Alma Mater Studiorum · Università di Bologna

DOTTORATO DI RICERCA IN MATEMATICA
Ciclo XXXI

Settore Concorsuale: 01/A3
Settore Scientifico Disciplinare: MAT 06

APPLICATIONS OF STOCHASTIC PROCESSES TO FINANCIAL RISK COMPUTATION

Presentata da: Anastasia Borovykh

Coordinatore Dottorato:
Chiar.ma Prof.ssa
Giovanna Citti

Supervisore:
Chiar.mo Prof.
Andrea Pascucci

Esame finale anno 2019

Acknowledgements

A big thank you to my supervisor Prof. Andrea Pascucci for providing an overall great guidance and helpful suggestions in all the topics we worked on, for allowing me to explore a wide variety of applications in risk management, and giving me the opportunity to present my work at numerous conferences; to Prof. Cornelis W. Oosterlee for all his help on and insights into the different topics we worked on, for giving me the opportunity to branch out to machine learning during my period at the CWI and for his overall guidance and support throughout the years, starting from my bachelor's degree and continuing to the present day; to Prof. Sander Bohté for introducing me to machine learning and for his great help and expertise in the projects we worked on; to Pietro Rossi and his colleagues at Prometeia for giving me the opportunity to present my work in industry, for introducing me to Prometeia and for the always interesting discussions; to Stefano la Rovere and his colleages at NIER Ingegneria for giving me the opportunity to understand the applications of my work in industry; and to the staff at the University of Bologna, in particular Lucia Vichi, Oscar Lo Surdo and Giovanna Citti for all their help with the administrative work. Last but not least, a million times thanks to my family and friends for their overall greatness, and in particular to my parents and grandparents for their endless encouragement, help and support.

List of commonly used abbreviations

GBM:	Geometric Brownian motion
EMM:	Equivalent martingale measure
CVA:	Credit valuation sdjustment
DVA:	Debt valuation adjustment
CCR:	Counterparty credit risk
LSM:	Least-squares Monte-Carlo method
MVA:	Margin valuation adjustment
CSA:	Credit support annex
FVA:	Funding value adjustment
KVA:	Capital value adjustment
PDE:	Partial differential equation
BSDE:	Backward Stochastic Differential equation
XVA:	Collective name for the valuation adjustments
CLT:	Central limit theorem
LLN:	Law of large numbers
FFT:	Fast Fourier transform
COS:	The Fourier-based method for option pricing
BCOS:	The Fourier-based method for BSDEs
HP:	Hawkes process
SDE:	Stochastic differential equation
RCLL:	Right-continuous with left limits

Contents

Summary

Ever since the financial crisis the focus on having efficient analytic and numerical methods in the field of financial risk computations has increased significantly. New regulations have been invoked that force banks and other financial entities to much more carefully monitor the various risks involved in their daily practices. One area in which the regulations have increased significantly has been that of the so-called valuation adjustments in derivative pricing. In particular, under the most recent regulatory framework Basel III banks are required to hold a particular amount of capital on their balance sheets. This amount of capital is determined by looking at the various portfolios of the bank and computing the risks involved in holding and trading those portfolios; in other words banks are now required to price *all* components of a trade. These additional factors are collectively called valuation adjustments. The analysis and valuation of these adjustments is crucial to banks, but in turn is also a complex task involving both accounting methodologies as well as the need for efficient mathematical methods.

A Fourier-based method such as the COS method has been both an efficient and accurate method for pricing derivatives, by making use of the characteristic function, i.e. the Fourier transform of the density. For flexible models that are able to incorporate many of the aspects of current market dynamics, e.g. a stochastic jump-intensity or volatility smiles and skews, there is no availability of an explicit expression of the density or characteristic function. One way of obtaining an approximation to the characteristic function is to make use of a Taylor expansion of the generator of the process. This allows to split the Cauchy problem for the density into simplified Cauchy problems and solve these explicitly in the Fourier space. In order to price derivatives with or without valuation adjustments, the arising (non-linear) partial differential equation can then be solved by means of a combination of the COS method and the approximated characteristic function, resulting in an efficient and easy-to-implement valuation method.

Another risk metric related to counterparty credit risk, whose importance has increased since the crisis is that of systemic risk. Systemic risk was an important contributor to the financial crisis, where the collapse of individual financial entities triggered a chain of defaults throughout the system. Carefully monitoring the banks that are most prone to triggering a large loss in the system after experiencing a loss themselves is of the essence in the prevention of such events. Due to the system being large with each individual entity having complex interactions with many other entities, methods for computing the risk in such a system are not trivial and again require the need for efficient mathematical models. Hawkes processes are able to incorporate an important feature of risk in interconnected systems, in particular that of a cross- and self-excitement in the monetary reserves of the banks. While modeling the full multivariate system in case of a large number of banks is time-consuming, using a weak convergence analysis in which the number of entities in the system tends to infinity allows us to obtain an expression for the behavior of the monetary reserve process in a large system, and quantify the systemic risk present in the system.

Chapter 1

Introduction

1.1 Background

This thesis deals with the applications of stochastic processes in finance, in particular focusing on risk computation. In the field of financial risk computations often efficient analytic and numerical methods are required to compute various risk measures. Ever since the global financial crisis the focus on having these kinds of methods has increased significantly, with new regulations in place that force banks and other financial entities to much more carefully monitor the various risks involved in their daily practices. Here we deal with two main subsets of risk measures, the first one being related to derivative valuation, the second one being risk computation in financial systems.

Both efficiency and accuracy are of the essence when valuing and risk-managing financial derivatives and portfolios comprising of these derivatives. For risk management purposes traders can be restricted to hold a particular maximum amount of risk in their portfolios, forcing the trader to efficiently and accurately compute both the value of the portfolio for calibration purposes as well as its sensitivities in order to quantify and minimize the portfolio risk. The price of an option without early-exercise features under the risk-neutral measure is given by the expected value of the discounted payoff of this option.

Under the most recent regulatory framework Basel III banks are required to hold a particular amount of capital on their balance sheets. This amount of capital is determined by looking at the various portfolios of the bank and computing the risks involved in holding and trading those portfolios; in other words banks are now required to price *all* components of a trade, and not just the option value. These additional factors are collectively called valuation adjustments. The most widely known valuation adjustment is the Credit Valuation Adjustment, or CVA, and it captures the risk that the counterparty in a particular deal is prone to default. One of the latest significant adjustments is the Funding Valuation Adjustment (FVA) capturing the funding consists of the cash needed to enter and hold the portfolios. The analysis and valuation of these adjustments is important for banks, but in turn it can also be a complex task involving both accounting methodologies as well as the need for efficient mathematical methods.

The last risk measure we consider is known as systemic risk. Related to counterparty credit risk, it is the risk that an event at bank level can cause further instability of the entire financial system. In financial systems, and in particular in large, *interconnected* financial systems which are prone to default propagation, the ability to measure the stability of the system is of great importance. Systemic risk was an important contributor to the global financial crisis, where the collapse of Lehman brothers, a bank with many connections to other banks and investors through e.g. loans, triggered a chain of defaults throughout the system [39]. In an interconnected banking system carefully monitoring the banks that are most prone to triggering a large loss at its creditors after experiencing a loss themselves is of the essence in the prevention of such systemic events. Due to these financial systems often being large with each individual bank having complex interactions with other banks (through e.g. loans, or common balance sheet holdings), methods for computing risk in such systems are not trivial, and again require the need for efficient mathematical models.

In the next chapters we will discuss several novel ways of valuing the above-mentioned risks i.e. derivative pricing and hedging, valuation adjustments and systemic risk. In particular our methods will be motivated by the real-world dynamics, and include significant improvements over the current state-of-the-art methodologies. The main tools for computing the above mentioned risks will be by means of stochastic differential equations, the solution of partial differential equations and weak convergence analysis. In the upcoming sections of this chapter we will introduce the main concepts that will be employed in the coming chapters, as well as provide a more detailed definition of the risks we will consider.

1.2 Stochastic processes

In this section we will briefly discuss the notion of stochastic processes as we will use it in the financial models in the coming chapters. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and I a real interval, also called the index set, of the form $\mathbb{R}_+$. A measurable stochastic process on the state space $\mathbb{R}^N$ is a collection of $(X_t)_{t\in I}$ of random variables with values in $\mathbb{R}^N$ such that the map

$$X : I \times \Omega \to \mathbb{R}^N, \quad X(t, \omega) = X_t(\omega),$$

is measurable with respect to the product σ-algebra $\mathcal{B}(I) \otimes \mathcal{F}$. The stochastic process associates for each $t \in I$, the random variable X_t in $\mathbb{R}^N$. For any point $\omega \in \Omega$, the mapping

$$X_t(\omega) : I \to \mathbb{R}^N,$$

is called a sample function, or when we interpret the index set I as time, a sample path of the stochastic process. In our setting the stochastic processes describe the evolution of a random phenomenon in time.

A stochastic process is called Markov if the conditional probability distribution of future states depends only on the present state, and not on the full history that preceded it.

Definition 1.2.1 (Markov process)**.** Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$ be the filtered probability space. The adapted stochastic process X has the Markov property if for every bounded $\mathcal{B}$-measurable function $\varphi : \mathbb{R}^N \to \mathbb{R}$ we have

$$\mathbb{E}[\varphi(X_T)|\mathcal{F}_t] = \mathbb{E}[\varphi(X_T)|X_t], \quad T \geq t.$$

In a financial setting with the stochastic process representing the asset price the Markov property translates to the weak form of the Efficient Market Hypothesis. In other words, the current stock price contains all the information of the past, motivating the random walk model for stock returns.

A stochastic process is said to be a martingale if the expected value of a future state is given by the present one.

Definition 1.2.2 (Martingale)**.** Let M be an integrable adapted stochastic process on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t))$. We say that M is a martingale with respect to the filtration $\mathcal{F}_t$ and to the measure $\mathbb{P}$ if

$$M_s = \mathbb{E}[M_t|\mathcal{F}_s], \quad \text{for every } 0 \leq s \leq t.$$

We will next introduce a common stochastic process used for modelling asset dynamics: the Lévy process, of which the geometric Brownian motion is a well-known example. Furthermore, we define the local Lévy process, an extension of the Lévy process, which is the particular stochastic process considered in the option valuation applications in this thesis.

1.2.1 Exponential Lévy processes

With exponential Lévy processes the asset price is modeled as an exponential function of a Lévy process X_t,

$$S_t = S_0 e^{X_t}.$$

The class of Lévy processes includes the Brownian motion and Poisson processes and preserves the property of independence and stationarity of the increments. The Brownian motion is the only Lévy process with continuous increments; on the other hand, the presence of jumps is one of the main motivations that has led to consider Lévy processes for modelling asset dynamics. We start with recalling the definition: an adapted stochastic process X_t on $(\Omega, \mathcal{F}, \mathbb{P})$ with $X_0 = 0$ a.s., is a Lévy process if

1. X has increments independent of the past, that is $X_t - X_s$ is independent of $\mathcal{F}_s$ for $0 \leq s \leq t$,

2. it has stationary increments, that is $X_t - X_s$ has the same distribution as X_{t-s} for $0 \leq s \leq t$,

3. it is stochastically continuous, that is for any $t \geq 0$ and $\epsilon > 0$, we have

$$\lim_{h \to 0} \mathbb{P}(|X_{t+h} - X_t| \geq \epsilon) = 0.$$

A simple and widely used exponential Lévy process is the Geometric Brownian Motion (GBM) model, whereby the logarithm of the asset price follows a Brownian motion. The asset price S_t has GBM dynamics if it satisfies the following stochastic differential equation,

$$dS_t = \mu S_t dt + \sigma S_t dW_t,$$

where W_t is the Brownian motion, μ is the so-called drift paramter, and σ is the volatility parameter. Under the GBM model, there exists the well-known Black-Scholes formula giving the price of European put and call options.

A remarkable property of the Lévy model is the availability of the explicit form of the characteristic function. The characteristic function is defined as

$$\hat{\Gamma}(t, x; T, \xi) = \mathcal{F}(\Gamma(t, x; T, y))(\xi) = \mathbb{E}[e^{i\xi X_t}],$$

where $\mathcal{F}(\cdot)$ denotes the Fourier transform with respect to the second set of variables (T, y), and equivalently the expected value is taken with respect to the density of X_t. Then we have

Theorem 1.2.3. *If X is a Lévy process, then there exists a unique function $\psi \in C(\mathbb{R}^d, \mathbb{C})$, the space of continuous functions from $\mathbb{R}^d$ into the space of complex numbers $\mathbb{C}$ such that $\psi(0) = 0$ and*

$$\hat{\Gamma}(t, x; T, \xi) = e^{t\psi(\xi)}, \quad t \geq 0, \ \xi \in \mathbb{R}^d.$$

The function ψ is called the characteristic exponent of X.

Since the distribution of a random variable is determined by its characteristic function, a consequence of the above it that the law of the Lévy process is fully determined by its characteristic exponent, or equivalently the distribution of X_t for a single time.

Let us consider several examples.

Example 1.2.4 (Brownian motion with drift)**.** Let $X_t = \mu t + \sigma W_t$, where W is a standard real Brownian motion. We can then find

$$\mathbb{E}[e^{i\xi X_t}] = e^{i\mu t\xi} \mathbb{E}\left[e^{i\xi\sigma W_t}\right] = e^{i\mu t\xi + \frac{1}{2}(i\xi\sigma)^2 t},$$

so that

$$\psi(\xi) = i\mu\xi - \frac{\sigma^2\xi^2}{2},$$

Example 1.2.5 (Jump-diffusion process)**.** Consider the jump-diffusion

$$X_t = \mu t + \sigma W_t + \sum_{n=1}^{N_t} Z_n,$$

where N_t represents a Poisson counting process with intensity λ. Assume the distribution of the Z_n, $\eta(x)$, follows a normal distribution, i.e. $Z_n \sim \mathcal{N}(\alpha, \beta^2)$ for all n. Then we can find

$$\begin{aligned}\mathbb{E}\left[e^{i\xi Z_1}\right] &= \int_{\mathbb{R}} e^{i\xi x}\eta(x)dx \\ &= e^{i\alpha\xi - \frac{1}{2}\beta^2\xi^2}.\end{aligned} \tag{1.1}$$

We have

$$\begin{aligned}\mathbb{E}\left[e^{i\zeta X_t}\right] &= e^{i\mu\xi t - \frac{1}{2}\sigma^2\zeta^2 t}\sum_{n>0}\mathbb{E}\left[e^{i\xi\sum_{k=1}^{n} Z_k}\mathbb{1}_{\{N_t=n\}}\right] \\ &= e^{i\mu\zeta - \frac{1}{2}\sigma^2\xi^2}\sum_{n>0}\mathbb{E}\left[e^{i\xi Z_1}\right]^2\mathbb{P}(N_t = n) \\ &= e^{i\mu\zeta - \frac{1}{2}\sigma^2\xi^2}e^{-\lambda t}\sum_{n>0}\frac{(\lambda t\hat{\eta}(\xi))^n}{n!} = e^{i\mu\zeta - \frac{1}{2}\sigma^2\xi^2}e^{-\lambda(1+\hat{\eta}(\xi))t},\end{aligned}$$

where $\hat{\eta}(\xi)$ is the characteristic function of $\eta(x)$. We have used the independence of $N_t, Z_1, ..., Z_n$ and the definition of the Poisson distribution. Using (1.1), we then obtain

$$\psi(\zeta) = i\mu\xi - \frac{1}{2}\sigma^2\zeta^2 + \lambda\left(e^{i\alpha\xi - \frac{1}{2}\beta^2\xi^2} - 1\right).$$

1.2.2 Local Lévy processes

Several problems arise when modelling the asset price using the geometric Brownian motion. First of all, the GBM is not able to reproduce the volatility skew or smile present in most financial markets, arising when trying to fit the Black-Scholes prices to the observed market prices. Furthermore, under the GBM model the paths of the asset prices are continuous functions of time. Empirical evidence however has shown that asset prices tend to contain jumps, appearing as discontinuities in the price path (see e.g. [21]). Lastly, it has been widely recognized that the empirical distribution of stock returns is not Gaussian, but tends to possess both skewness and heavy tails. These observations are the main motivation for practitioners to work with more general and flexible processes, a popular class of which are Lévy processes. The particular model we will consider here is what we call to be the local Lévy model, in which the term ‘local’ stems from the fact that we allow the coefficients to be dependent on the underlying process itself. We consider a defaultable asset S whose risk-neutral dynamics are given by:

$$S_t = \mathbb{1}_{\{t<\zeta\}}e^{X_t},$$

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t + \int_{\mathbb{R}} z d\tilde{N}_t(t, X_{t-}, dz),$$
$$d\tilde{N}_t(t, X_{t-}, dz) = dN_t(t, X_{t-}, dz) - \nu(t, X_{t-}, dz)dt,$$
$$\zeta = \inf\{t \geq 0 : \int_0^t \gamma(s, X_s)ds \geq \varepsilon\}, \tag{1.2}$$

where $\tilde{N}_t(t, x, dz)$ is a compensated random measure with state-dependent Lévy measure $\nu(t, x, dz)$. The default time ζ of S is defined in a canonical way as the first arrival time of a doubly stochastic Poisson process with local intensity function $\gamma(t, x) \geq 0$, and $\varepsilon \sim \text{Exp}(1)$ and is independent of X. Thus the model features,

- a local volatility function $\sigma(t, x)$: A local volatility function allows one to model the volatility smile or skew as observed in financial market;
- a local Lévy measure: Jumps in X arrive with a state-dependent intensity described by the local Lévy measure $\nu(t, x, dz)$. The jump intensity and jump distribution can thus change depending on the value of x. A state-dependent Lévy measure is an important feature because it allows to incorporate stochastic jump-intensity into the modeling framework;
- a local default intensity $\gamma(t, x)$: The asset S can default with a state-dependent default intensity.

This way of modeling default is also considered in a diffusive setting in [17] and for exponential Lévy models in [15].

When working with the above model we define the filtration of the market observer to be $\mathcal{G} = \mathcal{F}^X \vee \mathcal{F}^D$, where $\mathcal{F}^X$ is the filtration generated by X and $\mathcal{F}^D_t := \sigma(\{\zeta \leq u\}, u \leq t)$, for $t \geq 0$, is the filtration of the default. We assume

$$\int_{\mathbb{R}} e^{|z|}\nu(t, x, dz) < \infty.$$

If we impose that the discounted asset price $\tilde{S}_t := e^{-rt}S_t$ is a $\mathcal{G}$-martingale, we get the following restriction on the drift coefficient:

$$\mu(t, x) = \gamma(t, x) + r - \frac{\sigma^2(t, x)}{2} - \int_{\mathbb{R}} \nu(t, x, dz)(e^z - 1 - z).$$

1.3 Option pricing

In financial mathematics, the fast and accurate pricing of financial derivatives is an important branch of research. Depending on the type of financial derivative, the mathematical task is essentially the computation of integrals, and this sometimes needs to be performed in a recursive way in a time-wise direction, e.g. in the case of Bermudan options. Briefly, an option is an agreement between two parties which offers the buyer the right, but not the obligation, to buy or sell an underlying security at an agreed-upon price and at or upto an agreed-upon date. The price at which one can buy or sell is known as the strike price K, the expiration date as the maturity T and in case of the right to buy, respectively sell, the option is a called a call, respectively put. If the holder of the option can choose to exercise the option only at the maturity time T, the option is called a European option. If there is an infinite amount of exercise moments upto maturity it is an American option, while with a finite number of set dates upto maturity the option is of the Bermudan type.

Here we introduce the mathematical concept of pricing European options, in particular focusing on the Fourier-transform method known as the COS method. In a complete market the price of an option is said to be given by the discounted risk-neutral expectation of future payoffs. In particular, consider a stochastic process $\{X_t, t \in [0,T]\}$ defined on the usual probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ and governed by a stochastic differential equation of the Lévy type. The corresponding bank account evolves according to $dB_t = r_t B_t dt$, with r being the (for now deterministic) risk-free rate; B_t is thus the corresponding numeraire. Note that we work under the risk-neutral probability measure $\mathbb{Q}$, also known as the equivalent martingale measure (EMM). For completeness, the EMM is defined as follows:

Definition 1.3.1 (Equivalent martingale measure)**.** An equivalent martingale measure $\mathbb{Q}$ with numeraire $B_t := e^{\int_0^t r_s ds}$ is a probability measure on a measurable space $(\Omega, \mathcal{F})$ such that

1. $\mathbb{Q}$ is equivalent to the real-world measure $\mathbb{P}$,
2. the process of discounted prices $\tilde{S}_t = B_t^{-1} S_t$, is a strict $\mathbb{Q}$-martingale. In particular, the risk-neutral pricing formula

$$S_t = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r_s ds} S_T | \mathcal{F}_t \right], \quad t \in [0,T],$$

   holds.

Then, according to the risk-neutral valuation formula the price of a European option can be written as the expectation of the discounted (with the risk-free bank-account) payoff of this option

$$v(t, S_t) = e^{-rt} \mathbb{E}[\varphi(S_T)],$$

where to shorten notation we suppress the dependence of the expectation on the risk-neutral measure, v denotes the value of the option, t is the current time point, T is the maturity and $\varphi(S_T)$ will be used to denote the maturity payoff of the option. When computing option prices the evaluation of expectations of the type shown above is of the essence. It can be calculated via numerical integration or even analytical methods provided that the underlying density is known in closed form, which unfortunately is not the case for many models. What *is* commonly available is the characteristic function; in particular for Lévy models we have a closed-form expression (see Section 1.2.1). The characteristic function is the continuous Fourier transform of the density function, and can be used for obtaining the above expected value when transitioning into the Fourier domain. A method which makes use of this transformation is known as the COS method [30]. The COS method proposed by [30] is based on the insight that the Fourier-cosine series coefficients of the density $\Gamma(t,x;T,dy)$ (and therefore also of option prices) are closely related to the characteristic function of the underlying process, namely the following relationship holds:

$$\int_a^b e^{i\frac{k\pi}{b-a}}\Gamma(t,x;T,dy) \approx \hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right).$$

The COS method provides a way to calculating expected values (integrals) of the form

$$v(t,x) = \int_{\mathbb{R}} \varphi(T,y)\Gamma(t,x;T,dy),$$

and it consists of three approximation steps:

1. In the first step we truncate the infinite integration range to $[a,b]$ to obtain approximation v_1:

   $$v_1(t,x) := \int_a^b \varphi(T,y)\Gamma(t,x;T,dy).$$

   We assume this can be done due to the rapid decay of the distribution at infinity.

2. In the second step we replace the distribution with its cosine expansion and we get

   $$v_1(t,x) := \frac{b-a}{2}\sum_{k=0}^{\infty}{}' A_k(t,x;T)V_k(T),$$

   where $\sum'$ indicates that the first term in the summation is weighted by one-half and

   $$A_k(t,x;T) = \frac{2}{b-a}\int_a^b \cos\left(k\pi\frac{y-a}{b-a}\right)\Gamma(t,x;T,dy),$$
   $$V_k(T) = \frac{2}{b-a}\int_a^b \cos\left(k\pi\frac{y-a}{b-a}\right)\varphi(T,y)dy,$$

   are the Fourier-cosine series coefficients of the distribution and of the payoff function at time T respectively. Due to the rapid decay of the Fourier-cosine series coefficients, we truncate the series summation and obtain approximation v_2:

   $$v_2(t,x) := \frac{b-a}{2}\sum_{k=0}^{N-1}{}' A_k(t,x;T)V_k(T).$$

3. In the third step we use the fact that the coefficients A_k can be rewritten using the truncated characteristic function:

$$A_k(t,x;T) = \frac{2}{b-a}\mathrm{Re}\left(e^{-ik\pi\frac{a}{b-a}}\int_a^b e^{i\frac{k\pi}{b-a}y}\Gamma(t,x;T,dy)\right),$$

where $\mathrm{Re}(\cdot)$ denotes taking the real part of the input argument. The finite integration range can be approximated as

$$\int_a^b e^{i\frac{k\pi}{b-a}y}\Gamma(t,x;T,dy) \approx \int_{\mathbb{R}} e^{i\frac{k\pi}{b-a}y}\Gamma(t,x;T,dy) = \hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right).$$

Thus in the last step we replace A_k by its approximation:

$$\frac{2}{b-a}\mathrm{Re}\left(e^{-ik\pi\frac{a}{b-a}}\hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right)\right),$$

and obtain approximation v_3:

$$v_3(t,x) := \sum_{k=0}^{N-1}{}' \mathrm{Re}\left(e^{-ik\pi\frac{a}{b-a}}\hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right)\right)V_k(T). \tag{1.3}$$

In other words, *if* the characteristic function is available, the COS method allows us to compute the option value by transitioning into the Fourier domain.

1.3.1 Option pricing under the local Lévy model

When considering the local Lévy model as defined in Section 1.2.2, the computation of the above expected value is no longer trivial as there is no explicit formula for the density nor for the characteristic function. In Chapter 2 we discuss how one can use a so-called adjoint expansion method in order to compute these option prices, both for a European as well as a Bermudan derivative. Here, we briefly address the unavailability of explicit density and characteristic function expressions. Consider the asset dynamics for S_t as defined in (3.16) satisfying the martingale condition. The European option price with maturity T and payoff $\varphi(S_T)$ is then given by,

$$\begin{aligned}
v(t,x) &= \mathbb{E}\left[e^{r(T-t)}\varphi(S_T)|\mathcal{G}_t\right] \\
&= e^{r(T-t)}\mathbb{E}\left[\varphi(X_T)\mathbb{1}_{\{\zeta>T\}}|\mathcal{G}_t\right] + e^{r(T-t)}K\mathbb{E}\left[\mathbb{1}_{\{\zeta\le T\}}|\mathcal{G}_t\right] \\
&= e^{r(T-t)}\mathbb{E}\left[\varphi(X_T)\mathbb{1}_{\{\zeta>T\}}|\mathcal{G}_t\right] + e^{r(T-t)}K - e^{r(T-t)}K\mathbb{E}\left[\mathbb{1}_{\{\zeta>T\}}|\mathcal{G}_t\right] \\
&= e^{r(T-t)}K + e^{r(T-t)}\mathbb{1}_{\{\zeta>t\}}\mathbb{E}\left[e^{-\int_t^T\gamma(s,X_s)ds}(\varphi(X_T)-K)|\mathcal{G}_t\right],
\end{aligned} \tag{1.4}$$

with $\varphi(x) := \varphi(e^x)$, and $K := \varphi(0)$ and where we have used the result in [48, Section 2.2]. For $K=0$, we are thus interested in computing functions of the form

$$u(t,x) = \mathbb{E}\left[e^{-\int_t^T\gamma(s,X_s)ds}\varphi(X_T)|\mathcal{G}_t\right].$$

By a direct application of the Feynman-Kac representation theorem, $u(t,x)$ is the classical solution of the following Cauchy problem

$$\begin{cases} Lu(t,x) = 0, & t \in [0,T[,\ x \in \mathbb{R}, \\ u(T,x) = \varphi(x), & x \in \mathbb{R}, \end{cases} \tag{1.5}$$

where L is the integro-differential operator

$$\begin{aligned} Lu(t,x) &= \partial_t u(t,x) + r\partial_x u(t,x) + \gamma(t,x)(\partial_x u(t,x) - u(t,x)) + \frac{\sigma^2(t,x)}{2}(\partial_{xx} - \partial_x)u(t,x) \\ &\quad - \int_{\mathbb{R}} \nu(t,x,dz)(e^z - 1 - z)\partial_x u(t,x) + \int_{\mathbb{R}} \nu(t,x,dz)(u(t,x+z) - u(t,x) - z\partial_x u(t,x)). \end{aligned}$$

Denote by $\Gamma(t,x;T,dy)$ the fundamental solution of the operator L, which is defined as the solution of (1.5) with $\varphi = \delta_y$, where δ_y is the Dirac-delta function i.e.

$$\begin{cases} L\Gamma(t,x,T,y) = 0, & t \in [0,T[,\ x \in \mathbb{R}, \\ \Gamma(T,x,T,y) = \delta_y(x), & x \in \mathbb{R}, \end{cases} \tag{1.6}$$

Now, due to the state-dependency in the coefficients of the operator, there exists no explicit solution to the above Cauchy problem; similarly, for the characteristic function, which can be obtained by taking the Fourier transform of the above Cauchy problem, using $\mathcal{F}(\delta_y(x)) = e^{i\xi y}$, no explicit solutions exist. Therefore, in order to evaluate option prices under the local Lévy model we will introduce a so-called adjoint expansion method, based on a Taylor expansion of the coefficients of the operator in the Cauchy problem; in this way we split the single Cauchy problem into multiple problems, each of which will be simple to solve explicitly. For the full derivation we refer to Chapter 2. Note that this expansion method can be applied directly to compute option prices or densities using the Cauchy problem in (1.5) or (1.6); alternatively, one can compute the characteristic function by solving the Cauchy problem (1.6) in the Fourier domain and combine this with the above-mentioned COS method. In particular, when working under the local Lévy model the jumps are more easily handled in the Fourier domain, motivating our choice of employing the latter method.

1.4 Valuation adjustments

After the financial crisis which started in 2007, regulation required financial institutions to hold enough capital on their balance sheets to account for *all* components involved in a trade, and not just for the option value itself; in other words these valuation adjustments include the default risk and funding costs in the risk management (and pricing) of over-the-counter (OTC) derivatives. In particular, it was recognized that Counterparty Credit Risk (CCR) poses a substantial risk for financial institutions. In 2010 in the Basel III framework an additional capital charge requirement, called Credit Valuation Adjustment (CVA), was introduced to cover the risk of losses on a counterparty default event for OTC uncollateralized derivatives. The CVA is the expected loss arising from a default of the counterparty and can be defined as the difference between the risky value and the current risk-free value of a derivatives contract. CVA is calculated and hedged in the same way as derivatives by many banks, therefore having efficient ways of calculating the value and the Greeks of these adjustments is important.

One common way of pricing CVA is to use the concept of expected exposure, defined as the mean of the exposure distribution at a future date. Calculating these exposures typically involves computationally time-consuming Monte Carlo procedures, like nested Monte Carlo schemes or the more efficient least-squares Monte Carlo method (LSM)([49]). Recently the Stochastic Grid Bundling method (SGBM) was introduced as an improvement of the standard LSM ([45]). This method was extended to pricing CVA for Bermudan options in [32]. Another recently introduced alternative is the so-called finite-differences Monte Carlo method, see [25]. This method uses the scenario generation from the Monte Carlo method combined with finite-difference option valuation. In Section 3.4.2 we present an alternative analytical method for computing the CVA.

Besides CVA, many other valuation adjustments, collectively called XVA, have been introduced in derivative pricing in the recent years, causing a change in the way derivatives contracts are priced. For instance, a companies own credit risk is taken into account with a debt value adjustment (DVA). The DVA is the expected gain that will be experienced by the bank in the event that the bank defaults on its portfolio of derivatives with a counterparty. To reduce the credit risk in a derivatives contract, the parties can include a credit support annex (CSA), requiring one or both of the parties to post collateral. Valuation of derivatives under CSA was first done in [60]. A margin valuation adjustment (MVA) arises when the parties are required to post an initial margin. In this case the cost of posting the initial margin to the counterparty over the length of the contract is known as MVA. Funding value adjustments (FVA) can be interpreted as a funding cost or benefit associated to the hedge of market risk of an uncollateralized transaction through a collateralized market. While there is still a debate going on about whether to include or exclude this adjustment, see [42], [41] and [18] for an in-depth overview of the arguments, most dealers now seem to indeed take into account the FVA. The capital value adjustment (KVA) refers to the cost of funding the additional capital that is required for derivative trades. This capital acts as a buffer against unexpected losses and thus, as argued in [36], has to be included in derivative pricing.

When computing derivatives which take into account these above mentioned valuation adjustments, one often needs to redefine the hedging portfolio in order to account for the various cash flows resulting from the various risks taken into account, see e.g. [11] and [47]. In particular, we consider a derivative contract $\hat{v}$ on an asset S between a seller B and a counterparty C, both of which may default. In the case of default of either party, clearly the future payments that were supposed to be made by the defaulted party will no longer be possible, resulting in a loss in capital for the surviving party. The ability to quantify and hedge this risk is of the essence. Using replication arguments which include this credit risk and the various funding costs involved in the trading of the derivative, one can derive a pricing PDE, similar to the Black-Scholes framework. In particular, we hedge the derivative with a self-financing portfolio which will cover all the underlying risk factors of the model. This gives rise to non-linear partial differential equations as we will see in Chapter 3, and in the particular case of pricing under the local Lévy model we will obtain a non-linear PDE with state-dependent coefficients. By rewriting this non-linear PDE as a backward stochastic differential equation (BSDE), as will be defined in Section 3.2.2, we can use the techniques mentioned in the previous sections for computing the expected value arising in the valuation of derivatives without XVA, i.e. the COS method and adjoint expansion method, to compute the expected values arising from discretization and subsequent approximation of the BSDE in order to find the value of the derivatives *with* the various valuation adjustments.

1.4.1 Setting up a hedging portfolio

Here we briefly discuss the main idea behind setting up a hedging portfolio use to subsequently determine the corresponding derivative prices. We consider here the general Black-Scholes framework in which we assume one can lend and borrow at a single risk-free rate and the holder and its counterparty do not possess any default risk. In Chapter 3 we will extend this derivation to account for the various types of funding involved when holding a derivative, in other words there will no longer be a *single* rate at which the bank can borrow/lend, and both the bank itself and its counterparty will be prone to default.

We assume the asset S follows the Geometric Brownian motion $dS_t = \mu S_t dt + \sigma S_t dW_t$, and take an option v which is written on the underlying S. The idea is to construct a hedging portfolio such that the portfolio is riskless, i.e. in this case focussing on hedging the risk arising from the Brownian motion in the asset dynamics. The portfolio consists of the shorted derivative itself, Δ units of the underlying stock and come cash g,

$$\Pi(t) = -v(t, S) + \Delta S_t + g_t.$$

By the self-financing assumption, i.e. assuming that no money is extracted or added into the portfolio we obtain

$$\begin{aligned} d\Pi_t &= -dv(t, S) + \Delta dS_t + dg(t) \\ &= \left(-\frac{\partial v}{\partial t} - \mu S \frac{\partial v}{\partial S} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 v}{\partial S^2} + \Delta \mu S\right) dt + \left(\sigma S \frac{\partial v}{\partial S} + \Delta \sigma S\right) dW_t + dg(t), \end{aligned}$$

where we have applied Itô's lemma to $dv(t, S)$. The remaining cash in the portfolio will earn the risk-free rate

$$dg(t) = r(v(t, S) - \Delta S_t)dt.$$

In order to hedge the Brownian motion risk we set

$$\Delta = -\frac{\partial v}{\partial S}.$$

Then using the fact that the portfolio has to satisfy the martingale condition in the risk-neutral world, i.e. $\mathbb{E}[d\Pi] = 0$, we find the classic Black-Scholes option pricing formula

$$Lv(t, S) = rv(t, S),$$

where

$$Lv(t, S) = \frac{\partial v}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 v}{\partial S^2} + r\frac{\partial v}{\partial S}.$$

In Section 3.2.1 the set-up is different in the sense that we assume the asset itself to be risk-neutral, so the focus is on hedging the risk arising from the defaults of the bank and its counterparty, which we will do by including two defaultable bonds in our portfolio, corresponding to the bank and its counterparty. Furthermore, the portfolio does not simply earn the risk-free rate, but the various payments happen at different interest rates, corresponding to the particular type of lending/borrowing (e.g. collateralized, uncollateralized).

1.5 Systemic risk

The last notion of risk we will consider is systemic risk; systemic risk concerns itself with the risk of collapse in a large interconnected system triggered by an event at the level of the individual firm or bank. Studying the stability properties of such systems is of fundamental importance in modern economies, in which the global interconnectedness of financial institutions poses a significant systemic risk [39].

There are two main forms of linkage arising between financial institutions, for a full overview we refer to [13]. The first is via counterparty risk, arising from the fact that institutions trade derivatives and lend/borrow from other banks in the system. In this way, a default at one institution, and thus the failure to repay all or part of its loans, might result in the subsequent insolvency of its creditors, who were relying on these payments to fulfill their own obligations. A second form of contagion between banks arises from common balance sheet holdings, so that if the insolvency of one institution forces it to sell a large bulk of its assets, the subsequent lowering of the price of these assets can also hit the monetary reserves of other institutions holding similar assets. Being able to model these linkages is thus of importance when studying the stability of such an interconnected system, and in particular its susceptibility to a systemic risk event.

Many ways of modelling the financial network exist, but in Chapter 4 we focus on using a so-called mean-field model: the matrix of interbank borrowing/lending activities is exogenously specified and the dynamics of the banks' monetary reserves contain an interaction term, modelled through the empirical distribution of the system states. This empirical distribution thus captures the interaction of the node with the other nodes in the system. One way of studying these interacting systems is by investigating the behaviour of the system as the number of nodes approaches infinity (i.e. propagation of chaos). In [8] the authors consider an interacting model of the monetary reserve processes where the drift term represents interbank short-time lending and the monetary reserve is additionally subjected to a banking sector indicator which drives additional in-/out-flows of cash. By means of a detailed weak-convergence analysis they conclude that the underlying limit state process has purely diffusive dynamics and the contribution of the banking sector jump process is reflected only in the drift. In [56] the authors use the mean-field approach with an interaction through hitting times in estimating systemic failure.

Besides contagion through interbank lending agreements, it is important to also include the fact that contagion can occur through multiple other channels, in particular linked balance sheets that may result in fire sales (see e.g. [14] and [19]) and the so-called financial acceleration where a shock affecting the banks' portfolio can cause a reaction of its creditors to claim back even more funds of the bank in question. In [20] the authors argue that one should not ignore the compounded effect of both correlated market events and default contagion, since it can make the network considerably more vulnerable to default cascades. Motivated by the above mentioned research, in Chapter 4 we will model the (negative) effects of the self-exciting fire sales as well as the financial acceleration by including a self- and cross-exciting Hawkes counting process, as introduced in [38], in the dynamics for the monetary reserves of the bank. We combine this with the default propagation through interbank lending agreements to study the robustness of the network. Explicitly computing the systemic risk in such a large *multivariate* model is hard by means of the typical PDE methods as used in e.g. option pricing. We therefore will study the behavior of the system using the above-mentioned large system limit, derived by means of a weak convergence analysis. In the large system limit the interactions between individual nodes are simplified, and therefore computing systemic risk indicators can be done in an efficient manner.

1.5.1 Weak convergence

A weak convergence analysis allows one to analyze a particular system in a limiting setting. Besides the work of [8], this kind of large limit analysis has been applied to compute the risks in various settings. By means of a weak converence analysis the authors of [33] study the behavior of the default intensity in a large portfolio where the intensity is subjected to additional sources of clustering through exposure to a systematic risk factor and a contagion term. The law of large numbers (LLN) result is proven under the assumption that the systematic risk vanishes in the large-portfolio limit. In [34] the authors extend the previous result for general diffusion dynamics for the systemic risk factor without the vanishing assumption, producing a stochastic PDE for this density in the limit, as opposed to a PDE. In [66] the LLN result is extended by proving a central limit theorem

(CLT) in a similar setting, thus quantifying the fluctuations of the empirical measure (and thereby also the loss from default) around its large portfolio limits. In [12] the large portfolio limit for assets following a correlated diffusion is shown to approach a measure whose density satisfies an SPDE, while in [37] a similar result is proven for a stochastic volatility model for the asset price. Finally, [64] and [65] use mean-field and large portfolio approximation methods for the analysis of large pools of loans. We briefly discuss the main ingredients needed to show weak convergence in our setting; for a more complete introduction we refer to [7] for the definitions of weak convergence of probability measures and the work of the authors of [29] for the applications to Markov processes; in particular Chapter 3 for the weak convergence definitions and Chapter 4 for the needed tools to show the convergence.

For a metric space (V, d) (d denoting the metric), we let $\mathcal{V} = \mathcal{B}(V)$ be the σ-algebra of Borel subsets of V, and $\mathcal{P}(V)$ be the family of Borel probability measures on $\mathcal{V}$ of the metric space V. We start with the basic definitions of weak convergence of probability measures. Let $(\Sigma, \mathcal{F}, \mathbb{P})$ be the underlying probability space on which all random variables are defined. A probability measure on $\mathcal{V}$ is a non-negative, countably additive set function $\mathbb{Q}$ satisfying $\mathbb{Q}(V) = 1$. We define weak convergence as

Definition 1.5.1 (Weak convergence of probability measures)**.** If the probability measures $\mathbb{Q}^N$ and $\mathbb{Q}$ on $\mathcal{V}$ satisfy

$$\mathbb{Q}_N(f) \to \mathbb{Q}(f), \quad \text{where } \mathbb{Q}(f) = \int_V f d\mathbb{Q},$$

as $N \to \infty$ for every bounded continuous real function f on V, we say that $\mathbb{Q}^N$ converges weakly to $\mathbb{Q}$ and write $\mathbb{Q}^N \Rightarrow \mathbb{Q}$.

Equivalently, a sequence X_N of V-valued random variables is said to converge in distribution to the V-valued random variable X if

$$\lim_{N\to\infty} \mathbb{E}[f(X_N)] = \mathbb{E}[f(X)],$$

for $f \in C_K^\infty(V)$, where the distribution of the V-valued random variable X is an element of $\mathcal{P}(V)$.

In our situation we typically work with processes that contain jumps, and therefore we consider the weak convergence in the space D which includes certain discontinuous functions. In particular, $D_S(\mathbb{R}_+)$ is the space of real functions $x : \mathbb{R}_+ \to S$ that are right-continuous and have left-hand limits (RCLL). The space $D_S(\mathbb{R}_+)$ can be topologized by the Skorokhod metric which we denote by d_S. In particular, under this metric $D_S(\mathbb{R}_+)$ is separable if S is separable and $(D_S(\mathbb{R}_+), d_S)$ is complete if the metric space (S, d) is complete; see also Chapter 3.5 in [29].

More specifically we will consider stochastic processes X_t^i, $i = 1, ..., M$ taking values in some space $\mathcal{O}$. In order to study the behavior of the processes in the limit as $M \to \infty$ we define the sequence of empirical measures as

$$\nu_t^M := \frac{1}{M}\sum_{i=1}^{M} \mathbb{1}_{X_t^i}, \quad t \geq 0, \tag{1.7}$$

where $\mathbb{1}_x$ represents a unit mass placed at position $x \in \mathbb{R}$. Then ν_t^M take values in $V = D_S(\mathbb{R}_+)$, the space of RCLL processes on $[0, \infty)$ taking values in $S := \mathcal{P}(\mathcal{O})$, the collection of Borel probability measures on $\mathcal{O}$. We are interested in the weak convergence of elements of $\mathcal{P}(D_S(\mathbb{R}_+))$. Define

$$\mathbb{Q}_t^M := \mathbb{P}(\nu_t^M \in A),$$

for $A \in \mathcal{V}$ where $\mathcal{V} = \mathcal{B}(D_S(\mathbb{R}_+))$. The idea is to show that $\mathbb{Q}^M \Rightarrow \mathbb{Q}^*$, for some limit point $\mathbb{Q}^*$ which governs the behavior of the processes X_t^i, $i = 1, ..., M$ in the limit $M \to \infty$.

We will need the following two definitions.

Definition 1.5.2 (Relative compactness)**.** A family of probability measures Π on $(V, \mathcal{V})$ is defined to be relatively compact if every sequence of elements of Π contains a weakly convergent subsequence.

Definition 1.5.3 (Tightness)**.** A probability measure $\mathbb{P}$ on $(V, \mathcal{V})$ is tight if for each $\epsilon > 0$ there exists a compact set K such that $\mathbb{P}(K) > 1 - \epsilon$.

In order to show weak convergence we typically work on separable spaces, in particular we have the following lemma relating separability and weak convergence (see also Lemma 4.3 of Chapter 3.4 in [29])

Lemma 1.5.4. *Let the sequence $\{\mathbb{Q}^M\} \subset \mathcal{P}(V)$ be relatively compact, let $\mathbb{Q} \in \mathcal{P}(V)$, and let $\mathcal{M} \subset C_K^\infty(V)$ (a subset of differentiable, bounded functions on V) be separating. If*

$$\lim_{N\to\infty} \int_V f d\mathbb{Q}^N = \int_V f d\mathbb{Q},$$

holds for $f \in \mathcal{M}$, then $\mathbb{Q}^N \Rightarrow \mathbb{Q}$.

By Prohorov's theorem, if the space V is separable and complete under the corresponding metric d, a family of probability measures on $(V, \mathcal{V})$ is relatively compact if and only if it is tight.

Recall the following Lemma defining the tightness of distribution-valued processes in $D_S(\mathbb{R}_+)$ through tightness in real-valued process obtained by applying test functions (Chapter 4, Proposition 1.7 in [46])

Lemma 1.5.5. *A family of non-negative measure-valued processes $\{\nu^M : M \geq 1\}$ is tight in $D_S(\mathbb{R}_+)$ if $\{\nu^M(f) : M \geq 1\}$ is tight in $D_\mathbb{R}(\mathbb{R}_+)$ for $f \in C_K^\infty(\mathbb{R})$.*

Now fix $f \in C_K^\infty(\mathbb{R})$, a smooth function with compact support. To prove tightness for some measure $\nu_t^M(f)$ taking values in $D_\mathbb{R}(\mathbb{R}_+)$ equipped with the Skorokhod topology it is enough to verify the following two conditions

1. Compact containment: for any $t \in [0,T]$ and $\delta > 0$, there exists a compact set $K(t,\delta) \subset \mathbb{R}$ such that

$$\inf_M \mathbb{P}(\nu_t^M(f) \in K(t,\delta)) > 1-\delta. \tag{1.8}$$

2. For every $\delta > 0$

$$\lim_{\epsilon \to 0} \limsup_{M\to\infty} \mathbb{P}\left[\sup_{|s-t|\leq\epsilon} |\nu_t^M(f) - \nu_s^M(f)| > \delta\right] = 0. \tag{1.9}$$

If a sequence of measures satisfies the above requirements, it results in c-tightness of the sequence, a stronger result than just tightness since it means that it is tight with *continuous* sample paths.

1.5.2 Weak convergence for systemic risk

In order to see the basics of the derivation of weak convergence in a systemic risk application we will consider a simplified form of the set-up in [8] and briefly discuss how the limiting behavior of the network is derived. The monetary reserves of the k-th bank, $k = 1, ..., M$ are assumed to satisfy the following SDE:

$$dX_t^i = \frac{a^i}{M}\sum_{k=1}^{M}(X_t^k - X_t^i)dt + \sigma^i dW_t,$$

with initial monetary reserve level X_0^i, a^i being the positive constant governing the amount of money that is being either borrowed from bank k if $X_t^k > X_t^i$, or loaned to bank k if $X_t^i > X_t^k$, and $W_t = (W_t^i)_{i=1}^M$ a M-dimensional Brownian motion. The monetary reserve process governs the state of the bank, meaning that once $X_t^i \leq 0$ the bank is considered to be in a defaulted state. In our simplistic model we assume that if a bank has defaulted it still participates in the interbank network, meaning that if it receives a sufficient number of cash inflow it can come out of the defaulted state. An easy way to modify this, and to let a bank remain in defaulted state once the default level has been reached is to consider the Brownian motion to be scaled by X_t^i itself as well. Note that we can rewrite

$$dX_t^i = a^i(\bar{X}_t - X_t^i)dt + \sigma^i dW_t, \quad \bar{X}_t = \frac{1}{M}\sum_{k=1}^{M} X_t^k,$$

where the interaction is thus governed by the empirical density; therefore the above model is also called a mean-field model.

The idea is to analyze the behavior of the system as $M \to \infty$, i.e. the number of banks in the system tends to infinity. Assume for simplicity that all parameters in the model are equal, i.e. $p^i = (a^i, \sigma^i) = p^*$, for $i = 1, ..., M$. In order to analyze the behavior of the system in the limit we typically work with the sequence of empirical measures defined as in (1.7). In other words the empirical measures keep track of the empirical distribution of the monetary reserves. Let $S = \mathcal{P}(\mathbb{R})$ be the collection of Borel probability measures on $\mathbb{R}$. Then the empirical measure is an element of the space of RCLL functions $D_S(\mathbb{R}_+)$. Showing weak convergence of ν_t^M then amounts to (i) derive the form of the limiting martingale problem, i.e. understand the behavior of the processes ν_t^N as $N \to \infty$, (ii) show the existence of at least one limit point using the tightness, i.e. verify that $(\nu_t^N)_{N\in\mathbb{N}}$ is relatively compact as a $D_S(\mathbb{R}_+)$-valued random variable, and (iii) conclude uniqueness of the solution of the resulting limiting martingale problem. This is the structure we will typically follow when deriving weak convergence, in particular in Chapter 4.

We thus want to use the martingale problem (see Chapter 4 of [29]) to show that ν^M converges to a limiting process. The martingale problem is based on the fact that if a process $(X_t)_{t\geq 0}$ is a Markov process with infinitesimal generator $\mathcal{A}$, M_t defined as,

$$M_t := f(X_t) - f(X_0) - \int_0^t \mathcal{A}f(X_s)ds,$$

is a martingale. By a solution to the martingale problem we then mean a measurable stochastic process X with values in metric space S such that M_t is a martingale. Alternatively, a measurable process X is a solution of the martingale problem for generator $\mathcal{A}$ if and only if

$$\mathbb{E}\left[\left(f(X_{t_{m+1}}) - f(X_{t_m}) - \int_{t_m}^{t_{m+1}} \mathcal{A}f(X_s)ds\right)\prod_{j=1}^{M} h_j(X_{t_k})\right] = 0,$$

for $0 \leq t_1 < t_2 < ... < t_{m+1}$ and $h_1, ..., h_M \in C_K^\infty(S)$, the set of bounded, continuous functions. Generally speaking, if a sequence of processes X^M satisfies a martingale problem with generator $\mathcal{A}^M$, then we can say that the finite-dimensional distributions of X^M converge weakly to those of the process X with generator $\mathcal{A}$ if

$$\lim_{M\to\infty} \mathbb{E}\left[\left(f(X^M_{t_{m+1}}) - f(X^M_{t_m}) - \int_{t_m}^{t_{m+1}} \mathcal{A}f(X^M_s)ds\right)\prod_{j=1}^{M} h_j(X_{t_k})\right] = 0.$$

In this way, deriving the limiting martingale problem characterizes the limit point X, it being the solution of the above martingale problem. In more concrete terms the above is stated in Lemma 8.1 and Theorem 8.2 of Chapter 4 of [29].

To derive the limiting martingale problem in our case we will work with the empirical measure as applied to a test function $f \in C_K^\infty(\mathbb{R})$,

$$\nu_t^M(f) = \frac{1}{M}\sum_{i=1}^{M} f(X_t^k).$$

As mentioned in Lemma 1.5.4 in order to show weak convergence of the martingale problem we need to work with functions on separating subspaces. Define $\mathcal{S}$ to be the collection of bounded elements Φ on S defined by

$$\Phi(\nu) = \varphi(\nu(\mathbf{f})),$$

with $\mathbf{f} = (f_1, ..., f_N)$ and $f_n \in C_K^\infty(\mathbb{R})$. Then $\mathcal{S}$ separates S and it thus suffices to show weak convergence of the martingale problem for these functions $\Phi \in \mathcal{S}$.

We start with deriving the form of the limiting martingale problem. Applying Itô's formula to $\varphi(\nu(\mathbf{f}))$ we find

$$\Phi(\nu_u^M) = \Phi(\nu_t^M) + \int_t^u \left(\mathcal{C}_s^M + \mathcal{D}_s^M\right) ds + \mathcal{M}_u - \mathcal{M}_t,$$

with $\mathcal{M}_t$ a martingale and

$$\mathcal{C}_t^M := \sum_{n=1}^N \frac{\partial \varphi(\nu_t^M(\mathbf{f}))}{\partial f_n} \left(\nu_t^M(\mathcal{L}^1 f_n)\nu_t^M(I) - \nu_t^M(\mathcal{L}^2 f_n) + \nu_t^M(\mathcal{L}^3 f_n)\right),$$

$$\mathcal{D}_t^M := \frac{1}{M^2} \sum_{n,l=1}^N \frac{\partial^2 \varphi(\nu_t^M(\mathbf{f}))}{\partial f_n \partial f_l} \sum_{i=1}^M \left(\sigma^2 \frac{\partial f_n(X_t^i)}{\partial x} \frac{\partial f_l(X_t^i)}{\partial x}\right),$$

where we have defined

$$\mathcal{L}^1 f(x) = a\partial_x f(x), \quad \mathcal{L}^2 f(x) = ax\partial_x f(x), \quad \mathcal{L}^3 f(x) = \frac{1}{2}\sigma^2 \partial_{xx} f(x).$$

We need to understand what happens as $M \to \infty$. In particular, using the boundedness of the derivatives of $f \in C_K^\infty(\mathbb{R})$ and the scaling by M^{-2} we obtain that

$$\lim_{M\to\infty} \mathbb{E}\left[\int_t^u |\mathcal{D}_s^M| ds\right] = 0.$$

The limiting martingale problem is then given by

$$\lim_{M\to\infty} \mathbb{E}\Bigg[\Bigg(\Phi(\nu_{t_{m+1}}^M) - \Phi(\nu_{t_m}^M) - \int_{t_m}^{t_{m+1}} \sum_{n=1}^N \frac{\partial \varphi(\nu_u^M(\mathbf{f})}{\partial f_n} \left(\nu_u^M(\mathcal{L}^1 f_n)\nu_u^M(I) - \nu_u^M(\mathcal{L}^2 f_n) + \nu_u^M(\mathcal{L}^3 f_n)\right)\Bigg) \prod_{j=1}^M \Psi_j(\nu_{t_j}^M)\Bigg] = 0,$$

for $\Psi_j \in L^\infty(S)$.

Next, we need to show that the measure-valued process ν^M is relatively compact, when viewed as a sequence of stochastic processes in the Skorokhod space $D_S(\mathbb{R}_+)$. In other words, we need to verify the conditions for tightness: the compact containment given in (1.8) and regularity from (1.9). Then we have to verify that the solution of the resulting martingale problem is unique, for which we will use a duality argument (Chapter 4.4. in [29]). The full derivations for this will be discussed in Section 4.5.

Having shown the form of the martingale problem and derived existence and uniqueness of the limiting measure, we can finish the derivation by deriving the form of the limiting measure. Let $\mathbb{Q}_M$ be the $\mathbb{P}$-law of ν^M, that is

$$\mathbb{Q}^M := \mathbb{P}(\nu \in A),$$

for all $A \in \mathcal{B}(D_S(\mathbb{R}_+))$. Then we have $\mathbb{Q}^M \to \mathbb{Q}$, with $\mathbb{Q} = \delta_\nu$, where

$$\nu_t(f) = \mathbb{E}[f(X_t(\mathbf{p}))],$$

with

$$dX_t(\mathbf{p}) = a(x - X_t(\mathbf{p}))dt + \sigma dW_t.$$

In other words, in the limit of $M \to \infty$ the interaction between the banks in the system vanishes and the limiting behavior of each node is governed by the process $X_t(\mathbf{p})$; more specifically, the mean-field interaction term is represented by $(x - X_t(\mathbf{p}))$.

1.6 Organization of this thesis

The rest of this thesis is organized as follows. In Chapter 2 we show how one can value Bermudan options under the flexible local Lévy dynamics; the method is based on an adjoint expansion of the characteristic function in combination with the COS method and allows for an accurate and efficient valuation of both option prices and sensitivities required for hedging. In Chapter 3 we compute the Bermudan option value under the local Lévy model when in additional incorporating valuation adjustments. In particular these valuation adjustments require us to re-derive the hedging portfolio, resulting in a non-linear PDE, which we solve by means of a transformation to a backward stochastic differential equation and the adjoint expansion method; furthermore we extend the pricing under valuation adjustments to swaptions, one of the most common interest rate derivatives traded in the market. Chapter 4 is concerned with the computation of systemic risk, in which we model the banks capital by a stochastic process with a self-exciting jump term. This term accounts for additional sources of contagion observed in financial markets. We consider an asymptotic setting in which we let the number of banks in the system tend to infinity, and derive a limiting process through a weak convergence analysis. As opposed to numerical methods for computing the risk in a large system, the limiting process allows for a much faster computation of the necessary systemic risk indicators.

Chapter 2

Bermudan option valuation under local Lévy models

As discussed before, efficient methods for the computation of financial derivative prices are often required; the mathematical task of which boils down to the computation of integrals. For many stochastic processes that model the financial assets, these integrals can be most efficiently computed in the Fourier domain. However, for some relevant and recent stochastic models the Fourier domain computations are not at all straightforward, as these computations rely on the availability of the characteristic function of the stochastic process (read: the Fourier transform of the transitional probability distribution), which is not always known. This is especially true for state-dependent asset price processes, and for asset processes that include the notion of default in their definition. With the derivations and techniques in this chapter we make available the highly efficient pricing of so-called Bermudan options to the above mentioned classes of state-dependent asset dynamics, including jumps in asset prices and the possibility of default. In this sense, the class of asset models for which Fourier option pricing is highly efficient increases by the derivations presented here. Essentially, we approximate the characteristic function by an advanced Taylor-based expansion in such a way that the resulting characteristic function exhibits favorable properties for the pricing methods.

Fourier methods have often been among the winners in option pricing competitions such as BENCHOP [67]. In [31], a Fourier method called the COS method, as introduced in [30], was extended to the pricing of Bermudan options. The computational efficiency of the method was based on a specific structure of the characteristic function allowing to use the fast Fourier transform (FFT) for calculating the continuation value of the option. Fourier methods can readily be applied to solving problems under asset price dynamics for which the characteristic function is available. This is the case for exponential Lévy models, such as the Merton model developed in [55], the Variance-Gamma model developed in [54], but also for the Heston model [40]. However, in the case of local volatility, default and state-dependent jump measures there is no closed form characteristic function available and the COS method can not be readily applied.

Recently, in [58] the so-called *adjoint expansion method* for the approximation of the characteristic function in local Lévy models is presented. This method is worked out in the Fourier space by considering the adjoint formulation of the pricing problem, that is using a backward parametrix expansion as was also later done in [5]. Here, we generalize this method to include a defaultable asset whose risk-neutral pricing dynamics are described by an exponential Lévy-type martingale with a state-dependent jump measure, as has also been considered in [53] and in [44].

Having obtained the analytical approximation for the characteristic function we combine this with the COS method for Bermudan options. We show that this analytical formula for the characteristic function still possesses a structure that allows the use of an FFT-based method in order to calculate the continuation value. This results in an efficient and accurate computation of the Bermudan option value and of the Greeks. The characteristic function approximation used in the COS method is already very accurate for the 2nd-order approximation, meaning that the explicit formulas are simple. This makes the method easy and quick to implement. Finally, we present a theoretical justification of the accurate performance of the method by giving the error bounds for the approximated characteristic function.

The rest of this chapter is organized as follows. In Section 2.1 we present the general framework which includes a local default intensity, a state-dependent jump measure and a local volatility function. Then we derive the adjoint expansion of the characteristic function. In Section 2.2 we propose an efficient algorithm for calculating the Bermudan option value, which makes use of the Fast Fourier transform. In Section 2.3 we prove error bounds for the 0th- and 1st-order approximation, justifying the accuracy of the method. Finally, in Section 2.4 numerical examples are presented, showing the flexibility, accuracy and speed of the method.

2.1 General framework

Consider the Local Lévy model as defined in Section 1.2.2. As discussed before (see also [48, Section 2.2]) the price v of a European option with maturity T and payoff $\Phi(S_T)$ is given by

$$v(t, X_t) = \mathbb{1}_{\{\zeta > t\}} e^{-r(T-t)} E\left[e^{-\int_t^T \gamma(s, X_s) ds} \varphi(X_T) | X_t \right], \qquad t \leq T, \tag{2.1}$$

where $\varphi(x) = \Phi(e^x)$. Thus, in order to compute the price of an option, we must evaluate functions of the form

$$u(t, x) := E\left[e^{-\int_t^T \gamma(s, X_s) ds} \varphi(X_T) | X_t = x \right]. \tag{2.2}$$

Under standard assumptions, u can be expressed as the classical solution of the following Cauchy problem

$$\begin{cases} Lu(t,x) = 0, & t \in [0,T[,\ x \in \mathbb{R}, \\ u(T,x) = \varphi(x), & x \in \mathbb{R}, \end{cases}$$

where L is the integro-differential operator

$$\begin{aligned} Lu(t,x) &= \partial_t u(t,x) + r\partial_x u(t,x) + \gamma(t,x)(\partial_x u(t,x) - u(t,x)) + \frac{\sigma^2(t,x)}{2}(\partial_{xx} - \partial_x)u(t,x) \\ &\quad - \int_{\mathbb{R}} \nu(t,x,dz)(e^z - 1 - z)\partial_x u(t,x) + \int_{\mathbb{R}} \nu(t,x,dz)(u(t,x+z) - u(t,x) - z\partial_x u(t,x)). \end{aligned} \tag{2.3}$$

The function u in (2.2) can be represented as an integral with respect to the transition distribution of the defaultable log-price process $\log S$:

$$u(t,x) = \int_{\mathbb{R}} \varphi(y)\Gamma(t,x;T,dy). \tag{2.4}$$

Here we notice explicitly that $\Gamma(t,x;T,dy)$ is not necessarily a standard probability measure because its integral over $\mathbb{R}$ can be strictly less than one; nevertheless, with a slight abuse of notation, we say that its Fourier transform

$$\hat{\Gamma}(t,x;T,\xi) := \mathcal{F}(\Gamma(t,x;T,\cdot))(\xi) := \int_{\mathbb{R}} e^{i\xi y}\Gamma(t,x;T,dy), \qquad \xi \in \mathbb{R},$$

is the characteristic function of $\log S$.

2.1.1 Adjoint expansion of the characteristic function

In this section we generalize the results in [58] to our framework and develop an expansion of the coefficients

$$a(t,x) := \frac{\sigma^2(t,x)}{2}, \qquad \gamma(t,x), \qquad \nu(t,x,dz),$$

around some point $\bar{x}$. The coefficients $a(t,x)$, $\gamma(t,x)$ and $\nu(t,x,dz)$ are assumed to be continuously differentiable with respect to x, up to order $N \in \mathbb{N}$.

From now on, for simplicity, we assume that the coefficients are independent of t (see Remark 2.1.2 for the general case). First we introduce the nth-order approximation of L in (2.3):

$$\begin{aligned} L_n = L_0 + \sum_{k=1}^{n} \Big(&(x-\bar{x})^k a_k(\partial_{xx} - \partial_x) + (x-\bar{x})^k \gamma_k \partial_x - (x-\bar{x})^k \gamma_k \\ &- \int_{\mathbb{R}} (x-\bar{x})^k \nu_k(dz)(e^z - 1 - z)\partial_x + \int_{\mathbb{R}} (x-\bar{x})^k \nu_k(dz)(e^{z\partial_x} - 1 - z\partial_x)\Big), \end{aligned}$$

where

$$L_0 = \partial_t + r\partial_x + a_0(\partial_{xx} - \partial_x) + \gamma_0\partial_x - \gamma_0 - \int_{\mathbb{R}} \nu_0(dz)(e^z - 1 - z)\partial_x + \int_{\mathbb{R}} \nu_0(dz)(e^{z\partial_x} - 1 - z\partial_x),$$

and

$$a_k = \frac{\partial_x^k a(\bar{x})}{k!}, \qquad \gamma_k = \frac{\partial_x^k \gamma(\bar{x})}{k!}, \qquad \nu_k(dz) = \frac{\partial_x^k \nu(\bar{x}, dz)}{k!}, \qquad k \geq 0.$$

The basepoint $\bar{x}$ is a constant parameter which can be chosen freely. In general the simplest choice is $\bar{x} = x$ (the value of the underlying at initial time t): we will see that in this case the formulas for the Bermudan option valuation are simplified.

Let us assume for a moment that L_0 has a fundamental solution $G^0(t, x; T, y)$ that is defined as the solution of the Cauchy problem

$$\begin{cases} L_0 G^0(t, x; T, y) = 0 & t \in [0, T[, \ x \in \mathbb{R}, \\ G^0(T, \cdot; T, y) = \delta_y. \end{cases}$$

In this case we define the nth-order approximation of Γ as

$$\Gamma^{(n)}(t, x; T, y) = \sum_{k=0}^{n} G^k(t, x; T, y),$$

where, for any $k \geq 1$ and (T, y), $G^k(\cdot, \cdot; T, y)$ is defined recursively through the following Cauchy problem

$$\begin{cases} L_0 G^k(t, x; T, y) = -\sum\limits_{h=1}^{k} (L_h - L_{h-1}) G^{k-h}(t, x; T, y) & t \in [0, T[, \ x \in \mathbb{R}, \\ G^k(T, x; T, y) = 0, & x \in \mathbb{R}. \end{cases}$$

Notice that

$$\begin{aligned} L_h - L_{h-1} = {} & (x - \bar{x})^h a_h(\partial_{xx} - \partial_x) + (x - \bar{x})^h \gamma_h \partial_x - (x - \bar{x})^h \gamma_h \\ & - \int_{\mathbb{R}} (x - \bar{x})^h \nu_h(dz)(e^z - 1 - z)\partial_x + \int_{\mathbb{R}} (x - \bar{x})^h \nu_h(dz)(e^{z\partial_x} - 1 - z\partial_x). \end{aligned}$$

Correspondingly, the nth-order approximation of the characteristic function $\hat{\Gamma}$ is defined to be

$$\hat{\Gamma}^{(n)}(t, x; T, \xi) = \sum_{k=0}^{n} \mathcal{F}\left(G^k(t, x; T, \cdot)\right)(\xi) := \sum_{k=0}^{n} \hat{G}^k(t, x; T, \xi), \qquad \xi \in \mathbb{R}. \tag{2.5}$$

Now we remark that the operator L acts on (t, x) while the characteristic function is a Fourier transform taken with respect to y: in order to take advantage of such a transformation, in the following theorem we characterize $\hat{\Gamma}^{(n)}$ in terms of the Fourier transform of the adjoint operator $\tilde{L} = \tilde{L}^{(T,y)}$ of L, acting on (T, y).

Theorem 2.1.1 (Dual formulation)**.** *For any $(t,x) \in]0,T] \times \mathbb{R}$, the function $G^0(t,x;\cdot,\cdot)$ is defined through the following dual Cauchy problem*

$$\begin{cases} \tilde{L}_0^{(T,y)} G^0(t,x;T,y) = 0 & T > t, \ y \in \mathbb{R}, \\ G^0(T,x;T,\cdot) = \delta_x. \end{cases} \tag{2.6}$$

where

$$\tilde{L}_0^{(T,y)} = -\partial_T - r\partial_y + a_0(\partial_{yy} + \partial_y) - \gamma_0\partial_y - \gamma_0 + \int_{\mathbb{R}} \nu_0(dz)(e^z - 1 - z)\partial_y + \int_{\mathbb{R}} \bar{\nu}_0(dz)(e^{z\partial_y} - 1 - z\partial_y).$$

Moreover, for any $k \geq 1$, the function $G^k(t,x;\cdot,\cdot)$ is defined through the dual Cauchy problem as follows:

$$\begin{cases} \tilde{L}_0^{(T,y)} G^k(t,x;T,y) = -\sum_{h=1}^{k} \left(\tilde{L}_h^{(T,y)} - \tilde{L}_{h-1}^{(T,y)}\right) G^{k-h}(t,x;T,y) & T > t, \ y \in \mathbb{R}, \\ G^k(T,x;T,y) = 0 & y \in \mathbb{R}, \end{cases} \tag{2.7}$$

with

$$\begin{aligned} \tilde{L}_h^{(T,y)} - \tilde{L}_{h-1}^{(T,y)} &= a_h h(h-1)(y-\bar{x})^{h-2} + a_h (y-\bar{x})^{h-1} \left(2h\partial_y + (y-\bar{x})(\partial_{yy} + \partial_y) + h\right) \\ &\quad - \gamma_h h(y-\bar{x})^{h-1} - \gamma_h (y-\bar{x})^h (\partial_y + 1) \\ &\quad + \int_{\mathbb{R}} \nu_h(dz)(e^z - 1 - z) \left(h(y-\bar{x})^{h-1} + (y-\bar{x})^h \partial_y\right) \\ &\quad + \int_{\mathbb{R}} \bar{\nu}_h(dz) \left((y+z-\bar{x})^h e^{z\partial_y} - (y-\bar{x})^h - z\left(h(y-\bar{x})^{h-1} - (y-\bar{x})^h \partial_y\right)\right), \end{aligned}$$

where in defining the adjoint of the operator we use the notation

$$e^{z\partial_y} f(y) := \sum_{n=0}^{\infty} \frac{z^n}{n!} \partial_y^n f(y) = f(y+z).$$

Notice that the adjoint Cauchy problems (2.6) and (2.7) admit a solution in the Fourier space and can be solved explicitly; in fact, we have

$$\mathcal{F}\left(\tilde{L}_0^{(T,\cdot)} G^k(t,x;T,\cdot)\right)(\xi) = \psi(\xi)\hat{G}^k(t,x;T,\xi) - \partial_T \hat{G}^k(t,x;T,\xi),$$

where $\psi(\xi)$ is the characteristic exponent of the Lévy process with coefficients γ_0, a_0 and $\nu_0(dz)$, that is

$$\psi(\xi) = i\xi(r + \gamma_0) + a_0(-\xi^2 - i\xi) - \gamma_0 - \int_{\mathbb{R}} \nu_0(dz)(e^z - 1 - z)i\xi + \int_{\mathbb{R}} \nu_0(dz)(e^{iz\xi} - 1 - iz\xi).$$

Thus the solution (in the Fourier space) to problems (2.6) and (2.7) is given by

$$
\begin{aligned}
&\hat{G}^0(t,x;T,\xi) = e^{i\xi x+(T-t)\psi(\xi)},\\
&\hat{G}^k(t,x;T,\xi) = -\int_t^T e^{\psi(\xi)(T-s)}\mathcal{F}\left(\sum_{h=1}^{k}\left(\tilde{L}_h^{(s,\cdot)}-\tilde{L}_{h-1}^{(s,\cdot)}\right)G^{k-h}(t,x;s,\cdot)\right)(\xi)ds, \qquad k\geq 1.
\end{aligned} \tag{2.8}
$$

Now we consider the general framework and in particular we drop the assumption on the existence of the fundamental solution of L_0: in this case, we define the nth-order approximation of the characteristic function $\hat{\Gamma}$ as in (2.5), with $\hat{G}^k$ given by (2.8). We also notice that

$$
\begin{aligned}
&\mathcal{F}\left(\left(\tilde{L}_h^{(s,\cdot)}-\tilde{L}_{h-1}^{(s,\cdot)}\right)u(s,\cdot)\right)(\xi) =\\
&\quad\Big(a_h h(h-1)(-i\partial_\xi-\bar{x})^{h-2}+a_h(-i\partial_\xi-\bar{x})^{h-1}\left(-2hi\xi+(-i\partial_\xi-\bar{x})(-\xi^2-i\xi)+h\right)\Big)\hat{u}(s,\xi)\\
&\quad-\Big(\gamma_h h(-i\partial_\xi-\bar{x})^{h-1}-\gamma_h(-i\partial_\xi-\bar{x})^{h}(i\xi-1)\Big)\hat{u}(s,\xi)\\
&\quad+\int_{\mathbb{R}}\nu_h(dz)(e^z-1-z)\Big(h(-i\partial_\xi-\bar{x})^{h-1}-(-i\partial_\xi-\bar{x})^{h}i\xi\Big)\hat{u}(s,\xi)\\
&\quad+\int_{\mathbb{R}}\nu_h(dz)\Big((-i\partial_y-z-\bar{x})^{h}e^{i\xi z}-(-i\partial_y-\bar{x})^{h}+z\Big(h(-i\partial_\xi-\bar{x})^{h-1}-(-i\partial_\xi-\bar{x})^{h}i\xi\Big)\Big)\hat{u}(s,\xi).
\end{aligned}
$$

Remark 2.1.2. In case the coefficients γ, σ, ν depend on time, the solutions to the Cauchy problems are similar:

$$
\begin{aligned}
&\hat{G}^0(t,x;T,\xi) = e^{i\xi x}e^{\int_t^T\psi(s,\xi)ds},\\
&\hat{G}^k(t,x;T,\xi) = -\int_t^T e^{\int_s^T\psi(\tau,\xi)d\tau}\mathcal{F}\left(\sum_{h=1}^{k}\left(\tilde{L}_h^{(s,\cdot)}(s)-\tilde{L}_{h-1}^{(s,\cdot)}(s)\right)G^{k-h}(t,x;s,\cdot)\right)(\xi)ds,
\end{aligned}
$$

with

$$
\psi(s,\xi) = i\xi(r+\gamma_0(s))+a_0(s)(-\xi^2-i\xi)-\int_{\mathbb{R}}\nu_0(s,dz)(e^z-1-z)i\xi+\int_{\mathbb{R}}\nu_0(s,dz)(e^{iz\xi}-1-iz\xi),
$$

$$
\begin{aligned}
\tilde{L}_h^{(s,y)}(s)-\tilde{L}_{h-1}^{(s,y)}(s) &= a_h(s)h(h-1)(y-\bar{x})^{h-2}+a_h(s)(y-\bar{x})^{h-1}\left(2h\partial_y+(y-\bar{x})(\partial_{yy}+\partial_y)+h\right)\\
&\quad-\gamma_h(s)h(y-\bar{x})^{h-1}-\gamma_h(s)(y-\bar{x})^{h}(\partial_y+1)\\
&\quad+\int_{\mathbb{R}}\nu_h(s,dz)(e^z-1-z)\Big(h(y-\bar{x})^{h-1}+(y-\bar{x})^{h}\partial_y\Big)\\
&\quad+\int_{\mathbb{R}}\bar{\nu}_h(s,dz)\Big((y+z-\bar{x})^{h}e^{z\partial_y}-(y-\bar{x})^{h}-z\Big(h(y-\bar{x})^{h-1}-(y-\bar{x})^{h}\partial_y\Big)\Big).
\end{aligned}
$$

From these results one can already see that the dependency on x comes in through $e^{i\xi x}$ and after taking derivatives the dependency on x will take the form $(x-\bar{x})^m e^{i\xi x}$: this fact will be crucial in our analysis.

Example 2.1.3. To see the above dependency more explicitly for the second-order approximation of the characteristic function we consider, for ease of notation, a simplified model: a one-dimensional local Lévy model where the log-price solves the SDE

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + \int_{\mathbb{R}} d\tilde{N}_t(dz)z. \tag{2.9}$$

This model is a simplification of the original model, since we consider only a local volatility function, and no local default or state-dependent Lévy measure. Thus only a Taylor expansion of the local volatility coefficient is used. However, the dependency that we will see generalizes in the same way to the local default and state-dependent measure. By the martingale condition we have

$$\mu(x) = r - a(x) - \int_{\mathbb{R}} \nu(dz)(e^z - 1),$$

and therefore the Kolmogorov operator of (2.9) reads

$$\begin{aligned} Lu(t,x) = {} & \partial_t u(t,x) + r\partial_x u(t,x) + a(t,x)(\partial_{xx} - \partial_x)u(t,x) \\ & - \int_{\mathbb{R}} \nu(dz)(e^z - 1) + \int_{\mathbb{R}} \nu(dz)\left(u(t,x+z) - u(t,x)\right). \end{aligned}$$

In this case, we have the following explicit approximation formulas for the characteristic function $\hat{\Gamma}(t,x;T,\xi)$:

$$\hat{\Gamma}(t,x;T,\xi) \;\approx\; \hat{\Gamma}^{(n)}(t,x;T,\xi) := e^{i\xi x + (T-t)\psi(\xi)} \sum_{k=0}^{n} \hat{F}^k(t,x;T,\xi), \qquad n \geq 0, \tag{2.10}$$

with

$$\psi(\xi) = ir\xi - a_0(\xi^2 + i\xi) - \int_{\mathbb{R}} \nu(dz)(e^z - 1)i\xi + \int_{\mathbb{R}} \nu(dz)\left(e^{iz\xi} - 1\right),$$

and

$$\hat{F}^k(t,x;T,\xi) = \sum_{h=0}^{k} g_h^{(k)}(T-t,\xi)(x-\bar{x})^h; \tag{2.11}$$

here, for $k = 0, 1, 2$, we have

$$\begin{aligned} g_0^{(0)}(s,\xi) = {} & 1, \\ g_0^{(1)}(s,\xi) = {} & a_1 s^2(\xi^2 + i\xi)\frac{i}{2}\psi'(\xi), \\ g_1^{(1)}(s,\xi) = {} & -a_1 s(\xi^2 + i\xi), \\ g_0^{(2)}(s,\xi) = {} & \frac{1}{2}s^2 a_2 \xi(i+\xi)\psi''(\xi) - \frac{1}{6}s^3\xi(i+\xi)(a_1^2(i+2\xi)\psi'(\xi) - 2a_2\psi'(\xi)^2 + a_1^2\xi(i+\xi)\psi''(\xi)) \\ & - \frac{1}{8}s^4 a_1^2\xi^2(i+\xi)^2\psi'(\xi)^2, \\ g_1^{(2)}(s,\xi) = {} & \frac{1}{2}s^2\xi(i+\xi)(a_1^2(1-2i\xi) + 2ia_2\psi''(\xi)) - \frac{1}{2}s^3 i a_1^2\xi^2(i+\xi)^2\psi''(\xi), \\ g_2^{(2)}(s,\xi) = {} & -a_2 s\xi(i+\xi) + \frac{1}{2}s^2 a_1^2\xi^2(i+\xi)^2. \end{aligned}$$

Using the notation from above, we can write in the same way the approximation formulas for the general case. Here we present the results for $k = 0, 1$, since higher-order formulas are too long to include. We have:

$$\begin{aligned}
g_0^{(0)}(s,\xi) =& 1,\\
g_0^{(1)}(s,\xi) =& \frac{i}{2}a_1 s^2(\xi^2 + i\xi)\psi'(\xi) + \frac{1}{2}\gamma_1 s^2(i+\xi)\psi'(\xi) - \frac{1}{2}\int_{\mathbb{R}} \nu_1(dz)(e^z - 1 - z)s^2\xi\psi'(\xi) \qquad (2.12)\\
&- \frac{1}{2}\int_{\mathbb{R}} \nu_1(dz)(ie^{i\xi z} - i + \xi z)s^2\psi'(\xi),\\
g_1^{(1)}(s,\xi) =& -a_1 s(\xi^2 + i\xi) + \gamma_1 si(i+\xi) - \int_{\mathbb{R}} \nu_1(dz)(e^z - 1 - z)s\xi i\\
&+ \int_{\mathbb{R}} \nu_1(dz)(e^{i\xi z} - 1 - \xi iz)s.
\end{aligned}$$

Remark 2.1.4. From (2.10)-(2.11) and (3.18) we clearly see that the approximation of order n is a function of the form

$$\hat{\Gamma}^{(n)}(t,x;T,\xi) := e^{i\xi x}\sum_{k=0}^{n}(x-\bar{x})^k g_{n,k}(t,T,\xi), \tag{2.13}$$

where the coefficients $g_{n,k}$, with $0 \le k \le n$, depend only on t, T and ξ, but not on x. The approximation formula can thus always be split into a sum of products of functions depending only on ξ and functions that are linear combinations of $(x-\bar{x})^m e^{i\xi x}$, $m \in \mathbb{N}_0$.

2.2 Bermudan option valuation

A Bermudan option is a financial contract in which the holder can exercise at a predetermined finite set of exercise moments prior to maturity, and the holder of the option receives a payoff when exercising. Consider a Bermudan option with a set of M exercise moments $\{t_1, ..., t_M\}$, with $0 \le t_1 < t_2 < \cdots < t_M = T$. When the option is exercised at time t_m the holder receives the payoff $\Phi(t_m, S_{t_m})$. Recalling (2.1), the no-arbitrage value of the Bermudan option at time t is

$$v(t, X_t) = \mathbb{1}_{\{\zeta > t\}} \sup_{\tau\in\mathcal{T}_t} E\left[e^{-\int_t^\tau (r+\gamma(s,X_s))ds}\varphi(\tau, X_\tau)|X_t\right],$$

where $\varphi(t,x) = \Phi(t, e^x)$ and $\mathcal{T}_t$ is the set of all $\mathcal{G}$-stopping times taking values in $\{t_1, ..., t_M\} \cap [t, T]$. For a Bermudan Put option with strike price K, we simply have $\varphi(t,x) = (K - e^x)^+$. By the dynamic programming approach, the option value can be expressed by a backward recursion as

$$v(t_M, x) = \mathbb{1}_{\{\zeta > t_M\}}\varphi(t_M, x)$$

and

$$\begin{cases} c(t,x) = E\left[e^{\int_t^{t_m}(r+\gamma(s,X_s))ds} v(t_m, X_{t_m})|X_t = x\right], & t \in [t_{m-1}, t_m[\\ v(t_{m-1}, x) = \mathbb{1}_{\{\zeta > t_{m-1}\}}\max\{\varphi(t_{m-1}, x), c(t_{m-1}, x)\}, & m \in \{2, \dots, M\}. \end{cases} \tag{2.14}$$

In the above notation $v(t,x)$ is the option value and $c(t,x)$ is the so-called continuation value. The option value is set to be $v(t,x) = c(t,x)$ for $t \in]t_{m-1}, t_m[$, and, if $t_1 > 0$, also for $t \in [0, t_1[$.

Remark 2.2.1. Since the payoff of a Call option grows exponentially with the log-stock price, this may introduce significant cancellation errors for large domain sizes. For this reason we price Put options only using our approach and we employ the well-known Put-Call parity to price Calls via Puts. This is a rather standard argument (see, for instance, [68]).

2.2.1 An algorithm for pricing Bermudan Put options

In pricing the Bermudan option we will employ the COS method proposed by [31], and as explained in Section 1.3, which we briefly restate here for convenience. The COS method is based on the insight that the Fourier-cosine series coefficients of $\Gamma(t,x;T,dy)$ (and therefore also of option prices) are closely related to the characteristic function of the underlying process, namely the following relationship holds:

$$\int_a^b e^{i\frac{k\pi}{b-a}}\Gamma(t,x;T,dy) \approx \hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right).$$

The COS method provides a way to calculating expected values (integrals) of the form

$$v(t,x) = \int_{\mathbb{R}} \varphi(T,y)\Gamma(t,x;T,dy),$$

by (1) truncating the integration range to a finite region $[a,b]$, (2) replacing the distribution with its cosine expansion and truncating the series to N terms, and (3) using the above relation between the density and the characteristic function to rewrite the option price as

$$v(t,x) \approx {\sum_{k=0}^{N-1}}' \operatorname{Re}\left(e^{-ik\pi\frac{a}{b-a}}\hat{\Gamma}\left(t,x;T,\frac{k\pi}{b-a}\right)\right) V_k(T), \tag{2.15}$$

where

$$V_k(T) = \frac{2}{b-a}\int_a^b \cos\left(k\pi\frac{y-a}{b-a}\right)\varphi(T,y)dy,$$

are the Fourier-cosine series coefficients of the payoff function at time T respectively.

Next we return to the Bermudan Put pricing problem. Remembering that the expected value $c(t,x)$ in (2.14) can be rewritten in integral form as in (2.4), we have

$$c(t,x) = e^{-r(t_m-t)}\int_{\mathbb{R}} v(t_m,y)\Gamma(t,x;t_m,dy), \qquad t \in [t_{m-1}, t_m[.$$

Then we use the Fourier-cosine expansion (2.15), so that we get the approximation:

$$\hat{c}(t,x) = e^{-r(t_m-t)}{\sum_{k=0}^{N-1}}' \operatorname{Re}\left(e^{-ik\pi\frac{a}{b-a}}\hat{\Gamma}\left(t,x;t_m,\frac{k\pi}{b-a}\right)\right) V_k(t_m), \qquad t \in [t_{m-1}, t_m[\tag{2.16}$$

$$V_k(t_m) = \frac{2}{b-a}\int_a^b \cos\left(k\pi\frac{y-a}{b-a}\right)\max\{\varphi(t_m,y), c(t_m,y)\}dy,$$

with $\varphi(t,x) = (K-e^x)^+$.

Next we recover the coefficients $(V_k(t_m))_{k=0,1,...,N-1}$ from $(V_k(t_{m+1}))_{k=0,1,...,N-1}$. To this end, we split the integral in the definition of $V_k(t_m)$ into two parts using the early-exercise point x_m^*, which is the point where the continuation value is equal to the payoff, i.e. $c(t_m,x_m^*) = \varphi(t_m,x_m^*)$; thus we have

$$V_k(t_m) = F_k(t_m,x_m^*) + C_k(t_m,x_m^*), \qquad m = M-1, M-2, ..., 1,$$

where

$$\begin{aligned} F_k(t_m,x_m^*) &:= \frac{2}{b-a}\int_a^{x_m^*} \varphi(t_m,y)\cos\left(k\pi\frac{y-a}{b-a}\right)dy, \\ C_k(t_m,x_m^*) &:= \frac{2}{b-a}\int_{x_m^*}^b c(t_m,y)\cos\left(k\pi\frac{y-a}{b-a}\right)dy, \end{aligned} \tag{2.17}$$

and $V_k(t_M) = F_k(t_M, \log K)$.

Remark 2.2.2. Since we have a semi-analytic formula for $\hat{c}(t_m,x)$, we can easily find the derivatives with respect to x and use Newton's method to find the point x_m^* such that $c(t_m,x_m^*) = \varphi(t_m,x_m^*)$. A good starting point for the Newton method is $\log K$, since $x_m^* \le \log K$.

The coefficients $F_k(t_m,x_m^*)$ can be computed analytically using $x_m^* \le \log K$, so that we have

$$\begin{aligned} F_k(t_m,x_m^*) &= \frac{2}{b-a}\int_a^{x_m^*}(K-e^y)\cos\left(k\pi\frac{y-a}{b-a}\right)dy \\ &= \frac{2}{b-a}K\Psi_k(a,x_m^*) - \frac{2}{b-a}\chi_k(a,x_m^*), \end{aligned}$$

where

$$\begin{aligned} \chi_k(a,x_m^*) &= \int_a^{x_m^*} e^y\cos\left(k\pi\frac{y-a}{b-a}\right)dy \\ &= \frac{1}{1+\left(\frac{k\pi}{b-a}\right)^2}\left(e^{x_m^*}\cos\left(k\pi\frac{x_m^*-a}{b-a}\right) - e^a + \frac{k\pi e^{x_m^*}}{b-a}\sin\left(k\pi\frac{x_m^*-a}{b-a}\right)\right), \\ \Psi_k(a,x_m^*) &= \int_a^{x_m^*}\cos\left(k\pi\frac{y-a}{b-a}\right)dy = \begin{cases} \frac{b-a}{k\pi}\sin\left(k\pi\frac{x_m^*-a}{b-a}\right), & k \neq 0, \\ x_m^*-a, & k=0. \end{cases} \end{aligned}$$

On the other hand, by inserting the approximation (2.16) for the continuation value into the formula for $C_k(t_m,x_m^*)$ have the following coefficients $\hat{C}_k$ for $m = M-1, M-2, ..., 1$:

$$\hat{C}_k(t_m,x_m^*) = \frac{2e^{-r(t_{m+1}-t_m)}}{b-a}{\sum_{j=0}^{N-1}}' V_j(t_{m+1})\int_{x_m^*}^b \mathrm{Re}\left(e^{-ij\pi\frac{a}{b-a}}\hat{\Gamma}\left(t_m,x;t_{m+1},\frac{j\pi}{b-a}\right)\right)\cos\left(k\pi\frac{x-a}{b-a}\right)dx. \tag{2.18}$$

Thus the algorithm for pricing Bermudan options can then be summarized as follows:

1. For $k = 0, 1, ..., N-1$:
   - At time t_M, the coefficients are exact: $V_k(t_M) = F_k(t_M, \log K)$, as in (2.17).
2. For $m = M-1$ to 1:
   - Determine the early-exercise point x_m^* using Newton's method;
   - Compute $\hat{V}_k(t_m)$ using formula $\hat{V}_k(t_m) := F_k(t_m, x_m^*) + \hat{C}_k(t_m, x_m^*)$, (2.17) and (2.18). Use an FFT for the continuation value (see Section 3.2).
3. Final step: using $\hat{V}_k(t_1)$ determine the option price $\hat{v}(0, x) = \hat{c}(0, x)$ using (2.16).

Figure 2.1: Algorithm for Bermudan option valuation

2.2.2 An efficient algorithm for the continuation value

In this section we derive an efficient algorithm for calculating $\hat{C}_k(t_m, x_m^*)$ in (2.18). When considering an exponential Lévy process with constant coefficients as done in [31], the continuation value can be calculated using a Fast Fourier Transform (FFT). This can be done due to the fact that the characteristic function $\hat{\Gamma}(t, x; T, \xi)$ can be split into a product of a function depending only on ξ and a function of the form $e^{i\xi x}$. Note that we typically have $\xi = \frac{j\pi}{b-a}$. The integration over x results in a sum of a Hankel and Toeplitz matrix (with indices $(j+k)$ and $(j-k)$ respectively). The matrix-vector product, with these special matrices, can be transformed into a circular convolution which can be computed using FFTs.

From (3.18) we know that the nth-order approximation of the characteristic function is of the form:

$$\hat{\Gamma}^{(n)}(t_m, x; t_{m+1}, \xi) = e^{i\xi x} \sum_{k=0}^{n} (x - \bar{x})^k g_{n,k}(t_m, t_{m+1}, \xi),$$

where the coefficients $g_{n,k}(t, T, \xi)$, with $0 \le k \le n$, depend only on t, T and ξ, but not on x. Using (3.18) we write the continuation value as:

$$\hat{C}_k(t_m, x_m^*) = \sum_{h=0}^{n} e^{-r(t_{m+1}-t_m)} \sum_{j=0}^{N-1}{}' \operatorname{Re}\left(V_j(t_m) g_{n,h}\left(t_m, t_{m+1}, \frac{j\pi}{b-a}\right) M_{k,j}^h(x_m^*, b) \right),$$

where we have interchanged the sums and integral and defined:

$$M_{k,j}^h(x_m^*, b) = \frac{2}{b-a} \int_{x_m^*}^{b} e^{ij\pi \frac{x-a}{b-a}} (x - \bar{x})^h \cos\left(k\pi \frac{x-a}{b-a}\right) dx \tag{2.19}$$

This can be written in vectorized form as:

$$\hat{\mathbf{C}}(t_m, x_m^*) = \sum_{h=0}^{n} e^{-r(t_{m+1}-t_m)} \operatorname{Re}\left(\mathbf{V}(t_{m+1}) \mathcal{M}^h(x_m^*, b) \Lambda^h \right),$$

where $\mathbf{V}(t_{m+1})$ is the vector $[V_0(t_{m+1}), ..., V_{N-1}(t_{m+1})]^T$ and $\mathcal{M}^h(x_m^*, b)\Lambda^h$ is a matrix-matrix product with $\mathcal{M}^h$ being a matrix with elements $\{M_{k,j}^h\}_{k,j=0}^{N-1}$ and Λ^h is a diagonal matrix with elements

$$g_{n,h}\Big(t_m, t_{m+1}, \frac{j\pi}{b-a}\Big), \qquad j = 0, \dots, N-1.$$

We have the following theorem for calculating a generalized form of the integral in (2.19) which is used in the calculation of the continuation value.

Theorem 2.2.3. *The matrix $\mathcal{M}$ with elements $\{M_{k,j}\}_{k,j=0}^{N-1}$ such that:*

$$M_{k,j} = \int e^{jx}\cos(kx)x^m dx,$$

consists of sums of Hankel and Toeplitz matrices.

Proof. Using standard trigonometric identities we can rewrite the integral as:

$$\begin{aligned} M_{k,j} &= \int \cos(jx)\cos(kx)x^m dx + i\int \sin(jx)\cos(kx)x^m dx \\ &= M_{k,j}^H + iM_{k,j}^T, \end{aligned}$$

where we have defined:

$$\begin{aligned} M_{k,j}^H &= \frac{1}{2}\int \cos((j+k)x)x^m dx + \frac{1}{2}\int \sin((j+k)x)x^m dx, \\ M_{k,j}^T &= \frac{1}{2}\int \cos((j-k)x)x^m dx + \frac{1}{2}\int \sin((j-k)x)x^m dx. \end{aligned}$$

The following holds:

$$\begin{aligned} \int \cos(nx)x^m dx = &\frac{1}{n}x^m \sin(nx) + \sum_{i=1}^{\lceil m/2 \rceil}(-1)^{i+1}\frac{\prod_{j=0}^{2i-2}(m-j)}{n^{2i}}\cos(nx)x^{m-(2i-1)} \\ &- \sum_{i=1}^{\lfloor m/2 \rfloor}(-1)^{i+1}\frac{\prod_{j=0}^{2i-1}(m-j)}{n^{2i+1}}\sin(nx)x^{m-2i}, \\ \int \sin(nx)x^m dx = &-\frac{1}{n}x^m \cos(nx) + \sum_{i=1}^{\lceil m/2 \rceil}(-1)^{i+1}\frac{\prod_{j=0}^{2i-2}(m-j)}{n^{2i}}\sin(nx)x^{m-(2i-1)} \\ &- \sum_{i=1}^{\lfloor m/2 \rfloor}(-1)^{i+1}\frac{\prod_{j=0}^{2i-1}(m-j)}{n^{2i+1}}\cos(nx)x^{m-2i}. \end{aligned}$$

It follows that $\{M_{k,j}^H\}_{k,j=0}^{N-1}$ is a Hankel matrix with coefficient $(j+k)$ and $\{M_{k,j}^T\}_{k,j=0}^{N-1}$ is a Toeplitz matrix with coefficient $(j-k)$:

$$\mathcal{M}_H = \begin{pmatrix} M_0 & M_1 & M_2 & \dots & M_{N-1} \\ M_1 & M_2 & \dots & & M_N \\ \vdots & & & & \vdots \\ M_{N-2} & M_{N-1} & \dots & & M_{2N-3} \\ M_{N-1} & \dots & & M_{2N-3} & M_{2N-2} \end{pmatrix},$$

$$\mathcal{M}_T = \begin{pmatrix} M_0 & M_1 & \dots & M_{N-2} & M_{N-1} \\ M_{-1} & M_0 & M_1 & \dots & M_{N-2} \\ \vdots & & \ddots & & \vdots \\ M_{2-N} & \dots & M_{-1} & M_0 & M_1 \\ M_{1-N} & M_{2-N} & & M_{-1} & M_0 \end{pmatrix},$$

where we have defined

$$M_j = \frac{1}{2}\int \cos(jx)x^m dx + \frac{1}{2}\int \sin(jx)x^m dx.$$

□

From Theorem 2.2.3 we see that $\mathcal{M}^h(x_m^*, b)$ with elements $M_{k,j}^h$ consists of a sum of a Hankel and Toeplitz matrix.

Example 2.2.4. We derive explicitly the Hankel and Toeplitz matrices for $m = 0$ and $m = 1$. We calculate the indefinite integral

$$M_{k,j} = \frac{2}{b-a}\int e^{ij\pi\frac{x-a}{b-a}} \cos\left(k\pi\frac{x-a}{b-a}\right)(x-\bar{x})^m dx.$$

Suppose $m = 0$, in this case we have $M_{k,j} = M_{k,j}^H + M_{k,j}^T$, with:

$$M_{k,j}^H = -\frac{i\exp\left(i\frac{(j+k)\pi(x-a)}{b-a}\right)}{\pi(j+k)},$$
$$M_{k,j}^T = -\frac{i\exp\left(i\frac{(j-k)\pi(x-a)}{b-a}\right)}{\pi(j-k)},$$

where $\{M_{k,j}^H\}_{k,j=0}^{N-1}$ is a Hankel matrix and $\{M_{k,j}^T\}_{k,j=0}^{N-1}$ is a Toeplitz matrix with

$$M_j = \begin{cases} \frac{x}{b-a}, & j = 0, \\ \frac{i\exp\left(i\frac{j\pi(x-a)}{b-a}\right)}{\pi j}, & j \neq 0. \end{cases}$$

Suppose $m = 1$, in this case we have:

$$M_{k,j}^H = -\frac{a-b}{(j-k)^2\pi^2}\exp\left(i(j-k)\pi\frac{(x-a)}{b-a}\right) - \frac{x-\bar{x}}{(j-k)\pi}i\exp\left(i(j-k)\pi\frac{(x-a)}{b-a}\right),$$
$$M_{k,j}^T = -\frac{a-b}{(j+k)^2\pi^2}\exp\left(i(j+k)\pi\frac{(x-a)}{b-a}\right) - \frac{x-\bar{x}}{(j+k)\pi}i\exp\left(i(j+k)\pi\frac{(x-a)}{b-a}\right),$$

where $\{M_{k,j}^H\}_{k,j=0}^{N-1}$ is a Hankel matrix and $\{M_{k,j}^T\}_{k,j=0}^{N-1}$ is a Toeplitz matrix, with

$$M_j = \begin{cases} \frac{x(x-\bar{x})}{b-a}, & j = 0, \\ -\frac{a-b}{j^2\pi^2}\exp\left(ij\pi\frac{(x-a)}{b-a}\right) - \frac{x-\bar{x}}{j\pi}i\exp\left(ij\pi\frac{(x-a)}{b-a}\right), & j \neq 0. \end{cases}$$

Remark 2.2.5. If we take $\bar{x} = x$, which is most common in practice, the formulas are simplified significantly and only the case of $m = 0$ is relevant. In this case the characteristic function is simply $e^{i\xi x}$ times a sum of terms depending only on t_m, t_{m+1} and $\xi = \frac{j\pi}{b-a}$:

$$\hat{\Gamma}^{(n)}(t_m, x; t_{m+1}, \xi) = e^{i\xi x} g_{n,0}(t_m, t_{m+1}, \xi).$$

Using the split into sums of Hankel and Toeplitz matrices we can write the continuation value in matrix form as:

$$\hat{\boldsymbol{C}}(t_m, x_m^*) = \sum_{h=0}^{n} e^{-r(t_{m+1}-t_m)} \text{Re}\left((\mathcal{M}_H^h + \mathcal{M}_T^h) \boldsymbol{u}^h \right),$$

where $\mathcal{M}_H^h = \{M_{k,j}^{H,h}(x_m^*, b)\}_{k,j=0}^{N-1}$ is a Hankel matrix and $\mathcal{M}_T^l = \{M_{k,j}^{T,h}(x_m^*, b)\}_{k,j=0}^{N-1}$ is a Toeplitz matrix and $\boldsymbol{u}^h = \{u_j^h\}_{j=0}^{N-1}$, with $u_j^h = g_{n,h}\left(t_m, t_{m+1}, \frac{j\pi}{b-a}\right) V_j(t_{m+1})$ and $u_0^h = \frac{1}{2} g_{n,h}\left(t_m, t_{m+1}, 0\right) V_0(t_{m+1})$.

We recall that the circular convolution, denoted by $\circledast$, of two vectors is equal to the inverse discrete Fourier transform ($\mathcal{D}^{-1}$) of the products of the forward DFTs, $\mathcal{D}$, i.e.:

$$\mathbf{x} \circledast \mathbf{y} = \mathcal{D}^{-1}\{\mathcal{D}(\mathbf{x}) \cdot \mathcal{D}(\mathbf{y})\}.$$

For Hankel and Toeplitz matrices we have the following result (see [2] and Result 2.2 in [31]):

Theorem 2.2.6. *For a Toeplitz matrix $\mathcal{M}_T$, the product $\mathcal{M}_T \mathbf{u}$ is equal to the first N elements of $\mathbf{m}_T \circledast \mathbf{u}_T$, where $\mathbf{m}_T$ and $\mathbf{u}_T$ are $2N$ vectors defined by*

$$\mathbf{m}_T = [M_0, M_{-1}, M_{-2}, ..., M_{1-N}, 0, M_{N-1}, M_{N-2}, ..., M_1]^T,$$
$$\mathbf{u}_T = [u_0, u_1, ..., u_{N-1}, 0, ..., 0]^T.$$

For a Hankel matrix $\mathcal{M}_H$, the product $\mathcal{M}_H \mathbf{u}$ is equal to the first N elements of $\mathbf{m_H} \circledast \mathbf{u_H}$ in reversed order, where $\mathbf{m}_H$ and $\mathbf{u}_H$ are $2N$ vectors defined by

$$\mathbf{m}_H = [M_{2N-1}, M_{2N-2}, ..., M_1, M_0]^T$$
$$\mathbf{u}_H = [0, ..., 0, u_0, u_1, ..., u_{N-1}]^T.$$

Summarizing, we can calculate the continuation value $\hat{\mathbf{C}}(t_m, x_m^*)$ using the algorithm in Figure 2.2.

The continuation value requires five DFTs for each $h = 0, ..., n$, and a DFT is calculated using the FFT. In practice it is most common to have $\bar{x} = x$ and in this case we only need five FFTs. The computation of $F_k(t_m, x_m^*)$ is linear in N. The overall complexity of the method is dominated by the computation of $\hat{C}(t_m, x_m^*)$, whose complexity is $O(N \log_2 N)$ with the FFT. The complexity of the calculation for option value at time 0 is $O(N)$. If we have a Bermudan option with M exercise dates, the overall complexity will be $O((M-1)N \log_2 N)$.

1. For $h = 0, ..., n$:
    - Compute $M_j^h(x_1, x_2)$
    - Construct $\mathbf{m}_H^h$ and $\mathbf{m}_T^h$
    - Compute $\boldsymbol{u}^h(t_m) = \{u_j^h\}_{j=0}^{N-1}$
    - Construct $\mathbf{u}_T^h$ by padding N zeros to $\boldsymbol{u}^h(t_m)$
    - $\mathbf{MTu}^h$ = the first N elements of $\mathcal{D}^{-1}\{\mathcal{D}(\mathbf{m}_T^h) \cdot \mathcal{D}(\mathbf{u}_T^h)\}$
    - $\mathbf{MHu}^h$ = reverse{the first N elements of $\mathcal{D}^{-1}\{\mathcal{D}(\mathbf{m}_H^h) \cdot \mathbf{sgn} \cdot \mathcal{D}(\mathbf{u}_T^h)\}$}
2. Compute the continuation value using $\hat{\mathbf{C}}(t_m, x_m^*) = \sum_{h=0}^{n} e^{-r(t_{m+1}-t_m)}\text{Re}(\mathbf{MTu}^h + \mathbf{MHu}^h)$.

Figure 2.2: Algorithm for the computation of $\hat{\mathbf{C}}(t_m, x_m^*)$

Remark 2.2.7 (American options). The prices of American options can be obtained by applying a Richardson extrapolation (see, for instance, [50]) on the prices of a few Bermudan options with a small number of exercise dates. Let v_M denote the value of a Bermudan option with maturity T and a number M of early exercise dates that are $\frac{T}{M}$ years apart. Then, for any $d \in \mathbb{N}$, the following 4-point Richardson extrapolation scheme

$$\frac{1}{21}\left(64v_{2^{d+3}} - 56v_{2^{d+2}} + 14v_{2^{d+1}} - v_{2^d}\right)$$

gives an approximation of the corresponding American option price.

Remark 2.2.8 (The Greeks). The approximation method can also be used to calculate the Greeks at almost no additional cost. In the case of $\bar{x} = x$, we have the following approximation formulas for Delta and Gamma:

$$\hat{\Delta} = e^{-r(t_1-t_0)}\sum_{k=0}^{N-1}{}' \text{Re}\left(e^{ik\pi\frac{x-a}{b-a}}\left(\frac{ik\pi}{b-a}g_{n,0}\left(t_0, t_1, \frac{k\pi}{b-a}\right) + g_{n,1}\left(t_0, t_1, \frac{k\pi}{b-a}\right)\right)\right)\hat{V}_k(t_1),$$

$$\hat{\Gamma} = e^{-r(t_1-t_0)}\sum_{k=0}^{N-1}{}' \text{Re}\Bigg(e^{ik\pi\frac{x-a}{b-a}}\Bigg(-\frac{ik\pi}{b-a}g_{n,0}\left(t_0, t_1, \frac{k\pi}{b-a}\right) - g_{n,1}\left(t_0, t_1, \frac{k\pi}{b-a}\right)$$
$$+ 2\frac{ik\pi}{b-a}g_{n,1}\left(t_0, t_1, \frac{k\pi}{b-a}\right) + \left(\frac{ik\pi}{b-a}\right)^2 g_{n,0}\left(t_0, t_1, \frac{k\pi}{b-a}\right) + 2g_{n,2}\left(t_0, t_1, \frac{k\pi}{b-a}\right)\Bigg)\Bigg)\hat{V}_k(t_1).$$

2.3 Error estimates

The error in our approximation consists of the error of the COS method and the error in the adjoint expansion of the characteristic function. The error of the COS method depends on the truncation of the integration range $[a, b]$ and the truncation of the infinite summation of the Fourier-cosine expansion by N. The density rapidly decays to zero as $y \to \pm\infty$. Then the overall error can be bounded as follows:

$$\epsilon_1(x; N, [a, b]) \leq Q \left| \int_{\mathbb{R}\setminus[a,b]} \Gamma(t, x; T, dy) \right| + \left| \frac{P}{(N-1)^{\beta-1}} \right|,$$

where P and Q are constants not depending on N or $[a, b]$ and $\beta \geq n \geq 1$, with n being the algebraic index of convergence of the cosine series coefficients. For a sufficiently large integration interval $[a, b]$, the overall error is dominated by the series truncation error, which converges exponentially. The error in the backward propagation of the coefficients $V_k(t_m)$ is defined as $\epsilon_2(k, t_m) := V_k(t_m) - \hat{V}_k(t_m)$. With $[a, b]$ sufficiently large and a probability density function in $C_K^\infty([a, b])$, the error $\epsilon_1(k, t_m)$ converges exponentially in N. For a detailed derivation on the error of the COS method see [30] and [31].

We now present the error estimates for the adjoint expansion of the characteristic function at orders zero and one. We consider for simplicity a model with time-independent coefficients

$$X_t = x + \int_0^t \mu(X_s)ds + \int_0^t \sigma(X_s)dW_s + \int_0^t \int_{\mathbb{R}} \eta(X_{s-})zd\tilde{N}(s, dz), \tag{2.20}$$

where we have defined as usual $d\tilde{N}(t, dz) = dN(t, dz) - \nu(dz)dt$. This model is similar to the model we considered initially in (3.16); only now we deal with slightly simplified version and assume that the dependency on X_t in the measure can be factored out, which is often enough the case.

Let $\tilde{X}_t$ be the 0th-order approximation of the model in (2.20) with $\bar{x} = x$, that is

$$\tilde{X}_t = x + \int_0^t \mu(x)ds + \int_0^t \sigma(x)dW_s + \int_0^t \int_{\mathbb{R}} \eta(x)zd\tilde{N}(s, dz). \tag{2.21}$$

The characteristic exponent of $\tilde{X}_t - x$ is

$$\psi(\xi) = i\xi\mu(x) - \frac{\sigma(x)^2}{2}\xi^2 - \eta(x)\int_{\mathbb{R}} \nu(dz)(e^z - 1 - z)i\xi + \eta(x)\int_{\mathbb{R}} \nu(dz)(e^{iz\xi} - 1 - iz\xi). \tag{2.22}$$

Theorem 2.3.1. *Let $n = 0, 1$ and assume that the coefficients μ, σ, η are continuously differentiable with bounded derivatives up to order n. Let $\hat{\Gamma}^{(n)}(0, x; t, \xi)$ in (2.5) be the nth-order approximation of the characteristic function. Then, for any $T > 0$ there exists a positive constant C that depends only on T, on the norms of the coefficients and on the Lévy measure ν, such that*

$$\left| \hat{\Gamma}(0, x; t, \xi) - \hat{\Gamma}^{(n)}(0, x; t, \xi) \right| \leq C \left(1 + |\xi|^{1+3n}\right) t^{n+1}, \qquad t \in [0, T], \ \xi \in \mathbb{R}. \tag{2.23}$$

Proof. Let X and $\tilde{X}$ be as in (2.20) and (2.21) respectively. We first prove that

$$E[|X_t - \tilde{X}_t|^2] \leq C\left(\kappa_2 t^2 + \kappa_1^2 t^3\right), \qquad t \in [0,T], \tag{2.24}$$

for some positive constant C that depends only on T, on the Lipschitz constants of the coefficients μ, σ, η and on the Lévy measure ν. Here $\kappa_1 = -\psi'(0)$ and $\kappa_2 = -\psi''(0)$ where ψ in (2.22) is the characteristic exponent of the Lévy process $(\tilde{X}_t - x)$.

Using the Hölder inequality, the Itô isometry (see, for instance, [59]) and the Lipschitz continuity of η, μ and σ, the mean squared error is bounded by:

$$\begin{aligned} E\left[|X_t - \tilde{X}_t|^2\right] &\leq 3E\left[\left(\int_0^t (\mu(X_s) - \mu(x))ds\right)^2\right] + 3E\left[\left(\int_0^t (\sigma(X_s) - \sigma(x))dW_s\right)^2\right] \\ &\quad + 3E\left[\left(\int_0^t \int_{\mathbb{R}} (\eta(X_{s-}) - \eta(x))z d\tilde{N}(s,dz)\right)^2\right] \\ &\leq C\int_0^t E\left[|\tilde{X}_s - x|^2\right] ds + C\int_0^t E\left[|X_s - \tilde{X}_s|^2\right] ds, \end{aligned} \tag{2.25}$$

where

$$C = 6\left(\|\mu'\|_\infty^2 + \|\sigma'\|_\infty^2 + \|\eta'\|_\infty^2 \int_{\mathbb{R}} z^2 \nu(dz)\right).$$

Now we recall the following relationship between the first and second moment and cumulants

$$E[(\tilde{X}_s - x)] = c_1(s), \qquad E[(\tilde{X}_s - x)^2] = c_2(s) + c_1(s)^2,$$

where

$$c_n(s) = \frac{s}{i^n} \frac{\partial^n \psi(\xi)}{\partial \xi^n}\bigg|_{\xi=0},$$

and $\psi(\xi)$ is the characteristic exponent of $(\tilde{X}_s - x)$. Thus we have

$$E\left[|\tilde{X}_s - x|^2\right] = \kappa_2 s + \kappa_1^2 s^2. \tag{2.26}$$

Plugging (2.26) into (2.25) we get

$$E[|X_t - \tilde{X}_t|^2] \leq C\left(\frac{\kappa_2}{2}t^2 + \frac{\kappa_1^2}{3}t^3\right) + C\int_0^t E\left[|X_s - \tilde{X}_s|^2\right] ds,$$

and therefore estimate (2.24) follows by applying the Gronwall inequality in the form

$$\varphi(t) \leq \alpha(t) + C\int_0^t \varphi(s)ds \implies \varphi(t) \leq \alpha(t) + C\int_0^t \alpha(s)e^{C(t-s)}ds,$$

that is valid for any $C \geq 0$ and φ, α continuous functions.

From (2.24) and (2.26) we can also deduce that

$$E\left[|X_t - x|^2\right] \le 2E\left[\left|X_t - \tilde{X}_t\right|^2\right] + 2E\left[\left|\tilde{X}_t - x\right|^2\right] \le C\left(\kappa_2 t + \kappa_1^2 t^2\right), \qquad t \in [0,T]. \tag{2.27}$$

Moreover, from (2.24) we also get the following error estimate for the expectation of a Lipschitz payoff function v:

$$\left|E\left[v(X_t)\right] - E[v(\tilde{X}_t)]\right| \le C\sqrt{\kappa_2 t + \kappa_1^2 t^2}, \qquad t \in [0,T],$$

where now C also depends on the Lipschitz constant of v. In particular, taking $v(x) = e^{ix\xi}$, this proves (2.23) for $n = 0$.

Next we prove (2.23) for $n = 1$.

Proceeding as in the proof of Lemma 6.23 in [52] with $u(0,x) = \hat{\Gamma}(0,x;t,\xi)$ and $\bar{x} = x$, we find

$$\hat{\Gamma}(0,x;t,\xi) - \hat{\Gamma}^{(1)}(0,x;t,\xi) = \int_0^t E\left[(L - L_0)\hat{G}^1(s,X_s;t,\xi) + (L - L_1)\hat{G}^0(s,X_s;t,\xi)\right] ds,$$

where the 1st-order approximation is as usual

$$\hat{\Gamma}^{(1)}(s,X;t,\xi) = \hat{G}^0(s,X;t,\xi) + \hat{G}^1(s,X;t,\xi),$$

with

$$\begin{aligned}
\hat{G}^0(s,X;t,\xi) &= e^{iX\xi + (t-s)\psi(\xi)},\\
\hat{G}^1(s,X;t,\xi) &= e^{iX\xi + (t-s)\psi(\xi)} g_0^{(1)}(t-s,\xi),
\end{aligned}$$

and $g_0^{(1)}$ as in (2.12). Using the Lagrangian remainder of the Taylor expansion, we have

$$\begin{aligned}
L - L_0 &= \gamma'(\varepsilon')(X-x)(\partial_X - 1) + a'(\varepsilon')(X-x)(\partial_{XX} - \partial_X) + \eta'(\varepsilon')(X-x)\int_{\mathbb{R}} \nu(dz)(e^z - 1 - z)\partial_X\\
&\quad + \eta'(\varepsilon')(X-x)\int_{\mathbb{R}} \nu(dz)(e^{z\partial_X} - 1 - z\partial_X),\\
L - L_1 &= \frac{1}{2}\gamma''(\varepsilon'')(X-x)^2(\partial_X - 1) + \frac{1}{2}a''(\varepsilon'')(X-x)^2(\partial_{XX} - \partial_X)\\
&\quad + \frac{1}{2}\eta''(\varepsilon'')(X-x)^2\int_{\mathbb{R}} \nu(dz)(e^z - 1 - z)\partial_X + \frac{1}{2}\eta''(\varepsilon'')(X-x)^2\int_{\mathbb{R}} \nu(dz)(e^{z\partial_X} - 1 - z\partial_X),
\end{aligned}$$

for some $\varepsilon', \varepsilon'' \in [x, X]$. Now, $|\hat{G}^0| \le 1$ because $\hat{G}^0$ is the characteristic function of the process $\tilde{X}$ in (2.21); thus, we have

$$\left|(L - L_1)\hat{G}^0(s,X_s;t,\xi)\right| \le C(1 + |\xi|^2)\,|X_s - x|^2.$$

On the other hand, from (2.12) we have

$$\left|g_0^{(1)}(t-s,\xi)\right| \leq C(t-s)^2\left(1+|\xi|^4\right),$$

and therefore we get

$$\left|(L-L_0)\hat{G}^1(s,X_s;t,\xi)\right| \leq C(t-s)^2(1+|\xi|^4)\,|X_s-x|\,.$$

So we find

$$\left|\hat{\Gamma}(0,x;t,\xi)-\hat{\Gamma}^{(1)}(0,x;t,\xi)\right| \leq C(1+|\xi|^4)\int_0^t\left((t-s)^2E\left[|X_s-x|\right]+E\left[|X_s-x|^2\right]\right)ds$$

The thesis then follows from estimate (2.27) and integrating. □

Remark 2.3.2. The proof of Theorem 2.3.1 can be generalized to obtain error bounds for any $n\in\mathbb{N}$: however, one can see that, for $n\geq 2$, the order of convergence improves only in the diffusive part, according to the results proved in [52].

2.4 Numerical tests

For the numerical examples we use the second-order approximation of the characteristic function. We have found this to be sufficiently accurate by numerical experiments and theoretical error estimates. The formulas for the second-order approximation are simple, making the method easy to implement. For the COS method, unless otherwise mentioned, we use $N=200$ and $L=10$, where L is the parameter used to define the truncation range $[a,b]$ as follows:

$$[a,b]:=\left[c_1-L\sqrt{c_2+\sqrt{c_4}},c_1+L\sqrt{c_2+\sqrt{c_4}}\right], \tag{2.28}$$

where c_n is the nth cumulant of log-price process $\log S$, as proposed in [30]. The cumulants are calculated using the 0th-order approximation of the characteristic function. A larger N and L has little effect on the price, since a fast convergence is achieved already for small N and L. We compare the approximated values to a 95% confidence interval computed with a Longstaff-Schwartz method with 10^5 simulations and 250 time steps per year. Furthermore, in the expansion we always use $\bar{x}=x$.

2.4.1 Tests under CEV-Merton dynamics

Consider a process under the CEV-Merton dynamics:

$$dX_t=\left(r-a(x)-\lambda\left(e^{m+\delta^2/2}-1\right)\right)dt+\sqrt{2a(x)}dW_t+\int_{\mathbb{R}}d\tilde{N}_t(t,dz)z,$$

with

$$
\begin{aligned}
a(x) &= \frac{\sigma_0^2 e^{2(\beta-1)x}}{2}, \\
\nu(dz) &= \lambda \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(\frac{-(z-m)^2}{2\delta^2}\right) dz, \\
\psi(\xi) &= -a_0(\xi^2 + i\xi) + ir\xi - i\lambda\left(e^{m+\delta^2/2} - 1\right)\xi + \lambda\left(e^{mi\xi - \delta^2\xi^2/2} - 1\right).
\end{aligned}
$$

We use the following parameters $S_0 = 1$, $r = 5\%$, $\sigma_0 = 20\%$, $\beta = 0.5$, $\lambda = 30\%$, $m = -10\%$, $\delta = 40\%$ and compute the European and Bermudan option values.

Table 2.1: Prices for a European and a Bermudan Put option (expiry $T = 0.25$ with 3 exercise dates, expiry $T = 1$ with 10 exercise dates and expiry $T = 2$ with 20 exercise dates) in the CEV-Merton model for the 2nd-order approximation of the characteristic function, and a Monte Carlo method.

		European		Bermudan	
T	K	MC 95% c.i.	Value	MC 95% c.i.	Value
0.25	0.6	0.001240-0.001433	0.001326	0.001243-0.001431	0.001307
	0.8	0.005218-0.005679	0.005493	0.005314-0.005774	0.005421
	1	0.04222-0.04321	0.04275	0.04274-0.04371	0.04304
	1.2	0.1923-0.1938	0.1935	0.1979-0.1989	0.1981
	1.4	0.3856-0.3872	0.3866	0.3948-0.3958	0.3955
	1.6	0.5812-0.5829	0.5825	0.5940-0.5950	0.5941
1	0.6	0.006136-0.006573	0.006579	0.006307-0.006729	0.006096
	0.8	0.02526-0.02622	0.02581	0.02617-0.02711	0.02520
	1	0.08225-0.08395	0.08250	0.08480-0.08640	0.08593
	1.2	0.1965-0.1989	0.1977	0.2097-0.2115	0.2132
	1.4	0.3560-0.3589	0.3574	0.3946-0.3957	0.3954
	1.6	0.5341-0.5385	0.5364	0.5930-0.5941	0.5932
2	0.6	0.01444-0.01513	0.01529	0.01528-0.01594	0.01365
	0.8	0.04522-0.04655	0.04613	0.04596-0.04719	0.04659
	1	0.1046-0.1067	0.1077	0.1149-0.1168	0.1171
	1.2	0.2054-0.2083	0.2065	0.2319-0.2341	0.2345
	1.4	0.3351-0.3386	0.3382	0.3968-0.3987	0.3991
	1.6	0.4904-0.4944	0.4919	0.5927-0.5938	0.5935

We present the results in Table 2.1. The option value for both the Bermudan options as well as the

European options appears to be accurate. Since the COS method has a very quick convergence, already for $N = 64$ the error becomes stable. For at-the-money strikes we have $\log_{10}|\text{error}| \approx 3.5$. The use of the second-order approximation of the characteristic function is justified by the fact that the option value (and thus the error) stabilizes starting from the second-order approximation. Furthermore, it is noteworthy that the 0th-order approximation is already very accurate.

The computer used in the experiments has an Intel Core i7 CPU with a 2.2 GHz processor. The CPU time of the calculations depends on the number of exercise dates. Assuming we use the second-order approximation of the characteristic function, if we have M exercise dates the CPU time will be $5 \cdot M$ ms.

Remark 2.4.1. The method can be extended to include time-dependent coefficients. The accuracy and speed of the method will be of the same order as for time-independent coefficients.

Remark 2.4.2. The Greeks can be calculated at almost no additional cost using the formulas presented in Remark 2.2.8. Numerically, the order of convergence is algebraic and is the same for both the exact characteristic function as for the 2nd-order approximation.

2.4.2 Tests under the CEV-Variance-Gamma dynamics

Consider the jump process to be a Variance-Gamma process. The VG process, is obtained by replacing the time in a Brownian motion with drift θ and standard deviation ϱ, by a Gamma process with variance κ and unitary mean. The model parameters ϱ and κ allow to control the skewness and the kurtosis of the distribution of stock price returns. The VG density is characterized by a fat tail and is thus used as a model in situations where small and large asset values are more probable than would be the case for the lognormal distribution. The Lévy measure in this case is given by:

$$\nu(dx) = \frac{e^{-\lambda_1 x}}{\kappa x}\mathbb{1}_{\{x>0\}}dx + \frac{e^{\lambda_2 x}}{\kappa|x|}\mathbb{1}_{\{x<0\}}dx,$$

where

$$\lambda_1 = \left(\sqrt{\frac{\theta^2\kappa^2}{4} + \frac{\varrho^2\kappa}{2}} + \frac{\theta\kappa}{2}\right)^{-1}, \qquad \lambda_2 = \left(\sqrt{\frac{\theta^2\kappa^2}{4} + \frac{\varrho^2\kappa}{2}} - \frac{\theta\kappa}{2}\right)^{-1}.$$

Furthermore we have

$$\begin{aligned}
a(x) &= \frac{\sigma_0^2 e^{2(\beta-1)x}}{2}, \\
\mu(t,x) &= r + \frac{1}{\kappa}\log\left(1 - \kappa\theta - \frac{\kappa\varrho^2}{2}\right) - a(x), \\
\psi(\xi) &= -a_0(\xi^2 + i\xi) + ir\xi + i\frac{1}{\kappa}\log\left(1 - \kappa\theta - \frac{\kappa\varrho^2}{2}\right)\xi - \frac{1}{\kappa}\log\left(1 - i\kappa\theta\xi + \frac{\xi^2\kappa\varrho^2}{2}\right).
\end{aligned}$$

We use the following parameters $S_0 = 1$, $r = 5\%$, $\sigma_0 = 20\%$, $\beta = 0.5$, $\kappa = 1$, $\theta = -50\%$, $\varrho = 20\%$. The results for the European and Bermudan option are presented in Table 2.2.

Table 2.2: Prices for a European and a Bermudan Put option (10 exercise dates, expiry $T = 1$) in the CEV-VG model for the 2nd-order approximation of the characteristic function, and a Monte Carlo method.

	European		Bermudan	
K	MC 95% c.i.	Value	MC 95% c.i.	Value
0.6	0.03090-0.03732	0.03546	0.03756-0.03876	0.03749
0.8	0.08046-0.08247	0.08029	0.08290-0.08484	0.08395
1	0.1507-0.1531	0.1511	0.1572-0.1600	0.1594
1.2	0.2501-0.2538	0.2522	0.2634-0.2668	0.2685
1.4	0.3831-0.3876	0.3847	0.4073-0.4108	0.4137
1.6	0.5430-0.5479	0.5436	0.5920-0.5938	0.5937

2.4.3 CEV-like Lévy process with a state-dependent measure and default

In this section we consider a model similar to the one used in [44]. The model is defined with local volatility, local default and a state-dependent Lévy measure as follows:

$$\begin{aligned}
a(x) &= \frac{1}{2}(b_0^2 + \epsilon_1 b_1^2 \eta(x)), \\
\gamma(x) &= c_0 + \epsilon_2 c_1 \eta(x), \\
\nu(x, dz) &= \epsilon_3 \nu_N(dz) + \epsilon_4 \eta(x) \nu_N(dz), \\
\eta(x) &= e^{\beta x}.
\end{aligned} \tag{2.29}$$

We will consider Gaussian jumps, meaning that

$$\nu_N(dz) = \lambda \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(\frac{-(z-m)^2}{2\delta^2}\right) dz.$$

The regular CEV model has several shortcomings: the volatility for instance drops to zero as the underlying approaches infinity; also the model does not allow the underlying to experience jumps. This model tries to overcome these shortcomings, while still retaining CEV-like behaviour through $\eta(x)$. The local volatility function $\sigma(x)$ behaves asymptotically like the CEV model, $\sigma(x) \sim \sqrt{\epsilon_1} b_1 e^{\beta x/2}$ as $x \to -\infty$, reflecting the fact that the volatility tends to increase as the asset price drops (the leverage effect). Jumps of size dz arrive with a state-dependent intensity of $\nu(x, dz)$. Lastly, a default arrives with intensity $\gamma(x)$. The default function $\gamma(x)$ behaves asymptotically like $\epsilon_2 c_1 e^{\beta x}$ as $x \to -\infty$, reflecting the fact that a default is more likely to occur when the price goes down.

In Table 2.3 the results are presented for a model as defined in (2.29) without default, meaning that $c_0 = c_1 = 0$ and with a state-dependent jump measure, so $\nu(x, dz) = \eta(x)\nu_N(dz)$. In this case we have

$$\psi(\xi) = ir\xi - a_0(\xi^2 - i\xi) - \lambda\nu_0(e^{m+\delta^2/2} - 1)i\xi + \lambda\nu_0(e^{mi\xi-\delta^2\xi^2/2} - 1),$$

where $a_0 = \frac{1}{2}b_1^2 e^{\beta\bar{x}}$ and $\nu_0(dz) = e^{\beta\bar{x}}\nu_N(dz)$. The other parameters are chosen as: $b_1 = 0.15$, $b_0 = 0$, $\beta = -2$, $\lambda = 20\%$, $\delta = 20\%$, $m = -0.2$, $S_0 = 1$, $r = 5\%$, $\epsilon_1 = 1$, $\epsilon_3 = 0$, $\epsilon_4 = 1$, the number of exercise dates is 10 and $T = 1$. From the results for both the European option and the Bermudan option we see that the method performs very accurately, even for deeply in-the-money strikes.

In Table 2.4 the results are presented for the value of a defaultable Put option. In case of default prior to exercise the Put option payoff is 0, in case of no default the value is $(K - S_t)^+$, depending on the exercise time. We look at the model as defined in (2.29) with the possibility of default and consider state-independent jumps, meaning that we have $\gamma(x) = \eta(x)$ and $\nu(x, dz) = \nu_N(dz)$. We have

$$\psi(\xi) = ir\xi - a_0(\xi^2 - i\xi) + \gamma_0 i\xi - \gamma_0 - \lambda(e^{m+\delta^2/2} - 1)i\xi + \lambda(e^{mi\xi-\delta^2\xi^2/2} - 1),$$

where $a_0 = \frac{1}{2}b_1^2 e^{\beta\bar{x}}$ and $\gamma_0 = c_1 e^{\beta\bar{x}}$. The other parameters are $b_0 = 0$, $b_1 = 0.15$, $\beta = -2$, $c_0 = 0$, $c_1 = 0.1$, $S_0 = 1$, $r = 5\%$, $\epsilon_1 = 1$, $\epsilon_2 = 1$, $\epsilon_3 = 1$, $\epsilon_4 = 0$, the number of exercise dates is 10 and $T = 1$.

Table 2.3: Prices for a European and a Bermudan Put option (10 exercise dates, expiry $T = 1$) in the CEV-like model with state-dependent measure for the 2nd-order approximation characteristic function, and a Monte Carlo method.

	European		Bermudan	
K	MC 95% c.i.	Value	MC 95% c.i.	Value
0.8	0.01025-0.01086	0.009385	0.01068-0.01125	0.01024
1	0.04625-0.04745	0.04817	0.05141-0.05253	0.05488
1.2	0.1563-0.1582	0.1564	0.1942-0.1952	0.1952
1.4	0.3313-0.3334	0.3314	0.3927-0.3934	0.3930
1.6	0.5207-0.5229	0.5218	0.5919-0.5926	0.5920
1.8	0.7103-0.7124	0.7122	0.7906-0.7913	0.7910

Table 2.4: Prices for a European and a Bermudan Put option (10 exercise dates, expiry $T = 1$) in the CEV-like model with default for the 2nd-order approximation characteristic function, and a Monte Carlo method.

	European		Bermudan	
K	MC 95% c.i.	Value	MC 95% c.i.	Value
0.8	0.002905-0.003175	0.003061	0.005876-0.006245	0.006361
1	0.01845-0.01918	0.01893	0.03419-0.03506	0.03520
1.2	0.08148-0.08296	0.08297	0.1820-0.1827	0.1824
1.4	0.2184-0.2205	0.2173	0.3793-0.3801	0.3792
1.6	0.3867-0.3892	0.3841	0.5752-0.5763	0.5763
1.8	0.5597-0.5638	0.5556	0.7727-0.7739	0.7733

Chapter 3

Efficient XVA computation under local Lévy models

In the last Chapter we discussed how to price options, in particular Bermudan options, under the local Lévy model. The option prices were computed without any additional valuation adjustments; more precisely, we assumed that one could borrow and lend under the risk-free rate, and no further funding requirements or credit risks were involved. This resulted in a linear PDE, or equivalently by the Feynman-Kac theorem the expected value in (2.2), which we solved through the COS method and the approximate characteristic function. When computing prices in the presence of various other risks like funding and counterparty risks (i.e. under XVA), one needs to redefine the pricing partial differential equation (PDE) by constructing a hedging portfolio with cashflows that are consistent with the additional funding requirements. This has been done for unilateral CCR in [60], bilateral CCR and XVA in [11] and extended to stochastic rates in [47]. This results in a non-linear PDE.

Non-linear PDEs can be solved with e.g. finite-difference methods or the LSM for solving the corresponsing backward stochastic differential equation (BSDE). In [61] an efficient forward simulation algorithm that gives the solution of the non-linear PDE as an optimum over solutions of related but linear PDEs is introduced, with the computational cost being of the same order as one forward Monte Carlo simulation. The downside of these numerical methods is the computational time that is required to reach an accurate solution. An efficient alternative might be to use Fourier methods for solving the (non-)linear PDE or related BSDE, such as the COS method, as was introduced in [30], extended to Bermudan options in [31] and to BSDEs in [63]. In certain cases the efficiency of these methods is further increased due the ability to the use the fast Fourier transform (FFT).

We consider the exponential Lévy-type model with a state-dependent jump measure and propose an efficient Fourier-based method to solve for Bermudan derivatives, including options and swaptions, with XVA. We derive, in the presence of jumps, a non-linear partial integro-differential equation (PIDE) and its corresponding BSDE for an OTC derivative between the bank B and its counterparty C in the presence of CCR, bilateral collateralization, MVA, FVA and KVA. We extend the Fourier-based method known as the BCOS method, developed in [63], to solve the BSDE under Lévy models with non-constant coefficients. As this method requires the knowledge of the characteristic function of the forward process, which, in the case of the Lévy process with variable coefficients, is not known, we will use an approximation of the characteristic function obtained by the adjoint expansion method developed in [58], [53] and extended to the defaultable Lévy process with a state-dependent jump measure in the previous chapter (see also [9]). Compared to other state-of-the-art methods for calculating XVAs, like Monte Carlo methods and PDE solvers, our method is both more efficient and multipurpose. Furthermore we propose an alternative Fourier-based method for explicitly pricing the CVA term in case of unilateral CCR for Bermudan derivatives under the local Lévy model. The advantage of this method is that is allows us to use the FFT, resulting in a fast and efficient calculation. The Greeks, used for hedging CVA, can be computed at almost no additional cost.

The rest of the chapter is structured as follows. After a brief recap of the local Lévy model in Section 3.1, in Section 3.2 we derive the non-linear PIDE and corresponding BSDE for pricing contracts under XVA. In Section 3.3 we propose the Fourier-based method for solving this BSDE and in Section 3.4 this method is extended to pricing Bermudan contracts. In Section 3.4.2 an alternative FFT-based method for pricing and hedging the CVA term is proposed and Section 3.5 presents numerical examples validating the accuracy and efficiency of the proposed methods.

3.1 The model

We consider the model as defined in Section 1.2.2, but additionally assume the following factorization of the Lévy measure

$$\nu(t, X_{t-}, dq) = a(t, X_{t-})\nu(dq),$$

A small remark on notation: we have replaced dz with dq in order to avoid notational confusion with the BSDE notation in Section 3.2.2. Using the notion of default as defined in (3.16), the probability of default is

$$\mathrm{PD}(t) := \mathbb{P}(\zeta \leq t) = 1 - \mathbb{E}\left[e^{-\int_0^t \gamma(s,X_s)ds}\right]. \tag{3.1}$$

We assume furthermore

$$\int_{\mathbb{R}} e^{|q|} a(t,x)\nu(dq) < \infty.$$

Imposing that the discounted asset price $\tilde{S}_t := e^{-rt}S_t$ is a $\mathcal{G}$-martingale under the risk-neutral measure, we get the following restriction on the drift coefficient:

$$\mu(t,x) = \gamma(t,x) + r - \frac{\sigma^2(t,x)}{2} - a(t,x)\int_{\mathbb{R}} \nu(dq)(e^q - 1 - q), \tag{3.2}$$

with r being the risk-free (collateralized) rate. In the whole of the chapter we assume deterministic, constant interest rates, while the derivations can easily be extended to time-dependent rates. As before, the integro-differential operator of the process is given by,

$$\begin{aligned} Lu(t,x) =& \partial_t u(t,x) + \mu(t,x)\partial_x u(t,x) - \gamma(t,x)u(t,x) + \frac{\sigma^2(t,x)}{2}\partial_{xx}u(t,x) \\ &+ a(t,x)\int_{\mathbb{R}} \nu(dz)(u(t,x+q) - u(t,x) - z\partial_x u(t,x)). \end{aligned} \tag{3.3}$$

3.2 XVA computation

Consider a bank B and its counterparty C, both of them might default. Assume they enter into a contract paying $\Phi(S_t)$ at maturity. Let $\varphi(x) = \Phi(e^x)$, and assume the risk-neutral dynamics of the underlying as in (3.16) with the drift given by (3.2). Define $\hat{v}(t,x)$ to be the value to the bank of the (default risky) portfolio with valuation adjustments referred to as XVA and $v(t,x)$ to be the risk-free value. Note that the difference between these two values is called the *total valuation adjustment* and in our setting this consists of

$$\text{TVA} := \hat{v}(t,x) - v(t,x) = \text{CVA} + \text{DVA} + \text{KVA} + \text{MVA} + \text{FVA}. \tag{3.4}$$

The risk-free value $v(t,x)$ solves a linear PIDE:

$$\begin{aligned} Lv(t,x) &= rv(t,x), \\ v(T,x) &= \varphi(x), \end{aligned}$$

where L is given in (3.3). Assuming the dynamics in (3.16), this linear PIDE can be solved with the methods presented in [9].

3.2.1 Derivative pricing under CCR and bilateral CSA agreements

In [11], the authors derive an extension to the Black-Scholes PDE in the presence of a bilateral counterparty risk in a jump-to-default model with the underlying being a diffusion, using replication arguments that include the funding costs. In [47] this derivation is extended to a multivariate diffusion setting with stochastic rates in the presence of CCR, assuming that both parties B and C are subject to default. To mitigate the CCR, both parties exchange collateral consisting of the initial margin and the variation margin. The parties are obliged to hold regulatory capital, the cost of which is the KVA and face the costs of funding uncollateralized positions through collateralized markets, known as FVA. Both [11] and [47] extend the approach of [60], in which unilateral collateralization was considered. We extend the approach in Section 1.4.1 and the approach of the above authors to derive the value of $\hat{v}(t,x)$ when the underlying follows the jump-diffusion defined in (3.16). We assume a one-dimensional underlying diffusion and consider all rates to be deterministic and, for ease of notation, constant. We specify different rates, defined in Table 3.2.1, for different types of lending.

Rate	Definition	Rate	Definition
r	the risk-free rate	r_R	the rate received on funding secured by the underlying asset
r_D	the dividend rate in case the stock pays dividends	r_F	the rate received on unsecured funding
r_B	the yield on a bond of the bank B	r_C	the yield on the bond of the counterparty C
λ_B	$\lambda_B := r_B - r$	λ_C	$\lambda_C := r_C - r$
λ_F	$\lambda_F := r_F - r$	R_B	the recovery rate of the bank
R_C	the recovery rate of the counterparty		

Table 3.1: Definitions of the rates used throughout the chapter.

Assume that the parties B and C enter into a derivative contract on the spot asset that pays the bank B the amount $\varphi(X_T)$ at maturity T. The value of this derivative to the bank at time t is denoted by $\hat{v}(t,x,\mathcal{J}^B,\mathcal{J}^C)$ and depends on the value of the underlying X and the default states $\mathcal{J}^B$ and $\mathcal{J}^C$ of the bank B and counterparty C, respectively. Define I^{TC} to be the initial margin posted by the bank to the counterparty, I^{FC} the initial margin posted by the counterparty to the bank and $I^V(t)$ to be the variation margin on which a rate r_I is paid or received. The initial margin is constant throughout the duration of the contract. Let $K(t)$ be the regulatory capital on which a rate of r_K is paid/received.

The cashflows are viewed from the perspective of the bank B. At the default time of either the counterparty or the bank, the value of the derivative to the bank $\hat{v}(t,x)$ is determined with a mark-to-market rule M, which may be equal to either the derivative value $\hat{v}(t,x,0,0)$ prior to default or the risk-free derivative value $v(t,x)$, depending on the specifications in the ISDA master agreement. Denote by τ^B and τ^C the random default times of the bank and the counterparty respectively. We will use the notation $x^+ = \max(x,0)$ and $x^- = \min(x,0)$. In a situation in which the counterparty defaults, the bank is already in the possession of $I^V + I^{FC}$. If the outstanding value $M - (I^V + I^{FC})$ is negative, the bank has to pay the full amount $(M - I^V - I^{FC})^-$, while if the contract has a positive value to the bank, it will recover only $R_C(M - I^V - I^{FC})^+$. Using a similar argument in case the bank defaults, we find the following boundary conditions:

$$\theta_t^B := \hat{v}(t,x,1,0) = I^V(t) - I^{TC} + (M - I^V(t) + I^{TC})^+ + R^B(M - I^V(t) + I^{TC})^-,$$
$$\theta_t^C := \hat{v}(t,x,0,1) = I^V(t) + I^{FC} + R^C(M - I^V(t) - I^{FC})^+ + (M - I^V(t) - I^{FC})^-,$$

so that the portfolio value at default is given by

$$\theta_\tau = 1_{\tau^C<\tau^B}\theta_\tau^C + 1_{\tau^B<\tau^C}\theta_\tau^B,$$

with $\tau = \min(\tau^B, \tau^C)$. Further we introduce the default risky, zero-recovery, zero-coupon bonds (ZCBs) P^B and P^C with respective maturities T^B and T^C with face value one if the issuer has not defaulted, and zero otherwise. Assume the dynamics for P_t^B and P_t^C to be given by $P_t^B = \mathbb{1}_{\{\tau^B>t\}}e^{r_Bt}$ and $P_t^C = \mathbb{1}_{\{\tau^C>t\}}e^{r_Ct}$, so that

$$dP_t^B = r_B P_t^B dt - P_{t-}^B d\mathcal{J}_t^B,$$
$$dP_t^C = r_C P_t^C dt - P_{t-}^C d\mathcal{J}_t^C,$$

with $\mathcal{J}_t^B = \mathbb{1}_{\tau^B\leq t}$ and $\mathcal{J}_t^C = \mathbb{1}_{\tau^C\leq t}$, where the default times τ^B and τ^C are defined in a canonical way as the first arrival time of a doubly stochastic Poisson process with intensity functions γ^B and γ^C, respectively (see also the definition of the defaultable asset in (3.16)). We define the market interest rates for B and C to be $r_B = r + \gamma^B$ and $r_C = r + \gamma^C$, so that by the usual arguments (see, for instance, [48, Section 2.2]) the discounted bonds $e^{-rt}P_t^B$ and $e^{-rt}P_t^C$ are martingales under the risk-neutral measure.

We construct a hedging portfolio consisting of the shorted derivative, α_C units of P^C, α_B units of P^B and g units of cash:

$$\Pi(t) = -\hat{v}(t,x) + \alpha_B(t)P_t^B + \alpha_C(t)P_t^C + g(t).$$

In other words, since we assume both the underlying asset process and the tradeable bonds P_B and P_C to be risk-neutral, we focus on hedging the risk arising from the defaults of both B and C by means of the default-risky bonds.

If the value of the derivative is positive to B, it will incur a cost at the counterparties' default. To hedge this, B shorts P^C, i.e. $\alpha_C \leq 0$. If we assume B can borrow the bond close to the risk-free rate r (i.e. no haircut) through a repurchase agreement, it will incur financing costs of $r\alpha_C(t)P_t^C dt$. The cashflows from the collateralization follow from the rate r_{TC} received and r_{FC} paid on the initial margin and the rate r_I paid or received on the collateral, depending on whether $I^V > 0$, and the bank receives collateral, or $I^V < 0$, and the bank pays collateral respectively. From holding the regulatory capital we incur a cost of $r_K K(t)$. Finally, the rates r and r_F are respectively received or paid on the surplus cash in the account. This cash consists of the gap between the shorted derivative value and the collateral and the cost of buying α_B bonds P^B in order for B to hedge its own default, i.e. $-\hat{v}(t,x) - I^V(t) + I^{TC} - \alpha_B(t)P_t^B$. Thus, the total change in the cash account is given by

$$\begin{aligned} dg(t) =& [-r\alpha_C(t)P_t^C + r_{TC}I_{TC} - r_{FC}I_{FC} - r_I I^V(t) - r_K K(t) \\ & + r(-\hat{v}(t,x) - I^V(t) + I_{TC} - \alpha_B(t)P_t^B) + \lambda_F(-\hat{v}(t,x) - I^V(t) + I_{TC} - \alpha_B(t)P_t^B)^-]dt. \end{aligned}$$

Note that this is in contrast with the change in cash in a portfolio without the XVA arising from the different types of funding, i.e. where we assume the cash in the portfolio simply earns the risk-free rate

$$dg(t) = -r\hat{v}(t,x)dt.$$

Assuming the portfolio is self-financing we have

$$d\Pi(t) = -d\hat{v}(t,x) + \alpha_B(t)dP_t^B + \alpha_C(t)dP_t^C + dg(t).$$

Applying Itô's Lemma to $\hat{v}(t,x)$ gives us:

$$\begin{aligned} d\hat{v}(t,x) =& L\hat{v}(t,x)dt + \sigma(t,x)\partial_x\hat{v}(t,x)dW_t + \int_{\mathbb{R}}(\hat{v}(t,x+q) - \hat{v}(t,x))d\tilde{N}(t,x,dq) \\ & - (\theta^B - \hat{v}(t,x))d\mathcal{J}_t^B - (\theta^C - \hat{v}(t,x))d\mathcal{J}_t^C, \end{aligned}$$

with the operator L as in (3.3). Thus, we find,

$$\begin{aligned} d\Pi =& - L\hat{v}(t,x)dt - \sigma(t,x)\partial_x\hat{v}(t,x)dW_t - \int_{\mathbb{R}}(\hat{v}(t,x+q) - \hat{v}(t,x))d\tilde{N}(t,x,dq) \\ & + (\theta^B - \hat{v}(t,x))d\mathcal{J}_t^B + (\theta^C - \hat{v}(t,x))d\mathcal{J}_t^C - \alpha^B(t)P_{t-}^B d\mathcal{J}_t^B - \alpha^C(t)P_{t-}^C d\mathcal{J}_t^C \\ & + [\alpha^B(t)\lambda_B P_t^B + \alpha^C(t)\lambda_C P_t^C + (r_{TC} + r)I^{TC} - r_{FC}I^{FC} - (r_I + r)I^V(t) \\ & - r_K K(t) + r\hat{v}(t,x) + \lambda_F(-\hat{v}(t,x) - I^V(t) + I^{TC} - \alpha^B(t)P_t^B)^-]dt. \end{aligned}$$

By choosing

$$\alpha_B = -\frac{\theta^B - \hat{v}(t,x)}{P_B}, \quad \alpha_C = -\frac{\theta^C - \hat{v}(t,x)}{P_C},$$

we hedge the jump-to-default risk in the hedging portfolio, i.e.,

$$\begin{aligned} d\Pi =& - L\hat{v}(t,x)dt + \sigma(t,x)\partial_x\hat{v}(t,x)dW_t - \int_{\mathbb{R}}(\hat{v}(t,x+q) - \hat{v}(t,x))d\tilde{N}(t,X_{t-},dq) \\ & + [-(\theta^B - \hat{v}(t,x))\lambda_B - (\theta^C - \hat{v}(t,x))\lambda_C + (r_{TC} + r)I^{TC} - r_{FC}I^{FC} - (r_I + r)I^V(t) \\ & - r_K K(t) + r\hat{v}(t,x) + \lambda_F(\theta^B - I^V(t) + I^{TC})^-]dt. \end{aligned}$$

Then, using the fact that the portfolio has to satisfy the martingale condition in the risk-neutral world, i.e. $\mathbb{E}[d\Pi] = 0$, we find the non-linear pricing PIDE to be

$$L\hat{v}(t,x) = f(t,x,\hat{v}(t,x)), \tag{3.5}$$

where we have defined

$$f(t,x,\hat{v}(t,x)) = -(\theta^B(t)-\hat{v}(t,x))\lambda_B - (\theta^C(t)-\hat{v}(t,x))\lambda_C + (r_{TC}+r)I^{TC} - r_{FC}I^{FC} \\ - (r_I+r)I^V(t) - r_K K(t) + r\hat{v}(t,x) + \lambda_F(\theta^B - I^V(t) + I^{TC})^-.$$

3.2.2 BSDE representation

In this section we will cast the PIDE in (3.5) in the form of a Backward Stochastic Differential Equation. In the methods where we make use of BSDEs we assume $\gamma(t,x)=0$. We begin by recalling the non-linear Feynman-Kac theorem in the presence of jumps, see Theorem 4.2.1 in [27].

Theorem 3.2.1 (Non-linear Feynman-Kac Theorem)**.** *Consider X_t as in (3.16). We assume μ, σ and a to be Lipschitz continuous in x and additionally $|a(t,x)| \le K$. Consider the BSDE*

$$Y_t = \varphi(X_T) + \int_t^T f\left(s, X_s, Y_s, Z_s, a(s,X_{s-})\int_{\mathbb{R}} V_s(q)\delta(q)\nu(dq)\right)ds - \int_t^T Z_s dW_s \\ - \int_t^T \int_{\mathbb{R}} V_s(q) d\tilde{N}_s(s, X_s, q), \tag{3.6}$$

where the generator f is continuous and satisfies the Lipschitz condition in the space variables, δ is a measurable, bounded function and the terminal condition $\varphi(x)$ is measurable and Lipschitz continuous. Consider the non-linear PIDE

$$\begin{cases} Lv(t,x) = f(t,x,v(t,x), \partial_x v(t,x)\sigma(t,x), a(t,x)\int_{\mathbb{R}}(v(t,x+q)-v(t,x))\delta(q)\nu(dq)), \\ v(T,x) = \psi(x). \end{cases} \tag{3.7}$$

If the PIDE in (3.7) has a solution $v(t,x) \in C^{1,2}$, the FBSDE in (3.6) has a unique solution $(Y_t, Z_t, V_t(q))$ that can be represented as

$$\begin{aligned} Y_s^{t,x} &= u(s, X_s^{t,x}), \\ Z_s^{t,x} &= \partial_x u(s, X_s^{t,x})\sigma(s, X_s^{t,x}), \\ V_s^{t,x}(q) &= u(s, X_s^{t,x}+q) - u(s, X_s^{t,x}), \qquad q \in \mathbb{R}, \end{aligned}$$

for all $s \in [t,T]$, where Y is a continuous, real-valued and adapted process and where the control processes Z and V are continuous, real-valued and predictable.

In our case, the BSDE corresponding to the PIDE in (3.5) reads

$$Y_t = \varphi(X_T) + \int_t^T f(s, X_s, Y_s) ds - \int_t^T Z_s dW_s - \int_t^T \int_{\mathbb{R}} V_s(q) d\tilde{N}(s, X_s, dq), \tag{3.8}$$

where we have defined the driver function to be

$$f(t,x,y) = -\lambda_B(\theta^B - y) - \lambda_C(\theta^C - y) + (r_{TC}+r)I^{TC} - r_{FC}I^{FC} - (r_I+r)I^V(t)$$

$$- r_K K(t) + ry + \lambda_F(\theta^B - I^V(t) + I^{TC})^-.$$

We remark here, that when assuming that one can borrow and lend simply at the risk-free rate, and no additional credit risks are considered, the driver function would be given by

$$f(t, x, \hat{v}(t, x)) = rv(t, x),$$

in this way relating the above derivation to the PDE considered in Chapter 2.

3.2.3 A simplified driver function

Following [36], one can derive that the KVA is a function of trade properties (i.e. maturity, strike) and/or the exposure at default, which in turn is a function of the portfolio value, so that the cost of holding the capital can be rewritten as $r_K K(t) = r_K c_1 \hat{v}(t, x)$, with c_1 being a function of the trade properties. The collateral is paid when the portfolio has a negative value, and received when the portfolio has a positive value. Assuming the collateral is a multiple of the portfolio value we have $I^V(t) = c_2 \hat{v}(t, x)$, where c_2 is some constant. Then, the driver function is simply a function of the portfolio value.

Remark 3.2.2. Note that in the case of 'no collateralization' or 'perfect collateralization', the driver function reduces to $f(t, \hat{v}(t, x)) = r_u(t) \max(\hat{v}(t, x), 0)$, for a function r_u here left unspecified. In this case the BSDE is similar to the one considered in [61].

3.3 Solving FBSDEs

In this section we extend the BCOS method from [63] to solving FBSDEs under local Lévy models with variable coefficients and jumps (without default, i.e. $\gamma(t, x) = 0$). The conditional expectations resulting from the discretization of the FBSDE are approximated using the COS method. This requires the characteristic function, which we approximate using the Adjoint Expansion Method of [58] and Section 2.1.1.

3.3.1 Discretization of the BSDE

Consider the forward process X_t as in (3.16) and the BSDE Y_t as in (3.8) with a more general driver function $f(t, x, y, z)$. Define a partition $0 = t_0 < t_1 < ... < t_N = T$ of $[0, T]$ with a fixed time step $\Delta t = t_{n+1} - t_n$, for $n = N - 1, ...0$. Rewriting the set of FBSDEs we find,

$$X_{n+1} = X_n + \int_{t_n}^{t_{n+1}} \mu(s, X_s) ds + \int_{t_n}^{t_{n+1}} \sigma(s, X_s) dW_s + \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}} q d\tilde{N}_s(s, X_{s-}, dq),$$

$$Y_n = Y_{n+1} + \int_{t_n}^{t_{n+1}} f(s, X_s, Y_s, Z_s) ds - \int_{t_n}^{t_{n+1}} Z_s dW_s - \int_{t_n}^{t_{n+1}} \int_{\mathbb{R}} V_s(q) d\tilde{N}_s(s, X_{s-}, dq). \quad (3.9)$$

One can obtain an approximation of the process Y_t by taking conditional expectations with respect to the underlying filtration $\mathcal{G}_n$, using the independence of W_t and $\tilde{N}_t(t, X_{t-}, dq)$ and by approximating the integrals that appear with a theta-method, as first done in [69] and extended to BSDEs with jumps in [63]:

$$Y_n \approx \mathbb{E}_n[Y_{n+1}] + \Delta t \theta_1 f\left(t_n, X_n, Y_n, Z_n\right) + \Delta t(1-\theta_1)\mathbb{E}_n\left[f\left(t_{n+1}, X_{n+1}, Y_{n+1}, Z_{n+1}\right)\right].$$

Let $\Delta W_s := W_s - W_n$ for $t_n \leq s \leq t_{n+1}$. Multiplying both sides of equation (3.9) by ΔW_{n+1}, taking conditional expectations and applying the theta-method gives

$$\begin{aligned} Z_n \approx & -\theta_2^{-1}(1-\theta_2)\mathbb{E}_n[Z_{n+1}] + \frac{1}{\Delta t}\theta_2^{-1}\mathbb{E}_n[Y_{n+1}\Delta W_{n+1}] \\ & + \theta_2^{-1}(1-\theta_2)\mathbb{E}_n\left[f\left(t_{n+1}, X_{n+1}, Y_{n+1}, Z_{n+1}\right)\Delta W_{n+1}\right]. \end{aligned}$$

Since in our scheme the terminal values are functions of time t and the Markov process X, it is easily seen that there exist deterministic functions $y(t_n, x)$ and $z(t_n, x)$ so that

$$Y_n = y(t_n, X_n), \quad Z_n = z(t_n, X_n).$$

The functions $y(t_n, x)$ and $z(t_n, x)$ are obtained in a backward manner using the following scheme

$$\begin{aligned} y(t_N, x) =& \varphi(x), \quad z(t_N, x) = \partial_x\varphi(x)\sigma(t_N, x), \\ & \text{for } n = N-1, ..., 0: \\ y(t_n, x) =& \mathbb{E}_n[y(t_{n+1}, X_{n+1})] + \Delta t\theta_1 f\left(t_n, x\right) + \Delta t(1-\theta_1)\mathbb{E}_n\left[f(t_{n+1}, X_{n+1})\right], && (3.10) \\ z(t_n, x) =& -\frac{1-\theta_2}{\theta_2}\mathbb{E}_n[z(t_{n+1}, X_{n+1})] + \frac{1}{\Delta t}\theta_2^{-1}\mathbb{E}_n[y(t_{n+1}, X_{n+1})\Delta W_{n+1}] && (3.11) \\ & + \frac{1-\theta_2}{\theta_2}\mathbb{E}_n\left[f(t_{n+1}, X_{n+1})\Delta W_{n+1}\right], \end{aligned}$$

where we have simplified notations with

$$f(t, X_t) := f\left(t, X_t, y(t, X_t), z(t, X_t)\right).$$

In the case $\theta_1 > 0$ we obtain an implicit dependence on $y(t_n, x)$ in (3.10) and we use P Picard iterations starting with initial guess $\mathbb{E}_n[y(t_{n+1}, X_{n+1})]$ to determine $y(t_n, x)$.

3.3.2 The characteristic function

Using the derivation as in Section 2.1.1, and defining

$$s_k = \frac{\partial_x^k s(\cdot, \bar{x})}{k!}, \qquad \gamma_k = \frac{\partial_x^k \gamma(\cdot, \bar{x})}{k!}, \qquad \mu_k(dq) = \frac{\partial_x^k \mu(\cdot, \bar{x})}{k!}, \qquad a_k = \frac{\partial_x^k a(\cdot, \bar{x})}{k!} \qquad k \geq 0.$$

we find for the zeroth and $k \geq 1$-th order approximation of the characteristic function

$$\hat{G}^0(t, x; T, \xi) = e^{i\xi x} e^{\int_t^T \psi(s,\xi) ds},$$

$$\hat{G}^k(t,x;T,\xi) = -\int_t^T e^{\int_s^T \psi(\tau,\xi)d\tau}\mathcal{F}\left(\sum_{h=1}^k \left(\tilde{L}_h^{(s,\cdot)}(s) - \tilde{L}_{h-1}^{(s,\cdot)}(s)\right) G^{k-h}(t,x;s,\cdot)\right)(\xi)ds,$$

with

$$\psi(t,\xi) = i\xi\mu_0(t) + s_0(t)\xi^2 + \int_{\mathbb{R}} a_0\nu(t,dq)(e^{iz\xi} - 1 - iz\xi),$$

$$\begin{aligned}\tilde{L}_h^{(t,y)}(t) - \tilde{L}_{h-1}^{(t,y)}(t) &= \mu_h(t)h(y-\bar{x})^{h-1} + \mu_h(t)(y-\bar{x})^h\partial_y - \gamma_h(t)(y-\bar{x})^h \\ &+ s_h(t)h(h-1)(y-\bar{x})^{h-2} + s_h(t)(y-\bar{x})^{h-1}\left(2h\partial_y + (y-\bar{x})\partial_{yy}\right) \\ &+ \int_{\mathbb{R}} a_h(t)\bar{\nu}(dq)\left((y+q-\bar{x})^h e^{q\partial_y} - (y-\bar{x})^h - q\left(h(y-\bar{x})^{h-1} - (y-\bar{x})^h\partial_y\right)\right),\end{aligned}$$

where $\bar{\nu}(dq) = \nu(-dq)$.

3.3.3 The COS formulae

The conditional expectations will be approximated using the usual COS method, as explained in Section 1.3 and as has been applied to FBSDEs with jumps in [63]. The conditional expectations arising in the equations (3.10)-(3.11) are all of the form $\mathbb{E}_n[h(t_{n+1}, X_{n+1})]$ or $\mathbb{E}_n[h(t_{n+1}, X_{n+1})\Delta W_{n+1}]$. The COS formula for the first type of conditional expectation reads

$$\mathbb{E}_n^x[h(t_{n+1}, X_{n+1})] \approx {\sum_{j=0}^{J-1}}' H_j(t_{n+1})\mathrm{Re}\left(\hat{\Gamma}\left(t_n, x; t_{n+1}, \frac{j\pi}{b-a}\right)\exp\left(ij\pi\frac{-a}{b-a}\right)\right),$$

where $\sum'$ denotes an ordinary summation with the first term weighted by one-half, $J > 0$ is the number of Fourier-cosine coefficients we use, $H_j(t_{n+1})$ denotes the jth Fourier-cosine coefficients of the function $h(t_{n+1}, x)$ and $\hat{\Gamma}(t_n, x; t_{n+1}, \xi)$ is the conditional characteristic function of the process X_{n+1} given $X_n = x$. For the second type of conditional expectation, using integration by parts, we obtain

$$\begin{aligned}&\mathbb{E}_n^x[h(t_{n+1}, X_{n+1})\Delta W_n] \\ &\quad\approx \Delta t\sigma(t_n, x){\sum_{j=0}^{J-1}}' H_j(t_{n+1})\mathrm{Re}\left(i\frac{j\pi}{b-a}\hat{\Gamma}\left(t_n, x; t_{n+1}, \frac{j\pi}{b-a}\right)\exp\left(ij\pi\frac{-a}{b-a}\right)\right).\end{aligned}$$

See [63] for the full derivations.

Remark 3.3.1. Note that these formulas are obtained by using an Euler approximation of the forward process and using the 2nd-order approximation of the characteristic function of the actual process. We have found this to be more exact than using the characteristic function of the Euler process, which is equivalent to using just the 0th-order approximation of the characteristic function.

Finally we need to approximate the Fourier-cosine coefficients $H_j(t_{n+1})$ of $h(t_{n+1}, x)$ at time points t_n, where $n = 0, ..., N$. The Fourier-cosine coefficient of h at time t_{n+1} is defined by

$$H_j(t_{n+1}) = \frac{2}{b-a} \int_a^b h(t_{n+1}, x) \cos\left(j\pi \frac{x-a}{b-a}\right) dx.$$

Due to the structure of the approximated characteristic function of the local Lévy process, see (3.18), the coefficients of the functions $z(t_{n+1}, x)$ and the explicit part of $y(t_{n+1}, x)$ can be computed using the FFT algorithm, as done in Section 2.2.1, because of the matrix in (2.19) being of a certain form with constant diagonals. In order to determine $F_j(t_{n+1})$, the Fourier-Cosine coefficient of the function

$$f\left(t_{n+1}, x, y(t_{n+1}, x), z(t_{n+1}, x)\right),$$

due to the intricate dependence on the functions z and y we choose to approximate the integral in F_j by a discrete Fourier-Cosine transform (DCT). For the DCT we compute the integrand, and thus the functions $z(t_{n+1}, x)$ and $y(t_{n+1}, x)$, on an equidistant x-grid. Note that in this case we can easily approximate *all* Fourier-Cosine coefficients with a DCT (instead of the FFT). If we take J grid points defined by $x_i := a + (i + \frac{1}{2})\frac{b-a}{J}$ and $\Delta x = \frac{b-a}{J}$ we find, using the mid-point integration rule, the approximation

$$H_j(t_{n+1}) \approx \frac{2}{J} {\sum_{i=0}^{J-1}}' h(t_{n+1}, x_i) \cos\left(j\pi \frac{2i+1}{2J}\right),$$

which can be calculated using the DCT algorithm, with a computational complexity of $O(J \log J)$. Note that the truncation range is defined as in (2.28).

3.4 XVA computation for Bermudan derivatives

The method in Section 3.3 allows us to compute the XVA as in (3.4), consisting of CVA, DVA, MVA, KVA and FVA. In this section, we apply this method for computing Bermudan derivatives, as defined in Section 2.2 with XVA. The resulting method – the solution of the non-linear XVA PDE through a BSDE-type method – is an efficient alternative to finite-difference methods as well as to the Monte-Carlo based method developed in [61]. The efficiency is both due to the availability of the characteristic function in closed form through the adjoint expansion method and the fast convergence of the COS method. Furthermore, in finite difference methods complications may arise in the implementation of the scheme for jump diffusions. Since our proposed method works in the Fourier space, the jump component is easily handled by means of an additional term in the characteristic function and does not cause any further difficulties.

For the CVA component in the XVA we develop an alternative method, which due to the ability of the FFT, results in a particularly efficient computation.

3.4.1 XVA computation

Consider an OTC derivative contract between the bank B and the counterparty C on the underlying asset S_t given by (3.16) with $\gamma(t,x) = 0$ with a Bermudan-type exercise possibility: there is a finite set of so-called exercise moments $\{t_1, ..., t_M\}$ prior to the maturity, with $0 \leq t_1 < t_2 < \cdots < t_M = T$. The payoff from the point-of-view of bank B is given by $\varphi(t_m, X_{t_m})$. Denote $\hat{v}(t,x)$ to be the risky Bermudan option value and $c(t,x)$ the continuation value. By the dynamic programming approach, the value for a Bermudan derivative with XVA and M exercise dates $t_1, ..., t_M$ can be expressed by a backward recursion as

$$\hat{v}(t_M, x) = \varphi(t_M, x),$$

and the continuation value solves the non-linear PIDE defined in (3.5)

$$\begin{cases} \begin{cases} Lc(t,x) = f(t,x,c(t,x)), & t \in [t_{m-1}, t_m[\\ c(t_m, x) = \hat{v}(t_m, x) \end{cases} \\ \hat{v}(t_{m-1}, x) = \max\{\Phi(t_{m-1}, x), c(t_{m-1}, x)\}, \quad m \in \{2, \ldots, M\}. \end{cases}$$

The derivative value is set to be $\hat{v}(t,x) = c(t,x)$ for $t \in]t_{m-1}, t_m[$, and, if $t_1 > 0$, also for $t \in [0, t_1[$. The payoff function might take on various forms:

1. (Portfolio) Following [61], we can consider X_t to be the process of a portfolio which can take on both positive and negative values. Then, when exercised at time t_m, bank B receives the portfolio so that $\varphi(t_m, x) = e^x$.

2. (Bermudan option) In case the Bermudan contract is an option, the option value to the bank can not have a negative value for the bank. At the same time, in case of default of the bank itself, the counterparty loses nothing. In this case the framework simplifies to one with unilateral collateralization and default risk and the payoff at time t_m, if exercised, is given by $\varphi(t_m, x) = (K - e^x)^+$ for a put and $\varphi(t_m, x) = (e^x - K)^+$ for a call with K being the strike price.

3. (Swaptions) A swaption is an option in which the holder, bank B, has the right to exercise and enter into an underlying swap with fixed end date T_M. If the swaption is exercised at time T_m the underlying swap starts with payment dates $\mathcal{T}_m = \{T_{m+1}, ..., T_M\}$. We refer to the Appendix for more details on valuing this kind of instrument with XVA.

To solve for the continuation value we define a partition with N steps $t_{m-1} = t_{0,m} < t_{1,m} < t_{2,m} < ... < t_{n,m} < ... < t_{N,m} = t_m$ between two exercise dates t_{m-1} and t_m, with fixed time step $\Delta t_n := t_{n+1,m} - t_{n,m}$. Applying the method developed in Section 3.3, we find the following time iteration for the continuation value:

- At time $t_{N,m}$ set:

$$c(t_{N,m},x)=\hat{u}(t_m,x).$$

- For $n=N-1,...,0$ compute:

$$\begin{aligned}c(t_{n,m},x)\approx&\Delta t_n\theta_1 f(t_{n,m},x,c(t_{n,m},x))\\&+\sum_{j=0}^{J-1}{}' \Psi_j(x)(C_j(t_{n+1,m})+\Delta t_n(1-\theta_1)F_j(t_{n+1,m})),\end{aligned}\tag{3.12}$$

where we have defined

$$\Psi_j(x)=\mathrm{Re}\left(\hat{\Gamma}\left(t_{n,m},x;t_{n+1,m},\frac{j\pi}{b-a}\right)\exp\left(ij\pi\frac{-a}{b-a}\right)\right),$$

and the Fourier-cosine coefficients are given by

$$\begin{aligned}C_j(t_{n+1,m})&=\frac{2}{b-a}\int_a^b c(t_{n+1,m},x)\cos\left(j\pi\frac{x-a}{b-a}\right)dx,\\F_j(t_{n+1,m})&=\frac{2}{b-a}\int_a^b f(t_{n+1,m},x,c(t_{n+1,m},x))\cos\left(j\pi\frac{x-a}{b-a}\right)dx.\end{aligned}$$

In order to determine the function $c(t_n,x)$, we will perform P Picard iterations. To evaluate the coefficients with a DCT we need to compute the integrands $c(t_{n+1,m},x)$ and $f(t_{n+1,m},x,c(t_{n+1,m},x))$ on the equidistant x-grid with x_i, for $i=0,...,J-1$. In order to compute this at each time step $t_{n,m}$ we thus need to evaluate $c(t_{n,m},x)$ on the x-grid with J equidistant points using formula (3.12). The matrix-vector product in the formula results in a computational time of order $O(J^2)$.

Remark 3.4.1 (Convergence of the Picard iterations)**.** A Picard iteration is used to find the fixed-point c of $c=\Delta t\theta_1 f(t_{n,m},x,c)+h(t_{n,m},x)$, where $f(t,x,c)$ and $h(t,x)$ are respectively the implicit and explicit parts of the equation. Due to the computational domain of $c(t,x)$ being bounded by $[a,b]$, we can thus say that $f(t,x,c(t,x))$ is also bounded. If the driver function $f(t,x,c)$ is Lipschitz continuous in c, i.e. $\exists\ L^{Lipz}$ such that $|f(t,x,c_1)-f(t,x,c_2)|\leq L^{Lipz}|c_1-c_2|$, and Δt_n is small enough such that $\Delta t\theta_1 L^{Lipz}<1$, a unique fixed-point exists and the Picard iterations converge towards that point for any initial guess. In particular, for the XVA case the non-linearity is of the form $f(t,x,c)=-r\max(c,0)$, and this is Lipschitz continuous with $L^{Lipz}=1$. Thus for Δt sufficiently small, the Picard iteration converges to a unique fixed-point.

The total algorithm for computing the value of a Bermudan contract with XVA can be summarised as in Algorithm 1 in Figure 3.1. The total computational time for the algorithm is of order

$$O(M\cdot N(J+J^2+PJ+J\log_2 J)),\tag{3.13}$$

consisting of the computation for $M \cdot N$ times the computation of the characteristic function on the x-grid (due to the availability of the analytical approximation) of $O(J)$, computation of the matrix-vector multiplications in the formulas for $c(t_{n,m}, x)$ and $z(t_{n,m}, x)$ of $O(J^2)$, initialization of the Picard method with $\mathbb{E}_n[c(t_{n+1}, X_{n+1}]$ in $O(J^2)$ operations, computation of the P Picard approximations for $c(t_{n,m}, x)$ in $O(PJ)$ and computing the Fourier coefficients $F_j(t_n)$ and $C_j(t_n)$ with the DCT in $O(J \log_2 J)$ operations.

1. Define the x-grid with J grid points given by $x_i = a + (i + \frac{1}{2})\frac{b-a}{J}$ for $i = 0, ..., J-1$.
2. Calculate the final exercise date values $c(t_{N,M}, x) = \hat{u}(t_M, x)$ on the x-grid and compute the terminal coefficients $C_j(t_M)$ and $F_j(t_M)$ using the DCT.
3. Recursively for the exercise dates $m = M-1, ..., 0$ do:
   (a) For time steps $n = N-1, ..., 0$ do:
      i. Compute $c(t_{n,m}, x)$ using formula (3.12) and use this to determine $f(t_{n,m}, x, c(t_{n,m}, x))$ on the x-grid.
      ii. Subsequently, use these to determine $F_j(t_{n,m})$ and $C_j(t_{n,m})$ using the DCT.
   (b) Compute the new terminal condition $c(t_{N,m-1}, x) = \max\{\varphi(t_{0,m}, x), c(t_{0,m}, x)\}$ (either analytically or numerically) and the corresponding Fourier-cosine coefficient.
4. Finally $\hat{u}(t_0, x_0) = c(t_{0,0}, x_0)$.

Figure 3.1: Algorithm for Bermudan derivative valuation with XVA

3.4.2 An alternative for CVA computation

In this section we present an efficient alternative way of calculating the CVA term in (3.4) in the case of unilateral CCR using a Fourier-based method. Due to the ability of using the FFT this method is considerably faster for computing the CVA than the method presented in Section 4.1. We use the definition of CVA at time t given by

$$\mathrm{CVA}(t) = \hat{v}(t, X_t) - v(t, X_t),$$

where $v(t, X_t)$ is as usual the default-free value of the Bermudan option ($\gamma(t, x) = 0$), while $\hat{v}(t, X_t)$ is the value including default ($\gamma(t, x) \neq 0$). We consider the model as defined in (3.16). We will compute $v(t, X_t)$ and $\hat{v}(t, X_t)$ using the COS method and the approximation of the characteristic function (as derived in Section 2.2), without default and with default, respectively. In case of a default the payoff becomes zero. Note that the risky option value $\hat{v}(t,x)$ computed with the characteristic function for a defaultable underlying corresponds exactly to the option value in which the counterparty might default, with the probablity of default, $PD(t)$, defined as in (3.17). Thus, in this case we have unilateral CCR and $\zeta = \tau_C$, the default time of the counterparty.

Using the definition of the defaultable S_t, it is well-known (see, for instance, [48, Section 2.2]) that the risky no-arbitrage value of the Bermudan option on the defaultable asset S_t at time t is given by

$$\hat{u}\left(t, X_t\right) = \mathbb{1}_{\{\zeta > t\}} \sup_{\tau \in \{t_1, \ldots, t_M\}} \mathbb{E}\left[e^{-\int_t^\tau (r+\gamma(s,X_s))ds} \varphi(\tau, X_\tau) | X_t\right].$$

Remark 3.4.2 (Wrong-way risk)**.** By allowing the dependence of the default intensity on the underlying, a simplified form of wrong-way risk is already incorporated into the CVA valuation.

For a Bermudan put option with strike price K, we simply have $\varphi(t,x) = (K-x)^+$. By the dynamic programming approach, the option value can be expressed by a backward recursion as

$$\hat{u}(t_M, x) = \mathbb{1}_{\{\zeta > t_M\}} \max(\varphi(t_M, x), 0),$$

and

$$\begin{aligned} c(t,x) &= \mathbb{E}\left[e^{\int_t^{t_m}(r+\gamma(s,X_s))ds} \hat{u}(t_m, X_{t_m}) | X_t = x\right], \qquad t \in [t_{m-1}, t_m[\\ \hat{u}(t_{m-1}, x) &= \mathbb{1}_{\{\zeta > t_{m-1}\}} \max\{\varphi(t_{m-1}, x), c(t_{m-1}, x)\}, \qquad m \in \{2, \ldots, M\}. \end{aligned}$$

Thus to find the risky option price $\hat{v}(t, X_t)$ one uses the defaultable asset with $\gamma(t,x)$ representing the default intensity of the counterparty and in order to get the default-free value $v(t, X_t)$ one uses the default-free asset by setting $\gamma(t,x) = 0$. The CVA adjustment is calculated as the difference between the two. Both $\hat{v}(t,x)$ and $v(t,x)$ are calculated using the approximated characteristic function and the COS method applied to the continuation value. Due to the characteristic function being of the form (3.18), we are able to use the FFT in the matrix-vector multiplication when computing the continuation values of the Bermudan option with and without default, reducing this operation from $O(J^2)$ to $O(J \log_2 J)$. The total complexity of the calculation of the CVA value for a Bermudan option with M exercise dates is then $O(MJ \log_2 J)$. Comparing this to (3.13), in which the most time-consuming operations were indeed the matrix-vector products of order $O(J^2)$ that resulted from the computation of the functions on the x-grid of size J, we conclude that the method for CVA computation is indeed significantly faster due to the ability of using the FFT.

Remark 3.4.3 (The defaultable and default-free characteristic functions)**.** To find $v(t,x)$ we use

$$\hat{\Gamma}^r(t_m, x; t_{m+1}, \xi) := e^{i\xi x} \sum_{h=0}^{n} (x - \bar{x})^h g^r_{n,h}(t_m, t_{m+1}, \xi),$$

the characteristic function with $\gamma(t,x)=0$. For $\hat{u}(t,x)$ we use

$$\hat{\Gamma}^d(t_m,x;t_{m+1},\xi) := e^{i\xi x}\sum_{h=0}^{n}(x-\bar{x})^h g^d_{n,h}(t_m,t_{m+1},\xi),$$

where $\gamma(t,x)$ is chosen to be some specified function.

Hedging CVA

In practice CVA is hedged and thus practitioners require efficient ways to compute the sensitivity of the CVA with respect to the underlying. The widely used bump- and revalue- method, while resulting in precise calculations, might be slow to compute. Using the Fourier-based approach we find explicit formulas allowing for an easy computation of the first- and second-order derivatives of the CVA with respect to the underlying. For the first-order and second-order Greeks we have

$$\begin{aligned}
\Delta = {} & e^{-r(t_1-t_0)}{\sum_{j=0}^{J-1}}{}' \operatorname{Re}\left(e^{ij\pi\frac{x-a}{b-a}}\left(\frac{ij\pi}{b-a}g^d_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right)+g^d_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\right)\right)V^d_j(t_1)\\
& - e^{-r(t_1-t_0)}{\sum_{j=0}^{J-1}}{}' \operatorname{Re}\left(e^{ij\pi\frac{x-a}{b-a}}\left(\frac{ij\pi}{b-a}g^r_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right)+g^r_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\right)\right)V^r_j(t_1),\\
\frac{\partial\Delta}{\partial X} = {} & e^{-r(t_1-t_0)}{\sum_{j=0}^{J-1}}{}' \operatorname{Re}\Bigg(e^{ij\pi\frac{x-a}{b-a}}\Bigg(-\frac{ij\pi}{b-a}g^d_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right)-g^d_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\\
& + 2\frac{ij\pi}{b-a}g^d_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right)+\left(\frac{ij\pi}{b-a}\right)^2 g^d_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right)+2g^d_{n,2}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\Bigg)\Bigg)V^d_j(t_1)\\
& - e^{-r(t_1-t_0)}{\sum_{j=0}^{J-1}}{}' \operatorname{Re}\Bigg(e^{ij\pi\frac{x-a}{b-a}}\Bigg(-\frac{ij\pi}{b-a}g^r_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right)-g^r_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\\
& - 2\frac{ij\pi}{b-a}g^r_{n,1}\left(t_0,t_1,\frac{j\pi}{b-a}\right)+\left(\frac{ij\pi}{b-a}\right)^2 g^r_{n,0}\left(t_0,t_1,\frac{j\pi}{b-a}\right)+2g^r_{n,2}\left(t_0,t_1,\frac{j\pi}{b-a}\right)\Bigg)\Bigg)V_j(t_1)^r,
\end{aligned}$$

where V^d_k and V^r_k are the Fourier-cosine coefficients with the defaultable and default-free characteristic function terms, $g^d_{n,h}$ and $g^r_{n,h}$, respectively.

3.5 Numerical experiments

In this section we present numerical examples to justify the accuracy of the methods in practice. We compute the XVA with the method presented in Section 3.4.1 and the CVA in the case of unilateral CCR with the method from Section 3.4.2, which we show is more efficient for cases in which one only needs to compute the CVA. We compare the results of solving the BSDE with the COS method and the adjoint expansion of the characteristic function to the values obtained by using a least-squares Monte-Carlo method for computing the conditional expected values in the BSDE as done in e.g. [6].

The computer used in the experiments has an Intel Core i7 CPU with a 2.2 GHz processor. We use the second-order approximation of the characteristic function. We have found this to be sufficiently accurate by numerical experiments and theoretical error estimates. The formulas for the second-order approximation are simple, making the methods easy to implement.

3.5.1 A numerical example for XVA

Here, we check the accuracy of the method from Section 3.4.1. We will compute the Bermudan option value with XVA using a simplified driver function given by $f(t, \hat{v}(t, x)) = -r \max(\hat{v}(t, x), 0)$. Our method is easily extendible to the full driver function from Section 3.2. Consider X_t to be a portfolio process and the payoff, if exercised at time t_m, to be given by $\Phi(t_m, x) = x$. In this case the value we can receive at every exercise date is the value of the portfolio. Consider the model in Section 3.1 without default, with a local jump measure and a local volatility function with CEV-like dynamics and Gaussian jumps defined by

$$\sigma(x) = be^{\beta x}, \tag{3.14}$$

$$\nu(x, dq) = \lambda e^{\beta x} \frac{1}{\sqrt{2\pi\delta^2}} \exp\left(\frac{-(q-m)^2}{2\delta^2}\right) dq. \tag{3.15}$$

We assume the following parameters in equations (3.14)-(3.15), unless otherwise mentioned: $b = 0.15$, $\beta = -2$, $\lambda = 0.2$, $\delta = 0.2$, $m = -0.2$, $r = 0.1$, $K = 1$ and $X_0 = 0$ (so that $S_0 = 1$). In the LSM the number of time steps is taken to be 100 and we simulate 10^5 paths. In the COS method we take $J = 256$, $\theta_1 = 0.5$ and $N = 10$, $M = 10$, making the total number of time steps $N \cdot M = 100$. The truncation range is determined as in (2.28) with $L = 10$. Due to the state-dependent coefficients in the underlying dynamics in (3.14)-(3.15) we use the approximated characteristic function with the second-order approximation, i.e. $\hat{\Gamma}^{(2)}(t, x; T, \xi)$ and take $\bar{x} = x$, where $x = \{x_i\}_{i=0}^{J-1}$. Note that we thus compute the values, including those of the characteristic function, on the complete x-grid. In the final iteration when computing $\hat{u}(t_0, X_0)$ we use $\bar{x} = X_0$.

In Table 3.5.1 we analyse the error in the approximation of $\hat{u}(t_0, X_0)$ with $S_0 = 0.4$ for different values of the discretization parameter N and the number of grid points (and Fourier-cosine coefficients) J. We compare the approximated COS value to the 95% confidence interval obtained by an LSM. Accurate results are quickly obtained for small values of both J and N. In Figure 3.2 we plot the upper bound of the 95% confidence interval of the absolute error in the approximation for varying J and N. We observe approximately a linear convergence and note that the error stops decreasing at some point for increasing values of J and N. This can be due to the error being dominated by the approximated characteristic function. In particular we observe that $J = 32$ and $N = 10$ seem to be sufficient parameters to achieve a satisfactory accuracy in the approximation.

The results for $\hat{u}(t_0, X_0)$ of the COS approximation method compared to a 95% confidence interval of the value obtained through an LSM are presented in Table 4.2. These results show that our method is able to solve non-linear PIDEs accurately. The CPU time of the approximating method depends on the number of time steps $M \cdot N$ and is approximately $5 \cdot (N \cdot M)$ ms.

	$N = 1$	$N = 10$	$N = 20$	$N = 30$
$J = 8$	6.4E-03−6.9E-03	4.3E-03−4.8E-03	4.9E-03−5.3E-03	5.3E-03−5.8E-03
$J = 16$	2.3E-03−2.7E-03	8.8E-04−1.3E-03	6.2E-04−1.1E-03	5.4E-04−9.2E-04
$J = 32$	1.7E-03−2.0E-03	4.2E-04−8.3E-04	2.4E-04−6.3E-04	1.6E04−5.8E-04
$J = 64$	1.4E-03−1.9E-03	2.2E-04−6.5E-04	1.6E-04−2.3E-04	1.2E-04−2.9E-04
$J = 128$	1.7E-04−6.0E-04	2.1E-04−6.6E-04	2.3E-04−6.5E-04	1.9E-04−6.1E-04
$J = 256$	2.1E-04−6.6E-04	3.7E-04−7.7E-04	1.5E-04−5.7E-04	1.2E-04−3.1E-04

Table 3.2: The 95% confidence interval of the absolute error in the COS approximation of $\hat{u}(0, X_0)$ with $S_0 = 0.4$ compared to an LSM for varying parameters J and N.

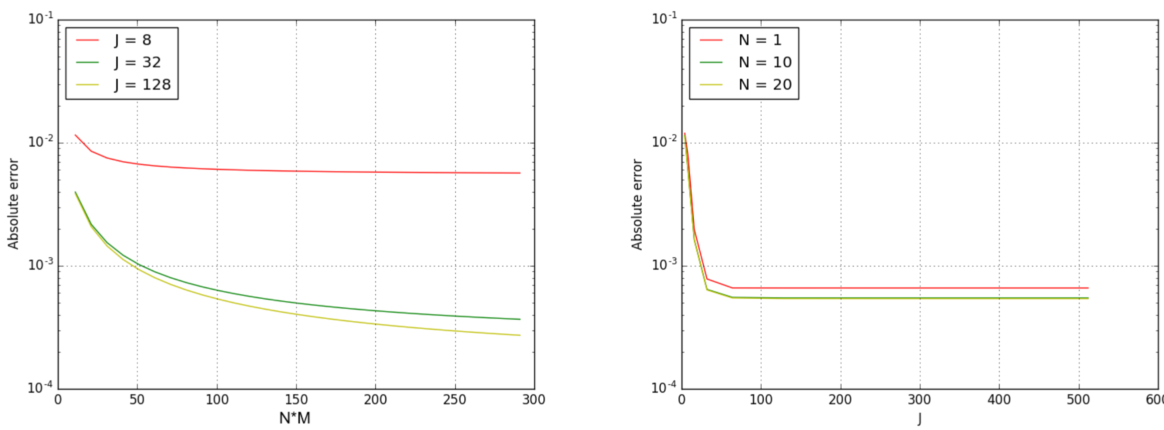


Figure 3.2: Convergence of the upper bound of the 95% confidence interval of the absolute error in the COS approximation $\hat{u}(0, X_0)$ with $S_0 = 0.4$ compared to a LSM for varying parameters J and N.

maturity T	S_0	MC value with XVA	COS value with XVA
0.5	0	0.03770−0.03838	0.03809
	0.2	0.2326−0.2330	0.2320
	0.4	0.4251−0.4254	0.4243
	0.6	0.6169−0.6171	0.6158
	0.8	0.8077−0.8079	0.8069
	1	1.000−1.000	1.0000
1	0	0.07374−0.07453	0.07228
	0.2	0.2611−0.2617	0.2606
	0.4	0.4461−0.4465	0.4454
	0.6	0.6288−0.6291	0.6288
	0.8	0.8126−0.8129	0.8113
	1	1.001−1.001	1.000

Table 3.3: A Bermudan put option with XVA (10 exercise dates, expiry $T = 0.5, 1$) in the CEV-like model for the 2nd-order approximation of the characteristic function, and an LSM comparison.

3.5.2 A numerical example for CVA

In this section we validate the accuracy of the method presented in Section 3.4.2 and compute the CVA in the case of unilateral CCR under the local Lévy dynamics with a local jump measure and a local volatility function with CEV-like dynamics, Gaussian jumps defined as in (3.15) and a local default function $\gamma(x) = ce^{\beta x}$. We assume the same parameters as in Section 3.5.1, except $r = 0.05$ and we take $c = 0.1$ in the default function. In the LSM the number of time steps is taken to be 100 and we simulate 10^5 paths. In the COS method we take $L = 10$ and $J = 100$. Again, due to the state-dependent coefficients in the underlying dynamics we use the approximated characteristic function as derived in Section 3.3.2 with the second-order approximation, i.e. $\hat{\Gamma}^{(2)}(t, x; T, \xi)$ and take $\bar{x} = X_0$.

The results for the CVA valuation with the FFT-based method and with LSM are presented in Table 3.4. The CPU time of the LSM is at least 5 times the CPU time of the approximating method, which for M exercise dates is approximately $3 \cdot M$ ms, thus more efficient than the computation of the XVA with the method in Section 4.1. The optimal exercise boundary in Figure 3.3 shows that the exercise region becomes larger when the probability of default increases; this is to be expected: in case of the default probability being greater, the option of exercising early is more valuable and used more often.

maturity T	strike K	MC CVA	COS CVA
0.5	0.6	$4.200 \cdot 10^{-4} - 4.807 \cdot 10^{-4}$	$1.113 \cdot 10^{-4}$
	0.8	0.001525−0.001609	$9.869 \cdot 10^{-4}$
	1	0.01254−0.01273	0.01138
	1.2	0.005908−0.005931	0.005937
	1.4	0.006657−0.06758	0.006898
	1.6	0.007795−0.008008	0.007883
1	0.6	8.673E-04−9.574E-04	4.463E-04
	0.8	0.005817−0.006040	0.003535
	1	0.02023−0.02054	0.01882
	1.2	0.01221−0.01222	0.1272
	1.4	0.01378−0.01391	0.01360
	1.6	0.01532−0.01502	0.01554

Table 3.4: CVA for a Bermudan put option (10 exercise dates, expiry $T = 0.5, 1$) in the CEV-like model for the 2nd-order approximation of the characteristic function, and an LSM comparison.

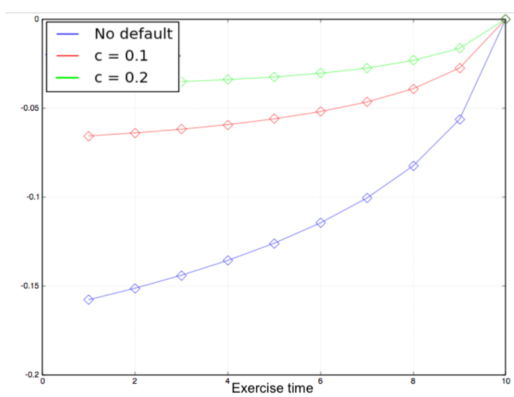


Figure 3.3: Optimal exercise boundary for a Bermudan put option (10 exercise dates, expiry $T = 1$) in the CEV-like model with varying default $c = 0, 0.1, 0.2$.

Appendix: Valuing swaptions with XVA under the local Lévy model

In the previous chapter we introduced a method for the computation of option prices including various valuation adjustments. In this Appendix we briefly show how the previous derivations can be extended to swaptions, in which the underlying interest rate is assumed to follow the local Lévy model. A swaption is a contract which gives the owner the right to enter into the underlying swap in which they either pay the fixed leg and receive the floating leg (a payer swaption), or receive the fixed leg and pay the floating leg (a receiver swaption). The buyer and seller of the swaption agree on the price of the swap, the length of the option period and the terms of the underlying swap, i.e. the notional amount on which the fixed and variable payments are computed, the fixed rate (also known as the strike of the swaption), and the frequency of observation of the floating leg. A common rate for the floating leg is the well-known Libor rate. Swaptions are used by financial institutions and banks to manage and hedge their interest rate risk, making them an actively traded and very liquid fixed income instrument.

To compute prices of these instruments, banks commonly use relatively simple models for the interest rate, which make both analytic and numerical methods for pricing the swaptions easy to implement, but these are not very flexible in modeling the real-world dynamics of the interest rate. In this chapter we therefore propose to model the interest rate using the flexible local Lévy model, and show that all maturity times T of the bond $P(t, T, x)$ depend on the same value of the short rate, which is similar to what happens in the Hull-White model and is a crucial part in an efficient evaluation of the swaption price. We furthermore extend the method of the previous chapter and show how one can efficiently compute the CVA for these derivatives.

The model We consider the stochastic credit risk-less interest rate r_t whose risk-neutral dynamics are given by the local Lévy model defined in Section 1.2.2 (without default), which we repeat here for convenience

$$r_t = e^{X_t},$$
$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t + \int_{\mathbb{R}} q d\tilde{N}_t(t, X_{t-}, dq),$$

$$d\tilde{N}_t(t, X_{t-}, dq) = dN_t(t, X_{t-}, dq) - a(t, X_{t-})\nu(t, dq)dt, \tag{3.16}$$

where $d\tilde{N}_t(t, X_{t-}, dq)$ is as usual the compensated random measure with state-dependent Lévy measure

$$\nu(t, X_{t-}, dq) = a(t, X_{t-})\nu(dq).$$

The default time ζ, which will be used for defining the defaultable zero-coupon bonds, is defined as in Section 1.2.2, so that the probability of default is

$$\mathrm{PD}(t) := \mathbb{P}(\zeta \leq t) = 1 - e^{-\int_0^t \alpha(s,x)ds}, \tag{3.17}$$

with $\alpha(t, x)$ now being the default intensity. Notice that the drift coefficient is not restricted. This drift coefficient takes into account the fixed equivalent martingale measure, so in the framework of martingale modeling, the selection of the equivalent martingale measure among all probability measures equivalent to $\mathbb{P}$ is an important task. Essentially, it can be considered a problem equivalent to calibration of the model. More precisely, since we know that the diffusion coefficient remains unchanged through a Girsanov change of measure, selecting $\mathbb{Q}$ is equivalent to estimating μ. Note however that μ represents the drift coefficient under the risk-neutral measure, so one cannot simply use the historical dynamics of r to estimate μ.

Default-free ZCB From now on we assume the coefficients are independent of t. We require an analytic formula for the ZCB which is related to the risk-neutral model for the underlying interest rate. Consider the money-market account given by

$$B(t) = e^{\int_0^t r_s ds}.$$

The value of a default-free zero-coupon bond between times t and T is defined as

$$P(t, T, x) = B(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{1}{B(T)}|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T \gamma(s,X_s)ds}|\mathcal{F}_t\right],$$

where we have defined $\gamma(s, X_s) := e^{X_s}$ and $P(t, T, x)$ is the price of the ZCB, conditional on $X_t = x$ associated with times t and T. Define furthermore the discount factor

$$D(t, T) = e^{-\int_t^T r_s ds},$$

that is unknown at time t since r is a progressively measurable stochastic process. Note the conceptual difference between $P(t, T, x)$ and $D(t, T)$: at maturity they have the same value $P(T, T, x) = D(T, T)$, but while $P(t, T, x)$ is a price, and as such is observable at time t (i.e. $\mathcal{F}_t$-measurable), the discount factor is $\mathcal{F}_T$-measurable and unobservable at time $t < T$. Note that as usual, the martingale property of the zero-coupon bond is given by

$$P(t, T, x) = \mathbb{E}^{Q}\left[e^{-\int_t^T r_s ds}p(T, T, x)|\mathcal{F}_t\right].$$

Assuming the risk-neutral dynamics (under $\mathbb{Q}$) for r_t, by definition of $P(t,T,x)$ and the Feynman-Kac representation, we must have that

$$LP(t,T,x) = 0, \qquad P(T,T,x) = 1.$$

with the generator L defined as in (3.3) with $\gamma(t,x) = e^x$. Define the Fourier transform of $P(t,T,x)$ to be

$$\hat{P}(t,T,\xi) := \mathcal{F}(P(t,T,x)) := \int_{\mathbb{R}} e^{i\xi x} P(t,T,x)dx.$$

Now, we will perform as usual the Taylor-based expansion of the operator L around a fixed point $\bar{x}$. Following [53], we can find that

$$\begin{aligned} &L_0 P^0(t,T,x) = 0, \qquad P^0(T,T,x) = 1 \\ &L_0 P^k(t,T,x) = -\sum_{h=1}^{k}(L_h - L_{h-1})P^{k-h}(t,T,x) \qquad P^0(T,T,x) = 0. \end{aligned}$$

Solving these in the Fourier space (note: no need for the adjoint formulation since the operators act on (t,x) and we take the Fourier transform also with respect to (t,x)) we find

$$\begin{aligned} &\hat{P}^0(t,T,\xi) = \delta_\xi e^{(T-t)\psi(\xi)}, \\ &\hat{P}^k(t,T,\xi) = -\int_t^T e^{(T-s)\psi(\xi)} \mathcal{F}\left(\sum_{h=1}^{k}\left(L_h(s) - L_{h-1}(s)\right)P^{k-h}(t,s,\cdot)\right)(\xi)ds, \end{aligned}$$

where

$$\psi(\xi) = i\xi\mu_0 - a_0\xi^2 - \gamma_0 + \int_{\mathbb{R}} \nu_0(dq)(e^{iz\xi} - 1 - iz\xi)$$

After some algebraic manipulations it can be shown, see also Chapter 2, that the n-th order approximation of the Fourier transform of the ZCB is of the form

$$\hat{P}^{(n)}(t,T,\xi) := \delta_\xi \sum_{k=0}^{n}(x-\bar{x})^k g_{n,k}(t,T,\xi). \tag{3.18}$$

Then, using the definition of the Dirac-delta function, we have

$$\int_{\mathbb{R}} \partial_\xi^n f(\xi)\delta_\xi d\xi = (-1)^n \partial_\xi^n f(\xi)|_{\xi=0}.$$

The inverse Fourier transform is thus given by

$$\begin{aligned} P^{(n)}(t,T,x) &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{i\xi x}\hat{P}^{(n)}(t,T,\xi)d\xi \\ &= \frac{1}{\sqrt{2\pi}}\sum_{k=0}^{n}(x-\bar{x})^k g_{n,k}(t,T;0), \end{aligned} \tag{3.19}$$

where the functions $g_{n,k}(t,T,0)$ depend on $\bar{x}$, which we take to be the value of x at time t. In this way, all maturity times T of the bond $P(t,T,x)$ depend on the same value of the short rate, namely $\bar{x} := x_t$. Note that the reason they do not depend on x_T is solely due to the fact that the interest rate is only represented through the integral. This is similar to what happens in the Hull-White model and is a crucial part in an efficient evaluation of the swaption price. Since r_t is a Markov process and the price of the zero-coupon bond between times t and T becomes non-stochastic at time t, we can say that it is a function or r_t, i.e. $P(t_m, t_k, r)$ is the price of a ZCB conditional on $r_{t_m} = r$ at time t_m with maturity t_k.

Defaultable ZCB The defaultable zero coupon bond is given by

$$\begin{aligned} P^D(t,T,x) &= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\zeta>T\}} e^{-\int_t^T r_s ds}\right] \\ &= \mathbb{1}_{\{\zeta>t\}} \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T (r_s+\alpha(s,X_s))ds}\right], \end{aligned}$$

where $\alpha(s, X_s)$ is the default intensity and we have used the standard identity as in (1.4). The same pricing can be performed on the defaultable ZCB, only now in the operator L we would have $\gamma(t,x) = \alpha(t,x) + e^x$.

Valuing swaptions A swaption gives the owner the right at the exercise date T_m to enter into the underlying swap with payment dates $T_{m+1}, ..., T_M$ (the tenor structure) and a nominal value N (the notional), which we for simplicity set equal to one. Let $\Delta_k := T_k - T_{k-1}$ and let the Libor rate at time t for a loan between times T_{k-1} and T_k be $F_k(t)$. The floating leg of the swap at time T_k has value $\Delta_k F_k(T_{k-1})$ and the fixed leg is $\Delta_k K$. Note that the rate to be applied for the floating leg at time T_k is fixed at reset time T_{k-1}. The time t discounted payoff of a payer swap can be expressed as

$$\sum_{k=m+1}^{M} P(t,T_k)\Delta_k(F_k(T_{k-1}) - K),$$

suppressing the dependence on the process x in $P(t,T,x)$. The arbitrage pricing law is given by

$$v(t,x) = N(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{V(T)}{N(T)}|\mathcal{F}_t\right],$$

where $N(t)$ is the numeraire, in our case the money-market accoung $B(t)$. The price of a European call swaption at time t with maturity T_m and strike K can be written as the risk-neutral expectation of the discounted future payoffs conditional on $X_t = x$:

$$v(t,x) = \mathbb{E}^{\mathbb{Q}}\left[\frac{B(t)}{B(T_m)}\left(\sum_{k=m+1}^{M}(\Delta_k F_k(T_m)P(T_m,T_k) - \Delta_k K P(T_m,T_k))\right)^+\right].$$

Note that

$$U(T_m, x) := \sum_{k=m+1}^{M} (\Delta_k F_k(T_m) P(T_m, T_k) - \Delta_k K P(T_m, T_k)),$$

is the payoff for entering the underlying swap at time T_m associated with payment times $T_{m+1}, ..., T_M$. Note that here we do not make a distinction between multiple curves, but use $P(t, T)$. We have,

$$F_m(t) = \frac{P(t, t_{m-1}) - P(t, t_m)}{\Delta_m P(t, t_m)}, \tag{3.20}$$

therefore the whole formula is a function of different zero-coupon bonds.

Remark 3.5.1. We can rewrite this using the swap rate, which makes the payoff equal to zero at time t and is given by

$$S(t) = \frac{P(t, T_m) - P(t, T_M)}{\sum_{i=m+1}^{M} \Delta_i P(t, T_i)},$$

in which case the swaption becomes a call option on the swap rate

$$v(t, x) = \mathbb{E}^{\mathbb{Q}} \left[P(t, T_m)(S(T_m) - K)^+ \sum_{i=m+1}^{M} \Delta_i P(T_m, T_i) \right].$$

Now, using (3.19) the equation under the expected value can be rewritten as a function of r_{T_m}, i.e. the interest rate at maturity time T_m. Therefore, we can use the method developed in Chapter 2, in which we combine the approximated characteristic function of the process r_t and the COS method in order to value functions of this form.

CVA computation Using the default-free and defaultable zero coupon bonds we can calculate two swap values $v(t, x)$ (risk-free) and $\hat{v}(t, x)$ (risky) respectively. In particular, suppose the counterparty is paying the floating leg, and the risky value corresponds to the value of the swap in which the counterparty might default. In a swap at time T_k we receive $F_k(T_{k-1})$, i.e. the Libor rate fixed at time T_{k-1}. The value of the discounted Libor payment is

$$\mathbb{E} \left[\frac{F_k(T_{k-1})}{B(T_k)} \right] \approx \mathbb{E} \left[\frac{F_k(T_{k-1}) P(T_{k-1}, T_k)}{B(T_{k-1})} \right],$$

where we have used $B(T_k) \approx B(T_{k-1})/P(T_{k-1}, T_k)$. Then assuming that the counterparty paying the Libor might default and using (3.20) we have

$$\begin{aligned} \mathbb{E} \left[\frac{F_k(T_{k-1}) P(T_{k-1}, T_k)}{B(T_{k-1})} \mathbb{1}_{\{\zeta > T_k\}} \right] =& \mathbb{E} \left[\frac{P(T_{k-1}, T_{k-1}) - P(T_{k-1}, T_k)}{B(T_{k-1})} \mathbb{1}_{\{\zeta > T_k\}} \right] \\ =& \mathbb{E} \left[\frac{\mathbb{1}_{\{\zeta > T_k\}}}{B(T_{k-1})} \right] - \mathbb{E} \left[\frac{\mathbb{1}_{\{\zeta > T_k\}}}{B(T_k)} \right] \end{aligned}$$

$$
\begin{aligned}
&\approx P^D(t, T_{k-1}) - P^D(t, T_k) \\
&\approx P^D(t, T_{k-1}) F^D(T_{k-1}, T_k) P(T_{k-1}, T_k),
\end{aligned} \tag{3.21}
$$

where we have defined,

$$
F_k^D(T_{k-1}) := \frac{P^D(T_{k-1}, T_{k-1}) - P^D(T_{k-1}, T_k)}{P(T_{k-1}, T_k)},
$$

which can be seen as a defaultable Libor rate. A more simple approximation might be,

$$
\mathbb{E}\left[F_k(T_{k-1}) P(T_{k-1}, T_k) \mathbb{1}_{\{\zeta > T_k\}}\right] \approx \mathbb{E}\left[F_k(T_{k-1}) P^D(T_{k-1}, T_k)\right], \tag{3.22}
$$

however clearly this approximation rests on the assumption of independence between $\mathbb{1}_{\{\zeta > T_k\}}$ and $P(T_{k-1}, T_k)$, in this way not incorporating the wrong-way risk.

Both the approximation in (3.21) and in (3.22) can be used to compute the CVA, which is computed as the difference between the swap value with the risky payoff (in which the floating leg payer can default) and the risk-free payoff (where we assume this can not happen). In particular, using e.g. the approximation in (3.22), the payoffs of the risky and risk-free swap, $\hat{U}$ and U, can be written respectively as

$$
\begin{aligned}
\hat{U}(t, x) &= \sum_{k=m+1}^{M} (\Delta_k F_k(T_m) P^D(T_m, T_k) - \Delta_k K P(T_m, T_k)), \\
U(t, x) &= \sum_{k=m+1}^{M} (\Delta_k F_k(T_m) P(T_m, T_k) - \Delta_k K P(T_m, T_k)),
\end{aligned}
$$

where P^D is the defaultable ZCB of the counterparty and P is the risk-free bond.

Chapter 4

Systemic risk in an interbank network with self-exciting shocks

Counterparty credit risk arises from the possibility that the counterparty in a particular transaction, e.g. in the form of a derivative, might default on the amount owed to the other party. In the last two chapters we discussed how banks can adjust the price of the derivative to account for this risk. The aim of this chapter is to investigate the systemic risk in an interconnected (e.g. through derivatives, loans) network. As discussed in Section 1.5, it is the risk that a default at an individual entity might cause liquidity problems not just for its counterparty but may result in the default propagating through the system as a whole. In particular we will incorporate both self- and cross-exciting shocks as well as interbank lending in the monetary reserve process of the bank. The excitement comes from the effect that *past* movements in the asset value of the bank and that of its counterparties have on the *current* variations in the banks' asset value. These effects are modelled using a Hawkes process [38]. These kinds of self-exciting processes have previously been used in portfolio credit risk computation in a top-down approach, see e.g. [1], [28] and [23]. In this work we model the monetary reserve process of a bank through a mean-field interaction diffusion with an additional Hawkes distributed jump term. We study the behavior of the system as the number of nodes approaches infinity by deriving the weak limit of the empirical measure of this interacting system.

In particular our convergence result is based on the analysis of [26], where the authors show that the intensity of a Hawkes process in the limit of a fully connected network tends to behave as that of a non-homogeneous Poisson process. We show that the underlying limit process for the monetary reserves of the nodes has purely diffusive dynamics and the effect of the Hawkes process is reflected in a *time-dependent* drift coefficient. Then we define several risk indicators and use the weak convergence analysis to derive the law of large numbers approximations to explicitly show the effects of the Hawkes process on the risk in a large interbank network. In the numerical section we then compare the LLN approximations with the actual values simulated through a Monte-Carlo method and conclude that in a model of interbank networks, the default risk is indeed higher when we incorporate the self- and cross-exciting shocks.

The rest of the chapter is structured as follows: in Section 4.1 we define the Hawkes process and give a motivation for incorporating it in the interbank network. In Section 4.2 we introduce the mean-field model for the log-monetary reserve process and study through simulations the effects of incorporating the self-exciting jump intensity and in particular compare it to the independent Poisson intensity. In Section 4.3 we derive the weak convergence of the empirical mean of monetary reserves, explicitly characterize the weak limit measure-valued process and provide several results for extensions of the model. Finally, in Section 4.4 we derive several measures of systemic risk in the network and numerically validate the accuracy of the derived limiting process.

4.1 The framework

4.1.1 Motivation

A known source of systemic risk in financial networks is the propagation of defaults due to interbank exposures such as loans. Due to these loans the failure of a borrowing node to repay its loans, may subsequently cause a loss in liquidity of the lenders as well. This then propagates the default through the network. Besides interbank exposures, another common cause of default propagation are fire sales. If one institution decides to liquidate a large part of its assets, depressing the price, this causes a loss at the institutions holding the same assets, creating a *cross-exciting* spiral across the institutions. Therefore, institutions that do not have mutual counterparty exposure can still suffer financial distress if they have holdings of common assets on their balance sheets. As illustrated by [35], the effects from these so-called fire sales can be even greater than the contagion effects due to counterparty exposures.

A *self-exciting* effect present in financial networks is known as financial acceleration and refers to the fact that current variations in the asset side of the balance sheet depend on past variations in the assets themselves. In other words, a shock affecting the banks portfolio can cause creditors to claim their funds back or tighten the credit conditions, in this way causing an additional shock for the bank.

As has been mentioned in [20], while interbank lending itself may not be a significant cause of default propagation, it is important to account for both the correlated effects of default contagion through lending agreements as well as exposure to common market events. Here, we choose to model the correlated effects of the fire sales, financial acceleration and the interbank lending structure on both the default propagation as well as on the overall loss in the network through a Hawkes counting process. The shocks affecting the portfolio of the institution arrive conditional on the infinite history of previous shocks to both the institutions own assets as well as those of the other nodes in the network provided that they share common assets.

4.1.2 Hawkes processes

Specific types of events that are observed in time do not always arrive in evenly spaced intervals, but can show signs of clustering, e.g. the arrival of trades in an order book, or the contagious default of financial institutions. Therefore, assuming that these events happen independently is not a valid assumption. A Hawkes process (HP), also known as a self-exciting process, has an intensity function whose current value, unlike in the Poisson process, is influenced by past events. In particular, if an arrival causes the conditional intensity to increase, the process is said to be self-exciting, causing a temporal clustering of arrivals. Hawkes processes can be used for modelling credit default events in a portfolio of securities, as has been done in e.g. [28] or for modelling asset prices using a mutually exciting jump component to model the contagion of financial shocks over different markets ([1]). An overview of other applications of Hawkes processes in finance, in particular in modelling the market microstructure, can be found in e.g. [3].

Let $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ be a complete filtered probability space where the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ satisfies the usual condition. Hawkes processes ([38]) are a class of multi-variate counting processes $(N_t^1, ..., N_t^M)_{t\geq 0}$ characterized by a stochastic intensity vector $(\lambda_t^1, ..., \lambda_t^M)_{t\geq 0}$ which describes the $\mathcal{F}_t$-conditional mean jump rate per unit of time, where $\mathcal{F}_t$ is the filtration generated by $(N^i)_{1\leq i\leq M}$ up to time t. Consider the set of nodes $I_M := \{1, \ldots, M\}$. Define the kernel $g(t) = (g^{i,j}(t), (i,j) \in I_M \times I_M)$ with $g^{i,j}(t) : \mathbb{R}_+ \to \mathbb{R}$ and the constant intensity $\mu = (\mu^i, i \in I_M)$ with $\mu^i \in \mathbb{R}_+$.

Definition 4.1.1 (Hawkes process)**.** A linear Hawkes process with parameters (g, μ) is a family of $\mathcal{F}_t$-adapted counting processes $(N_t^i)_{i\in I_M, t\geq 0}$ such that:

1. almost surely for all $i \neq j$, $(N_t^i)_{t\geq 0}$ and $(N_t^j)_{t\geq 0}$ never jump simultaneously,
2. for every $i \in I_M$, the compensator Λ_t^i of N_t^i has the form $\Lambda_t^i = \int_0^t \lambda_s^i ds$, where the intensity process $(\lambda_t^i)_{t\geq 0}$ is given by

$$\lambda_t^i = \mu^i + \sum_{j=1}^{M} \int_{[0,t[} g^{i,j}(t-s) dN_s^j. \tag{4.1}$$

In other words, $g^{i,j}$ denotes the influence of an event of type j on the arrival of i: each previous event dN_s^j raises the jump intensity $(\lambda_t^i)_{i\in I_M}$ of its neighbors through the function $g^{i,j}$. The compensated jump process $N_t - \int_0^t \lambda_s ds$ is a $\mathcal{F}_t$-local martingale. For g a positive and a decreasing function of time t, the influence of a jump decreases and tends to 0 as time evolves.

Following Proposition 3 in [26], one can rewrite the Hawkes process in the sense of Definition 4.1.1 as a Poisson-driven SDE with the i.i.d. family of $\mathcal{F}_t$-Poisson measures $(\pi^i(ds, dz), i \in I_M)$

with intensity measure (ds, dz):

$$N_t^i = \int_0^t \int_0^\infty \mathbb{1}_{\{z \le \mu_t + \sum_{j=1}^M \int_{[0,s[} g^{i,j}(t-s) dN_s^j\}} \pi^i(ds, dz). \tag{4.2}$$

Next we state a well-posedness result, based on Theorem 6 in [26]:

Lemma 4.1.2 (Existence and uniqueness)**.** *Let $g^{i,j}$ be locally integrable for all $(i,j) \in I_M \times I_M$; there exists a pathwise unique Hawkes process $(N_t^i)_{i \in I_M, t \ge 0}$, such that $\sum_{i=1}^M \mathbb{E}[N_t^i] < \infty$ for all $t \ge 0$.*

By introducing the pair $\{t_k, n_k\}_{k=1}^{K_t}$, where t_k denotes the time of event k, $n_k \in I_M$ is the event type and $K_t = \sum_{i=1}^M N_t^i$ is total number of event arrivals up to time t, we can rewrite the intensity as

$$\lambda_t^i = \mu^i + \sum_{k=1}^{K_t} g^{i,n_k}(t - t_k), \qquad i \in I_M.$$

A common choice for $g^{i,j}(t)$ is an exponential decay function defined as

$$g^{i,j}(t) = \alpha^{i,j} e^{-\beta^i t},$$

so that λ_t^i jumps by $\alpha^{i,j}$ when a shock in j occurs, and then decays back towards the mean level μ^i at speed β^i. Note that this function satisfies the local integrability property, i.e. $g^{i,j} \in L^1_{\text{loc}}(\mathbb{R}_+)$. If $g^{i,j}$ is exponential then the couple (N_t, λ_t) is a Markov process [3], and the intensity can be rewritten in a Markovian form as

$$d\lambda_t^i = \beta^i(\mu^i - \lambda_t^i)dt + \sum_{j=1}^M \alpha^{i,j} dN_t^j.$$

The simulation of a Hawkes process can be done using what is known as Ogata's modified thinning algorithm, see for more details [57] and [24].

If the Hawkes process $(N_t^i)_{i \in I_M, t \ge 0}$ satisfies certain conditions, we have the following stationarity result (see [10] and [4] for details), which will come in useful in the further sections.

Proposition 4.1.3. *Suppose that the matrix Φ with entries $\int_0^\infty |g^{i,j}(t)| dt$ has a spectral radius strictly less than one. Then there exists a unique multi-variate Hawkes process $(N_t^i)_{t \ge 0}$ for $i \in I_M$ with stationary increments and the associated intensity as in* (4.1) *is a stationary process. Moreover we have $\mathbb{E}[|\lambda_t|^2] < \infty$.*

Furthermore, we remark here that a multi-dimensional Hawkes process with stationary increments is uniquely defined by its first- and second-order statistics ([4]).

4.2 The mean-field model

In this section we define the mean-field model for the log-monetary reserves of each of the nodes in the model. The interaction between the nodes is defined through the drift term and additionally we consider the reserve process to be subjected to a self- and cross-exciting Hawkes distributed shock.

4.2.1 Definition

Define $\mathcal{F}_t = \sigma((W_s^i, N_s^i), 0 \leq s \leq t, i \in \mathbb{N})$. Assume that, for $i \in I_M$ the log-monetary reserves of the i-th bank satisfy the following stochastic differential equation (SDE)

$$dX_t^i = \frac{a^i}{M} \sum_{k=1}^{M} (X_t^k - X_t^i) dt + \sigma^i dW_t^i + c^i dN_t^i,$$

with $X_0^i \in \mathbb{R}_+$ the initial reserves for each bank and where $a^i \geq 0$, $\sigma^i \geq 0$ and $c^i := \hat{c}^i / M < 0$ are constants for each $i \in I_M$. The process $W(t) = \{W_t^i\}_{i=1}^M$ is an M-dimensional uncorrelated Brownian motion, and $N_t = \{N_t^i\}_{i=1}^M$ is the vector of Hawkes processes with self-exciting intensity λ_t^i as defined in Section 4.1.2. With the drift term defined in this way, if bank k has more (less) log-monetary reserves than bank i, i.e. $X_t^k > X_t^i$ ($X_t^k < X_t^i$), bank k is assumed to lend (borrow) a proportion of the surplus to (deficit from) bank i, with proportionality factor a^i/M. A jump in the Hawkes process i affects the corresponding X_t^i through the proportionality factor c^i and increases the intensity λ_t^j for $j \in I_M$ if $g^{i,j}(t) \neq 0$. In this way the jump activity varies over time resulting in a clustering of the arrival of the jumps and the shocks propagate through the network in a contagious manner through the contagion function $g^{i,j}(t)$. We thus interpret the jump term $c^i dN_t^i$ as a self- and cross-exciting *negative* effect that occurs due to financial acceleration and fire sales, resulting in a decrease in a banks monetary reserve. In [8] the authors considered a similar mean-field model for the monetary reserves but assumed the jumps to occur at independent Poisson distributed random times. However, not accounting for the clustering effect of the jumps might cause a significant underestimation of the systemic risk present in the network. We define a default level $D \leq 0$ and say that bank i is in a default state at time T if its log-monetary reserve reached the level D at time T. We remark that in our model even if bank i has defaulted, i.e. its monetary reserve reaches a negative level, it continues to participate in the interbank activities borrowing from the counterparties until it e.g. reaches a positive reserve level again. In other words, the level of monetary reserves takes values in $\mathbb{R}$. We will work in the following setting:

Assumption 4.2.1 (Parameters)**.** We collect the parameters associated with the dynamics of the i-th monetary reserve process $i \in I_M$ as

$$p^i := (a^i, \sigma^i, c^i) \in (\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_-).$$

We denote by $\mathbb{1}_x$ the Dirac-delta measure centered at x and we set

$$q^M = \frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{p_i}, \qquad \varphi_0^M = \frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{X_0^i}.$$

We assume $\lim_{M\to\infty} q^M = \mathbb{1}_{p^*}$, i.e. $p^i \to p^* := (a,\sigma,c)$ as $i\to\infty$ and $\lim_{M\to\infty}\varphi_0^M = \mathbb{1}_x$, i.e. $X_0^i \to x$ as $i\to\infty$. We take the exponential decay function for the contagion

$$g^{i,j}(t-s) = \frac{1}{M}g(t-s) := \frac{1}{M}\alpha e^{-\beta(t-s)},$$

which is a locally square-integrable function with $\alpha,\beta\in\mathbb{R}_+$. Finally, the parameters are assumed to all be bounded by a constant C_p.

We remark here that the results developed in this chapter hold also for more general distributions, i.e. $\lim_{M\to\infty} q^M = q$ and $\lim_{M\to\infty}\varphi_0^M = \varphi_0$, but for simplicity of the results we assume the parameter vector converges to a constant vector.

Defining the reserve average as

$$\bar{X}_t = \frac{1}{M}\sum_{i=1}^{M} X_t^i,$$

we can rewrite the SDE as a mean-field interaction SDE

$$dX_t^i = a^i(\bar{X}_t - X_t^i)dt + \sigma^i dW_t^i + c^i dN_t^i. \tag{4.3}$$

From (4.3) we see that the processes (X_t^i) are mean-reverting to their ensemble average $(\bar{X}_t)$ at rate a^i.

Lemma 4.2.2. *There exists a unique solution $(X_t^1,...,X_t^M)$ to the system of SDEs given by* (4.3) *for $i\in I_M$.*

Proof. The proof is similar to Theorem 9.1 in [43]. Define Y_t^i to be the solution of the SDE (4.3) without jumps. By Example 2 in [22], we know that the SDE has a unique strong solution $(Y_t^1,...,Y_t^M)$. By definition of a Hawkes process we have that $N^1,...,N^M$ never jump simultaneously: this implies the existence of an increasing sequence of jump times $(\tau_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty}\tau_n = +\infty$. Then we can define

$$X_t^{(i,1)} := \begin{cases} Y_t^i, & 0\le t<\tau_1, \\ Y_{\tau_1-}^i + \mathbb{1}_{k=i}c^i, & t=\tau_1, \text{ if there is a jump in } N^k. \end{cases} \tag{4.4}$$

From Lemma 4.1.2, we know that there exists a unique Hawkes process $(N_t^i)_{t\ge 0}$ for $i\in I_M$, thus we can say that $X_t^{(i,1)}$ is the unique solution to (4.3) for $t\in[0,\tau_1]$. Then we define $\bar{X}_t^{(i,2)}$ on $t\in$

$[0, \tau_2 - \tau_1]$ similar to (4.4) using as initial state $\bar{X}_0^i := X_{\tau_1}^{(i,2)}$ and driving factors $\bar{W}_t^i := W_{t+\tau_1}^i - W_{\tau_1}^i$ and $\bar{N}_t^i := N_{t+\tau_1}^i - N_{\tau_1}^i$. Then we set

$$X_t^i := \begin{cases} X_t^{i,1}, & 0 \le t < \tau_1, \\ \bar{X}_{t-\tau_1}^{(i,2)} & \tau_1 \le t \le \tau_2. \end{cases}$$

So that X_t^i, $t \in [0, \tau_2]$ is the unique solution to (4.3). Iterating the above process, we have that X_t^i is determined uniquely on the time interval $[0, \tau_n]$ for each $n \in \mathbb{N}$. □

4.2.2 Simulation

Consider, for the sake of illustration, the following SDE

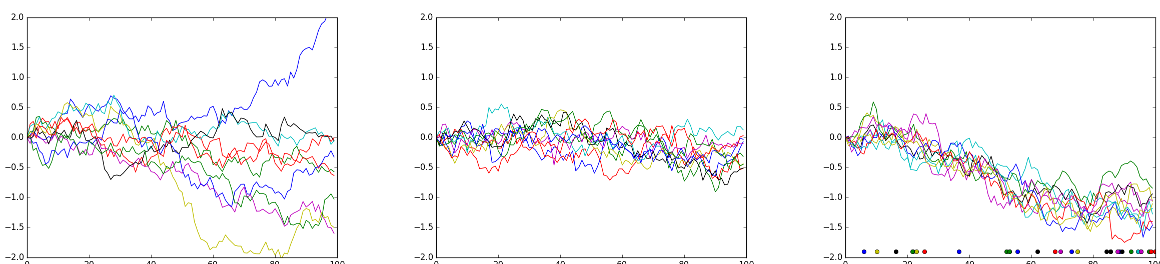
$$dX_t^i = a(\bar{X}_t - X_t^i)dt + \sigma d\tilde{W}_t^i + cdN_t^i,$$

with $\tilde{W}_t^i := \rho W_t^0 + \sqrt{1-\rho^2} W_t^i$, where W_t^i, $i = 0, ..., M$ are independent Brownian motions and W_t^0 represents common noise (similar to the setting in [16]). We keep the parameters of the constant intensity and the excitation function $g^{i,j} = \alpha^{i,j} e^{-\beta^i t}$ fixed at $\mu^i = 10/M$, $\beta^i = 2/M$ and $\alpha^{i,j} = 2/M$ and the initial reserve value is set at $X_0 = 0$.

Table 4.1: Parameters corresponding to the various scenarios of the realizations of $(X_t^i, i = 1, ..., 10)$.

Scenario	a	σ	c	ρ
No lending, independent BMs	0	1	0	0.2
Lending, independent BMs	10	1	0	0
No lending, correlated BMs	0	1	0	0.2
Lending and correlated BMs	10	1	0	0.2
Lending, correlated BMs and Poisson jumps	10	1	0.2	0.2
Lending, correlated BMs and Hawkes jumps	10	1	0.2	0.2

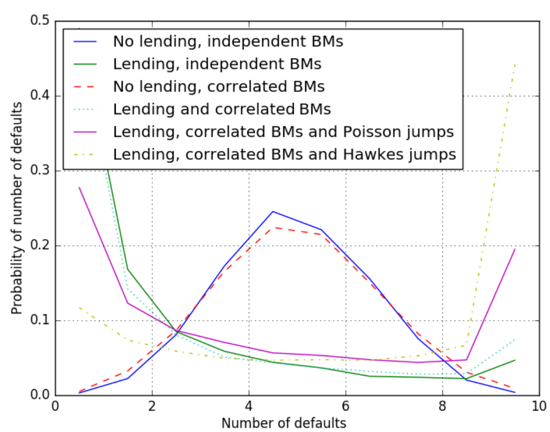
Figure 4.1: One realization of $(X_t^i, i = 1, ..., 10)$, $t = 1, ...100$ with no lending and independent Brownian motions (left), lending and correlated Brownian motions (center) and lending, correlated Brownian motions and the Hawkes distributed jump (with the jump times shown as dots) (right).



We consider several scenarios of the monetary reserve process denoted in Table 4.1. In Figure 4.1 we see that the trajectories generated by the correlated Brownian motions with lending are more grouped than the ones generated by independent Brownian motions without lending. The Hawkes shock, as expected, causes more trajectories to reach the default level, due to it being an additional source of default propagation.

Consider the default level $D = -0.7$. In Figure 4.2 we show the distributions of the number of defaults defined as $\mathbb{P}\left(\sum_{i=1}^{M}\left(\min_{0\leq t\leq T} X_t^i \leq D\right) = n\right)$, for the independent Brownian motion case, the dependent case and the cases including a Poisson process and a Hawkes process. We observe that the mean-field interbank lending causes most of the probability mass to be set around zero defaults, as opposed to the no lending case when the density function is centered at 5 defaults. However, the lending component also adds a non-negligible probability of all nodes defaulting at once. The correlation between the Brownian motions affects the loss distribution only slightly. As expected, adding the self-exciting and clustering Hawkes process increases the tail-risk even more so that the probability of all nodes reaching a default state rises significantly.

Figure 4.2: The distribution of the number of defaults in several different scenarios, as explained in Table 4.1. The parameters in the Monte Carlo simulated based on a discretized Euler-Maruyama scheme are $M = 10$, $T = 1$, 10000 simulations and 100 time steps.



4.2.3 Dependency

As we have already seen in Figure 4.2, the Hawkes process increases the probability of multiple defaults occurring at the same time as compared to an independent Poisson process. It is therefore of interest to study the dependence structure between the nodes in more detail. As is standard in multi-variate statistics, see [62], a tool for assessing the (not necessarily linear) dependency between variables is the measure $p(q)$ given by

$$p(q) = P\left(X^i > F_{X^i}^{-1}(q) | X^j > F_{X^j}^{-1}(q)\right), \qquad i,j \in I_M,$$

the probability of one of the variables X^i being above the qth quantile of its marginal distribution F_{X^i} conditional on the other variable X^j being above its qth quantile. To remove the influence of marginal aspects it is typical to transform the data to a common marginal distribution, with e.g. a transformation to unit Fréchet marginals (for details we refer to the methodology in [62]). In the presence of a dependence between two nodes in our model, the probability of default of one firm conditional on the default of the other will be significant. When computing the systemic risk present in interconnected financial networks, quantifying this dependence is clearly of key importance. Note that in our model we have two key dependencies present:

- Dependence through the drift term: a high X_t^1 results in a change in X_s^1 and X_s^2 for $s > t$ due to the interbank loans.

- Dependence through the Hawkes process: if $\Delta X_t^1 << 0$ represents the occurence of a jump at time t, then the likelihood of $\Delta X_s^1 << 0$ and $\Delta X_s^2 << 0$ for $s > t$ increases. We remark that the likelihood of seeing the shock decreases with a larger s due to the mean-reverting excitation function $g^{i,j}$ $i, j \in \{1, 2\}$.

Figure 4.3 shows the scatter plots for both an independent Poisson jump and a Hawkes jump. Already here we see that the Hawkes jump seems to reflect a more strong dependency in the tails. In Figure 4.4 we plot the measure $p(q)$ (for the left tail) compared to the $1-q$ function representing independence, for several different parameter sets. We see that the Hawkes process shows significantly more dependence between the two nodes for all quantiles compared to the Poisson process. In particular, we note that having only a jump term in the monetary reserve process results in a significant tail probability, where the tail probability of the Hawkes process is considerably higher than that of the Poisson process. This is to be expected since the self-exciting nature of the jumps causes the extreme events in one node to influence extreme events in the other node. Furthermore, incorporating the independent Brownian motion seems to reduce the tail risk almost to zero, while adding the interbank loans in turn causes a slight increase in the tail risk, due to the additional source of default propagation.

Figure 4.3: Scatter plots of X_t^1 and X_t^2 ($M = 2$) showcasing the dependence structure between the nodes in the presence of a Poisson jump (left) and a Hawkes jump (right).

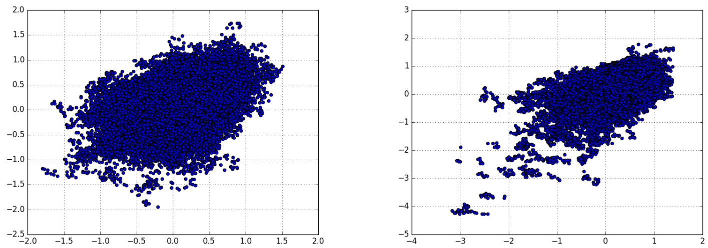
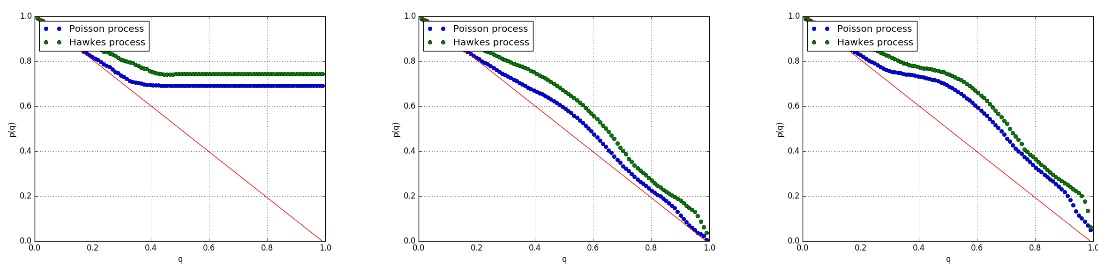


Figure 4.4: The measure $p(q)$ quantifying the dependence of X_t^1 and X_t^2 ($M = 2$) with Poisson and Hawkes jumps for the case of no Brownian motion, no interbank lending but only jumps (left, $\sigma = 0$, $a = 0$ and $c = -1$), Brownian motion, no lending and jumps (center, $\sigma = 0.1$, $a = 0$ and $c = -1$) and Brownian motion, lending and jumps (right, $\sigma = 0.1$, $a = 0.5$ and $c = -1$). The other parameters in the Monte Carlo simulation based on a Euler-Maruyama scheme are $T = 1$, 500 simulations, 100 time steps, $X_0^i = 0$, $\rho = 0$, with $\mu^i = 0.1$, $\beta^i = 1.2$, $\alpha^{i,j} = 1.2$.



4.3 Mean-field limit

We derive theoretical mean-field limits for the monetary reserve process with a Hawkes jump term to show the effects of considering this additional type of contagion on the total losses in the network in the case of the number of nodes tending to infinity. Our derivations are based on [16] and [8]. In other words, we wish to understand the behavior of the distribution of the process $X_t = (X_t^i)$, $i \in I_M$ as in (4.3) when $M \to \infty$. Let the vector (p^i, X_t^i) take on values in the space $\mathcal{O} := (\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_-) \times \mathbb{R}$. Define the sequence of empirical measures as

$$\nu_t^M := \frac{1}{M} \sum_{i=1}^{M} \delta_{(p^i, X_t^i)}, \qquad t \geq 0, \tag{4.5}$$

on the Borel space $\mathcal{B}(\mathcal{O})$. In other words we keep track of the empirical distribution of the type and monetary reserve for all nodes. Let $S := \mathcal{P}(\mathcal{O})$ be the collection of Borel probability measures on $\mathcal{O}$.

Then $(\nu_t^M)_{t\geq 0}$ is an element of the Skorokhod space $D_S[0,\infty)$, i.e. it can be viewed as an S-valued right-continuous, left-hand limited stochastic process. For any smooth function $f(p,x) \in C^\infty(\mathcal{O})$ defined for $(p,x) \in \mathcal{O}$ define the integral w.r.t. the measure ν by

$$\nu(f) := \int_{\mathcal{O}} f(p,x)\nu(dp \times dx), \tag{4.6}$$

so that

$$\nu_t^M(f) = \frac{1}{M}\sum_{i=1}^M f(p^i, X_t^i), \qquad t \geq 0. \tag{4.7}$$

Then we have $\bar{X}_t = \nu_t^M(I)$ where $I(x) = x$.

We wish to understand the dynamics for ν_t^M for large M. In deriving the limit of the process ν_t^M for $M \to \infty$ we use an argument similar to [8] and [33]. In particular, the focus here is on identifying the limiting dynamics, using the result of [26] on the behavior of Hawkes processes in a large system. We identify the limit through the generator of the limiting martingale problem in Section 4.3.1, and subsequently in Section 4.3.2 we identify the limit process.

4.3.1 Weak convergence

We want to use the martingale problem to show that ν_t^M converges to a limiting process. For notational convenience we will write $f(X_t^i) := f(p^i, X_t^i)$. By the definition of a Hawkes process we have that for all $i \neq j$, $(N_t^i)_{t\geq 0}$ and $(N_t^j)_{t\geq 0}$ never jump simultaneously and a jump in one of the processes dN_t^i results in only X_t^i having a jump of size c^i. Therefore, applying Itô's formula to the semimartingale X_t^i gives,

$$\begin{aligned} df(X_t^i) =& a^i \partial_x f(X_t^i)[\nu_t^M(I) - X_t^i]dt + \frac{1}{2}(\sigma^i)^2 \partial_{xx} f(X_t^i)dt + \sigma^i \partial_x f(X_t^i) dW_t^i \\ &+ (f(X_{t-}^i + c^i) - f(X_{t-}^i))dN_t^i, \end{aligned}$$

Then we have, using the definition of ν_t^M in (4.7),

$$\begin{aligned} \nu_t^M(f) =& \nu_0^M(f) + \int_0^t \nu_s^M(\mathcal{L}^1 f)\nu_s^M(I)ds - \int_0^t \nu_s^M(\mathcal{L}^2 f)ds + \frac{1}{M}\sum_{i=1}^M \int_0^t \sigma^i \partial_x f(X_s^i)dW_s^i \\ &+ \int_0^t \nu_s^M(\mathcal{L}^3 f)ds + \frac{1}{M}\sum_{i=1}^M \int_0^t \left(f(X_{t-}^i + c^i) - f(X_{t-}^i)\right)dN_s^i, \end{aligned} \tag{4.8}$$

where we have defined the operators $\mathcal{L}^*$ acting on $f(p^i, X_t^i)$ as

$$\mathcal{L}^1 f(p,x) := a\partial_x f(p,x), \qquad \mathcal{L}^2 f(p,x) := ax\partial_x f(p,x), \qquad \mathcal{L}^3 f(p,x) = \frac{1}{2}\sigma^2 \partial_{xx} f(p,x),$$

so that

$$\nu_t^M(\mathcal{L}^1 f) = \frac{1}{M}\sum_{i=1}^M a^i \partial_x f(p^i, X_t^i), \;\; \nu_t^M(\mathcal{L}^2 f) = \frac{1}{M}\sum_{i=1}^M a^i X_t^i \partial_x f(p^i, X_t^i),$$
$$\nu_t^M(\mathcal{L}^3 f) = \frac{1}{M}\sum_{i=1}^M \frac{1}{2}(\sigma^i)^2 \partial_{xx} f(p^i, X_t^i).$$

Define for any smooth function $\varphi \in C_K^\infty(\mathbb{R}^N)$ with $N \in \mathbb{N}$ and Borel measure $\nu \in S$

$$\Phi(\nu) = \varphi(\nu(\mathbf{f})), \tag{4.9}$$

with $\mathbf{f} = (f_1, ..., f_N)$ for $f_n \in C_K^\infty(\mathcal{O})$, $n = 1, ..., N$ and $\nu(\mathbf{f}) := (\nu(f_1), ..., \nu(f_N)) \in \mathbb{R}^N$. Let $\mathcal{S}$ be the collection of bounded measurable functions Φ on S. Then $\mathcal{S}$ separates S and it thus suffices to show convergence of the martingale problem for those functions. Then, by applying Itô's formula to $\varphi(\nu_t^M(\mathbf{f}))$ and using the fact that $d\tilde{N}_t^i := dN_t^i - \lambda_t^i dt$ and dW_t^i are martingales and X_{t-} and λ_{t-} are predictable, we find for $0 \le t < u$

$$\Phi(\nu_u^M) = \Phi(\nu_t^M) + \int_t^u \left(\mathcal{C}_s^M + \mathcal{D}_s^M + \mathcal{J}_s^M\right) ds + \mathcal{M}_u - \mathcal{M}_t,$$

where $(\mathcal{M}_t)_{t\ge 0}$ is an initial mean-zero martingale and

$$\mathcal{C}_t^M := \sum_{n=1}^N \frac{\partial \varphi(\nu_t^M(\mathbf{f}))}{\partial f_n}\left(\nu_t^M(\mathcal{L}^1 f_n)\nu_t^M(I) - \nu_t^M(\mathcal{L}^2 f_n) + \nu_t^M(\mathcal{L}^3 f_n)\right),$$
$$\mathcal{D}_t^M := \frac{1}{2M^2}\sum_{n,l=1}^N \frac{\partial^2 \varphi(\nu_t^M(\mathbf{f}))}{\partial f_n \partial f_l}\sum_{i=1}^M \left((\sigma^i)^2 \frac{\partial f_n(X_t^i, \lambda_t^i)}{\partial x}\frac{\partial f_l(X_t^i, \lambda_t^i)}{\partial x}\right)$$
$$\mathcal{J}_t^M := \sum_{i=1}^M \left[\varphi(\nu_t^M(\mathbf{f}) + J_t^{M,i}(\mathbf{f})) - \varphi(\nu_t^M(\mathbf{f}))\right]\lambda_t^i,$$

where $J_t^{M,i}(\mathbf{f}) = (J_t^{M,i}(f_1), ..., J_t^{M,i}(f_N))$ and

$$J_t^{M,i}(f) := \frac{1}{M}(f(X_{s-}^i + c^i) - f(X_s^i)).$$

We will need the following result given in Theorem 8 in [26]:

Theorem 4.3.1 (Propagation of chaos result for the Hawkes process)**.** *Consider the Hawkes process in the sense of* (4.2)*. For each $M \ge 1$ consider the complete graph with nodes I_M. Let $g : [0, \infty) \to \mathbb{R}$ be a locally square-integrable function and set $g^{i,j} = M^{-1} g$ for all $i, j \in I_M$. Define the limit equation*

$$\bar{N}_t = \int_0^t \int_0^\infty \mathbb{1}_{\{z \le (\mu_t + \int_0^s g(t-s) d\mathbb{E}[\bar{N}_s])\}}\pi(ds, dz), \tag{4.10}$$

where $\pi(ds,dz)$ *is a Poisson measure on* $[0,\infty)\times[0,\infty)$ *with intensity measure* $dsdz$*. Then we have* $d\mathbb{E}[\bar{N}_t]=\bar{\lambda}_t dt$ *and*

$$\bar{\lambda}_t := \mu + \int_0^t g(t-s)d\mathbb{E}[\bar{N}_s]. \tag{4.11}$$

In other words $\bar{N}=(\bar{N}_t)_{t\geq 0}$ *is an inhomogeneous Poisson process with intensity* $\bar{\lambda}_t$*. Let* $\bar{N}_t^i$ *be an i.i.d. family of solutions to (4.10) for* $i\in I_M$*. Define* $\Delta_M^i(t)=\int_0^t|d(\bar{N}_u^i-N_u^i)|$ *and* $\gamma_M(t)=\mathbb{E}[\Delta_M^i(t)]$*. Note that this* $\gamma_M(t)$ *does not depend on* i *due to exchangeability of both* $\bar{N}_t^i$ *and* N_t^i*. Then,*

$$\gamma_M(t)=\int_0^t \mathbb{E}\left[|\bar{\lambda}_t-\lambda_t^i|\right]ds,$$

and for $t\in[0,T]$ *we have*

$$\lim_{M\to\infty}\gamma_M(t)=0.$$

In other words, when all nodes interact in the same way in the limit of the number of nodes going to infinity, the Hawkes process reduces to an inhomogeneous Poisson process and we have for any $i\in I_M$ the following limit

$$\lim_{M\to\infty}\mathbb{E}\left[\int_t^u|\lambda_s^i-\bar{\lambda}_s|ds\right]=0. \tag{4.12}$$

The above convergence is thus presumed to be in the weak sense.

The task is now to find the generator of the limiting martingale problem which we will use to determine the process governing the dynamics of the monetary reserves in the limit, see e.g. Theorem 8.2 in Chapter 4 of [29]. For this we will use (4.12) and define a Taylor-based simplification of $\mathcal{J}_t^M$ as

$$\tilde{\mathcal{J}}_t^M := \sum_{n=1}^N \frac{\partial\varphi(\nu_t^M(\mathbf{f}))}{\partial x_n}\left[\frac{1}{M}\sum_{i=1}^M \bar{\lambda}_t\frac{\partial f_n(X_t^i)}{\partial x}c^i\right].$$

Using the triangle inequality we have

$$\begin{aligned}
&\mathbb{E}\left[\int_t^u|\mathcal{J}_s^M-\tilde{\mathcal{J}}_s^M|ds\right]\\
&\leq \mathbb{E}\left[\int_t^u\left|\sum_{i=1}^M\left[\varphi(\nu_s^M(\mathbf{f})+J_s^{M,i}(\mathbf{f}))-\varphi(\nu_t^M(\mathbf{f}))\right]\lambda_s^i-\sum_{i=1}^M\left[\sum_{n=1}^N\frac{\partial\varphi(\nu_s^M(\mathbf{f}))}{\partial x_n}J_s^{M,i}(\mathbf{f})\right]\lambda_s^i\right|ds\right]\\
&+\mathbb{E}\left[\int_t^u\left|\sum_{i=1}^M\left[\sum_{n=1}^N\frac{\partial\varphi(\nu_s^M(\mathbf{f}))}{\partial x_n}J_s^{M,i}(\mathbf{f})\right]\lambda_s^i-\sum_{i=1}^M\left[\sum_{n=1}^N\frac{\partial\varphi(\nu_s^M(\mathbf{f}))}{\partial x_n}\tilde{J}_s^{M,i}(\mathbf{f})\right]\lambda_s^i\right|ds\right]\\
&+\mathbb{E}\left[\int_t^u\left|\sum_{i=1}^M\left[\sum_{n=1}^N\frac{\partial\varphi(\nu_s^M(\mathbf{f}))}{\partial x_n}\tilde{J}_s^{M,i}(\mathbf{f})\right]\lambda_s^i-\sum_{i=1}^M\left[\sum_{n=1}^N\frac{\partial\varphi(\nu_s^M(\mathbf{f}))}{\partial x_n}\tilde{J}_s^{M,i}(\mathbf{f})\right]\bar{\lambda}_s\right|ds\right].
\end{aligned}$$

Applying a Taylor expansion to $f \in C_K^\infty(\mathcal{O})$ and using the boundedness of its derivatives and the definition $c^i = \hat{c}^i/M$, we find

$$J_t^{M,i}(f) \simeq \tilde{J}_t^{M,i}(f), \tag{4.13}$$

where $a^M \simeq b^M$ means $\lim_{M\to\infty} |a^M - b^M| = 0$ and

$$\tilde{J}_t^{M,i}(f) := \frac{1}{M}\frac{\partial f(X_t^i)}{\partial x}c^i.$$

Similarly, using the Taylor expansion of $\varphi \in C_K^\infty(\mathbb{R}^N)$ we have

$$\varphi(\nu_t^M(\mathbf{f}) + J_t^{M,i}(\mathbf{f})) - \varphi(\nu_t^M(\mathbf{f})) \simeq \sum_{n=1}^N \frac{\partial\varphi(\nu_t^M(\mathbf{f}))}{\partial x_n} J_t^{M,i}(\mathbf{f}). \tag{4.14}$$

Using the finiteness of λ_t^i from Proposition 4.1.3, equations (4.14) and (4.13), the boundedness of the derivatives of $f \in C_K^\infty(\mathcal{O})$ by their supremum, i.e. $||f|| = \sup_{(p,x)\in\mathcal{O}} |f(p,x)|$ and the bounds on the intensity given in (4.12) we have that

$$\lim_{M\to\infty} \mathbb{E}\left[\int_t^u |\mathcal{J}_s^M - \tilde{\mathcal{J}}_s^M| ds\right] = 0.$$

Similarly we have

$$\lim_{M\to\infty} \mathbb{E}\left[\int_t^u |\mathcal{D}_s^M| ds\right] = 0.$$

Define the operator $\mathcal{A}$ acting on the function $\Phi(\nu)$ defined in (4.9), as

$$\mathcal{A}\Phi(\nu) := \sum_{n=1}^N \frac{\partial\varphi(\nu_t^M(\mathbf{f}))}{\partial f_n}\left(\nu_t^M(\mathcal{L}^1 f_n)\nu_t^M(I) - \nu_t^M(\mathcal{L}^2 f_n) + \nu_t^M(\mathcal{L}^3 f_n) + \nu_t^M(\mathcal{L}^4 f_n)\right), \tag{4.15}$$

where $\mathcal{L}^4 := c\bar{\lambda}_t\partial_x$. Then we have the result:

Lemma 4.3.2 (Limiting martingale problem)**.** *For any $\Phi \in \mathcal{S}$ and $0 \le t_1 \le ... \le t_{m+1} \le \infty$, with $m \in \mathbb{N}$ and $\Psi_j \in L^\infty(S)$ we have that $\mathcal{A}$ is the generator of the limiting martingale problem, i.e.*

$$\lim_{M\to\infty} \mathbb{E}\left[\left(\Phi(\nu_{t_{m+1}}^M) - \Phi(\nu_{t_m}^M) - \int_{t_m}^{t_{m+1}} \mathcal{A}\Phi(\nu_u^M)du\right)\prod_{j=1}^m \Psi_j(\nu_{t_j}^M)\right] = 0. \tag{4.16}$$

4.3.2 Limiting process

Given the limiting martingale problem (4.16) and assuming the existence and uniqueness of a limit point, we wish to find the limiting process ν_t which satisfies equation (4.16). Let $\mathbf{p} = (p^*, x)$. Define the following measure-valued process by

$$\nu_t(A) := \mathbb{P}(X_t(\mathbf{p}) \in A), \tag{4.17}$$

where $A \in \mathcal{B}(\mathbb{R})$ and the underlying limiting state process $X(\mathbf{p}) = (X_t(\mathbf{p}))_{t\geq 0}$ is a Markovian diffusion with time-varying coefficients given by

$$X_t(\mathbf{p}) = x + \int_0^t \left(a\left(Q_1(s) - X_s(\mathbf{p})\right) + c\bar{\lambda}_s\right) ds + \sigma \int_0^t dW_s, \qquad t \geq 0, \tag{4.18}$$

with $\bar{\lambda}_t$ is defined in (4.11) and

$$Q_1(t) = x + c\int_0^t \bar{\lambda}_s ds. \tag{4.19}$$

Notice that $Q_1(t)$ satisfies the integral equation

$$Q_1(t) = e^{-at}\left(x + \int_0^t e^{as}\left(aQ_1(s) + c\bar{\lambda}_s\right) ds\right).$$

Using the definition of ν in (4.17) we have that

$$\nu_t(I) = \int_{\mathcal{O}} x\nu_t(dx) = \mathbb{E}\left[X_t(\mathbf{p})\right],$$

where the underlying state process $X_t(\mathbf{p})$ is given by (4.18). Notice that

$$\mathbb{E}\left[X_t(\mathbf{p})\right] = e^{-at}\left(x + \int_0^t e^{as}(aQ_1(s) + c\bar{\lambda}_s\right) ds,$$

from which it follows that

$$Q_1(t) = \nu_t(I), \tag{4.20}$$

where $I(x) = x$. We now prove that δ_ν indeed satisfies the martingale problem in Lemma 4.3.2:

Theorem 4.3.3 (Limiting process)**.** *The empirical measure-valued process ν^M admits the weak convergence $\nu^M \to \nu$, as $M \to \infty$, where ν is defined as in* (4.17)*. Furthermore, $\nu^M(I) \to Q_1$.*

Proof. Using the standard analysis of weak convergence as in Chapter 3 of [29], the weak convergence $\nu^M \to \nu$ as $M \to \infty$ follows from Lemma 4.3.2 and Lemmas 4.5.2, 4.5.3 and uniqueness of the limit point. In other words, if we define $\mathbb{Q}^M := \mathbb{P}(\nu^M \in \mathcal{B}(D_S[0,\infty)))$, we have that $\mathbb{Q}^M$ converges to the solution $\mathbb{Q}$ of the martingale problem generated by $\mathcal{A}$ in (4.15). Next we show that $\mathbb{Q} = \delta_\nu$, i.e. the limit measure-valued process ν can indeed be represented as in (4.17). We have for $f \in C_K^\infty(\mathcal{O})$ using the definition in (4.6) that

$$\nu_t(f) = \mathbb{E}[f(X_t(\mathbf{p}))]. \tag{4.21}$$

On the other hand, from (4.18) and using Itô's lemma, we have

$$f(X_t(\mathbf{p})) = f(x) + \int_0^t \frac{\partial f}{\partial x}(X_s(\mathbf{p}))(aQ_1(s) - aX_s(\mathbf{p}) + c\bar{\lambda}_s)ds + \frac{\sigma^2}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(X_s(\mathbf{p}))ds$$

$$+ \sigma \int_0^t \frac{\partial f}{\partial x}(X_s(\mathbf{p}))dW_s.$$

Then recalling the definition of the operators $\mathcal{L}^*$ and the equality $Q_1(t) = \nu_t(I)$ from (4.20) we have

$$\begin{aligned}\frac{\partial}{\partial t}\mathbb{E}[f(X_t(\mathbf{p}))] =& \frac{1}{2}\mathbb{E}\left[\sigma^2 \partial_{xx} f(X_t(\mathbf{p}))\right] + Q_1(t)\mathbb{E}[a\partial_x f(X_t(\mathbf{p}))] + \mathbb{E}[c\bar{\lambda}_t \partial_x f(X_t(\mathbf{p}))] \\ &- \mathbb{E}[aX_s(\mathbf{p})\partial_x f(X_t(\mathbf{p}))] \\ =& \mathbb{E}[\mathcal{L}^3 f(X_t(\mathbf{p}))] + \nu_t(I)\mathbb{E}[\mathcal{L}^1 f(X_t(\mathbf{p}))] + \mathbb{E}[\mathcal{L}^4 f(X_t(\mathbf{p}))] - \mathbb{E}[\mathcal{L}^2 f(X_t(\mathbf{p}))].\end{aligned}$$

Using (4.21) we find

$$\begin{aligned}\frac{d\Phi(\nu_t)}{dt} &= \sum_{n=1}^N \frac{\partial \varphi}{\partial x_n}(\nu_t(\mathbf{f}))\frac{d\nu_t(f_n)}{dt} \\ &= \sum_{n=1}^N \frac{\partial \varphi}{\partial x_n}\left(\nu_t(\mathcal{L}^3 f) + \nu_t(\mathcal{L}^1 f)\nu_t(I) + \nu_t(\mathcal{L}^4 f) - \nu_t(\mathcal{L}^2 f)\right) \\ &= \mathcal{A}\Phi(\nu_t).\end{aligned}$$

Then for all functions $\Phi(\cdot)$ of the form (4.9) we have

$$\Phi(\nu_t) = \Phi(\nu_s) + \int_s^t \mathcal{A}\Phi(\nu_u)du, \qquad 0 \leq s < t < \infty,$$

and hence δ_ν satisfies the martingale problem generated by $\mathcal{A}$. □

In other words, the propagation of chaos result from Theorem 4.3.3 tells us that the empirical mean ν^M converges to a measure ν whose underlying process $X_t(\mathbf{p})$ reflects the Hawkes process through a *time-dependent* drift.

4.3.3 Extensions of the model

In this section we shortly present results for several possible extensions of the results presented in Section 4.3. In particular we derive the limiting empirical distribution when including a compound Hawkes process in the monetary reserve model considered in (4.3); a systematic risk factor, where the derivation is based on the result from [34]; and furthermore prove a central limit theorem based on [66] which quantifies the fluctuation of the empirical distribution around its large system limit.

Compound Hawkes process

If we include a compound Hawkes process in the initial log-monetary reserve SDE, i.e.

$$dX_t^i = \frac{a^i}{M}\sum_{k=1}^M (X_t^k - X_t^i)dt + \sigma^i dW_t^i + c^i dS_t^i,$$

where

$$S_t^i = \sum_{j=1}^{N_t^i} Z_j^i,$$

where Z is an i.i.d. random variable with distribution function F, independent of N_t^i and W_t^i, such that $\lim_{M\to\infty} \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{Z^i} = y$. Then the limiting process is given by

$$X_t(\mathbf{p}) = x + \int_0^t \left(a\left(Q_1(s) - X_s(\mathbf{p})\right) + cy\bar{\lambda}_s\right) ds + \sigma \int_0^t dW_s, \qquad t \geq 0.$$

Systematic risk factor exposure

Similar to the analysis of [34] we can show that considering a non-vanishing systematic risk factor common to all the nodes in the system, we obtain a non-deterministic limiting behavior. Let $\mathcal{V}_t = \sigma(V_s, 0 \leq s \leq t)$ and $\mathcal{F}_t = \sigma((V_s, N_s^i, W_s^i), 0 \leq s \leq t, i \in \mathbb{N})$. Consider the following model for the log-monetary reserves

$$\begin{aligned} dX_t^i &= a^i(\bar{X}_t - X_t^i)dt + \sigma^i dW_t^i + c^i dN_t^i + \beta^i dY_t, \\ dY_t &= b_0(Y_t)dt + \sigma_0(Y_t)dV_t, \;\; Y_0 = y_0, \end{aligned}$$

where V_t is a standard Brownian motion independent of W_t^i and N_t^i. In other words, W_t^i represents a source of risk which is idiosyncratic to a specific name, while Y_t is a systematic risk factor driven by a Brownian motion that is common to all the nodes in the network with the parameter β^i representing the sensitivity of node i to the Y. The systematic risk factor causes correlated changes in the monetary reserve process and thus acts as an additional source of clustering. As usual we assume $p^i := (a^i, \sigma^i, c^i, \beta^i) \to p^* := (a, \sigma, c, \beta)$. Following the derivation in [34] and defining $\Phi(y, \nu) = \varphi_1(y)\varphi_2(\nu(\mathbf{f}))$, and applying Itô's lemma as in the derivations for the original model we obtain for $0 \leq t < u$

$$\begin{aligned} \Phi(Y_u, \nu_u^M) =& \Phi(Y_t, \nu_t^M) + \int_t^u (\varphi_1(Y_s)\mathcal{C}_s^M + \varphi_1(Y_s)\mathcal{D}_s^M + \varphi_1(Y_s)\mathcal{J}_s^M + \mathcal{B}_s^{M,1})ds \\ &+ \int_t^u \mathcal{B}^{M,2} dV_s + \mathcal{M}_u - \mathcal{M}_t, \end{aligned}$$

where we have defined

$$\begin{aligned} \mathcal{B}_t^{M,1} :=& \varphi_1(Y_t) \sum_{n=1}^{N} \frac{\partial \varphi_2(\nu_t^M(\mathbf{f}))}{\partial f_n} \nu_t^M(\mathcal{L}_{Y_t}^5 f_n) + \varphi_2(\nu(\mathbf{f})) \left(b_0(Y_t)\partial_y \varphi_1(Y_t) + \frac{1}{2}\sigma_0^2(Y_t)\partial_{yy}\varphi_1(Y_t) \right) \\ &+ \partial_y \varphi_1(Y_t) \sum_{n=1}^{N} \frac{\partial \varphi_2(\nu(\mathbf{f}))}{\partial f_n} \sigma_0(y) \nu_t^M(\mathcal{L}_{Y_t}^6 f_n) \end{aligned}$$

$$\mathcal{B}_t^{M,2} := \varphi_1(Y_t) \sum_{n=1}^{N} \frac{\partial \varphi_2(\nu_t^M(\mathbf{f}))}{\partial f_n} \nu_t^M(\mathcal{L}^6_{Y_t} f_n) + \sigma_0(Y_t)\partial_y \varphi_1(Y_t)\varphi_2(\nu(\mathbf{f})),$$

with $\mathcal{L}^5_y f(p,x) := \beta^i b_0(y)\partial_x f(p,x) + \frac{1}{2}(\beta^i)^2\sigma_0^2(y)\partial_x f(p,x)$ and $\mathcal{L}^6_y f(p,x) := \beta^i \sigma_0(y)\partial_x f(p,x)$. Taking the limit of $M \to \infty$, using the limits derived in Section 4.3.1, the vanishing of the martingale in the limit (see also Lemma 7.2 in [34]) and defining

$$\nu_t(f) = \mathbb{E}[f(X_t(\mathbf{p})|\mathcal{V}_t],$$

with

$$X_t(\mathbf{p}) = x + \int_0^t \left(a\left(\nu_t(I) - X_s(\mathbf{p})\right) + c\bar{\lambda}_s\right) ds + \sigma \int_0^t dW_s + \beta \int_0^t dY_s,$$

we obtain for the limiting process ν_t the following SPDE

$$\begin{aligned} d\nu_t(f(X_t)) = &\left(\nu_t(\mathcal{L}^1 f(X_t))\nu_t(I) - \nu_t(\mathcal{L}^2 f(X_t)) + \nu_t(\mathcal{L}^3 f(X_t)) + \nu_t(\mathcal{L}^4 f(X_t)) + \nu_t(\mathcal{L}^5_{Y_t} f(X_t))\right) dt \\ &+ \nu_t(\mathcal{L}^6_{Y_t} f(X_t)) dV_t, \end{aligned}$$

where we use Lemmas B.1 and B.2 in [34] to show that $\mathbb{E}\left[\int_0^t X_s dV_s|\mathcal{V}_t\right] = \int_0^t \mathbb{E}[X_s|\mathcal{V}_s]dV_s$. The systematic risk factor thus does not vanish in the limit, and results in the stochastic partial differential equation for the limiting process of the empirical measure, instead of the deterministic behavior in the original model.

A Central Limit Theorem result

Consider again the model defined in (4.3). In order to improve the first-order approximation of ν_t^M given in (4.17), we can analyze the fluctuations of ν^M around its large system limit ν. Following [66] we define

$$\Xi_t^M = \sqrt{M}(\nu_t^M - \nu_t).$$

The signed-measure-valued process Ξ^M weakly converges to the fluctuation limit $\bar{\Xi}$ in an appropriate space (in particular the convergence is considered in weighted Sobolev spaces in which the sequence Ξ^M, $M \in \mathbb{N}$ can be shown to be relatively compact; for discussion on this space, as well as the existence and uniqueness of the limiting point, we refer to Sections 7,8 and 9 in [66]). We start by deriving an expression for Ξ_t^M. Some terms in this expression will vanish in the limit of $M \to \infty$, and using the tightness of the processes (see Section 8 in [66]) and continuity of the operators in the expression for Ξ^M we can pass to the limit and find the expression that the limiting fluctuation process satisfies.

Subtracting ν_t from ν_t^M we find

$$d\Xi_t^M(f) = \left(\nu_t^M(\mathcal{L}^1 f)\Xi_t^M(I) + \nu_t(I)\Xi_t^M(\mathcal{L}^1 f) - \Xi_t^M(\mathcal{L}^2 f) + \Xi_t^M(\mathcal{L}^3 f) + \Xi_t^M(\mathcal{L}^4 f)\right) dt$$

$$+ d\mathcal{M}_t^M(f) + \sqrt{M}\frac{1}{M}\sum_{i=1}^{M}(f(X_t^i + c^i) - f(X_t^i))d\tilde{N}_t - \sqrt{M}\frac{1}{M}\sum_{i=1}^{M} c^i\frac{\partial f}{\partial x}\tilde{N}_t^i$$

$$+ \sqrt{M}\left(\frac{1}{M}\sum_{i=1}^{M}(f(X_t^i + c^i) - f(X_t^i))\lambda_t^i - \nu_t^M(\mathcal{L}^4 f)\right)dt,$$

where the martingale term is defined as

$$\mathcal{M}_t^M(f) = \sqrt{M}\left(\frac{1}{M}\sum_{i=1}^{M}\int_0^t \sigma^i\partial_x f dW_s^i + \int_0^t \frac{1}{M}\sum_{i=1}^{M} c^i\frac{\partial f}{\partial x}d\tilde{N}_s^i\right).$$

Using the limiting expressions for the Hawkes jump term and a Taylor approximation from Section 4.3.1, we have

$$\sqrt{M}\left|\frac{1}{M}\sum_{i=1}^{M}(f(X_t^i + c^i) - f(X_t^i)) - \frac{1}{M}\sum_{i=1}^{M} c^i\frac{\partial f}{\partial x}\right| \leq \frac{K^2}{M\sqrt{M}}\left\|\frac{\partial^2 f}{\partial x^2}\right\|. \tag{4.22}$$

Thus one can show by taking the limit $M \to \infty$, using (4.22) and Assumption 4.2.1 that the sequence $\{\Xi_t^M, t \in [0,T]\}_{M\in\mathbb{N}}$ converges in distribution to the limit point $\{\Xi_t \in [0,T]\}$ that satisfies

$$\begin{aligned}\Xi_t(f) =& \Xi_0(f) + \int_0^t \big(\nu_s^M(\mathcal{L}^1 f)\Xi_s(I) + \nu_s(I)\Xi_s(\mathcal{L}^1 f) - \Xi_s(\mathcal{L}^2 f) \\ &+ \Xi_s(\mathcal{L}^3 f) + \Xi_s(\mathcal{L}^4 f)\big)ds + \mathcal{M}_t(f),\end{aligned}$$

where $\{\mathcal{M}_t, t \in [0,T]\}$ is the distribution valued, continuous square integrable martingale with a deterministic quadratic variation to which the sequence $\{\mathcal{M}_t^M, t \in [0,T]\}_{M\in\mathbb{N}}$ converges in distribution (note: unlike in the LLN cases, the martingale term does not vanish in the CLT scaling case). By a martingale CLT (see 7.1.4 in [29]) $\mathcal{M}$ is Gaussian. This implies the following second-order approximation $\nu_t^M \overset{d}{\approx} \nu_t + \frac{1}{\sqrt{M}}\Xi_t$, giving a more accurate approximation for finite banking systems.

4.4 Systemic risk in a large network

In this section we introduce several systemic risk indicators to quantify the risk in our network and to show the particular dependence of the risk on the underlying parameters. We first remark on the difference between the monetary reserve with a Hawkes process and one with an independent Poisson process:

Remark 4.4.1 (Independent Poisson process versus Hawkes process)**.** Consider an independent Poisson process with intensity μ. It is straightforward to see that

$$\bar{\lambda}_t := \mu + \int_0^t \alpha e^{-\beta(t-s)}\bar{\lambda}_s ds \geq \mu,$$

since we assume $\alpha, \beta \geq 0$. Therefore, for $c < 0$ we have that $Q_1(t) \leq \tilde{Q}_1(t)$, with Q_1 and $\tilde{Q}_1$ being the averages from a Poisson jump with intensity λ_t and a jump with intensity μ respectively. Thus, in the limit $M \to \infty$, using $\nu^M(I) \to Q_1(t)$, we have as expected that the Hawkes process increases the default risk in the network.

4.4.1 Risk indicators

Here we show how one can measure the systemic risk in a large network using the limiting dynamics $X_t(\mathbf{p})$. We propose to compute systemic risk in the mean-field model based on the fraction of banks that have transitioned from a normal to a defaulted state. We define the risk indicator as the expected value of the fraction of banks that throughout time $t \in [0, T]$ have dropped below the default level D,

$$\mathrm{SR}^M := \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\left\{ \min\limits_{0 \leq t \leq T} X_t^i \leq D \right\}}.$$

Note that from Theorem 4.3.3 we have $\lim\limits_{M \to \infty} \nu_t^M = \nu_t$ for a continuous function f of X_t^i. For the indicator function over $t \in [0, T]$, we consider the approximate relationship to hold

$$\lim_{M \to \infty} \mathrm{SR}^M \approx \mathbb{E}\left[\mathbb{1}_{\left\{ \min\limits_{0 \leq t \leq T} X_t(\mathbf{p}) \leq D \right\}} \right],$$

in which the average over the indicator function of the M monetary reserve processes is thus replaced by the indicator of the limiting process.

Furthermore, similar to [8], we can define the average distance to default as

$$\mathrm{ADD}^M(t) := \mathbb{E}\left[\frac{1}{M} \sum_{i=1}^{M} X_t^i \right].$$

Note that $(\nu_t^M; M \in \mathbb{R})$ is uniformly integrable, i.e. for each $t \geq 0$

$$\sup_{M \in \mathbb{N}} \mathbb{E}\left[\left| \nu_t^M(I) \right|^2 \right] < \infty,$$

the proof of which is similar to the proof of Lemma 4.5.1 in Appendix 4.5 and Lemma B.2 in [8]. Then, for the average distance to default indicator we use the following limiting result

$$\lim_{M \to \infty} \mathrm{ADD}^M(t) = Q_1(t)$$

with $Q_1(t)$ as in (4.20). Note that in the case of independent Poisson jumps with intensity λ, the limit of the ADD indicator is given by $\lim\limits_{M \to \infty} \mathrm{ADD}^M(t) = x + c\lambda t$. This is in contrast to the case of the Hawkes jumps for which we have $\lim\limits_{M \to \infty} \mathrm{ADD}^M(t) = x + c \int_0^t \bar{\lambda}_s ds$.

4.4.2 Numerical results

We set $M = 300$, i.e. sufficiently large, and analyze how our approximation formulas for the various indicators of systemic risk compare to the corresponding Monte-Carlo estimate. The latter is obtained by simulating M interacting processes X_t^i, $i \in I_M$ using an Euler approximation of (4.3).

Remark 4.4.2 (Computation of $\bar{\lambda}_t$)**.** Define the partition of $[0, T]$ as $0 = t_0 < t_1 < ... < t_K = T$ with $\Delta t := t_i - t_{i-1}$. Then we approximate the integral in (4.11) as

$$\bar{\lambda}_{t_{i+1}} \approx \bar{\lambda}_{t_i} + \Delta t g(\Delta t)\bar{\lambda}_{t_i},$$

and $\bar{\lambda}_0 := \mu$. Using the approximated $\bar{\lambda}_t$ we compute $Q_1(t)$ as

$$Q_1(t_{i+1}) \approx Q_1(t_i) + \Delta t c\bar{\lambda}_{t_i},$$

where $Q_1(0) = x$.

Table 4.2: Monte Carlo estimates versus the LLN approximation for the systemic risk indicators with $\mu = 0.01$, $\alpha = 1$, $\beta = 1.2$, $a = 0.5$, $\sigma = 0.5$, $\hat{c} = -0.2$ and $D = 0$.

	Monte Carlo		Approximation	
x_0	SR	ADD(T)	SR	ADD(T)
0.002	0.945	0.007	0.949	0.007
0.1	0.821	0.096	0.816	0.096
0.2	0.658	0.197	0.652	0.197
0.5	0.252	0.497	0.261	0.497
0.8	0.057	0.797	0.058	0.797
1	0.016	0.998	0.017	0.997

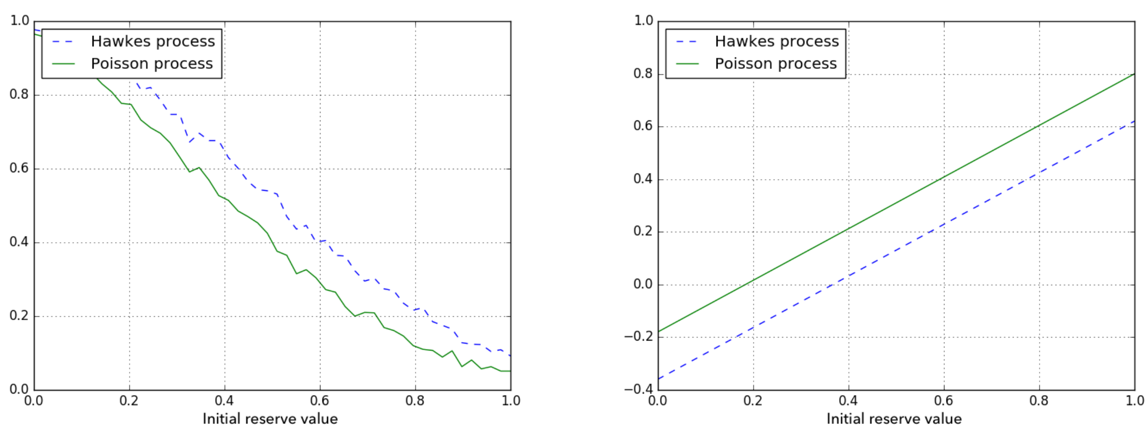
In Tables 4.2 and 4.3 we present the results for our approximation and the Monte-Carlo estimates for 5000 simulations, 100 time steps, $T = 1$ and $M = 300$. As expected the systemic risk in the network, as quantified by both SR and ADD, decreases as the initial monetary reserve value increases. Furthermore, a higher mean jump intensity μ results in a less stable network. In Figure 4.4.2 we show the LLN estimates for the systemic risk and the average distance to default for the Hawkes and Poisson process for different values of the initial reserve x_0. Our claims of the Hawkes process adding an additional default risk in the model are verified also in these numerical results, as the systemic risk indicator for the Hawkes process is considerably larger, while the average monetary reserves are consistently lower than for an independent Poisson process. Therefore, the self- and

Table 4.3: Monte Carlo estimates versus the LLN approximation for the systemic risk indicators with $\mu = 0.05$, $\alpha = 1$, $\beta = 1.2$, $a = 0.5$, $\sigma = 0.5$, $\hat{c} = -0.2$ and $D = 0$.

	Monte Carlo		Approximation	
x_0	SR	ADD(T)	SR	ADD(T)
0.01	0.947	-0.005	0.946	-0.007
0.1	0.826	0.085	0.830	0.083
0.2	0.669	0.186	0.653	0.183
0.5	0.262	0.486	0.269	0.483
0.8	0.061	0.785	0.061	0.783
1	0.017	0.985	0.016	0.0.983

cross-exciting shock modelled through the Hawkes process is an additional form of contagion in the network, resulting in the network being more prone to a systemic risk event.

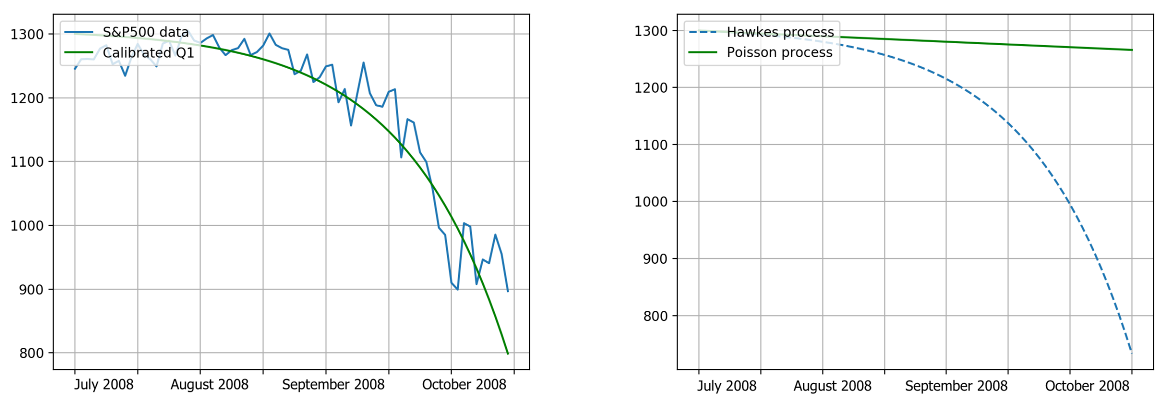
Figure 4.5: LLN estimates for the systemic risk (L) and LLN estimates for the average distance to default (R) at time $T = 1$ with $\mu = 0.2$, $\alpha = 1.2$, $\beta = 1.2$, $a = 0.5$, $\sigma = 0.5$, $c = -1$ and $D = 0$ for a independent Poisson process and the Hawkes process for $x_0 \in [0, 1]$



Calibrating the model

Calibration of the model considered in (4.3) with heterogeneous coefficients, in particular for the large banking system, is a complex task. In [1] the authors considered a calibration for a Hawkes diffusion model used to model asset returns and developed method of moments estimates for the parameters of the model. Even after making simplifying assumptions on the intensity, the model was fitted only on pairs of assets. The calibration of the mean-field SDE with Hawkes jumps for a large number of banks is therefore besides the scope of this chapter and left for further research. However, the limiting expression derived in Section 4.3 can be used to derive a simple and efficient way of calibrating the model. In particular, we can calibrate the average distance to default given by $Q_1(t)$ in (4.19) by fitting it to an average of a *sufficiently large* number of assets, resulting in the calibrated parameters x, c, μ, α and β. In particular, consider the asset price as a proxy for the monetary reserve process and consider the average of the components of the S&P500 index over the period of 2008-07-14 until 2008-10-21. Calibrating the deterministic expression for $Q_1(t)$ to the actual average distance to default we obtain the following set of parameters: $\mu = 0.3$, $x = 1300$, $\alpha = 0.07$, $\beta = 0.11$ and $c = -1.6$. It can be argued that the assumption of regularity of the parameters in the limit (see Assumption 4.2.1) is too strong and disables calibrating to actual excitation. Nevertheless, using this simple and efficient way of calibrating the model, we see from the left-hand side of Figure 4.6 that contagion is sufficiently captured; in particular note that the Poisson process is unable to model the necessary contagion as seen from the right-hand side of Figure 4.6, while the SDE with the Hawkes process provides a much better fit.

Figure 4.6: Calibrated model for Q_1 on the S&P500 data showing excitation effects (L) and the average of 5000 simulated SDE paths of $X_t(\mathbf{p})$ (R)



4.5 Existence and uniqueness

In this section we briefly discuss the ingredients needed to show existence and uniqueness of the limiting process. We first introduce an auxiliary Lemma which is a boundedness result of the moment estimate of the log-monetary reserve process.

Lemma 4.5.1. *For $n = 1, 2$ and $T \geq 0$ we have*

$$\sup_{0\leq t\leq T,\, M\in\mathbb{N}} \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}\left[\left|X_t^i\right|^n\right] < +\infty.$$

Proof. Let $n \in \{1, 2\}$. Recall the constant C_p bounding the parameters (p^i, X_0^i) from assumption 4.2.1. From Itô's formula we have

$$\begin{aligned}\mathbb{E}\left[|X_t^i|^n\right] =& \mathbb{E}\left[|X_0^i|^n\right] + a^i\mathbb{E}\left[\int_0^t n|X_s^i|^{n-1}(\bar{X}_s - X_s^i)ds\right] + \frac{1}{2}(\sigma^i)^2\mathbb{E}\left[\int_0^t n(n-1)|X_s^i|^{n-2}\right] \\ &+ \sigma^i\mathbb{E}\left[\int_0^t n|X_s^i|^{n-1}dW_s^i\right] + \mathbb{E}\left[\int_0^t \left[|X_s^i + c^i|^n - |X_s^i|^n\right]dN_s^i\right].\end{aligned}$$

Using Young's inequality we have

$$\begin{aligned}a^i n X_t^{i\,n-1}\bar{X}_t - a^i n|X_t^i|^{n-1}X_t^i &\leq a^i\frac{n}{M}\sum_{k=1}^{M}|X_t^i|^{n-1}|X_t^k| - a^i n|X_t^i|^n \\ &\leq C_p\frac{1}{M}\sum_{k=1}^{M}|X_t^k|^n + (2n-1)C_p|X_t^i|^n.\end{aligned}$$

Applying Young's inequality twice yields

$$\begin{aligned}\frac{n(n-1)}{2}(\sigma^i)2|X_t^i|^{n-2} &\leq \frac{n(n-1)}{2}\left(\frac{n-2}{n-1}|X_t^i|^{n-1} + \frac{1}{n-1}(\sigma^i)^{2n}\right) \\ &\leq \frac{n(n-1)}{2}\left(\frac{n-2}{n}|X_t^i|^n + \frac{1}{n} + \frac{1}{n-1}C_p\right).\end{aligned}$$

Finally, using Young's inequality and Proposition 4.1.3 there exists a constant C_n independent of M such that

$$\begin{aligned}\mathbb{E}\left[\int_0^t \left[|X_s^i + c^i|^n - |X_s^i|^n\right]dN_s^i\right] &= \mathbb{E}\left[\int_0^t \left[|X_s^i + c^i|^n - |X_s^i|^n\right]\lambda_s^i ds\right] \\ &\leq \frac{1}{2}\mathbb{E}\left[\int_0^t |c^i X_s^i|^{2(n-1)} + |c^i|^{2n}ds\right] + \frac{1}{2}\mathbb{E}\left[\int_0^t (\lambda_s^i)^2 ds\right] \\ &\leq C_n(1 + \mathbb{E}\left[\int_0^t |X_s^i|^n ds\right].\end{aligned}$$

The statement then follows from applying Gronwall's Lemma and the fact that the limiting constants are independent of M. □

In order to conclude weak convergence of the empirical measure ν_t^M to ν_t we need to determine the limiting martingale problem (as done in Section 4.3.1), show uniqueness of the limit point and its existence (i.e. tightness of the sequence of measure-valued processes). We first provide here a sketch of the proof for existence. We have to prove that the sequence of measure-valued processes

$\{\nu^M\}_{M\in\mathbb{N}}$ defined by (4.5) are relatively compact when viewed as a sequence of random processes on the Skorokhod space $D_S([0,\infty])$, the collection of càdlàg functions from $[0,\infty)$ to S. This is necessary to ensure that the laws of ν^M have at least one limit point (see also Chapters 2 and 3 of [29]). The complication arising from using a Hawkes process is the feedback loop in the intensity, however due to Theorem 4.3.1 we know that the intensity is bounded and thus the system will not explode. The relative compactness will be implied by the following two Lemmas: Lemma 4.5.2 on compact containment and Lemma 4.5.3 on the regularity of the ν^M's.

Lemma 4.5.2. *For every $T > 0$ and any smooth function $f \in C_K^\infty(\mathcal{O})$, we have*

$$\lim_{m\to\infty} \sup_{M\in\mathbb{N}} \mathbb{P}\left(\sup_{0\le t\le T} |\nu_t^M(f)| \ge m\right) = 0.$$

Proof. From (4.8) we have the following decomposition

$$\nu_t^M(f) = \nu_0^M(f) + A_t^M + B_t^M + C_t^M + D_t^M, \tag{4.23}$$

where we have defined

$$A_t^M := \frac{1}{M}\int_0^t \sum_{i=1}^M a^i \partial_x f(X_s^i)(\nu_s^M(I) - X_s^i)ds, \tag{4.24}$$

$$B_t^M := \frac{1}{2M}\int_0^t \sum_{i=1}^M (\sigma^i)^2 \partial_{xx} f(X_s^i)ds,$$

$$C_t^M := \frac{1}{M}\int_0^t \sum_{i=1}^M \left(\sigma^i \partial_x f(X_s^i) dW_s^i\right),$$

$$D_t^M := \int_0^t \left[\frac{1}{M}\sum_{i=1}^M (f(X_s^i + c^i) - f(X_{s-}^i))\right] dN_s^i.$$

Then we need to bound $\mathbb{E}\left[\sup_{0\le t\le T} |(\cdot)_t^M|\right]$ for each of the terms defined above. Denote for $f \in C_K^\infty(\mathcal{O})$ the supremum norm with $||f|| = \sup_{(p,x)\in\mathcal{O}} |f(p,x)|$. We will use the dominating constant C_p from Assumption 4.2.1. For A_t^M, B_t^M, C_t^M the estimates are similar to [8] and we omit the details here and just give the estimates

$$\mathbb{E}\left[\sup_{0\le t\le T} |A_t^M|\right] \le C_p \left\|\frac{\partial f}{\partial x}\right\| \int_0^T \frac{1}{M}\sum_{i=1}^M \mathbb{E}\left[|X_s^i|^2\right] ds + C_p \left\|\frac{\partial f}{\partial x}\right\|,$$

$$\mathbb{E}\left[\sup_{0\le t\le T} |B_t^M|\right] \le \frac{C_p}{2} \left\|\frac{\partial^2 f}{\partial x^2}\right\| T,$$

$$\mathbb{E}\left[\sup_{0\le t\le T} |C_t^M|\right] \le C_T C_p \left\|\frac{\partial f}{\partial x}\right\| (T+1).$$

Then we have by the mean-value theorem and using Proposition 4.1.3 which implies the existence of a constant C_λ such that $\mathbb{E}[\lambda_t^i] < C_\lambda$ that

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}|D_t^M|\right] &\le \sum_{i=1}^M \mathbb{E}\left[\int_0^T \frac{1}{M}|f(X_s^i+c^i)-f(X_{s-}^i)|dN_s^i\right]\\ &\le \left\|\frac{\partial f}{\partial x}\right\| \frac{1}{M}\sum_{i=1}^M c^i \int_0^T \mathbb{E}[\lambda_s^i]ds\\ &\le \left\|\frac{\partial f}{\partial x}\right\| C_p C_\lambda T.\end{aligned}$$

Using Lemma 4.5.1, we can find a positive constant C such that

$$\sup_{M\in\mathbb{N}} \mathbb{E}\left[\sup_{0\le t\le T}\left|\nu_t^M(f)\right|\right] < C.$$

□

Define $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot|\mathcal{F}_t]$.

Lemma 4.5.3. *Let $h(x,y) = |x-y| \wedge 1$ for any $x, y \in \mathbb{E}$. Then there exists a positive random variable $H_M(\gamma)$ with $\lim_{\gamma\to 0}\sup_{M\in\mathbb{N}} \mathbb{E}[H_M(\gamma)] = 0$ such that for all $0 \le t \le T$, $0 \le u \le \gamma$ and $0 \le v \le \gamma \wedge 1$, we have*

$$\mathbb{E}_t\left[h^2(\nu_{t+u}^M(f), \nu_t^M(f))h^2(\nu_t^M(f), \nu_{t-v}^M(f)\right] \le \mathbb{E}_t[H_M(\gamma)],$$

where the function $f \in C^\infty(\mathcal{O})$.

Proof. We have from (4.23)

$$(\nu_{t+u}^M - \nu_t^M)(f) = A_{t+u}^M - A_t^M + B_{t+u}^M - B_t^M + C_{t+u}^M - C_t^M + \mathcal{M}_{t+u}^M - \mathcal{M}_t^M + P_{t+u}^M - P_t,$$

where A_t^M, B_t^M, C_t^M are defined in (4.24) and

$$\begin{aligned}\mathcal{M}_t^M &:= \int_0^t \left[\frac{1}{M}\sum_{i=1}^M (f(X_s^i+c^i) - f(X_{s-}^i))\right] d\tilde{N}_s^i,\\ P_t^M &:= \int_0^t \left[\frac{1}{M}\sum_{i=1}^M (f(X_s^i+c^i) - f(X_{s-}^i))\right] \lambda_s^i ds,\end{aligned}$$

where we have used the fact that the compensated counting process $\tilde{N}_t^i := N_t^i - \int_0^t \lambda_s^i ds$ is a $\mathcal{F}_t$-local martingale. We have

$$\begin{aligned}h^2\left(\nu_{t+u}^M(f), \nu_t^M(f)\right) \le 16\big[&\left|A_{t+u}^M - A_t^M\right|^2 + \left|B_{t+u}^M - B_t^M\right|^2 + \left|C_{t+u}^M - C_t^M\right|^2\\ &+ \left|\mathcal{M}_{t+u}^M - \mathcal{M}_t^M\right|^2 + \left|P_{t+u}^M - P_t^M\right|^2\big].\end{aligned}$$

Let $0 \le u \le \gamma$. For the bounds on the first three differences we refer to Lemma 3.5 in [8]. For the fourth difference, using the martingale property and Itô Isometry for the martingale $(\mathcal{M}_t^M)$ with quadratic variation $[\tilde{N}_t, \tilde{N}_t] = N_t$, the mean-value theorem, Assumption 4.2.1 and Proposition 4.1.3, and the bound (6.1) in [33] we find

$$\begin{aligned}
\mathbb{E}_t \left[\left| \mathcal{M}_{t+u}^M - \mathcal{M}_t^M \right|^2 \right] &= \mathbb{E}_t \left[\left| \mathcal{M}_{t+u}^M \right|^2 - \left| \mathcal{M}_t^M \right|^2 \right] \\
&= \sum_{i=1}^M \mathbb{E}_t \left[\int_t^{t+u} \frac{1}{M} \left| f(X_s^i + c^i) - f(X_{s-}^i) \right|^2 dN_s^i \right] \\
&= \sum_{i=1}^M \mathbb{E}_t \left[\int_t^{t+u} \frac{1}{M} \left| f(X_s^i + c^i) - f(X_{s-}^i) \right|^2 \lambda_s^i ds \right] \\
&\le C_p \left\| \frac{\partial f}{\partial x} \right\|^2 \frac{1}{M} \sum_{i=1}^M \mathbb{E}_t \left[\int_t^{t+u} \lambda_s^i dt \right] \\
&\le C_p \frac{1}{2} \left\| \frac{\partial f}{\partial x} \right\|^2 \gamma^{\frac{1}{4}} \frac{1}{M} \sum_{i=1}^M \mathbb{E} \left[1 + \int_0^T (\lambda_s^i)^2 dt \right].
\end{aligned}$$

With the mean-value theorem and Assumption 4.2.1 we find

$$\begin{aligned}
\left| P_{t+u}^M - P_t^M \right| &= \left| \sum_{i=1}^M \int_t^{t+u} \left[\frac{1}{M} (f(X_s^i + c^i) - f(X_{s-}^i)) \right] \lambda_s^i ds \right| \\
&\le C_p \left\| \frac{\partial f}{\partial x} \right\| \frac{1}{M} \sum_{i=1}^M \int_t^{t+u} |\lambda_s^i| ds \\
&\le C_p \frac{1}{2} \left\| \frac{\partial f}{\partial x} \right\| \gamma^{\frac{1}{4}} \frac{1}{M} \sum_{i=1}^M \left(1 + \int_0^T (\lambda_s^i)^2 dt \right).
\end{aligned}$$

Then using Lemma 4.5.1 and Proposition 4.1.3 we can finish the proof. □

By Theorem 8.6 of Chapter 3 in [29], relative compactness of the sequence $\{\nu^M : M \ge 1\}$ in $D_S(\mathbb{R}_+)$ then follows directly from the above two lemmas. Then if uniqueness of the limit point ν_t holds, we can thus conclude that the sequence ν_t^M converges weakly to the limit point ν_t and we thus conclude that weak convergence holds. The proof for the uniqueness is similar to the proof of Lemma C.1 in [8].

Conclusion and Discussion

In this thesis we studied recent risk management problems using a variety of mathematical techniques. In particular we used stochastic processes and PDE theory to price European and Bermudan options in which the underlying follows a flexible state-dependent local Lévy model. Our approach allowed to efficiently compute the Greeks of the options for hedging purposes and to compute the prices both with and without valuation adjustments. Furthermore, using a weak convergence analysis we investigated the behavior of the systemic risk in a large interbank network in which the monetary reserve processes of the banks are connected through lending agreements as well as a self- and cross-excitement factor coming from common balance sheets and financial acceleration. This last term was modelled by means of a Hawkes process.

More specifically, in Chapter 2 and 3 we considered the underlying to follow a local Lévy model. This model extends other commonly used models such as the geometric Brownian motion, or the local volatility model, by including a jump process with a state-dependent measure and a local default intensity. Due to the state-dependency in the coefficients neither an explicit density nor characteristic function is available, so that pricing under this model is not trivial. We introduced a Taylor-based approximation to derive the approximate characteristic function, and used this in combination with a Fourier-based pricing method known as the COS method in order to compute both European and Bermudan options. Furthermore, we considered pricing Bermudan derivatives under the presence of XVA, consisting of CVA, DVA, MVA, FVA, and KVA. We derived the replicating portfolio with cashflows corresponding to the different rates for different types of lending, resulting in a non-linear PIDE. We proposed to solve the PIDE by rewriting it as BSDE and using the combination of the COS method and the expansion method for the characteristic function, this resulted in an efficient pricing method for both European and Bermudan derivatives *with* XVA. We presented an alternative for computing the CVA term in the case of unilateral collateralization (as is the case when the derivative is an option) without the use of BSDEs. This results in an even more efficient method due to the ability to use the FFT. Ideas for further research could be to include a way of incorporating wrong-way risk; in the current framework a simple form of wrong-way risk is included by allowing the dependence of the default intensity on the underlying. Nevertheless an interesting topic of research might be to develop a more rigorous framework to quantify the wrong-way risk in our model. Furthermore, the inclusion of stochastic interest rates could be an interesting further development. Lastly, the local Lévy model considered here was of a one-dimensional form. It would be useful to study the possibility of extending the results to multi-dimensional (be it stochastic volatility processes or correlated multi-asset portfolios). We remark that by e.g. [51] expressions for the density expansion in the case of multi-dimensional processes do exist, however combining these expressions with e.g. the COS method is not trivial. Nevertheless for both XVA portfolio computation, as stochastic volatility option pricing the above could be a very useful extension.

In Chapter 4 we studied the systemic risk in an interbank system. In particular we aimed to understand the effects of considering an additional self-exciting and clustering shock that impacts the monetary reserve or asset value of the nodes of the interbank system. The nodes are assumed to interact through the drift, and additionally are subjected to a Hawkes-distributed shock. In this way the jump activity varies over time resulting in jump clustering and the shocks propagate through the network in a contagious manner. This allows us to model both default propagation due to interbank loans as well as propagation due to linked balance sheets and financial acceleration. By a weak convergence analysis in which we studied the behavior of the system in the limiting setting of the number of banks going to infinity, we concluded that the clustering Hawkes jumps result in an additional and important source of default propagation in the network and should not be ignored. A potential extension might be to consider the Hawkes process shock size or intensity to be dependent on the monetary reserves, in this way the impact of a shock depends on the state of the system. In general, a more fragile state, in which firms are more susceptible to contagion, would then result in a larger impact. Furthermore, extending the work of [16] in which the authors consider a mean field game in a system without jumps, a game aspect could be introduced into our model. Each bank would be trying to optimize a particular borrowing rate and the effects of including the additional Poisson or Hawkes jump term on this optimal borrowing rate could be investigated. Concluding, the techniques and applications presented in this thesis are both novel and relevant for risk management applications after the financial crisis. The methods developed could be useful for practitioners as well provide a baseline for further research.

Bibliography

[1] Y. Aït-Sahalia, J. Cacho-Diaz, and R. Laeven, *Modeling financial contagion using mutually exciting jump processes*, Journal of Financial Economics, 117 (2015), pp. 585–606.

[2] A. Almendral and C. W. Oosterlee, *On american options under the variance gamma process*, Applied Mathematical Finance, 14 (2007), pp. 131–152.

[3] E. Bacry, I. Mastromatteo, and J.-F. Muzy, *Hawkes processes in finance*, Market Microstructure and Liquidity, 1 (2015).

[4] E. Bacry and J.-F. Muzy, *First-and second-order statistics characterization of Hawkes processes and non-parametric estimation*, IEEE Transactions on Information Theory, 62 (2016), pp. 2184–2202.

[5] V. Bally and A. Kohatsu-Higa, *A probabilistic interpretation of the parametrix method*, Ann. Appl. Probab., 25 (2015), pp. 3095–3138.

[6] C. Bender and J. Steiner, *Least-Squares Monte Carlo for Backward SDEs*, in Numerical methods in finance, Springer, 2012, pp. 257–289.

[7] P. Billingsley, *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons Inc., 1999.

[8] L. Bo and A. Capponi, *Systemic risk in interbanking networks*, SIAM Journal on Financial Mathematics, 6 (2015), pp. 386–424.

[9] A. Borovykh, A. Pascucci, and C. W. Oosterlee, *Pricing Bermudan options under local lévy models with default*, Journal of Mathematical Analysis and Applications, 450 (2017), pp. 929–953.

[10] P. Brémaud and L. Massoulié, *Stability of nonlinear Hawkes processes*, The Annals of Probability, (1996), pp. 1563–1588.

[11] C. Burgard and M. Kjaer, *Partial differential equation representations of derivatives with bilateral counterparty risk and funding costs*, Journal of Credit Risk, 7 (2011), pp. 75–93.

[12] N. Bush, B. M. Hambly, H. Haworth, L. Jin, and C. Reisinger, *Stochastic evolution equations in portfolio credit modelling*, SIAM Journal on Financial Mathematics, 2 (2011), pp. 627–664.

[13] A. Capponi, *Systemic risk, policies, and data needs*, in Optimization Challenges in Complex, Networked and Risky Systems, INFORMS, 2016, pp. 185–206.

[14] A. Capponi and M. Larsson, *Price contagion through balance sheet linkages*, Review of Asset Pricing Studies, 5 (2015), pp. 227–253.

[15] A. Capponi, S. Pagliarani, and T. Vargiolu, *Pricing vulnerable claims in a Lévy-driven model*, Finance Stoch., 18 (2014), pp. 755–789.

[16] R. Carmona, J.-P. Fouque, and L.-H. Sun, *Mean field games and systemic risk*, Communications in Mathematical Sciences, 13 (2015), pp. 911–933.

[17] P. Carr and V. Linetsky, *A jump to default extended CEV model: an application of Bessel processes*, Finance Stoch., 10 (2006), pp. 303–330.

[18] A. Castagna, *Yes, FVA is a cost for derivatives desks - A note on 'is FVA a cost for derivatives desks?' by Prof. Hull and Prof. White*, working paper, (2012).

[19] C. Chen, G. Iyengar, and C. C. Moallemi, *Asset-based contagion models for systemic risk*, Operations Research, (2014).

[20] R. Cont, A. Moussa, and E. Santos, *Network Structure and Systemic Risk in Banking Systems*, Cambridge University Press, 2013.

[21] R. Cont and P. Tankov, *Financial modelling with jump processes*, vol. 2, Chapman & Hall, 2004.

[22] T. J. Cox, A. Greven, and T. Shiga, *Finite and infinite systems of interacting diffusions*, Probability theory and related fields, 103 (1995), pp. 165–197.

[23] J. Cvitanić, J. Ma, and J. Zhang, *The law of large numbers for self-exciting correlated defaults*, Stochastic Processes and their Applications, 122 (2012), pp. 2781–2810.

[24] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes: Volume II: General theory and structure*, Springer Science & Business Media, 2007.

[25] C. de Graaf, Q. Feng, D. Kandhai, and C. Oosterlee, *Efficient computation of exposure profiles for counterparty credit risk*, International Journal of Theoretical and Applied Finance, 4 (2014).

[26] S. Delattre, N. Fournier, M. Hoffmann, et al., *Hawkes processes on large networks*, The Annals of Applied Probability, 26 (2016), pp. 216–261.

[27] Ł. Delong, *Backward stochastic differential equations with jumps and their actuarial and financial applications: BSDEs with jumps*, Springer Science & Business Media, 2013.

[28] E. Errais, K. Giesecke, and L. R. Goldberg, *Affine point processes and portfolio credit risk*, SIAM J. Financial Math., 1 (2010), pp. 642–655.

[29] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, John Wiley and Sons, 1986.

[30] F. Fang and C. W. Oosterlee, *A novel pricing method for European options based on Fourier-cosine series expansions*, SIAM J. Sci. Comput., 31 (2008/09), pp. 826–848.

[31] ——, *Pricing early-exercise and discrete Barrier options by Fourier-cosine series expansions*, Numer. Math., 114 (2009), pp. 27–62.

[32] Q. Feng and C. Oosterlee, *Calculation of exposure profiles and sensitivities of options under the heston and the Heston Hull-White models*, in Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science, Springer, 2017, pp. 265–301.

[33] K. Giesecke, K. Spiliopoulos, and R. Sowers, *Default clustering in large portfolios: Typical events*, The Annals of Applied Probability, (2013), pp. 348–385.

[34] K. Giesecke, K. Spiliopoulos, R. B. Sowers, and J. A. Sirignano, *Large portfolio asymptotics for loss from default*, Mathematical Finance, 25 (2015), pp. 77–114.

[35] P. Glasserman and H. P. Young, *How likely is contagion in financial networks?*, Journal of Banking, 50 (2015), pp. 383–399.

[36] A. Green, C. Kenyon, and C. Dennis, *KVA: Capital Valuation Adjustment*, Risk, 12 (2014).

[37] B. Hambly and N. Kolliopoulos, *Stochastic evolution equations for large portfolios of stochastic volatility models*, arXiv preprint arXiv:1701.05640, (2017).

[38] A. G. Hawkes, *Point spectra of some mutually exciting point processes*, J. R. Statist. Soc. B, 33 (1971), pp. 438–443.

[39] M. F. Hellwig, *Systemic risk in the financial sector: An analysis of the subprime-mortgage financial crisis*, De economist, 157 (2009), pp. 129–207.

[40] S. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, Rev. Financ. Stud., 6 (1993), pp. 327–343.

[41] J. Hull and A. White, *The FVA debate*, Risk, 7 (2012), pp. 83–85.

[42] ——, *XVAs: A gap between theory and practice*, Risk, (2016).

[43] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North Holland, 1981.

[44] A. Jacquier and M. Lorig, *The smile of certain Lévy-type models*, SIAM J. Financial Math., 4 (2013), pp. 804–830.

[45] S. Jain and C. W. Oosterlee, *The stochastic grid bundling method: Efficient pricing of Bermudan options and their Greeks*, Appl. Math. Comput., 269 (2015), pp. 412–432.

[46] C. Kipnis and C. Landim, *Scaling limits of interacting particle systems*, vol. 320, Springer Science & Business Media, 2013.

[47] A. Lesniewski and A. Richter, *Managing counterparty credit risk via BSDEs*, submitted, (2016).

[48] V. Linetsky, *Pricing equity derivatives subject to bankruptcy*, Math. Finance, 16 (2006), pp. 255–282.

[49] F. Longstaff and E. Schwartz, *Valuing American options by simulation: A simple least-squares approach*, Rev. Financ. Stud., 14 (2001), pp. 113–147.

[50] R. Lord, F. Fang, F. Bervoets, and C. W. Oosterlee, *A fast and accurate FFT-based method for pricing early-exercise options under Lévy processes*, SIAM J. Sci. Comput., 30 (2008), pp. 1678–1705.

[51] M. Lorig, S. Pagliarani, and A. Pascucci, *A taylor series approach to pricing and implied volatility for local–stochastic volatility models*, Journal of Risk, 17 (2014), pp. 3–19.

[52] ——, *Analytical expansions for parabolic equations*, SIAM J. Appl. Math., 75 (2015), pp. 468–491.

[53] ——, *A family of density expansions for Lévy-type processes*, Ann. Appl. Probab., 25 (2015), pp. 235–267.

[54] D. Madan and E. Seneta, *The variance gamma (VG) model for share market returns*, Journal of Business, 63 (1990), pp. 511–524.

[55] R. Merton, *Option pricing when underlying stock returns are discontinuous*, Journal of financial economics, 3 (1976), pp. 125–144.

[56] S. Nadtochiy and M. Shkolnikov, *Particle systems with singular interaction through hitting times: application in systemic risk modeling*, arXiv preprint arXiv:1705.00691, (2017).

[57] Y. Ogata, *On Lewis' simulation method for point processes*, IEEE Transactions on Information Theory, 27 (1981), pp. 23–31.

[58] S. Pagliarani, A. Pascucci, and C. Riga, *Adjoint expansions in local Lévy models*, SIAM J. Financial Math., 4 (2013), pp. 265–296.

[59] A. Pascucci, *PDE and martingale methods in option pricing*, vol. 2 of Bocconi & Springer Series, Springer, Milan; Bocconi University Press, Milan, 2011.

[60] V. Piterbarg, *Funding beyond discounting: Collateral agreements and derivatives pricing*, Risk, 2 (2010), pp. 97 –102.

[61] ——, *A non-linear PDE for XVA by forward Monte Carlo*, Risk, 10 (2015).

[62] S.-H. Poon, M. Rockinger, and J. Tawn, *Extreme value dependence in financial markets: Diagnostics, models, and financial implications*, The Review of Financial Studies, 17 (2003), pp. 581–610.

[63] M. J. Ruijter and C. W. Oosterlee, *A Fourier-cosine method for an efficient computation of solutions to BSDEs*, SIAM J. Sci. Comput., 37 (2015), pp. 859–889.

[64] J. Sirignano and K. Giesecke, *Risk analysis for large pools of loans*, Management Science, (2018).

[65] J. A. Sirignano, G. Tsoukalas, and K. Giesecke, *Large-scale loan portfolio selection*, Operations Research, 64 (2016), pp. 1239–1255.

[66] K. Spiliopoulos, J. A. Sirignano, and K. Giesecke, *Fluctuation analysis for the loss from default*, Stochastic Processes and their Applications, 124 (2014), pp. 2322–2362.

[67] L. von Sydow, L. J. Höök, E. Larsson, and et al., *BENCHOP—the BENCHmarking project in option pricing*, Int. J. Comput. Math., 92 (2015), pp. 2361–2379.

[68] B. Zhang and C. W. Oosterlee, *Fourier cosine expansions and put-call relations for Bermudan options*, in Numerical methods in finance, vol. 12 of Springer Proc. Math., Springer, Heidelberg, 2012, pp. 323–350.

[69] W. Zhao, Y. Li, and G. Zhang, *A generalized θ-scheme for solving backward stochastic differential equations.*, Discrete & Continuous Dynamical Systems-Series B, 17 (2012).