Numerical Methods for Heath-Jarrow-Morton Model of Interest Rates

by

Maria Krivko

A thesis submitted to The University of Leicester
for the degree of Doctor of Philosophy

March 2012

Abstract

Numerical Methods for Heath-Jarrow-Morton Model of Interest Rates

by

Maria Krivko

The celebrated HJM framework models the evolution of the term structure of interest rates through the dynamics of the forward rate curve. These dynamics are described by a multifactor infinite-dimensional stochastic equation with the entire forward rate curve as state variable. Under no-arbitrage conditions, the HJM model is fully characterized by specifying forward rate volatility functions and the initial forward curve. In short, it can be described as a unifying framework with one of its most striking features being the generality: any arbitrage-free interest rate model driven by Brownian motion can be described as a special case of the HJM model. The HJM model has closed-form solutions only for some special cases of volatility, and valuations under the HJM framework usually require a numerical approximation.We propose and analyze numerical methods for the HJM model. To construct the methods, we first discretize the infinite-dimensional HJM equation in maturity time variable using quadrature rules for approximating the arbitrage-free drift. This results in a finite-dimensional system of stochastic differential equations (SDEs) which we approximate in the weak and mean-square sense. The proposed numerical algorithms are highly computationally efficient due to the use of high-order quadrature rules which allow us to take relatively large discretization steps in the maturity time without affecting overall accuracy of the algorithms. They also have a high degree of flexibility and allow to choose appropriate approximations in maturity and calendar times separately. Convergence theorems for the methods are proved. Results of some numerical experiments with European-type interest rate derivatives are presented.

Acknowledgements

I would like to thank all of those who have supported me in making this work a success.

I want to thank my supervisor, Professor Michael Tretyakov, who has been a tremendous inspiration and support during the work on this thesis. Our frequent discussions made this an intellectually exhilarating, not to mention, highly joyful experience. He was always there to answer my questions – even on Skype – and encourage me when I needed it. Had it not been for his ingenious suggestions and indomitable dedication, this thesis would not have been possible.

I would like to thank Dr. David Packwood for his friendship throughout and for many helpful discussions.

I would also like to thank my parents Nikolay and Natalia whose love and support allowed me to reach my aims.

Finally, I would like to express my deepest gratitude to my friends and all my family. Dear Albina, Achim, David, Julia and Paul, I can never thank you enough for being so caring, patient and understanding.

Contents

Chapter 1

Introduction

1.1 Principles and aims of this thesis

In the seminal work [36] of Heath, Jarrow and Morton (HJM), the evolution of the term structure of interest rates is modelled via the dynamics of the forward rate curve. Specifically, the term structure dynamics are given by an infinite-dimensional multifactor stochastic differential equation taking the entire forward rate curve as a state variable. Under no-arbitrage conditions, the HJM model is fully characterized when the forward rate volatility process and the initial forward rate curve have been specified. The HJM model may best be described as a unifying framework for interest rate modeling. In fact, any arbitrage-free interest rate model driven by Brownian motion can be considered a special case of the HJM model. Its generality is one of the most striking features of the model.

While the original framework of HJM is applied to fixed income markets (see [36, 1, 15, 18, 26, 65, 28, 68, 80] and also references therein), more recent extensions of the HJM approach (see, e.g. the recent review [16]) have emerged. In [2, 7, 73] the HJM philosophy is implemented in the valuation of options on credit portfolios. Modeling the term structure of implied volatility in the spirit of the HJM approach is considered in [17, 72, 74]. The HJM philosophy has also been extended to modelling of mortality [5] and of financial electricity contracts [8].

In this thesis and [48], we deal with the standard HJM framework which models the dynamics of the forward curve

$$\{f(t,T), \quad t \leq T, \quad T \in [t_0, T^*], \quad t \in [t_0, t^*]\}.$$

Given an integrable deterministic initial forward curve

$$f(t_0, T) = f_0(T), \tag{1.1}$$

the arbitrage-free dynamics of the forward curve under the risk-neutral measure Q associated are modelled through an Ito process of the form

$$\begin{aligned} f(t,T) - f_0(T) &= \int_{t_0}^{t} \sigma^\top(s,T)\left(\int_s^T \sigma(s,u)du\right) ds \\ &+ \int_{t_0}^{t} \sigma^\top(s,T)dW(s), \quad t_0 \leq t \leq t^* \wedge T, \quad t_0 \leq T \leq T^*, \end{aligned} \tag{1.2}$$

where $W(t) = (W_1(t), \ldots, W_d(t))^\top$ is a d-dimensional standard Wiener process defined on a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t_0 \leq t \leq t^*}, Q\right)$ satisfying the usual hypotheses; $\sigma(t,T)$ is an $\mathbb{R}^d$-valued $\mathcal{F}_t$-progressively measurable stochastic process with $\int_{t_0}^T |\sigma(s,T)|^2 ds < \infty$; and $t^* \wedge T := \min(t^*, T)$.

In general, the volatility $\sigma(t,T) := \sigma(t,T,\omega)$ can depend on the current and past values of forward rates. In this thesis we restrict ourselves to the case in which σ depends on the current forward rate only, i.e.,

$$\sigma(t,T) := (\sigma_1(t,T,f(t,T)), \ldots, \sigma_d(t,T,f(t,T)))^\top,$$

where $\sigma_i(t,T,z)$, $i = 1, \ldots, d$, are deterministic functions defined on $[t_0, t^*] \times [t_0, T^*] \times \mathbb{R}$. Then the term $\int_s^T \sigma(s,u)du$ in (1.2) can be written as $\int_s^T \sigma(s,u,f(s,u))du$, and, consequently, (1.2)-(1.1) is an infinite-dimensional SDE.

The HJM model has closed-form solutions only for some special cases of the forward rate volatility process and pricing of interest rate derivatives under the HJM framework usually requires a numerical approximation. The literature on numerics for the HJM model is, to our knowledge, rather limited. In Section 3.3, we will provide an overview of approximation approaches for the HJM model.

In this thesis and [48], we propose and analyze a novel class of effective numerical methods for the HJM equation exploiting the idea of the method of lines. These methods can be used for simulating HJM models of various specifications. Our main focus is on the weak-sense numerical methods which can be used for valuing a broad range of interest rate products. To construct the numerical methods, we first discretize the infinite-dimensional HJM equation in maturity time variable T using quadrature rules for approximating the arbitrage-free drift. This results in a finite-dimensional system of stochastic differential equations (SDEs). As we show in the thesis and [48], if we take a quadrature rule of order p, the solution of this finite-dimensional system of SDEs converges to the HJM solution with mean-square order p in the maturity time discretization step Δ. From the method of lines point of view, we interpret the maturity time T as a "space" variable and the calendar time t as a "time" variable. To get fully discrete methods (i.e. discrete in both T and t), we approximate the obtained finite-dimensional system of SDEs in the weak and mean-square senses using the general theory of numerical integration of SDEs (see, e.g. [57, 58, 45]). The proposed numerical algorithms are computationally highly efficient due to the use of high-order quadrature rules which allow us to take relatively large discretization steps in the maturity time without affecting overall accuracy of the algorithms, i.e., the number of forward rates we need to approximate at each time moment t is significantly less than it is usually required in the case when the time grids for t and T coincide. Furthermore, since we exploit the method of lines, we have flexibility in choosing appropriate approximations in "space" and "time" separately. As we will see, in practice it is beneficial to use higher order rules for integration with respect to maturity time T and lower order numerical schemes for integration with respect to calendar time t.

We also prove convergence theorems for the methods constructed. We first prove convergence theorems for the HJM approximation discrete in the maturity time T

only. Then we analyze weak convergence of fully discrete methods to the approximations discrete in the maturity time. We show that this convergence is uniform in the maturity time discretization step Δ in order to obtain weak convergence of the fully-discrete numerical methods to the solution of the HJM equation. We note that in this thesis both the considered class of numerical methods and proof of their convergence include the known numerical schemes for the HJM model such as, e.g. those from [31].

We illustrate the introduced class of numerical methods by presenting some particular algorithms of various accuracy orders, which are ready for implementation. We test the proposed numerical algorithms on pricing European-type caps with the Vasicek and proportional volatility models for forward rates. The numerical tests demonstrate high computational efficiency of the proposed new algorithms.

1.2 Overview of chapters

Let us now provide a synopsis of the thesis chapters that are to follow.

Chapter 2 introduces the HJM framework and its properties. We begin by recalling various interest rates and associated products such as swaps, caps/floors and swaptions. We then explain how any arbitrage-free interest rate model driven by a Brownian motion can be described as a special case of the HJM model on examples of popular interest rate models. More specifically, we demonstrate the link between the HJM framework and selected models for LIBOR and short rates.

We begin *Chapter 3* by recalling the selected results from the theory of numerical integration of SDEs in weak and mean-square senses. Having covered the prerequisites required we then deal with the problem that initiated this research, i.e. the numerical approximation of the HJM model. The main objective of the chapter is to propose a new class of numerical methods for the HJM model that display a high degree of flexibility and computational efficiency. These methods are inspired by the idea of method of lines. We also present particular realisations of the introduced class of numerical methods along with their simulation results. This chapter is based on our paper [48].

In *Chapter 4*, we prove convergence theorems for the class of numerical methods introduced in Chapter 3. First, convergence results for the HJM approximation discrete in the maturity time T only are proved. Then, we prove convergence in weak and mean-square sense of the fully discrete methods to the approximations discrete in the maturity time.

In *Chapter 5*, we give the conclusions of the thesis and provide some remarks about possible directions of further research.

1.3 Probation of results

The results of this thesis and our paper [48] were presented at the following conferences/research seminars:

- Stochastic Analysis: A UK-China Workshop. Loughborough University, July 2011.
- Workshop on Stochastic Methods in Financial Markets. University of Ljubljana, August 2011.
- European Conference on Numerical Mathematics and Advanced Applications (ENUMATH). University of Leicester, September 2011.
- The second workshop on numerical methods for solving the filtering problem and high order methods for solving parabolic PDEs. Imperial College, London, September 2011.
- Seminar "Modern Methods in Applied Stochastics and Nonparametric Statistics". WIAS (Berlin), September 2011.
- Applied Mathematics seminar. University of Leicester, January 2012.
- Applied Mathematics and Statistics seminar (AMSTAT). University of Warwick, February 2012.

- Frontier Science Conference for Young Researchers on "Mathematics for Innovation: Large and Complex Systems". Tokyo, March 2012.

Chapter 2

The Heath-Jarrow-Morton framework

In this Chapter we shall begin by presenting well-known material from the interest rate theory based on sources such as [1, 15, 18, 26, 65, 28, 68, 80]. We concentrate mainly on the definitions and concepts required for the endeavours of this thesis. Interest rates such as instantaneous forward and LIBOR rates (Section 2.1.1) and related, simple products such as swaps, caps/floors and swaptions (Section 2.1.3) are introduced with the objective to fix and explain notation that is needed for later consideration. Selected results from the modern theory of asset pricing are reviewed in Section 2.1.2. We next introduce the classical HJM approach (Section 2.2) starting with the general framework under the objective probability measure. Then, we present the HJM dynamics under risk-neutral and forward measures. We also briefly review the LIBOR market model (Section 2.3) and one-factor short rate models (Section 2.4) in order to explain their relationship with the HJM framework, and thus emphasize the universality of the HJM modelling philosophy. The main goal of this chapter is to introduce the HJM framework and demonstrate that any arbitrage-free interest rate model driven by a Brownian motion can be described as a special case of the HJM model.

2.1 Introduction

2.1.1 Interest rates

Throughout this thesis we assume that there exists an arbitrage-free market with continuous and frictionless trading taking place inside a finite time horizon $[t_0, t^*]$ for maturities in $[t_0, T^*]$.

For convenience, let us recall some terminology of the theory of interest rates. For comprehensive studies on the interest rate term structure theory we refer to [1, 15, 18, 26, 65, 28, 68, 80].

The most basic interest contract is a default-free zero coupon bond, or simply bond, which has a single payment of one unit of cash at a fixed future maturity date T. Its price at time $t \leq T$ is denoted by $P(t,T)$. The term structure of bond prices $\{P(t,T) | T \geqslant t\}$ which we can observe at time t (today) is a (deterministic) non-increasing positive curve sufficiently smooth in T with $P(t,t) = 1$. Whereas for a fixed maturity T, $P(t,T)$ is a stochastic process, which hits the value one at its maturity, i.e. $P(T,T) = 1$.

A more informative measure of the bond market than its term structure is given by the implied interest rates. We shall introduce some of them.

A convenient, albeit a pure theoretical concept, the forward rate $f(t,T)$, $t \leq T$ represents the instantaneous continuously compounded rate prevailing at time t for riskless borrowing or lending over the infinitesimal time interval $[T, T + dT]$. The relation between zero-coupon bonds and instantaneous forward rates is given by

$$P(t,T) = \exp\left(-\int_t^T f(t,s)ds\right), \tag{2.1}$$

and

$$f(t,T) = -\frac{\partial \log P(t,T)}{\partial T}. \tag{2.2}$$

The current instantaneous rate, or so-called short rate, is

$$r(t) := f(t,t). \tag{2.3}$$

It is the current interest rate for a loan over the infinitesimal interval $[t, t + dt]$ prevailing at time t. Loosely speaking, $r(t)$ can be viewed as the overnight rate in effect at time t.

The money market account, $B(t)$, represents the accumulation factor and satisfies the following differential equation

$$dB(t) = r(t)B(t)dt,\ \ B(t_0) = 1, \tag{2.4}$$

with solution

$$B(t) = \exp\left(\int_{t_0}^{t} r(s)ds\right). \tag{2.5}$$

Intuitively, $B(t)$ represents the amount of cash accumulated up to time t starting with one unit of cash at time t_0 and continually reinvesting at the short rates $r(s)$, $s \in [t_0, t]$

Among the most important benchmark interest rates is the London Interbank Offered Rate (LIBOR). It is based on simple (or simply compounded) interest. The forward LIBOR rate $L(t, T, T+\delta)$ is the rate set at time t for the interval $[T, T+\delta]$, $t \leq T$. The accrual period of length δ is typically equal to three or six months. If we enter into a contract at time t to borrow one unit at time T and repay it with interest at time $T + \delta$, the interest due will be $\delta L(t, T, T + \delta)$.

Suppose we following this strategy: at time t sell one bond with maturity T and purchase $\frac{P(t,T)}{P(t,T+\delta)}$ of bonds with maturity $T + \delta$. The net effect is a forward investment of one unit of cash at time T yielding $\frac{P(t,T)}{P(t,T+\delta)}$ cash units at $T + \delta$ with certainty. This simple replication argument leads to the following identity between forward LIBOR rates and bond prices

$$L(t,T,T+\delta) = \frac{1}{\delta}\left(\frac{P(t,T)}{P(t,T+\delta)} - 1\right). \tag{2.6}$$

Remark 2.1.1 *A common informative measure of the current bond market at time t is the yield to maturity $y(t,T)$. It is the continuously compounding interest rate prevailing at time t for maturity T for which an investment of $P(t,T)$ at time t will produce a cash flow of one unit of cash at maturity T. It gives an indication of implied average interest rate offered by the bond. The yield $y(t,T)$ can be recovered from the price $P(t,T)$ via the formula*

$$y(t,T) = \frac{-\log P(t,T)}{T-t}. \tag{2.7}$$

Given a set of bond prices $\{P(t,T) | T \geqslant t\}$ for some fixed t, we can produce what is called the term structure of interest rate or zero coupon yield curve, a graph of $y(t,T)$ against time to maturity $T-t$ (typically in years).

2.1.2 Objective, risk-neutral and forward measures

For reference, in this part of the thesis we provide a brief review of selected results from the modern theory of asset pricing to introduce a terminology which will be helpful later on. For an introductory and complete account of the material below, see, e.g. [43, 37, 10, 20, 43].

We now fix a stochastic basis and assume that the uncertainty in the economy is characterised by a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t_0 \le t \le t^*}, \mathrm{P}\right)$ satisfying the usual hypotheses. We assume the existence of traded non-dividend paying assets with positive prices modelled by an m-dimensional vector-valued stochastic process $X(t) = (X_1(t), \ldots, X_m(t))^\top$ with $X_1(t) = B(t)$ defined in (2.4).

Let a *trading strategy* be an m-dimensional $\mathcal{F}_t-$adapted process

$$\phi(t,\omega) = (\phi_1(t,\omega), \ldots, \phi_m(t,\omega))^\top,$$

where $\phi_i(t, \omega)$ is interpreted as the number of units at time t held in asset $X_i(t)$. The value process at time t associated with the strategy ϕ is defined by

$$V(t) = \phi(t)^\top X(t).$$

The trading strategy is said to be *self-financing* if, for any $t \in [t_0, t^*]$,

$$V(t) - V(t_0) = \int_{t_0}^{t} \phi(s)^\top dX(s).$$

Intuitively, a strategy is self-financing if changes in the portfolio value are solely due to changes in the value of asset prices, with no funds being added or withdrawn.

The economic concept of *an arbitrage opportunity* in mathematical terms is defined as a self-financing strategy ϕ for which $V(t_0) = 0$ and, for some $t \in [t_0, t^*]$,

$$\mathrm{P}\left(V(t) \geq 0\right) = 1,$$

and

$$\mathrm{P}\left(V(t) > 0\right) > 0.$$

The important concept of a *numeraire* was first introduced in [27]. The idea is that one of m assets, the so-called numeraire, can be used to normalise all other assets. Let the numeraire be denoted by $D(t)$. Choosing the numeraire $D(t)$ implies the normalised asset process $X(t)/D(t) = (X_1(t)/D(t), \ldots, X_m(t)/D(t))^\top$ which is called the "discounted asset price" process. We say that a measure Q^D is an *equivalent martingale measure* induced by the numeraire D if the discounted asset price process $X(t)/D(t)$ is a martingale with respect to the measure Q^D.

The importance of an equivalent martingale measure follows from the sufficient condition for no-arbitrage: *the existence of an equivalent martingale measure implies the absence of arbitrage opportunities.* For the proof of this result, we refer to, e.g. [43].

A t^*-maturity *derivative security* (also known as a *contingent claim)* pays out an amount characterised by an $\mathcal{F}_{t^*}-$measurable random variable $H(t^*)$. We say that a contingent claim with payoff $H(t^*)$ is *attainable* if there exists a self-financing strategy ϕ such that $H(t^*) = V(t^*) = \phi(t^*)^\top X(t^*)$. We also say in this case that the trading strategy *replicates* the derivative security.

The following proposition provides mathematical characterization of the no-arbitrage price associated with any contingent claim.

Assume there exists an equivalent martingale measure Q^D induced by D and let H be a contingent claim. Then for each $t \in [t_0, t^*]$, we can price the derivative security by

$$H(t)/D(t) = E^{\mathrm{Q}}\left[H(t^*)/D(t^*)|\,\mathcal{F}_t\right], \tag{2.8}$$

where $H(t)$ is the price of contingent claim at time t. A corollary of (2.8) is that the price of any traded asset X_i, $i = 1, \ldots, m$ normalised by D is a martingale under Q^D. In practical applications, the choice of numeraire is often used as a tool to simplify calculations of the expectation in (2.8).

We say that a financial market is complete if every contingent claim is attainable. It was proved in [34] (see also [43, 37, 10, 20, 43]) that a financial market is *complete* if and only if there is a unique equivalent martingale measure. The existence of a unique martingale measure, therefore, not only eliminates the arbitrage opportunities but also guarantees the derivation of a unique price associated with any contingent claim.

We denote by Q (dropping the superscript B in notation for this particular measure) an equivalent martingale measure induced by continuously compounded money market account $B(t)$ defined in (2.4) as numeraire. This measure is often called the *risk-neutal measure.* The pricing formula (2.8) under Q is

$$H(t) = E^{\mathrm{Q}}\left[\exp\left(-\int_t^{t^*} r(s)ds\right)H(t^*)\middle|\,\mathcal{F}_t\right]. \tag{2.9}$$

This result is often referred as *the fundamental pricing formula.*

If we apply (2.9) under the risk neutral measure to the time t price $P(t,T)$ of a

T-maturity zero coupon bond, we obtain *the fundamental bond pricing formula*

$$P(t,T) = E^{\mathrm{Q}}\left[\exp\left(-\int_t^T r(s)ds\right)\middle|\, \mathcal{F}_t\right]. \tag{2.10}$$

Let us consider an equivalent martingale measure $\mathrm{Q}^{T+\delta}$ induced by the bond $P(t, T+\delta)$. This measure is called $T+\delta-$forward measure. The forward measure was introduced by Jamshidian [39] (see also [27]).

The pricing formula (2.8) under $\mathrm{Q}^{T+\delta}$ is

$$H(t)/P(t,T+\delta) = E^{\mathrm{Q}^{T+\delta}}\left[H(T)/P(T,T+\delta)|\, \mathcal{F}_t\right]. \tag{2.11}$$

The Radon-Nikodym derivative defining the measure $\mathrm{Q}^{T+\delta}$ is given by

$$\zeta_T = \frac{d\mathrm{Q}^{T+\delta}}{d\mathrm{Q}} = \frac{P(T,T+\delta)B(t_0)}{P(t_0,T+\delta)B(T)}. \tag{2.12}$$

The derivation of (2.12) is outlined as follows. By (2.11), we obtain

$$H(t_0)/P(t_0,T+\delta) = E^{\mathrm{Q}^{T+\delta}}\left[\frac{H(T)}{P(T,T+\delta)}\right],$$

and, by (2.9), we have

$$H(t_0)/P(t_0,T+\delta) = E^{\mathrm{Q}}\left[\frac{H(T)B(t_0)}{B(T)P(t_0,T+\delta)}\right].$$

Hence

$$E^{\mathrm{Q}^{T+\delta}}\left[\frac{H(T)}{P(T,T+\delta)}\right] = E^{\mathrm{Q}}\left[\frac{H(T)B(t_0)}{B(T)P(t_0,T+\delta)}\right]. \tag{2.13}$$

By definition of the Radon-Nikodym derivative, we also know that

$$E^{\mathrm{Q}^{T+\delta}}\left[\frac{H(T)}{P(T,T+\delta)}\right] = E^{\mathrm{Q}}\left[\frac{H(T)}{P(T,T+\delta)}\frac{d\mathrm{Q}^{T+\delta}}{d\mathrm{Q}}\right].$$

By comparing right hand sides of the last two equalities, we obtain (2.12).

For $t \leq T$, by the definition of the Radon-Nikodym process and (2.8), we have

$$\zeta_t = E^{\mathrm{Q}}\left[\zeta_T \middle| \mathcal{F}_t\right] = E^{\mathrm{Q}}\left[\frac{P(T, T+\delta)B(t_0)}{P(t_0, T+\delta)B(T)}\middle| \mathcal{F}_t\right] = \frac{P(t, T+\delta)B(t_0)}{P(t_0, T+\delta)B(t)}. \qquad (2.14)$$

Setting in (2.11) $\delta = 0$, we obtain the well-known result (see for e.g. [1, 15, 26, 80]) for valuing a claim under $T-$forward measure:

$$H(t) = P(t, T)E^{\mathrm{Q}^T}\left[H(T)\middle| \mathcal{F}_t\right].$$

A particularly appealing feature of this result is that it does not require knowledge of the joint distribution of $H(T)$ and $\frac{1}{B(T)}$ (cf. (2.9)) or $\frac{1}{P(T,T+\delta)}$ (cf. (2.11)).

Based on (2.11), one can show that the forward LIBOR rates $L(t, T, T+\delta)$ are martingales under $\mathrm{Q}^{T+\delta}$. Indeed, by putting $H(t) = L(t, T, T+\delta)P(t, T+\delta)$ in (2.11), we obtain

$$L(t, T, T+\delta) = E^{\mathrm{Q}^{T+\delta}}\left[L(T, T, T+\delta)\middle| \mathcal{F}_t\right]. \qquad (2.15)$$

2.1.3 Interest rate products

Let us review now some basic interest rate products such as swaps, caps/floors and swaptions [1, 15, 18, 26, 65, 28, 68, 80].

We specify a number of future dates $T_0 < T_1 < \ldots < T_n$ for notational simplicity equidistant with $\delta = T_i - T_{i-1}$, $i = 1, \ldots, n$. For convenience, we assume a unit notional value of all the contracts we introduce below.

Interest rate swap

An Interest-Rate Swap (IRS) is a contract allowing to exchange a payment stream at a fixed interest rate for a payment stream at a floating rate. Popularity of swaps reflects the fact that different companies can borrow at fixed or at floating rates in different markets. We shall consider here only plain vanilla IRS.

A payer (receiver) interest rate swap with a fixed rate K and a unit nominal value settled in arrears is a contract according to which its holder pays (receives) fixed payments of δK and receives (pays) floating payments of $\delta L(T_{i-1}, T_{i-1}, T_i)$ at the coupon dates T_i, $i = 1, \ldots, n$. At this description, we are considering for simplicity that fixed-rate payments and floating-rate payments occur at the same dates and with the same year fraction. Though the generalisation to different payment dates and day-count conventions is straightforward.

The net cash flow at time T_i is

$$\delta \left(L(T_{i-1}, T_{i-1}, T_i) - K \right).$$

By the fundamental evaluation result (2.9), the value of the swap at time $t \leq T_0$ is equal to the expected discounted value of its net cash flows, i.e.

$$\begin{aligned}
V_{swap}(t) &= B(t) \sum_{i=1}^{n} E^{\mathrm{Q}} \left(B^{-1}(T_i) \delta \left(L(T_{i-1}, T_{i-1}, T_i) - K \right) \middle| \mathcal{F}_t \right) \\
&= B(t) \sum_{i=1}^{n} E^{\mathrm{Q}} \Big[B^{-1}(T_{i-1}) \delta \left(L(T_{i-1}, T_{i-1}, T_i) - K \right) \\
&\qquad E^{\mathrm{Q}} \left(\exp \left(- \int_{T_{i-1}}^{T_i} r(s) ds \right) \middle| \mathcal{F}_{T_{i-1}} \right) \Big| \mathcal{F}_t \Big] \\
&= B(t) \sum_{i=1}^{n} E^{\mathrm{Q}} \left(B^{-1}(T_{i-1}) \delta \left(L(T_{i-1}, T_{i-1}, T_i) - K \right) P(T_{i-1}, T_i) \middle| \mathcal{F}_t \right),
\end{aligned}$$

in the second and third equalities we used the iterated conditioning and (2.10), correspondingly. Then, using (2.6), we obtain

$$\begin{aligned}
V_{swap}(t) &= \sum_{i=1}^{n} B(t) E^{\mathrm{Q}} \left(B^{-1}(T_{i-1}) \left(P(T_{i-1}, T_{i-1}) - P(T_{i-1}, T_i) - K \delta P(T_{i-1}, T_i) \right) \middle| \mathcal{F}_t \right) \\
&= \sum_{i=1}^{n} \left[P(t, T_{i-1}) - P(t, T_i) - K \delta P(t, T_i) \right] \\
&= P(t, T_0) - P(t, T_n) - K \delta \sum_{i=1}^{n} P(t, T_i), \qquad (2.16)
\end{aligned}$$

where in the second equality we used the fact that since $P(\cdot, T_i)$ are traded assets, their discounted prices by $B^{-1}(\cdot)$ are martingale.

Recalling again (2.6) , (2.16) can be rewritten as

$$V_{swap}(t) = \delta \sum_{i=1}^{n} P(t, T_i) \left(L(t, T_{i-1}, T_i) - K \right). \tag{2.17}$$

The interesting observation from (2.16) and (2.17) is that plain vanilla IRS can be valued at time $t \leq T_0$ using only the term structure observed at that time.

The fixed rate K that makes the IRS a fair contract at time $t \leq T_0$, i.e. for which $V_{swap}(t) = 0$ is called the forward swap rate $R_{swap}(t)$ and is given by

$$R_{swap}(t; T_0, T_n) = \frac{P(t, T_0) - P(t, T_n)}{\delta \sum_{i=1}^{n} P(t, T_i)}. \tag{2.18}$$

Thus, for $t \leq s \leq T_0$ the swap evaluation formula (2.17) can be rewritten as follows

$$V_{swap}(s) = \delta \sum_{i=s+1}^{n} P(s, T_i)(L(s, T_{i-1}, T_i) - R_{swap}(t; T_0, T_n)).$$

Caps and floors

An Interest Rate Cap (IRC) is a security that allows its holder to benefit from low floating rates and be protected from high ones. It can be viewed as a payer IRS, where each exchange payment is executed only if it has positive value. Similarly, an Interest Rate Floor (IRF) is an instrument designed to protect from low floating interest rates yet allow the holder to benefit from the high rates. It is equivalent to a receiver IRS, where exchange payments take place only if their values are positive.

Formally, a cap price is obtained by summing up the prices of the underlying caplets, call options on successive LIBOR rates. Consider a caplet set at time T_{i-1} with payment date at T_i, $i \geq 1$ with strike K and unit cap nominal value. Its price at time $t \leq T_0$ is given by

$$\begin{aligned} V_{caplet}(t) &= E^{\mathrm{Q}} \left(\exp \left(- \int_t^{T_i} r(u) du \right) \delta \left(L(T_{i-1}, T_{i-1}, T_i) - K \right)_+ \middle| \mathcal{F}_t \right) \\ &= E^{\mathrm{Q}} \left(\exp \left(- \int_t^{T_{i-1}} r(u) du \right) P(T_{i-1}, T_i) \delta \left(L(T_{i-1}, T_{i-1}, T_i) - K \right)_+ \middle| \mathcal{F}_t \right) \end{aligned}$$

$$= E^{\mathrm{Q}}\left(\exp\left(-\int_t^{T_{i-1}} r(u)du\right)(1-(1+K\delta)P(T_{i-1},T_i))_+\middle|\,\mathcal{F}_t\right), \tag{2.19}$$

where the second equality comes from iterated conditioning and (2.10). Therefore, the caplet price can be written as a multiple of the price of the European put with maturity T_{i-1}, strike $1/(1+K\delta)$, written on a zero-coupon bond with maturity T_i and $(1+K\delta)$ nominal amount.

Therefore, the caplet price can be written as a European put option with maturity T_{i-1} and unit strike written on $(1+K\delta)$ units of a zero coupon bond with maturity T_i.

We note that the caplet price in terms of instantaneous forward rates becomes

$$\begin{aligned} & V_{caplet}(t) \qquad (2.20)\\ = \; & E^{\mathrm{Q}}\exp\left(-\int_t^{T_{i-1}} r(u)du\right)\left[1-(1+K\delta)\exp\left(-\int_{T_{i-1}}^{T_i} f(T_{i-1},u)du\right)\right]_+ . \end{aligned}$$

By switching to the T_i-forward measure (cf. (2.11)), the valuation formula (2.19) (and analogously (2.20)) for caplets can be written in a more convenient form

$$V_{caplet}(t) = \delta P(t,T_i)E^{\mathrm{Q}^{T_i}}\left((L(T_{i-1},T_{i-1},T_i)-K)_+\middle|\,\mathcal{F}_t\right). \tag{2.21}$$

Similarly, a floor is a strip of floorlets, put options on successive LIBOR rates. Though we do not present a valuation formula for a floorlet price here, its form can be simply deduced from (2.19) or found in [1, 15, 18, 26, 65, 28, 68].

Swaptions

A European payer (receiver) swaption is an option that gives its holder a right, but not an obligation, to enter a payer (receiver) swap at a future date at a given fixed rate K. Usually, the swaption maturity coincides with the first reset date T_0 of the underlying swap. The underlying swap length $T_n - T_0$ is called the tenor of the swaption.

The value of the payer swaption at time $t \le T_0$ (cf. (2.17)) can be found as

$$\begin{aligned} V_{swaption}(t) &= E^{\mathrm{Q}}\left(\exp\left(-\int_t^{T_0} r(u)du\right)(V_{swap}(T_0))_+ \middle| \mathcal{F}_t\right) \\ &= E^{\mathrm{Q}}\left(\exp\left(-\int_t^{T_0} r(u)du\right)\right. \\ &\quad \times \left.\left[\sum_{i=1}^{n} P(T_0,T_i)\delta\left(L(T_0,T_{i-1},T_i)-K\right)\right]_+ \middle| \mathcal{F}_t\right), \end{aligned} \tag{2.22}$$

or alternatively in terms of instantaneous forward rates evaluates to

$$\begin{aligned} V_{swaption}(t) &= E^{\mathrm{Q}}\left(\exp\left(-\int_t^{T_0} r(u)du\right)\left[1-\exp\left(-\int_{T_0}^{T_n} f(T_0,u)du\right)\right.\right. \\ &\quad \left.\left.-\sum_{i=1}^{n} K\delta\exp\left(-\int_{T_0}^{T_i} f(T_0,u)du\right)\right]_+ \middle| \mathcal{F}_t\right). \end{aligned} \tag{2.23}$$

Using (2.18), we can rewrite (2.22) in a more compact form

$$V_{swaption}(t) = E^{\mathrm{Q}}\exp\left(-\int_t^{T_0} r(u)du\right)\delta\left(R_{swap}(T_0)-K\right)_+\sum_{i=1}^{n} P(T_0,T_i). \tag{2.24}$$

It is evident from (2.24) that a European payer (receiver) swaption is a call (put) option on the forward swap rate struck at the fixed rate of swap.

We note that $R_{swap}(t)\sum_{i=1}^{n} P(t,T_i)$ is a price of a tradable asset and that $\sum_{i=1}^{n} P(t,T_i)$, being a linear combination of zero coupon bonds, can be classified as a numeraire. The measure induced by this numeraire is known (see e.g. [1, 15, 26, 68]) as a swap measure with $R_{swap}(t)$ being a martingale under this measure.

2.2 The Heath-Jarrow-Morton framework of instantaneous forward rates

2.2.1 Forward curve dynamics

The Heath-Jarrow-Morton (HJM) model refers to a class of models that characterises the evolution of the term structure of interest rates through the dynamics of the forward rate curve [36]. These dynamics are described by a multifactor infinite-dimensional stochastic equation with the entire forward rate curve as state variable. It is important to note that HJM framework is a general setup which unifies all interest rate models driven by a Brownian motion.

We shall follow this setup pioneered by Heath, Jarrow and Morton in 1992 [36]. We assume that there exist an economy with a frictionlessly traded continuum of default-free zero-coupon bonds $\{P(t,T),\ t \leq T,\ T \in [t_0, T^*],\ t \in [t_0, t^*]\}$, where $P(t,T)$ denotes the price at time t of a bond with maturity T.

In the framework proposed by Heath, Jarrow, and Morton the forward curve dynamics under the objective probability measure P are modelled through an Ito process of the form

$$\begin{aligned} f(t,T) - f_0(T) &= \int_{t_0}^{t} \mu(s,T)ds + \int_{t_0}^{t} \sigma^\top(s,T)dW(s), \\ t_0 &\leq t \leq t^* \wedge T, \quad t_0 \leq T \leq T^*, \end{aligned} \tag{2.25}$$

where

- $W(t) = (W_1(t), \ldots, W_d(t))^\top$ is a d-dimensional standard Wiener process defined on $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t_0 \leq t \leq t^*}, \mathrm{P}\right)$;

- $f_0(T) := f(t_0, T)$ is a fixed, deterministic initial forward rate curve which is measurable as mapping $f(t_0, \cdot) : ([t_0, T^*], \mathcal{B}[t_0, T^*]) \to (\mathbb{R}, \mathcal{B})$, where $\mathcal{B}[t_0, T^*]$ is the Borel σ-algebra restricted to $[t_0, T^*]$;

- $\mu(s,T) := \mu(t,T,\omega)$ is an $\mathbb{R}$-valued $\mathcal{F}_t$-progressively measurable stochastic process, i.e. it is a measurable mapping:

  $$\mu : (\{(s,u) : t_0 \leq s \leq u \leq T\} \times \Omega, \mathcal{B}(\{(s,u) : t_0 \leq s \leq u \leq T\}) \otimes \mathcal{F}_s) \to (\mathbb{R}, \mathcal{B}),$$

  with

  $$\int_{t_0}^{T} |\mu(s,T,\omega)|\, dt < +\infty \text{ a.e. P, for all } T \in [t_0, T^*];$$

- $\sigma(s,T) := \sigma(t,T,\omega)$ is an $\mathbb{R}^d$-valued $\mathcal{F}_t$-progressively measurable stochastic process, i.e. σ_i is measurable as mapping:

$$\sigma_i : (\{(s,u) : t_0 \leq s \leq u \leq T\} \times \Omega, \mathcal{B}(\{(s,u) : t_0 \leq s \leq u \leq T\}) \otimes \mathcal{F}_s) \to (\mathbb{R}, \mathcal{B}),$$

with

$$\int_{t_0}^{T} \sigma_i^2(t,T,\omega)dt < +\infty \text{ a.e. P, for all } T \in [t_0, T^*], \ i = 1, \ldots, d;$$

and $t^* \wedge T := \min(t^*, T)$.

We note that this setup is quite general. The only substantive economic restrictions imposed on the forward rate process are that they have continuous sample paths and that they are driven by a finite number of random shocks.

Remark 2.2.1 *We follow the classical HJM framework and consider the case when the forward rate dynamics are driven by a finite-dimensional Wiener process, i.e. $d < \infty$ in (2.25). For an infinite-dimensional prospective on HJM framework we refer to [25, 18].*

Next, we are interested in the dynamics of the short rate process. From (2.25) with $T = t$, the dynamics of the short rate process in integral form is

$$r(t) = f_0(t) + \int_{t_0}^{t} \mu(s,t)ds + \int_{t_0}^{t} \sigma^\top(s,t)dW(s), \tag{2.26}$$

To obtain the differential form, we apply the Ito formula to $r(t) := f(t,t)$

$$\begin{aligned} dr(t) &= df(t,T)|_{T=t} + \frac{\partial}{\partial T} f(t,T)|_{T=t} dt, \\ r(t_0) &= f_0(t_0). \end{aligned}$$

This yields

$$dr(t) = \left[\frac{df_0(t)}{dt} + \mu(t,t) + \int_{t_0}^{t} \frac{\partial \mu(s,t)}{\partial t} ds \right. \tag{2.27}$$

$$
\begin{aligned}
&\quad + \int_{t_0}^{t} \frac{\partial \sigma^\top(s,t)}{\partial t} dW(s) \Bigg] dt + \sigma^\top(t,t) dW(t), \\
r(t_0) &= f_0(t_0).
\end{aligned}
$$

The sufficient regularity conditions (see [36, Conditions C.2]) to ensure that the savings account $B(t)$ satisfies

$$
0 < B(t) < +\infty \quad a.e.\mathrm{P}, t_0 \le t \le t^*,
$$

are

$$
\int_{t_0}^{t^*} |f_0(s)|\, ds < +\infty \text{ and } \int_{t_0}^{t^*} \left[\int_{t_0}^{t} |\mu(s,t)|\, du \right] ds < +\infty \text{ a.e. P.} \tag{2.28}
$$

The conditions (see [36, Conditions C.3]) imposed on the bond price process to ensure that it is well-behaved are

$$
\int_{t_0}^{t} \left[\int_{s}^{t} |\sigma_i(s,u)|\, du \right]^2 ds < +\infty \text{ a.e. P, for all } t \in [t_0, t^*],\ i = 1, \ldots, d; \tag{2.29}
$$

$$
\int_{t_0}^{t} \left[\int_{t}^{T} |\sigma_i(s,u)|\, du \right]^2 ds < +\infty \text{ a.e. P, for all } t \in [t_0, T], T \in [t_0, T^*],\ i = 1, \ldots, d; \tag{2.30}
$$

and

$$
t \to \int_{t}^{T} \left[\int_{t_0}^{t} |\sigma_i(s,u)|\, dW_i(s) \right]^2 du < +\infty \tag{2.31}
$$

is continuous a.e. P, for all $T \in [t_0, T^*]$, $i = 1, \ldots, d$.

Lemma 2.2.2 *Suppose conditions (2.28)-(2.31) are satisfied. Then the zero-coupon bond price process corresponding to (2.25) is an Ito process of the form*

$$
\begin{gathered}
P(t,T) = P(t_0,T) \\
+ \int_{t_0}^{t} P(s,T)(r(s) + b(s,T)) ds + \int_{t_0}^{t} P(s,T) \sigma_P^\top(s,T) dW(s), \\
t_0 \le t \le t^* \wedge T, \quad t_0 \le T \le T^*,
\end{gathered} \tag{2.32}
$$

where

$$\sigma_P(s,T) = -\int_s^T \sigma(s,u)du, \tag{2.33}$$

is the T-bond volatility and

$$b(s,T) = -\int_s^T \mu(s,u)du + \frac{1}{2}\sigma_P^\top(s,T)\sigma_P(s,T). \tag{2.34}$$

Proof. Based on (2.1) and (2.25) the log-dynamics of the bond price process is

$$\begin{aligned}
\log P(t,T) &= -\int_t^T f(t,u)du \\
&= -\int_t^T f_0(u)du - \int_t^T \int_{t_0}^t \mu(s,u)ds\ du - \int_t^T \int_{t_0}^t \sigma^\top(s,u)dW(s)\ du.
\end{aligned}$$

Using the classical and stochastic Fubini Theorems (Appendix B and also see [36, 26]) twice, we obtain

$$\begin{aligned}
&\log P(t,T) \\
= & -\int_t^T f_0(u)du - \int_{t_0}^t \int_t^T \mu(s,u)du\ ds - \int_{t_0}^t \int_t^T \sigma^\top(s,u)du\ dW(s) \\
= & -\int_{t_0}^T f_0(u)du - \int_{t_0}^t \int_s^T \mu(s,u)du\ ds - \int_{t_0}^t \int_s^T \sigma^\top(s,u)du\ dW(s) \\
& + \int_{t_0}^t f_0(u)du + \int_{t_0}^t \int_s^t \mu(s,u)du\ ds + \int_{t_0}^t \int_s^t \sigma^\top(s,u)du\ dW(s) \\
= & -\int_{t_0}^T f_0(u)du + \int_{t_0}^t \left(b(s,T) - \frac{1}{2}\sigma_P^\top(s,T)\sigma_P(s,T)\right) ds + \int_{t_0}^t \sigma_P(s,T)\ dW(s) \\
& + \int_{t_0}^t \left(f_0(u) + \int_{t_0}^u \mu(s,u)\ ds + \int_{t_0}^u \sigma^\top(s,u)\ dW(s)\right) du.
\end{aligned}$$

Then by (2.26) we have

$$\begin{aligned}
\log P(t,T) &= \log P(t_0,T) + \int_{t_0}^t \left(r(s) + b(s,T) - \frac{1}{2}\sigma_P^\top(s,T)\sigma_P(s,T)\right) ds \\
&\quad + \int_{t_0}^t \sigma_P(s,T)\ dW(s).
\end{aligned}$$

An application of Ito's formula to the above result implies (2.32). □

We also shall be interested in the dynamics for the discounted bond price process defined as $\frac{P(t,T)}{B(t)}$. This is the value of a T-maturity bond expressed in the units of the accumulating factor $B(t)$. A straightforward application of Ito's formula yields (cf. (2.4) and (2.32))

$$\begin{aligned} \frac{P(t,T)}{B(t)} &= P(t_0,T) \\ + \int_{t_0}^{t} \frac{P(s,T)}{B(s)} b(s,T)ds &+ \int_{t_0}^{t} \frac{P(s,T)}{B(s)} \sigma_P^\top(s,T)dW(s), \\ t_0 \le t \le t^* \wedge T, &\quad t_0 \le T \le T^*. \end{aligned} \tag{2.35}$$

2.2.2 HJM: Risk-neutral measure dynamics

We shall now investigate the restriction on the HJM dynamics (2.25) imposed by no-arbitrage argument under the risk-neutral measure Q associated with the numeraire $B(t)$. In what follows we let $\mathrm{Q} \sim \mathrm{P}$ be an equivalent probability measure. We denote by W^{Q} the d-dimensional Brownian motion under measure Q obtained by Girsanov transform.

Theorem 2.2.3 *(HJM drift condition) Given that the forward curve evolution under the objective probability measure* P *is of the form (2.25) and conditions (2.28)-(2.31) hold, the arbitrage-free dynamics of the forward curve under the risk neutral measure* Q *associated with the numeraire $B(t)$ are of the form*

$$\begin{aligned} f(t,T) - f_0(T) &= \int_{t_0}^{t} \sigma^\top(s,T) \left(\int_s^T \sigma(s,u)du \right) ds \\ &\quad + \int_{t_0}^{t} \sigma^\top(s,T)dW^{\mathrm{Q}}(s), \quad t_0 \le t \le t^* \wedge T, \quad t_0 \le T \le T^*. \end{aligned} \tag{2.36}$$

Proof. There is no arbitrage in the model if there is an equivalent martingale measure Q such that the discounted bond price process $\frac{P(t,T)}{B(t)}$ is a Q-martingale, as we already stated in Section 2.1.2 (see also [20, 1, 15, 26]), for $t_0 \le t \le t^* \wedge T,\ t_0 \le T \le T^*$. Hence, in view of (2.35) the Q-dynamics of $\frac{P(t,T)}{B(t)}$ are of the form

$$\frac{P(t,T)}{B(t)} = P(t_0,T) + \int_{t_0}^{t} \frac{P(s,T)}{B(s)} \sigma_P^\top(s,T)dW^{\mathrm{Q}}(s), \tag{2.37}$$

$$t_0 \le t \le t^* \wedge T, \quad t_0 \le T \le T^*,$$

and this obviously posits the dynamics of the bond price process (cf. (2.32)) to be

$$\begin{aligned} P(t,T) &= P(t_0,T) \\ &\quad + \int_{t_0}^{t} P(s,T)r(s)ds + \int_{t_0}^{t} P(s,T)\sigma_P^\top(s,T)dW^{\mathrm{Q}}(s). \end{aligned} \tag{2.38}$$

Through (2.1), the dynamics in (2.38) constrain the evolution of forward rates. By Ito's formula,

$$d\ln P(t,T) = \left[r(t) - \frac{1}{2}\sigma_P^\top(t,T)\sigma_P(t,T)\right]dt + \sigma_P^\top(t,T)dW^{\mathrm{Q}}(t). \tag{2.39}$$

Thus

$$\begin{aligned} df(t,T) &= -d\left(\frac{\partial \log P(t,T)}{\partial T}\right) = -\frac{\partial}{\partial T}\left(d\log P(t,T)\right) \\ &= -\frac{\partial}{\partial T}\left[-\frac{1}{2}\sigma_P^\top(t,T)\sigma_P(t,T)\right]dt - \frac{\partial}{\partial T}\sigma_P^\top(t,T)dW^{\mathrm{Q}}(t) \\ &= \sigma_P^\top(t,T)\frac{\partial}{\partial T}\sigma_P(t,T)dt - \frac{\partial}{\partial T}\sigma_P^\top(t,T)dW^{\mathrm{Q}}(s). \end{aligned} \tag{2.40}$$

Using (2.33) we can, therefore, rewrite the expression for the drift in (2.40) as

$$\sigma^\top(t,T)\int_t^T \sigma(t,u)du,$$

This is the risk-neutral drift of the forward curve process imposed by the absence of arbitrage. Substituting it back into (2.40), we arrive to (2.36). □

2.2.3 HJM: Forward measure dynamics

The HJM dynamics can be written under the T_i-forward measure Q^{T_i} instead of the risk-neutral measure. Let us denote a drift in the HJM dynamics under the T_i-forward measure by $\mu_f^{T_i}(t,T)$.

We introduce

$$\alpha_j(t,T) := -\left(\sigma_P(t,T)\right)_j - \frac{\left(\mu_f^{T_i}(t,T)\right)_j}{\sigma_j(t,T)},$$

where the subscript j denotes the corresponding component of the vector. Then, in view of (2.36), by Girsanov's theorem, we have

$$\zeta_t = \left.\frac{dQ^{T_i}}{dQ}\right|_{\mathcal{F}_t} = \exp\left(-\int_{t_0}^t \alpha^\top(s)dW^Q(s) - \frac{1}{2}\int_{t_0}^t \alpha^\top(s)\alpha(s)ds\right).$$

The process ζ_t is an exponential martingale and has the dynamics

$$d\zeta_t = -\alpha^\top(t)\zeta_t dW^Q(t). \tag{2.41}$$

The process W^{T_i} defined by

$$dW^{T_i} = \alpha(t)dt + dW^Q(t),$$

is a standard Brownian motion under Q^{T_i}.

Recall (2.14):

$$\zeta_t = \frac{P(t,T_i)}{P(t_0,T_i)B(t)}, \tag{2.42}$$

and, hence, taking into account (2.37) :

$$d\zeta_t = \frac{1}{P(t_0,T_i)}\frac{P(t,T_i)}{B(t)}\sigma_P(t,T_i)dW^Q(t). \tag{2.43}$$

Comparing (2.41) and (2.43), we deduce

$$-\alpha(t)\zeta_t = \frac{1}{P(t_0,T_i)}\frac{P(t,T_i)}{B(t)}\sigma_P(t,T_i).$$

Finally, taking into account (2.42), we obtain

$$\alpha(t,T) = -\sigma_P(t,T_i),$$

or, by definition of α,

$$\mu_f^T(t,T) = -\sigma^\top(t,T)\left(\int_T^{T_i}\sigma(t,u)du\right).$$

Hence, the corresponding forward rate dynamics have the form (cf. (2.36)):

$$\begin{aligned} f(t,T) - f_0(T) &= -\int_{t_0}^{t}\sigma^\top(s,T)\left(\int_T^{T_i}\sigma(s,u)du\right)ds \qquad (2.44)\\ +\int_{t_0}^{t}\sigma^\top(s,T)dW^{\mathrm{Q}^{T_i}}(s),\ t_0 &\le t\le t^*\wedge T,\quad T<T_i\le T^*. \end{aligned}$$

2.2.4 The HJM framework: summary and discussion

In the HJM setting we concern ourselves with modelling the evolution of the term structure of interest rates over time by describing the dynamics of the forward rate curve. The resulting class of models is very broad. In this sense, the HJM model is probably the best described as a unifying framework with one of its most striking features being the generality: any arbitrage-free interest rate model driven by a Brownian motion can be described as a special case of the HJM model.

To specify a particular no-arbitrage HJM dynamics (cf. $(2.36), (2.44)$), two inputs are required, namely, the forward rate volatility function, $\sigma(t,T)$, and the initial forward curve, $f_0(T)$. Thus, any arbitrage-free interest rate model defined in the filtration generated by a Brownian motion corresponds to a particular choice of $\sigma(t,T)$. Also, we note that the HJM models are automatically consistent with the initial bond prices $P(t_0,T)$ if the initial forward curve chosen relates to these bonds prices through (2.1). This contrasts with the short rate model approach (see Section 2.4 below) where choosing parameters of the drift is essential for calibrating models to observed bond prices.

The HJM dynamics (2.25) are described by a multifactor infinite-dimensional stochastic equation with the entire forward rate curve as state variable. We note that it can be transformed to a first-order hyperbolic SPDE using the Musiela parameterization [61], where the forward curve is parametrized by the time to maturity

$x = T - t$:

$$u(t,x) = f(t, t+x).$$

This results in the following SPDE

$$\begin{aligned} du(t,x) &= \left(\frac{\partial}{\partial x} u(t,x) + \tilde{\mu}(t,x)\right) dt + \tilde{\sigma}^{\top}(t,x) dW(t), \\ u(t_0, \cdot) &= f(t_0, \cdot), \\ t_0 &\leq t \leq t^*,\ 0 \leq x \leq T^* - t_0, \end{aligned}$$

where $\tilde{\mu}(t,x) := \mu(t, t+x)$ and $\tilde{\sigma}(t,x) := \sigma(t, t+x)$.

The HJM model in the framework of stochastic partial differential equation has attracted a significant amount of attention in the literature, see e.g. the monographs [18, 25] and references therein.

Here, we have considered exclusively models driven by Wiener processes. For the literature on the extended HJM methodology for the term structure models driven by Poisson measures see for instance the monographs [77, 11]. In [24] term structure models driven by a Lévy process are considered.

2.3 Market model dynamics of LIBOR rates

The models considered in this section are closely related to the HJM framework (Section 2.2) in that they describe the arbitrage-free dynamics of the term structure of interest rates through the evolution of forward rates. But they model the dynamics of simple interest rates, such as LIBOR rates, rather then continuously compounded forward rates. In this section we shall consider the interest rate market model developed primarily through the work of Brace, Gatarek, and Musiela (BGM) [14] and Miltersen, Sandmann, and Sondermann [42]. The term “market model” refers to the fact that the approach is based on modelling interest rates that are observable in the market. These models follow the spirit of HJM (see Sections 2.2.2, 2.2.3), where the drift conditions of the modelled rates are imposed by no-arbitrage considerations once the numeraire and volatility structure are specified. The LIBOR market model is a subclass of the general HJM framework that is derived based on the assumption of lognormal LIBOR rate dynamics. The motivation for its introduction was to derive a model consistent with the market practice to price caplets/floorlets with Black's formula.

We consider a class of models in which a finite set of maturities or *tenor* dates

$$t_0 = T_0 < \cdots < T_N = T^*, \quad T_i = i\delta, \quad i = 0, \ldots, N, \tag{2.45}$$

are fixed in advance, where

$$\delta = (T^* - t_0)/N,$$

denotes the fixed length of the interval between tenor dates and typically is set to either three or six months.

Let us denote, for simplicity of the presentation, the time t forward LIBOR rate (2.6) for the accrual period $[T_i, T_{i+1}]$ and the payment at T_{i+1} by

$$\begin{aligned} L^i(t) \quad &: \quad = L(t, T_i, T_{i+1}), \\ t_0 \quad &\leq \quad t \leq t^* \wedge T_i, \quad t_0 < T_i \leq T^*, \quad i = 0, \ldots, N-1. \end{aligned}$$

Also, denote by $\varrho(t)$ the auxiliary index dependent on time t so that

$$\varrho(t) = \min \{ i = 0, 1, \ldots, N : \quad t < T_i \}, \tag{2.46}$$

i.e., or $T_{\varrho(t)}$ is the closest maturity to the time t from the right.

Recall that the LIBOR rate as of time t is given in the terms of the continuous forward rate curve at time t by (cf. (2.6), (2.2))

$$1 + \delta L^i(t) = \exp\left(\int_{T_i}^{T_{i+1}} f(t, u) du \right). \tag{2.47}$$

2.3.1 LIBOR: Risk neutral measure dynamics

We shall follow the BGM original approach [14] which derives the LIBOR market model from the arbitrage-free dynamics (2.36) of the continuous forward curve under the risk-neutral measure Q. To specify an HJM model, or equivalently, to specify instantaneous forward rate volatility function, $\sigma(t,T)$, we assume that the LIBOR rate dynamics have lognormal volatility structure, i.e.

$$\frac{dL^i(t)}{L^i(t)} = \mu_{L^i}(t)dt + \lambda_i^\top(t)dW^{\mathbb{Q}}(t), \tag{2.48}$$

where $\lambda_i(t) := \lambda(t, T_i; \delta)$ is an $\mathbb{R}^d$-valued bounded and piecewise continuous deterministic function and $\mu_{L^i}(t) := \mu_L(t, T_i; \delta)$ is the drift corresponding to the dynamics (2.48) which we shall derive below.

Using the Ito formula, (2.47) and (2.39), we have

$$\begin{aligned}
dL^i(t) &= \frac{1}{\delta} d\exp\left(\int_{T_i}^{T_{i+1}} f(t,u)du\right)\\
&= \frac{1}{\delta}\exp\left(\int_{T_i}^{T_{i+1}} f(t,u)du\right) d\left(\int_{T_i}^{T_{i+1}} f(t,u)du\right)\\
&\quad + \frac{1}{2\delta}\exp\left(\int_{T_i}^{T_{i+1}} f(t,u)du\right) \left|\sigma_P(t,T_i) - \sigma_P(t,T_{i+1})\right|^2 dt\\
&= \frac{1}{\delta}\exp\left(\int_{T_i}^{T_{i+1}} f(t,u)du\right) \left[\frac{1}{2}\left(|\sigma_P(t,T_{i+1})|^2 - |\sigma_P(t,T_i)|^2\right)dt\right.\\
&\quad \left. + (\sigma_P(t,T_i) - \sigma_P(t,T_{i+1}))^\top dW^{\mathbb{Q}}(t)\right]\\
&\quad + \frac{1}{2\delta}\exp\left(\int_{T_i}^{T_{i+1}} f(t,u)du\right) \left|\sigma_P(t,T_i) - \sigma_P(t,T_{i+1})\right|^2 dt\\
&= \frac{1}{\delta}\left(1 + \delta L^i(t)\right)\left[\sigma_P^\top(t,T_i)\left(\sigma_P(t,T_i) - \sigma_P(t,T_{i+1})\right)dt\right.\\
&\quad \left. + \left(\sigma_P(t,T_i) - \sigma_P(t,T_{i+1})\right)^\top dW^{\mathbb{Q}}(t)\right].
\end{aligned}$$

In order for the LIBOR rate dynamics to have a lognormal volatility structure (2.48), we need to impose the following restriction on the volatility structure of the zero-coupon bond prices,

$$\sigma_P(t,T_i) - \sigma_P(t,T_{i+1}) = \frac{\delta L^i(t)}{(1 + \delta L^i(t))}\lambda_i(t), \tag{2.49}$$

or, equivalently,

$$\begin{aligned}\sigma_P(t,T_{i+1}) &= -\sum_{j=\varrho(t)}^{i}\frac{\delta L^j(t)}{(1+\delta L^j(t))}\lambda_j(t)+\sigma_P(t,T_{\varrho(t)}),\\ t_0 &\leq t\leq t^*\wedge T_i,\quad t_0\leq T_i\leq T^*.\end{aligned}$$

Consequently, the LIBOR rate process can be rewritten as

$$\begin{aligned}\frac{dL^i(t)}{L^i(t)} &= \lambda_i^\top(t)\left(\sum_{j=\varrho(t)}^{i}\frac{\delta L^j(t)}{(1+\delta L^j(t))}\lambda_j(t)-\sigma_P(t,T_{\varrho(t)})\right)dt\\ &\quad+\lambda_i^\top(t)dW^{\mathrm{Q}}(t), \qquad (2.50)\\ t_0 &\leq t\leq t^*\wedge T_i,\quad t_0\leq T_i\leq T^*.\end{aligned}$$

Remark 2.3.1 *(Relationship with the HJM analysis) From the derivation of the LIBOR rate model it is clear that it is a special case of the general HJM class of diffusive interest rate models. It follows from (2.33) and (2.49) that the LIBOR rate model volatility $L(t,T,T+\delta)\lambda(t,T;\delta)$ is related to the HJM instantaneous volatility function $\sigma(t,T)$ by*

$$\int_T^{T+\delta}\sigma(t,u)du=\frac{\delta L(t,T,T+\delta)}{(1+\delta L(t,T,T+\delta))}\lambda(t,T;\delta). \qquad (2.51)$$

We can rewrite the r.h.s of (2.51) using (2.47) as

$$\int_T^{T+\delta}\sigma(t,u)du=\left(1-\exp\left(-\int_T^{T+\delta}f(t,u)du\right)\right)\lambda(t,T;\delta).$$

Differentiating in T we obtain

$$\begin{aligned}\sigma(t,T+\delta)&=\sigma(t,T)+(f(t,T+\delta)-f(t,T))\exp\left(-\int_T^{T+\delta}f(t,u)du\right)\lambda(t,T;\delta)\\ &\quad+\left(1-\exp\left(-\int_T^{T+\delta}f(t,u)du\right)\right)\frac{\partial\lambda(t,T;\delta)}{\partial T}.\end{aligned}$$

This gives a recurrence relation. Once $\sigma(t,\cdot)$ is specified for $T\in[0,\delta)$ (typically *one put $\sigma(t,T)=0$ for $T \in [0,\delta)$) it can be solved by forward induction. This results in a complicated dependence of the HJM instantaneous volatility function $\sigma(t,T)$ on the forward rate curve. The question of whether the corresponding HJM equation has a unique and well-behaved solution was studied in [14] and also in [25].*

2.3.2 LIBOR: Forward measure dynamics

We are interested in finding the LIBOR rate dynamics $L^i(t)$ under the forward measure Q^{T_k} for the maturity T_k associated with the numeraire $P(\cdot, T_k)$.

Let us first consider the probability measure $\mathrm{Q}^{T_{i+1}}$ associated with $P(\cdot, T_{i+1})$, i.e. the price of the bond whose maturity coincides with the maturity of the LIBOR rate. Recall from (2.15) that the LIBOR rate $L^i(t)$ is a martingale under the measure $\mathrm{Q}^{T_{i+1}}$. Under the lognormal assumption (cf. (2.50)), we obtain the dynamics of $L^i(t)$ under $\mathrm{Q}^{T_{i+1}}$:

$$dL^i(t) = L^i(t)\lambda_i^\top(t)dW^{\mathrm{Q}^{T_{i+1}}}(t), \; t \le T_i. \tag{2.52}$$

Hence, the $\mathrm{Q}^{T_{i+1}}$-distribution of $\log L^i(T_i)$ conditional on $\mathcal{F}_t$ is Gaussian with mean

$$\log L^i(t) - \frac{1}{2}\int_t^{T_i} |\lambda_i(s)|^2 \, ds$$

and variance

$$\int_t^{T_i} |\lambda(s,T_i)|^2 \, ds.$$

Then, the time $t \le T_i$ price of a caplet (2.1.3) set at time T_i with payment date at T_{i+1}, with strike K and unit cap nominal value can be evaluated by Black's formula:

$$\begin{aligned} V_{caplet}(t) &= \delta P(t,T_{i+1}) E^{\mathrm{Q}^{T_{i+1}}} \left(L^i(T_i) - K\right)_+ \\ &= \delta P(t,T_{i+1}) \left(L^i(t)\Phi(d_1(t)) - K\Phi(d_2(t))\right), \end{aligned} \tag{2.53}$$

$$d_{1,2} = \frac{\log\left(\frac{L^i(t)}{K}\right) \pm \frac{1}{2}\int_t^{T_i} |\lambda(s,T_i)|^2 \, ds}{\sqrt{\int_t^{T_i} |\lambda(s,T_i)|^2 \, ds}},$$

where Φ is the standard Gaussian cumulative distribution function.

Let us now find the dynamics of $L^i(t)$ under a measure $\mathrm{Q}^{T_{k+1}}$ different from $\mathrm{Q}^{T_{i+1}}$. This can be done with the help of the Girsanov theorem. The systematic procedure of how asset price dynamics change when one changes the numeraire is described in [15].

First, notice that for $T_k > T_i$, $t \leq T_i$

$$\frac{P(t,T_{k+1})}{P(t,T_{i+1})} = 1/\prod_{j=i+1}^{k}\left(1+\delta L^j(t)\right),$$

and for $T_k < T_i$, $t \leq T_k$

$$\frac{P(t,T_{k+1})}{P(t,T_{i+1})} = \prod_{j=k+1}^{i}\left(1+\delta L^{s_j}(t)\right).$$

Let us denote the drift of the $L^i(t)$ dynamics under a measure $\mathrm{Q}^{T_{k+1}}$ by $\mu_L^{T_{k+1}}(t)$ and prove by backward iteration that

$$\mu_L^{T_{k+1}}(t) = \begin{cases} \sum\limits_{j=k+1}^{i} \frac{\delta L^j(t)}{1+\delta L^j(t)}\lambda_i^\top(t)\lambda_j(t), T_k < T_i, t \leq T_k, \\ -\sum\limits_{j=i+1}^{k} \frac{\delta L^j(t)}{1+\delta L^j(t)}\lambda_i^\top(t)\lambda_j(t),\ \ T_k > T_i,\ \ t \leq T_i. \end{cases} \tag{2.54}$$

We first show that (2.54) holds for the base case $k = i-1$. The Radon-Nykodim derivative process ζ_t relating the measure $\mathrm{Q}^{T_{i+1}}$ and Q^{T_i} (cf. (2.12)) is given by

$$\begin{aligned} \zeta_t &= E^{\mathrm{Q}^{T_{k+1}}}\left[\left.\frac{d\mathrm{Q}^{T_i}}{d\mathrm{Q}^{T_{i+1}}}\right|\mathcal{F}_t\right] \\ &= \frac{P(t,T_i)P(t_0,T_{i+1})}{P(t_0,T_i)P(t,T_{i+1})} = \left(1+\delta L^i(t)\right)\frac{P(t_0,T_{i+1})}{P(t_0,T_i)}, \end{aligned}$$

and hence by (2.52)

$$d\zeta_t = \frac{P(t_0,T_{i+1})}{P(t_0,T_i)}\delta dL^i(t) = \frac{P(t_0,T_{i+1})}{P(t_0,T_i)}L^i(t)\lambda_i^\top(t)dW^{\mathrm{Q}^{T_{i+1}}}(t),$$

or

$$\frac{d\zeta_t}{\zeta_t} = \frac{L^i(t)\lambda_i^\top(t)}{1+\delta L^i(t)} dW^{Q^{T_{i+1}}}(t).$$

From the Girsanov theorem, it follows that

$$dW^{Q^{T_i}}(t) = dW^{Q^{T_{i+1}}}(t) - \frac{L^i(t)\lambda_i(t)}{1+\delta L^i(t)} dt,$$

is a $Q^{T_i}-$ Brownian motion.

Hence,

$$\frac{dL^i(t)}{L^i(t)} = \frac{L^i(t)}{1+\delta L^i(t)}\lambda_i^\top(t)\lambda_i(t)dt + \lambda_i^\top(t)dW^{Q^{T_i}}(t),$$

and also

$$\frac{dL^{i-1}(t)}{L^{i-1}(t)} = -\frac{L^i(t)}{1+\delta L^i(t)}\lambda_{i-1}^\top(t)\lambda_i(t)dt + \lambda_{i-1}^\top(t)dW^{Q^{T_{i+1}}}(t).$$

For some general $k < i-1$, we have analogous to the basis case

$$d\zeta_t = \frac{P(t_0, T_{i+1})}{P(t_0, T_k)} \prod_{j=k+1}^{i} d(1+\delta L^j(t)).$$

Applying the product rule for Ito processes, we have

$$\begin{aligned}
&\prod_{j=k+1}^{i} d(1+\delta L^j(t)) \\
&= \prod_{j=k+1}^{i}(1+\delta L^j(t)) \sum_{j=k+1}^{i} \left(\frac{\delta dL^j(t)}{1+\delta L^j(t)} + \sum_{l=j+1}^{i} \frac{\delta dL^j(t)}{1+\delta L^j(t)} \frac{\delta dL^l(t)}{1+\delta L^l(t)} \right) \\
&= \prod_{j=k+1}^{i}(1+\delta L^j(t)) \sum_{j=k+1}^{i} \left\{ - \sum_{m=j+1}^{i} \frac{\delta L^j(t)\lambda_j^\top(t)}{1+\delta L^j(t)} \frac{L^m(t)\lambda_m(t)}{1+\delta L^m(t)} dt \right. \\
&\left. +\lambda_j^\top(t)dW^{Q^{T_{i+1}}}(t) + \sum_{l=j+1}^{i} \frac{\delta L^j(t)\lambda_j^\top(t)}{1+\delta L^j(t)} \frac{\delta L^l(t)\lambda_l(t)}{1+\delta L^l(t)} \right\}
\end{aligned}$$

$$= \prod_{j=k+1}^{i}(1+\delta L^j(t)) \sum_{j=k+1}^{i} \frac{\delta L^j(t)}{1+\delta L^j(t)} \lambda_j^\top(t)dW^{Q^{T_{i+1}}}(t).$$

From this we conclude that

$$\frac{d\zeta_t}{\zeta_t} = \sum_{j=k+1}^{i} \frac{\delta L^j(t)}{1+\delta L^j(t)} \lambda_j^\top(t) dW^{\mathrm{Q}^{T_{i+1}}}(t),$$

and hence

$$dW^{\mathrm{Q}^{T_{k+1}}}(t) = dW^{\mathrm{Q}^{T_{i+1}}}(t) - \sum_{j=k+1}^{i} \frac{\delta L^j(t)}{1+\delta L^j(t)} \lambda_j^\top(t) dW^{\mathrm{Q}^{T_{i+1}}}(t)dt.$$

Then, the dynamics of $L^i(t)$ under $\mathrm{Q}^{T_{k+1}}$ is given according to the following three cases

$$\frac{dL^i(t)}{L^i(t)} = \begin{cases} \sum\limits_{j=k+1}^{i} \frac{\delta L^j(t)}{1+\delta L^j(t)} \lambda_i^\top(t)\lambda_j(t)dt + \lambda_i^\top(t) dW^{\mathrm{Q}^{T_{k+1}}}(t), \ T_k < T_i, t \le T_k, \\ \lambda_i^\top(t) dW^{\mathrm{Q}^{T_{i+1}}}(t), \ T_k = T_i, \ t \le T_i, \\ -\sum\limits_{j=i+1}^{k} \frac{\delta L^j(t)}{1+\delta L^j(t)} \lambda_i^\top(t)\lambda_j(t)dt + \lambda_i^\top(t) dW^{\mathrm{Q}^{T_{k+1}}}(t), \ T_k > T_i, t \le T_i. \end{cases} \tag{2.55}$$

2.4 One factor short rate models

The earliest stochastic interest rate models were models of the short rates and they are still popular in the financial industry. They are all, however are HJM models and demonstrating that this is the case for a selected number of examples is the purpose of this section.

We recall from (2.10) that discount bond prices are given by the risk-neutral expectation

$$P(t,T) = E^{\mathrm{Q}}\left[\exp\left(-\int_t^T r(s)ds\right)\middle|\,\mathcal{F}_t\right]. \tag{2.56}$$

Thus, knowledge of the risk-neutral dynamics of the short rate process $r(t)$ is in principle, sufficient to compute time t discount bond prices for all maturities $T > t$.

One approach, for which evaluating expectation in (2.56) becomes particularly tractable, is to model the short rate as a Gaussian random process. We shall consider two examples of such models here, namely the Ho-Lee and the Vasicek models. One of the drawbacks of Gaussian short rate models is a positive probability of negative short rates. Another problem is the lack of interest rate dependence of the short rate volatility and hence no means to control the volatility skew implied by the model. We shall present one example of the model, namely the Cox-Ingersol-Ross model, which can address, at least partially, both of these shortcomings. We present below one-factor models only. For more details on the short-rate modeling approach (e.g. multifactor models) see e.g. [1, 15, 26, 65, 28, 68].

Example 2.4.1 *Ho-Lee model*

In the case of the Ho-Lee model, the dynamics of the short rate under the risk neutral measure Q are given by the stochastic differential equation

$$dr(t) = \theta(t)dt + \sigma_r dW^{\mathrm{Q}}(t), \;\; r(0) = r_{0,} \tag{2.57}$$

where $\theta(t)$ is a deterministic function and $\sigma_r > 0$ is constant.

It is clear that $-\int_t^T r(s)ds$ is Gaussian with mean conditional on $\mathcal{F}_t$ equal to

$$E\left(-\int_t^T r(s)ds \middle| \mathcal{F}_t\right) = -(T-t)r(t) - \int_t^T (T-u)\theta(u)du$$

and variance conditional on $\mathcal{F}_t$ equal to

$$Var\left(-\int_t^T r(s)ds \middle| \mathcal{F}_t\right) = \frac{1}{3}\sigma_r^2(T-t)^3.$$

Hence, from the basic moment properties of lognormal variables we have

$$P(t,T) = \exp\left(-(T-t)r(t) - \int_t^T (T-u)\theta(u)du + \frac{1}{6}\sigma_r^2(T-t)^3\right).$$

Then, using (2.2), we obtain

$$f(t,T) = r(t) + \int_t^T \theta(u)du - \frac{1}{2}\sigma_r^2(T-t)^2,$$

and hence

$$df(t,T) = \sigma_r^2(T-t)dt + \sigma_r dW^{\mathrm{Q}}(t). \tag{2.58}$$

We have, thus, established that the forward rate volatility in the Ho-Lee model is $\sigma(t,T) = \sigma_r$ (cf. (2.36)).

Example 2.4.2 *Vasicek model*

The Vasicek model assumes that the short rates follow a one-factor Ornstein-Uhlenbeck process with constant coefficients under the risk-neutral measure, that is

$$dr(t) = \kappa(\vartheta - r(t))dt + \sigma_r dW^{\mathrm{Q}}(t), \ r(0) = r_{0,} \tag{2.59}$$

where κ, ϑ and σ_r are positive constants.

Integrating equation (2.59), we obtain, for each $s \geq t$,

$$r(s) = r(t)\exp(-\kappa(s-t) + (1-\exp(-\kappa(s-t)))\,\vartheta + \sigma_r \int_t^s \exp(-\kappa(s-u)dW^{\mathrm{Q}}(u),$$

so that $r(s)$ conditional on $\mathcal{F}_t$ is normally distributed with mean and variance given by

$$E\left(r(s)|\,\mathcal{F}_t\right) = r(t)\exp(-\kappa(s-t)) + (1-\exp(-\kappa(s-t)))\,\vartheta, \tag{2.60}$$

$$Var\left(r(s)|\,\mathcal{F}_t\right) = \frac{\sigma_r^2}{2\kappa}\left[1-\exp(-2\kappa(s-t)\right]. \tag{2.61}$$

As a consequence of (2.60), the short rate is mean reverting in the sense that if an interest rate is high for historical reasons, it will most likely fall in the future (and vice versa if the interest rate is low). As $t \to \infty$, the mean of the short rate approaches ϑ and the variance goes to $\sigma_r^2/2\kappa$. Accordingly, ϑ is regarded as a *long term level* (also known as *mean reversion level*). The speed at which the short rate can be expected to revert to its long-term level is determined by κ, known as the *mean reversion speed.*

To establish a discount bond pricing formula (2.56) in the Vasicek model, we observe that $-\int_t^T r(s)ds$ is Gaussian with mean conditioned on $\mathcal{F}_t$ equal to

$$E\left(-\int_t^T r(s)ds \middle| \mathcal{F}_t\right) = -\vartheta(T-t) - (r(t)-\vartheta)\left(\exp(-\kappa(T-t)) - 1\right)/\kappa$$

and variance conditioned on $\mathcal{F}_t$ equal to

$$Var\left(-\int_t^T r(s)ds \middle| \mathcal{F}_t\right) = \frac{\sigma_r^2}{4\kappa^3}\left[-\exp(-2\kappa(T-t)) - 4\exp(-\kappa(T-t)) + 2(T-t)\kappa - 3\right].$$

From the standard properties of lognormal random variable, it follows that the discount bond prices (2.56) in the Vasicek model can be evaluated as

$$P(t,T) = \exp\left(A(t,T) + B(t,T)r(t)\right), \tag{2.62}$$

where

$$B(t,T) = \frac{(1-\exp(-\kappa(T-t)))}{\kappa},$$
$$A(t,T) = \left(\vartheta - \frac{\sigma_r^2}{2\kappa^2}\right)(B(t,T) - (T-t)) - \frac{\sigma_r^2 B(t,T)}{4\kappa^2}.$$

Finally, (2.2) and (2.62) yields the form of the initial forward curve

$$\begin{aligned} f_0(T) &= \exp(-\kappa(T-t_0))r_0 + (1-\exp(-\kappa(T-t_0)))\,\vartheta \\ &\quad -\frac{\sigma_r^2}{2\kappa^2}\left(1-\exp(-\kappa(T-t_0))\right)^2, \end{aligned} \tag{2.63}$$

and thus the forward rate volatility corresponding to the Vasicek model becomes

$$\sigma(t,T) = \sigma_r \exp(-\kappa(T-t)). \tag{2.64}$$

Example 2.4.3 *Cox-Ingersol-Ross (CIR) model*

A special case of time-homogeneous affine one-factor short rate models (i.e., models with the drift and square of the diffusion linear in $r(t)$) is the CIR model where the risk-neutral short rate dynamics are given by

$$dr(t) = \kappa(\theta - r(t))dt + \sigma_r \sqrt{r(t)} dW^{\mathbb{Q}}(t), \ r(0) = r_{0,} \tag{2.65}$$

where $\kappa > 0$, $\sigma_r > 0$, $r_{0,} > 0$ and θ are constants.

The regularity condition

$$2\kappa\theta > \sigma_r^2,$$

has to be imposed to ensure that zero is unattainable to the process (2.65), i.e. $r(t)$ stays strictly positive (see, e.g. [1, 15, 26]).

The process $r(t)$ features a non-central chi-squared distribution with $4\kappa\theta/\sigma_r^2$ degrees of freedom and non-centrality parameter $4\kappa r_0 \exp(-\kappa t)/\sigma_r^2(1 - \exp(-\kappa t))$.

By the Feynman-Kac formula, $P(t, x; T)$ satisfies the following PDE

$$\frac{\partial}{\partial t}P(t,x;T) + \kappa(\theta - x)\frac{\partial}{\partial x}P(t,x;T) + \frac{1}{2}\sigma_r^2 x \frac{\partial^2}{\partial x^2}P(t,x;T) - xP(t,x;T) = 0, \tag{2.66}$$

with the terminal condition

$$P(T, x; T) = 1.$$

The solution of this PDE problem is given by

$$P(t, x; T) = \exp\left(A(t, T) - B(t, T)x\right),$$

where $A(t, T)$ and $B(t, T)$ satisfy a system of Riccati ODEs (see details in [1, 15, 26])

$$\begin{aligned} \frac{dA}{dt} - \kappa\theta B &= 0, \\ -\frac{dB}{dt} + \frac{1}{2}\sigma_r^2 B^2 + \kappa B &= 1, \end{aligned}$$

with terminal conditions $A(T, T) = 0$, $B(T, T) = 0$.

Thus, by (2.2) the forward curve must evolve as

$$f(t, T) = r(t)\frac{\partial B}{\partial T}(t, T) - \kappa\theta \int_t^T \frac{\partial B}{\partial T}(s, T)ds,$$

and the corresponding forward rate volatility has the form

$$\sigma(t,T) = \sigma_r \sqrt{r(t)} \frac{\partial B}{\partial T}(t,T).$$

2.5 Summary

In this Chapter definitions and models from the interest rate theory which will be used throughout the thesis are presented. We introduce various interest rates such as forward, LIBOR and short rates (Section 2.1.1) along with some popular models for them (Sections 2.2, 2.3, 2.4 correspondingly). We recall the basic building blocks for most interest rate derivatives, namely, swaps, caps/floors and swaptions (Section 2.1.3). Some fundamental results from the theory of arbitrage-free pricing of contingent claims are also reviewed (Section 2.1.2).

The focus of the Chapter is on the Heath-Jarrow-Morton framework for modelling of instantaneous forward rates and the corresponding terminology. We highlight the key elements and assumptions of the HJM modelling philosophy. We start by presenting the general framework under the objective probability measure (Sections 2.2.1). Then, the arbitrage-free dynamics under risk-neutral (Sections 2.2.2) and forward measures (Sections 2.2.3) are derived.

We also explain how this HJM framework unifies other popular interest rates models. Namely, we demonstrate how the dynamics of LIBOR (Section 2.3.1) and short rates (Section 2.4) can be derived from the HJM framework.

Chapter 3

Numerical methods for Heath-Jarrow-Morton models

In this thesis we propose and analyze a new class of effective numerical methods for the HJM equation inspired by the idea of the method of lines (see, e.g. [71]). Our methods facilitate simulation of the HJM model under various specifications. The primary focus is weak-sense approximations which can be used for valuation of a broad class of interest rate products. To construct these numerical methods, we first discretise the HJM equation in the maturity time variable T applying quadrature rules to approximate the arbitrage-free drift (Section 3.4.1). In effect, this reduces the infinite-dimensional HJM equation to a finite-dimensional system of coupled stochastic differential equations (SDEs). In accordance with the method of lines, the maturity time T is interpreted as a "space" variable while the calendar time t is interpreted as a "time" variable. To obtain fully discrete methods (discrete in both T and t), we then approximate the obtained finite-dimensional system of SDEs in the weak (Section 3.4.2) and mean-square (Section 3.6) senses using the general theory of numerical integration of SDEs (Section (3.1.2) and also see, e.g. [45, 58, 57]). The proposed numerical algorithms (see Section 3.5 for realisations of the algorithms and Section 3.7 for numerical experiments) are computationally highly efficient due to the use of high-order quadrature rules. The quadrature rules of high orders allow us to use relatively large discretisation-steps in maturity time T without affecting the overall accuracy of the methods. More precisely, the number of forward rates that need to be approximated at each time moment t are significantly less in our algorithms than what is usually required when the time-grids for t and T coincide. Moreover, by capitalising on the method of lines, our numerical methods offer flexibility in choosing appropriate approximations in "space" and "time" separately. As we will see, in practice (see Remark 3.4.7), it is advantageous to use higher-order quadrature rules for integration with respect to maturity time T and lower-order numerical schemes for integration with respect to calendar time t.

We shall start this Chapter by presenting material on solutions of SDEs and their properties based on sources such as [3, 29, 49, 41, 70]. We consider "usual" SDEs and also a more general form of SDEs which initial value and coefficients depend on a parameter. We present results (see Section 3.1.1) on existence and uniqueness of the solutions of such equations and their differentiability with respect to a parameter if they dependent on it.

It is well-known that only some SDEs can be solved exactly and, in general, an approximation is required to obtain a numerical solution. In Section (3.1.2) we are going to be concerned with the numerical issues related to such approximations. We introduce criteria of mean-square and weak approximations and give examples of numerical schemes of various orders of accuracy. For a detailed and extensive study on numerical analysis of stochastic differential equations see [45, 58, 57].

3.1 Introduction to numerical methods for stochastic differential equations

We assume that the following assumptions hold throughout this section. As before let $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t_0 \leq t \leq t^*}, \mathrm{P}\right)$ be a filtered probability space satisfying the usual hypotheses and fix a time horizon $[t_0, t^*]$. For simplicity, we will drop the measure superscript on the expectation operator, so will write simply E for the expectation under measure P.

To keep this subsection of a manageable size, we will only present the material relevant to this thesis and forego to carry out the proofs of the theorems, as long as they can be found exactly as stated in the literature.

3.1.1 The solution of stochastic differential equations

We consider a vector-valued SDE in the Ito sense

$$X(t) = x_0 + \int_{t_0}^{t} a(s, X(s))ds + \sum_{i=1}^{d} \int_{t_0}^{t} b_i(s, X(s))dW_i(s), \tag{3.1}$$

or in an abbreviated differential notation

$$dX(t) = a(t, X(t))dt + \sum_{i=1}^{d} b_i(t, X(t))dW_i(t), \tag{3.2}$$

$$X(t_0) = x_0, \tag{3.3}$$

where $X(t)$ is an $\mathbb{R}^n$-valued stochastic process defined on $[t_0, t^*]$; $W(t) = (W_1(t), \ldots, W_d(t))^\top$ is a d-dimensional standard Wiener process; $a : [t_0, t^*] \times \mathbb{R}^n \to \mathbb{R}^n$ is a measurable function with $\int_{t_0}^{t} |a(s, X(s))| \, ds < \infty$ a.s. for any $t \in [t_0, t^*]$; $b_i : [t_0, t^*] \times \mathbb{R}^n \to \mathbb{R}^n$ are measurable functions with $\int_{t_0}^{t} |b_i(s, X(s))|^2 \, ds < \infty$, $i = 1, \ldots, d$ a.s. for any $t \in [t_0, t^*]$; x_0 is an $\mathcal{F}_{t_0}$-measurable $\mathbb{R}^n$-valued random variable.

The theory of such equations and their solutions being stochastic processes can be found, e.g. in [3, 29, 41, 70]. Let us give the definition of the solution of equation (3.2), (3.3) (see [41, pp. 48]).

Definition 3.1.1 *An $\mathbb{R}^n$-valued stochastic process $\{X(t)\}_{t_0 \le t \le t^*}$ is called a solution of equation* (3.2), (3.3) *if it has the following properties:*

(i) *$\{X(t)\}$ is continuous and $\mathcal{F}_t-$adapted;*

(ii) *equation* (3.1) *holds for every $t \in [t_0, t^*]$ with probability 1.*

A solution $\{X(t)\}$ is said to be unique if any other solution $\{\hat{X}(t)\}$ is indistinguishable from $\{X(t)\}$, that is

$$\mathrm{P}(X(t) = \hat{X}(t) \text{ for all } t_0 \leq t \leq t^*) = 1.$$

We shall now state the conditions that guarantee the existence and uniqueness of the solution to equation (3.2), (3.3). The proof of this result can be found, e.g. in ([41, pp. 51]).

Theorem 3.1.2 *(existence and uniqueness) Suppose $a(t,x)$ and $b(t,x)$ satisfy the Lipschitz and linear growth conditions*

$$|a(t,x) - a(t,y)| + |b(t,x) - b(t,y)| \leq K\,|x-y|, \tag{3.4}$$

$$|a(t,x)|^2 + |b(t,x)|^2 \leq K^2\left(1 + |x|^2\right), \tag{3.5}$$

for all $t \in [t_0, t^*]$ and $x, y \in R^n$, where K is a constant. Then the equation (3.2), (3.3) has a unique $\mathbb{R}^n$-valued solution $X(t)$ on $[t_0, t^*]$.

We shall now give estimates for the moments of the solution of (3.2), (3.3) that will be used in the thesis (see e.g. [3, pp. 116] for the proof).

Theorem 3.1.3 *Suppose that the assumptions of Theorem 3.1.2 are satisfied and that*

$$E\,|x_0|^{2m} < \infty,$$

where m is a positive integer. Then, for the solution $X(t)$ to equation (3.2), (3.3) on $[t_0, t^*]$,

$$E\,|X(t)|^{2m} \leq \left(1 + E\,|x_0|^{2m}\right) e^{C(t-t_0)},$$

where $C = 2m(2m+1)K^2$ and K from (3.4)-(3.5).

Let us consider SDE of the form (3.1) whose initial value x_0 and coefficients a and b_i depend on a parameter ϑ which varies through some set of numbers $\Theta \subset \mathbb{R}$, i.e., we consider SDE of the form

$$X(\vartheta,t) = x_0(\vartheta) + \int\limits_{t_0}^{t} a(\vartheta,s,X(\vartheta,s))ds + \sum_{i=1}^{d} \int\limits_{t_0}^{t} b_i(\vartheta,s,X(\vartheta,s))dW_i(s), \qquad (3.6)$$

where for each $\vartheta \in \Theta$, $a : \Theta \times [t_0, t^*] \times \mathbb{R}^n \to \mathbb{R}^n$ is a measurable function with $\int_{t_0}^{t} |a(\vartheta, s, X(s))| \, ds < \infty$ a.s. for any $t \in [t_0, t^*]$; $b_i : \Theta \times [t_0, t^*] \times \mathbb{R}^n \to \mathbb{R}^n$ are measurable functions with $\int_{t_0}^{t} |b_i(\vartheta, s, X(s))|^2 \, ds < \infty$, $i = 1, \ldots, d$ a.s. for any $t \in [t_0, t^*]$; $x_0(\vartheta)$ is an $\mathcal{F}_0$-measurable $\mathbb{R}^n$-valued random variable.

We fix $\vartheta \in \Theta$ and assume that conditions of Theorem 3.1.2 are satisfied. Then the equation (3.6) has on $[t_0, t^*]$ a unique $\mathbb{R}^n$-valued solution $X(t, \vartheta)$, continuous in t with probability 1.

We shall now be interested in the question of differentiability of the solution of SDE with respect to a parameter. We present the statement of the theorem which addresses this issue, for its proof see [49, pp. 105].

Theorem 3.1.4 *Suppose that the process $x_0(\vartheta)$ is j times (continuously) differentiable at a point $\vartheta_0 \in \Theta$, and that the functions $a(\vartheta, s, y)$, $b_i(\vartheta, s, y)$ are j times continuously differentiable with respect to ϑ, y. Furthermore, assume that all derivatives of the foregoing functions, up to order j inclusive, do not exceed $K(1 + |y|)^m$ for any $\vartheta \in \Theta$, y. Then the process $X(\vartheta, t)$ is j times (continuously) differentiable at the point ϑ_0.*

3.1.2 Mean-square and weak approximations for stochastic differential equations

The material presented in this part of the thesis follows closely [58] to which we refer for a comprehensive study on stochastic numerics.

We consider an equally-spaced grid for time t with step $h = (t^* - t_0)/M$:

$$t_0 < \cdots < t_M = t^*, \quad t_k = kh, \quad k = 0, \ldots, M. \qquad (3.7)$$

We use a constant step h to keep notation manageable. All results in this section, however, are easily extendable to the non-equidistant grids.

A numerical method for the approximation of the solution of (3.2), (3.3) on the grid (3.7) is an algorithm that produces a set of discrete values $\left\{\bar{X}(t_k)\right\}$, $k = 0, \ldots, M$ which approximate the solution $X(t)$ evaluated on the grid (3.7).

Definition 3.1.5 *(Mean-square convergence) We say that a numerical method has a mean-square order of accuracy $q > 0$ if there exists a positive constant K independent of k and h such that*

$$\left(E\left|X(t_k) - \bar{X}(t_k)\right|^2\right)^{\frac{1}{2}} \leq Kh^q. \tag{3.8}$$

Remark 3.1.6 *Often the notion of strong order of accuracy is used: if for some numerical method*

$$\left(E\left|X(t_k) - \bar{X}(t_k)\right|\right) \leq Kh^q, \tag{3.9}$$

where K is a positive constant independent of k and h, then we say that the strong order of accuracy of the method is equal to q. Clearly, if the mean-square order of a method is q, then the method has the same strong order.

In what follows, we will denote by $X(t_k)$ or X_k the solution of (3.2), (3.3) evaluated at t_k and an approximation of $X(t_k)$ by $\bar{X}(t_k)$ or simply by $\bar{X}_k$. Also, $X_{t_k,X}(t)$ denotes the solution of (3.2), (3.3) for $t_k \leq t \leq t^*$ satisfying the initial condition $X(t_k) = X$ at time t_k, where X is an $\mathcal{F}_{t_k}$–measurable random variable with finite second moment, i.e. $E\left|X\right|^2 < \infty$.

The *one-step approximation* $\bar{X}_{t,x}(t+h)$, $t_0 \leq t < t+h \leq t^*$, depends on x, t, h, and $\{W_1(\theta) - W_1(t), \ldots, W_d(\theta) - W_d(t),\ t \leq \theta \leq t+h\}$ and it is defined as follows:

$$\bar{X}_{t,x}(t+h) = x + A(t, x, h; W_i(\theta) - W_i(t),\ i = 1, \ldots, d,\ t \leq \theta \leq t+h), \tag{3.10}$$

where by $\bar{X}_{t,x}(t+h)$ we denote the approximation of the solution at step $t+h$ such that $\bar{X}(t) = x$ and A is a vector function of dimension n.

Now based on the one-step approximation we can recurrently construct the approximations $\left(\bar{X}_k, \mathcal{F}_{t_k}\right)$, $k = 0, \ldots, M$, $t_{k+1} - t_k = h_{k+1}$, $t_M = t^*$:

$$\bar{X}_0 = X(t_0) = x_0, \tag{3.11}$$
$$\bar{X}_{k+1} = \bar{X}_{t_k, \bar{X}_k}(t_{k+1}) = \bar{X}_k$$
$$+A(t_k, \bar{X}_k, h_{k+1}; W_i(\theta) - W_i(t),\ i = 1, \ldots, d,\ t_k \leq \theta \leq t_{k+1}),$$

where by $\bar{X}_{t_k,X}(t_{k+1})$ denotes the approximation of the solution at step $k+1$ satisfying the following condition at step $k : \bar{X}(t_k) = X$. Clearly,

$$\bar{X}_{k+1} = \bar{X}_{t_k, \bar{X}_k}(t_{k+1}) = \bar{X}_{t_0, \bar{X}_0}(t_{k+1}).$$

We shall now provide the statement of the theorem which relates properties of a one-step approximation with the mean-square order of convergence of the corresponding numerical scheme. The proof of this theorem can be found in [58, Chapter 1] (also see [56, 57]).

Theorem 3.1.7 *Suppose the one-step approximation $\bar{X}_{t,x}(t+h)$ has order of accuracy q_1 for the expectation and order q_2 for the mean-square deviation; more precisely, for arbitrary $t_0 \leq t \leq t^* - h$, $x \in \mathbb{R}^n$ the following inequalities hold*

$$\left| E\left(X_{t,x}(t+h) - \bar{X}_{t,x}(t+h)\right)\right| \leq K(1+|x|^2)^{\frac{1}{2}} h^{q_1}, \tag{3.12}$$

$$\left[E\left|X_{t,x}(t+h) - \bar{X}_{t,x}(t+h)\right|^2\right]^{\frac{1}{2}} \leq K(1+|x|^2)^{\frac{1}{2}} h^{q_2}, \tag{3.13}$$

also let

$$q_2 \geq \frac{1}{2},\ q_1 \geq q_2 + \frac{1}{2}.$$

Then for any M and $k = 0, \ldots, M$ the following inequality holds:

$$\left[E\left|X_{t_0,X_0}(t_k) - \bar{X}_{t_0,X_0}(t_k)\right|^2\right]^{\frac{1}{2}} \leq K(1+|X_0|^2)^{\frac{1}{2}} h^{q_2 - \frac{1}{2}},$$

i.e. the mean-square order of accuracy of the method constructed using the one-step approximation $\bar{X}_{t,x}(t+h)$ *is* $q = q_2 - \frac{1}{2}$.

Mean-square convergence is a typical criteria in applications concerned with scenario simulation where it is necessary to simulate approximate trajectories of solutions of SDEs. However, in finance quite often only moments of a function of SDE solutions are of interest. In pricing of contingent claims, it is the expectation of a certain payoff function that one wants to approximate (cf. (2.9), (2.11)). Thus, convergence in distribution of the numerical solutions will be suffice. In such applications the weak convergence is considered. The key results on weak approximations were obtained in [54, 55, 57, 78, 64].

We say that a function $g(x)$ belongs to class $\mathbf{F}$ if we can find constants $K > 0$, $\varkappa > 0$ such that for all $x \in \mathbb{R}^n$ the following inequality holds

$$|g(x)| \leq K\left(1 + |x|^{\varkappa}\right). \tag{3.14}$$

If a function $g(x, s)$ also depends on parameter $s \in \mathcal{S}$, then we say that $g(x, s)$ belongs to $\mathbf{F}$ (with respect to the variable x) if the inequality of the type (3.14) holds uniformly in $s \in \mathcal{S}$.

Definition 3.1.8 *(Weak convergence) We say that a numerical method has weak order of accuracy* $q > 0$ *if there exists a positive constant* K *independent of* k *and* h *such that*

$$\left|Eg(X(t_k)) - Eg(\bar{X}(t_k))\right| \leq Kh^q. \tag{3.15}$$

Note that numerical integration in the mean-square sense with some order of accuracy guarantees an approximation in the weak sense with the same order of accuracy, since if $\left(E\left|X(t_k) - \bar{X}(t_k\right|^2\right)^{\frac{1}{2}} = O(h^q)$ then for every function satisfying a Lipschitz condition we have $Eg(X(t_k)) - Eg(\bar{X}(t_k) = O(h^q)$.

The one step weak approximation $\bar{X}_{t,x}(t+h)$ of the solution $X_{t,x}(t+h)$ can be constructed by comparing moments up to sufficient order of the vector $\bar{X}_{t,x}(t+h)-x$ and the corresponding moments of the vector $X_{t,x}(t+h) - x$.

Let us along with (3.2) , (3.3) consider the one-step approximation of the form

$$\bar{X}_{t,x}(t+h) = x + A(t,x,h;\xi), \tag{3.16}$$

where ξ is a random variable (in general a vector) having moments up to sufficiently high order and A is a vector function of dimension n.

According to (3.16), we construct the approximating sequence

$$\bar{X}_0 = X_0 = X(t_0),$$

$$\bar{X}_{k+1} = \bar{X}_k + A(t,\bar{X}_k,h;\xi),\ \ k = 0,\ldots,M, \tag{3.17}$$

where ξ_0 is independent of $\bar{X}_0$, while ξ_k for $k > 0$ is independent of $\bar{X}_0,\ldots,\bar{X}_k$, $\xi_0,\ldots,\xi_{k-1}$.

We denote by $\delta X = X - x = X_{t,x}(t+h) - x$, $\delta\bar{X} = \bar{X} - x = \bar{X}_{t,x}(t+h) - x$ and by δX^i, $\delta\bar{X}^i$ the i-th component of the vectors δX and $\delta\bar{X}$, correspondingly.

Next, we give the statement of the theorem which relates properties of a one-step approximation with the weak order of convergence of the corresponding numerical scheme. The proof of this theorem can be found in [58, Chapter 2] (see also [55, 57, 78]).

Theorem 3.1.9 *Suppose that*

(a) the coefficients of equation (3.2) *are continuous, satisfy a Lipschitz condition (3.4) and together with their partial derivatives with respect to* x *of order up to* $2p+2$, *inclusively, belong to* $\mathbf{F}$;

(b) the method (3.16) is such that

$$\left| E\left(\prod_{j=1}^{s}\delta X^{i_j} - \prod_{j=1}^{s}\delta\bar{X}^{i_j}\right)\right| \le K(x)h^{q+1},\ \ s = 1,\ldots,2p+1,\ \ K(x)\in\mathbf{F} \tag{3.18}$$

$$E\prod_{j=1}^{2p+2}\left|\delta\bar{X}^{i_j}\right| \le K(x)h^{q+1},\ \ i_j = 1,\ldots,n,\ \ K(x)\in\mathbf{F}; \tag{3.19}$$

(c) the function $g(x)$ together with its partial derivatives of order up to $2p+2$, inclusively, belong to $\mathbf{F}$*;*

(d) for a sufficiently large m the expectation $E\left|\bar{X}_k\right|^{2m}$ exist and are uniformly bounded with respect to M and $k = 0, \ldots, M$.

Then, for all M and all $k = 0, \ldots, M$ the following inequality holds:

$$\left|Eg(X_{t_0,X_0}(t_k)) - Eg(\bar{X}_{t_0,X_0}(t_k))\right| \le Kh^q,$$

i.e., the method (3.17) has order of accuracy q in the sense of weak approximations.

We shall now give examples of some numerical algorithms for SDE (3.2) – (3.3). For this purpose let us recall the Ito-Taylor expansion of solutions of SDE (see, e.g. [45, 58, 57]). For the clarity of presentation, we start with the integral representation of (3.2) :

$$X_{t,x}(t+h) = x + \int\limits_t^{t+h} a(s, X(s))ds + \sum_{i=1}^{d} \int\limits_t^{t+h} b_i(s, X(s))dW_i(s). \tag{3.20}$$

Applying Ito's formula to $a(s, X(s))$ gives

$$a(s, X(s)) = a(t,x) + \int\limits_t^{s} La(u, X(u))du + \sum_{i=1}^{d} \int\limits_t^{s} \Lambda_i a(u, X(u))dW_i(u), \tag{3.21}$$

where the operators L and Λ_i, $i = 1, \ldots, d$ are given by

$$L = \frac{\partial}{\partial t} + a^\top \frac{\partial}{\partial x} + \frac{1}{2} \sum_{i=1}^{d} \sum_{m=1}^{n} \sum_{j=1}^{n} b_i^m b_i^j \frac{\partial^2}{\partial x_m \partial x_j},$$

$$\Lambda_i = \sum_{j=1}^{n} b_i^j \frac{\partial}{\partial x_j}.$$

Similarly,

$$b_i(s, X(s)) = b_i(t,x) + \int\limits_t^{s} Lb_i(u, X(u))du + \sum_{j=1}^{d} \int\limits_t^{s} \Lambda_j b_i(u, X(u))dW_j(u). \tag{3.22}$$

Plugging (3.21) and (3.22) into (3.20), we obtain

$$X_{t,x}(t+h) = x + a(t,x)h + \sum_{i=1}^{d} b_i(t,x)(W_i(t+h) - W_i(t)) + \rho_1, \tag{3.23}$$

where

$$\begin{aligned}
\rho_1 = & \int_t^{t+h} \left(\int_t^s La(u, X(u))du \right) ds + \sum_{i=1}^{d} \int_t^{t+h} \left(\int_t^s \Lambda_i a(u, X(u))dW_i(u) \right) ds \\
& + \sum_{i=1}^{d} \int_t^{t+h} \left(\int_t^s Lb_i(u, X(u))du \right) dW_i(s) \qquad (3.24)\\
& + \sum_{i=1}^{d} \sum_{j=1}^{d} \int_t^{t+h} \left(\int_t^s \Lambda_j b_i(u, X(u))dW_j(u) \right) dW_i(s).
\end{aligned}$$

We can repeat this procedure arbitrary many times assuming a and b_i are sufficiently smooth. Just applying it one more time we arrive at

$$\begin{aligned}
& X_{t,x}(t+h) \\
= \; & x + a(t,x)h + \sum_{i=1}^{d} b_i(t,x)(W_i(t+h) - W_i(t)) \\
& + La(t, X(t))\frac{h^2}{2} + \sum_{i=1}^{d} \Lambda_i a(t, X(t)) \int_t^{t+h} (W_i(s) - W_i(t)) \; ds \\
& + \sum_{i=1}^{d} Lb_i(t, X(t)) \int_t^{t+h} (s-t) \; dW_i(s) \\
& + \sum_{i=1}^{d} \sum_{j=1}^{d} \Lambda_j b_i(t, X(t)) \int_t^{t+h} (W_j(s) - W_j(t)) \, dW_i(s) + \rho_2, \qquad (3.25)
\end{aligned}$$

$$\begin{aligned}
\rho_2 = & \int_t^{t+h} \int_t^s \left(\int_t^u L^2 a(u_1, X(u_1))du_1 \right. \\
& \left. + \sum_{i=1}^{d} \int_t^u \Lambda_i La(u_1, X(u_1))dW_i(u_1) \right) du \; ds
\end{aligned}$$

$$+\sum_{i=1}^{d}\int_t^{t+h}\int_t^{s}\left(\int_t^{u} L\Lambda_i a(u_1,X(u_1))du_1\right.$$

$$\left.+\sum_{j=1}^{d}\int_t^{u}\Lambda_j\Lambda_i a(u_1,X(u_1))dW_j(u_1)\right)dW_i(u)\ ds$$

$$+\sum_{i=1}^{d}\int_t^{t+h}\int_t^{s}\left(\int_t^{u} L^2 b_i(u_1,X(u_1))du_1\right.$$

$$\left.+\sum_{j=1}^{d}\int_t^{u}\Lambda_j L b_i(u_1,X(u_1))dW_j(u_1)\right)du\ dW_i(s)$$

$$+\sum_{i=1}^{d}\sum_{j=1}^{d}\int_t^{t+h}\int_t^{s}\int_t^{u} L\Lambda_j b_i(u_1,X(u_1))\ du_1\ dW_j(u)\ dW_i(s)$$

$$+\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{m=1}^{d}\int_t^{t+h}\int_t^{s}\int_t^{u}\Lambda_i\Lambda_j b_k(u_1,X(u_1))\ dW_m(u_1)\ dW_j(u)\ dW_i(s).$$

The truncated Ito-Taylor expansion based on (3.23):

$$\bar{X}_{t,x}(t+h)=x+a(t,x)h+\sum_{i=1}^{d}b_i(t,x)(W_i(t+h)-W_i(t)), \tag{3.26}$$

corresponds to the *one-step Euler approximation.*

By (3.11), this approximation generates the *explicit Euler method:*

$$\bar{X}_0=x_0,$$

$$\bar{X}_{k+1}=\bar{X}_k+a_k h+\sum_{i=1}^{d}b_{ik}(W_i(t_k+h)-W_i(t_k)), \tag{3.27}$$

where a_k, b_{jk} are the values of the coefficients a and b_j evaluated at the point (t_k, X_k).

To establish the mean square order of accuracy, we shall find q_1 and q_2 as in Theorem 3.1.7. We assume that functions $a(t,x)$ and $b(t,x)$ have partial derivatives with respect to t up to order one and with respect to x up to order two which belong to $\mathbf{F}$.

From (3.23) and (3.26), we have

$$\left|E\left(X_{t,x}(t+h)-\bar{X}_{t,x}(t+h)\right)\right| = |E\rho_1|.$$

The form of the remainder ρ_1 (3.24) and the fact that $La \in \mathbf{F}$ imply that we can find an even number $2m$ and a number $K > 0$ such that

$$\begin{aligned} |E\rho_1| &= \left| E \int_t^{t+h} \left(\int_t^s La(u, X(u))du \right) ds \right| \\ &\leq \left| \int_t^{t+h} \left(\int_t^s K\left(1 + E\,|X(u)|^{2m}\right) \right) ds \right|, \end{aligned}$$

Moreover, with help of Theorem 3.1.3, we conclude

$$|E\rho_1| \leq K(x)h^2, \;\; K(x) \in \mathbf{F}, \tag{3.28}$$

i.e. $q_1 = 2$. Next, to find q_2, we estimate the mean-square deviation of the one-step approximation $\bar{X}_{t,x}(t+h)$ as follows:

$$E\left|X_{t,x}(t+h)-\bar{X}_{t,x}(t+h)\right|^2 = E\,|\rho_1|^2 \tag{3.29}$$

$$\begin{aligned} \leq K \Bigg(& E\left| \int_t^{t+h} \left(\int_t^s La(u, X(u))du \right) ds \right|^2 + E\left| \sum_{i=1}^d \int_t^{t+h} \left(\int_t^s \Lambda_i a(u, X(u))dW_i(u) \right) ds \right|^2 \\ & + E\left| \sum_{i=1}^d \int_t^{t+h} \left(\int_t^s Lb_i(u, X(u))du \right) dW_i(s) \right|^2 \\ & + E\left| \sum_{i=1}^d \sum_{j=1}^d \int_t^{t+h} \left(\int_t^s \Lambda_j b_i(u, X(u))dW_j(u) \right) dW_i(s) \right|^2 \Bigg) \end{aligned}$$

Let us estimate each term in (3.29) separately based on the assumption that $La \in \mathbf{F}$, $Lb_i \in \mathbf{F}$, $\Lambda_i a \in \mathbf{F}$, $\Lambda_j b_i \in \mathbf{F}$, also using the Ito isometry, Cauchy–Bunyakovsky inequality and Theorem 3.1.3:

$$
\begin{aligned}
&E\left|\int\limits_t^{t+h}\left(\int\limits_t^s La(u,X(u))du\right)ds\right|^2 \\
\le\ & hE\left(\int\limits_t^{t+h}\left(\int\limits_t^s\left(1+|X(u)|^\kappa\right)du\right)^2 ds\right) \le Kh^3\int\limits_t^{t+h}\int\limits_t^s\left(1+E\left|X(u)\right|^{2\kappa}\right)du\ ds \\
\le\ & K(x)h^5,\ K(x)\in\mathbf{F}; \qquad (3.30)
\end{aligned}
$$

$$
\begin{aligned}
&E\left|\sum_{i=1}^d\int\limits_t^{t+h}\left(\int\limits_t^s \Lambda_i a(u,X(u))dW_i(u)\right)ds\right|^2 \\
\le\ & KE\sum_{i=1}^d\left(\int\limits_t^{t+h}\left(\int\limits_t^s \Lambda_i a(u,X(u))dW_i(u)\right)ds\right)^2 \\
\le\ & KhE\sum_{i=1}^d\int\limits_t^{t+h}\left(\int\limits_t^s \Lambda_i a(u,X(u))dW_i(u)\right)^2 ds \\
\le\ & Kh\sum_{i=1}^d\int\limits_t^{t+h}\int\limits_t^s\left(1+E\left|X(u)\right|^{2\kappa}\right)du\ ds \le K(x)h^3,\ K(x)\in\mathbf{F}; \qquad (3.31)
\end{aligned}
$$

$$
\begin{aligned}
&E\left|\sum_{i=1}^d\int\limits_t^{t+h}\left(\int\limits_t^s Lb_i(u,X(u))du\right)dW_i(s)\right|^2 \\
\le\ & KE\sum_{i=1}^d\left(\int\limits_t^{t+h}\left(\int\limits_t^s Lb_i(u,X(u))du\right)dW_i(s)\right)^2 \\
\le\ & K\sum_{i=1}^d\left(\int\limits_t^{t+h}E\left(\int\limits_t^s Lb_i(u,X(u))du\right)^2 ds\right) \\
\le\ & Kh\sum_{i=1}^d\left(\int\limits_t^{t+h}\int\limits_t^s\left(1+E\left|X(u)\right|^{2\kappa}\right)du\ ds\right) \le K(x)h^3,\ K(x)\in\mathbf{F}; \qquad (3.32)
\end{aligned}
$$

$$
E\left|\sum_{i=1}^{d}\sum_{j=1}^{d}\int\limits_{t}^{t+h}\left(\int\limits_{t}^{s}\Lambda_j b_i(u,X(u))dW_j(u)\right)dW_i(s)\right|^2
$$
$$
\leq \; KE\sum_{i=1}^{d}\sum_{j=1}^{d}\left(\int\limits_{t}^{t+h}\left(\int\limits_{t}^{s}\Lambda_j b_i(u,X(u))dW_j(u)\right)dW_i(s)\right)^2
$$
$$
\leq \; K\sum_{i=1}^{d}\sum_{j=1}^{d}\left(\int\limits_{t}^{t+h}\int\limits_{t}^{s}\left(1+E\left|X(u)\right|^{2\kappa}\right)du\; ds\right) \leq K(x)h^2,\; K(x)\in\mathbf{F}. \quad (3.33)
$$

Based on estimates (3.30)-(3.33), we conclude that

$$
E\left|\rho_1\right|^2 \leq K(x)h^2,\; K(x)\in\mathbf{F}, \tag{3.34}
$$

i.e. $q_2 = 1$. Therefore, according to Theorem 3.1.7 the Euler method (3.27) has mean-square order of convergence equal to 1/2. In the case of additive noise, i.e. when $b_k(t,x) \equiv b_k(t)$, based on (3.31),(3.32) we find that $q_2 = 3/2$. This implies that the Euler method (3.27) for systems with additive noise is of the first mean-square order.

Additionally assuming that $a(t,x)$ and $b(t,x)$ have partial derivatives in x up to order four that belong to $\mathbf{F}$, we shall demonstrate that the Euler method has the first order of weak convergence. Specifically, by Theorem 3.1.9 we should establish that the following inequalities hold

$$
\left|E\left(\prod_{j=1}^{s}\delta X^{i_j}-\prod_{j=1}^{s}\delta\bar{X}^{i_j}\right)\right| \leq K(x)h^2,\; s=1,\ldots,3,\; K(x)\in\mathbf{F}; \tag{3.35}
$$

$$
E\prod_{j=1}^{4}\left|\delta\bar{X}^{i_j}\right| \leq K(x)h^2,\; i_j=1,\ldots,n,\; K(x)\in\mathbf{F},
$$

where

$$
\delta X^{i_j} \;=\; X^{i_j}_{t,x}(t+h)-x^{i_j} = a^{i_j}(t,x)h+\sum_{k=1}^{d}b^{i_j}_k(t,x)\int\limits_{t}^{t+h}dW_k(s)ds+\rho^{i_j}_1,
$$
$$
\delta\bar{X}^{i_j} \;=\; \bar{X}^{i_j}_{t,x}(t+h)-x^{i_j} = a^{i_j}(t,x)h+\sum_{k=1}^{d}b^{i_j}_k(t,x)\int\limits_{t}^{t+h}dW_k(s)ds,
$$

and ρ_1 is from (3.24).

To show that the inequalities under consideration hold, we will use the following estimate

$$\left| E\rho_1 \int\limits_t^{t+h} dW_k(s) \right| \leq K(x)h^2, \;\; K(x) \in \mathbf{F}, \tag{3.36}$$

To proof (3.36), we note that the expectation of the norm of the first three terms in ρ_1 multiplied by $\int\limits_t^{t+h} dW_k(s)$ has at least the second order of smallness with respect to h. This follows from the application of the Cauchy–Bunyakovsky inequality to this product and, subsequently, (3.30)-(3.32). To show that this also holds for the last term in ρ_1 multiplied by $\int\limits_t^{t+h} dW_k(s)$, we apply Ito's formula to this last term in ρ_1:

$$\begin{aligned} &\sum_{i=1}^{d}\sum_{j=1}^{d} \int\limits_t^{t+h} \left(\int\limits_t^{s} \Lambda_j b_i(u, X(u)) dW_j(u) \right) dW_i(s) \qquad (3.37) \\ = \;\; &\sum_{i=1}^{d}\sum_{j=1}^{d} \Lambda_j b_i(t, X(t)) \int\limits_t^{t+h}\int\limits_t^{s} dW_j(u)\; dW_i(s) \\ &+\sum_{i=1}^{d}\sum_{j=1}^{d} \int\limits_t^{t+h}\int\limits_t^{s}\int\limits_t^{u} L\Lambda_j b_i(u_1, X(u_1))\; du_1\; dW_j(u)\; dW_i(s) \\ &+\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{m=1}^{d} \int\limits_t^{t+h}\int\limits_t^{s}\int\limits_t^{u} \Lambda_i\Lambda_j b_k(u_1, X(u_1))\; dW_m(u_1)\; dW_j(u)\; dW_i(s). \end{aligned}$$

All terms in (3.37), except of the first one, have order of smallness at least $3/2$. Using the Cauchy–Bunyakovsky inequality, we can readily show that the expectation of absolute value of the product of each of such terms with $\int\limits_t^{t+h} dW_k(s)$ is smaller or equal to $K(x)h^2$, $K(x) \in \mathbf{F}$. Let us now prove that the expectation of the product of the first term in (3.37) with $\int\limits_t^{t+h} dW_k(s)$ is zero, i.e.

$$E\left(\sum_{i=1}^{d}\sum_{j=1}^{d} \Lambda_j b_i(t, X(t)) \int\limits_t^{t+h}\int\limits_t^{s} dW_j(u)\; dW_i(s) \cdot \int\limits_t^{t+h} dW_k(s) \right) = 0. \tag{3.38}$$

Indeed, by changing the variables

$$V_k = -W_k,\ k = 1, \ldots, d,$$

where V_k are independent Wiener processes. Since the number of Wiener processes participating in (3.38) is odd,. we have

$$\begin{aligned} & E\left(\int_t^{t+h} \int_t^{s} dW_j(u)\ dW_i(s) \cdot \int_t^{t+h} dW_k(s) \middle| \mathcal{F}_t \right) \\ = \; & -E\left(\int_t^{t+h} \int_t^{s} dV_j(u)\ dV_i(s) \cdot \int_t^{t+h} dV_k(s) \middle| \mathcal{F}_t \right), \end{aligned}$$

which implies (3.38). Establishing (3.38) completes the proof of (3.36)

For $s = 1$, i.e. we are considering first moments, (3.35) follows from (3.28). For $s = 2$, (3.35) holds, since we have

$$\begin{aligned} & \left| E\left(\prod_{j=1}^{2} \delta X^{i_j} - \prod_{j=1}^{2} \delta \bar{X}^{i_j} \right) \right| \\ = \; & \left| E \sum_{j=1}^{2} \left(\rho_1^{i_j} a^{i_j}(t,x)h + \rho_1^{i_j} \sum_{k=1}^{d} b_k^{i_j}(t,x) \int_t^{t+h} dW_k(s)ds \right) + \rho_1^{i_1}\rho_1^{i_2} \right| \\ \leq \; & K(x)h^3 + K(x)h^2 + K(x)h^2 \leq K(x)h^2,\ K(x) \in \mathbf{F}, \end{aligned}$$

where we used (3.28), (3.34), (3.36) and the Cauchy–Bunyakovsky inequality. This also already makes clear that for $s = 3$ all terms contain at least one component of ρ_1 as a factor and, hence, have order of smallness at least 2 with respect to h.

For weak convergence we only need to approximate the measure induced by the process, hence we can replace the Wiener process increments by other random variables with similar first four moment properties. Thus, by choosing more easily replicated increments we can obtain a simpler scheme. For instance, the method usually called as the *weak Euler scheme*

$$\bar{X}_0 = x_0,$$

$$\bar{X}_{k+1} = \bar{X}_k + a_k h + \sum_{i=1}^{d} b_{ik}\xi_{ik}\sqrt{h}, \tag{3.39}$$

where $\xi_{j,k}$ $k = 0, \dots, M-1$, are independent random variables distributed by the law $P(\xi = \pm 1) = 1/2$, also has first order of accuracy in the sense of weak approximation.

With addition of one extra term from the Ito-Taylor expansion (3.25) to the Euler scheme we obtain the one-step approximation

$$\begin{aligned} \bar{X}_{t,x}(t+h) &= x + a(t,x)h + \sum_{i=1}^{d} b_i(t,x)(W_i(t+h) - W_i(t)) \\ &+ \sum_{j=1}^{d}\sum_{i=1}^{d} \Lambda_j b_i(t, X(t)) \int_t^{t+h} (W_j(s) - W_j(t))\, dW_i(s). \end{aligned} \tag{3.40}$$

Iteratively progressing forward this results in the *Milstein scheme* (see [53, 45, 58, 57]):

$$\bar{X}_0 = x_0,$$

$$\begin{aligned} \bar{X}_{k+1} &= \bar{X}_k + a_k h + \sum_{i=1}^{d} b_{ik}(W_i(t_k+h) - W_i(t_k)) \\ &+ \sum_{i=1}^{d}\sum_{j=1}^{d} \Lambda_j b_{ik} \int_{t_k}^{t_k+h} (W_j(s) - W_j(t))\, dW_i(s). \end{aligned} \tag{3.41}$$

For this method we have

$$E\left|X_{t,x}(t+h) - \bar{X}_{t,x}(t+h)\right| = O(h^2),$$

$$E\left|X_{t,x}(t+h) - \bar{X}_{t,x}(t+h)\right|^2 = O(h^3),$$

i.e. $q_1 = 2$ and $q_2 = 3/2$, yielding first order of the mean-square convergence for (3.41) according to Theorem 3.1.7.

We remark that in the general case we will face with the difficulty of simulating multiple stochastic integrals appearing in scheme (3.41) since they cannot be easily expressed in terms of Wiener process increments. However, it is clearly possible for a single noise ($d=1$) since

$$\int\limits_{t_k}^{t_k+h} (W(s)-W(t))\,dW(s) = \frac{1}{2}\left(W(t_k+h)-W(t_k)\right)^2 - \frac{1}{2}h.$$

The Milstein scheme (3.41) has the same weak order of accuracy as the Euler method does. This illustrates the fact that an increase in the order of accuracy in the mean-square sense does not in general imply an increase of the weak order of accuracy.

Finally, we shall consider a scheme, which will be of the second weak order. The one-step approximation is obtained by truncating the remainder in (3.25). Using the identity

$$\int\limits_{t}^{t+h} (s-t)\;dW_i(s) = h\int\limits_{t}^{t+h} dW_i(s) - \int\limits_{t}^{t+h} (W_i(s)-W_i(t))ds,$$

the one-step approximation becomes

$$\begin{aligned}
\bar{X}_{t,x} \quad &= \quad x + ah + \sum_{i=1}^{d} b_i \int\limits_{t}^{t+h} dW_i(s) + La\frac{h^2}{2} \\
&\quad + \sum_{i=1}^{d} (\Lambda_i a - Lb_i) \int\limits_{t}^{t+h} (W_i(s)-W_i(t))\;\;ds + \sum_{i=1}^{d} Lb_i \int\limits_{t}^{t+h} dW_i(s)h \\
&\quad + \sum_{i=1}^{d}\sum_{j=1}^{d} \Lambda_j b_i \int\limits_{t}^{t+h} (W_j(s)-W_j(t))\,dW_i(s),
\end{aligned}$$

where the coefficients a, b_i, La, $\Lambda_i a$, Lb_i, $\Lambda_j b_i$ are evaluated at the point (t,x).

The corresponding method where we replace the Wiener process increments by simpler random variables has the form ([54, 57, 58, 78, 45]):

$$X_{k+1} = X_k + a_k h + \sum_{i=1}^{d} b_{ik}\xi_{ik}h^{1/2} + (La)_k\,\frac{h^2}{2}$$

$$+\frac{1}{2}\sum_{i=1}^{d}\left(\Lambda_i a+Lb_i\right)_k \xi_{ik}h^{3/2}+\sum_{i=1}^{d}\sum_{j=1}^{d}\left(\Lambda_j b_i\right)_k \xi_{ijk}, \tag{3.42}$$

where ξ_{ijk} satisfy

$$\xi_{ijk}=\frac{1}{2}\xi_{ik}\xi_{jk}-\frac{1}{2}\theta_{ij}\zeta_{ik}\zeta_{jk}, \quad \theta_{ij}=\begin{cases}-1, \ i<j, \\ \\ 1, \ i\geq j,\end{cases}$$

and ξ_{ik} and ζ_{ik} $k=0,\ldots,M-1$, are independent random variables distributed by the corresponding law $P(\xi=0)=2/3$, $P(\xi=\pm\sqrt{3})=1/6$ and $P(\zeta=\pm1)=1/2$.

3.2 The HJM framework: revisited

In this Section we shall revisit (see Section 2.2) the HJM model under risk-neutral measure Q in order to state the assumptions we impose on the volatility and initial forward curve. These assumptions guarantee the existence of the unique strong solution of the SDE corresponding to HJM model and are sufficient for construction of the class of numerical methods proposed in Section 3.4. We shall also formulate a pricing problem for European type interest rate products which will used to illustrate the numerical algorithms proposed.

As before (see Chapter 2), we assume that there exists an arbitrage-free market with a frictionlessly traded continuum of default-free zero-coupon bonds $\{P(t,T)$, $t\leq T$, $T\in[t_0,T^*]$, $t\in[t_0,t^*]\}$, where $P(t,T)$ denotes the price at calendar time t of a bond with maturity T. We require that $P(T,T)=1$ and $P(t,T)$ is sufficiently smooth in the maturity variable T.

The HJM framework [36] models the dynamics of the forward curve (see Section 2.2.1)

$$\{f(t,T), \quad t\leq T, \quad T\in[t_0,T^*], \quad t\in[t_0,t^*]\}.$$

We recall from Section 2.2.2, that given an integrable deterministic initial forward curve

$$f(t_0, T) = f_0(T),$$

the arbitrage-free dynamics of the forward curve under the risk-neutral measure Q associated with the numeraire $B(t)$ (cf. (2.4),(2.5)) are modelled through an Ito process of the form

$$\begin{aligned} f(t,T) - f_0(T) &= \int_{t_0}^{t} \sigma^\top(s,T) \left(\int_s^T \sigma(s,u) du \right) ds \\ &\quad + \int_{t_0}^{t} \sigma^\top(s,T) dW^{\mathrm{Q}}(s), \quad t_0 \le t \le t^* \wedge T, \quad t_0 \le T \le T^*, \end{aligned} \tag{3.43}$$

where $W(t) = (W_1(t), \dots, W_d(t))^\top$ is a d-dimensional standard Wiener process defined on a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t_0 \le t \le t^*}, \mathrm{Q}\right)$ satisfying the usual hypotheses; $\sigma(t,T)$ is an $\mathbb{R}^d$-valued $\mathcal{F}_t$-progressively measurable stochastic process with $\int_{t_0}^T |\sigma(s,T)|^2 ds < \infty$; and $t^* \wedge T := \min(t^*, T)$.

In general, the volatility $\sigma(t,T) := \sigma(t,T,\omega)$ can depend on the current and past values of forward rates. In this thesis we restrict ourselves to the case in which σ depends on the current forward rate only, i.e.,

$$\sigma(t,T) := (\sigma_1(t,T,f(t,T)), \dots, \sigma_d(t,T,f(t,T)))^\top, \tag{3.44}$$

where $\sigma_i(t,T,z)$, $i = 1, \dots, d$, are deterministic functions defined on $[t_0, t^*] \times [t_0, T^*] \times \mathbb{R}$. Then the term $\int_s^T \sigma(s,u)du$ in (3.43) can be written as $\int_s^T \sigma(s,u,f(s,u))du$, and, consequently, (3.43)-(3.44) is an infinite-dimensional SDE. We impose the following assumptions on the HJM model (3.43)-(3.44).

Assumption 3.2.1 The functions $\sigma_i(t,T,z)$, $i = 1, \dots, d$, are uniformly bounded, i.e., there is a constant $C > 0$ such that

$$|\sigma_i(t,T,z)| \le C, \ \ (t,T,z) \in [t_0, t^*] \times [t_0, T^*] \times \mathbb{R}. \tag{3.45}$$

Assumption 3.2.2 For sufficiently large p_1, $p_2 \geq 1$, the partial derivatives

$$\frac{\partial^{j+k+l}\sigma_i(t,T,z)}{\partial t^j \partial T^k \partial z^l}, \; 0 \leq j \leq p_1, \; 0 \leq k+l \leq p_2, \quad i = 1,\ldots,d, \tag{3.46}$$

are continuous and uniformly bounded in $[t_0, t^*] \times [t_0, T^*] \times \mathbb{R}$.

Assumption 3.2.3 The initial forward curve $f_0(T)$, $T \in [t_0, T^*]$, is deterministic and sufficiently smooth.

The imposed conditions are sufficient to ensure that the stochastic equation (3.43)-(3.44) has a unique strong solution $f(t,T)$ (cf. Theorem 3.1.2), which is sufficiently smooth in the last argument (cf. Theorem 3.1.4), (see [36, 79] and also [49, 29] for differentiating SDE solutions with respect to a parameter).

Further, it is not difficult to show (see Appendix A) that they imply boundedness of exponential moments of $f(t,T)$, i.e., for any $c \in \mathbb{R}$ there is a constant $C > 0$ such that

$$E^{\mathrm{Q}} \exp(c|f(t,T)|) < C \tag{3.47}$$

for all $(t,T) \in [t_0, t^*] \times [t_0, T^*]$. The constant C in (3.47) depends on the initial forward curve $f_0(T)$, volatility $\sigma(t,T,z)$, and on c.

Remark 3.2.4 *As it was shown in [79], for the SDE* (3.43)-(3.44) *to have the unique strong solution it suffices to require a weaker assumption than Assumption 3.2.1:*

$$|\sigma_i(t,T,z)| \leq C\left(1 + |z|^{1/2}\right).$$

However, in the thesis we restrict ourselves to the stronger set of conditions which allow us to consider methods of higher order. Assumptions 3.2.1-3.2.3 are sufficient for all the statements in this thesis. The choice of p_1 and p_2 depends on a particular algorithm (as usual, the more accurate an algorithm the more derivatives are needed). At the same time, the imposed conditions are not necessary and the proposed numerical methods themselves can be used under broader assumptions.

We pay attention that Assumptions 3.2.1-3.2.3 do not guarantee positiveness of $f(t,T)$ which could be a desirable property taking into account the financial context of the HJM model. One can notice that if we also require that

$$f_0(T) \geq 0 \text{ and } \sigma_i(t,T,0) = 0,\ i = 1,\ldots,d,\ (t,T) \in [t_0,t^*] \times [t_0,T^*],$$

then the forward rates are nonnegative $f(t,T) \geq 0$ for all $(t,T) \in [t_0,t^*] \times [t_0,T^*]$.

We will illustrate our numerical methods for the HJM model (3.43)-(3.44) by pricing interest rates derivatives of European-type which we described in Section 2.1.3. Among these instruments are interest rate caps, floors, and swaptions [1, 15, 18, 26, 65, 28, 68]. A cap price is obtained by summing up the prices of the underlying caplets. Consider a caplet set at time s_k with payment date at $s_i > s_k$, with strike K and unit cap nominal value. We recall from Section 2.1.3, that its price at time $t_0 \leq s_k$ is given by (cf. (2.20))

$$E^{\mathrm{Q}} \exp\left(-\int_{t_0}^{s_k} r(u)du\right)\left[1-(1+K(s_i-s_k))\exp\left(-\int_{s_k}^{s_i} f(s_k,u)du\right)\right]_+ . \quad (3.48)$$

Now consider a payer swaption of maturity s_k and with underlying swap maturity $s_i > s_k$. Its price at time $t_0 \leq s_k$ can be found as (cf. (2.23))

$$\begin{aligned} & E^{\mathrm{Q}} \exp\left(-\int_{t_0}^{s_k} r(u)du\right)\left[1-\exp\left(-\int_{s_k}^{s_i} f(s_k,u)du\right)\right. \\ & \left. -K\sum_{j=k+1}^{i}(s_j-s_{j-1})\exp\left(-\int_{s_k}^{s_j} f(s_k,u)du\right)\right]_+ . \end{aligned} \quad (3.49)$$

Let $G(z)$, $z \in \mathbb{R}$, be a payoff function satisfying the global Lipschitz condition, i.e.,

$$|G(z)-G(z')| \leq K\,|z-z'|\,, \quad z,z' \in \mathbb{R}. \quad (3.50)$$

In this thesis, motivated by the above examples, we consider the price of a generic interest rate contract under risk-neutral measure of the form

$$F(t_0, f_0(\cdot); s_k, s_i) = E\exp(-Y(s_k))G\left(P(s_k, s_i)\right), \tag{3.51}$$

where

$$Y(s_k) = \int_{t_0}^{s_k} r(u)du, \tag{3.52}$$

$$P(s_k, s_i) = \exp\left(-Z(s_k, s_i)\right), \tag{3.53}$$

and

$$Z(s_k, s_i) = \int_{s_k}^{s_i} f(s_k, u)du. \tag{3.54}$$

We note that (3.51) does not cover the case of swaptions (3.49). To include swaptions, the payoff G in (3.51) should be of the form $G\left(P(s_k, s_{k+1}), \ldots, P(s_k, s_i)\right)$ and

$$F(t_0, f_0(\cdot); s_k, s_{k+1}, \ldots, s_i) = E\exp(-Y(s_k))G\left(P(s_k, s_{k+1}), \ldots, P(s_k, s_i)\right). \tag{3.55}$$

We limit ourselves in the thesis to the payoff of the form (3.51) for the sake of transparent exposition. All the proposed numerical algorithms are applicable to the more general form of the payoff (3.55). Also, no additional ideas are required to extend our theoretical analysis to the case (3.55).

Remark 3.2.5 *(Forward measure pricing) The HJM dynamics can be written under the s_k-forward measure (see Section 2.2.3) instead of the risk-neutral measure. We recall that the corresponding SDE has the form* (*cf.* (3.43))*:*

$$\begin{aligned} f(t,T) - f_0(T) &= \int_{t_0}^{t} \sigma^\top(s,T)\left(\int_{s_k}^{T} \sigma(s,u)du\right)ds + \int_{t_0}^{t} \sigma^\top(s,T)dW^{s_k}(s), \\ &t_0 \le t \le t^* \wedge T \wedge s_k, \quad t_0 \le T \le T^*, \end{aligned} \tag{3.56}$$

with $W^{s_k}(s)$ being a d-dimensional standard Wiener process under the s_k-forward measure Q^{s_k}. The pricing formula for a generic interest rate contract with payoff

$G\left(P(s_k, s_i)\right)$ *under* Q^{s_k} *is* (*cf.* (2.11))*:*

$$F(t_0, f_0\left(\cdot\right); s_k, s_i) = P(t_0, s_k) E^{\mathrm{Q}^{s_k}}\left(G\left(P(s_k, s_i)\right)\right). \tag{3.57}$$

This form is computationally simpler than (3.51) *since it does not require evaluation of the short rate. At the same time we note that pricing of some interest rate products (e.g., Eurodollar futures) require the use of risk-neutral measure [15, 62]. In this thesis we construct numerical algorithms for approximating* (3.51). *Obviously, these algorithms are readily (actually more easily) applicable to* (3.57).

3.3 Review of existing numerical methods for HJM model

The HJM model has closed-form solutions only for some special cases of volatility, and valuations under the HJM framework usually require a numerical approximation. Before embarking on the construction of the numerical algorithms for HJM model, we are going to review the existing methods. As far as we know, the literature on numerics for the HJM model is rather sparse.

The common approach (see, e.g. [35, 40, 31, 32, 12] and the references therein) is to discretize the HJM equation itself taking coinciding grids in the calendar time t and in the maturity time T. As we will see, the requirement of taking the same steps in calendar and maturity times limits efficiency of numerical schemes. The known methods differ in the way they approximate the integral in the arbitrage-free drift of the HJM model while they all use Euler-type schemes for discretization in calendar time. In [35, 40, 31, 32] approximations of the arbitrage-free drift are chosen so that the overall discrete approximations of the HJM equation preserve the martingale property for the discretized discounted bond process. Moreover, in [32] a variance-reduction technique based on a combination of importance sampling and stratified sampling for pricing path-dependent European-style interest rate options in a multifactor HJM setting is analysed. The authors in [12] present two numerical methods for HJM model, Euler type finite difference and finite element methods, and prove their error estimates which can be useful for adaptive algorithms. Though, the setting considered is restricted to the volatility function depending the forward curve only through the current short rate.

A different numerical approach, based on Galerkin approximation, is considered in [52]. In this paper, the problem of valuing American type interest rate products is characterised as the solution of an infinite-dimensional Hamilton-Jacobi variational inequality which reduces to the Kolmogorov backward equation for the products exclusive of early exercise. The forward rate curve is approximated using a Fourier-Legendre expansion resulting in a finite-dimensional approximation of the infinite-dimensional HJM equation and associated valuing problem. This valuing problem is further approximated by a nonlinear partial differential equation trough the use of the penalization technique. An adaptive method-of-lines extrapolation finite element method based upon the penalised formulation of the valuing problem is proposed. The spacial semi-discretisation in terms of the linear finite element basis is introduced. Whereas, the semi-implicit discretization in time based upon the backward Euler method is employed to obtain fully discreet analogue of the valuing problem.

Using the Musiela parametrization (Section 2.2.4, see also [61, 25, 18]), the HJM equation can be re-written in the form of a first-order stochastic partial differential equation (SPDE) which then can be approximated using an SPDE solver. Such an approach was used, e.g., in [22, 23]. In [23], the authors consider the HJM SPDE with stochastic volatility chosen as a mean-reverting Ornstein-Uhlenbeck process. A symmetrically weighted sequential splitting scheme of weak order 2 along with Quasi-Monte Carlo algorithm is analysed within this setting. It is argued in this paper that higher-order weak approximation schemes can be used together with Quasi Monte Carlo algorithms to obtain an efficient pricing method, which is superior to Multi-level Monte Carlo. The efficiency of the numerical method is demonstrated by the calibration of the model to a caplet data. Also, a payer swaption is priced using the calibrated model.

We note that, in comparison with other works, the papers [12, 22, 23, 52] rigorously proof convergence results of the proposed numerical methods.

3.4 New class of numerical methods for HJM model

In this section we construct a numerical method for simulating (3.51) with the forward rates $f(t,T)$ satisfying the infinite-dimensional SDE (3.43)-(3.44). Examples of some particular algorithmic realizations of this method are given in Section 3.5.

This section is organized in the following way. We first introduce a maturity time discretization (T-discretization) and arrive at a finite-dimensional approximation of (3.43)-(3.44), i.e., at a finite system of SDEs (Section 3.4.1). Then (Section 3.4.2) we discretize time (t-discretization) and apply a weak-sense numerical integrator to the obtained finite system of SDEs. Finally, Section 3.4.3 deals with approximating the functionals Y and Z from (3.52)-(3.54) and the option price (3.51).

For the simplicity of presentation, we consider equally-spaced grids for maturity time T and time t. A nonuniform discretization might be needed in practical financial applications, and a generalization of the proposed algorithms to nonuniform time grids is straightforward.

3.4.1 T-discretization

Consider a uniform partition of the maturity time interval $[t_0, T^*]$ with a maturity time step (T-step) $\Delta = (T^* - t_0)/N$:

$$t_0 = T_0 < \cdots < T_N = T^*, \quad T_i = i\Delta, \quad i = 0, \ldots, N. \tag{3.58}$$

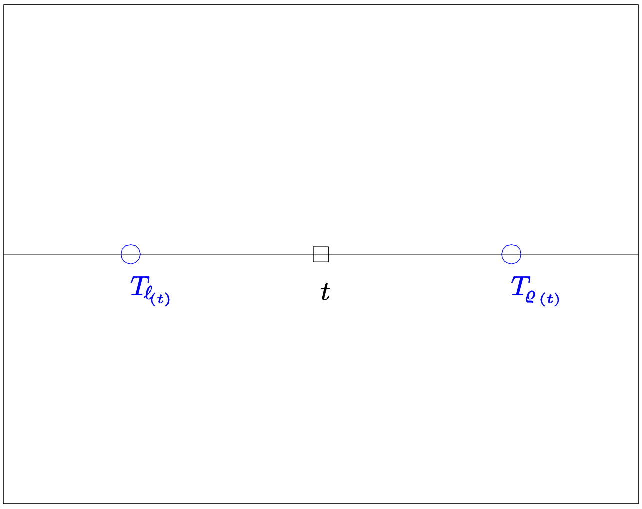

Figure 3.1: Index notation for the closest node on the grid (3.58) to the time t from the left (or from the right), denoted by $T_{\ell(t)}$ (or by $T_{\varrho(t)}$).

We note that the grid (3.58) has clearly different meaning to the partition introduced in Chapter 2, though we use a similar notation as it is standard for both cases.

We shall first introduce the index notation we are going to use. Denote by $\ell(t)$ the auxiliary index dependent on time t so that

$$\ell(t) = \max\left\{i = 0, 1, \ldots, N : \quad t \geq T_i\right\}, \tag{3.59}$$

and by $\varrho(t)$ the auxiliary index dependent on time t so that

$$\varrho(t) = \min\left\{i = 0, 1, \ldots, N : \quad t < T_i\right\}, \tag{3.60}$$

i.e., $T_{\ell(t)} \leq t < T_{\varrho(t)}$ and $T_{\ell(t)}$ (or $T_{\varrho(t)}$) is the closest node on the grid (3.58) to the time t from the left (or from the right). See Figure 3.1 for illustration of this index notation. We also point out that $\varrho(t) = \ell(t) + 1$.

Further, we require for simplicity that Δ is sufficiently small so that a number of nodes T_i between t^* and T^* is sufficient for realization of all the quadrature rules and interpolation/extrapolation used in the method which we introduce in this Section 3.4. We will pay attention to the required amount of nodes between t^* and T^* in the method's description. At the same time, if in practical realization the distance between t^* and T^* is relatively small in comparison with the chosen T-step Δ, then one would need to run simulation for a slightly longer maturity-time interval, extending it beyond T^* by a few steps of Δ (see further explanation in Section 3.5.3).

For a node T_i, $i = 0, \ldots, N$, on the maturity time grid (3.58), we approximate the integrals in (3.43):

$$I_j(s, T_i) := \int_s^{T_i} \sigma_j(s, u)du, \; j = 1, \ldots, d, \; t_0 \le s \le t^* \wedge T_i, \quad i = 1, \ldots, N, \quad (3.61)$$

by a composite quadrature rule $S_{I_j}(s, T_i, \Delta)$:

$$I_j(s, T_i) \approx S_{I_j}(s, T_i, \Delta) = \Delta \sum_{k=\varrho(s)}^{\kappa(s,T_i)} \gamma_k(s)\sigma_j(s, T_k), \quad (3.62)$$

where the quadrature rule's weights $\gamma_k(s)$ and the nodes $k = \varrho(s), \ldots, \kappa(s, T_i)$ are chosen so that under Assumptions 3.2.1-3.2.3 the approximation is of order $O(\Delta^p)$ for a given $p \ge 1$, i.e., the numerical integration error is estimated as

$$\left(E\left[S_{I_j}(s, T_i, \Delta) - I_j(s, T_i) \right]^2 \right)^{1/2} \le C\Delta^p \quad (3.63)$$

with a constant $C > 0$ independent of Δ, s, T_i, j. Some examples of such quadratures are given in Section 3.5. We note (see details in Section 3.5) that when s and T_i are close, to approximate $I_j(s, T_i)$ with a required accuracy the number $\kappa(s, T_i)$ in (3.62) can be chosen larger than i. We recall that since we assumed that there is a sufficient number of nodes between t^* and T^* the number $\kappa(s, T_i)$ does not exceed N. We will also use the vector notation $I(s, T_i) := (I_1(s, T_i), \ldots, I_d(s, T_i))^\top$ and $S_I(s, T_i, \Delta) := (S_{I_1}(s, T_i, \Delta), \ldots, S_{I_d}(s, T_i, \Delta))^\top$.

For a fixed $T = T_i$, it is convenient for later purposes (namely, for computing the short rate $r(t) = f(t, t)$ as it will become clear in Section 3.4.3) to consider the

SDE (3.43)-(3.44) on a slightly larger time interval: $t_0 \le t \le t^* \wedge T_{(i+1)\wedge N}$, i.e., for $T_i < t^*$ (note that $t^* \le T_N$) we would like to extend the definition of $f(t, T_i)$ from $t \in [t_0, T_i]$ to $t \in [t_0, T_{i+1}]$. Though from the point of view of financial applications the forward rate $f(t, T_i)$ is not defined on the interval $t \in (T_i, T_{i+1}]$, Assumptions 3.2.1-3.2.3 guarantee that (3.43)-(3.44) has the strong solution on the extended interval and, as it will be seen in future, this extension is beneficial from the computational prospective (see also Remarks 3.4.3 and 3.4.4). This extension requires from us to consider, in addition to (3.61), the integrals

$$I_j(s, T_{\ell(s)}) := \int_s^{T_{\ell(s)}} \sigma_j(s, u) du. \tag{3.64}$$

We approximate these integrals by a quadrature rule analogous to the one in (3.62) but with summation index k starting from $\ell(s)$:

$$I_j(s, T_{\ell(s)}) \approx S_{I_j}(s, T_{\ell(s)}, \Delta) = \Delta \sum_{k=\ell(s)}^{\kappa(s, T_{\ell(s)})} \gamma_k(s) \sigma_j(s, T_k), \tag{3.65}$$

and we require that its error satisfies (3.63). Combining (3.62) and (3.65), we will write in what follows that

$$S_{I_j}(s, T_i, \Delta) = \Delta \sum_{k=\ell(s)}^{\kappa(s, T_i)} \gamma_k(s) \sigma_j(s, T_k) \tag{3.66}$$

with the coefficient $\gamma_{\ell(s)}(s) = 0$ if $i > \ell(s)$.

Using (3.66), we approximate the solution $f(t, T)$ of the infinite-dimensional SDE (3.43)-(3.44) at the nodes $T = T_0, \ldots, T_N$, by the $N+1$-dimensional stochastic process $\tilde{f}^i(t) \approx f(t, T_i)$, $i = 0, \ldots, N$, which satisfies the finite system of coupled SDEs:

$$\begin{aligned} \tilde{f}^i(t) - f_0^i &= \int_{t_0}^t \tilde{\sigma}_i^\top(s) \tilde{S}_I(s, T_i, \Delta) ds + \int_{t_0}^t \tilde{\sigma}_i^\top(s) dW(s), \\ t_0 &\le t \le t^* \wedge T_{(i+1)\wedge N}, \; i = 0, \ldots, N, \end{aligned} \tag{3.67}$$

where

$$
\begin{aligned}
f_0^i &= f_0(T_i), && (3.68)\\
\tilde{\sigma}_i(s) &= (\tilde{\sigma}_{i,1}(s), \ldots, \tilde{\sigma}_{i,d}(s))^\top = (\sigma_1(s, T_i, \tilde{f}^i(s)), \ldots, \sigma_d(s, T_i, \tilde{f}^i(s)))^\top, \\
\tilde{S}_I(s, T_i, \Delta) &= \left(\tilde{S}_{I_1}(s, T_i, \Delta), \ldots, \tilde{S}_{I_d}(s, T_i, \Delta)\right)^\top
\end{aligned}
$$

and

$$\tilde{S}_{I_j}(s, T_i, \Delta) = \Delta \sum_{k=\ell(s)}^{\kappa(s,T_i)} \gamma_k(s)\tilde{\sigma}_{k,j}(s). \quad (3.69)$$

We emphasize again that we extended the time interval from $t \in [t_0, t^*]$ to $t \in [t_0, t^* \wedge T_{(i+1)\wedge N}]$.

Assumptions 3.2.1-3.2.3 guarantee the existence of the unique strong solution of (3.67)-(3.69). Further, it is not difficult to show that they also imply boundedness of exponential moments of $\tilde{f}^i(t)$, i.e., for any $c \in \mathbb{R}$ there is a constant $C > 0$ such that (cf. (3.47)):

$$E \exp(c|\tilde{f}^i(t)|) < C \quad (3.70)$$

for all $t \in [t_0, t^*] \wedge T_{(i+1)\wedge N}$, $i = 0, \ldots, N$. In connection with (3.70) we recall that due to Assumption 3.2.3 the initial forward rate curve $f_0(T)$, $t_0 \le T \le T^*$, is bounded by a finite constant. Hence, $\tilde{f}^i(0) = f_0^i$, $i = 0, \ldots, N$, are bounded by the same constant.

In Section 4.1.1 we prove (see Theorem 4.1.1) mean-square convergence of $\tilde{f}^i(t)$ to $f(t, T_i)$ when $\Delta \to 0$. We note that the system (3.67)-(3.69) plays only an auxiliary role in our consideration. It is used as guidance in constructing fully discrete numerical algorithms (i.e., discrete in both T and t) and also in proofs of their convergence.

3.4.2 t-discretization

In this section we discretize the finite system of coupled ordinary SDEs (3.67)-(3.69) with respect to calendar time t and thus arrive at a fully discrete method.

We introduce an equally-spaced grid for calendar time t with step (t-step) $h = (t^* - t_0)/M$:

$$t_0 < \cdots < t_M = t^*, \quad t_k = kh, \quad k = 0, \dots, M.$$

In what follows we use the notation (cf. (3.59) and (3.60)):

$$\ell_k := \ell(t_k), \quad \varrho_k := \varrho(t_k). \tag{3.71}$$

We consider an approximation $\bar{f}^i_{k+1}$ of $\tilde{f}^i(t_{k+1})$ from (3.67) (i.e., a full discretization of (3.43)-(3.44) in both T and t) of the form (cf. (3.17))

$$\bar{f}^i_0 = f^i_0, \; i = 0, \dots, N, \tag{3.72}$$

$$\bar{f}^i_{k+1} = \bar{f}^i_k + A^i(t_k, T_i; \bar{f}^j_k, \; j = \ell_{k+1}, \dots, \kappa(t_{k+1}, T_i) \vee i; h; \xi_k),$$

$$i = \ell_{k+1}, \dots, N, \; k = 0, \dots, M,$$

where the form of the functions A^i depends on the coefficients of (3.67)-(3.69), i.e., on σ and on a choice of the quadrature rule S_{I_j}; $\kappa(t_k, T_i)$ is as in the quadrature (3.69); $\xi_k, \; k = 0, \dots, M$, are some random vectors which have moments of a sufficiently high order and ξ_k for $k > 0$ are independent of $\bar{f}^i_j$, $i = \ell_j, \dots N$, $j = 0, \dots, k$, and of $\xi_0, \dots, \xi_{k-1}$.

To simplify the exposition of our theoretical analysis, in what follows we consider the extended $\tilde{f}^i(t)$ and $\bar{f}^i_k$. We put

$$\tilde{f}^i(t) = \tilde{f}^i(T_{i+1}), \quad T_{(i+1)\wedge N} \wedge t^* \leq t \leq t^*, \quad 0 \leq i \leq \ell(t^*) - 1,$$

and then the $N+1$-dimensional vector $\{\tilde{f}^i(t), \, i = 0, \dots, N\}$ is defined for all $t \in [t_0, t^*]$. We put

$$\bar{f}^i_k = \bar{f}^i_{\mathbf{m}}, \quad k = \mathbf{m} + 1, \dots, M, \quad 0 \leq i \leq \ell(t^*) - 1,$$

where $\mathbf{m} = \lceil (T_{i+1} - t_0)/h \rceil - 1$ (we recall that $\lceil \cdot \rceil$ denotes the ceiling function). Then the $N+1$-dimensional vector $\{\bar{f}_k^i,\, i=0,\ldots,N\}$ is defined for all $k=0,\ldots,M$. Let us emphasize that we do not use the extension of $\bar{f}_k^i$ in numerical algorithms and these extensions of $\tilde{f}^i(t)$ and $\bar{f}_k^i$ are done in order to use the vector notation $\tilde{f}(t)$ and $\bar{f}_k$ without need to adjust length of these vectors as t and k grow.

We assume that the A^i in (3.72) are such that $\bar{f}_k^i$ satisfy the following condition.

Assumption 3.4.1 For any $c \in \mathbb{R}$ there is a constant $C > 0$ such that

$$E\exp(c|\bar{f}_k^i|) < C \tag{3.73}$$

for all $i = 0,\ldots N,\;\; k=0,\ldots,M.$

This condition is satisfied by all sensible numerical schemes (i.e., sensible choices of A^i in (3.72)) thanks to the uniform boundedness of $\sigma_i(t,T,z)$ (see Assumption 3.2.1) and boundedness of the initial condition (see Assumption 3.2.3 and also the comment after (3.70)). In particular, it is satisfied by the weak Euler-type scheme (3.95) we use in the algorithms in Section 3.5.

We also require that the numerical method (3.72) for the SDEs (3.67)-(3.69) is of local weak order $q+1$, i.e., that the following assumption holds.

Assumption 3.4.2 We assume that the method *(3.72)* is such that for some positive constant C independent of Δ

$$|E(\prod_{j=1}^{s}\delta\tilde{f}^{i_j} - \prod_{j=1}^{s}\delta\bar{f}^{i_j})| \le Ch^{q+1},\; s=1,\ldots,2q+1, \tag{3.74}$$

$$E\prod_{j=1}^{2q+2}|\delta\bar{f}^{i_j}| \le Ch^{q+1}, \tag{3.75}$$

where

$$\delta\tilde{f}^i := \tilde{f}_{t,x}^i(t+h) - x^i,\quad \delta\bar{f}^i := \bar{f}_{t,x}^i(t+h) - x^i,$$

and $\tilde{f}_{t,x}^i(t+h)$ is the solution of the SDEs *(3.67)* with the initial condition x given at time $t:\ \tilde{f}_{t,x}^i(t) = x^i$, and $\bar{f}_{t,x}^i(t+h)$ is the one-step approximation of *(3.67)* found according to *(3.72)* with $\bar{f}_{t,x}^i(t) = x^i$.

Assumption 3.4.2 is similar to the one used in the standard theory of numerical integration of SDEs in the weak sense (cf. Theorem 3.1.9). As we will see in Section 4.1, Assumptions 3.2.1-3.2.3 and 3.4.1-3.4.2 guarantee weak convergence of the numerical method (3.72) to the solution of the auxiliary system of SDEs (3.67) with order h^q.

We note that C in (3.74)-(3.75) is independent of x while in the standard theory of numerical integration of SDEs one usually has C depending on x in such estimates (see Theorem 3.1.9). In our case it is natural to put C independent of x since the coefficients of (3.67) and their derivatives are uniformly bounded (see Assumptions 3.2.1-3.2.2). We also emphasize that the constants C in (3.74)-(3.75) are required not to depend on Δ.

Remark 3.4.3 *The numerical method* (3.72) *contains the approximation* $\bar{f}_k^{\ell_k}$ *of the forward rate* $f(t_k, T_{\ell_k})$ *(recall that* $t_k \geq T_{\ell_k}$*) which from the financial point of view does not exist unless* $t_k = T_{\ell_k}$*. However, from both theoretical and numerical points of view, it is not prohibiting to consider the values* $\bar{f}_k^{\ell_k}$ *which, as we will see later in Section 3.4.3, is computationally beneficial. We may interpret the points* (t_k, T_{ℓ_k}) *on our* (t, T)*-grid as fictitious nodes (see also Remark 3.4.4).*

The approximation (3.72) of the infinite-dimension stochastic equation (3.43)-(3.44) has two discretization steps: T-step Δ (i.e., step in maturity time) and t-step h (i.e., step in time). We can say that the T-step Δ controls the error of approximating (3.43)-(3.44) by (3.67)-(3.69) while the t-step h controls the error of approximating (3.67)-(3.69) by (3.72). We will later (see Remark 3.4.7) discuss how to choose Δ and h in practice.

3.4.3 Approximation of the price of an interest rate contract

In the previous section we introduced an approximation $\bar{f}_k^i$ of the solution to (3.43)-(3.44). Now we illustrate how this approximation can be used for evaluating the expectation (3.51)-(3.54), i.e.,

$$F(t_0, f_0(\cdot); s_k, s_i) = E \exp(-Y(s_k)) G(P(s_k, s_i)).$$

To find F, one has to compute $Z(s_k, s_i)$ from (3.52) and $Y(s_k)$ from (3.54). In this section we construct numerical approximations for $Z(s_k, s_i)$ and $Y(s_k)$. For clarity of the exposition, we assume in what follows that

$$s_k = t^* \quad \text{and} \quad s_i = T^* .$$

We approximate the maturity time integral from (3.54) by a quadrature rule $S_Z(t^*, T^*, \Delta)$:

$$Z(t^*, T^*) = \int_{t^*}^{T^*} f(t^*, u) du \approx S_Z(t^*, T^*, \Delta) = \Delta \sum_{j=\varrho_M}^{N} \tilde{\gamma}_j f(t^*, T_j), \qquad (3.76)$$

where the weights $\tilde{\gamma}_j$ are chosen so that the quadrature rule is of order $p > 0$, i.e., an inequality of the form (3.63) holds:

$$\left(E\left[S_Z(t^*, T^*, \Delta) - Z(t^*, T^*)\right]^2\right)^{1/2} \leq C\Delta^p. \qquad (3.77)$$

The assumption we made at the beginning of Section 3.4.1 that there is a sufficient number of nodes T_i between t^* and T^* ensures that we can find a quadrature rule (3.76) satisfying (3.77). Some examples of quadratures $S_Z(t^*, T^*, \Delta)$ are given in Section 3.5.

In general, T-discretization and t-discretization have different steps Δ and h, and approximate values of the short rate $r(t_k) = f(t_k, t_k)$ (cf. (2.3)) are not directly available among $\bar{f}_k^i$ which are defined on the (t, T)-grid. Then to numerically evaluate $Y(t^*)$, we need to construct an approximation of $f(t_k, t_k)$ based on the values $\bar{f}_k^i$, $i = \ell_k, \ldots, N$. To this end, let us first consider an approximation of the exact short rate $r(t) = f(t, t)$ using the values of $f(t, T_i)$, $i = \ell(t), \ldots, N$. We recall that thanks to Assumptions 3.2.2-3.2.3 the solution $f(t,T)$ of (3.43)-(3.44) is sufficiently smooth in the last argument. We approximate $r(t)$ by $\pi(t)$ as

$$\begin{aligned}
\pi(t) &= \pi(t; f(t,T_i),\ i = \ell(t), \ldots, \ell(t)+\theta) \qquad (3.78)\\
&= \sum_{l=0}^{\ell(t^*)} \sum_{i=0}^{\theta} \lambda_i(t) f(t, T_{l+i}) \chi_{t\in[T_l, T_{l+1})}\\
&= \sum_{i=0}^{\theta} \lambda_i(t) f(t, T_{\ell(t)+i}), \quad t \in [t_0, t^*],
\end{aligned}$$

where $\lambda_i(t)$ are coefficients independent of f, $|\lambda_i(t)|$ are bounded by a constant independent of Δ, θ is a non-negative integer independent of t and Δ, and χ_A is the indicator function of a set A. We choose the number θ and the coefficients $\lambda_i(t)$ so that the approximation (3.78) is of order p :

$$\left(E\left[r(t) - \pi(t)\right]^2\right)^{1/2} \le C\Delta^p,\ p > 0. \qquad (3.79)$$

The form of (3.78) covers both polynomial interpolation and extrapolation. For interpolation, we approximate $r(t)$ using the values $f(t,T_i)$, $i = \ell(t), \ldots, \ell(t)+\theta$. For extrapolation, the coefficient $\lambda_0(t) = 0$ and we approximate $r(t)$ using the values $f(t,T_i)$, $i = \varrho(t), \ldots, \varrho(t)+\theta-1$. Some particular examples of the approximation $\pi(t)$ are given in Section 3.5. Recall that in Section 3.4.1 we assumed that Δ is such that there is a sufficient number of nodes T_i between t^* and T^* which should, in particular, ensure that $\ell(t^*)+\theta \le N$.

Remark 3.4.4 *We note that we need fictitious points* (t_k, T_{ℓ_k}) *on our* (t,T)*-grid (see also Remark 3.4.3) for the interpolating form of* (3.78)*. The extrapolating form of* (3.78) *does not need the fictitious points as it is sufficient to compute* $\bar{f}_k^i$ *for* $i = \varrho_k, \ldots, \varrho_k+\theta-1,\ k = 0, \ldots, M,$ *all of which have the usual financial meaning. However, we reserve the possibility to use an interpolation (and, consequently, the fictitious points) for simulating short rates since interpolation is usually computationally preferable to extrapolation.*

Using the short rate approximation $\pi(s)$, we approximate the time integral in

(3.52) as

$$Y(t^*) = \int_{t_0}^{t^*} r(s)ds \approx \int_{t_0}^{t^*} \pi(s)ds \approx \tilde{Y}(t^*) := \int_{t_0}^{t^*} \tilde{\pi}(s)ds, \tag{3.80}$$

where $\tilde{\pi}(s)$ has the form of $\pi(s)$ from (3.78) but with $\tilde{f}^i(t)$ instead of $f(t, T_i)$:

$$\tilde{\pi}(s) = \pi(s; \tilde{f}^i(s),\ i = \ell(s), \dots, \ell(s) + \theta).$$

We extend the system of SDEs (3.67) by adding to it the auxiliary differential equation

$$d\tilde{Y} = \pi(s; \tilde{f}^i(s),\ i = \ell(s), \dots, \ell(s) + \theta)ds, \quad \tilde{Y}(t_0) = 0. \tag{3.81}$$

Recall that $\tilde{\pi}(s)$ for every $s \in [t_0, t^*]$ is a linear combination of $\tilde{f}^i(s)$, $i = \ell(s), \dots, \ell(s)+\theta$.

Let $\tilde{Y}_{t,x,y}(s)$, $s \geq t$, be the solution of (3.81) with the initial condition $\tilde{Y}_{t,x,y}(t) = y$ and with $\tilde{f}^i(s) = \tilde{f}^i_{t,x}(s)$ (recall that $\tilde{f}^i_{t,x}(s)$ are defined in Assumption 3.4.2). We observe that:

$$\tilde{Y}_{t,x,0}(t+h) = \sum_{l=\ell(t)}^{\ell(t+h)} \int_t^{t+h} \sum_{i=0}^{\theta} \lambda_i(s) \tilde{f}^{l+i}_{t,x}(s) \chi_{s \in [T_l, T_{l+1})} ds, \tag{3.82}$$

We can show (see Appendix A), that under $h \leq \alpha\Delta$ for some $\alpha > 0$ and for any positive integer m, we have

$$E\left|\tilde{Y}_{t,x,0}(t+h)\right|^m \leq Ch^m \left(1 + \sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l|^m \right), \tag{3.83}$$

where $C > 0$ is a constant independent of Δ and x. We note that the condition $h \leq \alpha\Delta$ guarantees that the number $\ell(t+h) - \ell(t)$ is independent of Δ, which ensures that the constant C in (3.83) is independent of Δ and the number of terms in the sum on the right-hand side of (3.83) is also independent of Δ. This will be essential for proving convergence Theorem 4.1.6.

Now we extend the fully discrete approximation (3.72) by adding to it an approximation of (3.81):

$$\bar{Y}_0 = 0,\ \bar{Y}_{k+1} = \bar{Y}_k + A^Y(t_k; \bar{f}_k^j,\ j = \ell_k, \ldots, \ell_{k+1} + \theta; h),\ k = 0, \ldots, M, \qquad (3.84)$$

where the form of $A^Y(t_k; \bar{f}_k^j,\ j = \ell_k, \ldots, \ell_{k+1} + \theta; h) = A^Y(t_k; h)$ depends on the form of $\pi(s)$ from (3.78) and the accuracy required.

We replace Assumption 3.4.2 on the one-step approximation by the assumption which is applicable to the extended system (3.67), (3.81) and the extended discretization (3.72), (3.84).

Assumption 3.4.5 Let $h \le \alpha\Delta$ for some $\alpha > 0$. We assume that the method *(3.72), (3.84)* is such that for some positive constant C independent of Δ

$$\left| E\left(\delta\tilde{Y}^m \prod_{j=1}^{s-m} \delta\tilde{f}^{i_j} - \delta\bar{Y}^m \prod_{j=1}^{s-m} \delta\bar{f}^{i_j} \right) \right| \le Ch^{q+1}\left(1 + \sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l|^m \right), \qquad (3.85)$$

$$m = 0, \ldots, s,\ s = 1, \ldots, 2q+1;$$

$$\left[E \max_{0\le m\le 2q+2, \{i_1, \ldots i_{2q+2-m}\} \in \{0,\ldots,N\}} \left| \delta\bar{Y}^m \prod_{j=1}^{2q+2-m} \delta\bar{f}^{i_j} \right|^2 \right]^{1/2} \qquad (3.86)$$
$$\le\ Ch^{q+1}\left(1 + \sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l|^{2q+2} \right),$$

where

$$\delta\tilde{f}^i = \tilde{f}_{t,x}^i(t+h) - x^i,\quad \delta\bar{f}^i = \bar{f}_{t,x}^i(t+h) - x^i,$$
$$\delta\tilde{Y} = \tilde{Y}_{t,x,y}(t+h) - y,\quad \delta\bar{Y} = \bar{Y}_{t,x,y}(t+h) - y,$$

$\tilde{f}_{t,x}^i(t+h)$ and $\bar{f}_{t,x}^i(t+h)$ are as in Assumption 3.4.2 and $\tilde{Y}_{t,x,y}(s)$, $s \ge t$, is the solution of *(3.81)* with the initial condition $\tilde{Y}_{t,x,y}(t) = y$, and $\bar{Y}_{t,x,y}(t+h)$ is its one-step approximation found according to *(3.84)* with $\bar{Y}_{t,x,y}(t) = y$.

Note that the constants C in (3.85)-(3.86) do not depend on x, y, and Δ. The dependence of the estimates (3.85)-(3.86) on x is consistent with (3.83). The condition $h \leq \alpha\Delta$ in Assumption 3.4.5 is not restrictive from the practical point of view since we aim to be constructing efficient numerical algorithms for the HJM model by allowing bigger T-steps Δ without losing accuracy. We also note that this condition arises only when we need to approximate the short rate (see also Remark 4.1.7).

Further, we make the following assumption.

Assumption 3.4.6 For some $c > 0$ and $C > 0$

$$E \exp(c|\bar{Y}_k|) < C \tag{3.87}$$

for all $k = 0, \ldots, M$.

As a rule, the condition (3.87) immediately follows from Assumption 3.4.1 which is the case, e.g., for the algorithms presented in Section 3.5.

Based on (3.76), (3.80) and using (3.72), (3.84), we arrive at the approximation $\bar{F}$ of F from (3.51):

$$F(t_0, f_0(\cdot); t^*, T^*) \approx \bar{F}(t_0, f_0; t^*, T^*) = E \exp(-\bar{Y}_M) G\left(\bar{P}(t^*, T^*)\right), \tag{3.88}$$

where $\bar{Y}_M$ is from (3.84);

$$\bar{P}(t^*, T^*) = \exp\left(-\bar{S}_Z(t^*, T^*, \Delta)\right), \tag{3.89}$$

$\bar{S}_Z(t^*, T^*, \Delta)$ is the quadrature rule of the form (3.76) with $f(t^*, T_j)$ replaced by $\bar{f}_M^j$:

$$\bar{S}_Z(t^*, T^*, \Delta) = \Delta \sum_{j=\varrho_M}^{N} \tilde{\gamma}_j \, \bar{f}_M^j; \tag{3.90}$$

and f_0 means the initial condition of (3.67), which is the $N+1$-dimensional vector $\left(f_0^0, \ldots, f_0^N\right)^\top = \left(f_0(T_0), \ldots, f_0(T_N)\right)^\top$.

Finally, the expectation of the discounted payoff in (3.88) is approximated by the Monte Carlo method, i.e.,

$$\begin{aligned} F(t_0, f_0(\cdot); t^*, T^*) &\approx \bar{F}(t_0, f_0; t^*, T^*) \\ &\approx \hat{F}(t_0, f_0; t^*, T^*) = \frac{1}{L}\sum_{l=1}^{L} \exp(-\bar{Y}_M^{(l)}) G\left(\bar{P}^{(l)}(t^*, T^*)\right), \end{aligned} \tag{3.91}$$

where $\bar{Y}_M^{(l)}$, $\bar{P}^{(l)}$ are computed using independent realizations $\bar{f}_k^{j,(l)}$, $j = \ell_k, \dots, N$, $k = 1, \dots, M$, of the random variables $\bar{f}_k^j$.

In (3.91) the first approximate equality corresponds to the error of numerical integration and the error in the second approximate equality comes from the Monte Carlo technique.

The Monte Carlo (i.e., statistical) error in (3.91) is evaluated by

$$\begin{aligned} \bar{\rho}_{MC} &= c\,\frac{\left[Var\left\{\exp(-\bar{Y}_M)G\left(\bar{P}(t^*, T^*)\right)\right\}\right]^{1/2}}{L^{1/2}} \\ &\approx c\,\frac{\left[Var\left\{\exp(-Y(t^*))G\left(P(t^*, T^*)\right)\right\}\right]^{1/2}}{L^{1/2}}, \end{aligned} \tag{3.92}$$

where, for example, the values $c = 1, 2, 3$ correspond to the fiducial probabilities 0.68, 0.95, 0.997, respectively. The Monte Carlo error can be decreased by variance reduction techniques (see, e.g. [31, 32, 58, 60] and references therein). In this paper we deal with the numerical integration error and numerical algorithms which are effective with regard to (t, T)-discretization.

The numerical integration error is analyzed in Section 4.1. The main result of this Section is stated in Theorem 4.1.8, that proves the convergence of the approximation $\bar{F}(t_0, f_0; t^*, T^*)$ to $F(t_0, f_0(\cdot); t^*, T^*)$ with order $p > 0$ in Δ and with order $q > 0$ in h. This provides the theoretical basis for the following remark.

Remark 3.4.7 *(Relationship between Δ and h) A higher order p, i.e., a higher order of an approximation $\tilde{F}(t_0, f_0; t^*, T^*)$ of $F(t_0, f_0(\cdot); t^*, T^*)$, can be achieved by using a higher-order quadrature rules in* (3.66) *and* (3.76) *and higher-order interpolation or extrapolation in* (3.78)*. For this purpose, we can use a large arsenal of* *effective quadrature rules and interpolation/extrapolations methods from the deterministic numerical analysis (see, e.g. [19, 50, 66]) which are directly applicable here* (*see Section 3.5*). *To achieve a higher order* q, *we need a higher-order weak-sense numerical scheme for* (3.67)-(3.69). *As it is known* (*see, e.g.[45, 58, 57]*), *this is a harder task, and, due to complexity of stochastic schemes, one usually restricts themselves to using weak methods of orders* 1 *or* 2. *As a result, in practice we will take* $p \geq q$. *Then, to balance the two errors in* (4.58), *we choose* $\Delta = \alpha h^{q/p}$ *for some* $\alpha > 0$ *to obtain the overall error to be of order* $O(h^q)$. *In other words, by increasing the order* p *we can take larger* T*-discretization steps* Δ *and, consequently, significantly improve computational efficiency of HJM simulation which, in particular, is illustrated in our numerical experiments in Section 3.7.*

3.5 Numerical algorithms

In this section we provide some particular examples of the generic numerical method introduced in Section 3.4. For simplicity of the presentation, we restrict ourselves in this section to the case of T-step being not larger than the t-step, i.e.,

$$h \leq \Delta, \tag{3.93}$$

which is a stronger condition than the one assumed in Theorems 4.1.6 and 4.1.8: $h \leq \alpha\Delta,\ \alpha > 0$. These Theorems we discuss in the next Section. This requirement is not particular restricting since our aim is to construct efficient algorithms for the HJM model by allowing bigger T-steps Δ without losing accuracy as it is discussed in the beginning of Section 3.4 and Remark 3.4.7. We note that there is no difficulty in constructing algorithms imposing $h \leq \alpha\Delta$ for some $\alpha > 0$ instead of (3.93). The condition (3.93) ensures that there cannot be more than one node T_i in any interval $[t_k, t_{k+1})$ hence there are only two cases possible: either $\ell_{k+1} = \ell_k$ or $\ell_{k+1} = \ell_k + 1$ (see Figure 3.2 for illustration of the nodes locations in both cases). This is used in constructing numerical algorithms of this section.

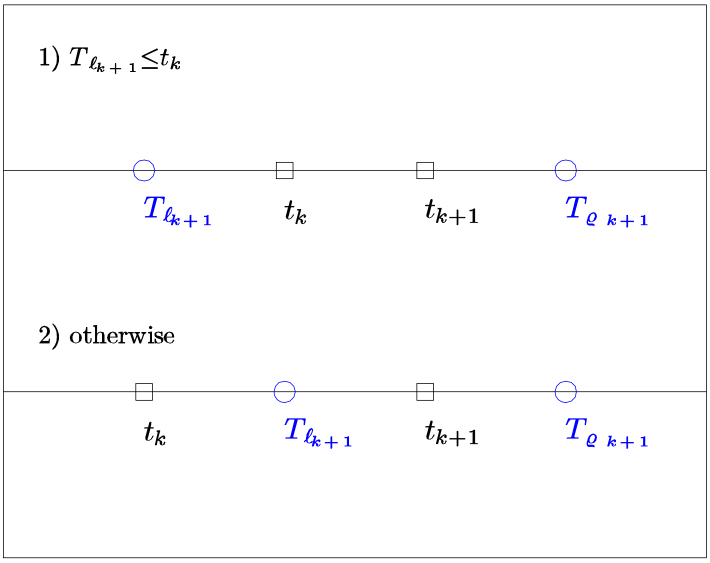

Figure 3.2: Two possible cases of the nodes location on the grid under condition $h \le \Delta$.

We need the following new notation in this section:

$$\Delta_{i,k} := T_i - t_k \tag{3.94}$$

and

$$t_{k+1/2} = \frac{t_k + t_{k+1}}{2} \ .$$

We shall limit the illustration (see also Remark 3.5.4) of the generic numerical method from Section 3.4 to considering only the weak Euler-type scheme (i.e., with $q = 1$) as a numerical approximation of the SDEs (3.67), (3.81), i.e., as an approximation of the t-dynamics. For approximations of higher order q see Remark 3.5.4. Based on the Euler-type approximation for the t-dynamics, the extended discretization (3.72), (3.84) takes the form

$$\bar{f}_0^i = f_0^i, \ i = 0, \dots, N, \quad \bar{Y}_0 = 0, \tag{3.95}$$

$$\bar{f}_{k+1}^i = \bar{f}_k^i + \sum_{j=1}^d \bar{\sigma}_{i,j}(t_k) \bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) + h^{1/2} \sum_{j=1}^d \bar{\sigma}_{i,j}(t_k) \xi_{j,k+1},$$

$$i = \ell_{k+1}, \ldots, N,$$

$$\bar{Y}_{k+1} = \bar{Y}_k + A^Y(t_k; \bar{f}_k^j,\ j = \ell_k, \ldots, \ell(t^*) + \theta; h),\ k = 0, \ldots, M-1,$$

where $\xi_{j,k+1}$ are independent random variables distributed by the law $P(\xi = \pm 1) = 1/2$,

$$(\bar{\sigma}_{i,1}(t_k), \ldots, \bar{\sigma}_{i,d}(t_k))^\top = (\sigma_1(t_k, T_i, \bar{f}_k^i), \ldots, \sigma_d(t_k, T_i, \bar{f}_k^i))^\top,$$

$\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h)$ depends on our choice of the quadrature rule (3.66), and A^Y is as in (3.84) and depends on the choice of approximation for the short rate (3.78).

In the remaining part of this section, we give three algorithms based on rectangle ($p = 1$), trapezoid ($p = 2$), and Simpson ($p = 4$) quadrature rules $S_{I_j}(t_k, T_i, \Delta)$ accompanied by short rate approximations of the corresponding orders. In all these cases it is not difficult to check that (3.95) satisfies Assumption 3.4.1 and that $\bar{Y}_k$ satisfy Assumption 3.4.6.

3.5.1 Algorithm of order $O(\Delta + h)$

The application of the composite rectangle rule to approximate the integrals $I_j(t_k, T_i)$ in (3.61) and $Z(t^*, T^*)$ in (3.52) yields

$$\begin{aligned}
\bar{\mathbb{S}}_{I_j}(t_k, T_{\ell_{k+1}}; \Delta, h) &= h\Delta_{\ell_{k+1},k}\bar{\sigma}_{\ell_{k+1},j}(t_k), \\
\bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h) &= \begin{cases} h\Delta_{\varrho_{k+1},k}\bar{\sigma}_{\varrho_{k+1},j}(t_k), & \text{if } T_{\ell_{k+1}} < t_k, \\ \bar{\mathbb{S}}_{I_j}(t_k, T_{\ell_{k+1}}; \Delta, h) + h\Delta\bar{\sigma}_{\varrho_{k+1},j}(t_k), & \text{otherwise,} \end{cases} \\
\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) &= \bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h) + h\Delta \sum_{m=\varrho_{k+1}+1}^{i} \bar{\sigma}_{m,j}(t_k), \\
i &= \varrho_{k+1} + 1, \ldots, N,\ j = 1, \ldots, d,
\end{aligned} \tag{3.96}$$

$$\bar{S}_Z(t^*, T^*, \Delta) = \bar{f}_M^{\varrho_M}\Delta_{\varrho_M,M} + \Delta \sum_{m=\varrho_{M+1}}^{N} \bar{f}_M^m. \tag{3.97}$$

By straightforward calculations one can show that the used rectangle rule satisfies the order conditions (3.63) and (3.77) with $p = 1$. We pay attention that we incorporated two cases in (3.96): when $\ell_{k+1} = \ell_k$ and hence $T_{\ell_{k+1}} \leq t_k$ and when (see also (3.93)) $\ell_{k+1} = \ell_k + 1$ and hence $T_{\ell_{k+1}} > t_k$.

We use the piecewise constant approximation of the short rate (cf. (3.78)):

$$\pi(t) = \sum_{l=0}^{\ell(t^*)} f(t, T_l)\chi_{t\in[T_l,T_{l+1})}, \quad t \in [t_0, t^*]. \tag{3.98}$$

The approximation (3.98) obviously satisfies the order condition (3.79) with $p = 1$. To satisfy Assumption 3.4.5 with $q = 1$, we, in particular, need to approximate the integral $\tilde{Y}_{t,x,0}(t + h)$ in (3.82) by $\bar{Y}_{t,x,0}(t + h)$ from (3.84) with local order $O(h^2)$. In the case of (3.98) the coefficient in the right-hand side of (3.81) $\pi(s; x^i,\ i = \ell(s)) = \sum_{l=0}^{\ell(t^*)} x^l \chi_{s\in[T_l,T_{l+1})} = x^{\ell(s)}$ is only piece-wise smooth. Further, according to the condition (3.93), we can have two cases: either an open interval (t_k, t_{k+1}) does not contain any node T_i of the T-grid or it contains a single node T_{ϱ_k}. In the former case we can approximate the integral $\tilde{Y}_{t,x,0}(t+h)$ in (3.82) by the left rectangle rule and we have $A^Y(t_k; \bar{f}_k^j,\ j = \ell_k, \ell_{k+1}; h) = h\bar{f}_k^{\ell_k}$ with the local error of order $O(h^2)$ as needed. In the second case to achieve the local error of order $O(h^2)$ despite lack of smoothness of $\pi(s; x^i,\ i = \ell(s))$, we split the integral $\tilde{Y}_{t,x,0}(t+h) = \tilde{Y}_{t,x,0}(T_{\ell_k+1}) + \tilde{Y}_{T_{\varrho_k},f(T_{\ell_k+1}),0}(t+h)$ and approximate the first integral by the left-rectangle rule and the second by the right-rectangle rule: $A^Y(t_k; \bar{f}_k^j,\ j = \ell_k, \ell_{k+1}; h) = \Delta_{\ell_{k+1},k}\bar{f}_k^{\ell_k} - \Delta_{\ell_k+1,k+1}\bar{f}_{k+1}^{\ell_{k+1}}$. Thus,

$$A^Y(t_k; \bar{f}_k^j, j = \ell_k, \ell_{k+1}; h) = \left(h \wedge \Delta_{\ell_{k+1},k}\right) \bar{f}_k^{\ell_k} - \left(0 \wedge \Delta_{\ell_k+1,k+1}\right) \bar{f}_{k+1}^{\ell_{k+1}}. \tag{3.99}$$

We note that despite the use of $\bar{f}_{k+1}^{\ell_{k+1}}$ in the right-hand side of (3.99) the method does not require to resolve any implicitness.

Assumption 3.4.5 with $q = 1$ can be checked for the scheme (3.95), (3.96), (3.99) following the standard, routine way (see Section 4.3 and also. [58, Chap. 2]).

The algorithm based on (3.95) and (3.96), (3.97), (3.99), we will call **Algorithm 3.5.1** for the option price (3.51)-(3.54). According to Theorem 4.1.8 from the next Chapter, this algorithm is of order $O(\Delta+h)$, which under the condition (see (3.93)) $\Delta = \alpha h$, $\alpha \geq 1$, resulting in $O(h)$. We also note that in the case $\Delta = h$ the short rate is readily available on the grid and its approximation is not needed. Algorithm 3.5.1 with $\Delta = h$ is analogous to the numerical methods for the HJM model considered in [35, 40, 31]. As it is shown in our numerical experiments (see Section 3.7), Algorithm 3.5.1 is less efficient than the new algorithms (Algorithms 3.5.2 and 3.5.3) which we propose in the next two sections.

Remark 3.5.1 *If we replace* A^Y *in* (3.99) *by*

$$A^Y(t_k; h) = h\bar{f}_k^{\ell_k} \tag{3.100}$$

then Assumption 3.4.5 with $q = 1$ *is not satisfied and we cannot guarantee closeness of* $\bar{Y}_k$ *and* $\tilde{Y}(t_k)$*. Nevertheless,* $\bar{Y}_k$ *from* (3.100) *still apparently approximates* $Y(t_k)$ *so that the overall algorithm for computing the option price* (3.51)-(3.54) *remains of weak order* $O(\Delta + h)$*. This can be justified by some nonrigorous arguments and this was also demonstrated in our numerical experiments. To obtain such a result rigorously, we need to conduct convergence proof without using the intermediate finite-dimensional SDEs* (3.67), (3.81)*. We do not pursue this direction in the thesis. At the same time, we note that in all our numerical tests the scheme using* A^Y *from* (3.99) *gave more accurate results than the scheme with* A^Y *from* (3.100) *in the cases when the* T_i *nodes do not belong to the* t*-grid. Otherwise* A^Y *in* (3.99) *and* A^Y *in* (3.100) *obviously coincide.*

3.5.2 Algorithm of order $O(\Delta^2 + h)$

In this section we use quadrature rules (3.66), (3.76) and a short rate approximation (3.78) of order $O(\Delta^2)$.

We aim at applying the standard composite trapezoid rule to the integrals $I_j(s, T_i)$ in (3.61) and (3.64). The trapezoid rule requires that each of the integration subintervals $[T_{\ell(s)}, s]$, $[s, T_{\varrho(s)}]$, $[T_{\varrho(s)}, T_{\varrho(s)+1}]$, $\ldots$, $[T_{i-1}, T_i]$ span at least two nodes on the T-grid. However, the integration intervals $[T_{\ell(s)}, s]$ and $[s, T_{\varrho(s)}]$ usually contain just a single node on the T-grid: $T_{\ell(s)}$ and $T_{\varrho(s)}$, respectively. We resolve this issue by applying the right and left rectangle rules on these two intervals, respectively. Thus, the quadrature rule $S_{I_j}(s, T_i, \Delta)$ takes the form for $s \in [t_0, t^*]$, $i = \ell(s), \ldots, T^*$:

$$S_{I_j}(s, T_{\ell(s)}, \Delta) = \big(T_{\ell(s)} - s\big)\, \sigma_j(s, T_{\ell(s)}), \tag{3.101}$$

$$S_{I_j}(s, T_i, \Delta) = (T_{\varrho(s)} - s)\sigma_j(s, T_{\varrho(s)}) + \frac{\Delta}{2} \sum_{m=\varrho(s)}^{i-1} [\sigma_j(s, T_m) + \sigma_j(s, T_{m+1})] \quad \text{for } s \le T_i,$$

$$j = 1, \ldots, d.$$

This quadrature rule satisfies the order condition (3.63) with $p = 2$. To this end, we recall that left and right rectangle rules have local order two and we use them here on one or two integration steps only while the trapezoid rule has local order three and the composite trapezoid rule is of order two.

To ensure that (3.95) satisfies Assumption 3.4.5 with $q = 1$, we, in particular, need to approximate the integral $\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta)ds$ by $\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h)$ on a single step with weak order $O(h^2)$. If the node $T_{\ell_{k+1}}$ is not between t_k and t_{k+1} (due to (3.93) it cannot be more than one T-node in (t_k, t_{k+1})), it is sufficient to approximate the integral by the left rectangle rule and put $\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) = h\bar{S}_{I_j}(t_k, T_i, \Delta)$, where $\bar{S}_{I_j}(t_k, T_i, \Delta)$ is of the form (3.101) but with $\bar{\sigma}_{m,j}(t_k)$ instead of $\sigma_j(s, T_m)$. However, if $T_{\ell_{k+1}} > t_k$ then due to (3.101) we apply one integration rule on $[t_k, T_{\ell_{k+1}}]$ and the other on $[T_{\ell_{k+1}}, T_{\varrho_{k+1}}]$, which causes loss of smoothness of the integrand $\tilde{S}_{I_j}(s, T_i, \Delta)$. To reach the required order $O(h^2)$, we construct the approximation using the following guidance. First, we split integrals and apply the corresponding integration rules to $\tilde{S}_{I_j}(s, T_i, \Delta)$ according to (3.101). Thus, for $i > \ell_{k+1}$, we obtain

$$\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta)ds \tag{3.102}$$

$$= \int_{t_k}^{T_{\ell_{k+1}}} \tilde{S}_{I_j}(s, T_i, \Delta)ds + \int_{T_{\ell_{k+1}}}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta)ds$$

$$
\begin{aligned}
&= \int_{t_k}^{T_{\ell_{k+1}}} \left((T_{\ell_{k+1}} - s)\tilde{\sigma}_{\ell_{k+1},j}(s) + \frac{\Delta}{2} \sum_{m=l_{k+1}}^{i-1} [\tilde{\sigma}_{m,j}(s) + \tilde{\sigma}_{m+1,j}(s)] \right) ds \\
&+ \int_{T_{\ell_{k+1}}}^{t_{k+1}} \left((T_{\varrho_{k+1}} - s)\tilde{\sigma}_{\varrho_{k+1},j}(s) + \frac{\Delta}{2} \sum_{m=\varrho_{k+1}}^{i-1} [\tilde{\sigma}_{m,j}(s) + \tilde{\sigma}_{m+1,j}(s)] \right) ds \\
= &\int_{t_k}^{T_{\ell_{k+1}}} \left[(T_{\ell_{k+1}} - s)\tilde{\sigma}_{\ell_{k+1},j}(s) + \frac{\Delta}{2}\tilde{\sigma}_{\ell_{k+1},j}(s) + \frac{\Delta}{2}\tilde{\sigma}_{\varrho_{k+1},j}(s) \right] ds \\
&+ \int_{T_{\ell_{k+1}}}^{t_{k+1}} (T_{\varrho_{k+1}} - s)\tilde{\sigma}_{\varrho_{k+1},j}(s) + \frac{\Delta}{2} \int_{t_k}^{t_{k+1}} \sum_{m=\varrho_{k+1}}^{i-1} [\tilde{\sigma}_{m,j}(s) + \tilde{\sigma}_{m+1,j}(s)] \, .
\end{aligned}
$$

Lets us consider each term of (3.102) along with the corresponding approximation separately:

$$
\begin{gathered}
\int_{t_k}^{T_{\ell_{k+1}}} \left[(T_{\ell_{k+1}} - s)\tilde{\sigma}_{\ell_{k+1},j}(s) + \frac{\Delta}{2}\tilde{\sigma}_{\ell_{k+1},j}(s) \right] \\
\approx \tilde{\sigma}_{\ell_{k+1},j}(t_k) \int_{t_k}^{T_{\ell_{k+1}}} \left(T_{\ell_{k+1}} - s + \frac{\Delta}{2} \right) ds = (T_{\ell_{k+1}} - t_k)\frac{T_{\ell_{k+1}} - t_k + \Delta}{2}\tilde{\sigma}_{\ell_{k+1},j}(t_k);
\end{gathered}
$$

$$
\begin{gathered}
\int_{t_k}^{T_{\ell_{k+1}}} \frac{\Delta}{2}\tilde{\sigma}_{\varrho_{k+1},j}(s)ds + \int_{T_{\ell_{k+1}}}^{t_{k+1}} (T_{\varrho_{k+1}} - s)\tilde{\sigma}_{\varrho_{k+1},j}(s) \\
\approx \tilde{\sigma}_{\varrho_{k+1},j}(t_k) \left(\int_{t_k}^{T_{\ell_{k+1}}} \frac{\Delta}{2} ds + \int_{T_{\ell_{k+1}}}^{t_{k+1}} (T_{\varrho_{k+1}} - s)ds \right) = \frac{\Delta}{2}[t_{k+2} - T_{\ell_{k+1}}]\tilde{\sigma}_{\varrho_{k+1},j}(t_k);
\end{gathered}
$$

$$
\frac{\Delta}{2} \int_{t_k}^{t_{k+1}} \sum_{m=\varrho_{k+1}}^{i-1} [\tilde{\sigma}_{m,j}(s) + \tilde{\sigma}_{m+1,j}(s)] \approx h\frac{\Delta}{2} \sum_{m=\varrho_{k+1}}^{i-1} [\tilde{\sigma}_{m,j}(t_k) + \tilde{\sigma}_{m+1,j}(t_k)] \, ;
$$

This yields the following approximation for $\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta)ds$:

$$
\begin{aligned}
\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta)ds \approx{}& (T_{\ell_{k+1}} - t_k)\frac{T_{\ell_{k+1}} - t_k + \Delta}{2}\tilde{\sigma}_{\ell_{k+1},j}(t_k) \\
&+ \frac{\Delta}{2}[t_{k+2} - T_{\ell_{k+1}}]\tilde{\sigma}_{\varrho_{k+1},j}(t_k) + h\frac{\Delta}{2} \sum_{m=\varrho_{k+1}}^{i-1} [\tilde{\sigma}_{m,j}(t_k) + \tilde{\sigma}_{m+1,j}(t_k)]
\end{aligned}
\tag{3.103}
$$

For $i = \ell_{k+1}$, we have

$$
\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_{\ell_{k+1}}, \Delta)ds \tag{3.104}
$$

$$
\begin{aligned}
&= \int_{t_k}^{T_{\ell_{k+1}}} \tilde{S}_{I_j}(s, T_{\ell_{k+1}}, \Delta) ds + \int_{T_{\ell_{k+1}}}^{t_{k+1}} \tilde{S}_{I_j}(s, T_{\ell_{k+1}}, \Delta) ds \\
&= \int_{t_k}^{T_{\ell_{k+1}}} (T_{\ell_{k+1}} - s)\tilde{\sigma}_{\ell_{k+1},j}(s) ds + \int_{T_{\ell_{k+1}}}^{t_{k+1}} (s - T_{\ell_{k+1}})\tilde{\sigma}_{\ell_{k+1},j}(s) ds \\
&\approx \left(T_{\ell_{k+1}} - t_k\right)\left(T_{\ell_{k+1}} - t_{k+1}\right)\tilde{\sigma}_{\ell_{k+1},j}(t_k).
\end{aligned}
$$

As a result $\bar{\mathbb{S}}_{I_j}(t_k, T_i, \Delta)$ in (3.95) is taken of the form:

$$
\bar{\mathbb{S}}_{I_j}(t_k, T_{\ell_{k+1}}; \Delta, h) = \begin{cases} h\Delta_{\ell_{k+1},k}\bar{\sigma}_{\ell_{k+1},j}(t_k), & \text{if } T_{\ell_{k+1}} \leq t_k, \\ \Delta_{\ell_{k+1},k}\Delta_{\ell_{k+1},k+1}\bar{\sigma}_{\ell_{k+1},j}(t_k), & \text{otherwise,} \end{cases}, \tag{3.105}
$$

$$
\bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h) = \begin{cases} h\Delta_{\varrho_{k+1},k}\bar{\sigma}_{\varrho_{k+1},j}(t_k), & \text{if } T_{\ell_{k+1}} \leq t_k, \\ \Delta_{\ell_{k+1},k}\ \frac{\Delta_{\varrho_{k+1},k}}{2}\bar{\sigma}_{\ell_{k+1},j}(t_k) - \Delta_{\ell_{k+1},k+2}\ \frac{\Delta}{2}\bar{\sigma}_{\varrho_{k+1},j}(t_k), & \text{otherwise,} \end{cases}
$$

$$
\begin{aligned}
\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) &= \bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h) \\
&\quad + h\frac{\Delta}{2}\left(\bar{\sigma}_{\varrho_{k+1},j}(t_k) + 2\sum_{m=\varrho_{k+1}+1}^{i-1}\bar{\sigma}_{m,j}(t_k) + \bar{\sigma}_{i,j}(t_k)\right), \\
&\qquad\qquad i = \varrho_{k+1}+1, \ldots, N, \quad j = 1, \ldots, d.
\end{aligned}
$$

Remark 3.5.2 *We note that in* (3.105) *we can substitute the rules* $\bar{\mathbb{S}}_{I_j}(t_k, T_{\ell_{k+1}}; \Delta, h)$ *and* $\bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h)$ *by the slightly simpler ones*

$$
\bar{\mathbb{S}}_{I_j}(t_k, T_{\ell_{k+1}}; \Delta, h) = h\Delta_{\ell_{k+1},k}\bar{\sigma}_{\ell_{k+1},j}(t_k),
$$

$$
\bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h) = \begin{cases} h\Delta_{\varrho_{k+1},k}\bar{\sigma}_{\varrho_{k+1},j}(t_k), & \textit{if } T_{\ell_{k+1}} \leq t_k, \\ \Delta_{\ell_{k+1},k}\ \frac{\Delta}{2}\ \left[\bar{\sigma}_{\ell_{k+1},j}(t_k) + \bar{\sigma}_{\varrho_{k+1},j}(t_k)\right], & \textit{otherwise,} \end{cases}
$$

without losing order of convergence (cf. (3.103), (3.104)*). But we suggest to use* *the expression in* (3.105) *due to its better symmetry properties in the case of the coinciding maturity and calendar time grids and also potentially better accuracy.*

By a similar reasoning used to derive (3.101), we obtain the corresponding quadrature rule $S_Z(t^*, T^*, \Delta)$ (see (3.76)). Namely, we apply the right-rectangle rule on the integration interval $\left[t_M, T_{\varrho_M}\right]$ and the composite trapezoid rule on the rest of the integration interval, i.e.,

$$\bar{S}_Z(t^*, T^*, \Delta) = \bar{f}_M^{\varrho_M} \Delta_{\varrho_M, M} + \frac{\Delta}{2}\left(\bar{f}_M^{\varrho_M} + 2\sum_{j=\varrho_M+1}^{N-1} \bar{f}_M^j + \bar{f}_M^N\right). \tag{3.106}$$

It is not difficult to show that the combination of rectangle and trapezoid rules used for deriving (3.106) satisfies the order condition (3.77) with $p = 2$.

We use linear interpolation for the short rate in (3.78):

$$\pi(t) = \sum_{l=0}^{\ell(t^*)} \left[\frac{t - T_l}{\Delta} f(t, T_{l+1}) + \frac{T_{l+1} - t}{\Delta} f(t, T_l)\right] \chi_{t \in [T_l, T_{l+1})}, \; t \in [t_0, t^*]. \tag{3.107}$$

The approximation (3.107) obviously satisfies the order condition (3.79) with $p = 2$. As in the case of Algorithm 3.5.1, the coefficient in the right-hand side of (3.81) is also only piece-wise smooth here. Consider first the case when the node $T_{\ell_{k+1}}$ is not between t_k and t_{k+1}. The application of the left-rectangle rule to the integral $\int_{t_k}^{t_{k+1}} \tilde{\pi}(s)ds$ has the error $O(h^2/\Delta)$, i.e., it does not lead to a uniform error estimate $O(h^2)$ required by Assumption 3.4.5. To ensure that the estimates $O(h^2)$ in Assumptions 3.4.5 are uniform in Δ, we use the following guidance:

$$\begin{aligned}
\int_{t_k}^{t_{k+1}} \tilde{\pi}(s)ds &= \int_{t_k}^{t_{k+1}} \left[\frac{s - T_{\ell_{k+1}}}{\Delta} \tilde{f}^{\varrho_{k+1}}(s) + \frac{T_{\varrho_{k+1}} - s}{\Delta} \tilde{f}^{\ell_{k+1}}(s)\right] ds \\
&\approx \tilde{f}^{\varrho_{k+1}}(t_k) \int_{t_k}^{t_{k+1}} \frac{s - T_{\ell_{k+1}}}{\Delta} ds + \tilde{f}^{\ell_{k+1}}(t_k) \int_{t_k}^{t_{k+1}} \frac{T_{\varrho_{k+1}} - s}{\Delta} ds \\
&= \tilde{f}^{\varrho_{k+1}}(t_k) h \frac{t_{k+1/2} - T_{\ell_{k+1}}}{\Delta} + \tilde{f}^{\ell_{k+1}}(t_k) h \frac{T_{\varrho_{k+1}} - t_{k+1/2}}{\Delta}.
\end{aligned}$$

So, in this case we put $A^Y(t_k; h) = h\left[\frac{\Delta_{\rho_{k+1}, k+1/2}}{\Delta} \bar{f}_k^{\ell_{k+1}} - \frac{\Delta_{\ell_{k+1}, k+1/2}}{\Delta} \bar{f}_k^{\varrho_{k+1}}\right]$. In the other case, i.e., if $T_{\ell_{k+1}} > t_k$, we split the integral $\int_{t_k}^{t_{k+1}} \tilde{\pi}(s)ds = \int_{t_k}^{T_{\ell_{k+1}}} \tilde{\pi}(s)ds +$

$\int_{T_{\ell_{k+1}}}^{t_{k+1}} \tilde{\pi}(s)ds$ and approximate each of them separately as we did in constructing (3.99),i.e.

$$\int_{t_k}^{T_{\ell_{k+1}}} \tilde{\pi}(s)ds = \int_{t_k}^{T_{\ell_{k+1}}} \left[\frac{s - T_{\ell_k}}{\Delta}\tilde{f}^{l_{k+1}}(s) + \frac{T_{l_{k+1}} - s}{\Delta}\tilde{f}^{\ell_k}(s)\right] ds$$
$$\approx \tilde{f}^{l_{k+1}}(t_k)\int_{t_k}^{T_{\ell_{k+1}}} \frac{s - T_{\ell_k}}{\Delta}ds + \tilde{f}^{\ell_k}(t_k)\int_{t_k}^{T_{\ell_{k+1}}} \frac{T_{l_{k+1}} - s}{\Delta}ds$$
$$= (T_{l_{k+1}} - t_k)\left(\frac{t_k - T_{l_k-1}}{2\Delta}\tilde{f}^{l_{k+1}}(t_k) + \frac{T_{l_{k+1}} - t_k}{2\Delta}\tilde{f}^{\ell_k}(t_k)\right);$$

$$\int_{T_{\ell_{k+1}}}^{t_{k+1}} \tilde{\pi}(s)ds = \int_{T_{\ell_{k+1}}}^{t_{k+1}} \left[\frac{s - T_{\ell_{k+1}}}{\Delta}\tilde{f}^{\varrho_{k+1}}(s) + \frac{T_{\varrho_{k+1}} - s}{\Delta}\tilde{f}^{\ell_{k+1}}(s)\right] ds$$
$$\approx \tilde{f}^{\varrho_{k+1}}(t_{k+1})\int_{T_{\ell_{k+1}}}^{t_{k+1}} \frac{s - T_{\ell_{k+1}}}{\Delta}ds + \tilde{f}^{\ell_{k+1}}(t_{k+1})\int_{t_k}^{T_{\ell_{k+1}}} \frac{T_{\varrho_{k+1}} - s}{\Delta}ds$$
$$= (t_{k+1} - T_{l_{k+1}})\left(\frac{t_{k+1} - T_{l_{k+1}}}{2\Delta}\tilde{f}^{l_{k+1}}(t_{k+1}) + \frac{T_{\varrho_{k+1}+1} - t_{k+1}}{2\Delta}\tilde{f}^{\ell_k}(t_{k+1})\right);$$

As a result, we arrive at

$$A^Y(t_k;h) = \begin{cases} h\left[\frac{\Delta_{\rho_{k+1},k+1/2}}{\Delta}\bar{f}_k^{\ell_{k+1}} - \frac{\Delta_{\ell_{k+1},k+1/2}}{\Delta}\bar{f}_k^{\varrho_{k+1}}\right], & \text{if } T_{\ell_{k+1}} \le t_k, \\ \Delta_{\ell_{k+1},k}\left[\frac{\Delta_{\ell_{k+1},k}}{2\Delta}\bar{f}_k^{\ell_k} - \frac{\Delta_{\ell_k-1,k}}{2\Delta}\bar{f}_k^{\ell_{k+1}}\right] & \\ -\Delta_{\ell_{k+1},k+1}\left[\frac{\Delta_{\varrho_{k+1}+1,k+1}}{2\Delta}\bar{f}_{k+1}^{\ell_{k+1}} - \frac{\Delta_{\ell_{k+1},k+1}}{2\Delta}\bar{f}_{k+1}^{\varrho_{k+1}}\right] & \text{otherwise.} \end{cases} \tag{3.108}$$

Assumption 3.4.5 with $q = 1$ can be checked for the scheme (3.95), (3.105), (3.108) following the standard way which we describe in Section 4.3.

The algorithm based on (3.95) and (3.105), (3.106), (3.108) we will call **Algorithm 3.5.2** for the option price (3.51)-(3.54). According to Theorem 4.1.8 from the next Chapter, this algorithm is of order $O(\Delta^2 + h)$. In practice (see Remark 3.4.7) we choose $\Delta = \alpha\sqrt{h}$ with $\alpha > 0$ such that (3.93) is satisfied, which results in the algorithm's accuracy $O(h)$. In our experiments (see Section 3.7) Algorithm 3.5.2 outperformed Algorithm 3.5.1.

3.5.3 Algorithm of order $O(\Delta^4 + h)$

At the beginning of Section 3.4.1 we made the assumption that there is a sufficient number of nodes T_i between t^* and T^* which ensures that we have enough nodes on the T-grid for using the quadrature rules (3.66) and (3.76) and the short rate approximations (3.78) of the required accuracy. This assumption gives an unnecessary restriction for using higher-order algorithms in practice and we now demonstrate how it can be relaxed. To this end, we introduce N' instead of N in the method (3.72) as the number of discretization nodes on T-grid:

$$N' := N \vee \max_{0<i\leq N} \kappa(t^*, T_i) \vee (\ell(t^*) + \theta)\,, \tag{3.109}$$

where $\kappa(t^*, T_i)$ and θ are as in (3.66) and (3.78), respectively. Also, in (3.76) we can put N' instead of N and if required increase N' further to be able to approximate the integral $Z(t^*, T^*)$ on the left-hand side of (3.76) with the prescribed accuracy. As a result, we avoid the restriction on how close t^* can be to T^*. It is clear that this extension of the T-grid by a fixed number of nodes in the case of large Δ does not influence our theoretical results.

Without re-writing the Euler-type scheme (3.95), we will assume in this section that we run it for $i = \ell_{k+1}, \ldots, N'$ instead of $i = \ell_{k+1}, \ldots, N$.

We are aiming at constructing an algorithm of order $O(\Delta^4 + h)$ and would like to exploit the standard composite Simpson rule for approximation of the integrals $I_j(s, T_i) = \int_s^{T_i} \sigma_j(s, u)du$ from (3.61) and (3.64). The Simpson rule needs three nodes per integration step. But the integrals $I_j(s, T_{\ell(s)})$, $I_j(s, T_{\varrho(s)})$, and $I_j(s, T_{\varrho(s)+1})$ are over the intervals which have just one or two nodes on the T-grid under (3.93).

We first consider the integrals $I_j(s, T_i) = \int_s^{T_i} \sigma_j(s, u)du$ with $T_i = T_{\ell(s)}$, $T_{\varrho(s)}$, and $T_{\varrho(s)+1}$, which we approximate by quadrature rules $S_{I_j}(s, T_i, \Delta)$ of the form

$$S_{I_j}(s, T_i, \Delta) = (T_i - s)\left[\beta_1^i \sigma_j(s, T_{\ell(s)}) + \beta_2^i \sigma_j(s, T_{\varrho(s)}) + \beta_3^i \sigma_j(s, T_{\varrho(s)+1})\right], \tag{3.110}$$

where the coefficients β_1^i, β_2^i, β_3^i depend on the value of T_i. We require that (3.110) is of order 4, i.e., that (3.63) is satisfied for these three integrals with $p = 4$. One can show that the following sets of coefficients satisfy this order requirement:

$$\beta_1^{\ell(s)} = \frac{5}{12} + \frac{5}{12}\frac{T_{\varrho(s)} - s}{\Delta} + \frac{1}{6}\frac{\left(T_{\varrho(s)} - s\right)^2}{\Delta^2}, \tag{3.111}$$

$$\beta_2^{\ell(s)} = \frac{2}{3} - \frac{1}{3}\frac{T_{\varrho(s)} - s}{\Delta} - \frac{1}{3}\frac{\left(T_{\varrho(s)} - s\right)^2}{\Delta^2}, \quad \beta_3^{\ell(s)} = -\frac{1}{12} - \frac{1}{12}\frac{T_{\varrho(s)} - s}{\Delta} + \frac{1}{6}\frac{\left(T_{\varrho(s)} - s\right)^2}{\Delta^2};$$

$$\begin{aligned} \beta_1^{\varrho(s)} &= \frac{1}{4}\frac{T_{\varrho(s)} - s}{\Delta} + \frac{1}{6}\frac{\left(T_{\varrho(s)} - s\right)^2}{\Delta^2}, \ \beta_2^{\varrho(s)} = 1 - \frac{1}{3}\frac{\left(T_{\varrho(s)} - s\right)^2}{\Delta^2}, \\ \beta_3^{\varrho(s)} &= -\frac{1}{4}\frac{T_{\varrho(s)} - s}{\Delta} + \frac{1}{6}\frac{\left(T_{\varrho(s)} - s\right)^2}{\Delta^2}; \end{aligned} \tag{3.112}$$

$$\beta_1^{\varrho(s)+1} = -\frac{1}{12} + \frac{1}{12}\frac{T_{\varrho(s)} - s}{\Delta} + \frac{1}{6}\frac{\left(T_{\varrho(s)} - s\right)^2}{\Delta^2}, \tag{3.113}$$

$$\beta_2^{\varrho(s)+1} = \frac{2}{3} + \frac{1}{3}\frac{T_{\varrho(s)} - s}{\Delta} - \frac{1}{3}\frac{\left(T_{\varrho(s)} - s\right)^2}{\Delta^2}, \ \beta_3^{\varrho(s)+1} = \frac{5}{12} - \frac{5}{12}\frac{T_{\varrho(s)} - s}{\Delta} + \frac{1}{6}\frac{\left(T_{\varrho(s)} - s\right)^2}{\Delta^2}.$$

Further, for $\varrho(s)+1 < i \leq N$ we write $I_j(s, T_i) = I_j(s, T_{\varrho(s)}) + I_j(T_{\varrho(s)}, T_i; s)$ with $I_j(T_{\varrho(s)}, T_i; s) := \int_{T_{\varrho(s)}}^{T_i} \sigma_j(s, u)du$; and we approximate the integral $I_j(T_{\varrho(s)}, T_i; s)$, $\varrho(s)+1 < i \leq N'$, by the composite Simpson rule $S_{I_j}(T_{\varrho(s)}, T_i, \Delta; s)$ if its integration interval spans an odd number of maturity time nodes:

$$\begin{aligned} S_{I_j}(T_{\varrho(s)}, T_i, \Delta; s) &= \frac{\Delta}{3}\left(\sigma_j(s, T_{\varrho(s)}) + 2\sum_{l=1}^{(i-\varrho(s))/2-1} \sigma_j(s, T_{\varrho(s)+2l}) \right. \\ &\quad \left. +4\sum_{l=1}^{(i-\varrho(s))/2} \sigma_j(s, T_{\varrho(s)+2l-1}) + \sigma_j(s, T_i)\right), \end{aligned} \tag{3.114}$$

and otherwise we apply the Simpson's 3/8 rule for the last four nodes:

$$\begin{aligned} S_{I_j}(T_{\varrho(s)}, T_i, \Delta; s) &= \frac{\Delta}{3}\left(\sigma_j(s, T_{\varrho(s)}) + 2\sum_{l=1}^{(i-\varrho(s)-1)/2-2} \sigma_j(s, T_{\varrho(s)+2l}) \right. \\ &\quad \left. +4\sum_{l=1}^{(i-\varrho(s)-1)/2-1} \sigma_j(s, T_{\varrho(s)+2l-1}) + \sigma_j(s, T_{i-3})\right) \\ &\quad +\frac{3\Delta}{8}\left(\sigma_j(s, T_{i-3}) + 3\sigma_j(s, T_{i-2}) + 3\sigma_j(s, T_{i-1}) + \sigma_j(s, T_i)\right). \end{aligned} \tag{3.115}$$

By straightforward calculations one can show that the quadrature rule (3.110), (3.114), and (3.115) satisfies the order condition (3.63) with $p = 4$.

To obtain $\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h)$ based on (3.110), (3.114), and (3.115), we need again to consider the two cases: when $\ell_{k+1} = \ell_k$ and hence $T_{\ell_{k+1}} \leq t_k$ and when (see also (3.93)) $\ell_{k+1} = \ell_k + 1$ and hence $T_{\ell_{k+1}} > t_k$. If $\ell_{k+1} = \ell_k$, the application of the left-rectangle rule to the integral $\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta; s)ds$ has the error $O(h^2)$ required by Assumption 3.4.5 with $\tilde{S}_{I_j}(t_k, T_i, \Delta; t_k)$ of the form (3.110), (3.114)-(3.115) with $\tilde{\sigma}_{l,j}(t_k)$ instead of $\sigma_j(t_k, T_l)$. Hence, we put $\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) := h\bar{S}_{I_j}(t_k, T_i, \Delta; t_k)$ with $\bar{S}_{I_j}(t_k, T_i, \Delta; t_k)$ having the form (3.110), (3.114)-(3.115) with $\sigma_j(t_k, T_l)$ replaced by $\bar{\sigma}_{l,j}(t_k)$. It is quite easy to show that this approximation has the required error $O(h^2)$. For instance, for $\bar{S}_{I_j}(t_k, T_i, \Delta; t_k)$ of the form (3.110), we need to show that this is true for the approximations of the following integrals (cf. (3.111)-(3.113)):

$$\int_{t_k}^{t_{k+1}} (T_i - s)ds = h\,(T_i - t_k) + O(h^2);$$

$$\int_{t_k}^{t_{k+1}} (T_i - s)\frac{T_{\varrho_k} - s}{\Delta}ds = \frac{h}{\Delta}\,(T_i - t_k)\left(T_{\varrho_k} - t_k\right) + R(h),$$

where for ϑ between t and $t + h$

$$R_1(h) = -\frac{h^2}{2}\frac{1}{\Delta}[(T_i - \vartheta) + \left(T_{\varrho_k} - \vartheta\right)],$$

and hence

$$|R_1(h)| \leq \frac{h^2}{2};$$

$$\int_{t_k}^{t_{k+1}} (T_i - s)\frac{\left(T_{\varrho_k} - s\right)^2}{\Delta^2}ds = \frac{h}{\Delta^2}(T_i - t_k)\left(T_{\varrho_k} - t_k\right)^2 + R_2(h),$$

where for ϑ between t and $t + h$

$$R_2(h) = -\frac{h^2}{2}\frac{1}{\Delta^2}[(T_i - \vartheta)^2 + \left(T_{\varrho_k} - \vartheta\right)(T_i - \vartheta)],$$

and hence

$$|R_2(h)| \leq \frac{h^2}{2};$$

As a result, when $\ell_{k+1} = \ell_k$, $\bar{\mathbb{S}}_{I_j}(t_k, T_i, \Delta)$ in (3.95) is taken of the form:

$$\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) = h\bar{S}_{I_j}(t_k, T_i, \Delta; t_k), \tag{3.116}$$

with $\bar{S}_{I_j}(t_k, T_i, \Delta; t_k)$ of the form (3.110), $i = \ell_{k+1},\ \varrho_{k+1},\ \varrho_{k+1}+1;$

$$\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) = \begin{cases} \bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h) + \bar{S}_{I_j}(T_{\varrho_{k+1}}, T_i, \Delta; s), \\ \text{ with } \bar{S}_{I_j}(T_{\varrho_k}, T_i, \Delta; s) \text{ of the form (3.114),} \\ i - \varrho_{k+1} + 1 \text{ is odd;} \\ \bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h) + \bar{S}_{I_j}(T_{\varrho_{k+1}}, T_i, \Delta; s), \\ \text{ with } \bar{S}_{I_j}(T_{\varrho_{k+1}}, T_i, \Delta; s) \text{ of the form (3.115),} \\ i - \varrho_{k+1} + 1 \text{ is even;} \\ \qquad i = \varrho_{k+1}+1, \ldots, N', \quad j = 1, \ldots, d. \end{cases}$$

In the case, when $\ell_{k+1} = \ell_k + 1$, we split the integral

$$\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta)ds = \int_{t_k}^{T_{\ell_{k+1}}} \tilde{S}_{I_j}(s, T_i, \Delta)ds + \int_{T_{\ell_{k+1}}}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta)ds$$

and plug in the corresponding approximations of $\tilde{S}_{I_j}(s, T_i, \Delta)$ according to (3.110), (3.114)-(3.115). Thus, when $\ell_{k+1} = \ell_k + 1$, we obtain

$$\begin{aligned} &\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_{\ell_{k+1}}, \Delta)ds \qquad (3.117) \\ &= \int_{t_k}^{T_{\ell_{k+1}}} \left(T_{\ell_{k+1}} - s\right) \left[\beta_1^{\varrho_k}\tilde{\sigma}_{\ell_k,j}(s) + \beta_2^{\varrho_k}\tilde{\sigma}_{\varrho_k,j}(s) + \beta_3^{\varrho_k}\tilde{\sigma}_{\varrho_k+1,j}(s)\right] ds \\ &+ \int_{T_{\ell_{k+1}}}^{t_{k+1}} \left(T_{\ell_{k+1}} - s\right) \left[\beta_1^{\ell_{k+1}}\tilde{\sigma}_{\ell_{k+1},j}(s) + \beta_2^{\ell_{k+1}}\tilde{\sigma}_{\varrho_{k+1},j}(s) + \beta_3^{\ell_{k+1}}\tilde{\sigma}_{\varrho_{k+1}+1,j}(s)\right] ds; \end{aligned}$$

$$
\begin{aligned}
&\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_{\varrho_{k+1}}, \Delta) ds \qquad (3.118)\\
&= \int_{t_k}^{T_{\ell_{k+1}}} \left(T_{\varrho_{k+1}} - s\right) \left[\beta_1^{\varrho_k+1} \tilde{\sigma}_{\ell_k,j}(s) + \beta_2^{\varrho_k+1} \tilde{\sigma}_{\varrho_k,j}(s) + \beta_3^{\varrho_k+1} \tilde{\sigma}_{\varrho_k+1,j}(s)\right] ds\\
&+ \int_{T_{\ell_{k+1}}}^{t_{k+1}} \left(T_{\varrho_{k+1}} - s\right) \left[\beta_1^{\varrho_{k+1}} \tilde{\sigma}_{\ell_{k+1},j}(s) + \beta_2^{\varrho_{k+1}} \tilde{\sigma}_{\varrho_{k+1},j}(s) + \beta_3^{\varrho_{k+1}} \tilde{\sigma}_{\varrho_{k+1}+1,j}(s)\right] ds;
\end{aligned}
$$

$$
\begin{aligned}
&\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_{\varrho_{k+1}+1}, \Delta) ds \qquad (3.119)\\
&= \int_{t_k}^{T_{\ell_{k+1}}} \Big(\left(T_{\ell_{k+1}} - s\right) \left[\beta_1^{\varrho_k} \tilde{\sigma}_{\ell_k,j}(s) + \beta_2^{\varrho_k} \tilde{\sigma}_{\varrho_k,j}(s) + \beta_3^{\varrho_k} \tilde{\sigma}_{\varrho_k+1,j}(s)\right]\\
&\qquad + \frac{\Delta}{3} \left(\tilde{\sigma}_{\ell_{k+1},j}(s) + 4\tilde{\sigma}_{\varrho_{k+1},j}(s) + \tilde{\sigma}_{\varrho_{k+1}+1,j}(s)\right) \Big) ds\\
&+ \int_{T_{\ell_{k+1}}}^{t_{k+1}} \left(T_{\varrho_{k+2}} - s\right) \left[\beta_1^{\varrho_{k+1}+1} \tilde{\sigma}_{\ell_{k+1},j}(s) + \beta_2^{\varrho_{k+1}+1} \tilde{\sigma}_{\varrho_{k+1},j}(s) + \beta_3^{\varrho_{k+1}+1} \tilde{\sigma}_{\varrho_{k+1}+1,j}(s)\right] ds;
\end{aligned}
$$

$$
\begin{aligned}
&\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta) ds \qquad (3.120)\\
&= \int_{t_k}^{T_{\ell_{k+1}}} \left(T_{\ell_{k+1}} - s\right) \left[\beta_1^{\varrho_k} \tilde{\sigma}_{\ell_k,j}(s) + \beta_2^{\varrho_k} \tilde{\sigma}_{\varrho_k,j}(s) + \beta_3^{\varrho_k} \tilde{\sigma}_{\varrho_k+1,j}(s)\right] ds\\
&\qquad + \int_{t_k}^{T_{\ell_{k+1}}} \frac{\Delta}{3} \Bigg(\tilde{\sigma}_{\ell_{k+1},j}(s) + 2 \sum_{l=1}^{(i-\ell_{k+1})/2-1} \tilde{\sigma}_{\ell_{k+1}+2l,j}(s)\\
&\qquad\qquad + 4 \sum_{l=1}^{(i-\ell_{k+1})/2} \tilde{\sigma}_{\ell_{k+1}+2l-1,j}(s) + \tilde{\sigma}_{i,j}(s) \Bigg) ds\\
&+ \int_{T_{\ell_{k+1}}}^{t_{k+1}} \left(T_{\varrho_{k+1}} - s\right) \left[\beta_1^{\varrho_{k+1}} \tilde{\sigma}_{\ell_{k+1},j}(s) + \beta_2^{\varrho_{k+1}} \tilde{\sigma}_{\varrho_{k+1},j}(s) + \beta_3^{\varrho_{k+1}} \tilde{\sigma}_{\varrho_{k+1}+1,j}(s)\right] ds\\
&\qquad + \int_{T_{\ell_{k+1}}}^{t_{k+1}} \frac{\Delta}{3} \Bigg(\tilde{\sigma}_{\ell_{k+1},j}(s) + 2 \sum_{l=1}^{(i-\ell_{k+1}-1)/2-2} \tilde{\sigma}_{\ell_{k+1}+2l,j}(s)\\
&\qquad\qquad + 4 \sum_{l=1}^{(i-\ell_{k+1}-1)/2-1} \tilde{\sigma}_{\ell_{k+1}+2l-1,j}(s) + \tilde{\sigma}_{i-3,j}(s) \Bigg) ds\\
&\qquad + \int_{T_{\ell_{k+1}}}^{t_{k+1}} \frac{3\Delta}{8} \left(\tilde{\sigma}_{i-3,j}(s) + 3\tilde{\sigma}_{i-2,j}(s) + 3\tilde{\sigma}_{i-1,j}(s) + \tilde{\sigma}_{i,j}(s)\right) ds,
\end{aligned}
$$

for $i > \varrho_{k+1} + 1$ and $i - \ell_{k+1} + 1$ is odd.

Analogously to (3.120), we can derive the expression for $\int_{t_k}^{t_{k+1}} \tilde{S}_{I_j}(s, T_i, \Delta)ds$ when $i > \varrho_{k+1} + 1$ and $i - \ell_{k+1} + 1$ is even.

Next, we approximate (3.117)-(3.120) to obtain the required $\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h)$ analogously to how we have proceeded in constructing Algorithm 3.5.2. As a result, we obtain the corresponding expressions of $\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h)$ in (3.95) when $\ell_{k+1} = \ell_k$ of the form:

$$\begin{aligned}\bar{\mathbb{S}}_{I_j}(t_k, T_{\ell_{k+1}}; \Delta, h) = \Delta^2_{\ell_{k+1},k}\left(\bar{\beta}_1^{\varrho_k}\bar{\sigma}_{\ell_k,j}(t_k) + \bar{\beta}_2^{\varrho_k}\bar{\sigma}_{\varrho_k,j}(t_k) + \bar{\beta}_3^{\varrho_k}\bar{\sigma}_{\varrho_k+1,j}(t_k)\right) \\ +\Delta^2_{\ell_{k+1},k+1}\left(\bar{\beta}_1^{\ell_{k+1}}\bar{\sigma}_{\ell_{k+1},j}(t_k) + \bar{\beta}_2^{\ell_{k+1}}\bar{\sigma}_{\varrho_{k+1},j}(t_k) + \bar{\beta}_3^{\ell_{k+1}}\bar{\sigma}_{\varrho_{k+1}+1,j}(t_k)\right),\end{aligned} \tag{3.121}$$

where

$$\begin{aligned}\bar{\beta}_1^{\varrho_k} &= \frac{1}{12}\frac{\Delta_{\ell_{k+1},k}}{\Delta} + \frac{1}{24}\frac{\Delta^2_{\ell_{k+1},k}}{\Delta^2},\ \bar{\beta}_2^{\varrho_k} = \frac{1}{2} - \frac{1}{12}\frac{\Delta^2_{\ell_{k+1},k}}{\Delta^2}, \\ \bar{\beta}_3^{\varrho_k} &= -\frac{1}{12}\frac{\Delta_{\ell_{k+1},k}}{\Delta} + \frac{1}{24}\frac{\Delta^2_{\ell_{k+1},k}}{\Delta^2};\end{aligned}$$

$$\begin{aligned}\bar{\beta}_1^{\ell_{k+1}} &= -\frac{1}{2} - \frac{1}{4}\frac{\Delta_{\ell_{k+1},k+1}}{\Delta} - \frac{1}{24}\frac{\Delta^2_{\ell_{k+1},k+1}}{\Delta^2},\ \bar{\beta}_2^{\ell_{k+1}} = \frac{1}{3}\frac{\Delta_{\ell_{k+1},k+1}}{\Delta} + \frac{1}{12}\frac{\Delta^2_{\ell_{k+1},k+1}}{\Delta^2}, \\ \bar{\beta}_3^{\ell_{k+1}} &= -\frac{1}{12}\frac{\Delta_{\ell_{k+1},k+1}}{\Delta} - \frac{1}{24}\frac{\Delta^2_{\ell_{k+1},k+1}}{\Delta^2};\end{aligned}$$

$$\begin{aligned}\bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h) = \Delta_{\ell_{k+1},k}\left(\bar{\beta}_1^{\varrho_k+1}\bar{\sigma}_{\ell_k,j}(t_k) + \bar{\beta}_2^{\varrho_k+1}\bar{\sigma}_{\varrho_k,j}(t_k) + \bar{\beta}_3^{\varrho_k+1}\bar{\sigma}_{\varrho_k+1,j}(t_k)\right) \\ +\Delta_{\ell_{k+1},k+1}\left(\bar{\beta}_1^{\varrho_{k+1}}\bar{\sigma}_{\ell_{k+1},j}(t_k) + \bar{\beta}_2^{\varrho_{k+1}}\bar{\sigma}_{\varrho_{k+1},j}(t_k) + \bar{\beta}_3^{\varrho_{k+1}}\bar{\sigma}_{\varrho_{k+1}+1,j}(t_k)\right),\end{aligned} \tag{3.122}$$

where

$$\begin{aligned}\bar{\beta}_1^{\varrho_k+1} &= -\frac{1}{12}\Delta + \frac{1}{12}\frac{\Delta^2_{\ell_{k+1},k}}{\Delta} + \frac{1}{24}\frac{\Delta^3_{\ell_{k+1},k}}{\Delta^2}, \\ \bar{\beta}_2^{\varrho_k+1} &= \frac{2}{3}\Delta + \frac{1}{2}\Delta_{\ell_{k+1},k} - \frac{1}{12}\frac{\Delta^3_{\ell_{k+1},k}}{\Delta^2}, \\ \bar{\beta}_3^{\varrho_k+1} &= \frac{5}{12}\Delta - \frac{1}{12}\frac{\Delta^2_{\ell_{k+1},k}}{\Delta} + \frac{1}{24}\frac{\Delta^3_{\ell_{k+1},k}}{\Delta^2};\end{aligned}$$

$$
\begin{aligned}
\bar{\beta}_1^{\varrho_{k+1}} &= -\frac{5}{12}\Delta - \frac{1}{2}\Delta_{\ell_{k+1},k+1} - \frac{1}{4}\frac{\Delta^2_{\ell_{k+1},k+1}}{\Delta} - \frac{1}{24}\frac{\Delta^3_{\ell_{k+1},k+1}}{\Delta^2},\\
\bar{\beta}_2^{\varrho_{k+1}} &= -\frac{2}{3}\Delta + \frac{1}{3}\frac{\Delta^2_{\ell_{k+1},k+1}}{\Delta} + \frac{1}{12}\frac{\Delta^3_{\ell_{k+1},k+1}}{\Delta^2},\\
\bar{\beta}_3^{\varrho_{k+1}} &= -\frac{1}{12}\Delta - \frac{1}{12}\frac{\Delta^2_{\ell_{k+1},k+1}}{\Delta} - \frac{1}{24}\frac{\Delta^3_{\ell_{k+1},k+1}}{\Delta^2};
\end{aligned}
$$

$$
\begin{aligned}
\bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}+1}; \Delta, h) = {} & \Delta_{\ell_k+1,k}\left(\bar{\beta}_1^{\varrho_k}\bar{\sigma}_{\ell_k,j}(t_k) + \bar{\beta}_2^{\varrho_k}\bar{\sigma}_{\varrho_k,j}(t_k) + \bar{\beta}_3^{\varrho_k}\bar{\sigma}_{\varrho_k+1,j}(t_k)\right) \qquad (3.123)\\
& +\Delta_{\ell_k+1,k}\frac{\Delta}{3}\left(\bar{\sigma}_{\ell_{k+1},j}(t_k) + 4\bar{\sigma}_{\varrho_{k+1},j}(t_k) + \bar{\sigma}_{\varrho_{k+1}+1,j}(t_k)\right)\\
& +\Delta_{\ell_{k+1},k+1}\left(\bar{\beta}_1^{\varrho_{k+1}+1}\bar{\sigma}_{\ell_{k+1},j}(t_k) + \bar{\beta}_2^{\varrho_{k+1}+1}\bar{\sigma}_{\varrho_{k+1},j}(t_k) + \bar{\beta}_3^{\varrho_{k+1}+1}\bar{\sigma}_{\varrho_{k+1}+1,j}(t_k)\right),
\end{aligned}
$$

where $\bar{\beta}_1^{\varrho_k}$, $\bar{\beta}_2^{\varrho_k}$, $\bar{\beta}_3^{\varrho_k}$ as in (3.121), and

$$
\begin{aligned}
\bar{\beta}_1^{\varrho_{k+1}+1} &= -\frac{1}{3} - \frac{1}{2}\Delta - \frac{1}{4}\frac{\Delta^2_{\ell_{k+1},k+1}}{\Delta} - \frac{1}{24}\frac{\Delta^3_{\ell_{k+1},k+1}}{\Delta^2},\\
\bar{\beta}_2^{\varrho_{k+1}+1} &= -\frac{4}{3}\Delta + \frac{1}{3}\frac{\Delta^2_{\ell_{k+1},k+1}}{\Delta} + \frac{1}{12}\frac{\Delta^3_{\ell_{k+1},k+1}}{\Delta^2},\\
\bar{\beta}_3^{\varrho_{k+1}+1} &= -\frac{1}{3}\Delta - \frac{1}{12}\frac{\Delta^2_{\ell_{k+1},k+1}}{\Delta} - \frac{1}{24}\frac{\Delta^3_{\ell_{k+1},k+1}}{\Delta^2};
\end{aligned}
$$

Finally, if $i > \varrho_{k+1}+1$ and $i - \ell_{k+1} + 1$ is odd, we have.

$$
\begin{aligned}
& \bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) \qquad (3.124)\\
= {} & \Delta_{\ell_k+1,k}\left(\bar{\beta}_1^{\varrho_k}\bar{\sigma}_{\ell_k,j}(t_k) + \bar{\beta}_2^{\varrho_k}\bar{\sigma}_{\varrho_k,j}(t_k) + \bar{\beta}_3^{\varrho_k}\bar{\sigma}_{\varrho_k+1,j}(t_k)\right)\\
& +\Delta_{\ell_k+1,k}\frac{\Delta}{3}\Bigg(\tilde{\sigma}_{\ell_{k+1},j}(t_k) + 2\sum_{l=1}^{(i-\ell_{k+1})/2-1}\tilde{\sigma}_{\ell_{k+1}+2l,j}(t_k)\\
& +4\sum_{l=1}^{(i-\ell_{k+1})/2}\tilde{\sigma}_{\ell_{k+1}+2l-1,j}(t_k) + \tilde{\sigma}_{i,j}(t_k)\Bigg)\\
& +\Delta_{\ell_{k+1},k+1}\left(\bar{\beta}_1^{\varrho_{k+1}}\tilde{\sigma}_{\ell_{k+1},j}(t_k) + \bar{\beta}_2^{\varrho_{k+1}}\tilde{\sigma}_{\varrho_{k+1},j}(t_k) + \bar{\beta}_3^{\varrho_{k+1}}\tilde{\sigma}_{\varrho_{k+1}+1,j}(t_k)\right)\\
& +\Delta_{\ell_k+1,k+1}\frac{\Delta}{3}\Bigg(\tilde{\sigma}_{\ell_{k+1},j}(t_k) + 2\sum_{l=1}^{(i-\ell_{k+1}-1)/2-2}\tilde{\sigma}_{\ell_{k+1}+2l,j}(t_k)
\end{aligned}
$$

$$+4\sum_{l=1}^{(i-\ell_{k+1}-1)/2-1}\tilde{\sigma}_{\ell_{k+1}+2l-1,j}(t_k)+\tilde{\sigma}_{i,j}(t_k)\Bigg)$$
$$\Delta_{\ell_k+1,k+1}\frac{3\Delta}{8}\left(\tilde{\sigma}_{i-3,j}(t_k)+3\tilde{\sigma}_{i-2,j}(t_k)+3\tilde{\sigma}_{i-1,j}(t_k)+\tilde{\sigma}_{i,j}(t_k)\right),$$

where $\bar{\beta}_1^{\varrho_k}$, $\bar{\beta}_2^{\varrho_k}$, $\bar{\beta}_3^{\varrho_k}$ as in (3.121) and $\bar{\beta}_1^{\varrho_{k+1}}$, $\bar{\beta}_2^{\varrho_{k+1}}$, $\bar{\beta}_3^{\varrho_{k+1}}$ as in (3.122). Analogously to (3.124), the expression for $\bar{\mathbb{S}}_{I_j}(t_k,T_i;\Delta,h)$ can be derived in the case if $i>\varrho_{k+1}+1$ and $i-\ell_{k+1}+1$ is even.

Using (3.110), (3.114), and (3.115), we construct the quadrature rule $S_Z(t^*,T^*,\Delta)$ (see (3.76)) and arrive at

$$\bar{S}_Z(t^*,T_i,\Delta)=\Delta_{i,M}\left[\beta_1^i\bar{f}_M^{\ell_M}+\beta_2^i\bar{f}_M^{\varrho_M}+\beta_3^i\bar{f}_M^{\varrho_M+1}\right]; \tag{3.125}$$

$$\bar{S}_Z(T_{\varrho_M},T^*,\Delta)=\frac{\Delta}{3}\left(\bar{f}_M^{\varrho_M}+2\sum_{l=1}^{(N-\varrho_M-1)/2}\bar{f}_M^{\varrho_M+2l}+4\sum_{l=1}^{(N-\varrho_M+1)/2}\bar{f}_M^{\varrho_M+2l-1}+\bar{f}_M^N\right); \tag{3.126}$$

$$\begin{aligned}\bar{S}_Z(T_{\varrho_M},T^*,\Delta) &= \frac{\Delta}{3}\left(\bar{f}_M^{\varrho_M}+2\sum_{l=1}^{(N-\varrho_M)/2-2}\bar{f}_M^{\varrho_M+2l}+4\sum_{l=1}^{(N-\varrho_M)/2-1}\bar{f}_M^{\varrho_M+2l-1}+\bar{f}_M^{N-3}\right)\\ &\quad+\frac{3\Delta}{8}\left(\bar{f}_M^{N-3}+3\bar{f}_M^{N-2}+3\bar{f}_M^{N-1}+\bar{f}_M^N\right).\end{aligned} \tag{3.127}$$

Then we define $\bar{S}_Z(t^*,T^*,\Delta)$ to be used in the algorithm as

$$\bar{S}_Z(t^*,T^*,\Delta)=\begin{cases}(3.125),\ (3.112)\text{ with } i=\varrho_M \ \text{ if } N=\varrho_M,\\[2ex] (3.125),\ (3.113)\text{ with } i=\varrho_M+1\ \text{ if } N=\varrho_M+1,\\[2ex] \bar{S}_Z(t_M,T_{\varrho_M},\Delta)+\bar{S}_Z(T_{\varrho_M},T_N,\Delta)\ \text{ if } N>\varrho_M+1,\end{cases} \tag{3.128}$$

where $\bar{S}_Z(t_M,T_{\varrho_M},\Delta)$ is from (3.125), (3.113) with $i=\varrho_M$ and

$$\bar{S}_Z(T_{\varrho_M},T_N,\Delta)=\begin{cases}(3.126)\text{ if } N-\varrho_M+1 \text{ is odd,}\\[2ex] (3.127)\text{ if } N-\varrho_M+1\text{ is even.}\end{cases}$$

It is not difficult to show that the quadrature rules used for deriving (3.128) satisfy the order condition (3.77) with $p = 4$.

For the short rate approximation $\pi(t)$ (see (3.78)), we use cubic polynomial interpolation which obviously satisfies the order condition (3.79) with $p = 4$:

$$\pi(t) = \sum_{j=0}^{3} L_j(t) f(t, T_{\ell(t)+j}), \tag{3.129}$$

where

$$L_j(t) = \prod_{\substack{i=0 \\ i \neq j}}^{3} \frac{t - T_{\ell(t)+i}}{T_{\ell(t)+j} - T_{\ell(t)+i}}.$$

To obtain the corresponding $A^Y(t_k; h) = A^Y(t_k; \bar{f}_k^j, j = \ell_k, \ldots, \ell_{k+1}+3; h)$, we follow a similar guidance as the one used to obtain (3.108).

$$\begin{aligned}\int_{t_k}^{t_{k+1}} \tilde{\pi}(s)ds = \int_{t_k}^{t_{k+1}\wedge T_{\varrho_k}} & \left[\frac{(s - T_{\varrho_k})(s - T_{\varrho_k+1})(s - T_{\varrho_k+2})}{-6\Delta^3} \tilde{f}^{l_k}(s)\right. \\ & + \frac{(s - T_{\ell_k})(s - T_{\varrho_k+1})(s - T_{\varrho_k+2})}{2\Delta^3} \tilde{f}^{\varrho_k}(s) \\ \frac{(s - T_{\ell_k})(s - T_{\varrho_k})(s - T_{\varrho_k+2})}{-2\Delta^3} \tilde{f}^{\varrho_{k+1}} & + \left.\frac{(s - T_{\ell_k})(s - T_{\varrho_k})(s - T_{\varrho_k+1})}{6\Delta^3} \tilde{f}^{\varrho_{k+2}}\right] ds \\ + \int_{t_{k+1}\wedge T_{\varrho_k}}^{t_{k+1}} & \left[\frac{(s - T_{\varrho_{k+1}})(s - T_{\varrho_{k+1}+1})(s - T_{\varrho_{k+1}+2})}{-6\Delta^3} \tilde{f}^{l_{k+1}}(s)\right. \\ & + \frac{(s - T_{\ell_{k+1}})(s - T_{\varrho_{k+1}+1})(s - T_{\varrho_{k+1}+2})}{2\Delta^3} \tilde{f}^{\varrho_{k+1}} \\ & \frac{(s - T_{\ell_{k+1}})(s - T_{\varrho_{k+1}})(s - T_{\varrho_{k+1}+2})}{-2\Delta^3} \tilde{f}^{\varrho_{k+1}+1} \\ & + \left.\frac{(s - T_{\ell_{k+1}})(s - T_{\varrho_{k+1}})(s - T_{\varrho_{k+1}+1})}{6\Delta^3} \tilde{f}^{\varrho_{k+1}+2}\right] ds.\end{aligned} \tag{3.130}$$

Next, each of $\tilde{f}^i(s)$ in (3.130) is approximated, e.g. as $\tilde{f}^{l_k}(s) \approx \tilde{f}^{l_k}(t_k)$ and $\tilde{f}^{l_{k+1}}(s) \approx \tilde{f}^{l_{k+1}}(t_{k+1})$. Then, the integrals in (3.130) are evaluated exactly. In our numerical experiments (see Section 3.7), we approximate these integrals with the standard composite Simpson rule (cf. (3.114)), since it is exact for cubic polynomials. We do not write the expression of $A^Y(t_k; h)$ here but there is no difficulty to restore it.

The algorithm presented in this section, we will call **Algorithm 3.5.3** for the option price (3.51)-(3.54). Assumption 3.4.5 with $q = 1$ can be checked for this algorithm following the standard way. According to Theorem 4.1.8 from the next Section, Algorithm 3.5.3 is of order $O(\Delta^4+h)$. In practice (see Remark 3.4.7) we will choose $\Delta = \alpha\sqrt[4]{h}$ with $\alpha > 0$ such that (3.93) holds, which results in the algorithm's accuracy $O(h)$.

Remark 3.5.3 (*Complexity of the algorithms*) *Let us estimate computational complexity of the algorithms considered in this section. The number of operations in these algorithms is of order $O(MN)$. Then running times of Algorithm 3.5.1 with $\Delta = \alpha h$, Algorithm 3.5.2 with $\Delta = \alpha\sqrt{h}$, and Algorithm 3.5.3 with $\Delta = \alpha\sqrt[4]{h}$ are proportional to M^2, $M\sqrt{M}$, and $M\sqrt[4]{M}$, respectively. Also, it should be taken into account that Algorithms 3.5.2 and 3.5.3 require approximately twice and four times number of operations per t-step, respectively, than Algorithm 3.5.1. Hence one can expect that in reaching a similar accuracy Algorithm 3.5.3 is approximately $M^{3/4}/4$ faster than Algorithm 3.5.1 and Algorithm 3.5.2 is $\sqrt{M}/2$ faster than Algorithm 3.5.1. This is confirmed in our numerical experiments* (*see Section 3.7*).

Remark 3.5.4 *If in Algorithms 3.5.2 and 3.5.3 we substitute the Euler scheme* (3.95) *by a second-order* (*i.e.,* $q = 2$) *weak scheme* (*see (3.42) and also more examples of such schemes can be found in, e.g. [57, 58, 45]*), *then these modified algorithms (they should satisfy Assumption 3.4.5 with $q = 2$) will become of order* $O(h^2 + \Delta^2)$ *and* $O(h^2 + \Delta^4)$, *respectively. Choosing $\Delta = h$ and $\Delta = \alpha\sqrt{h}$ in the modified Algorithms 3.5.2 and 3.5.3, respectively, their accuracy becomes of order* $O(h^2)$.

3.6 Mean-square method

In most of the financial applications weak numerical methods, which we have considered in the previous sections, are sufficient. At the same time, mean-square methods (see Section 3.1.2) can be useful for simulating scenarios. Also, mean-square convergence of fully discrete approximations for the HJM model is of theoretical interest.

In this section we consider a mean-square method for (3.67)-(3.69) and in Section 4.2 we prove its convergence.

We consider an approximation $\bar{f}^i_{k+1}$ of $\tilde{f}^i(t_{k+1})$ from (3.67) (i.e., a full discretization of (3.43)-(3.44) in both T and t) of the form

$$\begin{gathered}
\bar{f}^i_0 = f^i_0, i = 0, \ldots, N, \\
\bar{f}^i_{k+1} = \bar{f}^i_k + \\
+A^i(t_k, T_i; \bar{f}^j_k,\ j = \ell_{k+1}, \ldots, \kappa(t_{k+1}, T_i) \vee i; h; W_l(s) - W_l(t_k), \\
l = 1, \ldots, d,\ t_k \le s \le t_{k+1}), \\
i = \ell_{k+1}, \ldots, N,\ k = 0, \ldots, M,
\end{gathered} \tag{3.131}$$

where the form of function A^i depends on the coefficients of (3.67)-(3.69), i.e., on σ and on choice of the quadrature rule S_{I_j}; $\kappa(t_k, T_i)$ is as in the quadrature (3.69). Note that in this section we use the same notation $\bar{f}^i_k$ for the mean-square approximation as the one we use for weak approximations in all the other sections of this paper. Since mean-square approximations of (3.67) are considered in this section only, this abuse of notation does not lead to any confusion.

As before, we put

$$\bar{f}^i_k = \bar{f}^i_{\mathbf{m}}, \qquad k = \mathbf{m} + 1, \ldots, M,\ 0 \le i \le \ell(t^*) - 1,$$

where $\mathbf{m} = \lceil (T_{i+1} - t_0)/h \rceil - 1$. Then the $N+1$-dimensional vector $\{\bar{f}^i_k, i = 0, \ldots, N\}$ is defined for all $k = 0, \ldots, M$.

We impose the following assumption on the one-step approximation $\bar{f}^i_{t,x}(t+h)$ of the method (3.131) for the solution $\tilde{f}^i_{t,x}(t+h)$ of (3.67) with the initial condition x given at time $t : \tilde{f}^i_{t,x}(t) = x^i$.

Assumption 3.6.1 Let

$$q_2 \ge \frac{1}{2}\,,\ q_1 \ge q_2 + \frac{1}{2}\,. \tag{3.132}$$

Suppose the one-step approximation $\bar{f}^i_{t,x}(t+h)$ has order of accuracy q_1 for expectation of the deviation and order of accuracy q_2 for the mean-square deviation; more precisely, for arbitrary $t_0 \le t \le t^* - h$, $x \in \mathbb{R}^{N+1}$ the following inequalities hold:

$$|E(\tilde{f}^i_{t,x}(t+h) - \bar{f}^i_{t,x}(t+h))| \le Ch^{q_1}\,, \tag{3.133}$$

$$\left[E|\tilde{f}^i_{t,x}(t+h) - \bar{f}^i_{t,x}(t+h))|^2\right]^{1/2} \le Ch^{q_2}, \tag{3.134}$$

$$i = 0, \dots N,$$

where $C > 0$ is a constant independent of h, Δ, *and* x.

Assumption 3.6.1 is analogous to the conditions of the fundamental theorem of mean-square convergence in [58, p. 4] (see Theorem 3.1.7). We note that C in (3.133)-(3.134) are independent of x while in the fundamental theorem such C depend on x. In our case it is natural to put C independent of x since the coefficients of (3.67) and their derivatives are uniformly bounded (see Assumptions 3.2.1-3.2.2). We also emphasize that the constants C in (3.133)-(3.134) do not depend on Δ.

To illustrate the results of this section, let us present a mean-square algorithm for (3.43)-(3.44) based on the mean-square Euler-type scheme:

$$\bar{f}^i_0 = f^i_0,\ i = 0, \dots, N, \quad \bar{Y}_0 = 0, \tag{3.135}$$

$$\bar{f}^i_{k+1} = \bar{f}^i_k + \sum_{j=1}^{d} \bar{\sigma}_{i,j}(t_k)\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) + h^{1/2} \sum_{j=1}^{d} \bar{\sigma}_{i,j}(t_k)\xi_{j,k+1},$$

$$i = \ell_{k+1}, \dots, N,$$

where $\xi_{j,k+1}$ are independent Gaussian random variables with zero mean and unit variance,

$$(\bar{\sigma}_{i,1}(t_k), \dots, \bar{\sigma}_{i,d}(t_k))^\top = (\sigma_1(t_k, T_i, \bar{f}^i_k), \dots, \sigma_d(t_k, T_i, \bar{f}^i_k))^\top,$$

and $\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h)$ depends on our choice of the quadrature rule (3.66). If

$\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h)$ is taken from (3.96) or (3.105) or from Algorithm 3.5.3 then $p = 1$, $p = 2$ or $p = 4$, respectively, and $q_1 = 2$ and $q_2 = 1$ under $h \leq \Delta$. The overall error of these algorithms are $O(\Delta + h^{1/2})$, $O(\Delta^2 + h^{1/2})$, and $O(\Delta^4 + h^{1/2})$, respectively.

3.7 Numerical examples

In this section we demonstrate accuracy and convergence properties of the algorithms from Section 3.5. We also compare computational costs of the algorithms. This comparison illustrates that the algorithms with higher-order quadrature rules are more efficient.

For illustration, we price an interest rate caplet (see Section 2.1.3) which is an interest rate derivative providing protection against an increase in an interest rate for a single period. Suppose a caplet is set at time t^* with payment date at T^* and has the unit nominal value and a strike K. The arbitrage price of the caplet is given by (3.48) with $t^* = s_k$ and $T^* = s_i$. The caplet parameters chosen for the experiments are $t^* = 1.0$, $T^* = 6.0$, $K = 0.03$.

A particular model within the HJM framework (3.43)-(3.44) is specified by a choice of the volatility function and initial forward rate curve. Here we consider two examples: a one-factor model with deterministic exponential volatility function (Vasicek model, see, e.g. [6, 15] and cf. Section 2.4) and a two-factor model with proportional volatility function (see, e.g. [36, 31, 62, 33]). The former one admits a closed-form formula for the caplet price.

The algorithms were implemented using C++ with GCC 3.4.3 compiler. The experiments were run on ALICE HPC Computer nodes of the University of Leicester, each with dual quad-core 2.67GHz Intel Xeon X5550 processor, 12 GB RAM, and OS 64-bit Scientific Linux 5.4. Uniform random numbers were generated with the additive lagged Fibonacci generator, $F(1279, 418)$ (for implementation see [58, 66]).

3.7.1 Vasicek model

We consider the one-factor HJM model (3.43)-(3.44) with the deterministic volatility function given by

$$\sigma(t,T) = \sigma \exp(-\kappa(T-t)), \tag{3.136}$$

and the initial forward curve defined as

$$f_0(T) = \exp(-\kappa(T-t_0))r_0 + (1-\exp(-\kappa(T-t_0)))\,\vartheta - \frac{\sigma^2}{2\kappa^2}\left(1-\exp(-\kappa(T-t_0))\right)^2, \tag{3.137}$$

where σ, κ, r_0, and ϑ are given positive constants.

Table 3.1: *Algorithm 3.5.1 for the Vasicek model.* Performance of Algorithm 3.5.1 with $\Delta = h$ in the case of the Vasicek model (3.136), (3.137) with parameters $\sigma = 0.02$, $r_0 = 0.05$, $\kappa = 1$ and $\theta = 1$ for pricing a unit nominal caplet with parameters $t_0 = 0$, $t^* = 1.0$, $T^* = 6.0$, $K = 0.03$. L is the number of independent runs in the Monte Carlo simulation (see (3.91)).

h	L	*error*	*CPU time, min*
0.2	10^7	$4.22 \times 10^{-2} \pm 2.80 \times 10^{-6}$	4.00×10^{-1}
0.1	10^7	$2.04 \times 10^{-2} \pm 3.06 \times 10^{-6}$	9.00×10^{-1}
0.05	10^7	$1.00 \times 10^{-2} \pm 3.19 \times 10^{-6}$	2.45×10^{0}
0.025	10^7	$4.98 \times 10^{-3} \pm 3.26 \times 10^{-6}$	7.73×10^{0}
0.0125	10^9	$2.48 \times 10^{-3} \pm 3.29 \times 10^{-7}$	2.30×10^{3}
0.00625	10^9	$1.24 \times 10^{-3} \pm 3.31 \times 10^{-7}$	8.32×10^{3}

It is known ((2.19) and also see, e.g. [6, 15, 68]) that a caplet corresponds to a put option on a zero-coupon bond. In [38] analytic expressions for the European option prices on zero-coupon and coupon bearing bonds under the Vasicek model are derived. In particular, the price of the caplet set at time t^* with payment date at T^*, unit nominal value and strike K is given by

$$F(t_0, f_0(\cdot); t^*, T^*) = P(t_0,t^*)\Phi(-c_P + \sigma_P) - (1 + K(T^* - t^*))P(t_0,T^*)\Phi(-c_P), \tag{3.138}$$

Table 3.2: *Algorithm 3.5.2 for the Vasicek model.* Performance of Algorithm 3.5.2 with $\Delta = \sqrt{h}$ in the case of the Vasicek model (3.136), (3.137) with the same parameters as in Table 3.1.

h	L	*error*	*CPU time, min*
0.2	10^7	$6.53 \times 10^{-3} \pm 2.72 \times 10^{-6}$	3.00×10^{-1}
0.1	10^7	$3.32 \times 10^{-3} \pm 2.99 \times 10^{-6}$	5.33×10^{-1}
0.05	10^7	$1.65 \times 10^{-3} \pm 3.11 \times 10^{-6}$	1.03×10^{0}
0.025	10^7	$8.29 \times 10^{-4} \pm 3.22 \times 10^{-6}$	2.13×10^{0}
0.0125	10^9	$4.13 \times 10^{-4} \pm 3.27 \times 10^{-7}$	4.65×10^{2}
0.00625	10^9	$2.07 \times 10^{-4} \pm 3.30 \times 10^{-7}$	1.09×10^{3}

Table 3.3: *Algorithm 3.5.3 for the Vasicek model.* Performance of Algorithm 3.5.3 with $\Delta = \alpha\sqrt[4]{h}$ in the case of the Vasicek model (3.136), (3.137) with the same parameters as in Table 3.1.

h	L	*error*	*CPU time, min*	α
0.2	10^7	$1.25 \times 10^{-3} \pm 2.53 \times 10^{-6}$	2.59×10^{-1}	9.97×10^{-1}
0.1	10^7	$6.28 \times 10^{-4} \pm 2.87 \times 10^{-6}$	4.10×10^{-1}	9.70×10^{-1}
0.05	10^7	$3.18 \times 10^{-4} \pm 3.09 \times 10^{-6}$	8.11×10^{-1}	9.76×10^{-1}
0.025	10^7	$1.56 \times 10^{-4} \pm 3.20 \times 10^{-6}$	1.59×10^{0}	9.43×10^{-1}
0.0125	10^9	$9.62 \times 10^{-5} \pm 3.26 \times 10^{-7}$	3.13×10^{2}	9.97×10^{-1}
0.00625	10^9	$4.71 \times 10^{-5} \pm 3.30 \times 10^{-7}$	5.96×10^{2}	9.70×10^{-1}

where $\Phi(\cdot)$ denotes the standard normal cumulative distribution function and

$$\sigma_P = \frac{\sigma}{\kappa}\sqrt{\frac{1-\exp(-2\kappa(t^*-t_0))}{2\kappa}}\left[1-\exp(-2\kappa(T^*-t^*))\right],$$

$$c_P = \frac{1}{\sigma_P}\ln\frac{(1+K(T^*-t^*))P(t_0,T^*)}{P(t_0,t^*)} + \frac{\sigma_P}{2}.$$

The values of parameters chosen in the experiments are $t_0 = 0$, $\sigma = 0.02$, $r_0 = 0.05$, $\kappa = 1$, and $\vartheta = 1$. The values $\kappa = 1$ and $\vartheta = 1$ are rather unrealistic from the financial point of view and are chosen for illustrative purposes. Under a more realistic choice of parameters, simulations are done with a particular time step h (see Table 3.4).

The results of the experiments with Algorithm 3.5.1 of order $O(\Delta + h)$, Algorithm 3.5.2 of order $O(\Delta^2+h)$, and Algorithm 3.5.3 of order $O(\Delta^4+h)$ are presented in Tables 3.1, 3.2, and 3.3, respectively. For Algorithms 3.5.1 and 3.5.2, we set $\Delta = h$ and $\Delta = \sqrt{h}$, respectively. For Algorithm 3.5.3, we set $\Delta = \alpha\sqrt[4]{h}$ with $\alpha > 0$ so that T-grid remains equally spaced:

$$\alpha = \frac{1}{\sqrt[4]{h}} \begin{cases} \frac{T^*-t_0}{\left\lfloor \frac{T^*-t_0}{\sqrt[4]{h}} \right\rfloor + 1}, \text{ if } \frac{T^*-t_0}{\sqrt[4]{h}} - \left\lfloor \frac{T^*-t_0}{\sqrt[4]{h}} \right\rfloor \geq \frac{1}{2}, \\ \frac{T^*-t_0}{\left\lfloor \frac{T^*-t_0}{\sqrt[4]{h}} \right\rfloor}, \text{ otherwise,} \end{cases} \tag{3.139}$$

where $\lfloor \cdot \rfloor$ denotes the integer part of a real number. It is clear that $\alpha \approx 1$.

As a result, the errors of all three algorithms become of order $O(h)$. In the tables, the values before "$\pm$" are estimates of the bias computed as the difference between the exact caplet price (3.138) and its sampled approximation (see (3.91)), while the values after "$\pm$" give half of the size of the confidence interval for the corresponding estimator with probability 0.95. The number of Monte Carlo runs L is chosen here so that the Monte Carlo error is small in comparison with the bias. The results from the tables are visualised in Figure 3.3. This confirms that the experimentally observed convergence rates for Algorithms 3.5.1, 3.5.2 and 3.5.3 are in agreement with the theoretical first order convergence in h (reference line "Order one"). We note that in the analysis of convergence of Algorithm 3.5.3 one has to take into account not only values of h but also of α. As expected, the experiments demonstrate that Algorithm 3.5.3 is the most computationally efficient among the three algorithms tested and also Algorithm 3.5.2 outperforms Algorithm 3.5.1. As it follows from Tables 3.1-3.4, for a fixed time step h the ratios of running times of the considered algorithms is in agreement with the theoretical prediction (see Remark 3.5.3).

In Table 3.4 we present results for $h = 0.1$ and $L = 10^9$ in the case of the more realistic choice of parameters $\kappa = 0.178$ and $\vartheta = 0.086$ of the Vasicek model. With

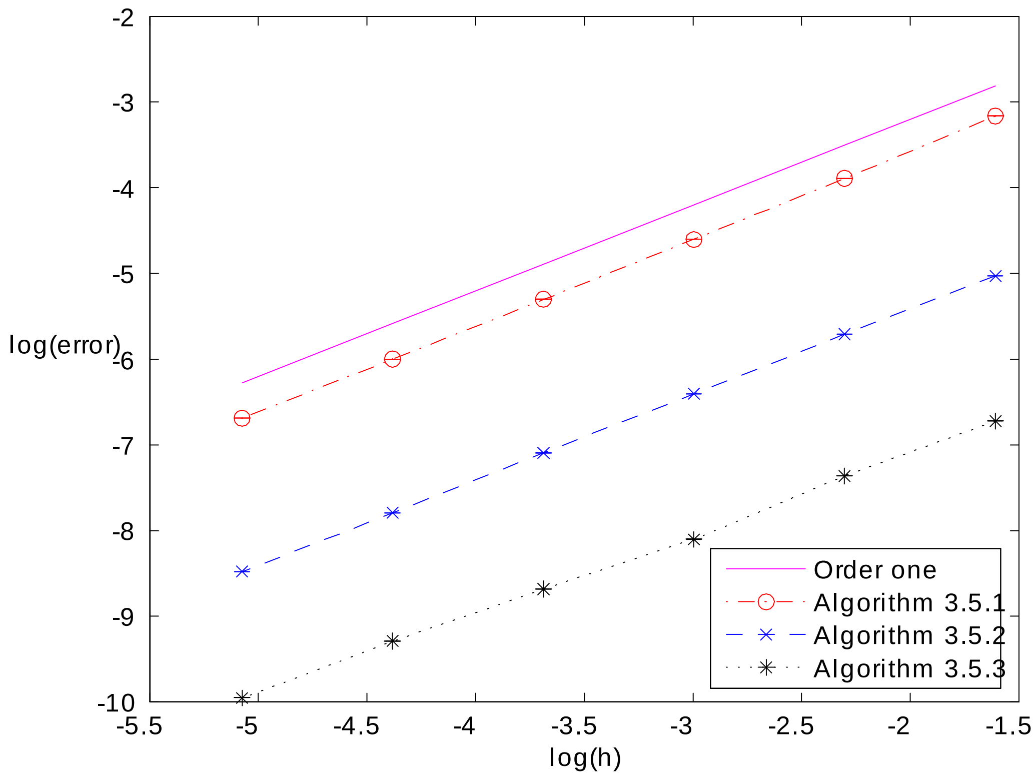

Figure 3.3: *Vasicek model:* Stepsize error diagram in log coordinates of the results from Tables 3.1, 3.2, and 3.3. $Log(error)$ is the log of estimates of the bias.

these parameters, the bias is very small, and if one would like to analyze it, e.g. for $h = 2^{-2} \times 5^{-1}$, then the number of Monte Carlo runs has to be increased up to 10^{11} in order to make the Monte Carlo error sufficiently smaller than the bias. We see from Table 3.4 that Algorithm 3.5.3 is more than twice faster and more accurate than Algorithm 3.5.1.

3.7.2 Proportional volatility model

Here we choose the volatility functions of the form

$$\sigma_j(t,T) = \sigma_j \exp(-\kappa_j(T-t)) \min\left(f(t,T), \Gamma\right), \tag{3.140}$$

where σ_j and κ_j are positive constants and Γ is a large positive number introduced to cap the proportional volatility in order to avoid an explosion of the forward-rate

Table 3.4: *Vasicek model.* Performance of the algorithms (3.5.1)-(3.5.3) with $h = 2^{-6}$ and $L = 10^9$ in the case of the Vasicek model (3.136), (3.137) with parameters $\sigma = 0.02$, $r_0 = 0.05$, $\kappa = 0.178$ and $\theta = 0.086$ for pricing a unit nominal caplet with parameters $t_0 = 0$, $t^* = 1.0$, $T^* = 6.0$, $K = 0.03$.

	h	L	*error*	*CPU time, min*
Algorithm 3.5.1	0.1	10^9	$-5.38 \times 10^{-4} \pm 2.90 \times 10^{-6}$	9.79×10^1
Algorithm 3.5.2	0.1	10^9	$1.75 \times 10^{-4} \pm 2.86 \times 10^{-6}$	5.51×10^1
Algorithm 3.5.3	0.1	10^9	$7.81 \times 10^{-5} \pm 2.89 \times 10^{-6}$	4.10×10^1

process (cf. Assumption 3.2.1 and also Remark 3.2.4). The volatility specification of the form (3.140) yields an approximately lognormal distribution of forward rates (cf. Section 2.3).

Table 3.5: *Algorithm 3.5.1 for the Proportional volatility model.* Performance of Algorithm 3.5.1 with $\Delta = h$ in the case of the proportional volatility model (3.140) with parameters (3.141) and with initial forward curve (3.142) for pricing a unit nominal caplet with parameters $t_0 = 0$, $t^* = 1.0$, $T^* = 6.0$, $K = 0.03$.

h	L	*error*	*CPU time, min*
0.2	10^9	$5.80 \times 10^{-4} \pm 2.57 \times 10^{-6}$	6.99×10^1
0.125	10^9	$3.64 \times 10^{-4} \pm 2.57 \times 10^{-6}$	1.21×10^2
0.1	10^9	$2.92 \times 10^{-4} \pm 2.57 \times 10^{-6}$	1.67×10^2
0.05	10^9	$1.48 \times 10^{-4} \pm 2.56 \times 10^{-6}$	4.84×10^2
0.025	10^9	$7.50 \times 10^{-5} \pm 2.56 \times 10^{-6}$	1.63×10^3

Let us note that in [33] a number of volatility models including one and two factors proportional volatility models are examined. The performance of the models is evaluated based on the accuracy of their out-of-sample price prediction and their ability to hedge caps and floors. This study reveals that in out-of-sample pricing accuracy the one- and two- factor proportional volatility models outperform the other competing one- and two- factor models, correspondingly. The one-factor LIBOR market model (see Section 2.3) outperforms the proportional volatility model only in pricing tests, which were not strictly out-of-sample. In terms of hedging performance, the two-factor models provides significantly better results than the one-factor models.

Table 3.6: *Algorithm 3.5.2 for the Proportional volatility model.* Performance of Algorithm 3.5.2 with $\Delta = \sqrt{h}$ in the case of the proportional volatility model (3.140) with the same parameters as in Table 3.5.

h	L	*error*	*CPU time, min*
0.2	10^9	$1.50 \times 10^{-4} \pm 2.56 \times 10^{-6}$	5.70×10^1
0.125	10^9	$8.91 \times 10^{-5} \pm 2.56 \times 10^{-6}$	7.88×10^1
0.1	10^9	$8.01 \times 10^{-5} \pm 2.56 \times 10^{-6}$	1.03×10^2
0.05	10^9	$3.47 \times 10^{-5} \pm 2.56 \times 10^{-6}$	2.16×10^2
0.025	10^9	$1.62 \times 10^{-5} \pm 2.54 \times 10^{-6}$	4.78×10^2

Table 3.7: *Algorithm 3,5.3 for the Proportional volatility model.* Performance of Algorithm 3.5.3 with $\Delta = \alpha\sqrt[4]{h}$ in the case of the proportional volatility model (3.140) with the same parameters as in Table 3.5.

h	L	*error*	*CPU time, min*	α
0.2	10^9	$7.04 \times 10^{-5} \pm 2.57 \times 10^{-6}$	5.32×10^1	9.97×10^{-1}
0.125	10^9	$4.59 \times 10^{-5} \pm 2.57 \times 10^{-6}$	6.34×10^1	9.17×10^{-1}
0.1	10^9	$3.66 \times 10^{-5} \pm 2.57 \times 10^{-6}$	7.40×10^1	9.70×10^{-1}
0.05	10^9	$1.74 \times 10^{-5} \pm 2.57 \times 10^{-6}$	1.54×10^2	9.76×10^{-1}

In our experiments we consider two factors, i.e., $d = 2$. We use the same parameters for (3.140) as those found in [33] by calibrating the model to the market prices of caps and floors across different maturities and strike rates:

$$\sigma_1 = 0.1043,\ \sigma_2 = 0.1719,\ \kappa_1 = 0.052,\ \kappa_2 = 0.035. \tag{3.141}$$

As the initial forward curve, we take the one used in numerical examples in [32]:

$$f_0(T) = \log(150 + 48T)/100. \tag{3.142}$$

Since the closed-form formula for caplet price is not available for the HJM model (3.43)-(3.44) with the volatility (3.140), we found the reference caplet price by evaluating the price using Algorithm 3.5.3 with $h = 0.00625$, $\Delta = \alpha\sqrt[4]{h}$ with α from (3.139), and taking the number of Monte Carlo runs $L = 10^9$. This reference value has the Monte Carlo error 2.56×10^{-6}, which gives half of the size of the confidence interval for the corresponding estimator with probability 0.95.

Tables 3.5, 3.6, and 3.7 report the results of our experiments for Algorithm 3.5.1 with $\Delta = h$, Algorithm 3.5.2 with $\Delta = \sqrt{h}$, and Algorithm 3.5.3 with $\Delta = \alpha\sqrt[4]{h}$, α is from (3.139). As in the previous tables, the error column values before "$\pm$" are estimates of the bias computed using the reference price value and the values after "$\pm$" reflect the Monte Carlo error with probability 0.95. As in the Vasicek model example, the Monte Carlo error was made relatively small in order to be able to analyze the bias. We visualise results from the tables in Fugure 3.4. It is clear that the results demonstrate first order of convergence which is in agreement with our theoretical results. The experiments also clearly illustrate the computational superiority of Algorithm 3.5.3 whereas Algorithm 3.5.1 is the slowest out of the three algorithms presented. For instance, Algorithm 3.5.1 with $h = 0.025$ and the error $7.5\times10^{-5}\pm2.56\times10^{-6}$ required 1.63×10^{3} minutes of computer time (see Table 3.5) while Algorithm 3.5.3 with $h = 0.2$ and the similar error $7.04 \times 10^{-5} \pm 2.57 \times 10^{-6}$ required 5.32×10^{1} minutes of computer time (see Table 3.7), i.e., Algorithm 3.5.3 is 30 time faster than Algorithm 3.5.1 in achieving the same level of accuracy.

We note that time measurements were made for comparison purposes only, every effort was made to realize all the algorithms in an analogous way but we did not aim at having the most efficient computer code. Further, to demonstrate convergence of the algorithms, we made the Monte Carlo error much smaller that the numerical integration error. In practice, one usually balances the two errors, e.g. in the case of Algorithm 3.5.3 with $h = 0.2$ it is sufficient to use $L = 10^{7}$ instead of $L = 10^{9}$ in order to get accuracy of order $10^{-4} - 10^{-5}$, this simulation requires just about 30 seconds of computer time.

3.8 Summary

We start this Chapter by revising well-known results from the theory of SDEs. In particular, we consider the "usual" SDE and also the SDE which depend on a pa-

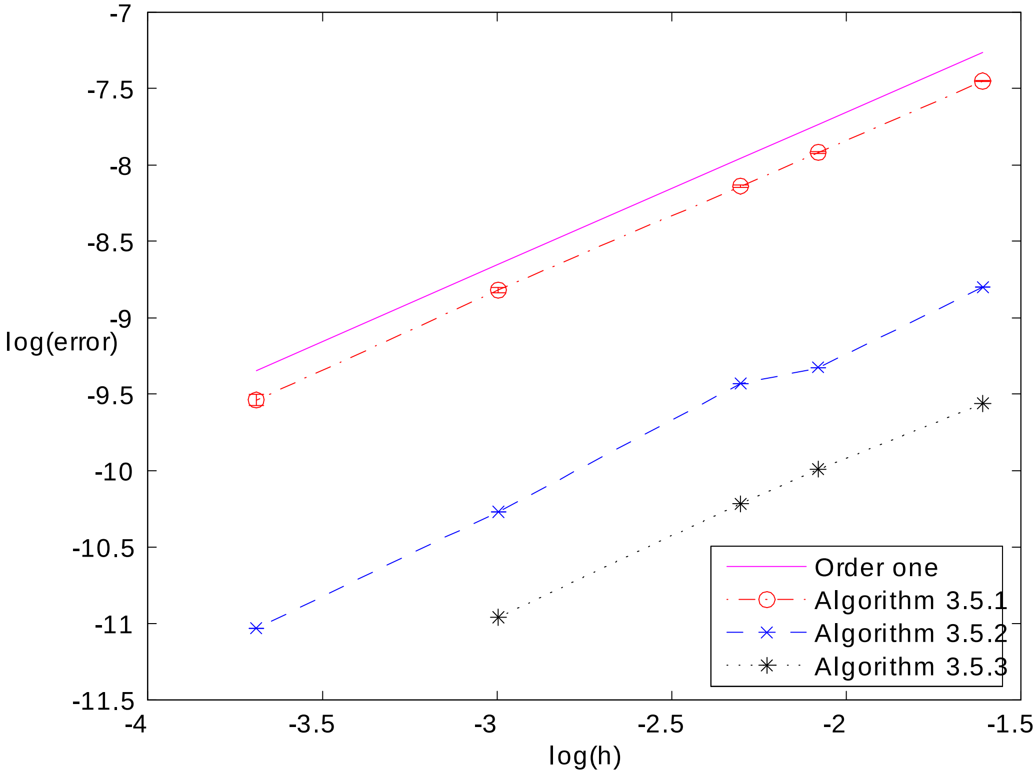

Figure 3.4: *Proportional volatility model:* Stepsize error diagram in log coordinates of the results from Tables 3.5, 3.6, and 3.7. $Log(error)$ is the log of estimates of the bias.

rameter. The results on existence and uniqueness of the solutions of such equations and their differentiability with respect to a parameter, if they dependent on such, are presented in Section 3.1.1. The HJM equation is an example of an infinite-dimensional SDE which initial value and coefficients depend on a parameter.

In Section 3.1.2, we review construction of numerical algorithms based on the Ito-Taylor expansion and present the mean-square and weak criteria for evaluating their order of convergence.

The main contribution of this Chapter is to propose and analyze a new class of efficient numerical methods for the HJM model, inspired by the method of lines (Section 3.4). This provides a rigorous framework for the construction of numerical algorithms that features a high degree of flexibility. The results of this Chapter are based on our paper [48]. In Section 3.4.1, we proceed by first discretising the arbitrage-free drift in the HJM equation in maturity-time variable via the use of high-order quadrature rules. This leads to a finite-dimensional system of coupled stochastic differential equations to which a weak sense numerical integrator is applied for discretisation in calendar time (see Section 3.4.2). Examples of some particular algorithmic realizations of this method are given in Section 3.5. Even though, the main focus of the thesis is weak-sense approximations for the HJM model, we consider mean-square approximations in Section 3.6. They present their own theoretical interest and make our account on HJM numerics complete. The use of high-order quadrature rules allows us to take relatively large steps in the maturity time approximation, preserving the overall accuracy of the algorithms. This is confirmed by our numerical experiments in Section 3.7. We also demonstrate accuracy and convergence properties of the algorithms from Section 3.5 and document the computational costs attached. We conclude that the algorithms with higher-order quadrature rules are more efficient.

Chapter 4

Proof of convergence theorems

In this Chapter we are going to be concerned with establishing a theoretical basis for validity of the results from Chapter 3. More specifically, in Section 4.1 we prove convergence theorems for the methods constructed in Section 3.4. We first establish convergence results for the HJM approximation discrete in the maturity time T only (Section 4.1.1). Then, we analyze weak convergence of fully discrete methods to the approximations discrete in the maturity time (Section 4.1.2). We show that this convergence is uniform in the maturity time discretization step Δ in order to obtain weak convergence of the fully-discrete numerical methods to the solution of the HJM equation. In Section 4.2 we prove convergence for the mean-square method defined in Section 3.6. Finally, in Section 4.3 we demonstrate the routine check of the assumptions imposed on the numerical method in Section 3.4 on the example of Algorithm 3.5.2 defined in Section 3.5.

4.1 Convergence theorems

The aim of this section is to prove the convergence of the approximation $\bar{F}(t_0, f_0; t^*, T^*)$ defined in (3.88) to $F(t_0, f_0(\cdot); t^*, T^*)$ from (3.51) as $h \to 0$ and $\Delta \to 0$. Recall that $F(t_0, f_0; t^*, T^*)$ and $\bar{F}(t_0, f_0(\cdot); t^*, T^*)$ have the form:

$$F(t_0, f_0(\cdot); s_k, s_i) = E \exp(-Y(s_k)) G(P(s_k, s_i)),$$

where

$$Y(s_k) = \int_{t_0}^{s_k} r(u)du,$$

$$P(s_k, s_i) = \exp\left(-Z(s_k, s_i)\right),$$

and

$$Z(s_k, s_i) = \int_{s_k}^{s_i} f(s_k, u)du;$$

while

$$\bar{F}(t_0, f_0; t^*, T^*) = E\exp(-\bar{Y}_M) G\left(\bar{P}(t^*, T^*)\right),$$

where $\bar{Y}_M$ is from (3.84):

$$\bar{Y}_0 = 0, \;\; \bar{Y}_{k+1} = \bar{Y}_k + A^Y(t_k; \bar{f}_k^j, \;\; j = \ell_k, \ldots, \ell_{k+1} + \theta; h), \;\; k = 0, \ldots, M,$$

$$\bar{P}(t^*, T^*) = \exp\left(-\bar{S}_Z(t^*, T^*, \Delta)\right),$$

and

$$\bar{S}_Z(t^*, T^*, \Delta) = \Delta \sum_{j=\varrho_M}^{N} \tilde{\gamma}_j \; \bar{f}_M^j.$$

Denote by $\tilde{F}(t_0, f_0; t^*, T^*)$ the approximation of $F(t_0, f_0(\cdot); t^*, T^*)$ from (3.51) resulting from approximating the solution $f(t, T_i)$ of (3.43)-(3.44) by $\tilde{f}^i(t)$ from (3.67), i.e.,

$$\tilde{F}(t_0, f_0; t^*, T^*) = E\exp(-\tilde{Y}(t^*)) G\left(\tilde{P}(t^*, T^*)\right), \tag{4.1}$$

where

$$\tilde{P}(t^*, T^*) = \exp\left(-\tilde{S}_Z(t^*, T^*, \Delta)\right), \tag{4.2}$$

$\tilde{Y}(t)$ is from (3.81) and $\tilde{S}_Z(t^*, T^*, \Delta)$ is the quadrature rule of the form (3.76) with $f(t^*, T_j)$ replaced by $\tilde{f}^j(t^*)$.

The error R of weak approximation of F by $\bar{F}$ can be written as a sum of two contributing terms:

$$
\begin{aligned}
R &= F(t_0, f_0(\cdot); t^*, T^*) - \bar{F}(t_0, f_0; t^*, T^*) \qquad (4.3)\\
&= \left[F(t_0, f_0(\cdot); t^*, T^*) - \tilde{F}(t_0, f_0; t^*, T^*)\right] + \left[\tilde{F}(t_0, f_0; t^*, T^*) - \bar{F}(t_0, f_0; t^*, T^*)\right]\\
: &= R_1 + R_2,
\end{aligned}
$$

where R_1 is the error due to T-discretization of (3.43)-(3.44) and R_2 is the error due to t-discretization of (3.67)-(3.69). The first error, R_1, is analyzed in Section 4.1.1 and the second error, R_2, is analyzed in Section 4.1.2.

Note that in this section we shall use the letters K, C, and c to denote various constants which are independent of Δ and h.

4.1.1 T-discretization error

In this section we analyze the error of the finite-dimensional approximation (3.67)-(3.69) for the infinite-dimensional stochastic equation (3.43)-(3.44). The plan of this section is as follows. First, we prove (see Theorem 4.1.1) that the approximation (3.67)-(3.69) has mean-square convergence of order Δ^p. This result plays an intermediate role for getting an estimate for the T-discretization error R_1 but, at the same time, it has its own theoretical value. Based on Theorem 4.1.1, we prove (see Lemma 4.1.2) the mean-square convergence of $\tilde{Y}$ from (3.80) to Y from (3.52). Finally, in Theorem 4.1.3 we prove that the weak-sense error R_1 (see (4.3)) of (3.67)-(3.69) is of order Δ^p.

Theorem 4.1.1 *Suppose Assumptions 3.2.1-3.2.3 are satisfied. Then the approximation* $\tilde{f}^i(t)$ *from* (3.67)-(3.69) *converges to* $f(t, T_i)$ *from* (3.43)-(3.44) *as* $\Delta \to 0$ *with the mean-square order* p*, i.e.,*

$$\left(E\left|\tilde{f}^i(t) - f(t, T_i)\right|^2\right)^{1/2} \leq K\Delta^p,\ t \in [t_0, t^* \wedge T_{(i+1)\wedge N}],\ i = 0, \ldots, N, \qquad (4.4)$$

where $K > 0$ *is a constant independent of* Δ*,* t*, and* i*.*

Proof. Denote by $\rho(t,T_i)$ the error of the approximation (3.67)-(3.69):

$$\rho(t,T_i) := \tilde{f}^i(t) - f(t,T_i),\ t \in [t_0, t^* \wedge T_{(i+1)\wedge N}],\ i = 0,\dots,N. \tag{4.5}$$

Clearly,

$$\rho(t_0,T_i) = 0. \tag{4.6}$$

Due to Assumption 3.2.2, $\sigma(s,T,z)$ is globally Lipschitz in z whence

$$|\tilde{\sigma}_i(s) - \sigma(s,T_i)| = \left|\sigma(s,T_i,\tilde{f}^i(s)) - \sigma(s,T_i,f(s,T_i))\right| \leq K\,|\rho(s,T_i)|\,, \tag{4.7}$$

and (cf. (3.66) and (3.69))

$$\begin{aligned} \left|\tilde{S}_I(s,T_i,\Delta) - S_I(s,T_i,\Delta)\right| &= \Delta\left|\sum_{k=\ell(s)}^{\kappa(s,T_i)} \gamma_k(s)\left(\tilde{\sigma}_k(s) - \sigma(s,T_k)\right)\right| \\ &\leq K\Delta\sum_{k=\ell(s)}^{\kappa(s,T_i)} |\tilde{\sigma}_k(s) - \sigma(s,T_k)| \leq K\Delta\sum_{k=\ell(s)}^{\kappa(s,T_i)} |\rho(s,T_k)|\,. \end{aligned} \tag{4.8}$$

We have from (3.67)-(3.69) and (3.43)-(3.44):

$$\begin{aligned} \rho(t,T_i) &= \int_{t_0}^t \left[\tilde{\sigma}_i^\top(s)\tilde{S}_I(s,T_i,\Delta) - \sigma^\top(s,T_i)I(s,T_i)\right] ds \\ &\quad + \int_{t_0}^t \left[\tilde{\sigma}_i(s) - \sigma(s,T_i)\right]^\top dW(s) \\ &= \int_{t_0}^t \tilde{\sigma}_i^\top(s)\left[\tilde{S}_I(s,T_i,\Delta) - S_I(s,T_i,\Delta)\right] ds \\ &\quad + \int_{t_0}^t \tilde{\sigma}_i^\top(s)\left[S_I(s,T_i,\Delta) - I(s,T_i)\right] ds \\ &\quad + \int_{t_0}^t \left[\tilde{\sigma}_i(s) - \sigma(s,T_i)\right]^\top I(s,T_i)ds + \int_{t_0}^t \left[\tilde{\sigma}_i(s) - \sigma(s,T_i)\right]^\top dW(s). \end{aligned}$$

By Ito's formula, we obtain

$$\begin{aligned} \rho^2(t,T_i) &= \int_{t_0}^t 2\rho(s,T_i)\tilde{\sigma}_i^\top(s)\left[\tilde{S}_I(s,T_i,\Delta) - S_I(s,T_i,\Delta)\right] ds \\ &\quad + \int_{t_0}^t 2\rho(s,T_i)\tilde{\sigma}_i^\top(s)\left[S_I(s,T_i,\Delta) - I(s,T_i)\right] ds \end{aligned} \tag{4.9}$$

$$
\begin{aligned}
&+\int_{t_0}^{t} 2\rho(s,T_i)\left[\tilde{\sigma}_i(s)-\sigma(s,T_i)\right]^\top I(s,T_i)ds \\
&+\int_{t_0}^{t} \left[\tilde{\sigma}_i(s)-\sigma(s,T_i)\right]^\top \left[\tilde{\sigma}_i(s)-\sigma(s,T_i)\right] ds \\
&+\int_{t_0}^{t} 2\rho(s,T_i)\left[\tilde{\sigma}_i(s)-\sigma(s,T_i)\right]^\top dW(s).
\end{aligned}
$$

Then

$$
\begin{aligned}
E\rho^2(t,T_i) &= 2\int_{t_0}^{t} E\rho(s,T_i)\tilde{\sigma}_i^\top(s)\left[\tilde{S}_I(s,T_i,\Delta)-S_I(s,T_i,\Delta)\right] ds \qquad (4.10)\\
&\quad +2\int_{t_0}^{t} E\rho(s,T_i)\tilde{\sigma}_i^\top(s)\left[S_I(s,T_i,\Delta)-I(s,T_i)\right] ds \\
&\quad +2\int_{t_0}^{t} E\rho(s,T_i)\left[\tilde{\sigma}_i(s)-\sigma(s,T_i)\right]^\top I(s,T_i)ds \\
&\quad +\int_{t_0}^{t} E\left[\tilde{\sigma}_i(s)-\sigma(s,T_i)\right]^\top \left[\tilde{\sigma}_i(s)-\sigma(s,T_i)\right] ds.
\end{aligned}
$$

Using the boundedness of $\sigma(s,T,z)$ (see (3.45)) and the inequality (4.8), the first term on the right-hand side of (4.10) is estimated as

$$
\begin{aligned}
&\left|2\int_{t_0}^{t} E\rho(s,T_i)\tilde{\sigma}_i^\top(s)\left[\tilde{S}_I(s,T_i,\Delta)-S_I(s,T_i,\Delta)\right] ds\right| \qquad (4.11)\\
\leq\ & K\Delta\left[\int_{t_0}^{t} E\left|\rho(s,T_i)\right| \sum_{k=\ell(s)}^{\kappa(s,T_i)} \left|\rho(s,T_k)\right| ds\right].
\end{aligned}
$$

Using the boundedness of $\sigma(s,T,z)$, the inequality $2ab \leq a^2+b^2$, and the condition (3.63) for the quadrature rule S_I, we obtain for the second term on right hand side of (4.10):

$$
\begin{aligned}
&\left|2\int_{t_0}^{t} E\rho(s,T_i)\tilde{\sigma}_i^\top(s)\left[S_I(s,T_i,\Delta)-I(s,T_i)\right] ds\right| \qquad (4.12)\\
\leq\ & K\int_{t_0}^{t} E|\rho(s,T_i)|\left|S_I(s,T_i,\Delta)-I(s,T_i)\right| ds \\
\leq\ & K\int_{t_0}^{t} \left[E\rho^2(s,T_i)+E\left|S_I(s,T_i,\Delta)-I(s,T_i)\right|^2\right] ds \\
\leq\ & K\int_{t_0}^{t} \left[E\rho^2(s,T_i)+\Delta^{2p}\right] ds.
\end{aligned}
$$

Using the inequality (4.7) and the boundedness of $\sigma(s,T,z)$, we get for the third term on the right-hand side of (4.10):

$$\left|2\int_{t_0}^{t} E\rho(s,T_i)\left[\tilde{\sigma}_i(s)-\sigma(s,T_i)\right]^{\top} I(s,T_i)ds\right| \le K\int_{t_0}^{t} E\rho^2(s,T_i)ds. \qquad (4.13)$$

By the inequality (4.7), the fourth term on the right-hand of (4.10) is estimated as

$$\int_{t_0}^{t} E\left[\tilde{\sigma}_i(s)-\sigma(s,T_i)\right]^{\top}\left[\tilde{\sigma}_i(s)-\sigma(t,T_i)\right]ds \le K\int_{t_0}^{t} E\rho^2(s,T_i)ds. \qquad (4.14)$$

Substituting (4.11)-(4.14) in (4.10) and using the inequality $2ab \le a^2+b^2$, we obtain

$$\begin{aligned}
E\rho^2(t,T_i) \;&\le\; K\int_{t_0}^{t}\left\{E\rho^2(s,T_i)+\Delta E\left[|\rho(s,T_i)|\sum_{k=\ell(s)}^{\kappa(s,T_i)}|\rho(s,T_k)|\right]+\Delta^{2p}\right\}ds\\
&\le\; K\int_{t_0}^{t}\Bigg\{\;(1+\Delta(\kappa(s,T_i)-\ell(s)+2))\,E\rho^2(s,T_i)\\
&\qquad +\Delta\sum_{k=\ell(s),\;k\neq i}^{\kappa(s,T_i)}E\rho^2(s,T_k)+\Delta^{2p}\Bigg\}ds\\
&\le\; K\int_{t_0}^{t}\left\{E\rho^2(s,T_i)+\Delta\sum_{k=\ell(s),\;k\neq i}^{\kappa(s,T_i)}E\rho^2(s,T_k)\right\}ds+K\Delta^{2p},\\
t \;&\in\; [t_0,t^*\wedge T_{(i+1)\wedge N}],\quad i=0,\ldots,N. \qquad (4.15)
\end{aligned}$$

We have used here that $\Delta(\kappa(s,T_i)-\ell(s)+2) \le T^*-t_0$.

Introduce $\rho_M(t) := \max_{\ell(t)\le i\le N} E\rho^2(t,T_i)$, $t\in[t_0,t^*]$. Clearly (see (4.6)), $\rho_M(t_0) = 0$. Then we get from (4.15):

$$\rho_M(t) \le K\int_{t_0}^{t}\rho_M(s)ds + K\Delta^{2p},$$

whence (4.4) follows by the Gronwall inequality. Theorem 4.1.1 is proved. □

Using Theorem 4.1.1, we prove the following lemma.

Lemma 4.1.2 *Suppose Assumptions 3.2.1-3.2.3 are satisfied. The approximation*

$\tilde{Y}(t)$ *from* (3.81) *converges to* $Y(t)$ *from* (3.52) *as* $\Delta \to 0$ *with the mean-square order* $p > 0$*, i.e.,*

$$\left(E\left[Y(t^*) - \tilde{Y}(t^*) \right]^2 \right)^{1/2} \leq K\Delta^p, \tag{4.16}$$

where $K > 0$ *is a constant independent of* Δ*.*

Proof. Consider the error of the approximation (3.81) for (3.52) (see also (3.80)):

$$Y(t^*) - \tilde{Y}(t^*) = \int_{t_0}^{t^*} f(s,s)ds - \int_{t_0}^{t^*} \tilde{\pi}(s)ds. \tag{4.17}$$

We rearrange the right-hand side of (4.17) to split this error into the error due to approximation of the short rate $r(t) = f(t,t)$ by $\pi(t)$ and the error due to approximation of $f(t, T_i)$ by $\tilde{f}^i(t)$:

$$Y(t^*) - \tilde{Y}(t^*) = \int_{t_0}^{t^*} (f(s,s) - \pi(s))\, ds + \int_{t_0}^{t^*} (\pi(s) - \tilde{\pi}(s))\, ds. \tag{4.18}$$

Due to the condition (3.79) imposed on our choice of the approximation $\pi(s)$, we have

$$E\left(\int_{t_0}^{t^*} (f(s,s) - \pi(s))\, ds \right)^2 \leq K\Delta^{2p}. \tag{4.19}$$

Recalling the form of the approximation $\pi(s)$ from (3.78), we get

$$\begin{aligned} & E\left(\int_{t_0}^{t^*} (\pi(s) - \tilde{\pi}(s))\, ds \right)^2 \\ = \quad & E\left[\int_{t_0}^{t^*} \sum_{l=0}^{\ell(t^*)} \sum_{i=0}^{\theta} \lambda_i(s)(f(s, T_{l+i}) - \tilde{f}^{l+i}(s))\chi_{s\in[T_l, T_{l+1})} ds \right]^2, \end{aligned}$$

where $\lambda_i(s)$ are bounded coefficients and the number θ is independent of Δ. Then, using (4.4), we obtain

$$E\left(\int_{t_0}^{t^*} (\pi(s) - \tilde{\pi}(s))\, ds \right)^2 \leq K\Delta^{2p}. \tag{4.20}$$

The relations (4.18)-(4.20) imply the required error estimate (4.16). $\square$

In the next theorem we obtain an estimate for the weak sense error R_1 from (4.3).

Theorem 4.1.3 *Suppose Assumptions 3.2.1-3.2.3 are satisfied. Assume that the payoff function $G(z)$ satisfies the global Lipschitz condition* (3.50). *Then the approximation $\tilde{F}(t_0, f_0;\, t^*, T^*)$ from* (4.1) *converges to $F(t_0, f_0\,(\cdot)\,; t^*, T^*)$ from* (3.51), (3.52)-(3.54) *with order $p > 0$, i.e.,*

$$\left|F(t_0, f_0\,(\cdot)\,; t^*, T^*) - \tilde{F}(t_0, f_0; t^*, T^*)\right| \leq K\Delta^p, \tag{4.21}$$

where $K > 0$ is a constant independent of Δ.

Proof. We have (see (3.51), (3.52)-(3.54) and (4.1)-(4.2)):

$$\begin{aligned} & R_1 = F(t_0, f_0\,(\cdot)\,; t^*, T^*) - \tilde{F}(t_0, f_0; t^*, T^*) \\ = \;& E\exp(-Y(t^*))G\left(P(t^*,T^*)\right) - E\exp(-\tilde{Y}(t^*))G\left(\tilde{P}(t^*,T^*)\right) \\ = \;& E\left[\exp\left(-Y(t^*)\right) - \exp\left(-\tilde{Y}(t^*)\right)\right] G\left(\tilde{P}(t^*,T^*)\right) \\ & + E\left[G\left(P(t^*,T^*)\right) - G\left(\tilde{P}(t^*,T^*)\right)\right] \exp(-Y(t^*)). \end{aligned} \tag{4.22}$$

Consider the first term on the right-hand side of (4.22). By the mean value theorem, we get

$$\exp\left(-Y(t^*)\right) - \exp\left(-\tilde{Y}(t^*)\right) = (\tilde{Y}(t^*) - Y(t^*))\exp(\vartheta), \tag{4.23}$$

where ϑ is a point between $-\tilde{Y}(t^*)$ and $-Y(t^*)$.

Due to the global Lipschitz condition (3.50) imposed on $G(z)$, we have (recall that $\tilde{S}_Z(t^*, T^*, \Delta)$ is the quadrature rule of the form (3.76) with $f(t^*, T_i)$ replaced by $\tilde{f}^i(t^*)$:

$$\begin{aligned} |G(\tilde{P}(t^*,T^*))| & \leq K\tilde{P}(t^*,T^*) = K\exp\left(-\tilde{S}_Z(t^*,T^*,\Delta)\right) \\ & = K\exp\left(-\Delta\sum_{j=\varrho_M}^{N}\tilde{\gamma}_j\;\tilde{f}^j(t^*)\right). \end{aligned} \tag{4.24}$$

Using (4.23), (4.24), and the Cauchy–Bunyakovsky inequality twice, we obtain

$$\begin{aligned}
&\left|E\left[\exp\left(-Y(t^*)\right)-\exp(-\tilde{Y}(t^*))\right]G(\tilde{P}(t^*,T^*))\right| \qquad (4.25)\\
\le\ & K\left|\left[E(\tilde{Y}(t^*)-Y(t^*))^2\right]^{1/2}\left[E\exp(4\vartheta)\right]^{1/4}\left[E\exp\left(-4\Delta\sum_{j=\varrho_M}^{N}\tilde{\gamma}_j\ \tilde{f}^j(t^*)\right)\right]^{1/4}\right|.
\end{aligned}$$

Thanks to (3.47) and (3.70), exponential moments of $-\tilde{Y}(t^*)$ and $-Y(t^*)$ are bounded and, consequently, for some $K>0$ we get $E\exp(4\vartheta)<K$. Due to (3.70), we also have

$$\begin{aligned}
E\exp\left(-4\Delta\sum_{j=\varrho_M}^{N}\tilde{\gamma}_j\ \tilde{f}^j(t^*)\right) &= E\left[\exp\left(-4T\sum_{j=\varrho_M}^{N}\tilde{\gamma}_j\ \tilde{f}^j(t^*)\right)\right]^{1/N}\\
&\le \frac{1}{N}\sum_{j=\varrho_M}^{N}E\exp\left(-4T\tilde{\gamma}_j\ \tilde{f}^j(t^*)\right)<K.
\end{aligned}$$

Then (4.25) together with (4.16) implies

$$|E\left[\exp\left(-Y(t^*)\right)-\exp(-\tilde{Y}(t^*))\right]G(\tilde{P}(t^*,T^*))|\le K\Delta^p. \qquad (4.26)$$

Let us now consider the second term on the right-hand side of (4.22). Due to the global Lipschitz condition (3.50) imposed on $G(z)$, we have

$$\left|G\left(P(t^*,T^*)\right)-G(\tilde{P}(t^*,T^*))\right|\le K\cdot\left|P(t^*,T^*)-\tilde{P}(t^*,T^*)\right|. \qquad (4.27)$$

Further, by the mean value theorem, we get

$$\begin{aligned}
P(t^*,T^*)-\tilde{P}(t^*,T^*) &= \exp\left(-Z\left(t^*,T^*\right)\right)-\exp\left(-\tilde{S}_Z(t^*,T^*,\Delta)\right)\\
&= \left(\tilde{S}_Z(t^*,T^*,\Delta)-Z\left(t^*,T^*\right)\right)\exp\left(\vartheta\right), \qquad (4.28)
\end{aligned}$$

where ϑ is between $-\tilde{S}_Z(t^*,T^*,\Delta)$ and $-Z\left(t^*,T^*\right)$.

Using (4.27), (4.28), and the Cauchy–Bunyakovsky inequality twice, we obtain

$$\begin{aligned} & \left| E\left[G\left(P(t^*,T^*)\right) - G(\tilde{P}(t^*,T^*))\right] \exp(-Y(t^*)) \right| \qquad (4.29) \\ \leq & \left[E\left(\tilde{S}_Z(t^*,T^*,\Delta) - Z\left(t^*,T^*\right)\right)^2 \right]^{1/2} \left[E\exp(-4Y(t^*))\right]^{1/4} \left[E\exp(4\vartheta)\right]^{1/4}. \end{aligned}$$

It is clear that (3.47) and (3.70) imply boundedness of the exponential moments present in the right-hand side of (4.29) and, hence,

$$\begin{aligned} & \left| E\left[G\left(P(t^*,T^*)\right) - G\left(\tilde{P}(t^*,T^*)\right)\right] \exp(-Y(t^*)) \right| \qquad (4.30) \\ \leq & \; K\left[E\left(\tilde{S}_Z(t^*,T^*,\Delta) - Z\left(t^*,T^*\right)\right)^2 \right]^{1/2}. \end{aligned}$$

We have

$$\begin{aligned} & E\left(\tilde{S}_Z(t^*,T^*,\Delta) - Z\left(t^*,T^*\right)\right)^2 \qquad (4.31) \\ = & \; E\left(\tilde{S}_Z(t^*,T^*,\Delta) - S_Z(t^*,T^*,\Delta) + S_Z(t^*,T^*,\Delta) - Z\left(t^*,T^*\right)\right)^2 \\ \leq & \; 2E\left[\tilde{S}_Z(t^*,T^*,\Delta) - S_Z(t^*,T^*,\Delta)\right]^2 + 2E\left[\tilde{S}_Z(t^*,T^*,\Delta) - Z\left(t^*,T^*\right)\right]^2. \end{aligned}$$

Due to the condition (3.77) imposed on the quadrature rule $S_Z(t^*,T^*,\Delta)$, the second term on the right-hand side of (4.31) is bounded from above by $K\Delta^{2p}$. Using (4.4), we obtain for the first term on the right-hand side of (4.31) (cf. (3.76)):

$$\begin{aligned} 2E\left[\tilde{S}_Z(t^*,T^*,\Delta) - S_Z(t^*,T^*,\Delta)\right]^2 &= 2\Delta^2 E\left[\sum_{j=\varrho_M}^{N} \tilde{\gamma}_j \left(\tilde{f}^j(t^*) - f(t^*,T_j)\right)\right]^2 \\ &\leq K\Delta \sum_{j=\varrho_M}^{N} E\left[\tilde{f}^j(t^*) - f(t^*,T_j)\right]^2, \\ &\leq K\Delta(N-\varrho_M+1)\Delta^{2p} \leq K\Delta^{2p}. \end{aligned}$$

Hence

$$E\left(\tilde{S}_Z(t^*,T^*,\Delta) - Z\left(t^*,T^*\right)\right)^2 \leq K\Delta^{2p}. \qquad (4.32)$$

The required estimate (4.21) follows from (4.22), (4.26), (4.30), and (4.32). Theorem 4.1.3 is proved. $\square$

4.1.2 t-discretization error

In this section we analyze the error R_2 (see (4.3)) due to t-discretization of (3.67)-(3.69):

$$R_2 = \tilde{F}(t_0, f_0; t^*, T^*) - \bar{F}(t_0, f_0; t^*, T^*).$$

Then combining its estimate with the estimate (4.21) for R_1 from Theorem 4.1.3, we prove convergence of the weak approximation $\bar{F}$ to F (see (4.3)). In the analysis of R_2 the key is to show that convergence of $\bar{F}(t_0, f_0; t^*, T^*)$ to $\tilde{F}(t_0, f_0; t^*, T^*)$ is uniform in Δ, which is the reason why we cannot just apply here the standard results of weak convergence of numerical methods for SDEs (see, e.g. [45, 57, 58]). The convergence theorem is proved under the assumption that the pay-off function $G(z)$ in (3.51) is sufficiently smooth. At the end of this section we also discuss how this assumption can be relaxed.

To prove the convergence theorem (Theorem 4.1.6) of $\bar{F}(t_0, f_0; t^*, T^*)$ to $\tilde{F}(t_0, f_0; t^*, T^*)$, we need the following technical lemma. We will use the multi-index notation:

$$\mathbf{i} = (i_0, \ldots, i_N)$$

with i_j being nonnegative integers, $|\mathbf{i}| = i_0 + \cdots + i_N$, and $\mathbf{i}! = i_0! \cdots i_N!$.

Lemma 4.1.4 *Let Λ^m be the m^{th}-order operator*

$$\Lambda^m = \Lambda^m_\mu = \sum_{|\mathbf{i}|=m} \mu^{\mathbf{i}} \frac{\partial^m}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}}$$

with any $\mu^{\mathbf{i}}$. Suppose Assumptions 3.2.1 and 2.2 are satisfied. Assume that the payoff function $G(z)$ has m_* bounded derivatives. Then for $m > 0$ up to the order m_*

$$\left|\Lambda^m \tilde{F}(t, x; t^*, T^*)\right| \leq K \mu_{Max} \exp(c\Delta |x|), \tag{4.33}$$

where $K > 0$ and $c > 0$ do not depend on Δ and $x \in \mathbb{R}^{N+1}$, and $\mu_{Max} := \max_{|\mathbf{i}|=m} |\mu^{\mathbf{i}}|$.

Remark 4.1.5 *To help with intuitive understanding of this lemma, we remark that $\Lambda^m \tilde{F}$ can be viewed as a Frechet derivative of the option price with respect to the discretized initial forward rate curve.*

Proof of Lemma 4.1.4. Recall the notation: $\tilde{f}^j_{t,x}(s),\ s \geq t$, is the solution of the system of SDEs (3.67)-(3.69) with the initial condition at $t \geq t_0 : \tilde{f}^j_{t,x}(t) = x^j$. We introduce a more detailed notation for $\tilde{S}_Z(t^*, T^*, \Delta)$ (cf. (3.76)):

$$\tilde{S}_Z(t, x; t^*, T^*, \Delta) = \Delta \sum_{j=\varrho_M}^{N} \tilde{\gamma}_j \tilde{f}^j_{t,x}(t^*), \tag{4.34}$$

which we can present as (cf. (3.67))

$$\tilde{S}_Z(t, x; t^*, T^*, \Delta) = \Delta \sum_{j=\varrho_M}^{N} \tilde{\gamma}_j x^j$$
$$+\Delta \sum_{j=\varrho_M}^{N} \tilde{\gamma}_j \left[\int_t^{t^*} \sigma^\top(s, T_j, \tilde{f}^j_{t,x}(s)) \tilde{S}_I(t, x; s, T_j, \Delta) ds + \int_t^{t^*} \sigma^\top(s, T_j, \tilde{f}^j_{t,x}(s)) dW(s) \right],$$

where (cf. (3.69))

$$\tilde{S}_I(t, x; s, T_j, \Delta) = \Delta \sum_{l=\ell(s)}^{\kappa(s,T_j)} \gamma_l(s) \sigma(s, T_l, \tilde{f}^l_{t,x}(s)). \tag{4.35}$$

Then, thanks to Assumption 3.2.1, we obtain for any positive integer m :

$$E\tilde{P}^m(t^*, T^*) = E \exp\left(-m\tilde{S}_Z(t, x; t^*, T^*, \Delta)\right) \tag{4.36}$$
$$= E \exp\left(-m\Delta \sum_{j=\varrho_M}^{N} \tilde{\gamma}_j x^j \right.$$
$$\left. -m\Delta \sum_{j=\varrho_M}^{N} \tilde{\gamma}_j \left[\int_t^{t^*} \sigma^\top(s, T_j, \tilde{f}^j_{t,x}(s)) \tilde{S}_I(t, x; s, T_j, \Delta) ds + \int_t^{t^*} \sigma^\top(s, T_j, \tilde{f}^j_{t,x}(s)) dW(s) \right] \right)$$
$$\leq K \exp\left(c\Delta \sum_{j=\varrho_M}^{N} |x^j| \right),$$

where $K > 0$ and $c > 0$ do not depend on Δ.

Further, recall that $\tilde{Y}_{t,x,y}(s),\ s \geq t$, is the solution of (3.81) with the initial condition $\tilde{Y}_{t,x,y}(t) = y$ and with $\tilde{f}^i(s) = \tilde{f}^i_{t,x}(s)$, i.e.,

$$\begin{aligned} \tilde{Y}_{t,x,y}(s) \quad &: \; = y + \int_t^s \tilde{\pi}(s')ds' && (4.37) \\ &= \; y + \int_t^s \pi(s'; \tilde{f}^i_{t,x}(s'),\ i = \ell(s'), \ldots, \ell(s') + \theta)ds' \\ &= \; y + \sum_{l=\ell(t)}^{\ell(s)} \sum_{i=0}^{\theta} \int_{t \vee T_l}^{s \wedge T_{l+1}} \lambda_i(s') \tilde{f}^{l+i}_{t,x}(s')\ ds' \\ &= \; y + \sum_{l=\ell(t)}^{\ell(s)} \sum_{m=l}^{l+\theta} \int_{t \vee T_l}^{s \wedge T_{l+1}} \lambda_{m-l}(s') \tilde{f}^m_{t,x}(s')ds', \quad t \geq t_0, \quad s \geq t, \end{aligned}$$

where (cf. (3.78)) θ and $\lambda_{m-l}(s')$ depend on our choice of the accuracy order of short rate approximation, and θ does not depend on Δ, and $|\lambda_{m-l}(s')|$ are bounded by a constant independent of Δ.

We also see that $\tilde{Y}_{t,x,y}(s) = y + \tilde{Y}_{t,x,0}(s)$. Using (4.37), (3.67), and Assumption 3.2.1, one can show that for any $\varkappa > 0$

$$\begin{aligned} &E\left[\exp(\varkappa|\tilde{Y}_{t,x,0}(t^*)|)\right] && (4.38) \\ = \;& E\exp\left(\varkappa\left|\sum_{l=\ell(t)}^{\ell(t^*)} \sum_{m=l}^{l+\theta} \int_{t \vee T_l}^{t^* \wedge T_{l+1}} \lambda_{m-l}(s') \tilde{f}^m_{t,x}(s')ds'\right|\right) \leq K\exp\left(c\Delta \sum_{l=\ell(t)}^{\ell(t^*)+\theta} |x^l|\right), \end{aligned}$$

where $K > 0$ and $c > 0$ do not depend on Δ.

Using smoothness of $G(z)$, we obtain

$$\begin{aligned} \Lambda^m \tilde{F}(t, x; t^*, T^*) &= E\Lambda^m \exp(-\tilde{Y}_{t,x,0}(t^*))G(\tilde{P}(t^*, T^*)) && (4.39) \\ = E \sum_{|\mathbf{i}|=m} \mu^{\mathbf{i}} &\frac{\partial^m}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}} \exp(-\tilde{Y}_{t,x,0}(t^*))G(\tilde{P}(t^*, T^*)) \\ = E\exp(-\tilde{Y}_{t,x,0}(t^*)) &\sum_{k_*=0}^{m} \sum_{n_*=0}^{m-k_*} \sum_{\alpha=0}^{m} \sum_{\beta=0}^{\alpha} \sum_{\bar{j}_{k_*} + \bar{l}_{n_*} = m} C(\alpha, \beta, j_1, \ldots, j_{k_*}, l_1, \ldots, l_{n_*}) \\ &\times \frac{d^\alpha}{dz^\alpha} G\left(\tilde{P}(t^*, T^*)\right) \tilde{P}^\beta(t^*, T^*) \end{aligned}$$

$$\times \sum_{i_1,\dots,i_{\bar{j}_{k_*}},r_1,\dots,r_{\bar{l}_{n_*}}=0}^{N} \mu^{\mathbf{i}} \prod_{k=1}^{k_*} \frac{\partial^{j_k}}{\partial x^{i_{1+\bar{j}_{k-1}}} \cdots \partial x^{i_{\bar{j}_k}}} \tilde{Y}_{t,x,0}(t^*)$$
$$\times \prod_{n=1}^{n_*} \frac{\partial^{l_n}}{\partial x^{r_{1+\bar{l}_{n-1}}} \cdots \partial x^{r_{\bar{l}_n}}} \tilde{S}_Z(t,x;t^*,T^*,\Delta),$$

where $C(\alpha,\beta,j_1,\dots,j_{k_*},l_1,\dots,l_{n_*})$ are constants independent of N; $\bar{j}_k = \sum_{r=1}^{k} j_r$, $\bar{l}_n = \sum_{r=1}^{n} l_r$; the sum $\sum_{\bar{j}_{k_*}+\bar{l}_{n_*}=m}$ is taken over all positive integers $j_1,\dots,j_{k_*}$ and $l_1,\dots,l_{n_*}$ such that $j_k \le j_{k+1}$, $k=1,\dots,k_*-1$, $l_n \le l_{n+1}$, $n=1,\dots,n_*-1$, and $\bar{j}_{k_*}+\bar{l}_{n_*}=m$; and in the right-hand side the multi-index $\mathbf{i}$ at $\mu^{\mathbf{i}}$ corresponds to the values taken by $i_1,\dots,i_{\bar{j}_{k_*}}$, $r_1,\dots,r_{\bar{l}_{n_*}}$.

We have (cf. (4.34)):

$$\frac{\partial^l}{\partial x^{i_1}\cdots\partial x^{i_l}} \tilde{S}_Z(t,x;t^*,T^*,\Delta) = \Delta \sum_{q=\varrho_M}^{N} \tilde{\gamma}_q \frac{\partial^l}{\partial x^{i_1}\cdots\partial x^{i_l}} \tilde{f}^q_{t,x}(t^*). \tag{4.40}$$

Using the Cauchy-Bunyakovsky inequality, the assumed boundedness of derivatives of $G(z)$, and the inequalities (4.36) and (4.38), we obtain from (4.39)-(4.40):

$$|\Lambda^m \tilde{F}(t,x;t^*,T^*)| \le K \exp\left(c\Delta \sum_{j=\varrho_M}^{N} |x^j| \right) \tag{4.41}$$
$$\times \sum_{k_*=0}^{m} \sum_{n_*=0}^{m-k_*} \sum_{\bar{j}_{k_*}+\bar{l}_{n_*}=m} \left(E\left[\sum_{i_1,\dots,i_{\bar{j}_{k_*}},r_1,\dots,r_{\bar{l}_{n^*}}=0}^{N} \mu^{\mathbf{i}} \prod_{k=1}^{k_*} \frac{\partial^{j_k}}{\partial x^{i_{1+\bar{j}_{k-1}}}\cdots\partial x^{i_{\bar{j}_k}}} \tilde{Y}_{t,x,0}(t^*) \right.\right.$$
$$\left.\left. \times \prod_{n=1}^{n_*} \left(\Delta \sum_{q=\varrho_M}^{N} \tilde{\gamma}_q \frac{\partial^{l_n}}{\partial x^{r_{1+\bar{l}_{n-1}}}\cdots\partial x^{r_{\bar{l}_n}}} \tilde{f}^q_{t,x}(t^*) \right) \right]^2 \right)^{1/2},$$

where $K>0$ and c are independent of Δ and x. Then, to complete the proof of this lemma, it is sufficient to show that for any $0 \le k_* \le m$ and $0 \le n_* \le m-k_*$, any combinations of $j_1,\dots,j_{k^*}$ and $l_1,\dots,l_{n^*}$ satisfying $\bar{j}_{k_*}+\bar{l}_{n^*}=m$, and any combination of $q_1, \ldots, q_{n^*}$ with $\varrho_M \leq q_i \leq N$:

$$E\left[\sum_{i_1,\ldots,i_{\bar{j}_{k_*}},r_1,\ldots,r_{\bar{l}_{n^*}}=0}^{N} \mu^{\mathbf{i}} \prod_{k=1}^{k_*} \frac{\partial^{j_k}}{\partial x^{i_{1+\bar{j}_{k-1}}} \cdots \partial x^{i_{\bar{j}_k}}} \tilde{Y}_{t,x,0}(t^*) \right. \tag{4.42}$$
$$\left. \times \prod_{n=1}^{n_*} \frac{\partial^{l_n}}{\partial x^{r_{1+\bar{l}_{n-1}}} \cdots \partial x^{r_{\bar{l}_n}}} \tilde{f}^{q_n}_{t,x}(t^*)\right]^2 \leq K\mu^2_{Max},$$

where $K > 0$ is independent of Δ and x.

We can obtain the following SDEs (see (3.81)):

$$\begin{aligned} d\frac{\partial^j}{\partial x^{i_1}\cdots\partial x^{i_j}}\tilde{Y}_{t,x,0}(s) &= \sum_{l=\ell(t)}^{\ell(t^*)}\sum_{r=l}^{l+\theta}\chi_{s\in[T_l,T_{l+1})}\cdot\lambda_{r-l}(s)\cdot\frac{\partial^j}{\partial x^{i_1}\cdots\partial x^{i_j}}\tilde{f}^r_{t,x}(s)ds, \\ \frac{\partial^j}{\partial x^{i_1}\cdots\partial x^{i_j}}\tilde{Y}_{t,x,0}(t) &= 0, \end{aligned}$$

and (see (3.67))

$$\begin{aligned} d\frac{\partial^l}{\partial x^{r_1}\cdots\partial x^{r_l}}\tilde{f}^q_{t,x}(s) = \sum_{\alpha=0}^{l}\sum_{\beta=0}^{l}\sum_{n_*=0}^{l-1}\sum_{\tau_*=1}^{l-n_*}\sum_{\bar{l}_{n_*}+\bar{p}_{\tau_*}=l} C(\alpha,\beta,n_*,\tau_*) \\ \times \sum_{\{k_1,\ldots,k_l\}=\{r_1,\ldots,r_l\}} \Delta\sum_{v=\ell(s)}^{\kappa(s,T_q)}\gamma_v\ \frac{d^\alpha}{dz^\alpha}\sigma^\top(s,T_q,\tilde{f}^q_{t,x}(s))\frac{d^\beta}{dz^\beta}\sigma^\top(s,T_v,\tilde{f}^v_{t,x}(s)) \\ \times\prod_{n=1}^{n_*}\frac{\partial^{l_n}}{\partial x^{k_{1+\bar{l}_{n-1}}}\cdots\partial x^{k_{\bar{l}_n}}}\tilde{f}^q_{t,x}(s)\ \prod_{\tau=1}^{\tau_*}\frac{\partial^{p_\tau}}{\partial x^{k_{1+\bar{l}_{n_*}+\bar{p}_{\tau-1}}}\cdots\partial x^{k_{\bar{l}_{n_*}+\bar{p}_\tau}}}\tilde{f}^v_{t,x}(s)\ ds \\ +\sum_{\alpha=1}^{l}\sum_{n_*=1}^{l}\sum_{\bar{l}_{n_*}=l}C(\alpha,n_*)\sum_{\{k_1,\ldots,k_l\}=\{r_1,\ldots,r_l\}}\frac{d^\alpha}{dz^\alpha}\sigma^\top(s,T_q,\tilde{f}^q_{t,x}(s)) \\ \times\prod_{n=1}^{n_*}\frac{\partial^{l_n}}{\partial x^{k_{1+\bar{l}_{n-1}}}\cdots\partial x^{k_{\bar{l}_n}}}\tilde{f}^q_{t,x}(s)\ dW(s), \\ \frac{\partial^l}{\partial x^{r_1}\cdots\partial x^{r_l}}\tilde{f}^q_{t,x}(t) = \chi_{l=1}, \end{aligned}$$

where $C(\alpha,\beta,n_*,\tau_*)$ and $C(\alpha,n_*)$ are constants independent of N, and $\sum_{\{k_1,\ldots,k_l\}=\{r_1,\ldots,r_l\}}$ means summation over all possible recombinations $\{k_1,\ldots,k_l\}$ of $r_1,\ldots,r_l$ (note that the number of terms in this sum depends on l but not on N).

To obtain (4.42), we first consider the case $m = 1$ for which it is sufficient to get an estimate for $E\left[\sum_{i=0}^{N}\mu^i\frac{\partial}{\partial x^i}\tilde{f}^j_{t,x}(s)\right]^2$. To this end, introduce the process

$\zeta_{t,x}(s) = (\zeta^0(s), \ldots, \zeta^N(s))^\top$ with $\zeta^j(s) := \sum_{i=0}^N \mu^i \frac{\partial}{\partial x^i} \tilde{f}^j_{t,x}(s),\ s \geq t$, which satisfies the following system of SDEs

$$d\zeta^j = \frac{d}{dz}\sigma^\top(s, T_j, \tilde{f}^j_{t,x}(s)) \cdot \tilde{S}_I(t, x; s, T_j, \Delta) \cdot \zeta^j\ ds$$
$$+\sigma^\top(s, T_j, \tilde{f}^j_{t,x}(s)) \cdot \Delta \sum_{l=\ell(s)}^{\kappa(s,T_j)} \gamma_l(s) \frac{d}{dz}\sigma^\top(s, T_l, \tilde{f}^l_{t,x}(s)) \cdot \zeta^l\ ds + \frac{d}{dz}\sigma^\top(s, T_i, \tilde{f}^j_{t,x}(s)) \cdot \zeta^j\ dW(s),$$
$$\zeta^j(t) = \mu^j, \quad j = 0, \ldots, N.$$

Then using Ito's formula and Assumptions 3.2.1 and 3.2.2, we obtain after some straightforward calculations:

$$E\left[\zeta^j_{t,x}(s)\right]^2 \leq K\left[\mu^j\right]^2 + K\int_t^s E\left[\zeta^j(s')\right]^2 ds' + K\int_t^s \Delta \sum_{l=\ell(s')}^{\kappa(s,T_j)} E\left[\zeta^l(s')\right]^2 ds'.$$

Let $\mathcal{E}(s) := \max_{0\leq j\leq N} E\left[\zeta^j_{t,x}(s)\right]^2$. Then

$$\mathcal{E}(s) \leq K\mu^2_{Max} + K\int_t^s \mathcal{E}(s')ds',$$

where $K > 0$ does not depend on Δ and x. Hence, by Gronwall's inequality

$$\mathcal{E}(s) \leq K\mu^2_{Max}, \quad t \leq s \leq t^*. \tag{4.43}$$

Next, we consider the $(N+1)^2$-dimensional process

$$\zeta^{j_1,j_2}(s) := \sum_{i_1,i_2=0}^N \mu^{i_1,i_2} \frac{\partial}{\partial x^{i_1}} \tilde{f}^{j_1}_{t,x}(s) \frac{\partial}{\partial x^{i_2}} \tilde{f}^{j_2}_{t,x}(s),\ j_1, j_2 = 0, \ldots, N,\ s \geq t.$$

Using the same recipe as in the case of estimating $\max_{0\leq j\leq N} E\left[\zeta^j_{t,x}(s)\right]^2$, we get that

$$\max_{0\leq j\leq N} E\left[\zeta^{j_1,j_2}_{t,x}(s)\right]^2 \leq K\mu^2_{Max}, \tag{4.44}$$

where $K > 0$ does not depend on Δ and x. Using (4.44) and repeating the same recipe again in the case of the processes

$$
\begin{aligned}
{}_2\zeta^j_{t,x}(s) \quad &:= \sum_{i_1,i_2=0}^{N} \mu^{i_1,i_2} \frac{\partial^2}{\partial x^{i_1}\partial x^{i_2}} \tilde{f}^j_{t,x}(s),\ j=0,\ldots,N, \\
\eta^j_{t,x}(s) \quad &:= \sum_{i_1,i_2=0}^{N} \mu^{i_1,i_2} \frac{\partial}{\partial x^{i_1}} \tilde{Y}_{t,x,0}(s) \frac{\partial}{\partial x^{i_2}} \tilde{f}^j_{t,x}(s),\ j=0,\ldots,N,\ s\geq t,
\end{aligned}
$$

we obtain

$$
\max_{0\leq j\leq N} E\left[{}_2\zeta^j_{t,x}(s)(s)\right]^2 \quad \leq \quad K\mu^2_{Max}, \tag{4.45}
$$

$$
\max_{0\leq j\leq N} E\left[\eta^j_{t,x}(s)\right]^2 \quad \leq \quad K\mu^2_{Max}, \tag{4.46}
$$

where $K>0$ does not depend on Δ and x. Using (4.45), it is not difficult to get that for the process

$$
{}_2\eta_{t,x}(s) := \sum_{i_1,i_2=0}^{N} \mu^{i_1,i_2} \frac{\partial^2}{\partial x^{i_1}\partial x^{i_2}} \tilde{Y}_{t,x,0}(s)
$$

the following estimate also holds

$$
E\left[{}_2\eta_{t,x}(s)\right]^2 \leq K\mu^2_{Max}. \tag{4.47}
$$

It is clear that (4.44)-(4.47) are sufficient for proving (4.42) with $m=2$. To show (4.42) for $m=3$, we need to obtain estimates for the second moments of the processes

$$
\begin{aligned}
\zeta^{j_1,j_2,j_3}(s) \quad &= \quad \sum_{i_1,i_2,i_3=0}^{N} \mu^{i_1,i_2,i_3} \frac{\partial}{\partial x^{i_1}} \tilde{f}^{j_1}_{t,x}(s) \frac{\partial}{\partial x^{i_2}} \tilde{f}^{j_2}_{t,x}(s) \frac{\partial}{\partial x^{i_3}} \tilde{f}^{j_3}_{t,x}(s),\ j_1,j_2,j_3=0,\ldots,N, \\
{}_2\zeta^{j_1,j_2}_{t,x}(s) \quad &= \quad \sum_{i_1,i_2,i_3=0}^{N} \mu^{i_1,i_2,i_3} \frac{\partial^2}{\partial x^{i_1}\partial x^{i_2}} \tilde{f}^{j_1}_{t,x}(s) \frac{\partial}{\partial x^{i_3}} \tilde{f}^{j_2}_{t,x}(s),\ j_1,j_2=0,\ldots,N, \\
{}_3\zeta^j_{t,x}(s) \quad &= \quad \sum_{i_1,i_2,i_3=0}^{N} \mu^{i_1,i_2,i_3} \frac{\partial^3}{\partial x^{i_1}\partial x^{i_2}\partial x^{i_3}} \tilde{f}^j_{t,x}(s),\ j=0,\ldots,N,\ s\geq t,
\end{aligned}
$$

and

$$\eta_{t,x}^{j_1,j_2}(s) = \sum_{i_1,i_2,i_3=0}^{N} \mu^{i_1,i_2,i_3} \frac{\partial}{\partial x^{i_1}} \tilde{f}_{t,x}^{j_1}(s) \frac{\partial}{\partial x^{i_2}} \tilde{f}_{t,x}^{j_2}(s) \frac{\partial}{\partial x^{i_3}} \tilde{Y}_{t,x,0}(s), \ j_1, j_2 = 0, \ldots, N,$$

$$_{2,1}\eta_{t,x}^{j}(s) = \sum_{i_1,i_2,i_3=0}^{N} \mu^{i_1,i_2,i_3} \frac{\partial^2}{\partial x^{i_1} \partial x^{i_2}} \tilde{f}_{t,x}^{j}(s) \frac{\partial}{\partial x^{i_3}} \tilde{Y}_{t,x,0}(s), \ j = 0, \ldots, N,$$

$$_{1,2}\eta_{t,x}^{j}(s) = \sum_{i_1,i_2,i_3=0}^{N} \mu^{i_1,i_2,i_3} \frac{\partial}{\partial x^{i_1}} \tilde{f}_{t,x}^{j}(s) \frac{\partial^2}{\partial x^{i_2} \partial x^{i_3}} \tilde{Y}_{t,x,0}(s), \ j = 0, \ldots, N,$$

$$_{3}\eta_{t,x}(s) = \sum_{i_1,i_2,i_3=0}^{N} \mu^{i_1,i_2,i_3} \frac{\partial^3}{\partial x^{i_1} \partial x^{i_2} \partial x^{i_3}} \tilde{Y}_{t,x,0}(s), \ s \geq t,$$

which can be done using the same recipe but with more laborious calculations. In the case of an arbitrary m one need to consider processes $\zeta^{j_1,\ldots,j_m}(s)$, $_2\zeta_{t,x}^{j_1,\ldots,j_{m-1}}(s)$, $\ldots$, $_m\zeta_{t,x}^{j}(s)$, $\eta_{t,x}^{j_1,\ldots,j_{m-1}}(s)$, $_{m-1,1}\eta_{t,x}^{j}(s), \ldots,$ $_m\eta_{t,x}(s)$ defined in the same fashion as we did in the cases $m = 2$ and 3. It is not difficult to see that employing the same recipe maxima of their second moments will be again bounded by $K\mu_{Max}^2$, from which (4.42) follows for an arbitrary m.

The required inequality (4.33) follows from (4.41) and (4.42). Lemma 4.1.4 is proved. $\square$

Using Lemma 4.1.4, we now prove convergence of $\bar{F}(t_0, f_0; t^*, T^*)$ to $\tilde{F}(t_0, f_0; t^*, T^*)$ in the case of smooth payoffs G.

Theorem 4.1.6 *Let $h \leq \alpha\Delta$ for some $\alpha > 0$. Suppose Assumptions 3.2.1-3.2.3 and Assumptions 3.4.1, 3.4.5, and 3.4.6 are satisfied. Assume that the payoff function $G(z)$ has bounded derivatives up to a sufficiently high order. Then the approximation $\bar{F}(t_0, f_0; t^*, T^*)$ defined by* (3.88)-(3.90), (3.72), (3.84) *converges to $\tilde{F}(t_0, f_0; t^*, T^*)$ from* (4.1) *with order $q > 0$, i.e.,*

$$\left|\tilde{F}(t_0, f_0; t^*, T^*) - \bar{F}(t_0, f_0; t^*, T^*)\right| \leq Kh^q, \tag{4.48}$$

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

Proof. Using the standard technique (see [58, p. 100]), we can write the difference R_2 in the form

$$\begin{aligned} R_2 &= \tilde{F}(t_0, f_0; t^*, T^*) - \bar{F}(t_0, f_0; t^*, T^*) \qquad (4.49) \\ &= E\exp(-\tilde{Y}_{t_0,f_0,0}(t_M))G(\exp(-\tilde{S}_Z(t_0, f_0; t^*, T^*, \Delta))) \\ &\quad -E\exp(-\bar{Y}_M)G(\exp(-\bar{S}_Z(t^*, T^*, \Delta))) \end{aligned}$$

$$\begin{aligned} &= \sum_{i=0}^{M-1} E\Big[\exp(-\tilde{Y}_{t_i,\bar{f}_i,\bar{Y}_i}(t_M))G(\exp(-\tilde{S}_Z(t_i, \bar{f}_i; t^*, T^*, \Delta))) \\ &\quad - \exp(-\tilde{Y}_{t_{i+1},\bar{f}_{i+1},\bar{Y}_{i+1}}(t_M))G(\exp(-\tilde{S}_Z(t_{i+1}, \bar{f}_{i+1}; t^*, T^*, \Delta)))\Big] \\ &= \sum_{i=0}^{M-1} E\{\exp(-\tilde{Y}_{t_i,\bar{f}_i,\bar{Y}_i}(t_{i+1}))E[\exp(-\tilde{Y}_{t_{i+1},\tilde{f}_{t_i,\bar{f}_i}(t_{i+1}),0}(t_M)) \\ &\quad \times G(\exp(-\tilde{S}_Z(t_{i+1}, \tilde{f}_{t_i,\bar{f}_i}(t_{i+1}); t^*, T^*, \Delta)))|\tilde{f}_{t_i,\bar{f}_i}(t_{i+1})] \\ &\quad - \exp(-\bar{Y}_{i+1})E[\exp(-\tilde{Y}_{t_{i+1},\bar{f}_{i+1},0}(t_M))G(\exp(-\tilde{S}_Z(t_{i+1}, \bar{f}_{i+1}; t^*, T^*, \Delta)))|\bar{f}_{i+1}]\} \end{aligned}$$

$$\begin{aligned} &= \sum_{i=0}^{M-1} E\Big\{\exp(-\tilde{Y}_{t_i,\bar{f}_i,\bar{Y}_i}(t_{i+1}))\tilde{F}(t_{i+1}, \tilde{f}_{t_i,\bar{f}_i}(t_{i+1}); t^*, T^*) \\ &\quad - \exp(-\bar{Y}_{i+1})\tilde{F}(t_{i+1}, \bar{f}_{i+1}; t^*, T^*)\Big\} \\ &= \sum_{i=0}^{M-1} E\exp(-\bar{Y}_i)E\Big[\exp(-\tilde{Y}_{t_i,\bar{f}_i,0}(t_{i+1}))\tilde{F}(t_{i+1}, \tilde{f}_{t_i,\bar{f}_i}(t_{i+1}); t^*, T^*) \\ &\quad - \exp(-\bar{Y}_{t_i,\bar{f}_i,0}(t_{i+1}))\tilde{F}(t_{i+1}, \bar{f}_{t_i,\bar{f}_i}(t_{i+1}); t^*, T^*)\,\Big|\bar{f}_i\Big] \\ &= \sum_{i=0}^{M-1} E\exp(-\bar{Y}_i)\rho(t_i, \bar{f}_i), \end{aligned}$$

where

$$\begin{aligned} \rho(t, x) &= E[\exp(-\tilde{Y}_{t,x,0}(t+h))\tilde{F}(t+h, \tilde{f}_{t,x}(t+h); t^*, T^*) \qquad (4.50) \\ &\quad - \exp(-\bar{Y}_{t,x,0}(t+h))\tilde{F}(t+h, \bar{f}_{t,x}(t+h); t^*, T^*)]. \end{aligned}$$

Now we write the Taylor expansion of the terms under expectation in (4.50) in powers of $\delta\tilde{Y} = -\tilde{Y}_{t,x,0}(t+h)$ and $\delta\tilde{f}^i = \tilde{f}^i_{t,x}(t+h) - x^i$ and in powers of $\delta\bar{Y} = -\bar{Y}_{t,x,0}(t+h)$ and $\delta\bar{f}^i_{t,x} = \bar{f}^i_{t,x}(t+h) - x^i$. As a result, we obtain

$$
\begin{aligned}
& \exp(-\tilde{Y}_{t,x,0}(t+h))\tilde{F}(t+h, \tilde{f}_{t,x}(t+h); t^*, T^*) \qquad (4.51)\\
= \; & \tilde{F}(t+h, x; t^*, T^*) \\
& + \sum_{|\mathbf{i}|+k=1}^{2q+1} \frac{1}{\mathbf{i}!k!} \frac{\partial^{|\mathbf{i}|}}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}} \tilde{F}(t+h, x; t^*, T^*) \left(\delta\tilde{f}^0\right)^{i_0} \cdots \left(\delta\tilde{f}^N\right)^{i_N} \delta\tilde{Y}^k \\
& + \sum_{|\mathbf{i}|+k=2q+2} \frac{1}{\mathbf{i}!k!} \frac{\partial^{|\mathbf{i}|}}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}} \tilde{F}(t+h, x + \tilde{\chi}(\tilde{f}_{t,x}(t+h) - x); t^*, T^*) \\
& \times \exp(-\tilde{\theta}\tilde{Y}_{t,x,0}(t+h)) \times \left(\delta\tilde{f}^0\right)^{i_0} \cdots \left(\delta\tilde{f}^N\right)^{i_N} \delta\tilde{Y}^k,
\end{aligned}
$$

where $\tilde{\chi}$ and $\tilde{\theta}$ are from $[0,1]$.

Further,

$$
\begin{aligned}
& \exp(-\bar{Y}_{t,x,0}(t+h))\tilde{F}(t+h, \bar{f}_{t,x}(t+h); t^*, T^*) \qquad (4.52)\\
= \; & \tilde{F}(t+h, x; t^*, T^*) \\
& + \sum_{|\mathbf{i}|+k=1}^{2q+1} \frac{1}{\mathbf{i}!k!} \frac{\partial^{|\mathbf{i}|}}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}} \tilde{F}(t+h, x; t^*, T^*) \left(\delta\bar{f}^0\right)^{i_0} \cdots \left(\delta\bar{f}^N\right)^{i_N} \delta\bar{Y}^k \\
& + \sum_{|\mathbf{i}|+k=2q+2} \frac{1}{\mathbf{i}!k!} \frac{\partial^{|\mathbf{i}|}}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}} \tilde{F}(t+h, x + \bar{\chi}(\bar{f}_{t,x}(t+h) - x); t^*, T^*) \\
& \times \exp(-\bar{\theta}\bar{Y}_{t,x,0}(t+h)) \left(\delta\bar{f}^0\right)^{i_0} \cdots \left(\delta\bar{f}^N\right)^{i_N} \delta\bar{Y}^k,
\end{aligned}
$$

with $\bar{\chi}$ and $\bar{\theta}$ being from $[0,1]$.

It is not difficult to check (see also (3.83)) that under the assumed condition $h \le \alpha\Delta$, $\alpha > 0$, the following inequality holds:

$$
\left[E \max_{0 \le m \le 2q+2, \{i_1, \ldots i_{2q+2-m}\} \in \{0,\ldots,N\}} \left| \delta\tilde{Y}^m \prod_{j=1}^{2q+2-m} \delta\tilde{f}^{i_j} \right|^2 \right]^{1/2} \le C h^{q+1} \left(1 + \sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l|^{2q+2} \right), \qquad (4.53)
$$

where $C > 0$ is independent of Δ. We note that the number of components x^l appearing in the right-hand side of (4.53) is not larger than $1+\theta$, which does not depend on Δ.

Using Lemma 4.1.4, the inequalities (4.53), (4.38) and (3.70), and the Cauchy-Bunyakovsky inequality, we obtain

$$
\begin{aligned}
&|E \sum_{|\mathbf{i}|+k=2q+2} \frac{1}{\mathbf{i}!k!} \frac{\partial^{|\mathbf{i}|}}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}} \tilde{F}(t+h, x+\tilde{\chi}(\tilde{f}_{t,x}(t+h)-x); t^*, T^*) \\
&\times \exp(-\tilde{\theta}\tilde{Y}_{t,x,0}(t+h)) \left(\delta\tilde{f}^0\right)^{i_0} \cdots \left(\delta\tilde{f}^N\right)^{i_N} \delta\tilde{Y}^k| \\
\leq\ & E\exp(|\tilde{Y}_{t,x,0}(t+h)|) \times \sum_{k=0}^{2q+2} \Bigg| \sum_{|\mathbf{i}|=2q+2-k} \frac{1}{\mathbf{i}!k!} \left(\delta\tilde{f}^0\right)^{i_0} \cdots \left(\delta\tilde{f}^N\right)^{i_N} \delta\tilde{Y}^k \\
&\times \frac{\partial^{|\mathbf{i}|}}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}} \tilde{F}(t+h, x+\tilde{\chi}(\tilde{f}_{t,x}(t+h)-x); t^*, T^*)\Bigg| \\
\leq\ & KE\Big[\exp(|\tilde{Y}_{t,x,0}(t+h)|) \max_{|\mathbf{i}|+k=2q+2} | \left(\delta\tilde{f}^0\right)^{i_0} \cdots \left(\delta\tilde{f}^N\right)^{i_N} \delta\tilde{Y}^k| \\
&\times \exp(c\Delta|x+\tilde{\chi}(\tilde{f}_{t,x}(t+h)-x)|)\Big] \\
\leq\ & K\exp(c\Delta|x|) \left(1+\sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l|^{2q+2}\right) h^{q+1},
\end{aligned}
\tag{4.54}
$$

where $K>0$ and $c>0$ independent of Δ, h, and x.

Analogously, using Lemma 4.1.4, the inequality (3.86) from Assumption 3.4.5, and Assumptions 3.4.1 and 3.4.6, we get

$$
\begin{aligned}
&|E \sum_{|\mathbf{i}|+k=2q+2} \frac{1}{\mathbf{i}!k!} \frac{\partial^{|\mathbf{i}|}}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}} \tilde{F}(t+h, x+\bar{\chi}(\bar{f}_{t,x}(t+h)-x); t^*, T^*) \\
&\qquad \times \exp(-\bar{\theta}\bar{Y}_{t,x,0}(t+h)) \left(\delta\bar{f}^0\right)^{i_0} \cdots \left(\delta\bar{f}^N\right)^{i_N} | \\
&\qquad \leq K\exp(c\Delta|x|) \left(1+\sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l|^{2q+2}\right) h^{q+1}.
\end{aligned}
\tag{4.55}
$$

We obtain from (4.50)-(4.52) and (4.54), (4.55):

$$
\begin{aligned}
|\rho(t,x)| \ \leq\ & \sum_{k=0}^{2q} \left| \sum_{|\mathbf{i}|=1}^{2q+1-k} \mu_k^{\mathbf{i}} \frac{\partial^{|\mathbf{i}|}}{(\partial x^0)^{i_0} \cdots (\partial x^N)^{i_N}} \tilde{F}(t+h, x; t^*, T^*) \right| \\
&+K\exp(c\Delta|x|) \left(1+\sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l|^{2q+2}\right) h^{q+1},
\end{aligned}
\tag{4.56}
$$

with $K > 0$ and $c > 0$ independent of Δ and

$$\mu_k^{\mathbf{i}} = \frac{1}{\mathbf{i}!k!}\left[E\left(\delta\tilde{f}^0\right)^{i_0}\cdots\left(\delta\tilde{f}^N\right)^{i_N}\delta\tilde{Y}^k - E\left(\delta\bar{f}^0\right)^{i_0}\cdots\left(\delta\bar{f}^N\right)^{i_N}\delta\bar{Y}^k\right].$$

Applying Lemma 4.1.4 and using the inequality (3.85) from Assumption 3.4.5, we obtain from (4.56):

$$|\rho(t,x)| \le K\exp(c\Delta|x|)\left(1+\sum_{l=\ell(t)}^{\ell(t+h)+\theta}|x^l|^{2q+2}\right)h^{q+1}, \tag{4.57}$$

where $K > 0$ and $c > 0$ do not depend on Δ and x.

Substituting (4.57) in (4.49) and using Assumptions 3.4.1 and 3.4.6 and the Cauchy-Bunyakovsky inequality, we arrive at the required (4.48). Theorem 4.1.6 is proved. □

Remark 4.1.7 *As it follows from the proof, in Theorem 4.1.6 the condition $h \le \alpha\Delta$ is used only for estimating the parts of the error involving the approximate discounting factor $\tilde{Y}_{t,x,0}(s)$. If pricing an interest rate derivative does not require a discounting factor (e.g., when one uses the forward measure pricing, cf. Remark 3.2.5) then a theorem analogous to Theorem 4.1.6 can be proved under Assumptions 3.2.1- 3.2.3 and Assumptions 3.4.1 and 3.4.2 without the restriction on h.*

Theorems 4.1.3 and 4.1.6 imply the following result.

Theorem 4.1.8 *Under the conditions of Theorems 4.1.1 and 4.1.6, the approximation $\bar{F}(t_0, f_0;\ t^*, T^*)$ defined by (3.88)-(3.90), (3.72), (3.84) converges to $F(t_0, f_0(\cdot); t^*, T^*)$ from (3.51)-(3.54), (3.43)-(3.44) with order $p > 0$ in Δ and with order $q > 0$ in h, i.e.,*

$$\left|F(t_0, f_0(\cdot); t^*, T^*) - \bar{F}(t_0, f_0; t^*, T^*)\right| \le K(\Delta^p + h^q), \tag{4.58}$$

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

According to the motivation examples considered in Section 2.2, the payoff $G(z)$ is usually globally Lipschitz (see (3.50)) but not sufficiently smooth function as it is required in Theorem 4.1.6 and, consequently, in Theorem 4.1.8. Let us discuss two ways how one can deal with this theoretical difficulty.

First, as it was noted in, e.g. [59], we can approximate the payoff function $G(z)$ by a smooth function $\breve{G}(z)$. Denote by ε an error of this approximation. The proposed numerical method can be applied to the smooth approximating function $\breve{G}(z)$ and Theorems 4.1.6 and 4.1.8 remain valid for F with $\breve{G}$ instead of G. In this case, the overall error in evaluating the price of an interest rate contract consists of the numerical integration errors estimated in Theorem 4.1.8 and the error ε of the smoothening of G.

Second, one can exploit the result of [4] which in application to our problem means that if the transition Markov function for the process $\tilde{f}(t)$ is sufficiently smooth and $\bar{f}_k^i$ is simulated by the strong Euler scheme then $\bar{F}(t_0, f_0; t^*, T^*)$ converges to $\tilde{F}(t_0, f_0; t^*, T^*)$ with order one in h even for nonsmooth G.

We remark that the computational practice (see our numerical experiments in Section 3.7) suggests that the error estimates of Theorems 4.1.6 and 4.1.8 are valid for the weak Euler-type scheme (see (3.95) below) in the case of nonsmooth $G(z)$. Further, it is natural to expect that for higher-order weak schemes the error estimates of Theorems 4.1.6 and 4.1.8 are also valid for nonsmooth payoffs $G(z)$. We note that to answer on these theoretical questions related to nonsmoothness of $G(z)$ further development of the general theory of numerical integration of ordinary SDEs is required which is outside the scope of the present thesis.

4.2 Mean-square convergence theorems

In this Section first we shall prove mean-square convergence of $\bar{f}_k^i$ defined in (3.131) to $\tilde{f}^i(t_k)$ from (3.67)-(3.69) is uniform in Δ. This result is then used to prove mean-square convergence of $\bar{f}_k^i$ to $f(t_k, T_i)$ from (3.43)-(3.44) exploiting in addition Theorem 4.1.1. We cannot use here the standard result, e.g. the fundamental theorem of mean-square convergence [58, p. 4], since we need to show that the convergence is uniform in Δ.

Theorem 4.2.1 *Suppose Assumptions 3.2.1-3.2.3 and Assumption 3.6.1 are satisfied. Then for any* M, N *and* $k = 0, 1, \dots, M$, $i = 0, 1, \dots, N$ *the following inequality holds:*

$$\left[E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2\right]^{1/2} \le Kh^{q_2 - 1/2}\,, \tag{4.59}$$

i.e., the order of mean-square accuracy of the method (3.131) *for* (3.67) *is* $q = q_2 - 1/2$.

Proof. We have (cf. [58, pp. 7-8])

$$\begin{aligned}\tilde{f}^i(t_{k+1}) - \bar{f}^i_{k+1} &= \tilde{f}^i_{t_0,f_0}(t_{k+1}) - \bar{f}^i_{t_0,f_0}(t_{k+1}) = \tilde{f}^i_{t_k,\tilde{f}(t_k)}(t_{k+1}) - \bar{f}^i_{t_k,\bar{f}_k}(t_{k+1}) \\ &= (\tilde{f}^i_{t_k,\tilde{f}(t_k)}(t_{k+1}) - \tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1})) + (\tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}) - \bar{f}^i_{t_k,\bar{f}_k}(t_{k+1}))\,,\end{aligned} \tag{4.60}$$

where the first difference in the right-hand side of (4.60) is the error of the solution arising due to the error in the initial data at time t_k, accumulated over k steps, and the second difference is the one-step error at the $(k+1)$-step. Taking the square of both sides of (4.60), we obtain

$$\begin{aligned}R^2_{i,k+1} &:= E|\tilde{f}^i(t_{k+1}) - \bar{f}^i_{k+1}|^2 \\ &= EE(|\tilde{f}^i_{t_k,\tilde{f}(t_k)}(t_{k+1}) - \tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1})|^2|\mathcal{F}_{t_k}) \\ &+ EE(|\tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}) - \bar{f}^i_{t_k,\bar{f}_k}(t_{k+1})|^2|\mathcal{F}_{t_k}) \\ &+ 2EE((\tilde{f}^i_{t_k,\tilde{f}(t_k)}(t_{k+1}) - \tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}))(\tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}) - \bar{f}^i_{t_k,\bar{f}_k}(t_{k+1}))|\mathcal{F}_{t_k})\,.\end{aligned} \tag{4.61}$$

Due to the condition (3.134), we get for the second term on the right-hand side of (4.61):

$$|EE(|\tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}) - \bar{f}^i_{t_k,\bar{f}_k}(t_{k+1})|^2|\mathcal{F}_{t_k})| \le Ch^{2q_2}. \tag{4.62}$$

Let us estimate the first term on the right-hand side of (4.61). Ito's formula implies that

$$\varepsilon_i^2(t_{k+1}) := E|\tilde{f}^i_{t_k,\tilde{f}(t_k)}(t_{k+1}) - \tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1})|^2$$
$$= E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2 + 2E\int_{t_k}^{t_{k+1}} (\tilde{f}^i_{t_k,\tilde{f}(t_k)}(s) - \tilde{f}^i_{t_k,\bar{f}_k}(s))$$
$$\times[\sigma^\top(s,T_i,\tilde{f}^i_{t_k,\tilde{f}(t_k)}(s))\tilde{S}_I(t_k,\tilde{f}(t_k);s,T_i,\Delta) - \sigma^\top(s,T_i,\tilde{f}^i_{t_k,\bar{f}_k}(s))\tilde{S}_I(t_k,\bar{f}_k;s,T_i,\Delta)]ds$$
$$+E\int_{t_k}^{t_{k+1}} |\sigma(s,T_i,\tilde{f}^i_{t_k,\tilde{f}(t_k)}(s)) - \sigma(s,T_i,\tilde{f}^i_{t_k,\bar{f}_k}(s))|^2 ds.$$

Then, recalling that $\sigma(s,T,z)$ is globally Lipschitz in z due to Assumption 3.2.2 and the form of $\tilde{S}_I(t_k,\bar{f}_k;s,T_i,\Delta)$ (see (4.35)), we obtain

$$\begin{aligned}\varepsilon_i^2(t_{k+1}) \;&\leq\; E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2 + K\int_{t_k}^{t_{k+1}} E|\tilde{f}^i_{t_k,\tilde{f}(t_k)}(s) - \tilde{f}^i_{t_k,\bar{f}_k}(s)|^2 ds \\ &\quad + K\Delta\int_{t_k}^{t_{k+1}} \sum_{l=\ell(s)}^{\kappa(s,T_i)} E|\tilde{f}^l_{t_k,\tilde{f}(t_k)}(s) - \tilde{f}^l_{t_k,\bar{f}_k}(s)|^2 ds,\end{aligned}$$

where $K > 0$ does not depend on Δ. Introduce $\varepsilon^2_{Max}(s) := \max_{0\leq i\leq M} \varepsilon_i^2(s)$. Then

$$\varepsilon^2_{Max}(t_{k+1}) \leq \max_{0\leq i\leq M} E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2 + K\int_{t_k}^{t_{k+1}} \varepsilon^2_{Max}(s)ds$$

which implies that for all $0 \leq i \leq M$ and all sufficiently small $h > 0$:

$$\varepsilon_i^2(t_{k+1}) \leq e^{Kh} \max_{0\leq i\leq M} E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2 \leq \max_{0\leq i\leq M} E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2 \cdot (1+Kh), \qquad (4.63)$$

where $K > 0$ does not depend on Δ and h.

Now let us estimate the third term on the right-hand side of (4.61). We have (cf. Lemma 1.1.3 in [58, p. 5]):

$$\tilde{f}^i_{t_k,\tilde{f}(t_k)}(t_{k+1}) - \tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}) = \tilde{f}^i(t_k) - \bar{f}^i_k + Z^i, \qquad (4.64)$$

where

$$
\begin{aligned}
Z^i &= \int_{t_k}^{t_{k+1}} [\sigma^\top(s, T_i, \tilde{f}^i_{t_k,\tilde{f}(t_k)}(s))\tilde{S}_I(t_k, \tilde{f}(t_k); s, T_i, \Delta) \\
&\quad -\sigma^\top(s, T_i, \tilde{f}^i_{t_k,\bar{f}_k}(s))\tilde{S}_I(t_k, \bar{f}_k; s, T_i, \Delta)]ds \\
&\quad + \int_{t_k}^{t_{k+1}} [\sigma^\top(s, T_i, \tilde{f}^i_{t_k,\tilde{f}(t_k)}(s)) - \sigma^\top(s, T_i, \tilde{f}^i_{t_k,\bar{f}_k}(s))]dW(s).
\end{aligned}
$$

Using (4.63), it is not difficult to get

$$
E\left(Z^i\right)^2 \le Kh \cdot \max_{0\le i\le M} E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2, \tag{4.65}
$$

where $K > 0$ does not depend on Δ and h. Using (4.64), (3.133), (3.134), (4.65), and (3.132), we obtain

$$
\begin{aligned}
& |EE((\tilde{f}^i_{t_k,\tilde{f}(t_k)}(t_{k+1}) - \tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}))(\tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}) - \bar{f}^i_{t_k,\bar{f}_k}(t_{k+1}))|\mathcal{F}_{t_k})| \qquad (4.66) \\
\le\ & |E(\tilde{f}^i(t_k) - \bar{f}^i_k)E(\tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}) - \bar{f}^i_{t_k,\bar{f}_k}(t_{k+1}))|\mathcal{F}_{t_k})|
\end{aligned}
$$

$$
\begin{aligned}
& +|EZ^i \cdot (\tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}) - \bar{f}^i_{t_k,\bar{f}_k}(t_{k+1}))| \\
\le\ & (E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2)^{1/2} \cdot Kh^{q_1} + \left(E\left(Z^i\right)^2\right)^{1/2}\left(E(\tilde{f}^i_{t_k,\bar{f}_k}(t_{k+1}) - \bar{f}^i_{t_k,\bar{f}_k}(t_{k+1})^2\right)^{1/2} \\
\le\ & Kh^{q_1}(E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2)^{1/2} + Kh^{q_2+1/2} \cdot \left(\max_{0\le i\le M} E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2\right)^{1/2} \\
\le\ & Kh^{q_2+1/2} \cdot \left(\max_{0\le i\le M} E|\tilde{f}^i(t_k) - \bar{f}^i_k|^2\right)^{1/2},
\end{aligned}
$$

where $K > 0$ does not depend on Δ and h.

Let $R^2_{Max,k} := \max_{0\le i\le M} R^2_{i,k}$. Then it follows from (4.61), (4.62), (4.63) and (4.66) that

$$
R^2_{Max,k+1} \le R^2_{Max,k} \cdot (1 + Kh) + Kh^{q_2+1/2}R_{Max,k} + Ch^{2q_2}.
$$

Using the elementary relation

$$h^{q_2+1/2} R_{Max,k} \leq \frac{R_{Max,k}^2 h}{2} + \frac{h^{2q_2}}{2} \,,$$

we get

$$R_{Max,k+1}^2 \leq R_{Max,k}^2 \cdot (1 + Kh) + Ch^{2q_2}$$

whence (4.59) follows taking into account Lemma 1.1.6 from [58, p. 7] and the fact that $R_{Max,0}^2 = 0$. Theorem 4.2.1 is proved. $\square$

Theorems 4.1.1 and 4.2.1 imply the following result.

Theorem 4.2.2 *Assume that the conditions of Theorems 4.1.1 and 4.2.1 hold. Then for any* M, N *and* $i = 0, 1, \ldots, N$, $k = 0, 1, \ldots, \lceil (T_i - t_0)/h \rceil - 1$ *the mean-square error is estimated as*

$$\left[E|f(t_k, T_i) - \bar{f}_k^i|^2 \right]^{1/2} \leq K \cdot (\Delta^p + h^{q_2 - 1/2}) \,, \tag{4.67}$$

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

4.3 Checking assumptions for Algorithm 3.5.2

In Section 3.5 we gave examples of particular algorithmic realisations of the generic numerical method defined in Section 3.4. In this Section we are going to demonstrate a routine check of the assumptions we imposed on the method on the example of one of the algorithms from Section 3.5. More specifically, we shall be checking that Assumptions 3.4.1, 3.4.5 and 3.4.6 are satisfied by Algorithm 3.5.2.

Recall (see Section 3.5) that Algorithm 3.5.2 is defined as follows

$$\bar{f}_0^i = f_0^i, \; i = 0, \ldots, N, \;\; \bar{Y}_0 = 0, \tag{4.68}$$

$$\bar{f}_{k+1}^i = \bar{f}_k^i + \sum_{j=1}^d \bar{\sigma}_{i,j}(t_k) \mathbb{S}_{I_j}(t_k, T_i; \Delta, h) + h^{1/2} \sum_{j=1}^d \bar{\sigma}_{i,j}(t_k) \xi_{j,k+1},$$

$$i = \ell_{k+1}, \ldots, N,$$

$$\bar{Y}_{k+1} = \bar{Y}_k + A^Y(t_k; \bar{f}_k^j,\ j = \ell_k, \ldots, \ell(t^*) + \theta; h),\ k = 0, \ldots, M-1,$$

where $\xi_{j,k+1}$ are independent random variables distributed by the law $P(\xi = \pm 1) = 1/2$,

$$(\bar{\sigma}_{i,1}(t_k), \ldots, \bar{\sigma}_{i,d}(t_k))^\top = (\sigma_1(t_k, T_i, \bar{f}_k^i), \ldots, \sigma_d(t_k, T_i, \bar{f}_k^i))^\top,$$

$$\bar{\mathbb{S}}_{I_j}(t_k, T_{\ell_{k+1}}; \Delta, h) = \begin{cases} h\Delta_{\ell_{k+1},k}\bar{\sigma}_{\ell_{k+1},j}(t_k), & \text{if } T_{\ell_{k+1}} \le t_k, \\ \Delta_{\ell_{k+1},k}\Delta_{\ell_{k+1},k+1}\bar{\sigma}_{\ell_{k+1},j}(t_k), & \text{otherwise,} \end{cases} \tag{4.69}$$

$$\bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h) = \begin{cases} h\Delta_{\varrho_{k+1},k}\bar{\sigma}_{\varrho_{k+1},j}(t_k), & \text{if } T_{\ell_{k+1}} \le t_k, \\ \Delta_{\ell_{k+1},k}\ \frac{\Delta_{\varrho_{k+1},k}}{2}\bar{\sigma}_{\ell_{k+1},j}(t_k) - \Delta_{\ell_{k+1},k+2}\ \frac{\Delta}{2}\bar{\sigma}_{\varrho_{k+1},j}(t_k), & \text{otherwise,} \end{cases}$$

$$\bar{\mathbb{S}}_{I_j}(t_k, T_i; \Delta, h) = \bar{\mathbb{S}}_{I_j}(t_k, T_{\varrho_{k+1}}; \Delta, h)$$
$$+h\frac{\Delta}{2}\left(\bar{\sigma}_{\varrho_{k+1},j}(t_k) + 2\sum_{m=\varrho_{k+1}+1}^{i-1}\bar{\sigma}_{m,j}(t_k) + \bar{\sigma}_{i,j}(t_k)\right),$$
$$i = \varrho_{k+1}+1, \ldots, N, \quad j = 1, \ldots, d.$$

$$A^Y(t_k; h) = \begin{cases} h\left[\frac{\Delta_{\rho_{k+1},k+1/2}}{\Delta}\bar{f}_k^{\ell_{k+1}} - \frac{\Delta_{\ell_{k+1},k+1/2}}{\Delta}\bar{f}_k^{\varrho_{k+1}}\right], & \text{if } T_{\ell_{k+1}} \le t_k, \\ \Delta_{\ell_{k+1},k}\left[\frac{\Delta_{\ell_{k+1},k}}{2\Delta}\bar{f}_k^{\ell_k} - \frac{\Delta_{\ell_k-1,k}}{2\Delta}\bar{f}_k^{\ell_{k+1}}\right] & \\ -\Delta_{\ell_{k+1},k+1}\left[\frac{\Delta_{\varrho_{k+1}+1,k+1}}{2\Delta}\bar{f}_{k+1}^{\ell_{k+1}} - \frac{\Delta_{\ell_{k+1},k+1}}{2\Delta}\bar{f}_{k+1}^{\varrho_{k+1}}\right] & \text{otherwise.} \end{cases} \tag{4.70}$$

Checking Assumption 3.4.1 We need to check that for a $c \in R$ there is a constant $C > 0$ such that

$$E\exp(c|\bar{f}_k^i|) < C$$

for all $i = 0, \ldots N,\ k = 0, \ldots, M$. This condition is clearly satisfied by Algorithm 3.5.2 thanks to the uniform boundedness of $\sigma_i(t,T,z)$ (see Assumption 3.2.1) and boundedness of the initial condition (see Assumption 3.2.3 and also the comment after (3.70)).

Checking Assumption 3.4.6. We need to show that $\bar{Y}_k$ satisfied the estimate

$$E\exp(c|\bar{Y}_k|) < C$$

for some $c > 0$ and $C > 0$ for all $k = 0,\dots,M$. This result clearly follows from (4.70) and Assumption 3.4.1.

Checking Assumption 3.4.5. We need to demonstrate that Algorithm 3.5.2 is such that for some independent of Δ positive constant C

$$\left|E\left(\delta\tilde{Y}^m\prod_{j=1}^{s-m}\delta\tilde{f}^{i_j} - \delta\bar{Y}^m\prod_{j=1}^{s-m}\delta\bar{f}^{i_j}\right)\right| \le Ch^2\left(1+\sum_{l=\ell(t)}^{\ell(t+h)+1}|x^l|^m\right), \tag{4.71}$$

$$m = 0,\dots,s,\ \ s = 1,\dots,3;$$

$$\left[E\max_{0\le m\le 4,\{i_1,\dots i_{2q+2-m}\}\in\{0,\dots,N\}}\left|\delta\bar{Y}^m\prod_{j=1}^{4-m}\delta\bar{f}^{i_j}\right|^2\right]^{1/2} \le Ch^2\left(1+\sum_{l=\ell(t)}^{\ell(t+h)}|x^l|^4\right), \tag{4.72}$$

where

$$\delta\tilde{f}^i = \tilde{f}^i_{t,x}(t+h) - x_i = \int_t^{t+h}\tilde{\sigma}_i^\top(s)\tilde{S}_I(s,T_i,\Delta)ds + \int_t^{t+h}\tilde{\sigma}_i^\top(s)dW(s), \tag{4.73}$$

$$\delta\bar{f}^i = \bar{f}^i_{t,x}(t+h) - x_i = \sigma^\top(t,T_i,x^i)\bar{\mathbb{S}}_{I_j}(t,T_i,\Delta) + h^{1/2}\sigma^\top(t,T_i,x^i)\xi_{j,k+1}, \tag{4.74}$$

where $\xi_{j,k+1}$ are independent random variables distributed by the law $P(\xi = \pm 1) = 1/2$,

$$\begin{aligned}\delta\tilde{Y} &= \tilde{Y}_{t,x,y}(t+h) - y \\ &= \int_t^{t+h\wedge T_{\varrho(t)}}\left[\frac{T_{\varrho(t)}-s}{\Delta}\tilde{f}^{\ell(t)}_{t,x}(s) + \frac{s-T_{\ell(t)}}{\Delta}\tilde{f}^{\varrho(t)}_{t,x}(s)\right]ds \\ &\quad + \int_{T_{\varrho(t)}\wedge t+h}^{t+h}\left[\frac{T_{\varrho(t+h)}-s}{\Delta}\tilde{f}^{\ell(t+h)}_{t,x}(s) + \frac{s-T_{\ell(t+h)}}{\Delta}\tilde{f}^{\varrho(t+h)}_{t,x}(s)\right]ds,\end{aligned} \tag{4.75}$$

$$
\begin{aligned}
\delta\bar{Y} &= \bar{Y}_{t,x,y}(t+h) - y \\
&= h \wedge \left(T_{\varrho(t)} - t\right)\left(\frac{T_{\ell(t+h)} - t}{2\Delta} \wedge \frac{T_{\ell(t+h)} - t - h/2}{\Delta} \cdot x^{\ell(t)}\right. \\
&\quad \left. + \frac{t + h/2 - T_{\ell(t)}}{\Delta} \wedge \frac{t + \Delta - T_{\ell(t)}}{2\Delta} \cdot x^{\varrho(t)}\right) \\
&\quad + 0 \vee \left(t + h - T_{\varrho(t)}\right)\left(\frac{T_{\varrho(t+h)} + \Delta - t - h}{\Delta} \bar{f}_{t,x}^{\ell(t+h)}(t+h)\right. \\
&\quad \left. + \frac{t + h - T_{\ell(t+h)}}{\Delta} \bar{f}_{t,x}^{\varrho(t+h)}(t+h)\right).
\end{aligned} \tag{4.76}
$$

In what follows, $O(h^2)$ are functions such that

$$
\left|O(h^2)\right| \leq Kh^2 \tag{4.77}
$$

where Kdoes not depend on h, Δ and x.

Let us first check condition (4.72). We have

$$
\begin{aligned}
\left|\delta\bar{Y}^m \prod_{j=1}^{4-m} \delta\bar{f}^{i_j}\right| = &\left|h \wedge \left(T_{\varrho(t)} - t\right)\left(\frac{T_{\ell(t+h)} - t}{2\Delta} \wedge \frac{T_{\ell(t+h)} - t - h/2}{\Delta} \cdot x^{\ell(t)}\right.\right. \\
&\left. + \frac{t + h/2 - T_{\ell(t)}}{\Delta} \wedge \frac{t + \Delta - T_{\ell(t)}}{2\Delta} \cdot x^{\varrho(t)}\right) \\
&+ \left(0 \vee \left(t + h - T_{\varrho(t)}\right) \frac{T_{\varrho(t+h)} + \Delta - t - h}{\Delta} \bar{f}_{t,x}^{\ell(t+h)}(t+h)\right. \\
&\left.\left. + \frac{t + h - T_{\ell(t+h)}}{\Delta} \bar{f}_{t,x}^{\varrho(t+h)}(t+h)^m\right)\right|^m \\
&\times \left|\prod_{j=1}^{4-m} \left(\sigma^\top(t, T_{i_j}, x^{i_j})\mathbb{S}_{I_j}(t, T_{i_j}, \Delta) + h^{1/2}\sigma^\top(t, T_{i_j}, x^{i_j})\xi\right)\right|
\end{aligned} \tag{4.78}
$$

$$
\begin{aligned}
&\leq h^m |x^{\ell(t)} + x^{\varrho(t)} + x^{\ell(t+h)} + x^{\varrho(t+h)} + \delta\bar{f}^{\ell(t+h)} + \delta\bar{f}^{\varrho(t+h)}|^m \prod_{j=1}^{4-m} (Kh + Kh^{1/2}) \\
&\leq Kh^{m+2-m/2}(|x^{\ell(t)}|^m + |x^{\varrho(t)}|^m + |x^{\ell(t+h)}|^m + |x^{\varrho(t+h)}|^m + |\delta\bar{f}^{\ell(t+h)}|^m + |\delta\bar{f}^{\varrho(t+h)}|^m) \\
&\leq Kh^{m/2+2}(|x^{\ell(t)}|^m + |x^{\varrho(t)}|^m + |x^{\ell(t+h)}|^m + |x^{\varrho(t+h)}|^m + Kh^{m/2})
\end{aligned}
$$

Hence

$$\max_{0\leq m\leq 4,\{i_1,\dots i_{2q+2-m}\}\in\{0,\dots,N\}}\left|\delta\bar{Y}^m\prod_{j=1}^{4-m}\delta\bar{f}^{i_j}\right| \tag{4.79}$$
$$\leq\ Kh^2(|x^{\ell(t)}|^4+|x^{\varrho(t)}|^4+|x^{\ell(t+h)}|^4+|x^{\varrho(t+h)}|^4+1).$$

We note that the estimate is not even random due to boundedness of σ and ξ. The range of indices, $m=0,\dots,s,\ s=1,\dots,3$, in (4.71) generates 9 cases (see below). We start with the case $m=0,\ s=1$, i.e. we want to establish

$$\left|E\left(\delta\tilde{f}^{i_j}-\delta\bar{f}^{i_j}\right)\right|\leq Ch^2, \tag{4.80}$$

for some independent of Δ positive constant C. By (4.73) and (4.74), we have

$$\left|E\left(\delta\tilde{f}^{i_j}-\delta\bar{f}^{i_j}\right)\right|=|E\int_t^{t+h}\tilde{\sigma}_i^\top(s)\tilde{S}_I(s,T_i,\Delta)ds-\sigma^\top(t,T_i,x^i)\bar{\mathbb{S}}_I(t,T_i,\Delta)|.$$

Applying Ito formula (cf. (3.67)), we obtain

$$\begin{aligned}\tilde{\sigma}_i(s)\ &=\ \sigma(s,T_i,\tilde{f}^i_{t,x}(s))=\sigma(t,T_i,x^i)\\&\quad+\int_t^s(\frac{\partial}{\partial s}+\tilde{\sigma}_i^\top(s)\tilde{S}_I(s,T_i,\Delta)\frac{\partial}{\partial x^i}+\frac{\sigma\sigma^\top}{2}\frac{\partial^2}{\partial x^i\partial x^i})\sigma(s',T_i,\tilde{f}^i_{t,x}(s'))ds'\\&\quad+\int_t^s\tilde{\sigma}_i^\top(s)\frac{\partial}{\partial x^i}\sigma(s',T_i,\tilde{f}^i_{t,x}(s'))dW(s').\end{aligned} \tag{4.81}$$

Then

$$\begin{aligned}&|E\int_t^{t+h}\tilde{\sigma}_i^\top(s)\tilde{S}_I(s,T_i,\Delta)ds-\sigma^\top(t,T_i,x^i)\bar{\mathbb{S}}_I(t,T_i,\Delta)|\\&=\ |\sigma^\top(t,T_i,x^i)\left[E\int_t^{t+h}\tilde{S}_I(s,T_i,\Delta)ds-\bar{\mathbb{S}}_I(t,T_i,\Delta)\right]|+O(h^2),\end{aligned} \tag{4.82}$$

where $O(h^2)$ as in (4.77) and we used above that $\bar{\mathbb{S}}_I(t,T_i,\Delta)$ is non-random, since (cf. 4.69) it is a linear combination of $\sigma(t,T_j,x^j)$, $j=\ell(t),\dots,i$.

Recall from Section 3.5 the procedure we used to construct $\bar{\mathbb{S}}$ based on the the quadrature rule defined in (3.101). For instance, for $T_i\geq t+h$, using (4.81), we have

$$E\int_t^{t+h}\tilde{S}_{I_j}(s,T_i,\Delta)ds = E\int_t^{t+h}[(T_{\varrho(s)}-s)\tilde{\sigma}_{\varrho(s),j}(s)$$
$$+\frac{\Delta}{2}\sum_{m=\varrho(s)}^{i-1}(\tilde{\sigma}_{m,j}(s)+\tilde{\sigma}_{m+1,j}(s))]ds = \int_t^{t+h}[(T_{\varrho(s)}-s)\sigma_j(t,T_{\varrho(t)},x^{\varrho(t)})$$
$$+\frac{\Delta}{2}\sum_{m=\varrho(s)}^{i-1}\big(\sigma_j(t,T_m,x^m)+\sigma_j(t,T_{m+1},x^{m+1})\big)]ds+O(h^2)$$

$$= \begin{cases} h(T_{\varrho(t)}-t-h/2)\sigma_j(t,T_{\varrho(t)},x^{\varrho(t)}) \\ +\frac{\Delta}{2}h\sum_{m=\varrho(t)}^{i-1}(\sigma_j(t,T_m,x^m)+\sigma_j(t,T_{m+1},x^{m+1})) \\ +O(h^2) \quad \text{if} \;\; T_{\varrho(t)} \ge t+h \\ \int_t^{T_{\varrho(t)}}[(T_{\varrho(t)}-s)\sigma_j(t,T_{\varrho(t)},x^{\varrho(t)})+\frac{\Delta}{2}\sum_{m=\varrho(t)}^{i-1}(\sigma_j(t,T_m,x^m) \\ +\sigma_j(t,T_{m+1},x^{m+1}))]ds+\int_{T_{\varrho(t)}}^{t+h}[(T_{\varrho(t)+1}-s)\sigma_j(t,T_{\varrho(t)+1},x^{\varrho(t)+1}) \\ +\frac{\Delta}{2}\sum_{m=\varrho(t)+1}^{i-1}(\sigma_j(t,T_m,x^m)+\sigma_j(t,T_{m+1},x^{m+1}))]ds \\ +O(h^2) \quad \text{if} \;\; T_{\varrho(t)} < t+h \end{cases}$$

$$= \begin{cases} h(T_{\varrho(t)}-t-h/2)\sigma_j(t,T_{\varrho(t)},x^{\varrho(t)})+\frac{\Delta}{2}h\sum_{m=\varrho(t)}^{i-1}(\sigma_j(t,T_m,x^m) \\ +\sigma_j(t,T_{m+1},x^{m+1})) \quad \text{if} \;\; T_{\varrho(t)} \ge t+h, \\ \frac{1}{2}(T_{\varrho(t)}-t)(T_{\varrho(t)}+\Delta-t)\sigma_j(t,T_{\varrho(t)},x^{\varrho(t)}) \\ +\frac{1}{2}\Delta(t+2h-T_{\varrho(t)})\sigma_j(t,T_{\varrho(t)+1},x^{\varrho(t)+1}) \\ +h\frac{\Delta}{2}\sum_{m=\varrho(t)+1}^{i-1}(\sigma_j(t,T_m,x^m)+\sigma_j(t,T_{m+1},x^{m+1}))+O(h^2) \quad \text{if} \;\; T_{\varrho(t)} < t+h \end{cases} \tag{4.83}$$

It is clear that $O(h^2)$ in (4.83) does not depend on Δ and x. Then, truncating the terms of order $O(h^2)$ in (4.83), we obtain $\bar{\mathbb{S}}$ as in (4.69). This gives us:

$$E\int_t^{t+h} \tilde{S}_I(s,T_i,\Delta)ds - \bar{\mathbb{S}}_I(t,T_i,\Delta) = O(h^2),\ T_i \geq t+h. \tag{4.84}$$

Analogously, we can establish the same estimate for the $t \leq T_i < t+h$. Thus, in view of (4.82) and (4.84), we establish (4.80).

Next, we consider the case $m=0$, $s=2$. Using (4.82), (4.84), (4.81), we obtain

$$\left|E\left(\prod_{j=1}^{2}\delta\tilde{f}^{i_j} - \prod_{j=1}^{2}\delta\bar{f}^{i_j}\right)\right| \tag{4.85}$$

$$= \left|E\left(\int_t^{t+h}\int_t^{t+h} \tilde{\sigma}_{i_1}^{\top}(s)\tilde{S}_I(s,T_{i_1},\Delta)\tilde{\sigma}_{i_2}^{\top}(\acute{s})\tilde{S}_I(\acute{s},T_{i_2},\Delta)dsd\acute{s} + \int_t^{t+h}\tilde{\sigma}_{i_1}^{\top}(s)\tilde{\sigma}_{i_2}(s)ds\right)\right.$$

$$\left. -\sigma^{\top}(t,T_{i_1},x^{i_1})\bar{\mathbb{S}}_I(s,T_{i_1},\Delta)\sigma^{\top}(t,T_{i_2},x^{i_1})\bar{\mathbb{S}}_I(\acute{s},T_{i_2},\Delta) - h\sigma^{\top}(t,T_{i_1},x^{i_1})\sigma(t,T_{i_2},x^{i_2})\right|$$

$$= O(h^2)$$

Analogously, we can show that (4.71) holds for $m=0$, $s=3$.

We shall now analyse the case $m=1$, $s=1$. We are going to consider each term in $\delta\tilde{Y}$ with its corresponding term in $\delta\bar{Y}$ separately. Using (4.80), (4.82) and (4.84), we obtain:

$$\left|E\int_t^{t+h\wedge T_{\varrho(t)}} \frac{T_{\varrho(t)}-s}{\Delta}\tilde{f}_{t,x}^{\ell(t)}(s)ds\right.$$

$$\left. -h\wedge\left(T_{\varrho(t)}-t\right)\left(\frac{T_{\varrho(t)}-t}{2\Delta}\wedge\frac{T_{\varrho(t)}-t-h/2}{\Delta}\right)x^{\ell(t)}\right|$$

$$= \left|E\int_t^{t+h\wedge T_{\varrho(t)}} \frac{T_{\varrho(t)}-s}{\Delta}\left[\int_t^s \tilde{\sigma}_{\ell(t)}^{\top}(s')\tilde{S}_I(s',T_{\ell(t)},\Delta)ds' + \int_t^s \tilde{\sigma}_{\ell(t)}^{\top}(s')dW(s')\right]ds\right|$$

$$= O(h^2);$$

$$\left|E\int_t^{t+h\wedge T_{\varrho(t)}} \frac{s-T_{\ell(t)}}{\Delta}\tilde{f}_{t,x}^{\varrho(t)}(s)ds\right.$$

$$\left. -h\wedge\left(T_{\varrho(t)}-t\right)\frac{t+h/2-T_{\ell(t)}}{\Delta}\wedge\frac{t+\Delta-T_{\ell(t)}}{2\Delta}\cdot x^{\varrho(t)}\right|$$

$$= \left| E \int_t^{t+h \wedge T_{\varrho(t)}} \frac{s - T_{\ell(t)}}{\Delta} \left[\int_t^s \tilde{\sigma}^\top_{\varrho(t)}(s') \tilde{S}_I(s', T_{\varrho(t)}, \Delta) ds' + \int_t^s \tilde{\sigma}^\top_{\varrho(t)}(s') dW(s') \right] ds \right|$$
$$= O(h^2);$$

$$\Bigg| E \int_{T_{\varrho(t)} \wedge t+h}^{t+h} \frac{T_{\varrho(t+h)} - s}{\Delta} \tilde{f}^{\ell(t+h)}_{t,x}(s) ds$$
$$- 0 \vee \left(t + h - T_{\varrho(t)}\right) \frac{T_{\varrho(t+h)} + \Delta - t - h}{\Delta} \bar{f}^{\ell(t+h)}_{t,x}(t+h) \Bigg|$$
$$= \Bigg| 0 \vee \left(t + h - T_{\varrho(t)}\right) E \left[\tilde{f}^{\ell(t+h)}_{t,x}(t+h) - \bar{f}^{\ell(t+h)}_{t,x}(t+h) \right]$$
$$- E \int_{T_{\varrho(t)} \wedge t+h}^{t+h} \frac{T_{\varrho(t+h)} - s}{\Delta} \left(\int_s^{t+h} \sigma^\top(s', T_{\ell(t+h)}, \tilde{f}^{\ell(t+h)}_{s, \tilde{f}_{t,x}(s)}) \tilde{S}_I(s', T_{\ell(t+h)}, \Delta) ds' \right.$$
$$\left. + \int_t^s \sigma^\top(s', T_{\ell(t+h)}, \tilde{f}^{\ell(t+h)}_{s, \tilde{f}_{t,x}(s)}) dW(s') ds \right) \Bigg| = O(h^2);$$

$$\Bigg| E \int_{T_{\varrho(t)} \wedge t+h}^{t+h} \frac{s - T_{\ell(t+h)}}{\Delta} \tilde{f}^{\varrho(t+h)}_{t,x}(s) ds$$
$$- 0 \vee \left(t + h - T_{\varrho(t)}\right) \frac{t + h - T_{\ell(t+h)}}{\Delta} \bar{f}^{\varrho(t+h)}_{t,x}(t+h) \Bigg|$$
$$= \Bigg| 0 \vee \left(t + h - T_{\varrho(t)}\right) E \left[\tilde{f}^{\varrho(t+h)}_{t,x}(t+h) - \bar{f}^{\varrho(t+h)}_{t,x}(t+h) \right]$$
$$- E \int_{T_{\varrho(t)} \wedge t+h}^{t+h} \frac{s - T_{\ell(t+h)}}{\Delta} \left(\int_s^{t+h} \sigma^\top(s', T_{\varrho(t+h)}, \tilde{f}^{\varrho(t+h)}_{s, \tilde{f}_{t,x}(s)}) \tilde{S}_I(s', T_{\varrho(t+h)}, \Delta) ds' \right.$$
$$\left. + \int_t^s \sigma^\top(s', T_{\varrho(t+h)}, \tilde{f}^{\varrho(t+h)}_{s, \tilde{f}_{t,x}(s)}) dW(s') ds \right) \Bigg| = O(h^2).$$

Finally, let us consider the case $m = 1$, $s = 2$. Thanks to (4.78), it is only left to analyse the following

$$\left| E \delta \tilde{Y} \delta \tilde{f}^i \right| = \left| E \delta \tilde{Y} \left[\int_t^{t+h} \tilde{\sigma}^\top_{i_j}(s) \tilde{S}_I(s, T_{i_j}, \Delta) ds + \int_t^{t+h} \tilde{\sigma}^\top_{i_j}(s) dW(s) \right] \right|.$$

We shall consider the product of each term in $\delta\tilde{Y}$ and $\delta\tilde{f}^i$ separately. We have

$$\begin{aligned}
&\Bigg|E\int_t^{t+h\wedge T_{\varrho(t)}}\frac{T_{\varrho(t)}-s}{\Delta}\tilde{f}_{t,x}^{\ell(t)}(s)ds\\
&\qquad\times\left[\int_t^{t+h}\tilde{\sigma}_{i_j}^{\top}(s)\tilde{S}_I(s,T_{i_j},\Delta)ds+\int_t^{t+h}\tilde{\sigma}_{i_j}^{\top}(s)dW(s)\right]\Bigg|\\
&\quad\leq K|x^{\ell(t)}h^2+Kh^3+E\int_t^{t+h\wedge T_{\varrho(t)}}\int_t^{t+h}\tilde{\sigma}_{\ell(t)}^{\top}(s)\tilde{\sigma}_{i_j}(s)dW(s)|\leq K|1+x^{\ell(t)}|h^2;
\end{aligned}$$

$$\begin{aligned}
&\Bigg|E\int_t^{t+h\wedge T_{\varrho(t)}}\frac{s-T_{\ell(t)}}{\Delta}\tilde{f}_{t,x}^{\varrho(t)}(s)ds\\
&\qquad\times\left[\int_t^{t+h}\tilde{\sigma}_{i_j}^{\top}(s)\tilde{S}_I(s,T_{i_j},\Delta)ds+\int_t^{t+h}\tilde{\sigma}_{i_j}^{\top}(s)dW(s)\right]\Bigg|\\
&\quad\leq K|x^{\varrho(t)}h^2+Kh^3+E\int_t^{t+h\wedge T_{\varrho(t)}}\int_t^{t+h}\tilde{\sigma}_{\varrho(t)}^{\top}(s)\tilde{\sigma}_{i_j}(s)dW(s)|\leq K|1+x^{\varrho(t)}|h^2;
\end{aligned}$$

$$\begin{aligned}
&\Bigg|E\int_{T_{\varrho(t)}\wedge t+h}^{t+h}\frac{T_{\varrho(t+h)}-s}{\Delta}\tilde{f}_{t,x}^{\ell(t+h)}(s)ds\\
&\quad\times\left[\int_t^{t+h}\tilde{\sigma}_{i_j}^{\top}(s)\tilde{S}_I(s,T_{i_j},\Delta)ds+\int_t^{t+h}\tilde{\sigma}_{i_j}^{\top}(s)dW(s)\right]\Bigg|\leq K|1+x^{\ell(t+h)}|h^2;
\end{aligned}$$

$$\begin{aligned}
&\Bigg|E\int_{T_{\varrho(t)}\wedge t+h}^{t+h}\frac{s-T_{\ell(t+h)}}{\Delta}\tilde{f}_{t,x}^{\varrho(t+h)}(s)ds\\
&\quad\times\left[\int_t^{t+h}\tilde{\sigma}_{i_j}^{\top}(s)\tilde{S}_I(s,T_{i_j},\Delta)ds+\int_t^{t+h}\tilde{\sigma}_{i_j}^{\top}(s)dW(s)\right]\Bigg|\leq K|1+x^{\varrho(t+h)}|h^2.
\end{aligned}$$

Using a similar approach and (4.78), (4.71) for the remaining cases, $m=1$, $s=3$; $m=2$, $s=2,3$; $m=3$, $s=3$ can be checked.

For Algorithm 3.5.1 and Algorithm 3.5.3, Assumptions 3.4.1, 3.4.5 and 3.4.6 can be checked via an analogous routine as we described in this Section.

4.4 Summary

This Chapter is devoted to proving and establishing convergence properties for the numerical methods from Chapter 3. We begin this Chapter with Section 4.1, in which the convergence theorems for the methods constructed in Section 3.4 are proved. More specifically, in Section 4.1.1 we prove convergence results for the maturity time approximations of the HJM model and the pricing problem under consideration. Then, in Section 4.1.2 we establish weak convergence of fully discrete methods to the approximations discrete in the maturity time. In spirit of the method of lines, this convergence is proved to be uniform in the maturity time discretization step Δ. Convergence results for the mean-square method defined in Section 3.6 are considered in Section 4.2. We finish this Chapter by showing that the assumptions imposed on the generic numerical method in Section 3.4 hold for the algorithms presented in Section 3.5.

Chapter 5

Conclusions and outlook

5.1 Conclusions

In this section we present the main conclusions from the preceding chapters.

The background to our endeavours

One of the most general platforms in the interest rate theory is the celebrated HJM framework [36] which models the entire forward curve directly. This is a very broad setup which covers all arbitrage-free interest rate models driven by a finite number of Brownian motions. More specifically, the representatives of the HJM framework are the popular LIBOR market models and short rate models The HJM model is mathematically described via an infinite-dimensional multifactor stochastic differential equation taking the entire forward rate curve as a state variable. Under no-arbitrage conditions, the HJM model is fully characterized by specifying the forward rate volatility process and the initial forward rate curve.

The original HJM framework is used for modelling fixed income markets (see [36, 15, 18, 28] and also references therein). Recently, the HJM philosophy has been extended to credit and equity markets (see, e.g. the recent review [16]) and modelling of mortality [5] and of financial electricity contracts [8].

The HJM framework possesses many interesting and powerful properties (see Section 2.2), though its mathematical description is rather complex. The numerical approximation of the HJM model with stochastic volatility functions remains a challenging task. As far as we know, the literature on numerics for the HJM model is rather sparse.

Novel class of numerical methods for HJM framework

In this thesis we proposed a novel class of numerical algorithms for the HJM model together with a rigorous numerical analysis. The idea of the method of lines served as inspiration for our approach. The proposed methods facilitate simulation of the HJM model under various specifications. The main focus of our research was put on the weak-sense numerical methods which can be used for valuing a broad range of interest rate products. The numerical methods were constructed with the following guidance. We first discretized the infinite-dimensional HJM equation in maturity time variable T by approximating the arbitrage-free drift with quadrature rules. As a result, we obtained a finite-dimensional system of stochastic differential equations (SDEs). This system played an intermediate role in our considerations. It was used as a guidance to construct the fully discrete numerical methods and in the proofs of the convergence results. We showed in this thesis and in [48] that if we take a quadrature rule of order p, the solution of this finite-dimensional system of SDEs converges to the HJM solution with mean-square order p in the maturity time discretization step Δ. The fully discrete methods (discrete in both T and t), were obtained by approximating the finite-dimensional system of SDEs in the weak and mean-square senses using the general theory of numerical integration of SDEs (see, e.g. [57, 58, 45]). We proved in this thesis and in [48] that if we take a mean-square numerical integrator of order q, the solution of this fully discrete method converges to the solution of the HJM approximation discrete in the maturity time T only with mean-square order q in the calendar time discretization step h. For illustration of our weak-sense numerical methods for the HJM model we considered approximation of the pricing problem of a generic interest rate contract of European-type, which covers a broad range of popular derivatives. We analyzed weak convergence of fully discrete methods to the approximations discrete in the maturity time. We showed that this convergence is uniform in the maturity time discretization step Δ in order to obtain weak convergence of the fully-discrete numerical methods to the solution of the HJM equation.

The introduced class of numerical methods was illustrated by presenting some particular algorithms of various accuracy orders, which are ready for implementation. We tested the proposed numerical algorithms on pricing European-type caps with the Vasicek and proportional volatility models for forward rates. The results of the numerical tests confirmed our theoretical conclusions that, within the proposed class, the numerical methods possess both computational efficiency and flexibility. The computational efficiency is due to the use of high-order quadrature rules which permits us to take large discretization steps in the maturity time without affecting overall accuracy of the algorithms. More precisely, the number of forward rates that need to be approximated at each time moment t are significantly less in our algorithms than what is usually required when the time-grids for t and T coincide. As illustrated in the numerical experiments, new algorithms (e.g., Algorithm 5.3) can considerably outperform the existing algorithms with coinciding grids (with Algorithm 5.1 being a typical representative). Fast numerical algorithms are the cornerstone of efficient calibration of the HJM model. In particular, fast calibration is crucial for the model's applicability in practice. In spirit of the method of lines, the proposed class of numerical methods displayed a high degree of flexibility providing freedom in choosing appropriate approximations in "space" and "time" separately. Based on our theoretical results and demonstrated in our numerical experiments, we concluded that, it is beneficial in practice to use higher order rules for integration with respect to maturity time T and lower order numerical schemes for integration with respect to calendar time t.

5.2 Outlook for research

In this section, we shall discuss some possible directions of future research to develop the work of this thesis further.

The numerical experiments in this thesis and in [48] demonstrated that new algorithms can considerably outperform the existing ones. Within the proposed approach, even more computationally efficient algorithms based on second order numerical integrators (see Remark 3.5.4) and/or multi-level Monte Carlo method [30] can be derived. From a financial engineering point of view, this means, in particular, that calibration of the HJM model can be done faster,which makes the HJM model more attractive to practitioners.

The proposed numerical approximation of the HJM model are tested on pricing European-type interest rate derivatives. Combining the developed numerical algorithms with, for instance, regression (see, e.g., [51] and also [32]), one can construct numerical procedures for pricing interest rate contracts of American and Bermudian-types which requires further study. The other extensions of the results presented in this thesis and [48] include computing Greeks (see algorithms for evaluation Greeks in the case of equity markets in, e.g., [32, 59] and in the references therein).

Though the original HJM framework is used for modelling fixed income markets. As we mentioned before, the applicability of HJM model goes far beyond interest rate modelling (see, e.g. the recent review [16]). Potentially, the ideas we developed in this thesis could be useful for constructing numerical approximation in these applications.

In addition, the HJM model can be transformed into a first-order hyperbolic SPDE using the Musiela parameterization and numerical algorithms exploiting SPDE solvers can be considered. Let us also note that in general not much attention has been paid yet to weak-sense numerical approximations of infinite-dimensional stochastic equations, including SPDEs, with exception of, e.g. [75, 13, 22, 21]. The approach to numerical analysis of weak methods for the HJM model considered in this thesis has a potential to be exploited for weak approximations of SPDEs.

Another challenging direction of research is to develop numerical algorithms for pricing interest rate barrier options within HJM framework. At the moment this problem is the main focus of our research.

Appendix A

Proofs of selected expressions

In this Appendix derivation of some expressions from Chapter 3 are presented.

Proof of expression (3.47): To show that (3.47) holds it is sufficient to demonstrate the following is true:

$$E\exp(|\int_{t_0}^{t}\sigma^{\top}(s,T)dW^{\mathrm{Q}}(s)|) \leq C. \tag{A.1}$$

We note that

$$\begin{aligned}
&\exp(|\int_{t_0}^{t}\sigma^{\top}(s,T)dW^{\mathrm{Q}}(s)|) \qquad (A.2)\\
\leq\ &\exp(\int_{t_0}^{t}\sigma^{\top}(s,T)dW^{\mathrm{Q}}(s)) + \exp(-\int_{t_0}^{t}\sigma^{\top}(s,T)dW^{\mathrm{Q}}(s)).
\end{aligned}$$

Thanks to the boundness of σ (3.45), we have

$$\int_{t_0}^{t}\sigma^{\top}(s,T)\sigma(s,T)ds) \leq C. \tag{A.3}$$

Using (A.3), we obtain

$$\begin{aligned}
&E^{\mathrm{Q}}\exp\left(\int_{t_0}^{t}\sigma^{\top}(s,T)dW^{\mathrm{Q}}(s)\right) \qquad (A.4)\\
\leq\ &CE^{\mathrm{Q}}\left[\exp\left(\int_{t_0}^{t}\sigma^{\top}(s,T)dW^{\mathrm{Q}}(s) - \frac{1}{2}\int_{t_0}^{t}\sigma^{\top}(s,T)\sigma(s,T)ds\right)\right],
\end{aligned}$$

where $C > 0$ is a constant.

Also, from (A.3), the Novikov condition

$$E^{\mathrm{Q}} \exp\left(\frac{1}{2}\int_{t_0}^{t^*} \sigma^\top(s,T)\sigma(s,T)ds\right) < \infty,$$

is evidently satisfied. This ensures that the expression under expectation operator in (A.4) is a martingale with initial value 1. Hence, we conclude that

$$E^{\mathrm{Q}} \exp(\int_{t_0}^{t} \sigma^\top(s,T)dW^{\mathrm{Q}}(s)) \le C.$$

Analogously, we can show that

$$E^{\mathrm{Q}} \exp(-\int_{t_0}^{t} \sigma^\top(s,T)dW^{\mathrm{Q}}(s)) \le C.$$

The last two estimates yield (A.1) and, hence, show that (3.47) holds.

Proof of expression (3.83): Recall that (cf. (3.67), (3.82))

$$\begin{aligned} \tilde{f}^i_{t,x}(s) &= x^i + \int_t^s \tilde{\sigma}_i^\top(u)\tilde{S}_I(u,T_i,\Delta)du + \int_t^s \tilde{\sigma}_i^\top(u)dW(u), \\ t_0 &\le t \le t^* \wedge T_{(i+1)\wedge N},\ i = 0,\dots,N, \end{aligned} \tag{A.4}$$

$$\tilde{Y}_{t,x,0}(t+h) = \sum_{l=\ell(t)}^{\ell(t+h)} \int_t^{t+h} \sum_{i=0}^{\theta} \lambda_i(s)\tilde{f}^{l+i}_{t,x}(s)\chi_{s\in[T_l,T_{l+1})}ds. \tag{A.5}$$

We shall show:

$$E\left|\tilde{Y}_{t,x,0}(t+h)\right|^m \le Ch^m\left(1 + \sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l|^m\right). \tag{A.6}$$

Let us first consider the case $m = 1$. By (A.4) and (A.5), we have

$$\begin{aligned} & E\left|\tilde{Y}_{t,x,0}(t+h)\right| \\ = \;& E\left|\sum_{l=\ell(t)}^{\ell(t+h)} \int_t^{t+h} \sum_{i=0}^{\theta} \lambda_i(s)\tilde{f}^{l+i}_{t,x}(s)\chi_{s\in[T_l,T_{l+1})}ds\right| \end{aligned}$$

$$
\begin{aligned}
&\leq \; C \sum_{l=\ell(t)}^{\ell(t+h)} \int_t^{t+h} \sum_{i=0}^{\theta} E \left| \left[x^{l+i} + \int_t^s \tilde{\sigma}_{l+i}^{\top}(u) \tilde{S}_I(u, T_{l+i}, \Delta) du \right. \right. \\
&\qquad \left. \left. + \int_t^s \tilde{\sigma}_{l+i}^{\top}(u) dW(u) \right] \chi_{s \in [T_l, T_{l+1})} ds \right| \\
&\leq \; C \sum_{l=\ell(t)}^{\ell(t+h)} \sum_{i=0}^{\theta} \int_t^{t+h} \left[\left| x^{l+i} \right| + h + h^{\frac{1}{2}} \right] ds \leq Ch \left(1 + \sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l| \right),
\end{aligned}
$$

where $C > 0$ is a constant independent of Δ and x. To obtain the last inequality, we use the condition that $h \leq \alpha \Delta$ for some $\alpha > 0$ which ensures, that the number $\ell(t+h) - \ell(t)$ is independent of Δ.

Next, we shall consider the case $m = 2$. Using the Cauchy–Bunyakovsky inequality, we have

$$
\begin{aligned}
& E \left| \tilde{Y}_{t,x,0}^2(t+h) \right| \\
&\leq CE \left| \sum_{l=\ell(t)}^{\ell(t+h)} \sum_{k=\ell(t)}^{\ell(t+h)} \int_t^{t+h} \int_t^{t+h} \sum_{i=0}^{\theta} \sum_{j=0}^{\theta} \tilde{f}_{t,x}^{l+i}(s) \tilde{f}_{t,x}^{k+j}(s') \chi_{s \in [T_l, T_{l+1})} \chi_{s' \in [T_k, T_{k+1})} ds ds' \right| \\
&\leq CE \left| \sum_{l=\ell(t)}^{\ell(t+h)} \sum_{k=\ell(t)}^{\ell(t+h)} \int_t^{t+h} \int_t^{t+h} \sum_{i=0}^{\theta} \sum_{j=0}^{\theta} \left[\left(\tilde{f}_{t,x}^{l+i}(s) \right)^2 \right. \right. \\
&\qquad \left. \left. + \left(\tilde{f}_{t,x}^{l+j}(s) \right)^2 \right] \chi_{s \in [T_l, T_{l+1})} \chi_{s' \in [T_k, T_{k+1})} ds ds' \right| \\
&\leq \sum_{l=\ell(t)}^{\ell(t+h)} \sum_{k=\ell(t)}^{\ell(t+h)} \int_t^{t+h} \int_t^{t+h} \sum_{i=0}^{\theta} \sum_{j=0}^{\theta} \left[\left| x^{l+i} \right|^2 + h^2 + h \right] ds ds' \leq Ch^2 \left(1 + \sum_{l=\ell(t)}^{\ell(t+h)+\theta} |x^l|^2 \right).
\end{aligned}
$$

Using the similar reasoning, we can prove (A.6) for general m.

Appendix B

Stochastic Calculus

In this Appendix we give a short overview of the tools from stochastic calculus frequently used in this thesis. For more extensive background on the topics covered in this Appendix see e.g. [9, 41, 44, 63, 67, 76].

We assume that we are working on a filtered probability space $(\Omega, \mathcal{F}, \mathrm{P}, \{\mathcal{F}_t\}_{t\geq 0})$ satisfying the usual conditions:

1. $\mathcal{F}$ is P-complete: if $B \subset A \in \mathcal{F}$ and $\mathrm{P}(A) = 0$ then $B \in \mathcal{F}$.

2. $\mathcal{F}_0$ contains all P-null sets;

3. The filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is *right continuous*, i.e. $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all $t \geq 0$.

A *filtration* is a family$\{\mathcal{F}_t\}_{t\geq 0}$ of increasing sub-σ-algebras, i.e. $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for all $0 \leq s < t < \infty$.

A n-dimensional function $X : \Omega \to \mathbb{R}^n$, $n \in \mathbb{N}$ is called $\mathcal{F}-$measurable if $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$ for all $B \in \mathcal{B}(\mathbb{R}^n)$, where $\mathcal{B}(\mathbb{R}^n)$ is the Borel sigma-algebra in $\mathbb{R}^n$. A n-dimensional measurable function $X : \Omega \to \mathbb{R}^n$, $n \in \mathbb{N}$ is called an $\mathbb{R}^n-$valued random variable.

If a random variable $X : \Omega \to \mathbb{R}$ is integrable with respect to the probability measure P, then the number $E^{\mathrm{P}}[X] = \int_\Omega X(\omega) d\mathrm{P}(\omega)$ is called the expectation of X with respect to the measure P.

The conditional expectation of an integrable random variable X given a sub-sigma algebra $\mathcal{G} \subset \mathcal{F}$ is defined to be a $\mathcal{G}-$measurable function $E^{\mathrm{P}}[X|\,\mathcal{G}]$ with

$$\int_G X(\omega)d\mathrm{P}(\omega) = \int_G E^{\mathrm{P}}[X|\,\mathcal{G}]\,d\mathrm{P}(\omega) \text{ for all } G \in \mathcal{G}.$$

For background on conditional expectation and its properties see [9, 76].

Definition B.1 (Stochastic process) *A stochastic process (vector process) is a family $\{X(t)\}_{t\geq 0}$ of random variables (vectors) defined on the filtered probability space $(\Omega, \mathcal{F}, \mathrm{P}, \{\mathcal{F}_t\}_{t\geq 0})$ and indexed by some set I. In our setting, I is taken as $[t_0, t^*]$ for some $t^* > t_0 \geq 0$. We say that*

(a) $\{X(t)\}_{t\geq 0}$ is *adapted* (to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$) if X is $\mathcal{F}_t$-measurable for every $t \geq 0$.

(b) $\{X(t)\}_{t\geq 0}$ is *measurable* if the stochastic process regarded as a function of two variables (t, ω) from $[0, \infty) \times \Omega$ to $\mathbb{R}^n$, $n \in \mathbb{N}$ is $\mathcal{B}([0,\infty)) \otimes \mathcal{F}$-measurable with $\mathcal{B}([0,\infty)) \otimes \mathcal{F}$ denoting the product sigma-algebra created by $\mathcal{B}([0,\infty))$ and $\mathcal{F}$, i.e. the smallest sigma algebra which contains all sets $G_1 \times .G_2 \in \mathcal{B}([0,\infty)) \times \mathcal{F}$.

(c) $\{X(t)\}_{t\geq 0}$ is *progressively measurable* if the mapping $(t,\omega) : [0, T] \times \Omega \to \mathbb{R}^n$, $n \in \mathbb{N}$ is $\mathcal{B}([0, T]) \otimes \mathcal{F}_T$-measurable.

Note, that each progressively measurable stochastic process is also adapted. Moreover, if stochastic process is adapted and all its paths are right-continuous then it is progressively measurable.

Let us now define a special stochastic process which may be considered as one of the atoms of modern finance.

Definition B.2 (Brownian motion) *A* one-dimensional (P$-$) Brownian motion *is a real-valued (*P*-a.s) continuous $\mathcal{F}_t$-adapted process $W(t) = \{W(t)\}_{t\geq 0}$ with the following properties*

1. $W(0) = 0$ P-*a.s.;*

2. *for* $0 \leq s < t < \infty$, *under* P, *the increment* $W(t) - W(s)$ *is normally distributed with mean zero and variance* $t - s$;

3. *for* $0 \leq s < t < \infty$, *the increment* $W(t) - W(s)$ *is independent of* $\mathcal{F}_s$.

A d-dimensional process $W(t) = (W_1(t), \ldots, W_d(t))$ defines a *d-dimensional Brownian motion* if every $W_i(t)$ is a one-dimensional Brownian motion, and $W_1(t), \ldots, W_d(t)$ are independent.

We note that we will write, when it is necessary, $W^{\mathrm{P}}(t)$ for Brownian motion with respect to measure P.

We shall always assume that Brownian motions are defined relative to the filtration given in the definition of the underlying filtered probability space.

One of the important concepts we need for modelling in finance is that of martingales.

Definition B.3 *An* $\mathbb{R}^n-$*valued* $\{\mathcal{F}_t\}_{t\geq 0}-$*adapted integrable stochastic process* $\{X(t)\}_{t\geq 0}$ *is called a martingale with respect to* $\{\mathcal{F}_t\}_{t\geq 0}$ *and the measure* P *if*

$$E^{\mathrm{P}}\left[X(t)|\,\mathcal{F}_s\right] = X(s) \quad \mathrm{P} - a.s \text{ for all } 0 \leq s < t < \infty.$$

Note that every martingale has a cadlag modification, i.e. another stochastic process $\{Y(t)\}_{t\geq 0}$ which is right continuous and has left limits and such that $\mathrm{P}(\omega : X(t,\omega) = Y(t,\omega) = 1)$.

Let $0 \leq t \leq t^*$. We shall now define the Itô integral

$$\int_0^t X(s)dW(s) \tag{B1}$$

for any stochastic process $X = \{X(t)\}_{0\leq t\leq t^*}$ in the space of all real-valued progressively measurable stochastic processes equipped with the norm

$$|X|^2 = \mathbb{E}\int_0^t |X(t)|^2 dt < \infty. \tag{B2}$$

To define this Itô integral, we first define the integral $\int_0^t H(s)dW(s)$ for a class of simple processes H. Then we show that each progressively measurable process X with $|X| < \infty$ can be approximated by such simple processes H's and we define the limit of $\int_0^t H(s)dW(s)$ as the integral $\int_0^t X(s)dW(s)$.

Definition B.4 (Simple processes) *A real-valued stochastic process* $H = \{H(t)\}_{0\le t\le t^*}$ *is called a* simple *process if there exists a partition* $0 = t_0 < t_1 < \ldots < t_n = t^*$ *and bounded random variables* ξ_i, $0 \le i \le n-1$, *such that* ξ_i *is* $\mathcal{F}_{t_i}$*-measurable and*

$$H(t) = \xi_0 \boldsymbol{\chi}_{[t_0,t_1]}(t) + \sum_{i=1}^{n-1} \xi_i \boldsymbol{\chi}_{[t_i,t_{i+1}]}(t), \quad t \in [0, t^*]. \tag{B3}$$

In the above definition, $\boldsymbol{\chi}_{[A]}(t)$ is the indicator function of the set A.

Definition B.5 (Itô integral with simple integrands) *If* H *is a simple process with the form* (B3), *the* Itô integral, *of* H *with respect to the Brownian motion* W *is the process defined by*

$$\int_0^t H(s)dW(s) := \sum_{i=0}^{n} \xi_i (W(t_{i+1}) - W(t_i)), \quad t \in [0, t^*].$$

Definition B.6 (Itô integral) *Let be* X *a progressively measurable process satisfying* $|X| < \infty$. *The Itô integral of* X *with respect to the Brownian motion* W *is the process defined by*

$$\int_0^t X(s)dW(s) := \lim_{n\to\infty} \int_0^t H_n(s)dW(s) \quad in\ L^2(\Omega, \mathbb{R}), \quad t \in [0, t^*],$$

where $\{H_n\}$ *is a sequence of simple processes such that*

$$\lim_{n\to\infty} \mathbb{E} \int_0^t |X(s) - H_n(s)|^2 ds = 0.$$

We shall now give some of the properties of the Itô integral.

Proposition B.7 *Let be* X, Y *a progressively measurable processes satisfying* $|X| < \infty$, $|Y| < \infty$. *Then*

(a) $\int_0^t X(s)dW(s)$ *is* $\mathcal{F}_t-$*measurable;*

(b) $\mathbb{E}\int_0^t X(s)dW(s) = 0$

(c) Itô isometry*:*

$$\mathbb{E}\left|\int_0^t X(s)dW(s)\right|^2 = \mathbb{E}\int_0^t |X(s)|^2\, ds; \tag{B5}$$

(c) for $\alpha,\, \beta \in \mathbb{R}$, $\int_0^t [\alpha X(s) + \beta Y(s)]\, dW(s) = \alpha\int_0^t X(s)dW(s) + \beta\int_0^t Y(s)dW(s)$.

Definition B.8 *An n-dimensional Itô process is an* $\mathbb{R}^n$*-valued continuous adapted process of the form*

$$X(t) = X(0) + \int\limits_{t_0}^{t} a(s, X(s))ds + \sum_{i=1}^{d}\int\limits_{t_0}^{t} b_i(s, X(s))dW_i^{\mathrm{P}}(s), \tag{B6}$$

where $W^{\mathrm{P}}(t) = \left(W_1^{\mathrm{P}}(t), \ldots, W_d^{\mathrm{P}}(t)\right)^\top$ *is a d-dimensional standard* $\mathrm{P}-$ *Wiener process;* a *is* $\mathbb{R}^n-$*valued adapted process with*$\int_{t_0}^t |a(s, X(s))|\, ds < \infty$ *a.s. for any* $t \in [t_0, t^*]$; b_i *are* $\mathbb{R}^n-$*valued adapted processes with* $\int_{t_0}^t |b_i(s, X(s))|^2\, ds < \infty$, $i = 1, \ldots, d$ *a.s. for any* $t \in [t_0, t^*]$. *We shall say that* $X(t)$ *has stochastic differential* $dX(t)$ *given by*

$$dX(t) = a(t, X(t))dt + \sum_{i=1}^{d} b_i(t, X(t))dW_i^{\mathrm{P}}(t), \tag{B7}$$

Proposition B.9 (Itô formula) *Let* $X(t)$ *be n-dimensional Itô process with the stochastic differential (B7). Let* $f : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}$, $(t, x) \to f(t, x)$ *be a function once differentiable with respect to* t *and twice with respect to* x. *Then*

$$f(t, X(t))$$

is again an Itô process with stochastic differential of the form

$$df(t, X(t)) = Lf(t, X(t))dt + \sum_{i=1}^{d} \Lambda_i f(t, X(t))dW_i^{\mathrm{P}}(t), \tag{B8}$$

where the operators L and Λ_i, $i = 1, \ldots, d$ are given by

$$L = \frac{\partial}{\partial t} + a^\top \frac{\partial}{\partial x} + \frac{1}{2} \sum_{i=1}^{d} \sum_{m=1}^{n} \sum_{j=1}^{n} b_i^m b_i^j \frac{\partial^2}{\partial x_m \partial x_j},$$

$$\Lambda_i = \sum_{j=1}^{n} b_i^j \frac{\partial}{\partial x_j}.$$

The Itô formula is the chain rule of stochastic calculus. It differs from the corresponding result in ordinary calculus through the appearance of second order derivatives in dt term.

We shall now give some results from the change of measure technique.

Definition B.10 (Equivalent measures) *Let* P *and Q be two measures defined on the same measurable space $(\Omega, \mathcal{F})$. We say that Q is absolutely continuous with respect to* P, *written $Q \ll \mathrm{P}$, if $Q(G) = 0$ whenever $P(G) = 0$, $G \in \mathcal{F}$. If both $Q \ll \mathrm{P}$ and $P \ll \mathrm{Q}$, we call* P *and Q equivalent measures and denote this by $Q \sim \mathrm{P}$.*

Theorem B.11 (Radon-Nikodym) *Let* P *and Q be measures on the measurable space $(\Omega, \mathcal{F})$. Then $Q \ll \mathrm{P}$ if and only if there exist an integrable function $g \geqslant 0$ Q-a.s. such that*

$$\mathrm{Q}(G) = \int_G g d\mathrm{P} \quad \textit{for all } G \in \mathcal{F}.$$

g is called the Radon-Nikodym derivative of Q with respect to P *and it is also written as $g = dQ/d\mathrm{P}$.*

For calculating conditional expectations it can be needed to change the point of view and use a different measure. The following theorem gives us a relation between conditional expectations with respect to different measures.

Theorem B.12 (Bayes formula) *Let* P *and Q be two measures defined on the same measurable space $(\Omega, \mathcal{F})$ and let $dQ/d\mathrm{P}$ be the Radon-Nikodym derivative of Q with respect to* P. *Furthermore, let X be an integrable random variable on the probability space $(\Omega, \mathcal{F}, Q)$ and $\mathcal{G} \subset \mathcal{F}$ be a sub-sigma algebra of $\mathcal{F}$. Then the following* *holds:*

$$E^{\mathrm{P}}\left[X\cdot\frac{d\mathrm{Q}}{d\mathrm{P}}\bigg|\,\mathcal{G}\right]=E^{\mathrm{Q}}\left[X|\,\mathcal{G}\right]\cdot E^{\mathrm{P}}\left[\frac{d\mathrm{Q}}{d\mathrm{P}}\bigg|\,\mathcal{G}\right].$$

Theorem B.13 (Girsanov theorem) *Let X be an n-dimensional Itô process introduced in Definition B.8. Suppose there exist an $\mathbb{R}^d$-valued $\mathcal{F}_t$-adapted process $\alpha(t)$ with $\int_0^t |\alpha(s)|^2 ds < \infty$ a.s. for every $t \in [0, t^*]$ and also an $\mathbb{R}^n$-valued $\mathcal{F}_t$-adapted process $v(t, X(t))$ with $\int_0^t |v(s, X(s))|ds < \infty$ such that*

$$b(t,X(t))\alpha(t)=a(t,X(t))-v(t,X(t))$$

and assume that $\alpha(t)$ satisfies the Novikov condition

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \alpha^2(s)ds\right)\right]<\infty. \tag{B9}$$

Define

$$\xi_t=\exp\left(-\int_0^t \alpha^{\intercal}(s)dW^{\mathrm{P}}(s)-\frac{1}{2}\int_0^t \alpha^2(s)ds\right),\quad t\le t^*, \tag{B10}$$

and

$$d\mathrm{Q}=\xi_T d\mathrm{P}\quad on\ \ \mathcal{F}_T.$$

Then

$$W^{\mathrm{Q}}(t)=\int_0^t \alpha(s)ds+W^{\mathrm{P}}(t),\quad t\le t^*, \tag{B11}$$

is a Brownian motion with respect to the probability measure Q and the process X has the stochastic differential of the form

$$dX(t)=v(t,X(t))dt+b(t,X(t))dW^{\mathrm{Q}}(t) \tag{B12}$$

in terms of $\tilde{W}(t)$.

We shall finally state the Fubini's theorem for stochastic integrals. The proof of this theorem can be found for e.g. in [26, pp. 99].

Theorem B.14 (Fubini's theorem for stochastic integrals) *Consider the $\mathbb{R}^d$-*

valued stochastic process $\phi = \phi(\omega, t, s)$ *with two indices,* $0 \leq t,\ s \leq T$*, satisfying the following properties*

1. ϕ *is progressively measurable;*

2. $\sup_{t,s} \|\phi(t, s)\| < \infty$.

Then

$$\int_0^T \left(\int_0^T \phi(t, s) dW(t) \right) ds = \int_0^T \left(\int_0^T \phi(t, s) ds \right) dW(t). \tag{B13}$$

Bibliography

[1] L. Andersen, V. Piterbarg. *Interest Rate Modeling.* Volumes 1-3. Atlantic Financial Press, 2010.

[2] L. Andersen, V. Piterbarg, J. Sidenius. A New Framework for Dynamic Credit Portfolio Loss Modeling. *Working paper*, 2005.

[3] L. Arnold. *Stochastic Differential Equations: Theory and Applications.* Wiley, 1974.

[4] V. Bally, D. Talay. The law of the Euler scheme for stochastic differential equations: I. Convergence rate of the distribution function. *Prob. Theory Rel. Fields*, **104** (1996), 43-60.

[5] D. Bauer, F.E. Benth, R. Kiesel. Modeling the forward surface of mortality. Preprint, Georgia State University, 2010.

[6] M. Baxter, A. Rennie. *Financial Calculus.* Cambridge University Press, 1996.

[7] N. Bennani. The forward loss model: a dynamic term structure approach for the pricing of portfolios of credit derivatives. Tech. report, November 2005.

[8] F.E. Benth, S. Koekebakker. Stochastic modeling of financial electricity contracts. *Energy Economics*, **30** (2008), 1116-1157.

[9] N. Bingham, R. Kiesel. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives.* Springer, 2000.

[10] T. Bjork. *Arbitrage Theory in Continuous Time.* Oxford University Press, 2009.

[11] T. Bjork, Y. Kabanov, and W. Runggaldier. Bond market structure in the presence of marked point processes. *Math. Finance*, **7**(1997), 211–239.

[12] T. Bjork, A. Szepessi, R. Tempone, G. Zouraris. Monte Carlo Euler approximation of HJM term structure financial models. *Working paper*, 2002.

[13] A. de Bouard, A. Debussche. Weak and strong order of convergence of a semi discrete scheme for the stochastic Nonlinear Schrödinger equation. *J. Appl. Maths Optim.*, **54** (2006), 369-399.

[14] A. Brace, D. Gatarek, M. Musiela. The market model of interest rate dynamics. *Math. Finance,* **7**(1997), 127–155

[15] D. Brigo, F. Mercurio. *Interest Rate Models: Theory and Practice.* Springer, 2006.

[16] R. Carmona. HJM: A unified approach to dynamic models for fixed income, credit and equity markets. *Lecture Notes in Mathematics* **1919**, Springer, 2007, pp. 3-45.

[17] R. Carmona, S. Nadtochiy. HJM dynamics for equity models, Tech. report, 2007.

[18] R. Carmona, M. Tehranchi. *Interest Rate Models: An Infinite-Dimensional Stochastic Analysis Perspective.* Springer, 2006.

[19] P. Davis, P. Rabinowitz. *Methods of Numerical Integration.* Dover, 2009.

[20] F. Delbaen, W.Schachermayer. A General Version of the Fundamental Theorem of Asset Pricing. *Mathematische Annalen* **300**(1994), 463-520.

[21] A. Debussche. Weak approximation of stochastic partial differential equations: the nonlinear case. *Math.Comp.*, **80** (2011), 89-117.

[22] P. Dörsek, J. Teichmann: A semigroup point of view on splitting schemes for stochastic (partial) differential equations. Preprint arXiv/1011.2651, 2010.

[23] P. Dörsek, J. Teichmann: Efficient simulation and calibration of general HJM models by splitting schemes. Preprint arXiv/:1112.5330, 2011.

[24] E. Eberlein, S. Raible. Term structure models driven by general Lévy processes. *Math. Finance* **9**(1999), 31–53.

[25] D. Filipovic. *Consistency problems for Heath-Jarrow-Morton interest rate models*. Springer, 2001.

[26] D. Filipovic. *Term-Structure Models: A Graduate Course*. Springer, 2009.

[27] H. Geman, N. El-Karoui, J.C. Rochet. Change of numeraire, change of probability measure, and option pricing. *Journal of Applied Probability* **32** (1995), 443-458.

[28] R. Gibson, F-S. Lhabitant, D. Talay. Modeling the term structure of interest rates: a survey. *Foundations and Trends in Finance* **5** (2010), 1–156.

[29] I.I. Gihman, A.V. Skorohod. *Stochastic Differential Equations*. Springer, 1972.

[30] M. B. Giles. Multi-level Monte Carlo path simulation. *Operations Research* **56** (2008), 981–986.

[31] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2003.

[32] P. Glasserman, P. Heidelberger, P. Shahabuddin. Importance sampling in the Heath-Jarrow-Morton framework. *J. Derivatives* **7** (1999), 32–50.

[33] A. Gupta, M.G. Subrahmanyam. Pricing and hedging interest rate options: evidence from cap-floor markets. *J. Banking and Finance* **29** (2005), 701-733.

[34] J. M. Harrison, S. R. Pliska. A stochastic calculus model of continuous trading: Complete markets. *Stoch. Proc. Applic.* **15** (1983), 313-316.

[35] D. Heath, R. Jarrow, A. Morton. Bond pricing and the term structure of interest rates: a discrete time approximation. *J. Finan. Quant. Anal.* **25** (1990), 419-440.

[36] D. Heath, R. Jarrow, A. Morton. Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica* **60** (1992), 77-105.

[37] P. Hunt, J. Kennedy. *Financial derivatives in theory and practice.* Wiley, 2000.

[38] F. Jamshidian. An exact bond option pricing formula. *J. Finance* **44** (1989), 205-209.

[39] F. Jamshidian. Bond and Options Evaluation in the Gaussian Interest Rate Model. *Res. Finance* **9** (1991), 131-170.

[40] R. Jarrow. *Modeling Fixed Income Securities and Interest Rate Options.* Stanford University Press, 2002.

[41] X. Mao. *Stochastic Differential Equations and Applications.* Horwood Publications, 2008.

[42] K. Miltersen, K. Sandmann, D. Sondermann. Closed form solutions for term structure derivatives with lognormal interest rates. *J. Finance* **52**(1997), 409–430.

[43] M. Musiela, M.Rutkowski. *Martingale Methods in Financial Modelling*, Spinger, 1998.

[44] I. Karatzas, S.E. Shreve. *Brownian Motion and Stochastic Calculus.* Springer, 1991.

[45] P. Kloeden, E. Platen. *The Numerical Solution of Stochastic Differential Equations.* Springer, 1999.

[46] A. N. Kolmogorov, S. V. Fomin. *Elements of the Theory of Functions and Functional Analysis.* Dover, 1957.

[47] M. Krivko. A hybrid model for plastic card fraud detection systems. *Expert Systems with Applications*, **37**(2010), 6070-6076.

[48] M. Krivko, M.V. Tretyakov (2011). Numerical integration of Heath-Jarrow-Morton model of interest rates. Research Report No. MA-11-001, University of Leicester, 48pp. (Preprint arXiv/1109.2557)

[49] N.V. Krylov. *Controlled Diffusion Processes*. Springer, 1980.

[50] N.V. Krylov. Approximate Calculation of Integrals. Fizmatgiz, Moscow, 1959. (English transl., Macmillan, 1962.)

[51] F.A. Longstaff , E.S. Schwartz. Valuing American options by simulation: A simple least-squares approach. *Rev. Fin. Studies* **14** (2001), 113–147.

[52] M. Marcozzi. On the valuation of interest rate products under multi-factor HJM term-structures. *Appl. Num. Math.* **59** (2009), 2873-2890

[53] G.N. Milstein. Approximate integration of stochastic differential equations. *Theory Probab. Appl.* **19**(1974), 583-588.

[54] G.N. Milstein. A method of second order accuracy integration of stochastic differential equations. *Theory Probab. Appl.* **23**(1978), 414–419.

[55] G.N. Milstein. Weak approximation of solutions of systems of stochastic differential equations. *Theor. Prob. Appl.* **30**(1985), 706-721.

[56] G.N. Milstein. A theorem on the order of convergence of mean-square approximations of solutions of systems of stochastic differential equations. *Theor. Prob. Appl.* **33**(1987), 809-811.

[57] G.N. Milstein. *Numerical Integration of Stochastic Differential Equations* Springer, 2010. (reprint of 1st ed. 1995 edition.)

[58] G.N. Milstein, M.V. Tretyakov. *Stochastic Numerics for Mathematical Physics*. Springer, 2004.

[59] G.N. Milstein, M.V. Tretyakov. Numerical analysis of Monte Carlo evaluation of Greeks by finite differences. *J. Comp. Fin.* **8** (2005), 1–33.

[60] G.N. Milstein, M.V. Tretyakov. Practical variance reduction via regression for simulating diffusions. *SIAM J. Numer. Anal.* **47** (2009), 887-910.

[61] M. Musiela. Stochastic PDEs and term structure models. Journées Internationales de Finance, IGR-AFFi, La Boule (1993).

[62] S.K. Nawalkha, N.A. Beliaeva, G.M. Soto. *Dynamic Term Structure Modeling: The Fixed Income Valuation Course.* Wiley, 2007.

[63] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications.* Springer, 2002.

[64] E. Pardoux, D. Talay. Discretization and simulation of stochastic differential equations. *Acta Appi. Math.* **3**(1985), 23-47.

[65] A. Pelsser. *Efficient Methods for Valuing Interest Rate Derivatives.* Springer, 2000.

[66] W. Press, S. Teukolsky, W. Vetterling, B. Flannery. *Numerical Recipes: The Art of Scientific Computing.* Cambridge University Press, 3 edition, 2007.

[67] P. E. Protter. *Stochastic Integration and Differential Equations.* Springer, 2004.

[68] R. Rebonato. *Interest-rate Option Models.* Wiley, 1998.

[69] P. Ritchken, L. Sankarasubramanian. Volatility structures of forward rates and the dynamics of the term structure. *Math. Finance*, **5**(1995), 55-72.

[70] L. C. G. Rogers, D. Williams. *Diffusions, Markov Processes and Martingales. Vol 2. Itô Calculus.* Wiley, 1994.

[71] W. E. Schiesser. The Numerical Methods of Lines: Integration of Partial Differential Equations. Academic Press Inc., 1992.

[72] P. Schönbucher. A market model for stochastic implied volatility. Phil. Trans. of the Royal Society, **357**(1999), 2071 – 2092.

[73] P. Schönbucher. Portfolio losses and the term structure of loss transition rates: a new methodology for the pricing of portfolio credit derivatives. Tech. report, 2005.

[74] M. Schweizer, J. Wissel. Term structures of implied volatilities: Absence of arbitrage and existence results. *Math. Finance*, **18**(2008), 77–114.

[75] T. Shardlow. Numerical methods for stochastic parabolic PDEs. *Numer. Func. Anal. Optimiz.,* **20** (1999), 121-145.

[76] A. N. Shiryaev. *Probability.* Springer, 1996. (2nd edition, first published in Russian: Nauka, 1989.)

[77] H. Shirakawa. Interest rate options pricing with Poisson-Gaussian forward rate curve processes. *Math. Finance*, **1**(1991), 77–94.

[78] D. Talay. How to discretize stochastic differential equations. Lect. notes in Math. **972**(1983), 276-292.

[79] J. Wissel. Some results on strong solutions of SDEs with application to interest rate models. *Stoch. Proc. Applic.* **117** (2007), 720-741.

[80] R. Zagst. *Interest Rate Management.* Springer, 2010.