Université catholique de Louvain

Institute of Statistics, Biostatistics and Actuarial Sciences

Self-exciting and Fractional Processes in Finance

John-John Ketelbuters

Thesis submitted in partial fulfilment
of the requirements of the degree of
Docteur en sciences

Composition of the jury:
Prof. Donatien Hainaut (UCLouvain), Supervisor
Prof. Christian Hafner (UCLouvain), Secretary
Prof. Carlo Sgarra (Politectnico di Milano)
Prof. Tommi Sottinen (University of Vaasa)
Prof. Pierre Devolder (UCLouvain), President

Louvain-la-Neuve, November 2022

Acknowledgements

This PhD thesis is the result of five years of work, while I was a teaching assistant at the Institute of Statistics, Biostatistics and Actuarial Sciences (ISBA), at UCLouvain. During these five years, I benefited from the help and support of numerous people.

First of all, I would like to thank my PhD supervisor, Prof. Donatien Hainaut. I thank him for his guidance, his availability, his advices and his encouragements along the way. I also thank him for his trust and his efforts in finding a topic that would both motivate me and lead to interesting results. I thank the members of my "Comité d'Accompagnement", Prof. Pierre Devolder and Prof. Frédéric Vrins for interesting suggestions and discussions, as well as for helping me when I needed.

I am grateful to Prof. Pierre Devolder, Prof. Christian Hafner, Prof. Carlo Sgarra and Prof. Tommi Sottinen for accepting to be part of the jury and for the interesting discussions we had at this occasion. In particular, I thank Prof. Pierre Devolder and Prof. Christian Hafner for accepting to be respectively the chairman and the secretary of the jury. I am thankful to Prof. Carlo Sgarra for coming in person from Italy to participate to my private defense. The modifications suggested by the jury certainly helped to improve this PhD thesis.

I also thank the administrative staff of the ISBA: Nancy Guillaume, Marguerite-Marie Hanon, Sophie Malali, Nadja Peiffer and Tatiana Regout, for their kindness and for always being ready to help.

During these five years, it has been a pleasure to be the teaching assistant of Vincent Bremhorst, Céline Bugli, Prof. Anouar El Ghouch, Prof. Marie-Paule Kestemont, Nathalie Lefèvre, Prof. Johan Segers, Cédric Taverne and Robin Van Oirbeek. I would like to give a special thank to Marie-Paule Kestemont for trusting me for the remedial courses and the preparatory courses in statistics, as well as for her help and for our (sometimes long but always very pleasant) phone calls.

I would like to thank Prof. Michel Denuit. I remember that he was the first (in July 2016) to tell me that several PhD positions were going to be available at the time I finish my master degree, one year later. I also would like to thank Prof. Bernadette Govaerts, Prof. Catherine Legrand, Prof. Eugen Pircalabelu, Prof. Christian Ritter and Prof. Rainer Von Sachs for the conversations we had. I am especially grateful to Eugen Pircalabelu for sharing with me his own video equipment for my private PhD defense.

I was lucky to be surrounded by many distinguished colleagues ("granprof" or "professeur", as I took the bad habit to call some of them). I hope not to forget anyone: Hassana Al-Hassan, Hugues Annoye, Ndeye Arame Seck, Stefka Kirilova Asenova, Antoine Barbieri, Oussama Belhouari, Aurélie Bertrand, Vincent Bremhorst, Joris Chau, Benjamin Colling, Mickaël De Backer, Benjamin Deketelaere, Morine Delhelle, Lieven Desmet, Keivan Diakite, Hortense Doms, Jean-Loup Dupret, Oswaldo Gressani, Alain Guillet, Séverine Guisset, Vanessa Hanna, Fanny Hoogstoel, Alexandre Jacquemain, Charlotte Jamotton, Chikeola Ladekpo, Lise Léonard, Stéphane Lhaut, Nathalie Lucas, Rebecca Marion, Manon Martin, Sophie Mathieu, Cheikh Mbaye, Kassu Mehari Beyene, Gilles Mordant, Hélène Morsomme, Anas Mourahib, Catherine Rasse, Ensiyeh Nezakati Rezazadeh, Pauline Ngugnie Diffouo, Charles-Guy Njike Leunga, Emmanuel Niyigena, Florian Pechon, Vincent Plassier, Christophe Poncin, Antoine Soetewey, Raphaël Tursis, Nathan Uyttendaele, Linqi Wang, Fadoua Zeddouk and Aigerim Zhuman.

Finally, I thank my family and friends for unwavering support.

Contents

List of Figures

List of Tables

Chapter 0

Introduction

Financial mathematics, also known as quantitative finance or mathematical finance, consist of tools that allow to price financial products and assess the risks that economic agents are subject to. These tools are quantitative methodologies that help to evaluate and manage financial risks. An important type of financial risk is known as credit risk, or counterparty risk. Credit risk is the risk that arises from the possibility that an economic agent will not be able to meet its financial commitments, in which case we say that it defaults. The most frequently used approaches for credit risk are the structural models and the reduced models. Structural models aim at representing the balance sheet of a firm, as in Merton (1974). The balance sheet evolves randomly over time and the default occurs when the assets fall below the liabilities. In reduced models, initiated by Duffie and Singleton (1999), the default of a company corresponds to a jump of a stochastic point process. As it is common to work with the intensity[1] associated with the point process, reduced models are also known as intensity-based models.

Two other important problems in financial mathematics are to determine the appropriate price of a financial commitment and to devise a proper investment strategy that allows to meet this commitment. These last two problems are respectively referred to as pricing and hedging of financial derivatives. In the easiest case, the financial commitment can be fully replicated by existing financial assets. A market in which any financial commitment can be replicated is said to be a complete market. However markets are not complete in practice. Generally speaking, the investment (or hedging) strategy fails to exactly replicate the payoff associated to the financial commitment at hand. Facing the absence of a perfect strategy, one aims at determining a reasonably good strategy.

[1]The intensity can be seen as an instantaneous probability of jump. The higher the intensity, the more chances we have of observing a default.

In order to tackle those issues, it is usually assumed that the financial markets behave according to a mathematical model which is determined by a set of parameters. Once the parameters are determined with the help of market data, the practitioner is able to perform computations delivering sensible answers to the problems at hand. The assumption that the financial markets are properly described by the chosen mathematical model means that this model should be able to take all the possible risks into account. This is of course a strong assumption, to the point that it is always breached to some extent. As a famous aphorism says, "all models are wrong, but some are useful". Therefore, a more reasonable requirement for a model is the ability to incorporate, if not all, the most important risks that arise in a particular situation. For example, the financial crisis of 2008 and the collapse of Lehman Brothers has highlighted that a mathematical model that is able to measure the risk of contagion between defaults could be particularly useful. We can also cite the Eurozone sovereign debt crisis (Aït-Sahalia et al. (2014)). The risk of contagion refers to the fact that the default of an economic operator is likely to affect other economic operators that usually trade with him. As a consequence, the default of an agent could lead other agents to default shortly afterwards. The important economical consequences of the default of Lehman Brothers indicates the existence of a propagation phenomena that is channeled through the complex economic system. The example of the risk of contagion is not chosen at random, since it will be ubiquitous in this thesis. The mathematical tool used in this work to deal with the contagion mechanism is called a self-exciting point process. Such processes are also known as Hawkes processes, introduced in Hawkes (1971a) and Hawkes (1971b). In such a model, the occurrence of an event, i.e. the default of an economic operator, triggers an increase of the likelihood of occurrence of other subsequent events through an increase of the intensity of the Hawkes process. This exactly corresponds to the phenomenon of clustering of defaults described above.

The fundamental financial instrument that is concerned with credit risk is the Credit Default Swap (CDS). We give the definition of Brigo and Mercurio (2013). A CDS is a contract that ensures protection against the default of a specified entity. More precisely, three entities are involved in such a contract. The first two entities are respectively called the protection buyer and the protection seller. The third is called the reference credit. In a CDS contract, it is agreed that if the reference credit defaults during a pre-specified time period, the protection seller pays a fixed amount of money to the protection buyer, thereby protecting the latter against the potential default of the reference credit. Of course, the protection seller does not agree to this for free: the protection buyer pays a premium to the protection seller. CDS that protect against the defaults of large companies are now actively traded and thus quite liquid. The price of such an instrument in the market reflects the anticipations regarding the ability of the reference credit to face its financial obligations. The more likely the reference credit is to default, the higher the price of the CDS on the market. The price of the CDS is thus inherently bound to the general perception of the probability of default of the reference credit. As a consequence, it is common practice to deduce survival probabilities from CDS prices, a survival probability being the probability that the reference credit will not default during a certain period of time. The deduced survival probabilities, say market survival probabilities, can then be compared to the survival probabilities implied by mathematical models. One is then able to choose the set of parameters that minimizes some kind of distance between the survival probabilities implied by the model and the market survival probabilities. This procedure is called the calibration of the model. However, given a particular curve of market survival probabilities and a mathematical model, it is not guaranteed that the parameters obtained from the calibration will give satisfactory results. On top of being able to take into account the risk of contagion described above, another sensible requirement for a candidate model is thus to be able to provide a good match with respect to the market survival probabilities.

The Covid-19 crisis gave rise to examples of survival probabilities that are not easily matched by the most common mathematical models used for credit risk. Some companies, such as airlines, found themselves in a situation in which their survival was threatened in the short run. However, provided that these companies survive the crisis, they will have a good viability afterwards. Such a situation leads to survival probability curves that exhibit a strong convexity: they decrease sharply for short maturities but become relatively flat for longer maturities. The second main topic of this thesis is to propose credit models that can reproduce this kind of shape for survival probability curves. To achieve this, we use the technique of time-change, which is well-known in finance. It consists in allowing time to pass in a random way. An intuitive way of thinking about time-changes is to represent a period of agitated highly volatile markets, i.e. with a lot of transactions and movements in the prices, as a moment during which time passes faster than usual. Conversely, a period with few transactions and few price changes can be seen as a period during which time flies rather slowly. Introducing such a distortion of time in a mathematical model can allow to obtain a better fit with respect to real market data.

The last topic addressed in this thesis is the modeling of illiquid stocks. These are stocks of smaller companies that are not heavily traded on the market, making them illiquid. Such a stock is characterized by the presence of periods of time during which its price remains constant. From the intuitive description of the technique of time-change given in the previous paragraph, it seems that this technique is suitable for modeling illiquid stocks. One would need a time-change that freezes time randomly, thereby representing the absence of transactions. It turns out that the time-change that allows to do this is precisely the same as the one that allows to obtain highly convex shapes of survival probability curves in a credit model, as described in the previous paragraph. The time-changed model allows us to study the pricing and the hedging of options whose underlying is an illiquid stock. The content of the thesis is discussed in more details in the following outline.

Outline of the Thesis

The four main chapters of this dissertation are self-contained research articles that can be read independently from each other. The outline of the dissertation is as follows.

Chapter 1: Time-Consistent Evaluation of Credit Risk with Contagion

This chapter aims at dealing with the inherent incompleteness of insurance markets in a context of credit insurance. The considered problem is the valuation of a credit insurance portfolio, i.e. an homogeneous collection of contracts in which the insurer makes the commitment to pay a fixed amount of money in case a given economic agent defaults within a pre-specified period of time. It is assumed in this chapter that the economic agent whose default is covered by the insurer is a small or medium company that is not listed on a stock exchange. As a consequence, financial instruments such as CDS are not available on the market. It is thus not possible to rely on the existence of a unique risk-neutral measure, as the market is not complete in this case.

In this chapter, we propose to use an actuarial evaluation under the real probability measure. An actuarial evaluation typically consists of an expected value to which a safety margin is added. The expected value component covers the average payments to be made by the insurer, whereas the safety margin allows to cover the other operating costs of the insurer, and hopefully to generate a profit for the shareholders. The risk margin can for example be based on the variance, which leads to an evaluation called the variance principle, or based on the standard-deviation, which leads to the standard-deviation principle. The inclusion of a risk margin unfortunately prevents the actuarial evaluation from satisfying the desirable property of time-consistency, according to which a risk that will be more expensive then another one at a future date should already be more expensive today. Following the approach of Pelsser and Ghalehjooghi (2016), starting from evaluations that are not time-consistent, we derive partial integro-differential equations (PIDE) that describe their time-consistent counterparts. We show that this can be achieved by using classical tools from stochastic calculus such as Ito's lemma and infinitesimal generators. Afterwards, we propose a numerical method based on finite differences to solve the obtained PIDE's.

As mentioned above in the general introduction, one of the central themes of this thesis is the risk of contagion, which is modeled with self-exciting stochastic processes. The credit model we use in this chapter is thus based on a self-exciting Hawkes process. It is a reduced model in which the intensity increases at each occurrence of a jump, thereby increasing the probability to observe subsequent jumps. Our numerical results allow us to bring to light the impact of using a time-consistent evaluation, as well as interactions between the risk of contagion and the use of time-consistent evaluations. More precisely, our conclusion is twofold. On the one hand, we show that time-consistent evaluations tend to give higher prices, compared to time-inconsistent evaluations. On the other hand, our results show that the time-consistency of evaluations allows to better take into account the risk of contagion in credit insurance, if such a risk exists. Finally, in order to allow for a practical use of our results, we propose a method to calibrate our model by likelihood maximization.

Chapter 2: CDS Pricing with Fractional Hawkes Processes

In this chapter, we propose a fractional (or subdiffusive) self-exciting reduced model for the risk of corporate default. A fractional model is obtained by performing a particular type of time-change, namely the inverse of an α-stable subordinator[2]. In a context of credit risk, we show that such a time-change gives the possibility of obtaining more convex shapes for the survival probability curve. That is, the obtained curves tend to decrease sharply for lower maturities and become almost flat for longer maturities. This feature corresponds to companies that are perceived as having a significant risk of bankruptcy in the short term, but a good long term financial viability provided that they survive the difficult period. The 2020 Covid-19 crisis gave concrete examples. We can cite airlines, that were well established strong businesses, but were endangered by the suspension of the air traffic.

We study the properties of an intensity-based model that is time-changed by the inverse of an α-stable subordinator. On top of obtaining more convex shapes for the survival probability curves, performing such a time-change allows to incorporate two particular features. Firstly, it introduces random periods of time where the survival probability is frozen, thereby modeling periods of time where the viability of the company is not threatened. Secondly, the time-change implies possible sharp drops in the survival probability. This feature corresponds to the occurrence of one-time events that threaten the creditworthiness of the company.

The chapter starts by a description of the model. It is first described and studied without performing the time-change. We show that the joint probability density function (PDF) and Laplace transform of the intensity and its integral are solutions of Fokker-Planck equations. This is a PDE that is also referred to as forward Kolmogorov equation. Next, after a brief introduction to fractional calculus, we show that the time-changed model leads to a similar Fokker-Planck equation. The only difference is that the partial derivative with respect to the time is replaced by a fractional Caputo derivative[3], leading to a fractional partial differential equation (FPDE). After a discussion regarding approximations of Caputo fractional derivatives, we describe a simple and fast numerical method to solve the Fokker-Planck

[2] A subordinator is an increasing Lévy process.
[3] Wich explains the name *fractional* model.

equation of the Laplace transform, thereby enabling the computation of the survival probabilities implied by the time-changed model. Finally, we use our results to calibrate the model to real market data and show that it leads to an improvement of the fit.

Chapter 3: A Recursive Method for Computing Moments of Hawkes Intensities: Application to the Potential Approach of Credit Risk

Chapter 3 explores an alternative to the common structural models and reduced models in credit risk. We propose to use another approach called the potential approach. This approach was introduced in Constantinides (1992). The name "potential approach" was coined by Rogers (1997). It was used in these papers for modeling the term structure of interest rates. As in the previous chapter, the driving process is both fractional and self-excited. Instead of modeling the assets and liabilities of the company or modeling the instantaneous probability of default, i.e. the intensity, the potential approach models the survival probability curve directly. More specifically, the probability of a firm survival is the ratio of the expectation of a supermartingale on its initial value. The supermartingale is ruled by a time-changed Hawkes jump-diffusion process. The self-excitation mechanism allows for sudden and repeated increases of default probabilities. The time-change is the inverse of an α-stable subordinator, as in Chapter 2.

With respect to the reduced model used in Chapter 2, the benefit of the potential approach is to be easier to handle, both mathematically and computationally. In Chapter 2, we had to rely on fractional calculus to derive a FPDE. This is rather complicated and requires a substantial amount of computations. In addition, solving numerically a FPDE is computationally intensive. It requires to approximate numerically Caputo fractional derivatives, for which it is necessary, at each computation step, to use all the values obtained at the previous steps. To avoid these shortcomings, we introduce a new method to compute the moments of a self-exciting Hawkes process intensity. We prove that all the moments of this intensity share the same form, i.e. a sum of exponential functions added to a constant, and that they can be computed recursively. Then, using a result from Piryatinska et al. (2004) that links the Laplace transforms of inverse α-stable processes with the Mittag-Leffler function, we show that this recursive method extends perfectly to the fractional time-changed setting. The moments of the time-changed intensity are then a sum of Mittag-Leffler functions added to a constant. Being able to compute the moments of the intensity, we can approximate its Laplace transform with the help of a Taylor-Maclaurin expansion. We show that this can be used to approximate the survival probabilities in the potential approach. Afterwards, we prove the convergence of our approximation and provide upper bounds on the numerical error.

Finally we propose some numerical experiments with the same data as in Chapter 2. These experiments indicate that this approach gives results that are as satisfactory as in Chapter 2, while being much more tractable both mathematically and numerically.

Chapter 4: Option Pricing and Hedging in Illiquid Markets in Presence of Jump Clustering

This last chapter differs from the previous ones as it is not concerned with credit risk. Instead, we focus here on the modeling of a particular type of stock prices. However, the tools we use are the same, since we will rely on self-exciting jump processes and time-change by the inverse of an α-stable subordinator. Our goal is to capture two important features that small capitalization stocks tend to exhibit. The first is the presence of motionless periods in their prices. It is a consequence of a lack of liquidity, since these stocks are not heavily traded. The second is the occurrence of clustered sudden moves in the price at the times the stock is traded.

A natural model to represent such assets is thus a model that exhibits both self-excited jumps and motionless periods. To achieve this, we propose to time-change a self-exciting Hawkes jump-diffusion process by the inverse of an α-stable subordinator. In the following, the α-stable subordinator is denoted by $(U_t)_{t\geqslant 0}$ and its inverse is denoted by $(S_t)_{t\geqslant 0}$. Our goal is to deal with the pricing and hedging of call options whose underlying is a small capitalization stock. To achieve this, we have to work in a dynamic setting, as the hedging strategy requires to adjust our position through time. A first issue arises: the inverse $(S_t)_{t\geqslant 0}$ of an α-stable subordinator does not satisfy the Markov property. Moreover, we are not able to compute the distribution of the inverse α-stable process conditional on its past values. A solution to this problem is proposed in Hainaut (2021) and consists in adding some information to the time-change. More precisely, instead of working with the univariate process $(S_t)_{t\geqslant 0}$, we work with the bivariate process $(S_t, U_{S_t})_{t\geqslant 0}$. Hainaut (2021) shows that despite the conditional distribution of S_t given S_s, $s < t$ is unknown, the conditional distribution of S_t given (S_s, U_{S_s}) can be obtained. The first contribution of this chapter is to show that the bivariate process $(S_t, U_{S_t})_{t\geqslant 0}$ satisfy the Markov property. This proves that the approach proposed in Hainaut (2021) does not suffer from any loss of information.

A second issue is the pricing of vanilla options in a fractional self-exciting model. We propose to compute call option prices by using the classical Fast-Fourier Transform (FFT) method. To this end, we derive a FPDE that characterizes the Fourier transform of the log-asset price and present a numerical method based on finite differences to solve it. This approach is similar to the one we used in Chapter 2 to obtain a Laplace transform that was described by a FPDE. However, the oscillatory character of complex exponentials leads to a FPDE that is trickier to solve numerically. To overcome this, we study finite differences approximations that have a higher order of convergence. The finite differences we use are inspired by Alikhanov (2015), who introduced a higher order approximation of the Caputo fractional derivative and used it to solve FPDE's. Once the Fourier transform is obtained, call prices can be computed with a FFT algorithm.

Thirdly, we describe a whole class of risk-neutral measures. Fourthly, the hedging of options is discussed and an optimal quadratic hedging strategy is derived for call options. The hedging strategy is obtained by applying Ito's lemma on the call option price process. Finally the last section of this chapter proposes some numerical experiments.

Published Material

The first chapter of this thesis is based on the article Ketelbuters and Hainaut (2022b) that is published in the Journal of Computational and Applied Mathematics. The second chapter is based on Ketelbuters and Hainaut (2022a) and is published in the European Journal of Operational Research.

Chapter 1

Time-Consistent Evaluation of Credit Risk with Contagion

Introduction

Insurance markets are incomplete by nature since most of actuarial risks are not hedgeable with financial instruments. Choosing an appropriate valuation method is therefore a challenging task. This problem was initially tackled in a static way with the concepts of evaluations or risk measures. We refer the reader to e.g. Artzner et al. (1999) for an introduction about coherent risk measures and Kaas et al. (2008) for examples. Unlike insurance pricing, financial pricing is treated in a dynamic way, focused on hedging and replicating portfolios, see for example Black and Scholes (1973). Delbaen and Schachermayer (1994) show that this intrinsically dynamic approach can be applied with various securities, as long as the asset prices are bounded semimartingales. Dynamic evaluation methods extend this dynamic nature of financial pricing to insurance pricing. This type of evaluations, as well as the concept of time-consistency (or dynamic-consistency), was introduced by Riedel (2003) in discrete time. Rosazza Gianin (2006) uses g-expectations to extend the concept of time-consistency to continuous time. Pelsser and Stadje (2014) introduce the concept of two step market evaluations, that allows to apply time-consistent evaluations on products that incorporate both a financial risk and an actuarial risk. Pelsser and Ghalehjooghi (2016) use the definition of time-consistency in order to turn time-inconsistent classical evaluation principles into time-consistent ones. They derive partial differential equations (PDE) that characterize time-consistent evaluations. This is the approach we use in this chapter.

Another possible approach would be to use g-expectations. A g-expectation is a generalization of the classical expectation operator and is defined as the solution of a backward stochastic differential equation. According to Rosazza Gianin (2006) (Proposition 19), g-expectations lead to time-consistent evaluations, so they are a substitute for the approach of Pelsser and Ghalehjooghi (2016).

Credit risk for transactions involving large quoted corporations can be hedged with derivatives like credit default swaps. Nevertheless financial markets do not propose similar covers for small or medium enterprises. For this category of companies, the only solution to limit credit exposure is provided by trade credit insurances. Insurers proposing these contracts rely on diversification to hedge themselves but are exposed to contagion between defaults. This motivates us to propose a time-consistent valuation framework managing both default and contagion risks.

As explained in the introduction to this thesis, the two main paradigms for credit risk valuation are the structural and reduced approaches. In the first one, the balance sheet of the firm value is explicitly modeled and the default occurs when the asset value falls below liabilities. For details, we refer the reader e.g. to Merton (1974), Black and Cox (1976), Geske (1977) and Longstaff and Schwartz (1995). The second approach is called the reduced-form or intensity-based credit model. The default is in this case triggered by the first jump of a point process. The reduced-form was initiated by the work of Duffie and Singleton (1999). Textbooks concerning this second kind of models are Bielecki and Rutkowski (2014), Duffie and Singleton (2003), Lando (2004), Jeanblanc et al. (2009) and Schoutens and Cariboni (2009). In this chapter, we opt for a reduced-form approach.

A model for the contagion between defaults of companies consists in using self-exciting point processes, one of the main tools we use in this thesis. In this approach, any jump of this process triggers an increase of the instantaneous expected number of jumps. Hawkes (1971a), Hawkes (1971b) and Hawkes and Oakes (1974) were among the first to introduce such point processes. Since then, Hawkes processes have been used to model earthquake occurrences, see e.g. Musmeci and Vere-Jones (1992) or Ogata (1998), or terrorist activity, see Porter and White (2012). They are also used for modeling financial transactions, see e.g. Bauwens and Hautsch (2006), Bacry et al. (2015), Hainaut (2017), Moraux and Hainaut (2019), Hainaut and Goutte (2019), or for credit risk in Ma and Xu (2016). For a survey of properties and other possible applications of Hawkes processes, the reader is referred to Reinhart (2018).

This chapter differs from Chapters 2 and 3. Since we are interested here in a credit portfolio, we can observe the defaults of several companies. In the subsequent chapters, we will be interested in the credit risk of a single company. Moreover, in this first chapter, the default of a company entails an increase of the intensity of default if and only if the company belongs to the credit portfolio. In that aspect, the jumps we observe in the processes correspond to defaults that are completely endogenous with respect to the credit portfolio. This is different in Chapters 2 and 3, where the intensity process is exogenous, i.e. it is only affected by the defaults of other companies.

The contributions of this chapter are organized as follows. Firstly, we introduce a new model based on self-exciting processes in which the default intensity of each company is an affine function of a common market intensity. We provide closed-form expressions for the moments of this intensity process. Secondly we find semi-closed form expressions for the Laplace transform of the intensity process. We use this Laplace transform to compute probabilities of default. Thirdly, we extend the approach of Pelsser and Ghalehjooghi (2016) to our setting in two directions. On the one hand, we evaluate in a time-consistent manner credit insurance portfolio. On the other hand, we take into account the risk of contagion. We obtain partial integro-differential equations (PIDE) for the time-consistent values and solve them numerically. Fourthly, we provide a numerical method for solving the PIDE. This method solves PIDE's for which we miss some of the boundary conditions. With the help of this method, we present a numerical illustration and analyze the impacts of contagion risk and time-consistency on the value of a credit insurance product. Finally, we propose a method to calibrate our model.

1.1 The Model

In the intensity approach to credit risk, defaults of companies are modeled by jumps of point processes. A jump of a point process corresponds to the default of a company. In this chapter, we are interested in the defaults observed in a credit insurance portfolio. The defaults correspond to the jumps of a point process $(P_t)_{t\geqslant 0}$. The losses induced by the defaults are modeled by the sizes of the jumps of $(P_t)_{t\geqslant 0}$. We start by describing precisely how the process $(P_t)_{t\geqslant 0}$ is built.

Let $(N_t)_{t\geqslant 0}$ be a point process on a filtered probability space $(\Omega, \mathbb{F}, P)$ where $\mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0}$ is the natural filtration of all the processes to be described in this section. We denote the intensity of $(N_t)_{t\geqslant 0}$ by $(\lambda_t)_{t\geqslant 0}$, another stochastic process. The intensity represents the instantaneous probability to observe a jump in the process $(N_t)_{t\geqslant 0}$. Let $\xi_1, \xi_2, \ldots$ be independent identically distributed positive square integrable random variables distributed according to the probability measure ν. We will often drop the subscript and write simply ξ for a ν-distributed random variable. We assume that the Laplace transform of ν exists and we denote it by $\tilde{\nu}$. That is, the integral $\tilde{\nu}(z) := \int_{\mathbb{R}^+} e^{-zx}\nu(\mathrm{d}x)$ exists at least for all z in a neighborhood of 0. We also assume that the values of the Laplace transform $\tilde{\nu}$, as well as the probabilities $P(a \leqslant \xi < b) = \nu([a, b))$ for any real numbers $a \leqslant b$, can be be obtained, at least numerically. The same assumption holds for the first and second moments of ν. Define the process $(P_t)_{t\geqslant 0}$ as

$$P_t = \sum_{k=1}^{N_t} \xi_k.$$

It is a pure positive jump process whose jump intensity is the stochastic process $(\lambda_t)_{t\geqslant 0}$. The dynamics of this intensity process is assumed to satisfy the stochastic differential equation (SDE)

$$\mathrm{d}\lambda_t = \kappa(\theta - \lambda_{t-})\mathrm{d}t + \eta \mathrm{d}P_t + \sqrt{\lambda_{t-}}\sigma \mathrm{d}W_t \tag{1.1}$$

where $(W_t)_{t\geqslant 0}$ is a standard Brownian motion, $\kappa, \theta, \eta > 0$, $\sigma \geqslant 0$ and $\lambda_0 \in [\theta, +\infty)$. As a consequence of SDE (1.1), the intensity depends upon the history of the process $(P_t)_{t\geqslant 0}$. This allows the process $(P_t)_{t\geqslant 0}$ to exhibit a self-exciting behaviour: a jump of the process $(P_t)_{t\geqslant 0}$ induces a jump in the intensity $(\lambda_t)_{t\geqslant 0}$, which in turn implies an increase of the probability to observe another jump in $(P_t)_{t\geqslant 0}$. The presence of $\sqrt{\lambda_{t-}}$ in the term $\sqrt{\lambda_{t-}}\sigma \mathrm{d}W_t$ at SDE (1.1) ensures that $\lambda_t \geqslant 0$ for all $t \geqslant 0$, almost surely. As early as λ_t gets close to 0, the noise term $\sqrt{\lambda_{t-}}\sigma \mathrm{d}W_t$ gets close to zero and the drift drives the value of the process upward. As a consequence, λ_t cannot fall below 0[1]. The parameter θ represents a level towards which the intensity tends to revert. The parameter κ in Equation (1.1) represents the speed of the reversion towards θ and σ is a volatility parameter.

Since the jumps of $(P_t)_{t\geqslant 0}$ correspond to defaults, our model introduces a phenomenon of contagion between defaults: a default of a company in the portfolio triggers an increase of the intensity of defaults, thereby increasing the chances to observe other defaults in the portfolio. The self-exciting behavior of $(\lambda_t, P_t)_{t\geqslant 0}$ thus aims to model the clustering of defaults observed in practice, see e.g. Ait-Sahalia et al. (2015).

It is important to note that the SDE (1.1) implies that the impact of past defaults dissipates at an exponential speed through time, with a rate κ. This is more visible at Equation (1.3), where $(\lambda_t)_{t\geqslant 0}$ is expressed in integral form. We observe at the last term of this equation that a default of cost $\mathrm{d}P_s$ at time $s \geqslant 0$ results in the term $\eta e^{-\kappa(t-s)}\mathrm{d}P_s$ at time $t \geqslant s$ so that defaults are forgotten at an exponential speed. This property is actually crucial in the use of stochastic calculus, and therefore in the mathematical developments throughout this thesis.

We close this section with a discussion about what differentiates our model from the one of Pelsser and Ghalehjooghi (2016). The main difference is the fact that the jump process used in Pelsser and Ghalehjooghi (2016) is a constant intensity compound Poisson process that is completely exogenous. More precisely, the jump

[1]The additional condition $2\kappa\theta \geqslant \sigma^2$ ensures that $\lambda_t > 0$ almost surely, for all t. See e.g. Section 6.3.1 in Jeanblanc et al. (2009). This is known as the Feller condition

diffusion process is

$$\mathrm{d}y_t = a(t, y_{t-})\mathrm{d}t + b(t, y_{t-})\mathrm{d}W_t + \mathrm{d}J_t$$

with $J_t = \sum_{i=1}^{N_t} \xi_i$ and $(N_t)_{t\geqslant 0}$ a Poisson process with constant intensity $\lambda > 0$. The process $(J_t)_{t\geqslant 0}$ is thus a compound Poisson process with stochastic jumps $\xi_1, \xi_2, \ldots$. This setting could be adapted to the modeling of a credit portfolio by taking $(y_t)_{t\geqslant 0}$ as the intensity of a Cox process $(D_t)_{t\geqslant 0}$ (i.e. a doubly stochastic Poisson process, it is a Poisson process with a stochastic intensity) whose jumps represent the defaults in the credit portfolio. In such a setting, the jumps of $(N_t)_{t\geqslant 0}$ represent defaults that are exogenous to the credit portfolio, whereas the jumps of $(D_t)_{t\geqslant 0}$ represent defaults that occur inside the credit portfolio (endogenous defaults). The difference with the model we propose in this chapter is the fact that the jumps of $(D_t)_{t\geqslant 0}$ would have no impact on its stochastic intensity $(y_t)_{t\geqslant 0}$, so that no phenomenon of contagion can exist inside the credit portfolio. In other words, in the jump model of Pelsser and Ghalehjooghi (2016), the occurrence of defaults in the credit portfolio can only be affected by exogenous defaults, but not by endogenous ones.

1.2 Expectation and Variance of the Intensity Process

This section is concerned with the computation of the expectation and variance of the intensity process described in Equation (1.1). These computations will allow us to derive non-explosion conditions for our model, that is, $\lim_{t\to+\infty}(\mathbb{E}[\lambda_t] \vee \mathbb{V}\mathrm{ar}(\lambda_t)) < +\infty$.

Before moving to the expectation and variance of the intensity, note that SDE (1.1) for $(\lambda_t)_{t\geqslant 0}$ can be written in integral form as

$$\lambda_t = \lambda_0 + \kappa \int_0^t (\theta - \lambda_{s-})\mathrm{d}s + \eta \int_0^t \mathrm{d}P_s + \int_0^t \sqrt{\lambda_{s-}}\sigma \mathrm{d}W_s \tag{1.2}$$

or even as

$$\lambda_t = \theta + e^{-\kappa t}(\lambda_0 - \theta) + \sigma \int_0^t e^{-\kappa(t-s)}\sqrt{\lambda_{s-}}\mathrm{d}W_s + \eta \int_0^t e^{-\kappa(t-s)}\mathrm{d}P_s. \tag{1.3}$$

To see that Equation (1.3) holds, define the process $(Z_t)_{t\geqslant 0}$ as $Z_t = \lambda_t e^{\kappa t}$. By Ito's lemma, we have

$$\mathrm{d}Z_t = e^{\kappa t}\left[\kappa(\theta - \lambda_{t-})\mathrm{d}t + \sigma\sqrt{\lambda_{t-}}\mathrm{d}W_t\right] + \lambda_{t-}\kappa e^{\kappa t}\mathrm{d}t + e^{\kappa t}\eta \mathrm{d}P_t.$$

An integration gives

$$\lambda_t e^{\kappa t} = \lambda_0 + \theta(e^{\kappa t} - 1) + \int_0^t e^{\kappa s}\sqrt{\lambda_{s-}}\sigma \mathrm{d}W_s + \eta \int_0^t e^{\kappa s}\mathrm{d}P_s.$$

It remains to multiply by $e^{-\kappa t}$ to obtain Equation (1.3).

Proposition 1.1. *Assume that $\kappa > \eta\mathbb{E}[\xi]$. For any $t \geqslant 0$, we have*

$$\mathbb{E}[\lambda_t] = e^{a_3 t}\lambda_0 + \frac{\kappa\theta}{a_3}(e^{a_3 t} - 1) \tag{1.4}$$

and

$$\mathbb{V}\mathrm{ar}(\lambda_t) = \frac{(e^{a_3 t} - 1)(2a_1\lambda_0 e^{a_3 t} + a_2(e^{a_3 t} - 1))}{2a_3} \tag{1.5}$$

where $a_1 = \sigma^2 + 2\eta^2\mathbb{V}\mathrm{ar}(\xi)$, $a_2 = \frac{\kappa\theta\left(\sigma^2+2\eta^2\mathbb{V}\mathrm{ar}(\xi)\right)}{\eta\mathbb{E}[\xi]-\kappa}$ and $a_3 = \eta\mathbb{E}[\xi] - \kappa$.

Proof. By taking the expectation on Equation (1.2) and differentiating with respect to t, we obtain the ordinary differential equation (ODE)

$$\frac{\mathrm{d}\mathbb{E}[\lambda_t]}{\mathrm{d}t} = \kappa(\theta - \mathbb{E}[\lambda_t]) + \eta\mathbb{E}[\xi]\mathbb{E}[\lambda_t]$$

whose solution is well given by Equation (1.4). For the variance, using Ito's lemma yields

$$\frac{\mathrm{d}\mathbb{E}[\lambda_t^2]}{\mathrm{d}t} = \mathbb{E}[\lambda_t]\left(2\kappa\theta + \sigma^2 + 2\eta^2\mathbb{V}\mathrm{ar}(\xi)\right) + 2\mathbb{E}[\lambda_t^2]\left(\eta\mathbb{E}[\xi] - \kappa\right). \tag{1.6}$$

By the chain rule, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbb{E}[\lambda_t])^2 = 2\kappa\theta\mathbb{E}[\lambda_t] + 2(\mathbb{E}[\lambda_t])^2\left(-\kappa + \eta\mathbb{E}[\xi]\right). \tag{1.7}$$

Putting together Equations (1.6) and (1.7), we obtain $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{V}\mathrm{ar}(\lambda_t)$. Then, replacing $\mathbb{E}[\lambda_t]$ by its expression at Equation (1.4) and performing some simplifications leads to the following ODE

$$\frac{\mathrm{d}\mathbb{V}\mathrm{ar}(\lambda_t)}{\mathrm{d}t} = e^{a_3 t}(a_1\lambda_0 + a_2) - a_2 + 2a_3\mathbb{V}\mathrm{ar}(\lambda_t)$$

whose solution is given by Equation (1.5). $\square$

From this proposition, we infer that the non-explosion condition $\lim_{t\to+\infty}(\mathbb{E}[\lambda_t] \vee \mathbb{V}\mathrm{ar}(\lambda_t)) < +\infty$ is equivalent to $a_3 < 0$, that is, $\eta\mathbb{E}[\xi] < \kappa$. This condition means that the speed of the mean reversion κ must be sufficiently large.

In the next section, we derive semi-closed form expressions for the Laplace transform of the intensity.

1.3 Laplace Transform

Let r be a real number that represents the constant interest rate and define the process $(P_t^r)_{t\geqslant 0}$ as the discounted version of $(P_t)_{t\geqslant 0}$, that is,

$$P_t^r := \int_0^t e^{-rs}\mathrm{d}P_s.$$

From this definition, $\mathbb{E}[P_t^r]$ represents the expected discounted cost of the credit portfolio up to time t. Moreover, we will denote by $(\Lambda_t)_{t\geqslant 0}$ the integral of $(\lambda_t)_{t\geqslant 0}$, i.e. $\Lambda_t := \int_0^t \lambda_u \mathrm{d}u$. Note that $(\Lambda_t)_{t\geqslant 0}$ is the compensator of $(N_t)_{t\geqslant 0}$. The next proposition will allow us to compute the Laplace transform of this process numerically, and thus to obtain its moments.

Proposition 1.2. *There exist functions A, B such that for all $0 \leqslant t \leqslant T$,*

$$\mathbb{E}[\exp\{-v_1(N_T-N_t)-v_2(P_T^r-P_t^r)-v_3(\Lambda_T-\Lambda_t)\}|\mathcal{F}_t] = \exp\{A(v,t,T)+\lambda_{t-}B(v,t,T)\}$$

where $v = (v_1, v_2, v_3)^\top$ and A, B satisfy the system of ODE's

$$\frac{\partial A}{\partial t}(v,t,T) = -\kappa\theta B(v,t,T)$$

$$\frac{\partial B}{\partial t}(v,t,T) = \kappa B(v,t,T) - \frac{\sigma^2}{2}B^2(v,t,T) + v_3 - \left(e^{-v_1}\tilde{\nu}(-\eta B(v,t,T) - ue^{-rt}) - 1\right).$$

Moreover A and B satisfy the terminal condition $A(v,T,T) = B(v,T,T) = 0$.

Proof. It suffices to show that there exist functions A, B such that the stochastic process

$$(M_t)_{t\in[0,T]} := \left(\exp\left\{A(v,t,T) + \lambda_{t-}B(v,t,T) - v_1 N_t - v_2 P_t^r - v_3\Lambda_t\right\}\right)_{t\in[0,T]}$$

is a martingale. Noting that $M_t := h(t, \lambda_t, N_t, P_t^r, \Lambda_t)$ where h is defined as

$$h(t, y, x_1, x_2, x_3) = \exp\left\{A(v,t,T) + yB(v,t,T) - v_1x_1 - v_2x_2 - v_3x_3\right\}$$

and applying Ito's Lemma allows to obtain

$$
\begin{aligned}
\mathrm{d}M_t = M_t\Bigg[&\left(\frac{\partial A}{\partial t}(v,t,T) + \lambda_{t-}\frac{\partial B}{\partial t}(v,t,T)\right)\mathrm{d}t \\
&+ B(v,t,T)\left(\kappa(\theta-\lambda_{t-})\mathrm{d}t + \sigma\sqrt{\lambda_{t-}}\mathrm{d}W_t\right) + \frac{\sigma^2}{2}B^2(v,t,T)\lambda_{t-}\mathrm{d}t \\
&- v_3\lambda_{t-}\mathrm{d}t + \left(\exp\{-v_1+\eta B(v,t,T)\xi - v_2e^{-rt}\xi\}-1\right)\mathrm{d}N_t\Bigg] \\
&= M_t\Bigg[\left(\frac{\partial A}{\partial t}(v,t,T) + \kappa\theta B(v,t,T)\right) \\
&+ \lambda_{t-}\left(\begin{aligned}&\frac{\partial B}{\partial t}(v,t,T) - \kappa B(v,t,T) - v_3 + \frac{\sigma^2}{2}B^2(v,t,T) \\ &\qquad + \exp\{\xi(\eta B(v,t,T) - v_2e^{-rt}) - v_1\} - 1\end{aligned}\right)\Bigg]\mathrm{d}t \\
&+ M_t\left[\exp\{\xi(\eta B(v,t,T)-v_2e^{-rt})-v_1\}-1\right](\mathrm{d}N_t - \lambda_t\mathrm{d}t) \\
&+ M_t\left[\sigma B(v,t,T)\sqrt{\lambda_{t-}}\right]\mathrm{d}W_t,
\end{aligned}
$$

where ξ is a random variable distributed according to ν. Since both $(W_t)_{t\geqslant 0}$ and $(N_t - \int_0^t \lambda_s \mathrm{d}s)_{t\geqslant 0}$ are martingales, a sufficient condition for $(M_t)_{t\geqslant 0}$ to be a martingale is

$$0 = \frac{\partial B}{\partial t}(v,t,T) - \kappa B(v,t,T) - v_3 + \frac{\sigma^2}{2}B^2(v,t,T) + e^{-v_1}\tilde{\nu}(-\eta B(v,t,T) + v_2e^{-rt}) - 1$$

and

$$0 = \frac{\partial A}{\partial t}(v,t,T) + \kappa\theta B(v,t,T).$$

The conclusion follows. □

It is now easy to characterize the Laplace transform of $(P_t^r)_{t\geqslant 0}$, which is the most interesting for our purpose.

Corollary 1.1. *There exist functions A, B such that for all $0 \leqslant t \leqslant T$,*

$$\mathbb{E}[\exp\{-u(P_T^r - P_t^r)\}|\mathcal{F}_t] = \exp\{A(u,t,T) + \lambda_{t-}B(u,t,T)\}$$

where A, B satisfy the system of ODE's

$$\frac{\partial A}{\partial t}(u,t,T) = -\kappa\theta B(u,t,T)$$

$$\frac{\partial B}{\partial t}(u,t,T) = \kappa B(u,t,T) - \frac{\sigma^2}{2}B^2(u,t,T) - \tilde{\nu}(-\eta B(u,t,T) + ue^{-rt}) + 1$$

and A, B satisfy the terminal condition $A(u,T,T) = B(u,T,T) = 0$.

Proof. Take $v_1 = v_3 = 0$ and $v_2 = u$ in Proposition 1.2. □

Note that in Proposition 1.2 as well as in Corollary 1.1, it holds that $A(0,t,T) = B(0,t,T) = 0$ for all t,T. The Laplace transform of $(P_t^r)_{t\geqslant 0}$ will be useful in the numerical application section to compute moments.

The next section is dedicated to the derivation of the time-consistent counterparts of well known actuarial evaluations, namely the variance principle, the standard-deviation principle and the mean value principle. The reader is referred to Kaas et al. (2008) for an introduction to these actuarial evaluation principles.

1.4 Time-Consistent Evaluations

This section extends the approach of Pelsser and Ghalehjooghi (2016) to the framework we introduced. A time-consistent evaluation principle is an evaluation principle that has the following property: if it is known today that, at a future date, a risk X will be more expensive than another risk Y, then the risk X should already be more expensive today. An equivalent characterization of time-consistency is the following: let $(\Pi_t)_{t\geqslant 0}$ be a dynamic evaluation. Such an evaluation is time-consistent if and only if

$$\Pi_t(X) = \Pi_t(e^{-r\Delta}\Pi_{t+\Delta}(X)) \tag{1.8}$$

for any risk X and any $t, \Delta \geqslant 0$. This equivalent characterization is the one we will use in the sequel.

We assume that the companies of the credit portfolio are small or medium companies for which no credit derivatives are available. These risks being not hedgeable by financial instruments, the market is incomplete and a unique risk-neutral measure does not exist. To manage this uncertainty about the risk-neutral measure, we use actuarial valuation methods and work under the real probability measure $\mathbb{P}$.

Variance Principle

For a risk X expressed in a certain monetary unit, the static variance principle is computed as

$$\mathbb{E}[X] + \alpha\mathbb{V}\mathrm{ar}(X). \tag{1.9}$$

The variance principle is derived in Pratt (1964). More precisely, it is shown that the risk premium asked by an economic agent for a small risk is approximately proportional to the variance of the risk. The parameter α is then a measure of risk aversion. Note that in Equation (1.9), the expectation $\mathbb{E}[X]$ is expressed in the monetary unit, whereas the term $\mathbb{V}\mathrm{ar}(X)$ is expressed in squared monetary units. Therefore the parameter α should be replaced by α/L, where $L > 0$ is a reference amount expressed in the monetary unit. Equivalently, this means that α is a quantity expressed in $1/m_u$, where m_u is the monetary unit.

Let Π^v be a function that is the time-consistent valuation function derived from the variance principle of the trade credit insurance portfolio. This function depends on (t, λ_{t-}). In the case of the variance principle, the time-consistency requirement (1.8) writes

$$\begin{aligned}\Pi^v(t,\lambda_{t-}) = \mathbb{E}\left[e^{-r\varepsilon}\Pi^v(t+\varepsilon,\lambda_{t+\varepsilon-})|\,\mathcal{F}_t\right] \\ + \alpha\mathbb{V}\mathrm{ar}\left(e^{-r\varepsilon}\Pi^v(t+\varepsilon,\lambda_{t+\varepsilon-})|\,\mathcal{F}_t\right),\end{aligned} \tag{1.10}$$

for any $t, \varepsilon \geqslant 0, t+\varepsilon \in [0,T]$, where $T > 0$ is the maturity of the contracts. This means that we should have

$$\begin{aligned}&\lim_{\varepsilon\downarrow 0}\frac{\mathbb{E}\left[e^{-r\varepsilon}\Pi^v(t+\varepsilon,\lambda_{t+\varepsilon-})|\,\mathcal{F}_t\right] + \alpha\mathbb{V}\mathrm{ar}\left(e^{-r\varepsilon}\Pi^v(t+\varepsilon,\lambda_{t+\varepsilon-})|\,\mathcal{F}_t\right) - \Pi^v(t,\lambda_{t-})}{\varepsilon} \\ &\quad = 0.\end{aligned} \tag{1.11}$$

We will separately look at the expectation term and the variance term at the right-hand side of Equation (1.10). For the expectation part, using Ito's lemma for semimartingales (see e.g. Chapter 2 of Protter (2005) or Chapter 6 Section 7 of Rogers and Williams (2000)) yields

$$\begin{aligned}&\lim_{\varepsilon\downarrow 0}\mathbb{E}\left[\frac{\Pi^v(t+\varepsilon,\lambda_{t+\varepsilon-}) - \Pi^v(t,\lambda_{t-})}{\varepsilon}\,|\,\mathcal{F}_t\right] \\ &\quad = \Pi^v_t(t,\lambda_{t-}) + \kappa(\theta-\lambda_{t-})\Pi^v_\lambda(t,\lambda_{t-}) \\ &\quad + \frac{\sigma^2\lambda_t}{2}\Pi_{\lambda\lambda}(t,\lambda_{t-}) + \mathbb{E}[\xi + \Pi^v(t,\lambda_{t-}+\eta\xi) - \Pi^v(t,\lambda_{t-})]\lambda_{t-}\end{aligned} \tag{1.12}$$

where Π^v_t, Π^v_λ and $\Pi^v_{\lambda\lambda}$ are shorthands for $\partial\Pi^v/\partial t, \partial\Pi^v/\partial\lambda$ and $\partial^2\Pi^v/\partial\lambda^2$ respectively and $\xi \sim \nu$. Note that the result of Equation (1.12) is the infinitesimal generator of Π^v, that we will denote by $\mathcal{A}\Pi^v$. For the variance part, we will use infinitesimal generators. First, let us recall the following elementary identity for variance

$$\mathbb{V}\mathrm{ar}(Y_{t+\varepsilon}|\mathcal{F}_t) = \mathbb{E}[Y^2_{t+\varepsilon}|\mathcal{F}_t] - Y_t^2 - (\mathbb{E}[Y_{t+\varepsilon}|\mathcal{F}_t] - Y_t)^2 - 2Y_t(\mathbb{E}[Y_{t+\varepsilon}|\mathcal{F}_t] - Y_t) \tag{1.13}$$

where $(Y_t)_{t\geqslant 0}$ is used instead of $(\Pi^v(t,\lambda_{t-}))_{t\geqslant 0}$ for shorter notations. Dividing Equation (1.13) by ε on both sides and letting $\varepsilon \downarrow 0$ leads to

$$\lim_{\varepsilon\to 0}\frac{\mathbb{V}\mathrm{ar}(\Pi^v(t+\varepsilon,\lambda_{t+\varepsilon-})|\mathcal{F}_t)}{\varepsilon} = \mathcal{A}\left(\Pi^{v2}\right) - 2\Pi^v\left(\mathcal{A}\Pi^v\right) \tag{1.14}$$

where Π^{v2} means $(\Pi^v)^2$ and is used to avoid too many parentheses. We also remind that Π^v is of course a shorthand for $\Pi^v(t,\lambda_{t-})$. The use of Ito's lemma for semimartingales and the chain rule allow us to get the following expression for the infinitesimal generator $\mathcal{A}\left(\Pi^{v2}\right)$:

$$\begin{aligned}\mathcal{A}\left(\Pi^{v2}\right) &= 2\Pi^v\Pi^v_t + 2\Pi^v\Pi^v_\lambda\kappa(\theta-\lambda_{t-}) + \frac{1}{2}\sigma^2\lambda_{t-}\left(2\left(\Pi^v_\lambda\right)^2 + 2\Pi^v\Pi^v_{\lambda\lambda}\right)\\ &\quad + \left(\mathbb{E}[(\xi+\Pi^v(t,\lambda_{t-}+\eta\xi))^2] - \Pi^{v2}\right)\lambda_{t-}\end{aligned} \tag{1.15}$$

where, again, $\xi\sim\nu$. Then, starting from Equation (1.11), using that $e^{-r\varepsilon} = 1-r\varepsilon+o(\varepsilon)$ and using our computations regarding the expectation and the variance, we obtain a PIDE for the time-consistent variance price

$$\begin{aligned}r\Pi^v &= \Pi^v_t + \kappa(\theta-\lambda_{t-})\Pi^v_\lambda + \frac{1}{2}\sigma^2\lambda_{t-}[\Pi^v_{\lambda\lambda} + 2\alpha(\Pi^v_\lambda)^2]\\ &\quad + \mathbb{E}[\xi + \Pi^v(t,\lambda_{t-}+\eta\xi) - \Pi^v(t,\lambda_{t-})]\lambda_{t-}\\ &\quad + \alpha\mathbb{E}\left[(\xi+\Pi^v(t,\lambda_{t-}+\eta\xi) - \Pi^v(t,\lambda_{t-}))^2\right]\lambda_{t-}.\end{aligned} \tag{1.16}$$

Given the computations we already performed, deriving a PIDE for the time-consistent standard-deviation principle is an easy matter.

Standard-Deviation Principle

For a risk X expressed in a certain monetary unit, the static standard-deviation principle is computed as

$$\mathbb{E}[X] + \beta\sqrt{\mathbb{V}\mathrm{ar}(X)}. \tag{1.17}$$

Let Π^s be a function that is the time-consistent valuation derived from the standard-deviation principle. Again, this function depends on (t,λ_{t-}). For any $t+\varepsilon\in[0,T]$, we should have

$$\begin{aligned}\Pi^s(t,\lambda_{t-}) &= \mathbb{E}\left[e^{-r\varepsilon}\Pi^s(t+\varepsilon,\lambda_{t+\varepsilon-})\,|\,\mathcal{F}_t\right]\\ &\quad + \beta\sqrt{\mathbb{V}\mathrm{ar}\left(e^{-r\varepsilon}\Pi^s(t+\varepsilon,\lambda_{t+\varepsilon-})\,|\,\mathcal{F}_t\right)}\end{aligned} \tag{1.18}$$

for a fixed constant $\beta>0$. As noted in Pelsser and Ghalehjooghi (2016), the terms $\mathbb{E}[X]$ and $\sqrt{\mathbb{V}\mathrm{ar}(X)}$ have different time scales. That is, the expectation term scales linearly with the time whereas the standard-deviation term scales with the square root of the time. Indeed, for a small $\Delta_t > 0$, Equation (1.12) shows that

$$\mathbb{E}[\Pi(t+\Delta_t, \lambda_{t+\Delta_t-}) - \Pi(t, \lambda_{t-})|\mathcal{F}_t] \approx (\mathcal{A}\Pi)\Delta_t$$

whereas Equation (1.14) shows that

$$\mathbb{V}\mathrm{ar}(\Pi(t+\Delta_t, \lambda_{t+\Delta_t-})|\mathcal{F}_t) \approx (\mathcal{A}(\Pi^2) - 2\Pi\mathcal{A}\Pi)\Delta_t$$

so that both the expectation and the variance scales linearly with time. It follows that the standard-deviation scales linearly with the square root of the time. Therefore, the β parameter must be replaced by $\sqrt{\varepsilon}\beta$ in Equation (1.18), which yields

$$\begin{aligned}\Pi^s(t, \lambda_{t-}) = {} & \mathbb{E}\left[e^{-r\varepsilon}\Pi^s(t+\varepsilon, \lambda_{t+\varepsilon-}) \,|\, \mathcal{F}_t\right] \\ & + \beta\sqrt{\varepsilon}\sqrt{\mathbb{V}\mathrm{ar}\left(e^{-r\varepsilon}\Pi^s(t+\varepsilon, \lambda_{t+\varepsilon-}) \,|\, \mathcal{F}_t\right)}.\end{aligned} \tag{1.19}$$

From an economic point of view, it means that if we want to compare a standard-deviation measured over a year with a standard-deviation measured over a month, the annual outcome should be scaled with $\sqrt{1/12}$ to become fairly comparable. From Equation (1.19), we proceed as for the time-consistent variance principle: we divide by ε and let ε tend to 0. Using the results we obtained for the variance principle, we have

$$\begin{aligned}& r\Pi^s(t, \lambda_{t-}) = \Pi^s_t(t, \lambda_{t-}) + \kappa(\theta - \lambda_{t-})\Pi^s_\lambda(t, \lambda_{t-}) + \frac{\sigma^2\lambda_{t-}}{2}\Pi^s_{\lambda\lambda}(t, \lambda_{t-}) \\ & + \mathbb{E}\left[\xi + \Pi^s(t, \lambda_{t-} + \eta\xi) - \Pi^s(t, \lambda_{t-})\right]\lambda_{t-} \\ & + \beta\lim_{\varepsilon\downarrow 0}\sqrt{\frac{\mathbb{V}\mathrm{ar}(\Pi^s(t+\varepsilon, \lambda_{t+\varepsilon-})|\mathcal{F}_t)}{\varepsilon}}.\end{aligned} \tag{1.20}$$

Equation (1.20) also provides a more mathematical reason to the multiplication of the standard-deviation by $\sqrt{\varepsilon}$: it is exactly what we need to obtain the convergence of the last term when $\varepsilon \to 0$. With the help of Equations (1.14) and (1.15) we can compute the limit of the standard-deviation term and arrive to the following PIDE

$$\begin{aligned}r\Pi^s(t, \lambda_{t-}) = {} & \Pi^s_t(t, \lambda_{t-}) + \kappa(\theta - \lambda_{t-})\Pi^s_\lambda(t, \lambda_{t-}) + \frac{\sigma^2\lambda_{t-}}{2}\Pi^s_{\lambda\lambda}(t, \lambda_{t-}) \\ & + \mathbb{E}\left[\xi + \Pi^s(t, \lambda_{t-} + \eta\xi) - \Pi^s(t, \lambda_{t-})\right]\lambda_{t-} \\ & + \beta\left[\mathbb{E}\left[(\xi + \Pi^s(t, \lambda_{t-} + \eta\xi) - \Pi^s)^2\right]\lambda_{t-} + \sigma^2\lambda_{t-}\left(\Pi_\lambda\right)^2\right]^{1/2}.\end{aligned} \tag{1.21}$$

In the next subsection, we derive the time-consistent counterpart of an evaluation known as the mean value principle.

Mean Value Principle

The static mean value principle is given by

$$\Pi(X) = v^{-1}\Big(\mathbb{E}[v(X)]\Big) \tag{1.22}$$

where v can be any convex increasing function. Jensen's inequality ensures that this principle gives a higher price than an expected value. The time-consistent evaluation operator Π^m derived from the static Mean Value principle should satisfy

$$\Pi^m(t, \lambda_{t-}) = v^{-1}(\mathbb{E}[v(e^{-r\varepsilon}\Pi^m(t+\varepsilon, \lambda_{t+\varepsilon-}))|\mathcal{F}_t])$$

for any $0 \leqslant t+\varepsilon \leqslant T$, with $t, \varepsilon \geqslant 0$. It implies that

$$\lim_{\varepsilon \downarrow 0} \frac{\mathbb{E}\left[v\left(\frac{\Pi^m(t+\varepsilon, \lambda_{t+\varepsilon-})}{e^{r\varepsilon}}\right)|\mathcal{F}_t\right] - v(\Pi^m(t, \lambda_{t-}))}{\varepsilon} = 0$$

for all $t \geqslant 0$. Using this condition and Ito's Lemma on the process $\left(v\left(\frac{\Pi^m(t+\varepsilon, \lambda_{t+\varepsilon-})}{e^{r\varepsilon}}\right)\right)_{\varepsilon \geqslant 0}$, we will derive a partial differential equation for the time-consistent price. In order to apply Ito's Lemma, we compute the partial derivatives of the function $(\varepsilon, \lambda) \mapsto f(\varepsilon, \lambda) := v\left(\frac{\Pi^m(t+\varepsilon, \lambda)}{e^{r\varepsilon}}\right)$ by the use of the chain rule. This leads to

$$\frac{\partial f}{\partial \varepsilon}(s, \lambda_{t+s-}) =$$
$$v'\left(\frac{\Pi^m(t+s, \lambda_{t+s-})}{e^{rs}}\right) \frac{\Pi^m_t(t+s, \lambda_{t+s-}) - r\Pi^m(t+s, \lambda_{t+s-})}{e^{rs}},$$

$$\frac{\partial f}{\partial \lambda}(s, \lambda_{t+s-}) = v'\left(\frac{\Pi^m(t+s, \lambda_{t+s-})}{e^{rs}}\right) \frac{\Pi^m_\lambda(t+s, \lambda_{t+s-})}{e^{rs}},$$

and

$$\begin{aligned}\frac{\partial^2 f}{\partial \lambda^2}(s, \lambda_{t+s-}) &= \left(\frac{\Pi^m_\lambda(t+s, \lambda_{t+s-})}{e^{rs}}\right)^2 v''\left(\frac{\Pi^m(t+s, \lambda_{t+s-})}{e^{rs}}\right) \\ &\quad + v\left(\frac{\Pi^m(t+s, \lambda_{t+s-})}{e^{rs}}\right) \frac{\Pi^m_{\lambda\lambda}(t+s, \lambda_{t+s-})}{e^{rs}}.\end{aligned}$$

As before, from Ito's Lemma, we find

$$
\begin{aligned}
0 = & \frac{\sigma^2}{2}\lambda_{t-}\left(\Pi^m_\lambda(t,\lambda_{t-})\right)^2 v''(\Pi^m(t,\lambda_{t-})) \\
& + \mathbb{E}\left[v(\xi + \Pi^m(t,\lambda_{t-}+\eta\xi)) - v(\Pi^m(t,\lambda_{t-}))\right]\lambda_{t-} \\
& + v'(\Pi^m(t,\lambda_{t-}))\Bigg[\Pi^m_t(t,\lambda_{t-}) - r\Pi^m(t,\lambda_{t-}) \\
& \qquad + \Pi^m_\lambda(t,\lambda_{t-})\kappa(\theta-\lambda_{t-}) + \frac{\sigma^2}{2}\lambda_{t-}\Pi^m_{\lambda\lambda}(t,\lambda_{t-})\Bigg],
\end{aligned}
\tag{1.23}
$$

which can be used as a PIDE for the time-consistent price.

The Exponential Principle

The exponential principle is a particular case of the mean value principle. It is obtained by defining the function v by $v(x) = e^{\gamma^{-1}x}$, $\gamma > 0$. A PIDE for the time-consistent exponential principle is obtained by replacing v, $v'(x) = \gamma^{-1}e^{\gamma^{-1}x}$ and $v''(x) = \gamma^{-2}e^{\gamma^{-1}x}$ in Equation (1.23). We obtain

$$
\begin{aligned}
0 = & \frac{\sigma^2}{2}\lambda_{t-}\left(\Pi^m_\lambda(t,\lambda_{t-})\right)^2\gamma^{-2}e^{\gamma^{-1}\Pi^m(t,\lambda_{t-})} \\
& + \frac{\Pi^m_t(t,\lambda_{t-}) - r\Pi^m(t,\lambda_{t-}) + \Pi^m_\lambda(t,\lambda_{t-})\kappa(\theta-\lambda_{t-}) + \frac{\sigma^2}{2}\lambda_{t-}\Pi^m_{\lambda\lambda}(t,\lambda_{t-})}{\gamma e^{-\gamma^{-1}\Pi^m(t,\lambda_{t-})}} \\
& + \mathbb{E}\left[e^{\gamma^{-1}(\xi+\Pi^m(t,\lambda_{t-}+\eta\xi))} - e^{\gamma^{-1}\Pi^m(t,\lambda_{t-})}\right]\lambda_{t-}.
\end{aligned}
\tag{1.24}
$$

Since $e^{\gamma^{-1}\Pi^m(t,\lambda_{t-})} > 0$, we can divide both sides of Equation (1.24) by $e^{\gamma^{-1}\Pi^m(t,\lambda_{t-})}$, so that we get

$$
\begin{aligned}
& \frac{\Pi^m_t(t,\lambda_{t-}) - r\Pi^m(t,\lambda_{t-}) + \Pi^m_\lambda(t,\lambda_{t-})\kappa(\theta-\lambda_{t-}) + \frac{\sigma^2}{2}\lambda_{t-}\Pi^m_{\lambda\lambda}(t,\lambda_{t-})}{\gamma} \\
& + \mathbb{E}\left[e^{\gamma^{-1}(\xi+\Pi^m(t,\lambda_{t-}+\eta\xi)-\Pi^m(t,\lambda_{t-}))} - 1\right]\lambda_{t-} + \frac{\sigma^2}{2}\lambda_{t-}\left(\Pi^m_\lambda(t,\lambda_{t-})\right)^2\gamma^{-2} \\
& = 0.
\end{aligned}
\tag{1.25}
$$

This plays a role in the numerical methods part. Indeed, the exponential of the evaluation $e^{\gamma^{-1}\Pi^m(t,\lambda_{t-})}$ can lead to huge numbers, causing problems in the numerical solving of the PDE. It is therefore important to get rid of it.

The last evaluation used in Pelsser and Ghalehjooghi (2016) is the cost-of-capital principle, which requires the computation of a value-at-risk. The stationary and independent increments in their processes are used to derive an approximation by a normal distribution and thereby find a PIDE for the time-consistent cost-of-capital principle. In our case, the self-exciting jumps prevent the increments of $(\lambda_t)_{t\geqslant 0}$ from being independent and stationary. As a consequence, we can't derive a PIDE in our setting for the time-consistent cost-of-capital principle.

1.5 Numerical Application

The goal of this numerical application section is to explore numerically the impact of the risk of contagion between defaults and the impact of the time-consistency of evaluations. We are also interested in observing possible interactions between time-consistency and contagion. In order to do so, we start by explaining how to compute actual evaluations on the basis of the results presented above. This is done firstly for the time-inconsistent evaluations and afterwards for their time-consistent counterparts. The numerical results are then presented and analyzed. Finally, we close this section with a calibration method for our model.

Time-inconsistent Setting

The time-inconsistent evaluations represented in the upcoming graphs correspond to the functions

$$[0,T] \ni t \mapsto \mathbb{E}[e^{rt}(P_T^r - P_t^r)|\mathcal{F}_t] + \alpha \mathbb{V}\mathrm{ar}(e^{rt}(P_T^r - P_t^r)|\mathcal{F}_t), \tag{1.26}$$

$$[0,T] \ni t \mapsto \mathbb{E}[e^{rt}(P_T^r - P_t^r)|\mathcal{F}_t] + \beta\sqrt{T-t}\sqrt{\mathbb{V}\mathrm{ar}(e^{rt}(P_T^r - P_t^r)|\mathcal{F}_t)} \tag{1.27}$$

and

$$[0,T] \ni t \mapsto \gamma \ln\left(\mathbb{E}\left[\exp\left\{\frac{e^{rt}(P_T^r - P_t^r)}{\gamma}\right\}|\mathcal{F}_t\right]\right) \tag{1.28}$$

where $e^{rt}(P_T^r - P_t^r) = \int_t^T e^{-r(s-t)}\mathrm{d}P_s$ is the remaining discounted cost of the portfolio seen from time $t \in [0,T]$. The expectation is computed with

$$\begin{aligned}
\mathbb{E}[(P_T^r - P_t^r)|\mathcal{F}_t] &= (-1)\frac{\partial}{\partial u}\mathbb{E}[\exp\{-u(P_T^r - P_t^r)\}|\mathcal{F}_t]\Big|_{u=0} \\
&= -\left(\frac{\partial A}{\partial u}(u,t,T) + \lambda_{t-}\frac{\partial B}{\partial u}(u,t,T)\right)\Big|_{u=0} \\
&\approx -\frac{A(\Delta_u,t,T) - A(-\Delta_u,t,T)}{2\Delta_u} - \lambda_{t-}\frac{B(\Delta_u,t,T) - B(-\Delta_u,t,T)}{2\Delta_u}
\end{aligned} \tag{1.29}$$

that holds for a small $\Delta_u > 0$. The functions A and B come from Corollary 1.1. Similarly, we have for the second moment

$$\begin{aligned}
&\mathbb{E}[(P_T^r - P_t^r)^2|\mathcal{F}_t] = \frac{\partial^2}{\partial u^2}\mathbb{E}[\exp\{-u(P_T^r - P_t^r)\}|\mathcal{F}_t]\Big|_{u=0} \\
&= \frac{\partial}{\partial u}\left[(-1)\left(\frac{\partial A}{\partial u}(u,t,T) + \lambda_{t-}\frac{\partial B}{\partial u}(u,t,T)\right)\exp\{A(u,t,T) + \lambda_{t-}B(u,t,T)\}\right]\Bigg|_{u=0} \\
&= \left[\left(\frac{\partial^2 A}{\partial u^2}(u,t,T) + \lambda_{t-}\frac{\partial^2 B}{\partial u^2}(u,t,T)\right) + \left(\frac{\partial A}{\partial u}(u,t,T) + \lambda_{t-}\frac{\partial B}{\partial u}(u,t,T)\right)^2\right]\Bigg|_{u=0} \\
&\approx \left(\frac{A(\Delta_u,t,T) + A(-\Delta_u,t,T)}{\Delta_u^2} + \lambda_{t-}\frac{B(\Delta_u,t,T) + B(-\Delta_u,t,T)}{\Delta_u^2}\right) \\
&+ \left(\frac{A(\Delta_u,t,T) - A(-\Delta_u,t,T)}{2\Delta_u} + \lambda_{t-}\frac{B(\Delta_u,t,T) - B(-\Delta_u,t,T)}{2\Delta_u}\right)^2
\end{aligned} \tag{1.30}$$

where we used the fact that $A(0,t,T) = B(0,t,T) = 0$ within the usual second order difference approximation $\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}(x) \approx (f(x+\Delta) - 2f(x) + f(x-\Delta))/\Delta^2$. By combining Equations (1.29) and (1.30), we get

$$\begin{aligned}
\mathbb{V}\mathrm{ar}((P_T^r - P_t^r)|\mathcal{F}_t) \approx &\frac{A(\Delta_u,t,T) + A(-\Delta_u,t,T)}{\Delta_u^2} \\
&+ \lambda_{t-}\frac{B(\Delta_u,t,T) + B(-\Delta_u,t,T)}{\Delta_u^2}.
\end{aligned} \tag{1.31}$$

We now explain how we approximate the values of the functions A and B with the help of Corollary 1.1. We proceed recursively starting from $A(u,T,T) = B(u,T,T) = 0$. We use a forward difference approximation, which gives for $k = 0, 1, \ldots$

$$\begin{aligned}
\frac{\partial B}{\partial t}&(u, T-(k+1)\Delta_t, T) \approx \frac{B(u, T-k\Delta_t, T) - B(u, T-(k+1)\Delta_t, T)}{\Delta_t} \\
&= \kappa B(u, T-(k+1)\Delta_t, T) - \frac{\sigma^2}{2}B^2(u, T-(k+1)\Delta_t, T) \\
&- \tilde{\nu}\left(-\eta B(u, T-(k+1)\Delta_t, T) + ue^{-r(T-(k+1))}\right) + 1
\end{aligned} \tag{1.32}$$

and we solve numerically for $B(u, T-(k+1)\Delta_t, T)$. Similarly, the values of A are found using

$$A(u, T-(k+1)\Delta_t, T) = A(u, T-k\Delta_t, T) + \Delta_t\kappa\theta B(u, T-(k+1)\Delta_t, T). \tag{1.33}$$

Note that the difference steps Δ_u and Δ_t were fixed to 10^{-6} and 2×10^{-5} respectively in all our computations of time-inconsistent evaluations.

Time-consistent Setting

In this section, we explain the numerical method that allows us to solve the 3 PIDE's derived in Section 1.4. To this end, we start by explaining the general setting of the method. Afterwards we address the problem of solving each PIDE specifically.

We work with a finite difference method. There are two variables to discretize: the intensity λ and the time t. The finite difference method is performed on a finite grid for the two variables. The limits for the time will be taken to be 0 and T for the time, T being the maturity of the contract whereas the limits for λ are taken to be 0 and $\lambda_{\max} > 0$. The boundary $\lambda_{\max}$ will be fixed so that it is large enough to be attained by $(\lambda_t)_{t\geqslant 0}$ with a very low probability. We define n_λ and n_t, strictly positive integers, to be the number of steps in the grid for λ and t respectively. These integers imply mesh sizes for both discretized variables, we will denote those mesh sizes by $\Delta_\lambda := \lambda_{\max}/n_\lambda$ and $\Delta_t := T/n_t$. The grid for the time is $(t_0, t_1, \dots, t_{n_t})$, where $t_i = i\Delta_t$ for each i. The grid for the intensity λ_t is $(\lambda_0, \lambda_1, \dots, \lambda_{n_\lambda})$ where $\lambda_j = j\Delta_\lambda$ for each j.

The approximated values of the function Π on the grid will be denoted by $p_{i,j} \approx \Pi(t_i, \lambda_j)$. The matrix that contains all approximated values $p_{i,j}$ will be denoted by $P = (p_{i,j})_{\substack{i=0,\dots,n_t \\ j=0,\dots,n_\lambda}}$ The approximations of partial derivatives will be denoted as

$$\partial_t p_{i,j} \approx \frac{\partial \Pi}{\partial t}(t_i, \lambda_j) \qquad \partial_\lambda p_{i,j} \approx \frac{\partial \Pi}{\partial \lambda}(t_i, \lambda_j) \qquad \partial_{\lambda\lambda} p_{i,j} \approx \frac{\partial^2 \Pi}{\partial \lambda^2}(t_i, \lambda_j) \tag{1.34}$$

so that $\partial_t P = (\partial_t p_{i,j})_{\substack{i=1,\dots,n_t+1 \\ j=1,\dots,n_\lambda+1}}$, $\partial_\lambda P = (\partial_\lambda p_{i,j})_{\substack{i=1,\dots,n_t+1 \\ j=1,\dots,n_\lambda+1}}$ and $\partial_{\lambda\lambda} P = (\partial_{\lambda\lambda} p_{i,j})_{\substack{i=1,\dots,n_t+1 \\ j=1,\dots,n_\lambda+1}}$ are matrices that approximate the needed partial derivatives of Π.

The only boundary condition in this problem is $\Pi(T, \lambda) = 0$ for any possible intensity λ. Indeed, at time T the contract is over and is thus worth 0. Consequently, we set $p_{n_t,j} = 0$ for all $j \in \{0, 1, \dots, n_\lambda\}$. All the other $p_{i,j}$'s will be found with a finite difference method based on the PIDE's. The approximations of the partial derivatives with respect to time are, for all $i \in \{0, 1, \dots, n_t - 1\}$ and $j \in \{0, 1, \dots, n_\lambda\}$, given by

$$\partial_t p_{i,j} = \frac{p_{i+1,j} - p_{i,j}}{\Delta_t}.$$

Concerning the approximations of the partial derivatives with respect to λ, we use the following formulas

$$\partial_\lambda p_{i,j} = \begin{cases} \theta_{\text{or}} \frac{p_{i,j+2}-p_{i,j}}{2\Delta_t} + (1-\theta_{\text{or}}) \frac{p_{i+1,j+2}-p_{i+1,j}}{2\Delta_t} & \text{if } j = 0 \\ \theta_{\text{or}} \frac{p_{i,j}-p_{i,j-2}}{2\Delta_\lambda} + (1-\theta_{\text{or}}) \frac{p_{i+1,j}-p_{i+1,j-2}}{2\Delta_\lambda} & \text{if } j = n_\lambda \\ \theta_{\text{or}} \frac{p_{i,j+1}-p_{i,j-1}}{2\Delta_\lambda} + (1-\theta_{\text{or}}) \frac{p_{i+1,j+1}-p_{i+1,j-1}}{2\Delta_\lambda} & \text{if } j \notin \{0, n_\lambda\} \end{cases} \tag{1.35}$$

$$\partial_{\lambda\lambda} p_{i,j} = \begin{cases} \theta_{\text{or}} \frac{p_{i,j+2}-2p_{i,j+1}+p_{i,j}}{(\Delta_\lambda)^2} + (1-\theta_{\text{or}}) \frac{p_{i+1,j+2}-2p_{i+1,j+1}+p_{i+1,j}}{(\Delta_\lambda)^2} & \text{if } j = 0 \\ \theta_{\text{or}} \frac{p_{i,j}-2p_{i,j-1}+p_{i,j-2}}{(\Delta_\lambda)^2} + (1-\theta_{\text{or}}) \frac{p_{i+1,j}-2p_{i+1,j-1}+p_{i+1,j-2}}{(\Delta_\lambda)^2} & \text{if } j = n_\lambda \\ \theta_{\text{or}} \frac{p_{i,j+1}-2p_{i,j}+p_{i,j-1}}{(\Delta_\lambda)^2} + (1-\theta_{\text{or}}) \frac{p_{i+1,j+1}-2p_{i+1,j}+p_{i+1,j-1}}{(\Delta_\lambda)^2} & \text{if } j \notin \{0, n_\lambda\} \end{cases} \tag{1.36}$$

where $\theta_{\text{or}} \in [0,1]$ is an over relaxation parameter (the subscript "or" stands for over relaxation). This parameter was fixed to 0.5 in our computations. Our approximations of partial derivatives with respect to λ present a special feature. At Equations (1.35) and (1.36) in the cases $j \in \{0, n_\lambda\}$, the approximations are modified so that we do not fall outside of the grid. This trick allows us to solve the PIDE's in spite of the lack of boundary conditions.

Variance Principle

In this subsection, we focus on solving PIDE (1.16) numerically. The partial derivatives will be approached according to Equations (1.34), (1.35) and (1.36). In addition to partial derivatives with respect to t and λ, this equation also contains two terms that correspond to the expected impact of a default on the price, that is, $\mathbb{E}[\Pi^v(t, \lambda_t + \eta\xi) - \Pi^v(t, \lambda_t)]$ and $\mathbb{E}\left[(\xi + \Pi^v(t, \lambda_{t-} + \eta\xi) - \Pi^v(t, \lambda_{t-}))^2\right]$. These terms will be approximated by discretization of the random variable ξ, which is equivalent to Riemann sum approximations. More precisely the first term is approximated by

$$\begin{aligned} &\mathbb{E}[\xi + \Pi^v(t_i, \lambda_j + \eta\xi) - \Pi^v(t_i, \lambda_j)] \\ &\quad\approx \mathbb{E}[\xi] + \sum_{k=j+1}^{n_\lambda} p^v_{i,k} \nu\left(\left[\frac{(k-j-1)}{\eta}\Delta_\lambda, \frac{(k-j)}{\eta}\Delta_\lambda\right)\right) \\ &\quad+ p^v_{i,n_\lambda} \nu\left(\left[\frac{n_\lambda - j}{\eta}\Delta_\lambda, +\infty\right)\right) - p^v_{i,j}. \end{aligned}$$

Before approaching the second term, it is rewritten as

$$\begin{aligned} &\mathbb{E}[(\xi + \Pi^v(t_i, t_j + \eta\xi) - \Pi^v(t_i, \lambda_j))^2] \\ &\quad\approx \mathbb{E}[\xi^2] + \mathbb{E}[\Pi^{v2}(t_i, \lambda_j + \eta\xi)] + 2\mathbb{E}[\xi\Pi^v(t_i, \lambda_j + \eta\xi)] \\ &\quad+ \Pi^{v2}(t_i, \lambda_j) - 2\Pi^v(t_i, \lambda_j)\mathbb{E}[\xi] - 2\Pi^v(t_i, \lambda_j)\mathbb{E}[\Pi^v(t_i, t_j + \eta\xi)]. \end{aligned} \tag{1.37}$$

In Equation (1.37), the term $\mathbb{E}[\Pi^{v2}(t_i, t_j + \eta\xi)]$ is approximated by

$$\begin{aligned}
&\mathbb{E}[\Pi^{v2}(t_i, t_j + \eta\xi)] \\
&\quad \approx \sum_{k=j+1}^{n_\lambda+1} (p_{i,k})^2 \nu \left(\left[\frac{(k-j-1)}{\eta}\Delta_\lambda, \frac{(k-j)}{\eta}\Delta_\lambda \right) \right) \\
&\quad + (p_{i,n_\lambda+1})^2 \nu \left(\left[\frac{n_\lambda + 1 - j}{\eta}\Delta_\lambda, +\infty \right) \right),
\end{aligned}$$

whereas $\mathbb{E}[\xi\Pi^v(t_i, t_j + \eta\xi)]$ is approximated by

$$\begin{aligned}
&\mathbb{E}[\xi\Pi^v(t_i, t_j + \eta\xi)] \\
&\quad \approx \sum_{k=j+1}^{n_\lambda+1} \frac{(k-j)}{\eta}\Delta_\lambda p_{i,k} \nu \left(\left[\frac{(k-j-1)}{\eta}\Delta_\lambda, \frac{(k-j)}{\eta}\Delta_\lambda \right) \right) \\
&\quad + \mathbb{E}[\xi] p_{i,n_\lambda+1} \nu \left(\left[\frac{n_\lambda + 1 - j}{\eta}\Delta_\lambda, +\infty \right) \right).
\end{aligned}$$

Let $p^v_{i,j}$ be the approximation of $\Pi^v(t_i, \lambda_j)$ and P^v be the matrix that contains all the approximations $p^v_{i,j}$. As already pointed out, the last column of P^v, denoted by[2] $P^v_{\star n_t}$, is filled with zeros. We can now replace the terms of PIDE (1.16) by their approximations described above. For each $k \in \{0, 1, \ldots, n_t - 1\}$, we can find $P^v_{\star k}$ from $P^v_{\star(k+1)}$ by solving numerically a system of equations.

Solving numerically PIDE (1.21), i.e. the one that describes the time-consistent standard-deviation principle, is done in the same way. That is, the method and the approximations are the same, but the system of equations that allows to find $P^s_{\star k}$ from $P^s_{\star(k+1)}$ is modified so that it correspond to PIDE (1.21).

Exponential Principle

We focus on the specificity of the numerical resolution of PIDE (1.25). Apart from the partial derivatives that are again approached by the formulas (1.34), (1.35) and (1.36), we need to approximate $\mathbb{E}\left[e^{\gamma^{-1}(\xi + \Pi^m(t_i, \lambda_j + \eta\xi))}\right]$. This is done with a discretization of the random variable ξ

[2]that is actually the column number $n_t + 1$ of P^v, since there are $n_t + 1$ points in the time grid.

$$
\begin{aligned}
&\mathbb{E}\left[e^{\gamma^{-1}(\xi+\Pi^m(t_i,\lambda_j+\eta\xi))}\right] \\
&\quad\approx \sum_{k=j+1}^{n_\lambda} \exp\left\{\gamma^{-1}\left(\frac{(k-j)}{\eta}\Delta_\lambda + p^m_{i,k}\right)\right\}\nu\left(\left[\frac{(k-j-1)}{\eta}\Delta_\lambda, \frac{(k-j)}{\eta}\Delta_\lambda\right)\right) \\
&\quad+\exp\left\{\gamma^{-1}\left(\frac{n_\lambda-j}{\eta}\Delta_\lambda + p^m_{i,n_\lambda}\right)\right\}\nu\left(\left[\frac{n_\lambda-j}{\eta}\Delta_\lambda, +\infty\right)\right).
\end{aligned}
$$

We can now proceed inductively as before: $P^m_{\star k}$ is deduced from $P^m_{\star(k+1)}$ by solving numerically a system of equations.

Results

This section studies numerically the convergence and stability of the numerical procedure described in Section 1.5 for the computation of the time-consistent evaluations. More precisely, the time-consistent evaluations we represent on the graphs are given by the functions $[0,T] \ni t \mapsto \Pi^v(t, 0.1)$, $[0,T] \ni t \mapsto \Pi^s(t, 0.1)$ and $[0,T] \ni t \mapsto \Pi^m(t, 0.1)$ respectively for the variance, standard-deviation and exponential principles. Note that the price at time t is computed with the assumption that λ_t equals the target level 0.1. In other words, the evaluation at time t is computed with the assumption that the process $(\lambda_t)_{t\geqslant 0}$ satisfies SDE (1.1) but that its value at time t happens to be the target level θ (which is set to 0.1 in the computations).

For the computations, it is also assumed that the jump measure ν satisfies

$$\nu(\mathrm{d}x) = \rho\exp\{-x\rho\}\mathrm{d}x,\ x \geqslant 0$$

so that the jump size ξ is exponentially distributed with mean ρ^{-1}. This is equivalent to assuming that the loss given default has the memoryless property, i.e. $P(\xi > t+s|\xi > s) = P(\xi > t)$ for all $s, t \geqslant 0$. For example, when a default occurs, the probability of losing more than 2000 monetary units given that the loss exceeds 1000 monetary units is the same as the probability of losing more than 1000 monetary units. The values of the parameters we fixed for our numerical experiments are given in Table 1.1. In the abbreviation “SE=yes”, SE stands for “Self-Exciting”, meaning that the model exhibits a phenomenon of contagion between defaults (that is, $\eta > 0$). The second line, “SE=no”, corresponds to a model without contagion ($\eta = 0$). Recall that the parameters α, β and γ are specific to the variance, standard-deviation and exponential principle respectively. Note that in the section devoted to the variance principle, we pointed out that the variance term should be divided by a reference amount L. This reference amount L is here chosen to be the average default cost, ρ^{-1}, as indicated in Table 1.1 (parameter α). The maturity is fixed to $T = 10$ years.

Convergence

We present and discuss the numerical results obtained by solving the PIDE's with the method described in Section 1.5 and the parameters of Table 1.1 (SE=yes). We start by checking the convergence and stability of the results. The convergence with respect to Δ_t is fast and the results are not modified by the choice of $\lambda_{\max}$, as we can observe on Figure 1.1. In this Figure, we represent the evolution of the time-consistent price for the 3 principles with an intensity kept constant and equal to $\lambda_0 = \theta = 0.1$. Note that in order to obtain Figure 1.1, Δ_λ was fixed to 0.0001.

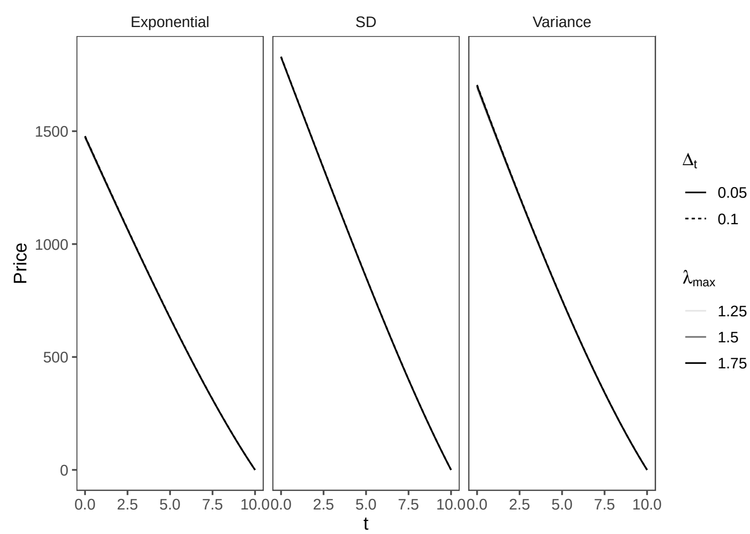

Figure 1.1: Evolution of the Time-Consistent Variance, Standard-Deviation (SD) and Exponential Principles.

Figure 1.2 shows the convergence with respect to Δ_λ, which is visibly slower than the convergence with respect to Δ_t.

Now that we have checked numerically the convergence of our method, we will use it to assess the impact of time-consistency and self-excitation on the price.

Joint Impact of Time-Consistency and Self-Excitation

In this section, the time-consistent evaluations are compared with their time-inconsistent equivalents. The time-inconsistent evaluations are given by the functions of Equations (1.26), (1.27) and (1.28). These functions are computed with the help of the method described in details in Section 1.5. Note that, as for the time-consistent evaluations, the value of λ_t is needed to compute the evaluation at time t. It is why we always assume that λ_t equals the target level 0.1 in the computations.

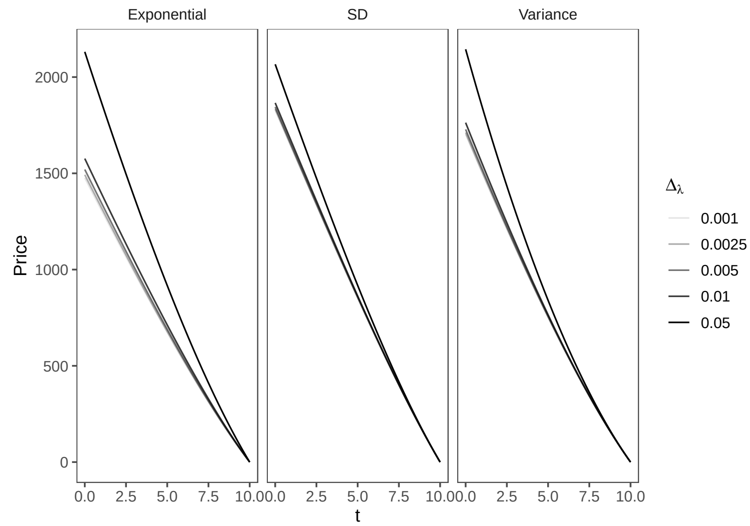

Figure 1.2: Convergence with Respect to Δ_λ.

Table 1.1: Parameters Used for the Computations.

	θ	λ_0	κ	σ	η	ρ	r	T	β	α	γ
SE = yes	0.1	0.1	0.24	0.05	0.0001	0.001	0.02	10	0.1	0.1ρ	10000
SE = no	0.152598	0.1	0.24	0.05	0	0.001	0.02	10	0.1	0.1ρ	10000

Computing prices that ignore contagion requires some comments. Cancelling the phenomenon of contagion in the portfolio can be simply done in our model. Indeed, it suffices to set $\eta = 0$. However, if we do so, the new model is not anymore fairly comparable to the previous one. As a matter of fact, setting $\eta = 0$ will lead to a lower expected number of defaults and therefore to lower prices. This is why we will adjust the parameter θ in order to keep the expected discounted cost $\mathbb{E}[P_T^r]$ unchanged. This is illustrated at Figure 1.3 and the new "corrected" value for θ is raised to 0.152898. More precisely, with the parameters of the first line of Table 1.1, we computed that $\mathbb{E}[P_T^r] = 1197.441$ with the method explained in Section 1.5. This corresponds to the horizontal line of Figure 1.3. This figure also displays the expected value of P_T^r with the same parameters, with the exception of $\eta = 0$, as a function of θ. Therefore, it illustrates that θ should be raised to 0.152598 (vertical line) to keep the expected value unchanged, as indicated in Table 1.1.

Figure 1.4 compares the prices in four different settings. The first setting combines contagion and time-consistency (TC = yes, SE = yes) and is represented by a

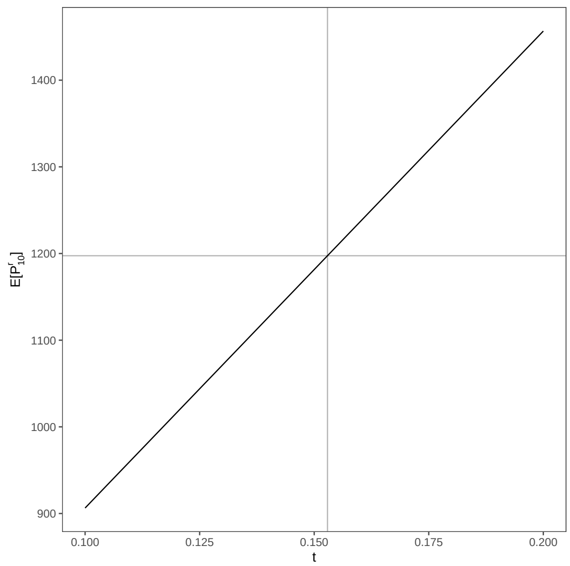

Figure 1.3: Expected Cost of Defaults $\mathbb{E}[P^r_{10}]$ in Function of θ with $\eta = 0$.

dashed dark curve. The second setting is time-consistent but ignores contagion (TC = yes, SE = no). It corresponds to the continuous dark curve. The third and fourth settings are not time-consistent. The third takes the phenomenon of contagion into account and is represented by a clear dashed curve whereas the fourth, represented by a clear continuous curve, does not.

Note that for all three cases, with the exception of the standard-deviation principle where the two continuous lines are indistinguishable on the graph, the dark curves are always above the clear curves. This means that using time-consistent evaluations leads to higher prices in most cases, compared to the corresponding time-inconsistent evaluations. Figure 1.5 makes it easier to see by showing the difference between the time-consistent evaluations and the time-inconsistent ones. A positive difference means that the time-consistent evaluation is more expensive. The standard-deviation principle exhibits the smallest differences, so that it is the least time-inconsistent in our setting. It also distinguishes itself because it is the only one that leads to higher (but very close) prices in the time-inconsistent setting, which happens in the absence of a risk of contagion only. In all other cases, the positive differences indicate that time-inconsistent evaluations tend to underestimate the cost of the credit portfolio. The fact that the dashed curves always lie above the continuous ones means that this underestimation problem is much more severe when there is a risk of contagion. Note however that all curves bond together when the time gets closer to the maturity $T = 10$. This means that for a short maturity such as one or two years, the difference between time-consistent evaluations and time-inconsistent evaluations is small. As a conclusion, the existence of a risk of contagion in a credit portfolio can be seen as an additional reason to work with time-consistent evaluations, when working with long-term contracts.

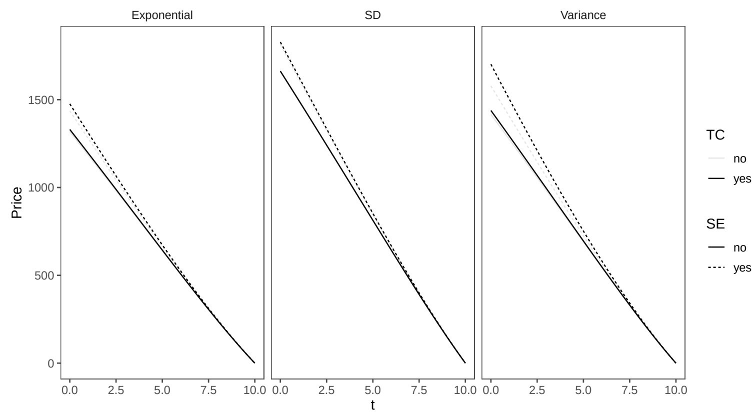

Figure 1.4: Evolution of the Time-Consistent and Time-Inconsistent Prices, with and without Contagion.

Calibration to Real Data

A common practice for calibrating a model in finance is to choose the set of parameters that minimizes the sum of the squared differences between the market prices and the prices implied by the model. However, this approach does not make sense in our case since no market prices are available. We must therefore calibrate the model on past losses that we assume to be representative of the future. Assume that we observed such data on the period of time $[0, T]$. The data consists of pairs $D := \{(t_i, m_i) \in [0, T] \times \mathbb{R}^+ : 1 \leqslant i \leqslant n\}$ that describe the times $0 \leqslant t_1 < \cdots < t_n \leqslant T$ of the defaults that occurred between 0 and T and the corresponding lost amounts $m_1, \ldots, m_n > 0$ in the $N_T = n$ case. Then, assuming that $\sigma = 0$, we can numerically obtain the parameters by maximizing the likelihood. The reason for setting $\sigma = 0$ is the fact that the Brownian term in the intensity $(\lambda_t)_{t \geqslant 0}$ cannot be observed in the data D. Proposition 7.2.III in Daley and Vere-Jones (2006) states that the log-likelihood function L associated to $(N_t)_{t \geqslant 0}$ is given by

$$L(D) = \sum_{j=1}^{n} \ln(\lambda_{t_j -}) - \int_0^T \lambda_{t-} \mathrm{d}t + T,$$

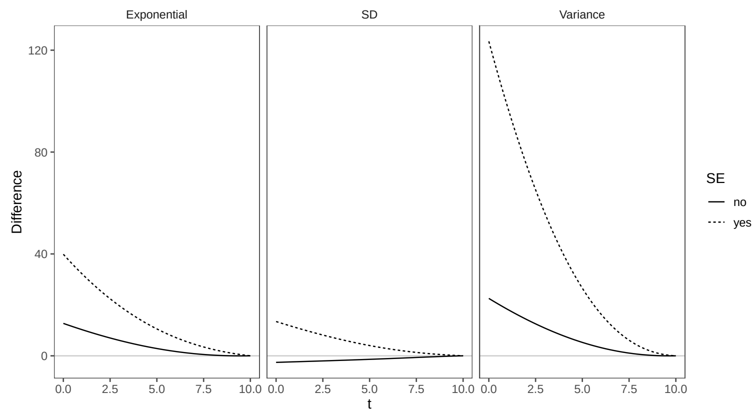

Figure 1.5: Time-Consistent Evaluation minus their Time-Inconsistent Counterparts.

where λ_{t-} denotes the left limit $\lim_{s\uparrow t}\lambda_s$. In particular, note that $\lambda_{t_j-}=\lambda_{t_j}-m_j$. Following Equation (1.3), for any $t\in[0,T]$, the observed past intensity satisfies

$$\lambda_t=\theta+e^{-\kappa t}(\lambda_0-\theta)+\eta\sum_{t_i\leqslant t}e^{-\kappa(t-t_i)}m_i \tag{1.38}$$

so that

$$\sum_{j=1}^{n}\ln(\lambda_{t_j-})=\sum_{j=1}^{n}\ln\left(\theta+e^{-\kappa t_j}(\lambda_0-\theta)+\eta\sum_{i=1}^{j-1}e^{-\kappa(t_j-t_i)}m_i\right).$$

Moreover,

$$\int_0^T\lambda_{t-}\mathrm{d}t=\theta T+(\lambda_0-\theta)\frac{1-e^{-\kappa T}}{\kappa}+\eta\int_0^T\sum_{t_i\leqslant t}e^{-\kappa(t-t_i)}m_i\mathrm{d}t.$$

The last term can be computed as

$$\begin{aligned}
&\int_0^T \sum_{t_i \leqslant t} e^{-\kappa(t-t_i)} m_i \mathrm{d}t \\
&\quad = \underbrace{\int_0^{t_1} \sum_{t_i \leqslant t} e^{-\kappa(t-t_i)} m_i \mathrm{d}t}_{=0} + \sum_{j=1}^{n-1} \int_{t_j}^{t_{j+1}} \sum_{t_i \leqslant t} e^{-\kappa(t-t_i)} m_i \mathrm{d}t \\
&\qquad\quad + \int_{t_n}^{T} \sum_{t_i \leqslant t} e^{-\kappa(t-t_i)} m_i \mathrm{d}t \\
&\quad = \sum_{j=1}^{n-1} \sum_{i=1}^{j} m_i \int_{t_j}^{t_{j+1}} e^{-\kappa(t-t_i)} \mathrm{d}t + \sum_{i=1}^{n} m_i \int_{t_n}^{T} e^{-\kappa(t-t_i)} \mathrm{d}t \\
&\quad = \sum_{j=1}^{n-1} \sum_{i=1}^{j} \frac{m_i e^{\kappa t_i}}{\kappa} \left(e^{-\kappa t_j} - e^{-\kappa t_{j+1}}\right) + \sum_{i=1}^{n} \frac{m_i e^{\kappa t_i}}{\kappa} \left(e^{-\kappa t_n} - e^{-\kappa T}\right).
\end{aligned}$$

Hence, we conclude that the log-likelihood L can be expressed as

$$\begin{aligned}
L(D) = {} & \sum_{j=1}^{n} \ln \left(\theta + e^{-\kappa t_j} (\lambda_0 - \theta) + \eta \sum_{i=1}^{j-1} e^{-\kappa(t_j - t_i)} m_i \right) - (\lambda_0 - \theta) \frac{1 - e^{-\kappa T}}{\kappa} \\
& - \frac{\eta}{\kappa} \left(\sum_{j=1}^{n-1} \sum_{i=1}^{j} m_i \left(e^{-\kappa(t_j - t_i)} - e^{-\kappa(t_{j+1} - t_i)} \right) \right) \\
& + (1-\theta) T - \frac{\eta}{\kappa} \sum_{i=1}^{n} m_i \left(e^{-\kappa(t_n - t_i)} - e^{-\kappa(T - t_i)} \right).
\end{aligned}$$

Considering L as a function of the parameters, the maximum likelihood estimators can be found by a numerical maximization. Then we can use the default costs $m_1, m_2, \ldots, m_n$ to determine the jump size distribution ν. For example, with our assumption that the jump sizes are exponentially distributed, the parameter $\rho > 0$ of the exponential distribution can be approximated by $\hat{\rho} = \left(\frac{1}{n} \sum_{i=1}^{n} m_i\right)^{-1}$, which is also the maximum likelihood estimator.

We assess the performance of the proposed calibration procedure with the help of simulations. In order to do so, we generated 10,000 paths of the processes $(\lambda_t)_{t \geqslant 0}$ and $(P_t)_{t \geqslant 0}$ via an Euler discretization scheme. More precisely, the discretized path $(\lambda_0, \lambda_\Delta, \lambda_{2\Delta}, \ldots, \lambda_{d\Delta} = \lambda_T)$ of $(\lambda_t)_{t \geqslant 0}$ is generated through the following procedure: we set $\lambda_0 \in \mathbb{R}$ and recursively compute

$$\lambda_{(j+1)\Delta} = \lambda_{j\Delta} + (\lambda_{j\Delta} - \theta)\Delta + \eta (X_{j+1} \wedge 1) Y_{j+1}$$

Table 1.2: Real parameters and average of 10,000 maximum likelihood estimations.

	η	κ	ρ	θ
Real value	0.002	5.000	0.0005	1.000
Average estimation	0.002485	5.149	0.0005077	1.067

where X_{j+1} is Poisson distributed with parameter $\Delta\lambda_{j\Delta}$ and Y_{j+1} is exponentially distributed with parameter ρ. All the Poisson and exponential random variables are of course independent. The discretized path $(P_0 = 0, P_\Delta, P_{2\Delta}, \ldots, P_{d\Delta} = P_T)$ of $(P_t)_{t\geqslant 0}$ is then simply

$$P_{j\Delta} = \sum_{i=1}^{j} (X_i \wedge 1) Y_i.$$

The observation period was set to $T = 10$ years and the number of discretization steps was set to $d = 3650$, giving a daily discretization. To simplify the calibrations, we assumed that $\lambda_0 = \theta$. This is not a strong assumption because Equation (1.3) implies that the impact of the gap $(\lambda_0 - \theta)$ between λ_0 and θ on the intensity $(\lambda_t)_{t\geqslant 0}$ is $e^{-\kappa t}(\lambda_0 - \theta)$ at time t, i.e. it decreases exponentially with time. Our 10,000 simulations gave us 10,000 values of the maximum likelihood estimator $(\hat{\eta}, \hat{\kappa}, \hat{\rho}, \hat{\theta})$ of $(\eta, \kappa, \rho, \theta)$. The average values of the 10,000 estimations are compared to the real parameters used for the simulations in Table 1.2. Figure 1.6 displays density plots based on the 10,000 values of the maximum likelihood estimators. We observe that all the estimators are reasonably close to the true values on average.

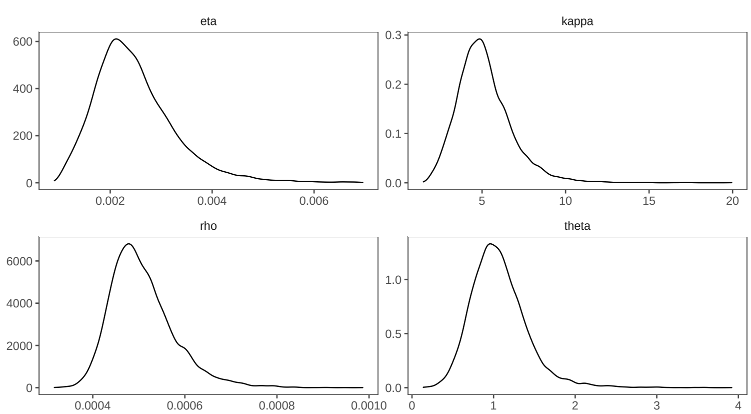

Figure 1.6: Density plots of the 10,000 maximum likelihood estimations.

We close this section by pointing out the interest of Equation (1.38) for a practitioner. Once the parameters are estimated, this equation allows to retrieve the value of the intensity at each time t on the basis of the history of past defaults. As a consequence, one is always able to compute the time-consistent price at all time so that our results are usable in practice.

Conclusions

In this chapter we propose a model for credit insurance portfolio allowing to take the risk of contagion into account. This is done with the help of self-exciting processes. We start by providing closed-form formula for the two first moments of such processes. Then we derive semi-closed form expressions for the Laplace transform of these self-exciting processes. These semi-closed form expressions are then used to derive the distribution of the time of default in a credit insurance model relying on self-exciting processes. We derive PIDE's for time-consistent evaluations of credit insurance portfolio, extending the method of Pelsser and Ghalehjooghi (2016) to models which incorporate self-excitation. In order to numerically solve these PIDE's, we describe a method allowing to find the solution, in spite of the lack of boundary conditions of our problem. By solving numerically these PIDE's, we are able to bring to light the effect of time-consistency on the prices and possible interactions between time-consistency of evaluations and the existence of a risk of contagion. More precisely, we showed that the prices tend to be higher when they are computed with time-consistent evaluations. We also showed that the difference is greater when a risk of contagion exists, motivating the use of time-consistent evaluations in such contexts. Finally, we propose a calibration of the model via likelihood maximization, which qualifies our results for practical use.

Chapter 2

CDS Pricing with Fractional Hawkes Processes

Introduction

The most common paradigms for credit risk modeling are structural and intensity-based models. The first approach is called firm value model, or structural default model. In this approach, the balance sheet of the firm is explicitly modeled and the default is triggered when assets fall below liabilities. For example of such models, we refer the reader to Merton (1974), Black and Cox (1976), Geske (1977), Longstaff and Schwartz (1995), Hainaut and Colwell (2014), Ayadi et al. (2016) and Ballotta et al. (2019). The second approach, the one that is used in this chapter, is called the reduced-form or intensity-based credit model. This second approach was initiated by the work of Duffie and Singleton (1999) and is for example used in Mari and Renò (2005), Hainaut and Le Courtois (2014) and Brigo and Vrins (2018). Textbooks concerning this second kind of models are Bielecki and Rutkowski (2014), Duffie and Singleton (2003), Lando (2004), Jeanblanc et al. (2009) and Schoutens and Cariboni (2009). This chapter deals with a new type of time-changed intensity-based credit model.

The intensity of our model is a mean-reverting process, a kind of process that is often used in finance, for example in Ekvall et al. (1997) or Wong and Lo (2009). The type of credit model used in this chapter has an additional feature: it takes into account the spillover between defaults, also called contagion. As noted in the the introduction to this thesis, the importance of this phenomenon has been emphasized by the financial crisis of 2008. As we emphasized in the general introduction and in the previous chapter, this is one of the main topics of this thesis. A way to model contagion is provided by self-exciting point processes. This kind of processes was already used in Chapter 1 and will also be used in Chapters 3 and 4. For these processes, the default arrival intensity at a given point in time (that is also the instantaneous probability to observe a jump) depends on the number of past defaults, as well as on the elapsed time since the occurence of these defaults. This approach finds its origin in Hawkes (1971a), Hawkes (1971b) and Hawkes and Oakes (1974). More recently, Errais et al. (2010) develop a self-exciting approach for modeling defaults in a credit portfolio. In the most common and simplest specification, the default arrival intensity process is persistent and suddenly increases when a default occurs in the portfolio. Ait-Sahalia et al. (2015) use a similar approach for modeling contagion between different markets. In Hainaut (2016b) and Hainaut (2016a), self-exciting jump-diffusions are used for modeling the term structure of interest rates.

This chapter additionally deals with the modeling of two phenomena within an intensity-based model. The first phenomenon is the presence of periods of time during which the survival of a company is not threatened. In order to properly model such periods of time, the survival probability of the company should remain constant. This contrasts with the usual intensity-based models, where the survival probabilities strictly decrease as time passes. The second phenomenon we aim to model is the random occurrences of one-time events that threaten the short-term viability of the company. The proper modeling of this second phenomenon would imply the possibility to observe very sharp drops in the survival probability of a company. This is again a feature that is not observed in classic intensity-based credit models. Obtaining those two features involves a non-Markov time-change built with the inverse of an α-stable subordinator. This time-change approach is applied to diffusions in physics for describing the movement of heavy particles that can get immobilized (see, e.g. Metzler and Klafter (2004) or Eliazar and Klafter (2004)). This type of time-changed Brownian motions, called sub-diffusions, are also popular in econophysics (see Scalas (2006)) and applied to financial derivatives by Magdziarz (2009a) for illiquid markets. As shown in, e.g. Magdziarz (2009b), the probability density function (PDF) of sub-diffusions is solution of a fractional Fokker-Planck equation. This equation is proposed in Barkai et al. (2000) and Metzler et al. (1999). Articles of Leonenko et al. (2013b) and Leonenko et al. (2013a) go a step further and explore fractional Pearson diffusions and their correlation. We will show that performing a time-change with the inverse of an α-stable subordinator allows to solve one of the issues raised in the general introduction to this thesis. As a matter of fact, it allows to better fit the survival probability curves that are highly convex, such as the curves for airline companies during the Covid-19 crisis.

Note that in this chapter, as well as in the next chapter, we are concerned with the credit risk of a single company. This differs from Chapter 1, where we were interested in a credit portfolio with several companies. Besides, in Chapter 1, the defaults and the phenomenon of contagion between them were completely endogenous with respect to the credit portfolio. In this second chapter, the contagion mechanism is purely exogenous.

This chapter is organized in 3 sections. The first section introduces our framework, without dealing with the time-change. It is also concerned with the derivation of some useful results regarding this framework. The second section deals with the first main contribution of this chapter: a model that incorporates the two aforementioned features and a study of its properties. To build such a model, a non-Markov time-change is added to the setting introduced in Section 1. Results of Section 1 are extended to this new setting with the help of fractional calculus. The third section deals with our second main contribution: a numerical method to compute probabilities of survival that works in both settings, with or without time-changing by the inverse of an α-stable subordinator. This method is similar to the one we used in Chapter 1, since it consists in solving numerically a PDE with the help of finite difference approximations for partial derivatives. Finally, our last main contribution is to show that our numerical method can be used to calibrate the model to real data and obtain a better fit.

2.1 Non-Fractional Setting

In the intensity approach to credit risk, defaults of companies are modeled by jumps of point processes. A jump of a point process corresponds to the default of a company. In this chapter, we are interested in the time of default of a single company. The time of default is the time of the first jump of the point process $(D_t)_{t\geqslant 0}$, that is the positive random variable $\inf\{t > 0 : D_t \geqslant 1\}$. We start by describing precisely how the process $(D_t)_{t\geqslant 0}$ is built.

Let $(N_t)_{t\geqslant 0}$ be a point process on a filtered probability space $(\Omega, \mathbb{F}, \mathbb{P})$ where $\mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0}$ is a filtration to be defined more precisely later on. We denote the intensity of $(N_t)_{t\geqslant 0}$ by $(\lambda_t)_{t\geqslant 0}$, another stochastic process. The intensity represents the instantaneous probability to observe a jump in the process $(N_t)_{t\geqslant 0}$. Let $\xi_1, \xi_2, \ldots$ be independent identically distributed (i.i.d.) exponential random variables with parameter $\rho > 0$. Consequently, if the distribution of $\xi_1, \xi_2, \ldots$ is given by the measure ν on $(\mathbb{R}^+, \mathcal{B}_{\mathbb{R}^+})$, we have

$$\mathrm{d}\nu(x) = \rho e^{-\rho x}\mathrm{d}x, \quad \rho > 0$$

for all $x \in \mathbb{R}^+$, and $\mathbb{E}[\xi_i] = \rho^{-1}$ for all i. We also define the process $(P_t)_{t\geqslant 0}$ as

$$P_t = \sum_{k=1}^{N_t} \xi_k.$$

It is a pure positive jump process whose jump intensity is the stochastic process $(\lambda_t)_{t\geqslant 0}$. The dynamics of this intensity process is assumed to satisfy the stochastic differential equation (SDE)

$$\mathrm{d}\lambda_t = \kappa(\theta - \lambda_{t-})\mathrm{d}t + \eta \mathrm{d}P_t \tag{2.1}$$

where $\kappa, \theta, \eta > 0$ and $\lambda_0 \in [\theta, +\infty)$. As a consequence of SDE (2.1), the intensity depends upon the history of the process $(P_t)_{t\geqslant 0}$. This allows the process $(P_t)_{t\geqslant 0}$ to exhibit a self-exciting behaviour: a jump of the process $(P_t)_{t\geqslant 0}$ induces a jump in the intensity $(\lambda_t)_{t\geqslant 0}$, which in turn implies an increase of the probability to observe another jump in $(P_t)_{t\geqslant 0}$. This is illustrated in Figure 2.1, where each jump of $(\lambda_t)_{t\geqslant 0}$ corresponds to a jump of $(P_t)_{t\geqslant 0}$ (that is not represented here).

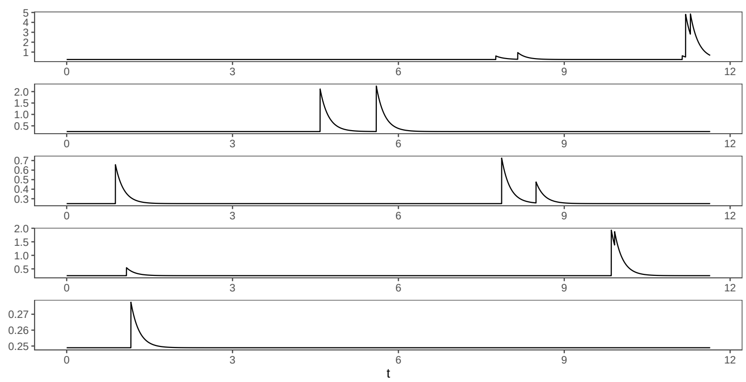

Figure 2.1: 5 Simulated Paths of $(\lambda_t)_{t\geqslant 0}$.

The parameter θ represents a minimum level of instantaneous jump probability, as well as a level to which this instantaneous probability tends to revert. The fact that $\lambda_t \geqslant \theta$ almost surely for all t plays a role throughout this chapter in the derivation of formulas, because it ensures that the PDF of λ_t equals 0 at the point 0. More precisely, if we denote by

$$p(t, x|s, y) = \frac{\partial}{\partial x}\mathbb{P}(\lambda_t \leqslant x|\lambda_s = y)$$

the PDF of λ_t given λ_s, $t \geqslant s$, we always have $p(t, 0|s, y) = 0$. The parameter κ in Equation (2.1) represents the speed of the reversion towards θ.

From an economic point of view, $(\lambda_t)_{t\geqslant 0}$ represents the systemic risk in the economy. Large values of this process represent bad times for companies, increasing the odds to observe defaults. A jump of $(P_t)_{t\geqslant 0}$ represents a bad economic event and induces an increase of the default intensity. The self exciting behavior of $(\lambda_t, P_t)_{t\geqslant 0}$ aims to model the clustering of defaults observed in practice, see, e.g. Ait-Sahalia et al. (2015).

The time of default of the company is modeled as the first jump of the point process $(D_t)_{t\geqslant 0}$, whose intensity is assumed to be $(\lambda_t)_{t\geqslant 0}$. Despite the processes $(D_t)_{t\geqslant 0}$ and $(P_t)_{t\geqslant 0}$ both share the same intensity $(\lambda_t)_{t\geqslant 0}$, there is a very important difference between them. In order to clarify this difference, let us say a few words about how we decompose the filtration $\mathbb{F}$. Let $\mathbb{H} = (\mathcal{H}_t)_{t\geqslant 0}$ be the filtration generated by the triplet $(N_t, P_t, \lambda_t)_{t\geqslant 0}$. The filtration $\mathbb{D} = (\mathcal{D}_t)_{t\geqslant 0}$ is defined as the natural filtration of $(D_t)_{t\geqslant 0}$. The filtration $\mathbb{F}$ is then defined as $\mathbb{F} = \mathbb{H} \vee \mathbb{D}$. We end this discussion about filtrations with an important assumption: conditional on $\mathbb{H}$, $(D_t)_{t\geqslant 0}$ is a non-homogeneous Poisson process, in particular for any $0 \leqslant s \leqslant t$ and $k \in \mathbb{N}$,

$$\mathbb{P}(D_s = k|\mathcal{H}_t) = \frac{\Lambda_s^k}{k!} e^{-\Lambda_s}$$

where $(\Lambda_t)_{t\geqslant 0} := (\int_0^t \lambda_u \mathrm{d}u)_{t\geqslant 0}$ is the compensator of $(D_t)_{t\geqslant 0}$. From this last assumption follows the fundamental difference between the processes $(D_t)_{t\geqslant 0}$ and $(N_t)_{t\geqslant 0}$: by definition of $(\lambda_t)_{t\geqslant 0}$, the jumps of $(N_t)_{t\geqslant 0}$ exactly correspond to the jumps of $(\lambda_t)_{t\geqslant 0}$, whereas it is absolutely not the case for $(D_t)_{t\geqslant 0}$. In other words, the jumps of $(N_t)_{t\geqslant 0}$ are observed in $(\lambda_t)_{t\geqslant 0}$, while the jumps of $(D_t)_{t\geqslant 0}$ are not. From an economic point of view, this means that the company whose default is modeled by the first jump of $(D_t)_{t\geqslant 0}$ is not systemic, but it is sensitive to defaults of systemic companies.

The first proposition is about the expectation and variance of the intensity $(\lambda_t)_{t\geqslant 0}$. It allows to derive some conditions on the parameters to ensure that the intensity process does not explode. The proof is in Appendix 1.

Proposition 2.1. *Assume that $\kappa > \frac{\eta}{\rho}$. For any $t \geqslant 0$, we have*

$$\mathbb{E}[\lambda_t] = e^{-a_1 t}\lambda_0 - \frac{\kappa\theta}{a_1}(e^{-a_1 t} - 1) \tag{2.2}$$

and

$$\mathbb{V}\mathrm{ar}(\lambda_t) = \frac{(1 - e^{-a_1 t})\,(2a_2\lambda_0 e^{-a_1 t} + a_3(1 - e^{-a_1 t}))}{2a_1} \tag{2.3}$$

where $a_1 := \kappa - \frac{\eta}{\rho}$, $a_2 := 2\left(\frac{\eta}{\rho}\right)^2$ and $a_3 := \kappa\theta\frac{a_2}{a_1}$.

From this result, we observe that the condition that ensures the non-explosion of the intensity is $\kappa > \frac{\eta}{\rho}$. That is, the speed of the mean reversion must be sufficient to prevent the process from exploding. If this condition is satisfied, we have the following long-term values for the expected value and variance

$$\lim_{t\to+\infty} \mathbb{E}[\lambda_t] = \frac{\kappa\theta}{\kappa - \frac{\eta}{\rho}}$$

$$\lim_{t\to+\infty} \mathbb{V}\mathrm{ar}(\lambda_t) = \kappa\theta\left(\frac{\eta}{\eta - \rho\kappa}\right)^2.$$

In the following, we denote

$$p(t, x_1, x_2|s, y_1, y_2) = \frac{\partial^2\mathbb{P}(\lambda_t \leqslant x_1, \Lambda_t \leqslant x_2|\lambda_s = y_1, \Lambda_s = y_2)}{\partial x_1 \partial x_2} \tag{2.4}$$

and

$$\phi(t, z_1, z_2|s, y_1, y_2) = \mathbb{E}\left[\exp\left\{-z_1\lambda_t - z_2\Lambda_t\right\}|\lambda_s = y_1, \Lambda_s = y_2\right]$$

where $t \geqslant s \geqslant 0$, $y_1 \geqslant \theta$, $y_2 \geqslant 0$ and $\Lambda_t := \int_0^t \lambda_{u-}\mathrm{d}u$. Note that p corresponds to the conditional PDF of (λ_t, Λ_t) given (λ_s, Λ_s) whereas ϕ is the corresponding Laplace transform. We may write $p(t, x_1, x_2|\cdot)$ and $\phi(t, x_1, x_2|\cdot)$ instead of respectively $p(t, x_1, x_2|s, y_1, y_2)$ and $\phi(t, z_1, z_2|s, y_1, y_2)$ to shorten the notations. The conditional PDF p satisfies a partial differential equation (PDE).

Proposition 2.2. *The joint PDF $p(t, x_1, x_2|s, y_1, y_2)$ of (λ_t, Λ_t) is solution of the Fokker-Planck equation*

$$\begin{aligned}\frac{\partial p(t, x_1, x_2|\cdot)}{\partial t} &= \kappa p(t, x_1, x_2|\cdot) - \frac{\partial p(t, x_1, x_2|\cdot)}{\partial x_2} x_1 \\ &- \frac{\partial p(t, x_1, x_2|\cdot)}{\partial x_1}\kappa(\theta - x_1) - \eta\mathbb{E}[\xi p(t, x_1 - \eta\xi, x_2|\cdot)] \\ &+ x_1\mathbb{E}[p(t, x_1 - \eta\xi, x_2|\cdot) - p(t, x_1, x_2|\cdot)]\end{aligned} \tag{2.5}$$

with boundary condition $p(s, x_1, x_2|s, y_1, y_2) = \delta_{\{x_1 - y_1, x_2 - y_2\}}$, where δ stands for the Dirac measure located at $(0, 0) \in \mathbb{R}^2$.

The proof can be found in Appendix 1. The next result is the fact that the Laplace transform of $p(t, z_1, z_2|s, y_1, y_2)$ also satisfies a PDE. The proof has also been relegated to Appendix 1 and relies essentially on PDE (2.5).

Proposition 2.3. *The Laplace transform $\phi(t, z_1, z_2|\cdot)$ of the joint PDF $p(t, x_1, x_2|\cdot)$ is solution of the following forward Kolmogorov equation*

$$\frac{\partial \phi(t, z_1, z_2|\cdot)}{\partial t} = -z_1\kappa\theta\phi(t, z_1, z_2|\cdot) + \frac{\partial \phi(t, z_1, z_2|\cdot)}{\partial z_1}\gamma(z_1, z_2)$$

where $\gamma(z_1, z_2) := 1 - \kappa z_1 - \mathbb{E}[e^{-z_1\eta\xi}] + z_2$. Moreover, the boundary conditions $\phi(s, z_1, z_2|s, y_1, y_2) = \exp\{-z_1 y_1 - z_2 y_2\}$ and $\phi(t, 0, 0|s, y_1, y_2) = 1$ hold.

The next section aims to extend the previous results to the time-changed framework.

2.2 Fractional Setting

This second section extends the model to the fractional case with the help of a change of time. This allows to obtain fractional Hawkes processes, as introduced in Hainaut (2020). We start this second section by describing the subordinator and some of its properties.

The Subordinator

We use a time-change with a subordinator that is the inverse of an α-stable Lévy process. Such processes are discussed in, e.g. the Chapter 3 of Sato (1999) and in

Janicki and Weron (1994) We briefly introduce this class of processes and derive some properties of their inverses.

Definition 2.1. Let $(U_t)_{t\geqslant 0}$ be a stochastic process and $\alpha \in (0,1)$. We say that the process $(U_t)_{t\geqslant 0}$ is an α-stable subordinator if it is an increasing Lévy process that satisfies

$$\mathbb{E}\left[e^{-uU_t}\right] = e^{-tu^{\alpha}}$$

for all $t \geqslant 0$.

Throughout all the upcoming developments, we assume that $(U_t)_{t\geqslant 0}$ is an α-stable subordinator. The inverse of $(U_t)_{t\geqslant 0}$ will be denoted by $(S_t)_{t\geqslant 0}$, and defined as

$$S_t := \inf\{s \geqslant 0 : U_s \geqslant t\}.$$

Figure 2.2 shows five simulated paths of $(U_t)_{t\geqslant 0}$ and the corresponding paths of $(S_t)_{t\geqslant 0}$. The algorithm used to obtain such simulations is from Magdziarz (2009, b) and is described in details in Appendix 2. The graphs are obtained with the fractional order $\alpha = 0.94151$ of Table 2.1 (fractional case). A property of interest in the paths of the process $(S_t)_{t\geqslant 0}$ is the presence of periods of time where these paths remain constant, as we can observe in Figure 2.2. The motionless periods correspond to the jumps of $(U_t)_{t\geqslant 0}$, which follows from the definition of $(S_t)_{t\geqslant 0}$. Figure 2.3 shows 50 simulations of survival probability curves. We call these processes survival probability because in an intensity model, the probability of survival until time t is $\mathbb{E}[\exp\{-\Lambda_t\}]$. The paths in the time-changed model are much more heterogeneous. In particular, the time-changed setting allows to model companies that do not encounter troubles during long periods of time, therefore having survival probabilities remaining constant during these periods. That property is absolutely not mimicable by a usual credit model, as we can see on the left graph of Figure 2.2. Indeed, the strictly positive intensity of these models necessarily imply decreasing survival probability curves.

The PDF of U_t will be denoted by $p_U(t,\tau) = \frac{\partial}{\partial\tau}\mathbb{P}(U_t \leqslant \tau)$ whereas the PDF of S_t will be denoted by $g(t,\tau) = \frac{\partial}{\partial\tau}\mathbb{P}(S_t \leqslant \tau)$. The relations between the functions p_U and g are described in the following result. Proofs can be found in Hainaut (2020).

Proposition 2.4. *The following relations hold for all* t, τ

(1) $S_t =^d \left(\frac{t}{U_1}\right)^{\alpha}$*, i.e. both random variables share the same probability distribution;*

(2) $p_U(\tau,t) = \tau^{-\frac{1}{\alpha}} p_U(1, \tau^{-\frac{1}{\alpha}} t)$*;*

(3) $\tau g(t,\tau) = \frac{t}{\alpha} p_U(\tau,t)$*;*

(4) $g(t,\tau) = -\frac{\partial}{\partial\tau}\int_0^t p_U(\tau,u)\mathrm{d}u.$

We now state a result about the Laplace transform of $(S_t)_{t\geqslant 0}$. It will be used later as a boundary condition.

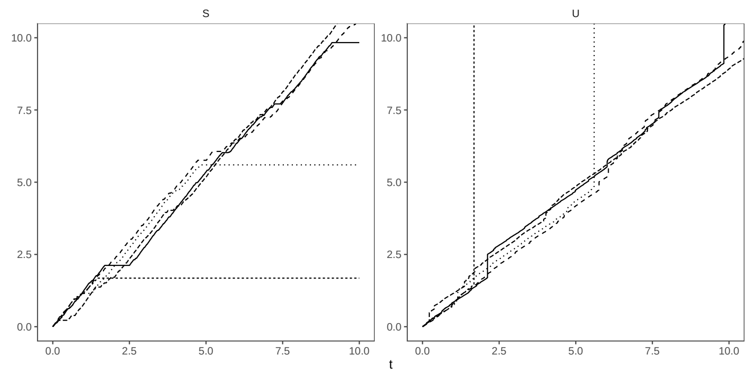

Figure 2.2: 5 Paths of $(U_t)_{t \geqslant 0}$ and Corresponding Paths of $(S_t)_{t \geqslant 0}$.

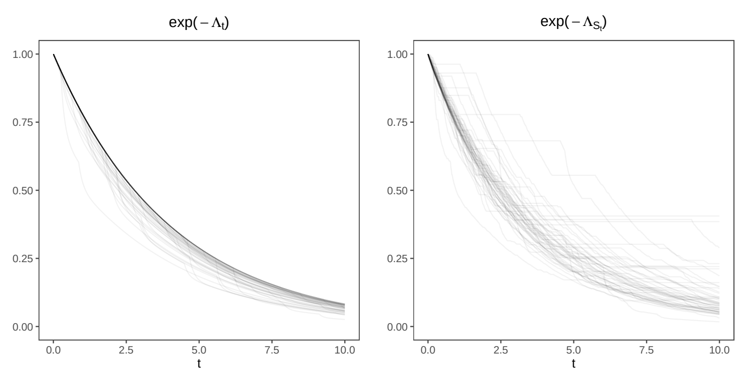

Figure 2.3: 50 Paths of $(\exp\{-\Lambda_t\})_{t \geqslant 0}$ and $(\exp\{-\Lambda_{S_t}\})_{t \geqslant 0}$.

Proposition 2.5. *The Laplace transform of* $(S_t)_{t\geqslant 0}$ *is given by*

$$\begin{aligned}\mathbb{E}[e^{-\omega S_t}] &:= \int_0^{+\infty} e^{-\omega\tau} g(t,\tau)\mathrm{d}\tau \\ &= \mathrm{E}_\alpha(-\omega t^\alpha)\end{aligned}$$

where E_α *denotes the Mittag-Leffler function*

$$\mathrm{E}_\alpha(z) := \sum_{k=0}^{+\infty} \frac{z^k}{\Gamma(k\alpha+1)}.$$

Moreover, by differentiation of the Laplace transform of S_t*, we have for any* $n \in \mathbb{N}$

$$\mathbb{E}[S_t^n] = \frac{n!t^{n\alpha}}{\Gamma(n\alpha+1)}.$$

The reader is referred to Piryatinska et al. (2004) for a proof.

Caputo Derivatives

The main difference between the results obtained in the fractional setting (i.e. time-changed setting) compared to the non-fractional one is the replacement of the partial derivative with respect to time by a fractional derivative in the PDE's, turning them into fractional PDE's (FPDE). In the following, we present the needed material about Caputo fractional derivatives. Textbooks about fractional calculus are, e.g. Podlubny (1999) and Diethelm (2010). We start with the definition of Caputo derivative.

Definition 2.2. For a differentiable function $h : \mathbb{R}^+ \to \mathbb{R}$, the Caputo derivative of h of order $\alpha \in (0,1)$ is defined as

$$\frac{\mathrm{d}^\alpha h}{\mathrm{d}t^\alpha}(t) := \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \frac{\mathrm{d}}{\mathrm{d}s} h(s)\mathrm{d}s. \tag{2.6}$$

Caputo derivatives are also defined for any non-integer order $\alpha > 1$. We give the definition for the $\alpha \in (0,1)$ case since it is the only one we need here. The interested reader may find a more geneal definition in the textbooks mentioned above. Note that all the values of the derivatives at times $s \in (0,t)$ are needed to compute the Caputo derivative at time t. The interpretation is that if we are interested in the change rate at time t, we need to take into account all the change rates at $s \in (0,t)$. From a modeling point of view, this introduces a memory effect. The following result will be useful regarding Caputo derivatives and Laplace transforms.

Proposition 2.6. *For any differentiable function* $h : \mathbb{R}^+ \to \mathbb{R}$*, it holds that*

$$\widetilde{\frac{\mathrm{d}^\alpha h}{\mathrm{d}t^\alpha}}(\omega) = \omega^\alpha \tilde{h}(\omega) - \omega^{\alpha-1} h(0), \tag{2.7}$$

where

$$\widetilde{\frac{\mathrm{d}^\alpha h}{\mathrm{d}t^\alpha}}(\omega) := \int_0^{+\infty} \frac{\mathrm{d}^\alpha h}{\mathrm{d}t^\alpha}(t) e^{-t\omega} \mathrm{d}t \textit{ and } \tilde{h}(\omega) := \int_0^{+\infty} h(s) e^{-s\omega} \mathrm{d}s.$$

This is a well known result in fractional calculus. A proof can be found in the chapter 2 of Podlubny (1999). We can now turn to the main results regarding the fractional setting. Those are exposed in the next subsection.

Fractional Fokker-Planck Equation for the Time-Changed Process

This subsection is concerned with the main theoretical results of this chapter, namely the FPDE for the joint PDF and joint Laplace transform of the time-changed process $(\lambda_{S_t}, \Lambda_{S_t})_{t\geqslant 0}$. Before proving that the joint PDF and Laplace transform satisfy FPDE's, we need a preliminary result about Laplace transforms. We denote

$$p_\alpha(t, x_1, x_2|y) := \frac{\partial^2 \mathbb{P}(\lambda_{S_t} \leqslant x_1, \Lambda_{S_t} \leqslant x_2 | \lambda_0 = y)}{\partial x_1 \partial x_2},$$

the joint PDF of $(\lambda_{S_t}, \Lambda_{S_t})_{t\geqslant 0}$. Also, recall that p is the joint PDF of $(\lambda_t, \Lambda_t)_{t\geqslant 0}$, as defined at Equation (2.4). The Laplace transform of this function satisfies the next lemma.

Lemma 2.1. *It holds that*

$$\tilde{p}_\alpha(\omega, x_1, x_2|y) = \omega^{\alpha-1} \tilde{p}(\omega^\alpha, x_1, x_2|y),$$

where

$$\tilde{p}_\alpha(\omega, x_1, x_2|y) := \int_0^{+\infty} e^{-\omega t} p_\alpha(t, x_1, x_2|y) \mathrm{d}t.$$

Proof. First note that

$$p_\alpha(t, x_1, x_2|y) = \int_0^{+\infty} p(\tau, x_1, x_2|0, y, 0) g(t, \tau) \mathrm{d}\tau,$$

where $g(t,\cdot)$ is the PDF of S_t. Therefore it follows that $\tilde{p}_\alpha$ satisfies

$$\begin{aligned}
\tilde{p}_\alpha(\omega,x_1,x_2|y) &:= \int_0^{+\infty} e^{-\omega t} p_\alpha(t,x_1,x_2|y)\mathrm{d}t \\
&= \int_0^{+\infty} e^{-t\omega} \int_0^{+\infty} p(\tau,x_1,x_2|0,y,0) g(t,\tau)\mathrm{d}\tau\mathrm{d}t \\
&= \int_0^{+\infty} p(\tau,x_1,x_2|0,y,0) \int_0^{+\infty} e^{-t\omega} g(t,\tau)\mathrm{d}t\mathrm{d}\tau \\
&= \omega^{\alpha-1} \int_0^{+\infty} e^{-\tau\omega^\alpha} p(\tau,x_1,x_2|0,y,0)\mathrm{d}\tau = \omega^{\alpha-1}\tilde{p}(\omega^\alpha,x_1,x_2|0,y,0),
\end{aligned}$$

as announced. □

We now have the tools to prove the first main theoretical result of this chapter, namely that the joint PDF p_α satisfies a FPDE.

Proposition 2.7. *The joint PDF p_α satisfies the FPDE*

$$\begin{aligned}
&\frac{\partial^\alpha p_\alpha(t,x_1,x_2|y)}{\partial t^\alpha} \\
&\quad = \kappa p_\alpha(t,x_1,x_2|y) - x_1 \frac{\partial}{\partial x_2} p_\alpha(t,x_1,x_2|y) - \kappa(\theta - x_1)\frac{\partial}{\partial x_1} p_\alpha(t,x_1,x_2|y) \\
&\quad - \eta\mathbb{E}[\xi p_\alpha(t,x_1-\eta\xi,x_2|y)] + x_1\mathbb{E}[p_\alpha(t,x_1-\eta\xi,x_2|y) - p_\alpha(t,x_1,x_2|y)]
\end{aligned}$$

with boundary condition $p_\alpha(s,x_1,x_2|y) = \delta_{\{x_1-y,x_2\}}$, where δ stands for the Dirac measure located at $(0,0)\in\mathbb{R}^2$.

Proof. In the following, $p(t,x_1,x_2|y)$ is used as a shorthand for $p(t,x_1,x_2|0,y,0)$. Firstly let us look at the integral

$$\int_0^{+\infty} \frac{\partial}{\partial t}\left(e^{-\omega t} p(t,x_1,x_2|y)\right)\mathrm{d}t. \tag{2.8}$$

From the fundamental theorem of calculus, it is clearly equal to $-p(0,x_1,x_2|y)$. However, by computing the derivative inside the integral, we obtain that the integral at Equation (2.8) is also equal to

$$-\omega\tilde{p}(\omega,x_1,x_2|y) + \int_0^{+\infty} e^{-\omega t}\frac{\partial}{\partial t} p(t,x_1,x_2|y)\mathrm{d}t$$

so that we infer

$$\omega\tilde{p}(\omega,x_1,x_2|y) - p(0,x_1,x_2|y) = \int_0^{+\infty} e^{-\omega t}\frac{\partial}{\partial t} p(t,x_1,x_2|y)\mathrm{d}t. \tag{2.9}$$

Secondly we use Proposition 2.2, i.e. the Fokker-Planck equation for p, on the right-hand side of Equation (2.9). After some simplifications, this gives

$$\begin{aligned}
&\omega\tilde{p}(\omega, x_1, x_2|y) - p(0, x_1, x_2|y) \\
&= - x_1 \frac{\partial}{\partial x_2}\tilde{p}(\omega, x_1, x_2|y) - \kappa(\theta - x_1)\frac{\partial}{\partial x_1}\tilde{p}(\omega, x_1, x_2|y) + \kappa\tilde{p}(\omega, x_1, x_2|y) \\
&\quad - \eta\mathbb{E}\left[\xi\tilde{p}(\omega, x_1 - \eta\xi, x_2|y)\right] + x_1\mathbb{E}\left[\tilde{p}(\omega, x_1 - \eta\xi, x_2|y) - \tilde{p}(\omega, x_1, x_2|y)\right].
\end{aligned} \tag{2.10}$$

Since this is valid for any ω, one can replace ω by ω^α in Equation (2.10). Doing this, multiplying each side by $\omega^{\alpha-1}$, and using the fact that $p_\alpha(0, x_1, x_2|y) = p(0, x_1, x_2|y) = 0$ provides us with

$$\begin{aligned}
&\omega^\alpha(\omega^{\alpha-1}\tilde{p}(\omega^\alpha, x_1, x_2|y)) - \omega^{\alpha-1}p_\alpha(0, x_1, x_2|y) \\
&= - x_1\frac{\partial}{\partial x_2}\left(\omega^{\alpha-1}\tilde{p}(\omega^\alpha, x_1, x_2|y)\right) - \kappa(\theta - x_1)\frac{\partial}{\partial x_1}\left(\omega^{\alpha-1}\tilde{p}(\omega^\alpha, x_1, x_2|y)\right) \\
&\quad + \kappa\omega^{\alpha-1}\tilde{p}(\omega^\alpha, x_1, x_2|y) - \eta\mathbb{E}\left[\xi\omega^{\alpha-1}\tilde{p}(\omega^\alpha, x_1 - \eta\xi, x_2|y)\right] \\
&\quad + x_1\mathbb{E}\left[\omega^{\alpha-1}\tilde{p}(\omega^\alpha, x_1 - \eta\xi, x_2|y) - \omega^{\alpha-1}\tilde{p}(\omega^\alpha, x_1, x_2|y)\right].
\end{aligned} \tag{2.11}$$

Using Lemma 2.1, we can rewrite Equation (2.11) as

$$\begin{aligned}
&\omega^\alpha\tilde{p}_\alpha(\omega, x_1, x_2|y) - \omega^{\alpha-1}p_\alpha(0, x_1, x_2|y) \\
&= - x_1\frac{\partial}{\partial x_2}\tilde{p}_\alpha(\omega, x_1, x_2|y) + -\kappa(\theta - x_1)\frac{\partial}{\partial x_1}\tilde{p}_\alpha(\omega, x_1, x_2|y) \\
&\quad + \kappa\tilde{p}_\alpha(\omega, x_1, x_2|y) - \eta\mathbb{E}\left[\xi\tilde{p}_\alpha(\omega, x_1 - \eta\xi, x_2|y)\mathrm{d}t\right] \\
&\quad + x_1\mathbb{E}\left[\omega^{\alpha-1}\tilde{p}_\alpha(\omega, x_1 - \eta\xi, x_2|y) - \tilde{p}_\alpha(\omega, x_1, x_2|y)\right].
\end{aligned} \tag{2.12}$$

An application of Proposition 2.6 to the function $t \in (0, +\infty) \mapsto p(t, x_1, x_2|y)$ allows to conclude that the left-hand side of Equation (2.12) equals $\widetilde{\frac{\partial^\alpha p}{\partial t^\alpha}}(\omega, x_1, x_2|y)$. $\square$

We also obtained a FPDE for the Laplace transform, that is the second main theoretical result of this chapter. In the following, the Laplace transform of $(S_t, \lambda_{S_t}, \Lambda_{S_t})_{t\geqslant 0}$ is denoted by

$$\phi_\alpha(t, z_1, z_2, z_3|y) := \mathbb{E}\left[\exp\left\{-z_1 S_t - z_2\lambda_{S_t} - z_3\Lambda_{S_t}\right\}|\lambda_0 = y\right].$$

Proposition 2.8. *The function ϕ_α satisfies the FPDE*

$$\begin{aligned}
\frac{\partial^\alpha \phi_\alpha}{\partial t^\alpha}(t, z_1, z_2, z_3|y) = &- z_1\phi_\alpha(t, z_1, z_2, z_3|y) - z_2\kappa\theta\phi_\alpha(t, z_1, z_2, z_3|y) \\
&+ \gamma(z_2, z_3)\frac{\partial\phi_\alpha}{\partial z_2}(t, z_1, z_2, z_3|y)
\end{aligned} \tag{2.13}$$

where $\gamma(z_2, z_3) := 1 - \kappa z_2 - \mathbb{E}[e^{-z_2 \eta \xi}] + z_3$. *Moreover, the following boundary conditions hold:*

$$\phi_\alpha(t, z_1, 0, 0|y) = \mathrm{E}_\alpha(-z_1 t^\alpha)$$
$$\phi_\alpha(0, z_1, z_2, z_3|y) - e^{-z_2 y} = \phi_\alpha(t, 0, 0, 0|y) - 1 = 0.$$

Proof. The first boundary condition follows from Proposition 2.5, the second and the third are straightforward given the definition of ϕ_α. In order to derive Equation (2.13), we start by observing the joint PDF p_α of the triple $(S_t, \lambda_{S_t}, \Lambda_{S_t})$, that is

$$p_\alpha(t, x_1, x_2, x_3|y) := \frac{\partial^3}{\partial x_1 \partial x_2 \partial x_3} \mathbb{P}\left(S_t \leqslant x_1, \lambda_{S_t} \leqslant x_2, \Lambda_{S_t} \leqslant x_3 | \lambda_0 = y\right).$$

From the definition of conditional PDF, we have

$$p_\alpha(t, x_1, x_2, x_3|y) = g(t, x_1) p(x_1, x_2, x_3|y). \tag{2.14}$$

See Equation (2.4) for the definition of p. From Equation (2.14), it follows that

$$\begin{aligned}
\tilde{p}_\alpha(\omega, x_1, x_2, x_3|y) &:= \int_0^{+\infty} e^{-\omega t} p_\alpha(t, x_1, x_2, x_3|y) \mathrm{d}t \\
&= p(x_1, x_2, x_3|y) \int_0^{+\infty} e^{-\omega t} g(t, x_1) \mathrm{d}t \\
&= \omega^{\alpha-1} e^{-x_1 \omega^\alpha} p(x_1, x_2, x_3|y).
\end{aligned}$$

Thanks to this result we can obtain

$$\begin{aligned}
&\tilde{\phi}_\alpha(\omega, z_1, z_2, z_3|y) \\
&:= \int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty e^{-\omega t - z_1 x_1 - z_2 x_2 - z_3 x_3} p_\alpha(t, x_1, x_2, x_3|y) \mathrm{d}t \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \\
&= \omega^{\alpha-1} \int_0^\infty \int_0^\infty \int_0^\infty e^{-z_1 x_1 - z_2 x_2 - z_3 x_3} e^{-x_1 \omega^\alpha} p(x_1, x_2, x_3|y) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \\
&= \omega^{\alpha-1} \int_0^{+\infty} e^{-(z_1 + \omega^\alpha) x_1} \phi(x_1, z_2, z_2|y) \mathrm{d}x_1,
\end{aligned} \tag{2.15}$$

where $\phi(x_1, z_2, z_3|y) = \int_0^{+\infty} \int_0^{+\infty} e^{-x_2 z_2 - x_3 z_3} p(x_1, x_2, x_3|y) \mathrm{d}x_2 \mathrm{d}x_3$. Next, an integration by parts yields

$$\begin{aligned}
&\tilde{\phi}_\alpha(\omega, z_1, z_2, z_3|y) \\
&\quad = \frac{\omega^{\alpha-1}}{z_1 + \omega^\alpha} \left[\phi(0, z_2, z_3|y) + \int_0^{+\infty} e^{-(z_1 + \omega^\alpha) t} \frac{\partial \phi}{\partial t}(t, z_1, z_2|y) \mathrm{d}t \right].
\end{aligned} \tag{2.16}$$

Proposition 2.3 gives an expression of $\frac{\partial\phi}{\partial t}(t, z_1, z_2|y)$ that we can use to obtain

$$\begin{aligned}\int_0^{+\infty} e^{-(z_1+\omega^\alpha)t}\frac{\partial\phi}{\partial t}(t, z_1, z_2|y)\mathrm{d}t = -z_2\kappa\theta\int_0^{+\infty} e^{-(z_1+\omega^\alpha)t}\phi(t, z_2, z_3|y)\mathrm{d}t \\ + \gamma(z_2, z_3)\int_0^{+\infty} e^{-(z_1+\omega^\alpha)t}\frac{\partial\phi}{\partial z_2}(t, z_2, z_3|y)\mathrm{d}t.\end{aligned} \tag{2.17}$$

By this last equation, as well as Equation (2.15), Equation (2.16) is equivalent to

$$\begin{aligned}(z_1 + \omega\alpha)\tilde{\phi}_\alpha(\omega, z_1, z_2, z_3|y) = \omega^{\alpha-1}\phi(0, z_1, z_2|y) - z_2\kappa\theta\tilde{\phi}_\alpha(\omega, z_1, z_2, z_3|y) \\ + \gamma(z_2, z_3)\frac{\partial\tilde{\phi}_\alpha}{\partial z_2}(\omega, z_1, z_2, z_3|y).\end{aligned} \tag{2.18}$$

Observe that $\phi(0, z_2, z_3|y) = e^{-z_2 y} = \phi_\alpha(0, z_1, z_2, z_3|y)$, so that we obtain

$$\begin{aligned}&\omega^\alpha\tilde{\phi}_\alpha(\omega, z_1, z_2, z_3|y) - \omega^{\alpha-1}\phi_\alpha(0, z_1, z_2, z_3|y) \\ &= - z_1\tilde{\phi}_\alpha(\omega, z_1, z_2, z_3|y) - z_2\kappa\theta\tilde{\phi}_\alpha(\omega, z_1, z_2, z_3|y) + \gamma(z_2, z_3)\frac{\partial}{\partial z_2}\tilde{\phi}_\alpha(\omega, z_1, z_2, z_3|y).\end{aligned}$$

The announced result then follows from Proposition 2.6. □

Corollary 2.1.

$$\phi_\alpha(t, z_1, z_2|y) := \mathbb{E}\left[e^{-z_1\lambda_{S_t} - z_2\int_0^{S_t}\lambda_u\mathrm{d}u}\big|\lambda_0 = y\right],$$

then it holds that

$$\frac{\partial^\alpha\phi_\alpha}{\partial t^\alpha}(t, z_1, z_2|y) = \phi_\alpha(t, z_1, z_2|y)\beta(z_1) + \frac{\partial}{\partial z_1}\phi_\alpha(t, z_1, z_2|y)\gamma(z_1, z_2) \tag{2.19}$$

Moreover, the boundary conditions $\phi_\alpha(t, 0, 0|y) = 1$ and $\phi_\alpha(0, z_1, z_2|y) = e^{-z_1 y}$ are satisfied.

Proof. Apply Proposition 2.8 with $z_1 = 0$. □

Note that this last PDE is nearly the same as in the non-fractional case, see Proposition 2.3. The only difference is that the partial derivative with respect to time is replaced by a fractional Caputo derivative.

2.3 Numerical Illustration

This section shows how we can use the FPDE of the Laplace transform ϕ_α to compute survival probabilities. Several methods have been developed recently to solve FPDE's numerically. We can cite, e.g. the variational iteration method in Odibat and Momani (2009), the homotopy perturbation method in Golbabai and Sayevand (2011), the Laplace transform method in Jafari et al. (2013), the Haar-wavelet method in Wang et al. (2014), the Adomian decomposition method in Jafari and Gejji (2006) and El-Sayed et al. (2010) and the spectral regularization method in Zheng and Wei (2010). In our case, the FPDE is solved numerically with the help of a finite difference method. Finite difference methods are of common use in finance for solving PDE's, see Zvan et al. (2000) and Kim et al. (2016) for examples and Duffy (2006) for a textbook. They have also been successfully used to solve FPDE's in, e.g. Murio (2008), fei Ding and xin Zhang (2011) and in Song and Weiguo (2013), where it is used to price European and American options in a fractional version of the Black and Scholes model. The finite difference approximation we use for the Caputo derivative is the same as in Murio (2008) and fei Ding and xin Zhang (2011). However, our numerical method differs in two aspects. Firstly the equation we solve is not the same as in the mentioned references. Secondly our method is adapted to solve a FPDE with very few boundary conditions, as explained later.

In our model, the time of default τ is the first jump of the time changed point process $(D_{S_t})_{t\geqslant 0}$, i.e.

$$\tau := \inf\{t \geqslant 0 | D_{S_t} > 0\}.$$

The process $(D_t)_{t\geqslant 0}$ is assumed to have intensity $(\lambda_t)_{t\geqslant 0}$. It follows that

$$\mathbb{P}(\tau \geqslant t|\mathcal{F}_0) = \mathbb{E}\left[\exp\left\{-\Lambda_{S_t}\right\}|\mathcal{F}_0\right] = \phi_\alpha(t,0,1|y).$$

From Corollary 2.1, we know that the function $(t,z) \mapsto \phi_\alpha(t,z,1|y)$ satisfies the FPDE

$$\frac{\partial^\alpha \phi_\alpha}{\partial t^\alpha}(t,z,1|y) = -z\kappa\theta\phi_\alpha(t,z,1|y) + \frac{\partial}{\partial z}\phi_\alpha(t,z,1|y)\gamma(z,1)$$

with the boundary condition $\phi_\alpha(0,z,1|y) = e^{-zy}$. We start by solving the simple non-fractional case ($\alpha = 1$) and then extend the method to the fractional case with $\alpha \in (0,1)$.

Non-Fractional Case

Recall that

$$\phi(t,z_1,z_2|s,y_1,y_2) = \mathbb{E}\left[\exp\left\{-z_1\lambda_t - z_2\Lambda_t\right\}|\lambda_s = y_1, \Lambda_s = y_2\right].$$

For the sake of shorter notations, we will denote $\phi(t,z_1,z_2)=\phi(t,z_1,z_2|0,y,0)$, for a fixed $y \geqslant \theta$. Recall that $\gamma(z_1,z_2)=1-\kappa z_1-\mathbb{E}[e^{-z_1\eta\xi}]+z_2$. The goal is to solve the PDE

$$\frac{\partial\phi}{\partial t}(t,z_1,z_2)=-z_1\kappa\theta\phi(t,z_1,z_2)+\frac{\partial\phi}{\partial z_1}\gamma(z_1,z_2) \tag{2.20}$$

for a fixed z_2. Then, setting $z_2=1$ allows to find the survival probabilities implied by the model. In order to solve numerically this equation, we will work on two grids. The first one is $\{t_0,t_1,\dots,t_{n_t}\}$, where $t_k=k\Delta_t$. The second one is $\{z_{-n_z},z_{-n_z+1},\dots,z_{-1},z_0,z_1,\dots,z_{n_z}\}$, where $z_j=j\Delta_z$. We write respectively $\hat{\phi}(k,j;z_2)$, $\frac{\partial\hat{\phi}}{\partial z_1}(k,j;z_2)$ and $\frac{\partial\hat{\phi}}{\partial z_1}(k,j;z_2)$ for our approximations of $\phi(k\Delta_t,j\Delta_z,z_2)$, $\frac{\partial\phi}{\partial t}(k\Delta_t,j\Delta_z,z_2)$ and $\frac{\partial\phi}{\partial z_1}(k\Delta_t,j\Delta_z,z_2)$. The partial derivatives are approximated as

$$\frac{\partial\hat{\phi}}{\partial t}(k,j;z_2):=\frac{\hat{\phi}(k,j;z_2)-\hat{\phi}(k-1,j;z_2)}{\Delta_t} \tag{2.21}$$

$$\frac{\partial\hat{\phi}}{\partial z_1}(k,j;z_2):=\begin{cases}\frac{\hat{\phi}(k,-n_z+1;z_2)-\hat{\phi}(k,-n_z;z_2)}{\Delta_z} & j=-n_z\\ \frac{\hat{\phi}(k,n_z;z_2)-\hat{\phi}(k,n_z-1;z_2)}{\Delta_z} & j=n_z\\ \frac{\hat{\phi}(k,j+1;z_2)-\hat{\phi}(k,j-1;z_2)}{2\Delta_z} & \text{otherwise.}\end{cases} \tag{2.22}$$

The special aspect of our method lies in Equation (2.22). Approximating the partial derivative with respect to z_1 in such a way, i.e. changing the approximation when $j\in\{-n_z,n_z\}$, allows us to obtain a solution even if the only boundary condition we know is $\phi(0,z_1,z_2|y)=e^{-z_1 y}$. Using these notations, PDE (2.20) translates into

$$\frac{\partial\hat{\phi}}{\partial t}(k,j;z_2)=\hat{\phi}(k,j;z_2)\beta_j+\frac{\partial\hat{\phi}}{\partial z_1}(k,j;z_2)\gamma_j(z_2), \tag{2.23}$$

where $\beta_j=-j\Delta_z\kappa\theta$ and $\gamma_j(z_2)=\gamma(j\Delta_z,z_2)$. Using Equations (2.21) and (2.22), we find

$$\hat{\phi}(k-1,j;z_2)=$$
$$\begin{cases}\left.\begin{array}{l}\left[1-\Delta_t\beta_{-n_z}+\frac{\Delta_t}{\Delta_z}\gamma_{-n_z}(z_2)\right]\hat{\phi}(k,-n_z;z_2)\\ -\frac{\Delta_t}{\Delta_z}\gamma_{-n_z}(z_2)\hat{\phi}(k,-n_z+1)\end{array}\right\} & j=-n_z\\ \left.\begin{array}{l}\left[1-\Delta_t\beta_{n_z}-\frac{\Delta_t}{\Delta_z}\gamma_{n_z}(z_2)\right]\hat{\phi}(k,n_z;z_2)\\ +\frac{\Delta_t}{\Delta_z}\gamma_{n_z}(z_2)\hat{\phi}(k,n_z-1+1)\end{array}\right\} & j=n_z\\ \left.\begin{array}{l}\frac{\Delta_t}{2\Delta_z}\gamma_j(z_2)(\hat{\phi}(k,j-1;z_2)-\hat{\phi}(k,j+1;z_2))\\ +\left[1-\Delta_t\beta_j\right]\hat{\phi}(k,j;z_2)\end{array}\right\} & j\notin\{-n_z,n_z\}\end{cases} \tag{2.24}$$

Equation (2.24) is then nothing more than a linear system of equations. Let us denote

$$\hat{\mathbf{\Phi}}_k(z_2) := \begin{pmatrix} \hat{\phi}(k, -n_z; z_2) \\ \hat{\phi}(k, -n_z + 1; z_2) \\ \vdots \\ \hat{\phi}(k, n_z; z_2) \end{pmatrix} \in \mathbb{R}^{2n_z+1}$$

and the matrix $\mathbf{A}$ by

$$\begin{pmatrix} b_{-n_z} + a_{-n_z} & -a_{-n_z} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \frac{a_{-n_z+1}}{2} & b_{-n_z+1} & \frac{-a_{-n_z+1}}{2} & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & \frac{a_{-n_z+2}}{2} & b_{-n_z+2} & \frac{-a_{-n_z+2}}{2} & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & \frac{a_{-n_z-1}}{2} & b_{n_z-1} & \frac{-a_{n_z-1}}{2} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & a_{n_z} & b_{n_z} - a_{n_z} \end{pmatrix}$$

where[1] $a_j := (\Delta_t/\Delta_z)\gamma_j(z_2)$ and $b_j := 1 - \Delta_t\beta_j$. Note that $\mathbf{A}(z_2) \in \mathbb{R}^{(2n_z+1)\times(2n_z+1)}$ for any z_2. Equation (2.24) then becomes

$$\hat{\mathbf{\Phi}}_{k-1}(z_2) = \mathbf{A}(z_2)\hat{\mathbf{\Phi}}_k(z_2),$$

so that given the (approximated) values $\hat{\mathbf{\Phi}}_{k-1}(z_2)$ of the function ϕ at time $(k-1)\Delta_t$, our problem reduces to inverting the matrix $\mathbf{A}(z_2)$.

From the boundary conditions of the PDE in Proposition 2.3, we know that for any z_1, z_2, $\phi(0, z_1, z_2|0, y_1, 0) = \exp\{-z_1 y_1\}$. It follows that, starting from the initial condition $\hat{\phi}(0, j; z_2) = \exp\{-j\Delta_z y_1\}$, i.e. the values in the vector $\hat{\mathbf{\Phi}}_0(z_2)$, we can determine all the vectors $\hat{\mathbf{\Phi}}_k(z_2)$ recursively using $\hat{\mathbf{\Phi}}_k(z_2) = (\mathbf{A}(z_2))^{-1}\hat{\mathbf{\Phi}}_{k-1}(z_2)$.

Approximating Caputo Fractional Derivatives

This subsection discusses the numerical approximation of the Caputo derivative of a function. Let us consider the grid $\{t_0, t_1, \dots, t_{n_t}\}$ of the previous section, that is $t_k = k\Delta_t$ for all k. The approximation formula is based on Equation (2.6) and the

[1] Obviously we should write $a_j(z_2)$ and $\mathbf{A}(z_2)$ but the notations needed to be shortened here.

following development

$$\begin{aligned}\frac{\mathrm{d}^\alpha h}{\mathrm{d}t^\alpha}(t_k) &= \frac{1}{\Gamma(1-\alpha)}\int_0^{t_k}(t_k-s)^{-\alpha}\frac{\mathrm{d}}{\mathrm{d}s}h(s)\mathrm{d}s\\ &= \frac{1}{\Gamma(1-\alpha)}\sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}}(t_k-s)^{-\alpha}\frac{\mathrm{d}}{\mathrm{d}s}h(s)\mathrm{d}s\\ &\approx \frac{1}{\Gamma(1-\alpha)}\sum_{i=0}^{k-1}\left(\frac{h(t_{i+1})-h(t_i)}{\Delta_t}\right)\int_{t_i}^{t_{i+1}}(t_k-s)^{-\alpha}\mathrm{d}s\\ &= \frac{-\Delta_t^{-\alpha}}{\Gamma(2-\alpha)}\sum_{i=0}^{k-1}(h(t_{i+1})-h(t_i))\left[(k-i-1)^{1-\alpha}-(k-i)^{1-\alpha}\right].\end{aligned}$$

Therefore the approximation formula for the Caputo fractional derivative is

$$\frac{\mathrm{d}^\alpha h}{\mathrm{d}t^\alpha}(t_k) \approx \frac{-\Delta_t^{-\alpha}}{\Gamma(2-\alpha)}\sum_{i=0}^{k-1}(h(t_{i+1})-h(t_i))\left[(k-i-1)^{1-\alpha}-(k-i)^{1-\alpha}\right]. \tag{2.25}$$

In the next subsection, we use this approximation to solve numerically the FPDE (2.19) of Corollary 2.1.

Fractional Case

We describe a numerical scheme to solve the FPDE of Corollary 2.1. The method is basically the same as in the non-fractional case, although it is a bit complicated by the fractional derivative with respect to time. We work with the same grids as in the non-fractional case. Our approximations of $\phi_\alpha(k\Delta_t, j\Delta_z, z_2)$, $\frac{\partial \phi_\alpha}{\partial z_1}(k\Delta_t, j\Delta_z, z_2)$, and $\frac{\partial^\alpha \phi_\alpha}{\partial t^\alpha}(k\Delta_t, j\Delta_z, z_2)$ are respectively denoted by $\hat{\phi}_\alpha(k, j; z_2)$, $\frac{\partial \hat{\phi}_\alpha}{\partial z_1}(k, j; z_2)$ and $\frac{\partial^\alpha \hat{\phi}_\alpha}{\partial t^\alpha}(k, j; z_2)$. The partial derivative with respect to z_1 is approximated with the formula we have seen in the non-fractional case, that is

$$\frac{\partial \hat{\phi}_\alpha}{\partial z_1}(k, j; z_2) := \begin{cases} \frac{\hat{\phi}_\alpha(k,-n_z+1;z_2)-\hat{\phi}_\alpha(k,-n_z;z_2)}{\Delta_z} & j=-n_z\\ \frac{\hat{\phi}_\alpha(k,n_z;z_2)-\hat{\phi}_\alpha(k,n_z-1;z_2)}{\Delta_z} & j=n_z\\ \frac{\hat{\phi}_\alpha(k,j+1;z_2)-\hat{\phi}_\alpha(k,j-1;z_2)}{2\Delta_z} & \text{otherwise.}\end{cases} \tag{2.26}$$

As discussed above, the Caputo derivative is approximated by

$$\frac{\partial^\alpha \hat{\phi}_\alpha}{\partial t^\alpha}(k,j;z_2) := \frac{-\Delta_t^{-\alpha}}{\Gamma(2-\alpha)} \sum_{i=0}^{k-1} [(k-i-1)^{1-\alpha} - (k-i)^{1-\alpha}](\hat{\phi}_\alpha(i+1,j;z_2) - \hat{\phi}_\alpha(i,j;z_2)). \tag{2.27}$$

Having introduced these notations, the FPDE in Corollary 2.1 translates into

$$\frac{\partial^\alpha \hat{\phi}_\alpha}{\partial t^\alpha}(k,j;z_2) = \hat{\phi}_\alpha(k,j;z_2)\beta_j + \frac{\partial \hat{\phi}_\alpha}{\partial z_1}(k,j;z_2)\gamma_j(z_2) \tag{2.28}$$

where β_j and γ_j are the same as in Equation (2.23). Starting from Equation (2.28), we can use approximations of Equations (2.27) and (2.26) to obtain

$$\begin{aligned}
&\frac{\Delta_t^{-\alpha}}{\Gamma(2-\alpha)}\Bigg[\left((k-1)^{1-\alpha} - k^{1-\alpha}\right)\hat{\phi}_\alpha(0,j;z_2) \\
&-\sum_{i=1}^{k-1} \hat{\phi}_\alpha(i,j;z_2)\left(2(k-i)^{1-\alpha} - (k-i+1)^{1-\alpha} - (k-i-1)^{1-\alpha}\right)\Bigg] \\
&= \begin{cases}
\left.\begin{array}{l}\hat{\phi}_\alpha(k,-n_z+1;z_2)\dfrac{\gamma_{-n_z}(z_2)}{\Delta_z} \\ \quad + \hat{\phi}_\alpha(k,-n_z;z_2)\left[\dfrac{-\Delta_t^{-\alpha}}{\Gamma(2-\alpha)} + \beta_{-n_z} - \dfrac{\gamma_{-n_z}(z_2)}{\Delta_z}\right]\end{array}\right\} & j=-n_z \\
\left.\begin{array}{l}\hat{\phi}_\alpha(k,n_z-1;z_2)\dfrac{-\gamma_{n_z}(z_2)}{\Delta_z} \\ \quad + \hat{\phi}_\alpha(k,n_z;z_2)\left[\dfrac{-\Delta_t^{-\alpha}}{\Gamma(2-\alpha)} + \beta_{n_z} + \dfrac{\gamma_{n_z}(z_2)}{\Delta_z}\right]\end{array}\right\} & j=n_z \\
\left.\begin{array}{l}\hat{\phi}_\alpha(k,j-1;z_2)\dfrac{-\gamma_j(z_2)}{2\Delta_z} + \hat{\phi}_\alpha(k,j+1;z_2)\dfrac{\gamma_j(z_2)}{2\Delta_z} \\ \quad + \hat{\phi}_\alpha(k,j;z_2)\left[\dfrac{-\Delta_t^{-\alpha}}{\Gamma(2-\alpha)} + \beta_j\right]\end{array}\right\} & \text{otherwise.}
\end{cases}
\end{aligned}$$

This is again a linear system of equations. Using vector and matrix notations, we have

$$\begin{aligned}
&\frac{\Delta_t^{-\alpha}}{\Gamma(2-\alpha)}\Bigg[\left((k-1)^{1-\alpha} - k^{1-\alpha}\right)\hat{\mathbf{\Phi}}_\alpha(0;z_2) \\
&-\sum_{i=1}^{k-1} \hat{\mathbf{\Phi}}_\alpha(i;z_2)\left(2(k-i)^{1-\alpha} - (k-i+1)^{1-\alpha} - (k-i-1)^{1-\alpha}\right)\Bigg] \\
&= \mathbf{A}_\alpha(z_2)\hat{\mathbf{\Phi}}_\alpha(k;z_2)
\end{aligned}$$

where[2] $\mathbf{A}_\alpha$ is

$$\begin{pmatrix} b_{-n_z} - a_{-n_z} & a_{-n_z} & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \frac{-a_{-n_z+1}}{2} & b_{-n_z+1} & \frac{a_{-n_z+1}}{2} & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \frac{-a_{-n_z+2}}{2} & b_{-n_z+2} & \frac{a_{-n_z+2}}{2} & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \frac{-a_{-n_z-1}}{2} & b_{n_z-1} & \frac{a_{n_z-1}}{2} \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & -a_{n_z} & b_{n_z} + a_{n_z} \end{pmatrix}$$

and

$$a_j(z_2) := \frac{\gamma_j(z_2)}{\Delta_z} \qquad b_j := \frac{-\Delta_t^{-\alpha}}{\Gamma(2-\alpha)} + \beta_j.$$

As in the non-fractional case, starting from the boundary condition $\hat{\phi}_\alpha(0, j; z_2) = \exp\{-j\Delta_z y_1\}$, we can obtain all the values on the grid using the recursion

$$\hat{\mathbf{\Phi}}_\alpha(k; z_2) = (\mathbf{A}_\alpha(z_2))^{-1} \frac{\Delta_t^{-\alpha}}{\Gamma(2-\alpha)} \Bigg[\left((k-1)^{1-\alpha} - k^{1-\alpha}\right) \hat{\mathbf{\Phi}}_\alpha(0; z_2) \\ - \sum_{i=1}^{k-1} \hat{\mathbf{\Phi}}_\alpha(i; z_2) \left(2(k-i)^{1-\alpha} - (k-i+1)^{1-\alpha} - (k-i-1)^{1-\alpha}\right) \Bigg].$$

In the next subsection, we find lower and upper bounds on the survival probabilities. This allows us to ensure the precision of our numerical procedure.

Bounds on the Survival Probabilities

The bounds on the probabilities in the non-fractional model are given in the next proposition.

Proposition 2.9. *Let $q : \mathbb{R}^+ \to [0,1], t \mapsto q(t) := \phi(t, 0, 1|0, y, 0)$ be the survival probability function implied by the non-fractional model (see Proposition 2.3) and $f_1, f_2 : \mathbb{R}^+ \to [0,1]$ be defined as*

$$f_1(t) := \exp\left\{ -\left(\left(y - \frac{\kappa\theta}{\kappa - \eta\rho^{-1}} \right) \frac{1 - e^{-(\kappa - \eta\rho^{-1})t}}{\kappa - \eta\rho^{-1}} + \frac{\kappa\theta t}{\kappa - \eta\rho^{-1}} \right) \right\}$$

and $f_2(t) := e^{-\theta t}$. Then it holds that $f_1 \leqslant q \leqslant f_2$.

Proof. The inequality $q \leqslant f_2$ follows from $\lambda_t \geqslant \theta$. For the other inequality, note that the convexity of the function $x \mapsto e^{-x}$ implies, by Jensen's inequality

$$q(t) = \mathbb{E}[\exp\{-\Lambda_t\}|\lambda_0 = y] \geqslant \exp\{-\mathbb{E}[\Lambda_t|\lambda_0 = y]\}.$$

The expression of f_1 is then obtained by integrating the expected value of λ_t, see Equation (2.2). □

[2] Again we should write $a_j(z_2)$ and $\mathbf{A}_\alpha(z_2)$ but shorter notations are needed.

We also have bounds in the fractional case.

Proposition 2.10. *Let* $q_\alpha : \mathbb{R}^+ \to [0,1], t \mapsto q_\alpha(t) := \phi_\alpha(t,0,1|y)$ *be the survival probability function implied by the fractional model (see Corollary 2.1) and* $f_{\alpha,1}, f_{\alpha,2} : \mathbb{R}^+ \to [0,1]$ *be defined as*

$$f_{\alpha,1}(t) := \exp\left\{-\left(\left(y - \frac{\kappa\theta}{\kappa - \eta\rho^{-1}}\right)\frac{1 - \mathrm{E}_\alpha(-(\kappa - \eta\rho^{-1})t^\alpha)}{\kappa - \eta\rho^{-1}} + \frac{\kappa\theta t^\alpha}{(\kappa - \eta\rho^{-1})\Gamma(\alpha+1)}\right)\right\}$$

and $f_{\alpha,2}(t) := E_\alpha(-\theta t^\alpha)$*, where* E_α *is the Mittag-Leffler function. Then it holds that* $f_{\alpha,1} \leqslant q_\alpha \leqslant f_{\alpha,2}$.

Proof. The inequality $q_\alpha \leqslant f_{\alpha,2}$ follows from $\Lambda_{S_t} \geqslant \theta S_t$ and Proposition 2.5. The other inequality follows from Jensen's inequality

$$\begin{aligned} q_\alpha(t) &\geqslant \exp\{-\mathbb{E}[\Lambda_{S_t}|\lambda_0 = y]\} \\ &= \exp\left\{-\mathbb{E}\left[\left(y - \frac{\kappa\theta}{\kappa - \eta\rho^{-1}}\right)\frac{1 - e^{-(\kappa - \eta\rho^{-1})S_t}}{\kappa - \eta\rho^{-1}} + \frac{\kappa\theta S_t}{\kappa - \eta\rho^{-1}}\right]\right\} \end{aligned}$$

where we used Equation (2.2). The proof is then easily completed with the help of Proposition 2.5. □

These two Propositions will be useful in the next section to show the accuracy of the numerical method.

Numerical Results

This section is devoted to the calibrations of both models (fractional and non-fractional) to real market data. The first data have been taken on the 25th of January 2021 from Bloomberg. They consist of survival probabilities for American Airlines. These data are displayed in the *Market* column of Table 2.2. The model is calibrated by choosing the set of parameters that minimizes the quantity

$$\sum_{T=1}^{10} (Prob_T^{\mathrm{Model}} - Prob_T^{\mathrm{Market}})^2$$

where $Prob_T^{\mathrm{Market}}$ denotes the market survival up to time T probability from Bloomberg and $Prob_T^{\mathrm{Model}}$ denotes the corresponding survival probability implied by the model. Referring to the notations from above, the computations are made with $\Delta_t = 0.002$, $n_t = 5000$, $\Delta_z = 0.01$, $n_z = 10$ and λ_0 was fixed at the target level θ. The parameters from the calibration are shown in Table 2.1. The graphical results of the calibrations are shown at Figure 2.4 and the numerical results are given in Table 2.2. In this table, we compare the market data to the survival

Table 2.1: Parameters of the Calibrated Model.

	θ	κ	η	ρ	α
Non-fractional	0.2382	37.6009	1.7284	2.5861	1.0000
Fractional	0.2489	6.6832	1.7284	2.5861	0.9415

Table 2.2: Numerical Results of the Calibration and Bounds on the Computed Probabilities. M stands for "Market" and corresponds to the market data.

		Non-fractional				Fractional			
T	M	f_1	q	f_2	$\lvert M-q\rvert$	$f_{\alpha,1}$	q_α	$f_{\alpha,2}$	$\lvert M-q_\alpha\rvert$
1	0.7954	0.7848	0.7849	0.7881	0.0105	0.7568	0.7604	0.7766	0.0350
2	0.5673	0.6158	0.6161	0.6210	0.0488	0.5831	0.5919	0.6176	0.0246
3	0.4576	0.4832	0.4835	0.4894	0.0259	0.4529	0.4664	0.4965	0.0088
4	0.3633	0.3791	0.3795	0.3857	0.0162	0.3535	0.3711	0.4024	0.0078
5	0.3011	0.2975	0.2978	0.3039	0.0033	0.2769	0.2977	0.3285	0.0034
6	0.2492	0.2334	0.2338	0.2395	0.0154	0.2175	0.2408	0.2700	0.0084
7	0.2032	0.1832	0.1835	0.1888	0.0197	0.1713	0.1963	0.2234	0.0069
8	0.1661	0.1437	0.1440	0.1488	0.0221	0.1351	0.1613	0.1860	0.0048
9	0.1339	0.1128	0.1130	0.1172	0.0209	0.1068	0.1335	0.1559	0.0004
10	0.1061	0.0885	0.0887	0.0924	0.0174	0.0845	0.1115	0.1315	0.0054

probabilities implied by the model and computed by solving the (F)PDE with the method described above. As we can observe on Figure 2.4, the non-fractional model is unable to properly fit the market data. Moving to the fractional model allows to introduce a greater degree of convexity and thus a better fit. To illustrate numerically the improvement of the fit, the absolute differences with respect to the market survival probabilities can also be found in Table 2.2. These differences between probabilities sum to 0.2002 in the non-fractional case whereas they sum to only 0.1055 in the fractional case. Thus the fractional model performs almost twice better according to the sum of absolute values criterion.

Table 2.2 displays the lower and upper bounds on the survival probabilities given at Propositions 2.9 and 2.10. Recall that f_1 and f_2 are respectively the lower and upper bounds in the non-fractional case and $f_{\alpha,1}$ and $f_{\alpha,2}$ are respectively the lower and upper bounds in the fractional case. Since the numerical method leads to values that are always between the lower and the upper bounds, we can deduce bounds on the error of the numerical results. Note that these bounds do not depend on the discretization parameters of the (F)PDE, but only on the parameters of the model. As a consequence, they do not prove the convergence of the numerical method. However, the convergence can be observed numerically, as shown in Table 2.3.

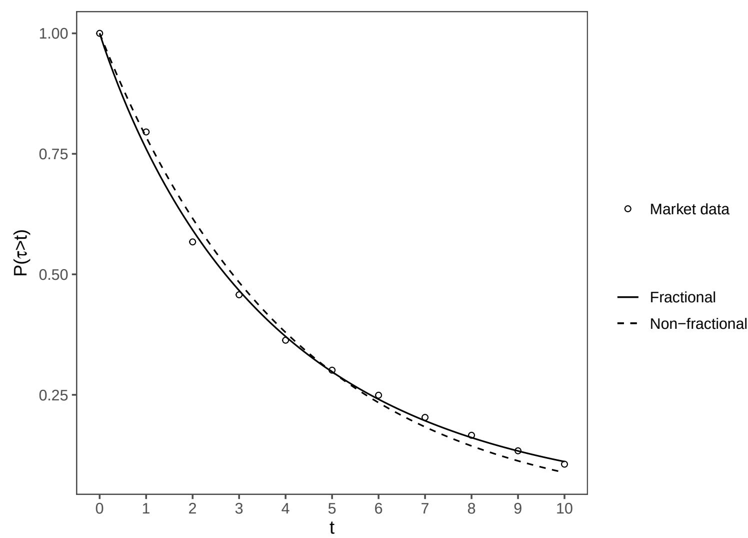

Figure 2.4: Market Survival Probabilities and Calibrated Survival Probabilities.

Table 2.3: Convergence of the Survival Probabilities Computed with the Fractional Model. The columns correspond to the different values of the time discretization step Δ_t.

Time	0.1	0.02	0.01	0.005	0.002	0.001	0.0005
1	0.7648	0.7612	0.7607	0.7604	0.7603	0.7602	0.7602
2	0.5968	0.5927	0.5922	0.5919	0.5918	0.5917	0.5917
3	0.4712	0.4672	0.4667	0.4664	0.4663	0.4662	0.4662
4	0.3755	0.3718	0.3713	0.3711	0.3709	0.3709	0.3709
5	0.3017	0.2983	0.2979	0.2977	0.2976	0.2976	0.2975
6	0.2443	0.2413	0.2410	0.2408	0.2407	0.2407	0.2406
7	0.1993	0.1968	0.1964	0.1963	0.1962	0.1962	0.1962
8	0.1638	0.1617	0.1614	0.1613	0.1612	0.1612	0.1612
9	0.1357	0.1339	0.1336	0.1335	0.1335	0.1334	0.1334
10	0.1133	0.1117	0.1115	0.1115	0.1114	0.1114	0.1114

Table 2.4: Numerical Solving of the FPDE and Monte-Carlo Simulations.

Time	Market	FPDE	Average	CI 99%	St. Dev.
1	0.7954	0.7604	0.7615	$[0.7555, 0.7675]$	0.0741
2	0.5673	0.5919	0.5928	$[0.5846, 0.6010]$	0.1011
3	0.4576	0.4664	0.4652	$[0.4557, 0.4748]$	0.1175
4	0.3633	0.3711	0.3697	$[0.3595, 0.3799]$	0.1249
5	0.3011	0.2977	0.2963	$[0.2859, 0.3067]$	0.1277
6	0.2492	0.2408	0.2402	$[0.2296, 0.2508]$	0.1301
7	0.2032	0.1963	0.1967	$[0.1861, 0.2074]$	0.1307
8	0.1661	0.1613	0.1625	$[0.1521, 0.1729]$	0.1277
9	0.1339	0.1335	0.1347	$[0.1247, 0.1447]$	0.1226
10	0.1061	0.1115	0.1127	$[0.1032, 0.1223]$	0.1177

Table 2.4 shows that the results obtained from our numerical method are congruent with Monte-Carlo simulations of the process $(\exp\{-\Lambda_{S_t}\})_{t\geqslant 0}$. The algorithm we use to obtain such simulations is described in details in Appendix 2. Table 2.4 is based on 1000 Monte-Carlo simulations and displays the market survival probabilities (Market), the survival probabilities obtained from solving the FPDE numerically (FPDE), a 99% confidence interval based on the standard deviation (CI 99% and St. Dev.) of the simulations and the survival probabilities computed from the simulations (Average). The survival probabilities and their standard deviations are respectively computed as the empirical means and standard deviations of the generated paths of $(\exp\{-\Lambda_{S_t}\})_{t\geqslant 0}$. Note that the numerical results from the FPDE are very close to the Monte-Carlo simulations and always lie in the 99% confidence interval.

From a computational cost point of view, the fractional model is of course more expensive. All the computations have been performed with the statistical software R. The non-fractional model is calibrated in 0.9234 second whereas the fractional model is calibrated in 4.9739 minutes. Once calibrated, the survival probabilities (up to 10 years maturity, as in Tables 2.2 and 2.4) in the fractional model can be computed in 1.4611 seconds (0.0239 second in the non-fractional case). This higher computational cost is inherent to mathematical models that exhibit memory: at each iteration, all the results of the previous iterations are taken into account, thereby substantially increasing the computational time.

Figure 2.5 shows an example of the different shapes we can obtain through the choice of the fractional order α. To obtain the survival probabilities of Figure 2.5, we used the parameters of the calibrated fractional model (see Table 2.1), with the exception of α that ranges between 0.1 and 1 by steps of 0.1. The main observable feature is the degree of convexity of the curves. With a low fractional order α, the survival probability drops sharply at the beginning, but decreases much slower

Table 2.5: Russia: market survival probabilities and results of the calibration. M, F and NF respectively denote the market data, the fractional model and the non-fractional model.

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$	$t=8$	$t=9$	$t=10$
M	0.3919	0.3735	0.3563	0.3401	0.3247	0.3099	0.2956	0.2818	0.2685	0.2555
F	0.4115	0.3676	0.3428	0.3258	0.3129	0.3026	0.2940	0.2867	0.2803	0.2747
NF	0.7776	0.6047	0.4702	0.3656	0.2843	0.2211	0.1719	0.1337	0.1040	0.0808

afterwards. This provides a good tool for modeling companies that have financial troubles in the short term, but have a good long term viability if they overcome their troubles.

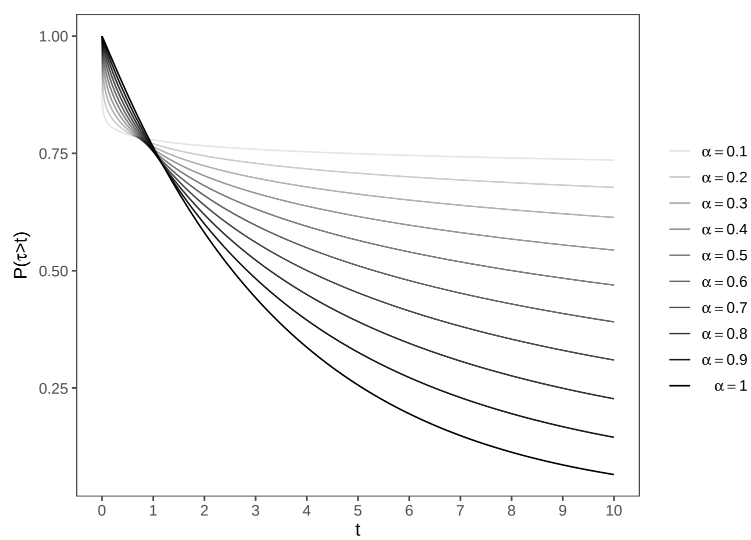

Figure 2.5: Survival Probability Curves in Function of α.

To conclude this section about numerical results, we propose two additional examples of calibrations to real data. In the first example, we show that the fractional model is also particularly suited for CDS on countries, and especially countries during war times. The first example is a calibration on default probabilities for the Russian Federation, taken from Bloomberg on the 20$^{\text{th}}$ September 2022. The results are displayed at Tables 2.5 and 2.6, as well as on Figure 2.6. We can observe here that the fractional model performs much better than the non-fractional model, the latter being completely unable to reproduce the survival probability shape.

Table 2.6: Calibration on survival probabilities for Russia.

	θ	κ	η	ρ	α
Non-fractional	0.2246	4503.011	205.9599	2.5872	1.0000
Fractional	1.0701	4726.523	205.9614	2.4648	0.2586

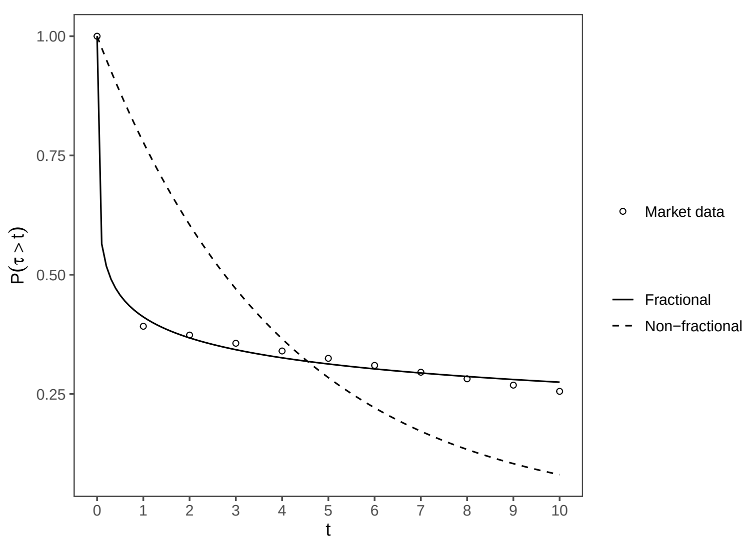

Figure 2.6: Survival probabilities for Russia and calibrated models.

Table 2.7: Microsoft: market survival probabilities and results of the calibration. M, F and NF respectively denote the market data, the fractional model and the non-fractional model.

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$	$t=8$	$t=9$	$t=10$
M	0.9975	0.9931	0.9860	0.9757	0.9630	0.9501	0.9354	0.9229	0.9096	0.8956
F	0.9904	0.9810	0.9716	0.9623	0.9530	0.9439	0.9349	0.9259	0.9171	0.9083
NF	0.9905	0.9810	0.9717	0.9624	0.9532	0.9441	0.9351	0.9262	0.9173	0.9086

Table 2.8: Calibration on survival probabilities for Microsoft.

	θ	κ	η	ρ	α
Non-fractional	0.0094	5.6258	0.2586	2.5861	1.0000
Fractional	0.0095	5.4025	0.2484	2.5871	0.9995

The second additional example shows a case in which neither the fractional model nor the non-fractional one can reproduce the survival probability shape. The data are default probabilities for Microsoft taken from Bloomberg on the 20$^{\text{th}}$ September 2022. The results are displayed at Tables 2.7 and 2.8, as well as on Figure 2.7. In this case, the survival probability curve on the market exhibits a concave shape, which the models can't reproduce. Note that from Table 2.6, we observe that the fractional model does not perform any better than the non-fractional one, the calibrated parameters being nearly the same for both models. This example shows that one should use an other model than the ones presented in this chapter if the survival probability curve of interest is concave.

2.4 Conclusions

This chapter proposes an application of fractional Hawkes jump processes to credit risk modeling. These processes present several interesting properties. Firstly, the intensity may be seen as a systematic factor driving the probability of failure of firms. The mechanism of self-excitation in this intensity ensures that the model explains periods during which we observe a persistent increase of default probabilities. Secondly, the fractional model allows to obtain desirable features that differ from usual intensity models. Indeed, we have seen that obtaining survival probability curves that are constant during long periods of time is not possible with classical intensity models. Thirdly, the model is numerically tractable as it is possible to perform a calibration on real market data. As shown by this calibration, the fractional model fits better. We have seen that survival probabilities can be obtained by solving a fractional Fokker-Planck Equation. In this equation, the derivative

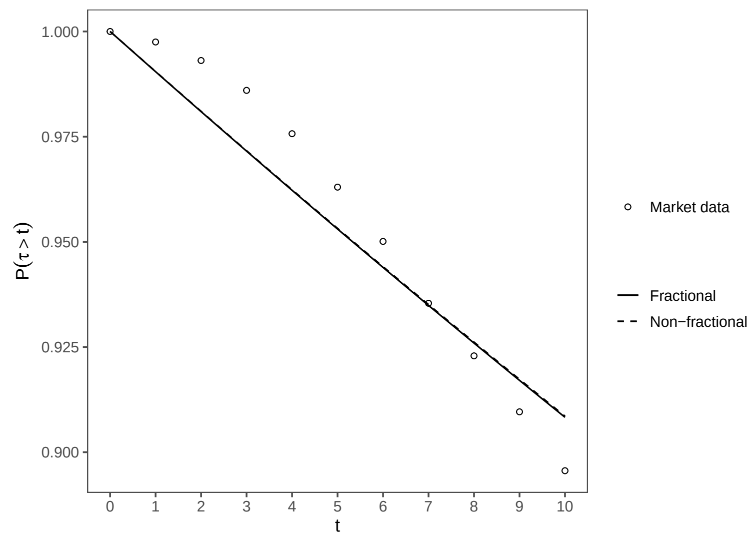

Figure 2.7: Survival probabilities for Microsoft and calibrated models.

with respect to time is replaced by a Caputo derivative. After a discussion about the proper approximation of Caputo derivatives, we have seen that the fractional model can explain a wide variety of term structures of survival probabilities.

2.5 Appendices

Appendix 1: Proofs of the Results of Section 2.1

Proof of Proposition 2.1. From SDE (2.1), we obtain the ordinary differential equation (ODE)

$$\frac{\mathrm{d}\mathbb{E}[\lambda_t]}{\mathrm{d}t} = \kappa(\theta - \mathbb{E}[\lambda_t]) + \frac{\eta}{\rho}\mathbb{E}[\lambda_t] \tag{2.29}$$

whose solution is well given by Equation (2.2). For the variance, noting that $\mathbb{E}[\xi^2] = \rho^{-2}$ and using Ito's lemma yields

$$\frac{\mathrm{d}\mathbb{E}[\lambda_t^2]}{\mathrm{d}t} = \mathbb{E}[\lambda_t]\left(2\kappa\theta + 2\left(\frac{\eta}{\rho}\right)^2\right) + 2\mathbb{E}[\lambda_t^2]\left(\frac{\eta}{\rho} - \kappa\right). \tag{2.30}$$

By the chain rule, we find

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbb{E}[\lambda_t])^2 = 2\kappa\theta\mathbb{E}[\lambda_t] + 2(\mathbb{E}[\lambda_t])^2\left(-\kappa + \frac{\eta}{\rho}\right). \tag{2.31}$$

Putting together Equations (2.30) and (2.31), we obtain $\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{V}\mathrm{ar}(\lambda_t)$. Then, replacing $\mathbb{E}[\lambda_t]$ by its expression at Equation (2.2) and performing some simplifications leads an ODE for the variance. The solution of this ODE is well given by Equation (2.3). □

Proof of Proposition 2.2. We use the bivariate Kramers-Moyal expansion on the function p. This expansion says that the bivariate PDF p of (λ_t, Λ_t) satisfies

$$\begin{aligned} &p(t+\Delta, x_1, x_2|\cdot) - p(t, x_1, x_2|\cdot) \\ &= \sum_{n=1}^{+\infty}\sum_{j=0}^{n} \frac{(-1)^n}{j!(n-j)!}\frac{\partial^j}{\partial x_1^j}\frac{\partial^{n-j}}{\partial x_2^{n-j}}\left(\mathbb{M}(j, n-j, \Delta|t, x_1, x_2)p(t, x_1, x_2|\cdot)\right) \end{aligned} \tag{2.32}$$

where $\mathbb{M}(j, n-j, \Delta|t, x_1, x_2) = \mathbb{E}[(\lambda_{t+\Delta} - \lambda_t)^j(\Lambda_{t+\Delta} - \Lambda_t)^{n-j}|\lambda_t = x_1, \Lambda_t = x_2]$. A proof of this result can be found in Hainaut (2020). We need to determine the values of

$$\mathbb{M}(j, n-j, \Delta|x_1, x_2) = \mathbb{E}[(\lambda_{t+\Delta} - \lambda_t)^j(\Lambda_{t+\Delta} - \Lambda_t)^{n-j}|\lambda_t = x_1, \Lambda_t = x_2]. \tag{2.33}$$

From Ito's Lemma applied to the function $h(x, y) = x^j y^{n-j}$ we find

$$\begin{aligned} &h(\lambda_{t+\Delta} - \lambda_t, \Lambda_{t+\Delta} - \Lambda_t) \\ &\quad = j\int_{s=0}^{\Delta}(\lambda_{t+s-} - \lambda_t)^{j-1}(\Lambda_{t+s} - \Lambda_t)^{n-j}\kappa(\theta - \lambda_{t+s-})\mathrm{d}s \\ &\quad + (n-j)\int_{s=0}^{\Delta}(\lambda_{t+s} - \lambda_t)^{j}(\Lambda_{t+s} - \Lambda_t)^{n-j-1}\lambda_{t+s-}\mathrm{d}s \\ &\quad + \int_{s=0}^{\Delta}(\Lambda_{t+s} - \Lambda_t)^{n-j}\Big((\lambda_{t+s} - \lambda_t)^j - (\lambda_{t+s-} - \lambda_t)^j\Big)\mathrm{d}N_{t+s}. \end{aligned} \tag{2.34}$$

Now looking at the expectation of each term in Equation (2.34), we get

$$\begin{aligned} &j\mathbb{E}\left[\int_{s=0}^{\Delta}(\lambda_{t+s-} - \lambda_t)^{j-1}(\Lambda_{t+s} - \Lambda_t)^{n-j}\kappa(\theta - \lambda_{t+s-})\mathrm{d}s\right] \\ &\quad = \begin{cases} \kappa(\theta - \lambda_t)\Delta + \mathcal{O}(\Delta^2) & \text{if } j = 1 \text{ and } n = 1 \\ \mathcal{O}(\Delta^2) & \text{otherwise,} \end{cases} \end{aligned} \tag{2.35}$$

$$\begin{aligned}(n-j)\mathbb{E}&\left[\int_{s=0}^{\Delta}(\lambda_{t+s-}-\lambda_t)^j(\Lambda_{t+s}-\Lambda_t)^{n-j-1}\lambda_{t+s-}\mathrm{d}s\right]\\&=\begin{cases}\lambda_t\Delta+\mathcal{O}(\Delta^2) & \text{if } j=0 \text{ and } n=1\\ \mathcal{O}(\Delta^2) & \text{otherwise,}\end{cases}\end{aligned}\tag{2.36}$$

$$\begin{aligned}\mathbb{E}&\left[\int_{s=0}^{\Delta}(\Lambda_{t+s}-\Lambda_t)^{n-j}\Big((\lambda_{t+s}-\lambda_t)^j-(\lambda_{t+s-}-\lambda_t)^j\Big)\mathrm{d}N_{t+s}\right]\\&=\begin{cases}\eta^j\mathbb{E}[\xi^j]\Delta\lambda_t+\mathcal{O}(\Delta^2) & \text{if } n=j\\ \mathcal{O}(\Delta^2) & \text{otherwise.}\end{cases}\end{aligned}\tag{2.37}$$

Then dividing Equation (2.32) by Δ and replacing the terms by what we just found leads to

$$\begin{aligned}&\frac{p(t+\Delta,x_1,x_2|\cdot)-p(t,x_1,x_2|\cdot)}{\Delta}\\&=-\kappa\Big(-p(t,x_1,x_2|\cdot)+(\theta-x_1)\frac{\partial p(t,x_1,x_2|\cdot)}{\partial x_1}\Big)\\&+\mathbb{E}\left[\sum_{j=1}^{+\infty}\frac{(-\eta\xi)^j}{j!}\frac{\partial^j}{\partial x_1^j}(x_1p(t,x_1,x_2|\cdot))\right]-x_1\frac{\partial p(t,x_1,x_2|\cdot)}{\partial x_2}+\mathcal{O}(\Delta).\end{aligned}\tag{2.38}$$

Recall that

$$\frac{\partial^j}{\partial x_1^j}(x_1p(t,x_1,x_2|\cdot))=j\frac{\partial^{j-1}p(t,x_1,x_2|\cdot)}{\partial x_1^{j-1}}+x_1\frac{\partial^j p(t,x_1,x_2|\cdot)}{\partial x_1^j},\tag{2.39}$$

which we can use to write

$$\begin{aligned}\mathbb{E}&\left[\sum_{j=1}^{+\infty}\frac{(-\eta\xi)^j}{j!}\frac{\partial^j}{\partial x_1^j}(x_1p(t,x_1,x_2|\cdot))\right]\\&=\mathbb{E}\left[-\eta\xi\sum_{j=0}^{+\infty}\frac{(-\eta\xi)^j}{j!}\left(\frac{\partial^j}{\partial x_1^j}p(t,x_1,x_2|\cdot)\right)\right]\\&\quad+\mathbb{E}\left[\sum_{j=1}^{+\infty}\frac{(-\eta\xi)^j}{j!}x_1\frac{\partial^j p(t,x_1,x_2|\cdot)}{\partial x_1^j}\right]\\&=x_1\mathbb{E}[p(t,x_1-\eta\xi,x_2|\cdot)-p(t,x_1,x_2|\cdot)]\\&\quad-\eta\mathbb{E}[p(t,x_1-\eta\xi,x_2|\cdot)].\end{aligned}\tag{2.40}$$

It follows that letting Δ tend to zero in Equation (2.38) leads to the announced result. $\square$

Proof of Proposition 2.3. We start by writing

$$\begin{aligned}\frac{\partial \phi(t,z_1,z_2|\cdot)}{\partial t} &= \lim_{\Delta\to 0}\frac{\phi(t+\Delta,z_1,z_2|\cdot)-\phi(t,z_1,z_2|\cdot)}{\Delta}\\ &= \int_0^{+\infty}\int_0^{+\infty} e^{-z_1x_1-z_2x_2}\frac{\partial p(t,x_1,x_2|\cdot)}{\partial t}\mathrm{d}x_1\mathrm{d}x_2.\end{aligned} \tag{2.41}$$

This allows us to use the Fokker-Planck Equation (2.5), which leads to

$$\begin{aligned}&\frac{\partial \phi(t,z_1,z_2|\cdot)}{\partial t} = \kappa\int_0^{+\infty}\int_0^{+\infty} e^{-z_1x_1-z_2x_2}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\\ &-\int_0^{+\infty}\int_0^{+\infty} e^{-z_1x_1-z_2x_2}\kappa(\theta-x_1)\frac{\partial p(t,x_1,x_2|\cdot)}{\partial x_1}\mathrm{d}x_1\mathrm{d}x_1\\ &-\int_0^{+\infty}\int_0^{+\infty} e^{-z_1x_1-z_2x_2}x_1\frac{\partial p(t,x_1,x_2|\cdot)}{\partial x_2}\mathrm{d}x_1\mathrm{d}x_2\\ &-\eta\int_0^{+\infty}\int_0^{+\infty} e^{-z_1x_1-z_2x_2}\mathbb{E}[\xi p(t,x_1-\eta\xi,x_1|\cdot)]\mathrm{d}x_1\mathrm{d}x_2\\ &+\int_0^{+\infty}\int_0^{+\infty} x_1e^{-z_1x_1-z_2x_2}\mathbb{E}[p(t,x_1-\eta\xi,x_2|\cdot)]\mathrm{d}x_1\mathrm{d}x_2\\ &-\int_0^{+\infty}\int_0^{+\infty} x_1e^{-z_1x_1-z_2x_2}\mathbb{E}[p(t,x_1,x_2|\cdot)]\mathrm{d}x_1\mathrm{d}x_2.\end{aligned} \tag{2.42}$$

The proof is then completed by computing all the terms in Equation (2.42). This can be done with integrations by parts. $\square$

Appendix 2: Monte-Carlo Simulations

The Monte-Carlo simulations of the processes $(\lambda_t)_{t\geqslant 0}$ are performed with a classic Euler discretization scheme. Let $\delta > 0$ be the discretization step size. This step size was set to 10^{-3} in our computations. In order to simulate a discretized path $(\lambda^\delta_{n\delta})_{n\geqslant 0}$ of $(\lambda_t)_{t\geqslant 0}$, we set $\lambda^\delta_0 := \lambda_0 \in \mathbb{R}$ with $\lambda_0 \geqslant \theta$ and for $n \geqslant 0$, we set

$$\lambda^\delta_{(n+1)\delta} := \lambda^\delta_{n\delta} + \kappa(\theta - \lambda^\delta_{n\delta})\delta + \eta\xi_n B_n$$

where ξ_n is exponentially distributed with mean ρ^{-1} and B_n is Poisson distributed with parameter $\delta\lambda^\delta_{n\delta}$. All the random variables ξ_n and B_n are independent. The integral $(\Lambda_t)_{t\geqslant 0}$ of $(\lambda_t)_{t\geqslant 0}$ is approximated by $(\Lambda^\delta_{n\delta})_{n\geqslant 0}$. The latter is computed with Riemann sums, that is $\Lambda^\delta_0 := 0$ and

$$\Lambda^\delta_{(n+1)\delta} := \Lambda^\delta_{n\delta} + \delta\lambda^\delta_{n\delta}.$$

Simulating the inverse α-stable process $(S_t)_{t\geqslant 0}$ is done with the algorithm proposed in Magdziarz (2009b) which we describe now. We still work with the same discretization step $\delta = 10^{-3}$. In order to simulate the discretized process $(S^\delta_{n\delta})_{n\geqslant 0}$, we start by simulating the discretized α-stable process $(U^\delta_{n\delta})_{n\geqslant 0}$. To do so, we set $U^\delta_0 := 0$ and

$$U^\delta_{(n+1)\delta} := U^\delta_{n\delta} + \delta^{1/\alpha}X_n$$

where X_n are i.i.d. totally skewed positive α-stable random variables. Realizations of such random variables are obtained with

$$X_n := \frac{\sin(\alpha(V_n + \frac{\pi}{2}))}{(\cos(V_n))^{1/\alpha}}\left(\frac{\cos(V_n - \alpha(V_n + \frac{\pi}{2})}{W_n}\right)^{\frac{1-\alpha}{\alpha}}$$

where V_n are i.i.d. random variables uniformly distributed on $(-\pi/2, \pi/2)$ and W_n are i.i.d. exponentially distributed random variables with mean 1. For more about simulations of α-stable random variables, see the Chapter 3 of Janicki and Weron (1994). It is then possible to approximate $(S_t)_{t\geqslant 0}$ with

$$S^\delta_t := (\min\{n \in \mathbb{N} : U^\delta_{n\delta} > t\} - 1)\delta.$$

By Theorem 2 in Magdziarz (2009b), $\sup_{0\leqslant s\leqslant T} |S_s - S^\delta_s| \leqslant \delta$ a.s. for all $T > 0$, so that the error in the approximation of $(S_t)_{t\geqslant 0}$ is always bounded by $\delta = 10^{-3}$. Given that $S^\delta_t \in \{\delta n : n \in \mathbb{N}\} \cup \{0\}$ for all $t \geqslant 0$, it is easy to combine $(\Lambda^\delta_{n\delta})_{n\geqslant 0}$ and $(S^\delta_t)_{t\geqslant 0}$ to obtain a simulated path of $(\Lambda_{S_t})_{t\geqslant 0}$ by setting $\Lambda^\delta_{S_t} := \Lambda^\delta_{S^\delta_t}$.

Chapter 3

A Recursive Method for Computing Moments of Hawkes Intensities: Application to the Potential Approach of Credit Risk

3.1 Introduction

As pointed out in the introductions to the first and second chapters of this thesis, the most common paradigms in credit risk management are structural and reduced models. In the first approach, the balance sheet of a firm is explicitly modeled and the default occurs when the assets fall below the liabilities, as e.g. in Capinski (2007), He and Chen (2018), Jeon et al. (2016) or Hainaut and Colwell (2014). In the reduced approach, the default is caused by the first jump of a point process, see e.g. Liang and Wang (2012), Hainaut and Le Courtois (2014) or Eckert et al. (2016). Chen and Panjer (2003) unify discrete structural and reduced models whereas Ballestra and Pacelli (2014) propose an hybrid approach. In this chapter, we investigate a third alternative called the potential approach. This approach dates back to Constantinides (1992) and has been frequently used in the literature for interest rate modeling.

In interest rate modeling, the most frequent approaches are respectively to model the spot rate, which is done in e.g. Vasicek (1977) and Cox et al. (1985), or to model the forward rate. Well known works for the forward rate approach are Ho and Lee (1986) and Heath et al. (1992). The spot rate approach consists in specifying a stochastic process $(r_t)_{t\geqslant 0}$, called the spot rate or instantaneous rate, so that for a contingent claim Y paid at time $T \geqslant 0$, the price at time $t \in [0,T]$ is given by

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left\{-\int_t^T r_s \mathrm{d}s\right\} Y \middle| \mathcal{F}_t\right],$$

where $\mathbb{Q}$ is a risk-neutral measure and $(\mathcal{F}_t)_{t\geqslant 0}$ is the relevant filtration. In case the contingent claim Y equals one with probability one, we obtain the important particular case of a zero-coupon bond, whose price is usually denoted by $P(t,T)$,

$$P(t,T) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{-\int_t^T r_s \mathrm{d}s\right\} \middle| \mathcal{F}_t\right]. \tag{3.1}$$

Constantinides (1992) proposed to directly model the state price process, also called the pricing kernel. This is a stochastic process $(\xi_t)_{t\geqslant 0}$ such that for a contingent claim Y, the price is given by

$$\frac{\mathbb{E}_{\mathbb{Q}}\left[\xi_T Y \middle| \mathcal{F}_t\right]}{\xi_t}.$$

This was called the potential approach by Rogers (1997), since $(\xi_t)_{t\geqslant 0}$ should be a potential[1] to guarantee satisfying properties[2] for the model. With the potential approach, Rutkowski (1997) derived valuation formulas for interest rate caps and swaptions. Since then, the potential approach for interest rates has been used in numerous studies, including Flesaker and Hughston (1999), Macrina (2014), Brody et al. (2020) and Da Fonseca et al. (2022).

In the reduced approach to credit risk, it appears that the credit model has the same structure as a spot rate model for interest rates. More precisely, the reduced approach to credit risk uses a positive process $(\lambda_t)_{t\geqslant 0}$, called the default intensity, such that

$$\mathbb{Q}(\tau > T | \tau > t) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left\{-\int_t^T \lambda_s \,\mathrm{d}s\right\} \middle| \mathcal{F}_t\right], \tag{3.2}$$

where τ is the time of default. Note the similarity between Equations (3.1) and (3.2). Given this similarity, we propose to use the potential approach for credit risk by computing the survival probabilities from a potential $(\xi_t)_{t\geqslant 0}$. Equation (3.2) would then become

$$\mathbb{Q}(\tau > T | \tau > t) = \frac{\mathbb{E}_{\mathbb{Q}}\left[\xi_T \middle| \mathcal{F}_t\right]}{\xi_t}.$$

[1] A positive supermartingale whose expectation converges to zero as the time goes to infinity.
[2] Such as the positivity of interest rates.

In this chapter, the potential approach is extended to credit risk in two directions. Firstly, we work with a self-excited process that allows sudden and repeated jumps of default probabilities, as those caused by contagion between defaults of several firms. In this framework, the intensity of default arrival depends on recent past shocks. This approach finds its origin in the articles of Hawkes (1971a), Hawkes (1971b) and Hawkes and Oakes (1974). More recently, Errais et al. (2010) develop a self-excited approach for modeling defaults in a credit portfolio. Ait-Sahalia et al. (2015) use a similar model to measure contagion between international markets. Hainaut (2016b) and Hainaut (2016a) propose a model explaining the dynamics of interest rates by self-excited jump-diffusions.

The second way in which we extend the potential approach is the use of the technique of time-change. As hinted in the introduction to this thesis, as well as in Chapter 2, this technique allows to obtain a better fit for survival probability curves that exhibit a high degree of convexity. Such curves are observed for companies that face financial difficulties in the short-term, but has a good long-term viability provided that they overcome their difficulties. Airline companies during the Covid-19 crisis were given as an example above. Obtaining such convex shapes involves a non-Markov time-change built with the inverse of an α-stable subordinator, as in Chapter 2. This approach is applied to diffusions in physics for describing the movements of heavy particles that can get immobilized (see e.g. Metzler and Klafter (2004) or Eliazar and Klafter (2004)). This type of time-changed Brownian motions, called subdiffusions, are also popular in econophysics (see Scalas (2006)) and applied to financial derivatives by Magdziarz (2009a) for illiquid markets. As shown e.g. in Magdziarz (2009b), the probability density function of sub-diffusions is solution of a Fractional Fokker-Planck equation. This equation is proposed in Barkai et al. (2000) and Metzler et al. (1999). Articles of Leonenko et al. (2013b) and Leonenko et al. (2013a) go a step further and explore fractional Pearson diffusions and their correlation.

Computing default probabilities in such a setting requires to be able to approximate the moment generating function of (fractional) Hawkes process intensities. To do this, the most common approach is to derive a (fractional partial) differential equation and to solve it numerically, as in e.g. Ketelbuters and Hainaut (2022a) (Chapter 2 of this thesis). Since this approach is rather complicated both mathematically and numerically, we propose here to use an other approach. Our idea is to use the moments of the Hawkes process intensity to approximate the moment generating function via the Maclaurin series of the exponential function. In a recent article, Cui et al. (2020) describe a method to derive the moments of Hawkes processes and their intensities. In this chapter, we go a step further and prove that the moments of the Hawkes process intensity always share the same form and can be computed recursively. This recursive procedure is new to the best of our knowledge and allows to efficiently compute default probabilities in our proposed model.

The structure of this chapter is as follows. Sections 2 and 3 review the potential approach and the properties of a Hawkes jump-diffusion. In Section 4, we extend the Hawkes jump-diffusion to the subdiffusive setting by using the inverse of an α-stable subordinator as time-change. Section 5 introduces our new numerical method to compute all the moments of a Hawkes process intensity. We show that this method also works in the subdiffusive time-changed setting. We end Section 5 by showing how this result can be used to approximate default probabilities and provide upper bounds on the numerical error. Finally, Section 6 presents a numerical analysis and shows in particular that our model performs as well as in Chapter 2, while being much simpler to deal with numerically.

3.2 The potential approach for credit risk modeling

All stochastic processes considered in this chapter are defined on a probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ where $\mathbb{Q}$ is the risk-neutral measure. The default time of a firm is a random time denoted by τ. As mentioned in the introduction, the two most common approaches in credit risk management are the structural and reduced models. In the first one, the default is triggered when the market value of assets falls below debts. In the second approach, the default is caused by the first jump of a point process. In this chapter, we investigate an alternative way, called the potential approach. Instead of modeling an event triggering the default, we propose a dynamics for the probability that the company survives until a certain maturity. This probability is of course in the interval $[0,1]$ and is a decreasing function of the time horizon. The potential approach is based upon an $\mathbb{F}^\xi = (\mathcal{F}^\xi_t)_{t\geqslant 0}$-adapted stochastic process $(\xi_t)_{t\geqslant 0}$. The filtration $\mathbb{F}^\xi$ is defined precisely in the next section. If the company is still in activity at time s, the survival probability until time $t \geqslant s$ is denoted by $\mathbb{Q}(\tau > t|\tau > s)$ and defined as

$$\mathbb{Q}(\tau > t|\tau > s) = \frac{1}{\xi_s}\mathbb{E}_{\mathbb{Q}}[\xi_t|\mathcal{F}^\xi_s].$$

To ensure that $\frac{1}{\xi_s}\mathbb{E}_{\mathbb{Q}}[\xi_t|\mathcal{F}^\xi_s]$ takes its values in $[0,1]$ and decreases with t, $(\xi_t)_{t\geqslant 0}$ needs no fulfill three conditions

(i) $\xi_t > 0$ for all t, $\mathbb{Q}$-a.s.;

(ii) $(\xi_t)_{t\geqslant 0}$ is a supermartingale, i.e. $\mathbb{E}_{\mathbb{Q}}[\xi_t|\mathcal{F}^\xi_s] \leqslant \xi_s$ for all $s \leqslant t$, $\mathbb{Q}$-a.s.;

(iii) the expectation satisfies $\lim_{t\to+\infty} \mathbb{E}_{\mathbb{Q}}[\xi_t] = 0$.

There exist various ways to build such a process in a Brownian setting. We refer to e.g. Rogers (1997) for examples. In this chapter, we consider a process $(\xi_t)_{t\geqslant 0}$ that is a Hawkes jump-diffusion supermartingale. The construction and properties of this process are developed in the next sections.

We emphasize that the model is directly specified under the risk-neutral measure $\mathbb{Q}$. It is the case because our model will be calibrated on default probabilities inferred from CDS quotes on the market. As a consequence, the resulting default probabilities do not correspond to the physical measure $\mathbb{P}$ but to the risk-neutral measure $\mathbb{Q}$. In particular, the numerical values for these probabilities can't be interpreted as being the probabilities of default in the real world.

3.3 Hawkes jump-diffusion supermartingales

This section is devoted to the construction of Hawkes jump-diffusion supermartingales. We consider a self-exciting counting process $(N_t)_{t\geqslant 0}$ whose intensity is given by

$$\mathrm{d}\lambda_t = \kappa(\theta - \lambda_{t-})\mathrm{d}t + \sigma\sqrt{\lambda_{t-}}\mathrm{d}W_t + \eta \mathrm{d}P_t$$

where $(W_t)_{t\geqslant 0}$ is a standard Brownian motion and

$$P_t = \sum_{k=1}^{N_t} J_k.$$

The jump sizes $J_1, J_2, \cdots =^{\text{dist}} J$ are i.i.d. random variables that are assumed to be distributed according to a measure ν defined as

$$\nu(B) := \int_{B\cap\mathbb{R}^+} \rho e^{-\rho x}\mathrm{d}x, \quad B \in \mathcal{B}_{\mathbb{R}}$$

with $\rho > 0$. As a consequence, $J_1, J_2, \dots$ have an exponential distribution with parameter ρ (and thus expected value $1/\rho$). The condition $\kappa > \eta/\rho$ that guarantees that $\lim_{t\to+\infty} \mathbb{E}_{\mathbb{Q}}[\lambda_t] < +\infty$ is assumed to hold and is referred to as the non-explosion condition.

We denote by $\mathbb{F}^{\xi} = (\mathcal{F}^{\xi}_t)_{t\geqslant 0}$ the filtration generated by the process $(\lambda_t, P_t, N_t, W_t)_{t\geqslant 0}$. We assume that the random variables $J_1, J_2, \dots$ are in some sense independent of the process $(\lambda_t, P_t, N_t, W_t)_{t\geqslant 0}$. More precisely, if we denote by $T_1, T_2, \dots$ the stopping times that represent the jump times of $(\lambda_t, N_t, P_t)_{t\geqslant 0}$, that is

$$T_k = \inf\{\tau > 0 : N_\tau = k\}$$

then we assume that J_k is independent of $\mathcal{F}^{\xi}_{T_k-}$ for any $k = 1, 2 \dots$. In other words, nothing can be said about the jump size J_k until just before the associated jump time T_k. Of course, the jump sizes $J_1, J_2, \dots$ are not independent from the process $(N_t)_{t\geqslant 0}$, as the size of a jump influences the occurrence of subsequent jumps of $(N_t)_{t\geqslant 0}$. However, past jumps do not have any impact on subsequent jump sizes. Using our assumptions, we can obtain an expression for the expected value of the exponential of $(\lambda_t)_{t\geqslant 0}$.

Lemma 3.1. *For any $z < \rho/\eta$ and $t \geqslant u \geqslant 0$, it holds that*

$$\begin{aligned}&\mathbb{E}_{\mathbb{Q}}[e^{z\lambda_t} - e^{z\lambda_u}]\\&\quad= \mathbb{E}_{\mathbb{Q}}\left[\int_u^t \left\{ z\kappa\theta e^{z\lambda_{s-}} + \left(-z\kappa + \frac{\sigma^2}{2}z^2 + \frac{\rho}{\rho - \eta z} - 1\right)\lambda_{s-}e^{z\lambda_{s-}}\right\}\mathrm{d}s\right].\end{aligned}$$

Proof. Fix $z < \rho/\eta$. Ito's lemma applied on the process $(e^{z\lambda_t})_{t\geqslant 0}$ yields

$$\begin{aligned}e^{z\lambda_t} - e^{z\lambda_u} = &\int_u^t ze^{z\lambda_{s-}}\kappa(\theta - \lambda_{s-})\mathrm{d}s + \int_u^t ze^{z\lambda_{s-}}\sigma\sqrt{\lambda_{s-}}\mathrm{d}W_s\\&+ \frac{\sigma^2}{2}\int_u^t z^2 e^{z\lambda_{s-}}\lambda_{s-}\mathrm{d}s + \sum_{u<s\leqslant t}(e^{z\lambda_s} - e^{z\lambda_{s-}}).\end{aligned} \tag{3.3}$$

The jump part can be rewritten as

$$\sum_{u<s\leqslant t}(e^{z\lambda_s} - e^{z\lambda_{s-}}) = \sum_{k\geqslant 1}\mathbf{1}_{\{u<T_k\leqslant t\}}e^{z\lambda_{T_k-}}\left(e^{z\eta J_k} - 1\right).$$

From this equation, we compute that

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}\left[\sum_{u<s\leqslant t}(e^{z\lambda_s} - e^{z\lambda_{s-}})\right] &= \sum_{k\geqslant 1}\mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{\{u<T_k\leqslant t\}}e^{z\lambda_{T_k-}}\left(e^{z\eta J_k} - 1\right)\right]\\&= \sum_{k\geqslant 1}\mathbb{E}_{\mathbb{Q}}\left[\mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{\{T_k\leqslant t\}}e^{z\lambda_{T_k-}}\left(e^{z\eta J_k} - 1\right)\big|\mathcal{F}^{\xi}_{T_k-}\right]\right]\\&= \sum_{k\geqslant 1}\mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{\{u<T_k\leqslant t\}}e^{z\lambda_{T_k-}}\mathbb{E}_{\mathbb{Q}}\left[\left(e^{z\eta J_k} - 1\right)\big|\mathcal{F}^{\xi}_{T_k-}\right]\right]\\&= \left(\frac{\rho}{\rho-\eta z} - 1\right)\sum_{k\geqslant 1}\mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}_{\{u<T_k\leqslant t\}}e^{z\lambda_{T_k-}}\right]\\&= \left(\frac{\rho}{\rho-\eta z} - 1\right)\mathbb{E}_{\mathbb{Q}}\left[\int_u^t e^{z\lambda_{s-}}\mathrm{d}N_s\right]\\&= \left(\frac{\rho}{\rho-\eta z} - 1\right)\mathbb{E}_{\mathbb{Q}}\left[\int_u^t \lambda_{s-}e^{z\lambda_{s-}}\mathrm{d}s\right].\end{aligned} \tag{3.4}$$

The first equality of Equation (3.4) follows from the monotone convergence theorem, as all the terms are non-negative. The second equality comes from the tower property. The third equality is a consequence of the $\mathcal{F}^{\xi}_{T_k-}$-measurability of the random variables T_k and λ_{T_k-}. The fourth equality is a consequence of our assumption of independence between $\mathcal{F}^{\xi}_{T_k-}$ and J_k, combined with the formula for the moment generating function of exponential random variables. The formula for this moment generating function is valid because of the assumption $\eta z < \rho$. Finally, the last two equalities are respectively consequences of the definition of stochastic integral and the fact that the process $(\int_0^t \lambda_{s-}\mathrm{d}s)_{t\geqslant 0}$ is the compensator of $(N_t)_{t\geqslant 0}$. By combining Equations (3.3) and (3.4), we find

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\left[e^{z\lambda_t}-e^{z\lambda_u}\right]\\
&=\mathbb{E}_{\mathbb{Q}}\left[\begin{aligned}\int_u^t ze^{z\lambda_{s-}}\kappa(\theta-\lambda_{s-})\mathrm{d}s+\frac{\sigma^2}{2}\int_u^t z^2e^{z\lambda_{s-}}\lambda_{s-}\mathrm{d}s\\
+\left(\frac{\rho}{\rho-\eta z}-1\right)\int_u^t\lambda_{s-}e^{z\lambda_{s-}}\mathrm{d}s\end{aligned}\right]\\
&=\mathbb{E}_{\mathbb{Q}}\left[\int_u^t\left\{z\kappa\theta e^{z\lambda_{s-}}+\left(-z\kappa+\frac{\sigma^2}{2}z^2+\frac{\rho}{\rho-\eta z}-1\right)\lambda_{s-}e^{z\lambda_{s-}}\right\}\mathrm{d}s\right],
\end{aligned}$$

which is what we had to prove. □

The condition $z<\rho/\eta$ in Lemma 3.1 is essential for the existence of the Laplace transform of the jump distribution ν and is therefore assumed to hold throughout this chapter. Lemma 3.1 can be used to prove that the random variables $(e^{z\lambda_t})_{t\geqslant 0}$ have finite expectation for some $z>0$. To this end, we study the function q

$$(-\infty,\rho/\eta)\ni z\mapsto q(z):=-z\kappa+\frac{\sigma^2}{2}z^2+\frac{\rho}{\rho-\eta z}-1$$

that appears in the expectation. Obviously we can restrict our attention to $z\in(0,\rho/\eta)$, as $e^{z\lambda_t}\leqslant 1$ implies that the expectation is finite when $z\leqslant 0$.

Lemma 3.2. *If $\sigma>0$ and if we define the constants*

$$\zeta=\left(\frac{\sigma^2}{2}\rho+\eta\kappa\right)^2+2\eta\sigma^2(\eta-\kappa\rho) \tag{3.5}$$

and

$$\gamma_{\pm}=\frac{\left(\frac{\sigma^2}{2}\rho+\eta\kappa\right)\pm\sqrt{\zeta}}{\eta\sigma^2}, \tag{3.6}$$

then $0<\gamma_-<\rho/\eta<\gamma_+$. Moreover, for $z\in(0,\rho/\eta)$,

$$q(z)<0\iff z\in(0,\gamma_-),$$

$$q(z)=0\iff z=\gamma_-,$$

and

$$q(z)>0\iff z\in(\gamma_-,\rho/\eta).$$

If $\sigma=0$ then $\frac{\kappa\rho-\eta}{\eta\kappa}>0$ and for $z\in(0,\rho/\eta)$,

$$q(z)<0\iff z\in(0,\frac{\kappa\rho-\eta}{\eta\kappa}),$$

$$q(z) = 0 \Longleftrightarrow z = \frac{\kappa\rho - \eta}{\eta\kappa},$$

and

$$q(z) > 0 \Longleftrightarrow z \in (\frac{\kappa\rho - \eta}{\eta\kappa}, \rho/\eta). \tag{3.7}$$

Proof. Since $\rho - \eta z > 0$ and $z > 0$, multiplying the inequality $q(z) \leqslant 0$ by $(\rho - \eta z)$ and dividing it by z yields

$$q(z) \leqslant 0 \Longleftrightarrow -\frac{\sigma^2}{2}\eta z^2 + \left(\frac{\sigma^2}{2}\rho + \kappa\eta\right) z + \eta - \kappa\rho \leqslant 0. \tag{3.8}$$

Assume that $\sigma > 0$. Then the determinant the second order equation in (3.8) is given by ζ in Equation (3.5). Some straightforward computations lead to

$$\zeta = \left(\frac{\sigma^2}{2}\rho - \eta\kappa\right)^2 + 2\eta^2\sigma^2, \tag{3.9}$$

so that ζ is strictly positive. As a consequence, $q(\gamma_\pm) = 0$, where $\gamma_\pm$ is given in Equation (3.6). Recall that the non-explosion condition for the Hawkes process intensity is the inequality $\eta - \kappa\rho < 0$. Therefore,

$$\zeta = \left(\frac{\sigma^2}{2}\rho + \eta\kappa\right)^2 + 2\eta\sigma^2(\eta - \kappa\rho) < \left(\frac{\sigma^2}{2}\rho + \eta\kappa\right)^2,$$

which implies that both quantities of Equation (3.6) are strictly positive. Moreover, Equation (3.9) implies that

$$\sqrt{\zeta} > \left|\frac{\sigma^2}{2}\rho - \eta\kappa\right| \geqslant \eta\kappa - \frac{\sigma^2}{2}\rho$$

and thus

$$\gamma_- = \frac{\left(\frac{\sigma^2}{2}\rho + \eta\kappa\right) - \sqrt{\zeta}}{\eta\sigma^2} < \frac{\left(\frac{\sigma^2}{2}\rho + \eta\kappa\right) - \left(\eta\kappa - \frac{\sigma^2}{2}\rho\right)}{\eta\sigma^2} = \frac{\rho}{\eta}. \tag{3.10}$$

In addition,

$$\begin{aligned}\gamma_+ &= \frac{\left(\frac{\sigma^2}{2}\rho + \eta\kappa\right) + \sqrt{\left(\frac{\sigma^2}{2}\rho - \eta\kappa\right)^2 + 2\eta^2\sigma^2}}{\eta\sigma^2} \\ &> \frac{\left(\frac{\sigma^2}{2}\rho + \eta\kappa\right) + \left|\frac{\sigma^2}{2}\rho - \eta\kappa\right|}{\eta\sigma^2} \\ &\geqslant \frac{\sigma^2\rho}{\eta\sigma^2} = \frac{\rho}{\eta}.\end{aligned}$$

All the claims of the $\sigma > 0$ case are proved. In the case $\sigma = 0$, the claims easily follow from Equation (3.8) and the non-explosion condition $\eta - \kappa\rho < 0$. $\square$

These lemmas allow to prove the integrability of $e^{z\lambda_t}$, as well as an upper bound for its expected value.

Corollary 3.1. *Define*

$$\gamma = \begin{cases} \gamma_- & \text{if } \sigma > 0, \\ \frac{\kappa\rho - \eta}{\eta\kappa} & \text{if } \sigma = 0. \end{cases}$$

Then for any $z \leqslant \gamma$ and any $t \geqslant 0$, the random variable $e^{z\lambda_t}$ is $\mathbb{Q}$-integrable. Moreover, we have the upper bound

$$\mathbb{E}_{\mathbb{Q}}[e^{z\lambda_t}|\mathcal{F}_s^{\xi}] \leqslant 1 \vee e^{z\lambda_s + z\kappa\theta(t-s)}$$

whenever $t \geqslant s \geqslant 0$. Moreover,

$$\mathbb{E}_{\mathbb{Q}}[e^{\gamma\lambda_t}|\mathcal{F}_s^{\xi}] = e^{\gamma\lambda_s + \gamma\kappa\theta(t-s)}.$$

Proof. Since $e^{z\lambda_t} \leqslant 1$ when z is nonpositive, we only have to consider the case $z \in (0, \gamma)$. Fix any $t \geqslant u \geqslant s \geqslant 0$. From Lemmas 3.1 and 3.2, we find

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}\left[e^{z\lambda_t} - e^{z\lambda_u}|\mathcal{F}_s^{\xi}\right] &= z\kappa\theta\mathbb{E}_{\mathbb{Q}}\left[\int_u^t e^{z\lambda_{\tau-}}\mathrm{d}\tau|\mathcal{F}_s^{\xi}\right] + q(z)\mathbb{E}_{\mathbb{Q}}\left[\int_u^t \lambda_{\tau-}e^{z\lambda_{\tau-}}\mathrm{d}\tau|\mathcal{F}_s^{\xi}\right] \\ &\leqslant z\kappa\theta\int_u^t \mathbb{E}_{\mathbb{Q}}\left[e^{z\lambda_{\tau-}}|\mathcal{F}_s^{\xi}\right]\mathrm{d}\tau, \end{aligned}$$

where we used Tonelli's theorem to swap the integral and the expectation. It implies that

$$\frac{\partial}{\partial t}\mathbb{E}_{\mathbb{Q}}\left[e^{z\lambda_t}|\mathcal{F}_s^{\xi}\right] \leqslant z\kappa\theta\mathbb{E}_{\mathbb{Q}}\left[e^{z\lambda_t}|\mathcal{F}_s^{\xi}\right],$$

$t \geqslant s$, and thus Grönwall's inequality yields

$$\mathbb{E}_{\mathbb{Q}}\left[e^{z\lambda_t}|\mathcal{F}_s^{\xi}\right] \leqslant e^{z\lambda_s + z\kappa\theta(t-s)}.$$

In the case $z = \gamma$, Lemma 3.2 yields

$$\frac{\partial}{\partial t}\mathbb{E}_{\mathbb{Q}}\left[e^{\gamma\lambda_t}|\mathcal{F}_s^{\xi}\right] = z\kappa\theta\mathbb{E}_{\mathbb{Q}}\left[e^{\gamma\lambda_t}|\mathcal{F}_s^{\xi}\right]$$

and hence

$$\mathbb{E}_{\mathbb{Q}}[e^{\gamma\lambda_t}|\mathcal{F}_s^{\xi}] = e^{\gamma\lambda_s + \gamma\kappa\theta(t-s)},$$

as announced. □

With the help of our previous results, we build a Hawkes jump-diffusion supermartingale. To this end, we rely on an eigenfunction of the infinitesimal generator of $(\lambda_t)_{t\geqslant 0}$. Let us consider a positive function $\psi \in C_+^2(\mathbb{R}^+)$ that satisfies the following differential equation

$$\mathcal{A}\psi(\lambda) = \beta\psi(\lambda),$$

where β is a strictly positive constant and $\mathcal{A}$ is the infinitesimal generator operator of $(\lambda_t)_{t\geqslant 0}$. By Ito's lemma, the infinitesimal generator $\mathcal{A}$ is given by

$$\mathcal{A}f(\lambda) = \kappa(\theta-\lambda)\frac{\mathrm{d}f}{\mathrm{d}\lambda} + \lambda\frac{\sigma^2}{2}\frac{\mathrm{d}^2 f}{\mathrm{d}\lambda^2} + \lambda\int_{\mathbb{R}^+}(f(\lambda+\eta z)-f(\lambda))\nu(\mathrm{d}z).$$

The function ψ is thus an eigenfunction of the infinitesimal generator $\mathcal{A}$, with associated eigenvalue β. By construction, the process $(M_t)_{t\geqslant 0}$ defined as $M_t = e^{-\beta t}\psi(\lambda_t)$ is a local-martingale. Indeed, according to Ito's lemma, we have for $t \geqslant u \geqslant 0$,

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[M_t - M_u|\mathcal{F}_u^{\xi}\right] &= \mathbb{E}_{\mathbb{Q}}\left[-\beta\int_u^t e^{-\beta s}\psi(\lambda_{s-})\mathrm{d}s + \int_u^t e^{-\beta s}\mathcal{A}\psi(\lambda_{s-})\mathrm{d}s\big|\mathcal{F}_u^{\xi}\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[-\beta\int_u^t e^{-\beta s}\psi(\lambda_{s-})\mathrm{d}s + \int_u^t e^{-\beta s}\beta\psi(\lambda_{s-})\mathrm{d}s\big|\mathcal{F}_u^{\xi}\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[-\beta\int_u^t M_{s-}\mathrm{d}s + \beta\int_u^t M_{s-}\mathrm{d}s\big|\mathcal{F}_u^{\xi}\right] = 0.
\end{aligned}$$

As a consequence, $\mathbb{E}_{\mathbb{Q}}[M_t|\mathcal{F}_u^{\xi}] = M_u$ whenever $u \leqslant t$ and thus

$$\mathbb{E}_{\mathbb{Q}}\left[\psi(\lambda_t)|\mathcal{F}_u^{\xi}\right] = e^{\beta(t-u)}\psi(\lambda_u).$$

It follows that $a > \beta$ implies that the process $(\xi_t)_{t\geqslant 0}$ that satisfies $\xi_t = e^{-at}\psi(\lambda_t)$ for all $t \geqslant 0$ is a supermartingale. As a matter of fact,

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[\xi_t|\mathcal{F}_u^{\xi}] &= \mathbb{E}_{\mathbb{Q}}\left[e^{-at}\psi(\lambda_t)|\mathcal{F}_u^{\xi}\right] \\
&= e^{-at+\beta(t-u)}\psi(\lambda_u) \\
&< e^{-au}\psi(\lambda_u) = \xi_u.
\end{aligned} \tag{3.11}$$

Moreover, it is also clear from Equation (3.11) that $\lim_{t\to+\infty}\mathbb{E}_{\mathbb{Q}}[\xi_t] = 0$. We call the process $(\xi_t)_{t\geqslant 0}$ a *Hawkes diffusion supermartingale*. Thanks to its properties, this process is eligible for defining the term structure of default probabilities. The next step is thus to find such an eigenfunction ψ. The following result shows that ψ can be taken to be an exponential function.

Proposition 3.1. *Let γ be the constant defined in Corollary 3.1. Then the function ψ defined by*

$$\psi(\lambda) = e^{\gamma\lambda} \tag{3.12}$$

is an eigenfunction of the infinitesimal generator $\mathcal{A}$ with associated eigenvalue $\beta = \gamma\kappa\theta$.

Proof. Assuming that ψ satisfies Equation (3.12), Lemma 3.1 yields

$$\mathcal{A}\psi(\lambda) = \gamma\kappa\theta\psi(\lambda) + q(\gamma)\lambda\psi(\lambda),$$

where the function q is given at Equation (3.7). According to Lemma 3.2, $q(\gamma) = 0$, so that we further have

$$\mathcal{A}\psi(\lambda) = \gamma\kappa\theta\psi(\lambda).$$

We conclude that the function ψ is an eigenfunction of $\mathcal{A}$ with eigenvalue $\gamma\kappa\theta$. $\square$

Corollary 3.2. *For any $z \in (0, \gamma)$ and $w > \beta = \gamma\kappa\theta$, the process $(\xi_t)_{t\geqslant 0}$ defined by $e^{-wt+z\lambda_t}$ is a positive $(\mathbb{Q}, \mathbb{F}^\xi)$-supermartingale whose expectation converges to 0.*

Proof. Corollary 3.1 implies that when $t \geqslant s \geqslant 0$, we have

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}[\xi_t|\mathcal{F}^\xi_s] &= \mathbb{E}_{\mathbb{Q}}[e^{-wt+z\lambda_t}|\mathcal{F}^\xi_s] \\ &\leqslant e^{-wt+z\lambda_s+\beta(t-s)} \\ &= \xi_s e^{-(w-\beta)(t-s)} \\ &< \xi_s \end{aligned} \tag{3.13}$$

so that $(\xi_t)_{t\geqslant 0}$ is a supermartingale. The convergence of its expectation follows from Equation (3.13) and $e^{-(w-\beta)(t-s)} \to 0$ when $t \to +\infty$, since $w - \beta > 0$. $\square$

If we consider the supermartingale $(\xi_t)_{t\geqslant 0}$ and write τ for the time of default, the potential approach consists in assuming that the term structure of default probabilities is determined by

$$\mathbb{Q}(\tau > t|\tau > s) = \frac{e^{-wt}\mathbb{E}_{\mathbb{Q}}[e^{z\lambda_t}|\mathcal{F}^\xi_s]}{e^{-ws+z\lambda_s}}. \tag{3.14}$$

The next section introduces the fractional version of the model.

3.4 Fractional Hawkes jump-diffusion supermartingales

In this section, we introduce a fractional Hawkes jump-diffusion supermartingale. This process is obtained by time-changing the supermartingale $(\xi_t)_{t\geqslant 0}$. We start by introducing the time-change process and the considered filtrations. Afterwards, we prove that the supermartingale $(\xi_t)_{t\geqslant 0}$ remains a supermartingale under the time-change and that it keeps all its desired properties.

Let $(U_t)_{t\geqslant 0}$ be an α-stable subordinator on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$, with $\alpha \in (0, 1)$. The parameter α will be referred to as the *fractional order* in the sequel. Define its inverse $(S_t)_{t\geqslant 0}$ as

$$S_t := \inf\{\tau \geqslant 0 : U_\tau \geqslant t\}.$$

We denote by $\mathbb{F}^U = (\mathcal{F}_t^U)_{t\geqslant 0}$ the natural filtration of $(U_t)_{t\geqslant 0}$, which we assume to be independent from $\mathbb{F}^\xi$. The supermartingale $(\xi_t)_{t\geqslant 0}$ will be time-changed by the inverse process $(S_t)_{t\geqslant 0}$. The inverse process $(S_t)_{t\geqslant 0}$ possesses the property of "randomly freezing time". This property can be observed on Figure 3.1. This figure displays corresponding sample paths of the α-stable process $(U_t)_{t\geqslant 0}$ and its inverse $(S_t)_{t\geqslant 0}$ for some levels of the fractional order α. One can see that the frozen time periods of $(S_t)_{t\geqslant 0}$ (i.e. the periods of time for which the process remains constant) correspond to the jumps of the original process $(U_t)_{t\geqslant 0}$. Also note that by definition of $(S_t)_{t\geqslant 0}$, representing the path $[0, 10] \ni t \mapsto S_t(\omega)$ only requires the values of $t \mapsto U_t(\omega)$ until it reaches the value 10. This is why the graphs on the right-side are truncated at $y = 10$. Finally, note that the values of the fractional order close to 0 give the wildest behaviour for $(U_t)_{t\geqslant 0}$, i.e. paths with very large jumps. This translates into a lot of long motionless periods for the inverse $(S_t)_{t\geqslant t}$ when α is close to 0.

Let us now introduce precisely the filtration we consider for the time-changed supermartingale $(\xi_{S_t})_{t\geqslant 0}$. Define the filtration $\mathbb{F}^{\xi,U} = (\mathcal{F}_t^{\xi,U})_{t\geqslant 0}$ as $\mathcal{F}_t^{\xi,U} = \mathcal{F}_t^\xi \vee \mathcal{F}_t^U$. By definition of $(S_t)_{t\geqslant 0}$, for any $u, s \geqslant 0$ it holds that $\{S_u \leqslant s\} = \{U_s \geqslant u\}$, which clearly belongs to $\mathcal{F}_s^U$. As a consequence, for any fixed $u \geqslant 0$, S_u is a $\mathbb{F}^U$-stopping-time, and thus also a $\mathbb{F}^{\xi,U}$-stopping-time. We can therefore define the filtration $\mathbb{G} = (\mathcal{G}_t)_{t\geqslant 0}$ as the collection of stopped σ-algebras

$$\mathcal{G}_t := \mathcal{F}_{S_t}^{\xi,U} = \{A \in \mathcal{F}_\infty^{\xi,U} : A \cap \{S_t \leqslant s\} \in \mathcal{F}_s^{\xi,U} \text{ for all } s \geqslant 0\},$$

with $\mathcal{F}_\infty^{\xi,U} = \bigvee_{t\geqslant 0} \mathcal{F}_t^{\xi,U}$. We show now that the obtained process $(\xi_{S_t})_{t\geqslant 0}$ remains a supermartingale and can therefore be used as in the previous section to define a term structure of default probabilities.

Proposition 3.2. *The process $(\xi_{S_t})_{t\geqslant 0}$ is a positive $(\mathbb{Q}, \mathbb{G})$-supermartingale that satisfies $\lim_{t\to+\infty} \mathbb{E}_\mathbb{Q}[\xi_{S_t}] = 0$.*

Proof. By Corollary 3.2, $(\xi_t)_{t\geqslant 0}$ is a $(\mathbb{Q}, \mathbb{F}^\xi)$-supermartingale. Since the filtrations $\mathbb{F}^\xi$ and $\mathbb{F}^U$ are independent, it holds that

$$\mathbb{E}_\mathbb{Q}[\xi_t | \mathcal{F}_s^\xi \vee \mathcal{F}_s^U] = \mathbb{E}_\mathbb{Q}[\xi_t | \mathcal{F}_s^\xi]$$

for any $s, t \geqslant 0$. The property we use in the previous equation can be found in the chapter 9.7 of Williams (1991), at property (k). As a consequence, $(\xi_t)_{t\geqslant 0}$ is also a $(\mathbb{Q}, \mathbb{F}^{\xi,U})$-supermartingale. Then, since for any $s \leqslant t$, we have $S_s \leqslant S_t$ $\mathbb{Q}$-a.s., it follows from the Optional Stopping Theorem for supermartingales (see e.g. Theorem 3.3 in the chapter 2 of Revuz and Yor (2005)) that

$$\xi_{S_s} \geqslant \mathbb{E}_\mathbb{Q}[\xi_{S_t} | \mathcal{F}_{S_s}^{U,\xi}] = \mathbb{E}_\mathbb{Q}[\xi_{S_t} | \mathcal{G}_s].$$

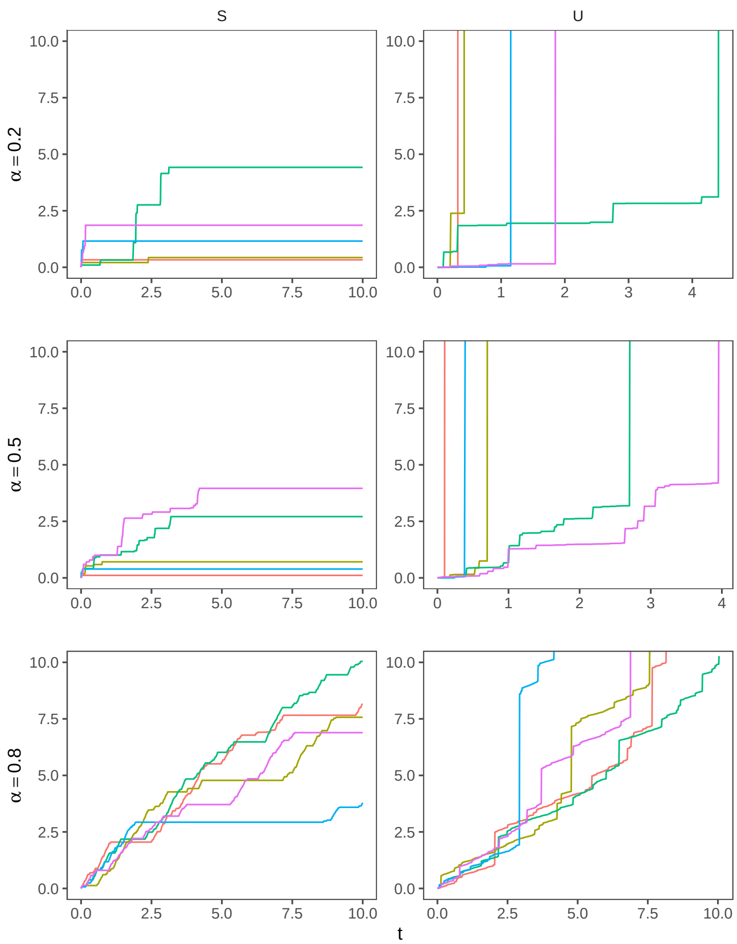

Figure 3.1: Five simulated paths of $(U_t)_{t\geqslant 0}$ and corresponding paths of $(S_t)_{t\geqslant 0}$ for fractional orders $\alpha \in \{0.2, 0.5, 0.8\}$.

This proves that $(\xi_{S_t})_{t\geqslant 0}$ is a $(\mathbb{Q},\mathbb{G})$-supermartingale. Recall that from Piryatinska et al. (2004),

$$\mathbb{E}_{\mathbb{Q}}[e^{-\omega S_t}] = \mathrm{E}_\alpha(-\omega t^\alpha),$$

where the function E_α is the Mittag-Leffler function

$$\mathrm{E}_\alpha(z) = \sum_{k=0}^{+\infty} \frac{z^k}{\Gamma(k\alpha+1)}.$$

As a consequence, from Corollary 3.1 and Tonelli's theorem,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[\xi_{S_t}] &= \int_{\mathbb{R}^+} \mathbb{E}_{\mathbb{Q}}[\xi_u] g(t,u)\mathrm{d}u \\ &\leqslant \int_{\mathbb{R}^+} e^{(\beta-w)u} g(t,u)\mathrm{d}u \\ &= \mathrm{E}_\alpha((\beta-w)t^\alpha),\end{aligned}$$

where $u \mapsto g(t,u)$ denotes the probability density function of S_t, $t>0$. This proves the convergence of the expectation to 0 and ends the proof. □

The supermartingale $(\xi_{S_t})_{t\geqslant 0}$ inherits from $(S_t)_{t\geqslant 0}$ the property of having motionless periods, during which the time appears to be frozen. The term structure of default probabilities is defined as follows: let $\tau^{(\alpha)}$ be the time of default in the fractional model. Then we assume that the term structure of default probabilities satisfies

$$\mathbb{Q}(\tau^{(\alpha)} > t | \tau^{(\alpha)} > s) = \frac{\mathbb{E}_{\mathbb{Q}}[e^{-wS_t + z\lambda_{S_t}} | \mathcal{G}_s]}{e^{-wS_s + z\lambda_{S_s}}}, \tag{3.15}$$

for $t \geqslant s \geqslant 0$.

The next section addresses the problem of computing the default probabilities, in both the non-fractional and fractional settings.

3.5 Computation of the default probabilities

The main challenge in computing the default probabilities of Equations (3.14) and (3.15) is the computation of $\mathbb{E}_{\mathbb{Q}}[e^{z\lambda_t}|\mathcal{F}_s]$ and $\mathbb{E}_{\mathbb{Q}}[e^{-wS_t+z\lambda_{S_t}}|\mathcal{F}_{S_s}]$, $t \geqslant s \geqslant 0$. To this end, several approaches can be thought of. A first possible approach is the one of Duffie et al. (2000). It consists in showing that the expected value $\mathbb{E}_{\mathbb{Q}}[e^{z\lambda_t}|\mathcal{F}_s]$ can be written as $\exp\{a(t-s) + \lambda_s b(t-s)\}$ for some functions a, b and that these functions satisfy a system of differential equations. The expected value can then be found by solving this system numerically. In our context, the main drawback of this approach is that it does not transpose very well to the fractional model. As a matter of fact, obtaining the default probabilities in the fractional setting would require to compute numerically the integral

$$\int_0^{+\infty} \mathbb{E}_{\mathbb{Q}}[e^{z\lambda_\tau}]g(t,\tau)\mathrm{d}\tau = \int_0^{+\infty} \exp\{a(\tau)+\lambda_0 b(\tau)\}g(t,\tau)\mathrm{d}\tau,$$

where $\tau \mapsto g(t,\tau)$ is the probability density function of S_t. Then one could use Proposition 3.2 in Gupta and Kumar (2022) that says that

$$g(t,\tau) = \frac{1}{\pi}\sum_{k=0}^{+\infty}\frac{(-\tau)^k}{k!}\Gamma((1+k)\alpha)t^{-(k+1)\alpha}\sin(\pi\alpha(k+1)), \quad \tau > 0$$

to evaluate the integral. This first approach seems to be computationally heavy, as it requires several layers of numerical approximations. A second approach is to derive a partial differential equation for the transform $(t,z) \mapsto \mathbb{E}_{\mathbb{Q}}[e^{z\lambda_t}]$. This approach is used in e.g. Ketelbuters and Hainaut (2022a) and has the benefit of having a natural extension to the fractional case. As a matter of fact, the transform $(t,z) \mapsto \mathbb{E}_{\mathbb{Q}}[e^{z\lambda_{S_t}}]$ satisfies a fractional partial differential equation. Compared to the non-fractional case, the only difference is that the partial derivative with respect to time is a Caputo fractional derivative of order α. In both cases, the (fractional) partial differential equation can be solved numerically with a finite difference method. However, this second approach suffers from two drawbacks. The first drawback is that we have to use fractional calculus, which is rather complicated from a mathematical standpoint. The second drawback is that fractional partial differential equations are far from being the easiest equations to solve numerically. Indeed, performing finite difference approximations for fractional derivatives requires, at each iteration, to use the computed values for all the previous iterations. This substantially increases the computation time. This is why we propose here a third approach that rely on a recursive method to compute all the moments of λ_t. To the best of our knowledge, this recursive method is new. Once the moments $\mathbb{E}_{\mathbb{Q}}[\lambda_t^k]$ are computed for all nonnegative integers $k \leqslant n$, we can approximate the expectation $\mathbb{E}_{\mathbb{Q}}[e^{z\lambda_t}]$ with the help of the MacLaurin series expansion of the exponential function, that is

$$\mathbb{E}_{\mathbb{Q}}[e^{z\lambda_t}] \approx \sum_{k=0}^{n}\frac{\mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^k]}{k!}.$$

This third approach avoids the shortcomings of the two other ones. Firstly, it is much faster computationally. Secondly, it has an easy extension to the fractional case, without relying on the complicated machinery of fractional calculus. That is, once the moments of λ_t are obtained, the moments of the time-changed process λ_{S_t} can be found very easily. Thirdly, the approximation can be shown to be convergent and a bound on the numerical error can be provided in both the non-fractional and fractional cases.

This section is organized as follows: we start by introducing and proving the recursive method that allows to compute the moments of λ_t. Then, as a corollary to this method, we show that the moments of the time-changed process λ_{S_t} can be obtained in a similar manner. Finally, after introducing the approximations for the survival probabilities of Equations (3.14) and (3.15), we show that these approximations are convergent and give an upper bound on the associated numerical error.

The next result gives the recursive method for the moments of λ_t.

Proposition 3.3. *Let $z > 0$. For each $n \geqslant 1$, the moment $\mathbb{R}^+ \ni t \mapsto \mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^n]$ satisfies*

$$\mathbb{E}_{\mathbb{Q}}[(z\lambda_t^n)] = \left(\sum_{j=1}^{n} d_{j,n} e^{-jat}\right) + c_n$$

with $a = \kappa - (\eta/\rho)$, for all $n \in \mathbb{N}$. Moreover, the constants $d_{k,n}, c_n$ can be computed recursively with the relations

$$d_{1,1} = z\lambda_0 - \frac{\kappa z\theta}{a}, \qquad c_1 = \frac{\kappa z\theta}{a},$$

$$d_{j,n} = \begin{cases} -\left(\frac{b_n d_{j,n-1}}{(j-n)a} + \sum_{p=2}^{n-j+1} z^{-1} m_p \binom{n}{p} \frac{d_{j,n-p+1}}{(j-n)a}\right) & j < n, \\ (z\lambda_0)^n - \left(\sum_{p=2}^{n} z^{-1} m_p \binom{n}{p} \frac{c_{n-p+1}}{na}\right) - \frac{b_n c_{n-1}}{na} & \\ \qquad + \sum_{k=1}^{n-1} \left(\frac{b_n d_{k,n-1}}{(k-n)a} + \sum_{p=2}^{n-k+1} z^{-1} m_p \binom{n}{p} \frac{d_{k,n-p+1}}{(k-n)a}\right) & j = n, \end{cases}$$

$$c_n = \frac{b_n c_{n-1}}{na} + \sum_{p=2}^{n} z^{-1} m_p \binom{n}{p} \frac{c_{n-p+1}}{na}.$$

The symbol b_n is used as a shorthand for

$$b_n = n\kappa z\theta + \frac{n(n-1)z\sigma^2}{2}$$

and $m_p = \mathbb{E}_{\mathbb{Q}}[(z\eta J)^p] = \left(\frac{z\eta}{\rho}\right)^p p!$.

Proof. Define the process $(\tilde{\lambda}_t)_{t\geqslant 0}$ as $\tilde{\lambda}_t := z\lambda_t$. This process satisfies the SDE

$$\mathrm{d}\tilde{\lambda}_t = \kappa(\tilde{\theta} - \tilde{\lambda}_{t-})\mathrm{d}t + \tilde{\sigma}\sqrt{\tilde{\lambda}_{t-}}\mathrm{d}W_t + \tilde{\eta}\mathrm{d}P_t$$

where $\tilde{\theta} = z\theta$, $\tilde{\sigma} = \sqrt{z}\sigma$ and $\tilde{\eta} = z\eta$. Note that the jump intensity of $(P_t)_{t\geqslant 0}$ remains $(\lambda_t)_{t\geqslant 0} = (z^{-1}\tilde{\lambda}_t)_{t\geqslant 0}$. As a consequence, Ito's lemma applied on the process $(\tilde{\lambda}_t^n)_{t\geqslant 0}$ yields

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[\tilde{\lambda}_t^n\right] - \tilde{\lambda}_0^n &= \mathbb{E}_{\mathbb{Q}}\Bigg[\int_0^t n\tilde{\lambda}_{s-}^{n-1}\left(\kappa(\tilde{\theta}-\tilde{\lambda}_{s-})\mathrm{d}s + \tilde{\sigma}\sqrt{\tilde{\lambda}_{s-}}\mathrm{d}W_s\right) \\
&\qquad + \frac{n(n-1)}{2}\int_0^t \tilde{\lambda}_{s-}^{n-2}\tilde{\lambda}_{s-}\tilde{\sigma}^2\mathrm{d}s + \int_0^t \left(\tilde{\lambda}_s^n - \tilde{\lambda}_{s-}^n\right)\mathrm{d}N_s\Bigg] \\
&= \mathbb{E}_{\mathbb{Q}}\Bigg[n\int_0^t \left(\kappa\tilde{\theta}\tilde{\lambda}_{s-}^{n-1} - \kappa\tilde{\lambda}_{s-}^n\right)\mathrm{d}s \\
&\qquad + \frac{n(n-1)}{2}\int_0^t \tilde{\lambda}_{s-}^{n-1}\tilde{\sigma}^2\mathrm{d}s + \int_0^t \left(\tilde{\lambda}_s^n - \tilde{\lambda}_{s-}^n\right)z^{-1}\tilde{\lambda}_{s-}\mathrm{d}s\Bigg]
\end{aligned}$$

It follows that

$$\begin{aligned}
&\frac{\mathrm{d}\mathbb{E}_{\mathbb{Q}}\left[\tilde{\lambda}_t^n\right]}{\mathrm{d}t} \\
&= \mathbb{E}_{\mathbb{Q}}[\tilde{\lambda}_t^{n-1}]\left(n\kappa\tilde{\theta} + \frac{n(n-1)\tilde{\sigma}^2}{2}\right) - n\kappa\mathbb{E}_{\mathbb{Q}}[\tilde{\lambda}_s^n] + \mathbb{E}_{\mathbb{Q}}\left[\left((\tilde{\lambda}_t + \tilde{\eta}J)^n - \tilde{\lambda}_t^n\right)z^{-1}\tilde{\lambda}_t\right] \\
&= \mathbb{E}_{\mathbb{Q}}[\tilde{\lambda}_t^{n-1}]\left(n\kappa\tilde{\theta} + \frac{n(n-1)\tilde{\sigma}^2}{2}\right) - n\kappa\mathbb{E}_{\mathbb{Q}}[\tilde{\lambda}_s^n] \\
&\qquad + \mathbb{E}_{\mathbb{Q}}\left[\left(\sum_{p=0}^n \binom{n}{p}(\tilde{\eta}J)^p\tilde{\lambda}_t^{n-p} - \tilde{\lambda}_t^n\right)z^{-1}\tilde{\lambda}_t\right] \\
&= \mathbb{E}_{\mathbb{Q}}[\tilde{\lambda}_t^{n-1}]\left(n\kappa\tilde{\theta} + \frac{n(n-1)\tilde{\sigma}^2}{2}\right) - n\kappa\mathbb{E}_{\mathbb{Q}}[\tilde{\lambda}_s^n] + \sum_{p=1}^n \binom{n}{p}z^{-1}m_p\mathbb{E}_{\mathbb{Q}}\left[\tilde{\lambda}_t^{n-p+1}\right].
\end{aligned}$$

Therefore, if we define the functions $(f_n)_{n\in\mathbb{N}}$, $f_n : \mathbb{R}^+ \to \mathbb{R}^+$ as $f_n(t) = \mathbb{E}_{\mathbb{Q}}\left[\tilde{\lambda}_t^n\right]$, then each of these functions satisfies the first order linear differential equation

$$f_n'(t) = -naf_n(t) + h_{n-1}(t), \tag{3.16}$$

where

$$h_{n-1}(t) = \begin{cases} b_n f_{n-1}(t) & \text{if } n = 1, \\ b_n f_{n-1}(t) + \sum_{p=2}^n \binom{n}{p} z^{-1} m_p f_{n-p+1}(t) & \text{if } n \geqslant 2. \end{cases}$$

Equation (3.16) is easy to solve. Multiplying both sides by e^{nat} leads to

$$e^{nat}f_n'(t) + nae^{nat}f_n(t) = e^{nat}h_{n-1}(t)$$

and since the left-hand side is $(e^{nat}f_n(t))'$, an integration yields

$$f_n(t) = e^{-nat}\tilde{\lambda}_0^n + e^{-nat}\int_0^t e^{nas}h_{n-1}(s)\mathrm{d}s. \tag{3.17}$$

We prove now by induction that the functions $(f_n)_{n\in\mathbb{N}}$ are of the form

$$f_n(t) = \left(\sum_{j=1}^{n} d_{j,n}e^{-jat}\right) + c_n \tag{3.18}$$

and that the numbers $d_{j,n}, c_n$ can be computed recursively, as in the statement. From Equation (3.17), we find that

$$f_1(t) = \left(\tilde{\lambda}_0 - \frac{\kappa\tilde{\theta}}{a}\right)e^{-at} + \frac{\kappa\tilde{\theta}}{a}$$

so that Equation (3.18) holds for $n = 1$ with $d_{1,1} = \tilde{\lambda}_0 - \kappa\tilde{\theta}/a$ and $c_1 = \kappa\tilde{\theta}/a$. Assume that Equation (3.18) holds for $1,\dots,n-1$ and that $n \geqslant 2$. Using Equations (3.17) and (3.18), we can compute

$$\begin{aligned}
\int_0^t e^{nas}h_{n-1}(s)\mathrm{d}s &= b_n\int_0^t e^{nas}\left[\left(\sum_{j=1}^{n-1} d_{j,n-1}e^{-jas}\right) + c_{n-1}\right]\mathrm{d}s \\
&\quad + \sum_{p=2}^{n} z^{-1}m_p\binom{n}{p}\int_0^t e^{nas}\left[\left(\sum_{j=1}^{n-p+1} d_{j,n-p+1}e^{-jas}\right) + c_{n-p+1}\right]\mathrm{d}s \\
&= -\sum_{j=1}^{n-1} e^{-(j-n)at}\left(\frac{b_n d_{j,n-1}}{(j-n)a} + \sum_{p=2}^{n-j+1} z^{-1}m_p\binom{n}{p}\frac{d_{j,n-p+1}}{(j-n)a}\right) \\
&\quad + e^{nat}\left[\frac{b_n c_{n-1}}{na} + \sum_{p=2}^{n} z^{-1}m_p\binom{n}{p}\frac{c_{n-p+1}}{na}\right] - \sum_{p=2}^{n} z^{-1}m_p\binom{n}{p}\frac{c_{n-p+1}}{na} \\
&\quad + \sum_{j=1}^{n-1}\left[\frac{b_n d_{j,n-1}}{(j-n)a} + \sum_{p=2}^{n-j+1} z^{-1}m_p\binom{n}{p}\frac{d_{j,n-p+1}}{(j-n)a}\right] - \frac{b_n c_{n-1}}{na}.
\end{aligned}$$

As a consequence, we find

$$\begin{aligned}
f_n(t) = &-\sum_{j=1}^{n-1} e^{-jat}\left(\frac{b_n d_{j,n-1}}{(j-n)a} + \sum_{p=2}^{n-j+1} z^{-1} m_p \binom{n}{p} \frac{d_{j,n-p+1}}{(j-n)a}\right) \\
&+ e^{-nat}\left[\tilde{\lambda}_0^n - \frac{b_n c_{n-1}}{na} + \sum_{j=1}^{n-1}\left(\frac{b_n d_{j,n-1}}{(j-n)a} + \sum_{p=2}^{n-j+1} z^{-1} m_p \binom{n}{p} \frac{d_{j,n-p+1}}{(j-n)a}\right)\right. \\
&\left. - \left(\sum_{p=2}^{n} z^{-1} m_p \binom{n}{p} \frac{c_{n-p+1}}{na}\right)\right] + \left[\frac{b_n c_{n-1}}{na} + \sum_{p=2}^{n} z^{-1} m_p \binom{n}{p} \frac{c_{n-p+1}}{na}\right] \\
= &\left(\sum_{j=1}^{n} d_{j,n} e^{-jat}\right) + c_n,
\end{aligned}$$

where

$$d_{j,n} = \begin{cases} -\left(\frac{b_n d_{j,n-1}}{(j-n)a} + \sum_{p=2}^{n-j+1} z^{-1} m_p \binom{n}{p} \frac{d_{j,n-p+1}}{(j-n)a}\right) & \text{if } j < n \\ \tilde{\lambda}_0^n - \left(\sum_{p=2}^{n} z^{-1} m_p \binom{n}{p} \frac{c_{n-p+1}}{na}\right) - \frac{b_n c_{n-1}}{na} & \\ \quad + \sum_{k=1}^{n-1}\left(\frac{b_n d_{k,n-1}}{(k-n)a} + \sum_{p=2}^{n-k+1} z^{-1} m_p \binom{n}{p} \frac{d_{k,n-p+1}}{(k-n)a}\right) & \text{if } j = n \end{cases}$$

and

$$c_n = \frac{b_n c_{n-1}}{na} + \sum_{p=2}^{n} z^{-1} m_p \binom{n}{p} \frac{c_{n-p+1}}{na}.$$

This concludes the proof. □

Note that in the proof of the previous result, the method of using differential equations to find the moments of a Hawkes process intensity is not new, as this approach is used in e.g. Cui et al. (2020). The originality of the result lies in the observation that all the moments share the same form, i.e. a linear combination of exponential functions plus a constant, and that the associated coefficients can be computed recursively. According to the next corollary, it turns out that this result extends particularly well to the fractional case.

Corollary 3.3. *The moment* $\mathbb{R}^+ \ni t \mapsto \mathbb{E}_{\mathbb{Q}}[(z\lambda_{S_t})^n]$ *satisfy*

$$\mathbb{E}_{\mathbb{Q}}[(z\lambda_{S_t})^n] = \left(\sum_{j=1}^{n} d_{j,n} \mathrm{E}_\alpha(-jat^\alpha)\right) + c_n,$$

for all $n \in \mathbb{N}$. The function E_α is the Mittag-Leffler function

$$\mathrm{E}_\alpha(z) = \sum_{k=0}^{+\infty} \frac{z^k}{\Gamma(k\alpha + 1)}.$$

The constants $a, c_n, d_{j,n}$ are given in Proposition 3.3.

Proof. The main tool of this proof is a result from Piryatinska et al. (2004), which states that

$$\mathbb{E}_\mathbb{Q}[e^{-\omega S_t}] = \mathrm{E}_\alpha(-\omega t^\alpha).$$

As a consequence, by Proposition 3.3 and the law of iterated expectations, we have

$$\begin{aligned}
\mathbb{E}_\mathbb{Q}[(z\lambda_{S_t})^n] &= \left(\sum_{j=1}^{n} d_{j,n} \mathbb{E}_\mathbb{Q}\left[e^{-jaS_t}\right]\right) + c_n \\
&= \left(\sum_{j=1}^{n} d_{j,n} \mathrm{E}_\alpha(-jat^\alpha)\right) + c_n,
\end{aligned}$$

as announced. □

We now have the tools to approximate the survival probabilities. In the non-fractional case, we approximate the survival probabilities $t \mapsto q(t) := \mathbb{Q}(\tau > t)$ by the sequence of functions $(q_n)_{n\in\mathbb{N}}$ defined as

$$q_n(t) := \sum_{k=0}^{n} \frac{e^{-wt}\mathbb{E}_\mathbb{Q}[(z\lambda_t)^k]}{e^{z\lambda_0} k!}. \tag{3.19}$$

We prove now the convergence of this approximation and give a bound for the numerical error.

Proposition 3.4. *Let $z \in [0, \gamma]$. For each $t \geqslant 0$, the approximation $q_n(t)$ converges towards $q(t)$ as $n \to +\infty$. Moreover, if $z \leqslant \gamma/2$, the error $q(t) - q_n(t)$ satisfies the following bounds*

$$\begin{aligned}
&0 \leqslant q(t) - q_n(t) \\
&\leqslant \frac{\mathbb{E}_\mathbb{Q}\left[(z\lambda_t)^{n+1}\right] e^{\gamma(\lambda_0+\kappa\theta t)} + \sqrt{\mathbb{E}_\mathbb{Q}[(z\lambda_t)^{2n+2}] - (\mathbb{E}_\mathbb{Q}[(z\lambda_t)^{n+1}])^2}\sqrt{e^{\gamma(\lambda_0+\kappa\theta t)}}}{e^{wt+z\lambda_0}(n+1)!}.
\end{aligned}$$

Proof. Note that by positivity of $(\tilde{\lambda}_t)_{t\geqslant 0}$, the sequence of random variables $(\sum_{j=0}^{n} \tilde{\lambda}_t^j)_{n\in\mathbb{N}}$ is increasing. As a consequence, the monotone convergence theorem states that

$$\mathbb{E}_{\mathbb{Q}}[e^{\tilde{\lambda}_t}] = \mathbb{E}_{\mathbb{Q}}\left[\lim_{n\to+\infty}\sum_{j=0}^{n}\frac{\tilde{\lambda}_t^j}{j!}\right] = \lim_{n\to+\infty}\sum_{j=0}^{n}\frac{\mathbb{E}_{\mathbb{Q}}\left[\tilde{\lambda}_t^j\right]}{j!},$$

which already proves the convergence of the approximation. We now derive a bound on the error we commit by keeping the $n+1$ first terms in the series. This error is the expected value of $\sum_{k\geqslant n+1}((z\lambda_t)^k/k!)$, which satisfies

$$\begin{aligned}\sum_{k=n+1}^{+\infty}\frac{(z\lambda_t)^k}{k!} &= \sum_{k=0}^{+\infty}\frac{(z\lambda_t)^{k+n+1}}{(k+n+1)!}\\ &\leqslant \frac{(z\lambda_t)^{n+1}}{(n+1)!}\sum_{k=0}^{+\infty}\frac{(z\lambda_t)^k}{k!}\\ &= \frac{(z\lambda_t)^{n+1}e^{z\lambda_t}}{(n+1)!}\end{aligned}$$

where the inequality is a consequence of $(k+n+1)! \geqslant k!(n+1)!$. The proof is completed by observing that

$$\begin{aligned}&\mathbb{E}_{\mathbb{Q}}\left[(z\lambda_t)^{n+1}e^{z\lambda_t}\right] = \mathbb{E}_{\mathbb{Q}}\left[(z\lambda_t)^{n+1}\right]\mathbb{E}_{\mathbb{Q}}\left[e^{z\lambda_t}\right] + \mathbb{C}\mathrm{ov}_{\mathbb{Q}}((z\lambda_t)^{n+1}, e^{z\lambda_t})\\ &\leqslant \mathbb{E}_{\mathbb{Q}}\left[(z\lambda_t)^{n+1}\right]\mathbb{E}_{\mathbb{Q}}\left[e^{z\lambda_t}\right] + \sqrt{\mathbb{V}\mathrm{ar}_{\mathbb{Q}}((z\lambda_t)^{n+1})\mathbb{V}\mathrm{ar}_{\mathbb{Q}}(e^{z\lambda_t})}\\ &\leqslant \mathbb{E}_{\mathbb{Q}}\left[(z\lambda_t)^{n+1}\right]\mathbb{E}_{\mathbb{Q}}\left[e^{z\lambda_t}\right] + \sqrt{\mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^{2n+2}] - (\mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^{n+1}])^2}\sqrt{\mathbb{E}_{\mathbb{Q}}[e^{2z\lambda_t}]}\\ &\leqslant \mathbb{E}_{\mathbb{Q}}\left[(z\lambda_t)^{n+1}\right]\mathbb{E}_{\mathbb{Q}}\left[e^{\gamma\lambda_t}\right] + \sqrt{\mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^{2n+2}] - (\mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^{n+1}])^2}\sqrt{\mathbb{E}_{\mathbb{Q}}[e^{\gamma\lambda_t}]}\\ &= \mathbb{E}_{\mathbb{Q}}\left[(z\lambda_t)^{n+1}\right]e^{\gamma(\lambda_0+\kappa\theta t)} + \sqrt{\mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^{2n+2}] - (\mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^{n+1}])^2}\sqrt{e^{\gamma(\lambda_0+\kappa\theta t)}},\end{aligned} \tag{3.20}$$

where the last equality follows from Corollary 3.1. □

Note that the moments $\mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^{2n+2}]$ and $\mathbb{E}_{\mathbb{Q}}[(z\lambda_t)^{n+1}]$ that appear in the upper bound of Proposition 3.4 can be computed with the help of the recursive method of Proposition 3.3.

In the fractional case, we approximate the survival probabilities $t \mapsto \mathbb{Q}(\tau^{(\alpha)} > t)$ by the functions $(q_n^{(\alpha)})_{n\in\mathbb{N}}$ as

$$
\begin{aligned}
q_n^{(\alpha)}(t) &:= \sum_{k=0}^{n} \frac{\mathbb{E}_{\mathbb{Q}}\left[e^{-wS_t}(z\lambda_{S_t})^k\right]}{e^{z\lambda_0}k!} \\
&= \sum_{k=0}^{n} \frac{1}{e^{z\lambda_0}k!}\mathbb{E}_{\mathbb{Q}}\left[e^{-wS_t}\mathbb{E}_{\mathbb{Q}}\left[(z\lambda_{S_t})^k|\mathcal{F}_t^S\right]\right] \\
&= \sum_{k=0}^{n} \frac{1}{e^{z\lambda_0}k!}\mathbb{E}_{\mathbb{Q}}\left[e^{-wS_t}\left(c_k + \sum_{j=1}^{k} d_{j,k}e^{-jaS_t}\right)\right] \\
&= \sum_{k=0}^{n} \frac{1}{e^{z\lambda_0}k!}\mathbb{E}_{\mathbb{Q}}\left[c_k e^{-wS_t} + \sum_{j=1}^{k} d_{j,k}e^{-(w+ja)S_t}\right] \\
&= \sum_{k=0}^{n} \frac{1}{e^{z\lambda_0}k!}\left(c_k \mathrm{E}_\alpha\left(-wt^\alpha\right) + \sum_{j=1}^{k} d_{j,k}\mathrm{E}_\alpha\left(-(w+ja)t^\alpha\right)\right).
\end{aligned}
\tag{3.21}
$$

In a similar way, we can also prove the convergence of the approximation in the fractional case, and give a bound on the error.

Proposition 3.5. *Let $z \in [0,\gamma]$. For each $t \geqslant 0$, the approximation $q_n^{(\alpha)}(t)$ converges towards $q^{(\alpha)}(t)$ as $n \to +\infty$. Moreover, if $z \leqslant \gamma/2$, the error $q^{(\alpha)}(t) - q_n^{(\alpha)}(t)$ satisfies the following bounds*

$$
\begin{aligned}
0 &\leqslant q^{(\alpha)}(t) - q_n^{(\alpha)}(t) \\
&\leqslant \frac{e^{-z\lambda_0}}{(n+1)!}\left(\begin{aligned}&\mathbb{E}_{\mathbb{Q}}\left[(z\lambda_{S_t})^{n+1}\right]e^{\gamma\lambda_0}\mathrm{E}_\alpha((\gamma\kappa\theta - w)t^\alpha) \\ &\qquad + \sqrt{\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left((z\lambda_{S_t})^{n+1}\right)e^{\gamma\lambda_0}\mathrm{E}_\alpha((\gamma\kappa\theta - 2w)t^\alpha)}\end{aligned}\right).
\end{aligned}
$$

Proof. The convergence of $q_n^{(\alpha)}(t)$ towards $q^{(\alpha)}(t)$ follows from the monotone convergence theorem. As in the proof of Proposition 3.4,

$$
\sum_{k=n+1}^{+\infty} \frac{(z\lambda_{S_t})^k}{k!} \leqslant \frac{(z\lambda_{S_t})^{n+1}e^{z\lambda_{S_t}}}{(n+1)!}.
$$

As a consequence,

$$\begin{aligned}
q^{(\alpha)}(t) - q_n^{(\alpha)}(t) &\leqslant \frac{e^{-z\lambda_0}}{(n+1)!}\mathbb{E}_{\mathbb{Q}}\left[(z\lambda_{S_t})^{n+1}e^{-wS_t+z\lambda_{S_t}}\right] \\
&\leqslant \frac{e^{-z\lambda_0}}{(n+1)!}\Bigg(\mathbb{E}_{\mathbb{Q}}\left[(z\lambda_{S_t})^{n+1}\right]\mathbb{E}_{\mathbb{Q}}\left[e^{-wS_t+z\lambda_{S_t}}\right] \\
&\qquad + \sqrt{\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left((z\lambda_{S_t})^{n+1}\right)\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left(e^{-wS_t+z\lambda_{S_t}}\right)}\Bigg) \\
&\leqslant \frac{e^{-z\lambda_0}}{(n+1)!}\Bigg(\mathbb{E}_{\mathbb{Q}}\left[(z\lambda_{S_t})^{n+1}\right]\mathbb{E}_{\mathbb{Q}}\left[e^{-wS_t+z\lambda_{S_t}}\right] \\
&\qquad + \sqrt{\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left((z\lambda_{S_t})^{n+1}\right)\mathbb{E}_{\mathbb{Q}}\left(e^{-2wS_t+2z\lambda_{S_t}}\right)}\Bigg) \\
&\leqslant \frac{e^{-z\lambda_0}}{(n+1)!}\Bigg(\mathbb{E}_{\mathbb{Q}}\left[(z\lambda_{S_t})^{n+1}\right]\mathbb{E}_{\mathbb{Q}}\left[e^{-wS_t+\gamma\lambda_{S_t}}\right] \\
&\qquad + \sqrt{\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left((z\lambda_{S_t})^{n+1}\right)\mathbb{E}_{\mathbb{Q}}\left[e^{-2wS_t+\gamma\lambda_{S_t}}\right]}\Bigg).
\end{aligned}$$

Moreover,

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[e^{-wS_t+\gamma\lambda_{S_t}}\right] &= \int_0^{+\infty} e^{-w\tau}\mathbb{E}_{\mathbb{Q}}\left[e^{\gamma\lambda_\tau}\right]g(t,\tau)\mathrm{d}\tau \\
&= e^{\gamma\lambda_0}\int_0^{+\infty} e^{(\gamma\kappa\theta-w)\tau}g(t,\tau)\mathrm{d}\tau \\
&= e^{\gamma\lambda_0}\mathrm{E}_\alpha((\gamma\kappa\theta-w)t^\alpha)
\end{aligned}$$

and similarly

$$\mathbb{E}_{\mathbb{Q}}\left[e^{-2wS_t+\gamma\lambda_{S_t}}\right] = e^{\gamma\lambda_0}\mathrm{E}_\alpha((\gamma\kappa\theta-2w)t^\alpha).$$

The result follows. □

Again, the moments that appear in the upper bound of Proposition 3.5 can be computed with the recursive method of Corollary 3.3.

3.6 Numerical experiments

This section presents the numerical results we obtained with the methods presented above. The model is calibrated on real data taken from Bloomberg on 25 January 2021. These data consist of survival probabilities for American Airlines. Note that these data are the same as in Chapter 2, so that we can compare the fit for both models. The model is calibrated in both settings, i.e. non-fractional and fractional. The parameters are found by minimizing the sum of squared errors, i.e.

$$\sum_{T=1}^{10}(Prob_T^{\mathrm{Model}} - Prob_T^{\mathrm{Market}})^2$$

Table 3.1: Parameters used for the computations. NF stands for the non-fractional model whereas F stands for the fractional model.

	θ	σ	η	ρ	κ	λ_0	γ	z	w	α
NF	0.0551	0.1051	9.5599	3.1870	7.7522	0.0551	0.2044	0.1022	0.2434	1.0000
F	0.6751	0.0999	9.9308	2.3752	6.1281	0.6751	0.0760	0.0380	0.3144	0.9083

where $Prob_T^{\text{Market}}$ denotes the market survival up to time T probability from Bloomberg and $Prob_T^{\text{Model}}$ denotes the corresponding survival probability implied by our model. The optimal parameters obtained for both models are given in Table 3.1 whereas Figure 3.2 shows the graphical result of the calibration. In particular, Figure 3.2 shows that the fit is as good as in Ketelbuters and Hainaut (2022a), see Figure 2.4 in the second chapter of this thesis. As explained above, this is achieved with a model that is easier to deal with both mathematically and numerically. Table 3.2 shows the numerical convergence of our method and Table 3.3 shows the values of the upper bounds that were given at Propositions 3.4 and 3.5. Finally, Figure 3.3 shows the different shapes we can achieve through the choice of the fractional order α, when all the other parameters are kept unchanged and equal to their value in Table 3.1.

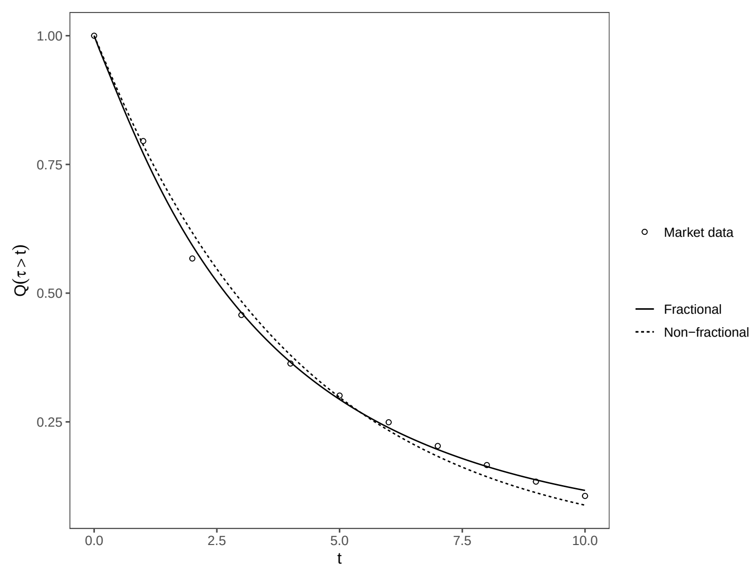

Figure 3.2: Results of the calibration.

To conclude this section about numerical results, we propose two additional examples

Table 3.2: Convergence of the computed probabilities and market survival probabilities.

	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$	$t = 9$	$t = 10$
Market data	0.795400	0.567300	0.457600	0.363300	0.301100	0.249200	0.203200	0.166100	0.133900	0.106100
Non-fractional										
$n = 2$	0.787450	0.617377	0.484019	0.379467	0.297500	0.233237	0.182856	0.143358	0.112392	0.088114
$n = 5$	0.787809	0.617661	0.484242	0.379642	0.297636	0.233345	0.182941	0.143424	0.112443	0.088155
$n = 10$	0.787835	0.617681	0.484258	0.379654	0.297646	0.233352	0.182947	0.143429	0.112447	0.088158
$n = 15$	0.787835	0.617682	0.484258	0.379655	0.297646	0.233353	0.182947	0.143429	0.112447	0.088158
$n = 20$	0.787835	0.617682	0.484258	0.379655	0.297646	0.233353	0.182947	0.143429	0.112447	0.088158
Fractional										
$n = 2$	0.766549	0.588248	0.458572	0.362939	0.291225	0.236678	0.194678	0.161986	0.136285	0.115891
$n = 5$	0.771007	0.592472	0.462013	0.365698	0.293448	0.238485	0.196162	0.163218	0.137318	0.116765
$n = 10$	0.771297	0.592780	0.462269	0.365904	0.293615	0.238621	0.196273	0.163310	0.137395	0.116830
$n = 15$	0.771301	0.592785	0.462274	0.365908	0.293618	0.238623	0.196276	0.163312	0.137396	0.116832
$n = 20$	0.771301	0.592785	0.462274	0.365908	0.293618	0.238623	0.196276	0.163312	0.137396	0.116832

Table 3.3: Bounds on the numerical error.

	$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$	$t = 6$	$t = 7$	$t = 8$	$t = 9$	$t = 10$
Non-fractional										
$n = 5$	0.0096	0.0079	0.0065	0.0053	0.0043	0.0036	0.0029	0.0024	0.0020	0.0016
$n = 10$	0.0061	0.0050	0.0041	0.0034	0.0028	0.0023	0.0019	0.0015	0.0012	0.0010
$n = 20$	0.0038	0.0031	0.0026	0.0021	0.0017	0.0014	0.0011	0.0009	0.0008	0.0006
$n = 30$	0.0028	0.0023	0.0019	0.0016	0.0013	0.0011	0.0009	0.0007	0.0006	0.0005
$n = 40$	0.0023	0.0019	0.0016	0.0013	0.0010	0.0009	0.0007	0.0006	0.0005	0.0004
Fractional										
$n = 5$	0.0338	0.0377	0.0350	0.0317	0.0286	0.0259	0.0236	0.0216	0.0199	0.0184
$n = 10$	0.0192	0.0239	0.0226	0.0205	0.0185	0.0168	0.0153	0.0140	0.0128	0.0119
$n = 20$	0.0098	0.0146	0.0142	0.0130	0.0118	0.0107	0.0097	0.0089	0.0082	0.0076
$n = 30$	0.0062	0.0108	0.0107	0.0099	0.0090	0.0081	0.0074	0.0068	0.0062	0.0058
$n = 40$	0.0064	0.0087	0.0078	0.0080	0.0073	0.0072	0.0060	0.0055	0.0049	0.0045

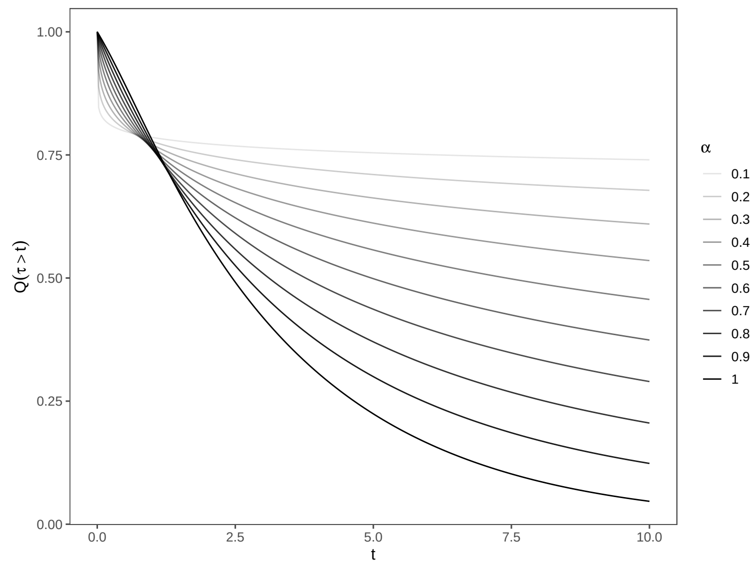

Figure 3.3: Survival probabilities for various fractional orders α.

of calibrations to real data. In the first example, we show that the fractional model is also particularly suited for CDS on countries, and especially countries during war times. The first example is a calibration on default probabilities for the Russian Federation, taken from Bloomberg on the 20^{th} September 2022. The results are displayed at Tables 3.4 and 3.5, as well as on Figure 3.4. We can observe here that the fractional model performs much better than the non-fractional model, the latter being completely unable to reproduce the survival probability shape.

The second additional example shows a case in which neither the fractional model nor the non-fractional one can reproduce the survival probability shape. The data are default probabilities for Microsoft taken from Bloomberg on the 20^{th} September

Table 3.4: Russia: market survival probabilities and results of the calibration. M, F and NF respectively stand for the market data, the fractional model and the non-fractional model.

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$	$t=8$	$t=9$	$t=10$
M	0.3919	0.3735	0.3563	0.3401	0.3247	0.3099	0.2956	0.2818	0.2685	0.2555
F	0.4115	0.3675	0.3427	0.3257	0.3129	0.3026	0.2940	0.2868	0.2804	0.2749
NF	0.7785	0.6049	0.4700	0.3652	0.2837	0.2204	0.1713	0.1331	0.1034	0.0803

Table 3.5: Calibration on survival probabilities for Russia. NF stands for the non-fractional model and F for the fractional model.

	θ	σ	η	ρ	κ	λ_0	γ	z	w	α
NF	0.0223	0.0999	9.3351	5.2387	7.6154	0.0223	0.4298	0.2149	0.2524	1.000
F	0.9574	0.1014	9.7745	3.6976	6.4232	0.9574	0.2226	0.1113	1.3687	0.253

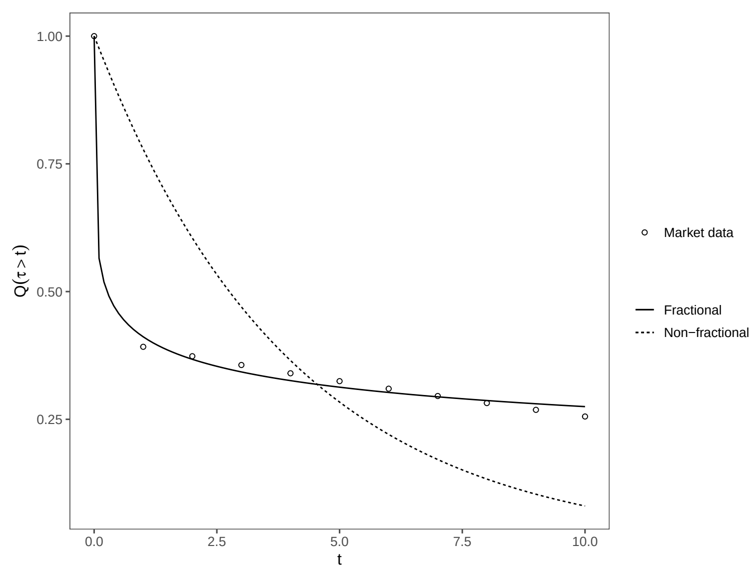

Figure 3.4: Survival probabilities for Russia and calibrated models.

Table 3.6: Microsoft: market survival probabilities and results of the calibration. M, F and NF respectively stand for the market data, the fractional model and the non-fractional model.

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$	$t=8$	$t=9$	$t=10$
M	0.9975	0.9931	0.9860	0.9757	0.9630	0.9501	0.9354	0.9229	0.9096	0.8956
F	0.9930	0.9851	0.9763	0.9668	0.9567	0.9463	0.9356	0.9247	0.9137	0.9027
NF	0.9930	0.9850	0.9763	0.9668	0.9567	0.9463	0.9356	0.9247	0.9137	0.9027

Table 3.7: Calibration on survival probabilities for Microsoft. NF stands for the non-fractional model and F for the fractional model.

	θ	σ	η	ρ	κ	λ_0	γ	z	w	α
NF	0.2156	0.1	10.0422	1.7981	5.9123	0.2156	0.0099	0.005	0.0126	1
F	0.2151	0.1	10.0425	1.7988	5.9106	0.2151	0.0099	0.005	0.0126	1

2022. The results are displayed at Tables 3.6 and 3.7, as well as on Figure 3.5. In this case, the survival probability curve on the market exhibits a concave shape, which the models can't reproduce. Note that from Table 3.5, we observe that the fractional model does not perform any better than the non-fractional one, the calibrated parameters being nearly the same for both models. This example shows that one should use an other model than the ones presented in this chapter if the survival probability curve of interest is concave.

Conclusions

In this chapter, we proposed to adapt the potential approach initiated by Rogers (1997) for interest rate modeling, to credit risk. This has been done with a fractional Hawkes jump diffusion supermartingale. In order to compute the survival probabilities implied by the model, we proposed a new recursive method to compute all the moments of a (fractional) Hawkes process intensity. We provided a proof that our numerical method converges and gave an upper bound on the numerical error. Finally, our numerical experiments showed that our model gives a good fit compared to similar models, but with a better numerical tractability.

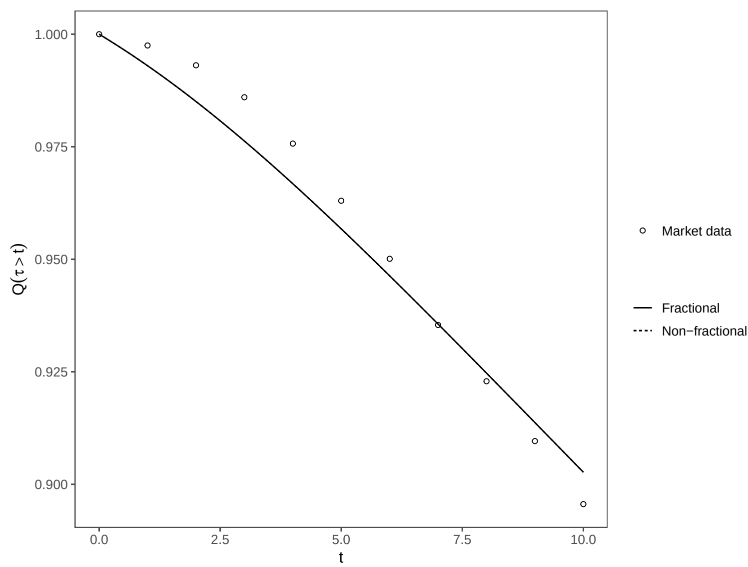

Figure 3.5: Survival probabilities for Microsoft and calibrated models.

Chapter 4

Option Pricing and Hedging in Illiquid Markets in Presence of Jump Clustering

Introduction

Illiquid assets, for example in emerging markets, are characterized by periods of time without any trade, and thus without movement in their prices. Properly taking into account this feature require stochastic processes that exhibit motionless periods. In this regard, usual models that involve Brownian motions or more general Lévy processes do not seem appropriate, as such processes move all the time. This problem also occurs in physical systems exhibiting subdiffusion. The periods without trades correspond to the trapping events in which the subdiffusive particle gets immobilized. Subdiffusion is thus a well known phenomenon in statistical physics. The usual mathematical tool for this phenomenon is fractional calculus. More precisely, the density of a subdiffusive model is characterized by a fractional partial differential equation (FPDE), as in Barkai et al. (2000), Metzler and Klafter (2004) and Metzler and Klafter (2000). That is, the partial derivative with respect to time is replaced by a fractional derivative. These particular equations are also called fractional Fokker-Planck equations. Due to the link between the subdiffusive behaviour and fractional calculus, we also refer to subdiffusive models as fractional models. Subdiffusive dynamics can also be obtained by the technique of time-change. This is what we have done in Chapters 2 and 3. A subdiffusion is obtained when a standard diffusion process is time-changed by the inverse of an α-stable subordinator. The inverse α-stable subordinator is characterized by the presence of motionless periods in its paths. From an asset modeling point of view, it allows to introduce illiquidity in the price of the asset, since the motionless periods correspond to the absence of trades. This approach is used in Magdziarz (2009a) to obtain a subdiffusive version of the famous Black and Scholes model and to show that vanilla option prices can be computed by numerical integration. It is also established that despite the Black and Scholes market being complete, its subdiffusive counterpart is not, since more than one risk-neutral measure exist. Hainaut and Leonenko (2021) propose an extension to a subdiffusive jump diffusion model, for which option prices can be computed by solving a fractional partial integro-differential equation. This equation is similar to Dupire's forward partial differential equation (PDE) for call prices.

The approach we propose in this chapter is not the only possibility to deal with illiquidity in asset price modeling. One may think of a high-frequency trading approach. Such an approach would consist in explaining the illiquidity of the asset by either the absence of buy or sell orders in the market, or by the existence of a large spread between the ask price and the bid price for existing orders. Although they do not particularly concentrate on illiquid assets, the papers of Cartea and Jaimungal (2013), Cont and de Larrard (2013), Cartea et al. (2014) and Lehalle et al. (2021) use this approach by modeling the arrival of buy or sell orders. In particular, Cartea et al. (2014) use self-exciting processes to model the clustering of arrival of buy and sell orders. A possible way to introduce illiquidity in a high-frequency approach could be to time-change the arrival of orders by the inverse of an α-stable subordinator.

A feature that often characterizes financial data is the clustering of certain events. Events that are typically clustered are jumps in the asset prices and defaults of companies, see Aït-Sahalia et al. (2014) and Ait-Sahalia et al. (2015). As an example, we can cite the 2008 economic crisis. As mentioned in the general introduction to this thesis, the collapse of Lehman Brothers brought the financial system to the brink of a breakdown. The important economical consequences indicates the existence of a propagation phenomena that is channeled through the complex economic system. Self-exciting point processes, also called Hawkes processes, allow to mathematically take into account this phenomenon of contagion. This approach finds its origins in the papers Hawkes (1971a) and Hawkes (1971b), as well as in Hawkes and Oakes (1974). The arrival of events is ruled by an intensity process which represents the instantaneous probability of occurrence. In the most common and simplest setting, the intensity process suddenly increases when an event occurs. Motivated by the 2008 crisis, the Hawkes process approach was used in Errais et al. (2010) to model the spillover effect between defaults. Self-exciting processes were also used in Ait-Sahalia et al. (2015) to model the clustering of jumps in the prices of financial assets or in Hainaut (2016b) and Hainaut (2016a) for modeling the term structure of interest rates. The pricing and hedging of vanilla options when the asset is driven by a self-exciting jump diffusion was studied in Moraux and Hainaut (2018). More precisely, this article shows that the Fourier transform of the log-asset price is described by a system of ordinary differential equations. After having solved this system numerically, the method of Carr and Madan (1999) is used to invert the transform with a fast Fourier transform (FFT) algorithm, leading to the option prices.

A self-exciting model that exhibits subdiffusive behaviour was introduced in Hainaut (2020) under the name of fractional Hawkes process. In Chapter 2, we have seen that such a process can be calibrated to credit default swaps by solving numerically a FPDE that characterizes the associated Laplace transform. The fractional Hawkes process is obtained by time-changing a Hawkes process, which is similar to the work of Magdziarz (2009a) with the Black and Scholes model. More precisely, the time-change used for subdiffusions is a process $(S_t)_{t\geqslant 0}$ defined as $S_t := \inf\{\tau > 0 : U_\tau > t\}$, where $(U_t)_{t\geqslant 0}$ is an α-stable subordinator, i.e. an increasing Lévy process to be described with more details hereinafter. In other words, $(S_t)_{t\geqslant 0}$ is the inverse of the α-stable subordinator $(U_t)_{t\geqslant 0}$. $(S_t)_{t\geqslant 0}$ is of course nondecreasing, since S_t is defined as the smallest time $(U_t)_{t\geqslant 0}$ goes above the level t. Note that this time-change is identical to the one we used in Chapters 2 and 3.

We refer to $(S_t)_{t\geqslant 0}$ as a time-change and not as a subordinator because it is not a Markov process, and thus not a Lévy process either. Related to that, the practical use in finance of such a time-change suffers from a twofold problem: nothing can be said about the distribution of S_t conditional on S_s, $s < t$. Even worse is the fact that even if the distribution of S_t given S_s could be known, the fact that $(S_t)_{t\geqslant 0}$ is not a Markov process would make this distribution irrelevant. This is of course a significant problem if, as often, one wants to re-evaluate a financial product after it is issued. A solution to the first part of this problem was proposed in Hainaut (2021). More specifically, Hainaut (2021) shows how to determine the distribution of S_t conditional on (S_s, U_{S_s}), $s < t$. That is, we have to adjunct the additional information on the last known value U_{S_s} of $(U_{S_t})_{t\geqslant 0}$, which represents the actual value reached by $(U_t)_{t\geqslant 0}$ on the first time it goes above the level s. This solution provides an insight for the second part of the problem, which becomes: *Is it relevant to use the distribution of S_t given the last known values (S_s, U_{S_s})?* As a matter of fact, it could be that the distribution S_t seen from time $s \in (0, t)$ depends on some values (S_u, U_{S_u}) with $u < s$, leading to a loss of information. The first main contribution of the present chapter is to show that the answer to the question in italics is *yes*. We show that even though $(S_t)_{t\geqslant 0}$ is not a Markov process, the bivariate process $(S_t, U_{S_t})_{t\geqslant 0}$ satisfies the Markov property, ensuring that the approach of Hainaut (2021) does not suffer from any loss of information.

The second main contribution of this chapter is a method for pricing vanilla options when the asset model is both self-exciting and subdiffusive. We obtain such a model by applying the technique of time-change mentioned above on the self-exciting jump diffusion asset model of Moraux and Hainaut (2018), leading to what we call a fractional (or subdiffusive) self-exciting jump diffusion model. We show that the Fourier transform of the log-price is ruled by a FPDE. Inspired by the finite difference method of Alikhanov (2015), we propose a numerical method to solve the aforementioned FPDE and show that its error is (at least) of order 2. Then, we compute vanilla option prices by numerical inversion of the transform with the help of a FFT algorithm.

The third contribution of this chapter is to tackle the question of the hedging of contingent claims in a fractional self-exciting jump diffusion setting. As the market is incomplete, a perfect replication of all the contingent claims is impossible. In this case, we aim at approximating the payoff of the contingent claim by a self-financing portfolio (or self-financing strategy) that minimizes the expected squared hedging error, as in e.g. Föllmer and Sondermann (1985) and Cont et al. (2012). Unlike approaches based on other loss functions, quadratic hedging yields linear hedging rules that are very convenient to implement, see Schweizer (2001). Relying on the predictable representation theorem for square integrable martingales of Kunita and Watanabe (1967), we derive the existence and uniqueness of an optimal quadratic hedging strategy for any square integrable contingent claim. Moreover, if the payoff of the contingent claim is not path dependent, i.e. if it depends on the final value of the asset only, we give an explicit formula for this optimal quadratic hedging strategy. In particular, we deduce the hedging strategies for call options.

The chapter is organized as follows. In the first section, we describe precisely the model we propose. In the second section, we deal with the Markov property of $(S_t, U_{S_t})_{t\geqslant 0}$. To do so, we derive some properties of the processes that are involved and study the different filtrations that arise in our setting. In the third section, we derive a fractional partial differential equation for the transform of the log-asset price. In the fourth section, we describe a whole class of risk-neutral measures. As a consequence, the market is arbitrage free and incomplete. In the fifth section, we propose a numerical method based on finite differences that allows to solve numerically the fractional partial differential equation derived in the third section. Furthermore, we show how the results of this method can be used to retrieve the price of a call option under a particular risk-neutral measure. In Section 6, we deal with the optimal quadratic hedging of contingent claims. Finally, the last section presents our numerical results that consist of call option prices and implied volatilities computed with the method we proposed.

4.1 The Model

This first section is dedicated to the precise introduction of the model we use. For a technical reason, the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is assumed to be the product of two complete probability spaces $(\Omega^{(1)}, \mathcal{F}^{(1)}, \mathbb{P}^{(1)})$ and $(\Omega^{(2)}, \mathcal{F}^{(2)}, \mathbb{P}^{(2)})$. The first probability space $(\Omega^{(1)}, \mathcal{F}^{(1)}, \mathbb{P}^{(1)})$ carries the non-fractional model which, as mentioned in the introduction, is the same as in Moraux and Hainaut (2018). The other probability space carries all the processes related to the time-change operation (the α-stable Lévy process and its inverse). The introduction of the model with two probability spaces is the reason why the different stochastic processes will be indexed by a superscript (i) indicating on which probability space they live $(i = 1, 2)$. This gives a little cumbersome notations at the beginning of this chapter, but the product space construction is crucial to simplify the proof of one of our results (this result is Proposition 4.17).

Let $(A_t^{(1)})_{t\geqslant 0}$ be a stochastic process on $(\Omega^{(1)}, \mathcal{F}^{(1)}, \mathbb{P}^{(1)})$ that satisfy the following stochastic differential equation (SDE)

$$\frac{\mathrm{d}A_t^{(1)}}{A_{t-}^{(1)}} = \mu \mathrm{d}t + \sigma \mathrm{d}W_t^{(1)} + \mathrm{d}D_t^{(1)} - \lambda_{t-}^{(1)}\mathbb{E}[e^{\xi} - 1]\mathrm{d}t, \tag{4.1}$$

where $(W_t^{(1)})_{t\geqslant 0}$ is a standard Brownian motion on $(\Omega^{(1)}, \mathcal{F}^{(1)}, \mathbb{P}^{(1)})$. The processes $(J_t^{(1)})_{t\geqslant 0}$ and $(\lambda_t^{(1)})_{t\geqslant 0}$ are defined as we explain now. Let $(N_t^{(1)})_{t\geqslant 0}$ be a pure jump process with jumps of size 1 and intensity $(\lambda_t^{(1)})_{t\geqslant 0}$ and $\xi_1, \xi_2, \cdots =_{\mathrm{dist}} \xi$ be independent and identically distributed random variables whose distribution is given by the probability measure ν on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. We assume that the jumps $\xi_1, \xi_2, \ldots$ are double-exponential random variables. That is, the probability measure ν is of the form

$$\nu(B) := p\int_{B\cap\mathbb{R}^+} \rho^+ e^{-\rho_+ z}\mathrm{d}z - (1-p)\int_{B\cap\mathbb{R}^-} \rho^- e^{-\rho_- z}\mathrm{d}z \tag{4.2}$$

where p and $(1-p)$ are respectively the probabilities to observe upward and downward jumps and $\rho_+ > 0$, $\rho_- < 0$ are the jump size parameters. The average upward jump size is $1/\rho_+ > 0$ whereas the average downward jump size is $1/\rho_- < 0$. Overall, the expected size of a jump is $\mathbb{E}[\xi] = p(1/\rho_+) + (1-p)(1/\rho_-)$. The joint moment-generating function of $(\xi, |\xi|)$ will be needed in the subsequent developments and is denoted by ψ. From Equation (4.2), we have

$$\begin{aligned}\psi(z_1, z_2) &:= \mathbb{E}\left[e^{z_1\xi + z_2|\xi|}\right] \\ &= p\frac{\rho_+}{\rho_+ - (z_1 + z_2)} + (1-p)\frac{\rho_-}{\rho_- - (z_1 - z_2)}\end{aligned} \tag{4.3}$$

provided that $(z_1 + z_2) < \rho_+$ and $(z_1 - z_2) > \rho_-$. We set

$$D_t^{(1)} := \sum_{k=1}^{N_t^{(1)}} (e^{\xi_k} - 1) \qquad J_t^{(1)} := \sum_{k=1}^{N_t^{(1)}} \xi_k \qquad L_t^{(1)} := \sum_{k=1}^{N_t^{(1)}} |\xi_k|, \tag{4.4}$$

and define the stochastic intensity process $(\lambda_t^{(1)})_{t\geqslant 0}$ as

$$\mathrm{d}\lambda_t^{(1)} = \kappa(\theta - \lambda_{t-}^{(1)})\mathrm{d}t + \eta \mathrm{d}L_t^{(1)}. \tag{4.5}$$

Finally, we define the log-price of the stock as the process $(X_t^{(1)})_{t\geqslant 0} := (\ln A_t^{(1)})_{t\geqslant 0}$. Using Ito's Lemma, we have

$$\mathrm{d}X_t^{(1)} = \left(\mu - \frac{\sigma^2}{2} - \lambda_{t-}^{(1)}\mathbb{E}[e^\xi - 1]\right)\mathrm{d}t + \sigma\mathrm{d}W_t^{(1)} + \mathrm{d}J_t^{(1)}. \tag{4.6}$$

This relation implies that

$$A_t^{(1)} = A_0^{(1)} \exp\left\{\left(\mu - \frac{\sigma^2}{2}\right)t - \mathbb{E}[e^\xi - 1]\int_0^t \lambda_{s-}^{(1)}\mathrm{d}s + \sigma W_t^{(1)} + J_t^{(1)}\right\}. \tag{4.7}$$

This completes the introduction of all the stochastic processes defined on $(\Omega^{(1)}, \mathcal{F}^{(1)}, \mathbb{P}^{(1)})$. Let $(U_t^{(2)})_{t\geqslant 0}$ be an α-stable subordinator on the probability space $(\Omega^{(2)}, \mathcal{F}^{(2)}, \mathbb{P}^{(2)})$. The process $(U_t^{(2)})_{t\geqslant 0}$ is thus a strictly increasing Lévy process whose Laplace transform satisfies

$$\mathbb{E}[e^{-\omega U_t^{(2)}}] = e^{-t\omega^\alpha}$$

for some $\alpha \in (0,1)$. We denote by $(S_t^{(2)})_{t\geqslant 0}$ the inverse of $(U_t^{(2)})_{t\geqslant 0}$, that is

$$S_t^{(2)} := \inf\{\tau > 0 : U_\tau^{(2)} > t\}$$

for each $t \geqslant 0$. The paths of $(S_t^{(2)})_{t\geqslant 0}$ are nondecreasing $\mathbb{P}^{(2)}$-a.s., but not strictly increasing. Note that $(S_t^{(2)})_{t\geqslant 0}$ has continuous paths $\mathbb{P}^{(2)}$-a.s., which is a consequence of the strictly increasing paths of $(U_t^{(2)})_{t\geqslant 0}$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be defined as follows: $\Omega = \Omega^{(1)} \times \Omega^{(2)}$, $\mathcal{F} = \mathcal{F}^{(1)} \otimes \mathcal{F}^{(2)}$ and $\mathbb{P}(B_1 \times B_2) = \mathbb{P}^{(1)}(B_1)\mathbb{P}^{(2)}(B_2)$ for any $B_1 \times B_2 \in \mathcal{F}^{(1)} \times \mathcal{F}^{(2)}$. The process $(A_t)_{t\geqslant 0}$ is defined on $(\Omega, \mathcal{F}, \mathbb{P})$ as $A_t(\omega_1, \omega_2) = A_t^{(1)}(\omega_1)$ for all $(\omega_1, \omega_2) \in \Omega$. The process $(U_t)_{t\geqslant 0}$ is defined as $U_t(\omega_1, \omega_2) = U_t^{(2)}(\omega_2)$. The other processes $(\lambda_t)_{t\geqslant 0}$, $(J_t)_{t\geqslant 0}$, $(S_t)_{t\geqslant 0}$ and the others are defined similarly. This construction obviously implies that the processes $(A_t)_{t\geqslant 0}$ and $(U_t)_{t\geqslant 0}$ are independent. The null sets of $(\Omega, \mathcal{F}, \mathbb{P})$ are denoted by $\mathcal{N}$. By the assumption of completeness of the probability space, we obviously have $\mathcal{N} \subset \mathcal{F}$.

The different jump processes will often be expressed through their random measure. Let $\Xi(\mathrm{d}z, \mathrm{d}t)$ be the random measure associated with $(J_t)_{t\geqslant 0}$. That is, for $A, B \in \mathcal{B}_\mathbb{R}$,

$$\Xi(A, B) = \sum_{n\geqslant 0} \delta_{(\xi_n, \tau_n)}(A, B)$$

where τ_n and ξ_n are respectively the time and size of the n^{th} jump of $(J_t)_{t\geqslant 0}$, and $\delta_{(x,y)}$ is the Dirac measure located at $(x, y) \in \mathbb{R}^2$. Then we have

$$J_t = \int_{s=0}^t \int_{z\in\mathbb{R}} z\ \Xi(\mathrm{d}z, \mathrm{d}s),$$

$$L_t = \int_{s=0}^{t} \int_{z\in\mathbb{R}} |z| \; \Xi(\mathrm{d}z, \mathrm{d}s)$$

and

$$D_t = \int_{s=0}^{t} \int_{z\in\mathbb{R}} (e^z - 1) \; \Xi(\mathrm{d}z, \mathrm{d}s).$$

The random measure Ξ can also be used to express the jump intensity $(\lambda_t)_{t\geqslant 0}$ as

$$\mathrm{d}\lambda_t = \kappa(\theta - \lambda_{t-})\mathrm{d}t + \eta \int_{z\in\mathbb{R}} |z| \; \Xi(\mathrm{d}z, \mathrm{d}t).$$

The compensator of the random measure Ξ is denoted by υ and satisfies

$$\upsilon(A, B) = \int_{s\in A\cap\mathbb{R}^+} \int_{z\in B} \lambda_{s-} \; \nu(\mathrm{d}z)\mathrm{d}s$$

for $A, B \in \mathcal{B}_{\mathbb{R}}$. Finally, the compensated random measure $\tilde{\Xi}$ is defined as $\tilde{\Xi}(\mathrm{d}z, \mathrm{d}t) = \Xi(\mathrm{d}z, \mathrm{d}t) - \upsilon(\mathrm{d}z, \mathrm{d}t)$.

The stock price is represented by the time-changed process $(A_{S_t})_{t\geqslant 0}$. Figure 4.1 displays examples of paths for the asset price $(A_{S_t})_{t\geqslant 0}$ (or $(A_t)_{t\geqslant 0}$ in the bottom right non fractional case) for different fractional orders α. The cases $\alpha < 1$ exhibit motionless periods, corresponding to the absence of trade for the asset. This is the feature that allows to take into account the potential illiquidity of the asset. Observe that when α is smaller, the periods of illiquidity tend to last longer.

4.2 Markov property of the time-change process

We denote by $(\mathcal{F}_t^U)_{t\geqslant 0}$ and $(\mathcal{F}_t^S)_{t\geqslant 0}$ the natural filtrations of $(U_t)_{t\geqslant 0}$ and $(S_t)_{t\geqslant 0}$. Note that all the natural filtrations are assumed to contain the null sets $\mathcal{N}$. The α-stable subordinator $(U_t)_{t\geqslant 0}$ admits a PDF that we denote by $p_U(t,\tau) = \frac{\partial}{\partial\tau}\mathbb{P}(U_t \leqslant \tau)$. This PDF satisfies the representation

$$p_U(t,\tau) = \frac{1}{\pi} \sum_{k=0}^{+\infty} \frac{(-1)^k}{k!} \Gamma(1 + k\alpha)\tau^{-(1+\alpha k)} t^k \sin(\pi\alpha k)$$

for $\tau > 0$. See e.g. Proposition 3.1 in Gupta and Kumar (2022) for a proof. As a consequence, the subordinator $(U_t)_{t\geqslant 0}$ has strictly increasing paths $\mathbb{P}$-a.s.. Note that being a Lévy process, $(U_t)_{t\geqslant 0}$ is also a Feller process, for which the reader is referred to the Chapter 3 of Revuz and Yor (2005). If we consider $(U_t)_{t\geqslant 0}$ as a Feller process, any initial distribution can be considered for U_0. For a random variable X and a $\bigvee_{t\geqslant 0} \mathcal{F}_t^U$-measurable random variable Z, we write $\mathbb{E}^X[Z]$ for the expected value obtained for Z if the initial distribution of U_0 is the distribution of

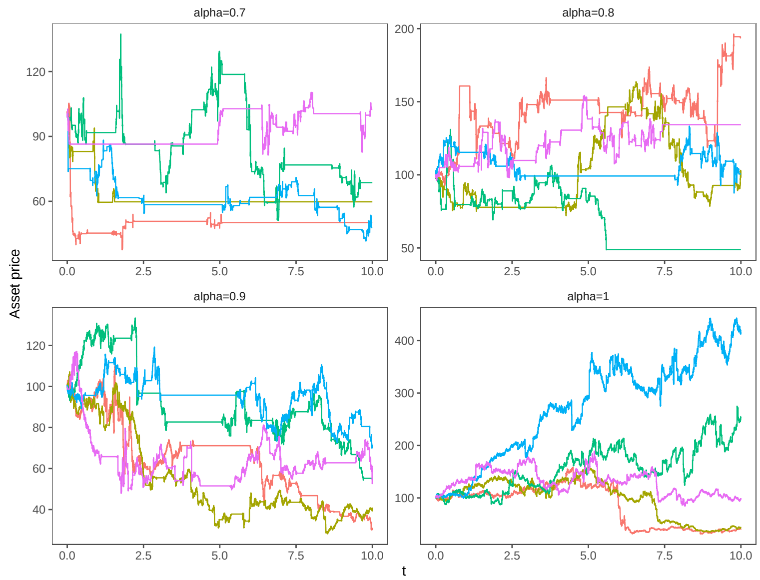

Figure 4.1: Five simulated paths for fractional order varying from 0.7 to 1. A fractional order equal to 1 corresponds to the non fractional model (no time change).

X. When no superscript is added on the expectation operator, we consider $U_0 = 0$ $\mathbb{P}$-a.s., i.e. $\mathbb{E}[Z] = \mathbb{E}^0[Z]$ and $(U_t)_{t\geqslant 0}$ is then a Lévy process.

The next proposition gives some further basic properties of $(U_t)_{t\geqslant 0}$ and its inverse $(S_t)_{t\geqslant 0}$.

Proposition 4.1. *(i) For any $t \geqslant 0$, $S_t \in \{\tau \geqslant 0 : U_\tau \geqslant t\}$ $\mathbb{P}$-a.s., so that $U_{S_t} \geqslant t$ $\mathbb{P}$-a.s..*

(ii) For any $t \geqslant 0$, S_t is a $\mathbb{P}$-a.s. finite stopping-time with respect to the filtration $(\mathcal{F}_t^U)_{t\geqslant 0}$. Moreover, $\sup_{u\geqslant 0} S_u < +\infty$ $\mathbb{P}$-a.s..

(iii) $(S_t)_{t\geqslant 0}$ has continuous paths $\mathbb{P}$-a.s..

Proof. (i) By definition of S_t and the definition of infimum, there is a random sequence $\tau_n \in \{\tau \geqslant 0 : U_\tau > t\}$ such that $\tau_n \downarrow S_t$. By the right-continuity of $(U_\tau)_{\tau\geqslant 0}$, we have $\lim_{n\to+\infty} U_{\tau_n} = U_{S_t}$. Since $U_{\tau_n} > t$ for all n, we also have $\lim_{n\to+\infty} U_{\tau_n} = U_{S_t} \geqslant t$.

(ii) That S_t is a stopping-time is a consequence of $\{S_t \leqslant s\} = \{U_s \geqslant t\} \in \mathcal{F}_s^U$ for all $s \geqslant 0$. Let us prove that it is finite a.s.. For any fixed $t > 0$,

$$\mathbb{P}(S_t < +\infty) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{S_t \leqslant n\}\right) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{U_n \geqslant t\}\right),$$

where we used that $\{S_t \leqslant n\} = \{U_n \geqslant t\}$. Moreover, since the process $(U_t)_{t\geqslant 0}$ is increasing, the sequence of events $(\{U_n \geqslant t\})_{n\in\mathbb{N}}$ is increasing. As a consequence, the continuity of probability measures implies that

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{U_n \geqslant t\}\right) = \lim_{n\to+\infty} \mathbb{P}(U_n \geqslant t).$$

Since $(U_s)_{s\geqslant 0}$ is an α-stable process, the random variables U_n and $n^{1/\alpha}U_1$ share the same probability distribution, thus

$$\lim_{n\to+\infty} \mathbb{P}(U_n \geqslant t) = \lim_{n\to+\infty} \mathbb{P}\left(\left\{U_1 \geqslant \frac{t}{n^{1/\alpha}}\right\}\right).$$

Again, the sequence of events $\left(\left\{U_1 \geqslant \frac{t}{n^{1/\alpha}}\right\}\right)_{n\in\mathbb{N}}$ is increasing so that the continuity of probability measures yields

$$\lim_{n\to+\infty} \mathbb{P}\left(\left\{U_1 \geqslant \frac{t}{n^{1/\alpha}}\right\}\right) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \left\{U_1 \geqslant \frac{t}{n^{1/\alpha}}\right\}\right) = \mathbb{P}(U_1 > 0) = 1.$$

This proves that S_t is a.s. finite. As a consequence of the nondecreasing paths of $(S_t)_{t\geqslant 0}$, the sequence of events $(\{S_n < +\infty\})_{n\in\mathbb{N}}$ is decreasing. Therefore, the continuity of probability measures implies

$$\mathbb{P}\left(\sup_{u\geqslant 0} S_u < +\infty\right) = \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} \{S_n < +\infty\}\right) = \lim_{n\to\infty} \mathbb{P}(S_n < +\infty) = 1,$$

which concludes the proof of (ii).

(iii) A proof of the right-continuity of $(S_t)_{t\geqslant 0}$ is given at Lemma 4.8 in Chapter 0 of Revuz and Yor (2005). We show now that the left-continuity can be obtained as a consequence of the strictly increasing paths of $(U_t)_{t\geqslant 0}$. Let

$$A := \{\omega \in \Omega : \text{the function } t \mapsto U_t(\omega) \text{ is strictly increasing}\}$$

and

$$B_l := \{\omega \in \Omega : \text{the function } t \mapsto S_t(\omega) \text{ is left-continuous}\}.$$

Let $\omega \in A$ and assume by contradiction that $\omega \notin B_l$. Then there would be some $t \geqslant 0$ and $\eta > 0$ such that for all $\varepsilon > 0$,

$$S_t(\omega) - \eta > S_{t-\varepsilon}(\omega).$$

As a consequence, since $\omega \in A$,

$$t > U_{S_t(\omega)-\eta}(\omega) > U_{S_{t-\varepsilon}(\omega)}(\omega) \geqslant t - \varepsilon,$$

the last inequality being a consequence of Proposition 4.1 (i). By letting ε decrease to 0, we obtain

$$t > U_{S_t(\omega)-\eta}(\omega) \geqslant t,$$

which is a contradiction. We deduce that $A \subset B_l$. Since $\mathbb{P}(A) = 1$, the conclusion follows. □

Revuz and Yor (2005) also note that $S_{t-} = \inf\{s \geqslant 0 : U_s \geqslant t\}$, but since $(S_t)_{t\geqslant 0}$ is left-continuous by (iii) in Proposition 4.1, we have

$$S_t = \inf\{s \geqslant 0 : U_s \geqslant t\} = \inf\{s \geqslant 0 : U_s > t\}.$$

As $(S_t)_{t\geqslant 0}$ is a nondecreasing process, Proposition 4.1 (ii) implies that it is has bounded variation. Since it has continuous (and thus càdlàg) paths, Proposition 4.1 (ii) also implies that $(S_t)_{t\geqslant 0}$ is a semimartingale. Moreover, having proved that S_t is a stopping-time for any $t \geqslant 0$, we can introduce the stopped sigma-algebras

$$\mathcal{F}^U_{S_t} := \left\{A \in \mathcal{F}^U_\infty : A \cap \{S_t \leqslant u\} \in \mathcal{F}^U_u \text{ for all } u \geqslant 0\right\},$$

where $\mathcal{F}^U_\infty = \bigvee_{t\geqslant 0} \mathcal{F}^U_t$. Since for any $0 \leqslant s \leqslant t$, S_s and S_t are stopping-times that satisfy $S_s \leqslant S_t$ a.s., it is clear that the collection $(\mathcal{F}^U_{S_t})_{t\geqslant 0}$ of sigma-algebras is a filtration. We also introduce the filtrations $(\mathcal{F}^{U\circ S}_t)_{t\geqslant 0}$ and $(\mathcal{F}^{S,U\circ S}_t)_{t\geqslant 0}$ to be the natural filtrations of respectively $(U_{S_t})_{t\geqslant 0}$ and $(S_t, U_{S_t})_{t\geqslant 0}$. The next proposition shows that $(\mathcal{F}^U_{S_t})_{t\geqslant 0}$ is finer than the natural filtration of $(S_t)_{t\geqslant 0}$.

Proposition 4.2. *For all $t \geqslant 0$, $\mathcal{F}^S_t \subset \mathcal{F}^U_{S_t}$.*

Proof. Fix any $r \in \mathbb{R}$ and $\tau \leqslant t$. We start by showing that $\{S_\tau \leqslant r\} \cap \{S_t \leqslant u\} \in \mathcal{F}^U_u$ for all $u \geqslant 0$. To this end, note that $\{S_\tau \leqslant r\} \cap \{S_t \leqslant u\} = \{U_r \geqslant \tau\} \cap \{U_u \geqslant t\}$. If $r > u$, then

$$\{U_u \geqslant t\} \subset \{U_r \geqslant t\} \subset \{U_r \geqslant \tau\}$$

so that $\{U_r \geqslant \tau\} \cap \{U_u \geqslant t\} = \{U_u \geqslant t\} \in \mathcal{F}_u^U$. If $r \leqslant u$, it is clear that $\{U_r \geqslant \tau\} \cap \{U_u \geqslant t\} \in \mathcal{F}_u^U$. This allows us to conclude that $\{S_\tau \leqslant r\} \in \mathcal{F}_{S_t}^U$. Since the intervals of the form $(-\infty, r]$ generate the Borel σ-algebra $\mathcal{B}_\mathbb{R}$, it implies that for any $B \in \mathcal{B}_\mathbb{R}$ and $0 \leqslant \tau \leqslant t$, we have

$$\{S_\tau \in B\} \in \mathcal{F}_{S_t}^U.$$

The conclusion follows. □

Similarly, we can define the σ-algebra of events that happen before the stopping-time S_t as

$$\mathcal{F}_{S_t-}^U := \mathcal{F}_0^U \vee \sigma\{A \cap \{s < S_t\} : A \in \mathcal{F}_s^U, s \in \mathbb{R}^+\}.$$

It is well known that $\mathcal{F}_{S_t-}^U \subset \mathcal{F}_{S_t}^U$, for all $t \geqslant 0$. Finally, we will also use the stopped processes $(U_s^{S_t})_{s\geqslant 0}$ that are defined as

$$U_s^{S_t} := \begin{Bmatrix} U_{S_t} & \text{if } s \geqslant S_t \\ U_s & \text{if } s < S_t \end{Bmatrix} = U_s \mathbf{1}_{\{s<S_t\}} + U_{S_t}\mathbf{1}_{\{s\geqslant S_t\}} = U_{s\wedge S_t}.$$

We give now some intermediary results that will be useful to prove that the bivariate process $(S_t, U_{S_t})_{t\geqslant 0}$ is a Markov process with respect to its natural filtration $(\mathcal{F}_t^{S,U\circ S})_{t\geqslant 0}$. Before stating and proving these results, let us stress their purpose. The proof of the Markov property of $(S_t, U_{S_t})_{t\geqslant 0}$ relies on the strong Markov property of the Feller process $(U_t)_{t\geqslant 0}$, which states that such processes renew themselves at stopping-times. When stopping $(U_t)_{t\geqslant 0}$ at the stopping-time S_t, the relevant σ-algebra is the stopped σ-algebra $\mathcal{F}_{S_t}^U$. However, as mentioned before, we want to prove that $(S_t, U_{S_t})_{t\geqslant 0}$ satisfies the Markov property with respect to its natural filtration $(\mathcal{F}_t^{S,U\circ S})_{t\geqslant 0}$. The first goal of our preliminary results is thus to establish that both filtration $(\mathcal{F}_t^{S,U\circ S})_{t\geqslant 0}$ and $(\mathcal{F}_{S_t}^U)_{t\geqslant 0}$ coincide. The other goal of the preliminary results is to express some events of interest as countable unions or intersections of more simple events to prove that they are measurable.

Lemma 4.1. *Let $\mathcal{N}$ denote the collection of all null sets in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then for each $t \geqslant 0$ it holds that*

$$\mathcal{F}_{S_t}^U = \sigma(S_t, U^{S_t}) \vee \mathcal{N},$$

where $\sigma(S_t, U^{S_t})$ denote the sigma-algebra generated by the random variable S_t and the stopped process $(U_s^{S_t})_{s\geqslant 0}$.

Proof. **First part.** The first step is to prove $\mathcal{F}_{S_t}^U \supset \sigma(S_t, U^{S_t}) \vee \mathcal{N}$. By Proposition 4.2, it is clear that $\mathcal{F}_{S_t}^U \supset \sigma(S_t)$. Since we assume that the probability space is complete and that the filtrations contain the null sets, it also holds that $\mathcal{F}_{S_t}^U \supset \mathcal{N}$.

Therefore, the first step will be completed by proving that $\mathcal{F}^U_{S_t} \supset \sigma(U^{S_t})$. Fix $s \geqslant 0$ and $B \in \mathcal{B}_\mathbb{R}$. For any $u \geqslant 0$,

$$
\begin{aligned}
&\{U^{S_t}_s \in B\} \cap \{S_t \leqslant u\} \\
&= \Big(\{U^{S_t}_s \in B\} \cap \{S_t \leqslant u\} \cap \{S_t \leqslant s\}\Big) \cup \Big(\{U^{S_t}_s \in B\} \cap \{S_t \leqslant u\} \cap \{S_t > s\}\Big) \\
&= \Big(\{U_{S_t} \in B\} \cap \{S_t \leqslant u \wedge s\}\Big) \cup \Big(\{U_s \in B\} \cap \{S_t \leqslant u\} \cap \{S_t > s\}\Big)
\end{aligned}
$$

Theorem 6 Chapter 1 in Protter (2005) implies that $\{U_{S_t} \in B\} \in \mathcal{F}^U_{S_t}$ and consequently that $\{U^{S_t}_s \in B\} \cap \{S_t \leqslant u \wedge s\} \in \mathcal{F}^U_{u\wedge s} \subset \mathcal{F}^U_u$. Moreover,

$$
\begin{aligned}
&\{U_s \in B\} \cap \{S_t \leqslant u\} \cap \{S_t > s\} \\
&\qquad = \left.\begin{cases} \emptyset & \text{if } s \geqslant u \\ \{U_s \in B\} \cap \overline{\{U_s \geqslant t\}} \cap \{U_u \geqslant t\} & \text{if } s < u \end{cases}\right\} \in \mathcal{F}^U_u .
\end{aligned}
$$

It follows that $\{U^{S_t}_s \in B\} \cap \{S_t \leqslant u\} \in \mathcal{F}^U_u$, which proves the first inclusion.

Second part. The second step is to prove $\mathcal{F}^U_{S_t} \subset \sigma(S_t, U^{S_t}) \vee \mathcal{N}$. Define

$$
\mathcal{A} := \left\{ A \in \mathcal{F}^U_\infty : \mathbb{E}[\mathbf{1}_A | \mathcal{F}^U_{S_t}] = \mathbb{E}[\mathbf{1}_A | \sigma(S_t, U^{S_t}) \vee \mathcal{N}] \right\}.
$$

Clearly $\mathcal{A}$ is a λ-system that contains $\mathcal{N}$. Let $\mathcal{C}$ be the collection of subsets

$$
\mathcal{C} := \left\{ \bigcap_{j=1}^n \{U_{t_j} \in B_j\} : n \in \mathbb{N}, t_j \in \mathbb{R}^+, B_j \in \mathcal{B}_\mathbb{R} \right\}.
$$

This collection is a π-system that satisfies $\sigma(\mathcal{C}) \vee \mathcal{N} = \mathcal{F}^U_\infty$. For $C \in \mathcal{C}$,

$$
\begin{aligned}
\mathbb{E}[\mathbf{1}_C | \mathcal{F}^U_{S_t}] &= \mathbb{E}\left[\prod_{j=1}^n \mathbf{1}_{\{U_{t_j} \in B_j\}} \middle| \mathcal{F}^U_{S_t} \right] \\
&= \sum_{k=1}^{n+1} \mathbb{E}\left[\mathbf{1}_{\{S_t \in [t_{k-1}, t_k)\}} \prod_{j=1}^n \mathbf{1}_{\{U_{t_j} \in B_j\}} \middle| \mathcal{F}^U_{S_t} \right],
\end{aligned}
$$

where t_0 and t_{n+1} have to be understood as 0 and $+\infty$ respectively. Note that

$$
\mathbf{1}_{\{S_t \in [t_{k-1}, t_k)\}} \mathbf{1}_{\{U_{t_j} \in B_j\}} = \begin{cases} \mathbf{1}_{\{S_t \in [t_{k-1}, t_k)\}} \mathbf{1}_{\{U_{t_j \wedge S_t} \in B_j\}} & \text{if } j < k \\ \mathbf{1}_{\{S_t \in [t_{k-1}, t_k)\}} \mathbf{1}_{\{U_{t_j \vee S_t} \in B_j\}} & \text{if } j \geqslant k. \end{cases}
$$

Moreover, by the $\mathcal{F}^U_{S_t}$-measurability of the random variables S_t (Proposition 4.2) and $U_{t_j \wedge S_t} = U^{S_t}_{t_j}$ (first part of this proof), we deduce that $\mathbf{1}_{\{S_t \in [t_{k-1}, t_k)\}} \mathbf{1}_{\{U_{t_j \wedge S_t} \in B_j\}}$

is $\mathcal{F}^U_{S_t}$-measurable whenever $j < k$. As a consequence,

$$\mathbb{E}[\mathbf{1}_C|\mathcal{F}^U_{S_t}] = \sum_{k=1}^{n+1} \mathbf{1}_{\{S_t\in[t_{k-1},t_k)\}} \prod_{j=1}^{k-1} \mathbf{1}_{\{U_{t_j\wedge S_t}\in B_j\}} \mathbb{E}\left[\prod_{j=k}^{n} \mathbf{1}_{\{U_{t_j\vee S_t}\in B_j\}} \middle| \mathcal{F}^U_{S_t}\right]$$

Let Z be the $\mathcal{F}^U_\infty$-measurable random variable $\prod_{j=k}^n \mathbf{1}_{\{U_{(t_j\vee S_t)-S_t}\in B_j\}}$. Then

$$\prod_{j=k}^{n} \mathbf{1}_{\{U_{t_j\vee S_t}\in B_j\}} = Z\circ\theta_{S_t},$$

where $(\theta_t)_{t\geqslant 0}$ denotes the translation operators defined as $\theta_t(U_s) = U_{s+t}$ (see Revuz and Yor (2005), Chapters 1 and 3). From Theorem 3.1 in Chapter 3 of Revuz and Yor (2005), we have

$$\mathbb{E}\left[Z\circ\theta_{S_t}|\mathcal{F}^U_{S_t}\right] = \mathbb{E}^{U_{S_t}}[Z],$$

and thus

$$\mathbb{E}[\mathbf{1}_C|\mathcal{F}^U_{S_t}] = \sum_{k=1}^{n+1} \mathbf{1}_{\{S_t\in[t_{k-1},t_k)\}} \prod_{j=1}^{k-1} \mathbf{1}_{\{U_{t_j\wedge S_t}\in B_j\}} \mathbb{E}^{U_{S_t}}[Z].$$

Since $\mathcal{F}^U_{S_t} \supset \sigma(S_t, U^{S_t}) \vee \mathcal{N}$, the tower property for conditional expectations gives

$$\begin{aligned}
&\mathbb{E}[\mathbf{1}_C|\sigma(S_t, U^{S_t})\vee\mathcal{N}] = \mathbb{E}[\mathbb{E}[\mathbf{1}_C|\mathcal{F}^U_{S_t}]|\sigma(S_t, U^{S_t})\vee\mathcal{N}] \\
&= \sum_{k=1}^{n+1} \mathbb{E}\left[\mathbf{1}_{\{S_t\in[t_{k-1},t_k)\}} \prod_{j=1}^{k-1} \mathbf{1}_{\{U_{t_j\wedge S_t}\in B_j\}} \mathbb{E}^{U_{S_t}}[Z] \middle| \sigma(S_t, U^{S_t})\vee\mathcal{N}\right] \\
&= \sum_{k=1}^{n+1} \mathbf{1}_{\{S_t\in[t_{k-1},t_k)\}} \prod_{j=1}^{k-1} \mathbf{1}_{\{U_{t_j\wedge S_t}\in B_j\}} \mathbb{E}^{U_{S_t}}[Z] \\
&= \mathbb{E}[\mathbf{1}_C|\mathcal{F}^U_{S_t}],
\end{aligned}$$

where the third equality comes from the $\sigma(S_t, U^{S_t})$-measurability of the random variable in the conditional expectation. From Dynkin's π-λ theorem, we conclude that $\sigma(\mathcal{C}) \subset \mathcal{A}$. Since $\sigma(\mathcal{C})\vee\mathcal{N} = \mathcal{F}^U_\infty$, it implies that $\mathcal{F}^U_\infty = \mathcal{A}$. As a conclusion, for any $A \in \mathcal{F}^U_{S_t}$, $A \in \mathcal{A}$ and therefore

$$\mathbf{1}_A = \mathbb{E}[\mathbf{1}_A|\mathcal{F}^U_{S_t}] = \mathbb{E}[\mathbf{1}_A|\sigma(S_t, U^{S_t})\vee\mathcal{N}],$$

so that $A \in \sigma(S_t, U^{S_t})\vee\mathcal{N}$, concluding the proof. □

In the following, we write $(\mathcal{F}^{U\circ S}_t)_{t\geqslant 0}$ for the natural filtration of the process $(U_{S_t})_{t\geqslant 0}$ and $(\mathcal{F}^{U\circ S,S}_t)_{t\geqslant 0}$ for the natural filtration of the bivariate process $(U_{S_t}, S_t)_{t\geqslant 0}$.

Lemma 4.2. *For all* $t \geqslant 0$,

(*i*) $\mathcal{F}^U_{S_t} = \sigma(\mathcal{F}^U_{S_t-}, U_{S_t})$,

(*ii*) $\mathcal{F}^U_{S_t-} \subset \mathcal{F}^S_t$.

Proof. (*i*) Recall that the $\mathcal{F}^U_{S_t}$-measurability of U_{S_t} follows from Theorem 6 (P.5) in Protter (2005). Therefore we just have to prove that $\mathcal{F}^U_{S_t} \subset \sigma(\mathcal{F}^U_{S_t-}, U_{S_t})$. By Lemma 4.1, $\mathcal{F}^U_{S_t} = \sigma(S_t, U^{S_t}) \vee \mathcal{N}$, so that we will prove that $\sigma(\mathcal{F}^U_{S_t-}, U_{S_t})$ contains both $\sigma(S_t)$ and $\sigma(U^{S_t})$. From the definition of $\mathcal{F}^U_{S_t-}$, clearly $\{S_t > s\} \in \mathcal{F}^U_{S_t-}$ for all $s \geqslant 0$, so that $\sigma(S_t) \subset \sigma(\mathcal{F}^U_{S_t-}, U_{S_t})$. For the second inclusion, note that for any $s \geqslant 0$ and $B \in \mathcal{B}_{\mathbb{R}}$, one has

$$\begin{aligned}\{U^{S_t}_s \in B\} &= \Big(\{U^{S_t}_s \in B\} \cap \{S_t \geqslant s\}\Big) \cup \Big(\{U^{S_t}_s \in B\} \cap \{S_t > s\}\Big) \\ &= \Big(\{U_{S_t} \in B\} \cap \{S_t \geqslant s\}\Big) \cup \Big(\{U_s \in B\} \cap \{S_t > s\}\Big).\end{aligned}$$

That the sets $\{S_t \geqslant s\}$ and $\{U_s \in B\} \cap \{S_t > s\}$ belong to $\mathcal{F}^U_{S_t-}$ is a direct consequence of the definition of $\mathcal{F}^U_{S_t-}$. We thus conclude that the second inclusion $\sigma(U^{S_t}) \subset \sigma(\mathcal{F}^U_{S_t-}, U_{S_t})$ holds.

(*ii*) We have to show that $A \cap \{S_t > s\} \in \mathcal{F}^S_t$ whenever $s \geqslant 0$ and $A \in \mathcal{F}^U_s$. Fix some $s \geqslant 0$. We start by showing that $\{U_u < b\} \cap \{S_t > s\} \in \mathcal{F}^S_t$ for any $b \in \mathbb{R}$ and $u \leqslant s$. Such sets satisfy

$$\{U_u < b\} \cap \{S_t > s\} = \{S_b > u\} \cap \{S_t > s\},$$

which proves that they are in $\mathcal{F}^S_t$ if $b \leqslant t$. In the case $b > t$, the nondecreasing paths of $(S_t)_{t \geqslant 0}$ entail that $\{S_b > u\} \cap \{S_t > s\} = \{S_t > s\}$, and thus the set also belongs to $\mathcal{F}^S_t$.

Now, let

$$\mathcal{A} = \{A \in \mathcal{F}^U_s : A \cap \{S_t > s\} \in \mathcal{F}^S_t\}.$$

The collection $\mathcal{A}$ clearly contains Ω and it is easy to check that it is closed under countable intersections. We prove that $\mathcal{A}$ is closed under complementation. Let $A \in \mathcal{A}$. Then,

$$\overline{A} \cap \{S_t > s\} = \overline{A \cup \{S_t \leqslant s\}}.$$

Moreover,

$$A \cup \{S_t \leqslant s\} = \{S_t \leqslant s\} \cup (A \cap \{S_t > s\}).$$

Since $\{S_t \leqslant s\} \in \mathcal{F}^S_t$ is trivial and $A \cap \{S_t > s\} \in \mathcal{F}^S_t$ follows from $A \in \mathcal{A}$, we have that $A \cup \{S_t \leqslant s\} \in \mathcal{F}^S_t$. We deduce that $\overline{A} \cap \{S_t > s\} \in \mathcal{F}^S_t$, and thus $\overline{A} \in \mathcal{A}$. The collection $\mathcal{A}$ is therefore a σ-algebra, and from our previous findings,

$$\{\{U_u < b\} : b \in \mathbb{R}, u \in [0, s]\} \subset \mathcal{A}.$$

As a consequence, $\mathcal{F}_s^U = \mathcal{A}$. The conclusion follows. $\square$

Lemma 4.3. *Let $t > 0$ and $x_1 \geqslant 0$. Then*

$$\{U_{S_t} \geqslant x_1\} = \begin{cases} \Omega & \text{if } x_1 \leqslant t \\ \{S_t = S_{x_1}\} & \text{if } x_1 > t. \end{cases}$$

Proof. The case $x_1 \leqslant t$ is obvious since $U_{S_t} \geqslant t$ (Proposition 4.1). Let $x_1 > t$ and $\omega \in \{S_{x_1} > S_t\}$. Then there exists $\xi \in (S_t(\omega), S_{x_1}(\omega))$. It follows that $U_{S_t(\omega)}(\omega) < U_\xi(\omega) \leqslant x_1$ and thus $\{U_{S_t} \geqslant x_1\} \cap \{S_{x_1} > S_t\} = \emptyset$. Since $(S_u)_{u\geqslant 0}$ is nondecreasing, we have

$$\begin{aligned} \{U_{S_t} \geqslant x_1\} &= \{U_{S_t} \geqslant x_1\} \cap \{S_{x_1} \geqslant S_t\} \\ &= (\{U_{S_t} \geqslant x_1\} \cap \{S_{x_1} > S_t\}) \cup (\{U_{S_t} \geqslant x_1\} \cap \{S_{x_1} = S_t\}) \\ &= \{U_{S_t} \geqslant x_1\} \cap \{S_{x_1} = S_t\} \end{aligned}$$

so that $\{U_{S_t} \geqslant x_1\} \subset \{S_{x_1} = S_t\}$. For the converse inclusion, let $\omega \in \{S_{x_1} = S_t\}$. Proposition 4.1 implies that $x_1 \leqslant U_{S_{x_1}(\omega)}(\omega)$. Since $U_{S_{x_1}(\omega)}(\omega) = U_{S_t(\omega)}(\omega)$, we conclude that $\omega \in \{U_{S_t} \geqslant x_1\}$. $\square$

Lemma 4.4. *Let $0 < t < x_1$. Then*

$$\{S_t = S_{x_1}\} = \bigcap_{n\in\mathbb{N}} \bigcup_{\substack{(q_1,q_2)\in\mathbb{Q}\times\mathbb{Q} \\ 0\leqslant q_1<q_2<q_1+\frac{1}{n}}} \Big[\{U_{q_1} < t\} \cap \{U_{q_2} \geqslant x_1\}\Big]. \tag{4.8}$$

Proof. We denote by C the set at the right-hand side of Equation (4.8). Let $\omega \in C$. We can find sequences $(q_{1,n})_{n\in\mathbb{N}}$, $(q_{2,n})_{n\in\mathbb{N}}$ such that for all n, $q_{1,n} < q_{2,n} < q_{1,n} + \frac{1}{n}$, $U_{q_{1,n}}(\omega) < t$ and $U_{q_{2,n}}(\omega) \geqslant x_1$. By definition of $(S_t)_{t\geqslant 0}$, we have

$$q_{1,n} \leqslant S_t(\omega) \leqslant S_{x_1}(\omega) \leqslant q_{2,n} < q_{1,n} + \frac{1}{n}.$$

It is then clear that both the sequences $(q_{1,n})_{n\in\mathbb{N}}$, $(q_{2,n})_{n\in\mathbb{N}}$ are convergent and converge towards the same limit. We get $\omega \in \{S_t = S_{x_1}\}$ as a conclusion.

Conversely, let $\omega \in \{S_t = S_{x_1}\}$. For each n, let $q_{1,n} \in (S_t(\omega) - \frac{1}{2n}, S_t(\omega)) \cap \mathbb{Q}^+$ and $q_{2,n} \in (S_{x_1}(\omega), S_{x_1}(\omega) + \frac{1}{2n}) \cap \mathbb{Q}^+$. We clearly have $U_{q_{1,n}}(\omega) < t$, $U_{q_{2,n}}(\omega) \geqslant x_1$, $q_{1,n} < q_{2,n}$ and $q_{2,n} - q_{1,n} < \frac{1}{n}$, the latter being a consequence of $S_t(\omega) = S_{x_1}(\omega)$. This proves that $\omega \in C$. $\square$

Corollary 4.1. *For any $t \geqslant 0$, $\mathcal{F}_t^{U\circ S} \subset \mathcal{F}_\infty^U$.*

Proof. This is a consequence of Lemmas 4.3 and 4.4, as they imply that $\{U_{S_s} \geqslant x\} \in \mathcal{F}^U_\infty$ for all $s \leqslant t$. $\square$

Lemma 4.5. *For any $t \geqslant 0$, we have $\mathcal{F}^{U\circ S,S}_t = \mathcal{F}^U_{S_t}$.*

Proof. We begin with the proof of $\mathcal{F}^{U\circ S,S}_t \subset \mathcal{F}^U_{S_t}$. Since Proposition 4.2 states that $\mathcal{F}^S_t \subset \mathcal{F}^U_{S_t}$, we only need to prove $\mathcal{F}^{U\circ S}_t \subset \mathcal{F}^U_{S_t}$. This second inclusion follows easily from Lemma 4.2 (i): for any $s \geqslant 0$, $\sigma(U_{S_s}) \subset \mathcal{F}^U_{S_s}$, but this implies that $\sigma(U_{S_s}) \subset \mathcal{F}^U_{S_t}$ whenever $s \leqslant t$. It follows that $\mathcal{F}^{U\circ S}_t \subset \mathcal{F}^U_{S_t}$.

We now prove $\mathcal{F}^{U\circ S,S}_t \supset \mathcal{F}^U_{S_t}$. By Lemma 4.2 (i), $\mathcal{F}^U_{S_t} = \sigma(\mathcal{F}^U_{S_t-}, U_{S_t})$. According to (ii) of the same lemma, $\mathcal{F}^U_{S_t-} \subset \mathcal{F}^S_t$, so that $\mathcal{F}^U_{S_t-} \subset \mathcal{F}^{U\circ S,S}_t$. Since the $\mathcal{F}^{U\circ S,S}_t$-measurability of U_{S_t} is trivial, the result follows. $\square$

We can now prove the Markov property for $(S_t, U_{S_t})_{t\geqslant 0}$.

Theorem 4.1. *The bivariate process $(S_t, U_{S_t})_{t\geqslant 0}$ is a Markov process with respect to its natural filtration.*

Proof. According to Theorem 1.2 and Remark 1.6 (ii) in Chapter 9 of Çınlar (2011), proving that the Markov property holds amounts to showing that

$$\mathbb{E}[f(S_t, U_{S_t})|\mathcal{F}^{S,U\circ S}_s] = \mathbb{E}[f(S_t, U_{S_t})|\sigma(S_s, U_{S_s})]$$

for any $t \geqslant s \geqslant 0$ and any bounded positive Borel measurable function $f : \mathbb{R}^2 \to \mathbb{R}^+$. Let

$$\mathcal{A} = \left\{A \in \sigma(S_t, U_{S_t}) : \mathbb{E}[\mathbf{1}_A|\mathcal{F}^{S,U\circ S}_s] = \mathbb{E}[\mathbf{1}_A|\sigma(S_s, U_{S_s})]\right\}$$

and

$$\mathcal{C} = \left\{\{S_t \leqslant x_1\} \cap \{U_{S_t} \geqslant x_2\} : x_1, x_2 \in \mathbb{R}\right\}.$$

Note that $\mathcal{C}$ is a π-system and $\mathcal{A}$ is a λ-system. Since the collection $\{(-\infty, x_1] \times [x_2, +\infty) : x_1, x_2 \in \mathbb{R}\}$ generates the Borel σ-algebra $\mathcal{B}_{\mathbb{R}^2}$, it is clear that $\sigma(\mathcal{C}) = \sigma(S_t, U_{S_t})$. We prove now that $\mathcal{C} \subset \mathcal{A}$. To this end, we start by observing that

$$\begin{aligned}
&\{S_t \leqslant x_1\} \cap \{U_{S_t} \geqslant x_2\} \\
&\quad = \Big(\{S_t \leqslant x_1\} \cap \{U_{S_t} \geqslant x_2\} \cap \{S_s = S_t\}\Big) \cup \Big(\{S_t \leqslant x_1\} \cap \{U_{S_t} \geqslant x_2\} \cap \{S_s < S_t\}\Big) \\
&\quad = \Big(\{S_s \leqslant x_1\} \cap \{U_{S_s} \geqslant x_2\} \cap \{S_s = S_t\}\Big) \cup \Big(\{S_t \leqslant x_1\} \cap \{U_{S_t} \geqslant x_2\} \cap \{S_s < S_t\}\Big) \\
&\quad = \Big(\{S_s \leqslant x_1\} \cap \{U_{S_s} \geqslant x_2 \vee t\}\Big) \cup \Big(\{S_t \leqslant x_1\} \cap \{U_{S_t} \geqslant x_2\} \cap \{S_s < S_t\}\Big).
\end{aligned}$$

Since $\{S_s \leqslant x_1\} \cap \{U_{S_s} \geqslant x_2 \vee t\}$ is in $\sigma(S_s, U_{S_s})$, it remains to show that

$$\mathbb{E}[\mathbf{1}_{\{S_t \leqslant x_1\} \cap \{U_{S_t} \geqslant x_2\} \cap \{S_s < S_t\}} | \mathcal{F}_s^{S, U\circ S}] = \mathbb{E}[\mathbf{1}_{\{S_t \leqslant x_1\} \cap \{U_{S_t} \geqslant x_2\} \cap \{S_s < S_t\}} | \sigma(S_s, U_{S_s})].$$

To do so, we proceed in two steps. In the first step we will assume that $x_2 \leqslant t$, so that $\{U_{S_t} \geqslant x_2\} = \Omega$. This case is the most simple. In the second step, we will work with $x_2 > t$.

Let $x_2 \leqslant t$. Then $\{S_t \leqslant x_1\} \cap \{U_{S_t} \geqslant x_2\} \cap \{S_s < S_t\} = \{S_t \leqslant x_1\} \cap \{S_s < S_t\}$. We will establish that

$$\begin{aligned} &\{S_t \leqslant x_1\} \cap \{S_s < S_t\} \\ &= \left(\bigcap_{n \in \mathbb{N}} \bigcup_{\substack{q \in \mathbb{Q} \\ q > 0}} \left[\{U_{S_s + q} \geqslant t\} \cap \{S_s + q < x_1 + n^{-1}\} \right] \right) \cap \{S_s < S_t\}. \end{aligned} \tag{4.9}$$

Let ω be a member of the set at the left-hand side of Equation (4.9). Let $n \in \mathbb{N}$ and $q \in (S_t(\omega) - S_s(\omega), x_1 + n^{-1} - S_s(\omega)) \cap \mathbb{Q}$. Such a q exists because

$$S_t(\omega) - S_s(\omega) \leqslant x_1 - S_s(\omega) < x_1 + n^{-1} - S_s(\omega).$$

Moreover, $q + S_s(\omega) > S_t(\omega)$ ensures that $U_{q+S_s(\omega)}(\omega) > t$, so that $\omega \in \{U_{S_s+q} \geqslant t\}$. Conversely, let ω be a member of the set at the right-hand side of Equation (4.9). For each $n \in \mathbb{N}$, let $q_n \in \mathbb{Q}, q_n > 0$ satisfy $U_{S_s(\omega)+q_n}(\omega) \geqslant t$ and $q_n < x_1 + n^{-1} - S_s(\omega)$. Since $S_s(\omega) + q_n < x_1 + n^{-1}$ for any n, we have

$$t \leqslant U_{S_s(\omega)+q_n}(\omega) < U_{x_1+n^{-1}}(\omega).$$

By the right-continuity of $(U_t)_{t\geqslant 0}$, $\lim_{n\to+\infty} U_{x_1+n^{-1}}(\omega) = U_{x_1}(\omega) \geqslant t$. By definition of $(S_t)_{t\geqslant 0}$, this is equivalent to $S_t(\omega) \leqslant x_1$. This proves that $\omega \in \{S_t \leqslant x_1\}$ and finishes the proof of Equation (4.9). Since the bivariate process $(U_t, t)_{t\geqslant 0}$ is a Feller process[1] with respect to its natural filtration $(\mathcal{F}_t^U)_{t\geqslant 0}$, it follows that if we define the $\mathcal{F}_\infty^U$-measurable indicator random variable Z as

$$Z := \mathbf{1}\left\{ \bigcap_{n \in \mathbb{N}} \bigcup_{\substack{q \in \mathbb{Q} \\ q > 0}} \left[\{U_q \geqslant t\} \cap \{q < x_1 + n^{-1}\} \right] \right\}$$

then Equation (4.9) implies that

$$\mathbf{1}_{\{S_t \leqslant x_1\} \cap \{S_s < S_t\}} = \mathbf{1}_{\{S_s < S_t\}} Z \circ \theta_{S_s}.$$

By using the $\mathcal{F}_{S_s}^U$-measurablility of $\mathbf{1}_{\{S_s < S_t\}}$ and Theorem 3.1 from the Chapter 3 of Revuz and Yor (2005), we find

[1] this follows from the fact that it is a Lévy process.

$$\begin{aligned}\mathbb{E}[\mathbf{1}_{\{S_t\leqslant x_1\}\cap\{S_s<S_t\}}|\mathcal{F}^U_{S_s}] &= \mathbf{1}_{\{S_s<S_t\}}\mathbb{E}[Z\circ\theta_{S_s}|\mathcal{F}^U_{S_s}]\\ &= \mathbf{1}_{\{S_s<S_t\}}\mathbb{E}^{(U_{S_s},S_s)}[Z].\end{aligned} \tag{4.10}$$

Similarly, Lemma 4.3 entails that $\mathbf{1}_{\{S_s<S_t\}}$ is $\sigma(S_s, U_{S_s})$-measurable and thus by the tower property,

$$\begin{aligned}\mathbb{E}[\mathbf{1}_{\{S_t\leqslant x_1\}\cap\{S_s<S_t\}}|\sigma(S_s,U_{S_s})] &= \mathbb{E}[\mathbf{1}_{\{S_s<S_t\}}\mathbb{E}[Z\circ\theta_{S_s}|\mathcal{F}^U_{S_s}]|\sigma(S_s,U_{S_s})]\\ &= \mathbf{1}_{\{S_s<S_t\}}\mathbb{E}[\mathbb{E}^{(U_{S_s},S_s)}[Z]|\sigma(S_s,U_{S_s})]\\ &= \mathbf{1}_{\{S_s<S_t\}}\mathbb{E}^{(U_{S_s},S_s)}[Z].\end{aligned} \tag{4.11}$$

By combining Equations (4.10) and (4.11), it follows that

$$\mathbb{E}[\mathbf{1}_{\{S_t\leqslant x_1\}\cap\{S_s<S_t\}}|\mathcal{F}^U_{S_s}] = \mathbb{E}[\mathbf{1}_{\{S_t\leqslant x_1\}\cap\{S_s<S_t\}}|\sigma(S_s,U_{S_s})]$$

whenever $x_2 \leqslant t$.

We now consider the case $x_2 > t$. We will establish that when $x_2 > t$,

$$\begin{aligned}&\{S_t\leqslant x_1\}\cap\{U_{S_t}\geqslant x_2\}\cap\{S_s<S_t\}\\ &= \left(\bigcap_{n\in\mathbb{N}}\ \bigcup_{\substack{(q_1,q_2)\in\mathbb{Q}^2\\ 0<q_1<q_2<q_1+\frac{1}{n}}}\begin{bmatrix}\{U_{S_s+q_1}<t\}\cap\{U_{S_s+q_2}\geqslant x_2\}\\ \cap\,\{S_s+q_2<x_1+n^{-1}\}\end{bmatrix}\right)\cap\{S_s<S_t\}.\end{aligned} \tag{4.12}$$

Let ω be a member of the first set of Equation (4.12) and $n\in\mathbb{N}$. Let

$$q_1 \in (0\vee(S_t(\omega)-S_s(\omega)-(2n)^{-1}), S_t(\omega)-S_s(\omega))\cap\mathbb{Q}$$

and

$$q_2 \in \big(S_t(\omega)-S_s(\omega), (S_t(\omega)-S_s(\omega)+(2n)^{-1})\wedge(x_1+n^{-1}-S_s(\omega))\big)\cap\mathbb{Q}.$$

The existence of q_1 follows from $S_t(\omega) > S_s(\omega)$ whereas the existence of q_2 follows from $S_t(\omega)\leqslant x_1 < x_1+n^{-1}$. The inequality $S_s(\omega)+q_1 < S_t(\omega)$ implies that $U_{S_s(\omega)+q_1}(\omega) < t$. By Lemma 4.3, we have $S_{x_2}(\omega) = S_t(\omega)$, so that $S_s(\omega)+q_2 > S_t(\omega)$ implies that $S_s(\omega)+q_2 > S_{x_2}(\omega)$, which in turn implies that $U_{S_s(\omega)+q_2}(\omega)\geqslant x_2$. Finally, note that by construction, it also holds that $0<q_1<q_2<q_1+n^{-1}$. This proves that ω is also a member of the second set of Equation (4.12). We prove now the reverse inclusion: assume that ω is a member of the second set of Equation (4.12). For each n, let $q_{1,n}, q_{2,n}$ be numbers in $\mathbb{Q}$ that satisfy $0<q_{1,n}<q_{2,n}<q_{1,n}+n^{-1}$, $U_{S_s(\omega)+q_{1,n}}(\omega) < t$, $U_{S_s(\omega)+q_{2,n}}(\omega)\geqslant x_2 > t$ and $q_{2,n} < x_1+n^{-1}-S_s(\omega)$. By construction, $S_s(\omega)+q_{1,n} < S_t(\omega)$, thus

$$x_2 \leqslant U_{S_s(\omega)+q_{2,n}}(\omega) \leqslant U_{S_s(\omega)+q_{1,n}+n^{-1}}(\omega) \leqslant U_{S_t(\omega)+n^{-1}}(\omega)$$

and the right-continuity of $(U_t)_{t\geqslant 0}$ implies that $U_{S_t(\omega)}(\omega) \geqslant x_2$. Similarly, $S_s(\omega) + q_{2,n} < x_1 + n^{-1}$ implies that

$$t < x_2 < U_{S_s(\omega)+q_{2,n}}(\omega) < U_{x_1+n^{-1}}(\omega)$$

and $U_{x_1}(\omega) > t$ follows from the right-continuity of $(U_t)_{t\geqslant 0}$. We get $S_t(\omega) \leqslant x_1$ as a consequence, which ends the proof of Equation (4.12). As a consequence, the $\mathcal{F}^U_\infty$-measurable random variable Z defined as

$$Z := \mathbf{1}\left\{\bigcap_{n\in\mathbb{N}} \bigcup_{\substack{(q_1,q_2)\in\mathbb{Q}^2 \\ 0<q_1<q_2<q_1+\frac{1}{n}}} \left[\{U_{q_1} < t\} \cap \{U_{q_2} \geqslant x_2\} \cap \{q_2 < x_1 + n^{-1}\}\right]\right\}$$

satisfies

$$\mathbf{1}_{\{S_t\leqslant x_1\}\cap\{U_{S_t}\geqslant x_2\}\cap\{S_s<S_t\}} = \mathbf{1}_{\{S_s<S_t\}} Z\circ\theta_{S_s}.$$

Again, by using the $\mathcal{F}^U_{S_s}$-measurability of $\mathbf{1}_{\{S_s<S_t\}}$ and Theorem 3.1 from the chapter 3 of Revuz and Yor (2005), we find

$$\begin{aligned}\mathbb{E}[\mathbf{1}_{\{S_t\leqslant x_1\}\cap\{U_{S_t}\geqslant x_2\}\cap\{S_s<S_t\}}|\mathcal{F}^U_{S_s}] &= \mathbf{1}_{\{S_s<S_t\}}\mathbb{E}[Z\circ\theta_{S_s}|\mathcal{F}^U_{S_s}] \\ &= \mathbf{1}_{\{S_s<S_t\}}\mathbb{E}^{(U_{S_s},S_s)}[Z].\end{aligned} \tag{4.13}$$

Similarly, Lemma 4.3 entails that $\mathbf{1}_{\{S_s<S_t\}}$ is $\sigma(S_s, U_{S_s})$-measurable and thus by the tower property,

$$\begin{aligned}&\mathbb{E}[\mathbf{1}_{\{S_t\leqslant x_1\}\cap\{U_{S_t}\geqslant x_2\}\cap\{S_s<S_t\}}|\sigma(S_s,U_{S_s})] \\ &\quad= \mathbb{E}[\mathbf{1}_{\{S_s<S_t\}}\mathbb{E}[Z\circ\theta_{S_s}|\mathcal{F}^U_{S_s}]|\sigma(S_s,U_{S_s})] \\ &\quad= \mathbf{1}_{\{S_s<S_t\}}\mathbb{E}[\mathbb{E}^{(U_{S_s},S_s)}[Z]|\sigma(S_s,U_{S_s})] \\ &\quad= \mathbf{1}_{\{S_s<S_t\}}\mathbb{E}^{(U_{S_s},S_s)}[Z].\end{aligned} \tag{4.14}$$

By combining Equations (4.13) and (4.14), it follows that

$$\begin{aligned}&\mathbb{E}[\mathbf{1}_{\{S_t\leqslant x_1\}\cap\{U_{S_t}\geqslant x_2\}\cap\{S_s<S_t\}}|\mathcal{F}^U_{S_s}] \\ &\quad= \mathbb{E}[\mathbf{1}_{\{S_t\leqslant x_1\}\cap\{U_{S_t}\geqslant x_2\}\cap\{S_s<S_t\}}|\sigma(S_s,U_{S_s})]\end{aligned} \tag{4.15}$$

whenever $x_2 > t$. It follows that $\mathcal{C} \subset \mathcal{A}$ and thus by Dynkin's π-λ theorem, for any $A \in \sigma(\mathcal{C}) = \sigma(S_t, U_{S_t})$,

$$\mathbb{E}[\mathbf{1}_A|\mathcal{F}^{S,U\circ S}_s] = \mathbb{E}[\mathbf{1}_A|\sigma(S_s,U_{S_s})]. \tag{4.16}$$

Let $f : \mathbb{R}^2 \to \mathbb{R}^+$ be a bounded positive Borel measurable function. Define the sequence of functions $(f_n : \mathbb{R}^2 \to \mathbb{R}^+)_{n\in\mathbb{N}}$ as

$$f_n(x,y) := \sum_{k=0}^{2^{2n}} \frac{k}{2^n} \mathbf{1}_{\{f(x,y)\in[k/2^n,(k+1)/2^n)\}}.$$

We clearly have that $0 \leqslant f_n \leqslant f_{n+1} \leqslant f$ for all $n \in \mathbb{N}$ and $f_n \to f$. Therefore, Equation (4.16) and two applications of the monotone convergence theorem yield

$$\begin{aligned}
\mathbb{E}[f(S_t, U_{S_t})|\mathcal{F}_s^{S,U\circ S}] &= \mathbb{E}\left[\lim_{n\to+\infty} f_n(S_t, U_{S_t})|\mathcal{F}_s^{S,U\circ S}\right] \\
&= \lim_{n\to+\infty} \mathbb{E}[f_n(S_t, U_{S_t})|\mathcal{F}_s^{S,U\circ S}] \\
&= \lim_{n\to+\infty} \sum_{k=0}^{2^{2n}} \frac{k}{2^n} \mathbb{E}\left[\mathbf{1}_{\left\{(S_t,U_{S_t})\in f^{-1}\left[\frac{k}{2^n},\frac{k+1}{2^n}\right)\right\}}|\mathcal{F}_s^{S,U\circ S}\right] \\
&= \lim_{n\to+\infty} \sum_{k=0}^{2^{2n}} \frac{k}{2^n} \mathbb{E}\left[\mathbf{1}_{\left\{(S_t,U_{S_t})\in f^{-1}\left[\frac{k}{2^n},\frac{k+1}{2^n}\right)\right\}}|\sigma(S_s, U_{S_s})\right] \\
&= \lim_{n\to+\infty} \mathbb{E}[f_n(S_t, U_{S_t})|\sigma(S_s, U_{S_s})] \\
&= \mathbb{E}\left[\lim_{n\to+\infty} f_n(S_t, U_{S_t})|\sigma(S_s, U_{S_s})\right] \\
&= \mathbb{E}[f(S_t, U_{S_t})|\sigma(S_s, U_{S_s})],
\end{aligned}$$

which is what we had to prove. □

In the next section, we study the transform of our asset process. This will be useful to compute option prices.

4.3 FPDE's for densities and Fourier transforms

This section is devoted to the derivation of FPDE's for the joint density and transform of the bivariate process $(X_{S_t}, \lambda_{S_t})_{t\geqslant 0}$[2]. Before diving into the main topic of this section, let us start by introducing the filtrations we consider for this process. In the non-fractional market, the filtration of the asset price $(A_t)_{t\geqslant 0}$ is denoted by $(\mathcal{F}_t^A)_{t\geqslant 0}$ and is defined as

$$\mathcal{F}_t^A = \mathcal{N} \vee \sigma\{W_s : 0 \leqslant s \leqslant t\} \vee \sigma\{J_s : 0 \leqslant s \leqslant t\} = \mathcal{F}_t^W \vee \mathcal{F}_t^J,$$

[2]Recall that $(X_t)_{t\geqslant 0}$ is the logarithm of the asset price, i.e. $X_t = \ln A_t$.

where $(\mathcal{F}_t^W)_{t\geqslant 0}$ and $(\mathcal{F}_t^J)_{t\geqslant 0}$ are the completed natural filtrations of the Brownian motion $(W_t)_{t\geqslant 0}$ and the jump process $(J_t)_{t\geqslant 0}$ respectively. In other words, we assume that the impacts on the asset price of the Brownian component (i.e. the noise component) and the jump component can be isolated from one another. Recall that $\mathcal{N}$ denotes the null sets of the probability space and are in $\mathcal{F}_t^W \vee \mathcal{F}_t^J$ because we assumed that the natural filtrations are completed. Note that the natural filtration $(\mathcal{F}_t^J)_{t\geqslant 0}$ carries the information about the other jump processes $(D_t)_{t\geqslant 0}$ and $(L_t)_{t\geqslant 0}$ and about the intensity $(\lambda_t)_{t\geqslant 0}$, that is

$$\sigma\{D_s : 0 \leqslant s \leqslant t\} \vee \sigma\{L_s : 0 \leqslant s \leqslant t\} \vee \sigma\{\lambda_s : 0 \leqslant s \leqslant t\} \subset \mathcal{F}_t^J,$$

for each $t \geqslant 0$. Let us now introduce the filtration for the fractional market. Let $(\mathcal{F}_t^{A,U})_{t\geqslant 0}$ be defined as $\mathcal{F}_t^{A,U} = \mathcal{F}_t^A \vee \mathcal{F}_t^U$. Since S_s is a $(\mathcal{F}_t^U)_{t\geqslant 0}$-stopping-time for all $s \geqslant 0$, it is also a $(\mathcal{F}_t^{A,U})_{t\geqslant 0}$-stopping-time. Therefore, we can introduce the filtration $(\mathcal{F}_t^{A\circ S})_{t\geqslant 0}$ defined as the time-changed filtration $\mathcal{F}_t^{A\circ S} = \mathcal{F}_{S_t}^{A,U}$. It is well known from the chapter 10 of Jacod (1979) that if a process $(A_t)_{t\geqslant 0}$ is a $(\mathcal{F}_t^{A,U})_{t\geqslant 0}$-semimartingale, then the time-changed process $(A_{S_t})_{t\geqslant 0}$ is a $(\mathcal{F}_{S_t}^{A,U})_{t\geqslant 0}$-semimartingale[3]. Therefore, the filtration of the asset price $(A_{S_t})_{t\geqslant 0}$ can be chosen to be $(\mathcal{F}_t^{A\circ S})_{t\geqslant 0}$.

We move now to the main topic of this section. We first need to derive PDE's for the joint density and transform of the process $(X_t, \lambda_t)_{t\geqslant 0}$. Afterwards, we use the PDE of $(X_t, \lambda_t)_{t\geqslant 0}$ to show that the joint density and transform of the time-changed process $(X_{S_t}, \lambda_{S_t})_{t\geqslant 0}$ satisfy similar equations, with the difference that the partial derivative with respect to time is replaced by a fractional Caputo derivative. In the non-fractional case, we work conditionally on the last known value for $(X_t, \lambda_t)_{t\geqslant 0}$. This is justified because self-exciting Hawkes processes satisfy the Markov property provided that we keep track of their intensity. More precisely, $(J_t, \lambda_t)_{t\geqslant 0}$ is a Markov process whereas the jump process $(J_t)_{t\geqslant 0}$ alone is not.

Proposition 4.3. *Let $p(t, x_1, x_2|s, y_1, y_2)$ be the bivariate probability density function (PDF) of (X_t, λ_t) given $(X_s, \lambda_s) = (y_1, y_2)$, $s < t$, i.e.*

$$p(t, x_1, x_2|s, y_1, y_2) := \frac{\partial^2}{\partial x_1 \partial x_2}\mathbb{P}(X_t \leqslant x_1, \lambda_t \leqslant x_2 | X_s = y_1, \lambda_s = y_2).$$

It will be denoted by $p(t, x_1, x_2|\cdot)$ for shorter notations. Then p satisfies the PDE

$$\begin{aligned}\frac{\partial}{\partial t}p(t,x_1,x_2|\cdot) &= -\left(\mu - \frac{\sigma^2}{2} - x_2\mathbb{E}[e^{\xi}-1]\right)\frac{\partial}{\partial x_1}p(t,x_1,x_2|\cdot)\\ &+ \frac{\sigma^2}{2}\frac{\partial^2}{\partial x_1^2}p(t,x_1,x_2|\cdot) - \kappa(\theta - x_2)\frac{\partial}{\partial x_2}p(t,x_1,x_2|\cdot) + \kappa p(t,x_1,x_2|\cdot)\\ &- \eta\mathbb{E}[|\xi|p(t,x_1-\xi,x_2-\eta|\xi||\cdot)] + \mathbb{E}[p(t,x_1-\xi,x_2-\eta|\xi||\cdot) - p(t,x_1,x_2|\cdot)],\end{aligned}$$

[3]This is stated in Theorem 10.16, which can be applied to our case because our time-change $(S_t)_{t\geqslant 0}$ is both finite and continuous. We shall come back more precisely to this point in the section about changes of measures.

with initial condition $p(s, x_1, x_2|s, y_1, y_2) = \delta_{\mathbb{R}^2}(x_1 - y_1, x_2 - y_2)$, $\delta_{\mathbb{R}^2}$ *being the Dirac measure located at* $(0,0) \in \mathbb{R}^2$.

Proof. This proof relies on the Cramers-Moyal expansion. It states that the bivariate PDF p satisfies

$$\begin{aligned} &p(t+\Delta, x_1, x_2|\cdot) - p(t, x_1, x_2|\cdot) \\ &= \sum_{n=1}^{+\infty} \sum_{j=0}^{n} \frac{(-1)^n}{j!(n-j)!} \frac{\partial^j}{\partial x_1^j} \frac{\partial^{n-j}}{\partial x_2^{n-j}} (\mathbb{M}(j, n-j, \Delta|t, x_1, x_2) p(t, x_1, x_2|\cdot)) \end{aligned} \tag{4.17}$$

where $\mathbb{M}(j, n-j, \Delta|t, x_1, x_2) = \mathbb{E}[(X_{t+\Delta} - X_t)^j (\lambda_{t+\Delta} - \lambda_t)^{n-j} | X_t = x_1, \lambda_t = x_2]$. The proof of this statement can be found in Hainaut (2020). Applying Ito's Lemma on the functions $\mathbb{R}^2 \ni (x, y) \mapsto h_{j,n}(x, y) := x^j y^{n-j}$ leads to

$$\begin{aligned} &h_{j,n}(X_{t+\Delta} - X_t, \lambda_{t+\Delta} - \lambda_t) \\ &= j \int_{s=t}^{t+\Delta} (X_{s-} - X_t)^{j-1} (\lambda_{s-} - \lambda_t)^{n-j} \left(\left[\mu - \frac{\sigma^2}{2} - \lambda_{s-} \mathbb{E}[e^{\xi} - 1] \right] \mathrm{d}s + \sigma \mathrm{d}W_s \right) \\ &+ \frac{j(j-1)\sigma^2}{2} \int_{s=t}^{t+\Delta} (X_{s-} - X_t)^{j-2} (\lambda_{s-} - \lambda_t)^{n-j} \mathrm{d}s \\ &+ (n-j) \int_{s=t}^{t+\Delta} (X_{s-} - X_t)^{j} (\lambda_{s-} - \lambda_t)^{n-j-1} \kappa(\theta - \lambda_{s-}) \mathrm{d}s \\ &+ \int_{s=t}^{t+\Delta} \int_{z \in \mathbb{R}} \left((X_{s-} + z - X_t)^j - (X_{s-} - X_t)^j \right) \\ &\qquad \left((\lambda_{s-} + \eta|z| - \lambda_t)^j - (\lambda_{s-} - \lambda_t)^j \right) \Xi(\mathrm{d}z, \mathrm{d}s) \end{aligned} \tag{4.18}$$

Let us now examine the expectation of each term of Equation (4.18). We have for the first term

$$\begin{aligned} &\mathbb{E}\left[j \int_{s=t}^{t+\Delta} (X_{s-} - X_t)^{j-1} (\lambda_{s-} - \lambda_t)^{n-j} \left(\left[\mu - \frac{\sigma^2}{2} - \lambda_{s-} \mathbb{E}[e^{\xi} - 1] \right] \mathrm{d}s + \sigma \mathrm{d}W_s \right) \right] \\ &= \begin{cases} \left(\mu - \frac{\sigma^2}{2} - \lambda_t \mathbb{E}[e^{\xi} - 1] \right) \Delta + \mathcal{O}(\Delta^2) & \text{if } n = j = 1 \\ \mathcal{O}(\Delta^2) & \text{otherwise,} \end{cases} \end{aligned} \tag{4.19}$$

whereas the other terms can be respectively written as

$$\begin{aligned}&\mathbb{E}\left[\frac{j(j-1)\sigma^2}{2}\int_{s=t}^{t+\Delta}(X_{s-}-X_t)^{j-2}(\lambda_{s-}-\lambda_t)^{n-j}\mathrm{d}s\right]\\&=\begin{cases}\sigma^2\Delta+\mathcal{O}(\Delta^2) & \text{if } n=j=2\\ \mathcal{O}(\Delta^2) & \text{otherwise,}\end{cases}\end{aligned}\tag{4.20}$$

$$\begin{aligned}&\mathbb{E}\left[(n-j)\int_{s=t}^{t+\Delta}(X_{s-}-X_t)^{j}(\lambda_{s-}-\lambda_t)^{n-j-1}\kappa(\theta-\lambda_{s-})\mathrm{d}s\right]\\&=\begin{cases}\kappa(\theta-\lambda_t)\Delta+\mathcal{O}(\Delta^2) & \text{if } n=1, j=0\\ \mathcal{O}(\Delta^2) & \text{otherwise,}\end{cases}\end{aligned}\tag{4.21}$$

and

$$\begin{aligned}&\mathbb{E}\left[\begin{array}{l}+\displaystyle\int_{s=t}^{t+\Delta}\int_{z\in\mathbb{R}}\Big((X_{s-}+z-X_t)^j-(X_{s-}-X_t)^j\Big)\\ \qquad\qquad\Big((\lambda_{s-}+\eta|z|-\lambda_t)^j-(\lambda_{s-}-\lambda_t)^j\Big)\,\Xi(\mathrm{d}z,\mathrm{d}s)\end{array}\right]\\&=\lambda_t\eta^{n-j}\mathbb{E}[\xi^j|\xi|^{n-j}]\Delta+\mathcal{O}(\Delta^2).\end{aligned}\tag{4.22}$$

Using Equations (4.17) to (4.22), we obtain

$$\begin{aligned}&\frac{p(t+\Delta,x_1,x_2|\cdot)-p(t,x_1,x_2|\cdot)}{\Delta}\\&=-\frac{\partial}{\partial x_1}\left[\left(\mu-\frac{\sigma^2}{2}-x_2\mathbb{E}[e^\xi-1]\right)p(t,x_1,x_2|\cdot)\right]\\&+\frac{\sigma^2}{2}\frac{\partial^2}{\partial x_1^2}p(t,x_1,x_2|\cdot)-\frac{\partial}{\partial x_2}\left[\kappa(\theta-x_2)p(t,x_1,x_2|\cdot)\right]\\&+\mathbb{E}\left[\sum_{n=1}^{+\infty}\sum_{j=0}^{n}\frac{(-\xi)^j(-|\eta\xi|)^{n-j}}{j!(n-j)!}\frac{\partial^j}{\partial x_1^j}\frac{\partial^{n-j}}{\partial x_2^{n-j}}\left(x_2p(t,x_1,x_2|\cdot)\right)\right]+\mathcal{O}(\Delta)\\&=-\left(\mu-\frac{\sigma^2}{2}-x_2\mathbb{E}[e^\xi-1]\right)\frac{\partial}{\partial x_1}p(t,x_1,x_2|\cdot)\\&+\frac{\sigma^2}{2}\frac{\partial^2}{\partial x_1^2}p(t,x_1,x_2|\cdot)-\kappa(\theta-x_2)\frac{\partial}{\partial x_2}p(t,x_1,x_2|\cdot)\\&+\kappa p(t,x_1,x_2|\cdot)\\&+\mathbb{E}\left[\sum_{n=1}^{+\infty}\sum_{j=0}^{n}\frac{(-\xi)^j(-|\eta\xi|)^{n-j}}{j!(n-j)!}\frac{\partial^j}{\partial x_1^j}\frac{\partial^{n-j}}{\partial x_2^{n-j}}\left(x_2p(t,x_1,x_2|\cdot)\right)\right]+\mathcal{O}(\Delta).\end{aligned}\tag{4.23}$$

The expectation term of Equation (4.23) can be rewritten as

$$\begin{aligned}
&\mathbb{E}\left[\sum_{n=1}^{+\infty}\sum_{j=0}^{n}\frac{(-\xi)^j(-|\eta\xi|)^{n-j}}{j!(n-j)!}\frac{\partial^j}{\partial x_1^j}\frac{\partial^{n-j}}{\partial x_2^{n-j}}\left(x_2p(t,x_1,x_2|\cdot)\right)\right]\\
&=\mathbb{E}\left[\sum_{n=1}^{+\infty}\sum_{j=0}^{n-1}\frac{(-\xi)^j(-|\eta\xi|)^{n-j}}{j!(n-j)!}\frac{\partial^j}{\partial x_1^j}\frac{\partial^{n-j}}{\partial x_2^{n-j}}\left(x_2p(t,x_1,x_2|\cdot)\right)\right]\\
&+\mathbb{E}\left[\sum_{n=1}^{+\infty}\frac{(-\xi)^n}{n!}\frac{\partial^n}{\partial x_1^n}(x_2p(t,x_1,x_2|\cdot))\right]\\
&=\mathbb{E}\left[-\eta|\xi|\sum_{n=1}^{+\infty}\sum_{j=0}^{n-1}\frac{(-\xi)^j(-|\eta\xi|)^{n-j-1}}{j!(n-j-1)!}\frac{\partial^j}{\partial x_1^j}\frac{\partial^{n-j-1}}{\partial x_2^{n-j-1}}p(t,x_1,x_2|\cdot)\right]\\
&+\mathbb{E}\left[\sum_{n=1}^{+\infty}\sum_{j=0}^{n-1}\frac{(-\xi)^j(-|\eta\xi|)^{n-j}}{j!(n-j)!}x_2\frac{\partial^j}{\partial x_1^j}\frac{\partial^{n-j}}{\partial x_2^{n-j}}p(t,x_1,x_2|\cdot)\right]\\
&+\mathbb{E}\left[\sum_{n=1}^{+\infty}\frac{(-\xi)^n}{n!}x_2\frac{\partial^n}{\partial x_1^n}p(t,x_1,x_2|\cdot)\right]\\
&=\mathbb{E}\left[-\eta|\xi|\sum_{n=0}^{+\infty}\sum_{j=0}^{n}\frac{(-\xi)^j(-|\eta\xi|)^{n-j}}{j!(n-j)!}\frac{\partial^j}{\partial x_1^j}\frac{\partial^{n-j}}{\partial x_2^{n-j}}p(t,x_1,x_2|\cdot)\right]\\
&+\mathbb{E}\left[\sum_{n=1}^{+\infty}\sum_{j=0}^{n}\frac{(-\xi)^j(-|\eta\xi|)^{n-j}}{j!(n-j)!}x_2\frac{\partial^j}{\partial x_1^j}\frac{\partial^{n-j}}{\partial x_2^{n-j}}p(t,x_1,x_2|\cdot)\right]\\
&=-\eta\mathbb{E}[|\xi|p(t,x_1-\xi,x_2-\eta|\xi||\cdot)]\\
&+x_2\mathbb{E}[p(t,x_1-\xi,x_2-\eta|\xi||\cdot)-p(t,x_1,x_2|\cdot)].
\end{aligned}\tag{4.24}$$

The proof is then completed by combining Equations (4.23) and (4.24). □

For $t \geqslant 0$, $z_1 \in \mathbb{C}$ and $z_2 \in \mathbb{R}$, define the conditional transform φ as

$$\varphi(t,z_1,z_2|s,y_1,y_2) := \mathbb{E}[e^{z_1X_t+\mathrm{i}z_2\lambda_t}|X_s=z_1,\lambda_s=y_2].$$

It will also be denoted by $\varphi(t,z_1,z_2|\cdot)$ for shorter notations. The function φ satisfies a PDE.

Proposition 4.4. *The function φ satisfies the PDE*

$$\begin{aligned}
\frac{\partial\varphi}{\partial t}(t,z_1,z_2|\cdot) &= \left(\mathrm{i}z_1\mathbb{E}[e^{\xi}-1]-\kappa z_2+\mathrm{i}-\mathrm{i}\mathbb{E}\left[e^{z_1\xi+\mathrm{i}z_2\eta|\xi|}\right]\right)\frac{\partial\varphi}{\partial z_2}(t,z_1,z_2|\cdot)\\
&+\left(\left(\mu-\frac{\sigma^2}{2}\right)z_1+\frac{(\sigma z_1)^2}{2}+\mathrm{i}\kappa\theta z_2\right)\varphi(t,z_1,z_2|\cdot)
\end{aligned}$$

with boundary conditions $\varphi(s, z_1, z_2|s, y_1, y_2) = e^{z_1 y_1 + \mathrm{i} z_2 y_2}$ *and* $\varphi(s, 0, 0|s, y_1, y_2) = 1$.

Proof. The partial derivative of φ satisfies

$$\begin{aligned}
&\frac{\partial \varphi}{\partial t}(t, z_1, z_2|\cdot) \\
&\quad = \lim_{\Delta \to 0} \frac{\int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{z_1 x_1 + \mathrm{i} z_2 x_2} (p(t+\Delta, x_1, x_2|\cdot) - p(t, x_1, x_2|\cdot)) \mathrm{d}x_1 \mathrm{d}x_2}{\Delta} \\
&\quad = \int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{z_1 x_1 + \mathrm{i} z_2 x_2} \frac{\partial p}{\partial t}(t, x_1, x_2|\cdot) \mathrm{d}x_1 \mathrm{d}x_2.
\end{aligned}$$

Thanks to Proposition 4.3, we find

$$\begin{aligned}
&\frac{\partial \varphi}{\partial t}(t, z_1, z_2|\cdot) \\
&\quad = -\int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{z_1 x_1 + \mathrm{i} z_2 x_2} \left(\mu - \frac{\sigma^2}{2} - x_2 \mathbb{E}[e^{\xi} - 1]\right) \frac{\partial p}{\partial x_1}(t, x_1, x_2|\cdot) \mathrm{d}x_1 \mathrm{d}x_2 \\
&\quad + \frac{\sigma^2}{2} \int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{z_1 x_1 + \mathrm{i} z_2 x_2} \frac{\partial^2 p}{\partial x_1^2}(t, x_1, x_2|\cdot) \mathrm{d}x_1 \mathrm{d}x_2 \\
&\quad - \kappa \int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{z_1 x_1 + \mathrm{i} z_2 x_2} (\theta - x_2) \frac{\partial p}{\partial x_2}(t, x_1, x_2|\cdot) \mathrm{d}x_1 \mathrm{d}x_2 \\
&\quad \kappa \int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{z_1 x_1 + \mathrm{i} z_2 x_2} p(t, x_1, x_2|\cdot) \mathrm{d}x_1 \mathrm{d}x_2 \\
&\quad - \eta \mathbb{E}\left[|\xi| \int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{z_1 x_1 + \mathrm{i} z_2 x_2} p(t, x_1 - \xi, x_2 - \eta|\xi||\cdot) \mathrm{d}x_1 \mathrm{d}x_2\right] \\
&\quad \mathbb{E}\left[\int_{\mathbb{R}^+} \int_{\mathbb{R}} x_2 e^{z_1 x_1 + \mathrm{i} z_2 x_2} (p(t, x_1 - \xi, x_2 - \eta|\xi||\cdot) - p(t, x_1, x_2|\cdot)) \mathrm{d}x_1 \mathrm{d}x_2\right]
\end{aligned}$$

We can now compute all the terms with the help of integrations by parts. To this end, note that if $\omega \in \mathbb{R}$ and $\mathbb{E}[e^{\omega X_t}] < +\infty$, then

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}} e^{\omega x_1} p(t, x_1, x_2|\cdot) \mathrm{d}x_1 \mathrm{d}x_2 < +\infty$$

so that

$$\int_{\mathbb{R}} e^{\omega x_1} p(t, x_1, x_2|\cdot) \mathrm{d}x_1 < +\infty \tag{4.25}$$

for almost all x_2. Fix any $x_2 \in \mathbb{R}$ such that (4.25) holds. Since the function $x_1 \mapsto e^{\omega x_1} p(t, x_1, x_2|\cdot)$ is positive and continuous, Equation (4.25) implies that

$$\lim_{x_1 \to \pm\infty} e^{\omega x_1} p(t, x_1, x_2|\cdot) = 0.$$

Also note that $p(t, x_1, x_2|\cdot)$ vanishes whenever $x_2 \to +\infty$. Moreover, since by definition of $(\lambda_t)_{t\geqslant 0}$, $\lambda_t \geqslant \theta$ for all t, it follows that $p(t, x_1, 0|\cdot) = 0$ for all t, x_1. Bearing these facts in mind, the first term is rewritten as

$$
\begin{aligned}
&- \int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1 + \mathrm{i}z_2x_2}\left(\mu - \frac{\sigma^2}{2} - x_2\mathbb{E}[e^{\xi}-1]\right)\frac{\partial p}{\partial x_1}(t, x_1, x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\\
&= -\left(\mu - \frac{\sigma^2}{2}\right)\int_{\mathbb{R}^+} z_1 e^{\mathrm{i}z_2x_2}\begin{bmatrix} p(t,x_1,x_2|\cdot)\dfrac{e^{z_1x_1}}{z_1} \\ \quad - \displaystyle\int e^{z_1x_1}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\end{bmatrix}_{x_1=-\infty}^{+\infty}\mathrm{d}x_2\\
&+ \mathbb{E}[e^{\xi}-1]\int_{\mathbb{R}^+} x_2 e^{\mathrm{i}z_2x_2} z_1\begin{bmatrix} p(t,x_1,x_2|\cdot)\dfrac{e^{z_1x_1}}{z_1} \\ \quad - \displaystyle\int e^{z_1x_1}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\end{bmatrix}_{x_1=-\infty}^{+\infty}\mathrm{d}x_2\\
&= z_1\left(\mu - \frac{\sigma^2}{2}\right)\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\\
&+ \mathrm{i}z_1\mathbb{E}[e^{\xi}-1]\frac{\partial}{\partial z_2}\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\\
&= z_1\left(\mu - \frac{\sigma^2}{2}\right)\varphi(t,z_1,z_2|\cdot) + \mathrm{i}z_1\mathbb{E}[e^{\xi}-1]\frac{\partial \varphi}{\partial z_2}(t,z_1,z_2|\cdot).
\end{aligned}
\tag{4.26}
$$

The second second term is computed with two integrations by parts, that is

$$
\begin{aligned}
&\frac{\sigma^2}{2}\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}\frac{\partial^2 p}{\partial x_1^2}(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\\
&= \frac{\sigma^2}{2}\int_{\mathbb{R}^+} e^{\mathrm{i}z_2x_2} z_1\begin{bmatrix} \dfrac{e^{z_1x_1}}{z_1}\dfrac{\partial p}{\partial x_1}(t,x_1,x_2|\cdot) \\ \quad - \displaystyle\int e^{z_1x_1}\frac{\partial p}{\partial x_1}(t,x_1,x_2|\cdot)\mathrm{d}x_1\end{bmatrix}_{x_1=-\infty}^{+\infty}\mathrm{d}x_2\\
&= -\frac{\sigma^2}{2}z_1\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}\frac{\partial p}{\partial x_1}(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\\
&= -\frac{\sigma^2}{2}z_1\int_{\mathbb{R}^+} e^{\mathrm{i}z_2x_2} z_1\begin{bmatrix} \dfrac{e^{z_1x_1}}{z_1}p(t,x_1,x_2|\cdot) \\ \quad - \displaystyle\int e^{z_1x_1}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\end{bmatrix}_{x_1=-\infty}^{+\infty}\mathrm{d}x_2\\
&= \frac{\sigma^2}{2}z_1^2\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\\
&= \frac{\sigma^2}{2}z_1^2\varphi(t,z_1,z_2|\cdot).
\end{aligned}
\tag{4.27}
$$

The third term is

$$
\begin{aligned}
&-\kappa\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}(\theta-x_2)\frac{\partial p}{\partial x_2}(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\\
&=-\kappa\theta\int_{\mathbb{R}} e^{z_1x_1}\mathrm{i}z_2\int_{\mathbb{R}^+}\frac{e^{\mathrm{i}z_2x_2}}{\mathrm{i}z_2}\frac{\partial p}{\partial x_2}(t,x_1,x_2|\cdot)\mathrm{d}x_2\mathrm{d}x_1\\
&-\kappa\mathrm{i}\int_{\mathbb{R}} e^{z_1x_1}\int_{\mathbb{R}^+}\mathrm{i}x_2e^{\mathrm{i}z_2x_2}\frac{\partial p}{\partial x_2}(t,x_1,x_2|\cdot)\mathrm{d}x_2\mathrm{d}x_1\\
&=-\kappa\theta\mathrm{i}z_2\int_{\mathbb{R}} e^{z_1x_1}\left[\begin{matrix}p(t,x_1,x_2|\cdot)\\ \quad -\int e^{\mathrm{i}x_2z_2}p(t,x_1,x_2|\cdot)\mathrm{d}x_2\end{matrix}\right]_{x_2=0}^{+\infty}\mathrm{d}x_1\\
&-\kappa\mathrm{i}\frac{\partial}{\partial z_2}\int_{\mathbb{R}} e^{z_1x_1}\int_{\mathbb{R}^+}e^{\mathrm{i}z_2x_2}\frac{\partial p}{\partial x_2}(t,x_1,x_2|\cdot)\mathrm{d}x_2\mathrm{d}x_1\\
&=\kappa\theta\mathrm{i}z_2\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\\
&-\kappa\mathrm{i}\frac{\partial}{\partial z_2}\int_{\mathbb{R}} e^{z_1x_1}\mathrm{i}z_2\left[\begin{matrix}\frac{e^{\mathrm{i}z_2x_2}}{\mathrm{i}z_2}p(t,x_1,x_2|\cdot)\\ \quad -\int e^{\mathrm{i}z_2x_2}p(t,x_1,x_2|\cdot)\mathrm{d}x_2\end{matrix}\right]_{x_2=0}^{+\infty}\mathrm{d}x_1\\
&=\kappa\theta\mathrm{i}z_2\varphi(t,z_1,z_2|\cdot)-\kappa\frac{\partial}{\partial z_2}\left(z_2\varphi(t,z_1,z_2|\cdot)\right)\\
&=\kappa\varphi(t,z_1,z_2|\cdot)[\theta\mathrm{i}z_2-1]-z_2\kappa\frac{\partial\varphi}{\partial z_2}(t,z_1,z_2|\cdot).
\end{aligned}
\tag{4.28}
$$

The fourth term is simply

$$
\kappa\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2=\kappa\varphi(t,z_1,z_2|\cdot).
\tag{4.29}
$$

The fifth term is computed as

$$
\begin{aligned}
&-\eta\mathbb{E}\left[|\xi|\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}p(t,x_1-\xi,x_2-\eta|\xi||\cdot)\mathrm{d}x_1\mathrm{d}x_2\right]\\
&=-\eta\mathbb{E}\left[|\xi|\int_{-\eta|\xi|}^{+\infty}\int_{\mathbb{R}} e^{z_1(x_1+\xi)+\mathrm{i}z_2(x_2+\eta|\xi|)}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\right]\\
&=-\eta\mathbb{E}\left[|\xi|e^{z_1\xi+\mathrm{i}z_2\eta|\xi|}\int_{\mathbb{R}^+}\int_{\mathbb{R}} e^{z_1x_1+\mathrm{i}z_2x_2}p(t,x_1,x_2|\cdot)\mathrm{d}x_1\mathrm{d}x_2\right]\\
&=-\eta\mathbb{E}\left[|\xi|e^{z_1\xi+\mathrm{i}z_2\eta|\xi|}\right]\varphi(t,z_1,z_2|\cdot).
\end{aligned}
\tag{4.30}
$$

Similarly, the last term is rewritten as

$$\begin{aligned}
&\mathbb{E}\left[\int_{\mathbb{R}+}\int_{\mathbb{R}} x_2 e^{z_1x_1+\mathrm{i}z_2x_2}(p(t,x_1-\xi,x_2-\eta|\xi||\cdot)-p(t,x_1,x_2|\cdot))\mathrm{d}x_1\mathrm{d}x_2\right]\\
&=\varphi(t,z_1,z_2|\cdot)\eta\mathbb{E}\left[|\xi|e^{z_1\xi+\mathrm{i}z_2\eta|\xi|}\right]\\
&\qquad\qquad+\mathrm{i}\left(1-\mathbb{E}\left[e^{z_1\xi+\mathrm{i}z_2\eta|\xi|}\right]\right)\frac{\partial\varphi}{\partial z_2}(t,z_1,z_2|\cdot).
\end{aligned}\tag{4.31}$$

Adding all the terms of Equations (4.26) to (4.31) gives the announced result. □

Define $\varphi^{(0)}$ as

$$\varphi^{(0)}(t,z_1,z_2|s,y_1,y_2):=\mathbb{E}[e^{z_1(X_t-y_1)+\mathrm{i}z_2(\lambda_t-y_2)}|X_s=y_1,\lambda_s=y_2],$$

which is a variant of φ where the initial values of the processes $(X_t)_{t\geqslant 0}$ and $(\lambda_t)_{t\geqslant 0}$ are substracted. In short, we also write $\varphi^{(0)}(t,z_1,z_2|\cdot)=\varphi^{(0)}(t,z_1,z_2|s,y_1,y_2)$. The PDE of Proposition 4.4 easily imply a PDE for $\varphi^{(0)}$.

Corollary 4.2. *The function $\varphi^{(0)}$ satisfies the PDE*

$$\frac{\partial\varphi^{(0)}}{\partial t}(t,z_1,z_2|\cdot)=\gamma(z_1,z_2)\frac{\partial\varphi^{(0)}}{\partial z_2}(t,z_1,z_2|\cdot)+\beta(z_1,z_2)\varphi^{(0)}(t,z_1,z_2|\cdot)$$

where

$$\gamma(z_1,z_2):=\mathrm{i}z_1\mathbb{E}[e^{\xi}-1]-\kappa z_2+\mathrm{i}-\mathrm{i}\mathbb{E}\left[e^{z_1\xi+\mathrm{i}z_2\eta|\xi|}\right]$$

and

$$\begin{aligned}
\beta(z_1,z_2):=&\left(r-\frac{\sigma^2}{2}\right)z_1+\frac{(\sigma z_1)^2}{2}+\mathrm{i}\kappa(\theta-y_2)z_2-z_1y_2\mathbb{E}[e^{\xi}-1]\\
&-y_2+y_2\mathbb{E}\left[e^{z_1\xi+\mathrm{i}z_2\eta|\xi|}\right].
\end{aligned}$$

Moreover the boundary conditions $\varphi(s,z_1,z_2|\cdot)=\varphi(s,0,0|\cdot)=1$ hold.

Proof. It follows from

$$\varphi^{(0)}(t,z_1,z_2|\cdot)=\exp\{-z_1x_1-\mathrm{i}z_2y_2\}\varphi(t,z_1,z_2|\cdot)$$

and Proposition 4.4. □

The interest of Corollary 4.2 is that the PDE of this corollary happened to be slightly easier to solve numerically than the PDE of Proposition 4.4.

We now turn to the derivation of fractional counterparts of the PDE's. In the following, for $s\leqslant t$, p_α will denote either the conditional PDF of (X_{S_t},λ_{S_t}) given that $(X_{S_s},\lambda_{S_s})=(y_1,y_2)$, $S_s=v$ and $U_{S_s}=u$, in which case it is written as

$$\begin{aligned}
&p_\alpha(t,x_1,x_2|s,y_1,y_2,v,u)\\
&:=\frac{\partial^2}{\partial x_1\partial x_2}\mathbb{P}\left(X_{S_t}\leqslant x_1,\lambda_{S_t}\leqslant x_2|X_{S_s}=y_1,\lambda_{S_s}=y_2,S_s=v,U_{S_s}=u\right),
\end{aligned}$$

the conditional PDF of $(S_t, X_{S_t}, \lambda_{S_t})$ given the same information, in which case we denote it by

$$
\begin{aligned}
&p_\alpha(t, x_0, x_1, x_2|s, y_1, y_2, v, u) := \\
&\frac{\partial^3}{\partial x_0 \partial x_1 \partial x_2} \mathbb{P}\left(S_t \leqslant x_0, X_{S_t} \leqslant x_1, \lambda_{S_t} \leqslant x_2 | X_{S_s} = y_1, \lambda_{S_s} = y_2, S_s = v, U_{S_s} = u\right).
\end{aligned}
$$

or finally the conditional PDF of (S_t, X_{S_t}). In this last case, we write

$$
\begin{aligned}
&p_\alpha(t, x_0, x_1|s, y_1, y_2, v, u) \\
&\quad := \frac{\partial^2}{\partial x_0 \partial x_1} \mathbb{P}\left(S_t \leqslant x_0, X_{S_t} \leqslant x_1 | X_{S_s} = y_1, \lambda_{S_s} = y_2, S_s = v, U_{S_s} = u\right).
\end{aligned}
$$

The letter p still denotes the conditional PDF of (X_t, λ_t) knowing (X_s, λ_s). The function g will denote the PDF of S_t given $S_s = v$ and $U_{S_s} = u$, i.e.

$$
g(t, \tau|s, v, u) := \frac{\partial}{\partial \tau} \mathbb{P}(S_t - S_s \leqslant \tau | S_s = v, U_{S_s} = u),
$$

where $t \geqslant s$. It is shown in Hainaut (2021) that the Laplace transform of g with respect to time satisfies

$$
\int_u^{+\infty} e^{-\omega t} g(t, \tau|s, v, u) \mathrm{d}t = \omega^{\alpha-1} e^{-\omega u - \tau \omega^\alpha}
$$

so that

$$
\begin{aligned}
\tilde{g}(\omega, \tau|s, v, u) &:= \int_{\mathbb{R}^+} e^{-\omega t} g(u + t, \tau|s, v, u) \mathrm{d}t \\
&= \omega^{\alpha-1} e^{-\tau \omega^\alpha}
\end{aligned}
\tag{4.32}
$$

for $t \geqslant u$. Note that the conditional PDF's are related by the following identity

$$
p_\alpha(t, x_1, x_2|s, y_1, y_2, v, u) = \int_{\mathbb{R}^+} p(v + \tau, x_1, x_2|v, y_1, y_2) g(t, \tau|s, v, u) \mathrm{d}\tau \tag{4.33}
$$

for $t > u$. This identity follows from

$$
\begin{aligned}
&\mathbb{P}\left(X_{S_t} \leqslant x_1, \lambda_{S_t} \leqslant x_2 | X_{S_s} = y_1, \lambda_{S_s} = y_2, S_s = v, U_{S_s} = u\right) \\
&= \int_{\mathbb{R}^+} \mathbb{P}(X_{v+\tau} \leqslant x_1, \lambda_{v+\tau} \leqslant x_2 | X_{S_s} = y_1, \lambda_{S_s} = y_2, S_s = v, U_{S_s} = u) g(t, \tau|s, v, u) \mathrm{d}\tau \\
&= \int_{\mathbb{R}^+} \mathbb{P}(X_{v+\tau} \leqslant x_1, \lambda_{v+\tau} \leqslant x_2 | X_v = y_1, \lambda_v = y_2) g(t, \tau|s, v, u) \mathrm{d}\tau.
\end{aligned}
$$

The reason for imposing $t > u$ is the fact that if $s \leqslant t \leqslant u = U_{S_s}$, then we have $U_{S_t} = U_{S_s} = u$, that is U_{S_t} is not random anymore.

The next lemma comes as a consequence of the relation between p and p_α at Equation (4.33).

Lemma 4.6. *The Laplace transforms of* p *and* p_α *satisfy*

$$\tilde{p}_\alpha(\omega, x_1, x_2|s, y_1, y_2, v, u) = \omega^{\alpha-1}\tilde{p}(\omega^\alpha, x_1, x_2|v, y_1, y_2)$$

where

$$\tilde{p}_\alpha(\omega, x_1, x_2|s, y_1, y_2, v, u) := \int_{\mathbb{R}^+} e^{-\omega t} p_\alpha(u+t, x_1, x_2|s, y_1, y_2, v, u)\mathrm{d}t$$

and

$$\tilde{p}(\omega, x_1, x_2|v, y_1, y_2) := \int_{\mathbb{R}^+} e^{-t\omega} p(v+t, x_1, x_2|v, y_1, y_2)\mathrm{d}t.$$

Proof. From Equation (4.33) and Fubini's Theorem,

$$\begin{aligned}
\tilde{p}_\alpha(\omega, x_1, x_2|s, y_1, y_2, v, u) &:= \int_{\mathbb{R}^+} e^{-\omega t} p_\alpha(u+t, x_1, x_2|s, y_1, y_2, v, u)\mathrm{d}t \\
&= \int_{\mathbb{R}^+} e^{-\omega t} \int_{\mathbb{R}^+} p(v+\tau, x_1, x_2|v, y_1, y_2) g(u+t, \tau|s, v, u)\mathrm{d}\tau\mathrm{d}t \\
&= \int_{\mathbb{R}^+} p(v+\tau, x_1, x_2|v, y_1, y_2) \int_{\mathbb{R}^+} e^{-\omega t} g(u+t, \tau|s, v, u)\mathrm{d}t\mathrm{d}\tau \\
&= \omega^{\alpha-1} \int_{\mathbb{R}^+} e^{-\tau\omega^\alpha} p(v+\tau, x_1, x_2|v, y_1, y_2)\mathrm{d}\tau \\
&= \omega^{\alpha-1}\tilde{p}(\omega^\alpha, x_1, x_2|v, y_1, y_2),
\end{aligned}$$

as announced. □

In the next lemma, we recall another useful result.

Lemma 4.7. *For any function* $h : \mathbb{R}^+ \to \mathbb{R}$*, it holds that*

$$\widetilde{\frac{\mathrm{d}^\alpha h}{\mathrm{d}t^\alpha}}(\omega) = \omega^\alpha \tilde{h}(\omega) - \omega^{\alpha-1} h(0)$$

where

$$\widetilde{\frac{\mathrm{d}^\alpha h}{\mathrm{d}t^\alpha}}(\omega) := \int_{\mathbb{R}^+} \frac{\mathrm{d}^\alpha h}{\mathrm{d}t^\alpha}(t) e^{-\omega t}\mathrm{d}t \text{ and } \tilde{h}(\omega) := \int_{\mathbb{R}^+} h(t) e^{-\omega t}\mathrm{d}t.$$

The proof can be found in Podlubny (1999). Thanks to these preliminary results about fractional calculus, we can extend our PDE's to the fractional case. The PDF $p_\alpha(t, x_1, x_2|s, y_1, y_2, v, u)$ may be written as $p_\alpha(t, x_1, x_2|\cdot)$ for shorter notations.

Proposition 4.5. *For $t \geqslant u$, the joint conditional PDF p_α of (X_{S_t}, λ_{S_t}) satisfies the following fractional partial differential equation (FPDE)*

$$\begin{aligned}\frac{\partial^\alpha p_\alpha}{\partial t^\alpha}(t, x_1, x_2|\cdot) = &-\left(\mu - \frac{\sigma^2}{2} - x_2\mathbb{E}[e^\xi - 1]\right)\frac{\partial}{\partial x_1}p_\alpha(t, x_1, x_2|\cdot)\\ &+ \frac{\sigma^2}{2}\frac{\partial^2}{\partial x_1^2}p_\alpha(t, x_1, x_2|\cdot) - \kappa(\theta - x_2)\frac{\partial}{\partial x_2}p_\alpha(t, x_1, x_2|\cdot)\\ &+ \kappa p_\alpha(t, x_1, x_2|\cdot) - \eta\mathbb{E}[|\xi|p_\alpha(t, x_1 - \xi, x_2 - \eta|\xi||\cdot)]\\ &+ \mathbb{E}[p_\alpha(t, x_1 - \xi, x_2 - \eta|\xi||\cdot) - p_\alpha(t, x_1, x_2|\cdot)],\end{aligned}$$

with initial condition $p_\alpha(s, x_1, x_2|s, y_1, y_2, v, u) = \delta_{\{x_1 - y_1, x_2 - y_2\}}$.

Proof. Recall that for $v \leqslant t$, $p(t, x_1, x_2|v, y_1, y_2)$ denotes the PDF of (X_t, λ_t) given that $(X_v, \lambda_v) = (y_1, y_2)$ evaluated at $(x_1, x_2) \in \mathbb{R}^2$. The identity

$$\int_{\mathbb{R}^+} \frac{\partial}{\partial t}\left(e^{-\omega t}p(v + t, x_1, x_2|v, y_1, y_2)\right)\mathrm{d}t = -p(v, x_1, x_2|v, y_1, y_2)$$

follows from the fundamental theorem of calculus while an explicit computation of the integral gives

$$\begin{aligned}&\int_{\mathbb{R}^+} \frac{\partial}{\partial t}\left(e^{-\omega t}p(v + t, x_1, x_2|v, y_1, y_2)\right)\mathrm{d}t\\ &\qquad = -\omega\tilde{p}(\omega, x_1, x_2|v, y_1, y_2) + \int_{\mathbb{R}^+} e^{-\omega t}\frac{\partial}{\partial t}p(v + t, x_1, x_2|v, y_1, y_2)\mathrm{d}t\end{aligned}$$

with $\tilde{p}(\omega, x_1, x_2|v, y_1, y_2) := \int_{\mathbb{R}^+} e^{-\omega t}p(v + t, x_1, x_2|v, y_1, y_2)\mathrm{d}t$. As a consequence, we have

$$\begin{aligned}&\omega\tilde{p}(\omega, x_1, x_2|v, y_1, y_2) - p(v, x_1, x_2|v, y_1, y_2)\\ &\qquad = \int_{\mathbb{R}^+} e^{-\omega t}\frac{\partial}{\partial t}p(v + t, x_1, x_2|v, y_1, y_2)\mathrm{d}t.\end{aligned} \tag{4.34}$$

From the PDE of p (see Proposition 4.3), the right-hand side of Equation (4.34) becomes

$$
\begin{aligned}
&-\int_{\mathbb{R}+} e^{-\omega t}\left(\mu-\frac{\sigma^2}{2}-x_2\mathbb{E}[e^{\xi}-1]\right)\frac{\partial}{\partial x_1}p(v+t,x_1,x_2|\cdot)\mathrm{d}t\\
&+\frac{\sigma^2}{2}\int_{\mathbb{R}+} e^{-\omega t}\frac{\partial^2}{\partial x_1^2}p(v+t,x_1,x_2|\cdot)\mathrm{d}t-\kappa(\theta-x_2)\int_{\mathbb{R}+} e^{-\omega t}\frac{\partial}{\partial x_2}p(v+t,x_1,x_2|\cdot)\mathrm{d}t\\
&+\kappa\int_{\mathbb{R}+} e^{-\omega t}p(v+t,x_1,x_2|\cdot)\mathrm{d}t-\eta\mathbb{E}\left[|\xi|\int_{\mathbb{R}+} e^{-\omega t}p(v+t,x_1-\xi,x_2-\eta|\xi||\cdot)\mathrm{d}t\right]\\
&+\mathbb{E}\left[\int_{\mathbb{R}+} e^{-\omega t}[p(v+t,x_1-\xi,x_2-\eta|\xi||\cdot)-p(v+t,x_1,x_2|\cdot)]\mathrm{d}t\right]\\
&=-\left(\mu-\frac{\sigma^2}{2}-x_2\mathbb{E}[e^{\xi}-1]\right)\frac{\partial}{\partial x_1}\tilde{p}(\omega,x_1,x_2|\cdot)+\frac{\sigma^2}{2}\frac{\partial^2}{\partial x_1^2}\tilde{p}(\omega,x_1,x_2|\cdot)\\
&-\kappa(\theta-x_2)\frac{\partial}{\partial x_2}\tilde{p}(\omega,x_1,x_2|\cdot)+\kappa\tilde{p}(\omega,x_1,x_2|\cdot)\\
&-\eta\mathbb{E}[|\xi|\tilde{p}(\omega,x_1-\xi,x_2-\eta|\xi||\cdot)]+\mathbb{E}[\tilde{p}(\omega,x_1-\xi,x_2-\eta|\xi||\cdot)-\tilde{p}(\omega,x_1,x_2|\cdot)],
\end{aligned}
$$

where $p(v+t,x_1,x_2|\cdot)$ denotes $p(v+t,x_1,x_2|v,y_1,y_2)$ and $\tilde{p}(\omega,x_1,x_2|\cdot)$ denotes $\tilde{p}(\omega,x_1,x_2|v,y_1,y_2)$. Since this is valid for any ω, we can replace ω by ω^α. After multiplying both sides by $\omega^{\alpha-1}$, it gives

$$
\begin{aligned}
&\omega^\alpha(\omega^{\alpha-1}\tilde{p}(\omega^\alpha,x_1,x_2|\cdot))-\omega^{\alpha-1}p(v,x_1,x_2|\cdot)\\
&=-\left(\mu-\frac{\sigma^2}{2}-x_2\mathbb{E}[e^{\xi}-1]\right)\frac{\partial}{\partial x_1}(\omega^{\alpha-1}\tilde{p}(\omega^\alpha,x_1,x_2|\cdot))+\frac{\sigma^2}{2}\frac{\partial^2}{\partial x_1^2}(\omega^{\alpha-1}\tilde{p}(\omega^\alpha,x_1,x_2|\cdot))\\
&-\kappa(\theta-x_2)\frac{\partial}{\partial x_2}(\omega^{\alpha-1}\tilde{p}(\omega^\alpha,x_1,x_2|\cdot))+\kappa(\omega^{\alpha-1}\tilde{p}(\omega^\alpha,x_1,x_2|\cdot))\\
&-\eta\mathbb{E}[|\xi|(\omega^{\alpha-1}\tilde{p}(\omega^\alpha,x_1-\xi,x_2-\eta|\xi||\cdot))]\\
&+\mathbb{E}[(\omega^{\alpha-1}\tilde{p}(\omega^\alpha,x_1-\xi,x_2-\eta|\xi||\cdot))-(\omega^{\alpha-1}\tilde{p}(\omega^\alpha,x_1,x_2|\cdot))]
\end{aligned}
$$

Using Lemma 4.6 and the fact that $p(v,x_1,x_2|v,y_1,y_2)=p_\alpha(s,x_1,x_2|s,y_1,y_2,v,u)=\delta_{\{x_1-y_1,x_2-y_2\}}$, we get

$$
\begin{aligned}
&\omega^\alpha\tilde{p}_\alpha(\omega,x_1,x_2|s,y_1,y_2,v,u)-\omega^{\alpha-1}p_\alpha(s,x_1,x_2|s,y_1,y_2,v,u)\\
&\quad=\widetilde{\frac{\partial^\alpha p_\alpha}{\partial t^\alpha}}(\omega,x_1,x_2|s,y_1,y_2,v,u)\\
&\quad=-\left(\mu-\frac{\sigma^2}{2}-x_2\mathbb{E}[e^{\xi}-1]\right)\frac{\partial}{\partial x_1}\tilde{p}_\alpha(\omega,x_1,x_2|s,y_1,y_2,v,u)\\
&\quad+\frac{\sigma^2}{2}\frac{\partial^2}{\partial x_1^2}\tilde{p}_\alpha(\omega,x_1,x_2|s,y_1,y_2,v,u)\\
&\quad-\kappa(\theta-x_2)\frac{\partial}{\partial x_2}\tilde{p}_\alpha(\omega,x_1,x_2|s,y_1,y_2,v,u)+\kappa\tilde{p}_\alpha(\omega,x_1,x_2|s,y_1,y_2,v,u)\\
&\quad-\eta\mathbb{E}[|\xi|\tilde{p}_\alpha(\omega,x_1-\xi,x_2-\eta|\xi||s,y_1,y_2,v,u)]\\
&\quad+\mathbb{E}[\tilde{p}_\alpha(\omega,x_1-\xi,x_2-\eta|\xi||s,y_1,y_2,v,u)-\tilde{p}_\alpha(\omega,x_1,x_2|s,y_1,y_2,v,u)],
\end{aligned}
$$

and the result follows. □

From the definition of conditional PDF, we have

$$p_\alpha(t, x_0, x_1, x_2|s, y_1, y_2, v, u) = g(t, x_0 - v|s, v, u)p(x_0, x_1, x_2|v, y_1, y_2).$$

The Laplace transform of p_α therefore satisfies

$$\begin{aligned}
&\tilde{p}_\alpha(\omega, x_0, x_1, x_2|s, y_1, y_2, v, u)\\
&\quad := \int_{\mathbb{R}^+} e^{-\omega t} p_\alpha(u+t, x_0, x_1, x_2|s, y_1, y_2, v, u)\mathrm{d}t\\
&\quad = p(x_0, x_1, x_2|s, y_1, y_2, v, u)\int_{\mathbb{R}^+} e^{-\omega t} g(u+t, x_0 - v|s, v, u)\mathrm{d}t\\
&\quad = p(x_0, x_1, x_2|s, y_1, y_2, v, u)\tilde{g}(\omega, x_0 - v|s, v, u)\\
&\quad = \omega^{\alpha-1} p(x_0, x_1, x_2|v, y_1, y_2)e^{-(x_0-v)\omega^\alpha},
\end{aligned} \tag{4.35}$$

where the last equality follows from Equation (4.32). For $z_0, z_2 \in \mathbb{R}$ and $z_1 \in \mathbb{C}$, define the transform φ_α as

$$\begin{aligned}
&\varphi_\alpha(t, z_0, z_1, z_2|s, y_1, y_2, v, u)\\
&\quad := \mathbb{E}[e^{-z_0 S_t + z_1 X_{S_t} + \mathrm{i} z_2 \lambda_{S_t}}|X_{S_s} = y_1, \lambda_{S_s} = y_2, S_s = v, U_{S_s} = u],
\end{aligned}$$

with $t \geqslant u$. Then the function φ_α satisfies a fractional PDE.

Proposition 4.6. *The following FPDE holds*

$$\begin{aligned}
&\frac{\partial^\alpha \varphi_\alpha}{\partial t^\alpha}(t, z_0, z_1, z_2|s, y_1, y_2, v, u)\\
&\quad = \left(\mathrm{i} z_1 \mathbb{E}[e^\xi - 1] - \kappa z_2 + \mathrm{i} - \mathbb{E}\left[e^{z_1\xi + \mathrm{i} z_2 \eta|\xi|}\right]\right)\frac{\partial \varphi_\alpha}{\partial z_2}(t, z_0, z_1, z_2|s, y_1, y_2, v, u)\\
&\quad + \left(\left(\mu - \frac{\sigma^2}{2}\right)z_1 + \frac{(\sigma z_1)^2}{2} + \mathrm{i}\kappa\theta z_2 - z_0\right)\varphi_\alpha(t, z_0, z_1, z_2|s, y_1, y_2, v, u)
\end{aligned}$$

with boundary conditions

$$\varphi_\alpha(t, z_0, 0, 0|s, y_1, y_2, v, u) = \mathrm{E}_\alpha(-z_0(t-u)^\alpha)$$

$$\varphi_\alpha(v, z_0, z_1, z_2|s, y_1, y_2, v, u) = e^{-z_0 v + z_1 y_1 + \mathrm{i} z_2 y_2}$$

and

$$\varphi_\alpha(t, 0, 0, 0|s, y_1, y_2, v, u) = 1.$$

Proof. Note that the time-Laplace transform of $\varphi^{\alpha}_{X,\lambda}$ satisfies

$$\begin{aligned}
&\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega, z_0, z_1, z_2|s, y_1, y_2, v, u) := \int_{\mathbb{R}+} e^{-\omega t}\varphi_{\alpha}(u+t, z_0, z_1, z_2|s, y_1, y_2, v, u)\mathrm{d}t \\
&= \int_{\mathbb{R}+}\int_{\mathbb{R}+}\int_{\mathbb{R}}\int_{\mathbb{R}+} e^{-\omega t - z_0 x_0 + z_1 x_1 + \mathrm{i} z_2 x_2} p_{\alpha}(u+t, x_0, x_1, x_2|s, y_1, y_2, v, u)\mathrm{d}x_0\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}t \\
&= \int_{\mathbb{R}+}\int_{\mathbb{R}}\int_{v}^{+\infty} e^{-z_0 x_0 + z_1 x_1 + \mathrm{i} z_2 x_2}\tilde{p}_{\alpha}(\omega, x_0, x_1, x_2|s, y_1, y_2, v, u)\mathrm{d}x_0\mathrm{d}x_1\mathrm{d}x_2 \\
&= \omega^{\alpha-1}e^{v\omega^{\alpha}}\int_{\mathbb{R}+}\int_{\mathbb{R}}\int_{v}^{+\infty} e^{-x_0[z_0+\omega^{\alpha}] + z_1 x_1 + \mathrm{i} z_2 x_2} p(x_0, x_1, x_2|v, y_1, y_2)\mathrm{d}x_0\mathrm{d}x_1\mathrm{d}x_2 \\
&= \omega^{\alpha-1}e^{v\omega^{\alpha}}\int_{v}^{+\infty} e^{-x_0[z_0+\omega^{\alpha}]}\varphi(x_0, z_1, z_2|v, y_1, y_2)\mathrm{d}x_0,
\end{aligned} \tag{4.36}$$

where the penultimate equality comes as a consequence of Equation (4.35). Next, an integration by parts yields

$$\begin{aligned}
&\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega, z_0, z_1, z_2|s, y_1, y_2, v, u) \\
&\quad = \frac{-\omega^{\alpha-1}e^{v\omega^{\alpha}}}{z_0+\omega^{\alpha}}\Bigg[\varphi(x_0, z_1, z_2|v, y_1, y_2)e^{-[z_0+\omega^{\alpha}]x_0} \\
&\qquad\qquad - \int e^{-[z_0+\omega^{\alpha}]x_0}\frac{\varphi}{\partial x_0}(x_0, z_1, z_2|v, y_1, y_2)\mathrm{d}x_0\Bigg]_{x_0=v}^{+\infty} \qquad (4.37) \\
&\quad = \frac{\omega^{\alpha-1}e^{-z_0 v}}{z_0+\omega^{\alpha}}\varphi(v, z_1, z_2|v, y_1, y_2) \\
&\qquad + \frac{\omega^{\alpha-1}e^{v\omega^{\alpha}}}{z_0+\omega^{\alpha}}\int_{v}^{+\infty} e^{-[z_0+\omega^{\alpha}]t}\frac{\partial\varphi}{\partial t}(t, z_1, z_2|v, y_1, y_2)\mathrm{d}t.
\end{aligned}$$

From Proposition 4.4, we know that

$$\begin{aligned}
&\omega^{\alpha-1}e^{v\omega^{\alpha}}\int_{v}^{+\infty} e^{-[z_0+\omega^{\alpha}]t}\frac{\partial\varphi}{\partial t}(t,z_1,z_2|v,y_1,y_2)\mathrm{d}t\\
&=\omega^{\alpha-1}e^{v\omega^{\alpha}}\left(\mathrm{i}z_1\mathbb{E}[e^{\xi}-1]-\kappa z_2\right)\int_{v}^{+\infty} e^{-[z_0+\omega^{\alpha}]t}\frac{\partial\varphi}{\partial z_2}(t,z_1,z_2|v,y_1,y_2)\mathrm{d}t\\
&+\frac{\left(\mu-\frac{\sigma^2}{2}\right)z_1+\frac{(\sigma z_1)^2}{2}+\mathrm{i}\kappa\theta z_2+\mathbb{E}\left[(1-\eta|\xi|)e^{z_1\xi+\mathrm{i}z_2|\xi|}\right]-1}{\omega^{1-\alpha}e^{-v\omega^{\alpha}}}\\
&\qquad\qquad\int_{v}^{+\infty} e^{-[z_0+\omega^{\alpha}]t}\varphi(t,z_1,z_2|v,y_1,y_2)\mathrm{d}t\\
&=\left(\mathrm{i}z_1\mathbb{E}[e^{\xi}-1]-\kappa z_2\right)\frac{\partial}{\partial z_2}\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u)\\
&+\left(\left(\mu-\frac{\sigma^2}{2}\right)z_1+\frac{(\sigma z_1)^2}{2}+\mathrm{i}\kappa\theta z_2+\mathbb{E}\left[(1-\eta|\xi|)e^{z_1\xi+\mathrm{i}z_2|\xi|}\right]-1\right)\\
&\qquad\qquad\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u)
\end{aligned}$$

where the second equality is a consequence of Equation (4.36). Equation (4.37) is then rewritten as

$$\begin{aligned}
&(z_0+\omega^{\alpha})\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u)-\omega^{\alpha-1}e^{-z_0v}\varphi(v,z_1,z_2|v,y_1,y_2)\\
&=\left(\mathrm{i}z_1\mathbb{E}[e^{\xi}-1]-\kappa z_2\right)\frac{\partial}{\partial z_2}\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u)\\
&+\left(\left(\mu-\frac{\sigma^2}{2}\right)z_1+\frac{(\sigma z_1)^2}{2}+\mathrm{i}\kappa\theta z_2+\mathbb{E}\left[(1-\eta|\xi|)e^{z_1\xi+\mathrm{i}z_2|\xi|}\right]-1\right)\\
&\qquad\qquad\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u).
\end{aligned}$$

By noting that

$$\varphi_{\alpha}(u,z_0,z_1,z_2|s,y_1,y_2,v,u)=e^{-z_0v+z_1y_1+\mathrm{i}z_2y_2}=e^{-z_0v}\varphi(v,z_1,z_2|v,y_1,y_2),$$

it follows that the last equation becomes

$$\begin{aligned}
&\omega^{\alpha}\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u)-\omega^{\alpha-1}\varphi_{\alpha}(u,z_0,z_1,z_2|s,y_1,y_2,v,u)\\
&=-z_0\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u)\\
&+\left(\mathrm{i}z_1\mathbb{E}[e^{\xi}-1]-\kappa z_2\right)\frac{\partial}{\partial z_2}\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u)\\
&+\left(\left(\mu-\frac{\sigma^2}{2}\right)z_1+\frac{(\sigma z_1)^2}{2}+\mathrm{i}\kappa\theta z_2+\mathbb{E}\left[(1-\eta|\xi|)e^{z_1\xi+\mathrm{i}z_2|\xi|}\right]-1\right)\\
&\qquad\qquad\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u).
\end{aligned}$$

Since Lemma 4.7 implies that

$$\begin{aligned}
\widetilde{\frac{\partial\varphi_{\alpha}}{\partial t^{\alpha}}}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u)=&\,\omega^{\alpha}\tilde{\varphi}^{\alpha}_{X,\lambda}(\omega,z_0,z_1,z_2|s,y_1,y_2,v,u)\\
&-\omega^{\alpha-1}\varphi_{\alpha}(0,z_0,z_1,z_2|s,y_1,y_2,v,u),
\end{aligned}$$

the conclusion follows. □

As in the non-fractional case, we derive an equation for the variant $\varphi_\alpha^{(0)}$ of φ_α where the initial values of the processes are substracted. More precisely, if

$$\begin{aligned}
&\varphi_\alpha^{(0)}(t, z_0, z_1, z_2|s, y_1, y_2, v, u) \\
&\quad := \mathbb{E}[e^{-z_0 S_t + z_1(X_{S_t} - y_1) + \mathrm{i} z_2(\lambda_{S_t} - y_2)} | X_{S_s} = y_1, \lambda_{S_s} = y_2, S_s = v, U_{S_s} = u],
\end{aligned}$$

then $\varphi_\alpha^{(0)}$ satisfies a FPDE.

Corollary 4.3. *The function $\varphi_\alpha^{(0)}$ satisfies the FPDE*

$$\begin{aligned}
\frac{\partial^\alpha \varphi_\alpha^{(0)}}{\partial t^\alpha}(t, z_1, z_2|s, y_1, y_2, v, u) =& \gamma(z_1, z_2) \frac{\partial \varphi_\alpha^{(0)}}{\partial z_2}(t, z_0, z_1, z_2|s, y_1, y_2, v, u) \\
&+ \beta(z_0, z_1, z_2) \varphi_\alpha^{(0)}(t, z_0, z_1, z_2|s, y_1, y_2, v, u)
\end{aligned}$$

where

$$\gamma(z_1, z_2) := \mathrm{i} z_1 \mathbb{E}[e^\xi - 1] - \kappa z_2 + \mathrm{i} - \mathrm{i}\mathbb{E}\left[e^{z_1 \xi + \mathrm{i} z_2 \eta |\xi|}\right]$$

and

$$\begin{aligned}
\beta(z_0, z_1, z_2) :=& \left(r - \frac{\sigma^2}{2}\right) z_1 + \frac{(\sigma z_1)^2}{2} + \mathrm{i}\kappa(\theta - y_2) z_2 \\
&- z_1 y_2 \mathbb{E}[e^\xi - 1] - y_2 + y_2 \mathbb{E}\left[e^{z_1 \xi + \mathrm{i} z_2 \eta |\xi|}\right] - z_0.
\end{aligned}$$

Moreover the boundary conditions

$$\varphi_\alpha^{(0)}(t, z_0, 0, 0|s, y_1, y_2, v, u) = \mathrm{E}_\alpha(-z_0 (t-u)^\alpha),$$

$$\varphi_\alpha^{(0)}(v, z_0, z_1, z_2|s, y_1, y_2, v, u) = e^{-z_0 v}$$

and

$$\varphi_\alpha^{(0)}(t, 0, 0, 0|s, y_1, y_2, v, u) = 1$$

hold.

The proof uses Corollary 4.2 and is essentially the same as Proposition 4.6. It is therefore omitted.

4.4 Changes of measure

In this section, we extend the results of the changes of measure section of Moraux and Hainaut (2018) to the subdiffusive model. To this end, we use a result of Jacod (1979) that states that under a condition called *adaptation to a time change*, sometimes also referred to as *synchronization with a time-change*, a time-changed local-martingale remains a local-martingale. We begin this section by stating this result precisely. Then, we show that it applies to our setting and thereby obtain risk-neutral measures for the time-changed model.

For a probability measure $\mathbb{Q}$ and a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0}$ on $(\Omega, \mathcal{F})$, we define $\mathcal{M}(\mathbb{Q}, \mathbb{F})$ and $\mathcal{M}_{\text{loc}}(\mathbb{Q}, \mathbb{F})$ to be respectively the set of all $(\mathbb{Q}, \mathbb{F})$-martingales and the set of all $(\mathbb{Q}, \mathbb{F})$-local martingales. Of course, it holds that $\mathcal{M}(\mathbb{Q}, \mathbb{F}) \subset \mathcal{M}_{\text{loc}}(\mathbb{Q}, \mathbb{F})$. A subscript 0 is added to these notations when the collection is restricted to the (local-) martingales $(M_t)_{t\geqslant 0}$ that start at 0, i.e. $M_0 = 0$ $\mathbb{Q}$-a.s.. A superscript c is added when the collection is restricted to the (local-) martingales that have continuous paths $\mathbb{Q}$-a.s.. For example, the set $\mathcal{M}_0^{\text{c}}(\mathbb{Q}, \mathbb{F})$ (resp. $\mathcal{M}_{0,\text{loc}}(\mathbb{Q}, \mathbb{F})$) thus contains the continuous martingales which start at 0 (resp. the local-martingales which start at 0 but are not necessarily continuous). We give the definition of two orthogonal local-martingales.

Definition 4.1. Let $(M_t)_{t\geqslant 0}$ and $(H_t)_{t\geqslant 0}$ be two processes in $\mathcal{M}_{\text{loc}}(\mathbb{Q}, \mathbb{F})$. We say that the local-martingales $(M_t)_{t\geqslant 0}$ and $(H_t)_{t\geqslant 0}$ are *orthogonal* if the process $(M_t H_t)_{t\geqslant 0}$ belongs to $\mathcal{M}_{0,\text{loc}}(\mathbb{Q}, \mathbb{F})$. This is written as $(M_t)_{t\geqslant 0} \perp (H_t)_{t\geqslant 0}$.

This definition can be found in e.g. Jacod (1979) (Definition 2.10). It allows us to further define the collection

$$\begin{aligned}&\mathcal{M}_{\text{loc}}^{\text{d}}(\mathbb{Q}, \mathbb{F})\\ &:= \{(M_t)_{t\geqslant 0} \in \mathcal{M}_{\text{loc}}(\mathbb{Q}, \mathbb{F}) : (M_t)_{t\geqslant 0} \perp (H_t)_{t\geqslant 0} \text{ for all } (H_t)_{t\geqslant 0} \in \mathcal{M}_{0,\text{loc}}^{\text{c}}(\mathbb{Q}, \mathbb{F})\}.\end{aligned}$$

For convenience, we will write $\mathbb{F}^U$, $\mathbb{F}^A$ and $\mathbb{F}^{A\circ S}$ instead of $(\mathcal{F}_t^U)_{t\geqslant 0}$, $(\mathcal{F}_t^A)_{t\geqslant 0}$ and $(\mathcal{F}_t^{A\circ S})_{t\geqslant 0}$, see the beginning of Section 4.3 for the definitions of these filtrations. In particular, the collection $\mathcal{M}_{\text{loc}}^{\text{d}}(\mathbb{Q}, \mathbb{F})$ is known to contain the compensated jump processes, as shown in the next proposition.

Proposition 4.7. *Let $g : \mathbb{R} \to \mathbb{R}$ be a function that satisfies $\int_{\mathbb{R}} |g(z)| \nu(\mathrm{d}z) < +\infty$. Recall that $\tilde{\Xi}(\mathrm{d}z, \mathrm{d}t)$ denotes the compensated random measure*

$$\tilde{\Xi}(\mathrm{d}z, \mathrm{d}t) = \Xi(\mathrm{d}z, \mathrm{d}t) - \upsilon(\mathrm{d}z, \mathrm{d}t) = \Xi(\mathrm{d}z, \mathrm{d}t) - \nu(\mathrm{d}z)\lambda_{t-}\mathrm{d}t.$$

The compensated jump process $(\tilde{J}_t^g)_{t\geqslant 0}$ defined as

$$\tilde{J}_t^g := \int_0^t \int_{z\in\mathbb{R}} g(z)\, \tilde{\Xi}(\mathrm{d}z, \mathrm{d}s)$$

is in $\mathcal{M}_{\text{loc}}^{\text{d}}(\mathbb{P}, \mathbb{F}^A)$.

Proof. The compensated jump process can also be expressed as

$$\tilde{J}_t^g = \sum_{k=1}^{N_t} g(\xi_k) - \mathbb{E}[g(\xi_1)]\Lambda_t,$$

where $\xi_1, \xi_2, \ldots$ are the i.i.d. jump sizes of $(J_t)_{t\geqslant 0}$ distributed according to ν and $\Lambda_t = \int_0^t \lambda_{s-}\mathrm{d}s$. Let $(H_t)_{t\geqslant 0} \in \mathcal{M}^{\mathrm{c}}_{0,\mathrm{loc}}(\mathbb{P}, \mathbb{F}^A)$. We have to show that $(\tilde{J}^g_t H_t)_{t\geqslant 0}$ belongs to $\mathcal{M}_{0,\mathrm{loc}}(\mathbb{P}, \mathbb{F}^A)$. From the definition of quadratic covariation,

$$\tilde{J}^g_t H_t = \int_0^t H_{s-}\mathrm{d}\tilde{J}^g_s + \int_0^t \tilde{J}^g_{s-}\mathrm{d}H_s + [\tilde{J}^g, H]_t.$$

Since the compensated process $(\tilde{J}^g_t)_{t\geqslant 0}$ and $(H_t)_{t\geqslant 0}$ are in $\mathcal{M}_{\mathrm{loc}}(\mathbb{P}, \mathbb{F}^A)$, it is enough to show that $([\tilde{J}^g, H]_t)_{t\geqslant 0}$ is indistinguishable from the null process. Since $(\sum_{k=1}^{N_t} g(\xi_k))_{t\geqslant 0}$ is quadratic pure jump,

$$[\tilde{J}^g, H]_t = \sum_{0<s\leqslant t} g(\Delta J_s)\Delta H_s - \mathbb{E}[g(\xi_1)][\Lambda, H]_t$$

but the continuity of $(H_t)_{t\geqslant 0}$ implies that $\sum_{0<s\leqslant t} g(\Delta J_s)\Delta H_s = 0$. Moreover, if we denote by $(E_t)_{t\geqslant 0}$ the time process $(t)_{t\geqslant 0}$, then it is a continuous quadratic pure jump semimartingale[4]. As a consequence $([E, H]_t)_{t\geqslant 0}$ is indistinguishable from the null process. Then, Theorem 29 in Protter (2005) (Chapter 2) yields

$$[\Lambda, H]_t = [\lambda \cdot E, H]_t = \int_0^t \lambda_{s-}\mathrm{d}[E, H]_s = 0,$$

where $((\lambda \cdot E)_t)_{t\geqslant 0}$ denotes the stochastic integral $(\int_0^t \lambda_{s-}\mathrm{d}E_s) = (\Lambda_t)_{t\geqslant 0}$. □

The second important notion is the one of *adaptation to a time-change*.

Definition 4.2. An $\mathbb{F}$-time-change $(S_t)_{t\geqslant 0}$ is a nondecreasing process such that for any $t \geqslant 0$, S_t is an $\mathbb{F}$-stopping-time. Moreover, a stochastic process $(X_t)_{t\geqslant 0}$ is *adapted to the time-change* $(S_t)_{t\geqslant 0}$ if $(X_t)_{t\geqslant 0}$ is constant on any interval $[S_{t-}, S_t]$ for all $t \geqslant 0$, $\mathbb{Q}$-a.s..

It is clear that if $(S_t)_{t\geqslant 0}$ is as described in the previous sections, it is an $(\mathbb{F}^U \vee \mathbb{F}^A)$-time-change. Moreover, its continuity entails that any process is adapted to it, as $S_{t-} = S_t$ for all t. Adaptation to time-changes is discussed in details in the Chapter 10 of Jacod (1979). It is also referred to as *synchronization with a time-change*, as e.g. in Kobayashi (2011).

Recall that a *random set* is a subset of $\Omega \times \mathbb{R}^+$. A particular case of random set is given by *random intervals*. Given two random variables $S, T : \Omega \to \mathbb{R}^+ \cup \{+\infty\}$, the random interval $[\![S, T]\!]$ is the random set defined as

$$[\![S, T]\!] := \{(\omega, t) \in \Omega \times \mathbb{R}^+ : S(\omega) \leqslant t \leqslant T(\omega)\}.$$

[4]Because it is càdlàg and has finite variation on compacts, see Theorem 26 in the chapter 2 of Protter (2005).

To an $\mathbb{F}$-time-change $(S_t)_{t\geqslant 0}$, we can associate the random set J_S as

$$J_S = \{(\omega, t) \in \Omega \times \mathbb{R}^+ : S_{t-}(\omega) < +\infty\}.$$

Let $\mathcal{P}$ be a class of stochastic processes and define

$$\mathcal{P}^{J_S} := \{(X_t)_{t\geqslant 0} : (X_{t\wedge T})_{t\geqslant 0} \in \mathcal{P} \text{ for any } \mathbb{F}\text{-stopping-time } T \text{ that satisfies } [\![0, T]\!] \subset J_S\}.$$

With all these notations introduced, we can state the result that is of interest to us, namely that under some conditions, a time-changed local-martingale remains a local-martingale. This result is contained[5] in Theorem 10.16 in Jacod (1979).

Lemma 4.8. *Let $(M_t)_{t\geqslant 0}$ be a process that belongs to $\mathcal{M}^{\mathrm{c}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{F})$ (resp. a process that belongs to $\mathcal{M}^{\mathrm{d}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{F})$) and that is adapted to an $\mathbb{F}$-time-change $(S_t)_{t\geqslant 0}$. Then the time-changed process $(M_{S_t})_{t\geqslant 0}$ belongs to $\mathcal{M}^{\mathrm{c}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{G})^{J_S}$ (resp. belongs to $\mathcal{M}^{\mathrm{d}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{G})^{J_S}$), where $\mathbb{G}$ denotes the time-changed filtration $(\mathcal{F}_{S_t})_{t\geqslant 0}$.*

If the time-change $(S_t)_{t\geqslant 0}$ is the inverse of an α-stable subordinator, Proposition 4.1 implies that $\mathbb{P}(\{S_t < +\infty \text{ for all } t \geqslant 0\}) = 1$ so that by ignoring a null set, we can assume that $\{S_t < +\infty \text{ for all } t \geqslant 0\} = \Omega$. It follows that the associated random set J_S is simply $\Omega \times \mathbb{R}^+$, and thus that $[\![0, +\infty]\!] = \Omega \times \mathbb{R}^+ \subset J_S$. As a consequence, the inclusions $(M_{S_t})_{t\geqslant 0} \in \mathcal{M}^{\mathrm{c}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{G})^{J_S}$ (resp. $(M_{S_t})_{t\geqslant 0} \in \mathcal{M}^{\mathrm{d}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{G})^{J_S}$) in particular means that $(M_{S_t})_{t\geqslant 0} \in \mathcal{M}^{\mathrm{c}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{G})$ (resp. $(M_{S_t})_{t\geqslant 0} \in \mathcal{M}^{\mathrm{d}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{G})$). Finally, note that since the inverse of an α-stable process has continuous paths a.s., we can drop the assumption of a process adapted to the time-change in Lemma 4.8. This is summarized in the next corollary.

Corollary 4.4. *Let $(M_t)_{t\geqslant 0}$ be a process that belongs to $\mathcal{M}^{\mathrm{c}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{F})$ (resp. a process that belongs to $\mathcal{M}^{\mathrm{d}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{F})$) and an $\mathbb{F}$-time-change $(S_t)_{t\geqslant 0}$ that is the inverse of an α-stable subordinator Then the time-changed process $(M_{S_t})_{t\geqslant 0}$ belongs to $\mathcal{M}^{\mathrm{c}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{G})$ (resp. belongs to $\mathcal{M}^{\mathrm{d}}_{\mathrm{loc}}(\mathbb{Q}, \mathbb{G})$), where $\mathbb{G}$ denotes the time-changed filtration $(\mathcal{F}_{S_t})_{t\geqslant 0}$.*

We will also use use a change of variable formula for time-changed stochastic integral. This formula is stated at Proposition 10.21 of Jacod (1979). Let $(Z_t)_{t\geqslant 0}$ and $(H_t)_{t\geqslant 0}$ be two a semimartingales, with $(Z_t)_{t\geqslant 0}$ in synchronization with a time-change $(S_t)_{t\geqslant 0}$ and $(H_t)_{t\geqslant 0}$ predictable. If the stochastic integral $\int_0^t H_s \mathrm{d}Z_s$ exists, then the integral $\int_0^{S_t} H_s \mathrm{d}Z_s$ also exists and satisfies the change of variable formula

$$\int_0^{S_t} H_s \mathrm{d}Z_s = \int_0^t H_{S_{s-}} \mathrm{d}Z_{S_s}. \tag{4.38}$$

[5] This theorem also contains the same result for other class of processes such as semimartingales for example, as noted at the beginning of Section 4.3, but we are only interested with local-martingales here.

This formula can also be found in Kobayashi (2011) (Lemma 2.3). Again, since the inverse of an α-stable subordinator is continuous, we can drop the assumption of synchronization with the time-change.

In the following, we focus on a family of changes of measure induced by exponential martingales $(M_t(x,\zeta))_{t\geqslant 0}$ that satisfy

$$M_t(x,\zeta) = \exp\left\{b_1\lambda_t + xL_t + b_2 t - \frac{1}{2}\int_0^t \zeta_s^2 \mathrm{d}s - \int_0^t \zeta_s \mathrm{d}W_s\right\}, \tag{4.39}$$

where $(\zeta_t)_{t\geqslant 0}$ is a predictable adapted process of the form $\zeta_t = \zeta_0 + \zeta_1\lambda_{t-}$ and ζ_0, ζ_1 and x are constants. The constants b_1, b_2 correspond to the price of jump risk. To see this, note first that if we remove the jumps from our model, we obtain the Black and Scholes model and thus a complete market. As a consequence, the incompleteness of the market in our model, and thus the existence of an infinity of risk-neutral measures, is due to the presence of jumps in the asset price. The choice of the risk-neutral measure[6] is therefore the choice of how to price the jump risk.

The next proposition gives sufficient conditions under which $(M_t(x,\zeta))_{t\geqslant 0}$ is in $\mathcal{M}_{\mathrm{loc}}(\mathbb{P}, \mathbb{F}^{A,U})$.

Proposition 4.8. *If the relations*

$$\kappa b_1 - (\psi(0, \eta b_1 + x) - 1) = 0$$

and

$$b_2 + \kappa\theta b_1 = 0$$

hold, then the process $(M_t(x,\zeta))_{t\geqslant 0}$ is in $\mathcal{M}_{\mathrm{loc}}(\mathbb{P}, \mathbb{F}^{A,U})$.

Proof. Proposition 4.1 in Moraux and Hainaut (2018) states that the two conditions of the statement are sufficient for $(M_t(x,\zeta))_{t\geqslant 0}$ to be in $\mathcal{M}_{\mathrm{loc}}(\mathbb{P}, \mathbb{F}^{A})$. It remains to be shown that it is also in $\mathcal{M}_{\mathrm{loc}}(\mathbb{P}, \mathbb{F}^{A,U})$. By definition of local-martingale, there exists a sequence $(T_n)_{n\in\mathbb{N}}$ of $\mathbb{F}^A$-stopping-times that satisfies $T_n \uparrow +\infty$ $\mathbb{P}$-a.s. and such that the stopped process $(M_{t\wedge T_n}(x,\zeta))_{t\geqslant 0}$ is a $(\mathbb{P}, \mathbb{F}^A)$-martingale for any $n \in \mathbb{N}$. Recall that the filtrations $\mathbb{F}^A$ and $\mathbb{F}^U$ are independent. We will use the following property: for an integrable random variable X and two σ-algebras $\mathcal{H}$ and $\mathcal{G}$, if $\mathcal{H}$ is independent of $\sigma(X) \vee \mathcal{G}$, then

$$\mathbb{E}[X|\mathcal{G} \vee \mathcal{H}] = \mathbb{E}[X|\mathcal{G}]$$

$\mathbb{P}$-a.s.. This is property (k) in the chapter 9.7 of Williams (1991). As a consequence, we find for $t \geqslant s$,

$$\mathbb{E}[M_{t\wedge T_n}(x,\zeta)|\mathcal{F}_s^{A,U}] = \mathbb{E}[M_{t\wedge T_n}(x,\zeta)|\mathcal{F}_s^{A}] = M_{s\wedge T_n}(x,\zeta).$$

It follows that $(M_t(x,\zeta))_{t\geqslant 0} \in \mathcal{M}_{\mathrm{loc}}(\mathbb{P}, \mathbb{F}^{A,U})$. ☐

[6]Which amounts to choosing b_1 and b_2, as will be shown in the next propositions.

We can use the local-martingales $(M_t(x,\zeta))_{t\geqslant 09}$ as densities to define equivalent measures $\mathbb{Q}^{x,\zeta}$ as

$$\mathbb{Q}^{x,\zeta}(A|\mathcal{F}_s^{A,U}) = \mathbb{E}\left[\mathbf{1}_A \frac{M_t(x,\zeta)}{M_s(x,\zeta)}|\mathcal{F}_s^{A,U}\right] \tag{4.40}$$

which is also written as

$$\left.\frac{\mathrm{d}\mathbb{Q}^{x,\zeta}}{\mathrm{d}\mathbb{P}}\right|_{\mathcal{F}_t} = \frac{M_t(x,\zeta)}{M_0(x,\zeta)}.$$

The next proposition shows that such measures preserve the structure of the model we specified.

Proposition 4.9. *Under the measure $\mathbb{Q}^{x,\zeta}$, the intensity $(\lambda_{S_t})_{t\geqslant 0}$ satisfies the SDE*

$$\mathrm{d}\lambda_{S_t} = \kappa(\theta^{x,\zeta} - \lambda_{S_{t-}})\mathrm{d}S_t + \eta^{x,\zeta}\mathrm{d}L_{S_t}$$

where

$$\frac{\theta^{x,\zeta}}{\theta} = \frac{\eta^{x,\zeta}}{\eta} = \psi(0, \eta b_1 + x).$$

The distribution of the α-stable subordinator remains unchanged under $\mathbb{Q}^{x,\zeta}$. The jumps of the processes $(D_t)_{t\geqslant 0}$, $(J_t)_{t\geqslant 0}$ and $(L_t)_{t\geqslant 0}$ defined at Equation (4.4) *have intensity $(\lambda_t)_{t\geqslant 0}$ under $\mathbb{Q}^{x,\zeta}$. Moreover, the jump sizes $\xi_1, \xi_2, \ldots$ remain double exponential under this measure and their distribution is given by*

$$\nu^{x,\zeta}(B) = p^{x,\zeta}\int_{B\cap\mathbb{R}^+} \rho_+^{x,\zeta} e^{-\rho_+^{x,\zeta}z}\mathrm{d}z + (1-p^{x,\zeta})\int_{B\cap\mathbb{R}^-} \rho_-^{x,\zeta} e^{-\rho_-^{x,\zeta}z}\mathrm{d}z,$$

with

$$\rho_+^{x,\zeta} = \rho_+ - (\eta b_1 + x),$$
$$\rho_-^{x,\zeta} = \rho_- + (\eta b_1 + x)$$

and

$$p^{x,\zeta} = \frac{p\rho_+\rho_-^{x,\zeta}}{p\rho_+\rho_-^{x,\zeta} + (1-p)\rho_+^{x,\zeta}\rho_-}.$$

Proof. Propositions 4.2 and 4.3 in Moraux and Hainaut (2018) establish that under $\mathbb{Q}^{x,\zeta}$,

$$\lambda_t = \lambda_0 + \int_0^t \kappa(\theta^{x,\zeta} - \lambda_{s-})\mathrm{d}s + \eta^{x,\zeta} L_t.$$

Thanks to the change of variable formula (4.38), it follows that

$$\lambda_{S_t} = \lambda_0 + \int_0^t \kappa(\theta^{x,\zeta} - \lambda_{S_{s-}})\mathrm{d}S_s + \eta^{x,\zeta} L_{S_t}.$$

The fact that the law of $(U_t)_{t\geqslant t}$ remains unchanged follows from the independence of $\mathbb{F}^A$ and $\mathbb{F}^U$. The remainder of the statement corresponds to Propositions 4.2 and 4.3 in Moraux and Hainaut (2018). $\square$

Proposition 4.10. *If the process* $(\zeta_t)_{t\geqslant 0}$ *satisfy*

$$\zeta_t = \frac{\mu + \lambda_{t-}[\psi(0,\eta b_1 + x)(\psi^{x,\zeta}(1,0)-1) - (\psi(1,0)-1)] - r}{\sigma},$$

and if the relations of Proposition 4.8 are satisfied, then the equivalent measure $\mathbb{Q}^{x,\zeta}$ *of Equation* (4.40) *is a risk-neutral measure. Under this measure, the asset price* $(A_{S_t})_{t\geqslant 0}$ *satisfies the SDE*

$$\frac{\mathrm{d}A_{S_t}}{A_{S_{t-}}} = r\mathrm{d}S_t + \sigma\mathrm{d}W_{S_t} + \left(\mathrm{d}D_{S_t} - \mathbb{E}_{\mathbb{Q}^{x,\zeta}}[e^{\xi}-1]\lambda_{S_{t-}}\mathrm{d}S_t\right)$$

whereas its logarithm $(X_{S_t})_{t\geqslant 0} = (\ln A_{S_t})_{t\geqslant 0}$ *satisfies*

$$\mathrm{d}X_{S_t} = \left(r - \frac{\sigma^2}{2} - \mathbb{E}_{\mathbb{Q}^{x,\zeta}}[e^{\xi}-1]\lambda_{S_{t-}}\right)\mathrm{d}S_t + \sigma\mathrm{d}W_{S_t} + \mathrm{d}J_{S_t}.$$

Proof. It is proved in Moraux and Hainaut (2018) (Proposition 4.4) that under $\mathbb{Q}^{x,\zeta}$,

$$A_t = r\int_0^t A_{s-}\mathrm{d}s + \sigma\int_0^t A_{s-}\mathrm{d}W_s + \int_0^t A_{s-}\left(\mathrm{d}D_s - \mathbb{E}_{\mathbb{Q}^{x,\zeta}}[e^{\xi}-1]\lambda_{s-}\mathrm{d}s\right) \quad (4.41)$$

and

$$X_t = \int_0^t \left(r - \frac{\sigma^2}{2} - \mathbb{E}_{\mathbb{Q}^{x,\zeta}}[e^{\xi}-1]\lambda_{s-}\right)\mathrm{d}s + \sigma W_t + J_t.$$

The dynamics of $(A_{S_t})_{t\geqslant 0}$ and $(X_{S_t})_{t\geqslant 0}$ then follow from the change of variable formula (4.38). Moreover, note that by Equation (4.41) and Ito's lemma applied on the process $(\tilde{A}_t)_{t\geqslant 0} = (e^{-rt}A_t)_{t\geqslant 0}$, we have

$$\tilde{A}_t = \sigma\int_0^t \tilde{A}_{s-}\mathrm{d}W_s + \int_0^t \tilde{A}_{s-}\left(\mathrm{d}D_s - \mathbb{E}_{\mathbb{Q}^{x,\zeta}}[e^{\xi}-1]\lambda_{s-}\mathrm{d}s\right)$$

or using the random measure Ξ,

$$\tilde{A}_t = \sigma\int_0^t \tilde{A}_{s-}\mathrm{d}W_s + \int_0^t \tilde{A}_{s-}\int_{z\in\mathbb{R}}(e^z-1)(\Xi(\mathrm{d}z,\mathrm{d}s) - \lambda_{s-}\nu(\mathrm{d}z)\mathrm{d}s)$$

under $\mathbb{Q}^{x,\zeta}$. Again, from the change of variable formula (4.38),

$$\tilde{A}_{S_t} = \sigma\int_0^t \tilde{A}_{S_{s-}}\mathrm{d}W_{S_s} + \int_0^t \tilde{A}_{S_{s-}}\int_{z\in\mathbb{R}}(e^z-1)(\Xi(\mathrm{d}z,\mathrm{d}S_s) - \lambda_{s-}\nu(\mathrm{d}z)\mathrm{d}S_s). \quad (4.42)$$

Since the Brownian motion $(W_t)_{t\geqslant 0}$ belongs to $\mathcal{M}^{\mathrm{c}}_{\mathrm{loc}}(\mathbb{Q}^{x,\zeta}, \mathbb{F}^A)$, Corollary 4.4 implies that the time-changed Brownian motion $(W_{S_t})_{t\geqslant 0}$ belongs to $\mathcal{M}^{\mathrm{c}}_{\mathrm{loc}}(\mathbb{Q}^{x,\zeta}, \mathbb{F}^{A\circ S})$. Moreover, Proposition 4.7 states that

$$\left(\int_0^t \int_{z\in\mathbb{R}} (e^z - 1)\left(\Xi(\mathrm{d}z, \mathrm{d}s) - \lambda_{s-}\nu(\mathrm{d}z)\mathrm{d}s\right)\right)_{t\geqslant 0}$$

is in $\mathcal{M}^{\mathrm{d}}_{\mathrm{loc}}(\mathbb{Q}^{x,\zeta}, \mathbb{F}^A)$. Therefore, Corollary 4.4 and the change of variable formula (4.38) entail that the time-changed compensated jump process

$$\left(\int_0^t \int_{z\in\mathbb{R}} (e^z - 1)\left(\Xi(\mathrm{d}z, \mathrm{d}S_s) - \lambda_{S_{s-}}\nu(\mathrm{d}z)\mathrm{d}S_s\right)\right)_{t\geqslant 0}$$

is in $\mathcal{M}^{\mathrm{d}}_{\mathrm{loc}}(\mathbb{Q}^{x,\zeta}, \mathbb{F}^{A\circ S})$. Since stochastic integrals with respect to local-martingales are local-martingales, Equation (4.42) implies that $(\tilde{A}_{S_t})_{t\geqslant 0}$ is in $\mathcal{M}_{\mathrm{loc}}(\mathbb{Q}^{x,\zeta}, \mathbb{F}^{A\circ S})$, which ends the proof. □

After having introduced a class of risk-neutral measures, the following question arises: which risk-neutral measure should we choose? A possible approach, described e.g. in Section 10.5 of Cont and Tankov (2004), is to choose the risk-neutral measure $\mathbb{Q}$ that minimizes some kind of distance with respect to the physical measure $\mathbb{P}$. A first possibility is to choose the risk-neutral measure that minimizes the relative entropy, also called the Kullback-Leibler distance. The measure is

$$\mathbb{Q}_{\mathrm{E}} = \arg \min_{\mathbb{Q}^{x,\zeta}\in\mathcal{Q}} \mathbb{E}_{\mathbb{P}}\left[\frac{\mathrm{d}\mathbb{Q}^{x,\zeta}}{\mathrm{d}\mathbb{P}} \ln \frac{\mathrm{d}\mathbb{Q}^{x,\zeta}}{\mathrm{d}\mathbb{P}}\right],$$

where $\mathcal{Q}$ denotes the class of risk-neutral measures described above. This leads to the minimal entropy risk-neutral measure. The minimal entropy risk-neutral measure for exponential Lévy asset prices has been studied in Chan (1999), Fujiwara and Miyahara (2003), and Hubalek and Sgarra (2006). Another possible criterion is the variance, in which case the measure is

$$\mathbb{Q}_{\mathrm{V}} = \arg \min_{\mathbb{Q}^{x,\zeta}\in\mathcal{Q}} \mathbb{E}_{\mathbb{P}}\left[\left(\frac{\mathrm{d}\mathbb{Q}^{x,\zeta}}{\mathrm{d}\mathbb{P}}\right)^2\right] = \arg \min_{\mathbb{Q}^{x,\zeta}\in\mathcal{Q}} \mathbb{V}\mathrm{ar}_{\mathbb{P}}\left(\frac{\mathrm{d}\mathbb{Q}^{x,\zeta}}{\mathrm{d}\mathbb{P}}\right)$$

where the second equality holds because we necessarily have $\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}^{x,\zeta}}{\mathrm{d}\mathbb{P}}\right] = 1$. The obtained measure $\mathbb{Q}_{\mathrm{V}}$ is then called the minimum variance risk-neutral measure.

We derive measures $\mathbb{Q}_{\mathrm{V}}$ and $\mathbb{Q}_{\mathrm{E}}$ that allow to reasonably minimize the criteria of variance and relative entropy while remaining inside the measure class $\mathcal{Q}$ we introduced. To lighten the notations, we write M_t for $M_t(x, \zeta)$ and m_t for $\ln M_t(x, \zeta)$. In the minimum variance case, our goal is to minimize $\mathbb{E}[M_t^2]$. From the definition of $(M_t)_{t\geqslant 0}$, it is easily seen that

$$\begin{aligned}\mathrm{d}m_t &= b_1\mathrm{d}\lambda_t + x\mathrm{d}L_t + b_2\mathrm{d}t - \frac{1}{2}\zeta_t^2\mathrm{d}t - \zeta_t\mathrm{d}W_t\\ &= \left[b_1\kappa(\theta-\lambda_{t-}) + b_2 - \frac{1}{2}\zeta_t^2\right]\mathrm{d}t - \zeta_t\mathrm{d}W_t + (x+b_1\eta)\int_{z\in\mathbb{R}} |z|\ \Xi(\mathrm{d}z,\mathrm{d}t).\end{aligned}$$

Since $M_t^2 = \exp\{2m_t\}$, Ito's lemma for semimartingales yields

$$\begin{aligned}\mathrm{d}M_t^2 &= 2M_{t-}^2\left(\left[b_1\kappa(\theta-\lambda_{t-}) + b_2 + \frac{1}{2}\zeta_t^2\right]\mathrm{d}t - \zeta_t\mathrm{d}W_t\right)\\ &\quad + M_{t-}^2\int_{z\in\mathbb{R}}(\exp\{2(b_1\eta+x)|z|\}-1)\ \Xi(\mathrm{d}z,\mathrm{d}t)\\ &= 2M_{t-}^2\left[\lambda_{t-}\left(\frac{1}{2}[\psi(0,2(b_1\eta+x))-1]-b_1\kappa\right) + \left(b_1\kappa\theta + b_2 + \frac{1}{2}\zeta_t^2\right)\right]\mathrm{d}t\\ &\quad - 2M_{t-}^2\zeta_t\mathrm{d}W_t + M_{t-}^2\int_{z\in\mathbb{R}}(\exp\{2(b_1\eta+x)|z|\}-1)\ \tilde{\Xi}(\mathrm{d}z,\mathrm{d}t),\end{aligned}$$

where we have used

$$\begin{aligned}\int_{z\in\mathbb{R}}(\exp\{2(b_1\eta+x)|z|\}-1)\ v(\mathrm{d}z,\mathrm{d}t) &= \int_{z\in\mathbb{R}}(\exp\{2(b_1\eta+x)|z|\}-1)\ \nu(\mathrm{d}z)\lambda_{t-}\mathrm{d}t\\ &= [\psi(0,2(b_1\eta+x))-1]\lambda_{t-}\mathrm{d}t.\end{aligned}$$

Using the martingale conditions $b_1\kappa\theta + b_2 = 0$ and $b_1\kappa = \psi(0, b_1\eta + x) - 1$ from Proposition 4.8, the SDE becomes

$$\begin{aligned}\mathrm{d}M_t^2 &= 2M_{t-}^2\left[\lambda_{t-}\left(\frac{1}{2}\psi(0,2(b_1\eta+x)) - \psi(0,b_1\eta+x) + \frac{1}{2}\right) + \frac{1}{2}\zeta_t^2\right]\mathrm{d}t\\ &\quad - 2M_{t-}^2\zeta_t\mathrm{d}W_t + M_{t-}^2\int_{z\in\mathbb{R}}(\exp\{2(b_1\eta+x)|z|\}-1)\ \tilde{\Xi}(\mathrm{d}z,\mathrm{d}t).\end{aligned}\tag{4.43}$$

From Equation (4.43), we observe that we have to minimize

$$\frac{1}{2}\psi(0,2(b_1\eta+x)) - \psi(0,b_1\eta+x).$$

Since this quantity only depends on $b_1\eta + x$, set $a := b_1\eta + x$. The function

$$a \mapsto \frac{1}{2}\psi(0,2a) - \psi(0,a)$$

admits a unique global minimum at $a = 0$, for a value of $\psi(0,0) = -1/2$. Indeed, from Jensen's inequality,

$$\frac{1}{2}\psi(0,2a) - \psi(0,a) \geqslant \frac{1}{2}\psi^2(0,a) - \psi(0,a).$$

Since the strictly convex function $x \mapsto \frac{1}{2}x^2 - x$ reaches its global minimum at $x = 1$, we get

$$\frac{1}{2}\psi(0, 2a) - \psi(0, a) \geqslant -1/2,$$

proving that $a = 0$ corresponds to the global minimum of $a \mapsto \frac{1}{2}\psi(0, 2a) - \psi(0, a)$.

Note that by Proposition 4.9, the law of the stock price under the risk-neutral measure depends only on $a = b_1\eta + x$. We can thus choose to set $x = b_1 = b_2 = 0$.

We move to the minimal entropy risk-neutral measure, for which we have to minimize $\mathbb{E}[M_t \ln M_t] = \mathbb{E}[m_t e^{m_t}]$. The function $f : \mathbb{R} \to \mathbb{R}$ defined as $f(x) = xe^x$ satisfies

$$f'(x) = e^x + xe^x$$

and

$$f''(x) = 2e^x + xe^x.$$

As a consequence, Ito's lemma yields

$$\begin{aligned}
&\mathrm{d}(m_t e^{m_t}) = \mathrm{d}f(m_t) \\
&= e^{m_{t-}}\left[b_1\kappa(\theta - \lambda_{t-}) + b_2 - \frac{1}{2}\zeta_t^2\right]\mathrm{d}t - e^{m_{t-}}\zeta_t \mathrm{d}W_t \\
&\quad + m_{t-}e^{m_{t-}}\left[b_1\kappa(\theta - \lambda_{t-}) + b_2 - \frac{1}{2}\zeta_t^2\right]\mathrm{d}t - m_{t-}e^{m_{t-}}\zeta_t \mathrm{d}W_t \\
&\quad + \frac{1}{2}(2e^{m_{t-}} + m_{t-}e^{m_{t-}})\zeta_t^2 \mathrm{d}t \\
&\quad + \int_{z\in\mathbb{R}}\left[(m_{t-} + a|z|)e^{m_{t-}+a|z|} - m_{t-}e^{m_{t-}}\right]\Xi(\mathrm{d}z, \mathrm{d}t) \\
&= (1 + m_{t-})e^{m_{t-}}(b_1\kappa(\theta - \lambda_{t-}) + b_2)\mathrm{d}t + \frac{1}{2}\zeta_t^2 e^{m_{t-}}\mathrm{d}t - (e^{m_{t-}} + m_{t-}e^{m_{t-}})\zeta_t \mathrm{d}W_t \\
&\quad + m_{t-}e^{m_{t-}}[\psi(0, a) - 1]\lambda_{t-}\mathrm{d}t + ae^{m_{t-}}\psi'(0, a)\lambda_{t-}\mathrm{d}t \\
&\quad + \int_{z\in\mathbb{R}}\left[(m_{t-} + a|z|)e^{m_{t-}+a|z|} - m_{t-}e^{m_{t-}}\right]\tilde{\Xi}(\mathrm{d}z, \mathrm{d}t),
\end{aligned}$$

where

$$\psi'(0, a) = \frac{\partial\psi}{\partial z_2}(z_1, z_2)\bigg|_{(z_1,z_2)=(0,a)} = \mathbb{E}[|\xi|e^{a|\xi|}].$$

By using the martingale conditions $b_2 = -\kappa\theta b_1$ and $\psi(0, a) - 1 = \kappa b_1$ of Proposition 4.8, we obtain after some simplifications

$$\begin{aligned}
\mathrm{d}(m_t e^{m_t}) = e^{m_{t-}}&\left[\lambda_{t-}\left(a\psi'(0, a) - \psi(0, a) + 1\right) + \frac{1}{2}\zeta_t^2\right]\mathrm{d}t - (e^{m_{t-}} + m_{t-}e^{m_{t-}})\mathrm{d}W_t \\
&+ \int_{z\in\mathbb{R}}\left[(m_{t-} + a|z|)e^{m_{t-}+a|z|} - m_{t-}e^{m_{t-}}\right]\tilde{\Xi}(\mathrm{d}z, \mathrm{d}t).
\end{aligned} \tag{4.44}$$

From Equation (4.44), we deduce that we have to minimize the function

$$a \mapsto a\psi'(0,a) - \psi(0,a) = \mathbb{E}[a|\xi|e^{a|\xi|} - e^{a|\xi|}].$$

Since the function $x \mapsto xe^x - e^x$ reaches its unique global minimum at $x = 0$, the function $a \mapsto a\psi'(0,a) - \psi(0,a)$ admits a unique global minimum at $a = 0$. This gives the same result as in the variance minimization case, so that $\mathbb{Q}_{\mathrm{V}} = \mathbb{Q}_{\mathrm{E}}$. We conclude from Proposition 4.9 that both under $\mathbb{Q}_{\mathrm{V}}$ and $\mathbb{Q}_{\mathrm{E}}$, all the parameters (with of course the exception of the drift of the stock price μ) remain as under the physical measure $\mathbb{P}$.

To obtain this result, we have ignored the term ζ_t^2 in Equations (4.43) and (4.44). However, by the definition of $(\zeta_t)_{t\geqslant 0}$ (Proposition 4.10), this process depends on $b_1\eta + x$ and therefore should be part of the quantity to minimize to find the optimal measures. The term ζ_t^2 was not included in the analysis because its inclusion would lead to tedious computations, and the resulting measure would not be part of the class of measures we introduced.

4.5 Pricing of call options

In this section, we show how to price call and put options in our framework. We proceed by numerically inverting the Fourier transform of the option prices.

Numerical computation of the Fourier transform

This subsection is devoted to a numerical method that allows to solve the FPDE of Corollary 4.3. We work with a grid on $[0,T] \times [-z_{\max}, z_{\max}]$ where $T, z_{\max} > 0$ are fixed constants. The discretization step sizes are $\Delta_t := T/n_t$ and $\Delta_z := z_{\max}/n_z$ with $n_t, n_z \in \mathbb{N}$. The grid is defined as $\{t_0, \ldots, t_{n_t}\} \times \{z_{-n_z}, z_{-n_z+1}, \ldots, z_{n_z}\}$ where $t_x := x\Delta_t$ and $z_y := y\Delta_z$. This grid contains $(n_t+1) \times (2n_z+1)$ points. We set $\sigma_\alpha := 1 - \frac{\alpha}{2}$. Note that for any $\alpha \in (0,1)$, we have $\sigma_\alpha \in (\frac{1}{2}, 1)$, so that the inequalities $t_j < t_j + \frac{1}{2}\Delta_t < t_{j+\sigma_\alpha} < t_{j+1}$ always hold.

We use $\hat{\varphi}_\alpha^{(0)}$ to denote the approximated values of the function $\varphi_\alpha^{(0)}$ from Corollary 4.3. That is, for fixed s, y_1, y_2, v, u and $\omega \in \mathbb{C}$, we write

$$\beta\hat{\varphi}_\alpha^{(0)}(t,z) \approx \beta(\omega,z)\varphi_\alpha^{(0)}(t,r,\omega,z|s,y_1,y_2,v,u)$$

to express the fact that $\beta\hat{\varphi}_\alpha^{(0)}(t,z)$ is the quantity that aims to approximate

$$\beta(\omega,z_i)\varphi_\alpha^{(0)}(t,r,\omega,z|s,y_1,y_2,v,u).$$

Similarly, we also write

$$\frac{\Delta^\alpha\hat{\varphi}_\alpha^{(0)}}{\Delta t^\alpha}(t,z) \approx \frac{\partial^\alpha\varphi_\alpha^{(0)}}{\partial t^\alpha}(t,r,\omega,z|s,y_1,y_2,v,u)$$

and

$$\gamma \frac{\Delta \hat{\varphi}_\alpha^{(0)}}{\Delta z}(t, z) \approx \gamma(r, \omega, z) \frac{\partial \varphi_\alpha^{(0)}}{\partial z}(t, r, \omega, z | s, y_1, y_2, v, u).$$

The FPDE of Corollary 4.3 translates into

$$\frac{\Delta^\alpha \hat{\varphi}_\alpha^{(0)}}{\Delta t^\alpha}(t_{j+\sigma_\alpha}, z_i) = \beta \hat{\varphi}_\alpha^{(0)}(t_{j+\sigma_\alpha}, z_i) + \gamma \frac{\Delta \hat{\varphi}_\alpha^{(0)}}{\Delta z}(t_{j+\sigma_\alpha}, z_i). \tag{4.45}$$

We will now give the formulas for our approximations and study their rates of convergence.

The approximation used for the Caputo fractional derivative is the one introduced in Alikhanov (2015). The Caputo derivative approximation for a function f is

$$\frac{\mathrm{d}^\alpha f}{\mathrm{d}t^\alpha}(t_{j+\sigma_\alpha}) \approx \frac{\Delta^\alpha f}{\Delta t^\alpha}(t_{j+\sigma_\alpha}) := \frac{\Delta_t^\alpha}{\Gamma(1-\alpha)} \sum_{s=0}^{j} c_{j-s}^{(\alpha,j)} \frac{f(t_{s+1}) - f(t_s)}{\Delta_t} \tag{4.46}$$

where $c_0^{(\alpha,0)} := \sigma_\alpha^{1-\alpha}$ and

$$c_s^{(\alpha,j)} := \begin{cases} a_0^{(\alpha)} + b_1^{(\alpha)} & \text{if } s = 0, \\ a_s^{(\alpha)} + b_{s+1}^{(\alpha)} - b_s^{(\alpha)} & \text{if } 1 \leqslant s \leqslant j-1, \\ a_j^{(\alpha)} - b_j^{(\alpha)} & \text{if } s = j, \end{cases} \tag{4.47}$$

when $j \in \{1, \ldots, n_t\}$. The constants $a_s^{(\alpha)}$ and $b_s^{(\alpha)}$ are defined as

$$a_s^{(\alpha)} := (s + \sigma_\alpha)^{1-\alpha} - (s - 1 + \sigma_\alpha)^{1-\alpha} \tag{4.48}$$

and

$$\begin{aligned} b_s^{(\alpha)} :=& \frac{1}{2-\alpha} \left[(s+\sigma_\alpha)^{2-\alpha} - (s-1+\sigma_\alpha)^{2-\alpha} \right] \\ &- \frac{1}{2} \left[(s+\sigma_\alpha)^{1-\alpha} + (s-1+\sigma_\alpha)^{1-\alpha} \right]. \end{aligned} \tag{4.49}$$

The approximation formula of Equations (4.46)-(4.49) is based on a quadratic interpolation of f on the grid $\{0 = t_0, t_1, \ldots, t_{n_t} = T\}$ and is derived in Alikhanov (2015). This approximation has the benefit of a high order error, that is $\mathcal{O}(\Delta_t^{3-\alpha})$, as stated in the next proposition.

Proposition 4.11. *For any $\alpha \in (0,1)$ and $f \in \mathcal{C}^3[0, t_{j+1}]$,*

$$\left| \frac{\mathrm{d}^\alpha f}{\mathrm{d}t^\alpha}(t_{j+\sigma_\alpha}) - \frac{\Delta^\alpha f}{\Delta t^\alpha}(t_{j+\sigma_\alpha}) \right| = \mathcal{O}(\Delta_t^{3-\alpha}).$$

Proof. See Lemma 2 in Alikhanov (2015). □

The fractional derivative of Corollary 4.3 is thus approximated by

$$\frac{\Delta^\alpha \hat{\varphi}_\alpha^{(0)}}{\Delta t^\alpha}(t_{j+\sigma_\alpha}, z_i) := \frac{\Delta_t^\alpha}{\Gamma(1-\alpha)} \sum_{s=0}^{j} c_{j-s}^{(\alpha,j)} \frac{\hat{\varphi}_\alpha^{(0)}(t_{s+1}, z_i) - \hat{\varphi}_\alpha^{(0)}(t_s, z_i)}{\Delta_t}. \tag{4.50}$$

We move to the approximation $\beta\hat{\varphi}_\alpha^{(0)}(t,z)$. The next proposition introduces the formula we use.

Proposition 4.12. *Let $f \in \mathcal{C}^2[0,T]$. Then,*

$$f(t_{j+\sigma_\alpha}) = (1-\sigma_\alpha) f(t_j) + \sigma_\alpha f(t_{j+1}) + \mathcal{O}(\Delta_t^2).$$

Proof. The proof is based on Taylor's theorem. This theorem implies that

$$\begin{aligned}
&\sigma_\alpha f(t_{j+1}) \\
&= \sigma_\alpha \left(f(t_{j+\sigma_\alpha}) + (t_{j+1} - t_{j+\sigma_\alpha}) f'(t_{t_{j+\sigma_\alpha}}) + \frac{(t_{j+1} - t_{j+\sigma_\alpha})^2}{2} f''(\tilde{t}_{j+1}) \right) \\
&= \sigma_\alpha \left(f(t_{j+\sigma_\alpha}) + (1-\sigma_\alpha)\Delta_t f'(t_{t_{j+\sigma_\alpha}}) + \frac{((1-\sigma_\alpha)\Delta_t)^2}{2} f''(\tilde{t}_{j+1}) \right)
\end{aligned} \tag{4.51}$$

and

$$\begin{aligned}
&(1-\sigma_\alpha) f(t_j) \\
&\quad = (1-\sigma_\alpha) \left(f(t_{j+\sigma_\alpha}) + (t_j - t_{j+\sigma_\alpha}) f'(t_{j+\sigma_\alpha}) + \frac{(t_{j+\sigma_\alpha} - t_j)^2}{2} f''(\tilde{t}_j) \right) \\
&\quad = (1-\sigma_\alpha) \left(f(t_{j+\sigma_\alpha}) - \sigma_\alpha \Delta_t f'(t_{j+\sigma_\alpha}) + \frac{(\sigma_\alpha \Delta_t)^2}{2} f''(\tilde{t}_j) \right)
\end{aligned} \tag{4.52}$$

for some $\tilde{t}_{j+1} \in (t_{j+\sigma_\alpha}, t_{j+1})$ and $\tilde{t}_j \in (t_j, t_{j+\sigma_\alpha})$. Summing Equations (4.51) and (4.52) yields the announced result. □

The term $\beta\hat{\varphi}_\alpha^{(0)}(t,z)$ is therefore computed as

$$\beta\hat{\varphi}_\alpha^{(0)}(t_{j+\sigma_\alpha}, z_i) := \beta(r,\omega,z_i) \left[\sigma_\alpha \hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_i) + (1-\sigma_\alpha)\hat{\varphi}_\alpha^{(0)}(t_j, z_i) \right] \tag{4.53}$$

for all $i \in \{-n_z, \ldots, n_z\}$. The approximation $\gamma \frac{\Delta \hat{\varphi}_\alpha^{(0)}}{\Delta z}(t,z)$ is more tricky. This happens for two reasons. The first is that we have to control the rates of convergence with respect to both Δ_t and Δ_z. The second is that we have to rely on 3 different approximations in order not to fall outside of the grid $\{z_{-n_z}, \ldots, z_{n_z}\}$ when approximating the partial derivative with respect to z. The falling outside of the grid problem will be explained more precisely and addressed later on, in the case $i \in \{-n_z, n_z\}$. For now, we assume that $i \in \{-n_z+1, \ldots, n_z-1\}$. The approximation is in this case based on the following proposition.

Proposition 4.13. *Let $f, k \in \mathcal{C}^3[-z_{\max}, z_{\max}]$. Then,*

$$k(z_i)f'(z_i) = \frac{k(z_{i+\frac{1}{2}})f(z_{i+1}) - \Delta_z k'(z_i)f(z_i) - k(z_{i-\frac{1}{2}})f(z_{i-1})}{2\Delta_z} + \mathcal{O}(\Delta_z^2).$$

Proof. By Taylor's theorem,

$$k(z_{i+\frac{1}{2}}) = k(z_i) + \frac{\Delta_z}{2}k'(z_i) + \frac{1}{2}\left(\frac{\Delta_z}{2}\right)^2 k''(z_i) + \frac{1}{6}\left(\frac{\Delta_z}{2}\right)^3 k'''(\tilde{z}_{i+\frac{1}{2}}) \tag{4.54}$$

for some $\tilde{z}_{i+\frac{1}{2}} \in (z_i, z_{i+\frac{1}{2}})$ and

$$k(z_{i-\frac{1}{2}}) = k(z_i) - \frac{\Delta_z}{2}k'(z_i) + \frac{1}{2}\left(\frac{\Delta_z}{2}\right)^2 k''(z_i) - \frac{1}{6}\left(\frac{\Delta_z}{2}\right)^3 k'''(\tilde{z}_{i-\frac{1}{2}}) \tag{4.55}$$

for some $\tilde{z}_{i-\frac{1}{2}} \in (z_{i-\frac{1}{2}}, z_i)$. Similarly,

$$f(z_{i+1}) = f(z_i) + \Delta_z f'(z_i) + \frac{\Delta_z^2}{2}f''(z_i) + \frac{\Delta_z^3}{6}f'''(\tilde{z}_{i+1}) \tag{4.56}$$

for some $\tilde{z}_{i+1} \in (z_i, z_{i+1})$ and

$$f(z_{i-1}) = f(z_i) - \Delta_z f'(z_i) + \frac{\Delta_z^2}{2}f''(z_i) - \frac{\Delta_z^3}{6}f'''(\tilde{z}_{i-1}) \tag{4.57}$$

for some $\tilde{z}_{i-1} \in (z_{i-1}, z_i)$. Multiplying Equation (4.54) by Equation (4.56) and Equation (4.55) by Equation (4.57) leads to

$$k(z_{i+\frac{1}{2}})f(z_{i+1}) - k(z_{i-\frac{1}{2}})f(z_{i-1}) = 2\Delta_z k(z_i)f'(z_i) + \Delta_z k'(z_i)f(z_i) + \mathcal{O}(\Delta_z^3).$$

A division by Δ_z then implies that

$$\frac{k(z_{i+\frac{1}{2}})f(x_{i+1}) - \Delta_z k'(z_i)f(x_i) - k(z_{i-\frac{1}{2}})f(x_{i-1})}{2\Delta_z} = k(z_i)f'(z_i) + \mathcal{O}(\Delta_z^2),$$

as announced. □

Let k be a function that belongs to $\mathcal{C}^3[-z_{\max}, z_{\max}]$ and $g : [0,T] \times [-z_{\max}, z_{\max}] \to \mathbb{R} : (t,z) \mapsto g(t,z)$ be a function that is twice continuously differentiable with respect to its first argument and three times continuously differentiable with respect to its second argument. Proposition 4.12 implies that

$$k(z_i)\frac{\partial g}{\partial z}(t_{j+\sigma_\alpha}, z_i) = \sigma_\alpha k(z_i)\frac{\partial g}{\partial z}(t_{j+1}, z_i) + (1-\sigma_\alpha)k(z_i)\frac{\partial g}{\partial z}(t_j, z_i) + \mathcal{O}(\Delta_t^2). \tag{4.58}$$

Moreover, Proposition 4.13 applied on each term at the right-hand side of Equation (4.58) provides us with

$$\begin{aligned} k(z_i)\frac{\partial g}{\partial z}(t_{j+\sigma_\alpha}, z_i) &= \mathcal{O}(\Delta_z^2 + \Delta_t^2) \\ &+ \sigma_\alpha \frac{k(z_{i+\frac{1}{2}})g(t_{j+1}, z_{i+1}) - k'(z_i)g(t_{j+1}, z_i) - k(z_{i-\frac{1}{2}})g(t_{j+1}, z_{i-1})}{2\Delta_z} \\ &+ (1-\sigma_\alpha)\frac{k(z_{i+\frac{1}{2}})g(t_j, z_{i+1}) - k'(z_i)g(t_j, z_i) - k(z_{i-\frac{1}{2}})g(t_j, z_{i-1})}{2\Delta_z} \end{aligned} \tag{4.59}$$

The approximation $\gamma\frac{\Delta\hat{\varphi}_\alpha^{(0)}}{\Delta z}(t,z)$ then follows from Equation (4.59), that is

$$\begin{aligned} &\gamma\frac{\Delta\hat{\varphi}_\alpha^{(0)}}{\Delta z}(t_{j+\sigma_\alpha}, z_i) \\ &:= \sigma_\alpha\frac{\gamma(\omega, z_{i+\frac{1}{2}})\hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_{i+1}) - \Delta_z\frac{\partial\gamma}{\partial z_i}(\omega, z_i)\hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_i) - \gamma(\omega, z_{i-\frac{1}{2}})\hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_{i-1})}{2\Delta_z} \\ &+ (1-\sigma_\alpha)\frac{\gamma(\omega, z_{i+\frac{1}{2}})\hat{\varphi}_\alpha^{(0)}(t_j, z_{i+1}) - \Delta_z\frac{\partial\gamma}{\partial z_i}(\omega, z_i)\hat{\varphi}_\alpha^{(0)}(t_j, z_i) - \gamma(\omega, z_{i-\frac{1}{2}})\hat{\varphi}_\alpha^{(0)}(t_j, z_{i-1})}{2\Delta_z}. \end{aligned} \tag{4.60}$$

Unfortunately, Approximation (4.60) cannot be used when $i \in \{-n_z, n_z\}$. The reason is what we previously referred to as the falling outside of the grid problem. As a matter of fact, the z_{i+1} (resp. z_{i-1}) that appears in Equation (4.60) prevents us from using this formula for $i = n_z$ (resp. $i = -n_z$) without falling outside of the grid $\{z_{-n_z}, \dots, z_{n_z}\}$. The next proposition introduces the approximations we use when $i \in \{-n_z, n_z\}$.

Proposition 4.14. *Let $f \in \mathcal{C}^3[-z_{\max}, z_{\max}]$. Then we have*

(*i*)

$$f'(-z_{\max}) = \frac{-f(-z_{\max}+2\Delta_z) + 4f(-z_{\max}+\Delta_z) - 3f(-z_{\max})}{2\Delta_z} + \mathcal{O}(\Delta_z^2)$$

(*ii*)

$$f'(z_{\max}) = \frac{3f(z_{\max}) - 4f(z_{\max}-\Delta_z) + f(z_{\max}-2\Delta_z)}{2\Delta_z} + \mathcal{O}(\Delta_z^2).$$

Proof. By Taylor's theorem,

$$\frac{f(-z_{\max}+\Delta_z) - f(-z_{\max})}{\Delta_z} = f'(-z_{\max}) + \frac{\Delta_z}{2}f''(-z_{\max}) + \mathcal{O}(\Delta_z^2) \tag{4.61}$$

and

$$\begin{aligned} f''(-z_{\max}) &= f''(-z_{\max}+\Delta_z) + \mathcal{O}(\Delta_z) \\ &= \frac{f(-z_{\max}+2\Delta_z) - 2f(-z_{\max}+\Delta_z) + f(-z_{\max})}{\Delta_z^2} + \mathcal{O}(\Delta_z). \end{aligned} \tag{4.62}$$

Substituting Equation (4.62) into Equation (4.61) gives

$$\begin{aligned}&\frac{f(-z_{\max}+\Delta_z)-f(-z_{\max})}{\Delta_z}\\&\qquad=f'(-z_{\max})+\frac{f(-z_{\max}+2\Delta_z)-2f(-z_{\max}+\Delta_z)+f(-z_{\max})}{2\Delta_z}+\mathcal{O}(\Delta_z^2).\end{aligned}$$

Result (i) is obtained by isolating $f'(-z_{\max})$. The proof of (ii) is similar and therefore omitted. □

Proposition 4.14 motivates the approximations

$$\begin{aligned}&\gamma\frac{\Delta\hat{\varphi}_\alpha^{(0)}}{\Delta z}(t_{j+\sigma_\alpha},z_i)\\&:=\sigma_\alpha\gamma(\omega,z_i)\frac{-\hat{\varphi}_\alpha^{(0)}(t_{j+1},z_{i+2})+4\hat{\varphi}_\alpha^{(0)}(t_{j+1},z_{i+1})-3\hat{\varphi}_\alpha^{(0)}(t_{j+1},z_i)}{2\Delta_z}\\&\quad+(1-\sigma_\alpha)\gamma(\omega,z_i)\frac{-\hat{\varphi}_\alpha^{(0)}(t_j,z_{i+2})+4\hat{\varphi}_\alpha^{(0)}(t_j,z_{i+1})-3\hat{\varphi}_\alpha^{(0)}(t_j,z_i)}{2\Delta_z}\end{aligned}\tag{4.63}$$

when $i=-n_z$ and

$$\begin{aligned}&\gamma\frac{\Delta\hat{\varphi}_\alpha^{(0)}}{\Delta z}(t_{j+\sigma_\alpha},z_i)\\&:=\sigma_\alpha\gamma(\omega,z_i)\frac{3\hat{\varphi}_\alpha^{(0)}(t_{j+1},z_i)-4\hat{\varphi}_\alpha^{(0)}(t_{j+1},z_{i-1})+\hat{\varphi}_\alpha^{(0)}(t_{j+1},z_{i-2})}{2\Delta_z}\\&\quad+(1-\sigma_\alpha)\gamma(\omega,z_i)\frac{3\hat{\varphi}_\alpha^{(0)}(t_j,z_i)-4\hat{\varphi}_\alpha^{(0)}(t_j,z_{i-1})+\hat{\varphi}_\alpha^{(0)}(t_j,z_{i-2})}{2\Delta_z}\end{aligned}\tag{4.64}$$

when $i=n_z$. From Equation (4.45), we replace the approximations by their formulas of Equations (4.50), (4.53), (4.60), (4.63) and (4.64), and we express the values of time t_{j+1} in function of the values up to time t_j. This leads to

$$
\begin{aligned}
&\hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_{i+1}) \left[\frac{-\sigma_\alpha \Delta_t^{1-\alpha} \gamma(\omega, z_{i+\frac{1}{2}})}{2\Delta_z} \right] \\
&\quad + \hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_i) \left[\frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)} - \Delta_t^{1-\alpha} \sigma_\alpha \beta(r, \omega, z_i) + \frac{\Delta_t^{1-\alpha} \sigma_\alpha \frac{\partial \gamma}{\partial z_i}(\omega, z_i)}{2\Delta_z} \right] \\
&\quad + \hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_{i-1}) \left[\frac{\sigma_\alpha \Delta_t^{1-\alpha} \gamma(\omega, z_{i-\frac{1}{2}})}{2\Delta_z} \right] \\
&= \hat{\varphi}_\alpha^{(0)}(t_j, z_i) \left[(1-\sigma_\alpha) \Delta_t^{1-\alpha} \beta(r, \omega, z_i) + \frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)} \right] \qquad (4.65) \\
&\quad - \frac{1}{\Gamma(1-\alpha)} \sum_{s=0}^{j-1} c_{j-s}^{(\alpha,j)} \left(\hat{\varphi}_\alpha^{(0)}(t_{s+1}, z_i) - \hat{\varphi}_\alpha^{(0)}(t_s, z_i) \right) \\
&\quad + \frac{(1-\sigma_\alpha)\Delta_t^{1-\alpha}}{2\Delta_z} \left[\gamma(\omega, z_{i+\frac{1}{2}}) \hat{\varphi}_\alpha^{(0)}(t_j, z_{i+1}) - \Delta_z \frac{\partial \gamma}{\partial z_i}(\omega, z_i) \hat{\varphi}_\alpha^{(0)}(t_j, z_i) \right. \\
&\qquad\qquad \left. - \gamma(\omega, z_{i-\frac{1}{2}}) \hat{\varphi}_\alpha^{(0)}(t_j, z_{i-1}) \right]
\end{aligned}
$$

in the case $i \in \{-n_z+1, \ldots, n_z-1\}$. Otherwise, it gives

$$
\begin{aligned}
&\hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_{i+2}) \left[\frac{\sigma_\alpha \Delta_t^{1-\alpha} \gamma(\omega, z_i)}{2\Delta_z} \right] + \hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_{i+1}) \left[\frac{-2\Delta_t^{1-\alpha} \sigma_\alpha \gamma(\omega, z_i)}{\Delta_z} \right] \\
&+ \hat{\varphi}_\alpha^{(0)}(t_{j+1}, z_{i-1}) \left[\frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)} - \Delta_t^{1-\alpha} \sigma_\alpha \beta(r, \omega, z_i) + 3 \frac{\sigma_\alpha \Delta_t^{1-\alpha} \gamma(\omega, z_i)}{2\Delta_z} \right] \\
&= \hat{\varphi}_\alpha^{(0)}(t_j, z_i) \left[(1-\sigma_\alpha) \Delta_t^{1-\alpha} \beta(r, \omega, z_i) + \frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)} \right] \qquad (4.66) \\
&- \frac{1}{\Gamma(1-\alpha)} \sum_{s=0}^{j-1} c_{j-s}^{(\alpha,j)} \left(\hat{\varphi}_\alpha^{(0)}(t_{s+1}, z_i) - \hat{\varphi}_\alpha^{(0)}(t_s, z_i) \right) \\
&+ \frac{(1-\sigma_\alpha)\Delta_t^{1-\alpha} \gamma(\omega, z_i)}{2\Delta_z} \left[-\hat{\varphi}_\alpha^{(0)}(t_j, z_{i+2}) + 4\hat{\varphi}_\alpha^{(0)}(t_j, z_{i+1}) - 3\hat{\varphi}_\alpha^{(0)}(t_j, z_i) \right]
\end{aligned}
$$

when $i = -n_z$ and

$$\begin{aligned}
&\hat{\varphi}^{(0)}_{\alpha}(t_{j+1}, z_i)\left[\frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)} - \Delta_t^{1-\alpha}\sigma_\alpha\beta(r,\omega,z_i) - 3\frac{\sigma_\alpha\Delta_t^{1-\alpha}\gamma(\omega,z_i)}{2\Delta_z}\right] \\
&+ \hat{\varphi}^{(0)}_{\alpha}(t_{j+1}, z_{i-1})\left[\frac{2\Delta_t^{1-\alpha}\sigma_\alpha\gamma(\omega,z_i)}{\Delta_z}\right] \\
&+ \hat{\varphi}^{(0)}_{\alpha}(t_{j+1}, z_{i-2})\left[\frac{-\sigma_\alpha\Delta_t^{1-\alpha}\gamma(\omega,z_i)}{2\Delta_z}\right] \\
&= \hat{\varphi}^{(0)}_{\alpha}(t_j, z_i)\left[(1-\sigma_\alpha)\Delta_t^{1-\alpha}\beta(r,\omega,z_i) + \frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)}\right] \\
&- \frac{1}{\Gamma(1-\alpha)}\sum_{s=0}^{j-1} c_{j-s}^{(\alpha,j)}\left(\hat{\varphi}^{(0)}_{\alpha}(t_{s+1}, z_i) - \hat{\varphi}^{(0)}_{\alpha}(t_s, z_i)\right) \\
&+ \frac{(1-\sigma_\alpha)\Delta_t^{1-\alpha}\gamma(\omega,z_i)}{2\Delta_z}\left[3\hat{\varphi}^{(0)}_{\alpha}(t_j, z_i) - 4\hat{\varphi}^{(0)}_{\alpha}(t_j, z_{i-1}) + \hat{\varphi}^{(0)}_{\alpha}(t_j, z_{i-2})\right]
\end{aligned} \tag{4.67}$$

when $i = n_z$. Equations (4.65)-(4.67) actually describe a linear system of equations. Namely, if we define the matrix $\boldsymbol{\gamma} \in \mathbb{R}^{(2n_z+1)\times(2n_z+1)}$ as

$$\begin{pmatrix}
3\gamma_{n_z} & -4\gamma_{n_z} & \gamma_{n_z} & \cdots & 0 & 0 & 0 \\
\gamma_{n_z-1+\frac{1}{2}} & -\Delta_z\partial\gamma_{n_z-1} & -\gamma_{n_z-1-\frac{1}{2}} & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & \gamma_{-n_z+1+\frac{1}{2}} & -\Delta_z\partial\gamma_{-n_z+1} & -\gamma_{-n_z+1-\frac{1}{2}} \\
0 & 0 & 0 & \cdots & -\gamma_{-n_z} & 4\gamma_{-n_z} & -3\gamma_{-n_z}
\end{pmatrix}$$

where γ_i and $\partial\gamma_i$ are respectively used as shorthands for $\gamma(\omega, z_i)$ and $\frac{\partial\gamma}{\partial z_i}(\omega, z_i)$, and the matrix $\boldsymbol{\beta} \in \mathbb{R}^{(2n_z+1)\times(2n_z+1)}$ as

$$\boldsymbol{\beta} := \operatorname{diag}\left(\beta(r,\omega,z_{n_z}), \beta(r,\omega,z_{n_z-1}), \ldots, \beta(r,\omega,z_{-n_z})\right),$$

then we can rewrite Equations (4.65)-(4.67) in the matrix form

$$\begin{aligned}
&\left[-\sigma_\alpha\Delta_t^{1-\alpha}\left(\boldsymbol{\beta} + \frac{\boldsymbol{\gamma}}{2\Delta_z}\right) + \mathbf{I}\frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)}\right]\hat{\boldsymbol{\varphi}}^{\alpha}(t_{j+1}) \\
&= \left[(1-\sigma_\alpha)\Delta_t^{1-\alpha}\left(\boldsymbol{\beta} + \frac{\boldsymbol{\gamma}}{2\Delta_z}\right) + \mathbf{I}\frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)}\right]\hat{\boldsymbol{\varphi}}^{\alpha}(t_j) \\
&+ \frac{1}{\Gamma(1-\alpha)}\sum_{s=0}^{j-1} c_{j-s}^{(\alpha,j)}\left(\hat{\boldsymbol{\varphi}}^{\alpha}(t_{s+1}) - \hat{\boldsymbol{\varphi}}^{\alpha}(t_s)\right)
\end{aligned}$$

where $\mathbf{I} \in \mathbb{R}^{(2n_z+1)\times(2n_z+1)}$ is an identity matrix and

$$\hat{\boldsymbol{\varphi}}^{\alpha}(t_j) = (\hat{\varphi}^{(0)}_{\alpha}(t_j, z_{n_z}), \hat{\varphi}^{(0)}_{\alpha}(t_j, z_{n_z-1}), \ldots, \hat{\varphi}^{(0)}_{\alpha}(t_j, z_{-n_z}))^\top \in \mathbb{R}^{(2n_z+1)\times 1}.$$

From the initial condition $\hat{\boldsymbol{\varphi}}_\alpha(t_0) = 1$, it is thus possible to compute recursively all the $\hat{\boldsymbol{\varphi}}_\alpha(t_j)$'s by using

$$\begin{aligned}
&\hat{\boldsymbol{\varphi}}^\alpha(t_{j+1}) \\
&= \left[-\sigma_\alpha \Delta_t^{1-\alpha}\left(\beta + \frac{\gamma}{2\Delta_z}\right) + \mathbf{I}\frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)}\right]^{-1}\left[(1-\sigma_\alpha)\Delta_t^{1-\alpha}\left(\beta + \frac{\gamma}{2\Delta_z}\right) + \mathbf{I}\frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)}\right]\hat{\boldsymbol{\varphi}}^\alpha(t_j) \\
&+ \frac{1}{\Gamma(1-\alpha)}\left[-\sigma_\alpha \Delta_t^{1-\alpha}\left(\beta + \frac{\gamma}{2\Delta_z}\right) + \mathbf{I}\frac{c_0^{(\alpha,j)}}{\Gamma(1-\alpha)}\right]^{-1}\sum_{s=0}^{j-1} c_{j-s}^{(\alpha,j)}\left(\hat{\boldsymbol{\varphi}}^\alpha(t_{s+1}) - \hat{\boldsymbol{\varphi}}^\alpha(t_s)\right).
\end{aligned}$$

Inversion of the Fourier transform via FFT

Now that we are armed to compute the values of the transform φ_α, this subsection describes in details how to obtain the call prices from this transform by inverting it. As in Carr and Madan (1999), this inversion is performed numerically with the help of a Fast Fourier Transform (FFT) algorithm. For a fixed maturity $T > 0$, and an evaluation time $t \in [0, T)$, we wish to compute the call price

$$\mathbb{E}_{\mathbb{Q}}[e^{-rS_T}(e^{X_{S_T}} - e^k)_+ | \mathcal{G}_t],$$

where $(X_t)_{t\geqslant 0}$ is the logarithm of the stock price and $k = \ln K$ is the log-strike. By the Markov property, this call price can be considered as the function

$$C(T, k|t, y_1, y_2, v, u) = \mathbb{E}_{\mathbb{Q}}[e^{-rS_T}(e^{X_{S_T}} - e^k)_+ | (A_{S_t}, \lambda_{S_t}, S_t, U_{S_t}) = (y_1, y_2, v, u)].$$

Since this function is not square integrable (with respect to the log-strike k), Fourier theory does not apply. However, if we fix $\varepsilon > 1$, then the function

$$c(T, k|t, y_1, y_2, v, u) := e^{\varepsilon k}\mathbb{E}_{\mathbb{Q}}[e^{-rS_T}(e^{X_{S_T}} - e^k)_+ | (A_{S_t}, \lambda_{S_t}, S_t, U_{S_t}) = (y_1, y_2, v, u)]$$

is square integrable and its Fourier transform exists. Denote by $\hat{c}$ the Fourier transform with respect to the log-strike k,

$$\hat{c}(T, \omega|t, y_1, y_2, v, u) := \int_{\mathbb{R}} e^{\mathrm{i}\omega k} c(T, k|t, y_1, y_2, v, u)\mathrm{d}k.$$

This Fourier transform can be expressed in terms of the Fourier-Laplace transform of the couple (S_t, X_{S_t}) through the following computations

$$\begin{aligned}
\hat{c}(T,\omega|t,y_1,y_2,v,u) &= \int_{\mathbb{R}} e^{(\mathrm{i}\omega+\varepsilon)k} C(T,k|t,y_1,y_2,v,u)\mathrm{d}k \\
&= \int_{\mathbb{R}} e^{(\mathrm{i}\omega+\varepsilon)k} \int_{\mathbb{R}+} \int_k^{+\infty} e^{-rx_0} \left(e^{x_1} - e^k\right) p_\alpha(T,x_0,x_1|t,y_1,y_2,v,u)\mathrm{d}x_1\mathrm{d}x_0\mathrm{d}k \\
&= \int_{\mathbb{R}+} e^{-rx_0} \int_{\mathbb{R}} e^{x_1} p_\alpha(T,x_0,x_1|t,y_1,y_2,v,u) \int_{-\infty}^{x_1} e^{(\mathrm{i}\omega+\varepsilon)k}\mathrm{d}k\mathrm{d}x_1\mathrm{d}x_0 \\
&- \int_{\mathbb{R}+} e^{-rx_0} \int_{\mathbb{R}} p_\alpha(T,x_0,x_1|t,y_1,y_2,v,u) \int_{-\infty}^{x_1} e^{(\mathrm{i}\omega+\varepsilon+1)k}\mathrm{d}k\mathrm{d}x_1\mathrm{d}x_0 \\
&= \left(\frac{1}{\mathrm{i}\omega+\varepsilon} - \frac{1}{\mathrm{i}\omega+\varepsilon+1}\right) \int_{\mathbb{R}+} \int_{\mathbb{R}} e^{-rx_0+(\mathrm{i}\omega+\varepsilon+1)x_1} p_\alpha(T,x_0,x_1|t,y_1,y_2,v,u)\mathrm{d}x_1\mathrm{d}x_0 \\
&= \frac{1}{(\mathrm{i}\omega+\varepsilon)^2 + (\mathrm{i}\omega+\varepsilon)} \varphi_\alpha(T,r,\mathrm{i}\omega+\varepsilon+1,0|t,y_1,y_2,v,u).
\end{aligned}$$

Then we can invert the Fourier transform to obtain the call price

$$\begin{aligned}
&C(T,k|t,y_1,y_2,v,u) \\
&= \frac{e^{-\varepsilon k}}{\pi} \int_{\mathbb{R}+} e^{-\mathrm{i}\omega k} \frac{\varphi_\alpha(T,r,\mathrm{i}\omega+\varepsilon+1,0|t,y_1,y_2,v,u)}{(\mathrm{i}\omega+\varepsilon)^2+(\mathrm{i}\omega+\varepsilon)}\mathrm{d}\omega.
\end{aligned} \tag{4.68}$$

We will approximate this integral with the help of Riemann sums and perform the computations with the help of a fast Fourier transform algorithm. Following Carr and Madan (1999), the approximation for the integral of Equation (4.68) is of the form

$$\begin{aligned}
&C(T,k_n|t,y_1,y_2,v,u) \\
&\approx \frac{e^{-\varepsilon k_n}}{\pi} \mathrm{Re} \sum_{j=1}^{N} e^{\mathrm{i}\omega_j k_n} \frac{\varphi_\alpha(T,r,\mathrm{i}\omega_j+\varepsilon+1,0|t,y_1,y_2,v,u)}{(\mathrm{i}\omega_j+\varepsilon)^2+(\mathrm{i}\omega_j+\varepsilon)} \left[\frac{3+(-1)^j-\zeta_{j-1}}{3}\right] \Delta_\omega
\end{aligned}$$

for the range of log-strikes $k_n = -k_{\max} + \Delta_k(n-1)$, $k_{\max} > 0$, $n = 1, \ldots, N \in \mathbb{N}$ and $\Delta_k = 2k_{\max}/N$. The discretization of the variable ω is of the form $\omega_j = \Delta_\omega(j-1)$ for some $\Delta_\omega > 0$ and $\zeta_{j-1} = 0$ if $j = 1$ and $\zeta_{j-1} = 0$ when $j \neq 1$. Imposing $\Delta_k\Delta_\omega = 2\pi/N$ and performing some computations yields

$$\begin{aligned}
&C(T,k_n|t,y_1,y_2,v,u) \approx \Delta_\omega \frac{e^{-\varepsilon k_n}}{\pi} \\
&\mathrm{Re} \sum_{j=1}^{N} \exp\left\{-\frac{2\mathrm{i}\pi(j-1)(n-1)}{N}\right\} \frac{\varphi_\alpha(T,r,\mathrm{i}\omega_j+\varepsilon+1,0|\cdot)}{(\mathrm{i}\omega_j+\varepsilon)^2+(\mathrm{i}\omega_j+\varepsilon)} (-1)^{j-1} \left[\frac{3+(-1)^j-\zeta_{j-1}}{3}\right],
\end{aligned}$$

where $\varphi_\alpha(T,r,\mathrm{i}\omega_j+\varepsilon+1,0|t,y_1,y_2,v,u) = \varphi_\alpha(T,r,\mathrm{i}\omega_j+\varepsilon+1,0|\cdot)$. This can be computed with a fast Fourier transform algorithm.

4.6 Hedging of a call option

In this section, we discuss the hedging strategy for contingent claims, and in particular for call options. A first strategy could be to "Delta" hedge the call option, that is buying a quantity $\partial C/\partial A_{S_t}$ of the underlying asset $(A_{S_t})_{t\geqslant 0}$ to mimic the first order behavior of the call option price with respect to the asset. However, such an approach fails to properly take into account the jumps that occur in the price. To better take into account the occurrences of jumps, we propose a hedging strategy that is determined through a variance minimization problem, as pioneered in Föllmer and Sondermann (1985). This approach is also referred to as quadratic hedging and is used in numerous studies. We can cite for example Föllmer and Schweizer (1991), Pham (2000) and Moraux and Hainaut (2018).

The variance minimization of the hedging error will be performed under a fixed risk-neutral measure $\mathbb{Q}$, as described in Section 4.4. The choice of such a measure is questionable. As a matter of fact, this approach suffers from two drawbacks. The first is that the chosen risk-neutral measure $\mathbb{Q}$ becomes an input of the optimization procedure that leads to the hedging strategy. Consequently, the hedging strategy provides neither a risk-neutral measure nor a price for the contingent claim. The second drawback is that the profits and losses that will occur in practice are ruled by the real-world probability measure. From this point of view, the optimization should be performed under the real probability measure and not under a risk-neutral one. However, hedging under a risk-neutral measure has some benefits. On the one hand, quadratic hedging with discontinuous processes under probability measures that are not risk-neutral does not admit a solution in general. On the other hand, the non-uniqueness of a risk-neutral measure allows to adjust the parameters that appear in the optimization problem. In particular, the parameters could then be adjusted to reflect the uncertainty over the evolution of prices and the risk aversion of traders. As noted in Hansen and Sargent (1995) and Hansen and Sargent (2001), the risk-neutral measure can be chosen so that the obtained hedging strategy is more conservative than any other strategy built under $\mathbb{P}$.

The quadratic hedging approach relies on predictable representation theorems for martingales. Such theorems allow to write contingent claims as stochastic integrals of predictable processes with respect to the risk factors that drive the underlying asset price. The first representation theorem is often called the Kunita-Watanabe decomposition theorem and was introduced in Kunita and Watanabe (1967). Similar results can be found in Kunita (2004) and in Chapter 3 of Jacod and Shiryaev (2003). We start with the general case of the hedging of a square integrable contingent claims Y. More precisely, we establish the existence and uniqueness the quadratic hedging strategy under each fixed risk-neutral measure $\mathbb{Q}$. Afterwards, we determine the particular hedging strategy when the contingent claim Y is the payoff of a call option.

In order to rely on the Kunita-Watanabe decomposition theorem for square integrable martingales, we work with the discounted asset prices. Recall that in the subdiffusive case, the discounted asset price is $(\tilde{A}_{S_t})_{t\geqslant 0}$, where $\tilde{A}_t = e^{-rt}A_t, t \geqslant 0$. Fix a risk-neutral measure $\mathbb{Q}$ as in Section 4.4. Note that when talking about square integrable martingales, we refer to the terminology of Protter (2005) (P.180), i.e. an $(\mathcal{F}_t)_{t\geqslant 0}$-adapted stochastic process $(M_t)_{t\geqslant 0}$ that satisfies $\mathbb{E}[M_t^2] < +\infty$ and $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$ for all $t \geqslant s \geqslant 0$. In particular, we do not necessarily mean that $\sup_{t\geqslant 0}\mathbb{E}[M_t^2] < +\infty$ (the processes that satisfy this additional condition are referred to as L^2 martingales in Protter (2005) and are of course square integrable martingales).

Let $T > 0$ be a deterministic maturity and Y be an $\mathcal{F}_T^{A\circ S}$-measurable positive and square integrable random variable that represents the payoff of a contingent claim to be paid at time T. We aim at finding a portfolio (or strategy) in order to hedge the contingent claim Y. A portfolio consists of a vector valued predictable stochastic process $(\pi_t^{(0)}, \pi_t^{(1)}, \ldots, \pi_t^{(d)})_{t\in[0,T]}$ where d is the number of risky assets in the market and $\pi_t^{(i)}$ is the quantity of assets i to be hold at time t. Asset $i = 0$ is the risk-free asset. In our case, $d = 1$ and the present value $(V_t(\boldsymbol{\pi}))_{t\in[0,T]}$ of the portfolio $\boldsymbol{\pi} = (\pi_t^{(0)}, \pi_t^{(1)})_{t\in[0,T]}$ is given by

$$V_t(\boldsymbol{\pi}) = \pi_t^{(0)} + \pi_t^{(1)}\tilde{A}_{S_t} \tag{4.69}$$

at time t. On top of being $(\mathcal{F}_t^{A\circ S})_{t\in[0,T]}$-predictable, we require the portfolio $\boldsymbol{\pi}$ to be self-financing. This concept refers to the fact that the instantaneous changes in the value $(V_t(\boldsymbol{\pi}))_{t\in[0,T]}$ are due to changes in the prices of the assets, and not to instantaneous rebalancing of the asset quantities $(\pi_t^{(0)}, \pi_t^{(1)})_{t\in[0,T]}$ we hold. Mathematically, this condition is expressed as

$$V_t(\boldsymbol{\pi}) = V_0(\boldsymbol{\pi}) + \int_0^t \pi_u^{(1)}\mathrm{d}\tilde{A}_{S_u}.$$

For more material about self-financing portfolios, we refer to the chapter 2 of Jeanblanc et al. (2009). The goal is thus to find

$$\boldsymbol{\pi}^\star = \arg\min_{\boldsymbol{\pi}\in\boldsymbol{\Pi}} \mathbb{E}_{\mathbb{Q}}\left[(V_T(\boldsymbol{\pi}) - e^{-rS_T}Y)^2\right] \tag{4.70}$$

where $\boldsymbol{\Pi}$ is the set of all self-financing portfolios. The next result requires the notion of orthogonal martingales, given at Definition 4.1. Let us state the Kunita-Watanabe decomposition theorem.

Lemma 4.9. *Let $(B_t)_{t\geqslant 0}, (M_t)_{t\geqslant 0}$ be two square integrable $(\mathbb{Q}, \mathbb{G})$-martingales, where $\mathbb{G} = (\mathcal{G}_t)_{t\geqslant 0}$ is a filtration that satisfies the usual conditions. There exists a unique choice of square integrable $(\mathbb{Q}, \mathbb{G})$-martingales $(M_t')_{t\geqslant 0}$ and $(M_t'')_{t\geqslant 0}$ such that* *(i)* $(M'_t)_{t\geqslant 0} \in \mathbf{\Phi}(B)$, *where*

$$\mathbf{\Phi}(B) := \left\{ \left(\int_0^t \phi_u \mathrm{d}B_u \right)_{t\geqslant 0} : (\phi_u)_{u\geqslant 0} \text{ is a } \mathbb{G}\text{-predictable process} \right\},$$

(ii) $(M''_t)_{t\geqslant 0}$ *is orthogonal to any process that belongs to* $\mathbf{\Phi}(B)$,

(iii) $(M_t)_{t\geqslant 0}$ *and* $(M'_t + M''_t)_{t\geqslant 0}$ *are indistinguishable.*

The statement and the proof can be found at Proposition 4.1 in Kunita and Watanabe (1967). In order to apply Lemma 4.9, we need the filtration $(\mathcal{G}_t)_{t\geqslant 0}$ to satisfy what are called the usual conditions. The first usual condition is that $\mathcal{G}_0$ must contain all the null sets of the probability space, which is here the case by assumption. The second is that the σ-algebras $\mathcal{G}_{t+} := \bigcap_{\varepsilon>0} \mathcal{G}_{t+\varepsilon}$ and $\mathcal{G}_t$ should coincide for any $t \geqslant 0$. This property is called the right-continuity of the filtration. The two next lemmas show that the filtration $\mathbb{F}^{A\circ S}$ that we have chosen for the discounted asset price $(\tilde{A}_{S_t})_{t\geqslant 0}$ is right-continuous.

Lemma 4.10. *Let* $(\mathcal{H}_t)_{t\geqslant 0}$ *and* $(\mathcal{F}_t)_{t\geqslant 0}$ *be two right-continuous filtrations on the same probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*. Then the filtration* $(\mathcal{H}_t \vee \mathcal{F}_t)_{t\geqslant 0}$ *is right-continuous.*

Proof. Fix $t \geqslant 0$. We begin by showing the inclusion

$$\bigcap_{\varepsilon>0} (\mathcal{H}_{t+\varepsilon} \cup \mathcal{F}_{t+\varepsilon}) \subset \mathcal{H}_t \cup \mathcal{F}_t. \tag{4.71}$$

Note that this inclusion is trivial in the case $\mathcal{H}_t \cup \mathcal{F}_t = \mathcal{F}$, so that we assume $\mathcal{H}_t \cup \mathcal{F}_t \neq \mathcal{F}$, that is $\mathcal{H}_t \cup \mathcal{F}_t$ is strictly included in $\mathcal{F}$. In this proof, set complements are taken with respect to $\mathcal{F}$, i.e. for $\mathcal{A} \subset \mathcal{F}$, $\overline{\mathcal{A}} = \mathcal{F} \setminus \mathcal{A}$. Let $B \in \overline{\mathcal{H}_t \cup \mathcal{F}_t}$. By the right-continuity of the filtrations and two applications of De Morgan's laws, we have

$$\begin{aligned} B &\in \overline{\left(\bigcap_{\varepsilon>0} \mathcal{H}_{t+\varepsilon} \right) \cup \left(\bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} \right)} \\ &= \overline{\left(\bigcap_{\varepsilon>0} \mathcal{H}_{t+\varepsilon} \right)} \cap \overline{\left(\bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon} \right)} \\ &= \left(\bigcup_{\varepsilon>0} \overline{\mathcal{H}_{t+\varepsilon}} \right) \cap \left(\bigcup_{\varepsilon>0} \overline{\mathcal{F}_{t+\varepsilon}} \right) \end{aligned}$$

so that there are $\varepsilon_{\mathcal{H}}, \varepsilon_{\mathcal{F}} > 0$ such that $B \in \overline{\mathcal{H}_{t+\varepsilon_{\mathcal{H}}}} \cap \overline{\mathcal{F}_{t+\varepsilon_{\mathcal{F}}}}$. Setting $\varepsilon = \varepsilon_{\mathcal{H}} \wedge \varepsilon_{\mathcal{F}}$, we have that $B \in \overline{\mathcal{H}_{t+\varepsilon}} \cap \overline{\mathcal{F}_{t+\varepsilon}}$, which implies

$$B \in \overline{\bigcap_{\varepsilon>0} (\mathcal{H}_{t+\varepsilon} \cup \mathcal{F}_{t+\varepsilon})}$$

and thereby proves the inclusion at Equation (4.71). As a consequence,

$$\sigma\left(\bigcap_{\varepsilon>0}(\mathcal{H}_{t+\varepsilon}\cup\mathcal{F}_{t+\varepsilon})\right)\subset\mathcal{H}_t\vee\mathcal{F}_t.$$

It remains to show that

$$\bigcap_{\varepsilon>0}(\mathcal{H}_{t+\varepsilon}\vee\mathcal{F}_{t+\varepsilon})=\sigma\left(\bigcap_{\varepsilon>0}(\mathcal{H}_{t+\varepsilon}\cup\mathcal{F}_{t+\varepsilon})\right).$$

To this end, note that

$B\in\bigcap_{\varepsilon>0}(\mathcal{H}_{t+\varepsilon}\vee\mathcal{F}_{t+\varepsilon})$

$\Leftrightarrow$ For any $\varepsilon>0$, $B\in\mathcal{S}$ whenever $\mathcal{S}$ is a σ-algebra that satisfies $\mathcal{S}\supset\mathcal{H}_{t+\varepsilon}\cup\mathcal{F}_{t+\varepsilon}$

$\Leftrightarrow$ For any $\varepsilon>0$, $B\in\mathcal{S}$ whenever $\mathcal{S}$ is a σ-algebra that satisfies $\mathcal{S}\supset\mathcal{H}_{t+\varepsilon}\vee\mathcal{F}_{t+\varepsilon}$

$\Leftrightarrow$ For any $\varepsilon>0$, $B\in\mathcal{H}_{t+\varepsilon}\vee\mathcal{F}_{t+\varepsilon}$

$\Leftrightarrow$ $B\in\bigcap_{\varepsilon>0}(\mathcal{H}_{t+\varepsilon}\vee\mathcal{F}_{t+\varepsilon})$,

concluding the proof. □

Lemma 4.11. *The filtration* $\mathbb{F}^{A\circ S}$ *is right-continuous.*

Proof. Recall that $\mathcal{F}_t^{A\circ S}=\mathcal{F}_{S_t}^{A,U}$, where $\mathcal{F}_t^{A,U}=\mathcal{F}_t^{A}\vee\mathcal{F}_t^{U}=\mathcal{F}_t^{W}\vee\mathcal{F}_t^{J}\vee\mathcal{F}_t^{U}$.

First we show that $\mathbb{F}^{A,U}$ is right-continuous. According to Lemma 4.10, it is enough to show that $\mathbb{F}^{W}$, $\mathbb{F}^{J}$ and $\mathbb{F}^{U}$ are right-continuous. The right-continuity of both $\mathbb{F}^{U}$ and $\mathbb{F}^{W}$ is guaranteed by Theorem 31 in the chapter 1 of Protter (2005), which states that the natural filtrations of Lévy processes completed with the null sets are right-continuous. The proof of the right-continuity of natural filtrations of counting processes (Theorem 25 in the same chapter) is also valid for the right-continuity of $\mathbb{F}^{J}$. It follows that $\mathbb{F}^{A,U}$ is right-continuous. It remains to show that this conclusion extends to the time-changed filtration $\mathbb{F}^{A\circ S}$.

Fix $t\geqslant 0$ and assume that $A\in\bigcap_{n\in\mathbb{N}}\mathcal{F}^{A\circ S}_{t+\frac{1}{n}}$. Then for each $n\in\mathbb{N}$ and all $s\geqslant 0$, $A\cap\{S_{t+\frac{1}{n}}\leqslant s\}\in\mathcal{F}_s^{A,U}$. It implies that for any $k\in\mathbb{N}$,

$$\bigcap_{m\geqslant k}\bigcup_{n\in\mathbb{N}}A\cap\left\{S_{t+\frac{1}{n}}\leqslant s+\frac{1}{m}\right\}\in\mathcal{F}^{A,U}_{s+\frac{1}{k}},\tag{4.72}$$

and thus the right-continuity of $\mathbb{F}^{A,U}$ entails that the set of Equation (4.72) is in $\mathcal{F}_s^{A,U}$. Finally, since the paths of $(S_t)_{t\geqslant 0}$ are nondecreasing and right-continuous, the set of Equation (4.72) equals $A\cap\{S_t\leqslant s\}$. This proves that $A\in\mathcal{F}_{S_t}^{A,U}=\mathcal{F}_t^{A\circ S}$. □

We can now state the existence and uniqueness of the quadratic hedging strategy.

Proposition 4.15. *Let Y be a $\mathcal{F}_T^{A\circ S}$- measurable square integrable random variable. Then there exists a unique choice of a $\mathbb{F}^{A\circ S}$-predictable process $(y_t)_{t\geqslant 0}$ and a martingale $(M_t)_{t\geqslant 0}$ such that*

$$e^{-rS_T}Y = \int_0^T y_s \mathrm{d}\tilde{A}_{S_s} + M_T \tag{4.73}$$

with $(M_t)_{t\geqslant 0}$ being orthogonal to any process in $\mathbf{\Phi}(\tilde{A}\circ S)$. Moreover, there exists a unique solution to the quadratic hedging problem (4.70) *given by*

$$\boldsymbol{\pi}^\star = (\mathbb{E}_{\mathbb{Q}}[e^{-rS_T}Y] - y_0\tilde{A}_{S_0}, y_t)_{t\in[0,T]}.$$

Proof. Define the process $(Y_t)_{t\geqslant 0}$ as $Y_t := \mathbb{E}_{\mathbb{Q}}[e^{-rS_T}Y|\mathcal{F}_t^{A\circ S}]$. It is then clear that $(Y_t)_{t\geqslant 0}$ is a square integrable $(\mathbb{Q},\mathbb{F}^{A\circ S})$-martingale. Lemma 4.9 entails the existence of a unique choice of a $\mathbb{F}^{A\circ S}$-predictable process $(y_t)_{t\geqslant 0}$ and a square integrable $(\mathbb{Q},\mathbb{F}^{A\circ S})$-martingale $(M_t)_{t\geqslant 0}$ such that

$$Y_t = \int_0^t y_s \mathrm{d}\tilde{A}_{S_s} + M_t.$$

Moreover, the stochastic integral $(\int_0^t y_s \mathrm{d}\tilde{A}_{S_s})_{t\geqslant 0}$ is a square integrable $(\mathbb{Q},\mathbb{F}^{A\circ S})$-martingale. Equation (4.73) follows. The orthogonality of $(M_t)_{t\geqslant 0}$ to any process in $\mathbf{\Phi}(\tilde{A}\circ S)$ corresponds to the second statement of Lemma 4.9. Then, for any self-financing portfolio $\boldsymbol{\pi} = (\pi_t^{(0)}, \pi_t^{(1)})$, we can write

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}[(V_T(\boldsymbol{\pi}) - e^{-rS_T}Y)^2] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\left(V_0(\boldsymbol{\pi}) + \int_0^T (\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u} - M_T\right)^2\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[(V_0(\boldsymbol{\pi}) - M_T)^2\right] + \mathbb{E}_{\mathbb{Q}}\left[\left(\int_0^T (\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u}\right)^2\right] \\
&\quad - 2\mathbb{E}_{\mathbb{Q}}\left[M_T \int_0^T (\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u}\right].
\end{aligned} \tag{4.74}$$

Note that $\left(\int_0^t (\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u}\right)_{t\geqslant 0} \in \mathbf{\Phi}(\tilde{A}\circ S)$ and thus this martingale is orthogonal to $(M_t)_{t\geqslant 0}$. It implies that

$$\mathbb{E}_{\mathbb{Q}}\left[M_T \int_0^T (\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u}\right] = 0.$$

Therefore, with some more straightforward computations, Equation (4.74) becomes

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[(V_T(\boldsymbol{\pi}) - e^{-rS_T}Y)^2] =&(V_0(\boldsymbol{\pi}) - M_0)^2 + \mathbb{E}_{\mathbb{Q}}[(M_T - M_0)^2] \\ &+ \mathbb{E}_{\mathbb{Q}}\left[\left(\int_0^T (\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u}\right)^2\right].\end{aligned}$$

To this point, it is already clear that the optimal initial value for the hedging strategy is $V_0(\boldsymbol{\pi}) = M_0 = \mathbb{E}_{\mathbb{Q}}[e^{-rS_T}Y]$. It is also clear that the optimal $(\pi_t^{(1)})_{t\geqslant 0}$ is determined through the minimization of

$$\mathbb{E}_{\mathbb{Q}}\left[\left(\int_0^T (\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u}\right)^2\right].$$

Moreover,

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}&\left[\left(\int_0^T (\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u}\right)^2\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[\left[(\pi^{(1)} - y)\cdot(\tilde{A}\circ S), (\pi^{(1)} - y)\cdot(\tilde{A}\circ S)\right]_T\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[\int_0^T \left(\pi_t^{(1)} - y_t\right)^2 \mathrm{d}\left[\tilde{A}\circ S, \tilde{A}\circ S\right]_t\right]\end{aligned} \tag{4.75}$$

where $((\pi^{(1)} - y)\cdot(\tilde{A}\circ S))_t$ denotes the stochastic integral $\int_0^t(\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u}$. The first and second equalities in Equation (4.75) respectively follow from Corollary 3 and Theorem 29 in Protter (2005), since $\left(\int_0^t(\pi_u^{(1)} - y_u)\mathrm{d}\tilde{A}_{S_u}\right)_{t\geqslant 0}$ is a martingale. Quadratic variations being increasing processes, the quantity at Equation (4.75) is minimized when $\pi_t^{(1)} = y_t$ for all $t \geqslant 0$. The result follows by combining our findings with Equation (4.69). □

Unfortunately, Proposition 4.15 does not give an explicit way to compute the optimal strategy $\boldsymbol{\pi}^\star$. The reason is that the Kunita-Watanabe decomposition theorem does not give an explicit formula for the predictable process $(y_t)_{t\geqslant 0}$ but only asserts its existence and uniqueness. However, for any square-integrable contingent claim whose payoff depends only on the terminal value $\tilde{A}_{S_T}$ of the stock, one can use Ito's lemma to compute explicitly the solution of the quadratic hedging problem. The remainder of this section is thus dedicated to the explicit quadratic hedging portfolio when $Y = f(A_{S_T})$. By the Markov property, there exists a function g such that the $(\mathcal{F}_t^{A\circ S})_{t\geqslant 0}$-martingale $(Y_t)_{t\geqslant 0} = (\mathbb{E}_{\mathbb{Q}}[e^{-rS_T}f(A_{S_T})|\mathcal{F}_t^{A\circ S}])_{t\geqslant 0}$ satisfies

$$\mathbb{E}_{\mathbb{Q}}[e^{-rS_T}f(A_{S_T})|\mathcal{F}_t^{A\circ S}] = g(t, \tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}), \tag{4.76}$$

or

$$g(t, y_1, y_2, v, u) = \mathbb{E}[e^{-rS_T} f(A_{S_T}) | (\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}) = (y_1, y_2, v, u)]. \qquad (4.77)$$

We now show that the argument t is actually irrelevant in the function g. The intuition behind this result is that in the time-changed setting, the relevant time is given by the last observed value of $(U_{S_t})_{t\geqslant 0}$, ruling out the original time scale $(t)_{t\geqslant 0}$. To prove this, we rely on two intermediary results.

Lemma 4.12. *For all $t \geqslant 0$, it holds that $S_{U_t} = t$ a.s.. Moreover $S \circ U = \mathrm{id}_{\mathbb{R}^+}$ a.s., that is the processes $((S \circ U)_t)_{t\geqslant 0}$ and $(t)_{t\geqslant 0}$ are indistinguishable.*

Proof. The first part of the statement follows readily from

$$S_{U_t(\omega)}(\omega) = \inf\{\tau > 0 : U_\tau \geqslant U_t(\omega)\}$$

and the fact that the paths of $(U_t)_{t\geqslant 0}$ are strictly increasing a.s.. The second part of the statement is a consequence of the càdlàg paths of $((S \circ U)_t)_{t\geqslant 0}$, combined with the first statement. $\square$

Lemma 4.13. *For each $u \geqslant s \geqslant 0$, we have*

$$\{S_s = v\} \cap \{U_{S_s} = u\} = \{S_u = v\} \cap \{U_{S_s} = u\}. \qquad (4.78)$$

As a consequence,

$$\{S_s = v\} \cap \{U_{S_s} = u\} = \{S_u = v\} \cap \{U_{S_s} = u\} \cap \{U_{S_u} = u\}. \qquad (4.79)$$

Proof. Let ω be a member of the set at the left-hand side of Equation (4.78). Since $U_v(\omega) = u$, it is clear that

$$S_u(\omega) := \inf\{\tau > 0 : U_\tau(\omega) \geqslant u\} \leqslant v.$$

Assume by contradiction that $\inf\{\tau > 0 : U_\tau(\omega) \geqslant u\} < v$. Then there exists an $\varepsilon > 0$ such that $U_{v-\varepsilon}(\omega) \geqslant u$. Hence, as $u \geqslant s$, this leads to $v - \varepsilon \geqslant S_s(\omega) = v$, which is a contradiction. We conclude that $S_u(\omega) = v$.

Conversely, let ω be a member of the set at the right-hand side of Equation (4.78). The equality $U_{S_s(\omega)}(\omega) = u$ implies that

$$S_u(\omega) = (S \circ U \circ S)_s(\omega) = v.$$

Moreover, Lemma 4.12 says that $(S \circ U \circ S)_s(\omega) = S_s(\omega)$ a.s., which proves that $S_s(\omega) = v$.

That the set at the left-hand side of Equation (4.79) contains the set at the right-hand side is an obvious consequence of Equation (4.78) that we have just established. Let $\omega \in \{S_s = v\} \cap \{U_{S_s} = u\}$. By Equation (4.78), $S_s(\omega) = S_u(\omega) = v$, so that $U_{S_u(\omega)}(\omega) = U_v(\omega) = u$. $\square$

We are now able to prove the irrelevancy of the t argument in the function g of Equation (4.77).

Proposition 4.16. *Let g be the function of Equation (4.77). Then for any $t \in [0, T]$, $u \geqslant t$, $v \geqslant 0$, we have*

$$g(t, y_1, y_2, v, u) = g(u \wedge T, y_1, y_2, v, u)$$

so that g does not depend on t.

Proof. By Lemma 4.13,

$$\begin{aligned} g(t, y_1, y_2, v, u) &= \mathbb{E}_{\mathbb{Q}}[e^{-rS_T} f(A_{S_T})|(\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}) = (y_1, y_2, v, u)] \\ &= \mathbb{E}_{\mathbb{Q}}\left[e^{-rS_T} f(A_{S_T})|(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_s}, U_{S_u}) = (y_1, y_2, v, u, u)\right]. \end{aligned}$$

If $u \leqslant T$, the Markov property of $(S_t, U_{S_t})_{t \geqslant 0}$ implies that

$$\begin{aligned} &\mathbb{E}_{\mathbb{Q}}\left[e^{-rS_T} f(A_{S_T})|(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_s}, U_{S_u}) = (y_1, y_2, v, u, u)\right] \\ &= \mathbb{E}_{\mathbb{Q}}\left[e^{-rS_T} f(A_{S_T})|(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_u}) = (y_1, y_2, v, u)\right] \\ &= g(u, y_1, y_2, v, u) \end{aligned}$$

whereas if $u > T$, one has

$$\begin{aligned} &\mathbb{E}_{\mathbb{Q}}\left[e^{-rS_T} f(A_{S_T})|(A_{S_u}, \lambda_{S_u}, S_u, U_{S_s}, U_{S_u}) = (y_1, y_2, v, u, u)\right] \\ &= e^{-rS_T} f(A_{S_T}) \\ &= g(T, y_1, y_2, v, u). \end{aligned}$$

Indeed, in that case, since $U_{S_s} = U_{S_u} = u$ and $s \leqslant T < u$, we have $v = S_s \leqslant S_T \leqslant S_u = v$. $\square$

The previous result motivates that in the following, the function g will be defined as

$$g(y_1, y_2, v, u) = \mathbb{E}_{\mathbb{Q}}[e^{-rS_T} f(A_{S_T})|(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_u}) = (y_1, y_2, v, u)].$$

Assuming that g has continuous second order partial derivatives, we apply Ito's lemma for semimartingales on the process $(g(\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}))_{t \geqslant 0}$. To do so, note that all the processes that are involved are indeed semimartingales. As a matter of fact, since $(\tilde{A}_t)_{t \geqslant 0}$ and $(\lambda_t)_{t \geqslant 0}$ are semimartingales, Theorem 10.16 in Jacod (1979) implies that their time-changed counterparts are also semimartingales. Moreover, since $(S_t)_{t \geqslant 0}$ and $(U_{S_t})_{t \geqslant 0}$ have finite variation (because they are nondecreasing) and are càdlàg, Theorem 26 in the second chapter of Protter (2005) entails that they are quadratic pure jump semimartingales. In order to use Ito's lemma, we need to determine the quadratic (co)variations of (or between) the processes that are involved in the function g. We begin with the proof that the quadratic covariation between $(U_t)_{t \geqslant 0}$ and $(A_t)_{t \geqslant 0}$ or $(\lambda_t)_{t \geqslant 0}$ is zero. The proof consists in showing that these processes never jump together.

Proposition 4.17. *The quadratic covariation processes* $([A,U]_t)_{t\geqslant 0}$ *and* $([\lambda,U]_t)_{t\geqslant 0}$ *are indistinguishable from the null process.*

Proof. We give the proof for $([A,U]_t)_{t\geqslant 0}$ only, as the case of $([\lambda,U]_t)_{t\geqslant 0}$ is essentially the same.

The proof rely on the assumption regarding the structure of the probability space $(\Omega,\mathcal{F},\mathbb{P})$ explained in the introduction of the model. Since $(U_t)_{t\geqslant 0}$ is a quadratic pure jump semimartingale[7], it holds that

$$[A,U]_t = \sum_{0<s\leqslant t} \Delta A_s \Delta U_s.$$

Let us define $\mathcal{J}_t^n(\omega) = \{s \in (0,t] : \Delta U_s(\omega) > n^{-1}\}$ and $\mathcal{J}_t(\omega) = \bigcup_{n\geqslant 1} \mathcal{J}_t^n(\omega) = \{s \in (0,t] : \Delta U_s(\omega) > 0\}$. Since $(U_t)_{t\geqslant 0}$ is almost surely strictly increasing and $U_t(\omega) < +\infty$ for almost all ω, the set $\mathcal{J}_t^n(\omega)$ is almost surely finite for all $n \geqslant 0$ and $t > 0$ (which also implies that the set of jumps $\mathcal{J}_t(\omega)$ is countable, as it is a countable union of finite sets). Similarly, we define $\tilde{\mathcal{J}}_t^n(\omega_2) = \{s \in (0,t] : \Delta U_s^{(2)}(\omega_2) > n^{-1}\}$ and $\tilde{\mathcal{J}}_t(\omega_2) = \bigcup_{n\geqslant 1} \tilde{\mathcal{J}}_t^n(\omega_2)$. Note that for all n, $\tilde{\mathcal{J}}_t^n(\omega_2) = \mathcal{J}_t^n(\omega_1,\omega_2)$. From the monotone convergence theorem and Tonelli's theorem, it follows that

$$\begin{aligned}
\mathbb{E}[|[A,U]_t|] &\leqslant \mathbb{E}\left[\sum_{0<s\leqslant t} |\Delta A_s| \Delta U_s\right] \\
&= \int_\Omega \sum_{s\in\mathcal{J}_t(\omega)} |\Delta A_s(\omega)| \Delta U_s(\omega) \mathbb{P}(\mathrm{d}\omega) \\
&= \lim_{n\to+\infty} \int_\Omega \sum_{s\in\mathcal{J}_t^n(\omega)} |\Delta A_s(\omega)| \Delta U_s(\omega) \mathbb{P}(\mathrm{d}\omega) \\
&= \lim_{n\to+\infty} \int_{\Omega_2}\int_{\Omega_1} \sum_{s\in\mathcal{J}_t^n(\omega_1,\omega_2)} |\Delta A_s(\omega_1,\omega_2)| \Delta U_s(\omega_1,\omega_2) \mathbb{P}^{(1)}(\mathrm{d}\omega_1)\mathbb{P}^{(2)}(\mathrm{d}\omega_2) \\
&= \lim_{n\to+\infty} \int_{\Omega_2}\int_{\Omega_1} \sum_{s\in\tilde{\mathcal{J}}_t^n(\omega_2)} |\Delta A_s^{(1)}(\omega_1)| \Delta U_s^{(2)}(\omega_2) \mathbb{P}^{(1)}(\mathrm{d}\omega_1)\mathbb{P}^{(2)}(\mathrm{d}\omega_2) \\
&= \lim_{n\to+\infty} \int_{\Omega_2} \sum_{s\in\tilde{\mathcal{J}}_t^n(\omega_2)} \Delta U_s^{(2)}(\omega_2) \left(\int_{\Omega_1} |\Delta A_s^{(1)}(\omega_1)| \mathbb{P}^{(1)}(\mathrm{d}\omega_1)\right) \mathbb{P}^{(2)}(\mathrm{d}\omega_2) \\
&= \lim_{n\to+\infty} \int_{\Omega_2} \sum_{s\in\tilde{\mathcal{J}}_t^n(\omega_2)} \Delta U_s^{(2)}(\omega_2) [\psi(1,0)-1]\mathbb{E}[A_{s-}]\mathbb{E}[\Delta N_s] \mathbb{P}^{(2)}(\mathrm{d}\omega_2) \\
&= 0,
\end{aligned}$$

[7] Again, this process is càdlàg and since it is strictly increasing, it has finite variation. Theorem 26 in the second chapter of Protter (2005) therefore implies that it is a quadratic pure jump semimartingale.

where the last equality is a consequence of $\Delta N_s = 0$ $\mathbb{P}$-a.s., for all $s \geqslant 0$. We conclude that for all $t \geqslant 0$, $[A,U]_t = 0$ a.s., that is $([A,U]_t)_{t\geqslant 0}$ and the null process are modifications. The conclusion then follows from the fact that $([A,U]_t)_{t\geqslant 0}$ is a càdlàg process (which is a consequence of Theorem 22 in Protter (2005), combined with the polarization identity for quadratic variations). □

Corollary 4.5. *The quadratic covariation process* $([A \circ S, U \circ S]_t)_{t\geqslant 0}$ *and* $([\lambda \circ S, U \circ S]_t)_{t\geqslant 0}$ *are indistinguishable from the null process.*

Corollary 4.6. *The processes* $(|\Delta(A\circ S)_t \Delta(U\circ S)_t|)_{t\geqslant 0}$ *and* $(|\Delta(\lambda\circ S)_t \Delta(U\circ S)_t|)_{t\geqslant 0}$ *are indistinguishable from the null process.*

Proof. This follows from the proof of Proposition 4.17. It is shown that the process $(\sum_{0<s\leqslant t} |\Delta A_s|\Delta U_s)_{t\geqslant 0}$ is indistinguishable from the zero process. The result then follows from the inequalities

$$0 \leqslant |\Delta(A \circ S)_t \Delta(U \circ S)_t| \leqslant \sum_{0<s\leqslant S_t} |\Delta A_s| \Delta U_s$$

and

$$0 \leqslant |\Delta(\lambda \circ S)_t \Delta(U \circ S)_t| \leqslant \sum_{0<s\leqslant S_t} |\Delta \lambda_s| \Delta U_s.$$

□

Proposition 4.18. *For any semimartingale* $(X_t)_{t\geqslant 0}$*, the quadratic covariation process* $([X,S]_t)_{t\geqslant 0}$ *is indistinguishable from the null process.*

Proof. The quadratic covariation processes being càdlàg, it is enough to show that they are modifications of the null process. As observed above, the process $(S_t)_{t\geqslant 0}$ is a quadratic pure jump semimartingale. As a consequence, Theorem 28 of Protter (2005) (Chapter 2) yields

$$[X,S]_t = X_0 S_0 + \sum_{0<s\leqslant t} \Delta X_s \Delta S_s,$$

which is 0 a.s., since $(S_t)_{t\geqslant 0}$ is continuous and $S_0 = 0$. □

Proposition 4.19. *Assume that the function g of Equation* (4.76) *has continuous second order partial derivatives. Then, we have that*

$$
\begin{aligned}
g(\tilde{A}_{S_T}, \lambda_{S_T}, S_T, U_{S_T}) &= g(\tilde{A}_0, \lambda_0, 0, 0) + \sigma \int_0^T \tilde{A}_{S_{t-}} \frac{\partial g}{\partial y_1} \mathrm{d}W_{S_t} \\
&+ \int_0^T \left[\frac{\partial g}{\partial v} + \kappa(\theta - \lambda_{S_{t-}}) \frac{\partial g}{\partial y_2} - \tilde{A}_{S_{t-}} \mathbb{E}[e^{\xi} - 1] \lambda_{S_{t-}} \frac{\partial g}{\partial y_1} + \frac{\sigma^2}{2} \tilde{A}^2_{S_{t-}} \frac{\partial^2 g}{\partial y_1^2} \right] \mathrm{d}S_t \\
&+ \int_0^T \int_{z \in \mathbb{R}} (g(e^z \tilde{A}_{S_{t-}}, \lambda_{S_{t-}} + \eta|z|, S_t, U_{S_t}) - g(\tilde{A}_{S_{t-}}, \lambda_{S_{t-}}, S_t, U_{S_t})) \Xi(\mathrm{d}z, \mathrm{d}S_t) \\
&+ \sum_{0 < t \leqslant T} (g(\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}) - g(\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_{t-}})).
\end{aligned}
$$

Proof. An application of Ito's lemma for semimartingales yields

$$
\begin{aligned}
& g(t, \tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}) - g(\tilde{A}_{S_0}, \lambda_{S_0}, S_0, U_{S_0}) \\
&= \int_0^t \frac{\partial g}{\partial y_1} (\tilde{A}_{S_{s-}}, \lambda_{S_{s-}}, S_{s-}, U_{S_{s-}}) \tilde{A}_{S_{s-}} \left[\sigma \mathrm{d}W_{S_s} + \lambda_{S_{s-}} \mathbb{E}[e^{\xi} - 1] \mathrm{d}S_s \right] \\
&\quad + \int_0^t \frac{\partial g}{\partial y_2} (\tilde{A}_{S_{s-}}, \lambda_{S_{s-}}, S_{s-}, U_{S_{s-}}) \kappa(\theta - \lambda_{S_{s-}}) \mathrm{d}S_s \\
&\quad + \int_0^t \frac{\partial g}{\partial v} (\tilde{A}_{S_{s-}}, \lambda_{S_{s-}}, S_{s-}, U_{S_{s-}}) \mathrm{d}S_s \\
&\quad + \frac{\sigma^2}{2} \int_0^t \frac{\partial^2 g}{\partial y_1^2} (\tilde{A}_{S_{s-}}, \lambda_{S_{s-}}, S_{s-}, U_{S_{s-}}) \tilde{A}^2_{S_{s-}} \mathrm{d}S_s \\
&\quad + \sum_{0 < s \leqslant t} \left[g(\tilde{A}_{S_s}, \lambda_{S_s}, S_s, U_{S_s}) - g(\tilde{A}_{S_{s-}}, \lambda_{S_{s-}}, S_{s-}, U_{S_-}) \right].
\end{aligned}
\tag{4.80}
$$

The jump part can be decomposed as follows

$$
\begin{aligned}
& \sum_{0 < s \leqslant t} \left[g(\tilde{A}_{S_s}, \lambda_{S_s}, S_s, U_{S_s}) - g(\tilde{A}_{S_{s-}}, \lambda_{S_{s-}}, S_{s-}, U_{S_{s-}}) \right] \\
&= \sum_{0 < s \leqslant t} \left[g(\tilde{A}_{S_s}, \lambda_{S_s}, S_s, U_{S_s}) - \lim_{u \uparrow s} g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_u}) \right] \mathbf{1}_{\{\Delta(U \circ S)_s = 0\} \cap \{\Delta(\tilde{A} \circ S)_s > 0\}} \\
&+ \sum_{0 < s \leqslant t} \left[g(\tilde{A}_{S_s}, \lambda_{S_s}, S_s, U_{S_s}) - \lim_{u \uparrow s} g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_u}) \right] \mathbf{1}_{\{\Delta(U \circ S)_s > 0\} \cap \{\Delta(\tilde{A} \circ S)_s = 0\}} \\
&+ \sum_{0 < s \leqslant t} \left[g(\tilde{A}_{S_s}, \lambda_{S_s}, S_s, U_{S_s}) - \lim_{u \uparrow s} g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_u}) \right] \mathbf{1}_{\{\Delta(U \circ S)_s > 0\} \cap \{\Delta(\tilde{A} \circ S)_s > 0\}}
\end{aligned}
$$

and Corollary 4.6 implies that the last term is indistinguishable from the zero process. The continuity of g is then used to obtain

$$\begin{aligned}
&\sum_{0<s\leqslant t}\left[g(\tilde{A}_{S_s},\lambda_{S_s},S_s,U_{S_s})-g(\tilde{A}_{S_{s-}},\lambda_{S_{s-}},S_{s-},U_{S_{s-}})\right]\\
&=\sum_{0<s\leqslant t}\left[g(\tilde{A}_{S_s},\lambda_{S_s},S_s,U_{S_s})-\lim_{u\uparrow s}g(\tilde{A}_{S_u},\lambda_{S_u},S_u,U_{S_u})\right]\mathbf{1}_{\{\Delta(U\circ S)_s=0\}\cap\{\Delta(\tilde{A}\circ S)_s>0\}}\\
&\quad+\sum_{0<s\leqslant t}\left[g(\tilde{A}_{S_s},\lambda_{S_s},S_s,U_{S_s})-\lim_{u\uparrow s}g(\tilde{A}_{S_u},\lambda_{S_u},S_u,U_{S_u})\right]\mathbf{1}_{\{\Delta(U\circ S)_s>0\}\cap\{\Delta(\tilde{A}\circ S)_s=0\}}\\
&=\sum_{0<s\leqslant t}\left[g(\tilde{A}_{S_s},\lambda_{S_s},S_s,U_{S_s})-g(\lim_{u\uparrow s}(\tilde{A}_{S_u},\lambda_{S_u},S_u,U_{S_u}))\right]\mathbf{1}_{\{\Delta(U\circ S)_s=0\}\cap\{\Delta(\tilde{A}\circ S)_s>0\}}\\
&\quad+\sum_{0<s\leqslant t}\left[g(\tilde{A}_{S_s},\lambda_{S_s},S_s,U_{S_s})-g(\lim_{u\uparrow s}(\tilde{A}_{S_u},\lambda_{S_u},S_u,U_{S_u}))\right]\mathbf{1}_{\{\Delta(U\circ S)_s>0\}\cap\{\Delta(\tilde{A}\circ S)_s=0\}}\\
&=\sum_{0<s\leqslant t}\left[g(\tilde{A}_{S_s},\lambda_{S_s},S_s,U_{S_s})-g(\tilde{A}_{S_{s-}},\lambda_{S_{s-}},S_s,U_{S_s})\right]\mathbf{1}_{\{\Delta(U\circ S)_s=0\}\cap\{\Delta(\tilde{A}\circ S)_s>0\}}\\
&\quad+\sum_{0<s\leqslant t}\left[g(\tilde{A}_{S_s},\lambda_{S_s},S_s,U_{S_s})-g(\tilde{A}_{S_s},\lambda_{S_s},S_s,U_{S_{s-}})\right]\mathbf{1}_{\{\Delta(U\circ S)_s>0\}\cap\{\Delta(\tilde{A}\circ S)_s=0\}}.
\end{aligned}$$

The result is finally obtained by rearranging the terms in Equation (4.80) and rewriting the first jumps term with the help of the random measure Ξ, which gives

$$\begin{aligned}
&\sum_{0<s\leqslant t}\left[g(\tilde{A}_{S_s},\lambda_{S_s},S_s,U_{S_s})-g(\tilde{A}_{S_{s-}},\lambda_{S_{s-}},S_s,U_{S_s})\right]\mathbf{1}_{\{\Delta(U\circ S)_s=0\}\cap\{\Delta(\tilde{A}\circ S)_s>0\}}\\
&=\int_0^T\int_{z\in\mathbb{R}}(g(e^z\tilde{A}_{S_{t-}},\lambda_{S_{t-}}+\eta|z|,S_t,U_{S_t})-g(\tilde{A}_{S_{t-}},\lambda_{S_{t-}},S_t,U_{S_t}))\Xi(\mathrm{d}z,\mathrm{d}S_t)
\end{aligned}$$

and concludes the proof. □

Using Proposition 4.19, we can derive the optimal hedging strategy $\boldsymbol{\pi}^\star$ of Equation (4.70) when $Y_T=f(A_{S_T})$. This strategy is given in the next proposition.

Proposition 4.20. *Assume that the function g of Equation (4.76) has continuous second order partial derivatives. Then if $Y=f(A_{S_T})$ is square integrable, the hedging problem (4.70) admits the solution $\boldsymbol{\pi}^\star=(\pi_t^{(0\star)},\pi_t^{(1\star)})_{t\geqslant 0}$, where*

$$\begin{aligned}
\pi_u^{(1\star)}&=\frac{\sigma^2\frac{\partial g}{\partial y_1}}{\left[\sigma^2+\lambda_{S_{u-}}\int_{\mathbb{R}}(e^z-1)^2\nu(\mathrm{d}z)\right]}\\
&+\frac{\lambda_{S_{u-}}\int_{\mathbb{R}}[g(e^z\tilde{A}_{S_{u-}},\lambda_{S_{u-}}+\eta|z|,S_u,U_{S_u})-g(\tilde{A}_{S_{u-}},\lambda_{S_{u-}},S_u,U_{S_u})](e^z-1)\nu(\mathrm{d}z)}{\tilde{A}_{S_{u-}}\left[\sigma^2+\lambda_{S_{u-}}\int_{\mathbb{R}}(e^z-1)^2\nu(\mathrm{d}z)\right]}
\end{aligned}$$

and

$$\pi_u^{(0\star)}=\mathbb{E}_{\mathbb{Q}}[e^{-rS_T}f(A_{S_T})]-\pi_0^{(1\star)}\tilde{A}_{S_0}.$$

Proof. Recall that the function g and the $\mathbb{F}^{A\circ S}$-martingale $(Y_t)_{t\geqslant 0} = (\mathbb{E}_{\mathbb{Q}}[e^{-rS_T}f(A_{S_T})|\mathcal{F}_t^{A\circ S}])_{t\geqslant 0}$ satisfy

$$\mathbb{E}_{\mathbb{Q}}[e^{-rS_T}f(A_{S_T})|\mathcal{F}_t^{A\circ S}] = g(\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}). \tag{4.81}$$

For any self-financing portfolio $\boldsymbol{\pi} = (\pi_t^{(0)}, \pi_t^{(1)})_{t\in[0,T]}$, we have the following decomposition of the quadratic hedging error

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\left[\left(Y_T - V_0(\boldsymbol{\pi}) - \int_0^T \pi_t^{(1)}\mathrm{d}\tilde{A}_{S_t}\right)^2\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\left(Y_T - Y_0 + (Y_0 - V_0(\boldsymbol{\pi})) - \int_0^T \pi_t^{(1)}\mathrm{d}\tilde{A}_{S_t}\right)^2\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\begin{array}{r}(Y_0 - V_0(\boldsymbol{\pi}))^2 + \left(Y_T - Y_0 - \displaystyle\int_0^T \pi_t^{(1)}\mathrm{d}\tilde{A}_{S_t}\right)^2 \\ + 2(Y_0 - V_0(\boldsymbol{\pi}))\left(Y_T - Y_0 - \displaystyle\int_0^T \pi_t^{(1)}\mathrm{d}\tilde{A}_{S_t}\right)\end{array}\right] \\
&= (Y_0 - V_0(\boldsymbol{\pi}))^2 + \mathbb{E}_{\mathbb{Q}}\left[\left(Y_T - Y_0 - \int_0^T \pi_t^{(1)}\mathrm{d}\tilde{A}_{S_t}\right)^2\right],
\end{aligned}$$

where the last equality comes from

$$\mathbb{E}_{\mathbb{Q}}[Y_T - Y_0] = Y_0 - Y_0 = 0$$

and

$$\mathbb{E}_{\mathbb{Q}}\left[\int_0^T \pi_t^{(1)}\mathrm{d}\tilde{A}_{S_t}\right] = 0,$$

as $(Y_t)_{t\geqslant 0}$ and $(\tilde{A}_{S_t})_{t\geqslant 0}$ are $\mathbb{F}^{A\circ S}$-martingales. It is then clear that the optimal initial amount is $V_0(\boldsymbol{\pi}^\star) = Y_0 = \mathbb{E}_{\mathbb{Q}}[e^{-rS_T}f(A_{S_T})]$. Next, Proposition 4.19 implies that

$$\begin{aligned}
&Y_T - Y_0 - \int_0^T \pi_t^{(1)}\mathrm{d}\tilde{A}_{S_t} = \sigma\int_0^T \tilde{A}_{S_{t-}}\left(\frac{\partial g}{\partial y_1} - \pi_t^{(1)}\right)\mathrm{d}W_{S_t} \\
&+ \int_0^T \left[\frac{\partial g}{\partial v} + \kappa(\theta - \lambda_{S_{t-}})\frac{\partial g}{\partial y_2} - \tilde{A}_{S_{t-}}\mathbb{E}_{\mathbb{Q}}[e^{\xi} - 1]\lambda_{S_{t-}}\left(\frac{\partial g}{\partial y_1} - \pi_t^{(1)}\right) + \frac{\sigma^2}{2}\tilde{A}_{S_{t-}}^2\frac{\partial^2 g}{\partial y_1^2}\right]\mathrm{d}S_t \\
&+ \int_0^T\int_{z\in\mathbb{R}}\left(\begin{array}{l}g(e^z\tilde{A}_{S_{t-}}, \lambda_{S_{t-}} + \eta|z|, S_t, U_{S_t}) \\ \qquad - g(\tilde{A}_{S_{t-}}, \lambda_{S_{t-}}, S_t, U_{S_t}) - (e^z - 1)\tilde{A}_{S_{t-}}\pi_t^{(1)}\end{array}\right)\Xi(\mathrm{d}z, \mathrm{d}S_t) \\
&+ \sum_{0<t\leqslant T}\left(g(\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}) - g(\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_{t-}})\right)
\end{aligned}$$

In order to shorten the notations, we define

$$H_t^{(1)} := \sigma \tilde{A}_{S_{t-}} \left(\frac{\partial g}{\partial y_1}(\tilde{A}_{S_{t-}}, \lambda_{S_{t-}}, S_{t-}, U_{S_{t-}}) - \pi_t^{(1)} \right),$$

$$\begin{aligned}
H_t^{(2)} := &\frac{\partial g}{\partial v}(\tilde{A}_{S_{t-}}, \lambda_{S_{t-}}, S_{t-}, U_{S_{t-}}) + \kappa(\theta - \lambda_{S_{t-}}) \frac{\partial g}{\partial y_2}(\tilde{A}_{S_{t-}}, \lambda_{S_{t-}}, S_{t-}, U_{S_{t-}}) \\
&- \tilde{A}_{S_{t-}} \mathbb{E}_{\mathbb{Q}}[e^{\xi} - 1] \lambda_{S_{t-}} \left(\frac{\partial g}{\partial y_1}(\tilde{A}_{S_{t-}}, \lambda_{S_{t-}}, S_{t-}, U_{S_{t-}}) - \pi_t^{(1)} \right) \\
&+ \frac{\sigma^2}{2} \tilde{A}_{S_{t-}}^2 \frac{\partial^2 g}{\partial y_1^2}(\tilde{A}_{S_{t-}}, \lambda_{S_{t-}}, S_{t-}, U_{S_{t-}}), \\
X_t^{(3)} := &\int_0^t \int_{z \in \mathbb{R}} \Big(g(e^z \tilde{A}_{S_{u-}}, \lambda_{S_{u-}} + \eta |z|, S_u, U_{S_u}) - g(\tilde{A}_{S_{u-}}, \lambda_{S_{u-}}, S_u, U_{S_u}) \\
&\qquad - (e^z - 1) \tilde{A}_{S_{u-}} \pi_u^{(1)} \Big) \Xi(\mathrm{d}z, \mathrm{d}S_u)
\end{aligned}$$

and

$$X_t^{(4)} := \sum_{0 < u \leqslant t} (g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_u}) - g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_{u-}})).$$

Moreover we denote

$$X_t^{(1)} := \int_0^t H_u^{(1)} \mathrm{d}W_{S_u} \qquad X_t^{(2)} := \int_0^t H_u^{(2)} \mathrm{d}S_u$$

and $(H_t)_{t \in [0,T]} := (Y_t - Y_0 - \int_0^t \pi_u^{(1)} \mathrm{d}\tilde{A}_{S_u})_{t \geqslant 0}$, which is a martingale. Using the notations we introduced, we find

$$\begin{aligned}
H_T &= Y_T - Y_0 - \int_0^T \pi_t^{(1)} \mathrm{d}\tilde{A}_{S_t} \\
&= \int_0^T H_t^{(1)} \mathrm{d}W_{S_t} + \int_0^T H_t^{(2)} \mathrm{d}S_t + X_T^{(3)} + X_T^{(4)} \qquad (4.82) \\
&= \sum_{k=1}^4 X_T^{(k)}.
\end{aligned}$$

Since $(H_t)_{t \in [0,T]}$ is a martingale, we have

$$\mathbb{E}_{\mathbb{Q}} \left[\left(Y_T - Y_0 - \int_0^T \pi_t^{(1)} \mathrm{d}\tilde{A}_{S_t} \right)^2 \right] = \mathbb{E}_{\mathbb{Q}}[H_T^2] = \mathbb{E}_{\mathbb{Q}}[[H, H]_T],$$

where the second equality follows from Corollary 3 in Protter (2005) (Chapter 2). Combining Equation (4.82) and the bilinearity of quadratic covariations, it follows that

$$[H,H]_t = \sum_{k=1}^{4}\sum_{j=1}^{4}[X^{(k)},X^{(j)}]_t.$$

Note that the process $(X^{(k)}_t)_{t\in[0,T]}$ is continuous when $k \in \{1,2\}$ and quadratic pure jump when $k \neq 1$. Indeed, Theorem 29 in Protter (2005) (Chapter 2) implies that

$$[X^{(2)},X^{(2)}]_t = \int_0^t \left(H^{(2)}_{u-}\right)^2 \mathrm{d}[S,S]_u$$

which is indistinguishable from the null process, as implied by Proposition 4.18. It follows that if $k \neq 1$ and $j \in \{1,2,3,4\}$ then

$$[X^{(k)},X^{(j)}]_t = [X^{(j)},X^{(k)}]_t = \sum_{0<u\leqslant t} \Delta X^{(k)}_u \Delta X^{(j)}_u.$$

Since $(X^{(k)}_t)_{t\in[0,T]}$ is continuous for $k \in \{1,2\}$, we infer that $([X^{(k)},X^{(j)}]_t)_{t\in[0,T]}$ is indistinguishable from the null process whenever $k \in \{1,2\}$ and $j \neq 1$. In addition, Theorem 29 in the second chapter of Protter (2005) also entails that

$$[X^{(1)},X^{(1)}]_t = \int_0^t \left(H^{(1)}_u\right)^2 \mathrm{d}[W\circ S, W\circ S]_u = \int_0^t \left(H^{(1)}_u\right)^2 \mathrm{d}S_u.$$

Next we have

$$\begin{aligned}
[X^{(3)},X^{(3)}]_t = \int_0^t \int_{z\in\mathbb{R}} \Big(& g(e^z \tilde{A}_{S_{u-}}, \lambda_{S_{u-}} + \eta|z|, S_u, U_{S_u}) \\
& - g(\tilde{A}_{S_{u-}}, \lambda_{S_{u-}}, S_u, U_{S_u}) - (e^z-1)\tilde{A}_{S_{u-}}\pi^{(1)}_u \Big)^2 \Xi(\mathrm{d}z, \mathrm{d}S_u),
\end{aligned}$$

$$[X^{(4)},X^{(4)}]_t = \sum_{0<u\leqslant t} (g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_u}) - g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_{u-}}))^2$$

and

$$[X^{(3)},X^{(4)}]_t = [X^{(4)},X^{(3)}]_t = 0,$$

which follows from Corollary 4.6. As a consequence,

$$\begin{aligned}
[H,H]_T = & \int_0^t \left(H^{(2)}_{u-}\right)^2 \mathrm{d}S_u \\
& + \sum_{0<u\leqslant t} (g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_u}) - g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_{u-}}))^2 \\
& + \int_0^t \int_{z\in\mathbb{R}} \Big(g(e^z \tilde{A}_{S_{u-}}, \lambda_{S_{u-}} + \eta|z|, S_u, U_{S_u}) - g(\tilde{A}_{S_{u-}}, \lambda_{S_{u-}}, S_u, U_{S_u}) \\
& \qquad - (e^z-1)\tilde{A}_{S_{u-}}\pi^{(1)}_u \Big)^2 \Xi(\mathrm{d}z, \mathrm{d}S_u).
\end{aligned}$$

and thus the strategy $(\pi_t^{(1)})_{t\in[0,T]}$ will be chosen so as to minimize the quantity

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}[H_T^2] = \mathbb{E}_{\mathbb{Q}}&\left[\int_0^T \left(\sigma \tilde{A}_{S_{u-}}\left(\frac{\partial g}{\partial y_1}(\tilde{A}_{S_{u-}}, \lambda_{S_{u-}}, S_{u-}, U_{S_{u-}}) - \pi_u^{(1)}\right)\right)^2 \mathrm{d}S_u\right] \\
+ \mathbb{E}_{\mathbb{Q}}&\left[\int_0^T \int_{z\in\mathbb{R}} \Big(g(e^z\tilde{A}_{S_{u-}}, \lambda_{S_{u-}} + \eta|z|, S_u, U_{S_u}) - g(\tilde{A}_{S_{u-}}, \lambda_{S_{u-}}, S_u, U_{S_u})\right. \\
&\qquad\left. - (e^z-1)\tilde{A}_{S_{u-}}\pi_u^{(1)}\Big)^2 \nu(\mathrm{d}z)\lambda_{S_u}\mathrm{d}S_u\right] \\
+ \mathbb{E}_{\mathbb{Q}}&\left[\sum_{0<u\leqslant T}(g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_u}) - g(\tilde{A}_{S_u}, \lambda_{S_u}, S_u, U_{S_{u-}}))^2\right].
\end{aligned}$$

Since the third term does not depend on $\pi_u^{(1)}$ and $(S_t)_{t\geqslant 0}$ is nondecreasing (thus $\mathrm{d}S_u \geqslant 0$), the problem amounts to finding

$$\begin{aligned}
\pi_u^{(1\star)} = \arg\min_{\pi_u^{(1)}} &\left\{\left(\sigma \tilde{A}_{S_{u-}}\left(\frac{\partial g}{\partial y_1}(\tilde{A}_{S_{u-}}, \lambda_{S_{u-}}, S_{u-}, U_{S_{u-}}) - \pi_u^{(1)}\right)\right)^2\right. \\
&+ \int_{z\in\mathbb{R}} \Big(g(e^z\tilde{A}_{S_{u-}}, \lambda_{S_{u-}} + \eta|z|, S_u, U_{S_u}) - g(\tilde{A}_{S_{u-}}, \lambda_{S_{u-}}, S_u, U_{S_u}) \\
&\qquad\left. - (e^z-1)\tilde{A}_{S_{u-}}\pi_u^{(1)}\Big)^2 \nu(\mathrm{d}z)\lambda_{S_u}\right\}.
\end{aligned}$$

It is then easy to check the first and second derivatives with respect to $\pi_u^{(1)}$ to conclude that

$$\begin{aligned}
\pi_u^{(1\star)} = &\frac{\sigma^2 \frac{\partial g}{\partial y_1}}{\left[\sigma^2 + \lambda_{S_{u-}}\int_{\mathbb{R}}(e^z-1)^2\nu(\mathrm{d}z)\right]} \\
&+ \frac{\lambda_{S_{u-}}\int_{\mathbb{R}}(g(e^z\tilde{A}_{S_{u-}}, \lambda_{S_{u-}} + \eta|z|, S_u, U_{S_u}) - g(\tilde{A}_{S_{u-}}, \lambda_{S_{u-}}, S_u, U_{S_u})(e^z-1)\nu(\mathrm{d}z)}{\tilde{A}_{S_{u-}}\left[\sigma^2 + \lambda_{S_{u-}}\int_{\mathbb{R}}(e^z-1)^2\nu(\mathrm{d}z)\right]},
\end{aligned}$$

as announced. □

The optimal quadratic hedging strategy for a call option comes as a corollary.

Corollary 4.7. *The call price process* $(Y_t)_{t\in[0,T]} = (\tilde{C}(t, \tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}))_{t\in[0,T]}$

is a martingale that satisfies the following SDE

$$\begin{aligned}\tilde{C}(\tilde{A}_{S_T}, \lambda_{S_T}, S_T, U_{S_T}) &= \tilde{C}(\tilde{A}_0, \lambda_0, 0, 0) + \sigma \int_0^T \tilde{A}_{S_{t-}} \frac{\partial \tilde{C}}{\partial y_1} \mathrm{d}W_{S_t} \\ &+ \int_0^T \left[\frac{\partial \tilde{C}}{\partial v} + \kappa(\theta - \lambda_{S_{t-}}) \frac{\partial \tilde{C}}{\partial y_2} - \tilde{A}_{S_{t-}} \mathbb{E}_{\mathbb{Q}}[e^{\xi} - 1] \lambda_{S_t} \frac{\partial \tilde{C}}{\partial y_1} + \frac{\sigma^2}{2} \tilde{A}^2_{S_{t-}} \frac{\partial^2 \tilde{C}}{\partial y_1^2} \right] \mathrm{d}S_t \\ &+ \int_0^T \int_{z \in \mathbb{R}} (\tilde{C}(e^z \tilde{A}_{S_{t-}}, \lambda_{S_{t-}} + \eta |z|, S_t, U_{S_t}) - \tilde{C}(\tilde{A}_{S_{t-}}, \lambda_{S_{t-}}, S_t, U_{S_t})) \Xi(\mathrm{d}z, \mathrm{d}S_t) \\ &+ \sum_{0 < t \leqslant T} (\tilde{C}(\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_t}) - \tilde{C}(\tilde{A}_{S_t}, \lambda_{S_t}, S_t, U_{S_{t-}})).\end{aligned}$$

Moreover, in the case of a call option, the hedging problem (4.70) *admits the solution* $\boldsymbol{\pi}^\star = (\pi_t^{(0\star)}, \pi_t^{(1\star)})_{t \geqslant 0}$*, where*

$$\begin{aligned}\pi_u^{(1\star)} &= \frac{\sigma^2 \frac{\partial \tilde{C}}{\partial y_1}}{\left[\sigma^2 + \lambda_{S_{u-}} \int_{\mathbb{R}} (e^z - 1)^2 \nu(\mathrm{d}z)\right]} \\ &+ \frac{\lambda_{S_{u-}} \int_{\mathbb{R}} (\tilde{C}(e^z \tilde{A}_{S_{u-}}, \lambda_{S_{u-}} + \eta |z|, S_u, U_{S_u}) - \tilde{C}(\tilde{A}_{S_{u-}}, \lambda_{S_{u-}}, S_u, U_{S_u})(e^z - 1) \nu(\mathrm{d}z)}{\tilde{A}_{S_{u-}} \left[\sigma^2 + \lambda_{S_{u-}} \int_{\mathbb{R}} (e^z - 1)^2 \nu(\mathrm{d}z)\right]}\end{aligned}$$

and

$$\pi_u^{(0\star)} = \mathbb{E}_{\mathbb{Q}}[e^{-rS_T}(A_{S_T} - K)_+] - \pi_0^{(1\star)} \tilde{A}_{S_0}.$$

We close this section with a theoretical result about stopping-times.

Lemma 4.14. *For any a.s. finite* $(\mathcal{F}^U_{S_t})$*-stopping-time* τ*, the random variables* U_{S_τ} *and* $U_{S_{\tau-}}$ *are* $(\mathcal{F}^U_{S_t})_{t \geqslant 0}$*-stopping times.*

Proof. To prove that U_{S_τ} is a $(\mathcal{F}^U_{S_t})_{t \geqslant 0}$-stopping time, we need to show that for any $t \geqslant 0$, $\{U_{S_\tau} \leqslant t\} \in \mathcal{F}^U_{S_t}$. We have that

$$\{U_{S_\tau} \leqslant t\} = (\{U_{S_\tau} \leqslant t\} \cap \{\tau \leqslant t\}) \cup \underbrace{(\{U_{S_\tau} \leqslant t\} \cap \{\tau > t\})}_{=\emptyset} \tag{4.83}$$

where the emptiness of the second set follows from the inequality $U_{S_\tau} \geqslant \tau$. The right-continuous paths of $(U_{S_t})_{t \geqslant 0}$ imply that it is a progressively measurable process, from what it follows that U_{S_τ} is $\mathcal{F}^U_{S_\tau}$-measurable. Note that the two assertions of the previous sentence are respectively Propositions 4.8 and 4.9 in the chapter 1 of Revuz and Yor (2005). The $\mathcal{F}^U_{S_\tau}$-measurability of U_{S_τ} precisely means that $\{U_{S_\tau} \in B\} \cap \{\tau \leqslant t\} \in \mathcal{F}^U_{S_t}$ for all $t \geqslant 0$ and any Borel set $B \in \mathcal{B}_{\mathbb{R}}$. It is then obvious that Equation (4.83) implies that $\{U_{S_\tau} \leqslant t\} \in \mathcal{F}^U_{S_t}$. We conclude that U_{S_τ} is a $(\mathcal{F}^U_{S_t})_{t \geqslant 0}$-stopping time.

Let us prove now that $U_{S_{\tau-}}$ is a $(\mathcal{F}^U_{S_t})_{t\geqslant 0}$-stopping time. One has

$$\begin{aligned}\{U_{S_{\tau-}} \leqslant t\} &= \big(\{U_{S_{\tau-}} \leqslant t\} \cap \{\tau = 0\}\big) \cup \big(\{U_{S_{\tau-}} \leqslant t\} \cap \{\tau > 0\}\big)\\ &= \{\tau = 0\} \cup \big(\{U_{S_{\tau-}} \leqslant t\} \cap \{\tau > 0\}\big)\,.\end{aligned}$$

As $\{\tau = 0\}$ is in $\mathcal{F}^U_{S_t}$ for any $t \geqslant 0$, it remains to show that it is also the case for $\{U_{S_{\tau-}} \leqslant t\} \cap \{\tau > 0\}$. To this end, we will establish that for any $k \in \mathbb{N}$,

$$\{U_{S_{\tau-}} \leqslant t\} \cap \{\tau > 0\} = \bigcap_{\substack{n\in\mathbb{N}\\ n\geqslant k}} \bigcup_{q\in\mathbb{Q}\cap(0,t]} \left\{q < \tau \leqslant q + \frac{1}{n}\right\} \cap \{U_{S_q} \leqslant t\}. \tag{4.84}$$

Let ω be a member of the set at the left-hand side of Equation (4.84) and fix $n \in \mathbb{N}$. Let $q \in ([\tau(\omega) - \frac{1}{2n}] \vee 0, \tau(\omega)) \cap \mathbb{Q}$. Then $\omega \in \big\{q < \tau \leqslant q + \frac{1}{n}\big\}$. Moreover, $q < \tau(\omega)$ implies that $U_{S_q}(\omega) \leqslant U_{S_{\tau-}}(\omega) \leqslant t$ and thus $\omega \in \{U_{S_q} \leqslant t\}$. Finally, note that $U_{S_q}(\omega) \geqslant q$ entails the inequality $q \leqslant t$. This proves that ω is also a member of the second set.

Let ω be a member of the set at the right-hand side of Equation (4.84). Then for each $n \geqslant k$, we have a $q_n \in (0, t] \cap \mathbb{Q}$ that satisfies $q_n < \tau(\omega) \leqslant q_n + \frac{1}{n}$ and $U_{S_{q_n}}(\omega) \leqslant t$. Up to a subsequence, we obtain that $q_n \uparrow \tau(\omega)$. Therefore,

$$\lim_{n\to+\infty} U_{S_{q_n}}(\omega) = U_{S_{\tau-}}(\omega) \leqslant t.$$

The second inclusion follows.

It remains to observe that Equation (4.84) establishes that $\{U_{S_{\tau-}} \leqslant t\} \cap \{\tau > 0\} \in \mathcal{F}^U_{S_{t+\frac{1}{k}}}$ for any $k \in \mathbb{N}$. The $\mathcal{F}^U_{S_t}$-measurability of this set is then a consequence of the right-continuity of $(\mathcal{F}^U_{S_t})_{t\geqslant 0}$ (see Lemma 4.11). □

4.7 Numerical results

This section presents the results of our numerical experiments. These results consist of call prices and associated implied volatilities computed with the help of the method described in Section 4.5. The values of the parameters are from Moraux and Hainaut (2018) and are reported in Table 4.1.

Figure 4.2 displays the prices of call options for various strikes, fractional orders α and maturities from one to six months. Note that the case $\alpha = 1$ corresponds to the non-fractional case. All the call prices used in this section are reported in Tables 4.2 and 4.3. We observe that a lower fractional order leads to higher call prices for the considered maturities. This observation is reflected on Figure 4.3, which presents the implied volatilities computed with those call prices. As a matter of fact,

Table 4.1: Parameters used for the computations.

θ	ρ_+	ρ_-	p	η	κ	σ	λ_0	r	S_0
6.44	30.47	-33.9	0.37	337.08	14.71	0.12	5.62	0.01	100

the implied volatilities are always higher for smaller fractional orders. Figure 4.4 presents the corresponding implied volatility surfaces. From these last two figures, we notice that the fractional model allows to introduce a special feature: the implied volatilities are especially high for the lowest maturities. This seems to indicate that such models correspond to a situation in which most of the movements of the price of the stock occur very shortly. A fractional model thus seems appropriate for use in a period of turbulent financial markets.

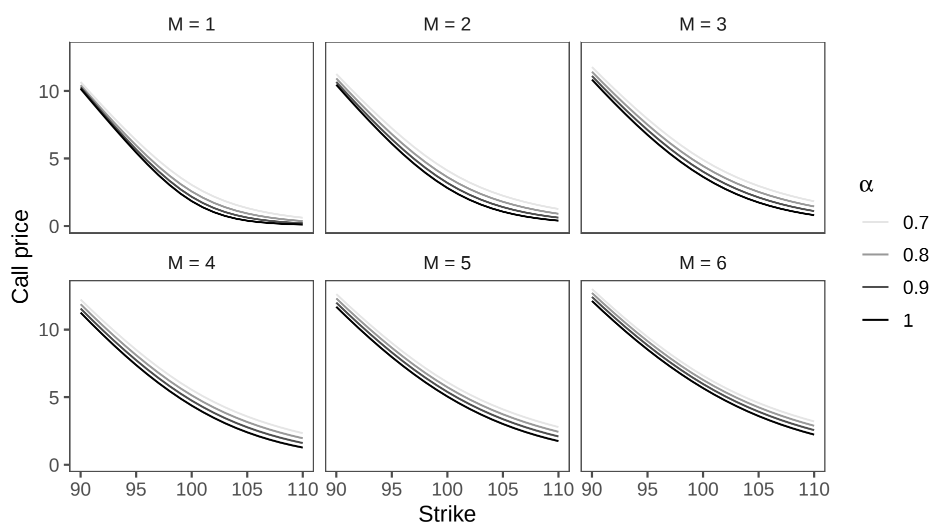

Figure 4.2: Comparison of the call option prices for various fractional orders α and for the non-fractional case ($\alpha = 1$). M corresponds to the maturity expressed in months.

We compare the results of the model proposed in this chapter with the well-known model of Heston (1993). In the Heston model, the stock price is assumed to satisfy, under the risk-neutral measure, the following SDE

$$\frac{\mathrm{d}A_t}{A_t} = r\mathrm{d}t + \sqrt{v_t}\mathrm{d}W_t \tag{4.85}$$

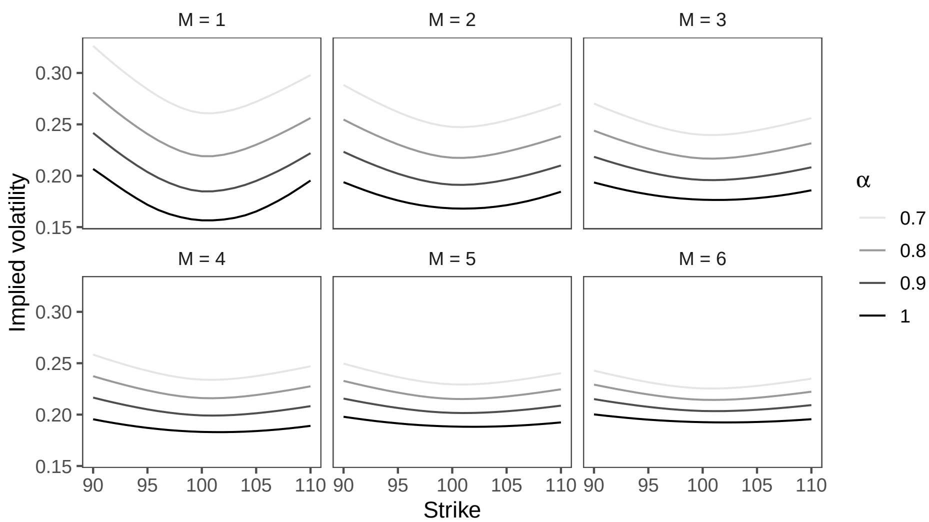

Figure 4.3: Comparison of the implied volatility smiles for various fractional orders α and for the non-fractional case ($\alpha = 1$). M corresponds to the maturity expressed in months.

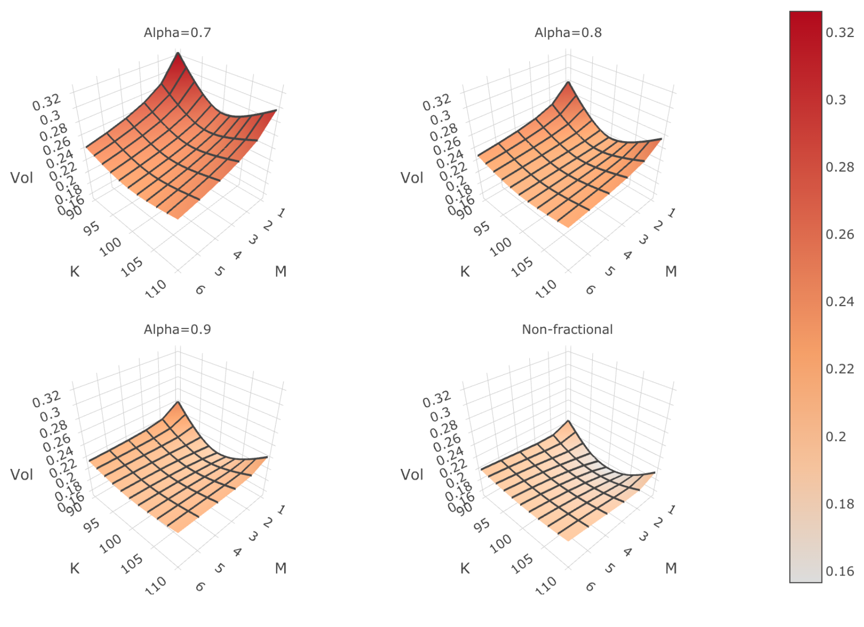

Figure 4.4: Comparison of the implied volatilities surfaces for various fractional orders α and for the non-fractional case. The M axis corresponds to the maturity expressed in months and the K axis corresponds to the strikes.

Table 4.2: Call prices for strikes from 90 to 110.

	$\alpha = 0.7$						$\alpha = 0.8$					
	$M = 1$	$M = 2$	$M = 3$	$M = 4$	$M = 5$	$M = 6$	$M = 1$	$M = 2$	$M = 3$	$M = 4$	$M = 5$	$M = 6$
90	10.65	11.25	11.76	12.22	12.63	13.00	10.42	10.92	11.42	11.88	12.31	12.71
91	9.75	10.40	10.95	11.43	11.86	12.25	9.49	10.05	10.58	11.07	11.53	11.95
92	8.86	9.57	10.15	10.66	11.11	11.51	8.57	9.20	9.77	10.29	10.77	11.21
93	8.00	8.76	9.38	9.91	10.38	10.80	7.67	8.37	8.99	9.54	10.04	10.50
94	7.16	7.98	8.64	9.19	9.68	10.12	6.80	7.57	8.23	8.81	9.33	9.81
95	6.35	7.23	7.92	8.50	9.00	9.45	5.97	6.80	7.50	8.11	8.65	9.15
96	5.58	6.52	7.24	7.84	8.36	8.82	5.17	6.07	6.80	7.44	8.00	8.51
97	4.86	5.84	6.59	7.21	7.74	8.21	4.42	5.38	6.15	6.80	7.38	7.90
98	4.19	5.22	5.98	6.61	7.16	7.64	3.73	4.74	5.53	6.21	6.80	7.33
99	3.58	4.64	5.42	6.06	6.61	7.10	3.11	4.15	4.96	5.65	6.25	6.79
100	3.04	4.11	4.90	5.55	6.10	6.59	2.56	3.62	4.44	5.13	5.74	6.28
101	2.58	3.64	4.43	5.07	5.63	6.12	2.10	3.15	3.97	4.66	5.27	5.81
102	2.18	3.22	4.00	4.64	5.20	5.69	1.72	2.74	3.55	4.23	4.83	5.37
103	1.85	2.86	3.62	4.25	4.80	5.28	1.40	2.38	3.17	3.84	4.43	4.97
104	1.57	2.53	3.28	3.90	4.43	4.91	1.14	2.07	2.83	3.48	4.07	4.59
105	1.33	2.25	2.97	3.57	4.10	4.57	0.94	1.80	2.53	3.16	3.73	4.25
106	1.14	2.00	2.69	3.28	3.79	4.25	0.77	1.57	2.26	2.87	3.43	3.93
107	0.97	1.78	2.44	3.01	3.51	3.96	0.64	1.36	2.02	2.61	3.14	3.64
108	0.83	1.59	2.22	2.76	3.25	3.69	0.53	1.19	1.81	2.37	2.89	3.37
109	0.72	1.42	2.02	2.54	3.01	3.44	0.44	1.04	1.62	2.16	2.65	3.12
110	0.62	1.27	1.83	2.34	2.79	3.21	0.37	0.91	1.45	1.96	2.44	2.89

Table 4.3: Call prices for strikes from 90 to 110.

	$\alpha = 0.9$						$\alpha = 1$					
	$M = 1$	$M = 2$	$M = 3$	$M = 4$	$M = 5$	$M = 6$	$M = 1$	$M = 2$	$M = 3$	$M = 4$	$M = 5$	$M = 6$
90	10.26	10.66	11.11	11.56	11.99	12.41	10.16	10.45	10.83	11.25	11.69	12.12
91	9.30	9.76	10.25	10.74	11.21	11.65	9.19	9.53	9.96	10.42	10.89	11.35
92	8.36	8.89	9.43	9.95	10.44	10.91	8.22	8.63	9.11	9.62	10.11	10.60
93	7.44	8.03	8.62	9.18	9.70	10.19	7.28	7.75	8.29	8.84	9.37	9.88
94	6.54	7.21	7.85	8.44	8.99	9.50	6.35	6.91	7.50	8.09	8.65	9.19
95	5.67	6.42	7.11	7.73	8.31	8.84	5.44	6.10	6.75	7.37	7.97	8.52
96	4.84	5.67	6.41	7.06	7.66	8.20	4.58	5.33	6.04	6.70	7.31	7.89
97	4.06	4.97	5.74	6.42	7.04	7.60	3.78	4.61	5.37	6.06	6.69	7.29
98	3.35	4.32	5.12	5.82	6.45	7.02	3.05	3.95	4.74	5.46	6.11	6.72
99	2.72	3.73	4.55	5.26	5.90	6.48	2.40	3.36	4.17	4.90	5.56	6.18
100	2.17	3.20	4.03	4.75	5.39	5.97	1.84	2.82	3.64	4.38	5.05	5.67
101	1.71	2.73	3.55	4.27	4.92	5.50	1.39	2.35	3.17	3.90	4.57	5.19
102	1.34	2.32	3.13	3.84	4.48	5.06	1.03	1.94	2.74	3.47	4.13	4.75
103	1.04	1.96	2.75	3.45	4.08	4.65	0.75	1.60	2.37	3.07	3.73	4.34
104	0.81	1.66	2.42	3.09	3.71	4.28	0.55	1.31	2.04	2.72	3.36	3.96
105	0.63	1.41	2.12	2.77	3.37	3.93	0.40	1.07	1.75	2.40	3.02	3.60
106	0.50	1.19	1.86	2.49	3.07	3.61	0.30	0.87	1.50	2.11	2.71	3.28
107	0.39	1.01	1.63	2.23	2.79	3.31	0.23	0.71	1.28	1.86	2.43	2.98
108	0.31	0.86	1.44	2.00	2.53	3.04	0.18	0.59	1.10	1.64	2.18	2.71
109	0.25	0.73	1.26	1.79	2.30	2.79	0.14	0.49	0.94	1.44	1.95	2.46
110	0.21	0.63	1.11	1.61	2.09	2.56	0.11	0.40	0.81	1.27	1.75	2.23

Table 4.4: Parameters obtained for the Heston model.

	σ	κ	η	v_0	ρ
$\alpha = 0.7$	2.1129	5.4290	0.0611	0.0955	-0.0586
$\alpha = 0.8$	1.6831	5.3945	0.0614	0.0612	-0.0572
$\alpha = 0.9$	1.3224	5.3998	0.0584	0.0379	-0.0504
$\alpha = 1$	1.0244	5.3940	0.0533	0.0226	-0.0368

where $(v_t)_{t \geqslant 0}$ is the volatility process, assumed to satisfy the SDE

$$\mathrm{d}v_t = \kappa(\eta - v_t)\mathrm{d}t + \sigma\sqrt{v_t}\mathrm{d}\tilde{W}_t. \tag{4.86}$$

The processes $(W_t)_{t \geqslant 0}$ and $(\tilde{W}_t)_{t \geqslant 0}$ are two Brownian motions with correlation $\rho \in [-1, 1]$. The Heston model can be calibrated by using its characteristic function, for which we benefit from a closed-form formula, see Albrecher et al. (2007).

In order to compare our results with the Heston model, we calibrated the Heston model on the implied volatilities computed with our model. We have performed four calibrations: one for each fractional order $\alpha \in \{0.7, 0.8, 0.9, 1\}$. For each calibration, we have taken the implied volatilities for calls with strike in $\{90, 91, \dots, 110\}$ and maturity (in months) in $\{1, 2, \dots, 6\}$. We thus have $6 \times 21 = 126$ implied volatilities to calibrate each Heston model. The parameters have been chosen so as to minimize

$$\sum_{M=1}^{6} \sum_{K=90}^{110} \left(\sigma_{\mathrm{I}}^{\mathrm{Reference}}(M, K) - \sigma_{\mathrm{I}}^{\mathrm{Heston}}(M, K)\right)^2,$$

where $\sigma_{\mathrm{I}}^{\mathrm{Reference}}(M, K)$ and $\sigma_{\mathrm{I}}^{\mathrm{Heston}}(M, K)$ respectively denote the implied volatility from our model and the implied volatility from the Heston model, both for a strike K and a maturity M months.

The results of the calibrations can be found in Table 4.4 and 4.5. We observe that the Heston model achieves a good fit, although some small mismatches can be observed for the lowest or highest strikes. This is not surprising, as the Heston model is known to be able to reproduce a decent amount of implied volatility smiles (Albrecher et al. (2007)). As a conclusion, the Heston model seems a good approximation of our model regarding call option prices and implied volatilities. However, the Heston model would fail both taking into account the illiquidity of an asset via motionless periods in the price path and reproducing self-exciting shocks in the asset price.

Conclusions

In this chapter, we proposed a model for small capitalization stocks that are characterized by illiquidity periods. The motionless periods in their prices were

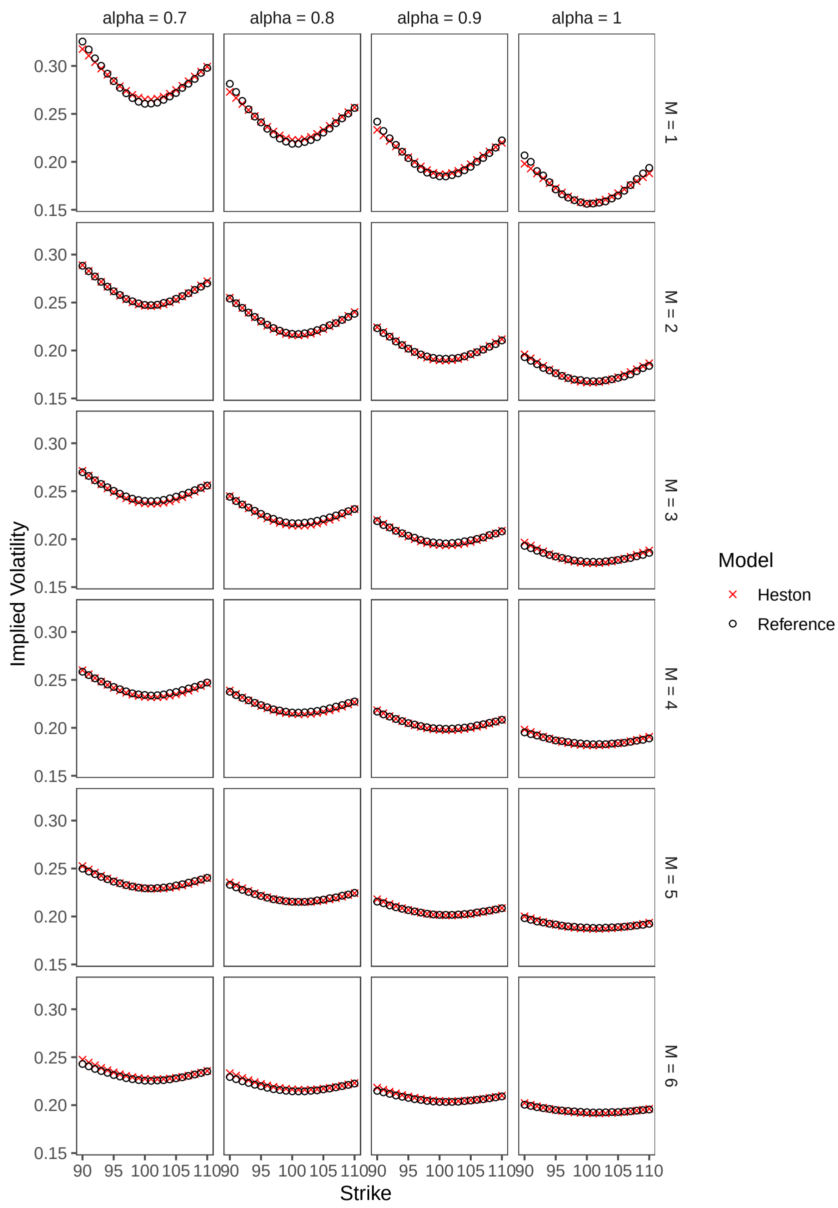

Figure 4.5: Comparison of implied volatilities between the (non-)fractional models and the classical Heston model. M is the maturity expressed in months.

modeled through the inverse of an α-stable subordinator. The occurrence of clustered sudden moves in the prices were captured by the use of self-exciting Hawkes processes. In a first section, we have showed that if we add some information about the path of the α-stable subordinator, its inverse satisfy the Markov property. In the next sections, we have presented a class of risk-neutral measures for our model and described a numerical method to compute call option prices. Afterwards, we derived quadratic hedging strategy for contingent claims, and in particular for call options. We concluded this chapter by giving the numerical results we obtained when computing call option prices with the method we proposed.

Conclusions

In this thesis, we proposed new methods for mathematical finance. These contributions were organized around two main topics. The first is the risk of contagion, which refers to the phenomenon of clustering of events in financial data. We have seen that a clustering of events can for example occur in credit risk: the default of an economic agent may lead to defaults of other economic agents through a propagation phenomenon. The other example that was of interest in this thesis was the clustering of shocks in asset prices. The second main topic we treated was the use of fractional processes in finance. We have used this type of processes in two areas of quantitative finance: the credit risk modeling and stock price modeling. In the case of credit risk, we have shown that using a fractional model leads to an improvement of the fit for companies whose survival is threatened in the short term. In the case of stock price modeling, we have seen that a fractional model is appropriate for the modeling of illiquid stocks. The prices of such stocks exhibit motionless periods of time, a feature that is well reproduced by the fractional model we used.

In a first chapter, we were concerned with time-consistent actuarial evaluations of a credit insurance portfolio subject to the risk of contagion. A time-consistent evaluation is an evaluation according to which a risk that will be cheaper than another one at a future time should already be cheaper today. The time-consistency is thus a desirable criterion that regards a proper use of the information available at each point in time. We have seen that the risk margins we find in classical actuarial evaluations prevents them from being time-consistent. Using tools of stochastic calculus, we derived PIDE to describe the time-consistent counterparts of the variance, standard-deviation and exponential principles. A numerical method based on finite differences was proposed to solve these PIDE. The numerical results of this procedure allowed us to draw several conclusions regarding the risk of contagion and the time-consistency of evaluations. The first main conclusion was that for all the actuarial evaluations considered, the time-consistency leads to higher prices, suggesting that time-inconsistent evaluations tend to underestimate the value of the credit portfolio. The second main conclusion was that the price underestimation was more severe in case a risk of contagion exists in the portfolio. This suggests that the existence of a risk of contagion in a credit portfolio can be seen as a reason to opt for a time-consistent evaluation. Finally, in order to allow for practical use, we proposed a method to calibrate our model on past data by relying on likelihood maximization.

In the second chapter, we used a fractional reduced self-exciting model for obtaining a better fit for survival probabilities in credit risk. The fractional model is obtained by time-changing the original model by the inverse of an α-stable subordinator. We have seen that the survival probabilities implied by the model can be obtained by computing the Laplace transform of the integral of the intensity. We started with the derivation of a PDE for the joint PDF of the intensity and its integral. A PDE was also derived for their joint Laplace transform. Then, using tools from fractional calculus, we have shown that these two PDE turn into FPDE in the time-changed model. The PDE and FPDE are almost identical, the only difference is that the partial derivative with respect to time is replaced by a fractional Caputo derivative in the time-changed setting. In order to obtain the survival probability curve implied by the model, we propose a finite difference method to solve the (F)PDE. Our results allowed us to calibrate both the original and fractional model to real market data and show that the fractional model provides a better fit.

The third chapter proposed a more simple alternative to the method of Chapter 2. The goal was here to avoid the use of fractional calculus and to simplify the computations, while maintaining the quality of the fit. We used the potential approach that was introduced in Rogers (1997) for the modeling of the term structure of interest rates. The potential approach consists in using a supermartingale process in order to directly model the survival probability curve. In that regard, it is a substitute for the more common reduced and structural approaches to credit risk. The risk of contagion was taken into account by using a Hawkes jump-diffusion exponential supermartingale. As in Chapter 2, the model was time-changed by the inverse of an α-stable process. In order to compute the survival probabilities implied by the model, we needed the Laplace transform of a Hawkes process intensity. By introducing a new method to compute all the moments of the Hawkes process intensity, we were able to approximate the Laplace transform with the help of a Taylor-Maclaurin expansion. We proved the convergence of this approximation and gave upper bounds on the numerical error. Finally, we calibrated the model on real data and were able to obtain a fit that is as good as in Chapter 2.

The fourth and last chapter used similar tools for the modeling of the prices of illiquid stocks, as well as the pricing and hedging of options on these stocks. Inverses of α-stable subordinators exhibit motionless periods in their sample paths, which makes them good candidates to model stocks that are not heavily traded on the market. Option pricing and hedging requires to work in a dynamic framework. To this end we established that, when coupled with another adequate process, the inverse α-stable subordinator satisfies the Markov property. Therefore, it makes sense to use its conditional distribution based on the last observed value of these processes. Next, we described a procedure to compute vanilla option prices in our model, taking the call option as an example. We relied on a FFT algorithm to compute call prices from the Fourier transform of the log asset price. To obtain this Fourier transform, we showed that it satisfies a FPDE. We solved this equation with the help of finite differences approximations for which the orders of convergence were discussed in details. Afterwards, we established a class of risk-neutral measures. Next, we derived an investment strategy that minimizes the hedging error. Finally, we proposed some numerical experiments to illustrate our method.

References

Ait-Sahalia, Y., Cacho-Diaz, J., and Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3):585–606.

Albrecher, H., Mayer, P., Schoutens, W., and Tistaert, J. (2007). The little heston trap. *Wilmott Magazine*, pages 83–92.

Alikhanov, A. A. (2015). A new difference scheme for the time fractional diffusion equation. *Journal of Computational Physics*, 280:424–438.

Artzner, P., Delbaen, F., Jean-Marc, E., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

Ayadi, M. A., Ben-Ameur, H., and Fakhfakh, T. (2016). A dynamic program for valuing corporate securities. *European Journal of Operational Research*, 249(2):751–770.

Aït-Sahalia, Y., Laeven, R. J., and Pelizzon, L. (2014). Mutual excitation in eurozone sovereign cds. *Journal of Econometrics*, 183(2):151–167.

Bacry, E., Mastromatteo, I., and Muzy, J. (2015). Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(1).

Ballestra, L. and Pacelli, G. (2014). Valuing risky debt: A new model combining structural information with the reduced-form approach. *Insurance: Mathematics and Economics*, 55:261–271.

Ballotta, L., Fusai, G., and Marazzina, D. (2019). Integrated structural approach to credit value adjustment. *European Journal of Operational Research*, 272(3):1143–1157.

Barkai, E., Metzler, R., and Klafter, J. (2000). From continuous time random walks to the fractional fokker-planck equation. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 61:132–8.

Bauwens, L. and Hautsch, N. (2006). Modelling financial high frequency data using point processes. *in: "Handbook of Financial Time Series", Springer.*

Bielecki, T. and Rutkowski, M. (2014). *Credit Risk: Modeling, Valuation and Hedging*. Springer Finance. Springer Berlin Heidelberg.

Black, F. and Cox, J. C. (1976). Valuing corporate securities: Some effects of bond indenture provisions. *Journal of Finance*, 31(2):351–367.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Brigo, D. and Mercurio, F. (2013). *Interest Rate Models - Theory and Practice*. Springer Finance. Springer, second edition.

Brigo, D. and Vrins, F. (2018). Disentangling wrong-way risk: pricing credit valuation adjustment via change of measures. *European Journal of Operational Research*, 269(3):1154–1164.

Brody, D., Hughston, L., and Meister, B. (2020). Theory of cryptocurrency interest rates. *SIAM Journal on Financial Mathematics*, 11(1):148–168.

Capinski, M. (2007). A model of credit risk based on cash flow. *Computers and Mathematics with Applications*, 54:499–506.

Carr, P. and Madan, D. B. (1999). Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2:61–73.

Cartea, A. and Jaimungal, S. (2013). Modelling asset prices for algorithmic and high-frequency trading. *Applied Mathematical Finance*, 20(6):512–547.

Cartea, A., Jaimungal, S., and Ricci, J. (2014). Buy low, sell high: A high frequency trading perspective. *SIAM Journal on Financial Mathematics*, 5(1):415–444.

Chan, T. (1999). Pricing contingent claims on stocks driven by Lévy processes. *The Annals of Applied Probability*, 9(2):504– 528.

Chen, C.-J. and Panjer, H. (2003). Unifying discrete structural models and reduced-form models in credit risk using a jump-diffusion process. *Insurance: Mathematics and Economics*, 33:357–380.

Çınlar, E. (2011). *Probability and Stochastics*. Graduate Texts in Mathematics. Springer.

Constantinides, G. M. (1992). A theory of the nominal term structure of interest rates. *The Review of Financial Studies*, 5(4):531–552.

Cont, R. and de Larrard, A. (2013). Price dynamics in a markovian limit order market. *SIAM Journal on Financial Mathematics*, 4(1):1–25.

Cont, R. and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC Financial Mathematics Series. Taylor & Francis.

Cont, R., Tankov, P., and Voltchkova, E. (2012). Hedging with options in models with jumps. *Abel Symposia on Stochastic Analysis and Applications*, 2.

Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407.

Cui, L., Hawkes, A., and Yi, H. (2020). An elementary derivation of moments of hawkes processes. *Advances in Applied Probability*, 52:102–137.

Da Fonseca, J., Dawui, K., and Malevergne, Y. (2022). A linear-rational multi-curve term structure model with stochastic spread. *SSRN Electronic Journal*.

Daley, D. and Vere-Jones, D. (2006). *An Introduction to the Theory of Point Processes: Volume I: Elementary Theory and Methods*. Probability and Its Applications. Springer.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300:463–520.

Diethelm, K. (2010). *The Analysis of Fractional Differential Equations*. Springer.

Duffie, D., Pan, J., and Singleton, K. (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica*, 68(6):1343–1376.

Duffie, D. and Singleton, K. (1999). Modeling term structure of defaultable bonds. *Review of Financial Studies*, 12:687–720.

Duffie, D. and Singleton, K. (2003). *Credit Risk: Pricing Measurement and Management*. Princeton University Press.

Duffy, D. (2006). *Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach*. John Wiley and Sons.

Eckert, J., Jakob, K., and Fischer, M. (2016). A credit portfolio framework under dependent risk parameters: Probability of default, loss given default and exposure at default. *Journal of Credit Risk*, 12:97–119.

Ekvall, N., Jennergren, L. P., and Näslund, B. (1997). Currency option pricing with mean reversion and uncovered interest parity: A revision of the garman-kohlhagen model. *European Journal of Operational Research*, 100(1):41–59.

El-Sayed, A., Behiry, S., and Raslan, W. (2010). Adomian's decomposition method for solving an intermediate fractional advection-dispersion equation. *Computers and Mathematics with Applications*, 59:1759–1765.

Eliazar, I. and Klafter, J. (2004). Spatial gliding, temporal trapping, and anomalous transport. *Physica D: Nonlinear Phenomena*, 187:30–50.

Errais, E., Giesecke, K., and Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1):642–665.

fei Ding, H. and xin Zhang, Y. (2011). Notes on implicit finite difference approximation for time fractional diffusion equations. *Computers and Mathematics with Applications*, 61(9):2924–2928.

Flesaker, B. and Hughston, L. (1999). Positive interest. *Risk*, 9.

Fujiwara, T. and Miyahara, Y. (2003). The minimal entropy martingale measures for geometric lévy processes. *Finance and Stochastics*, 7:509–531.

Föllmer, H. and Schweizer, M. (1991). Hedging of contingent claims under incomplete information. *Applied Stochastic Analysis*, pages 389–414.

Föllmer, H. and Sondermann, D. (1985). Hedging of non-redundant contingent claims. *Contributions to Mathematical Economics*, pages 205–223.

Geske, R. (1977). The valuation of corporate liabilities as compound options. *The Journal of Financial and Quantitative Analysis*, 12(4):541–552.

Golbabai, A. and Sayevand, K. (2011). Analytical modelling of fractional advection-dispersion equation defined in a bounded space domain. *Mathematical and Computer Modelling*, 53:1708–1718.

Gupta, N. and Kumar, A. (2022). Inverse tempered stable subordinators and related processes with mellin transform. *Statistics and Probability Letters*, 186:109465.

Hainaut, D. (2016a). A bivariate hawkes process for interest rate modeling. *Economic Modelling*, 57:180–196.

Hainaut, D. (2016b). A model for interest rates with clustering effects. *Quantitative Finance*, 16(8):1203–1218.

Hainaut, D. (2017). Clustered lévy processes and their financial applications. *Journal of Computational and Applied Mathematics*, 319:117–140.

Hainaut, D. (2020). Fractional hawkes processes. *Physica A: Statistical Mechanics and its Applications*, 549:124330.

Hainaut, D. (2021). A fractional multi-states model for insurance. *Insurance: Mathematics and Economics*, 98:120–132.

Hainaut, D. and Colwell, D. (2014). A structural model for credit risk with switching processes and synchronous jumps. *The European Journal of Finance*, 22:1–23.

Hainaut, D. and Goutte, S. (2019). A switching microstructure model for stock prices. *Mathematics and Financial Economics*, 13:459–490.

Hainaut, D. and Le Courtois, O. (2014). An intensity model for credit risk with switching lévy processes. *Quantitative Finance*, 14(8):1453–1465.

Hainaut, D. and Leonenko, N. (2021). Option pricing in illiquid markets: A fractional jump-diffusion approach. *Journal of Computational and Applied Mathematics*, 381:112995.

Hansen, L. and Sargent, T. (1995). Discounted linear exponential quadratic gaussian control. *IEEE Transactions on Automatic Control*, 40(5):968–971.

Hansen, L. P. and Sargent, T. J. (2001). Robust control and model uncertainty. *The American Economic Review*, 91(2):60–66.

Hawkes, A. G. (1971a). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society. Series B (Methodological)*, 33(3):438–443.

Hawkes, A. G. (1971b). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90.

Hawkes, A. G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11(3):493–503.

He, X.-J. and Chen, W. (2018). A monte-carlo based approach for pricing credit default swaps with regime switching. *Computers and Mathematics with Applications*, 76(7):1758–1766.

Heath, D., Jarrow, R., and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1):77–105.

Heston, S. L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies*, 6(2):327–343.

Ho, T. S. Y. and Lee, S.-B. (1986). Term structure movements and pricing interest rate contingent claims. *The Journal of Finance*, 41(5):1011–1029.

Hubalek, F. and Sgarra, C. (2006). Esscher transforms and the minimal entropy martingale measure for exponential lévy models. *Quantitative Finance*, 6(2):125–145.

Jacod, J. (1979). *Calcul stochastique et problèmes de martingales*. Lecture Notes in Mathematics. Springer, Berlin.

Jacod, J. and Shiryaev, A. N. (2003). *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin, second edition.

Jafari, H. and Gejji, V. (2006). Solving linear and nonlinear fractional diffusion and wave equations by adomian decomposition. *Applied Mathematics and Computation*, 180:488–497.

Janicki, A. and Weron, A. (1994). *Simulation and Chaotic Behavior of Alpha-stable Stochastic Processes*. CRC Press.

Jeanblanc, M., Yor, M., and Chesney, M. (2009). *Mathematical Methods for Financial Markets*. Springer-Verlag.

Jeon, J., Yoon, J.-H., and Kang, M. (2016). Valuing vulnerable geometric asian options. *Computers and Mathematics with Applications*, 71(2):676–691.

Kaas, R., Goovaerts, M., Dhaene, J., and Denuit, M. (2008). *Modern Actuarial Risk Theory: Using R*. Springer.

Ketelbuters, J.-J. and Hainaut, D. (2022a). Cds pricing with fractional hawkes processes. *European Journal of Operational Research*, 297(3):1139–1150.

Ketelbuters, J.-J. and Hainaut, D. (2022b). Time-consistent evaluation of credit risk with contagion. *Journal of Computational and Applied Mathematics*, 403:113848.

Kim, J., Kim, T., Jo, J., Choi, Y., Lee, S., Hwang, H., Yoo, M., and Jeong, D. (2016). A practical finite difference method for the three-dimensional black–scholes equation. *European Journal of Operational Research*, 252(1):183–190.

Kobayashi, K. (2011). Stochastic calculus for a time-changed semimartingale and the associated stochastic differential equations. *Journal of Theoretical Probability*, 24:789–820.

Kunita, H. (2004). Representation of martingales with jumps and applications to mathematical finance. *Stochastic Analysis and Related Topics in Kyoto. In honour of Kiyosi Itô*, page 209–233.

Kunita, H. and Watanabe, S. (1967). On square integrable martingales. *Nagoya Mathematical Journal*, 30:209–245.

Lando, D. (2004). *Credit Risk Modeling: Theory and Applications*. Princeton University Press, United States.

Lehalle, C.-A., Mounjid, O., and Rosenbaum, M. (2021). Optimal liquidity-based trading tactics. *Stochastic Systems*, 11(4):368–390.

Leonenko, N., Meerschaert, M., and Sikorskii, A. (2013a). Correlation structure of fractional pearson diffusions. *Computers and Mathematics with Applications*, 66(5):737–745. Fractional Differentiation and its Applications.

Leonenko, N. N., Meerschaert, M. M., and Sikorskii, A. (2013b). Fractional pearson diffusions. *Journal of Mathematical Analysis and Applications*, 403(2):532–546.

Liang, X. and Wang, G. (2012). A reduced form credit risk model with common shock and regime switching. *Insurance: Mathematics and Economics*, 51(3):567–575.

Longstaff, F. and Schwartz, E. S. (1995). A simple approach to valuing risky fixed and floating rate debt. *The Journal of Finance*, 50(3):789–819.

Ma, Y. and Xu, W. (2016). Structural credit risk modelling with hawkes jump diffusion processes. *Journal of Computational and Applied Mathematics*, 303:69–80.

Macrina, A. (2014). Heat kernel models for asset pricing. *International Journal of Theoretical and Applied Finance*, 17:1450048.

Magdziarz, M. (2009a). Black-scholes formula in subdiffusive regime. *Journal of Statistical Physics*, 136:553–564.

Magdziarz, M. (2009b). Stochastic representation of subdiffusion processes with time-dependent drift. *Stochastic Processes and their Applications*, 119(10):3238–3252.

Mari, C. and Renò, R. (2005). Credit risk analysis of mortgage loans: An application to the italian market. *European Journal of Operational Research*, 163:83–93.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, 29(2):449–470.

Metzler, R., Barkai, E., and Klafter, J. (1999). Anomalous diffusion and relaxation close to thermal equilibrium: A fractional fokker-planck equation approach. *Physical Review Letters*, 82:3563–3567.

Metzler, R. and Klafter, J. (2000). The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Physics Reports*, 339(1):1–77.

Metzler, R. and Klafter, J. (2004). The restaurant at the end of the random walk: Recent developments in the description of anomalous transport by fractional dynamics. *Journal of Physics A: Mathematical and General*, 3737.

Moraux, F. and Hainaut, D. (2018). Hedging of options in presence of jump clustering. *Journal of Computational Finance*, 22:1–35.

Moraux, F. and Hainaut, D. (2019). A switching self-exciting jump diffusion process for stock prices. *Annals of Finance*, 15(2):267–306.

Murio, D. A. (2008). Implicit finite difference approximation for time fractional diffusion equations. *Computers and Mathematics with Applications*, 56(4):1138–1145.

Musmeci, F. and Vere-Jones, D. (1992). A space-time clustering model for historical earthquakes. *Annals of the Institute of Statistical Mathematics*, 44:1–11.

Odibat, Z. and Momani, S. (2009). The variational iteration method: An efficient scheme for handling fractional partial differential equations in fluid mechanics. *Computers and Mathematics with Applications*, 58(11):2199–2208. New Analytical Methods and Applications.

Ogata, Y. (1998). Space-time point-process models for earthquake occurrences. *Annals of the Institute of Statistical Mathematics*, 50(2):379–402.

Pelsser, A. and Ghalehjooghi, A. S. (2016). Time-consistent actuarial valuations. *Insurance: Mathematics and Economics*, 66:97–112.

Pelsser, A. and Stadje, M. (2014). Time-consistent and market-consistent evaluations. *Mathematical Finance*, 24:25–65.

Pham, H. (2000). On quadratic hedging in continuous time. *Mathematical Methods of Operational Research*, 51:315–339.

Piryatinska, A., Saichev, A., and Woyczynski, W. (2004). Models of anomalous diffusion: The subdiffusive case. *Physica A: Statistical and Theoretical Physics*, 349 (2005):375–420.

Podlubny, I. (1999). *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*. Mathematics in science and engineering. Academic Press, London.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32(1):122–136.

Protter, P. (2005). *Stochastic Integration and Differential Equation*. Springer-Verlag, Berlin, Heidelberg, second edition.

Reinhart, A. (2018). A review of self-exciting spatio-temporal point processes and their applications. *Statistical Science*, 33(3):299–318.

Revuz, D. and Yor, M. (2005). *Continuous Martingales and Brownian Motion*. A Series of Comprehensive Studies in Mathematics. Springer, third edition.

Riedel, F. (2003). Dynamic coherent risk measures. *Stochastic Processes and their Applications*, 112(2):185–200.

Rogers, L. and Williams, D. (2000). *Diffusions, Markov Processes and Martingales: Volume 2, Itô Calculus*. Cambridge Mathematical Library. Cambridge University Press.

Rogers, L. C. G. (1997). The potential approach to the term structure of interest rates and foreign exchange rates. *Mathematical Finance*, 7:157–176.

Rosazza Gianin, E. (2006). Risk measures via g-expectations. *Insurance: Mathematics and Economics*, 39(1):19–34.

Rutkowski, M. (1997). A note on the flesaker-hughston model of the term structure of interest rates. *Applied Mathematical Finance*, 4(3):151–163.

Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge studies in advanced mathematics, revised edition.

Scalas, E. (2006). *Five Years of Continuous-time Random Walks in Econophysics*, pages 3–16. Springer.

Schoutens, W. and Cariboni, J. (2009). *Levy Processes in Credit Risk*. John Wiley and Sons.

Schweizer, M. (2001). *A Guided Tour through Quadratic Hedging Approaches*, page 538–574. Cambridge University Press.

Song, L. and Weiguo, W. (2013). Solution of the fractional black-scholes option pricing model by finite difference method. *Abstract and Applied Analysis*, 2013:1–10.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188.

Wang, L., Ma, Y., and Meng, Z. (2014). Haar wavelet method for solving fractional partial differential equations numerically. *Applied Mathematics and Computation*, 227:66–76.

Williams, D. (1991). *Probability with Martingales.* Cambridge mathematical textbooks. Cambridge University Press.

Wong, H. Y. and Lo, Y. W. (2009). Option pricing with mean reversion and stochastic volatility. *European Journal of Operational Research*, 197(1):179–187.

Zheng, G. H. and Wei, T. (2010). Spectral regularization method for a cauchy problem of the time fractional advection-dispersion equation. *Journal of Computational and Applied Mathematics*, 233:2631–2640.

Zvan, R., Vetzal, K. R., and Forsyth, P. A. (2000). Pde methods for pricing barrier options. *Journal of Economic Dynamics and Control*, 24(11-12):1563–1590.