University of Nottingham

School of Mathematical Sciences

Topics in financial and computational mathematics

Simon Maurer, MSc.

A thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

August 2021

Abstract

In this thesis, we consider two different aspects in financial option pricing.

In the first part, we consider stochastic differential equations driven by general Lévy processes (SDEs) with finite and infinite activity and the related, via the Feynman-Kac formula, Dirichlet problem for integro-partial differential equation (IPDE). We approximate the solution of IPDE using a numerical method for the SDEs. The method is based on three ingredients: (i) we approximate small jumps by a diffusion; (ii) we use restricted jump-adaptive time-stepping; and (iii) between the jumps we exploit a weak Euler approximation. We prove weak convergence of the considered algorithm and present an in-depth analysis of how its error and computational cost depend on the jump activity level. We present the results of a range of numerical experiments including application of the suggested numerical scheme in the context of Foreign Exchange (FX) options, where we present an example on barrier basket currency option pricing in a multi-dimensional setting.

In the second part of the thesis, we suggest an intermediate currency approach that allows us to price options on all FX markets simultaneously under the same risk-neutral measure which ensures consistency of FX option prices across all markets. In particular, it is sufficient to calibrate a model to the volatility smile on the domestic market as, due to the consistency of pricing formulas, the model automatically reproduces the correct smile for the inverse pair (the foreign market). We first consider the case of two currencies and then the multi-currency setting. We illustrate the intermediate currency approach by applying it to the Heston and SABR stochastic volatility models, to the model in which exchange rates are described by an extended skewed normal distribution, and also to the model-free approach of option pricing.

Acknowledgements

I would like to thank all of those who have supported and encouraged me throughout my research studies.

Firstly, I would like to thank the collaborators of my research projects. I would like to thank Professor George Deligiannidis from the University of Oxford for the interesting discussions and suggestions. I thank Dr. Tim Sharp from Credit Suisse for providing an interesting topic to work on and providing practical financial market insights.

Secondly, I would like to thank my supervisor Prof. Michael Tretyakov for his constant advice, supportive feedback, professional guidance, challenging questions and generous support throughout my time at the University of Nottingham. I am very grateful for the discussions we have had over the years and hope to stay in touch in the future.

I would also like to thank all my friends, all the postgraduates students and academic staff, that I have met during my time in Nottingham, who have made this a truly great experience. I also acknowledge that I have been financially supported through a scholarship awarded by the School of Mathematics at the University of Nottingham.

Finally, I would like to thank my whole family and in particular, my girlfriend Valentina, who have supported me continuously throughout this journey and unquestionably helped me to finish this thesis. This accomplishment would not have been possible without them.

Publications

The following publications are the results of this research:

- S. Maurer, T. E. Sharp, and M. V. Tretyakov. Pricing FX Options under Intermediate Currency. https://arxiv.org/abs/1912.01387, 2021
- G. Deligiannidis, S. Maurer, and M. V. Tretyakov. Random walk algorithm for the Dirichlet problem for parabolic integro-differential equation. *BIT Numerical Mathematics*, April 2021. https://doi.org/10.1007/s10543-021-00863-2

CONTENTS

INTRODUCTION

1

AIMS OF THIS THESIS

One important aim of mathematical research is to find suitable models which are able to describe and capture observations made in the real world and try to explain and model this observed behaviour. The financial markets are a particular large, quickly changing and exciting field of interest for mathematical research, and Mathematics and Technology play an increasingly important role. Especially, in a world where interactions between countries, economies and currencies become more connected and interdependent, the interactions can be significant and need to be modelled adequately.

In this thesis we consider two different aspects of financial option pricing. The first part of this thesis focuses on the computational aspect and it is dedicated to a new numerical method for SDEs driven by Lévy processes with finite and infinite activity. The introduced restricted jump-adaptive time-stepping scheme to solve Dirichlet IPDE problems is analysed in full depth for different cases of jump activity. In particular, the theoretical convergence behaviour of the numerical scheme in the case of infinite activity is discussed in detail and a wide range of numerical examples are presented. We end this part by demonstrating the use of the introduced numerical scheme and we apply it to estimate the price of a foreign exchange (FX) barrier basket option involving five different currencies, where the underlying exchange rates are modelled by exponential Lévy processes.

In the second part of the thesis, we present a novel framework for pricing derivatives on the FX market. We explore a very simple but practically valuable approach, where we focus on finding a numeraire with respect to which we can price all FX derivatives traded on any of the domestic markets simultaneously under the same measure. Thanks to this approach, models for different currency pairs can be calibrated to all volatility smiles in a consistent manner. For example, in the case of two currencies, it is sufficient to calibrate a model for the GBPEUR exchange rate on e.g. the GBP domestic market and the smile on the EUR domestic market is automatically reproduced without any need of additional calibration, whereas following traditional approaches, this is not always the case. Then, we extend this methodology to the multi-dimensional setting and explore different pricing models. We end this part of the thesis with a range of numerical examples, where we calibrate some of the considered models to real market data in two and three dimensions.

Part I

NEW NUMERICAL METHODS FOR SDES DRIVEN BY LÉVY PROCESSES WITH INFINITE ACTIVITY

2

OVERVIEW

In this Part of the thesis, we present a new restricted jump-adaptive time-stepping scheme to solve Dirichlet IPDE problems with underlying SDEs driven by finite and infinite Lévy processes. The following chapters are based on the paper [30].

Stochastic differential equations (SDEs) are used to model various phenomena in different fields such as Biology [2], Physics [9, 23, 28, 95, 111] and Finance [1, 15, 55, 104, 107]. Typically, SDEs contain some sort of noise, which is driven by a stochastic process. The most commonly used noise is Brownian motion, which is one of the best known Lévy processes and hence widely studied [39, 106].

This research considers SDEs driven by a more general class of Lévy processes [4, 12, 105] and we are interested in including jump processes as noise, therefore we make use of another important building block: Poisson processes. Considering a wider class of Lévy processes also means that the numerical methods to solve problems involving these type of processes have to be developed further and other (more complex) techniques have to be used. We will make use of existing results [9, 23, 4] on existence and uniqueness of solutions of such SDE systems. As solutions of SDEs can rarely be found exactly, numerical methods are needed to solve SDEs driven by Lévy processes. Numerical methods for ordinary differential equations (ODEs) and results about convergence, consistency and stability are commonly known (see e.g. [75, 19, 51]). When dealing with stochastic differential equations (SDEs) instead of ODEs, these methods used in the same fashion work poorly for SDEs. There is a broad literature on methods for SDEs [69, 92, 50, 97], which adapt the methods for ODEs to SDEs and give other methods. When dealing with SDEs, the main difference is, that there is some source of randomness involved, which makes these methods more complicated. Due to this randomness, the question arises of how to measure closeness between the exact solution and an approximation. For this, we will define mean-square and weak convergence and mention their use.

This part of the thesis begins with the Preliminaries in Chapter 3, where we first define characteristic functions and moments. We then give definitions for stochastic processes in general and specifically for Lévy Processes in Section 3.2 and further define the concept of infinite divisibility, which is important for descriptions of distributional properties of Lévy processes. In Section 3.3, we introduce the Lévy–Khintchine representation, which gives a formula for the characteristic function of any infinitely distributed random variable. We mention the Lévy–Ito decomposition, that helps us to understand the structure of Lévy processes better and give some examples of commonly known Lévy processes. This is then followed by Section 3.4 about SDEs driven by Lévy processes, a short introduction about existence and uniqueness of solutions of such SDE systems and a description of how to measure numerical convergence for SDEs. Moreover, we focus on weak approximation and introduce a general IPDE problem. In Chapter 4 we introduce the main topic of this research: a boundary value IPDE problem. We then discuss possible numerical methods and questions associated with this problem and review existing literature in Section 4.2. In particular, we highlight the differences between our research and existing literature with particular focus on the case of Lévy processes with infinite activity. The most important chapter of this part is Chapter 5, where we introduce the suggested restricted time-stepping algorithm and give proofs for numerical convergence. We want to highlight the extensive discussion with respect to infinite activity in Section 5.3.4. In the last chapter of this part, we showcase the theoretical results in a range of theoretical and practical numerical examples.

3

PRELIMINARIES ON LÉVY PROCESSES AND SDES DRIVEN BY LÉVY PROCESSES

In this chapter we introduce some selected useful tools and characteristics when working with Lévy processes such as characteristic functions (CFs), moment generating functions (MGFs) and the characteristic exponent (sometimes also known as cumulant generating function [23] or Lévy exponent [4]). Furthermore, we explain their useful properties and how they are connected. We then have a look at the concept of infinite divisibility, the Lévy-Ito decomposition and introduce the Lévy-Khintchine formula, which shows that characteristic functions of Lévy processes have a specific form. We end this chapter by briefly introducing mean–square and weak approximations of SDEs. This chapter is mainly based on [4, 23].

3.1 CHARACTERISTIC FUNCTION AND MOMENTS

We will give some basic definition of σ–algebras and measures followed by some introduction into characteristic functions.

Definition 3.1.1. Let Ω be a non-empty set and $\mathcal{F}$ a collection of subsets of Ω. We then call $\mathcal{F}$ a σ**–algebra** if the following hold:

1. $\emptyset \in \mathcal{F}$, where $\emptyset$ is the empty set.
2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$, where $A \in \Omega$.

3. If $(A_n, n \in \mathbb{N})$ is a sequence of subsets in $\mathcal{F}$ then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

The pair $(\Omega, \mathcal{F})$ is called a **measurable space**. A **measure** on $(\Omega, \mathcal{F})$ is a mapping $\mu : \mathcal{F} \to [0, \infty]$ that satisfies

1. $\mu(\emptyset) = 0$,
2. $$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

   for every sequence $(A_n, n \in \mathbb{N})$ of mutually disjoint sets in $\mathcal{F}$.

Then $(\Omega, \mathcal{F}, \mu)$ is called a **measure space**. One commonly known example for a measure is the *Borel measure*. Let $\mathcal{B}(\mathbb{R}^d)$ be the Borel σ–algebra of $\mathbb{R}^d$, which is the smallest σ–algebra of subsets of $\mathbb{R}^d$ that contain all the open sets. If $S \in \mathcal{B}(\mathbb{R}^d)$ then its Borel σ–algebra is defined as follows

$$\mathcal{B}(S) = \{E \cup S; E \in \mathcal{B}(\mathbb{R}^d)\},$$

where E is a set in $\mathcal{B}(\mathbb{R}^d)$. Any measure on $(S, \mathcal{B}(S))$ is then called a Borel measure. In general, a measure does not need to be finite. A measure μ defined on $(\Omega, \mathcal{F})$ is finite, if $\mu(\Omega) < \infty$. A more flexible measure is the Radon measure defined as follows:

Definition 3.1.2. Let $A \subset \mathbb{R}^d$ and $(A, \mathcal{A})$ be a measurable space. A measure μ is called a **Radon measure**, if for every compact measurable set $B \in \mathcal{A}$, $\mu(B) < \infty$.

Another measure, which is also useful is a so-called Lebesgue measure, defined as follows:

Definition 3.1.3. Let $\mathcal{B}(S)$ be the Borel-σ algebra, then we can define a measure $\lambda(\omega)$ called the **Lebesgue measure** on $\mathbb{R}^d$ as such:

$$\lambda(A) := \int_A ds.$$

A point measure, which is also useful, is defined as follows:

Definition 3.1.4. A measure $\delta_x(\omega)$ associated with a point $x \in E$ is defined as follows:

$$\delta_x(A) := \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{otherwise,} \end{cases}$$

for $x \in \mathbb{R}^d$. Such a measure is then called **Dirac measure**.

In the case, where a measure $P : \mathcal{F} \to [0, 1]$ and $P(\Omega) = 1$, we call this measure P a **probability measure** and $(\Omega, \mathcal{F}, P)$ a **probability space**.
If $(\Omega, \mathcal{F}, P)$ is a given probability space, then a function $f : \Omega \to \mathbb{R}^n$ is called **$\mathcal{F}$-measurable** if

$$f^{-1}(A) := \{\omega \in \Omega; f(\omega) \in A\} \in \mathcal{F}$$

for all open sets $A \in \mathbb{R}^d$.

In the theory of stochastic processes, we are particularly interested in complete probability spaces.

Definition 3.1.5. A probability space $(\Omega, \mathcal{F}, P)$ is called **complete** if for all events $A \subset B$ for $B \in \mathcal{F}$ with $P(B) = 0$ implies that $A \in \mathcal{F}$.

Remark 3.1.6. Throughout this work, we make the assumption that all probability spaces $(\Omega, \mathcal{F}, P)$ are complete (if not otherwise stated). This can be referred to as the "standard assumptions".

Definition 3.1.7. Let X be a $\mathbb{R}^d$-valued random variable defined on $(\Omega, \mathcal{F}, P)$. The **characteristic function** (CF) of X is the function $\Phi_X : \mathbb{R}^d \to \mathbb{C}$ defined by

$$\Phi_X(t) := \mathbb{E}\left[e^{i\langle t,X\rangle}\right] = \int_\Omega e^{i\langle t,X(\omega)\rangle} dP(\omega),$$

where $t \in \mathbb{R}^d$ and (t, X) are the scalar products of t and X.

Note, that the characteristic function completely determines the distribution of the random variable X and always exists, as it is the Fourier transform of the probability measure in respect to X.

Definition 3.1.8. A measure μ_2 is said to be **absolutely continuous** with respect to a measure μ_1 if for any measurable set A

$$\mu_1(A) = 0 \Rightarrow \mu_2(A) = 0.$$

Further, if X is absolutely continuous, then the probability density function (PDF) f_X exists and $\int_\Omega e^{i\langle t, X(\omega)\rangle} dP(\omega) = \int_{\mathbb{R}^d} e^{i\langle t,y\rangle} df_X(y)$. If the moment generating function (MGF) exists as well then it is defined as follows.

Definition 3.1.9. Let X be a $\mathbb{R}^d$-valued random variable defined on $(\Omega, \mathcal{F}, P)$. The **moment generating function** (MGF) of X is defined (assuming $E[e^{\langle t,X\rangle}]$ exists) by

$$M_X(t) := \Phi(-it) = \mathbb{E}[e^{\langle t,X\rangle}],$$

where $t \in \mathbb{R}^d$.

Both the CF and the MGF have useful properties. We will limit our presentation to the properties of the CF, as in the general case the CF always exists (while the MGF not necessarily does):

- If $\mathbb{E}\left[|X_j|^n\right] < \infty$ for some $1 \leq j \leq d$ and $k \in \mathbb{N}$ and $k \leq n$, then the kth moment $\mathbb{E}\left[X_j^k\right]$ can be calculated by differentiating the CF:

$$\mathbb{E}\left[X_j^k\right] = i^{-k} \frac{\partial^k \Phi_X(t)}{\partial t_j^k}\bigg|_{t=0}.$$

- If $(X_i, i = 1, \dots, d)$ are independent random variables, then the CF of $Y = X_1 + \dots + X_d$ is given by:

$$\Phi_Y(t) = \prod_{i=1}^{d} \Phi_{X_i}(t). \qquad (3.1.1)$$

The definition of moments can be found in the Appendix B.1.3. As the characteristic function of Lévy processes can be expressed in a specific way, we will look at the log-characteristic function. This leads to the definition of the characteristic exponent.

Definition 3.1.10. Let X be a $\mathbb{R}^d$-valued random variable defined on $(\Omega, \mathcal{F}, P)$ with characteristic function $\Phi_X(t), t \in \mathbb{R}^d$. The **characteristic exponent** of X is the function $\Psi_X(t)$ such that

$$\Phi_X(t) = e^{\Psi_X(t)},$$

where $t \in \mathbb{R}^d$.

Note that similarly to (3.1.1), the characteristic exponent of $Y = X_1 + ... + X_n$, with independent random variables $(X_i, i = 1, ..., n)$ can be written as follows:

$$\Psi_Y(t) = \Psi_{X_1+...+X_n}(t) = \sum_{i=1}^n \Psi_{X_i}(t).$$

For a selection of the most known distributions, the characteristic function and also the characteristic exponent are commonly known and can be found in Table 3.1.1.

Table 3.1.1: Some probability distributions, their characteristic functions and characteristic exponents

distribution	Exponential $(\lambda > 0)$	Poisson $(\lambda > 0)$	Normal
$f_X(x)$	$\lambda e^{-\lambda x} \mathbb{1}_{x \geq 0}$	$e^{-\lambda} \frac{\lambda^x}{x!}, x \in \mathbb{N}$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
$\Phi_X(t)$	$\frac{\lambda}{\lambda - t}$	$e^{\lambda(e^{it}-1)}$	$e^{i\mu t - \frac{1}{2}\sigma^2 t^2}$
$\Psi_X(t)$	$\log\left(\frac{\lambda}{\lambda-t}\right)$, for $t < \lambda$	$\lambda(e^{it} - 1)$	$i\mu t - \frac{1}{2}\sigma^2 t^2$

3.2 LÉVY PROCESSES

Stochastic processes describe the random evolution of a dynamical system over time [64]. Opposed to a deterministic process, the evolution of this process has uncertainty in regard to its outcome. Lévy processes are stochastic processes with specific properties, which we will define in this chapter. We will introduce the concept of infinite divisibility and how it is connected to Lévy processes. Further, we explain the notion of random measures and in particular Poisson random measures.

Definition 3.2.1. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space and let $X = (X(t), t \geq 0) = (X_t)_{t\geq 0}$ be a family of random variables defined on that probability space and $t \in \mathbb{R}$ is interpreted as time. Then we say that X is a **stochastic process**.

Note that one can class stochastic processes in discrete-time (for more information see e.g. [41]) and continuous-time processes. Stochastic processes can be used to model various phenomena in Biology [2], Physics (see [111] or [9]), Social Sciences [43] and Finance (see [107] or [15]).

When looking at stochastic processes over time, we are also interested in the information flow over time. As time goes on more information is progressively available. This feature can be added to the structure of the probability space and is defined as follows.

Definition 3.2.2. Let $(\Omega, \mathcal{F}, P)$ be a complete probability space. A **filtration** on $(\Omega, \mathcal{F}, P)$ is an increasing family of σ-algebras $(\mathcal{F}_t), t \in [0, T]$ for which the following holds:

$$\forall\, t \geq s \geq 0 \qquad \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}.$$

Note that a probability space equipped with a filtration is called a **filtered probability space** and denoted by $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$.

We are now prepared to introduce a Lévy process, which is a class of stochastic processes with certain properties.

Definition 3.2.3. A stochastic process $(X_t)_{t\geq 0}$ defined on $(\Omega, \mathcal{F}, P)$ with $X_0 = 0$ is said to be a **Levy process** if it has the following properties:

1. The paths of X are right continuous with left-hand limits (i.e. càdlàg).

2. Independent increments: For $0 \leq s \leq t$, $X_t - X_s$ is independent of the σ-algebra $\mathcal{F}_t = \{X_u : u \leq t\}$.

3. Stationary (time-homogeneous) increments: For $0 \leq s \leq t$, $X_t - X_s$ has the same distribution as X_{t-s}.

4. Stochastic continuity: $\forall \epsilon > 0, \lim_{t \to s} P(|X_t - X_s| > \epsilon) = 0$.

Note that sample paths of Lévy processes do not necessarily have to be continuous, this is only true for some subclasses such as Wiener processes. We are interested in distributions of independent increments of Lévy processes and therefore, the concept of infinite divisibility is important, as it will put a constraint on possible choices.

Definition 3.2.4. Let X be a random variable valued in $\mathbb{R}^d$ with probability measure P_X. We then say that X is **infinitely divisible**, if $\forall n \in \mathbb{N}$ there exist n i.i.d. random variables $Y_1, \dots, Y_n$ such that $Y_1 + \dots + Y_n$ has the same distribution as X.

Note that it can be shown that the equality in distribution of $Y_1 + \dots + Y_n$ and X_t (denoted by $X_t \stackrel{d}{=} Y_1 + \dots + Y_n$) can be expressed in the following ways:

- In terms of the characteristic function:

$$\Phi_X(t) = \Phi_{X_1} \times \cdots \times \Phi_{X_n}.$$

- As a convolution of identical measures μ_i (see [4]):

$$\mu_X = \mu_1 * \cdots * \mu_n.$$

Before we look at some examples of infinitely divisible distributions, we can look at the characteristic function $\Phi(t)$ and the characteristic exponent $\Psi(t)$, which easily allow us to determine, whether a distribution is infinitely divisible or not. This property is formulated in the following proposition (for a proof see [4]).

Proposition 3.2.5. *Let X be a random variable on $(\Omega, \mathcal{F}, P)$. Then the following statements are equivalent:*

- *X is infinitely divisible.*
- *The n–th root of the characteristic function $\Phi_X(t)$ is the characteristic function of n i.i.d. random variables $X_i, i \in \mathbb{N}$:*

$$\Phi_X(t) = \Phi_{X_1} \times \cdots \times \Phi_{X_n} = [\Phi_{X_1}]^n .$$

- *The characteristic exponent $\Psi_X(t)$ can be written as a sum of the characteristic exponents $\Psi_{X_i}(t)$ of n i.i.d. random variables:*

$$\Psi_X(t) = \Psi_{X_1} + \cdots + \Psi_{X_n} = n\Psi_{X_1}.$$

Remark 3.2.6. It should be easy to see that the concept of infinite divisibility gives us an idea of the structure the characteristic function of a Lévy process should have. The characteristic function of a Levy process $(X_t)_{t\geq 0}$ on $\mathbb{R}^d$ has the following form:

$$\Phi_{X_t}(z) = e^{t\Psi(z)}, \quad z \in \mathbb{R}^d,$$

where $\Psi(z)$ is the characteristic exponent of $X_1 = X(1)$ with $t = 1$. This shows us that characteristic exponent of a Lévy process varies linearly in t. Therefore, the knowledge over distribution of X_1 will be sufficient to specify the distribution of X_t. A rough outline of the proof can be found in Appendix A.1.1. Note that, we use this form for the characteristic exponent, when talking about Lévy processes rather than the more general form from Definition 3.1.10

Example 3.2.1 (Normal distribution)**.** Let us only look at the simple case where $d = 1$ and let Y be a random variable which follows the normal distribution $Y \sim \mathcal{N}(\mu, \sigma^2)$, hence it is obvious that Y must be a infinitely divisible random variable as its CF $\Phi_Y(t)$ (as seen above in Table 3.1.1) can be written as

$$\begin{aligned}\Phi_Y(t) &= \exp\left(i\mu t - \frac{1}{2}\sigma^2 t^2\right) \\ &= \left[\exp\left(i\frac{\mu}{n}t - \frac{1}{2}\frac{\sigma^2}{n}t^2\right)\right]^n \\ &= [\Phi_{Y_1}(t)]^n, \quad t \in \mathbb{R},\end{aligned}$$

where $\Phi_{Y_1}(t)$ is the CF of a normal distributed random variable $Y_1 \sim \mathcal{N}\left(\frac{\mu}{n}, \frac{\sigma^2}{n}\right)$. Similarly, we can deduct the following statement for the characteristic exponent:

$$\Psi_Y(t) = i\mu t - \frac{1}{2}\sigma^2 t^2 = n\left(i\frac{\mu}{n}t - \frac{1}{2}\frac{\sigma^2}{n}t^2\right) = n\Psi_{Y_1}(t).$$

Similar statements can be shown for the multivariate case.

Example 3.2.2 (Poisson distribution)**.** In a similar fashion for $d = 1$, let Z be a random variable which follows a Poisson distribution: $Z \sim Poi(\lambda)$. Then it is also easy to see that Z is infinitely divisible as

$$\begin{aligned}\Phi_Z(t) &= \exp\left(\lambda(e^{it} - 1)\right) \\ &= \left[\exp\left(\frac{\lambda}{n}(e^{it} - 1)\right)\right]^n \\ &= [\Phi_{Z_1}(t)]^n, \quad t \in \mathbb{R},\end{aligned}$$

where $\Phi_{Z_1}(t)$ is the CF of a Poisson distributed random variable $Z_1 \sim Poi\left(\frac{\lambda}{n}\right)$.

Lévy processes can be used to describe jumps occurring in different observed phenomena [23, 9]. A convenient tool to analyse these jumps are random measures and in particular Poisson random measures.

Definition 3.2.7. Let $E \in \mathbb{R}^d$ and $(E, \mathcal{E})$ be a measurable space and $(\Omega, \mathcal{F}, P)$ a probability space. A measure M on E is called **random measure** if the following holds:

$$M : \Omega \times \mathcal{E} \to \mathbb{R}_+, \text{ i.e. } (\omega, A) \mapsto M(\omega, A),$$

such that

1. For (almost all) $\omega \in \Omega$, $M(\omega, \cdot)$ is measure on E.

2. For each measurable set $A \subset E$, $M(\cdot, A) = M(A)$ is a random variable.

To understand the notion of random measures better we show its connection to a Poisson process.

Example 3.2.3 (Random jump measure)**.** Let $(\tau_i)_{i \geq 1}$ be a sequence of independent exponential random variables with parameter $\lambda > 0$ and $T_n = \sum_{i=1}^{n} \tau_i$. The process $(N_t)_{t \geq 0}$ defined by

$$N_t = \sum_{n \geq 1} \mathbb{1}_{t \geq T_n} = \#\{n \geq 1, t \geq T_n\}$$

is called a Poisson process with intensity λ. It is a counting process of the number of jumps which occur in the time $[0, t]$.

This counting procedure defines a random measure M on $[0, \infty)$, for any $A \subset \mathbb{R}_+$ let

$$M(\omega, A) = \#\{i \geq 1, T_i(\omega) \in A\}.$$

Note that $M(\omega, A)$ depends on ω, hence it is a *random* measure. If we fix ω, then all $T_i(\omega)$ are deterministic, hence if $A = \emptyset$, i.e. all $T_i(\omega) \notin A$, then $M(\omega, \emptyset) = 0$. Moreover, let $(A_n, n \in \mathbb{N})$ be a sequence of subsets in $\mathcal{F}$, such that $\cup_{n=1}^{\infty} A_n \in \mathcal{F}$. Then, $M(\omega, \cup_{n=1}^{\infty} A_n) = \#\{i \geq 1, T_i(\omega) \in \cup_{n=1}^{\infty} A_n\} = \sum_{n=1}^{\infty} \#\{i \geq 1, T_i(\omega) \in A_n\} = \sum_{n=1}^{\infty} M(\omega, A_n)$.

Further, the intensity λ of the Poisson process determines the average value of this random measure: $\mathbb{E}[M(A)] = \lambda |A|$, where $|A|$ is the Lebesgue measure of A. M can also be called the random jump measure associated to the Poisson process N_t, which allows us to express N_t in the following way:

$$N_t(\omega) = M(\omega, [0,t]) = \int_{[0,t]} M(\omega, ds).$$

Definition 3.2.8. Let $E \in \mathbb{R}^d$ and μ be a given Radon measure on $(E, \mathcal{E})$ and $(\Omega, \mathcal{F}, P)$ a probability space. An integer valued random measure N on E with intensity measure μ is called **Poisson random measure** if the following holds:

$$N : \Omega \times \mathcal{E} \to \mathbb{N}, \text{ i.e. } (\omega, A) \mapsto N(\omega, A),$$

such that

1. For (almost all) $\omega \in \Omega$, $N(\omega, \cdot)$ is an integer valued Radon measure on E: for any bounded measurable $A \subset E$, $N(A) < \infty$ is an integer valued random variable.

2. For each measurable set $A \subset E$, $N(\cdot, A) = M(A)$ is a Poisson random variable with parameter $\mu(A)$:

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

3. For disjoint measurable sets $A_1, \ldots, A_n \in \mathcal{E}$, the random variables $N(A_1), \ldots, N(A_n)$ are independent.

For an easier understanding, we can look at two examples of a Poisson random measure [21, 23].

Example 3.2.4 (Particles in boxes)**.** Let E be a countable set with n number of elements (i.e. $|E| = n$). Let $\mathcal{E} = \mathcal{P}(E)$ be the power set of E, then $|\mathcal{E}| = 2^n$. Let μ be a measure on it. For each $x \in E$, let N_x be independent Poisson distributed random variables with mean $\mu(\{x\})$. We may think of E as a countable collection of boxes and of N_x as the number of particles in the box x. Then,

$$N(\omega, A) = \sum_{x \in E} N_x(\omega) \mathbb{1}_x(A) \quad \omega \in \Omega, A \in \mathcal{E},$$

defines a Poisson random measure M on $(E, \mathcal{E})$ with intensity measure μ. It can be easily seen, that for $\omega \in \Omega$ (meaning for any realisation of random number of particles in the boxes $N_x(\omega)$), $N(\omega, \cdot)$ is a Radon measure on E for $A \in E$ as $N(A) < \infty$ (as for any set of boxes, the number of total particles will be finite). The Poisson random measure can be seen as a total counter of particles in all boxes for a certain set of boxes. Moreover, let us fix $A \subset E$, then $M(\omega, A)$ is clearly a random variable dependent on ω. Due to the independence of N_x and $N_y, x \neq y$ and the fact that the sum of independent Poisson distributed random variables is also Poisson distributed, $N(\cdot, A) = \sum_{x \in E} N_x(\omega) \mathbb{1}_x(A) = N_{\sum_{x \in A} x}$. Lastly, for disjoint measurable sets $A_1, \ldots, A_n \in \mathcal{E}$, the random variables $N(A_1), \ldots, N(A_n)$ are clearly independent as $\forall k \in \mathbb{N}, \quad P(N(A_i) = k) = e^{-\mu(A_i)} \frac{(\mu(A_i))^k}{k!}$ are independent.

Example 3.2.5 (Poisson random measure as jump processes)**.** Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ be a filtered probability space. We can now consider a Poisson random measure N on $E = [0, T] \times \mathbb{R}^d \backslash \{0\}$ with parameter $\mu(A), A \in E$. It can be described as the counting measure associated to a random configuration of points $(T_n, Y_n) \in E$:

$$N(\omega, A) = \sum_{n \geq 1} \delta_{(T_n(\omega), Y_n(\omega))}(A) \quad A \subset E,$$

where $\delta_x(A)$ is the Dirac measure of a point $x = (T_n, Y_n) \in E$. Each point $(T_n(\omega), Y_n(\omega))$ corresponds to an observation made at time T_n and described by a random variable $Y_n(\omega) \in \mathbb{R}^d$. We can interpret the first coordinate t as time and we will say that N is a non–anticipating (or adapted to the filtration $(\mathcal{F}_t)_{t \geq 0}$) Poisson random measure, if $(T_n)_{n \geq 1}$ are non–anticipating random times and Y_n is known at T_n. The Poisson random measure $N(\omega, A)$ can be seen as a counting process (in regard to time and space) for all points in a certain area A.

Introducing a real valued measurable function $f : E \mapsto \mathbb{R}$, for which the following holds:

$$\mu(|f|) = \int_{[0,T]} \int_{\mathbb{R}^d \setminus \{0\}} |f(s,y)| \mu(ds\, dy) < \infty,$$

then we can create a stochastic process which corresponds to the Poisson random measure $N(A)$ and the function f. The intensity of that process can be described as the expectation of a random variable $N(f)$,

$$\mathbb{E}[N(f)] = \mu(f) = \int_{[0,T]} \int_{\mathbb{R}^d \setminus \{0\}} f(s,y) \mu(ds, dy).$$

If we now integrate f with respect to M up to time t, this gives a non–anticipating (or adapted) stochastic process:

$$X_t = X_t(f) = \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} f(s,y) N(ds, dy) = \sum_{\{n, T_n \in [0,t]\}} f(T_n, Y_n).$$

This example shows, that the Poisson random measure contains all information about the discontinuities (jumps) of the process X_t, whose jumps occur at random times T_n and have jump size $f(T_n, Y_n)$.

3.3 DISTRIBUTION AND STRUCTURE OF LÉVY PROCESSES

In the previous Section 3.2, we have introduced Lévy processes and random measures (see 3.2.3 and 3.2.7). Now, we will introduce the Lévy-Khintchine representation [4] which gives a formula for the characteristic function for infinitely divisible random variables and hence for Lévy processes. Then, we can state the Lévy-Ito decomposition, that allows us to describe the sample path structure.

First, we introduce the Lévy measure which is used in the Lévy–Khintchine formula. It allows us to deal with finite and infinite activity of Lévy processes.

Definition 3.3.1. Let ν be a Borel measure defined on $\mathbb{R}^d \backslash \{0\}$ for which the following condition holds

$$\int_{\mathbb{R}^d \backslash \{0\}} (|x^2| \wedge 1)\nu(dx) < \infty. \tag{3.3.1}$$

We call ν a **Lévy measure**.

In the first instance the above condition (3.3.1) does not seem to be obvious. However, it will ensure that all integrals in the Lévy-Khintchine formula exist. The Lévy measure is an important part in the definition of a Lévy process, especially in regard to the finite and infinite activity.

Definition 3.3.2. A Lévy process with Lévy measure ν is said to be of **finite** activity if $\nu(\mathbb{R}^d \backslash \{0\}) < \infty$. Otherwise, it is called of **infinite** activity.

The differentiation between Lévy processes of finite and infinite activity is important, as dealing with Lévy processes of infinite activity is usually more complex and also they are more difficult to simulate. Therefore, the two cases have to be treated separately, particularly with regards to numerical methods, which we will see in Chapter 5 and Chapter 6.

We can now proceed to the Lévy–Khintchine formula, which provides a characterization of random variables with infinitely divisible distributions using their characteristic functions. We will present it here without the proof which can be found in literature on Lévy processes such as [4], [23] or [12].

Theorem 3.3.3 (Lévy-Khintchine formula)**.** *Let X be a random variable valued on $\mathbb{R}^d$ with probability measure μ_X and CF $\Phi(z)$. Then μ_X is infinitely divisible if and only if there exists a triplet (b, A, ν), with $b \in \mathbb{R}^d$, a positive definite matrix $A \in \mathbb{R}^{d\times d}$ and a Lévy measure ν on $\mathbb{R}^d \backslash \{0\}$, such that for all $z \in \mathbb{R}^d$,*

$$\Phi(z) = \exp\left\{ i\langle b, z\rangle - \frac{1}{2}\langle z, Az\rangle + \int_{\mathbb{R}^d \backslash \{0\}} [e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle \mathbb{1}_{|x|<1}]\nu(dx) \right\}. \tag{3.3.2}$$

We can now take this idea further and formalise what we mentioned in Remark 3.2.6 and use it for Lévy processes which leads to the following theorem.

Theorem 3.3.4 (Lévy-Khintchine formula for Lévy processes). *Let $X = (X_t)_{t\geq 0}$ on $\mathbb{R}^d$ be a Lévy process, then the CF has the following form*

$$\begin{aligned}\Phi_{X_t}(z) &= e^{t\Psi(z)} \\ &= \exp\left\{t\left(i\langle b,z\rangle - \frac{1}{2}\langle z,Az\rangle \right.\right. \\ &\qquad \left.\left. + \int_{\mathbb{R}^d\setminus\{0\}}[e^{i\langle z,x\rangle} - 1 - i\langle z,x\rangle\mathbb{1}_{|x|<1}]\nu(dx)\right)\right\}, \quad z\in\mathbb{R}^d,\end{aligned}$$

where $\Psi(z)$ is the characteristic exponent of $X_1 = X(1)$ and (b,A,ν) is a triplet with $b \in \mathbb{R}^d$, a positive definite matrix $A \in \mathbb{R}^{d\times d}$ and a Lévy measure ν on $\mathbb{R}^d\setminus\{0\}$.

A proof can be found in [4].

Remark 3.3.5. Note that such a triplet (b,A,ν) is sometimes called the Lévy-Khintchine triplet and it can be shown that it is sufficient to characterize any Lévy process uniquely (see [4, Corollary 2.4.21] or [105, Theorem 8.1]).

To get a better understanding of the Lévy-Khintchine formula, we give some examples of commonly used finite Lévy processes and their Lévy-Khintchine triplet (if it exists).

Example 3.3.1 (Brownian Motion with drift). Let $(B_t)_{t\geq 0}$ with $B_0 = 0$ be a Brownian motion in $\mathbb{R}^m$, let $b \in \mathbb{R}^d$ be a vector, let $A \in \mathbb{R}^{d\times d}$ be a positive definite matrix and $\sigma \in \mathbb{R}^{d\times m}$ be such that $\sigma\sigma^T = A$. The (Gaussian) process $(C_t)_{t\geq 0}$ in $\mathbb{R}^d$ defined by

$$C_t = bt + \sigma B_t \tag{3.3.3}$$

has then the characteristic exponent of following form:

$$\Psi_C(z) = i\langle b,z\rangle - \frac{1}{2}\langle z,Az\rangle, \quad z\in\mathbb{R}^d.$$

The Lévy-Khintchine triplet of C_t can then be written as $(b, A, 0)$. Note that it can be shown, that a Lévy process is of the form (3.3.3) if and only if it has continuous sample paths. Note that a stochastic process is said to be Gaussian, if all its finite-dimensional distributions are Gaussian.

Example 3.3.2 (Poisson process)**.** Let $(N_t)_{t\geq 0}$ be a Poisson process of intensity $\lambda > 0$. As we have seen in example 3.2.3, N_t can be expressed in the form of its associated random jump measure M:

$$N_t = M([0,t]) = \int_{[0,t]} M(ds).$$

It is a Lévy process which takes values in $\mathbb{N} \cup 0$ and for any $t > 0$. N_t follows a Poisson distribution with parameter λt:

$$P(N(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!},$$

for each $n = 0, 1, 2, \ldots$. The characteristic exponent of N_t has the following form:

$$\Psi_N(z) = \lambda(e^{i\langle z,y\rangle} - 1), \quad z \in \mathbb{R}^d.$$

The Lévy-Khintchine triplet of N_t can then be written as $(0, 0, \lambda\delta_1)$, where δ_1 is the Dirac measure on 1. It can be clearly seen that the Poisson process is a simple jump process.

Example 3.3.3 (Compound Poisson process)**.** Let $Y_i, i = 1, 2, \ldots,$ be i.i.d. random variables in $\mathbb{R}^d$ with common probability distribution $F(y)$ and let (N_t) be a Poisson process with intensity $\lambda > 0$, that is independent of all Y_i. Then the compound Poisson process J_t with intensity λ is defined as follows:

$$J_t = \sum_{i=1}^{N_t} Y_i,$$

for each $t > 0$. As seen in Example 3.2.5, the compound Poisson process can also be expressed in the form of its associated Poisson random measure $N(dt, dx)$ with intensity λ in the following way:

$$J_t = \int_0^t \int_{\mathbb{R}^d \setminus \{0\}} xN(ds, dx).$$

The characteristic exponent of J_t has the following form:

$$\Psi_J(z) = \lambda \int_{\mathbb{R}^d \setminus \{0\}} (e^{i\langle z,y \rangle} - 1)F(dy), \quad z \in \mathbb{R}^d. \tag{3.3.4}$$

A derivation for the characteristic function and exponent of J_t can be found in the Appendix A.1.2. We can introduce a measure $\nu(A) = \lambda F(A)$ for which (3.3.4) then takes the following form:

$$\Psi_J(z) = \int_{\mathbb{R}^d} (e^{i\langle z,y \rangle} - 1)\nu(dy), \quad z \in \mathbb{R}^d.$$

The Lévy-Khintchine triplet of J_t can then be written as $(0, 0, \lambda F(\cdot))$.

So far we have looked at examples for Lévy processes of finite activity such as Examples 3.3.1, 3.3.2 and 3.3.3. When we want to look at examples for Lévy processes of infinite activity, we first have to introduce the idea of a compensated measure.

Definition 3.3.6. Let $N(A)$ be a Poisson random measure with intensity measure μ. Then we define a **compensated Poisson random measure** as

$$\hat{N}(A) := N(A) - \mu(A), \tag{3.3.5}$$

where $A \subset E$, where E is a set in $\mathbb{R}^d$.

We can now look at an example of a Lévy process of infinite activity.

Example 3.3.4 (Lévy process with infinite activity)**.** Let $X = (X_t)_{t \geq 0}$ on $\mathbb{R}^d$ be a Lévy process without Brownian motion and drift with infinite activity. We can describe X_t using its associated Poisson random measure $N(ds, dx)$

and its associated compensated Poisson random measure $\hat{N}(ds, dx)$ in the following way:

$$X_t = \int_0^t \int_{|x|<1} x\hat{N}(ds, dx) + \int_0^t \int_{|x|>1} xN(ds, dx).$$

Note that the idea is to split up the jumps of the Lévy process into jumps of size lower than one and bigger than one (this level is chosen arbitrarily). The existence of infinite activity implies that there can be an infinite number of small jumps. Therefore, we need to use the compensated Poisson random measure for the part of small jumps, which ensures analytical properties of the integrals. Due to the properties of the Lévy measure, there can only be a finite number of jumps bigger than one. Hence, it is sufficient to describe these jumps using the Poisson random measure. A more technical explanation can be found in Remark 3.3.8.

The characteristic exponent has the form

$$\Psi_X(z) = \int_{\mathbb{R}^d \setminus \{0\}} [e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle \mathbb{1}_{|x|<1}]\nu(dx), \quad z \in \mathbb{R}^d,$$

and the Lévy-Khintchine triplet of X_t is $(0, 0, \nu)$.

As we have seen in Theorems 3.3.3 and 3.3.4, the Lévy-Khintchine formula gives us information about the distribution of infinitely divisible random variables and Lévy processes. To understand the structure of the paths of Lévy processes, we will formulate a famous result known as the *Lévy–Ito decomposition*.

Theorem 3.3.7 (Lévy–Ito decomposition)**.** *If $X = (X_t)_{t\geq 0}$ on $\mathbb{R}^d$ is a Lévy process, then there exists $b \in \mathbb{R}^d$, a d–dimensional Brownian motion $(B_A(t))_{t\geq 0}$ with covariance matrix $A \in \mathbb{R}^{d\times d}$ and an independent Poisson random measure N on $[0, \infty) \times \mathbb{R}^d$, with intensity measure $\nu(dx) \times dt$ and Lévy measure ν, such that*

$$X_t = bt + B_A(t) + \int_0^t \int_{|x|<1} x\hat{N}(ds, dx) + \int_0^t \int_{|x|>1} xN(ds, dx). \qquad (3.3.6)$$

A proof of Theorem 3.3.7 can be found in [4] or [23].

Remark 3.3.8. This theorem allows us to understand the structure of Lévy processes and it hints at an idea of how to construct general Lévy processes by independently combining special cases of Lévy processes. When combining different parts of Examples 3.3.1, 3.3.3 and 3.3.4, X_t can be seen as a combination of a Brownian motion with constant drift, a compound Poisson process which describes the jumps of size larger than 1 and a compensated compound Poisson process.

To continue further, one can check, if all Lévy processes can be represented in a similar way, which is not as intuitively understandable on the first inspection. To emphasize that, this representation also describes Lévy processes with infinite activity, i.e. an infinite number of jumps, where most of the jumps are small. This suggests that the Lévy measure is not necessarily a finite measure. By the condition (3.3.1) on the Lévy measure, we ensure that the integral $\int_{\mathbb{R}^d\setminus\{0\}}[e^{i\langle z,x\rangle}-1-i\langle z,x\rangle\mathbb{1}_{|x|<1}]\nu(dx)$ exists for any $|x|$, as for small jumps ($|x|<1$) $|e^{i\langle z,x\rangle}-1-i\langle z,x\rangle|$ behaves like $|x^2|$ and the condition (3.3.1) ensures that $\int_{|x|<1}|x^2|\nu(dx)$ exists. Similarly, the same condition means that X only has a finite number of jumps, which are larger then 1 ($|x|>1$). This means, that an infinite amount of small jumps can occur, however the integrals in the Lévy-Khintchine formula still converge due to the mentioned condition on the Lévy measure.

3.4 SDES DRIVEN BY LÉVY PROCESSES AND EXISTENCE AND UNIQUENESS OF SOLUTIONS

The existence of (unique) solutions for SDEs driven by Lévy processes depends on regularity conditions on the parameters used in the SDE. We first introduce the concepts of Lipschitz continuity and polynomial growth for general functions. Then we proceed with introducing a suitable SDE system and state a Theorem on existence and uniqueness of a solution under certain regularity conditions.

Definition 3.4.1. A function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is said to be (globally) **Lipschitz (continuous)**, if there exists a positive constant $L \in \mathbb{R}$, such that

$$\|f(t,x) - f(t,y)\| \leq L\|x-y\| \tag{3.4.1}$$

holds for all $t \in [0,T]$ and $x, y \in \mathbb{R}^d$.

Definition 3.4.2. A function $f : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is said to have **polynomial growth**, if for some n there exists a positive constant $C \in \mathbb{R}$, such that

$$\sup_{t\in[0,T]} \|f(t,x)\| \leq C(1+\|x\|^n) \tag{3.4.2}$$

holds for all $x \in \mathbb{R}^d$. We then say, that the function f belongs to the class of functions $\mathbf{F}$, written $f \in \mathbf{F}$. The function f is said to have **linear growth**, if (3.4.2) holds for $n = 1$.

Let us now introduce a general system of SDEs driven by Lévy processes. Suppose that we are given a d-dimensional standard $\mathcal{F}_t$-adapted Brownian motion process $w = (w(t), t \geq 0)$ with $w(t) = (w^1(t), \dots, w^d(t))^T$ for each $t \geq 0$ and an independent $\mathcal{F}_t$-adapted Poisson random measure N defined on $[0, \infty) \times \mathbb{R}^m$ with compensator $\hat{N}(dt, dy) = N(dt, dy) - \nu(dy)dt$ where ν is a Lévy measure. This system of SDEs can be described as follows:

$$\begin{aligned} dX = b(t, X(t-))dt + \sigma(t, X(t-))dw(t) + \int_{|y|<1} F(t, X(t-))y\hat{N}(dt, dy) \\ + \int_{|y|>1} F(t, X(t-))yN(dt, dy), \end{aligned} \tag{3.4.3}$$

where X and b are vectors of dimension d and σ is a $d \times d$ matrix, $X(t-)$ denotes X just before a jump time (assuming a jump time is occurring at t). Further, $F(t,x) = (F^{ij}(t,x))$ is a $d \times m$-matrix. We can now consider (3.4.3) as an initial value problem with a fixed initial condition $X(t_0) = x$, where $x \in \mathbb{R}^d$. Further, we make the following assumptions on the coefficients of (3.4.3):

(i) **Lipschitz condition:** There exists a constant $L > 0$ such that,

$$\|b(t,x) - b(t,y)\|^2 + \|\sigma(t,x) - \sigma(t,y)\|^2 + \int_{\mathbb{R}^d} \|F(t,x) - F(t,y)\|^2 \|z\|^2 \nu(dz) \leq L\|x-y\|^2 \tag{3.4.4}$$

holds for all $t \in [0,T]$ and $x, y \in \mathbb{R}^d$ and $\|\cdot\|$ is the $n-$dimensional Euclidean norm.

(ii) **Growth condition:** There exists a constant $C > 0$ such that,

$$\|b(t,x)\|^2 + \|\sigma(t,x)\|^2 + \int_{\mathbb{R}^d} \|F(t,x)\|^2 \|z\|^2 \nu(dz) \leq C(1 + \|x\|^2) \tag{3.4.5}$$

holds for all $x \in \mathbb{R}^d$.

Let $X_{t_0,x}$ be a Lévy procces solving the SDE system (3.4.3) and let us assume that the conditions (3.4.4) and (3.4.5) hold. Then we can describe X_{0,X_0} as follows:

$$\begin{aligned} X_{t_0,x}(t) = x &+ \int_{t_0}^{t} b(s, X(s-))\mathrm{d}s + \int_{t_0}^{t} \sigma(s, X(s-))\mathrm{d}w(s) \\ &+ \int_{t_0}^{t}\int_{|y|<1} F(s, X(t-))y\hat{N}(ds,dy) \\ &+ \int_{t_0}^{t}\int_{|y|>1} F(s, X(t-))yN(ds,dy). \end{aligned} \tag{3.4.6}$$

We can now formulate the following theorem similar to [4, Theorem 6.2.3] about the existence of a unique strong solution for the problem (3.4.3), which we state here without proof, which can be found in [4].

Theorem 3.4.3. *Assume that the coefficients of the SDE (3.4.3) follow the Lipschitz and growth conditions (3.4.4) and(3.4.5). Then there exists a unique solution* $X = (X_t)_{0\leq t}$ *to the SDE (3.4.3) with initial condition* $X(t_0) = x$. *Further,* X *is adapted to the filtration* $(\mathcal{F}_t, t \geq 0)$ *and càdlàg.*

3.5 MEAN–SQUARE AND WEAK ORDER CONVERGENCE FOR NUMERICAL METHODS

When simulating solutions of a SDE system such as (3.4.3), one is usually interested in either the path of the trajectory or in the expected value of some functional of the process. Numerical simulation methods are usually based on discrete approximations of the continuous solution and, when speaking about convergence in regard to these numerical methods, there are commonly two main criteria used: *mean–square* and *weak convergence* (see [92]).

Definition 3.5.1. Let $(X(t), \mathcal{F}_t)$, $t \in [t_0, T]$ be a solution of the SDE system (3.4.3) and let X_k be its numerical approximation for the (time-)steps $t_k \in [t_0, T], k = 0, \dots, N$ with the fixed step size $h = t_{k+1} - t_k$. Then we say that the **mean–square order of convergence** of this method is equal to p, if

$$\left(\mathbb{E}|X(t_k) - X_k|^2\right)^{\frac{1}{2}} \leq Kh^p, \tag{3.5.1}$$

where K is a positive constant independent of k and h.

Definition 3.5.2. Let $(X(t), \mathcal{F}_t)$, $t \in [t_0, T]$ be a solution of the SDE system (3.4.3) and let X_k be its numerical approximation for the (time-)steps $t_k \in [t_0, T], k = 0, \dots, N$ with the fixed step size $h = t_{k+1} - t_k$. Then we say that the **weak order of convergence** of this method is equal to p, if

$$|\mathbb{E}(g(X(T))) - \mathbb{E}(g(X_N))| \leq Kh^p, \tag{3.5.2}$$

for g from a class of functions **F** and where K is a positive constant independent of h.

We will focus on the *weak order of convergence*, as the main goal of this part of the thesis is about solving a Dirichlet IPDE problem (which we introduce in Chapter 4). Weak methods are sufficient in this case, as we are only interested in using numerical methods to find an approximation to the expectation of a functional of the corresponding system of SDE to solve the IPDE problem.

4

NUMERICAL METHODS FOR IPDE PROBLEMS AND LITERATURE REVIEW

In this chapter, we focus on the probabilistic representation of the parabolic solution of integro-partial differential equation (IPDE) problems. We are interested in extending existing numerical schemes, in particular research used for Dirichlet PDE problems [92, Algorithm 2 in Chapter 6.2], to the Dirichlet IPDE case. For that reason, we explore on how to deal with finite and infinite activity for Dirichlet (and Cauchy) IPDE problems. Firstly, we will introduce the relevant problem in Section 4.1 and then review other existing literature in Section 4.2.

4.1 THE IPDE PROBLEM

The main topic of this research is about solving the following IPDE problem.

Let G be a bounded domain in $\mathbb{R}^d$, $Q = [t_0, T) \times G$ be a cylinder in $\mathbb{R}^{d+1}$, $\Gamma = \bar{Q} \setminus Q$ be the part of the cylinder's boundary consisting of the upper base and lateral surface, $G^c = \mathbb{R}^d \setminus Q$ be the complement of G and $Q^c := (t_0, T] \times G^c \cup \{T\} \times \bar{G}$. Consider the Dirichlet problem for the integro-partial differential equation (IPDE):

$$\begin{aligned} \frac{\partial u}{\partial t} + Lu + c(t,x)u + g(t,x) = 0, \quad (t,x) \in Q, \\ u(t,x) = \varphi(t,x), \quad (t,x) \in Q^c, \end{aligned} \tag{4.1.1}$$

where the integro-differential operator L is of the form

$$\begin{aligned} Lu(t,x) :&= \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}(t,x)\frac{\partial^2 u}{\partial x^i \partial x^j}(t,x) + \sum_{i=1}^{d} b^i(t,x)\frac{\partial u}{\partial x^i}(t,x) \\ &+ \int_{\mathbb{R}^m} \Big\{ u\big(t, x + F(t,x)z\big) - u(t,x) \\ &- \langle F(t,x)z, \nabla u(t,x)\rangle \mathbf{I}(|z| \le 1) \Big\} \nu(\mathrm{d}z); \end{aligned} \tag{4.1.2}$$

$a(t,x) = \big(a^{ij}(t,x)\big)$ is a $d \times d$-matrix; $b(t,x) = (b^1(t,x), \dots, b^d(t,x))^\top$ is a d-dimensional vector; $c(t,x)$, $g(t,x)$, and $\varphi(t,x)$ are scalar functions; $F(t,x) = \big(F^{ij}(t,x)\big)$ is a $d \times m$-matrix; and $\nu(z)$, $z \in \mathbb{R}^m$, is a Lévy measure such that $\int_{\mathbb{R}^m} (|z|^2 \wedge 1)\nu(\mathrm{d}z) < \infty$. We allow ν to be of infinite intensity, i.e. we may have $\nu\big(B(0,r)\big) = \infty$ for some $r > 0$, where as usual for $x \in \mathbb{R}^d$ and $s > 0$ we write $B(x,s)$ for the open ball of radius s centred at x.

When the solution u of (4.1.1) is regular enough, for example when

$$u \in C^{1,2}\left([t_0, T] \times \mathbb{R}^d\right),$$

the Feynman-Kac formula (see [23][Proposition 12.6] or references therein) assures a probabilistic representation of the solution $u(t,x)$ to (4.1.1) in terms of the following system of Lévy-driven SDEs:

$$u(t,x) = \mathbb{E}\left[\varphi\left(\tau_{t,x}, X_{t,x}(\tau_{t,x})\right) Y_{t,x,1}(\tau_{t,x}) + Z_{t,x,1,0}(\tau_{t,x})\right], \quad (t,x) \in Q, \tag{4.1.3}$$

where $(X_{t,x}(s), Y_{t,x,y}(s), Z_{t,x,y,z}(s))$ for $s \ge t$, solves the system of SDEs consisting of (5.1.1) and

$$dY = c(s, X(s-))Y ds, \quad Y_{t,x,y}(t) = y, \tag{4.1.4}$$

$$dZ = g(s, X(s-))Y ds, \quad Z_{t,x,y,z}(t) = z, \tag{4.1.5}$$

and $\tau_{t,x} = \inf\{s \ge t : (s, X_{t,x}(s)) \notin Q\}$ is the first exit-time of the space-time Lévy process $(s, X_{t,x}(s))$ from the space-time cylinder Q. To see why this holds, one may apply Ito's lemma, see e.g. [4, Theorem 4.4.7], and the fact that u solves (4.1.1) to prove that the process

$$u\left(t \wedge \tau_{t,x}, X_{t,x}(t \wedge \tau_{t,x})\right) Y_{t,x,1}(t \wedge \tau_{t,x}) + Z_{t,x,1,0}(t \wedge \tau_{t,x}),$$

is a martingale. The claimed formula follows by letting $t \to \infty$.

A weak-sense approximation (as described in Section 3.5) of the SDEs together with the Monte Carlo technique gives us a numerical approach to evaluating $u(t,x)$, which is especially useful for higher-dimensional problems.

This introduced Dirichlet problem (4.1.1) is particularly interesting in the context of financial products, when assuming that the underlying asset follows some sort of jump process and there is various research into the field of option pricing based on Lévy-type models [87, 23], and we illustrate this in Section 6.4.

4.2 A LITERATURE REVIEW ON SOLUTIONS FOR IPDE PROBLEMS

In this section, we compare existing literature for different numerical methods to approximate solutions of problems where the underlying SDEs are driven by Lévy processes with finite and infinite activity. We will first briefly summarize existing methods for SDEs with noise driven by Wiener processes and then give an extensive overview of literature in regard to finding weak approximations for the solution of Dirichlet IPDE problems and the corresponding SDEs driven by Lévy processes with finite and infinite activity.

First of all, there is an extensive range of research on numerical schemes for SDEs [88, 69, 82, 91, 100, 60, 103, 93, 18, 72, 97, 70, 71, 90, 79] and references therein, and it is necessary to highlight that our research focuses on Dirichlet IPDE problems, whereas there is also some references made to the Cauchy problem. The literature we review here covers both.

There is a wide range of literature addressing numerical methods and approximations for SDEs, [69] gives numerous schemes (e.g. Euler Scheme, Order 2.0 Weak Taylor scheme) which deal with the case of SDEs driven by a wide range of Lévy processes and give a broad overview of strong and weak approximations. The authors of [92] describe a whole range of numerical schemes for weak (and strong) approximations for SDEs driven by Wiener processes. Moreover, they introduce various numerical methods for SDEs which are suitable to solve Dirichlet problems (boundary value problems) and also Chauchy problems (initial value problems). In particular, an algorithm (Algorithm 2.1) for a simple random walk [91], which we will extend to a general IPDE problem where the underlying noise is modelled by Lévy processes with finite and infinite activity (see Chapter 5).

There has been a considerable amount of research in weak-sense numerical methods for Lévy-type SDEs of finite and infinite activity (see [100, 60, 103, 93, 18, 72, 70, 71] and references therein), which follow similar approaches as our research. A brief summary of literature for numerical methods for IPDE problems can be found in Table 4.2.1. Protter and Talay [100] and Jacod et al. [60] follow the traditional approach approximating the Lévy process using an Euler scheme with a uniform grid. One problem of that approach is that for general Lévy processes, there are no efficient algorithms to simulate the increments of the Lévy process and secondly due to the use of a fixed grid, the discretization error between two points can become large in cases of large jumps. For Lévy processes of finite activity, Rubenthaler [103], Mordecki at el [93] and other authors [18] introduce the idea of replacing the jump part of the Lévy process with a suitable compound Poisson approximation. In addition, they place the discretization points of the Euler scheme at the jump times of the compound process to solve this problem. However, for Lévy processes with large jump intensity this can cause problems due to the singularity of the Lévy measure at zero. For the case of a Lévy process with no diffusion part, Kohatsu-Higa and Tankov [72] develop the idea of Rubenthaler [103] further and make use of the approach of Asmussen and Rosinski [5], which approximates small jumps with an appropriate Brownian motion between the jump times. Additionally, they replace the approximation of the solution of the continuous SDE between the jump times with a suitable approximation, which leads to a lower discretization error compared to [103].

Our approach is most closely related to [71], where Kohatsu-Higa et al. introduce a more general class of high order approximation schemes for Lévy processes with infinite activity, with the objective to design optimal compound Poisson approximations. In contrast to previous work [93, 18, 72], they introduce a scheme for Lévy process with a non-degenerate Brownian motion part and combine developments for high order approximations of the Brownian component of weak approximations for continuous SDEs (such as [94]) with suitable jump adapted approximation schemes for pure jump SDEs. Moreover, instead of following the idea of Asmussen and Rosinsiki [5], they follow a moment-matching approach of [110] which introduces an additional compound Poisson term to approximate the Lévy process by a finite intensity Lévy process which incorporates all jumps lower than a certain threshold. The main error estimate in [71] is dependent on the intensity of the compound Poisson process λ_ϵ, which only considers jumps larger than ϵ. However, this dependency on epsilon is not explicitly given. As we will see in Section 5, for certain choices of ϵ, in particular in the case for infinite activity, the convergence can be very slow.

In this work, we combine the probabilistic representation approach (Feynman-Kac formula) for the Dirichlet IPDE problem with the development of a new numerical scheme (Algorithm 1) for a general class of Lévy-type processes. We extend the ideas of [92] for Brownian motion to find weak approximation schemes for the Lévy case. The idea of solving Dirichlet IPDE problems following the Feynman-Kac approach, means that we need to simulate trajectories for the corresponding SDEs driven by Lévy processes with finite and infinite activity. As in [5, 72, 71], we replace small jumps (smaller than ϵ) with an appropriate Brownian motion, which makes the numerical solution of SDEs with infinite activity of the Lévy measure feasible in practice, this also allows us to overcome the computationally difficulty of simulating trajectories for these SDEs in the infinite activity case.

There are three main differences between our approach and that of [71]. First, we use restricted jump-adapted time-stepping while in [71] jump-adapted time-stepping was used. When speaking about jump-adapted time-stepping, we mean that time discretization points are located at jump times τ_k and between the jumps the remaining diffusion process is effectively approximated [72, 71]. By restricted jump-adapted time-stepping, we understand the following. We fix a time-discretization step $h > 0$. If the jump time increment δ for the next time step is less than h, we set the time increment $\theta = \delta$, otherwise $\theta = h$, i.e., our time steps are defined as $\theta = \delta \wedge h$. We highlight that this is a different time-stepping strategy to commonly used ones in the literature including the finite-activity case (i.e., jump-diffusion). For example, in the finite activity case it is common [82, 93, 97] to simulate τ_k before the start of simulations and then superimpose those random times on a grid with some constant or variable finite, small time-step h. Our time-stepping approach is more natural for the problem under consideration than both commonly used strategies; its benefits are discussed in Section 5.3, with the infinite activity case considered in more detail in Subsections 5.3.4 and 6.3. It is beneficial for accuracy restricting δ by h, when jumps are rare (e.g. in the jump-diffusion case) and it is also beneficial for convergence rates (measured in the average number of steps) in the case of α-stable Lévy measures with $\alpha \in (1, 2)$ (see Sections 5.3 and 6). Additionally, by assuring that the maximum step-size for one step is limited by h, we can avoid large discretization errors in the Brownian motion part when jump times are far apart.

Second, in comparison to [71, 70] we explicitly show (singular) dependence of the numerical integration error of our algorithm on the parameter ϵ which is the cut-off for small jumps replaced by the Brownian motion. While the dependency of their error estimate mentions that it is dependent on ϵ, this dependency is not specified further. Here, we highlight the dependency clearly in Section 5.3 and also analyse this singularity in depth numerically in Chapter 6.

Third, in comparison with the literature we consider the Dirichlet problem for IPDEs, though we also comment briefly on the Cauchy case on some occasions. While the majority of the mentioned literature [100, 60, 103, 93] do not address the connection of IPDE problems via the Feynman-Kac formula, practically, it is a very useful connection, in particular with regard to related problems in Finance, which we also discuss in Section 6.4.

Additionally, instead of following the approach using the Feynman-Kac formula, one could also solve the IPDE problem (4.1.1) following the approach of finite difference methods [3, 24, 84, 114]. The methods developed in Andersen and Andreasen [3] introduce extensions to the approach of Dupire [37] which shows important model improvements, in particular in regard to the implied volatility surface modelling. However, their approach is limited to jump-diffusion models with finite activity. Cont and Volchkova [24] describe a finite difference scheme for an IPDE problem in regard to option pricing theory. They propose an implicit finite difference scheme, which in contrast to [3] shows a more rigorous analysis of consistency, stability, and convergence and further, they also look at Lévy processes involving infinite activity.

Another numerical method to approximate the solution to the initial IPDE problem is by applying Fourier transform algorithms [74, 80]. The authors [74] suggest that their approach using Wiener-Hopf factorization and Fast Fourier Transform algorithms works well for Lévy processes with finite and infinite activity. Compared to this, where we focus on a more general case and handling higher dimensional problems, they focus in particular on problems without diffusion process and on the case of lower dimension (i.e. $d = 1$).

Depending on the underlying Lévy process, the Monte Carlo simulation can be computationally expensive, one could also explore numerical optimization approaches [13, 99, 66, 67, 68] which aim to find bounds for the corresponding expectation introduced through the Feynman-Kac formula. In [66], the authors propose an optimization approach, where they use a mathematical programming framework to compute upper and lower bounds of the target expectation, whereby they first bound the considered expectation at maturation T from one side and then optimise (i.e. minimise or maximise) the other bound. While the authors need to make some limiting assumptions on the functional form of the underlying functions and require the Lévy measure to be in closed form, they do not require to simulate sample paths of the underlying Lévy process, and hence do not need to have the exact knowledge of increment distributions. In contrast, our approach requires knowledge of the underlying distributions to simulate the sample paths with jumps. Using their approach, it is particularly useful in settings where it is too expensive computationally or not possible to run MC simulations. The authors of [67], also extend their general approach [66], by employing tempering on bounding functions to avoid the polynomial explosion to overcome a required moment condition, which ruled out SDEs driven by stable Lévy processes, which can be dealt with in the approach presented in this thesis.

Numerical schemes for Lévy type SDEs									
Literature	time step			Lévy type			IPDE problem		Small jump **T**runcation/ **R**eplacement
	fixed (h)	jump times (δ)	jump–adaptive $(h \wedge \delta)$	diffusion	jump (finite act.)	jump (infinite act.)	Cauchy	Dirichlet	
Kloeden and Platen [69]	X			X			X		
Milstein and Tretyakov [92]	X			X			X	X	
Protter and Talay [100]	X				X				
Jacod et al. [60]	X			X	X	X			T
Liu and Li[82]	X			X	X			X	
Rubenthaler [103]			X	X	X	X			T
Mordecki et al [93]			X	X	X				
Platen and Bruti-Liberati [18]			X	X	X				
Kohatsu-Higa and Tankov [72]		X			X	X			R
Tankov [110]			X		X	X			R
Kohatsu-Higa and Ngo [70]		X	X	X	X	X			R
Kohatsu-Higa et al. [71]			X	X	X	X			R
Deligiannidis et al. (our approach) [30]			X	X	X	X	ideas given	X	R

Table 4.2.1: A structured summary of literature for numerical methods for IPDE problems.'X' suggests that the research paper covers the topic.

5

A SIMPLEST RANDOM WALK FOR SOLVING THE DIRICHLET PROBLEM FOR IPDES

In this chapter, we introduce, prove and discuss the main theoretical convergence results of the proposed restricted jump-adaptive time stepping algorithm for the introduced IPDE problem in Section 4.1.

We first introduce some necessary preliminaries in Section 5.1. This is then followed by an important ingredient of the method in Section 5.2, where we approximate small jumps by a diffusion process. In the last Section 5.3 of this chapter, we present the introduced algorithm, investigate the one-step error and global error of said algorithm and prove weak convergence with particular focus on the case of infinite intensity of jumps. The suggested algorithm and in-depth analysis are published in [30].

5.1 PRELIMINARIES TO THE IPDE PROBLEM (4.1.1)

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t_0 \leq t \leq T}, P)$ be a filtered probability space. The operator L defined in (4.1.2), on a bounded domain, is the generator of the d-dimensional process $X_{t_0,x}(t)$ given by

$$X_{t_0,x}(t) = x + \int_{t_0}^{t} b(s, X(s-))\mathrm{d}s + \int_{t_0}^{t} \sigma(s, X(s-))\mathrm{d}w(s) \qquad (5.1.1)$$
$$+ \int_{t_0}^{t} \int_{\mathbb{R}^d} F(s, X(s-))z\hat{N}(\mathrm{d}z, \mathrm{d}s),$$

where the $d \times d$ matrix $\sigma(s,x)$ is defined through $\sigma(s,x)\sigma^\top(s,x) = a(s,x)$; $w(t) = (w^1(t), \dots, w^d(t))^\top$ is a standard d-dimensional Wiener process; and $\hat{N}$ is a Poisson random measure on $[0,\infty) \times \mathbb{R}^m$ with intensity measure $\nu(\mathrm{d}z) \times \mathrm{d}s$, $\int_{\mathbb{R}^m} (|z|^2 \wedge 1)\nu(\mathrm{d}z) < \infty$, and compensated small jumps, i.e.,

$$\hat{N}\left([0,t] \times B\right) = \int_{[0,t]\times B} N(\mathrm{d}z, \mathrm{d}s) - t\nu(B \cap \{|z| \le 1\}),$$

for all $t \ge 0$ and $B \in \mathcal{B}\big(\mathbb{R}^m\big)$.

Remark 5.1.1. Often [4, 100] a simpler model of the form

$$X(t) = x + \int_{t_0}^{t} F(s, X(s-))\mathrm{d}Z(s), \tag{5.1.2}$$

where $Z(t)$, $t \ge t_0$, is an m-dimensional Lévy process with the characteristic exponent

$$\begin{aligned}\Psi(\xi) = \mathrm{i}\langle \mu, \xi\rangle - \frac{1}{2}\langle \xi, \sigma\xi\rangle + \int_{|z|\le 1} \left[\mathrm{e}^{\mathrm{i}\langle \xi, z\rangle} - 1 - \mathrm{i}\langle \xi, z\rangle\right]\nu(\mathrm{d}z) \\ + \int_{|z|>1} \left[\mathrm{e}^{\mathrm{i}\langle \xi, z\rangle} - 1\right]\nu(\mathrm{d}z),\end{aligned}$$

is considered instead of the general SDEs (5.1.1). The equation (5.1.2) is obtained as a special case of (5.1.1) by setting $b(t,x) = \mu F(t,x)$ and $\sigma(t,x) = \sigma F(t,x)$.

Therefore, we can see that if one can simulate trajectories of

$$\{(s, X_{t,x}(s), Y_{t,x,1}(s), Z_{t,x,1,0}(s)); s \ge 0\}$$

then the solution of the Dirichlet problem for IPDE (4.1.1) can be estimated by applying the Monte Carlo technique to (4.1.3). This approach however is not generally implementable for Lévy measures of infinite intensity, that is when $\nu\big(B(0,r)\big) = \infty$ for some $r > 0$. The difficulty arises from the presence of an infinite number of small jumps in any finite time interval, and can be overcome by replacing these small jumps by an appropriate diffusion exploiting the idea of the method developed in [72, 5], which we apply here. Alternatively, the issue can be overcome if one can simulate directly from the increments of Lévy processes. We will not discuss this case in this research as we only assume that one has access to the Lévy measure.

5.2 APPROXIMATION OF SMALL JUMPS BY DIFFUSION

We will now consider the approximation of (5.1.1) discussed above, where small jumps are replaced by an appropriate diffusion. In the case of the whole space (the Cauchy problem for a IPDE) such an approximation was considered in [72, 5], or see also [30][Sec. 3.4.], but we only consider the Dirichlet problem here.

Let γ_ϵ be an m-dimensional vector with the components

$$\gamma_\epsilon^i = \int_{\epsilon \le |z| \le 1} z^i \nu(\mathrm{d}z); \tag{5.2.1}$$

and B_ϵ is an $m \times m$ matrix with the components

$$B_\epsilon^{ij} = \int_{|z|<\epsilon} z^i z^j \nu(\mathrm{d}z), \tag{5.2.2}$$

while β_ϵ be obtained from the formula $\beta_\epsilon \beta_\epsilon^\top = B_\epsilon$. Note that $|B_\epsilon^{ij}|$ (and hence also the elements of β_ϵ) are bounded by a constant independent of ϵ thanks to the Lévy measure definition.

Remark 5.2.1. In many practical situations (see e.g. [23]), where the dependence among the components of $X(t)$ introduced through the structure of the SDEs is enough, we can allow the components of the driving Poisson measure to be independent. This amounts to saying that ν is concentrated on the axes, and as a result B_ϵ will be a diagonal matrix.

We shall consider the modified jump-diffusion $\widetilde{X}_{t_0,x}(t) = \widetilde{X}^{\epsilon}_{t_0,x}(t)$ defined as

$$\widetilde{X}_{t_0,x}(t) = x + \int_{t_0}^{t} \left[b(s, \widetilde{X}(s-)) - F(s, \widetilde{X}(s-))\gamma_\epsilon \right] \mathrm{d}s + \int_{t_0}^{t} \sigma(s, \widetilde{X}(s-))\mathrm{d}w(s) \tag{5.2.3}$$
$$+ \int_{t_0}^{t} F(s, \widetilde{X}(s-))\beta_\epsilon \mathrm{d}W(s) + \int_{t_0}^{t} \int_{|z|\geq\epsilon} F(s, \widetilde{X}(s-)) z N(\mathrm{d}z, \mathrm{d}s),$$

where $W(t)$ is a standard m-dimensional Wiener process, independent of N and w. We observe that, in comparison with (5.1.1), in (5.2.3) jumps less than ϵ in magnitude are replaced by the additional diffusion part. In this way, the new Lévy measure has finite activity allowing us to simulate its events exactly, i.e. in a practical way.

Therefore, we can approximate the solution of $u(t,x)$ the IPDE (4.1.1) by

$$\begin{aligned} u(t,x) &\approx u^{\epsilon}(t,x) \\ &:= \mathbb{E}\left[\varphi\left(\widetilde{\tau}_{t,x}, \widetilde{X}_{t,x}(\widetilde{\tau}_{t,x}) \right) \widetilde{Y}_{t,x,1}(\widetilde{\tau}_{t,x}) + \widetilde{Z}_{t,x,1,0}(\widetilde{\tau}_{t,x}) \right], \quad (t,x) \in Q, \end{aligned} \tag{5.2.4}$$

where $\widetilde{\tau}_{t,x} = \inf\{s \geq t : (s, \widetilde{X}_{t,x}(s)) \notin Q\}$ is the first exit time of the space-time Lévy process $(s, \widetilde{X}_{t,x}(s))$ from the space-time cylinder Q and $\left(\widetilde{X}_{t,x}(s), \widetilde{Y}_{t,x,y}(s), \widetilde{Z}_{t,x,y,z}(s)\right)_{s\geq 0}$ solves the system of SDEs consisting of (5.2.3) along with

$$d\widetilde{Y} = c(s, \widetilde{X}(s-))\widetilde{Y} ds, \quad \widetilde{Y}_{t,x,y}(t) = y, \tag{5.2.5}$$
$$d\widetilde{Z} = g(s, \widetilde{X}(s-))\widetilde{Y} ds, \quad \widetilde{Z}_{t,x,y,z}(t) = z. \tag{5.2.6}$$

Since the new Lévy measure has finite activity, we can derive a constructive weak scheme for (5.2.3), (5.2.5)-(5.2.6) (see Section 5.3). By using this method together with the Monte Carlo technique, we will arrive at an implementable approximation of $u^{\epsilon}(t,x)$ and hence of $u(t,x)$.

We will next show that indeed u^ϵ defined in (5.2.4) is a good approximation to the solution of (4.1.1). Before proceeding, we need to formulate appropriate assumptions.

5.2.1 *Assumptions*

To begin with, we make the following assumptions on the coefficients of the problem (4.1.1) which will guarantee, see e.g. [4], that the SDEs (5.1.1), (4.1.4)-(4.1.5) and (5.2.3), (5.2.5)-(5.2.6) have unique adapted, càdlàg solutions with finite moments.

Assumption 5.2.1. (*Lipschitz condition*) *There exists a constant* $K > 0$ *such that for all* $x_1, x_2 \in \mathbb{R}^d$ *and all* $t \in [t_0, T]$,

$$\begin{aligned}\left\|b(t,x_1) - b(t,x_2)\right\|^2 + \left\|\sigma(t,x_1) - \sigma(t,x_2)\right\|^2 \\ + \|c(t,x_1) - c(t,x_2)\|^2 + \|g(t,x_1) - g(t,x_2)\|^2 \\ + \int_{\mathbb{R}^d} \|F(t,x_1) - F(t,x_2)\|^2 |z|^2 \nu(\mathrm{d}z) \le K\|x_1 - x_2\|^2. \end{aligned} \tag{5.2.7}$$

Assumption 5.2.2. (*Growth condition*) *There exists a constant* $K > 0$ *such that for all* $x \in \mathbb{R}^d$ *and all* $t \in [t_0, T]$,

$$\left\|b(t,x)\right\|^2 + \left\|\sigma(t,x)\right\|^2 + \|g(t,x)\|^2 + \int_{\mathbb{R}^d} \|F(t,x)\|^2 |z|^2 \nu(\mathrm{d}z) \le K(1 + \|x\|)^2, \tag{5.2.8}$$

$$\|c(t,x)\| \le K. \tag{5.2.9}$$

Remark 5.2.2. Since G is bounded, in practice the above assumptions in the space variable are only required in $\bar{G}$. We chose to impose them in $\mathbb{R}^d$ to simplify the presentation as it allows us to construct a global solution to the SDEs (5.2.3), rather than having to deal with local solutions built up to the exit time from the domain. In practice the assumption can be bypassed by multiplying the coefficients with a bump function that vanishes outside G, without affecting the value of (4.1.3).

In order to streamline the presentation and avoid lengthy technical discussions (see Remarks 5.2.3 and 5.2.4), we will make the following assumption regarding the regularity of solutions to (4.1.1).

Assumption 5.2.3. *The Dirichlet problem* (4.1.1) *admits a classical solution* $u(\cdot,\cdot) \in C^{l,n}([t_0,T] \times \mathbb{R}^d)$ *with some* $l \geq 1$ *and* $n \geq 2$.

In addition to the IPDE problem (4.1.1), we also consider the IPDE problem for u^ϵ from (5.2.4):

$$\frac{\partial u^\epsilon}{\partial t} + L_\epsilon u^\epsilon + c(t,x)u^\epsilon + g(t,x) = 0, \quad (t,x) \in Q, \qquad (5.2.10)$$
$$u^\epsilon(t,x) = \varphi(t,x), \quad (t,x) \in Q^c,$$

where

$$L_\epsilon v(t,x) := \frac{1}{2}\sum_{i,j=1}^{d}\left[a^{ij}(t,x) + \Big(F(t,x)B_\epsilon(t,x)F^\top(t,x)\Big)^{ij}\right]\frac{\partial^2 v}{\partial x^i \partial x^j}(t,x) \qquad (5.2.11)$$
$$+ \sum_{i=1}^{d}\Big(b^i(t,x) - \sum_{j=1}^{m} F^{ij}(t,x)\gamma_\epsilon^j\Big)\frac{\partial v}{\partial x^i}(t,x)$$
$$+ \int_{|z|\geq\epsilon}\Big\{v\big(t,x+F(t,x)z\big) - v(t,x)\Big\}\nu(\mathrm{d}z).$$

Again, for simplicity (but see Remark 5.2.3), we impose the following conditions on the solution u^ϵ of the above Dirichlet problem.

Assumption 5.2.4. *The auxiliary Dirichlet problem* (5.2.10) *admits a classical solution* $u^\epsilon(\cdot,\cdot) \in C^{l,n}([t_0,T] \times \mathbb{R}^d)$ *with some* $l \geq 1$ *and* $n \geq 2$.

Finally, we also require that u^ϵ and its derivatives do not grow faster than a polynomial function at infinity.

Assumption 5.2.5 (*Smoothness and growth*)**.** There exist constants $K > 0$ and $q \geq 1$ such that for all $x \in \mathbb{R}^d$, all $t \in [t_0, T]$ and $\epsilon > 0$, the solution u^ϵ of the IPDE problem (5.2.10) and its derivatives satisfy

$$\left\| \frac{\partial^{l+j}}{\partial t^l \partial x^{i_1} \cdots \partial x^{i_j}} u^\epsilon(t,x) \right\| \leq K(1 + \|x\|^q), \tag{5.2.12}$$

where $0 \leq 2l + j \leq 4$, $\sum_{k=1}^{j} i_k = j$, and i_k are integers from 0 to j.

Remark 5.2.3. Sufficient conditions guaranteeing Assumptions 5.2.3, 5.2.4 and 5.2.5 consist in sufficient smoothness of the coefficients, the boundary ∂G, and the function φ and in appropriate compatibility of φ and g and also of the integral operator (see e.g. [45, 59, 89]).

Remark 5.2.4. The main goal of this research is to present the numerical method and study its convergence under 'good' conditions when its convergence rates are optimal (i.e., highest possible). As usual, in these circumstances, the conditions (here Assumptions 5.2.3, 5.2.4, and 5.2.5) are somewhat restrictive. See Theorem 3.3 in [45, p. 93], which indicates sufficient conditions for Assumption 5.2.3 to hold. If one drops the compatibility condition (3.11) in Theorem 3.3 of [45, p. 93], then, as in the diffusion case, the smoothness of the solution will be lost through the boundary of Q at the terminal time T. This affects only the last step of the method and the proof can be modified (see such a recipe in the case of the Neumann problem and diffusion in e.g. [79]), but we do not include such complications here for transparency of the proofs. Further, in the case of an α-stable Lévy process with $\alpha \in (1, 2)$ spatial derivatives of $u(t, x)$ may blow up near the boundary ∂G, the blow up is polynomial with the power dependent on α if the integral operator does not satisfy some compatibility conditions (see the discussion in [45, p. 96]). This situation requires further analysis of the proposed method, which is beyond the scope of this research.

5.2.2 *Closeness of $u^\epsilon(t,x)$ and $u(t,x)$*

In this section, we now state and refer to a proof for the theorem on closeness of $u^\epsilon(t,x)$ and $u(t,x)$. In what follows we use the same letters K and C for various positive constants independent of x, t, and ϵ.

Theorem 5.2.5. *Let Assumptions 5.2.1, 5.2.2 and 5.2.3 hold, the latter with $l = 1$ and $m = 3$. Then for $0 \le \epsilon < 1$*

$$|u^\epsilon(t,x) - u(t,x)| \le K \int_{|z|\le\epsilon} |z|^3 \nu(\mathrm{d}z), \quad (t,x) \in Q, \tag{5.2.13}$$

where $K > 0$ does not depend on t, x, ϵ.

We omit the proof for this theorem here, but refer to it in [30][Thm 2.1].

Example 5.2.1 (Tempered α-stable Process)**.** For $\alpha \in (0,2)$ and $m = 1$, consider an α-stable process with Lévy measure given by $\nu(\mathrm{d}z) = |z|^{-1-\alpha}\mathrm{d}z$. Then

$$\int_{|z|\le\epsilon} |z|^3 \nu(\mathrm{d}z) = 2\frac{\epsilon^{3-\alpha}}{3-\alpha}.$$

Similarly, for a tempered stable distribution which has Lévy measure given by

$$\nu(\mathrm{d}z) = \Big(\frac{C_+ \mathrm{e}^{-\lambda_+ z}}{z^{1+\alpha}} \mathbf{I}(z > 0) + \frac{C_- \mathrm{e}^{-\lambda_- |z|}}{|z|^{1+\alpha}} \mathbf{I}(z < 0)\Big)\mathrm{d}z,$$

for $\alpha \in (0,2)$ and C_+, C_-, λ_+, $\lambda_- > 0$ we find that the error from approximating the small jumps by diffusion as in Theorem 5.2.5 is of the order $O(\epsilon^{3-\alpha})$.

5.3 WEAK APPROXIMATION OF JUMP-DIFFUSIONS IN BOUNDED DOMAINS

In this section we suggest and investigate a numerical algorithm which weakly approximates the solutions of the jump-diffusion (5.2.3), (5.2.5)-(5.2.6) with finite intensity of jumps in a bounded domain, i.e. approximates $u^\epsilon(t,x)$

from (5.2.4). In the first part, in Section 5.3.1 we formulate the algorithm based on a simplest random walk. Then, we analyse the one-step error of the algorithm in Section 5.3.2 and the global error in Section 5.3.3. In Section 5.3.4 we combine the convergence result of Section 5.3.3 with Theorem 5.2.5 to get error estimates in the case of infinite activity of jumps.

5.3.1 *Algorithm*

In what follows we also require the following to hold.

Assumption 5.3.1 (Lévy measure). *There exists a constant $K > 0$*

$$\int_{\mathbb{R}^m} |z|^p \nu(dz) \leq K$$

for up to a sufficiently large $p \geq 2$.

This is a natural assumption since Lévy measures of practical interest (see e.g. [23] and examples here in Example 5.2.1 and Section 6) have this property.

Let us describe an algorithm for simulating a Markov chain that approximates a trajectory of (5.2.3), (5.2.5)-(5.2.6). In what follows we assume that we can exactly sample the intervals δ between consecutive jump times with the intensity

$$\lambda_\epsilon := \int_{|z|>\epsilon} \nu(\mathrm{d}z) \tag{5.3.1}$$

and jump sizes J_ϵ distributed according to the density

$$\rho_\epsilon(z) := \frac{\nu(z)\mathbf{I}(|z| > \epsilon)}{\lambda_\epsilon}. \tag{5.3.2}$$

Remark 5.3.1. There are known methods for simulating jump times and sizes for many standard distributions. In general, if there exists an explicit expression for the jump size density, one can construct a rejection method to sample jump sizes. An overview with regard to simulation of jump times and sizes can be found in [23, 32].

Thanks to Assumption 5.3.1, we have

$$\mathbb{E}\left[|J_\epsilon|^p\right] \equiv \frac{1}{\lambda_\epsilon}\int_{|z|>\epsilon}|z|^p\nu(dz) \leq \frac{K}{\lambda_\epsilon} \tag{5.3.3}$$

with $K > 0$ being independent of ϵ and $p \geq 2$. We also note that

$$\frac{|\gamma_\epsilon|^2}{\lambda_\epsilon} \leq K, \tag{5.3.4}$$

where $K > 0$ is a constant independent of ε, since by the Cauchy-Schwarz inequality

$$\begin{aligned}\frac{|\gamma_\epsilon|^2}{\lambda_\epsilon} &\leq \left(\int_{\epsilon<|z|<1}\frac{|z|}{\sqrt{\lambda_\epsilon}}\nu(\mathrm{d}z)\right)^2 \leq \int_{\epsilon<|z|<1}\frac{|z|^2}{\lambda_\epsilon}\nu(\mathrm{d}z)\times\lambda_\epsilon\\ &\leq \int_{0<|z|<1}|z|^2\nu(\mathrm{d}z) < \infty\end{aligned}$$

thanks to the Lévy measure definition.

We can now describe the algorithm. Fix a time-discretization step $h > 0$ and suppose the current position of the chain is (t, x, y, z). If the jump time increment $\delta < h$, we set $\theta = \delta$, otherwise $\theta = h$, i.e. $\theta = \delta \wedge h$.

In the case $\theta = h$, we apply the weak explicit Euler approximation with the simplest simulation of noise to the system (5.2.3), (5.2.5)-(5.2.6) with no jumps:

$$\widetilde{X}_{t,x}(t+\theta) \approx X = x + \theta\cdot\left(b(t,x) - F(t,x)\gamma_\epsilon\right) + \sqrt{\theta}\cdot\left(\sigma(t,x)\,\xi + F(t,x)\beta_\epsilon\,\eta\right), \tag{5.3.5}$$

$$\widetilde{Y}_{t,x,y}(t+\theta) \approx Y = y + \theta\cdot c(t,x)\,y\,, \tag{5.3.6}$$

$$\widetilde{Z}_{t,x,y,z}(t+\theta) \approx Z = z + \theta\cdot g(t,x)\,y\,, \tag{5.3.7}$$

where $\xi = (\xi^1,\ldots,\xi^d)^\intercal$, $\eta = (\eta^1,\ldots,\eta^m)^\intercal$, with $\xi^1,\ldots,\xi^d$ and $\eta^1,\ldots,\eta^m$ mutually independent random variables, taking the values ±1 with equal probability. In the case of $\theta < h$, we replace (5.3.5) by the following explicit Euler approximation

$$\widetilde{X}_{t,x}(t+\theta) \approx X = x + \theta \cdot (b(t,x) - F(t,x)\gamma_\epsilon) + \sqrt{\theta} \cdot (\sigma(t,x)\,\xi + F(t,x)\beta_\epsilon\,\eta) + F(t,x)J_\epsilon. \tag{5.3.8}$$

Let $(t_0, x_0) \in Q$. We aim to find the value $u^\epsilon(t_0, x_0)$, where $u^\epsilon(t,x)$ solves the problem (5.2.10). Introduce a discretization of the interval $[t_0, T]$, for example the equidistant one:

$$h := (T - t_0)/L.$$

To approximate the solution of the system (5.2.3), we construct a Markov chain $(\vartheta_k, X_k, Y_k, Z_k)$ which stops at a random step $\varkappa$ when (ϑ_k, X_k) exits the domain Q. The algorithm is formulated as Algorithm 1 below.

Remark 5.3.2. If λ_ϵ is large so that $1 - e^{-\lambda_\epsilon h}$ is close to 1, then $I_k = 1$ (i.e., jump happens) is almost on every time step. In this situation it is computationally beneficial to modify Algorithm 1 in the following way: instead of sampling both I_k and θ_k, sample δ_k according to the exponential distribution with parameter λ_ϵ and set $\theta_k = \delta_k \wedge h$ and $I_k = 0$ if $\theta_k < h$, else $I_k = 0$.

Remark 5.3.3. We note [91, 92] that in the diffusion case (i.e., when there is no jump component in the noise which drives SDEs) solving Dirichlet problems for parabolic or elliptic PDEs requires to complement a random walk inside the domain G with a special approximation near the boundary ∂G. In contrast, in the case of Dirichlet problems for IPDEs we do not need a special construction near the boundary since the boundary condition is defined on the whole complement G^c. Here, when the chain X_k exits G, we know the exact value of the solution $u^\epsilon(\bar{\vartheta}_\varkappa, X_\varkappa) = \varphi(\bar{\vartheta}_\varkappa, X_\varkappa)$ at the exit point $(\bar{\vartheta}_\varkappa, X_\varkappa)$, while in the diffusion case when a chain exits G, we do not know the exact value of the solution at the exit point and need an

Algorithm 1 Algorithm for (5.2.3), (5.2.5)-(5.2.6).

Output: $\bar{\vartheta}_\varkappa, X_\varkappa, Y_\varkappa, Z_\varkappa$

Initialize: $\vartheta_0 = t_0,\ X_0 = x_0,\ Y_0 = 1,\ Z_0 = 0,\ k = 0.$
while $\vartheta_k < T$ or $X_k \in G$ **do**
  Simulate: ξ_k and η_k with i.i.d. components taking values ± 1 with probability $1/2$ and independently $I_k \sim \text{Bernoulli}(1 - e^{-\lambda_\epsilon h})$.
  if $I_k = 0$, **then**
    Set: $\theta_k = h$
    Evaluate: $X_{k+1},\ Y_{k+1},\ Z_{k+1}$ according to $(5.3.5) - (5.3.7)$ with $t = \vartheta_k, \theta = \theta_k, \xi = \xi_k, \eta = \eta_k, x = X_k, y = Y_k, z = Z_k$.
  else
    Sample: θ_k according to the density $\dfrac{\lambda_\epsilon e^{-\lambda_\epsilon x}}{1 - e^{-\lambda_\epsilon h}}$ with finite support $[0, h]$.
    Sample: jump size $J_{\epsilon,k}$ according to the density (5.3.2).
    Evaluate: X_{k+1}, Y_{k+1} and Z_{k+1} according to (5.3.8), (5.3.6), (5.3.7) with $t = \vartheta_k, \theta = \theta_k, \xi = \xi_k, \eta = \eta_k, J_\epsilon = J_{\epsilon,k}, x = X_k, y = Y_k, z = Z_k$.
  end if
  Set: $\vartheta_{k+1} = \vartheta_k + \theta_k$ and $k = k + 1$.
end while
Set: $X_\varkappa = X_k, Y_\varkappa = Y_k, Z_\varkappa = Z_k, \varkappa = k, \vartheta_\varkappa = \vartheta_k$.
if $\vartheta_\varkappa < T$ **then Set:** $\bar{\vartheta}_\varkappa = \vartheta_\varkappa$
else Set: $\bar{\vartheta}_\varkappa = T$
end if

approximation. Due to this fact, Algorithm 1 is somewhat simpler than algorithms for Dirichlet problems for parabolic or elliptic PDEs (cf. [91, 92] and references therein).

5.3.2 *One-step error*

In this section we consider the one-step error of Algorithm 1. The one step of this algorithm takes the form for $(t, x) \in Q$:

$$X = x + \theta\,(b(t,x) - F(t,x)\gamma_\epsilon) + \sqrt{\theta}\,(\sigma(t,x)\xi + F(t,x)\beta_\epsilon\eta) + \mathbf{I}(\theta < h)F(t,x)J_\epsilon, \tag{5.3.9}$$

$$Y = y + \theta c(t,x)y, \tag{5.3.10}$$

$$Z = z + \theta g(t,x)y. \tag{5.3.11}$$

Before we state and prove an error estimate for the one-step of Algorithm 1, we need to introduce some additional notation. For the ease of notation let us write $b = b(t,x)$, $\sigma = \sigma(t,x)$, $F = F(t,x)$, $g = g(t,x)$, $c = c(t,x)$, $J = J_\epsilon$. Let us define the intermediate points Q_i and their differences Δ_i, for $i = 1, \ldots, 4$:

$$\begin{aligned}
\Delta_1 &= \theta^{1/2} \left[\sigma\xi + F\beta_\epsilon \eta \right], \qquad (5.3.12)\\
\Delta_2 &= \theta \left[b - F\gamma_\epsilon \right], \\
\Delta_3 &= \mathbf{I}(\theta < h) F J, \\
Q_1 &= x + \Delta_1 + \Delta_2 + \Delta_3 = X, \\
Q_2 &= x + \Delta_2 + \Delta_3, \\
Q_3 &= x + \Delta_3, \\
Q_4 &= x,
\end{aligned}$$

where $x \in G$. Note that Q_i, $i = 1, \ldots, 3$, can be outside G.

Lemma 5.3.4 (Moments of intermediate points Q_i)**.** *Under Assumptions 5.2.1 and 5.3.1, there is $K > 0$ independent of ϵ and h such that for $p \geq 1$:*

$$\mathbb{E}\left[|Q_i|^{2p} | \theta, t, x \right] \leq K(1 + \theta^{2p} |\gamma_\epsilon|^{2p}), \ i = 1, 2, \qquad (5.3.13)$$

$$\mathbb{E}\left[|Q_i|^{2p} | \theta, t, x \right] \leq K, \ i = 3, 4, \qquad (5.3.14)$$

where Q_i are defined in (5.3.12).

Proof. It is not difficult to see that the points Q_i, $i = 1, 2$, are of the following form

$$Q_i = x + c_1 \theta^{1/2} \left[\sigma(t,x)\xi + F(t,x)\beta_\epsilon \eta \right] + \theta \left[b(t,x) - F(t,x)\gamma_\epsilon \right] + \mathbf{I}(\theta < h) F(t,x) J_\epsilon,$$

where c_1 is either 0 or 1. It is obvious that ξ and η and their moments are all bounded. The functions $b(t,x)$, $\sigma(t,x)$ and $F(t,x)$ are bounded as $(t,x) \in Q$, and for $x \in G$, $|x|^{2p}$ is also bounded. Recall that sufficiently high moments of J_ϵ are bounded as in (5.3.3). Then, using the Cauchy-Schwarz inequality, we can show that

$$\begin{aligned}\mathbb{E}\left[|Q_i|^{2p}|\theta, t, x\right] &\leq |x|^{2p} + K\theta^p + K\theta^{2p}\left[1 + |\gamma_\epsilon|^{2p}\right] + K\mathbf{I}(\theta < h)\mathbb{E}\left[|J_\epsilon|^{2p}\right] \\ &\leq K(1 + \theta^{2p}|\gamma_\epsilon|^{2p}).\end{aligned}$$

Hence, we obtained (5.3.13). The bound (5.3.14) is shown analogously. □

We will need the following technical lemma.

Lemma 5.3.5 (Moments of θ)**.** *For integer $p \geq 2$, we have*

$$\mathbb{E}\left[\theta^p\right] \leq K\frac{1 - e^{-\lambda_\epsilon h}(1 + \lambda_\epsilon h)}{\lambda_\epsilon^p}, \tag{5.3.15}$$

where $K > 0$ depends on p but is independent of λ_ϵ and h.

Proof. The proof is by induction. By straightforward calculations, we get

$$\begin{aligned}E\left[\theta^2\right] &= \lambda_\epsilon \int_0^h t^2 e^{-\lambda_\epsilon t}dt + \lambda_\epsilon \int_h^\infty h^2 e^{-\lambda_\epsilon t}dt \\ &= \left[-h^2 e^{-\lambda_\epsilon t}\right]_{t=h}^{t=\infty} + \left[-t^2 e^{-\lambda_\epsilon t}\right]_{t=0}^{t=h} + 2\int_0^h t e^{-\lambda_\epsilon t}dt \\ &= 2\left(\left[-\frac{t}{\lambda_\epsilon}e^{-\lambda_\epsilon t}\right]_{t=0}^{t=h} + \frac{1}{\lambda_\epsilon}\int_0^h e^{-\lambda_\epsilon t}dt\right) \\ &= 2\left(-\frac{h}{\lambda_\epsilon}e^{-\lambda_\epsilon h} - \frac{1}{(\lambda_\epsilon)^2}\left[e^{-\lambda_\epsilon t}\right]_{t=0}^{t=h}\right) \\ &= 2\frac{1 - e^{-\lambda_\epsilon h}(1 + \lambda_\epsilon h)}{\lambda_\epsilon^2}.\end{aligned}$$

Then assuming that (5.3.15) is true for some integer $p \geq 2$, we obtain

$$\begin{aligned}\mathbb{E}\left[\theta^{p+1}\right] &= \lambda_\epsilon \int_0^h t^{p+1}e^{-\lambda_\epsilon t}dt + h^{p+1}\lambda_\epsilon \int_h^\infty e^{-\lambda_\epsilon t}dt \\ &= (p+1)\int_0^h t^p e^{-\lambda_\epsilon t}dt \leq \frac{p+1}{\lambda_\epsilon}\left[\lambda_\epsilon \int_0^h t^p e^{-\lambda_\epsilon t}dt + h^p \lambda_\epsilon \int_h^\infty e^{-\lambda_\epsilon t}dt\right] \\ &= \frac{p+1}{\lambda_\epsilon}\mathbb{E}\left[\theta^p\right] \leq K\,(p+1)\,\frac{1 - e^{-\lambda_\epsilon h}(1 + \lambda_\epsilon h)}{\lambda_\epsilon^{p+1}}.\end{aligned}$$

Therefore, as (5.3.15) holds for $p+1$, hence by induction it holds for all $p \geq 2$. □

Now we prove an estimate for the one-step error.

Theorem 5.3.6 (One–step error of Algorithm 1). *Under Assumption 5.2.4 with $l = 2, m = 4$ and Assumptions 5.2.1, 5.2.5 and 5.3.1 the one–step error of Algorithm 1 given by*

$$R(t,x,y,z) := u^{\epsilon}(t+\theta, X)Y + Z - u^{\epsilon}(t,x)y - z$$

satisfies the bound

$$\left|\mathbb{E}[R(t,x,y,z)]\right| \leq K(1+|\gamma_\epsilon|^2)\frac{1-e^{-\lambda_\epsilon h}(1+\lambda_\epsilon h)}{\lambda_\epsilon^2}y, \tag{5.3.16}$$

where $K > 0$ is a constant independent of h and ϵ.

Proof. For any smooth function $v(t,x)$, we write $D_l v_n = (D_l v)(t, Q_n)$ for the l-th time derivative and $(D_l^k v)(t,x)[f_1,\dots,f_k]$ for the l-th time derivative of the k-th spatial derivative evaluated in the directions f_j. For example, if $k=2$ and $l=1$,

$$D_1^2 v[f_1,f_2] = \sum_{i=1}^{d}\sum_{j=1}^{d} f_{1,i}f_{2,j}\frac{\partial^3 v}{\partial t \partial x_i \partial x_j}.$$

We will also use the following short notation

$$D_l^k v_i[f_1,\dots,f_k] := (D_l^k v)(t,Q_i)[f_1,\dots,f_k].$$

The aim of this theorem is to achieve an error estimate explicitly capturing the (singular) dependence of the one-step error on ϵ. Therefore, we split the error into several parts according to the intermediate points Q_i defined in (5.3.12).

Using (5.3.9) and (5.3.12), we have

$$u^{\epsilon}(t+\theta, X) = u^{\epsilon}(t+\theta, Q_1)$$

$$= u^\epsilon\Big(t+\theta, x + \mathbf{I}(\theta < h)FJ + \theta(b - F\gamma_\epsilon) + \theta^{1/2}(\sigma\xi + F\beta_\epsilon\eta)\Big)$$
$$= u^\epsilon\Big(t+\theta, x + \Delta_1 + \Delta_2 + \Delta_3\Big).$$

To precisely account for the factor γ_ϵ and powers of θ in the analysis of the one-step error, we use multiple Taylor expansions of $u^\epsilon(t+\theta, X)$. We obtain

$$\begin{aligned}
u^\epsilon(t+\theta, X) &= u^\epsilon(t, Q_1) + \theta D_1 u_1^\epsilon + R_{11} \qquad (5.3.17)\\
&= u^\epsilon(t, Q_2) + D^1 u_2^\epsilon[\Delta_1] + \frac{1}{2} D^2 u_2^\epsilon[\Delta_1, \Delta_1]\\
&\quad + \frac{1}{6} D^3 u_2^\epsilon[\Delta_1, \Delta_1, \Delta_1] + \theta D_1 u_2^\epsilon + \theta D_1^1 u_2^\epsilon[\Delta_1]\\
&\quad + R_{11} + R_{12} + R_{13}\\
&= u^\epsilon(t, Q_3) + D^1 u_3^\epsilon[\Delta_2] + D^1 u_2^\epsilon[\Delta_1] + \frac{1}{2} D^2 u_3^\epsilon[\Delta_1, \Delta_1]\\
&\quad + \frac{1}{6} D^3 u_2^\epsilon[\Delta_1, \Delta_1, \Delta_1] + \theta D_1 u_3^\epsilon + \theta D_1^1 u_2^\epsilon[\Delta_1] + R_{11} + R_{12}\\
&\quad + R_{13} + R_{14} + R_{15} + R_{16}\\
&= u^\epsilon(t, Q_3) + D^1 u_4^\epsilon[\Delta_2] + D^1 u_2^\epsilon[\Delta_1] + \frac{1}{2} D^2 u_4^\epsilon[\Delta_1, \Delta_1]\\
&\quad + \frac{1}{6} D^3 u_2^\epsilon[\Delta_1, \Delta_1, \Delta_1] + \theta D_1 u_4^\epsilon + \theta D_1^1 u_2^\epsilon[\Delta_1] + R_1,
\end{aligned}$$

where the remainders are as follows

$$\begin{aligned}
R_{11} &= \frac{1}{2}\theta^2 \int_0^1 s D_2 u^\epsilon\Big(t + (1-s)\theta, Q_1\Big) ds,\\
R_{12} &= \frac{1}{24} \int_0^1 s^3 D^4 u^\epsilon(t, sQ_2 + (1-s)Q_1)[\Delta_1, \Delta_1, \Delta_1, \Delta_1] ds,\\
R_{13} &= \frac{1}{2}\theta \int_0^1 s^2 D_1^2 u^\epsilon(t, sQ_2 + (1-s)Q_1)[\Delta_1, \Delta_1] ds,\\
R_{14} &= \frac{1}{2} \int_0^1 s D^2 u^\epsilon(t, s(Q_3 + (1-s)Q_2)[\Delta_2, \Delta_2] ds,\\
R_{15} &= \frac{1}{2} \int_0^1 s^2 D^3 u^\epsilon(t, s(Q_3) + (1-s)Q_2)[\Delta_1, \Delta_1, \Delta_2] ds,\\
R_{16} &= \theta \int_0^1 s D_1^1 u^\epsilon(t, s(Q_3) + (1-s)Q_2)[\Delta_2] ds,\\
R_{17} &= \int_0^1 s D^2 u^\epsilon(t, s(Q_4) + (1-s)Q_3)[\Delta_2, \Delta_3] ds,\\
R_{18} &= \frac{1}{2} \int_0^1 s D^3 u^\epsilon(t, s(Q_4) + (1-s)Q_3)[\Delta_1, \Delta_1, \Delta_3] ds,
\end{aligned}$$

$$R_{19} = \theta \int_0^1 s D_1^1 u^\epsilon(t, s(Q_4) + (1-s)Q_3)[\Delta_3] ds,$$
$$R_1 = R_{11} + R_{12} + R_{13} + R_{14} + R_{15} + R_{16} + R_{17} + R_{18} + R_{19}.$$

Using (5.3.17), (5.3.10)-(5.3.11), and the fact that ξ and η have mean zero and that components of ξ, η, θ, J are mutually independent, we obtain

$$\begin{aligned} &\mathbb{E}[u^\epsilon(t+\theta, X)Y + Z] \qquad (5.3.18)\\ &= \mathbb{E}\Big[\left(u^\epsilon(t, Q_3) + D^1 u_4^\epsilon[\Delta_2] + \frac{1}{2} D^2 u_4^\epsilon[\Delta_1, \Delta_1] + \theta D_1 u_4^\epsilon \right) (y + \theta c y) \\ &\quad + z + \theta g y + y(1+\theta c) R_1 \Big]. \end{aligned}$$

The following elementary formulas are needed for future calculations:

$$\begin{aligned} \mathbb{E}\left[D^2 u^\epsilon[\Delta_1, \Delta_1] | \theta \right] & \qquad (5.3.19)\\ &= \theta \sum_{i,j=1}^d \left[a^{ij}(t,x) + \left(F(t,x) B_\epsilon(t,x) F^\top(t,x) \right)^{ij} \right] \frac{\partial^2 u^\epsilon}{\partial x^i \partial x^j} \\ &=: \theta(a + F B_\epsilon F^T) : \nabla\nabla u^\epsilon, \\ u^\epsilon(t, Q_3) - u^\epsilon(t,x) &= u^\epsilon(t, x + \mathbf{I}(\theta < h) F J) - u^\epsilon(t,x) \\ &= \mathbf{I}(\theta < h)[u^\epsilon(t, x + FJ) - u^\epsilon(t,x)], \\ \mathbb{E}[\theta] &= \frac{1 - e^{-\lambda_\epsilon h}}{\lambda_\epsilon}, \\ \mathbb{E}[\theta^2] &= 2\frac{1 - e^{-\lambda_\epsilon h}(1 + \lambda_\epsilon h)}{\lambda_\epsilon^2}, \\ \mathbb{E}[\mathbf{I}(\theta < h)] &= 1 - e^{-\lambda_\epsilon h}, \\ \mathbb{E}[\mathbf{I}(\theta < h)\theta] &= \frac{1 - e^{-\lambda_\epsilon h}(1 + \lambda_\epsilon h)}{\lambda_\epsilon}. \end{aligned}$$

Also, $\mathbb{E}[v(J)]$ for some $v(z)$ will mean

$$\mathbb{E}[v(J)] = \mathbb{E}[v(J_\epsilon)] = \frac{1}{\lambda_\epsilon} \int_{|s| > \epsilon} v(s) \nu(ds).$$

Noting that $u_4^\epsilon = u^\epsilon(t,x) = u^\epsilon$ and using (5.3.18), (5.3.12), (5.3.19) and (5.2.10), we obtain

$$
\begin{aligned}
\mathbb{E}\,[R] &:= \mathbb{E}\big[u^\epsilon(t+\theta,X)Y + Z - u^\epsilon y - z\big] \\
&= \mathbb{E}[\theta\Big(D_1u^\epsilon + D^1u^\epsilon[b - F\gamma_\epsilon] + \frac{1}{2}(a + FB_\epsilon F^T) : \nabla\nabla u^\epsilon\Big)(y + \theta c y) + \theta g y \\
&\quad + u^\epsilon(t, x + \mathbf{I}(\theta < h)FJ)(y + \theta c y) - u^\epsilon y] + y\mathbb{E}[(1 + \theta c)R_1] \\
&= \mathbb{E}[\theta\Big(D_1u^\epsilon + D^1u^\epsilon[b - F\gamma_\epsilon] + \frac{1}{2}(a + FB_\epsilon F^T) : \nabla\nabla u^\epsilon + cu^\epsilon + g\Big)y \\
&\quad + [u^\epsilon(t, x + \mathbf{I}(\theta < h)FJ) - u^\epsilon)]y \\
&\quad + \theta^2\Big(D_1u^\epsilon + D^1u^\epsilon[b - F\gamma_\epsilon] + \frac{1}{2}(a + FB_\epsilon F^T) : \nabla\nabla u^\epsilon\Big)cy \\
&\quad + \theta\,[u^\epsilon(t, x + \mathbf{I}(\theta < h)FJ) - u^\epsilon]\,cy\Big] + y\mathbb{E}[(1 + \theta c)R_1] \\
&= \mathbb{E}[\theta\Big(D_1u^\epsilon + D^1u^\epsilon[b - F\gamma_\epsilon] + \frac{1}{2}(a + FB_\epsilon F^T) : \nabla\nabla u^\epsilon + cu^\epsilon + g\Big)y \\
&\quad + \mathbf{I}(\theta < h)[u^\epsilon(t, x + FJ) - u^\epsilon)]y \\
&\quad + \theta^2\Big(D_1u^\epsilon + D^1u^\epsilon[b - F\gamma_\epsilon] + \frac{1}{2}(a + FB_\epsilon F^T) : \nabla\nabla u^\epsilon\Big)cy \\
&\quad + \theta\mathbf{I}(\theta < h)[u^\epsilon(t, x + FJ) - u^\epsilon]cy] + y\mathbb{E}[(1 + \theta c)R_1] \\
&= \mathbb{E}[\theta\Big(D_1u^\epsilon + D^1u^\epsilon[b - F\gamma_\epsilon] + \frac{1}{2}(a + FB_\epsilon F^T) : \nabla\nabla u^\epsilon + cu^\epsilon + g\Big)y] \\
&\quad + \mathbb{E}\,[\mathbf{I}(\theta < h)[u^\epsilon(t, x + FJ) - u^\epsilon(t,x)]y] + y\mathbb{E}[R_1(1 + \theta c) + R_2] \\
&= \frac{1 - e^{-\lambda_\epsilon h}}{\lambda_\epsilon}\Big(D_1u^\epsilon + D^1u^\epsilon[b - F\gamma_\epsilon] + \frac{1}{2}(a + FB_\epsilon F^T) : \nabla\nabla u^\epsilon \\
&\quad + cu^\epsilon(t,x) + g\Big)y \\
&\quad + \Big(1 - e^{-\lambda_\epsilon h}\Big)\,\mathbb{E}\,[u^\epsilon(t, x + FJ) - u^\epsilon(t,x)]\,y + y\mathbb{E}[R_0] \\
&= \frac{1 - e^{-\lambda_\epsilon h}}{\lambda_\epsilon}\Big(D_1u^\epsilon + D^1u^\epsilon[b - F\gamma_\epsilon] + \frac{1}{2}(a + FB_\epsilon F^T) : \nabla\nabla u^\epsilon \\
&\quad + cu^\epsilon(t,x) + g\Big)y \\
&\quad + \frac{1 - e^{-\lambda_\epsilon h}}{\lambda_\epsilon}\int\limits_{|s|\ge\epsilon}\{u^\epsilon(t, x + Fs) - u^\epsilon(t,x)\}\nu(ds)y + y\mathbb{E}[R_0] \\
&= y\mathbb{E}[R_0],
\end{aligned}
$$

where

$$R_0 = R_1(1+\theta c) + R_2,$$

$$R_2 = R_{21} + R_{22},$$

and

$$R_{21} = \theta^2\Big(D_1 u^\epsilon + D^1 u^\epsilon[b - F\gamma_\epsilon] + \frac{1}{2}(a + FB_\epsilon F^T) : \nabla\nabla u^\epsilon\Big)c,$$
$$R_{22} = \theta\mathbf{I}(\theta < h)[u^\epsilon(t, x + FJ) - u^\epsilon(t,x)]c.$$

It is clear that many of the terms in R are only non–zero in the case $\theta < h$, i.e. when a jump occurs. We rearrange the terms in R_0 according to their degree in θ:

$$\begin{aligned} R_0 = &\underbrace{R_{17} + R_{18} + R_{19} + R_{22}}_{\mathbf{I}(\theta<h)\theta\text{-terms}} + \underbrace{R_{11} + R_{12} + R_{13} + R_{14} + R_{15} + R_{16} + R_{21}}_{\theta^2\text{ - terms}} \\ &+ \underbrace{\theta c(R_{17} + R_{18} + R_{19})}_{(\mathbf{I}(\theta<h)\theta^2\text{-terms}} + \underbrace{\theta c(R_{11} + R_{12} + R_{13} + R_{14} + R_{15} + R_{16})}_{\theta^3\text{ - terms}} \end{aligned}$$

Now to estimate the terms in the error R_0, we observe that

(i) $\int_{|s|>\epsilon} s\nu(ds) = \gamma_\epsilon + \int_{|s|>1} s\nu(ds)$ with the latter integral bounded and, in particular, $|\mathbb{E}[J]| \leq K(1 + |\gamma_\epsilon|)/\lambda_\epsilon$;

(ii) $\mathbb{E}\left[|J|^{2p}\right]$, $p \geq 1$, are bounded by K/λ_ϵ (see (5.3.3));

(iii) the terms R_{17}, R_{18}, R_{19}, R_{21} and R_{22} contain derivatives of u^ϵ evaluated at or between the points Q_3 and Q_4 and in their estimation Assumption 5.2.5 and (5.3.14) from Lemma 5.3.4 are used;

(iv) the terms R_{11}, R_{12}, R_{13}, R_{14}, R_{15} and R_{16} contain derivatives of u^ϵ evaluated at or between the points Q_1 and Q_2 and in their estimation Assumption 5.2.5, (5.3.13) from Lemma 5.3.4, and Lemma 5.3.5 are used;

(v) $\gamma_\epsilon^2/\lambda_\epsilon$ is bounded by a constant independent of ϵ.

As a result, we obtain

$$\begin{aligned}
\left|\mathbb{E}\left[R_{17}+R_{18}+R_{19}+R_{22}\right]\right| &\leq K_1\frac{(1+|\gamma_\epsilon|^2)}{\lambda_\epsilon}\mathbb{E}\left[\mathbf{I}(\theta<h)\theta\right], \\
\left|\mathbb{E}\left[\theta(R_{17}+R_{18}+R_{19})\right]\right| &\leq K_2\frac{(1+|\gamma_\epsilon|^2)}{\lambda_\epsilon}\mathbb{E}\left[\mathbf{I}(\theta<h)\theta^2\right] \\
&\leq K_3\frac{(1+|\gamma_\epsilon|^2)}{\lambda_\epsilon}\mathbb{E}\left[\mathbf{I}(\theta<h)\theta\right], \\
\left|\mathbb{E}\left[(R_{11}+R_{12}+R_{13}+R_{14}+R_{15}+R_{16}+R_{21})\right]\right| & \\
&\leq K_4(1+|\gamma_\epsilon|^2)(\mathbb{E}\left[\theta^2\right]+|\gamma_\epsilon|^q\mathbb{E}\left[\theta^{q+2}\right]) \\
&\leq K_5(1+|\gamma_\epsilon|^2)\frac{1-e^{-\lambda_\epsilon h}(1+\lambda_\epsilon h)}{\lambda_\epsilon^2},
\end{aligned}$$

and

$$\begin{aligned}
&\left|\mathbb{E}\left[\theta(R_{11}+R_{12}+R_{13}+R_{14}+R_{15}+R_{16})\right]\right| \\
&\leq K_6(1+|\gamma_\epsilon|^2)(\mathbb{E}\left[\theta^3\right]+|\gamma_\epsilon|^q\mathbb{E}\left[\theta^{q+3}\right]) \\
&\leq K_7(1+|\gamma_\epsilon|^2)\frac{1-e^{-\lambda_\epsilon h}(1+\lambda_\epsilon h)}{\lambda_\epsilon^3} \leq K_8(1+|\gamma_\epsilon|^2)\frac{1-e^{-\lambda_\epsilon h}(1+\lambda_\epsilon h)}{\lambda_\epsilon^2},
\end{aligned}$$

where all constants $K_i > 0$ are independent of h and ϵ and $q \geq 1$.

Overall we obtain

$$\begin{aligned}
\left|\mathbb{E}[R]\right| &\leq (K_1+K_3)\frac{(1+|\gamma_\epsilon|^2)}{\lambda_\epsilon}y\mathbb{E}\left[\mathbf{I}(\theta<h)\theta\right] \\
&\quad+(K_5+K_8)(1+|\gamma_\epsilon|^2)y\frac{1-e^{-\lambda_\epsilon h}(1+\lambda_\epsilon h)}{\lambda_\epsilon^2} \\
&\leq K\left\{\frac{1}{\lambda_\epsilon}\mathbb{E}\left[\mathbf{I}(\theta<h)\theta\right]+\frac{1-e^{-\lambda_\epsilon h}(1+\lambda_\epsilon h)}{\lambda_\epsilon^2}\right\}(1+|\gamma_\epsilon|^2)y \\
&= 2K(1+|\gamma_\epsilon|^2)\frac{1-e^{-\lambda_\epsilon h}(1+\lambda_\epsilon h)}{\lambda_\epsilon^2}y.
\end{aligned}$$

□

Remark 5.3.7. We note the following two asymptotic regimes for the one-step error (5.3.16). For $\lambda_\epsilon h < 1$ (in practice, this occurs only when λ_ϵ is small or moderate like it is in jump-diffusions), we can expand the exponent in (5.3.16) and obtain that the one-step error is of order $O(h^2)$:

$$\left|\mathbb{E}[R(t,x,y,z)]\right| \le K(1+|\gamma_\epsilon|^2)h^2 y.$$

In the case, where λ_ϵ is very large (e.g., for small ϵ in the infinite activity case) then the term with $e^{-\lambda_\epsilon h}$ can be neglected and we get

$$\left|\mathbb{E}[R(t,x,y,z)]\right| \le K\frac{1+|\gamma_\epsilon|^2}{\lambda_\epsilon^2}y.$$

The usefulness of a more precise estimate (5.3.16) is that it includes situations in between these two asymptotic regimes and also allows to consider an interplay between h and ϵ (see Section 5.3.4).

5.3.3 *Global error*

In this section we obtain an estimate for the global weak-sense error of Algorithm 1. We first estimate average number of steps $\mathbb{E}[\varkappa]$ of Algorithm 1.

Lemma 5.3.8 (*Number of steps*)**.** *The average number of steps $\varkappa$ for the chain X_k from Algorithm 1 satisfies the following bound*

$$E[\varkappa] \le \frac{(T-t_0)\lambda_\epsilon}{1-e^{-\lambda_\epsilon h}} + 1.$$

Proof. It is obvious that if we replace the bounded domain G in Algorithm 1 with the whole space $\mathbb{R}^d$ (i.e., replace the Dirichlet problem by the Cauchy one), then the corresponding number of steps $\varkappa'$ of Algorithm 1 is not less than $\varkappa$. Hence it is sufficient to get an estimate for $E[\varkappa']$. Let $\delta_1, \delta_2, \ldots$ be the interarrival times of the jumps, $\theta_i = \delta_i \wedge h$ for $i \ge 0$, and $S_k = \sum_{i=0}^{k-1}\theta_i$ for $k \ge 0$. Then

$$\varkappa \le \varkappa' := \inf\{l : S_l \ge T - t_0\}.$$

Introduce the martingale: $\widetilde{S}_0 = 0$ and $\widetilde{S}_k := S_k - k\mathbb{E}[\theta]$ for $k \geq 1$. Since $\theta_i \leq h$ we have that $\widetilde{S}_{\varkappa'-1} \leq S_{\varkappa'-1} < T - t_0$ almost surely and thus by the optional stopping theorem we obtain

$$\mathbb{E}\left[\widetilde{S}_{\varkappa'-1}\right] = \mathbb{E}\left[\widetilde{S}_0\right] = 0.$$

Therefore

$$\mathbb{E}\left[S_{\varkappa'-1}\right] = \mathbb{E}[\varkappa' - 1] \cdot \mathbb{E}[\theta]$$

and we conclude

$$\begin{aligned}\mathbb{E}[\varkappa] &\leq \mathbb{E}\left[\varkappa'\right] = \mathbb{E}[\varkappa' - 1] + 1 \\ &= \frac{\mathbb{E}\left[S_{\varkappa'-1}\right]}{\mathbb{E}[\theta]} + 1 \leq \frac{(T - t_0)\lambda_\epsilon}{1 - \mathrm{e}^{-\lambda_\epsilon h}} + 1.\end{aligned}$$

□

We also need the following auxiliary lemma.

Lemma 5.3.9 (Boundedness of Y_k in Algorithm 1)**.** *The chain Y_k defined in (5.3.6) is uniformly bounded by a deterministic constant:*

$$Y_k \leq e^{\bar{c}(T - t_0 + h)},$$

where $\bar{c} = \max_{(t,x)\in\bar{Q}} c(t,x)$.

Proof. From (5.3.6), we can express Y_k via previous Y_{k-1} and get the required estimate as follows:

$$\begin{aligned}Y_k &= Y_{k-1}(1 + \theta_k c(t_{k-1}, x_{k-1})) \leq Y_{k-1}(1 + \theta_k \bar{c}) \\ &\leq Y_{k-1} e^{\bar{c}\theta_k} \leq Y_{k-2} e^{\bar{c}(\theta_k + \theta_{k-1})} \leq Y_0 e^{\bar{c}(\vartheta_k - t_0)} \leq e^{\bar{c}(T - t_0 + h)}.\end{aligned}$$

□

Now we prove the convergence theorem for Algorithm 1.

Theorem 5.3.10 (Global error of Algorithm 1)**.** *Under Assumption 5.2.4 with $l = 2$, $m = 4$ and Assumptions 5.2.1, 5.2.5 and 5.3.1, the global error of Algorithm 1 satisfies the following bound*

$$\left|\mathbb{E}[\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa + Z_\varkappa] - u^\epsilon(t_0, x_0)\right| \le K(1 + |\gamma_\epsilon|^2)\left(\frac{1}{\lambda_\epsilon} - h\frac{e^{-\lambda_\epsilon h}}{1 - e^{-\lambda_\epsilon h}}\right) + K\frac{1 - e^{-\lambda_\epsilon h}}{\lambda_\epsilon}, \tag{5.3.20}$$

where $K > 0$ is a constant independent of h and ϵ.

Proof. Recall (see (5.2.4)):

$$u^\epsilon(t, x) = \mathbb{E}\left[\varphi\left(\widetilde{\tau}_{t,x}, \widetilde{X}_{t,x}(\widetilde{\tau}_{t,x})\right)\widetilde{Y}_{t,x,1}(\widetilde{\tau}_{t,x}) + \widetilde{Z}_{t,x,1,0}(\widetilde{\tau}_{t,x})\right].$$

The global error

$$\mathbf{R} := \left|\mathbb{E}[\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa + Z_\varkappa] - u^\epsilon(t_0, x_0)\right|$$

can be written as

$$\begin{aligned}\mathbf{R} &= \big|\mathbb{E}[\mathbf{I}(\vartheta_\varkappa \ge T)\left(\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa - u^\epsilon(\vartheta_\varkappa, X_\varkappa)Y_\varkappa\right) + u^\epsilon(\vartheta_\varkappa, X_\varkappa)Y_\varkappa \\ &\quad + Z_\varkappa - u^\epsilon(t_0, x_0)]\big| \\ &\le \left|\mathbb{E}[\mathbf{I}(\vartheta_\varkappa \ge T)\left(\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa - u^\epsilon(\vartheta_\varkappa, X_\varkappa)Y_\varkappa\right)]\right| \\ &\quad + \left|\mathbb{E}[u^\epsilon(\vartheta_\varkappa, X_\varkappa)Y_\varkappa + Z_\varkappa - u^\epsilon(t_0, x_0)]\right|.\end{aligned} \tag{5.3.21}$$

Using Lemma 5.3.9, Assumption 5.2.5 and Lemmas 5.3.4 and 5.3.5 as well as that $\bar{\vartheta}_\varkappa - \vartheta_\varkappa \le \theta_\varkappa$, we have for the first term in (5.3.21):

$$\begin{aligned}\mathbb{E}[\mathbf{I}(\vartheta_\varkappa \ge T)\left(\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa - u^\epsilon(\vartheta_\varkappa, X_\varkappa)Y_\varkappa\right)] &\le KE\left[\theta_\varkappa(1 + |\gamma_\epsilon|^q\theta_\varkappa^q)\right] \\ &\le K\frac{1 - e^{-\lambda_\epsilon h}}{\lambda_\epsilon},\end{aligned} \tag{5.3.22}$$

where $K > 0$ does not depend on h or ε.

For the second term in (5.3.21), we exploit ideas from [92] to re-express the global error. We use Theorem 5.3.6 and Lemmas 5.3.9 and 5.3.8:

$$
\begin{aligned}
&|\mathbb{E}[u^{\epsilon}(\vartheta_{\varkappa}, X_{\varkappa})Y_{\varkappa} + Z_{\varkappa} - u^{\epsilon}(t_0, x_0)]| \qquad (5.3.23)\\
&= \left|\mathbb{E}\left[\sum_{k=0}^{\varkappa-1}\mathbb{E}\left[u^{\epsilon}(\vartheta_{k+1}, X_{k+1})Y_{k+1} + Z_{k+1} - u^{\epsilon}(\vartheta_k, X_k)Y_k - Z_k \middle| \vartheta_k, X_k, Y_k, Z_k\right]\right]\right|\\
&= \left|\mathbb{E}\left[\sum_{k=0}^{\varkappa-1}\mathbb{E}\left[R(\vartheta_k, X_k, Y_k, Z_k)\middle|\vartheta_k, X_k, Y_k, Z_k\right]\right]\right|\\
&\le \mathbb{E}\left[\sum_{k=0}^{\varkappa-1}\frac{1-e^{-\lambda_\epsilon h}(1+\lambda_\epsilon h)}{\lambda_\epsilon^2}K(1+|\gamma_\epsilon|^2)Y_k\right]\\
&\le K\frac{1+|\gamma_\epsilon|^2}{\lambda_\epsilon^2}\left(1-e^{-\lambda_\epsilon h}(1+\lambda_\epsilon h)\right)\mathbb{E}\left[\varkappa\right]\\
&\le K(1+|\gamma_\epsilon|^2)\left(\frac{1}{\lambda_\epsilon(1-e^{-\lambda_\epsilon h})} - h\frac{e^{-\lambda_\epsilon h}}{1-e^{-\lambda_\epsilon h}}\right)(T-t_0)\\
&\le K(1+|\gamma_\epsilon|^2)\left(\frac{1}{\lambda_\epsilon} - h\frac{e^{-\lambda_\epsilon h}}{1-e^{-\lambda_\epsilon h}}\right),
\end{aligned}
$$

where, as usual constants $K > 0$ are changing from line to line. Combining (5.3.21)-(5.3.23), we arrive at (5.3.20). □

Remark 5.3.11 (Error estimate and convergence)**.** Note that the error estimate in Theorem 5.3.10 gives us the expected results in the limiting cases (see also Remark 5.3.7). If $\lambda_\epsilon h < 1$, we obtain:

$$\mathbf{R} \le K(1+|\gamma_\epsilon|^2)h,$$

which is expected for weak convergence in the jump-diffusion case.

If λ_ϵ is large (meaning that almost always $\theta < h$), the error is tending to

$$\mathbf{R} \le K(1+|\gamma_\epsilon|^2)\frac{1}{\lambda_\epsilon},$$

as expected (cf. [72]).

We also remark that for any fixed λ_ϵ, we have first order convergence when $h \to 0$.

Remark 5.3.12. In the case of symmetric measure $\nu(z)$ we have $\gamma_\epsilon = 0$ and hence the global error (5.3.20) becomes

$$\begin{aligned} &\left|\mathbb{E}[\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa + Z_\varkappa] - u^\epsilon(t_0, x_0)\right| \\ &\leq K\left(\frac{1}{\lambda_\epsilon} - h\frac{e^{-\lambda_\epsilon h}}{1-e^{-\lambda_\epsilon h}}\right) + K\frac{1-e^{-\lambda_\epsilon h}}{\lambda_\epsilon}. \end{aligned} \tag{5.3.24}$$

Remark 5.3.13 (Cauchy case)**.** In general, it might be possible to show that the global error estimate (5.3.20) for Algorithm 1 also holds in the Cauchy problem case. Some suggestions and initial thoughts are given in [30].

5.3.4 *The case of infinite intensity of jumps*

In this section we combine the previous results, Theorem 5.2.5 and 5.3.10, to obtain an overall error estimate for solving the problem (4.1.1) in the case of infinite intensity of jumps by Algorithm 1. We obtain

$$\begin{aligned} &\left|\mathbb{E}[\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa + Z_\varkappa] - u(t_0, x_0)\right| \\ &\leq K(1+|\gamma_\epsilon|^2)\left(\frac{1}{\lambda_\epsilon} - h\frac{e^{-\lambda_\epsilon h}}{1-e^{-\lambda_\epsilon h}}\right) + K\frac{1-e^{-\lambda_\epsilon h}}{\lambda_\epsilon} + K\int_{|z|\leq\epsilon}|z|^3\nu(dz), \end{aligned} \tag{5.3.25}$$

where $K > 0$ is independent of h and ϵ.

Let us consider an α-stable process in which the Lévy measure has the following singular behaviour near zero

$$\nu(dz) \sim |z|^{-m-\alpha}dz, \ \alpha \in (0,2), \tag{5.3.26}$$

i.e., we are focusing our attention here on the singularity near zero only and the sign $\sim$ means that the limit of the ratio of both sides equals to some positive constant. Consequently, all calculations are done in this section up to positive constant factors independent of ϵ and h. The behaviour (5.3.26) is typical for m-dimensional Lévy measures near zero (see e.g. [4, p. 37] and also the one-dimensional Example 5.2.1). Then

$$\lambda_\epsilon = \int_{|z|\geq\epsilon} \nu(\mathrm{d}z) \sim \epsilon^{-\alpha},$$
$$\gamma_\epsilon^2 = \sum_{i=1}^m \left[\int_{\epsilon\leq|z|\leq 1} z^i \nu(\mathrm{d}z)\right]^2 \qquad \sim \epsilon^{2-2\alpha} \text{ for } \alpha \neq 1$$
$$\text{and } \gamma_\epsilon^2 \sim |\ln\epsilon|^2 \text{ for } \alpha = 1,$$
$$\int_{|z|\leq\epsilon} |z|^3 \nu(\mathrm{d}y) \sim \epsilon^{3-\alpha}.$$

Hence

$$\begin{aligned}&\left|\mathbb{E}[\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa + Z_\varkappa] - u(t_0,x_0)\right| \\ &\leq K\left[(1+\gamma_\epsilon^2)\left(\epsilon^\alpha - h\frac{e^{-\epsilon^{-\alpha}h}}{1-e^{-\epsilon^{-\alpha}h}}\right) + \epsilon^\alpha\left(1-e^{-\epsilon^{-\alpha}h}\right) + \epsilon^{3-\alpha}\right].\end{aligned} \tag{5.3.27}$$

Let us measure the computational cost of Algorithm 1 in terms of the average number of steps (see Lemma 5.3.8). Since

$$E\left[\varkappa\right] \leq \frac{(T-t_0)\lambda_\epsilon}{1-e^{-\lambda_\epsilon h}} \leq K\frac{\epsilon^{-\alpha}}{1-e^{-\epsilon^{-\alpha}h}},$$

we choose to use the cost associated with the average number of steps as

$$C := \frac{\epsilon^{-\alpha}}{1-e^{-\epsilon^{-\alpha}h}}.$$

We fix a tolerance level ρ_{tol} and require ϵ and h to be so that

$$\rho_{tol} = \rho(\epsilon,h) := (1+\gamma_\epsilon^2)\left(\epsilon^\alpha - \frac{he^{-\epsilon^{-\alpha}h}}{1-e^{-\epsilon^{-\alpha}h}}\right) + \epsilon^\alpha\left(1-e^{-\epsilon^{-\alpha}h}\right) + \epsilon^{3-\alpha}.$$

Note that since we are using the Euler scheme for SDEs' approximation, the decrease of ρ_{tol} in terms of cost cannot be faster than linear. We now consider three cases of α.

The case $\alpha \in (0,1)$. We have

$$\rho(\epsilon,h) \le \epsilon^{2-\alpha} + 2\epsilon^{\alpha} + \epsilon^{3-\alpha} = O(\epsilon^{\alpha})$$

and, by choosing sufficiently small ϵ, we can reach the required ρ_{tol}. It is optimal to take $h = \infty$ (in practice, taking $h = T - t_0$) and the cost is then $C = 1/\epsilon^{\alpha}$. Hence ρ_{tol} is inversely proportional to C, and convergence is linear in cost (to reduce ρ_{tol} twice, we need to double C).

The case $\alpha = 1$. We have

$$\rho(\epsilon,h) = (1 + |\ln\epsilon|^2)\left(\epsilon - \frac{he^{-\epsilon^{-1}h}}{1 - e^{-\epsilon^{-1}h}}\right) + \epsilon\left(1 - e^{-\epsilon^{-1}h}\right) + \epsilon^2 = O(\epsilon|\ln\epsilon|^2),$$

i.e. convergence is almost linear in cost.

The case $\alpha \in (1,2)$. If we take $h = \infty$, then $\rho(\epsilon,h) = O(\epsilon^{2-\alpha})$ and the convergence order in terms of cost is $2/\alpha - 1$, which is very slow (e.g., for $\alpha = 3/2$, the order is $1/3$ and for $\alpha = 1.9$, the order is ≈ 0.05). Let us now take $h = \epsilon^{\ell}$ with $\ell \ge \alpha$. Then

$$\begin{aligned}\rho(\epsilon,h) &= \rho(\epsilon,\epsilon^{\ell}) = (1 + \epsilon^{2-2\alpha})\left(\epsilon^{\alpha} - \frac{\epsilon^{\ell}e^{-\epsilon^{\ell-\alpha}}}{1 - e^{-\epsilon^{\ell-\alpha}}}\right) + \epsilon^{\alpha}\left(1 - e^{-\epsilon^{\ell-\alpha}}\right) + \epsilon^{3-\alpha} \\ &\le (1 + \epsilon^{2-2\alpha})\epsilon^{\ell} + \epsilon^{\ell} + \epsilon^{3-\alpha} = \epsilon^{2-2\alpha+\ell} + 2\epsilon^{\ell} + \epsilon^{3-\alpha}\end{aligned}$$

and $C \approx 1/h = \epsilon^{-\ell}$. The optimal $\ell = 1 + \alpha$, for which $\rho(\epsilon,h) = O(\epsilon^{3-\alpha})$ and the convergence order in terms of cost is $(3-\alpha)/(1+\alpha)$, which is much better (e.g., for $\alpha = 3/2$, the order is $3/5$ and it cannot be smaller than $1/3$ for any $\alpha \in (1,2)$). Note that in the case of symmetric measure $\nu(z)$ (see Remark 5.3.12), convergence is linear in cost for $\alpha \in (1,2)$.

To summarise, for $\alpha \in (0,1)$ we have first order convergence and there is no benefit of restricting jump adapted steps by h . However, in the case of $\alpha \in (1,2)$, it is beneficial to use restricted jump-adapted steps to get the order of $(3-\alpha)/(1+\alpha)$. We also recall that restricted jump-adapted steps should typically be used for jump-diffusions (the finite activity case when there is no singularity of λ_ϵ and γ_ϵ) because jump time increments δ typically take too large values and to control the error at every step we should truncate those times at a sufficiently small $h > 0$ for a satisfactory accuracy.

6

NUMERICAL EXPERIMENTS USING THE INTRODUCED ALGORITHM

In this chapter we illustrate the theoretical results of Section 5.3. In particular, we showcase the behaviour in the case of infinite intensity of jumps for different regimes of α. In Section 6.1 we introduce the common Monte Carlo technique and its notation used throughout this chapter. We display numerical tests of Algorithm 1 in four different examples:

(i) a non-singular Lévy measure (Example 6.2.1),

(ii) a singular Lévy measure which is similar to that of Example 5.2.1 (see Example 6.3.1),

(iii) pricing a foreign-exchange (FX) barrier basket option where the underlying model is of exponential Lévy-type (Example 6.4.1),

(iv) pricing a FX barrier option showing that the convergence orders hold (Example 6.4.2).

6.1 MONTE CARLO TECHNIQUE

As it is common for weak approximation (see e.g. [92]), in simulations we complement Algorithm 1 by the Monte Carlo techniques and evaluate $u(t_0, x)$ or $u^\epsilon(t_0, x)$ as

$$\begin{aligned}\bar{u}(t_0, x) &:= \mathbb{E}\left[\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa + Z_\varkappa\right] \qquad (6.1.1)\\ &\simeq \hat{u} = \frac{1}{M}\sum_{m=1}^{M}\left[\varphi(\bar{\vartheta}_\varkappa^{(m)}, X_\varkappa^{(m)})Y_\varkappa^{(m)} + Z_\varkappa^{(m)}\right],\end{aligned}$$

where $(\bar{\vartheta}_\varkappa^{(m)}, X_\varkappa^{(m)}, Y_\varkappa^{(m)}, Z_\varkappa^{(m)})$ are independent realisations of $(\bar{\vartheta}_\varkappa, X_\varkappa, Y_\varkappa, Z_\varkappa)$. The Monte Carlo error of (6.1.1) is

$$\sqrt{D_M} := \frac{\left(\mathrm{Var}\left[\varphi(\bar{\vartheta}_\varkappa, X_\varkappa)Y_\varkappa + Z_\varkappa\right]\right)^{1/2}}{M^{1/2}} \simeq \sqrt{\bar{D}_M},$$

where

$$\bar{D}_M = \frac{1}{M}\left[\frac{1}{M}\sum_{m=1}^{M}\left[\Xi^{(m)}\right]^2 - \left(\frac{1}{M}\sum_{m=1}^{M}\Xi^{(m)}\right)^2\right],$$

and $\Xi^{(m)} = \varphi\left(\bar{\vartheta}_\varkappa^{(m)}, X_\varkappa^{(m)}\right)Y_\varkappa^{(m)} + Z_\varkappa^{(m)}$. Then $\bar{u}(t_0, x)$ falls in the corresponding confidence interval $\hat{u} \pm 2\sqrt{\bar{D}_M}$ with probability 0.95.

6.2 EXAMPLE WITH A NON-SINGULAR LÉVY MEASURE

In this section, we illustrate Algorithm 1 in the case of a simple non-singular Lévy measure (i.e., the jump-diffusion case), where there is no need to replace small jumps and hence we directly approximate $u(t_0, x)$ rather than $u^\epsilon(t_0, x)$. Consequently, the numerical integration error does not depend on ϵ. We recall (see Theorem 5.3.10) that Algorithm 1 has first order of convergence in h.

Example 6.2.1 (Non-singular Lévy measure)**.** To construct this and the next example, we use the same recipe as in [91, 92]: we choose the coefficients of the problem (4.1.1) so that we can write down its solution explicitly. Having the exact solution is very useful for numerical tests.

Consider the problem (4.1.1) with $d = 3$, $G = U_1$ which is the open unit ball centred at the origin in $\mathbb{R}^3$, and with the coefficients

$$a^{11}(t,x) = 1.21 - x_2^2 - x_3^2, \quad a^{22} = 1, \quad a^{33} = 1, \quad a^{ij} = 0,\ i \neq j, \quad b = 0, \tag{6.2.1}$$

$$F(t,x) = (f,f,f)^T,\ f \in \mathbb{R}, \tag{6.2.2}$$

$$g(t,x) := \frac{1}{2}e^{t-T}(1.21 - x_1^4 - x_2^4) + 6(1 - \frac{1}{2}e^{t-T}) \left[x_1^2(1.21 - x_2^2 - x_3^2) + x_2^2\right] \tag{6.2.3}$$
$$+ (1 - \frac{1}{2}e^{t-T})\Big[(C_+ - C_-)\frac{4f}{\mu^2}(x_1^3 + x_2^3) + (C_+ + C_-)\frac{12f^2}{\mu^3}(x_1^2 + x_2^2)$$
$$+ (C_+ - C_-)\frac{24f^3}{\mu^4}(x_1 + x_2) + (C_+ + C_-)\frac{48f^4}{\mu^5}\Big],$$

with the boundary condition

$$\varphi(t,x) = (1 - \tfrac{1}{2}\mathrm{e}^{t-T})(1.21 - x_1^4 - x_2^4) \tag{6.2.4}$$

and with the Lévy measure density

$$\nu(\mathrm{d}z) = \begin{cases} C_- \mathrm{e}^{-\mu|z|}\mathrm{d}z, & \text{if } z < 0, \\ C_+ \mathrm{e}^{-\mu|z|}\mathrm{d}z, & \text{if } z > 0, \end{cases}$$

where C_- and C_+ are some positive constants. Note that, keeping in mind Remark 5.2.2, the coefficients from (6.2.1)-(6.2.3) satisfy Assumptions 5.2.1-5.2.2.

It is not difficult to verify that this problem has the solution

$$u(t,x) = (1 - \tfrac{1}{2}\mathrm{e}^{t-T})(1.21 - x_1^4 - x_2^4).$$

and we also find

$$\lambda = \int_{|z|>0} \nu(\mathrm{d}z) = \int_{\mathbb{R}} \nu(\mathrm{d}z) = \frac{C_+ + C_-}{\mu},$$
$$\rho(z) = \frac{C_- \mathrm{e}^{-\mu|z|}\mathbf{I}(z<0) + C_+ \mathrm{e}^{-\mu|z|}\mathbf{I}(z>0)}{\lambda}.$$

We simulated jump sizes by analytically inverting the cumulative distribution function corresponding to the density $\rho(z)$ and making use of uniform random numbers in the standard manner.

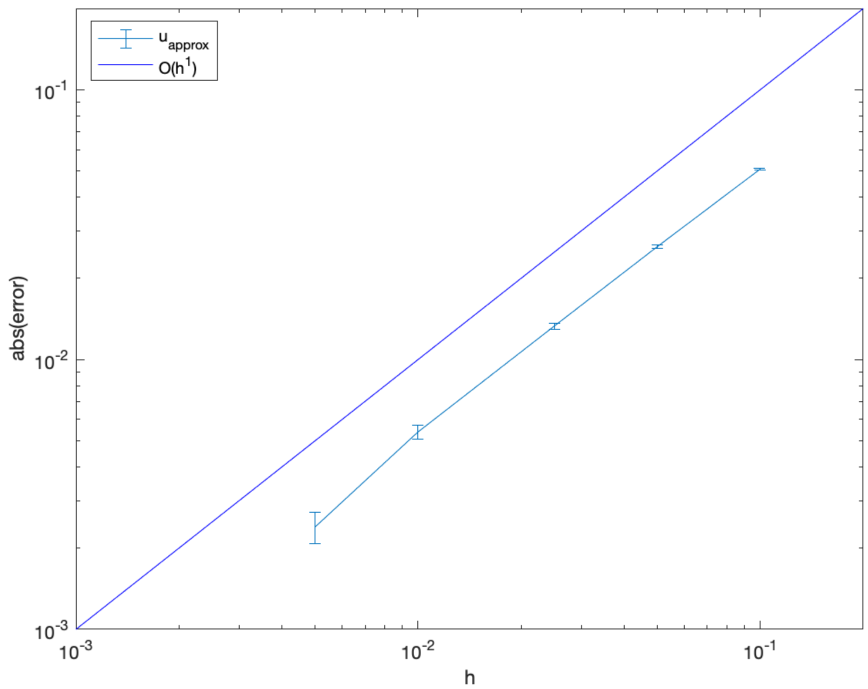

Figure 6.2.1: Non-singular Lévy measure example: dependence of the error e on h, the error bars show the Monte Carlo error. The parameters used are $T = 1, C_+ = 30, C_- = 1.0, \mu = 3.0, f = 0.1, M = 40000000$ and $\hat{u}$ is evaluated at the point $(0, 0)$.

Here the absolute error e is given by

$$e = |\hat{u} - u|, \tag{6.2.5}$$

Table 6.2.1: Non-singular Lévy measure example. The parameters are the same as in Figure 6.2.1. The column $\hat{\varkappa}$ gives the sample average of the number of steps together with its Monte Carlo error.

h	$\hat{u}$	$2\sqrt{\hat{D}_M}$	e	$\hat{\varkappa}$
0.1	0.9367	0.0004	0.0507	7.72 ± 0.0037
0.05	0.961244	0.0004	0.0262	11.04 ± 0.0056
0.025	0.9742	0.0004	0.0133	17.85 ± 0.0096
0.01	0.9821	0.0003	0.0054	37.85 ± 0.0217
0.005	0.9850	0.0003	0.0024	70.90 ± 0.0416

where the true solution for the point $(0,0)$ is $u = u(0,0) \approx 0.987433$. The expected convergence order $O(h)$ can be clearly seen in Figure 6.2.1 and Table 6.2.1.

6.3 EXAMPLE WITH A SINGULAR LÉVY MEASURE

In this section, we confirm dependence of the error of Algorithm 1 on the cut-off parameter ϵ for jump sizes and on the parameter α of the Lévy measure as well as associated computational costs which were derived in Section 5.3.4.

Example 6.3.1 (Singular Lévy measure)**.** Consider the problem (4.1.1) with $d = 3$, $G = U_1$ which is the open unit ball centred at the origin in $\mathbb{R}^3$, and with the coefficients as in (6.2.1), (6.2.2), and

$$g(t,x) := \frac{1}{2}e^{t-T}(1.21 - x_1^4 - x_2^4) + 6(1 - \frac{1}{2}e^{t-T})\left[x_1^2(1.21 - x_2^2 - x_3^2) + x_2^2\right] \tag{6.3.1}$$

$$\begin{aligned}
&+ (1 - \frac{1}{2}e^{t-T})\Big[(C_+ - C_-)f\left(\frac{4}{\mu} + \frac{4}{\mu^2}\right)(x_1^3 + x_2^3) \\
&+ (C_+ + C_-)f^2\left(\frac{6}{2-\alpha} + \frac{6}{\mu} + \frac{12}{\mu^2} + \frac{12}{\mu^3}\right)(x_1^2 + x_2^2) \\
&+ (C_+ - C_-)f^3\left(\frac{4}{3-\alpha} + \frac{4}{\mu} + \frac{12}{\mu^2} + \frac{24}{\mu^3} + \frac{24}{\mu^4}\right)(x_1 + x_2) \\
&+ (C_+ + C_-)f^4\left(\frac{2}{4-\alpha} + \frac{2}{\mu} + \frac{8}{\mu^2} + \frac{24}{\mu^3} + \frac{48}{\mu^4} + \frac{48}{\mu^5}\right)\Big],
\end{aligned}$$

with the boundary condition (6.2.4), and with the Lévy measure density

$$\nu(\mathrm{d}z) = \begin{cases} C_- \mathrm{e}^{-\mu(|z|-1)}\mathrm{d}z, & \text{if } z < -1, \\ C_- |z|^{-(\alpha+1)}\mathrm{d}z, & \text{if } -1 \le z < 0, \\ C_+ |z|^{-(\alpha+1)}\mathrm{d}z, & \text{if } 0 < z \le 1, \\ C_+ \mathrm{e}^{-\mu(|z|-1)}\mathrm{d}z, & \text{if } z > 1, \end{cases} \tag{6.3.2}$$

where C_-, C_+, and μ are some positive constants and $\alpha \in (0, 2)$.

We observe that $C_- \neq C_+$ gives an asymmetric jump measure and the Lévy process has infinite activity and, if $\alpha \in [1, 2)$, infinite variation. Note that, keeping in mind Remark 5.2.2, the coefficients from (6.2.1), (6.2.2), (6.3.1) satisfy Assumptions 5.2.1-5.2.2.

It is not difficult to verify that this problem has the following solution

$$u(t, x) = (1 - \tfrac{1}{2}\mathrm{e}^{t-T})(1.21 - x_1^4 - x_2^4).$$

Other quantities needed for the algorithm take the form

$$\gamma_\epsilon = (C_+ - C_-)\frac{1 - \epsilon^{1-\alpha}}{1-\alpha}, \quad \alpha \neq 1,$$
$$B_\epsilon = (C_+ + C_-)\frac{\epsilon^{2-\alpha}}{2-\alpha},$$
$$\beta_\epsilon = \sqrt{B_\epsilon} = \sqrt{(C_+ + C_-)\frac{\epsilon^{2-\alpha}}{2-\alpha}},$$

and moreover,

$$\lambda_\epsilon = \int_{|z|>\epsilon} \nu(\mathrm{d}z) = (C_+ + C_-)\left(\frac{1}{\mu} + \frac{\epsilon^{-\alpha} - 1}{\alpha}\right),$$
$$\rho_\epsilon(z) = \frac{1}{\lambda_\epsilon}[C_- \mathrm{e}^{-\mu(|z|-1)}\mathbf{I}(z < -1) + C_- |z|^{-(\alpha+1)}\mathbf{I}(-1 \le z < -\epsilon)$$
$$+ C_+ |z|^{-(\alpha+1)}\mathbf{I}(\epsilon < z \le 1) + C_+ \mathrm{e}^{-\mu(|z|-1)}\mathbf{I}(z > 1)].$$

In this example, the absolute error e is given by

$$e = |\hat{u}^{\epsilon} - u|. \tag{6.3.3}$$

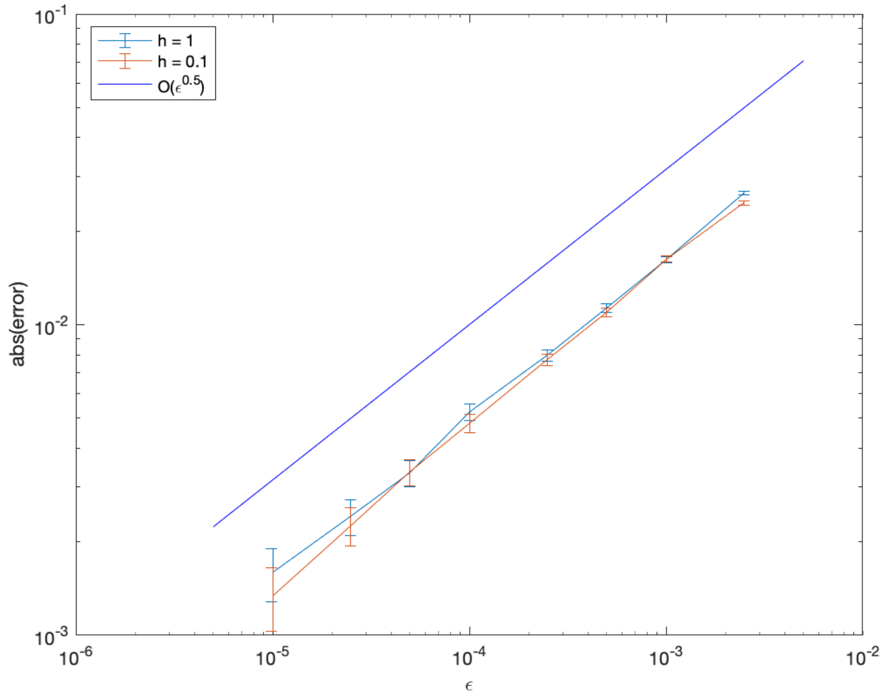

Figure 6.3.1: Singular Lévy measure example, the case $\alpha = 0.5$: dependence of the error e on ϵ, the error bars show the Monte Carlo error. The parameters used are $T = 1, C_+ = 0.1, C_- = 1.0, \mu = 3.0, f = 0.2, M = 40000000$ and $\hat{u}$ is evaluated at the point $(0, 0)$.

For the case of $\alpha = 0.5$, we can clearly see in Figure 6.3.1 and Table 6.3.1 that the error is of order $O(\epsilon^{\alpha}) = O(\epsilon^{0.5})$ as expected. We also observe linear convergence as shown in Figure 6.3.2 in computational cost (measured in average number of steps). Furthermore, we note that choosing a smaller time step, e.g. $h = 0.1$, does not change the behaviour in this case which is in accordance with our prediction of Section 5.3.4

Numerical results for the case $\alpha = 1.5$ are given in Figures 6.3.3 and 6.3.4 and Tables 6.3.2 and 6.3.3. As is shown in Section 5.3.4, convergence (in terms of computational costs) can be improved in the case of $\alpha \in (1, 2)$ by choosing $h = \epsilon^{1+\alpha}$. In Figure 6.3.4, for all ϵ it can be seen that choosing a smaller (but

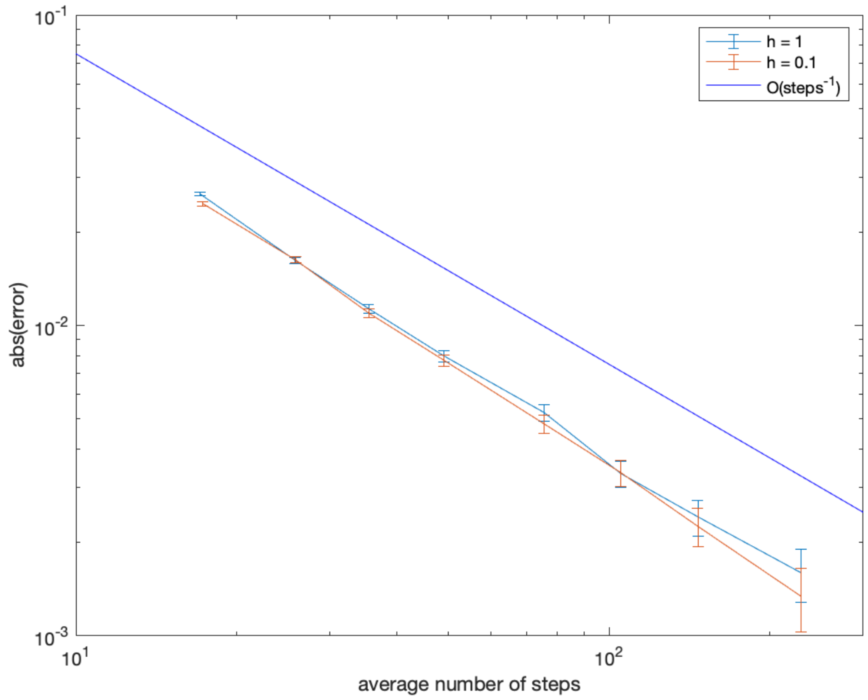

Figure 6.3.2: Singular Lévy measure example, the case $\alpha = 0.5$: dependence of the error e on the average number of steps (computational costs). The parameters are the same as in Figure 6.3.1.

optimally chosen) step parameter h results in quicker convergence (i.e., for the same cost, we can achieve a better result if h is chosen in an optimal way) and naturally in a smaller error.

We recall that if the jump measure is symmetric, i.e. $C_- = C_+$ in the considered example, then $\gamma_\epsilon = 0$ and the numerical integration error of Algorithm 1 is no longer singular (see Theorem 5.3.10 and Remark 5.3.12). Consequently (see Section 5.3.4), in this case the computational cost depends linearly on ϵ even for $\alpha = 1.5$, which is confirmed on Figure 6.3.5.

6.4 FX OPTION PRICING UNDER A LÉVY-TYPE CURRENCY EXCHANGE MODEL

In this section, we demonstrate the use of Algorithm 1 for pricing financial derivatives where the underlying follow a Lévy process. We apply the

Table 6.3.1: Singular Lévy measure example for $\alpha = 0.5$ and $h = 1$. The parameters are the same as in Figure 6.3.1. The column $\hat{\varkappa}$ gives the sample average of the number of steps together with its Monte Carlo error.

ϵ	$\hat{u}$	$2\sqrt{\hat{D}_M}$	e	λ_ϵ	γ_ϵ	$\hat{\varkappa}$
0.0025	0.9610	0.0004	0.0265	42.2	-1.71	17.10 ± 0.0096
0.001	0.9713	0.0004	0.0162	67.7	-1.74	25.78 ± 0.0149
0.0005	0.9761	0.0004	0.0113	96.6	-1.76	35.45 ± 0.0208
0.00025	0.9795	0.0003	0.0080	137.3	-1.77	48.96 ± 0.0290
0.0001	0.9822	0.0003	0.0052	218.2	-1.78	75.53 ± 0.0452
0.00005	0.9841	0.0003	0.0033	309.3	-1.79	105.32 ± 0.0633
0.000025	0.9850	0.0003	0.0024	438.2	-1.79	147.07 ± 0.0888
0.00001	0.9858	0.0003	0.0016	693.9	-1.79	229.51 ± 0.1393

Table 6.3.2: Singular Lévy measure example for $\alpha = 1.5$ and $h = 1$. The parameters are the same as in Figure 6.3.3 and Figure 6.3.4. The column $\hat{\varkappa}$ gives the sample average of the number of steps together with its Monte Carlo error.

ϵ	$\hat{u}$	$2\sqrt{\hat{D}_M}$	e	λ_ϵ	γ_ϵ	$\hat{\varkappa}$
0.05	1.0862	0.0011	0.0988	1541.7	-166.7	15.47 ± 0.002
0.04	1.0814	0.0011	0.0939	2158.0	-192.0	20.38 ± 0.003
0.03	1.0683	0.0010	0.0809	3327.1	-229.1	29.53 ± 0.005
0.02	1.0499	0.0010	0.0625	6119.6	-291.4	51.02 ± 0.008
0.01	1.0216	0.0010	0.0342	17324.7	-432.0	135.63 ± 0.022
0.009	1.0187	0.0010	0.0313	20292.4	-458.0	157.88 ± 0.026
0.008	1.0158	0.0010	0.0284	24215.4	-488.7	187.25 ± 0.030

algorithm to estimate the price of a foreign exchange (FX) barrier basket option. A barrier basket option gives the holder the right to buy or sell a certain basket of assets (here foreign currencies) at a specific price K at maturity T in the case when a certain barrier event has occurred. The most used barrier-type options are knock-in and knock-out options. This type of option becomes active (or inactive) in the case of the underlying price $S(t)$ reaching a certain threshold (the barrier) B before reaching its maturity. In most cases barrier option prices cannot be given explicitly and therefore have to be approximated. We illustrate that the algorithm successfully works in the multidimensional case in Example 6.4.1 and also experimentally demonstrate the convergence orders in Example 6.4.2, where Assumptions 5.2.3-5.2.5 do not hold.

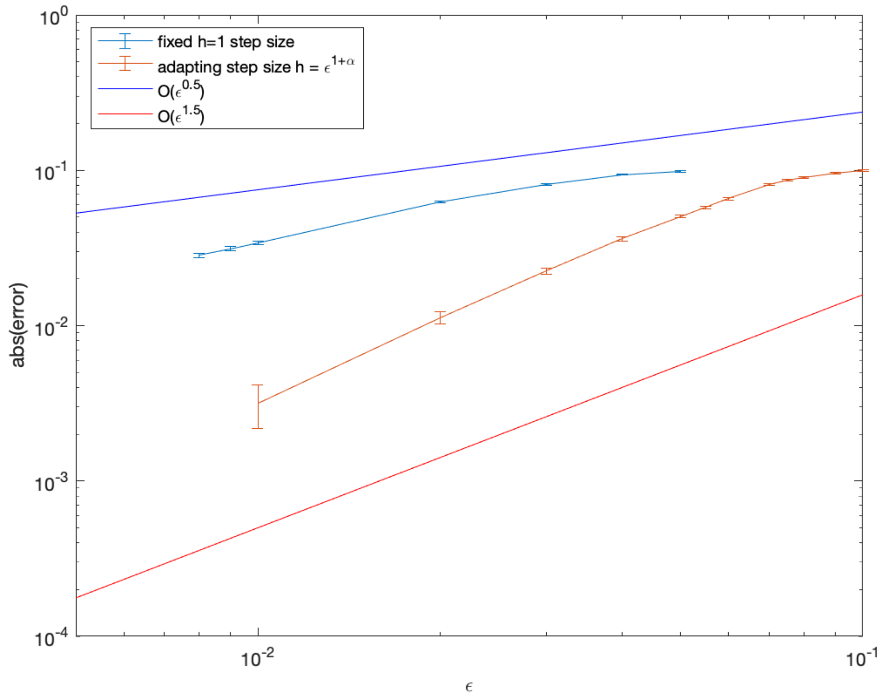

Figure 6.3.3: Singular Lévy measure example, the case $\alpha = 1.5$: dependence of the error e on ϵ, the error bars show the Monte Carlo error. The parameters used are $T = 1, C_+ = 1.0, C_- = 25.0, \mu = 3.0, f = 1.0, M = 100000000$ and $\hat{u}$ is evaluated at the point $(0, 0)$.

Example 6.4.1 (Barrier basket option pricing)**.** Let us consider the case with five currencies: GBP, USD, EUR, JPY and CHF, and let us assume that the domestic currency is GBP. We denote the corresponding spot exchange rates as

$$S_1(t) = S_{USDGBP}(t),\ S_2(t) = S_{EURGBP}(t),$$
$$S_3(t) = S_{JPYGBP}(t),\ S_4(t) = S_{CHFGBP}(t),$$

where $S_{FORDOM}(t)$ describes the amount of domestic currency DOM one pays/receives for one unit of foreign currency FOR (for more details see Section 9 or [113, 20]). We assume that under a risk-neutral measure $\mathbb{Q}$ the dynamics for the spot exchange rates can be written as

$$S_i(t) = S_i(t_0)\exp((r_{GBP} - r_i)(t - t_0) + X_i(t)), \qquad i = 1, 2, 3, 4,$$

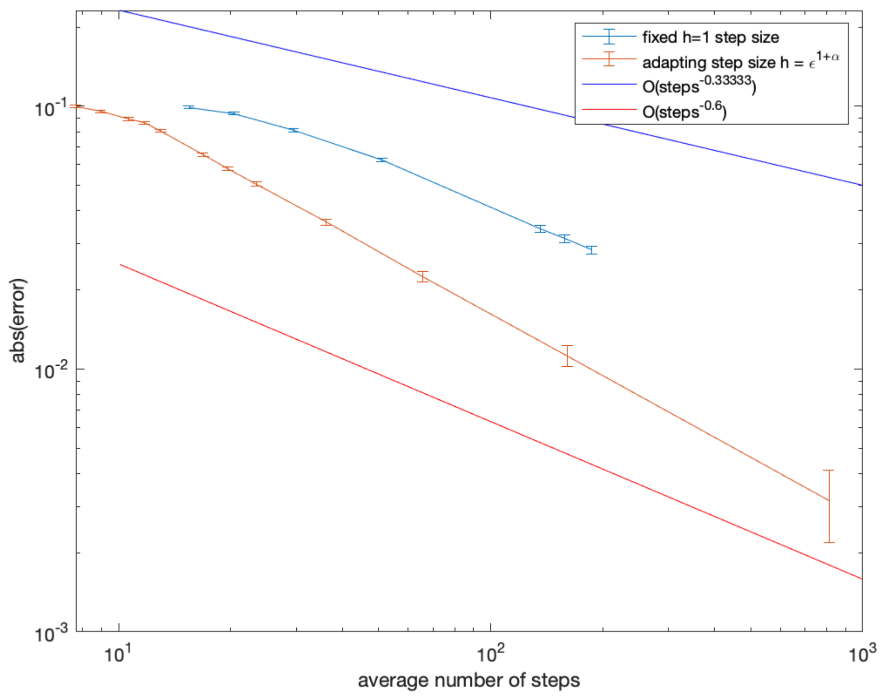

Figure 6.3.4: Singular Lévy measure example, the case $\alpha = 1.5$: dependence of the error e on the average number of steps (computational costs), the error bars show the Monte Carlo error. The parameters are the same as in Figure 6.3.3.

where r_i are the corresponding short rates of USD, EUR, JPY, CHF and r_{GBP} is the short rate for GBP, which are for simplicity assumed to be constant; and $X(t)$ is a 4-dimensional Lévy process similar to (5.1.1) with a single jump noise:

$$X(t) = \int_{t_0}^{t} b(t, X(s-))ds + \int_{t_0}^{t} \sigma(s, X(s-))dw(s) + \int_{t_0}^{t} \int_{\mathbb{R}} F(s, S(s-))z\hat{N}(dz, ds). \tag{6.4.1}$$

Here $w(t) = (w_1(t), w_2(t), w_3(t), w_4(t))^\top$ is a 4-dimensional standard Wiener process. As $\nu(z)$, we choose the Lévy measure with density (6.3.2) as in Example 6.3.1 and we take $F(t,x) = (f_1, f_2, f_3, f_4)^\top$. We also assume that $\sigma(s,x)$ is a constant 4×4 matrix.

The risky asset for a domestic GBP business are the foreign currencies $Y_i(t) = B_i(t) \cdot S_i(t)$, where $B_i(t)$ denotes the foreign currency (account).

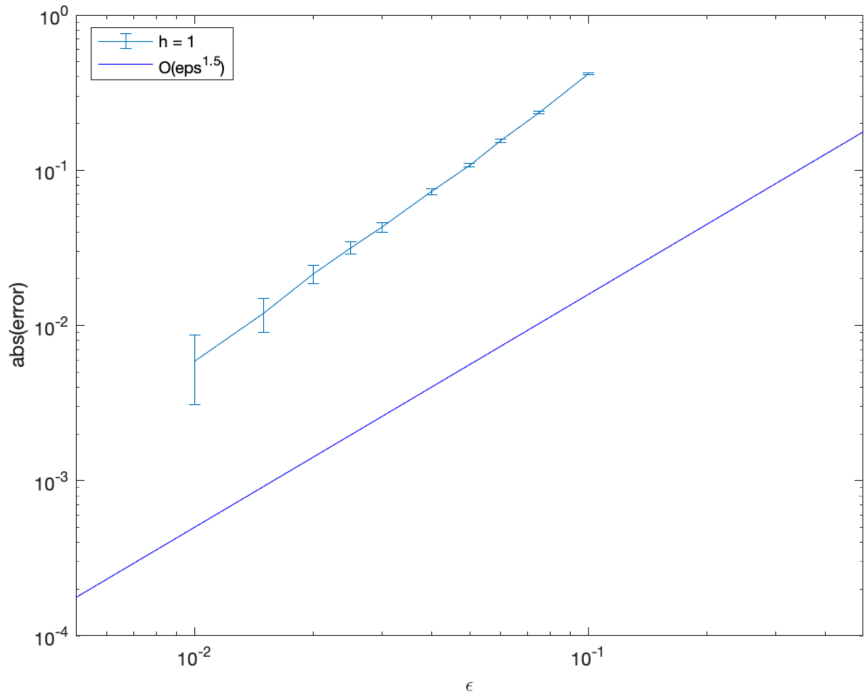

Figure 6.3.5: Dependency of ϵ on *error* plot for a simulation example with symmetric singular Lévy measure for $\alpha = 1.5$. The parameters used are $T = 1, C_+ = 0.5, C_- = 0.5, \mu = 3.0, f = 1.0, M = 100000000$ and $\hat{u}$ is evaluated at the point $(0,0)$.

Under the measure $\mathbb{Q}$ all the discounted assets $\widetilde{Y}_i(t) = e^{(r_i - r_{GBP})(t-t_0)} S_i(t) = S_i(t_0)\exp(X_i(t))$ have to be martingales on the domestic market (therefore discounted by the domestic interest rate) to avoid arbitrage. Using the Ito formula for Lévy processes, we can derive the SDEs for $\widetilde{Y}_i$ (see e.g. [4, p.288]):

$$\frac{d\widetilde{Y}_i}{\widetilde{Y}_i} = \left[b_i(t, X(s-)) + \frac{1}{2}\sum_{j=1}^{4} \sigma_{ij}^2 + \int\limits_{|z|<1} \left(e^{f_i z} - 1 - f_i z \right) \nu(dz) \right] dt \qquad (6.4.2)$$
$$+ \sum_{j=1}^{4} \sigma_{ij} dw_j(s) + \int\limits_{\mathbb{R}} \left(e^{f_i z} - 1 \right) \hat{N}(dz, ds).$$

Therefore, for all $\widetilde{Y}_i$ to be martingales, the drift component b_i has to be so that

$$b_i = -\frac{1}{2}\sum_{j=1}^{4} \sigma_{ij}^2 - \int\limits_{\mathbb{R}} \left(e^{f_i z} - 1 - f_i z \mathbf{I}_{|z|<1} \right) \nu(dz) \qquad (6.4.3)$$

Table 6.3.3: Singular Lévy measure example for $\alpha = 1.5$ and adjusted $h = \epsilon^{1+\alpha}$. The parameters are the same as in Figures 6.3.3 and 6.3.4.

ϵ	h	$\hat{u}$	$2\sqrt{\hat{D}_M}$	e	λ_ϵ	γ_ϵ	$\hat{\varkappa}$
0.10	3.16×10^{-3}	1.0872	0.0011	0.0998	540	−104	7.68 ± 0.001
0.09	2.43×10^{-03}	1.0829	0.0011	0.0955	633	−112	8.97 ± 0.001
0.08	1.81×10^{-03}	1.0769	0.0011	0.0895	757	−122	10.62 ± 0.002
0.075	1.54×10^{-03}	1.0739	0.0011	0.0864	835	−127	11.69 ± 0.002
0.07	1.29×10^{-03}	1.0680	0.0011	0.0806	927	−133	13.00 ± 0.002
0.06	8.82×10^{-04}	1.0530	0.0011	0.0655	1171	−148	16.92 ± 0.003
0.055	7.09×10^{-04}	1.0453	0.0011	0.0579	1335	−157	19.70 ± 0.003
0.05	5.59×10^{-04}	1.0380	0.0011	0.0506	1542	−167	23.50 ± 0.004
0.04	3.20×10^{-04}	1.0236	0.0010	0.0362	2158	−192	36.19 ± 0.006
0.03	1.56×10^{-04}	1.0099	0.0010	0.0225	3327	−229	65.66 ± 0.011
0.02	5.66×10^{-05}	0.9987	0.0010	0.0112	6120	−291	160.57 ± 0.026
0.01	1.00×10^{-05}	0.9906	0.0010	0.0032	17325	−432	812.35 ± 0.132

$$= -\frac{1}{2}\sum_{j=1}^{4}\sigma_{ij}^2 - \frac{C_-}{\mu + f_i}e^{-f_i} - \frac{C_+}{\mu - f_i}e^{f_i} - \frac{C_+ - C_-}{\mu} - I_i(\alpha, C_+, C_-),$$

where

$$I_i(\alpha, C_+, C_-) = \sum_{n=2}^{\infty}\frac{(C_+ + C_-(-1)^n)f_i^n}{n!(n-\alpha)}.$$

We also note that

$$\int_{|z|>1} e^{f_i z}\nu(dz) < \infty$$

is satisfied by (6.3.2) if $f_i < \mu$.

Let us consider a down-and-out (DAO) put option, which can be written as

$$P_{t_0}(T, K) = \exp^{-r_{GBP}(T-t_0)} \qquad (6.4.4)$$
$$\mathbb{E}\left[\mathbf{I}\left(\min_{t_0 \le t \le T} S(t) > B\right)\max\left(K - \sum_{i=1}^{4} w_i S_i(T), 0\right)\right],$$

where $\mathbf{I}\left(\min_{t_0 \le t \le T} S(t) > B\right) = 1$ if for all of the underlying exchange rates $S_i(t) > B_i,\ t_0 \le t \le T$, otherwise it is zero.

We use Algorithm 1 (the algorithm is applied to X from (6.4.1) and then S is computed as $\exp(X)$ to achieve higher accuracy) together with the

Monte Carlo technique to evaluate this barrier basket option price (6.4.4). In Table 6.4.1, market data for the 4 currency pairs are given, and in Table 6.4.2 the option and model parameters are provided, which are used in simulations here.

Table 6.4.1: Market data for 4 currency pairs. Here σ_i are volatilities for the corresponding pairs and ρ_{ij} are the correlation coefficients for the corresponding two pairs.

Market data				Correlation data ρ_{ij}		
currency pair i	$S_i(0)$	r_i	σ_i	USDGBP	EURGBP	JPYGBP
USDGBP	0.81	0.02	0.095			
EURGBP	0.88	0.00	0.089	0.87		
JPYGBP	0.0075	−0.011	0.071	0.94	0.77	
CHFGBP	0.90	0.075	0.110	0.86	0.93	0.96
	r_{GBP}	0.01				

Table 6.4.2: Option and model parameters for Example 6.4.1

Option parameter					Model parameter		
currency pair	Barrier B_i	w_i			jump factor f_i	α	1.5
USDGBP	0.50	0.20	t_0	0.0	0.10	C_+	0.3
EURGBP	0.60	0.25	T	1.0	0.15	C_-	1.2
JYNGBP	0.0045	0.45	K	0.5	0.05	μ	3.0
CHFGBP	0.55	0.10			0.12	M	10^6

To find the matrix $\sigma = \{\sigma_{ij}\}$ used in the model (6.4.1), we form the matrix a using the volatility σ_i and correlation coefficient data from Table 6.4.1 in the usual way, i.e., $a_{ii} = \sigma_i^2$ and $a_{ij} = \sigma_i \sigma_j \rho_{ij}$ for $i \neq j$. Then the matrix σ is the solution of $\sigma\sigma^\top = a$ obtained by the Cholesky decomposition.

The results of the simulations are presented in Figure 6.4.1 for different choices of ϵ and different choices of h. In Figure 6.4.2, it can be seen that (similar to Example 6.3.1) by choosing the step size h optimally results in a better approximation for the same cost.

In this example we demonstrated that Algorithm 1 can be successfully used to price a FX barrier basket option involving 4 currency pairs following an exponential Lévy model despite the considered problem not satisfying Assumptions 5.2.3-5.2.5 of Section 5.2.1. In particular, we note that the

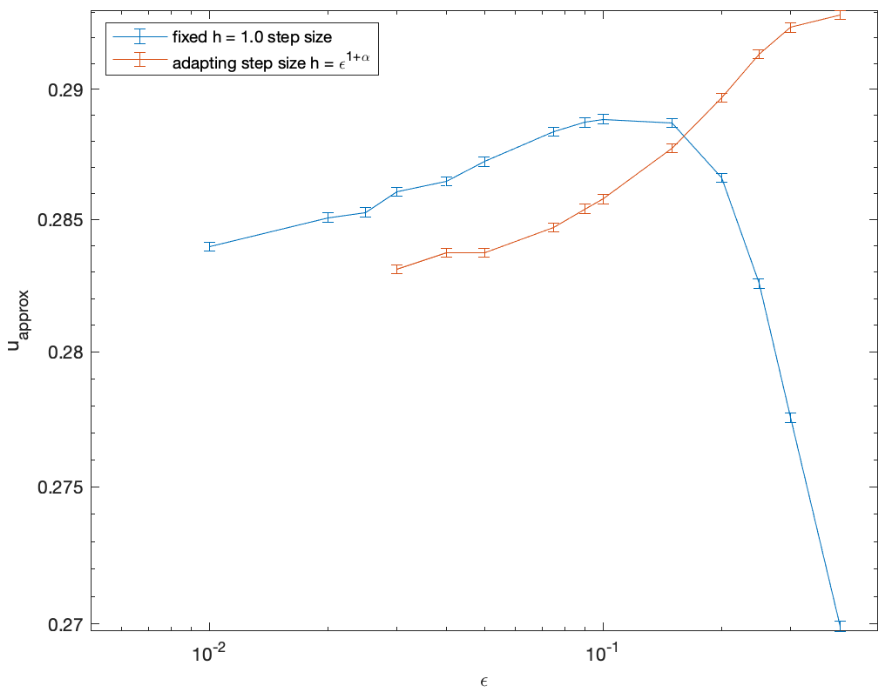

Figure 6.4.1: Dependence of the approximate price of the FX barrier basket option on ϵ for different choices of h. The error bars show the Monte Carlo error.

algorithm is easy to implement and it gives sufficient accuracy with relatively small computational costs. Moreover, application of Algorithm 1 can be easily extended to other multi-dimensional barrier option (and other types of options and not only on FX markets), while other approximation techniques such as finite difference methods or Fourier transform methods typically cannot cope with higher dimensions.

Example 6.4.2 (Barrier option pricing: one currency pair)**.** In this example, we demonstrate that the convergence orders and computational costs discussed in Section 5.3.4 appear to hold, despite the considered problem not satisfying Assumptions 5.2.3-5.2.5 of Section 5.2.1.

Let us consider the case with two currencies: GBP and USD. As before, we assume that the domestic currency is GBP. The corresponding spot exchange rate is

$$S(t) = S_{USDGBP}(t).$$

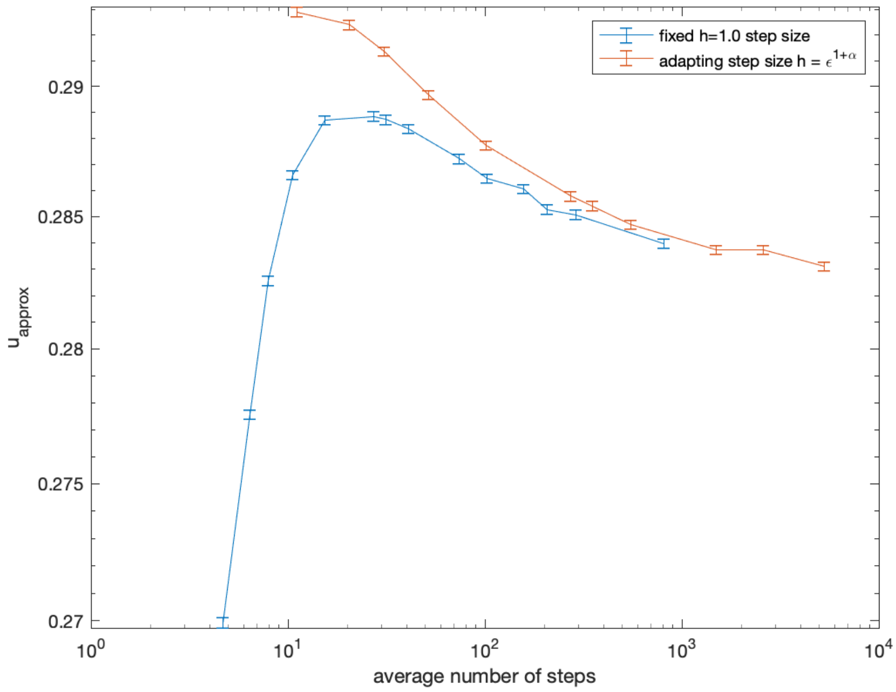

Figure 6.4.2: Dependence of the approximate price of the FX barrier basket option on average number of steps (computational costs) for different choices of h. The error bars show the Monte Carlo error.

We assume the same dynamics under a risk-neutral measure $\mathbb{Q}$ for the spot exchange rates as in Example 6.4.1. Moreover, $X(t)$ is a 1-dimensional Lévy process as defined in (6.4.1) but for one dimension only. Following the same fashion as in Example 6.4.1, the risky asset for a domestic GBP business is the foreign currency $Y(t) = B(t) \cdot S(t)$, where $B(t)$ denotes the foreign currency (account) and under the measure $\mathbb{Q}$ the discounted asset $\widetilde{Y}(t)$ has to be a martingale on the domestic market to avoid arbitrage. Using the Ito formula for Lévy processes, we can derive the SDE for $\widetilde{Y}$ as we did in (6.4.2)-(6.4.3). We compute the value for a DAO put option (cf. (6.4.4)):

$$P_{t_0}(T,K) = \exp^{-r_{GBP}(T-t_0)} \mathbb{E}\left[\mathbf{I}\left(\min_{t_0 \le t \le T} S(t) > B\right) \max\left(K - S(T), 0\right)\right]. \tag{6.4.5}$$

The approximate solution $\hat{P} = \hat{P}_{t_0}(T,K)$ is obtained by applying Algorithm 1 directly to the SDE for $S(t)$. To study the dependence of the error of Algorithm 1 on the cut-off parameter ϵ for jump sizes and on the parameter

α of the Lévy measure as well as associated computational costs, we need to compare the approximation $\hat{P}$ with the true price $P_{t_0}(T,K)$. However, in this example, we do not have the exact price, and therefore need to accurately simulate a reference solution. To this end, as in Example 6.4.1, we apply Algorithm 1 to $X(t)$ and use a sufficiently small ϵ and h and also a large number of Monte Carlo simulations M (see Tables 6.4.5 and 6.4.9). We denote this reference solution as $\hat{P}^{ref} = \hat{P}^{ref}_{t_0}(T,K)$. In this example the absolute error e_{ref} of Algorithm 1 is evaluated as

$$e_{ref} = |\hat{P} - \hat{P}^{ref}|.$$

In Table 6.4.3, market data for the currency pair are given, and in Table 6.4.4 the option and model parameters are provided, which are used in simulations here.

Table 6.4.3: Market data for the currency pair. Here σ is the volatility.

Market data			
currency pair	$S(0)$	r_{USD}	σ
USDGBP	0.81	0.02	0.095
	r_{GBP}	0.01	

Table 6.4.4: Option and model parameters for Example 6.4.2

Option parameter				
currency pair	Barrier B	t_0	T	K
USDGBP	0.50	0.0	1.0	0.5

Model parameter					
jump factor f	α	C_+	C_-	μ	M
0.10	0.5	0.3	1.2	3.0	10^8
0.10	1.5	0.3	1.2	3.0	10^8

The results of the simulations for $\alpha = 0.5$ are presented in Figures 6.4.3 and 6.4.4 and in Tables 6.4.6 and 6.4.7 for different choices of ϵ and fixed $h = 1.0$ and $h = 0.1$. We can clearly see that the error is of order $O(\epsilon^{\alpha}) = O(\epsilon^{0.5})$

as expected. We also observe linear convergence in computational cost (measured in average number of steps).

Table 6.4.5: Reference solution $\hat{P}^{ref}$ for singular Lévy measure example for $\alpha = 0.5$.

M	ϵ	h	$\hat{u}$
10^8	5×10^{-5}	1×10^{-5}	0.28951

$2\sqrt{\hat{D}_M}$	λ_ϵ	γ_ϵ	$\hat{\varkappa}$
8.7×10^{-6}	421.8	-1.7873	98223.5 ± 5.7

Table 6.4.6: FX barrier option example for $\alpha = 0.5$ and $h = 1$.

ϵ	$\hat{u}$	$2\sqrt{\hat{D}_M}$	e_{ref}	λ_ϵ	γ_ϵ	$\hat{\varkappa}$
0.002	0.29053	2.7×10^{-5}	0.00102	64.6	-1.72	65.39 ± 0.002
0.0015	0.29040	2.7×10^{-5}	0.00089	75.0	-1.73	75.58 ± 0.002
0.001	0.29027	2.7×10^{-5}	0.00076	92.4	-1.74	92.67 ± 0.002
0.0009	0.29024	2.7×10^{-5}	0.00073	97.5	-1.75	97.71 ± 0.002
0.0008	0.29021	2.7×10^{-5}	0.00070	103.6	-1.75	103.67 ± 0.002
0.0007	0.29015	2.7×10^{-5}	0.00064	110.9	-1.75	110.86 ± 0.003
0.0006	0.29012	2.7×10^{-5}	0.00061	120.0	-1.76	119.78 ± 0.003
0.0005	0.29006	2.8×10^{-5}	0.00055	131.7	-1.76	131.25 ± 0.003

Table 6.4.7: FX barrier option example for $\alpha = 0.5$ and $h = 0.1$.

ϵ	$\hat{u}$	$2\sqrt{\hat{D}_M}$	e_{ref}	λ_ϵ	γ_ϵ	$\hat{\varkappa}$
0.002	0.29054	2.7×10^{-5}	0.00103	64.6	-1.72	65.48 ± 0.002
0.0015	0.29043	2.7×10^{-5}	0.00092	75.0	-1.73	75.62 ± 0.002
0.001	0.29027	2.7×10^{-5}	0.00076	92.4	-1.74	92.68 ± 0.002
0.0008	0.29020	2.7×10^{-5}	0.00069	103.6	-1.75	103.67 ± 0.002
0.0007	0.29015	2.7×10^{-5}	0.00064	110.9	-1.75	110.86 ± 0.003
0.0006	0.29011	2.7×10^{-5}	0.00060	120.0	-1.76	119.78 ± 0.003
0.0005	0.29005	2.7×10^{-5}	0.00054	131.7	-1.76	131.26 ± 0.003

Numerical results for the case $\alpha = 1.5$ are given in Figures 6.4.5 and 6.4.6 and in Tables 6.4.9 and 6.4.10. We observe the expected orders of convergence as given in Section 5.3.4. In this example, we experimentally demonstrated that convergence orders and computational cost for Algorithm 1 are consistent with predictions of Section 5.3.4 despite the considered problem not satisfying assumptions of Section 5.2.1.

Table 6.4.8: Reference solution $\hat{P}^{ref}$ for singular Lévy measure example for $\alpha = 1.5$.

M	ϵ	h	$\hat{u}$
10^8	0.001	1×10^{-5}	0.24301

$2\sqrt{\hat{D}_M}$	λ_ϵ	γ_ϵ	$\hat{\varkappa}$
1.0×10^{-5}	31622.3	-55.1	110969.3 ± 2.5

Table 6.4.9: FX barrier option example for $\alpha = 1.5$ and $h = 1$.

ϵ	$\hat{u}$	$2\sqrt{\hat{D}_M}$	e_{ref}	λ_ϵ	γ_ϵ	$\hat{\varkappa}$
0.1	0.24842	3.2×10^{-5}	0.00541	31.1	-3.9	31.78 ± 0.001
0.08	0.24793	3.2×10^{-5}	0.00492	43.7	-4.6	43.83 ± 0.001
0.07	0.24758	3.2×10^{-5}	0.00451	53.5	-5.0	53.23 ± 0.003
0.06	0.24721	3.2×10^{-5}	0.00420	67.5	-5.5	66.69 ± 0.002
0.05	0.24674	3.2×10^{-5}	0.00372	88.9	-6.2	87.20 ± 0.003
0.04	0.24621	3.2×10^{-5}	0.00320	124.5	-7.2	121.26 ± 0.003

Table 6.4.10: FX barrier option example for $\alpha = 1.5$ and adapting step size $h = \epsilon^{1+\alpha}$.

ϵ	h	$\hat{u}$	$2\sqrt{\hat{D}_M}$	e_{ref}	λ_ϵ	γ_ϵ	$\hat{\varkappa}$
0.4	1.01×10^{-1}	0.24634	3.3×10^{-5}	0.00333	3.5	-1.0	12.69 ± 0.0003
0.35	7.25×10^{-2}	0.24678	3.3×10^{-5}	0.00377	4.3	-1.2	16.85 ± 0.0004
0.3	4.93×10^{-2}	0.24682	3.3×10^{-5}	0.00381	5.6	-1.5	23.65 ± 0.0006
0.25	3.13×10^{-2}	0.24636	3.3×10^{-5}	0.00335	7.5	-1.8	35.72 ± 0.0009
0.2	1.79×10^{-2}	0.24549	3.3×10^{-5}	0.00248	10.7	-2.2	59.99 ± 0.0015
0.15	8.71×10^{-3}	0.24468	3.3×10^{-5}	0.00167	16.7	-2.8	118.81 ± 0.0031

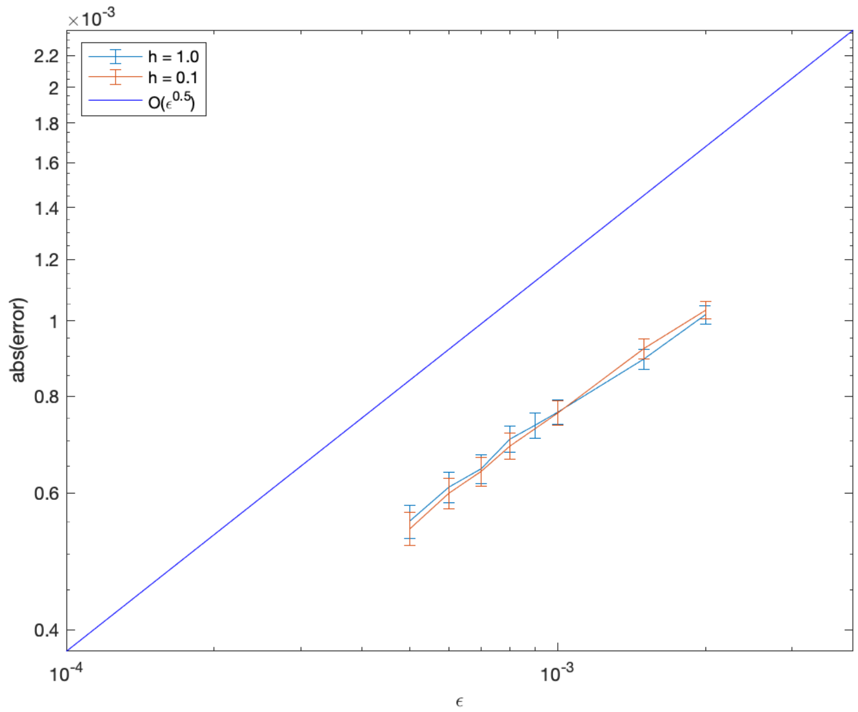

Figure 6.4.3: FX barrier option example, the case $\alpha = 0.5$: dependence of the error on ϵ for different choices of h. The error bars show the Monte Carlo error.

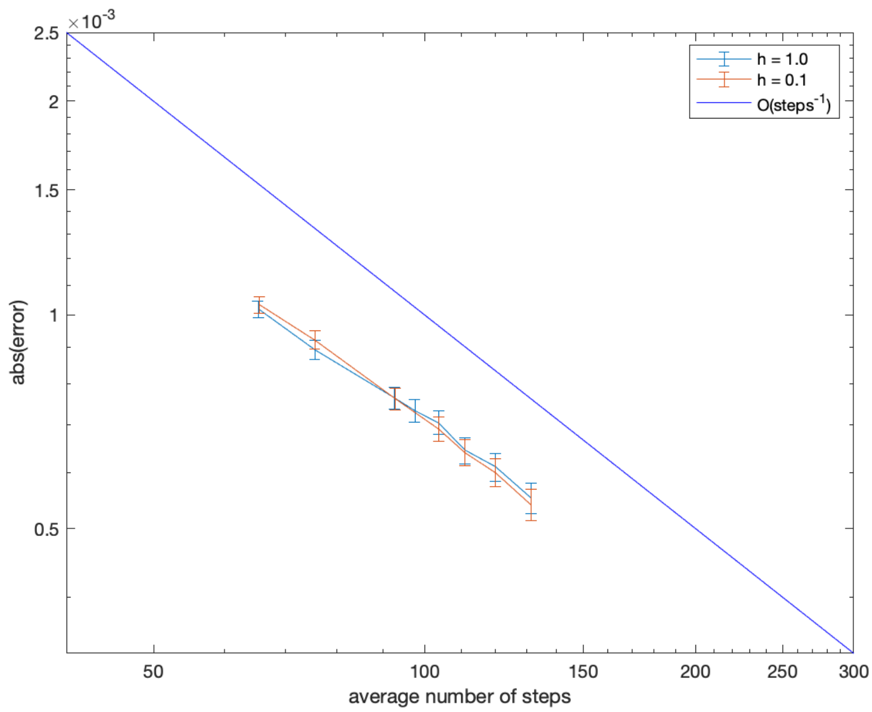

Figure 6.4.4: FX barrier option example, the case $\alpha = 0.5$: dependence of the error e on the average number of steps.

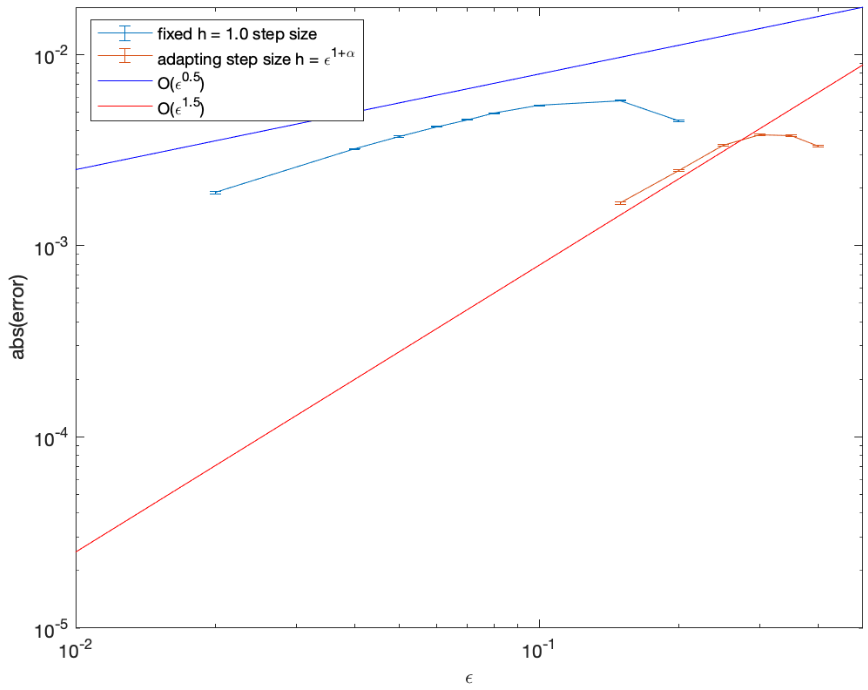

Figure 6.4.5: FX barrier option example, the case $\alpha = 1.5$: dependence of the error e on ϵ, the error bars show the Monte Carlo error.

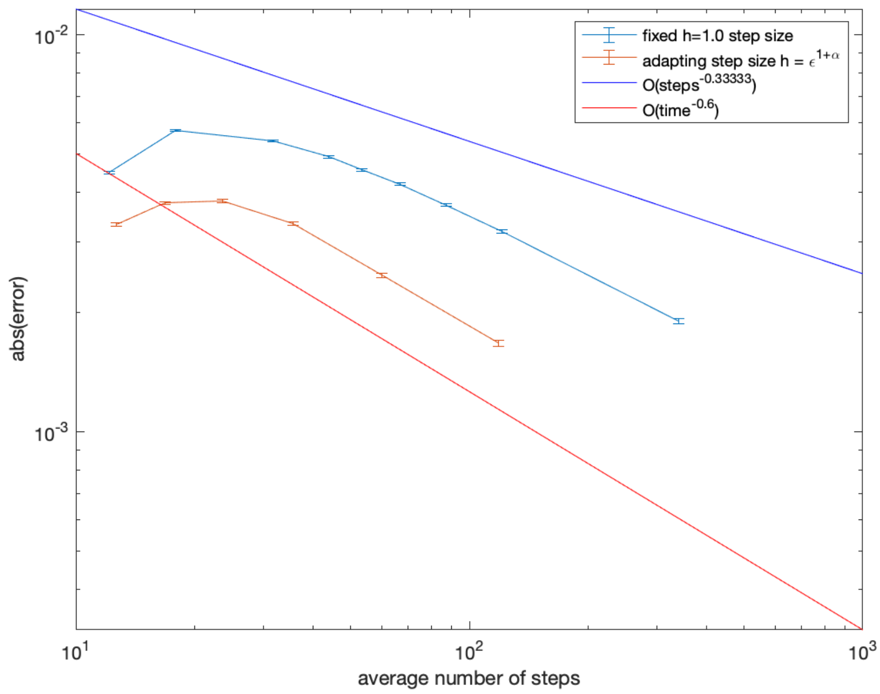

Figure 6.4.6: FX barrier option example, the case $\alpha = 1.5$: dependence of the error e on the average number of steps.

7

CONCLUSIONS AND FUTURE WORK

In this part of the thesis, we have introduced a new algorithm (Algorithm 1), a restricted jump-adaptive numerical scheme, for weak-sense approximations of stochastic differential equations driven by general Lévy processes with infinite activity. In Chapter 5, we highlight the usefulness of the introduced scheme by the connection of a probabilistic representation of the solution of a IPDE problem, as finding the solution comes down to being able to simulate systems of Lévy-driven SDEs efficiently and accurately. This research covers two main ingredients needed to discuss the convergence behaviour of weak approximations of Lévy-driven SDEs with infinite activity. One is that we replace small jumps with an appropriate Brownian motion, which is presented in Section 5.2, where we follow the same approach as [5]. This assures that the numerical approximation to the IPDE problem is computationally feasible in the infinite activity case. Naturally, replacing the small jumps introduces a numerical error (see Theorem 5.2.5), which is part of the total error of the introduced numerical scheme. The second ingredient, and the main part of this research is the in-depth analysis and discussion of the weak-sense error estimate for the algorithm in Section 5.3. This includes the derivation of the error bounds and an analysis of the one-step error with resulting Theorem 5.3.6, followed by the global error estimate in Theorem 5.3.10. It is important to note, that the resulting one-step and global error estimates explicitly show the (singular) dependence of the error on the parameter ϵ, which is the cut-off level for small jumps replaced by the

Brownian motion. We also see, that choice of the Lévy measure is also a key consideration, in particular, with respect to symmetry.

Furthermore, in Section 5.3.4, we give an overall error estimate for solving the IPDE problem described in 4.1 and consider an α-stable process, in which the Lévy measure has a singular behaviour near zero. We also showcase the different possible worst-case convergence orders of the error estimate and computational costs for different levels of activity, which can be described by different regimes of the parameter α.

Finally, we illustrate the theoretical convergence results from Chapter 5 in a range of different examples in Chapter 6. We complement the suggested algorithm with the Monte Carlo technique to get approximations of the corresponding expectations. It is assuring to see, that we are able to produce the exact convergence orders in two theoretical examples. Moreover, we look at the performance of our algorithm in a practical example on pricing (multidimensional) FX barrier options, and its noteworthy, that the introduced numerical scheme also works when our initial model assumptions are not satisfied. We note, that our suggested algorithm combined with Monte Carlo techniques are fairly easy to implement and the extension to problems with higher dimensions uncomplicated and computationally unproblematic.

Overall, this work is focused on dealing with the Dirichlet problem for IPDEs, but some considerations and remarks are made on how some results possibly hold or could be extended to the Cauchy case. Additionally, rather than using the numerical schemes suggested here, one could also explore finite difference methods or Fourier transform methods, although, there might be computational limitations for high-dimensional problems.

Another interesting field to explore, could be the computational aspect of the resulting numerical schemes. In particular in the case of infinite activity, we needed a large amount of Monte Carlo simulations to produce the presented convergence results. Therefore, one might consider programming frameworks to optimise the computational costs, depending on the needed accuracy.

Part II

FX OPTION PRICING USING INTERMEDIATE CURRENCY

8

FX LITERATURE REVIEW AND OVERVIEW

In this Part of the thesis, we present a novel framework for pricing derivatives on the foreign exchange (FX) market. The following sections are part of the paper [86].

As it is well known (see e.g. [15, 113] and also Section 9 here), in the case of a foreign exchange (FX) for two currencies (say, GBP and EUR) no measure is simultaneously risk-neutral for the market on which GBP is the domestic currency and for the market on which EUR is the domestic currency. This can be seen as an asymmetry between the different market views due to the different choice of numeraires. In practice currency pair conventions are usually used in order to standardize option price quotations for each specific currency pair [81, 113]. But this can lead to calibration difficulties. Each of the domestic markets has its own volatility smile curve for options on the corresponding foreign currency. Suppose we want to use a stochastic volatility model given under a risk-neutral measure on the GBP domestic market which we calibrate to the smile for options on EUR. If we re-write this model for the inverse pair, i.e., where options on GBP are traded, in a risk-neutral fashion, we need to calibrate it to the smile on this market as the previously found parameters of the model typically do not match the smile for the inverse pair. This is inconvenient. This situation becomes even more complicated in a multi–currency setting while it is of practical importance to be able to price options on the global FX market in a consistent fashion. With a large number N of currencies, the existence of a consistent FX model is not trivial as a suitable model must preserve relationships between all N currencies and consistency of volatility smiles between all $N(N-1)/2$ cross pairs.

To address these problems of consistent FX modelling, in [33] (see also [35, 34, 26]) the concept of intrinsic currency [33, 34] or artificial currency [26] was introduced. The approach of [33] is based on the idea that each currency has an 'intrinsic value', which is a description of the value of a currency in relation to other currencies. In the intrinsic currency-valuation framework of [33, 34] one models the N intrinsic values of N currencies rather than modelling the $N-1$ exchange rates. In [34] Doust extends his original idea of the intrinsic currency-valuation framework to a SABR-type model, captures the observed volatility smile on the FX market in a multi-currency setting. On a FX market with N currencies, he describes the market with N intrinsic currency values and chooses one (without loss of generality) as the valuation currency and its associated risk–neutral measure, which produces the usual risk–neutral processes for all exchange rates. For option pricing, this approach results in a closed form solution similar to the original SABR model by Hagan et al. [54] adapted to the intrinsic currency–valuation framework, which allows the pricing of FX vanilla options on one currency pair considering the correlation effects of all N currencies. In [26] N exchange rates between an artificial currency and N real currencies are modelled under a risk-neutral measure associated with the artificial currency so that all relationships (in particular, the inversion property that the exchange rate for a pair of real currencies and for their inverse satisfy SDEs of a similar form) between N currencies are satisfied.

Here we explore a very simple but very valuable from the practical angle idea: find a numeraire with respect to which we can price all FX derivatives traded on any of the domestic markets simultaneously under the same measure. This resolves the issue highlighted above: models for different currency pairs can be calibrated to all smiles in a consistent manner. For instance, in the case of two currencies, it is sufficient to calibrate a model on e.g. the GBP domestic market and the smile on the EUR domestic market is automatically reproduced without any need of additional calibration.

We show that such a numeraire exists via introducing the concept of an intermediate pseudo-currency. The main difference with [33, 35, 34, 26] is that the pseudo-currency is explicitly defined via exchange rates of real currencies, while in [33, 35, 34, 26] exchange rates of real currencies are described via an artificial currency. Consequently, we naturally model $N-1$ exchange rates, not N as in [33, 35, 34, 26]. Further, we can use three modelling approaches. The first one is the traditional modelling way in Financial Mathematics, where we start from a stochastic model for $N-1$ exchange rates under a 'market' measure and then we introduce a pseudo-currency market which, as we show, has a risk-neutral measure. Under this risk-neutral measure (the intermediate pseudo-currency is used as the numeraire) we can price FX products on all currency markets simultaneously which guarantees consistency of volatility smiles and other natural relationships between currencies (e.g., the foreign-domestic symmetry). This approach allows us to start with popular stochastic volatility models (e.g., Heston or SABR) written under a 'market' measure and derive the corresponding consistent models on the pseudo-currency market. Alternatively, in the second approach, from the start we model exchange rates under a risk-neutral measure or under a forward measure associated with the pseudo-currency market. The third approach is model-free (see [6, 40, 7, 27] and references therein), where we reconstruct a risk-neutral measure or a forward measure from volatility smiles. We note that the intermediate pseudo-currency in comparison with the intrinsic currency of [33] does not have a financial interpretation, but our focus here is solely on consistent calibration and modelling of exchange rates.

The rest of this Part is organized as follows. In Chapter 9 we present some standard FX market conventions, recall that there is no measure which is simultaneously risk-neutral for both domestic and foreign FX markets and also recall the foreign-domestic symmetry. A convenient numeraire and the associated intermediate pseudo-currency market are introduced in Section 10, where the corresponding pricing formulas for FX options are also derived. This is done for clarity of the exposition in the case of a single currency pair. We extend the intermediate pseudo-currency concept to the multi–currency setting in Section 10.2. In Section 11 we illustrate the concept by first applying it to the Heston model [56, 11] and SABR [54]. Then, for further illustration, we model the spot exchange rate using an extended skewed normal distribution. This exchange rate model is an illustration of how one can describe the observed fat-tailed distribution of the log exchange rate (compared to the assumption of log normal). The considered extended skewed normal distribution is constructed by combining one normal and two shifted half-normal distributed random variables and it allows a flexible control of the tails of the spot exchange rate distribution. We note that the use of the extended skewed normal distribution in pricing FX options is somewhat new. Further, we illustrate our FX option pricing mechanism on the model-free approach. We provide some calibration examples in Chapter 12.

9

PRELIMINARIES ABOUT OPTION PRICING ON THE FX MARKET

In this chapter we give a short introduction to the Foreign Exchange market, including currency forwards and options. We will look at the standard currency option pricing formula based on Garman and Kohlhagen [44] and its version in terms of the currency forward price similar to [14]. Further, we describe the occurrence of the volatility smile and its behaviour on the FX market [112, 20]. Lastly, we recall that there is no measure which is simultaneously risk-neutral for both the domestic and the foreign market and also state the foreign-domestic symmetry.

9.1 FOREIGN EXCHANGE MARKET AND FINANCIAL DERIVATIVES

The foreign exchange market (also known as FX or currency market) is the largest financial market in the world. The trading volume in the FX market was an estimated average of \$6.6 trillion per day in 2019 (see [8] for more details). It is a global market where different currencies are traded 24 hours. Interest rate swaps with a trading volume of \$3.2 trillion per day are the most common transaction worldwide. With a trading volume of \$2.0 trillion per day, the second most common transaction was the spot transaction, which is defined as follows.

Definition 9.1.1 (FX spot rate)**.** A **spot FX transaction** is an agreement between two parties, where one party agrees to buy one currency while the other party agrees to sell another currency at an agreed price at the spot date t. The current spot exchange rate is denoted as $S_{c_1/c_2}(t)$, where c_1 (c_2) denotes currency 1 (2).

For example, the spot exchange rate to exchange EUR(€) to USD($) at time t is denoted as

$$S_{€/\$}(t)$$

and is quoted as

$$\frac{\text{units of USD}}{\text{one EUR}},$$

$$S_{\$/€}(t) = \frac{1}{S_{€/\$}(t)}.$$

In currency pairs (e.g. EUR-USD), the first mentioned currency is known as the foreign (or base) currency, while the second is known as the domestic currency (or numeraire) [20, 113].

Holding on to or trading in a foreign currency over time can bare the risk of losing money due to different spot rates. For this reason one method to deal with this foreign exchange risk is a forward contract.

Definition 9.1.2 (FX forward)**.** A **forward FX contract** (or short forward) is a contract between two parties agreed at time t. One party agrees to buy one currency while the other party agrees to sell another currency in the future at time T at a price agreed beforehand. This price is called the **forward exchange rate** and denoted by $F_{c_1/c_2}(t, T)$.

To price a forward fairly, we need to look at the interest rates of both countries which currencies are traded. In this case, we assume that both interest rates are constant. The connection between the interest rates and the forward rate is stated in the following theorem (see [14]), which can be seen as a no-arbitrage condition for exchanges between different currencies.

Theorem 9.1.3. *Let us assume $S_{c_1/c_2}(t)$ is the current spot rate and $F_{c_1/c_2}(t,T)$ is the current forward exchange rate between two currencies c_1 and c_2. Then the following equation holds:*

$$F_{c_1/c_2}(t,T) = S_{c_1/c_2}(t)e^{(r_2-r_1)(T-t)}, \tag{9.1.1}$$

where r_2 and r_1 denote the interest short rates for the domestic and foreign markets, respectively, t denotes the current time and T is the maturity date of the forward.

Proof. Assume an investor can invest money at the foreign interest rate r_1 and at the domestic interest rate r_2. Investing one unit at the domestic rate r_2 at time t will give them $e^{r_2(T-t)}$ at time T. Exchanging one unit to the foreign currency and then investing at the foreign rate r_1 at time t will give them $\frac{e^{r_1(T-t)}}{S_{c_1/c_2}(t)}$ at time T. Currently the investor is dependent on the spot exchange rate $S_{c_1/c_2}(T)$ to exchange his foreign currency back to domestic currency. However they could hedge that risk by buying a forward. Then $F_{c_1/c_2}(t,T)$ is the fair forward exchange rate, which has to hold so that no arbitrage is possible.

$$\begin{aligned} \frac{F_{c_1/c_2}(t,T)}{S_{c_1/c_2}(t)}e^{r_1(T-t)} &= e^{r_2(T-t)} \\ \Leftrightarrow \quad F_{c_1/c_2}(t,T) = S_{c_1/c_2}(t)\frac{e^{r_2(T-t)}}{e^{r_1(T-t)}} &= S_{c_1/c_2}(t)e^{(r_2-r_1)(T-t)} \end{aligned}$$

□

Another financial instrument to hedge risk in regard to exchange rates are options.

Definition 9.1.4 (European FX option)**.** A (plain vanilla European) **FX option** gives the holder the right but not the obligation to buy (call option) or to sell (put option) an amount of one currency for another currency at an agreed currency rate, also called option strike K, at a specific time T.

Currency options are often denoted in put/call pairs, e.g. a put option to sell USD and buy Euro is denoted as USD/EUR.

9.2 FX OPTION PRICING

Before we proceed to find the fair price of a FX option, we will look at the no-arbitrage condition for put and call prices – the put-call parity for FX options. For definiteness, in this chapter we use the EUR-USD and USD-EUR pairs, where we assume that EUR is the foreign currency, while USD is the domestic currency. Obviously, all statements also hold for the inverse pair or any other currency pair.

The value of a put, $P_{\$}(t,T)$, and a call, $C_{\$}(t,T)$, at the maturity time T and stated in the domestic currency USD can be written as follows:

$$\begin{aligned} C_{\$}(T,T) &= (S_{€/\$}(T) - K, 0)_+ := \max(S_{€/\$}(T) - K, 0) \qquad (9.2.1) \\ P_{\$}(T,T) &= (K - S_{€/\$}(T), 0)_+ := \max(K - S_{€/\$}(T), 0) \end{aligned}$$

where $S_{€/\$}(T)$ denotes the spot exchange rate at time T. We can now state the following theorem which links the option price for FX put and calls with the same strike K, maturity T and time left to maturity $T-t$.

Theorem 9.2.1. *The **Put–Call Parity** for FX options is*

$$P_{\$}(t,T) - C_{\$}(t,T) = Ke^{-r_{\$}(T-t)} + S_{€/\$}^{-r_{€}(T-t)}.$$

Proof. The proof is similar to the proof for the put-call parity for stock options and can be found in [31, 20]. The idea is that payoffs of two portfolios with the same payoffs at maturity must have the same price. For FX options, portfolio 1 consists of a long put and a short call, while portfolio 2 consists of a long zero-coupon bond in domestic currency and a short zero-coupon bond in foreign currency. □

To illustrate the idea of option pricing, we can make the following (classical) model assumptions:

- no taxes, no transaction costs, no restrictions on long or short positions,

- the domestic and the foreign interest rate are riskless and constant over time,
- let $(\Omega, \mathcal{F}, P, \mathcal{F}_t)$ be a filtered probability space,
- the spot exchange rate follows a Geometric Brownian Motion (GBM) with the SDE[1]

$$dS = \mu S dt + \sigma S dW, \tag{9.2.2}$$

  where μ is the drift, σ the volatility of the spot exchange rate process $S(t)$ and $W(t)$ a standard Wiener process,
- the solution to (9.2.2) is commonly known (see [57, 65]) and is given as follows:

$$S(T) = S(0)e^{(\mu - \frac{\sigma^2}{2})T + \sigma W(t)}.$$

Proposition 9.2.2 (Garman and Kohlhagen option price)**.** *The arbitrage-free price for a FX call/put option with the payoff function* (9.2.1) *can be written as*

$$C_{\$}(t,T) = e^{-r_{€}(T-t)} S_t \cdot N\left(\frac{\log(\frac{S_{€/\$}(t)}{K}) + (r_{\$} - r_{€} + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\right) \tag{9.2.3}$$

$$- e^{-r_{\$}(T-t)} K \cdot N\left(\frac{\log(\frac{S_{€/\$}(t)}{K}) + (r_{\$} - r_{€} - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\right),$$

$$P_{\$}(t,T) = e^{-r_{\$}(T-t)} K \cdot N\left(\frac{\log(\frac{K}{S_{€/\$}(t)}) - (r_{\$} - r_{€} - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\right) \tag{9.2.4}$$

$$- e^{-r_{€}(T-t)} S_t \cdot N\left(\frac{\log(\frac{K}{S_{€/\$}(t)}) - (r_{\$} - r_{€} + \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\right).$$

To derive the pricing formula (9.2.3) and (9.2.4) for currency options, one can follow the SDE approach which can be found in [15] based on the original paper of Garman and Kohlhagen [44] or see also Appendix B.2.

1 Note that we omit the currency pair notation. We use $S(t)$ here for $S_{€/\$}(t)$.

Further, we can rewrite the call option price in terms of the forward exchange rate defined in (9.1.1):

$$C_{\$}(t,T) = e^{-r_{\$}(T-t)} F_{€/\$}(t,T) \cdot N\left(\frac{\log(\frac{F_{€/\$}(t,T)}{K}) + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}\right) \tag{9.2.5}$$

$$- e^{-r_{\$}(T-t)} K \cdot N\left(\frac{\log(\frac{F_{€/\$}(t,T)}{K}) - \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}\right),$$

$$P_{\$}(t,T) = e^{-r_{\$}(T-t)} K \cdot N\left(\frac{\log(\frac{K}{F_{€/\$}(t,T)}) + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}\right)$$

$$- e^{-r_{\$}(T-t)} F_{€/\$}(t,T) \cdot N\left(\frac{\log(\frac{K}{F_{€/\$}(t,T)}) - \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}\right).$$

9.3 VOLATILITY SMILE

One of the main assumptions in the Black-Scholes (BS) model is, that the volatility is constant over time. However, empirical studies [15, 57, 107] show that this assumption does not hold on financial markets.

The Black-Scholes model [16, 76] and its extension to FX market options (see Section 9.2) result in a closed-form formula, which describes the call or put price of an option. As we can see in Proposition 9.2.2, the option price depends on the spot exchange rate S, interest rates r, strike K, maturity T and the volatility σ. All of those parameters, except the volatility, can be observed on the market. A common market practice is to look at *implied (market) volatility*, which can be derived using current available market price data, on the same forward pair $F_{€/\$}(0,T)$ with the same parameters strike K and maturity T, which we denote as $\bar{C}_{\$}(0,T)$. We can now solve the Black-Scholes pricing formula (9.2.5), such that

$$\bar{C}_{\$}(0,T) = C_{\$}(0,T,K,F_{€/\$}(0,T),\sigma_{impl}) \tag{9.3.1}$$

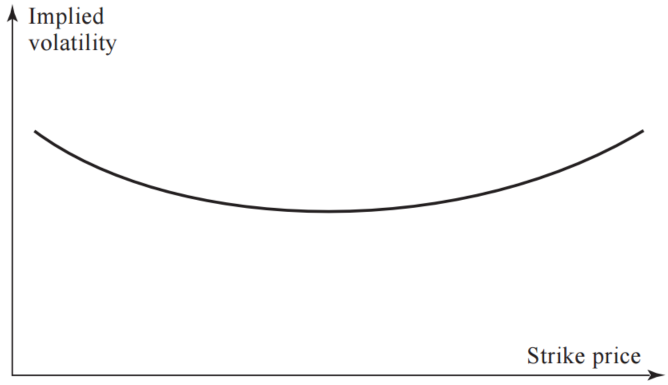

Figure 9.3.1: Typical implied Volatility graph for FX options [57, Chapter 19.2]

holds for the implied volatility σ_{impl}. However, note that the pricing formula (9.3.1) cannot be solved analytically for the implied volatility. It is hence necessary to find an approximation for σ_{impl} numerically and there exist various numerical methods, e.g. Bisection, Secant or Newton method [98, 58], to find an approximation for the solution. Instead of using implied volatility, one could look at *historical volatility*, which can be obtained analysing historical spot exchange rate data over a fixed period of time. Volatility is non-constant over time, hence it makes sense to use data over the same time interval as the time to maturity. A known disadvantage of using historical volatility is that it reflects market expectations in the past. The current market prices represent the past market information and the current market expectations, hence the implied volatility is known to be a better estimate for the future.

As empirical evidence has shown [15, 57, 107], the implied volatility for options on the FX market varies with strike (and time) and rather looks like a volatility smile, which can be seen in Figure 9.3.1. Therefore, it is possible to find the so called *volatility surface* $\sigma_{impl}(K, T)$, which reflects the dependency of the implied volatility over strike K and maturity T. The observed volatility smile is a characteristic of the FX market [15]. A reason for its appearance is that the far in-the-money (ITM) and far out-of-the-money (OTM) options are more expensive than expected under the Black-Scholes model [57]. This suggests, that the real distribution for the log spot exchange rate deviates from the assumed normal distribution. The existence of a smile suggests fatter tails and higher central peaks compared to the normal distribution, which indicates that very small and large moves in the spot exchange rate are more likely to occur. There are many different and well-known approaches to solve these short-comings:

- Use of stochastic volatility models: e.g. Heston [56, 46, 11], SABR [54].
- Use of local volatility models [37, 76].
- Use of Lévy processes as underlying stochastic process [23].

In Chapter 11, we will illustrate some of the above approaches under the intermediate currency pricing idea. Additionally, we will introduce the method of using a random variable following a extended skew normal distribution with the possibility to adjust the weights in the resulting distribution tails, which enables us to capture the volatility smile.

9.4 RISK-NEUTRAL MEASURES ON FX MARKETS AND THE FOREIGN-DOMESTIC SYMMETRY

We recall (see e.g. [15, 113, 65, 107]) that there is no measure which is simultaneously risk-neutral for both the domestic and the foreign market. Let us denote the EUR-USD spot exchange rate at time t as

$$f(t) := S_{€/\$}(t).$$

The inverse spot exchange rate can then be expressed as

$$S_{\$/€}(t) = \frac{1}{S_{€/\$}(t)} = \frac{1}{f(t)}.$$

Within the standard option pricing setting, we assume that the currency market under a 'market' measure is described by the system:

$$\begin{aligned} dB_{\$} &= r_{\$}(t)B_{\$}dt, \qquad (9.4.1)\\ dB_{€} &= r_{€}(t)B_{€}dt,\\ df &= \mu(t)fdt + \sigma(t)fdW(t), \end{aligned}$$

where $B_{\$}(t)$, $B_{€}(t)$ and $r_{\$}(t)$, $r_{€}(t)$ are USD and EUR bank accounts with their short interest rates, respectively; $\sigma(t) > 0$ is a volatility, $\mu(t)$ is a drift; and $W(t)$ is a standard Wiener process. It is assumed that the coefficients $r_{\$}(t)$, $r_{€}(t)$, $\sigma(t)$, and $\mu(t)$ are stochastic processes adapted to a filtration $\mathcal{F}_t$ to which $W(t)$ is also adapted (typically, in stochastic volatility models $\mathcal{F}_t$ is larger than the natural filtration of $W(t)$), and they have bounded second moments. We also require that $\sigma(t)$ satisfies Novikov's condition.

On the USD market, the foreign currency EUR is paid for by USD (the domestic currency) and the risky asset is

$$Y_{€/\$}(t) = S_{€/\$}(t)B_{€}(t),$$

while on the EUR market the risky asset is

$$Y_{\$/€}(t) = S_{\$/€}(t)B_{\$}(t).$$

Following the classical theory of pricing, we have to find equivalent (local) martingale measures (EMMs) $Q^{\$}$ and $Q^{€}$ under which the corresponding discounted risky assets are (local) martingales (see [65]). By standard arguments we arrive at the SDEs for $f(t)$ and $g(t) := 1/f(t)$ written under the corresponding EMMs:

$$df = (r_{\$}(t) - r_{€}(t))fdt + \sigma(t)fdW^{Q^{\$}}(t), \qquad (9.4.2)$$

$$dg = (r_{€}(t) - r_{\$}(t))gdt - \sigma(t)gdW^{Q^{€}}(t), \qquad (9.4.3)$$

where $W^{Q^{\$}}(t)$ is a standard Wiener process under $Q^{\$}$ and $W^{Q^{€}}(t)$ is a standard Wiener process under $Q^{€}$. We can see (cf. (9.4.1) and (9.4.2)-(9.4.3)) that the market prices of risk on the two markets differ:

$$\gamma_{€}(t) = \frac{\mu(t) + r_{€}(t) - r_{\$}(t)}{\sigma(t)} \neq \frac{\sigma^2(t) - \mu(t) + r_{\$}(t) - r_{€}(t)}{-\sigma(t)} = \gamma_{\$}(t)$$

(recall that $\sigma(t) > 0$). Thus,

$$Q^{\$} \neq Q^{€}, \tag{9.4.4}$$

i.e., there is no measure which is simultaneously risk-neutral for the EUR domestic market and for the USD domestic market in this rather general setting.

Note that the SDE (9.4.2) for f under the measure $Q^{€}$ takes the form

$$df = (r_{\$}(t) - r_{€}(t) + \sigma^2(t))f dt + \sigma(t) f dW^{Q^{€}}. \tag{9.4.5}$$

Intuitively, one could think that the drift for the exchange rate $g(t) = 1/f(t)$ in (9.4.3) should be the negative of the drift of $f(t)$ under the same measure, i.e. $-(r_{€}(t) - r_{\$}(t)) = r_{\$}(t) - r_{€}(t)$. However, as we can see in (9.4.5), this is not the case. This is related to the phenomenon known as Siegel's paradox [108], which is due to the convexity of the function $1/f$.

Let us also recall [49, 81, 113, 38] that under the no-arbitrage assumption (and other standard conditions like no transaction costs, etc.), there is the so-called foreign-domestic symmetry for FX options which we formulate in the following theorem. This symmetry is the key requirement for a model to be consistent for a currency pair and its inverse pair (see e.g. [33, 34, 26, 48] and references therein and also Appendix B.3).

Theorem 9.4.1. *Under the no-arbitrage assumption, there is the following relationship (called **Foreign-Domestic Symmetry**) for FX options*

$$C_{€/\$}(0, T, K) = S_{€/\$}(0)\, K\, P_{\$/€}\left(0, T, \frac{1}{K}\right), \tag{9.4.6}$$

where $C_{€/\$}(0, T, K)$ *is the call option price (in \$) at time* 0 *to buy one EUR for* $\$K$ *at time* T*;* $P_{\$/€}(0, T, 1/K)$ *is the put option price (in €) at time* 0 *to sell one USD for* $€\frac{1}{K}$ *at time* T*.*

Let us emphasise that the proof of this theorem is solely based on the no-arbitrage argument, and hence it states a fundamental property of the FX market. Suppose we take a stochastic volatility model (e.g., a popular model such as the Heston and SABR) and calibrate it using option data on the USD market. If we rewrite this model with the obtained parameters for the inverse pair €/\$, then option prices computed by this model on the EUR market would not match the data on this market and the property (9.4.6) would not be satisfied, i.e. we would get an arbitrage. Instead, if we calibrate the inverse pair model again but using option data on the EUR market, then the property (9.4.6) is obviously satisfied, but it is inconvenient that the model needs to be calibrated twice despite the fact that the two smiles are consistent with each other due to absence of arbitrage and the symmetry (9.4.6). We note in passing (see e.g. [29, 26]) that for the SABR and Heston models there are mappings between the parameters obtained for USD-EUR and the parameters of the inverted world (i.e., EUR-USD), still the parameters are different for the direct and inverted worlds, which we illustrate in Example 9.4.1.

Example 9.4.1 (Illustration of standard Heston parameters)**.** In this example, we illustrate this difference in parameters, when we calibrate a standard Heston model to option price data for the GBP-EUR and the inverse EUR-GBP pair. In Figure 9.4.1, we can see the calibrated smiles following the MATLAB code and approach of [62]. It is important to note, that in the calibration, the parameters for v_0 and κ are fixed and we only optimise the parameters (δ, θ, ρ) to match the market data. The difference in the resulting parameters between the GBP-EUR world and inverted EUR-GBP world can be clearly seen in Table 9.4.1.

Table 9.4.1: The results of standard Heston model calibration for GBP-EUR and EUR-GBP data.

parameter	GBP-EUR smile	EUR-GBP smile
v_0	0.011979	0.011979
κ	1.500	1.500
δ	0.32792	0.31406
θ	0.018072	0.016805
ρ	-0.40828	0.40912

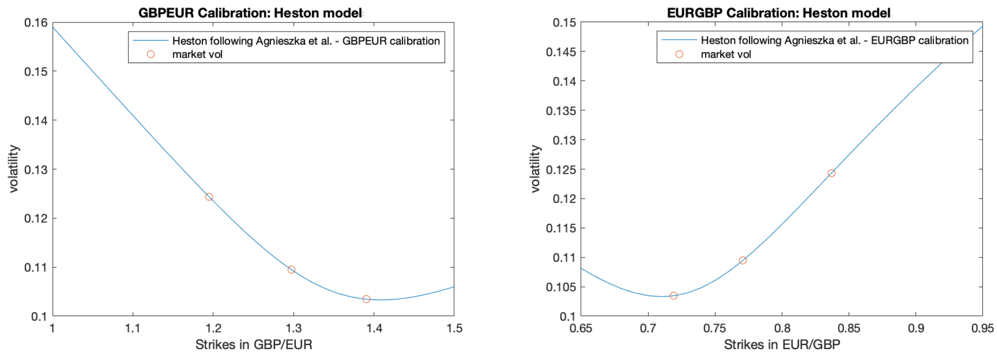

Figure 9.4.1: Illustration of GBP-EUR calibration and EUR-GBP calibration and their parameters.

In the next chapter we find a numeraire allowing to price options on USD and EUR markets simultaneously after a single calibration. In particular, within the proposed approach, calibration of a stochastic volatility model using FX data from one of the domestic markets guarantees replication of volatility smiles by the model on both domestic markets.

10

PRICING FX OPTIONS UNDER INTERMEDIATE PSEUDO-CURRENCY

In this chapter we propose a suitable candidate to be used as numeraire for which options on USD and EUR markets can be priced simultaneously under the same measure. We will illustrate in Section 10.1, that it is convenient to introduce such a numeraire using the notion of an artificial currency, which we call an intermediate pseudo-currency in this Chapter to distinguish it from the intrinsic currency of [33, 34] and the artificial currency of [26]. In Section 10.2 we extend this new idea to the case of multi-currency markets.

10.1 FX OPTION PRICING VIA INTERMEDIATE PSEUDO-CURRENCY

In this section, we start by introducing the intermediate pseudo-currency market, then (Section 10.1.1) we consider pricing under an EMM Q^X on the pseudo-market and (Section 10.1.2) – under the T-forward measure Q_T^X equivalent to Q^X. We note that the intermediate currency market is virtual and is only used as a proxy to find a suitable numeraire and write down the corresponding pricing formulas, while calibration is done using the usual FX data.

Definition 10.1.1. Let $S_{€/\$}(t) = f(t)$ be the EUR-USD exchange rate at time t. An **intermediate pseudo-currency X** is a currency with exchange rate EUR-X, $S_{€/X}(t) = \sqrt{f(t)}$, and the exchange rate USD-X, $S_{\$/X}(t) = \frac{1}{\sqrt{f(t)}}$.

We note the natural relationship for the intermediate currency

$$S_{€/X}(t) \cdot \frac{1}{S_{\$/X}(t)} = f(t). \tag{10.1.1}$$

We remark the following symmetry:

$$S_{€/X}(t) = \sqrt{f(t)} = \frac{1}{\frac{1}{\sqrt{f(t)}}} = \frac{1}{S_{\$/X}(t)} = S_{X/\$}(t)$$

and

$$S_{\$/X}(t) = \frac{1}{\sqrt{f(t)}} = \frac{1}{S_{€/X}(t)} = S_{X/€}(t).$$

We also introduce the money market account B_X for the intermediate currency X with its respective interest rate $r_X(t)$:

$$dB_X = r_X(t) B_X dt. \tag{10.1.2}$$

In the next section we first establish that for a sufficiently broad class of models for $f(t)$ there is an EMM Q^X on the pseudo-market and then, assuming existence of an EMM Q^X, we derive a pricing formula.

Remark 10.1.2. We can introduce $S_{€/X}(t) = f^{\alpha}(t)$ with any $\alpha \in (0,1)$, then $S_{X/\$}(t) = f^{\alpha-1}(t)$. Each particular α leads to the corresponding numeraire suitable for the stated purposes (but note that the numeraires associated with the original USD and EUR markets are not suitable for the set objective as discussed in Section 9.4). Arbitrariness of α can potentially be used for calibration purposes but we do not consider this aspect here. For clarity and also for the sake of symmetry, we choose to use $\alpha = 1/2$ in this paper.

Remark 10.1.3. We do not attach any economic interpretation to the intermediate pseudo-currency. Our interest is purely motivated by calibration aspects. We also note that we model a single exchange rate which is natural, not 2 rates as in [33, 35, 34, 26].

10.1.1 *An EMM for the intermediate market*

Consider the virtual market where the domestic currency is X. On this market we have two risky assets: USD paid by X and EUR paid by X:

$$Y_{€/X}(t) = S_{€/X}(t)B_{€}(t), \quad Y_{\$/X}(t) = S_{\$/X}(t)B_{\$}(t). \tag{10.1.3}$$

Assume that EUR-USD exchange rate $f(t)$ satisfies the model (9.4.1). Based on (9.4.1), we can write the SDEs under market measure for $Y_{€/X}(t)$ and $Y_{\$/X}(t)$:

$$\begin{aligned} dY_{€/X} &= \frac{1}{2}\left(\mu(t) + 2r_{€}(t) - \frac{\sigma^2(t)}{4}\right) Y_{€/X}dt + \frac{\sigma(t)}{2} Y_{€/X}dW(t), \\ dY_{\$/X} &= \frac{1}{2}\left(-\mu(t) + 2r_{\$}(t) + \frac{3\sigma^2(t)}{4}\right) Y_{\$/X}dt - \frac{\sigma(t)}{2} Y_{\$/X}dW(t). \end{aligned}$$

If we choose the intermediate currency interest rate r_X equal to

$$r_X(t) = \frac{r_{\$}(t) + r_{€}(t)}{2} + \frac{\sigma^2(t)}{8}, \tag{10.1.4}$$

then there is an EMM Q^X for the pseudo-currency market with the following market price of risk $\gamma(t)$:

$$\gamma(t) = \frac{\mu(t) - \frac{\sigma^2(t)}{2} + r_{€}(t) - r_{\$}(t)}{\sigma(t)}, \tag{10.1.5}$$

i.e.

$$\begin{aligned} dY_{€/X} &= r_X(t)Y_{€/X}dt - \frac{\sigma(t)}{2} Y_{€/X}dW^{Q^X}, \\ dY_{\$/X} &= r_X(t)Y_{\$/X}dt + \frac{\sigma(t)}{2} Y_{\$/X}dW^{Q^X}, \end{aligned}$$

where W^{Q^X} is the standard Wiener process under Q^X. So, we have shown that the intermediate pseudo-currency market can be arbitrage free within this setting. We summarise this result in the following statement.

Theorem 10.1.4. *Assume that the EUR-USD currency market under a 'market' measure is described by the model (9.4.1). Then there is the unique intermediate currency interest rate $r_X(t)$ defined in (10.1.4) and an EMM Q^X for the intermediate pseudo-currency market with the market price of risk $\gamma(t)$ from (10.1.5), i.e., under (10.1.4) the market is arbitrage-free.*

We observe from (10.1.4) that even if the short rates $r_\$(t)$ and $r_€(t)$ are assumed to be constant, the intermediate currency interest rate $r_X(t)$ is non-constant if the volatility $\sigma(t)$ is time-dependent. Especially, if $\sigma(t)$ is a stochastic process, then so is the short rate $r_X(t)$.

Example 10.1.1 (*An analogue of the Garman-Kohlhagen formula*)**.** Assume that the exchange rate between EUR and USD $f(t) = S_{€/\$}(t)$ satisfies the model (9.4.1) with constant coefficients: $\sigma(t) = \sigma$, $r_€(t) = r_€$ and $r_\$(t) = r_\$$. Note that in this simplified case (the geometric Brownian motion case) the intermediate currency interest rate r_X is constant. Analogously, to the standard derivation of the Garman-Kohlhagen formula, we can find option prices for a pseudo-currency market investor. For a European floating-strike call option (priced in X) to buy 1 EUR for $\frac{K}{\sqrt{f(T)}}$ X, we have

$$\begin{aligned}
\mathcal{E}^C_{€/X}(0,T,f(0),K,r_\$,r_€) &= e^{-r_X T}\mathbb{E}_{Q^X}\left[\left(\sqrt{f(T)} - \frac{K}{\sqrt{f(T)}}\right)_+\right] \qquad (10.1.6)\\
&= \sqrt{f(0)}e^{-r_€ T}N\left(\frac{\log\frac{f(0)}{K} + (r_\$ - r_€ + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right)\\
&\quad - \frac{K}{\sqrt{f(0)}}e^{-r_\$ T}N\left(\frac{\log\frac{f(0)}{K} + (r_\$ - r_€ - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right).
\end{aligned}$$

And, similarly for a European floating strike put option (priced in X) to sell 1 USD for $\frac{\sqrt{f(T)}}{K}$ X we have:

$$\begin{aligned}
\mathcal{E}^P_{\$/X}(0,T,\frac{1}{f(0)},\frac{1}{K},r_€,r_\$) &= e^{-r_X T}\mathbb{E}_{Q^X}\left[\left(\frac{\sqrt{f(t)}}{K} - \frac{1}{\sqrt{f(t)}}\right)_+\right] \qquad (10.1.7)\\
&= \frac{\sqrt{f(0)}}{K}e^{-r_\$ T}N\left(\frac{\log\frac{f(0)}{K} + (r_\$ - r_€ + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right)
\end{aligned}$$

$$- \frac{1}{\sqrt{f(0)}} e^{-r_€ T} N \left(\frac{\log \frac{f(0)}{K} + (r_\$ - r_€ - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} \right).$$

From (10.1.6) and (10.1.7), we can deduce prices for the call $C_{€/\$}$ and put $P_{\$/€}$. To this end, we first observe that the in-the-money payoff of the call $\mathcal{E}^C_{€/X}$ (priced in X) is equivalent to buying €1 for $\$K$. Indeed, this call's payoff is equal to the amount of X

$$\left(\sqrt{f(T)} - \frac{K}{\sqrt{f(T)}} \right)_+$$

which is equivalent to the amount of USD

$$\sqrt{f(T)} \left(\sqrt{f(T)} - \frac{K}{\sqrt{f(T)}} \right)_+ = (f(T) - K)_+$$

as we can exchange X for USD at the rate $\sqrt{f(T)}$. Analogously, the in-the-money payoff of $\mathcal{E}^P_{\$/X}$ is equivalent to selling \$1 USD for €$1/K$.

Further, by multiplying the price of $\mathcal{E}^C_{€/X}$ priced in X by $\sqrt{f(0)}$, we convert its option price in X to the price in USD, and by multiplying the price of $\mathcal{E}^P_{\$/X}$ priced in X by $1/\sqrt{f(0)}$, we convert its price to EUR. Hence

$$\begin{aligned} C_{€/\$}(0,T,f(0),K,r_\$,r_€) &= \sqrt{f(0)}\mathcal{E}^C_{€/X}(0,T,f(0),K,r_\$,r_€), \quad (10.1.8) \\ P_{\$/€}\left(0,T,\frac{1}{f(0)},\frac{1}{K},r_€,r_\$\right) &= \frac{1}{\sqrt{f(0)}}\mathcal{E}^P_{\$/X}\left(0,T,\frac{1}{f(0)},\frac{1}{K},r_€,r_\$\right). \end{aligned}$$

Comparing the resulting formulas for $C_{€/\$}$ and $P_{\$/€}$, it is not difficult to show that the foreign-domestic symmetry (9.4.6) holds.

Now let us look at a general FX option pricing formula based on the intermediate currency. Let $S_{€/\$}(t) = f(t)$ be the EUR-USD exchange rate at time t defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, Q^X)$, where Q^X is an EMM corresponding to the virtual market for which the intermediate currency X is domestic (note that at the start of this subsection we demonstrated that there is a broad class of models for which Q^X exists). Assume that the distribution of $f(t)$ is such that $f(t)$ and $1/f(t)$ have second moments. We observe that we do not assume a particular model for $f(t)$ in the pricing part of this section. For simplicity, let the interest rates for the USD and EUR money markets, $r_\$$ and $r_€$, be constant. As we noticed earlier, the intermediate currency interest rate $r_X(t)$ is, in general, not constant even when $r_\$$ and $r_€$ are constant. We assume that $r_X(t)$ is adapted to the same filtration $\mathcal{F}_t$ and

$$B_X(t) = \exp\left(\int_0^t r_X(s)ds\right). \tag{10.1.9}$$

Introduce the discounting factor $D_X(t,T)$ related to the intermediate currency interest rate and the intermediate currency non-defaultable zero-coupon bond price $P_X(t,T)$:

$$D_X(t,T) = \exp\left(-\int_t^T r_X(s)ds\right) \tag{10.1.10}$$

and

$$P_X(t,T) = \mathbb{E}_{Q^X}\left[D_X(t,T)|\mathcal{F}_t\right], \tag{10.1.11}$$

where we assumed that $D_X(t,T)$ has finite moments. Since Q^X is an EMM, the discounted $Y_{€/X}(t)$ and $Y_{\$/X}(t)$,

$$D_X(0,t)Y_{€/X}(t) = D_X(0,t)S_{€/X}(t)B_€(t) = D_X(0,t)\sqrt{f(t)}B_€(t)$$

and

$$D_X(0,t)Y_{\$/X}(t) = D_X(0,t)S_{\$/X}(t)B_\$(t) = D_X(0,t)\frac{1}{\sqrt{f(t)}}B_\$(t),$$

are Q^X-martingales. Hence we obtain for any $t \geq 0$

$$\begin{aligned} \sqrt{f(0)} &= e^{r_€ t}\mathbb{E}_{Q^X}\left[D_X(0,t)\sqrt{f(t)}\right], \\ \frac{1}{\sqrt{f(0)}} &= e^{r_\$ t}\mathbb{E}_{Q^X}\left[\frac{D_X(0,t)}{\sqrt{f(t)}}\right]. \end{aligned}$$

Thus, to obey the no-arbitrage condition, the distribution of $f(t)$, $t \geq 0$, under Q^X should be so that

$$\frac{\mathbb{E}_{Q^X}\left[D_X(0,t)\sqrt{f(t)}\right]}{\mathbb{E}_{Q^X}\left[\frac{D_X(0,t)}{\sqrt{f(t)}}\right]} = e^{(r_\$ - r_€)t} f(0). \tag{10.1.12}$$

In option pricing we will consider the following natural class of payoff functions $g(x;K)$, where $x > 0$ denotes the price of the underlier and $K \geq 0$ has the meaning of a strike.

Assumption 10.1.1. Let payoff functions $g(x;K)$ be homogeneous functions of order 1, i.e. for any $a > 0$:

$$a \cdot g(x;K) = g(ax;aK). \tag{10.1.13}$$

It is clear that e.g. plain vanilla puts and calls satisfy (10.1.13). For definiteness, assume that $g(x;K)$ is a payoff of an option written on one EUR, where x has the meaning of EUR-USD exchange rate, and K and g are denominated in USD. As in the case of a call (see Example 10.1.1), the amount of USD $g(x;K)$ is equivalent to the amount $G(x;K)$ in X:

$$G(x;K) := \frac{1}{\sqrt{x}} g(x;K) = g\left(\sqrt{x}; \frac{K}{\sqrt{x}}\right),$$

where $1/\sqrt{x}$ has the meaning of the exchange rate USD-X (cf. Definition 10.1.1) and $G(x;K)$ and $K/\sqrt{x}$ are denominated in X. According to the risk-neutral pricing theory, we can write the value of the European option $V_{€/X}(t)$ with payoff $g(\sqrt{x}; \frac{K}{\sqrt{x}})$ and maturity T at time $t \leq T$ as

$$V_{€/X}(t) = \mathbb{E}_{Q^X}\left[D_X(t,T) g\left(\sqrt{f(T)}; \frac{K}{\sqrt{f(T)}}\right) \middle| \mathcal{F}_t \right].$$

Note that this is an option on EUR priced in X. The price in dollars for this option is

$$V_{€/\$}(t) = \sqrt{f(t)}\mathbb{E}_{Q^X}\left[D_X(t,T)g\left(\sqrt{f(T)};\frac{K}{\sqrt{f(T)}}\right)\middle|\mathcal{F}_t\right]. \qquad (10.1.14)$$

Analogously, we can derive a formula for an option on USD priced in EUR:

$$V_{\$/€}(t) = \frac{1}{\sqrt{f(t)}}\mathbb{E}_{Q^X}\left[D_X(t,T)g\left(\frac{1}{\sqrt{f(T)}};\sqrt{f(T)}K\right)\middle|\mathcal{F}_t\right], \qquad (10.1.15)$$

where $g(y;K)$ is a payoff of an option written on one USD, y has the meaning of USD-EUR exchange rate, and K and g are denominated in EUR. We summarise this result in the following theorem.

Theorem 10.1.5. *Assume that the EUR-USD exchange rate $f(t)$ satisfies a model for which the no-arbitrage condition (10.1.12) holds. Then the arbitrage price of a European option on EUR with a payoff $g(x;K)$ and maturity time T is given by (10.1.14) and the arbitrage price of an option on USD is given by (10.1.15).*

It is not difficult to show that the foreign-domestic symmetry (9.4.6) holds when we use the pricing formulas (10.1.14) and (10.1.15) based on the intermediate currency.

10.1.2 *T-forward measure for the intermediate market*

Introduce the T-forward measure Q_T^X equivalent to Q^X on $\mathcal{F}_T$ with the Radon-Nikodym derivative

$$\frac{Q_T^X}{Q^X} = \frac{1}{P_X(0,T)B_X(T)} \qquad (10.1.16)$$

and for $t > 0$

$$E_Q\left[\frac{Q_T^X}{Q^X}\middle|\mathcal{F}_t\right] = \frac{P_X(t,T)}{P_X(0,T)B_X(t)}. \qquad (10.1.17)$$

Under this forward measure, we get [61, 47] (see also [15]):

$$\begin{aligned} \sqrt{f(0)} &= e^{r_€ T}\mathbb{E}_{Q^X}\left[D_X(0,T)\sqrt{f(T)}\right] \\ &= e^{r_€ T}P_X(0,T)\mathbb{E}_{Q^X_T}\left[\sqrt{f(T)}\right], \\ \frac{1}{\sqrt{f(0)}} &= e^{r_\$ T}\mathbb{E}_{Q^X}\left[\frac{D_X(0,T)}{\sqrt{f(T)}}\right] = e^{r_\$ T}P_X(0,T)\mathbb{E}_{Q^X_T}\left[\frac{1}{\sqrt{f(T)}}\right]. \end{aligned} \tag{10.1.18}$$

Then the no-arbitrage condition (10.1.12) becomes

$$\frac{\mathbb{E}_{Q^X_T}\left[\sqrt{f(T)}\right]}{\mathbb{E}_{Q^X_T}\left[\frac{1}{\sqrt{f(T)}}\right]} = e^{(r_\$ - r_€)T}f(0). \tag{10.1.19}$$

Further, (10.1.18) implies that the bond price $P_X(0,T)$ should satisfy

$$P_X(0,T) = e^{-r_€ T}\frac{\sqrt{f(0)}}{\mathbb{E}_{Q^X_T}\sqrt{f(T)}} = e^{-r_\$ T}\frac{1}{\sqrt{f(0)}\mathbb{E}_{Q^X_T}\left[\frac{1}{\sqrt{f(T)}}\right]}. \tag{10.1.20}$$

Notice that $f(0)$ is the current EUR-USD exchange rate and hence it is observable as well as $r_\$$ and $r_€$. The current forward EUR-USD exchange rate

$$F_{€/\$}(0,T) = e^{(r_\$ - r_€)T}f(0) \tag{10.1.21}$$

is also observable on the USD market.

We note that the forward EUR-X and USD-X exchange rates,

$$F_{€/X}(t,T) = e^{-r_€(T-t)}\frac{\sqrt{f(t)}}{P_X(t,T)} \text{ and } F_{\$/X}(t,T) = e^{-r_\$(T-t)}\frac{1}{P_X(t,T)\sqrt{f(t)}}, \tag{10.1.22}$$

are both Q^X_T-martingales. For convenience, we recall that if $r_X(t)$ is deterministic then the two measures Q^X and Q^X_T coincide.

It is also not difficult to show that

$$\sqrt{f(t)} = e^{r_€(T-t)}\mathbb{E}_{Q^X}\left[D_X(t,T)\sqrt{f(T)}\middle|\mathcal{F}_t\right]$$

$$
\begin{aligned}
&= e^{r_{€}(T-t)} P_X(t,T) \mathbb{E}_{Q_T^X}\left[\sqrt{f(T)} \middle| \mathcal{F}_t \right], \\
\frac{1}{\sqrt{f(t)}} &= e^{r_{\$}(T-t)} \mathbb{E}_{Q^X}\left[\frac{D_X(t,T)}{\sqrt{f(T)}} \middle| \mathcal{F}_t \right] \\
&= e^{r_{\$}(T-t)} P_X(t,T) \mathbb{E}_{Q_T^X}\left[\frac{1}{\sqrt{f(T)}} \middle| \mathcal{F}_t \right].
\end{aligned}
$$

Then

$$
\begin{aligned}
P_X(t,T) &= e^{-r_{€}(T-t)} \frac{\sqrt{f(t)}}{\mathbb{E}_{Q_T^X}\left[\sqrt{f(T)} \middle| \mathcal{F}_t \right]} \\
&= e^{-r_{\$}(T-t)} \frac{1}{\sqrt{f(t)} \mathbb{E}_{Q_T^X}\left[\frac{1}{\sqrt{f(T)}} \middle| \mathcal{F}_t \right]}.
\end{aligned}
\tag{10.1.23}
$$

The pricing formula (10.1.14) under the T-forward measure Q_T^X becomes

$$
\begin{aligned}
V_{€/\$}(t) &= \sqrt{f(t)} \mathbb{E}_{Q^X}\left[D_X(t,T) g\left(\sqrt{f(T)}; \frac{K}{\sqrt{f(T)}} \right) \middle| \mathcal{F}_t \right] \\
&= \sqrt{f(t)} P_X(t,T) \mathbb{E}_{Q_T^X}\left[g\left(\sqrt{f(T)}; \frac{K}{\sqrt{f(T)}} \right) \middle| \mathcal{F}_t \right] \\
&= \frac{e^{-r_{\$}(T-t)}}{\mathbb{E}_{Q_T^X}\left[\frac{1}{\sqrt{f(T)}} \middle| \mathcal{F}_t \right]} \mathbb{E}_{Q_T^X}\left[g\left(\sqrt{f(T)}; \frac{K}{\sqrt{f(T)}} \right) \middle| \mathcal{F}_t \right],
\end{aligned}
\tag{10.1.24}
$$

where in the last line we used (10.1.23). Analogously, we have (see (10.1.15)):

$$
V_{\$/€}(t) = \frac{e^{-r_{€}(T-t)}}{\mathbb{E}_{Q_T^X}\left[\sqrt{f(T)} \middle| \mathcal{F}_t \right]} \mathbb{E}_{Q_T^X}\left[g\left(\frac{1}{\sqrt{f(t)}}; \sqrt{f(T)} K \right) \middle| \mathcal{F}_t \right]. \tag{10.1.25}
$$

We summarize this result in the next theorem.

Theorem 10.1.6. *Assume that the EUR-USD exchange rate $f(t)$ satisfies a model for which the no-arbitrage condition (10.1.12) or (10.1.19) holds. Then the arbitrage price of an option on EUR with a payoff $g(x;K)$ and maturity time T is given by (10.1.24) and the arbitrage price of an option on USD is given by (10.1.25).*

The advantage of (10.1.24) and (10.1.25) vs (10.1.14) and (10.1.15) is that in (10.1.24) and (10.1.25) we do not need to compute the intermediate currency interest rate $r_X(t)$.

Example 10.1.2. The prices of the call for buying €1 for \$$K$ and of the put for selling \$1 for €$1/K$ are equal to

$$C_{€/\$}(0,T,K) = \frac{e^{-r_\$ T}}{\mathbb{E}_{Q_T^X}\left[\frac{1}{\sqrt{f(T)}}\right]}\mathbb{E}_{Q_T^X}\left[\left(\sqrt{f(T)} - \frac{K}{\sqrt{f(T)}}\right)_+\right], \quad (10.1.26)$$

$$P_{\$/€}\left(0,T,\frac{1}{K}\right) = \frac{e^{-r_€ T}}{\mathbb{E}_{Q_T^X}\sqrt{f(T)}}\mathbb{E}_{Q_T^X}\left[\left(\frac{\sqrt{f(T)}}{K} - \frac{1}{\sqrt{f(T)}}\right)_+\right].$$

We see that these pricing formulas satisfy the foreign-domestic symmetry (9.4.6):

$$C_{€/\$}(0,T,K) = f(0)\cdot K \cdot P_{\$/€}\left(0,T,\frac{1}{K}\right).$$

To summarise, we derived the consistent pricing formulas for FX options. Although the new pricing formulas are derived using the virtual X market, their evaluation depends on parameters of the USD and EUR markets only. When we are interested in option prices at the current time $t = 0$, they are valid for any distribution (i.e., we do not need to explicitly define the process $f(t)$) of the exchange rate $f(T)$ which satisfies (10.1.19). We will demonstrate this observation in illustrations of the new pricing formulas in Section 11.

10.2 EXTENSION TO THE MULTI–CURRENCIES CASE

Let us assume we have N currencies c_i, where $i = 1,\dots,N$. Fixing one currency, for definiteness $i = N$, we can introduce the $N-1$ exchange rates

$$f_j = S_{c_j/c_N} > 0, \quad j = 1,\dots,N-1, \quad (10.2.1)$$

which denote the exchange rates between the currency c_N to all other currencies c_i, $i = 1,\dots,N-1$.

Now we introduce the intermediate currency X by defining the N exchange rates $S_{c_i/X}$ as follows

$$S_{c_i/X} = f_1^{b_{i1}} \times f_2^{b_{i2}} \times \cdots \times f_{N-1}^{b_{iN-1}}, \quad i = 1, \ldots, N, \tag{10.2.2}$$

where $b_{ij} \in \mathbb{R}$ are so that

$$\begin{aligned} b_{jj} &= 1 - \alpha_j, \ j = 1, \ldots, N-1, \\ b_{ij} &= -\alpha_j, \ i \neq j, \ \ i = 1, \ldots, N \ \ j = 1, \ldots, N-1. \end{aligned}$$

By symmetry arguments (see also Remark 10.2.2 below), we choose

$$\alpha_i = \frac{1}{N}, \ \ i = 1, \ldots, N-1. \tag{10.2.3}$$

Note that $S_{c_i/X}$ is the exchange rate between the observable currency c_i and the introduced intermediate currency X and hence it is the worth of 1 unit of currency c_i in the intermediate currency X. In the case (10.2.1)-(10.2.2), (10.2.3), the exchange rate $S_{c_i/X}$ can be written in the concise form via geometric mean $GM(f_j)$ of the sequence of f_j :

$$S_{c_i/X} = f_i \left[\left(\prod_{j=1}^{N-1} \frac{1}{f_j} \right)^{1/(N-1)} \right]^{(N-1)/N} := f_i \left[GM(f_j) \right]^{(N-1)/N}.$$

We assume that the currency market under a 'market' measure P is described by the system:

$$\begin{aligned} df_j &= \mu_j(t) f_j dt + \sigma_j(t) f_j d\widetilde{W}_j, \quad j = 1, \ldots, N-1, \\ d\widetilde{W}_l d\widetilde{W}_k &= d\widetilde{W}_k d\widetilde{W}_l = \rho_{lk}(t) dt, \quad l, k = 1, \ldots, N-1, \end{aligned} \tag{10.2.4}$$

and

$$dB_i = r_i(t) B_i dt, \quad i = 1, \ldots, N, \tag{10.2.5}$$

where $B_i(t)$ describes the bank account of currency c_i with its short rate $r_i(t)$; $\sigma_j(t) > 0$ is the volatility of the exchange rate $f_j(t)$, $\mu_j(t)$ is its drift; and $\widetilde{W}(t) = (\widetilde{W}_1(t), \ldots, \widetilde{W}_{N-1}(t))^T$ is an $N-1$–dimensional correlated Wiener process with the correlation matrix $R(t) \in \mathbb{R}^{N-1\times N-1}$ which components we denote by $\rho_{ij}(t)$ (obviously $\rho_{ii} = 1$). It is assumed that $r_i(t)$, $\sigma_j(t)$, $\mu_j(t)$ are stochastic processes adapted to a filtration $\mathcal{F}_t$ to which $\widetilde{W}(t)$ is also adapted, and they have bounded second moments and $\sigma_j(t)$ satisfy Novikov's condition. Furthermore, let us assume that the matrix R is symmetric strictly positive definite. Then using the Cholesky decomposition, we can represent $R = LL^T$, where $L \in \mathbb{R}^{N-1\times N-1}$ is a lower triangular matrix with entries $L_{i,j}$. Using this decomposition, we can rewrite the SDEs (10.2.4) as

$$df_j = \mu_j(t) f_j dt + \sigma_j(t) f_j \sum_{k=1}^{j} L_{jk}(t) dW_k, \quad j = 1, \ldots, N-1, \tag{10.2.6}$$

where

$$L_{ii}(t) = \sqrt{1 - \sum_{k=1}^{i-1} L_{ik}^2(t)}, \quad L_{ji}(t) = \frac{\rho_{ij} - \sum_{k=1}^{i-1} L_{jk}(t) L_{ik}(t)}{L_{ii}(t)}, \quad \text{for } i < j,$$

and $W(t) = (W_1(t), \ldots, W_{N-1}(t))^T$ is an $N-1$–dimensional standard Wiener process. We first show that the intermediate currency introduced in (10.2.2) permits an arbitrage-free market involving all N currencies.

Theorem 10.2.1. *Assume that $N-1$ exchange rates f_j between the currency c_N to all other currencies c_i, $i = 1, \ldots, N-1$, under a 'market' measure are described by the model (10.2.6) together with (10.2.5). Consider the intermediate currency X introduced in (10.2.2). There is the unique intermediate currency interest rate $r_X(t)$ defined by*

$$r_X(t) = \frac{1}{N} \sum_{i=1}^{N} r_i(t) + \frac{1}{2N}\left(1 - \frac{1}{N}\right) \sum_{i=1}^{N-1} \sigma_i^2(t) - \frac{1}{N^2} \sum_{j=1}^{N-1} \sum_{k=1}^{j-1} \sigma_j(t) \sigma_k(t) \rho_{jk}(t) \tag{10.2.7}$$

and there is an EMM Q^X for the intermediate pseudo-currency market, i.e., under (10.2.7) this market is arbitrage-free.

Proof. Applying the Ito formula to (10.2.2), we obtain the SDEs for the exchange rates $S_{c_i/X}$:

$$\begin{aligned}\frac{dS_{c_i/X}}{S_{c_i/X}} &= \Bigg[\frac{1}{N}\sum_{j=1}^{N-1}\left(-\mu_j + \frac{\frac{1}{N}+1}{2}\sigma_j^2\right) + \mu_i \mathbb{1}_{i\neq N} \\ &\quad + \frac{1}{N^2}\sum_{j=1}^{N-1}\sum_{k=1}^{j-1}\sigma_j\sigma_k\rho_{jk} - \frac{1}{N}\sigma_i\sum_{k=1}^{N-1}\mathbb{1}_{i\neq N}\sigma_k\rho_{ik}\Bigg]dt \\ &\quad - \frac{1}{N}\sum_{j=1}^{N-1}\sum_{k=1}^{j}\sigma_j L_{jk}dW_k + \sigma_i\sum_{k=1}^{i}\mathbb{1}_{i\neq N}L_{ik}dW_k \\ &= \Bigg[\frac{1}{N}\sum_{j=1}^{N-1}\left(\frac{1}{2}\left(\frac{1}{N}+1\right)\sigma_j^2 - \sigma_i\mathbb{1}_{i\neq N}\sigma_j\rho_{ij} + \frac{1}{N}\sigma_j\sum_{k=1}^{j-1}\sigma_k\rho_{kj} - \mu_j\right) \\ &\quad + \mu_i\mathbb{1}_{i\neq N}\Bigg]dt \\ &\quad - \frac{1}{N}\sum_{j=1}^{N-1}\sum_{k=1}^{j}\sigma_j L_{jk}dW_k + \sigma_i\mathbb{1}_{i\neq N}\sum_{k=1}^{i}L_{ij}dW_k, \quad i = 1,\dots,N.\end{aligned}$$

On the considered market the risky assets have the prices $Y_{c_i/X} = S_{c_i/X}B_i$, $i = 1,\dots,N$. Introduce the discounted risky assets' prices in the usual way:

$$\widetilde{Y}_{c_i/X}(t) = \frac{S_{c_i/X}(t)B_i(t)}{B_X(t)}, \quad i = 1,\dots,N. \tag{10.2.8}$$

The discounted prices satisfy the SDEs

$$\begin{aligned}\frac{d\widetilde{Y}_{c_i/X}}{\widetilde{Y}_{c_i/X}} &= [r_i - r_X]\,dt \\ &\quad + \Bigg[\frac{1}{N}\sum_{j=1}^{N-1}\left(\frac{1}{2}\left(\frac{1}{N}+1\right)\sigma_j^2 - \sigma_i\mathbb{1}_{i\neq N}\sigma_j\rho_{ij} + \frac{1}{N}\sigma_j\sum_{k=1}^{j-1}\sigma_k\rho_{kj} - \mu_j\right) \\ &\quad + \mu_i\mathbb{1}_{i\neq N}\Bigg]dt \\ &\quad - \frac{1}{N}\sum_{j=1}^{N-1}\sum_{k=1}^{j}\sigma_j L_{jk}dW_k + \sigma_i\mathbb{1}_{i\neq N}\sum_{k=1}^{i}L_{ik}dW_k, \quad i = 1,\dots,N.\end{aligned}$$

The no-arbitrage condition requires existence of an EMM Q^X under which all $\widetilde{Y}_{c_i/X}$ are martingales. This implies that for Q^X to exist the following system of N simultaneous linear algebraic equations in N unknown variables (which are the market prices of risk γ_k, $k = 1, \ldots, N-1$, and r_X) should have a solution:

$$r_i - r_X + \frac{1}{N}\sum_{j=1}^{N-1}\left(\frac{1}{2}\left(\frac{1}{N}+1\right)\sigma_j^2 - \sigma_i \mathbb{1}_{i\neq N}\sigma_j\rho_{ij} + \frac{1}{N}\sigma_j\sum_{k=1}^{j-1}\sigma_k\rho_{kj} - \mu_j\right) \tag{10.2.9}$$

$$+ \mu_i \mathbb{1}_{i\neq N}$$
$$= -\frac{1}{N}\sum_{j=1}^{N-1}\sum_{k=1}^{j}\sigma_j L_{jk}\gamma_k + \sigma_i \mathbb{1}_{i\neq N}\sum_{k=1}^{i} L_{ik}\gamma_k, \quad i = 1, \ldots, N.$$

Subtracting the equation (10.2.9) with $i = N$ from the equations (10.2.9) for $i \neq N$, we obtain

$$r_i - r_N + \mu_i - \frac{1}{N}\sigma_i\sum_{k=1}^{N-1}\sigma_k\rho_{ik} = \sigma_i\sum_{k=1}^{i} L_{ik}\gamma_k, \quad i = 1, \ldots, N-1. \tag{10.2.10}$$

Using (10.2.10), we recurrently find the market prices of risk:

$$\gamma_i = \frac{r_i - r_N + \mu_i - \frac{1}{N}\sigma_i\sum\limits_{k=1}^{N-1}\sigma_k\rho_{ik} - \sigma_i\sum\limits_{k=1}^{i-1} L_{i,k}\gamma_k}{\sigma_i L_{i,i}}, \quad i = 1, \ldots, N-1, \tag{10.2.11}$$

which are well defined due to our assumptions $\sigma_i > 0$ and $L_{i,i} > 0$. Moreover, sum up (10.2.10) over i from $i = 1$ to $N-1$ and substitute the result in (10.2.9) with $i = N$ to confirm (10.2.7):

$$r_N - r_X + \frac{1}{N}\sum_{j=1}^{N-1}\left(\frac{1}{2}\left(\frac{1}{N}+1\right)\sigma_j^2 + \frac{1}{N}\sigma_j\sum_{k=1}^{j-1}\sigma_k\rho_{kj} - \mu_j\right)$$
$$= -\frac{1}{N}\sum_{j=1}^{N-1}\left(r_j - r_N + \mu_j - \frac{1}{N}\sigma_j\sum_{k=1}^{N-1}\sigma_k\rho_{jk}\right)$$
$$\Leftrightarrow \quad r_X = \frac{1}{N}\sum_{j=1}^{N} r_j + \frac{1}{2N}\left(1-\frac{1}{N}\right)\sum_{j=1}^{N-1}\sigma_j^2 - \frac{1}{N^2}\sum_{j=1}^{N-1}\sum_{k=1}^{j-1}\sigma_j\sigma_k\rho_{jk}.$$

The found γ_i, $i = 1, \ldots, N-1$, from (10.2.11) and r_X from (10.2.7) together with Girsanov's theorem ensure that there is an EMM Q^X under which all $\widetilde{Y}_{c_i/X}$ are martingales. Thus, the considered market is arbitrage free. Theorem 10.2.1 is proved.

Remark 10.2.2. Recall that we chose to use $\alpha_1 = \cdots = \alpha_{N-1} = \frac{1}{N}$ in (10.2.2). If we repeat the proof of Theorem 10.2.1 for arbitrary $0 < \alpha_j < 1$ (see Appendix B.4) then we arrive at the following intermediate currency interest rate r_X :

$$r_X = \left(1 - \sum_{j=1}^{N-1} \alpha_j\right) r_N + \sum_{j=1}^{N-1} \alpha_j r_j + \sum_{j=1}^{N-1} \frac{\alpha_j(1-\alpha_j)}{2}\sigma_j^2 - \sum_{j=1}^{N-1}\sum_{k=1}^{j-1} \alpha_j \alpha_k \sigma_j \sigma_k \rho_{jk}, \tag{10.2.12}$$

ensuring that there is an EMM in this market. We see that the choice $\alpha_j = \frac{1}{N}$ results in the symmetry so that each r_j enters (10.2.12) with the same weight. Other choices of α_j give a 'preference' to a particular currency.

Analogously to Assumption 10.1.1, we will consider payoffs as first-order homogeneous functions in the multi-currencies case.

Assumption 10.2.1. Let payoff functions $g(x_1, \ldots, x_{N-1}; K)$ be homogeneous functions of order 1, i.e. for any $a > 0$

$$a \cdot g(x_1, \ldots, x_{N-1}; K) = g(ax_1, \ldots, ax_{N-1}; aK). \tag{10.2.13}$$

Most multi-currency options (e.g. basket options [22]) have pay-offs belonging to this class. Consider a European-type option with maturity time T and payoff in the currency c_N:

$$g(T) := g(f_1(T), \ldots, f_{N-1}(T); K).$$

Its equivalent value in the intermediate currency X is equal to (see (10.2.2)):

$$G(T) := S_{c_N/X}(T) \cdot g(T)$$

$$
\begin{aligned}
&= g\Big(f_1(T)S_{c_N/X}(T), \dots, f_{N-1}(T)S_{c_N/X}(T); K\cdot S_{c_N/X}(T)\Big) \qquad (10.2.14)\\
&= g\Big(S_{c_1/X}(T), \dots, S_{c_{N-1}/X}(T); K\cdot S_{c_N/X}(T)\Big)\\
&= g\Big(S_{c_1/X}(T), \dots, S_{c_N/X}(T); K'\Big),
\end{aligned}
$$

where $K' = K \cdot S_{c_N/X}(t)$ is the equivalent strike in X. It is not difficult to see that at the maturity time T the option holder is indifferent between receiving $g(T)$ in currency c_N or $G(T)$ in currency X as they can obtain the same amount by exchanging $G(T)$ to c_N:

$$
\begin{aligned}
\frac{G(T)}{S_{c_N/X}(T)} &= \frac{1}{S_{c_N/X}(T)} g\Big(S_{c_1/X}(T), \dots, S_{c_{N-1}/X}(T); K\cdot S_{c_N/X}(T)\Big)\\
&= g\Big(f_1(T), \dots, f_{N-1}(T); K\Big).
\end{aligned}
$$

Example 10.2.1 (Basket option)**.** Consider a basket option on the c_N market written on all $N-1$ exchange rates $f_i(t)$, $i = 1, \dots, N-1$, which has the pay-off function of the form [22]:

$$
g(x_1, \dots, x_{N-1}; K) = \left(\sum_{i=1}^{N-1} \omega_i x_i - K\right)_+,
$$

where x_i, $i = 1, \dots, N-1$, and K are denominated in the currency c_N and $\omega_i \geq 0$, $i = 1, \dots, N-1$, are some weights. The equivalent pay-off on the X currency market at the maturity T is equal to

$$
\begin{aligned}
G(T) &= S_{c_N/X}(T)\cdot g(f_1(T), \dots, f_{N-1}(T); K) \qquad (10.2.15)\\
&= S_{c_N/X}(T)\left(\sum_{i=1}^{N-1} \omega_i f_i(T) - K\right)_+ = \left(\sum_{i=1}^{N-1} \omega_i S_{c_i/X}(t) - K\cdot S_{c_N/X}(t)\right)_+\\
&= \left(\sum_{i=1}^{N-1} \omega_i S_{c_i/X}(t) - K'\right)_+,
\end{aligned}
$$

where $S_{c_i/X}(t)$ and K' are denominated in the intermediate currency X.

As in the case of a single FX pair (see Theorem 10.1.4), we have shown by Theorem 10.2.1 that there is a sufficiently broad class of models for which there is an EMM Q^X with an appropriate choice of the intermediate currency interest rate $r_X(t)$. We now generalize the pricing formulas of Theorems 10.1.5 and 10.1.6 from a single FX pair to the multi-currency case.

Let the exchange rates $f_i(t)$ between the currency c_N to all other currencies c_i, $i = 1, \dots, N-1$, be defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, Q^X)$, where Q^X is an EMM corresponding to the virtual market for which the intermediate currency X is domestic. Assume that $f_i(t)$, $i = 1, \dots, N-1$, and the exchange rates $S_{c_i/X}$ between the pseudo-currency X to all the currencies c_i, $i = 1, \dots, N$, defined in (10.2.2), (10.2.3) have second moments. Further, assume that $r_X(t)$ is adapted to the same filtration $\mathcal{F}_t$ and recall the expressions and assumptions for the money market account $B_X(t)$ (see (10.1.9)), the discounting factor $D_X(t,T)$ related to the intermediate currency interest rate (see (10.1.10)) and the intermediate currency zero-coupon bond price $P_X(t,T)$ (see (10.1.11)).

As Q^X is an EMM, the discounted $Y_{c_i/X}(t)$ for all $i = 1, \dots, N$,

$$\widetilde{Y}_{c_i/X} = D_X(0,t) Y_{c_i/X}(t) = D_X(0,t) S_{c_i/X}(t) B_{c_i}(t), \quad i = 1, \dots, N,$$

are Q^X-martingales. Hence we obtain

$$S_{c_i/X}(0) = e^{r_i t} \mathbb{E}_{Q^X}\left[D_X(0,t) S_{c_i/X}(t)\right], \quad i = 1, \dots, N.$$

Therefore, for all $i = 1, \dots, N-1$ and $t > 0$, we have

$$\frac{\mathbb{E}_{Q^X}\left[D_X(0,t) S_{c_i/X}(t)\right]}{\mathbb{E}_{Q^X}\left[D_X(0,t) S_{c_N/X}(t)\right]} = e^{(r_i - r_N)t} \frac{S_{c_i/X}(0)}{S_{c_N/X}(0)} = e^{(r_N - r_i)t} f_i(0). \tag{10.2.16}$$

Hence, to obey the no-arbitrage condition, the distributions of $S_{c_i/X}(t)$, $t > 0$, under Q^X should be so that (10.2.16) holds.

Consider a European option with maturity T and pay-off function $G(T)$ on the intermediate currency market. Its price in X is equal to

$$V_X(t) = \mathbb{E}_{Q^X}\left[D_X(0,T) G(T) | \mathcal{F}_t\right]. \tag{10.2.17}$$

Using (10.2.14), we obtain the price for this option in the currency c_N:

$$V_{c_N}(t) = \frac{1}{S_{c_N/X}(t)} \mathbb{E}_{Q^X}\left[D_X(0,T)\cdot g(S_{c_1/X}(T),\dots,S_{c_{N-1}/X}(T);KS_{c_N/X}(T))|\mathcal{F}_t\right]. \tag{10.2.18}$$

Then the analog of Theorem 10.1.5 is as follows.

Theorem 10.2.3. *Assume that the exchange rates $f_i(t)$, $i = 1,\dots,N-1$, (or $S_{c_i/X}(t), i = 1,\dots,N$) satisfy a model for which the no-arbitrage condition (10.2.16) holds. Then on the c_N market the arbitrage price $V_{c_N}(t)$ of a European option on $c_1, \dots, c_{N-1}$ currencies with a payoff $g(x_1,\dots,x_{N-1};K)$ and maturity time T is given by (10.2.18).*

We want to change measure again now, so therefore introduce the T-forward measure Q^X_T equivalent to Q^X on $\mathcal{F}_T$ with the Radon-Nikodym derivative as in (10.1.16) (see also (10.1.17)). Under this forward measure, we get for $i = 1,\dots,N$,

$$S_{c_i/X}(0) = e^{r_iT}\mathbb{E}_{Q^X}\left[D_X(0,T)S_{c_i/X}(T)\right] = e^{r_iT}P_X(0,T)\mathbb{E}_{Q^X_T}\left[S_{c_i/X}(T)\right]. \tag{10.2.19}$$

Then the no-arbitrage conditions (10.2.16) become

$$\frac{\mathbb{E}_{Q^X_T}\left[S_{c_i/X}(T)\right]}{\mathbb{E}_{Q^X_T}\left[S_{c_N/X}(T)\right]} = e^{(r_N-r_i)T}f_i(0), \quad i=1,\dots,N-1. \tag{10.2.20}$$

Here $F_{c_i/c_N}(0) = e^{(r_N-r_i)T}f_i(0)$ is the current forward c_i-c_N exchange rate. It follows from (10.2.20) that for any $j = 1,\dots,N$:

$$\frac{\mathbb{E}_{Q^X_T}\left[S_{c_i/X}(T)\right]}{\mathbb{E}_{Q^X_T}\left[S_{c_j/X}(T)\right]} = e^{(r_j-r_i)T}\frac{f_i(0)}{f_j(0)} = e^{(r_j-r_i)T}S_{c_i/c_j}(0), \quad i=1,\dots,N,\ i\neq j. \tag{10.2.21}$$

We remark that the no-arbitrage condition does not depend on the choice of c_N used in (10.2.1).

Moreover, (10.2.19) implies that the bond price $P_X(0,T)$ should satisfy

$$P_X(0,T) = e^{-r_i T} \frac{S_{c_i/X}(0)}{\mathbb{E}_{Q_T^X}\left[S_{c_i/X}(T)\right]}, \quad i = 1, \dots, N. \tag{10.2.22}$$

We observe that the relationships (10.2.20) ensure that (10.2.22) holds for all $i = 1, \dots, N$. Note that $f_i(0)$ are the current c_i/c_N exchange rates and hence $S_{c_i/X}(0)$ (see (10.2.2)) are observable as well as all r_i. Similarly to (10.1.23), we also have

$$P_X(t,T) = e^{-r_i(T-t)} \frac{S_{c_i/X}(t)}{\mathbb{E}_{Q_T^X}\left[S_{c_i/X}(T)|\mathcal{F}_t\right]}, \quad i = 1, \dots, N.$$

Analogously to (10.1.25), the pricing formula (10.2.18) under the T-forward measure Q_T^X becomes

$$\begin{aligned} V_{c_N}(t) = {} & \frac{e^{-r_N(T-t)}}{\mathbb{E}_{Q_T^X}\left[S_{c_N/X}(T)|\mathcal{F}_t\right]} \\ & \times \mathbb{E}_{Q_T^X}\left[g(S_{c_1/X}(T), \dots, S_{c_{N-1}/X}(T); K S_{c_N/X}(T))|\mathcal{F}_t\right]. \end{aligned} \tag{10.2.23}$$

Then the analog of Theorem 10.1.6 is as follows.

Theorem 10.2.4. *Assume that the exchange rates $f_i(t)$, $i = 1, \dots, N-1$, (or $S_{c_i/X}(t), i = 1, \dots, N$) satisfy a model for which the no-arbitrage condition (10.2.20) (or (10.2.16)) holds. Then on the c_N market the arbitrage price $V_{c_N}(t)$ of a European option on $c_1, \dots, c_{N-1}$ currencies with a payoff $g(x_1, \dots, x_{N-1}; K)$ and maturity time T is given by (10.2.23).*

It is clear that the pricing formula (10.2.23) remains true if we replace the currency c_N with any other c_j and the scalable payoff $g(x_1, \dots, x_{j-1}, x_j \dots, x_N; K)$ is denominated in c_j. We now return to Example 10.2.1.

Example 10.2.2 (Basket option pricing)**.** Let us make the same assumptions as in Example 10.2.1 and find the arbitrage price of a European option with pay-off $G(T)$ at maturity T on the X currency market given by

$$G(t) = \left(\sum_{i=1}^{N-1} \omega_i S_{c_i/X}(t) - K \right)_+ ,$$

where $S_{c_i/X}(t)$ and K are denominated in X. Following Theorem 10.2.4, the price of this basket option at time 0, denominated in currency c_N, is equal to

$$BasketOption_{c_N}(0) = \frac{e^{-r_N T}}{\mathbb{E}_{Q_T^X}\left[S_{c_N/X}(T)\right]} \mathbb{E}_{Q_T^X}\left[\left(\sum_{i=1}^{N-1} \omega_i S_{c_i/X}(T) - K \right)_+ \right].$$

To summarise, we derived consistent pricing formulas (10.2.18) and (10.2.23) for FX options in the multi-currency case. As it was in Chapter 10.1 for a single FX pair, here in the multi-currency case, although the pricing formulas (10.2.18) and (10.2.23) are derived using the virtual X market, their evaluation depends on parameters of the real c_i, $i = 1, \ldots, N$, markets only. The distinguishing feature of our approach in comparison with the others is that we can price all FX options regardless from their domestic market using the same measure which in turn guarantees that all natural relationships between exchange rates and FX options are automatically fulfilled.

11

ILLUSTRATIONS

For illustrative purposes, we consider four examples in this chapter. The first example (Section 11.1) illustrates the use of FX pricing from Chapter 10 in the case when the EUR-USD exchange rate $f(t)$ is described by the Heston model [56] whereas the second example (Section 11.2) deals with the SABR model [54]. In these two examples we follow the traditional route: we start with models written under a 'market' measure, then find an EMM Q^X on the intermediate currency market and use Theorem 10.1.5 for pricing FX options. The third example presented in Section 11.3 follows a different route: we propose a distribution for an exchange rate at maturity time T, e.g. for EUR-USD, under a forward measure Q_T^X on the intermediate currency market so that the no-arbitrage condition (10.1.19) is satisfied. Then, we use Theorem 10.1.6 or Theorem 10.2.4 for pricing FX options. To this end, in Section 11.3 we assume that the EUR-USD exchange rate $f(T)$ has a skew normal distribution. We remark that the use of the considered extended skew normal model for FX pricing is novel. In Section 11.4, we illustrate the results of Sections 10.1 and 10.2 in the case of the model-free approach.

11.1 HESTON MODEL

For simplicity, let the interest rates for the USD and EUR money markets, $r_\$$ and $r_€$, be constant. Consider the Heston stochastic volatility model for the

EUR-USD exchange rate $S_{€/\$}(t) = f(t)$ written under a 'market' measure [56]:

$$\begin{aligned} df &= \mu f\, dt + \sqrt{v} f \left(\sqrt{1-\rho^2}\, dW_1(t) + \rho dW_2(t) \right), \qquad (11.1.1) \\ dv &= \kappa\,(\theta - v) dt + \delta \sqrt{v} dW_2(t), \\ & f(0) = f_0, \quad v(0) = v_0, \qquad (11.1.2) \end{aligned}$$

where $W_1(t)$ and $W_2(t)$ are independent standard Wiener processes; $\sigma(t) = \sqrt{v(t)}$ is a (stochastic) volatility; θ, κ, δ, f_0 and v_0 are positive constants, satisfying

$$2\kappa\theta \geq \delta^2; \qquad (11.1.3)$$

and the correlation coefficient $\rho \in (-1, 1)$. Recall [63] that the condition (11.1.3) guarantees that zero is unattainable by $v(t)$ in finite time.

Following Section 10.1.1, to re-write (11.1.1) under Q^X, we need to find the market prices of risk, $\gamma_1(t)$ and $\gamma_2(t)$, so that (cf. (10.1.5)):

$$\sqrt{1-\rho^2}\gamma_1(t) + \rho\gamma_2(t) = \frac{\mu - v(t)/2 + r_{€} - r_{\$}}{\sqrt{v(t)}}. \qquad (11.1.4)$$

As it is standard for the Heston model [56], to deal with incompleteness of the market, we choose

$$\gamma_2(t) = \lambda\sqrt{v(t)}, \qquad (11.1.5)$$

where λ is a constant. Thus, we have

$$\begin{aligned} d\sqrt{f} &= (r_X(t) - r_{€})\sqrt{f} dt \\ &\quad + \frac{\sqrt{v}}{2}\sqrt{f}\left(\sqrt{1-\rho^2}\, dW_1^{Q^X}(t) + \rho\, dW_2^{Q^X}(t)\right), \qquad (11.1.6) \\ d\frac{1}{\sqrt{f}} &= (r_X(t) - r_{\$})\frac{1}{\sqrt{f}} dt - \frac{\sqrt{v}}{2}\frac{1}{\sqrt{f}}\left(\sqrt{1-\rho^2}\, dW_1^{Q^X}(t) + \rho\, dW_2^{Q^X}(t)\right), \\ dv &= \kappa\,(\theta - v) dt + \delta\sqrt{v} dW_2^{Q^X}, \; v(0) = v_0, \end{aligned}$$

where, as before (see (10.1.4)),

$$r_X(t) = \frac{r_\$ + r_€}{2} + \frac{v(t)}{8} \tag{11.1.7}$$

and, without changing the notation, the new κ and θ in (11.1.6) are equal to $\kappa + \lambda\delta$ and $\kappa\theta/(\kappa + \lambda\delta)$, respectively, in terms of the old κ and θ from (11.1.1). Then (see Theorem 10.1.5), e.g. the price of the call (in USD) for buying €1 for $K is equal to

$$C_{€/\$}(0,T,K) = \sqrt{f(0)}\mathbb{E}_{Q^X}\left[D_X(0,T)\left(\sqrt{f(T)} - \frac{K}{\sqrt{f(T)}}\right)_+\right], \tag{11.1.8}$$

where $\sqrt{f(T)}$ and $1/\sqrt{f(T)}$ are from (11.1.6).

Now, we can rewrite (11.1.6) under the T-forward measure Q^X_T using the results of Section 10.1.2. By (11.1.7), we have

$$\begin{aligned} P_X(t,T) &= \mathbb{E}_{Q^X}\left[D_X(t,T)|\mathcal{F}_t\right] \\ &= \mathbb{E}_{Q^X}\left[\exp\left(-\int_t^T r_X(s)ds\right)\middle|\mathcal{F}_t\right] \\ &= \exp\left(-\frac{r_\$ + r_€}{2}(T-t)\right)\mathbb{E}_{Q^X}\left[\exp\left(-\int_t^T \frac{v(s)}{8}ds\right)\middle|v(t)\right]. \end{aligned}$$

The stochastic X short rate $r_X(t)$ defined by (11.1.7) with $v(t)$ from (11.1.6) possesses an affine term structure (see e.g. [15, 17, 36]):

$$P_X(t,T) = \exp\left(-\frac{r_\$ + r_€}{2}(T-t) + A(T-t) - C(T-t)v(t)\right). \tag{11.1.9}$$

The PDE problem for $P_X(t,T) = p(t,v)$ can be written by applying the Feymann-Kac formula:

$$\begin{aligned} &\frac{\partial}{\partial t}p(t,v) + \kappa(\theta - v)\frac{\partial}{\partial v}p(t,v) + \frac{\delta^2 v}{2}\frac{\partial^2}{\partial v^2}p(t,v) - \frac{v}{8}p(t,v) \\ &\quad - \frac{r_\$ + r_€}{2}p(t,v) = 0, \end{aligned} \tag{11.1.10}$$

where $p(T,v) = 1$. Now let us assume that the solution to 11.1.10 can be written as given in 11.1.9. We can derive the partial derivatives of $p(t,v)$:

$$\begin{aligned}
\frac{\partial}{\partial t}p(t,v) &= \left(\frac{r_{\$}+r_{€}}{2} - \frac{\partial}{\partial t}A(T-t) + \frac{\partial}{\partial t}C(T-t)v(t)\right)p(t,v),\\
\frac{\partial}{\partial v}p(t,v) &= -C(T-t)p(t,v),\\
\frac{\partial^2}{\partial v^2}p(t,v) &= [C(T-t)]^2\,p(t,v).
\end{aligned}$$

This leaves us with the following PDE:

$$\begin{aligned}
&\left(\frac{\partial}{\partial t}A(T-t) + \frac{\partial}{\partial t}C(T-t)v\right)p(t,v) + \kappa(\theta - v)\left(-C(T-t)\right)p(t,v)\\
&\quad + \frac{\delta^2 v}{2}\left[C(T-t)\right]^2 p(t,v) - \frac{v}{8}p(t,v) = 0\\
\Leftrightarrow\quad & p(t,v)\left[-\frac{\partial}{\partial t}A(T-t) - \kappa\theta C(T-t)\right]\\
&\quad + p(t,v)v(t)\left[\frac{\partial}{\partial t}C(T-t) + \frac{\delta^2}{2}\left[C(T-t)\right]^2 + \kappa C(T-t) - \frac{1}{8}\right] = 0,
\end{aligned}$$

which should be true for any $v(t)$. Therefore,

$$\begin{aligned}
\frac{\partial}{\partial t}A(T-t) &= -\kappa\theta C(T-t), \quad A(0) = 0,\\
\frac{\partial}{\partial t}C(T-t) &= -\frac{\delta^2 v}{2}[C(T-t)]^2 - \kappa C(T-t) + \frac{1}{8}, \quad C(0) = 0.
\end{aligned}$$

We can rewrite the equation by replacing $C(T-t)$ by $C(t)$. Moreover, this equation for $C(t)$ is a Riccati equation, therefore let us do the following substitution:

$$C(t) = \frac{2\frac{\partial}{\partial t}u(t)}{\delta^2 u(t)}, \qquad \frac{\partial}{\partial t}C(t) = \frac{2\frac{\partial^2}{\partial t^2}u(t)}{\delta^2 u(t)} - \frac{2\delta^2\left(\frac{\partial}{\partial t}u(t)\right)^2}{[\delta^2 u(t)]^2}. \tag{11.1.11}$$

This will leave us with the following equation:

$$\frac{2\frac{\partial^2}{\partial t^2}u(t)}{\delta^2 u(t)} - \frac{2\delta^2\left(\frac{\partial}{\partial t}u(t)\right)^2}{[\delta^2 u(t)]^2} = -\frac{\delta^2}{2}\left[\frac{2\frac{\partial}{\partial t}u(t)}{\delta^2 u(t)}\right]^2 - \kappa\left[\frac{2\frac{\partial}{\partial t}u(t)}{\delta^2 u(t)}\right] + \frac{1}{8}$$

$$\Leftrightarrow \quad 16\frac{\partial^2}{\partial t^2}u(t) + 16\kappa\frac{\partial}{\partial t}u(t) - \delta^2 u(t) = 0.$$

This second order ODE has the general solution $u(t)$ and derivative $\frac{\partial}{\partial t}u(t)$ as follows:

$$\begin{aligned} u(t) &= B_1 \exp\left(\frac{-\kappa+\beta}{2}t\right) + B_2 \exp\left(\frac{-\kappa-\beta}{2}t\right), \\ \frac{\partial}{\partial t}u(t) &= B_1\frac{-\kappa+\beta}{2}\exp\left(\frac{-\kappa+\beta}{2}t\right) + B_2\frac{-\kappa-\beta}{2}\exp\left(\frac{-\kappa-\beta}{2}t\right), \end{aligned}$$

where $\beta = \sqrt{\kappa^2 + \frac{1}{4}\delta^2}$. Resubstituting this back into (11.1.11) gives us the following solution for $C(t)$:

$$C(t) = \frac{2}{\delta^2}\frac{B_1\frac{-\kappa+\beta}{2}\exp\left(\frac{-\kappa+\beta}{2}t\right) + B_2\frac{-\kappa-\beta}{2}\exp\left(\frac{-\kappa-\beta}{2}t\right)}{B_1\exp\left(\frac{-\kappa+\beta}{2}t\right) + B_2\exp\left(\frac{-\kappa-\beta}{2}t\right)}.$$

As $C(0) = 0$, we can choose $B_1 = \frac{-\kappa-\beta}{-\kappa+\beta} = \frac{\beta+\kappa}{\beta-\kappa}$ and $B_2 = 1$. This leaves us with the following formula for $C(t)$:

$$\begin{aligned} C(t) &= \frac{1}{\delta^2}\frac{(\kappa+\beta)\exp\left(\frac{-\kappa+\beta}{2}t\right) - (\kappa+\beta)\exp\left(\frac{-\kappa-\beta}{2}t\right)}{\frac{\beta+\kappa}{\beta-\kappa}\exp\left(\frac{-\kappa+\beta}{2}t\right) + \exp\left(\frac{-\kappa-\beta}{2}t\right)} \\ &= \frac{(\beta+\kappa)(\beta-\kappa)}{\delta^2}\frac{\exp\left(\frac{-\kappa+\beta}{2}t\right) - \exp\left(\frac{-\kappa-\beta}{2}t\right)}{(\beta+\kappa)\exp\left(\frac{-\kappa+\beta}{2}t\right) + (\beta-\kappa)\exp\left(\frac{-\kappa-\beta}{2}t\right)} \\ &= \frac{1}{4}\frac{(\exp(\beta t) - 1)}{(\beta+\kappa)\exp(\beta t) + (\beta-\kappa)}. \end{aligned}$$

Now, we find the solution for $A(t)$:

$$\frac{\partial}{\partial t}A(t) = -\kappa\theta C(t), \quad A(0) = 0,$$

and obtain

$$A(t) = -\int_0^t \kappa\theta C(t)dt$$

$$
\begin{aligned}
&= -\frac{\kappa\theta}{4}\int_0^t \frac{(\exp(\beta t)-1)}{(\beta+\kappa)\exp(\beta t)+(\beta-\kappa)}dt \\
&= -\frac{\kappa\theta}{4}\Bigg\{\frac{1}{\beta}\int_1^{e^{\beta t}} \frac{u}{(\beta+\kappa)u+(\beta-\kappa)}\frac{1}{u}du \\
&\quad + \frac{1}{\beta}\int_1^{e^{-\beta t}} \frac{1}{(\beta+\kappa)/u+(\beta-\kappa)}\frac{1}{u}du\Bigg\} \\
&= -\frac{\kappa\theta}{4\beta}\Bigg\{\frac{1}{\beta+\kappa}\Big[\ln\Big((\beta+\kappa)u+(\beta-\kappa)\Big)\Big]_1^{e^{\beta t}} \\
&\quad + \frac{1}{\beta-\kappa}\Big[\ln\Big((\beta-\kappa)u+(\beta+\kappa)\Big)\Big]_1^{e^{-\beta t}}\Bigg\} \\
&= -\frac{\kappa\theta}{4\beta}\Bigg\{\frac{1}{\beta^2-\kappa^2}\left[-(\beta+\kappa)\ln(2\beta)-(\beta-\kappa)\ln(2\beta)\right] \\
&\quad + \frac{1}{\beta+\kappa}\ln\Big((\beta+\kappa)e^{\beta t}+(\beta-\kappa)\Big) \\
&\quad + \frac{1}{\beta-\kappa}\ln\Big((\beta+\kappa)e^{\beta t}+(\beta-\kappa)\Big)+\frac{\ln(e^{\beta t})}{\beta-\kappa}\Bigg\} \\
&= \frac{\kappa\theta}{4\beta}\frac{1}{\beta^2-\kappa^2}\Bigg\{2\beta\ln(2\beta)-(\beta-\kappa)\ln\Big((\beta+\kappa)e^{\beta t}+(\beta-\kappa)\Big) \\
&\quad -(\beta+\kappa)\ln\Big((\beta+\kappa)e^{\beta t}-(\beta-\kappa)\Big)+(\beta+\kappa)\ln(e^{\beta t})\Bigg\} \\
&= \frac{2\kappa\theta}{\delta^2}\left\{\ln\left(\frac{2\beta}{(\beta+\kappa)e^{\beta t}+(\beta-\kappa)}+\frac{\beta+\kappa}{2\beta}\ln(e^{\beta t})\right)\right\} \\
&= \frac{2\kappa\theta}{\delta^2}\ln\left(\frac{2\beta e^{\frac{\beta+\kappa}{2}t}}{(\beta+\kappa)e^{\beta t}+(\beta-\kappa)}\right).
\end{aligned}
$$

Hence, $P(t,X)$ can be written as

$$
P_X(t,T) = \exp\left(-\frac{r_\$+r_€}{2}(T-t)+A(T-t)-C(T-t)v(t)\right),
$$

where

$$
\begin{aligned}
A(t) &= \frac{2\kappa\theta}{\delta^2}\ln\left(\frac{2\beta e^{(\beta+\kappa)t/2}}{(\beta+\kappa)\left(e^{\beta t}-1\right)+2\beta}\right), \\
C(t) &= \frac{1}{4}\frac{e^{\beta t}-1}{(\beta+\kappa)\left(e^{\beta t}-1\right)+2\beta},
\end{aligned}
$$

with

$$\beta = \sqrt{\kappa^2 + \delta^2/4}.$$

We note that

$$dP_X = r_X(t)P_X dt - \delta C(T-t)\sqrt{v}P_X dW_2^{Q^X}(t).$$

Next, we obtain

$$\begin{aligned}
&\mathbb{E}_{Q^X}\left[\frac{dQ_T^X}{dQ^X}\middle| \mathcal{F}_t\right] \\
&\quad= \frac{P_X(t,T)}{P_X(0,T)B_X(t)} \\
&\quad= \exp\left(-\frac{1}{2}\int_0^t C^2(T-s)\delta^2 v(s)ds - \int_0^t C(T-s)\delta\sqrt{v(s)}dW_2^{Q^X}(s)\right).
\end{aligned} \tag{11.1.12}$$

Hence

$$dW_2^{Q_T^X} = dW_2^{Q^X} + C(T-t)\delta\sqrt{v(t)}dt.$$

To complete the change of measure, we need to look at $W_1^{Q_T^X}(t)$. To this end, we recall that both forward EUR-X and USD-X exchange rates,

$$F_{€/X}(t,T) = e^{-r_€(T-t)}\frac{\sqrt{f(t)}}{P_X(t,T)}$$

and

$$F_{\$/X}(t,T) = e^{-r_\$(T-t)}\frac{1}{P_X(t,T)\sqrt{f(t)}},$$

should be Q_T^X-martingales. It is not difficult to check that to achieve the above no-arbitrage requirement, we need

$$dW_1^{Q^X}(t) = dW_1^{Q_T^X}(t),$$

which is natural since the change of measure (11.1.12) does not depend on $W_1^{Q^X}(t)$.

Hence, applying Theorem 10.1.6 to the Heston model setting, we can price, e.g. the call option as (see (10.1.26)):

$$\begin{aligned} C_{€/\$}(0,T,K) &= \frac{e^{-r_\$ T}}{\mathbb{E}_{Q_T^X}\left[\frac{1}{\sqrt{f(T)}}\right]}\mathbb{E}_{Q_T^X}\left[\left(\sqrt{f(T)}-\frac{K}{\sqrt{f(T)}}\right)_+\right] \\ &= \frac{e^{-r_\$ T}}{\mathbb{E}_{Q_T^X}\left[F_{\$/X}(T,T)\right]}\mathbb{E}_{Q_T^X}\left[\left(F_{€/X}(T,T)-KF_{\$/X}(T,T)\right)_+\right], \end{aligned} \tag{11.1.13}$$

and

$$C_{\$/€}(0,T,K) = \frac{e^{-r_€ T}}{\mathbb{E}_{Q_T^X}\left[F_{€/X}(T,T)\right]}\mathbb{E}_{Q_T^X}\left[\left(F_{\$/X}(T,T)-KF_{€/X}(T,T)\right)_+\right], \tag{11.1.14}$$

where

$$\begin{aligned} dF_{€/X}(t,T) &= \frac{\sqrt{v}}{2}F_{€/X}(t,T)\left(\sqrt{1-\rho^2}\,dW_1^{Q^X}(t)+\rho\,dW_2^{Q_T^X}\right) \\ &\quad +\delta C(T-t)\sqrt{v}F_{€/X}(t,T)dW_2^{Q_T^X}, \\ dF_{\$/X}(t,T) &= -\frac{\sqrt{v}}{2}F_{\$/X}(t,T)\left(\sqrt{1-\rho^2}\,dW_1^{Q^X}(t)+\rho\,dW_2^{Q_T^X}\right) \\ &\quad +\delta C(T-t)\sqrt{v}F_{\$/X}(t,T)dW_2^{Q_T^X}, \\ dv &= \left(\kappa-C(T-t)\delta^2\right)\left(\frac{\kappa\theta}{\kappa-C(T-t)\delta^2}-v\right)dt+\delta\sqrt{v}dW_2^{Q_T^X}, \\ &\quad v(0)=v_0, \end{aligned} \tag{11.1.15}$$

and we require that for $0 \le t \le T$

$$\kappa/\delta^2 > C(t). \tag{11.1.16}$$

The prices (11.1.13) and (11.1.14) satisfy the foreign-domestic symmetry (see Theorem 9.4.1).

We note that in comparison with the classical Heston model (11.1.1), the model (11.1.15) has time dependence in the coefficients. For other time-dependent Heston models, see e.g. [10, 53] and references therein.

11.2 SABR MODEL

For simplicity again, let the interest rates for the USD and EUR money markets, $r_\$$ and $r_€$, be constant. Following Section 10.1.1, we can re-write the classical SABR model [54] for EUR-USD exchange rate $f(t)$ under the measure Q^X, and the corresponding SDEs for $S_{€/X} = \sqrt{f}$ and $S_{\$/X} = 1/\sqrt{f}$ take the form

$$\begin{aligned} d\sqrt{f} &= (r_X(t) - r_€)\sqrt{f}dt \\ &\quad + \frac{\sigma(t)}{2}\sqrt{f}\left(\sqrt{1-\rho^2}\,dW_1^{Q^X}(t) + \rho\,dW_2^{Q^X}(t)\right), \qquad (11.2.1) \\ d\frac{1}{\sqrt{f}} &= (r_X(t) - r_\$)\frac{1}{\sqrt{f}}dt - \frac{\sigma(t)}{2}\frac{1}{\sqrt{f}}\left(\sqrt{1-\rho^2}\,dW_1^{Q^X}(t) + \rho\,dW_2^{Q^X}(t)\right), \end{aligned}$$

$$d\sigma = \nu\sigma dW_2^{Q^X}(t),\ \sigma(0) = \alpha, \qquad (11.2.2)$$

where $W_1^{Q^X}(t)$ and $W_2^{Q^X}$ are independent standard Wiener processes under Q^X, $\rho \in (-1, 0]$ is the correlation coefficient, $\nu > 0$ is the volatility of the volatility $\sigma(t)$, α is a positive constant, and (see (10.1.4))

$$r_X(t) = \frac{r_\$ + r_€}{2} + \frac{\sigma^2(t)}{8}\,.$$

We note that the parameter known as β in the classical SABR model is taken to be equal to 1 here, which is the typical requirement for FX modelling as it ensures that the SDE for the exchange rate for the inverse pair $1/f$ has the same form as for f.

By Theorem 10.1.5, e.g. the price of the call (in USD) for buying €1 for \$$K$ is equal to

$$C_{€/\$}((0, T, K) = \sqrt{f(0)}\mathbb{E}_{Q^X}\left[D_X(0,T)\left(\sqrt{f(T)} - \frac{K}{\sqrt{f(T)}}\right)_+\right], \qquad (11.2.3)$$

where $\sqrt{f(T)}$ and $1/\sqrt{f(T)}$ satisfy (11.2.1), (11.2.2).

11.3 EXTENDED SKEW NORMAL MODEL

In this section we look at another illustration of Theorem 10.1.6. We start not with a model under a 'market' measure but with a direct assumption on the distribution of the exchange rate under a forward measure Q_T^X on the intermediate market.

Here, we assume that under a T-forward measure Q_T^X the EUR-USD exchange rate $f(T)$ can be written as

$$f(T) = \bar{F}e^Z, \tag{11.3.1}$$

where $\bar{F} > 0$ is a constant and Z is a random variable such that $\mathbb{E}\left[e^Z\right]$ exists and the no-arbitrage condition (10.1.19) is satisfied by $f(T)$. Here the no-arbitrage condition (10.1.19) implies that

$$\bar{F} = F\,\frac{\mathbb{E}\left[e^{-Z/2}\right]}{\mathbb{E}\left[e^{Z/2}\right]}, \tag{11.3.2}$$

where we neglect the full notation $\mathbb{E}_{Q_T^X}[\cdot]$ and write $\mathbb{E}[\cdot]$ instead as in this section we work with the measure Q_T^X only. Additionally, we write here F instead of $F_{€/\$}(0,T)$ for the current forward EUR-USD exchange rate (see (10.1.21)). We use this simplified notation throughout this section, which should not cause any confusion. The interest rates for the USD and EUR money markets, $r_\$$ and $r_€$, are assumed to be constant.

Also, (10.1.26) (i.e. Theorem 10.1.6), (11.3.1) and (11.3.2) imply that the price (in USD) of the European call to buy €1 for $\$K$ at the maturity T is

$$\begin{aligned} C_{€/\$}(0,T,K) &= \frac{e^{-r_\$ T}}{\mathbb{E}\left[\frac{1}{\sqrt{f(T)}}\right]}\mathbb{E}\left[\left(\sqrt{f(T)} - \frac{K}{\sqrt{f(T)}}\right)_+\right] \\ &= e^{-r_\$ T}\,\frac{\sqrt{\bar{F}}}{\mathbb{E}\left[e^{-Z/2}\right]}\mathbb{E}\left[\left(\sqrt{f(T)} - \frac{K}{\sqrt{f(T)}}\right)\mathbb{1}_{\bar{F}e^Z > K}\right] \\ &= e^{-r_\$ T}\,\frac{\sqrt{\bar{F}}}{\mathbb{E}\left[e^{-Z/2}\right]}\left(\sqrt{\bar{F}}\,\mathbb{E}\left[e^{Z/2}\mathbb{1}_{Z>z_0}\right] - \frac{K}{\sqrt{\bar{F}}}\,\mathbb{E}\left[e^{-Z/2}\mathbb{1}_{Z>z_0}\right]\right) \end{aligned} \tag{11.3.3}$$

$$
\begin{aligned}
&= e^{-r_\$ T}\left(\bar{F}\,\frac{\mathbb{E}\left[e^{Z/2}\mathbb{1}_{Z>z_0}\right]}{\mathbb{E}\left[e^{-Z/2}\right]} - K\,\frac{\mathbb{E}\left[e^{-Z/2}\mathbb{1}_{Z>z_0}\right]}{\mathbb{E}\left[e^{-Z/2}\right]}\right)\\
&= e^{-r_\$ T}\left(F\,\frac{\mathbb{E}\left[e^{Z/2}\mathbb{1}_{Z>z_0}\right]}{\mathbb{E}\left[e^{Z/2}\right]} - K\,\frac{\mathbb{E}\left[e^{-Z/2}\mathbb{1}_{Z>z_0}\right]}{\mathbb{E}\left[e^{-Z/2}\right]}\right)\\
&= e^{-r_\$ T}\left(F\,\frac{M(1/2,z_0)}{M(1/2)} - K\,\frac{M(-1/2,z_0)}{M(-1/2)}\right),
\end{aligned}
$$

where $z_0 = \log\left(K/\bar{F}\right)$ and

$$
M(t) = \mathbb{E}[e^{tZ}] \text{ and } M(t,z_0) = \mathbb{E}\left[e^{tZ}\mathbb{1}_{Z>z_0}\right], \tag{11.3.4}
$$

which are the moment generating function (MGF) and the restricted MGF for Z, respectively.

Analogous to (11.3.3), we can derive the pricing formulas for the put and also for the call and put for the inverse pair:

$$
\begin{aligned}
P_{€/\$}(0,T,K) &= \frac{e^{-r_\$ T}}{\mathbb{E}\left[\frac{1}{\sqrt{f(t)}}\right]}\mathbb{E}\left[\left(\frac{K}{\sqrt{f(T)}} - \sqrt{f(T)}\right)_+\right]\\
&= e^{-r_\$ T}\left(K\,\frac{M^*(-1/2,z_0)}{M(-1/2)} - F\,\frac{M^*(1/2,z_0)}{M(1/2)}\right),\\
C_{\$/€}\left(0,T,\frac{1}{K}\right) &= \frac{e^{-r_€ T}}{\mathbb{E}\left[\sqrt{f(t)}\right]}\mathbb{E}\left[\left(\frac{1}{\sqrt{f(T)}} - \frac{\sqrt{f(T)}}{K}\right)_+\right]\\
&= e^{-r_€ T}\left(\frac{1}{F}\,\frac{M^*(-1/2,z_0)}{M(-1/2)} - \frac{1}{K}\,\frac{M^*(1/2,z_0)}{M(1/2)}\right),\\
P_{\$/€}\left(0,T,\frac{1}{K}\right) &= \frac{e^{-r_€ T}}{\mathbb{E}\left[\sqrt{f(t)}\right]}\mathbb{E}\left[\left(\frac{\sqrt{f(T)}}{K} - \frac{1}{\sqrt{f(T)}}\right)_+\right]\\
&= e^{-r_€ T}\left(\frac{1}{K}\,\frac{M(1/2,z_0)}{M(1/2)} - \frac{1}{F}\,\frac{M(-1/2,z_0)}{M(-1/2)}\right),
\end{aligned}
$$

where

$$
M^*(t,z_0) = \mathbb{E}\left[e^{tZ}\mathbb{1}_{Z<z_0}\right].
$$

It is easy to show that these pricing formulas satisfy the Foreign-Domestic symmetry (see Theorem 9.4.1):

$$\begin{aligned} C_{€/\$}(0,T,K) &= e^{-r_\$ T}\left(F\frac{M(1/2,z_0)}{M(1/2)} - K\frac{M(-1/2,z_0)}{M(-1/2)}\right) \\ &= e^{-r_\$ T}\cdot F\cdot K\left(\frac{1}{K}\frac{M(1/2,z_0)}{M(1/2)} - \frac{1}{F}\frac{M(-1/2,z_0)}{M(-1/2)}\right) \\ &= e^{-r_\$ T} f(0) e^{(r_\$ - r_€)T} K\left(\frac{1}{K}\frac{M(1/2,z_0)}{M(1/2)} - \frac{1}{F}\frac{M(-1/2,z_0)}{M(-1/2)}\right) \\ &= e^{-r_€ T}\left(\frac{1}{K}\frac{M(1/2,z_0)}{M(1/2)} - \frac{1}{F}\frac{M(-1/2,z_0)}{M(-1/2)}\right) \\ &= f(0) K e^{-r_€ T}\left(\frac{1}{K}\frac{M(1/2,z_0)}{M(1/2)} - \frac{1}{F}\frac{M(-1/2,z_0)}{M(-1/2)}\right) \\ &= f(0)\cdot K\cdot P_{\$/€}\left(0,T,\frac{1}{K}\right). \end{aligned}$$

The same clearly also holds for:

$$P_{€/\$}(0,T,K) = f(0)\cdot K\cdot C_{\$/€}\left(0,T,\frac{1}{K}\right).$$

Let us now propose a skew normal model for the random variable Z. To this end, we start by introducing a new random variable V, which is a combination of one normal and two shifted half-normal distributed random variables:

$$V := X + \alpha_1 \max(\beta_1 - Y, 0) + \alpha_2 \max(Y - \beta_2, 0), \tag{11.3.5}$$

where X and Y are independent random variables with the standard normal distribution and α_1, α_2, β_1, $\beta_2 \in \mathbb{R}$ are parameters. The support domains of the two half-normal distributions which, from the modelling perspective, should not overlap and can be described by the parameters β_1 and β_2. Therefore, we are only interested in the following case:

$$0 < \beta_1 \leq \beta_2.$$

As we use the random variable Z in (11.3.1) similarly to a Gaussian random variable is used in the geometric Brownian motion model for $f(T)$, we then define it as follows

$$Z = aV, \tag{11.3.6}$$

where $a = \sigma\sqrt{T}$ with σ having the meaning of volatility and T of the maturity time. The advantage of using Z instead of a Gaussian random variable is that Z can have heavier tails and can be successfully used for describing the volatility smile effect, which we mentioned in Section 9.3. Meanwhile, Z still has a very simple distribution which makes the model (11.3.1), (11.3.5), (11.3.6) very practical as it allows fast calibration. Indeed, the MGFs (11.3.4), which we need for pricing calls and puts (see (11.3.3)), can be found analytically for this Z. The corresponding expressions are given in the next proposition.

Proposition 11.3.1. *For $0 < \beta_1 \leq \beta_2$, the MGF $M(t)$ and the restricted MGF $M(t, z_0)$ from (11.3.4) are equal to*

$$M(t) = e^{\frac{(at)^2}{2}}(N(\beta_2) - N(\beta_1)) + e^{\frac{t}{2}(ta^2(1+\alpha_2^2) - 2a\alpha_2\beta_2)} N(ta\alpha_2 - \beta_2)$$
$$+ e^{\frac{t}{2}(ta^2(1+\alpha_1^2)+2a\alpha_1\beta_1)} N(ta\alpha_1 + \beta_1), \tag{11.3.7}$$
$$M(t, z_0) = e^{\frac{(at)^2}{2}} N\left(at - \frac{z_0}{a}\right)(N(\beta_2) - N(\beta_1)) + e^{\frac{t}{2}(ta^2(1+\alpha_1^2)+2a\alpha_1\beta_1)} \tag{11.3.8}$$
$$\times \left\{ N(ta\alpha_1 + \beta_1) - N_2\left(ta\alpha_1 + \beta_1, \frac{\frac{z_0}{a} - at - \alpha_1(\beta_1 + ta\alpha_1)}{\sqrt{1+\alpha_1^2}}; \frac{-\alpha_1}{\sqrt{1+\alpha_1^2}}\right)\right\}$$
$$+ e^{\frac{t}{2}(ta^2(1+\alpha_2^2)-2a\alpha_2\beta_2)}$$
$$\times \left\{ N(ta\alpha_2 - \beta_2) - N_2\left(ta\alpha_2 - \beta_2, \frac{\frac{z_0}{a} - t + \alpha_2(\beta_2 - ta\alpha_2)}{\sqrt{1+\alpha_2^2}}; \frac{-\alpha_2}{\sqrt{1+\alpha_2^2}}\right)\right\},$$

where $N(\cdot)$ is the cdf of the standard normal distribution and $N_2(\cdot, \cdot; \rho)$ is the cdf of the bivariate normal distribution with zero mean, unit variance, and correlation ρ.

The proof for Proposition 11.3.1 can be found in Apendix B.5 (and also in [85]).

The distribution of Z, defined in (11.3.5), (11.3.6), is dependent on the five parameters a, α_1, α_2, β_1, and β_2. All of them can be used to manipulate the distribution of Z and, in particular, its skew and kurtosis, which are equal to

$$skew_Z = skew_V = \frac{M_V^{(3)}(0) - 3M_V^{(1)}(0)M_V^{(2)}(0) + 2\left[M_V^{(1)}(0)\right]^3}{\left(M_V^{(2)}(0) - \left[M_V^{(1)}(0)\right]^2\right)^{3/2}},$$

$$kurtosis_Z = kurtosis_V$$

$$= \frac{M_V^{(4)}(0) - 4M_V^{(1)}(0)M_V^{(3)}(0) + 6\left[M_V^{(1)}(0)\right]^2 M_V^{(2)}(0) - 3\left[M_V^{(1)}(0)\right]^4}{\left(M_V^{(2)}(0) - \left[M_V^{(1)}(0)\right]^2\right)^2},$$

where $M_V^{(i)}(0)$ are i-th derivatives of the MGF for the random variable V. The analytical formulae for these can be found in Appendix B.6.

By putting $\alpha_1 = \alpha_2 = 0$ in (11.3.5), the random variable Z becomes normal with zero mean and variance a^2, and the considered model (11.3.1), (11.3.5), (11.3.6) is reduced to the geometric Brownian motion whose one of the critical deficiencies is a flat (constant) volatility. In this case, Z has $skew = 0$ and $kurtosis = 3$. In Figure 11.3.1, one can see the difference of distribution of Z (blue area) compared to a standardized normal distribution (red line). It can be seen that a parameter set with $\alpha_1 < 0$ and $\alpha_2 = 0$ results in a bigger left tail in distribution and a skew in the resulting volatility smile ($skew \approx -1.6$). Similarly, in Figure 11.3.2, it can be observed that using $\alpha_1 < 0$ to adjust the left tail and $\alpha_2 > 0$ to adjust the right tail of the smile, we can get an asymmetric distribution and an asymmetric smile. A smaller α_2 results in a smaller right tail and therefore in a flatter smile. As seen in these figures, by adjusting the parameters α_1, α_2, β_1, β_2, the shape of the distribution and of the smile can be changed in various ways and it can be associated with the resulting skew and kurtosis of the log exchange rate. Therefore, after calibrating the parameters of Z to FX market data, we can compare $skew_Z$ with zero skew and $kurtosis_Z$ with the kurtosis of 3 in the geometric

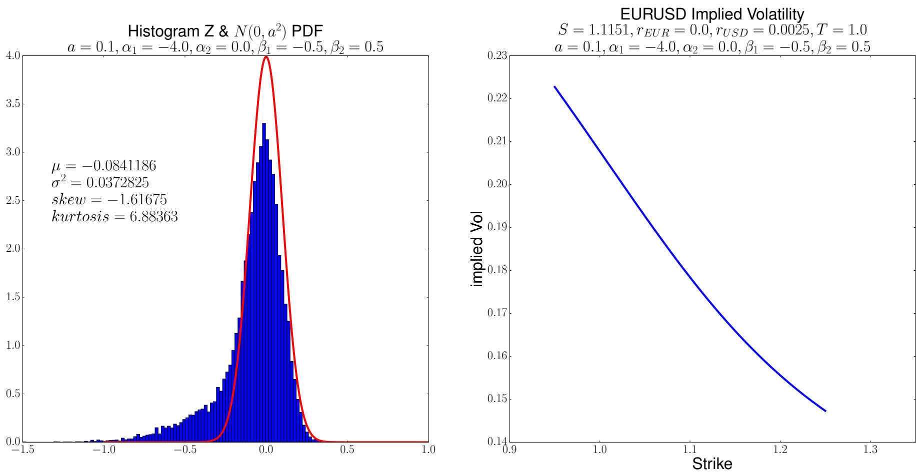

Figure 11.3.1: Effect of the parameters on the distribution of Z and the corresponding smile: the case of $\alpha_1 < 0$ and $\alpha_2 = 0$.

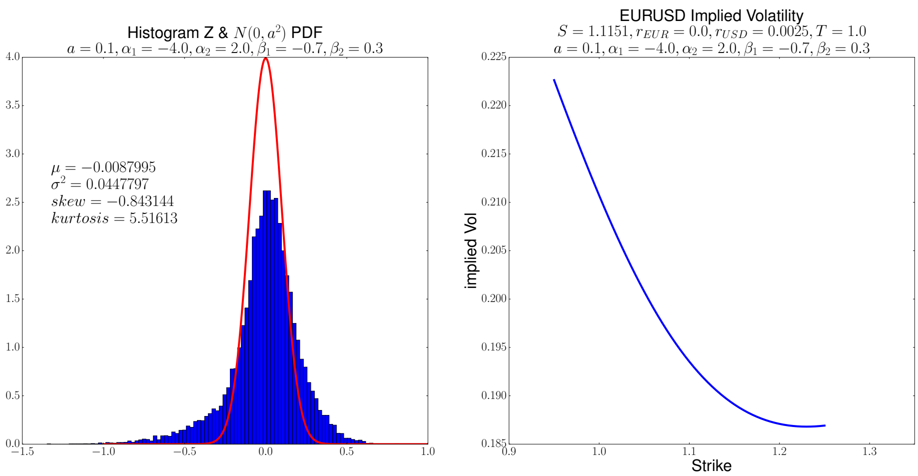

Figure 11.3.2: Effect of the parameters on the distribution of Z and the corresponding smile: the case of $\alpha_1 < 0$ and $\alpha_2 > 0$.

Brownian motion case and make a conclusion about how far volatility is from a constant.

11.4 MODEL-FREE APPROACH

The model-free approach to pricing derivatives has become popular in recent years [6, 40, 7, 27] (see also references therein). The main idea of the approach is to construct a density or distribution function of risky assets under a risk-neutral measure using observed prices of plain-vanilla options. For clarity of the exposition how this approach works within our intermediate currency framework, we start with the case of two currencies in Subsection 11.4.1. Then we will extend the consideration to the three-currencies case where we will exploit ideas from [6] (see also [7]) in Subsection 11.4.2.

11.4.1 *Model-free approach in two dimensions*

In this subsection we will work under a T–forward measure Q_T^X. Assume that we know prices of call options $C_{€/\$}(0;K)$ for all strikes $K > 0$ and let $\rho(x;T)$ be the density of the EUR-USD exchange rate $f(T)$ under Q_T^X. According to (10.1.24), we have

$$\begin{aligned} V_{€/\$}(0) &= \frac{e^{-r_\$ T}}{\mathbb{E}_{Q_T^X}\left[\sqrt{f(T)}\right]}\mathbb{E}_{Q_T^X}\left[g\left(\sqrt{f(T)};\frac{K}{\sqrt{f(T)}}\right)\right] \qquad (11.4.1) \\ &= \frac{e^{-r_\$ T}}{\mathbb{E}_{Q_T^X}\left[\sqrt{f(T)}\right]}\int_0^\infty \frac{1}{\sqrt{x}} g\left(x;K\right)\rho(x;T)dx \\ &= \int_0^\infty g\left(x;K\right)\frac{\partial^2}{\partial x^2}C_{€/\$}(0;x)dx, \end{aligned}$$

as simple calculations show that the first and second derivative with respect to the strike K of a call option price (see (10.1.26)) can be written as follows:

$$\begin{aligned}
\frac{\partial C_{€/\$}(0;K)}{\partial K} &= -\frac{e^{-r_\$ T}}{\mathbb{E}_{Q_T^X}\left[\frac{1}{\sqrt{f(T)}}\right]} \int_K^\infty \frac{1}{\sqrt{x}}(x-K)\rho(x;T)dx, \\
\frac{\partial^2 C_{€/\$}(0;K)}{\partial K^2} &= \frac{e^{-r_\$ T}}{\mathbb{E}_{Q_T^X}\left[\frac{1}{\sqrt{f(T)}}\right]} \frac{\partial^2}{\partial K^2} \int_\infty^K \frac{1}{\sqrt{x}}(x-K)\rho(x;T)dx \\
&= \frac{e^{-r_\$ T}}{\mathbb{E}_{Q_T^X}\left[\frac{1}{\sqrt{f(T)}}\right]} \frac{1}{\sqrt{K}}\rho(K;T),
\end{aligned}$$

hence

$$\frac{e^{-r_\$ T}}{\sqrt{K}\mathbb{E}_{Q_T^X}\left[\sqrt{f(T)}\right]}\rho(K;T) = \frac{\partial^2}{\partial K^2} C_{€/\$}(0;K). \tag{11.4.2}$$

Typically, observed data are expressed via volatility smile data $\sigma(K)$ and from (10.1.8) we have

$$\begin{aligned}
C_{€/\$}(0;K) = F_{€/\$} e^{-r_\$ T} N\left(\frac{\log \frac{F_{€/\$}}{K} + \sigma^2(K)T/2}{\sigma(K)\sqrt{T}}\right) \\
-Ke^{-r_\$ T} N\left(\frac{\log \frac{F_{€/\$}}{K} - \sigma^2(K)T/2}{\sigma(K)\sqrt{T}}\right).
\end{aligned} \tag{11.4.3}$$

Combining (11.4.1) with (11.4.3) and given $\sigma(K)$, we can price any FX derivative $V_{€/\$}(0)$ and analogously any derivative $V_{\$/€}(0)$ based on a smile from one of the markets. Note that the smile data computed from $C_{€/\$}(0;K)$ coincide with smile data computed from $C_{\$/€}(0;K)$ and that the prices $V_{€/\$}(0)$ and $V_{\$/€}(0)$ are consistent with each other thanks to using the intermediate currency framework, as introduced in Section 10.1.

11.4.2 *Model-free approach in three dimensions*

We can now progress to the three-currencies case, where we follow the same idea as in Subsection 11.4.1 to construct a density of distributon function, but we will see that the derivations are more complicated due to the third dimension as expected.

Let us assume that we are interested in the GBP-USD-EUR currency triangle, where we denote GBP as currency 1, USD as currency 2, and EUR as currency 3. As before, the interest rates for the GBP, USD and EUR money markets, $r_{£}$, $r_{\$}$ and $r_{€}$, are assumed to be constant.

Let us now look at a best-of option on the EUR market which payoff is equal to

$$b(T) = \max\left\{\frac{(S_{£/€}(T)-K_1)_+}{K_1}, \frac{(S_{\$/€}(T)-K_2)_+}{K_2}\right\}. \tag{11.4.4}$$

As it is known [6], the value of a best-of option is arbitrary close to values of plain-vanilla calls on $S_{£/€}(T)$ or $S_{\$/€}(T)$ or to a vanilla option on the cross $S_{£/\$}(T)$. Hence, a model used for FX pricing should price a best-of option and plain-vanilla options in a consistent manner.

By (10.2.2) we have

$$\begin{aligned} S_{£/X} &= S_{£/€}^{2/3}(T)S_{\$/€}^{-1/3}(T), \\ S_{\$/X} &= S_{£/€}^{-1/3}(T)S_{\$/€}^{2/3}(T), \\ S_{€/X} &= S_{£/€}^{-1/3}(T)S_{\$/€}^{-1/3}(T). \end{aligned} \tag{11.4.5}$$

Using the pricing formula (10.2.23), we get

$$V_{€}(0) = \frac{e^{-r_{€}T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]}\mathbb{E}_{Q_T^X}G(T) = \frac{e^{-r_{€}T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]}\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)g(T)\right], \tag{11.4.6}$$

where $g(T)$ is an arbitrary payoff on the EUR market. Therefore, for the best-of option we have

$$\begin{aligned} v_{€}(0) &= \frac{e^{-r_{€}T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} \\ &\quad \times \mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\ \max\left\{\frac{(S_{£/€}(T)-K_1)_+}{K_1}, \frac{(S_{\$/€}(T)-K_2)_+}{K_2}\right\}\right] \end{aligned}$$

$$= \frac{e^{-r_€ T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} \mathbb{E}_{Q_T^X}\left[S_{£/€}^{-1/3}(T) S_{\$/€}^{-1/3}(T)\right.$$
$$\left. \max\left\{\frac{(S_{£/€}(T)-K_1)_+}{K_1}, \frac{(S_{\$/€}(T)-K_2)_+}{K_2}\right\}\right]$$
$$= \frac{e^{-r_€ T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]}$$
$$\times \int_0^\infty \int_0^\infty (xy)^{-1/3} \max\left\{\frac{(x-K_1)_+}{K_1}, \frac{(y-K_2)_+}{K_2}\right\} \rho(x,y;T) dx dy,$$

where $\rho(x,y;T)$ is the joint density of the exchange rates $S_{£/€}(T)$ and $S_{\$/€}(T)$ under Q_T^X.

Differentiation of $b(T)$ (see (B.7.1) in Appendix B.7.1) gives:

$$\left[1 + K_1 \frac{\partial}{\partial K_1} + K_2 \frac{\partial}{\partial K_2}\right] b(T) = \mathbf{I}(S_{£/€}(T) < K_1, S_{£/€}(T) < K_2) - 1.$$

Therefore

$$\frac{\partial^2}{\partial K_1 \partial K_2}\left[1 + K_1 \frac{\partial}{\partial K_1} + K_2 \frac{\partial}{\partial K_2}\right] v_€(0)$$
$$= \frac{e^{-r_€ T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} K_1^{-1/3} K_2^{-1/3} \rho(K_1, K_2; T). \tag{11.4.7}$$

Note that we have from (10.2.20)

$$\frac{\mathbb{E}_{Q_T^X}\left[S_{c_i/X}(T)\right]}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} = F_{c_i/€}(0), \quad i = 1, 2, \tag{11.4.8}$$

and from (10.2.23)

$$V_{c_N}(0) = \frac{e^{-r_N T}}{\mathbb{E}_{Q_T^X}\left[S_{c_N/X}(T)\right]} \mathbb{E}_{Q_T^X}\left[S_{c_N/X}(T) g(T)\right], \tag{11.4.9}$$

where N can be any of the three currencies with $g(T)$ being in the currency N. Therefore (taking also into account (11.4.5)), if we can evaluate (11.4.7) from market data, then we can price any FX derivatives on any of the three markets using the same $\rho(K_1, K_2; T)$ and, thus, ensuring consistency of

FX option pricing across different markets. Note that in comparison with (10.2.23) we do not assume in (11.4.9) that $g(T)$ is scalable.

Market data in the case of three currencies are typically presented via three volatility smiles: $\sigma_1(K)$ and $\sigma_2(K)$ from vanilla options on GBP-EUR and USD-EUR, respectively, and $\sigma_3(K)$ from the cross, GBP-USD. To compute values on the smile curves from observed option prices in the context of our intermediate currency approach, the Garman-Kohlhagen formulas given below for completeness of the exposition should be used.

To complete, the model-free pricing, we need to express the current price $v(0)$ of the best-of option via the three volatility smiles. To this end, we need to find $v(0)$ assuming that the exchange rates follow geometrical Brownian motions under a T–forward measure Q^X_T, which coincides with the EMM Q^X.

We can set the exchange rates $S_{£/€}(T)$ and $S_{\$/€}(T)$ as follows:

$$\begin{aligned} S_{£/€}(T) &= F_{£/€} \exp\left(-aT + \sigma_1 \sqrt{T} X_1\right), \\ S_{\$/€}(T) &= F_{\$/€} \exp\left(-bT + \sigma_2 \sqrt{T} X_2\right), \end{aligned} \tag{11.4.10}$$

where $F_{£/€} = F_{£/€}(0,T)$ and $F_{\$/€} = F_{\$/€}(0,T)$ are the current forward GBR-EUR and USD-EUR exchange rates, respectively, and $X_i \sim N(0,1)$ with correlation coefficient ρ_{12}, and where we need to find $a, b \in \mathbb{R}$, so that the no-arbitrage conditions (11.4.8) hold.

Let us use (11.4.10) to rewrite the equations $S_{c_i/X}(T)$ (11.4.5) which lead to the following:

$$\begin{aligned} S_{£/X}(T) &= F_{£/€}^{2/3} F_{\$/€}^{-1/3} \exp\left(-\frac{2}{3}aT + \frac{2}{3}\sigma_1\sqrt{T}X_1 + \frac{1}{3}bT - \frac{1}{3}\sigma_2\sqrt{T}X_2\right), \\ S_{\$/X}(T) &= F_{£/€}^{-1/3} F_{\$/€}^{2/3} \exp\left(\frac{1}{3}aT - \frac{1}{3}\sigma_1\sqrt{T}X_1 - \frac{2}{3}bT + \frac{2}{3}\sigma_2\sqrt{T}X_2\right), \\ S_{€/X}(T) &= F_{£/€}^{-1/3} F_{\$/€}^{-1/3} \exp\left(\frac{1}{3}aT - \frac{1}{3}\sigma_1\sqrt{T}X_1 + \frac{1}{3}bT - \frac{1}{3}\sigma_2\sqrt{T}X_2\right). \end{aligned}$$

The no-arbitrage (11.4.8) condition for $i = 1$ then leads to:

$$F_{£/€} = \frac{\mathbb{E}_{Q_T^X}\left[S_{£/X}(T)\right]}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]}.$$

Hence

$$F_{£/€} = \frac{\mathbb{E}_{Q_T^X}\left[F_{£/€}^{2/3} F_{\$/€}^{-1/3} \exp\left(-\frac{2}{3}aT + \frac{2}{3}\sigma_1\sqrt{T}X_1 + \frac{1}{3}bT - \frac{1}{3}\sigma_2\sqrt{T}X_2\right)\right]}{\mathbb{E}_{Q_T^X}\left[F_{£/€}^{-1/3} F_{\$/€}^{-1/3} \exp\left(\frac{1}{3}aT - \frac{1}{3}\sigma_1\sqrt{T}X_1 + \frac{1}{3}bT - \frac{1}{3}\sigma_2\sqrt{T}X_2\right)\right]},$$

which we can rewrite as follows:

$$\begin{aligned}
\exp(aT) &= \frac{\mathbb{E}_{Q_T^X}\left[\exp\left(\frac{2}{3}\sigma_1\sqrt{T}X_1 - \frac{1}{3}\sigma_2\sqrt{T}X_1\rho - \frac{1}{3}\sigma_2\sqrt{T}X_2\sqrt{1-\rho^2}\right)\right]}{\mathbb{E}_{Q_T^X}\left[\exp\left(-\frac{1}{3}\sigma_1\sqrt{T}X_1 - \frac{1}{3}\sigma_2\sqrt{T}X_1\rho - \frac{1}{3}\sigma_2\sqrt{T}X_2\sqrt{1-\rho^2}\right)\right]} \\
&= \frac{\exp\left(\frac{T}{2}\left[\frac{2}{3}\sigma_1 - \frac{1}{3}\sigma_2\rho\right]^2 + \frac{T}{18}\left[\sigma_2\sqrt{1-\rho^2}\right]^2\right)}{\exp\left(\frac{T}{2}\left[\frac{1}{3}\sigma_1 + \frac{1}{3}\sigma_2\rho\right]^2 + \frac{T}{18}\left[\sigma_2\sqrt{1-\rho^2}\right]^2\right)} \\
&= \frac{\exp\left(\frac{T}{2}\left[\frac{4}{9}\sigma_1^2 - \frac{4}{9}\sigma_1\sigma_2\rho + \frac{1}{9}\sigma_2^2\right]\right)}{\exp\left(\frac{T}{2}\left[\frac{1}{9}\sigma_1^2 + \frac{2}{9}\sigma_1\sigma_2\rho + \frac{1}{9}\sigma_2^2\right]\right)} = \exp\left(\frac{T}{6}\left[\sigma_1^2 - 2\sigma_1\sigma_2\rho\right]\right).
\end{aligned}$$

And we can solve for a to end up with

$$a = \frac{1}{6}\left[\sigma_1^2 - 2\sigma_1\sigma_2\rho\right].$$

In the same way for $i = 2$ we can find b:

$$b = \frac{1}{6}\left[\sigma_2^2 - 2\sigma_1\sigma_2\rho\right].$$

Thus

$$\begin{aligned}
S_{£/€}(T) &= F_{£/€}\exp\left(-\frac{T}{6}\left[\sigma_1^2 - 2\sigma_1\sigma_2\rho_{12}\right] + \sigma_1\sqrt{T}X_1\right), \\
S_{\$/€}(T) &= F_{\$/€}\exp\left(-\frac{T}{6}\left[\sigma_2^2 - 2\sigma_1\sigma_2\rho_{12}\right] + \sigma_2\sqrt{T}X_2\right).
\end{aligned} \tag{11.4.11}$$

Then, the GBP-USD exchange rate is equal to

$$S_{£/\$}(T) = \frac{S_{£/€}(T)}{S_{\$/€}(T)} = F_{£/\$} \exp\left(-\frac{T}{6}\left[\sigma_3^2 - 2\sigma_2\sigma_3\rho_{23}\right] + \sigma_3\sqrt{T}X_3\right), \tag{11.4.12}$$

where

$$\begin{aligned} F_{£/\$} &= \frac{F_{£/€}}{F_{\$/€}}, \\ \sigma_3^2 &= \sigma_1^2 - 2\sigma_1\sigma_2\rho_{12} + \sigma_2^2, \end{aligned}$$

and $X_3 \sim N(0,1)$ with the correlation coefficients

$$\rho_{13} = \frac{\sigma_1^2 + \sigma_3^2 - \sigma_2^2}{2\sigma_1\sigma_3}, \;\; \rho_{23} = \frac{\sigma_2^2 + \sigma_3^2 - \sigma_1^2}{2\sigma_2\sigma_3},$$

with X_1 and X_2, respectively.

We have

$$\begin{aligned} &\frac{S_{€/X}(T)}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} \\ &= \frac{\exp\left(-\frac{1}{3}\sigma_1\sqrt{T}X_1 - \frac{1}{3}\sigma_2\sqrt{T}X_2\right)}{\mathbb{E}_{Q_T^X}\left[\exp\left(-\frac{1}{3}\sigma_1\sqrt{T}X_1 - \frac{1}{3}\sigma_2\sqrt{T}X_2\right)\right]} \\ &= \exp\left(-\frac{T}{18}\left[\sigma_1^2 + 2\sigma_1\sigma_2\rho + \sigma_2^2\right] - \frac{1}{3}\sigma_1\sqrt{T}X_1 - \frac{1}{3}\sigma_2\sqrt{T}X_2\right), \end{aligned}$$

and it is not difficult to show that the above expression is the Radon-Nikodym derivative $\frac{dQ_T^€}{dQ_T^X}$ of the T–forward measure $Q_T^€$ on the EUR market with respect to Q_T^X. Then

$$V_€(0) = \frac{e^{-r_€ T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]}\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)g(T)\right] = e^{-r_€ T}\mathbb{E}_{Q_T^€}\left[g(T)\right].$$

Therefore, the corresponding Garman-Kohlhagen formulas for calls are given by (see, e.g. [15]):

$$C_{£/€}(0;K) = F_{£/€}e^{-r_€ T}N\left(\frac{\ln(F_{£/€}/K)+\sigma_1^2T/2}{\sigma_1\sqrt{T}}\right) - Ke^{-r_€ T}N\left(\frac{\ln(F_{£/€}/K)-\sigma_1^2T/2}{\sigma_1\sqrt{T}}\right),$$

$$C_{\$/€}(0;K) = F_{\$/€}e^{-r_€ T}N\left(\frac{\ln(F_{\$/€}/K)+\sigma_2^2T/2}{\sigma_2\sqrt{T}}\right) - Ke^{-r_€ T}N\left(\frac{\ln(F_{\$/€}/K)-\sigma_2^2T/2}{\sigma_2\sqrt{T}}\right).$$

Similarly

$$\begin{aligned} C_{£/\$}(0;K) &= \frac{e^{-r_£ T}}{\mathbb{E}_{Q_T^X}\left[S_{£/X}(T)\right]}\mathbb{E}_{Q_T^X}\left[S_{£/X}(T)(S_{\$/£}(T)-K)_+\right] \\ &= e^{-r_£ T}\mathbb{E}_{Q_T^£}\left[(S_{\$/£}(T)-K)_+\right] \\ &= F_{£/\$}e^{-r_\$ T}N\left(\frac{\ln(F_{£/\$}/K)+\sigma_3^2T/2}{\sigma_3\sqrt{T}}\right) \\ &\quad -Ke^{-r_\$ T}N\left(\frac{\ln(F_{£/\$}/K)-\sigma_3^2T/2}{\sigma_3\sqrt{T}}\right). \end{aligned}$$

To proceed, we want to make use of (11.4.7). We have the price of best–of option on the EUR market (see [83, 109, 6]):

$$\begin{aligned} v_€(0) &= \frac{e^{-r_€ T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} \\ &\quad \times\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\ \max\left\{\frac{(S_{£/€}(T)-K_1)_+}{K_1},\frac{(S_{\$/€}(T)-K_2)_+}{K_2}\right\}\right] \\ &= e^{-r_€ T}\mathbb{E}_{Q_T^€}\left[\max\left\{\frac{(S_{£/€}(T)-K_1)_+}{K_1},\frac{(S_{\$/€}(T)-K_2)_+}{K_2}\right\}\right] \\ &= e^{-r_€ T}\left[\frac{F_{£/€}}{K_1}N(d_1^+,d_3^+;\rho_{13})+\frac{F_{\$/€}}{K_2}N(d_2^+,d_3^-;\rho_{23})\right. \end{aligned} \quad (11.4.13)$$

$$+N(-d_1^-, -d_2^-; \rho_{12}) - 1\Big],$$

where

$$d_i^{\pm} = \frac{\ln(F_i/K_i) \pm \sigma_i^2 T/2}{\sigma_i\sqrt{T}}$$

and $F_1 = F_{£/€}$, $F_2 = F_{\$/€}$, and $F_3 = F_{£/\$}$.

Now, we put the implied volatility smiles $\sigma_i(K_i)$, $i = 1, 2, 3$, with $K_3 = K_1/K_2$, in (11.4.13) and evaluate the left-hand side of (11.4.7). As a result, we obtain a function of the strikes K_i for $i = 1, 2, 3$, for $v_{€}(0) = v_{€}(0; K_1, K_2, \sigma_1(K_1), \sigma_2(K_2), \sigma_3(K_1/K_2))$ from (11.4.13), and we can find the appropriate derivative to obtain the following function (see Appendix B.7.2 or [7, Ch. 11]):

$$U(K_1, K_2) := \left[1 + K_1\frac{\partial}{\partial K_1} + K_2\frac{\partial}{\partial K_2}\right] v_{€} + e^{-r_{€}T} \tag{11.4.14}$$

$$= e^{-r_{€}T}\left(N(-d_1^-, -d_2^-; \rho_{12}) + \left[K_1\sigma_1'(K_1)\frac{\partial}{\partial\sigma_1} + K_2\sigma_2'(K_2)\frac{\partial}{\partial\sigma_2}\right] v_{€} + 1\right)$$

$$= e^{-r_{€}T}\left[N(-d_1^-, -d_2^-; \rho_{12}) + K_1\sqrt{T}\sigma_1'(K_1)N'(d_1^-)N\left(\frac{d_1^-\rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right)\right.$$

$$\left. + K_2\sqrt{T}\sigma_2'(K_2)N'(d_2^-)N\left(\frac{d_2^-\rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}}\right)\right].$$

Note that $U(0, K_2) = U(K_1, 0) = 0$ is straight forward to calculate, as $N(-\infty, b; \rho) = 0$ and also $N(-\infty) = 0$.

Let us summarise how the model-free approach can be used in practice:

(i) for observed plain-vanilla prices, compute values of the implied volatilities $\sigma_i(K_i)$ by inverting the Garman-Kohlhagen formulas;

(ii) smoothly interpolate the implied values to obtain three smiles $\sigma_i(K_i)$;

(iii) plug-in the smiles in (11.4.14);

(iv) use $U(K_1, K_2)$ (cf. (11.4.7) and (11.4.14)) together with (11.4.8) to price options on all the three markets by the pricing formula (11.4.9).

Note that the final step (iv) can be either realized via integration by parts (see Example 11.4.1 below) or by further differentiation (see Appendix B.7.3) to get

$$\frac{e^{-r_{€}T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} K_1^{-1/3} K_2{}^{-1/3}\ \rho(K_1, K_2; T) = \frac{\partial^2}{\partial K_1 \partial K_2} U(K_1, K_2).$$

We highlight that thanks to the intermediate currency approach we can consistently price products for all the six pairs based on a single calibration.

We note that the no arbitrage condition imposes the following asymptotic requirements on smiles [78, 6, 7] (see also B.7.5)

$$\sigma_i^2(K) = o(|\ln K|) \text{ as } K \to 0,\ \infty. \tag{11.4.15}$$

Also, to ensure that $-1 < \rho_{ij}(K_1, K_1) < 1$ the smiles should satisfy [6, 7] (see also B.7.5):

$$\begin{aligned} \sigma_1(K_1) + \sigma_2(K_2) &> \sigma_3(K_1/K_2), \\ \sigma_2(K_2) + \sigma_3(K_1/K_2) &> \sigma_1(K_1), \\ \sigma_1(K_1) + \sigma_3(K_1/K_2) &> \sigma_2(K_2). \end{aligned} \tag{11.4.16}$$

Note, that if the requirements (11.4.15) and (11.4.16) are not satisfied by the volatility smiles, the constructed distribution function $U(K_1, K_2)$ or the corresponding pdf $\frac{\partial^2}{\partial K_1 \partial K_2} U(K_1, K_2)$ may not be real or positive.

Example 11.4.1. Consider basket pricing as in Example 10.2.2. Doing integration by parts twice and the fact that $U(0, y) = U(y, 0) = 0$, we get [7, Ch. 11]:

$$\begin{aligned} &BasketOption_{€}(0) \\ &\quad = \frac{e^{-r_{€}T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} \\ &\qquad \times \mathbb{E}_{Q_T^X}\left[S_{£/€}^{-1/3}(T) S_{\$/€}^{-1/3}(T) \left(K - \omega_1 S_{£/€}(T) - \omega_2 S_{\$/€}(T)\right)_+\right] \end{aligned} \tag{11.4.17}$$

$$
\begin{aligned}
&= \int_0^\infty \int_0^\infty (K - \omega_1 x - \omega_2 y)_+ \frac{\partial^2}{\partial x \partial y} U(x,y) dx dy \\
&= \int_0^{K/\omega_2} \int_0^{\frac{K-\omega_2 y}{\omega_1}} (K - \omega_1 x - \omega_2 y) \frac{\partial^2}{\partial x \partial y} U(x,y) dx dy \\
&= \int_0^{K/\omega_2} [(K - \omega_1 x - \omega_2 y) \frac{\partial}{\partial y} U(x,y)]_{x=0}^{x=\frac{K-\omega_2 y}{\omega_1}} dy \\
&\quad + \int_0^{K/\omega_2} \int_0^{\frac{K-\omega_2 y}{\omega_1}} \omega_1 \frac{\partial}{\partial y} U(x,y) dx dy \\
&= 0 + \int_0^{K/\omega_1} \int_0^{\frac{K-\omega_1 x}{\omega_2}} \omega_1 \frac{\partial}{\partial y} U(x,y) dy dx \\
&= \int_0^{K/\omega_1} \left([\omega_1 U(x,y)]_{y=0}^{y=\frac{K-\omega_1 x}{\omega_2}} - 0 \right) dx \\
&= \int_0^{K/\omega_1} \left[\omega_1 U(x, \frac{K - \omega_1 x}{\omega_2}) - \omega_1 U(x,0) \right] dx \\
&= \int_0^K U\left(\frac{z}{\omega_1}, \frac{K-z}{\omega_2} \right) dz.
\end{aligned}
$$

We also obtain

$$
\begin{aligned}
&BasketOption_{£}(0) \\
&\quad = \frac{e^{-r_{£} T}}{\mathbb{E}_{Q_T^X}\left[S_{£/X}(T)\right]} \mathbb{E}_{Q_T^X}\left[S_{£/X}(T)\left(K - \omega_1 S_{€/£}(T) - \omega_2 S_{\$/£}(T)\right)_+\right] \qquad (11.4.18) \\
&\quad = \frac{S_{€/£}(0) e^{-r_{€} T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} \times \\
&\qquad \mathbb{E}_{Q_T^X}\left[S_{£/€}^{2/3}(T) S_{\$/€}^{-1/3}(T) \left(K - \frac{\omega_1}{S_{£/€}(T)} - \omega_2 \frac{S_{\$/€}(T)}{S_{£/€}(T)} \right)_+ \right] \\
&\quad = S_{€/£}(0) \int_0^\infty \int_0^\infty x \left(K - \frac{\omega_1}{x} - \omega_2 \frac{y}{x} \right)_+ \frac{\partial^2}{\partial x \partial y} U(x,y) dx dy \\
&\quad = S_{€/£}(0) \int_0^\infty \left[U\left(\infty, \frac{z}{\omega_2}\right) - U\left(\frac{z\omega_2 + \omega_1}{K}, \frac{z}{\omega_2} \right) \right] dz,
\end{aligned}
$$

in a similar fashion, where

$$
U(\infty, K_2) = e^{-r_{€} T} \left[N(-d_2^-) + K_2 \sqrt{T} \sigma_2'(K_2) N'(d_2^-) - 1 \right].
$$

We note that if we set one of ω_i to zero in (11.4.17), then the formula gives the EUR price of a put on GBP or USD. Substituting U from (11.4.14) in (11.4.17) with one of ω_i being zero, we can recover the Black-Sholes price of the corresponding put which means that the pricing formula (11.4.17) (or what is the same, (11.4.9)) exactly reproduces the plain vanilla data to which the calibration is made. See a calibration illustration in Chapter 12. The difference with the approach of [6, 7] is that here we obtain the density which can be used to price on all the three markets.

Remark 11.4.1. We note that given a smile we get the exact pricing density (11.4.1) for a single pair. This allows us to combine marginals for each pair together with a copula to get a joint density for the triangle instead of using the approach based on the best-of option as considered here. In our case, we do not explore the use of copulas here, but for the corresponding discussion see [7].

12

NUMERICAL ILLUSTRATIONS AND CALIBRATION

In this chapter we present calibration examples for the models from Section 11.1 and 11.3 and we illustrate the model-free approach of Section 11.4.

We recall [22, 101] that the FX market is different to other financial markets in terms of volatility smile construction and quoting mechanisms used. FX options are quoted in implied volatility σ, delta Δ instead of strike K, and maturity T. The market convention is to quote three currency pair-specific most commonly traded options. Their choice depends on a delta hedging and ATM convention [101, 25] and typically 25Δ options are among the considered options. Occasionally, one also uses 10Δ put/call options, as they are widely available but not as liquid as 25Δ options [22]. The option prices are inverted to calculate the corresponding volatility values, which are used for constructing the volatility smile. The data we use in this chapter for calibration are given in Table 12.0.1.

Table 12.0.1: FX market data for 1 year maturity options, Bloomberg 03/06/2016.

	GBP-EUR	USD-EUR	GBP-USD
$\sigma_{25Put\Delta}$	12.435%	9.005%	11.000%
σ_{ATM}	10.945%	9.250%	13.072%
$\sigma_{25Call\Delta}$	10.345%	10.265%	9.972%

12.1 CALIBRATION: EXTENDED SKEW NORMAL MODEL

In this section we calibrate the model (11.3.1), (11.3.5), (11.3.6) from Section 11.3 to market data for two currency pairs. The use of just three options in calibration of volatility smiles leads to another typical (and which is in contrast to other markets) feature of the FX market that the volatility smile should interpolate the given three data points. Therefore, FX calibration is usually done via a root-finding numerical algorithm, while on other markets, where a large number of option prices are available for constructing volatility curves, one normally uses least-square type algorithms for this purpose.

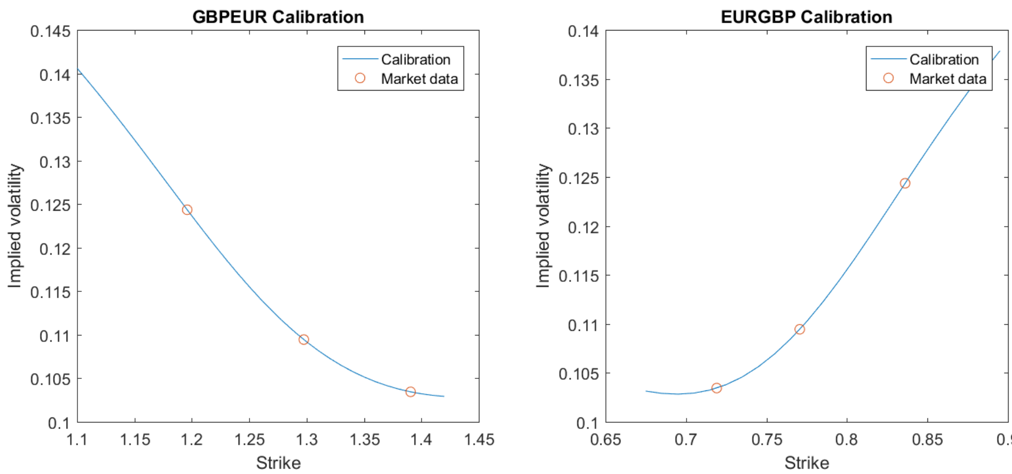

Figure 12.1.1: Calibration results for the GBP-EUR currency pair (left) and the inverse pair EUR-GBP (right) with $T = 1$, $r_{£} = 0.0025$, $r_{€} = 0.00$, $S_{£/€}(0) = 1.2935$.

Table 12.1.1: The results of calibration for GBP-EUR and EUR-GBP.

parameter	GBP-EUR/EUR-GBP
a	0.06297173
α_1	-3.18990817
α_2	1.57557895
β_1	-0.5
β_2	0.5

	GBP-EUR	EUR-GBP
skew	-0.87012308	0.87012308
kurtosis	4.94244079	4.94244079

The calibration was done in MATLAB R2016a, where we use the MATLAB function *fsolve* (which by default uses the built-in *trust-region-dogleg* algorithm) to match the option price data (three points per currency pair). We fixed the (free) parameters $\beta_1 = -0.5$ and $\beta_2 = 0.5$. For the calibration of the GBP-EUR pair, we use $a = \sigma_{ATM}, \alpha_1 = -3.0$ and $\alpha_2 = 1.0$ as initial values, as the negative skew of the volatility smile suggests a larger left tail (of the the distribution of Z). The calibration on a standard Desktop computer (Windows 7, 64-bit, Intel(R) Core(TM) i5-6500 CPU@3.20GHz, 16GB RAM) takes 0.11 seconds.

The calibration results for the GBP-EUR pair are given in Figure 12.1.1 and Table 12.1.1. One can see that the proposed pricing mechanism (see Theorem 10.1.6 and also (11.3.3)) together with the exchange rate model (11.3.1), (11.3.5), (11.3.6) preserves the volatility smile symmetry as skew, kurtosis (neglecting natural sign changes) and the model parameters stay the same. We also confirm that it is sufficient to calibrate the model using the GBP-EUR data and that the model reproduces both GBP-EUR and EUR-GBP smiles with the same parameters a, α_1, α_2, β_1, β_2. Moreover, it can been seen that the resulting skew of 0.870 and kurtosis of 4.942 indicate the difference of the resulting distribution Z to a normal distribution ($skew = 0$, $kurtosis = 3.0$).

The calibration results for the USD-EUR pair are given in Figure 12.1.2 and Table 12.1.2. The same observations as above for the GBP-EUR pair can be made here as well.

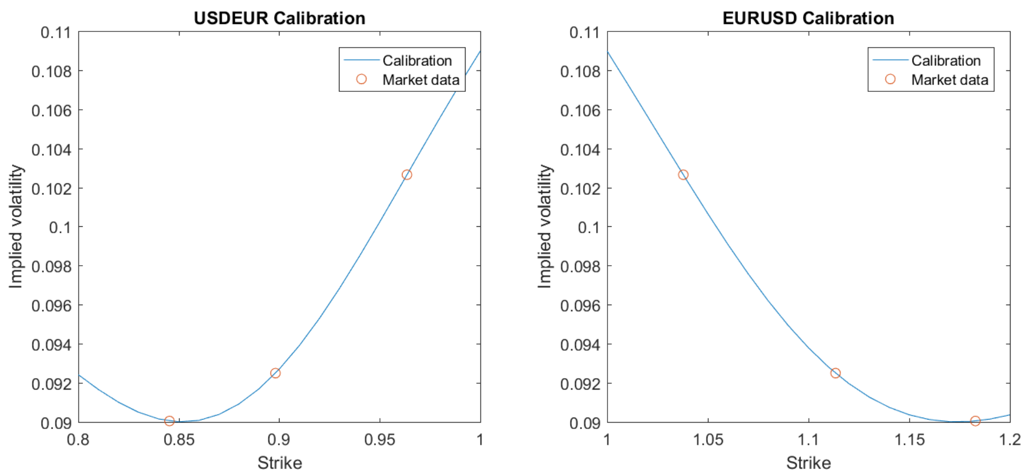

Figure 12.1.2: Calibration for the USD-EUR currency pair (left) and the inverse pair EUR-USD (right) with $T = 1.0$, $r_{\$} = 0.0025$, $r_{€} = 0.00$, $S_{\$/€}(0) = 0.8968$.

Table 12.1.2: The results of calibration for USD-EUR and EUR-USD.

parameter	USD-EUR/EUR-USD
a	0.05259980
α_1	−1.94011846
α_2	2.90433341
β_1	−0.5
β_2	0.5

	USD-EUR	EUR-USD
skew	0.53740761	−0.53740761
kurtosis	4.52666183	4.52666183

12.2 CALIBRATION: HESTON MODEL

In this section we calibrate (i.e., find a parameter set of $v_0, \kappa, \delta, \theta, \rho$) the Heston model (11.1.15) from Section 11.1 to market data for the GBP-EUR currency pairs as shown in Table 12.0.1. Similar to Section 12.1, we used a root-finding method. We fix the parameters for v_0 and κ at suitable levels. To compute option prices, we applied the Monte Carlo technique to the pricing formulas (11.1.13) and (11.1.14). To simulate the Heston model under Q^{X_T} (11.1.15), we used an Euler discretisation scheme for the log forward prices using a fixed time discretisation $h > 0$ and starting time $t = 0$, with the following initial values:

- $P_X(0,T) = \exp(-\frac{r_€ + r_£}{2} T + A(T) - C(T)v(0))$ (following (11.1.9)),
- $\widetilde{F}_{£/X}(0) = F_{£/X}(0,T) = e^{-r_£ T} \frac{\sqrt{f(0)}}{P_X(0,T)}$,
- $\widetilde{F}_{€/X}(0) = F_{€/X}(0,T) = e^{-r_€ T} \frac{1}{P_X(0,T)\sqrt{f(0)}}$,
- $\widetilde{v}(0) = v_0$.

The Euler approximation from a current point $(t, \widetilde{F}_{£/X}(t), \widetilde{F}_{€/X}(t), \widetilde{v}(t))$, can be written as follows:

$$\widetilde{F}_{£/X}(t+h) \approx \widetilde{F}_{£/X}(t) \exp\Bigg(\left[\frac{\widetilde{v}(t)}{8} + \delta^2 \frac{(C(T-(t+h)))^2}{2} \right] h \\ + \sqrt{\widetilde{v}(t)} \left[\delta C(T-(t+h))\eta + \frac{\sqrt{1-\rho^2}\xi + \rho\eta}{2} \right] \sqrt{h} \Bigg),$$

$$\widetilde{F}_{€/X}(t+h) \approx \widetilde{F}_{€/X}(t) \exp\Bigg(\left[\frac{\widetilde{v}(t)}{8} + \delta^2 \frac{(C(T-(t+h)))^2}{2} \right] h$$

$$+\sqrt{\widetilde{v}(t)}\left[\delta C(T-(t+h))\eta - \frac{\sqrt{1-\rho^2}\xi + \rho\eta}{2}\right]\sqrt{h}\Bigg),$$

where ξ and η are mutually independent standard normal random variables. We also use the moment-matching scheme for the volatility process [102], which preserves positivity of the volatility process and can be written as follows:

$$\widetilde{v}(t+h) = (\widetilde{\theta}_{t+h} + (\widetilde{v}(t) - \widetilde{\theta}_{t+h})e^{-h\widetilde{\kappa}_{t+h}})\exp\left(-\frac{1}{2}(\Gamma_{t+h})^2 + \Gamma_{t+h}\eta\right),$$

where

$$\begin{aligned}
\widetilde{\kappa}_{t+h} &= \kappa - C(T-(t+h))\delta^2, \\
\widetilde{\theta}_{t+h} &= \frac{\theta\kappa}{\kappa - C(T-(t+h))\delta^2}, \\
\Gamma_{t+h} &= \log\left(1 + \frac{\widetilde{v}(t)\delta^2(1-e^{-2h\widetilde{\kappa}_{t+h}})}{2\widetilde{\kappa}_{t+h}(\widetilde{\theta}_{t+h} + [\widetilde{v}(t) - \widetilde{\theta}_{t+h}]e^{-h\widetilde{\kappa}_{t+h}})^2}\right).
\end{aligned}$$

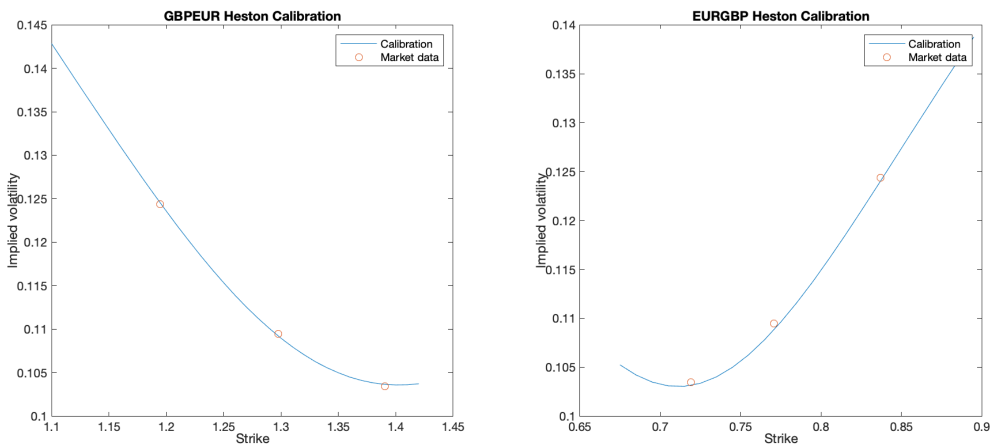

Figure 12.2.1: Heston model calibration results for the GBP-EUR currency pair (left) and the inverse pair EUR-GBP (right) with $T = 1$, $r_£ = 0.0025$, $r_€ = 0.00$, $S_{£/€}(0) = 1.2935$.

The calibration results for the GBP-EUR pair are given in Figure 12.2.1 and Table 12.2.1. Again, it can be seen that the proposed pricing mechanism (see Theorem 10.1.6) together with the Heston model (11.1.13) is flexible enough

Table 12.2.1: The results of Heston model calibration for GBP-EUR and EUR-GBP.

parameter	GBP-EUR/EUR-GBP
N_{MC}	10^9
h	0.05
v_0	0.0086
κ	1.500
δ	0.71020580946071
θ	0.02949445852250
ρ	−0.40966532579627

to match the GBP-EUR smile. Moreover, it is sufficient to calibrate the model to the GBP-EUR smile, which results in the inverse smile EUR-GBP smile to be automatically calibrated automatically (using the appropriate pricing formula (11.1.14)).

12.3 ILLUSTRATION OF THE MODEL-FREE APPROACH

In this section, we illustrate how we can approximate the scaled density function in the model-free approach of Section 11.4 from market data for three currencies. We recall that thanks to the intermediate currency approach we can use the same density function to price options on all three markets. We retrieve the scaled density by differentiating $U(K_1, K_2)$ twice:

$$\frac{\partial^2}{\partial K_1 \partial K_2} U(K_1, K_2) = \frac{e^{-r_{€} T}}{\mathbb{E}_{Q_T^X}\left[S_{€/X}(T)\right]} K_1^{-1/3} K_2{}^{-1/3}\, \rho(K_1, K_2; T).$$

We use the same market data as before, for the three currency pairs GBP-EUR, USD-EUR and GBP-USD, which can be found in Table 12.0.1. We can find the corresponding strikes by inverting the Garman-Kohlhagen formula for all three pairs. As we need the volatility smiles to satisfy the growth condition (11.4.15), we fit a 2nd order polynomial with the three

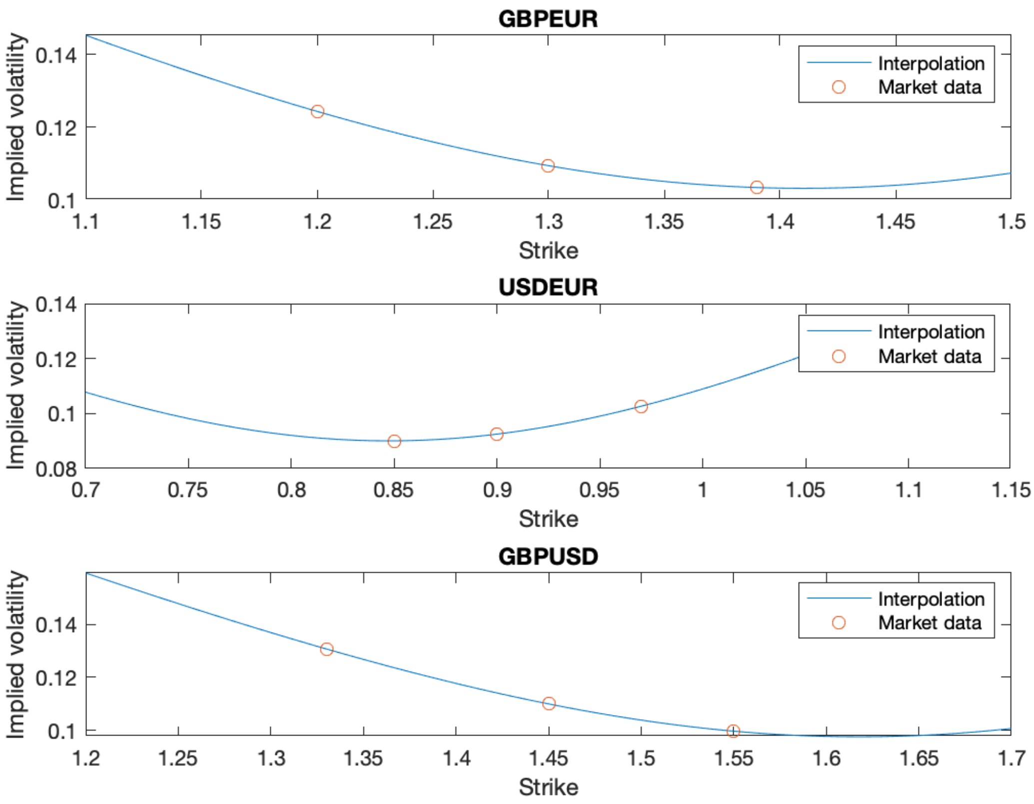

Figure 12.3.1: Implied volatility interpolation for GBP-EUR, USD-EUR and GBP-USD pairs with $T = 1.0$, $r_\$ = 0.0025$,$r_\$ = 0.0025$, $r_€ = 0.00$, $r_£ = 0.0025$, $r_£ = 0.0025$, $S_{\$/€}(0) = 0.8968$, $S_{£/€}(0) = 1.2935$, $S_{£/\$} = 1.4423$.

parameters $p_i^{(j)} \in \mathbb{R}, \quad j = 1, 2, 3$, to the implied volatility data transformed by $\exp\left[\sigma_i^2(K)\right]$. Then we obtain the interpolated implied volatilities as

$$\tilde{\sigma}_i(K) = \sqrt{\log\left[p_i^{(1)}K^2 + p_i^{(2)}K + p_i^{(3)}\right]}.$$

The results of the interpolation for the implied volatility smiles can be seen in Figure 12.3.1. The partial derivative with respect to K_1 and K_2 of $U(K_1, K_2)$ can be found by numerically differentiating (11.4.14) on a fine grid of K_1 and K_2. We use the MATLAB function *diff* to compute the point-wise $\frac{\partial^2}{\partial K_1 \partial K_2} U(K_1, K_2)$ surface for a range of strikes K_1 and K_2. Note that $K_3 = \frac{K_1}{K_2}$. The resulting surface and contour plots are given in Figure 12.3.2. We remark that $\frac{\partial^2}{\partial K_1 \partial K_2} U(K_1, K_2)$ is positive for the whole range of strikes considered as required.

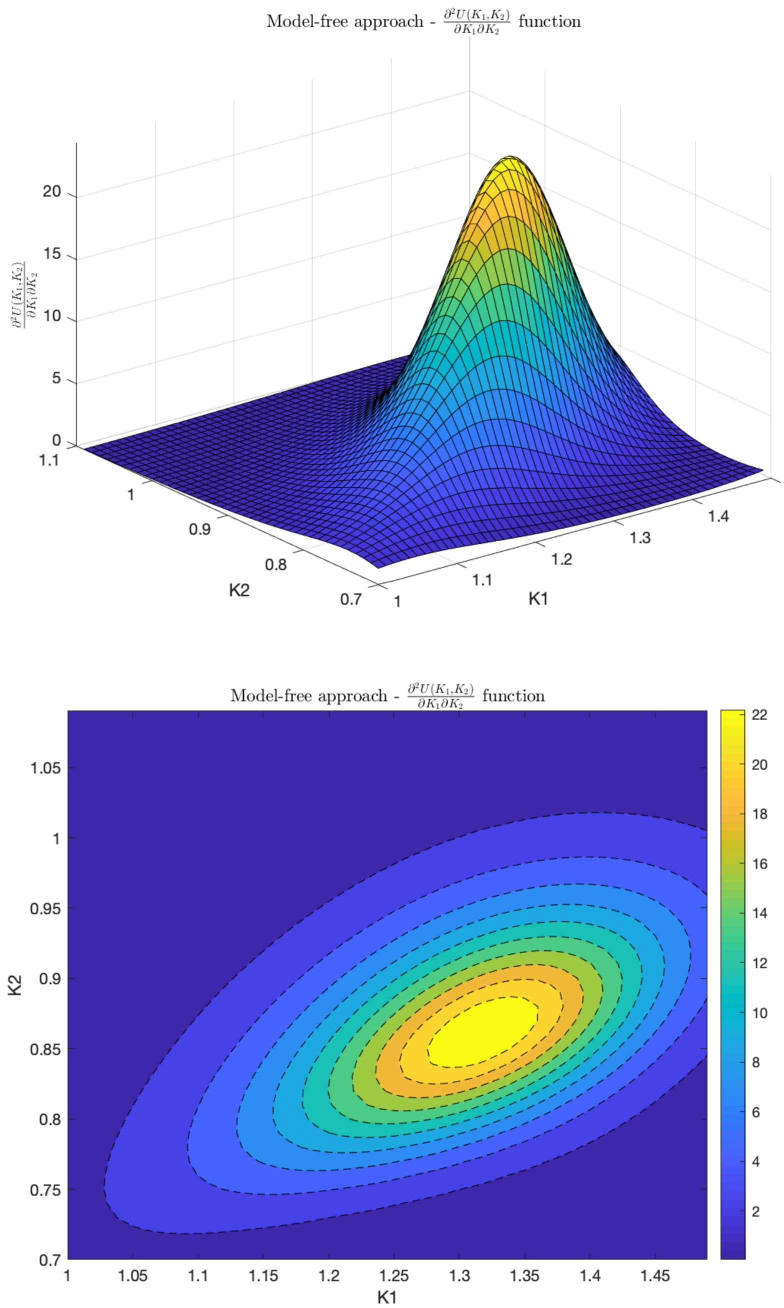

Figure 12.3.2: Implied scaled density surface (top) and contour plot (bottom) for the three currency pairs for a range of strikes K_1 and K_2.

13

CONCLUSIONS AND FUTURE WORK

In this part of the thesis, we have presented a novel idea to option pricing on the FX market. The main result is the introduction of a new framework of an intermediate currency market presented in Section 10, which can be seen as a solution to overcome the practical inconvenience of having multiple sets of parameters when working with foreign exchange rate models, see Example 9.4.1 as an illustration for that when working with a Heston model. In particular, in the multi-dimensional setting this is useful. We reviewed that there is no measure which is simultaneously risk-neutral for both the domestic and the foreign market and also stated the foreign-domestic symmetry (see Theorem 9.4.1) in our setting. While in existing models, there is a natural preference to a certain market due to market conventions or geographic location, current models implicitly have a bias towards a certain base currency or numeraire. The idea presented here, overcomes this inconsistency, which can be observed as the Foreign-Domestic symmetry holds under this new model idea, as shown in Section 10.1 and also when illustrated in different examples in Section 11. Moreover, we showed that an equivalent market measure (EMM) exists on this intermediate currency market (see 10.1.1) and derived pricing formulas under the EMM Q^X for a general class of scalable payoff functions $g(x;K)$ in Theorem 10.1.5 and further, pricing formulas under the forward measure Q_T^X in Theorem 10.1.6. These pricing formulas satisfy the foreign-domestic symmetry and are generally true.

Overall, we highlight the usefulness of this framework in Chapter 11, where we illustrate the framework applied to a range of different models: In Chapter 11.1, we apply the intermediate market idea for the case, where the underlying exchange rate is described by a Heston model and by a SABR model in Chapter 11.2. In both cases, we follow the classical option pricing route, where we start under a 'market' measure and then derive an EMM on the intermediate currency market. Moreover, we derive the resulting pricing formulas for European options. We present another illustrative example in Chapter 11.3, where we propose a distribution for an exchange rate at maturity time T under a forward measure on the intermediate currency market. Assuming a skew normal distribution and using the general pricing formulas derived in Theorem 10.1.6 and Theorem 10.2.4, we can derive closed-form pricing formulas. In Chapter 11.4, we follow a different approach, which we denote as 'model-free', in the sense that no assumptions on the form of the underlying exchange rate are made. The main idea of this approach is to construct a density function of risky assets under a risk-neutral measure using the observed prices of plain-vanilla options. We illustrate the idea in two and three dimensions and derive the corresponding useful pricing formulas.

As one of the aims of the suggested framework was to simplify calibration, i.e. finding the right model parameters to match market data quotes, we demonstrate the results for different models with numerical examples in Section 12. As we are trying to find the parameters using a root-finding method, we do not necessarily have unique solutions, as we have more parameters than data points. The results of the calibration of the extended skew normal model shown in Figure 12.1.1 and Figure 12.1.2, show that the model is flexible enough to model a range of different volatility smiles and skews. In Chapter 12.2, the same is true for the time-dependent Heston model approach, however, it is worth noting that as the pricing of each option is done using a Monte Carlo simulation, the calibration process is not very time-efficient and for time-sensitive tasks, it might be better to use other pricing approaches. It might be interesting in the future to improve the Monte Carlo technique used by using variance reduction methods such as antithetic variates or importance sampling, which might help with computational limitations. In the final numerical illustration in Chapter 12.3, we demonstrate how we can approximate the scaled density function in the model-free approach using market data for three currencies.

Overall, the numerical examples show that the suggested framework can be used in different scenarios. In the future, it would be very interesting to evaluate the performance of these pricing methods compared to other traditional models, especially in a multi-dimensional setting. In particular, it would be good to compare our achieved results to traditional Heston and SABR models and possibly some other machine learning techniques from more recent research projects. This might give a better understanding of the introduced framework in terms of model flexibility and also computational performance.

Part III

CONCLUSION

14

CONCLUSION

In this thesis, we have presented two different aspects in financial option pricing. In the first part, we have focused on the computational aspect and analysed a new numerical method for SDEs driven by Lévy processes with infinite activity. The introduced algorithm, a restricted jump-adaptive numerical scheme, for weak-sense approximations is a useful method to solve Dirichlet IPDE problems. It is worth highlighting, that we were able to replicate the theoretical convergence results in different numerical examples. In the second part, we have presented a novel framework for pricing derivatives on the FX market. We show that this framework can be used with different pricing models and that it can be useful in calibration tasks. In both instances, we have been able to demonstrate that the suggested approaches can be used in applications involved in Financial Mathematics.

In general, the presented results, especially from Part I, are not limited to this interesting research area, but more importantly can be used in a wide range of fields such as Physics [42] and Biology [77], where one can use the suggested algorithm to simulate the Lévy process dynamics. As long as it is possible to make the same underlying assumptions, it is possible to apply the suggested numerical and theoretical results.

While in Financial Mathematics, with constant development of new models and the advancement of technology, the importance of being able to describe real market observations has not changed. The research presented in this thesis offers practical approaches with a view to modelling and solving multi-dimensional problems under particular sets of model assumptions.

With a broader view towards the future of Financial Mathematics, it would be very interesting to compare the presented results with recent advancements in Machine Learning, where causality and interpretability are often not as easy compared to more classical modelling approaches as presented in this thesis.

Another interesting area of research, from a computational aspect, would be to compare the results achieved in this thesis to models used in financial institutions, such as investment banks and hedge funds. In particular, it would be interesting to compare model performance in terms of computational costs compared to accuracy but also volatility smile behaviour outside of the market data quotes given.

Additionally, another challenging direction of research, in particular with respect to Part II, would be to see if the models are in some way better in terms of option pricing and if such an advantage could be exploited systematically. The main focus of this research was the introduction of the intermediate currency market framework in the first place and it's possible application to existing pricing models.

BIBLIOGRAPHY

[1] F. Abergel and R. Tachet. A nonlinear partial integro-differential equation from mathematical finance. *Discrete and Continuous Dynamical Systems-Series A*, 27(3):907–917, 2010.

[2] L. J. S. Allen. *An introduction to stochastic processes with applications to biology*. CRC Press, 2010.

[3] L. Andersen and J. Andreasen. Jump-diffusion processes: Volatility smile fitting and numerical methods for option pricing. *Review of Derivatives Research*, 4(3):231–262, 2000.

[4] D. Applebaum. *Lévy processes and stochastic calculus*. Cambridge university press, 2009.

[5] S. Asmussen and J. Rosiński. Approximations of small jumps of Lévy processes with a view towards simulation. *Journal of Applied Probability*, 38(02):482–493, 2001.

[6] P. Austing. Repricing the cross smile: an analytic joint density. *Risk*, July:72–75, 2011.

[7] P. Austing. *Smile Pricing Explained*. Palgrave Macmillan, 2014.

[8] Bank of International Settlements. Triennial central bank survey of foreign exchange and derivatives market activity in 2019. Technical report, Sep 2019.

[9] O. E. Barndorff-Nielsen, T. Mikosch, and S. I. Resnick. *Lévy processes: theory and applications*. Springer Science & Business Media, 2012.

[10] E. Benhamou, E. Gobet, and M. Miri. Time dependent Heston model. *SIAM J. Finan. Math.*, 1:289–325, 2010.

[11] L. Bergomi. *Stochastic volatility modeling*. CRC press, 2015.

[12] J. Bertoin. *Lévy processes*, volume 121. Cambridge university press, 1998.

[13] D. Bertsimas, I. Popescu, and J. Sethuraman. Moment problems and semidefinite programming. *Handbook on semidefinite programming: Theory, Algorithms, and Applications*, pages 469–509, 2000.

[14] N. Biger and J. Hull. The valuation of currency options. *Financial Management*, pages 24–28, 1983.

[15] T. Bjork. *Arbitrage Theory in Continuous Time*. Oxford University Press, 2009.

[16] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–54, 1973.

[17] D. Brigo and F. Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*. Springer Science & Business Media, 2007.

[18] N. Bruti-Liberati and E. Platen. Strong approximations of stochastic differential equations with jumps. *Journal of Computational and Applied Mathematics*, 205(2):982–1001, 2007.

[19] J. C. Butcher. *Numerical methods for ordinary differential equations*. John Wiley & Sons, 2016.

[20] A. Castagna. *FX Options and Smile Risk*. Wiley, 2010.

[21] E. Çınlar. *Probability and stochastics*, volume 261. Springer Science & Business Media, 2011.

[22] I. Clark. *Foreign Exchange Option Pricing: A Practitioners Guide*. Wiley, 2011.

[23] R. Cont and P. Tankov. *Financial modelling with jump processes*. Chapman & Hall/CRC, Boca Raton, FL, 2004.

[24] R. Cont and E. Voltchkova. A finite difference scheme for option pricing in jump diffusion and exponential lévy models. *SIAM Journal on Numerical Analysis*, 43(4):1596–1626, 2005.

[25] Z. Dadachanji. *FX Barrier Options: A Comprehensive Guide for Industry Quants*. Springer, 2016.

[26] A. De Col, A. Gnoatto, and M. Grasselli. Smiles all around: FX joint calibration in a multi-Heston model. *J. Banking & Fin.*, 37(10):3799–3818, 2013.

[27] S. De Marco and C. Martini. Moment generating functions and normalized implied volatilities: unification and extension via Fukasawa's pricing formula. *Quant. Fin.*, 18:609–622, 2018.

[28] M. Dehghan and F. Shakeri. Solution of parabolic integro-differential equations arising in heat conduction in materials with memory via he's variational iteration technique. *International Journal for Numerical Methods in Biomedical Engineering*, 26(6):705–715, 2010.

[29] S. del Baño Rollin. Spot inversion in the Heston model, CRM Prepint 837, 2008.

[30] G. Deligiannidis, S. Maurer, and M. V. Tretyakov. Random walk algorithm for the Dirichlet problem for parabolic integro-differential equation. *BIT Numerical Mathematics*, April 2021. https://doi.org/10.1007/s10543-021-00863-2.

[31] D.F. DeRosa. *Options on Foreign Exchange (Wiley Finance)*. Wiley, 3 edition, 2011.

[32] L. Devroye. *Non-uniform random variate generation*. Springer, New York, 1986.

[33] P. Doust. The intrinsic currency valuation framework. *Risk*, pages 89–95, July 2008.

[34] P. Doust. The stochastic intrinsic currency volatility model: a consistent framework for multiple FX rates and their volatilities. *Appl. Math. Fin.*, 19(5):381–445, 2012.

[35] P. Doust and J. Chen. Estimating intrinsic currency values. *RISK*, pages 76–81, 2007.

[36] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.

[37] B. Dupire. Pricing with a smile. *Risk*, 7:18–20, 1994.

[38] E. Eberlein, A. Papapantoleon, and A. N. Shiryaev. On the duality principle in option pricing: semimartingale setting. *Finance Stoch.*, 12:265–292, 2008.

[39] D. Freedman. *Brownian motion and diffusion*. Springer Science & Business Media, 2012.

[40] M. Fukasawa. The normalizing transformation of the implied volatility smile. *Math. Fin.*, 22:753–762, 2012.

[41] R. G. Gallager. *Discrete stochastic processes*, volume 321. Springer Science & Business Media, 2012.

[42] P. Garbaczewski, J. Klauder, and R. Olkiewicz. Schrödinger problem, lévy processes, and noise in relativistic quantum mechanics. *Physical Review E*, 51:4114, 06 1995.

[43] C. W. Gardiner. *Stochastic methods*. Springer-Verlag, Berlin–Heidelberg–New York–Tokyo, 1985.

[44] M.B. Garman and S.W. Kohlhagen. Foreign currency option values. *J. International Money and Finance*, 2(3):231–237, 1983.

[45] M. G. Garroni and J.-L. Menaldi. *Green functions for second order parabolic integro-differential problems*, volume 275 of *Pitman Research Notes in Mathematics Series*. Longman Scientific & Technical, Harlow, 1992.

[46] J. Gatheral. *The volatility surface: a practitioner's guide*. John Wiley & Sons, 2011.

[47] H. Geman, N. El Karoui, and J.-C. Rochet. An exact bond option pricing formula. *J. Appl. Prob.*, 32(2):443–458, 1995.

[48] A. Gnoatto and M. Grasselli. An affine multicurrency model with stochastic volatility and stochastic interest rates. *SIAM J. Finan. Math.*, 5(1):493–531, 2014.

[49] J.O. Grabbe. The pricing of call and put options on foreign exchange. *J. Int. Money Financ.*, 2:239–253, 1983.

[50] C. Graham and D. Talay. *Stochastic simulation and Monte Carlo methods: mathematical foundations of stochastic simulation*, volume 68. Springer Science & Business Media, 2013.

[51] D. Griffiths and D. J. Higham. *Numerical methods for ordinary differential equations: initial value problems*. Springer Science & Business Media, 2010.

[52] G. Grimmett and D. Welsh. *Probability: An Introduction*. OUP Oxford, 2014.

[53] L.A. Grzelak and C.W. Oosterlee. On the Heston model with stochastic interest rates. *SIAM J. Financial Math.*, 2:255–286, 2011.

[54] P. S. Hagan, D. Kumar, A. S. Lesniewski, and D. E. Woodward. Managing smile risk. *The Best of Wilmott*, page 249, 2002.

[55] P. Hepperger. Hedging electricity swaptions using partial integro-differential equations. *Stochastic Processes and their Applications*, 122(2):600–622, 2012.

[56] S. L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Finan. Studies*, 6(2):327–343, 1993.

[57] J. C. Hull. *Options, Futures, and Other Derivatives*. Prentice Hall, 8 edition, February 2011.

[58] P. Jäckel. Let's be rational. *Wilmott*, 2015(75):40–53, 2015.

[59] N. Jacob. *Pseudo differential operators and Markov processes. Vol. III*. Imperial College Press, London, 2005.

[60] J. Jacod, T. G Kurtz, S. Méléard, and P. Protter. The approximate euler method for lévy driven stochastic differential equations. 41(3):523–558, 2005.

[61] F. Jamshidian. An exact bond option pricing formula. *J. Finance*, 44(1):205–209, 1989.

[62] A. Janek, T. Kluge, R. Weron, and U. Wystup. FX Smile in the Heston Model. MPRA Paper 25491, University Library of Munich, Germany, September 2010.

[63] A. Janek, T. Kluge, R. Weron, and U. Wystup. Fx smile in the heston model. Mpra paper, University Library of Munich, Germany, 2010.

[64] O. Kallenberg. *Foundations of modern probability*. Springer Science & Business Media, 2006.

[65] I. Karatzas and S. E. Shreve. Brownian motion and stochastic calculus. 1998.

[66] K. Kashima and R. Kawai. An optimization approach to weak approximation of stochastic differential equations with jumps. *Applied numerical mathematics*, 61(5):641–650, 2011.

[67] K. Kashima and R. Kawai. A weak approximation of stochastic differential equations with jumps through tempered polynomial optimization. *Stochastic models*, 27(1):26–49, 2011.

[68] K. Kashima and R. Kawai. On weak approximation of stochastic differential equations through hard bounds by mathematical programming. *SIAM Journal on Scientific Computing*, 35(1):A.1–A.21, 2013.

[69] P.E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Applications of Mathematics. Springer-Verlag, 1992.

[70] A. Kohatsu-Higa and H.-L. Ngo. Weak approximations for sde's driven by lévy processes. In *Seminar on Stochastic Analysis, Random Fields and Applications VII*, pages 131–169. Springer, 2013.

[71] A. Kohatsu-Higa, S. Ortiz-Latorre, and P. Tankov. Optimal simulation schemes for Lévy driven stochastic differential equations. *Mathematics of Computation*, 83(289):2293–2324, 2014.

[72] A. Kohatsu-Higa and P. Tankov. Jump-adapted discretization schemes for Lévy-driven SDEs. *Stochastic Process. Appl.*, 120(11):2258–2285, 2010.

[73] S. Kotz. *Continuous multivariate distributions.* Wiley series in probability and statistics. Applied probability and statistics. Wiley, New York ; Chichester, 2nd ed. edition, 2000.

[74] O. Kudryavtsev and S. Levendorskiĭ. Fast and accurate pricing of barrier options under lévy processes. *Finance and Stochastics*, 13(4):531–562, 2009.

[75] J. D. Lambert. Computational methods in ordinary differential equations. 1973.

[76] D. Lamberton and B. Lapeyre. *Introduction to stochastic calculus applied to finance*. CRC press, 2007.

[77] M. Landis, J. Schraiber, and M. Liang. Phylogenetic analysis using levy processes: Finding jumps in the evolution of continuous traits. *Systematic biology*, 62, 10 2012.

[78] R. Lee. The moment formula for implied volatility at extreme strikes. *Math. Fin.*, 14:469–480, 2004.

[79] B. Leimkuhler, A. Sharma, and M. V. Tretyakov. Simplest random walk for approximating Robin boundary value problems and ergodic limits of reflected diffusions. *arXiv: 2006.15670*, 2020.

[80] Sergei Levendorskii, Oleg Kudryavtsev, and Vadim Zherder. The relative efficiency of numerical methods for pricing american options under lévy processes. *Journal of Computational Finance*, 9(2):69, 2005.

[81] A. Lipton. *Mathematical Methods for Foreign Exchange: a Financial Engineer's Approach*. WSPC, 2001.

[82] X. Q. Liu and C. W. Li. Weak approximation and extrapolations of stochastic differential equations with jumps. *SIAM J. Numer. Anal.*, 37(6):1747–1767, 2000.

[83] W. Margrabe. The value of an option to exchange one asset for another. *J. Finance*, 33:177–186, 1978.

[84] A.-M. Matache, T. Von Petersdorff, and C. Schwab. Fast deterministic pricing of options on lévy driven assets. *ESAIM: Mathematical Modelling and Numerical Analysis*, 38(1):37–71, 2004.

[85] S. Maurer. Extension of the Skew Normal Distribution and Applications in Financial Options Pricing. *MSc. Dissertation in Financial and Computational Mathematics, University of Nottingham*, 2016.

[86] S. Maurer, T. E. Sharp, and M. V. Tretyakov. Pricing FX Options under Intermediate Currency. https://arxiv.org/abs/1912.01387, 2021.

[87] R.C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144, 1976.

[88] R. Mikulevicius and E. Platen. Time discrete Taylor approximations for Ito processes with jump component. *Math. Nachr.*, 138(6):93–104, 1988.

[89] R. Mikulevicius and H. Pragarauskas. On Cauchy-Dirichlet problem in half-space for linear integro-differential equations in weighted Hölder spaces. *Electron. J. Probab.*, 10:1398–1416, 2005.

[90] R. Mikulevicius and C. Zhang. Weak Euler scheme for Levy-driven stochastic differential equations. *Theory Probab. Applic.*, 63:346–366, 11 2018.

[91] G.N. Milstein and M.V. Tretyakov. The simplest random walks for the Dirichlet problem. *Theory Probab. Applic.*, 47(1):53–68, 2002.

[92] G.N. Milstein and M.V. Tretyakov. *Stochastic numerics for mathematical physics*. Springer, Berlin, 2004.

[93] E. Mordecki, A. Szepessy, R. Tempone, and G. E. Zouraris. Adaptive weak approximation of diffusions with jumps. *SIAM Journal on Numerical Analysis*, 46(4):1732–1768, 2008.

[94] K. Oshima, J. Teichmann, D. Veluščček, et al. A new extrapolation method for weak approximation schemes with applications. *The Annals of Applied Probability*, 22(3):1008–1045, 2012.

[95] C.V. Pao. Solution of a nonlinear integrodifferential system arising in nuclear reactor dynamics. *Journal of Mathematical Analysis and Applications*, 48(2):470–492, 1974.

[96] J. K. Patel and C. B. Read. *Handbook of the normal distribution*. Marcel Dekker Inc, 1996.

[97] E. Platen and N. Bruti-Liberati. *Numerical solution of stochastic differential equations with jumps in finance*, volume 64. Springer Science & Business Media, 2010.

[98] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and Brian P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, 3 edition, September 2007.

[99] J.A Primbs. Optimization based option pricing bounds via piecewise polynomial super- and sub-martingales. In *2008 American Control Conference*, pages 363–368. IEEE, 2008.

[100] P. Protter and D. Talay. The Euler scheme for Lévy driven stochastic differential equations. *The Annals of Probability*, 25(1):393–423, 1997.

[101] D. Reiswich and U. Wystup. FX volatility smile construction. *Wilmott*, 2012(60):58–69, 2012.

[102] F. D. Rouah. *The Heston Model and its Extensions in Matlab and C*. John Wiley & Sons, 2013.

[103] S. Rubenthaler. Numerical simulation of the solution of a stochastic differential equation driven by a lévy process. *Stochastic processes and their applications*, 103(2):311–349, 2003.

[104] E.W. Sachs and A.K. Strauss. Efficient solution of a partial integro-differential equation in finance. *Applied Numerical Mathematics*, 58(11):1687–1703, 2008.

[105] K. Sato. *Lévy processes and infinitely divisible distributions*. Cambridge University Press, 1999.

[106] R.L. Schilling and L. Partzsch. *Brownian motion: an introduction to stochastic processes*. Walter de Gruyter GmbH & Co KG, 2014.

[107] S. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance, 1st edition, June 2004.

[108] J. Siegel. Risk, interest rates and the forward exchange. *Quart. J. Econ.*, 86(2):303–309, 1972.

[109] R. Stulz. Options on the minimum or maximum of two assets: Analysis and applications. *J. Fin. Econ.*, 10:161–185, 1982.

[110] P. Tankov. High order weak approximation schemes for lévy-driven sdes. In *Monte Carlo and Quasi-Monte Carlo Methods 2010*, pages 667–683. Springer, 2012.

[111] N. van Kampen. *Stochastic processes in physics and chemistry*. 1992, 1995.

[112] P. Wilmott. *Paul Wilmott Introduces Quantitative Finance*. Wiley-Interscience, New York, NY, USA, 2 edition, 2007.

[113] U. Wystup. *FX Options and Structured Products*. Wiley, 2007.

[114] X. Zhang. Valuation of american options in a jump-diffusion model. *Numerical Methods in Finance, Cambridge University Press, Cambridge, UK*, pages 93–114, 1997.

Part IV

APPENDIX

A

APPENDIX A

A.1 PROOFS AND OTHER DERIVATIONS

A.1.1 *Proof form of characteristic function of Lévy process*

In Remark 3.2.6 we give the form the characteristic function of a Lévy process $(X_t)_{t\geq 0}$ on $\mathbb{R}^d$ should have the following form:

$$\Phi_{X_t}(z) = e^{t\Psi(z)}, \quad z \in \mathbb{R}^d,$$

where $\Psi(z)$ is the characteristic exponent of $X_1 = X(1)$.

Proof. Suppose that X is a Lévy process. Due to its independent increments we get that the characteristic function of a X has to be a multiplicative function for $s \geq 0$:

$$\begin{aligned}
\Phi_{X_{t+s}}(z) &= \mathbb{E}\left(e^{i\langle z, X_{t+s}\rangle}\right) \\
&= \mathbb{E}\left(e^{i\langle z, X_{t+s}-X_s\rangle} e^{i\langle z, X_s\rangle}\right) \\
&= \mathbb{E}\left(e^{i\langle z, X_{t+s}-X_s\rangle}\right)\mathbb{E}\left(e^{i\langle z, X_s\rangle}\right) \\
&= \Phi_{X_t}(z)\Phi_{X_s}(z).
\end{aligned}$$

Given this and the fact that $\Phi_{X_0}(z) = 1$ and the fact that $t \to \Phi_{X_t}(z)$ is continuous (see Lemma 1.3.2 in [4]), $\Phi_{X_t}(z)$ has to be an exponential function. □

A.1.2 *Derivation of characteristic function of a compound Poisson process*

In Example 3.3.3 we claim hat the characteristic exponent of J_t has the following form:

$$\Psi_J(z) = \lambda \int_{\mathbb{R}^d} (e^{i\langle z,y\rangle} - 1)F(dy), \quad z \in \mathbb{R}^d.$$

This can be easiest seen by deriving the characteristic function for J_t, where we denote the common CF for all Y_i, $i = 1, 2, \dots,$ by Φ_Y.

$$\begin{aligned}
\Phi_{J_t}(z) &= \mathbb{E}\left[e^{i(z,J_t)}\right] = \mathbb{E}\left[\exp i(z, \sum_{i=1}^{N_t} Y_i)\right] = \mathbb{E}\left[\mathbb{E}\left[\exp i(z, \sum_{i=1}^{N_t} Y_i) \mid N_t\right]\right] \\
&= \sum_{i=0}^{\infty} \mathbb{E}\left[\exp i(z, \sum_{i=1}^{N_t} Y_i) \mid N_t = n\right] P(N_t = n) \\
&= \sum_{i=0}^{\infty} \mathbb{E}\left[\exp i(z, \sum_{i=1}^{n} Y_i)\right] e^{-\lambda t}\frac{(\lambda t)^n}{n!} = e^{-\lambda t}\sum_{i=0}^{\infty} \Phi_Y^n \frac{(\lambda t)^n}{n!} = e^{t\lambda(\Phi_Y - 1)} \\
&= \exp\left(t\lambda \int_{\mathbb{R}^d} (e^{i\langle z,y\rangle} - 1)F(dy)\right).
\end{aligned}$$

The stated result for the characteristic exponent then easily follows.

B

APPENDIX B

B.1 NORMAL AND BIVARIATE DISTRIBUTIONS AND RELATED MGFS

We give a short overview over normal and bivariate normal distributions, followed by a broad review of moment generating functions and some of its useful properties based on [96].

Definition B.1.1 (Normal distribution)**.** A random variable Y is said to have a **normal distribution** with mean μ and standard deviation σ, if its probability density function (pdf) can be written as

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y-\mu)^2}{2\sigma^2}},$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$.
In the case that $\mu = 0$ and $\sigma = 1$ we can say that distribution of Y is **standard normal** and its pdf can be written as

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}. \tag{B.1.1}$$

For an integral of a standard normal distributed random variable Y we will use the following notation for the cumulative density function (CDF)

$$N(a) := \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy. \tag{B.1.2}$$

Definition B.1.2 (Bivariate normal distribution)**.** Let us assume that X_1 and X_2 are both normally distributed with means μ_1 and μ_2 and standard deviations σ_1 and σ_2, respectively. Further, X_1 and X_2 are correlated with a correlation coefficient ρ. We can then say that the joint distribution of X_1 and X_2 is **bivariate normal** and their pdf can be written as

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right]},$$

where $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1, \sigma_2 > 0$ and $\rho \in (-1, 1)$.

In the case that $\mu_1 = \mu_2 = 0$ and $\sigma_1 = \sigma_2 = 1$ we can say that the joint distribution of X_1 and X_2 is **standard bivariate normal** with correlation coefficient ρ and the pdf can be written as

$$f_{X_1,X_2}(x_1,x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x_1^2+x_2^2-2\rho x_1 x_2}{2(1-\rho^2)}}. \tag{B.1.3}$$

For an integral of two standard bivariate normal distributed random variables X_1 and X_2 with correlation coefficient ρ we will use the following notation

$$N_2(a,b;\rho) := \int_{-\infty}^{a} \int_{-\infty}^{b} \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x_1^2+x_2^2-2\rho x_1 x_2}{2(1-\rho^2)}} dx_2 dx_1.$$

Definition B.1.3 (Moments)**.** The k^{th} **raw moment** of a random variable Y with pdf f_Y is defined as

$$\mathbb{E}[Y^k] = \int_{-\infty}^{\infty} y^k f_Y(y) dy, \tag{B.1.4}$$

if $\mathbb{E}[Y^k] < \infty$. The k^{th} **central moment** of random variable Y with pdf f_Y and mean μ is defined as

$$\mathbb{E}[(Y-\mu)^k] = \int_{-\infty}^{\infty} (y-\mu)^k f_Y(y)dy, \tag{B.1.5}$$

if $\mathbb{E}[(Y-\mu)^k] < \infty$. The k^{th} **standardised moment** of random variable Y with pdf f_Y, mean μ and variance σ^2 is defined as

$$\frac{\mathbb{E}[(Y-\mu)^k]}{\sigma^k} = \int_{-\infty}^{\infty} \frac{(y-\mu)}{\sigma^k}^k f_Y(y)dy, \tag{B.1.6}$$

if $\frac{\mathbb{E}[(Y-\mu)^k]}{\sigma^k} < \infty$.

For $k = 1$ in (B.1.4), $k = 2$ in (B.1.5) and $k = 3, 4$ in (B.1.6) we usually say that

$$\begin{aligned}
\mathbb{E}[Y] = \mu &= \text{mean of Y}, \\
\mathbb{E}[(Y-\mu)^2)] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \sigma^2 &= \text{variance of Y}, \\
\frac{\mathbb{E}[(Y-\mu)^3]}{\sigma^3} = \frac{\mathbb{E}[Y^3] - 3\mu\mathbb{E}[Y^2] + 2\mu^2}{\sigma^3} &= \text{skewness of Y},
\end{aligned} \tag{B.1.7}$$

$$\frac{\mathbb{E}[(Y-\mu)^4]}{\sigma^4} = \frac{\mathbb{E}[Y^4] - 4\mu\mathbb{E}[Y^3] + 6\mu^2\mathbb{E}[Y^2] - 3\mu^4}{\sigma^4} = \text{curtosis of Y}, \tag{B.1.8}$$

which are useful measures of distributions and can easily be computed via moments.

Definition B.1.4 (Moment generating function (mgf))**.** The **moment generating function** of a random variable Y is defined (assuming $E[e^{tY}]$ exists) as

$$M_Y(t) := \mathbb{E}[e^{tY}], \quad t \in \mathbb{R}.$$

For example, the mgf of a normal distributed random variable X with mean μ and standard deviation σ is given by

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2},$$

and hence for a standard normal distributed random variable Y (with $\mu = 0, \sigma = 1$)

$$M_Y(t) = e^{\frac{1}{2}t^2}.$$

Mgfs are in particular useful to calculate moments of random variables:

$$\left.\frac{d^k M_Y(t)}{dt^k}\right|_{t=0} = \mathbb{E}[Y^k].$$

Let us give further properties of mgfs which we will use in this dissertation.

Theorem B.1.5. ***(Convolution Theorem)***
The random variables $Y_1, Y_2, ..., Y_n$ are independent if and only if the moment generating function of $Y_1 + Y_2 + ... + Y_n$ is given by:

$$M_{Y_1+Y_2+...+Y_n}(t) = M_{Y_1}(t) \cdot M_{Y_2}(t) \cdots M_{Y_n}(t)$$

Proof. We state this theorem without proof, which can be found in Chapter 7.4 of [52]. □

Further we will need the shifting and the scaling property for mgfs, which can be seen as follows.
Multiplication by a constant factor:

$$\begin{aligned} M_{aY}(t) = \mathbb{E}[e^{taY}] &= \int_{-\infty}^{\infty} e^{taY} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2 - 2atY + (at)^2}{2}} e^{\frac{(at)^2}{2}} dy \\ &= e^{\frac{(at)^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-at)^2}{2}} dy \\ &= e^{\frac{(at)^2}{2}} = M_Y(at). \end{aligned} \tag{B.1.9}$$

Shifting by a constant:

$$\begin{aligned} M_{Y+b}(t) = \mathbb{E}[e^{t(Y+b)}] &= \int_{-\infty}^{\infty} e^{ty+bt} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\ &= e^{bt} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2-2ty+t^2}{2}} e^{\frac{t^2}{2}} dy \end{aligned} \tag{B.1.10}$$

$$
\begin{aligned}
&= e^{bt} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-t)^2}{2}} e^{\frac{t^2}{2}} dy \\
&= e^{bt} e^{\frac{t^2}{2}} \\
&= e^{bt} M_Y(t).
\end{aligned}
$$

B.2 PROOF OF PROPOSITION 9.2.2

For a domestic investor whose domestic currency is USD, the extended market can be described as follows:

$$
\begin{aligned}
dB_{\$} &= r_{\$} B_{\$} dt, \\
dB_{€} &= r_{€} B_{€} dt, \\
dS &= \mu S dt + \sigma S dW,
\end{aligned}
$$

where $B_{\$}, B_{€}$ and $r_{\$}, r_{€}$ are a domestic and a foreign currency bank account with its interest rates respectively. Then the risky asset for this investor is the foreign currency (EUR) in domestic currency (USD), hence

$$
Y(t) = S(t) B_{€}(t).
$$

For risk-free pricing, we need to find an Equivalent Martingale Measure (EMM) $Q^{\$}$, under which the discounted $Y(t)$ should be a martingale.

$$
\begin{aligned}
\widetilde{Y}(t) &= \frac{S(t) B_{€}(t)}{B_{\$}(t)}, \\
d\widetilde{Y} &= (\mu + r_{€} - r_{\$}) \widetilde{Y} dt + \sigma \widetilde{Y} dW, \\
\gamma(t) &= \frac{\mu + r_{€} - r_{\$}}{\sigma} \text{ and } dW^{Q^{\$}} = dW + \gamma(t) dt, \\
d\widetilde{Y} &= \sigma \widetilde{Y} dW^{Q^{\$}}, \\
dS &= (r_{\$} - r_{€}) S dt + \sigma S dW^{Q^{\$}}.
\end{aligned}
$$

We can now derive the risk-free option price for a domestic investor.

$$\begin{aligned} C_{\$}(0,T) &= v(S,T) = e^{-r_{\$}T}\mathbb{E}_{Q^{\$}}\left[(S(T)-K)_+\right] \\ &= \mathbb{E}_{Q^{\$}}\left[(e^{-r_{\$}T}S(T) - e^{-r_{\$}T}K)_+\right]. \end{aligned}$$

This expectation is only positive if$(\cdot)_+ > 0$, hence

$$\begin{aligned} & e^{-r_{\$}T}S(T) - e^{-r_{\$}T}K > 0 \\ \Leftrightarrow \quad & S(0)e^{(-r_{€}-\frac{\sigma^2}{2})T-\sigma\sqrt{T}\xi} > e^{-r_{\$}T}K \\ \Leftrightarrow \quad & \xi < \frac{\log\left(\frac{S(0)}{K}\right) + (r_{\$} - r_{€} - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}} := b, \text{ where } \xi \sim \mathcal{N}(0,1). \end{aligned}$$

$$\begin{aligned} C_{\$}(0,T) &= e^{-r_{€}T}S(0)\int_{\infty}^{b}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2+2\sigma\sqrt{T}z+\sigma^2T}{2}}dz \\ &\quad - e^{-r_{\$}T}K\int_{\infty}^{b}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz \\ &= e^{-r_{€}T}S(0)\int_{\infty}^{b+\sigma\sqrt{T}}\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}dz - e^{-r_{\$}T}K\cdot N(b) \\ &= e^{-r_{€}T}S(0)\cdot N\left(\frac{\log\left(\frac{S(0)}{K}\right)+(r_{\$}-r_{€}+\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right) \\ &\quad - e^{-r_{\$}T}K\cdot N\left(\frac{\log\left(\frac{S(0)}{K}\right)+(r_{\$}-r_{€}-\frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right). \end{aligned}$$

We can now extend to a general version for the price at time t instead of 0. The same approach can be used to derived the put option price.

B.3 PROOF FOR THEOREM 9.4.1

On the FX market there exists another important relationship between calls and puts, known as **foreign domestic symmetry**. This section is based on [81]. The idea is that on the FX market a call on one currency pair, e.g. EURUSD, is the same as a put on the inverse pair, e.g. USDEUR. With the right scaling, both options should have the same value, otherwise this would allow arbitrage. Further we make the assumptions, that there are no transaction or other costs. Therefore, we show that the following equation holds as stated in Theorem 9.4.1:

$$C_{€/\$}(0,T,K) = S_{€/\$}(0)\,K\,P_{\$/€}\left(0,T,\frac{1}{K}\right),$$

where $C_{€/\$}(0,T,K)$ is the call option price (in \$) at time 0 to buy one EUR for \$$K$ at time T; $P_{\$/€}(0,T,1/K)$ is the put option price (in €) at time 0 to sell one USD for €$\frac{1}{K}$ at time T.

Let us assume we are looking at a call option on $S_{€/\$}(T)$ at time T. The owner of this option has the right to buy €1 for the exchange rate $K\,\frac{USD}{EUR}$. Let us denote the value of this option

$$C_{€/\$}(0,T,S_{€/\$}(T),K,r_{\$},r_{€}).$$

On the other hand this means, he also has the right to sell \$$K$ for the exchange rate $\frac{1}{K}\frac{EUR}{USD}$. A put option with the right to sell \$1, is a put option on $S_{\$/€}(T) = \frac{1}{S_{€/\$}(T)}$ and has the value

$$P_{\$/€}\left(0,T,\frac{1}{S_{€/\$}(T)},\frac{1}{K},r_{€},r_{\$}\right).$$

Note that we take the view of a foreign investor to denote the price of this put option, which means that $r_{€}$ denotes the domestic interest rate and $r_{\$}$ the foreign interest rate respectively.

We need to scale the put option so that both investment strategies give the investor the same rights. K put options have the value

$$K\,P_{\$/€}\left(0,T,\frac{1}{S_{€/\$}(T)},\frac{1}{K},r_{€},r_{\$}\right),$$

expressed in Euro and the owner has the right to sell \$$K$ for the exchange rate $1\frac{EUR}{USD}$.

As mentioned above, since the investor has the same right in both cases, they have to have the same value in the same currency to avoid arbitrage, which leads to the following relationship.

$$C_{€/\$}(0,T,S_{€/\$}(T),K,r_\$,r_€) = S_{€/\$}(0)\,K \times P_{\$/€}\left(0,T,\frac{1}{S_{€/\$}(T)}\frac{1}{K},r_€,r_\$\right).$$

Note, that an alternative proof involving change of measure can be found in [107][Ch 9.3]. It also emphasizes the importance of the relevant measure when pricing currency options.

B.4 PROOF OF THEOREM 10.2.1 FOR GENERAL α

As mentioned in Remark 10.2.2, it is possible to choose to use a general $\alpha_1 = \cdots = \alpha_{N-1}$ in (10.2.2). We therefore repeat the proof of Theorem 10.2.1 for arbitrary $0 < \alpha_j < 1$:

Theorem B.4.1. *Assume that $N-1$ exchange rates f_j between the currency c_N to all other currencies c_i, $i = 1,\ldots,N-1$, under a 'market' measure are described by the model (10.2.6) together with (10.2.5). Consider the intermediate currency X introduced in (10.2.2), with arbitrary α_i. There is the unique intermediate currency interest rate $r_X(t)$ defined by*

$$r_X = \left(1 - \sum_{j=1}^{N-1}\alpha_j\right) r_N + \sum_{j=1}^{N-1}\alpha_j r_j + \sum_{j=1}^{N-1}\frac{\alpha_j(1-\alpha_j)}{2}\sigma_j^2 - \sum_{j=1}^{N-1}\sum_{k=1}^{j-1}\alpha_j\alpha_k\sigma_j\sigma_k\rho_{jk} \tag{B.4.1}$$

and there is an EMM Q^X for the intermediate pseudo-currency market.

Proof. Applying the Ito formula to (10.2.2), we obtain the SDEs for the exchange rates $S_{c_i/X}$:

$$\frac{dS_{c_i/X}}{S_{c_i/X}} = \left[\sum_{j=1}^{N-1}\left(-\alpha_j\mu_j + \frac{\alpha_j(\alpha_j+1)}{2}\sigma_j^2\right) + (\mu_i - \alpha_i\sigma_i^2)\mathbb{1}_{i\neq N}\right.$$

$$
\begin{aligned}
&+ \sum_{j=1}^{N-1} \sum_{k=1}^{j-1} \alpha_j \alpha_k \sigma_j \sigma_k \rho_{jk} + \left(\alpha_i \sigma_i^2 - \sigma_i \sum_{j=1}^{N-1} \alpha_j \sigma_j \rho_{ij} \right) \mathbb{1}_{i \neq N} \Bigg] dt \\
&- \sum_{j=1}^{N-1} \alpha_j \sigma_j \sum_{k=1}^{j} L_{jk} dW_k + \mathbb{1}_{i \neq N} \sigma_i \sum_{k=1}^{i} L_{ik} dW_k \\
= &\Bigg[\sum_{j=1}^{N-1} \Bigg(- \alpha_j \mu_j + \frac{\alpha_j(\alpha_j + 1)}{2} \sigma_j^2 + \sum_{k=1}^{j-1} \alpha_j \alpha_k \sigma_j \sigma_k \rho_{jk} \\
&- \sigma_i \mathbb{1}_{i \neq N} \alpha_j \sigma_j \rho_{ij} \Bigg) + \mu_i \mathbb{1}_{i \neq N} \Bigg] dt \\
&- \sum_{j=1}^{N-1} \sum_{k=1}^{j} \alpha_j \sigma_j L_{jk} dW_k + \mathbb{1}_{i \neq N} \sigma_i \sum_{k=1}^{i} L_{ik} dW_k, \quad i = 1, \ldots, N.
\end{aligned}
$$

On the considered market the risky assets have the prices $Y_{c_i/X} = S_{c_i/X} B_i$, $i = 1, \ldots, N$. Let us introduce the discounted risky assets' prices in the usual way:

$$
\widetilde{Y}_{c_i/X}(t) = \frac{S_{c_i/X}(t) B_i(t)}{B_X(t)}, \quad i = 1, \ldots, N.
$$

The discounted prices satisfy the SDEs

$$
\begin{aligned}
\frac{d\widetilde{Y}_{c_i/X}}{\widetilde{Y}_{c_i/X}} = &[r_i - r_X] dt \\
&+ \Bigg[\sum_{j=1}^{N-1} \Bigg(- \alpha_j \mu_j + \frac{\alpha_j(\alpha_j + 1)}{2} \sigma_j^2 + \sum_{k=1}^{j-1} \alpha_j \alpha_k \sigma_j \sigma_k \rho_{jk} \\
&- \sigma_i \mathbb{1}_{i \neq N} \alpha_j \sigma_j \rho_{ij} \Bigg) + \mu_i \mathbb{1}_{i \neq N} \Bigg] dt \\
&- \sum_{j=1}^{N-1} \sum_{k=1}^{j} \alpha_j \sigma_j L_{jk} dW_k + \mathbb{1}_{i \neq N} \sigma_i \sum_{k=1}^{i} L_{ik} dW_k, \quad i = 1, \ldots, N.
\end{aligned}
$$

The no-arbitrage condition requires existence of an EMM Q^X under which all $\widetilde{Y}_{c_i/X}$ are martingales. This implies that for Q^X to exist the following system of N simultaneous linear algebraic equations in N unknown variables (which are the market prices of risk γ_k, $k = 1, \ldots, N-1$, and r_X) should have a solution:

$$r_i - r_X + \sum_{j=1}^{N-1} \left(-\alpha_j \mu_j + \frac{\alpha_j(\alpha_j+1)}{2}\sigma_j^2 + \sum_{k=1}^{j-1} \alpha_j \alpha_k \sigma_j \sigma_k \rho_{jk} - \sigma_i \mathbb{1}_{i \neq N} \alpha_j \sigma_j \rho_{ij} \right)$$
$$+ \mu_i \mathbb{1}_{i \neq N} \tag{B.4.2}$$
$$= -\sum_{j=1}^{N-1} \sum_{k=1}^{j} \alpha_j \sigma_j L_{jk} \gamma_k + \mathbb{1}_{i \neq N} \sigma_i \sum_{k=1}^{i} L_{ik} \gamma_k, \quad i = 1, \ldots, N.$$

Subtracting the equation (B.4.2) with $i = N$ from the equations (B.4.2) for $i \neq N$, we obtain

$$r_i - r_N + \mu_i - \sigma_i \sum_{j=1}^{N-1} \alpha_j \sigma_j \rho_{ij} = \sigma_i \sum_{k=1}^{i} L_{ik} \gamma_k, \quad i = 1, \ldots, N-1. \tag{B.4.3}$$

Using (B.4.3), we recurrently find the market prices of risk:

$$\gamma_i = \frac{r_i - r_N + \mu_i - \sigma_i \sum_{j=1}^{N-1} \alpha_j \sigma_j \rho_{ij} - \sigma_i \sum_{k=1}^{i-1} L_{i,k} \gamma_k}{\sigma_i L_{i,i}}, \quad i = 1, \ldots, N-1, \tag{B.4.4}$$

which are well defined because due to our assumptions $\sigma_i > 0$ and $L_{i,i} > 0$. Moreover, sum up (B.4.3) over i from $i = 1$ to $N-1$ and substitute the result in (B.4.2) with $i = N$ to confirm (B.4.1):

$$r_N - r_X + \sum_{j=1}^{N-1} \left(-\alpha_j \mu_j + \frac{\alpha_j(\alpha_j+1)}{2}\sigma_j^2 + \sum_{k=1}^{j-1} \alpha_j \alpha_k \sigma_j \sigma_k \rho_{jk} \right)$$
$$= -\sum_{j=1}^{N-1} \left(\alpha_j r_j - \alpha_j r_N + \alpha_j \mu_j - \alpha_j \sigma_j \sum_{k=1}^{N-1} \alpha_k \sigma_k \rho_{jk} \right)$$
$$\Leftrightarrow \quad r_X = \left(1 - \sum_{j=1}^{N} \alpha_j\right) r_N + \sum_{j=1}^{N} \alpha_j r_j + \sum_{j=1}^{N-1} \frac{\alpha_j(\alpha_j+1)}{2}\sigma_j^2$$
$$+ \sum_{j=1}^{N-1} \sum_{k=1}^{j-1} \alpha_j \alpha_k \sigma_j \sigma_k \rho_{jk} - \sum_{j=1}^{N-1} \alpha_j \sigma_j \sum_{k=1}^{N-1} \alpha_k \sigma_k \rho_{jk}$$
$$= \left(1 - \sum_{j=1}^{N} \alpha_j\right) r_N + \sum_{j=1}^{N} \alpha_j r_j + \sum_{j=1}^{N-1} \frac{\alpha_j(\alpha_j+1)}{2}\sigma_j^2$$

$$
\begin{aligned}
&- \sum_{j=1}^{N-1} \sum_{k=j}^{N-1} \alpha_j \alpha_k \sigma_j \sigma_k \rho_{jk} \\
&= \left(1 - \sum_{j=1}^{N} \alpha_j\right) r_N + \sum_{j=1}^{N} \alpha_j r_j + \sum_{j=1}^{N-1} \frac{\alpha_j(1-\alpha_j)}{2} \sigma_j^2 \\
&- \sum_{j=1}^{N-1} \sum_{k=1}^{j-1} \alpha_j \alpha_k \sigma_j \sigma_k \rho_{jk}.
\end{aligned}
$$

The found γ_i, $i = 1, \ldots, N-1$, from (B.4.4) and r_X from (B.4.1) together with Girsanov's theorem ensure that there is an EMM Q^X under which all $\widetilde{Y}_{c_i/X}$ are martingales. Thus, the considered market is arbitrage free. Theorem B.4.1 is proved.

B.5 PROOF OF PROPOSITION 11.3.1

Derivation of the MGF $M(t)$. Consider the MGF

$$M_{\alpha_1 \max(\beta_1 - Y, 0) + \alpha_2 \max(Y - \beta_2, 0)}(t)$$

for the random variable $\alpha_1 \max(\beta_1 - Y, 0) + \alpha_2 \max(Y - \beta_2, 0)$. We have for $\beta_1 \le \beta_2$:

$$
\begin{aligned}
M_{\alpha_1 \max(\beta_1 - Y, 0) + \alpha_2 \max(Y - \beta_2, 0)}(t) &= e^{\frac{t}{2}(t\alpha_1^2 + 2\alpha_1\beta_1)} N(t\alpha_1 + \beta_1) + N(\beta_2) - N(\beta_1) \\
&+ e^{\frac{t}{2}(t\alpha_2^2 - 2\alpha_2\beta_2)} N(t\alpha_2 - \beta_2).
\end{aligned}
\tag{B.5.1}
$$

Using (B.5.1), the fact that V is a combination of two independent random variables, X and $\alpha_1 \max(\beta_1 - Y, 0) + \alpha_2 \max(Y - \beta_2, 0)$, and the convolution theorem, we obtain the MGF for V:

$$
\begin{aligned}
&M_V(t) = M_X(t) \times M_{\alpha_1 \max(\beta_1 - Y, 0) + \alpha_2 \max(Y - \beta_2, 0)}(t) \\
&= e^{\frac{t^2}{2}} \left(e^{\frac{t}{2}(t\alpha_1^2 + 2\alpha_1\beta_1)} N(t\alpha_1 + \beta_1) + N(\beta_2) - N(\beta_1) + e^{\frac{t}{2}(t\alpha_2^2 - 2\alpha_2\beta_2)} N(t\alpha_2 - \beta_2) \right)
\end{aligned}
$$

$$= e^{\frac{t^2}{2}}(N(\beta_2) - N(\beta_1)) + e^{\frac{t}{2}(t(1+\alpha_2^2)-2\alpha_2\beta_2)} N(t\alpha_2 - \beta_2)$$
$$+ e^{\frac{t}{2}(t(1+\alpha_1^2)+2\alpha_1\beta_1)} N(t\alpha_1 + \beta_1).$$

Making use of basic properties of MGFs leads to the resulting formula (11.3.7):

$$\begin{aligned} M(t) &= M_Z(t) = M_{aV}(t) = M_V(at) \\ &= e^{\frac{(at)^2}{2}}(N(\beta_2) - N(\beta_1)) + e^{\frac{t}{2}(ta^2(1+\alpha_2^2)-2a\alpha_2\beta_2)} N(ta\alpha_2 - \beta_2) \\ &+ e^{\frac{t}{2}(ta^2(1+\alpha_1^2)+2a\alpha_1\beta_1)} N(ta\alpha_1 + \beta_1). \end{aligned}$$

□

Derivation of the restricted MGF $M(t, z_0)$. To obtain the formula (11.3.8) for $M(t, z_0)$, we consider the following restricted MGF for Z:

$$M_Z^*(t, z_0) := \mathbb{E}[e^{tZ}\mathbb{1}_{Z<z_0}],$$

which can be viewed as a complement to $M(t, z_0)$ as $M(t, z_0) = M(t) - M_Z^*(t, z_0)$ (note that $M_Z^*(t, z_0)$ is naturally used for pricing puts). We start with deriving the restricted MGF for V:

$$M_V(t, v_0) := \mathbb{E}[e^{tV}\mathbb{1}_{V<v_0}].$$

By splitting up the integration domain into three regions and calculating each integral separately, we obtain for $\beta_1 \leq \beta_2$:

$$\begin{aligned} &M_V(t, v_0) \\ &= \mathbb{E}[e^{tV}\mathbb{1}_{\{V<v_0\}}] \\ &= \mathbb{E}[e^{t(X+\alpha_1 \max(\beta_1-Y,0)+\alpha_2 \max(Y-\beta_2,0))}\mathbb{1}_{\{X+\alpha_1 \max(\beta_1-Y,0)+\alpha_2 \max(Y-\beta_2,0)<v_0\}}] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t(x+\alpha_1 \max(\beta_1-y,0)+\alpha_2 \max(y-\beta_2,0))} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \\ &\quad \times \mathbb{1}_{\{x+\alpha_1 \max(\beta_1-y,0)+\alpha_2 \max(y-\beta_2,0)<v_0\}} dydx \end{aligned}$$

$$
\begin{aligned}
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\beta_1} e^{t(x+\alpha_1(\beta_1-y))} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathbb{1}_{\{x+\alpha_1(\beta_1-y)<v_0\}} dydx \\
&\quad + \int_{-\infty}^{\infty} \int_{\beta_1}^{\beta_2} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathbb{1}_{\{x<v_0\}} dydx \\
&\quad + \int_{-\infty}^{\infty} \int_{\beta_2}^{\infty} e^{t(x+\alpha_2(y-\beta_2))} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathbb{1}_{\{x+\alpha_2(y-\beta_2)<v_0\}} dydx \\
&= e^{\frac{t^2}{2}} N(v_0 - t)\big(N(\beta_2) - N(\beta_1)\big) \\
&\quad + e^{\frac{t}{2}(t(1+\alpha_1^2)+2\alpha_1\beta_1)} N_2\left(t\alpha_1 + \beta_1, \frac{v_0 - t - \alpha_1(\beta_1 + \alpha_1 t)}{\sqrt{1+\alpha_1^2}}; \frac{-\alpha_1}{\sqrt{1+\alpha_1^2}}\right) \\
&\quad + e^{\frac{t}{2}(t(1+\alpha_2^2)-2\alpha_2\beta_2)} N_2\left(t\alpha_2 - \beta_2, \frac{v_0 - t + \alpha_2(\beta_2 - t\alpha_2)}{\sqrt{1+\alpha_2^2}}; \frac{-\alpha_2}{\sqrt{1+\alpha_2^2}}\right).
\end{aligned}
$$

Using basic properties of MGFs, we get

$$
\begin{aligned}
&M^*(t,z_0) = M_Z^*(t,z_0) = \mathbb{E}\left[e^{tZ}\mathbb{1}_{Z<z_0}\right] = \mathbb{E}\left[e^{atV}\mathbb{1}_{V<\frac{z_0}{a}}\right] = M_V\left(at, \frac{z_0}{a}\right) \\
&= e^{\frac{(at)^2}{2}} N\left(\frac{z_0}{a} - at\right)\big(N(\beta_2) - N(\beta_1)\big) \\
&\quad + e^{\frac{t}{2}(ta^2(1+\alpha_1^2)+2a\alpha_1\beta_1)} N_2\left(ta\alpha_1 + \beta_1, \frac{\frac{z_0}{a} - at - \alpha_1(\beta_1 + ta\alpha_1)}{\sqrt{1+\alpha_1^2}}; \frac{-\alpha_1}{\sqrt{1+\alpha_1^2}}\right) \\
&\quad + e^{\frac{t}{2}(ta^2(1+\alpha_2^2)-2a\alpha_2\beta_2)} N_2\left(ta\alpha_2 - \beta_2, \frac{\frac{z_0}{a} - t + \alpha_2(\beta_2 - ta\alpha_2)}{\sqrt{1+\alpha_2^2}}; \frac{-\alpha_2}{\sqrt{1+\alpha_2^2}}\right).
\end{aligned}
$$

We can simplify the following expression

$$
\begin{aligned}
&M(t,z_0) = M(t) - M^*(t,z_0) \\
&= e^{\frac{(at)^2}{2}} N\left(at - \frac{z_0}{a}\right)\big(N(\beta_2) - N(\beta_1)\big) \\
&\quad + e^{\frac{t}{2}(ta^2(1+\alpha_1^2)+2a\alpha_1\beta_1)} \\
&\quad \times \left\{ N(ta\alpha_1 + \beta_1) - N_2\left(ta\alpha_1 + \beta_1, \frac{\frac{z_0}{a} - at - \alpha_1(\beta_1 + ta\alpha_1)}{\sqrt{1+\alpha_1^2}}; \frac{-\alpha_1}{\sqrt{1+\alpha_1^2}}\right)\right\} \\
&\quad + e^{\frac{t}{2}(ta^2(1+\alpha_2^2)-2a\alpha_2\beta_2)}
\end{aligned}
$$

$$\times \left\{ N(ta\alpha_2 - \beta_2) - N_2\left(ta\alpha_2 - \beta_2, \frac{\frac{z_0}{a} - t + \alpha_2(\beta_2 - ta\alpha_2)}{\sqrt{1+\alpha_2^2}}; \frac{-\alpha_2}{\sqrt{1+\alpha_2^2}} \right) \right\},$$

which gives (11.3.8). □

B.6 MOMENTS OF THE RANDOM VARIABLE V

The derivatives of the MGF $M_V^{(i)}(0)$ the random variable V from (11.3.5) (i.e., the first four moments of V) are equal to

$$\begin{aligned}
M_V^{(1)}(0) &= \alpha_1\beta_1 N(\beta_1) - \alpha_2\beta_2 N(-\beta_2) + \frac{\alpha_1}{\sqrt{2\pi}} e^{-\frac{\beta_1^2}{2}} + \frac{\alpha_2}{\sqrt{2\pi}} e^{-\frac{\beta_2^2}{2}}, \\
M_V^{(2)}(0) &= N(\beta_1)\left[(\alpha_1\beta_1)^2 + \alpha_1^2\right] + N(\beta_2) + N(-\beta_2)\left[(\alpha_2\beta_2)^2 + 1 + \alpha_2^2\right] \\
&\quad + \alpha_1\beta_1 \frac{\alpha_1}{\sqrt{2\pi}} e^{-\frac{\beta_1^2}{2}} - \alpha_2\beta_2 \frac{\alpha_2}{\sqrt{2\pi}} e^{-\frac{\beta_2^2}{2}}, \\
M_V^{(3)}(0) &= N(\beta_1)\left[3\alpha_1\beta_1(1+\alpha_1^2) + (\alpha_1\beta_1)^3\right] \\
&\quad + N(-\beta_2)\left[-3\alpha_2\beta_2(1+\alpha_2^2) - (\alpha_2\beta_2)^3\right] \\
&\quad + \frac{\alpha_1}{\sqrt{2\pi}} e^{-\frac{\beta_1^2}{2}} \left[(\alpha_1\beta_1)^2 + 3 + 2\alpha_1^2\right] + \frac{\alpha_2}{\sqrt{2\pi}} e^{-\frac{\beta_2^2}{2}} \left[(\alpha_2\beta_2)^2 + 3 + 2\alpha_2^2\right], \\
M_V^{(4)}(0) &= 3N(\beta_2) - 3N(\beta_1) \\
&\quad + N(-\beta_2)\left[3(1+\alpha_2^2)(2(\alpha_2\beta_2)^2 + 1 + \alpha_2^2) + (\alpha_2\beta_2)^4\right] \\
&\quad + N(\beta_1)\left[3(1+\alpha_1^2)(2(\alpha_1\beta_1)^2 + 1 + \alpha_1^2) + (\alpha_1\beta_1)^4\right] \\
&\quad + \frac{\alpha_1}{\sqrt{2\pi}} e^{-\frac{\beta_1^2}{2}} \left[\alpha_1\beta_1(6 + 5\alpha_1^2 + (\alpha_1\beta_1)^2)\right] \\
&\quad + \frac{\alpha_2}{\sqrt{2\pi}} e^{-\frac{\beta_2^2}{2}} \left[-\alpha_2\beta_2(6 + 5\alpha_2^2 + (\alpha_2\beta_2)^2)\right].
\end{aligned}$$

B.7 USEFUL FORMULAS AND DERIVATIONS IN REGARD TO BEST-OF OPTIONS

B.7.1 *Derivation of best-of option pay-off function*

Similar to [6], we can rewrite the payoff of a best-off option on two assets S_1 and S_2 can be written as follows

$$\begin{aligned} b(T) &= \frac{S_1}{K_1}\mathbf{I}\left(S_1 > K_1, \frac{S_1}{K_1} > \frac{S_2}{K_2}\right) + \frac{S_2}{K_2}\mathbf{I}\left(S_2 > K_2, \frac{S_1}{K_1} < \frac{S_2}{K_2}\right) \\ &\quad + \mathbf{I}\left(S_1 < K_1, S_2 < K_2\right) - 1 \\ &= \frac{S_1}{K_1}\mathbf{I}\left(S_1 > K_1, S_3 > K_3\right) + \frac{S_2}{K_2}\mathbf{I}\left(S_2 > K_2, S_3 < K_3\right) \\ &\quad + \mathbf{I}\left(S_1 < K_1, S_2 < K_2\right) - 1, \end{aligned}$$

where $S_3 = \frac{S_1}{S_2}$ and $K_3 = \frac{K_1}{K_2}$.

While, we do not try to give a full analytical proof here, the following holds (see [7]):

$$\begin{aligned} \left[1 + K_1\frac{\partial}{\partial K_1} + K_2\frac{\partial}{\partial K_2}\right] b(T) &= b(T) - \frac{S_1}{K_1}\mathbf{I}\left(S_1 > K_1, S_3 > K_3\right) \\ &\quad - \frac{S_2}{K_2}\mathbf{I}\left(S_2 > K_2, S_3 < K_3\right) \\ &= \mathbf{I}\left(S_1 < K_1, S_2 < K_2\right) - 1, \end{aligned}$$

and rearranging leads to

$$\left[1 + K_1\frac{\partial}{\partial K_1} + K_2\frac{\partial}{\partial K_2}\right] b(T) + 1 = \mathbf{I}\left(S_1 < K_1, S_2 < K_2\right). \qquad \text{(B.7.1)}$$

B.7.2 *Derivation of partial derivatives of the cumulative distribution function of $v_€$*

We have the price in a convenient form (see [83, 109, 6]) as follows:

$$v_€(0) = e^{-r_€ T}\left[\frac{F_{£/€}}{K_1} N(d_1^+, d_3^+; \rho_{13}) + \frac{F_{\$/€}}{K_2} N(d_2^+, d_3^-; \rho_{23}) \right.$$
$$\left. + N(-d_1^-, -d_2^-; \rho_{12}) - 1\right],$$

We now assume that the volatilites are a function of the strikes: $\sigma_i(K_i)$, $i = 1, 2, 3$, with $K_3 = K_1/K_2$ as in (11.4.13). First, note that the following derivatives and useful equivalencies hold, which we need later:

$$d_i^\pm = \frac{\log \frac{F_i}{K_i} \pm \sigma_i^2 T/2}{\sigma_i \sqrt{T}},$$
$$d_i^+ = d_i^- + \sigma_i\sqrt{T},$$
$$\rho_{12} = \frac{\sigma_1^2 + \sigma_2^2 - \sigma_3^2}{2\sigma_1\sigma_2},$$
$$\rho_{13} = \frac{\sigma_1^2 + \sigma_3^2 - \sigma_2^2}{2\sigma_1\sigma_3} = \frac{\rho_{12}\sigma_1\sigma_2 - \sigma_2^2 + \sigma_3^2}{\sigma_1\sigma_3},$$
$$\rho_{23} = \frac{\sigma_2^2 + \sigma_3^2 - \sigma_1^2}{2\sigma_2\sigma_3} = \frac{\rho_{12}\sigma_1\sigma_2 - \sigma_1^2 + \sigma_3^2}{\sigma_2\sigma_3},$$
$$N(a, b, \rho) = \int_{-\infty}^{a}\int_{-\infty}^{b} \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2+y^2-2\rho xy}{2(1-\rho^2)}} dxdy.$$

And also:

$$N'(d_1^+) = \frac{K_1}{F_1} N'(d_1^-),$$
$$\frac{\partial}{\partial d_1^-} N(-d_1^-, -d_2^-; \rho_{12})$$
$$= \frac{\partial}{\partial d_1^-} \int_{-\infty}^{-d_1^-}\int_{-\infty}^{-d_2^-} \frac{1}{2\pi\sqrt{1-\rho_{12}^2}} \exp\left(-\frac{x^2+y^2-2\rho_{12}xy}{2(1-\rho_{12}^2)}\right) dxdy$$

$$
\begin{aligned}
&= -\int_{-\infty}^{-d_2^-} \frac{1}{2\pi\sqrt{1-\rho_{12}^2}} \exp\left(-\frac{(d_1^-)^2 + y^2 + 2\rho_{12} d_1^- y}{2(1-\rho_{12}^2)}\right) dy \\
&= -\exp\left(-\frac{(d_1^-)^2}{2}\right) \int_{-\infty}^{-d_2^- + \rho_{12} d_1^-} \frac{1}{2\pi\sqrt{1-\rho_{12}^2}} \exp\left(-\frac{u^2}{2(1-\rho_{12}^2)}\right) du \\
&= -\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(d_1^-)^2}{2}\right) \int_{-\infty}^{\frac{-d_2^- + \rho_{12} d_1^-}{\sqrt{1-\rho_{12}^2}}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right) dv \\
&= -N'(d_1^-) N\left(\frac{-d_2^- + \rho_{12} d_1^-}{\sqrt{1-\rho_{12}^2}}\right),
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial d_1^+}{\partial \sigma_1} &= -\frac{\log \frac{F_1}{K_1} - \sigma_1^2 T/2}{\sigma_1^2 \sqrt{T}} = -\frac{\log \frac{F_1}{K_1}}{\sigma_1^2 \sqrt{T}} + \sqrt{T}/2, \\
\frac{\partial d_1^-}{\partial \sigma_1} &= -\frac{\log \frac{F_1}{K_1} + \sigma_1^2 T/2}{\sigma_1^2 \sqrt{T}} = -\frac{\log \frac{F_1}{K_1}}{\sigma_1^2 \sqrt{T}} - \sqrt{T}/2, \\
\sqrt{1-\rho_{13}^2} &= \sqrt{1 - \left(\frac{\rho_{12}\sigma_1\sigma_2 - \sigma_2^2 + \sigma_3^2}{\sigma_1\sigma_3}\right)^2} \\
&= \sqrt{\frac{\sigma_1^2\sigma_3^2 - (\rho_{12}\sigma_1\sigma_2)^2 - 2\rho_{12}\sigma_1\sigma_2(\sigma_3^2 - \sigma_2^2) - (\sigma_3^2 - \sigma_2^2)^2}{\sigma_1^2\sigma_3^2}} \\
&= \sqrt{\frac{\sigma_1^2\sigma_2^2(1-\rho_{12}^2)}{\sigma_1^2\sigma_3^2}} \\
&= \frac{\sigma_2}{\sigma_3}\sqrt{1-\rho_{12}^2}, \\
\frac{\sigma_3^2}{\sigma_2^2}\rho_{13}^2 &= -\left[1 - \rho_{12}^2 - \frac{\sigma_3^2}{\sigma_2^2}\right], \\
\sqrt{1-\rho_{23}^2} &= \frac{\sigma_1}{\sigma_3}\sqrt{1-\rho_{12}^2} \\
\frac{\sigma_3^2}{\sigma_1^2}\rho_{23}^2 &= -\left[1 - \rho_{12}^2 - \frac{\sigma_3^2}{\sigma_1^2}\right],
\end{aligned}
$$

$$
\frac{d_3^+ - \rho_{13} d_1^+}{\sqrt{1-\rho_{13}^2}}
$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{1-\rho_{12}^2}} \frac{\sigma_3}{\sigma_2} \\
&\quad \times \left(\frac{\log \frac{F_1}{K_1} - \log \frac{F_2}{K_2} + \sigma_3^2 T/2}{\sigma_3 \sqrt{T}} - \left(\rho_{12} \frac{\sigma_2}{\sigma_3} - \frac{\sigma_2^2}{\sigma_1 \sigma_3} + \frac{\sigma_3}{\sigma_1} \right) \frac{\log \frac{F_1}{K_1} + \sigma_1^2 T/2}{\sigma_1 \sqrt{T}} \right) \\
&= \frac{1}{\sqrt{1-\rho_{12}^2}} \\
&\quad \times \left(\frac{\log \frac{F_1}{K_1} - \log \frac{F_2}{K_2}}{\sigma_2 \sqrt{T}} + \frac{\sigma_3^2}{\sigma_2} \sqrt{T}/2 + \left(-\rho_{12} + \frac{\sigma_2}{\sigma_1} - \frac{\sigma_3^2}{\sigma_1 \sigma_2} \right) \frac{\log \frac{F_1}{K_1} + \sigma_1^2 T/2}{\sigma_1 \sqrt{T}} \right) \\
&= \frac{1}{\sqrt{1-\rho_{12}^2}} \\
&\quad \times \Big(- \frac{\log \frac{F_2}{K_2}}{\sigma_2 \sqrt{T}} + \sigma_2 \sqrt{T}/2 - \rho_{12} \frac{\log \frac{F_1}{K_1} + \sigma_1^2 T/2}{\sigma_1 \sqrt{T}} + \frac{\log \frac{F_1}{K_1}}{\sigma_2 \sqrt{T}} - \frac{\sigma_3^2}{\sigma_1 \sigma_2} \frac{\log \frac{F_1}{K_1}}{\sqrt{T}} \\
&\quad + \frac{\sigma_2}{\sigma_1^2} \frac{\log \frac{F_1}{K_1}}{\sqrt{T}} \Big) \\
&= \frac{1}{\sqrt{1-\rho_{12}^2}} \left(-d_2^- - \rho_{12} d_1^+ + \frac{\log \frac{F_1}{K_1}}{\sqrt{T}} \left(-\frac{\sigma_3^2}{\sigma_1^2 \sigma_2} + \frac{1}{\sigma_2} + \frac{\sigma_2}{\sigma_1^2} \right) \right) \\
&= \frac{1}{\sqrt{1-\rho_{12}^2}} \left(-d_2^- - \rho_{12} d_1^- - 2\rho_{12} \sigma_1 \sqrt{T}/2 + \frac{\log \frac{F_1}{K_1}}{\sqrt{T}} \frac{2\rho_{12}}{\sigma_1} \right) \\
&= \frac{1}{\sqrt{1-\rho_{12}^2}} \left(-d_2^- - \rho_{12} d_1^- + 2\rho_{12} d_1^- \right) = \frac{-d_2^- + \rho_{12} d_1^-}{\sqrt{1-\rho_{12}^2}}, \\
&\frac{d_3^- - \rho_{23} d_2^+}{\sqrt{1-\rho_{23}^2}} = \frac{-d_1^- + \rho_{12} d_2^-}{\sqrt{1-\rho_{12}^2}},
\end{aligned}
$$

$$
\begin{aligned}
\left(\frac{d_3^+ - \rho_{13} d_1^+}{\sqrt{1-\rho_{13}^2}} \right)^2 &= \frac{(d_3^+)^2 - 2\rho_{13} d_1^+ d_3^+ + (\rho_{13})^2 (d_1^+)^2}{1-\rho_{13}^2} = \frac{(\rho_{12} d_1^- - d_2^-)^2}{1-\rho_{12}^2}, \\
\frac{(d_3^+)^2 - 2\rho_{13} d_1^+ d_3^+}{1-\rho_{13}^2} &= \frac{(\rho_{12} d_1^- - d_2^-)^2 - \frac{\sigma_3^2}{\sigma_2^2} (\rho_{13})^2 (d_1^+)^2}{1-\rho_{12}^2}, \\
\left(\frac{d_3^- - \rho_{23} d_2^+}{\sqrt{1-\rho_{23}^2}} \right)^2 &= \frac{(d_3^-)^2 - 2\rho_{23} d_2^+ d_3^- + (\rho_{23})^2 (d_2^+)^2}{1-\rho_{23}^2} = \frac{(\rho_{12} d_2^- - d_1^-)^2}{1-\rho_{12}^2},
\end{aligned}
$$

$$\begin{aligned}
\frac{(d_3^-)^2 - 2\rho_{23} d_2^+ d_3^-}{1-\rho_{23}^2} &= \frac{(\rho_{12} d_2^- - d_1^-)^2 - \frac{\sigma_3^2}{\sigma_1^2}(\rho_{23})^2 (d_2^+)^2}{1-\rho_{12}^2}, \\
(d_1^+)^2 &= (d_1^- + \sigma_1\sqrt{T})^2 \\
&= (d_1^-)^2 + 2d_1^-\sigma_1\sqrt{T} + \sigma_1^2 T) = (d_1^-)^2 + 2\log\frac{F_1}{K_1}, \\
(d_2^+)^2 &= (d_2^- + \sigma_2\sqrt{T})^2 \\
&= (d_2^-)^2 + 2d_2^-\sigma_2\sqrt{T} + \sigma_2^2 T) = (d_2^-)^2 + 2\log\frac{F_2}{K_2},
\end{aligned}$$

$$\begin{aligned}
&\frac{(d_1^+)^2 - 2\rho d_1^+ d_3^+ + (d_3^+)^2}{2(1-\rho_{13}^2)} \\
&= \frac{\frac{\sigma_3^2}{\sigma_2^2}(d_1^+)^2 + (\rho_{12} d_1^- - d_2^-)^2 - \frac{\sigma_3^2}{\sigma_2^2}(\rho_{13})^2 (d_1^+)^2}{2(1-\rho_{12}^2)} \\
&= \frac{\frac{\sigma_3^2}{\sigma_2^2}(d_1^+)^2 + \rho_{12}^2 (d_1^-)^2 - 2\rho_{12} d_1^- d_2^- + (d_2^-)^2 + \left[1-\rho_{12}^2 - \frac{\sigma_3^2}{\sigma_2^2}\right](d_1^+)^2}{2(1-\rho_{12}^2)} \\
&= \frac{\rho_{12}^2 (d_1^-)^2 - 2\rho_{12} d_1^- d_2^- + (d_2^-)^2 + \left[1-\rho_{12}^2\right]\left[(d_1^-)^2 + 2\log\frac{F_1}{K_1}\right]}{2(1-\rho_{12}^2)} \\
&= \frac{(d_1^-)^2 - 2\rho_{12} d_1^- d_2^- + (d_2^-)^2 + 2(1-\rho_{12}^2)\log\frac{F_1}{K_1}}{2(1-\rho_{12}^2)} \\
&= \log\frac{F_1}{K_1} + \frac{(d_1^-)^2 - 2\rho_{12} d_1^- d_2^- + (d_2^-)^2}{2(1-\rho_{12}^2)}
\end{aligned}$$

$$\frac{\partial\rho_{13}}{\partial\sigma_1} = \frac{\sigma_1^2+\sigma_2^2-\sigma_3^2}{2\sigma_1^2\sigma_3}, \quad \frac{\partial\rho_{23}}{\partial\sigma_1} = -\frac{\sigma_1}{\sigma_2\sigma_3}, \quad \frac{\partial\rho_{12}}{\partial\sigma_1} = \frac{\sigma_1^2-\sigma_2^2+\sigma_3^2}{2\sigma_1^2\sigma_2}.$$

Note that the following is true for the partial of the bivariate cumulative normal distribution with respect to the correlation ρ (see [73][Eq. (46.16)]).

$$\frac{\partial N(a,b,\rho)}{\partial\rho} = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left(-\frac{a^2-2\rho ab+b^2}{2(1-\rho^2)}\right).$$

This and use of some of the equivalences leads to the following:

$$\begin{aligned}
\frac{\partial N(d_1^+, d_3^+, \rho_{13})}{\partial \rho_{13}} &= \frac{1}{2\pi\sqrt{1-\rho_{13}^2}} \exp\left(-\frac{(d_1^+)^2 - 2\rho_{13} d_1^+ d_3^+ + (d_3^+)^2}{2(1-\rho_{13}^2)}\right) \\
&= \frac{K_1}{F_1}\frac{\sigma_3}{\sigma_2}\frac{1}{2\pi\sqrt{1-\rho_{12}^2}} \exp\left(-\frac{(d_1^-)^2 - 2\rho_{12} d_1^- d_2^- + (d_2^-)^2}{2(1-\rho_{12}^2)}\right), \\
\frac{\partial N(d_2^+, d_3^-, \rho_{23})}{\partial \rho_{23}} &= \frac{1}{2\pi\sqrt{1-\rho_{23}^2}} \exp\left(-\frac{(d_2^+)^2 - 2\rho_{23} d_2^+ d_3^- + (d_3^-)^2}{2(1-\rho_{23}^2)}\right) \\
&= \frac{K_2}{F_2}\frac{\sigma_3}{\sigma_1}\frac{1}{2\pi\sqrt{1-\rho_{12}^2}} \exp\left(-\frac{(d_1^-)^2 - 2\rho_{12} d_1^- d_2^- + (d_2^-)^2}{2(1-\rho_{12}^2)}\right).
\end{aligned}$$

$$\frac{\partial}{\partial \sigma_1} N(d_1^+, d_3+, \rho_{13}) = \frac{\partial \rho_{13}}{\partial \sigma_1}\frac{\partial N(d_1^+, d_3+, \rho_{13})}{\partial \rho_{13}}$$

Now the use of the above leads to the following:

$$\begin{aligned}
&K_1 \sigma_1' \frac{F_1}{K_1}\frac{\partial}{\partial \sigma_1} N(d_1^+, d_3^+, \rho_{13}) \\
&\quad = \sigma_1' F_1 N'(d_1^+) N\left(\frac{d_3^+ - \rho_{13} d_1^+}{\sqrt{1-\rho_{13}^2}}\right)\left(-\frac{\log\frac{F_1}{K_1}}{\sigma_1^2\sqrt{T}} + \sqrt{T}/2\right) \\
&\qquad + F_1 \sigma_1' \frac{\partial \rho_{13}}{\partial \sigma_1}\frac{\partial N(d_1^+, d_3^+, \rho_{13})}{\partial \rho_{13}}, \\
&K_1 \sigma_1' \frac{F_2}{K_2}\frac{\partial}{\partial \sigma_1} N(d_2^+, d_3^-, \rho_{23}) \\
&\quad = K_1 \frac{F_2}{K_2} \sigma_1' \frac{\partial \rho_{23}}{\partial \sigma_1}\frac{\partial N(d_2^+, d_3^-, \rho_{23})}{\partial \rho_{23}}, \\
&K_1 \sigma_1' \frac{\partial}{\partial \sigma_1} N(-d_1^-, -d_2^-; \rho_{12}) \\
&\quad = -K_1 \sigma_1' N'(d_1^-) N\left(\frac{-d_2^- + \rho_{12} d_1^-}{\sqrt{1-\rho_{12}^2}}\right)\left(-\frac{\log\frac{F_1}{K_1}}{\sigma_1^2\sqrt{T}} - \sqrt{T}/2\right) \\
&\qquad + K_1 \sigma_1' \frac{\partial \rho_{12}}{\partial \sigma_1}\frac{\partial N(-d_1^-, -d_2^-; \rho_{12})}{\partial \rho_{12}},
\end{aligned}$$

which we can combine as follows:

$$F_1\sigma_1' N'(d_1^+) N\left(\frac{d_3^+ - \rho_{13} d_1^+}{\sqrt{1-\rho_{13}^2}}\right)\left(-\frac{\log\frac{F_1}{K_1}}{\sigma_1^2\sqrt{T}} + \sqrt{T}/2\right)$$
$$+ K_1\sigma_1' N'(d_1^-) N\left(\frac{-d_2^- + \rho_{12} d_1^-}{\sqrt{1-\rho_{12}^2}}\right)\left(\frac{\log\frac{F_1}{K_1}}{\sigma_1^2\sqrt{T}} + \sqrt{T}/2\right)$$
$$= K_1\sigma_1' N'(d_1^-) N\left(\frac{-d_2^- + \rho_{12} d_1^-}{\sqrt{1-\rho_{12}^2}}\right)\sqrt{T}.$$

And also,

$$F_1\sigma_1'\frac{\partial\rho_{13}}{\partial\sigma_1}\frac{\partial N(d_1^+, d_3^+, \rho_{13})}{\partial\rho_{13}} + K_1\frac{F_2}{K_2}\sigma_1'\frac{\partial\rho_{23}}{\partial\sigma_1}\frac{\partial N(d_2^+, d_3^-, \rho_{23})}{\partial\rho_{23}}$$
$$+ K_1\sigma_1'\frac{\partial\rho_{12}}{\partial\sigma_1}\frac{\partial N(-d_1^-, -d_2^-; \rho_{12})}{\partial\rho_{12}}$$
$$= \sigma_1' F_1\frac{\sigma_1^2+\sigma_2^2-\sigma_3^2}{2\sigma_1^2\sigma_3}\frac{K_1}{F_1}\frac{\sigma_3}{\sigma_2}\frac{1}{2\pi\sqrt{1-\rho_{12}^2}}\exp\left(-\frac{(d_1^-)^2 - 2\rho_{12}d_1^-d_2^- + (d_2^-)^2}{2(1-\rho_{12}^2)}\right)$$
$$- \sigma_1'\frac{F_2}{K_2}\frac{\sigma_1}{\sigma_2\sigma_3}\frac{K_2}{F_2}\frac{\sigma_3}{\sigma_1}\frac{1}{2\pi\sqrt{1-\rho_{12}^2}}\exp\left(-\frac{(d_1^-)^2 - 2\rho_{12}d_1^-d_2^- + (d_2^-)^2}{2(1-\rho_{12}^2)}\right)$$
$$+ \sigma_1' K_1\frac{\sigma_1^2-\sigma_2^2+\sigma_3^2}{2\sigma_1^2\sigma_2}\frac{1}{2\pi\sqrt{1-\rho_{12}^2}}\exp\left(-\frac{(d_1^-)^2 - 2\rho_{12}d_1^-d_2^- + (d_2^-)^2}{2(1-\rho_{12}^2)}\right)$$
$$= 0.$$

The same equivalences hold for the partial derivatives with respect to σ_2, hence, all together simplifies to the following:

$$U(K_1, K_2) := \left[1 + K_1\frac{\partial}{\partial K_1} + K_2\frac{\partial}{\partial K_2}\right] v_{€} + e^{-r_{€}T}$$
$$= e^{-r_{€}T}\left(N(-d_1^-, -d_2^-; \rho_{12}) + \left[K_1\sigma_1'(K_1)\frac{\partial}{\partial\sigma_1} + K_2\sigma_2'(K_2)\frac{\partial}{\partial\sigma_2}\right] v_{€} + 1\right)$$
$$= e^{-r_{€}T}\left[N(-d_1^-, -d_2^-; \rho_{12}) + K_1\sqrt{T}\sigma_1'(K_1) N'(d_1^-) N\left(\frac{d_1^-\rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right)\right.$$

$$+K_2\sqrt{T}\sigma_2'(K_2)N'(d_2^-)N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)\Bigg].$$

B.7.3 *Derivation of second order partial derivatives of the cumulative distribution function of $v_€$*

We have the distribution function $U(K_1, K_2)$ (see (11.4.14)):

$$\begin{aligned}U(K_1,K_2) = e^{-r_€T}\Big[&N(-d_1^-,-d_2^-;\rho_{12})\\ &+K_1\sqrt{T}\sigma_1'(K_1)N'(d_1^-)N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)\\ &+K_2\sqrt{T}\sigma_2'(K_2)N'(d_2^-)N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)\Bigg].\end{aligned}$$

To make the following presentation clearer, let us start with the following partial derivatives:

$$\begin{aligned}&\frac{\partial N(-d_1^-,-d_2^-;\rho_{12})}{\partial K_i}\\ &=\frac{\partial N(-d_1^-,-d_2^-;\rho_{12})}{\partial d_i^-}\frac{\partial d_i^-}{\partial\sigma_i}\frac{\partial\sigma_i(K_i)}{K_i}+\frac{\partial N(-d_1^-,-d_2^-;\rho_{12})}{\partial\rho_{12}}\frac{\partial\rho_{12}}{\partial\sigma_i}\frac{\partial\sigma_i(K_i)}{\partial K_i},\\ &\frac{\partial N(-d_1^-,-d_2^-;\rho_{12})}{\partial K_1}\\ &=\frac{\partial N(-d_1^-,-d_2^-;\rho_{12})}{\partial d_1^-}\frac{\partial d_1^-}{\partial\sigma_1}\frac{\partial\sigma_1(K_1)}{K_1}+\frac{\partial N(-d_1^-,-d_2^-;\rho_{12})}{\partial\rho_{12}}\frac{\partial\rho_{12}}{\partial\sigma_1}\frac{\partial\sigma_1(K_1)}{\partial K_1},\\ &\frac{\partial^2 N(-d_1^-,-d_2^-;\rho_{12})}{\partial K_1\partial K_2}\\ &=\frac{\partial^2 N(-d_1^-,-d_2^-;\rho_{12})}{\partial d_1^-\partial d_2^-}\frac{\partial d_1^-}{\partial\sigma_1}\frac{\partial\sigma_1(K_1)}{K_1}\frac{\partial d_2^-}{\partial\sigma_2}\frac{\partial\sigma_2(K_2)}{K_2}\\ &\quad+\frac{\partial^2 N(-d_1^-,-d_2^-;\rho_{12})}{\partial\rho_{12}\partial d_2^-}\frac{\partial\rho_{12}}{\partial\sigma_1}\frac{\partial\sigma_1(K_1)}{\partial K_1}\frac{\partial d_2^-}{\partial\sigma_2}\frac{\partial\sigma_2(K_2)}{K_2}\\ &\quad+\frac{\partial^2 N(-d_1^-,-d_2^-;\rho_{12})}{\partial d_1^-\partial\rho_{12}}\frac{\partial d_1^-}{\partial\sigma_1}\frac{\partial\sigma_1(K_1)}{K_1}\frac{\partial\rho_{12}}{\partial\sigma_2}\frac{\partial\sigma_2(K_2)}{\partial K_2}\end{aligned}$$

$$
\begin{aligned}
&+ \frac{\partial^2 N(-d_1^-, -d_2^-; \rho_{12})}{\partial \rho_{12} \partial \rho_{12}} \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial \sigma_1(K_1)}{\partial K_1} \frac{\partial \rho_{12}}{\partial \sigma_2} \frac{\partial \sigma_2(K_2)}{\partial K_2} \\
&= \sigma_1'(K_1) \sigma_2'(K_2) \\
&\times \left[\frac{\partial^2 N(-d_1^-, -d_2^-; \rho_{12})}{\partial d_1^- \partial d_2^-} \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2} + \frac{\partial^2 N(-d_1^-, -d_2^-; \rho_{12})}{\partial \rho_{12} \partial d_2^-} \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2} \right. \\
&\left. + \frac{\partial^2 N(-d_1^-, -d_2^-; \rho_{12})}{\partial d_1^- \partial \rho_{12}} \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial \rho_{12}}{\partial \sigma_2} + \frac{\partial^2 N(-d_1^-, -d_2^-; \rho_{12})}{\partial \rho_{12} \partial \rho_{12}} \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial \rho_{12}}{\partial \sigma_2} \right],
\end{aligned}
$$

where $\sigma_i'(K_i) = \frac{\partial \sigma_i(K_i)}{K_i}$ and $\sigma_i''(K_i) = \frac{\partial \sigma_i'(K_i)}{K_i}$. Moreover, we have the following:

$$
\begin{aligned}
&\frac{\partial \left(K_1 \sqrt{T} \sigma_1'(K_1) N'(d_1^-) N\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right) \right)}{\partial K_1} \\
&= \sqrt{T} N'(d_1^-) \left(\sigma_1'(K_1) + \frac{\partial \sigma_1'(K_1)}{\partial K_1} K_1 - d_1^- \sigma_1'(K_1) K_1 \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial \sigma_1(K_1)}{\partial K_1} \right) \\
&\quad \times N\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right) \\
&\quad + K_1 \sqrt{T} \sigma_1'(K_1) N'(d_1^-) \frac{\partial N\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right)}{\partial d_1^-} \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial \sigma_1(K_1)}{\partial K_1} \\
&\quad + K_1 \sqrt{T} \sigma_1'(K_1) N'(d_1^-) \frac{\partial N\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right)}{\partial \rho_{12}} \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial \sigma_1(K_1)}{\partial K_1},
\end{aligned}
$$

$$
\begin{aligned}
&\frac{\partial^2 \left(K_1 \sqrt{T} \sigma_1'(K_1) N'(d_1^-) N\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right) \right)}{\partial K_1 \partial K_2} \\
&= \sqrt{T} N'(d_1^-) \left(\sigma_1'(K_1) + \frac{\partial \sigma_1'(K_1)}{\partial K_1} K_1 - d_1^- K_1 \sigma_1'(K_1) \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial \sigma_1(K_1)}{\partial K_1} \right) \frac{\partial \sigma_2(K_2)}{\partial K_2} \\
&\quad \times \left[\frac{\partial N\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right)}{\partial d_2^-} \frac{\partial d_2^-}{\partial \sigma_2} + \frac{\partial N\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right)}{\partial \rho_{12}} \frac{\partial \rho_{12}}{\partial \sigma_2} \right] \\
&\quad + K_1 \sqrt{T} \sigma_1'(K_1) N'(d_1^-) \frac{\partial \sigma_1(K_1)}{\partial K_1} \frac{\partial \sigma_2(K_2)}{\partial K_2} \frac{\partial d_1^-}{\partial \sigma_1}
\end{aligned}
$$

$$\times\left[\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-\partial d_2^-}\frac{\partial d_2^-}{\partial\sigma_2}+\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-\partial\rho_{12}}\frac{\partial\rho_{12}}{\partial\sigma_2}\right]$$

$$+K_1\sqrt{T}\sigma_1'(K_1)N'(d_1^-)\frac{\partial\sigma_1(K_1)}{\partial K_1}\frac{\partial\sigma_2(K_2)}{\partial K_2}\frac{\partial\rho_{12}}{\partial\sigma_1}$$

$$\times\left[\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial\rho_{12}\partial d_2^-}\frac{\partial d_2^-}{\partial\sigma_2}+\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial\rho_{12}\partial\rho_{12}}\frac{\partial\rho_{12}}{\partial\sigma_2}\right]$$

$$=\sqrt{T}N'(d_1^-)\sigma_2'(K_2)\left\{\left(\sigma_1'(K_1)+\sigma_1''(K_1)K_1-d_1^-(\sigma_1'(K_1))^2K_1\frac{\partial d_1^-}{\partial\sigma_1}\right)\right.$$

$$\times\left[\frac{\partial N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_2^-}\frac{\partial d_2^-}{\partial\sigma_2}+\frac{\partial N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial\rho_{12}}\frac{\partial\rho_{12}}{\partial\sigma_2}\right]$$

$$+K_1(\sigma_1'(K_1))^2\frac{\partial d_1^-}{\partial\sigma_1}\left[\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-\partial d_2^-}\frac{\partial d_2^-}{\partial\sigma_2}+\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-\partial\rho_{12}}\frac{\partial\rho_{12}}{\partial\sigma_2}\right]$$

$$\left.+K_1(\sigma_1'(K_1))^2\frac{\partial\rho_{12}}{\partial\sigma_1}\left[\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial\rho_{12}\partial d_2^-}\frac{\partial d_2^-}{\partial\sigma_2}+\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial\rho_{12}\partial\rho_{12}}\frac{\partial\rho_{12}}{\partial\sigma_2}\right]\right\},$$

$$\frac{\partial\left(K_2\sqrt{T}\sigma_2'(K_2)N'(d_2^-)N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)\right)}{\partial K_1}$$

$$=K_2\sqrt{T}\sigma_2'(K_2)N'(d_2^-)\sigma_1'(K_1)\left[\frac{\partial N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-}\frac{\partial d_1^-}{\partial\sigma_1}+\frac{\partial N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial\rho_{12}}\frac{\partial\rho_{12}}{\partial\sigma_1}\right],$$

$$\frac{\partial^2\left(K_2\sqrt{T}\sigma_2'(K_2)N'(d_2^-)N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)\right)}{\partial K_1\partial K_2}$$

$$=\sqrt{T}N'(d_2^-)\sigma_1'(K_1)\left(\sigma_2'(K_2)+\sigma_2''(K_2)K_2-d_2^-K_2(\sigma_2'(K_2))^2\frac{\partial d_2^-}{\partial\sigma_2}\right)$$

$$\times \left[\frac{\partial N\left(\frac{d_2^- \rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-} \frac{\partial d_1^-}{\partial \sigma_1} + \frac{\partial N\left(\frac{d_2^- \rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial \rho_{12}} \frac{\partial \rho_{12}}{\partial \sigma_1} \right]$$

$$+ K_2 \sqrt{T} \sigma_2'(K_2) N'(d_2^-) \sigma_1'(K_1) \sigma_2'(K_2) \frac{\partial d_1^-}{\partial \sigma_1}$$

$$\times \left[\frac{\partial^2 N\left(\frac{d_2^- \rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^- \partial d_2^-} \frac{\partial d_2^-}{\partial \sigma_2} + \frac{\partial^2 N\left(\frac{d_2^- \rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^- \partial \rho_{12}} \frac{\partial \rho_{12}}{\partial \sigma_2} \right]$$

$$+ K_2 \sqrt{T} \sigma_2'(K_2) N'(d_2^-) \sigma_1'(K_1) \sigma_2'(K_2) \frac{\partial \rho_{12}}{\partial \sigma_1}$$

$$\times \left[\frac{\partial^2 N\left(\frac{d_2^- \rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial \rho_{12} \partial d_2^-} \frac{\partial d_2^-}{\partial \sigma_2} + \frac{\partial^2 N\left(\frac{d_2^- \rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial \rho_{12} \partial \rho_{12}} \frac{\partial \rho_{12}}{\partial \sigma_2} \right].$$

The following derivatives will also be helpful:

$$\frac{\partial d_i^\pm}{\partial \sigma_i} = -\frac{\log \frac{F_i}{K_i}}{\sigma_i^2 \sqrt{T}} \pm \sqrt{T}/2,$$

$$\frac{\partial N(f(d_i^-))}{\partial d_i^-} = \frac{\partial f(d_i^-)}{\partial d_i^-} N'(f(d_i^-)),$$

$$\frac{\partial N'(f(d_i^-))}{\partial d_i^-} = -f(d_i^-) \frac{\partial f(d_i^-)}{\partial d_i^-} N'(f(d_i^-)),$$

$$\frac{\partial \rho_{12}}{\partial \sigma_1} = \frac{\sigma_1^2 - \sigma_2^2 + \sigma_3^2}{2\sigma_1^2 \sigma_2}, \qquad \frac{\partial \rho_{12}}{\partial \sigma_2} = \frac{\sigma_2^2 - \sigma_1^2 + \sigma_3^2}{2\sigma_1 \sigma_2^2},$$

$$N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right) = \exp\left(-\frac{(d_1^-)^2 - 2\rho_{12} d_1^- d_2^- + (d_2^-)^2}{2(1-\rho_{12}^2)}\right) N'(d_1^-),$$

$$N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right) = N'\left(\frac{d_2^- \rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}}\right) \exp\left(\frac{(d_1^-)^2 + (d_2^-)^2}{2}\right),$$

$$\frac{\partial}{\partial \rho_{12}} \frac{d_1^- - \rho_{12} d_2^-}{\sqrt{1-\rho_{12}^2}} = \frac{d_1^- \rho_{12} - d_2^-}{(1-\rho_{12}^2)^{3/2}},$$

$$\frac{\partial}{\partial \rho_{12}} \frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} = \frac{d_1^- - d_2^- \rho_{12}}{(1-\rho_{12}^2)^{3/2}},$$

$$\frac{\partial}{\partial\rho_{12}}\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}=\frac{d_2^--d_1^-\rho_{12}}{(1-\rho_{12}^2)^{3/2}}.$$

We will also need the partial derivatives of $N(-d_1^-,-d_2^-;\rho_{12})$:

$$\begin{aligned}
\frac{\partial}{\partial d_1^-}N(-d_1^-,-d_2^-;\rho_{12})&=-N'(d_1^-)N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right),\\
\frac{\partial^2}{\partial d_1^-\partial d_2^-}N(-d_1^-,-d_2^-;\rho_{12})&=-\frac{N'(d_1^-)}{\sqrt{1-\rho_{12}^2}}N'\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right),\\
\frac{\partial^2}{\partial d_1^-\partial\rho_{12}}N(-d_1^-,-d_2^-;\rho_{12})&=-N'(d_1^-)\frac{d_1^--\rho_{12}d_2^-}{(1-\rho_{12}^2)^{3/2}}N'\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right),
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial N(-d_1^-,-d_2^-;\rho_{12})}{\partial\rho_{12}}\\
&=\frac{1}{2\pi\sqrt{1-\rho_{12}^2}}\exp\left(-\frac{(d_1^-)^2-2\rho_{12}d_1^-d_2^-+(d_2^-)^2}{2(1-\rho_{12}^2)}\right)\\
&=\frac{1}{2\pi N'(d_1^-)\sqrt{1-\rho_{12}^2}}N'\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right),\\
&\frac{\partial^2 N(-d_1^-,-d_2^-;\rho_{12})}{\partial\rho_{12}\partial d_2^-}\\
&=\frac{d_1^-\rho_{12}-d_2^-}{2\pi(1-\rho_{12}^2)^{3/2}}\exp\left(-\frac{(d_1^-)^2-2\rho_{12}d_1^-d_2^-+(d_2^-)^2}{2(1-\rho_{12}^2)}\right)\\
&=\frac{d_1^-\rho_{12}-d_2^-}{2\pi N'(d_1^-)(1-\rho_{12}^2)^{3/2}}N'\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right),\\
&\frac{\partial^2 N(-d_1^-,-d_2^-;\rho_{12})}{\partial\rho_{12}\partial\rho_{12}}\\
&=\left[\frac{\rho_{12}}{2\pi(1-\rho_{12}^2)^{3/2}}+\frac{d_1^-d_2^-(1+\rho_{12}^2)-(d_1^-)^2\rho_{12}-(d_2^-)^2\rho_{12}}{2\pi(1-\rho_{12}^2)^{5/2}}\right]\\
&\quad\times\exp\left(-\frac{(d_1^-)^2-2\rho_{12}d_1^-d_2^-+(d_2^-)^2}{2(1-\rho_{12}^2)}\right)
\end{aligned}$$

$$= \frac{1}{2\pi N'(d_1^-)} \left[\frac{\rho_{12}}{(1-\rho_{12}^2)^{3/2}} + \frac{d_1^- d_2^- (1+\rho_{12}^2) - (d_1^-)^2 \rho_{12} - (d_2^-)^2 \rho_{12}}{(1-\rho_{12}^2)^{5/2}} \right]$$
$$\times N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right).$$

Then, we also need the following partial derivatives of $N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)$:

$$\frac{\partial N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-} = \frac{\rho_{12}}{\sqrt{1-\rho_{12}^2}} N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right),$$
$$\frac{\partial N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_2^-} = \frac{-1}{\sqrt{1-\rho_{12}^2}} N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right),$$
$$\frac{\partial N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial \rho_{12}} = \frac{d_1^- - \rho_{12} d_2^-}{(1-\rho_{12}^2)^{3/2}} N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right),$$
$$\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^- \partial d_2^-} = \frac{d_1^- \rho_{12}^2 - d_2^- \rho_{12}}{(1-\rho_{12}^2)^{3/2}} N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right),$$

$$\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^- \partial \rho_{12}} = N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right)$$
$$\times \left[\frac{1}{(1-\rho_{12}^2)^{3/2}} + \frac{\rho_{12} \left\{ d_1^- d_2^- (1+\rho_{12}^2) - [(d_1^-)^2 + (d_2^-)^2] \rho_{12} \right\}}{(1-\rho_{12}^2)^{5/2}} \right],$$
$$\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial \rho_{12} \partial d_2^-} = N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right)$$
$$\times \left[\frac{-\rho_{12}}{(1-\rho_{12}^2)^{3/2}} + \frac{[(d_1^-)^2 + (d_2^-)^2] \rho_{12} - d_1^- d_2^- (1+\rho_{12}^2)}{(1-\rho_{12}^2)^{5/2}} \right],$$
$$\frac{\partial^2 N\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial \rho_{12} \partial \rho_{12}} = N'\left(\frac{d_1^- \rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}} \right) \left[\frac{-4 d_2^- \rho_{12}^2 + 3 d_1^- \rho_{12} + d_2^-}{(1-\rho_{12}^2)^{5/2}} \right.$$

$$+\frac{-(d_1^-)^3\rho_{12} + (d_2^-)^3\rho_{12}^2 + (d_1^-)^2 d_2^- [2\rho_{12}^2 + 1] - d_1^-(d_2^-)^2[\rho_{12}^3 + 2\rho_{12}]}{(1-\rho_{12}^2)^{7/2}}\Bigg],$$

and the partial derivatives of $N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)$:

$$\frac{\partial N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-} = \frac{-1}{\sqrt{1-\rho_{12}^2}} N'\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right),$$

$$\frac{\partial N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial \rho_{12}} = \frac{d_2^- - \rho_{12}d_1^-}{(1-\rho_{12}^2)^{3/2}} N'\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right),$$

$$\frac{\partial^2 N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-\partial d_2^-} = \frac{d_2^-\rho_{12}^2 - d_1^-\rho_{12}}{(1-\rho_{12}^2)^{3/2}} N'\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right),$$

$$\frac{\partial^2 N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial d_1^-\partial \rho_{12}} = N'\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)$$
$$\times\left[\frac{-\rho_{12}}{(1-\rho_{12}^2)^{3/2}} + \frac{[(d_1^-)^2 + (d_2^-)^2]\rho_{12} - d_1^- d_2^-(1+\rho_{12}^2)}{(1-\rho_{12}^2)^{5/2}}\right],$$

$$\frac{\partial^2 N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial \rho_{12}\partial d_2^-} = N'\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)$$
$$\times\left[\frac{1}{(1-\rho_{12}^2)^{3/2}} + \frac{\rho_{12}\left\{d_1^- d_2^-(1+\rho_{12}^2) - [(d_1^-)^2 - (d_2^-)^2]\rho_{12}\right\}}{(1-\rho_{12}^2)^{5/2}}\right],$$

$$\frac{\partial^2 N\left(\frac{d_2^-\rho_{12}-d_1^-}{\sqrt{1-\rho_{12}^2}}\right)}{\partial \rho_{12}\partial \rho_{12}} = N'\left(\frac{d_1^-\rho_{12}-d_2^-}{\sqrt{1-\rho_{12}^2}}\right)\Bigg[\frac{4d_1^-\rho_{12}^2 + 3d_2^-\rho_{12} - d_1^-}{(1-\rho_{12}^2)^{5/2}}$$
$$+\frac{(d_1^-)^3\rho_{12}^2 - (d_2^-)^3\rho_{12} - (d_1^-)^2 d_2^-[\rho_{12}^3 + 2\rho_{12}] + d_1^-(d_2^-)^2[2\rho_{12} + 1]}{(1-\rho_{12}^2)^{7/2}}\Bigg].$$

Now, we can put all the above together to get the following:

$$
\begin{aligned}
P_1 &= \frac{\partial^2 N(-d_1^-, -d_2^-; \rho_{12})}{\partial K_1 \partial K_2} \\
&= \sigma_1'(K_1)\sigma_2'(K_2) \left\{ -\frac{N'(d_1^-)}{\sqrt{1-\rho_{12}^2}} N'\left(\frac{d_1^-\rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right) \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2} \right. \\
&\quad + \frac{d_1^-\rho_{12} - d_2^-}{2\pi N'(d_1^-)(1-\rho_{12}^2)^{3/2}} N'\left(\frac{d_1^-\rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right) \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2} \\
&\quad - N'(d_1^-) \frac{d_1^- - \rho_{12} d_2^-}{(1-\rho_{12}^2)^{3/2}} N'\left(\frac{d_1^-\rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right) \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial \rho_{12}}{\partial \sigma_2} \\
&\quad + \frac{1}{2\pi N'(d_1^-)} \left[\frac{\rho_{12}}{(1-\rho_{12}^2)^{3/2}} + \frac{d_1^- d_2^-(1+\rho_{12}^2) - (d_1^-)^2\rho_{12} - (d_2^-)^2\rho_{12}}{(1-\rho_{12}^2)^{5/2}} \right] \\
&\quad \left. \times N'\left(\frac{d_1^-\rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right) \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial \rho_{12}}{\partial \sigma_2} \right\} \\
&= \sigma_1'(K_1)\sigma_2'(K_2) N'\left(\frac{d_1^-\rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right) \left\{ -\frac{N'(d_1^-)}{\sqrt{1-\rho_{12}^2}} \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2} \right. \\
&\quad + \frac{d_1^-\rho_{12} - d_2^-}{2\pi N'(d_1^-)(1-\rho_{12}^2)^{3/2}} \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2} - N'(d_1^-) \frac{d_1^- - \rho_{12} d_2^-}{(1-\rho_{12}^2)^{3/2}} \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial \rho_{12}}{\partial \sigma_2} \\
&\quad + \frac{1}{2\pi N'(d_1^-)} \\
&\quad \left. \times \left[\frac{\rho_{12}}{(1-\rho_{12}^2)^{3/2}} + \frac{d_1^- d_2^-(1+\rho_{12}^2) - (d_1^-)^2\rho_{12} - (d_2^-)^2\rho_{12}}{(1-\rho_{12}^2)^{5/2}} \right] \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial \rho_{12}}{\partial \sigma_2} \right\},
\end{aligned}
$$

and also

$$
\begin{aligned}
P_2 &= \frac{\partial^2 \left(K_1 \sqrt{T} \sigma_1'(K_1) N'(d_1^-) N\left(\frac{d_1^-\rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right)\right)}{\partial K_1 \partial K_2} \\
&= \sqrt{T} N'(d_1^-) \sigma_2'(K_2) N'\left(\frac{d_1^-\rho_{12} - d_2^-}{\sqrt{1-\rho_{12}^2}}\right) \\
&\quad \left\{ \left(\sigma_1'(K_1) + \sigma_1''(K_1) K_1 - d_1^- (\sigma_1'(K_1))^2 K_1 \frac{\partial d_1^-}{\partial \sigma_1} \right) \right.
\end{aligned}
$$

$$
\begin{aligned}
&\times \left[\frac{-1}{\sqrt{1-\rho_{12}^2}} \frac{\partial d_2^-}{\partial \sigma_2} + \frac{d_1^- - \rho_{12} d_2^-}{(1-\rho_{12}^2)^{3/2}} \frac{\partial \rho_{12}}{\partial \sigma_2} \right] \\
&+ K_1(\sigma_1'(K_1))^2 \left[\frac{d_1^- \rho_{12}^2 - d_2^- \rho_{12}}{(1-\rho_{12}^2)^{3/2}} \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2} \right. \\
&+ \left(\frac{1}{(1-\rho_{12}^2)^{3/2}} + \frac{\rho_{12} \left\{ d_1^- d_2^- (1+\rho_{12}^2) - [(d_1^-)^2 + 0(d_2^-)^2]\rho_{12} \right\}}{(1-\rho_{12}^2)^{5/2}} \right) \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial \rho_{12}}{\partial \sigma_2} \\
&\times \left(\frac{-\rho_{12}}{(1-\rho_{12}^2)^{3/2}} + \frac{[(d_1^-)^2 + (d_2^-)^2]\rho_{12} - d_1^- d_2^- (1+\rho_{12}^2)}{(1-\rho_{12}^2)^{5/2}} \right) \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2} \\
&+ \left(\frac{-4 d_2^- \rho_{12}^2 + 3 d_1^- \rho_{12} + d_2^-}{(1-\rho_{12}^2)^{5/2}} \right. \\
&+ \left. \frac{-(d_1^-)^3 \rho_{12} + (d_2^-)^3 \rho_{12}^2 + (d_1^-)^2 d_2^- [2\rho_{12}^2 + 1] - d_1^- (d_2^-)^2 [\rho_{12}^3 + 2\rho_{12}]}{(1-\rho_{12}^2)^{7/2}} \right) \\
&\left. \times \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial \rho_{12}}{\partial \sigma_2} \right] \Bigg\},
\end{aligned}
$$

and finally

$$
\begin{aligned}
P_3 &= \frac{\partial^2 \left(K_2 \sqrt{T} \sigma_2'(K_2) N'(d_2^-) N \left(\frac{d_2^- \rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}} \right) \right)}{\partial K_1 \partial K_2} \\
&= \sqrt{T} N'(d_2^-) \sigma_1'(K_1) N' \left(\frac{d_2^- \rho_{12} - d_1^-}{\sqrt{1-\rho_{12}^2}} \right) \\
&\quad \Bigg\{ \left(\sigma_2'(K_2) + \sigma_2''(K_2) K_2 - d_2^- K_2 (\sigma_2'(K_2))^2 \frac{\partial d_2^-}{\partial \sigma_2} \right) \\
&\quad \times \left[\frac{-1}{\sqrt{1-\rho_{12}^2}} \frac{\partial d_1^-}{\partial \sigma_1} + \frac{d_2^- - \rho_{12} d_1^-}{(1-\rho_{12}^2)^{3/2}} \frac{\partial \rho_{12}}{\partial \sigma_1} \right] \\
&\quad + K_2 (\sigma_2'(K_2))^2 \\
&\quad \times \left[\frac{d_2^- \rho_{12}^2 - d_1^- \rho_{12}}{(1-\rho_{12}^2)^{3/2}} \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2} \right. \\
&\quad + \left(\frac{-\rho_{12}}{(1-\rho_{12}^2)^{3/2}} + \frac{[(d_1^-)^2 + (d_2^-)^2]\rho_{12} - d_1^- d_2^- (1+\rho_{12}^2)}{(1-\rho_{12}^2)^{5/2}} \right) \frac{\partial d_1^-}{\partial \sigma_1} \frac{\partial \rho_{12}}{\partial \sigma_2} \\
&\quad + \left(\frac{1}{(1-\rho_{12}^2)^{3/2}} + \frac{\rho_{12} \left\{ d_1^- d_2^- (1+\rho_{12}^2) - [(d_1^-)^2 - (d_2^-)^2]\rho_{12} \right\}}{(1-\rho_{12}^2)^{5/2}} \right) \frac{\partial \rho_{12}}{\partial \sigma_1} \frac{\partial d_2^-}{\partial \sigma_2}
\end{aligned}
$$

$$+\left(\frac{4d_1^-\rho_{12}^2+3d_2^-\rho_{12}-d_1^-}{(1-\rho_{12}^2)^{5/2}}\right.$$
$$\left.+\frac{(d_1^-)^3\rho_{12}^2-(d_2^-)^3\rho_{12}-(d_1^-)^2d_2^-[\rho_{12}^3+2\rho_{12}]+d_1^-(d_2^-)^2[2\rho_{12}+1]}{(1-\rho_{12}^2)^{7/2}}\right)$$
$$\left.\left.\times\frac{\partial\rho_{12}}{\partial\sigma_1}\frac{\partial\rho_{12}}{\partial\sigma_2}\right]\right\}.$$

Therefore, we get the following result:

$$\frac{\partial^2}{\partial K_1\partial K_2}U(K_1,K_2)=e^{-r_€T}\left[P_1+P_2+P_3\right].$$

Note, that a numerical evaluation of the partial derivatives might be easier, especially if extended to higher dimensions.

B.7.4 *Proof that the analytic pricing formula for a best-of option simplifies to a Vanilla option*

In this section, we show that the best-of option pricing formula correctly reprices all vanilla options. Therefore, we first take $K_2 \to \infty$ in the best-of pricing formula (11.4.13):

$$\begin{aligned} v_€(0,K_1,K_2) &= e^{-r_€T}\left[\frac{F_{£/€}}{K_1}N(d_1^+,d_3^+;\rho_{13})+\frac{F_{\$/€}}{K_2}N(d_2^+,d_3^-;\rho_{23})\right. \\ &\qquad \left.+N(-d_1^-,-d_2^-;\rho_{12})-1\right]. \end{aligned}$$

It can be shown (see [73]) that the standard bivariate normal distribution $N_2(a,b;\rho)$ has the following property:

$$N_2(a,\infty;\rho)=N_1(a) \quad \text{and} \quad N_2(\infty,b;\rho)=N_1(b),$$

where $N_1(x)$ denotes the standard normal distribution. We also know the following for $K_i \to \infty$:

$$\begin{aligned} d_i^- &= \frac{\log \frac{F_i}{K_i} - \sigma_i^2 T/2}{\sigma_i \sqrt{T}} \to -\infty, \\ -d_i^- &= \frac{\log \frac{K_i}{F_i} + \sigma_i^2 T/2}{\sigma_i \sqrt{T}} \to \infty, \end{aligned}$$

and further for $K_2 \to \infty$:

$$d_3^{\pm} = \frac{\log \frac{F_1 K_2}{F_2 K_1} \pm \sigma_i^2 T/2}{\sigma_i \sqrt{T}} \to \infty,$$

if the implied volatilities $\sigma_i(K_i)$ do not grow too fast as we require in (11.4.15). Hence for $K_2 \to \infty$:

$$\begin{aligned} v_{€}(0, K_1, \infty) &= e^{-r_{€} T} \left[\frac{F_{£/€}}{K_1} N(d_1^+) + N(-d_1^-) - 1 \right] \\ &= e^{-r_{€} T} \left[\frac{F_{£/€}}{K_1} N(d_1^+) - N(d_1^-) \right], \end{aligned}$$

as it is known that $N(a) = 1 - N(-a)$. This leads us to the Garman-Kohlhagen call option price for:

$$\begin{aligned} C_{£/\$}(0; K_1) &= K_1 v_{€}(0, K_1, \infty) \\ &= F_{£/€} e^{-r_{€} T} N(d_1^+) - K_1 e^{-r_{€} T} N(d_1^-), \end{aligned}$$

which matches the formulas in Section 11.4.2. The same holds for the other currency pairs, and we omit the derivations at this point.

B.7.5 *Restrictions on volatility smiles in the model-free framework*

To ensure that the distribution function $U(K_1, K_2)$ is valid, the following (natural) restrictions on the correlation coefficient

$$\rho_{12}(K_1, K_2) = \frac{\sigma_1^2(K_1) + \sigma_2^2(K_2) - \sigma_3^2(K_3)}{2\sigma_1(K_1)\sigma_2(K_2)}$$

are necessary:

$$-1 < \rho_{ij}(K_1, K_2) < 1, \quad \forall K_1, K_2.$$

This restriction can be rewritten in terms of the corresponding volatilities σ_1, σ_2 and σ_3, which are easier to work with, when working with volatility data:

$$\begin{aligned} \frac{\sigma_1^2 + \sigma_2^2 - \sigma_3^2}{2\sigma_1\sigma_2} &< 1 \\ \Leftrightarrow \quad \sigma_1^2 + \sigma_2^2 - \sigma_3^2 &< 2\sigma_1\sigma_2 \\ \Leftrightarrow \quad \sigma_1^2 + \sigma_2^2 - 2\sigma_1\sigma_2 &< \sigma_3^2 \\ \Leftrightarrow \quad (\sigma_1 - \sigma_2)^2 &< \sigma_3^2 \\ \Leftrightarrow \quad \sigma_1 - \sigma_2 &< \sigma_3 \\ \Leftrightarrow \quad \sigma_1 &< \sigma_2 + \sigma_3. \end{aligned}$$

In the same way we get

$$\begin{aligned} -1 &< \frac{\sigma_1^2 + \sigma_2^2 - \sigma_3^2}{2\sigma_1\sigma_2} \\ \Leftrightarrow \quad -2\sigma_1\sigma_2 &< \sigma_1^2 + \sigma_2^2 - \sigma_3^2 \\ \Leftrightarrow \quad \sigma_3^2 &< \sigma_1^2 + \sigma_2^2 + 2\sigma_1\sigma_2 \\ \Leftrightarrow \quad \sigma_3^2 &< (\sigma_1 + \sigma_2)^2 \\ \Leftrightarrow \quad \sigma_3 &< \sigma_1 + \sigma_2, \end{aligned}$$

and finally

$$\begin{aligned}
&\frac{\sigma_1^2+\sigma_2^2-\sigma_3^2}{2\sigma_1\sigma_2} < 1 \\
\Leftrightarrow \quad & \sigma_1^2+\sigma_2^2-\sigma_3^2 < 2\sigma_1\sigma_2 \\
\Leftrightarrow \quad & \sigma_1^2+\sigma_2^2-2\sigma_1\sigma_2 < \sigma_3^2 \\
\Leftrightarrow \quad & (\sigma_2-\sigma_1)^2 < \sigma_3^2 \\
\Leftrightarrow \quad & \sigma_2-\sigma_1 < \sigma_3 \\
\Leftrightarrow \quad & \sigma_2 < \sigma_1+\sigma_3.
\end{aligned}$$