Università degli Studi di Padova

Department of Mathematics "Tullio Levi-Civita"
Doctoral Program in Mathematics
Curriculum: Computational Mathematics
XXXIII Cycle

TOPICS IN FINANCIAL MATHEMATICS

Direttore della Scuola: Ch.mo Prof. Martino Bardi

Supervisor: Ch.mo Prof. Giorgia Callegaro

Co-supervisor: Ch.mo Prof. Martino Grasselli

Candidate: Andrea Mazzoran

Abstract

This Dissertation is devoted to the study of three different problems in mathematical finance: the first refers to energy markets, the second is related to a smile modelling problem and the third is relative to stochastic analysis and PDEs.

In the first chapter we give an introduction to the main topics covered in the Thesis and the main achievements.

The second chapter deals with a class of self-exciting models for the forward prices, arising in power markets. In order to capture the whole forward curve evolution a Heath-Jarrow-Morton approach is adopted. In particular we propose and investigate two model classes for forward power price dynamics, based on continuous branching processes with immigration, and on Hawkes processes with exponential kernel, respectively, with an application to the French power futures market.

The third chapter provides a general framework for a joint calibration of a general Exchange-Traded-Product (ETP) and its underlying futures contracts. We apply this methodology to the specific case of VXX Exchange-Traded-Notes (ETN) and VIX futures. Thanks to a stochastic-local volatility model and an application of the Markovian projection method, we are able to outline a general framework to perform a calibration on VXX and VIX futures in a parsimonious manner. Moreover, we discuss numerical calibration results on real-market data by highlighting the impact of model parameters on implied volatilities.

Finally, in the fourth chapter, we adopt the Lie symmetries approach framework to investigate a family of boundary conditions for PDEs associated to some barrier contracts in mathematical finance. We adopt the Black-Scholes framework to solve some barrier contract problems, explicitly solving the corresponding PDEs and computing the transition densities in some particular cases.

Sommario

Questa Tesi si occupa dello studio di tre problemi nell'ambito della finanza matematica: il primo proviene dal mercato dell'energia, il secondo da problemi di modelizzazione dello smile e il terzo è un problema relativo all'analisi stocastica e alle PDE.

Nel primo capitolo diamo un'introduzione generale sugli argomenti trattati nella Tesi e i principali risultati ottenuti.

Nel secondo capitolo proponiamo una classe di modelli self-exciting per il prezzo forward, che sorgono nel mercato energetico. Per riuscire a descrivere in maniera consistente l'evoluzione della curva forward, adottiamo l'approccio di Heath-Jarrow-Morton. In particolare, proponiamo e studiamo due modelli per la dinamica dei prezzi forward, uno basato sui processi continui di branching con immigrazione e l'altro su processi di Hawkes con kernel esponenziale. Questi modelli verranno poi testati sul mercato energetico francese dei futures.

Nel terzo capitolo proponiamo un framework generale dove poter attuare una calibrazione congiunta di un Exchange-Traded-Product (ETP) e dei suoi contratti futures sottostanti. Applichiamo questa metodologia al caso specifico dell'Exchange-Traded-Note (ETN) VXX e ai futures sul VIX. Grazie a un modello a volatilità locale stocastica e a un'applicazione del metodo della proiezione markoviana, siamo in grado di proporre un framework dove poter attuare una calibrazione congiunta del VXX e dei futures sul VIX in maniera parsimoniosa. Inoltre vengono discussi i risultati numerici su opzioni di mercato, evidenziando l'impatto che hanno i parametri di modello sulla volatilità implicita.

Infine, nel quarto capitolo, grazie alle tecniche provenienti dall'ambito delle simmetrie di Lie, studiamo alcune PDE con condizioni al bordo associate a dei contratti a barriera che si incontrano in finanza matematica. Adottando il modello di Black-Scholes, siamo in grado di risolvere esplicitamente le PDE corrispondenti e in alcuni casi siamo in grado di calcolare la densità di transizione.

Alla mia bellissima famiglia

Acknowledgements

It is unlikely that I could fully express my gratitude to all the people that positively impacted my journey throughout my academic career. Despite this, I will try my best.

First and foremost, I would like to express my deepest and warmest gratitude to my supervisors Prof. Giorgia Callegaro and Prof. Martino Grasselli. In the last three years, Giorgia and Martino have been my supervisors, my mentors and my advisors and I am most thankful for they have also become trustable friends. Giorgia relentlessly encouraged me, especially at the beginning of this adventure. She was always enthusiastic to help in any way she could throughout the whole process. She pushed me towards new experiences and challenges – attending conferences or joining workshops – in order to put myself and my research out there. Each of these challenges has marked my growth as a scientist as well as a human being. I hope she will continue inspiring me as she has done so far. Martino has always been willing to share some of his time with me. Not even once has he refused me his help, time or advice. His insight and knowledge of the subject steered me through this research. Our inspiring conversations represented the starting point from which I have been able to improve my critical thinking and value the importance of having multiple perspectives. Throughout the last three years, my supervisors have been an invaluable resource of advice, continuous support and patience. Their deep knowledge and plentiful experience have encouraged me in my studies and academic research.

I would also like to express my sincere appreciation to Prof. Carlo Sgarra, with whom I had the pleasure to start a fruitful collaboration. Professor Sgarra brought his insightful view into our work, adding an essential resource of knowledge to the project. The friendly, familiar and exciting environment that he created will always be a sweet memory of mine.

Many thanks also to Prof. Simone Scotti, for his precious advice and willingness to help. He gave me the priceless opportunity to participate to "my very first conference" - the Data Science and Finance Conference in Siem Reap (Cambodia) – where I had the chance to give my first research talk. Both Simone and Carlo's help in preparing the presentation was invaluable. Thank you.

A special mention to Dr. Andrea Pallavicini, for he has given me the chance to work with a brilliant professional. His work ethic, rigour and thoroughness on his job were impressive. His remarkable expertise in the field and technical support have been an essential addition to my research journey.

I must thank Prof. Mark Craddock as well, as he introduced me to the fascinating world of Lie Symmetries. The field has intrigued me from the very beginning.

My scientific and personal growth throughout the last three years have been the results of

the amazing opportunity to work with brilliant, experienced and knowledgeable scientists from different backgrounds.

My studies and my life during this Ph.D. would have never been the same without the kind support, friendship and encouragement of all my colleagues at the Department of Mathematics and all the members of the Mathematical Finance group at the University of Padova. To them goes all my sincere gratefulness since they have made this journey wonderful and enjoyable.

Finally, but not less importantly, I must thank all the members of my amazing and beloved family. Since the very beginning, they have encouraged every choice I made during this pathway, feeding me with support, love and sympathy. Surely, I would have not achieved this goal of mine without them.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and main achievements

As the title suggests, this Thesis deals with some problems arising in the mathematical finance environment.
In the second chapter, we shall start with a comparison of two models class arising in power markets, one based on *Continuous Branching process with Immigration* (CBI) and the other one on *Hawkes process with exponential kernel*. We then test these two models on real market data, outlining a parameters estimation procedure and a jumps detection analysis, since both models exhibit jumps clustering features. We conclude with statistical tests to show which one of the two classes is more suitable for our modelling description.
Then, inspired by recent developments in commodity smile modelling, in the third chapter we provide a general framework for a joint calibration of an *Exchange Traded Product* (ETP) and its underlying futures contracts. We analyze the VIX futures market with a focus on the *Exchange Traded Notes* (ETN) written on such contracts, in particular, we investigate the VXX notes tracking the short-end part of the futures term structure. Thanks to an application of *Gyöngy Lemma* and to a tailor-made *stochastic-local volatility* model, we are able to provide a framework for a joint calibration of plain vanilla options both on VIX futures and VXX.
Finally, in the fourth chapter, we shall deal with some problem arising in the abstract pricing framework, more precisely, we investigate a family of boundary conditions for PDEs associated to some barrier contracts in mathematical finance. We apply different techniques coming from the Lie Symmetries approach, that involve, among others, Laplace-based inversion technique and Sine and Cosine transforms, and we exploit the Lie symmetries of the corresponding PDEs, characterizing, in some cases, the density function of the corresponding process.
We now present a brief description of the contents of each of the three chapters of the Thesis. At the beginning of each chapter, however, one may refer to the introductory sections for a more comprehensive discussion as well as for more references to the related literature.

Overview of Chapter 2

Energy markets are commodity markets that deal specifically with the trade and supply of energy. Energy market may refer to a power market, but can also relate to other sources of energy, like oil and natural gas. We are going to focus on power since we are working on it in Chapter 2, even though the characteristics we are going to outline relate to the other energy-related commodities as well. Power prices show empirical features that are not common to classical financial markets. As a consequence of its non-storability, power cannot be stored in warehouses like the great majority of the other commodities, and at any moment demand has to be met by power produced at the same time. Therefore, power prices are primarily driven by the equilibrium of demand and supply. The demand can generally depend on many factors, such as on temperature and on sunlight and therefore power prices follow seasonal patterns at different time scales. This implies that they tend to oscillate around a seasonally varying equilibrium level. Uncertainty in prices comes also from the supply-side. In fact, production costs vary substantially between different types of installation: at one extreme, wind, sun and hydropower are practically at zero cost while gas turbines are at the other end of the scale. As soon as demand increases, more expensive plants enter into production and unexpected outages or disruption in transmission make the supply-demand curve shift upwards (positive spikes) and prices jump, even in a considerable way, followed by a fast decrease to an average level as soon as the shock is absorbed by the network. In a similar manner, sharp downward movements (negative spikes) can happen as well, thanks, for example, to the combination of low demand with particularly favorable weather conditions (e.g. strong wind or strong sun radiation) for energy renewable sources. For a more general treatment to energy markets, the reader can refer to Benth et al. (2008).
These stylized features imply that historical series of both futures and spot prices include seasonality, mean-reversion, spikes and small fluctuations.

When it comes to describe the power price dynamics, one can opt for modelling either the spot or the forward price. In the former case, we mention the paper by Jiao et al. (2019), where a model for power spot prices based on Continuous Branching Processes with Immigration (CBI) was proposed and where the forward prices were computed with respect to a suitable structure preserving equivalent martingale measure. We also mention the work of Eyjolfsson and Tjøstheim (2018), where the authors describe a class of Hawkes processes and present an empirical investigation based on data from UK power market supporting Hawkes-type models for spot prices.
In the latter case, we cite the work of Benth et al. (2019) and Filimonov et al. (2014), who attempt to provide a description of the whole forward curves dynamics, focusing on Heath-Jarrow-Morton type models. We want to mention also the work of Benth and Paraschiv (2018), where the authors propose a random field approach based on Gaussian random fields by adopting the Musiela parametrization in order to describe the forward curve dynamics.

Our work focus on modelling the forward prices. The main contribution and novelty of it is to investigate the performances of two self-exciting model classes, namely the CBI processes and the Hawkes processes, in describing the forward price dynamics in power markets. The reasons behind the choice of these two class of models is their natural predisposition to describe the stylized factors arising in power markets we have summarized earlier.

In the first part of Chapter 2, we introduce and investigate the aforementioned models for forward power price dynamics. For both of these model we assume that the price of a forward at time t with maturity T is described by an additive model defined as

$$f(t,T) = \Lambda(t) - \Lambda(0) + \sum_{i=1}^{n} X_i(t,T),$$

where $\Lambda(t)$ is a deterministic function, the so-called *seasonality function*, and the $X_i(t,T)$ terms satisfy a stochastic differential equation that depends on the model we consider. We deliberately skip the technical details here, but they are provided in Section 2.2. In the CBI-based case, each of $X_i(t,T)$ term satisfies

$$\begin{aligned} X_i(t,T) = X_i(0,T) - \int_0^t a_i X_i(s,T)ds + \sigma_i \int_0^t \int_0^{X_i(s,T)} W_i(ds,du) + \\ + \gamma_i \int_0^t \int_0^{X_i(s-,T)} \int_{\mathbb{R}_+} z\widetilde{N}_i(ds,du,dz), \end{aligned}$$

where a_i, σ_i and γ_i are constant model parameters, $W(ds,du)$ is a white noise on $\mathbb{R}_+^2$ with unit covariance and $\widetilde{N}(ds,du,d\zeta)$ is an independent compensated Poisson random measure on $\mathbb{R}_+^3$ with intensity $dsdu\mu(d\zeta)$, with $\mu(d\zeta)$ being a Lévy measure on $\mathbb{R}_+$ and satisfying $\int_0^\infty (\zeta \wedge \zeta^2)\mu(d\zeta) < \infty$.
In the Hawkes-based case, instead, each $X_i(t,T)$ term satisfies

$$X_i(t,T) = X_i(0,T) - \int_0^t c_i X_i(s,T)ds + \int_0^t \sigma_i \sqrt{X_i(s,T)} dW_i(s) + \int_0^t \int_0^\infty z\tilde{J}_i(dz,ds),$$

where c_i and γ_i are constant model parameters and $\tilde{J}_i(dz,ds)$ are compensated marked point process with intensity $\lambda_i(t)$, satisfying the SDE:

$$\lambda_i(t) = \lambda_i(0) - \beta_i \int_0^t \lambda_i(s)ds + \alpha_i \int_0^t \int_0^\infty z J_i(ds,dz),$$

where, $\beta_i > 0$ is the rate of exponential decay of the influence of previous jumps on the intensity level and α the amplitude of the memory kernel of each factor X_i. We assume the jump size distributed according to an exponential density with parameter δ_i for each (λ_i, X_i), so we can write:

$$\tilde{J}_i(ds,dz) = J_i(ds,dz) - \lambda_i(s)\mu(dz)ds = J_i(ds,dz) - \lambda_i(s)\delta_i \exp(-\delta_i z)(dz)ds.$$

We also outline the main properties and features of these two model and the reasons why they are suitable for a modelling point of view.

In the second part of the chapter, we test the two models against electricity futures daily data in the French power market. Futures are fundamental hedging instruments for producers and consumers, as well as trading instruments for financial speculators. Electricity futures guarantee the

delivery of a given amount of power for a specified future time period, for example a week (W), a month (M), a quarter (Q) or a whole calendar year (Y). Some contracts prescribe physical delivery, while others are financially settled. This requires a non-standard set of arbitrage relations that must hold among contracts with overlapping delivery periods. For example, one can buy a futures for delivery of power over the first quarter of the year, i.e. January, February and March or one can buy three monthly futures for delivery over each corresponding month. Notably, futures constitute almost the total of regulated derivative trading.

Since power is a flow commodity, instantaneous forward contracts are not directly traded on the market, but futures are, as previously stated. Therefore, we need to be able to collect information included in the market futures prices in order to describe accordingly the forward dynamics. Representing forward prices by one continuous term structure curve is a necessary input when modelling the forward price dynamics, thus one needs methods to extract a smooth curve from quoted prices. The literature on possible methods to do so is vast. We adopt and modify the approach followed by Benth et al. (2008), whose idea was basically to assume that the forward curve can be represented by the sum of two continuous functions, namely

$$f(u) = \Lambda(u) - \Lambda(0) + \varepsilon(u),$$

where $\Lambda(u)$ is the seasonal function that we suppose to be described by a trigonometric function and $\varepsilon(u)$ is a smooth function with prescribed properties, called *adjustment function*.

Part of the work is devoted to the implementation of the algorithm to extract the forward dynamics curves from the futures prices. This required a strong pre-processing procedure to address the issue of overlapping contracts.

After detecting the jumps and their distribution via a iterative weighted least square algorithm, we propose a parameters estimation method for both models. Then, we estimate the intensity of both models. In the CBI case, it is known that the intensity is proportional to the process itself, while in the Hawkes case, a more careful approach was needed. Indeed, in this latter model, we rely on the widely known paper of Ozaki (1979), where a maximum likelihood estimation procedure of Hawkes' self-exciting point process was presented alongside with an explicit formula for the log-likelihood. We conclude with a statistical analysis, mainly relying on the *Kolmogorov-Smirnov test*, to check the adequacy of the two models to describe the observed forward prices evolution. A result of Lallouache and Challet (2016) for Hawkes processes and a straight application of the *Kolmogorov-Smirnov test* for the CBI processes, enable us to evaluate which of the models best describes the forward price dynamics evolution.

This is a joint work with G. Callegaro (*University of Padova*) and C. Sgarra (*Politecnico of Milano*), see Callegaro et al. (2019).

Overview of Chapter 3

During the last two decades several models have been introduced with the aim of developing and generalizing the well-known Black-Scholes framework for equity derivatives pricing. In particular, two main strands of research have been widely developed and used: Local Volatility (LV) models and Stochastic Volatility (SV) models.
Local Volatility models were introduced for the first time by Dupire et al. (1994) and Derman and Kani (1994) and they assume that the diffusion coefficient of the underlying asset is no longer a constant value but instead a deterministic function of time and of the underlying asset itself, namely

$$dS_t = (r-d)S_t dt + \eta_S\left(S_t, t\right) S_t dW_t,$$

where r and d are the risk-free interest rate and the dividend rate, respectively, and the function η_S is the so-called *local volatility function*.
In the Stochastic Volatility models instead, like in the highly celebrated Heston (1993), the volatility itself is considered to be a stochastic process v_t with its own dynamics. Thus, this is a two-factor model, driven by two correlated Wiener processes W_t and Z_t with correlation ρ, namely

$$\begin{aligned} dS_t &= (r-d)S_t dt + f\left(v_t\right) S_t dW_t \\ dv_t &= a\left(v_t, t\right) dt + c\left(v_t, t\right) dZ_t \\ dW_t dZ_t &= \rho dt, \end{aligned}$$

where r and d are the risk-free interest rate and the dividend rate, respectively, and f, a, c are functions of v_t.

Both of these models present many advantages, but also some drawbacks. The main advantage of local volatility models is their capability of a theoretically perfect fit of the market quoted plain vanilla options. If a good calibration of a local volatility model is performed, the model can reproduce the market prices. Unfortunately, one of its main drawbacks is that its implied volatility dynamics is inconsistent with the observed one. On the other hand, a stochastic volatility model is able to reproduce a consistent dynamics, but not to fit exactly the market prices.
One of the possible answers to overcome such limitations is to look for a model with a stochastic volatility dynamics which can well replicate the market prices at the same time. The key ingredient to reach this is to merge the LV and the SV models characteristics in a suitable and consistent manner. More precisely, the main idea is to model the diffusion coefficient of the underlying asset process S_t as the product between a stochastic component $f(v_t)$ and a deterministic function $\ell(S_t, t)$. Thus, this generalized stochastic-local volatility (SLV) model, that dates back to Lipton (2002), is described by the following dynamics

$$\begin{aligned} dS_t &= (r-d)S_t dt + \ell\left(S_t, t\right) f\left(v_t\right) S_t dW_t \\ dv_t &= a\left(v_t, t\right) dt + c\left(v_t, t\right) dZ_t \\ dW_t dZ_t &= \rho dt, \end{aligned}$$

where r and d are the risk-free interest rate and the dividend rate, respectively, f,a,c are functions of v_t and ρ is the correlation between the Wiener processes W_t and Z_t. The function ℓ_S is the

so-called *leverage function*.
The pivotal element of SLV models is the function ℓ_S that can be computed via an application of the Gyöngy lemma, as we shall see in Chapter 3.

Chapter 3 is based on stochastic-local volatility models applied to the market of volatility products. More precisely, it deals with the problem of studying the volatility smile of Exchange-traded products (ETPs) based on futures strategies. Exchange traded products are a variety of financial instruments that are generally created in order to track another financial instrument, whether a share-price index, a currency, a commodity or an interest rate. They are traded throughout the day on national exchanges.
The goal of Chapter 3 is to provide a general framework to study the volatility smile of a generic ETP based on futures strategies. The main problem to solve is the fact that the price process of the ETP strategy is a non Markovian function of the futures contract prices. We first start by modelling the marginal distribution of futures prices and of the ETP price and then we proceed to describe the full joint model. Options on futures can be calculated thanks to a mean-reverting local-volatility model, while options on the ETP spot price can be directly calibrated by a standard local-volatility model, that is without a mean-reversion term. The next step is to consider a generic SLV model for future dynamics under the risk-neutral measure, namely

$$dF_t(T_i) = \nu_t(T_i) \cdot dW_t,$$

where $F_t(T_i)$ is the value of the futures at time t with maturity T_i, $\nu_t(T_i)$ are vector processes, possibly depending on the futures price itself, and W_t is a vector of standard Brownian motions under the risk-neutral measure.
The key idea is to notice that since we only need to know the marginal densities of the futures and ETP prices, one can greatly simplify the problem by computing the Markov projections of the underlying processes. The process ν needs to be chosen in a way to allow to reprice correctly the options on futures and on the ETP, as described in Subsection 3.3.2. As we shall see, one of the main ingredients to reach the target is to include a (local) correlation function that links the futures dynamics with the one of the ETP through their local volatility functions, that are computed according to Nastasi et al. (2020). Even if the methodology can be applied in principle to general ETPs, we focus our attention to the specific case of the VXX Exchange Traded Notes (ETN) and VIX futures. The VXX is defined as a strategy on the first and second nearby VIX futures, namely it replicates the behavior of the one month futures on the VIX, see for example Gehricke and Zhang (2020) for a general overview.
As a first attempt, one may try to complete the task by using only standard local volatility models for VIX and VXX. After some numerical simulations, however, we have realized that this simpler choice was not enough to reproduce both quoted smiles. Therefore, we required the model to include a stochastic volatility component, chosen as a *Cox-Ingersol-Ross* process, leading to a full stochastic-local volatility model, as described in Subsection 3.4.3. It turns out that this more complex and elaborate model was the right setting to work with. Indeed, thanks to this brand-new model, we were able to jointly reproduce both the VIX futures and the VXX quoted smile. The quality of the fit for options quotes on VIX and VXX is fully satisfactory.

In order to calibrate the local volatility and the leverage functions, one needs a procedure to compute conditional expectations. We adapt to our case a technique developed in Guyon and Henry-Labordère (2012), where the core idea is to write an approximation to evaluate the expectation of a process X_t given a second process Y_t, using smooth integration kernels. Therefore, we can evaluate

$$\mathbb{E}\left[X_t \mid Y_t = K\right] \approx \frac{\mathbb{E}\left[X_t \delta^{\varepsilon}\left(Y_t - K\right)\right]}{\mathbb{E}\left[\delta^{\varepsilon}\left(Y_t - K\right)\right]},$$

where δ^{ε} is a suitably defined mollifier of the Dirac delta depending on a smoothing coefficient ε, that in our case study is a Gaussian kernel.

This is a joint work with M. Grasselli (*University of Padova*) and A. Pallavicini (*Imperial College London and Banca IMI Milan*), see Grasselli et al. (2020).

Overview of Chapter 4

A symmetry of a differential equation is a transformation that maps solutions to solutions. Continuous symmetries which have group properties are called Lie symmetries, since the means for computing them were developed in the late nineteenth century by Sophus Lie. The book of Lie, see Lie (1912), contains a collection of its original papers. Modern accounts of the theory can be found in Olver (1993) and in Bluman and Kumei (1989). The range of applications of Lie symmetry groups is extensive. Symmetries have been applied in fields ranging from epidemiology to finance and the applications continue to grow, particularly in stochastic analysis, where it is particularly useful for computing transition probability densities and functionals.

Most applications of symmetry methods rely on the construction of group invariant solutions, reduction of PDEs to canonical forms and simple mappings of trivial solutions to non trivial solutions. For example, if u is a solution of the one dimensional heat equation, $u_t = u_{xx}$, then so is $u(\varepsilon x, \varepsilon^2 t)$, when $\varepsilon > 0$, as one can easily check. This is a scaling symmetry and scaling symmetries are extremely useful. We refer the reader to Craddock (2014) for a detailed exposition of scaling symmetries in the solution of Cauchy problems on the line. A less obvious but well known result states that if we start with the one-dimensional heat equation then, for ε sufficiently small, we have that

$$\frac{1}{\sqrt{1+4\varepsilon t}} \exp\left\{\frac{-\varepsilon x^2}{1+4\varepsilon t}\right\} u\left(\frac{x}{1+4\varepsilon t}, \frac{t}{1+4\varepsilon t}\right)$$

is a solution, whenever $u(x,t)$ is. This means that even a trivial constant solution, $u(x,t) = c$, $c \in \mathbb{R}$, can be mapped into a non trivial one.

Our work outlines the use of a new and promising methodology which unites Lie symmetry analysis of linear partial differential equations (PDEs) with harmonic analysis. In this framework, a solution obtained by application of a symmetry from a known solution is to be regarded as the kernel of an integral operator, which maps suitable initial data to new solutions. This leads to the computation of fundamental solutions by inversion of integral transforms obtained by symmetries, rather than through the construction of an invariant solution to a boundary value problem. The rigorous justification of these new methods relies on the representation theory of the underlying symmetry group and we refer the reader to the works of Craddock and coauthors in the references, e.g. Craddock and Platen (2004), Craddock and Lennox (2007), Craddock and Lennox (2009), Craddock (2019a) Craddock (2019b) Craddock and Grasselli (2019).

In Chapter 4, we investigate a family of boundary conditions for PDEs associated to some barrier contracts arising in mathematical finance. We exploit the Lie symmetries of the corresponding PDEs and we characterize the density functions of the corresponding absorbed and reflected processes.

We start from the classic Black-Scholes PDE

$$u_t = \frac{1}{2}\sigma^2 S^2 u_{SS} + rSu_S - ru, \quad S \in \mathcal{D} \subseteq \mathbb{R},\ t > 0, \tag{1.1}$$

where $\mathcal{D}$ is an open interval of the real line and t is the time to maturity. This is the Kolmogorov

backward equation associated to a diffusion S satisfying

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

killed at rate r, where W is standard Brownian motion defined in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathcal{P})$. We consider, as boundary conditions, a terminal boundary condition associated to the payoff at maturity T along with the so-called *Robin boundary condition*

$$\alpha u(b,t) + \beta u_S(b,t) = g(t), \quad b > 0,$$

which is a condition that involves the derivative of the value function. This value is related to a condition involving the Delta of the option when the underlying hits the boundary.
After suitable change of variables, we reduce to study the following problem

$$\begin{aligned} w_t &= \frac{1}{2}\sigma^2 S^2 w_{SS} + rSw_S, \quad S > b > 0 \\ w(S,0) &= f(S), \\ \alpha w(b,t) + \beta w_S(b,t) &= g(t), \end{aligned} \tag{1.2}$$

where the functions $f(\cdot)$ and $g(\cdot)$ are supposed to have sufficient regularity. All the details and technicalities are presented in Chapter 4.

First we are going to solve the problem in the case of an homogeneous Robin boundary condition, namely the case $g(t) = 0$ and then we move to the more general case. We decided to treat both cases separately because we are going to exploit different techniques that can have separate interest in itself. Indeed, in the homogeneous case, we adopt particular techniques coming from the theory of ODEs together with inversion of general integral transforms. We also provide a representation of the fundamental solution of problem (1.2) in the homogeneous case.

In the non-homogeneous case, we follow a bottom-up approach and we present two methods to compute a solution. Both methods rely on the assumption that the solution can be expressed in terms of a sum of two functions, where the first one solves a slightly modified version of the homogeneous Robin problem.
Finally, we provide a general representation result for the solution of problem (1.2) as well.

This is a joint work with M. Grasselli (*University of Padova*) and M. Craddock (*School of Mathematical and Physical Sciences, University of Technology Sydney*).

Chapter 2

A Self-Exciting Modelling Framework for Forward Prices in Power Markets

2.1 Introduction

Energy markets, and in particular, electricity markets, exhibit very peculiar features that are not common to classical financial markets. The historical series of both futures and spot prices include seasonality, mean-reversion, spikes and small fluctuations. One can alternatively describe the power price dynamics by modelling the spot or the forward price. In the former case, the spot price can be obtained as a limit of the forward price when the maturity is close to the current time, in the latter case it is possible to derive the forward price from the spot by computing the conditional expectation with respect to a suitable risk-neutral measure of the spot price at the maturity.

After the pioneering paper by Schwartz (1997), where an Ornstein-Uhlenbeck dynamics is assumed to describe the spot price behaviour, several different approaches have been investigated in order to describe the power price evolution. A comprehensive literature review until 2008 is offered in the book by Benth et al. (2008). A similar effort has been devoted to identify reliable models for the forward price dynamics, and a huge amount of literature is available focusing on Heath-Jarrow-Morton type models as in Benth et al. (2019) and in Filimonov et al. (2014), in the attempt to provide a description of the whole forward curves dynamics, in analogy with forward interest rates in fixed-income markets as in Heath et al. (1992). Some of the classical models proposed include jumps and/or stochastic volatility. Benth and Paraschiv (2018) propose a random field approach based on Gaussian random fields by adopting the Musiela parametrization in order to describe the forward curve dynamics. Empirical evidence suggests that in many assets jumps appear in cluster, thus requiring the introduction of jump processes exhibiting a clustering or self-exciting behaviour. Kiesel and Paraschiv (2017) recently presented a systematic empirical investigation of electricity intraday prices and suggested an approach based on Hawkes processes in order to describe the power price dynamics with jump clustering features. Self-exciting features in electricity prices attracted already some attention by several authors: Herrera and González (2014) proposed a self-excited model for electricity spot prices, while Christensen et al. (2009), Clements et al.

(2013) pointed out that time between spikes has a significant impact on the likelihood of future occurrences, thus providing a strong support to models including self-exciting properties.

The large class of models available in the literature describing the power price dynamics is then widening in order to include models exhibiting self-exciting features.

We also mention the paper by Jiao et al. (2019), where a model based on continuous branching processes with immigration for power spot prices was proposed, and the forward prices computed with respect to a suitable structure preserving equivalent martingale measure.

Eyjolfsson and Tjøstheim (2018) describe a class of Hawkes processes and present an empirical investigation based on data from UK power market supporting Hawkes-type models for spot prices.

The purpose of the present paper is to investigate if self-exciting features can arise in the power forward prices evolution as well, and in order to perform this investigation we shall focus on two different model classes: the Continuous Branching Processes with Immigration (CBI henceforth) and the Hawkes processes. While CBI processes are always affine, Hawkes processes in general are not, but when the kernel describing the intensity dynamics is of exponential type they are, and this feature makes the Hawkes processes with exponential kernel appealing from the modelling point of view. By considering the two model classes mentioned before, i.e. CBI and Hawkes processes, we then want to provide the description of the full term structure of power forward prices, following a Heath-Jarrow-Morton approach.

Power is a flow commodity, this meaning that instantaneous forward contracts are not directly traded on the market, but futures (sometimes called flow forward) are. So, in order to perform any kind of inference on the model proposed, it is necessary to extract the relevant information on the forward dynamics included in the futures prices. This can be done by applying suitable optimization procedures proposed in the literature, eventually modified in order to provide the best performances in the case under examination. These procedures are far from trivial from the computational point of view and require a careful implementation of the optimization step. We deliberately chose to work on daily data, in order to show how self-exciting effects can arise not only on a small time scale, but also at a coarser level.

The chapter is organized as follows. In Section 2.2 we introduce the processes on which our models are based and in Section 2.3 we present and discuss the models proposed for the forward power price dynamics. In Section 2.4 we discuss the dynamics of Futures contracts when the forward dynamics is assumed to be given by the models introduced. From Section 2.5 to Section 2.8, we provide the theoretical background and numerical results relative to the calibration/parameters' estimation for the model proposed. In Section 2.9 we provide some concluding remarks and discuss future extensions of the present work.

2.2 The Modelling Framework

2.2.1 Continuous Branching Processes with Immigration

We now introduce our modelling framework for the electricity price, which is based on stochastic differential equations driven by Lévy random fields. We consider a Lévy random field, which is a combination of a Gaussian random measure W and a compensated Poisson random measure N

independent of W. For a background on such general stochastic equations with jumps, we refer the readers e.g. to Dawson et al. (2012), Li and Ma (2015) and Walsh (1986).

Let us now briefly introduce all the relevant ingredients of our work and recall some preliminary results. We fix a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.

A *white noise* W on $\mathbb{R}_+^2$ is a Gaussian random measure such that, for any Borel set $A \in \mathcal{B}(\mathbb{R}_+^2)$ with finite Lebesgue measure $|A|$, $W(A)$ is a normal random variable of mean zero and variance $|A|$ and if $A_1, \cdots, A_n$ are disjoint Borel sets in $\mathcal{B}(\mathbb{R}_+^2)$, then $W(A_1), \cdots, W(A_n)$ are mutually independent.

We denote by N the Poisson random measure on $\mathbb{R}_+^3$ with *intensity* λ which is a Borel measure on $\mathbb{R}_+^3$ defined as the product of the Lebesgue measure on $\mathbb{R}_+ \times \mathbb{R}_+$ with a Borel measure μ on $\mathbb{R}_+$ such that $\int_0^\infty (z \wedge z^2)\mu(dz) < +\infty$. Note that μ is a Lévy measure since $\int_0^\infty (1 \wedge z^2)\mu(dz) < +\infty$.

Recall that for each Borel set $B \in \mathcal{B}(\mathbb{R}_+^3)$ with $\lambda(B) < +\infty$, the random variable $N(B)$ has the Poisson distribution with parameter $\lambda(B)$. Moreover, if $B_i, i = 1, \ldots, n$ are disjoint Borel sets in $\mathcal{B}(\mathbb{R}_+^3)$, then $N(B_1), \cdots, N(B_n)$ are mutually independent. We let $\widetilde{N} = N - \lambda$ be the compensated Poisson random measure on $\mathbb{R}_+^3$ associated to N.

We introduce the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geqslant 0}$ as the natural filtration generated by the Lévy random field (see Dawson et al. (2012)) and satisfying the usual conditions, namely, for any Borel subset $A \in \mathcal{B}(\mathbb{R}_+)$ and $B \in \mathcal{B}(\mathbb{R}_+^2)$ of finite Lebesgue measure, the processes $(W([0,t] \times A), t \geq 0)$ and $(\widetilde{N}([0,t] \times B), t \geq 0)$ are $\mathbb{F}$-martingales.

We consider the following stochastic differential equation written in the integral form. Let $a, b, \sigma, \gamma \in \mathbb{R}_+$ be constant parameters. Consider the equation:

$$Y(t) = Y(0) + \int_0^t a\,(b - Y(s))\,ds + \sigma \int_0^t \int_0^{Y(s)} W(ds, du) + \gamma \int_0^t \int_0^{Y(s-)} \int_{\mathbb{R}_+} z\widetilde{N}(ds, du, dz), \tag{2.1}$$

where $W(ds, du)$ is a white noise on $\mathbb{R}_+^2$ with unit covariance, $\widetilde{N}(ds, du, dz)$ is an independent compensated Poisson random measure on $\mathbb{R}_+^3$ with intensity $ds\ du\ \mu(dz)$ with $\mu(dz)$ being a Lévy measure on $\mathbb{R}_+$ and satisfying $\int_0^\infty (z \wedge z^2)\mu(dz) < \infty$. We recall that by $Y(t-) = \lim_{s\uparrow t} Y(s)$ we denote the jump at time t of the process $Y(t)$, the limit being from the left of the process.

The integrals appearing in Equation (2.1) (and in the following) are both in the sense of Walsh (1986). It follows from Dawson et al. (2012, Theorem 3.1) or Li and Ma (2015, Theorem 2.1) that Equation (2.1) has a unique strong solution.

Our model actually belongs to the family of CBI processes. Continuous Branching Processes with Immigration (CBI) are a class of stochastic processes commonly used in modelling population dynamics, see for example Pardoux (2016) for a general overview of these models. The self-exciting features, arising from the integrals in Equation (2.1) extended on the domain $[0, Y(s))$ with respect to the integration variable u, describe the growth of the population due to the reproduction of the previous generations. In the present modelling framework they just describe jumps generated by previous jumps. We briefly recall the definition by Kawazu and Watanabe (1971, Definition 1.1).

Definition 2.2.1. *A Markov process Y with state space $\mathbb{R}_+$ is called a CBI process characterized by branching mechanism $\Psi(\cdot)$ and immigration rate $\Phi(\cdot)$, if its characteristic representation is*

given, for $p \geq 0$, by:

$$\mathbb{E}_y\left[e^{-pY(t)}\right] = \exp\left(-yv(t,p) - \int_0^t \Phi\big(v(s,p)\big)ds\right),$$

where $\mathbb{E}_y$ denotes the conditional expectation with respect to the initial value $Y(0) = y$. The function $v : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ satisfies the following differential equation:

$$\frac{\partial v(t,p)}{\partial t} = -\Psi(v(t,p)), \quad v(0,p) = p$$

and Ψ and Φ are functions of the variable $q \geq 0$ given by

$$\Psi(q) = aq + \frac{1}{2}\sigma^2 q^2 + \gamma\int_0^\infty (e^{-qu} - 1 + qu)\pi(du),$$
$$\Phi(q) = abq + \int_0^\infty (1 - e^{-qu})\nu(du),$$

with $\sigma, \gamma \geq 0$, $\beta \in \mathbb{R}$ and π, ν being two Lévy measures such that

$$\int_0^\infty (u \wedge u^2)\pi(du) < \infty, \quad \int_0^\infty (1 \wedge u)\nu(du) < \infty.$$

It is proved in Dawson et al. (2012, Theorem 3.1) that the process in Equation (2.1) is a CBI process with the branching mechanism Ψ given by

$$\Psi(q) = aq + \frac{1}{2}\sigma^2 q^2 + \int_0^\infty (e^{-q\gamma z} - 1 + q\gamma z)\mu(dz),$$

while the immigration rate is given by

$$\Phi(q) = abq.$$

The link between CBI processes and the affine term structure models has been established by Filipović (2001). If the process Y takes values in $\mathbb{R}_+$, he proves equivalence between the two classes. We recall that the joint Laplace transform of a CBI process Y and its integrated process, which is given in Filipović (2001, Theorem 5.3), is defined as follows: for non-negative real numbers ξ and θ, we have:

$$\mathbb{E}_y\left[e^{-\xi Y(t) - \theta\int_0^t Y(s)ds}\right] = \exp\Big\{-yv(t,\xi,\theta) - \int_0^t \Phi\big(v(s,\xi,\theta)\big)ds\Big\},$$

where $v(t,\xi,\theta)$ is the unique solution of

$$\frac{\partial v(t,\xi,\theta)}{\partial t} = -\Psi(v(t,\xi,\theta)) + \theta, \quad v(0,\xi,\theta) = \xi.$$

2.2.2 Hawkes Processes

A Hawkes process is a special counting process with a random intensity function. We introduce now the Hawkes processes with exponential kernel. They can be written as follows :

$$Y(t) = Y(0) + \sum_{i=1}^{N_t} Z_i = Y(0) + \int_0^t \int_0^\infty z J(ds, dz), \tag{2.2}$$

where the last term is an Ito integral, N_t is the number of jumps in the interval between 0 and t and $J(dz, ds)$ is a Poisson random measure with intensity $\lambda(t)$, satisfying the SDE:

$$\begin{aligned} \lambda(t) &= \lambda(0) - \beta \int_0^t \lambda(s) ds + \alpha \int_0^t \int_0^\infty z J(ds, dz) \\ &= \exp(-\beta t)\lambda(0) + \alpha \sum_{i=1}^{N_t} \exp[-\beta(t - t_i)] Z_i. \end{aligned} \tag{2.3}$$

Here $\beta > 0$ is the rate of exponential decay of the influence of previous jumps on the intensity level and α the amplitude of the memory kernel, t_i are the jumps times and Z_i the jump sizes, which we shall assume distributed according to an exponential density with parameter δ, so that only positive jumps appear in both Equations (2.2) and (2.3), and we can write $\tilde{J}(ds, dz) = J(ds, dz) - \lambda(s)\mu(dz)ds$ and $\mu(dz) = \delta \exp(-\delta z)dz$, where $\tilde{J}(ds, dz)$ denotes the compensated version of the Poisson measure $J(ds, dz)$. We assume the following condition holds: $\beta - \alpha/\delta > 0$, granting the non-explosiveness of the Hawkes process (see e.g. Bernis et al. (2018)).

Hawkes processes with exponential kernel are the only class of Hawkes processes exhibiting both the Markov property and an affine structure (see e.g. Errais et al. (2010)). They have been extensively used in order to describe the dynamics of several asset classes, including equities as in Hainaut and Moraux (2019), commodities as in Eyjolfsson and Tjøstheim (2018), exchange rates as in Rambaldi et al. (2015) and credit risk as in Errais et al. (2010).

2.3 Forward Prices Modelling

In this section we are going to introduce the two alternative models for the forward prices, that we are going to test against electricity market data. In both cases the price at time t of a forward contract with maturity $T \geq t$ is additive and it can be defined as follows

$$f(t, T) = \Lambda(t) - \Lambda(0) + \sum_{i=1}^{n} X_i(t, T), \tag{2.4}$$

where $\Lambda(t)$ is a deterministic seasonality function that will be made precise later on, n is the number of factors used and each of the terms X_i is an underlying factor, whose dynamics will be specified in the following Subsections 2.3.1 and 2.3.2.

2.3.1 The Forward Model based on CBI

Let n be the number of factor in model (2.4) and let $a_i, \sigma_i, \gamma_i \in \mathbb{R}_+,\ i = 1, \dots, n$ be constant parameters. Our first model assumes the following dynamics for the factors $X_i, i = 1, \dots, n$:

$$\begin{aligned} X_i(t,T) = X_i(0,T) - \int_0^t a_i X_i(s,T)ds + \sigma_i \int_0^t \int_0^{X_i(s,T)} W_i(ds,du) \\ + \gamma_i \int_0^t \int_0^{X_i(s-,T)} \int_{\mathbb{R}_+} z\widetilde{N}_i(ds,du,dz). \end{aligned} \tag{2.5}$$

Namely, the X_i's evolve in time with respect to the historical measure $\mathbb{P}$ according to Equation (2.1) with immigration rate $b_i = 0$. By recalling that the intensity of the Poisson random measure $\widetilde{N}_i(ds,du,dz)$ is given by $dsdu\mu_i(dz)$, we assume $\mu_i(dz) = \delta_i \exp(-\delta_i z)dz$ with $\delta_i > 0$, for $i = 1, \cdots, n$, $z > 0$.

Assuming the dynamics for the factors X_i as in (2.5), it is possible to re-write Equation (2.4) as follows:

$$\begin{aligned} f(t,T) = \Lambda(t) - \Lambda(0) + \sum_{i=1}^n X_i(0,T) - \sum_{i=1}^n \int_0^t a_i X_i(s,T)ds + \sum_{i=1}^n \sigma_i \int_0^t \int_0^{X_i(s,T)} W_i(ds,du) \\ + \sum_{i=1}^n \gamma_i \int_0^t \int_0^{X_i(s-,T)} \int_{\mathbb{R}_+} z\widetilde{N}_i(ds,du,dz), \end{aligned}$$

or, equivalently, as

$$\begin{aligned} f(t,T) = \Lambda(t) - \Lambda(0) + f(0,T) - \sum_{i=1}^n \int_0^t a_i X_i(s,T)ds + \sum_{i=1}^n \sigma_i \int_0^t \int_0^{X_i(s,T)} W_i(ds,du) \\ + \sum_{i=1}^n \gamma_i \int_0^t \int_0^{X_i(s-,T)} \int_{\mathbb{R}_+} z\widetilde{N}_i(ds,du,dz), \end{aligned} \tag{2.6}$$

where $f(0,T) = \sum_{i=1}^n X_i(0,T)$, as one can easily check from (2.4).

The relation between the dynamics of the forward price with respect to the historical measure $\mathbb{P}$ and the risk-neutral dynamics, written with respect to $\mathbb{Q}$, can be easily obtained by applying the following result, proved in the paper by (Jiao et al., 2019, Proposition 4.1).

Proposition 2.3.1. *Let $X_1, X_2, \cdots, X_n$ be independent CBI processes where for each $i \in \{1, \cdots, n\}$, X_i is a CBI process under the probability measure $\mathbb{P}$, with dynamics given by 2.5. Assume that the filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is generated by the random fields $W_1, W_2, \cdots, W_n$ and by the compensated Poisson random measures $\widetilde{N}_1, \widetilde{N}_2, \cdots, \widetilde{N}_n$. For each i, fix $\eta_i \in \mathbb{R}$ and $\xi_i \in \mathbb{R}_+$ and define*

$$U_t := \sum_{i=1}^n \eta_i \int_0^t \int_0^{X_i(s)} W_i(ds,du) + \sum_{i=1}^n \int_0^t \int_0^{X_i(s-)} \int_0^\infty (e^{-\xi_i z} - 1)\widetilde{N}_i(ds,du,dz).$$

Then the Doléans-Dade exponential $\mathcal{E}(U)$ is a martingale under $\mathbb{P}$ and the probability measure $\mathbb{Q}$ defined by

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \mathcal{E}(U)_t,$$

is equivalent to $\mathbb{P}$. Moreover, under $\mathbb{Q}$, X_i is a CBI process with parameters $(a_i^{\mathbb{Q}}, b_i^{\mathbb{Q}}, \sigma_i^{\mathbb{Q}}, \gamma_i^{\mathbb{Q}}, \mu_i^{\mathbb{Q}})$, where:

$$\begin{gathered} a_i^{\mathbb{Q}} = a_i^{\mathbb{P}} - \sigma_i^{\mathbb{P}}\eta_i - \int_0^\infty z(e^{-\theta_i z} - 1)\mu_i^{\mathbb{P}}(dz), \\ b_i^{\mathbb{Q}} = a_i^{\mathbb{P}} b_i^{\mathbb{P}}/a_i^{\mathbb{Q}}, \quad \sigma_i^{\mathbb{Q}} = \sigma_i^{\mathbb{P}}, \quad \gamma_i^{\mathbb{Q}} = \gamma_i^{\mathbb{P}} \\ \mu_i^{\mathbb{Q}}(dz) = e^{-\theta_i z}\mu_i^{\mathbb{P}}(dz), \quad \delta_i^{\mathbb{Q}} = \delta_i^{\mathbb{P}} \end{gathered}$$

Remark 2.3.2. *In this context, the parameters η_i, ξ_i can be interpreted as the Market Price of Risk associated with the diffusion/jump part of $X_i, i = 1, \ldots, n$, respectively.*

Remark 2.3.3. *In order to avoid arbitrage opportunities we shall assume that the de-seasonalized dynamics of every factor X_i is a local martingale under $\mathbb{Q}$ and this will automatically imply that $a_i = 0$ under $\mathbb{Q}$. Since the first integral is defined with respect to the Gaussian white noise $W_i(ds, du)$ and the second integral is defined with respect to the compensated Poisson random measure $\widetilde{N}_i(ds, du, dz)$, each process $X_i(t, T)$ is in fact a local martingale with respect to $\mathbb{Q}$.*

Remark 2.3.4. *From Proposition 2.3.1, specifying the relations between the model parameters under the risk-neutral measure $\mathbb{Q}$ and the historical measure $\mathbb{P}$, it is clear that in the CBI modelling framework, for each factor X_i, a mean reversion speed coefficient a_i can be non-null under $\mathbb{P}$ and zero under $\mathbb{Q}$. As far as the immigration term b_i is concerned, if it vanishes under $\mathbb{Q}$, it will be zero under any equivalent probability measure.*

Assumption 2.3.5. *In the estimation procedure applied to the real market data we shall assume that only one process of the type introduced in (2.5) will drive the forward curve dynamics.*

2.3.2 The Forward Model Based on Hawkes Processes

Let n be the number of factor in model (2.4) and let $c_i, \sigma_i \in \mathbb{R}_+$, $i = 1, \ldots, n$ be constant parameters. As alternative to the model proposed in the previous subsection, we consider, under $\mathbb{P}$, the dynamics (2.5), where now each X_i, $i = 1, \ldots, n$, satisfies a SDE of the following form:

$$X_i(t, T) = X_i(0, T) - \int_0^t c_i X_i(s, T)ds + \sigma_i \int_0^t \sqrt{X_i(s, T)}dW_i(s) + \int_0^t \int_0^\infty z\tilde{J}_i(dz, ds), \quad (2.7)$$

where $\tilde{J}_i(dz, ds)$ are compensated marked point process with intensity $\lambda_i(t)$, satisfying the SDE:

$$\lambda_i(t) = \lambda_i(0) - \beta_i \int_0^t \lambda_i(s)ds + \alpha_i \int_0^t \int_0^\infty z J_i(ds, dz), \quad (2.8)$$

where we recall that $\beta_i > 0$ is the rate of exponential decay of the influence of previous jumps on the intensity level and α_i the amplitude of the memory kernel of each factor X_i. We assume the

jump size distributed according to an exponential density with parameter δ_i for each (λ_i, X_i), so we can write:

$$\tilde{J}_i(ds, dz) = J_i(ds, dz) - \lambda_i(s)\mu(dz)ds = J_i(ds, dz) - \lambda_i(s)\delta_i \exp(-\delta_i z)(dz)ds. \tag{2.9}$$

Remark 2.3.6. *The choice of a square-root process for the diffusion part of the forward curves dynamics is motivated by the positivity requirement as well as the choice of the exponential distribution for the jumps size.*

Therefore, Equation (2.4) in this case takes the following form:

$$\begin{aligned} f(t,T) = \Lambda(t) - \Lambda(0) + \sum_{i=1}^{n} X_i(0,T) - \sum_{i=1}^{n} \int_0^t c_i X_i(s,T)ds + \sum_{i=1}^{n} \int_0^t \sigma_i \sqrt{X_i(s,T)} dW_i(s) \\ + \sum_{i=1}^{n} \int_0^t \int_0^\infty z\tilde{J}_i(dz, ds), \end{aligned}$$

or, equivalently, as

$$\begin{aligned} f(t,T) = \Lambda(t) - \Lambda(0) + f(0,T) - \sum_{i=1}^{n} \int_0^t c_i X_i(s,T)ds + \sum_{i=1}^{n} \int_0^t \sigma_i \sqrt{X_i(s,T)} dW_i(s) \\ + \sum_{i=1}^{n} \int_0^t \int_0^\infty z\tilde{J}_i(dz, ds), \end{aligned} \tag{2.10}$$

where, again, $f(0,T) = \sum_{i=1}^{n} X_i(0,T)$.

In order to make the presentation of the two model classes more homogeneous, we can introduce the Dawson-Li representation for the Hawkes-type dynamics as well and write the SDE governing the dynamics of forward prices under the historical measure $\mathbb{P}$ as follows:

$$\begin{aligned} X_i(t,T) = X_i(0,T) - \int_0^t c_i X_i(s,T)ds + \int_0^t \int_0^{X_i(s,T)} \sigma_i W_i(du, ds) \\ + \int_0^t \int_0^{X_i(s-,T)} \int_{\mathbb{R}_+} z\tilde{N}_i(dz, du, ds), \end{aligned} \tag{2.11}$$

where the definition of the integrals and the notations are the same as in Subsection 2.2.1 and the $\lambda_i(t)$ evolve according to Eq. (2.8).

It is immediate to remark that the dynamics described by the two model classes look almost identical when written in the Dawson-Li representation, the main difference being the specification of the equation governing the evolution of the intensity processes. This is one of the reasons behind the choice of these two alternative models to describe the forward prices' evolution.

The dynamics just described is given with respect to the historical probability measure $\mathbb{P}$. In order to obtain a description with respect to the risk-neutral measure $\mathbb{Q}$ we need to introduce a measure change, as we have already done for the CBI processes in Proposition 2.3.1. The following proposition provides a measure change preserving the Hawkes-type dynamics. A proof can be found in Bernis et al. (2020).

Proposition 2.3.7. *Let (λ_i, X_i) be described by Equations (2.3) and (2.11) under the historical probability $\mathbb{P}$. Fix $(\eta, \xi) \in \mathbb{R} \times (-\delta_i, \infty)$ and define:*

$$U_t := \sum_{i=1}^{n} \eta_i \sigma_i \int_0^t \int_0^{X_i(s)} W_i(ds, du) + \sum_{i=1}^{n} \int_0^t \int_0^{\lambda_i(s-)} \int_{\mathbb{R}_+} \left(e^{-\xi_i z} - 1\right) \widetilde{J}_i(ds, du, dz).$$

Then the Doléans-Dade exponential $\mathcal{E}(U)$ is a martingale under $\mathbb{P}$ and the probability measure $\mathbb{Q}$ defined by $\frac{d\mathbb{Q}}{d\mathbb{P}}\big|_{\mathcal{F}_t} := \mathcal{E}(U)_t$ is equivalent to $\mathbb{P}$. The dynamics with respect to $\mathbb{Q}$ takes the following form:

$$\begin{aligned} X_i(t,T) &= X_i(0,T) + \int_0^t \int_0^{X_i(s,T)} \sigma_i^{\mathbb{Q}} W_i(du, ds) + \int_0^t \int_0^{X_i(s_-,T)} z \tilde{J}_i^{\mathbb{Q}}(dz, du, ds), \\ \lambda_i(t) &= \lambda_i(0) - \int_0^t \beta_i^{\mathbb{Q}} \lambda(s) ds + \alpha_i^{\mathbb{Q}} \int_0^t \int_0^{\infty} \exp\left[-\beta^{\mathbb{Q}}(t-s)\right] J_i^{\mathbb{Q}}(dz, ds), \end{aligned}$$

where

$$\begin{gathered} c_i^{\mathbb{Q}} = c_i^{\mathbb{P}} - \sigma_i^{\mathbb{P}} \eta_i - \textstyle\int_0^{\infty} z(e^{-\theta_i z} - 1) \mu_i^{\mathbb{P}}(dz), \quad \sigma_i^{\mathbb{Q}} = \sigma_i^{\mathbb{P}} \\ \alpha_i^{\mathbb{Q}} = \alpha_i^{\mathbb{P}}, \quad \beta_i^{\mathbb{Q}} = \beta_i^{\mathbb{P}}, \quad \mu_i^{\mathbb{Q}}(dz) = e^{-\theta_i z} \mu_i^{\mathbb{P}}(dz). \end{gathered}$$

Remark 2.3.8. *In this context, the parameters η_i, ξ_i can be interpreted as the Market Price of Risk associated with the diffusion/jump part of the $X_i, i = 1, \ldots, n$, respectively.*

Remark 2.3.9. *We shall assume, as for the previous model, that the de-seasonalized dynamics of X_i is a local martingale under $\mathbb{Q}$ and this will automatically imply that the mean reversion speed c_i of any X_i must vanish under $\mathbb{Q}$. Both the diffusion and the jump terms are in fact local martingales with respect to $\mathbb{Q}$.*

Remark 2.3.10. *From Proposition 2.3.7, specifying the relations between the model parameters under the risk-neutral measure $\mathbb{Q}$ and the historical measure $\mathbb{P}$, it is clear that in the Hawkes modelling framework, for each factor X_i, a mean reversion speed coefficient c_i can be nonzero under $\mathbb{P}$ and zero under $\mathbb{Q}$. A non zero mean-reverting term can then appear in the dynamics written with respect to the historical measure $\mathbb{P}$, although this term vanishes under $\mathbb{Q}$.*

Assumption 2.3.11. *In the estimation procedure applied to the real market data we shall assume that only one process of the type introduced in (2.7) will drive the forward curve dynamics.*

2.4 The Futures Dynamics

We focus here rigorously on forward contracts delivering a quantity of energy over a finite period of time. We shall refer to them as futures, even if in the literature they are sometimes called swaps or flow forwards.

Definition 2.4.1. *The price at time $t \geq 0$ of a futures contract with delivery period $[T_1, T_2]$ with $t \leq T_1 \leq T_2$ si given by*

$$F(t, T_1, T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(t, x)\, dx, \tag{2.12}$$

where $f(t, \cdot)$ is the price at time t of the forward contract to be paid upon delivery.

Remark 2.4.2. *From Definition 2.4.1 it is clear why futures are sometimes called flow forwards: the owner of a futures with delivery period over $[T_1, T_2]$ would substantially receive a constant flow of the commodity over this period. Notice also that a futures contract delivering the commodity over a time period which collapses into a single point coincides with a forward.*

The value at time t of a Futures contract with delivery period $[T_1, T_2]$ is given, in our modelling framework, by (recall Equation (2.4)):

$$F(t, T_1, T_2) = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(t, x) dx = \Lambda(t) - \Lambda(0) + \frac{1}{T_2 - T_1} \left(\sum_{i=1}^{n} \int_{T_1}^{T_2} X_i(t, x) dx \right). \tag{2.13}$$

By introducing the dynamics of the factors X_i into the above equation, we get the following equation describing the futures' dynamics under the risk-neutral probability $\mathbb{Q}$ both in the CBI framework (recall Equation (2.5)):

$$\begin{aligned} F(t, T_1, T_2) = \Lambda(t) - \Lambda(0) &+ \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(0, x) dx \\ &+ \frac{1}{T_2 - T_1} \sum_{i=1}^{n} \sigma_i \int_{T_1}^{T_2} \int_0^t \int_0^{X_i(s,x)} W_i(ds, dy) dx \\ &+ \frac{1}{T_2 - T_1} \sum_{i=1}^{n} \gamma_i \int_{T_1}^{T_2} \int_0^t \int_0^{X_i(s-,x)} \int_{\mathbb{R}_+} z \widetilde{N}_i(ds, dy, dz) dx. \end{aligned}$$

and in the Hawkes setting (recall Equation (2.7)):

$$\begin{aligned} F(t, T_1, T_2) = \Lambda(t) - \Lambda(0) &+ \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} f(0, x) dx \\ &+ \frac{1}{T_2 - T_1} \sum_{i=1}^{n} \sigma_i \int_{T_1}^{T_2} \int_0^t \sqrt{X_i(s, x)} dW_i(s) dx \\ &+ \frac{1}{T_2 - T_1} \sum_{i=1}^{n} \int_{T_1}^{T_2} \int_0^t \int_{\mathbb{R}_+} z \tilde{J}_i(dz, ds) dx. \end{aligned}$$

Assumption 2.4.3. *From now on, in view of our numerical analysis, we will assume that one driving factor is sufficient. Namely, we will consider the case $n = 1$.*

In order to rule out arbitrage opportunities the prices of futures with different delivery periods must satisfy specific time-consistency relations. In particular, the value of a futures contract

with delivery period $[T_1, T_n]$ is linked to the values of the contracts with delivery on intervals $[T_i, T_{i+1}], i = 1, \ldots, n-1$, where $[T_i, T_{i+1}]$ represents a partition of the interval $[T_1, T_n]$, by the following relation:

$$F(t, T_1, T_n) = \frac{1}{T_n - T_1} \sum_{i=1}^{n-1} (T_{i+1} - T_i) F(t, T_i, T_{i+1}). \tag{2.14}$$

Generally, in power market, one can observe energy-related futures contracts deliver the underlying spot over a certain time period, for example a week (W), a month (M), a quarter (Q) or a whole calendar year (Cal). Therefore, one can find at the same time several futures covering the same time period (e.g. a quarter covered by the corresponding contract as well as a quarter covered by the three corresponding monthly contracts). However, these contracts are not all present in the markets at the same time. In fact, usually only the shortest-maturity months are present, while the other months are still packed in quarterly or calendar contracts. As time goes by, these long contracts are unpacked in a cascade mechanism, as shown in Figure 2.1.

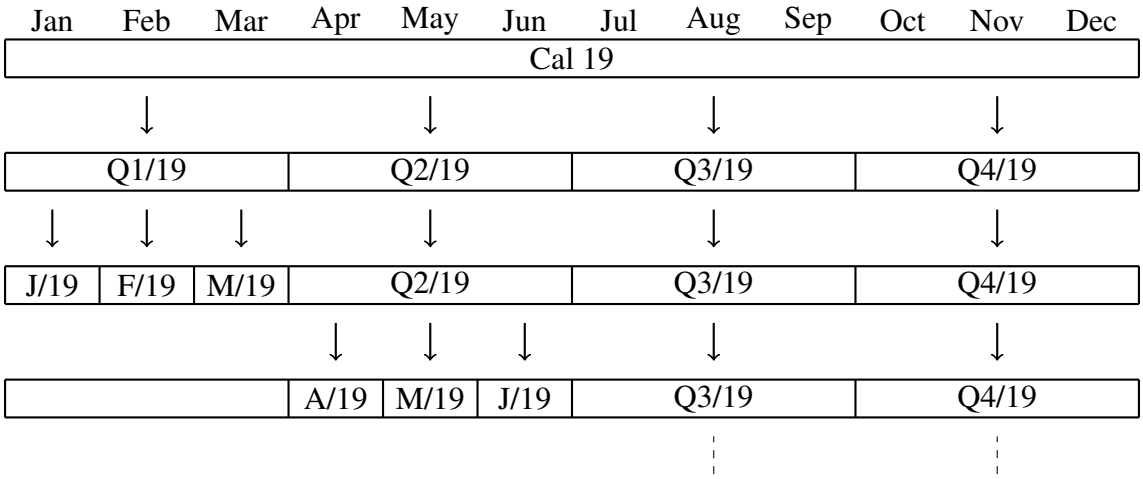


Figure 2.1: Cascade unpacking mechanism of futures contracts. For each given calendar year, as time passes by, futures are unpacked first in quarters, then in the corresponding months. It may happen that the same delivery period is covered by different contracts, e.g. one simultaneously finds quotes for the monthly contracts J/19, F/19, M/19 and for the quarterly Q1/19.

2.5 Data Analysis: from Futures Prices to Forward Curves

From a theoretical point of view the contracts are settled continuously over the delivery period, as you can see from Equation (2.12), but in practice they are settled at discrete times. Assuming settlement at N points in time $u_1 < u_2 < \ldots < u_N$, with $u_1 = T_1$, $u_N = T_2$, then the discrete

version of Equation (2.12) becomes

$$F(t, T_1, T_2) = \frac{1}{T_2 - T_1} \sum_{i=1}^{N} f(t, u_i)\, \Delta_i,$$

where $\Delta_i = u_{i+1} - u_i$.

The main goal of what follows is to provide a forward dynamics formulation starting from the futures prices that we observe in the market. What we are going to do is to build a smooth curve describing today's forward prices from quoted futures prices, according to the Heath-Jarrow-Morton framework outlined before.

This is a well studied problem in literature and there are basically two ways to proceed: either fitting a parametric function to the entire yield curve by regression, or fitting all observed yields with a spline (see for example Anderson (1996) for a survey on different methods for constructing yield curves). Here we opt for the latter approach, mainly because of its flexibility and easiness to implement. In addition, it can be used both for interpolation and smoothing, as we shall see in Subsection 2.5.1.
Throughout the paper we will also use the notation T_i^s, T_i^e to denote the first (start) and the last (end) day of the delivery period of the i-th contract, and T^s, T^e in case of no ambiguity.

Data set description

We are working on financially settled base load daily futures contract based upon the daily prices of the French power market. The term *base load* indicates that the delivery of electricity takes place for each hour of the day, in contrast to *peak load* contracts, that instead settle the delivery only for specified hours (from 8 to 20 typically). These contracts prescribe the delivery of 1 MW per day in the contract period (week, month, quarter or year) and their price is in Euros and Euro cents per MWh.

More precisely, our data set consists of French futures closing prices downloaded from Thomson Reuters, that span over a period of 17 years, from 2002 to 2019. These contracts are divided with respect to the length of the delivery period into: weekly (tickers F7B1-B5), monthly (ticker F7BM), quarterly (ticker F7BQ) and yearly (ticker F7BY) contracts. Each of them presents 4 typologies of rolling contracts, namely $c1$, $c2$, $c3$ and $c4$, where $c1$ and $c4$ are the ones with the closest and the farthest delivery period, respectively (Figure 2.3 in Section 2.5.2 shows an example to see how these rolling contracts work).

On the market we observe the quantity $F(0, T^s, T^e)$ for every contract, for different choices of T^s, T^e with $T^e - T^s$ = 7, 30, 90 and 365 days according to the weekly, monthly, quarterly and yearly contracts, respectively, where "0" is the current date, the first available being July $1, 2002$, while the last available being March $15, 2019$, for a total number of 4234 trading days. More precisely, on each day, that we label as the "0" day, we have different number of contracts with different delivery periods, depending on the market data availability of that day. Our goal is to extract the curve $f(0, u)$ for all the different choices of the "0" date, as we are going to see in Subsection 2.5.1.

Notation 2.5.1. *When possible from now on we will write $f(u)$ instead of $f(0,u)$ and $F(T_1,T_2)$ instead of $F(0,T_1,T_2)$ to shorten the notation.*

2.5.1 Extracting Smooth Forward Curves from market data

Obtaining a smooth curve of forward prices from futures prices is a well studied problem in the literature, see for example Fleten and Lemming (2003). The initial condition for using a Heath-Jarrow-Morton approach when modelling forwards is a smooth curve describing today's forward prices, which must be extracted from the futures prices observed in the market. We will follow the approach of Benth et al. (2008, Chapter 7) by imposing the following

Assumption 2.5.2. *The forward curve can be represented as the sum of two continuous functions $\Lambda(u)$ and $\varepsilon(u)$:*

$$f(u) = \Lambda(u) - \Lambda(0) + \varepsilon(u), \quad u \in [T^s, T^e], \tag{2.15}$$

where T^s is the starting day of the settlement period for the contract with the closest delivery period and T^e is the last day of the settlement period for the contract with the farthest delivery period. We interpret $\Lambda(u)$ as a seasonality function and $\varepsilon(u)$ to be an adjustment function that captures the forward curve's deviation from the seasonality.

For the specification of the seasonality function we follow the approach of Benth et al. (2008), namely we define

$$\Lambda(u) = a\cos\left((u-b)\cdot\frac{2\pi}{365}\right). \tag{2.16}$$

The parameter $a \in \mathbb{R}_+$ is obtained by finding the minimum of the prices over all the contracts, while b is the normalized distance between the end of the last day of the year from the day when the minimum occurs.

This procedure leads to:

$$a = 13.600, \quad b = 1358.038. \tag{2.17}$$

There are several other methods for extracting the seasonality function from the data (see for example Paraschiv (2013), and Kiesel et al. (2019) for an application to hourly data), but since this topics is not the main focus of our study, we prefer to stick to the well known method proposed by Benth et al. (2007) and systematically described in (Benth et al., 2008, Chapter 7).

We shall see now how the adjustment function ε is obtained.

The function ε: a maximum smooth forward curve

Sticking to the approach of Benth et al. (2008), we perform a maximum smoothness criterion applied to the adjustment function ε.

Remark 2.5.3. *One may ask why the maximum smoothness criterion is applied only to the adjustment function ε and not to the entire forward function f. This ensures the presence of a seasonality pattern that, otherwise, would have possibly been smoothed out.*

The properties we require for the adjustment function are that it is twice continuously differentiable and horizontal at time T^e, i.e.

$$\varepsilon'(T^e) = 0. \tag{2.18}$$

This flatness condition reflects that the long end of the curve may be several years ahead, and obviously the market's view on risk become less and less sensitive as time goes by.

Let us denote by $C_0^2([T^s, T^e])$ the set of real-valued functions on the interval $[T^s, T^e]$ which are twice continuously differentiable with zero derivative in T^e. We consider $\mathcal{C}$ as the set of polynomial spline functions of order four which belong to $C_0^2([T^s, T^e])$.

Definition 2.5.4. *We define the smoothest possible forward curve on an interval $[T^s, T^e]$ as the function which minimizes, over $\mathcal{C}$, the integral*

$$\int_{T^s}^{T^e} [\varepsilon''(u)]^2 du$$

and such that the closing prices matching condition holds (this is made precise in Equation (2.24)).

We interpret the smoothest forward curve (2.15) to be the one for which ε solves the minimization problem above, with Λ chosen as in (2.16) and a, b as in (2.17).

A smooth forward curve constrained by closing prices

We present here the procedure to extract the forward dynamics in a general situation with a fixed number of contracts from the market, even though we will often make references to our own case. Before presenting the algorithm we need to introduce a procedure in order to deal with overlapping periods. Let

$$\mathcal{T} = \{(T_1^s, T_1^e), \ldots, (T_m^s, T_m^e)\}$$

be a list of start and end dates for the settlement periods of m different futures contracts for a given day (in our case, $m = 16$, typically). We need to be able to handle the problem of overlapping settlement periods to rule out arbitrage opportunities, as we have pointed out in Section 2.4.
This was a concrete issue working with our data because it happens that, in a given day, two or more contracts have delivery periods that overlap. To overcome this, we create a new list of dates $\widetilde{\mathcal{T}}$, namely

$$\widetilde{\mathcal{T}} = \{T_0, T_1, \ldots, T_n\},$$

where overlapping contracts are split into sub-periods. In our case n is typically 24, T_0 denotes the starting day of the contract with the closest delivery period, while T_n denotes the last day of the contract with the farthest delivery period. Figure 2.2 outlines how this procedure works, namely the elements of the list $\widetilde{\mathcal{T}}$ are basically the elements in $\mathcal{T}$ sorted in ascending order, with duplicate dates removed.

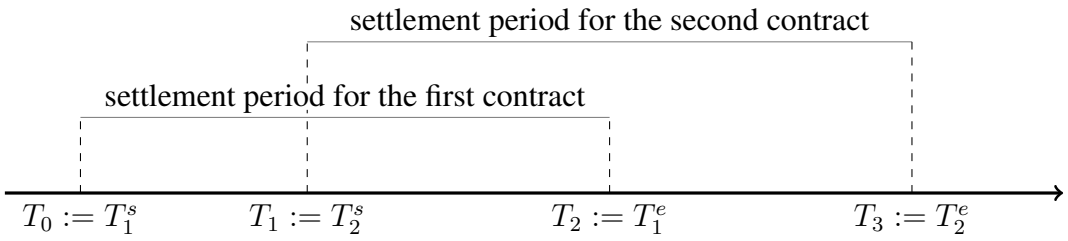


Figure 2.2: Dealing with overlapping delivery time windows.

The futures prices could be taken into account either by exact matching or by a constraint on the bid-ask spread prices. Dealing with closing prices, here we impose an exact matching prices on them (see Equation (2.24)). From now on we denote with F_i^C the closing price for the future i, $i \in \{1, \dots, m\}$. The adjustment functions ε is chosen in the class $\mathcal{C}$, that we recall to be the set of polynomial spline functions of order four, namely

$$\varepsilon(u; \boldsymbol{x}) = \begin{cases} a_1 u^4 + b_1 u^3 + c_1 u^2 + d_1 u + e_1, & u \in [T_0, T_1], \\ a_2 u^4 + b_2 u^3 + c_2 u^2 + d_2 u + e_2, & u \in [T_1, T_2], \\ \vdots & \\ a_n u^4 + b_n u^3 + c_n u^2 + d_n u + e_n, & u \in [T_{n-1}, T_n]. \end{cases}$$

where $\boldsymbol{x}' = [a_1, b_1, c_1, d_1, e_1, \dots, a_n, b_n, c_n, d_n, e_n]$ is the row vector of the coefficients of the splines we want to find and where, with a slight abuse of notation, we use $\varepsilon(u; \boldsymbol{x})$ instead of $\varepsilon(u)$ to stress the dependence on the vector coefficients $\boldsymbol{x}$.

In this way we have, roughly speaking, a spline for every settlement period. To find the unknown parameters $\boldsymbol{x}' = [a_1, b_1, c_1, d_1, e_1, \dots, a_n, b_n, c_n, d_n, e_n]$ in order to fully recover the adjustment function, we need to solve the following equality constrained convex quadratic programming problem

$$\min_{\boldsymbol{x} \in \mathbb{R}^{5n}} \int_{T_0}^{T_n} \left[\varepsilon''(u; \boldsymbol{x})\right]^2 du, \tag{2.19}$$

subject to the following constraints:

i) continuity of the derivatives up to second order at the knots, for $j = 1, \dots, n-1$,

$$a_{j+1} T_j^4 + b_{j+1} T_j^3 + c_{j+1} T_j^2 + d_{j+1} T_j + e_{j+1} = a_j T_j^4 + b_j T_j^3 + c_j T_j^2 + d_j T_j + e_j, \tag{2.20}$$

$$4a_{j+1} T_j^3 + 3b_{j+1} T_j^2 + 2c_{j+1} T_j + d_{j+1} = 4a_j T_j^3 + 3b_j T_j^2 + 2c_j T_j + d_j, \tag{2.21}$$

$$12a_{j+1} T_j^2 + 6b_{j+1} T_j + 2c_{j+1} = 12a_j T_j^2 + 6b_j T_j + 2c_j, \tag{2.22}$$

ii) flatness at the end (see Equation (2.18))

$$\varepsilon'(T_n; \boldsymbol{x}) = 0, \tag{2.23}$$

iii) matching of the closing prices (see Equation (2.12)), for $i = 1, \dots, m$,

$$F_i^C = \frac{1}{T_i^e - T_i^s} \int_{T_i^s}^{T_i^e} [\Lambda(u) - \Lambda(0) + \varepsilon(u; \boldsymbol{x})]\, du. \tag{2.24}$$

In this way the minimisation problem (2.19) has a total of $3n + m - 2$ constraints (i.e., $3(n-1)$ constraints from (2.20)-(2.22), one constraint from (2.23) and m constraints from (2.24)).
In the next Proposition we show how to re-write problem (2.19) in a more compact way to fully exploit it.

Proposition 2.5.5. *The minimisation problem* (2.19) *can be equivalently re-written as*

$$\min_{\boldsymbol{x} \in \mathbb{R}^{5n}} \boldsymbol{x}' \mathbf{H} \boldsymbol{x}, \tag{2.25}$$

where

$$\mathbf{H} = \begin{bmatrix} h_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & h_n \end{bmatrix} \quad \textit{with} \quad h_j = \begin{bmatrix} \frac{144}{5}\Delta_j^5 & 18\Delta_j^4 & 8\Delta_j^3 & 0 & 0 \\ 18\Delta_j^4 & 12\Delta_j^3 & 6\Delta_j^2 & 0 & 0 \\ 8\Delta_j^3 & 6\Delta_j^2 & 4\Delta_j^1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{2.26}$$

and

$$\Delta_j^l = T_j^l - T_{j-1}^l, \tag{2.27}$$

for $j = 1, \dots, n$, *and* $l = 1, \dots, 5$.

Proof. Without loss of generality, we compute h_1, the first block of $\mathbf{H}$ in (2.26), i.e. the block referring to the first contract that has delivery period $[T_0, T_1]$, the other blocks can be computed in a similar manner. We have to evaluate the integral in (2.19) for the first settlement period. After computing the second derivatives of $\varepsilon(u; \boldsymbol{x})$ and expanding the integrals we get

$$\begin{aligned} &\int_{T_0}^{T_1} (12a_1u^2 + 6b_1u + 2c_1)^2\, du = \\ &= \int_{T_0}^{T_1} (144a_1^2u^4 + 36b_1^2u^2 + 4c_1^2 + 144a_1b_1u^3 + 48a_1c_1u^2 + 24b_1c_1u)\, du = \\ &= \left[\frac{144}{5}a_1^2u^5 + 12b_1^2u^3 + 4c_1^2u + 36a_1b_1u^4 + 16a_1c_1u^3 + 12b_1c_1u^2\right]_{T_1}^{T_2} = \\ &= \frac{144}{5}a_1^2(T_1^5 - T_0^5) + 12b_1^2(T_1^3 - T_0^3) + 4c_1^2(T_1 - T_0) + 36a_1b_1(T_1^4 - T_0^4) + \\ &+ 16a_1c_1(T_1^3 - T_0^3) + 12b_1c_1(T_1^2 - T_0^2). \end{aligned} \tag{2.28}$$

Since (2.28) is an homogeneous polynomial of degree two w.r.t. a_1, b_1, c_1, d_1, e_1, we can re-write it as a quadratic form. Indeed, if we call $\boldsymbol{x}_1' = [a_1, b_1, c_1, d_1, e_1]$, we can rewrite (2.28) as

$$\boldsymbol{x}_1' \boldsymbol{h}_1 \boldsymbol{x}_1$$

where

$$\boldsymbol{h}_1 = \begin{bmatrix} \frac{144}{5}(T_1^5 - T_0^5) & 18(T_1^4 - T_0^4) & 8(T_1^3 - T_0^3) & 0 & 0 \\ 18(T_1^4 - T_0^4) & 12(T_1^3 - T_0^3) & 6(T_1^2 - T_0^2) & 0 & 0 \\ 8(T_1^3 - T_0^3) & 6(T_1^2 - T_0^2) & 4(T_1 - T_0) & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and this is according to (2.26) and (2.27). □

As far as the constraints (2.20) - (2.24) are concerned, we clearly see they are linear w.r.t. $\boldsymbol{x}$. Hence, we can gather them in a matrix form as $\mathbf{A}\boldsymbol{x} = \mathbf{b}$, where $\mathbf{A}$ is a $(3n + m - 2) \times 5n$-dimensional matrix, and $\mathbf{b}$ is a $(3n + m - 2)$-dimensional vector. Solving the problem (2.19) with the constraints (2.20) - (2.24) is equivalent to solving (2.25) with the constraints written in the form $\mathbf{A}\boldsymbol{x} = \mathbf{b}$. Let $\lambda' = [\lambda_1, \lambda_2, \ldots, \lambda_{3n+m-2}]$ be the corresponding Lagrange multiplier vector to the constraints (2.20) - (2.24). Therefore, we can now express (2.19) as the following unconstrained minimization problem

$$\min_{\boldsymbol{x} \in \mathbb{R}^{5n}, \lambda \in \mathbb{R}^{3n+m-2}} \boldsymbol{x}'\mathbf{H}\boldsymbol{x} + \lambda'(\mathbf{A}\boldsymbol{x} - \mathbf{b}). \tag{2.29}$$

Remark 2.5.6. *The advantage of dealing with problem* (2.29)*, instead of problem* (2.19) *with the constraints* (2.20) - (2.24)*, is that* (2.29) *is a unconstrained problem that can be simply solved. Indeed, the solution* $[\bar{\boldsymbol{x}}, \bar{\lambda}]$ *is obtained just solving the linear system*

$$\begin{bmatrix} 2\mathbf{H} & \mathbf{A}' \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b} \end{bmatrix}. \tag{2.30}$$

The dimension of the left matrix is $(8n + m - 2) \times (8n + m - 2)$*. Solving* (2.30) *numerically is standard, and can be done using various techniques (e.g. QR or LU factorisation)*

All the code and the computations have been implemented in MATLAB R2018a, on a CPU 2.6 GHz and 12 Gb of RAM HP Notebook with Windows 10.

2.5.2 Numerical Results

Recall that we are working with French Futures closing prices from 2002 to 2019. The yearly contracts span from 2002 to 2019, the quarterly from 2011 to 2019, the monthly from 2011 to 2019 and the weekly from 2010 to 2019. We are working with *rolling contracts*, called $c1$, $c2$, $c3$, and $c4$. Figure 2.3 shows how these contracts roll in the case of a monthly contract.

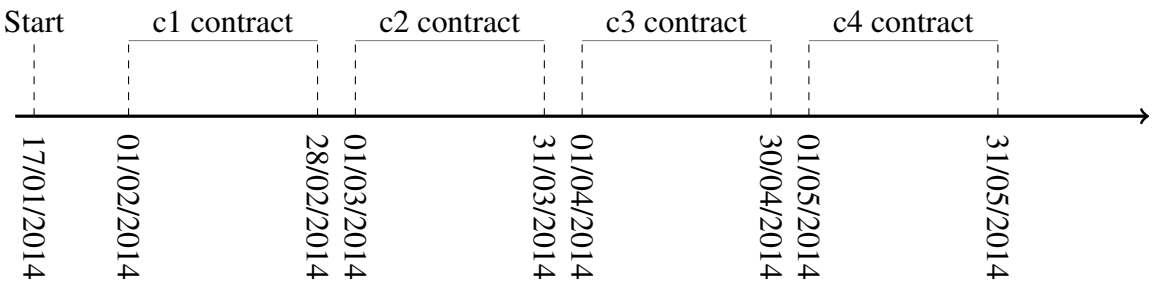
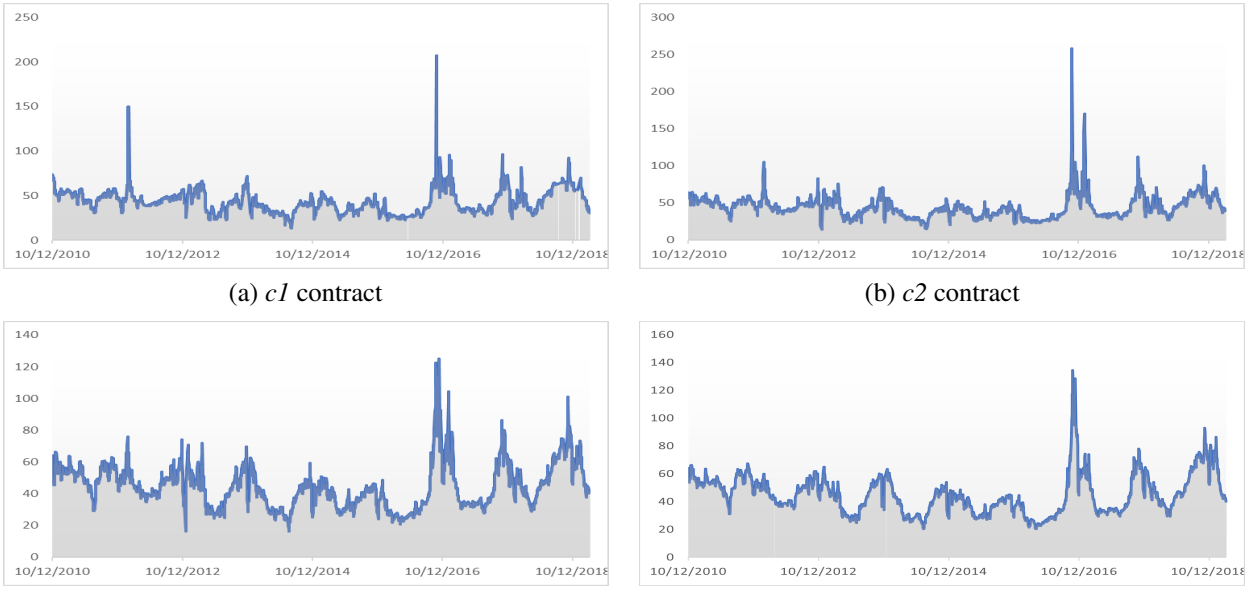


Figure 2.3: An example of how rolling contracts work, in the case when Today, the "0" date, is January $17, 2014$.

As you can see from Figure 2.3, the c1 contract is the closer one to the current date and its delivery period spans from 01/02/2014 to 28/02/2014. After 28/02/2014, there is a rollover from the c1 contract to the c2 contract and so on and so forth.

Figures 2.4, 2.5, 2.6 and 2.7 show the plot of the futures closing prices in Euros of the $c1, c2, c3$ and $c4$ contracts for the weekly, quarterly, monthly and yearly contracts, respectively.

In the x-axis there are the different dates, while in the y-axis there is the price. As you can see, the presence of seasonality is pretty strong, especially in quarterly contracts.



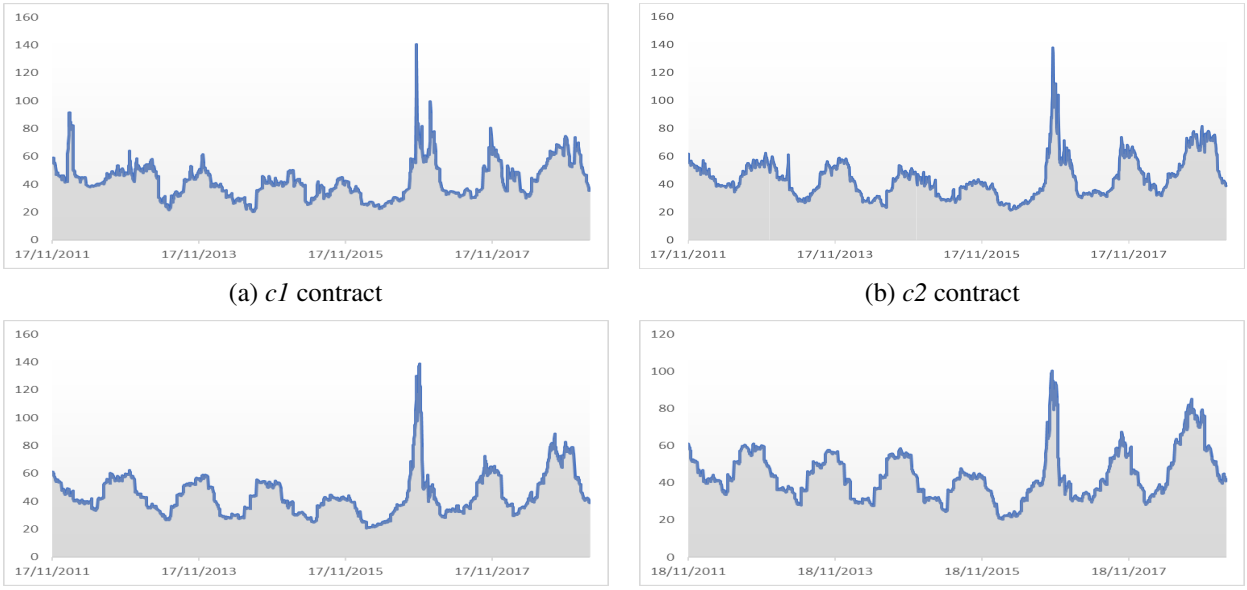
(a) *c1* contract

(b) *c2* contract

(c) *c3* contract

(d) *c4* contract

Figure 2.4: Weekly contract prices: the *c1* and *c4* contracts start on December $10, 2010$ and end on March $14, 2019$, while the *c2* and *c3* contracts start on December $10, 2010$ and end on March $15, 2019$.



(a) *c1* contract

(b) *c2* contract

(c) *c3* contract

(d) *c4* contract

Figure 2.5: Monthly contract prices: the *c1* contract starts on November $17, 2011$ and ends on March $14, 2019$, the *c2* and *c3* contracts start on November $17, 2011$ and end on March $15, 2019$ while the *c4* contract starts on November $18, 2011$ and ends on March $15, 2019$

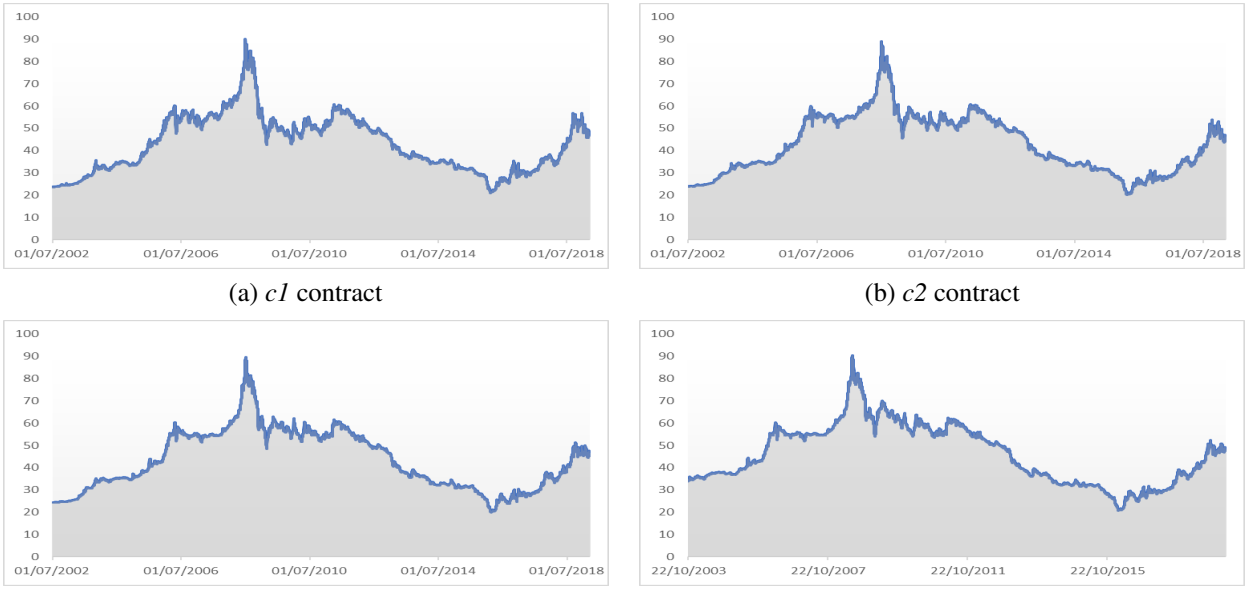
(a) *c1* contract

(b) *c2* contract

(c) *c3* contract

(d) *c4* contract

Figure 2.6: Quarterly contract prices: the *c1*, *c2* and *c3* contracts start on November $28, 2011$ and end on March $15, 2019$ while the *c4* contract starts on November $28, 2011$ and ends on March $14, 2019$



(a) *c1* contract

(b) *c2* contract

(c) *c3* contract

(d) *c4* contract

Figure 2.7: Yearly contract prices: the *c1* contract starts on July $1, 2002$ and ends on March $15, 2019$, the *c2* and *c3* contracts start on July $1, 2002$ and end on March $14, 2019$ while the *c4* contract starts on October $22, 2003$ and ends on March $14, 2019$

Since we were worried that our analysis could have been affected by the presence of the quarterly contracts being sensible to the seasonality pattern, we did the analysis in both cases, with and without considering the quarterly contracts. After the analysis was performed, we noticed that the results were coherent in both cases, so from now on we will focus only on contracts different from quarterly.

The algorithm presented takes approximately 40 seconds to extract the 4234 different curves, i.e. the de-seasonalized curves $f(0, u)$, for all the 4234 different values of "0".

Figure 2.8 shows, for every day from 2002 to 2019, a different de-seasonalized curve $f(0, u)$, for a total number of 4234 curves. In the x-axis we have the time to maturity while in the y-axis we have the prices. Note that different colours in the curves mean different type of contracts. Note also that the further we move on the x-axis, the flatter the curves become. This is in line with the flatness constraint in Equation (2.18). The reason behind the difference in the shapes of the curves is to be investigated in the price constraint but also in the nature of the contracts, since for each day the number and the type of available contracts were different, having to deal also with overlapping settlement periods. From now on, with a slightly abuse of notation, we refer to f as the de-seasonalized forward curves.

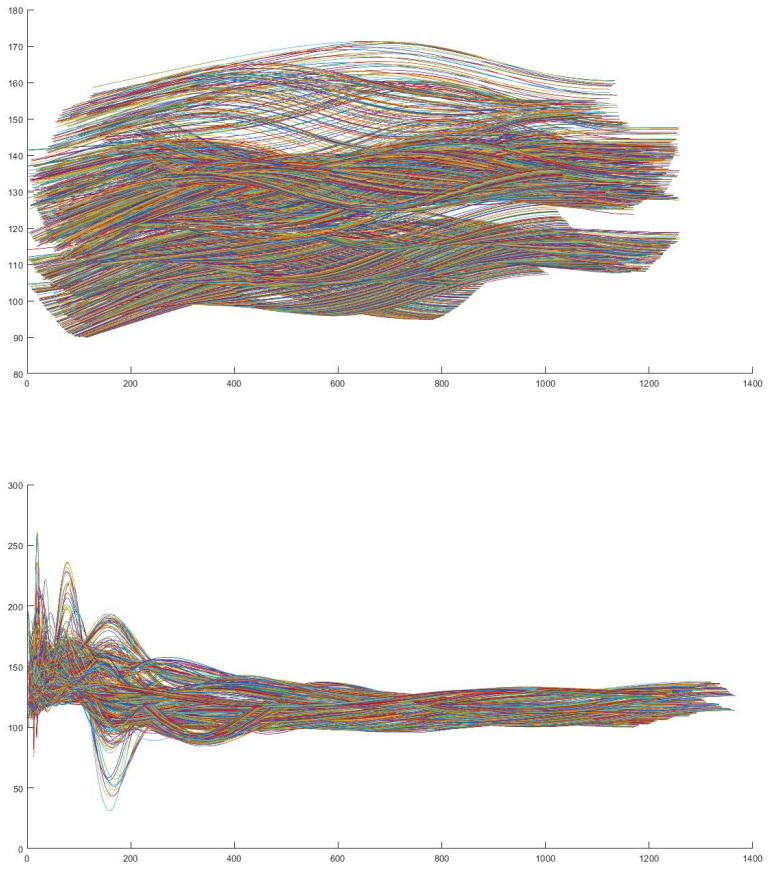


Figure 2.8: Some of the de-seasonalized forward curves extracted without considering quarterly contracts. In the x-axis there is the time to maturity while in the y-axis there are the prices in Euros.

2.6 Jump Detection

We now want to detect the jumps. We will be only dealing with positive jumps since we have supposed that the jump size is distributed according to an exponential density, as already described in Section 2.3. In the next subsection we will describe an algorithm that allows to detect jumps, and, as a by-product, which also gives their size.

2.6.1 Description of the Algorithm

In order to detect jumps we proceed in the following way: for a fixed maturity T, we define

$$V_t = f(t, T),$$

the vertical section at maturity T, where the parameter t ranges through all the curves, i.e. $t = 1, \dots, 4234$. Roughly speaking, looking at Figure 2.8, this is nothing but the intersection between the vertical line $x = T$ and the curves. There are several ways to detect jumps from the data in the literature, the more natural one can think of is to fix a threshold $\Theta \in \mathbb{R}_+$ and to identify a jump at time $\tilde{t}$ if the condition $|V_{\tilde{t}+1} - V_{\tilde{t}}| \geq \Theta$ is satisfied.
Here we follow an iterative weighted least square approach, inspired by the work of Callegaro et al. (2017). Define $n = 4234$ the total number of curves and $\mathcal{N} = \{1, 2, \dots, n-1\}$. The algorithm to detect jumps reads as follows:

1. Define $\sigma_1^2 = \frac{1}{n-2} \sum_{t \in \mathcal{N}} \frac{(V_{t+1} - V_t)^2}{V_t}$;
2. Identify all the $t \in \mathcal{N}$ such that $\frac{V_{t+1} - V_t}{\sqrt{V_t}} \geq 3\sigma_1$ and denote by $\mathcal{M}_1 \subseteq \mathcal{N}$ this family of indices, so that $m_1 = |\mathcal{M}_1|$;
3. Define $\sigma_2^2 = \frac{1}{n - m_1 - 1} \sum_{t \in \mathcal{N} \setminus \mathcal{M}_1} \frac{(V_{t+1} - V_t)^2}{V_t}$;
4. Identify all the $t \in (\mathcal{N} \setminus \mathcal{M}_1)$ such that $\frac{V_{t+1} - V_t}{\sqrt{V_t}} \geq 3\sigma_2$ and denote by $\mathcal{M}_2 \subseteq (\mathcal{N} \setminus \mathcal{M}_1)$ this family of indices, so that $m_2 = |\mathcal{M}_2|$;
5. Iterate the procedure updating $\sigma_i^2 = \frac{1}{n - (\sum_{j=1}^{i-1} m_j) - 1} \sum_{t \in \mathcal{N} \setminus (\bigcup_{j=1}^{i-1} \mathcal{M}_j)} \frac{(V_{t+1} - V_t)^2}{V_t}$ and $\mathcal{M}_i \subseteq \mathcal{N} \setminus (\bigcup_{j=1}^{i-1} \mathcal{M}_j)$;
6. Stop when finding $k \in \mathbb{N}$ such that $m_k = |\mathcal{M}_k| = 0$ (no new jumps are detected).

This procedure finds, at every iterations, new jumps. Clearly, as the number of iterations increases, σ_i decreases and so the jumps detected become smaller and smaller. After several tests on the data, we noticed that stopping at $k \in \mathbb{N}$ such that $m_k = 0$ would lead to too many jumps, the last detected being much smaller compared to the ones discovered at the first iterations. Indeed, as you can see from Table 2.5, after the second iteration, the value of sigma drops considerably, leading to an unreasonable number of small size jumps. Therefore, we chose, as a good compromise,

to stop the algorithm after the first two iterations, because, otherwise, the presence of the jumps detected after the second iteration could have severely biased the analysis.

T	200	400	700
1° iteration	0.4581	0.2743	0.0909
2° iteration	0.2187	0.1295	0.0664
3° iteration	0.0712	0.0212	~ 0

Table 2.1: Values of sigma at different maturities and at different iterations.

2.6.2 Jumps Analysis

We selected different values of T, namely $T = 200$, $T = 400$ and $T = 700$ days. This covers all the different shapes of the curves and so it represents a good sampling of our data. Figure 2.9 shows the size and distribution of the jumps detected at $T = 200$, $T = 400$ and $T = 700$, after running two iterations of the algorithm described in Subsection 2.6.1.

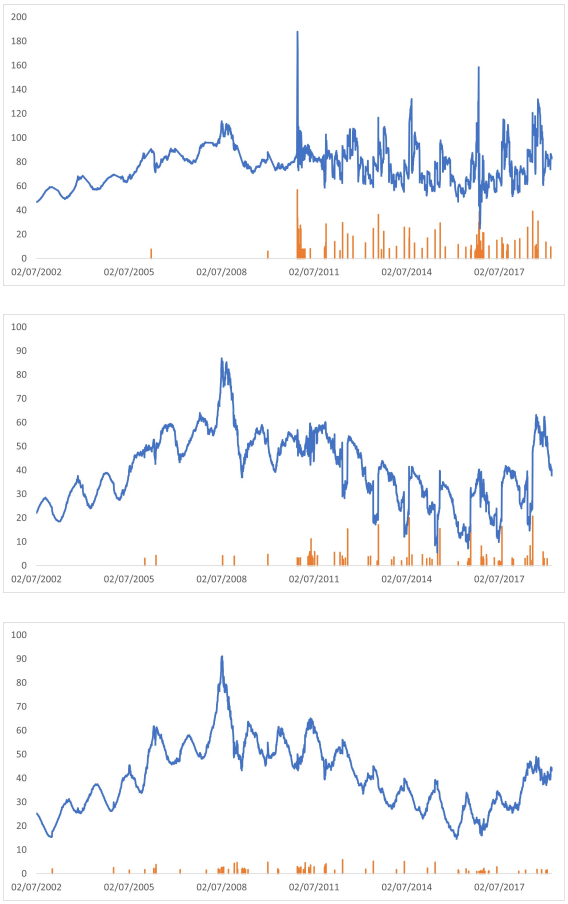


Figure 2.9: Jumps detected (vertical orange lines) and price plot (blue line) at $T = 200$, $T = 400$ and $T = 700$, from the top to the bottom, respectively.

Table 2.2 shows the number of jumps detected at each iteration, according to the different value of T:

T	200	400	700
1° iteration	38	19	43
2° iteration	48	55	36
Total	86	74	79

Table 2.2: Number of jumps detected at different maturities and at different iterations.

As one can see from Figure 2.9, at $T = 200$ the jumps detected are the bigger ones with respect to their amplitude, and this is not surprising looking at Figure 2.8, where one can clearly see that the price movements are pretty significant at $T = 200$. On the other hand, when $T = 700$, the jumps detected are quite small, the reason being that at $T = 700$ the curves are pretty flatten, leading to prices which are close to each other.

2.7 Parameters Estimation

Before starting with the statistical tests on the two models, we still have to estimate: the size of the jumps and the parameters characterizing the drift and the volatility coefficients. We start with δ, the jumps' size. Recall that in Subsections 2.3.1 and 2.3.2 we assumed for both models the jumps' size to be distributed like an exponential random variable with parameter $\delta > 0$. Let z_i be the size of the i-th jump, where $i = 1, \dots, L$ and L is the number of jumps at the chosen maturity T (see Table 2.2). Then δ can be estimated e.g. via its Maximum Likelihood Estimator:

$$\hat{\delta} = \frac{L}{\sum_{i=1}^{n} z_i}. \tag{2.31}$$

Table 2.3 shows the estimated value of δ (the fact that we have the smallest values of $\hat{\delta}$ at $T = 200$ is not surprising at all, because at the beginning jumps are bigger, as said before):

T	200	400	700
$\widehat{\delta}$	0.064282682	0.194300598	0.430239733

Table 2.3: Parameter estimation for δ.

We now need to estimate the parameter appearing in the drift coefficient of our forward dynamics.

Remark 2.7.1. *Notice that neither in Equation* (2.6) *(stated in the CBI framework), nor in Equation* (2.11) *(stated in the Hawkes framework) the mean-reversion term appears, but it does if we*

pass under the measure $\mathbb{P}$ by exploiting, respectively, Propositions 2.3.1 and 2.3.7. In both cases we end up with the following dynamics under $\mathbb{P}$, for a fixed T:

$$\widetilde{X}(t+1,T) = \widetilde{X}(t,T) - \int_t^{t+1} \widetilde{a}\widetilde{X}(s,T)ds$$

where $\widetilde{X}$ denotes the factor appearing in the forward dynamics without the seasonality and with no jumps (i.e., we remove all the times at which a jump has occurred).

In order to estimate $\widetilde{a}$ we simply discretize the above equation, leading to

T	200	400	700
$\widehat{\widetilde{a}}$	-0.001387342	-0.001771845	-0.000237952

Table 2.4: Parameters estimation for $\tilde{a}$

As you can see from Table 2.4, the estimated value of $\widetilde{a}$ in all the three cases is really small, very close to 0.

Now it remains to estimate the volatility parameter σ, appearing in both Equations (2.6) and in Equation (2.11). By recalling the iterative algorithm presented in Subsection 2.6.1 and taking as $\widehat{\sigma}$ the value of σ_2 (namely, the estimation after the second iteration), we get

T	200	400	700
$\widehat{\sigma}$	0.218667945	0.129560813	0.066361067

Table 2.5: Parameter estimation for σ.

2.8 Testing the Models

In this section we want to perform statistical tests concerning the intensity of the jumps. We want to check what is the best process modelling the jumps we have detected in Section 2.6. We test the two models based on Hawkes and branching processes, plus the Poisson, which is a toy-model:

(0) Poisson process;

(1) Hawkes process;

(2) Self exciting branching process.

We will mainly rely on the Kolmogorov-Smirnov (KS) test, namely we will test the null hypothesis H_0, stating that the data have the same cumulative distribution function as the one coming from one of the above models, against the alternative H_1. We fix a significance level equal to 0.05.

2.8.1 Jump Intensity Estimation

Before using the KS test to check whether the jumps distribution comes from one of the three models, we need to estimate the intensity from our data. The input in all the cases will be the time occurrences of the jumps over $[0, T]$ (for the three different values of T), $0 < \tau_1 < \tau_2 < \cdots < \tau_N = T$, where N can take the values 86, 74 and 79 depending on the chosen maturity T, as you can check from Table 2.2.

(0) [Poisson] The (constant) intensity, $\lambda^P > 0$, is estimated as the ratio between the total number of (positive) jumps and the sum of the inter-times between two consecutive jumps.

(1) [Hawkes] Here the intensity is given by Equation (2.8), so this case will be treated in a separate subsection.

(2) [Branching] In this case the stochastic intensity $\lambda^B(t) \propto X(t, T)$ and the constant of proportionality γ is estimated, for a fixed T, as the ratio between the total number of (positive) jumps and the cumulative (de-seasonalized) forward prices.

The Hawkes Setting: Estimating λ

Recalling Equation (2.8), it is clear that we have to estimate three parameters: $\lambda(0)$, α and β. We mainly rely on the paper by Ozaki (1979) to find a Maximum Likelihood Estimation (MLE).
The log-likelihood of a Hawkes process whose response function is of the form $\alpha e^{-\beta t}$, is given by

$$\log L(\tau_1, \cdots, \tau_N) = -\lambda(0)\tau_N + \sum_{i=1}^{N} \frac{\alpha}{\beta}\left(e^{-\beta(\tau_N - \tau_i)} - 1\right) + \sum_{i=1}^{N} \log\left(\lambda(0) + \alpha A(i)\right), \qquad (2.32)$$

where $A(i) = \sum_{\tau_j < \tau_i} e^{-\beta(\tau_i - \tau_j)}$ for $i \geq 2$ and $A(1) = 0$.

In order to estimate the parameters $\lambda(0), \alpha, \beta$, we need to find the maximum of the function in (2.32), which is a real value function of three variables. After adapting the procedure to our framework, the maximum was found using MATLAB, via the *fminsearch* function.

Tables 2.6, 2.7 and 2.8 show the parameters estimated for the jump intensity of the three different models at maturities $T = 200$, $T = 400$ and $T = 700$, respectively. As far as the branching model is concerned, $\lambda^B(t)$ is proportional to the process $X(t, T)$ (for a fixed T), and the parameter to be estimated is γ, which is the constant ratio between the two processes.

Model	Parameters				
	$\lambda(0)$	α	β	λ^P	γ
Poisson	–	–	–	0.023	–
Hawkes	0.017	0.074	0.094	–	–
Branching	–	–	–	–	0.00028

Table 2.6: Parameters estimation for the three models at $T = 200$.

Model	Parameters				
	$\lambda(0)$	α	β	λ^P	γ
Poisson	–	–	–	0.018	–
Hawkes	0.0026	0.012	0.016	–	–
Branching	–	–	–	–	0.00021

Table 2.7: Parameters estimation for the three models at $T = 400$.

Model	Parameters				
	$\lambda(0)$	α	β	λ^P	γ
Poisson	–	–	–	0.0019	–
Hawkes	0.040	0.059	0.085	–	–
Branching	–	–	–	–	0.00032

Table 2.8: Parameters estimation for the three models at $T = 700$.

2.8.2 KS test for the models

We perform a Kolmogorov-Smirnov test in order to check which of the proposed distributions best models the jumps in our data. At the end of the subsection we will provide the p-values to conclude.

(0) [Poisson] We check whether the jumps inter-times are drawn from an exponential distribution with parameter λ^P, where λ^P is the one given in Tables 2.6, 2.7 and 2.8. The p-value is automatically given by MATLAB via the *kstest* function.

(1) [Hawkes] We mainly adapt the methods in Lallouache and Challet (2016) to our purpose. In particular, we check if the time-deformed series of durations $\{\theta_i\}_{i=1,\dots,N}$, defined by

$$\theta_i = \int_{\tau_{i-1}}^{\tau_i} \widehat{\lambda}_t dt, \tag{2.33}$$

has an exponential distribution of parameter 1, where $\widehat{\lambda}_t$ is the intensity estimated before (the estimated parameters can be found in Tables 2.6, 2.7 and 2.8), and where recall that the τ_i's are the jumps arrival times. The p-value is automatically given by Matlab via the *kstest* function.

(2) [Branching] The procedure here is quite different from the ones adopted before and it is the object of the following subsection.

Setting the KS Test for the Branching Model

In order to perform the KS test in this case, we rely on the following classical result.

Proposition 2.8.1. *Let $(N_t)_{t\geq 0}$ be a non homogeneous Poisson process with continuous expectation function. If n events have occured in $(0,T]$, then the arrival times $\tau_1,\dots,\tau_n$ are distributed as the order statistics from a sample with cumulative distribution function*

$$F(t) = \frac{\int_0^t \lambda^B(s)ds}{\int_0^T \lambda^B(s)ds}, \quad 0 \leq t \leq T, \tag{2.34}$$

where in our case $\lambda^B(s) \propto f(s,T)$, for a fixed T.

Remark 2.8.2. *Recall that $\lambda^B(s) \propto f(s,T)$, for any fixed T. It is crucial to notice, from Proposition 2.8.1, that the distribution of the arrival times is independent of the factor of proportionality connecting λ^B and $f(\cdot,T)$.*

Therefore, in this case we perform a KS test, comparing the cumulative distribution function $F(t)$ in Equation (2.34) with the empirical one relative to the jump times.
Since $F(t)$ given in Proposition (2.8.1) is not a priori associated to a known distribution, we have to rely on the classical theory on the KS test. The KS statistics for the test is

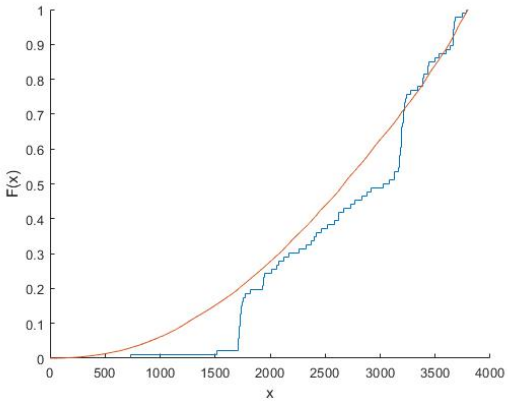
$$D_n = \sup_{x\in\mathbb{R}} |S_n(x) - F(x)|, \tag{2.35}$$

where n is the number of our data (recall Table 2.2), F is the cumulative distribution function in Equation (2.34) and $S_n(x)$ is the empirical cumulative distribution function of the jump arrival times. Table 2.9 shows the results.

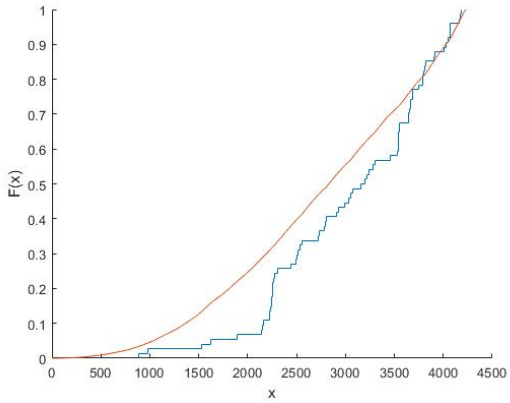
T	200	400	700
D_n	0.2151	0.2276	0.2513

Table 2.9: Maximum distance between the empirical distribution function and the theoretical one for the Branching case.

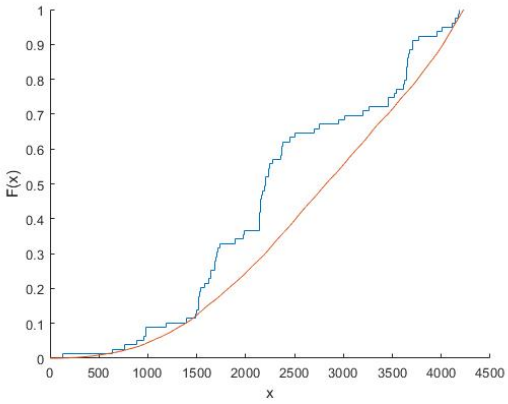
Figure 2.10 graphically compares the two cumulative distribution functions at the three different values of T.



(a) T=200



(b) T=400



(c) T=700

Figure 2.10: Comparison between the cumulative distribution function of F (red line) and the empirical cumulative distribution function of the jumps arrival times (blue line) for the three different values of T.

In order to obtain the p-value in the branching case, we apply the asymptotic results in Facchinetti (2009) in the case when the dataset is greater than 35. Table 2.10 provide the p-values for each of the three models.

Values of T	200	400	700
Poisson	~ 0	~ 0	~ 0
Branching	0.041	0.018	0.042
Hawkes	0.23	0.31	0.13

Table 2.10: p-value test for the three models for different T

As Table 2.10 suggests, the hypothesis that the intensity follows a Poisson process is highly rejected, as we expected. In the branching case the hypothesis is also rejected, even if the p-value in this case was much closer to the acceptance level of 0.05. For the Hawkes case the test fails to reject the hypothesis since all the three values are above our level of acceptance. Therefore, we can state that, up to our numerical experience, the Hawkes model seems to be the one that best describe the jumps found in our numerical analysis.

2.9 Conclusion

As a conclusion we can resume the main achievement presented in the present chapter. We proposed two alternative models for power forward prices evolution, based on a HJM approach, extending to forward prices dynamics two models already proposed for the spot price dynamics by Eyjolfsson and Tjøstheim (2018) and Jiao et al. (2019). After extracting the forward curves from quoted futures prices, we proposed a parameters estimation method for both models and then we performed a test on the adequacy of the two models in describing the observed forward prices evolution. The final conclusion of our test is that the hypothesis that forward prices follow a CBI-type dynamics is rejected, while the hypothesis of a Hawkes type dynamics is not. This conclusion suggests that self-exciting effects can arise in power forward dynamics as well as in the spot dynamics, and that an approach based on Hawkes processes can capture these effects in a natural and parsimonious way.

Chapter 3

A general framework for a joint calibration of VIX and VXX options

3.1 Introduction

Since the last decades, a significant and increasing number of volatility-related products has appeared in financial markets. The great financial crisis accelerated the demand for volatility instruments, in a context of pronounced market uncertainty, see Sussman and Morgan (2012). Created by the CBOE[1] in 1993, the VIX index is by far one of the most popular volatility instruments and it is considered as a landmark by most market players. VIX is a real-time market index that represents the market's expectation of 30-day forward-looking volatility. Derived from the price inputs of the S&P 500 index options, it provides a measure of market risk and investors' sentiments. Investors, research analysts and portfolio managers look to VIX values as a way to measure market risk, fear and stress before they take investment decisions, this being a reason why it is also known as the "Fear Index".

In the beginning, VIX was computed as a weighted measure of the 30-day implied volatility of eight S&P 100 at-the-money put and call options, when the derivatives market had limited activity and was in growing stages. Ten years later, changes in its calculation were made following the introduction of a wider set of options based on the broader S&P 500 index, an expansion that allowed for a more accurate view of Investors' expectations on future market volatility. Like all indexes, one cannot buy the VIX directly. Instead, investors can either take position in VIX through futures or options contracts, or through VIX-based exchange-traded products. Indeed, VIX became tradable via futures in 2004 and via options in 2006, hence providing market players new instruments to trade implied volatility in a direct manner, see Whaley (2009). Since then, volumes on VIX futures and options have drastically increased. For instance, the VIX became the second most traded underlying in the CBOE options market[2] in 2015, right after the S&P 500

[1] See the white paper of CBOE in 2014 titled *Cboe Volatility Index*, available at https://www.cboe.com/micro/vix/vixwhite.pdf.

[2] See the white paper of CBOE in 2016 titled *CBOE holdings reports trading volume for 2015*, available at http://ir.cboe.com/~/media/Files/C/CBOE-IR-V2/press-release/2016/

itself. As derivatives on the VIX may be inaccessible to non-institutional players, primarily due to the large notional sizes of the contracts as currently designed by the CBOE, some exchange-traded notes (ETNs) on the VIX were introduced. In 2009, Barclays launched the first two ETNs on the VIX: VXX and VXZ. The VXX ETN was the first exchange-traded product (ETP) on VIX futures, issued in 2009 by Barclays Capital iPath[3], shortly after the inception of the VIX futures indices. The VXX is a non-securitized debt obligation, similar to a zero-coupon bond, but with a redemption value that depends on the level of the S&P 500 (SPX) VIX Short-Term Futures Total Return index (SPVXSTR). The SPVXSTR tracks the performance of a position in the nearest and second-nearest maturing VIX futures contracts, which is rebalanced daily to create a nearly constant 1-month maturity. In other words, VXX mimicks the behaviour of the VIX, namely of the 30-day forward-looking volatility, using a replicable strategy based on futures on VIX.

Since 2009 then, ETNs has then flourished on the VIX: currently more than thirty of them exist, with several billion dollars in market caps and daily volumes (see Alexander and Korovilas (2013) for a comprehensive empirical study on VIX ETNs). While the introduction of ETNs on VIX enabled new investors to trade market volatility while enjoying the benefits of small nominals and bid-ask spreads, the performance of those products raised concerns among market players and regulators[4]. Even though VXX is very heavily traded, it has lost 99.84% of its value since inception and Figure 3.1 shows the remarkable underperformance of VXX compared to VIX. Whereas many market players use VXX as a proxy for trading VIX in a cheap manner, in the VXX Prospectus[3] Barclays warns VXX holders that "Your ETN is not Linked to the VIX index and the value of your ETN may be less than it would have been had your ETN been linked to the VIX Index".

Attempts to model VIX can be divided in two strands: on one hand, the consistent-pricing approach models the joint (risk-neutral) dynamics of the S&P 500 and the VIX mostly with the aim of pricing derivatives on the two indices in a consistent manner, on the other hand, the stand-alone approach directly models the dynamics of the VIX.

In the former case, we can include the work of Cont and Kokholm (2013) and references therein, while more applications can be found in the work of Gehricke and Zhang (2018), where they propose a consistent framework for S&P 500 (modelled with a Heston-like specification) and the VIX derived from the square root of the variance swap contract. In addition, they construct the VXX contract from the Prospectus and show that the roll yield of VIX futures drives the difference between the VXX and VIX returns on time series. More recently, in the paper of Gatheral et al. (2020), a joint calibration of SPX and VIX smile has been successfully obtained using the quadratic rough Heston model, while Guyon (2020b) builds up a non-parametric discrete-time model that jointly and exactly calibrates to the prices of SPX options, VIX futures and VIX options. We also mention the work of Guyon (2020a), where the author investigates conditions for the existence of a continuous model on the S&P 500 index (SPX) that calibrates the full surface of SPX and VIX implied volatilities, and the paper of De Marco and Henry-Labordere (2015), where the authors

cboe-holdings-volume-report-december-2015.pdf.

[3]See the VXX Prospectus of Barclays available at https://www.ipathetn.com/US/16/en/instruments.app?searchType=text&dataType=html&searchTerm=VXX

[4]See the white paper of Bloomberg in 2012 titled *SEC said to review Credit Suisse VIX Note* available at https://www.bloomberg.com/news/articles/2012-03-29/sec-said-to-review-credit-suisse-vix-note.

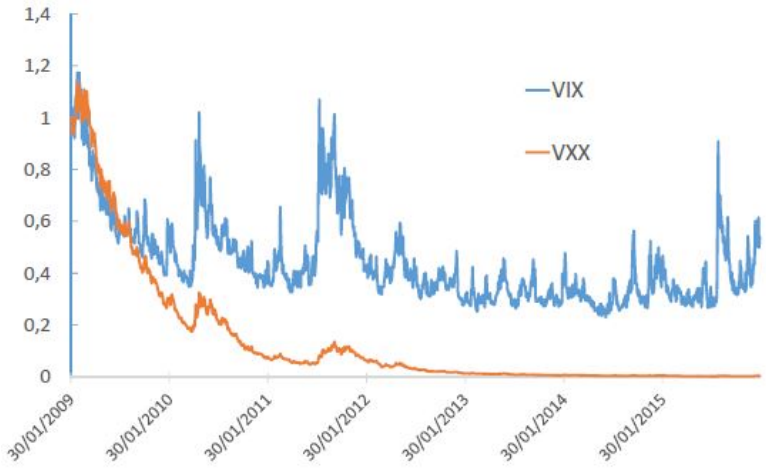


Figure 3.1: Prices in dollars of VIX (in blue) and VXX (in red), normalized to 1 on 30 January 2009.

bound VIX options from vanilla options and VIX futures, which leads to a new martingale optimal transport problem that they solve numerically.

In the latter case, the first stand-alone model for the VIX was proposed by Whaley, see Whaley (1993), where he models the spot VIX as a geometric Brownian motion, hence enabling to price options on the VIX in the Black and Scholes framework, but failing to capture the smile effect and the mean-reversion of volatility processes, which is necessary in order to fit the term structure of futures prices (Schwert (1990), Pagan and Schwert (1990) and Schwert (2011)). We cite also the attempt of Drimus and Farkas (2013) based on local volatility. Since then, numerous stand-alone models have been proposed for the VIX. We refer to Bao et al. (2012) and Lin (2013) for a detailed review on that strand of literature.

The literature on VXX is scarce. To our knowledge, there are only a few theoretical studies which propose a unified framework for VIX and VXX. Existing studies on the VXX (Bao et al. (2012)) typically model the VXX in a stand-alone manner, which does not take into account the link between VIX and VXX. As a consequence, the loss for VXX when the VIX is in contango cannot be anticipated in such models. In the recent paper of Gehricke and Zhang (2020), they analyze the daily implied volatility curves of the VXX options market alone (where they consider the implied volatility as a function of the moneyness through quadratic polynomial regressions), providing a necessary benchmark for developing a VXX option pricing model.

We also mention the work of Grasselli and Wagalath (2018), where the authors were able to link the properties of VXX to those of the VIX in a tractable way. In particular, they quantify the systematic loss observed empirically for VXX when the VIX futures term-structure is in contango and they derive option prices, implied volatilities and skews of VXX from those of VIX in infinitesimal developments. The affine stochastic volatility model of Grasselli and Wagalath (2018) exploits the FFT methodology in order to price futures and options on VIX, while a MonteCarlo simulation is performed in order to compare the resulting smile on the VXX implied by the calibrated model

with the one provided by real market data. The results are not completely satisfactory, which is not surprising, since the VXX can be seen as an exotic path-dependent product on VIX, so that the information regarding the smile on vanillas (on VIX) are well known to be not enough to recover the whole information concerning path dependent products (see e.g. Bergomi (2015)).

In the present paper we leave the pure stochastic volatility model for the VIX and we introduce a stochastic-local volatility model (SLV), that is able to jointly calibrate VIX and VXX options in a consistent manner. The same methodology can be easily extended to other pairs of futures contracts and notes. SLV models are the de-facto industry standard in option pricing since they allow to join the advantages of stochastic volatility models with a good fit of plain options. They were first introduced in Lipton (2002), while a first efficient PDE-based calibration procedure was presented in Ren et al. (2007). In the present paper, since we investigate multi-factor models, we follow the calibration approach of Guyon and Henry-Labordère (2012), which is based on a Monte Carlo simulation. The SLV model we introduce for the VIX futures starts from the work of Nastasi et al. (2020), where the problem of modelling futures smiles for commodity assets in a flexible but parsimonious way is discussed. In particular, here we investigate a multi-factor extension of such model which is taylored to describe VIX futures prices, and to allow a joint calibration of VIX and VXX plain vanilla options.

The key feature of our model is the possibility to decouple the problems of calibrating VIX plain vanilla options, which are recovered by a suitable choice of the local-volatility function, from the more complex task of calibrating the plain vanilla options on the VXX notes. We recall that in our approach the VXX notes are exotic path-dependent product written on VIX futures. We solve the second problem by introducing a local-correlation function between the factors driving the futures prices, and by properly selecting the remaining free dynamics paramters. The calibration of the local correlation function is inspired by the work of Guyon (2017).

The chapter is organized as follows. In Section 3.2 we give a brief overview of SLV models, highlighting their properties and the main advantages in using these models. In Section 3.3, we present the general setting of our model. We start from the evolution of a general ETP and their underlying futures contracts, then we describe how to calibrate futures prices and options on futures. In Section 3.4 we adapt the general framework of the previous section to the case of the VXX ETN strategy. In Section 3.5 a calibration exercise based on real data along with numerical results and comparisons are presented, involving a joint calibration and a parameters sensitivity analysis. In Section 3.6 we draw some conclusion and remarks.

3.2 Stochastic-Local Volatility Models

One of the main advantages of using local volatility (LV) models is their natural modeling of plain-vanilla market volatilities. Indeed, a LV model can be calibrated with extreme precision to any given set of arbitrage-free European vanilla option prices. LV model were first introduced by Dupire et al. (1994) and Derman and Kani (1994). Although well-accepted, LV models have certain limitations, for example, they generate flattening implied forward volatilities, see Rebonato (1999), that may lead to a mispricing of financial products like forward-starting options.

On the other hand, stochastic volatility (SV) models, like the well known Heston model, see Heston (1993), are considered to be more accurate choices for pricing forward volatility sensitive derivatives, see Gatheral (2011). Although the SV models have desired features for pricing, they often cannot be very well calibrated to a given set of arbitrage-free European vanilla option prices. In particular, the accuracy of the Heston model for pricing short-maturity options in the equity market is typically unsatisfactory.

One possible improvement to the previous issues is considering stochastic-local volatility (SLV) models, that take advantages of both LV and SV models properties. SLV models, going back to Lipton (2002), are still an intricate task, both from a theoretical as well as a practical point of view. The main advantage of using SLV models is that one can achieve both a good fit to market series data and in principle a perfect calibration to the implied volatility smiles. In such models the discounted price process $(S_t)_{t\geq 0}$ of an asset follows the stochastic differential equation (SDE)

$$dS_t = S_t \ell(t, S_t) v_t dW_t, \tag{3.1}$$

where $(v_t)_{t\geq 0}$ is some stochastic process taking real values, and $\ell(t, K)$ is a sufficiently regular deterministic function, the so-called leverage function, depending on time and on the current value of the underlying asset. W_t, as usual, is a one-dimensional Brownian motion, possibly correlated with the noise driving the process v. Obviously, SV and LV models are recovered by setting $\ell(t, K) \equiv 1$ or $v \equiv 1$, respectively. At this stage, the stochastic volatility process $(v_t)_{t\geq 0}$ can be very general and by a slight abuse of terminology we call $(v_t)_{t\geq 0}$ stochastic volatility as in the SV model.

The leverage function $\ell(t, K)$ is the crucial part in this model. It allows in principle to perfectly calibrate the market implied volatility surface. In order to achieve this goal, ℓ needs to satisfy the following relation:

$$\ell^2(t, K) = \frac{\sigma^2_{\text{Dup}}(t, K)}{\mathbb{E}\left[v_t^2 \,\middle|\, S_t = K\right]}, \tag{3.2}$$

where σ_{Dup} denotes Dupire's local volatility function, see Dupire (1996). The familiar reader would recognize in Equation (3.2) an application of the Markovian projection procedure, a concept originating from the celebrated Gyöngy Lemma, see Gyöngy (1986), whose idea basically lies in finding a diffusion which "mimicks" the fixed-time marginal distributions of an Itô process.

For the derivation of Equation (3.2), we refer to Guyon and Henry-Labordere (2013). Notice that Equation (3.2) is an implicit equation for ℓ as it is needed for the computation of $\mathbb{E}\left[v_t^2 \mid S_t = K\right]$. It is a standard thing in SLV models and this in turn means that the SDE for the price process $(S_t)_{t\geq 0}$ is actually a McKean-Vlasov SDE, since the probability distribution of S_t enters the characteristics of the equation. Existence and uniqueness results for this equation are not obvious in principle, as the coefficients do not typically satisfy standard conditions like for instance Lipschitz continuity with respect to the so-called Wasserstein metric. In fact, deriving the set of stochastic volatility parameters for which SLV models exist uniquely, for a given market implied volatility surface, is a very challenging and still open problem. In the sequel, we will limit to investigate numerically the validity of the assumption just stated in our specific framework.

3.3 Smile Modelling for ETP on Futures Strategies

We start this section with a wider look to study the volatility smile of a generic ETP based on futures strategies. Then, the results shown in this section will be applied to our specific VXX case in Section 3.4.

3.3.1 ETP on Futures Strategies

We analyze an ETP based on a strategy of N futures contracts $F_t(T)$. If we assume that the ETP strategy itself is not collateralized and it requires a proportional fee payment ϕ_t, we can write the strategy price process V_t as given by

$$\frac{dV_t}{V_t} = (r_t - \phi_t)dt + \sum_{i=1}^{N} \omega_t^i \frac{dF_t(T_i)}{F_t(T_i)}, \tag{3.3}$$

where r_t is the deterministic risk-free rate. Moreover, we assume that the futures contracts and the ETP are expressed in the same currency as the bank account based on r_t. The processes ω_t^i represent the investment percentage (or weights) in the futures contracts, and they are defined by the ETP term-sheet. Notice that the weights may depend on the whole vector of futures prices. It is straightforward to extend the analysis to more general ETP strategies based on total-return or dividend-paying indices. Moreover, we can introduce securities in foreign currencies.

We can consider as a first example the family of S&P GSCI commodity indices. ETC's tracking these indices are liquid in the market, and they represent an investment in a specific commodity or commodity class. The investment is in the first nearby one-month future, but for the last five days of the contract, where a roll over procedure is implemented to sell the first nearby and to buy the second one. A second example are the volatility-based ETN's VXX and VXZ. They represent respectively an investment in short-term or in the long-term structure of VIX futures. The investment is in a portfolio of futures mimicking a rolling one-month contract. In Section 3.4 we will focus on the VXX case.

3.3.2 Modeling the ETP Smile

We wish to jointly model the volatility smile of an ETP and of its underlying futures contracts. Indeed, market quotes for ETP plain-vanilla options can be found on the market. The main problem to solve is the fact that the price process of the ETP strategy is non-Markov, since it depends on the futures contract prices, namely the vector process $[V_t, F_t(T_1), \ldots, F_t(T_N)]$ is Markov.

We first start by modelling the marginal distribution of futures prices, then we proceed to describe the full model. The marginal distribution is modelled by following the procedure developed in Nastasi et al. (2020) and originally implemented for the commodity market. We consider that the futures prices $F_t^s(T_i)$ can be modelled in term of a common process s_t, which we can identify with the price of a rolling futures contract. Futures prices are defined in term of s_t as given by

$$F_t^s(T_i) = F_0(T_i)\left(1 - (1 - s_t)\, e^{-\int_t^{T_i} a(u)du}\right) \tag{3.4}$$

where $F_0(T_i)$ are the futures prices as observed today in the market, and a is a non-negative function of time. The value of the futures at maturity times t, which are not quoted in the market, can be obtained by interpolation, but they are never used in actual calculations. The driving risk process s_t is modelled by a mean-reverting local-volatility dynamics, namely

$$ds_t = a(t)\,(1 - s_t)\;dt + \eta\,(t, s_t)\,s_t\,dW_t^s, \quad s_0 = 1, \tag{3.5}$$

where η is a positive and bounded function of time and price, to be determined by the calibration of options of futures, and W_t^s is a standard Brownian motion under the risk-neutral measure. As a consequence, the futures prices follows a local-volatility dynamics given by

$$dF_t^s(T_i) = \eta_F\,(t, T_i, F_t^s(T_i))\;dW_t^s, \tag{3.6}$$

where we define

$$\eta_F\,(t, T_i, K) := \left(K - F_0\,(T_i)\left(1 - e^{-\int_t^{T_i} a(u)du}\right)\right)\eta\,(t, k_F\,(t, T_i, K))\,, \tag{3.7}$$

$$k_F\,(t, T_i, K) := 1 - \left(1 - \frac{K}{F_0\,(T_i)}\right)e^{\int_t^{T_i} a(u)du}. \tag{3.8}$$

Thus, plain-vanilla options on futures can be calculated as plain-vanilla options on the process s_t, which, in turn, can be easily computed by means of the Dupire equation. More precisely, given the dynamics in (3.5), we have that the normalized call price

$$c(t, k) := \mathbb{E}\left[(s_t - k)^+\right] \tag{3.9}$$

satisfies the parabolic PDE

$$\partial_t c(t, k) = \left(-a(t) - a(t)(1 - k)\partial_k + \frac{1}{2}k^2\eta^2(t, k)\partial_k^2\right)c(t, k), \tag{3.10}$$

with boundary conditions

$$c(t, 0) = 1, \quad c(t, \infty) = 0, \quad c(0, k) = (1 - k)^+. \tag{3.11}$$

A more detailed analysis and a proof of this (extended) Dupire equation can be found in Nastasi et al. (2020, Proposition 4.1). Given the prices of options on futures from the market, we can plug their corresponding normalized call prices in Equation (3.10), and then solve it for the function η. Then, we can compute the futures local volatility function $\eta_F(t, T_i, K)$, thanks to the mapping in Equation (3.7), and eventually recover the prices of options on futures.

There is a vast literature on how to solve Equation (3.10) with the boundary conditions (3.11), here we use an implicit PDE discretization method, as usually done for solving the Dupire equation, see for example Gatheral (2011). We allow a mean-reversion $a(t)$ possibly different from zero when we are calibrating the model to options on futures. On the other hand, options on the ETP spot price can be directly calibrated by a standard local-volatility model, that is without a mean-reversion term. Hence, once $a(t)$ is chosen, we can calibrate to the market $\eta_F(t, T_i, K)$ and $\eta_V(t, K)$, as previously described.

Thanks to the normalized spot dynamics (3.5), we can price all futures options by evaluating the PDE (3.10), and ETP options as well. In this way we implicitly obtain the marginal probability densities of futures prices and ETP prices.

We are now looking for the joint probability densities of futures prices, in order to be able to jointly model the dynamics of both the ETP strategy and its underlying futures. In order to do that, we need to specify a non trivial correlation structure among the futures. Notice that, for example, a one factor model based on the normalised spot dynamics, as in (3.5), would not be able to produce a non-trivial final correlation among futures, see Nastasi et al. (2020, Section 5.1). Hence, we have to enrich our model by adding new risk factors.

We start by introducing a generic stochastic volatility dynamics for futures prices under the risk-neutral measure, as given by

$$dF_t(T_i) = \nu_t\left(T_i\right) \cdot dW_t, \tag{3.12}$$

where $\nu_t\left(T_i\right)$ are vector processes, possibly depending on the futures price itself, and W_t is a vector of standard Brownian motions under the risk-neutral measure. The notation $a \cdot b$ refers to the inner product between vectors a and b.

In general, we can numerically solve together Equations (3.3) and (3.12) to calculate options prices on the ETP or on the futures contracts. Our aim is at calibrating both the futures and the ETP quoted smile. In the following, we always assume that interest rates and fees are deterministic functions of time, so that we write: $r_t := r(t)$ and $\phi_t := \phi(t)$. We can simplify the problem by noticing that we need only to know the marginal densities of the futures and ETP prices, so that we can focus our analysis on the Markovian projection of the dynamics (3.3) and (3.12). In this way, the model can match the marginal densities of futures prices and ETP prices predicted by the local volatility model.

Proposition 3.3.1. *Given the dynamics* (3.12) *and* (3.3)*, the Markov projections of the underlying processes F_t and V_t are given by*

$$d\tilde{F}_t\left(T_i\right) = \eta_F\left(t, T_i, \tilde{F}_t\left(T_i\right)\right) d\tilde{W}_t^i, \quad \eta_F\left(t, T_i, K\right) := \sqrt{\mathbb{E}\left[\left\|\nu_t\left(T_i\right)\right\|^2 \,\middle|\, F_t\left(T_i\right) = K\right]}, \tag{3.13}$$

and

$$\frac{d\tilde{V}_t}{\tilde{V}_t} = \left(r_t - \phi_t\right) dt + \eta_V(t, \tilde{V}_t) d\tilde{W}_t^0, \quad \eta_V(t, K) := \sqrt{\mathbb{E}\left[\sum_{i,j=1}^{N} \hat{\omega}_t^i \nu_t\left(T_i\right) \cdot \nu_t^T\left(T_j\right) \hat{\omega}_t^j \,\middle|\, V_t = K\right]}, \tag{3.14}$$

respectively, where $\tilde{W}_t^0$ and $\tilde{W}_t^i$, $i = 1, \ldots, N$ are scalar Brownian motions and where

$$\hat{\omega}_t^i \doteq \frac{\omega_t^i}{F_t(T_i)}, \quad \hat{\omega}_t^j \doteq \frac{\omega_t^j}{F_t(T_j)}. \tag{3.15}$$

The main ingredient for proving Proposition 3.3.1 is the Gyöngy Lemma, see Gyöngy (1986), that for the sake of clarity we recall here below.

Lemma 3.3.2 (Gyöngy). *Let $X(t)$ be given by*

$$X_t = X_0 + \int_0^t \alpha(s)\, ds + \int_0^t \beta(s)\, dW_s, \quad t \geq 0 \tag{3.16}$$

where W is a Brownian motion under some probability measure $\mathbb{P}$, where α, β are adapted bounded stochastic processes such that (3.16) admits a unique solution. If we define $a(t,x)$, $b(t,x)$ by

$$a(t,x) = \mathbb{E}\left[\alpha(t) \,|\, X(t) = x\right],$$
$$b^2(t,x) = \mathbb{E}\left[\beta^2(t) \,|\, X(t) = x\right],$$

then there exists a filtered probability space $(\widetilde{\Omega}, \tilde{\mathcal{F}}, \{\tilde{\mathcal{F}}\}_{t\geq 0}, \widetilde{\mathbb{P}})$ where $\tilde{W}$ is a $\widetilde{\mathbb{P}}-$Brownian motion, such that the SDE

$$Y_t = Y_0 + \int_0^t a(s, Y(s))\, ds + \int_0^t b(s, Y(s))\, d\tilde{W}_s, \quad t \geq 0$$

admits a weak solution that has the same one-dimensional distributions as X. The process Y is called the Markovian projection of the process X.

Proof. (Proposition 3.3.1) Equation (3.13) follows immediately from equation (3.12) and from Lemma 3.3.2. Dynamics (3.3) can be rewritten as

$$\frac{dV_t}{V_t} = (r_t - \phi_t)dt + \sum_{i=1}^{N} \omega_t^i \, \frac{\nu_t\left(T_i\right) \cdot dW_t}{F_t(T_i)}, \tag{3.17}$$

and applying Lemma 3.3.2 we get

$$\eta_V(t,K)^2 = \mathbb{E}\left[\sum_{i,j=1}^{N} \frac{\omega_t^i \nu_t\left(T_i\right)}{F_t(T_i)} \cdot \frac{\nu_t^T\left(T_j\right)\omega_t^j}{F_t(T_j)} \,\middle|\, V_t = K\right], \tag{3.18}$$

while the drift part remains unaffected since both r_t and ϕ_t are deterministic functions. The conclusion follows immediately. □

The functions $\eta_F(t,T,K)$ and $\eta_V(t,K)$ are respectively the futures and ETP local volatility functions (we do not show the dependency of $\eta_V(t,K)$ on all the futures maturity to lighten the notation). As already discussed previously, we can calibrate the model on plain-vanilla market quotes using the approach of Nastasi et al. (2020), that is $\eta_F(t,T,K)$ will be calibrated using options on VIX futures and $\eta_V(t,K)$ using options on the ETP strategy.

So far, we have preserved the marginal densities of futures prices and ETP prices, by applying the Gyöngy lemma directly on the futures and ETP dynamics, as shown in Proposition (3.3.1). Thus, we can calibrate our model to match the marginal densities of futures and ETP prices predicted by the local volatility model, and in turn to match the plain vanilla option-on-futures prices and on ETP prices quoted on the market, by requiring that the process ν satisfies (3.13) and (3.14). We recall again that the local volatility for futures prices is defined in Equation (3.7), while the ETP

local volatility can be recovered by using a standard local volatility model with mean-reversion $a = 0$.

Hence, we have to ensure the two constraints given by the Markov projection (3.13) and (3.14) by properly choosing the vector processes $\nu_t(T_i)$. More precisely, we need to define a suitable dynamics for the process $\nu_t(T_i)$ satisfying Equation (3.13) and (3.14), which in turn will enable us to preserve the calibration to plain vanilla options. In the following, we will focus on a specific form, allowing to perform the calibration of the futures local volatility independently of the ETP dynamics. A simple way to do that is selecting the vector process $\nu_t(T_i)$ as

$$\nu_t\left(T_i\right) \doteq \ell_F\left(t, T_i, F_t\left(T_i\right)\right) \sqrt{v_t} R\left(t, T_i, V_t\right), \tag{3.19}$$

where v_t is a scalar process independent of the futures and ETP dynamics. Here, $\ell_F(t, T_i, K)$ is the leverage function referred to the maturity T_i, that can be computed as

$$\ell_F(t, T_i, K) = \frac{\eta_F(t, T_i, K)}{\sqrt{\mathbb{E}\left[v_t \,\middle|\, F_t = K\right]}}, \tag{3.20}$$

where, again, $\eta_F(t, T_i, K)$ is the futures local volatility function. Here, the vector R plays the role of local correlation and satisfies the condition $\|R\| = 1$. It is the main ingredient to model the dependency between different futures and it enables us to correlate them in a natural way. In fact, Equation (3.19) satisfies by construction Equation (3.13), so that the model reprices plain-vanilla options on futures correctly. Notice that, as in standard stochastic-local volatility models, the above equation is not a definition since $\nu_t(T_i)$ appears implicitly also on the right-hand side within the conditional law of the expectation. We will investigate numerically in practical cases the validity of this assumption. In Section 3.4 we shall introduce a simple specification for R that will be enough for our purposes. Then, we shall analyze the constraint on the ETP dynamics by substituting Equation (3.19) into Equation (3.14). In other words, we are looking for conditions on R under which the model is able to reprice the ETP plain vanilla options correctly. In this way R becomes a (vector) function of the ETP price. Straightforward computations yield

$$\eta_V^2(t, K) = \sum_{i,j=1}^{N} R\left(t, T_i, K\right) \cdot R\left(t, T_j, K\right) \mathbb{E}\left[v_t \hat{\omega}_t^i \hat{\omega}_t^j \ell_F\left(t, T_i, F_t\left(T_i\right)\right) \ell_F\left(t, T_j, F_t\left(T_j\right)\right) \,\middle|\, V_t = K \right]. \tag{3.21}$$

In conclusion, to complete the calibration to ETP plain vanilla prices, we have to solve the above equation for an (admissible) unknown local function $R(t, T_i, K)$.

3.4 The VXX ETN Strategy

We consider the specific case of the VIX futures and the VXX ETN strategy.

3.4.1 Definition of the ETN Strategy

The VXX is defined as a strategy on the first and second nearby VIX futures, see Gehricke and Zhang (2018), with the following weights:

$$\omega_t^1 \doteq \frac{\alpha_1(t)F_t\left(T_1\right)}{\alpha_1(t)F_t\left(T_1\right)+\alpha_2(t)F_t\left(T_2\right)}, \quad \omega_t^2 \doteq \frac{\alpha_2(t)F_t\left(T_2\right)}{\alpha_1(t)F_t\left(T_1\right)+\alpha_2(t)F_t\left(T_2\right)}, \tag{3.22}$$

where we define

$$\alpha_1\left(t,T_0,T_1\right) := \frac{\varsigma\left(T_1,-1\right)-\varsigma(t,1)}{\varsigma\left(T_1,-1\right)-T_0}, \quad \alpha_2\left(t,T_0,T_1\right) := 1-\alpha_1\left(t,T_0,T_1\right), \tag{3.23}$$

with T_0 being the settlement date of the futures contract just expired (the one preceding the first nearby), T_1 the settlement date of the first nearby contract, and $\varsigma(t,d)$ a calendar date shift of d business days. In this formulae we are considering $t < T_1$, otherwise we have to consider the following pair of future contracts.

Remark 3.4.1. *Grasselli and Wagalath (2018) adopted a slightly different definition, using the year fraction between the second and the first nearby futures as reference time interval, instead of the year fraction between the first nearby and the previous futures contract (without calendar adjustments), leading to*

$$\alpha_1(t,T_1,T_2) := \frac{T_1-t}{T_2-T_1}. \tag{3.24}$$

3.4.2 The Local Correlation

We now implement the modelling framework of the previous sections, by exploiting the structure of the local correlation vector R. A simple way to do that consists in using the following parametrization:

$$R\left(t,T_1,K\right) := [1,0], \quad R\left(t,T_2,K\right) := [\rho(t,K), \sqrt{1-\rho(t,K)^2}], \tag{3.25}$$

which satisfies $\|R\| = 1$ by construction. That is, we express the dependence in terms of a local correlation coefficient $\rho(t,K)$ that has to be determined by solving Equation (3.21).

Remark 3.4.2. *The choice in (3.25) relates to the case $N = 2$, namely two futures, but it can be easily extended to an arbitrary number N. In that case, the local correlation vector $R(t,T_i,K)$ can be found by performing the Cholesky decomposition to the corresponding local correlation matrix.*

Proposition 3.4.3. *Given the parametrization in (3.25), the local correlation coefficient $\rho(t,K)$ is given by*

$$\rho(t,K) = \frac{\eta_V^2(t,K)-A_1(t,K)-A_2(t,K)}{2A_{12}(t,K)}, \tag{3.26}$$

provided that t is not on a futures maturity date, where

$$
\begin{aligned}
A_1(t,K) &\doteq \mathbb{E}\left[(\hat{\omega}_t^1)^2 \ell_F(t,T_1,F_t(T_1))^2 v_t \,\middle|\, V_t = K\right], \\
A_2(t,K) &\doteq \mathbb{E}\left[(\hat{\omega}_t^2)^2 \ell_F(t,T_2,F_t(T_2))^2 v_t \,\middle|\, V_t = K\right], \\
A_{12}(t,K) &\doteq \mathbb{E}\left[\hat{\omega}_t^1 \hat{\omega}_t^2 \ell_F(t,T_1,F_t(T_1))\ell_F(t,T_2,F_t(T_2)) v_t \,\middle|\, V_t = K\right],
\end{aligned}
$$

and where

$$
\hat{\omega}_t^1 = \frac{\omega_t^1}{F_t(T_1)} = \frac{\alpha_1(t)}{\alpha_1(t)F_t(T_1) + \alpha_2(t)F_t(T_2)}, \quad \hat{\omega}_t^2 = \frac{\omega_t^2}{F_t(T_2)} = \frac{\alpha_2(t)}{\alpha_1(t)F_t(T_1) + \alpha_2(t)F_t(T_2)}. \tag{3.27}
$$

We recall, once again, that the leverage function $\ell_F(t,T_i,K),\ i=1,2$ is given by Equation (3.20). Notice that when t corresponds to a futures maturity date, Equation (3.26) becomes meaningless, because the weights in (3.27) vanish.

Proof. Exploiting the local volatility $\eta_V(t,K)$ in Equations (3.14) and (3.21), we get

$$
\begin{aligned}
\eta_V(t,K)^2 &= \mathbb{E}\left[(\hat{\omega}_t^1)^2 \left\|\nu_t(T_1)\right\|^2 + (\hat{\omega}_t^2)^2 \left\|\nu_t(T_2)\right\|^2 + 2\hat{\omega}_t^1\hat{\omega}_t^2 \nu_t(T_1)\cdot\nu_t^T(T_2) \,\middle|\, V_t = K\right] \\
&= \mathbb{E}\left[(\hat{\omega}_t^1)^2 \ell_F(t,T_1,F_t(T_1))^2 v_t + (\hat{\omega}_t^2)^2 \ell_F(t,T_2,F_t(T_2))^2 v_t \,\middle|\, V_t = K\right] + \\
&+ 2\mathbb{E}\left[\hat{\omega}_t^1\hat{\omega}_t^2 \ell_F(t,T_1,F_t(T_1))\sqrt{v_t}\ell_F(t,T_2,F_t(T_2))\sqrt{v_t}\rho(t,K) \,\middle|\, V_t = K\right] \\
&= \mathbb{E}\left[(\hat{\omega}_t^1)^2 \ell_F(t,T_1,F_t(T_1))^2 v_t \,\middle|\, V_t = K\right] + \mathbb{E}\left[(\hat{\omega}_t^2)^2 \ell_F(t,T_2,F_t(T_2))^2 v_t \,\middle|\, V_t = K\right] + \\
&+ 2\rho(t,K)\mathbb{E}\left[\hat{\omega}_t^1\hat{\omega}_t^2 \ell_F(t,T_1,F_t(T_1))\ell_F(t,T_2,F_t(T_2)) v_t \,\middle|\, V_t = K\right].
\end{aligned} \tag{3.28}
$$

Therefore, if we define

$$
\begin{aligned}
A_1(t,K) &\doteq \mathbb{E}\left[(\hat{\omega}_t^1)^2 \ell_F(t,T_1,F_t(T_1))^2 v_t \,\middle|\, V_t = K\right], \\
A_2(t,K) &\doteq \mathbb{E}\left[(\hat{\omega}_t^2)^2 \ell_F(t,T_2,F_t(T_2))^2 v_t \,\middle|\, V_t = K\right], \\
A_{12}(t,K) &\doteq \mathbb{E}\left[\hat{\omega}_t^1 \hat{\omega}_t^2 \ell_F(t,T_1,F_t(T_1))\ell_F(t,T_2,F_t(T_2)) v_t \,\middle|\, V_t = K\right],
\end{aligned}
$$

then, when t is not on a futures maturity date, we can obtain $\rho(t,K)$ from Equation (3.28) and get the result. □

We can estimate the conditional expectations by means of the techniques developed in Guyon and Henry-Labordère (2012), namely approximating conditional expectations using smooth integration kernels. Thus, we can evaluate the expectation of a process X_t given a second process Y_t as

$$\mathbb{E}\left[X_t \mid Y_t = K\right] \approx \frac{\mathbb{E}\left[X_t \delta^{\varepsilon}\left(Y_t - K\right)\right]}{\mathbb{E}\left[\delta^{\varepsilon}\left(Y_t - K\right)\right]}, \tag{3.29}$$

where δ^{ε} is a suitably defined mollifier of the Dirac delta depending on a smoothing coefficient ε. We can use such approximation within the Monte Carlo simulation for $[V_t, F_t(T_1), \ldots, F_t(T_N)]$ to evaluate the diffusion coefficients of the processes. Alternatively we could use the collocation method developed in Van der Stoep et al. (2014). We stress again that the existence of a solution for equation (3.26) with the above coefficients is an open question, and, moreover, even if the solution exists, we have to check against market quotes if the chosen parameterization is able to produce a local correlation $\rho(t, K) \in [-1, 1]$ given the plain-vanilla option prices on futures contract and ETP strategy quoted in the market.

Remark 3.4.4. *If we assume that $v_t = 1$, namely if we remove the stochastic volatility component, then the formulae simplify and we obtain that the options on futures can be calibrated without any further calculations, since we get*

$$\nu_t\left(T_i\right) \doteq \ell_F\left(t, T_i, F_t\left(T_i\right)\right) R\left(t, T_i, V_t\right), \tag{3.30}$$

while the local coefficients entering the calculation of the local correlation become

$$\begin{aligned}
A_1(t,K) &\doteq \mathbb{E}\left[(\hat{\omega}_t^1)^2 \eta_F(t, T_1, F_t(T_1))^2 \,\middle|\, V_t = K\right], \\
A_2(t,K) &\doteq \mathbb{E}\left[(\hat{\omega}_t^2)^2 \eta_F(t, T_2, F_t(T_2))^2 \,\middle|\, V_t = K\right], \\
A_{12}(t,K) &\doteq \mathbb{E}\left[\hat{\omega}_t^1 \hat{\omega}_t^2 \eta_F(t, T_1, F_t(T_1)) \eta_F(t, T_2, F_t(T_2)) \,\middle|\, V_t = K\right].
\end{aligned}$$

Indeed, in this particular case, the leverage function $\ell_F(t, T_i, K)$ reduces to the local volatility function $\eta_F(t, T_i, K)$.

Remark 3.4.5 (VXZ Strategy). *The above reasoning can be extended to the VXZ ETN strategy where four future contracts in the long part of the futures term structure are considered, in a similar way as the VXX is considering the short term, although this time we need to parametrize a larger local correlation matrix.*

3.4.3 Model Specification

We start from the stochastic-local volatility dynamics for the futures defined in Section 3.3, namely

$$dF_t(T_i) = \ell_F(t, T_i, F_t(T_i)) \sqrt{v_t} \cdot dW_t, \tag{3.31}$$

where the stochastic volatility component v_t follows a simple one-dimensional CIR process, i.e.

$$dv_t = \kappa\,(\theta - v_t)\,dt + \xi\sqrt{v_t}\,dZ_t, \tag{3.32}$$

with initial condition $v(0) = v_0$ and where κ is the speed of the mean-reversion of the variance, θ is the long-run average variance, ξ is the volatility-of-volatility parameter and ρ^v is the correlation between W and Z. For the sake of simplicity, we assume the same Brownian motion Z_t for the volatility of each futures $F_t(T_i)$, $i = 1, \ldots, N$, so we basically assume that all futures share the (common stochastic volatility) CIR factor v_t.

We describe the (static) correlation structure across futures dynamics through a positive semi-definite matrix $\Sigma_W \in \mathbb{R}^{N\times N}$, where N refers to the number of futures. Therefore, the correlation structure for the tuple of all the Brownian factors (W_t, Z_t) takes the following form:

$$\Sigma_{W,Z} = \begin{pmatrix} \Sigma_W & \rho_{WZ}^T \\ \rho_{WZ} & 1 \end{pmatrix} \in \mathbb{R}^{(N+1)\times(N+1)}, \quad \Sigma_W \in \mathbb{R}^{N\times N}, \quad \rho_{WZ} \in \mathbb{R}^{1\times N}, \tag{3.33}$$

where $\rho_{WZ} = (\rho^v, \ldots, \rho^v)$ is the correlation between the Brownian motions W driving the futures and the noise Z driving the (common) stochastic volatility component v, so that the components of the vector ρ_{WZ} are all equal to the constant ρ^v.

he parameter ρ^v should be chosen in order to ensure the positive semi-definiteness of the correlation matrix $\Sigma_{W,Z}$. For example, in the case $N = 2$ and W_1 independent of W_2, it turns out that the parameter ρ^v needs to belong to the interval $\left[-1/\sqrt{2}, 1/\sqrt{2}\right]$.

Remark 3.4.6. *Our one factor specification for the stochastic volatility turns out to be enough for our purposes, though, of course, it is not difficult to generalise our model to a multifactor stochastic volatility framework. What is more, this apparently more general choice does not lead to better results in the quality of the fit, up to our numerical experience.*

3.5 Numerical Illustration on a Real Data Set

In this section we perform a numerical exercise, based on real data quoted by the market for options on the VIX futures and the VXX ETN strategy, where we provide a set of admissible parameters, that is leading to an admissible local correlation vector in (3.26) satisfying $\rho(t, K) \in [-1, 1]$. In this way we build up a coherent framework that correctly reprices plain vanilla options on futures contract and VXX strategy quoted in the market. We underline that the set of parameters we are going to find does not necessarily represent the optimal choice, namely we are just providing an admissible framework, not a full calibration. We leave for future research the complete exploration of this issue.

3.5.1 VIX and VXX Data Set

We test our model against a set of three maturites for VXX call options with a corresponding set of four maturities for VIX futures call options. All data were downloaded by Bloomberg on

November 7, 2019. Table 3.1 (resp. Table 3.2) describes the features of the VIX (resp. VXX) options data set.

VIX futures call options		
Maturities	Number of options	Strikes range (min − max)
T_1: November 20, 2019	40	$10 - 80$
T_2: December 18, 2019	40	$10 - 80$
T_3: January 22, 2020	35	$10 - 80$
T_4: February 19, 2020	35	$10 - 80$

Table 3.1: VIX futures call options data set as on T_0: November 7, 2019.

VXX call options		
Maturities	Number of options	Strikes range (min − max)
T_1: November 20, 2019	65	$8 - 60$
T_2: December 18, 2019	59	$2 - 60$
T_3: January 22, 2020	53	$10 - 90$

Table 3.2: VXX call options data set as on T_0: November 7, 2019.

The different maturities of VIX and VXX options as on November 7, 2019 involved in the numerical exercise are described in Figure 3.2.

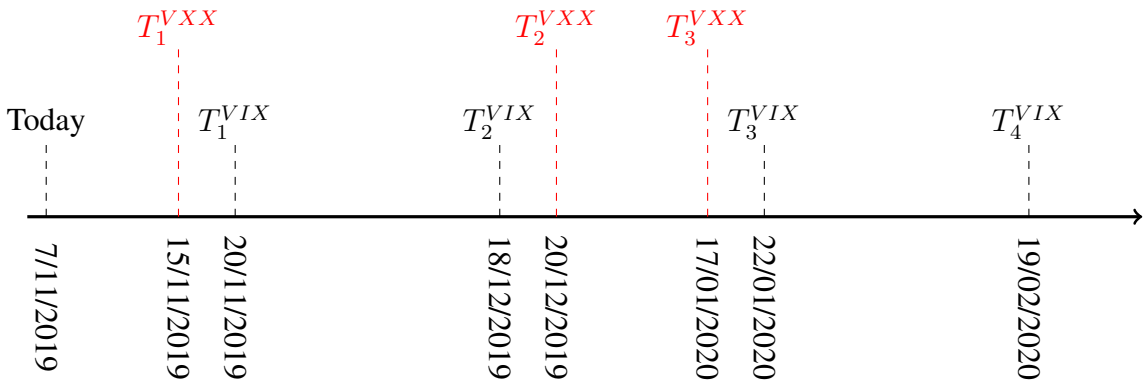


Figure 3.2: Maturities of options on VIX futures and VXX as on November 7, 2019 in our data set. $T_i^{VIX}, i = 1, \ldots, 4$ (resp. $T_j^{VXX}, j = 1, \ldots, 3$) refer to VIX futures (resp. VXX) options maturities.

The values of VIX futures and VXX spot price, as observed in the market, are shown in Table 3.3, while we show in Figure 3.3 (resp. Figure 3.4) the bid, mid and ask market quotes for call options written on VIX (resp. VXX futures), for the four (resp. three) maturities of the data set.

Futures and Spot prices	Market Value	Maturities
$F_0(T_1)$	14.60	T_1: November 20, 2019
$F_0(T_2)$	16.15	T_2: December 18, 2019
$F_0(T_3)$	17.45	T_3: January 22, 2020
$F_0(T_4)$	18.15	T_4: February 19, 2020
Spot V_0	19.22	T_0: November 7, 2019

Table 3.3: Market values of VIX futures $F_0(T_1), \ldots, F_0(T_4)$ and of the VXX spot V_0 as on T_0 (November 7, 2019).

Figure 3.3: Bid, mid and ask market quotes for VIX futures call options, as on November 7, 2019, for the four maturities in the data set: November 20, 2019, December 18, 2019, January 22, 2020 and February 19, 2020, from the top left corner going clockwise.

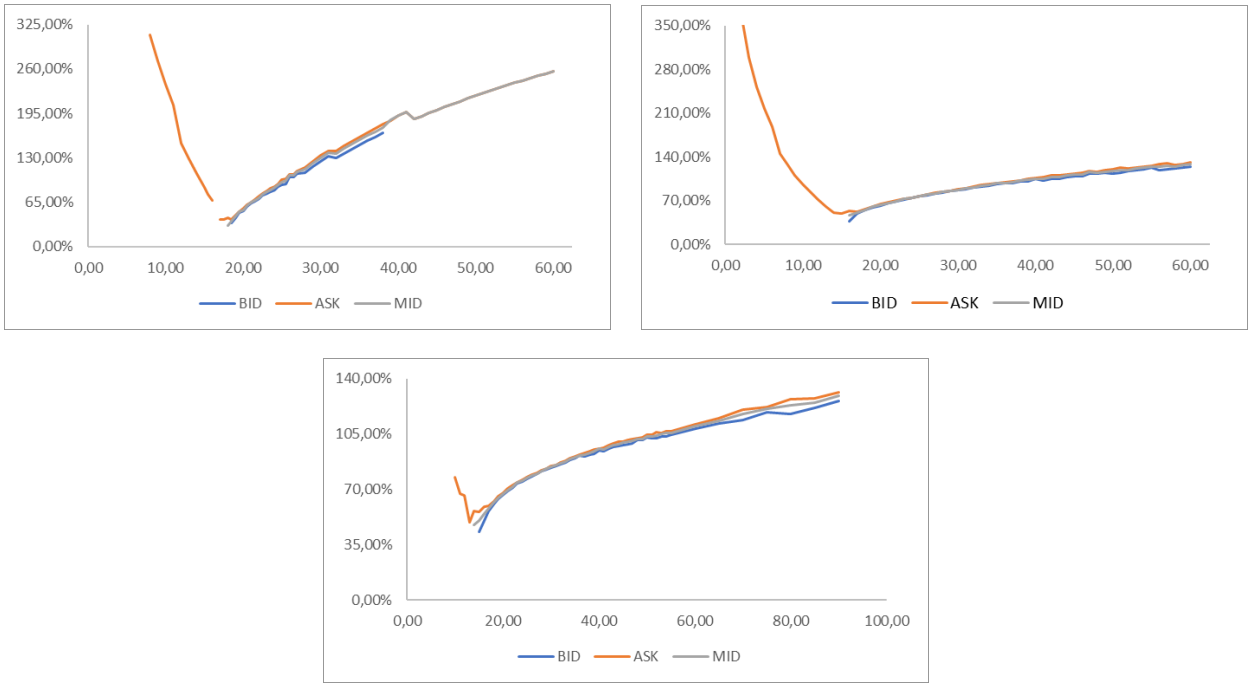


Figure 3.4: Bid, mid and ask market quotes for VXX call options, as on November 7, 2019, for the three maturities in the data set: November 15, 2019, December 20, 2019 and January 17, 2020, from the top to the bottom.

An additional difficulty comes from the peculiar nature of the VIX-VXX options market data. In general, one can calibrate on market quotes or choose to implement a smoothing technique, like the widely-used SVI parametrization of the implied volatility smile, see Gatheral and Jacquier (2014) and the references therein. With this method, we are guaranteed that the interpolation is arbitrage free. We applied the SVI methodology to some market quotes of VIX and VXX Call and Put options, as on May 24, 2019. Figure 3.5 shows that VIX market tends to be more irregular for shorter maturities, where a SVI application seems to be more needed, while it tends to be more regular as the maturities go further. On the other hand, the VXX option market seems to be more balanced overall, as Figure 3.6 suggests, so the SVI smoothing is not really necessary.

3.5.2 Parameters Sensitivity

In this subsection, we perform a sensitivity analysis that will highlight the impact of the model parameters on the VXX smiles. The aim of this analysis is to find a range of parameters that guarantees a good fit to the VXX market smile, while maintaining the fit on the VIX market smile. Then, we test if these parameters generate a local correlation $\rho(t, K) \in [-1, 1]$, by solving Equation (3.26). More precisely, we let each parameter vary while keeping the others freezed. In this way, we can check the overall impact of each parameter on the shape of the VXX smile. This will give us an insight on the set of reasonable parameters that can generate a good fit, provided that the local correlation in (3.26) is admissible.

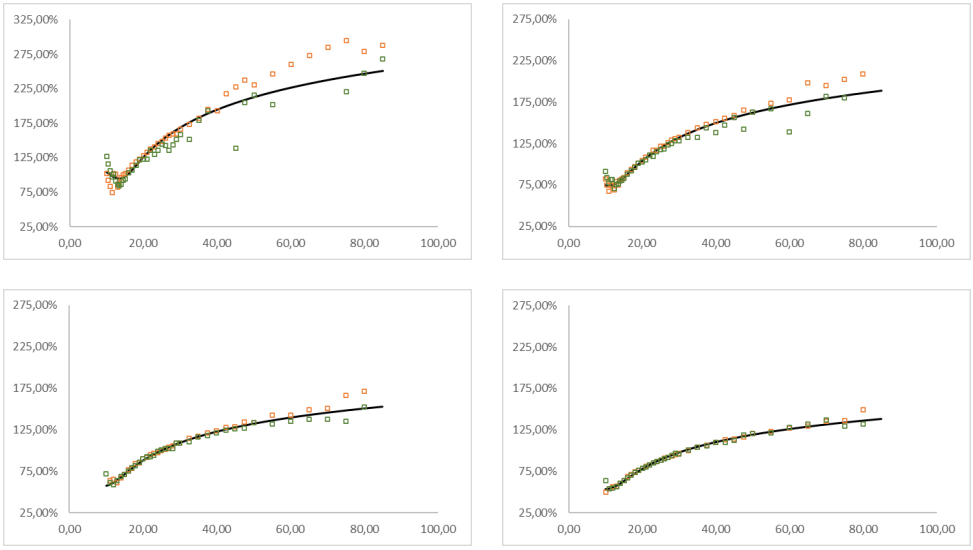


Figure 3.5: Application of SVI methodology to VIX options: the continuous black line represents the volatility smile using the SVI formulae. Orange squares represent Call market quotes, while the green squares represent Put market quotes. From the top left corner, as on May 24, 2019, pictures refer to maturities June 19, 2019, July 17, 2019, August 21, 2019 and September 18, 2019, respectively.

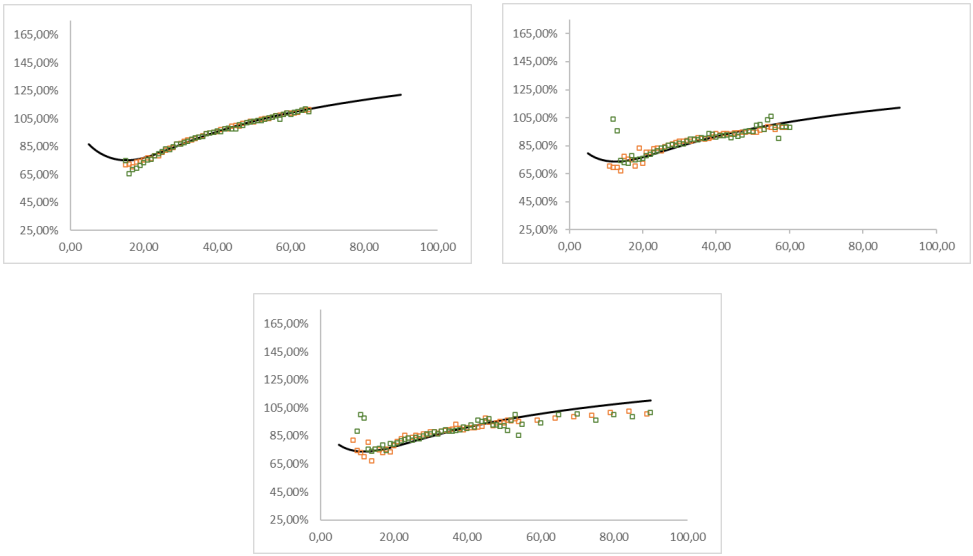


Figure 3.6: Application of SVI methodology to VXX options: the continuous black line represents the volatility smile using the SVI formulae. Orange squares represent Call market quotes, while the green squares represent Put market quotes. From the top left corner, as on May 24, 2019, pictures refer to maturities September 20, 2019, December 20, 2019 and January 17, 2020, respectively.

We shall investigate the impact of the mean-reversion speed a, the correlation between futures, ρ_{W_i,W_j}, $i,j = 1,\dots,4$, the volatility-of-volatility ξ and the correlation ρ^v between futures and their (common) volatility factor v. Concerning the stochastic volatility component v, we will provide an admissible choice for the parameters κ, θ and v_0 (that is, a specification such that the Feller condition holds and the local correlation $\rho(t,K)$ always remains in $[-1,1]$) without performing a sensitivity analysis, as it turns out that they do not play an important role in the smile modeling. As we shall see, the impact of the mean-reversion speed a plays a crucial role in order to recover the ATM model implied volatility from the VXX options market. The impact of the correlations between futures ρ_{W_1,W_2} is also significant, while the other parameters do not really affect the shape of the VXX smile.

We start by analysing the effect of the mean-reversion a. In Figure 3.7, we present the impact of the mean reversion parameter a on the VXX smile at the first maturity, namely November 15, 2019, according to Figure 3.2. The results are representative also for the other maturities. In order to do that, we fixed the volatility-of-volatility $\xi = 1.1$, the correlation between futures $\rho_{W_1,W_2} = 0.85$ and $\rho^v = 0.75$. We then let the mean-reversion speed take the values $a = 0,2,4,6,8,9$. Figure 3.7 shows the market and the model smiles with the different choices of the mean-reversion a. We get that the higher the mean-reversion, the lower the smile. Thus, in order reproduce the market ATM implied volatility level for the VXX smile, we need a high mean-reversion speed, say around $a = 8$, according to Figure 3.7.

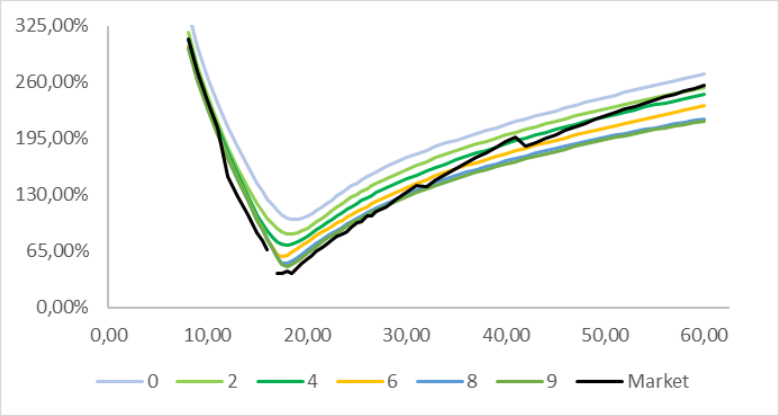


Figure 3.7: Impact of the mean-reversion speed $a = 0,2,4,6,8,9$ in the shape of the VXX smile. The other parameters are settled as follows: $\xi = 1.1$, $\rho_{W_1,W_2} = 0.85$ and $\rho^v = 0.75$.

The (high) value for the mean-reversion speed could seem misconceiving. However, we should keep in mind that we are jointly modeling two markets, the VIX and VXX option books, that display different peculiarities, therefore a somehow non standard level for the mean reversion seems a reasonable price to pay in order to reach the goal.

The impact of the correlation between futures ρ_{W_1,W_2} seems to play a pretty important role too in the shape of the VXX smile. To see this, we fix $\rho^v = 0.75$, $\xi = 1.1$, $a = 8$ and let ρ_{W_1,W_2} varying according to $-0.5, 0, 0.25, 0.55, 0.85$ and 1. In Figure 3.8 we see that the impact of the correlation between futures tends to affect only the right tail of the smile, while leaving unaffected

the left one.

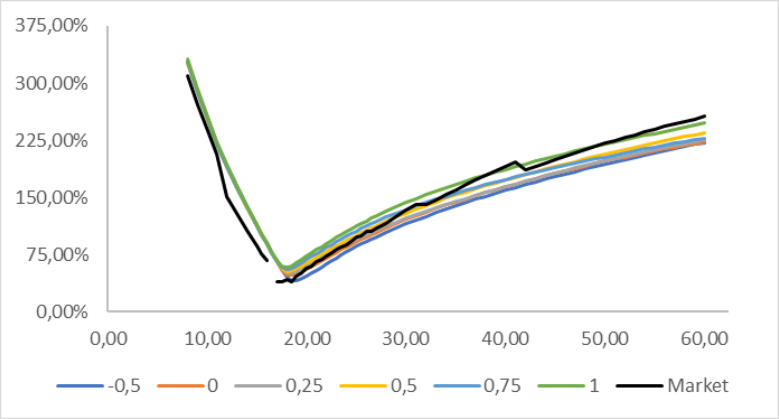


Figure 3.8: Impact of the correlation between futures $\rho_{W_1,W_2} = -0.5, 0, 0.25, 0.55, 0.85, 1$ on the shape of the VXX smile. The other parameters are settled as follows: $\rho^v = 0.75$, $\xi = 1.1$ and $a = 8$.

Next we consider the impact of the volatility-of-volatility ξ and the correlation between futures and their volatilities, namely ρ^v. As we are going to see, these two parameters, especially ρ^v, seem to play a less pivotal role than the two ones already considered. We start with the volatility-of-volatility and as shown before, we keep the mean-reversion $a = 8$ fixed, $\rho_{W_1,W_2} = 0.85$, $\rho^v = 0.75$ and let ξ varying within 0.1 and 1.6, with a step of 0.5. Figure 3.9 shows that the (positive) impact of the volatility-of-volatility is pretty significant only on the tails of the VXX smile (especially on the left ones).

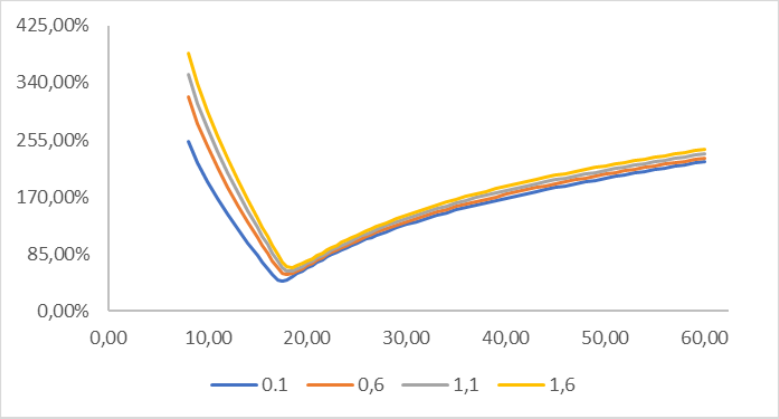


Figure 3.9: Impact of the volatility-of-volatility $\xi = 0.1, 0.6, 1.1, 1.6$ in the shape of the VXX smile. The other parameters are settled as follows: $\rho_{W_1,W_2} = 0.85$, $\rho^v = 0.75$ and $a = 8$.

Lastly, we check the impact of the correlation between futures and their volatilities, namely ρ^v. We fix $\rho_{W_1,W_2} = 0.85$, $\xi = 1.1$ and $a = 8$ and let $\rho_{F,Z}$ varying with $0.25, 0.5$ and 0.75. Figure 3.10

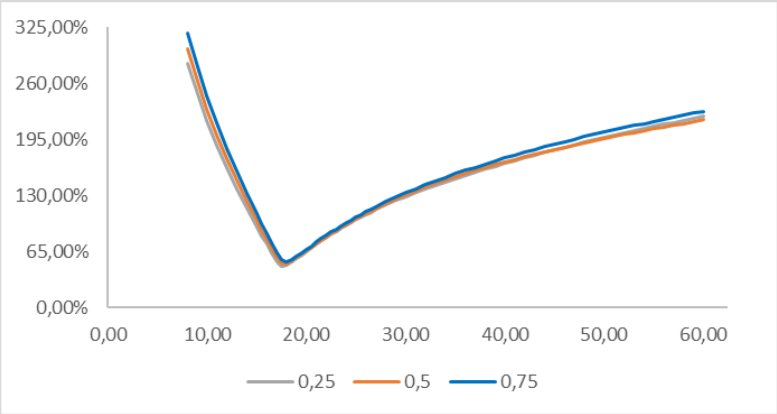


Figure 3.10: Impact of the correlation between futures and their volatilities $\rho^v = 0.25, 0.5, 0.75$ on the shape of the VXX smile. The other parameters are settled as follows: $\rho_{W_1,W_2} = 0.85$, $\xi = 1.1$ and $a = 8$.

suggests that the correlations between futures and their volatilities, $\rho_{F_i,Z}$ have a very limited impact on the shape of the VXX smile.

For the sensitivity analysis with respect to the mean-reversion, the correlation between futures, the volatility of volatility and the correlation between futures and their volatilities, we get results that are qualitatively similar as the ones presented before, even if the parameters are made varying in a different manner, namely the general behaviour of the former quantities evolves according to Figures 3.7, 3.8, 3.9 and 3.10, respectively, up to our data set.

In the next subsection, we are going to build up a consistent framework in order to fix an admissible choice for all parameters that allows us to reproduce the market smiles of both VIX and VXX.

3.5.3 Joint Fit of VIX and VXX Market Smiles

Thanks to the analysis performed in Subsection 3.5.2, we are now able to pick a cocktail of parameters that enables us to get a good fit to both the VIX and VXX market smiles and that leads to an admissible correlation parameter $\rho(t, K) \in [-1, 1]$. That is, the model gives the proper correlation we need to put in the VIX and VXX dynamics, in order to reprice correctly the corresponding options, as described in Section 3.4.

Let us summarize the parameters that we have fixed in order to perform the joint fit.

- The mean-reversion speed used in the joint fit is $a = 8$ (this value revealed to be the optimal trade-off in order to recover the implied volatility ATM level for the VXX options, while keeping a good fit for the VIX market).
- The volatility-of-volatility parameter is fixed as $\xi = 1.1$.
- The (static) correlation structure between the Brownian motions driving the VIX futures and the parameters describing the leverages (i.e. the correlation between futures and their

common stochastic volatility CIR process) are given by the following correlation matrix between futures their volatilities: the positions $(i,j),\ i=1,\dots,4,\ j=1,\dots,4,\ i<j$ represent the correlation between futures, ρ_{W_i,W_j}, while the positions $(i,5),\ i=1,\dots,4$ represent the correlation between futures and their volatilities, ρ^v.

$$\Sigma_{W,Z} = \begin{array}{c} \begin{matrix} W_1 & W_2 & W_3 & W_4 & Z \end{matrix} \\ \begin{pmatrix} 1 & 0.85 & 0.85 & 0.85 & 0.75 \\ * & 1 & 0.85 & 0.85 & 0.75 \\ * & * & 1 & 0.85 & 0.75 \\ * & * & * & 1 & 0.75 \\ * & * & * & * & 1 \end{pmatrix} \end{array} \begin{matrix} W_1 \\ W_2 \\ W_3 \\ W_4 \\ Z \end{matrix}$$

- The remaining parameters used in the CIR-like volatility process are chosen by taking into account the non-violation of the Feller condition. The complete parameter list, where we repeat the entry already set for sake of clarity, is given by:

$$\kappa = 2.5\,,\ \theta = 2.5\,,\ \xi = 1.1\,,\ v_0 = 1\,,\ \rho^v = 0.75$$

We have now all the ingredients to perform the joint calibration of the implied volatilities of call options on VIX and VXX for different strikes and maturities around the ATM via the SLV method described previously which employs a Monte Carlo simulation with $5\cdot 10^5$ paths. In Table 3.4 (resp. Table 3.5) we show the model implied volatilities on VIX (resp. VXX) around ATM (according to Table 3.3), that are generated by the SLV model, and the market ones, with the relative errors.

Maturities	Strike	ASK	BID	MID	Model	Relative error
November 20, 2019	$K=14$	1.0911	0.8042	0.9482	1.0739	0.01576
	$K=14.5$	1.0936	0.8650	0.9792	1.0911	0.00229
	$K=15$	1.1567	0.9297	1.0433	1.1392	0.01513
December 18, 2019	$K=15$	0.8779	0.7519	0.8151	0.8642	0.01560
	$K=16$	0.8995	0.7819	0.8406	0.9030	0.00389
	$K=17$	0.8406	0.8354	0.8702	0.9170	0.09088
January 22, 2020	$K=16$	0.7297	0.6255	0.6778	0.7263	0.00466
	$K=17$	0.7619	0.6644	0.7131	0.7540	0.01036
	$K=18$	0.7861	0.6911	0.7386	0.7839	0.00279
February 19, 2020	$K=17$	0.6992	0.6020	0.6507	0.6896	0.01372
	$K=18$	0.7157	0.6361	0.6759	0.7103	0.00754
	$K=19$	0.7425	0.6513	0.6969	0.7360	0.00875

Table 3.4: Implied volatility comparison between market quotes as on November 7, 2019 and the implied volatility obtained via our SLV models for VIX call options, for the four maturities and around ATM, according to Table 3.3.

Maturities	Strike	BID	ASK	MID	Model	Relative error
November 15, 2019	$K = 19$	0.4254	0.4602	0.4429	0.4602	~ 0
	$K = 19.5$	0.5027	0.5111	0.5070	0.5030	0.01584
	$K = 20$	0.5212	0.5633	0.5394	0.5473	0.02840
December 20, 2019	$K = 19$	0.6002	0.6078	0.6040	0.6050	0.00461
	$K = 19.5$	-	-	-	0.6154	-
	$K = 20$	0.6267	0.6455	0.6361	0.6329	0.01951
January 17, 2020	$K = 19$	0.6378	0.6528	0.6453	0.6450	0.01194
	$K = 19.5$	-	-	-	0.6514	-
	$K = 20$	0.6649	0.6796	0.6723	0.6735	0.00838

Table 3.5: Implied volatility comparison between market quotes as on November 7, 2019 and the implied volatility obtained via our SLV models for VXX call options, for the three maturities and around ATM, according to Table 3.3.

In Figure 3.11 (resp. Figure 3.12) we show the whole implied volatilities for all strikes obtained by our SLV model and the market volatilities of VIX futures (resp. VXX) call options. Some remarks are in order: first, we calibrated on ask option quotes as they were more regular and complete, namely we considered real quotes instead of e.g. SVI interpolations. In fact, ask quotes are much more informative on the VXX market, according to Figure 3.4. Despite the irregular shape of the market VXX smiles, the fit is overall good and even remarkably good for the VIX smiles. Recall that this market is quite irregular, with large bid-ask spreads, according to Figure 3.3. This means that our SLV model reproduces VIX and VXX smiles that fall within the market bid-ask spreads even after a perturbation of (some of) the parameters, according to the sensitivity analysis performed in the previous subsection. What is more, our parameters specification leads to an admissible local correlation vector $\rho(t, K)$ that falls in the interval $[-1, 1]$, as required in Proposition 3.4.3.

All the numerical computations and the Monte Carlo simulations were written in MATLAB R2018a on a CPU 2.6 GHz and 12 Gb of RAM HP Notebook with Windows 10.

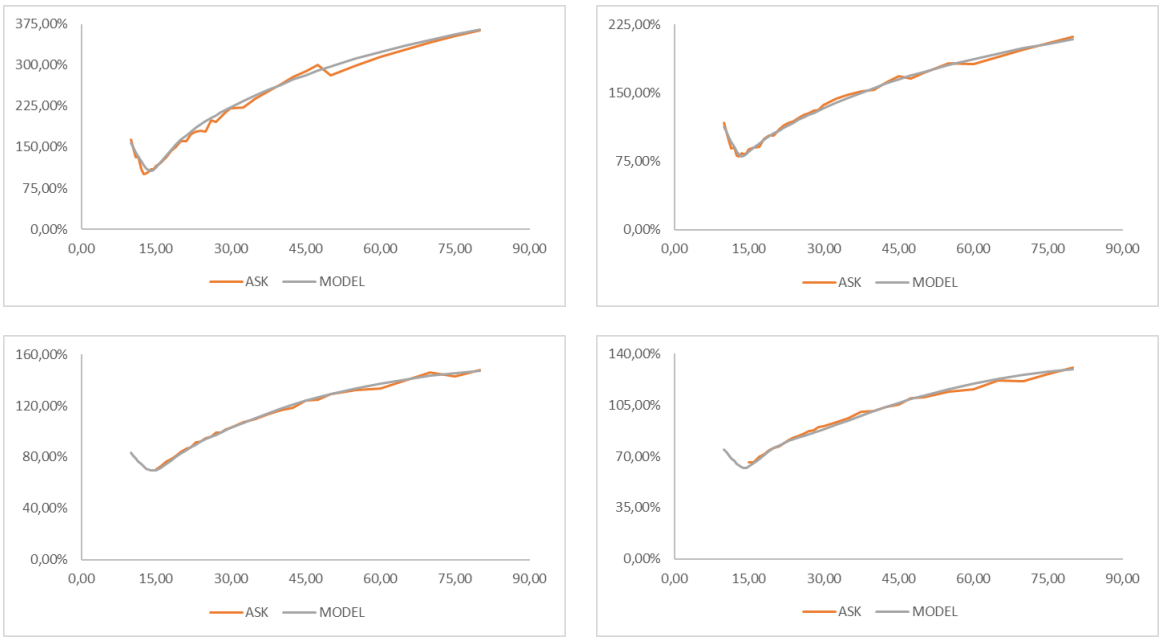


Figure 3.11: VIX futures model implied volatility against ask VIX call options market quotes, as on November 7, 2019, for the four maturities November 20, 2019, December 18, 2019, January 22, 2020 and February 19, 2020, from the top left corner going clockwise.

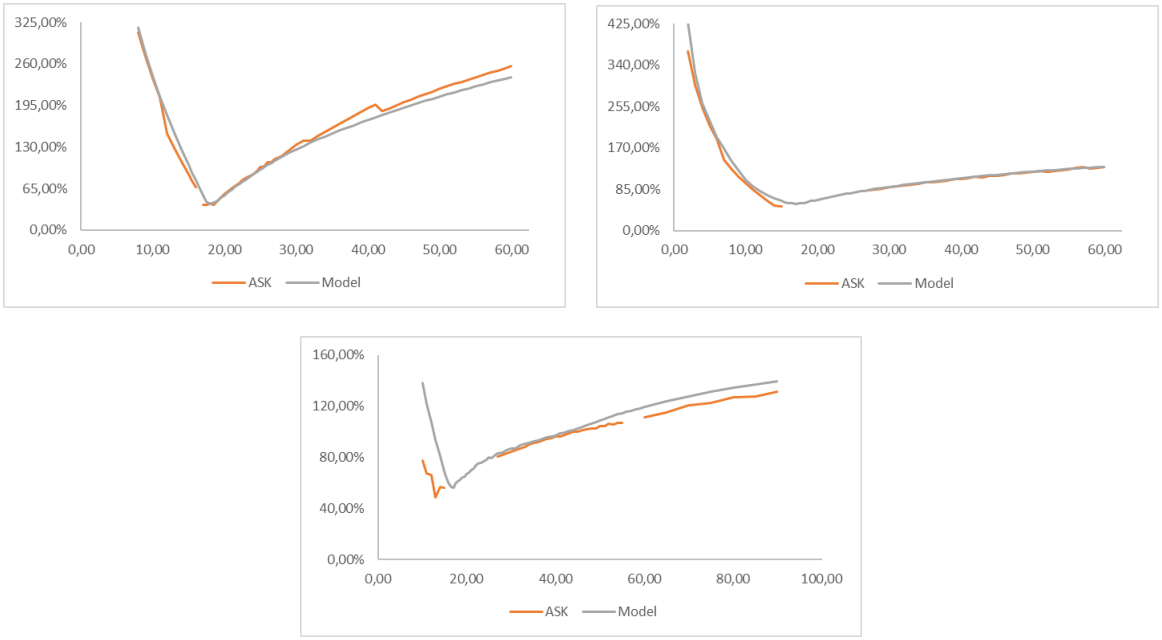


Figure 3.12: VXX model implied volatility against ask VXX call options market quotes, as on November 7, 2019, for the three maturities November 15, 2019, December 20, 2019 and January 17, 2020, from the top to the bottom.

3.6 Conclusion and Further Developments

In this chapter, we have presented a general framework for a calibration of Exchange-Traded Products (ETP's) based on futures strategies. In a numerical simulation based on real data we have considered as an example of ETP the VXX ETN, which is a strategy on the nearest and second-nearest maturing VIX futures contracts. The main ingredient to achieve our goal was a full stochastic-local volatility model that can be calibrated on a book of options on a general ETP and its underlying futures contracts in a parsimonious manner. A parameters sensitivity analysis allowed us to fix suitable values for the model parameters in order to fit the smile of the VIX and VXX ETN, with a good level of accuracy. Moreover, our specification led to an admissible local correlation function that justifies the consistency of the whole procedure. The implementation of a full calibration algorithm (in order to find the optimal value of all model parameters) based on neural networks techniques is currently under investigation. Last but not least, the methodology proposed has been applied to the case of VIX futures and VXX ETN, but the framework is flexible enough to tackle also more general ETP's.

Chapter 4

Pricing with Absorbed and Reflected Processes: a Lie Symmetry Approach

4.1 Introduction

A symmetry of a differential equation can be defined as a transformation that maps solutions to solutions. Continuous symmetries which have group properties are called Lie symmetries, developed in Lie (1912). The theory has been more recently addressed in Olver (1993) and Blumans's works such as Bluman and Kumei (1989). The range of applications of Lie symmetry groups is massive. Symmetries have been applied in several fields, from epidemiology to finance. Further applications are also found in stochastic analysis, where it is employed for the computation of transition probability densities and functionals. For an overview of the possible applications, please refer, for example, to Craddock and Lennox (2007), Craddock and Lennox (2009), Craddock and Platen (2004), Lennox (2011), Craddock (2009), Craddock and Lennox (2012), Baldeaux and Platen (2013), Goard (2011), Goard (2000), Goard and Mazur (2013), Andersen (2011), Lipton (2002), Zuhlsdorff (2002), Carr et al. (2006), Carr et al. (2013), Cordoni and Di Persio (2014), Grasselli (2017), Craddock and Grasselli (2019) and Itkin (2013).

Applications of symmetry methods mostly rely on developing group invariant solutions, reduction of PDEs to canonical forms and simple mappings of trivial solutions to non trivial solutions. In this chapter we adopt a new and still developing methodology which unites Lie symmetry analysis of linear partial differential equations (PDEs) with harmonic analysis. In this framework, a solution obtained by application of a symmetry from a known solution is to be regarded as the kernel of an integral operator, which maps suitable initial data to new solutions. The result is the computation of fundamental solutions by inversion of integral transforms obtained by symmetry, rather than through the construction of an invariant solution to a boundary value problem. The rigorous justification of these new methods relies on the representation theory of the underlying symmetry group and we refer the reader to the literature in the references; e.g. Craddock and Lennox (2007), Craddock and Lennox (2009), Craddock and Platen (2004), Lennox (2011), Craddock (2009), Craddock and Dooley (2010) and Craddock and Lennox (2012).

If we have at hand a Lie group that leaves the PDE invariant, then we can apply it to any solution

to obtain another solution. The key result for finding the Lie groups admitted by the PDE is Lie's theorem. For a PDE of order n, we write down the infinitesimal generator $\mathbf{v}$ of the symmetry and obtain its so-called n-th *prolongation* $\mathrm{pr}^n\mathbf{v}$, as we shall see next. Lie's theorem states that $\mathbf{v}$ generates a symmetry of the PDE

$$P(x, D^\alpha u) = 0,\ x \in \Omega \subseteq \mathbb{R}^m,\ D^\alpha u = \partial_{x_1^{\alpha_1} \cdots x_n^{\alpha_m}} u,\ \alpha_1 + \cdots + \alpha_m \leq n,$$

if and only if $\mathrm{pr}^n\mathbf{v}[P(x, D^\alpha u)] = 0$ whenever $P(x, D^\alpha u) = 0$. All the details can be found in Olver (1993, Chapter 2).

This condition leads to an equation in a set of independent functions of the derivatives of u. As the equation must be true for arbitrary values of these independent functions, their coefficients must vanish, leading to a linear system of equations known as the Determining Equations. Once these equations are solved, one can find the corresponding Lie group admitted by the PDE.

The main goal of the present chapter is to present a method for the direct solution of an inhomogeneous boundary value problem arising in the study of diffusions, more specifically arising in the Black-Scholes framework. In principle, the method we present do not require a change of variables or rely on the invariance of the PDE, however, we combine symmetry analysis and integral equations to construct an explicit solutions. We analyse the so-called Robin boundary condition, which is somehow a combination between an absorbed and reflected boundary condition. To make this more precise, consider an Itô process $\{S_t\}_{t\geq 0}$ satisfying the SDE

$$dS_t = \mu\left(S_t, t\right) dt + \sigma\left(S_t, t\right) dW_t, S(0) = S_0, \quad S_0 \in \mathcal{D} \subseteq \mathbb{R},$$

where $\mathcal{D}$ is an open interval of the real line and $\{W_t\}_{t\geq 0}$ is a standard Brownian motion. The transition probability density function for the diffusion is a positive fundamental solution of the Kolmogorov backwards equation

$$u_t + \frac{1}{2}\sigma(S,t)^2 u_{SS} + \mu(S,t)u_S = 0, S \in \mathcal{D}, t > 0,$$

that is also a probability density function. Given $b > 0$, we say that the process S is absorbed at $S = b$ if the boundary condition at b is $u(b,t) = 0$, while the process S is reflected at b if the boundary condition at b is $u_S(b,t) = 0$. The mixed case, which is known as Robin boundary condition, has the form $\alpha u(b,t) + \beta u_S(b,t) = 0$, where α and β are constants.

The chapter is organized as follows. In Section 4.2 we give a brief theoretical background on Lie symmetries for PDEs, introducing the prolongation formula and applying it to the heat equation. In Section 4.3, we present the general setting of our problem, introducing the Black-Scholes PDE with the initial and the Robin boundary conditions. In Section 4.4, we solve the problem in the case of an homogeneous Robin boundary condition, using techniques coming from Fourier and Laplace-based transforms. In Section 4.5 we solve the problem in the case of an inhomogeneous Robin boundary condition. We start by solving the problem using two different approach, one more focused on inversion of transforms and one more algebraic-oriented. In the end, we prove a broader result, that can be applied to general positive second order operators. In Section 4.6 we draw some conclusion and remarks.

4.2 Lie Symmetries for PDEs

In this Section, Lie groups transformation will be briefly introduced with a particular focus on their applications to the solution of PDEs. The introduction will closely follow the presentation given in Bluman and Kumei (1989). Another standard reference for this topic is Olver (1993).

4.2.1 Invariance of a Differential Equation

Let us now take into account the second order PDE in two independent variables (x, t) of the form

$$F(x, t, u, u_x, u_t, u_{xx}, u_{xt}, u_{tt}) = 0. \tag{4.1}$$

We work exclusively on a second order PDE in two independent variables as it is the setting of interest. Surely, the mentioned results can be easily extended to systems of partial differential equations of any order in any dimension.

We now give a definition that will be crucial in the following. The PDE (4.1) is said to admit the Lie group of transformations $(X_\epsilon, T_\epsilon, U_\epsilon)_\epsilon$, for $\epsilon > 0$, if the family of solutions of (4.1) is an invariant family of surfaces for $(X_\epsilon, T_\epsilon, U_\epsilon)_\epsilon$.

In other words, this implies that if we have at hand a Lie group G that leaves the PDE invariant, then we can apply it to any solution to get another solution. If the solution we start from is itself an invariant surface, then by applying the Lie group to it, we will end up with the solution itself. Here the crucial concept is that the Lie group must act on the space where the solutions of the PDE lives, in our case the (x, t, u)-space.

We introduce a vector field

$$\mathbf{v} = \xi(x, t, u)\partial_x + \tau(x, t, u)\partial_t + \phi(x, t, u)\partial_u, \tag{4.2}$$

which is the infinitesimal generator of G and it is also known as the infinitesimal symmetry. We can extend the action of G in a natural way to act also on the derivatives of u up to a given order. The only requirement is that if G acts on the triple (x, t, u) via $G : (x, t, u) \to (\tilde{x}, \tilde{t}, \tilde{u})$, then the *first prolongation* $G^{(1)}$ of G acts via

$$G^{(1)} : (x, t, u, u_x, u_t) \to (\tilde{x}, \tilde{t}, \tilde{u}, \tilde{u}_{\tilde{x}}, \tilde{u}_{\tilde{t}}).$$

In other words G sends $u(x, t)$ to a transformed function $\tilde{u}(\tilde{x}, \tilde{t})$. If we apply $G^{(1)}$ to the derivative $\frac{\partial u}{\partial x}$, then it must give $\frac{\partial \tilde{u}}{\partial \tilde{x}}$ and similarly for the t derivatives. The general rules is that the second prolongation must also send derivatives up to second order to the corresponding derivatives of the transformed function, the third prolongation sends derivatives up to third order to the corresponding derivatives of the transformed function and so forth. In fact, one can derive a formula for the prolongation with the core principle being only the requirement that the chain rule holds. We denote the second prolongation of G by $G^{(2)}$. The generator of $G^{(2)}$ is the second prolongation of the vector field $\mathbf{v}$. This is given by (see e.g. Bluman and Kumei (1989) and Olver (1993)):

$$\mathrm{pr}^2\mathbf{v} = \mathbf{v} + \phi^x \frac{\partial}{\partial u_x} + \phi^t \frac{\partial}{\partial u_t} + \phi^{xx} \frac{\partial}{\partial u_{xx}} + \phi^{xt} \frac{\partial}{\partial u_{xt}} + \phi^{tt} \frac{\partial}{\partial u_{tt}}.$$

Expressions for ϕ^x, ϕ^t an so on are given by the prolongation formula (4.3) given below.

The main result due to Lie (see e.g. Bluman and Kumei (1989)) is that the PDE (4.1) admits the Lie group of transformations $(X_\epsilon, T_\epsilon, U_\epsilon)$ if and only if

$$F(x,t,u,u_x,u_t,u_{xx},u_{xt},u_{tt}) = 0 \rightarrow (G^{(2)}F)(x,t,u,u_x,u_t,u_{xx},u_{xt},u_{tt}) = 0,$$

namely the surface $\{F(x,t,u,u_x,u_t,u_{xx},u_{xt},u_{tt}) = 0\}$ is invariant for the second prolongation $G^{(2)}$ of the group action G. We refer e.g. to Olver (1993) or Bluman and Kumei (1989) for the construction of the second prolongation for a group. At the level of vector fields, this leads to Lie's Theorem which states that $\mathbf{v}$ generates a one parameter group of symmetries if and only if

$$\mathrm{pr}^2\mathbf{v}[F(x,t,u,u_x,u_t,u_{xx},u_{xt},u_{tt})] = 0,$$

whenever $F(x,t,u,u_x,u_t,u_{xx},u_{xt},u_{tt}) = 0$.

The coefficients ϕ^x, ϕ^t and so on are given by the prolongation formula (4.3), first published by Olver (1978): [1] We introduce the notation

$$\phi^{\mathbf{J}}_j(x,u^{(n)}) = D_{\mathbf{J}}\big(\phi_j - \sum_{i=1}^{p}\xi^i u^j_i\big) + \sum_{i=1}^{p}\xi^i u^j_{\mathbf{J},i}, \tag{4.3}$$

where $u^j_i = \dfrac{\partial u^j}{\partial x^i}$, and $u^j_{\mathbf{J},i} = \dfrac{\partial u^j_{\mathbf{J}}}{\partial x^i}$, and $D_{\mathbf{J}}$ is the total differentiation operator. So that for example $D_xF(x,t,u(x,t)) = F_x + u_xF_u$.

In practice, we use the variable names rather than the multi-indices in the exponents. So we write ϕ^{xx} rather than $\phi^{1,1}$. From Chapter 2 of Olver (1993) we find that the coefficient functions ϕ^x and ϕ^t are

$$\begin{aligned}
\phi^x &= D_x\left(\phi - \xi u_x - \tau u_t\right) + \xi u_{xx} + \tau u_{xt} \\
&= \left(\phi_x + \phi_u u_x - \xi_x u_x - \xi_u u_x^2 - \xi u_{xx} - \tau_x u_t - \tau_u u_x u_t - \tau u_{xt}\right) + \xi u_{xx} + \tau u_{xt} \\
&= \phi_x + (\phi_u - \xi_x)u_x - \xi_u u_x^2 - \tau_x u_t - \tau_u u_x u_t.
\end{aligned}$$

In a similar way

$$\begin{aligned}
\phi^t &= D_t\left(\phi - \xi u_x - \tau u_t\right) + \xi u_{xt} + \tau u_{tt} \\
&= \phi_t - \xi_t u_x + (\phi_u - \tau_t)u_t - \xi_u u_x u_t - \tau_u u_t^2,
\end{aligned}$$

and

$$\begin{aligned}
\phi^{xx} &= D_{xx}\left(\phi - \xi u_x - \tau u_t\right) + \xi u_{xxx} + \tau u_{xxt} \\
&= \phi_{xx} + (2\phi_{xu} - \xi_{xx})u_x - \tau_{xx}u_t + (\phi_{uu} - 2\xi_{xu})u_x^2 - 2\tau_{xu}u_x u_t \\
&\quad - \xi_{uu}u_x^3 - \tau_{uu}u_x^2 u_t + (\phi_u - 2\xi_x)u_{xx} - 2\tau_x u_{xt} - 3\xi_u u_x u_{xx} \\
&\quad - \tau_u u_t u_{xx} - 2\tau_u u_x u_{xt}.
\end{aligned}$$

The coefficients ϕ^{xt} and ϕ^{tt} can be obtained in the same manner, but we do not need them for our calculations.

[1] Lie derived a recursive formula for $\mathrm{pr}^n\mathbf{v}$ in terms of $\mathrm{pr}^{n-1}\mathbf{v}$.

Partial derivatives of u	Coefficients
1	$\phi_\tau = \phi_{xx}$
u_x	$-\xi_t = 2\phi_{xu} - \xi_{xx}$
u_x^2	$0 = \phi_{uu} - 2\xi_{xu}$
u_x^3	$0 = -\xi_{uu}$
u_{xx}	$\phi_u - \tau_t = -\tau_{xx} + \phi_u - 2\xi_x$
$u_x u_{xx}$	$-\xi_u = -2\tau_{xu} - 3\xi_u$
$u_x^2 u_{xx}$	$0 = -\tau_{uu}$
u_{xx}^2	$-\tau_u = -\tau_u$
u_{xt}	$0 = -2\tau_x$
$u_x u_{xt}$	$0 = -2\tau_u$

Table 4.1: Determining equations for the symmetry group of the Heat equation.

The Heat Equation

We begin with what in many ways is the canonical example of Lie symmetry analysis, the one dimensional heat equation

$$u_t = u_{xx}. \tag{4.4}$$

This was one of the first PDEs whose symmetries were studied by Lie himself. To compute the symmetries of (4.4), let $\mathbf{v}$ be given by (4.2). We seek to determine all possible coefficient functions ξ, τ and ϕ such that the corresponding one-parameter group $\exp(\epsilon\mathbf{v})$ is a symmetry group of the heat equation.

The heat equation is a second order PDE, so we need the second prolongation of $\mathbf{v}$. By Lie's Theorem, $\mathbf{v}$ generates symmetries of the heat equation if and only if

$$\mathrm{pr}^2\mathbf{v}[u_{xx} - u_t] = 0,$$

whenever $u_t = u_{xx}$. Applying $\mathrm{pr}^2\mathbf{v}$ to (4.4), we obtain

$$\phi^t = \phi^{xx}, \tag{4.5}$$

which must be satisfied whenever $u_t = u_{xx}$. Substituting the formulae for ϕ^t and ϕ^{xx} which were derived above into (4.5), and replacing u_t with u_{xx}, we obtain the following system of defining equations for the functions ξ, τ and ϕ.

$$\begin{aligned}
&\phi_t - \xi_t u_x + (\phi_u - \tau_t)u_{xx} - \xi_u u_x u_{xx} - \tau_u u_{xx}^2 = \phi_{xx} + (2\phi_{xu} - \xi_{xx})u_x \\
&- \tau_{xx}u_{xx} + (\phi_{uu} - 2\xi_{xu})u_x^2 - 2\tau_{xu}u_x u_{xx} - \xi_{uu}u_x^3 - \tau_{uu}u_x^2 u_{xx} \\
&+ (\phi_u - 2\xi_x)u_{xx} - 2\tau_x u_{xt} - 3\xi_u u_x u_{xx} - \tau_u u_{xx}^2 - 2\tau_u u_x u_{xt}.
\end{aligned}$$

To solve this system, we equate the coefficients of the partial derivatives of u. We then obtain a set of determining equations for the symmetry group of the heat equation, that are shown in Table 4.1.

First notice that $\phi_t = \phi_{xx}$ must hold. Looking at the coefficient of $u_x u_{xt}$, gives $\tau_u = 0$. Also, the coefficient of u_{xt} tells us that $2\tau_x = 0$. Hence τ cannot depend upon x or u.

From the coefficient of $u_x u_{xx}$, we further have $\xi_u = 0$, and equating the coefficients of u_{xx} gives $\xi_x = \frac{1}{2}\tau_t$. Thus ξ depends only on x and t, and integration gives

$$\xi(x,t) = \frac{1}{2}x\tau_t + \rho(t),$$

for some undetermined function ρ of t. From the coefficient of u_x^2 we see that $\phi_{uu} = 0$, which tells us that ϕ is linear in u. Hence we can write

$$\phi(x,t,u) = \alpha(x,t)u + \beta(x,t),$$

where α and β are to be determined. Since $\phi_t = \phi_{xx}$ we conclude that $\beta_t = \beta_{xx}$ is a solution of the heat equation. We have no further information about this function. We also know that $\alpha_t = \alpha_{xx}$. We have more information about α. Comparing the coefficients of u_x we find that

$$-\xi_t = 2\alpha_x - \xi_{xx}. \tag{4.6}$$

So

$$\alpha_x = -\frac{1}{4}x\tau_{tt} - \frac{1}{2}\rho_t.$$

Integrating this equation gives

$$\alpha = -\frac{1}{8}x^2\tau_{tt} - \frac{1}{2}x\rho_t + \sigma(t). \tag{4.7}$$

We see then that

$$\alpha_t = -\frac{1}{8}x^2\tau_{ttt} - \frac{1}{2}x\rho_{tt} + \sigma_t \tag{4.8}$$

and

$$\alpha_{xx} = -\frac{1}{4}\tau_{tt}. \tag{4.9}$$

We can then determine the forms of the functions τ, ρ and σ from the requirement that $\alpha_t = \alpha_{xx}$. This tells us that $\tau_{ttt} = 0, \rho_{tt} = 0$, and $\sigma_t = -\frac{1}{4}\tau_{tt}$. So τ is quadratic in t, ρ is linear and σ is linear. In this way, we are able to determine all the coefficient functions ξ, τ and ϕ for the infinitesimal symmetry of the heat equation. A convenient representation is

$$\begin{aligned}
\xi &= c_1 + c_4 x + 2c_5 t + 4c_6 xt, \\
\tau &= c_2 + 2c_4 t + 4c_6 t^2 \\
\phi &= (c_3 - c_5 x - 2c_6 t - c_6 x^2)u + \beta(x,t),
\end{aligned}$$

where $c_1, \ldots, c_6$ are arbitrary constants of integration, and $\beta(x,t)$ is an arbitrary solution of the heat equation. Writing out the most general form of $\mathbf{v}$ with these coefficients, and using the

numbering of the constants to determine the grouping, we now see that a basis for the Lie algebra of symmetries is given by the six vector fields

$$\mathbf{v}_1 = \partial_x, \quad \mathbf{v}_2 = \partial_t, \quad \mathbf{v}_3 = u\partial_u, \ \mathbf{v}_4 = x\partial_x + 2t\partial_t - \frac{1}{2}u\partial_u,$$
$$\mathbf{v}_5 = 2t\partial_x - xu\partial_u, \ \mathbf{v}_6 = 4xt\partial_x + 4t^2\partial_t - (x^2 + 2t)u\partial_u.$$

There is also an infinite dimensional ideal within the Lie algebra consisting of vector fields of the form $\mathbf{v}_\beta = \beta(x,t)\partial_u$. Here β is an arbitrary solution of the heat equation. The symmetry produced by $\mathbf{v}_\beta$ is simply a superposition of solutions. That is, if u is a solution of the heat equation, and β is a solution, then so is $u + \epsilon\beta$.

In Appendix A we present an application of this technology to the Black-Scholes PDE, exponentiating the corresponding Lie algebra of symmetries.

4.3 The Problem

We start this section by introducing the general setting of the problem we are going to tackle. Let $\mathcal{D}$ be an open interval of the real line and consider the Black-Scholes PDE:

$$u_t = \frac{1}{2}\sigma^2 S^2 u_{SS} + rSu_S - ru, \quad S \in \mathcal{D} \subseteq \mathbb{R}, \ t > 0, \tag{4.10}$$

with initial condition given by

$$u(S,0) = f(S),$$

and subject to the general Robin boundary condition

$$\alpha u(b,t) + \beta u_S(b,t) = g(t), \quad b > 0, \tag{4.11}$$

where σ, r, b and α, β are constants. The PDE in (4.10) is the Kolmogorov backward equation associated to a diffusion S satisfying

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

killed at rate r, where W is a standard Brownian motion defined in a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$.

Example 4.3.1. *We can consider* (4.11) *as the PDE arising from the problem of pricing a contract whose payoff involves the presence of a lower barrier ($S > b$). The case $\beta = 0$ corresponds to an absorbing barrier, namely the process is absorbed at $S = b$, while the case $\alpha = 0$ is related to a reflection at b. The general mixed case is referred as the Robin boundary condition. Of course, we could also consider $S < b$ in the case of an upper barrier problem.*

We get rid of the killing by the change of variable $u = e^{-rt}w$, giving

$$w_t = \frac{1}{2}\sigma^2 S^2 w_{SS} + rSw_S.$$

The initial boundary conditions is the same since

$$u(S,0) = f(S) = e^0 w(S,0) = w(S,0).$$

After the change of variable the Robin boundary condition reads as

$$\alpha e^{-rt} w(b,t) + \beta e^{-rt} w_S(b,t) = g(t), \quad b > 0,$$

namely

$$\alpha w(b,t) + \beta w_S(b,t) = e^{rt} g(t), \quad b > 0.$$

Hence, up to rename the function $g(t)$, we can also assume the same Robin boundary condition

$$\alpha w(b,t) + \beta w_S(b,t) = g(t), \quad b > 0.$$

Therefore, from now on, we can consider the equivalent problem

$$\begin{aligned} w_t &= \frac{1}{2}\sigma^2 S^2 w_{SS} + rSw_S, \quad S > b > 0 \\ w(S,0) &= f(S), \\ \alpha w(b,t) + \beta w_S(b,t) &= g(t). \end{aligned} \tag{4.12}$$

Remark 4.3.2. *In mathematical finance the option pricing problem is typically formulated in terms of a PDE with a terminal boundary associated to the payoff at maturity T, while in our approach we considered an initial condition. Of course, one can perform the deterministic change of variable $\tilde{t} = T - t$ in order to transform the problem into our framework. We skip the details for sake of brevity.*

In the next sections we are going to solve problem (4.12), first in the case where $g(t) = 0$ and then in the more general case. We decide to treat separately the homogeneous and inhomogeneous Robin condition problems because we are going to exploit different techniques that can have separate interest in itself. Finally, we provide a general representation for the solution of problem (4.12) as well.

4.4 Homogeneous Robin Condition

In this section we solve the problem (4.12) in the case $g(t) = 0$[2], namely

$$\begin{aligned} w_t &= \frac{1}{2}\sigma^2 S^2 w_{SS} + rSw_S, \quad S > b > 0 \\ w(S,0) &= f(S), \\ \alpha w(b,t) + \beta w_S(b,t) &= 0. \end{aligned} \tag{4.13}$$

[2]In Appendix A, we are going to provide another method to find a solution of the PDE in (4.13) via Lie symmetry methods.

We look for a separable solution of the form $w(S,t) = e^{\lambda t} v(S)$, so let us solve the ODE

$$\frac{1}{2}\sigma^2 S^2 v''(S) + rSv'(S) - \lambda v(S) = 0,$$

with (the constant) λ and (the function) v to be determined. Setting

$$v(S) = S^\alpha,$$

we get

$$\frac{1}{2}\sigma^2 S^2 \alpha(\alpha-1) S^{\alpha-2} + rS\alpha S^{\alpha-1} - \lambda S^\alpha = 0,$$

that is

$$\frac{1}{2}\sigma^2 \alpha(\alpha-1) + r\alpha - \lambda = 0,$$

which gives

$$\alpha = \frac{-\left(r - \frac{\sigma^2}{2}\right) \pm \sqrt{\left(r - \frac{\sigma^2}{2}\right)^2 + 2\sigma^2\lambda}}{\sigma^2}.$$

Now put

$$\left(r - \frac{\sigma^2}{2}\right)^2 + 2\sigma^2\lambda = -\xi^2\sigma^4,$$

that is, we reparametrize λ in terms of ξ

$$\lambda = -\frac{\xi^2\sigma^2}{2} - \frac{1}{2\sigma^2}\left(r - \frac{\sigma^2}{2}\right)^2, \tag{4.14}$$

so that we can express α as follows

$$\alpha = \mu \pm i\xi, \tag{4.15}$$

where

$$\mu = \frac{\sigma^2 - 2r}{2\sigma^2}. \tag{4.16}$$

The general solution of the ODE can be written as $w(S,t,\lambda) = e^{\lambda t}\left(c_1 v_1 + c_2 v_2\right)$, where v_1, v_2 are independent solutions of the ODE, that is

$$w(S,t,\lambda) = e^{\lambda t}\left(c_1 S^{\mu+i\xi} + c_2 S^{\mu-i\xi}\right).$$

The next step is finding the solution that satisfies the Robin boundary condition in terms of sines and cosines. From

$$S^{\mu\pm i\xi} = S^\mu\left(\cos(\xi \ln S) \pm i \sin(\xi \ln S)\right)$$

and taking real and complex parts, we get that the solution can be written as

$$\begin{aligned} w^{\xi}(S,t) &= e^{\lambda t} S^{\mu} \left(A\cos(\xi \ln S) + B\sin(\xi \ln S)\right) \\ &= e^{-\frac{\xi^2\sigma^2}{2}t + ct} S^{\mu} \left(A\cos(\xi \ln S) + B\sin(\xi \ln S)\right) \end{aligned} \tag{4.17}$$

where $c = -\frac{1}{2\sigma^2}\left(r - \frac{\sigma^2}{2}\right)^2$ and A, B are constants to be determined by the Robin boundary condition. Exploiting the Robin condition, we get

$$\begin{aligned} &\alpha b^{\mu} e^{-\frac{\xi^2\sigma^2}{2}t + ct} \left(A\cos\left(\xi \ln b\right) + B\sin\left(\xi \ln b\right)\right) + \\ &\beta b^{\mu-1} e^{-\frac{\xi^2\sigma^2}{2}t + ct} \left(\mu\left(A\cos(\xi \ln b) + B\sin(\xi \ln b)\right) + B\xi\cos(\xi \ln b) - A\xi\sin(\xi \ln b)\right) = 0, \end{aligned}$$

that, after some algebraic manipulations, leads to the linear relation

$$\begin{aligned} 0 =& A\left(b\alpha\cos(\xi \ln b) + \beta\mu\cos(\xi \ln b) - \beta\xi\sin(\xi \ln b)\right) \\ &+ B\left(\beta\xi\cos(\xi \ln b) + b\alpha\sin(\xi \ln b) + \beta\mu\sin(\xi \ln b)\right), \end{aligned}$$

from which we have

$$A = \beta\xi\cos(\xi \ln b) + b\alpha\sin(\xi \ln b) + \beta\mu\sin(\xi \ln b) \tag{4.18}$$

$$B = -b\alpha\cos(\xi \ln b) - \beta\mu\cos(\xi \ln b) + \beta\xi\sin(\xi \ln b). \tag{4.19}$$

With this choice of the coefficients A and B, the function (4.17) solves the PDE (4.13) subject to the homogeneous Robin condition.

In order to complete the solution to the problem (4.12), we have to ensure the general boundary condition, namely including $w(S,0) = f(S)$. To do that, we use the following lemma, whose proof can be found e.g. in Olver (1993).

Lemma 4.4.1. *Suppose that we have a linear PDE*

$$P\left(x, D^{\alpha}\right) u = \sum_{|\alpha| \leqslant n} a_{\alpha}(x) D^{\alpha} u, \quad x \in \Omega \subseteq \mathbb{R}^m \tag{4.20}$$

with $\alpha = (\alpha_1, \ldots, \alpha_m)\,,\ \alpha_i \in \mathbb{N},\ |\alpha| = \alpha_1 + \cdots + \alpha_m$ *and* $D^{\alpha} = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_m^{\alpha_m}}$. *Suppose that* $w^{\xi}(x)$ *is a continuous one parameter family of solutions of the PDE in* (4.20)*, which holds for* $\xi \in I \subseteq \mathbb{R}$. *Suppose further that* $\varphi\colon I \to \mathbb{R}$ *is a function with sufficiently rapid decay. Then*

$$w(x) = \int_I \varphi(\xi) w^{\xi}(x)\, d\xi$$

is also a solution of the PDE in (4.20)*. Furthermore, if the PDE in* (4.20) *is time dependent and* $w^{\xi}(x,t)$ *is the family of symmetry solution, then* $w(x,t) = \int_I \varphi(\xi) w^{\xi}(x,t)\, d\xi$ *is also a solution of the PDE in* (4.20)*.*

Hence, we let

$$w(S,t) = \int_0^\infty \varphi(\xi) w^\xi(S,t) d\xi, \tag{4.21}$$

and thanks to Lemma 4.4.1, if the function φ has sufficient decay, it turns out that also $w(S,t)$ is a solution of the PDE in (4.12). Moreover, the function (4.21) with A and B as in (4.18) and (4.19), satisfies the Robin boundary condition, by construction. We have now to choose φ in order to satisfy the initial boundary condition, namely $w(S,0) = f(S)$. Taking $t = 0$ we get the following integral equation

$$f(S) = \int_0^\infty \varphi(\xi) w^\xi(S,0) d\xi,$$

namely

$$\int_0^\infty \varphi(\xi) \left(A\cos(\xi \ln S) + B\sin(\xi \ln S)\right) d\xi = S^{-\mu} f(S). \tag{4.22}$$

Using the value of A and B found in (4.18), (4.19) and using some elementary trigonometric identities, we can rewrite (4.22) as

$$\int_0^\infty \varphi(\xi) \left(\beta\xi \cos\left(\xi \ln\left(\frac{b}{S}\right)\right) + (b\alpha + \beta\mu)\sin\left(\xi \ln\left(\frac{b}{S}\right)\right)\right) d\xi = S^{-\mu} f(S).$$

We now reduce this integral transform to a Fourier sine transform by following an argument from Craddock (2019b). Let $\zeta(x) = \int_0^\infty \varphi(\xi)\sin(\xi x) d\xi$, so that the integral transform becomes a linear ODE that, after the change of variable $x = \ln\left(\frac{b}{S}\right)$, reads

$$\beta\zeta'(x) + (b\alpha + \beta\mu)\zeta(x) = b^{-\mu} e^{\mu x} f(be^{-x}), \tag{4.23}$$

with initial condition $\zeta(0) = 0$. The ODE admits the solution

$$\begin{aligned}\zeta(x) &= \frac{e^{-\frac{\tilde{A}x}{\beta}}}{\beta} \int_0^x b^{-\mu} e^{\frac{(\tilde{A}+\beta\mu)z}{\beta}} f(be^{-z}) dz \\ &= \frac{b^{-\mu}}{\beta} \int_0^x e^{\frac{\tilde{A}}{\beta}(z-x)+\mu z} f(be^{-z}) dz,\end{aligned} \tag{4.24}$$

where, for sake of ease, we set $\tilde{A} = b\alpha + \beta\mu$.
The Fourier sine transform $\zeta(x) = \int_0^\infty \phi(\xi)\sin(\xi x) d\xi$ can be inverted by the formula

$$\phi(\xi) = \frac{2}{\pi} \int_0^\infty F(\eta)\sin(\xi\eta) d\eta, \tag{4.25}$$

where

$$F(\eta) = \frac{b^{-\mu}}{\beta} \int_0^\eta e^{\frac{\tilde{A}}{\beta}(z-\eta)+\mu z} f(be^{-z}) dz. \tag{4.26}$$

The solution of the boundary value problem is then

$$\begin{aligned} w(S,t) &= \int_0^\infty \phi(\xi) w^\xi(S,t) d\xi \\ &= \int_0^\infty \frac{2}{\pi} \int_0^\infty F(\eta) \sin(\xi\eta)\, w^\xi(S,t) d\eta d\xi \\ &= \int_0^\infty \int_0^\infty \frac{2}{\pi} F(\eta) \sin(\xi\eta) w^\xi(S,t) d\xi d\eta \\ &= \int_0^\infty F(\eta) \mathcal{G}(S,t,\eta) d\eta, \end{aligned} \tag{4.27}$$

where

$$\begin{aligned} \mathcal{G}(S,\eta,t;A,B) &= \frac{2}{\pi} \int_0^\infty \sin(\xi\eta) w^\xi(S,t) d\xi \\ &= \frac{2}{\pi} \int_0^\infty \sin(\xi\eta) e^{-\frac{\xi^2\sigma^2}{2}t + ct} S^\mu \left(A\cos(\xi \ln S) + B \sin(\xi \ln S)\right) d\xi \\ &= \frac{e^{ct - \frac{\left(\eta - \ln\left(\frac{b}{S}\right)\right)^2}{2\sigma^2 t}}}{\sqrt{2\pi}\left(\sigma^2 t\right)^{3/2}} S^\mu k(S,\eta,t), \end{aligned} \tag{4.28}$$

where A and B are defined in (4.18) and (4.19), respectively, $\mu = \frac{\sigma^2 - 2r}{2\sigma^2}$, $c = -\frac{1}{2\sigma^2}\left(r - \frac{\sigma^2}{2}\right)^2$ and where

$$k(S,\eta,t) = e^{-\frac{2\eta \ln\left(\frac{b}{S}\right)}{\sigma^2 t}} \left(\beta\eta - t\sigma^2\tilde{A} + \beta \ln\left(\frac{b}{S}\right)\right) + \left(\beta\eta + t\sigma^2\tilde{A} - \beta\ln\left(\frac{b}{S}\right)\right).$$

Finally, recall that by definition we have

$$w(S,t) = \int_0^\infty f(y) p(S,y,t) dy,$$

then using Fubini we can deduce the form of the transition density of the Geometric Brownian Motion (GBM) subject to the Robin boundary conditions. In fact, we can rewrite the solution as

$$\begin{aligned} w(S,t) &= \int_0^\infty F(\eta)\mathcal{G}(S,t,\eta) d\eta \\ &= \int_0^\infty \frac{b^{-\mu}}{\beta} \int_0^\eta e^{\frac{\tilde{A}}{\beta}(z-\eta) + \mu z} f(be^{-z}) \mathcal{G}(S,t,\eta) dz d\eta \\ &= \int_0^\infty f(be^{-z}) \int_z^\infty \frac{b^{-\mu}}{\beta} e^{\frac{\tilde{A}}{\beta}(z-\eta)+\mu z} \mathcal{G}(S,t,\eta) d\eta dz \\ &= \int_0^\infty f(be^{-z}) \bar{p}(S,z,t) dz \\ &= \int_0^b f(y) p(S,y,t) dy. \end{aligned} \tag{4.29}$$

Therefore, we deduce the transition density

$$p(S,y,t) = \frac{1}{y}\bar{p}\left(S, \ln\left(\frac{b}{y}\right), t\right), \tag{4.30}$$

where

$$\bar{p}(S,z,t) = \int_z^\infty \frac{b^{-\mu}}{\beta} e^{\frac{\tilde{A}}{\beta}(z-\eta)+\mu z}\mathcal{G}(S,t,\eta)d\eta.$$

Quite surprisingly, the integral determining the kernel in (4.30) can be computed explicitly in term of Gaussians and error functions. Indeed, we can actually compute $p(S,y,t)$ as

$$p(S,y,t) = -\frac{e^{ct+\frac{\sigma^2 t\tilde{A}^2}{2\beta^2}} S^{\mu-\frac{\tilde{A}}{\beta}} y^{-\mu-1-\frac{\tilde{A}}{\beta}} \tilde{A} b^{\frac{2\tilde{A}}{\beta}}}{\beta} \operatorname{erfc}\left(\frac{\sigma^2 t\tilde{A} + \beta\left(\ln\frac{b}{y} + \ln b - \ln S\right)}{\sqrt{2}\beta\sigma\sqrt{t}}\right) +$$
$$+ \frac{e^{ct} S^\mu y^{-\mu-1}}{\sqrt{2\pi}\sqrt{\sigma^2 t}}\left(b^{\frac{2\ln\frac{b}{y}+\ln S}{\sigma^2 t}} + S^{\frac{2\ln\frac{b}{y}+\ln b}{\sigma^2 t}}\right) \exp\left(-\frac{\left(\ln\frac{b}{y}+\ln S\right)^2 + 2\ln b\ln\frac{b}{y} + \ln^2 b}{2\sigma^2 t}\right),$$

where erfc is the complementary error function, defined as

$$\operatorname{erfc}(x) := 1 - \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}\mathrm{d}t$$

and provided that $be^{-\frac{\tilde{A}t\sigma^2}{\beta}} < S < be^{\frac{\tilde{A}t\sigma^2}{\beta}}$, in order to grant integrability. To sum up, we found a solution to problem (4.12), which involves error functions and a one-dimensional integral evaluation. There is a wide range of functions $f(S)$ for which we are actually able to give an explicit solution of problem (4.12), for example if $f(S)$ belongs to the class of the polynomials. In Figure 4.1, we plot the solution (4.29) of problem (4.12) for a particular choice of the parameters $r, \sigma^2, \alpha, \beta$ and b, in the case of $f(S) = S$.

4.5 Inhomogeneous Robin Condition

In this subsection we solve Problem (4.12) in the case of a generic function $g(t)$. We rewrite here the problem for the sake of ease

$$\begin{aligned} w_t &= \frac{1}{2}\sigma^2 S^2 w_{SS} + rSw_S, \quad S > b > 0 \\ w(S,0) &= f(S), \\ \alpha w(b,t) + \beta w_S(b,t) &= g(t). \end{aligned} \tag{4.31}$$

In order to attack Problem (4.31), we are going to follow a bottom-up approach and we present two methods to compute a solution. Both methods assume that the solution can be expressed in

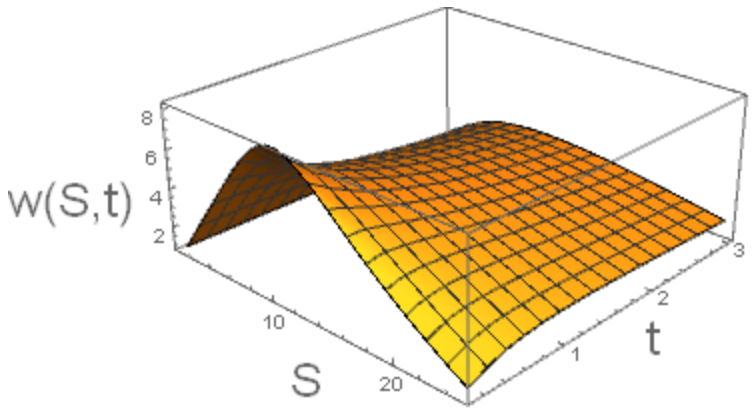


Figure 4.1: Plot of the solution $w(S,t)$, where S ranges between 1 and 25, t between 0.1 and 3. The parameters are settled as follows: $r = 0.01$, $\sigma = 0.75$, $b = 10$, $\alpha = 0$ and $\beta = 10$. This corresponds to a contract involving a Delta constraint when the underlying hits the lower barrier b.

terms of a sum of two functions, where the first one solves a slightly modified version of the homogeneous Robin problem that we have already solved, namely problem (4.13). More precisely, in Subsection 4.5.1 we describe the first method that provides a solution based on the inversion of a Laplace transform, while in Subsection 4.5.2, we describe a more algebraic-based approach starting from the Robin condition. As we are going to see, this latter method can only be exploited for a particular choice of the function $g(t)$. Eventually, in Subsection 4.5.3 we combine the methods of the previous two Subsection to generalise a well known result on the representation of solutions of Parabolic PDEs to Problem (4.31).

The starting point to tackle Problem (4.31) is to set $w(S,t) = v(S,t) + h(S,t)$, where we require that $v(S,t)$ satisfies the condition

$$\alpha v(b,t) + \beta v_S(b,t) = 0,$$

so that $h(S,t)$ must satisfy

$$\alpha h(b,t) + \beta h_S(b,t) = g(t). \tag{4.32}$$

The corresponding PDE for $v(S,t), h(S,t)$ reads as

$$v_t = \frac{1}{2}\sigma^2 S^2 v_{SS} + rSv_S + \frac{1}{2}\sigma^2 S^2 h_{SS} + rSh_S - h_t, \quad S > b > 0. \tag{4.33}$$

In order for $w(S,t)$ to satisfy the boundary condition at $(S,0)$, we need $v(S,t)$ (and $h(S,t)$) to satisfy

$$v(S,0) = f(S) - h(S,0).$$

Collecting all things together, we end up with the following problem in terms of the function $v(S,t)$

$$\begin{aligned} v_t &= \frac{1}{2}\sigma^2 S^2 v_{SS} + rSv_S + \frac{1}{2}\sigma^2 S^2 h_{SS} + rSh_S - h_t, \quad S > b > 0 \\ v(S,0) &= f(S) - h(S,0), \\ \alpha v(b,t) + \beta v_S(b,t) &= 0, \end{aligned} \tag{4.34}$$

where we require the function $h(S,t)$ to satisfy the constraint (4.32). If we are able to solve problem (4.34) for the function $v(S,t)$ (and $h(S,t)$) with the condition (4.32), then, by superposition, we are also able to solve problem (4.31) by simply setting $w(S,t) = v(S,t) + h(S,t)$, as stated in the next proposition whose obvious proof is omitted.

Proposition 4.5.1. *Let $(v(S,t), h(S,t))$ be functions that solve problem (4.34), with the condition (4.32). Then, by setting $w(S,t) = v(S,t) + h(S,t)$, the function $w(S,t)$ solves the problem (4.31).*

4.5.1 A Laplace-based approach

One possible way to proceed is to find a function $h(S,t)$ that satisfies the problem

$$\begin{aligned} h_t &= \frac{1}{2}\sigma^2 S^2 h_{SS} + rSh_S, \quad S > b > 0 \\ \alpha h(b,t) + \beta h_S(b,t) &= g(t), \end{aligned} \tag{4.35}$$

as suggested in Proposition 4.5.1. To accomplish this, we start with the solution of the PDE (4.17) that we rewrite here for the sake of ease,

$$h^\xi(S,t) = e^{\lambda t} S^\mu \left(A\cos(\xi \ln S) + B\sin(\xi \ln S)\right), \tag{4.36}$$

where λ and μ are defined as in (4.14) and (4.16), respectively, and A and B are arbitrary constants. We recall that function (4.36) solves the PDE in (4.35) for every choice of the constants A, B. We set then

$$h(S,t) = \int_0^\infty \psi(\xi) h^\xi(S,t) d\,\xi, \tag{4.37}$$

and, thanks to Lemma 4.4.1, if ψ is a function with sufficient decay, then (4.37) solves the PDE in (4.35) as well. Our goal is to find a suitable ψ that makes the function (4.37) to satisfy also the Robin condition in (4.35). In the next lines we are going to see that we can get the desired result by inverting a Laplace transform. After exploiting the Robin condition in (4.35) for the function (4.37), we get

$$\begin{aligned} g(t) = \alpha &\int_0^\infty b^\mu \psi(\xi) e^{-\frac{1}{2}\sigma^2 t\left(\mu^2+\xi^2\right)} \left(A\cos(\xi \ln b) + B\sin(\xi \ln b)\right)\, d\xi + \\ + \beta &\int_0^\infty b^{\mu-1} \psi(\xi) e^{-\frac{1}{2}\sigma^2 t\left(\mu^2+\xi^2\right)} \left((B\mu - A\xi)\sin(\xi \ln b) + (A\mu + B\xi)\cos(\xi \ln b)\right)\, d\xi \end{aligned}$$

that can be rewritten as

$$\begin{aligned} b^{1-\mu} g(t) e^{\frac{1}{2}t\sigma^2\mu^2} = &\int_0^\infty \psi(\xi) e^{-\frac{1}{2}t\sigma^2\xi^2} \left(\alpha b A + \beta(A\mu + B\xi)\right)\cos(\xi \ln b)\, d\xi + \\ &+ \int_0^\infty \psi(\xi) e^{-\frac{1}{2}t\sigma^2\xi^2} \left(\alpha b B + \beta(B\mu - A\xi)\right)\sin(\xi \ln b)\, d\xi. \end{aligned} \tag{4.38}$$

By choosing $A = -\beta\xi$ and $B = b\alpha + \beta\mu$ we simplify equation (4.38), namely

$$b^{1-\mu} g(t) e^{\frac{1}{2}t\sigma^2\mu^2} = \int_0^\infty \psi(\xi) e^{-\frac{1}{2}t\sigma^2\xi^2} \sin(\xi \ln b) \left(\beta^2\xi^2 + \tilde{A}^2\right)\, d\xi, \tag{4.39}$$

where we recall that $\tilde{A} = \alpha b + \beta\mu$.

If we let $s = \frac{1}{2}t\sigma^2$ and we use the change of variable $z = \xi^2$, then Equation (4.39) reduces to a Laplace transform, namely

$$b^{1-\mu} g\left(\frac{2s}{\sigma^2}\right) e^{s\mu^2} = \int_0^\infty \tilde{\psi}(\sqrt{z}) e^{-sz} \frac{dz}{\sigma\sqrt{z}},$$

where

$$\tilde{\psi}(\xi) := \psi(\xi) \sin(\xi \ln b) \left(\beta^2\xi^2 + \tilde{A}^2\right).$$

Hence, we can invert it to find

$$\tilde{\psi}(\sqrt{z}) = \sigma\sqrt{z} b^{1-\mu} W(z), \tag{4.40}$$

where $W(z) = \mathcal{L}^{-1}\left[g\left(\frac{2s}{\sigma^2}\right) e^{s\mu^2}; z\right]$ indicates the inverse Laplace transform of g with inverse transform variable z, meaning that $\mathcal{L}\left[W(z); s\right] = g\left(\frac{2s}{\sigma^2}\right) e^{s\mu^2}$. Relabelling the parameter we eventually get

$$\psi(\xi) = \frac{\sigma\xi b^{1-\mu} W(\xi^2)}{\sin\left(\xi \ln b\right)\left(\beta^2\xi^2 + \tilde{A}^2\right)}. \tag{4.41}$$

Hence, if we set

$$w(S,t) = v(S,t) + h(S,t),$$

where v is the solution of the homogeneous Robin problem, namely problem (4.13) with boundary condition $\tilde{f}(S) = f(S) - h(S,0)$ and where $h(S,t)$ is defined by (4.37) with ψ as in (4.41), then $w(S,t)$ solves problem (4.31). Therefore, the solution $w(S,t)$ can be written in the form

$$w(S,t) = \int_0^b \tilde{f}(y) p(S,y,t) dy + \int_0^\infty \psi(\xi) S^\mu e^{\lambda t} (\tilde{A}\sin(\xi\log(S)) - \beta\xi\cos(\xi\log(S)))\, d\xi, \tag{4.42}$$

where $\tilde{f}(y) = f(y) - h(y,0)$, $p(S,y,t)$ is defined as in (4.30) and ψ is defined as in (4.41).

Example 4.5.2. *For some choices of the function $g(t)$, we are actually able to compute explicitly the inverse Laplace transform in (4.41). Let $g(t)$ the function defined as $g(t) = e^t t^n$, where $n \in \mathbb{N}$. Then*

$$\mathcal{L}^{-1}\left[g\left(\frac{2s}{\sigma^2}\right) e^{s\mu^2}; z\right] = \frac{\left(\frac{2}{\sigma^2}\right)^n \theta\left(\mu^2 + z + \frac{2}{\sigma^2}\right)\left(\mu^2 + \frac{2}{\sigma^2} + z\right)^{-n-1}}{\Gamma(-n)},$$

where $\Gamma(\cdot)$ is the Gamma function, while $\theta(\cdot)$ is the Heaviside step function. Hence

$$\psi(\xi) = \frac{2^n \xi b^{1-\mu} \theta\left(\mu^2 + \xi^2 + \frac{2}{\sigma^2}\right)\left(\mu^2 + \frac{2}{\sigma^2} + \xi^2\right)^{-n-1}}{\sigma^{2n-1}\sin\left(\xi\ln b\right)\left(\beta^2\xi^2 + (\alpha b + \beta\mu)^2\right)\Gamma(-n)}. \tag{4.43}$$

Solution of problem (4.35) is then

$$h(S,t) = \int_0^\infty \psi(\xi) S^\mu e^{\lambda t} (\tilde{A}\sin(\xi\log(S)) - \beta\xi\cos(\xi\log(S)))\, d\xi \tag{4.44}$$

where ψ is defined as in (4.43), $\tilde{A} = \alpha b + \beta\mu$ and λ and μ as in (4.14) and (4.16), respectively. The solution of the problem (4.31) is then

$$w(S,t) = v(S,t) + h(S,t), \tag{4.45}$$

where $v(S,t)$ is the function defined as before and the function $h(S,t)$ is defined in (4.44) with ψ as in (4.43). Also in this case, the solution $w(S,t)$ can be numerically evaluated, as long as the boundary function $f(S)$ is chosen and the parameters $r, \sigma^2, \alpha, \beta$ and b are settled.

4.5.2 An algebraic-based approach

Another way to tackle the problem (4.31) is to first start from the Robin condition and then focus on the PDE. More precisely, we start by finding a (class of) functions $h(S,t)$ satisfying the Robin condition in (4.35). By algebraic manipulation, it is easy to see that such function $h(S,t)$ can be represented as

$$h(S,t) = \frac{1}{2}(S-b)^2 h_2(S,t) + \frac{g(t) - \beta h_1(t)}{\alpha} + (S-b)h_1(t), \tag{4.46}$$

with $h_2(S,t)$ and $h_1(t)$ arbitrary functions. Now we want to select the functions $h_2(S,t)$ and $h_1(t)$ such that the function (4.46) satisfies the PDE in (4.35). The PDE for $h(S,t)$ reads

$$\begin{aligned} &2\alpha S h_2(S,t)\left(S\left(2r+\sigma^2\right) - 2br\right) + 2\alpha S(b-S){h_2}^{(1,0)}(S,t)\left(br - S\left(r+2\sigma^2\right)\right) + \\ &\alpha\sigma^2 S^2 (b-S)^2 {h_2}^{(2,0)}(S,t) + 4h_1'(t)(\alpha b + \beta - \alpha S) - 2\alpha(b-S)^2 {h_2}^{(0,1)}(S,t) + \\ &+ 4\alpha r S h_1(t) = 4g'(t), \end{aligned} \tag{4.47}$$

where ${h_2}^{(0,1)}(S,t) = \frac{\partial h_2(S,t)}{\partial t}$, ${h_2}^{(1,0)}(S,t) = \frac{\partial h_2(S,t)}{\partial S}$, ${h_2}^{(2,0)}(S,t) = \frac{\partial h_2(S,t)}{\partial^2 S}$.

The PDE (4.47) looks more complicate than the one in (4.35), but we have to keep in mind that we have freedom on the functions $h_2(S,t)$ and $h_1(t)$. One simple choice is setting $g'(t) = c$, $c \in \mathbb{R}$ and $h_1(t) = 0$, which allows us to reduce the PDE (4.47) to an ODE. In this case (we consider now $c = 0$ for sake of ease) we find

$$h_1(t) = 0 \tag{4.48}$$

$$h_2(S,t) = h_2(S) = \frac{x^{-\frac{|\sigma^2-2r|+2r-\sigma^2}{2\sigma^2}}\left(c_1|\sigma^2-2r| + c_2\sigma^2 S^{\frac{|\sigma^2-2r|}{\sigma^2}}\right)}{(b-S)^2|\sigma^2-2r|}, \tag{4.49}$$

where $c_1, c_2 \in \mathbb{R}$.

With this choice it is easy to check that the function $h(S,t)$ in (4.46) satisfies both the PDE (4.47) and the Robin condition in (4.35). The solution of the problem (4.31) is then

$$w(S,t) = v(S,t) + h(S,t), \tag{4.50}$$

where $v(S,t)$ is again the solution of the homogeneous Robin problem, namely problem (4.13) with boundary condition $\tilde{f}(S) = f(S) - h(S,0)$.

4.5.3 Representations of Solutions of Parabolic PDEs

In this subsection we provide a general result that gives a representation for the solution of our problem. The procedure we are going to present can be applied to problems which involve any positive second order operator.

We study the following problem:

$$\begin{cases} u_t(S,t) = Lu(S,t) \\ u(S,0) = f(S) \\ \alpha u(b,t) + \beta u_S(b,t) = g(t), \end{cases} \tag{4.51}$$

with $S > 0, t > 0, b > 0$ and where L is a positive second order operator in S. In our case, L is the second order Black-Scholes operator, namely

$$Lu(S,t) = \frac{1}{2}\sigma^2 S^2 u_{SS}(S,t) + rSu_S(S,t).$$

In order to solve Problem (4.51), we set $u(S,t) = v(S,t) + h(S,t)$, where

$$\alpha v(b,t) + \beta v_S(b,t) = 0,$$

and

$$\alpha h(b,t) + \beta h_S(b,t) = g(t). \tag{4.52}$$

This gives

$$\begin{cases} v_t(S,t) = Lv(S,t) + K(S,t) \\ v(S,0) = f(S) - h(S,0) \\ \alpha v(b,t) + \beta v_S(b,t) = g(t), \end{cases} \tag{4.53}$$

where $K(S,t) = Lh - h_t(S,t)$ and $S > 0, t > 0$. We assume $g(t)$ to be differentiable and integrable. Let $q(x,y,t)$ be a fundamental solution of the problem $u_t = Lu$, $\alpha u(b,t) + \beta u_S(b,t) = 0$, that is $q_t(S,y,t) = Lq(S,y,t)$ and $\alpha q(b,y,t) + \beta q_S(b,y,t) = 0$.

Proposition 4.5.3. *Let $h(S,t)$ be a function that satisfies (4.52) and let $q(x,y,t)$ a fundamental solution of $q_t = Lq$, such that $\alpha q(b,y,t) + \beta q_x(b,y,t) = 0$. Then the solution of the boundary value problem (4.51) can be written as*

$$\begin{aligned} u(S,t) =& h(S,t) + \int_0^\infty (f(y) - h(y,0))q(S,y,t)dy \\ &+ \int_0^t \int_0^\infty \left(Lh(y,\tau) - \frac{\partial h}{\partial t}(y,\tau) \right) q(S,y,t-\tau)dyd\tau. \end{aligned}$$

In general, the function $h(S,t)$ is not unique, as we have seen in Subsection 4.5.2, and different choices of it may lead to different representations of the solution to Problem (4.51).

Proof. We begin by noticing that $q(S,y,0)=\delta(S-y)$, with $\delta(\cdot)$ being the Dirac function, since $q(S,y,t)$ is a fundamental solution. Hence

$$\begin{aligned}u(S,0)=&h(S,0)+\int_0^\infty (f(y)-h(y,0))\delta(S-y)dy\\&+\int_0^0\int_0^\infty \left(Lh(y,\tau)-\frac{\partial h}{\partial t}(y,\tau)\right)q(S,y,0-\tau)dyd\tau\\=&h(S,0)+f(S)-h(S,0)=f(S).\end{aligned}$$

Moreover, we have

$$\begin{aligned}u(b,t)=&h(b,t)+\int_0^\infty (f(y)-h(y,0))q(b,y,t)dy\\&+\int_0^t\int_0^\infty \left(Lh(y,\tau)-\frac{\partial h}{\partial t}(y,\tau)\right)q(b,y,t-\tau)dyd\tau\end{aligned}$$

and

$$\begin{aligned}u_S(b,t)=&h_S(b,t)+\int_0^\infty (f(y)-h(y,0))q_S(b,y,t)dy\\&+\int_0^t\int_0^\infty \left(Lh(y,\tau)-\frac{\partial h}{\partial t}(y,\tau)\right)q_S(b,y,t-\tau)dyd\tau,\end{aligned}$$

then

$$\begin{aligned}\alpha u(b,t)+\beta u_S(b,t)=&\alpha h(b,t)+\beta h_S(b,t)+\\&+\int_0^\infty (f(y)-h(y,0))(\alpha q(b,y,t)+\beta q_S(b,y,t))dy+\\&+\int_0^t\int_0^\infty \left(Lh(y,\tau)-\frac{\partial h}{\partial t}(y,\tau)\right)(\alpha q(b,y,t-\tau)+\beta q_S(b,y,t-\tau))dyd\tau\\=&g(t),\end{aligned}$$

thanks to the previous assumptions on $h(S,t)$ and $q(S,y,t)$. Hence, the Robin condition is satisfied as well. We next check that u satisfies the PDE. We assume that we have sufficient regularity to allow us to differentiate under the integral sign, then

$$\begin{aligned}u_t=&h_t(S,t)+\int_0^\infty (f(y)-h(y,0))q_t(S,y,t)dy\\&+\int_0^t\int_0^\infty \left(Lh(y,\tau)-\frac{\partial h}{\partial t}(y,\tau)\right)q_t(S,y,t-\tau)dyd\tau\\&+\int_0^\infty \left(Lh(y,t)-\frac{\partial h}{\partial t}(y,t)\right)\delta(S-y)dy\\=&h_t(S,t)+\int_0^\infty (f(y)-h(y,0))q_t(S,y,t)dy+Lh(S,t)-h_t(S,t)\\&+\int_0^t\int_0^\infty \left(Lh(y,\tau)-\frac{\partial h}{\partial t}(y,\tau)\right)q_t(S,y,t-\tau)dyd\tau,\end{aligned}$$

so that

$$\begin{aligned} u_t =& Lh(S,t) + \int_0^\infty (f(y) - h(y,0)) Lq(S,y,t)dy \\ &+ \int_0^t \int_0^\infty \left(Lh(y,\tau) - \frac{\partial h}{\partial t}(y,\tau) \right) Lq(S,y,t-\tau)dyd\tau = Lu, \end{aligned}$$

where we used the fact that $Lq = q_t$. This concludes the proof. □

We can now write the solution, following the characterization of Proposition (4.5.3). It takes the form

$$\begin{aligned} w(S,t) =& h(S,t) + \int_0^\infty (f(y) - h(y,0)) q(S,y,t)dy \\ &+ \int_0^t \int_0^\infty \left(Lh(y,\tau) - \frac{\partial h}{\partial t}(y,\tau) \right) q(S,y,t-\tau)dyd\tau, \end{aligned}$$

where the function $h(S,t)$ can be chosen from (4.46) and $p(S,y,t)$ is defined as in (4.30).

Example 4.5.4. *If we take* $h(S,t) = \frac{g(t)}{\alpha}$*, namely* $h_1(t) = 0$ *and* $h_2(S,t) = 0$ *in* (4.46)*, then the solution becomes*

$$w(S,t) = \frac{g(t)}{\alpha} + \int_0^\infty \left(f(y) - \frac{g(0)}{\alpha} \right) q(S,y,t)dy - \int_0^t \int_0^\infty \frac{g'(\tau)}{\alpha} q(S,y,t-\tau)dyd\tau,$$

where, again, $p(S,y,t)$ *is defined as in* (4.30)*.*

4.6 Conclusion

In this chapter we have dealt with a boundary pricing problem in the Black-Scholes framework. More precisely, we studied the option pricing PDE with a boundary value of Robin type, which involves also the derivative of the value function. This value is related to a boundary condition involving the Delta of the option. We have first tackled the problem in the case of an homogeneous Robin condition, where we were able to characterize the solution in terms of its transition density. A numerical representation of the solution has been presented, for a given parameters choice. Then we have switched to the general case, namely the inhomogeneous one, where we have started with two tune-up methods before proving the more general result on the representation of the solution of our problem, that somehow includes and extends the previous cases. We leave for future research the interesting task of incorporating the study of other types of boundary conditions, for example double barrier contracts, in the spirit of Skachkov (2002), that could be a natural extension of this work.

Appendix A

Derivation of the symmetries

In the next Proposition, we explicitly give the complete set of symmmetries of PDE (A.1).

Proposition 1. *The PDE*

$$u_t = \frac{1}{2}\sigma^2 x^2 u_{xx} + rxu_x, \quad x \in \mathbb{R}, \tag{A.1}$$

has the following symmetries

$$\begin{aligned}
&U_\varepsilon^{(1)}(x,t) = u(x, t-\varepsilon),\\
&U_\varepsilon^{(2)}(x,t) = u\left(e^{-\varepsilon}x, t\right),\\
&U_\varepsilon^{(3)}(x,t) = e^{\varepsilon}u(x,t),\\
&U_\varepsilon^{(4)}(x,t) = \exp\left(\frac{\varepsilon\left(t\left(-2r+\sigma^2+2\varepsilon\right)-2\log(x)\right)}{\sigma^2}\right)u\left(xe^{-2t\varepsilon}, t\right),\\
&U_\varepsilon^{(5)}(x,t) = \exp\left(-\frac{e^{-8\varepsilon}\left(e^{4\varepsilon}-1\right)\left(\sigma^2-2r\right)\left(-t\left(e^{4\varepsilon}+1\right)\left(2r-\sigma^2\right)-4e^{4\varepsilon}\log(x)\right)}{8\sigma^2}\right)u\left(x^{e^{-4\varepsilon}}, te^{-8\varepsilon}\right),\\
&U_\varepsilon^{(6)}(x,t) = \frac{1}{\sqrt{1+8t\varepsilon}}\exp\left(-\frac{\varepsilon\left(2rt-\sigma^2+2\ln(x)\right)^2}{\sigma^2(1+8t\varepsilon)}\right)u\left(x^{1/1+8t\varepsilon}, \frac{t}{1+8t\varepsilon}\right),
\end{aligned} \tag{A.2}$$

that is, for ε sufficiently small, $U_\varepsilon^{(i)}(x,t)$, $i = 1\dots6$ is a solution of the PDE (A.1) whenever $u(x,t)$ is.

Proof. The PDE (A.1) admits a Lie symmetry group whose finite dimensional part has dimension 6. For sake of brevity, we skip the computations to find the Lie algebra, since they basically mimic the ones we have done in Subsection 4.2.1 for the heat equation.

The corresponding Lie algebra for (A.1) is generated by the following infinitesimal symmetries:

$$
\begin{aligned}
\mathbf{v}_1 &= \partial_t \\
\mathbf{v}_2 &= x\partial_x \\
\mathbf{v}_3 &= u\partial_u \\
\mathbf{v}_4 &= 2tx\partial_x + \left(t - \frac{2rt}{\sigma^2} - \frac{2\ln(x)}{\sigma^2}\right) u\partial_u \\
\mathbf{v}_5 &= 8t\partial_t + (4x\ln(x))\,\partial_x + \left(-\frac{(\sigma^2-2r)\,(t\,(\sigma^2-2r) - 2\ln(x))}{\sigma^2}\right) u\partial_u \\
\mathbf{v}_5 &= 8t\partial_t + (4x\ln(x))\,\partial_x + \left(-\frac{4r^2t}{\sigma^2} + 4rt - \frac{4r\log(x)}{\sigma^2} - \sigma^2 t + 2\log(x)\right) u\partial_u \\
\mathbf{v}_6 &= 8t^2\partial_t + (8tx\ln(x))\,\partial_x + \\
&\quad + \left(-\frac{4r^2t^2}{\sigma^2} + 4rt^2 - \frac{8rt\ln(x)}{\sigma^2} - \sigma^2t^2 + 4t\ln(x) - 4t - \frac{4\ln^2(x)}{\sigma^2}\right) u\partial_u
\end{aligned}
\tag{A.3}
$$

All other symmetries of (A.1) have infinitesimal generators of the form $\mathbf{v}_\beta = \beta\partial_u$, where β is an arbitrary solution of the PDE. These simmetries correspond to the superposition of solutions, i.e. $u \longrightarrow u + \varepsilon\beta$ is a solution whenever u and β are solutions. We exponentiate, for example, the symmetry $\mathbf{v}_6$, namely we are looking for the solution to the following system of ODE's:

$$\frac{d\tilde{t}}{d\epsilon} = 8\tilde{t}^2, \quad \tilde{t}(0) = t \tag{A.4}$$

$$\frac{d\tilde{x}}{d\epsilon} = 8\tilde{t}\tilde{x}\ln(\tilde{x}), \quad \tilde{x}(0) = x \tag{A.5}$$

$$\frac{d\tilde{u}}{d\epsilon} = \left(-\frac{4r^2\tilde{t}^2}{\sigma^2} + 4r\tilde{t}^2 - \frac{8r\tilde{t}\ln(\tilde{x})}{\sigma^2} - \sigma^2\tilde{t}^2 + 4\tilde{t}\ln(\tilde{x}) - 4\tilde{t} - \frac{4\ln^2(\tilde{x})}{\sigma^2}\right)\tilde{u}, \quad \tilde{u}(0) = u. \tag{A.6}$$

From Equation (A.4) we immediately get

$$\tilde{t}(\varepsilon) = \frac{t}{1 - 8t\varepsilon},$$

so Equation (A.5) becomes

$$\frac{d\tilde{x}}{d\epsilon} = \frac{8t}{1-8t\varepsilon}\tilde{x}\ln(\tilde{x}), \quad \tilde{x}(0) = x,$$

that can be rewritten as

$$\frac{d\tilde{x}}{\tilde{x}\ln(\tilde{x})} = \frac{8t}{1-8t\varepsilon}d\epsilon, \quad \tilde{x}(0) = x,$$

and noting that $\frac{d}{dz}\left(\ln\left(\ln(z)\right)\right) = \frac{1}{z\ln z}$ we immediately get

$$\tilde{x}(\varepsilon) = x^{1/1-8t\varepsilon}.$$

Equation (A.6) then becomes

$$\frac{d\tilde{u}}{\tilde{u}} = \left(-\frac{4r^2t^2}{\sigma^2(1-8t\varepsilon)^2} + \frac{4rt^2}{(1-8t\varepsilon)^2} - \frac{8rt\ln(x^{\frac{1}{1-8t\varepsilon}})}{\sigma^2(1-8t\varepsilon)} - \frac{\sigma^2t^2}{(1-8t\varepsilon)^2} \right) d\varepsilon +$$
$$+ \left(\frac{4t\ln(x^{\frac{1}{1-8t\varepsilon}})}{1-8t\varepsilon} - \frac{4t}{1-8t\varepsilon} - \frac{4\ln^2(x^{\frac{1}{1-8t\varepsilon}})}{\sigma^2} \right) d\varepsilon, \quad \tilde{u}(0) = u,$$

whose solution is given by

$$\tilde{u} = u\sqrt{1-8t\varepsilon}\exp\left(\varepsilon \frac{(2rt - \sigma^2t + 2\log(x))^2}{\sigma^2(8t\varepsilon - 1)} \right), \tag{A.7}$$

with $u = u(x,t)$.
Relabeling the parameters using the inverse of $(\tilde{t}, \tilde{x})$ completes the proof. □

Bibliography

Alexander, C. and Korovilas, D. (2013). Volatility exchange-traded notes: curse or cure? *The Journal of Alternative Investments*, 16(2):52–70.

Andersen, L. (2011). Option pricing with quadratic volatility: a revisit. *Finance and Stochastics*, 15(2):191–219.

Anderson, N. (1996). *Estimating and interpreting the yield curve*, volume 6. John Wiley & Sons Incorporated.

Baldeaux, J. and Platen, E. (2013). *Functionals of multidimensional diffusions with applications to finance*, volume 5. Springer Science & Business Media.

Bao, Q., Li, S., and Gong, D. (2012). Pricing vxx option with default risk and positive volatility skew. *European Journal of Operational Research*, 223(1):246–255.

Benth, F. E., Benth, J. S., and Koekebakker, S. (2008). *Stochastic modelling of electricity and related markets*, volume 11. World Scientific.

Benth, F. E., Koekkebakker, S., and Ollmar, F. (2007). Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation. *The Journal of Derivatives*, 15(1):52–66.

Benth, F. E. and Paraschiv, F. (2018). A space-time random field model for electricity forward prices. *Journal of Banking & Finance*, 95(10):203–216.

Benth, F. E., Piccirilli, M., and Vargiolu, T. (2019). Mean-reverting additive energy forward curves in a heath–jarrow–morton framework. *Mathematics and Financial Economics*, 13(4):543–577.

Bergomi, L. (2015). *Stochastic volatility modeling*. CRC press.

Bernis, G., Brignone, R., Scotti, S., and Sgarra, C. (2020). A gamma ornstein-uhlenbeck model driven by a hawkes process. *Available at SSRN 3370304*.

Bernis, G., Salhi, K., and Scotti, S. (2018). Sensitivity analysis for marked hawkes processes: application to clo pricing. *Mathematics and Financial Economics*, 12(4):541–559.

Bluman, G. and Kumei, S. (1989). *Symmetry and Differential Equations*. Springer, New York.

Callegaro, G., Gaïgi, M., Scotti, S., and Sgarra, C. (2017). Optimal investment in markets with over and under-reaction to information. *Mathematics and Financial Economics*, 11(3):299–322.

Callegaro, G., Mazzoran, A., and Sgarra, C. (2019). A self-exciting modelling framework for forward prices in power markets. *arXiv preprint arXiv:1910.13286*.

Carr, P., Fisher, T., and Ruf, J. (2013). Why are quadratic normal volatility models analytically tractable? *SIAM Journal on Financial Mathematics*, 4(1):185–202.

Carr, P., Laurence, P., and Wang, T.-H. (2006). Generating integrable one dimensional driftless diffusions. *Comptes Rendus Mathematique*, 343(6):393–398.

Christensen, T., Hurn, S., and Lindsay, K. (2009). It never rains but it pours: modeling the persistence of spikes in electricity prices. *The Energy Journal*, 30(1):25–48.

Clements, A., Fuller, J., and Hurn, S. (2013). Semi-parametric forecasting of spikes in electricity prices. *Economic Record*, 89(287):508–521.

Cont, R. and Kokholm, T. (2013). A consistent pricing model for index options and volatility derivatives. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 23(2):248–274.

Cordoni, F. and Di Persio, L. (2014). Lie symmetry approach to the cev model. *International Journal of Differential Equations and Applications*, 13(3):109–127.

Craddock, M. (2009). Fundamental Solutions, Transition Densities and the Integration of Lie Symmetries. *Journal of Differential Equations*, 246(1):2538–2560.

Craddock, M. (2014). On an integral arising in mathematical finance. In *Nonlinear Economic Dynamics and Financial Modelling*, pages 355–370. Springer.

Craddock, M. (2019a). Integral transform and Lie symmetry methods for scalar diffusions. *Submitted*.

Craddock, M. (2019b). The Solution of Absorption, Reflection and Robin Problems for Diffusions via Symmetries and Integral Equations. *Submitted*.

Craddock, M. and Dooley, A. (2010). On the Equivalence of Lie Symmetries and Group Representations. *Journal of Differential Equations*, 249(1):621–653.

Craddock, M. and Grasselli, M. (2019). Lie Symmetry Methods for Local Volatility Models. *Stochastic Processes and Their Applications, to appear*.

Craddock, M. and Lennox, K. (2007). Lie group symmetries as integral transforms of fundamental solutions. *Journal of Differential Equations*, 232(2):652–674.

Craddock, M. and Lennox, K. (2009). The Calculation of Expectations for Classes of Diffusion Processes by Lie Symmetry Methods. *Annals of Applied Probability*, 19(1):127–157.

Craddock, M. and Lennox, K. (2012). Lie Symmetry Methods for Multidimensional Parabolic PDEs and Diffusions. *Journal of Differential Equations*, 252(1):56–90.

Craddock, M. and Platen, E. (2004). Symmetry Group Methods for Fundamental Solutions. *Journal of Differential Equations*, 207(2):285–302.

Dawson, D. A., Li, Z., et al. (2012). Stochastic equations, flows and measure-valued processes. *The Annals of Probability*, 40(2):813–857.

De Marco, S. and Henry-Labordere, P. (2015). Linking vanillas and vix options: a constrained martingale optimal transport problem. *SIAM Journal on Financial Mathematics*, 6(1):1171–1194.

Derman, E. and Kani, I. (1994). Riding on a smile. *Risk*, 7(2):32–39.

Drimus, G. and Farkas, W. (2013). Local volatility of volatility for the VIX market. *Review of Derivatives Research*, 3(16):267–293.

Dupire, B. (1996). A unified theory of volatility. *Derivatives pricing: The classic collection*, pages 185–196.

Dupire, B. et al. (1994). Pricing with a smile. *Risk*, 7(1):18–20.

Errais, E., Giesecke, K., and Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics*, 1(1):642–665.

Eyjolfsson, H. and Tjøstheim, D. (2018). Self-exciting jump processes with applications to energy markets. *Annals of the Institute of Statistical Mathematics*, 70(2):373–393.

Facchinetti, S. (2009). A procedure to find exact critical values of kolmogorov-smirnov test. *Statistica Applicata*, 21:337–359.

Filimonov, V., Bicchetti, D., Maystre, N., and Sornette, D. (2014). Quantification of the high level of endogeneity and of structural regime shifts in commodity markets. *Journal of international Money and finance*, 42:174–192.

Filipović, D. (2001). A general characterization of one factor affine term structure models. *Finance and Stochastics*, 5(3):389–412.

Fleten, S.-E. and Lemming, J. (2003). Constructing forward price curves in electricity markets. *Energy Economics*, 25(5):409–424.

Gatheral, J. (2011). *The volatility surface: a practitioner's guide*, volume 357. John Wiley & Sons.

Gatheral, J. and Jacquier, A. (2014). Arbitrage-free svi volatility surfaces. *Quantitative Finance*, 14(1):59–71.

Gatheral, J., Jusselin, P., and Rosenbaum, M. (2020). The quadratic rough heston model and the joint s&p 500/vix smile calibration problem. *arXiv preprint arXiv:2001.01789*.

Gehricke, S. A. and Zhang, J. E. (2018). Modeling vxx. *Journal of Futures Markets*, 38(8):958–976.

Gehricke, S. A. and Zhang, J. E. (2020). The implied volatility smirk in the vxx options market. *Applied Economics*, 52(8):769–788.

Goard, J. (2000). New solutions to the bond-pricing equation via lie's classical method. *Mathematical and Computer Modelling*, 32(3-4):299–313.

Goard, J. (2011). Pricing of volatility derivatives using 3/2-stochastic models. *Proceedings of the World Congress on Engineering 2011*, 1:443–438.

Goard, J. and Mazur, M. (2013). Stochastic volatility models and the pricing of VIX options. *Mathematical Finance*, 23(3):139–158.

Grasselli, M. (2017). The 4/2 stochastic volatility model: a unified approach for the heston and the 3/2 model. *Mathematical Finance*, 27(4):1013–1034.

Grasselli, M., Mazzoran, A., and Pallavicini, A. (2020). A general framework for a joint calibration of vix and vxx options. *arXiv preprint arXiv:2012.08353*.

Grasselli, M. and Wagalath, L. (2018). Vix vs vxx: A joint analytical framework. *International Journal of Theoretical and Applied Finance, to appear*.

Guyon, J. (2017). Calibration of local correlation models to basket smiles. *Journal of Computational Finance*, 21(1):1–51.

Guyon, J. (2020a). Inversion of convex ordering in the vix market. *Quantitative Finance*, 20(10):1597–1623.

Guyon, J. (2020b). The joint s&p 500/vix smile calibration puzzle solved. *Risk, April*.

Guyon, J. and Henry-Labordère, P. (2012). Being particular about calibration. *Risk*, 25(1):88–93.

Guyon, J. and Henry-Labordere, P. (2013). *Nonlinear option pricing*. CRC Press.

Gyöngy, I. (1986). Mimicking the one-dimensional marginal distributions of processes having an itô differential. *Probability theory and related fields*, 71(4):501–516.

Hainaut, D. and Moraux, F. (2019). A switching self-exciting jump diffusion process for stock prices. *Annals of Finance*, 15(2):267–306.

Heath, D., Jarrow, R., and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society*, 60(1):77–105.

Herrera, R. and González, N. (2014). The modeling and forecasting of extreme events in electricity spot markets. *International Journal of Forecasting*, 30(3):477–490.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The review of financial studies*, 6(2):327–343.

Itkin, A. (2013). New solvable stochastic volatility models for pricing volatility derivatives. *Review of Derivatives Research*, 16(2):111–134.

Jiao, Y., Ma, C., Scotti, S., and Sgarra, C. (2019). A branching process approach to power markets. *Energy Economics*, 79:144–156.

Kawazu, K. and Watanabe, S. (1971). Branching processes with immigration and related limit theorems. *Theory of Probability & Its Applications*, 16(1):36–54.

Kiesel, R. and Paraschiv, F. (2017). Econometric analysis of 15-minute intraday electricity prices. *Energy Economics*, 64:77–90.

Kiesel, R., Paraschiv, F., and Sætherø, A. (2019). On the construction of hourly price forward curves for electricity prices. *Computational Management Science*, 16(1-2):345–369.

Lallouache, M. and Challet, D. (2016). The limits of statistical significance of hawkes processes fitted to financial data. *Quantitative Finance*, 16(1):1–11.

Lennox, K. (2011). Lie Symmetry Methods for Multi-dimensional Linear Parabolic PDEs and Diffusions. *PhD Thesis, UTS, Sydney*.

Li, Z. and Ma, C. (2015). Asymptotic properties of estimators in a stable cox–ingersoll–ross model. *Stochastic Processes and their Applications*, 125(8):3196–3233.

Lie, S. (1912). *Vorlesungen über Differentialgleichungen mit bekannten infinitesimalen Tranformatioen*. Teubner.

Lin, Y.-N. (2013). Vix option pricing and cboe vix term structure: A new methodology for volatility derivatives valuation. *Journal of Banking & Finance*, 37(11):4432–4446.

Lipton, A. (2002). The vol smile problem. *Risk Magazine*, 15:61–65.

Nastasi, E., Pallavicini, A., and Sartorelli, G. (2020). Smile modeling in commodity markets. *International Journal of Theoretical and Applied Finance*, 23(3):1–28.

Olver, P. (1978). Symmetry groups and conservation laws in the formal variational calculus. *Working paper, University of Oxford*.

Olver, P. J. (1993). *Applications of Lie Groups to Differential Equations*. Graduate Texts in Mathematics, Springer, New York.

Ozaki, T. (1979). Maximum likelihood estimation of hawkes' self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31(1):145–155.

Pagan, A. R. and Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of econometrics*, 45(1-2):267–290.

Paraschiv, F. (2013). Price dynamics in electricity markets. In *Handbook of Risk Management in Energy Production and Trading*, pages 47–69. Springer.

Pardoux, É. (2016). Probabilistic models of population evolution. *Mathematical Biosciences Institute Lecture Series. Stochastics in Biological Systems. Springer, Berlin*.

Rambaldi, M., Pennesi, P., and Lillo, F. (2015). Modeling foreign exchange market activity around macroeconomic news: Hawkes-process approach. *Physical Review E*, 91(1):012819.

Rebonato, R. (1999). *Volatility and correlation: In the pricing of equity, FX and interest-rate options*. John Wiley & Sons.

Ren, Y., Madan, D., and Qian Qian, M. (2007). Calibrating and pricing with embedded local volatility models. *Risk Magazine*, 20(9):138–143.

Schwartz, E. S. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of finance*, 52(3):923–973.

Schwert, G. W. (1990). Stock returns and real activity: A century of evidence. *The Journal of Finance*, 45(4):1237–1257.

Schwert, G. W. (2011). Stock volatility during the recent financial crisis. *European Financial Management*, 17(5):789–805.

Skachkov, I. (2002). Black-scholes equation in laplace transform domain. *Wilmott Technical Article (online) March*.

Sussman, A. and Morgan, C. (2012). Risk measurement on demand: Complexity, volatility and regulatory uncertainty. *TABB Group Research*.

Van der Stoep, A. W., Grzelak, L. A., and Oosterlee, C. W. (2014). The heston stochastic-local volatility model: efficient monte carlo simulation. *International Journal of Theoretical and Applied Finance*, 17(7):1450045.

Walsh, J. B. (1986). An introduction to stochastic partial differential equations. In *École d'Été de Probabilités de Saint Flour XIV-1984*, pages 265–439. Springer.

Whaley, R. E. (1993). Derivatives on market volatility: Hedging tools long overdue. *The journal of Derivatives*, 1(1):71–84.

Whaley, R. E. (2009). Understanding the vix. *The Journal of Portfolio Management*, 35(3):98–105.

Zuhlsdorff, C. (2002). Extended Libor Market Models with Affine and Quadratic Volatility. *Working Paper, University of Bonn*.