Y \in \mathcal{Y};\ -1, 0 \notin Y\}$, the reference filtration $\mathcal{F}$ is given as

$$\mathcal{F}_t = \bigvee_{Y \in \mathcal{Y}_\circ} \big(\sigma(\eta_Y \wedge t) \vee \sigma(\{\eta_Y > t\})\big),\ t \geq 0.$$

As shown in Section 7.2 of Crépey and Song (2015a), in the DMO model with $\mathcal{G}$ and $\mathcal{F}$ as above, the condition (C) holds for $\mathbb{P} = \mathbb{Q}$. Let $J^Y = \mathbf{1}_{[0,\eta_Y)}$. Similar to $(\mathbf{m}, \mathbf{k})$ (resp. $(\mathbf{m}, \tilde{\mathbf{k}})$) in the DGC model, the process

$$X = I = (J^Y)_{Y\in\mathcal{Y}} \text{ (resp. } \tilde{X} = (\mathbf{1}_{Y\in\mathcal{Y}_\circ} J^Y)_{Y\in\mathcal{Y}}) \tag{5.5.42}$$

plays the role of a $(\mathcal{G}, \mathbb{Q})$ (resp. $(\mathcal{F}, \mathbb{P})$) Markov factor in the DMO model.

TVA model

A DMO setup can be used as a TVA model for credit derivatives, with

$$E_b = \mathcal{Y}_b := \{Y \in \mathcal{Y};\, -1 \in Y\},\ E_c = \mathcal{Y}_c := \{Y \in \mathcal{Y};\, 0 \in Y\},\ E = \mathcal{Y}_\bullet := \mathcal{Y}_b \cup \mathcal{Y}_c$$

and

$$\tau_b = \tau_{-1} = \min_{Y\in\mathcal{Y}_b} \eta_Y,\ \tau_c = \tau_0 = \min_{Y\in\mathcal{Y}_c} \eta_Y,$$

hence

$$\tau = \min_{Y\in\mathcal{Y}_\bullet} \eta_Y,\ \gamma = \mathbf{1}_{[0,\tau)}\tilde{\gamma} \text{ with } \tilde{\gamma} = \sum_{Y\in\mathcal{Y}_\bullet} \gamma_Y. \tag{5.5.43}$$

By Crépey, Bielecki and Brigo (2014, Proposition 8.3.1 page 205), in the case of CDS contracts and CDO tranches, for every shock $Y \in \mathcal{Y}$ and process $U = P$ or Δ, there exists a continuous, explicit function $\tilde{U}_Y$ such that

$$U_\tau = \tilde{U}_Y(\tau, \mathbf{J}_{\tau-}), \tag{5.5.44}$$

or $\tilde{U}^Y_\tau$ in a shorthand notation, on the event $\{\tau = \eta_Y\}$. The coefficient $\bar{f}_t(\vartheta)$ in (5.3.9) is then given, for $t \in [0, \bar{\tau}]$, by

$$\begin{aligned}\bar{f}_t(\vartheta) + r_t\vartheta = (1-R_c)\sum_{Y\in\mathcal{Y}_c} \gamma^Y_t \big(\tilde{P}^Y_t + \tilde{\Delta}^Y_t\big)^+ - (1-R_b)\sum_{Y\in\mathcal{Y}_b} \gamma^Y_t \big(\tilde{P}^Y_t + \tilde{\Delta}^Y_t\big)^- \\ + \bar{\lambda}_t (P_t - \vartheta)^+.\end{aligned} \tag{5.5.45}$$

Assuming that the processes r and $\bar{\lambda}$ are given before τ as continuous functions of (t, X_t), which also holds for P in case of vanilla credit derivatives on the reference names, then

$$\bar{f}_t(\vartheta) = \bar{f}(t, X_t, \vartheta), \tilde{f}_t(\vartheta) = \bar{f}_t(\vartheta) - \tilde{\gamma}\vartheta = \tilde{f}(t, \tilde{X}_t, \vartheta) \tag{5.5.46}$$

(cf. (5.5.43)), so that we are again in a Markovian setup where the FT and the PHL schemes are valid and, in principle, applicable.

5.5.3 Strong versus weak dynamic copula model

However, one peculiarity of the TVA BSDEs in our credit portfolio models is that, even though full and reduced Markov structures have been identified, which is required for justifying the validity of the FT and/or PHL numerical schemes, and the

corresponding generators $\mathcal{A}$ or $\tilde{\mathcal{A}}$ can be written explicitly (see Crépey and Song (2015b)), the Markov structures are too heavy for being of any practical use in the numerics. Instead, fast and exact simulation and clean pricing schemes are available based on the dynamic copula structures.

Moreover, in the case of the DGC model, we lose the Gaussian copula structure after a branching point in the PHL scheme. In fact, as visible on Crépey, Bielecki and Brigo (2014, Formula (7.7) p. 175), the DGC conditional multivariate survival probability function is stated in terms of a ratio of Gaussian survival probability functions, which is explicit but does not simplify into a single Gaussian survival probability function. It's only in the DMO model that the conditional multivariate survival probability function, which arises as a ratio of exponential survival probability functions (see Crépey, Bielecki and Brigo (2014, Formula (8.11) p. 197 and Section 8.2.1.1)), simplifies into a genuine exponential survival probability function. Hence, the PHL scheme is not applicable in the DGC model.

The FT scheme based on (III) is not practical either because the Gaussian copula structure is only under $\mathbb{Q}$ and, again, the (full or reduced) Markov structures are not practical. In the end, the only practical scheme in the DGC model is the FT scheme based on the partially reduced BSDE (II). Eventually, it's only in the DMO model that the FT and the PHL schemes are both practical and can be compared numerically.

5.6 Numerics

For the numerical implementation, we consider stylized CDS contracts and protection legs of CDO tranches corresponding to dividend processes D of the respective form, for $0 \leq t \leq T$:

$$\begin{aligned} D_t = D_t^i &= \big((1-R_i)\mathbf{1}_{t\geq\tau_i} - S_i(t\wedge\tau_i)\big)Nom_i \\ D_t = D_t^\star &= \Big(\big((1-R_\star)\sum_{j\in N}\mathbf{1}_{t\geq\tau_j} - (n+2)a\big)^+ \wedge (n+2)(b-a)\Big)Nom_\star, \end{aligned} \tag{5.6.47}$$

where all the recoveries R_i and $R_\star$ (resp. nominals Nom_i and $Nom_\star$) are set to 40% (resp. to 100). The contractual spreads S_i of the CDS contracts are set such that the corresponding prices are equal to 0 at time 0. Protection legs of CDO tranches, where the attachment and detachment points a and b are such that $0 \leq a \leq b \leq 100\%$, can also be seen as CDO tranches with upfront payment. Note that credit derivatives traded as swaps or with upfront payment coexist since the crisis. Unless stated otherwise, the following numerical values are used:

$$r = 0, R_b = 1, R_c = 40\%, \bar{\lambda} = 100 \text{ bp} = 0.01, \mu = \frac{2}{T}, m = 10^4.$$

5.6.1 Numerical results in the DGC model

First we consider DGC random times τ_i defined by (5.5.37), where the function h_i is chosen so that τ_i follows an exponential distribution with parameter $\gamma_i^\star$ (which

in practice can be calibrated to a related CDS spread or a suitable proxy). More precisely, let Φ and Ψ_i be the survival functions of a standard normal distribution and an exponential distribution with intensity $\gamma_i^\star$. We choose $h_i = \Phi^{-1} \circ \Psi_i$, so that (cf. (5.5.37))

$$\mathbb{P}(\tau_i \geq t) = \mathbb{P}\left(\Psi_i^{-1}\left(\Phi\left(\varepsilon_i\right)\right) \geq t\right) = \mathbb{P}\Big(\Phi\left(\varepsilon_i\right) \leq \Psi_i(t)\Big) = \Psi_i(t),$$

for $\Phi\left(\varepsilon_i\right)$ has a standard uniform distribution. Moreover, we use a function $\varsigma(\cdot)$ in (5.5.37) constant before a time horizon $U > T$ and null after U, so that $\varsigma(0) = \frac{1}{\sqrt{U}}$ (given the constraint that $\nu^2(0) = \int_0^\infty \varsigma^2(s)ds = 1$) and

$$\nu^2(t) = \int_t^\infty \varsigma^2(s)ds = \frac{U-t}{U},\ m_t^i = \int_0^t \varsigma(u)dB_u^i = \frac{1}{\sqrt{U}}B_t^i,\ \int_0^\infty \varsigma(u)dB_u^i = \frac{1}{\sqrt{U}}B_U^i.$$

In the case of the DGC model, the only practical TVA numerical scheme is the FT scheme (5.4.24) based on the partially reduced BSDE (II), which can be described by the following steps:

1. Draw an time ζ_1 following an exponential law of parameter μ. If $\zeta_1 < T$, then simulate $\mathbf{m}_{\zeta_1} = (\frac{1}{\sqrt{U}}B_{\zeta_1}^i)_{l\in N} \sim \mathcal{N}(0, \frac{\zeta_1}{U}I_n(1,\varrho))$, where $I_n(1,\varrho)$ is a $n \times n$ matrix with diagonal equal to 1 and all off-diagonal entries equal to ϱ, and go to Step 2. Otherwise, go to Step 4.

2. Draw a second time ζ_2, independent from ζ_1, following an exponential law of parameter μ. If $\zeta_1 + \zeta_2 < T$, then obtain the vector $\mathbf{m}_{\zeta_1+\zeta_2}$ as $\mathbf{m}_{\zeta_1} + (\mathbf{m}_{\zeta_1+\zeta_2} - \mathbf{m}_{\zeta_1})$, where $\mathbf{m}_{\zeta_1+\zeta_2} - \mathbf{m}_{\zeta_1} = (\frac{1}{\sqrt{U}}(B_{\zeta_1+\zeta_2}^i - B_{\zeta_1}^i))_{l\in N} \sim \mathcal{N}(0, \frac{\zeta_2}{U}I_n(1,\varrho))$, and go to Step 3. Otherwise, go to Step 4.

3. Draw a third time ζ_3, independent from ζ_1 and ζ_2, following an exponential law of parameter μ. If $\zeta_1 + \zeta_2 + \zeta_3 < T$, then obtain the vector $\mathbf{m}_{\zeta_1+\zeta_2+\zeta_3}$ as $\mathbf{m}_{\zeta_1+\zeta_2} + (\mathbf{m}_{\zeta_1+\zeta_2+\zeta_3} - \mathbf{m}_{\zeta_1+\zeta_2})$, where $\mathbf{m}_{\zeta_1+\zeta_2+\zeta_3} - \mathbf{m}_{\zeta_1+\zeta_2} = (\frac{1}{\sqrt{U}}(B_{\zeta_1+\zeta_2+\zeta_3}^i - B_{\zeta_1+\zeta_2}^i))_{l\in N} \sim \mathcal{N}(0, \frac{\zeta_3}{U}I_n(1,\varrho))$. Go to Step 4.

4. Simulate the vector $\mathbf{m}_U$ from the last simulated vector $\mathbf{m}_t$ ($t = 0$ by default) as $\mathbf{m}_t + (\mathbf{m}_U - \mathbf{m}_t)$, where $\mathbf{m}_U - \mathbf{m}_t = (\frac{1}{\sqrt{U}}(B_U^i - B_t^i))_{i\in N} \sim \mathcal{N}(0, \frac{U-t}{U}I_n(1,\varrho))$. Deduce $(B_U^i)_{i\in N}$, hence $\tau_i = \Psi_i^{-1} \circ \Phi\left(\frac{1}{\sqrt{U}}B_U^i\right)$, $i \in N$, and in turn the vectors $\mathbf{k}_{\zeta_1}$ (if $\zeta_1+\zeta_2+\zeta_3 < T$), $\mathbf{k}_{\zeta_1+\zeta_2}$ (if $\zeta_1+\zeta_2 < T$) and $\mathbf{k}_{\zeta_1+\zeta_2+\zeta_3}$ (if $\zeta_1+\zeta_2+\zeta_3 < T$). Eventually compute $\bar{f}_{\zeta_1}$, $\bar{f}_{\zeta_1+\zeta_2}$, and $\bar{f}_{\zeta_1+\zeta_2+\zeta_3}$ for the three orders of the FT scheme.

We perform TVA computations on CDS contracts with maturity $T = 10$ years, choosing for that matter $U = T + 1 = 11$ years, hence $\varsigma = \frac{\mathbf{1}_{[0,11]}}{\sqrt{11}}$, for $\varrho = 0.6$ unless otherwise stated. Table 5.1 displays the contractual spreads of the CDS contracts used in these experiments. In Figure 5.3, the left graph shows the TVA on a CDS on name 1, computed in a DGC model with $n = 1$ by FT scheme of order 1 to 3, for different levels of nonlinearity represented by the value of the unsecured borrowing

i	-1	0	1
S_i	36	41	47

i	-1	0	1	2	3	4	5	6	7	8	9	10
S_i	39	40	47	36	41	48	54	54	27	30	36	50

Table 5.1: Time-0 bp CDS spreads of names -1 (the bank), 0 (the counterparty) and of the reference names 1 to n used when $n = 1$ (*left*) and $n = 10$ (*right*).

spread $\bar{\lambda}$. The right graph shows similar results regarding a portfolio comprising one CDS contract per name $i = 1, \ldots, 10$. The time-0 clean value of the default leg of the CDS in case $n = 1$, respectively the sum of the ten default legs in case $n = 10$, is 4.52, respectively 40.78 (of course $P_0 = 0$ in both cases by definition of fair contractual spreads). Hence, in relative terms, the TVA numbers visible in Figure 5.3 are quite high, much greater for instance than in the cases of counterparty risk on interest rate derivatives considered in Crépey, Gerboud, Grbac, and Ngor (2013). This is explained by the wrong-way risk feature of the DGC model, namely, the default intensities of the surviving names and the value of the CDS protection spike at defaults in this model. When $\bar{\lambda}$ increases (for $\bar{\lambda} = 0$ that's a case of linear TVA where FT higher order terms equal 0), the second (resp. third) FT term may represent in each case up to 5% to 10% of the first (resp. second) FT term, from which we conclude that the first FT term can be used as a first order linear estimate of the TVA, with a nonlinear correction that can be estimated by the second FT term.

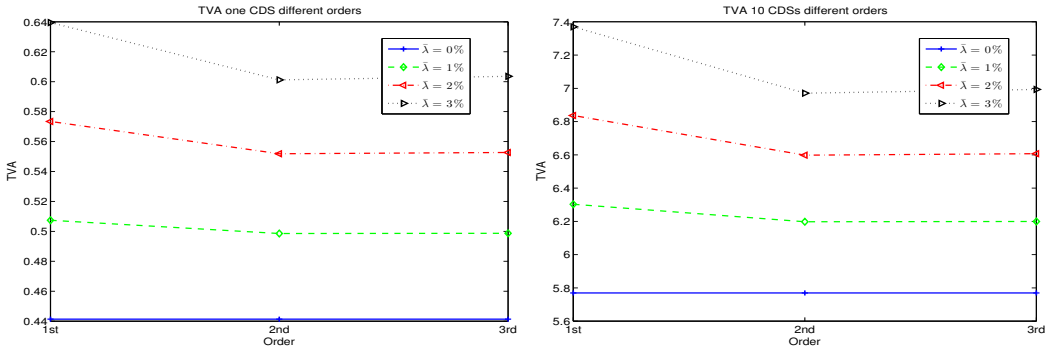

Figure 5.3: *Left*: DGC TVA on one CDS computed by FT scheme of order 1 to 3, for different levels of nonlinearity (unsecured borrowing spread $\bar{\lambda}$). *Right*: Similar results regarding the portfolio of CDS contracts on ten names.

In Figure 5.4, the left graph shows the TVA on one CDS computed by FT scheme of order 3 as a function of the DGC correlation parameter ϱ, with other parameters set as before. The right graph shows the analogous results regarding the portfolio of ten CDS contracts. In both cases, the TVA numbers increase (roughly linearly) with ϱ, including for high values of ϱ, as desirable from the financial interpretation point of view, whereas it has been noted in Brigo and Chourdakis (2008) (see the blue curve in Figure 1 of the ssrn version of the paper) that for high levels of the

correlation between names, other models may show some pathological behaviours.

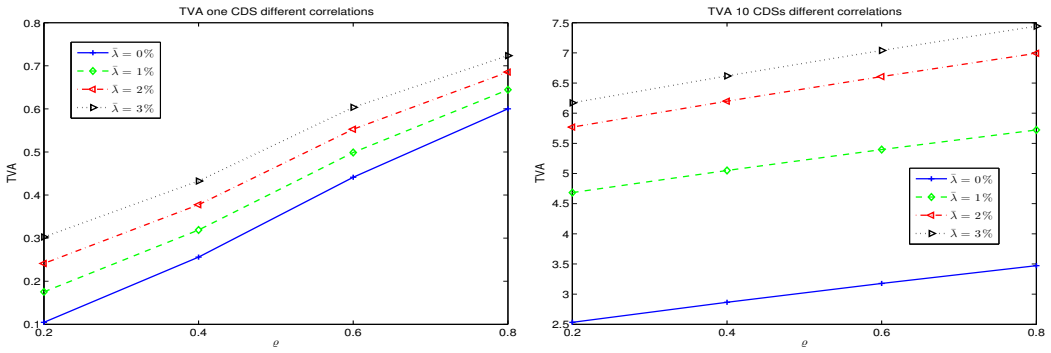

Figure 5.4: *Left*: TVA on one CDS computed by FT scheme of order 3 as a function of the DGC correlation parameter ϱ. *Right*: Similar results regarding a portfolio of CDS contracts on ten different names.

In Figure 5.5, the left graph shows that the errors, in the sense of the % relative standard errors (% rel. SE), of the different orders of the FT scheme don't explode with the dimension (number of credit names that underlie the CDS contracts). The middle graph, produced with $n = 1$, shows that the errors don't explode with the level of nonlinearity represented by the unsecured borrowing spread $\bar{\lambda}$. Consistent with the fact that the successive FT terms are computed by purely forward Monte Carlo schemes, their computation times are essentially linear in the number of names, as visible in the right graph.

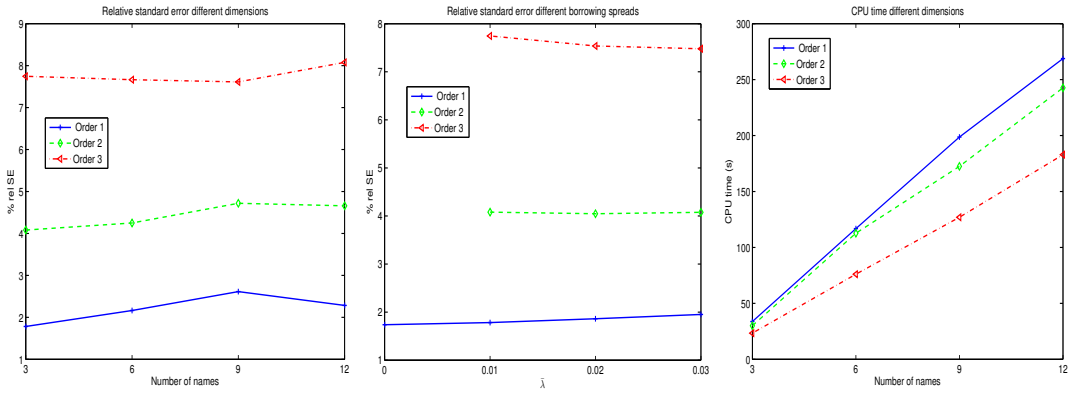

Figure 5.5: *Left*: The % relative standard errors of the different orders of the expansions don't explode with the number of names ($\bar{\lambda} = 100$ bp). *Middle*: The % relative standard errors of the different orders of the expansions don't explode with the level of nonlinearity represented by the unsecured borrowing spread $\bar{\lambda}$ ($n = 1$). *Right*: Since FT terms are computed by purely forward Monte Carlo schemes, their computation times are linear in the number of names ($\bar{\lambda} = 100$ bp).

To conclude this section, we compare the linear approximation (5.4.14) corresponding to the first FT term in (5.4.24) (FT1 in Table 5.2) with the linear approximations (5.4.12)-(5.4.13) (LA in Table 5.2). One can see from Table 5.2 that the LA and FT1 estimates are consistent (at least in the sense of their 95% confidence intervals, which always intersect each other). But the LA standard errors are larger than the FT1 ones. In fact, using the formula for the intensity γ of τ in FT1 can be viewed as a form of variance reduction with respect to LA, where τ is simulated. Of course, for $\bar{\lambda} \neq 0$ (case of the right tables where $\bar{\lambda} = 3\%$), both linear approximations are biased as compared with the complete FT estimate (with nonlinear correction, also shown in Table 5.2), particularly in the high dimensional case with 10 CDS contracts (see the bottom panels in Table 5.2). Figure 5.6 completes these results by showing the LA, FT1 and FT standard errors computed for different levels of nonlinearity and different dimensions.

Summarizing, in the DGC model, the PHL is not practical. The FT scheme based on the partially reduced TVA BSDE (II) gives an efficient way of estimating the TVA. The nonlinear correction with respect to the linear approximations (5.4.14) or (5.4.15) amounts up to 5% in relative terms, depending on the unsecured borrowing spread $\bar{\lambda}$.

Method	TVA	95% CI	Rel. SE
LA	0.65	$[0.57, 0.73]$	6.08 %
FT1	0.61	$[0.59, 0.63]$	1.66%
FT	0.60	$[0.58, 0.62]$	1.64 %

Method	TVA	95% CI	Rel. SE
LA	0.66	$[0.60, 0.72]$	4.39%
FT1	0.62	$[0.59, 0.64]$	1.96%
FT	0.60	$[0.58, 0.63]$	1.84%

Method	TVA	95% CI	Rel. SE
LA	6.17	$[5.43, 6.92]$	6.03%
FT1	6.24	$[5.77, 6.72]$	3.78%
FT	6.17	$[5.66, 6.68]$	4.15%

Method	TVA	95% CI	Rel. SE
LA	6.81	$[6.16, 7.45]$	4.76%
FT1	7.82	$[7.39, 8.25]$	2.73%
FT	6.99	$[6.67, 7.31]$	2.28%

Table 5.2: LA, FT1 and FT estimates: 1 CDS (*top*) and 10 CDSs (*bottom*), with parameters $\bar{\lambda} = 0\%$, $\varrho = 0.8$ (*left*) and $\bar{\lambda} = 3\%$, $\varrho = 0.6$ (*right*).

5.6.2 Numerical results in the DMO model

In the DMO model, the FT scheme (5.4.18) for the fully reduced BSDE (5.4.23) can be implemented through following steps:

1. Simulate the time η_Y of each (individual or joint) shock following an independent exponential law of parameter γ_Y, $Y \in \mathcal{Y}$, then retrieve the τ_i through the formula (5.5.41).

2. Draw a time ζ_1 following an exponential law of parameter μ. If $\zeta_1 < T$, compare the default time of each name with ζ_1 to obtain the reduced Markov

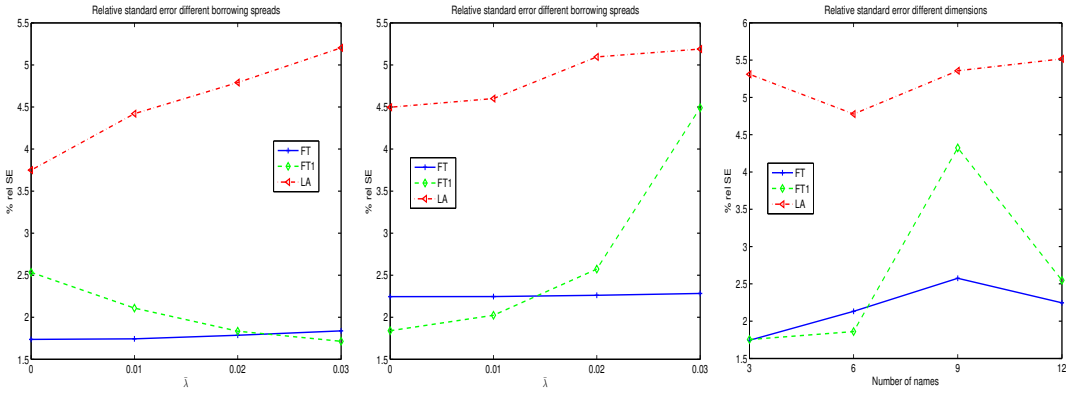

Figure 5.6: The % relative standard errors of the different tranches don't explode with the level of nonlinearity represented by the unsecured borrowing spread $\bar{\lambda}$. *Left*: 1 CDS. *Middle*: 10 CDSs. *Right*: the % relative standard errors of the different schemes (LA, FT1, FT in figures) don't explode with the number of names ($\bar{\lambda} = 100$ bp, $\varrho = 0.6$).

factor $\tilde{X}_{\zeta_1}$ as of (5.5.42) and in turn $\tilde{f}_{\zeta_1}$ as of (5.5.45)-(5.5.46), then go to Step 3. Otherwise stop.

3. Draw a second time ζ_2 following an independent exponential law of parameter μ. If $\zeta_1 + \zeta_2 < T$, compare the default time τ_i of each name with $\zeta_1 + \zeta_2$ to obtain the Markov factor $\tilde{X}_{\zeta_1+\zeta_2}$ and $\tilde{f}_{\zeta_1+\zeta_2}$ then go to Step 4. Otherwise stop.

4. Draw a third time ζ_3 following an independent exponential law of parameter μ. If $\zeta_1+\zeta_2+\zeta_3 < T$, compare the default time of each name with $\zeta_1+\zeta_2+\zeta_3$ to obtain the Markov factor $\tilde{X}_{\zeta_1+\zeta_2+\zeta_3}$ and $\tilde{f}_{\zeta_1+\zeta_2+\zeta_3}$.

We can also consider the PHL scheme (5.4.31) based on the partially reduced BSDE (II) with

$$\mathcal{D} = \{x = (x^Y)_{Y\in\mathcal{Y}} \in \{0,1\}^{\mathcal{Y}} \text{ such that } x^Y = 1 \text{ for } Y \in \mathcal{Y}_\bullet\}.$$

To simulate the random tree $\overline{\mathcal{T}}$ in (5.4.31), we follow the approach sketched before (5.4.31) where, in order to evolve $X = \mathbf{J}$ according to the DMO generator $\mathcal{A}$ during a time interval ζ, a particle born from a node $x = (j_Y)_{Y\in\mathcal{Y}} \in \{0,1\}^{\mathcal{Y}}$ at time t, all one needs is, for each Y such that $j_Y = 1$, draw an independent exponential random variable θ_Y of parameter γ_Y and then set $x' = (j_Y \mathbf{1}_{[0,\theta_Y)}(\zeta))_{Y\in\mathcal{Y}}$. Rephrasing in more algorithmic terms:

1. To simulate the random tree $\overline{\mathcal{T}}$ under the expectation in (5.4.31), we repeat the following step (generation of particles, or segments between consecutive nodes of the tree) until a generation of particles dies without children:

   For each node $(t, x = (j_Y)_{Y\in\mathcal{Y}}, k)$ issued from the previous generation of particles (starting with the root-node $(0, X_0, k = 1)$), for each

of the k new particles, indexed by l, issued from that node, simulate an independent exponential random variable ζ_l and set

$$(t'_l, x'_l, k'_l) = (t + \zeta_l, (j_Y \mathbf{1}_{[0,\theta^l_Y)}(\zeta_l))_{Y\in\mathcal{Y}}, \mathbf{1}_{x'_l\in\mathcal{D}}\nu_l),$$

where, for each l, the θ^l_Y are independent exponential-γ_Y random draws and ν_l is an independent draw in the finite set $\{0, 1, \cdots, d\}$ with some fixed probabilities $p_0, p_1, \ldots, p_d$.

2. To compute the random variable Φ under the expectation in (5.4.31), we loop over the nodes of the tree $\overline{\mathcal{T}}$ thus constructed (if $\overline{\mathcal{T}} \subset [0,T]\times\mathcal{D}$, otherwise $\Phi = 0$ in the first place) and we form the product in (5.4.31), where the $\bar{a}_k(t,x)$ are retrieved as in (5.4.30).

The PHL schemes (5.4.34) based on the full BSDE (I) or (5.4.36) based on the fully reduced BSDE (III) can be implemented along similar lines.

We perform TVA computations in a DMO model with $n = 120$, for individual shock intensities taken as $\gamma_{\{i\}} = 10^{-4}\times(100+i)$ (increasing from ~ 100 bps to 220 bps as i increases from 1 to 120) and four nested groups of common shocks $I_1 \subset I_2 \subset I_3 \subset I_4$, respectively consisting of the riskiest 3%, 9%, 21% and 100% (i.e. all) names, with respective shock intensities $\gamma_{I_1} = 20$ bp, $\gamma_{I_2} = 10$ bp, $\gamma_{I_3} = 6.67$ bp and $\gamma_{I_4} = 5$ bp. The counterparty (resp. the bank) is taken as the eleventh (resp. tenth) safest name in the portfolio. In the model thus specified, we consider CDO tranches with upfront payment, i.e. credit protection bought by the bank from the counterparty at time 0, with nominal 100 for each obligor, maturity $T = 2$ years and attachment (resp. detachment) points are 0%, 3% and 14% (resp. 3%, 14% and 100%). The respective value of P_0 (upfront payment) for the equity, mezzanine and senior tranche is 229.65, 5.68 and 2.99. Accordingly, the ranges of approximation chosen for $pol(y) \approx y^+$ in the respective PHL schemes are 250, 200 and 10. We use polynomial approximation of order $d = 4$ with $(p_0, p_1, p_2, p_3, p_4) = (0.5, 0.3, 0.1, 0.09, 0.01)$. We set $\mu = 0.1$ in all PHL schemes and $\mu = 2/T = 0.2$ in all FT schemes.

Figure 5.7 shows the TVA computed by the FT scheme (5.4.23) based on the fully reduced BSDE (III), for different levels of nonlinearity (unsecured borrowing basis $\bar{\lambda}$). We observe that, in all cases, the third order term is negligible. Hence, in further FT computations, we only compute the orders 1 (linear part) and 2 (nonlinear correction).

Table 5.3 compares the results of the above FT scheme (5.4.23) based on the fully reduced BSDE (III) with those of the PHL schemes (5.4.36) based on (III) again ($\widetilde{\text{PHL}}$ in the tables), (5.4.31) based on the partially reduced BSDE (II) ($\overline{\text{PHL}}$ in the tables) and (5.4.34) based on the full BSDE (I) (PHL in the tables), for the three CDO tranches and two sets of parameters. The three PHL schemes are of course slightly biased, but the first two, based on the BSDEs with null terminal condition (III) or (II), exhibit much less variance than the third one, based on the full BSDE with terminal condition ξ. This is also visible in Figure 5.9 (note the different scales of the y axes going from left to right in the picture), which also

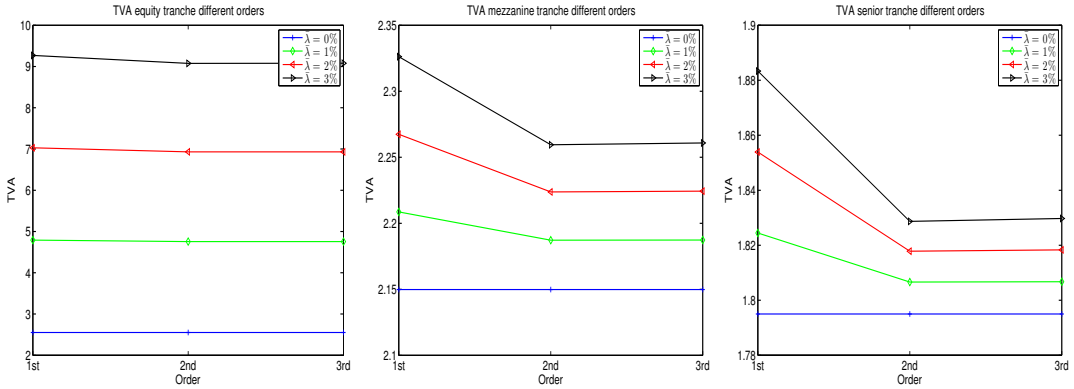

Figure 5.7: TVA on CDO tranches with 120 underlying names computed by FT scheme of order 1 to 3 for different levels of nonlinearity (unsecured borrowing basis $\bar{\lambda}$). *Left*: Equity tranche. *Middle*: Mezzanine tranche. *Right*: Senior Tranche.

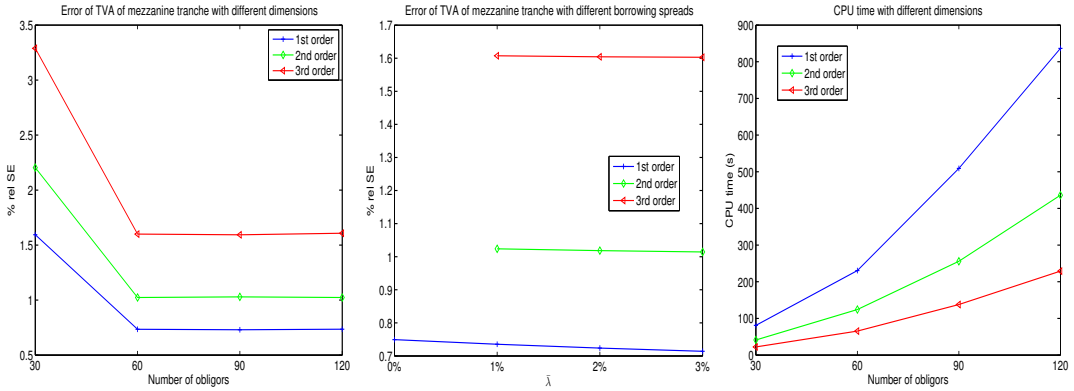

Figure 5.8: Analog of Fig. 5.5 for the CDO tranche of Fig. 5.7 in the DMO model ($\bar{\lambda} = 0.01$).

shows that, for any of these schemes, the relative standard errors don't explode with the level of nonlinearity or the number of reference names in the CDO (the results for the $\overline{\text{PHL}}$ scheme are not shown on the figure as very similar to those of the $\widetilde{\text{PHL}}$ scheme). In comparing the TVA values on the left and the right hand side of Table 5.3, we see that the intensities of the common shocks, which play a role similar to the correlation ϱ in the DGC model, have a more important impact on the higher tranches (mezzanine and senior tranche), whereas the equity tranche is more sensitive to the level of the unsecured borrowing spread $\bar{\lambda}$.

Method	TVA	95% CI	Rel. SE
FT	3.13	[3.10 , 3.16]	0.48 %
$\widetilde{\text{PHL}}$	3.07	[2.87 , 3.28]	3.35 %
$\overline{\text{PHL}}$	3.16	[2.94 , 3.37]	3.37 %
PHL	2.53	[2.13 , 2.94]	8.02%

Method	TVA	95% CI	Rel. SE
FT	9.08	[9.00 , 9.16]	0.46 %
$\widetilde{\text{PHL}}$	9.05	[8.40 , 9.70]	3.58 %
$\overline{\text{PHL}}$	9.28	[8.63 , 9.94]	3.51 %
PHL	12.59	[6.92 , 18.27]	22.54%

Method	TVA	95% CI	Rel. SE
FT	6.43	[6.33 , 6.53]	0.75 %
$\widetilde{\text{PHL}}$	6.34	[5.93 , 6.75]	3.22 %
$\overline{\text{PHL}}$	6.34	[5.93 , 6.75]	3.25 %
PHL	4.86	[2.84 , 6.89]	20.82%

Method	TVA	95% CI	Rel. SE
FT	2.29	[2.25 , 2.32]	0.77 %
$\widetilde{\text{PHL}}$	2.51	[2.35 , 2.67]	3.17 %
$\overline{\text{PHL}}$	2.68	[2.52 , 2.85]	3.12 %
PHL	1.93	[0.79 , 3.08]	29.57%

Method	TVA	95% CI	Rel. SE
FT	5.32	[5.24 , 5.40]	0.75 %
$\widetilde{\text{PHL}}$	5.24	[4.90 , 5.58]	3.22 %
$\overline{\text{PHL}}$	5.25	[4.90 , 5.58]	3.25 %
PHL	4.01	[2.32 , 5.70]	21.03%

Method	TVA	95% CI	Rel. SE
FT	1.83	[1.80 , 1.86]	0.78 %
$\widetilde{\text{PHL}}$	1.80	[1.69 , 1.92]	3.13 %
$\overline{\text{PHL}}$	1.87	[1.75 , 1.99]	3.11 %
PHL	1.36	[0.41 , 2.31]	35.05%

Table 5.3: FT, PHL, $\overline{\text{PHL}}$ and $\widetilde{\text{PHL}}$ schemes applied to the equity (*top*), mezzanine (*middle*) and senior (*bottom*) tranche, for the parameters $\bar{\lambda} = 0\%$, $\lambda_{I_j} = 60bp/j$ (*left*) or $\bar{\lambda} = 3\%$, $\lambda_{I_j} = 20bp/j$ (*right*).

5.7 Conclusion

Under mild assumptions, three equivalent TVA BSDEs are available. The original "full" BSDE (I) is stated with respect to the full model filtration $\mathcal{G}$ and the original pricing measure $\mathbb{Q}$. It does not involve the intensity γ of the counterparty first-to-default time τ. The partially reduced BSDE (II) is also stated with respect to $(\mathcal{G}, \mathbb{Q})$ but it involves both τ and γ. The fully reduced BSDE (III) is stated with respect to a smaller "reference filtration" $\mathcal{F}$ and it only involves γ. Hence, in principle, the full BSDE (I) should be preferred for models with a "simple" τ whereas the fully reduced BSDE (III) should be preferred for models with a "simple" γ. But, in nonimmersive

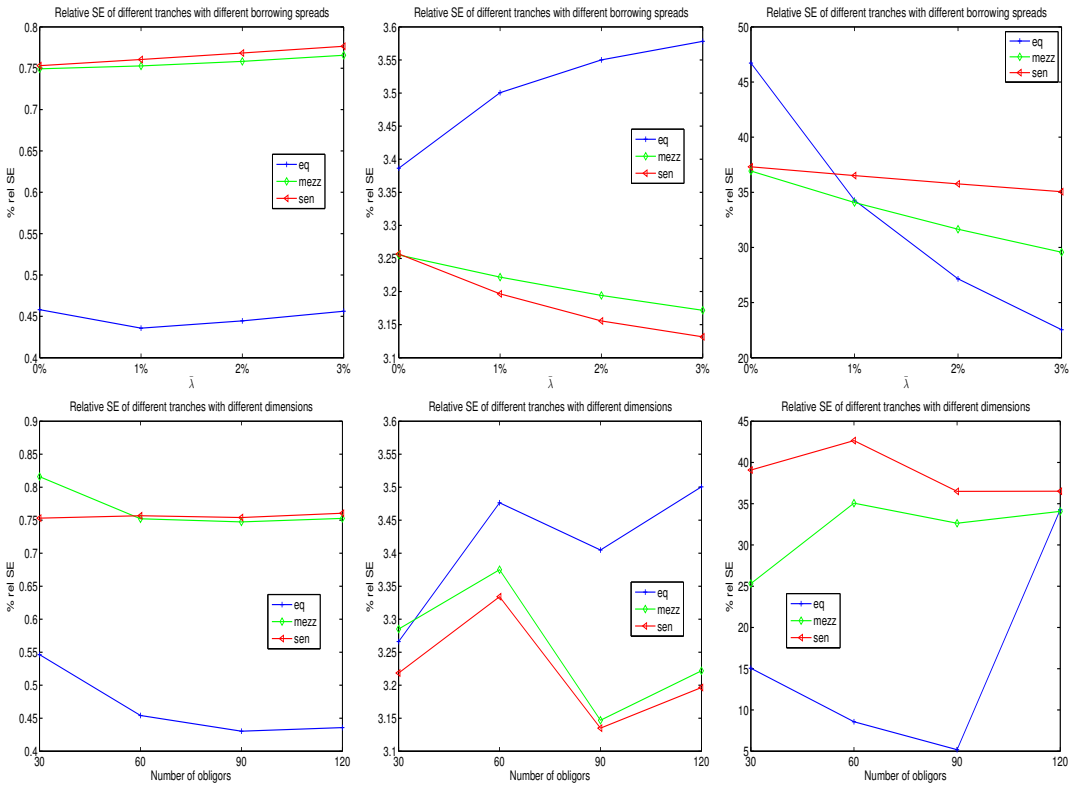

Figure 5.9: *Bottom*: The % relative standard errors of the different tranches don't explode with the number of names ($\bar{\lambda} = 100$ bp). *top*: The % relative standard errors of the different tranches don't explode with the level of nonlinearity represented by the unsecured borrowing spread $\bar{\lambda}$ ($n = 120$). *Left*: FT scheme. *Middle*: $\widetilde{\text{PHL}}$ scheme. *Right*: PHL scheme.

setups, the fully reduced BSDE (III) is stated with respect to a modified probability measure $\mathbb{P}$. Even though switching from $(\mathcal{G}, \mathbb{Q})$ to $(\mathcal{F}, \mathbb{P})$ is transparent in terms of the generator of related Markov factor processes, this can be an issue in situations where the Markov structure is important in the theory to guarantee the validity of the numerical schemes, but is not really practical from an implementation point of view. This is for instance the case with the credit portfolio models that we use for illustrative purposes in our numerics, where the Markov structure that emerges from the dynamic copula model is too heavy and it's only the copula features that can be used in the numerics—copula features under the original stochastic basis $(\mathcal{G}, \mathbb{Q})$, which do not necessarily hold under a reduced basis $(\mathcal{F}, \mathbb{P})$ (especially when $\mathbb{P} \neq \mathbb{Q}$). As for the partially reduced BSDE (II), as compared with the full BSDE (I), its interest is its null terminal condition, which is key for the FT scheme as recalled below. But of course (II) can only be used when one has an explicit formula for γ.

For nonlinear and very high-dimensional problems such as counterparty risk on credit derivatives, the only feasible numerical schemes are purely forward simulation schemes, such as the linear Monte Carlo expansion of Fujii and Takahashi (2012a,2012b) or the branching particles scheme of Henry-Labordère (2012), respectively dubbed "FT scheme" and "PHL scheme" in the chapter. In our setup, the PHL scheme involves a nontrivial and rather sensitive fine-tuning for finding a polynomial in ϑ that approximates the terms $(P_t - \vartheta)^{\pm}$ in $fva_t(\vartheta)$ in a suitable range for ϑ. This fine-tuning requires a preliminary knowledge on the solution obtained by running another approximation (linear approximation or FT scheme) in the first place. Another limitation of the PHL scheme in our case is that it is more demanding than the FT scheme in terms of the structural model properties that it requires. Namely, in our credit portfolio problems, both a Markov structure and a dynamic copula are required for the PHL scheme. But, whereas a "weak" dynamic copula structure in the sense of simulation and forward pricing by copula means is sufficient for the FT scheme, a dynamic copula in the stronger sense that the copula structure is preserved in the future is required in the case of the PHL scheme. This strong dynamic copula property is satisfied by our common-shock model but not in the Gaussian copula model. In conclusion, the FT schemes applied to the partially or fully reduced BSDEs (II) or (III) (a null terminal condition is required so that the full BSDE (I) is not eligible for this scheme) appears as the method of choice on these problems.

An important message of the numerics is that, even for realistically high levels of nonlinearity, i.e. an unsecured borrowing spread $\bar{\lambda} = 3\%$, the third order FT correction was always found negligible and the second order FT correction less than 5% to 10% of the first order, linear FT term. In conclusion, a first order FT term can be used for obtaining "the best linear approximation" to our problem, whereas a nonlinear correction, if wished, can be computed by a second order FT term.

Bibliography

Akahori, J. , Y. Hishida, J. Teichmann and T. Tsuchiya (2014). A Heat Kernel Approach to Interest Rate Models. *Japan Journal of Industrial and Applied Mathematics*, DOI 10.1007/s13160-014-0147-3 .

Applebaum D. (2008). Universel Malliavin calculus in Fock space and Lévy-Itô spaces. Preprint.

Bally, V. and E. Clément (2011). Integration by parts formula with respect to jump times for stochastic differential equations. *Stochastic Analysis 2010*, pp. 7–29. Springer Berlin Heidelberg.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics 24*, 1–13.

Bertoin, J. (1998). *Lévy Processes*. Cambridge Tracts in Mathematics. Cambridge University Press.

Bianchetti, M. (2010). Two curves, one price. *Risk Magazine*, August, 74–80.

Bianchetti, M. and M. Morini (Eds.) (2013). *Interest Rate Modelling After the Financial Crisis*. (Risk Books).

Biane, P. (1990). Chaotic representation for finite Markov chains. *Stochastics and Stochastic Reports 30*, 61–68.

Bichteler, K., J. B. Gravereaux, and J. Jacod (1987). *Malliavin Calculus for Processes with Jumps*. Gordon and Breach Science Publishers.

Bielecki, T. R., D. Brigo, and S. Crépey (2009). Markov chain models of porfolio credit risk. *Oxford Handbook of Credit Derivatives*.

Bielecki, T. R., S. Crépey, and A. Herbertsson (2009). Markov Chain Models of Porfolio Credit Risk. *Oxford Handbook of Credit Derivatives*.

Bielecki, T. R., M. Jeanblanc, and M. Rutkowski (2009). *Credit Risk Modeling*. Osaka University CSFI Lecture Notes Series 2. Osaka University Press.

Bielecki, T. R., A. Vidozzi, and L. Vidozzi (2008b). A Markov copulae approach to pricing and hedging of credit index derivatives and ratings triggered step-up bonds. *Journal of Credit Risk 4*(1), 47–76.

Bouleau N. (1984). *Décomposition de l' Energie par Niveau de Potentiel*. Lect. Notes in M. 1096. Springer.

Bouleau N. (2003). *Error Calculus for Finance and Physics, the Language of Dirichlet Forms*. De Gruyter.

Bouleau N. and L. Denis (2009). Energy image density property and the lent particle method for Poisson measures. *Journal of Functional Analysis 257*, 1144–1174.

Bouleau N. and L. Denis (2015). *Dirichlet Forms Methods for Poisson Point Measures and Lévy Processes*. Springer. Forthcoming.

Bouleau, N. and F. Hirsch (1986). Formes de Dirichlet générales et densité des variables aléatoires réelles sur l'espace de Wiener. *J. Funct. Analysis 69*(2), 229–259.

Bouleau, N. and F. Hirsch (1991). *Dirichlet Forms and Analysis on Wiener Space*. De Gruyter.

Brigo, D. and K. Chourdakis (2008). Counterparty risk for credit default swaps: Impact of spread volatility and default correlation. *International Journal of Theoretical and Applied Finance 12*(7), 1007–1026.

Brigo, D. and F. Mercurio (2006). *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit*. Springer Verlag.

Brigo, D., M. Morini, and A. Pallavicini (2013). *Counterparty Credit Risk, Collateral and Funding: with Pricing Cases for all Asset Classes*. Wiley.

Brody, D. C., and L. P. Hughston (2005). Chaos and coherence: a new framework for interest-rate modelling. *Proceedings of the Royal Society A460*, 2041, 85–110.

Brody, D. C., L. P. Hughston, E. Mackie (2012). General theory of geometric Lévy models for dynamic asset pricing. *Proceedings of the Royal Society A468*, 1778–1798.

Carlen, E.A. and E. Pardoux (1990). Differential calculus and integration by parts on Poisson space. *Stochastics, Algebra and Ananlysis in Classical and Quantum Dynamics*, pp. 63–73. Kluwer.

Cheng, S. and M. R. Tehranchi (2014). Polynomial Models for interest rates and stochastic volatility. *arXiv*, 1404.6190.

Chighouby, F., B. Djehichez, and B. Mezerdix (2009). *The Stochastic Maximum Principle*. Random Operators.

Cont, R. and P. Tankov (2003). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC Press.

Coquio, A. (1993). Formes de Dirichlet sur l'espace canonique de Poisson et application aux équations différentielles stochastiques. *Ann. Inst. Henri Poincaré 19*(1), 1–36.

Crépey, S. (2012a). Bilateral Counterparty risk under funding constraints. Part I: Pricing. *Mathematical Finance 25*(1), 1–22.

Crépey, S. (2012b). Bilateral counterparty risk under funding constraints - Part II: CVA. *Mathematical Finance 25*(1), 23–50.

Crépey, S. (2013). *Financial Modeling: A Backward Stochastic Differential Equations Perspective*. Springer Finance Textbooks. Springer.

Crépey, S., Bielecki, T. R. and D. Brigo (2014). *Counterparty Risk and Funding - A Tale of Two Puzzles*. Chapman and Hall/CRC Financial Mathematics Series.

Crépey, S., R. Gerboud, Z. Grbac, and N. Ngor (2013). Counterparty risk and funding: the four wings of the TVA. *International Journal of Theoretical and Applied Finance 16*(2), 1350006 (31 pages).

Crépey, S., Z. Grbac, N. Ngor, and D. Skovmand (2015). A Lévy HJM multiple-curve model with application to CVA computation. *Quantitative Finance 15*(3).

Crépey, S., M. Jeanblanc, and D. L. Wu (2013). Informationally dynamized Gaussian copula. *International Journal of Theoretical and Applied Finance 16*(2), 1350008 (29 pages).

Crépey, S., A. Macrina, T. M. Nguyen, and D. Skovmand (2015). Rational multi-curve models with counterparty-risk valuation adjustments. Submitted.

Crépey, S. and T. M. Nguyen (2014). Greeks without resimulation in spatially homogeneous Markov Chain Models. In *Proceedings of the Actuarial and Financial Mathematics Conference Interplay between finance and insurance*, M. Vanmaele, G. Deelstra, A. De Schepper, J. Dhaene, W. Schoutens, S.Vanduffel and D. Vyncke (Eds.), VWK, pp.17-28 (2014).

Crépey, S. and T. M. Nguyen (2015). Nonlinear Monte Carlo schemes for counterparty risk on credit derivatives. Working paper.

Crépey, S. and A. Rahal (2014). Simulation/regression pricing schemes for CVA computations on CDO tranches. *Communications in Statistics - Theory and Methods* 43(7), 1390–1408.

Crépey, S. and S. Song (2014a). Invariant times. Submitted.

Crépey, S. and S. Song (2014b). BSDEs of counterparty risk. Forthcoming in *Stochastic Processes and their Applications*.

Crépey, S. and S. Song (2015). Counterparty risk and funding: Immersion and beyond. hal-00989062v2.

Crépey, S. and S. Song (2015b). Invariant times, Markov processes and BSDEs with implicit terminal condition. Working paper.

Cuchiero, C., M. Keller-Ressel, J. Teichmann (2012). Polynomial processes and their applications to mathematical finance. *arXiv* 0812.4740v2.

Cuchiero, C., C. Fontana, A. Gnoatto (2014). A General HJM Framework for multiple yield curve modeling. *arXiv* 1406.4301v1.

Dellacherie C. and P. A. Meyer (1975). *Probabilités et Potentiel* Chap I à IV. Hermann.

Denis, L. (2000). A criterion of density for solutions of Poisson-driven SDEs. *Probab. Theory Relat. Fields 118*, 406–426.

Denis, L. and T. M. Nguyen (2015). Malliavin calculus for Markov chains using perturbations of time. Submitted.

Dermoune, A., P. Krée, and L. Wu (1988). Calcul stochastique non-adapté par rapport à une measure aléatoire de Poisson. *Sém. Proba. XXII*, LNM 1321, 477–484.

Di Nunno, G., B. K. Øksendal, and F. Proske (2008) *Malliavin Calculus for Lévy Processes with Applications to Finance*. Springer-Verlag Berlin Heidelberg.

Döberlein, F. and M. Schweizer (2001). On Savings Accounts in Semimartingale Term Structure Models. *Stochastic Analysis and Applications 19*, 605–626.

Eberlein, E., K. Glau and A. Papapantoleon (2010). Analysis of Fourier transform valuation formulas and applications. *Applied Mathematical Finance 17*, 211–240.

El Karoui, N., S. Peng, and M.-C. Quenez (1997). Backward stochastic differential equations in finance. *Mathematical Finance 7*, 1–71.

El-Khatib, Y. and N. Privault (2004). Computations of Greeks in a market with jumps via the Malliavin calculus. *Finance and Stochastics 8*(2), 161–179.

Elliott, R. J. (1991). Martingales Associated with Finite Markov Chains. *Seminar on Stochastic Processes 1990 - Progress in Probability 24*, 161–172.

Filipović, D. and A. B. Trolle (2013). The term structure of interbank risk. *Journal of Financial Economics 109*, 707–733.

Filipović, D., M. Larsson and A. B. Trolle (2014.) Linear-rational term stucture models. *SSRN:2397898*.

Flesaker, B. and L. P. Hughston (1996). Positive Interest, *Risk Magazine 9*, 46–49.

Fujii, M., Y. Shimada, and A. Takahashi (2010). A note on construction of multiple swap curves with and without collateral. *FSA Research Review 6*, 139–157.

Fujii, M., Y. Shimada, and A. Takahashi (2011), A market model of interest rates with dynamic basis spreads in the presence of collateral and multiple currencies. *Wilmott Magazine 54*, 61–73.

Fujii, M. and A. Takahashi (2012a). Analytical approximation for non-linear FBSDEs with perturbation scheme. *International Journal of Theoretical and Applied Finance 15*(5), 1250034(24).

Fujii, M. and A. Takahashi (2012b). Perturbative expansion of FBSDE in an incomplete market with stochastic volatility. *Quarterly Journal of Finance 2*(3), 1250015(24).

He, S.-W., J.-G. Wang, and J.-A. Yan (1992). *Semimartingale Theory and Stochastic Calculus*. CRC Press.

Heath, D., R. Jarrow, A. Morton (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica 60*, 77–105.

Henrard, M. (2007). The irony in the derivatives discounting. *Wilmott Magazine 30*, 92–98.

Henrard, M. (2010). The irony in the derivatives discounting part II: the crisis. *Wilmott Magazine 2*, 301–316.

Henrard, M. (2014). *Interest Rate Modelling in the Multi-curve Framework*. (Palgrave Macmillan).

Henry-Labordère, P. (2012). Cutting CVA's complexity. *Risk Magazine* (July), 67–73.

Hoyle, E., L. P. Hughston, A. Macrina (2011). Lévy Random Bridges and the Modelling of Financial Information. *Stochastic Processes and their Applications 121*, 856–884.

Hughston, L. P. and A. Rafailidis (2005). A chaotic approach to interest rate modelling. *Finance and Stochastics 9*, 43-65.

Hull, J. C., A. Sokol and A. White (2014). Modelling the short rate: the real and risk-neutral worlds. *http://dx.doi.org/10.2139/ssrn.2403067*.

Hunt, P. and J. Kennedy (2004). *Financial Derivatives in Theory and Practice, revised edition.*(Wiley).

Ishikawa, Y. and H. Kunita (2006). Malliavin calculus on the Wiener-Poisson space and its application to canonical SDE with jumps. *Stoch. Processes and their App. 116*, 1743–1769.

Itô, K. (1951). Multiple Wiener Integral. *Journal of the Mathematical Society of Japan 3*(1), 157–169.

Itô, K. (1956). Spectral type of the shift transformation of differential processes with stationary increments. *JTrans. Amer. Math. 81*, 253–263.

Jacod, J. (1979). *Calcul Stochastique et Problème de Martingales*. Lecture Notes in Math 714. Springer.

Kenyon, C. (2010). Short-rate pricing after the liquidity and credit shocks: including the basis. *Risk Magazine*, November, 83–87.

Kijima, M., K. Tanaka and T. Wong (2009). A multi-quality model of interest rates. *Quantitative Finance 9*(2), 133–145.

Kruse, T. and A. Popier (2014). BSDEs with jumps in a general filtration. *arXiv:1412.4622.*

Li, D. (2000). On default correlation: A copula function approach. *Journal of Fixed Income 9*(4), 43–54.

Longstaff, F. A. and E. S. Schwartz (2001). Valuing American options by simulations: a simple least-squares approach, *Review of Financial Studies 14*(1), 113–147.

Ma and M. Röckner (1992). *Introduction to the Theory of (Non-Symmetric) Dirichlet Forms*. Springer.

Ma and M. Röckner (2000). Construction of diffusion on configuration spaces. *Osaka J. Math. 37*, 273–314.

Macrina, A. (2014). Heat Kernel Models for Asset Pricing. *International Journal of Theoretical and Applied Finance 17* (7), 1-34.

Macrina, A. and P. A. Parbhoo (2014). Randomised Mixture Models for Pricing Kernels. *Asia-Pacific Financial Markets 21*, 281-315.

Malliavin, P. (1997). *Stochastic Analysis*. Springer.

Marshall, A. W. and I. Olkin (1967). A multivariate exponential distribution. *Journal of the American Statistical Association 2*, 84–98.

McKean, H. P. (1975). Application of Brownian motion to the equation of Kolmogorov-Petrovskii-Piskunov. *Communications on Pure an Applied Mathematics 28*(3) May, 323–331.

Mercurio, F. (2010). A LIBOR market model with stochastic basis. *Risk Magazine*, December, 84–89.

Mercurio, F. (2010). Interest rates and the credit crunch: new formulas and market models. *Technical report, Bloomberg Portfolio Research Paper No. 2010-01-FRONTIERS*.

Mercurio, F. (2010). Modern LIBOR Market Models: Using Different Curves for Projecting Rates and for Discounting. *International Journal of Theoretical and Applied Finance 13*, 113-137.

Moreni, N. and A. Pallavicini (2013). Parsimonious multi-curve HJM modelling with stochastic volatility. In M. Bianchetti and M. Morini (Eds.), *Interest Rate Modelling After the Financial Crisis*, (Risk Books).

Moreni, N. and A. Pallavicini (2014). Parsimonious HJM modelling for multiple yield-curve dynamics. *Quantitative Finance 14*(2), 199–210.

Nguyen, T. and F. Seifried (2014). The multi-curve potential model. http://ssrn.com/abstract=2502374.

Nourdin, I. and T. Simon (2006). On the absolute continuity of Lévy processes with drift. *Ann. Probab. 34*, 1035–1051.

Nualart, D. (1995). *Malliavin Calculus and Related Topics*. Springer.

Nualart, D. (2010). *PDEs: Regularity of the probability law of the solution*. Univ. of Texas at Austin.

Nualart, D. and L. Quer-Sardanyons (2007). Existence and smoothness of the density for spatially homo- geneous SPDEs. *Potential Anal. 27*, 281–299.

Nualart, D. and J. Vives (1990). Anticipative calculus for the Poisson process based on the Fock space. *Sém. Prob. XXIV*. Lect. Notes in M. 1426. Springer.

Parbhoo, P. A (2013). Information-driven pricing kernel models. PhD Thesis (University of the Witwatersrand).

Picard, J. (1996). On the existence of smooth densities for jump processes. *Probab. Theory Relat. Fields 105* , 481–511.

Privault, N. (2009). *Stochastic Analysis in Discrete and Continuous Settings with Normal Martingales*. Springer.

Rogers, L. C. G. (1997). The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates. *Mathematical Finance 7*, 157-176.

Rutkowski, M. (1997). A note on the Flesaker-Hughston model of the term structure of interest rates. *Applied Mathematical Finance 4*, 151-163.

Shigekawa, I. (2004). *Stochastic Analysis*. AMS.

Siu, T. K. (2014). Integration by parts and martingale representation for a Markov Chain. *Abstract and Applied Analysis*, 1–11. Hindawi Publishing Corporation

Tsitsiklis, J.N. and B. Van Roy. (2001). Regression methods for pricing complex American-style options. *IEEE Transactions on Neural Networks 12*, 694–703.

Watanabe, S. (1987). Analysis on Wiener functionals (Malliavin calculus) and its applications to heat kernels. *The Annals of Prob. 15*(1), 1–39.