UNIVERSITÉ CATHOLIQUE DE LOUVAIN

LOUVAIN INSTITUTE OF DATA ANALYSIS AND MODELING IN ECONOMICS AND STATISTICS

INSTITUT DE STATISTIQUE, BIOSTATISTIQUE ET SCIENCES ACTUARIELLES

Financial and actuarial derivatives pricing with self-exciting processes

Charles Guy NJIKE LEUNGA

Thesis submitted in partial fulfilment
of the requirements of the degree of
Doctor of sciences

Thesis committee:
Prof. Michel DENUIT (UCLouvain), Chairman
Prof. Donatien HAINAUT (UCLouvain), Supervisor
Prof. Pierre DEVOLDER (UCLouvain), Supervisor
Prof. Leonardo IANIA (UCLouvain), Secretary
Prof. Jérôme BARBARIN (UCLouvain)
Prof. Maggie CHEN (Université de Cardiff)
Prof. Griselda DEELSTRA (Université Libre de Bruxelles)

Louvain-la-Neuve, February 2023

To my family, especially my parents and wife

Acknowledgments

This endeavor would not have been possible without the support of many people. I thus would like to take this opportunity to express my gratitude to some of them.

Foremost, I would like to express my sincere gratitude to my PhD supervisors, Prof. Donatien Hainaut and Prof. Pierre Devolder, for agreeing to mentor me and guide me throughout my work. Prof. Hainaut, thank you for the continuous support, your patience, your motivation, your enthusiasm, your availability and your valuable assistance till the realisation of this thesis. Prof Devolder thank you for introducing me to the field of Mathematical Finance, for your encouragements and for the insightful discussions we had.

I would also like to extend my gratitude and appreciation to the jury members, Prof. Michel Denuit, Prof. Leonardo Iania, Prof. Jérôme Barbarin, Prof. Griselda Deelstra and Prof. Maggie Chen for the time that they have taken to read the thesis and for making constructive comments and feedback which led to a significant improvement in the quality of the final version.

I want to specially thank Prof. Griselda Deelstra, Prof. Christian Hafner that kindly accepted to be part of my accompanying committee. I am very grateful to their useful comments and encouragements during the confirmation test.

Many thanks to the professors Johan Segers, Catherine Legrand, Christian Hafner, Cédric Heuchenne, Eugen Pircalabelu and Marie-Paule Kestemont, Nathan Uyttendaele, Aurélie Bertrand and Anouar El Ghouch, with whom I had the opportunity to collaborate and develop teaching skills. I am also grateful to all the members of the SMCS for introducing me to consulting work. A special thanks goes to the administrative team and IT staff who were always available to help me whenever I encountered administrative difficulties or technology related issues.

I am thankful to my PhD colleagues of ISBA, CORE and LFIN for having made the workplace a wonderful environment and for the challenging and funny discussions. I take the opportunity to thank my office neighbour Stefka and John-John for passing stimulating moments with them and sharing research distress.

Last but not the least, my family members and loved ones have been the most ardent supporters, and have also been the main victims of the time I have devoted to the realization of this thesis. I thank you all.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Most of financial products exchanged by market participants are derivative products, which are financial contracts whose value hinges on another financial asset, underlying it. For instance, underlying assets can be stocks, interest rates, commodities and foreign exchange rates. In order to price and measure risks carried by such financial products, quantitative finance relies on mathematical models such as continuous-time stochastic models. The latter ones are used in this thesis to capture the dynamics of the underlying assets. One of the most famous stochastic process for the modelling of asset prices was developed by Black and Scholes (1973). This work initiated various research projects in view of the criticism it received from the empirical studies on financial time series. Similarly to the Black-Scholes model, the model proposed by Uhlenbeck and Ornstein (1930), mainly used to model interest rate for its mean-reverting property launched a quest of models that best reproduce the dynamics of market interest rates. It is worth noting that the models aforementioned are diffusion processes made up of a drift component and a normally distributed Brownian motion component at a given time. An important step in the field of finance was taken thanks to Merton (1976), with the introduction in finance of processes including a diffusion part and a jump part driven by a point process such as the Poisson process. This jump part makes it possible to replicate sudden and unexpected changes in the asset's price. In the meantime, Hawkes (1971 a) and Hawkes (1971 b) had introduced a new family of point processes called Hawkes processes, that hold self-excitation and mutually excitation properties. At that time, no one guessed point processes with these features could be useful for financial applications. In the last 20 years, Hawkes processes have received much more attention in finance for the modelling of contagion effects in credit risk and asset price modelling (see e.g. Errais et al. (2010), Bacry and Muzy (2014), Aït-Sahalia et al. (2015), Hainaut and Colwell (2016) and Hainaut and Moraux (2018 a)). Already used in seismology (see e.g. Ogata (1998)) for its key property that the occurrence of any event increases the probability of further events occurring, this late use of Hawkes processes in finance reveals the recent emergence of contagion effects in this field and the upsurge of this type of phenomenon. Facing dramatic consequences of shocking events like the financial crisis of 2008, and recently, the COVID-19 pandemic or the war in Ukraine, setting models that accurately account the contagion patterns underlying financial products is a challenging task.

The main objective of this thesis is to investigate how incorporating a Hawkes process or a Hawkes process with modified mathematical properties from the original one, into a diffusion framework may improve the model fit, the pricing of derivatives and their hedging. In order to achieve this goal, we have explored four directions. First, the default risk with possible clustering of defaults of interbank market members observed during the credit crunch of 2007 calls for a new interest rate theory consistent with market rates and based on credit risk theory. Indeed, most of the previous research focuses either on a way to make consistent the interest rate theory with market rates by the so-called multiple curve approach (see Henrard (2014), Morini (2009), Grbac and Runggaldier (2015)) or on the change on credit quality within the bank panel determining the interbank rate (see Filipovíc and Trolle (2012)). While the first solution above came from practitioners and lacked of theoretical foundation, the second proposition originated from researchers. Although the latter one seems realistic, it relies strongly on the composition of panel banks which changes dynamically according to the selection criteria. In order to explain the observed market rates quotes after 2007, Chapter 2 explores a new way to reconcile the classical interest rate theory and market quotes by combining a multiple curve approach with an intensity based approach defining the arrivals rate of default within the interbank market as a process of Hawkes type and of the form of Cox-Ingersoll-Ross processes (Cox et al., 1985).

Second, low interest rates environment prompts investors to consider riskier assets or long-term investment strategies to obtain higher returns. For instance, a pre- or post-retirement investment strategy consisting in buying living and/or death annuity guarantees protects from market downturn and allows for extra returns in conjunction with growth. The valuation of such insurance products requires a suitable modelling of the underlying account value which may experience cluster jumps scaled according to the economic regime. The pricing of life insurance contracts with guarantees has been a topic of research interest for many years. Several authors have tried to address this pricing issue based on financial models (see Brennan and Schwartz (1976), Brennan and Schwartz (1979), Boyle and Schwartz (1977) or Milvesky and Posner (2001) with some considering also the policyholder behaviour or profit (see Bauer et al. (2008), Bernard et al. (2014), Cui et al. (2017), Bernard and Moenig (2018)) and others investigating the fair value of equity-linked insurance with regard to the level of solvency capital required (see Hardy (2003), Barbarin and Devolder (2005) and Feng and Volkmer (2012)). All these studies however have in common the limit of not incorporating switches in economic regimes into the variable annuities valuation framework, although it might help reducing the risk of losses for VA issuers, as these products offer protection of capital during periods of recession and provide additional returns in conjunction with growth to their holders. Regime-switching techniques have been used in the field of actuarial sciences in a continuous-time Markov switching framework to model interest rates and/or reference risky assets (see e.g. Hardy

(2001), Siu (2005), Lin et al. (2009), Hainaut (2014), Fan et al. (2015), Ignatieva et al. (2016)). However, to the best of our knowledge, self-exciting ones have not been explored for the valuation of equity-linked life insurance contracts such as VAs, although Hainaut and Moraux (2018 b) showed that the self-excitation property of the Hawkes process combined with regime switching models enables to better reproduce the dynamics of financial markets. Moreover, most studies rely on the assumption of a fixed interest rate to discount the cash flows, which is seems unrealistic for VAs as increase in interest rates is likely to have an adverse effect on the issuers commitments. In an attempt to address these limitations in the VAs valuation, Chapter 3 extends the work of Hainaut and Moraux (2018 b) to propose a switching Hawkes jump diffusion process for the modelling of the account values, and discusses the impact of the switching jump clustering on the value of variable annuities.

Third, the way huge changes in stock price may affect future price in the short and long term invites to investigate new classes of relaxation functions for stock price models that allow the future price to remember past jumps for a longer period of time. Starting from Black and Scholes (1973) reference model for asset pricing, the geometric Brownian motion with constant drift and volatility, several researchers have suggested incorporating a jump component to better reflect the asset returns features. Merton (1976), Kou (2002), Andersen et al. (2002), Carr and Wu (2003), Eraker (2004) proposed jump diffusion models which enable to reproduce the skewness and the excess kurtosis observed in assets returns, but do not allow for clustered jumps as they rely on the assumption that the arrivals of shocks are independently and/or identically distributed. To overcome this limitation highlighted by Maheu and McCurdy (2004), self-exciting point processes of Hawkes type were introduced in financial literature. For instance, Errais et al. (2010), Aït-Sahalia et al. (2015) and Hainaut and Moraux (2018 a) have used Hawkes processes with an exponential kernel to incorporate the clustering of jumps, with the advantage of having Hawkes process's moments with a closed form. However Bacry and Muzy (2014) and Bacry et al. (2016) provide empirical evidence of slow decay kernel importance for price and trades in high-frequency and conclude that the intensity of market events should depend on their history and on the state of the order book. Blanc et al. (2016) generalize the Hawkes price models introduced in Bacry and Muzy (2014), by allowing all feedback effects in the jump intensity that are linear and quadratic in past returns. It is worth noting that, the foundation of the popular rough volatility models comes from Hawkes processes with kernel of power-law type, as in the quadratic Hawkes processes (see e.g. EL Euch et al. (2018) and Dandapani et al. (2021)). Chen et al. (2021) design another range of Hawkes processes that they named fractional Hawkes processes by replacing the exponential kernel with a sign-changed first derivative of a Mittag-Leffler function firstly defined in Mittag-Leffler (1903) as a power series. Although, Mittag-Leffler functions make the Hawkes process non-Markov, in Chapter 4 we consider a Mittag-Leffler function which is sub-exponential and investigate the use of a Hawkes process with a relaxation function or kernel function of the form of Mittag-Leffler functions for stock price modelling and option pricing. To achieve our goal, we use a technique similar to the one used in the literature of stochastic rough volatility models (see e.g. Abi Jaber and El Euch (2019) and Bäuerle and Desmettre (2020)) to recover a Markov structure.

Many empirical findings on asset's return time series such as heavy tails, skewness and volatility clustering highlight the lack of diffusion processes to explain the volatility of asset's return. In view of that, the fourth and final direction of the thesis is devoted to the expansion of the asset model described in Chapter 4, by considering a stochastic volatility, and to expand the Heston model (Heston, 1993) by including a long memory feature. In Chapter 5, while the diffusion volatility of the asset price is a Cox-Ingersoll-Ross process, we consider jumps prices driven by Hawkes processes whose kernel functions are Fourier transforms of measures and we discuss the impact of the volatility memory implied by these kernel functions in option pricing.

1.1 Structure of the thesis

The structure of the thesis is organized as follows.
Chapter 2 puts forward one way to resolve the discrepancies between interest rate practice and theory after the 2007 credit crisis by modifying the arbitrage-free condition. In this framework, the forward Libor rate is no longer considered as a risk-free rate and the credit and liquidity risks within the interbank market partly drive its dynamics. In a similar manner to the multiple-curve approach, we model the evolution of default-free rates, assimilated to overnight interest swap rates, and the default times of an interbank market segment determined by its tenor. For each segment, we use the reduced form approach to model the arrival rate of defaults with a self-exciting jump-diffusion process. Then, we deduce the dynamics of the spot forward Libor rates and provide closed-form approximation pricing formulae for options on forward Libor rates and swap rates. Even in a context of negative interest rates and compared to other forms of intensity processes such as a Cox-Ingersoll-Ross process, the self-excitation property allows a better understanding of the spread OIS-IRS and provides information about the interbank credit risk. Furthermore, our framework enables to parse the impact of the interbank credit risk on forward Libor as well as on interest rates derivatives like caps, floors, and swaptions.

Chapter 3 revisits the valuation of variable annuities under a regime-switching model when the dynamics of the underlying stock price follows a self-exciting switching jump-diffusion process. In this framework, we add a jump component to a regime-switching geometric Brownian for large shocks on the stock price. Unlike Chapter 2, the intensity of shocks arrival is a Hawkes process modulated by a continuous time hidden Markov chain with a finite number of states corresponding to the different economic regimes. The features of shocks are thus modulated by economic regimes. The interest rate used for discounting is stochastic and correlated to the stock market. After deriving the theoretical moment generating function under an equivalent martingale measure, we approximate the call and put options prices by Fourier inversion.

Together with the mortality model, we discuss the valuation of variable annuity contract that guarantees a minimum living or death benefit. We conclude this chapter by some numerical results that highlight the impact of self-exciting jumps and economic regimes on the valuation of guarantees.

Chapter 4 proposes a model for asset prices driven by a self-exciting jump diffusion process similar to the one presented in Chapter 3 but not modulated by a hidden Markov chain. The novel feature of our model is that shocks are ruled by a Hawkes process with a Mittag-Leffler memory kernel. The Mittag-Leffler kernel is a relaxation function used in fractional calculus to describe complex memory effect. In particular, it generalizes the exponential memory that ensures Hawkes process has the Markov property. Despite its interesting characteristics, the Hawkes process with the Mittag-Leffler kernel is non-stationary and a non-Markov process. Nevertheless, we can infer a closed form expression for the moment generating function of log-returns. It is obtained by representing the Hawkes process with Mittag-Leffler kernel as an infinite dimensional Markov process. We also provide a change of measure and price European options by exploiting the fast Fourier transform technique. Applied to the times series of S&P 500 daily values, we demonstrate the efficiency of the proposed model to produce excess of kurtosis, skewness compared to models with exponential memory. Finally, we show the impact of the memory parameter on prices of call option on S&P 500.

Chapter 5 proposes an expansion of the stochastic volatility model of Heston by adding to the instantaneous asset prices, a jump component driven by a Hawkes process with a kernel function or memory kernel that is a Fourier transform of a probability measure. This kernel function defines the memory of the asset prices process. For instance, if it is fast decreasing, the contagion effect between asset price jumps is limited in time. Otherwise, the processes remember the history of asset price jumps for a long period. To investigate the impact of different rates of decay or types of memory, we consider four probability measures, the Laplace, the Gaussian, the Logistic and the Cauchy measures. Unlike Hawkes processes with exponential kernel, we lost the Markov property but preserve the stationarity property which ensures that the unconditional expected arrival rate of jump does not explode. Moreover, these kernel functions solve the non-stationary problem of the Hawkes process with Mittag-Leffler kernel employed in Chapter 4. In absence of the Markov property, we take advantage of the Fourier transform representation to derive a closed form expression of European call option price based on characteristic functions. The numerical illustration shows that our extensions of the Heston model achieve a better fit of the Euro Stoxx 50 option data than the standard version. To conclude, we analyse the sensitivity of European call option to memory and self-excitation parameters, underlying price, volatility and jump risks.

1.2 Published materials

The thesis is composed of four chapters, each chapter being either published or submitted.

Chapter 2 is published in International Journal of Theoretical and Applied Finance, Volume No. 23, Issue No. 06, Article No. 2050039, Year 2020. (Njike and Hainaut 2020)

Chapter 3 is published in Journal of Methodology and Computing in Applied Probability, Volume No. 24, Pages 963–990, Year 2022. (Njike and Hainaut 2022)

Chapter 4 has been submitted for publication and is at the second round of a submission process for publication.

Chapter 5 is submitted for publication.

Chapter 2

Interbank credit risk modelling with self-exciting jump processes

2.1 Introduction

To the best of our knowledge, up to now there is not a standard theory on the market of interbank rates which explains the observed market rates quotes since the credit crunch of 2007. As the interest rate derivatives market is the largest derivatives market, modelling interbank rates with a realistic and analytically tractable approach is a great challenge for the financial industry. Most of the previous research focuses either on a way to make consistent the interest rate theory with market rates by the so-called multiple curve approach (see Henrard 2014, Morini 2009, Grbac and Runggaldier 2015), or on the change on credit quality within the bank panel determining the interbank rate (see Filipovíc and Trolle 2012). However, as underlined by Mercurio (2009), model the default time of the interbank market remains a non-trivial task when it comes to a generic counterparty.

From the 1980s to nowadays, many integrated frameworks dedicated to interest rates modelling were proposed in the literature. The common aim of these frameworks is to explain changes in the behavior of market rates quotes, but also to capture emerging risks within the interest rates market. This chapter focuses on the modelling of forward Libor rates and proposes another framework. Before 2007, interest rate swaps (IRSs) were priced with a single interest rate curve that was totally independent from the frequency of floating payments. Libor was considered as a good proxy for the risk free rate. The deposit rates deduced from Libor, and Overnight indexed swap rates (OIS) for the same maturity chased each other with a small distance of few basis points. Earlier in 2000, authors like Schöenbucher (2001) integrate the default risk in the Libor market model to care about the risk introduced by changes in the level of the default-free interest rates or in the market's assessment of the credit quality of the obligor. Although neglected, the presence of credit and liquidity risk within the interbank market seemed to be a good justification for the small distance observed between deposit rates from Libor and OIS. After the credit crunch, we witnessed a serious widening of the distance between equivalent rates, the explosion of basis swap rates based on Libor rates with different tenors and the divergence between forward rate agreements (FRAs) and corresponding forward rates implied by deposits. Moreover, FRAs were not necessary hedgeable with a deposit and a loan, leading traditional no arbitrage conditions unsatisfied in practice.

Consequently, the discount curve has to be modeled differently and Libor for different tenors must be analyzed separately. From a single interest rate curve framework we move to a multiple curves framework where the classical interest rate theory as detailed by Brigo and Mercurio (2006) can not be applied. In the literature two strands of solution have emerged. The first one, called multiple curves approach by practitioners consists in dealing with the discrepancies between rates by segmenting market rates and labeling them differently according to their application period. In this setting Bianchetti (2008) uses two yield curves for market coherent estimation of discount factors and forward rates with different underlying tenor. Mercurio (2009) follows the multiple curves approach, and proposes a new Libor market model based on the joint evolution of FRA rates and forward rates belonging to the discount curve. Ametrano and Bianchetti (2013) provide the theoretical background necessary for bootstrapping interest curves in the multiple curves framework, while Fries (2016) focuses on the interpolation method. Hull and White (2015) validates the use of multiple curves by showing that the risk-free rates when pricing derivatives should be bootstrapped from OIS instead of Libor swap rate. Henrard (2014) explains from a practical point of view, foundations and evolution of the multiple curves framework, and its implementation in finance. Then, to be consistent with the new feature of interest rates, Eberlein et al. (2019) model negative rates or extremely low rates in a multiple curve setup.

The second strand of solutions comes from researchers and consists in modelling credit and liquidity risks of panel banks using the counterparties credit risk theory. For instance, Filipovíc and Trolle (2012) infer a term structure of interbank risks from spreads between rates on IRS indexed to Libor and OIS. They model the time when a deterioration in the credit quality of a bank within the initial Libor banks panel occurs . Their model is fitted to CDS spread term structures for every individual banks panel. Even if this approach is realistic, it remains cumbersome as the composition of panel banks changes dynamically according to the selection criteria. On the other hand, since the credit crunch of 2007-2008 the feedback phenomena (see Hurd 2018) observed between bankruptcies have motivated the use of self-exciting processes in credit risk modelling. The theory about self-exciting processes was introduced by Hawkes (1971 b). In the financial literature, these processes were firstly used in high frequency trading around the years 2000s. In credit risk, Errais et al. (2010) use a self-exciting process to model the cumulative losses of a credit portfolio taking into account the fact that when a loss occurs, the probability that another loss occurs, increases. Dassios and Zhao (2017) meanwhile introduce a new class of self-exciting processes that are driven by external factors . They use this process to calculate default probabilities and price zero coupon bonds with a high default probability. Coonjobeharry et al. (2016) also price convertible bonds with default risk using jump-diffusion processes. Moreover, in the literature, self-exciting point processes are used for the modelling of interest rates. Hainaut (2016b) models short term rates via a self-exciting jump process and reproduces the clustering of shocks on the Euro overnight index average. Hainaut (2016a) models again the short rate by taking into account the aggregate supply and demand for fixed income instruments modeled via a bi-variate self-exciting point process.

This chapter explores another way to reconcile the classical interest rate theory and market quotes by combining a multiple curve approach with a counterparty credit risk model. In our framework, we segment the interbank market into sub-markets according to the tenor of the interbank rate applied. The interbank credit risk of a sub-market, is the risk that a loss occurs from the default of one counterparty in a transaction between two banks. Motivated by the possible succession of default losses due to interconnections between banks, as Fontana and Runggaldier (2010) we consider a reduced-form credit risk model where the arrival rate of credit events in a sub-market follows a self-exciting jump-diffusion process of Hawkes type. The jump part in the intensity is a point process with positive jumps when a credit event occurs. This ensures that a default raises the probability of the arrival of the next credit event.

The rest of the chapter is structured as follows: in section (2.2) we introduce notations and definitions. Then, we explain our framework and motivate the choice of self-exciting jump diffusion processes for the default intensity. In section (2.3) we derive the forward Libor dynamics under the risk neutral world and study through simulations the impact of jumps on our default intensity. In section (2.4), we price options on forward Libor rates and swap rates. For this purpose, we describe the change of measure respectively towards the forward measure and the annuity measure. Then, we use a bivariate Inverse Fourier transform to infer under these measures the joint probability distribution of the short rate and of the intensity process. In section (2.5), we consider a bicurves framework involving the discount curve and the six months forward Libor. We calibrate our intensity model and compare it to other models in terms of goodness of fit. Finally, we analyze the sensitivity of the survival probability of an interbank market segment to parameters defining the self-excitation.

2.2 Notation and basic definitions

Let consider a filtered probability space $(\Omega, \mathscr{G}, \mathbb{Q})$ where the filtration $(\mathscr{G}_t)_{t\geq 0}$ satisfies the usual conditions and $\mathbb{Q}$ is a spot martingale measure. For sake of convenience, quantities that refer to defaultable bond prices or interest rates carry an index "c". We shall write $\mathbb{E}^{\mathbb{Q}}(\cdot) = \mathbb{E}(\cdot)$, and unless otherwise specified all expectations are determined under the risk neutral measure $\mathbb{Q}$. We assume that the instantaneous risk-free rate process is adapted to the subfiltration $(\mathscr{F}_t)_{t\geq 0}$ of $(\mathscr{G}_t)_{t\geq 0}$ and is ruled by the Hull and White model:

$$\mathrm{d}r_t = (\theta(t) - ar_t)\mathrm{d}t + \sigma_r \mathrm{d}W_t^r, \tag{2.1}$$

where W_t^r is a Brownian motion, $a \geq 0$, and $\sigma_r \geq 0$. $\theta(t)$ is a deterministic function of time t satisfying:

$$\theta(t) = \frac{\partial f^M(0,t)}{\partial t} + a f^M(0,t) + \frac{\sigma_r^2}{2a}\left(1 - e^{-2at}\right), \tag{2.2}$$

where $f^M(0,t) = \frac{\partial P^M(0,t)}{\partial t}$ is the instantaneous forward rate deduced from the discount curve observed. Furthermore, our framework may be extended to most of alternative interest rate models.

Transactions in the interbank market are exposed to the risk of default of counterparties. Thus, we define the default time of the interbank market as a stopping time, denoted by τ. For instance, this default time may occur when a credit event hits one of the interbank market participants or after a succession of credit events hitting interbank market participants. Let $(\mathscr{T}_t)_{t\in\mathbb{R}_+}$ be the filtration generated by τ and defined by:

$$\mathscr{T}_t = \sigma(\{\tau \leq u\} : u \leq t).$$

For a sake of convenience, we assume that τ satisfies the following distributional constraints:

$$\begin{cases} \mathbb{Q}(\tau < +\infty) = 1 \\ \mathbb{Q}(\tau = 0) = 0 \\ \mathbb{Q}(\tau \geq t) > 0 \text{ for every } t \in \mathbb{R}_+ \end{cases}.$$

The default of the interbank market is triggered by the first jump of a doubly stochastic process $(N_t^\tau)_{t\geq 0}$ that has the same intensity process $(\lambda_t)_{t\geq 0}$ of a self-exciting jump process $(N_t)_{t\geq 0}$. The point process $(N_t)_{t\geq 0}$ records the number of negative shocks that impact the interbank market. Whereas $(\lambda_t)_{t\geq 0}$ and $(N_t)_{t\geq 0}$ are dependent, $(\lambda_t)_{t\geq 0}$ and $(N_t^\tau)_{t\geq 0}$ are independent. This intensity process is adapted to the subfiltration $(\mathscr{H}_t)_{t\geq 0}$. The augmented filtration $(\mathscr{G}_t)_{t\geq 0}$ is defined by:

$$\mathscr{G}_t = \mathscr{F}_t \vee \mathscr{H}_t \vee \mathscr{T}_t, \text{ for any } t \in \mathbb{R}_+.$$

Unlike the process N_t^τ, the process N_t is a self-exciting process which is not an inhomogeneous Poisson process conditionally to the filtration of the intensity up to time t. The information about default in the interbank market will be recovered from the term structures of swap rates. According to the definition of a doubly stochastic process, for all $s > t$,

$$\mathbb{Q}\left(N_s^\tau - N_t^\tau = 0 \mid \mathscr{G}_t \vee \mathscr{H}_s\right) = \exp\left(-\int_t^s \lambda_u \mathrm{d}u\right). \tag{2.3}$$

The conditional survival probability is defined by

$$\mathbb{Q}(\tau \geq T \mid \tau \geq t, \mathscr{H}_t) = \mathbb{E}\left(e^{-\int_t^T \lambda_u \mathrm{d}u} | \mathscr{H}_t\right). \tag{2.4}$$

Interconnections between participants in the interbank market can lead to a succession of credit events that are hard to explain if the intensity is modeled by a diffusion process. For this reason, we work with an intensity process ruled by the following stochastic differential equation:

$$d\lambda_t = \kappa(\gamma-\lambda_t)dt+\sigma_\lambda\sqrt{\lambda_t}dW_t^\lambda+\eta dN_t \tag{2.5}$$

where κ is the rate of mean reversion, γ is the long-run level, σ_λ is the volatility coefficient, η is the jump in the intensity when a credit event occurs and W_t^λ is a standard Brownian motion. The diffusion part of this process is the Cox Ingersoll Ross (CIR) model which ensures the non-negativity of the intensity. Since N_t has λ_t for intensity, a negative shock increases the arrival rate of the next negative shock proportionally to η. This intensity process is driven by the SDE (2.5):

$$\lambda_T = \gamma+(\lambda_t-\gamma)e^{-\kappa(T-t)}+\sigma_\lambda\int_t^T e^{-\kappa(T-u)}\sqrt{\lambda_u}dW_u^\lambda +\eta\int_t^T e^{-\kappa(T-u)}dN_u. \tag{2.6}$$

The analytical tractability of this model allows us to deduce its first and second order moments by solving the partial differential equations derived from the Dynkin formula (see Hainaut 2017 for details),

$$\mathbb{E}[\lambda_T\mid\mathscr{G}_t] = \lambda_t-\frac{\kappa\gamma}{\eta-\kappa}\left(e^{-(\eta-\kappa)T}-e^{-(\eta-\kappa)t}\right), \tag{2.7}$$

$$\begin{aligned}\mathbb{E}[\lambda_T^2\mid\mathscr{G}_t] = {} & \lambda_t^2e^{-a_1(T-t)} \\ & +\frac{a_2}{a_1}\left(e^{-a_1(T+t)}-e^{-2a_1T}\right)\left(\lambda_t+\frac{\kappa\gamma}{\eta-\kappa}e^{-(\eta-\kappa)t}\right) \\ & -\frac{a_2\kappa\gamma}{(\eta-\kappa)(3\eta+\kappa)}\left(e^{-a_1(T+t)-(\eta-\kappa)t}-e^{-(5\eta+3\kappa)T}\right),\end{aligned} \tag{2.8}$$

where $a_1=\eta+\kappa$ and $a_2=\left(2\kappa\gamma+\sigma_\lambda^2+\eta^2\right)$. Equation (2.7) involves to choose $\eta\geq\kappa$, in order to ensure that the expected arrival rate of credit events is finite when $T\to\infty$. The figure (2.1) shows the simulated path of processes $(\lambda_t)_{t\geq0}, (N_t)_{t\geq0}$ and $(N_t^\tau)_{t\geq0}$ for $\kappa=1$, $\gamma=1.5$, $\sigma=1$, $\eta=2$ and $\lambda_0=0.75$. From this graph, we see that the default of the interbank market is not directly linked to the default of one of its members. Moreover, a default of an interbank member raises the probability of the arrival of the next default among members as well as the default of interbank market himself. The next proposition presents the moment generating function of the cumulated arrival rate of credit events over a time interval called the compensator and denoted by $\int_t^T\lambda_s ds$.

Proposition 1. *For $\omega\in\mathbb{C}$, let $\psi(t,\lambda_t,N_t)=\mathbb{E}\left(e^{\omega\int_t^T\lambda_s ds}|\mathscr{F}_t\right)$ be the moment generating function of $\int_t^T\lambda_s ds$. There exist two multivariate functions C, and D such that*

$$\psi(t,\lambda_t,N_t)=\exp\left(C(t,T)+D(t,T)\lambda_t\right) \tag{2.9}$$

where C and D are solutions of the following partial differential equations (PDE)

$$\begin{cases}0=\partial_t C(t,T)+\kappa\gamma D(t,T)\\ 0=\partial_t D(t,T)+\omega-\kappa D(t,T)+\frac{1}{2}\sigma_\lambda^2(D(t,T))^2\\ \qquad +\exp(D(t,T)\eta)-1\end{cases}, \tag{2.10}$$

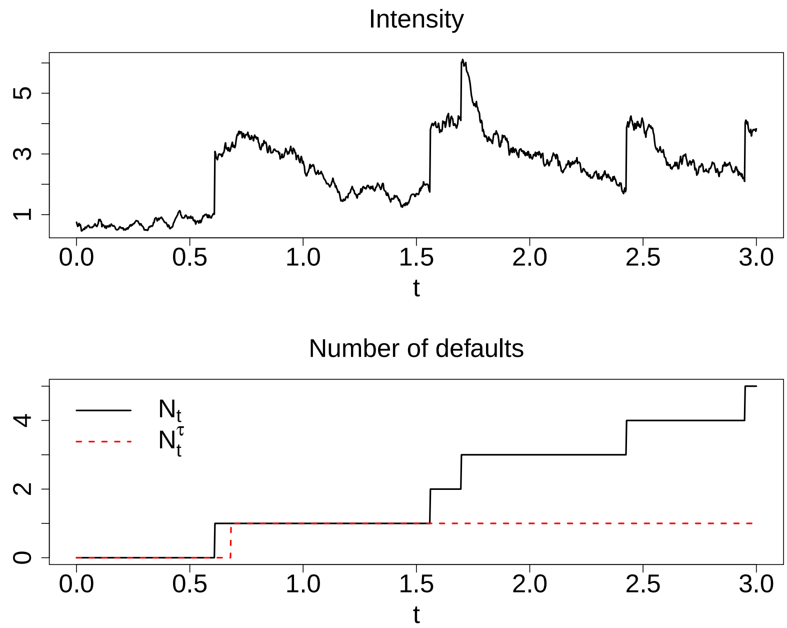


Figure 2.1: Simulated path of the intensity process and the counting processes over a period three years. The intensity parameters considered are $\kappa = 1$, $\gamma = 1.5$, $\sigma = 1$, $\eta = 2$ and $\lambda_0 = 0.75$. The arrival rate of credit events increases with an amplitude of 2 at each jump of the process N_t.

with the terminal conditions $C(T,T) = 0$ and $D(T,T) = 0$.

Proof. Let us define

$$Y_t \quad := \quad \mathbb{E}\left(e^{\omega \int_t^T \lambda_s \mathrm{d}s} | \mathscr{H}_t\right) \tag{2.11}$$

As $\mathscr{H}_t \subset \mathscr{H}_u$ for any $u \geq t$, applying the rule of conditional expectations leads to:

$$\begin{aligned} Y_t &= \mathbb{E}\left(e^{\omega \int_t^u \lambda_s \mathrm{d}s} \mathbb{E}\left(e^{\omega \int_u^T \lambda_s \mathrm{d}s} \,|\, \mathscr{H}_u\right) | \, \mathscr{H}_t\right) \\ &= \mathbb{E}\left(e^{\omega \int_t^u \lambda_s \mathrm{d}s} Y_u \,|\, \mathscr{H}_t\right). \end{aligned} \tag{2.12}$$

Then, the following limit converges to zero:

$$\lim_{u \to t} \frac{\mathbb{E}\left(e^{\omega \int_t^u \lambda_s \mathrm{d}s} Y_u \,|\, \mathscr{H}_t\right) - Y_t}{u - t} \quad = \quad 0. \tag{2.13}$$

By assuming enough regularity and using the first order Taylor approximation of the exponential, this limit is rewritten as follows:

$$\lim_{u \to t} \frac{\mathbb{E}\left(Y_u \,|\, \mathscr{H}_t\right) - Y_t}{u - t} = -\omega \lambda_t Y_t \,, \tag{2.14}$$

in which the left term is the infinitesimal generator of the moment generating function. So as to lighten future calculations, let us denote $f(t,\lambda_t,N_t) := Y_t$ and f_t , f_λ, $f_{\lambda\lambda}$ the partial derivatives of f with respect to time, intensity and intensity twice. The Equation (2.14) is equal to $\mathscr{A} f = \omega\lambda_t f$, or after developments to:

$$\begin{aligned} \mathscr{A} f &= f_t + \kappa(\gamma - \lambda_t) f_\lambda + \frac{1}{2}\sigma_\lambda^2 \lambda_t f_{\lambda\lambda} \\ &\quad + \left(f\left(t,\lambda_t+\eta,N_t+1\right) - f\left(t,\lambda_t,N_t\right)\right)\lambda_t, \end{aligned} \tag{2.15}$$

with terminal condition

$$f(T,\lambda_T,N_T) = 1.$$

We must then solve the next SDE

$$\begin{aligned} -\omega\lambda_t f &= f_t + \kappa(\gamma - \lambda_t) f_\lambda + \frac{1}{2}\sigma_\lambda^2 \lambda_t f_{\lambda\lambda} \\ &\quad + \left(f\left(t,\lambda_t+\eta,N_t+1\right) - f\left(t,\lambda_t,N_t\right)\right)\lambda_t. \end{aligned} \tag{2.16}$$

We do the assumption that $f = \exp\left(C(t,T) + D(t,T)\lambda_t + H(t,T)N_t\right)$. Then we rewrite Equation (2.17) as follows

$$\begin{aligned} 0 &= \left(\partial_t C(t,T) + \partial_t D(t,T)\lambda_t + \partial_t H(t,T) N_t\right) \\ &\quad + \omega\lambda_t + \kappa(\gamma - \lambda_t) D(t,T) \\ &\quad + \frac{1}{2}\sigma_\lambda^2 \lambda_t D(t,T)^2 + \left(\exp\left(D(t,T)\eta + H(t,T)\right) - 1\right)\lambda_t, \end{aligned} \tag{2.17}$$

equivalent to the following PDE system:

$$\begin{cases} 0 = \partial_t C(t,T) + \kappa\gamma D(t,T) = 0 \\ 0 = \partial_t D(t,T) + \omega - \kappa D(t,T) + \frac{1}{2}\sigma_\lambda^2 D(t,T)^2 \\ \qquad + \left(\exp\left(D(t,T)\eta + H(t,T)\right) - 1\right) \\ 0 = \partial_t H(t,T) \end{cases}.$$

Given that $\partial_t H(t,T) = 0$, and $f(T,\lambda_T,N_T) = 1$ then $H = 0$. □

Hereafter, we define the different type of bond prices used in the sequel. Let $\mathscr{L}^F = \{T_{\alpha_1}, T_{\alpha_1+1}, \cdots, T_{\beta_1}\}$ and $\mathscr{L}^S = \{S_{\alpha_2}, S_{\alpha_2+1}, \cdots, S_{\beta_2}\}$ be two collections of time. The distances between two successive times are denoted respectively by δ^F and δ^S for the time structures $\mathscr{L}^F$ and $\mathscr{L}^S$. Let $R \in [0,1]$ be the recovery rate, which is the fraction of the bond's face value repaid to the bondholder in case of default. The price of a default risk-free discount bond at time t with maturity T_k is defined by:

$$P(t,T_k) = \mathbb{E}\left(e^{-\int_t^{T_k} r_s \mathrm{d}s} | \mathscr{F}_t\right). \tag{2.18}$$

From Brigo and Mercurio (2006) part (3.3.2), we know that:

$$P(t,T_k) = \exp\left(A(t,T_k) - B(t,T_k) r_t\right), \tag{2.19}$$

where the functions A and B are given by:

$$A = \ln\left(\frac{P^M(0,T_k)}{P^M(0,t)}\right) + Bf^M(0,t) - \frac{\sigma_r^2\left(1-e^{-2at}\right)}{4a}B^2, \tag{2.20}$$

$$B = \frac{1-\exp(-a(T_k-t))}{a}. \tag{2.21}$$

In practice, the prices of default risk-free discount bonds are obtained by bootstrapping the OIS swap curve, as done by Smith (2013) and implemented in the last section of this chapter.

The price of a defaultable discount bond at time t with maturity T_k is denoted by $P^c(t,T_k)$. Under the assumption that the time of default τ and the risk free rate process r_t are independent, this price is equal to the following product:

$$P^c(t,T_k) = P(t,T_k)\left(R + 1_{\tau\geq t}(1-R)\,\mathbb{Q}(\tau \geq T_k \,|\, \tau \geq t\,,\mathscr{H}_t)\right). \tag{2.22}$$

We now define forward rates and swap rates. The default-free simple forward Libor rate over $[T_k, T_{k+1}]$ as seen from time t is the rate of interest that must apply between times T_k and T_{k+1} so that the price at T_k of a discount bond maturing at T_{k+1} be equal to the forward price. This default-free simple forward Libor is:

$$F(t,T_k,T_{k+1}) = \frac{1}{\delta_k^F}\left(\frac{P(t,T_k)}{P(t,T_{k+1})} - 1\right), \tag{2.23}$$

whereas the defaultable simple forward Libor rate over $[T_k, T_{k+1}]$ seen from time t is:

$$F^c(t,T_k,T_{k+1}) = \frac{1}{\delta_k^F}\left(\frac{P^c(t,T_k)}{P^c(t,T_{k+1})} - 1\right). \tag{2.24}$$

The swap rate associates to an interest rate swap defined on the time structure $\mathscr{L}^F$ and $\mathscr{L}^S$ is the value of the fixed rate such that the corresponding swap value is zero. This swap rate is given by:

$$S_{\alpha,\beta}(t) = \frac{\sum_{i=\alpha_1+1}^{\beta_1} \delta^F P(t,T_i) F^c(t,T_{i-1},T_i)}{\sum_{j=\alpha_2+1}^{\beta_2} P(t,S_j)}. \tag{2.25}$$

We end up this section by a brief review of the modification of the arbitrage-free condition in presence of counterparty risk. A good no-arbitrage relation takes into account the fact that transactions in the interbank market are not riskless. Under the assumption of no-arbitrage, the unitary FRA payoff can be replicated at time t with the following operations: the payer of this FRA borrows

$$\left(1 + K^{FRA}(T_k - T_{k+1})\right) P^c(t,T_{k+1})$$

till T_{k+1}, lends $P^c(t,T_k)$ till T_k and reinvests from T_k to T_{k+1} at the defaultable simple forward Libor. Then, the price of FRA is then

$$FRA(t;T_k;T_{k+1};K) = P^c(t,T_k) - \left(1 + K^{FRA}(T_k - T_{k+1})\right) P^c(t,T_{k+1}), \tag{2.26}$$

where the rate K^{FRA} that cancels the FRA at the inception t is given by

$$K^{FRA} = F^c(t,T_k,T_{k+1}) = \frac{1}{\delta_k^F}\left(\frac{P^c(t,T_k)}{P^c(t,T_{k+1})} - 1\right). \tag{2.27}$$

Therefore, it follows that

$$F^c(t,T_k,T_{k+1}) = \frac{1}{\delta_k^F}\left(\frac{P(t,T_k)\left(R + 1_{\tau\geq t}(1-R)\,\mathbb{Q}(\tau\geq T_k\,|\,\tau\geq t\,,\mathscr{H}_t)\right)}{P(t,T_{k+1})\left(R + 1_{\tau\geq t}(1-R)\,\mathbb{Q}(\tau\geq T_{k+1}\,|\,\tau\geq t\,,\mathscr{H}_t)\right)} - 1\right).$$

In the rest of this chapter, we assume that the recovery rate is null, meaning that transactions are totally non collateralized. Since

$$\frac{\mathbb{Q}\left(\tau\geq T_{k+1}\,|\,\tau\geq t\,,\mathscr{H}_t\right)}{\mathbb{Q}\left(\tau\geq T_k\,|\,\tau\geq t\,,\mathscr{H}_t\right)} = \mathbb{Q}\left(\tau\geq T_{k+1}\,|\,\tau\geq T_k\,,\mathscr{H}_t\right), \tag{2.28}$$

it follows that

$$F^c(t,T_k,T_{k+1}) = \frac{1}{\delta_k^F}\left(\frac{P(t,T_k)}{P(t,T_{k+1})}\frac{1}{\mathbb{Q}\left(\tau\geq T_{k+1}\,|\,\tau\geq T_k\,,\mathscr{H}_t\right)} - 1\right). \tag{2.29}$$

If market FRA rates and theoretical FRA rates are respectively denoted by $F^c(t,T_k,T_{k+1})$ and $F(t,T_k,T_{k+1})$, we observe that
$F^c(t,T_k,T_{k+1}) \geq F(t,T_k,T_{k+1})$ since $\frac{1}{\mathbb{Q}(\tau\geq T_{k+1}\,|\,\tau\geq T_k\,,\mathscr{H}_t)} \geq 1$. With a proof by contradiction, we deduce as Bianchetti (2008) that an equality between our market FRA rates and the theoretical FRA rates leads to an arbitrage. This equality holds only for a transaction with a collateral or an overnight tenor. In the sequel, under the assumption that there is no collateral we work with F^c. Since, FRA contracts and OIS bonds or default risk-free discount bonds form the basic traded assets within the interbank market, the new no arbitrage condition is such that under the risk neutral measure $\mathbb{Q}$ the discounted prices of OIS bonds and FRA contracts are martingales. As we set our models under the martingale measure $\mathbb{Q}$, discounted OIS bond prices are martingales under $\mathbb{Q}$. Furthermore, show that the discounted FRA contract prices are martingales under $\mathbb{Q}$ is equivalent to show that FRA contract prices are martingales under forward measures. Let consider a FRA contract concluded at time t over $[T_k,T_{k+1}]$ where the nominal is equal to one, the fix rate is K^{FRA} and the floating rate is $L^c\left(T_k,T_{k+1}\right)$. The price of this FRA contract is given by :

$$\begin{aligned} FRA\left(t;T_k;T_{k+1};K^{FRA}\right) &= \delta_k^F P\left(t,T_{k+1}\right) \\ &\quad \times\mathbb{E}^{\mathbb{Q}_{T_{k+1}}}\left[L^c\left(T_k,T_{k+1}\right) - K^{FRA}\,|\,\mathscr{G}_t\right], \end{aligned} \tag{2.30}$$

where $\mathbb{Q}_{T_{k+1}}$ is the forward measure. The fix rate K^{FRA} that cancels the FRA at the inception t is then:

$$F^c\left(t,T_k,T_{k+1}\right) = \mathbb{E}^{\mathbb{Q}_{T_{k+1}}}\left[L^c\left(T_k,T_{k+1}\right)\,|\,\mathscr{G}_t\right]. \tag{2.31}$$

From the two last equations the interbank market is arbitrage-free if the ratio $\frac{P^c(t,T_k)}{P^c(t,T_{k+1})}$ is martingale under $\mathbb{Q}_{T_{k+1}}$. This, yields to the following condition:

$$\exp\left(-\eta\Delta_k D(t)\right) - 1 = \frac{-\frac{1}{2}\Delta_k D(t)^2\eta^2 + \sigma_\lambda^2 D(t,T_k)\Delta_k D(t)\lambda_t}{\left(\lambda_t e^{\eta D(t,T_{k+1})} + 1\right)}, \tag{2.32}$$

where $\Delta_k D(t) = D(t,T_{k+1}) - D(t,T_k)$. In appendix (2.A) we detail how we derive this arbitrage-free condition.

2.3 Forward Libor dynamics

The forward Libor rate differs from one tenor to another. This is due to the difference in exposure times to default risk. But also by the demand for loans of each tenor. Then, we subdivide the interbank market into sub-markets according to the tenor of the transactions. The forward Libor in equation (2.29) depends on the tenor representing an interbank market segment. Instead of a multiplicative spread between forward and discounting curves as proposed by Henrard (2014), we have in Equation (2.29) an additional term defined as follows:

$$\beta(t,T_k,T_{k+1}) = \frac{1}{\mathbb{Q}\left(\tau \geq T_{k+1} \,|\, \tau \geq T_k \,,\, \mathscr{H}_t\right)}, \tag{2.33}$$

that is always greater or equal to one. From equations (2.33) and (2.29), when the probability of default over a time period $\left[T_k,T_{k+1}\right]$ rises, both the factor β and the forward Libor rate increases.
In the sequel of this chapter, the term structure of the interbank credit risk for a tenor δ is defined by $\{T \longmapsto \beta(t,T,T+\delta),\ T > t\}$. Hence, at time t, the measure of the credit risk within the interbank market segment associated to the tenor δ for a time period $[T,T+\delta]$ is given by:

$$\begin{aligned}
\beta(t,T,T+\delta) &= \frac{1}{\mathbb{Q}(\tau \geq T+\delta \,|\, \tau \geq T \,,\, \mathscr{H}_t)} \\
&= \frac{\mathbb{Q}(\tau \geq T \,|\, \tau \geq t \,,\, \mathscr{H}_t)}{\mathbb{Q}(\tau \geq T+\delta \,|\, \tau \geq t \,,\, \mathscr{H}_t)} \\
&= \frac{\mathbb{E}\left(e^{-\int_t^T \lambda_u \mathrm{d}u} | \mathscr{H}_t\right)}{\mathbb{E}\left(e^{-\int_t^{T+\delta} \lambda_u \mathrm{d}u} | \mathscr{H}_t\right)}.
\end{aligned}$$

From equation (2.9), we infer that

$$\beta(t,T,T+\delta) = \exp\left(C(t,T) - C(t,T+\delta) + (D(t,T) - D(t,T+\delta))\,\lambda_t\right). \tag{2.34}$$

Thus, the term structure of the interbank risk is fully determined by the bivariates functions C and D. The simple[1] spot forward Libor for $[t,T]$ is given by:

$$\begin{aligned}
L^c(t,T) &= F^c(t,t,T) \\
&= \frac{1}{\delta(t,T)}\left[\frac{1}{P^c(t,T)} - 1\right],
\end{aligned}$$

[1] We work with simple rate because markets use the simple capitalization.

where $P^c(t,T)$ is the defaultable bond price:

$$
\begin{aligned}
P^c(t,T) &= 1_{\tau\geq t}P(t,T)\,\mathbb{Q}\left(\tau\geq T\,|\,\tau\geq t\,,\,\mathscr{H}_t\right)\\
&= 1_{\tau\geq t}\mathbb{E}\left(e^{-\int_t^T r_s \mathrm{d}s}|\mathscr{F}_t\right)\mathbb{E}\left(e^{-\int_t^T \lambda_u \mathrm{d}u}|\mathscr{H}_t\right)\\
&= 1_{\tau\geq t}\exp\left(A(t,T)-B(t,T)r_t+C(t,T)+D(t,T)\lambda_t\right). \qquad (2.35)
\end{aligned}
$$

Hence, if $\tau>t$, the Libor rate is

$$
L^c(t,T)=\frac{1}{\delta\,(t,T)}\left[\exp\left(-A(t,T)-C(t,T)+B(t,T)r_t-D(t,T)\lambda_t\right)-1\right]. \qquad (2.36)
$$

Whereas the simple forward Libor is given by:

$$
\begin{aligned}
F^c(t,T,S) &= \left[\exp\left(A(t,T)-A(t,S)+C(t,T)-C(t,S)+\left(B(t,S)-B(t,T)\right)r_t\right)\right.\\
&\quad\left.\times\exp\left(-\left(D(t,S)-D(t,T)\right)\lambda_t\right)-1\right]\times\frac{1}{\delta\,(T,S)}. \qquad (2.37)
\end{aligned}
$$

In equations (2.36) and (2.37) we note that if D is a bivariate function with negatives values, a jump in the intensity process increases both the simple spot forward Libor and the simple forward Libor. Applying the Itô's lemma for jump processes to equations (2.36) and (2.37) allows us to establish the dynamics of these rates in the next proposition.

Proposition 2. *Under the risk neutral measure,*
(i) the stochastic differential equation of a simple spot Libor is given by

$$
\begin{aligned}
dL^c(t,T) &= \left(\frac{1}{\delta\,(t,T)}+L^c(t,T)\right)\left[\left(H_1\,(t,T)+H_2\,(t,T)\,r_t+H_3\,(t,T)\,\lambda_t\right)dt\right.\\
&\quad+B\,(t,T)\,\sigma_r dW_t^r-\sigma_\lambda D\,(t,T)\,\sqrt{\lambda_t}dW_t^\lambda\\
&\quad\left.+\eta D\,(t,T)\left(\frac{1}{2}\eta D\,(t,T)-1\right)dN_t\right]+\left(L^c\,(t,T)-L^c(t^-,T)\right)dN_t\,, \qquad (2.38)
\end{aligned}
$$

where H_1, H_2 and H_3 are solutions of the following PDE system:

$$
\begin{cases}
H_1\,(t,T)=-\partial_t A\,(t,T)-\partial_t C\,(t,T)+\theta\,(t)\,B\,(t,T)+\frac{1}{2}\sigma_r^2 B^2\,(t,T)\\
\qquad\qquad -\kappa\gamma D\,(t,T)\,,\\
H_2\,(t,T)=\partial_t B\,(t,T)-aB\,(t,T)\,,\\
H_3\,(t,T)=\frac{1}{2}\sigma_\lambda^2 D^2\,(t,T)-\partial_t D\,(t,T)+\kappa D\,(t,T)\,.
\end{cases} \qquad (2.39)
$$

(ii) the defaultable forward Libor dynamics is given by

$$
\begin{aligned}
dF^c(t,T,S) &= \left[\left(R_1\,(t,T,S)+R_2\,(t,T,S)\,r_t+R_3\,(t,T,S)\cdot\lambda_t\right)dt\right.\\
&\quad+\left(B(t,S)-B(t,T)\right)\sigma_r dW_t^r-\sigma_\lambda\left(D(t,S)-D(t,T)\right)\sqrt{\lambda_t}dW_t^\lambda\\
&\quad\left.+\eta\left(D(t,S)-D(t,T)\right)\left(\frac{1}{2}\eta\left(D(t,S)-D(t,T)\right)-1\right)dN_t\right]\\
&\quad\times\left(\frac{1}{\delta}+F^c(t,T,S)\right)+\left(F^c(t,T,S)-F^c(t^-,T,S)\right)dN_t\,, \qquad (2.40)
\end{aligned}
$$

where R_1, R_2 and R_3 are solutions of the following PDE system

$$\begin{cases} R_1(t,T,S) = \partial_t A(t,T) - \partial_t A(t,S) + \partial_t C(t,T) - \partial_t C(t,S) \\ \qquad +\theta(t)(B(t,S)-B(t,T)) + \frac{1}{2}\sigma_r^2(B(t,S)-B(t,T))^2 \\ \qquad\qquad -\kappa\gamma(D(t,S)-D(t,T)), \\ R_2(t,T,S) = (\partial_t B(t,S) - \partial_t B(t,T)) - a(B(t,S)-B(t,T)), \\ R_3(t,T,S) = \frac{1}{2}\sigma_\lambda^2(D(t,S)-D(t,T))^2 - (\partial_t D(t,S) - \partial_t D(t,T)) \\ \qquad\qquad +\kappa(D(t,S)-D(t,T)). \end{cases} \tag{2.41}$$

Proof. (i) We define the function $g \in \mathscr{C}^{1,2,2}$ such as

$$g(t,y,z) = \frac{1}{\delta(t,T)}\left[\exp\left(-A(t,T) - C(t,T) + B(t,T)y - D(t,T)z\right) - 1\right], \tag{2.42}$$

by multidimensional Ito's lemma we can write

$$\begin{aligned} \mathrm{d}L^c(t,T) = {} & \frac{\partial g}{\partial t}(t,r_t,\lambda_t)\mathrm{d}t + \frac{\partial g}{\partial y}(t,r_t,\lambda_t)\mathrm{d}r_t + \frac{\partial g}{\partial z}(t,r_t,\lambda_t)\mathrm{d}\lambda_t \\ & + \frac{1}{2}\frac{\partial^2 g}{\partial y^2}(t,r_t,\lambda_t)\mathrm{d}[r_t,r_t]^c + \frac{1}{2}\frac{\partial^2 g}{\partial z^2}\mathrm{d}[\lambda_t,\lambda_t]^c + \left(L^c(t,T) - L^c(t^-,T)\right)\mathrm{d}N_t \end{aligned} \tag{2.43}$$

with

$$\begin{aligned} & \frac{\partial g}{\partial t}(t,r_t,\lambda_t) = (-A_t - C_t + B_t r_t - D_t\lambda_t)\left(\frac{1}{\delta} + L^c(t,T)\right), \\ & \frac{\partial g}{\partial y}(t,r_t,\lambda_t) = B\left(\frac{1}{\delta} + L^c(t,T)\right), \ \frac{\partial g}{\partial z}(t,r_t,\lambda_t) = -D\left(\frac{1}{\delta} + L^c(t,T)\right), \\ & \frac{\partial^2 g}{\partial y^2}(t,r_t,\lambda_t) = B^2\left(\frac{1}{\delta} + L^c(t,T)\right), \ \frac{\partial^2 g}{\partial z^2}(t,r_t,\lambda_t) = D^2\left(\frac{1}{\delta} + L^c(t,T)\right), \\ & \mathrm{d}[r_t,r_t]^c = \sigma_r^2\mathrm{d}t \ \text{ and } \mathrm{d}[\lambda_t,\lambda_t]^c = \sigma_\lambda^2\lambda_t\mathrm{d}t + \eta^2\mathrm{d}N_t. \end{aligned}$$

Then by replacing each term of equation (2.43) we retrieve the equation (2.38). Similarly, we can proof the point (ii). □

The results in proposition (2) is important for pricing of option on forward Libor. In the next, section we discuss in detail the pricing of some interest rates derivatives.

2.4 The pricing of option on spot forward Libor, and swap

This section illustrates how our model can be used for the pricing of interest rates derivatives under a forward measure or an annuity measure. Let us define by $\Pi(L^c(T,S))$ the payoff paid at time $S \geq T$ by a European option written on $L^c(T,S)$, which is driven by Equation (2.36). The price of such an option under the risk neutral measure is given by:

$$\mathrm{C}(t,r_t,\lambda_t) = \mathbb{E}\left[e^{-\int_t^S r_u \mathrm{d}u}\Pi(L^c(T,S)) \mid \mathscr{G}_t\right]. \tag{2.44}$$

However, due to the dependence between the discount factor $e^{-\int_t^S r_u \mathrm{d}u}$ and the payoff $\Pi\left(L^c(T,S)\right)$, we change of numeraire. This technique allows us to switch to the forward measure and simplifies the calculation of the price. Since under the forward measure the price of any financial assets divided by the numeraire $P(\cdot,S)$ is a martingale, therefore the option price becomes

$$\begin{aligned}
\mathrm{C}(t,r_t,\lambda_t) &= P(t,S)\,\mathbb{E}^{\mathbb{Q}_S}\left[\Pi\left(L^c(T,S)\right) \mid \mathscr{G}_t\right] \\
&= P(t,S)\int_0^{+\infty}\int_{-\infty}^{+\infty} \Pi(y_1,y_2)\, f_{r_T,\lambda_T}(y_1,y_2)\,\mathrm{d}y_1\mathrm{d}y_2 \qquad (2.45)
\end{aligned}$$

where f_{r_T,λ_T} is the joint density function of $r_T \mid \mathscr{G}_t$ and $\lambda_T \mid \mathscr{G}_t$ under the forward measure. The Radon-Nikodym density associated to this change of measure is given by:

$$\frac{\mathrm{d}\mathbb{Q}_S}{\mathrm{d}\mathbb{Q}} = e^{-\int_0^S r_u \mathrm{d}u}\frac{1}{P(0,S)},$$

where

$$\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}_S}{\mathrm{d}\mathbb{Q}} \mid \mathscr{G}_t\right] = e^{-\int_0^t r_u \mathrm{d}u}\frac{P(t,S)}{P(0,S)}.$$

The calculation of the expected payoff under $\mathbb{Q}_S$ requires to determine the joint density function f_{r_T,λ_T}. The joint density function f_{r_T,λ_T} does not admit any closed form expression. However, a numerical solution can be found with the discrete Fourier transform (DFT) method. To implement this method we need the moment generating function of $r_T + \lambda_T$ that is detailed in the next proposition.

Proposition 3. *Under the measure $\mathbb{Q}_S$ the moment generating function of $r_T + \lambda_T$ denoted by $\Phi_{T,S}^{r,\lambda}$ is equal to*

$$\begin{aligned}
\Phi_{T,S}^{r,\lambda}(\omega_1,\omega_2) &= \mathbb{E}^{\mathbb{Q}_S}\left[e^{\omega_1 r_T + \omega_2 \lambda_T} \mid \mathscr{G}_t\right] \\
&= e^{G_1(t,T) - A(t,S) + (G_2(t,T)+B(t,S))r_t + G_3(t,T)\lambda_t}, \qquad (2.46)
\end{aligned}$$

where $(\omega_1,\omega_2)\in\mathbb{C}$, the bivariate functions $A(\cdot,\cdot)$, $B(\cdot,\cdot)$are defined by Equations (2.21). Whereas $G_1(\cdot,\cdot)$, $G_2(\cdot,\cdot)$ and $G_3(\cdot,\cdot)$ are solution of the following PDE system:

$$\begin{cases}
\partial_t G_1(t,T) + \theta(t)\,G_2(t,T) + \kappa\gamma G_3(t,T) + \frac{1}{2}\sigma_r^2\left(G_2(t,T)\right)^2 & = 0 \\
\partial_t G_2(t,T) - aG_2(t,T) & = 1 \\
\partial_t G_3(t,T) - \kappa G_3(t,T) + \frac{1}{2}\sigma_\lambda^2\left(G_3(t,T)\right)^2 + \left(e^{\eta G_3(t,T)} - 1\right) & = 0
\end{cases}, \qquad (2.47)$$

and satisfy the terminal conditions:

$$\begin{cases}
G_1(T,T) & = A(T,S) \\
G_2(T,T) & = \omega_1 - B(T,S) \\
G_3(T,T) & = \omega_2
\end{cases}.$$

Proof. By definition of the forward measure and using the fact that $\mathscr{G}_t \subset \mathscr{G}_T$, the moment generating function of $r_T + \lambda_T$ is given by:

$$\begin{aligned}\phi_{T,S}^{r,\lambda}(\omega_1,\omega_2) &= \frac{\mathbb{E}\left[\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}_S}{\mathrm{d}\mathbb{Q}} \mid \mathscr{G}_T\right] e^{\omega_1 r_T+\omega_2\lambda_T} \mid \mathscr{G}_t\right]}{\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}_S}{\mathrm{d}\mathbb{Q}} \mid \mathscr{G}_t\right]} \\ &= \frac{\mathbb{E}\left[e^{-\int_t^T r_u \mathrm{d}u + A(T,S)+(\omega_1-B(T,S))r_T+\omega_2\lambda_T} \mid \mathscr{G}_t\right]}{e^{A(t,S)-B(t,S)r_t}}. \end{aligned} \tag{2.48}$$

In the sequel of this proof, we focus on the evaluation of

$$Y_t = \mathbb{E}\left[e^{-\int_t^T r_u \mathrm{d}u + A(T,S)+(\omega_1-B(T,S))r_T+\omega\lambda_T} \mid \mathscr{G}_t\right], \tag{2.49}$$

and find its closed form thanks to the infinitesimal generator properties for Markov processes. As in the proposition (1) we can proof that for $Y_t = f(t,r_t,\lambda_t,N_t)$

$$\lim_{v\to t} \frac{\mathbb{E}\left(Y_v \,|\, \mathscr{G}_t\right) - Y_t}{v-t} = r_t Y_t = \mathscr{A} f.$$

Then, using the Itô Lemma the infinitesimal generator is

$$\begin{aligned}\mathscr{A} f &= \partial_t f + (\theta(t) - a r_t)\,\partial_r f + \kappa(\gamma - \lambda_t)\,\partial_\lambda f + \frac{1}{2}\sigma_r^2 \partial_r^2 f + \frac{1}{2}\sigma_\lambda^2 \lambda_t \partial_\lambda^2 f \\ &\quad + \lambda_t \left(f(t,r_t,\lambda_t+\eta,N_t+1) - f(t,r_t,\lambda_t,N_t)\right). \end{aligned} \tag{2.50}$$

Assuming that $f(t,r_t,\lambda_t,N_t) = e^{G_1(t,T)+G_2(t,T)r_t+G_3(t,T)\lambda_t}$ we deduce the following ODEs system

$$\begin{cases} \partial_t G_1(t,T) + \theta(t)\, G_2(t,T) + \kappa\gamma G_3(t,T) + \frac{1}{2}\sigma_r^2 \left(G_2(t,T)\right)^2 & = 0 \\ \partial_t G_2(t,T) - a G_2(t,T) & = 1 \\ \partial_t G_3(t,T) - \kappa G_3(t,T) + \frac{1}{2}\sigma_\lambda^2 \left(G_3(t,T)\right)^2 + \left(e^{\eta G_3(t,T)} - 1\right) & = 0 \end{cases},$$

with the following terminal conditions

$$\begin{cases} G_1(T,T) & = A(T,S) \\ G_2(T,T) & = \omega_1 - B(T,S) \\ G_3(T,T) & = \omega_2 \end{cases}.$$

It is sufficient to replace the value of Y_t in our last expression of this moment generating function to complete the proof. □

The next proposition introduces the discretization framework used in numerical applications to construct the joint density function f_{r_T,λ_T} under the forward measure.

Proposition 4. *Let M be the number of steps used in the two dimensional DFT and $\Delta_{y_l} = \frac{\left(1+1_{\{l=1\}}\right)y_l^{max}}{M-1}$ be this step of discretization. Let us denote $\Delta_{z_l} = \frac{2\pi}{M\Delta_{y_l}}$ and*

$$z_{j_1,j_2} = \left(-\frac{M}{2}\Delta_{z_1} + (j_1-1)\Delta_{z_1},\, -\frac{M}{2}\Delta_{z_2} + (j_2-1)\Delta_{z_2}\right),$$

for $j_l = 1...M$. *Let* $\Phi^{r,\lambda}_{T,S}$ *be the moment generating function of* $r_T + \lambda_T \mid \mathcal{G}_t$ *under the measure* $\mathbb{Q}_S$. *Let defined by* $\Pi\left(L^c(T,S)\right)$ *the payoff paid at time* $S \geq T$ *by a European option written on* $L^c(T,S)$. *Let* $C(t, r_t, \lambda_t)$ *the price of this option. The values of* $f_{r_T,\lambda_T}(\cdot,\cdot)$ *at points*

$$y_{k_1,k_2} = \left(-\frac{M}{2}\Delta_{y_1} + (k_1-1)\Delta_{y_1}, -\frac{M}{2}\Delta_{y_2} + (k_2-1)\Delta_{y_2}\right)$$

are approached by the sum:

$$\begin{aligned} f_{r_T,\lambda_T}\left(y_{k_1,k_2}\right) \approx & \frac{e^{-iM\pi + i\sum_{l=1}^{2}(k_l-1)\pi}}{M^2\Delta_{y_1}\Delta_{y_2}} \\ & \times \sum_{j_1=1}^{M}\sum_{j_2=1}^{M} \Upsilon_{j_1,j_2}\Phi^{r,\lambda}_{T,S}\left(iz_{j_1,j_2}\right) e^{i(j_1-1)\pi} e^{-i\sum_{l=1}^{2}(j_l-1)(k_l-1)\frac{2\pi}{M}}, \end{aligned} \tag{2.51}$$

where $\Upsilon_{j_1,j_2} = \left(\frac{1}{2}\right)^{\left(1_{\{j_1=1\}}+1_{\{j_2=1\}}\right)}\left(1 - 1_{\{j_1\neq 1, j_2\neq 1\}}\right) + 1_{\{j_1\neq 1, j_2\neq 1\}}$ *and* $\Phi^{r,\lambda}_{T,S}$ *is defined in equation (2.46).*

Proof. The joint probability density of $r_T \mid \mathcal{G}_t$ and $\lambda_T \mid \mathcal{G}_t$ is retrieved by calculating the Fourier Transform of the moment generating function $\Phi^{r,\lambda}_{T,S}$ as follows:

$$\begin{aligned} f_{r_T,\lambda_T}(y) &= \frac{1}{(2\pi)^2}\mathscr{F}\left[\Phi^{r,\lambda}_{T,S}(iz)\right](y) \\ &= \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\Phi^{r,\lambda}_{T,S}(iz)\, e^{-iz^t y}\mathrm{d}z_1\mathrm{d}z_2. \end{aligned} \tag{2.52}$$

At points y_{k_1,k_2}, this last integral is approached with the trapezoid rule and leads to the following estimate for f_{r_T,λ_T}:

$$f_{r_T,\lambda_T}\left(y_{k_1,k_2}\right) \approx \frac{1}{(2\pi)^2}\sum_{j_1=1}^{M}\sum_{j_2=1}^{M}\Upsilon_{j_1,j_2}\Phi^{r,\lambda}_{T,S}\left(iz_{j_1,j_2}\right)e^{-iz^t_{j_1,j_2}y_{k_1,k_2}}\Delta_{z_1}\Delta_{z_2}, \tag{2.53}$$

and given that $\Delta_{z_l}\Delta_{y_l} = \frac{2\pi}{M}$

$$\begin{aligned} z^t_{j_1,j_2}y_{k_1,k_2} &= \sum_{l=1}^{2}\left(\frac{M^2}{4}\Delta_{z_l}\Delta_{y_l} - \frac{M}{2}(j_l-1)\Delta_{z_l}\Delta_{y_l} - \frac{M}{2}(k_l-1)\Delta_{z_l}\Delta_{y_l}\right) \\ &\quad + \sum_{l=1}^{2}(j_l-1)(k_l-1)\Delta_{z_l}\Delta_{y_l} \\ &= \sum_{l=1}^{2}\left(\frac{M}{2}\pi - (j_l-1)\pi - (k_l-1)\pi + (j_l-1)(k_l-1)\frac{2\pi}{M}\right), \end{aligned}$$

and

$$\begin{aligned} e^{-iz^t_{j_1,j_2}y_{k_1,k_2}} &= e^{-iM\pi + i\sum_{l=1}^{2}(k_l-1)\pi}e^{i\sum_{l=1}^{2}(j_l-1)\pi} \\ &\quad \times e^{-i\sum_{l=1}^{2}(j_l-1)(k_l-1)\frac{2\pi}{M}}. \end{aligned}$$

□

In particular, based on the previous result we are able to compute the price of caplet. A caplet with price $\mathrm{capl}(t,T,S)$ purchased at date t with expiry date T, pays to the holder the positive difference between a simple spot forward Libor market rate $L^c(T,S)$ and the strike rate κ at settlement date S. If we denote by N the principal of this contract, the caplet price is given by

$$\begin{aligned}\mathrm{capl}(t,T,S) &= N\delta(T,S)P(t,S)\mathbb{E}^{\mathbb{Q}_S}\left[(L^c(T,S)-\kappa)^+ \mid \mathscr{G}_t\right],\\ &\approx N\delta(T,S)P(t,S)\sum_{k_1=1}^{M}\sum_{k_2=1}^{M}\Pi\left(y_{k_1,k_2}\right)f_{r_T,\lambda_T}\left(y_{k_1,k_2}\right)\Delta_{y_1}\Delta_{y_2}. \end{aligned} \tag{2.54}$$

where

$$y_{k_1,k_2}=\left(-\frac{M}{2}\Delta_{y_1}+(k_1-1)\Delta_{y_1},-\frac{M}{2}\Delta_{y_2}+(k_2-1)\Delta_{y_2}\right),$$

$\Delta_{y_l}=\frac{\left(1+1_{\{l=1\}}\right)y_l^{max}}{M-1}$, $\Pi(r_T,\lambda_T)=(L^c(T,S)-\kappa)^+$ and $L^c(t,T)$ satisfies the equation (2.36). Apart from the round-off error, this approximation entails a truncation error and a sampling error which vanish when the number of steps M increases. Moreover, when we consider the following model parameters $a=0.75$, $\sigma_r=0.1$, $r_0=-8.33\times10^{-5}$, $\kappa=0.12$, $\gamma=\lambda_0=0.001$, $\sigma_\lambda=6.9\times10^{-5}$ and $\eta=0.13$, the figure (2.2) emphasizes that the approximated caplet price converges to a value around 0.0076 as from $M=2^8$.

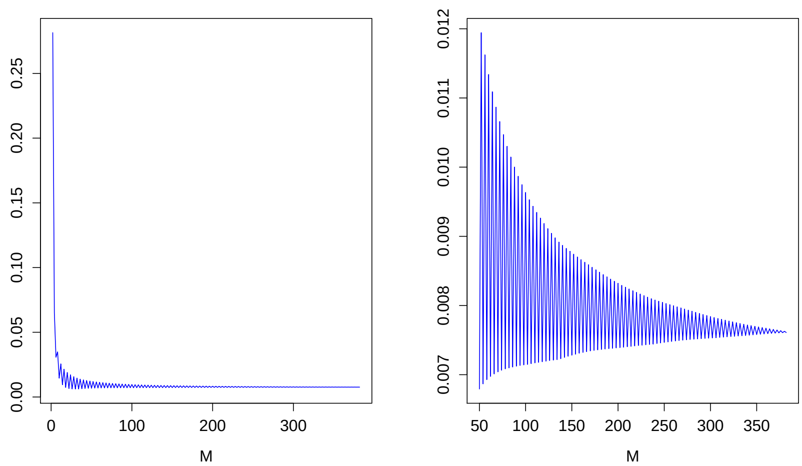


Figure 2.2: Convergence study of the caplet price approximation in equation (2.54). The number of steps M ranges from 2 to 382, and $y_1^{max}=y_2^{max}=1$. The considered model parameters are $a=0.75$, $\sigma_r=0.1$, $r_0=-8.33\times10^{-5}$, $\kappa=0.12$, $\gamma=\lambda_0=0.001$, $\sigma_\lambda=6.9\times10^{-5}$ and $\eta=0.13$. The approximated caplet price converges to a value around 0.0076 as M increases.

We next derive a pricing formula for a cap, which is a portfolio of caplets. Let us consider a cap with nominal N, defined on a set of increasing dates $\{T_0, T_1, T_2, ..., T_n\}$. At each time T_i with $i = 1, ..., n$, the holder of a cap receives the following cash flow

$$N\delta\left(T_{i-1}, T_i\right)\left(L^c(T_{i-1}, T_i) - \kappa\right)^+.$$

Then, the price at time $t \leq T_0$ of a cap is given by the following sum

$$\begin{aligned}
\operatorname{cap}(t) &= \sum_{i=1}^{n} \operatorname{capl}(t, T_{i-1}, T_i) \\
&= N\sum_{i=1}^{n} \delta\left(T_{i-1}, T_i\right) P\left(t, T_i\right) \mathbb{E}^{\mathbb{Q}_{T_i}}\left[\left(L^c(T_{i-1}, T_i) - \kappa\right)^+ \mid \mathscr{G}_t\right] \\
&\approx N\sum_{i=1}^{n} P\left(t, T_i\right) \delta\left(T_{i-1}, T_i\right) \sum_{k=1}^{M}\sum_{l=1}^{M} \Pi\left(x_k, y_l\right) f_{r_{T_{i-1}}, \lambda_{T_{i-1}}}\left(x_k, y_l\right) \Delta_x \Delta_y.
\end{aligned} \tag{2.55}$$

Since interest rates may be negative, we use the Bachelier normal model for quoting interest rate caps in terms of implied volatility. Within this model the Euro cap price at time t is defined as follows:

$$\begin{aligned}
\operatorname{cap}^{\text{Normal}}\left(t, \mathscr{L}^F, \mathscr{L}^S, \delta^F, N, K, \sigma_{0,n}\right) &= N\sum_{i=1}^{n} \delta_i^F P\left(t, T_i\right)\left(F^c\left(t, T_{i-1}, T_i\right) - K\right)\Phi(d) \\
&\quad + \sigma_{0,n} N \sum_{i=1}^{n} \delta_i^F P\left(t, T_i\right)\sqrt{T_{i-1}}\varphi(d)
\end{aligned} \tag{2.56}$$

where $d = \frac{F^c(t, T_{i-1}, T_i) - K}{\sigma_{0,n}\sqrt{T_{i-1}}}$, and $\sigma_{0,n}$ is the implied volatility parameter.

In a similar way, we derive a pricing formula for a swaption. By definition, a swaption is an option that gives the right to enter into a payer or receiver swap at some future date. Let us consider an interest rate swap with m annual fixed payments $\{S_0, S_1, S_2, ..., S_m\}$ and the floating leg payments on set of time $\{T_0, T_1, T_2, ..., T_n\}$. The value of a payer swaption at time T_0 is given by

$$\begin{aligned}
\operatorname{PS}_{0,n}\left(T_0\right) &= \max\left(\operatorname{IRS}\left(T_0\right), 0\right) \\
&= N\sum_{i=1}^{n} \delta^F\left(T_{i-1}, T_i\right) P\left(T_0, T_i\right)\left(S_{0,n}\left(T_0\right) - \kappa\right)^+,
\end{aligned} \tag{2.57}$$

where the swap rate $S_{0,n}(t)$ is given by the equation (2.25). Then, at time $t \leq T_0$

$$\operatorname{PS}_{0,n}(t) = N\mathbb{E}^{\mathbb{Q}^{IRS}}\left[\left(S_{0,n}\left(T_0\right) - \kappa\right)^+ \mid \mathscr{G}_t\right]\sum_{i=1}^{n} \delta^F\left(T_{i-1}, T_i\right) P\left(t, T_i\right), \tag{2.58}$$

where $\mathbb{Q}^{IRS}$ is the equivalent martingale measure of $\mathbb{Q}$, obtained by choosing

$$\sum_{i=1}^{n} \delta^F\left(T_{i-1}, T_i\right) P\left(\cdot, T_i\right)$$

as a numeraire. In order to calculate the expectation in the previous formula, we need to determine the swap rate dynamics under the measure $\mathbb{Q}^{IRS}$.
Let $\Pi : \left(r_{T_0}, \lambda_{T_0}\right) \longmapsto \Pi\left(r_{T_0}, \lambda_{T_0}\right)$ be a function such that:

$$\begin{aligned}
\Pi\left(r_{T_0}, \lambda_{T_0}\right) &= \left(S_{0,n}\left(T_0\right) - \kappa\right)^+ \\
&= \left(\frac{\sum_{i=1}^n \delta^F P\left(T_0, T_i\right) F^c\left(T_0, T_{i-1}, T_i\right)}{\sum_{j=1}^m P\left(T_0, S_j\right)} - \kappa\right)^+ \\
&= \left(\frac{\sum_{i=1}^n \left(P(T_0, T_{i-1}) \beta\left(T_0, T_{i-1}, T_i\right) - P\left(T_0, T_i\right)\right)}{\sum_{j=1}^m P\left(T_0, S_j\right)} - \kappa\right)^+ .
\end{aligned} \tag{2.59}$$

It follows that

$$\mathbb{E}^{\mathbb{Q}^{IRS}}\left[\Pi\left(r_{T_0}, \lambda_{T_0}\right) \mid \mathscr{G}_t\right] = \int_0^{+\infty} \int_{-\infty}^{+\infty} \Pi(x, y) f_{r_{T_0}, \lambda_{T_0}}(x, y)\, \mathrm{d}x\mathrm{d}y, \tag{2.60}$$

where $f_{r_{T_0}, \lambda_{T_0}}$ is the joint density function of $r_{T_0} \mid \mathscr{G}_t$ and $\lambda_{T_0} \mid \mathscr{G}_t$ under the measure $\mathbb{Q}^{IRS}$. Let $E_t = \sum_{i=1}^n \delta^F\left(T_{i-1}, T_i\right) P\left(t, T_i\right)$ be a numeraire chosen for defining the measure $\mathbb{Q}^{IRS}$. The Radon-Nykodym density associated is defined by:

$$\frac{\mathrm{d}\mathbb{Q}^{IRS}}{\mathrm{d}\mathbb{Q}} = e^{-\int_0^{T_0} r_u \mathrm{d}u} \frac{E_{T_0}}{E_0},$$

where

$$\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}^{IRS}}{\mathrm{d}\mathbb{Q}} \mid \mathscr{G}_t\right] = e^{-\int_0^t r_u \mathrm{d}u} \frac{E_t}{E_0}.$$

In the next proposition, we use this Radon-Nykodym density to find the moment generating function of $r_{T_0} + \lambda_{T_0}$ under the measure $\mathbb{Q}^{IRS}$.

Proposition 5. *Under the measure $\mathbb{Q}^{IRS}$ the moment generating function of $r_{T_0} + \lambda_{T_0}$ denoted by $\Phi_{T0}^{r,\lambda}$ is equal to*

$$\begin{aligned}
\Phi_{T0}^{r,\lambda}\left(\omega_1, \omega_2\right) = \mathbb{E}^{\mathbb{Q}^{IRS}}&\left[e^{\omega_1 r_{T_0} + \omega_2 \lambda_{T_0}} \mid \mathscr{G}_t\right] = \\
&\frac{1}{E_t} \sum_{i=1}^n \delta^F\left(T_{i-1}, T_i\right) e^{U_1^i(t, T_0) + U_2^i(t, T_0) r_t + U_3^i(t, T_0) \lambda_t},
\end{aligned} \tag{2.61}$$

where $U_j^i(\cdot,\cdot)$, for $j = 1,2,3$ are bivariate functions satisfying the following PDE system:

$$\begin{cases}
\partial_t U_1^i\left(t, T_0\right) + \theta(t) U_2^i\left(t, T_0\right) + \kappa\gamma U_3^i\left(t, T_0\right) + \frac{1}{2}\sigma_r^2\left(U_2^i\left(t, T_0\right)\right)^2 & = 0 \\
\partial_t U_2^i\left(t, T_0\right) - a U_2^i\left(t, T_0\right) & = 1 \\
\partial_t U_3^i\left(t, T_0\right) - \kappa U_3^i\left(t, T_0\right) + \frac{1}{2}\sigma_\lambda^2\left(U_3^i\left(t, T_0\right)\right)^2 + \left(e^{\eta U_3^i(t, T_0)} - 1\right) & = 0
\end{cases}, \tag{2.62}$$

and satisfy the following terminal conditions:

$$\begin{cases} U_1^i(T_0,T_0) &= A(T_0,T_i) \\ U_2^i(T_0,T_0) &= \omega_1 - B(T_0,T_i) \\ U_3^i(T_0,T_0) &= \omega_2 \end{cases}.$$

Proof. As we only know the dynamics of these processes under the risk neutral, we first rewrite the moment generating function as function of expectation under the risk neutral world. Indeed,

$$\begin{aligned} \Phi_{T0}^{r,\lambda}(\omega_1,\omega_2) &= \frac{\mathbb{E}\left[\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}^{IRS}}{\mathrm{d}\mathbb{Q}} \mid \mathscr{G}_T\right] e^{\omega_1 r_{T_0}+\omega_2\lambda_{T_0}} \mid \mathscr{G}_t\right]}{\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}^{IRS}}{\mathrm{d}\mathbb{Q}} \mid \mathscr{G}_t\right]} \\ &= \frac{1}{E_t}\sum_{i=1}^{n}\delta^F\left(T_{i-1},T_i\right) \\ &\times \mathbb{E}\left[e^{-\int_t^{T_0} r_u \mathrm{d}u + A(T_0,T_i)+(\omega_1-B(T_0,T_i))r_{T_0}+\omega_2\lambda_{T_0}} \mid \mathscr{G}_t\right]. \end{aligned}$$

Next, we focus on the evaluation of

$$Y_t^i = \mathbb{E}\left[e^{-\int_t^{T_0} r_u \mathrm{d}u + A(T_0,T_i)+(\omega_1-B(T_0,T_i))r_{T_0}+\omega_2\lambda_{T_0}} \mid \mathscr{G}_t\right], \tag{2.63}$$

and find its closed form thanks to the infinitesimal generator properties for Markov processes. As in the proposition (1) we can proof that for $Y_t^i = f_i(t,r_t,\lambda_t,N_t)$,

$$\lim_{v\to t}\frac{\mathbb{E}\left(Y_v^i|\mathscr{G}_t\right)-Y_t^i}{v-t} = r_t Y_t^i = \mathscr{A} f_i.$$

Then, using Itô Lemma the infinitesimal generator is expressed as follows

$$\begin{aligned} \mathscr{A} f_i &= \partial_t f_i + (\theta(t)-ar_t)\partial_r f_i + \kappa(\gamma-\lambda_t)\partial_\lambda f_i + \frac{1}{2}\sigma_r^2\partial_r^2 f_i + +\frac{1}{2}\sigma_\lambda^2\lambda_t\partial_\lambda^2 f_i \\ &+ \lambda_t\left(f_i(t,r_t,\lambda_t+\eta,N_t+1)-f_i(t,r_t,\lambda_t,N_t)\right). \end{aligned} \tag{2.64}$$

Assuming that $f_i(t,r_t,\lambda_t,N_t) = e^{U_1^i(t,T_0)+U_2^i(t,T_0)r_t+U_3^i(t,T_0)\lambda_t}$ we deduce the following ODEs system

$$\begin{cases} \partial_t U_1^i(t,T_0) + \theta(t)U_2^i(t,T_0) + \kappa\gamma U_3^i(t,T_0) + \frac{1}{2}\sigma_r^2\left(U_2^i(t,T_0)\right)^2 &= 0 \\ \partial_t U_2^i(t,T_0) - aU_2^i(t,T_0) &= 1 \\ \partial_t U_3^i(t,T_0) - \kappa U_3^i(t,T_0) + \frac{1}{2}\sigma_\lambda^2\left(U_3^i(t,T_0)\right)^2 + \left(e^{\eta U_3^i(t,T_0)}-1\right) &= 0 \end{cases},$$

with the following terminal conditions

$$\begin{cases} U_1^i(T_0,T_0) &= A(T_0,T_i) \\ U_2^i(T_0,T_0) &= \omega_1 - B(T_0,T_i) \\ U_3^i(T_0,T_0) &= \omega_2 \end{cases}.$$

Thus,

$$\Phi_{T0}^{r,\lambda}(\omega_1,\omega_2) = \frac{1}{E_t}\sum_{i=1}^{n}\delta^F\left(T_{i-1},T_i\right)e^{U_1^i(t,T_0)+U_2^i(t,T_0)r_t+U_3^i(t,T_0)\lambda_t}. \tag{2.65}$$

□

Then, by a convenient discretization framework as in proposition (4) we find numerically the joint probability density $f_{r_{T_0},\lambda_{T_0}}(\cdot,\cdot)$ of $r_{T_0}\mid\mathscr{G}_t$ and $\lambda_{T_0}\mid\mathscr{G}_t$ under the measure $\mathbb{Q}^{IRS}$. Hence, equation (2.60) becomes

$$\mathbb{E}^{\mathbb{Q}^{IRS}}\left[\Pi\left(r_{T_0},\lambda_{T_0}\right)\mid\mathscr{G}_t\right] \approx \sum_{k_1=1}^{M}\sum_{k_2=1}^{M}\Pi\left(y_{k_1,k_2}\right)f_{r_{T_0},\lambda_{T_0}}\left(y_{k_1,k_2}\right)\Delta y_1\Delta y_2. \tag{2.66}$$

From equations (2.58) and (2.66), the price of the European swaption defined above is approximated as follows

$$\begin{aligned}\mathrm{PS}_{0,n}(t) \approx N&\left(\sum_{k_1=1}^{M}\sum_{k_2=1}^{M}\Pi\left(y_{k_1,k_2}\right)f_{r_{T_0},\lambda_{T_0}}\left(y_{k_1,k_2}\right)\Delta y_1\Delta y_2\right)\\ &\times\sum_{i=1}^{n}\delta^F\left(T_{i-1},T_i\right)P(t,T_i).\end{aligned} \tag{2.67}$$

Then, based on Bachelier normal model, the implied volatility of this swaption satisfies the following equation

$$\mathrm{PS}^{\mathrm{Normal}}\left(t,\mathscr{L}^F,\mathscr{L}^S,\delta^F,N,K,\sigma_{0,n}\right) = \mathrm{PS}_{0,n}(t),$$

where

$$\begin{aligned}\mathrm{PS}^{\mathrm{Normal}}\left(t,\mathscr{L}^F,\mathscr{L}^S,\delta^F,N,K,\sigma_{0,n}\right) = N&\left[\left(S_{0,n}(t)-K\right)\Phi(d)+\sigma_{0,n}\sqrt{T_0}\varphi(d)\right]\\ &\times\sum_{i=1}^{n}\delta_i^F P(t,T_i),\end{aligned} \tag{2.68}$$

$d = \frac{S_{0,n}(t)-K}{\sigma_{0,n}\sqrt{T_0}}$ and $\sigma_{0,n}$ is the implied volatility parameter.

2.5 Calibration and illustration

In this section, we calibrate our model and illustrate the impact of parameters defining the level of self-excitation on the conditional survival probability and forward Libor rates. Considering a bi-curve framework formed by the discount rate and rates with six months tenor, estimated parameters found hereafter, come from a calibration process detailed throughout this section. Since the short rate is independent from the default intensity, short rate parameters are found in a way to fit at best the discount curve and the implied volatility surface of options with OIS discounting. The data treatment and the bootstraping procedure for interest rates in a multiple curve framework is based on the work of Hull and White (2015) and Ametrano and Bianchetti (2013). For sake of convenience OISs rates of tenor up to a year, have a single fixed payment on the days after its maturity date. While, OISs of tenors greater than one year, have fixed payment annually. Then, the default free bond price with maturity date T_β is given by :

$$P\left(t,T_\beta\right)=\begin{cases}\frac{1}{1+\delta\left(t,T_\beta\right)S_{0,\beta}^{\mathrm{OIS}}(t)} & \text{if } t<T_\beta\leq t+1\\ \left(1-S_{0,\beta}^{\mathrm{OIS}}(t)\sum_{j=1}^{\beta-1}P(t,T_j)\right)\frac{1}{1+S_{0,\beta}^{\mathrm{OIS}}(t)} & \text{if } T_\beta>t+1\end{cases}, \tag{2.69}$$

where $S_{0,\beta}^{\mathrm{OIS}}(t)$ is the swap rate for collateralized transaction and using OIS discounting, with reset date T_0, maturity date T_β, and fixed payment times schedule $\left\{T_0,T_1,\ldots,T_\beta\right\}$. The value of the discount factor for any other intermediate maturity is obtained by interpolation of zero coupon rates, with monotonic cubic spline method. This interpolation method leads to the following closed expressions of the observed discount factor and instantaneous forward rate for other maturity $T\in\left[T_i,T_{i+1}\right]$ with $i\in[0,\beta]$:

$$P^M(t,T)=a_i+b_i(T-T_i)+c_i(T-T_i)^2+d_i(T-T_i)^3\ , \tag{2.70}$$

$$f^M(t,T)=-\frac{b_i+2c_i(T-T_i)+3d_i(T-T_i)^2}{P^M(t,T)}\ , \tag{2.71}$$

where $a_i,b_i,c_i\in\mathbb{R}$. Notice that the function $\theta(\cdot)$ in equation (2.2) depend only on the parameters set $\{a,\sigma_r,r_0\}$ and fits the observed discount curve derived from the OISs rates curve. The last step of our short rate calibration, consists to estimate the parameters set $\{a,\sigma_r,r_0\}$ minimizing the sum of squared errors between observed implied volatility surface of swaptions using OIS discounting and those given by our model as described hereafter.

Swaptions for collateralized transaction and using OIS discounting refer to the single curve framework. Here we do not use the Libor rate instead OIS rate or mixed them in our calculation. The price of this swaption at the inception $t\geq 0$ is given by:

$$\begin{aligned}\mathrm{PS}_{\alpha,\beta}^{\mathrm{OIS}}(t)&=N\mathbb{E}^{\mathbb{Q}^{IRS}}\left[\left(S_{\alpha,\beta}^{\mathrm{OIS}}(T_\alpha)-\kappa\right)^+\mid\mathscr{G}_t\right]\sum_{i=\alpha+1}^{\beta}\delta^F(T_{i-1},T_i)P(t,T_i)\\&=N\left(\sum_{i=\alpha+1}^{\beta}\delta^F(T_{i-1},T_i)P(t,T_i)\right)\int_{-\infty}^{+\infty}\Pi(x)f_{r_{T_\alpha}}(x)\,\mathrm{d}x,\end{aligned} \tag{2.72}$$

where

$$\Pi:\ x\longmapsto\Pi(x)=\left(\frac{1-e^{A\left(T_\alpha,T_\beta\right)-B\left(T_\alpha,T_\beta\right)x}}{\sum_{i=\alpha+1}^{\beta}\delta^F(T_{i-1},T_i)\,e^{A(T_\alpha,T_i)-B(T_\alpha,T_i)x}}-\kappa\right)^+,$$

and $f_{r_{T_\alpha}}$ is the probability density function of r_{T_α} under measure $\mathbb{Q}^{\mathrm{IRS}}$. In appendix (2.B) we find a close-form approximation of this density function. Then, it follows

	Short rate process		
Model	a	σ_r	r_0
HW	$7.487799e-01$	$9.851328e-02$	$-8.337071e-05$

Table 2.1: Estimated short rate parameters based on the term structure of OIS rate, and the implied volatility surface of option on OIS rates observed within the interbank market the 11 February 2019.

that the estimated parameters set of $\{a, \sigma_r, r_0\}$ minimizes $\sum_{(\alpha,\beta)} \left(\sigma_{\alpha,\beta}^{\text{Market}} - \sigma_{\alpha,\beta}^{\text{Model}}\right)^2$, where $\sigma_{\alpha,\beta}^{\text{Model}}$ satisfies the following equation:

$$\text{PS}^{\text{Normal}}\left(t, \mathscr{L}^F, \mathscr{L}^S, \delta^F, N, K, \sigma_{\alpha,\beta}^{\text{Model}}\right) = \text{PS}_{\alpha,\beta}^{\text{OIS}}(t).$$

It remains to estimate the parameters set $\{\kappa, \gamma, \eta, \sigma_\lambda, \lambda_0\}$ of our intensity process. This parameters set is estimated by minimizing the sum of squared errors between the observed IRS six months curve and those given by our model. By doing this, we expect to capture the pattern of the spread OIS-IRS six months observed in the corresponding interbank market segment, or the credit risk within the interbank market where products with six months tenor are exchanged.
We complete our calibration process based on Bloomberg data. These data include the term structure of OIS rate, the implied volatility surface of swaptions with OIS discounting and collateral, and the term structure of IRS six months observed within the interbank market on 11 February 2019. The estimated parameters found for different models are in the following tables (2.1,2.2). Although, a weak value of λ_0 and σ_λ can lead to refute the presence of self-excitation, these values probably simply reflect the current state of the market or that the market does not include this risk in the pricing. In this context, compared to a CIR intensity process, a Hawkes intensity process with an exponential decay, or a deterministic mean-reverting (DMR[2]) function, the best fit of the IRS six months curve as shown in table (2.3) (see also figure (2.5)) is obtained with our self-exciting jump process denoted by SEJCIR[3] and the Hawkes process. In term of OIS-IRS 6 months spread, the SEJCIR model allows for a better fit of this spread than the other approaches. As shown in the figure (2.6), the SEJCIR captures global changes in the dynamic of this OIS-IRS six months spread with a mean squared errors below 0.1 basis points.

[2]Deterministic Mean reverting
[3]SEJCIR stands to our intensity process which is a CIR process plus a double stochastic jump process

Model	Intensity process κ	γ	σ_λ	η	λ_0
CIR	0.0763197	0.0017700	0.0000739	0	0.0015061
Hawkes	0.1242973	0.0010715	0	0.1243053	0.0013269
DMR	0.088905	0.003889	0	0	0
SEJCIR	0.12463426	0.00107537	0.00006976	0.12466822	0.00132491

Table 2.2: Estimated intensity process parameters based on the term structure of OIS rate and IRS 6-months observed within the interbank market the 11 February 2019.

	CIR	Hawkes	DMR	SEJCIR
RSE	$3.896808e-04$	$5.994997e-05$	$2.07622e-02$	$5.977578e-05$

Table 2.3: Relative squared errors when the intensity model is either a CIR process, a Hawkes process, a deterministic mean reverting function or a SEJCIR model.

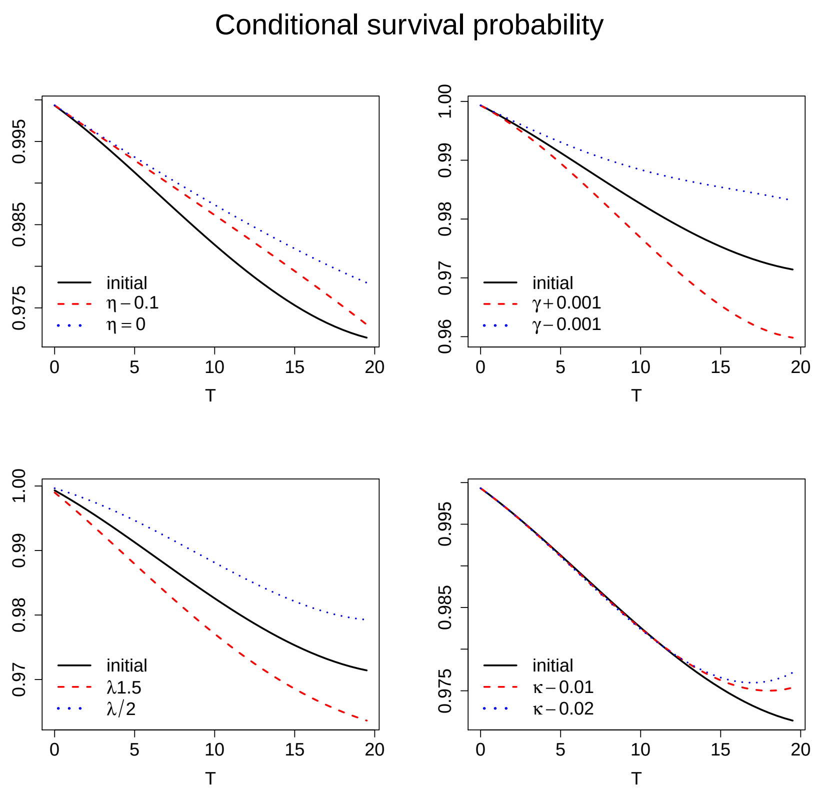


Figure 2.3: Conditional survival probability of the six months interbank market segment $\mathbb{Q}(\tau \geq T \mid \tau \geq 0, \mathscr{H}_0)$ for $T \in [0,20]$. For each plot, we consider the initial set of SEJCIR estimated parameters and we vary one of them. This conditional survival probability decreases faster when the self-excitation parameter increases.

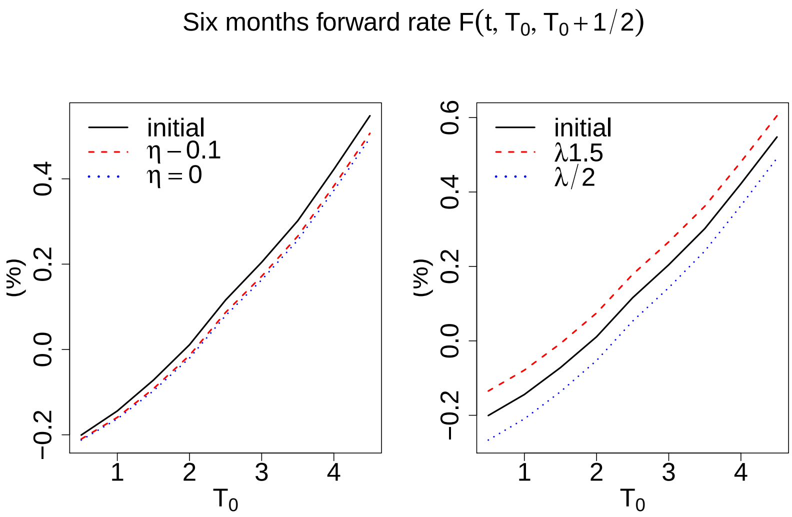


Figure 2.4: Six months forward rate sensitivity on self-excitation parameters η and λ_0. Given a reset date T_0, the forward Libor increases when the initial arrival rate of credit events λ_0 or the amplitude of credit events η increases.

To wrap up this section, based on the estimated parameters of the SEJCIR model in table (2.2), we analyze the impact of the self-excitation parameter on default probabilities. For this purpose, we draw in figure (2.3) the conditional survival probability of six months interbank market segment. This probability decreases slowly and remains close to one for maturity up to 20 years. However, a change in the self-excitation parameter impacts the rate of decreasing. From the graph at the top left, the conditional survival probability rises for long term when the jump size decreases. This result comes from the fact that a credit event is less likely. A similar conclusion can be drawn from the graph at the top right and bottom left, a decrease in the long-run level γ or the initial value λ_t makes credit events less likely in this interbank market segment. On the contrary, the bottom right graph reveals that when the rate of mean reversion rises credit events are more frequent and hint a decrease of the conditional survival probability. Thus, the self-excitation property in our framework increases the conditional survival probability when a credit events occur.

In figure (2.4) we draw the six months forward rates for different reset dates, when changing self-excitation parameters. As the uncertainty rises with the reset date, the six months forward rates curve increases. From the graphs, when the jump size and the initial value of arrival rate of default rise the value of the six months forward rate increases. That is coherent with our model, since from equation (2.29), the forward rate increases when the survival default probability decrease. Then, the self-excitation property in our framework increases the forward Libor rate when credit events occur.

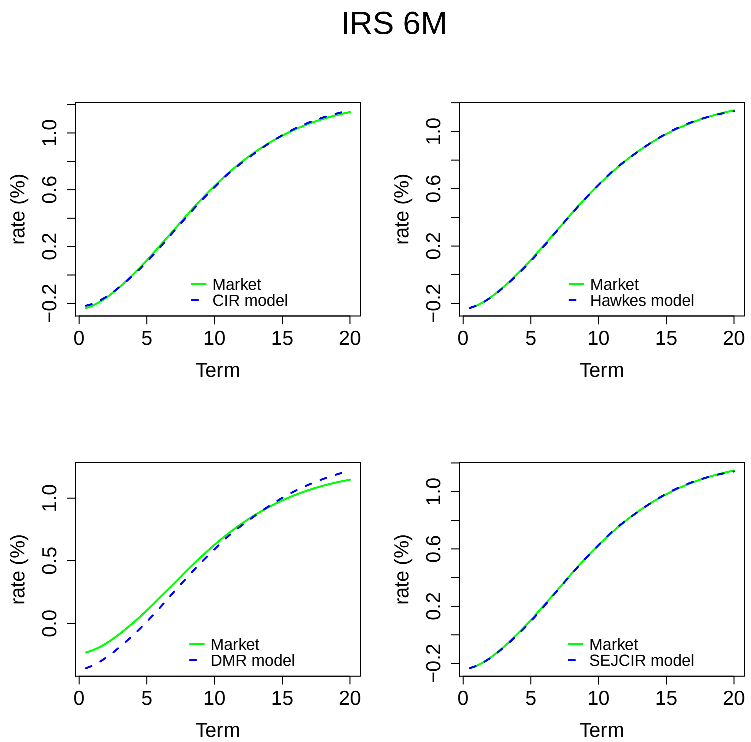


Figure 2.5: IRS curve with six months tenor for maturity from 6 months to 20 years observed in the market on 11/02/2019, and from the CIR model, the Hawkes model, the DMR model, and the SEJCIR with the estimated parameters of the tables (2.1,2.2).

2.6 Conclusions

In this chapter, we have explored a new use of self-exciting processes to model the interbank credit risk in a multiple curve framework. In particular, we have segmented the interbank market according to tenor. Further, we have modeled the credit risk of each segment using an intensity-based approach, where the intensity is a self-exciting jump Cox Ingersoll Ross (SEJCIR) process. In a realistic way, we have used the self-exciting feature to introduce contagion between the interbank market participants.

In this setting, the forward Libor rate is a non-decreasing function of an interbank risk factor that modified the arbitrage-free condition. Thanks to the tractability and flexibility of the SEJCIR process, we have obtained the arbitrage-free condition, and we have derived closed-form approximation of pricing formulas for caplets, caps, as well as swaptions.

Numerical applications have shown that the SEJCIR process enables to reproduce the IRS curve and to explain the deviance between the IRS and OIS curve as a result of the interbank credit risk. The results obtained in the sensitivity analysis confirm the importance of the interbank credit risk to reflect the confidence level within the interbank market. In our view, this framework can be extended to incorporate the economic regime, so that this confidence level varies more in a crisis period.

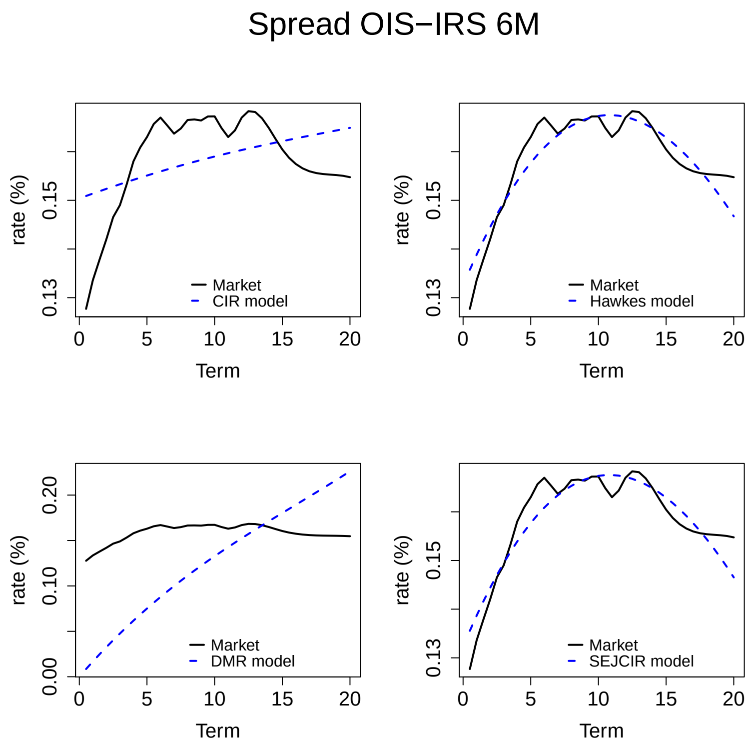


Figure 2.6: Spread between the OIS curve and the IRS curve with six months tenor for maturity from 6 months to 20 years observed in the market on 11/02/2019, and from the CIR model, the Hawkes model, the DMR model and the SEJCIR with the estimated parameters of the tables (2.1,2.2).

Appendices

2.A Arbitrage free condition

The arbitrage-free condition is such that the ratio $\frac{P^c(t,T_k)}{P^c(t,T_{k+1})}$ is martingale under $\mathbb{Q}_{T_{k+1}}$. The Radon-Nikodym density associated to the change of measure $\mathbb{Q}$ to $\mathbb{Q}_{T_{k+1}}$ is defined as follows:

$$\Lambda_t := \frac{\mathrm{d}\mathbb{Q}_{T_{k+1}}}{\mathrm{d}\mathbb{Q}} \mid \mathscr{G}_t = e^{-\int_0^t r_u \mathrm{d}u} \frac{P\left(t, T_{k+1}\right)}{P\left(0, T_{k+1}\right)}.$$

Then, the ratio $\frac{P^c(t,T_k)}{P^c(t,T_{k+1})}$ is martingale under $\mathbb{Q}_{T_{k+1}}$ if $\Lambda_t \frac{P^c(t,T_k)}{P^c(t,T_{k+1})}$ is a martingale under the risk neutral measure. We have

$$\begin{aligned} \mathrm{d}\left(\Lambda_t \frac{P^c(t,T_k)}{P^c(t,T_{k+1})}\right) &= \mathrm{d}\left(e^{-\int_0^t r_u \mathrm{d}u} \frac{P(t,T_k)}{P(0,T_{k+1})} \beta(t,T_k,T_{k+1})\right) \\ &= e^{-\int_0^t r_u \mathrm{d}u} \frac{P(t,T_k)}{P(0,T_{k+1})} \left(\mathrm{d}\beta(t,T_k,T_{k+1})\right. \\ &\quad \left. -\sigma_r B(t,T_k)\beta(t,T_k,T_{k+1})\mathrm{d}W_t^r\right), \end{aligned} \tag{2.73}$$

where the second line comes from the fact that $e^{-\int_0^t r_u \mathrm{d}u} P(t,T_k)$ is a martingale under $\mathbb{Q}$ and $\mathrm{d}[W^r, W^\lambda]_t = 0$. Note that $\beta(t,T_k,T_{k+1})$ is defined in equation (2.34). Next, given the Ito differentiation rule, we obtain :

$$\begin{aligned} \frac{\mathrm{d}\beta(t,T_k,T_{k+1})}{\beta(t,T_k,T_{k+1})} &= \mu_t^\beta \mathrm{d}t + \sigma_\lambda \sqrt{\lambda_t} \Delta_k D(t) \mathrm{d}W_t^\lambda \\ &\quad + \left(\exp\left(-\eta \Delta_k D(t)\right) + \frac{1}{2}\Delta_k D(t)^2 \eta^2 - 1\right) \\ &\quad \times \left(\mathrm{d}N_t - \lambda_t \mathrm{d}t\right), \end{aligned} \tag{2.74}$$

where $\Delta_k D(t) = D(t,T_{k+1}) - D(t,T_k)$ and

$$\begin{aligned} \mu_t^\beta &= \left(e^{\eta D(t,T_k)} - e^{\eta D(t,T_{k+1})} - \sigma_\lambda^2 D(t,T_k)\Delta_k D(t)\right)\lambda_t \\ &\quad + \exp\left(-\eta \Delta_k D(t)\right) + \frac{1}{2}\Delta_k D(t)^2 \eta^2 - 1. \end{aligned} \tag{2.75}$$

From equations (2.73,2.74) the market is arbitrage-free if the drift $\mu_t^\beta = 0$. The arbitrage free condition can rewrite as condition on the change in the factor β when a credit event occurs:

$$\exp\left(-\eta \Delta_k D(t)\right) - 1 = \frac{-\frac{1}{2}\Delta_k D(t)^2 \eta^2 + \sigma_\lambda^2 D(t,T_k)\Delta_k D(t)\lambda_t}{\left(\lambda_t e^{\eta D(t,T_{k+1})} + 1\right)}. \tag{2.76}$$

2.B Conditional density of the short rate under the annuity measure

As we only know the dynamics of these processes under the risk neutral, we first rewrite the moment generating function as function of expectation under the risk neutral world. Indeed,

$$
\begin{aligned}
\phi^r_{T_\alpha}(\omega) &= \mathbb{E}^{\mathbb{Q}^{IRS}}\left[e^{\omega r_{T_\alpha}} \mid \mathscr{G}_t\right] \\
&= \frac{\mathbb{E}\left[\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}^{IRS}}{\mathrm{d}\mathbb{Q}} \mid \mathscr{G}_T\right] e^{\omega r_{T_\alpha}} \mid \mathscr{G}_t\right]}{\mathbb{E}\left[\frac{\mathrm{d}\mathbb{Q}^{IRS}}{\mathrm{d}\mathbb{Q}} \mid \mathscr{G}_t\right]} \\
&= \frac{1}{E_t}\mathbb{E}\left[e^{-\int_t^{T_\alpha} r_u \mathrm{d}u} E_{T_\alpha} e^{\omega r_{T_\alpha}} \mid \mathscr{G}_t\right] \\
&= \frac{1}{E_t}\sum_{i=1}^{n} \delta^F\left(T_{i-1},T_i\right)\mathbb{E}\left[e^{-\int_t^{T_\alpha} r_u \mathrm{d}u + A(T_\alpha,T_i)+(\omega-B(T_\alpha,T_i))r_{T_\alpha}} \mid \mathscr{G}_t\right].
\end{aligned}
$$

Next, we focus on the evaluation of

$$
Y^i_t = \mathbb{E}\left[e^{-\int_t^{T_\alpha} r_u \mathrm{d}u + A(T_\alpha,T_i)+(\omega-B(T_\alpha,T_i))r_{T_\alpha}} \mid \mathscr{G}_t\right]
$$

and find its closed form thanks to the infinitesimal generator properties for Markov processes. As in the proposition (1) we can proof that for $Y^i_t = f_i(t,r_t)$,

$$
\lim_{v\to t} \frac{\mathbb{E}\left(Y^i_v \mid \mathscr{G}_t\right) - Y^i_t}{v-t} = r_t Y^i_t = \mathscr{A} f_i.
$$

Then, using Itô Lemma the infinitesimal generator is

$$
\mathscr{A} f_i = \partial_t f_i + (\theta(t) - a r_t)\,\partial_r f_i + \frac{1}{2}\sigma_r^2 \partial_r^2 f_i.
$$

Assuming that $f_i(t,r_t,\lambda_t,N_t) = e^{V^i_1(t,T_\alpha)+V^i_2(t,T_\alpha)r_t}$ we deduce the following ODEs system

$$
\begin{cases}
\partial_t V^i_1(t,T_\alpha) + \theta(t) V^i_2(t,T_\alpha) + \frac{1}{2}\sigma_r^2\left(V^i_2(t,T_\alpha)\right)^2 & = 0 \\
\partial_t V^i_2(t,T_\alpha) - a V^i_2(t,T_\alpha) & = 1
\end{cases},
$$

with the following terminal conditions

$$
\begin{cases}
V^i_1(T_\alpha,T_\alpha) & = A(T_\alpha,T_i) \\
V^i_2(T_\alpha,T_\alpha) & = \omega - B(T_\alpha,T_i)
\end{cases}.
$$

Then, by solving this ODEs system for each i, the moment generating function of r_{T_α} is

$$
\phi^r_{T_\alpha}(\omega) = \frac{1}{E_t}\sum_{i=\alpha+1}^{\beta} \delta^F\left(T_{i-1},T_i\right) e^{V^i_1(t,T_\alpha)+V^i_2(t,T_\alpha)r_t},
$$

where

$$V_1^i(t,T_\alpha) = A(T_\alpha,T_i) + \mu\left[e^{au} f^M(0,u) + \left(\frac{\sigma_r}{a}\right)^2\left(e^{-au}+e^{au}\right)\right]_{u=t}^{u=T_\alpha}$$
$$-\frac{1}{a}\left[f^M(0,u) + aP^M(0,u) + \frac{\sigma_r^2}{a}\left(u+\frac{e^{-2au}}{2a}\right)\right]_{u=t}^{u=T_\alpha}$$
$$+\left(\frac{\sigma_r}{2a}\right)^2\left[a\mu^2 e^{2au} - 4\mu e^{au} + 2u\right]_{u=t}^{u=T_\alpha},$$

$$V_2^i(t,T_\alpha) = \mu e^{at} - \frac{1}{a},$$

with $\mu = \left(\omega + \frac{1}{a} - B(T_\alpha,T_i)\right)e^{-aT_\alpha}$. Then, using inverse DFT the values of $f_{r_{T_\alpha}}(\cdot)$ at points $x_k = -\frac{M}{2}\Delta_x + (k-1)\Delta_x$ are approached by the sum:

$$f_{r_{T_\alpha}}(x_k) \approx \frac{2}{M\Delta_x} Re\left(\sum_{j=1}^{M} \Upsilon_j \phi_{T_\alpha}^r\left(iz_j\right)(-1)^{(j-1)} e^{-i\frac{2\pi}{M}(j-1)(k-1)}\right),$$

where $\Upsilon_j = \frac{1}{2}1_{\{j=1\}} + 1_{\{j\neq 1\}}$, M is the number of steps used in the DFT, $\Delta_x = \frac{2x_{max}}{M-1}$ is this step of discretization, and $\Delta_z = \frac{2\pi}{M\Delta_x}$ and $z_j = (j-1)\Delta_z$ for $j = 1...M$.

Chapter 3

Valuation of annuity guarantees under a self-exciting switching jump Model

3.1 Introduction

Over the past decade, the design of life insurance products has been progressively adapted to customer needs, from classic pure endowments, death deferred capital or unit-linked insurance products to variable annuities (VAs). A variable annuity is a unit-linked annuity product with guarantees or riders, which are mainly living and death benefits. Such a product allows for the allocation of savings to either a pre- or post-retirement investment strategy, as it guarantees lifetime income or death capital and protects VA holders from market downturns. In addition to these advantages, the popularity of VAs stems from the fact that they are indirect tax-deferred investments. Unlike in the United States, where sales volumes of VAs amounts to more than \$20 billion per quarter[1], in Europe the VA market is still expanding according to the annual reports of major European VA issuers. However, due to rider features of VAs, any VA issuer faces various risks, including financial, mortality or longevity and surrender risk. Furthermore, the low or negative interest rates environment, the interconnection of the markets and the interplay between risks are features that make the fair valuation of life insurance contracts with guarantees challenging.

The pricing of life insurance contracts with guarantees has been a topic of research interest for many years. Brennan and Schwartz (1976), Brennan and Schwartz (1979) and Boyle and Schwartz (1977) made important contributions to the academic understanding of the pricing of equity-linked products and life insurance policies using the modern option price theory. According to Milvesky and Posner (2001) and Milvesky and Salisbury (2006) research, various annuities are priced under the assumption of a constant interest rate and a geometric Brownian motion to model the underlying stock price. Considering the same setting, Bauer et al. (2008) extended the work of

[1] According to the Insured Retirement Institute. https://www.myirionline.org/newsroom/newsroom-detail-view/iri-issues-second-quarter-2019-annuity-sales-report

Milvesky and Salisbury (2006) by incorporating policyholder behavior. Considering stochastic interest rates, Lin and Tan (2003) and Kijima and Wong (2007) proposed a valuation approach for equity-indexed annuities. To insure that VA contracts remain profitable for the policyholder and the issuer, Bernard et al. (2014), Cui et al. (2017), Bernard and Moenig (2018) and Landriault et al. (2021) incorporated different fee structures into the pricing framework. In addition to the pricing issue associated with the financial approach, researchers have also investigated the fair valuation of equity-linked insurance with regard to the level of solvency capital required (see e.g., Hardy 2003; Barbarin and Devolder 2005 and Feng and Volkmer 2012). Using simulation or analytical methods where financial models permit, they analyzed liabilities for equity-linked insurance contracts and determined the fair values of parameters of these contracts using risk measures. The present chapter applies a risk-neutral pricing approach to VAs with minimum death and minimum life guarantees. An overview of other equity-linked life insurance policies is provided by Hardy (2003) and Bacinello et al. (2011). For VAs with minimum death and life guarantees, a benefit is paid either at the time of death or upon the maturity of the contract. This benefit is capped due to a guaranteed rate and depends on the performance of the underlying investment fund. Hence, VA holders receive at least the guaranteed amount during periods of crisis and earn additional expected income in periods of economic growth due to the good performance of investment funds. In contrast, as observed during the credit crunch that occured between October 2007 and March 2009, the main VA issuers incurred large losses. These losses were explained with reference to rising guarantee values, the collapse of earnings from mutual fund fees, negative hedging results and exploding hedging costs. As VAs offer protection of capital during periods of recession and provide an additional returns in conjunction with growth to their holders, incorporating switches in economic regimes into the VAs valuation framework is thus a good way of reducing the risk of losses for VA issuers.

In the finance literature, Goldfeld et al. (1973) developed the regime-switching technique. Hamilton (1989) popularized this technique in economics and econometrics. In finance, regime-switching models are based on the simple and natural notion that the economic environment is not stable but subject to regular changes at certain non-predictable stopping times. Various authors in the field of actuarial science have explored the use of regime-switching models in a continuous-time Markov switching framework. Without intending to present an exhaustive list, we refer to Hardy (2001), Siu (2005) Lin et al. (2009), Hainaut (2014), Fan et al. (2015) and Ignatieva et al. (2016). The common point of these articles is the use of diffusion processes modulated by a continuous-time Markov chain to model the interest rate and/or the reference risky asset. To the best of our knowledge, self-exciting jump-diffusion processes modulated by a continuous time hidden Markov chain have not been explored in the valuation framework of equity-linked life insurance contracts such as VAs. The literature on switching jump-diffusion is less abundant. For instance, Hainaut and Colwell (2016) extend the framework of Siu (2005) by using Markov regime-switching jump diffusion models. Fan et al. (2015) price annuity guarantees under double regime-switching models in which the underlying stock price was modulated by a Markov chain and its jump part depends on the number of state changes recorded. Hainaut and Moraux (2018 b) showed that the self-excitation property of the Hawkes process combined with regime-switching models enables one to better reproduce the dynamics of financial markets. In this framework, the features of shocks are modulated by economic cycles. The authors' empirical analysis of the S&P 500, reveals that self-excited jumps occur mainly during economic recessions and nearly disappear in periods of economic growth. The present chapter considers the setting of Hainaut and Moraux (2018 b) which integrated specific features of stock price dynamics such as jumps, heavy-tails of the return's distribution, the time-varying volatility, the regime-switching and, for option price, implied volatility smiles.

In the cited articles, the interest rates used for discounting cash flows are mostly considered constant. This assumption seems unrealistic for VAs, as potential increases in interest rates could have adverse effects on their commitments, mainly due to declining asset values. The literature features many stochastic interest rates models (e.g., Hull-White model (1990), Cox-Ingersoll-Ross (CIR) model (1985), Vasicek model (1977)) that allow for the replication of interest rate features such as time consistency, negativity, mean reversion, time-varying volatility and the clustering effect. Without losing generality, we consider a Hull-White model correlated to the stock price model, but any other stochastic model that enables one to reproduce the characteristics of interest rates can be included in our VAs valuation framework. As VAs are long-term commitments, under the real measure, we fit the interest rate model using the Kalman filter 1960 to reflect the long-term trend.

This chapter proposes an integrated framework for the valuation of VAs that guarantee a minimum benefit in case of death or life. It makes four contributions to the actuarial literature. First, this chapter introduces self-exciting switching jump-diffusion (SESJD) models into the standard pricing framework of VAs. Second, we design a valuation framework with a stochastic interest rate model correlated to the stock price process, and from a management perspective, we consider a portfolio consisting of a bond and a risky asset as the reference investment fund. Third, we define an equivalent martingale measure and derive an approximate closed-form expression for the pricing of VAs that guarantee a minimum living and death benefit. Finally, we present the detailed procedure used for the econometric calibration of our models based on the filtering technique.

The remainder of this chapter is structured as follows: in Section 3.2, we specify the hidden Markov chain model, the SESJD model and the interest rates model. Thereafter, we define a Radon Nikodym density for our equivalent martingale pricing measure. Under this pricing measure we find the dynamics of the investment portfolio and the moment generating function of the couple formed by the cumulated short-rate and the portfolio price. We conclude this section with a pricing formula of a European call option written on the investment portfolio obtained by an inverse Fourier transform. In Section 3.3, we specify our VA contract, and based on this contract, we provide an approximate closed-form expression of its price. In Section 3.4, we first discuss about the calibration of our models in the real world and then present a numerical implementation of our valuation framework, focusing on the impact of VA features.

3.2 The financial risk

We consider a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ equipped with a filtration $\{\mathscr{F}_t\}_{t\geq 0}$ that is generated by a continuous Markov chain and other processes we specify hereafter. Here, the probability measure $\mathbb{P}$ on $(\Omega, \mathscr{F})$ denotes the real world probability measure. The time set is $\mathrm{T} = [0, T^*]$, where $T^* < \infty$ is the final trading date. We consider a financial market in which security prices are adapted to the subfiltration $(\mathscr{H}_t)_{t\geq 0}$. Throughout this chapter, the economy is categorized into N states or regimes indexed by a set of integers $\mathscr{N} := \{1, 2, ..., N\}$. The information about the state of economy over time, is carried by a (continuous) hidden Markov chain $\zeta(t)$ taking values from a set of unit vectors $E = \{e_1, e_2, ..., e_N\}$, where $e_j = (0, ..., 0, 1, 0, ...0)^\top$ with 1 at the j^{th} position. The natural filtration generated by $\{\zeta(t)\}_{t\geq 0}$ is denoted by $\{\mathscr{G}_t\}_{t\geq 0}$, where $\mathscr{G}_t = \sigma(\zeta(s) : s \in \mathrm{T}, s \leq t)$. The generator of $\zeta(t)$ is an $N \times N$ matrix $Q_0 := \left[q_{i,j}\right]_{i,j=1,2,...,N}$, whose elements satisfy the following standard conditions:

$$q_{i,j} \geq 0, \qquad \forall i \neq j, \quad \text{and} \quad \sum_{j=1}^{N} q_{i,j} = 0, \qquad \forall i \in \mathscr{N}.$$

Each $q_{i,j}$ for $i \neq j$ is the instantaneous transition rate from state i to state j, while $-q_{i,i} = \sum_{j\neq i, j=1}^{N} q_{i,j} > 0$ is the instantaneous exit rate from state i. Hence,

$$\begin{aligned}
\lim_{\Delta\to 0} \frac{P\left(\zeta(t+\Delta) = e_j \mid \zeta(t) = e_i\right)}{\Delta} &= q_{i,j}, \\
\lim_{\Delta\to 0} \frac{1 - P(\zeta(t+\Delta) = e_i \mid \zeta(t) = e_i)}{\Delta} &= -q_{i,i}.
\end{aligned}$$

The matrix of transition probabilities over the time interval $[t, s]$ is denoted as $P(t, s)$ and is defined by:

$$P(t,s) = \left(p_{i,j}(t,s)\right)_{1\leq i,j\leq N} = \exp\left(Q_0(s-t)\right), \qquad s \geq t,$$

where for $1 \leq i, j \leq N$, $p_{i,j}(t, s)$ is the probability that the chain switches from state i at time t to state j at time s. Moreover, the probability that the chain is in state i at time s denoted by $p_i(s)$ depends upon the initial probabilities $p_k(0)$ and the transition probabilities $p_{k,i}(0, s)$, where $k = 1, 2, ..., N$. $p_i(s)$ is defined as follows:

$$p_i(s) = P(\zeta(s) = e_i) = \sum_{k=1}^{N} p_k(0)\, p_{k,i}(0,s), \qquad \text{for all } 1 \leq i \leq N.$$

From the semi-martingale representation theorem for $\zeta(t)$ provided by Elliott et al. (2005) and given that $\mathscr{G}_t$ is the filtration of $\zeta(t)$, we can write $\zeta(t)$ as the sum of a $\mathscr{G}_t$ predictable-adapted process and of a $\mathscr{G}_t$ martingale increment process, $\{M(t)\}_{t\geq 0}$:

$$\zeta(t) = \zeta(0) + \int_0^t Q_0^\top \zeta(s)\,\mathrm{d}s + M(t). \tag{3.1}$$

In this market, investors purchase three assets: cash, bonds and a stock. The price of this risky asset over time is a process $(S(t))_{t\geq 0}$ modulated by the Markov chain. We assume that the insurer invests the premium from the VA contract in a portfolio of bonds and stocks. The process $(S(t))_{t\geq 0}$ follows a modified version of the stock price model proposed by Hainaut and Moraux (2018 b) in which the instantaneous return is the sum of a drift, Brownian motions and a compensated jump process:

$$\begin{aligned}\frac{\mathrm{d}S(t)}{S(t^-)} &= \mu(t)\,\mathrm{d}t + \rho\sigma(t)\,\mathrm{d}W_r(t) + \sqrt{1-\rho^2}\sigma(t)\,\mathrm{d}W_S(t) \\ &\quad + \mathrm{d}\left(\sum_{i=1}^{N(t)}\left(e^{J_i}-1\right)\right) - \lambda(t)\,\mathbb{E}\left(e^{J}-1\right)\mathrm{d}t.\end{aligned} \tag{3.2}$$

We assume that the instantaneous average growth rate $\mu(t)$ of the stock and the Brownian volatility $\sigma(t)$ are modulated by the Markov chain $\zeta(t)$. In detail, $\mu(t) = \zeta(t)^{\top}\bar{\mu}$ and $\sigma(t) = \zeta(t)^{\top}\bar{\sigma}$, where $\bar{\mu} = (\bar{\mu}_1, \bar{\mu}_2, ..., \bar{\mu}_N)^{\top} \in \mathbb{R}^N_+$ and $\bar{\sigma} = (\bar{\sigma}_1, \bar{\sigma}_2, ..., \bar{\sigma}_N)^{\top} \in \mathbb{R}^N_+$. $W_r(t)$ and $W_S(t)$ are two independent Brownian motions, and ρ is the correlation coefficient between the short rate and the stock price. The amplitude of shocks J_i are i.i.d. random double exponential variables (DEJ) with density $\nu(z)$ on $\mathbb{R}$. Further details about the jump size distribution and moments are provided in Appendix A 3.A. It is important to note that most of the results presented in this chapter are applicable to any other jump size distributions that enables both negative and positive jumps. The counting process $(N(t))_{t\geq 0}$ with intensity $(\lambda(t))_{t\geq 0}$ counts the number of observed shocks independent of their amplitude up to time t. To replicate the clustering of shocks observed in stocks markets, the intensity is stochastic and defined as follows :

$$\lambda(t) = \lambda(0) - \beta\int_0^t e^{\beta(s-t)}(\lambda(0)-\eta(s))\,\mathrm{d}s + \int_0^t \gamma e^{\beta(s-t)}\mathrm{d}L_s,$$

or in terms of stochastic differential equation (SDE) by :

$$\mathrm{d}\lambda(t) = \beta(\eta(t)-\lambda(t))\,\mathrm{d}t + \gamma\mathrm{d}L(t), \tag{3.3}$$

where $\beta \geq 0$ is the speed of reversion, $L(t) = \sum_{i=1}^{N(t)}|J_i|$ enables to increase the intensity $\lambda(t)$ of $\gamma|J|$ when a jump occurs. The long-run level $\eta(t)$ is modulated by the Markov chain $\zeta(t) : \eta(t) = \zeta(t)^{\top}\bar{\eta}$, where $\bar{\eta} = (\bar{\eta}_1, \bar{\eta}_2, ..., \bar{\eta}_N)^{\top} \in \mathbb{R}^N_+$. The information about this counting process is contained in the subfiltration $\mathscr{H}$. Hainaut and Moraux, 2018 b derive from their proposition 2.1 that conditionally to $\mathscr{H}_0$ the expectation of $\lambda(t)$ converges when t tends to infinity only if the speed of mean reversion is larger than $\gamma\mathbb{E}(|J|)$. In the calibration section, this stability condition ensures that the number of shocks hitting the stock price process $S(t)$ does not explode as the maturity increases. To preserve the analytical tractability, we assume that the risk-free rate $(r(t))_{t\geq 0}$ is not modulated by the Markov chain and satisfies, under $\mathbb{P}$ this SDE:

$$\mathrm{d}r(t) = \left(\theta^{\mathbb{P}}(t) - a_r r(t)\right)\mathrm{d}t + \sigma_r\mathrm{d}W_r(t), \tag{3.4}$$

where the level of mean reversion $\theta^{\mathbb{P}}(t)$ is a process, $a_r \geq 0$ is the speed of mean-reversion and $\sigma_r \geq 0$ is the volatility of the interest rates market. In order to price bonds, we define a risk neutral measure. However, the market is incomplete due to regime switches and jumps. The risk-neutral measure derived under the martingale condition is not unique, which is also the case with the price of a given product in this market. Different equivalent measures such as the Esscher transform measure, the minimal entropy measure or the variance-optimal measure, could be used, but, in this chapter, we do not discuss the selection criteria of an equivalent martingale measure. Instead, we consider that the equivalent martingale measure $\mathbb{Q}$ of $\mathbb{P}$ with the Radon Nykodym density defined by:

$$\begin{aligned}\Lambda(t) = \left.\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right|_{\mathscr{F}_t} =: \exp\Big(-\frac{1}{2}\int_0^t \left(\sigma_r^2\theta_r^2 + \sigma^2(u)\,\theta_S^2(u)\right)\mathrm{d}u \\ -\int_0^t \sigma_r\theta_r \mathrm{d}W_r(u) - \int_0^t \sigma(u)\,\theta_S(u)\,\mathrm{d}W_S(u)\Big)\end{aligned} \tag{3.5}$$

where $\theta_S(t) = \zeta(t)^{\top}\bar{\theta}_S$, $\theta_r \in \mathbb{R}$ and $\bar{\theta}_S \in \mathbb{R}^N$. Using the Itô's differentiation rule, we find that $\mathrm{d}\Lambda(t) = -\sigma_r\theta_r \mathrm{d}W_r(t) - \sigma(t)\,\theta_S(t)\,\mathrm{d}W_S(t)$. Then, as $\Lambda(t)$ is an $\mathscr{F}$-martingale under the probability measure $\mathbb{P}$ and $\mathbb{E}(\Lambda(t)) = 1$, our equivalent measure $\mathbb{Q}$ is well-defined. If θ_r is a constant and $\theta_S(t)$ is such that

$$\theta_S(t) = \frac{\mu(t) - r(t) - \rho\sigma_r\theta_r\sigma(t)}{\sqrt{1-\rho^2}\sigma^2(t)}, \tag{3.6}$$

then the discounted price process of the risky asset $\exp\left(-\int_0^t r(u)\,\mathrm{d}u\right)S(t)$ is a $\mathscr{F}$-martingale under the probability measure $\mathbb{Q}$. Under this risk neutral measure, $(S(t))_{t\in\mathrm{T}}$ and $(r(t))_{t\in\mathrm{T}}$ satisfy the following SDEs :

$$\mathrm{d}r(t) = \left(\theta^{\mathbb{Q}}(t) - a_r r(t)\right)\mathrm{d}t + \sigma_r \mathrm{d}W_r^{\mathbb{Q}}(t), \tag{3.7}$$

$$\begin{aligned}\frac{\mathrm{d}S(t)}{S(t^-)} = r(t)\,\mathrm{d}t + \rho\sigma(t)\,\mathrm{d}W_r^{\mathbb{Q}}(t) + \sqrt{1-\rho^2}\sigma(t)\,\mathrm{d}W_S^{\mathbb{Q}}(t) \\ +\mathrm{d}\left(\sum_{i=1}^{N(t)}\left(e^{J_i}-1\right)\right) - \lambda(t)\,\mathbb{E}\left(e^{J}-1\right)\mathrm{d}t,\end{aligned} \tag{3.8}$$

where the processes $\left(W_r^{\mathbb{Q}}(t)\right)_{t\in\mathrm{T}}$ and $\left(W_S^{\mathbb{Q}}(t)\right)_{t\in\mathrm{T}}$ are two Brownian motions under $\mathbb{Q}$ such that:

$$\begin{aligned}W_r^{\mathbb{Q}}(t) &= W_r(t) + \sigma_r\theta_r t, \\ W_S^{\mathbb{Q}}(t) &= W_S(t) + \int_0^t \sigma(u)\,\theta_S(u)\,\mathrm{d}u.\end{aligned}$$

$\theta^{\mathbb{Q}}(t) = \theta^{\mathbb{P}}(t) - \sigma_r^2\theta_r$ is a deterministic function of the time, where $-\sigma_r\theta_r$ is the interest rate's risk price. We note that the market price of risk $-\theta_S(t)\,\sigma(t)$ in our framework depends on the state of the economy and the instantaneous short rate. For the sake of simplicity, we assume that the distribution of the hidden Markov chain ζ and the modulated jump process (λ, N) remain unchanged after the measure change from $\mathbb{P}$ to $\mathbb{Q}$. Thus, the semimartingale representation of the hidden Markov chain in Equation (3.1) still holds.

In this setting, asset prices are valued as the sum of their expected discounted cash flows under the risk-neutral measure. From Equation (3.7), the short-rate follows a Hull-White process (1990) under $\mathbb{Q}$. The price of a zero-coupon bond at time t with maturity T is given by:

$$ZC(t,T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(s)\mathrm{d}s} \mid \mathscr{F}_t\right] = e^{-A^r(t,T)-B^r(t,T)r(t)}, \tag{3.9}$$

where

$$B^r(t,T) = \frac{1}{a_r}\left(1-e^{-a_r(T-t)}\right), \tag{3.10}$$

and

$$A^r(t,T) = \int_t^T f^M(0,s)\,\mathrm{d}s - f^M(0,t)\,B^r(t,T) + \frac{\sigma_r^2}{4a_r}\left(1-e^{-2a_r t}\right)B^r(t,T)^2, \tag{3.11}$$

with $f^M(0,t)$ the instantaneous forward rate observed at time 0. From Equation (3.9), we have the SDE of this zero coupon bond :

$$\frac{\mathrm{d}ZC(t,T)}{ZC(t^-,T)} = r(t)\,\mathrm{d}t - \sigma_r B^r(t,T)\,\mathrm{d}W_r^{\mathbb{Q}}(t). \tag{3.12}$$

It follows that $-\theta_r\sigma_r^2 B^r(t,T)$ represents the instantaneous risk premium for an investment in such a zero-coupon bond.

The equations obtained under the measure $\mathbb{Q}$ are used later to value the portfolio in which the insurer invests the unique premium. Composed of a zero coupon bond and a risky asset, the value of this portfolio denoted by $V(t)$ at time t is defined as follows :

$$\frac{\mathrm{d}V(t)}{V(t^-)} = \omega_S\frac{\mathrm{d}S(t)}{S(t^-)} + (1-\omega_S)\frac{\mathrm{d}ZC(t,T)}{ZC(t^-,T)}, \tag{3.13}$$

where ω_S is the proportion of the premium invested in the risky asset. Without loss of generality, we assume that our portfolio is self-financed and continuously re-balanced in order to keep this proportion constant over the time. It is worth noting that one may considered the proportion ω_S invested in the risky asset as a function of the time. For instance, it could decreases with the age of the VA's holder who is risk adverse. Under the risk neutral measure from Equation (3.13) the log-return of the investment portfolio $X(t) := \ln\frac{V(t)}{V(0)}$ satisfies the following SDE :

$$\begin{aligned}
\mathrm{d}X(t) = &\left(r(t) + \omega_S(1-\omega_S)\rho\sigma(t)\sigma_r B^r(t,T) - \frac{1}{2}\omega_S^2\sigma^2(t)\right)\mathrm{d}t \\
&-\left(\frac{1}{2}(1-\omega_S)^2\sigma_r^2(B^r(t,T))^2 + \omega_S\lambda(t)\,\mathbb{E}\left(e^J-1\right)\right)\mathrm{d}t \\
&+(\omega_S\rho\sigma(t) - (1-\omega_S)\sigma_r B^r(t,T))\,\mathrm{d}W_r^{\mathbb{Q}}(t) \\
&+\omega_S\sqrt{1-\rho^2}\sigma(t)\,\mathrm{d}W_S^{\mathbb{Q}}(t) + \ln\left(1+\omega_S\left(e^J-1\right)\right)\mathrm{d}N(t).
\end{aligned} \tag{3.14}$$

We conclude this section with a proposition used to price our VA contract. To price embedded options in the VA, we need the moment generating function (mgf) of the bivariate random variable $(\int_t^s r(u)\,\mathrm{d}u, X(s))$ given the filtration $\mathscr{F}_t$ with $t \le s$.

Proposition 6. *The mgf $M_{r,X}^{t,s}$ of the pair of random variables $\left(\int_t^s r(u)\,du,\ X(s)\right)$ given the filtration $\mathscr{F}_t$ with $\bar{\omega} = (\omega_1, \omega_2) \in \mathbb{C}_-^2$ and $t \le s$ is equal to :*

$$\begin{aligned} M_{r,X}^{t,s}(\omega_1,\omega_2) &= \mathbb{E}^{\mathbb{Q}}\left(e^{\omega_1 \int_t^s r(u)du + \omega_2 X(s)} \mid \mathscr{F}_t\right) \\ &= \left(\frac{V(t)}{V(0)}\right)^{\omega_2} \zeta^{\top}(t)\tilde{A}(\bar{\omega},t,s) \\ &\quad \times \exp\left((\omega_2+\omega_1)B^r(t,s)r(t) + C(\bar{\omega},t,s)\lambda(t)\right), \end{aligned} \tag{3.15}$$

where $\tilde{A}(\cdot)$ is a vector of N functions, and the function $C(\cdot)$ satisfy the terminal value problems' partial differential equations (PDEs) system:

$$\begin{cases} 0 = \frac{\partial \tilde{A}(\bar{\omega},t,s)}{\partial t} + (diag(D(\bar{\omega},t,s)) + Q_0)\tilde{A}(\bar{\omega},t,s) \\ 0 = C_t - \omega_2\omega_S \mathbb{E}^{\mathbb{Q}}\left(e^J - 1\right) - \beta C + \mathbb{E}^{\mathbb{Q}}(g(J,C)) - 1 \end{cases} \tag{3.16}$$

with $g : (x,C) \longmapsto e^{\omega_2 \ln(1+\omega_S(e^x-1)) + C\gamma|x|}$, a real valued bivariate function,

$$\begin{aligned} D(\bar{\omega},t,s) &= -\frac{\omega_2}{2}(1-\omega_2)\left(D_1^2 + \omega_S^2\left(1-\rho^2\right)\bar{\sigma}^2\right) + \theta^{\mathbb{Q}}(t)D_2 + \frac{1}{2}\sigma_r^2 D_2^2 \\ &\quad + \omega_2\sigma_r D_1 D_2 + \beta\bar{\eta}C, \end{aligned}$$

$D_1 = \omega_S\rho\bar{\sigma} - (1-\omega_S)\sigma_r B^r(t,T)$, $D_2 = (\omega_2+\omega_1)B^r(t,s)$, $C(\bar{\omega},s,s) = 0$ and $\tilde{A}(\bar{\omega},s,s) = 1_{\mathbb{R}^N}$.

Proof. In Appendix B 3.B □

Note that while the PDE system (3.16) does not admit an analytical solution, a numerical solutions can be found using a numerical method. The joint distribution of $(\int_t^s r(u)\,\mathrm{d}u, X(s))$ defined by this mgf allows us to price by Fourier transform inversion, a European put option written on the portfolio value V with maturity T and strike $V(t)e^k$. Let $p_T(k)$ be the value of this European put option and ψ the joint density of $(\int_t^s r(u)\,\mathrm{d}u, X(s))$ under $\mathbb{Q}$. Thus,

$$p_T(k) = V(t)\int_{-\infty}^{k}\int_{-\infty}^{+\infty} e^{-y_1}\left(e^k - e^{y_2}\right)\psi(y_1,y_2)\,\mathrm{d}y_1\mathrm{d}y_2. \tag{3.17}$$

This European put option is not square integrable since p_T tends to infinity instead of zero when the log strike k tends to $+\infty$. To obtain a square integrable function, we consider the modified put option price:

$$\bar{p}_T(k) := \exp(-\alpha k)\,p_T(k), \tag{3.18}$$

where $\alpha > 0$ is chosen such that $\bar{p}_T$ is a square-integrable function. Then the Fourier transform of $\bar{p}_T$ exists and is defined by

$$\mathscr{F}(\bar{p}_T)(v) = \int_{-\infty}^{+\infty} e^{ivk}\bar{p}_T(k)\,\mathrm{d}k, \tag{3.19}$$

and, by inverse Fourier transform, the option price is obtained numerically using the following equation :

$$p_T(k) = \frac{e^{\alpha k}}{\pi} \int_0^{+\infty} e^{-ivk} \mathscr{F}(\bar{p}_T)(v)\, \mathrm{d}v. \tag{3.20}$$

Applying Simpson's rule to the integral in Equation (3.20), we approximate put options prices in the following proposition:

Proposition 7. *Let M be the number of steps used in the discrete Fourier transform and $\Delta_k = \frac{2k_{max}}{M}$ the step of discretization. Let us denote $\delta_j = \frac{1}{3}\left(3 + (-1)^j - 1_{\{j=1\}}\right)$, $\Delta_v = \frac{2\pi}{M\Delta_k}$ and $v_j = (j-1)\Delta_v$. The value of $p_T(k)$ at points $k_j = -kmax + (j-1)\Delta_k$ is approximated by*

$$p_T\left(k_j\right) \approx \frac{2V(t)\, e^{\alpha k_j}}{M\Delta_k} \sum_{l=1}^{M} \delta_l e^{-\frac{2\pi i}{M}(j-1)(l-1)} (-1)^{l-1} \frac{M_{r,X}^{t,T}(-1, 1-\alpha+iv_l)}{\alpha^2 - \alpha - v_l^2 + i(-2\alpha+1)v_l}. \tag{3.21}$$

for all $j = 1,...,M$ and any $\alpha > 0$ such that $M_{r,X}^{t,T}(-1, 1-\alpha) < +\infty$. This last relation can be computed with a fast Fourier transform algorithm.

Proof. In Appendix D 3.D ☐

In the next section, we consider several of the different forms of VA contract and price them within our modelling framework.

3.3 Valuation of the variable annuity

3.3.1 Definitions

The VA contract valued under our models setting consists of a guaranteed minimum accumulation benefit (GMAB) and guaranteed minimum death benefit (GMDB). Any holder of this VA contract pays a unique premium $\Pi > 0$ in exchange for benefits that depend on the return of the investment portfolio. The policy charges fees for management expenses such as the mortality fee, administrative costs, management fees and other fees linked to the contract riders. These fees are charged annually to the account value of the VA contract. If we denote by φ the proportional fee rate applied to the account value in order to recover the cost of all the options, then the account value A satisfies the following SDE:

$$\frac{\mathrm{d}A(t)}{A(t^-)} = \frac{\mathrm{d}V(t)}{V(t^-)} - \varphi \mathrm{d}t, \tag{3.22}$$

where its initial value $A(0)$ is equal to the unique premium Π. Let $\delta \geq 0$ be the guaranteed interest rate used to reassess the premium during the lifetime of the VA contract. Our VA contract provides a minimum guaranteed benefit in case of death and a lump sum payment upon expiry. If withdrawal and surrender are not allowed during the contract's lifetime, benefits are defined as follows

- GMDB: In case of death prior to the stated maturity T, the insurer pays the maximum between the account value and a guaranteed amount. At the death time τ_d, $t < \tau_d \leq T$, the benefit is $b_D(\tau_d) = \max\{A(\tau_d), G(\tau_d)\}$, where $G(0) \in (0, \Pi]$ and the common guaranteed amount $G(\tau_d)$ on the market is
  - either the roll-up of the premium at the guaranteed interest rate, called the roll-up guaranteed
  $$G(\tau_d) = G(t)\, e^{\delta(\tau_d - t)}, \tag{3.23}$$
  - or the highest account value recorded every year before the death, called the ratchet guaranteed,
  $$G(\tau_d) = \max_{j \in \{1, \cdots, \lfloor \tau_d \rfloor\}} (G(t), A(t+j)). \tag{3.24}$$
- GMAB: At expiry date T (typically the end of the accumulation period), the insured, if alive, receives the maximum between the policy account value and the guaranteed amount. The benefit is
  $$b_L(T) = \max\{A(T), G(T)\} \tag{3.25}$$
  where the guaranteed amount $G(T)$ is expressed similarly as in the death case 3.24.

The fair proportional fee rate φ is the solution of the following equation:

$$\Pi \quad := \quad L(t, T, \varphi), \tag{3.26}$$

where $L(t, T, \varphi)$ denotes the total liabilities of the insurer at time t. This total liability is the net of the proportional fee φ which is deducted from the investment fund value.

3.3.2 Valuation

The theorem of asset pricing and the linearity of the expectation allow us to define the total liabilities of the insurer at time t toward an insured aged x at time t as follows:

$$\begin{aligned} L(t, T, \varphi) &:= \mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^{\tau_d} r(u)\mathrm{d}u} 1_{\{\tau_d < T\}} b_D(\tau_d) + e^{-\int_t^T r(u)\mathrm{d}u} 1_{\{\tau_d > T\}} b_L(T)\right) \\ &= L_D(t, T, \varphi) + L_L(t, T, \varphi), \end{aligned} \tag{3.27}$$

where t and T are integers. $L_D(t, T, \varphi)$ and $L_L(t, T, \varphi)$ are respectively the present values of the death benefit and of the survival benefit at time t. These quantities are evaluated under the assumption that mortality risk is independent from financial risks. We also assume that individuals die at the end of the year and that the death benefit is paid at the end of the death year. This latter assumption is arguable but can be adapted according to the terms of the contract. The death liability in Equation (3.27) is then given by:

$$
\begin{aligned}
L_D(t,T,\varphi) &= \mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^{\tau_d} r(u)\mathrm{d}u}1_{\{\tau_d<T\}}b_D(\tau_d) \mid \mathscr{F}_t\right) && (3.28)\\
&= \sum_{j=0}^{T-t-1} {}_jp_{x+t}q_{x+t+j}\mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^{t+j+1} r(u)\mathrm{d}u}b_D(t+j+1) \mid \mathscr{F}_t\right),
\end{aligned}
$$

where $t+1, t+2, ..., T$ are the possible dates of payment of death benefits until the maturity T, whereas the life liability $L_L(t,T,\varphi)$, is given by:

$$
\begin{aligned}
L_L(t,T,\varphi) &= \mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^{T} r(u)\mathrm{d}u}1_{\{\tau_d>T\}}b_L(T) \mid \mathscr{F}_t\right) && (3.29)\\
&= {}_T\, p_{x+t}\mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^{T} r(u)\mathrm{d}u}b_L(T) \mid \mathscr{F}_t\right), && (3.30)
\end{aligned}
$$

where the life benefit $b_L(T) = \max\{A(T), G(T)\}$ changes according to the two type of the guaranteed amount considered hereafter:

- Roll-up guaranteed
  In this case, the guaranteed amount $G(T)$ satisfies Equation (3.23) and the life benefit is the maximum between a random variable and a constant. This life benefit can be rewritten as a function of the payoff of a put option and enables us to simplify the life liability formula as follows:

$$
\begin{aligned}
L_L(t,T,\varphi) &= {}_T\, p_{x+t}\mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^{T} r(u)\mathrm{d}u}\max\left\{A(T), G(t)e^{\delta(T-t)}\right\} \mid \mathscr{F}_t\right) && (3.31)\\
&= {}_T\, p_{x+t}\mathbb{E}^{\mathbb{Q}}\Bigg(e^{-\int_t^{T} r(u)\mathrm{d}u}\Bigg(A(t)e^{X(T)-X(t)-\varphi(T-t)}\\
&\qquad + \left(G(t)e^{\delta(T-t)} - A(t)e^{X(T)-X(t)-\varphi(T-t)}\right)_+\Bigg) \mid \mathscr{F}_t\Bigg)\\
&= {}_T\, p_{x+t}\mathbb{E}^{\mathbb{Q}}\Bigg(e^{-\int_t^{T} r(u)\mathrm{d}u}\Bigg(A(0)e^{X(T)-\varphi T}\\
&\qquad + \left(G(0)e^{\delta T} - A(0)e^{X(T)-\varphi T}\right)_+\Bigg) \mid \mathscr{F}_t\Bigg)\\
&= {}_T\, p_{x+t}e^{-\varphi T}A(0)\Bigg(\mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^{T} r(u)\mathrm{d}u+X(T)} \mid \mathscr{F}_t\right)\\
&\qquad + \mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^{T} r(u)\mathrm{d}u}\left(\frac{G(0)}{A(0)}e^{(\varphi+\delta)T} - e^{X(T)}\right)_+ \mid \mathscr{F}_t\right)\Bigg)\\
&= {}_T\, p_{x+t}A(0)e^{-\varphi T}\left(M_{r,X}^{t,T}(-1,1) + p_T(t)\right).
\end{aligned}
$$

where $p_T(t)$ is the price of a European put option with maturity T and strike $\frac{G(0)}{A(0)}e^{(\delta+\varphi)T}$. Recall that the initial account value $A(0)$ equals Π and the initial amount guaranteed

$G(0)$ is in the interval $(0,\Pi]$. From Equation (3.31), the insurer commitment of a GMAB is then hedgeable by investing in cash and a put option.

- Annual ratchet guaranteed
  The life benefit is the highest account value A recorded every year until the maturity T. Since the account value at any time depends on the state of the chain, the property of the independence of increments does not hold. While we can not deduce a closed-form expression in this case, the life liability for a guaranteed annual ratchet:

$$L_L(t,T,\varphi) = {}_T\, p_{x+t}\mathbb{E}^{\mathbb{Q}}\Bigg(e^{-\int_t^T r(u)\mathrm{d}u} \times \max\left\{A(T), \max_{j\in\{1,\cdots,T-t-1\}}\{G(t), A(t+j)\}\right\} \mid \mathscr{F}_t\Bigg) \tag{3.32}$$

can be evaluated using the Monte Carlo approach.
In addition to the life liability, the death liability 3.28 can be rewrite as a weighted sum of life commitments in this way :

$$\begin{aligned} L_D(t,T,\varphi) &= \sum_{j=0}^{T-t-1} {}_j p_{x+t}\, q_{x+j} \frac{{}_{j+1}p_{x+t}}{{}_{j+1}p_{x+t}} \\ &\quad \times \mathbb{E}^{\mathbb{Q}}\left(e^{-\int_t^{t+j+1}(r_u+\nu_u)\mathrm{d}u} b_L(t+j+1) \mid \mathscr{F}_t\right) \\ &= \sum_{j=0}^{T-t-1}\left(\frac{{}_j p_{x+t}}{{}_{j+1}p_{x+t}} - 1\right) L_L(t, t+j+1, \varphi). \end{aligned} \tag{3.33}$$

It follows that for a roll-up guaranteed, the death liability

$$\begin{aligned} L_D(t,T,\varphi) &= A(0)\sum_{j=0}^{T-t-1}\left(\frac{{}_j p_{x+t}}{{}_{j+1}p_{x+t}} - 1\right) {}_{t+j+1}p_{x+t}\, e^{-\varphi(t+j+1)} \\ &\quad \times \left(M_{r,X}^{t,t+j+1}(-1,1) + p_{t+j+1}(t)\right). \end{aligned}$$

is also hedgeable by investing in cash and in put options with different maturities.

3.4 Numerical implementation

This section is divided into four subsections. The first subsection presents the mortality rate model and its estimated parameters, while the second subsection details the econometric calibration of our interest rate model via the Kalman filter approach. The third subsection presents the estimation procedure of our stock models when the hidden Markov chain ζ has three states: e_1 for "boom" or "trend up", e_2 for "recession" or "trend down" and e_3 for "crisis" or "trend sidway". Finally, we perform a sensitivity analysis to highlight the impact of self-exciting jumps and economic regimes on the valuation of the VA defined in the previous section.

3.4.1 Estimation of the mortality rate

The mortality rate model follows the Gompertz Makeham distribution (Gompertz (1825), Matthew (1860)). This setting enables one to derive a closed-form expression of the survival probability. For an individual aged x at time zero, the mortality rate $\mu(x)$ is given by:

$$\mu(x) = a_\mu + b_\mu c_\mu^x, \quad a_\mu = -\ln(s_\mu), \quad b_\mu = -\ln(g_\mu)\ln\left(c_\mu^x\right), \tag{3.34}$$

where a_μ is the mortality rate due to accidents and $b_\mu c_\mu^x$ is the term related to the ageing. The chosen parameters in Table 3.1 maximize the likelihood function when we consider the mortality tables for both sexes provided by the Belgian regulator.[2] The national regulator for life annuities and death products recommends these mortality tables as a baseline.

s_μ	0.9995465
g_μ	0.9998611
c_μ	1.106383

Table 3.1: The Gompertz Makeham parameters for both sexes Belgian mortality table, obtained by likelihood maximization.

If τ_d denotes the death time, we evaluate the survival probability of an individual aged x at time zero as follows :

$$_\varepsilon p_x = P(\tau_d > \varepsilon) = e^{-\int_0^\varepsilon \mu(x+u)\mathrm{d}u}, \ \forall \varepsilon \geq 0. \tag{3.35}$$

3.4.2 Econometric calibration of the short rates model

As VAs are long-term commitments, we fit the interest rate model under the real measure $\mathbb{P}$ using the Kalman filter 1960 to reflect the long-term trend. The speed of mean reversion and volatility found in this method are next used under the risk neutral condition. We believe these parameters better reflect the long-term trend than those we could obtain by fitting the model to, for example swaptions. Our estimation procedure is divided into two steps and based on a time series of daily ECB yield bonds[3] data from September 6, 2004, to April 16, 2020 for 10 maturities. Instead of using the overnight rate as a proxy for the short-term rate, in the first step we use the Kalman filter approach 1960 to guess its dynamics for the time series of daily ECB bond yields. Babbs and Nowman (1999) adopted a similar approach. Let $\{t_1, \cdots, t_{3991}\}$ be the set of times at which we observe these yield bonds curves. Table 3.2 summarizes the descriptive statistics of the yield bonds.

[2] The Belgian mortality tables for both sexes are available via the following link http://www.ejustice.just.fgov.be/eli/arrete/2013/01/29/2013011073/justel

[3] Source: http://sdw.ecb.europa.eu/browse.do?node=9691126

Maturity	1	2	3	4	5	6	7	8	9	10
Mean	0.84	0.97	1.13	1.31	1.48	1.65	1.81	1.95	2.08	2.19
Std. dev.	1.59	1.61	1.63	1.62	1.61	1.61	1.6	1.59	1.59	1.58
Min	-0.91	-0.97	-1.	-1.01	-0.996	-0.97	-0.93	-0.89	-0.87	-0.82
Max	4.53	4.71	4.73	4.73	4.73	4.72	4.73	4.74	4.76	4.78

Table 3.2: Descriptive statistics of the yield bonds in percent, from September 6, 2004, to April 16, 2020.

The bond price $ZC(t,T)$ (3.9) is the exponential of an affine function in $r(t)$: $-A^r(t,T) - r(t)B^r(t,T)$, where functions A^r and B^r are defined in Equation (3.11) and (3.10) respectively. Therefore, the observable spot rate or yield bond $R(t,T) = -\frac{\ln(ZC(t,T))}{T-t}$ is an affine function. We consider a linear Kalman filter. The measurement equation is defined as follows :

$$R(t_k) \quad = z(t_k)\, r(t_k) + d(t_k) + \varepsilon(t_k) \qquad \varepsilon(t_k) \sim N(0,H) \tag{3.36}$$

where $R(t_k) = (R(t_k, t_k + \tau_i))_{1\le i\le 10}$ is the vector of yield bond for the ten maturities $(\tau_i)_{1\le i\le 10}$ ranging from one year to ten years, and $z(t_k) = \left(\frac{B^r(t_k,t_k+\tau_i)}{\tau_i}\right)_{1\le i\le 10}$ and $d(t_k) = \left(\frac{A^r(t_k,t_k+\tau_i)}{\tau_i}\right)_{1\le i\le 10}$ are vectors derived from the price of zero-coupon bonds. $\tau_i \in \{1,2,...,10\}$ is the maturity of the considered bond. The $\varepsilon(t_k)$ are the measurement errors assumed normally distributed with a covariance matrix $H = \operatorname{diag}(h_1,\cdots,h_{10})$. Note that the variance of the measurement errors depends on maturities. For a step size $\Delta = \frac{1}{365}$, we have 3991 values of bond yields for each maturity τ_i. Moreover, the transition equation of the state variable r according to this step size is given by

$$r(t_k) = f(t_k)\, r(t_{k-1}) + g(t_k) + w(t_k) \tag{3.37}$$

where

$$\begin{aligned}
f(t_k) &= e^{-a_r\Delta},\\
g(t_k) &= \theta^{\mathbb{P}} B^r(t_{k-1},t_k),\\
w(t_k) &\sim N\left(0, \frac{\sigma_r^2}{2a_r}\left(1-e^{-2a_r\Delta}\right)\right).
\end{aligned}$$

Let us define by $\Theta = \left\{a_r, \sigma_r, \theta_r, \theta^{\mathbb{P}}, h_1, \cdots, h_{10}\right\}$ the set of parameters of the short-rate model including the parameter from the distribution of measurement errors. Note that $-\sigma_r\theta_r$ is the interest rate's risk price. We assume that error terms of the measurement and the transition equations are not correlated. The distribution of $r(t_k)$ then depends only on the value of $r(t_{k-1})$, while the distribution of $R(t_k)$ depends only on $r(t_{k-1})$. With the Kalman filter, we can guess the unobservable values $r(t_k)$ and estimate the filtering error. The likelihood function of the log likelihood for the observations $R(t_1), \cdots, R(t_{3991})$ is

$$\begin{aligned}
l_{3991} &:= \ln L(R(t_1), \cdots, R(t_{3991}); \Theta)\\
&= \ln p(R(t_1)) + \sum_{k=2}^{3991} \ln p\left(R(t_k) \mid \mathscr{H}_{t_{k-1}}\right),
\end{aligned} \tag{3.38}$$

where $p\left(R(t_k) \mid \mathscr{H}_{t_{k-1}}\right)$ is the conditional distribution of $R(t_k)$ given the information set, $\mathscr{H}$ at time t_{k-1}. $p(R(t_1))$ is the initial distribution of $R(t_1)$ assumed as the stationary distribution of the instantaneous spot rate $r(t_1) \sim N\left(f(t_1)\, r(0) + g(t_1), \frac{\sigma_r^2}{2a_r}\right)$. We find an estimate $\hat{\Theta}$ of Θ by maximizing the log likelihood in Equation (3.38).

In the second step, we estimate the mean reversion level $\theta^{\mathbb{Q}}$ in the short-rate model such that we fit the initial forward curve. For this purpose, we assume that the initial forward curve follows a Nelson-Siegel parametric curve (1987). The Nelson-Siegel model is a parsimonious model of yield curves widely used in practice to replicate at best the initial term structure of interest rates. Thus, we define the initial forward rate as follows:

$$f^M(0,s) = z_1 + z_2 e^{-\alpha s} + z_3 s e^{-\alpha s}, \quad s > 0 \tag{3.39}$$

where z_1, z_2, z_3 and $\alpha > 0$ are fixed real numbers. As a proxy of $r(0) = f(0,0)$, we use the value of instantaneous forward rate 10-month residual maturity from ECB data equal to -0.65%. Next, we deduce a closed form of the moving target at time $s > 0$:

$$\begin{aligned} \theta^{\mathbb{Q}}(s) &= a_r z_1 + \frac{\sigma_r^2}{2a_r} + (z_3 + (a_r - \alpha) z_2) e^{-\alpha s} + (a_r - \alpha) z_3 s e^{-\alpha s} \\ &\quad - \frac{\sigma_r^2}{2a_r} e^{-2a_r s}. \end{aligned} \tag{3.40}$$

Given the estimated values of a_r, σ_r and θ_r, we find z_1, z_2, z_3 and $\alpha > 0$ that minimize the mean square error between the observed yield curve as of April 16, 2020, and the model yield curve where the mean reversion level satisfies Equation (3.40).
Table 3.3 contains estimated parameters of the short-rate model and the estimated variances of measurement errors per maturity based on the estimation procedure described previously. Figure 3.1 compares the observed yield curve as of April 16,

a_r	1.61%	θ_r	5.06%
σ_r	0.69%	r_0	−0.65%
z_1	−1.07%	z_2	0.42%
z_3	−0.75%	α	1.51

h_1	0.33%	h_2	0.15%
h_3	0%	h_4	0.13%
h_5	0.23%	h_6	0.32%
h_7	0.39%	h_8	0.45%
h_9	0.51%	h_{10}	0.55%

Table 3.3: In the left side, the estimated parameters of the short-rate model. In the right side, the estimated variance of the measurement errors.

2020, to the yield curve obtained from the Hull-White model where the initial forward curve follows a Nelson-Siegel model with the estimated parameters in Table 3.3.

3.4.3 Econometric calibration of the stock price model

This subsection describes the procedure for estimating the SESJD model from a time series. This estimation procedure is divided into three steps. In the first step, an estimation of the switching-diffusion part is done with a Hamilton filter. The second and third steps involve detecting jumps with the peak over threshold (POT) and estimating parameters of the switching Hawkes intensity $\lambda(t)$ by log-likelihood maximization. Hainaut and Moraux (2018 b) go beyond this procedure by using their results as the initial set of parameters of a particle Markov chain Monte Carlo (PMCMC). However, in terms of log likelihood, there is no significant difference between these two

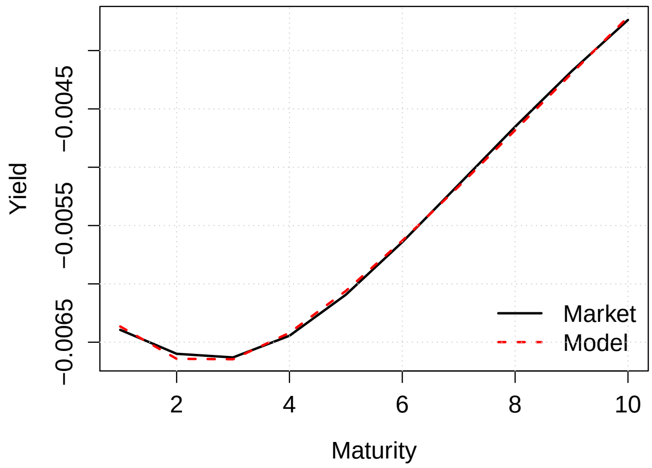


Figure 3.1: The observed yield curve as of April 16, 2020, and the yield curve derived from the Hull-White model where the initial forward curve follows a Nelson-Siegel model with the estimated parameters in Table 3.3.

procedures. We choose the POT procedure because it is less time consuming than the PMCMC procedure. The following paragraphs present, in detail, each step of our estimation procedure.
Let us denote by $y = \{y(t_1), ..., y(t_m)\}$ the time series of log returns of Euronext 100, measured at times $t_1, ..., t_m$ and equally spaced by a lag Δ of one day of trading:

$$y\left(t_j\right) = \ln\left(\frac{S\left(t_j\right)}{S\left(t_{j-1}\right)}\right) \qquad j = 1, ..., m. \tag{3.41}$$

We first estimate a switching geometric Brownian motion (SGBM) process with the Hamilton filter developed by Hamilton (1989). In fact, the Kalman filter previously used for the Hull-White process has been extended by Hamilton to the case where the unobservable state variable is a Markov chain $\zeta(t)$ and where we need the probabilities that the chain will be in each state at a given time. The vector of probabilities of being at time t_j in a certain state is computed recursively as follows:

$$\pi_{j+1} = \frac{f\left(y\left(t_j\right)\right) * \left(\pi_j^\top P_\Delta\right)}{\left\langle f\left(y\left(t_j\right)\right) * \left(\pi_j^\top P_\Delta\right), 1_{\mathbb{R}^N}\right\rangle}, \tag{3.42}$$

where the initial vector π_0 is defined such that the chain starts with its stationary distribution which is the eigen vector of the matrix P_Δ associated to the eigen value equal to 1. $f\left(y\left(t_j\right)\right)$ is the vector of Gaussian density for $y\left(t_j\right)$ in each of the N regimes, and $P_{\Delta} = \exp(Q_0 \Delta)$ is the daily matrix of transition probabilities. Note that in Equation (3.42), $*$ stands for the Hadamard product. Using the vectors of probabilities, the Hamilton filter's algorithm allows us to derive the following log-likelihood function for the m observations of the log return:

$$\ln L(y(t_1), ..., y(t_m)) = \sum_{j=1}^{m} \ln \left\langle f\left(y\left(t_j\right)\right), \left(\pi_j^{\top} P_{\Delta}\right)\right\rangle. \tag{3.43}$$

By maximizing the log-likelihood in Equation (3.43), we obtain (as displayed in Table 3.4) the estimated parameters of the SGBM process. The SESJD process extends the SGBM process by having a modified jump part driven by the hidden Markov chain. Given the state of the chain, we expect that this jump component captures the effect of rare events on the stock prices. To detect jumps, we apply the POT procedure in the second step.

We assume that it is likely that a jump will occur when the vector $y\left(t_j\right) - \tilde{\mu}^{ML}\left(t_j\right)\Delta$ is above or below some thresholds. These thresholds are defined by:

$$\sqrt{\Delta}\mathring{\sigma}\Phi^{-1}(\alpha_k),\ k = 1,2$$

where $\tilde{\mu}^{ML}\left(t_j\right)$ is the drift of the SGBM in the most likely state obtained by the Hamilton filter, $\mathring{\sigma}$ is the standard deviation of the whole sample and $\Phi^{-1}(\alpha_1)$ and $\Phi^{-1}(\alpha_2)$ are respectively the α_1 and α_2 percentiles of the standard normal distribution. These percentiles are such that $y\left(t_j\right) - \tilde{\mu}^{ML}\left(t_j\right)\Delta$ follows a Gaussian distribution when we consider the sample without jumps or the values of log returns limited to $\Phi^{-1}(\alpha_1)$ and $\Phi^{-1}(\alpha_2)$:

$$\alpha_1, \alpha_2 = \arg \min_{\text{sample without jumps}} \left((Skewness)^2 + (Kurtosis - 3)^2\right). \tag{3.44}$$

The thresholds obtained by solving the Equation (3.44) are presented in Figure 3.2, where the robustness of our choice is checked with the Jarque Bera test. Figure 3.2 enables us to link the upper or lower values of the log return to specific events, such as the credit crunch that occured between 2007 and 2009, the European sovereign debt crisis that occured between late 2009 and 2012 and the current COVID-19 crisis.

Furthermore, the POT procedure enables us to distinguish positive from negative jumps in the historical Euronext 100 prices, as in Figure 3.3. Regarding our observation period, there are often negative jumps on the Euronext 100 as a result of the negative effects of crises.

Once jumps are detected, the distribution of the return on the Euronext 100 prices is approximated by $\tilde{\mu}^{ML}\left(t_j\right)\Delta + J_j$ when there is a jump at t_j. Otherwise, the return on the stock is normally distributed as

$$\tilde{\mu}^{ML}\left(t_j\right)\Delta + \sigma^{ML}\left(t_j\right)\left(\hat{\rho} + \sqrt{1-\hat{\rho}^2}\right)W(\Delta),$$

where $\hat{\rho}$ is the sample correlation between the Euronext 100 prices and the Eonia rates taken as the risk-free rate proxy. $\sigma^{ML}\left(t_j\right)$ is the standard deviation at the most

$\bar{\sigma}_1$	9.45% (0.0026)	$\bar{\mu}_1$	26.62% (0.03814)
$\bar{\sigma}_2$	19.89% (0.0061)	$\bar{\mu}_2$	−6.908% (0.0826)
$\bar{\sigma}_3$	47.78% (0.0306)	$\bar{\mu}_3$	−82.90% (0.46)
$\left(p_{ij}(0,1\,\text{day})\right)_{i,j=1,2,3}$	state 1	state 2	state 3
state 1	0.9587 (0.014)	0.0413 (0.0135)	0.0000 (0.0005)
state 2	0.0074 (0.0023)	0.9585 (0.0077)	0.0343 (0.0054)
state 3	0.0000 (1.15×10^{-6})	0.0300 (0.0048)	0.9700 (0.0049)
Log.Lik.	12749.41	AIC	−25474.82
BIC	−25399.31		

Table 3.4: Parameters of a switching geometric Brownian motion, fitted with the Hamilton filter to the Euronext 100 time series and their standard errors in bracket.

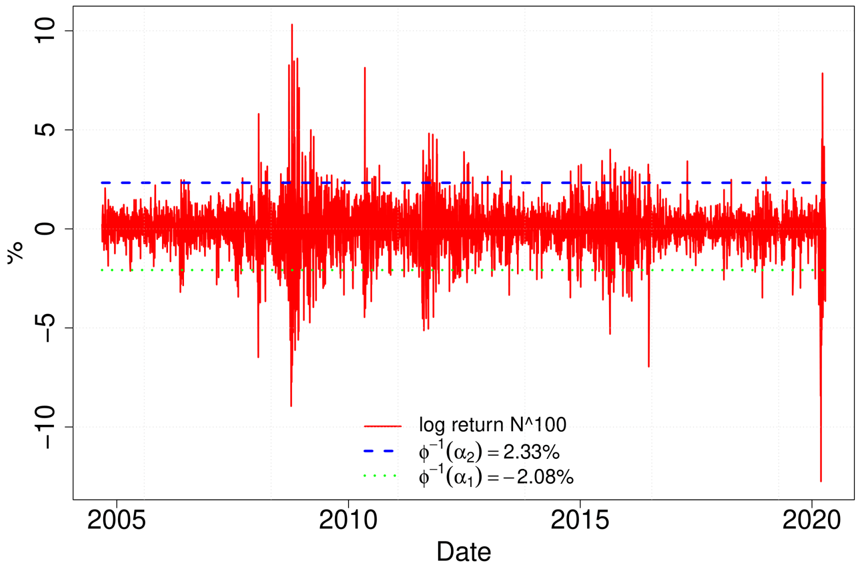


Figure 3.2: The historical log return of the Euronext 100 and the thresholds between which the log return follows a Gaussian distribution. The historical log-return of the Euronext 100 prices is over the thresholds mainly during crises such as the credit crunch between 2007 and 2009, the European sovereign debt crisis between late 2009 and 2012 and currently the COVID-19 crisis.

likely state obtained from the Hamilton filter procedure. We then estimate the jump parameters ρ^- , ρ^+ and p by maximizing the log likelihood of the jump probability density function for the log return observations where jumps occur. It follows that :

$$\rho^-,\rho^+,p \quad = \quad \arg\max \sum_{j=1}^{m} \ln \nu\left(y\left(t_j\right) - \tilde{\mu}^{ML}\left(t_j\right)\Delta \mid \rho^-,\rho^+,p\right) 1_{\{jump\,at\,t_j\}}.$$

The values found are reported in Table 3.5. The upward exponential jumps are less likely with a probability of 36%, but the average size of positive shocks $\frac{1}{\rho^+}$ is higher than that of negative shocks $-\frac{1}{\rho^-}$.
In the last step, we estimate the intensity parameters $\lambda(t)$. From Equation (3.3), over a time interval of length Δ, the variation of the intensity is given by

$$\Delta\lambda\left(t_j\right) = \beta\left(\eta^{ML}\left(t_{j-1}\right) - \lambda\left(t_{j-1}\right)\right)\Delta + \gamma J_j 1_{\{jump\,at\,t_j\}}.$$

We then find the estimated parameters of the intensity in Table 3.5 by log likelihood maximization as follows:

$$\beta,\gamma,\bar{\eta} \quad = \quad \arg\max \sum_{j=1}^{m} \ln \nu\left(y\left(t_j\right) \mid \rho^-,\rho^+,p\right) 1_{\{jump\,at\,t_j\}}.$$

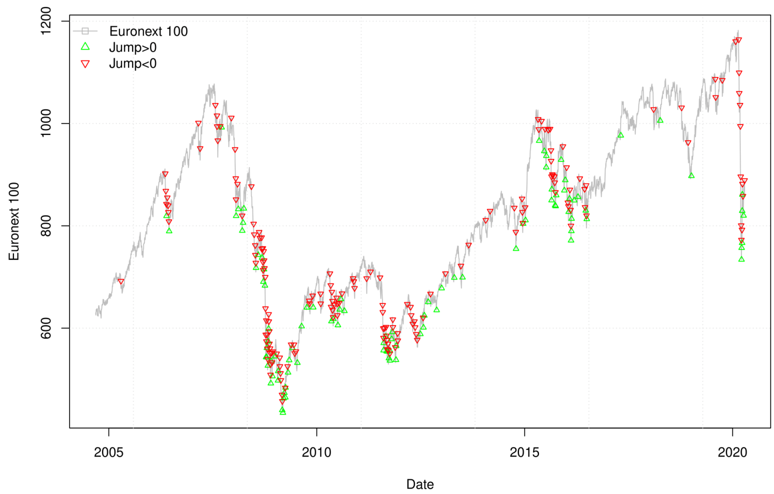
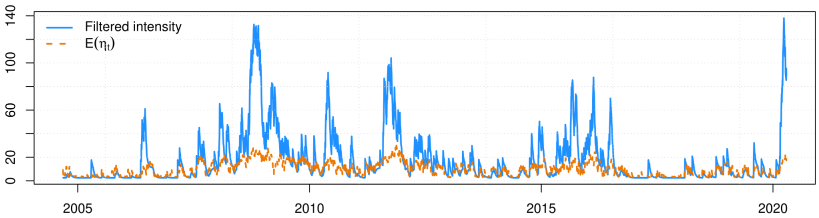


Figure 3.3: The filtered jumps in the historical Euronext 100 prices. There are mainly negative jumps on the Euronext 100 prices over our observation period.

$\bar{\sigma}_1$	8.924% (0.0012)	$\bar{\mu}_1$	26.62% (0.0404)
$\bar{\sigma}_2$	16.49% (0.0107)	$\bar{\mu}_2$	-6.908% (0.0855)
$\bar{\sigma}_3$	18.76% (0.0083)	$\bar{\mu}_3$	-82.90% (0.3328)
Log.Lik.	13075.3		
$\bar{\eta}_1$	2.59 (0.7406)	γ	15.12 (3.2936)
$\bar{\eta}_2$	11.03 (2.6134)	β	30.09 (7.587)
$\bar{\eta}_3$	44.82 (17.7978)		
Log.Lik.	849.22		
ρ^+	31.87 (3.2191)	ρ^-	-36.07 (2.7192)
p	0.36 (0.0289)		
Log.Lik.	473.62		

Table 3.5: Estimated parameters of the SESJD obtained based on the log return Euronext 100 time series and their standard errors in bracket.

In Figure 3.4, the upper graph depicts the average reversion level of the intensity and the filtered intensity obtained based on the historical Euronext 100 prices. We observe high jump intensities during crisis periods such as the credit crunch of 2008, the European sovereign debt crisis that occured between late 2009 and 2012 and the current COVID-19 crisis. During these periods, there is a large spread between jump intensities and the average reversion levels. As expected, during periods of economic growth this spread is close to zero, as is the jump intensity. Another feature is the clustering of peaks due to the self-excitation property. In the bottom graph in Figure 3.4, we draw the filtered state in the case of the SGBM process and the SESJD process. Their filtered state paths seem to be the same, but, during the credit crunch of 2008, the SGBM filter later detects the crisis regime, and this crisis regime lasts for a shorter period than with the SESJD filter.



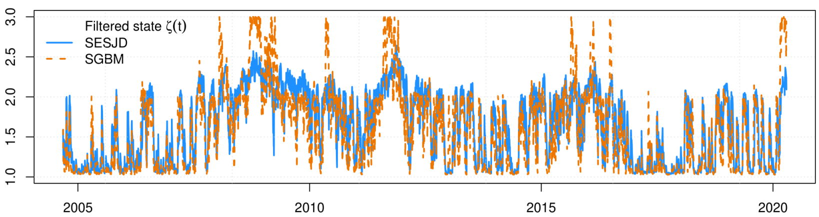


Figure 3.4: The upper graph depicts the filtered intensities of the SESJD process, while the lower graph depicts the filtered states of both the SESJD and SGBM processes.

3.4.4 Numerical analysis

In this subsection, we price the value of a VA specified in Table 3.6 and study the impact of our setting on this price. The choice of a guaranteed rate δ equal to zero is motivated by the low interest rate level with a spot rate $r(0) = -0.65\%$ as of April 16, 2020. The following analysis considers an insured 40-year-old person, who paid a unique premium of €1 for a GMAB and GMDB contract. To evaluate this contract, we consider the initial guaranteed amount $G(0)$ equal to the premium and we find the fair fee rate φ solution of the Equation (3.26).

Insured age (x)	40 years	guaranteed rate (δ)	0%
Premium (Π)	1	Maturity (T)	$\{1,2,...,20\}$

Table 3.6: Specification of variable annuity contract.

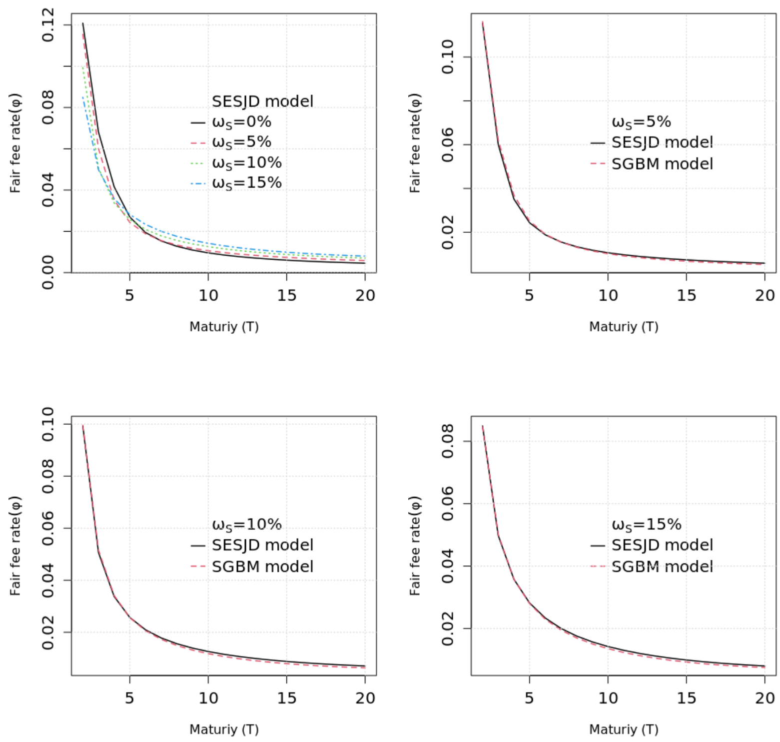


Figure 3.5: Fair fee rate φ of the GMAB and GMDB contract when we consider the roll-up guaranteed base. The guaranteed rate δ is assumed equal to zero, while the unique premium $A(0)$ is assumed equal to the initial guaranteed amount $G(0)$. The other parameters used are in Tables 3.4 and 3.5. Given the financial environment as of April 16, 2020, long term contract and a partial investment in risky asset allow to reduce the fair fee rate. For long-term contract, the SESJD model leads to a fair fee rate higher than the one from SGBM model.

Figure 3.5 presents the fair fee rate φ for different contract maturities. From the top left graph, the fair fee rate is not a strict monotonous function of the investment proportion in stock ω_S, while the fair fee rate is a decreasing function of the contract maturity. The GMAB and GMDB contract is then interesting for long term contract and a small investment in the Euronext 100. This is explained by positive yields on long term bonds and the cost of investment in risky assets. The top right graph and the bottom graphs compare the fair fee rates obtained using the SESJD model to those from the SGBM model for three investment scenarios, where the proportions invested in the Euronext 100 are $\omega_S = 5\%$, 10% and 15%, respectively. For higher maturities, these graphs show that the SESJD model leads to more fees than the SGBM

T	5	10	15	20
φ (%)	2.57	1.26	0.88	0.71
$L(t,T,\varphi)$	1.05	1.10	1.16	1.22

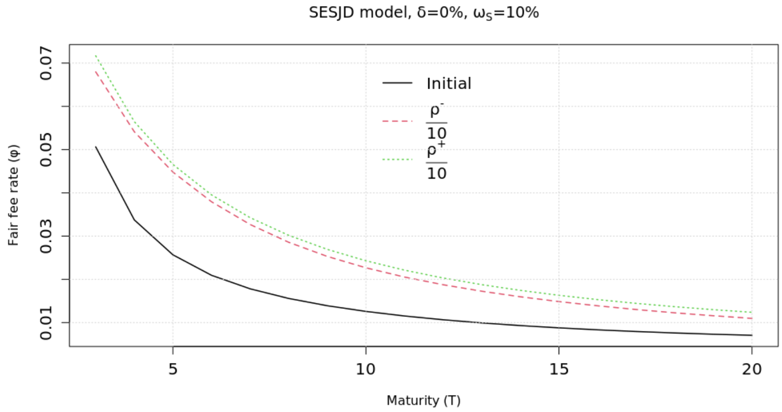
Table 3.7: Liability of the GMAB and GMDB contract when we consider the annual ratchet guaranteed. For each maturity T, we consider the fair fee rate φ found in roll-up guaranteed base. The liability in the roll-up guaranteed base is then equal to unique premium $\Pi = 1$. The liability in the annual ratchet guaranteed base obtained by Monte Carlo is higher than in roll-up guaranteed base due to the cost of additional options.

model. This is due to the jump component, which allows one to modulate the Euronext 100 volatility according to the economic regime. To analyze the effect of the jump component, the upper graph of Figure 3.6 depicts the fair fee rates when we invest 10% in the Euronext 100, let other parameters go unchanged and successively vary the double exponential jump parameters ρ^- and ρ^+. Recall that expectations of positive and negatives jumps are respectively equal to $\frac{1}{\rho^-}$ and $\frac{1}{\rho^+}$. Divide, respectively, ρ^- and ρ^+ by 10 to increase the probability of observing large downward and upward exponential jumps, which cause more self-excitation. From the upper graph in Figure 3.6, we conclude that the fair fee rate φ increases with the level of self-excitation. The lower graph depicts, for a 3% fee rate, the liability when the initial arrival rate of shocks $\lambda(0) = \zeta(0)^\top \bar{\eta}$ changes given the regime $\zeta(0) = e_1$ to e_3. In regime 3, which corresponds to a depressed economic context, the liability decreases dramatically with maturity due to an initial arrival rate of shocks higher than in regimes 1 and 2.

Figure 3.5 also reveals the difficulties of offering higher guarantees in a negative interest rate environment. From this figure, the fair fee rates range from 12% to 0.7% when we consider a guaranteed rate δ equal to zero and an initial guaranteed amount $G(0)$ equal to the unique premium. To reduce the fee rate and make the contract attractive despite low interest rates, a review of the guaranteed amount is considered. Figure 3.7 shows that the fair fee rate φ decreases as $G(0)$ goes down. The negative yields in the market constrain to charge more as the guaranteed amount is higher. For short term contracts, the reduction of the initial guaranteed amount significantly reduces the fair fee rate φ due to the uncertainty horizon.

For the sake of comparison between roll-up guaranteed and the annual ratchet guaranteed, we compute, using the Monte Carlo approach, the liability of the GMAB and GMDB contract for a maturity $T = 5, 10, 15, 20$ and the corresponding fair fee rate found in the roll-up guaranteed. The results presented in Table 3.7 show that the liability in the annual ratchet guaranteed base is higher than in the roll-up guaranteed base due to the cost of additional options.

We conclude this numerical analysis with a study of the impact of the correlation between the stock market and the yield curve. We assume a 40-year-old insured person underwriting a 20-year GMAB and GMDB contract with a 1.5% fee rate and consider



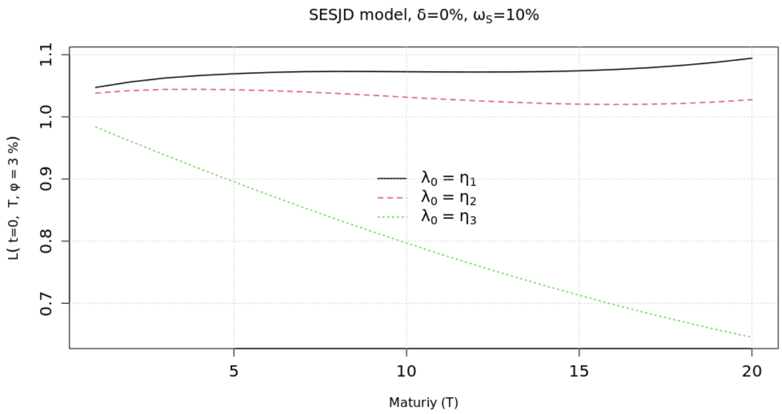


Figure 3.6: Impact of the self-excitation parameters on the fair fee rate and the liability of a GMAB and GMDB contract. From the top graph, the fair fee rate is shifted up when the exponential jump parameters ρ^- and ρ^+ decrease and create more self-excitation, whereas the bottom graph shows that the liability is shifted down when the initial arrival rate of shocks $\lambda(0) = \zeta(0)^\top \bar{\eta}$ increases as we consider the initial regime $\zeta(0) = e_1$ to e_3.

that 10% of the contract premium is invested in the Euronext 100. Figure 3.8 shows the impact of the correlation between interest rates and stock prices. In a context of low interest rates, this correlation has a small impact on liabilities, even if, from Equation (3.6), the market price of risk increases with the correlation ρ. In contrast, we observe that the value of liabilities decreases as Eonia rates and Euronext 100 prices are correlated.

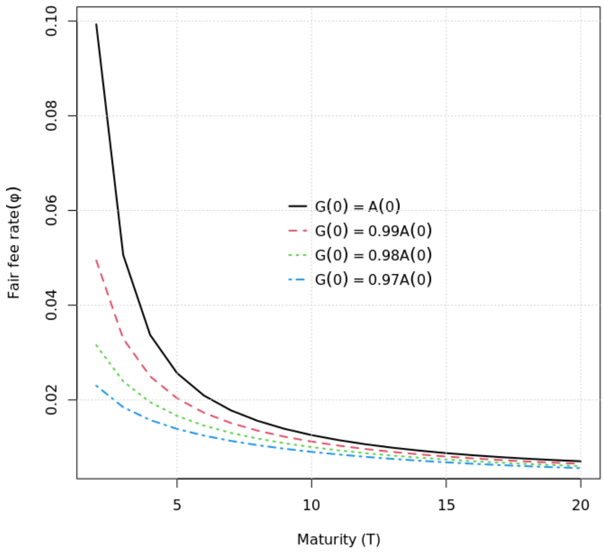


Figure 3.7: Impact of the initial guaranteed amount $G(0)$ on the fee rate φ for the GMAB and GMDB contract. We consider the roll-up guaranteed base with a guaranteed rate $\delta = 0\%$ and a proportion invested in the Euronext 100 $\omega_S = 10\%$; the other parameters used are presented in Tables 3.4 and 3.5. Given the market condition as of April 16, 2020, where the spot rate $r(0) = -0.65\%$, reducing the initial guaranteed amount allows one to reduce the fair fee rate φ.

3.5 Conclusions

In this chapter, we combined regime-switching and self-excitation features to model the underlying investment fund price of VA contracts with a minimum death and accumulated life benefit. This underlying investment fund is a portfolio of a bond and a risky asset correlated to the interest rate. Despite the dependency on the chain and the correlation between the stock and the interest rate process, we rewrote the commitment of such a VA issuer with a roll-up guaranteed base as a linear combination of put option prices. For other guaranteed bases, the dependency on the chain requires the use of the Monte Carlo approach.

This chapter presented a detailed economic calibration of our models based on the historical data of the Eonia rates and the Euronext 100 prices from September 06, 2004, to April 16, 2020. The estimated parameters found for the SGBM and the SESJD processes enable us to draw three conclusions. First, ignoring the self-excitation modulated by the Markov chain leads to an underestimation of the fair fee rate of the VA contract. The features of the SESJD process are thus recommended, as they allow one to not only consider the economic regime, but also to capture random shocks in the

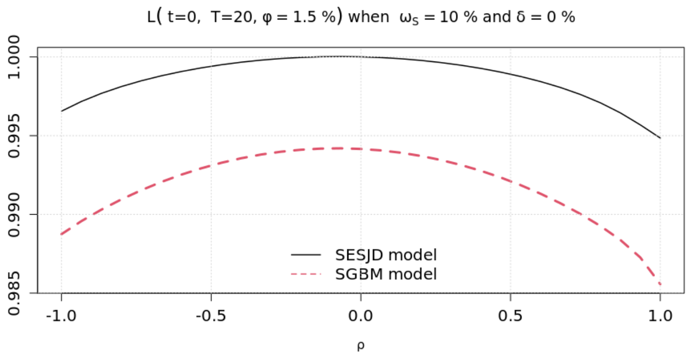


Figure 3.8: Influence of the correlation between Eonia rates and Euronext 100 prices on the liability of a 20 years GMAB and GMDB contract with a 1.5% fee rate, with 10% of the premium being invested in the Euronext 100. Based on our data, the value of the liability decreases as Eonia rates and Euronext 100 prices are correlated.

stock market and to scale their amplitude given the regime. Second, in a low interest rate environment, reducing the guarantee allows one to decrease the fair fee rate. Finally, the GMAB and GMDB contracts are sensitive to the economic regime and the self-excitation level.

From a broader perspective, the chapter calls for consideration of other guarantees such as the surrender benefit or the withdrawal benefit. From a practical perspective, it would be interesting to extend this work to the valuation of a large portfolio of VA contracts based on risk measures.

Appendices

3.A Appendix A

The probability function (pdf) $v(z)$ is defined by the three parameters $\rho^+ \in \mathbb{R}^+$, $\rho^- \in \mathbb{R}^-$ and $p \in (0,1)$:

$$v(z) = p\rho^+ e^{-\rho^+ z} 1_{\{z \geq 0\}} - (1-p)\rho^- e^{-\rho^- z} 1_{\{z<0\}},$$

where p and $(1-p)$ are respectively the probabilities of observing upward and downward exponential jumps.

The average sizes of positive and negative shocks are equal to $\frac{1}{\rho^+}$ and $\frac{1}{\rho^-}$. The expectation of J_i is the weighted sum of the expected average jumps:

$$\mathbb{E}(J_i) = p\frac{1}{\rho^+} + (1-p)\frac{1}{\rho^-}.$$

In later developments, the moment-generating function for the sum of the J_i and of its absolute value is needed. We define it as follows:

$$\begin{aligned}\psi_J(z_1,z_2) &:= \mathbb{E}\left(e^{z_1 J + z_2 |J|}\right)\\ &= p\frac{\rho^+}{\rho^+ - (z_1+z_2)} + (1-p)\frac{\rho^-}{\rho^- - (z_1 - z_2)},\end{aligned}$$

under the condition that $(z_1+z_2) < \rho^+$ and $(z_1+z_2) > \rho^-$. In particular,

$$\mathbb{E}\left(e^J - 1\right) = p\frac{\rho^+}{\rho^+ - 1} + (1-p)\frac{\rho^-}{\rho^- - 1} - 1.$$

3.B Appendix B

If we note $f(t,s,r(t),\lambda(t),X(t),\zeta(t)) = \mathbb{E}^{\mathbb{Q}}\left(e^{\omega_1 \int_t^s r(u)\mathrm{d}u + \omega_2 X(s)} \mid \mathscr{F}_t\right)$, f is the solution for an Ito's equation for the semi-martingale. Let $\mathscr{A}$ the infinitesimal generator of $(r(t),\lambda(t),X(t),\zeta(t))$. By definition, this infinitesimal generator equals

$$\lim_{u\to t} \frac{\mathbb{E}^{\mathbb{Q}}(f(u,s,r(u),\lambda(u),X(u),\zeta(u)) \mid \mathscr{F}_t) - f(t,s,r(t),\lambda(t),X(t),\zeta(t))}{u-t}.$$

With enough regularity we have:

$$\mathscr{A} f(t,s,r(t),\lambda(t),X(t),\zeta(t)) = -\omega_1 r(t) f(t,s,r(t),\lambda(t),X(t),\zeta(t)). \tag{3.45}$$

Since $f(t,s,r(t),\lambda(t),X(t),\zeta(t))$ is the solution of an Ito's equation for the semi-martingale, if $\zeta(t)=e_i, i\in\mathcal{N}$, then:

$$\begin{aligned}
-\omega_1 r(t) f = {} & f_X\left(r(t)-\omega_S\lambda(t)\mathbb{E}^{\mathbb{Q}}\left(e^J-1\right)\right.\\
& \left.-\frac{1}{2}\left(\left(\omega_S\rho\bar{\sigma}_i-(1-\omega_S)\sigma_r B^r(t,T)\right)^2+\omega_S^2\left(1-\rho^2\right)\bar{\sigma}_i^2\right)\right)+f_t\\
& +\left(\theta^{\mathbb{Q}}(t)-a_r r(t)\right)f_r+\frac{1}{2}\sigma_r^2 f_{rr}-\sigma_r(1-\omega_S)\sigma_r B^r(t,T) f_{rX}\\
& +\sigma_r\omega_S\rho\bar{\sigma}_i f_{rX}+\frac{1}{2}\left(\left(\omega_S\rho\bar{\sigma}_i-(1-\omega_S)\sigma_r B^r(t,T)\right)^2\right)f_{XX}\\
& +\frac{1}{2}\omega_S^2\left(1-\rho^2\right)\bar{\sigma}_i^2 f_{XX}+\beta\left(\bar{\eta}_i-\lambda(t)\right)f_\lambda+\lambda(t)\\
& \times\int_{-\infty}^{+\infty}\left(f\left(t,s,r(t),\lambda(t)+\gamma|z|,X(t)+\ln\left(1+\omega_S\left(e^z-1\right)\right),e_i\right)-f(\cdot)\right)\nu(\mathrm{d}z)\\
& +\sum_{j\neq i}^{\mathcal{N}} q_{i,j}\left(f\left(t,s,r(t),\lambda(t),X(t),e_j\right)-f\left(t,s,r(t),\lambda(t),X(t),e_i\right)\right).
\end{aligned} \tag{3.46}$$

[4]Assume that f is an exponential affine function of $r(t)$, $\lambda(t)$ and $X(t)$:

$$\begin{aligned}
f(t,s,r(t),\lambda(t),X(t),e_i)=\exp\Big(& A(\bar{\omega},t,s,e_i)+B(\bar{\omega},t,s)\,r(t)\\
& +C(\bar{\omega},t,s)\,\lambda(t)+\omega_2 X(t)\Big),
\end{aligned} \tag{3.47}$$

where $\bar{\omega}=(\omega_1,\omega_2)$. $A(\bar{\omega},t,s,e_i)$ for $i\in\mathcal{N}$, $B(\bar{\omega},t,s)$ and $C(\bar{\omega},t,s)$ are time dependent functions. Since $f(s,s,r(s),\lambda(s),X(s),\zeta(s))=e^{\omega_2 X(s)}$ we have the following terminal condition $A(\bar{\omega},s,s,e_i)=0$ for $i\in\mathcal{N}$, $B(\bar{\omega},s,s)=0$ and $C(\bar{\omega},s,s)=0$. We can simplify the two last terms of Equation (3.46) as follows:

$$\begin{aligned}
\int_{-\infty}^{+\infty}\left(f\left(t,s,r(t),\lambda(t)+\gamma|z|,X(t)+\ln\left(1+\omega_S\left(e^z-1\right)\right),e_i\right)-f(\cdot)\right)\nu(\mathrm{d}z)= \\
f\left(t,s,r(t),\lambda(t),X(t),e_i\right)\left(\mathbb{E}^{\mathbb{Q}}\left(g(J,C)\right)-1\right)
\end{aligned}$$

where $g:(x,C)\longmapsto e^{\omega_2\ln(1+\omega_S(e^x-1))+C\gamma|x|}$ is a real valued bivariate function. As shown in Appendix C 3.C, this function admits a closed-form expression when $\omega_2=1$. Since $q_{ii}=-\sum_{j\neq i,j\in\mathcal{N}} q_{i,j}$, the last term becomes:

$$\begin{aligned}
\sum_{j\neq i,j\in\mathcal{N}} q_{i,j}\left(f\left(t,s,r(t),\lambda(t),X(t),e_j\right)-f\left(t,s,r(t),\lambda(t),X(t),e_i\right)\right)= \\
\sum_{j\in\mathcal{N}} q_{i,j} f\left(t,s,r(t),\lambda(t),X(t),e_j\right).
\end{aligned}$$

[4]In this chapter we simplify the notation of partial derivatives as follows $\frac{\partial f}{\partial t}=f_t$, $\frac{\partial f}{\partial r}=f_r$, $\frac{\partial^2 f}{\partial r^2}=f_{rr}$...

Then, from Equation (3.46), we derive this PDEs system:

$$\begin{cases} 0 = A_t + \theta^{\mathbb{Q}}(t)B + \frac{1}{2}\sigma_r^2 B^2 + \sigma_r\left(\omega_S\rho\bar{\sigma}_i - (1-\omega_S)\sigma_r B^r(t,T)\right)\omega_2 B \\ \qquad -\frac{\omega_2}{2}\left(\left(\omega_S\rho\bar{\sigma}_i - (1-\omega_S)\sigma_r B^r(t,T)\right)^2 + \omega_S^2\left(1-\rho^2\right)\bar{\sigma}_i^2\right) \\ \ +\frac{\omega_2^2}{2}\left(\left(\omega_S\rho\bar{\sigma}_i - (1-\omega_S)\sigma_r B^r(t,T)\right)^2 + \omega_S^2\left(1-\rho^2\right)\bar{\sigma}_i^2\right) + \beta\bar{\eta}_i C \\ \qquad +\sum_{j\in\mathcal{N}} q_{i,j}\exp\left(A\left(\bar{\omega},t,s,e_j\right) - A\left(\bar{\omega},t,s,e_i\right)\right) \quad for \quad i\in\mathcal{N} \\ 0 = B_t - a_r B + \omega_2 + \omega_1 \\ 0 = C_t - \omega_2\omega_S\mathbb{E}^{\mathbb{Q}}\left(e^J - 1\right) - \beta C + \mathbb{E}^{\mathbb{Q}}\left(g(J,C)\right) - 1 \end{cases} \tag{3.48}$$

Let $\tilde{A}(\bar{\omega},t,s) = \left(e^{A(\bar{\omega},t,s,e_1)},\ldots,e^{A(\bar{\omega},t,s,e_N)}\right)$ be a vector of functions. Then, the first equation of the System (3.48) has the following matrix form:

$$\frac{\partial\tilde{A}(\bar{\omega},t,s)}{\partial t} + (\operatorname{diag}(D(\bar{\omega},t,s)) + Q_0)\tilde{A}(\bar{\omega},t,s) = 0,$$

where $D(\bar{\omega},t,s)$ is a vector of functions defined by

$$\begin{aligned} D(\bar{\omega},t,s) = &-\frac{\omega_2}{2}\left(\left(\omega_S\rho\bar{\sigma} - (1-\omega_S)\sigma_r B^r(t,T)\right)^2 + \omega_S^2\left(1-\rho^2\right)\bar{\sigma}^2\right) \\ &+\theta^{\mathbb{Q}}(t)B + \frac{1}{2}\sigma_r^2 B^2 + \sigma_r\left(\omega_S\rho\bar{\sigma} - (1-\omega_S)\sigma_r B^r(t,T)\right)\omega_2 B \\ &+\frac{\omega_2^2}{2}\left(\left(\omega_S\rho\bar{\sigma} - (1-\omega_S)\sigma_r B^r(t,T)\right)^2 + \omega_S^2\left(1-\rho^2\right)\bar{\sigma}^2\right) + \beta\bar{\eta}C. \end{aligned}$$

We complete the proof of this proposition with the solution of the second equation of the PDEs system (3.48):

$$B(\bar{\omega},t,s) = (\omega_2 + \omega_1)B^r(t,s).$$

3.C Appendix C

For a real-valued function $C(\cdot)$ and $\omega_2 = 1$, if

$$g : (x,C) \longmapsto e^{\omega_2\ln(1+\omega_S(e^x-1)) + C\gamma|x|}$$

then

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}(g(J,C)) &= \mathbb{E}^{\mathbb{Q}}\left(e^{\ln\left(1+\omega_S\left(e^J-1\right)\right) + C(\bar{\omega},t,s)\gamma|J|}\right) \\ &= \mathbb{E}^{\mathbb{Q}}\left(\left(1+\omega_S\left(e^J-1\right)\right)e^{C(\bar{\omega},t,s)\gamma|J|}\right) \\ &= (1-\omega_S)\psi_J(0,\gamma C(\bar{\omega},t,s)) + \omega_S\psi_J(1,\gamma C(\bar{\omega},t,s)). \end{aligned}$$

3.D Appendix D

Using Simpson's rule, we approximate the integral in Equation (3.20) as follows :

$$p_T(k) \approx \frac{e^{\alpha k}}{\pi}\sum_{l=1}^{M}\delta_l e^{-iv_l k}\mathscr{F}(\bar{p}_T)(v_l)\Delta_v \tag{3.49}$$

where $\delta_l = \frac{1}{3}\left(3+(-1)^j - 1_{\{j=1\}}\right)$. The Fourier transform of the modified put option $\mathscr{F}(\bar{p}_T)(v_l)$ is determined as follows:

$$\begin{aligned}
\frac{\mathscr{F}(\bar{p}_T)(v_l)}{V(t)} &= \int_{-\infty}^{+\infty} e^{iv_l k} \int_{-\infty}^{k} \int_{-\infty}^{+\infty} e^{-y_1} \left(e^{(1-\alpha)k} - e^{y_2-\alpha k}\right) \psi(y_1, y_2)\, \mathrm{d}y_1 \mathrm{d}y_2 \mathrm{d}k \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \int_{y_2}^{+\infty} e^{iv_l k - y_1} \left(e^{(1-\alpha)k} - e^{y_2-\alpha k}\right) \psi(y_1, y_2)\, \mathrm{d}k \mathrm{d}y_1 \mathrm{d}y_2 \\
&= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-y_1} \psi(y_1, y_2) \left(\frac{e^{(1-\alpha+iv_l)y_2}}{iv_l - \alpha} - \frac{e^{(1-\alpha+iv_l)y_2}}{1+iv_l-\alpha}\right) \mathrm{d}y_1 \mathrm{d}y_2 \\
&= \frac{M_{r,X}^{t,T}(-1, 1-\alpha+iv_l)}{\alpha^2 - \alpha - v_l^2 + i(-2\alpha+1)v_l}. \qquad (3.50)
\end{aligned}$$

From Equation (3.19), $\mathscr{F}(\bar{p}_T)(0)$ should be finite, and α is non-negative. It follows that the put option price is well-defined for any $\alpha > 0$ such that $M_{r,X}^{t,T}(-1, 1-\alpha) < +\infty$. In contrast, for $k_j = -k_{max} + (j-1)\Delta_k$, $\Delta_k = \frac{2k_{max}}{M}$, $\Delta_v = \frac{2\pi}{M\Delta_k}$ and $v_l = (l-1)\Delta_v$, the product $v_l k_j$ is equal to

$$\begin{aligned}
v_l k_j &= -\frac{M}{2}(l-1)\Delta_v\Delta_k + (j-1)(l-1)\Delta_v\Delta_k \qquad (3.51) \\
&= -(l-1)\pi + (j-1)(l-1)\frac{2\pi}{M},
\end{aligned}$$

and

$$e^{-iv_l k} = (-1)^{l-1} e^{-\frac{2\pi i}{M}(j-1)(l-1)} \qquad (3.52)$$

Equations (3.50) and (3.52) to (3.49) allow us to end this proof.

Chapter 4

Long memory self-exciting jump diffusion for asset prices modelling

4.1 Introduction

Empirical studies on asset returns (see e.g. Bollerslev et al. 1992, Pagan 1996 and Cont 2001) suggest that any suitable asset return model should reproduce the following non-exhaustive stylized facts: 1. asset log-returns have a leptokurtic distribution; 2. the Taylor effect: autocorrelations of powers of absolute returns are higher than at power one; 3. asset prices are non stationary times series and possess a local trend at the least; 4. after a large or small change of either sign a large or small price change tends to occur, it is called volatility clustering. From these facts, the geometric Brownian motion with constant drift and volatility of the famous Black and Scholes model (1973) is unappropriated due to its thin tail distribution which makes it unable to capture large changes on the asset return. The seek of stochastic processes having these four realistic features of asset price has conducted researchers to modify the Black and Scholes model by incorporating a jump component.

In a pioneering work, Merton (1976) adds to the geometric Brownian motion a jump component driven by a Poisson process with jump sizes log normally distributed. This jump component explains abnormal changes in the asset price due to the arrival of important new information about the stock. Unlike the framework of Merton (1976), Kou (2002) incorporates the asymmetric leptokurtic features of asset price by using jump sizes that are double exponential distributed. Considering Poisson jumps of time-varying intensity, Andersen et al. (2002) extend the class of stochastic volatility diffusions (see e.g. Melino and Turnbull 1990 and Heston 1993) for asset returns and improve the model capacity to replicate stylized facts of asset returns. To validate the use of a jump diffusion, Carr and Wu (2003) develop a robust tool that confirms the presence of a jump component in the S&P 500 index. Eraker (2004) studies the empirical performance of jump diffusion models for asset price and concludes that they better fit options and stock markets data simultaneously. Meanwhile jump diffusion models of above works enable one to reproduce the skewness and the excess kurtosis observed in asset returns, they rely on the assumption that the arrivals of shocks are independently and/or identically distributed. Such models do not allow for clustered jumps, where the occurrence of one event increases the probability of a new event occurring. As Maheu and McCurdy (2004) point out, jump diffusion models should reproduce the clustering of jumps in asset prices due to news releases.

To deal with the clustering of jumps, self-exciting point processes of Hawkes type (Hawkes 1971 a, Hawkes 1971 b) have been recently introduced in the financial and economics literature. For instance, Errais et al. (2010) use Hawkes process (HP) to model the default clustering in the context of portfolio credit risk. Aït-Sahalia et al. (2015) introduce a multivariate HP for simultaneously modelling multiple asset returns. They conclude that self- and cross-exciting processes capture the clustering patterns observed in stock markets. Hainaut and Moraux (2018 a) hedge stock options when the dynamics of the underlying stock combine a diffusion process with a jump HP called self-exciting jump diffusion. Hainaut and Moraux (2018 b) extend Hainaut and Moraux (2018 a) by modulating the self-exciting jump diffusion with a Hidden Markov chain that represents the evolution of economics regimes. The common feature of these works is the use of HPs with an exponential kernel in order to incorporate the clustering of jumps. While this enables that such a HP is markovian and so guarantees HP's moments to take a closed form, the use of an exponential kernel implies that the length of jump effect on the value of today's intensity decreases exponentially with time. Brignone and Sgarra (2019) employ the Hawkes modelling setting of Hainaut and Moraux (2018 a), to propose a method for pricing Asian options in market models with the risky asset dynamics driven by a Hawkes process with exponential kernel. In the stochastic volatility class, Bernis et al. (2021) propose a G-OU Hawkes model with exponential kernel, that extends the G-OU Barndorff-Nielsen and Shephard model (2001; 2001) by taking into account jump clustering phenomena.

In the present chapter, we replace the exponential kernel by a Mittag-Leffler (ML) function firstly defined in Mittag-Leffler (1903) as a power series. Our ML function is sub-exponential and may be seen as a generalization of the exponential kernel. In this setting, the HP intensity remembers past jumps for a longer period than in the exponential framework. Chen et al. (2021) use a similar approach in their work to obtain what they call a fractional HP. Their ML type kernel is a probability density function of a positive random measure that decays as a power law and the HP is asymptotically stationary. Bacry and Muzy (2014) and Bacry et al. (2016) provide empirical evidence of slow decay kernel importance for high frequency order book modelling and conclude that the intensity of market events should depend on their history and on the state of the order book.

We contribute to the existing literature in four directions. We first propose a HP with a non-exponential kernel that is an infinite dimensional Markov process and allows for longer effects of past HP's jumps on future intensity than a HP with exponential kernel. Secondly, we introduce a way to incorporate the clustering of jumps in asset returns dynamics by combining to the geometric Brownian motion a HP jump with a ML type kernel. Thirdly, we provide a closed form expression of the conditional moment

generating function for log-returns and derive European options prices. Finally, we provide an estimation procedure of our model based on an asset prices time series data.
The present chapter is organized as follows. Section 4.2 introduces the main definitions and properties of Hawkes with the ML type kernel that we use along the chapter. In this section our HP with ML kernel is rewritten as an infinite dimensional Markov process and the conditional moment generating function of the intensity of our HP and its integral is derived through a discretisation approximation. Section 4.3 focuses on the application of our HP with ML kernel type to asset prices modelling and European options pricing. It begins by presenting the self- and cross-exciting framework that we use, and its estimation procedure. Subsequently, it provides the conditional moment generating function for log-returns. Finally, it presents a closed form expression for European call option prices based on Fourier transform and provides a sensitivity analysis to the memory kernel parameter based on the S&P 500 data.

4.2 Hawkes processes with Mittag-leffler type kernel

4.2.1 Hawkes processes

The HP (Hawkes 1971 a; Hawkes 1971 b) is a self-exciting point process in which the occurrence of an event increases the probability of further events occurring. Let us consider a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ on which is defined a HP (N_t, λ_t). Ω is the sample space, the filtration $(\mathscr{F}_t)_{t\geq 0}$ is a family of σ-algebras generated by all the processes defined in this chapter. $\mathbb{P}$ is the real world probability measure. The following equation gives a general definition of a HP (N_t, λ_t),

$$\lambda_t|\mathscr{F}_{t\_} \quad = \lim_{h\to 0} \frac{\mathbb{E}(N_{t+h}|\mathscr{F}_t) - N_t}{h} = \quad \lambda_0 + \eta \int_0^{t-} f(t-u)\, \mathrm{d}N_u \tag{4.1}$$

where the filtration $\mathscr{F}_{t\_}$ stands for the information available on the HP up to time t but not including. $\lambda_0 > 0$ is the background intensity, describing the arrival of events triggered by external sources, η scales the jump size of the HP intensity when a jump occurs and $f : \mathbb{R}_+ \mapsto \mathbb{R}_+$ is called the self-excitation kernel or excitation function of the process (N_t, λ_t). In a small time interval of length h, conditionally to the information until time t the intensity λ_t is also linked to the number of event N_t by the probability of observing m events over the time interval $[t, t+h]$:

$$P\left(N_{t+h} - N_t = m|\mathscr{F}_t\right) = \begin{cases} \lambda_t h + o(h) & \text{if } m = 1 \\ o(h) & \text{if } m > 1 \\ 1 - \lambda_t h + o(h) & \text{if } m = 0 \end{cases}$$

where $o(h)$ is such that $\lim_{h\to 0} \frac{o(h)}{h} = 0$.
The excitation function f in equation (4.1)provides the contribution to the intensity λ_t of an event that occurred at a previous time $u < t$. Hence, each event increases the intensity by $\eta \lim_{h\to 0^+} f(h)$, which then decays according to the excitation function

f until the next event occurs to push it up again. The excitation is decreasing so that more recent events have higher influence on the current intensity compared to events having occurred further away in the time. Among the most popular excitation function we distinguish :

- Exponential kernel
$$f(u) = e^{-\alpha u}, u > 0 \tag{4.2}$$
where α is the decay rate to a stable situation after an event. The HP with exponential kernel has the Markov property since conditional on the present, the future is independent of the past :
$$\lambda_{t_2} = \left(1 - e^{-\alpha(t_2-t_1)}\right)\lambda_0 + e^{-\alpha(t_2-t_1)}\lambda_{t_1} + \eta \int_{t_1}^{t_2-} e^{-\alpha(t_2-u)}\mathrm{d}N_u,$$
where $t_2 > t_1$. This kernel is widely used since it provides the Markov property to the HP (N_t, λ_t), and thus simplifies a lot of calculations. The exponential kernel is indicated for the modelling of short term effects as the effect of past jumps on future intensity decay quickly to zero.

- Power law kernel
$$f(u) = \frac{1}{(1+\alpha u)^{1+p}}, u > 0 \tag{4.3}$$
with $\alpha > 0$ and $p > 0$, makes the HP non-Markov and allows for a slower form of decay than the exponential kernel. This kernel function is used for modelling of long term effects.

To model short and long term effects, it is common to use a kernel that handles in an unified way properties of the exponential kernel and the power law kernel.

4.2.2 Mittag-Leffler type kernel

In this chapter, we consider an excitation function by using the Mittag-Leffler (ML) function of order $\alpha \in [0,1]$ defined as follows:

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(\alpha n + 1)}, z \in \mathbb{C} \tag{4.4}$$

where the Gamma function Γ is :

$$\Gamma(z) = \int_0^\infty e^{-x} x^{z-1} \mathrm{d}x,$$

with $\mathrm{Re}(z) > 0$. The ML function decays in a hybrid way between exponential and power law function as :

$$E_\alpha(-u) = \begin{cases} \frac{1}{1+u} & \text{for } \alpha = 0 \text{ and } 0 < u < 1, \\ e^{-u} & \text{for } \alpha = 1 \text{ and all } u > 0. \end{cases} \tag{4.5}$$

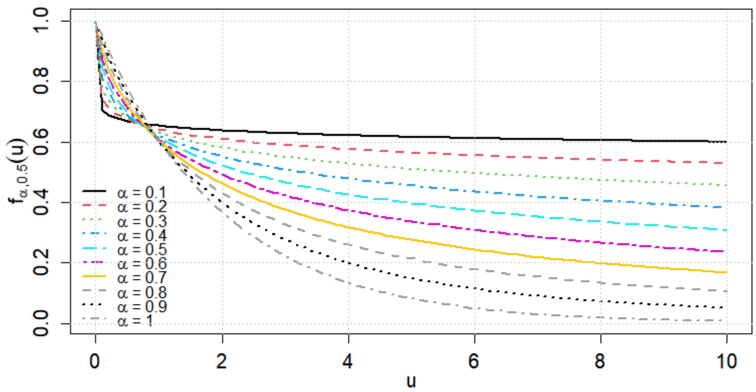


Figure 4.1: Curve of the Mittag-Leffler function $u \mapsto f_{\alpha,\beta=0.5}(u)$ for different $\alpha \in (0,1]$.

We refer the reader to Gorenflo et al. (2014) for properties of the ML function. In this chapter, we take into account the impact of past jump sizes on future intensity and define the HP intensity λ_t through the following relation:

$$\lambda_t|\mathscr{F}_{t_-} \quad = \quad \lambda_0 + \eta \int_0^{t-} f_{\alpha,\beta}(t-u)\,\mathrm{d}L_u \tag{4.6}$$

where $L_t = \sum_{i=1}^{N_t} |J_i|$ is the total jump sizes till time t. $\{J_i, i \geq 1\}$ are the jump sizes assumed to be independent and identically distributed random variables with density $\nu(z)$ on $\mathbb{R}$. The HP's excitation function $f_{\alpha,\beta}$ of ML type is

$$f_{\alpha,\beta}(u) \quad := \quad E_\alpha\left(-\beta u^\alpha\right),\ \alpha \in (0,1],\ \beta > 0. \tag{4.7}$$

Note that for $\alpha = 1$, we recover the exponential case $f_{1,\beta}(u) = e^{-\beta u}$. The choice of this kernel is motivated by its behavior in short and long term as shown in the figure 4.1. At short term $u \to 0^+$, Mainardi (2020) shows that $f_{\alpha,\beta}$ tends to a stretched exponential function for $\alpha \in (0,1)$,

$$f_{\alpha,\beta}(u) \quad \sim \quad \exp\left(-\frac{\beta u^\alpha}{\Gamma(1+\alpha)}\right). \tag{4.8}$$

While from Erdélyi et al. (1955), we know that when $u \to \infty$, the function $f_{\alpha,\beta}$ behaves as a negative power law function for $\alpha \in (0,1)$,

$$f_{\alpha,\beta}(u) \quad \sim \quad \frac{\left(\beta^{\frac{1}{\alpha}} u\right)^{-\alpha}}{\Gamma(1+\alpha)}. \tag{4.9}$$

From Hawkes (1971 b), we have that, if $\left|\left|f_{\alpha,\beta}\right|\right| := \int_0^\infty f_{\alpha,\beta}(u)\,\mathrm{d}u > 1$ then the HP is non-stationary. Given the asymptotic equivalences of our ML kernel (4.8) and (4.9), for $\alpha \in (0,1)$ its follows that the HP with kernel $f_{\alpha,\beta}$ (4.7) is non-stationary.
On the other hand, from Gorenflo et al. (2014) (Equation 3.7.8), the Laplace transform of $f_{\alpha,\beta}$ is given by:

$$F_{\alpha,\beta}(s) := \mathscr{L} f_{\alpha,\beta}(u)(s) \quad = \quad \frac{s^{\alpha-1}}{s^{\alpha}+\beta}, \tag{4.10}$$

where $\alpha > 0$, $\beta > 0$ and $Re(s) > |\beta|^{\frac{1}{\alpha}}$. Gorenflo et al. (2014) prove in Theorem 3.10 that the ML function of negative argument $E_{\alpha}(-\beta u^{\alpha})$ is completely monotonic for all $0 \leq \alpha \leq 1$ and $\beta \in \mathbb{R}_{+}$. Meaning that $f_{\alpha,\beta}$ possesses derivatives $f^{(n)}_{\alpha,\beta}$ of any order $n = 0,1,2,...$ and the derivatives ares alternating in sign:

$$(-1)^n f^{(n)}_{\alpha,\beta}(u) \quad \geq \quad 0, \forall u \in \mathbb{R}_{+}.$$

From Widder (1946) page 161, this is equivalent to the existence of a representation of the function $f_{\alpha,\beta}$ in the form of a Laplace-Stieltjes integral with non-decreasing density and non-negative measure $H_{\alpha,\beta}$:

$$f_{\alpha,\beta}(u) \quad = \quad \int_0^\infty e^{-uv} H_{\alpha,\beta}(v)\,\mathrm{d}v, \tag{4.11}$$

with

$$H_{\alpha,\beta}(v) \quad = \quad \frac{1}{\pi}\frac{\beta v^{\alpha-1}\sin(\alpha\pi)}{v^{2\alpha}+2\beta v^{\alpha}\cos(\alpha\pi)+\beta^2}, \tag{4.12}$$

for every $\alpha \in (0,1)$ and $\beta > 0$. The function $H_{\alpha,\beta}$ is derived by inverting the Laplace transform of $f_{\alpha,\beta}$ (4.10) as in Gorenflo and Mainardi (1997). For $\alpha \in (0,1)$, $\beta > 0$ and $v > 0$, the numerator of $H_{\alpha,\beta}(v)$ is always non- negative since $\sin(\alpha\pi) \geq 0$. Its denominator is as well non-negative since it is greater than $(v^{\alpha}-\beta)^2$. Given the asymptotic equivalence (4.8), $f_{\alpha,\beta}(0^+) = 1$ and from equation (4.11) we find $\int_0^\infty H_{\alpha,\beta}(v)\,\mathrm{d}v = 1$ which demonstrates that $H_{\alpha,\beta}$ is indeed a non-negative measure on $\mathbb{R}_{+}$. From equation (4.12), we note that $H_{\alpha,\beta}$ is undefined at point zero. Figure 4.2 depicts the non-negative measure $H_{\alpha,\beta}$ associated to the excitation function $f_{\alpha,\beta}$. The left graph, the tail of the distribution $H_{\alpha,\beta}$ becomes fatter as the parameter α decreases. While when the parameter β decreases, the tail of the distribution $H_{\alpha,\beta}$ is getting thinner. The excitation function $f_{\alpha,\beta}$ is then a moment generating function of a positive random variable U with probability density function $H_{\alpha,\beta}$.
Equation (4.11) into equation (4.6) allows us to rewrite our intensity process as follows :

$$\begin{aligned}
\lambda_t|\mathscr{F}_{t_-} &= \lambda_0 + \eta \int_0^{t-} f_{\alpha,\beta}(t-u)\,\mathrm{d}L_u \\
&= \lambda_0 + \eta \int_0^{t-}\int_0^{\infty} e^{-v(t-u)} H_{\alpha,\beta}(v)\,\mathrm{d}v\mathrm{d}L_u \\
&= \lambda_0 + \eta \int_0^{\infty}\left(\int_0^{t-} e^{-v(t-u)}\mathrm{d}L_u\right) H_{\alpha,\beta}(v)\,\mathrm{d}v \\
&= \lambda_0 + \eta \int_0^{\infty} Y_t^{(v)} H_{\alpha,\beta}(v)\,\mathrm{d}v.
\end{aligned} \tag{4.13}$$

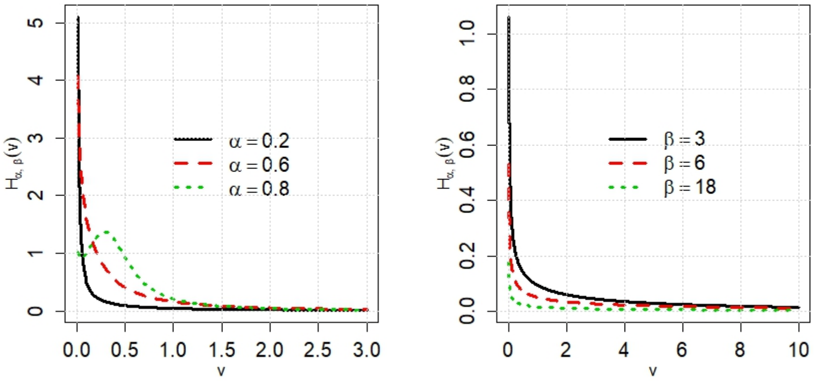


Figure 4.2: The non-negative measure $H_{\alpha,\beta}$. The left graph considers that $\beta = 0.5$ and α are equal to 0.2, 0.6 and 0.8 respectively. While the right graph considers that $\alpha = 0.5$ and β are equal to 3, 6 and 18 respectively. The tail of the distribution $H_{\alpha,\beta}$ becomes fatter as the parameter α decreases. While when the parameter β decreases, the support of the distribution $H_{\alpha,\beta}$ is expanded.

The third line is derived from the Fubini's theorem as the bivariate function $h : (v,u) \mapsto e^{-v(t-u)} H_{\alpha,\beta}(v)$ is measurable as a product of two measurable functions, and for each individual $w \in \Omega$,

$$\int_0^\infty \int_0^{t-} e^{-v(t-u)} H_{\alpha,\beta}(v)\,\mathrm{d}L_u(w)\,\mathrm{d}v < L_{t_-}(w) < \infty$$

since the function $u \mapsto e^{-v(t-u)}$ is bounded between 0 and 1 for every $u \in [0,t]$ and $H_{\alpha,\beta}$ is a probability measure on $\mathbb{R}_+$. In the fourth line, we denote by $Y_t^{(v)}$ a Markov process satisfying the following SDE:

$$\mathrm{d}Y_t^{(v)} = -vY_t^{(v)}\mathrm{d}t + \mathrm{d}L_t, \tag{4.14}$$

with the solution $Y_t^{(v)}|\mathscr{F}_{t_-} = \int_0^{t-} e^{-v(t-u)}\mathrm{d}L_u$. From equation (4.13), the intensity at time t is then a sum of the background rate λ_0 and the expectation of arrival rates $\left(Y_t^{(v)}\right)_{v \in \mathbb{R}^+}$ with respect of the measure $H_{\alpha,\beta}$. $\left(\lambda_t, \left(Y_t^{(v)}\right)_{v\in\mathbb{R}^+}, N_t\right)_{t\in\mathbb{R}^+}$ is then an infinite dimensional Markov process. This characterization allows us to use stochastic calculus in later developments in order to infer the moments of HP events.

4.2.3 Expected intensity

Using Fubini's theorem in equation (4.1) and (4.13), the expected intensity is

$$\mathbb{E}(\lambda_t|\mathscr{F}_0) = \lambda_0 + \eta\mathbb{E}(|J|)\int_0^t f_{\alpha,\beta}(t-u)\,\mathbb{E}(\lambda_u|\mathscr{F}_0)\,\mathrm{d}u, \tag{4.15}$$

or

$$\mathbb{E}(\lambda_t|\mathscr{F}_0) = \lambda_0 + \eta \int_0^\infty \mathbb{E}\left(Y_t^{(v)}|\mathscr{F}_0\right) H_{\alpha,\beta}(v)\,\mathrm{d}v. \tag{4.16}$$

Deriving the former equation (4.15) leads to

$$\begin{aligned}
\frac{\partial \mathbb{E}(\lambda_t|\mathscr{F}_0)}{\partial t} &= \eta\mathbb{E}(|J|)\mathbb{E}(\lambda_t|\mathscr{F}_0) + \eta\mathbb{E}(|J|)\int_0^t f^{(1)}_{\alpha,\beta}(t-u)\mathbb{E}(\lambda_u|\mathscr{F}_0)\,\mathrm{d}u \\
&= \eta\mathbb{E}(|J|)\mathbb{E}(\lambda_t|\mathscr{F}_0) \\
&\quad + \eta\mathbb{E}(|J|)\int_0^t g_{\alpha,\beta}(t-u) f_{\alpha,\beta}(t-u)\mathbb{E}(\lambda_u|\mathscr{F}_0)\,\mathrm{d}u,
\end{aligned} \tag{4.17}$$

where $f^{(1)}_{\alpha,\beta}(u) = \underbrace{-\frac{\beta u^{\alpha-2}}{\Gamma(\alpha-1)}}_{=g_{\alpha,\beta}(u)} f_{\alpha,\beta}(u)$ is the first order derivative of $f_{\alpha,\beta}$ derived using the Laplace transform properties. Note that in the case of the exponential kernel the function $g_{\alpha,\beta}$ is constant and allows one to derive an ordinary differential equation in $\mathbb{E}(\lambda_t|\mathscr{F}_0)$. Since $g_{\alpha,\beta}$ is not constant, we use the Laplace transform of equation (4.16) to derive the Laplace transform of expected intensity as shown in the following proposition.

Proposition 1. *The expected intensity conditionally to the information at time* 0 *is given by*

$$\mathbb{E}(\lambda_t|\mathscr{F}_0) = \lambda_0\left(1 + \eta\mathbb{E}(|J|)\int_0^t \sum_{r=0}^{\infty}(-\beta)^r u^{\alpha r} E^{r+1}_{1,\alpha r+1}(\eta\mathbb{E}(|J|)u)\,du\right),$$

where $0<\alpha<1$. $E^{r+1}_{1,\alpha r+1}$ *is the three parameters ML function*[1] *defined as follows :*

$$E^{\gamma}_{\alpha,\vartheta}(z) = \textstyle\sum_{n=0}^{\infty} \frac{(\gamma)_n}{n!\Gamma(\alpha n+\vartheta)} z^n, \quad \alpha>0,\ \vartheta>0,\ \gamma>0$$

with $(\gamma)_n = \gamma(\gamma+1)\cdots(\gamma+n-1)$.

Proof. The Laplace transform of equation (4.16) is

$$\begin{aligned}
\mathscr{L}\mathbb{E}(\lambda_t|\mathscr{F}_0)(s) &= \lambda_0\int_0^\infty e^{-us}\mathrm{d}u + \eta\mathbb{E}(|J|)F_{\alpha,\beta}(s)\mathscr{L}\mathbb{E}(\lambda_t|\mathscr{F}_0)(s) \\
&= \frac{\lambda_0}{s}\frac{1}{1-\eta\mathbb{E}(|J|)F_{\alpha,\beta}(s)} \\
&= \frac{\lambda_0}{s}\frac{s^\alpha+\beta}{s^\alpha+\beta-\eta\mathbb{E}(|J|)s^{\alpha-1}} \\
&= \frac{\lambda_0}{s}+\frac{\lambda_0}{s}\frac{\eta\mathbb{E}(|J|)s^{\alpha-1}}{s^\alpha+\beta-\eta\mathbb{E}(|J|)s^{\alpha-1}} \\
&= \frac{\lambda_0}{s}+\frac{\lambda_0}{s}\frac{\eta\mathbb{E}(|J|)}{s+\beta s^{1-\alpha}-\eta\mathbb{E}(|J|)}
\end{aligned} \tag{4.18}$$

[1]Note that for $\gamma=\vartheta=1$ we recover the classical ML function.

where in the first line we use the linearity of the Laplace transform and replace the Laplace transform of the convolution integral by the product of Laplace transforms. From Haubold et al. (2011) equation (17.6), we have:

$$\mathscr{L}^{-1}\left(\frac{\eta\mathbb{E}(|J|)}{s+\beta s^{1-\alpha}-\eta\mathbb{E}(|J|)}\right)(t) \quad = \quad \eta\mathbb{E}(|J|)\sum_{r=0}^{\infty}(-\beta)^r t^{\alpha r}E_{1,\alpha r+1}^{r+1}(\eta\mathbb{E}(|J|)t),$$

where $0<\alpha<1$ and $\left|\beta s^{1-\alpha}/(s-\eta\mathbb{E}(|J|))\right|<1$. Notice that this former inverse Laplace transform can be estimated numerically. By inverse Laplace transform of equation (4.18), the expected intensity is

$$\begin{aligned}\mathbb{E}(\lambda_t|\mathscr{F}_0) &= \lambda_0+\lambda_0\mathscr{L}^{-1}\left(\frac{1}{s}\frac{\eta\mathbb{E}(|J|)}{s+\beta s^{1-\alpha}-\eta\mathbb{E}(|J|)}\right)(t) \\ &= \lambda_0+\lambda_0\eta\mathbb{E}(|J|)\int_0^t\sum_{r=0}^{\infty}(-\beta)^r u^{\alpha r}E_{1,\alpha r+1}^{r+1}(\eta\mathbb{E}(|J|)u)\,\mathrm{d}u\end{aligned} \tag{4.19}$$

where $E_{1,\alpha r+1}^{r+1}$ is the three parameters ML function. □

Exploiting equation (4.19), it is tough to evaluate the long term expected intensity. Since $\frac{\eta\mathbb{E}(|J|)}{s+\beta s^{1-\alpha}-\eta\mathbb{E}(|J|)}$ has poles on the right side of s plane for all $0<\alpha\leq 1$, the existence of long term intensity is not guaranteed and we could not use the final value theorem[2] for Laplace transform to compute the long term expected intensity. In practice, for reasonable range of time $[0,T]$, the expected intensity $\mathbb{E}(\lambda_t|\mathscr{F}_0)$ is finite for $0\leq t\leq T$. This is the case in figure 4.3, where we draw the expected intensity $\mathbb{E}(\lambda_t|\mathscr{F}_0)$ obtained numerically by Laplace transform inversion for $\alpha\in(0,1]$, $\beta=0.5$, $\eta=0.25$, $\lambda_0=0.1$ and the time $t\geq 0$. For the other moments, we can approximate their values by using the conditional moment generating function of the intensity derived thanks to the Markov property of the process $\left(\lambda_t,\left(Y_t^{(v)}\right)_{v\in\mathbb{R}^+},N_t\right)_{t\in\mathbb{R}_+}$ in the next subsection. To the best of our knowledge the expression of the expected intensity in proposition 1 does not admit a more closed form expression and can be evaluated by a numerical approach. Subsequently, the expected number of jumps can be computed numerically using the following formula

$$\mathbb{E}(N_t|\mathscr{F}_0) \quad = \quad \int_0^t \mathbb{E}(\lambda_t|\mathscr{F}_0)\,\mathrm{d}u,$$

and as $Y_t^{(v)}=\int_0^{t-}e^{-v(t-u)}\mathrm{d}L_u$, its expected value is given by

$$\mathbb{E}\left(Y_t^{(v)}|\mathscr{F}_0\right) \quad = \quad \mathbb{E}(|J|)\int_0^t e^{-v(t-u)}\mathbb{E}(\lambda_u|\mathscr{F}_0)\,\mathrm{d}u. \tag{4.20}$$

[2] If f is continuously differentiable and both f and $f^{(1)}$ have a Laplace transform; $f^{(1)}$ is absolutely integrable and final value $f(\infty)$ exists then $\lim_{u\to\infty}f(u)=\lim_{s\to 0}s\mathscr{L}f(u)(s)$. Notice that $f(\infty)$ exists if $s\mathscr{L}(f(t))(s)$ has poles only on the left side of s plane. In other words, not on origin, the jw axis and the right side of s plane.

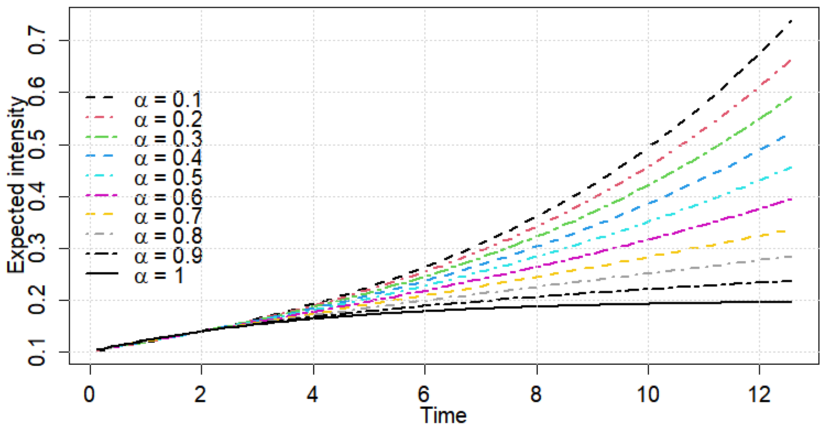


Figure 4.3: Expected intensity $\mathbb{E}(\lambda_t|\mathscr{F}_0)$ obtained numerically by Laplace transform inversion for $\alpha \in (0,1]$, $\beta = 0.5$, $\eta = 0.25$, $\lambda_0 = 0.1$ and the time $t \geq 0$.

4.2.4 Conditional moment generating function of the intensity and its integral

In this subsection, we derive the conditional mgf of the intensity and its integral, through an approximation of the HP with ML kernel $\left(\lambda_t, \left(Y_t^{(v)}\right)_{v\in\mathbb{R}^+}, N_t\right)_{t\in\mathbb{R}_+}$ obtained by means of a discretisation scheme. This discretisation scheme enables us to switch from a problem in infinite dimension to a problem in finite dimension for which Itô calculus is applicable. Let us recall that $\left(\lambda_t, \left(Y_t^{(v)}\right)_{v\in\mathbb{R}^+}, N_t\right)_{t\in\mathbb{R}_+}$ is an infinite dimensional Markov process when we rewrite the intensity as follows:

$$\lambda_t = \lambda_0 + \eta \int_0^\infty Y_t^{(v)} H_{\alpha,\beta}(v)\,\mathrm{d}v, \tag{4.21}$$

with $\mathrm{d}Y_t^{(v)} = -vY_t^{(v)}\mathrm{d}t + \mathrm{d}L_t$. Approximate the measure H as a discrete measures with a finite numbers of atoms allows to move to a framework with finite dimensional Markov process. For this purpose, we consider a partition $\mathscr{E}^{(n)} := \{0 < \xi_0^{(n)} < \xi_1^{(n)} <$

$... < \xi_n^{(n)} < \infty\}$. For $i = 0,...,n-1$, on each interval $(\xi_i^{(n)}, \xi_{i+1}^{(n)})$, the mass of any atom is

$$\begin{aligned} m_{i+1}^{(n)} &= \int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} H_{\alpha,\beta}(z)\,\mathrm{d}z, \\ &= \frac{1}{\alpha\pi}\arctan\left(\frac{\left(\xi_{i+1}^{(n)}\right)^{\alpha} + \beta\cos(\alpha\pi)}{\beta\sin(\alpha\pi)}\right) \\ &\quad - \frac{1}{\alpha\pi}\arctan\left(\frac{\left(\xi_{i}^{(n)}\right)^{\alpha} + \beta\cos(\alpha\pi)}{\beta\sin(\alpha\pi)}\right) \end{aligned} \tag{4.22}$$

and the barycenter with respect of the measure $H_{\alpha,\beta}(.)$ is defined as

$$b_{i+1}^{(n)} = \frac{\int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} zH_{\alpha,\beta}(z)\,\mathrm{d}z}{\int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} H_{\alpha,\beta}(z)\,\mathrm{d}z},$$

where the numerator does not admit a closed form expression but is approximated by :

$$\begin{aligned} \int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} zH_{\alpha,\beta}(z)\,\mathrm{d}z \approx & \left(\frac{\xi_{i+1}^{(n)} + \xi_i^{(n)}}{2\alpha\pi}\right)\left[\frac{1}{\alpha\pi}\arctan\left(\frac{\left(\xi_{i+1}^{(n)}\right)^{\alpha} + \beta\cos(\alpha\pi)}{\beta\sin(\alpha\pi)}\right)\right. \\ & \left. - \frac{1}{\alpha\pi}\arctan\left(\frac{\left(\xi_{i}^{(n)}\right)^{\alpha} + \beta\cos(\alpha\pi)}{\beta\sin(\alpha\pi)}\right)\right]. \end{aligned} \tag{4.23}$$

The result in equation (4.22) comes from the fact that

$$\begin{aligned} \frac{\beta\sin(\alpha\pi)}{\pi}\int \frac{z^{\alpha-1}}{z^{2\alpha} + 2\beta z^{\alpha}\cos(\alpha\pi) + \beta^2}\,\mathrm{d}z = & \frac{1}{\alpha\pi} \\ & \times\arctan\left(\frac{z^{\alpha} + \beta\cos(\alpha\pi)}{\beta\sin(\alpha\pi)}\right), \end{aligned} \tag{4.24}$$

while equation (4.23) is obtained with an integration by parts in which the integral is approximated with the trapezoidal method. Indeed, if we denote by $h_{\alpha,\beta}$ the primitive of $H_{\alpha,\beta}$, then by equation (4.24) $h_{\alpha,\beta}$ is given by

$$h_{\alpha,\beta}(z) = \frac{1}{\alpha\pi}\arctan\left(\frac{z^{\alpha} + \beta\cos(\alpha\pi)}{\beta\sin(\alpha\pi)}\right) - \frac{1}{\alpha\pi}\arctan(\cot(\alpha\pi)). \tag{4.25}$$

Since $H_{\alpha,\beta}(z) = \frac{dh_{\alpha,\beta}(z)}{dz}$, the integration by parts and the approximation of an integral by the trapezoidal method lead to

$$\begin{aligned}
\int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} zH_{\alpha,\beta}(z)\,dz &= \left[zh_{\alpha,\beta}(z)\right]_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} - \int_{\xi_i^{(n)}}^{\xi_{i+1}^{(n)}} h_{\alpha,\beta}(z)\,dz \\
&\approx \xi_{i+1}^{(n)} h_{\alpha,\beta}\left(\xi_{i+1}^{(n)}\right) - \xi_i^{(n)} h_{\alpha,\beta}\left(\xi_i^{(n)}\right) \\
&\quad - \frac{\xi_{i+1}^{(n)} - \xi_i^{(n)}}{2}\left(h_{\alpha,\beta}\left(\xi_{i+1}^{(n)}\right) + h_{\alpha,\beta}\left(\xi_i^{(n)}\right)\right) \\
&= \frac{\xi_{i+1}^{(n)} + \xi_i^{(n)}}{2}\left(h_{\alpha,\beta}\left(\xi_{i+1}^{(n)}\right) - h_{\alpha,\beta}\left(\xi_i^{(n)}\right)\right).
\end{aligned} \tag{4.26}$$

By construction, $\lim_{n\to\infty}\sum_{i=1}^{n} m_i^{(n)} = 1$ and the discrete measure for a partition of size n is

$$H_{\alpha,\beta}^{(n)}(z) = \sum_{i=1}^{n} m_i^{(n)} \delta_{b_i^{(n)}}(z) \tag{4.27}$$

where $\delta_{b_i^{(n)}}(z)$ is the Dirac measure. We consider that the following assumptions hold for the partition $\mathscr{E}^{(n)}$:

(i) $\xi_0^{(n)} \to 0$ and $\xi_n^{(n)} \to \infty$ when $n \to \infty$,

(ii) $\max|\xi_{i+1}^{(n)} - \xi_i^{(n)}| \to 0$ when $n \to \infty$,

(iii) $\mathscr{E}^{(n)} \subset \mathscr{E}^{(n+1)}$.

In practice, we take the $\mathscr{E}^{(n)}$ partition made up of points $\xi_i^{(n)} = q_{\frac{i+1}{n+2}}$, for $i \in \{0,...,n\}$, where the quantity q_γ is defined as follows

$$q_\gamma = \sup\left\{v \in \mathbb{R}_+ \cup \{\infty\} \mid h_{\alpha,\beta}(v) \leq \gamma\right\}, \tag{4.28}$$

for all probability $\gamma \in [0,1]$. In particular, $q_0 = 0$ and $q_1 = \infty$. For instance, the points $\xi_i^{(n)} = q_{\frac{i+1}{n+2}}$, for $i \in \{0,...,n\}$ constitute a partition $\mathscr{E}^{(n)}$. Indeed, the first assumption is satisfied as when $n \to \infty$, $\mathscr{E}_0^{(n)} = q_{\frac{1}{n+2}} \to q_0 = 0$ and $\mathscr{E}_n^{(n)} = q_{\frac{n+1}{n+2}} \to q_1 = \infty$. The second assumption $\max\left|\xi_{i+1}^{(n)} - \xi_i^{(n)}\right| \to 0$ when $n \to \infty$ holds since for each $i \in \{0,...,n\}$ the distance between their corresponding probabilities equals $\frac{1}{n+2}$ and vanish when n tends to ∞. The last assumption $\mathscr{E}^{(n)} \subset \mathscr{E}^{(n+1)}$ is derived from

$$\mathscr{E}_0^{(n)} = q_{\frac{1}{n+2}} > q_{\frac{1}{n+3}} = \mathscr{E}_0^{(n+1)},$$

and

$$\mathscr{E}_n^{(n)} = q_{\frac{n+1}{n+2}} < q_{\frac{n+2}{n+3}} = \mathscr{E}_{n+1}^{(n+1)}.$$

Figure 4.4 shows that the discrete approximation $dH_{\alpha,\beta}^{(n)}$ of the cdf $dH_{\alpha,\beta}$ becomes ac-

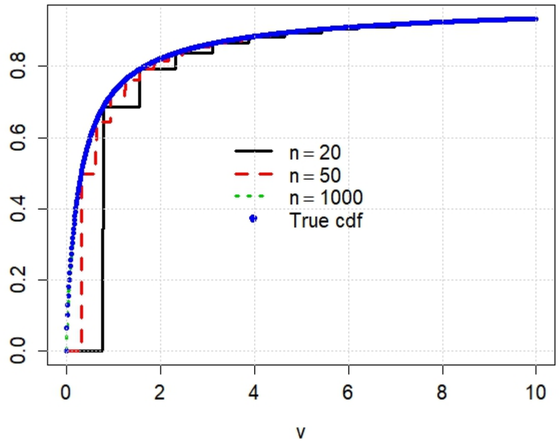


Figure 4.4: The true cdf $\mathrm{d}H_{\alpha,\beta}$ versus the discrete cdf $\mathrm{d}H^{(n)}_{\alpha,\beta}$ when $\alpha = 0.6$, $\beta = 0.5$ and $n \in \{20, 50, 1000\}$.

curate as the partition's size increases. Considering the previous discretisation setting, the next proposition connects the discretised measure to the true measure for any measurable function in the set $L^1\left(H_{\alpha,\beta}\right)$ of Lebesgue integrable functions with respect to the measure $H_{\alpha,\beta}$.

Proposition 2. *For $n \in \mathbb{N}$, if a partition $\mathscr{E}^{(n)}$ satisfies (i)–(iii) above and a function $f : \mathbb{R} \to \mathbb{C}$ is in $L^1\left(H_{\alpha,\beta}\right)$, then,*

$$\lim_{n\to\infty} \int_0^\infty f(z) H^{(n)}_{\alpha,\beta}(z)\,dz = \int_0^\infty f(z)\, H_{\alpha,\beta}(z)\,dz. \tag{4.29}$$

Proof. For $n \in \mathbb{N}$, using (4.27) and given the assumptions (i)–(iii) we obtain

$$\begin{aligned}
\int_0^\infty f(z) H^{(n)}_{\alpha,\beta}(z)\,\mathrm{d}z = \sum_{v=1}^{n} m^{(n)}_v f\left(\delta_{b^{(n)}_v}(z)\right) &= \sum_{v=1}^{n} m^{(n)}_v f\left(b^{(n)}_v\right) \\
&= \sum_{v=1}^{n} \int_{\xi^{(n)}_{v-1}}^{\xi^{(n)}_v} H_{\alpha,\beta}(z)\,\mathrm{d}z\, f\left(b^{(n)}_v\right) \\
&= \sum_{v=1}^{n} \int_{\xi^{(n)}_{v-1}}^{\xi^{(n)}_{v1}} H_{\alpha,\beta}(z)\, f\left(b^{(n)}_v\right)\mathrm{d}z \\
&= \int_{\xi^{(n)}_0}^{\xi^{(n)}_n} H_{\alpha,\beta}(z)\, f\left(t^{(n)}(z)\right)\mathrm{d}z
\end{aligned}$$

where

$$t^{(n)}(z) \quad = \quad \sum_{v=1}^{n} b^{(n)}_v 1_{\left[\xi^{(n)}_{v-1},\xi^{(n)}_v\right]}(z).$$

Given that the function f is continuous and $\lim_{n\to\infty} t^{(n)}(z) = z$, passing to the limit we obtain the expected result. □

We choose the size n of the partition. From equation (4.21) the intensity is given by

$$\begin{aligned}\lambda_t^{(n)} &:= \lambda_0 + \eta \int_0^\infty Y_t^{(v)} H^n_{\alpha,\beta}(z)\,\mathrm{d}z \\ &= \lambda_0 + \eta \sum_{v=1}^{n} m_v^{(n)} Y_t^{(b_v^{(n)})}\end{aligned} \tag{4.30}$$

with $N_t^{(n)}$ a point process of intensity $\lambda_t^{(n)}$ and

$$\mathrm{d}Y_t^{(b_v^{(n)})} = -b_v^{(n)} Y_t^{(b_v^{(n)})}\mathrm{d}t + \mathrm{d}L_t^{(n)} \qquad v = 1,...,n$$

with $L_t^{(n)} = \sum_{i=1}^{N_t^{(n)}} |J_i|$. Notice that

$$\begin{aligned}\mathrm{d}\lambda_t^{(n)} &= \eta \sum_{v=1}^{n} m_v^{(n)} \mathrm{d}Y_t^{(b_v^{(n)})} \\ &= -\eta \sum_{v=1}^{n} m_v^{(n)} b_v^{(n)} Y_t^{(b_v^{(n)})}\mathrm{d}t + \eta \sum_{v=1}^{n} m_v^{(n)} \mathrm{d}L_t^{(n)}\end{aligned} \tag{4.31}$$

Proposition 3. *The moment generating function of $\lambda_s^{(n)}$ conditionally to the information at time $t \le s$ is given by*

$$\begin{aligned}\mathbb{E}\left(e^{\omega_1 \lambda_s^{(n)} + \omega_2 \int_t^s \lambda_u^{(n)} du} | \mathscr{F}_t\right) &= \exp\left(B(t,s,\bar{\omega})\,\lambda_t^{(n)}\right) \\ &\quad \times \exp\left(\sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\bar{\omega}\right) m_v^{(n)} Y_t^{(b_v^{(n)})}\right)\end{aligned}$$

for all $\bar{\omega} = (\omega_1,\omega_2) \in \mathbb{C}_-^2$ that satisfies the regularity condition

$$\psi\left(0, \eta\omega_1 \sum_{v=1}^{n} m_v^{(n)}\right) < \infty$$

. The functions C are defined as follows

$$C\left(t,s,b_v^{(n)},\bar{\omega}\right) \quad = \quad -b_v^{(n)} \eta \int_t^s e^{-b_v^{(n)}(u-t)} B(u,s,\bar{\omega})\,du, \forall v = 1,\cdots,n$$

where B is a time dependent function that solves the following ODE:

$$\frac{\partial B}{\partial t}(t,s,\bar{\omega}) = 1 - \omega_2 - \psi\left(0, \eta B(t,s,\bar{\omega}) \sum_{v=1}^{n} m_v^{(n)} + \sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\bar{\omega}\right) m_v^{(n)}\right) \tag{4.32}$$

with final condition $B(s,s,\bar{\omega}) = \omega_1$. ψ is the moment generating function of the pair of the random jump sizes J and its absolute value : $\psi(z_1,z_2) := \mathbb{E}\left(e^{z_1 J + z_2 |J|}\right)$.

Proof. For fixed time s and $\bar{\omega} = (\omega_1, \omega_2) \in \mathbb{C}_-^2$ with $\mathbb{C}_-$ the set of complex numbers with non-positive real part, we define the conditional moment generating function by

$$f \equiv f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{\left(b_v^{(n)}\right)}\right)_{v=1,\ldots,n}, N_t^{(n)}\right) := \mathbb{E}\left(e^{\omega_1 \lambda_s^{(n)} + \omega_2 \int_t^s \lambda_u^{(n)} \mathrm{d}u} | \mathscr{F}_t\right),$$

where f is a complex-valued function on $[0,s] \times G$ with $G = \mathbb{R}_+ \times \mathbb{R}_+^n \times \mathbb{N}$ the state space of the Markov process $\left(\lambda_t^{(n)}, \left(Y_t^{\left(b_v^{(n)}\right)}\right)_{v=1,\ldots,n}, N_t^{(n)}\right)_{t \geq 0}$. By the tower property, for any $l \geq t$,

$$f = \mathbb{E}\left(e^{\omega_2 \int_t^l \lambda_u^{(n)} \mathrm{d}u} f\left(l, s, \lambda_l^{(n)}, \left(Y_l^{\left(b_v^{(n)}\right)}\right)_{v=1,\ldots,n}, N_l^{(n)}\right) | \mathscr{F}_t\right),$$

and

$$\lim_{l \to t} \frac{\mathbb{E}\left(e^{\omega_2 \int_t^l \lambda_u^{(n)} \mathrm{d}u} f\left(l, s, \lambda_l^{(n)}, \left(Y_l^{\left(b_v^{(n)}\right)}\right)_{v=1,\ldots,n}, N_l^{(n)}\right) | \mathscr{F}_t\right) - f}{l - t} = 0.$$

The following Taylor expansion with a Lagrange remainder

$$e^{\omega_2 \int_t^l \lambda_u^{(n)} \mathrm{d}u} = 1 + \omega_2 \lambda_t^{(n)} (l - t) + R_1(l, t)$$

allows us to obtain that the infinitesimal generator $\mathscr{A}$ of the Markov process $\left(\lambda_t^{(n)}, \left(Y_t^{\left(b_v^{(n)}\right)}\right)_{v=1,\ldots,n}, N_t^{(n)}\right)$ evaluated on f, defined by

$$\lim_{l \to t} \frac{\mathbb{E}\left(f\left(l, s, \lambda_l^{(n)}, \left(Y_l^{\left(b_v^{(n)}\right)}\right)_{v=1,\ldots,n}, N_l^{(n)}\right) | \mathscr{F}_t\right) - f}{l - t} \tag{4.33}$$

is equal to $-\omega_2 \lambda_t^{(n)} f$. Since $\mathscr{A} f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{\left(b_v^{(n)}\right)}\right)_{v=1,\ldots,n}, N_t^{(n)}\right)$ exists the function f is in the domain of $\mathscr{A}$ which contains the set of twice differentiable functions such that all partial derivatives of up to order two are continuous function (see e.g. Duffie et al. 2003). Using the Itô's lemma for semi-martingales, we infer that f satisfies the following differential equation :

$$\begin{aligned} -\omega_2 \lambda_t^{(n)} f &= \frac{\partial f}{\partial t} - \sum_{v=1}^{n} \frac{\partial f}{\partial Y^{(v)}} b_v^{(n)} Y_t^{\left(b_v^{(n)}\right)} - \eta \frac{\partial f}{\partial \lambda^{(n)}} \sum_{v=1}^{n} m_v^{(n)} b_v^{(n)} Y_t^{\left(b_v^{(n)}\right)} \\ &\quad + \lambda_t^{(n)} \int_{-\infty}^{\infty} \Delta f(z) \, \nu(\mathrm{d}z), \end{aligned} \tag{4.34}$$

where $\Delta f(z)$ is the variation of f when a jump of size z occurs and the time just before. Notice that the right hand of equation (4.17) corresponds to the infinitesimal generator $\mathscr{A}$ of the Markov process $\left(\lambda_t^{(n)}, \left(Y_t^{\left(b_v^{(n)}\right)}\right)_{v=1,\dots,n}, N_t^{(n)}\right)$ evaluated on f obtain using the Itô's lemma for semi-martingales.

Since $\left(\lambda_t^{(n)}, \left(Y_t^{\left(b_v^{(n)}\right)}\right)_{v=1,\dots,n}, N_t^{(n)}\right)$ is an affine jump process, from Duffie D. (2000) and Errais et al. (2010), under technical regularity condition we conjecture that f is an exponential affine function of $\lambda_t^{(n)}$ and $\left(Y_t^{\left(b_v^{(n)}\right)}\right)_{v=1,\dots,n}$:

$$f \quad := \quad \exp\left(A(t,s,\bar{\omega}) + B(t,s,\bar{\omega})\lambda_t^{(n)} + \sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\bar{\omega}\right) m_v^{(n)} Y_t^{\left(b_v^{(n)}\right)}\right)$$

where A and B are time dependent functions such that $A(s,s,\bar{\omega}) = 0$ and $B(s,s,\bar{\omega}) = \omega$. While C is a function that depends on the time and the position $v = 1,\cdots,n$ within the partition $\mathscr{E}^{(n)}$. Its final value $C\left(s,s,b_v^{(n)},\bar{\omega}\right)$ is equal to zero. From (Hubalek et al. 2016), the technical regularity condition needed is equivalent to the integrability of the jumps, that is

$$\left.\frac{\partial B}{\partial t}(t,s,\bar{\omega})\right|_{t=s} = \mathbb{E}\left(e^{\eta\omega_1\sum_{v=1}^{n} m_v^{(n)}|J|}\right) < \infty,$$

where the random jump size J has density $\nu(z)$ on $\mathbb{R}$. The partial derivatives of f are then given by :

$$\frac{\partial f}{\partial t} = \left(\frac{\partial A}{\partial t}(t,s,\bar{\omega}) + \frac{\partial B}{\partial t}(t,s,\bar{\omega})\lambda_t^{(n)} + \sum_{v=1}^{n}\frac{\partial C}{\partial t}\left(t,s,b_v^{(n)},\bar{\omega}\right) m_v^{(n)} Y_t^{\left(b_v^{(n)}\right)}\right) f,$$

$$\frac{\partial f}{\partial \lambda^{(n)}} = B(t,s,\bar{\omega}) f, \quad \frac{\partial f}{\partial Y^{\left(b_v^{(n)}\right)}} = C\left(t,s,b_v^{(n)},\bar{\omega}\right) m_v^{(n)} f \quad \forall v = 1,\cdots,n,$$

and the variation of f when a jump of size z occurs and the time just before is equal to

$$\left(\exp\left(|z|\left(\eta B(t,s,\bar{\omega})\sum_{v=1}^{n} m_v^{(n)} + \sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\bar{\omega}\right) m_v^{(n)}\right)\right) - 1\right) f.$$

Injecting all these expressions into equation (4.17) leads to :

$$\begin{aligned} -\omega_2\lambda_t^{(n)} \quad = \quad & \frac{\partial A}{\partial t}(t,s,\bar{\omega}) + \frac{\partial B}{\partial t}(t,s,\bar{\omega})\lambda_t^{(n)} \\ & + \sum_{v=1}^{n}\frac{\partial C}{\partial t}\left(t,s,b_v^{(n)},\bar{\omega}\right) m_v^{(n)} Y_t^{(v)} \\ & - \sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\bar{\omega}\right) m_v^{(n)} b_v^{(n)} Y_t^{\left(b_v^{(n)}\right)} \\ & - \eta B(t,s,\bar{\omega})\sum_{v=1}^{n} m_v^{(n)} b_v^{(n)} Y_t^{(v)} + + \lambda_t^{(n)} \\ & \times\left(\psi\left(0, \eta B(t,s,\bar{\omega})\sum_{v=1}^{n} m_v^{(n)} + \sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\bar{\omega}\right) m_v^{(n)}\right) - 1\right), \end{aligned} \tag{4.35}$$

where $\psi(z_1,z_2)=\mathbb{E}\left(e^{z_1 J+z_2|J|}\right)$. From which we find the following ODE system:

$$\begin{cases} 0=\frac{\partial A}{\partial t}(t,s,\bar{\omega}) \\ -\omega_2=\frac{\partial B}{\partial t}(t,s,\bar{\omega})+\psi\left(0,\eta B(t,s,\bar{\omega})\sum_{v=1}^{n} m_v^{(n)}+\sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\bar{\omega}\right)m_v^{(n)}\right)-1 \\ 0=\frac{\partial C}{\partial t}\left(t,s,b_v^{(n)},\bar{\omega}\right)-C\left(t,s,b_v^{(n)},\bar{\omega}\right)b_v^{(n)}-\eta B(t,s,\bar{\omega})b_v^{(n)},\ \forall v=1,\cdots,n. \end{cases}$$

As $A(s,s,\bar{\omega})=0$, the solution of the first ODE is $A(t,s,\bar{\omega})=0$. The last ODE admits the following solution:

$$C\left(t,s,b_v^{(n)},\bar{\omega}\right) = -b_v^{(n)}\eta\int_t^s e^{-b_v^{(n)}(u-t)}B(u,s,\bar{\omega})\,\mathrm{d}u.$$

□

In the next proposition, passing to the limit we derive the conditional moment generating function of the HP intensity and its integral.

Proposition 4. *For $n\in\mathbb{N}$, if a partition $\mathscr{E}^{(n)}$ satisfies assumptions (i)–(iii), for $t\geq 0$,*
a) $\left(\lambda_t^{(n)},\left(Y_t^{\left(b_v^{(n)}\right)}\right)_{v=1,\ldots,n},N_t^{(n)}\right)$ converges to $\left(\lambda_t,\left(Y_t^{(v)}\right)_{v\in\mathbb{R}_+},N_t\right)$ almost surely when n tends to ∞.
b) The moment generating function of $(\lambda_s,\int_t^s\lambda_u du)$ conditionally to the information at time $t\leq s$ is given by

$$\begin{aligned}\mathbb{E}\left(e^{\omega_1\lambda_s+\omega_2\int_t^s\lambda_u du}|\mathscr{F}_t\right) &= \exp\left(B(t,s,\bar{\omega})\lambda_t\right)\exp\Bigg(-\eta \\ &\quad\times\int_0^\infty\int_t^s ve^{-v(u-t)}B(u,s,\bar{\omega})Y_t^{(v)}H_{\alpha,\beta}(v)\,dudv\Bigg),\end{aligned} \tag{4.36}$$

for all $\bar{\omega}=(\omega_1,\omega_2)\in\mathbb{C}_-^2$ that satisfies the regularity condition $\psi(0,\eta\omega_1)<\infty$. The functionB is a time dependent function that solves the following integro-differential equation:

$$\begin{aligned}\frac{\partial B}{\partial t}(t,s,\bar{\omega}) &= 1-\psi\left(0,\eta B(t,s,\bar{\omega})+\eta\int_t^s B(u,s,\bar{\omega})f_{\alpha,\beta}^{(1)}(u-t)\,du\right) \\ &\quad -\omega_2\end{aligned} \tag{4.37}$$

with final condition $B(s,s,\bar{\omega})=\omega_1$.

Proof. a) For fixed $w\in\Omega$ and fixed $t\geq 0$,

$$\begin{aligned}\lim_{n\to\infty}\lambda_t^{(n)}(w) &= \lim_{n\to\infty}\left(\lambda_0+\eta\int_0^\infty Y_t^{(z)}(w)H_{\alpha,\beta}^{(n)}(z)\,\mathrm{d}z\right) \\ &= \lambda_0+\eta\lim_{n\to\infty}\int_0^\infty Y_t^{(z)}(w)H_{\alpha,\beta}^{(n)}(z)\,\mathrm{d}z,\end{aligned}$$

where $Y_t^{(z)} = \int_0^t e^{-z(t-u)}\mathrm{d}L_u^{(n)}$. The function $z \mapsto Y_t^{(z)}(w)$ is integrable with respect to the measure $H_{\alpha,\beta}$ as a sum of non-increasing exponential functions. Thus by proposition 2 we have

$$\lim_{n\to\infty}\int_0^\infty Y_t^{(z)}(w)\,H_{\alpha,\beta}^{(n)}(z)\,\mathrm{d}z = \int_0^\infty Y_t^{(z)}(w)\,H_{\alpha,\beta}(z)\,\mathrm{d}z,$$

which allows us to conclude that $\lambda_t^{(n)}$ converges to λ_t almost surely when n tends to ∞. It follows that $N_t^{(n)} = \int_0^t \lambda_u^{(n)}\mathrm{d}u$ converges almost surely to N_t. For $v \in \{1,...,n\}$, $Y_t^{(b_v^{(n)})} = \int_0^t e^{-b_v^{(n)}(t-u)}\mathrm{d}L_u^{(n)}$. If the partition $\mathscr{E}^{(n)}$ satisfies the assumptions (i)-(iii), when n tends to ∞, a barycenter $b_v^{(n)}$ tends to be an element of $\mathbb{R}_+$. Hence the sequence $\left(Y_t^{\left(b_v^{(n)}\right)}\right)_{v=1,...,n}$ tends to $\left(Y_t^{(v)}\right)_{v\in\mathbb{R}_+}$ almost surely.

b) The dominated convergence provides that

$$\mathbb{E}\left(e^{\omega_1\lambda_s+\omega_2\int_t^s\lambda_u\mathrm{d}u}|\mathscr{F}_t\right) = \lim_{n\to\infty}\mathbb{E}\left(e^{\omega_1\lambda_s^{(n)}+\omega_2\int_t^s\lambda_u^{(n)}\mathrm{d}u}|\mathscr{F}_t\right),$$

where from the proposition 3

$$\begin{aligned}\mathbb{E}\left(e^{\omega_1\lambda_s^{(n)}+\omega_2\int_t^s\lambda_u^{(n)}\mathrm{d}u}|\mathscr{F}_t\right) &= \exp\Bigg(B(t,s,\bar\omega)\,\lambda_t^{(n)} \\ &\quad + \sum_{v=1}^n C\left(t,s,b_v^{(n)},\bar\omega\right)m_v^{(n)}Y_t^{(b_v^{(n)})}\Bigg)\end{aligned}$$

and

$$\sum_{v=1}^n C\left(t,s,b_v^{(n)},\bar\omega\right)m_v^{(n)} = -\sum_{v=1}^n b_v^{(n)}\eta\int_t^s e^{-b_v^{(n)}(u-t)}B(u,s,\bar\omega)\,\mathrm{d}u\,m_v^{(n)}.$$

Taking into account the definition of the discretised measure $H_{\alpha,\beta}^{(n)}$ (4.27) we obtain

$$\int_0^\infty C(t,s,v,\bar\omega)\,H_{\alpha,\beta}^{(n)}(v)\,\mathrm{d}v = -\eta\int_0^\infty\left(\int_t^s ve^{-v(u-t)}B(u,s,\bar\omega)\,\mathrm{d}u\right)H_{\alpha,\beta}^{(n)}(v)\,\mathrm{d}v.$$

Similarly to point i), passing to the limit we have

$$\begin{aligned}\int_0^\infty C(t,s,v,\bar\omega)\,H_{\alpha,\beta}(v)\,\mathrm{d}v &= -\eta\int_0^\infty\int_t^s ve^{-v(u-t)}B(u,s,\bar\omega)\,H_{\alpha,\beta}(v)\,\mathrm{d}u\mathrm{d}v \\ &= -\eta\int_t^s B(u,s,\bar\omega)\int_0^\infty e^{-v(u-t)}vH_{\alpha,\beta}(v)\,\mathrm{d}v\mathrm{d}u \\ &= \eta\int_t^s B(u,s,\bar\omega)\,f_{\alpha,\beta}^{(1)}(u-t)\,\mathrm{d}u.\end{aligned}$$

The second and last lines of the latter equation stem from the Fubini's theorem and the frequency differentiation property of the Laplace transform. Likewise, we show that the limit of $\sum_{v=1}^n C\left(t,s,b_v^{(n)},\bar\omega\right)m_v^{(n)}Y_t^{(b_v^{(n)})}$ is

$$\int_0^\infty C(t,s,v,\bar\omega)\,Y_t^{(v)}H_{\alpha,\beta}(v)\,\mathrm{d}v = -\eta\int_0^\infty\int_t^s ve^{-v(u-t)}B(u,s,\bar\omega)\,Y_t^{(v)}H_{\alpha,\beta}(v)\,\mathrm{d}u\mathrm{d}v,$$

and concludes the proof. □

4.3 Option pricing model

4.3.1 Models specification

The asset price is a process $(S_t)_{t\geq 0}$ defined on the probability space Ω, endowed with its natural filtration $(\mathscr{F}_t)$ and the probability measure $\mathbb{P}$. $\mathbb{P}$ is the real historical measure. For the sake of simplicity, we assume the short rate interest constant and equal to r. The log-return of S_t noted X_t, is such that

$$S_t = S_0 \exp(X_t). \tag{4.38}$$

We assume that X_t is driven by the following jump-diffusion :

$$dX_t = \left(\mu - \frac{\sigma^2}{2} - \mathbb{E}\left(e^J - 1\right)\lambda_t\right)dt + \sigma dW_t + d\sum_{i=1}^{N_t} J_i, \tag{4.39}$$

where μ, σ are respectively the expected return and the instantaneous volatility part of the log-return explained by the Brownian motion W_t. $\sum_{i=1}^{N_t} J_i$, is the sum of the N_t jumps size up to time t. The jump sizes $\{J_i, i \geq 1\}$ are assumed to be independent and identically distributed Double exponential random variables with density $\nu(z)$ on $\mathbb{R}$:

$$\nu(z) = p\rho^+ \exp\left(-\rho^+ z\right) 1_{\{z\geq 0\}} - (1-p)\rho^- \exp\left(-\rho^- z\right) 1_{\{z<0\}}, \tag{4.40}$$

with $p \in [0,1]$, $\rho^+ > 0$ and $\rho^- < 0$. The term $\mathbb{E}\left(e^J - 1\right)\lambda_t$ in equation (4.39) comes from the compensator $\int_0^t \mathbb{E}\left(e^J - 1\right)\lambda_t dt$ of the jump process $\sum_{i=1}^{N_t}\left(e^{J_i} - 1\right)$. Its presence ensures that on average, the log-return is equal to μ.
In the sequel, we denote by $X_t^{(n)}$ the approximated log-return process according to the partition $\mathscr{E}^{(n)}$:

$$dX_t^{(n)} = \left(\mu - \frac{\sigma^2}{2} - \mathbb{E}\left(e^J - 1\right)\lambda_t^{(n)}\right)dt + \sigma dW_t + d\sum_{i=1}^{N_t^{(n)}} J_i. \tag{4.41}$$

The process $X_t^{(n)}$ is then a Markov process with respect to the intensity $\lambda_t^{(n)}$, the n factors $\left(Y_t^{\left(b_\nu^{(n)}\right)}\right)_{\nu=1,\dots,n}$ and the counter $N_t^{(n)}$. We can therefore deal with pricing and hedging problems using standard methods developed for stochastics models. Proposition 4 a) allows one to establish the convergence almost surely of $X_t^{(n)}$ to X_t.

4.3.2 Data description and calibration

This subsection describes the data set used in the numerical applications and the procedure for estimating the log-return model parameters in equation (4.41). The sample of data consists of S&P 500 daily values from 05$^{\text{th}}$ May 2010 to 30$^{\text{th}}$June 2021 (2811 observations). Table 4.1 provides summary statistics of daily log-returns. A kurtosis higher than 3 indicates that the distribution of log-returns has fat-tails. Whereas

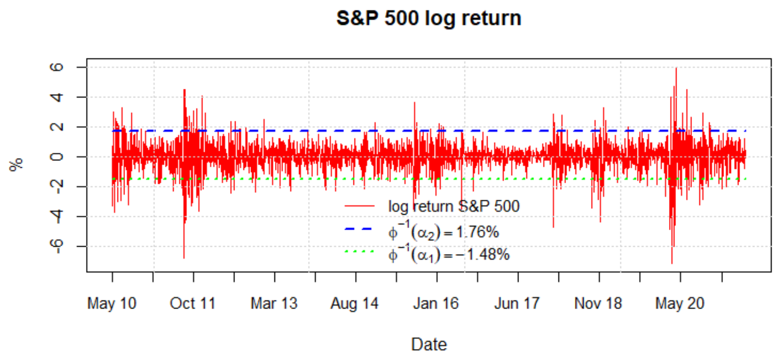


Figure 4.5: S&P 500 log-return from 02/06/2011 to 14/06/2021. There are mainly negative jumps on the S&P 500 prices over our observation period.

a negative skew indicates that returns are left-tailed. The Jarque Bera test rejects the assumption of normality.

	Value
Mean	0.05%
Standard deviation	0.96%
Skewness	-0.65
Kurtosis	9.06
Jarque Bera p-value	$< 2.2e-16$

Table 4.1: This table reports the mean, the standard deviation, the skewness, the kurtosis and the Jarque Bera p-value of the normality test on daily log-returns of the S&P 500 from 05$^{\text{th}}$ May 2010 to 30$^{\text{th}}$June 2021.

Figure 4.5 shows that the daily log-return randomly fluctuates around zero and displays volatility clustering with a serial dependence. The first cluster is observable from September 2011 to February 2012 coincides with the period of recession in euro area due to the European sovereign debt crisis. A second cluster is identified between June 2015 and June 2016 due to the fall in petroleum prices, the Greek debt default and the Brexit. Around March 2020, we observe a third cluster explained by the lockdown in US, UK and Europe to stop the covid-19 spreading. Shocks during these periods do not display any clear trend: negative abrupt movements follow by abrupt movements of either sign. These observations justify to add a jump component in the log-return model and to link the frequency of jumps to their size.
Among the available estimation methods, we choose the Peaks Over Threshold (POT) procedure that is used by Hainaut and Moraux (2018 b) in regime switching frame-

work where the HP has an exponential kernel. The POT method models extreme values of the daily log-returns through the jump component of the log-return process. This is done in pratice by setting the thresholds of negative and positive daily log-return. Our estimation procedure based on the POT method is divided into three steps. In the first step, from the time series of m log-returns $x = \{x_1, ..., x_m\}$ at times $\{t_1, ..., t_m = T\}$ equally spaced by a lag Δ of one day of trading, we find μ and σ that maximize the log-likelihood of a normal distribution corresponding to the distribution of the log-return without jump component. This log-likelihood function is defined as follows:

$$\ln l(x_1, .., x_m) = \sum_{i=1}^{m} \ln\left(f_{\mathcal{N}\left(\mu - \frac{\sigma^2}{2}, \sigma^2 \Delta\right)}(x_i) \right),$$

where $f_{\mathcal{N}\left(\mu - \frac{\sigma^2}{2}, \sigma^2 \Delta\right)}$ is the density of the Normal distribution with mean $\mu - \frac{\sigma^2}{2}$ and variance $\sigma^2 \Delta$. The second step consists in detecting jumps given that jumps occur when the return is above or below some thresholds. These thresholds, are Normal quantiles[3] $\Phi^{-1}(\alpha_1)$ and $\Phi^{-1}(\alpha_2)$ that guarantee that the sample without jumps

$$\left\{ x_i \mid x_i \in \left[\Phi^{-1}(\alpha_1), \Phi^{-1}(\alpha_2) \right] \right\}$$

has a skewness and a kurtosis as closed as possible to 0 and 3 respectively :

$$\alpha_1, \alpha_2 = \arg \min_{\text{sample without jumps}} \left((Skewness)^2 + (Kurtosis - 3)^2 \right).$$

Considering these thresholds, we estimate by log-likelihood maximization σ such that the sample without jumps has a Normal distribution with a mean $\hat{\mu}\Delta$ and a standard deviation $\sigma\sqrt{\Delta}$: $x_i \sim \hat{\mu}\Delta + \sigma W_\Delta$, for all x_i in $\left[\Phi^{-1}(\alpha_1), \Phi^{-1}(\alpha_2) \right]$. Here $\hat{\mu}$ is the estimator of μ found at the previous calibration step. When jumps are detected, the variation of prices is assumed equal to the jump size. The jumps parameters are then estimated by log-likelihood maximization:

$$\rho^-, \rho^+, p = \arg\max \sum_{j=1}^{m} \ln \nu \left(x_j - \hat{\mu}\Delta \mid \rho^-, \rho^+, p \right) 1_{\{jump\, at\, t_j\}}, \tag{4.42}$$

where ν is the probability density of a double exponential jump defined in equation (4.40). Letting $\{\tau_1, \cdots, \tau_{n_T}\}$ the set of n_T observed jump times over the interval $[0, T]$, the log-likelihood of the HP with intensity

$$\lambda_t = \lambda_0 + \eta \sum_{j=1}^{N_t} \left| J_j \right| f_{\alpha,\beta}\left(t - \tau_j \right), \text{ for } t \in [0, T],$$

[3] Notice that Φ is the cumulative distribution function of the standard Normal distribution.

is given by

$$\begin{aligned}\log L\left(\tau_1,\cdots,\tau_{n_T}\mid\lambda_0,\eta,\alpha,\beta\right) &= -\int_0^T \lambda_s \mathrm{d}s + \sum_{l=1}^{n_T}\log\left(\lambda_{\tau_l}\right)\\ &= -\lambda_0 T - \eta\sum_{l=1}^{n_T-1}\sum_{j=1}^{l}\int_{\tau_l}^{\tau_{l+1}}\left|J_j\right| f_{\alpha,\beta}\left(s-\tau_j\right)\mathrm{d}s\\ &\quad + \sum_{l=1}^{n_T}\log\left(\lambda_0+\eta\sum_{j<l}\left|J_j\right| f_{\alpha,\beta}\left(\tau_l-\tau_j\right)\right).\end{aligned}$$

The estimated intensity parameters are then such that :

$$\lambda_0,\eta,\alpha,\beta \quad = \quad \arg\max \log L\left(\tau_1,\cdots,\tau_{n_T}\mid\lambda_0,\eta,\alpha,\beta\right) \tag{4.43}$$

The estimated parameters obtained at each step of our estimation procedure and their corresponding log-likelihood are reported in table 4.2. The table 4.2 also contains the estimated parameters of the Merton (1976) jump diffusion (MJD) and the Kou (2002) jump diffusion (KJD) models, some key results of which are reported in appendix A 4.A. The analysis of estimated parameters we observe that the Brownian volatility in our geometric Brownian motion self-exciting jump (GBMSJ) modelwith ML kernel is less than the geometric Brownian motion (GBM) one. The upward exponential jumps are less likely with a probability of 34%, but the average size of positive shocks $\frac{1}{\rho^+}$ is almost equal to the average size of negative shocks $-\frac{1}{\rho^-}$. The comparison of likelihoods in the right table of table 4.2 enables to concludes that the KJD model better fits the S&P 500 time series than the GBM and the MJD
models. Unlike the GBMSJ model with ML kernel, these last three models admit a closed-form expression for the log-return density. As notice in the introduction our GBMSJ model extend the KJD model by considering a stochastic arrival rate of jump. In figure 4.6, we compare the kernel density estimation obtained using the S&P 500 times series to the density of the log-return at time Δ for the fitted GBM, MJD and KJD models. We find that the density function of log-returns modeled by the jump diffusion models are quite close to the kernel density estimation of empirical log-returns. In paricular, the KJD model and the GBMSJ model with ML kernel better reproduce the skewness of the kernel density estimation of empirical log-returns than the GBM, MJD models. We assess the quality of the fit using the QQ graphs in figure 4.7. From figure 4.7 we conclude that the GBMSJ model with ML kernel fits well the S&P 500 time series.
Figure 4.8 draws sample path of filtered intensities of the fitted GBMSJ with ML kernel. We observe high jump intensities during crisis periods such as the credit crunch of 2008, the European sovereign debt crisis between late 2009 and 2012 and currently the covid-19. During these periods, there is a large spread between jump intensities and the average reversion levels. Due to the self-excitation property, we observe clustering of peaks that last for a longer period than when the HP kernel is an exponential function ($\alpha = 1$). Another feature induces by the memory ($\alpha < 1$) is the slow reversion

	GBMSJ with ML kernel
σ	0.1 (0.001)
μ	0.128 (0.03)
Log-likelihood	9402
ρ^-	-41.88 (3.34)
ρ^+	40.61 (4.51)
p	0.34 (0.031)
Log-likelihood	496
λ_0	25 (1.63)
η	0.28 (0.019)
α	$0.81\ \left(4.19\times 10^{-4}\right)$
β	$0.88\ \left(2.96\times 10^{-6}\right)$
Log-likelihood	503

	GBM	MJD	KJD
σ	0.153 (0.002)	0.102 (0.002)	0.09 (0.004)
μ	0.128 (0.046)	0.277 (0.046)	0.140 (0.046)
λ		48.89 (18.62)	70 (9.381)
σ_j		0.0178 (0.004)	
μ_j		6.13×10^{-4} (0.003)	
ρ^-			-88.44 (6.712)
ρ^+			99.37 (11.018)
p			0.37 (0.05)
Log-likelihood	9066	9365	9393

Table 4.2: The upper table shows the estimated parameters and their standard errors in bracket, obtained based on the POT procedure. The lower table presents the estimated parameters and their standard errors in bracket for the geometric Brownian motion (GBM), the Merton jump diffusion (MJD) and the Kou jump diffusion (KJD) models, obtained by log-likelihood maximization.

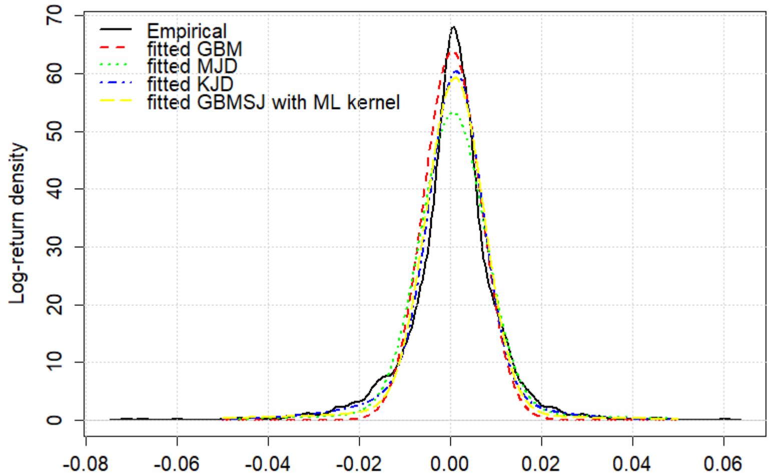


Figure 4.6: Kernel density estimation obtained using the S&P 500 times series and the density of the log-return at time Δ for the fitted GBM, MJD and KJD models and the GBMSJ model with ML kernel.

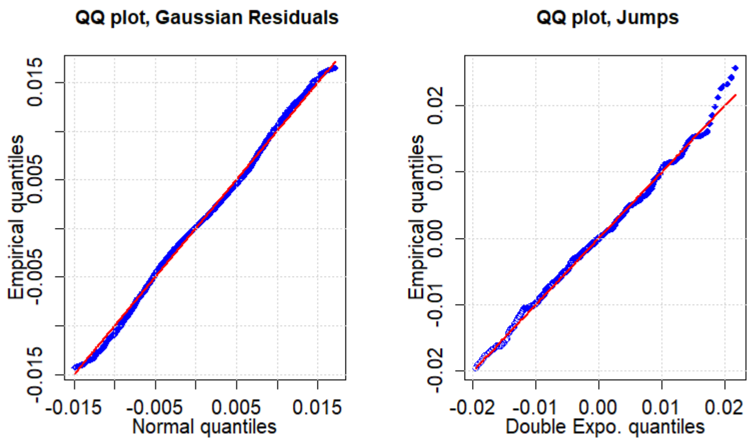


Figure 4.7: The left graph shows the QQ plot of the Gaussian part in our jump diffusion model versus the empirical distribution of log-returns on days when no jump is detected by the POT method. The right graph shows the QQ plots of truncated double exponential jumps identified by the POT method, versus the empirical distribution.

to the initial intensity λ_0. It is worth remembering that the self-excitation property, induced by the intensity of the Hawkes process is absent in the MJD and KJD models which have a constant arrival rate of jumps.

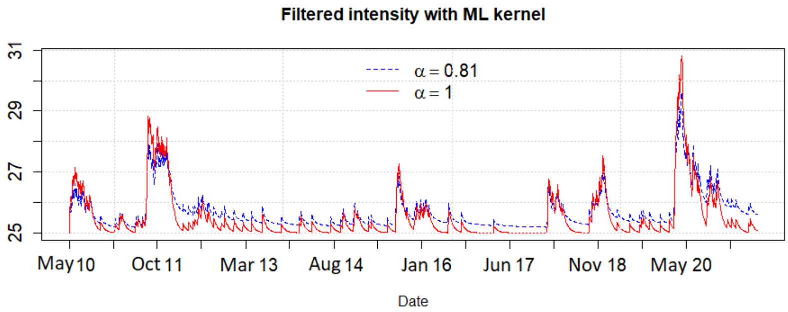


Figure 4.8: Sample path of filtered intensities of the GBMSJ with ML kernel based on the estimated parameters in table 4.2. Increase the parameter α from 0.81 to 1 corresponds to a loss of memory. Without memory the filtered intensities revert to the initial intensity λ_0 faster and clustering of peaks last for a shorter period.

4.3.3 Change of measure

Jump diffusion models lead to incomplete markets (see e.g. Kou 2002 and Tankov 2003). A consequence of this incompleteness is the existence of several equivalent measures that are all potential candidates for the definition of a risk neutral one. In this chapter, we focus on a family of changes of measure that are induced by an exponential martingale of the form:

$$\begin{aligned} M_t^{(n)}(\gamma,\varphi) \quad &:= \quad \exp\left(\kappa^{(n)}(\gamma)\int_0^t \lambda_s^{(n)}\mathrm{d}s+\gamma L_t^{(n)}\right) \\ &\qquad \times \exp\left(-\frac{1}{2}\int_0^t \varphi(s)^2\mathrm{d}s-\int_0^t \varphi(s)\,\mathrm{d}W_s\right) \end{aligned} \tag{4.44}$$

for all $0\leq t\leq T$, where $\varphi(.)$ is the market price for interest risk and the asset price of risk, assumed to be a linear function of the intensity of jumps.

Proposition 5. *If for any parameter γ,*

$$\kappa^{(n)}(\gamma) \quad = \quad 1-\psi(0,\gamma), \tag{4.45}$$

where $\psi(z_1,z_2):=\mathbb{E}\left(e^{z_1 J+z_2|J|}\right)$. Then $\left(M_t^{(n)}(\gamma,\varphi)\right)_{t\geq 0}$ is a local martingale. $\left(M_t^{(n)}(\gamma,\varphi)\right)_{t\geq 0}$ is a martingale on $[0,T]$ if $\mathbb{E}\left(\int_0^T \lambda_s^{(n)}ds\right)$ and $\mathbb{E}\left(\int_0^T \varphi(s)^2\,ds\right)$ are finite.

Proof. We denote by $U_t^{(n)}$ the exponent of $M_t^{(n)}$. Its SDE is given by

$$\mathrm{d}U_t^{(n)} = \kappa^{(n)}(\gamma)\lambda_t^{(n)}\mathrm{d}t - \frac{1}{2}\varphi(t)^2\mathrm{d}t - \varphi(t)\mathrm{d}W_t + \gamma \mathrm{d}L_t^{(n)} \tag{4.46}$$

Applying the Itô lemma semi-martingales to $M_t^{(n)}$ leads to

$$\begin{aligned}\mathrm{d}M_t^{(n)} &= M_t^{(n)}\mathrm{d}U_t^{(n)} + \frac{1}{2}M_t^{(n)}\mathrm{d}\left[U^{(n)},U^{(n)}\right]_t^c \\ &\quad + M_t^{(n)}\int_{-\infty}^{\infty}\left(e^{\gamma|z|}-1-\gamma|z|\right)\chi(\mathrm{d}z)\mathrm{d}N_t^{(n)} \\ &= M_t^{(n)}\kappa^{(n)}(\gamma)\lambda_t^{(n)}\mathrm{d}t - \varphi(t)M_t^{(n)}\mathrm{d}W_t + M_t^{(n)}\int_{-\infty}^{\infty}\left(e^{\gamma|z|}-1\right)\chi(\mathrm{d}z)\mathrm{d}N_t^{(n)},\end{aligned}$$

where $\left[U^{(n)},U^{(n)}\right]_t^c$ is the quadratic variation of the diffusion component of the process $U_t^{(n)}$.
Introducing the compensator of the jump process we have

$$\begin{aligned}\mathrm{d}M_t^{(n)} &= M_t^{(n)}\left(\kappa^{(n)}(\gamma) + \int_{-\infty}^{\infty}\left(e^{\gamma|z|}-1\right)\nu(\mathrm{d}z)\right)\lambda_t^{(n)}\mathrm{d}t - \varphi(t)M_t^{(n)}\mathrm{d}W_t \\ &\quad + M_t^{(n)}\int_{-\infty}^{\infty}\left(e^{\gamma|z|}-1\right)\left(\chi(\mathrm{d}z)\mathrm{d}N_t^{(n)} - \lambda_t^{(n)}\nu(\mathrm{d}z)\mathrm{d}t\right)\end{aligned}$$

where $J = \int_{-\infty}^{\infty}\chi(\mathrm{d}z)$. Since the integral with respect to $\chi(\mathrm{d}z)\mathrm{d}N_t^{(n)} - \lambda_t^{(n)}\nu(\mathrm{d}z)\mathrm{d}t$ is a local martingale, $M_t^{(n)}$ is also a local martingale if and only if

$$0 = \kappa^{(n)}(\gamma) + \psi(0,\gamma) - 1 \tag{4.47}$$

where $\psi(z_1,z_2) := \mathbb{E}\left(e^{z_1 J + z_2|J|}\right)$. Its follows that under the condition (4.47), $\left(M_t^{(n)}(\gamma,\varphi)\right)_{t\geq 0}$ is solution of a driftless SDE which is a martingale under the interval $[0,T]$ if $\mathbb{E}\left(\int_0^T \lambda_s^{(n)}\mathrm{d}s\right) < \infty$ and $\mathbb{E}\left(\int_0^T \varphi(s)^2\mathrm{d}s\right) < \infty$. The first condition ensures that the compensate counting process is a martingale on $[0,T]$, while the second condition guarantees that the diffusion component of $M_t^{(n)}$ is a martingale on $[0,T]$ (see e.g. Revuz and Yor 1999). □

Assuming the existence of $\kappa^{(n)}(\gamma)$ in equation (4.45), $\left(M_t^{(n)}(\gamma,\varphi)\right)_{t\geq 0}$ is a local martingale. When $\mathbb{E}\left(\int_0^T \lambda_s^{(n)}\mathrm{d}s\right)$ and $\mathbb{E}\left(\int_0^T \varphi(s)^2\mathrm{d}s\right)$ are finite, $\left(M_t^{(n)}(\gamma,\varphi)\right)_{t\geq 0}$ is a martingale on $[0,T]$ and we may therefore define an equivalent measure $\mathbb{Q}^{\gamma,\varphi}$ by the ratio :

$$\left.\frac{\mathrm{d}\mathbb{Q}^{\gamma,\varphi}}{\mathrm{d}\mathbb{P}}\right|_{\mathscr{F}_t} = M_t^{(n)}(\gamma,\varphi). \tag{4.48}$$

Proposition 6. *Under the equivalent measure $\mathbb{Q}^{\gamma,\varphi}$, we define by $N_t^{(n),Q}$ the counting process with the following intensity*

$$\lambda_t^{(n),Q} = \psi(0,\gamma)\lambda_t^{(n)}. \tag{4.49}$$

we also defined the double exponential distribution random variables J^{Q} of parameters :

$$\begin{aligned}
\rho^{+Q} &= \rho^{+}-\gamma, & (4.50)\\
\rho^{-Q} &= \rho^{-}+\gamma, & (4.51)\\
p^{Q} &= \frac{p\rho^{+}\rho^{-Q}}{p\rho^{+}\rho^{-Q}+(1-p)\rho^{-}\rho^{+Q}}, & (4.52)
\end{aligned}$$

with the following characteristic function :

$$\psi^{Q}(z_1,z_2):=\mathbb{E}\left(e^{z_1 J^{Q}+z_2\left|J^{Q}\right|}\right) = \frac{\psi(z_1,z_2+\gamma)}{\psi(0,\gamma)}. \tag{4.53}$$

and the process $L_t^{(n),Q}=\sum_{l=1}^{N_t^{(n),Q}}\left|J_l^{Q}\right|$. Then the dynamics of $\left(\lambda_t^{(n)}\right)_{t\geq 0}$ under $\mathbb{Q}^{\gamma,\varphi}$ is ruled by the following SDE

$$d\lambda_t^{(n),Q} = \eta^{Q}\sum_{v=1}^{n} m_v^{(n)} dY_t^{(b_v^{(n)}),Q}, \tag{4.54}$$

where $\lambda_0^{Q}=\psi(0,\gamma)\lambda_0$, $\eta^{Q}=\eta\psi(0,\gamma)$ and

$$dY_t^{(b_v^{(n)}),Q}=-b_v^{(n)}Y_t^{(b_v^{(n)}),Q}dt+dL_t^{(n),Q}. \tag{4.55}$$

Proof. If $\lambda_t^{(n),Q}=\psi(0,\gamma)\lambda_t^{(n)}$, the conditional mgf of $\lambda_t^{(n),Q}$ under the equivalent measure $\mathbb{Q}^{\gamma,\varphi}$ induced by the exponential martingale $M_t^{(n)}$ is given by

$$\begin{aligned}
\mathbb{E}^{Q}\left(e^{\omega\lambda_s^{(n),Q}}|\mathscr{F}_t\right) &= \frac{\mathbb{E}\left(\mathbb{E}\left(\frac{\mathrm{d}\mathbb{Q}^{\gamma,\varphi}}{\mathrm{d}\mathbb{P}}|\mathscr{F}_s\right)e^{\omega\psi(0,\gamma)\lambda_s^{(n)}}|\mathscr{F}_t\right)}{\mathbb{E}\left(\frac{\mathrm{d}\mathbb{Q}^{\gamma,\varphi}}{\mathrm{d}\mathbb{P}}|\mathscr{F}_t\right)}\\
&= e^{-U_t^{(n)}}\mathbb{E}\left(e^{U_s^{(n)}+\omega\psi(0,\gamma)\lambda_s^{(n)}}|\mathscr{F}_t\right).
\end{aligned}$$

As in the proof of proposition 3, if we note

$$f\equiv f\left(t,s,U_t^{(n)},\lambda_t^{(n)},\left(Y_t^{(b_v^{(n)})}\right)_{v=1,\dots,n},N_t^{(n)}\right) := \mathbb{E}\left(e^{U_s^{(n)}+\omega\psi(0,\gamma)\lambda_s^{(n)}}|\mathscr{F}_t\right)$$

and we conjecture that f is an exponential affine function of the exponent of $M_t^{(n)}$ denoted by $U_t^{(n)}$, $\lambda_t^{(n)}$ and $\left(Y_t^{(v)}\right)_{v=1,\dots,n}$:

$$\begin{aligned}
f &:= \exp\left(A(t,s,\omega)+\psi(0,\gamma)B(t,s,\omega)\lambda_t^{(n)}\right)\\
&\quad\times\exp\left(\sum_{v=1}^{n}C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}Y_t^{(b_v^{(n)})}+D(t,s,\omega)U_t^{(n)}\right),
\end{aligned}$$

where $A(s,s,\omega)=0$, $B(s,s,\omega)=\omega$, $C\left(s,s,b_v^{(n)},\omega\right)=0$ and $D(s,s,\omega)=1$.Then , according to the Itô's lemma for the semi-martingales, f solves the following differential equation :

$$\begin{aligned} 0 \quad = \quad & \frac{\partial f}{\partial t}+\frac{\partial f}{\partial U^{(n)}}\left(\kappa^{(n)}(\gamma)\lambda_t^{(n)}-\frac{1}{2}\varphi(t)^2\right) \\ & +\frac{1}{2}\frac{\partial^2 f}{\partial U^{(n)2}}\varphi(t)^2-\sum_{v=1}^{n}\frac{\partial f}{\partial Y^{(v)}}b_v^{(n)}Y_t^{(b_v^{(n)})}-\frac{\partial f}{\partial \lambda^{(n)}}\eta\sum_{v=1}^{n}m_v^{(n)}b_v^{(n)}Y_t^{(b_v^{(n)})} \\ & +\lambda_t^{(n)}\int_{-\infty}^{\infty}\Delta f(z)\,\nu(\mathrm{d}z) \end{aligned} \tag{4.56}$$

where $\Delta f(z)$ is the variation of f when a jump of size z occurs and the time just before.

$$\frac{\partial f}{\partial U^{(n)}}=D(t,s,\omega)f,\ \frac{\partial f}{\partial \lambda^{(n)}}=\psi(0,\gamma)B(t,s,\omega)f,\ \frac{\partial f}{\partial Y^{(v)}}=C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}f,$$

$$\frac{\partial^2 f}{\partial U^{(n)2}}=D(t,s,\omega)^2 f,\ \frac{\partial f}{\partial t}=f\left(\frac{\partial A}{\partial t}(t,s,\omega)+\psi(0,\gamma)\frac{\partial B}{\partial t}(t,s,\omega)\lambda_t^{(n)}\right)$$

$$+\left(\sum_{v=1}^{n}\frac{\partial C}{\partial t}\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}Y_t^{(b_v^{(n)})}+\frac{\partial D}{\partial t}(t,s,\omega)U_t^{(n)}\right)f,$$

and the variation of f when a jump of size z occurs and the time just before is

$$\begin{aligned} & \left(\exp\left(\left(\eta\psi(0,\gamma)B(t,s,\omega)\sum_{v=1}^{n}m_v^{(n)}\right.\right.\right. \\ & \qquad \left.\left.\left.+\sum_{v=1}^{n}C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}+D(t,s,\omega)\gamma\right)|z|\right)-1\right)f \end{aligned}$$

Injecting all these expressions into equation (5.73) leads to :

$$\begin{aligned} 0 \quad = \quad & \frac{\partial A}{\partial t}(t,s,\omega)+\psi(0,\gamma)\frac{\partial B}{\partial t}(t,s,\omega)\lambda_t^{(n)}+\sum_{v=1}^{n}\frac{\partial C}{\partial t}\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}Y_t^{(b_v^{(n)})} \\ & +\frac{\partial D}{\partial t}(t,s,\omega)U_t^{(n)}+D(t,s,\omega)\left(\kappa^{(n)}(\gamma)\lambda_t^{(n)}-\frac{1}{2}\varphi(t)^2\right)+\frac{1}{2}D(t,s,\omega)^2\varphi(t)^2 \\ & -\sum_{v=1}^{n}\left(C\left(t,s,b_v^{(n)},\omega\right)+\eta\psi(0,\gamma)B(t,s,\omega)\right)m_v^{(n)}b_v^{(n)}Y_t^{(b_v^{(n)})}+\lambda_t^{(n)}\Bigg(\\ & \psi\left(0,\sum_{v=1}^{n}\left(\eta\psi(0,\gamma)B(t,s,\omega)+C\left(t,s,b_v^{(n)},\omega\right)\right)m_v^{(n)}+D(t,s,\omega)\gamma\right)-1\Bigg) \end{aligned}$$

or

$$0 = \frac{\partial A}{\partial t}(t,s,\omega) - \frac{1}{2}D(t,s,\omega)\varphi(t)^2 + \frac{1}{2}D(t,s,\omega)^2\varphi(t)^2 \quad (4.57)$$

$$0 = \psi(0,\gamma)\frac{\partial B}{\partial t}(t,s,\omega) + D(t,s,\omega)\kappa^{(n)}(\gamma) - 1 \quad (4.58)$$
$$+\psi\left(0, \eta\psi(0,\gamma)B(t,s,\omega)\sum_{v=1}^{n} m_v^{(n)} + \sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)} + D(t,s,\omega)\gamma\right),$$

$$0 = \frac{\partial C}{\partial t}\left(t,s,b_v^{(n)},\omega\right) - \left(C\left(t,s,b_v^{(n)},\omega\right) + \eta\psi(0,\gamma)B(t,s,\omega)\right)b_v^{(n)} \quad (4.59)$$
$$\forall v \in \{1,\cdots,n\}$$

$$0 = \frac{\partial D}{\partial t}(t,s,,\omega). \quad (4.60)$$

$D(t,s) = 1$ solves the last ODE as $D(s,s) = 1$. It follows that $A(t,s) = 0$ solves the first ODE and that the second ODE becomes

$$0 = \psi(0,\gamma)\frac{\partial B}{\partial t}(t,s,\omega) + \kappa^{(n)}(\gamma) - 1 \quad (4.61)$$
$$+\psi\left(0, \eta\psi(0,\gamma)B(t,s,\omega)\sum_{v=1}^{n} m_v^{(n)} + \sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)} + \gamma\right)$$

Using equation (4.45) and given the double exponential distribution random variable J^Q with ψ^Q the characteristic function of the pair $\left(J^Q, \left|J^Q\right|\right)$ defined by

$$\psi^Q(z_1,z_2) = \frac{\psi(z_1,z_2+\gamma)}{\psi(0,\gamma)},$$

for $\omega_2 = 0$, we have a ODE similar to the one obtained under the real measure (4.32) :

$$\frac{\partial B}{\partial t}(t,s,\omega) = 1 - \psi^Q\left(0, \eta^Q B(t,s,\omega)\sum_{v=1}^{n} m_v^{(n)} + \sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}\right)$$

where for all $v \in \{1,\cdots,n\}$

$$C\left(t,s,b_v^{(n)},\omega\right) = -b_v^{(n)}\eta^Q\int_t^s e^{-b_v^{(n)}(u-t)}B(u,s,\omega)\,\mathrm{d}u,$$

solves equation (4.59) and $\eta^Q = \eta\psi(0,\gamma)$. Hence,

$$\mathbb{E}^Q\left(e^{\omega\lambda_s^{(n),Q}}|\mathscr{F}_t\right) = \exp\left(B(t,s,\bar{\omega})\lambda_t^{(n),Q}\right.$$
$$\left.+\sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\bar{\omega}\right)m_v^{(n)}Y_t^{(b_v^{(n)}),Q}\right)$$

and we conclude by comparing with the results of proposition 3 for $\omega_1 = \omega$ and $\omega_2 = 0$. □

Proposition 7. *If the $\mathscr{F}_t$ adapted process defining the martingale (4.44) on $[0,T]$, is equal to*

$$\varphi(t) = \frac{\mu - r}{\sigma} + \lambda_t^{(n)} \frac{\psi(0,\gamma)\left(\psi^Q(1,0)-1\right) - (\psi(1,0)-1)}{\sigma}, \tag{4.62}$$

then the equivalent measure is risk neutral and the log-return corresponding to the partition $\mathscr{E}^{(n)}$, $X_t^{(n)} := \ln \frac{S_t^{(n)}}{S_0}$, is driven by the following dynamics under the measure $\mathbb{Q}^{\gamma,\varphi}$:

$$\begin{aligned} dX_t^{(n)} &= \left(r - \frac{\sigma^2}{2} - \mathbb{E}\left(e^{J^Q} - 1\right)\lambda_t^{(n),Q}\right)dt + \sigma dW_t^Q \\ &\quad + d\sum_{i=1}^{N_t^{(n),Q}} J_i^Q, \end{aligned} \tag{4.63}$$

and the asset price is ruled by the following SDE :

$$\begin{aligned} dS_t^{(n)} &= rS_t^{(n)}dt + \sigma S_t^{(n)} dW_t^Q \\ &\quad + S_t^{(n)}\left(d\left(\sum_{i=1}^{N_t^{(n),Q}}\left(e^{J_i^Q} - 1\right)\right) - \mathbb{E}\left(e^{J^Q} - 1\right)\lambda_t^{(n),Q} dt\right) \end{aligned} \tag{4.64}$$

Proof. The discounted asset price $\widetilde{S}_t^{(n)} = e^{-rt}S_t^{(n)}$ is a $\mathscr{F}_t$- martingale on $[0,T]$, under the probability measure $\mathbb{Q}^{\gamma,\varphi}$ if and only if $M_t^{(n)}\widetilde{S}_t^{(n)}$ is a $\mathscr{F}_t$- martingale on $[0,T]$, under the probability measure $\mathbb{P}$. If $I_t^{(n)} = U_t^{(n)} - rt + X_t^{(n)}$ then $M_t^{(n)}\widetilde{S}_t^{(n)} = S_0 e^{I_t^{(n)}}$ and

$$\begin{aligned} \mathrm{d}I_t^{(n)} &= \left(\kappa^{(n)}(\gamma)\lambda_t^{(n)} - \frac{1}{2}\varphi(t)^2 - r + \mu - \frac{\sigma^2}{2} - \mathbb{E}\left(e^J - 1\right)\lambda_t^{(n)}\right)\mathrm{d}t \\ &\quad (\sigma - \varphi(t))\,\mathrm{d}W_t + \mathrm{d}\left(\sum_{i=1}^{N_t^{(n)}} J_i\right) + \gamma \mathrm{d}L_t^{(n)}. \end{aligned}$$

Using Itô's differentiation rule the SDE of the $\mathscr{F}_t$ adapted process $M_t^{(n)}\widetilde{S}_t^{(n)}$ is given by :

$$\begin{aligned} \mathrm{d}\left(M_t^{(n)}\widetilde{S}_t^{(n)}\right) &= S_0 \mathrm{d}e^{I_t^{(n)}} \\ &= S_0 e^{I_t^{(n)}}\left(\mathrm{d}I_t^{(n),c} + \frac{1}{2}(\sigma - \varphi(t))^2\,\mathrm{d}t\right) + S_0 e^{I_t^{(n)}}\lambda_t^{(n)}(\psi(1,\gamma)-1)\,\mathrm{d}t \\ &\quad + S_0 e^{I_t^{(n)}}\int_{-\infty}^{\infty}(\exp(z+\gamma|z|)-1)\left(\chi(\mathrm{d}z)\,\mathrm{d}N_t^{(n)} - \lambda_t^{(n)}\nu(\mathrm{d}z)\,\mathrm{d}t\right), \end{aligned} \tag{4.65}$$

where $I_t^{(n),c}$ is the diffusion part of the process $I_t^{(n)}$. When $\mathbb{E}\left(\int_0^T \lambda_s^{(n)}\mathrm{d}s\right)$ is finite, it follows that $M_t^{(n)}\widetilde{S}_t^{(n)}$ is a $\mathscr{F}_t$- martingale on $[0,T]$, under the probability measure $\mathbb{P}$ if and only if

$$
\begin{aligned}
0 &= \mu - r - \sigma\varphi(t) \\
&\quad + \lambda_t^{(n)} \left(\kappa^{(n)}(\gamma) - (\psi(1,0) - 1) + \psi(1,\gamma) - 1 \right).
\end{aligned}
$$

Taking into account equations (4.45) and (4.53) this last equation is rewritten to obtain equation (4.62) that concludes this proof. □

4.3.4 Option pricing

In this subsection we show how to price European call option of maturity T and strike K, written on the terminal spot S_T of some underlying asset. We start by determining the conditional mgf of the log return process under the equivalent measure $\mathbb{Q}^{\gamma,\varphi}$. Next, we will illustrate the Carr-Madan approach (1999) for computing the value of a call option given the conditional mgf of X_T. To this aim, the next proposition sets out the conditional mgf of the approximated log-return process.

Proposition 8. *The moment generating of $X_s^{(n)}$ under $\mathbb{Q}^{\gamma,\varphi}$ conditionally to the information at time $t \leq s$ is given by*

$$
\begin{aligned}
\Upsilon_{t,s}^{(n)}(\omega) &:= \mathbb{E}^Q \left(e^{\omega X_s^{(n)}} \mid \mathscr{F}_t \right) \qquad (4.66)\\
&= \exp\left(\omega X_t^{(n)} + A(t,s,\omega) + B(t,s,\omega)\,\lambda_t^{(n),Q} \right) \\
&\quad \times \exp\left(\sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\omega\right) m_v^{(n)} Y_t^{\left(b_v^{(n)}\right),Q} \right),
\end{aligned}
$$

for all $\omega \in \mathbb{C}_-$ that satisfies the regularity condition $\psi\left(\omega, \eta^Q \sum_{v=1}^n m_v^{(n)}\right) < \infty$. The functions A, C are such that

$$
A(t,s,\omega) = \omega \left(r + (\omega - 1)\frac{\sigma^2}{2} \right)(s-t), \tag{4.67}
$$

$$
C\left(t,s,b_v^{(n)},\omega\right) = -b_v^{(n)} \int_t^s e^{-b_v^{(n)}(u-t)} B(u,s,\omega)\,\eta^Q du, \tag{4.68}
$$

and the function B solves the following ODE:

$$
\begin{aligned}
\frac{\partial B}{\partial t}(t,s,\omega) &= 1 - \omega + \omega\psi^Q(1,0) \qquad (4.69)\\
&\quad - \psi^Q\left(\omega, \sum_{v=1}^{n} C\left(t,s,b_v^{(n)},\omega\right) m_v^{(n)} + \eta^Q B(t,s,\omega) \sum_{v=1}^{n} m_v^{(n)} \right)
\end{aligned}
$$

with the terminal condition $B(s,s,\omega) = 0$.

Proof. Let $\omega \in \mathbb{C}_-$. As in proposition 3, if we note

$$f \equiv f\left(t,s,X_t^{(n)},\lambda_t^{(n),Q},\left(Y_t^{\left(b_\nu^{(n)}\right),Q}\right)_{\nu=1,\ldots,n},N_t^{(n),Q}\right) \quad := \quad \mathbb{E}^Q\left(e^{\omega X_s^{(n)}} \mid \mathscr{F}_t\right)$$

and under the regularity condition $\psi\left(\omega,\eta^Q\sum_{\nu=1}^n m_\nu^{(n)}\right)<\infty$, we conjecture that f is an exponential affine function of $X_t^{(n)}$, $\lambda_t^{(n),Q}$ and $\left(Y_t^{\left(b_\nu^{(n)}\right),Q}\right)_{\nu=1,\ldots,n}$,

$$\begin{aligned} f \quad := \quad & \exp\left(\omega X_t^{(n)}+A(t,s,\omega)\right) && (4.70)\\ & \times\exp\left(B(t,s,\omega)\lambda_t^{(n),Q}+\sum_{\nu=1}^n C\left(t,s,b_\nu^{(n)},\omega\right)m_\nu^{(n)}Y_t^{\left(b_\nu^{(n)}\right),Q}\right) \end{aligned}$$

where $A(s,s,\omega)=0$ $B(s,s,\omega)=0$ and $C\left(s,s,b_\nu^{(n)}\right)=0$,for all $\nu\in\{1,\cdots,n\}$. Then, according to the Itô's lemma, f solves the following stochastic differential equation :

$$\begin{aligned} 0 \quad = \quad & \frac{\partial f}{\partial t}-\sum_{\nu=1}^n\frac{\partial f}{\partial Y^{(\nu),Q}}b_\nu^{(n)}Y_t^{\left(b_\nu^{(n)}\right),Q} && (4.71)\\ & +\frac{\partial f}{\partial X^{(n)}}\left(r-\frac{\sigma^2}{2}-\mathbb{E}\left(e^{J^Q}-1\right)\lambda_t^{(n),Q}\right) && (4.72)\\ & +\frac{1}{2}\frac{\partial^2 f}{\partial X^{(n)2}}\sigma^2-\frac{\partial f}{\partial\lambda^{(n),Q}}\eta^Q\sum_{\nu=1}^n m_\nu^{(n)}b_\nu^{(n)}Y_t^{\left(b_\nu^{(n)}\right),Q}+\lambda_t^{(n),Q}\\ & \times\int_{-\infty}^{\infty}\Delta f(z)\,\nu^Q(\mathrm{d}z) \end{aligned}$$

where

$$\frac{\partial f}{\partial Y^{(\nu)}}=C\left(t,s,b_\nu^{(n)},\omega\right)m_\nu^{(n)}f,\quad \frac{\partial f}{\partial X^{(n)}}=\omega f,\quad \frac{\partial^2 f}{\partial X^{(n)2}}=\omega^2 f,\ \frac{\partial f}{\partial\lambda^{(n),Q}}=B(t,s,\omega),$$

$$\frac{\partial f}{\partial t}=\left(\frac{\partial A}{\partial t}(t,s,\omega)+\sum_{\nu=1}^n\frac{\partial C}{\partial t}\left(t,s,b_\nu^{(n)},\omega\right)m_\nu^{(n)}Y_t^{\left(b_\nu^{(n)}\right),Q}+\frac{\partial B}{\partial t}(t,s,\omega)\lambda_t^{(n),Q}\right)f,$$

and $\Delta f(z)$ is the variation of f when a jump of size z occurs and the time just before :

$$\Delta f(z)=\left(\exp\left(\omega z+\sum_{\nu=1}^n C\left(t,s,b_\nu^{(n)},\omega\right)m_\nu^{(n)}|z|+\eta^Q B(t,s,\omega)\sum_{\nu=1}^n m_\nu^{(n)}|z|\right)-1\right)f$$

Injecting all these expressions into equation (4.71) leads to :

$$\begin{aligned}
0 &= \frac{\partial A}{\partial t}(t,s,\omega)+\sum_{v=1}^{n}\frac{\partial C}{\partial t}\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}Y_t^{\left(b_v^{(n)}\right),Q}+\frac{\partial B}{\partial t}(t,s,\omega)\lambda_t^{(n),Q} \\
&\quad -\sum_{v=1}^{n}C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}b_v^{(n)}Y_t^{\left(b_v^{(n)}\right),Q}-B(t,s,\omega)\eta^{Q}\sum_{v=1}^{n}m_v^{(n)}b_v^{(n)}Y_t^{\left(b_v^{(n)}\right),Q} \\
&\quad +\omega\left(r-\frac{\sigma^2}{2}-\mathbb{E}\left(e^{J^Q}-1\right)\lambda_t^{(n),Q}\right)+\frac{1}{2}\omega^2\sigma^2+\lambda_t^{(n),Q}\int_{-\infty}^{\infty}\left(\exp\left(\omega z\right.\right. \\
&\quad \left.\left.+\sum_{v=1}^{n}C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}|z|+\eta^{Q}B(t,s,\omega)\sum_{v=1}^{n}m_v^{(n)}|z|\right)-1\right)\nu^{Q}(\mathrm{d}z),
\end{aligned}$$

or

$$0 = \frac{\partial A}{\partial t}(t,s,\omega)+\omega\left(r+(\omega-1)\frac{\sigma^2}{2}\right), \tag{4.73}$$

$$\begin{aligned}
0 &= \frac{\partial B}{\partial t}(t,s,\omega)-\omega\mathbb{E}\left(e^{J^Q}-1\right)+\int_{-\infty}^{\infty}\left(\exp\left(\omega z\right.\right. \\
&\quad \left.\left.+\sum_{v=1}^{n}C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}|z|+\eta^{Q}B(t,s,\omega)\sum_{v=1}^{n}m_v^{(n)}|z|\right)-1\right)\nu^{Q}(\mathrm{d}z),
\end{aligned} \tag{4.74}$$

$$\begin{aligned}
0 &= \sum_{v=1}^{n}\frac{\partial C}{\partial t}\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}-\sum_{v=1}^{n}C\left(t,s,b_v^{(n)},\omega\right)m_v^{(n)}b_v^{(n)} \\
&\quad -B(t,s,\omega)\eta^{Q}\sum_{v=1}^{n}m_v^{(n)}b_v^{(n)}.
\end{aligned} \tag{4.75}$$

Then

$$A(t,s,\omega) = \omega\left(r+(\omega-1)\frac{\sigma^2}{2}\right)(s-t),$$

and the last equation is

$$0 = \frac{\partial C}{\partial t}\left(t,s,b_v^{(n)},\omega\right)-\left(C\left(t,s,b_v^{(n)},\omega\right)+B(t,s,\omega)\eta^{Q}\right)b_v^{(n)},$$

that admits as solution

$$C\left(t,s,b_v^{(n)},\omega\right) = -b_v^{(n)}\int_t^s e^{-b_v^{(n)}(u-t)}B(u,s,\omega)\eta^{Q}\mathrm{d}u$$

for all $v\in\{1,\cdots,n\}$.
Using the definition of the characteristic function ψ^Q (4.53) we retrieve the ODE of the function B (5.37) that concludes this proof. □

The next proposition establishes the conditional mgf of the log-return process considering that the approximated log-return process tends in distribution to the true log-return process.

Proposition 9. *The mgf of X_s under $\mathbb{Q}^{\gamma,\varphi}$ conditionally to the information at time $t \leq s$ is given by*

$$\begin{aligned} \Upsilon_{t,s}(\omega) &:= \mathbb{E}^{Q}\left(e^{\omega X_s} \mid \mathscr{F}_t\right) \qquad (4.76)\\ &= \exp\left(\omega X_t + A(t,s,\omega)\right) \\ &\quad \times \exp\left(-\eta \int_0^{\infty} \int_t^s v e^{-v(u-t)} B(u,s,\omega)\, Y_t^{(v),Q} H_{\alpha,\beta}(v)\, du dv \right. \\ &\qquad \left. + B(t,s,\omega)\,\lambda_t^{Q}\right), \end{aligned}$$

for all $\omega \in \mathbb{C}_-$ that satisfies the regularity condition $\psi\left(\omega,\eta^{Q}\right) < \infty$. The function A is defined as follows

$$A(t,s,\omega) = \omega\left(r + (\omega-1)\frac{\sigma^2}{2}\right)(s-t), \qquad (4.77)$$

and B is a time dependent function that solves the following integro-differential equation:

$$\begin{aligned} \frac{\partial B}{\partial t}(t,s,\omega) &= \omega\left(\psi^{Q}(1,0) - 1\right) \\ &\quad - \left(\psi^{Q}\left(\omega, \eta^{Q}\int_t^s B(u,s,\omega)\, f^{(1)}_{\alpha,\beta}(u-t)\,du + \eta^{Q} B(t,s,\omega)\right) - 1\right) \end{aligned}$$

with the terminal condition $B(s,s,\omega) = 0$.

Proof. This proof is similar to the one of proposition 4, as $X_s^{(n)}$ converges to X_s almost surely and by the dominated convergence the conditional mgf of the log-return process is equal to the limit with respect of n of the conditional mgf of the approximated log-return process. By some grouping and using the frequency derivative property of Laplace transform we retrieve the results of proposition 9. □

Let us focus on the pricing of the European call option of maturity T, written on S_t. Its payoff and its strike are expressed as function of the log-return X_t and of the log-strike k. The price $C(k)$ of this European call option is then a function of the log-strike k. If we denote by $f_{t,T}$ the risk neutral density at time $t \leq T$ of the log-return $X_T \mid \mathscr{F}_t$, the price of this call option is equal to the expected discounted payoffs :

$$C(k) = S_t \int_k^{+\infty} e^{-r(T-t)}\left(e^x - e^k\right) f_{t,T}(x)\,\mathrm{d}x. \qquad (4.78)$$

The next proposition approximates this European call price using the trapezoidal quadrature approximation of the European call price formula rewritten as in Carr and Madan (1999).

Proposition 10. *Let M be the number of steps used in the fast Fourier Transform (FFT) and $\Delta_k = \frac{2k_{max}}{M}$, be the step of discretisation. Let us denote $\gamma_l = \frac{1}{2}1_{\{l=1\}} +$*

$1_{\{l\neq 1\}}$, $\Delta_w = \frac{2\pi}{M\Delta_k}$ *and* $w_l = (l-1)\Delta_w$. *The value of* $C(k)$ *at points* $k_j = -\frac{M}{2}\Delta_k + \Delta_k(j-1)$ *are approximated by*

$$C\left(k_j\right) \approx \frac{2S_t e^{-ak_j - r(T-t)}}{M\Delta_k} \times \sum_{l=1}^{M} e^{-i\frac{2\pi}{M}(l-1)(u-1)}(-1)^{l-1}\left(\gamma_l \frac{\Upsilon_{t,T}(1+a+i\omega_l)}{a^2+a-\omega_l^2+i(2a+1)\omega_l}\right), \tag{4.79}$$

for all $j = 1,\cdots,M$, *with* $\Upsilon_{t,s}(\omega) = \mathbb{E}^{Q}\left(e^{\omega X_s} \mid \mathscr{F}_t\right)$ *and* $a > 0$ *such that* $\Upsilon_{t,T}(a+1)$ *is finite.*

Proof. This proof relies on the work of Carr and Madan (1999). The authors value call options in an efficient way with fast Fourier transform (FFT). According to authors, for a given dampening parameter $a > 0$ the call option price in our setting is given by

$$C(k) = \frac{S_t e^{-ak - r(T-t)}}{\pi} \int_0^{+\infty} \frac{e^{-i\omega k}\Upsilon_{t,T}(1+a+i\omega)}{a^2+a-\omega^2+i(2a+1)\omega} \mathrm{d}\omega \tag{4.80}$$

with an optimal dampening parameter a such that $\Upsilon_{t,T}(a+1) < +\infty$. The trapezoidal quadrature approximation of this latter call option price formula for M discrete log-strikes $k_j = -k_{max} + (j-1)\Delta_k$, $j = 1,\ldots,M$, equally spaced $\Delta_k = \frac{2k_{max}}{M}$ is given by :

$$C\left(k_j\right) \approx \frac{S_t e^{-ak_j - r(T-t)}}{\pi} \sum_{l=1}^{M} e^{-i\omega_l k_j}\left(\gamma_l \frac{\Upsilon_{t,T}(1+a+i\omega_l)}{a^2+a-\omega_l^2+i(2a+1)\omega_l}\Delta\omega\right) \tag{4.81}$$

where $\gamma_l = \frac{1}{2}1_{\{l=1\}} + 1_{\{l\neq 1\}}$ is the trapezoidal quadrature weights, $\Delta_\omega = \frac{2\pi}{M\Delta_k}$ ensures that the absolute value of the integrand function is sufficiently small for all $\omega > \omega_M$ with $\omega_l = (l-1)\Delta_\omega$ for $l = 1,\ldots,M$. Substituting the formulae of k_u and ω_l in equation (4.81) yields to

$$C\left(k_j\right) \approx \frac{S_t e^{-ak_j - r(T-t)}}{\pi} \times \sum_{l=1}^{M} e^{-i\frac{2\pi}{M}(l-1)(j-1)}(-1)^{l-1}\left(\gamma_l \frac{\Upsilon_{t,T}(1+a+i\omega_l)}{a^2+a-\omega_l^2+i(2a+1)\omega_l}\Delta\omega\right). \tag{4.82}$$

Next, the form of equation (4.82) allows to evaluate the approximated European call price using the FFT algorithm. Replacing $\frac{\Delta\omega}{\pi}$ by $\frac{2}{M\Delta_k}$ in equation (4.82) we conclude the proof. □

As main advantage, for a given maturity, the proposition 10 enables one to simultaneously evaluate call options with different strikes. When interested with European put options, their prices can be deduced using the put call parity.

4.3.5 Numerical analysis

This subsection highlights the impact of our framework on the S&P 500 log-return's distribution. It also shows the sensitivity of S&P 500 call options to the memory parameter.

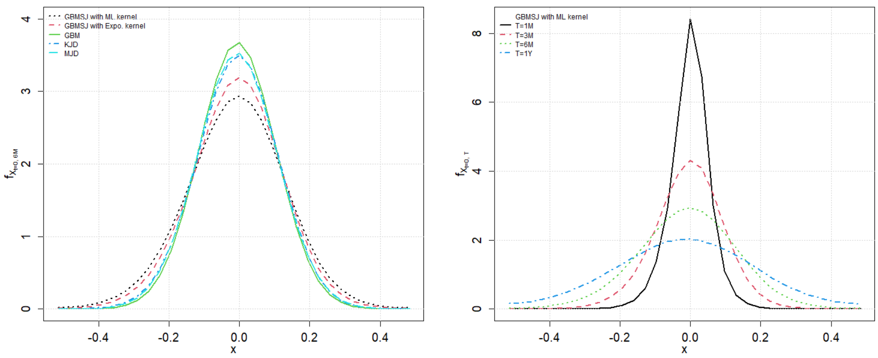


Figure 4.9: The left graph shows three S&P 500 log-return's distributions that represent the differences among the GBMSJ model with ML kernel, the GBMSJ model with exponential kernel, the GBM, the KJD and MJD models. The right graph shows the S&P 500 log-return's distributions at different times ranging from 1 month to 1 year. The parameters used for this illustration are presented in Table 4.2. The parameters of the GBMSJ with exponential kernel are the same as the one with ML kernel, except for α which is equal to 1.

Figure 4.9 depicts the risk neutral log-return's distribution obtained by fast Fourier transform as in Njike and Hainaut (2020) (Proposition 4.2). The left graph compares the risk neutral distributions derived respectively from the GBMSJ model with ML kernel, GBMSJ model with Expo. kernel and the GBM model. We observe a wide difference between the Gaussian distribution shape of the GBM model and distribution under the GBMSJ model with ML kernel. This difference is explained by a high value of the kurtosis and the skewness of the GBMSJ model with ML kernel. A similar finding is made when we compare the distribution of the MJD and the KJD to the one of the GBMSJ model with ML kernel. When comparing the log-return's distribution under the GBMSJ model with ML kernel to the one with Expo. kernel, this difference is especially at the level of the distribution tails. Taking into account past jumps on the S&P 500 prices thickens the log-return distributions tails. In the right graph, we examine the effect of the time horizon on the log-return distribution. It turns on that the tails of the distribution become thicker. So that, as the maturity increases we expect positive returns, but large negative returns may also occur.

To inspect how recalling past jumps for a long period influences options prices, we evaluate call options on S&P 500 with parameters of Table 4.2 and invert the Black & Scholes formula to retrieve the smile of volatilities. The surface of implied volatilities for different maturities and moneyness is plotted in the left graph of Figure 4.10. The curvature of the smile is more pronounced for short term maturities than for long term ones. The right graph illustrates the impact of a reducing memory of past jumps on the volatility. Remembering past jumps for a long period increases implied volatilities. Its

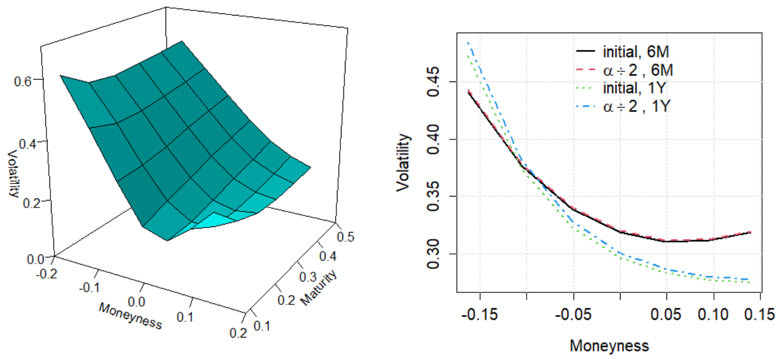


Figure 4.10: The left graph shows the surface of implied volatilities for call options on S&P 500. The right graph illustrates the impact of a modification of the memory parameter on the 6 months and 1 year volatility smiles. The parameters used for this illustration are in Table 4.2.

influence is limited at the short term and increases with the maturity.

4.4 Conclusions

In this chapter, a self-exciting jump diffusion process with long memory is designed for asset prices modelling. The self-excitation and the memory features are introduced via a Hawkes process (HP) with a Mittag-Leffler (ML) kernel. In addition to replicating the jump clustering phenomenon observed in stock markets, this setting allows to remember past jumps on asset prices for longer period than in the standard exponential kernel framework.

Although our HP with ML kernel is non-stationary and a non-Markov process, our ML kernel is the Laplace transform of a positive random measure. This allowed us to establish a closed form expression of the first order moment of the HP's intensity and rewrite the HP with ML kernel as an infinite dimensional Markov process. By a discretisation scheme and passing to the limit, we also derived a closed form expression for the conditional moment generating function of log-returns. To price option on asset, we found a family of affine changes of measure that preserves the dynamics of prices under the risk neutral measure.

An estimation procedure based on a Peak-Over-Threshold approach is applied to a time series of S&P 500 daily values from 05$^{\text{th}}$ May 2010 to 30$^{\text{th}}$June 2021. The estimated parameters of our model enabled us to show that long memory self-exciting jump diffusion process produces excess of kurtosis and skewness compared to model with exponential kernel, the Geometric Brownian Motion process, the Merton or Kou jump diffusion processes. Further, we showed that implied volatilities increase as the memory expands. But this influence is reduced at the short term.

From a broader perspective, it would be interesting to extend this chapter to the stochastic modelling of volatility where the memory feature is shown to be consitent with empirical findings reported in the literature on stock returns. Other possible application of HPs with ML kernel could be find in credit risk modelling with intensity based approach. From a theoretical point of view, it would be also interesting to investigate an approach to compute the second order statistics of our non-stationary HP.

Appendices

4.A Appendix A: Merton and Kou jump diffusion models

We consider that the Merton (1976) and the Kou (2002) jump diffusion models for the log-return are solution of the following SDE:

$$\mathrm{d}X_t = \left(\mu - \frac{\sigma^2}{2} - \mathbb{E}\left(e^J - 1\right)\lambda\right)\mathrm{d}t + \sigma \mathrm{d}W_t + \mathrm{d}\sum_{i=1}^{N_t} J_i, \tag{4.83}$$

where $\mathrm{d}N_t \sim$ follows a Poisson random variable with mean $\lambda \mathrm{d}t$. For the Merton jump diffusion (MJD) model, the jump sizes J_i are independent and identically normal distributed with mean μ_J and variance σ_J^2. Whereas, for the Kou jump diffusion (KJD) model, the jump sizes J_i are independent and identically distributed double exponential random variables with density $\nu(z)$ defined in equation (4.40).
From equation (4.83), we deduce that for all small length of time Δ,the probability density function of the log-return is given by:

$$\begin{aligned} f_{X_\Delta}(x) &= \lambda\Delta f_{\mathcal{N}\left(\left(\mu-\frac{\sigma^2}{2}-\mathbb{E}(e^J-1)\lambda\right)\Delta,\sigma^2\Delta\right)+J}(x) \\ &\quad + (1-\lambda\Delta) f_{\mathcal{N}\left(\left(\mu-\frac{\sigma^2}{2}-\mathbb{E}(e^J-1)\lambda\right)\Delta,\sigma^2\Delta\right)}(x) + o(\Delta), \end{aligned}$$

where o is such that $\lim_{\Delta\to 0}\frac{o(\Delta)}{\Delta} = 0$. J follows a normal distribution with mean μ_J and variance σ_J^2 in the MJD model. In the KJD model, J has a density $\nu(.)$. This derive from the following Poisson distribution definition

$$\begin{aligned} P\left(N_{t+h} - N_t = m | \mathscr{F}_t\right) &= \frac{(\lambda\Delta)^m e^{-\lambda\Delta}}{m!} \\ &= \begin{cases} \lambda\Delta + o(\Delta) & \text{if } m = 1 \\ o(\Delta) & \text{if } m > 1 \\ 1 - \lambda\Delta + o(\Delta) & \text{if } m = 0 \end{cases}. \end{aligned}$$

From a time series of m log-returns $x_1, ..., x_m$, equally spaced by a lag Δ of one day of trading, we may find estimated parameters of these two two jump diffusion models by maximization of the following log-likelihood function:

$$\ln l(x_1,..,x_m) = \sum_{i=1}^{m} \ln\left(f_{X_\Delta}(x_i)\right).$$

The conditional moment generation function of the log-return process at time $T \geq t$ is given by:

$$\mathbb{E}\left(e^{\omega X_T} \mid \mathscr{F}_t\right) = e^{G(\omega)(T-t)},$$

where the function G is defined by

$$G(\omega) = \omega\left(\mu - \frac{\sigma^2}{2} - \mathbb{E}\left(e^J - 1\right)\lambda\right) + \frac{\omega^2\sigma^2}{2} + \lambda\left(\mathbb{E}\left(e^{\omega J}\right) - 1\right),$$

with

$$\mathbb{E}\left(e^{\omega J}\right) = \begin{cases} e^{\omega\mu_J + \frac{\omega^2\sigma_J^2}{2}} & \text{if } J \sim \mathcal{N}\left(\mu_J, \sigma_J^2\right) \\ \frac{p\rho^+}{\rho^+ - \omega} + \frac{(1-p)\rho^-}{\rho^- - \omega} & \text{if } J \text{ has density } \nu(.) \end{cases}.$$

Chapter 5

Affine Heston model style with self-exciting jumps and long memory

5.1 Introduction

Over the past decade, a number of stochastic processes designed for option pricing have been introduced in the financial literature since the seminal work of Black and Scholes (1973). A common way to communicate option prices consists in computing the implied volatility (IV) resulting from the inversion of the Black and Scholes option price formula. IVs measures the uncertainty about how low or high an underlying asset might fall or rise over a given period. Unlike the historical asset's volatility, IVs are somehow expectations for future volatility. As a result, IVs play an important role in financial markets and option pricing models aim to replicate at least the main empirical features of asset time series. For a wide set of financial assets, Cont (2001) stated five stylized statistical facts about the asset's volatility. First, high volatility events tend to cluster in time. Second, there is a long memory effect in volatility which occurs when the effects of volatility shocks decay slowly. Third, the asset's volatility is negatively correlated with the asset's return. Fourth, trading volume is correlated with the asset's volatility. Fifth, there is an asymmetry in time scales of the volatility which is also known as the time reversal asymmetry of financial time series data. This has been emphasized by Zumbach and Lynch (2001) and called the "Zumbach effect" by Blanc et al. (2016). Beside these empirical findings in asset's volatility, common features of the IV surfaces are the volatility term structure effects, the at-the-money skew and the smile (see e.g. Aït-Sahalia and Lo (1998) Aït-Sahalia et al. (2021)). Designing an option pricing model that includes most of these characteristics is a challenging task.

In quest of the suitable option pricing model, this chapter extends previous works on affine stochastic volatility (SV) models stemming from the Black and Scholes framework. While simple and easy to use, the Black and Scholes model relies on strong assumptions regarding the financial behavior observed throughout market data. Merton (1976) and Bates (1991) relax the assumption of the Black and Scholes (1973) that the asset's return is normally distributed by inserting a jump component driven by a

Poisson process. Later, Heston (1993) proposes a closed form solution for an option on an asset with a SV model, which allows to relax the assumption of constant variance of the Black and Scholes model. He also introduces a correlation between the variance and the asset price in order to replicate the volatility smile and to explain the skewness of the asset's return. The Heston's framework is extended in several directions. Among them, we can cite stochastic volatility jump (SVJ) models, introduced by Bakshi et al. (1997), Duffie D. (2000), Bates (2000) and Eraker (2004) to mention a few pioneering works. The common point of these models, known as affine stochastic volatility models is the heteroskedastic property of the instantaneous change in variance through the square root term of the variance in the diffusion volatility. However, in the literature, other authors like Christoffersen et al. (2006) and Ignatieva et al. (2009) have studied non affine stochastic volatility models in which the instantaneous change in variance is proportional to any power of the variance in the diffusion volatility. They conclude that the choice between affine and non affine stochastic volatility models corresponds to find the right balance between fit flexibility and estimation properties. This motivate us to employ an affine SV model style for its tractability and economic interpretation documented in the literature (see e.g. Heston (1993), Bakshi et al. (1997), Duffie D. (2000)).

Notice that in view of expanding affine SVJ models, some authors like Gatheral (2012) investigated stochastic volatility models with jumps in asset prices and in the variance, denoted by SVJJ model. Although, this setting enables to maintain the clustering property of SV models, SVJJ models are less parsimonious than SVJ models, and are harder to fit to observed option prices. Another popular family of SV models considers that the log volatility behaves as a fractional Brownian motion with a small Hurst exponent. Among seminal contributions about this family of SV models we list the fractional SV model of Comte et al. (2010), and the rough volatility model of Gatheral et al. (2014) that in contrast to fractional SV, has a Hurst exponent of order less than 0.5. Even if these models are remarkably consistent with major features of financial time series, as stated by Blanc et al. (2016), they fail to induce a tail of distribution in the return fat enough even with fluctuating volatility since the asset's return in these models is conditionally Gaussian distributed. Moreover, these models are unable to explain in a realistic way how they generate fat tails, long memory in volatility and volatility clustering.

Nevertheless, most of affine SV models do not allow for the Zumbach effect. This motivates, authors like Zumbach et al. (2013), Jaisson and Rosenbaum (2015) and Blanc et al. (2016) to modify SV models in order to integrate this Zumbach effect. The two last works are based on quadratic Hawkes processes (HPs) which extend the self-exciting mechanism of HPs introduced by Hawkes (1971 a). In particular, Blanc et al. (2016) reformulate the general HP in a way that intra-day price changes, have an intensity proportional to powers of past price variations. Motivated by the empirical results, authors truncate this intensity to the second exponent and ignore the first exponent term. However, recent work of Hainaut and Moraux (2018 a), Hainaut and Moraux (2018 b) and Njike and Hainaut (2022) have shown the role of linear HP when it comes to replicate features of asset's return time series. In combination with the self-excitation property of the linear HP, Hainaut (2021) has designed a non-markov linear HP for claims process that has memory. Author has generated this memory effect by considering a linear HP with a kernel function that admits a Fourier's transform representation. Althougth, the HP with such long memory kernels are non-Markov, the author employed a similar procedure of the rough volatility literature (see e.g. Abi Jaber and El Euch (2019),Bäuerle and Desmettre (2020)) to recover a Markov structure and derive the conditional moment generating function of the Hawkes intensity. In this chapter, it was of interest to explore the benefit of such a setup in an option pricing model. In doing so, the linear HP will increase the probability to observe a jump of the asset price and therefore the future volatility after a jump occurrence, whilst the memory kernel of the linear HP will enable us to obtain a longer effect of past prices into future prices.

This chapters contributes to the financial literature in three ways. First, we introduce an option pricing model where the underlying asset price dynamics is self-excited and allows for a longer effect of past prices through a linear HP. Second, although our model is non Markov, we derive a closed-form expression for European call option prices. For hedging applications, Greeks of the three sources of risk: price risk, volatility risk and jump risk are obtained straightforward by taking derivatives of our formula with respect to the asset price, the variance or the arrival rate of jump. Third, we provide a calibration procedure of our model to real market option data.

The rest of the chapter proceeds as follows. Section 5.2 presents the affine Heston model style with self-exciting jumps and long memory. This section focuses on distributional properties of the arrival rate of jumps and the log-return through an infinite dimensional Markov representation of our asset price model and a discretisation scheme. Section 5.3 develops the option pricing model under our setting. Section 5.4 provides a descriptions of the Euro Stoxx 50 option data. Therein, we present our estimation procedure, discuss the estimated parameter and assess the sensitivity of option prices to memory and self-excitation parameters, underlying price, volatility and jump risks. Concluding comments are developped in Section 5.5.

5.2 Model specification

We define all the processes on a probability space Ω endowed with a risk neutral measure $\mathbb{Q}$ and a filtration $(\mathscr{F}_t)_{t\geq 0}$. We exclusively focus in the option pricing, and assumes that the pricing measure can be retrieve from European option prices. For sake of simplicity we assume a constant interest rate r.

Hereafter we consider a SVJ model with self-excited jump in return driven by a

Hawkes process with memory kernel:

$$\begin{aligned}\frac{dS_t}{S_t} &= r\mathrm{d}t + \sqrt{V_t}\left(\rho \mathrm{d}W_t^1 + \sqrt{1-\rho^2}\mathrm{d}W_t^2\right) \\ &\quad + \mathrm{d}\left(\sum_{l=1}^{N_t}\left(e^{J_l}-1\right)\right) - \lambda_t \mathbb{E}\left(e^J - 1\right)\mathrm{d}t \\ \mathrm{d}V_t &= \left(\theta_V - k_V V_t\right)\mathrm{d}t + \sigma_V \sqrt{V_t}\mathrm{d}W_t^1\end{aligned} \tag{5.1}$$

where k_V, $\frac{\theta_V}{k_V}$ and σ_V are respectively the speed of adjustment, long-run and variation coefficients of the diffusion volatility of V_t. W_t^1 and W_t^2 are two independent standard Brownian motions. ρ is the correlation between diffusion volatility of stock prices and variances. The variance process is the square root process introduced by Cox et al. (1985) and referred to as CIR. To ensure that zero is an unattainable for the diffusion volatility V_t, we assume the Feller condition $2\theta_V > \sigma_V^2$ holds. Let recall that affine stochastic model implies that the instantaneous change in variance is heteroskedastic via the $\sqrt{V_t}$ term in the diffusion volatility. We introduce memory and contagion in the dynamic of stock prices by considering an arrival intensity λ_t driven by the following equation:

$$\begin{aligned}\lambda_t|\mathscr{F}_t &= \lim_{h\to 0}\frac{\mathbb{E}\left(N_{t+h}|\mathscr{F}_t\right) - N_t}{h} \\ &= \alpha + (\lambda_0 - \alpha)\, g(t) + \eta \int_0^t g(t-u)\,\mathrm{d}L_u,\end{aligned} \tag{5.2}$$

with α the constant reversion level, λ_0 the initial intensity at time $t=0$, $L_t = \sum_{l=1}^{N_t}|J_l|$ the sum of the absolute values of the jump sizes up to time t. The intensity jumps by $\eta|J_l|$ when the l^{th} shock in the asset's returns dynamics occurs. Therefore, the occurrence of an event increases the probability of a further shock. The jump sizes J are double exponentially identically and independently distributed with density $\nu(z)$ on $\mathbb{R}$:

$$\nu(z) = p\rho^+ \exp\left(-\rho^+ z\right) 1_{\{z\geq 0\}} - (1-p)\,\rho^- \exp\left(-\rho^- z\right) 1_{\{z<0\}}.$$

Using the previous equation, it follows that the moment generating function (mgf) $\psi(.,.)$ of the pair $(J,|J|)$ is given by :

$$\psi(z_1,z_2) := \mathbb{E}\left(e^{z_1 J + z_2|J|}\right) = \frac{p\rho^+}{\rho^+ - (z_1+z_2)} + \frac{(1-p)\,\rho^-}{\rho^- - (z_1 - z_2)}.$$

In particular, the expectation of J and $|J|$ are weighted sum of the average sizes of positive and negative shocks :

$$\mathbb{E}(J) = p\frac{1}{\rho^+} + (1-p)\frac{1}{\rho^-}, \tag{5.3}$$

$$\mathbb{E}(|J|) = p\frac{1}{\rho^+} - (1-p)\frac{1}{\rho^-}. \tag{5.4}$$

We consider double exponential jumps as in Kou (2002) in order to replicate the asymmetry of the log-return distribution. Note that our SVJ model turns back to the Heston model if the jump size is null. In addition, we have the Bates model when the number of jumps is a Poisson process with a constant jump arrival rate λ, and normally distributed jump sizes with mean μ_J and variance σ_J^2.

In equation (5.2), the memory kernel function g, is a positive definite function such that $g(0) = 1$ and $\lim_{t\to\infty} g(t) = 0$. In this setting, the speed of reversion depends upon the decay rate of $g(.)$. The function $g(.)$ is the kernel function defining the memory of the stock prices process. If it is fast decreasing, the contagion effect between stock price jumps is limited in time. Otherwise, the processes remember the history of stock price jumps for a long period. In this chapter, we consider four functions $g(.)$ that are Fourier transform of positives measure on $\mathbb{R}$:

- The first function $g_L(.)$ is the Fourier transform of the Laplace or double exponential measure denoted by $v_L(.)$. Let $\beta \in \mathbb{R}^+$ be constant. The symmetric exponential measure $v_L(.)$ defined by the following relation :

$$v_L(\mathrm{d}u) = \frac{1}{2}\left(\beta\exp(-\beta u)\,1_{\{u\geq 0\}} + \beta\exp(\beta u)\,1_{\{u<0\}}\right)\mathrm{d}u, \tag{5.5}$$

  generates the following memory kernel:

$$g_L(h) = \int_{\mathbb{R}} v_L(u)\,e^{ihu}\mathrm{d}u = \frac{\beta^2}{\beta^2+h^2}. \tag{5.6}$$

  The mean of $v_L(.)$ is $\mu_L = 0$ and its variance is $\sigma_L^2 = \frac{2}{\beta^2}$. $g_L(.)$ is a power decreasing function. Since $\int_0^\infty g_L(u)\,\mathrm{d}u = \frac{\pi\beta}{2}$, in this case the pair (N,λ) is a stationary HP if $\beta < \frac{2}{\pi\eta\mathbb{E}(|J|)}$.

- The second function $g_G(.)$ is the Fourier transform of a Gaussian spectral measure $v_G(.)$

$$v_G(\mathrm{d}u) = \sqrt{\frac{\beta}{\pi}}e^{-\beta u^2}\mathrm{d}u. \tag{5.7}$$

  The mean of $v_G(.)$ is $\mu_G = 0$ and its variance is $\sigma_G^2 = \frac{1}{2\beta}$. Given that $v_G(.)$ is symmetric and $\int_{\mathbb{R}^+} e^{-\beta u^2}\cos(hu)\,\mathrm{d}u = \frac{1}{2}\sqrt{\frac{\pi}{\beta}}e^{-\frac{h^2}{4\beta}}$, the memory kernel is equal to

$$g_G(h) = e^{-\frac{h^2}{4\beta}}, \tag{5.8}$$

  where $\beta \in \mathbb{R}^+$. Since $\int_0^\infty g_G(u)\,\mathrm{d}u = \sqrt{\pi\beta}$, in this case the pair (N,λ) is a stationary HP if $\sqrt{\beta} < \frac{1}{\sqrt{\pi}\eta\mathbb{E}(|J|)}$.

- The third function $g_{Log}(.)$ is the Fourier transform of the logistic distribution defined as follows:

$$v_{Log}(\mathrm{d}u) = \frac{e^{-\frac{u}{\beta}}}{\beta\left(1+e^{-\frac{u}{\beta}}\right)^2}\mathrm{d}u, \tag{5.9}$$

where $\beta \in \mathbb{R}^+$. The mean of $v_{Log}(.)$ is $\mu_{Log} = 0$ and its variance is $\sigma^2_{Log} = \frac{\beta^2\pi^2}{3}$. The memory kernel is equal to

$$g_G(h) = \begin{cases} \frac{\pi\beta h}{\sinh(\pi\beta h)} & \text{if } h > 0 \\ 1 & \text{if } h = 0 \end{cases}. \tag{5.10}$$

Since $\int_0^\infty g_{Log}(u)\,\mathrm{d}u = \frac{\pi}{4\beta}$, in this case the pair (N,λ) is a stationary HP if $\beta > \frac{\pi\eta\mathbb{E}(|J|)}{4}$.

- The last memory kernel $g_C(.)$ is the Fourier transform of a Cauchy spectral measure $v_C(.)$

$$v_C(\mathrm{d}u) = \frac{1}{\pi}\frac{\beta}{\beta^2+u^2}\mathrm{d}u, \tag{5.11}$$

where $\beta \in \mathbb{R}^+$and the moments of $v_C(.)$ are undefined. Given that this measure is symmetric, the memory kernel $g_C(.)$ is an exponentially decreasing function

$$g_C(h) = e^{-\beta|h|}, \tag{5.12}$$

considered in the following as the exponential kernel because the delays after the arrival of jumps are non-negative. Since $\int_0^\infty g_C(u)\,\mathrm{d}u = \frac{1}{\beta}$, in this case the pair (N,λ) is a stationary HP if $\beta > \eta\mathbb{E}(|J|)$.

Within the literature, the Hawkes process with exponential kernel has an intensity λ_t usually defined as:

$$\begin{aligned} \lambda_t|\mathscr{F}_t &= \lim_{h\to 0}\frac{\mathbb{E}\left(N_{t+h}|\mathscr{F}_t\right)-N_t}{h} \\ &= \alpha+(\lambda_0-\alpha)e^{-\beta t}+\eta\int_0^t e^{-\beta(t-u)}\mathrm{d}L_u, \end{aligned} \tag{5.13}$$

where β is the constant rate of mean reversion. From Hawkes (1971 b), the pair (N,λ) is a stationary Hawkes process if the following inequality holds

$$\eta\mathbb{E}(|J|)\int_0^\infty e^{-\beta u}\mathrm{d}u < 1.$$

This latter condition is then met by choosing the intensity parameters η and β such as $\eta\mathbb{E}(|J|) < \beta$. This stationarity property of the Hawkes process is required to ensure that the long-term expectation, noted $\lambda_\infty = \lim_{t\to\infty}\mathbb{E}(\lambda_t)$ is finite and equal to $\lambda_\infty = \frac{\beta\alpha}{\beta-\eta\mathbb{E}(|J|)}$.

The first order moment of the HP intensity for the kernel functions in equations (5.6), (5.8) and (5.10) can be obtained by inverse Fourier transform. In fact, from equation (5.2), we use the properties of the expectation and apply the Fourier transform to obtain:

$$\begin{aligned} \mathscr{F}\left(\mathbb{E}(\lambda_t \mid \mathscr{F}_0)\right)(\omega) &= \alpha\mathscr{F}(1)(\omega)+(\lambda_0-\alpha)\mathscr{F}(g(t))(\omega) \\ &\quad +\eta\mathbb{E}(|J|)\mathscr{F}\left(\int_0^t g(t-u)\mathbb{E}(\lambda_u \mid \mathscr{F}_0)\,\mathrm{d}u\right)(\omega) \\ &= 2\alpha\pi\delta(\omega)+(\lambda_0-\alpha)v_k(\omega) \\ &\quad +\eta\mathbb{E}(|J|)v_k(\omega)\mathscr{F}\left(\mathbb{E}(\lambda_t \mid \mathscr{F}_0)\,\mathrm{d}u\right)(\omega) \end{aligned} \tag{5.14}$$

where $\delta(.)$ is the Dirac Delta function. Then, given that

$$\begin{aligned}\mathscr{F}\left(\mathbb{E}\left(\lambda_t \mid \mathscr{F}_0\right)\right)(\omega) &= \frac{2\alpha\pi\delta(\omega)+(\lambda_0-\alpha)\nu_k(\omega)}{1-\eta\mathbb{E}(|J|)\nu_k(\omega)} \\ &= \frac{2\alpha\pi\delta(\omega)}{1-\eta\mathbb{E}(|J|)\nu_k(\omega)}+\frac{(\lambda_0-\alpha)\nu_k(\omega)}{1-\eta\mathbb{E}(|J|)\nu_k(\omega)}\end{aligned} \tag{5.15}$$

it follows that

$$\begin{aligned}\mathbb{E}\left(\lambda_t \mid \mathscr{F}_0\right) &= \alpha\mathscr{F}^{-1}\left(\frac{2\pi\delta(\omega)}{1-\eta\mathbb{E}(|J|)\nu_k(\omega)}\right)(t)+(\lambda_0-\alpha) \\ &\quad \times\mathscr{F}^{-1}\left(\frac{\nu_k(\omega)}{1-\eta\mathbb{E}(|J|)\nu_k(\omega)}\right)(t) \\ &= \alpha\frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\omega t}\frac{2\pi\delta(\omega)}{1-\eta\mathbb{E}(|J|)\nu_k(\omega)}\mathrm{d}\omega+(\lambda_0-\alpha) \\ &\quad \times\frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\omega t}\frac{\nu_k(\omega)}{1-\eta\mathbb{E}(|J|)\nu_k(\omega)}\mathrm{d}\omega \\ &= \frac{\alpha}{1-\eta\mathbb{E}(|J|)\nu_k(0)}+(\lambda_0-\alpha)\frac{1}{2\pi}\int_{\mathbb{R}} e^{-i\omega t}\frac{\nu_k(\omega)}{1-\eta\mathbb{E}(|J|)\nu_k(\omega)}\mathrm{d}\omega.\end{aligned} \tag{5.16}$$

The previous stability conditions ensure that this latter integral exists. Notice that there is no closed form expression of this expected value of the HP intensity, except for the HP intensity with exponential kernel. Through either the Laplace transform technique or ODE technique of Fonseca and Zaatour (2013), the first order moment of the HP intensity with exponential kernel is given by,

$$\mathbb{E}\left(\lambda_t \mid \mathscr{F}_0\right) = \lambda_0 e^{(\eta\mathbb{E}(|J|)-\beta)t}-\frac{\beta\alpha}{\eta\mathbb{E}(|J|)-\beta}\left(1-e^{(\eta\mathbb{E}(|J|)-\beta)t}\right). \tag{5.17}$$

To the best of our knowledge, it is not possible to derive analytically moments of HP with non-exponential kernel listed above using the ordinary differential equation (ODE) technique of Fonseca and Zaatour (2013). As an alternative, we will derive a quasi-analytic formula of the conditional moment generating function of these HP intensities with non-exponential kernel.

To compare shapes of memory kernel we range above memory kernels by degree of decay and consider that their spectral density have the same variance σ^2. It follows that the parameter β of memory kernels $(g_k)_{k\in\{L,G,Log\}}$ is a function of σ^2. As the variance of the Cauchy spectral measure is not defined, we find an β by minimizing the sum of least square between the Log and Cauchy memory kernels. Figure 5.1 compares these four memory kernels when the variance $\sigma^2 = 16$. The left plot reveals that the Laplace kernel has a slower decay than Gaussian and logistic kernels. Further, in long run the exponential kernel seems to decay more slowly than the Gaussian and logistic kernels, while the opposite situation occurs in the short run. The right graph shows that the decay rate of the kernel increases with the thickness of the distribution tails.

The common point of these four memory kernels is the stability of the Hawkes intensity λ_t under a certain condition. Unlike the exponential kernel, the other kernels can

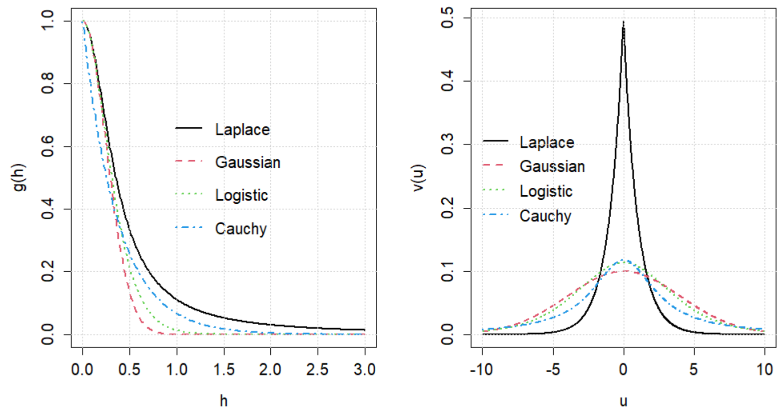


Figure 5.1: Comparison of Laplace, Gaussian and Logistic kernels for $\sigma_L^2 = \sigma_G^2 = \sigma_{Log}^2 = 16$. The Cauchy kernel is computed with $\beta_C = 2.71$.

enable the Hawkes intensity λ_t to remember past events for a long time according to their speed of decay. Considering the exponential kernel (5.12), the pair $(N_t, \lambda_t)_{t \geq 0}$ is a Markov process and the intensity λ_t is solution of the following stochastic differential equation (SDE):

$$d\lambda_t = \beta(\alpha - \lambda_t)\,dt + \eta\, dL_t. \tag{5.18}$$

Taking into account kernel functions in equations (5.6), (5.8) and (5.10), the pair $(N_s, \lambda_s)_{s \geq 0}$ is not a Markov process and we have very little information on its behavior conditionally to the filtration $\mathscr{F}_t$ for $t \leq s$. For non-Markov processes we cannot rely on Itô calculus to determine quantity like the mgf of the arrival rate of shocks in the asset's returns dynamics. However, by using the spectral representation of kernel $g_k(.)$ for $k \in \{L, G, Log\}$, we reformulate $(N_s, \lambda_s)_{s \geq 0}$ as an infinite dimensional Markov process in the complex plane. To do so, we rewrite $g_k(.)$ in Equation (5.2) as a Fourier transform of the measure $\nu_k(.)$ and change the order of integration:

$$\begin{aligned} \lambda_t &= \alpha + (\lambda_0 - \alpha)\, g_k(t) + \eta \int_0^t \int_{\mathbb{R}} e^{i(t-u)\xi} \nu_i(d\xi)\, dL_u \\ &= \alpha + (\lambda_0 - \alpha)\, g_k(t) + \eta \int_{\mathbb{R}} \left(\int_0^t e^{i(t-u)\xi} dL_u \right) \nu_k(d\xi). \end{aligned} \tag{5.19}$$

It follows that the process $Y_t^\xi = \int_0^t e^{i(t-u)\xi} dL_u$, defined on the complex plane $\mathbb{C}$ is solution of the following SDE:

$$dY_t^\xi = i\xi Y_t^\xi\, dt + dL_t, \tag{5.20}$$

and the SDE of the Hawkes intensity is given by

$$d\lambda_t = (\lambda_0 - \alpha)\frac{dg_k(t)}{dt} dt + \eta \int_{\mathbb{R}} dY_t^\xi \nu_k(d\xi). \tag{5.21}$$

Remark that $\int_{\mathbb{R}} \mathrm{d}Y_t^{\xi} \nu_k(\mathrm{d}\xi)$ is real since the measure $\nu_k(.)$ is symmetric and by rewriting Y_t^{ξ} as $\sum_{l=1}^{N_t^l} e^{i(t-u_l)} |J_l|$ with u_l the time of the l^{th} shock in the asset's returns dynamics.

To benefit from the reformulation of the process $(N_t, \lambda_t)_{t\geq 0}$ like an infinite dimensional Markov process $\left(N_t, \lambda_t, \left(Y_t^{\xi}\right)_{\xi\in\mathbb{R}}\right)_{t\geq 0}$, we use the discretisation scheme described below, and passing to the limit we infer the mgf of the arrival rate of shocks in the asset's returns dynamics. The key point of this discretisation scheme consists of an approximation of the measure $\nu_k(.)$ as a discrete measures with a finite numbers of atoms. For this purpose, we consider a symmetric partition $\mathscr{E}^{(n)} := \{-\infty < \xi_{-n}^{(n)} < ... < \xi_0^{(n)} = 0 < \xi_1^{(n)} < ... < \xi_n^{(n)} < \infty\}$, with $\xi_{-n}^{(n)} = -\xi_n^{(n)}$. For each interval $(\xi_l^{(n)}, \xi_{l+1}^{(n)})$, the barycenter is defined as

$$\begin{aligned} b_{l+1}^{(n)} &= \frac{\xi_l^{(n)} + \xi_{l+1}^{(n)}}{2}, \text{ for } l \in \{0, ..., n-1\}, \\ b_l^{(n)} &= \frac{\xi_l^{(n)} + \xi_{l+1}^{(n)}}{2}, \text{ for } l \in \{-n, ..., -1\}, \end{aligned} \tag{5.22}$$

while the mass of any atom with respect of the measure $\nu_k(.)$ is

$$\begin{aligned} m_{l+1}^{(k,n)} &= \int_{\xi_l^{(n)}}^{\xi_{l+1}^{(n)}} \nu_k(\mathrm{d}u), \text{ for } l \in \{0, ..., n-1\}, \\ m_l^{(k,n)} &= \int_{\xi_l^{(n)}}^{\xi_{l+1}^{(n)}} \nu_k(\mathrm{d}u), \text{ for } l \in \{-n, ..., -1\}. \end{aligned} \tag{5.23}$$

By construction $\lim_{n\to\infty} \sum_{l=-n}^{n} m_l^{(k,n)} = 1$ and the discrete approximation of the measure $\nu_k(.)$ given a partition of size n, denoted by $\tilde{\nu}_k^{(n)}(.)$ is defined as follows

$$\tilde{\nu}_k^{(n)}(u) = \sum_{l=-n}^{n} m_l^{(k,n)} \delta_{b_l^{(n)}}(u), \tag{5.24}$$

where $\delta_{b_l^{(n)}}(u)$ is the Dirac measure located at the barycenter $b_l^{(n)}$. According to Equations (5.22) and (5.23), and given that the measure $\nu_k(.)$ is symmetric,
for $l \in \{-n, ..., -n\} \setminus \{0\}$, $m_{-l}^{(k,n)} = m_l^{(k,n)}$ and $b_{-l}^{(n)} = -b_l^{(n)}$. It follows that the discrete measure $\tilde{\nu}_k^{(n)}(.)$ is also symmetric. We consider that the following assumptions holds for the partition $\mathscr{E}^{(n)}$:

(i) $\xi_{-n}^{(n)} \to -\infty$ and $\xi_n^{(n)} \to \infty$ when $n \to \infty$.

(ii) $\max |\xi_{l+1}^{(n)} - \xi_l^{(n)}| \to 0$ when $n \to \infty$.

(iii) $\mathscr{E}^{(n)} \subset \mathscr{E}^{(n+1)}$.

In this setting, for any integrable function, $f(.)$ with respect to the measure $v_k(.)$, we have that

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} f(u)\,\tilde{v}_k^{(n)}(\mathrm{d}u) = \int_{-\infty}^{\infty} f(u)\,v_k(\mathrm{d}u). \tag{5.25}$$

To lighten further developments, we choose n and adopt the following notations:

- $Y_t^{(l)} := Y_t^{\left(b_l^{(n)}\right)}$ for $l \in \{-n,...,n\}$ and $Y_t^{(l)}$ is solution of the following SDE

$$\mathrm{d}Y_t^{(l)} = ib_l^{(n)} Y_t^{(l)}\mathrm{d}t + \mathrm{d}L_t^{(n)}, \tag{5.26}$$

  with $L_t^{(n)} = \sum_{l=1}^{N_t^{(n)}} |J_l|$;

- the approximated intensity in the discrete framework is

$$\lambda_t^{(n)} = \alpha + (\lambda_0 - \alpha)\,g_k(t) + \eta \sum_{l=-n}^{n} m_l^{(k,n)} Y_t^{(l)} \tag{5.27}$$

  and has the following SDE

$$\mathrm{d}\lambda_t^{(n)} = (\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t}\mathrm{d}t + \eta\sum_{l=-n}^{n} m_l^{(k,n)} ib_l^{(n)} Y_t^{(l)}\mathrm{d}t + \eta\sum_{l=-n}^{n} m_l^{(k,n)}\mathrm{d}L_t^{(n)}. \tag{5.28}$$

Proposition 11. *The moment generating function of $\lambda_s^{(n)}$ conditionally to the information at time $t \le s$ is given by*

$$\begin{aligned}\mathbb{E}\left(e^{\omega\lambda_s^{(n)}}|\mathscr{F}_t\right) &= \exp\Big(F(t,s,\omega) + G(t,s,\omega)\,\lambda_t^{(n)} \\ &\quad + \sum_{l=-n}^{n} H_l\left(t,s,b_l^{(n)},\omega\right) m_l^{(k,n)} Y_t^{(l)}\Big)\end{aligned}$$

for all $\omega \in \mathbb{C}_-$ that satisfies the regularity condition $\psi\left(0,\eta\omega\sum_{l=-n}^{n} m_l^{(k,n)}\right) < \infty$. The functions H_l are defined as follows

$$H_l\left(t,s,b_l^{(n)},\omega\right) = b_l^{(n)}\eta i\int_t^s e^{ib_l^{(n)}(u-t)}G(u,s,\omega)\,du,\ \forall l \in \{-n,...,n\},$$

F and G are time dependent functions that solve the following ODEs system

$$\begin{aligned}\frac{\partial F}{\partial t}(t,s,\omega) &= -(\lambda_0-\alpha)\frac{dg_k(t)}{dt}G(t,s,\omega) \\ \frac{\partial G}{\partial t}(t,s,\omega) &= 1-\psi\Big(0,\eta G(t,s,\omega)\sum_{l=-n}^{n} m_l^{(k,n)} \\ &\quad + \sum_{l=-n}^{n} H_l\left(t,s,b_l^{(n)},\omega\right) m_l^{(k,n)}\Big).\end{aligned} \tag{5.29}$$

with final conditions $F(s,s,\omega)=0$, $G(s,s,\omega)=\omega$. ψ is the moment generating function of the pair of the random jump sizes J and its absolute value : $\psi(z_1,z_2) := \mathbb{E}\left(e^{z_1 J + z_2|J|}\right)$. $k \in \{L,G,Log,C\}$ indicates the type of HP memory kernel

Proof. In Appendix A 5.A. □

Passing to the limit with respect of n and using the frequency derivative property of Fourier transform, we have the following proposition on the conditional mgf of the HP intensity.

Proposition 12. *The moment generating function of λ_s conditionally to the information at time $t \leq s$ is given by*

$$\begin{aligned}\mathbb{E}\left(e^{\omega\lambda_s}|\mathscr{F}_t\right) &= \exp\left(F\left(t,s,\omega\right)+G\left(t,s,\omega\right)\lambda_t\right) \quad (5.30)\\ &\times\exp\left(i\eta\int_t^s G\left(u,s,\omega\right)\int_{\mathbb{R}} l e^{il(u-t)}Y_t^l v_k\left(l\right)dl\,du\right),\end{aligned}$$

for all $\omega \in \mathbb{C}_-$ that satisfies the regularity condition $\psi(0,\eta\omega) < \infty$. The functions F and G solve the following ODE and partial integro-differential equation (PIDE):

$$\frac{\partial F}{\partial t}\left(t,s,\omega\right) = -\left(\lambda_0-\alpha\right)\frac{dg_k\left(t\right)}{dt}G\left(t,s,\omega\right) \quad (5.31)$$

$$\begin{aligned}\frac{\partial G}{\partial t}\left(t,s,\omega\right) &= 1-\psi\Bigg(0,\eta G\left(t,s,\omega\right) \quad (5.32)\\ &+\eta\int_t^s G\left(u,s,\omega\right)\frac{dg_k\left(u-t\right)}{dt}du\Bigg)\end{aligned}$$

with final conditions $F(s,s,\omega) = 0$, $G(s,s,\omega) = \omega$. $k \in \{L,G,Log,C\}$ indicates the type of HP memory kernel.

Proof. The approximated intensity $\lambda_t^{(n)}$ defined in equation (5.27), converges almost surely to λ_t since for fixed $w \in \Omega$, the function $z \mapsto Y_t^{(z)}(w)$ is integrable with respect to the measure v_k and for any integrable function, $f(.)$ with respect to the measure $v_k(.)$, we have that

$$\lim_{n\to\infty}\int_{-\infty}^{\infty} f\left(u\right)\tilde{v}_k^{(n)}\left(\mathrm{d}u\right) = \int_{-\infty}^{\infty} f\left(u\right)v_k\left(\mathrm{d}u\right). \quad (5.33)$$

It follows that $N_t^{(n)} = \int_0^t \lambda_u^{(n)}\mathrm{d}u$ converges almost surely to N_t. For $l \in \{-n,...,n\}$, $Y_t^{(l)} = \int_0^t e^{-b_l^{(n)}(t-u)}\mathrm{d}L_u^{(n)}$. If the partition $\mathscr{E}^{(n)}$ satisfies the assumptions (i)-(iii), when n tends to ∞, a barycenter $b_l^{(n)}$ tends to be an element of $\mathbb{R}_+$. Hence the sequence $\left(Y_t^{\left(b_l^{(n)}\right)}\right)_{l=-n,...,n}$ tends to $\left(Y_t^{(l)}\right)_{l\in\mathbb{R}}$ almost surely. Next, the dominated convergence provides that

$$\mathbb{E}\left(e^{\omega\lambda_s u}|\mathscr{F}_t\right) = \lim_{n\to\infty}\mathbb{E}\left(e^{\omega\lambda_s^{(n)}}|\mathscr{F}_t\right),$$

where from proposition 11

$$\begin{aligned}\mathbb{E}\left(e^{\omega\lambda_s^{(n)}}|\mathscr{F}_t\right) &= \exp\Bigg(F\left(t,s,\omega\right)+G\left(t,s,\omega\right)\lambda_t^{(n)}\\ &+\sum_{l=-n}^{n} H_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}Y_t^{(l)}\Bigg)\end{aligned}$$

Figure 5.2: Expected value of the HP intensity $\lambda_t \mid \mathscr{F}_0$. For comparison purpose, the parameter β of the Laplace, the Gaussian and the Logistic kernels are such that $\sigma_L^2 = \sigma_G^2 = \sigma_{Log}^2 = 16$. While the Cauchy kernel is computed with $\beta_C = 2.71$. The other parameters are : $\alpha = 0.4$, $\eta = 0.2$, $\lambda_0 = 50$, $p = 0.4$, $\rho^+ = 30$ and $\rho^- = -37$.

and

$$\sum_{l=-n}^{n} H_l\left(t,s,b_l^{(n)},\omega\right) m_l^{(k,n)} = \sum_{l=-n}^{n} b_l^{(n)} \eta i \int_t^s e^{ib_l^{(n)}(u-t)} G(u,s,\omega)\,\mathrm{d}u\, m_l^{(k,n)}.$$

Taking into account the definition of the discretised measure $\tilde{\nu}_k^{(n)}$ (5.24) we obtain

$$\int_{-\infty}^{\infty} H_l(t,s,l,\omega)\,\tilde{\nu}_k^{(n)}(l)\,\mathrm{d}l = \eta i \int_{-\infty}^{\infty} \left(\int_t^s l e^{-il(u-t)} G(u,s,\omega)\,\mathrm{d}u\right) \tilde{\nu}_k^{(n)}(l)\,\mathrm{d}l.$$

Passing to the limit we have

$$\begin{aligned}
\int_{-\infty}^{\infty} H_l(t,s,l,\omega)\,\nu_k(l)\,\mathrm{d}l &= \eta i \int_{-\infty}^{\infty} \left(\int_t^s l e^{-il(u-t)} G(u,s,\omega)\,\mathrm{d}u\right) \nu_k(l)\,\mathrm{d}l \\
&= \eta i \int_t^s G(u,s,\omega) \int_{-\infty}^{\infty} l e^{-il(u-t)} \nu_k(l)\,\mathrm{d}l\mathrm{d}u \\
&= \eta \int_t^s G(u,s,\omega)\, \frac{\mathrm{d}g_k(u-t)}{\mathrm{d}t}\,\mathrm{d}u
\end{aligned}$$

The second and last lines of the latter equation stem from the Fubini's theorem and the frequency differentiation property of the Fourier transform. Likewise, we show that the limit of $\sum_{l=-n}^{n} H_l\left(t,s,b_l^{(n)},\omega\right) Y_t^{(l)} m_l^{(k,n)}$ is

$$\int_{-\infty}^{\infty} H_l(t,s,l,\omega)\, Y_t^l \nu_k(l)\,\mathrm{d}l = i\eta \int_t^s G(u,s,\omega) \int_{\mathbb{R}} l e^{il(u-t)} Y_t^l \nu_k(l)\,\mathrm{d}l\,\mathrm{d}u,$$

and concludes the proof. $\square$

Applying a finite difference approach to the mgf of the HP intensity in proposition 12, we draw the expected value of the HP intensity given the time in figure 5.2. From this figure, we conclude that the expected HP intensity decreases as time increases at the same rate of decay as the corresponding kernel function.

Let $X_t := \log S_t$ the log of the asset price at time t. Considering the discretisation scheme on the partition $\mathscr{E}^{(n)}$, the asset price noted $S_t^{(n)}$ is solution of the following PDE:

$$\begin{aligned}
\frac{\mathrm{d}S_t^{(n)}}{S_t^{(n)}} &= r\mathrm{d}t + \sqrt{V_t}\left(\rho \mathrm{d}W_t^1 + \sqrt{1-\rho^2}\mathrm{d}W_t^2\right) \\
&\quad + \mathrm{d}\left(\sum_{l=1}^{N_t^{(n)}} \left(e^{J_l} - 1\right)\right) - \lambda_t^{(n)} \mathbb{E}\left(e^J - 1\right)\mathrm{d}t.
\end{aligned} \tag{5.34}$$

By Itô formula the discretised version of the process X_t , noted $X_t^{(n)} := \log S_t^{(n)}$ admits the next SDE

$$\begin{aligned} \mathrm{d}X_t^{(n)} &= \left(r - \frac{1}{2}V_t - \lambda_t^{(n)}\mathbb{E}\left(e^J - 1\right)\right)\mathrm{d}t \qquad (5.35) \\ &\quad + \sqrt{V_t}\left(\rho \mathrm{d}W_t^1 + \sqrt{1-\rho^2}\mathrm{d}W_t^2\right) + \mathrm{d}\left(\sum_{l=1}^{N_t^{(n)}} J_l\right). \end{aligned}$$

The process $X_t^{(n)}$ is then a Markov process with respect to the intensity $\lambda_t^{(n)}$, the $2n+1$ factors $\left(Y_t^{(l)}\right)_{l=-n,\dots,n}$ and the counter $N_t^{(n)}$. We can therefore deal with pricing and hedging problems using standard methods developed for stochastic models. It is shown in the proof of proposition 12 that $\left(\lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right)$ converges almost surely to $\left(\lambda_t,\ \left(Y_t^l\right)_{l=-n,\dots,n},\ N_t\right)$. Consequently, the process $X_t^{(n)}$ converges almost surely to X_t. The next proposition establishes the characteristic function of $X_s^{(n)}$ conditionally to the information at time $t \le s$.

Proposition 13. *For all $\omega \in \mathbb{R}$ that satisfies the regularity condition*

$$\psi\left(i\omega, \eta \sum_{l=-n}^{n} m_l^{(k,n)}\right) < \infty,$$

the characteristic function of $X_s^{(n)}$ conditionally to the information at time $t \le s$ is given by

$$\begin{aligned} \Upsilon_{t,s}^{(n)}(\omega) &:= \mathbb{E}^Q\left(e^{i\omega X_s^{(n)}} \mid \mathscr{F}_t\right) \\ &= \exp\left(i\omega\left(X_t^{(n)} + r(s-t)\right)\right) \\ &\quad \times \exp\left(\theta_V \sigma_V^{-2}\left(k_V - i\omega\rho\sigma_V - \bar{d}\right)(s-t) - 2\log\left(\frac{1-\bar{g}e^{-\bar{d}(s-t)}}{1-\bar{g}}\right)\right) \\ &\quad \times \exp\left(V_t \sigma_V^{-2}\left(k_V - i\omega\rho\sigma_V - \bar{d}\right)\frac{1-e^{-\bar{d}(s-t)}}{1-\bar{g}e^{-\bar{d}(s-t)}}\right) \\ &\quad \times \exp\left(A_J(t,s,\omega) + \sum_{l=-n}^{n} C_l\left(t,s,b_l^{(n)},\omega\right) m_l^{(k,n)} Y_t^{(l)} + D(t,s,\omega)\lambda_t^{(n)}\right) \end{aligned}$$

where

$$\begin{aligned} \bar{d} &= \sqrt{\left(i\omega\rho\sigma_V - k_V\right)^2 + \sigma_V^2\left(i\omega + \omega^2\right)}, \\ \bar{g} &= \frac{k_V - i\omega\rho\sigma_V - \bar{d}}{\left(k_V - i\omega\rho\sigma_V + \bar{d}\right)}, \end{aligned}$$

$k \in \{L, G, Log, C\}$ indicates the type of HP memory kernel. For all $l \in \{-n, \dots, n\}$,

$$C_l\left(t,s,b_l^{(l)},\omega\right) = i b_l^{(n)} \eta \int_t^s e^{i b_l^{(n)}(u-t)} D(u,s,\omega)\,du. \qquad (5.36)$$

A_J and D are functions that solves the following ODEs system :

$$\begin{cases} \frac{\partial A_J}{\partial t}(t,s,\omega) & = -D(t,s,\omega)(\lambda_0-\alpha)\frac{dg_k(t)}{dt} \\ \frac{\partial D}{\partial t}(t,s,\omega) & = i\omega(\psi(1,0)-1) \\ & -\left(\psi\left(i\omega,\Sigma_{l=-n}^{n}C_l\left(t,s,b_l^{(l)},\omega\right)m_l^{(k,n)}+\eta D(t,s,\omega)\right)-1\right) \end{cases} \tag{5.37}$$

with the terminal condition $A_J(s,s,\omega)=0$ and $D(s,s,\omega)=0$.

Proof. In Appendix B 5.B. □

The previous proposition presents the characteristic function as the product of the exponent of an affine function related to the jump component and the characteristic function of the standard Heston model (see Schoutens et al. 2004 and Albrecher et al. (2006) for more details). Passing to the limit with respect of n , we have the following proposition on the characteristic function of the process X_s under the risk neutral measure $\mathbb{Q}$.

Proposition 14. *For all $\omega\in\mathbb{R}$ that satisfies the regularity condition $\psi(i\omega,\eta)<\infty$, the characteristic function of X_s conditionally to the information at time $t\le s$ is given by*

$$\begin{aligned} \Upsilon_{t,s}(\omega) &:= \mathbb{E}^{Q}\left(e^{i\omega X_s}\mid\mathscr{F}_t\right) \\ &= \exp\left(i\omega\left(X_t^{(n)}+r(s-t)\right)\right) \\ &\times\exp\left(\theta_V\sigma_V^{-2}\left(k_V-i\omega\rho\sigma_V-\bar d\right)(s-t)-2\log\left(\frac{1-\bar g e^{-\bar d(s-t)}}{1-\bar g}\right)\right) \\ &\times\exp\left(V_t\sigma_V^{-2}\left(k_V-i\omega\rho\sigma_V-\bar d\right)\frac{1-e^{-\bar d(s-t)}}{1-\bar g e^{-\bar d(s-t)}}\right) \\ &\times\exp\left(i\eta\int_t^s D(u,s,\omega)\int_{\mathbb{R}} l e^{il(u-t)}Y_t^l v_k(l)\,dl\,du+D(t,s,\omega)\lambda_t\right. \\ &\left.+A_J(t,s,\omega)\right) \end{aligned} \tag{5.38}$$

where

$$\begin{aligned} \bar d &= \sqrt{(i\omega\rho\sigma_V-k_V)^2+\sigma_V^2\left(i\omega+\omega^2\right)}, \\ \bar g &= \frac{k_V-i\omega\rho\sigma_V-\bar d}{\left(k_V-i\omega\rho\sigma_V+\bar d\right)}. \end{aligned}$$

$k\in\{L,G,Log,C\}$ indicates the type of HP memory kernel. A_J and D satisfy the next ODE and PIDE

$$\begin{cases} \frac{\partial A_J}{\partial t}(t,s,\omega) & = -D(t,s,\omega)(\lambda_0-\alpha)\frac{dg_k(t)}{dt} \\ \frac{\partial D}{\partial t}(t,s,\omega) & = i\omega(\psi(1,0)-1) \\ & -\left(\psi\left(i\omega,\eta\int_t^s\frac{dg_k}{dh}(u-t)D(u,s,\omega)du+\eta D(t,s,\omega)\right)-1\right) \end{cases} \tag{5.39}$$

with the terminal condition $A_J(s,s,\omega)=0$ and $D(s,s,\omega)=0$

Proof. This proof is similar to the one of proposition 12, as $X_s^{(n)}$ converges to X_s almost surely and by the dominated convergence the conditional characteristic function of the log-return process is equal to the limit with respect of n of the conditional characteristic function of the approximated log-return process. By some grouping and using the frequency derivative property of Fourier transform we retrieve the result of proposition 14. □

Our setting takes into account several features. It allows for stochastic evolution of the asset price's variance, includes a jump risk component into the asset price dynamics and replicate the jumps clustering. We also generalize some existing valuation models. For instance, we retrieve the Black and Scholes option price formula by setting $J_l = 0$ and $\theta_V = k_V = \sigma_V = 0$, the Heston option price formula by setting $J_l = 0$ and the Bates and Merton option price formula by setting $\alpha = \lambda_0$, $\eta = 0$ and the jump size J_l log-normally distributed.
For comparison purpose, figure 5.3 draws the density of the log-return at 50 days X_{50d} under the risk neutral measure for our four HP kernels. From these graphs, we conclude that the log-return density at 50 days with the HP Cauchy kernel has the thinnest tail. The log-return density at 50 days for ours extension of the Heston model with the Laplace, Gaussian and Logistic kernels look identical but the zoom in reveals the differences due to the decay rate of these kernel functions. Hence, the log-return density with the Laplace kernel which has a slower decay than the Gaussian and logistic ones, has the fattest tail.

5.3 Option Pricing Models

In this section, we develop the option pricing model under our modeling framework outlined in the previous section.
As many authors (see e.g. Heston (1993), Bakshi et al. (1997), Gatheral (2012)) we will express the call option price using the following delta probability decomposition of Black-Sholes formula (1973). For this aim, in the next proposition we introduce a convenient measure $\mathbb{Q}^\star$.

Proposition 15. *Let $\mathbb{Q}^\star$ an equivalent measures to $\mathbb{Q}$ on the measure space $(\Omega, \mathscr{F})$, and define by*

$$\begin{aligned}\left.\frac{d\mathbb{Q}^\star}{d\mathbb{Q}}\right|_{\mathscr{F}_0} &= \exp\left(-\frac{1}{2}\int_0^t V_u du + \int_0^t \sqrt{V_u}\left(\rho dW_u^1 + \sqrt{1-\rho^2} dW_u^2\right)\right) \\ &\quad \times \exp\left(-(\psi(1,0)-1)\int_0^t \lambda_u du + \sum_{l=1}^{N_u} J_l\right).\end{aligned}$$

Then, under the measure $\mathbb{Q}^\star$,
i) $W_t^{\star 1}$ and $W_t^{\star 2}$ are Brownian motions defined as follows:

$$dW_t^{\star 1} = dW_t^1 - \rho\sqrt{V_t}dt,$$

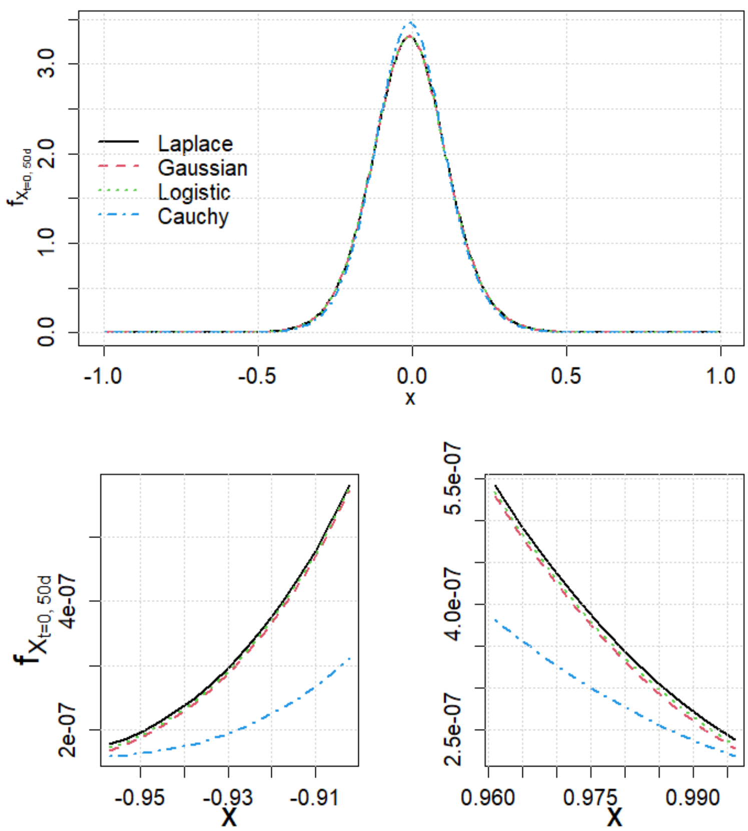


Figure 5.3: Distribution of $X_{50d} \mid \mathscr{F}_0$ under the risk neutral measure. For comparison purpose, the parameter β of the Laplace, the Gaussian and the Logistic kernels are such that $\sigma_L^2 = \sigma_G^2 = \sigma_{Log}^2 = 16$. While the Cauchy kernel is computed with $\alpha_C = 2.71$. The other parameters are : $S_0 = 1$, $r = 0.01$, $\theta_V = 0.4$, $k_V = 1$, $\rho = -0.7$, $\sigma_V = 0.01$, $v_0 = 0.01$, $\alpha = 0.4$, $\eta = 0.2$, $\lambda_0 = 50$, $p = 0.4$, $\rho^+ = 30$ and $\rho^- = -37$.

$$dW_t^{\star 2} \quad = \quad dW_t^2 - \sqrt{1-\rho^2}\sqrt{V_t}dt.$$

ii) $N_t^\star$ is a counting process with the following intensity

$$\lambda_t^\star \quad = \quad \psi(1,0)\,\lambda_t. \tag{5.40}$$

Considering the discretisation scheme on the partition $\mathscr{E}^{(n)}$, this intensity noted $\lambda_t^{(n),\star}$ is solution of next SDE

$$\begin{aligned} d\lambda_t^{(n),\star} \quad &= \quad \left(\lambda_0^\star - \alpha^\star\right)\frac{dg_k(t)}{dt}dt + \eta^\star \sum_{l=-n}^{n} m_l^{(k,n)} ib_l^{(n)} Y_t^{(l),\star}dt \\ &\quad + \eta^\star dL_t^{(n),\star}, \end{aligned} \tag{5.41}$$

where $\lambda_0^\star = \psi(1,0)\,\lambda_0$, $\eta^\star = \psi(1,0)\,\eta$, $L_t^{(n),\star} = \sum_{l=1}^{N_t^{(n),\star}} \left|J_l^\star\right|$ and for all $l \in \{-n,...,n\}$

$$dY_t^{(l),\star} = ib_l^{(n)} Y_t^{(l),\star}dt + dL_t^{(n),\star}. \tag{5.42}$$

$J^\star$ is a double exponential distribution random variables of parameters :

$$\rho^{+,\star} \quad = \quad \rho^+ - 1, \tag{5.43}$$

$$\rho^{-,\star} \quad = \quad \rho^- + 1, \tag{5.44}$$

$$p^\star \quad = \quad \frac{p\rho^+\rho^{-,\star}}{p\rho^+\rho^{-,\star} + (1-p)\,\rho^-\rho^{+,\star}}, \tag{5.45}$$

with the following joint characteristic function :

$$\psi^\star(z_1,z_2) := \mathbb{E}\left(e^{z_1 J^\star + z_2 |J^\star|}\right) \quad = \quad \frac{\psi(z_1+1,z_2)}{\psi(1,0)}. \tag{5.46}$$

iii) the processes X_t and V_t satisfy the following SDEs :

$$\begin{aligned} dX_t \quad &= \quad \left(r + \frac{1}{2}V_t - \lambda_t^\star\left(\psi^\star(1,0)-1\right)\right)dt \\ &\quad + \sqrt{V_t}\left(\rho dW_t^{\star 1} + \sqrt{1-\rho^2}dW_t^{\star 2}\right) + d\left(\sum_{l=1}^{N_t^\star} J_l^\star\right), \end{aligned} \tag{5.47}$$

$$dV_t \quad = \quad \left(\theta_V - (k_V - \rho\sigma_V)V_t\right)dt + \sigma_V\sqrt{V_t}dW_t^{\star 1}. \tag{5.48}$$

Proof. In Appendix C 5.C. □

From to this proposition and using the characteristic function of the log-return in proposition 13 we value an European call option of maturity T and strike K, written on the terminal spot price S_T of some underlying asset in the next proposition.

Proposition 16. *The price $C(t,T,S,V,\lambda,N,Y,K)$ of an European call option of maturity T, written on S_t is given by*

$$\begin{aligned} C(t,T,S,V,\lambda,N,Y,K) \quad &= \quad S_t P_0(t,T,S,V,\lambda,N,Y,K) \\ &\quad - Ke^{-r(T-t)}P_1(t,T,S,V,\lambda,N,Y,K), \end{aligned} \tag{5.49}$$

where K is the strike price, $P_j \equiv P_j(t,T,S,V,\lambda,N,Y,K)$ are the probability of the call expiring in-the-money under the forward measure $\mathbb{Q}^\star$ and the risk neutral measure $\mathbb{Q}$ respectively. For $j \in \{0,1\}$, these probabilities are of the form

$$P_0 = \frac{1}{2} + \frac{1}{\pi}\int_{\mathbb{R}_+} Re\left(\exp(-i\omega\log K)\frac{\Upsilon_{t,s}(\omega - i)}{i\omega e^{r(T-t)}S_t}\right)d\omega, \tag{5.50}$$

$$P_1 = \frac{1}{2} + \frac{1}{\pi}\int_{\mathbb{R}_+} Re\left(\exp(-i\omega\log K)\frac{\Upsilon_{t,s}(\omega)}{i\omega}\right)d\omega \tag{5.51}$$

where $\Upsilon_{t,s}$ is the characteristic function of the process $X_t$under the measure $\mathbb{Q}$ defined in the corollary 14.

Proof. Considering the discretisation scheme on the partition $\mathscr{E}^{(n)}$, the European call option price is obtained as

$$\begin{aligned}
C\left(t,T,S^{(n)},V,\lambda^{(n)},N^{(n)},Y,K\right) &= C\left(t,T,X^{(n)},V,\lambda^{(n)},N^{(n)},Y,K\right) \\
&= e^{-r(T-t)}\mathbb{E}\left(\max\left(e^{X_T^{(n)}} - K, 0\right)|\mathscr{F}_t\right) \\
&= e^{X_t^{(n)}}\mathbb{E}\left(\frac{e^{-rT}e^{X_T^{(n)}}}{e^{-rt}e^{X_t^{(n)}}}\mathbb{I}_{X_T^{(n)}>\log K}|\mathscr{F}_t\right) \\
&\quad -Ke^{-r(T-t)}\mathbb{E}\left(\mathbb{I}_{X_T^{(n)}>\log K}|\mathscr{F}_t\right) \\
&= e^{X_t^{(n)}}\mathbb{E}^\star\left(\mathbb{I}_{X_T^{(n)}>\log K}|\mathscr{F}_t\right) \\
&\quad -Ke^{-r(T-t)}\mathbb{E}\left(\mathbb{I}_{X_T^{(n)}>\log K}|\mathscr{F}_t\right) \\
&= S_t^{(n)}\mathbb{Q}^\star\left(X_T^{(n)} > \log K\right) \\
&\quad -Ke^{-r(T-t)}\mathbb{Q}\left(X_T^{(n)} > \log K\right) \\
&= S_t^{(n)}P_0^{(n)} - Ke^{-r(T-t)}P_1^{(n)}
\end{aligned} \tag{5.52}$$

where as in the Black-Sholes formula (1973)

$$P_0^{(n)} \equiv P_0\left(t,T,X^{(n)},V,\lambda^{(n)},N^{(n)},Y,K\right)$$

and

$$P_1^{(n)} \equiv P_1\left(t,T,X^{(n)},V,\lambda^{(n)},N^{(n)},Y,K\right)$$

are the probability of the call expiring in-the-money under the measure $\mathbb{Q}^\star$ and the risk neutral measure $\mathbb{Q}$ respectively. Since

$$C\left(T,T,X^{(n)},V,\lambda^{(n)},N^{(n)},Y,K\right) = \max\left(e^{X_T^{(n)}} - K, 0\right)$$

, for $j \in \{0,1\}$, with the terminal condition

$$P_j\left(T,T,X^{(n)},V,\lambda^{(n)},N^{(n)},Y,K\right) = \begin{cases} 1 & \text{if } X_T^{(n)} > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{5.53}$$

Given that the P_j are survival probabilities with respect of $X_T^{(n)} \mid \mathscr{F}_t$ under the measure $\mathbb{Q}^\star$ and the risk neutral measure $\mathbb{Q}$ respectively. We apply the inversion theorem of GIL-PELAEZ (1951) to get:

$$P_0^{(n)} = \frac{1}{2} + \frac{1}{\pi} \int_{\mathbb{R}_+} \text{Re}\left(\frac{e^{-i\omega \log K} \Upsilon_{t,T}^{(n),\star}(\omega)}{i\omega} \right) d\omega \tag{5.54}$$

$$P_1^{(n)} = \frac{1}{2} + \frac{1}{\pi} \int_{\mathbb{R}_+} \text{Re}\left(\frac{e^{-i\omega \log K} \Upsilon_{t,T}^{(n)}(\omega)}{i\omega} \right) d\omega, \tag{5.55}$$

where $\Upsilon_{t,T}^{(n),\star}(.)$ and $\Upsilon_{t,T}^{(n)}(.)$ are the characteristic functions of $X_T^{(n)} \mid \mathscr{F}_t$ under the measure $\mathbb{Q}^\star$ and the risk neutral measure $\mathbb{Q}$ respectively. From the definition of the measure $\mathbb{Q}^\star$ we have

$$\begin{aligned}
\Upsilon_{t,T}^{(n),\star}(\omega) &= \mathbb{E}^\star\left(e^{i\omega X_T^{(n)}} \mid \mathscr{F}_t \right) \\
&= \mathbb{E}\left(\frac{e^{-rT} e^{X_T^{(n)}}}{e^{-rt} e^{X_t^{(n)}}} e^{i\omega X_T^{(n)}} \mid \mathscr{F}_t \right) \\
&= \frac{1}{e^{r(T-t)} S_t^{(n)}} \mathbb{E}\left(e^{i(\omega - i) X_T^{(n)}} \mid \mathscr{F}_t \right) \\
&= \frac{\Upsilon_{t,T}^{(n)}(\omega - i)}{e^{r(T-t)} S_t^{(n)}}.
\end{aligned}$$

Replacing this into Equation (5.54), passing to the limit with respect of n and using the frequency derivative property of Fourier transform, allows us to complete this proof. □

The main problem for implementing this latter call option formula is the numerical approximation of integrals in equations (5.50) and (5.51). It worth noting that a wrong implementation involves a weak quality of fitness on real market option data. One may think about Fourier transform, but the characteristic function $\Upsilon_{t,s}$ is only evaluated numerically with possible discontinuities and the integrands have a point of singularity at $\omega = 0$. However, these integrals can be evaluate efficiently via a trapezoidal quadrature approximation. Note that the Lemma 1 in Witkovský (2001) enables one to compute the limits of integrands in equations (5.50) and (5.51) when ω vanish :

$$\text{Re}\left(\frac{e^{-i\omega \log K}}{i\omega} \frac{\Upsilon_{t,T}(\omega - i)}{e^{r(T-t)} S_t} \right) \to \mathbb{E}^\star(X_T \mid \mathscr{F}_t) - \log K \tag{5.56}$$

$$\text{Re}\left(\frac{e^{-i\omega \log K}}{i\omega} \Upsilon_{t,T}(\omega) \right) \to \mathbb{E}(X_T \mid \mathscr{F}_t) - \log K \tag{5.57}$$

where

$$\begin{aligned}\mathbb{E}^{\star}(X_T \mid \mathscr{F}_t) &= r(T-t)+\frac{1}{2}\frac{\theta_V}{k_V-\rho\sigma_V}(T-t) \\ &\quad +\left(V_t-\frac{\theta_V}{k_V-\rho\sigma_V}\right)\frac{1}{2(k_V-\rho\sigma_V)}\left(1-e^{-(k_V-\rho\sigma_V)(T-t)}\right), \end{aligned} \tag{5.58}$$

$$\begin{aligned}\mathbb{E}(X_T \mid \mathscr{F}_t) &= r(T-t)-\frac{1}{2}\frac{\theta_V}{k_V}(T-t) \\ &\quad -\frac{1}{2}\left(V_t-\frac{\theta_V}{k_V}\right)\frac{1}{k_V}\left(1-e^{-k_V(T-t)}\right), \end{aligned} \tag{5.59}$$

are results proven in Appendix D 5.D.

The trapezoidal quadrature approximation of the probabilities of the call expiring in-the-money for M discrete log-strikes $\log K_u := k_u = -k_{max}+(u-1)\Delta_k$, $u=1,...,M$, equally spaced $\Delta_k = \frac{2k_{max}}{M}$ are given by :

$$P_0 \approx \frac{1}{2}+\frac{1}{\pi}\sum_{l=1}^{M}\gamma_l \mathrm{Re}\left(\frac{e^{-i\omega_l k_u}}{i\omega_l}\frac{\Upsilon_{t,T}(\omega_l-i)}{e^{r(T-t)}S_t}\right)\Delta\omega$$

$$P_1 \approx \frac{1}{2}+\frac{1}{\pi}\sum_{l=1}^{M}\gamma_l \mathrm{Re}\left(\frac{e^{-i\omega_l k_u}}{i\omega_l}\Upsilon_{t,T}(\omega_l)\right)\Delta\omega$$

where $\gamma_l = \frac{1}{2}1_{\{l=1\}}+1_{\{l\neq 1\}}$ is the trapezoidal quadrature weights, $\Delta\omega = \frac{2\pi}{M\Delta_K}$ ensures that the absolute value of the integrand function is sufficiently small for all $\omega > \omega_M$ with $\omega_l = (l-1)\Delta_\omega$ for $l=1,...,M$.

From our call option price formula (5.49) we can derive Greeks for hedging purpose given the three sources of risk: the price risk S_t the volatility risk V_t and the jump risk λ_t. The analytical expressions of the three deltas are given by

$$\begin{aligned}\Delta S &= \frac{\partial C}{\partial S} = P_0 + S_t\frac{\partial P_0}{\partial S} - Ke^{-r(T-t)}\frac{\partial P_1}{\partial S}, \\ \Delta V &= \frac{\partial C}{\partial V} = S_t\frac{\partial P_0}{\partial V} - Ke^{-r(T-t)}\frac{\partial P_1}{\partial V}, \\ \Delta\lambda &= \frac{\partial C}{\partial \lambda} = S_t\frac{\partial P_0}{\partial \lambda} - Ke^{-r(T-t)}\frac{\partial P_1}{\partial \lambda}, \end{aligned} \tag{5.60}$$

where for $j \in \{0,1\}$

$$\begin{aligned}
\frac{\partial P_0}{\partial S} &= \frac{1}{\pi}\int_{\mathbb{R}_+} \mathrm{Re}\left(\frac{e^{-i\omega\log K}\Upsilon_{t,T}(\omega-i)}{e^{r(T-t)}S_t^2}\right)\mathrm{d}\omega,\\
\frac{\partial P_1}{\partial S} &= \frac{1}{\pi}\int_{\mathbb{R}_+} \mathrm{Re}\left(\frac{e^{-i\omega\log K}\Upsilon_{t,T}(\omega)}{S_t}\right)\mathrm{d}\omega,\\
\frac{\partial P_0}{\partial V} &= \frac{1}{\pi}\int_{\mathbb{R}_+} \mathrm{Re}\left(B(t,T,\omega-i)\frac{e^{-i\omega\log K}\Upsilon_{t,T}(\omega-i)}{i\omega e^{r(T-t)}S_t}\right)\mathrm{d}\omega,\\
\frac{\partial P_1}{\partial V} &= \frac{1}{\pi}\int_{\mathbb{R}_+} \mathrm{Re}\left(B(t,T,\omega)\frac{e^{-i\omega\log K}\Upsilon_{t,T}(\omega)}{i\omega}\right)\mathrm{d}\omega,\\
\frac{\partial P_0}{\partial \lambda} &= \frac{1}{\pi}\int_{\mathbb{R}_+} \mathrm{Re}\left(D(t,T,\omega-i)\frac{e^{-i\omega\log K}\Upsilon_{t,T}(\omega-i)}{i\omega e^{r(T-t)}S_t}\right)\mathrm{d}\omega,\\
\frac{\partial P_1}{\partial \lambda} &= \frac{1}{\pi}\int_{\mathbb{R}_+} \mathrm{Re}\left(D(t,T,\omega)\frac{e^{-i\omega\log K}\Upsilon_{t,T}(\omega)}{i\omega}\right)\mathrm{d}\omega,
\end{aligned}$$

with

$$B(t,T,\omega) = \sigma_V^{-2}\left(k_V - i\omega\rho\sigma_V - \bar{d}\right)\frac{1-e^{-\bar{d}(s-t)}}{1-\bar{g}e^{-\bar{d}(s-t)}},$$

and D which satisfies the ODE in equation (5.39). Similarly we can obtain the second and third order Greeks letter, gamma and vega which are respectively the second and third order partial derivatives of the call option price with respect of the asset price, the variance and the arrival rate of jumps. Thus, the sensitivity of our call option price with respect of these three state variables has an analytic form that relies on the computation of integrals.

5.4 Illustration

This section enables us to achieve three goals. Firstly, we present the method for calibrating our Heston model extension to an implied volatility surface. Secondly, we show that our model achieves a better fit of implied volatilities than the standard Heston. Finally, we study the sensitivity of European call option price to the memory parameter β.

5.4.1 Data description

The data used in this section is a result of a filtering process applied to the implied volatility surface of the Euro Stoxx 50 as of the 26^{th} of September 2019 in figure 5.4, with the spot price $S_0 = 3541$ and the dividend rate $q = 0.00225$. To our raw data obtain from the Bloomberg's Option Monitor we exclude options with maturity greater than one as there are less sensitive to volatility. The constant risk free rate r is approximated by the three-month instantaneous forward rate on the 26^{th} of the September 2019 implied in the yield curve for the euro area and taking into account issuers whose rating is triple A.

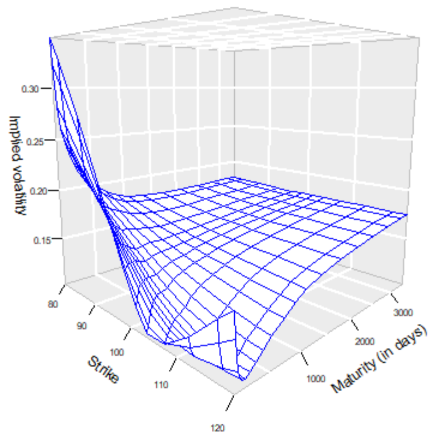


Figure 5.4: Implied volatility surface of the Euro Stoxx 50 as of the 26^{th} of september 2019, with the spot price $S_0 = 3541$ and the dividend rate $q = 0.00225$.

5.4.2 Calibration and sensitivity analysis

To calibrate our SVJ model under the risk neutral measure, we use the standard approach that consists in minimizing the sum of relative squares errors (SRSE) between IVs derived from our model and market IVs. More specifically, If we denote by Θ the parameters set of our SVJ model, one evaluate the vector of parameters $\hat{\Theta}$ for which our model gives the closest prices to those observed in the market as follows:

$$\hat{\Theta} = \underset{\Theta,\gamma}{\operatorname{argmin}} \left(\mathrm{SRSE}(\Theta) + \gamma \max\left(\sigma_V^2 - 2\theta_V, 0\right)\right), \tag{5.61}$$

where $\mathrm{SRSE}(\Theta)$ is defined by

$$\sum_s^M \left(\frac{\sigma_{IV}^{Market}(t, T_s, S, V, \lambda, N, Y, K_s) - \sigma_{IV}(t, T_s, S, V, \lambda, N, Y, K_s, \Theta)}{\sigma_{IV}^{Market}(t, T_s, S, V, \lambda, N, Y, K_s)} \right)^2. \tag{5.62}$$

K_s is the strike price of the s-th call option, T_s is the maturity of the call option, σ_{IV}^{Market} is the market implied volatility of the observed call option and M is the number of call options used. Note that the objective function is not convex and is not of any particular structure. Thus, by using a gradient based optimizing procedure we are not sure to obtain a global minimum. As proposed by Hamida and Cont (2005), we add a convex penalization term $\gamma \max\left(\sigma_V^2 - 2\theta_V, 0\right)$ to the objective function (5.62) in order to ensure that a gradient based algorithm can be used.

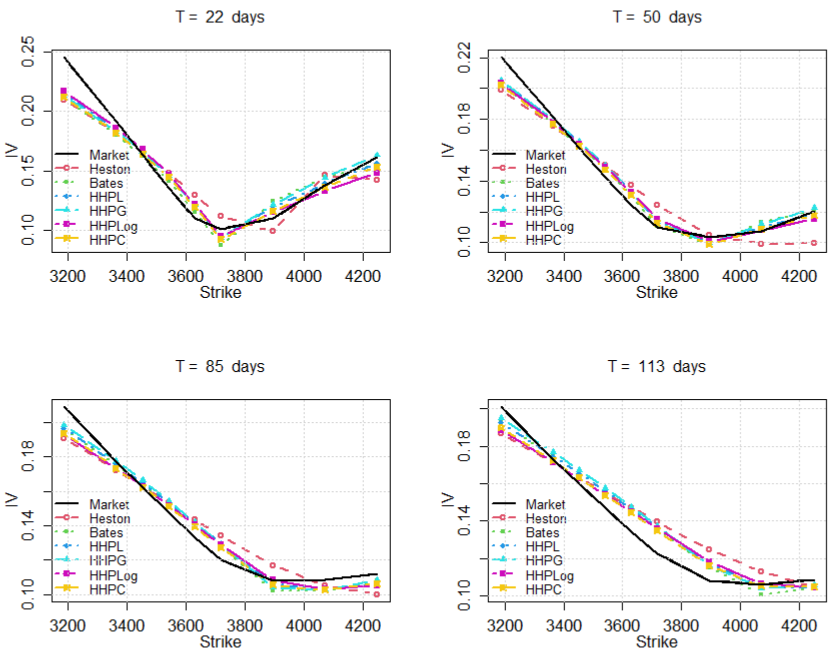


Figure 5.5: Market implied volatilities versus replicated implied volatilities from the Heston, Bates, HHPL, HHPG, HHPLog and HHPC models at maturities T equal to 22, 50, 85 and 113 days.

Based on the filtered IV surface of the Euro Stoxx 50 as of the 26$^{\text{th}}$ September 2019, we find the set of model parameters minimizing the objective function (5.62). Table 5.1 presents the result of our calibration procedure applied to the Bates model, the Heston model, and to its extensions with a self-exciting jumps induce by a Laplace (5.6) (HHPL) , Gaussian (5.8) (HHPG), Logistic (5.10) (HHPLog) or Cauchy (5.12) (HHPC) kernels. In table 5.1 the SRSE measures the quality of fit of theses models. We find that Heston models with jumps in the asset price dynamics are better, with SRSEs lower than the SRSE of the Heston model. In figure 5.5, we plot the market implied volatility and the volatility replicated through the Heston, Bates, HHPL, HHPG, HHPLog and HHPC models at maturities T equal to 22, 50, 85 and 113 days. We see that the implied volatilities derived from the Bates, HHPL, HHPG, HHPLog and HHPC models are the closest to market one and they outperform the Heston model, particularly for small maturities.

Another important quality check, when it comes to calibrate an implied volatility surface, is the replication of the term structure of at-the-money (ATM) volatility skews. Denoted by $\Psi(T)$, the ATM volatility skew at expiry time $T \geq t$ is the derivative of

	Heston	Bates	HHPL	HHPG	HHPLog	HHPC
θ_V	0.2183	0.2567	0.2475	0.2223	0.2744	0.2320
k_V	7.3697	7.9872	8.0511	6.7373	9.7406	7.8421
σ_V	0.6575	0.7068	0.70201	0.6632	0.7406	0.6808
ρ	-0.8136	-0.9895	-0.9983	-0.9969	-0.9989	-0.9913
V_0	0.0217	0.0184	0.0208	0.0201	0.0218	0.0202
λ_0			1.5484	1.1158	1.3813	1.3013
α			2.48734	8.2820	2.7708	4.5728
η			3.1635	0.2860	9.4444	4.3008
β			0.1362	0.3683	7.9078	4.9346
p			0.5106	0.5452	0.7765	0.4081
ρ^+			53.5226	46.6534	64.0890	55.3728
ρ_-			-144.0291	-137.6665	-97.4525	-149.4079
λ		0.1316	-144.0291	-137.6665	-97.4525	-149.4079
μ_J		0.0422	-144.0291	-137.6665	-97.4525	-149.4079
σ_J		0.0387	-144.0291	-137.6665	-97.4525	-149.4079
SRSE	0.5256	0.2873	0.2595	0.2647	0.2786	0.3005

Table 5.1: Estimated parameters.

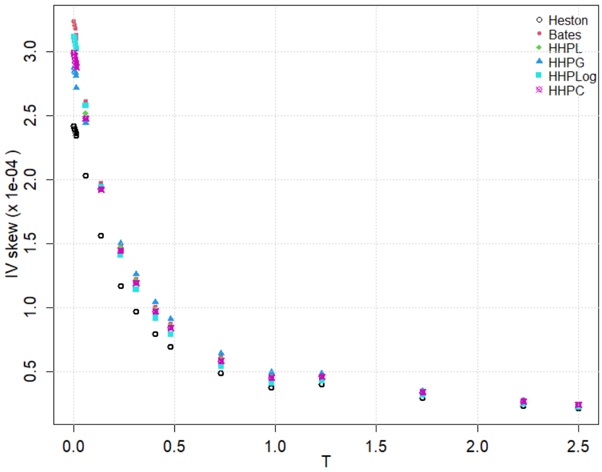


Figure 5.6: The slope of implied volatility skew as function of the maturity over a range from one day to 2.5 years. Model Parameters used are from table 5.1.

implied volatility with respect to the money strike:

$$\Psi(T) \quad := \quad \left| \frac{\partial \sigma_{IV}(t,T,S,V,\lambda,N,Y,K,\Theta)}{\partial K} \right|_{K=S_t}$$

For each model and for ε small enough, we approximate $\Psi(T)$ by

$$\hat{\Psi}(T) \quad = \quad \left| \frac{\sigma_{IV}\left(t,T,S,V,\lambda,N,Y,K+\varepsilon,\hat{\Theta}\right) - \sigma_{IV}\left(t,T,S,V,\lambda,N,Y,K-\varepsilon,\hat{\Theta}\right)}{2\varepsilon} \right|_{K=S_t},$$

where the approximation holds for small enough ε. Empirical studies (see e.g. Gatheral et al. 2014) show that the observed term structure of ATM volatility skew is well approximated by power-law functions of the form $T^{H-\frac{1}{2}}$ with $H \in \left(0, \frac{1}{2}\right)$ the Hurst exponent . Gatheral et al. (2014) shows that a value of H close to zero allows to generate volatility surface with a reasonable shape. From the table 5.2, we deduce that all models considered have an estimated Hurst exponent $\hat{H}$ between 0.09 and 0.14. The HHPLog model exhibits the smallest estimated Hurst exponent with an R-squared of the linear regression model that explains $\log \hat{\Psi}(T)$ given $\log T$ equal to 0.88. From figure 5.6, we conclude that compared to the Heston model, the other models better replicate the empirical short term explosion of the ATM volatility skew. As argued by Errais et al. 2010, a homogeneous Poisson process fails to reproduce the time clustering effect of jumps. To capture this effect, our extensions of the Heston model with self-excited jumps are a good alternative to the Bates model.
We now focus on the differences between our extensions of the Heston model. To distinguish them in term of memory length, we draw in figure 5.7 their corresponding

	Heston	Bates	HHPL	HHPG	HHPLog	HHPC
$\hat{H}$	0.13	0.11	0.12	0.14	0.09	0.12
R^2	0.87	0.88	0.87	0.85	0.88	0.87

Table 5.2: Estimated Hurst exponent $\hat{H}$ of the ATM volatility skew and the R-squared of the linear regression model that explains the log-ATM volatility skew $\log \hat{\Psi}(T)$ given $\log T$.

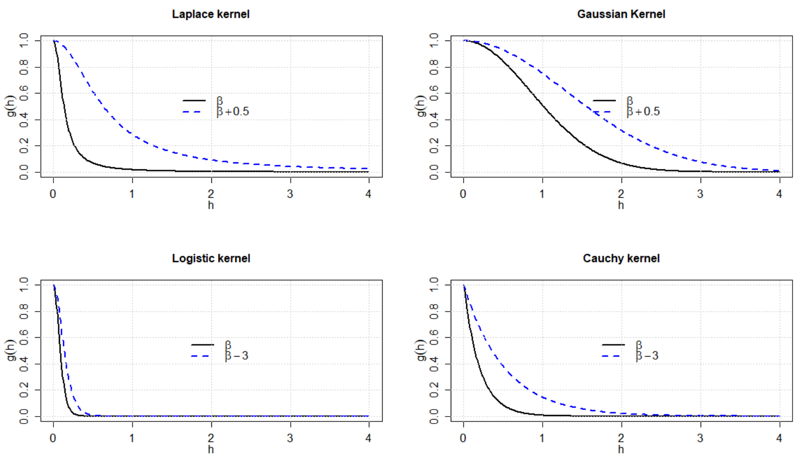


Figure 5.7: Sensitivity of kernel functions to the memory parameter β. The initial value of β for each model is in table 5.1.

kernel functions given the fitted parameters in table 5.1. We observe that the Gaussian kernel has the slowest decay, while the Logistic kernel has the fastest decay. Thus, the HHPG and HHPLog are respectively the models with the longest and shortest memory among our extensions of the Heston model. From the figure 5.7 also, we find that the memory of the HHPL and HHPG models increases with the parameter β. Whereas, the one of HHPLog and HHPC models increases when the parameter β decreases. The figure 5.8 plot the probability density function of the log-return at 113 days. This figure allows to make link between the memory length of models and the thickness of the tail of the log-return distribution. In agreement of the result of the section 5.2, the thickness of the tail of the log-return distribution for our extension of the Heston model are ordered by their memory lengths.

The impact of the memory parameter β of our affine Heston style models on option pricing is studied through the implied volatilities and the price risk ΔS (5.60). To inspect how recalling past price changes for a long period influences option prices, we vary the memory parameter β of our extensions of the Heston model. Figure 5.9 reveals that an increase in memory indicated by a slower decaying kernel function $g_{\cdot}(.)$, induces an increase in implied volatilities of out-of-the-money (OTM) call op-

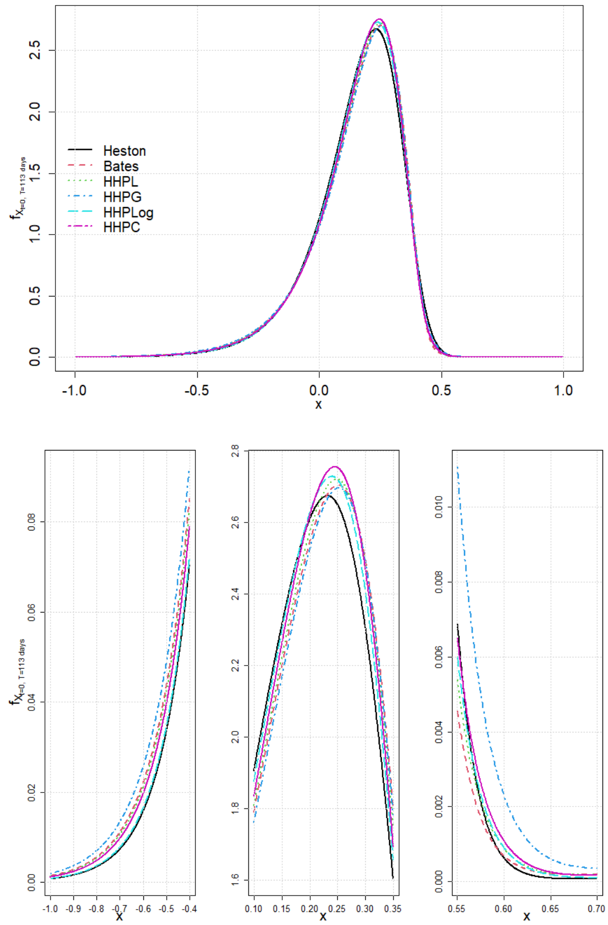


Figure 5.8: Probability density function of the log-return $f_{X_{0,113\,\mathrm{days}}}$.

tions ($S < K$). This increase in implied volatilities becomes more pronounced as the expiry rises. The figure 5.10 confirms this findings and brings out two points. Firstly, the impact of the memory parameter β is negligible in the short term, but increases with the expiry time T of the call option on Euro Stoxx 50. Secondly, for the HHPLog model with the shortest memory, the impact of a small increase or decrease in memory remains negligible in the long term.

The price risk ΔS measures the sensitivity of the option to the stock price. Let us recall that for options it should be interpreted as the amount of stocks to hold ($\Delta S > 0$) or to short sale ($\Delta S < 0$) for hedging an option. Called the "delta hedging", this is a very common strategy to do the arbitrage and minimize risk of portfolio in the option market. From the figure 5.11, the price risk ΔS calculated from the Heston model is

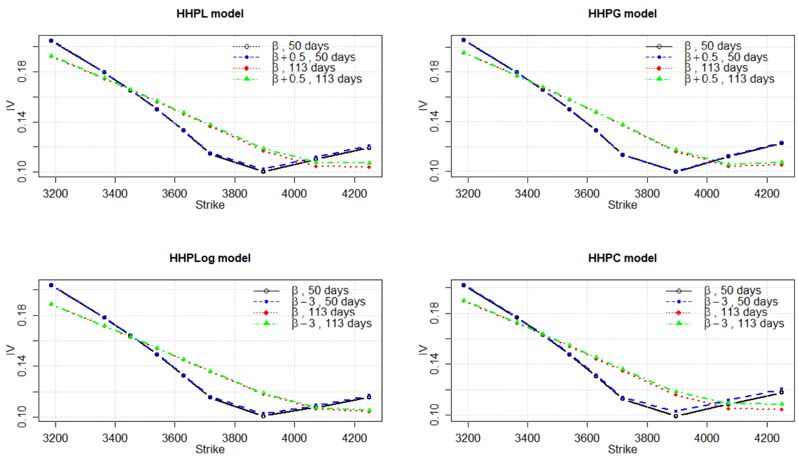


Figure 5.9: Sensitivity to the memory parameter β of the HHPL model. Model parameters used are in table 5.1.

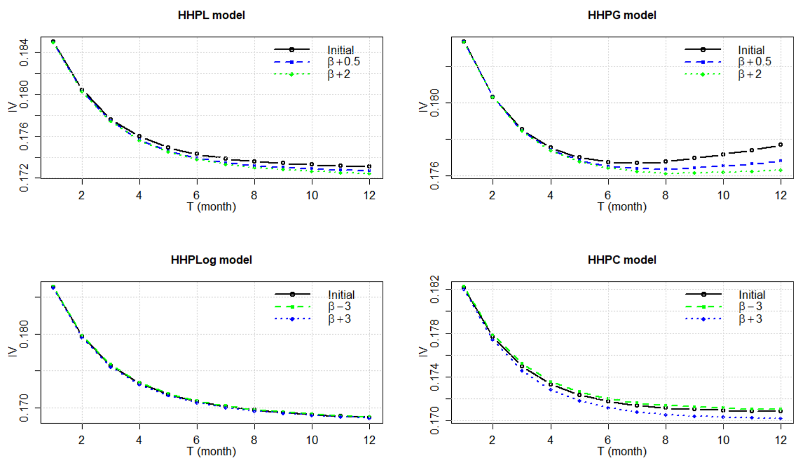


Figure 5.10: Sensitivity of the implied volatility to the memory parameter β regarding the expiry time T in month of the call option on Euro Stoxx 50. Model parameters used are in table 5.1.

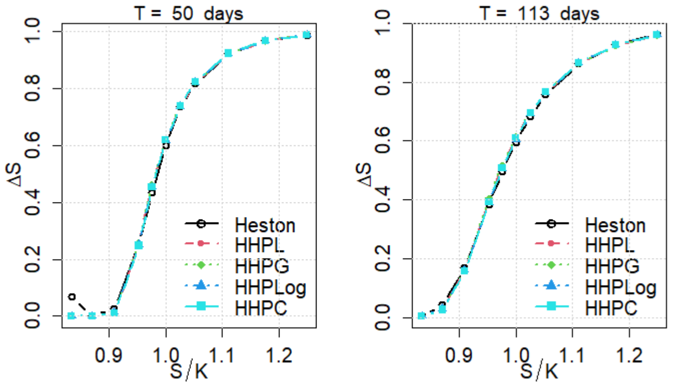


Figure 5.11: The price risk ΔS for the Heston, HHPL, HHPG, HHPLog and HHPC models. Model parameters used are in 5.1.

slightly higher than the one given by the other models for OTM call options, while, for in-the-money (ITM) call options, the price risk ΔS in the Heston model is almost the same. The figure 5.12 highlights the sensitivity of the price risk ΔS to the self-excitation parameter λ_0. We find that the increase in self-excitation caused by a rising λ_0, implies a higher price risk ΔS for OTM call options and a lower one for ITM call options. Unlike the price risk ΔS, from figure 5.13, the implied volatilities derived from our extensions of the Heston model are more sensitive to the self-excitation parameter λ_0 for OTM call options.

Although call option prices from the Heston model and its extensions have roughly the same sensitivity to changes in underlying price, we find that these models induce different shapes of sensitivity to volatility and jump risks in the option price. The upper graphs in figure 5.14 draw the sensitivity of call option prices expiring at 50 and 113 days to the volatility V. Known as the greek "vega", for all models considered, it takes its greatest value around the ATM. Except for the Heston model, the vega remains positive and decreases exponentially to zero as the strike moves away from the spot price. In the lower graphs, we observe a similar shape of the call option price change in case of a jump. Denoted by $\Delta\lambda$, it takes its greatest value at the OTM region and its curve widens when the expiry increases. It is worth nothing that a higher memory length does not necessary results in a wider $\Delta\lambda$ curve.

5.5 Conclusions

In this chapter, we propose an extension of the Heston model that allows for self-excited jumps in asset price and long memory of past price changes. This expansion corresponds to an asset price process with a stochastic volatility and a jump component driven by a linear Hawkes process (HP). The memory feature of our model stems from the decay rate of the HP kernel function. Considering HP kernel function that are Fourier transform of the Laplace, Gaussian, Logistic, and Cauchy measures, we generate different memory ranges, from the longest to the shortest. The self-excitation and memory properties of our framework enables us to have a realistic option pricing model that encompass major characteristic of asset's return times series such as the volatility clustering and the long memory in volatility. However, the memory effect leads to a dependence on the past that makes our asset price process non Markov.

The HP with kernel function which is a Fourier transform of the Cauchy measure corresponds to the tractable exponential kernel. Except the stochastic volatility model with jumps self-excited by this exponential kernel, our Heston extensions are non Markov processes. We use the Fourier transform representation of their kernel functions to establish a closed form expression of the conditional moment generating functions of the HP's intensity and the log-returns. Based on these results, we write the call option price according to the delta probability decomposition and derive Greeks for hedging purpose given the three sources of risk: the change in underlying price, volatility and the occurrence of a jump. By inverse Fourier transform of the conditional moment generating function of the log-returns, we find that with equal parameters, the thickness of the tails of its probability density function is fatter if the length of the memory of the HP kernel function is greater.

An estimation of our models to implied volatilities of the Euro Stoxx 50 as of the 26^{th} of September 2019 achieves better calibration performance than the one with the standard Heston model, particularly in the short term. They also provide a good fit of the term structure of at-the-money volatility skew, approximted by power-law functions of the form $T^{\hat{H}-\frac{1}{2}}$ with an estimated Hurst exponent $\hat{H}$ between 0.09 and 0.14. Considering the estimated parameters found and varying the memory parameter, we see that implied volatilities derived from our extensions of the Heston model , increase as the memory length expands. This sensitivity to the memory length appears more pronounced for out-of-the money call option prices. A similar conclusion is drawn, with respect to the sensitivity of the implies volatilities to other self-excitation parameters. The analysis of the Greeks delta, vega and the jump risk reveals that our extensions of the Heston model have roughly the same delta curve than the standard Heston. But these models induce different shapes of sensitivity to volatility and jump risks in the option price, which does not seem to be related only to their memory lengths.

A next step to this work will be to explore hedging strategies in our setting, to manage options whose underlying asset may experience some clustering of jumps. From a broader perspective, it would be interesting to extend this work to the pricing of bond, currency, exotic options or exotic derivatives which are path dependent. From an empirical point of view, it would be worthy to study the empirical relevance of such Hawkes memory kernels and develop accurated calibration procedure to financial time series.

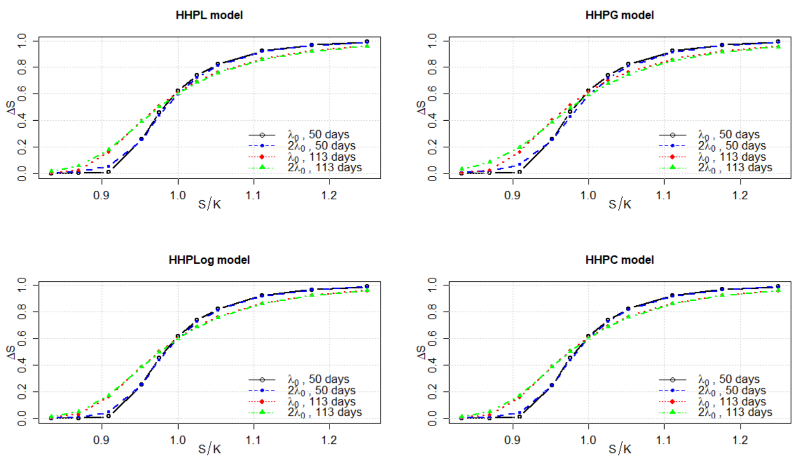


Figure 5.12: Sensitivity of the price risk ΔS to the initial intensity λ_0 of Hawkes processes. Model parameters used are in table 5.1.

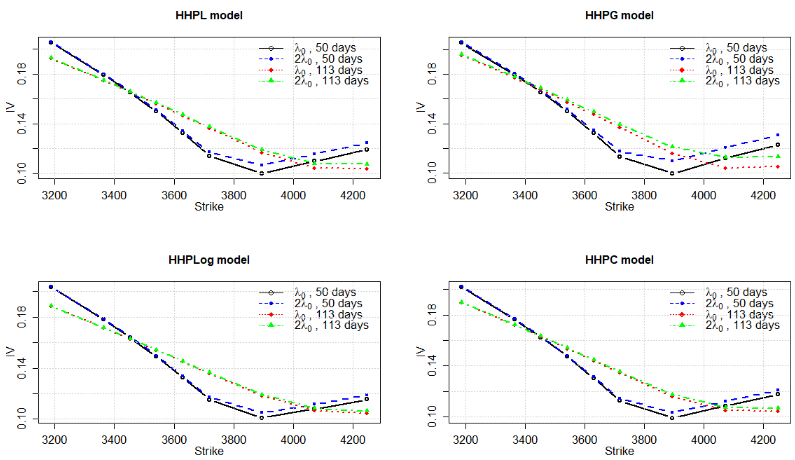


Figure 5.13: Sensitivity to the memory parameter λ_0 of the HHPL model. Model parameters used are in table 5.1.

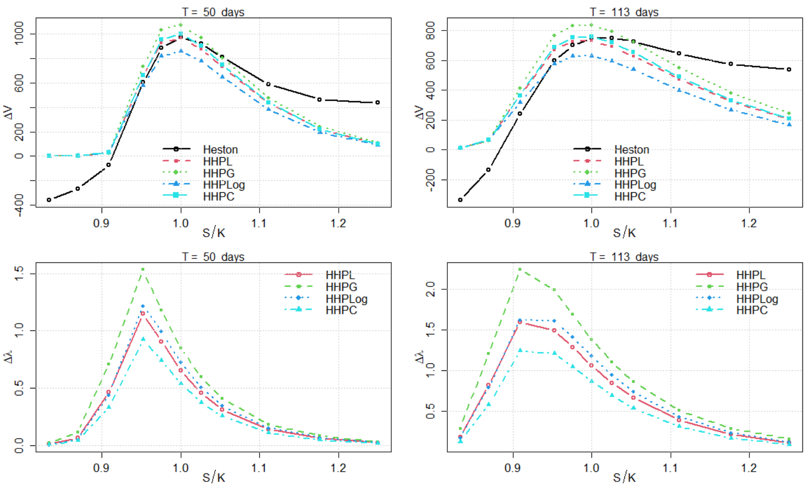


Figure 5.14: Vega and jump risk sensitivity of call prices expiring at 50 and 113 days. Model parameters used are in table 5.1.

Appendices

5.A Appendix A: proof of proposition 11

For fixed time s and $\omega \in \mathbb{C}_-$ with $\mathbb{C}_-$ the set of complex numbers with non-positive real part, we define the conditional moment generating function by

$$f \equiv f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\ldots,n}, N_t^{(n)}\right) \quad := \quad \mathbb{E}\left(e^{\omega \lambda_s^{(n)}} | \mathscr{F}_t\right),$$

where f is a complex-valued function on $[0,s] \times G$ with $G = \mathbb{R}_+ \times \mathbb{R}_+^n \times \mathbb{N}$ the state space of the Markov process $\left(\lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\ldots,n}, N_t^{(n)}\right)_{t \geq 0}$. By the tower property, for any $\tau \geq t$,

$$f \quad = \quad \mathbb{E}\left(f\left(\tau, s, \lambda_\tau^{(n)}, \left(Y_\tau^{(l)}\right)_{l=-n,\ldots,n}, N_\tau^{(n)}\right) | \mathscr{F}_t\right),$$

and

$$\lim_{\tau \to t} \frac{\mathbb{E}\left(f\left(\tau, s, \lambda_\tau^{(n)}, \left(Y_\tau^{(l)}\right)_{l=-n,\ldots,n}, N_\tau^{(n)}\right) | \mathscr{F}_t\right) - f}{\tau - t} \quad = \quad 0.$$

It follows that the infinitesimal generator $\mathscr{A}$ of the Markov process $\left(\lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\ldots,n}, N_t^{(n)}\right)$ evaluated on f is equal to zero:

$$\mathscr{A} f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\ldots,n}, N_t^{(n)}\right) \quad = \quad 0. \tag{5.63}$$

Since $\mathscr{A} f\left(t, s, \lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right)$ exists the function f is in the domain of $\mathscr{A}$ which contains the set of twice differentiable functions such that all partial derivatives of up to order two are continuous function (see e.g. Duffie et al. (2003)). On the other hand, the Itô's lemma for the semi-martingales allows us to infer that f satisfies the following differential equation:

$$\begin{aligned} 0 \quad = \quad & \frac{\partial f}{\partial t} + i b_l^{(n)} Y_t^{(l)} \sum_{l=-n}^{n} \frac{\partial f}{\partial Y^{(l)}} \\ & + \left((\lambda_0 - \alpha) \frac{\mathrm{d} g_k(t)}{\mathrm{d}t} + \eta \sum_{l=-n}^{n} m_l^{(k,n)} i b_l^{(n)} Y_t^{(l)} \right) \frac{\partial f}{\partial \lambda^{(n)}} + \lambda_t^{(n)} \int_{\mathbb{R}} \Bigg(\\ & f\left(t, s, \lambda_t^{(n)} + \eta |z| \sum_{l=-n}^{n} m_l^{(k,n)}, \left(Y_t^{(l)} + |z|\right)_{l=-n,\dots,n}, N_t^{(n)} + 1\right) - f(.) \Bigg) \nu(\mathrm{d}z). \end{aligned} \tag{5.64}$$

Given that $\left(\lambda_t^{(n)}, \left(Y_t^{(l)}\right)_{l=-n,\dots,n}, N_t^{(n)}\right)$ is an affine jump process, from Duffie D. (2000) and Errais et al. (2010), under technical regularity condition, we conjecture that f is an exponential affine function of $\lambda_t^{(n)}$ and $\left(Y_t^{(l)}\right)_{l=-n,\dots,n}$:

$$f \quad := \quad \exp\left(F(t,s,\omega) + G(t,s,\omega)\lambda_t^{(n)} + \sum_{l=-n}^{n} H_l\left(t, s, b_l^{(n)}, \omega\right) m_l^{(k,n)} Y_t^{(l)} \right)$$

where F and G are time dependent functions such that $F(s,s,\omega) = 0$ and $G(s,s,\omega) = \omega$. While H_l is a function that depends on the time and the position $l \in \{-n, \dots, n\}$ within the partition $\mathscr{E}^{(n)}$. Its final value $H_l\left(s, s, b_l^{(n)}, \omega\right)$ is equal to zero. From (Hubalek et al. 2016), the technical regularity condition needed is equivalent to the integrability of the jumps, that is

$$\left.\frac{\partial G}{\partial t}(t,s,\omega)\right|_{t=s} = \mathbb{E}\left(e^{\eta \omega \sum_{l=-n}^{n} m_l^{(k,n)} |J|} \right) < \infty,$$

where the random jump size J has density $\nu(z)$ on $\mathbb{R}$. The partial derivatives of f are then given by :

$$\begin{aligned} \frac{\partial f}{\partial t} = \left(\frac{\partial F}{\partial t}(t,s,\omega) + \frac{\partial G}{\partial t}(t,s,\omega)\lambda_t^{(n)} + \sum_{l=-n}^{n} \frac{\partial H_l}{\partial t}\left(t,s,b_l^{(n)},\omega\right) m_l^{(n)} Y_t^{(l)} \right) f, \\ \frac{\partial f}{\partial \lambda^{(n)}} = G(t,s,\omega) f, \; \frac{\partial f}{\partial Y^{(l)}} = H_l\left(t,s,b_l^{(n)},\omega\right) m_l^{(k,n)} f \quad \forall l \in \{-n,\dots,n\}, \end{aligned}$$

and the variation of f when a jump of size z occurs is equal to

$$\left(\exp\left(|z| \left(\eta G(t,s,\omega) \sum_{l=-n}^{n} m_l^{(k,n)} + \sum_{l=-n}^{n} H_l\left(t,s,b_l^{(n)},\omega\right) m_l^{(k,n)} \right) \right) - 1 \right) f.$$

Injecting all these expressions into Equation (5.64) lead to :

$$\begin{aligned}
0 &= \frac{\partial F}{\partial t}(t,s,\omega)+\frac{\partial G}{\partial t}(t,s,\omega)\lambda_t^{(n)}+\sum_{l=-n}^{n}\frac{\partial H_l}{\partial t}\left(t,s,b_l^{(n)},\omega\right)m_l^{(n)}Y_t^{(l)} \\
&\quad +ib_l^{(n)}Y_t^{(l)}\sum_{l=-n}^{n}H_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)} \\
&\quad +\left((\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t}+\eta\sum_{l=-n}^{n}m_l^{(k,n)}ib_l^{(n)}Y_t^{(l)}\right)G(t,s,\omega) \\
&\quad +\lambda_t^{(n)}\left(\psi\left(0,\eta G(t,s,\omega)\sum_{l=-n}^{n}m_l^{(k,n)}+\sum_{l=-n}^{n}H_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}\right)-1\right)
\end{aligned} \tag{5.65}$$

where $\psi(z_1,z_2):=\mathbb{E}\left(e^{z_1J+z_2|J|}\right)$. From which we find the following ODE system:

$$\begin{cases}
0=\frac{\partial F}{\partial t}(t,s,\omega)+(\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t}G(t,s,\omega)\\
0=\frac{\partial G}{\partial t}(t,s,\omega)+\psi\left(0,\eta G(t,s,\omega)\sum_{l=-n}^{n}m_l^{(k,n)}+\sum_{l=-n}^{n}H_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}\right)-1\\
0=\frac{\partial H_l}{\partial t}\left(t,s,b_l^{(n)},\omega\right)+ib_l^{(n)}H_l\left(t,s,b_l^{(n)},\omega\right)+\eta ib_l^{(n)}G(t,s,\omega),\ \forall l\in\{-n,...,n\}.
\end{cases}$$

For all $l\in\{-n,...,n\}$, the last ODE admits the following solution:

$$H_l\left(t,s,b_l^{(n)},\omega\right) = b_l^{(n)}\eta i\int_t^s e^{ib_l^{(n)}(u-t)}G(u,s,\omega)\,\mathrm{d}u,$$

and the proof is complete.

5.B Appendix B: proof of proposition 13

Let $\omega\in\mathbb{R}$. As in proposition 11, if we note

$$f\equiv f\left(t,s,X_t^{(n)},\lambda_t^{(n)},\left(Y_t^{(l)}\right)_{l=-n,\dots,n},N_t^{(n)}\right) := \mathbb{E}\left(e^{i\omega X_s^{(n)}}\mid\mathscr{F}_t\right)$$

and under the regularity condition $\psi\left(i\omega,\eta\sum_{l=-n}^{n}m_l^{(k,n)}\right)<\infty$, we conjecture that f is an exponential affine function of $X_t^{(n)}$, $\lambda_t^{(n)}$ and $\left(Y_t^{(l)}\right)_{l=-n,\dots,n}$,

$$\begin{aligned}
f &:= \exp\left(i\omega X_t^{(n)}+A(t,s,\omega)+B(t,s,\omega)V_t\right) \\
&\quad\times\exp\left(\sum_{l=-n}^{n}C_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}Y_t^{(l)}+D(t,s,\omega)\lambda_t^{(n)}\right)
\end{aligned} \tag{5.66}$$

where $A(s,s,\omega)=0$, $B(s,s,\omega)=0$, $D(s,s,\omega)=0$ and $C_l\left(s,s,b_l^{(n)},\omega\right)=0$,for all $l\in\{-n,\cdots,n\}$. Then, by the Itô's lemma for the semi-martingales, f solves the following differential equation:

$$0 = \frac{\partial f}{\partial t} + \left(r - \frac{1}{2}V_t - \lambda_t^{(n)}(\psi(1,0)-1)\right)\frac{\partial f}{\partial X^{(n)}} + \frac{1}{2}V_t\frac{\partial^2 f}{\partial\left(X^{(n)}\right)^2} \tag{5.67}$$
$$+\rho\sigma_V V_t\frac{\partial^2 f}{\partial X\partial V} + (\theta_V - k_V V_t)\frac{\partial f}{\partial V} + \frac{1}{2}\sigma_V^2 V_t\frac{\partial^2 f}{\partial V^2} + \sum_{l=-n}^{n}\frac{\partial f}{\partial Y^{(l)}} ib_l^{(n)} Y_t^{(l)}$$
$$+\frac{\partial f}{\partial \lambda^{(n)}}\left((\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t} + \eta\sum_{l=-n}^{n} m_l^{(k,n)} ib_l^{(n)} Y_t^{(l)}\right)$$
$$+\lambda_t^{(n)}\int_{\mathbb{R}}\Delta f(z)\,\nu(\mathrm{d}z),$$

where

$\frac{\partial f}{\partial t} = \left(\frac{\partial A}{\partial t} + \frac{\partial B}{\partial t}V_t + \sum_{l=-n}^{n}\frac{\partial C_l}{\partial t}m_l^{(k,n)}Y_t^{(l)} + \frac{\partial D}{\partial t}\lambda_t^{(n)}\right)f,$ $\quad \frac{\partial f}{\partial X^{(n)}} = i\omega f,$

$\frac{\partial^2 f}{\partial\left(X^{(n)}\right)^2} = -\omega^2 f,$ $\quad \frac{\partial^2 f}{\partial X^{(n)}\partial V} = i\omega B f,$

$\frac{\partial f}{\partial V} = Bf,$ $\quad \frac{\partial^2 f}{\partial V^2} = B^2 f,$

$\frac{\partial f}{\partial Y^{(l)}} = C_l m_l^{(k,n)} f,$ $\quad \frac{\partial f}{\partial \lambda^{(n)}} = Df,$

and the variation of f, $\Delta f(z)$, when a jump of size z occurs and the time just before is equal to

$$\left(\exp\left(i\omega z + \left(\sum_{l=-n}^{n} C_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)} + \eta D(t,s,\omega)\right)|z|\right) - 1\right)f \tag{5.68}$$

Injecting all these expressions into equation (5.67) lead to :

$$\begin{cases} \frac{\partial A}{\partial t} &= -i\omega r - \theta_V B - D(\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t} \\ \frac{\partial D}{\partial t} &= i\omega(\psi(1,0)-1) - \int_{\mathbb{R}}\left(\exp\left(i\omega z + \left(\sum_{l=-n}^{n} C_l m_l^{(k,n)} + \eta D\right)|z|\right) - 1\right)\nu(\mathrm{d}z) \\ \frac{\partial B}{\partial t} &= \bar{a} - \bar{b}B + \bar{c}B^2 \\ \frac{\partial C_l}{\partial t} &= -(C_l + \eta D)\, ib_l^{(n)}, \forall l \in \{-n,...n\} \end{cases} \tag{5.69}$$

where $\bar{a} = \frac{\omega}{2}(\omega + i)$, $\bar{b} = -k_V + i\omega\rho\sigma_V$, $\bar{c} = -\frac{1}{2}\sigma_V^2$. In order to separate the part of the characteristic function linked to the jump component we assume that $A(t,s,\omega) = A_D(t,s,\omega) + A_J(t,s,\omega)$, with terminal condition $A_D(s,s,\omega) = 0 = A_J(s,s,\omega)$. It follows that the ODEs system (5.69) can be subdivided into two ODEs systems :

$$\begin{cases} \frac{\partial A_D}{\partial t} &= -i\omega r - \theta_V B \\ \frac{\partial B}{\partial t} &= \bar{a} - \bar{b}B + \bar{c}B^2 \end{cases} \tag{5.70}$$

and

$$\begin{cases} \frac{\partial A_J}{\partial t} &= -D(\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t} \\ \frac{\partial D}{\partial t} &= i\omega(\psi(1,0)-1) \\ & \quad - \int_{\mathbb{R}}\left(\exp\left(i\omega z + \left(\sum_{l=-n}^{n} C_l m_l^{(k,n)} + \eta D\right)|z|\right) - 1\right)\nu(\mathrm{d}z) \\ \frac{\partial C_l}{\partial t} &= -(C_l + \eta D)\, ib_l^{(n)}, \forall l \in \{-n,...n\} \end{cases} . \tag{5.71}$$

The ODEs system (5.70) corresponds at the one of the standard Heston model. From Schoutens et al. (2004)), $\exp\left(i\omega X_t^{(n)}+A_D(t,s,\omega)+B(t,s,\omega)V_t\right)$ is equal to

$$
\begin{aligned}
&= \exp\left(i\omega X_t^{(n)}+r(s-t)\right)\\
&\quad\times\exp\left(\theta_V\sigma_V^{-2}\left(k_V-i\omega\rho\sigma_V-\bar{d}\right)(s-t)-2\log\left(\frac{1-\bar{g}e^{-\bar{d}(s-t)}}{1-\bar{g}}\right)\right)\\
&\quad\times\exp\left(V_t\sigma_V^{-2}\left(k_V-i\omega\rho\sigma_V-\bar{d}\right)\frac{1-e^{-\bar{d}(s-t)}}{1-\bar{g}e^{-\bar{d}(s-t)}}\right)
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{d} &= \sqrt{\left(i\omega\rho\sigma_V-k_V\right)^2+\sigma_V^2\left(i\omega+\omega^2\right)},\\
\bar{g} &= \frac{k_V-i\omega\rho\sigma_V-\bar{d}}{\left(k_V-i\omega\rho\sigma_V+\bar{d}\right)}.
\end{aligned}
$$

Albrecher et al. (2006) prove that this form of the characteristic function of the standard Heston model is stable for Fourier inversion or numerical integration.
Let us focus on the ODEs system (5.71). Its last ODE admits the following solution:

$$
C_l\left(t,s,b_l^{(n)},\omega\right) = i\eta b_l^{(n)}\int_t^s e^{ib_l^{(n)}(u-t)}D(u,s,\omega)\,\mathrm{d}u. \tag{5.72}
$$

This last formula allows us to conclude the proof.

5.C Appendix C: proof of proposition 15

The derivative $\left.\frac{\mathrm{d}\mathbb{Q}^\star}{\mathrm{d}\mathbb{Q}}\right|_{\mathscr{F}_0}$, equal to

$$
\begin{aligned}
\exp\Bigg(&-\frac{1}{2}\int_0^t V_u\mathrm{d}u+\int_0^t\sqrt{V_u}\left(\rho\mathrm{d}W_u^1+\sqrt{1-\rho^2}\mathrm{d}W_u^2\right)\\
&-\left(\psi(1,0)-1\right)\int_0^t\lambda_u\mathrm{d}u+\sum_{l=1}^{N_u}J_l\Bigg)
\end{aligned}
$$

is an exponential martingale as a solution of a driftless SDE and given that $\mathbb{E}\left(\int_0^\tau\lambda_s\mathrm{d}s\right)<\infty$ and $\mathbb{E}\left(\int_0^\tau V_s^2\mathrm{d}s\right)<\infty$ for all $\tau\geq 0$.
i) is straightforward. Hence, we start by the proof of ii). The moment generating function of $\lambda_s^{(n),\star}$ conditionally to the information at time $t\leq s$ is given by

$$
\begin{aligned}
\mathbb{E}^\star\left(e^{\omega\lambda_s^{(n),\star}}|\mathscr{F}_t\right) &= e^{U_t^{(n)}}\mathbb{E}\left(e^{U_s^{(n)}+\omega\psi(1,0)\lambda_s^{(n)}}|\mathscr{F}_t\right),\\
&\qquad\exp\left(F(t,s)+G(t,s)\lambda_t^{(n)}+\sum_{l=-n}^{n}H_l\left(t,s,b_l^{(n)}\right)m_l^{(k,n)}Y_t^{(l)}\right)
\end{aligned}
$$

where

$$
\begin{aligned}
U_t^{(n)} &= -\frac{1}{2}\int_0^t V_u\mathrm{d}u+\int_0^t\sqrt{V_u}\left(\rho\mathrm{d}W_u^1+\sqrt{1-\rho^2}\mathrm{d}W_u^2\right)\\
&\qquad-\mathbb{E}\left(e^J-1\right)\int_0^t\lambda_u^{(n)}\mathrm{d}u+\sum_{l=1}^{N_u^{(n)}}J_l^{(n)}
\end{aligned}
$$

is the approximated exponent of the exponential martingale $\left.\frac{\mathrm{d}\mathbb{Q}^{\star}}{\mathrm{d}\mathbb{Q}}\right|_{\mathscr{F}_0}$.
If we note

$$f \equiv f\left(t,s,U_t^{(n)},\lambda_t^{(n)},\left(Y_t^{(l)}\right)_{l=-n,\dots,n},N_t^{(n)}\right) \quad := \quad \mathbb{E}\left(e^{U_s^{(n)}+\omega\psi(1,0)\lambda_s^{(n)}}|\mathscr{F}_t\right)$$

and we conjecture that f is an exponential affine function of $U_t^{(n)}$ $\lambda_t^{(n)}$ and $\left(Y_t^{(v)}\right)_{v=1,\dots,n}$
,

$$\begin{aligned} f \quad := \quad & \exp\left(A(t,s,\omega)+\psi(1,0)B(t,s,\omega)\lambda_t^{(n)}\right), \\ & \times\exp\left(\sum_{l=-n}^{n}C_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}Y_t^{(l)}+D(t,s,\omega)U_t^{(n)}\right) \end{aligned}$$

where $A(t,s,\omega)=0$, $B(s,s,\omega)=\omega$, $C_l\left(s,s,b_l^{(n)},\omega\right)=0$ for all $l\in\{-n,...,n\}$and $D(s,s,\omega)=1$. According to the Itô's lemma, f solves the following differential equation :

$$\begin{aligned} 0 \quad = \quad & \frac{\partial f}{\partial t}+\frac{\partial f}{\partial U^{(n)}}\left(-(\psi(1,0)-1)\lambda_t^{(n)}-\frac{1}{2}V_t\right) \\ & +\frac{1}{2}\frac{\partial^2 f}{\partial U^{(n)2}}V_t+i\sum_{l=-n}^{n}\frac{\partial f}{\partial Y^{(l)}}b_l^{(n)}Y_t^{(l)} \\ & +\frac{\partial f}{\partial\lambda^{(n)}}\left((\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t}+\eta\sum_{l=-n}^{n}m_l^{(k,n)}ib_l^{(n)}Y_t^{(l)}\right)+\lambda_t^{(n)}\int_{-\infty}^{\infty}\Delta f(z)\,\nu(\mathrm{d}z) \end{aligned} \tag{5.73}$$

where

$$\frac{\partial f}{\partial U^{(n)}}=D(t,s,\omega)f,\ \frac{\partial f}{\partial\lambda^{(n)}}=\psi(1,0)B(t,s,\omega)f,\ \frac{\partial f}{\partial Y^{(l)}}=C_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}f,$$

$$\frac{\partial^2 f}{\partial U^{(n)2}}=D(t,s,\omega)^2 f,\ \frac{\partial f}{\partial t}=\left(\frac{\partial A}{\partial t}(t,s,\omega)+\psi(1,0)\frac{\partial B}{\partial t}(t,s,\omega)\lambda_t^{(n)}\right)f$$

$$+\left(\sum_{l=-n}^{n}\frac{\partial C_l}{\partial t}\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}Y_t^{(v)}+\frac{\partial D}{\partial t}(t,s,\omega)U_t^{(n)}\right)f,$$

and the variation of f, $\Delta f(z)$, when a jump of size z occurs and the time just before is

$$\begin{aligned} \Bigg(\exp\Bigg(D(t,s,\omega)z+\Bigg(\eta\psi(1,0)B(t,s,\omega)\sum_{l=-n}^{n}m_l^{(k,n)} \\ +\sum_{l=-n}^{n}C_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}\Bigg)|z|\Bigg)-1\Bigg)f \end{aligned}$$

Injecting all these expressions into equation (5.73) lead to :

$$
\begin{aligned}
0 \;=\; & \frac{\partial A}{\partial t}(t,s,\omega)+\psi(1,0)\frac{\partial B}{\partial t}(t,s,\omega)\lambda_t^{(n)}+\sum_{l=-n}^{n}\frac{\partial C_l}{\partial t}\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}Y_t^{(v)} \\
& +\frac{\partial D}{\partial t}(t,s,\omega)U_t^{(n)}+D(t,s,\omega)\left(-(\psi(1,0)-1)\lambda_t^{(n)}-\frac{1}{2}V_t\right) \\
& +\frac{1}{2}D(t,s,\omega)^2V_t+i\sum_{l=-n}^{n}C_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}b_l^{(n)}Y_t^{(l)} \\
& +\psi(1,0)B(t,s,\omega)\left((\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t}+\eta\sum_{l=-n}^{n}m_l^{(k,n)}ib_l^{(n)}Y_t^{(l)}\right) \\
& +\lambda_t^{(n)}\Bigg(\psi\Bigg(D(t,s,\omega),\eta\psi(1,0)B(t,s,\omega)\sum_{l=-n}^{n}m_l^{(k,n)} \\
& \qquad +\sum_{l=-n}^{n}C_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}\Bigg)-1\Bigg)
\end{aligned}
$$

or

$$
\begin{aligned}
0 \;=\; & \frac{\partial A}{\partial t}(t,s,\omega)-\frac{1}{2}D(t,s,\omega)V_t+\frac{1}{2}D(t,s,\omega)^2V_t \qquad (5.74)\\
& +\psi(1,0)B(t,s,\omega)(\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t} \\
0 \;=\; & \psi(1,0)\frac{\partial B}{\partial t}(t,s,\omega)-D(t,s,\omega)(\psi(1,0)-1) \qquad (5.75)\\
& +\psi\Bigg(D(t,s,\omega),\eta\psi(1,0)B(t,s,\omega)\sum_{l=-n}^{n}m_l^{(k,n)} \\
& \qquad +\sum_{l=-n}^{n}C_l\left(t,s,b_l^{(n)},\omega\right)m_l^{(k,n)}\Bigg)-1, \\
0 \;=\; & \frac{\partial C_l}{\partial t}\left(t,s,b_l^{(n)},\omega\right)+i\left(C_l\left(t,s,b_l^{(n)},\omega\right)+\eta\psi(1,0)B(t,s,\omega)\right)b_l^{(n)} \\
& \forall l\in\{-n,\dots,n\} \qquad (5.76)\\
0 \;=\; & \frac{\partial D}{\partial t}(t,s,\omega). \qquad (5.77)
\end{aligned}
$$

$D(t,s,\omega)=1$ solves the last PDE as $D(s,s,\omega)=1$. Given $D(t,s,\omega)=1$, the functions A and B satisfy the following PDE system :

$$
\begin{cases}
\frac{\partial A}{\partial t}(t,s,\omega) & =-\psi(1,0)B(t,s,\omega)(\lambda_0-\alpha)\frac{\mathrm{d}g_k(t)}{\mathrm{d}t}, \\
\frac{\partial B}{\partial t}(t,s,\omega) & =1-\frac{\psi\left(1,\eta\psi(1,0)B(t,s,\omega)\sum_{l=-n}^{n}m_l^{(k,n)}+\sum_{l=-n}^{n}C_l\left(t,s,b_l^{(n)}\right)m_l^{(k,n)}\right)}{\psi(1,0)}.
\end{cases}
\qquad (5.78)
$$

Let $\lambda_0^\star=\psi(1,0)\lambda_0$, $\alpha^\star=\psi(1,0)\alpha$, $\eta^\star=\psi(1,0)\eta$ and

$$
\psi^\star(z_1,z_2) \;=\; \frac{\psi(z_1+1,z_2)}{\psi(1,0)}. \qquad (5.79)
$$

Taking into account these definitions, the latter ODE system is similar to the one obtained under risk neutral measure $\mathbb{Q}$ in proposition 11. On the other hand, for all

$l \in \{-n, ..., n\}$,

$$C_l\left(t, s, b_l^{(n)}, \omega\right) = ib_v^{(n)}\eta^\star \int_t^s e^{ib_l^{(n)}(u-t)}B(u,s,\omega)\,\mathrm{d}u,$$

solves equation 5.76. By a suitable rearrangement, we can show that this definition of the joint moment generating function $\psi^\star$ corresponds at the one of the pair $J^\star$and $|J^\star|$where the jump size $J^\star$ follows a double exponential with parameters

$$\begin{aligned}
\rho^{+,\star} &= \rho^+ - 1, \\
\rho^{-,\star} &= \rho^- - 1, \\
p^\star &= \frac{p\rho^+\rho^{-,\star}}{p\rho^+\rho^{-,\star} + \rho^{+,\star}(1-p)\rho^-}.
\end{aligned}$$

All the previous results allows us to complete the proof of ii). The SDE of the variance V_t under $\mathbb{Q}^\star$ is derived from i). While for the process X_t, similarly to the proof of ii), the dynamics of X under the measure $\mathbb{Q}^\star$ is obtained by comparison with the results of proposition 13.

5.D Appendix D: Expected value of $X_T \mid \mathscr{F}_t$

Given the dynamics of X_T under the measure $\mathbb{Q}$, we have

$$\begin{aligned}
\mathbb{E}(X_T \mid \mathscr{F}_t) &= \mathbb{E}\left(\int_t^T \left(r - \frac{1}{2}V_u\right) \mid \mathscr{F}_t\right) \\
&= r(T-t) - \frac{1}{2}\mathbb{E}\left(\int_t^T V_u \mathrm{d}u \mid \mathscr{F}_t\right),
\end{aligned} \tag{5.80}$$

where the solution of the SDE of the variance process is given by

$$V_T = \frac{\theta_V}{k_V} + \left(V_t - \frac{\theta_V}{k_V}\right)e^{-k_V(T-t)} + \sigma_V \int_t^T e^{-k_V(T-s)}\sqrt{V_s}\mathrm{d}W_s^1. \tag{5.81}$$

Applying the Fubini rule leads to

$$\begin{aligned}
\int_t^T V_u \mathrm{d}u &= \frac{\theta_V}{k_V}(T-t) + \left(V_t - \frac{\theta_V}{k_V}\right)\frac{1}{k_V}\left(1 - e^{-k_V(T-t)}\right) \\
&\quad + \frac{\sigma_V}{k_V}\int_t^u \sqrt{V_s}\left(e^{-k_V(t-s)} - e^{-k_V(T-s)}\right)\mathrm{d}W_s^1.
\end{aligned} \tag{5.82}$$

Hence

$$\begin{aligned}
\mathbb{E}(X_T \mid \mathscr{F}_t) &= r(T-t) - \frac{1}{2}\frac{\theta_V}{k_V}(T-t) \\
&\quad - \frac{1}{2}\left(V_t - \frac{\theta_V}{k_V}\right)\frac{1}{k_V}\left(1 - e^{-k_V(T-t)}\right)
\end{aligned} \tag{5.83}$$

Under the measure $\mathbb{Q}^\star$, the SDE of X_t and V_t are given by :

$$\begin{aligned}
\mathrm{d}X_t &= \left(r + \frac{1}{2}V_t - \lambda_t^\star \psi^\star(1,0)\right)\mathrm{d}t \\
&\quad + \sqrt{V_t}\left(\rho \mathrm{d}W_t^{\star 1} + \sqrt{1-\rho^2}\mathrm{d}W_t^{\star 2}\right) + \mathrm{d}\left(\sum_{l=1}^{N_t^\star} J_l^\star\right),
\end{aligned} \tag{5.84}$$

$$\mathrm{d}V_t = (\theta_V - (k_V - \rho\sigma_V)V_t)\mathrm{d}t + \sigma_V\sqrt{V_t}\mathrm{d}W_t^{\star 1}. \tag{5.85}$$

It follows that

$$
\begin{aligned}
\mathbb{E}^{\star}\left(X_T \mid \mathscr{F}_t\right) &= \mathbb{E}^{\star}\left(\int_t^T \left(r+\frac{1}{2}V_u\right) \mid \mathscr{F}_t\right) \\
&= r\left(T-t\right)+\frac{1}{2}\mathbb{E}^{\star}\left(\int_t^T V_u \mathrm{d}u \mid \mathscr{F}_t\right) \\
&= r\left(T-t\right)+\frac{1}{2}\frac{\theta_V}{k_V-\rho\sigma_V}\left(T-t\right) \\
&\quad +\left(V_t-\frac{\theta_V}{k_V-\rho\sigma_V}\right)\frac{1}{2\left(k_V-\rho\sigma_V\right)}\left(1-e^{-\left(k_V-\rho\sigma_V\right)\left(T-t\right)}\right).
\end{aligned}
\tag{5.86}
$$

where as under the measure $\mathbb{Q}$

$$
\begin{aligned}
\int_t^T V_u \mathrm{d}u &= \frac{\theta_V}{k_V-\rho\sigma_V}\left(T-t\right) \\
&\quad \left(V_t-\frac{\theta_V}{k_V-\rho\sigma_V}\right)\frac{1}{k_V-\rho\sigma_V}\left(1-e^{-\left(k_V-\rho\sigma_V\right)\left(T-t\right)}\right) \\
&\quad +\frac{\sigma_V}{k_V-\rho\sigma_V}\int_t^u \sqrt{V_s}\left(e^{-\left(k_V-\rho\sigma_V\right)\left(t-s\right)}-e^{-\left(k_V-\rho\sigma_V\right)\left(T-s\right)}\right)\mathrm{d}W_s^{\star 1}.
\end{aligned}
\tag{5.87}
$$

Chapter 6

Conclusion and perspectives

Much of the recent research in financial mathematics is concerned with extensions and alternatives that aim to improve the modelling of the financial market dynamics. In this thesis, we have been interested in the contagion patterns underlying financial products. We have dealt with this issue mainly by using different settings of self-exciting point processes of Hawkes type, and applying it to the pricing of derivatives.

The first research direction undertaken in Chapter 2 focused on the modelling of the interbank credit risk to explain deviations between equivalent interbank rates observed during and after the credit crunch of 2007. In this respect, we have segmented the interbank market according to tenors and modeled the credit risk of each segment using an intensity-based approach. Aiming to allow a default or an exogenous event to trigger other defaults within an interbank market segment, we have assumed the arrival rate of default is determined by a self-exciting jump diffusion process which is the intensity of the Hawkes process with an exponential kernel plus a Brownian term proportional to the square root of the intensity as in the Cox Ingersoll Ross process. Unlike the classical interest rate theory, in this set up, the forward Libor rate is a non-decreasing function of an interbank risk factor. The intensity processes proposed being Markov and Hawkes processes with an exponential kernel tractable, we found a new arbitrage free condition and derived closed-form approximation of pricing formulas for caplets, caps, as well as swaptions. After presenting the parameter estimation procedures of the proposed models to real market data, the fitted model enables us to reproduce the IRS curve and to explain the deviance between the IRS and OIS curve. Although, we found weak values for the estimated initial arrival rate of default and its estimated Brownian volatility that could lead to refute the presence of self-excitation, these values probably simply reflect the state of the market or that the market does not include this risk in the pricing. The sensitivity analyses performed confirmed the importance of the interbank credit risk to reflect the confidence level within the interbank market.

The matching of the level of self-excitation to the economic state led us to the second research direction developed in Chapter 3. In this chapter we considered a switching Hawkes process with exponential kernel and proposed a valuation framework for variable annuities contracts with a minimum death and accumulated life benefit. More precisely, we used a self-exciting switching jump diffusion model for the underlying investment fund price of variable annuities and the Hull-White model with correlation for the reference risk-free rate. In such a Markov framework, we derived the conditional moment generating function needed to approximate by Fourier inversion the call and put option prices that determine the liability of the issuer of the variable annuities. We ended this chapter by a numerical implementation divided into two parts. First, we provided a detailed economic calibration of the proposed models based on the historical data. Second, with the mortality model, we conducted numerical analyses that revealed the sensitivity of variable annuities to the economic regime and the self-excitation level.

The kernel function of the Hawkes processes defines the memory of the Hawkes process. For instance, if it is fast decreasing, the contagion effect between jumps is limited in time. When using an exponential kernel as in the first two chapters of this thesis, the Hawkes process has the Markov property that enables the use of Itô's stochastic calculus for computations. In Chapter 4, we took another direction by introducing an asset prices model which is a long memory self-exciting jump diffusion process. The self-excitation and the memory features of this process is provided by a Hawkes process with a Mittag-Leffler kernel which is the Laplace transform of a non-negative measure. These features enabled us to obtain a realistic option pricing model that encompass major characteristic of asset's return times series such as the volatility clustering and the long memory in volatility. Althougth, the Hawkes process with this Mittag-Leffler function is non-Markov we employed a similar procedure of the rough volatility literature (see e.g. Abi Jaber and El Euch 2019) to derive the conditional moment generating function of the Hawkes intensity and and the log-returns. While we lost in tractability with the Mittag-Leffler kernel, its memory feature allowed to produce required excess of kurtosis and skewness compared to the self-exciting Hawkes jump diffusion model with exponential kernel, the Kou jump diffusion model, the Merton jump diffusion model or the geometric Brownian motion model. It is worth noting that, in this set up, when the time horizon is big enough, the number of price jumps could explode as the Hawkes with the chosen Mittag-leffler function is non-stationary.

Finally, in Chapter 5, we expanded the Heston model by introducing self-excited jumps in asset prices and a long memory of past price jumps. This expansion corresponds to an asset price process with a stochastic volatility and a jump component driven by a Hawkes process. To ensure the stationary property of the Hawkes process we considered kernel functions that are Fourier transform of the Laplace, Gaussian, Logistic, and Cauchy measures. The last kernel is actually a kernel of exponential form. By doing this we generated different memory ranges, from the longest to the shortest. Except the stochastic volatility model with jumps self-excited by the exponential kernel, our Heston extensions are non Markov processes. As in Chapter 4 we used the Fourier transform representation of these kernel functions to establish a closed form expression of the conditional moment generating functions of the HP's intensity and the log-returns. We found that with equal parameters, the thickness of the tails of the probability density function of the log-returns is fatter if the length of the memory of the Hawkes process kernel function is greater. An estimation of our Heston extensions to implied volatilities of the Euro Stoxx achieved better calibration performance than the one with the standard Heston model, particularly in the short term. They also provided a good fit of the term structure of at-the-money volatility skew. Considering the estimated parameters found and varying the memory parameter, we saw that implied volatilities derived from our extensions of the Heston model , increase as the memory length expands. This sensitivity to the memory length appears more pronounced for out-of-the money call option prices. A similar conclusion was drawn, with respect to the sensitivity of the implies volatilities to other self-excitation parameters. The analysis of the Greeks delta, vega and the jump risk revealed that our extensions of the Heston model have roughly the same delta curve than the standard Heston. But these models induce different shapes of sensitivity to volatility and jump risks in the option price, which does not seemed to be related only to their memory lengths.

A lot remains to be done and the Hawkes process framework certainly forms the basis for many other applications in finance and actuarial science, beyond those we develop in this thesis. Other applications in finance would be the pricing under a Hawkes with non exponential framework of currency, exotic options or exotic derivatives which are path dependent. Exploring hedging strategies in the context of our options pricing models, to manage options whose underlying asset may experience some clustering of jumps. The current work applies Hawkes processes with non-exponential kernels for modelling of stock and option prices, but these processes could be used in insurance risk theory to model claims arrival which might display a clustering effect and long-term memory. The results of Swishchuk et al. (2021) could serve as a starting point for the application of a Hawkes process with a non-exponential kernel in this context. In general, we believe Hawkes processes with non-exponential kernels merit further theoretical exploration and detailed empirical application. Furthermore, in order to make full use of these processes in practice, it is of great interest to develop an estimation procedure that is less computationally expensive than those used in this work. The issue of change of measure is also worth studying, as it is relevant for the pricing of financial products.

Bibliography

Abi Jaber, Eduardo and Omar El Euch (2019). “Multifactor Approximation of Rough Volatility Models”. *SIAM Journal on Financial Mathematics* 10.2, pp. 309–349.

Aït-Sahalia, Yacine, Julio Cacho-Diaz, and Roger J.A. Laeven (2015). “Modeling financial contagion using mutually exciting jump processes”. *Journal of financial economics* 117.3, pp. 585–606.

Aït-Sahalia, Yacine, Chenxu Li, and Chen Xu Li (2021). “Closed-form implied volatility surfaces for stochastic volatility models with jumps”. *Journal of Econometrics* 222.1, pp. 364–392.

Aït-Sahalia, Yacine and Andrew W. Lo (1998). “Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices”. *The Journal of Finance* 53.2, pp. 499–547.

Albrecher, Hansjörg, Philipp A. Mayer, Wim Schoutens, and Jurgen Tistaert (2006). “The little Heston trap”. *Wilmott*, pp. 83–92.

Ametrano, Ferdinando M. and Marco Bianchetti (2013). “Everything You Always Wanted to Know About Multiple Interest Rate Curve Bootstrapping But Were Afraid To Ask”. *Working paper*.

Andersen, Torben G., Luca Benzoni, and Jesper Lund (2002). “An Empirical Investigation of Continuous-Time Equity Return Models”. *Journal of Finance* 57.3, pp. 1239–1284.

Babbs, S. and K. Nowman (1999). “Kalman Filtering of Generalized Vasicek Term Structure Models”. *Journal of Financial and Quantitative Analysis* 34.

Bacinello, Anna Rita, Pietro Millossovich, Annamaria Olivieri, and Ermanno Pitacco (2011). “Variable Annuities: A Unifying Valuation Approach”. *Insurance: Mathematics and Economics* 49, pp. 285–297.

Bacry, Emmanuel, Thibault Jaisson, and Jean–François Muzy (2016). “Estimation of slowly decreasing Hawkes kernels: application to high-frequency order book dynamics”. *Quantitative Finance* 16.8, pp. 1179–1201.

Bacry, Emmanuel and Jean-François Muzy (2014). “Hawkes model for price and trades high-frequency dynamics”. *Quantitative Finance* 14.7, pp. 1147–1166.

Bakshi, Gurdip, Charles Cao, and Zhiwu Chen (1997). “Empirical Performance of Alternative Option Pricing Models”. *The Journal of Finance* 52.5, pp. 2003–2049.

Barbarin, J. and P. Devolder (2005). “Risk measure and fair valuation of an investment guarantee in life insurance”. *Insurance: Mathematics and Economics* 37, pp. 297–323.

Barndorff-Nielsen, Ole E. and Neil Shephard (2001a). “Modelling by Lévy Processess for Financial Econometrics”. *Lévy Processes*, pp. 283–318.

– (2001b). “Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63.2, pp. 167–241.

Bates, David S. (1991). “The Crash of ’87: Was It Expected? The Evidence from Options Markets”. *The Journal of Finance* 46.3, pp. 1009–1044.

– (2000). “Post-'87 crash fears in the S&P 500 futures option market”. *Journal of Econometrics* 94.1-2, pp. 181–238.

Bauer, D., A. Kling, and J. Russ (2008). “A Universal Pricing Framework for Guaranteed Minimum Benefits in Variable Annuities”. *ASTIN Bulletin* 38.2, pp. 621–651.

Bäuerle, Nicole and Sascha Desmettre (2020). “Portfolio Optimization in Fractional and Rough Heston Models”. *SIAM Journal on Financial Mathematics* 11.1, pp. 240–273.

Bernard, Carole, Mary Hardy, and Anne Mackay (2014). “STATE- DEPENDENT FEES FOR VARIABLE ANNUITY GUARANTEES”. *ASTIN Bulletin* 44.3, pp. 559–585.

Bernard, Carole and Thorsten Moenig (2018). “Where Less Is More: Reducing Variable Annuity Fees to Benefit Policyholder and Insurer”. *Journal of Risk and Insurance* 86.3, pp. 761–782.

Bernis, Guillaume, Riccardo Brignone, Simone Scotti, and Carlo Sgarra (2021). “A Gamma Ornstein–Uhlenbeck model driven by a Hawkes process”. *Mathematics and Financial Economics* 15.4, pp. 747–773.

Bianchetti, Marco (2008). “Two Curves, One Price: Pricing Hedging Interest Rate Derivatives Decoupling Forwarding and Discounting Yield Curves”. *SSRN Electronic Journal*.

Black, Fischer and Myron Scholes (1973). “The Pricing of Options and Corporate Liabilities”. *Journal of Political Economy* 81.3, pp. 637–654.

Blanc, P., J. Donier, and J.-P. Bouchaud (2016). “Quadratic Hawkes processes for financial prices”. *Quantitative Finance* 17.2, pp. 171–188.

Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner (1992). “ARCH modeling in finance”. *Journal of Econometrics* 52.1-2, pp. 5–59.

Boyle, P. P. and E. S. Schwartz (1977). “Equilibrium prices of guarantees under equity-linked contracts”. *The Journal of Risk and Insurance* 44.4, p. 639.

Brennan, M. J. and E. S. Schwartz (1976). “The pricing of equity-linked life insurance policies with an asset value guarantee”. *Journal of Financial Economics* 3.3, pp. 195–213.

– (1979). “Alternative investment stragtegies for the issuers of equity-linked life insurance with an asset value guarantee”. *Journal of Business* 52, pp. 63–93.

Brignone, Riccardo and Carlo Sgarra (2019). “Asian options pricing in Hawkes-type jump-diffusion models”. *Annals of Finance* 16.1, pp. 101–119.

Brigo, Damiano and Fabio Mercurio (2006). *Interest Rate Models Theory and Practice*. Springer Finance. Springer Berlin Heidelberg.

Carr, Peter and Dilip B. Madan (1999). “Option valuation using the fast Fourier transform”. *Journal of computational finance* 2, pp. 61–73.

Carr, Peter and Liuren Wu (2003). “What Type of Process Underlies Options? A Simple Robust Test”. *Journal of Finance* 58.6, pp. 2581–2610.

Chen, J., Alan G. Hawkes, and E. Scalas (2021). “A Fractional Hawkes Process”. *Nonlocal and Fractional Operators*. Ed. by Luisa Beghin, Francesco Mainardi, and Roberto Garrappa. Vol. 26. Springer International Publishing, pp. 121–131.

Christoffersen, Peter, Kris Jacobs, and Karim Mimouni (2006). “An Empirical Comparison of Affine and Non-Affine Models for Equity Index Options”. *SSRN Electronic Journal*.

Comte, F., L. Coutin, and E. Renault (2010). “Affine fractional stochastic volatility models”. *Annals of Finance* 8.2-3, pp. 337–378.

Cont, Rama (2001). “Empirical properties of asset returns: stylized facts and statistical issues”. *Quantitative finance* 1.2, p. 223.

Coonjobeharry, Radha K., Désiré Y. Tangman, and Muddun Bhuruth (2016). “A two-factor jump-diffusion model for pricing convertible bonds with default risk”. *International Journal of Theoretical and Applied Finance* 19.06.

Cox, John. C., Jr. Jonathan E. Ingersoll, and Stephen A. Ross (1985). “A theory of the term structure of interest rates”. *Econometrica* 53, pp. 385–408.

Cui, Zhenyu, Runhuan Feng, and Anne MacKay (2017). “Variable Annuities with VIX-Linked Fee Structure under a Heston-Type Stochastic Volatility Model”. *North American Actuarial Journal* 21.3, pp. 458–483.

Dandapani, Aditi, Paul Jusselin, and Mathieu Rosenbaum (2021). “From quadratic Hawkes processes to super-Heston rough volatility models with Zumbach effect”. *Quantitative Finance* 21.8, pp. 1235–1247.

Dassios, Angelos and Hongbiao Zhao (2017). “A generalised contagion process with an application to credit risk”. *International Journal of Theoretical and Applied Finance* 20.01, p. 1750003.

Duffie, D., D. Filipović, and W. Schachermayer (2003). “Affine processes and applications in finance”. *The Annals of Applied Probability* 13.3.

Duffie D. Pan J., Singleton K. (2000). “Transform analysis and asset pricing for affine jump diffusions”. *Econometrica* 68, pp. 1343–1376.

Eberlein, Ernst, Christoph Gerhart, and Zorana Grbac (2019). “Multiple curve Lévy forward price model allowing for negative interest rates”. *Mathematical Finance* 30.1, pp. 167–195.

EL Euch, Omar, Masaaki Fukasawa, and Mathieu Rosenbaum (2018). “The microstructural foundations of leverage effect and rough volatility”. *Journal of Finance and Stochastics* 22, pp. 241–280.

Elliott, R. J., L. Chan, and T. K. Siu (2005). “Option pricing and Esscher transformation under regime regime switching”. *Annals of Finance* 1.4, pp. 423–432.

Eraker, Bjørn (2004). “Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Option Prices”. *Journal of Finance* 59.3, pp. 1367–1403.

Erdélyi, A., W. Magnus, F. Oberhettinger, and F. Tricomi (1955). *Higher Transcendental Functions*. Vol. 3. McGraw-Hill.

Errais, Eymen, Kay Giesecke, and Lisa R. Goldberg (2010). “Affine Point Processes and Portfolio Credit Risk”. *SIAM Journal on Financial Mathematics* 1.1, pp. 642–665.

Fan, Kun, Yang Shen, Tak Kuen Siu, and Rongming Wang (2015). “Pricing Annuity Guarantees under a Double Regime-switching Model”. *Insurance: Mathematics and Economics* 62, pp. 62–78.

Feng, Runhuan and Hans W. Volkmer (2012). “Analytical calculation of risk measures for variable annuity guaranteed benefits”. *Insurance: Mathematics and Economics* 51.3, pp. 636–648.

Filipovíc, Damir and Anders B. Trolle (2012). “The Term Structure of interbank Risk”. *Working paper*.

Fonseca, José Da and Riadh Zaatour (2013). “Hawkes Process: Fast Calibration, Application to Trade Clustering and Diffusive Limit”. *SSRN Electronic Journal*.

Fontana, Claudio and Wolfgang J. Runggaldier (2010). “Credit risk and incomplete information: Filtering and em parameter estimation”. *International Journal of Theoretical and Applied Finance* 13.05, pp. 683–715.

Fries, Christian P. (2016). “Multi-curve Construction”. *Innovations in Derivatives Markets*. Springer International Publishing, pp. 227–250.

Gatheral, Jim (2012). *The Volatility Surface*. Wiley.

Gatheral, Jim, Thibault Jaisson, and Mathieu Rosenbaum (2014). “Volatility is Rough”. *SSRN Electronic Journal*.

GIL-PELAEZ, J. (1951). “Note on the inversion theorem”. *Biometrika* 38.3-4, pp. 481–482.

Goldfeld, M. Stephen, and Richard E. Quandt (1973). “A Markov model for switching regressions”. *Journal of econometrics* 1.1, pp. 3–15.

Gompertz, Benjamin (1825). “On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies.” *Philosophical Transactions of the Royal Society of London* 115, pp. 513–583.

Gorenflo, R. and F. Mainardi (1997). “Fractional Calculus”. *Fractals and Fractional Calculus in Continuum Mechanics*. Springer Vienna, pp. 223–276.

Gorenflo, Rudolf, Anatoly A. Kilbas, Francesco Mainardi, and Sergei V. Rogosin (2014). *Mittag-Leffler Functions, Related Topics and Applications*. Springer Berlin Heidelberg.

Grbac, Zorana and Wolfgang J. Runggaldier (2015). *Interest Rate Modeling: Post-Crisis Challenges and Approaches*. Springer International Publishing.

Hainaut, D. and D. Colwell (2016). “A structural model for credit risk with Markov modulated Lévy processes and synchronous jumps.” *European Journal of Finance* 22.11, pp. 1040–1062.

Hainaut, Donatien (2014). “Impulse control of pension fund contributions, in a regime switching economy.” *European Journal of Operational Research* 239, pp. 810–819.

– (2016a). “A bivariate Hawkes process for interest rate modeling”. *Economic Modelling* 57, pp. 180–196.

– (2016b). “A model for interest rates with clustering effects”. *Quantitative finance* 16 (8), pp. 1203–1218.

– (2017). “Clustered Lévy processes and their financial applications”. *Journal of Computational and Applied Mathematics* 319, pp. 117–140.

– (2021). “Moment generating function of non-Markov self-excited claims processes”. *Insurance: Mathematics and Economics* 101, pp. 406–424.

Hainaut, Donatien and Franck Moraux (2018 a). “Hedging of options in the presence of jump clustering”. *Journal of Computational Finance*.

– (2018 b). “A switching self-exciting jump diffusion process for stock prices”. *Annals of Finance*.

Hamida, Sana Ben and Rama Cont (2005). “Recovering volatility from option prices by evolutionary optimization”. *The Journal of Computational Finance* 8.4, pp. 43–76.

Hamilton, J.D. (1989). “A new approach to economic analysis of non-stationnary time series and the business cycle”. *Econometrica* 57, pp. 357–384.

Hardy, M. R. (2001). “A Regime-Switching Model of Long-Term Stock Returns”. *North American Actuarial Journal* 5.2, pp. 41–53.

Hardy, Mary (2003). *Investment Guarantees: Modeling and Risk Management for Equity-Linked Life Insurance*. Ed. by John Wiley & Sons. Wiley Finance.

Haubold, H. J., A. M. Mathai, and R. K. Saxena (2011). “Mittag-Leffler Functions and Their Applications”. *Journal of Applied Mathematics* 2011, pp. 1–51.

Hawkes, Alan G. (1971 a). “Point spectra of some mutually exciting point processes”. *Journal of the Royal Statistical Society Series B* 33, pp. 438–443.

– (1971 b). “Spectra of some self-exciting and mutually exciting point processes”. *Biometrika* 58, pp. 83–90.

Henrard, Marc (2014). *Interest Rate Modelling in the Multi-Curve Framework : Foundations, Evolution and Implementation*. Palgrave macmillan.

Heston, Steven L. (1993). “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. *Review of Financial Studies* 6.2, pp. 327–343.

Hubalek, Friedrich, Martin Keller-Ressel, and Carlo Sgarra (2016). “Geometric Asian option pricing in general affine stochastic volatility models with jumps”. *Quantitative Finance* 17.6, pp. 873–888.

Hull, John and Alan White (1990). “Pricing interest-rate derivatives securities”. *Review of Financial Studies* 3, pp. 573–592.

Hull, John and Alan White (2015). “OIS discounting, interest rate derivatives, and the modeling of stochastic interest rate spreads”. *Journal of investment management* 13.1, pp. 64–83.

Hurd, T. R. (2018). “BANK PANICS AND FIRE SALES, INSOLVENCY AND ILLIQUIDITY”. *International Journal of Theoretical and Applied Finance* 21.06, p. 1850040.

Ignatieva, K., Andrew Song, and J. Ziveyi (2016). “Pricing and Hedging of Guaranteed Minimum Benefits under Regime-Switching and Stochastic Mortality”. *Insurance: Mathematics and Economics* 70, pp. 286–300.

Ignatieva, Katja, Paulo Rodrigues, and Norman Seeger (2009). “Stochastic Volatility and Jumps: Exponentially Affine Yes or No? An Empirical Analysis of S&P500 Dynamics”. *SSRN Electronic Journal*.

Jaisson, Thibault and Mathieu Rosenbaum (2015). “Limit theorems for nearly unstable Hawkes processes”. *The Annals of Applied Probability* 25.2.

Kalman, R. E. (1960). “A New Approach to Linear Filtering and Prediction Problems”. *Journal of Basic Engineering* 82.1, pp. 35–45.

Kijima, M. and T. Wong (2007). “Pricing of Ratchet Equity-indexed Annuities under Stochastic Interest Rates.” *Insurance: Mathematics and Economics* 41.3, pp. 317–338.

Kou, S. G. (2002). “A Jump-Diffusion Model for Option pricing”. *Management Science* 48.8, pp. 1086–1101.

Landriault, David, Bin Li, Dongchen Li, and Yumin Wang (2021). “High-water mark fee structure in variable annuities”. *Journal of Risk and Insurance*.

Lin, Sheldon, Tan Ken Seng, and Hailiang Yang (2009). “Pricing Annuity Guarantees Under a Regime-Switching Model”. *North American Actuarial Journal* 13.3, pp. 316–332.

Lin, X. S. and K. S. Tan (2003). “Valuation of equity-indexed Annuities under Stochastic Interest Rates”. *North American Actuarial Journal* 7.4, pp. 72–91.

Maheu, John M. and Thomas H. McCurdy (2004). “News Arrival, Jump Dynamics, and Volatility Components for Individual Stock Returns”. *The Journal of Finance* 59.2, pp. 755–793.

Mainardi, Francesco (2020). “Why the Mittag-Leffler Function Can Be Considered the Queen Function of the Fractional Calculus?” *Entropy* 22.12, p. 1359.

Matthew, Makeham William (1860). “On the Law of Mortality and the Construction of Annuity Tables.” *The Assurance Magazine, and Journal of the Institute of Actuaries* 8.6, pp. 301–310.

Melino, Angelo and Stuart M. Turnbull (1990). “Pricing foreign currency options with stochastic volatility”. *Journal of Econometrics* 45.1-2, pp. 239–265.

Mercurio, Fabio (2009). “Interest Rates and The Credit Crunch: New Formulas and Market Models”. *Bloomberg: Frontiers (Topic)*.

Merton, Robert C. (1976). “Option pricing when underlying stock returns are discontinuous”. *Journal of Financial Economics* 3.1-2, pp. 125–144.

Milvesky, M. and S. E. Posner (2001). "The Titanic Option: Valuation of the Guaranteed Minimum Death Benefit in Variable Annuities and Mutual Funds". *Journal of Risk and Insurance* 68.1, pp. 91–126.

Milvesky, M. and T. S. Salisbury (2006). "Financial Valuation of the Guaranteed Minimum Withdrawal Benefits". *Insurance: Mathematics and Economics* 38, pp. 21–38.

Mittag-Leffler, G. M. (1903). "Sur la nouvelle fonction $E_\alpha(x)$". *C.R. Acad. Sci. Paris* (ser. II).137, pp. 554–558.

Morini, Massimo (2009). "Solving the Puzzle in the Interest Rate Market". *SSRN Electronic Journal*.

Nelson, Charles R. and Andrew F. Siegel (1987). "Parsimonious Modeling of Yield Curves". *The Journal of Business* 60.4, p. 473.

Njike, Charles Guy Leunga and Donatien Hainaut (2020). "Interbank credit risk modeling with self-exciting jump processes". *International Journal of Theoretical and Applied Finance* 23.06, p. 2050039.

– (2022). "Valuation of Annuity Guarantees Under a Self-Exciting Switching Jump Model". *Methodology and Computing in Applied Probability*.

Ogata, Yosihiko (1998). "Space-Time Point-Process Models for Earthquake Occurrences". *Annals of the Institute of Statistical Mathematics* 50.2, pp. 379–402.

Pagan, Adrian (1996). "The econometrics of financial markets". *Journal of Empirical Finance* 3.1, pp. 15–102.

Revuz, Daniel and Marc Yor (1999). *Continuous Martingales and Brownian Motion*. Springer Berlin Heidelberg.

Schöenbucher, Philipp J. (2001). "A Libor Market Model with Default Risk". *SSRN Electronic Journal*.

Schoutens, Wim, Erwin Simons, and Jurgen Tistaert (2004). "A perfect calibration! now what?" *Wilmott* 2004, pp. 66–78.

Siu, T.K. (2005). "Fair Valuation of Participating Policies with surrender Options and Regime Regime Switching". *Insurance: Mathematics and Economics* 37.3, pp. 553–552.

Smith, Donald J. (2013). "Valuing Interest Rate Swaps UsingOvernight Indexed Swap (OIS) Discounting". *The Journal of Derivatives* 20.4, pp. 49–59.

Swishchuk, Anatoliy, Rudi Zagst, and Gabriela Zeller (2021). "Hawkes processes in insurance: Risk model, application to empirical data and optimal investment". *Insurance: Mathematics and Economics* 101, pp. 107–124.

Tankov, Peter (2003). *Financial Modelling with Jump Processes*. Chapman and Hall/CRC.

Uhlenbeck, George Eugene and Leonard Salomon Ornstein (1930). "On the Theory of the Brownian Motion". *Physical Review* 36.5, pp. 823–841.

Vasicek, Oldrich (1977). "An equilibrium characterization of the term structure". *Journal of financial economics* 5.2, pp. 177–188.

Widder, David Vernon (1946). *Laplace transform*. Princeton university press.

Witkovský, Viktor (2001). “On the exact computation of the density and of the quantiles of linear combinations of t and F random variables”. *Journal of Statistical Planning and Inference* 94.1, pp. 1–13.

Zumbach, Gilles, Luis Fernández, and Caroline Weber (2013). “Processes for stocks capturing their statistical properties from one day to one year”. *Quantitative Finance* 14.5, pp. 849–861.

Zumbach, Gilles and Paul Lynch (2001). “Heterogeneous volatility cascade in financial markets”. *Physica A: Statistical Mechanics and its Applications* 298.3-4, pp. 521–529.