Computational Challenges in Risk Calculations for XVA

Likelihood Ratio Method to compute XVA Greeks for OTC derivatives

Kilian Buis

Delft University of Technology

TU Delft

Computational Challenges in Risk Calculations for XVA

Likelihood Ratio Method to compute XVA Greeks for OTC derivatives

by

Kilian Buis

Student Name	Student Number
Kilian Buis	4720806

Responsible TU Delft Supervisor:	Prof.dr.ir. C. Vuik,
Daily TU Delft Supervisor:	Dr. F. Fang,
Rabobank Supervisor:	J. Van der Linden,
Rabobank Supervisor:	S. Borst,
Defense date:	13-12-2023
Faculty:	Electrical Engineering, Mathematics & Computer Science, Delft

Acknowledgements

I would like to take this opportunity to express my gratitude to a number of people who helped me in the completion of this thesis. First of all, I would like to thank my two Rabobank supervisors, Jan van der Linden and Sebastiaan Borst, for constantly making themselves available to answer my questions. They both have a lot of knowledge about the topics that were part of my thesis, which was really helpful for me. Furthermore, I want to thank my TU Delft supervisor, Dr. Fang Fang, for her excellent guidance throughout the thesis. She always challenged me during our update meetings by asking critical questions and suggesting improvements. Lastly, I would also like to thank the whole Rabobank PMV Team. Especially during the draft presentations I gave to the team, their questions and suggestions were really helpful.

Kilian Buis
Delft, November 2023

Abstract

After the financial crisis, the standards for the valuation of financial derivatives were reviewed and several adjustments were made to these valuations, of which Credit Value Adjustment (CVA) is the most important one. CVA represents the price of counterparty credit risk that should be added to the default-free fair price of a financial derivative. Nowadays, banks and other financial institutions have dedicated CVA desks that are responsible for estimating CVA and its sensitivities to many market parameters for each counterparty. It is therefore important to have accurate and efficient methods to compute CVA and its sensitivities. This can, however, be a challenge.

Multiple methods to compute CVA sensitivities already exist in the literature, but these methods have certain drawbacks, such as a high computational expense or a payoff dependency. An efficient approach for computing CVA sensitivities is the Likelihood Ratio Method. One of the main advantages is its payoff independence, making the method applicable to sensitivities of any payoff, even payoffs with discontinuities. Furthermore, the Likelihood Ratio Method stands out by its ability to derive multiple sensitivities within a single Monte Carlo simulation. In this research, an innovative approach is developed, which involves using the Hull-White short rate model to model interest rates while applying the Likelihood Ratio Method. A drawback of the Likelihood Ratio Method is its susceptibility to high variance. To mitigate this issue, variance reduction techniques are explored, including antithetic sampling, control variates, and Quasi-Monte Carlo methods.

The performance of the Likelihood Ratio Method is compared to the Bump & Reprice method in computing first-order CVA sensitivities of three different over-the-counter derivatives. It is shown that for a certain range of model parameter values, the Likelihood Ratio Method is able to match the Bump & Reprice sensitivities. However, the Likelihood Ratio Method exhibits high variance for some extreme parameter values.

Keywords: XVA, CVA, Greeks, Monte Carlo, Quasi-Monte Carlo, Hull-White, Likelihood Ratio Method, Interest rates, Foreign exchanges

Nomenclature

In the nomenclature, we provide the reader with a list of symbols and definitions, to improve the readability of the thesis. We have split the symbols and definitions into different groups for a better overview.

Probability and statistics framework

Symbol	Definition
Ω	set of all possible outcomes representing the market.
$\mathcal{F}(t)$	filtration at time t.
$\mathbb{Q}$	risk-neutral measure.
$\mathbb{E}_s[\cdot]$	expectation operator under $\mathbb{Q}$ conditional on $\mathcal{F}(s)$.
$\text{Var}_s[\cdot]$	variance operator under $\mathbb{Q}$ conditional on $\mathcal{F}(s)$.
$\text{Cov}_s[\cdot]$	covariance operator under $\mathbb{Q}$ conditional on $\mathcal{F}(s)$.
$\Phi(\cdot)$	standard normal cumulative distribution function.
$\phi(\cdot)$	standard normal probability density function.

Measures

Symbol	Definition
$\mathbb{Q}^T$	T-forward measure.
$\lambda_{\mathbb{Q}^T}^{\mathbb{Q}}(t)$	Radon-Nikodym derivative to change from $\mathbb{Q}$ to $\mathbb{Q}^T$.
$\mathbb{Q}^f$	risk-neutral foreign measure.
$\lambda_{\mathbb{Q}}^{\mathbb{Q}^f}(t)$	Radon-Nikodym derivative to change from $\mathbb{Q}^f$ to $\mathbb{Q}$.

Interest rate framework

Symbol	Definition
$r(t)$	short rate at time t.
$V(t)$	the price of a financial derivative at time t.
$M(t)$	value at time t of one unit of currency on the money savings account.
$D(t,T)$	discount factor between two time instants $t \leq T$.
$P(t,T)$	value at time t of a zero-coupon bond with maturity T.
$F(t,T,S)$	forward rate at time t for the period $[T,S]$.
$f(t,T)$	instantaneous forward rate at time t for a future date T.
$f^M(0,T)$	instantaneous forward rate at time 0 for a future date T, observable in the market.
$P^M(0,T)$	value at time 0 of a zero-coupon bond with maturity T, observable in the market.
q	number of pillar dates and zero rates, excluding the zero rate at time 0.
Ω_{yc}	set of spine points of the yield curve.
T_j^p	j^{th} pillar date of the yield curve.
z_j	j^{th} zero rate of the yield curve.

Hull-White model

Symbol	Definition
$\chi(t)$	function to replicate the yield curve.
a	speed of mean reversion.
σ	volatility of the interest rate.
$W(t)$	Brownian motion at time t.
$x(t)$	stochastic part of $r(t)$.
$\alpha(t)$	deterministic part of $r(t)$.
$V(t,T)$	variance of the integral of both $x(t)$ and $r(t)$ from time t to time T.
$A(t,T)$	used for the notation of the zero-coupon bond price, it concerns all variables related to discounting and the variance of the integral of $x(t)$ from time t to time T.
$B(t,T)$	used for the notation of the zero-coupon bond price, it concerns the mean reversion between time t and time T.
$\hat{A}(t,T)$	alternative way of notation of the zero-coupon bond price, it concerns all variables related to discounting and the instantaneous forward rate.

Foreign exchange framework

Symbol	Definition
$y(t)$	FX spot rate at time t.
N_f	notional amount in the foreign currency of an FX digital option or FX swap.
$y_F(t,T)$	FX forward rate at time t, with maturity T.

FX-HW model

Symbol	Definition
σ_y	volatility of the FX spot rate.
$\rho_{y,d}$	correlation coefficient between the FX spot rate and the domestic short rate.
$\rho_{y,f}$	correlation coefficient between the FX spot rate and the foreign short rate.
$\rho_{y,d}$	correlation coefficient between the domestic and foreign short rate.
$L_y(t)$	logarithm of the FX spot rate.

IR and FX products

Symbol	Definition
$T*$	maximum time of the trading time horizon.
τ_i	tenor given by $T_i - T_{i-1}$ for some i.
$\ell(t;T_{i-1},T_i)$	Libor forward rate at time t for a period $[T_{i-1},T_i]$.
K	fixed interest rate of an IR/FX swap/digital option.
N_s	notional amount of an IR swap.
β	indicator whether the swap is a payer $(\beta = 1)$ or receiver $(\beta = -1)$ swap.
γ	indicator whether the option is a call $(\gamma = 1)$ or put $(\gamma = -1)$ option.
$\Sigma^2(t,T)$	instantaneous variance of the FX forward rate.
$\sigma^2_{\text{imp}}(t,T)$	implied variance of the FX forward rate.

Exposure and CVA framework

Symbol	Definition
$E(t)$	positive exposure at time t.
$\text{EE}(t_0, t)$	expected positive exposure at time t_0.
m	the number of exposure times.
T_D	time of default.
R_C	recovery rate.
$\bar{q}(T_k, T_{k-1})$	probability of default in the period (T_{k-1}, T_k).
ξ	hazard rate in the survival probability model.

Computing sensitivities

Symbol	Definition
θ	parameter of differentiation.
$Y(\theta)$	output of the model at parameter θ.
N	number of Monte Carlo simulations.
$\hat{\Delta}_{\text{C}}$	central difference estimator.
h	bump size of the central difference estimator.
$\mathbf{X}$	vector of risk factors, chosen such that $Y(\theta) = f(\mathbf{X})$ for some f.
$g_\theta(X)$	probability density function of $\mathbf{X}$, depending on θ.
$\dot{g}_\theta / g_\theta$	score function.
$I(T_k)$	integral of the short rate $r(t)$ between time 0 and T_k.
$\mu(\theta)$	expectation vector of $\mathbf{X}$.
Σ	covariance matrix of $\mathbf{X}$.
$\dot{\mu}(\theta)$	vector/matrix of derivatives of the components of μ with respect to θ.
$C(T_k, T_i)$	constant used in the derivative of $P^M(0, T_i)/P^M(0, T_k)$.
$D(T_k, T_i)$	constant used in the derivative of $P(T_k, T_i)$.

Contents

1 Introduction

In 2007, a financial crisis unfolded, originating in the United States' credit and housing market, but rapidly spreading around the world. It reached its climax in September 2008 when the investment bank Lehman Brothers filed for bankruptcy. Many financial analysts and institutions believe the crisis was mainly due to inadequate management of different risks [1]. After the crisis, regulators initiated a reassessment of the standards for the valuation of financial derivatives, and the concept of counterparty credit risk became an important element to include in these valuations [2]. Counterparty credit risk is the risk that a counterparty with whom one has entered into a financial contract will default before the expiration of the contract, and will fail to meet their side of the contractual agreement. Typical financial derivatives where counterparty credit risk is a common occurrence are over-the-counter (OTC) derivatives. These are derivatives that are privately negotiated between the two involved parties and are not traded on a centralized exchange. Examples of OTC derivatives are options and swaps.

The review of the standards of the valuation of financial derivatives resulted in several adjustments to the valuation of these derivatives, of which Credit Value Adjustment (CVA) is one of the most important and common adjustments. CVA represents the price of counterparty credit risk that should be added to the default-free fair price of a financial derivative to incorporate the counterparty's default risk into the price of the derivative. The value of CVA of a financial derivative at some time t before the expiration date T can be expressed as an expectation:

$$\text{CVA}(t) = (1 - R_C)\mathbb{E}_t \left[\frac{M(t)}{M(T_D)} \mathbb{1}_{T_D \leq T} \max(V(T_D), 0) \right], \tag{1.1}$$

where $(1 - R_C)$ is called the loss-given default, T_D is the time of the counterparty's default, the fraction $M(t)/M(T_D)$ is the discount factor between times t and T_D, and $V(T_D)$ represents the value of a derivative or portfolio at time T_D. Nowadays, managing CVA has become a key topic for banks and other financial institutions. If CVA is managed actively, then CVA sensitivities (or Greeks) are essential for hedging, explanation, and risk management. Banks have to compute CVA sensitivities to a large number of market parameters for each of its counterparties, and CVA sensitivities are also needed for Regulatory Capital under Basel III, when full revaluation is not feasible. It is therefore important to develop methods that can accurately and efficiently compute CVA and its sensitivities. Since (1.1) is a complex payoff function, CVA can usually not be priced analytically. Instead, Monte Carlo (MC) techniques are often employed to calculate CVA and its sensitivities.

Multiple methods to compute CVA sensitivities, relying on MC techniques, already exist in the literature. The simplest, the Bump & Reprice (B&R) method, employs a finite difference calculation with small shocks on each market parameter. Its strength lies in its simplicity and ease of implementation, providing a transparent way to calculate CVA sensitivities. However, for large portfolios with many underlying market factors, the B&R method can be computationally expensive. Besides the B&R method, pathwise derivative estimates are common approaches to compute CVA sensitivities. These methods provide sensitivities by direct simulation instead of finite difference approximations. A lot of research on these methods has already been done, for example by Broadie and Glasserman [3]. The advantage of these methods is their ability to calculate sensitivities of complex portfolios with many trades and underlying market factors at a much smaller computational cost than the B&R method. The main drawbacks, however, are that the pathwise derivative methods are payoff-dependent, and not always suitable for the computation of sensitivities of derivatives with discontinuous payoffs.

The goal of this thesis is to use a different method for computing first-order CVA sensitivities of OTC derivatives: the Likelihood Ratio Method (LRM). The main advantage of the method is that it is not necessary to differentiate the payoffs of financial derivatives. Instead, LRM is based on the differentiation of probability density functions. This enables LRM to compute sensitivities of derivatives with any type of payoff, including discontinuous payoffs. Moreover, multiple sensitivities can be computed in one single MC simulation, whereas other methods require separate computations for each sensitivity. However, as highlighted by Glasserman [4], a drawback of LRM is the potential for variance problems, especially when the parameter of differentiation is included in many elements required to simulate a MC path. Also, its application requires the user to choose a model, such that the probability density function of the risk factors is known. It is not guaranteed in all financial models, however, that the probability density function can be derived, posing a challenge to the practical application of the method.

Palazzi, Conti, and Pioppi [2] were the first to use LRM in the context of CVA sensitivities, focusing on giving possible solutions for CVA-related problems that may occur in real practice. In their CVA framework, they associated a normal process to each zero rate whenever stochastic interest rates were needed. After their study, they were confident that the method could be applied to the computation of CVA sensitivities, and they believe that more realistic dynamics for the underlying risk factors could be added as well. In this thesis, we extend their approach to include stochastic short rate models. As far as we know, these models have not yet been used in the CVA framework together with LRM. Much research has already been devoted to stochastic short rate models. In this thesis, we will make use of the Hull-White model, proposed in Hull and White [5], to describe the dynamics of the short rate. The main advantage of this model is that it is able to fit a given term structure of interest rates. Furthermore, this model provides closed-form solutions for various interest rate derivatives.

As previously mentioned, the complexity in the payoff of CVA (1.1) requires us to use MC techniques to compute CVA sensitivities. One of the main disadvantages of using MC techniques is its computational expense: quadrupling the number of paths only cuts the error in half. In other words, the convergence rate of ordinary MC methods is $\mathcal{O}(1/\sqrt{N})$. This means that in order to obtain a fast convergence with low variance, we would need a large number of MC simulations. Especially in the context of CVA, in which each evaluation of the payoff is computationally expensive, this poses a significant drawback. On top of that, LRM itself may introduce high variance in the simulation. To address these challenges of high variance and computational expense, multiple variance reduction techniques can be applied. In this thesis, we will employ classical techniques, such as antithetic sampling and control variates to enhance the efficiency of the simulation. Moreover, to achieve further variance reduction, we will introduce Quasi-Monte Carlo (QMC) methods. These methods have the potential to accelerate convergence from $\mathcal{O}(1/\sqrt{N})$ to almost $\mathcal{O}(1/N)$.

The remainder of this thesis is organized as follows. Section 2 provides the mathematical framework and the theory of financial markets and products. In Section 3, we outline the CVA framework and present a way to model CVA. In Section 4, we discuss the main ideas about the B&R method and LRM. Section 5 is devoted to the mathematics behind the estimation of multiple CVA sensitivities. A profound analysis of the theoretical errors and high variances that we encounter in the numerical tests is provided in Section 6. Techniques aimed at reducing these variances are presented in Section 7. Section 8 provides the results of this thesis, wherein we compare the performance of LRM to the B&R method in the computation of first-order CVA sensitivities of foreign exchange (FX) digital options and swaps, and interest rate (IR) swaps. Lastly, Section 9 concludes this thesis and outlines potential future research.

2
Mathematical Foundation and Financial Theory

This section is devoted to the mathematical framework and the theory of financial markets and products. In Section 2.1, we provide important mathematical definitions and theorems. Section 2.2 outlines important definitions regarding the interest rate (IR). We also state the Risk Neutral Pricing Theorem in this section. The dynamics of the Hull-White model, which will be used for short rate modeling, are presented in Section 2.3. In Section 2.4, we introduce the principles of foreign exchange (FX) markets and rates. Then in Section 2.5, we show the ideas behind measure changes, a concept that will be used when we describe the dynamics of the FX rates in Section 2.6 via the FX-HW model. Lastly, Section 2.7 presents the payoffs of different IR and FX derivatives.

2.1. Mathematical Foundation

In this section, we present definitions and theorems that are needed as background knowledge when reading this thesis. For some readers, this might be basic knowledge, but even then this section is important to read since it clarifies the notation that we use in this thesis. We start by stating a few definitions and theorems about stochastic processes.

2.1.1. Stochastic Processes

A stochastic process, $S(t)$, is a collection of random variables indexed by a time variable t, that represents the time. Examples of stochastic processes are stock prices or interest rates starting at some date in the past. If we look at these stochastic processes today, then we know exactly the path of the past. However, we do not know the precise path of the future. We can only simulate the future according to some distribution of our process. According to Oosterlee and Grzelak [6], the mathematical tool describing a stochastic process up to a certain time t is known as a σ-algebra[1]. The ordered sequence of σ-algebras is called a filtration. The mathematical definition is given as follows, from Redig [7].

Definition 2.1.1 (**Filtration**). *Let $(\Omega, \mathcal{F}, \mathbb{Q})$ be a probability space[2] where Ω is the set of all possible outcomes, $\mathcal{F}(t)$ is the $\sigma-$algebra, and $\mathbb{Q}$ is a probability measure. We call a sequence of σ-algebras $\{\mathcal{F}(t), t \in \mathbb{N}\}$ a filtration if*

(a) It is an increasing sequence $\mathcal{F}(0) \subset \mathcal{F}(1) \subset ... \subset \mathcal{F}(t) \subset ...$,
(b) $\mathcal{F}(t) \in \mathcal{F}$ for all $t \in \mathbb{N}$.

We call a stochastic process $S(t)$ adapted to the filtration $\{\mathcal{F}(t), t \in \mathbb{N}\}$ if for all $t \in \mathbb{N}$, $S(t)$ is a random variable which is $\mathcal{F}(t)$ measurable. In practice, a process being $\mathcal{F}(t)$ measurable means the realizations of the process are known up to time t.

[1] See Redig [7] for the definition of a σ-algebra.
[2] See Redig [7] for the definition of a probability space.

In the context of stochastic processes, the notion of martingales is important. A martingale is a process for which the best prediction of the expectation of a martingale's future value is its present value. The mathematical definition of a martingale is given as follows, from Redig [7].

Definition 2.1.2 (Martingale). *Let $(\Omega, \mathcal{F}, \mathbb{Q})$ be a probability space equipped with a filtration $\{\mathcal{F}(t), t \in \mathbb{N}\}$ under the probability measure $\mathbb{Q}$. Let $\{S(t), t \in \mathbb{N}\}$ denote an adapted process of real-valued random variables. Then $S(t)$ is a martingale w.r.t. to $\mathcal{F}(t)$ if*

(a) For all $t \in \mathbb{N}$, $S(t)$ is integrable,
(b) $\{S(t), t \in \mathbb{N}\}$ satisfies the martingale property which reads

$$\mathbb{E}^{\mathbb{Q}}[S(t) \mid \mathcal{F}(s)] = S(s) \text{ with } s \leq t.$$

Remark 2.1.1. *In definition 2.1.2, the superscript $\mathbb{Q}$ in the expectation operator indicates that the expectation is taken under the risk-neutral measure $\mathbb{Q}$. We will assume that all the expectations in this thesis are taken under the risk-neutral measure, unless stated differently, meaning for any process $S(t)$:*

$$\mathbb{E}[S(t) \mid \mathcal{F}(s)] := \mathbb{E}^{\mathbb{Q}}[S(t) \mid \mathcal{F}(s)].$$

Remark 2.1.2. *In Definition 2.1.2 we make use of a conditional expectation. Henceforth, we will use a different, more convenient, and concise notation for the conditional expectation, as well as for the conditional variance and covariance. For a stochastic process $S(t)$ and for any $s \leq t \leq T$, we write*

$$\begin{aligned}
\mathbb{E}_s[S(t)] &= \mathbb{E}[S(t) \mid \mathcal{F}(s)], \\
\text{Var}_s[S(t)] &= \text{Var}[S(t) \mid \mathcal{F}(s)], \\
\text{Cov}_s[S(t), S(T)] &= \text{Cov}[S(t), S(T) \mid \mathcal{F}(s)].
\end{aligned}$$

One of the most important martingale processes is the Wiener process or Brownian Motion. This process is commonly used in the construction of stochastic differential equations to describe asset price or interest rate movements. The definition of a Brownian Motion is given as follows, from Oosterlee and Grzelak [6].

Definition 2.1.3 (Brownian Motion). *A real-valued stochastic process, $W(t)$, is called a Brownian Motion, under the risk-neutral measure $\mathbb{Q}$, if*

(a) $W(t_0) = 0$,
(b) $W(t)$ is almost surely continuous,
(c) $W(t)$ has independent increments, i.e. for all $t_1 \leq t_2 \leq t_3 \leq t_4, W(t_2) - W(t_1) \perp W(t_4) - W(t_3)$, with distribution $W(t) - W(s) \sim \mathcal{N}(0, t-s)$ for $t_0 \leq s < t$ under the risk neutral measure $\mathbb{Q}$.

Proposition 2.1.3. *A Brownian Motion $W(t), t \in [0, T]$ is a martingale.*

Proof. For the proof of the proposition, we refer to Oosterlee and Grzelak [6]. □

2.1.2. Itô Calculus

In this section, we present two very important concepts that we will encounter frequently throughout this thesis, namely Itô's Lemma and Itô's Isometry. Itô's Lemma can be used to derive a solution for stochastic differential equations and is given as follows.

Lemma 2.1.4 (Itô's Lemma). *Suppose a process $S(t)$ follows the following dynamics*

$$dS(t) = \bar{\mu}(t, S(t))dt + \bar{\sigma}(t, S(t))dW(t), \quad \text{with } S(t_0) = S_0,$$

where $\bar{\mu}(t, S(t))$ and $\bar{\sigma}(t, S(t))$ satisfy the standard Lipschitz conditions[3]. Let $g(t, S)$ be a function of $S = S(t)$ and time t, with continuous partial derivatives $\partial g/\partial S, \partial^2 g/\partial S^2, \partial g/\partial t$. A stochastic variable $Y(t) := g(t, S)$ then has the following dynamics

$$dY(t) = \left(\frac{\partial g}{\partial t} + \bar{\mu}(t, S)\frac{\partial g}{\partial S} + \frac{1}{2}\frac{\partial^2 g}{\partial S^2}\bar{\sigma}^2(t, S)\right)dt + \frac{\partial g}{\partial S}\bar{\sigma}(t, S)dW(t).$$

Proof. For the proof of the lemma, we refer either to Oosterlee and Grzelak [6] or Shreve [8]. □

[3] See Oosterlee and Grzelak [6] for the standard Lipschitz conditions.

Itô's Isometry helps us to determine the integral with respect to a Brownian Motion. It is given as follows.

Theorem 2.1.5 (Itô's Isometry). *Let $S(t)$ be a stochastic process that is adapted to the filtration $\{\mathcal{F}(t), t \in \mathbb{N}\}$, and let $W(t)$ denote a Brownian Motion. Then for $t_0 < T$*

$$\mathbb{E}\left[\left(\int_{t_0}^{T} S(t)dW(t)\right)^2\right] = \int_{t_0}^{T} \mathbb{E}\left[S^2(t)\right] dt.$$

Proof. For the proof of the lemma, we refer to Oosterlee and Grzelak [6]. □

2.2. Interest Rate Theory

In this section, important definitions related to the interest rate are presented. We also state the Risk Neutral Pricing Theorem, and we devote attention to yield curves.

2.2.1. Short rates and Risk Neutral Pricing Theorem

There are multiple ways to describe the interest rate, but in regard to pricing interest rate derivatives, we will need the definition of the short rate.

Definition 2.2.1 (Short rate). *The short rate $r(t)$ is the (continuously compounded) interest rate at which one can borrow money for an infinitesimally period of time t.*

In the context of pricing financial derivatives, the Risk Neutral Pricing Theorem is a very important theorem. The theorem assumes that a financial derivative, with payoff function $V(T)$ at maturity time T, is defined under the risk-neutral measure $\mathbb{Q}$, and it aims to find the value of the derivative at time $t \leq T$. It does that by discounting the payoff from time T to time t. Discounting can be done using the money savings account. Money in a savings account grows at each time t at a rate $r(t)$. More formally, the definition of the money savings account is as follows.

Definition 2.2.2 (Money savings account). *Let $M(t)$ be the value at time t of one unit of currency on the money savings account, then*

$$dM(t) = r(t)M(t)dt.$$

Using the boundary condition $M(0) = 1$, gives

$$M(t) = \exp\left(\int_0^t r(s)ds\right). \tag{2.1}$$

The Risk Neutral Pricing Theorem is then given as follows.

Theorem 2.2.1 (Risk Neutral Pricing Theorem). *Let $V(T)$ be a $\mathcal{F}(T)$ measurable random variable under the risk-neutral measure $\mathbb{Q}$, that represents the payoff of a financial derivative at time T. Let $M(t)$ be the value of the money savings account at time t. Then the price of the financial derivative at time $t \leq T$ is given by*

$$V(t) = M(t)\mathbb{E}_t^{\mathbb{Q}}\left[\frac{V(T)}{M(T)}\right]. \tag{2.2}$$

Proof. For the proof of the theorem, we refer either to Shreve [8] or Oosterlee and Grzelak [6]. □

In (2.2) we have already seen the concept of stochastic discount factors. The formal definition, stated by Brigo and Mercurio [9], is as follows.

Definition 2.2.3 (Discount factor). *The stochastic discount factor $D(t,T)$ between two time instants $t \leq T$ is the amount at time t that is equivalent to one unit of currency payable at time T and is given by*

$$D(t,T) = \frac{M(t)}{M(T)} = \exp\left(-\int_t^T r(s)ds\right). \tag{2.3}$$

A fundamental concept in pricing interest rate derivatives is the zero-coupon bond. According to Brigo and Mercurio [9], a zero-coupon bond is a financial contract with the following definition.

Definition 2.2.4 (**Zero-coupon bond**). *A T-maturity zero-coupon bond (ZCB) is a contract that guarantees its holder the payment of one unit of currency at time T with no intermediate payments. The contract value at time $t \leq T$ is denoted by $P(t,T)$. Clearly, we have that $P(T,T) = 1$ for all T.*

Using the Risk Neutral Pricing Theorem and the fact that $P(T,T) = 1$, it is easy to see that

$$P(t,T) = \mathbb{E}\left[D(t,T)\right] = \mathbb{E}\left[\exp\left(-\int_t^T r(s)ds\right)\right], \tag{2.4}$$

meaning that, for stochastic interest rates, $D(t,T)$ is a random variable that depends on the evolution of the short rate between t and T, while the ZCB price $P(t,T)$ is a value of a contract at time t and hence has to be deterministic at time t.

As previously mentioned, there are multiple ways to describe interest rates. Besides the short rate, we can also define the so-called forward rate. It is related to the ZCB price and can be thought of as the rate at which interest grows from time T to time S, based on the market conditions at time t. The formal definition is given as follows, as given by Brigo and Mercurio [9].

Definition 2.2.5 (**Forward rate**). *Let $0 \leq t < T < S$. Then the forward rate $F(t,T,S)$ is defined as*

$$F(t,T,S) = \frac{1}{P(t,S)}\frac{P(t,T) - P(t,S)}{S-T} = \frac{1}{S-T}\left(\frac{P(t,T)}{P(t,S)} - 1\right).$$

The instantaneous forward rate $f(t,T)$ is defined as

$$f(t,T) = \lim_{S \to T^+} F(t,T,S) = -\frac{\partial}{\partial T}\log(P(t,T)). \tag{2.5}$$

We can rewrite (2.5) as

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right),$$

meaning that the short rate $r(t)$ can be defined as the limit of the instantaneous forward rate $r(t) = f(t,t)$. Note that for any $T > 0$, the ZCB price at time $t = 0$ can be obtained via (2.4), but can also be directly related to today's yield curve via

$$f^M(0,T) = -\frac{\partial}{\partial T}\log\big(P^M(0,T)\big). \tag{2.6}$$

Here we added the superscript M to indicate that these quantities can be obtained from the market.

2.2.2. Yield curve

A yield curve is a curve that consists of bond yields. The bond yield corresponds with the interest rate at which bonds are traded in the market. The yield is different for each bond maturity, giving rise to the existence of the yield curve.

Oosterlee and Grzelak [6] point out that a yield curve is one of the most important elements needed to determine the present value of future cash flows. A yield curve is constructed by mapping a discrete set of instruments, depending on the market, to a discrete set of input nodes of the curve Ω_{yc}. The input of the curve is a set of $q+1$ pillar dates $\{T_0^p, ..., T_q^p\}$ corresponding to a set of zero rates $\{z_0, ..z_q\}$. The yield curve is then represented by a set of pairs as

$$\Omega_{yc} = \{(T_0^p, P^M(0,T_0^p)), (T_1^p, P^M(0,T_1^p))..., (T_q^p, P^M(0,T_q^p))\},$$

where $P^M(0,t)$ represents the ZCB price at time 0 with maturity t. We know that $P^M(0,T_0^p) = 1$. For the other bond prices, we have a relation between the ZCB prices and the zero rates:

$$P^M(0,T_j^p) = \exp\left(-z_j T_j^p\right), \quad j \in \{1, ..., q\}. \tag{2.7}$$

The discrete set Ω_{yc} is called the set of spine points of the yield curve. For points that are not part of the discrete set, we can use an interpolation scheme to obtain a continuous function $P(0,t)$ for each t. For some point T_l that lies between two pillar dates, T_{j-1}^p and T_j^p with corresponding zero rates z_{j-1} and z_j, we use the convention that

$$P^M(0,T_l) = \exp\left(-\left(z_{j-1} + \frac{z_j - z_{j-1}}{T_j^p - T_{j-1}^p}\left(T_l - T_{j-1}^p\right)\right)(T_l - T_0)\right). \tag{2.8}$$

Then, the yield curve spine points together with the interpolation function define the yield curve. It is important to have an accurate yield curve since it is used for the pricing of financial derivatives, hedging, and risk management.

Recall the relation between the instantaneous forward rate and the ZCB prices that we observe in the market (2.6). By applying the interpolation scheme (2.8) it is possible for the instantaneous forward rate (2.6) to be analytically known.

Theorem 2.2.2. *Let $q+1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T_0^p, T_1^p, ..., T_q^p$. We assume that $T_{j-1} \leq t < T_j$ for some $j \in \{1, ..., q\}$ and we let $P^M(0,t)$ be the ZCB price with maturity at t valued at time 0. Then*

$$\begin{aligned} f^M(0,t) &= -\frac{\partial}{\partial t}\log\left(P^M(0,t)\right), \\ &= z_{j-1} + \frac{z_j - zj - 1}{T_j^p - T_{j-1}^p}(2t - T_{j-1}^p - T_0). \end{aligned}$$

Proof. For the proof of the theorem, we refer to Appendix A.1.1. □

2.3. Hull-White Model

In this section, we present the dynamics of the Hull-White model. We show that we can closed-form solutions for the short rate, its integrated part, but also for the ZCB prices.

Since the short rate is a stochastic process over time, a model is needed that describes the path of the short rate over time. One of the first and most successful short rate models was developed by Vasicek [10]. Vasicek assumed that the short rate $r(t)$ followed an Ornstein-Uhlenbeck process with constant coefficients. The main drawback of the Vasicek model, as mentioned by Brigo and Mercurio [9], is the poor fitting of the initial term structure of interest rates. Many extensions of the Vasicek model have been developed. By considering a time-varying parameter, Hull and White [5] proposed a model that is able to fit a given term structure of interest rates. That implies that the model uses observable market data, such as today's yield curve, as input and always provides us with the ZCB prices that precisely match those as given in the market. Furthermore, the Hull-White model is driven by a mean-reversion parameter, which aligns with the economic concept that interest rates have the tendency to return to a long-term average. This avoids extreme movements in the interest rate. These two features make the Hull-White Model very attractive to model short rate processes.

In the Hull-White model the short rate $r(t)$ is driven by the following stochastic differential equation (SDE):

$$dr(t) = \left[\chi(t) - a(t)r(t)\right]dt + \sigma(t)dW(t),$$

where $\chi(t)$ is a function to replicate the initial yield curve, $a(t)$ denotes the speed of mean reversion, $\sigma(t)$ is the volatility of the interest rate, and $W(t)$ denotes a standard Brownian Motion. Note that $\chi(t), a(t)$ and $\sigma(t)$ are functions of time. For simplicity, we focus on the extension where $a(t)$ and $\sigma(t)$ are constants, i.e. $a(t) \equiv a$ and $\sigma(t) \equiv \sigma$. Therefore, we rewrite the SDE as

$$dr(t) = \left[\chi(t) - ar(t)\right]dt + \sigma dW(t). \tag{2.9}$$

Solving this SDE provides us with the dynamics of the short rate $r(t)$.

2.3.1. Short rate expression

The aim of this section is to find the dynamics of the short rate $r(t)$. We start by solving (2.9) and we find the following.

Theorem 2.3.1. *Let $0 \leq s \leq t$. Then the solution of* (2.9)*, and hence the closed-form solution for the short rate, is given by*

$$r(t) = r(s)e^{-a(t-s)} + \int_s^t e^{-a(t-u)}\chi(u)du + \int_s^t \sigma e^{-a(t-u)}dW(u). \tag{2.10}$$

Proof. For the proof of the theorem, we refer to Appendix A.1.2. □

If we define a filtration $\mathcal{F}(s)$ for some $s \geq 0$, then we can find the distribution of the short rate process $r(t)$ given $\mathcal{F}(s)$.

Theorem 2.3.2. *Let $0 \leq s \leq t$. The short rate process $r(t)$* (2.10) *given $\mathcal{F}(s)$ is normally distributed with mean and variance given by*

$$\mathbb{E}_s\left[r(t)\right] = r(s)e^{-at} + \int_s^t e^{-a(t-u)}\chi(u)du, \tag{2.11}$$

$$\begin{aligned} \text{Var}_s\left[r(t)\right] &= \int_s^t \sigma^2 e^{-2a(t-u)}du \\ &= \frac{\sigma^2}{2a}(1 - e^{-2a(t-s)}). \end{aligned} \tag{2.12}$$

Proof. Since the increments of the Brownian Motion $W(t)$ are independent and follow a normal distribution, we get that $r(t)$ follows a normal distribution as well. Using Itô's Isometry (Theorem 2.1.5), finding the mean and variance is trivial. □

Note that the expectation of the short rate (2.11) contains an integral, that is not easy to solve due to the term $\chi(t)$. Therefore, we proceed in a different way to find the dynamics of $r(t)$. We rewrite and obtain

$$\begin{aligned} r(t) &= x(t) + \alpha(t), \\ dx(t) &= -ax(t)dt + \sigma dW(t), \quad x(0) = 0. \end{aligned}$$

Then $r(t)$ is split into a stochastic part $x(t)$ and a deterministic part $\alpha(t)$, that is given by an exact fitting to the term structure. In a similar way as we did for $r(t)$ we can find that $x(t)$ is given by

$$x(t) = x(s)e^{-a(t-s)} + \int_s^t \sigma e^{-a(t-u)}dW(u). \tag{2.13}$$

The stochastic part $x(t)$ also follows a normal distribution, with a different mean, but with the same variance as $r(t)$.

Theorem 2.3.3. *Let $0 \leq s \leq t$. The process $x(t)$* (2.13) *given $\mathcal{F}(s)$ is normally distributed with mean and variance given by*

$$\begin{aligned} \mathbb{E}_s\left[x(t)\right] &= x(s)e^{-a(t-s)}, \\ \text{Var}_s\left[x(t)\right] &= \int_s^t \sigma^2 e^{-2a(t-u)}du \\ &= \frac{\sigma^2}{2a}(1 - e^{-2a(t-s)}). \end{aligned}$$

Moreover, for any $0 \leq s \leq t, T$, the covariance of $x(t)$ and $x(T)$ given $\mathcal{F}(s)$ is given by

$$\text{Cov}_s\left[x(t), x(T)\right] = \begin{cases} \frac{\sigma^2}{2a}\left(e^{-a(T-t)} - e^{-a(t+T-2s)}\right) & t \leq T, \\ \frac{\sigma^2}{2a}\left(e^{-a(t-T)} - e^{-a(T+t-2s)}\right) & t > T. \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.1.3. □

Besides the dynamics of $x(t)$, we also want a closed-form solution for the deterministic part $\alpha(t)$.

Theorem 2.3.4. *For $t \geq 0$, the deterministic part $\alpha(t)$ is given by*

$$\begin{aligned}\alpha(t) &= f^M(0,t) + \int_0^t \frac{\sigma^2}{a}\left(1-e^{-a(t-s)}\right)e^{-a(t-s)}ds \\ &= f^M(0,t) + \frac{\sigma^2}{2a^2}\left(1-e^{-at}\right)^2,\end{aligned} \tag{2.14}$$

where $f^M(0,t)$ is given in (2.5) *as*

$$f^M(0,t) = -\frac{\partial}{\partial t}\log\left(P^M(0,t)\right).$$

Proof. For the proof of the theorem, we refer to Di Francesco [11]. □

We can now combine the dynamics of $x(t)$ and the expression for $\alpha(t)$, to obtain a second expression for the short rate, and its expectation. Note that the expression for the variance (2.12) does not change.

Theorem 2.3.5. *Let $0 \leq s \leq t$. The short rate process $r(t)$* (2.10) *is given by*

$$r(t) = x(s)e^{-a(t-s)} + \int_s^t \sigma e^{-a(t-u)}dW(u) + f^M(0,t) + \frac{\sigma^2}{2a^2}\left(1-e^{-at}\right)^2. \tag{2.15}$$

The short rate process $r(t)$ given $\mathcal{F}(s)$ is normally distributed with mean and variance given by

$$\begin{aligned}\mathbb{E}_s\left[r(t)\right] &= x(s)e^{-a(t-s)} + f^M(0,t) + \frac{\sigma^2}{2a^2}\left(1-e^{-at}\right)^2, \\ \text{Var}_s\left[r(t)\right] &= \frac{\sigma^2}{2a}(1-e^{-2a(t-s)}).\end{aligned}$$

Moreover, for any $0 \leq s \leq t, T$, the covariance of $x(t)$ and $x(T)$ given $\mathcal{F}(s)$ is given by

$$\text{Cov}_s\left[r(t), r(T)\right] = \begin{cases} \frac{\sigma^2}{2a}\left(e^{-a(T-t)} - e^{-a(t+T-2s)}\right) & t \leq T, \\ \frac{\sigma^2}{2a}\left(e^{-a(t-T)} - e^{-a(T+t-2s)}\right) & t > T. \end{cases}$$

Proof. Simply combine the expressions in Theorem 2.3.3 and Theorem 2.3.4. □

2.3.2. Integral expression

The aim of this section is to find the dynamics of the integral of the short rate $\int_t^T r(s)ds$ for $0 \leq t \leq T$. This requires us to find the dynamics and solution of $\int_t^T x(s)ds$ and $\int_t^T \alpha(s)ds$. We start by stating the dynamics of $\int_t^T x(s)ds$.

Theorem 2.3.6. *The integral of the process $x(t)$ can be expressed as*

$$\int_t^T x(s)ds = x(t)\frac{1-e^{-a(T-t)}}{a} + \int_t^T \frac{\sigma}{a}\left(1-e^{-a(T-u)}\right)dW(u). \tag{2.16}$$

Then $\int_t^T x(s)ds$ given $\mathcal{F}(t)$ is normally distributed with variance and mean given by

$$\begin{aligned}V(t,T) = \text{Var}_t\left[\int_t^T x(s)ds\right] &= \int_t^T \frac{\sigma^2}{a^2}\left(1-e^{-a(T-u)}\right)^2 du \\ &= \frac{\sigma^2}{a^2}\left((T-t) - \frac{2}{a}\left(1-e^{-a(T-t)}\right) + \frac{1}{2a}\left(1-e^{-2a(T-t)}\right)\right),\end{aligned} \tag{2.17}$$

$$\mathbb{E}_t\left[\int_t^T x(s)ds\right] = x(t)\frac{1-e^{-a(T-t)}}{a}.$$

Proof. For the proof of the theorem, we refer to Di Francesco [11]. □

Theorem 2.3.7. *The integral of the process $\alpha(t)$ (2.14) be expressed as*

$$\int_t^T \alpha(s)ds = \log\left(\frac{P^M(0,t)}{P^M(0,T)}\right) + \frac{V(0,T)-V(0,t)}{2}, \tag{2.18}$$

where $V(t,T)$ is given by (2.17).

Proof. For the proof of the theorem, we refer to Appendix A.1.4. □

Combining (2.16) and (2.18), provides us with a closed form of $\int_t^T r(s)ds$ as follows:

$$\int_t^T r(s)ds = x(t)\frac{1-e^{-a(T-t)}}{a} + \int_t^T \frac{\sigma}{a}\left(1-e^{-a(T-u)}\right)dW(u) + \log\left(\frac{P^M(0,t)}{P^M(0,T)}\right) + \frac{V(0,T)-V(0,t)}{2}. \tag{2.19}$$

Finding the dynamics of the integral is then easy.

Theorem 2.3.8. *The integral of the short rate process $\int_t^T r(s)ds$ given $\mathcal{F}(t)$ is normally distributed with mean and variance given by*

$$\mathbb{E}_t\left[\int_t^T r(s)ds\right] = x(t)\frac{1-e^{-a(T-t)}}{a} + \log\left(\frac{P^M(0,t)}{P^M(0,T)}\right) + \frac{V(0,T)-V(0,t)}{2}, \tag{2.20}$$

$$\text{Var}_t\left[\int_t^T r(s)ds\right] = V(t,T), \tag{2.21}$$

where $V(t,T)$ is given by (2.17).

Proof. Since $\int_t^T x(s)ds$ follows a normal distribution and $\int_t^T \alpha(s)ds$ is deterministic, we conclude that $\int_t^T r(s)ds$ follows a normal distribution as well. Using Itô's Isometry (Theorem 2.1.5), finding the mean and variance is trivial. □

We will now state the structure of the covariance of integrated short rates between different time points. We will fix the filtration to $\mathcal{F}(0)$, meaning we only look at the information available today. The covariance is then given as follows.

Theorem 2.3.9. *For any $0 \leq t, T$, the covariance of $\int_0^t r(u)du$ and $\int_0^T r(u)du$ given $\mathcal{F}(0)$ is given by*

$$\text{Cov}_0\left[\int_0^t r(u)du, \int_0^T r(u)du\right] = \begin{cases} \frac{\sigma^2}{a^2}\left(t - \frac{1-e^{-at}+e^{-a(T-t)}-e^{-aT}}{a} + \frac{e^{-a(T-t)}-e^{-a(t+T)}}{2a}\right) & t \leq T, \\ \frac{\sigma^2}{a^2}\left(T - \frac{1-e^{-aT}+e^{-a(t-T)}-e^{-at}}{a} + \frac{e^{-a(t-T)}-e^{-a(T+t)}}{2a}\right) & t > T, \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.1.5. □

We will see later that it is important to have analytical formulas for the cross-covariance terms between the short rate and its integrated part. The cross-covariance between different time points is given as follows.

Theorem 2.3.10. *For any $0 \leq t, T$, the covariance of $r(t)$ and $\int_0^T r(u)du$ given $\mathcal{F}(0)$ is given by*

$$\text{Cov}_0\left[r(t), \int_0^T r(u)du\right] = \begin{cases} \frac{\sigma^2}{a}\left(\frac{1}{a}\left(1-e^{-at}\right) - \frac{1}{2a}\left(e^{-a(T-t)}-e^{-a(t+T)}\right)\right) & t \leq T, \\ \frac{\sigma^2}{a}\left(\frac{1}{a}\left(e^{-a(t-T)}-e^{-at}\right) - \frac{1}{2a}\left(e^{-a(t-T)}-e^{-a(T+t)}\right)\right) & t > T, \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.1.6. □

2.3.3. Zero-coupon bond expression

The aim of this section is to present a closed-form expression for the ZCB price $P(t,T)$, defined in (2.4). The fact that the ZCB prices can be represented as a closed-form expression makes the computation of these bonds very efficient. The ZCB price can be expressed in terms of the stochastic part $x(t)$ (2.13) as follows.

Theorem 2.3.11. *For $0 \leq t < T$, the ZCB price $P(t,T)$ is given by*

$$\begin{aligned} P(t,T) &= \mathbb{E}_t\left[e^{-\int_t^T r(s)ds}\right] \\ &= A(t,T)e^{-B(t,T)x(t)}, \end{aligned} \tag{2.22}$$

where

$$\begin{aligned} A(t,T) &= \frac{P^M(0,T)}{P^M(0,t)} \exp\left(\frac{1}{2}(V(t,T) - V(0,T) + V(0,t))\right), \\ B(t,T) &= \frac{1-e^{-a(T-t)}}{a}, \end{aligned}$$

and $V(t,T)$ is given by (2.17)*.*

Proof. For the proof of the theorem, we refer to Di Francesco [11]. □

On the other hand, we can also express the ZCB price in terms of the short rate $r(t)$ (2.15) as follows.

Theorem 2.3.12. *For $0 \leq t < T$, the ZCB price $P(t,T)$ is given by*

$$\begin{aligned} P(t,T) &= \mathbb{E}_t\left[e^{-\int_t^T r(s)ds}\right] \\ &= \hat{A}(t,T)e^{-B(t,T)r(t)}, \end{aligned} \tag{2.23}$$

where

$$\begin{aligned} \hat{A}(t,T) &= \frac{P^M(0,T)}{P^M(0,t)} \exp\left(B(t,T)f^M(0,t) - \frac{\sigma^2}{4a}(1-e^{-2at})B(t,T)^2\right), \\ B(t,T) &= \frac{1-e^{-a(T-t)}}{a}, \end{aligned}$$

and $f^M(0,t)$ and $V(t,T)$ are given by (2.5) *and* (2.17)*, respectively.*

Proof. For the proof of the theorem, we refer to Brigo and Mercurio [9]. □

2.4. Foreign Exchange Rate Theory

This section is devoted to the theory of FX rates. This theory is needed to formulate a model to simulate FX rates. The notation in this section is in line with the notation of Oosterlee and Grzelak [6].

2.4.1. FX spot rate and FX forward rate

The FX market is the financial market that serves as a global platform where different currencies can be bought and sold. The primary purpose of the FX market is to enable the conversion of one currency into another currency. The pricing of currencies in the FX market is determined by supply and demand dynamics, influenced by e.g. interest rates. Exchange rates, which represent the relative value of one currency against another, fluctuate constantly as participants react to new information and adjust their trading strategies accordingly.

In order to model the exchange rates, we consider a domestic currency (e.g., the euro €) and a foreign currency (e.g., the dollar $). We define the FX spot rate as $y(t) := y_f^d(t)$. With this notation, we express the units of the domestic currency per unit of a foreign currency. If we define the notional in foreign currency by N_f, then this amount equals $N_f \cdot y(t)$ in the domestic currency.

According to Oosterlee and Grzelak [6] the FX forward is one of the most liquidly traded FX products. This contract can be seen as a mechanism for effectively borrowing and lending the same value in two different currencies, with the resulting proceeds converted into domestic currency. Using the forward contract we define the FX forward rate as follows.

Definition 2.4.1. *Let $r_d(t)$ be the short rate of the domestic currency, and $r_f(t)$ be the short rate of the foreign currency. Suppose that T is the maturity date of an FX forward contract, then we define the FX forward rate at time t by*

$$y_F(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[y(T)\right] = y(t)\frac{P_f(t,T)}{P_d(t,T)}, \tag{2.24}$$

where $P_f(t,T)$ and $P_d(t,T)$ are the ZCB prices in the foreign and domestic currencies. $P_d(t,T)$ is given by (2.23)*. We see later that $P_f(t,T)$ will be given by* (2.42)*.*

That means that the forward exchange rate $y_F(t,T)$ is defined as the expectation of the future spot exchange rate $y(T)$.

2.4.2. Multiple yield curves

In Section 2.2.2 we already explained the concept of a yield curve and showed how a yield curve is constructed. In the realm of FX markets, where two currencies are involved, we encounter two distinct yield curves. The construction is exactly the same, only the notation is slightly different.

For both yield curves we use the same set of $q+1$ pillar dates $\{T_0^p,...,T_q^p\}$. For the domestic currency, we use a set of zero rates $\{z_0^d,..z_q^d\}$. The yield curve for the domestic currency is then represented by a set of pairs as

$$\Omega_{yc}^d = \{(T_0^p, P_d^M(0,T_0^p)),(T_1^p, P_d^M(0,T_1^p))...,(T_q^p, P_d^M(0,T_q^p))\},$$

where $P_d^M(0,t)$ represents the domestic ZCB price at time 0 with maturity t. We know that $P_d^M(0,T_0^p) = 1$. For the other bond prices, we have a relation between the ZCB prices and the zero rates:

$$P_d^M(0,T_j^p) = \text{exp}\left(-z_j^d T_j^p\right), \quad j \in \{1,...,q\}.$$

For the foreign yield curve we use again a set of zero rates, labelled $\{z_0^f,..z_q^f\}$. The yield curve for the foreign currency is then represented by a set of pairs as

$$\Omega_{yc}^f = \{(T_0^p, P_f^M(0,T_0^p)),(T_1^p, P_f^M(0,T_1^p))...,(T_q^p, P_f^M(0,T_q^p))\},$$

where $P_f^M(0,t)$ represents the domestic ZCB price at time 0 with maturity t. We know that $P_f^M(0,T_0^p) = 1$. Again, for the other bond prices, we have a relation between the ZCB prices and the zero rates:

$$P_f^M(0,T_j^p) = \text{exp}\left(-z_j^f T_j^p\right), \quad j \in \{1,...,q\}.$$

For points that are not part of the discrete sets, we can use an interpolation scheme to get a continuous function. We use for some point T_l that lies between two pillar dates, T_{j-1}^p and T_j^p with corresponding zero rates given by either z_{j-1}^d and z_j^d, or z_{j-1}^f and z_j^f, that

$$P_d^M(0,T_l) = \text{exp}\left(-\left(z_{j-1}^d + \frac{z_j^d - z_{j-1}^d}{T_j^p - T_{j-1}^p}\left(T_l - T_{j-1}^p\right)\right)(T_l - T_0)\right), \tag{2.25}$$

$$P_f^M(0,T_l) = \text{exp}\left(-\left(z_{j-1}^f + \frac{z_j^f - z_{j-1}^f}{T_j^p - T_{j-1}^p}\left(T_l - T_{j-1}^p\right)\right)(T_l - T_0)\right). \tag{2.26}$$

2.5. Change of Measure Theory

In this section, we show how to change from one measure to another measure. The concept of changing measures has applications in many different fields, especially in the financial setting, where it is often used in the pricing of derivatives. Here, we specifically show how to change from the risk-neutral measure $\mathbb{Q}$ to the T-forward measure $\mathbb{Q}^T$, and from the risk-neutral foreign measure $\mathbb{Q}^f$ to the risk-neutral domestic measure $\mathbb{Q}^d$. Our aim is to use the change of measures to find today's price of a derivative.

The price of a derivative, V, is usually known at the maturity date T. However, in most cases, we are interested in the price today. We have already shown in the Risk Neutral Pricing Theorem (Theorem 2.2.1) that we can find today's price $V(t_0)$ by discounting the future price $V(T)$ under the risk-neutral measure:

$$V(t_0) = M(t_0)\mathbb{E}^{\mathbb{Q}}_{t_0}\left[\frac{V(T)}{M(T)}\right],$$

where $M(t)$ is given by the money savings account (2.1). In this case, we call the money savings account $M(t)$ the numérarie of the risk-neutral measure $\mathbb{Q}$. With numéraire we mean the basic standard by which values are measured.

In some cases, the payoff function function $V(T)$ is known under a different measure, such as the T-forward measure $\mathbb{Q}^T$. It is then convenient to discount the future price $V(T)$ under this T-forward measure. Consequently, the ability to seamlessly switch between the T-forward measure and the risk-neutral measure measures becomes essential.

The T-forward measure uses the zero coupon bond price $P(t,T)$ as numéraire. This implies that the relation between today's price $V(t_0)$ and future price $V(T)$ is expressed as

$$V(t_0) = P(t_0,T)\mathbb{E}^{T}_{t_0}\left[\frac{V(T)}{P(T,T)}\right],$$

where the superscript T in the expectation operator indicates that the expectation is taken under the T-forward measure $\mathbb{Q}^T$.

Changing from the risk-neutral measure to T-forward measure, or vice-versa, requires us to use the Radon-Nikodym derivative. According to Oosterlee and Grzelak [6], the Radon-Nikodym derivative is defined by the ratio of the numéraires of the two measures:

$$\lambda^{\mathbb{Q}}_{\mathbb{Q}^T}(t) := \left.\frac{d\mathbb{Q}}{d\mathbb{Q}^T}\right|_{\mathcal{F}(t)} = \frac{M(T)P(t,T)}{M(t)P(T,T)}.$$

Then if we want to compute today's price of the financial derivative under the risk-neutral measure $\mathbb{Q}$, but the future price $V(T)$ is known under the T-forward measure, we can use the Radon-Nikodym derivative to change the measure under which the expectation is taken:

$$\begin{aligned} V(t_0) = \mathbb{E}^{\mathbb{Q}}_{t_0}\left[\frac{M(t_0)}{M(T)}V(T)\right] &= \mathbb{E}^{T}_{t_0}\left[\frac{M(t_0)}{M(T)}V(T)\lambda^{\mathbb{Q}}_{\mathbb{Q}^T}(t_0)\right] \\ &= \mathbb{E}^{T}_{t_0}\left[\frac{M(t_0)}{M(T)}V(T)\frac{M(T)P(t_0,T)}{M(t_0)P(T,T)}\right] \\ &= P(t_0,T)\mathbb{E}^{T}_{t_0}\left[V(T)\right]. \end{aligned}$$

In this way, today's price $V(t_0)$ can still be computed, although the future value $V(T)$ is unknown under the risk-neutral measure $\mathbb{Q}$.

A similar change of measure can be implemented in the FX markets, where multiple currencies are involved, each governed by its own risk-neutral measure. We consider two currencies; the domestic currency with corresponding risk-neutral domestic measure $\mathbb{Q}$, and the foreign currency with corresponding risk-neutral foreign measure $\mathbb{Q}^f$. Typically, processes in the foreign currency are discounted under the risk-neutral foreign measure. However, these processes are sometimes known under the risk-neutral domestic measure. Therefore, it is convenient to be able to change between the risk-neutral foreign measure and the risk-neutral domestic measure.

We again define the Radon-Nikodym derivative as the ratio of the numéraires of both measures. According to Brigo and Mercurio [9], the Radon-Nikodym derivative to change between the foreign measure $\mathbb{Q}^f$ and the domestic measure $\mathbb{Q}$ is given by

$$\lambda_{\mathbb{Q}}^{\mathbb{Q}^f}(t) := \frac{d\mathbb{Q}^f}{d\mathbb{Q}}\bigg|_{\mathcal{F}(t)} = \frac{y(T)M_f(T)M_d(t)}{y(t)M_d(T)M_f(t)}.$$

Then if we want to compute today's price of a financial derivative under the risk-neutral foreign measure $\mathbb{Q}^f$, but the future price $V(T)$ is known under the risk-neutral domestic measure $\mathbb{Q}$, we can use the Radon-Nikodym derivative to change the measure under which the expectation is taken:

$$\begin{aligned}
V(t_0) = \mathbb{E}_{t_0}^{\mathbb{Q}^f}\left[\frac{M_f(t_0)}{M_f(T)}V(T)\right] &= \mathbb{E}_{t_0}^{\mathbb{Q}}\left[\frac{M_f(t_0)}{M_f(T)}V(T)\lambda_{\mathbb{Q}}^{\mathbb{Q}^f}(t_0)\right] \\
&= \mathbb{E}_{t_0}^{\mathbb{Q}}\left[\frac{M_f(t_0)}{M_f(T)}V(T)\frac{y(T)M_f(T)M_d(t_0)}{y(t_0)M_d(T)M_f(t_0)}\right] \\
&= \frac{M_d(t_0)}{y(t_0)}\mathbb{E}_{t_0}^{\mathbb{Q}}\left[V(T)\frac{y(T)}{M_d(T)}\right].
\end{aligned}$$

In this way, today's price $V(t_0)$ can still be computed, although the future value $V(T)$ is unknown under the risk-neutral foreign measure $\mathbb{Q}^f$.

We will present applications of measure changes in the following two chapters.

2.6. FX-HW model

In this section, we will find the dynamics of the FX spot rate. To do so, we will make use of the FX-HW (Foreign Exchange Hull-White) model. In this model, we have two short rate processes, $r_d(t)$ and $r_f(t)$, that follow a Hull-White process, and that are correlated with the FX spot process, $y(t)$. This results in a system of three processes. The model is given as follows.

$$\begin{aligned}
dy(t) &= (r_d(t) - r_f(t))y(t)dt + \sigma_y y(t)dW_y^{\mathbb{Q}}(t), \\
dr_d(t) &= [\chi_d(t) - a_d r_d(t)]\,dt + \sigma_d dW_d^{\mathbb{Q}}(t), \\
dr_f(t) &= [\chi_f(t) - a_f r_f(t)]\,dt + \sigma_f dW_f^{\mathbb{Q}^f}(t),
\end{aligned}$$

where the parameters a_d and a_f represent the speed of mean reversion parameters to the time-dependent functions $\chi_d(t)$ and $\chi_f(t)$, and the volatilities of the domestic, foreign, and FX processes are given by σ_d, σ_f and σ_y, respectively.

Moreover, $W_y^{\mathbb{Q}}(t)$ and $W_d^{\mathbb{Q}}(t)$ are Brownian motions under the risk-neutral domestic measure $\mathbb{Q}$, while $W_f^{\mathbb{Q}^f}(t)$ is a Brownian motion under the risk-neutral foreign measure $\mathbb{Q}^f$. However, we aim to express the SDE for $r_f(t)$ using a Brownian motion defined under the domestic risk-neutral measure $\mathbb{Q}$. As we explained in Section 2.5 we accomplish that by making a change from the risk-neutral foreign measure to the risk-neutral domestic measure. We also define the matrix of correlations between the Brownian motions

$$\mathbf{W}(t) = \begin{pmatrix} W_y^{\mathbb{Q}}(t) & W_d^{\mathbb{Q}}(t) & W_f^{\mathbb{Q}}(t) \end{pmatrix}$$

under the risk-neutral domestic measure as

$$\mathbf{W}(t)(d\mathbf{W}(t))^T = \begin{pmatrix} 1 & \rho_{y,d} & \rho_{y,f} \\ \rho_{y,d} & 1 & \rho_{d,f} \\ \rho_{y,f} & \rho_{d,f} & 1 \end{pmatrix} dt,$$

where $\rho_{y,d}$ is the correlation coefficient between the FX spot rate process $y(t)$ and the domestic short rate process $r_d(t)$, $\rho_{y,f}$ is the correlation coefficient between the FX spot rate process $y(t)$ and the foreign short rate process $r_f(t)$, and $\rho_{d,f}$ is the correlation coefficient between the domestic short rate process $r_d(t)$ and the foreign short rate process $r_f(t)$.

Given this correlation structure, it can be shown that under the risk-neutral domestic measure, the FX-HW model is given by

$$dy(t) = (r_d(t) - r_f(t))y(t)dt + \sigma_y y(t)dW_y^{\mathbb{Q}}(t), \tag{2.27}$$

$$dr_d(t) = \left[\chi_d(t) - a_d r_d(t)\right] dt + \sigma_d dW_d^{\mathbb{Q}}(t), \tag{2.28}$$

$$dr_f(t) = \left[\chi_f(t) - a_f r_f(t) - \sigma_f \sigma_y \rho_{y,f}\right] dt + \sigma_f dW_f^{\mathbb{Q}}(t). \tag{2.29}$$

the aim is to find the dynamics of the three processes.

2.6.1. Dynamics of domestic and foreign currency

Formulating the dynamics of the domestic currency is easy. Since the domestic currency process (2.28) is given by the Hull-White process from Section 2.3, we can use all the theory about the dynamics given in that section, with the notion that almost all parameters get an extra subscript, i.e. $r(t) \to r_d(t)$.

However, the foreign currency process (2.29) is not a Hull-White process anymore. Therefore, the dynamics under the risk-neutral domestic measure $\mathbb{Q}$ undergo slight modifications, which we outline in this section. We will not state all the proofs, because all the proofs are similar to the proofs that we have seen in Section 2.3. We start by rewriting (2.29) and we obtain

$$\begin{aligned} r_f(t) &= x_f(t) + \alpha_f(t), \\ dx_f(t) &= (-a_f x_f(t) - \sigma_f \sigma_y \rho_{y,f})dt + \sigma_f dW_f^{\mathbb{Q}}(t),\ x_f(0) = 0. \end{aligned}$$

Although $x_f(t)$ is not completely stochastic anymore due to the term $\sigma_f \sigma_y \rho_{y,f} dt$, we still will call $x_f(t)$ the stochastic part of $r_f(t)$, and $\alpha_f(t)$ the deterministic part of $r_f(t)$. It is easy to see that for $s \le t$

$$x_f(t) = x_f(s)e^{-a_f(t-s)} - \frac{\sigma_f \sigma_y \rho_{f,y}}{a_f}\left(1 - e^{-a_f(t-s)}\right) + \int_s^t \sigma_f e^{-a_f(t-u)} dW_f^{\mathbb{Q}}(u). \tag{2.30}$$

Using this expression, we find that $x_f(t)$ follows a normal distribution.

Theorem 2.6.1. *Let $0 \le s \le t$. The process $x_f(t)$ (2.29) given $\mathcal{F}(s)$ is normally distributed with mean and variance given by*

$$\begin{aligned} \mathbb{E}_s\left[x_f(t)\right] &= x_f(s)e^{-a_f(t-s)} - \frac{\sigma_f \sigma_y \rho_{f,y}}{a_f}\left(1 - e^{-a_f(t-s)}\right), \\ \text{Var}_s\left[x_f(t)\right] &= \int_s^t \sigma_f^2 e^{-2a_f(t-u)} du \\ &= \frac{\sigma_f^2}{2a_f}(1 - e^{-2a_f(t-s)}). \end{aligned}$$

Moreover, for any $0 \le s \le t, T$, the covariance of $x_f(t)$ and $x_f(T)$ given $\mathcal{F}(s)$ is given by

$$\text{Cov}_s\left[x_f(t), x_f(T)\right] = \begin{cases} \frac{\sigma_f^2}{2a_f}\left(e^{-a_f(T-t)} - e^{-a_f(t+T-2s)}\right) & t \le T, \\ \frac{\sigma_f^2}{2a_f}\left(e^{-a_f(t-T)} - e^{-a_f(T+t-2s)}\right) & t > T. \end{cases}$$

Proof. The proof of the theorem is similar to the proof of Theorem 2.3.3, therefore we refer to Appendix A.1.3. □

After finding the dynamics of $x_f(t)$, we look for an expression for the deterministic part $\alpha_f(t)$.

Theorem 2.6.2. *For $t \ge 0$, the deterministic part $\alpha_f(t)$ is given by*

$$\begin{aligned} \alpha_f(t) &= f_f^M(0,t) + \int_0^t \frac{\sigma_f^2}{a_f}\left(1 - e^{-a_f(t-s)}\right)e^{-a_f(t-s)} ds \\ &= f_f^M(0,t) + \frac{\sigma_f^2}{2a_f^2}\left(1 - e^{-a_f t}\right)^2, \end{aligned} \tag{2.31}$$

where $f_f^M(0,t)$ is given in (2.5) *as*

$$f_f^M(0,t) = -\frac{\partial}{\partial t}\log\big(P_f^M(0,t)\big). \tag{2.32}$$

Proof. For the proof of the theorem, we slightly changed the derivation in Di Francesco [11]. □

We can now combine the dynamics of $x_f(t)$ and the expression for $\alpha_f(t)$, to obtain the dynamics of the short rate process $r_f(t)$.

Theorem 2.6.3. *Let $0 \leq s < t$. The short rate process $r_f(t)$ is given by*

$$r_f(t) = x_f(s)e^{-a_f(t-s)} - \frac{\sigma_f \sigma_y \rho_{f,y}}{a_f}\left(1 - e^{-a_f(t-s)}\right) + \int_s^t \sigma_f e^{-a_f(t-u)} dW_f^{\mathbb{Q}}(u) + f_f^M(0,t) + \frac{\sigma_f^2}{2a_f^2}\left(1 - e^{-a_f t}\right)^2. \tag{2.33}$$

The short rate process $r_f(t)$ (2.33) given $\mathcal{F}(s)$ is normally distributed with mean and variance given by

$$\begin{aligned}
\mathbb{E}_s\left[r_f(t)\right] &= x_f(s)e^{-a_f(t-s)} - \frac{\sigma_f \sigma_y \rho_{f,y}}{a_f}\left(1 - e^{-a_f(t-s)}\right) + f_f^M(0,t) + \frac{\sigma_f^2}{2a_f^2}\left(1 - e^{-a_f t}\right)^2, \\
\text{Var}_s\left[r(t)\right] &= \frac{\sigma_f^2}{2a_f}(1 - e^{-2a_f(t-s)}).
\end{aligned}$$

Moreover, for any $0 \leq s \leq t, T$, the covariance of $r_f(t)$ and $r_f(T)$ given $\mathcal{F}(s)$ is given by

$$\text{Cov}_s\left[x_f(t), x_f(T)\right] = \begin{cases} \frac{\sigma_f^2}{2a_f}\left(e^{-a_f(T-t)} - e^{-a_f(t+T-2s)}\right) & t \leq T, \\ \frac{\sigma_f^2}{2a_f}\left(e^{-a_f(t-T)} - e^{-a_f(T+t-2s)}\right) & t > T. \end{cases}$$

Proof. Simply combine the expressions in Theorem 2.6.1 and theorem 2.6.3. The covariance structure is the same as for the domestic currency. □

Now we want to find the dynamics of $\int_t^T r_f(s)ds$ for $0 \leq t \leq T$. Therefore, we would need to find the dynamics of $\int_t^T x_f(s)ds$ and we need to find an expression of $\int_t^T \alpha_f(s)ds$.

Theorem 2.6.4. *The integral of the process $x_f(t)$ can be expressed as*

$$\int_t^T x_f(s)ds = x_f(t)\frac{1 - e^{-a_f(T-t)}}{a_f} - \frac{\sigma_f \sigma_y \rho_{y,f}}{a_f}\left(T - t - \frac{1 - e^{-a_f(T-t)}}{a_f}\right) + \int_t^T \frac{\sigma_f}{a_f}\left(1 - e^{-a_f(T-u)}\right) dW_f^{\mathbb{Q}}(u). \tag{2.34}$$

Then $\int_t^T x_f(t)dt$ given $\mathcal{F}(t)$ is normally distributed with variance and mean given by

$$\begin{aligned}
V_f(t,T) = \text{Var}_t\left[\int_t^T x_f(s)ds\right] &= \int_t^T \frac{\sigma_f^2}{a_f^2}\left(1 - e^{-a_f(T-u)}\right)^2 du \\
&= \frac{\sigma_f^2}{a_f^2}\left((T-t) - \frac{2}{a_f}\left(1 - e^{-a_f(T-t)}\right) + \frac{1}{2a_f}\left(1 - e^{-2a_f(T-t)}\right)\right),
\end{aligned} \tag{2.35}$$

$$\mathbb{E}_t\left[\int_t^T x_f(s)ds\right] = x_f(t)\frac{1 - e^{-a_f(T-t)}}{a_f} - \frac{\sigma_f \sigma_y \rho_{y,f}}{a_f}\left(T - t - \frac{1 - e^{-a_f(T-t)}}{a_f}\right).$$

Proof. For the proof of the theorem, we slightly changed the derivation in Di Francesco [11]. □

Since the expression of $\alpha_f(t)$ is similar to the expression of $\alpha_d(t)$, the same result holds for its integral.

Theorem 2.6.5. *The integral of the process $\alpha_f(t)$ (2.31) be expressed as*

$$\int_t^T \alpha_f(s)ds = \log\left(\frac{P_f^M(0,t)}{P_f^M(0,T)}\right) + \frac{V_f(0,T) - V_f(0,t)}{2}, \tag{2.36}$$

where $V_f(t,T)$ is given by (2.35).

Proof. The proof of the theorem is similar to the proof of Theorem 2.3.7, therefore we refer to Appendix A.1.4. □

Combining (2.34) and (2.36), provides us with a closed form of $\int_t^T r_f(s)ds$ as follows:

$$\begin{aligned}\int_t^T r_f(s)ds &= x_f(t)\frac{1-e^{-a_f(T-t)}}{a_f} - \frac{\sigma_f\sigma_y\rho_{y,f}}{a_f}\left(T-t-\frac{1-e^{-a_f(T-t)}}{a_f}\right) \\ &\quad + \int_t^T \frac{\sigma_f}{a_f}\left(1-e^{-a_f(T-u)}\right)dW_f^{\mathbb{Q}}(u) + \log\left(\frac{P_f^M(0,t)}{P_f^M(0,T)}\right) + \frac{V_f(0,T)-V_f(0,t)}{2}.\end{aligned} \tag{2.37}$$

The dynamics of the integrated foreign short rate are then easy to find.

Theorem 2.6.6. *The integral of the short rate process $\int_t^T r_f(s)ds$ given $\mathcal{F}(t)$ is normally distributed with mean and variance given by*

$$\begin{aligned}\mathbb{E}_t\left[\int_t^T r_f(s)ds\right] &= x_f(t)\frac{1-e_f^{-a_f(T-t)}}{a_f} - \frac{\sigma_f\sigma_y\rho_{y,f}}{a_f}\left(T-t-\frac{1-e^{-a_f(T-t)}}{a_f}\right) \\ &\quad + \log\left(\frac{P_f^M(0,t)}{P_f^M(0,T)}\right) + \frac{V_f(0,T)-V_f(0,t)}{2},\end{aligned} \tag{2.38}$$

$$\text{Var}_t\left[\int_t^T r_f(s)ds\right] = V_f(t,T), \tag{2.39}$$

where $V_f(t,T)$ is given by (2.35)*.*

Proof. Since $\int_t^T x(s)ds$ is normally distributed and $\int_t^T \alpha(s)ds$ is deterministic, we get that $\int_t^T r(s)ds$ is normally distributed as well. Using Itô's Isometry (Theorem 2.1.5), finding the mean and variance is trivial. □

In a similar way as we did for the domestic currency, we can also find the structure of the covariances of the integrated foreign short rates between different time points. Again, we will fix the filtration to $\mathcal{F}(0)$, meaning we only at the information available today. The covariance is then given as follows.

Theorem 2.6.7. *For any $0 \leq t, T$, the covariance of $\int_0^t r_f(u)du$ and $\int_0^T r_f(u)du$ given $\mathcal{F}(0)$ is given by*

$$\text{Cov}_0\left[\int_0^t r_f(u)du, \int_0^T r_f(u)du\right] = \begin{cases} \frac{\sigma_f^2}{a_f^2}\left(t - \frac{1-e^{-a_f t}+e^{-a_f(T-t)}-e^{-a_f T}}{a_f} + \frac{e^{-a_f(T-t)}-e^{-a_f(t+T)}}{2a_f}\right) & t \leq T, \\ \frac{\sigma_f^2}{a_f^2}\left(T - \frac{1-e^{-a_f T}+e^{-a_f(t-T)}-e^{-a_f t}}{a_f} + \frac{e^{-a_f(t-T)}-e^{-a_f(T+t)}}{2a_f}\right) & t > T. \end{cases}$$

Proof. The proof of the theorem is similar to the proof of Theorem 2.3.9, therefore we refer to Appendix A.1.5. □

Lastly, we want to present a closed-form expression for the ZCB price in the foreign currency, $P_f(t,T)$. It is important to note that this price needs to be determined under the risk-neutral domestic measure $\mathbb{Q}$ instead of the risk-neutral foreign measure $\mathbb{Q}^f$. Therefore, we need to make a change from the risk-neutral foreign measure to the risk-neutral domestic measure. We explained this procedure in Section 2.6. However, after we have changed measures, we need the dynamics of the FX spot rate $y(t)$ to calculate the ZCB price. Therefore, we will first discuss these dynamics, and afterward return to closed-form expression of the foreign ZCB price.

2.6.2. Dynamics of FX spot rate

In this section, we derive the dynamics for the FX spot rate. Note that the process $y(t)$, with SDE given by (2.27), is a Geometric Brownian Motion (GBM), meaning $y(t)$ is log-normally distributed. Therefore, we know that for $0 \leq s < t$

$$\begin{aligned}y(t) &= y(s)\exp\left(\int_s^t\left(r_d(u) - r_f(u) - \frac{1}{2}\sigma_y^2\right)du + \int_s^t \sigma_y dW_y^{\mathbb{Q}}(u)\right) \\ &= y(s)\exp\left(\int_s^t r_d(u)du - \int_s^t r_f(u)du - \frac{1}{2}\sigma_y^2(t-s) + \int_s^t \sigma_y dW_y^{\mathbb{Q}}(u)\right).\end{aligned} \tag{2.40}$$

We also know that the logarithm of the FX spot rate,

$$L_y(t) := \log y(t),$$

follows a normal distribution. By using Itô's Lemma (Theorem 2.1.4), it is easy to show that the SDE of $L_y(t)$ is given by

$$dL_y(t) = (r_d(t) - r_f(t) - \frac{1}{2}\sigma_y^2)dt + \sigma_y dW_y^{\mathbb{Q}}(t),$$

or equivalently for some $0 \leq s \leq t$

$$\begin{aligned} L_y(t) &= L_y(s) + \int_s^t \left(r_d(u) - r_f(u) - \frac{1}{2}\sigma_y^2 \right) du + \int_s^t \sigma_y dW_y^{\mathbb{Q}}(u) \\ &= \log y(s) + \int_s^t r_d(u)du - \int_s^t r_f(u)du - \frac{1}{2}\sigma_y^2(t-s) + \int_s^t \sigma_y dW_y^{\mathbb{Q}}(u). \end{aligned} \tag{2.41}$$

Since we have already computed the dynamics of the integrals of the domestic and foreign short rates, it is easy to find the dynamics of $L_y(t)$.

Theorem 2.6.8. *The logarithm of $y(t)$, $L_y(t)$, given $\mathcal{F}(s)$ is normally distributed with mean given by*

$$\begin{aligned} \mathbb{E}_s\left[L_y(t)\right] &= \log y(s) + \mathbb{E}_s\left[\int_s^t r_d(u)du\right] - \mathbb{E}_s\left[\int_s^t r_f(u)du\right] - \frac{1}{2}\sigma_y(t-s) \\ &= \log y(s) + x_d(s)B_d(s,t) + \log\left(\frac{P_d^M(0,s)}{P_d^M(0,t)}\right) + \frac{V_d(0,t) - V_d(0,s)}{2} - x_f(s)B_f(s,t) \\ &\quad + \frac{\sigma_f\sigma_y\rho_{y,f}}{a_f}\left(t - s - B_f(s,t)\right) - \log\left(\frac{P_f^M(0,s)}{P_f^M(0,t)}\right) - \frac{V_f(0,t) - V_f(0,s)}{2} - \frac{1}{2}\sigma_y^2(t-s) \end{aligned}$$

Proof. The proof is trivial and follows from Theorem 2.3.8 and Theorem 2.6.6. □

We also present the expression of the covariance structure of the logarithm of the FX spot rate between two time points. It is given as follows.

Theorem 2.6.9. *For any $0 \leq s \leq t \leq T$, the covariance of $L_y(t)$ and $L_y(T)$ given $\mathcal{F}(s)$ is given by*

$$\begin{aligned} \mathrm{Cov}_s[L_y(t), L_y(T)] = &\frac{\sigma_d^2}{a_d^2}\left(t - s - B_d(s,t) - \frac{e^{-a_d(T-t)} - e^{-a_d(T-s)}}{a_d} + \frac{e^{-a_d(T-t)} - e^{-a_d(t+T-2s)}}{2a_d}\right) \\ &+ \frac{\sigma_f^2}{a_f^2}\left(t - s - B_f(s,t) - \frac{e^{-a_f(T-t)} - e^{-a_f(T-s)}}{a_f} + \frac{e^{-a_f(T-t)} - e^{-a_f(t+T-2s)}}{2a_f}\right) \\ &- \frac{\sigma_d\sigma_f}{a_da_f}\rho_{d,f}\left(t - s - B_d(s,t) - \frac{e^{-a_f(T-t)} - e^{-a_f(T-s)}}{a_f} + \frac{e^{-a_f(T-t)} - e^{-a_d(t-s)-a_f(T-s)}}{a_d + a_f}\right. \\ &\quad \left. + t - s - B_f(s,t) - \frac{e^{-a_d(T-t)} - e^{-a_d(T-s)}}{a_d} + \frac{e^{-a_d(T-t)} - e^{-a_f(t-s)-a_d(T-s)}}{a_d + a_f}\right) \\ &+ \frac{\sigma_d\sigma_y}{a_d}\rho_{y,d}\left(t - s - B_d(s,t) + t - s - \frac{e^{-a_d(T-t)} - e^{-a_d(T-s)}}{a_d}\right) \\ &- \frac{\sigma_f\sigma_y}{a_d}\rho_{y,f}\left(t - s - B_f(s,t) + t - s - \frac{e^{-a_f(T-t)} - e^{-a_f(T-s)}}{a_f}\right) + \sigma_y^2(t-s). \end{aligned}$$

Proof. For the proof of the theorem, we refer to Appendix A.2.1. □

Note that, in case $0 \leq s \leq T \leq t$, then the covariance of $L_y(t)$ and $L_y(T)$ given $\mathcal{F}(s)$ is given by the formula in Theorem 2.6.9 but with switched positions for t and T.

As promised, we will present a closed-form expression for the foreign ZCB price, $P_f(t,T)$. The expression is given as follows.

Theorem 2.6.10. *For $0 \leq t < T$, the ZCB price $P(t,T)$ is given by*

$$\begin{aligned} P_f(t,T) &= \mathbb{E}_t^{\mathbb{Q}^f}\left[e^{-\int_t^T r_f(s)ds}\right] \\ &= \frac{1}{y(t)}\mathbb{E}_t^{\mathbb{Q}}\left[y(T)e^{-\int_t^T r_d(s)ds}\right] \\ &= \hat{A}_f(t,T)e^{-B_f(t,T)x_f(t)}, \end{aligned} \tag{2.42}$$

where

$$\begin{aligned} A_f(t,T) &= \frac{P_f^M(0,T)}{P_f^M(0,t)} \exp\left(\frac{1}{2}\left(V_f(t,T) - V_f(0,T) + V_f(0,t)\right)\right), \\ B_f(t,T) &= \frac{1-e^{-a_f(T-t)}}{a_f}, \end{aligned}$$

and $f_f^M(0,t)$ and $V(t,T)$ are given by (2.32) *and* (2.35)*, respectively.*

Proof. For the proof of the theorem, we refer to Appendix A.2.2. ☐

2.7. Interest rate and foreign exchange rate derivatives

In this section, we present three different financial derivatives. We also give formulas for pricing these derivatives. We start with the principles of an IR swap. Throughout this section, the notation aligns with the notation of Oosterlee and Grzelak [6].

2.7.1. IR swap

A swap is a financial derivative that involves two parties that exchange one or multiple cash flows over time. One of the most commonly traded swaps is the plain vanilla IR swap, where one party agrees to pay fixed cash flows that are equal to the interest rate at a predetermined, fixed rate on a notional amount. The times at which the cash flows are paid are predetermined as well. The other party pays, at the same predetermined times, a floating interest rate on the same notional amount, based on a benchmark rate, such as the Libor forward rate. If we define a trading time horizon $[0, T^*]$ and we define a set of times as $\{T_1, ..., T_n\}$ with tenor $\tau_i = T_i - T_{i-1}$ for $i = 1, ..., n$, then the Libor forward rate is defined as follows.

Definition 2.7.1. *For a given tenor τ_i, the Libor forward rate $\ell(t; T_{i-1}, T_i)$ for a period $[T_{i-1}, T_i]$ is defined as*

$$\ell_i(t) \equiv \ell(t; T_{i-1}, T_i) = \frac{1}{\tau_i}\frac{P(t, T_{i-1}) - P(t, T_i)}{P(t, T_i)}, \tag{2.43}$$

where $P(t,T)$ represents the ZCB price at time t maturing at time T (2.23).

The plain vanilla IR swap consists of two legs. The fixed leg is a series of fixed rate payments, at fixed rate K at each future payment time in $\{T_{j+1}, ..., T_n\}$. The floating leg is a series of floating forward Libor rates.

The two parties that are involved in the IR swap are called the IR swap payer and receiver. Usually, the payer receives a floating leg and pays a fixed leg, and the receiver pays the floating leg and receives the fixed leg. The payoff of the IR swap is given by

$$\bar{\beta}\sum_{j+1}^{n} N_s\tau_i\left(\ell(T_{i-1}; T_{i-1}, T_i) - K\right),$$

with N_s being the notional amount of the swap, K the fixed interest rate, $(\ell(T_{i-1}; T_{i-1}, T_i) - K)$ the Libor rate for a period $[T_{i-1}, T_i]$, and where $\bar{\beta} = 1$ is used for a payer and $\bar{\beta} = -1$ for a receiver.

Valuing the swap at some point time T_k can be done as follows.

Definition 2.7.2. *Let $\{T_{j+1}, ..., T_n\}$ be the set of future payment times with tenor τ_i for each $i \in \{j+1, ..., n\}$. Let N_s be the notional amount of the swap, K the fixed interest rate, and $\ell(T_{i-1}; T_{i-1}, T_i)$ be the forward Libor rate over the period $[T_{i-1}, T_i]$, as in* (2.43)*. Then the value of the swap at time T_k is given by*

$$V(T_k) = \bar{\beta} N_s \sum_{i=j+1}^{n} \tau_i \mathbb{E}_{T_k} \left[\frac{M(T_k)}{M(T_i)} \left(\ell(T_{i-1}; T_{i-1}, T_i) - K \right) \right], \tag{2.44}$$

with $\bar{\beta} = 1$ for a payer and $\bar{\beta} = -1$ for a receiver, and where $M(t)$ is the money-savings account (2.1)*.*

The expectation in (2.44) is taken under the risk-neutral measure. However, this expectation can not easily be solved. Therefore we make a change to the T-forward measure, see Section 2.5. Then the price of the IR swap is given as follows.

Theorem 2.7.1. *Let $\{T_{j+1}, ..., T_n\}$ be the set of future payment times with tenor τ_i for each $i \in \{j+1, ..., n\}$. Let N_s be the notional amount of the swap, K the fixed interest rate, and $\ell(T_{i-1}; T_{i-1}, T_i)$ be the forward Libor rate over the period $[T_{i-1}, T_i]$, as in* (2.43)*. Then the value of the swap at time T_k is given by*

$$\begin{aligned} V(T_k) &= \bar{\beta} N_s \sum_{i=j+1}^{n} \tau_i P(T_k, T_i) \left(\ell_i(T_k; T_{i-1}, T_i) - K \right), \\ &= \bar{\beta} N_s (P(T_k, T_j) - P(T_k, T_n)) - \bar{\beta} N_s K \sum_{i=j+1}^{n} \tau_i P(T_k, T_i), \end{aligned} \tag{2.45}$$

with $\bar{\beta} = 1$ for a payer and $\bar{\beta} = -1$ for a receiver.

Proof. For the proof of the theorem, we refer to Oosterlee and Grzelak [6]. □

2.7.2. FX swap

A different kind of swap contract is an FX swap. In an FX swap, a fixed FX rate is usually exchanged with a floating FX rate. The exchange of these cash flows only happens at maturity time T. One party receives a floating FX rate, $y(T)$, while the other party receives a fixed rate K. The amount that needs to be paid is based on the foreign notional amount N_f.

We call the two parties that are involved in the FX swap the payer and receiver. Usually, the payer receives a floating leg and pays a fixed leg, and the receiver pays the floating leg and receives the fixed leg. At the maturity date T of the FX swap, the value of the FX swap in the domestic currency is given by

$$V(T) = \bar{\beta} N_f (y(T) - K),$$

where $\bar{\beta} = 1$ for a payer and $\bar{\beta} = -1$ for a receiver. Valuing the FX swap at time T_k can be done as follows.

Definition 2.7.3. *Let $y(T)$ be the FX spot rate at maturity T, N_f be the foreign notional amount of the swap, and K the fixed FX rate. Then the value of the swap at time T_k is given by*

$$V(T_k) = \bar{\beta} M_d(T_k) \mathbb{E}_{T_k} \left[\frac{N_f}{M_d(T)} \left(y(T) - K \right) \right],$$

with $\bar{\beta} = 1$ for a payer and $\bar{\beta} = -1$ for a receiver, and where $M_d(t)$ is the money-savings account (2.1) *in the domestic currency.*

Again, the expectation is taken under the risk-neutral measure. Similarly as we did for the IR swap, we make the switch to the T-forward measure in order to solve the expectation. Again, we refer to Section 2.5 for all the details. We end up with the following theorem.

Theorem 2.7.2. *Let $y(T)$ be the FX spot rate at maturity T, N_f be the foreign notional amount of the swap, and K the fixed FX rate. Then the value of the swap at time T_k is given by*

$$V(T_k) = \bar{\beta} P_d(T_k, T) N_f (y_F(T_k, T) - K), \tag{2.46}$$

with $\bar{\beta} = 1$ for a payer and $\bar{\beta} = -1$ for a receiver. Furthermore, $P_d(t, T)$ is the ZCB price (2.23) *in the domestic currency, and $y_F(t, T)$ is the FX forward rate* (2.24)*.*

Proof. For the proof of the lemma, we refer to Oosterlee and Grzelak [6]. □

2.7.3. FX digital option

Lastly we present the principles of FX digital options. An FX Digital option, also called an FX cash-or-nothing option, is available in many different forms and is popular for hedging and speculation. At maturity time T the option is either worth 0 in case the forward exchange rate $y_F(T,T) \leq K$ for some strike price K, or the option is worth the notional amount N_f in case the forward exchange rate $y_F(T,T) > K$. Therefore, the value of the FX digital option at time T is given by

$$V(T) = N_f \mathbb{1}(\bar{\gamma}(y_F(T,T) - K) > 0),$$

where $\bar{\gamma} = 1$ for a call option and $\bar{\gamma} = -1$ for a put option, and N_f is the notional in the foreign currency. Before we can value the FX digital, we need a theorem that provides us the dynamics of the FX forward rate (2.24).

Theorem 2.7.3. *Under the T-forward measure $\mathbb{Q}^T$, the FX forward rate $y_F(t,T)$ follows a log-normal process given by*

$$dy_F(t,T) = y_F(t,T)\Sigma(t,T)dW_F^T(t),$$

where $W_F^T(t)$ is a Brownian Motion related to the FX forward rate, and $\Sigma(t,T)$ is the instantaneous variance of the FX forward, given by

$$\begin{aligned}\Sigma^2(t,T) &= \text{Var}\left[\sigma_y dW_y^T(t) - \sigma_f B_f(t,T)dW_f^T(t) + \sigma_d B_d(t,T)dW_d^T(t)\right] \\ &= \sigma_y^2 + \sigma_f^2 B_f^2(t,T) + \sigma_d^2 B_d^2(t,T) - 2\rho_{y,f}\sigma_y\sigma_f B_f(t,T) \\ &\quad + 2\rho_{y,d}\sigma_y\sigma_d B_d(t,T) - 2\rho_{d,f}\sigma_f B_f(t,T)\sigma_d B_d(t,T).\end{aligned} \tag{2.47}$$

Proof. For the proof of the theorem, we refer either to Green [12] or Oosterlee and Grzelak [6]. □

The value of the digital option is given as follows.

Theorem 2.7.4. *Let $y_F(t,T)$ be the FX forward rate* (2.24) *at time t with maturity T, N_f be the foreign notional amount of the digital option, and K the strike price. Then the value of the digital option at time T_k is given by*

$$V(T_k) = \bar{\gamma} N_f P_d(T_k,T)\Phi(d_2(T_k)), \tag{2.48}$$

where $\bar{\gamma} = 1$ for a call option and $\bar{\gamma} = -1$ for a put option, $\Phi(\cdot)$ is the standard normal cumulative distribution function, $P_d(t,T)$ is the ZCB price given by (2.23)*, and*

$$d_2(T_k) = \frac{\log\left(y_F(T_k,T)/K\right) - \frac{1}{2}\sigma_{\textit{imp}}^2(T_k,T)(T-T_k)}{\sigma_{\textit{imp}}(T_k,T)\sqrt{T-T_k}},$$

where

$$\sigma_{\textit{imp}}^2(t,T) = \frac{1}{T-t}\int_t^T \Sigma^2(s,T)ds,$$

and where $\Sigma(t,T)$ is the instantaneous variance of the FX forward rate (2.47).

Proof. For the proof of the theorem, we refer to Appendix A.2.3. □

3
CVA Framework

Credit Value Adjustment (CVA) is a concept in financial risk management that gained importance after the 2008 crisis. It refers to the adjustment made to the price of a financial derivative to account for counterparty credit risk (CCR). Even broader, Green [12] calls CVA the market value of CCR on a financial derivative. CCR is the risk that a counterparty with whom one has entered into a financial contract will default before the expiration of the contract and will fail to meet their side of the contractual agreement. This type of risk mainly arises in the context of OTC derivatives, like the swaps and digital options that we have seen in Section 2.7. The default of a counterparty can have significant consequences for the other party, leading to financial losses, legal disputes, and reputational damage. As a result, understanding and managing counterparty credit risk is crucial for financial institutions, investors, and traders.

As previously mentioned, CVA is the market value of CCR on a financial derivative. It represents the price of CCR that should be added to the default-free fair price of a financial derivative to incorporate the counterparty's default risk into the price of the derivative. In this way, the risk that the counterparty would miss payment obligations is compensated by the other party in the contract. The pricing of financial derivatives under CCR requires sophisticated models relying on the concept of exposure, which defines the loss in the event of a counterparty default.

3.1. Exposure

The concept of (positive) exposure is defined as follows.

Definition 3.1.1. *Let $V(t)$ be the value of a financial derivative at time t. Then the (positive) exposure, $E(t)$, is defined as the maximum of the value of the derivative and zero:*

$$E(t) := \max(V(t), 0).$$

Then, we define the expected (positive) exposure as follows.

Definition 3.1.2. *Let $E(t)$ be the (positive) exposure at time t. Then the expected (positive) exposure at time $t_0 < t$ is given by*

$$EE(t_0, t) = \mathbb{E}_{t_0}\left[\frac{M(t_0)}{M(t)}E(t)\right],$$

where $M(t)$ is the money-savings account that we introduced in (2.1) *given by*

$$M(t) = \exp\left(\int_0^t r(s)ds\right).$$

Since the values of financial derivatives are often analytically known, the expected exposure of financial derivatives can sometimes also be analytically derived, as well as their derivatives with respect to some market parameters. We present here the analytically derived expected exposures and their derivatives with respect to the FX spot rate $y(t_0)$ for an FX swap and an FX digital option.

3.1.1. Expected exposure for an FX swap

We start by finding the analytical expression of the expected exposure and its derivative with respect to the FX spot rate $y(t_0)$ for an FX swap. If we let N_f be the foreign notional amount of the swap, and K the fixed FX rate, then recall from Theorem 2.7.2, that the value of an FX swap at time T_k is given by

$$V(T_k) = \bar{\beta} P_d(T_k, T) N_f (y_F(T_k, T) - K),$$

with $\bar{\beta} = 1$ for a payer and $\bar{\beta} = -1$ for a receiver. Furthermore, $P_d(t, T)$ is the ZCB price (2.23) in the domestic currency, and $y_F(t, T)$ is the forward FX rate (2.24).

The expected exposure of an FX swap and its derivative with respect to the FX spot rate is then given as follows.

Theorem 3.1.1. *Let N_f be the foreign notional amount of the swap, and K the fixed FX rate. Then the expected exposure of the FX payer swap at time T_k and its derivative with respect to $y(t_0)$ are given by*

$$\begin{aligned}
\textit{EE}(t_0, T_k) &= N_f P_d(t_0, T) \left(y_F(t_0, T) \Phi(d_1(T_k)) - K \Phi(d_2(T_k)) \right), \\
\frac{\partial \textit{EE}(t_0, T_k)}{\partial y(t_0)} &= N_f P_d(t_0, T) \left(\frac{P_f(t_0, T)}{P_d(t_0, T)} \left(\Phi(d_1(T_k)) + \frac{\phi(d_1(T_k))}{\sigma_{\textit{imp}}(t_0, T_k)\sqrt{T_k - t_0}} \right) - K \frac{\phi(d_2(T_k))}{y(t_0)\sigma_{\textit{imp}}(t_0, T_k)\sqrt{T_k - t_0}} \right),
\end{aligned}$$

where $P_d(t, T)$ and $P_f(t, T)$ are the ZCB prices in the domestic and foreign currency, respectively, $y_F(t, T)$ is the forward FX rate (2.24)*, $\Phi(\cdot)$ is the standard normal cumulative distribution function with corresponding probability density function $\phi(\cdot)$, $d_1(T_k)$ and $d_2(T_k)$ are given by*

$$d_1(T_k) = \frac{\log\left(y_F(t_0, T)/K\right) + \frac{1}{2}\sigma^2_{\textit{imp}}(t_0, T_k)(T_k - t_0)}{\sigma_{\textit{imp}}(t_0, T_k)\sqrt{T_k - t_0}}, \quad d_2(T_k) = d_1(T_k) - \sigma_{\textit{imp}}(t_0, T_k)\sqrt{T_k - t_0},$$

and

$$\sigma^2_{\textit{imp}}(t, T) = \frac{1}{T - t} \int_t^T \Sigma^2(s, T) ds,$$

where $\Sigma(t, T)$ is the instantaneous variance of the FX forward rate (2.47)*. Similarly, the expected exposure of the FX payer receiver at time T_k and its derivative with respect to $y(t_0)$ are given by*

$$\begin{aligned}
\textit{EE}(t_0, T_k) &= N_f P_d(t_0, T) \left(K \Phi(-d_2(T_k)) - y_F(t_0, T) \Phi(-d_1(T_k)) \right), \\
\frac{\partial \textit{EE}(t_0, T_k)}{\partial y(t_0)} &= N_f P_d(t_0, T) \left(-K \frac{\phi(-d_2(T_k))}{y(t_0)\sigma_{\textit{imp}}(t_0, T_k)\sqrt{T_k - t_0}} - \frac{P_f(t_0, T)}{P_d(t_0, T)} \left(\Phi(-d_1(T_k)) - \frac{\phi(-d_1(T_k))}{\sigma_{\textit{imp}}(t_0, T_k)\sqrt{T_k - t_0}} \right) \right).
\end{aligned}$$

Proof. For the proof of the theorem, we refer to Appendix A.2.4. □

3.1.2. Expected exposure for an FX Digital option

We now present the analytical expression of the expected exposure and its derivative with respect to the FX spot rate $y(t_0)$ for an FX digital option. Recall that the value of the FX digital option at time T is given by

$$V(T) = N_f \mathbb{1}(\bar{\gamma}(y_F(T, T) - K) > 0),$$

where $\bar{\gamma} = 1$ for a call option and $\bar{\gamma} = -1$ for a put option, and N_f is the notional in the foreign currency.

The expected exposure of an FX swap and its derivative with respect to the FX spot rate is then given as follows.

Theorem 3.1.2. *Let N_f be the foreign notional amount of the swap, and K the fixed FX rate. Then the expected exposure of the FX digital call option at time T_k and its derivative with respect to $y(t_0)$ are given by*

$$\begin{aligned} \mathit{EE}(t_0, T_k) &= N_f P_d(t_0, T)\Phi(d_2), \\ \frac{\partial \mathit{EE}(t_0, T_k)}{\partial y(t_0)} &= N_f P_d(t_0, T)\frac{\phi(d_2)}{y(t_0)\sigma_{\mathit{imp}}(t_0, T)\sqrt{T - t_0}}, \end{aligned}$$

where $P_d(t, T)$ is the ZCB price (2.23) *in the domestic currency, $y_F(t, T)$ is the forward FX rate* (2.24)*, $\Phi(\cdot)$ is the standard normal cumulative distribution function with corresponding probability density function $\phi(\cdot)$, d_2 is given by*

$$d_2 = \frac{\log\left(y_F(t_0, T)/K\right) - \frac{1}{2}\sigma^2_{\mathit{imp}}(t_0, T)(T - t_0)}{\sigma_{\mathit{imp}}(t_0, T)\sqrt{T - t_0}},$$

and

$$\sigma^2_{\mathit{imp}}(t, T) = \frac{1}{T - t}\int_t^T \Sigma^2(s, T)ds,$$

where $\Sigma(t, T)$ is the instantaneous variance of the FX forward rate (2.47)*. Similarly, the expected exposure of the FX digital put option at time T_k is given by*

$$\begin{aligned} \mathit{EE}(t_0, T_k) &= -N_f P_d(t_0, T)\Phi(-d_2), \\ \frac{\partial \mathit{EE}(t_0, T_k)}{\partial y(t_0)} &= N_f P_d(t_0, T)\frac{\phi(-d_2)}{y(t_0)\sigma_{\mathit{imp}}(t_0, T)\sqrt{T - t_0}}. \end{aligned}$$

Proof. The proof of the theorem is similar to the proof of Theorem 3.1.1, therefore we refer to Appendix A.2.4. ☐

3.2. Credit Value Adjustment

In this section, we will derive the formula for unilateral CVA. In the unilateral context, we consider ourselves as a default-free company, that faces the risk that the counterparty, with whom the company has a financial contract, has the possibility to default, which means the counterparty is not able to meet its financial obligations. We assume that this risky contract is one of the IR or FX products that we introduced in Section 2.7. We assume that the maturity date of these products is T, and we denote the value of the financial derivatives at time $T_k \leq T$ with $V(T_k)$. Note that for the IR swap, the value at time T_k depends on a set of future payment times $\{T_{j+1}, ..., T_n\}$ with $T_n \leq T$, while for the FX swap and FX digital option, the value only depends on the times T_k and T.

Now we let T_D be the time of default. According to Oosterlee and Grzelak [6], the CVA value at time t_0 is given by

$$\text{CVA}(t_0) = \mathbb{E}_{t_0}\left[(1 - R_C)\frac{M(t_0)}{M(T_D)}\mathbb{1}_{T_D \leq T}\max(V(T_D), 0)\right],$$

where R_C is the recovery fraction $\in (0, 1]$, $M(t)$ is the money savings account (2.1), the expression $\mathbb{1}_{T_D \leq T}$ serves as indicator denoting whether a default of the counterparty has occurred before maturity, and $V(T_D)$ is the value at time T_D of the financial derivative that we consider.

If we assume independence between the default time T_D, the discount factors, and the market risks incorporated into the exposures, we can exclude wrong/right-way risks [2], and then CVA can be written as an integral

$$\text{CVA}(t_0) = (1 - R_C)\int_{t_0}^T \mathbb{E}\left[\frac{M(t_0)}{M(T_D)}\max(V(T_D), 0) \mid T_D = z\right] dF_{T_D}(z), \tag{3.1}$$

where F_{T_D} is the cumulative distribution function of the default time T_D. Finding a closed-form expression for the integral (3.1) is difficult. Therefore, we make an approximation. We specify a set of exposure times $\{T_1, ..., T_m\}$ with $T_m \leq T$. Then CVA at time t_0 can be approximated by

$$\text{CVA}(t_0) = (1 - R_C)\sum_{k=1}^m \text{EE}(t_0, T_k)\bar{q}(T_{k-1}, T_k),$$

with R_C the recovery fraction $\in [0,1]$, the expected positive exposure given by

$$\text{EE}(t, T_k) = \mathbb{E}_t \left[\frac{M(t)}{M(T_k)} \max(V(T_k), 0) \right],$$

and the probability of default in (T_{k-1}, T_k) given by

$$\bar{q}(T_{k-1}, T_k) = \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{T_{k-1} < T_D \leq T_k} \right].$$

Hence we see that CVA is a sum over all exposure times of the product of the loss given default (LGD = $1 - R_C$), the expected positive exposure (EE), and the probability of default (PD).

The probability of default is associated with survival probabilities. With $\bar{q}(T_{k-1}, T_k)$ we represent the chance that the counterparty will default during $[T_{k-1}, T_k]$. To model survival probabilities, we need a deterministic hazard rate ξ. The hazard rate can normally be based on credit derivatives, but in this thesis, the hazard rate is chosen as a constant. The probability of default in (T_{k-1}, T_k) is then given by

$$\begin{aligned}
\bar{q}(T_{k-1}, T_k) &= \mathbb{E}^{\mathbb{Q}} \left[\mathbb{1}_{T_{k-1} < T_D \leq T_k} \right] \\
&= \mathbb{Q}\left(T_D \leq T_k\right) - \mathbb{Q}\left(T_D \leq T_{k-1}\right). \\
&= e^{-\xi(T_k - t_0)} - e^{-\xi(T_{k-1} - t_0)}.
\end{aligned}$$

We have left out all the mathematical details since modeling the probability of default is not the scope of this thesis. For more details, we refer to Gregory [13].

Since the LGD and PD are deterministic, we can extend the expectation operator in the expected positive exposure formula over the LGD and the PD. We end up with the value of CVA as an expectation:

$$\text{CVA}(t_0) = \mathbb{E}_{t_0} \left[(1 - R_C) \sum_{k=1}^{m} \frac{M(t_0)}{M(T_k)} \max(V(T_k), 0) \bar{q}(T_{k-1}, T_k) \right]. \tag{3.2}$$

4

Computing First-Order CVA sensitivities

In this section, we will explain the main ideas behind two methods to compute CVA sensitivities; the Bump & Reprice (B&R) method, and the Likelihood Ratio Method (LRM). First, we introduce some notation, similar to the notation of Glasserman [4]. We consider a model that depends on a parameter θ ranging over some interval of the real line. For each value of θ, we can generate a random variable $Y(\theta)$, which represents the output of the model with parameter value θ. The goal of this section is to compute $\frac{d}{d\theta}\mathbb{E}[Y(\theta)]$. Therefore, we let $\mathbb{E}[Y(\theta)]$ equal the unilateral CVA formula (3.2), given by

$$\mathsf{CVA}(t_0,\theta) = \mathbb{E}_{t_0}\left[(1-R_C)\sum_{k=1}^{m}\frac{M_\theta(t_0)}{M_\theta(T_k)}\max(V_\theta(T_k),0)\bar{q}(T_{k-1},T_k)\right].$$

Note that for the discount factors, and for $V(T_k)$ we added dependency on θ. That means that only the discount factors and the values of the financial derivatives depend on θ, while the recovery rate R_C and the probability of default $\bar{q}(T_{k-1},T_k)$ are independent of θ. Dropping the expectation operator, we obtain the expression of $Y(\theta)$:

$$Y(\theta) = (1-R_C)\sum_{k=1}^{m}\frac{M_\theta(t_0)}{M_\theta(T_k)}\max(V_\theta(T_k),0)\bar{q}(T_{k-1},T_k). \tag{4.1}$$

We will present two methods to compute CVA sensitivities. In Section 4.1, we will outline the B&R method, and in Section 4.2 we will show the ideas behind LRM.

4.1. Bump & Reprice Method

The simplest method to estimate first-order CVA sensitivities is the B&R method. This is a recalculation of CVA using small shocks on each market parameter in order to compute finite difference approximations of the derivative of $\mathbb{E}[Y(\theta)]$. The main advantage of this method is that it is completely independent of the simulation framework, making it easy to understand and easy to implement. However, the number of calculations grows linearly with the number of underlying market risk factors, making the method computationally demanding [2].

The derivative can be estimated using a forward, backward, or central difference estimator. In this thesis, we opt for the central difference estimator, since Glasserman [4] claims that this estimator dominates the other estimators in terms of convergence.

We proceed as follows. We first choose a bump size $h > 0$ and we denote by N the number of MC simulations. Then we simulate N independent replications $\mathbf{Y}(\theta-h) = (Y_1(\theta-h),...,Y_N(\theta-h))$ and N independent replications $\mathbf{Y}(\theta+h) = (Y_1(\theta+h),...,Y_N(\theta+h))$. We average each set of replications to obtain

$$\bar{Y}(\theta-h) = \frac{1}{N}\sum_{i=1}^{N}Y_i(\theta-h) \quad \text{and} \quad \bar{Y}(\theta+h) = \frac{1}{N}\sum_{i=1}^{N}Y_i(\theta+h).$$

Then the central difference estimator is given by

$$\hat{\Delta}_{\mathsf{C}}(\theta) \equiv \hat{\Delta}_{\mathsf{C}}(\theta,N,h) = \frac{\bar{Y}(\theta+h)-\bar{Y}(\theta-h)}{2h}. \tag{4.2}$$

From (4.2) we conclude that we need at least two full evaluations of Y in order to obtain the B&R estimator. Since each evaluation requires us to do a MC simulation, it becomes clear why the method is computationally demanding.

Choosing a suitable bump size h can be a challenge. We want to choose h in such a way that the estimator has a small bias and a small variance. Glasserman [4] shows that

$$\text{Bias}[\hat{\Delta}_{\text{C}}(\theta)] = \frac{1}{6}\mathbb{E}[Y]'''(\theta)h^2 + \mathcal{O}(h^2). \tag{4.3}$$

This would imply that choosing a smaller h results in a small bias, and therefore a more accurate estimator. However, the effect of h must be weighed against the effect of h on the variance. For the variance the following holds.

Theorem 4.1.1. *Let* $\mathbf{Y}(\theta - h) = (Y_1(\theta - h), ..., Y_N(\theta - h))$ *and* $\mathbf{Y}(\theta + h) = (Y_1(\theta + h), ..., Y_N(\theta + h))$ *be i.i.d. Let* $h > 0$*, then the variance of the central difference estimator* (4.2) *is given by*

$$\text{Var}\left[\hat{\Delta}_{\text{C}}(\theta)\right] = \frac{1}{4Nh^2}\,\text{Var}\left[\mathbf{Y}(\theta + h) - \mathbf{Y}(\theta - h)\right]. \tag{4.4}$$

Proof. By using the properties of the variance, this proof is trivial. □

This implies that increasing the number of simulations N will decrease the variance. On the other hand, in contrast to the bias, we see that lowering h will increase the variance. Therefore, we need to make a trade-off by choosing h. We analyse this trade-off by looking at the dependence of Var $[\mathbf{Y}(\theta + h) - \mathbf{Y}(\theta - h)]$ on h. Glasserman [4] presents three cases regarding this dependence:

1. First of all, we can simulate $\mathbf{Y}(\theta - h)$ and $\mathbf{Y}(\theta + h)$ independently. Under the assumption that Var $[\mathbf{Y}(\theta)]$ is continuous in h, the variance of $\mathbf{Y}(\theta + h) - \mathbf{Y}(\theta - h)$ is of order $\mathcal{O}(1)$. This implies that the variance of the central difference estimator (4.4) is of order $\mathcal{O}(N^{-1}h^{-2})$.
2. In the second case we simulate $\mathbf{Y}(\theta - h)$ and $\mathbf{Y}(\theta + h)$ using common random numbers. That means that they are generated with the same seed for the random number generator. Then it follows that the variance of $\mathbf{Y}(\theta + h) - \mathbf{Y}(\theta - h)$ is of order $\mathcal{O}(h^{-1})$, meaning that the variance of the central difference estimator (4.4) is of order $\mathcal{O}(N^{-1}h^{-1})$.
3. For the last case, we do not only need that $\mathbf{Y}(\theta - h)$ and $\mathbf{Y}(\theta + h)$ are simulated with the same seed, but we also need that the output $Y(\cdot)$ is continuous in θ. It follows that the variance of $\mathbf{Y}(\theta + h) - \mathbf{Y}(\theta - h)$ is of order $\mathcal{O}(h^{-2})$, resulting in a variance of the central difference estimator (4.4) of order $\mathcal{O}(N^{-1})$.

Glasserman mentions that, regardless of the output $Y(\cdot)$, the second case outperforms the first case in terms of convergence, meaning it is better to simulate using random numbers than to simulate independently. However, the variance still depends on h, meaning we have to be careful with choosing h. A small h can lead to extreme variances, while a large h can result in a big bias. However, if our output $Y(\cdot)$ is continuous, the order of the variance of the central difference estimator (4.4) does not depend on h anymore, meaning no trade-off is needed between the bias and the variance. Therefore, we can choose h as small as possible, as long as we take care of the machine precision.

In the context of continuous outputs $Y(\theta)$, we will always fix the bump size at $h = 10^{-5}$. However, for discontinuous outputs, we will find an optimal bump size by comparing the bias and the variance for multiple values of h.

4.2. Likelihood Ratio Method

As an alternative to the B&R method, LRM is a way to estimate CVA sensitivities. The computation of CVA sensitivities under LRM involves the differentiation of a joint probability density function [2]. We require no knowledge of derivatives of the payoffs of financial payoffs. The payoffs can even be discontinuous in θ, which is an advantage over the B&R method, as the performance of the B&R method can be weak for discontinuous payoffs. Moreover, LRM allows the computation of multiple sensitivities simultaneously in a single MC simulation, while the B&R method needs to compute all the sensitivities separately.

Mathematically, we consider a discounted payoff $Y(\theta)$, in our case (4.1), and we assume the existence of a vector of risk factors $\mathbf{X} = (X_1, ..., X_m)$ such that for a function f we have $Y(\theta) = f(X_1, ..., X_m)$. Suppose that $\mathbf{X}$ has probability density function $g(\cdot)$. We assume that θ is a parameter of this density, such that the whole dependency of f on θ is incorporated in the density function of $\mathbf{X}$. For this reason, we write $g_\theta(\cdot)$ for the probability density function of $\mathbf{X}$. In this setting we have

$$\mathbb{E}[Y(\theta)] = \mathbb{E}[f(X_1, ..., X_m)] = \int_{\mathbb{R}^m} f(x) g_\theta(x) dx. \tag{4.5}$$

To find the LRM estimator, we need to assume that the order of differentiation and integration can be interchanged[1]. Glasserman [4] mentions that it is safe to assume that this assumption holds because most probability density functions are smooth functions with respect to their parameters. Using the assumption that integration and differentiation can be interchanged, we get

$$\begin{aligned}
\frac{d}{d\theta}\mathbb{E}[Y] &= \int_{\mathbb{R}^m} f(x) \frac{d}{d\theta} g_\theta(x) dx, \\
&= \int_{\mathbb{R}^m} f(x) \frac{\dot{g}_\theta(x)}{g_\theta(x)} g_\theta(x) dx, \\
&= \mathbb{E}\left[f(X) \frac{\dot{g}_\theta(X)}{g_\theta(X)} \right],
\end{aligned}$$

where $\dot{g} = dg_\theta/d\theta$. From this, it follows that

$$f(\mathbf{X}) \frac{\dot{g}_\theta(\mathbf{X})}{g_\theta(\mathbf{X})} \tag{4.6}$$

is an unbiased estimator of the derivative of $\mathbb{E}[Y]$. The expression $\dot{g}_\theta/g_\theta$, or equivalently $d \log g_\theta/d\theta$ is often called the score function.

To implement LRM in practice, it is essential to find a vector of risk factors $\mathbf{X} = (X_1, ..., X_m)$ with a known density function g, depending on θ. Once the risk factors have been chosen and the density function is known, the score function should be evaluated. This score function does not depend on the details of the discounted payoff. The score function, then, can be multiplied by any payoff $f(\mathbf{X})$ to get an estimate for the derivative of $\mathbb{E}[Y]$.

The fact that we can multiply any payoff with the score function to get an estimate for the derivative sounds promising. However, the risk factors, that need to be selected in advance, must be chosen with care. Their joint probability density function must be known, and the dependence of the function f on θ must be fully captured in this density function. Therefore, according to Palazzi, Conti, and Pioppi [2], the implementation of LRM in practice calls for the need to introduce some knowledge from the outside in the selection of risk factors. In Section 2.3, we have shown for the Hull-White model that most interest rate processes are normally distributed, implying that the density function in our framework is known. However, bear in mind that the density function is not always known, which can lead to difficulties in applying LRM.

The main drawback of LRM is that the LRM estimators are particularly affected by poor variance properties. In Section 6 we explain the nature of these variance problems. Furthermore, in Section 7 we present techniques to reduce the variance.

[1] In order to allow the interchanging of differentiation and integration we need that with probability 1, the derivatives of $X_i, i = 1, ..., m$ w.r.t. θ exists, and the derivative of Y w.r.t. to θ exists. Next to that, we need that Y is almost surely Lipschitz in θ. For more details see Broadie and Glasserman [3].

5 Estimation of CVA Deltas

In this section, we explain the mathematical details of the B&R method and LRM to compute three CVA sensitivities. First, in Section 5.1 we will explain how to compute one or multiple CVA zero deltas of an IR swap. Then, in Section 5.2 we will present the mathematics behind the computation of the CVA FX spot delta of an FX swap and an FX digital option.

5.1. CVA zero delta of an IR swap

The first sensitivity that we consider is the CVA zero delta of an IR swap. We let $Y(\theta)$ be given by (4.1), and we specify a set of exposure times by $\{T_1, ..., T_m\}$. We let the value of the FX swap at time T_k be given by (2.45).

Furthermore, we specify $q+1$ zero rates as inputs for our model. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T_0^p, T_1^p, ..., T_q^p$. Using these zero rates and pillar dates, we construct a yield curve, as explained in Section 2.2.2. The interpolation scheme that we use in this construction is given by (2.8).

Lastly, we choose the parameter of differentiation. In the one-dimensional case, we choose $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$. In the multi-dimensional case, we choose θ as a vector of p zero rates, i.e.

$$\theta = \begin{pmatrix} \theta_1 & \theta_2 & \cdots & \theta_p \end{pmatrix},$$

where for $j = 1, ..., p : \theta_j = z_i$ for some $z_i \in \{z_1, ..., z_q\}$.

We now present how to compute one or multiple CVA zero deltas of an IR swap using both the B&R method and LRM.

5.1.1. CVA zero delta - Bump & Reprice Method

In order to determine CVA zero deltas for an IR swap using the B&R method, we must select a bump size $h > 0$. For more information about the choice of bump sizes, we refer to Section 4.1. Once h has been chosen, we select the parameter of differentiation θ. In case $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$, then we estimate the CVA zero delta as

$$\hat{\Delta}_C(\theta) = \frac{\bar{Y}(\theta + h) - \bar{Y}(\theta - h)}{2h}.$$

In case θ is a vector of p zero rates, i.e. $\theta = \begin{pmatrix} \theta_1 & \theta_2 & \cdots & \theta_p \end{pmatrix}$, where for $j = 1, ..., p : \theta_j = z_i$ for some $z_i \in \{z_1, ..., z_d\}$, then we estimate all p CVA zero deltas separately, and combine them in one vector:

$$\hat{\Delta}_C(\theta) = \begin{pmatrix} \hat{\Delta}_C(\theta_1) & \hat{\Delta}_C(\theta_2) & \cdots & \hat{\Delta}_C(\theta_p) \end{pmatrix}.$$

5.1.2. CVA zero delta - Likelihood Ratio Method

In order to determine CVA zero deltas for an IR swap using LRM, we start by modifying $Y(\theta)$ (4.1) slightly:

$$
\begin{aligned}
Y(\theta) &= (1-R_C)\sum_{k=1}^{m}\frac{M_\theta(t_0)}{M_\theta(T_k)}\max(V_\theta(T_k),0)\bar{q}(T_{k-1},T_k),\\
&= (1-R_C)\sum_{k=1}^{m}\exp\left(-\int_{t_0}^{T_k} r_\theta(s)ds\right)\max(V_\theta(T_k),0)\bar{q}(T_{k-1},T_k).
\end{aligned}
$$

We added a subscript to the short rate process, meaning the short rate depends on θ. Consequently, the integral of the short rate also depends on θ.

We now fix $t_0 = 0$, which allows us to rewrite $Y(\theta)$ as

$$
Y(\theta) = (1-R_C)\sum_{k=1}^{m}\exp\left(-\int_{0}^{T_k} r_\theta(s)ds\right)\max(V_\theta(T_k),0)\bar{q}(T_{k-1},T_k)). \tag{5.1}
$$

The application of LRM relies on finding a vector of risk factors $\mathbf{X} = (X_1,...,X_m)$ such that $Y(\theta) = f(\mathbf{X})$ for some $f(\cdot)$, and such that the whole dependency of f on θ is incorporated in the density function of $\mathbf{X}$. This is a crucial assumption, and therefore we should always check if there is no hidden information of θ inside the payoff of a financial derivative. However, as we see in (2.7), the ZCB prices, that are part of the IR swap payoff (2.45), depend on θ. This results in our function f being dependent on θ. One way to overcome this is to choose multiple ZCB prices as risk factors. This incorporates the ZCB prices into the vector $\mathbf{X}$, ensuring the dependency of f on the ZCB prices is fully captured in the density function. However, just as Palazzi, Conti, and Pioppi [2] mentioned, this would only be possible if we would consider independent Hull White dynamics for each zero rate. This would not reflect sensible curve dynamics anymore, and hence this will not be considered in this thesis. Another way to overcome the issue of f being dependent on θ is simply to assume that f depends both on X and on θ. We then write $f := f(\mathbf{X},\theta)$ and we rewrite (4.5) as:

$$
\begin{aligned}
\frac{d}{d\theta}\mathbb{E}[Y] &= \frac{d}{d\theta}\int_{\mathbb{R}^m} f(x,\theta)g_\theta(x)dx\\
&= \int_{\mathbb{R}^m}\frac{d}{d\theta}\left(f(x,\theta)g_\theta(x)\right)dx\\
&= \int_{\mathbb{R}^m}\left(f(x,\theta)\frac{dg_\theta(x)}{d\theta} + \frac{df(x,\theta)}{d\theta}g_\theta(x)\right)dx\\
&= \int_{\mathbb{R}^m} f(x,\theta)\frac{\dot{g}_\theta(x)}{g_\theta(x)}g_\theta(x)dx + \int_{\mathbb{R}^m}\frac{df(x,\theta)}{d\theta}g_\theta(x)dx\\
&= \mathbb{E}\left[f(\mathbf{X},\theta)\frac{\dot{g}_\theta(\mathbf{X})}{g_\theta(\mathbf{X})}\right] + \mathbb{E}\left[\frac{df(\mathbf{X},\theta)}{d\theta}\right].
\end{aligned}
$$

This results in a LRM estimator that consists of two parts. In the first part, the whole dependency of f on θ is given in the density function of $\mathbf{X}$, therefore f only depends on $\mathbf{X}$. This part is independent of the discounted payoff and can be evaluated using LRM. For the second part, we need to differentiate f with respect to θ, creating a dependency on the discounted payoff. The vector of risk factors, $\mathbf{X}$, although depending on θ, will remain constant. That means, that once we found an analytical formula for $df(\mathbf{X},\theta)/d\theta$, then

$$
f(\mathbf{X},\theta)\frac{\dot{g}_\theta(\mathbf{X})}{g_\theta(\mathbf{X})} + \frac{df(\mathbf{X},\theta)}{d\theta}
$$

is an unbiased estimator of the CVA zero delta(s). Hence, despite the dependence of f on θ, LRM can still be applied for the computation of the CVA zero delta of an IR swap. However, one of the advantages of LRM - specifically, the avoidance of the differentiation of payoffs - is compromised.

We will now present the mathematical details of both the payoff-independent part and the payoff-dependent part of the LRM estimator.

Part 1: Payoff independent LRM

We use LRM to compute the payoff-independent part, hence we need to find a vector of risk factors **X** such that $f(\mathbf{X}) = Y(\theta)$. We define for $k = 1, ..., m$:

$$I(T_k) = \int_0^{T_k} r(u)du,$$

and we select the risk factors as:

$$\mathbf{X} = \begin{pmatrix} r(T_1) & r(T_2) & \cdots & r(T_m) & I(T_1) & I(T_2) & \cdots & I(T_m) \end{pmatrix}^T \tag{5.2}$$

The applicability of LRM relies on the dependency of the density function of **X** on θ. Indeed, using Theorem 2.3.5, Theorem 2.3.8, Theorem 2.3.9 and Theorem 2.3.10 we know that **X** given $\mathcal{F}(0)$ is multivariate normally distributed with mean $\mu(\theta)$ and covariance matrix Σ, i.e.,

$$\mathbf{X} \mid \mathcal{F}(0) \sim \mathcal{N}(\mu(\theta), \Sigma).$$

The vector of expectations $\mu(\theta)$ is given by

$$\mu(\theta) = \begin{pmatrix} \mu_{r,1}(\theta) & \mu_{r,2}(\theta) & \cdots \mu_{r,m}(\theta) & \mu_{I,1}(\theta) & \mu_{I,2}(\theta) & \cdots \mu_{I,m}(\theta) \end{pmatrix}^T \in \mathbb{R}^{2m \times 1},$$

where for $k = 1, ..., m$

$$\mu_{r,k}(\theta) = \mathbb{E}_0\left[r(T_k)\right] = f^M(0, T_k) + \frac{\sigma^2}{2a^2}\left(1 - e^{-aT_k}\right)^2, \tag{5.3}$$

$$\mu_{I,k}(\theta) = \mathbb{E}_0\left[I(T_k)\right] = \log\left(\frac{1}{P^M(0, T_k)}\right) + \frac{V(0, T_k)}{2}. \tag{5.4}$$

The covariance matrix Σ is given by

$$\Sigma = \begin{pmatrix} \Sigma^{r,r} & \Sigma^{r,I} \\ \Sigma^{I,r} & \Sigma^{I,I} \end{pmatrix} \in \mathbb{R}^{2m \times 2m},$$

where $\Sigma^{I,r} = (\Sigma^{r,I})^T$, and

$$\Sigma^{r,r} = \begin{pmatrix} \Sigma^{r,r}_{11} & \Sigma^{r,r}_{12} & \cdots & \Sigma^{r,r}_{1m} \\ \Sigma^{r,r}_{21} & \Sigma^{r,r}_{22} & \cdots & \Sigma^{r,r}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma^{r,r}_{m1} & \Sigma^{r,r}_{m2} & \cdots & \Sigma^{r,r}_{mm} \end{pmatrix}, \Sigma^{r,I} = \begin{pmatrix} \Sigma^{r,I}_{11} & \Sigma^{r,I}_{12} & \cdots & \Sigma^{r,I}_{1m} \\ \Sigma^{r,I}_{21} & \Sigma^{r,I}_{22} & \cdots & \Sigma^{r,I}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma^{r,I}_{m1} & \Sigma^{r,I}_{m2} & \cdots & \Sigma^{r,I}_{mm} \end{pmatrix}, \Sigma^{I,I} = \begin{pmatrix} \Sigma^{I,I}_{11} & \Sigma^{I,I}_{12} & \cdots & \Sigma^{I,I}_{1m} \\ \Sigma^{I,I}_{21} & \Sigma^{I,I}_{22} & \cdots & \Sigma^{I,I}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma^{I,I}_{m1} & \Sigma^{I,I}_{m2} & \cdots & \Sigma^{I,I}_{mm} \end{pmatrix}.$$

in which for $k, l = 1, ..., m$ with $k \leq l$ we have

$$\begin{aligned}
\Sigma^{r}_{kl} &= \text{Cov}_0\left[r(T_k), r(T_l)\right] \\
&= \frac{\sigma^2}{2a}\left(e^{-a(T_l - T_k)} - e^{-a(T_k + T_l)}\right), \\
\Sigma^{r,I}_{kl} &= \text{Cov}_0\left[r(T_k), \int_0^{T_l} r(u)du\right] \\
&= \frac{\sigma^2}{a}\left(\frac{1}{a}\left(1 - e^{-aT_k}\right) - \frac{1}{2a}\left(e^{-a(T_l - T_k)} - e^{-a(T_k + T_l)}\right)\right), \\
\Sigma^{I,I}_{kl} &= \text{Cov}_0\left[\int_0^{T_k} r(u)du, \int_0^{T_l} r(u)du\right] \\
&= \frac{\sigma^2}{a^2}\left(T_k - \frac{1 - e^{-aT_k} + e^{-a(T_l - T_k)} - e^{-aT_l}}{a} + \frac{e^{-a(T_l - T_k)} - e^{-a(T_k + T_l)}}{2a}\right),
\end{aligned}$$

and for $k > l$ we have

$$
\begin{aligned}
\Sigma^r_{kl} &= \mathrm{Cov}_0\left[r(T_k), r(T_l)\right] \\
&= \frac{\sigma^2}{2a}\left(e^{-a(T_k-T_l)} - e^{-a(T_l+T_k)}\right), \\
\Sigma^{r,I}_{kl} &= \mathrm{Cov}_0\left[r(T_k), \int_0^{T_l} r(u)du\right] \\
&= \frac{\sigma^2}{a}\left(\frac{1}{a}\left(e^{-a(T_k-T_l)} - e^{-aT_k}\right) - \frac{1}{2a}\left(e^{-a(T_k-T_l)} - e^{-a(T_l+T_k)}\right)\right), \\
\Sigma^{I,I}_{kl} &= \mathrm{Cov}_0\left[\int_0^{T_k} r(u)du, \int_0^{T_l} r(u)du\right] \\
&= \frac{\sigma^2}{a^2}\left(T_l - \frac{1 - e^{-aT_l} + e^{-a(T_k-T_l)} - e^{-aT_k}}{a} + \frac{e^{-a(T_k-T_l)} - e^{-a(T_l+T_k)}}{2a}\right).
\end{aligned}
$$

Once the risk factors are selected and simulated, we can find an expression for the score function of **X**. Since **X** is multivariate normally distributed with mean vector $\mu(\theta)$ and covariance matrix Σ, the density function of **X** is given by:

$$
g_\theta(\mathbf{X}) = \frac{1}{(2\pi)^{m/2}\,\mathrm{det}(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{X} - \mu(\theta))^T\Sigma^{-1}(\mathbf{X} - \mu(\theta))\right).
$$

Finding the score function is now easy, and is given by

$$
\frac{d}{d\theta}\log g_\theta(\mathbf{X}) = (\mathbf{X} - \mu(\theta))^T\Sigma^{-1}\dot{\mu}(\theta),
$$

where $\dot{\mu}(\theta)$ represents the vector or matrix of derivatives of the components of μ with respect to θ.

The fact that LRM is able to compute multiple CVA sensitivities in one single MC simulation is rooted in the structure of $\dot{\mu}(\theta)$. In case only one sensitivity is required, $\dot{\mu}(\theta)$ will be a vector in $\mathbb{R}^{2m\times 1}$. However, if p sensitivities are to be computed, $\dot{\mu}(\theta)$ will be a matrix in $\mathbb{R}^{2m\times p}$. It is therefore important to present the construction of $\dot{\mu}(\theta)$ in both cases. In the case of just one sensitivity, $\dot{\mu}(\theta)$ is given as follows.

Theorem 5.1.1. *Let $q+1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T^p_0, T^p_1, ..., T^p_q$. Assume $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$, and let for $k = 1, ..., m$*

$$
\begin{aligned}
\mu_{r,k} &= f^M(0, T_k) + \frac{\sigma^2}{2a^2}\left(1 - e^{-aT_k}\right)^2, \\
\mu_{I,k} &= \log\left(\frac{1}{P^M(0, T_k)}\right) + \frac{V(0, T_k)}{2},
\end{aligned}
$$

as in (5.3) *and* (5.4). *Then*

$$
\dot{\mu}(\theta) = \begin{pmatrix}\dot{\mu}_r(\theta) \\ \dot{\mu}_I(\theta)\end{pmatrix} \in \mathbb{R}^{2m\times 1},
$$

where

$$
\begin{aligned}
\dot{\mu}_r(\theta) &= \begin{pmatrix}\frac{\partial}{\partial\theta}\mu_{r,1} & \cdots & \frac{\partial}{\partial\theta}\mu_{r,j-1} & \frac{\partial}{\partial\theta}\mu_{r,j} & \frac{\partial}{\partial\theta}\mu_{r,j+1} & \cdots & \frac{\partial}{\partial\theta}\mu_{r,m}\end{pmatrix}^T, \\
\dot{\mu}_I(\theta) &= \begin{pmatrix}\frac{\partial}{\partial\theta}\mu_{I,1} & \cdots & \frac{\partial}{\partial\theta}\mu_{I,j-1} & \frac{\partial}{\partial\theta}\mu_{I,j} & \frac{\partial}{\partial\theta}\mu_{I,j+1} & \cdots & \frac{\partial}{\partial\theta}\mu_{I,m}\end{pmatrix}^T.
\end{aligned}
$$

If we take T_k such that $T^p_{j-1} \leq T_k < T^p_j$, then

$$\frac{\partial}{\partial \theta}\mu_{r,k} = \frac{2T_k - T^p_{j-1} - T^p_0}{T^p_j - T^p_{j-1}},$$
$$\frac{\partial}{\partial \theta}\mu_{I,k} = \frac{(T_k - T^p_{j-1})(T_k - T^p_0)}{T^p_j - T^p_{j-1}}.$$

On the other hand, if we take T_k such that $T^p_j \leq T_k < T^p_{j+1}$, then

$$\frac{\partial}{\partial \theta}\mu_{r,k} = 1 - \frac{2T_k - T^p_j - T^p_0}{T^p_{j+1} - T^p_j},$$
$$\frac{\partial}{\partial \theta}\mu_{I,k} = T_k - T^p_0 - \frac{(T_k - T^p_j)(T_k - T^p_0)}{T^p_{j+1} - T^p_j}.$$

Lastly, for all T_k such that $T_k < T^p_{j-1}$ or $T_k \geq T^p_{j+1}$, we obtain

$$\frac{\partial}{\partial \theta}\mu_{r,k} = 0,$$
$$\frac{\partial}{\partial \theta}\mu_{I,k} = 0.$$

Proof. For the proof of the theorem, we refer to Appendix A.3.1. □

In a more general case, θ will be the vector of $p \leq q$ zero rates. In that case $\dot{\mu}(\theta)$ is given as follows.

Theorem 5.1.2. *Let $d+1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_d$, with corresponding pillar dates $T^p_0, T^p_1, ..., T^p_d$. Assume θ is a vector of p zero rates, i.e. $\theta = \begin{pmatrix}\theta_1 & \theta_2 & \cdots & \theta_p\end{pmatrix}$, where for $j = 1, ..., p : \theta_j = z_i$ for some $z_i \in \{z_1, ..., z_d\}$. If we let for $k = 1, ..., m$*

$$\mu_{r,k} = f^M(0, T_k) + \frac{\sigma^2}{2a^2}\left(1 - e^{-aT_k}\right)^2,$$
$$\mu_{I,k} = \log\left(\frac{1}{P^M(0, T_k)}\right) + \frac{V(0, T_k)}{2},$$

as in (5.3) *and* (5.4)*, then*

$$\dot{\mu}(\theta) = \begin{pmatrix}\dot{\mu}_r(\theta_1) & \dot{\mu}_r(\theta_2) & \cdots & \dot{\mu}_r(\theta_p) \\ \dot{\mu}_I(\theta_1) & \dot{\mu}_I(\theta_2) & \cdots & \dot{\mu}_I(\theta_p)\end{pmatrix} \in \mathbb{R}^{2m \times p},$$

where for $j = 1, ..., p$

$$\dot{\mu}_r(\theta_j) = \begin{pmatrix}\frac{\partial}{\partial \theta_j}\mu_{r,1} & \cdots & \frac{\partial}{\partial \theta_j}\mu_{r,j-1} & \frac{\partial}{\partial \theta_j}\mu_{r,j} & \frac{\partial}{\partial \theta_j}\mu_{r,j+1} & \cdots & \frac{\partial}{\partial \theta_j}\mu_{r,m}\end{pmatrix}^T,$$
$$\dot{\mu}_x(\theta_j) = \begin{pmatrix}\frac{\partial}{\partial \theta_j}\mu_{I,1} & \cdots & \frac{\partial}{\partial \theta_j}\mu_{I,j-1} & \frac{\partial}{\partial \theta_j}\mu_{I,j} & \frac{\partial}{\partial \theta_j}\mu_{I,j+1} & \cdots & \frac{\partial}{\partial \theta_j}\mu_{I,m}\end{pmatrix}^T$$

are given as in Theorem 5.1.1

Proof. For the proof of the theorem, we can simply repeat the proof of Theorem 5.1.1 p times. □

Part 2: Payoff dependent differentiation

In order to determine the payoff-dependent part of the LRM estimator, we need to differentiate the function $f(\mathbf{X}, \theta)$ with respect to θ. The function $f(\mathbf{X}, \theta)$ is given as in (5.1):

$$f(\mathbf{X}, \theta) = (1 - R_C) \sum_{k=1}^{m} \text{exp} \left(- \int_0^{T_k} r(s) ds \right) \text{max}(V_\theta(T_k), 0) \bar{q}(T_{k-1}, T_k)).$$

Note that this derivative is only taken with respect to θ, meaning that the process **X** (5.2) will be fixed. Therefore we have removed the dependency of $r(t)$ on θ. Differentiating $f(\mathbf{X}, \theta)$ with respect to θ therefore yields

$$\frac{\partial f(\mathbf{X}, \theta)}{\partial \theta} = (1 - R_C) \sum_{k=1}^{m} \text{exp} \left(- \int_0^{T_k} r(s) ds \right) \frac{\partial}{\partial \theta} \left(\text{max}(V_\theta(T_k), 0) \right) \bar{q}(T_{k-1}, T_k).$$

To obtain a closed-form solution of this derivative, we introduce a few theorems. Each theorem is only given for the single sensitivity case, but can easily be expanded to the multiple sensitivity case. We start by stating the derivative of the ratio of two ZCBs from the market. This derivative depends on the choice of the parameter of differentiation θ. Therefore, we distinguish 7 cases.

Theorem 5.1.3. *Let $q + 1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T_0^p, T_1^p, ..., T_q^p$. Assume $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$. Lastly we assume $T_k \in \{T_1, ..., T_m\}$, and $T_i > T_k$. Then*

$$\frac{\partial}{\partial \theta} \left(\frac{P^M(0, T_i)}{P^M(0, T_k)} \right) = C(T_k, T_i) \frac{P^M(0, T_i)}{P^M(0, T_k)},$$

where

$$C(T_k, T_i) = \begin{cases} -\frac{(T_i - T_{j-1}^p)(T_i - T_0)}{T_j^p - T_{j-1}^p} & \text{for } T_k < T_{j-1}^p \leq T_i < T_j^p, \\ -(T_i - T_0) + \frac{(T_i - T_j^p)(T_i - T_0)}{T_{j+1}^p - T_j^p} & \text{for } T_k < T_{j-1}^p < T_j^p \leq T_i < T_{j+1}^p, \\ -\frac{(T_i - T_{j-1}^p)(T_i - T_0) - (T_k - T_{j-1}^p)(T_k - T_0)}{T_j^p - T_{j-1}^p} & \text{for } T_{j-1}^p \leq T_k < T_i < T_j^p, \\ -(T_i - T_0) + \frac{(T_i - T_j^p)(T_i - T_0)}{T_{j+1}^p - T_j^p} + \frac{(T_k - T_{j-1}^p)(T_k - T_0)}{T_j^p - T_{j-1}^p} & \text{for } T_{j-1}^p \leq T_k < T_j^p \leq T_i < T_{j+1}^p, \\ \frac{(T_k - T_{j-1}^p)(T_k - T_0)}{T_j^p - T_{j-1}^p} & \text{for } T_{j-1}^p \leq T_k < T_j^p < T_{j+1}^p \leq T_i, \\ -(T_i - T_k) + \frac{(T_i - T_j^p)(T_i - T_0) - (T_k - T_j^p)(T_k - T_0)}{T_{j+1}^p - T_j^p} & \text{for } T_j^p \leq T_k < T_i < T_{j+1}^p, \\ T_k - T_0 - \frac{(T_k - T_j^p)(T_k - T_0)}{T_{j+1}^p - T_j^p} & \text{for } T_j^p \leq T_k < T_{j+1}^p \leq T_i, \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.3.2. □

Now we present the derivative of future ZCBs.

Theorem 5.1.4. *Let $q + 1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T_0^p, T_1^p, ..., T_q^p$. Assume $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$. Lastly we assume $T_k \in \{T_1, ..., T_m\}$, and $T_i > T_k$. If we let $P(T_k, T_i)$ be given, as in Theorem 2.23, by*

$$P(T_k, T_i) = \frac{P^M(0, T_i)}{P^M(0, T_k)} \text{exp} \left(B(T_k, T_i)(f^M(0, T_k) - r(T_k)) - \frac{\sigma^2}{4a}(1 - e^{-2aT_k}) B(T_k, T_i)^2 \right),$$

then

$$\frac{\partial}{\partial \theta} P(T_k, T_i) = P(T_k, T_i) \left(C(T_k, T_i) + B(T_k, T_i) D(T_k) \right)),$$

where $C(T_k, T_i)$ is given as in Theorem 5.1.3 and

$$D(T_k) = \begin{cases} \frac{2T_k - T_{j-1}^p - T_0}{T_j^p - T_{j-1}^p} & \text{for } T_{j-1}^p \leq T_k < T_j^p, \\ 1 - \frac{2T_k - T_j^p - T_0}{T_{j+1}^p - T_j^p} & \text{for } T_j^p \leq T_k < T_{j+1}^p. \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.3.3. □

Using the derivative of the ZCBs, we determine the derivative of the payoff of an IR swap.

Theorem 5.1.5. *Let $q+1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T_0^p, T_1^p, ..., T_q^p$. Assume $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$ and let $T_k \in \{T_1, ..., T_m\}$. If we let $V_\theta(T_k)$ equal the value of the IR swap at time T_k* (2.45)*, and we denote the future payment dates by $\{T_{j+1}, ..., T_n\}$, then*

$$\frac{\partial V_\theta(T_k)}{\partial \theta} = \bar{\beta} Q \left(\frac{\partial P(T_k, T_j)}{\partial \theta} - \frac{\partial P(T_k, T_n)}{\partial \theta} \right) - \bar{\beta} Q K \sum_{i=j+1}^{n} \tau_i \frac{\partial P(T_k, T_i)}{\partial \theta},$$

where for $k = 1, ..., m$ and $i = j, ..., n$ the derivative $\frac{\partial P(T_k, T_i)}{\partial \theta}$ is given as in Theorem 5.1.4.

Proof. For the proof of the theorem, we refer to Appendix A.3.4. □

We end this section with the derivative of $f(\mathbf{X}, \theta)$ with respect to θ.

Corollary 5.1.6. *Let $f(\mathbf{X}, \theta)$ be given by*

$$f(\mathbf{X}, \theta) = (1 - R_C) \sum_{k=1}^{m} \exp\left(- \int_0^{T_k} r(s) ds \right) \max(V_\theta(T_k), 0) \bar{q}(T_{k-1}, T_k)).$$

Then

$$\frac{\partial f(\theta)}{\partial \theta} = (1 - R_C) \sum_{k=1}^{m} \exp\left(- \int_0^{T_k} r(s) ds \right) \frac{\partial}{\partial \theta} \left(\max(V_\theta(T_k), 0) \right) \bar{q}(T_{k-1}, T_k),$$

where

$$\frac{\partial}{\partial \theta} \left(\max(V_\theta(T_k), 0) \right) = \begin{cases} \frac{\partial}{\partial \theta} V_\theta(T_k) & \text{if } V_\theta(T_k) > 0, \\ 0 & \text{if } V_\theta(T_k) \leq 0. \end{cases}$$

The derivative $\frac{\partial}{\partial \theta} V_\theta(T_k)$ is given by Theorem 5.1.5.

Proof. For the proof of the corollary, we simply write out the maximum operator as

$$\max(V_\theta(T_k), 0) = \begin{cases} V_\theta(T_k) & \text{if } V_\theta(T_k) > 0, \\ 0 & \text{if } V_\theta(T_k) \leq 0. \end{cases}$$

Then we take the derivative on both sides. □

5.2. CVA FX spot delta of an FX swap and an FX digital option

The second sensitivity that we consider is the CVA FX spot delta of FX swaps and FX digital options. We specify two currencies, the domestic currency (e.g., euro €), and the foreign currency (e.g., dollar $). Then with $y(t)$ we express the units of the euro per unit of dollar. To determine the CVA FX spot delta for an FX swap or FX digital option, we take $\theta = y(t_0)$.

Furthermore, we again let $Y(\theta)$ be given by (4.1), and we specify a set of exposure times by $\{T_1, ..., T_m\}$. We let the value of the FX swap at time T_k be given by (2.46) and the value of the FX digital option be given by (2.48).

Then we construct two yield curves. We specify $2(q+1)$ zero rates as inputs for our model, with corresponding pillar dates $T_0^p, T_1^p, ..., T_q^p$. For the domestic currency, we use a set of zero rates $\{z_0^d, ..z_q^d\}$, and for the foreign currency, we use a set of zero rates $\{z_0^f, ..z_q^f\}$. Using these zero rates we construct two yield curves, see Section 2.4.2 for the details. The interpolation schemes that we use in this construction are given by (2.25) and (2.26).

We now present how to compute the CVA FX spot deltas of an FX swap and an FX digital option using both the B&R method and LRM.

5.2.1. CVA FX spot delta - Bump & Reprice method

In order to determine the CVA FX spot delta of an FX swap or FX digital option using the B&R method, we choose a bump size $h > 0$. For more information on the choice of bump sizes, we refer to Section 4.1. Once h has been chosen, we select the parameter of differentiation, $\theta = y(t_0)$, and we estimate the CVA FX spot delta as

$$\hat{\Delta}_C(\theta) = \frac{\bar{Y}(\theta + h) - \bar{Y}(\theta - h)}{2h}.$$

5.2.2. CVA FX spot delta - Likelihood Ratio Method

In a similar way as we did for the CVA zero deltas of IR swaps, we start by modifying $Y(\theta)$ (4.1) slightly, to obtain

$$\begin{aligned} Y(\theta) &= (1 - R_C) \sum_{k=1}^{m} \frac{M_{d,\theta}(t_0)}{M_{d,\theta}(T_k)} \max(V_\theta(T_k), 0)\bar{q}(T_{k-1}, T_k), \\ &= (1 - R_C) \sum_{k=1}^{m} \exp\left(-\int_{t_0}^{T_k} r_{d,\theta}(s)ds\right) \max(V_\theta(T_k), 0)\bar{q}(T_{k-1}, T_k), \end{aligned}$$

where we added a subscript to the domestic currency process, meaning the domestic currency depends on θ. We also added the subscript 'd' to indicate that the values are quoted in the domestic currency.

We now fix $t_0 = 0$, which allows us to rewrite $Y(\theta)$ as

$$Y(\theta) = (1 - R_C) \sum_{k=1}^{m} \exp\left(-\int_{0}^{T_k} r_{d,\theta}(s)ds\right) \max(V_\theta(T_k), 0)\bar{q}(T_{k-1}, T_k)). \tag{5.5}$$

The application of LRM relies on finding a vector of risk factors **X** such that $Y(\theta) = f(\mathbf{X})$. We define for $k = 1, ..., m$:

$$I_d(T_k) = \int_0^{T_k} r_d(u)du,$$

and we select the risk factors as:

$$\mathbf{X} = \begin{pmatrix} r_d(T_1) & \cdots & r_d(T_m) & r_f(T_1) & \cdots & r_f(T_m) & I_d(T_1) & \cdots & I_d(T_m) & L_y(T_1) & \cdots & L_y(T_m) \end{pmatrix}^T \in \mathbb{R}^{4m\times 1}.$$

The applicability of LRM relies on the dependency of the density function of **X** on θ. We can prove that **X** given $\mathcal{F}(0)$ is multivariate normally distributed with mean $\mu(\theta)$ and covariance matrix Σ, i.e.,

$$\mathbf{X} \mid \mathcal{F}(0) \sim \mathcal{N}(\mu(\theta), \Sigma).$$

Using Theorem 2.3.5, Theorem 2.6.3, Theorem 2.3.8 and Theorem 2.6.8, the vector of expectations $\mu(\theta)$ is given by

$$\mu(\theta) = \begin{pmatrix}\mu_{r_d,1}(\theta) & \cdots & \mu_{r_d,m}(\theta) & \mu_{r_f,1}(\theta) & \cdots & \mu_{r_f,m}(\theta) & \mu_{I_d,1}(\theta) & \cdots & \mu_{I_d,m}(\theta) & \mu_{L_y,1}(\theta) & \cdots & \mu_{L_y,m}(\theta)\end{pmatrix}^T,$$

where for $k = 1, ..., m$

$$\mu_{r_d,k}(\theta) = \mathbb{E}_0\left[r_d(T_k)\right] = f_d^M(0,T_k) + \frac{\sigma_d^2}{2a_d^2}\left(1 - e^{-a_d T_k}\right)^2, \tag{5.6}$$

$$\mu_{r_f,k}(\theta) = \mathbb{E}_0\left[r_f(T_k)\right] = f_f^M(0,T_k) + \frac{\sigma_f^2}{2a_f^2}\left(1 - e^{-a_f T_k}\right)^2 - \sigma_f\sigma_y\rho_{y,f}B_f(0,T_k), \tag{5.7}$$

$$\mu_{I_d,k}(\theta) = \mathbb{E}_0\left[I_d(T_k)\right] = \log\left(\frac{1}{P_d^M(0,T_k)}\right) + \frac{V_d(0,T_k)}{2}, \tag{5.8}$$

$$\mu_{L_y,k}(\theta) = \mathbb{E}_0\left[L_y(T_k)\right] = \log y(0) + \log\left(\frac{1}{P_d^M(0,T_k)}\right) + \frac{V_d(0,T_k)}{2} + \frac{\sigma_f\sigma_y\rho_{y,f}}{a_f}\left(T_k - B_f(0,T_k)\right) \tag{5.9}$$

$$- \log\left(\frac{1}{P_f^M(0,T_k)}\right) - \frac{V_f(0,T_k)}{2} - \frac{1}{2}\sigma_y^2 T_k. \tag{5.10}$$

The covariance matrix is given by

$$\Sigma = \begin{pmatrix} \Sigma^{r_d,r_d} & \Sigma^{r_d,r_f} & \Sigma^{r_d,I_d} & \Sigma^{r_d,L_y} \\ \Sigma^{r_f,r_d} & \Sigma^{r_f,r_f} & \Sigma^{r_f,I_d} & \Sigma^{r_f,L_y} \\ \Sigma^{I_d,r_d} & \Sigma^{I_d,r_f} & \Sigma^{I_d,I_d} & \Sigma^{I_d,L_y} \\ \Sigma^{L_y,r_d} & \Sigma^{L_y,r_f} & \Sigma^{L_y,I_d} & \Sigma^{L_y,L_y} \end{pmatrix} \in \mathbb{R}^{4m\times 4m},$$

where for $z_1, z_2 \in \{r_d, r_f, I_d, L_y\}$: $\Sigma^{z_2,z_1} = (\Sigma^{z_1,z_2})^T$, and

$$\Sigma^{z_1,z_2} = \begin{pmatrix} \Sigma_{11}^{z_1,z_2} & \Sigma_{12}^{z_1,z_2} & \cdots & \Sigma_{1m}^{z_1,z_2} \\ \Sigma_{21}^{z_1,z_2} & \Sigma_{22}^{z_1,z_2} & \cdots & \Sigma_{2m}^{z_1,z_2} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{m1}^{z_1,z_2} & \Sigma_{m2}^{z_1,z_2} & \cdots & \Sigma_{mm}^{z_1,z_2} \end{pmatrix}.$$

By Theorem 2.3.5, Theorem 2.3.9, Theorem 2.3.10, Theorem 2.6.3, and Theorem 2.6.9 we know that for $k, l = 1, ..., m$ with $k \leq l$ we have

$$\begin{aligned}
\Sigma_{kl}^{r_d,r_d} &= \text{Cov}_0\left[r_d(T_k), r_d(T_l)\right] \\
&= \frac{\sigma_d^2}{2a_d}\left(e^{-a_d(T_l - T_k)} - e^{-a_d(T_k + T_l)}\right), \\
\Sigma_{kl}^{r_f,r_f} &= \text{Cov}_0\left[r_f(T_k), r_f(T_l)\right] \\
&= \frac{\sigma_f^2}{2a_f}\left(e^{-a_f(T_l - T_k)} - e^{-a_f(T_k + T_l)}\right), \\
\Sigma_{kl}^{I_d,I_d} &= \text{Cov}_0\left[\int_0^{T_k} r_d(u)du, \int_0^{T_l} r_d(u)du\right] \\
&= \frac{\sigma_d^2}{a_d^2}\left(T_k - \frac{1 - e^{-a_d T_k} + e^{-a_d(T_l - T_k)} - e^{-a_d T_l}}{a_d} + \frac{e^{-a_d(T_l - T_k)} - e^{-a_d(T_k + T_l)}}{2a_d}\right), \\
\Sigma_{kl}^{L_y,L_y} &= \text{Cov}_0[L_y(T_k), L_y(T_l)] \\
&= \frac{\sigma_d^2}{a_d^2}\left(T_k - B_d(0,T_k) - \frac{e^{-a_d(T_l - T_k)} - e^{-a_d T_l}}{a_d} + \frac{e^{-a_d(T_l - T_k)} - e^{-a_d(T_k + T_l)}}{2a_d}\right) \\
&\quad + \frac{\sigma_f^2}{a_f^2}\left(T_k - B_f(0,T_k) - \frac{e^{-a_f(T_l - T_k)} - e^{-a_f T_l}}{a_f} + \frac{e^{-a_f(T_l - T_k)} - e^{-a_f(T_k + T_l)}}{2a_f}\right)
\end{aligned}$$

$$
\begin{aligned}
&- \frac{\sigma_d \sigma_f}{a_d a_f}\rho_{d,f}\Bigg(T_k - B_d(0,T_k) - \frac{e^{-a_f(T_l-T_k)} - e^{-a_f T_l}}{a_f} + \frac{e^{-a_f(T_l-T_k)} - e^{-a_d T_k - a_f T_l}}{a_d + a_f} \\
&+ T_k - B_f(0,T_k) - \frac{e^{-a_d(T_l-T_k)} - e^{-a_d T_l}}{a_d} + \frac{e^{-a_d(T_l-T_k)} - e^{-a_f T_k - a_d T_l}}{a_d + a_f}\Bigg) \\
&+ \frac{\sigma_d \sigma_y}{a_d}\rho_{y,d}\left(T_k - B_d(0,T_k) + T_k - \frac{e^{-a_d(T_l-T_k)} - e^{-a_d T_l}}{a_d}\right) \\
&- \frac{\sigma_f \sigma_y}{a_f}\rho_{y,f}\left(T_k - B_f(0,T_k) + T_k - \frac{e^{-a_f(T_l-T_k)} - e^{-a_f T_l}}{a_f}\right) + \sigma_y^2 T_k,
\end{aligned}
$$

$$
\begin{aligned}
\Sigma_{kl}^{r_d,I_d} &= \mathsf{Cov}_0\left(r_d(T_k), \int_0^{T_l} r_d(u)du\right) \\
&= \frac{\sigma_d^2}{a_d}\left(\frac{1-e^{-a_d T_k}}{a_d} - \frac{e^{-a_d(T_l-T_k)} - e^{-a_d(T_k+T_l)}}{2a_d}\right),
\end{aligned}
$$

and that for $k > l$ we have

$$
\begin{aligned}
\Sigma_{kl}^{r_d,r_d} &= \mathsf{Cov}_0\left[r_d(T_k), r_d(T_l)\right] \\
&= \frac{\sigma_d^2}{2a_d}\left(e^{-a_d(T_k-T_l)} - e^{-a_d(T_l+T_k)}\right), \\
\Sigma_{kl}^{r_f,r_f} &= \mathsf{Cov}_0\left[r_f(T_k), r_f(T_l)\right] \\
&= \frac{\sigma_f^2}{2a_f}\left(e^{-a_f(T_k-T_l)} - e^{-a_f(T_l+T_k)}\right), \\
\Sigma_{kl}^{I_d,I_d} &= \mathsf{Cov}_0\left[\int_0^{T_k} r_d(u)du, \int_0^{T_l} r_d(u)du\right] \\
&= \frac{\sigma_d^2}{a_d^2}\left(T_l - \frac{1-e^{-a_d T_l} + e^{-a_d(T_k-T_l)} - e^{-a_d T_k}}{a_d} + \frac{e^{-a_d(T_k-T_l)} - e^{-a_d(T_l+T_k)}}{2a_d}\right), \\
\Sigma_{kl}^{L_y,L_y} &= \mathsf{Cov}_0[L_y(T_k), L_y(T_l)] \\
&= \frac{\sigma_d^2}{a_d^2}\left(T_l - B_d(0,T_l) - \frac{e^{-a_d(T_k-T_l)} - e^{-a_d T_k}}{a_d} + \frac{e^{-a_d(T_k-T_l)} - e^{-a_d(T_l+T_k)}}{2a_d}\right) \\
&+ \frac{\sigma_f^2}{a_f^2}\left(T_l - B_f(0,T_l) - \frac{e^{-a_f(T_k-T_l)} - e^{-a_f T_k}}{a_f} + \frac{e^{-a_f(T_k-T_l)} - e^{-a_f(T_l+T_k)}}{2a_f}\right) \\
&- \frac{\sigma_d \sigma_f}{a_d a_f}\rho_{d,f}\Bigg(T_l - B_d(0,T_l) - \frac{e^{-a_f(T_k-T_l)} - e^{-a_f T_k}}{a_f} + \frac{e^{-a_f(T_k-T_l)} - e^{-a_d T_l - a_f T_k}}{a_d + a_f} \\
&+ T_l - B_f(0,T_l) - \frac{e^{-a_d(T_k-T_l)} - e^{-a_d T_k}}{a_d} + \frac{e^{-a_d(T_k-T_l)} - e^{-a_f T_l - a_d T_k}}{a_d + a_f}\Bigg) \\
&+ \frac{\sigma_d \sigma_y}{a_d}\rho_{y,d}\left(T_l - B_d(0,T_l) + T_l - \frac{e^{-a_d(T_k-T_l)} - e^{-a_d T_k}}{a_d}\right) \\
&- \frac{\sigma_f \sigma_y}{a_f}\rho_{y,f}\left(T_l - B_f(0,T_l) + T_l - \frac{e^{-a_f(T_k-T_l)} - e^{-a_f T_k}}{a_f}\right) + \sigma_y^2 T_l, \\
\Sigma_{kl}^{r_d,I_d} &= \mathsf{Cov}_0\left(r_d(T_k), \int_0^{T_l} r_d(u)du\right) \\
&= \frac{\sigma_d^2}{a_d}\left(\frac{e^{-a_d(T_k-T_l)} - e^{-a_d T_k}}{a_d} - \frac{e^{-a_d(T_k-T_l)} - e^{-a_d(T_l+T_k)}}{2a_d}\right).
\end{aligned}
$$

For the other covariance structures we introduce a few theorems. In these theorems, we fix the filtration to $\mathcal{F}(0)$, meaning we only look at the information available today.

Theorem 5.2.1. *For any* $0 \leq T_k, T_l$*, the covariance of* $r_d(T_k)$ *and* $r_f(T_l)$ *given* $\mathcal{F}(0)$ *is given by*

$$\text{Cov}_0\left[r_d(T_k), r_f(T_l)\right] = \begin{cases} \frac{\sigma_d\sigma_f}{a_d+a_f}\rho_{d,f}\left(e^{-a_f(T_l-T_k)} - e^{-a_dT_k-a_fT_l}\right) & T_k \leq T_l, \\ \frac{\sigma_d\sigma_f}{a_d+a_f}\rho_{d,f}\left(e^{-a_d(T_k-T_l)} - e^{-a_fT_l-a_dT_k}\right) & T_k > T_l. \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.4.1. □

Theorem 5.2.2. *For any* $0 \leq T_k, T_l$*, the covariance of* $r_d(T_k)$ *and* $L_y(T_l)$ *given* $\mathcal{F}(0)$ *is given by*

$$\text{Cov}_0\left[r_d(T_k), L_y(T_l)\right] = \begin{cases} \frac{\sigma_d^2}{a_d}\left(\frac{1-e^{-a_dT_k}}{a_d} - \frac{e^{-a_d(T_l-T_k)}-e^{a_d(T_k+T_l)}}{2a_d}\right) & \\ \quad -\frac{\sigma_d\sigma_f}{a_f}\rho_{d,f}\left(\frac{1-e^{-a_dT_k}}{a_d} - \frac{e^{-a_f(T_l-T_k)}-e^{-a_dT_k-a_fT_l}}{a_d+a_f}\right) & T_k \leq T_l, \\ \quad +\sigma_d\sigma_y\rho_{y,d}\frac{1-e^{-a_dT_k}}{a_d} & \\ \frac{\sigma_d^2}{a_d}\left(\frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} - \frac{e^{-a_d(T_k-T_l)}-e^{a_d(T_l+T_k)}}{2a_d}\right) & \\ \quad -\frac{\sigma_d\sigma_f}{a_f}\rho_{d,f}\left(\frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} - \frac{e^{-a_d(T_k-T_l)}-e^{-a_fT_l-a_dT_k}}{a_d+a_f}\right) & T_k > T_l. \\ \quad +\sigma_d\sigma_y\rho_{y,d}\frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} & \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.4.2. □

Theorem 5.2.3. *For any* $0 \leq T_k, T_l$*, the covariance of* $r_f(T_k)$ *and* $\int_0^{T_l} r_d(u)du$ *given* $\mathcal{F}(0)$ *is given by*

$$\text{Cov}_0\left[r_f(T_k), \int_0^{T_l} r_d(u)du\right] = \begin{cases} \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{1-e^{-a_fT_k}}{a_f} - \frac{e^{-a_d(T_l-T_k)}-e^{-a_fT_k-a_dT_l}}{a_d+a_f}\right) & T_k \leq T_l, \\ \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_k-T_l)}-e^{-a_dT_l-a_fT_k}}{a_d+a_f}\right) & T_k > T_l. \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.4.3. □

Theorem 5.2.4. *For any* $0 \leq T_k, T_l$*, the covariance of* $r_f(T_k)$ *and* $L_y(T_l)$ *given* $\mathcal{F}(0)$ *is given by*

$$\text{Cov}_0\left[r_f(T_k), L_y(T_l)\right] = \begin{cases} \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{1-e^{-a_fT_k}}{a_f} - \frac{e^{-a_d(T_l-T_k)}-e^{a_fT_k-a_dT_l}}{a_d+a_f}\right) & \\ \quad -\frac{\sigma_f^2}{a_f}\left(\frac{1-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_l-T_k)}-e^{-a_f(T_k+T_l)}}{2a_f}\right) & T_k \leq T_l, \\ \quad +\sigma_f\sigma_y\rho_{y,f}\frac{1-e^{-a_fT_k}}{a_f} & \\ \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_k-T_l)}-e^{a_dT_l-a_fT_k}}{a_d+a_f}\right) & \\ \quad -\frac{\sigma_f^2}{a_f}\left(\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_k-T_l)}-e^{-a_f(T_l+T_k)}}{2a_f}\right) & T_k > T_l. \\ \quad +\sigma_f\sigma_y\rho_{y,f}\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} & \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.4.4. □

Theorem 5.2.5. *For any* $0 \leq T_k, T_l$*, the covariance of* $\int_0^{T_k} r_d(u)du$ *and* $L_y(T_l)$ *given* $\mathcal{F}(0)$ *is given by*

$$\text{Cov}\left[\int_0^{T_k} r_d(s)ds, L_y(T_l)\right] = \begin{cases} \frac{\sigma_d^2}{a_d^2}\left(T_k - \frac{1-e^{-a_dT_k}+e^{-a_d(T_l-T_k)}-e^{-a_dT_l}}{a_d} + \frac{e^{-a_d(T_l-T_k)}-e^{-a_d(T_k+T_l)}}{2a_d}\right) & \\ \quad -\frac{\sigma_d\sigma_f}{a_da_f}\rho_{d,f}\left(T_k - \frac{1-e^{-a_dT_k}}{a_d} - \frac{e^{-a_f(T_l-T_k)}-e^{-a_fT_l}}{a_f}\right. & T_k \leq T_l, \\ \quad \left.+\frac{e^{-a_f(T_l-T_k)}-e^{-a_dT_k-a_fT_l}}{a_d+a_f}\right) + \frac{\sigma_d\sigma_y\rho_{y,d}}{a_d}\left(T_k - \frac{1-e^{-a_dT_k}}{a_d}\right) & \\ \frac{\sigma_d^2}{a_d^2}\left(T_l - \frac{1-e^{-a_dT_l}+e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} + \frac{e^{-a_d(T_k-T_l)}-e^{-a_d(T_l+T_k)}}{2a_d}\right) & \\ \quad -\frac{\sigma_d\sigma_f}{a_da_f}\rho_{d,f}\left(T_l - \frac{1-e^{-a_fT_l}}{a_f} - \frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d}\right. & T_k \leq T_l. \\ \quad \left.+\frac{e^{-a_d(T_k-T_l)}-e^{-a_fT_l-a_dT_k}}{a_d+a_f}\right) + \frac{\sigma_d\sigma_y\rho_{y,d}}{a_d}\left(T_l - \frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d}\right) & \end{cases}$$

Proof. For the proof of the theorem, we refer to Appendix A.4.5. □

Once the risk factors are selected and simulated, we can find an expression for the score function of **X**. Since **X** is multivariate normally distributed with mean vector $\mu(\theta)$ and covariance matrix Σ, the density function is given by:

$$g_\theta(\mathbf{X}) = \frac{1}{(2\pi)^{m/2}\,\mathrm{det}(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{X}-\mu(\theta))^T\Sigma^{-1}(\mathbf{X}-\mu(\theta))\right).$$

Finding the score function is again easy to find, and is given by

$$\frac{d}{d\theta}\log g_\theta(\mathbf{X}) = (\mathbf{X}-\mu(\theta))^T\Sigma^{-1}\dot{\mu}(\theta),$$

where $\dot{\mu}(\theta)$ represents the vector of derivatives of the components of μ with respect to θ. For $\theta = \log y(0)$, $\dot{\mu}(\theta)$ is given as follows.

Theorem 5.2.6. *Let* $\theta = \log y(0)$*. Let for* $k = 1, ..., m$

$$\begin{aligned}
\mu_{r_d,k}(\theta) &= f_d^M(0,T_k) + \frac{\sigma_d^2}{2a_d^2}\left(1-e^{-a_dT_k}\right)^2,\\
\mu_{r_f,k}(\theta) &= f_f^M(0,T_k) + \frac{\sigma_f^2}{2a_f^2}\left(1-e^{-a_fT_k}\right)^2 - \sigma_f\sigma_y\rho_{y,f}B_f(0,T_k),\\
\mu_{I_d,k}(\theta) &= \log\left(\frac{1}{P_d^M(0,T_k)}\right) + \frac{V_d(0,T_k)}{2},\\
\mu_{L_y,k}(\theta) &= \log y(0) + \log\left(\frac{1}{P_d^M(0,T_k)}\right) + \frac{V_d(0,T_k)}{2} + \frac{\sigma_f\sigma_y\rho_{y,f}}{a_f}\left(T_k - B_f(0,T_k)\right)\\
&\quad - \log\left(\frac{1}{P_f^M(0,T_k)}\right) - \frac{V_f(0,T_k)}{2} - \frac{1}{2}\sigma_yT_k,
\end{aligned}$$

as in (5.6)*,* (5.7)*,* (5.8) *and* (5.10)*. Then*

$$\dot{\mu}(\theta) = \begin{pmatrix}\dot{\mu}_{r_d}(\theta)\\ \dot{\mu}_{r_f}(\theta)\\ \dot{\mu}_{I_d}(\theta)\\ \dot{\mu}_{L_y}(\theta)\end{pmatrix} \in \mathbb{R}^{4m\times 1},$$

where

$$\begin{aligned}
\dot{\mu}_{r_d}(\theta) &= \begin{pmatrix}\frac{\partial}{\partial\theta}\mu_{r_d,1} & \cdots & \frac{\partial}{\partial\theta}\mu_{r_d,m}\end{pmatrix}^T,\\
\dot{\mu}_{r_f}(\theta) &= \begin{pmatrix}\frac{\partial}{\partial\theta}\mu_{r_f,1} & \cdots & \frac{\partial}{\partial\theta}\mu_{r_f,m}\end{pmatrix}^T,\\
\dot{\mu}_{I_d}(\theta) &= \begin{pmatrix}\frac{\partial}{\partial\theta}\mu_{I_d,1} & \cdots & \frac{\partial}{\partial\theta}\mu_{I_d,m}\end{pmatrix}^T,\\
\dot{\mu}_{L_y}(\theta) &= \begin{pmatrix}\frac{\partial}{\partial\theta}\mu_{L_y,1} & \cdots & \frac{\partial}{\partial\theta}\mu_{L_y,m}\end{pmatrix}^T,
\end{aligned}$$

and where for $k = 1, ..., m$

$$\begin{aligned}
\frac{\partial}{\partial\theta}\mu_{r_d,k} &= 0,\\
\frac{\partial}{\partial\theta}\mu_{r_f,k} &= 0,\\
\frac{\partial}{\partial\theta}\mu_{I_d,k} &= 0,\\
\frac{\partial}{\partial\theta}\mu_{L_y,k} &= \frac{1}{y(0)}.
\end{aligned}$$

Proof. The proof is trivial. It is easy to see that the derivatives are all zero, except for the derivative with respect to θ of the mean of $L_y(t)$ for $t \in \{T_1, ..., T_m\}$. □

6

Error Analysis

In the computation of CVA sensitivities, the implementation is susceptible to a few theoretical errors. These errors can be split into three categories:

- Numerical integration errors,
- MC errors,
- LRM errors.

We will briefly address the first two errors in Section 6.1 and Section 6.2. Then in Section 6.3, we will delve into a more detailed examination of the errors that are introduced in the implementations due to LRM.

6.1. Numerical integration errors

Regarding the integration error, we recall that the value of CVA at time t_0 can be written as

$$\text{CVA}(t_0) = (1 - R_C) \int_{t_0}^{T} \mathbb{E}\left[\frac{M(t_0)}{M(T_D)} \max(V(T_D), 0) \mid T_D = z\right] dF_{T_D}(z), \tag{6.1}$$

where R_C is the recovery fraction $\in [0, 1]$, $M(t)$ is the value of the money savings account at time t, T_D denotes the time of default of the counterparty, F_{T_D} is the cumulative distribution function of the default time T_D, and $V(T_k)$ denotes the value of a financial derivative at time T_D. Finding a closed-form expression for this integral is difficult. Therefore, we make an approximation. We specify a set of exposure times $\{T_1, ..., T_m\}$ with $T_m \leq T$. Then CVA at time t_0 can be approximated by

$$\text{CVA}(t_0) = (1 - R_C) \sum_{k=1}^{m} \text{EE}(t_0, T_k)\bar{q}(T_{k-1}, T_k), \tag{6.2}$$

with R_C the recovery fraction $\in [0, 1]$, the expected positive exposure given by

$$\text{EE}(t, T_k) = \mathbb{E}_t\left[\frac{M(t)}{M(T_k)} \max(V(T_k), 0)\right],$$

and the PD in (T_{k-1}, T_k) given by

$$\bar{q}(T_{k-1}, T_k) = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{T_{k-1} < T_D \leq T_k}\right].$$

The effectiveness of the approximation in (6.2) relies on the number of exposure dates m. Choosing the number of exposure dates requires us to make a trade-off between precision and computational efficiency. A small number of exposure dates leads to a loss of accuracy in the approximation. A large number of exposure dates can lead to a computationally expensive implementation. Striking a balance between precision and computational efficiency, we opt for 10 exposure dates per year in our simulations. It is difficult to justify whether this number of exposure dates is large enough to accurately estimate the integral (6.1). However, we use the same number of exposure dates in each simulation and for each method to compute the CVA sensitivities. This ensures that any integration errors introduced are systematic, affecting all methods and simulations uniformly. Therefore, the comparison between the methods still gives useful insights.

6.2. MC simulation errors

The second type of error is a result of MC simulations, which are essential for the computation of the expectation in (6.2) to obtain the value of CVA and its sensitivities. In general, in order to compute the expectation of random variable C, a MC method simulates a sample of N i.i.d. realizations $C_1, ..., C_N$ and uses the average of the N realizations, denoted as $\hat{C}_N$, as an estimate for the expectation of C. Each C_i has expected value $\mathbb{E}[C_i] = C$ and a variance of Var$[C_i] = \sigma_C^2$. The central limit theorem tells us that

$$\hat{C}_N - C \approx \mathcal{N}(0, \sigma_C^2/N).$$

Hence, the standard deviation of $\hat{C}_N - C$ is approximately $\sigma_C/\sqrt{N}$, resulting in a MC convergence rate of $\mathcal{O}(1/\sqrt{N})$.

This rate signifies that quadrupling the number of MC realizations, N, results in an error half as big. It is therefore important to use a large number of replications to obtain a good estimate of the expectation of C. To improve the convergence, we could make use of QMC methods. These methods have the possibility to accelerate the convergence from $\mathcal{O}(1/\sqrt{N})$ to nearly $\mathcal{O}(1/N)$. Nevertheless, it is important to note that a small error will persist despite these improvements.

6.3. LRM errors

The last type of error arises from LRM. Recall that LRM requires the choice of a vector of risk factors **X** with probability density function $g_\theta(\mathbf{X})$. The LRM estimator is then given by

$$f(\mathbf{X})\frac{\dot{g}_\theta(\mathbf{X})}{g_\theta(\mathbf{X})}.$$

This estimator serves as an unbiased estimator of the CVA sensitivity. As we previously mentioned, the main drawback of LRM is that the LRM estimators are affected by poor variance properties. One of the reasons for these poor variance properties is that the expectation of the score function is equal to 0:

$$\begin{aligned}\mathbb{E}\left[\frac{\dot{g}_\theta(\mathbf{X})}{g_\theta(\mathbf{X})}\right] &= \int \frac{\dot{g}_\theta(x)}{g_\theta(x)} g_\theta(x)dx \\ &= \frac{d}{d\theta}\int 1 \cdot g_\theta(x)dx = 0.\end{aligned}$$

This implies that the score function is a martingale process, and thus a process of increasing variance. Furthermore, the fact that the expectation of the score function is zero implies that the LRM estimator has no definite sign. According to Capriotti [14] and Palazzi, Conti, and Pioppi [2], this can lead to poor variance properties due to the cancellation effects in the MC average of MC configurations that have opposite signs but comparable sizes. Hence, the MC average is established by adding up small residues that result from these cancellation effects. As a result, the variance of the LRM estimator is often not only difficult to predict, but can also become very large.

Besides this general reason for the poor variance properties of LRM, there are a few model parameters that significantly impact the variance of the LRM estimator. The following analysis demonstrates how various model parameters affect the variance of the LRM estimator, beginning with an examination of the impact of high volatility values of the short rate processes.

High volatility values of the short rate

As previously mentioned, the MC average is established by adding up small residues that result from cancellation effects. In the case of high volatility, the MC average is even more affected by these cancellation effects. In particular, high volatility values result in increased absolute values for both the risk factors and the MC configurations. Consequently, the cancellation effects give rise to larger residues. The sum of these residues contributes to a higher variance in the MC averages. As a consequence, the variance of the LRM estimator experiences an increase in the case of high volatility.

Highly correlated risk factors

Highly correlated risk factors also have a significant impact on the variance of the LRM estimator. Since the risk vectors in this thesis are normally distributed, we consider a normally distributed process **X** with mean vector $\mu(\theta)$ and covariance matrix Σ. We have seen in Section 5 that the score function is given by

$$\frac{d}{d\theta}\log g_\theta(\mathbf{X}) = (\mathbf{X}-\mu(\theta))^T\Sigma^{-1}\dot{\mu}(\theta),$$

where $\dot{\mu}(\theta)$ represents the derivatives of the components of μ with respect to θ. It is important to note that the computation of the score function requires the inversion of the covariance matrix.

In most cases, the inversion of the covariance matrix poses no issues. However, according to Palazzi, Conti, and Pioppi [2], the presence of highly correlated risk factors can cause problems for the inversion of the covariance matrix. In particular, for extreme (positive/negative) correlations between two processes, the joint probability density function of the processes becomes almost degenerate. LRM then loses its ability to distinguish between the two contributions. As a result, the determinant of the covariance matrix tends to go to zero or even equals zero. Matrices with a determinant of zero are called singular, and for these matrices, it is not possible to find an inverse matrix. In case the determinant of the covariance matrix is not zero, but is very small, then the inversion of the covariance matrix is possible. However, the resulting inverse matrix will have very large values due to the small determinant. Consequently, extreme correlations result in an inverse covariance matrix with very large values, causing an explosion of the variance of the score function. As a result, we can expect a higher variance of the LRM estimator in case of extreme correlations.

Low values of FX volatility

A choice of low FX volatility can have a great impact on the variance of the LRM estimator. We demonstrate this by concentrating on the log-normal model of the FX spot rate, and performing a similar analysis as Glasserman [4] did in Example 7.3.1. For the log-normal model of a single asset, he showed that the variance of the score function diverges in the limit of small volatility. We will show here a similar conclusion.

We define the FX spot rate at time T as $y(T)$, with volatility parameter σ_y. Furthermore, we denote with $r_d(t)$ and $r_f(t)$ the domestic and foreign currency, respectively. From (2.40), it can be shown that for $t_0 \leq T$ we have

$$\begin{aligned} y(T) &= \exp\left(L_y(T)\right) \\ &= y(t_0)\exp\left(\bar{\mu}(t_0,T) + Z\sqrt{\bar{\sigma}^2(t_0,T)}\right), \end{aligned}$$

where Z is a standard normal random variable, and

$$\begin{aligned} \bar{\mu}(t_0,T) &= \mathbb{E}_{t_0}\left[L_y(T)\right] - \log(y(t_0)), \\ \bar{\sigma}^2(t_0,T) &= \mathrm{Var}_{t_0}\left[L_y(T)\right]. \end{aligned}$$

If we view $y(t_0)$ as a parameter of $y(T)$, then the log-normal probability density function of $y(T)$ is

$$g_\theta(x) = \frac{1}{x\sqrt{\bar{\sigma}^2(t_0,T)}}\phi(\zeta(x)), \quad \zeta(x) = \frac{\log(x/y(t_0)) - \bar{\mu}(t_0,T)}{\sqrt{\bar{\sigma}^2(t_0,T)}},$$

with ϕ being the standard normal density. We can show that the score function is given by

$$\frac{\partial \log g_\theta(y(T))}{\partial y(t_0)} = \zeta(y(T))\frac{\partial \zeta(y(T))}{\partial y(t_0)} = \frac{\log(y(T)/y(t_0)) - \bar{\mu}(t_0,T)}{y(t_0)\bar{\sigma}^2(t_0,T)}.$$

Since $y(T)$ is generated from $y(t_0)$, we have $\zeta(y(T)) = Z$ for the standard normal variable Z. Therefore the score function can be simplified to:

$$\frac{\partial \log g_\theta(y(T))}{\partial y(t_0)} = \frac{Z}{y(t_0)\sqrt{\bar{\sigma}^2(t_0,T)}}. \tag{6.3}$$

We know that the mean of the score function equals zero. For the variance of the score function, we obtain:

$$\begin{aligned}\text{Var}\left[\frac{\partial \log g_\theta(y(T))}{\partial y(t_0)}\right] &= \text{Var}\left[\frac{Z}{y(t_0)\sqrt{\bar{\sigma}^2(t_0,T)}}\right]\\ &= \mathbb{E}\left[\frac{Z^2}{y^2(t_0)\bar{\sigma}^2(t_0,T)}\right] - \mathbb{E}\left[\frac{Z}{y(t_0)\sqrt{\bar{\sigma}^2(t_0,T)}}\right]^2\\ &= \frac{1}{y^2(t_0)\bar{\sigma}^2(t_0,T)} \xrightarrow[\bar{\sigma}^2(t_0,T)\to 0]{} \infty.\end{aligned}$$

We conclude that the variance of the score function diverges when $\bar{\sigma}^2(t_0,T)$ is small. Note that $\bar{\sigma}^2(t_0,T)$ equals $\text{Var}_{t_0}[L_y(T)]$, or equivalently $\text{Cov}_{t_0}[L_y(T), L_y(T)]$, as in Theorem 2.6.9. It is evident that $\bar{\sigma}^2(t_0,T)$ approaches zero when the FX volatility σ_y goes to zero. Therefore, we conclude that the variance of the score function diverges in the limit of small FX volatility. As a result, we also expect the variance of the LRM estimator to explode when the FX volatility goes to zero.

Number of exposure dates

Lastly, we examine the impact of the number of exposure dates on the variance of the LRM estimator. According to Glasserman [4], when we define a process $\{S(T_1),...,S(T_m)\}$ and want to compute the derivative with respect to θ of the expectation $\mathbb{E}[f(S(T_1),...,S(T_m))]$ for some function $f(\cdot)$, then the score function will be given by the sum of all the individual score functions:

$$\frac{\log g_\theta(S(T_1),...,S(T_m))}{\partial\theta} = \sum_{j=1}^{m}\frac{\log g_\theta(S(T_j))}{\partial\theta}.$$

Since each separate score function has mean zero, the mean of the sum of all score functions also equals zero. The sum is therefore a martingale, and thus the sum is a process of increasing variance. This observation holds for any process $\{S(T_1),...,S(T_m)\}$, and therefore also for the process of risk vectors that we employ in this thesis. We therefore conclude that the variance of the score function increases, once the number of exposure dates m increases. That means that, if we fix the spacings $T_j - T_{j-1}$ constant, we will observe a higher variance of the LRM estimators once the maturities of financial derivatives get longer.

For the log-normal model of the FX spot rate we present a more detailed analysis of the above observation by a decomposition of the score function. We already showed the score function of $y(T)$ in the case of one exposure date, see (6.3). When using m exposure dates, we use the fact that the score function is given by the sum of all separate m score functions. Therefore we specify m independent standard normal random variables, $Z_1,...,Z_m$, and rewrite the score function (6.3) as

$$\begin{aligned}\frac{\log g_\theta(y(T_1),...,y(T_m))}{\partial y(t_0)} &= \sum_{j=1}^{m}\frac{\log g_\theta(y(T_j)\mid y(t_0))}{\partial y(t_0)}\\ &= \sum_{j=1}^{m}\frac{Z_j}{y(t_0)\sqrt{\bar{\sigma}^2(t_0,T_j)}}.\end{aligned}$$

The mean of the score function equals zero again. Its variance is given by

$$\begin{aligned}\text{Var}\left[\frac{\log g_\theta(y(T_1),...,y(T_m))}{\partial y(t_0)}\right] &= \text{Var}\left[\sum_{j=1}^{m}\frac{Z_j}{y(t_0)\sqrt{\bar{\sigma}^2(t_0,T_j)}}\right]\\ &= \sum_{j=1}^{m}\text{Var}\left[\frac{Z_j}{y(t_0)\sqrt{\bar{\sigma}^2(t_0,T_j)}}\right]\\ &= \frac{1}{y^2(t_0)}\sum_{j=1}^{m}\frac{1}{\bar{\sigma}^2(t_0,T_j)}.\end{aligned}$$

Since $\bar{\sigma}^2(t_0, T_j)$ equals $\text{Var}_{t_0}[L_y(T_j)]$, this term will always be positive. Hence, the variance of the score function grows once the number of exposure dates m increases. Therefore, we conclude that the variance of the LRM estimator increases once the maturities of financial derivatives get longer.

We can arrive at the same conclusion through an alternative approach. Instead of computing score functions, we examine the covariance matrix of the vector of risk factors. This matrix encapsulates the covariances among the risk factors between different exposure times. We can observe that for two exposure dates, that are closely positioned but far in the future, the variance of the risk factors at both dates and the covariance of the risk factors between those dates exhibit minimal differences, which results in a high correlation between those risk factors. This observation implies that extending the maturity dates leads to heightened correlation among the risk factors, and hence a higher variance of the LRM estimator.

7

Variance Reduction Techniques

In this section, we present three variance reduction techniques. As we have shown in Section 6.2, MC methods exhibit a slow convergence rate, primarily due to the fact that the error scales as the square root of the number of simulations. Moreover, as outlined in Section 6.3, LRM is susceptible to high variance problems. To mitigate these issues, variance reduction techniques can be used. These techniques aim to reduce the variance of the LRM estimators, allowing for the attainment of a specified level of accuracy with a reduced number of MC simulations. Among a diverse range of techniques available, two standard techniques are antithetic sampling and control variates. These methods improve the efficiency of the simulation. On the other hand, we can use QMC methods to reduce the variance. These are methods that have the ability to achieve a faster convergence rate than ordinary MC methods.

In Section 7.1 we show how antithetic sampling can be applied in order to reduce the variance. Section 7.2 provides insights into the use of control variates. Lastly, Section 7.3 explains what QMC methods are, and demonstrates their application in financial settings to obtain lower variances.

7.1. Antithetic Sampling

In this section, we show the ideas behind antithetic sampling to reduce the variance. According to Palazzi, Conti, and Pioppi [2], antithetic sampling is an efficient variance reduction tool when the LRM antithetic estimators feature high (negative) correlation.

Suppose we are estimating $\mathbb{E}[Y]$. The method of antithetic sampling simulates two sets of replications, $Y_1, ..., Y_N$ and $\tilde{Y}_1, ..., \tilde{Y}_N$, where all Y_i and $\tilde{Y}_i$ depend on a sequence of multivariate standard normal random variables. The first N replications $Y_1, ..., Y_N$ are constructed using a sequence of i.i.d. multivariate normal random variables $Z_1, ..., Z_N \sim \mathcal{N}(0, I)$. To introduce antithetic sampling, this sequence is paired with its mirror images $-Z_1, ..., -Z_N$ of i.i.d $\mathcal{N}(0, I)$ variables to generate N estimates $\tilde{Y}_1, ..., \tilde{Y}_N$.

In more detail, the pairs $(Y_1, \tilde{Y}_1), ..., (Y_N, \tilde{Y}_N)$ are simulated i.i.d., where for each i, Y_i is constructed using the sequence $Z_1, ..., Z_N$ and $\tilde{Y}_i$ is constructed using the sequence $-Z_1, ..., -Z_N$. The antithetic sampling estimator is then obtained by averaging all $2N$ observations:

$$\hat{Y}_{\text{AV}} = \frac{1}{2N}\left(\sum_{i=1}^{N} Y_i + \sum_{i=1}^{N} \tilde{Y}_i\right) = \frac{1}{N}\sum_{i=1}^{N} \frac{Y_i + \tilde{Y}_i}{2}.$$

That makes $\hat{Y}_{\text{AV}}$ the sample mean of the $2N$ observations.

7.2. Control Variates

In this section, the method of control variates is outlined. According to Glasserman [4], this method is one of the most effective methods for improving the efficiency of MC simulations.

The method uses the information about the errors in estimates of known quantities to reduce the error in an estimate of an unknown quantity. Formally, besides simulating the unknown quantities $Y_1, ..., Y_N$, control variates involve the simulation of known quantities $X_1, ..., X_N$. Each X_i is a vector consisting of m elements: $X_i = \left(X_i^{(1)}, ..., X_i^{(m)}\right)^T$. It is assumed that the vector of expectations $\mathbb{E}[X]$ of the X_i is known, and that the pairs $(X_1, Y_1), ..., (X_N, Y_N)$ are i.i.d. with covariance matrix given by

$$\begin{pmatrix} \Sigma_X & \Sigma_{XY} \\ \Sigma_{XY}^T & \sigma_Y^2 \end{pmatrix},$$

where Σ_X is the $m \times m$ covariance matrix of the X_i, Σ_{XY} is the $m \times 1$ vector of covariances between X_i and Y_i, and the scalar σ_Y^2 is the variance of the Y_i. Then for fixed $b \in \mathbb{R}^m$ the control variate estimator is given by

$$\bar{Y}(b) = \bar{Y} - b^T(\bar{X} - \mathbb{E}[X]),$$

where the part $\bar{X} - \mathbb{E}[X]$ serves as a control in estimating $\mathbb{E}[Y]$.

It is easy to see that $\bar{Y}(b)$ is an unbiased estimator of $\mathbb{E}[Y]$. Furthermore we can show that b^* minimizes the variance of $\bar{Y}(b)$ and is given by

$$b^* = \Sigma_X^{-1}\Sigma_{XY}.$$

In practice, Σ_{XY} is usually unknown. However, Glasserman [4] states that we can replace Σ_{XY} with its sample counterpart to get

$$\hat{b}_N = \Sigma_X^{-1} S_{XY},$$

where S_{XY} is the m-vector with j-th entry

$$\frac{1}{N-1}\sum_{i=1}^{N}\left(X_i^{(j)}Y_i - N\bar{X}^{(j)}\bar{Y}\right).$$

Employing a control variate appears to be a promising way for variance reduction. However, when combining a normally distributed control variate with antithetic sampling, the impact on variance reduction is minimal to none. This observation becomes apparent by considering the antithetic pairs $(Y_1, \tilde{Y}_1), ..., (Y_N, \tilde{Y}_N)$, as explained in Section 7.1, and by defining two antithetic control variates, X and $\tilde{X}$.

We assume that these control variates are normally distributed with mean vector μ and covariance matrix Σ, such that $\Sigma = AA^T$. Then for a standard normally distributed vector Z we can write X and $\tilde{X}$ as

$$X = \mu + AZ, \quad \tilde{X} = \mu - AZ.$$

The control variate estimators are then given by:

$$\begin{aligned} Y(b) &= Y - b^T(X - \mathbb{E}[X]) = Y - b^T(\mu + AZ - \mu) = Y - b^T AZ, \\ \tilde{Y}(b) &= \tilde{Y} - b^T(\tilde{X} - \mathbb{E}[\tilde{X}]) = \tilde{Y} - b^T(\mu - AZ - \mu) = \tilde{Y} + b^T AZ. \end{aligned}$$

Combining these expressions, we obtain the mixed estimator as:

$$\begin{aligned} \hat{Y}_{\text{AV + CV}}(b) &= \frac{1}{N}\sum_{i=1}^{N}\frac{Y_i(b) + \tilde{Y}_i(b)}{2} \\ &= \frac{1}{N}\sum_{i=1}^{N}\frac{Y_i - (b^T AZ)_i + \tilde{Y}_i + (b^T AZ)_i}{2} \\ &= \frac{1}{N}\sum_{i=1}^{N}\frac{Y_i + \tilde{Y}_i}{2}. \end{aligned}$$

We observe that the combined estimator is exactly the same as the antithetic sampling estimator. Therefore, the impact of using a normally distributed control variate is very limited if antithetic sampling is also employed.

7.3. QMC methods

In this section, we present QMC methods and explain how these methods can be used in financial settings to accelerate the convergence of a MC simulation. QMC methods are similar to ordinary MC methods. However, the pseudo-random numbers are replaced with low-discrepancy sequences. According to Green [12], discrepancy is a measure of the deviation from uniformity, and hence low-discrepancy sequences have the property of low deviation from uniformity. This makes these sequences more evenly distributed than pseudo-random numbers. It can be shown that low-discrepancy methods can accelerate convergence from $\mathcal{O}(1/\sqrt{N})$ to almost $\mathcal{O}(1/N)$. Out of all low-discrepancy methods, Sobol' sequences have been shown to perform the best for financial applications. For more details about Sobol' sequences, we refer to Chapter 5.2.3 in Glasserman [4]. Despite their promising convergence rates, low-discrepancy methods come with a drawback: it is a challenge to obtain an estimation of the standard deviation, especially in high dimensions. Since the number of exposure dates in the CVA framework is usually quite high, accurately estimating the standard deviations in our simulation framework can become difficult.

The application of the Koksma-Hlawka inequality allows us to establish an upper bound for the convergence rate [4]. The upper bound is shown to be a convergence rate of $\mathcal{O}(\log(N)^d/N)$, where d is the dimension of the problem. Clearly, for a small d and large N, this convergence rate is much better than the $\mathcal{O}(1/\sqrt{N})$ convergence rate of an ordinary MC method. However, for most practical applications the dimension d is too large, leading to a gross overestimation of the convergence rate.

Therefore, various research is done with the aim of finding a better convergence rate than the one given by the Koksma-Hlawka inequality. Bianchetti, Kucherenko, and Scoleri [15] stress that low-discrepancy sequences are deterministic, meaning that statistical measures like variances are not valid. To overcome this limitation, Owen [16] came up with the idea of scrambled low-discrepancy sequences. This is a way of making the deterministic low-discrepancy sequences random. It seems odd to randomize QMC points, but Glasserman [4] mentions that in this way the best features of ordinary MC and QMC methods are combined. Even better, randomizing the sequences can provide us with a measure of the standard deviation of the QMC methods.

Renzitti, Bastani, and Sivorot [17] mention that in most numerical tests aiming to estimate the standard deviation of QMC methods the power law is observed, which states that

$$\sigma_{\text{QMC}} = \frac{c_1}{N^{c_2}}, \tag{7.1}$$

where c_1, c_2 are unknown constants that need to be estimated empirically. In case c_2 is close to 1/2, the QMC method will perform similarly to the ordinary Method Carlo methods. But when $c_2 > 1/2$, the QMC method will outperform the ordinary MC methods.

The constants c_1 and c_2 can be determined using the root mean squared error (RMSE). We denote with V the exact value of $\partial\mathbb{E}[Y]/\partial\theta$, either given by a reference value or estimated through the B&R method with a very large number of MC simulations. Then, for a total of J times, we will estimate the value of V with a QMC method using N low-discrepancy sequences. With $V_N^{(j)}$ we denote the jth run performed using these N low-discrepancy sequences. In each jth run, a different seed is used. After determining V and the list of values $V_N^{(1)}, ..., V_N^{(J)}$ is obtained, we estimate the empirical standard deviation using N simulations as:

$$\hat{\sigma}_{\text{QMC},N} = \sqrt{\frac{1}{J-1}\sum_{j=1}^{J}\left(V - V_N^{(j)}\right)^2}.$$

By repeating this procedure for L values of $N_1, ..., N_L$, then c_1 and c_2 can be estimated using a linear least squares regression of the logarithm of the L standard deviations $\{\hat{\sigma}_{\text{QMC},N_1}, ..., \hat{\sigma}_{\text{QMC},N_L}\}$ onto the logarithm of the number of low-discrepancy sequences $\{N_1, ..., N_L\}$. Once c_1 and c_2 are found, an empirical estimate of the standard deviation of the QMC method can be obtained by (7.1).

8

Numerical Results

In this section, we present the application of LRM to the computation of CVA deltas of several OTC derivatives. We start by showing the performance of LRM in computing CVA FX spot deltas of FX digital options in Section 8.1. Section 8.2 presents LRM's capability in the computation of CVA FX spot deltas of FX swaps, and we end with demonstrating the effectiveness of LRM in the CVA zero delta computation of IR swaps in Section 8.3. Throughout these three sections, the B&R method will serve as a benchmark method to which we compare the performance of LRM. The goal is the show LRM's ability to align with the B&R values.

Since the application of LRM can lead to high variance, we will consistently make use of variance reduction techniques in order to reduce the variance. Specifically, we will apply Antithetic Sampling (see Section 7.1) and we introduce a Control Variate (see Section 7.2) in order to reduce the variance. However, it is worth noting that the impact of the control variate on the variance reduction is limited, as discussed in more detail in Section 7.2. Additionally, for LRM, we replace the random numbers with low-discrepancy sequences, employing a QMC method. The use of Sobol' sequences has the potential to drastically reduce the variance. However, as we explained in Section 7.3, the downside of using this technique is the inability to determine a precise value for the standard deviation. Therefore, we rely on empirical estimates of the standard deviation that are found by the power law. Again, we refer to Section 7.3 for more details about this estimation. This implies that all the QMC standard deviations of the sensitivities that we present in the upcoming sections are purely empirical estimates. The 'true' standard deviations can in fact be slightly lower. This is, however, not easy to justify.

8.1. CVA FX spot deltas of FX digital options

In this section, we present the performance of LRM in the computation of CVA FX spot deltas of FX digital options, and we compare this performance against the B&R method and against a reference value. The reference value is obtained by summing the analytical derivatives of the expected exposures at all exposure times, as in Theorem 3.1.2, and then multiplying the sum by the LGD and the PD.

As previously mentioned, we will employ LRM with a QMC method, while the B&R method serves as a benchmark without a QMC method. For the specific comparison of CVA FX spot deltas of FX digital options, we also present the results for the B&R method with a QMC method. We believe this is necessary due to the non-continuous payoff of FX digital options (see Section 2.7.3), which may impact the B&R method's performance. This leaves us with three methods to compare: LRM + QMC, B&R, and B&R + QMC.

In Section 8.1.1, we will present the set of parameters used in the comparisons. To implement the B&R method, the selection of a bump size h is crucial. Since the payoff of FX digital options is not continuous, the variance of the B&R estimator depends on the chosen bump size h, and a trade-off has to be made between the bias and the variance. In Section 8.1.2, we will find an optimal bump size by comparing the errors and standard deviations of multiple CVA FX spot deltas among multiple values of h. Then in Section 8.1.3, we will check the MC errors of our implementations and compare them with the theoretical MC errors. Lastly, in Section 8.1.4, we provide a comparison between the performance of LRM and the B&R method in computing CVA FX spot deltas of FX digital options.

8.1.1. Set of parameters

In this section, we present the set of parameters that we will use in our analysis. Our parameter selection aligns closely with the choices made by Oosterlee and Grzelak [6] in their numerical experiments. A few parameters, however, were chosen based on available data. According to the European Central Bank [18] the short rate of EUR on 25-10-2023 equals 0.03904. Additionally, Bloomberg [19] indicates a short rate of USD on the same date at 0.0538, with an exchange rate between USD en EUR at 0.9433.

Recall that in order to approximate the integral in the CVA formula (3.1) through the summation in (3.2), we specify a set of exposure times $\{T_1, ..., T_m\}$. Choosing the number of exposure dates, m, requires us to make a trade-off between precision and computational efficiency. As we explained in Section 6.1, we opt for 10 exposure dates per year.

With this information, we create the set $\mathcal{P}_0$ containing all the parameter values as follows:

$$\mathcal{P}_0 = \left\{ \begin{array}{lcrlcr} T & = & 3, & m & = & 30, \\ t_0 & = & 0, & y(t_0) & = & 0.9433, \\ r_d(t_0) & = & 0.03904, & r_f(t_0) & = & 0.0538, \\ \sigma_y & = & 0.1, & \rho_{y,d} & = & -0.25, \\ \rho_{y,f} & = & -0.25, & \rho_{d,f} & = & 0.25, \\ a_d & = & 0.02, & \sigma_d & = & 0.01, \\ a_f & = & 0.04, & \sigma_f & = & 0.015, \\ N_f & = & 10^6, & K & = & 0.9433, \\ R_c & = & 0, & \xi & = & 0.002. \end{array} \right\}. \tag{8.1}$$

In the upcoming sections, we consistently employ the set $\mathcal{P}_0$ as the set of parameters for our comparisons, unless stated otherwise. Furthermore, for the LRM simulations, we will always make use of $N = 2^{18}$ MC simulations. For a fair comparison between LRM and the B&R method, we will adjust the number of B&R simulations so that the CPU times of both methods are approximately equal.

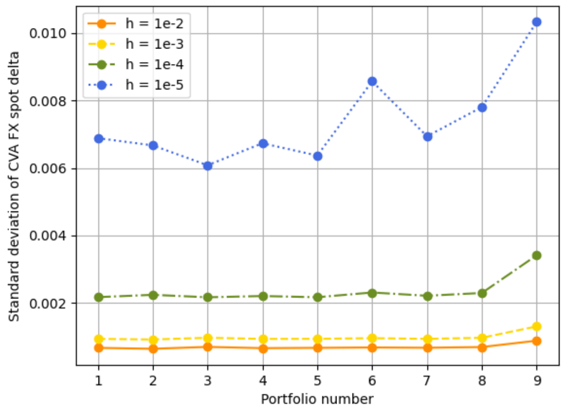
8.1.2. Bump size selection

Given the non-continuous nature of the payoff of FX digital options (see Section 2.7.3), the variance of the central difference estimator is influenced by the bump size h, see Section 4.1 for more details. A small h can lead to extreme variances, while a large h can result in a big bias. Therefore, a trade-off should be made between the bias and the variance. In this section, we find an optimal bump size. Given the influence of MC errors on our simulations, finding a value of the bias of our estimators is challenging. Instead, we focus on the errors introduced in our simulations. We will compare the relative errors and standard deviations of multiple FX digital call options for different values of h. With relative errors and standard deviations, we refer to the values of these quantities divided by the reference value that we computed. To facilitate a meaningful comparison, we evaluate the relative errors and standard deviations, computed by both B&R and B&R + QMC.

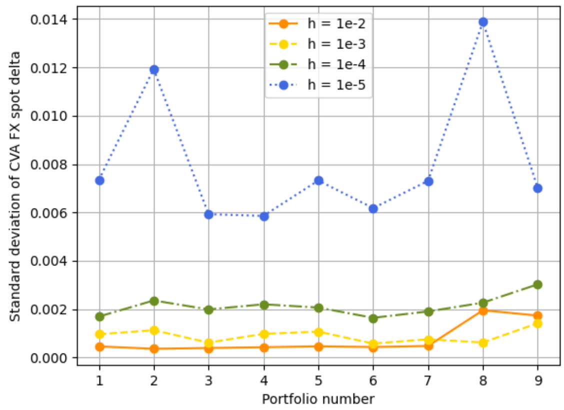
We examine four different bump sizes: $[10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}]$, and we select nine portfolios, each featuring a single FX digital call option. All nine options use the parameter set $\mathcal{P}_0$, with a modification made to one of the parameters. In Figure 8.1a and 8.1b we present for both methods the relative standard deviations of the CVA FX spot deltas of the nine FX digital call options, considering four different bump sizes for each method. It is evident that both methods exhibit high relative standard deviations for a bump size of $h = 10^{-4}$ and $h = 10^{-5}$. In contrast, a choice of $h = 10^{-2}$ or $h = 10^{-3}$ clearly outperforms the other options, aligning with the expectation that a smaller h generally leads to a higher variance.

In Figure 8.2a and 8.2b we present the relative errors of the CVA FX spot deltas of the nine FX digital options, considering the same four bump sizes for both methods. Interestingly, we observe a dominance for $h = 10^{-2}$ or $h = 10^{-3}$ in terms of relative error. Typically, smaller values of h are expected to yield lower bias, and therefore a lower relative error might be expected. However, the high relative errors for small values of h can be explained by examining the payoff of FX digital call options. The payoff at maturity of the FX digital call option equals zero when the FX forward rate does not exceed the strike price. In the scenario where one of the bumped MC paths results in zero value at maturity, there is a high probability that the other bumped path will also yield zero value at maturity, especially with a very small bump size. This results in many simulations giving zero value. However, for the few cases where one of the bumped paths remains below the strike price

K, and the other bumped path exceeds it, a non-zero payoff value is observed. Dividing this value by a small value of h in the B&R estimator will result in high errors. This is exactly the behavior observed for the choices of $h = 10^{-4}$ and $h = 10^{-5}$.

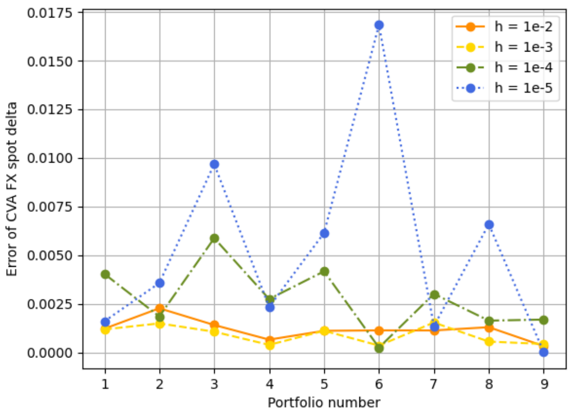


(a) Relative standard deviation computed by B&R.

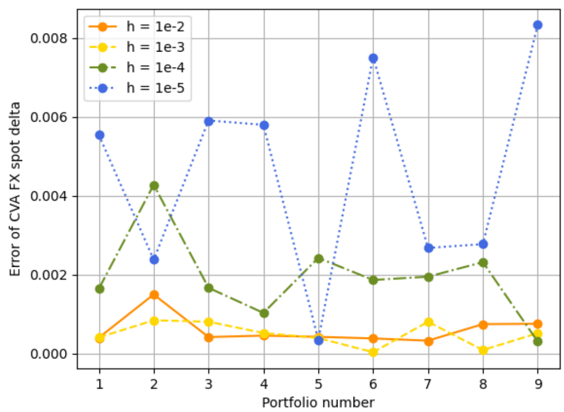


(b) Relative standard deviation computed by B&R + QMC.

Figure 8.1: Relative standard deviation of the CVA FX spot deltas of the FX digital options, computed by B&R and B&R + QMC for four different bump sizes.



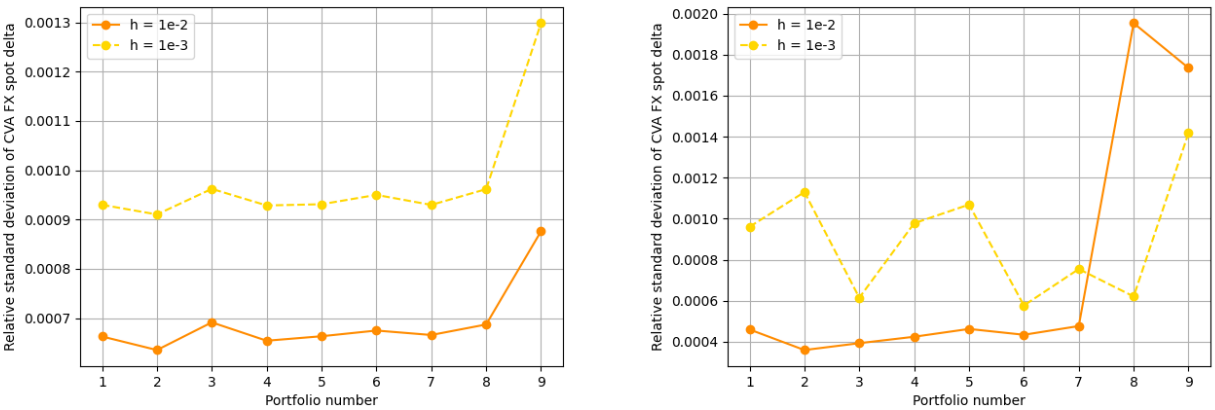
(a) Relative error computed by B&R.



(b) Relative error computed by B&R + QMC.

Figure 8.2: Relative error of the CVA FX spot deltas of the FX digital options, computed by B&R and B&R + QMC for four different bump sizes.

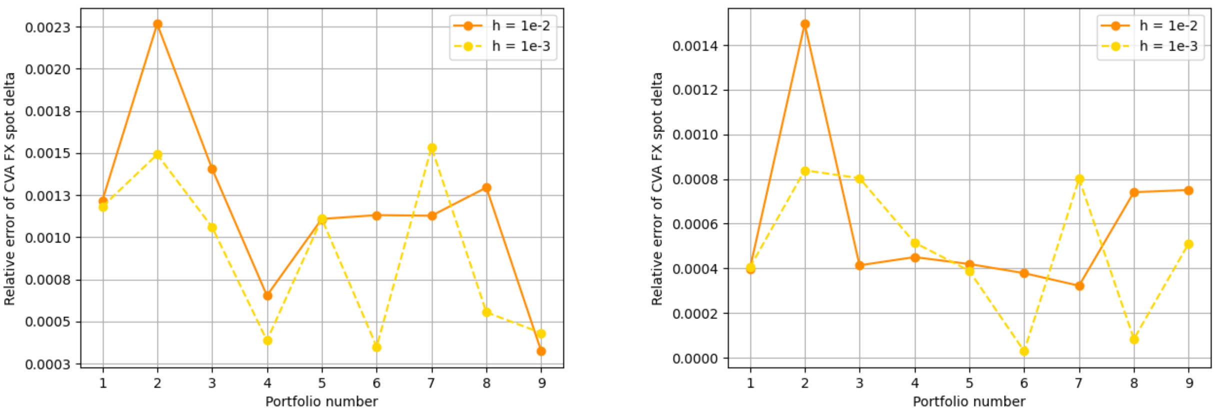
Based on the high relative errors and standard deviations, we conclude that $h = 10^{-4}$ and $h = 10^{-5}$ are not suitable choices as bump sizes. We proceed to compare the performance of $h = 10^{-2}$ and $h = 10^{-3}$. In Figure 8.3a and 8.3b we display the relative standard deviations of the FX spot CVA deltas of the nine FX digital options, only for $h = 10^{-2}$ and $h = 10^{-3}$. Similarly, we present the relative errors in Figure 8.4a and 8.4b. We see that for the B&R method, $h = 10^{-2}$ clearly performs better than $h = 10^{-3}$ in terms of standard deviation. Similarly, for the B&R + QMC method, $h = 10^{-2}$ demonstrates slightly better performance, with the exception of 2 portfolios. In terms of errors, $h = 10^{-2}$ and $h = 10^{-3}$ yield comparable results, both for B&R as for B&R + QMC. Based on these observations, we conclude that $h = 10^{-2}$ appears to be the most optimal bump size, and hence will be used whenever we compute CVA FX spot deltas of FX digital options using the B&R method.



(a) Relative standard deviation computed by B&R.

(b) Relative standard deviation computed by B&R + QMC.

Figure 8.3: Relative standard deviation of the CVA FX spot deltas of the FX digital options, computed by B&R and B&R + QMC for $h = 10^{-2}$ and $h = 10^{-3}$.



(a) Relative error computed by B&R.

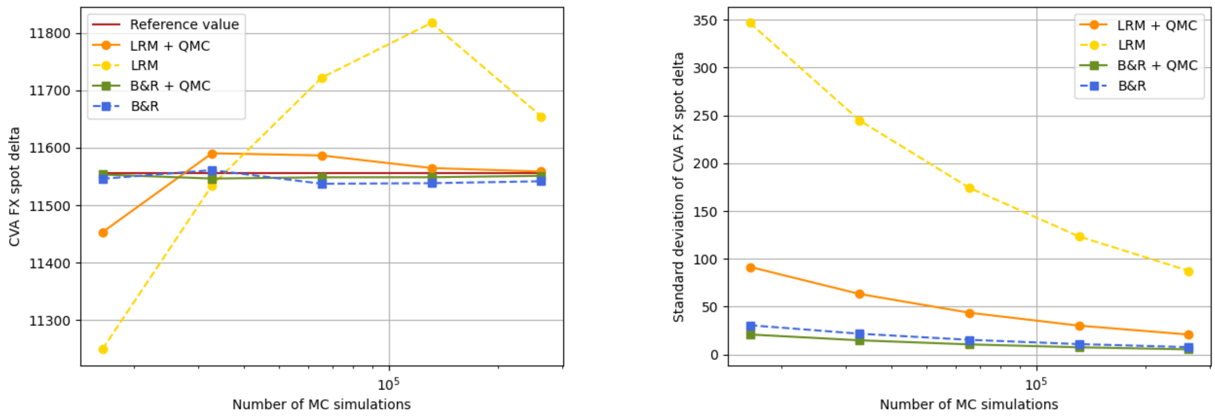
(b) Relative error computed by B&R + QMC.

Figure 8.4: Relative error of the CVA FX spot deltas of the FX digital options, computed by B&R and B&R + QMC for $h = 10^{-2}$ and $h = 10^{-3}$.

8.1.3. Convergence

In this section, we assess the convergence of the CVA FX spot deltas of FX digital options towards the reference value for all methods. We also clarify our choice of N for each method, based on CPU times.

Figure 8.5a presents the CVA FX spot delta estimates for the different methods using the set of parameters $\mathcal{P}_0$. Additionally, Figure 8.5b presents the standard deviations of the CVA FX spot deltas, where we varied the value of N from 2^{14} to 2^{18}. It is worth noting that LRM without the use of a QMC method is also included in these figures. It is obvious that all four methods are converging towards the reference value. For $N = 2^{18}$, LRM + QMC, B&R, and B&R + QMC yield accurate estimates for the CVA FX spot delta. However, LRM without a QMC method does not provide accurate results. While increasing the number of simulations could improve the accuracy, it will result in computational problems. Consequently, we choose to exclude LRM without a QMC method from our comparisons.

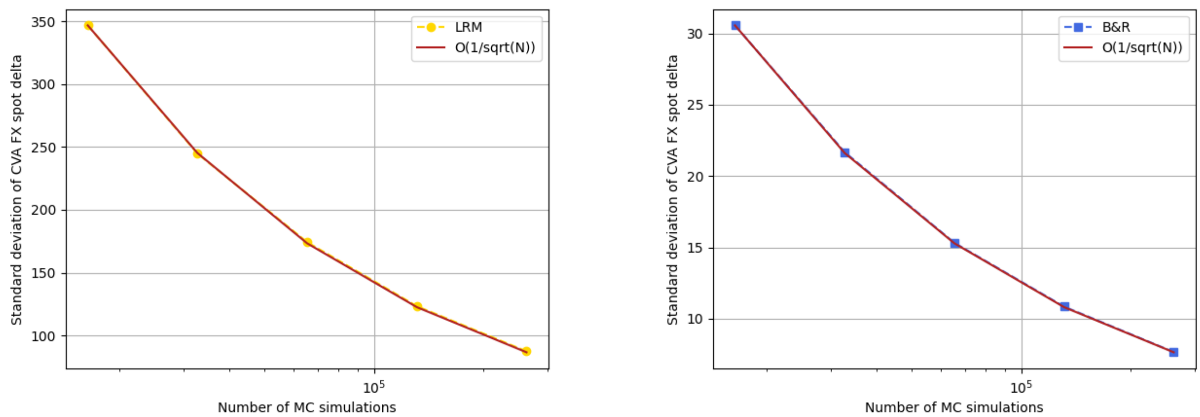


(a) CVA FX spot delta, compared with the reference value.

(b) Standard deviation of the CVA FX spot delta.

Figure 8.5: CVA FX spot delta, and its standard deviation of an FX digital option for different numbers of MC simulations.

Nevertheless, we can assess the convergence rate for the LRM estimates, and compare it with the theoretical rate of an ordinary MC method, which is of order $\mathcal{O}(1/\sqrt{N})$. A similar comparison can be made for the B&R method. Figures 8.6a and 8.6b depict this comparison between the convergence rate of LRM and the B&R method, along with a convergence rate of $\mathcal{O}(1/\sqrt{N})$. As expected, the ordinary MC methods align with the theoretical convergence rate. For the QMC methods, we can not perform a similar comparison, since the convergence rates are computed differently.



(a) Convergence rate of LRM.

(b) Convergence rate of the B&R method.

Figure 8.6: Convergence rates of LRM and the B&R method compared with the theoretical rates of $\mathcal{O}(1/\sqrt{N})$.

As previously observed in Figure 8.5a and 8.5b, LRM + QMC, B&R, and B&R + QMC all provide accurate estimates of the CVA FX spot delta. To obtain a fair comparison between the performances, we examine the CPU times of the different methods. In Table 8.1 we present the CPU times (in seconds) of the three methods across different numbers of MC simulations. It is clear that, in comparison to LRM + QMC, B&R is observed to be roughly 1.5 times slower, while B&R + QMC takes roughly double the amount of time. Additionally, we observe that the CPU times scale (almost) linearly with the number of MC simulations; doubling the number of simulations results in a CPU time twice as large.

Method / N	2^{14}	2^{15}	2^{16}	2^{17}	2^{18}
Time LRM + QMC	1.12	1.72	3.11	6.01	11.74
Time B&R	1.78	2.68	4.70	9.08	17.87
Time B&R + QMC	2.44	3.98	7.25	14.34	29.74

Table 8.1: CPU times (in seconds) for the three methods for different numbers of MC simulations.

For a fair comparison of the three methods in our analysis, we take different simulation numbers so that all methods have similar CPU times. For that reason, we fix the number of MC simulations for LRM + QMC at 2^{18} and adjust the number of simulations for the other methods accordingly. This is done by dividing the number of simulations for B&R by 1.5, and by dividing the number of simulations for B&R + QMC by 2. The overview of these simulation numbers is presented in Table 8.2. These numbers are used in the sensitivity analysis hereafter.

Method	N
LRM + QMC	2^{18}
B&R	$2^{18}/1.5$
B&R + QMC	2^{17}

Table 8.2: Choice of simulation numbers for the three methods.

8.1.4. Sensitivity Analysis

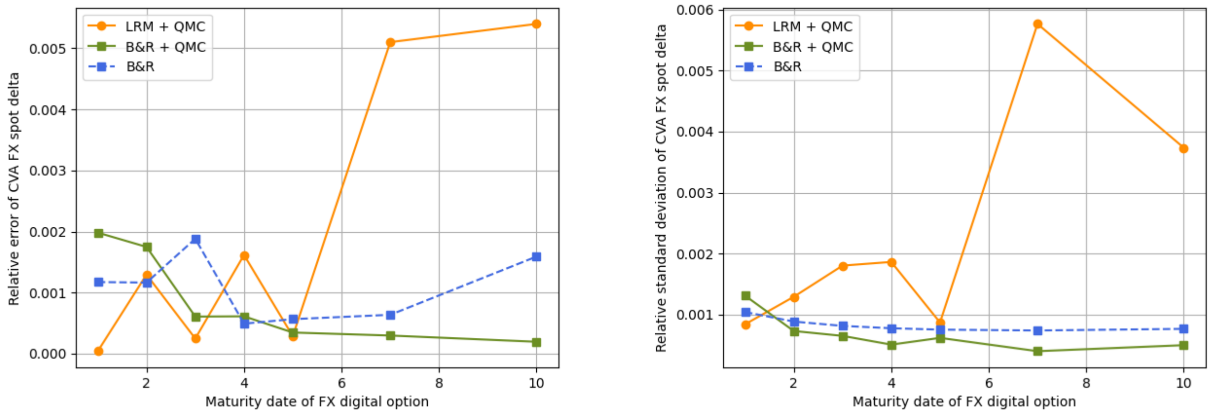
In this section, we will conduct a sensitivity analysis of LRM and the B&R method, for the computation of CVA FX spot deltas of FX digital options. We will consider several FX digital call options, all using the parameter set $\mathcal{P}_0$, but varying one of the parameters. For each parameter variation, we compute the reference value and obtain estimates of the CVA FX spot delta using the three methods. Subsequently, we compare the relative errors and standard deviations of the estimates of all the methods. To ensure a valid comparison, we show the CPU times of the methods and all the simulations in Appendix B.1. It is evident that all these CPU times are similar.

Maturity dates

We define a set of maturity dates, ranging from one year to 10 years, and we check the impact of these maturity dates on the errors and standard deviations of the CVA FX spot deltas. Figure 8.7a and 8.7b present the impact of varying the maturity dates on the relative errors and standard deviations of the CVA FX spot deltas of FX digital options. We observe from the figures that for short maturities (up to 5 years), LRM + QMC is able to match the B&R and B&R + QMC values quite well. However, as the maturity increases from 5 years up to 7 and 10 years, the performance of LRM diminishes due to high standard deviations of the LRM estimator. This aligns with the expected increase in the variance of the LRM estimator as the maturity date increases, as detailed in Section 6.3.

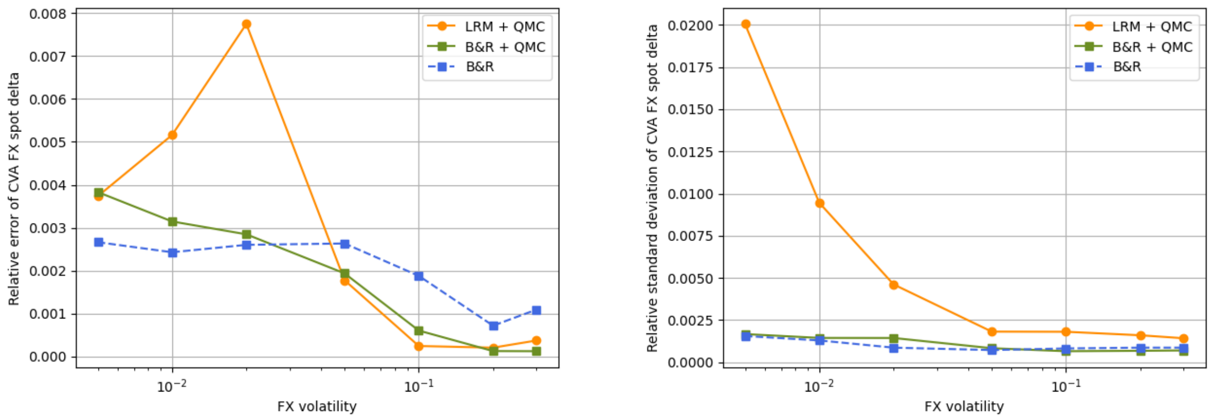
Volatility of the FX process

We define a set of FX volatilities ranging from 0.005 to 0.3 and we check the impact of these volatilities on the relative errors and standard deviations of the CVA FX spot deltas. Figure 8.8a and 8.8b show these impacts on the CVA FX spot deltas of the FX digital options, on a logarithm scale. Evidently, for volatilities exceeding 5%, LRM + QMC closely matches the B&R and B&R + QMC values. However, as the FX volatility decreases towards zero, LRM + QMC exhibits poor performance due to high standard deviations, which is in line with the expectation. As we outlined in Section 6.3, the variance of the score function diverges for low FX volatilities, thereby contributing to the observed high variance of LRM.



(a) Impact of T on the relative errors. **(b)** Impact of T on the relative standard deviations.

Figure 8.7: Impact of different maturity dates, T, on the relative errors and standard deviations of the CVA FX spot deltas of FX digital options.

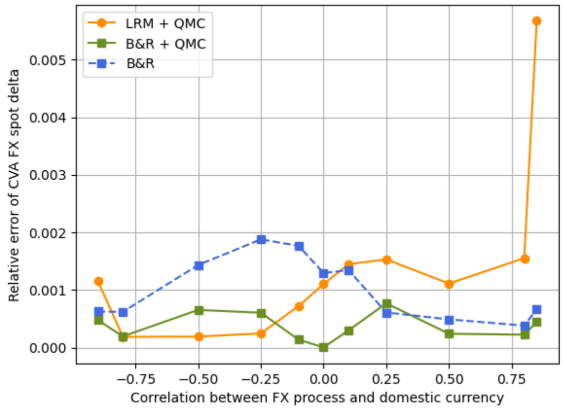


(a) Impact of σ_y on the relative errors. **(b)** Impact of σ_y on the relative standard deviations.

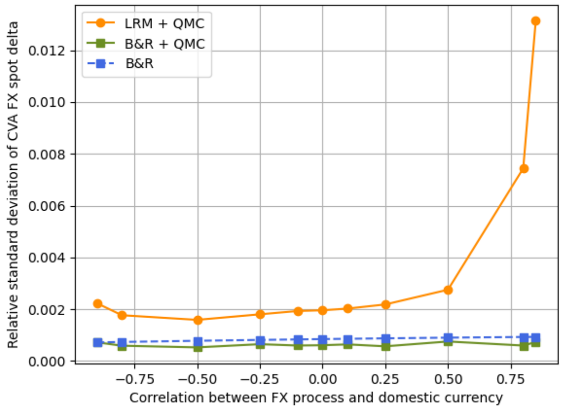
Figure 8.8: Impact of different FX volatility values, σ_y, on the relative errors and standard deviations of the CVA FX spot deltas of FX digital options.

Correlations between FX process, domestic currency, and foreign currency

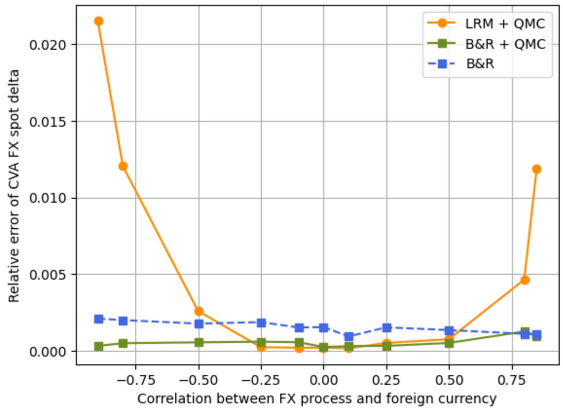
We define a set of correlations ranging from -1 to 1 and check the impact on the relative errors and standard deviations of the CVA FX spot deltas for these different values of correlations $\rho_{y,d}, \rho_{y,f}$ and $\rho_{d,f}$. Due to the covariance structures, choices of $\rho_{y,d} > 0.85, \rho_{y,f} > 0.85$ and $|\rho_{d,f}| > 0.85$ lead to a covariance matrix that is not semi-symmetric positive definite (SPD). Hence, we exclude these cases from our comparisons. In Figure 8.9a, 8.9c and 8.9e we present the sensitivities of the relative error of the CVA FX spot deltas for different values of $\rho_{y,d}, \rho_{y,f}$ and $\rho_{d,f}$, respectively. Similarly, in Figure 8.9b, 8.9d, and 8.9f we showcase the sensitivities of relative standard deviations.



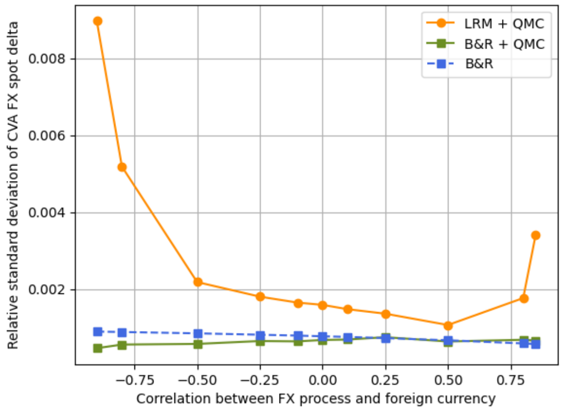
(a) Impact of $\rho_{y,d}$ on the relative errors.



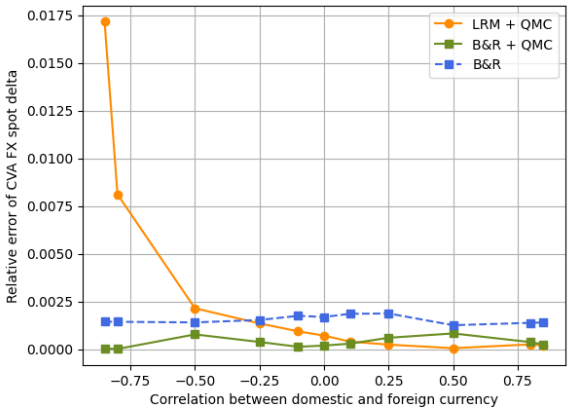
(b) Impact of $\rho_{y,d}$ on the relative standard deviations.



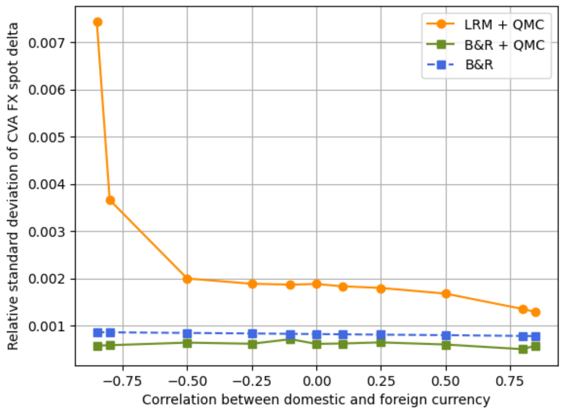
(c) Impact of $\rho_{y,f}$ on the relative errors.



(d) Impact of $\rho_{y,f}$ on the relative standard deviations.



(e) Impact of $\rho_{d,f}$ on the relative errors.



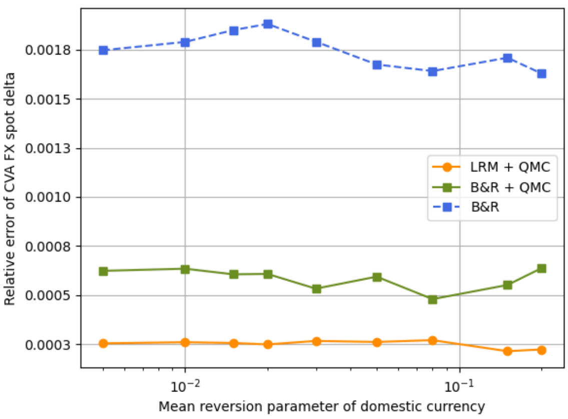
(f) Impact of $\rho_{d,f}$ on the relative standard deviations.

Figure 8.9: Impact of different $\rho_{y,d}$, $\rho_{y,f}$ and $\rho_{d,f}$ values on the relative errors and standard deviations of the CVA FX spot deltas of FX digital options.

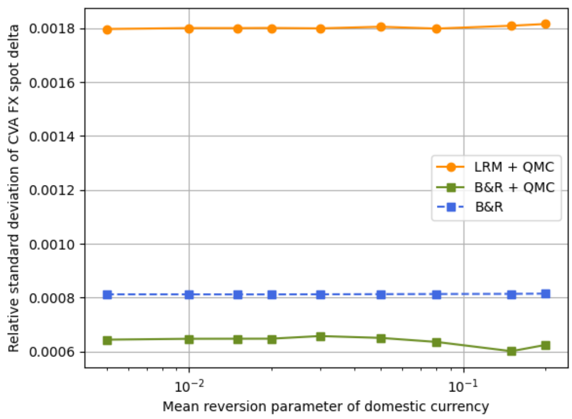
The figures clearly illustrate that for correlations around zero, LRM is able to fairly reproduce the B&R and B&R + QMC values. However, for higher correlation values, the variance of the LRM estimator increases. In particular, this variance increase is noticeable for high (negative/positive) $\rho_{y,d}$-values, for high (negative/positive) $\rho_{y,f}$-values, and for high negative $\rho_{d,f}$-values. The cases where $\rho_{y,d}$ and $\rho_{y,f}$ are equal to -1 have not been presented in the figures due to the potential distortion of the y-axis caused by an even worse performance of the LRM estimator. The poor performance of LRM in the presence of high correlations is again in line with our expectations. As explained in Section 6.3, high correlations lead to an explosion of the variance of the score function, due to problems with the inversion of the covariance matrix. This results in the observed high variance of the LRM estimator in the figures.

Mean reversion parameters of domestic currency and foreign currency

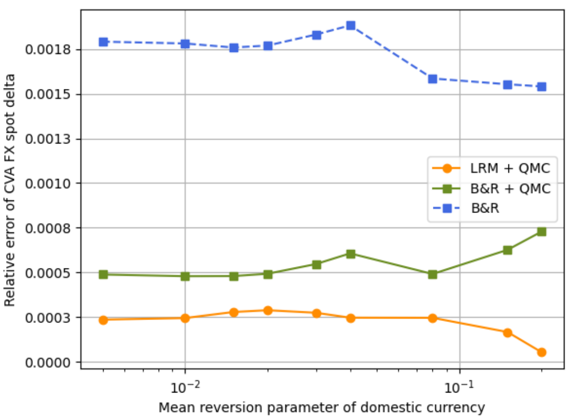
We define a set of mean reversion parameters ranging from 0.005 to 0.2 and we check the impact on the relative errors and standard deviations of the CVA FX spot deltas of these mean reversion parameters for the domestic currency, a_d, and the foreign currency, a_f. In Figure 8.10a and 8.10c we present, on a logarithmic scale, the impact on the relative error of the CVA FX spot deltas. Similarly, Figure 8.10b and 8.10d illustrate the impact on the relative standard deviations. It is clear that LRM performs well for all values of the mean reversion parameters.



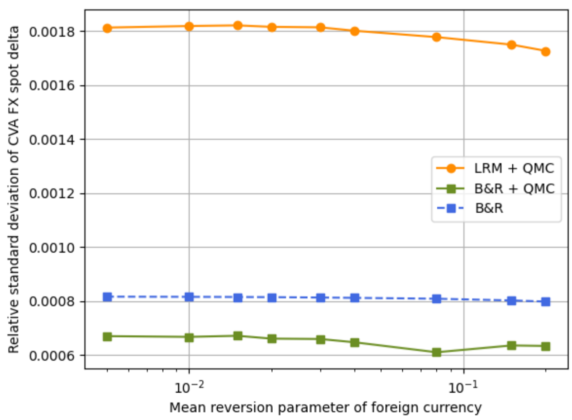
(a) Impact of a_d on the relative errors.



(b) Impact of a_d on the relative standard deviations.



(c) Impact of a_f on the relative errors.

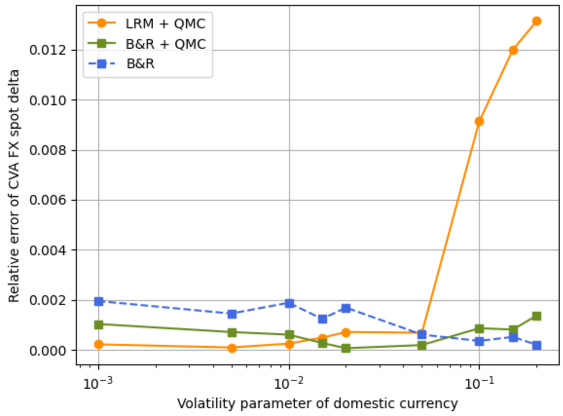


(d) Impact of a_f on the relative standard deviations.

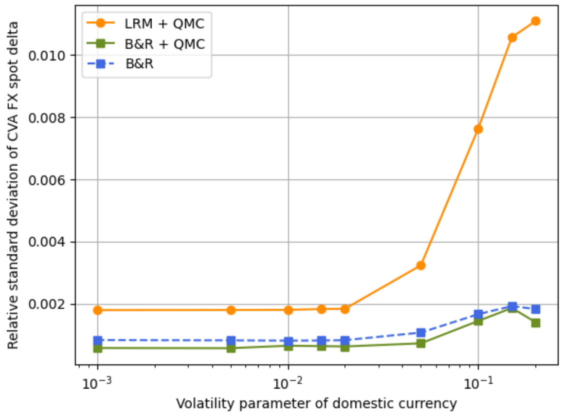
Figure 8.10: Impact of different a_d and a_f values on the relative errors and standard deviations of the CVA FX spot deltas of FX digital options.

Volatility of domestic currency and foreign currency

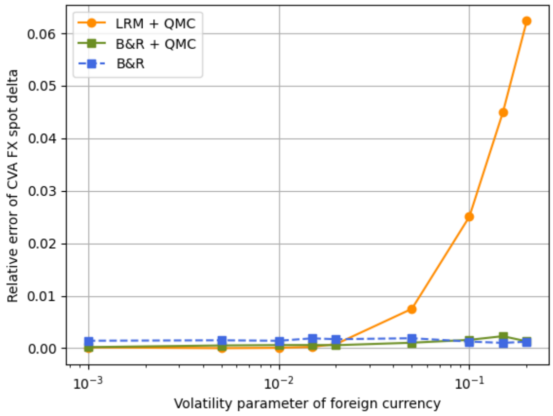
We define a set of volatility parameters ranging from 0.001 to 0.2 and we check the impact of varying the volatilities of the domestic currency, σ_d, and the foreign currency, σ_f, on the relative errors and standard deviations of the CVA FX spot deltas. Figure 8.11a and 8.11c present this impact on the relative errors of the CVA FX spot deltas, on a logarithmic scale. Similarly, the impact on the relative standard deviations is shown in Figure 8.11b and 8.11d. Notably, LRM demonstrates a close alignment with the B&R methods up to volatilities of 2%. However, in line with our expectations in Section 6.3, LRM exhibits a high variance for large values of volatility. This is because the high volatility values lead to a MC average with high variance due to summing large residues resulting from cancellation effects.



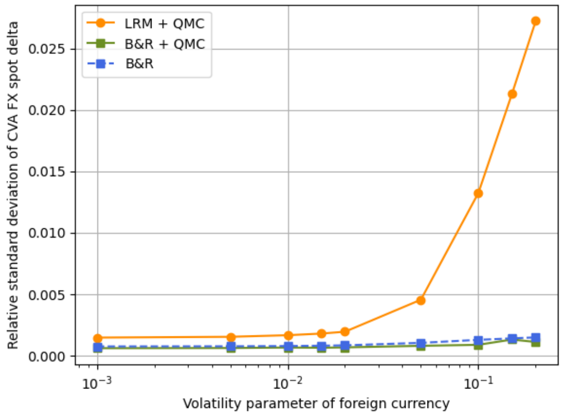
(a) Impact of σ_d on the relative errors.



(b) Impact of σ_d on the relative standard deviations.



(c) Impact of σ_f on the relative errors.



(d) Impact of σ_f on the relative standard deviations.

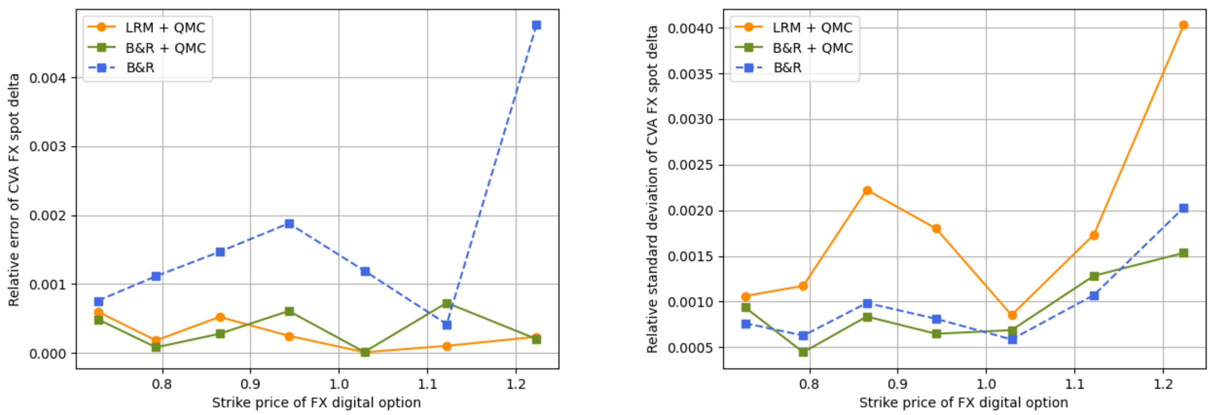
Figure 8.11: Impact of different σ_d and σ_f values on the relative errors and standard deviations of the CVA FX spot deltas of FX digital options.

Strike price of the FX digital option

We define the set of strike prices as

$$\left[y(t_0)\,\text{exp}\left(0.1c\sqrt{T}\right), c \in \{-1.5, -1, -0.5, 0, 0.5, 1, 1.5\}\right],$$

where T is the time-to-maturity of the digital option, set at 3 years. In this way, we create multiple FX digital options, some out-of-the-money, one at-the-money, and some in-the-money. Figure 8.12a and 8.12b illustrate the impact of these different strike prices on the relative errors and standard deviations of the CVA FX spot deltas. Across all different strike prices, LRM is able to fairly reproduce the B&R and B&R + QMC values. We do observe, however, a small increase in the variance once the digital option moves further out of the money. Importantly, this variance increase is not attributed to the characteristics of LRM, as the strike price K does not influence LRM's score function. Instead, the increase in the variance is linked to the payoff of the FX digital option. Specifically, the payoff of an FX digital option at maturity equals N_f in case $y_F(T,T) > K$, and zero otherwise. As the strike price K increases, the chance of $y_F(T,T)$ exceeding K diminishes, resulting in many MC paths assigning zero value to the payoff $V(T)$. Nevertheless, some paths will still exceed K, and therefore, they assign the value N_f to the payoff $V(T)$. This results in an increase in the variance of the payoff of the FX digital option. Therefore, the variance of both the LRM estimator and the B&R estimator will also increase, which is exactly what we observe in the figure.



(a) Impact of K on the relative errors.

(b) Impact of K on the relative standard deviations.

Figure 8.12: Impact of different strike prices of the FX digital options on the relative errors and standard deviations of the CVA FX spot deltas of FX digital options.

From this sensitivity analysis, we can deduce that for a certain range of parameters, LRM + QMC is proficient in reproducing the B&R and B&R + QMC values. However, for long maturities, low FX volatilities, relatively large correlation values, and high model volatilities, the well-known high-variance problems associated with LRM become evident. Additionally, for deep out-of-the-money FX digital options, an increase in the variance of the LRM estimator is observed. However, since the relative errors remain very low, and the rise in variance is not attributed to LRM itself, we can alleviate the concerns related to high strike prices.

Metric / Method	LRM + QMC	B&R	B&R + QMC
Relative error	$44/93$	$14/93$	$35/93$
Relative std. dev.	$1/93$	$12/93$	$80/93$

Table 8.3: Frequency of outperformance in relative error and standard deviation for all three methods.

To draw a comprehensive conclusion about the overall performance of LRM compared to the B&R methods, we compare the relative errors and standard deviations of all the 93 CVA FX spot deltas that we have computed in this section. In Table 8.3 we show for each method the number of times it outperformed the other two methods in terms of relative error and standard deviation in the computation of the deltas. The results indicate that LRM + QMC excels in terms of relative error, meaning the mean values of the LRM + QMC estimates are often very close to the reference values. However, in terms of relative standard deviations, LRM falls behind, suggesting that the LRM + QMC estimates are quite volatile. We can, however, relax this observation. Since LRM is used together with a QMC method, the observed standard deviations are empirically estimated using the power law, see Section 7.3 for more details. This implies that the observed standard deviations are not 'true' standard deviations, but rather empirical ones. Therefore, it remains possible for the true standard deviations to be slightly smaller. Consequently, we find it more meaningful to focus on the relative errors in this comparison.

By solely examining the errors, we conclude that for a certain range of parameters, LRM performs well and is able to match or outperform the B&R method. However, for long maturity dates, low FX volatilities, highly correlated risk factors, and large values of model volatilities, the variance of LRM is too high, leading to a better performance of the B&R method.

8.2. CVA FX spot deltas of FX swaps

In this section, we present the performance of LRM in the computation of CVA FX spot deltas of FX swaps, and we compare this performance against the B&R method. Similar to the approach taken for FX digital options, we establish a reference value by summing the analytical derivatives of the expected exposures at all exposure times, as in Theorem 3.1.1, and then multiplying the sum with the LGD and PD. However, in our implementation, the CVA FX spot deltas computed by LRM and the B&R method did not align with the reference values. Consequently, we will proceed in a slightly different way than we did for the CVA deltas of the FX digital options. Instead of using the real reference value, we compute the B&R CVA delta using a large number of MC simulations and treat this value as the reference value of the CVA delta of an FX swap. Then we compare the performance of LRM and the B&R method to this reference value.

To implement the B&R method, a bump size must be selected again. Since the payoff of an FX swap is continuous (see Section 2.7.2), no trade-off is needed between the bias and the variance. Therefore, we can choose h as small as possible, as long as we take care of the machine precision. We choose $h = 10^{-5}$.

In Section 8.2.1, we will present the set of parameters that we will use in the comparisons. Then in Section 8.2.2, we will check the MC errors of our implementations and compare them with the theoretical MC errors. Lastly, in Section 8.2.3, we compare the performance of LRM and the B&R method in computing the CVA FX spot deltas of FX swaps.

8.2.1. Set of parameters

In this section, we present the set of parameters that we will use in our analysis. We make use of the exact same set of parameters as we did for the FX digital options. Therefore, our set $\mathcal{P}_0$, containing all the parameter values, is given as follows:

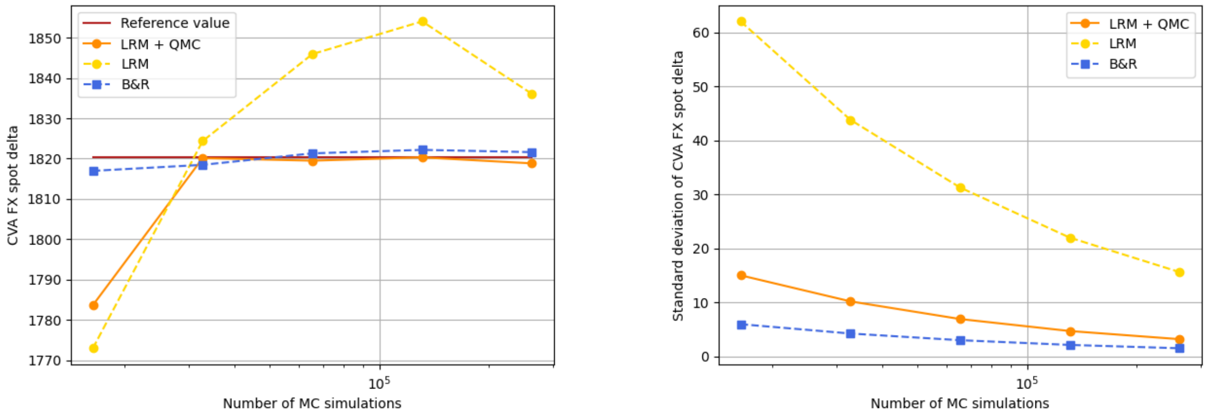
$$\mathcal{P}_0 = \left\{ \begin{array}{lcrlcr} T & = & 3, & m & = & 30, \\ t_0 & = & 0, & y(t_0) & = & 0.9433, \\ r_d(t_0) & = & 0.03904, & r_f(t_0) & = & 0.0538, \\ \sigma_y & = & 0.1, & \rho_{y,d} & = & -0.25, \\ \rho_{y,f} & = & -0.25, & \rho_{d,f} & = & 0.25, \\ a_d & = & 0.02, & \sigma_d & = & 0.01, \\ a_f & = & 0.04, & \sigma_f & = & 0.015, \\ N_f & = & 10^6, & K & = & 0.9433, \\ R_c & = & 0, & \xi & = & 0.002. \end{array} \right\}. \tag{8.2}$$

In the upcoming sections, we consistently employ the set $\mathcal{P}_0$ as the set of parameters for our comparisons, unless stated otherwise. Furthermore, for the LRM simulations, we will always make use of $N = 2^{18}$ MC simulations. For a fair comparison between LRM and the B&R method, we will adjust the number of B&R simulations so that the CPU times of both methods are approximately equal.

8.2.2. Convergence

In this section, we assess the convergence of the CVA FX spot deltas of FX swaps towards the reference value for all methods. We also clarify our choice of N for each method, based on CPU times.

In Figure 8.13a, the estimates of the CVA FX spot delta for the different methods are shown. Similarly, in Figure 8.13b we present the standard deviation of the CVA FX spot deltas, where we varied the value of N from 2^{14} to 2^{18}. It is worth noting that LRM without the use of a QMC method is also included in these figures.

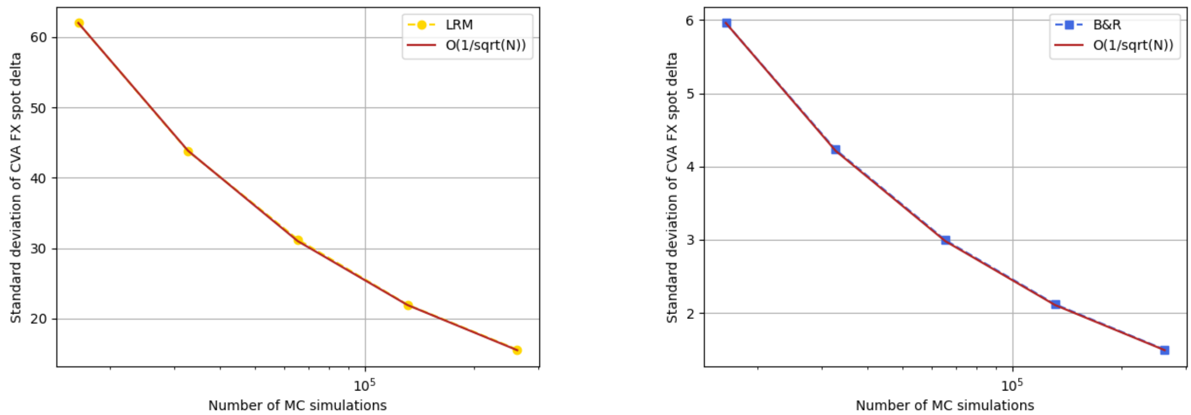


(a) CVA FX spot delta, compared with the reference value.

(b) Standard deviation of the CVA FX spot delta.

Figure 8.13: CVA FX spot delta, and its standard deviation of an FX Swap for different numbers of MC simulations.

It is clear that all three methods are converging towards the reference value. For $N = 2^{18}$, LRM + QMC and B&R provide accurate estimates for the CVA FX spot delta. However, similar to what we observed for the FX digital options, LRM without a QMC method fails to provide accurate results. Again, increasing the number of simulations could enhance the result, but this approach poses computational problems. Therefore, we opt to exclude LRM without a QMC method from our comparisons.



(a) Convergence rate of LRM.

(b) Convergence rate of the B&R method.

Figure 8.14: Convergence rates of LRM and the B&R method compared with the theoretical rates of $\mathcal{O}(1/\sqrt{N})$.

Nevertheless, we can assess the convergence rate for the LRM estimates, and compare it with the theoretical rate of an ordinary MC method, which is of order $\mathcal{O}(1/\sqrt{N})$. A similar comparison can be made for the B&R method. In Figure 8.14a and 8.14b we present this comparison between the convergence rate of LRM and the B&R method, along with a convergence rate of $\mathcal{O}(1/\sqrt{N})$. Aligning with our expectations, we see that the ordinary MC methods match the theoretical convergence rate. For the QMC methods, we can not perform a similar comparison, since the convergence rates are computed differently.

As previously observed in Figure 8.13a and 8.13b, LRM + QMC and B&R both provide accurate estimates of the CVA FX spot delta. To obtain a fair comparison between the performances, we examine the CPU times of the different methods. In Table 8.4 we present the CPU times (in seconds) of the two methods across different numbers of MC simulations.

Method / N	2^{14}	2^{15}	2^{16}	2^{17}	2^{18}
Time LRM + QMC	1.16	1.82	3.20	6.53	12.63
Time B&R	1.84	2.71	5.02	10.02	18.43

Table 8.4: CPU times (in seconds) for both methods for different numbers of MC simulations.

We observe a similar pattern as for the FX digital options. In comparison to LRM + QMC, B&R is observed to be roughly 1.5 times slower. Furthermore, we notice again that the CPU times scale (almost) linearly with the number of MC simulations; doubling the number of simulations results in a CPU time twice as large.

Method	N
LRM + QMC	2^{18}
B&R	$2^{18}/1.5$

Table 8.5: Choice of simulation numbers for both methods.

For a fair comparison of the two methods, we take different simulation numbers so that both methods have similar CPU times. For that reason, we will fix the number of MC simulations for LRM + QMC at 2^{18}, and adjust the number of simulations for the B&R method accordingly. We achieve this by dividing the number of simulations for the B&R method by 1.5. The overview of these numbers is presented in Table 8.5. These numbers are used in the sensitivity analysis hereafter.

8.2.3. Sensitivity Analysis

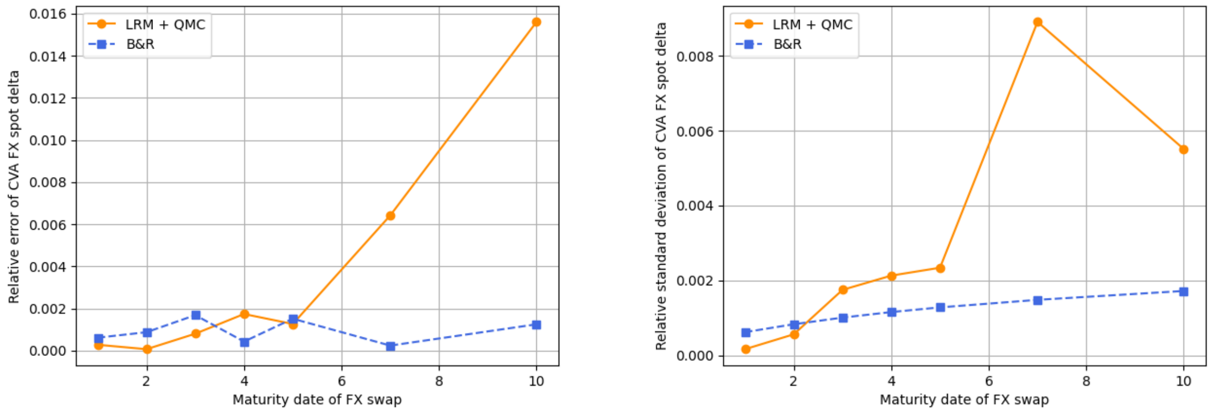
In this section, we will conduct a sensitivity analysis of LRM and the B&R method, for the computation of CVA FX spot deltas of FX swaps. We will consider several FX swaps, all using the parameter set $\mathcal{P}_0$, but varying one of the parameters. For each variation of parameters, we compute the reference value and obtain estimates of the CVA FX spot delta using both LRM and the B&R method. Then, we compare the relative errors and standard deviations of the estimates of both methods. To ensure a valid comparison, we show the CPU times of both methods and all the simulations in Appendix B.2. It is clear that all these CPU times are similar.

Maturity dates

We define a set of maturity dates, ranging from one year to 10 years, and we check the impact of these maturity dates on the errors and standard deviations of the CVA FX spot deltas. Figure 8.15a and 8.15b present the impact of varying the maturity dates on the relative errors and standard deviations of the CVA FX spot deltas of FX swaps. We observe exactly the same results as for the FX digital options: for short maturities (up to 5 years), LRM + QMC is able to fairly match the B&R values, but as the maturity increases from 5 years up to 7 and 10 years, we observe higher standard deviations, leading to a decline in the performance of LRM. As we have discussed in Section 6.3, the variance increase for longer maturities is consistent with the expectations.

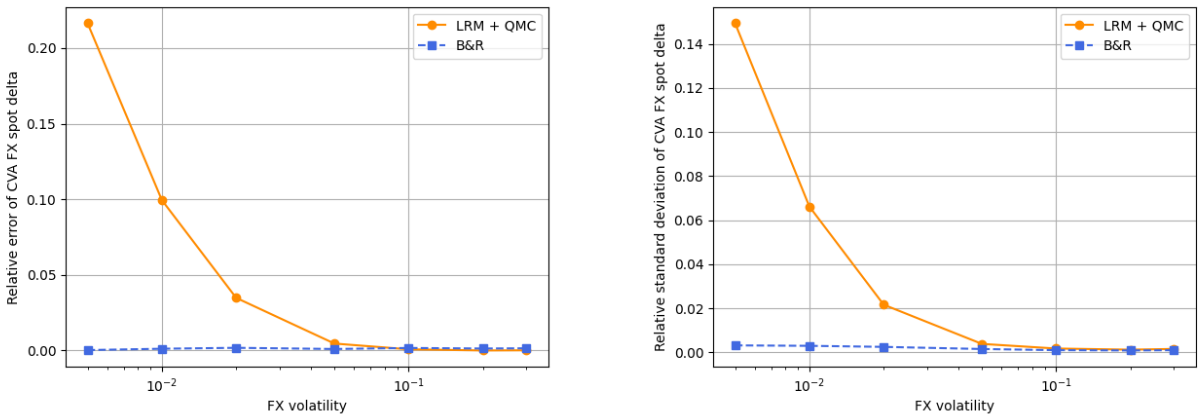
Volatility of the FX process

We define a set of FX volatilities ranging from 0.005 to 0.3 and we examine their impact on the relative errors and standard deviations of the CVA FX spot deltas. Figure 8.16a and 8.16b show these impacts on the CVA FX spot deltas of the FX swaps, on a logarithm scale. Similar to the observations for the FX digital options, for volatilities exceeding 5%, the LRM values are closely aligned with the B&R values. However, as the FX volatility decreases towards zero, LRM + QMC is not able to match the B&R values anymore. This is in line with our expectations. As we explained in Section 6.3, the variance of the score function diverges for low FX volatilities, resulting in the high variance that we observe for LRM.



(a) Impact of T on the relative errors. **(b)** Impact of T on the relative standard deviations.

Figure 8.15: Impact of different maturity dates, T, on the relative errors and standard deviations of the CVA FX spot deltas of FX swaps.

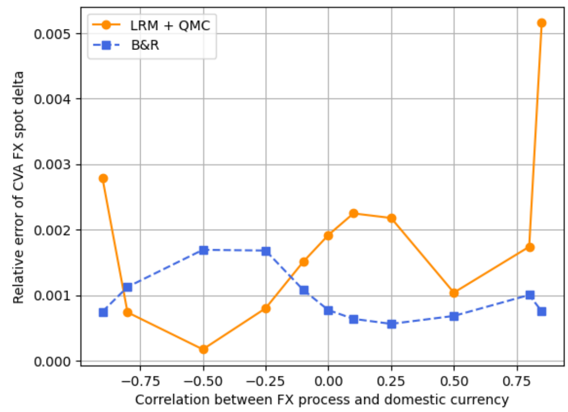


(a) Impact of σ_y on the relative errors. **(b)** Impact of σ_y on the relative standard deviations.

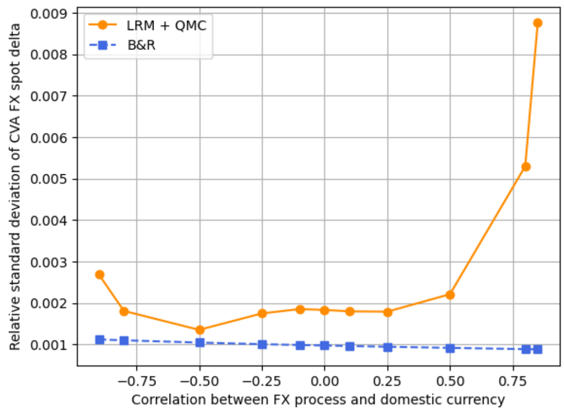
Figure 8.16: Impact of different FX volatility values, σ_y, values on the relative errors and standard deviations of the CVA FX spot deltas of FX swaps.

Correlations between FX process, domestic currency, and foreign currency

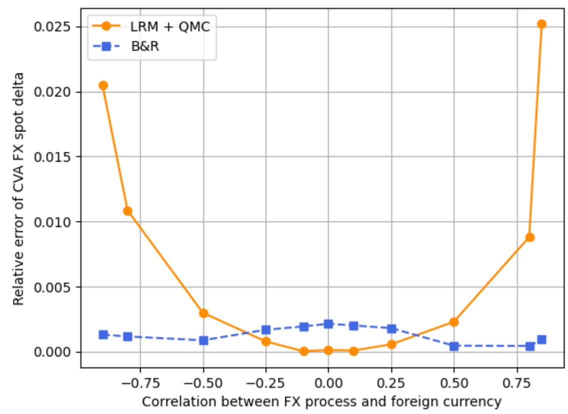
We define a set of correlations ranging from −1 to 1 and check the impact on the relative errors and standard deviations of the CVA FX spot deltas for these different values of correlations $\rho_{y,d}, \rho_{y,f}$ and $\rho_{d,f}$. Similar to what we observed for the FX digital options, choices of $\rho_{y,d} > 0.85, \rho_{y,f} > 0.85$ and $|\rho_{d,f}| > 0.85$ lead to a covariance matrix that is not semi-symmetric positive definite (SPD). Hence we exclude these cases from our comparisons. In Figure 8.17a, 8.17c and 8.17e we present the sensitivities of the relative error of the CVA FX spot deltas of FX with respect to different values of $\rho_{y,d}, \rho_{y,f}$ and $\rho_{d,f}$, respectively. Similarly, in Figure 8.17b, 8.17d and 8.17f we present the sensitivities of the relative standard deviations.



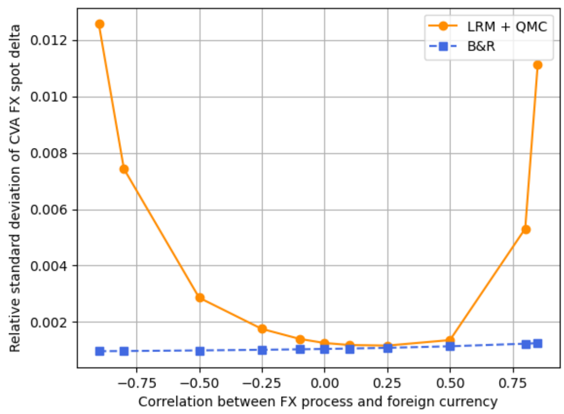
(a) Impact of $\rho_{y,d}$ on the relative errors.



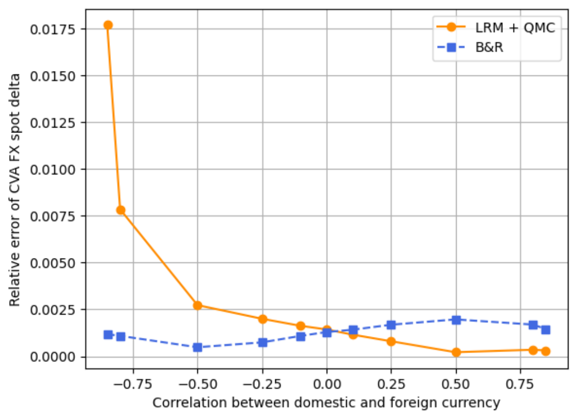
(b) Impact of $\rho_{y,d}$ on the relative standard deviations.



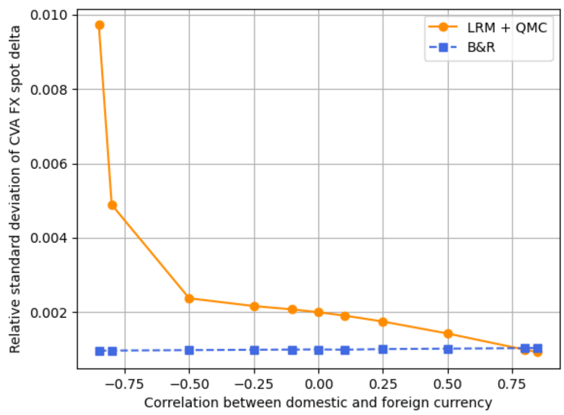
(c) Impact of $\rho_{y,f}$ on the relative errors.



(d) Impact of $\rho_{y,f}$ on the relative standard deviations.



(e) Impact of $\rho_{d,f}$ on the relative errors.



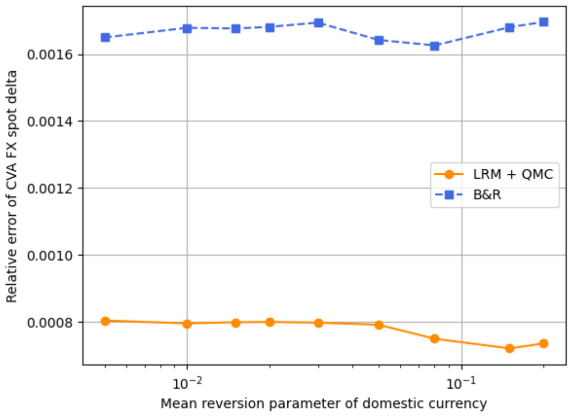
(f) Impact of $\rho_{d,f}$ on the relative standard deviations.

Figure 8.17: Impact of different $\rho_{y,d}$, $\rho_{y,f}$ and $\rho_{d,f}$ values on the relative errors and standard deviations of the CVA FX spot deltas of FX swaps.

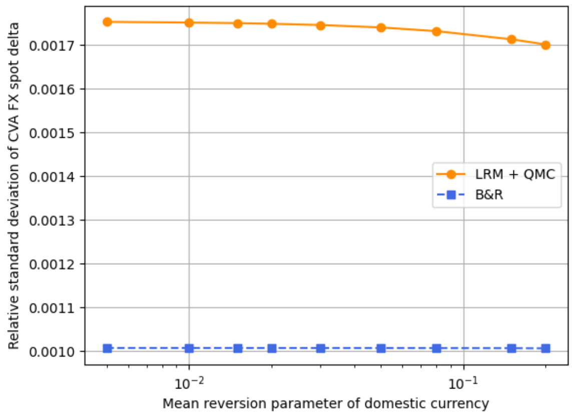
For correlations around zero, it is clear that LRM is able to reproduce the B&R values quite well. However, for higher correlation values, the variance of the LRM estimator increases. This is particularly observed for high (negative/positive) $\rho_{y,d}$-values, for high (negative/positive) $\rho_{y,f}$-values, and for high negative $\rho_{d,f}$-values. The cases where $\rho_{y,d}$ and $\rho_{y,f}$ are equal to -1 have not been presented in the figures due to their poor performance, which would distort the y-axis in the figures. The poor performance of LRM in the presence of high correlations aligns again with our expectations, outlined in Section 6.3, where we discussed that the highly correlated risk factors result in an explosion of the variance of the score function, due to problems with the inversion of the covariance matrix.

Mean reversion parameters of domestic currency and foreign currency

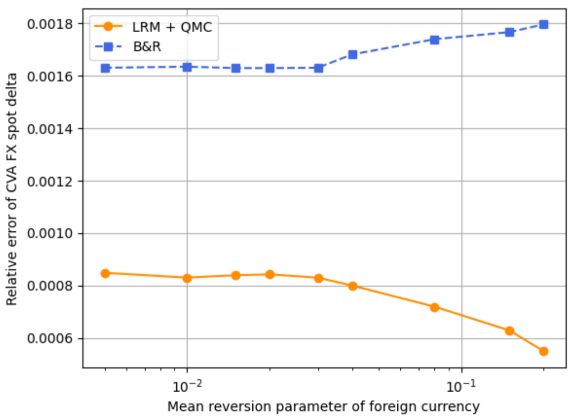
We define a set of mean reversion parameters ranging from 0.005 to 0.2 and we check the impact of varying the mean reversion parameters of the domestic currency, a_d, and the foreign currency, a_f, on the relative errors and standard deviations of the CVA FX spot deltas. Figure 8.18a and 8.18c presents the impact on the relative errors of the CVA FX spot deltas, on a logarithmic scale, for several values of a_d and a_f, respectively. Similarly, in Figure 8.18b and 8.18d the impact on the relative standard deviations is shown. Evidently, for all values of the mean reversion parameters, LRM performs well in the computation of the CVA FX spot deltas.



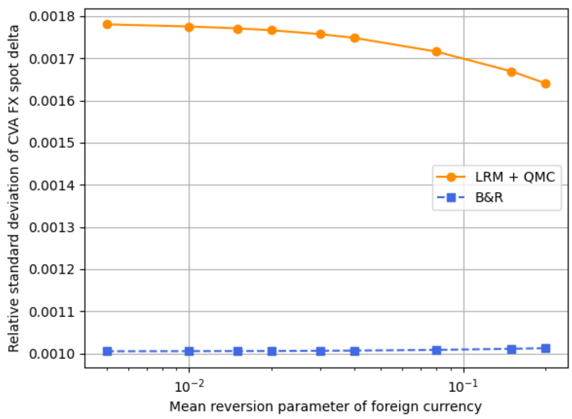
(a) Impact of a_d on the relative errors.



(b) Impact of a_d on the relative standard deviations.



(c) Impact of a_f on the relative errors.

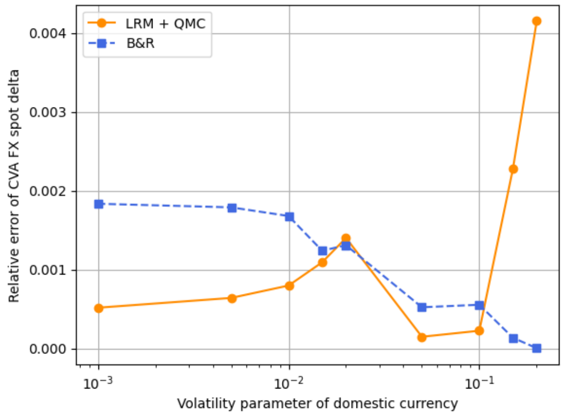


(d) Impact of a_f on the relative standard deviations.

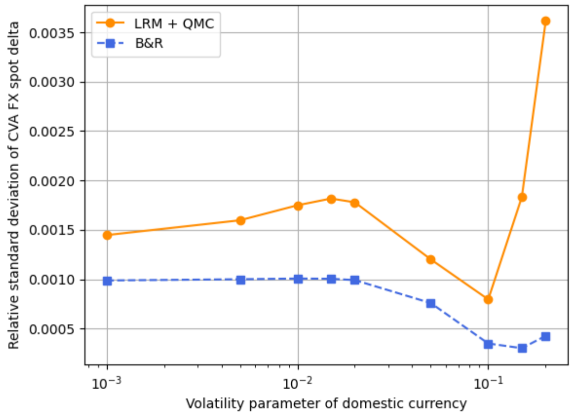
Figure 8.18: Impact of different a_d and a_f values on the relative errors and standard deviations of the CVA FX spot deltas of FX swaps.

Volatility of domestic currency and foreign currency

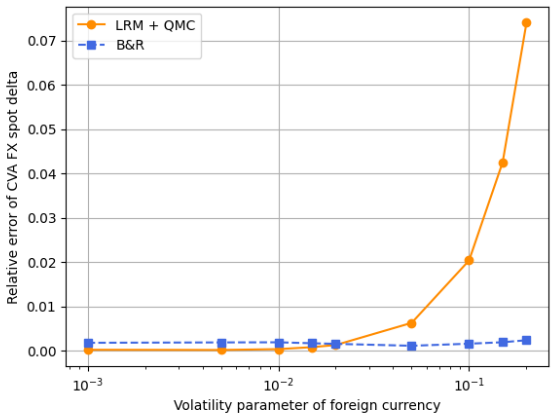
We define a set of volatility parameters ranging from 0.001 to 0.2 and we check the impact on the relative errors and standard deviations of the CVA FX spot deltas for these volatilities of the domestic currency, σ_d, and the foreign currency, σ_f. In Figure 8.19a and 8.19c we present the sensitivities of the relative errors with respect to σ_d and σ_f, on a logarithmic scale. Similarly, in Figure 8.19b and 8.19d we provide the sensitivities of the relative standard deviations with respect to the volatilities. For volatilities up to 2%, it is clear that LRM is able to reproduce the B&R values quite well. However, for higher volatility values, we observe an increase in the variance of the LRM estimators. This is in line with our expectations. As discussed in Section 6.3, the high volatility values lead to a MC average with high variance due to summing large residues resulting from cancellation effects.



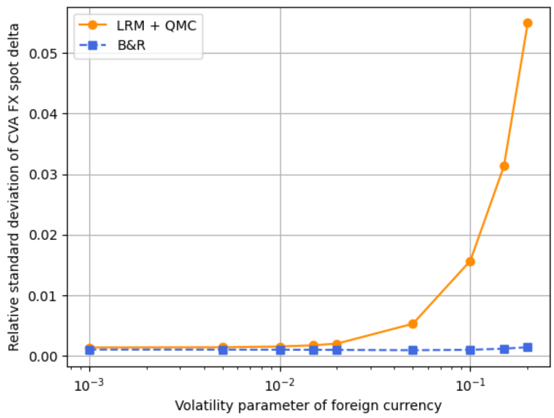
(a) Impact of σ_d on the relative errors.



(b) Impact of σ_d on the relative standard deviations.



(c) Impact of σ_f on the relative errors.



(d) Impact of σ_f on the relative standard deviations.

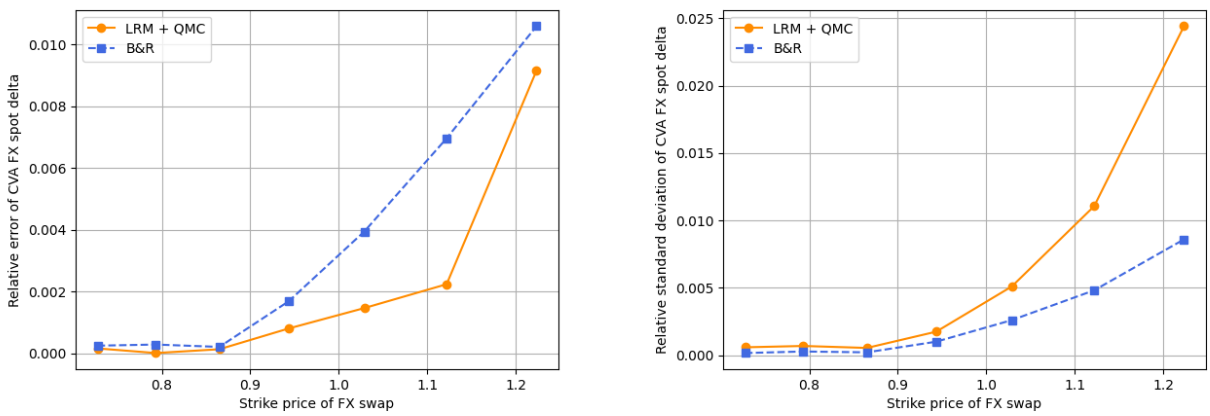
Figure 8.19: Impact of different σ_d and σ_f values on the relative errors and standard deviations of the CVA FX spot deltas of FX swaps.

Strike price of the FX swap

We define the set of strike prices as

$$\left[y(t_0)\,\exp\left(0.1c\sqrt{T}\right), c \in \{-1.5, -1, -0.5, 0, 0.5, 1, 1.5\}\right],$$

where T is the maturity date of the swap, set at 3 years. In this way, we create multiple FX swaps, some out-of-the-money, one at-the-money, and some in-the-money. In Figure 8.20a and 8.20b we illustrate the impact of different strike prices on the relative errors and standard deviations of the CVA FX spot deltas. For all the different strike prices, LRM is able to reproduce the B&R values quite well. However, just as for the FX digital options, we see an increase in the variance once the swap moves more out of the money. Importantly, this variance increase is again not attributed to the characteristics of LRM, as the strike price K does not influence LRM's score function. Instead, the increase in the variance is linked to the payoff of the FX swap. Specifically, the payoff of an FX swap at some time T_k is positive in case $y_F(T_k, T) > K$, and the payoff is negative in case $y_F(T, T) < K$. As the strike price K increases, the chance of $y_F(T_k, T)$ exceeding K diminishes, resulting in many MC paths assigning a negative value to the payoff $V(T_k)$, and therefore a zero value to the positive expected exposure at time T_k. Nevertheless, some paths will still exceed K and therefore, they assign a positive value to the payoff $V(T_k)$ and to the positive expected exposure. This results in an increase of the variance of the payoff of the FX swap. Therefore, the variance of both the LRM estimator and the B&R estimator will also increase. This is exactly what we see in the figure.



(a) Impact of K on the relative errors. **(b)** Impact of K on the relative standard deviations.

Figure 8.20: Impact of different strike prices of the FX swaps on the relative errors and standard deviations of the CVA FX spot deltas of FX swaps.

From this analysis, we can come up with a similar conclusion as for the FX digital options. For a certain range of parameters, LRM + QMC can reproduce the B&R values very well. However, for long maturities, low FX volatility values, relatively high correlation values, and high model volatility values, we encounter the well-known high-variance problems associated with LRM. Additionally, for deep out-of-the-money swaps, we observe an increase in the variance of LRM. However, given that the relative errors remain low and the increase in variance is not inherent to LRM itself, we can mitigate the concerns associated with high strike prices.

Metric / Method	LRM + QMC	B&R
Bias	$56/93$	$37/93$
Standard deviation	$4/93$	$89/93$

Table 8.6: Frequency of outperformance in relative error and standard deviation for both methods.

To draw a comprehensive conclusion about the overall performance of LRM compared to the B&R method, we compare the relative errors and standard deviations of all the 93 CVA FX spot deltas that we have computed in this section. In Table 8.6 we show for both methods the number of times it outperformed the other method in terms of relative error and standard deviation in the computation of the deltas. We observe similar results as for the FX digital options: LRM + QMC performs best in terms of relative error, indicating that the mean values of the LRM + QMC estimates are often very close to the reference values. However, in terms of relative standard deviations, LRM performs worse, suggesting that the LRM + QMC estimates exhibit significant volatility. Importantly, this observation can be relaxed again. Since LRM is used together with a QMC method, the standard deviations are empirically estimated using the power law, see Section 7.3 for more details. That implies that the standard deviations that we observe are not 'true' standard deviations, but rather empirical ones. It is therefore still possible for the true standard deviations to be slightly smaller. For this reason, we find it more useful to focus on the relative errors in this comparison.

Focusing solely on the errors, we conclude that for a certain range of parameters, LRM performs well and is able to match or outperform the B&R method. However, the performance for long maturity dates, low FX volatilities, highly correlated risk factors, and large values of model volatilities, remains poorly due to too high variance of the LRM estimators.

8.3. CVA zero deltas of IR swaps

In this section, we present the performance of LRM in the computation of CVA zero deltas of IR swaps, and we compare this performance against the B&R method. Ideally, we would also compare both methods with a reference value. However, due to the complex payoff of an IR swap, given in (2.45), we have not computed a reference value. Therefore, similar to the approach taken for FX swaps, we compute the B&R CVA delta using a large number of MC simulations and treat this value as the reference value of the CVA delta of an IR swap. Then we compare the performance of LRM and the B&R method with this reference value.

In this comparison, we will exploit one extra feature of LRM; the fact that LRM is able to compute multiple sensitivities in one single MC simulation. In this section, we will compute multiple CVA zero deltas of IR swaps. We specify $q+1$ zero rates as inputs for our model. In our model, we have chosen 10 zero rates with 10 corresponding pillar dates. The notation of these pillar dates and zero rates is presented in Table 8.7.

Maturity of the bond	Pillar Date	Zero Rate
1 month	T_1^p	z_1
2 months	T_2^p	z_2
3 months	T_3^p	z_3
6 months	T_4^p	z_4
12 months	T_5^p	z_5
18 months	T_6^p	z_6
24 months	T_7^p	z_7
36 months	T_8^p	z_8
48 months	T_9^p	z_9
60 months	T_{10}^p	z_{10}

Table 8.7: Notation of the pillar dates and zero rates that are used in the computation of CVA zero deltas of IR swaps.

For all the CVA sensitivities that we will compute in this section, we choose the parameter of differentiation, θ as a vector of multiple zero rates. This makes it possible for an IR swap, with for example a maturity of 2 years, to compute a total of 7 sensitivities, since the corresponding pillar dates of the first 7 zero rates are all smaller or equal to the maturity date of 2 years. Computing the CVA zero deltas of IR swaps has, however, also a downside; as we showed in Section 5.1, the application of LRM in computing CVA zero deltas requires us to split the LRM estimator into two parts, with one part being dependent on the payoff. This dependency on the payoff results in a loss of one of the advantages of LRM.

In the previous sections, we presented the CVA FX spot deltas of FX digital options and FX swaps across multiple sets of parameters. However, conducting a similar analysis for CVA zero deltas of IR swaps, involving multiple sensitivities for each parameter set, would be computationally intensive. Consequently, our procedure for the CVA zero deltas of IR swaps differs slightly from that for the FX digital options and FX swaps. We leverage the insights gained from the results obtained for FX digital options and FX swaps, where LRM demonstrated good performance for a certain range of parameters. However, LRM was faced with variance problems in cases of long maturity dates, high correlations, and high volatility values. Therefore, we focus our attention on these specific scenarios in Section 8.3.3. We also observed that the variance of LRM was a bit higher for the CVA deltas of deep out-of-the-money FX options/swaps. However, since the relative errors remained low, and the rise in variance was not due to LRM itself, we we already alleviated the concerns related to high strike prices. Still, we will devote some attention to different strike prices in Section 8.3.3. In summary, that means that we will compare the performance of LRM and the B&R method in the computation of CVA zero deltas for different maturity dates, volatility values, and strike prices. In these comparisons, we will perform similar analyses as we did for the FX digital and FX swap deltas.

To implement the B&R method, a bump size must be selected again. Since the payoff of an IR swap is continuous (see Section 2.7.1), no trade-off is needed between the bias and the variance. Therefore, we can choose h as small as possible, as long as we take care of the machine precision. We choose $h = 10^{-5}$.

In Section 8.3.1, we will present the set of parameters used in the comparisons. Then in Section 8.3.2, we will check the MC errors of our implementations and compare them with the theoretical MC errors. Lastly, in Section 8.3.3, we compare the performance of LRM and the B&R method in computing CVA zero deltas of IR swaps.

8.3.1. Set of parameters

In this section, we present the set of parameters that we will use in our analysis. Our parameter selection is mostly in line with the choices made by Oosterlee and Grzelak [6] in their numerical experiments. Only one parameter was chosen based on available data. According to the European Central Bank [18] the short rate of EUR at 25-10-2023 equals 0.03904.

Recall that in order to approximate the integral in the CVA formula (3.1) through the summation in (3.2), we specify a set of exposure times $\{T_1, ..., T_m\}$. We are faced with the same trade-off between precision and computational efficiency as we have seen for the FX digital options. Again, as we explained in Section 6.1, we choose to include 10 exposure dates per year.

With this information, we create the set $\mathcal{P}_1$ containing all the parameter values as follows:

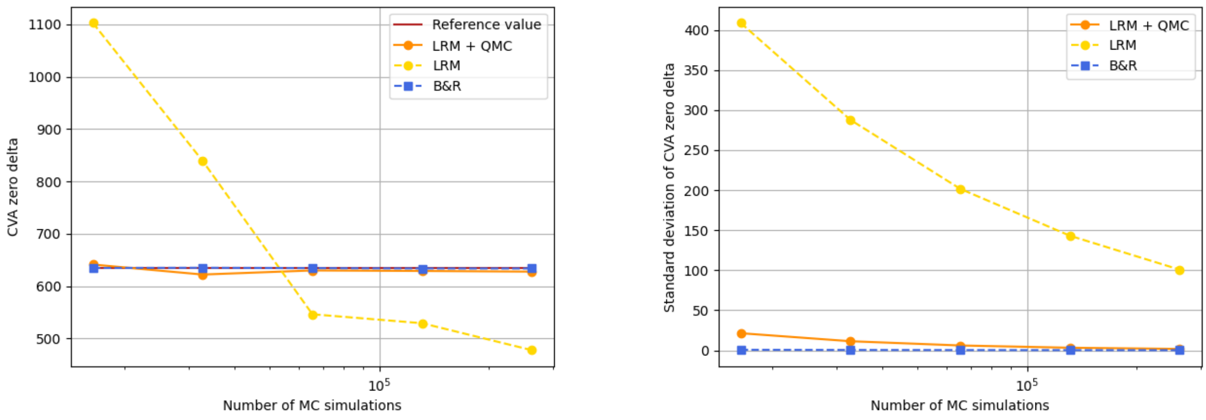
$$\mathcal{P}_1 = \left\{ \begin{array}{lcrlcr} T & = & 2, & m & = & 20, \\ t_0 & = & 0, & r(t_0) & = & 0.03904, \\ a & = & 0.02, & \sigma & = & 0.01, \\ Q & = & 10^6, & K & = & 0.03904, \\ R_c & = & 0, & \xi & = & 0.002. \end{array} \right\}. \tag{8.3}$$

Furthermore, the swaps that we consider have half-yearly payments. In the upcoming sections, we consistently employ the set $\mathcal{P}_1$ as the set of parameters for our comparisons, unless stated otherwise. Furthermore, for the LRM simulations, we will always make use of $N = 2^{18}$ MC simulations. For a fair comparison between LRM and the B&R method, we will adjust the number of B&R simulations so that the CPU times of both methods are approximately equal.

8.3.2. Convergence

In this section, we assess the convergence of the CVA zero deltas of IR swaps towards the reference value for all methods We also clarify our choice of N for each method, based on CPU times.

In Figure 8.21a we present the estimates of the CVA zero delta for the different methods using the set of parameters $\mathcal{P}_1$. Similarly, 8.21b provides the standard deviation of the CVA zero delta for the methods, where we varied the value of N from 2^{14} to 2^{18}. It is worth noting that LRM without the use of a QMC method is also included in these figures. We observe that all three methods are converging towards the analytical solution. For $N = 2^{18}$, LRM + QMC and B&R provide accurate estimates for the CVA zero delta. However, just like we observed for the FX digital options and FX swaps, LRM without a QMC method does not provide accurate results. Again, increasing the number of simulations could be considered as a potential solution, but it would likely lead to computational challenges. For that reason, we will exclude LRM without a QMC method from our comparisons.

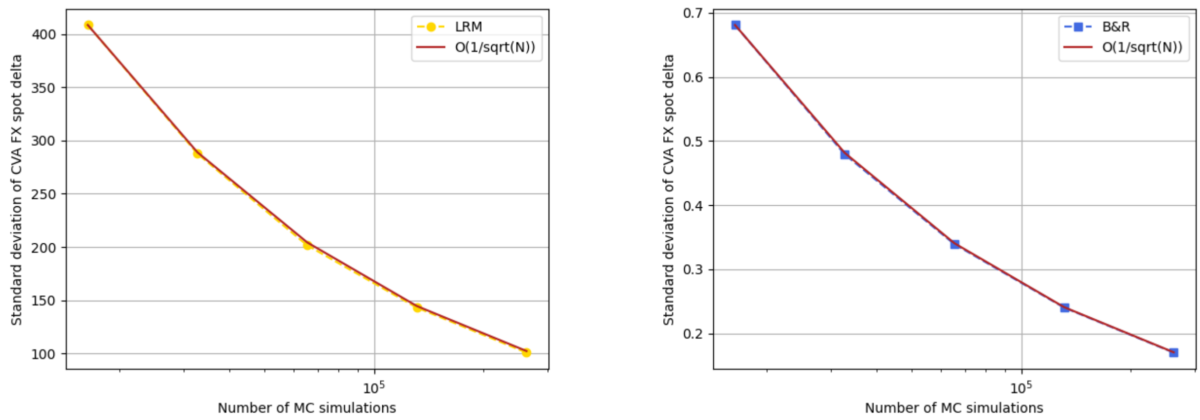


(a) CVA zero delta, compared with the reference value.

(b) Standard deviation of the CVA zero delta.

Figure 8.21: CVA zero delta, and its standard deviation of an IR Swap for different numbers of MC simulations.

Nevertheless, we can assess the convergence rate for the LRM estimates, and compare it with the theoretical rate of an ordinary MC method, which is of order $\mathcal{O}(1/\sqrt{N})$. A similar comparison can be made for the B&R method. In 8.22a and 8.22b we present this comparison between the convergence rate of LRM and the B&R method, and a convergence rate of $\mathcal{O}(1/\sqrt{N})$. Aligned with our expectation, we observe that the ordinary MC methods match the theoretical convergence rate. For the QMC methods, we can not perform a similar comparison, since the convergence rates are computed differently.



(a) Convergence rate of LRM.

(b) Convergence rate of the B&R method.

Figure 8.22: Convergence rates of LRM and the B&R method compared with rates of $\mathcal{O}(1/\sqrt{N})$.

As we previously observed in Figure 8.21a and 8.21b, LRM + QMC and B&R both provide accurate estimates of the CVA zero delta. To obtain a fair comparison between the performances, we examine the CPU times of both methods. However, instead of assessing the CPU time required to compute a single sensitivity, we leverage the advantage of LRM, which allows the computation of multiple CVA zero deltas in a single MC simulation. In this context, we use LRM to compute all 7 CVA zero deltas of the IR swap with parameter set $\mathcal{P}_1$. Additionally, we compute 7 separate sensitivities using the B&R method. Both methods make use of $N = 2^{18}$ simulations. The total CPU time (in seconds) for both methods to compute 7 sensitivities is presented in Table 8.8.

Method	CPU time (7 sensitivities)
Time LRM + QMC	32.15
Time B&R	49.38

Table 8.8: CPU times (in seconds) of both methods to compute 7 CVA zero deltas of the IR swap.

In comparison to LRM + QMC, we observe that B&R is roughly 1.5 times slower. We also know from our experience of the FX digital and FX swap deltas that the CPU time (almost) scales linearly with the number of MC simulations; doubling the number of simulations results in a CPU time twice as large. For a fair comparison of the methods, we take different simulation numbers so that both methods have similar CPU times. For that reason, we will fix the number of MC simulations for LRM + QMC at 2^{18}, and adjust the number of simulations for the B&R method accordingly by dividing the number of simulations by 1.5. The overview of these numbers is presented in Table 8.9. These numbers are used in the sensitivity analysis hereafter.

Method	N
LRM + QMC	2^{18}
B&R	$2^{18}/1.5$

Table 8.9: Choice of simulation numbers for both methods.

8.3.3. Sensitivity Analysis

In this section, we will conduct a sensitivity analysis of LRM and the B&R method for the computation of CVA zero deltas of IR swaps. We will consider several IR swaps, all using the parameter set $\mathcal{P}_1$, but varying one of the parameters. For each parameter variation, we compute the reference value and obtain estimates of the CVA zero deltas using the two methods. Subsequently, we compare the relative errors and standard deviations of the estimates of both methods. To ensure a valid comparison, we show the CPU times of all the methods and all the simulations in Appendix B.3. It is evident that all these CPU times are similar.

Maturity date

We define a set of maturity dates, ranging from one year to 10 years. For each maturity date, we compute many CVA zero deltas and we check the impact of these maturity dates on the relative errors and standard deviations of the CVA zero deltas. Table 8.10 shows these impacts of the CVA zero deltas of the IR swaps. We observe that, for all maturities and for all sensitivities, the relative errors and standard deviations of the B&R method are consistently lower than those computed by LRM. Furthermore, when examining a fixed zero rate with a corresponding pillar date, the relative errors and standard deviations of LRM increase as the maturity date increases. This aligns with the discussion in Section 6.3.

Volatility

We define a set of volatilities, ranging from 0.001 to 0.2. For each volatility, we compute many CVA zero deltas and we check the impact of these volatilities on the relative errors and standard deviations of the CVA zero deltas. Table 8.11 presents these impacts of the CVA zero deltas of the IR swaps. We observe consistently lower relative errors and standard deviations for the B&R method, for all volatilities and for all sensitivities. Additionally, when examining a fixed zero rate with a corresponding pillar date, the relative errors and standard deviations of LRM increase as the volatility parameter increases. This is in line with the discussion in Section 6.3.

		Relative error		Relative std. dev.	
T	θ	LRM + QMC	B&R	LRM + QMC	B&R
1	z_1	0.006101	0.000507	0.002837	0.000825
	z_2	0.002436	0.000063	0.001231	0.000380
	z_3	0.000148	0.000003	0.000770	0.000096
	z_4	0.000846	0.000068	0.000742	0.000170
	z_5	0.000475	0.000095	0.000298	0.000157
2	z_1	0.007036	0.000980	0.722966	0.000911
	z_2	0.011167	0.000635	0.621172	0.000756
	z_3	0.008892	0.000842	0.039621	0.000558
	z_4	0.015159	0.000230	0.005723	0.000332
	z_5	0.009576	0.000146	0.002698	0.000328
	z_6	0.002197	0.000217	0.002875	0.000221
	z_7	0.000413	0.000307	0.000467	0.000217
3	z_1	0.219082	0.001379	0.369042	0.001722
	z_2	0.060420	0.000118	0.184875	0.000309
	z_3	0.008233	0.000394	0.057870	0.000591
	z_4	0.014639	0.000061	0.011839	0.000355
	z_5	0.011502	0.000394	0.007193	0.000295
	z_6	0.003724	0.000254	0.006101	0.000324
	z_7	0.003930	0.000082	0.001633	0.000297
	z_8	0.000220	0.000142	0.000503	0.000185
4	z_1	0.624244	0.002524	0.474028	0.002122
	z_2	0.101681	0.001961	0.250525	0.000894
	z_3	0.082295	0.000461	0.045072	0.000437
	z_4	0.002298	0.000182	0.011653	0.000433
	z_5	0.020742	0.000012	0.007762	0.000299
	z_6	0.060048	0.000031	0.012852	0.000259
	z_7	0.024748	0.000297	0.007102	0.000221
	z_8	0.004010	0.000066	0.003173	0.000286
	z_9	0.000982	0.000025	0.000493	0.000139
5	z_1	0.219993	0.003286	0.692018	0.002393
	z_2	0.059912	0.000361	0.827957	0.001203
	z_3	0.086092	0.000255	0.201579	0.000372
	z_4	0.014737	0.000152	0.097092	0.000541
	z_5	0.059713	0.000124	0.047100	0.000397
	z_6	0.051874	0.000127	0.030875	0.000353
	z_7	0.008028	0.000121	0.017988	0.000239
	z_8	0.000099	0.000112	0.003527	0.000168
	z_9	0.003895	0.000474	0.003134	0.000273
	z_{10}	0.001175	0.000116	0.000478	0.000145
7	z_1	1.975675	0.001382	2.325940	0.002692
	z_2	1.598150	0.001077	1.077830	0.001531
	z_3	0.138737	0.000527	0.507366	0.000652
	z_4	0.389952	0.000225	0.183948	0.000714
	z_5	0.325010	0.000606	0.109473	0.000579
	z_6	0.269113	0.000351	0.154896	0.000549
	z_7	0.084934	0.000243	0.036441	0.000459
	z_8	0.058318	0.000065	0.038108	0.000393
	z_9	0.096621	0.000112	0.050432	0.000326
	z_{10}	0.013053	0.000335	0.007783	0.000350

Table 8.10: The relative errors and the standard deviations of multiple CVA zero deltas of IR swaps for different maturity dates, T.

σ	θ	Relative error LRM + QMC	Relative error B&R	Relative std. dev. LRM + QMC	Relative std. dev. B&R
0.001	z_1	0.660948	0.004276	1.624451	0.006506
	z_2	0.032374	0.000031	0.164619	0.000064
	z_3	0.001554	0.000002	0.013345	0.000001
	z_4	0.000529	0.000003	0.004011	0.000099
	z_5	0.001796	0.000030	0.001387	0.000082
	z_6	0.002043	0.000010	0.001786	0.000047
	z_7	0.000319	0.000022	0.000206	0.000046
0.005	z_1	0.213783	0.000222	0.504618	0.001366
	z_2	0.030847	0.000550	0.335383	0.000678
	z_3	0.000032	0.000090	0.022239	0.000275
	z_4	0.006814	0.000455	0.004030	0.000226
	z_5	0.003945	0.000082	0.001775	0.000186
	z_6	0.000542	0.000006	0.002079	0.000150
	z_7	0.000029	0.000093	0.000291	0.000124
0.01	z_1	0.007036	0.000980	0.722966	0.000911
	z_2	0.011167	0.000635	0.621172	0.000756
	z_3	0.008892	0.000842	0.039621	0.000558
	z_4	0.015159	0.000230	0.005723	0.000332
	z_5	0.009576	0.000146	0.002698	0.000328
	z_6	0.002197	0.000217	0.002875	0.000221
	z_7	0.000413	0.000307	0.000467	0.000217
0.02	z_1	0.042981	0.000768	0.372904	0.000643
	z_2	0.016472	0.000181	0.385043	0.000660
	z_3	0.005630	0.001239	0.030799	0.000623
	z_4	0.012430	0.000453	0.006842	0.000335
	z_5	0.009641	0.000115	0.003039	0.000368
	z_6	0.003969	0.000002	0.003041	0.000335
	z_7	0.000140	0.000282	0.000420	0.000239
0.05	z_1	0.028636	0.000250	0.183234	0.000469
	z_2	0.013225	0.000624	0.272955	0.000422
	z_3	0.005547	0.001029	0.021620	0.000458
	z_4	0.003774	0.000300	0.004650	0.000242
	z_5	0.003259	0.000050	0.002006	0.000231
	z_6	0.000710	0.000104	0.002496	0.000283
	z_7	0.000169	0.000117	0.000278	0.000146
0.1	z_1	0.098403	0.000251	1.152205	0.000467
	z_2	0.106827	0.000012	1.652447	0.000116
	z_3	0.023092	0.000290	0.098606	0.000235
	z_4	0.011454	0.000189	0.007158	0.000258
	z_5	0.009795	0.000050	0.003222	0.000272
	z_6	0.005438	0.000064	0.004822	0.000257
	z_7	0.000510	0.000090	0.001013	0.000228
0.15	z_1	0.201208	0.000423	2.554315	0.000519
	z_2	0.174458	0.000298	3.627934	0.000370
	z_3	0.034177	0.000305	0.250673	0.000209
	z_4	0.025075	0.000727	0.013881	0.000373
	z_5	0.023850	0.000506	0.006694	0.000466
	z_6	0.011573	0.000189	0.010559	0.000501
	z_7	0.000538	0.000403	0.002726	0.000416
0.2	z_1	0.265987	0.000730	4.094638	0.000579
	z_2	0.222690	0.000381	5.841438	0.000493
	z_3	0.041196	0.000613	0.451663	0.000367
	z_4	0.038411	0.000951	0.024052	0.000490
	z_5	0.036498	0.000729	0.011877	0.000651
	z_6	0.016190	0.000305	0.018435	0.000778
	z_7	0.000467	0.000696	0.005148	0.000603

Table 8.11: The relative errors and the standard deviations of multiple CVA zero deltas of IR swaps for different volatility values, σ.

Strike price of the IR swap

We define a set of strike prices as

$$\left[0.03094\,\text{exp}\Big(0.1c\sqrt{T}\Big), c \in \{-1.5, 0, 1.5\}\right],$$

where T is the maturity date of the swap, set at 2 years. In this way, we create multiple IR swaps, one out-of-the-money, one at-the-money, and one in-the-money. For each swap strike, we compute many CVA zero deltas, and we check the impact on the relative errors and standard deviations of the CVA zero deltas. Table 8.12 shows these impacts of the CVA zero deltas of the IR swaps. We see that for all three strike prices, the overall performance of LRM is decent. Importantly, however, the relative errors and standard deviations are consistently lower for the B&R method.

		Relative error		Relative std. dev.	
K	θ	LRM + QMC	B&R	LRM + QMC	B&R
0.03158	z_1	4.763861	0.016513	1.574974	0.027007
	z_2	0.463364	0.002834	0.915921	0.005851
	z_3	0.015146	0.003330	0.283786	0.001752
	z_4	0.017759	0.002191	0.009305	0.001290
	z_5	0.006449	0.000124	0.003100	0.000574
	z_6	0.000843	0.000049	0.005685	0.000195
	z_7	0.000411	0.000638	0.002516	0.000452
0.03904	z_1	0.007036	0.000980	0.722966	0.000911
	z_2	0.011167	0.000635	0.621172	0.000756
	z_3	0.008892	0.000842	0.039621	0.000558
	z_4	0.015159	0.000230	0.005723	0.000332
	z_5	0.009576	0.000146	0.002698	0.000328
	z_6	0.002197	0.000217	0.002875	0.000221
	z_7	0.000413	0.000307	0.000467	0.000217
0.04827	z_1	0.003348	0.000110	0.419345	0.000073
	z_2	0.004381	0.000138	0.552137	0.000089
	z_3	0.007480	0.000008	0.045144	0.000061
	z_4	0.009150	0.000502	0.004718	0.000297
	z_5	0.005600	0.000137	0.002340	0.000275
	z_6	0.000195	0.000076	0.002653	0.000202
	z_7	0.000352	0.000196	0.000409	0.000163

Table 8.12: The relative errors and the standard deviations of multiple CVA zero deltas for different strike prices of the IR swaps.

From this analysis, we can conclude that for all choices of parameters, LRM is not able to match the B&R values. We observe consistently lower relative errors and standard deviations for the B&R method compared to LRM. For long maturities, and large volatility values, the differences between the LRM and B&R values become even bigger due to the well-known high variance problems. Consequently, we would conclude that for all sets of parameters, the B&R method is preferred over LRM in the computation of CVA zero deltas of IR swaps.

Furthermore, examining all three tables, it becomes evident that the performance of LRM and the B&R method improves as the pillar date corresponding to the zero rate is growing more towards maturity. This is an interesting observation. However, it is crucial to note that each different zero rate represents a different CVA sensitivity, making a direct comparison between these sensitivities invalid. Consequently, drawing definitive conclusions regarding the difference in performance with respect to different pillar dates and zero rates is not feasible based on these observations.

9
Conclusion and Future Research

In this thesis, the application of LRM is extended to compute CVA sensitivities under the Hull-White short rate model. We have examined the performance of LRM in the computation of CVA sensitivities, with a focus on first-order sensitivities of three different OTC derivatives: FX digital options, FX swaps, and IR swaps..

In the computation of CVA sensitivities, LRM presents itself as an interesting alternative to other methods, such as the B&R method and pathwise sensitivity methods. In comparison to the B&R method, LRM is able to compute the same sensitivities at a lower computation time. Also, unlike the B&R method, LRM does not call for the selection of a bump size, offering a practical advantage, especially in the case of discontinuous payoffs. Furthermore, LRM only differentiates probability density functions, contrasting with pathwise sensitivity methods that necessitate the differentiation of payoffs. For the pathwise sensitivity methods, this can lead to problems for discontinuous payoffs, since the assumption of interchangeability between the differentiation and expectation operators may not hold. LRM, however, remains applicable to the computation of sensitivities of any type of payoff. Additionally, LRM stands out by enabling the computation of multiple sensitivities within one single MC simulation.

LRM comes, however, with some limitations. A practical limitation of LRM is that the user is required to choose a model, such that the probability density function of the risk factors is known, which is not guaranteed. Furthermore, in instances like short rate models, such as Hull-White, LRM still might need to differentiate payoff functions, due to the dependency of simulated forward ZCBs on the parameter of differentiation. In addition, LRM may encounter high variance problems, especially when the parameter of differentiation is included in many elements required to simulate a MC path. To overcome these high variance properties, we have employed variance reduction techniques. Specifically, we have used antithetic sampling, control variates, and QMC methods. As expected, the application of these techniques leads to significant variance reductions.

In our analysis we started with the comparison of LRM and the B&R method in the computation of CVA FX spot deltas of FX digital options and FX swaps. We observed that, for a certain range of parameters, LRM performs effectively and is capable of aligning with B&R values in the computation of the CVA FX spot deltas. In the majority of our experiments, we even found LRM estimates with a lower error than the B&R estimators. This suggests that, for a broad set of parameters, LRM serves as a computationally efficient alternative for the B&R method. However, for long maturity dates, low FX volatilities, high values of correlations, and high model volatilities, the variance of LRM is observed to be high, resulting in less favorable performance of LRM in comparison with the B&R method. In those cases, LRM does not prove to be a valid alternative to the B&R method.

In the last part of our analysis we exploited one of the advantages of LRM; the fact that LRM is able to compute multiple sensitivities in a single MC simulation. We have compared the performance of LRM and the B&R method in the computation of multiple CVA zero deltas of IR swaps. For a wide range of parameters, LRM did not manage to closely match the B&R values. In all the conducted experiments, we consistently observed better performance of the B&R method in terms of error and standard deviation. Additionally, we noted high variances of the LRM estimators for long maturities and high values of volatility. Consequently, we conclude that LRM is not an accurate alternative for the computation of CVA zero deltas of IR swaps.

Despite the promising results in the computation of the CVA FX spot deltas, we would still recommend further research. Firstly, it would be a valuable addition to make sure that the implementation of the reference value of the CVA FX spot delta of an FX swap matches the implementation of LRM and the B&R method. This alignment would facilitate a better comparison between the methods, contributing to a more conclusive result. Secondly, valuable insights could be obtained by expanding the assessment of LRM's performance to include CVA sensitivities of more financial derivatives, especially derivatives for which a reference value is available. As an example, examining the performance of sensitivities of caplet options could further enhance the understanding of LRM's capabilities.

The conclusion drawn from the analysis of CVA zero deltas of IR swaps indicates that LRM is not a suitable alternative to the B&R method, due to higher observed errors and variances. While antithetic sampling, control variates, and QMC methods were already employed as variance reduction techniques, more advanced methods that might yield further variance reduction could be explored. Stratified sampling and importance sampling are two such techniques, known for their potential to reduce the variance. However, these techniques are considerably more complex than the techniques that we used, making their practical application challenging. Nevertheless, despite their complexity, it could be worthwhile to investigate the effectiveness of these methods in improving LRM's performance.

In the analysis of multiple CVA zero deltas of IR swaps, we observed an improvement in the performance of LRM and the B&R method as the pillar date corresponding to the zero rate approaches maturity. However, the comparison involving different CVA sensitivities led us to the deduction that drawing a definitive conclusion about this observation is challenging. Despite this complexity, the observations remain interesting. Hence, we find it useful to delve deeper into these observations for a more nuanced understanding.

Furthermore, it might be useful to explore the performance of LRM in the computation of multiple CVA zero deltas for FX digital options and FX swaps. Given the simpler payoffs of FX digital options and FX swaps compared to IR swaps, there is potential for better results with LRM in this context. However, we will be faced with the same challenge as we encountered for the IR swaps. The payoffs of the FX digitals and FX swaps will depend on the zero rates via the ZCBs, making the LRM procedure payoff-dependent. This will result in a loss of one of the advantages of LRM.

Another addition to this research could be to insert more realistic scenarios. In this thesis, the focus has been on the computation of CVA sensitivities of a portfolio consisting either of a single swap or single option. However, in practice, banks have to compute CVA sensitivities to many more products and risk factors simultaneously. It might therefore be useful to add more products to the portfolios to provide a more holistic view of LRM's applicability and effectiveness in realistic scenarios.

Lastly, it might be useful to explore the impact of non-equidistant grids and the choice of risk factors. Throughout this thesis, the focus has been on using equidistant grids of exposure times, with the risk factors chosen as quantities known at each of these exposure times. In practice, banks often use grids that reflect the uneven distribution of risk over time, with a higher density of grid points in the earlier years where risk might be more concentrated. Exploring the possibility of selecting more risk factors to be evaluated in the earlier years, compared to later years, could therefore offer valuable insights into how well the method adapts to realistic risk factor distributions. This approach does, however, also has a limitation. Choosing a non-equidistant grid may result in the need for an interpolation scheme for time points where the risk factor quantities are unknown. This may contribute to an increase in the variance.

References

[1] I. Arregui, A. Leitao, B. Salvador, and C. Vázquez, “Efficient parallel Monte-Carlo techniques for pricing American options including counterparty credit risk,” *, International Journal of Computer Mathematics*, 2023.

[2] R. Palazzi, F. Conti, and M. Pioppi, “XVA Greeks via Likelihood Ratio Method,” 2022.

[3] M. Broadie and P. Glasserman, “Estimating Security Price Derivatives Using Simulation,” *Management Science*, pp. 269–285, 1996.

[4] P. Glasserman, *Monte Carlo Methods in Financial Engineering*. Springer, 2003.

[5] J. Hull and A. White, “Pricing Interest Rate Derivative Securities,” *Review of Financial Studies*, pp. 573–592, 1990.

[6] C. Oosterlee and L. A. Grzelak, *Mathematical Modeling and Computation in Finance*. World Scientific Publishing Europe Limited, 2019.

[7] F. Redig, *Lecture notes: Martingales in Discrete Time*, 2021.

[8] S. E. Shreve, *Stochastic Calculus for Finance II*. Springer Finance, 2004.

[9] D. Brigo and F. Mercurio, *Interest Rate Models - Theory and Practice With Smile, Inflation and Credit*. Springer Finance, 2006.

[10] O. A. Vasicek, “An equilibrium characterization of the term structure,” *Journal of Financial Economics*, pp. 177–188, 1997.

[11] M. Di Francesco, “A General Gaussian Interest Rate Model Consistent with the Current Term Structure,” *ISRN Probability and Statistics*, pp. 1–16, 2012.

[12] A. Green, *XVA: Credit, Funding and Capital Valuation Adjustments*, 1st ed. John Wiley & Sons Ltd, 2016.

[13] J. Gregory, *xVA Challange: Counterparty Credit Risk, Funding, Collateral and Capital*. Wiley, 2015.

[14] L. Capriotti, “Reducing the variance of likelihood ratio greeks in Monte Carlo,” *Proceedings of the 2008 Winter Simulation Conference*, pp. 587–593, 2008.

[15] M. Bianchetti, S. Kucherenko, and S. Scoleri, “Pricing and Risk Management with High-Dimensional Quasi-Monte Carlo and Global Sensitivity Analysis,” *Wilmott*, pp. 46–70, 2015.

[16] A. Owen, “Variance and discrepancy with alternative scramblings,” *ACM Transactions on Modeling and Computer Simulation*, vol. 13, pp. 363–378, 1993.

[17] S. Renzitti, P. Bastani, and S. Sivorot, “Accelerating CVA and CVA Sensitivities Using Quasi-Monte Carlo Methods,” *SSRN Electronic Journal*, pp. 1–51, 2020.

[18] *Euro short-term rate*. European Central Bank, 25-10-2023.

[19] *USD-EUR X-RATE*. Bloomberg, 25-10-2023.

A

Proofs

A.1. Proofs IR and HW dynamics

Theorem A.1.1. *Let $q+1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T_0^p, T_1^p, ..., T_q^p$. We assume that $T_{j-1} \leq t < T_j$ for some $j \in \{1, ..., q\}$ and we let $P^M(0,t)$ be the zero coupon bond with maturity at t valued at time 0. Then*

$$
\begin{aligned}
f^M(0,t) &= -\frac{\partial}{\partial t} \log \left(P^M(0,t)\right), \\
&= z_{j-1} + \frac{z_j - zj - 1}{T_j^p - T_{j-1}^p}(2t - T_{j-1}^p - T_0).
\end{aligned}
$$

Proof. By applying the interpolation scheme (2.8), we know that for $T_{j-1} \leq t < T_j$

$$
P^M(0,t) = \exp\left(-\left(z_{j-1} + \frac{z_j - z_{j-1}}{T_j^p - T_{j-1}^p}\left(t - T_{j-1}^p\right)\right)(t - T_0)\right),
$$

or equivalently

$$
\log P^M(0,t) = -z_{j-1}(t - T_0) - \frac{z_j - z_{j-1}}{T_j^p - T_{j-1}^p}\left(t - T_{j-1}^p\right)(t - T_0).
$$

We then get that

$$
\begin{aligned}
f^M(0,t) &= -\frac{\partial}{\partial t} \log \left(P^M(0,t)\right) \\
&= \frac{\partial}{\partial t}\left(z_{j-1}(t - T_0) + \frac{z_j - z_{j-1}}{T_j^p - T_{j-1}^p}\left(t - T_{j-1}^p\right)(t - T_0)\right) \\
&= z_{j-1} + \frac{z_j - zj - 1}{T_j^p - T_{j-1}^p}(2t - T_{j-1}^p - T_0).
\end{aligned}
$$

□

Theorem A.1.2. *Let $0 \leq s \leq t$. Then the solution of* (2.9)*, and hence the closed-form solution for the short rate, is given by*

$$
r(t) = r(s)e^{-a(t-s)} + \int_s^t e^{-a(t-u)}\chi(u)du + \int_s^t \sigma e^{-a(t-u)}dW(u).
$$

Proof. Rewriting the SDE for $e^{at}r(t)$ yields

$$
\begin{aligned}
d\left(e^{at}r(t)\right) &= ae^{at}r(t)dt + e^{at}dr(t) \\
&= ae^{at}r(t)dt + e^{at}\left[\theta(t) - ar(t)\right]dt + \sigma e^{at}dW(t) \\
&= \theta(t)e^{at}dt + \sigma e^{at}dW(t).
\end{aligned}
$$

Then by integrating both sides over $0 \leq s < t$, we obtain

$$\begin{aligned} e^{at}r(t) - e^{as}r(s) &= \int_s^t e^{au}\theta(u)du + \int_s^t \sigma e^{au}dW(u), \\ r(t) &= r(s)e^{-a(t-s)} + \int_s^t e^{-a(t-u)}\theta(u)du + \int_s^t \sigma e^{-a(t-u)}dW(u). \end{aligned}$$

□

Theorem A.1.3. *Let $0 \leq s < t$. The process $x(t)$ (2.13) given $\mathcal{F}(s)$ is normally distributed with mean and variance given by*

$$\begin{aligned} \mathbb{E}_s\left[x(t)\right] &= x(s)e^{-a(t-s)}, \\ \text{Var}_s\left[x(t)\right] &= \int_s^t \sigma^2 e^{-2a(t-u)}du \\ &= \frac{\sigma^2}{2a}(1 - e^{-2a(t-s)}). \end{aligned}$$

Moreover, for any $0 \leq s < t, T$, the covariance of $x(t)$ and $x(T)$ given $\mathcal{F}(0)$ is given by

$$\text{Cov}_s\left[x(t), x(T)\right] = \begin{cases} \frac{\sigma^2}{2a}\left(e^{-a(T-t)} - e^{-a(t+T-2s)}\right) & t \leq T, \\ \frac{\sigma^2}{2a}\left(e^{-a(t-T)} - e^{-a(T+t-2s)}\right) & t > T. \end{cases}$$

Proof. Since the increments of the Brownian Motion $W(t)$ are independent and follow a normal distribution, we get that $x(t)$ follows a normal distribution as well. Using Itô's Isometry (Theorem 2.1.5), finding the mean and variance is trivial. For the covariance, we start by assuming $0 \leq s \leq t \leq T$. Then if we split one of the integrals, we get

$$\begin{aligned} \text{Cov}_s\left[x(t), x(T)\right] &= \mathbb{E}_s\left[\left(x(t) - \mathbb{E}_s[x(t)]\right)\left(x(T) - \mathbb{E}_s[x(T)]\right)\right] \\ &= \mathbb{E}_s\left[\int_s^t \sigma e^{-a(t-u)}dW(u)\int_s^T \sigma e^{-a(T-u)}dW(u)\right] \\ &= \mathbb{E}_s\left[\int_s^t \sigma e^{-a(t-u)}dW(u)\left(\int_s^t \sigma e^{-a(T-u)}dW(u) + \int_t^T \sigma e^{-a(T-u)}dW(u)\right)\right] \\ &= \int_s^t \sigma^2 e^{-a(t+T-2u)}du + \mathbb{E}_s\left[\int_s^t \sigma e^{-a(t-u)}dW(u)\int_t^T \sigma e^{-a(T-u)}dW(u)\right]. \end{aligned}$$

Since the integrals in the expectation are independent, the expectation can be split into two expectations, which both are equal to zero. We end up with

$$\begin{aligned} \text{Cov}_s\left[x(t), x(T)\right] &= \int_s^t \sigma^2 e^{-a(t+T-2u)}du \\ &= \frac{\sigma^2}{2a}\left(e^{-a(T-t)} - e^{-a(t+T-2s)}\right). \end{aligned}$$

In a similar way, we find for $0 \leq s \leq T < t$ that

$$\begin{aligned} \text{Cov}_s\left[x(t), x(T)\right] &= \int_s^T \sigma^2 e^{-a(T+t-2u)}du \\ &= \frac{\sigma^2}{2a}\left(e^{-a(t-T)} - e^{-a(T+t-2s)}\right). \end{aligned}$$

□

Theorem A.1.4. *The integral of the process $\alpha(t)$ (2.14) be expressed as*

$$\int_t^T \alpha(s)ds = \log\left(\frac{P^M(0,t)}{P^M(0,T)}\right) + \frac{V(0,T) - V(0,t)}{2},$$

where $V(t,T)$ is given by (2.17)

Proof. We integrate the expression of $\alpha(t)$ (2.14) and we find

$$\begin{aligned}
\int_t^T \alpha(s)ds &= \int_t^T \left(f^M(0,s) + \frac{\sigma^2}{2a^2}\left(1-e^{-as}\right)^2 \right) ds \\
&= \int_t^T -\frac{\partial}{\partial s}\log\left(P^M(0,s)\right) + \frac{\sigma^2}{2a^2}\int_t^T \left(1-e^{-as}\right)^2 ds \\
&= \log\left(\frac{P^M(0,t)}{P^M(0,T)}\right) + \frac{1}{2}\frac{\sigma^2}{a^2}\left((T-t) + \frac{2}{a}\left(e^{-aT}-e^{-at}\right) - \frac{1}{2a}\left(e^{-2aT}-e^{-2at}\right)\right) \\
&= \log\left(\frac{P^M(0,t)}{P^M(0,T)}\right) + \frac{1}{2}\frac{\sigma^2}{a^2}\left((T-t) - 2\frac{\left(1-e^{-aT}\right)-\left(1-e^{-at}\right)}{a} + \frac{\left(1-e^{-2aT}\right)-\left(1-e^{-2at}\right)}{2a}\right) \\
&= \log\left(\frac{P^M(0,t)}{P^M(0,T)}\right) + \frac{V(0,T)-V(0,t)}{2}.
\end{aligned}$$

□

Theorem A.1.5. *For any $0 \le t, T$, the covariance of $\int_0^t r(u)du$ and $\int_0^T r(u)du$ given $\mathcal{F}(0)$ is given by*

$$\mathrm{Cov}_0\left[\int_0^t r(u)du, \int_0^T r(u)du\right] = \begin{cases} \frac{\sigma^2}{a^2}\left(t - \frac{1-e^{-at}+e^{-a(T-t)}-e^{-aT}}{a} + \frac{e^{-a(T-t)}-e^{-a(t+T)}}{2a}\right) & t \le T, \\ \frac{\sigma^2}{a^2}\left(T - \frac{1-e^{-aT}+e^{-a(t-T)}-e^{-at}}{a} + \frac{e^{-a(t-T)}-e^{-a(T+t)}}{2a}\right) & t > T, \end{cases}$$

Proof. We start by assuming $0 \le t \le T$. Then

$$\begin{aligned}
\mathrm{Cov}_0\left[\int_0^t r(u)du, \int_0^T r(u)du\right] &= \mathbb{E}_0\left[\left(\int_0^t r(u)du - \mathbb{E}_0\left[\int_0^t r(u)du\right]\right)\left(\int_0^T r(u)du - \mathbb{E}_0\left[\int_0^T r(u)du\right]\right)\right] \\
&= \mathbb{E}_0\left[\left(\int_0^t \frac{\sigma}{a}\left(1-e^{-a(t-u)}\right)dW(u)\right)\left(\int_0^T \frac{\sigma}{a}\left(1-e^{-a(T-u)}\right)dW(u)\right)\right].
\end{aligned}$$

In exactly the same way as in the proof of Theorem 2.3.3, we can split the latter integral into two integrals, of which one will be canceled out due to the independence with the left integral. We end up with

$$\begin{aligned}
\mathrm{Cov}_0\left[\int_0^t r(u)du, \int_0^T r(u)du\right] &= \int_0^t \frac{\sigma^2}{a^2}\left(1-e^{-a(t-u)}\right)\left(1-e^{-a(T-u)}\right)du \\
&= \frac{\sigma^2}{a^2}\left(t - \frac{1}{a}\left(1-e^{-at}+e^{-a(T-t)}-e^{-aT}\right) + \frac{1}{2a}\left(e^{-a(T-t)}-e^{-a(t+T)}\right)\right).
\end{aligned}$$

In a similar way, we find for $0 \le T < t$ that

$$\begin{aligned}
\mathrm{Cov}_0\left[\int_0^t r(u)du, \int_0^T r(u)du\right] &= \int_0^T \frac{\sigma^2}{a^2}\left(1-e^{-a(t-u)}\right)\left(1-e^{-a(T-u)}\right)du \\
&= \frac{\sigma^2}{a^2}\left(T - \frac{1}{a}\left(1-e^{-aT}+e^{-a(t-T)}-e^{-at}\right) + \frac{1}{2a}\left(e^{-a(t-T)}-e^{-a(T+t)}\right)\right).
\end{aligned}$$

□

Theorem A.1.6. *For any* $0 \leq t, T$*, the covariance of* $r(t)$ *and* $\int_0^T r(u)du$ *given* $\mathcal{F}(s)$ *is given by*

$$\text{Cov}_0\left[r(t), \int_0^T r(u)du\right] = \begin{cases} \frac{\sigma^2}{a}\left(\frac{1}{a}\left(1-e^{-at}\right) - \frac{1}{2a}\left(e^{-a(T-t)} - e^{-a(t+T)}\right)\right) & t \leq T, \\ \frac{\sigma^2}{a}\left(\frac{1}{a}\left(e^{-a(t-T)} - e^{-at}\right) - \frac{1}{2a}\left(e^{-a(t-T)} - e^{-a(T+t)}\right)\right) & t > T. \end{cases}$$

Proof. We start by assuming $0 \leq t \leq T$. Then

$$\begin{aligned} \text{Cov}_0\left[r(t), \int_0^T r(u)du\right] &= \mathbb{E}_0\left[\left(r(t) - \mathbb{E}_0\left[r(t)\right]\right)\left(\int_0^T r(u)du - \mathbb{E}_0\left[\int_0^T r(u)du\right]\right)\right] \\ &= \mathbb{E}_0\left[\left(\int_0^t \sigma e^{-a(t-u)}dW(u)\right)\left(\int_0^T \frac{\sigma}{a}\left(1-e^{-a(T-u)}\right)dW(u)\right)\right]. \end{aligned}$$

In exactly the same way as in the proof of Theorem 2.3.3, we can split the latter integral into two integrals, of which one will be canceled out due to the independence with the left integral. We end up with

$$\begin{aligned} \text{Cov}_0\left[r(t), \int_0^T r(u)du\right] &= \int_0^t \frac{\sigma^2}{a}e^{-a(t-u)}\left(1-e^{-a(T-u)}\right)du \\ &= \frac{\sigma^2}{a}\left(\frac{1}{a}\left(1-e^{-at}\right) - \frac{1}{2a}\left(e^{-a(T-t)} - e^{-a(t+T)}\right)\right). \end{aligned}$$

In a similar way, we find for $0 \leq T < t$ that

$$\begin{aligned} \text{Cov}_0\left[r(t), \int_0^T r(u)du\right] &= \int_0^T \frac{\sigma^2}{a}e^{-a(t-u)}\left(1-e^{-a(T-u)}\right)du \\ &= \frac{\sigma^2}{a}\left(\frac{1}{a}\left(e^{-a(t-T)} - e^{-at}\right) - \frac{1}{2a}\left(e^{-a(t-T)} - e^{-a(T+t)}\right)\right). \end{aligned}$$

$\square$

A.2. Proofs FX dynamics

Theorem A.2.1. *For any $0 \le s \le t \le T$, the covariance of $L_y(t)$ and $L_y(T)$ given $\mathcal{F}(s)$ is given by*

$$
\begin{aligned}
\text{Cov}_s[L_y(t), L_y(T)] = {} & \frac{\sigma_d^2}{a_d^2}\left(t - s - B_d(s,t) - \frac{e^{-a_d(T-t)} - e^{-a_d(T-s)}}{a_d} + \frac{e^{-a_d(T-t)} - e^{-a_d(t+T-2s)}}{2a_d}\right) \\
& - \frac{\sigma_f^2}{a_f^2}\left(t - s - B_f(s,t) - \frac{e^{-a_f(T-t)} - e^{-a_f(T-s)}}{a_f} + \frac{e^{-a_f(T-t)} - e^{-a_f(t+T-2s)}}{2a_f}\right) \\
& - \frac{\sigma_d\sigma_f}{a_d a_f}\rho_{d,f}\Bigg(t - s - B_d(s,t) - \frac{e^{-a_f(T-t)} - e^{-a_f(T-s)}}{a_f} + \frac{e^{-a_f(T-t)} - e^{-a_d(t-s)-a_f(T-s)}}{a_d + a_f} \\
& + t - s - B_f(s,t) - \frac{e^{-a_d(T-t)} - e^{-a_d(T-s)}}{a_d} + \frac{e^{-a_d(T-t)} - e^{-a_f(t-s)-a_d(T-s)}}{a_d + a_f}\Bigg) \\
& + \frac{\sigma_d\sigma_y}{a_d}\rho_{y,d}\left(t - s - B_d(s,t) + t - s - \frac{e^{-a_d(T-t)} - e^{-a_d(T-s)}}{a_d}\right) \\
& - \frac{\sigma_f\sigma_y}{a_d}\rho_{y,f}\left(t - s - B_f(s,t) + t - s - \frac{e^{-a_f(T-t)} - e^{-a_f(T-s)}}{a_f}\right) \\
& + \sigma_y^2(t-s).
\end{aligned}
$$

Proof. We start by writing the covariance as follows.

$$
\begin{aligned}
\text{Cov}_s[L_y(t), L_y(T)] &= \mathbb{E}_s\left[\left(L_y(t) - \mathbb{E}_s(L_y(t))\right)\left(L_y(T) - \mathbb{E}_s(L_y(T))\right)\right] \\
&= \mathbb{E}_s\Bigg[\left(\int_s^t \frac{\sigma_d}{a_d}\left(1 - e^{-a_d(t-u)}\right)dW_d^{\mathbb{Q}}(u) - \int_s^t \frac{\sigma_f}{a_f}\left(1 - e^{-a_f(t-u)}\right)dW_f^{\mathbb{Q}}(u) + \int_s^t \sigma_y dW_y^{\mathbb{Q}}(u)\right) \\
&\left(\int_s^T \frac{\sigma_d}{a_d}\left(1 - e^{-a_d(T-u)}\right)dW_d^{\mathbb{Q}}(u) - \int_s^T \frac{\sigma_f}{a_f}\left(1 - e^{-a_f(T-u)}\right)dW_f^{\mathbb{Q}}(u) + \int_s^T \sigma_y dW_y^{\mathbb{Q}}(u)\right)\Bigg].
\end{aligned}
$$

In exactly the same way as in the proof of Theorem 2.3.3, we can split the three latter integrals into six integrals. Three of them will be canceled out due to the independence with the left integrals, leaving us again with three integrals. Then after multiplying all the integrals, we end up with

$$
\begin{aligned}
\text{Cov}_s[L_y(t), L_y(T)] = {} & \int_s^t \frac{\sigma_d^2}{a_d^2}\left(1 - e^{-a_d(t-u)}\right)\left(1 - e^{-a_d(T-u)}\right)du + \int_s^t \frac{\sigma_f^2}{a_f^2}\left(1 - e^{-a_f(t-u)}\right)\left(1 - e^{-a_f(T-u)}\right)du \\
& - \frac{\sigma_d\sigma_f}{a_d a_f}\rho_{d,f}\Bigg(\int_s^t \left(1 - e^{-a_d(t-u)}\right)\left(1 - e^{-a_f(T-u)}\right)du \\
& + \int_s^t \left(1 - e^{-a_f(t-u)}\right)\left(1 - e^{-a_d(T-u)}\right)du\Bigg) \\
& + \frac{\sigma_d\sigma_y}{a_d}\rho_{y,d}\left(\int_s^t \left(1 - e^{-a_d(t-u)}\right)du + \int_s^t \left(1 - e^{-a_d(T-u)}\right)du\right) \\
& - \frac{\sigma_f\sigma_y}{a_f}\rho_{y,f}\left(\int_s^t \left(1 - e^{-a_f(t-u)}\right)du + \int_s^t \left(1 - e^{-a_f(T-u)}\right)du\right) + \int_s^t \sigma_y^2 du \\
= {} & \frac{\sigma_d^2}{a_d^2}\left(t - s - B_d(s,t) - \frac{e^{-a_d(T-t)} - e^{-a_d(T-s)}}{a_d} + \frac{e^{-a_d(T-t)} - e^{-a_d(t+T-2s)}}{2a_d}\right) \\
& - \frac{\sigma_f^2}{a_f^2}\left(t - s - B_f(s,t) - \frac{e^{-a_f(T-t)} - e^{-a_f(T-s)}}{a_f} + \frac{e^{-a_f(T-t)} - e^{-a_f(t+T-2s)}}{2a_f}\right) \\
& - \frac{\sigma_d\sigma_f}{a_d a_f}\rho_{d,f}\Bigg(t - s - B_d(s,t) - \frac{e^{-a_f(T-t)} - e^{-a_f(T-s)}}{a_f} + \frac{e^{-a_f(T-t)} - e^{-a_d(t-s)-a_f(T-s)}}{a_d + a_f} \\
& + t - s - B_f(s,t) - \frac{e^{-a_d(T-t)} - e^{-a_d(T-s)}}{a_d} + \frac{e^{-a_d(T-t)} - e^{-a_f(t-s)-a_d(T-s)}}{a_d + a_f}\Bigg)
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{\sigma_d\sigma_y}{a_d}\rho_{y,d}\left(t-s-B_d(s,t)+t-s-\frac{e^{-a_d(T-t)}-e^{-a_d(T-s)}}{a_d}\right)\\
&-\frac{\sigma_f\sigma_y}{a_d}\rho_{y,f}\left(t-s-B_f(s,t)+t-s-\frac{e^{-a_f(T-t)}-e^{-a_f(T-s)}}{a_f}\right)+\sigma_y^2(t-s).
\end{aligned}
$$

□

Theorem A.2.2. *For $0 \leq t < T$, the ZCB price $P(t,T)$ is given by*

$$
\begin{aligned}
P_f(t,T) &= \mathbb{E}_t^{\mathbb{Q}^f}\left[e^{-\int_t^T r_f(s)ds}\right]\\
&= \frac{1}{y(t)}\mathbb{E}_t^{\mathbb{Q}}\left[y(T)e^{-\int_t^T r_f(s)ds}\right]\\
&= \hat{A}_f(t,T)e^{-B_f(t,T)x_f(t)},
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{A}_f(t,T) &= \frac{P_f^M(0,T)}{P_f^M(0,t)}\,\text{exp}\left(\frac{1}{2}\left(V_f(t,T)-V_f(0,T)+V_f(0,t)\right)\right),\\
B_f(t,T) &= \frac{1-e^{-a_f(T-t)}}{a_f},
\end{aligned}
$$

and $f_f^M(0,t)$ and $V(t,T)$ are given by (2.32) *and* (2.35)*, respectively.*

Proof. Since the ZCB price is an expectation under the risk-neutral foreign measure, we need to change it to the risk-neutral domestic measure. We saw in Section 2.6 how to do this. Therefore,

$$
\begin{aligned}
P_f(t,T) &= \mathbb{E}_t^{\mathbb{Q}^f}\left[e^{-\int_t^T r_f(s)ds}\right]\\
&= \mathbb{E}_t^{\mathbb{Q}^f}\left[\frac{M_f(t)}{M_f(T)}\right]\\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\frac{M_f(t)}{M_f(T)}\frac{y(T)M_f(T)M_d(t)}{y(t)M_d(T)M_f(t)}\right]\\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\frac{y(T)M_d(t)}{y(t)M_d(T)}\right]\\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\frac{y(T)}{y(t)}e^{-\int_t^T r_d(s)ds}\right].
\end{aligned}
$$

Since $y(t)$ is log-normally distributed, we can use (2.40) to get

$$
\begin{aligned}
P_f(t,T) &= \mathbb{E}_t^{\mathbb{Q}}\left[\frac{y(T)}{y(t)}e^{-\int_t^T r_d(s)ds}\right]\\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\text{exp}\left(\int_t^T r_d(s)ds-\int_t^T r_f(s)ds-\frac{1}{2}\sigma_y^2(T-t)+\int_t^T\sigma_y dW_y^{\mathbb{Q}}(s)-\int_t^T r_d(s)ds\right)\right]\\
&= \mathbb{E}_t^{\mathbb{Q}}\left[\text{exp}\left(-\int_t^T r_f(s)ds-\frac{1}{2}\sigma_y^2(T-t)+\int_t^T\sigma_y dW_y^{\mathbb{Q}}(s)\right)\right].
\end{aligned}
$$

If we denote

$$
\begin{aligned}
Y(t,T) &= -\int_t^T r_f(s)ds-\frac{1}{2}\sigma_y^2(T-t)+\int_t^T\sigma_y dW_y^{\mathbb{Q}}(s)\\
&= -x_f(t)\frac{1-e^{-a_f(T-t)}}{a_f}+\frac{\sigma_f\sigma_y\rho_{y,f}}{a_f}\left(T-t-\frac{1-e^{-a_f(T-t)}}{a_f}\right)-\log\left(\frac{P_f^M(0,t)}{P_f^M(0,T)}\right)\\
&\quad-\frac{V_f(0,T)-V_f(0,t)}{2}-\int_t^T\frac{\sigma_f}{a_f}\left(1-e^{-a_f(T-u)}\right)dW_f^{\mathbb{Q}}(u)--\frac{1}{2}\sigma_y^2(T-t)+\int_t^T\sigma_y dW_y^{\mathbb{Q}}(s),
\end{aligned}
$$

then it is easy to see that, under the risk-neutral domestic measure, $Y(t,T)$ given $\mathcal{F}(t)$ is normally distributed with mean and variance given by

$$
\begin{aligned}
\mathbb{E}_t^{\mathbb{Q}}\left[Y(t,T)\right] &= -x_f(t)\frac{1-e_f^{-a_f(T-t)}}{a_f} + \frac{\sigma_f\sigma_y\rho_{y,f}}{a_f}\left(T-t-\frac{1-e^{-a_f(T-t)}}{a_f}\right) \\
&\quad - \log\left(\frac{P_f^M(0,t)}{P_f^M(0,T)}\right) - \frac{V_f(0,T)-V_f(0,t)}{2} - \frac{1}{2}\sigma_y^2(T-t), \\
\mathrm{Var}_t^{\mathbb{Q}}\left[Y(t,T)\right] &= V_f(t,T) + \sigma_y^2(T-t) - 2\frac{\sigma_f\sigma_y\rho_{y,f}}{a_f}\int_t^T\left(1-e^{-a_f(T-u)}\right)du \\
&= V_f(t,T) + \sigma_y^2(T-t) - 2\frac{\sigma_f\sigma_y\rho_{y,f}}{a_f}\left(T-t-\frac{1-e^{-a_f(T-t)}}{a_f}\right)
\end{aligned}
$$

We know that for a random variable Y that is normally distributed with mean μ and variance σ^2, then the expectation of e^{-Y} is given by $\mathbb{E}(e^{-Y}) = e^{-\mu+\sigma^2/2}$. Therefore,

$$
\begin{aligned}
P_f(t,T) &= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_f(s)ds - \frac{1}{2}\sigma_y^2(T-t) + \int_t^T \sigma_y dW_y^{\mathbb{Q}}(s)\right)\right] \\
&= \hat{A}_f(t,T)e^{-B_f(t,T)x_f(t)}
\end{aligned}
$$

where

$$
\begin{aligned}
\hat{A}_f(t,T) &= \frac{P^M(0,T)}{P^M(0,t)}\exp\left(\frac{1}{2}\left(V_f(t,T)-V_f(0,T)+V_f(0,t)\right)\right), \\
B_f(t,T) &= \frac{1-e^{-a_f(T-t)}}{a_f}.
\end{aligned}
$$

□

Theorem A.2.3. *Let $y_F(t,T)$ be the FX forward rate* (2.24) *at time t with maturity T, N_f be the foreign notional amount of the digital option, and K the strike price. Then the value of the digital option at time T_k is given by*

$$
V(T_k) = \bar{\gamma}N_f P_d(T_k,T)\Phi(d_2(T_k)),
$$

where $\bar{\gamma}=1$ for a call option and $\bar{\gamma}=-1$ for a put option, $\Phi(\cdot)$ is the standard normal cumulative distribution function, $P_d(t,T)$ is the ZCB price given by (2.23)*, and*

$$
d_2(T_k) = \frac{\log\left(y_F(T_k,T)/K\right) - \frac{1}{2}\sigma_{imp}^2(T_k,T)(T-T_k)}{\sigma_{imp}(T_k,T)\sqrt{T-T_k}},
$$

where

$$
\sigma_{imp}^2(t,T) = \frac{1}{T-t}\int_t^T \Sigma^2(s,T)ds,
$$

and where $\Sigma(t,T)$ is the instantaneous variance of the FX forward rate (2.47).

Proof. The price of the FX digital at time T_k is

$$
\begin{aligned}
V(T_k) &= \mathbb{E}_{T_k}^{\mathbb{Q}^d}\left[\frac{M_d(T_k)}{M_d(T)}V(T)\right] \\
&= \mathbb{E}_{T_k}^{\mathbb{Q}^d}\left[\frac{M_d(T_k)}{M_d(T)}N_f\mathbb{1}(\bar{\gamma}(y_F(T,T)>K))\right] \\
&= \bar{\gamma}N_f P_d(T_k,T)\mathbb{E}_{T_k}^{\mathbb{Q}^{T,d}}\left[\mathbb{1}(y_F(T,T)>K)\right] \\
&= \bar{\gamma}N_f P_d(T_k,T)\mathbb{Q}^{T,d}\left[y_F(T,T)>K\right].
\end{aligned}
$$

We know from Theorem 2.7.3 that

$$dy_F(t,T) = y_F(t,T)\Sigma(t,T)dW_F^T(t),$$

where $\Sigma(t,T)$ is the instantaneous variance of the FX forward given by 2.47. Using Ito's Lemma (Theorem 2.1.4), it is easy to see that

$$d\log y_F(t,T) = -\frac{1}{2}\Sigma^2(t,T)dt + \Sigma(t,T)dW_F^T(t).$$

Therefore, for $0 \leq T_k \leq T$, we get

$$y_F(T,T) = y_F(T_k,T)\exp\left(-\frac{1}{2}\int_{T_k}^T \Sigma^2(s,T)ds + \int_{T_k}^T \Sigma(s,T)dW_F^T(s)\right).$$

Since the Brownian Motion $W_F^T(t)$ is normally distributed, it is easy to see that the integral $\int_{T_k}^T \Sigma(s,T)dW_F^T(s)$ has mean zero and variance

$$\mathrm{Var}_{T_k}^{T,d}\left[\int_{T_k}^T \Sigma(s,T)dW_F^T(s)\right] = \int_{T_k}^T \Sigma^2(s,T)ds.$$

Thus, $\int_{T_k}^T \Sigma(s,T)dW_F^T(s)$ has the same distribution as $Z\cdot\sqrt{\int_{T_k}^T \Sigma^2(s,T)ds}$, with $Z \sim \mathcal{N}(0,1)$, and we can rewrite $y_F(T,T)$ as

$$y_F(T,T) = y_F(T_k,T)\exp\left(-\frac{1}{2}\int_{T_k}^T \Sigma^2(s,T)ds + Z\cdot\sqrt{\int_{T_k}^T \Sigma^2(s,T)ds}\right).$$

But then we can restructure some terms to get

$$\begin{aligned} y_F(T,T) > K &\iff y_F(T_k,T)\exp\left(-\frac{1}{2}\int_{T_k}^T \Sigma^2(s,T)ds + Z\cdot\sqrt{\int_{T_k}^T \Sigma^2(s,T)ds}\right) > K \\ &\iff Z \leq \frac{\log\left(y_F(T_k,T)/K\right) - \frac{1}{2}\int_{T_k}^T \Sigma^2(s,T)ds}{\sqrt{\int_{T_k}^T \Sigma^2(s,T)ds}} := d_2(T_k) \end{aligned}$$

We conclude that

$$\begin{aligned} V(T_k) &= \bar{\gamma}N_f P_d(T_k,T)\mathbb{Q}^{T,d}\left[y_F(T,T) > K\right] \\ &= \bar{\gamma}N_f P_d(T_k,T)\Phi(d_2(T_k)), \end{aligned}$$

with

$$d_2(T_k) = \frac{\log\left(y_F(T_k,T)/K\right) - \frac{1}{2}\int_{T_k}^T \Sigma^2(s,T)ds}{\sqrt{\int_{T_k}^T \Sigma^2(s,T)ds}} = \frac{\log\left(y_F(T_k,T)/K\right) - \frac{1}{2}\sigma_{\mathrm{imp}}^2(T_k,T)(T-T_k)}{\sigma_{\mathrm{imp}}(T_k,T)\sqrt{T-T_k}},$$

and with

$$\sigma_{\mathrm{imp}}^2(t,T) = \frac{1}{T-t}\int_t^T \Sigma^2(s,T)ds.$$

□

Theorem A.2.4. *Let N_f be the foreign notional amount of the swap, and K the fixed FX rate. Then the expected exposure of the FX payer swap at time T_k and its derivative with respect to $y(t_0)$ are given by*

$$\textit{EE}(t_0,T_k) = N_f P_d(t_0,T)\left(y_F(t_0,T)\Phi(d_1(T_k)) - K\Phi(d_2(T_k))\right),$$
$$\frac{\partial \textit{EE}(t_0,T_k)}{\partial y(t_0)} = N_f P_d(t_0,T)\left(\frac{P_f(t_0,T)}{P_d(t_0,T)}\left(\Phi(d_1(T_k)) + \frac{\phi(d_1(T_k))}{\sigma_{\textit{imp}}(t_0,T_k)\sqrt{T_k-t_0}}\right) - K\frac{\phi(d_2(T_k))}{y(t_0)\sigma_{\textit{imp}}(t_0,T_k)\sqrt{T_k-t_0}}\right),$$

where $P_d(t,T)$ and $P_f(t,T)$ are the ZCB prices in the domestic and foreign currency, respectively, $y_F(t,T)$ is the FX forward rate (2.24)*, $\Phi(\cdot)$ is the standard normal cumulative distribution function with corresponding probability density function $\phi(\cdot)$, $d_1(T_k)$ and $d_2(T_k)$ are given by*

$$d_1(T_k) = \frac{\log\left(y_F(t_0,T)/K\right) + \frac{1}{2}\sigma^2_{\textit{imp}}(t_0,T_k)(T_k-t_0)}{\sigma_{\textit{imp}}(t_0,T_k)\sqrt{T_k-t_0}}, \quad d_2(T_k) = d_1(T_k) - \sigma_{\textit{imp}}(t_0,T_k)\sqrt{T_k-t_0},$$

and

$$\sigma^2_{\textit{imp}}(t,T) = \frac{1}{T-t}\int_t^T \Sigma^2(s,T)ds,$$

where $\Sigma(t,T)$ is the instantaneous variance of the FX forward rate (2.47)*. Similarly, the expected exposure of the FX payer receiver at time T_k and its derivative with respect to $y(t_0)$ are given by*

$$\textit{EE}(t_0,T_k) = N_f P_d(t_0,T)\left(K\Phi(-d_2(T_k)) - y_F(t_0,T)\Phi(-d_1(T_k))\right),$$
$$\frac{\partial \textit{EE}(t_0,T_k)}{\partial y(t_0)} = N_f P_d(t_0,T)\left(-K\frac{\phi(-d_2(T_k))}{y(t_0)\sigma_{\textit{imp}}(t_0,T_k)\sqrt{T_k-t_0}} - \frac{P_f(t_0,T)}{P_d(t_0,T)}\left(\Phi(-d_1(T_k)) - \frac{\phi(-d_1(T_k))}{\sigma_{\textit{imp}}(t_0,T_k)\sqrt{T_k-t_0}}\right)\right).$$

Proof. By Theorem 2.7.2 we know that the value of the FX swap at time T_k is given by

$$V(T_k) = \bar{\beta} P_d(T_k,T) N_f (y_F(T_k,T) - K),$$

Then for the expected exposure for the payer swap, we use the definition of the ZCB (2.4) and switch to the T-forward measure to get

$$\begin{aligned}
\text{EE}(t_0,T_k) &= \mathbb{E}^{\mathbb{Q},d}_{t_0}\left[\frac{M(t_0)}{M(t)}\max\left(V(T_k)\right)\right] \\
&= N_f \mathbb{E}^{\mathbb{Q},d}_{t_0}\left[\frac{M(t_0)}{M(t)}P_d(T_k,T)\max\left(y_F(T_k,T)-K\right)\right] \\
&= N_f \mathbb{E}^{\mathbb{Q},d}_{t_0}\left[\frac{M(t_0)}{M(T)}\max\left(y_F(T_k,T)-K\right)\right] \\
&= N_f P_d(t_0,T)\mathbb{E}^{\mathbb{T},d}_{t_0}\left[\max\left(y_F(T_k,T)-K\right)\right] \\
&= N_f P_d(t_0,T)\left(\mathbb{E}^{\mathbb{T},d}_{t_0}\left[y_F(T_k,T)\mathbb{1}(y_F(T_k,T)>K)\right] - K\mathbb{E}^{\mathbb{T},d}_{t_0}\left[\mathbb{1}(y_F(T_k,T)>K)\right]\right)
\end{aligned}$$

In a similar way as we did in the Proof of Theorem A.2.3, we derive for, $Z \sim \mathcal{N}(0,1)$, that

$$y_F(T_k,T) = y_F(t_0,T)\exp\left(-\frac{1}{2}\int_{t_0}^{T_k}\Sigma^2(s,T)ds + Z\cdot\sqrt{\int_{t_0}^{T_k}\Sigma^2(s,T)ds}\right),$$

where $\Sigma(t,T)$ is the instantaneous variance of the FX forward given by 2.47. Then

$$\begin{aligned}
y_F(T_k,T) > K &\iff y_F(t_0,T)\exp\left(-\frac{1}{2}\int_{t_0}^{T_k}\Sigma^2(s,T)ds + Z\cdot\sqrt{\int_{t_0}^{T_k}\Sigma^2(s,T)ds}\right) > K \\
&\iff Z \leq \frac{\log\left(y_F(t_0,T)/K\right) - \frac{1}{2}\int_{t_0}^{T_k}\Sigma^2(s,T)ds}{\sqrt{\int_{t_0}^{T_k}\Sigma^2(s,T)ds}} := d_2(T_k)
\end{aligned}$$

Then we conclude that

$$
\begin{aligned}
\mathsf{EE}(t_0,T_k) &= N_f P_d(t_0,T)\left(\mathbb{E}^{\mathbb{T},d}_{t_0}\left[y_F(T_k,T)\mathbb{1}(y_F(T_k,T)>K)\right] - K\mathbb{E}^{\mathbb{T},d}_{t_0}\left[\mathbb{1}(y_F(T_k,T)>K)\right]\right)\\
&= N_f P_d(t_0,T)\Bigg(\mathbb{E}^{\mathbb{T},d}_{t_0}\Bigg[y_F(t_0,T)\exp\left(-\frac{1}{2}\int_{t_0}^{T_k}\Sigma^2(s,T)ds + Z\cdot\sqrt{\int_{t_0}^{T_k}\Sigma^2(s,T)ds}\right)\mathbb{1}(Z\leq d_2(T_k))\Bigg]\\
&\quad - K\mathbb{E}^{\mathbb{T},d}_{t_0}\left[\mathbb{1}(Z\leq d_2(T_k))\right]\Bigg)\\
&= N_f P_d(t_0,T)\left(y_F(t_0,T)\frac{1}{2\pi}\int_{d_2(T_k)}^{\infty} e^{-\frac{1}{2}\left(z+\sqrt{\int_{t_0}^{T_k}\Sigma^2(s,T)ds}\right)^2}dz - K\Phi(d_2(T_k))\right)\\
&= N_f P_d(t_0,T)\left(y_F(t_0,T)\Phi(d_1(T_k)) - K\Phi(d_2(T_k))\right),
\end{aligned}
$$

with

$$
d_1(T_k) = \frac{\log\left(y_F(t_0,T)/K\right) + \frac{1}{2}\sigma^2_{\mathsf{imp}}(t_0,T_k)(T_k-t_0)}{\sigma_{\mathsf{imp}}(t_0,T_k)\sqrt{T_k-t_0}}, \quad d_2(T_k) = d_1(T_k) - \sigma_{\mathsf{imp}}(t_0,T_k)\sqrt{T_k-t_0},
$$

and

$$
\sigma^2_{\mathsf{imp}}(t,T) = \frac{1}{T-t}\int_t^T \Sigma^2(s,T)ds,
$$

where $\Sigma(t,T)$ is the instantaneous variance of the FX forward rate (2.47). For the derivative of the exposure at time T_k with respect to $y(t_0)$, we simply use the chain rule and use the fact that the derivative of the cumulative distribution function of a normal distribution, $\Phi(\cdot)$, is the probability density function of a normal distribution, $\phi(\cdot)$. The proof for the FX receiver swap is similar. $\square$

A.3. Proofs CVA zero delta estimation

Theorem A.3.1. *Let $q+1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T_0^p, T_1^p, ..., T_q^p$. Assume $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$, and let for $k = 1, ..., m$*

$$\mu_{r,k} = f^M(0, T_k) + \frac{\sigma^2}{2a^2}\left(1 - e^{-aT_k}\right)^2,$$
$$\mu_{I,k} = \log\left(\frac{1}{P^M(0,T_k)}\right) + \frac{V(0,T_k)}{2},$$

as in (5.3) *and* (5.4)*. Then*

$$\dot{\mu}(\theta) = \begin{pmatrix}\dot{\mu}_r(\theta)\\ \dot{\mu}_I(\theta)\end{pmatrix} \in \mathbb{R}^{2m\times 1},$$

where

$$\dot{\mu}_r(\theta) = \begin{pmatrix}\frac{\partial}{\partial\theta}\mu_{r,1} & \cdots & \frac{\partial}{\partial\theta}\mu_{r,j-1} & \frac{\partial}{\partial\theta}\mu_{r,j} & \frac{\partial}{\partial\theta}\mu_{r,j+1} & \cdots & \frac{\partial}{\partial\theta}\mu_{r,m}\end{pmatrix}^T,$$
$$\dot{\mu}_I(\theta) = \begin{pmatrix}\frac{\partial}{\partial\theta}\mu_{I,1} & \cdots & \frac{\partial}{\partial\theta}\mu_{I,j-1} & \frac{\partial}{\partial\theta}\mu_{I,j} & \frac{\partial}{\partial\theta}\mu_{I,j+1} & \cdots & \frac{\partial}{\partial\theta}\mu_{I,m}\end{pmatrix}^T.$$

If we take T_k such that $T_{j-1}^p \leq T_k < T_j^p$, then

$$\frac{\partial}{\partial\theta}\mu_{r,k} = \frac{2T_k - T_{j-1}^p - T_0}{T_j^p - T_{j-1}^p},$$
$$\frac{\partial}{\partial\theta}\mu_{I,k} = \frac{(T_k - T_{j-1}^p)(T_k - T_0)}{T_j^p - T_{j-1}^p}.$$

On the other hand, if we take T_k such that $T_j^p \leq T_k < T_{j+1}^p$, then

$$\frac{\partial}{\partial\theta}\mu_{r,k} = 1 - \frac{2T_k - T_j^p - T_0}{T_{j+1}^p - T_j^p},$$
$$\frac{\partial}{\partial\theta}\mu_{I,k} = T_k - T_0 - \frac{(T_k - T_j^p)(T_k - T_0)}{T_{j+1}^p - T_j^p}.$$

Lastly, for all T_k such that $T_k < T_{j-1}^p$ or $T_k > T_{j+1}^p$, we obtain

$$\frac{\partial}{\partial\theta}\mu_{r,k} = 0,$$
$$\frac{\partial}{\partial\theta}\mu_{I,k} = 0.$$

Proof. We start by assuming that $T_{j-1}^p \leq T_k < T_j^p$,. Due to our interpolation scheme, we know that

$$P^M(0,T_k) = \exp\left(-\left(z_{j-1} + \frac{z_j - z_{j-1}}{T_j^p - T_{j-1}^p}\left(T_k - T_{j-1}^p\right)\right)(T_k - T_0)\right),$$

or equivalently

$$\log P^M(0,T_k) = -z_{j-1}(T_k - T_0) - \frac{z_j - z_{j-1}}{T_j^p - T_{j-1}^p}\left(T_k - T_{j-1}^p\right)(T_k - T_0).$$

We then get that

$$\begin{aligned}\frac{\partial}{\partial z_j}\mu_{r_k} &= \frac{\partial}{\partial z_j}\left(f^M(0,T_k) + \frac{\sigma^2}{2a^2}\left(1 - e^{-aT_k}\right)^2\right)\\ &= \frac{\partial}{\partial z_j}\left(f^M(0,T_k)\right)\end{aligned}$$

$$
\begin{aligned}
&= \frac{\partial}{\partial z_j}\left(-\frac{\partial}{\partial T_k}\left(\log P^M(0,T_k)\right)\right) \\
&= \frac{\partial}{\partial z_j}\left(\frac{\partial}{\partial T_k}\left(z_{j-1}(T_k-T_0)+\frac{z_j-z_{j-1}}{T_j^p-T_{j-1}^p}\left(T_k-T_{j-1}^p\right)(T_k-T_0)\right)\right) \\
&= \frac{\partial}{\partial z_j}\left(z_{j-1}+\frac{z_j-z_{j-1}}{T_j^p-T_{j-1}^p}\left(2T_k-T_{j-1}^p-T_0\right)\right) \\
&= \frac{2T_k-T_{j-1}^p-T_0}{T_j^p-T_{j-1}^p}, \\
\frac{\partial}{\partial z_j}\mu_{I_k} &= \frac{\partial}{\partial z_j}\left(\log\left(\frac{1}{P^M(0,T_k)}\right)+\frac{V(0,T_k)}{2}\right) \\
&= -\frac{\partial}{\partial z_j}\log\left(P^M(0,T_k)\right) \\
&= \frac{\partial}{\partial z_j}\left(z_{j-1}(T_k-T_0)+\frac{z_j-z_{j-1}}{T_j^p-T_{j-1}^p}\left(T_k-T_{j-1}^p\right)(T_k-T_0)\right) \\
&= \frac{\left(T_k-T_{j-1}^p\right)(T_k-T_0)}{T_j^p-T_{j-1}^p}.
\end{aligned}
$$

Now we assume that $T_j^p \leq T_k < T_{j+1}^p$,. Due to our interpolation scheme, we know that

$$
P^M(0,T_k) = \exp\left(-\left(z_j+\frac{z_{j+1}-z_j}{T_{j+1}^p-T_j^p}\left(T_k-T_j^p\right)\right)(T_k-T_0)\right),
$$

or equivalently

$$
\log P^M(0,T_k) = -z_j(T_k-T_0)-\frac{z_{j+1}-z_j}{T_{j+1}^p-T_j^p}\left(T_k-T_j^p\right)(T_k-T_0).
$$

We then get that

$$
\begin{aligned}
\frac{\partial}{\partial z_j}\mu_{r_k} &= \frac{\partial}{\partial z_j}\left(f^M(0,T_k)+\frac{\sigma^2}{2a^2}\left(1-e^{-aT_k}\right)^2\right) \\
&= \frac{\partial}{\partial z_j}\left(f^M(0,T_k)\right) \\
&= \frac{\partial}{\partial z_j}\left(-\frac{\partial}{\partial T_k}\left(\log P^M(0,T_k)\right)\right) \\
&= \frac{\partial}{\partial z_j}\left(\frac{\partial}{\partial T_k}\left(z_j(T_k-T_0)+\frac{z_{j+1}-z_j}{T_{j+1}^p-T_j^p}\left(T_k-T_j^p\right)(T_k-T_0)\right)\right) \\
&= \frac{\partial}{\partial z_j}\left(z_j+\frac{z_{j+1}-z_j}{T_{j+1}^p-T_j^p}\left(2T_k-T_j^p-T_0\right)\right) \\
&= 1-\frac{2T_k-T_j^p-T_0}{T_{j+1}^p-T_j^p}, \\
\frac{\partial}{\partial z_j}\mu_{I_k} &= \frac{\partial}{\partial z_j}\left(\log\left(\frac{1}{P^M(0,T_k)}\right)+\frac{V(0,T_k)}{2}\right) \\
&= -\frac{\partial}{\partial z_j}\log\left(P^M(0,T_k)\right) \\
&= \frac{\partial}{\partial z_j}\left(z_j(T_k-T_0)+\frac{z_{j+1}-z_j}{T_{j+1}^p-T_j^p}\left(T_k-T_j^p\right)(T_k-T_0)\right) \\
&= T_k-T_0-\frac{\left(T_k-T_j^p\right)(T_k-T_0)}{T_{j+1}^p-T_j^p}.
\end{aligned}
$$

Lastly, for all T_k such that $T_k < T^p_{j-1}$ or $T_k \geq T^p_{j+1}$, we get that the expression of $\log P^M(0,T_k)$ via the interpolation scheme is independent of z_j. Therefore

$$\frac{\partial}{\partial\theta}\mu_{r,k} = 0,$$
$$\frac{\partial}{\partial\theta}\mu_{I,k} = 0.$$

□

Theorem A.3.2. *Let $q+1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T^p_0, T^p_1, ..., T^p_q$. Assume $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$. Lastly we assume $T_k \in \{T_1, ..., T_m\}$, and $T_i > T_k$. Then*

$$\frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right) = C(T_k,T_i)\frac{P^M(0,T_i)}{P^M(0,T_k)},$$

where

$$C(T_k,T_i) = \begin{cases} -\frac{(T_i-T^p_{j-1})(T_i-T_0)}{T^p_j-T^p_{j-1}} & \text{for } T_k < T^p_{j-1} \leq T_i < T^p_j, \\ -(T_i-T_0) + \frac{(T_i-T^p_j)(T_i-T_0)}{T^p_{j+1}-T^p_j} & \text{for } T_k < T^p_{j-1} < T^p_j \leq T_i < T^p_{j+1}, \\ -\frac{(T_i-T^p_{j-1})(T_i-T_0)-(T_k-T^p_{j-1})(T_k-T_0)}{T^p_j-T^p_{j-1}} & \text{for } T^p_{j-1} \leq T_k < T_i < T^p_j, \\ -(T_i-T_0) + \frac{(T_i-T^p_j)(T_i-T_0)}{T^p_{j+1}-T^p_j} + \frac{(T_k-T^p_{j-1})(T_k-T_0)}{T^p_j-T^p_{j-1}} & \text{for } T^p_{j-1} \leq T_k < T^p_j \leq T_i < T^p_{j+1}, \\ \frac{(T_k-T^p_{j-1})(T_k-T_0)}{T^p_j-T^p_{j-1}} & \text{for } T^p_{j-1} \leq T_k < T^p_j < T^p_{j+1} \leq T_i, \\ -(T_i-T_k) + \frac{(T_i-T^p_j)(T_i-T_0)-(T_k-T^p_j)(T_k-T_0)}{T^p_{j+1}-T^p_j} & \text{for } T^p_j \leq T_k < T_i < T^p_{j+1}, \\ T_k - T_0 - \frac{(T_k-T^p_j)(T_k-T_0)}{T^p_{j+1}-T^p_j} & \text{for } T^p_j \leq T_k < T^p_{j+1} \leq T_i, \end{cases}$$

Proof. Just as in the proof of Theorem 5.1.1, we will first specify the interpolation scheme. For a general time point T_l between pillar dates T^p_j and T^p_{j+1} with corresponding zero rates z_j and z_{j+1}, we have that

$$P^M(0,T_l) = \exp\left(-\left(z_j + \frac{z_{j+1}-z_j}{T^p_{j+1}-T^p_j}\left(T_l - T^p_j\right)\right)(T_l - T_0)\right).$$

Then, if we start with assuming $T_k < T^p_{j-1} \leq T_i < T^p_j$. In this case we suppose that $T^p_f \leq T_k < T^p_{f+1}$ for some $T^p_f \in \{T^p_0, ..., T^p_{j-2}\}$. Then

$$\begin{aligned}\frac{P^M(0,T_i)}{P^M(0,T_k)} &= \frac{\exp\left(-z_{j-1}(T_i-T_0) - \frac{z_j-z_{j-1}}{T^p_j-T^p_{j-1}}\left(T_i-T^p_{j-1}\right)(T_i-T_0)\right)}{\exp\left(-z_f(T_k-T_0) - \frac{z_{f+1}-z_f}{T^p_{f+1}-T^p_f}\left(T_k-T^p_f\right)(T_k-T_0)\right)} \\ &= \exp\left(-z_{j-1}(T_i-T_0) + z_f(T_k-T_0) - \frac{z_j-z_{j-1}}{T^p_j-T^p_{j-1}}(T_i-T^p_{j-1})(T_i-T_0) + \frac{z_{f+1}-z_f}{T^p_{f+1}-T^p_f}(T_k-T^p_f)(T_k-T_0)\right),\end{aligned}$$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right) = \left(-\frac{(T_i-T^p_{j-1})(T_i-T_0)}{T^p_j-T^p_{j-1}}\right)\frac{P^M(0,T_i)}{P^M(0,T_k)}.$$

Now if $T_k < T^p_{j-1} < T^p_j \leq T_i < T^p_{j+1}$, we again suppose that $T^p_f \leq T_k < T^p_{f+1}$ for some $T^p_f \in \{T^p_0, ..., T^p_{j-2}\}$. Then

$$\frac{P^M(0,T_i)}{P^M(0,T_k)} = \frac{\exp\left(-z_j(T_i-T_0) - \frac{z_{j+1}-z_j}{T^p_{j+1}-T^p_j}\left(T_i-T^p_j\right)(T_i-T_0)\right)}{\exp\left(-z_f(T_k-T_0) - \frac{z_{f+1}-z_f}{T^p_{f+1}-T^p_f}\left(T_k-T^p_f\right)(T_k-T_0)\right)}$$

$$= \exp\left(-z_j(T_i-T_0)+z_f(T_k-T_0)-\frac{z_{j+1}-z_j}{T^p_{j+1}-T^p_j}(T_i-T^p_j)(T_i-T_0)+\frac{z_{f+1}-z_f}{T^p_{f+1}-T^p_f}(T_k-T^p_f)(T_k-T_0)\right),$$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right)=\left(-(T_i-T_0)+\frac{(T_i-T^p_j)(T_i-T_0)}{T^p_{j+1}-T^p_j}\right)\frac{P^M(0,T_i)}{P^M(0,T_k)}.$$

Now suppose $T^p_{j-1}\le T_k<T_i<T^p_j$, then

$$\frac{P^M(0,T_i)}{P^M(0,T_k)}=\frac{\exp\left(-z_{j-1}(T_i-T_0)-\frac{z_j-z_{j-1}}{T^p_j-T^p_{j-1}}\left(T_i-T^p_{j-1}\right)(T_i-T_0)\right)}{\exp\left(-z_{j-1}(T_k-T_0)-\frac{z_j-z_{j-1}}{T^p_j-T^p_{j-1}}\left(T_k-T^p_{j-1}\right)(T_k-T_0)\right)}$$
$$=\exp\left(-z_{j-1}(T_i-T_k)-\frac{z_j-z_{j-1}}{T^p_j-T^p_{j-1}}\left((T_i-T^p_{j-1})(T_i-T_0)-((T_k-T^p_{j-1})(T_k-T_0)\right)\right),$$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right)=\left(-\frac{(T_i-T^p_{j-1})(T_i-T_0)-(T_k-T^p_{j-1})(T_k-T_0)}{T^p_j-T^p_{j-1}}\right)\frac{P^M(0,T_i)}{P^M(0,T_k)}.$$

Now assume that $T^p_{j-1}\le T_k<T^p_j\le T_i<T^p_{j+1}$. Then

$$\frac{P^M(0,T_i)}{P^M(0,T_k)}=\frac{\exp\left(-z_j(T_i-T_0)-\frac{z_{j+1}-z_j}{T^p_{j+1}-T^p_j}\left(T_i-T^p_j\right)(T_i-T_0)\right)}{\exp\left(-z_{j-1}(T_k-T_0)-\frac{z_j-z_{j-1}}{T^p_j-T^p_{j-1}}\left(T_k-T^p_{j-1}\right)(T_k-T_0)\right)}$$
$$=\exp\left(-z_j(T_i-T_0)+z_{j-1}(T_k-T_0)-\frac{z_{j+1}-z_j}{T^p_{j+1}-T^p_j}(T_i-T^p_j)(T_i-T_0)+\frac{z_j-z_{j-1}}{T^p_j-T^p_{j-1}}(T_k-T^p_{j-1})(T_k-T_0)\right),$$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right)=\left(-(T_i-T_0)+\frac{(T_i-T^p_j)(T_i-T_0)}{T^p_{j+1}-T^p_j}+\frac{(T_k-T^p_{j-1})(T_k-T_0)}{T^p_j-T^p_{j-1}}\right)\frac{P^M(0,T_i)}{P^M(0,T_k)}.$$

Now let $T^p_{j-1}\le T_k<T^p_j<T^p_{j+1}\le T_i$. We suppose that $T^p_f\le T_i<T^p_{f+1}$ for some $T^p_f\in\{T^p_{j+1},...,T^p_{d-1}\}$. Then

$$\frac{P^M(0,T_i)}{P^M(0,T_k)}=\frac{\exp\left(-z_f(T_i-T_0)-\frac{z_{f+1}-z_f}{T^p_{f+1}-T^p_f}\left(T_i-T^p_f\right)(T_i-T_0)\right)}{\exp\left(-z_{j-1}(T_k-T_0)-\frac{z_j-z_{j-1}}{T^p_j-T^p_{j-1}}\left(T_k-T^p_{j-1}\right)(T_k-T_0)\right)}$$
$$=\exp\left(-z_f(T_i-T_0)+z_{j-1}(T_k-T_0)-\frac{z_{f+1}-z_f}{T^p_{f+1}-T^p_f}(T_i-T^p_f)(T_i-T_0)+\frac{z_j-z_{j-1}}{T^p_j-T^p_{j-1}}(T_k-T^p_{j-1})(T_k-T_0)\right),$$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right)=\left(\frac{(T_k-T^p_{j-1})(T_k-T_0)}{T^p_j-T^p_{j-1}}\right)\frac{P^M(0,T_i)}{P^M(0,T_k)}.$$

Now let $T^p_j\le T_k<T_i<T^p_{j+1}$. Then

$$\frac{P^M(0,T_i)}{P^M(0,T_k)}=\frac{\exp\left(-z_j(T_i-T_0)-\frac{z_{j+1}-z_j}{T^p_{j+1}-T^p_j}\left(T_i-T^p_j\right)(T_i-T_0)\right)}{\exp\left(-z_j(T_k-T_0)-\frac{z_{j+1}-z_j}{T^p_{j+1}-T^p_j}\left(T_k-T^p_j\right)(T_k-T_0)\right)}$$
$$=\exp\left(-z_j(T_i-T_k)-\frac{z_{j+1}-z_j}{T^p_{j+1}-T^p_j}\left((T_i-T^p_j)(T_i-T_0)-(T_k-T^p_j)(T_k-T_0)\right)\right),$$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right) = \left(-(T_i - T_k) + \frac{(T_i - T_j^p)(T_i - T_0) - (T_k - T_j^p)(T_k - T_0)}{T_{j+1}^p - T_j^p}\right)\frac{P^M(0,T_i)}{P^M(0,T_k)}.$$

Lastly, let $T_j^p \leq T_k < T_{j+1}^p \leq T_i$. We suppose that $T_f^p \leq T_i < T_{f+1}^p$ for some $T_f^p \in \{T_{j+1}^p, ..., T_{d-1}^p\}$. Then

$$\begin{aligned}
\frac{P^M(0,T_i)}{P^M(0,T_k)} &= \frac{\exp\left(-z_f(T_i - T_0) - \frac{z_{f+1}-z_f}{T_{f+1}^p - T_f^p}\left(T_i - T_f^p\right)(T_i - T_0)\right)}{\exp\left(-z_j(T_k - T_0) - \frac{z_{j+1}-z_j}{T_{j+1}^p - T_j^p}\left(T_k - T_j^p\right)(T_k - T_0)\right)} \\
&= \exp\left(-z_f(T_i - T_0) + z_j(T_k - T_0) - \frac{z_{f+1}-z_f}{T_{f+1}^p - T_f^p}(T_i - T_f^p)(T_i - T_0) + \frac{z_{j+1}-z_j}{T_{j+1}^p - T_j^p}(T_k - T_j^p)(T_k - T_0)\right),
\end{aligned}$$

$$\Rightarrow \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right) = \left(T_k - T_0 - \frac{(T_k - T_j^p)(T_k - T_0)}{T_{j+1}^p - T_j^p}\right)\frac{P^M(0,T_i)}{P^M(0,T_k)}.$$

□

Theorem A.3.3. *Let $q+1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T_0^p, T_1^p, ..., T_q^p$. Assume $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$. Lastly we assume $T_k \in \{T_1, ..., T_m\}$, and $T_i > T_k$. If we let $P(T_k, T_i)$ be given, as in Theorem 2.23, by*

$$P(T_k,T_i) = \frac{P^M(0,T_i)}{P^M(0,T_k)}\exp\left(B(T_k,T_i)(f^M(0,T_k) - r(T_k)) - \frac{\sigma^2}{4a}(1 - e^{-2aT_k})B(T_k,T_i)^2\right),$$

then

$$\frac{\partial}{\partial\theta}P(T_k,T_i) = P(T_k,T_i)\left(C(T_k,T_i) + B(T_k,T_i)D(T_k)\right),$$

where $C(T_k,T_i)$ is given as in Theorem 5.1.3 and $D(T_k)$ is given by

$$D(T_k) = \begin{cases} \frac{2T_k - T_{j-1}^p - T_0}{T_j^p - T_{j-1}^p} & \text{for } T_{j-1}^p \leq T_k < T_j^p, \\ 1 - \frac{2T_k - T_j^p - T_0}{T_{j+1}^p - T_j^p} & \text{for } T_j^p \leq T_k < T_{j+1}^p. \end{cases}$$

Proof. Applying the product rule yields

$$\begin{aligned}
\frac{\partial}{\partial\theta}P(T_k,T_i) &= \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\exp\left(B(T_k,T_i)(f^M(0,T_k) - r(T_k)) - \frac{\sigma^2}{4a}(1 - e^{-2aT_k})B(T_k,T_i)^2\right)\right) \\
&= \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right)\exp\left(B(T_k,T_i)(f^M(0,T_k) - r(T_k)) - \frac{\sigma^2}{4a}(1 - e^{-2aT_k})B(T_k,T_i)^2\right) \\
&\quad + \frac{P^M(0,T_i)}{P^M(0,T_k)}\frac{\partial}{\partial\theta}\left(\exp\left(B(T_k,T_i)(f^M(0,T_k) - r(T_k)) - \frac{\sigma^2}{4a}(1 - e^{-2aT_k})B(T_k,T_i)^2\right)\right).
\end{aligned}$$

The latter derivative can easily be determined. If we call

$$\begin{aligned}
h(T_k,T_i) &= \exp\left(B(T_k,T_i)(f^M(0,T_k) - r(T_k)) - \frac{\sigma^2}{4a}(1 - e^{-2aT_k})B(T_k,T_i)^2\right) \\
&= \exp\left(\tilde{h}(T_k,T_i)\right),
\end{aligned}$$

then

$$\begin{aligned}
\frac{\partial}{\partial\theta}h(T_k,T_i) &= h(T_k,T_i)\frac{\partial}{\partial\theta}\tilde{h}(T_k,T_i) \\
&= h(T_k,T_i)B(T_k,T_i)\frac{\partial}{\partial\theta}\left(f^M(0,T_k)\right).
\end{aligned}$$

As we proved in the proof of Theorem A.3.1, we see that

$$\frac{\partial}{\partial\theta}\left(f^M(0,T_k)\right) = D(T_k) = \begin{cases} \frac{2T_k - T^p_{j-1} - T_0}{T^p_j - T^p_{j-1}} & \text{for } T^p_{j-1} \le T_k < T^p_j, \\ 1 - \frac{2T_k - T^p_j - T_0}{T^p_{j+1} - T^p_j} & \text{for } T^p_j \le T_k < T^p_{j+1}. \end{cases}$$

Thus we conclude that

$$\frac{\partial}{\partial\theta}(h(T_k,T_i)) = B(T_k,T_i)D(T_k)\,\text{exp}\left(B(T_k,T_i)(f^M(0,T_k) - r(T_k)) - \frac{\sigma^2}{4a}(1-e^{-2aT_k})B(T_k,T_i)^2\right).$$

But then using Theorem 5.1.3

$$\begin{aligned}\frac{\partial}{\partial\theta}P(T_k,T_i) &= \frac{\partial}{\partial\theta}\left(\frac{P^M(0,T_i)}{P^M(0,T_k)}\right)\text{exp}\left(B(T_k,T_i)(f^M(0,T_k) - r(T_k)) - \frac{\sigma^2}{4a}(1-e^{-2aT_k})B(T_k,T_i)^2\right), \\ &+ \frac{P^M(0,T_i)}{P^M(0,T_k)}\frac{\partial}{\partial\theta}\left(\text{exp}\left(B(T_k,T_i)(f^M(0,T_k) - r(T_k)) - \frac{\sigma^2}{4a}(1-e^{-2aT_k})B(T_k,T_i)^2\right)\right), \\ &= P(T_k,T_i)\left(C(T_k,T_i) + B(T_k,T_i)D(T_k)\right).\end{aligned}$$

□

Theorem A.3.4. *Let $q+1$ be the number of zero rates our model uses as input. We label these zero rates as $z_0, z_1, ..., z_q$, with corresponding pillar dates $T^p_0, T^p_1, ..., T^p_q$. Assume $\theta = z_j$ for some $z_j \in \{z_1, ..., z_q\}$ and let $T_k \in \{T_1, ..., T_m\}$. If we let $V_\theta(T_k)$ equal the value of the IR swap at time T_k* (2.45)*, and we denote the future payment dates by $\{T_{j+1}, ..., T_n\}$, then*

$$\frac{\partial V_\theta(T_k)}{\partial\theta} = \bar{\beta}Q\left(\frac{\partial P(T_k,T_j)}{\partial\theta} - \frac{\partial P(T_k,T_n)}{\partial\theta}\right) - \bar{\beta}QK\sum_{i=j+1}^{n}\tau_i\frac{\partial P(T_k,T_i)}{\partial\theta},$$

where for $k = 1, ..., m$ and $i = j, ..., n$ the derivative $\frac{\partial P(T_k,T_i)}{\partial\theta}$ is given as in Theorem 5.1.4.

Proof. The payoff of a swap at time T_k is given by

$$\begin{aligned}V_\theta(T_k) &= \bar{\beta}Q\sum_{i=j+1}^{n}\tau_i P(T_k,T_i)\left(\ell_i(T_k;T_{i-1},T_i) - K\right) \\ &= \bar{\beta}Q(P(T_k,T_j) - P(T_k,T_n)) - \bar{\beta}QK\sum_{i=j+1}^{n}\tau_i P(T_k,T_i).\end{aligned}$$

We then get that

$$\begin{aligned}\frac{\partial V_\theta(T_k)}{\partial\theta} &= \frac{\partial}{\partial\theta}\left(\bar{\beta}Q(P(T_k,T_j) - P(T_k,T_n)) - \bar{\beta}QK\sum_{i=j+1}^{n}\tau_i P(T_k,T_i)\right) \\ &= \bar{\beta}Q\left(\frac{\partial P(T_k,T_j)}{\partial\theta} - \frac{\partial P(T_k,T_n)}{\partial\theta}\right) - \bar{\beta}QK\sum_{i=j+1}^{n}\tau_i\frac{\partial P(T_k,T_i)}{\partial\theta}.\end{aligned}$$

□

A.4. Proofs CVA FX spot delta estimation

Theorem A.4.1. *For any $0 \leq T_k, T_l$, the covariance of $r_d(T_k)$ and $r_f(T_l)$ given $\mathcal{F}(0)$ is given by*

$$\text{Cov}_0\left[r_d(T_k), r_f(T_l)\right] = \begin{cases} \frac{\sigma_d\sigma_f}{a_d+a_f}\rho_{d,f}\left(e^{-a_f(T_l-T_k)} - e^{-a_dT_k-a_fT_l}\right) & T_k \leq T_l, \\ \frac{\sigma_d\sigma_f}{a_d+a_f}\rho_{d,f}\left(e^{-a_d(T_k-T_l)} - e^{-a_fT_l-a_dT_k}\right) & T_k > T_l. \end{cases}$$

Proof. We start by assuming $0 \leq T_k \leq T_l$. We rewrite the covariance as follows.

$$\begin{aligned} \text{Cov}_0\left[r_d(T_k), r_f(T_l)\right] &= \mathbb{E}_0\left[(r_d(T_k) - \mathbb{E}_0(r_d(T_k)))(r_d(T_l) - \mathbb{E}_0(r_d(T_l)))\right] \\ &= \mathbb{E}_0\left[\int_0^{T_k} \sigma_d e^{-a_d(T_k-u)}dW_d^{\mathbb{Q}}(u) \int_0^{T_l} \sigma_f e^{-a_f(T_l-u)}dW_f^{\mathbb{Q}}(u)\right]. \end{aligned}$$

Note that we can split the integrals with bounds 0 and T_l into two integrals, one ranging from 0 to T_k, and one ranging from T_k to T_l. In a similar way as we did in earlier proofs, we can neglect the integrals ranging from T_k to T_l, since they are independent of the integrals ranging from 0 to T_k. Using this we get

$$\begin{aligned} \text{Cov}_0\left[r_d(T_k), r_f(T_l)\right] &= \sigma_d\sigma_f\rho_{d,f}\int_0^{T_k} e^{-a_d(T_k-u)-a_f(T_l-u)}du \\ &= \frac{\sigma_d\sigma_f}{a_d+a_f}\rho_{d,f}\left(e^{-a_f(T_l-T_k)} - e^{-a_dT_k-a_fT_l}\right). \end{aligned}$$

In a similar way, we find for $0 \leq T_l < T_k$ that

$$\begin{aligned} \text{Cov}_0\left[r_d(T_k), r_f(T_l)\right] &= \sigma_d\sigma_f\rho_{d,f}\int_0^{T_l} e^{-a_f(T_l-u)-a_d(T_k-u)}du \\ &= \frac{\sigma_d\sigma_f}{a_d+a_f}\rho_{d,f}\left(e^{-a_d(T_k-T_l)} - e^{-a_fT_l-a_dT_k}\right). \end{aligned}$$

□

Theorem A.4.2. *For any $0 \leq T_k, T_l$, the covariance of $r_d(T_k)$ and $L_y(T_l)$ given $\mathcal{F}(0)$ is given by*

$$\text{Cov}_0\left[r_d(T_k), L_y(T_l)\right] = \begin{cases} \begin{array}{l} \frac{\sigma_d^2}{a_d}\left(\frac{1-e^{-a_dT_k}}{a_d} - \frac{e^{-a_d(T_l-T_k)}-e^{a_d(T_k+T_l)}}{2a_d}\right) \\ \quad -\frac{\sigma_d\sigma_f}{a_f}\rho_{d,f}\left(\frac{1-e^{-a_dT_k}}{a_d} - \frac{e^{-a_f(T_l-T_k)}-e^{-a_dT_k-a_fT_l}}{a_d+a_f}\right) \\ \quad +\sigma_d\sigma_y\rho_{y,d}\frac{1-e^{-a_dT_k}}{a_d} \end{array} & T_k \leq T_l, \\ \begin{array}{l} \frac{\sigma_d^2}{a_d}\left(\frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} - \frac{e^{-a_d(T_k-T_l)}-e^{a_d(T_l+T_k)}}{2a_d}\right) \\ \quad -\frac{\sigma_d\sigma_f}{a_f}\rho_{d,f}\left(\frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} - \frac{e^{-a_d(T_k-T_l)}-e^{-a_fT_l-a_dT_k}}{a_d+a_f}\right) \\ \quad +\sigma_d\sigma_y\rho_{y,d}\frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} \end{array} & T_k > T_l. \end{cases}$$

Proof. We start by assuming $0 \leq T_k \leq T_l$. We rewrite the covariance as follows.

$$\begin{aligned} \text{Cov}_0\left[r_d(T_k), L_y(T_l)\right] &= \mathbb{E}_0\left[(r_d(T_k) - \mathbb{E}_0(r_d(T_k)))(L_y(T_l) - \mathbb{E}_0(L_y(T_l)))\right] \\ &= \mathbb{E}_0\Bigg[\int_0^{T_k} \sigma_d e^{-a_d(T_k-u)}dW_d^{\mathbb{Q}}(u)\Bigg(\int_0^{T_l} \frac{\sigma_d}{a_d}\left(1 - e^{-a_d(T_l-u)}\right)dW_d^{\mathbb{Q}}(u) \\ &\quad - \int_0^{T_l} \frac{\sigma_f}{a_f}\left(1 - e^{-a_f(T_l-u)}\right)dW_f^{\mathbb{Q}}(u) + \int_0^{T_l} \sigma_y dW_y^{\mathbb{Q}}(u)\Bigg)\Bigg]. \end{aligned}$$

Now we can take the products of all the integrals. Note that we can split the integrals with bounds 0 and T_l into two integrals, one ranging from 0 to T_k, and one ranging from T_k to T_l. In a similar way as we did in earlier proofs, we can neglect the integrals ranging from T_k to T_l, since they are independent of the integrals ranging from 0 to T_k. Using this we get

$$
\begin{aligned}
\mathrm{Cov}_0\left[r_d(T_k), L_y(T_l)\right] &= \frac{\sigma_d^2}{a_d}\int_0^{T_k} e^{-a_d(T_k-u)}\left(1-e^{-a_d(T_l-u)}\right)du \\
&\quad - \frac{\sigma_d\sigma_f}{a_f}\rho_{d,f}\int_0^{T_k} e^{-a_d(T_k-u)}\left(1-e^{-a_f(T_l-u)}\right)du + \sigma_d\sigma_y\rho_{y,d}\int_0^{T_k} e^{-a_d(T_k-u)}du \\
&= \frac{\sigma_d^2}{a_d}\left(\frac{1-e^{-a_dT_k}}{a_d} - \frac{e^{-a_d(T_l-T_k)}-e^{a_d(T_k+T_l)}}{2a_d}\right) \\
&\quad - \frac{\sigma_d\sigma_f}{a_f}\rho_{d,f}\left(\frac{1-e^{-a_dT_k}}{a_d} - \frac{e^{-a_f(T_l-T_k)}-e^{-a_dT_k-a_fT_l}}{a_d+a_f}\right) + \sigma_d\sigma_y\rho_{y,d}\frac{1-e^{-a_dT_k}}{a_d}.
\end{aligned}
$$

In a similar way, we find for $0 \leq T_l < T_k$ that

$$
\begin{aligned}
\mathrm{Cov}_0\left[r_d(T_k), L_y(T_l)\right] &= \frac{\sigma_d^2}{a_d}\int_0^{T_l} e^{-a_d(T_k-u)}\left(1-e^{-a_d(T_l-u)}\right)du \\
&\quad - \frac{\sigma_d\sigma_f}{a_f}\rho_{d,f}\int_0^{T_l} e^{-a_d(T_k-u)}\left(1-e^{-a_f(T_l-u)}\right)du + \sigma_d\sigma_y\rho_{y,d}\int_0^{T_l} e^{-a_d(T_k-u)}du \\
&= \frac{\sigma_d^2}{a_d}\left(\frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} - \frac{e^{-a_d(T_k-T_l)}-e^{a_d(T_l+T_k)}}{2a_d}\right) \\
&\quad - \frac{\sigma_d\sigma_f}{a_f}\rho_{d,f}\left(\frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} - \frac{e^{-a_d(T_k-T_l)}-e^{-a_fT_l-a_dT_k}}{a_d+a_f}\right) \\
&\quad + \sigma_d\sigma_y\rho_{y,d}\frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d}.
\end{aligned}
$$

□

Theorem A.4.3. *For any $0 \leq T_k, T_l$, the covariance of $r_f(T_k)$ and $\int_0^{T_l} r_d(u)du$ given $\mathcal{F}(0)$ is given by*

$$
\mathrm{Cov}_0\left[r_f(T_k), \int_0^{T_l} r_d(u)du\right] = \begin{cases} \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{1-e^{-a_fT_k}}{a_f} - \frac{e^{-a_d(T_l-T_k)}-e^{-a_fT_k-a_dT_l}}{a_d+a_f}\right) & T_k \leq T_l, \\ \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_k-T_l)}-e^{-a_dT_l-a_fT_k}}{a_d+a_f}\right) & T_k > T_l. \end{cases}
$$

Proof. We start by assuming $0 \leq T_k \leq T_l$. We rewrite the covariance as follows.

$$
\begin{aligned}
\mathrm{Cov}_0\left[r_f(T_k), \int_0^{T_l} r_d(u)du\right] &= \mathbb{E}_0\left[\left(r_f(T_k) - \mathbb{E}_0(r_f(T_k))\right)\left(\int_0^{T_l} r_d(u)du - \mathbb{E}_0\left(\int_0^{T_l} r_d(u)du\right)\right)\right] \\
&= \mathbb{E}_0\left[\int_0^{T_k}\sigma_f e^{-a_f(T_k-u)}dW_f^{\mathbb{Q}}(u)\int_0^{T_l}\frac{\sigma_d}{a_d}\left(1-e^{-a_d(T_l-u)}\right)dW_d^{\mathbb{Q}}(u)\right].
\end{aligned}
$$

Note that we can split the integrals with bounds 0 and T_l into two integrals, one ranging from 0 to T_k, and one ranging from T_k to T_l. In a similar way as we did in earlier proofs, we can neglect the integrals ranging from T_k to T_l, since they are independent of the integrals ranging from 0 to T_k. Using this we get

$$
\begin{aligned}
\mathrm{Cov}_0\left[r_f(T_k), \int_0^{T_l} r_d(u)du\right] &= \frac{\sigma_f\sigma_d}{a_d}\rho_{d,f}\int_0^{T_k} e^{-a_f(T_k-u)}\left(1-e^{-a_d(T_l-u)}\right)du \\
&= \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{1-e^{-a_fT_k}}{a_f} - \frac{e^{-a_d(T_l-T_k)}-e^{-a_fT_k-a_dT_l}}{a_d+a_f}\right).
\end{aligned}
$$

In a similar way, we find for $0 \leq T_l < T_k$ that

$$
\begin{aligned}
\mathrm{Cov}_0\left[r_f(T_k), \int_0^{T_l} r_d(u)du\right] &= \frac{\sigma_f\sigma_d}{a_d}\rho_{d,f}\int_0^{T_l} e^{-a_f(T_k-u)}\left(1-e^{-a_d(T_l-u)}\right)du \\
&= \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_k-T_l)}-e^{-a_dT_l-a_fT_k}}{a_d+a_f}\right).
\end{aligned}
$$

□

Theorem A.4.4. *For any $0 \leq T_k, T_l$, the covariance of $r_f(T_k)$ and $L_y(T_l)$ given $\mathcal{F}(0)$ is given by*

$$\text{Cov}_0\left[r_f(T_k), L_y(T_l)\right] = \begin{cases} \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{1-e^{-a_fT_k}}{a_f} - \frac{e^{-a_d(T_l-T_k)}-e^{a_fT_k-a_dT_l}}{a_d+a_f}\right) \\ \quad -\frac{\sigma_f^2}{a_f}\left(\frac{1-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_l-T_k)}-e^{-a_f(T_k+T_l)}}{2a_f}\right) & T_k \leq T_l, \\ \quad +\sigma_f\sigma_y\rho_{y,f}\frac{1-e^{-a_fT_k}}{a_f} \\ \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_k-T_l)}-e^{a_dT_l-a_fT_k}}{a_d+a_f}\right) \\ \quad -\frac{\sigma_f^2}{a_f}\left(\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_k-T_l)}-e^{-a_f(T_l+T_k)}}{2a_f}\right) & T_k > T_l. \\ \quad +\sigma_f\sigma_y\rho_{y,f}\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} \end{cases}$$

Proof. We start by assuming $0 \leq T_k \leq T_l$. We rewrite the covariance as follows.

$$\begin{aligned} \text{Cov}_0\left[r_f(T_k), L_y(T_l)\right] &= \mathbb{E}_0\left[(r_f(T_k) - \mathbb{E}_0(r_f(T_k)))(L_y(T_l) - \mathbb{E}_0(L_y(T_l)))\right] \\ &= \mathbb{E}_0\Bigg[\int_0^{T_k}\sigma_f e^{-a_f(T_k-u)}dW_f^{\mathbb{Q}}(u)\Bigg(\int_0^{T_l}\frac{\sigma_d}{a_d}\left(1-e^{-a_d(T_l-u)}\right)dW_d^{\mathbb{Q}}(u) \\ &\quad - \int_0^{T_l}\frac{\sigma_f}{a_f}\left(1-e^{-a_f(T_l-u)}\right)dW_f^{\mathbb{Q}}(u) + \int_0^{T_l}\sigma_y dW_y^{\mathbb{Q}}(u)\Bigg)\Bigg]. \end{aligned}$$

Now we can take the products of all the integrals. Note that we can split the integrals with bounds 0 and T_l into two integrals, one ranging from 0 to T_k, and one ranging from T_k to T_l. In a similar way as we did in earlier proofs, we can neglect the integrals ranging from T_k to T_l, since they are independent of the integrals ranging from 0 to T_k. Using this we get

$$\begin{aligned} \text{Cov}_0\left[r_f(T_k), L_y(T_l)\right] &= \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\int_0^{T_k}e^{-a_f(T_k-u)}\left(1-e^{-a_d(T_l-u)}\right)du \\ &\quad - \frac{\sigma_f^2}{a_f}\int_0^{T_k}e^{-a_f(T_k-u)}\left(1-e^{-a_f(T_l-u)}\right)du + \sigma_f\sigma_y\rho_{y,f}\int_0^{T_k}e^{-a_f(T_k-u)}du \\ &= \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{1-e^{-a_fT_k}}{a_f} - \frac{e^{-a_d(T_l-T_k)}-e^{a_fT_k-a_dT_l}}{a_d+a_f}\right) \\ &\quad - \frac{\sigma_f^2}{a_f}\left(\frac{1-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_l-T_k)}-e^{-a_f(T_k+T_l)}}{2a_f}\right) + \sigma_f\sigma_y\rho_{y,f}\frac{1-e^{-a_fT_k}}{a_f}. \end{aligned}$$

In a similar way, we find for $0 \leq T_l < T_k$ that

$$\begin{aligned} \text{Cov}_0\left[r_f(T_k), L_y(T_l)\right] &= \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\int_0^{T_l}e^{-a_f(T_k-u)}\left(1-e^{-a_d(T_l-u)}\right)du \\ &\quad - \frac{\sigma_f^2}{a_f}\int_0^{T_l}e^{-a_f(T_k-u)}\left(1-e^{-a_f(T_l-u)}\right)du + \sigma_f\sigma_y\rho_{y,f}\int_0^{T_l}e^{-a_f(T_k-u)}du \\ &= \frac{\sigma_d\sigma_f}{a_d}\rho_{d,f}\left(\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_k-T_l)}-e^{a_dT_l-a_fT_k}}{a_d+a_f}\right) \\ &\quad - \frac{\sigma_f^2}{a_f}\left(\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f} - \frac{e^{-a_f(T_k-T_l)}-e^{-a_f(T_l+T_k)}}{2a_f}\right) \\ &\quad + \sigma_f\sigma_y\rho_{y,f}\frac{e^{-a_f(T_k-T_l)}-e^{-a_fT_k}}{a_f}. \end{aligned}$$

$\square$

Theorem A.4.5. *For any $0 \le T_k, T_l$, the covariance of $\int_0^{T_k} r_d(u)du$ and $L_y(T_l)$ given $\mathcal{F}(0)$ is given by*

$$\text{Cov}\left[\int_0^{T_k} r_d(s)ds, L_y(T_l)\right] = \begin{cases} \frac{\sigma_d^2}{a_d^2}\left(T_k - \frac{1-e^{-a_dT_k}+e^{-a_d(T_l-T_k)}-e^{-a_dT_l}}{a_d} + \frac{e^{-a_d(T_l-T_k)}-e^{-a_d(T_k+T_l)}}{2a_d}\right) & \\ -\frac{\sigma_d\sigma_f}{a_da_f}\rho_{d,f}\Big(T_k - \frac{1-e^{-a_dT_k}}{a_d} - \frac{e^{-a_f(T_l-T_k)}-e^{-a_fT_l}}{a_f} & T_k \le T_l, \\ +\frac{e^{-a_f(T_l-T_k)}-e^{-a_dT_k-a_fT_l}}{a_d+a_f}\Big) + \frac{\sigma_d\sigma_y\rho_{y,d}}{a_d}\left(T_k - \frac{1-e^{-a_dT_k}}{a_d}\right) & \\ \frac{\sigma_d^2}{a_d^2}\left(T_l - \frac{1-e^{-a_dT_l}+e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} + \frac{e^{-a_d(T_k-T_l)}-e^{-a_d(T_l+T_k)}}{2a_d}\right) & \\ -\frac{\sigma_d\sigma_f}{a_da_f}\rho_{d,f}\Big(T_l - \frac{1-e^{-a_fT_l}}{a_f} - \frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} & T_k \le T_l. \\ +\frac{e^{-a_d(T_k-T_l)}-e^{-a_fT_l-a_dT_k}}{a_d+a_f}\Big) + \frac{\sigma_d\sigma_y\rho_{y,d}}{a_d}\left(T_l - \frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d}\right) & \end{cases}$$

Proof. We start by assuming $0 \le T_k \le T_l$. We rewrite the covariance as follows.

$$\begin{aligned} \text{Cov}_0\left[\int_0^{T_k} r_d(u)du, L_y(T_l)\right] &= \mathbb{E}_0\left[\left(\left(\int_0^{T_k} r_d(u)du - \mathbb{E}_0\left(\int_0^{T_k} r_d(u)du\right)\right)(L_y(T_l) - \mathbb{E}_0(L_y(T_l)))\right)\right] \\ &= \mathbb{E}_0\Bigg[\int_0^{T_k} \frac{\sigma_d}{a_d}\left(1-e^{-a_d(T_k-u)}\right)dW_d^{\mathbb{Q}}(u)\Bigg(\int_0^{T_l} \frac{\sigma_d}{a_d}\left(1-e^{-a_d(T_l-u)}\right)dW_d^{\mathbb{Q}}(u) \\ &\quad - \int_0^{T_l} \frac{\sigma_f}{a_f}\left(1-e^{-a_f(T_l-u)}\right)dW_f^{\mathbb{Q}}(u) + \int_0^{T_l} \sigma_y dW_y^{\mathbb{Q}}(u)\Bigg)\Bigg]. \end{aligned}$$

Now we can take the products of all the integrals. Note that we can split the integrals with bounds 0 and T_l into two integrals, one ranging from 0 to T_k, and one ranging from T_k to T_l. In a similar way as we did in earlier proofs, we can neglect the integrals ranging from T_k to T_l, since they are independent of the integrals ranging from 0 to T_k. Using this we get

$$\begin{aligned} \text{Cov}_0\left[r_f(T_k), L_y(T_l)\right] &= \int_0^{T_k} \frac{\sigma_d^2}{a_d^2}\left(1-e^{-a_d(T_k-u)}\right)\left(1-e^{-a_d(T_l-u)}\right)du \\ &\quad - \frac{\sigma_d\sigma_f}{a_da_f}\rho_{d,f}\int_0^{T_k}\left(1-e^{-a_d(T_k-u)}\right)\left(1-e^{-a_f(T_l-u)}\right)du + \frac{\sigma_d\sigma_y}{a_d}\rho_{y,d}\int_0^{T_k}\left(1-e^{-a_d(T_k-u)}\right)du \\ &= \frac{\sigma_d^2}{a_d^2}\left(T_k - \frac{1-e^{-a_dT_k}+e^{-a_d(T_l-T_k)}-e^{-a_dT_l}}{a_d} + \frac{e^{-a_d(T_l-T_k)}-e^{-a_d(T_k+T_l)}}{2a_d}\right) \\ &\quad - \frac{\sigma_d\sigma_f}{a_da_f}\rho_{d,f}\left(T_k - \frac{1-e^{-a_dT_k}}{a_d} - \frac{e^{-a_f(T_l-T_k)}-e^{-a_fT_l}}{a_f} + \frac{e^{-a_f(T_l-T_k)}-e^{-a_dT_k-a_fT_l}}{a_d+a_f}\right) \\ &\quad + \frac{\sigma_d\sigma_y\rho_{y,d}}{a_d}\left(T_k - \frac{1-e^{-adT_k}}{a_d}\right). \end{aligned}$$

In a similar way, we find for $0 \le T_l < T_k$ that

$$\begin{aligned} \text{Cov}_0\left[r_f(T_k), L_y(T_l)\right] &= \int_0^{T_l} \frac{\sigma_d^2}{a_d^2}\left(1-e^{-a_d(T_k-u)}\right)\left(1-e^{-a_d(T_l-u)}\right)du \\ &\quad - \frac{\sigma_d\sigma_f}{a_da_f}\rho_{d,f}\int_0^{T_l}\left(1-e^{-a_d(T_k-u)}\right)\left(1-e^{-a_f(T_l-u)}\right)du + \frac{\sigma_d\sigma_y}{a_d}\rho_{y,d}\int_0^{T_l}\left(1-e^{-a_d(T_k-u)}\right)du \\ &= \frac{\sigma_d^2}{a_d^2}\left(T_l - \frac{1-e^{-a_dT_l}+e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} + \frac{e^{-a_d(T_k-T_l)}-e^{-a_d(T_l+T_k)}}{2a_d}\right) \\ &\quad - \frac{\sigma_d\sigma_f}{a_da_f}\rho_{d,f}\Big(T_l - \frac{1-e^{-a_fT_l}}{a_f} - \frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d} + \frac{e^{-a_d(T_k-T_l)}-e^{-a_fT_l-a_dT_k}}{a_d+a_f}\Big) \\ &\quad + \frac{\sigma_d\sigma_y\rho_{y,d}}{a_d}\left(T_l - \frac{e^{-a_d(T_k-T_l)}-e^{-a_dT_k}}{a_d}\right). \end{aligned}$$

□

B

CPU times

In this section, we will show the CPU times of the experiments we conducted.

B.1. CPU times of CVA FX spot deltas of FX digital options

Method / T	1	2	3	4	5	7	10
LRM	3.20	9.88	9.39	16.27	16.69	41.29	73.82
LRM + QMC	3.75	8.49	12.61	16.58	22.80	33.23	68.41
B&R	3.58	7.68	11.76	17.56	21.65	34.68	54.63
B&R + QMC	4.49	9.51	14.82	19.94	27.49	37.36	58.76

Table B.1: CPU times (in sec) for the four methods for different maturities T.

Method / σ_y	0	0.005	0.01	0.02	0.05	0.1	0.2	0.3
LRM	11.29	10.78	10.51	10.84	9.31	10.60	9.13	11.09
LRM + QMC	12.00	13.60	11.93	14.03	11.84	11.96	11.74	12.37
B&R	12.89	12.23	12.21	12.46	11.87	12.26	11.84	12.07
B&R + QMC	14.47	15.10	14.46	14.50	14.39	14.20	14.18	14.18

Table B.2: CPU times (in sec) for the four methods for different volatilities of the FX process (σ_y).

Method / $\rho_{y,d}$	-1	-0.9	-0.8	-0.5	-0.25	-0.1	0	0.1	0.25	0.5	0.8	1
LRM	10.5	12.7	9.1	10.9	9.0	10.9	9.5	10.3	9.7	9.6	11.0	10.0
LRM + QMC	13.0	11.9	11.7	11.8	11.5	14.7	11.8	15.3	11.9	11.9	12.5	12.0
B&R	11.9	11.8	12.0	12.0	11.8	12.5	12.0	12.7	12.3	12.7	12.2	12.2
B&R + QMC	14.1	14.5	14.6	14.4	14.7	14.3	14.3	14.8	14.5	14.3	14.7	14.3

Table B.3: CPU times (in sec) for the four methods for different correlations between the FX process and the domestic currency ($\rho_{y,d}$).

Method / $\rho_{y,f}$	-1	-0.9	-0.8	-0.5	-0.25	-0.1	0	0.1	0.25	0.5	0.8	1
LRM	10.1	12.1	9.5	11.7	9.4	10.7	9.6	9.6	9.5	9.8	9.4	10.07
LRM + QMC	12.1	12.5	12.2	12.6	11.8	15.7	11.8	11.9	11.7	13.3	11.8	11.9
B&R	12.0	12.5	12.2	12.1	12.0	12.8	11.9	12.5	11.9	12.2	12.9	12.1
B&R + QMC	14.9	15.1	14.7	14.9	14.9	15.1	14.4	14.4	15.3	14.3	15.1	14.7

Table B.4: CPU times (in sec) for the four methods for different correlations between the FX process and the foreign currency ($\rho_{y,f}$).

Method / $\rho_{d,f}$	-0.85	-0.80	-0.5	-0.25	-0.1	0	0.1	0.25	0.5	0.80	0.85
LRM	10.69	9.93	10.04	9.18	9.93	9.42	10.30	9.28	9.56	9.38	9.43
LRM + QMC	12.85	12.22	12.25	11.59	12.62	11.83	11.64	11.86	12.94	11.71	12.94
B&R	12.65	12.18	12.30	12.08	12.78	11.85	12.52	11.86	11.95	11.79	13.69
B&R + QMC	15.32	14.92	14.64	14.65	14.59	14.52	14.33	14.47	14.30	14.59	14.72

Table B.5: CPU times (in sec) for the four methods for different correlations between the domestic and the foreign currency ($\rho_{d,f}$).

Method / a_d	0.005	0.01	0.015	0.02	0.03	0.05	0.08	0.15	0.2
LRM	9.81	10.53	9.19	9.37	9.19	10.51	9.45	9.67	9.17
LRM + QMC	12.89	13.63	11.65	14.00	11.73	13.74	11.78	12.46	11.71
B&R	12.81	12.05	11.87	12.85	11.93	12.64	12.15	11.98	11.82
B&R + QMC	15.29	14.79	14.23	14.30	14.61	14.53	14.78	14.36	16.21

Table B.6: CPU times (in sec) for the four methods for different mean reversion parameters of the domestic currency (a_d).

Method / σ_d	0.001	0.005	0.01	0.015	0.02	0.05	0.1	0.15	0.2
LRM	12.45	14.18	9.25	12.04	9.38	12.05	9.30	10.12	9.44
LRM + QMC	13.17	15.57	11.90	13.60	11.97	15.58	11.92	15.03	11.72
B&R	12.22	12.24	12.12	12.70	13.93	14.10	12.23	12.22	12.06
B&R + QMC	14.83	14.58	14.61	14.25	15.35	15.78	15.00	14.58	14.44

Table B.7: CPU times (in sec) for the four methods for different volatilities of the domestic currency (σ_d).

Method / a_f	0.005	0.01	0.015	0.02	0.03	0.04	0.08	0.15	0.2
LRM	10.33	15.28	9.35	9.60	9.17	9.84	9.39	10.04	9.43
LRM + QMC	12.13	12.05	11.82	12.33	11.69	12.03	11.94	13.40	11.64
B&R	12.11	13.04	11.99	12.63	11.98	12.25	11.89	12.75	11.89
B&R + QMC	14.33	14.78	14.32	14.34	14.88	14.17	14.43	14.36	14.21

Table B.8: CPU times (in sec) for the four methods for different mean reversion parameters of the foreign currency (a_f).

Method / σ_f	0.001	0.005	0.01	0.015	0.02	0.05	0.1	0.15	0.2
LRM	10.37	11.39	9.21	9.49	9.19	9.51	9.67	9.59	9.50
LRM + QMC	12.08	12.67	11.72	12.00	11.65	13.72	12.05	14.86	11.89
B&R	11.87	12.57	11.86	12.12	11.97	12.28	12.56	12.62	12.01
B&R + QMC	14.25	14.71	14.19	14.57	14.54	14.25	14.44	14.65	14.34

Table B.9: CPU times (in sec) for the four methods for different volatilities of the foreign currency (σ_f).

Method / K	0.7275	0.7933	0.8650	0.9433	1.0286	1.1217	1.2232
LRM	10.64	11.43	9.31	9.51	9.31	9.49	9.75
LRM + QMC	12.98	13.62	11.67	11.93	11.76	12.95	12.41
B&R	12.46	12.07	11.82	12.21	11.85	13.08	12.42
B&R + QMC	14.32	15.35	14.76	14.71	14.55	14.47	14.72

Table B.10: CPU times (in sec) for the four methods for different strike prices of the FX digital option (K).

B.2. CPU times of CVA FX spot deltas of FX swaps

Method / T	1	2	3	4	5	7	10
LRM	3.02	7.14	11.09	14.49	19.29	30.23	55.43
LRM + QMC	3.75	8.29	12.57	17.09	23.79	36.40	72.15
B&R	3.64	8.02	13.95	17.34	22.96	35.57	56.94

Table B.11: CPU times (in sec) for the three methods for different maturities T.

Method / σ_y	0	0.005	0.01	0.02	0.05	0.1	0.2	0.3
LRM	12.50	11.19	10.89	10.92	10.33	9.89	10.70	11.46
LRM + QMC	12.96	13.04	12.71	13.51	12.45	12.61	12.54	12.56
B&R	12.45	12.51	13.06	12.44	12.71	12.82	12.68	12.48

Table B.12: CPU times (in sec) for the three methods for different volatilities of the FX process (σ_y).

Method / $\rho_{y,d}$	-1	-0.9	-0.8	-0.5	-0.25	-0.1	0	0.1	0.25	0.5	0.8	1
LRM	11.1	12.3	10.8	12.1	11.6	11.1	12.4	11.5	10.6	10.6	9.9	10.7
LRM + QMC	12.9	14.1	12.9	12.6	12.5	12.7	13.8	13.1	12.8	12.6	12.5	12.8
B&R	12.7	12.6	12.5	12.6	12.4	12.7	12.8	12.7	12.7	12.5	12.5	12.6

Table B.13: CPU times (in sec) for the three methods for different correlations between the FX process and the domestic currency ($\rho_{y,d}$).

Method / $\rho_{y,f}$	-1	-0.9	-0.8	-0.5	-0.25	-0.1	0	0.1	0.25	0.5	0.8	1
LRM	11.3	11.8	10.6	10.8	9.9	10.8	11.4	11.4	11.3	11.6	11.1	11.0
LRM + QMC	12.8	13.5	12.8	12.7	12.7	13.7	12.9	12.3	13.6	12.7	14.2	12.8
B&R	12.7	14.0	12.5	12.9	12.6	12.7	12.9	12.6	13.4	13.0	12.6	12.7

Table B.14: CPU times (in sec) for the three methods for different correlations between the FX process and the foreign currency ($\rho_{y,f}$).

Method / $\rho_{d,f}$	-0.85	-0.80	-0.5	-0.25	-0.1	0	0.1	0.25	0.5	0.80	0.85
LRM	11.06	11.46	10.71	11.51	10.49	12.01	10.78	10.06	10.76	10.75	10.69
LRM + QMC	12.90	14.24	12.99	12.44	13.20	12.79	13.08	12.74	12.67	13.08	12.70
B&R	12.57	14.20	12.46	13.13	12.60	12.49	12.66	12.96	12.62	13.76	12.62

Table B.15: CPU times (in sec) for the three methods for different correlations between the domestic and the foreign currency ($\rho_{d,f}$).

Method / a_d	0.005	0.01	0.015	0.02	0.03	0.05	0.08	0.15	0.2
LRM	11.45	13.16	11.33	10.46	12.56	10.60	11.18	13.06	11.06
LRM + QMC	12.92	14.01	12.64	12.77	13.20	12.27	13.62	12.75	13.24
B&R	12.65	13.48	12.46	13.59	12.48	12.69	12.84	13.02	13.43

Table B.16: CPU times (in sec) for the three methods for different mean reversion parameters of the domestic currency (a_d).

Method / σ_d	0.001	0.005	0.01	0.015	0.02	0.05	0.1	0.15	0.2
LRM	11.12	11.16	11.26	11.64	10.32	11.23	10.33	10.69	11.61
LRM + QMC	12.98	13.57	13.40	12.66	12.55	12.78	12.94	12.91	12.68
B&R	13.62	13.00	12.82	12.62	12.69	12.65	12.85	12.56	12.47

Table B.17: CPU times (in sec) for the three methods for different volatilities of the domestic currency (σ_d).

Method / a_f	0.005	0.01	0.015	0.02	0.03	0.04	0.08	0.15	0.2
LRM	11.12	11.26	10.94	10.04	11.03	10.76	10.85	10.35	11.53
LRM + QMC	12.96	14.13	12.47	12.73	12.76	12.64	12.80	12.63	13.29
B&R	12.85	13.28	12.46	12.41	12.77	12.57	12.63	12.70	13.22

Table B.18: CPU times (in sec) for the three methods for different mean reversion parameters of the foreign currency (a_f).

Method / σ_f	0.001	0.005	0.01	0.015	0.02	0.05	0.1	0.15	0.2
LRM	10.46	12.22	10.82	9.53	9.76	9.27	10.36	10.34	10.54
LRM + QMC	12.32	12.67	12.49	11.94	11.89	11.85	11.97	13.13	12.91
B&R	11.84	13.05	12.05	12.08	12.09	11.86	12.11	13.21	12.41

Table B.19: CPU times (in sec) for the three methods for different volatilities of the foreign currency (σ_f).

Method / K	0.7275	0.7933	0.8650	0.9433	1.0286	1.1217	1.2232
LRM	10.98	11.16	9.92	10.15	10.75	9.95	10.72
LRM + QMC	12.26	12.71	12.21	12.10	12.09	11.98	12.73
B&R	12.33	12.39	11.94	11.97	12.18	12.28	12.09

Table B.20: CPU times (in sec) for the three methods for different strike prices of the FX digital option (K).

B.3. CPU times of CVA zero deltas of IR swaps

Method / T	1	2	3	4	5	7
LRM	9.52	34.57	72.41	137.76	227.82	413.78
LRM + QMC	10.35	35.15	73.91	135.16	224.20	416.63
B&R	9.86	32.68	72.63	134.58	223.28	408.73

Table B.21: CPU times (in sec) for the three methods for different maturities T to compute all the sensitivities with respect to the zero rates that have corresponding pillar dates smaller or equal to T.

Method / σ	0.001	0.005	0.01	0.02	0.05	0.1	0.15	0.2
LRM	33.58	33.42	34.56	33.89	33.68	34.19	33.94	33.77
LRM + QMC	34.117	34.96	35.14	35.63	32.65	34.37	35.57	34.64
B&R	38.96	38.81	32.68	34.38	38.21	35.22	34.61	32.61

Table B.22: CPU times (in sec) for the three methods for different volatilities σ to compute all the sensitivities with respect to the zero rates that have corresponding pillar dates smaller or equal to 2.

Method / K	0.03158	0.03904	0.04827
LRM	35.66	34.56	35.47
LRM + QMC	36.47	35.14	35.55
B&R	34.66	32.68	33.31

Table B.23: CPU times (in sec) for the three methods for different strike prices K to compute all the sensitivities with respect to the zero rates that have corresponding pillar dates smaller or equal to 2.