Fourier methods for numerical solution of FBSDEs
with applications in mathematical finance

submitted by

Polynice Oyono Ngou

A thesis in the department of
Mathematics and Statistics

Presented in partial fulfillment of the requirements for the degree of
Doctor in Philosophy (Mathematics) at

CONCORDIA UNIVERSITY
Montreal, Quebec, Canada

January 2014

CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Polynice OYONO NGOU

Entitled: Fourier methods for numerical solution of FBSDEs:

with applications in mathematical finance

and submitted in partial fulfillment of the requirements for the degree of

Doctor in Philosophy (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

Dr. Helena Osana (Concordia University, Education) Chair

Dr. René Ferland (UQAM, Mathematics) External Examiner

Dr. Prosper Dovonon (Concordia University, Economics) External to Program

Dr. Lea Popovic Examiner

Dr. Xiaowen Zhou Examiner

Dr. Cody B. Hyndman Thesis Supervisor

Approved by

Chair of Department or Graduate Program Director

Dean of Faculty

Abstract

Fourier methods for numerical solution of FBSDEs.

Polynice Oyono Ngou, PhD.
Concordia University, 2014.

We present a Fourier analysis approach to numerical solution of forward-backward stochastic differential equations (FBSDEs) and propose two implementations. Using the Euler time discretization for backward stochastic differential equations (BSDEs), Fourier analysis allows to express the conditional expectations included in the time discretization in terms of Fourier integrals. The space discretization of these integrals then leads to expressions involving discrete Fourier transforms (DFTs) so that the FFT algorithm can be used. We quickly presents the convolution method on a uniform space grid. Locally, this first implementation produces a truncation error, a space discretization error and an additional extrapolation error. Even if the extrapolation error is convergent in time, the resulting absolute error may be high at the boundaries of the uniform space grid. In order to solve this problem, we propose a tree-like grid for the space discretization which suppresses the extrapolation error leading to a globally convergent numerical solution for the BSDE. The method is then extended to FBSDEs with bounded coefficients, reflected FBSDEs and higher order time discretizations of FBSDEs. Numerical examples from finance illustrate its performance.

To my family...

Acknowledgments

First of all, I would like to express my gratitude to my supervisor Dr. Cody Hyndman for his patience, his useful comments and his financial support during my four year studies at Concordia University. These last years would not have been the same for me without his assistance and this work would not have been possible without his guidance.

I would like to thank the staff of the Mathematics and Statistics department at Concordia University for their continuous help. I also had the pleasure to work with several faculty members who have been good influence through their advice and understanding. I will particularly thank Dr. Jose Garrido for the trust he put in me in various situations.

My deepest appreciation goes to my committee members Dr. Lea Popovic, Dr. Wei Sun, Dr. Xiaowen Zhou and the external examiners Dr. Prosper Dovonon and Dr. René Ferland for their valuable comments. Each one of them has been an inspiration in many ways and their participation to this work is only one of them.

I shared great moments of joy with so many friends these last years. I fail to mention names for fear of forgetting some of them. I will only hope that life will continue to bless me with their kindness.

I am forever indebted to my family whose love and support have always been unconditional. My fathers, my mothers, my grand-mothers, my brothers and my sisters... you have given more than you can imagine. In all my modest achievements, from my first steps to the completion of this work, the will of making you proud has guided me.

Contents

List of Figures

List of Tables

Introduction

Since 1990, when Pardoux and Peng [96] proved the existence and uniqueness of solutions of backward stochastic differential equations (BSDEs), research on the subject has been prolific both in theory and applications. Pardoux and Peng [97] generalized the well-known Feynman-Kac formula by showing the relationship between BSDEs and quasilinear partial differential equations (PDEs) whereas Antonelli [4] introduced forward-backward stochastic differential equations (FBSDEs) and also established their well-posedness. Another article by Pardoux and Tang [98] studies FBSDEs well-posedness and generalizes the result of Pardoux and Peng [97] to coupled FBSDEs. The reader may find an introductory theory of FBSDEs in Appendix A. This study addresses specifically the problem of numerical resolution of FBSDEs which has been an active area of research for the last two decades. The interest in numerical solutions for FBSDEs mainly stems from their various applications, especially in mathematical finance, and the lack of general closed form solutions.

This short introduction first provides indications on the notation used throughout the document. Then a presentation is given on the Euler time discretization for BSDEs since we mostly rely on this scheme in this study. The introduction ends with a summary of the thesis content.

Notation

For a fixed terminal time $T > 0$, consider the complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$, so that $\mathbb{F} = \{\mathcal{F}_t : t \in [0,T]\}$ where the filtration is generated by a d-dimensional Brownian motion $\{W_t\}_{t\in[0,T]}$ and $\mathcal{F}_0$ contains the $\mathbf{P}$-null sets of the σ-algebra $\mathcal{F}$. We shall make use of the following operators and spaces through our presentation:

- For any vector $x \in \mathbb{R}^n$, we note the Euclidean norm as $|x| = \left(\sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}$. For a matrix $b \in \mathbb{R}^{n\times m}$, we use the Frobenius norm which verifies $|b|^{\frac{1}{2}} = \mathrm{Tr}(bb^*)$ where b^* is the transpose matrix of b. Also, $\|b\|_2$ represents the spectral norm of b when b is a square matrix.
- $\mathcal{C}^n(\mathbb{R}^m)$ is the set of n-times differentiable real valued functions on $\mathbb{R}^m$, and $C_b^n(\mathbb{R}^m)$ denotes that the derivatives are absolutely bounded up to order n.
- $\mathcal{B}(\mathbb{R}^m)$ is the Borel set on $\mathbb{R}^m$.
- $\mathcal{S}$ is the Skorohod set of real valued càdlàg functions on $[0,T]$.
- $L^p(\mathbb{R}^n)$ is the space of $\mathcal{F}_T$-measurable $\mathbb{R}^n$-valued random variables X such that $\|X\|_{L^p} = \mathbf{E}\left[|X|^p\right]^{\frac{1}{p}} < \infty$ for $p \in \mathbb{N}$ and L^∞ denotes the space of bounded random variables.
- $L_S^p(\mathbb{R}^n)$ is the space of $\mathbb{R}^n$-valued adapted processes X such that

$$\|X\|_{L_S^p} = \mathbf{E}\left[\sup_{t\in[0,T]} |X_t|^p\right]^{\frac{1}{p}} < \infty,$$

L_S^∞ denotes the space of bounded processes.

- $L_I^p(\mathbb{R}^n)$ is the space of predictable $\mathbb{R}^n$-valued processes X such that

$$\|X\|_{L_I^p} = \mathbf{E}\left[\left(\int_0^T |X_t|^2\, dt\right)^{\frac{p}{2}}\right]^{\frac{1}{p}} < \infty.$$

For a given stochastic process X on $(\Omega, \mathcal{F})$, we simplify the notation of the conditional expectation with respect to $\mathcal{F}_t$ as

$$\mathbf{E}_t^x\left[u(X_T)\right] = \mathbf{E}\left[u(X_T)|X_t = x\right] \tag{I.1}$$

for any deterministic function u where the conditional expectation is taken under the physical probability measure $\mathbf{P}$. Additional indications will be given when using an equivalent probability measure. Also, C and K will denote generic constants in our various inequalities.

Finally, all numerical results in this thesis are produced using a Pentium (R) Dual-Core, T4200 model processor with 2.0 GHz.

The Euler scheme

Many numerical methods for FBSDEs with Lipschitz coefficients presented in Chapter 1 (and particularly Monte Carlo regression and space discretization methods) are based on a time discretization called the Euler scheme. Those methods only differ in their approximations of the conditional expectations involved in the time discretization. This section intends to present the main time discretization algorithms for (F)BSDEs and the associated convergence result.

The Euler scheme is a time stepping method designed for SDEs and applicable to BSDEs and decoupled FBSDEs with Lipschitz coefficients. We shall consider the decoupled FBSDE of equation (A.5) with deterministic coefficients. The discrete time procedure that we present was first introduced by Zhang [123, 124] followed by Bouchard and Touzi [20] and Hu, Nualart and Song [61] among others. A concise summary of the method also figures in the review paper of Bouchard, Elie and Touzi [19].

Because of the assumption of Lipschitz coefficients, the forward SDE discretization is usually performed with an Euler scheme using a partition $\pi = \{0 = t_0 < t_1 < ... < t_n = T\}$ of the interval $[0, T]$. The numerical solution $\{X_t^\pi\}_{t\in[0,T]}$ relies on the values at times nodes defined as

$$\begin{cases} X_0^\pi = x_0 \\ X_{t_{i+1}}^\pi = X_{t_i}^\pi + a(t_i, X_{t_i}^\pi)\Delta_i + \sigma(t_i, X_{t_i}^\pi)\Delta W_i \end{cases} \tag{I.2}$$

where $\Delta_i = t_{i+1} - t_i$ and $\Delta W_i = W_{t_{i+1}} - W_{t_i}$ for $i = 0, 1, ..., n-1$. We get

$$X_t^\pi = X_{t_i}^\pi,\ t \in [t_i, t_{i+1}) \tag{I.3}$$

and this discretization yields a strong $\frac{1}{2}$-order convergent solution since the error $E_{X,\pi}$ on the forward process satisfies

$$E_{X,\pi}^2 := \mathbf{E}\left[\sup_{t\in[0,T]} |X_t - X_t^\pi|^2\right] = \mathcal{O}(|\pi|) \tag{I.4}$$

with Landau notation as shown in Kloeden and Platen [69] where $|\pi|$ is the maximal time step

$$|\pi| = \max_{0\le i<n} \Delta_i. \tag{I.5}$$

An Euler scheme also helps in discretizing the backward process and leads to the following time stepping

$$Y_{t_i}^{\pi} = Y_{t_{i+1}}^{\pi} + f(t_i, X_{t_i}^{\pi}, Y_{t_i}^{\pi}, Z_{t_i}^{\pi})\Delta_i - (Z^{\pi})_{t_i}^{*}\Delta W_i. \tag{I.6}$$

Taking the conditional expectations on both sides of equation (I.6) yields

$$Y_{t_i}^{\pi} = \mathbf{E}\left[Y_{t_{i+1}}^{\pi}|\mathcal{F}_{t_i}\right] + f(t_i, X_{t_i}^{\pi}, Y_{t_i}^{\pi}, Z_{t_i}^{\pi})\Delta_i$$

and if one first multiplies both sides of equation (I.6) by the Brownian increment ΔW_i and takes the conditional expectation after, we get

$$Z_{t_i}^{\pi} = \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^{\pi}\Delta W_i|\mathcal{F}_{t_i}\right].$$

Those last two equations define the backward algorithm for numerical solution of BSDEs

$$\begin{cases} Z_{t_n}^{\pi} = 0,\ Y_{t_n}^{\pi} = \xi^{\pi} \\ Z_{t_i}^{\pi} = \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^{\pi}\Delta W_i|\mathcal{F}_{t_i}\right] \\ Y_{t_i}^{\pi} = \mathbf{E}\left[Y_{t_{i+1}}^{\pi}|\mathcal{F}_{t_i}\right] + f(t_i, X_{t_i}^{\pi}, Y_{t_i}^{\pi}, Z_{t_i}^{\pi})\Delta_i \end{cases} \tag{I.7}$$

known as the implicit Euler scheme since the value of the approximate forward process $Y_{t_i}^{\pi}$ appears on both sides of the system last equation. We choose ξ^{π} such that

$$\|\xi - \xi^{\pi}\|_{L^2}^2 = \mathcal{O}(|\pi|) \tag{I.8}$$

which is possible, for instance, in the Markovian case by taking $\xi^{\pi} = g(X_T^{\pi})$ for continuous terminal conditions g.

In order to avoid solving a non-linear equation to recover the backward process values $Y_{t_i}^{\pi}$, one may consider an alternative scheme which is explicit in the backward process values

$$\begin{cases} Z_{t_n}^{\pi} = 0,\ Y_{t_n}^{\pi} = \xi^{\pi} \\ Z_{t_i}^{\pi} = \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^{\pi}\Delta W_i|\mathcal{F}_{t_i}\right] \\ Y_{t_i}^{\pi} = \mathbf{E}\left[Y_{t_{i+1}}^{\pi} + f(t_i, X_{t_i}^{\pi}, Y_{t_{i+1}}^{\pi}, Z_{t_i}^{\pi})\Delta_i|\mathcal{F}_{t_i}\right]. \end{cases} \tag{I.9}$$

and called the explicit Euler scheme. The approximate backward and control processes are then defined as done previously for the forward process as

$$Y_t^{\pi} = Y_{t_i}^{\pi},\ Z_t^{\pi} = Z_{t_i}^{\pi} \text{ for } t \in [t_i, t_{i+1}). \tag{I.10}$$

Another explicit scheme consists in replacing the conditional expectation of the driver by the driver values at conditional expectations from the explicit scheme of equation (I.9). This procedure leads to the scheme

$$\begin{cases} Z_{t_n}^{\pi} = 0,\ Y_{t_n}^{\pi} = \xi^{\pi} \\ Z_{t_i}^{\pi} = \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^{\pi}\Delta W_i|\mathcal{F}_{t_i}\right],\ 0 \leq i < n \\ Y_{t_i}^{\pi} = \mathbf{E}\left[Y_{t_{i+1}}^{\pi}|\mathcal{F}_{t_i}\right] + f(t_i, \mathbf{E}\left[Y_{t_{i+1}}^{\pi}|\mathcal{F}_{t_i}\right], Z_{t_i}^{\pi})\Delta_i,\ 0 \leq i < n. \end{cases} \tag{I.11}$$

We will denote the scheme of equation (I.11) the explicit Euler scheme 1 and the scheme of equation (I.9) is denoted the explicit Euler scheme 2.

The global discretization error E_{π} in the backward and control processes is defined as

$$E_{\pi}^2 := \max_{0 \leq i < n} \mathbf{E}\left[\sup_{t \in [t_i, t_{i+1}]} \left|Y_t - Y_{t_i}^{\pi}\right|^2\right] + \sum_{i=0}^{n-1} \mathbf{E}\left[\int_{t_i}^{t_{i+1}} \left|Z_s - Z_{t_i}^{\pi}\right|^2 ds\right] \tag{I.12}$$

and, under Lipschitz conditions on the FBSDE coefficients, admits the bound

$$E_\pi^2 \leq C\left(|\pi| + \sum_{i=0}^{n-1} \mathbf{E}\left[\int_{t_i}^{t_{i+1}} \left|Z_s - \bar{Z}_{t_i}^\pi\right|^2 ds\right]\right) \tag{I.13}$$

where

$$\bar{Z}_{t_i}^\pi = \frac{1}{\Delta_i}\mathbf{E}_{t_i}\left[\int_{t_i}^{t_{i+1}} Z_s ds\right] \tag{I.14}$$

and the constant C depends exponentially on the Lipschitz constants due to the usage of the Gronwall's inequality in the proof. From this equation, a regularity property on the control process Z is needed to prove the convergence of the Euler scheme.

This regularity result is proved by Zhang [124, 123] in the context of Lipschitz coefficients. Knowing that $\bar{Z}_{t_i}^\pi$ is the projection on the space $\mathcal{F}_{t_i}$-measurable random variables, and hence the best $\mathcal{F}_{t_i}$-measurable approximation, of $\{Z_t\}_{t\in[0,T]}$ on the interval $[t_i, t_{i+1}]$ we have

$$\begin{aligned}\sum_{i=0}^{n-1} \mathbf{E}\left[\int_{t_i}^{t_{i+1}} \left|Z_s - \bar{Z}_{t_i}^\pi\right|^2 ds\right] &\leq \sum_{i=0}^{n-1} \mathbf{E}\left[\int_{t_i}^{t_{i+1}} \left|Z_s - Z_{t_i}\right|^2 ds\right] \\ &\leq C\,|\pi|\end{aligned} \tag{I.15}$$

from equation (A.18). Consequently, the Euler schemes have a discretization error of

$$E_\pi = \mathcal{O}(|\pi|^{\frac{1}{2}}). \tag{I.16}$$

Summary

Chapter 1 gives an overview of existing numerical methods divided in three groups: partial differential equation (PDE) based methods, spatial discretization based methods and Monte Carlo regression based methods. Overall, PDE based methods require strong regularity condition on the FBSDE coefficient whereas Monte Carlo methods are time consuming. Spatial discretization methods may be seen as a tradeoff between the two other groups of methods.

The purpose of this thesis is the development of a spatial discretization method using analytic and numerical Fourier techniques. These techniques have already proved very efficient in various areas such as in numerical methods for PDEs or in mathematical finance. In the context of numerical methods for FBSDEs, numerical Fourier techniques have the advantage and particularity of being spectral methods especially in the group of spatial discretization methods for FBSDEs.

Chapter 2 gives the representation of BSDE numerical solutions as Fourier integrals and proposes a first implementation of a fast Fourier transform (FFT) based method for BSDEs. The numerical resolution is performed on a uniform grid and a local error analysis reveals a consistent method with truncation errors under smoothness conditions on the BSDE coefficients. Numerical examples then illustrate the accuracy of the method.

Even though the procedure in Chapter 2 is consistent and suitable for various applications, the presence of truncation errors makes it less accurate. Chapter 3, which is the central chapter of this thesis, introduces an alternative tree-like grid that removes the truncation error leading to a conditionally stable and globally convergent method. A simulation method for BSDEs is then developed and the convergence of the approximation is proved.

Chapter 4 focuses on the space discretization of FBSDEs as an extension of the material of Chapter 3. In the context of FBSDEs, we show that the Fourier representation of the BSDE numerical solution is possible and then build a conditionally stable and globally convergent

method. Nonetheless, the computations have to be performed through matrix multiplication instead of the FFT algorithm as in Chapter 2 and 3 so that Fourier methods are less efficient when applied on FBSDEs .

Finally, Chapter 5 visits the very recently proposed Runge-Kutta scheme for (F)BSDEs. Runge-Kutta schemes are higher order time convergent time stepping methods that improve the half order convergence of the Euler scheme used in the previous chapters. Here also, we characterize the numerical solutions and apply the method developed in Chapter 4 (or Chapter 3 in the simple BSDE case). In the more general setting of Runge-Kutta schemes, the conditional stability and global convergence of the method is established under mild conditions on some characteristic functions.

Chapter 1

A literature review of numerical methods for FBSDEs

Solutions to FBSDEs can be found through either the *four-step scheme* or the *method of continuation*[1]. In the former method, proposed by Ma, Protter and Yong [79], the forward and control processes are expressed explicitly in terms of a function which solves a Cauchy problem for a quasilinear parabolic PDE. Yong [114], for the latter method, gives implicit solutions using the notion of *bridge* which identifies new solutions of FBSDEs from known ones. Hence, on one hand solving a FBSDE explicitly by means of the *four-step scheme* requires an explicit solution to a quasilinear PDE, which is not an easy task, and on the other hand the *method of continuation* only gives implicit solutions. For this reason, numerical methods are being developed in order to seek approximate solutions to FBSDEs. The first numerical methods to appear were the PDE based method of Douglas, Ma and Protter [40] along with the random time partition scheme of Bally [7] and the tree based scheme of Chevance [28].

The present chapter gives an overview of existing numerical methods for FBSDEs and is organized as follows. Section 1.1 deals with the four-step scheme based methods that numerically solve the quasilinear PDE. Spatial discretization methods, including quantization tree methods and multinomial trees, are presented in Section 1.2. Section 1.3 is about Monte Carlo regression based methods and the final section summarizes numerical methods for BSDEs with non-Lipschitz coefficients.

1.1 Four-step scheme based methods

Even if all numerical methods for FBSDEs can be linked to the *four-step scheme*, we can clearly distinguish those treating the associated quasilinear PDE directly. These methods impose differentiability and boundedness constrains on the FBSDE coefficients in order to ensure the convergence of the numerical scheme for the PDE. Also, their extension to multidimensional FBSDEs remains an open problem due to the lack of algorithms for multidimensional PDEs. Nonetheless, four-step scheme based methods can compute numerical solutions to the most general coupled FBSDEs. Here, we present those methods according to the type of algorithm used to solve the PDE numerically.

[1] See the book of Ma and Yong [81]

1.1.1 Finite difference methods

In 1996 Douglas, Ma and Protter [40] initiated numerical methods for FBSDEs. In that paper, the authors rely on the *four-step scheme* and solve the quasilinear parabolic PDE using a combination of characteristics and finite difference method.

Two similar types of FBSDE with nonlinear coefficients are considered: a "general case" FBSDE where the diffusion coefficient of the forward process does not depend on the control process and a "special case" one where, in addition to the diffusion coefficient of the forward process, the forward and backward drifts do not depend on the control process. Thus, the stochastic equations are coupled in both cases and all involved processes (namely the forward and backward processes, the control process and the Brownian motion) are assumed to be one-dimensional.

When solving the "special case" FBSDE

$$\begin{cases} dX_t = a(t, X_t, Y_t)dt + \sigma(t, X_t, Y_t)dW_t \\ -dY_t = f(t, X_t, Y_t)dt - Z_t dW_t \\ X_0 = x_0 \ , \ Y_T = \xi \end{cases} \tag{1.1.1}$$

the PDE reduces to

$$\frac{\partial u}{\partial t} + a(t, x, u)\frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2(t, x, u)\frac{\partial^2 u}{\partial x^2} + f(t, x, u) = 0 \tag{1.1.2}$$

which is first modified using a time change of variable to get an initial value value problem. A method of characteristics is then applied and reduces the quasilinear advection-diffusion PDE into a diffusion PDE. A finite difference scheme then discretizes the resulting PDE with a first order forward difference in time along the characteristic and a second order central difference in space. For a time step of Δt and a space step of Δx, this gives a solution u_i^k at time mesh t_k and grid point x_i of the form

$$\frac{u_i^k - \bar{u}_i^{k-1}}{\Delta t} = \frac{1}{2}\sigma^2(t_k, x_i, u_i^{k-1})D^2[u_i^k] + f(t_k, x_i, u_i^{k-1}) \tag{1.1.3}$$

where $\bar{u}_i^{k-1} = u^{k-1}(x_i - a(t_k, x_i, u_i^{k-1})\Delta t)$ is the solution value along the characteristic obtained by interpolation, and

$$D^2[u_i^k] = \frac{u_{i+1}^k - 2u_i^k + u_{i-1}^k}{(\Delta x)^2}$$

stands for the second order finite difference for the second derivative. The scheme produces a convergent solution with a first order local truncation error in both time and space.

The availability of a numerical solution to the PDE then allows the construction of numerical solutions for the forward and backward SDEs using the *four-step scheme* representation of their solutions and this construction is made possible since the approximate solution to the PDE satisfies the Lipschitz condition. A forward (or explicit) Euler scheme[2] gives a numerical solution to the forward SDE and the approximate solution to the backward SDE is obtained by interpolating the PDE approximate solution at the values of the forward SDE numerical solution. This procedure reproduces the convergence rates of the underlying Euler scheme. More precisely, we get a half ($\frac{1}{2}$) order strongly convergent scheme for the solutions of both the forward and backward SDE and a first order weakly convergent scheme for the forward SDE solution.

The numerical solution for the "general case" FBSDE

$$\begin{cases} dX_t = a(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t)dW_t \\ -dY_t = f(t, X_t, Y_t, Z_t)dt - Z_t dW_t \\ X_0 = x_0 \ , \ Y_T = \xi \end{cases} \tag{1.1.4}$$

[2]See Kloeden and Platen[69].

where the PDE takes the form

$$\frac{\partial u}{\partial t} + a\left(t, x, u, \sigma(t,x,u)\frac{\partial u}{\partial x}\right)\frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2(t,x,u)\frac{\partial^2 u}{\partial x^2} + f\left(t, x, u, \sigma(t,x,u)\frac{\partial u}{\partial x}\right) = 0, \quad (1.1.5)$$

seeks not only approximate solutions for the forward and backward processes but also for the control process. This requires the implementation of a numerical scheme for the derivative of the PDE solution. Hence, the authors deduce a second quasilinear advection-diffusion PDE describing the aforementioned derivative by differentiating the initial PDE. The resulting system of PDEs is then solved by applying the numerical scheme of the "special case" to each PDE. Interpolations on the numerical solution of the PDE system help in solving the forward SDE with an Euler scheme and computing values for the backward SDE solution and the control process. Again, the method is half ($\frac{1}{2}$) order strongly convergent for the triple of processes solution to the FBSDE and first order weakly convergent for the solution of the forward SDE and the control process.

Milstein and Tretyakov [89, 90] also used a finite difference method on PDEs to solve FBSDEs numerically. The algorithms in both papers are based on the *four-step scheme* but apply to a less general type of FBSDEs. In the first paper [89], the authors consider a coupled FBSDE where the forward coefficients do not depend on the control process and the driver depends only linearly on it. The second paper [90] generalizes the first by introducing the control process in the forward drift and a nonlinear driver in the control process. The driver has the form

$$f(t,x,y,z) = f_1(t,x,y,z) + f_2(t,x,y,z)y + f_3(t,x,y,z)z \quad (1.1.6)$$

and this structure leads to the PDE

$$\begin{aligned} &\frac{\partial u}{\partial t} + \tilde{a}\left(t, x, u, \sigma(t,x,u)\frac{\partial u}{\partial x}\right)\frac{\partial u}{\partial x} \\ &+ \quad \frac{1}{2}\sigma^2(t,x,u)\frac{\partial^2 u}{\partial x^2} + f_2\left(t, x, u, \sigma(t,x,u)\frac{\partial u}{\partial x}\right)u \\ &+ \quad f_1\left(t, x, u, \sigma(t,x,u)\frac{\partial u}{\partial x}\right) = 0 \end{aligned} \quad (1.1.7)$$

where $\tilde{a} = a + \sigma f_3$.

One of the major differences between the approach of Milstein and Tretyakov [89, 90] and the one of Douglas, Ma and Protter [40] is in the numerical resolution of the quasilinear PDE. Indeed, Milstein and Tretyakov [89, 90] chose a layer method which is a first order numerical algorithm for parabolic PDEs grounded on these PDEs probabilistic representation. At time node t_k and for a time step of Δt, the numerical solution u_k of the PDE in equation (1.1.2) is given, for instance, by

$$u_k(x) = \frac{u_{k+1}(x_k^+) + u_{k+1}(x_k^-)}{2} + (\Delta t) f\left(t_k, x, u_{k+1}(x)\right) \quad (1.1.8)$$

where

$$x_k^{\pm} \quad = \quad x + (\Delta t) a\left(t_k, x, u_{k+1}(x)\right) \pm \sqrt{\Delta t}\sigma\left(t_k, x, u_{k+1}(x)\right). \quad (1.1.9)$$

As to the stochastic part, the forward SDE resolution may consist of an Euler or Milstein scheme and the backward and control processes are obtained by interpolating the PDE numerical solution. The second major difference with Douglas, Ma and Protter [40] approach being that a finite difference on the PDE numerical solution approximates the derivative of the PDE solution. Indeed, the discrete operator D_l defined as

$$D_l[u_{k+1}](x) = \frac{u_{k+1}(x_k^+) - u_{k+1}(x_k^-)}{2\sigma\left(t_k, x, u_{k+1}(x)\right)\sqrt{\Delta t}}, \quad (1.1.10)$$

which is a central finite difference for the first derivative in the layer method, may be used to approximate the derivative in various situations.

The method's rate of mean-square convergence is $\frac{1}{2}$ for the FBSDE triple solution when an Euler scheme discretizes the forward SDE and 1 for the forward and backward solutions when the Milstein scheme is used. The Milstein scheme on the forward SDE also yields a first order mean-square convergence for the control process if the finite difference on the PDE numerical solution is first order accurate.

1.1.2 The spectral method

It can be noted that, except for the Milstein and Tretyakov [89, 90] method under the Milstein scheme with its first order convergence, numerical methods for FBSDEs yield a half order convergence[3]. In their paper, Ma, Shen and Zhao [80] propose a four-step based numerical method with an enhanced rate of convergence.

A uni-dimensional fully coupled FBSDE with continuously differentiable coefficients is considered. These coefficients must satisfy further growth and boundedness conditions. In particular, the forward process volatility, which does not depend on the control process, must have a bounded second derivative.

The algorithm is applied to the quasilinear PDE expressed in divergence form

$$\frac{\partial u}{\partial t} + \tilde{a}\left(t, x, u, \frac{\partial u}{\partial x}\right)\frac{\partial u}{\partial x} + \frac{\partial}{\partial x}\left(\frac{1}{2}\sigma^2(t, x, u)\frac{\partial u}{\partial x}\right) + f\left(t, x, u\right) = 0 \tag{1.1.11}$$

for some Lipschitz and bounded function $\tilde{a}$. The discretization of this PDE starts with a first order implicit time stepping that leads to an uni-dimensional elliptic equation in space at each time node, but other time discretizations may be chosen[4] to reach higher orders of accuracy. At time step t_k, the solution u_k is expressed as

$$\begin{aligned} u_k - \left(\frac{\Delta t}{2}\right)\frac{\partial}{\partial x}\left(\sigma^2(t_k, x, u_{k+1})\frac{\partial u_k}{\partial x}\right) = \\ u_{k+1} + (\Delta t)\left(\tilde{a}\left(t_k, x, u_{k+1}, \frac{\partial u_{k+1}}{\partial x}\right)\frac{\partial u_{k+1}}{\partial x} + f(t_k, x, u_{k+1})\right). \end{aligned} \tag{1.1.12}$$

The structure of the space domain, and more precisely the fact that it is the whole real line, imposes the usage of Hermite polynomials[5] in the spectral method. Hence, at each time step, the PDE solution u_k and the term $\sigma^2(t_k, x, u_{k+1})\frac{\partial u_k}{\partial x}$ are interpolated with weighted polynomials with weight function

$$w(x) = e^{-\frac{1}{2}x^2} \tag{1.1.13}$$

and an integration by parts of the elliptic ordinary differential equation (ODE) gives a variational equation using Hermite quadrature. The last step for solving the PDE is to find the interpolation weights from the variational equation. The procedure produces a numerical solution[6] with first order accuracy in time and spectral convergence in space.

Concerning the numerical solution to the forward process, Ma, Shen and Zhao [80] propose three different highly accurate schemes including the first order Milstein scheme and two $\frac{3}{2}$-strongly consistent schemes in addition to the usual explicit Euler scheme. In particular, the Euler scheme leads to an approximation of the triple of processes which is half order accurate. The idea that the method convergence rate can be improved when using a higher time stepping

[3] This includes four-step and Monte Carlo regression based methods as we will see later in this review.
[4] The authors mention an alternative Adam–Bashforth scheme.
[5] See Appendix C for a definition of Hermite polynomials and more precisely equation (C.10).
[6] Including the solution to the PDE and its derivative.

method for the PDE and a higher order scheme for the forward SDE is illustrated with numerical examples. The authors are able to reach a $\frac{3}{2}$-order of convergence for the FBSDE solution.

1.2 Spatial discretization based methods

Quantization based methods aim to weaken the regularity conditions imposed on the FBSDE coefficients in PDE based methods. Those methods consist of replacing a (continuous) random variable by a discrete one in order to estimate the expectations involved in the local representation of the FBSDE. Hence, the methods avoid the direct numerical treatment of the associated quasi-linear PDE.

1.2.1 Quantization tree methods

After a time discretization of some local representation of FBSDE, quantization tree methods typically discretize the forward process space at each time node using a quantization grid. A function, called the quantizer, projects the values of the forward process on the grid. Calculating the conditional expectations involved in the local representation of the FBSDE then depends on the availability of transition probabilities at each time step.

In order to avoid solving a PDE numerically, Bally [7] developed a method for BSDEs with a driver depending on the control process but the resolution of the BSDE has to be done at Poisson random times to prove the method convergence. This inconvenience was mainly due to the unknown path regularity of the control process. Chevance [28] overcame the time discretization randomness by proposing the first quantization tree and convergent method for decoupled FBSDEs.

Chevance [28] considers Markovian cases where the backward process terminal value depends only on the forward process terminal value and his method applies if the BSDE driver does not depend on the control process. Also, the decoupled FBSDE coefficients (forward SDE coefficients, the driver and the terminal function) have to satisfy further differentiability and boundedness conditions.

The forward SDE is discretized (over a deterministic time grid) by a weak Euler scheme where the Brownian increments are replaced by a discrete random variable[7]. As to the backward SDE, its discrete local representation uses the driver's upper value. This discretization yields a first order absolute error (in time).

In order to compute an approximate solution for the backward equation, Chevance [28] uses a fixed space grid for the forward SDE at each time step and the forward process values on the space grid are determined with the closest neighbor rule as the quantizer. More precisely, at each time step the weak solution values of the forward process (using the previous time step projections) are projected on the grid and the projection itself is defined for any number as the least value among the closest grid point values to the number. The backward process values are then computed (backward in time) by evaluating the BSDE discrete local representation using the forward process projected values and the transition probabilities provided by the aforementioned discrete random variable. The procedure yields a weak solution for the BSDE that is first order accurate in time and space.

Other quantization tree methods include the methods of Bally and Pages [8] and Bally, Pages and Printems [9]. The methods are designed for reflected BSDEs (RBSDEs) and particularly for

[7] The discrete random variable must have zero mean, zero third moment, an unit variance and finite fifth moment.

the problem of pricing multidimensional American options and require only the usual Lipschitz conditions needed for the decoupled RFBSDE well-posedness. Along with the method, the authors address the problem of optimal grid and transition probability choice. Finally, Delarue and Menozzi [35] developed a quantization method for solving parabolic PDEs via their probabilistic representation through coupled FBSDEs. An interpolation procedure was then proposed by the same authors in [36] as an improvement of the method.

1.2.2 Multinomial trees

Multinomial trees differ from quantization trees in two main fashions. First, the BSDE is itself discretized in multinomial methods whereas quantization methods only discretize the BSDE local representation. Also, multinomial trees replace the Brownian increments that appear in the discrete version of the FBSDE with discrete random variables: in quantization methods, the forward process is directly targeted.

Indeed, note that Chevance discretizes only the local representation of the BSDE but not the BSDE itself because of the unknown regularity of the control process. Hence, he does not provide a numerical solution for the control process. The direct discretization of the BSDE[8] is later made possible by Briand, Delyon and Memin [21] along with Ma, Protter, San Martin and Torres [78] with the proof of its convergence with quasi-minimal conditions on the BSDE driver and terminal random variable.

Briand, Delyon and Memin [21] and Ma, Protter, San Martin and Torres [78] both discretize Brownian increments with a symmetric Bernoulli distribution leading to a binomial tree. The method applies to one dimensional problems but can be generalized to multidimensional[9] BSDEs. Nonetheless, one can note differences between the papers regarding the conditions of the BSDE coefficients, the converging objects, the type of convergence and the induced algorithms. Indeed, the BSDE driver of Briand, Delyon and Memin [21] depends on the control process whereas in [78] it does not but requires continuity. Also, the first paper [21] proves convergence in probability for the forward and control processes whereas the latter [78] proves convergence in distribution for the forward process and the stochastic integral of the control process with respect to the Brownian motion. Finally, Briand, Delyon and Memin [21] propose only an implicit method whereas Ma, Protter, San Martin and Torres[78] put forth a suitable explicit algorithm.

In general, the BSDE is discretized on an uniform partition with time step Δ as

$$Y_{t_i} = Y_{t_{i+1}} + \Delta f(t_i, Y_{t_i}, Z_{t_i}) - \sqrt{\Delta} Z_{t_i} \epsilon_{i+1} \tag{1.2.1}$$

where the ϵ_i are a sequence of independent, symmetric and discrete random variables. Equation (1.2.1) can be numerically solved with the implicit backward algorithm

$$Y_{t_i} = \mathbf{E}\left[Y_{t_{i+1}} | \mathcal{G}_{t_i}\right] + \Delta f(t_i, Y_{t_i}, Z_{t_i}) \tag{1.2.2}$$

$$Z_{t_i} = \frac{1}{\sqrt{\Delta}} \mathbf{E}\left[Y_{t_{i+1}} \epsilon_{i+1} | \mathcal{G}_{t_i}\right] \tag{1.2.3}$$

where $\mathcal{G}$ is the discrete filtration generated by the sequence of ϵ_i's. For the explicit scheme, one may consider

$$Y_{t_i} = \mathbf{E}\left[Y_{t_{i+1}} | \mathcal{G}_{t_i}\right] + \Delta f(t_i, \mathbf{E}\left[Y_{t_{i+1}} | \mathcal{G}_{t_i}\right], Z_{t_i}) \tag{1.2.4}$$

instead of the expression of equation (1.2.2).

[8] By replacing the Brownian increments in the BSDE with random steps.

[9] Bouchard and Touzi [20] consider "questionable" the extension of tree methods to high dimension problems due to the difficulty of such extensions.

Building on Briand, Delyon and Memin [21, 22] and Ma, Protter, San Martin and Torres [78], Peng and Xu [99] worked on implicit and explicit algorithms for one dimensional BSDEs and reflected BSDEs (RBSDEs) based on the binomial method and proved their convergence for the backward and control processes under minimal conditions on the BSDE driver. Finally, in the paper of Briand, Delyon and Memin [22], the approximation of the Brownian motion is generalized from (scaled) random walks used in [21] to martingales.

1.3 Monte Carlo regression based methods

Simulation methods mostly apply to decoupled FBSDEs and essentially started with the work of Zhang [124] and his proof of the control process path regularity. This path regularity allows the implementation of strongly convergent (simulation based) algorithms with a deterministic time discretization and avoid time node randomization as in Bally [7], numerical treatment of the quasilinear PDE as in Douglas, Ma and Protter [40] or high regularity conditions on the BSDE coefficients as in [40] or [28]. Moreover, the main advantage of simulation based methods lies in their efficiency for multidimensional problems.

1.3.1 Backward schemes

The numerical implementation of Zhang's [124] method requires only the usual Lipschitz conditions on the FBSDE coefficients needed for the problem well-posedness. In particular, some Lipschitz regularities[10] are defined for the BSDE terminal value function that is allowed to be non-Markovian, in the sense that it may depend on the whole forward process path.

Given a (deterministic) time partition, the forward SDE describing a (non-homogeneous) Ito diffusion is numerically integrated with an Euler scheme which yields a half ($\frac{1}{2}$) order strongly convergent solution. The mean square error (MSE), when valuing the BSDE terminal function with the Euler solution of the forward process, is first order convergent in time in case the terminal function is Markovian or L^1-Lipschitz. The MSE is of the order of $\log(\Delta t^{-1})\Delta t$ for time step Δt in the L^∞-Lipschitz case.

As to the BSDE itself, Zhang [124] discretizes it with an explicit Euler scheme and the backward and control process values are recovered backward in time using the BSDE local representation. More precisely, expectations taken on the BSDE discretization compute the approximate adapted solutions of the BSDE. The scheme is proved to be convergent in the L^2-sense of Zhang [124] and its convergence for the backward and control processes is strongly related to the Lipschitz regularity of the BSDE terminal value function. Indeed, the scheme yields a squared error of the order of $\log(\Delta t^{-1})\Delta t$ if either the driver does not depend of the control process or the time partition is *K-uniform*[11]. If in addition the BSDE terminal value function is Markovian or L^1-Lipschitz, then the convergence is half order in time (i.e a squared error of order one). Zhang gives the following general error bound

$$\sup_{t\in[0,T]} \left\| Y_t - \hat{Y}_t \right\|_{L^2}^2 + \left\| Z - \hat{Z} \right\|_{L_I^2}^2 \leq C \left(|\pi| + \|\xi - \xi^\pi\|_{L^2}^2 \right) \tag{1.3.1}$$

for a given time partition π, where $\hat{Y}$ and $\hat{Z}$ are the BSDE piecewise constant numerical solutions.

When giving orders of convergence for his algorithm, Zhang [124] does not provide any method for the valuation of the expectations involved in the BSDE local discrete representation and numerical solution so that the order of convergence are valid when valuing exact expectations. Thus, not only does one still need an approximate method for the expectations in order

[10] Definition 2.1 of Zhang [124].
[11] Definition 5.2 of Zhang [124].

to implement the Zhang [124] algorithm but the orders of convergence have to be updated to take into account the error induced by the approximation. Two main papers appeared to solve this problem.

Bouchard and Touzi [20] consider a (decoupled) FBSDE where the forward process is a multidimensional time homogeneous Ito diffusion. The backward process is taken uni-dimensional for simplicity. As in Zhang [124], an explicit Euler scheme numerically integrates the forward SDE but the backward SDE is discretized with an explicit Euler scheme. Nonetheless, conditional expectations still need to be computed in order to approximate the backward and control processes discrete local representations.

In order to approximate the expectations, Bouchard and Touzi [20] apply a Malliavin calculus based regression method which forces the simulated forward process values at each time node to be independent. Hence, the forward process values at each time node have to be simulated independently of its values at any other time node leading to a particularly time consuming procedure.

For a fixed number N of paths (in the forward process simulation), the algorithm yields an L^p-error of the order Δt^{-1} for the backward process. Hence, the method has the undesired property that errors tend to explode for a fixed number of simulated paths when reducing the time step. The L^p-error due to the expectation approximation multiplies this time stepping error on the backward process and is of the order of $\Delta t^{-1-\frac{d}{4p}} N^{-\frac{1}{2p}}$, where the forward process has dimension d, for a global L^p-error of the order of $\Delta t^{-2-\frac{d}{4p}} N^{-\frac{1}{2p}}$. From this result, the authors point out that if the number of simulated paths is taken to be $n^{3p+\frac{d}{2}}$, where n is the number of time steps in an uniform time grid, then one achieves a half order global L^p-convergence in time for the backward process. The article ends with an extension of the method to RBSDEs and a numerical example on an American option pricing problem.

One of the main disadvantages of the Malliavin weights regression method is its computational complexity. Crisan, Manolarakis and Touzi [33] address this problem and propose a simplification of the algorithm in Bouchard and Touzi [20]. The simplification consists in using alternative Malliavin weights and preserves the method's convergence features.

Using similar ideas, Gobet, Lemor and Warin [53] proposed another simulation method for multidimensional decoupled FBSDEs that does not require independent simulations of the forward process. Conditions on the FBSDE coefficients are those needed for well-posedness and, particularly, the terminal function is assumed to have the L^∞-Lipschitz property in the non-Markovian case. As usual, a forward Euler scheme approximates for the forward process.

The backward algorithm for the BSDE is also built on a backward Euler discretization and uses least squares regression to approximate the BSDE solution. At each time step, the backward and control processes are represented as linear combinations of some basis functions. The projection coefficients for both processes are then determined by minimizing the regression mean square error with Picard iterations. The authors suggest orthogonal polynomials or hypercubes as choices of basis functions.

The (partial) convergence study in [53], made on the backward and control process, gives a rather complex expression for the upper bound of the method L^2-squared error. Nonetheless, it highlights another drawback of simulation-based backward schemes: the fact that errors accumulate through the iterations since, at any time step, the approximations are computed using previous ones.

Beside the most used explicit and implicit Euler schemes for BSDEs, other discretization

schemes can be found in the literature. Zhao, Shen and Peng [125] introduced θ-schemes where the backward SDE is locally discretized with the driver and the control process weighted upper and lower values. The numerical solutions Y_{t_i} and Z_{t_i} at mesh time t_i then solve the following system of equations

$$\begin{aligned} Y_{t_i} &= \mathbf{E}_{t_i}\left[Y_{t_{i+1}}\right] + \Delta_i\left\{(1-\theta_1)\mathbf{E}_{t_i}\left[f(t_{i+1}, Y_{t_{i+1}}, Z_{t_{i+1}})\right]\right. \\ &\quad \left. + \theta_1 f(t_i, Y_{t_i}, Z_{t_i})\right\} \qquad (1.3.2) \\ 0 &= \mathbf{E}_{t_i}\left[\Delta W_i Y_{t_{i+1}}\right] + \Delta_i(1-\theta_2)\mathbf{E}_{t_i}\left[f(t_{i+1}, Y_{t_{i+1}}, Z_{t_{i+1}})\Delta W_i\right] \\ &\quad -\Delta_i\left\{(1-\theta_2)\mathbf{E}_{t_i}\left[Z_{t_{i+1}}\right] + \theta_2 Z_{t_i}\right\} \qquad (1.3.3) \end{aligned}$$

where θ_1 and θ_2 are chosen in $[0,1]$ and may depend on the mesh time t_i. The trapezoidal rule which consists in setting $\theta_1 = \frac{1}{2}$ and $\theta_2 = 1$ is known to be second order accurate for the backward process and first order accurate for the control process under differentiability and boundedness conditions on the driver and the terminal condition. One may refer to Zhao, Wang and Peng [126] or Li and Zhao [74].

More recently, Zhao, Zhang and Ju [127] carried out a multistep scheme but the conditional expectations in their method are valued using Gauss-Hermite quadrature instead of Monte Carlo regression.

1.3.2 Forward schemes

Forward schemes were designed by Bender and Denk [10] to address the two problems inherent in simulation-based backward methods: the error explosion for small time steps and the error accumulation due to embedded expectation approximations.

As in Gobet, Lemor and Warin [53], Bender and Denk [10] develop a method for multi-dimensional decoupled FBSDEs with a non-homogeneous diffusion as the forward process, a non-Markovian L^∞-Lipschitz terminal function and the usual (and minimal) conditions on the FBSDE coefficients.

Instead of discretizing the BSDE locally, the authors employ Riemann type sums and Picard iterations on the backward and control processes to approximate the backward integral at each time step. More precisely, the approximate backward process is expressed in terms of the previous Picard iteration processes and the obtained backward stochastic integral is discretized with lower Riemann sums. Taking the expectation from the subsequent expression preserves the adaptedness of the numerical solutions and gives formulas for the values of the current Picard iteration backward and control processes at the different time nodes.

The n-th Picard iteration on a time partition $\pi = \{0 = t_0 < t_1 < ... < t_N = T\}$ then takes the form

$$Y^n_{t_i} = \mathbf{E}_{t_i}\left[\xi^\pi + \sum_{j=i}^{N-1} f(t_j, X^\pi_{t_j}, Y^{n-1}_{t_j}, Z^{n-1}_{t_j})\Delta_j\right] \qquad (1.3.4)$$

$$Z^n_{t_i} = \mathbf{E}_{t_i}\left[\frac{\Delta W_i}{\Delta_i}\left(\xi^\pi + \sum_{j=i}^{N-1} f(t_j, X^\pi_{t_j}, Y^{n-1}_{t_j}, Z^{n-1}_{t_j})\Delta_j\right)\right] \qquad (1.3.5)$$

where $Y^{n-1}_{t_j}$ and $Z^{n-1}_{t_j}$ are the values of the numerical solutions obtained in the previous Picard iteration for the backward and the control processes respectively.

Even though Picard iterations are the main feature in the method, the algorithm remains quite efficient since, in general, very few iterations are needed to have satisfactory results.

The forward process can be simulated with a forward Euler scheme and the conditional expectations involved at each Picard iteration evaluated with a regression method forward through

time nodes. The authors use a least squares regression with orthogonal basis functions, as in Gobet, Lemor and Warin [53], to approximate the conditional expectations.

The algorithm produces a numerical solution with a half order L^2-convergence in time when the conditional expectations are exact which is similar to the result of Zhang [124] for backward methods. More specifically, the quadratic error is given by

$$\sup_{t\in[0,T]} \|Y_t - Y_t^n\|_{L^2}^2 + \|Z - Z^n\|_{L_I^2}^2 \leq C\left(|\pi| + \|\xi - \xi^\pi\|_{L^2}^2 + \left(\frac{1}{2} + C|\pi|\right)^n\right) \tag{1.3.6}$$

Finally, the procedure's convergence is also proved in the L^2-sense without specification of the rate of convergence.

Finally, an algorithm using Picard iterations was recently constructed by Bender and Zhang [11] for weakly coupled FBSDEs with coefficients satisfying monotonicity conditions.

1.4 Methods for quadratic growth BSDEs

If numerical methods for FBSDEs with Lipschitz coefficients have mainly attracted researchers, the non-Lipschitz cases are becoming a growing interest. Among BSDEs with non-Lipschitz coefficients, those with quadratic growth are certainly the most studied. Since the proof of their well-posedness by Kobylanski [70] for bounded terminal conditions, the studies of Briand and Hu [24] extended the result to unbounded terminal conditions with exponential moments and convex driver.

Nonetheless, numerical methods for quadratic BSDEs are available only in the bounded terminal value case. This comes from the fact that the martingale defined by the Ito-integral of the control process $(Z \bullet W)_t = \int_0^t Z_s dW_s$ does not necessary conserve its the BMO property when the terminal condition is unbounded, making the derivation of regularity and convergence results arduous in that case.

Furthermore, numerical resolution of quadratic BSDEs faces a major obstacle even in the bounded terminal condition case. Indeed, the existing time discretization methods, and particularly the Euler scheme, fail to converge for general quadratic BSDEs.

Imkeller and Dos Reis [65] try to overcome those difficulties by proposing a method that applies to decoupled FBSDEs with bounded terminal condition and differentiable coefficients. First, the authors prove the required path regularities for the solutions of quadratic BSDEs. Their numerical method relies on a truncation of the driver which reduces the problem from a quadratic framework to a Lipschitz one. Hence, the classical time discretization can be used on the modified BSDE.

For one-dimensional BSDEs, a family of differentiable truncation functions $h_n : \mathbb{R} \to \mathbb{R}$ is defined such that $h_n(z)$ is simultaneously bounded by $|z|$ and the integer $n+1$. Then, any numerical method for FBSDEs with Lipschitz coefficients can be applied to the approximated BSDE

$$-dY_t^n = f(t, X_t, Y_t^n, h_n(Z_t^n))dt - Z_t^n dW_t$$

where the control process is replaced by its truncated value in the BSDE driver.

Since classical algorithms already converge for the truncated BSDE, the method's global convergence depends on the convergence of the truncated BSDE solution to the actual solution. Thanks to the BMO property of the process $Z \bullet W$, a convergence rate is provided and is given, for any $\beta > 0$ and $p \in \mathbb{N}^*$, by the following error bound

$$\|Y^n - Y\|_{L_S^{2p}}^{2p} + \|Z^n - Z\|_{L_I^{2p}}^{2p} \leq D_\beta n^{-\frac{\beta}{2q}} \tag{1.4.1}$$

where D_β and q are given constants.

The authors point out the lack of efficiency of the truncation approach. Indeed, the method's convergence is assured when the time step Δt is inversely proportional to the exponential of the Lipschitz constant related to the truncated driver. More precisely, the global error, including the discretization error, is bounded by

$$D_\beta \left(\frac{1}{n^\beta} + \frac{e^{Cn^2}}{N} \right)$$

where N is the number of time steps and $C > 0$ is related to the Lipschitz constant of the driver. Hence, one needs to consider more time steps when this Lipschitz constant slightly increases, so to say when one tries to get a slightly better approximation through truncation.

Recently, another numerical algorithm was proposed by Richou [103]. It solves for decoupled FBSDE with bounded terminal values and a forward volatility that is only a function of time on an nonuniform time grid. The main feature of Richou's approach is the Hölder continuity of the terminal value function.

Chapter 2

Convolution method for BSDEs

In this chapter, we implement a quadrature method for numerical solution of BSDEs. An Euler scheme discretizes the equation and we apply an FFT (Fast Fourier Transform) algorithm to value the conditional expectations induced by the time discretization. Hence, the approach is a typical spatial discretization method with the feature that the FFT algorithm serves as an alternative to trees and quantization when computing the quadratures.

The FFT algorithm computes the (inverse) discrete Fourier transform (DFT) of a given function and is widely used in various fields, and particularly in spectral and pseudo-spectral methods for PDEs. In addition to its flexibility, one of its main advantages is its efficiency as the FFT algorithm computes the values of the discrete Fourier transform of n function values in $\mathcal{O}(n\log(n))$ operations.

Two interesting financial applications of the algorithm are those of Carr and Madan [26] and Lord et al. [76]. These two papers employ the FFT algorithm to compute quadratures in the context of option pricing under Lévy processes. The first paper deals with European options whereas the latter treats American options. Our approach is much closer to the one of Lord et al. [76] since the numerical resolution of the BSDE is also made by dynamic programming through the Euler scheme.

2.1 Preliminaries

In this chapter, we study the numerical solution of BSDEs of the form

$$Y_t = g(W_T) + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s^* dW_s \tag{2.1.1}$$

where W is a d-dimensional Brownian motion with driver $f : [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and terminal condition $g : \mathbb{R}^d \to \mathbb{R}$. Conditions on f and g are given in section A.1 of appendix A. In addition, Proposition A.2 assures the well-posedness of such BSDEs.

It is known[1] from Pardoux and Peng [96] that if the Cauchy problem to the diffusion PDE

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{1}{2}\sum_{i=1}^d \frac{\partial^2 u}{\partial x_i^2} + f(t, u, \nabla u) = 0,\ (t,x) \in [0,T) \times \mathbb{R}^d \\ u(T,x) = g(x) \end{cases} \tag{2.1.2}$$

has a unique solution then the solution (Y, Z) for the BSDE admits the representation

$$Y_t \quad = \quad u(t, W_t) \tag{2.1.3}$$

[1] See section A.2.2 in the appendix.

$$Z_t = \nabla u(t, W_t). \tag{2.1.4}$$

This representation plays an important role in the sequel, since the convolution method implicitly solves the PDE of equation (2.1.2).

2.2 Convolution for BSDEs

In this section, we present the main ideas behind the convolution method. More precisely, we build numerical approximations to the BSDE of equation (2.1.1), using a time discretization of the BSDE. Thereafter, we give some useful properties of these solutions. For simplicity, we develop the method in the one dimensional case, i.e $d = 1$.

2.2.1 The approximate solutions and their properties

The starting point of the convolution method for BSDEs is the Euler scheme. We will mainly consider its explicit version which takes the form

$$\begin{cases} Z^\pi_{t_n} = 0,\ Y^\pi_{t_n} = \xi^\pi \\ Z^\pi_{t_i} = \frac{1}{\Delta_i}\mathbf{E}\left[Y^\pi_{t_{i+1}}\Delta W_i|\mathcal{F}_{t_i}\right] \\ Y^\pi_{t_i} = \mathbf{E}\left[Y^\pi_{t_{i+1}}|\mathcal{F}_{t_i}\right] + f(t_i, \mathbf{E}\left[Y^\pi_{t_{i+1}}|\mathcal{F}_{t_i}\right], Z^\pi_{t_i})\Delta_i \end{cases} \tag{2.2.1}$$

on a time mesh $\pi = \{t_0 = 0 < t_1 < ... < t_n = T\}$ and refer to it as the Euler scheme 1. A similar version of the Euler scheme was already proposed by Peng and Xu [99] in the context of binomial trees with the difference that the authors compute the expectations after the space discretization and, hence, with a discrete filtration.

Since the Brownian motion W is a Markov process, we define the approximate gradient $\dot{u}_i : \mathbb{R} \to \mathbb{R}$ at time mesh t_i, $i = 0, 1, ..., n-1$ as

$$\dot{u}_i(x) = \frac{1}{\Delta_i}\mathbf{E}\left[Y^\pi_{t_{i+1}}\Delta W_i|W_{t_i} = x\right] \tag{2.2.2}$$

so that the approximate control process is given by

$$Z^\pi_{t_i} = \dot{u}_i(W_{t_i}). \tag{2.2.3}$$

We let the intermediate solution $\tilde{u}_i : \mathbb{R} \to \mathbb{R}$ at time t_i take the form

$$\tilde{u}_i(x) = \mathbf{E}\left[Y^\pi_{t_{i+1}}|W_{t_i} = x\right]. \tag{2.2.4}$$

Consequently, an approximate solution of the PDE of equation (2.1.2) at mesh time t_i consists of a real-valued function $u_i : \mathbb{R} \to \mathbb{R}$ satisfying

$$Y^\pi_{t_i} = u_i(W_{t_i}) \tag{2.2.5}$$

can be defined as

$$u_i(x) = \tilde{u}_i(x) + \Delta_i f(t_i, \tilde{u}_i(x), \dot{u}_i(x)) \tag{2.2.6}$$

where

$$\dot{u}_i(x) = \frac{1}{\Delta_i}\int_{-\infty}^{\infty}(y - x)u_{i+1}(y)h(y|x)dy \tag{2.2.7}$$

$$\tilde{u}_i(x) = \int_{-\infty}^{\infty}u_{i+1}(y)h(y|x)dy \tag{2.2.8}$$

for $i = 0, 1, ..., n-1$ and $u_n(x) = g(x)$. Note that $\dot{u}_i$ is the approximate solution for the gradient of the PDE solution ∇u at mesh time t_i. Similar expressions are obtained by Delarue and Menozzi [35] in the context of FBSDEs and quasi-linear PDEs.

Also, the function h is the density function of $W_{t_{i+1}}$ conditional on the value of W_{t_i}. By the stationary and the independence of Brownian increments, we have that

$$h(y|x) \quad = \quad h(y-x). \tag{2.2.9}$$

As in Lord et al. [76], the relation of equation (2.2.9) plays a central role in the convolution method since it allows us to express the functions $\tilde{u}_i$ and $\dot{u}_i$ as convolutions. This, in addition, simplifies the application of Fourier transforms and hence the computation of the integrals of equations (2.2.7) and (2.2.8) via the DFT as we shall see in the sequel. Indeed

$$h(x) = (2\pi\Delta_i)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2\Delta_i}\right) \tag{2.2.10}$$

since increments of a Brownian motion are normally distributed.

The implicit Euler scheme can also be considered in this analysis, only the expression for the approximate solution differs. Indeed when the BSDE is discretized with the time stepping

$$\begin{cases} Z^\pi_{t_n} = 0,\, Y^\pi_{t_n} = \xi^\pi \\ Z^\pi_{t_i} = \frac{1}{\Delta_i}\mathbf{E}\left[Y^\pi_{t_{i+1}}\Delta W_i|\mathcal{F}_{t_i}\right] \\ Y^\pi_{t_i} = \mathbf{E}\left[Y^\pi_{t_{i+1}}|\mathcal{F}_{t_i}\right] + f(t_i, Y^\pi_{t_i}, Z^\pi_{t_i})\Delta_i \end{cases} \tag{2.2.11}$$

the approximate solution has u_i the implicit form

$$u_i(x) \quad = \quad \tilde{u}_i(x) + \Delta_i f(t_i, u_i(x), \dot{u}_i(x)). \tag{2.2.12}$$

The approximate gradient $\dot{u}_i$ and the intermediate solution $\tilde{u}_i$ satisfy equations (2.2.7) and (2.2.8) respectively. The implicit representation of equation (2.2.12) is solvable for the approximate solution u_i when

$$|\pi|\, K < 1 \tag{2.2.13}$$

where K is the Lipschitz constant of the driver f. In this case, the explicit and implicit schemes produce solutions with similar properties. For this reason, we mainly focus the analysis on the explicit Euler scheme 1.

The solution (Y^π_i, Z^π_i) and $(u_i, \dot{u}_i)$, $i = 0, 1, ..., n-1$, display important properties that are worth mentioning. These properties are easily extensible to the multidimensional case even though we present them in the one-dimensional setting. The next lemma describes the integrability property of the solution (Y^π_i, Z^π_i).

Lemma 2.1. *Suppose the conditions of Assumption A.2 are satisfied and $p \geq 2$. Then there exists a positive constant $C_\pi > 0$ depending on the Lipschitz constant K, the time horizon T, the partition π such that*

$$\sup_i \|Y^\pi_i\|_{L^p} + \sup_i \Delta_i^{\frac{1}{2}} \|Z^\pi_i\|_{L^p} \leq C_\pi \left(\|\xi\|_{L^p} + T \sup_{t\in[0,T]} |f(t,0,0)|\right). \tag{2.2.14}$$

Hence, if $\xi = g(W_T) \in L^p$ and $f(.,0,0) \in \mathcal{C}$

$$Y^\pi_i \in L^p \text{ and } Z^\pi_i \in L^p.$$

Proof. Let's first note that for $2 \leq p < \infty$

$$\|Z^\pi_i\|^p_{L^p} \quad = \quad \frac{1}{\Delta_i^p}\mathbf{E}\left[\left|\mathbf{E}_{t_i}\left[Y^\pi_{i+1}\Delta W_i\right]\right|^p\right]$$

$$
\begin{aligned}
&\leq \frac{1}{\Delta_i^p}\mathbf{E}\left[\Delta_i^{\frac{p}{2}}\left|\mathbf{E}_{t_i}\left[(Y_{i+1}^\pi)^2\right]\right|^{\frac{p}{2}}\right] \text{ (by the Cauchy-Schwartz inequality),}\\
&\leq \Delta_i^{-\frac{p}{2}}\mathbf{E}\left[\mathbf{E}_{t_i}\left[\left|Y_{i+1}^\pi\right|^p\right]\right] \text{ (by Jensen's inequality),}\\
&= \Delta_i^{-\frac{p}{2}}\mathbf{E}\left[\left|Y_{i+1}^\pi\right|^p\right]\\
&= \Delta_i^{-\frac{p}{2}}\left\|Y_{i+1}^\pi\right\|_{L^p}^p. \qquad (2.2.15)
\end{aligned}
$$

Also, letting $\tilde{Y}_i^\pi = \mathbf{E}_{t_i}\left[Y_{i+1}^\pi\right]$, we have

$$
\begin{aligned}
\left\|\tilde{Y}_i^\pi\right\|_{L^p}^p &= \mathbf{E}\left[\left|\mathbf{E}_{t_i}\left[Y_{i+1}^\pi\right]\right|^p\right]\\
&\leq \mathbf{E}\left[\mathbf{E}_{t_i}\left[\left|Y_{i+1}^\pi\right|^p\right]\right] \text{ (by Jensen's inequality),}\\
&= \left\|Y_{i+1}^\pi\right\|_{L^p}^p. \qquad (2.2.16)
\end{aligned}
$$

Since, by the Lipschitz property of the driver f,

$$
|Y_i^\pi| \leq \left|\tilde{Y}_i^\pi\right| + \Delta_i K\left(\left|\tilde{Y}_i^\pi\right| + |Z_i^\pi|\right) + \Delta_i \sup_{t\in[0,T]}|f(t,0,0)|
$$

where K is the Lipschitz constant of f. Hence, we have that

$$
\begin{aligned}
\|Y_i^\pi\|_{L^p} &\leq \left\|\tilde{Y}_i^\pi\right\|_{L^p} + \Delta_i K\left(\left\|\tilde{Y}_i^\pi\right\|_{L^p} + \|Z_i^\pi\|_{L^p}\right) + \Delta_i \sup_{t\in[0,T]}|f(t,0,0)|\\
&\quad \text{(using Minkowsky inequality),}\\
&\leq (1+\Delta_i K + \Delta_i^{\frac{1}{2}}K)\left\|Y_{i+1}^\pi\right\|_{L^p} + \Delta_i \sup_{t\in[0,T]}|f(t,0,0)|\\
&\quad \text{(using inequalities (2.2.15) and (2.2.16)),}\\
&\leq \prod_{j=i}^{n-1}(1+\Delta_j K + \Delta_i^{\frac{1}{2}}K)\left(\|\xi\|_{L^p} + T\sup_{t\in[0,T]}|f(t,0,0)|\right)\\
&\quad \text{(by Gronwall's Lemma),}\\
&\leq \exp\left(KT + K\sum_{i=0}^{n-1}\Delta_i^{\frac{1}{2}}\right)\left(\|\xi\|_{L^p} + T\sup_{t\in[0,T]}|f(t,0,0)|\right)\\
&\leq \frac{1}{2}C_\pi\left(\|\xi\|_{L^p} + T\sup_{t\in[0,T]}|f(t,0,0)|\right) \qquad (2.2.17)
\end{aligned}
$$

From the inequalities of equation (2.2.17) and (2.2.15) we have

$$
\Delta_i^{\frac{1}{2}}\|Z_i^\pi\|_{L^p} \leq \frac{1}{2}C_\pi\left(\|\xi\|_{L^p} + T\sup_{t\in[0,T]}|f(t,0,0)|\right). \qquad (2.2.18)
$$

Taking the supremum on the left hand side of (2.2.17) and (2.2.18) then leads to the result of equation (2.2.14) for $2 \leq p < \infty$. Finally, taking the limit as $p \to \infty$ gives the result in the case $p = \infty$ and completes the proof. □

As to the approximate solution $(u_i, \dot{u}_i)$, we first present their differentiability properties. This property will be used in the error analysis of the convolution method.

Lemma 2.2. *If the driver $f \in \mathcal{C}^{1,2,2}$ is twice differentiable in the backward and control variables and the driver $g \in \mathcal{C}^2$ is also twice differentiable then the approximate solution $u_i \in \mathcal{C}^2$ and the approximate gradient $\dot{u}_i \in \mathcal{C}^2$ are both twice differentiable for $i = 0, 1, ..., n-1$.*

Proof. The result follows by applying Leibniz's integral rule successively. □

It is important to ensure the approximate solution $(u_i, \dot{u}_i)$ is close enough to the PDE solution u and its gradient ∇u. The next lemma describes the error induced by the BSDE time discretization through equations (2.2.6), (2.2.8) and (2.2.7). It holds since we know from Zhang [124, 123] that the time discretization of the BSDEs yields a first order (quadratic) error term.

Lemma 2.3. *Under the conditions of Assumption A.2, we have that*

$$\max_{0\leq i<n} \boldsymbol{E}\left[\sup_{t\in[t_i,t_{i+1}]} |u(t,W_t)-u_i(W_{t_i})|^2\right] + \sum_{i=0}^{n-1} \boldsymbol{E}\left[\int_{t_i}^{t_{i+1}} |\nabla u(t,W_t)-\dot{u}_i(W_{t_i})|^2\,ds\right] = \mathcal{O}(|\pi|) \tag{2.2.19}$$

for any time discretization $\pi = \{0 = t_0 < t_1 < ... < t_n = T\}$.

Equations (2.2.6) and (2.2.8), along with the expression of the density of equation (2.2.9), show that the intermediate solution $\tilde{u}_i$ and the approximate gradient $\dot{u}_i$ are successive convolution transformations. More specifically, they are (generalized) Gauss-Weierstrass transforms since the kernel h is the Gaussian density. We give a quick introduction to the Gauss-Weierstrass transform in Appendix D. This presentation stems from the impressive literature on convolution transformations. The books of Hirshman and Widder [60] or Zemanian [120], for instance, are dedicated to the subject and also to the particular case of Gauss-Weierstrass transform. In addition, Appendix E contains an introduction to Fourier analysis that is used in the sequel.

2.2.2 Construction of the convolution method

For a dampening parameter $\alpha \in \mathbb{R}$ and any function f, we define the function f^α as

$$f^\alpha(x) = e^{-\alpha x} f(x). \tag{2.2.20}$$

Taking the Fourier transform of $\tilde{u}_i^\alpha$ gives

$$\begin{aligned}
\mathfrak{F}[\tilde{u}_i^\alpha](\nu) &= \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu x} e^{-\alpha x} \int_{-\infty}^{\infty} u_{i+1}(y) h(y-x)\,dy\,dx \\
&= \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu x} \int_{-\infty}^{\infty} u_{i+1}^\alpha(y) e^{\alpha(y-x)} h(y-x)\,dy\,dx \\
&= \mathfrak{F}[u_{i+1}^\alpha](\nu)\mathfrak{F}[e^{-\alpha z}h(-z)](\nu)
\end{aligned} \tag{2.2.21}$$

using the convolution theorem of Proposition E.5. Moreover

$$\begin{aligned}
\mathfrak{F}[e^{-\alpha z}h(-z)](\nu) &= \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu z} e^{-\alpha z} h(-z)\,dz \\
&= \int_{-\infty}^{\infty} e^{\mathbf{i}(\nu-\mathbf{i}\alpha)x} h(x)\,dx \\
&\quad \text{after the change of variable } x = -z, \\
&= \phi(\nu - \mathbf{i}\alpha)
\end{aligned} \tag{2.2.22}$$

where

$$\phi(\nu) = \exp\left(-\frac{1}{2}\Delta_i \nu^2\right) \tag{2.2.23}$$

is the characteristic function of the density h.

The equality of equation (2.2.22) is well-defined since $|\phi(\nu - \mathbf{i}\alpha)| < \infty$ for any $\alpha \in \mathbb{R}$. Nonetheless, the structure of the terminal condition g (and more generally, the preceding approximation u_{i+1}) will have a major impact in the choice of the dampening parameter α. Indeed, the equations (2.2.21) and (2.2.22) then lead to

$$\mathfrak{F}[\tilde{u}_i^\alpha](\nu) = \mathfrak{F}[u_{i+1}^\alpha](\nu)\phi(\nu - \mathbf{i}\alpha) \tag{2.2.24}$$

and hence the parameter α must be chosen so that the dampened functions u_i^α, $i = 0, 1, ..., n$, are integrable and admit Fourier transforms.

Using the same approach, the Fourier transform of $\dot{u}_i^\alpha$ gives

$$\begin{aligned}
\mathfrak{F}[\dot{u}_i^\alpha](\nu) &= -\frac{1}{\Delta_i}\mathfrak{F}[u_{i+1}^\alpha](\nu)\mathfrak{F}[ze^{-\alpha z}h(-z)](\nu) \\
&= -\frac{\mathbf{i}}{\Delta_i}\mathfrak{F}[u_{i+1}^\alpha](\nu)\frac{\partial}{\partial\nu}\mathfrak{F}[e^{-\alpha z}h(-z)](\nu) \\
&= -\frac{\mathbf{i}}{\Delta_i}\mathfrak{F}[u_{i+1}^\alpha](\nu)\frac{\partial}{\partial\nu}\phi(\nu-\mathbf{i}\alpha) \\
&= (\alpha+\mathbf{i}\nu)\mathfrak{F}[u_{i+1}^\alpha](\nu)\phi(\nu-\mathbf{i}\alpha). \qquad (2.2.25)
\end{aligned}$$

where the second equality holds by Proposition E.4.

From equations (2.2.24) and (2.2.25), we recover the functions $\tilde{u}_i$ and $\dot{u}_i$ by taking the inverse Fourier transform and adjusting for the dampening factor

$$\begin{aligned}
\tilde{u}_i(x) &= e^{\alpha x}\mathfrak{F}^{-1}\left[\mathfrak{F}[u_{i+1}^\alpha](\nu)\phi(\nu-\mathbf{i}\alpha)\right](x) \qquad (2.2.26) \\
\dot{u}_i(x) &= e^{\alpha x}\mathfrak{F}^{-1}\left[(\alpha+\mathbf{i}\nu)\mathfrak{F}[u_{i+1}^\alpha](\nu)\phi(\nu-\mathbf{i}\alpha)\right](x). \qquad (2.2.27)
\end{aligned}$$

It is possible to construct an alternative explicit characterization of the BSDE numerical solution. One may consider directly the explicit Euler scheme

$$\begin{cases}
Z_{t_n}^\pi = 0,\ Y_{t_n}^\pi = \xi^\pi \\
Z_{t_i}^\pi = \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^\pi \Delta W_i | \mathcal{F}_{t_i}\right] \\
Y_{t_i}^\pi = \mathbf{E}\left[Y_{t_{i+1}}^\pi + f(t_i, Y_{t_{i+1}}^\pi, Z_{t_i}^\pi)\Delta_i | \mathcal{F}_{t_i}\right]
\end{cases} \qquad (2.2.28)$$

to define the approximate solution of the PDE of equation (2.1.2). We refer to this scheme as the explicit Euler scheme 2. In this case, the approximate solution and the approximate gradient consist of functions v_i and $\dot{v}_i$ at mesh time t_i which take the form

$$v_i(x) = \int_{-\infty}^{\infty} \tilde{v}_{i+1}(y)h(y-x)dy \qquad (2.2.29)$$

where

$$\begin{aligned}
\tilde{v}_{i+1}(x) &= v_{i+1}(x) + \Delta_i f(t_i, v_{i+1}(x), \dot{v}_i(x)), \qquad (2.2.30) \\
\dot{v}_i(x) &= \int_{-\infty}^{\infty} (y-x)v_{i+1}(y)h(y-x)dy \qquad (2.2.31)
\end{aligned}$$

for $i = 0, 1, ..., n-1$ and $v_n(x) = g(x)$. Following the steps of the previous characterization, these equations naturally lead to

$$\begin{aligned}
v_i(x) &= e^{\alpha x}\mathfrak{F}^{-1}\left[\mathfrak{F}[\tilde{v}_{i+1}^\alpha](\nu)\phi(\nu-\mathbf{i}\alpha)\right](x) \qquad (2.2.32) \\
\dot{v}_i(x) &= e^{\alpha x}\mathfrak{F}^{-1}\left[(\alpha+\mathbf{i}\nu)\mathfrak{F}[v_{i+1}^\alpha](\nu)\phi(\nu-\mathbf{i}\alpha)\right](x). \qquad (2.2.33)
\end{aligned}$$

In this case, both v_i^α and $\tilde{v}_i^\alpha$ for $i = 0, 1, ..., n-1$ along with the dampened terminal condition are assumed to be integrable so that they admit Fourier transforms.

2.3 Implementation

As seen in Section 2.2, the numerical approximations of the BSDE solution can be expressed in terms of convolutions representing the conditional expectations involved in the explicit and implicit Euler schemes. We present, in this section, the numerical techniques which will allow us to compute the quadratures in the solution expressions.

2.3.1 Space discretization

From equations (2.2.26), (2.2.27),(2.2.32) and (2.2.33) one notices that computing the approximate solutions u_i and v_i at mesh time t_i reduces to computing a function $\theta : \mathbb{R} \to \mathbb{R}$ depending of two functions $\psi : \mathbb{C} \to \mathbb{C}$ and $\eta : \mathbb{R} \to \mathbb{R}$ in the following manner

$$\theta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathbf{i}\nu x} \hat{\eta^{\alpha}}(\nu)\psi(\nu)d\nu \tag{2.3.1}$$

if we drop the dampening factor $e^{\alpha x}$. [2]

This integral is numerically computed by discretizing the Fourier space with an uniform grid of $N+1$ points $\{\nu_i\}_{i=0}^{N}$ on the interval $[-\frac{L}{2}, \frac{L}{2}]$ of length L , where N is even, such that

$$\nu_i = \nu_0 + i\Delta\nu \tag{2.3.2}$$

where $\nu_0 = -\frac{L}{2}$ and $\Delta\nu = \frac{L}{N}$. Hence, for any $x \in \mathbb{R}$

$$\begin{aligned} \theta(x) &\approx \frac{1}{2\pi} \int_{-\frac{L}{2}}^{\frac{L}{2}} e^{\mathbf{i}\nu x} \hat{\eta^{\alpha}}(\nu)\psi(\nu)d\nu \\ &\approx \frac{\Delta\nu}{2\pi} \sum_{i=0}^{N-1} e^{\mathbf{i}\nu_i x} \hat{\eta^{\alpha}}(\nu_i)\psi(\nu_i) \end{aligned} \tag{2.3.3}$$

where the integral is approximated using lower Riemann sums and

$$\hat{\eta^{\alpha}}(\nu_i) = \int_{-\infty}^{\infty} e^{-\mathbf{i}x\nu_i} \eta^{\alpha}(x)dx = \int_{-\infty}^{\infty} e^{-\mathbf{i}x\nu_i} e^{-\alpha x}\eta(x)dx. \tag{2.3.4}$$

This last integral is also computed using a uniform grid of $N+1$ points $\{x_j\}_{j=0}^{N}$ such that

$$x_j = x_0 + j\Delta x \tag{2.3.5}$$

where Δx is chosen so that the Nyquist relation[3] is satisfied, i.e

$$\Delta x = \frac{2\pi}{L}. \tag{2.3.6}$$

We approximate the integral of equation (2.3.4) by first restricting the integration interval to $[x_0, x_N] = [-\frac{l}{2}, \frac{l}{2}]$ and then applying a composite quadrature rule with weights $\{w_i\}_{i=0}^{N}$ so that

$$\begin{aligned} \hat{\eta^{\alpha}}(\nu_i) &\approx \int_{x_0}^{x_N} e^{-\mathbf{i}x\nu_i} \eta^{\alpha}(x)dx \\ &\approx \Delta x \sum_{j=0}^{N} w_j e^{-\mathbf{i}x_j\nu_i} \eta^{\alpha}(x_j) \\ &= \Delta x \cdot e^{-\mathbf{i}x_0\nu_i} \sum_{j=0}^{N} w_j e^{-\mathbf{i}ji\frac{2\pi}{N}} e^{-\mathbf{i}j\nu_0\Delta x} \eta^{\alpha}(x_j) \\ &= \Delta x \cdot e^{-\mathbf{i}x_0\nu_i} \left(\sum_{j=0}^{N-1} w_j e^{-\mathbf{i}ji\frac{2\pi}{N}} e^{-\mathbf{i}j\nu_0\Delta x} \eta^{\alpha}(x_j) + w_N \eta^{\alpha}(x_N) \right) \end{aligned} \tag{2.3.7}$$

since N is even. Assuming that

$$\eta^{\alpha}(x_0) = \eta^{\alpha}(x_N), \tag{2.3.8}$$

[2] For a suitable function η, the transform defining the function θ can be interpreted as dampened conditional expectation of η or $\nabla\eta$ depending of the function ψ.

[3] In its explicit form, we have $\Delta\nu \cdot \Delta x = \frac{2\pi}{N}$. One may then choose Δx first and retrieve $\Delta\nu$ form the Nyquist relation.

we have, for $i = 0, 1, ..., N-1$,

$$\hat{\eta^\alpha}(\nu_i) \approx \frac{2\pi}{\Delta\nu} e^{-\mathbf{i}x_0\nu_i} \mathfrak{D}\left[\{(-1)^j \tilde{w}_j \eta^\alpha(x_j)\}_{j=0}^{N-1}\right]_i \tag{2.3.9}$$

since $e^{-\mathbf{i}\nu_0\Delta x} = -1$ with

$$\tilde{w}_j = w_j + \delta_{N-j,N} w_N \tag{2.3.10}$$

where $\delta_{i,j}$ stands for the Kronecker's delta.

A similar approach can be found in Lord et al. [76] who enhance the discrete Fourier transform with a composite trapezoidal quadrature rule to compute this last integral. However, the authors omit the assumption of equation (2.3.8) leading to considerable numerical errors, especially around the boundaries of the restricted domain $[x_0, x_N]$.

Note that

$$(\alpha + \mathbf{i}\nu)\mathfrak{F}[\eta^\alpha](\nu) = \mathfrak{F}\left[\alpha\eta^\alpha + \frac{\partial \eta^\alpha}{\partial x}\right](\nu)$$

and hence, when using $\psi(\nu) = (\alpha + \mathbf{i}\nu)\phi(\nu - \mathbf{i}\alpha)$, the Fourier coefficients of the derivative $\frac{\partial \eta^\alpha}{\partial x}$ are implicitly considered. For this reason, we also assume that

$$\frac{\partial \eta^\alpha}{\partial x}(x_0) = \frac{\partial \eta^\alpha}{\partial x}(x_N) \tag{2.3.11}$$

for a differentiable function η.

The values of the function θ are computed at the grid points $\{x_k\}_{k=0}^{N-1}$ by combining equations (2.3.3) and (2.3.9)

$$\begin{aligned}
\theta(x_k) &\approx \sum_{j=0}^{N-1} e^{\mathbf{i}\nu_j x_k} \psi(\nu_j) e^{-\mathbf{i}x_0\nu_j} \mathfrak{D}\left[\{(-1)^i \tilde{w}_i \eta^\alpha(x_i)\}_{i=0}^{N-1}\right]_j \\
&= e^{\mathbf{i}k\nu_0\Delta x} \sum_{j=0}^{N-1} e^{\mathbf{i}jk\frac{2\pi}{N}} \psi(\nu_j) \mathfrak{D}\left[\{(-1)^i \tilde{w}_i \eta^\alpha(x_i)\}_{i=0}^{N-1}\right]_j \\
&= (-1)^k \mathfrak{D}^{-1}\left[\left\{\psi(\nu_j)\mathfrak{D}\left[\{(-1)^i \tilde{w}_i \eta^\alpha(x_i)\}_{i=0}^{N-1}\right]_j\right\}_{j=0}^{N-1}\right]_k .
\end{aligned} \tag{2.3.12}$$

Since we use the DFT, the underlying trigonometric (and hence periodic) interpolation allows us to set

$$\theta(x_N) = \theta(x_0). \tag{2.3.13}$$

We shall see, in the following sub-section how to adjust the method to consider functions that do not satisfy the conditions of equations (2.3.8) and (2.3.11).

2.3.2 Numerical considerations

The integrability condition on the dampened approximate solutions (u_i^α, v_i^α and $\tilde{v}_i^\alpha$), particularly on the terminal value function g^α, may seem too constraining since it narrows the scope of BSDEs that can be numerically solved by the method. In fact, the condition has no effect on the numerical method since we are performing a domain truncation when valuing numerically the Fourier transform of the dampened approximate solutions by the DFT in equation (2.3.7). Indeed, this integral truncation imposes an integrability condition only on the truncated function

$$\Xi(x) = \eta^\alpha(x)\mathbf{1}_{[x_0,x_n]}(x) \tag{2.3.14}$$

where

$$\mathbf{1}_A(x) = \begin{cases} 1 \text{ if } x \in A \\ 0 \text{ if } x \notin A \end{cases} \tag{2.3.15}$$

is the indicator function. The procedure is equivalent to solving the BSDE on the restricted domain $[x_0, x_N]$.

An important step in the convolution method for BSDEs presented here is the choice of the parameter α used for dampening purposes. As already mentioned, the optimal dampening parameter strongly depends on the structure of the function η and will considerably improve the accuracy of the approximation of equation (2.3.9). Indeed, the DFT algorithm gives better results when α is chosen such that

$$\eta^\alpha(x_0) = \eta^\alpha(x_N). \tag{2.3.16}$$

Since we intend to solve not only for the PDE solution u but also for its derivative ∇u, we also need

$$\frac{\partial \eta^\alpha}{\partial x}(x_0) = \frac{\partial \eta^\alpha}{\partial x}(x_N) \tag{2.3.17}$$

where we assume that the function η is differentiable at least at the boundaries of the restricted domain.

Note that, even for very simple terminal value functions g the condition of equation (2.3.16) may be impossible to satisfy. A straightforward example is provided by the function $g(x) = x$ on the interval $[x_0, x_N] = [-1, 1]$. In order to address this problem, we slightly modify the function η by adding a linear function to get the modified dampened function $\eta^\alpha_{\beta,\kappa}$ defined as

$$\eta^\alpha_{\beta,\kappa}(x) = e^{-\alpha x}(\eta(x) + \beta x + \kappa). \tag{2.3.18}$$

The following lemma gives the optimal choice for the dampening parameter $\alpha \in \mathbb{R}$, and the coefficients $\beta \in \mathbb{R}$ and $\kappa \in \mathbb{R}$.

Lemma 2.4. *Suppose the real function $\eta \in \mathcal{C}^1[a, b]$ is differentiable with*

$$\frac{\partial \eta}{\partial x}(a) \neq \frac{\partial \eta}{\partial x}(b)$$

and let $\eta^\alpha_{\beta,\kappa}$ be its dampened and modified function as defined in equation (2.3.18). Then

$$\alpha = \frac{1}{b-a} \log\left(\frac{\frac{\partial \eta}{\partial x}(b) + \beta}{\frac{\partial \eta}{\partial x}(a) + \beta}\right), \tag{2.3.19}$$

$$\kappa = \frac{e^{-\alpha b}(\eta(b) + \beta b) - e^{-\alpha a}(\eta(a) + \beta a)}{e^{-\alpha a} - e^{-\alpha b}} \tag{2.3.20}$$

solve the system of nonlinear equations

$$\begin{cases} \eta^\alpha_{\beta,\kappa}(a) = \eta^\alpha_{\beta,\kappa}(b) \\ \frac{\partial \eta^\alpha_{\beta,\kappa}}{\partial x}(a) = \frac{\partial \eta^\alpha_{\beta,\kappa}}{\partial x}(b) \end{cases} \tag{2.3.21}$$

for any $\beta \notin \{\frac{\partial \eta}{\partial x}(a), \frac{\partial \eta}{\partial x}(b)\}$. If, in addition,

$$\beta > \max\left(|\frac{\partial \eta}{\partial x}(b)|, |\frac{\partial \eta}{\partial x}(a)|\right) \tag{2.3.22}$$

then also $\alpha \in \mathbb{R}$ and $\kappa \in \mathbb{R}$.

Proof. The first equation of the system (2.3.21) gives (2.3.20) in a straightforward manner. Since η is differentiable, $\eta^\alpha_{\beta,\kappa}$ is also differentiable and

$$\frac{\partial \eta^\alpha_{\beta,\kappa}}{\partial x}(x) = -\alpha \eta^\alpha_{\beta,\kappa}(x) + e^{-\alpha x}\left(\frac{\partial \eta}{\partial x}(x) + \beta\right)$$

and the system (2.3.21) leads to (2.3.19). Clearly, if the inequality (2.3.22) holds then both $\frac{\partial \eta}{\partial x}(b) + \beta$ and $\frac{\partial \eta}{\partial x}(a) + \beta$ are strictly positive and $\alpha \in \mathbb{R}$. □

The transform of equation (2.3.18) may seem over parametrized since we use three parameters to satisfy only two conditions. However, using only two parameters may lead to complex parameters or to an inconsistent system.

Remark 2.5. When implementing the method, the values of derivative $\frac{\partial \eta}{\partial x}$ at x_0 and x_N can be approximated by finite difference. We use the second order forward (resp. backward) finite difference when estimating $\frac{\partial \eta}{\partial x}(x_0)$ (resp. $\frac{\partial \eta}{\partial x}(x_N)$) as follows

$$\frac{\partial \eta}{\partial x}(x_0) = \frac{-3\eta(x_0) + 4\eta(x_1) - \eta(x_2)}{2\Delta x} + \mathcal{O}(\Delta x^2) \tag{2.3.23}$$

$$\frac{\partial \eta}{\partial x}(x_N) = \frac{3\eta(x_N) - 4\eta(x_{N-1}) + \eta(x_{N-2})}{2\Delta x} + \mathcal{O}(\Delta x^2). \tag{2.3.24}$$

Also, one needs a positive constant, which represent the minimal slope allowed in the linear transform $\beta x + \kappa$, say $\epsilon > 0$, as an input. Set

$$\beta = \epsilon + \max\left(|\frac{\partial \eta}{\partial x}(x_N)|, |\frac{\partial \eta}{\partial x}(x_0)| \right). \tag{2.3.25}$$

Whenever $\frac{\partial \eta}{\partial x}(x_N) = \frac{\partial \eta}{\partial x_0}(b)$, one can set $\alpha = \kappa = 0$ and

$$\beta = -\frac{\eta(x_N) - \eta(x_0)}{x_N - x_0}. \tag{2.3.26}$$

Under the transformation of equation (2.3.18), the computation of our approximate solution is not significantly more complex. One just has to make simple adjustments for the coefficient $\beta \in \mathbb{R}$ and $\kappa \in \mathbb{R}$. For both Euler schemes of equations (2.2.1) and (2.2.28), properties of the conditional expectation allows the adjustments and the following theorem gives their essence.

Theorem 2.6. *Let $\eta : [a, b] \to \mathbb{R}$ be an integrable function and define $\eta_{\beta,\alpha} : [a, b] \to \mathbb{R}$ as*

$$\eta_{\beta,\kappa}(x) = \eta(x) + \beta x + \kappa$$

such that $\eta^{\alpha}_{\beta,\kappa}$ is the dampened and modified function of η according to equation (2.3.18). Then the function $\theta : [a, b] \to \mathbb{R}$ of equation (2.3.1) admits the alternative representation

$$\theta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathbf{i}\nu x} \widehat{\eta^{\alpha}_{\beta,\kappa}}(\nu)\psi(\nu)d\nu - e^{-\alpha x}\beta \tag{2.3.27}$$

if $\psi(\nu) = (\alpha + \mathbf{i}\nu)\phi(\nu - \mathbf{i}\alpha)$ or

$$\theta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathbf{i}\nu x} \widehat{\eta^{\alpha}_{\beta,\kappa}}(\nu)\psi(\nu)d\nu - e^{-\alpha x}(\beta x + \kappa) \tag{2.3.28}$$

if $\psi(\nu) = \phi(\nu - \mathbf{i}\alpha)$.

Proof. First, let $\psi(\nu) = (\alpha + \mathbf{i}\nu)\phi(\nu - \mathbf{i}\alpha)$. By definition, we know that

$$\begin{aligned}
\theta(x) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathbf{i}\nu x} \hat{\eta^{\alpha}}(\nu)\psi(\nu)d\nu \\
&= \frac{e^{-\alpha x}}{\Delta_i} \mathbf{E}\left[\eta(W_{t_{i+1}})\Delta W_i | W_{t_i} = x\right] \\
&= \frac{e^{-\alpha x}}{\Delta_i} \left(\mathbf{E}\left[\left(\eta(W_{t_{i+1}}) + \beta W_{t_{i+1}} + \kappa\right)\Delta W_i | W_{t_i} = x\right] - \beta\Delta_i\right) \\
&= \frac{e^{-\alpha x}}{\Delta_i} \mathbf{E}\left[\eta_{\beta,\kappa}(W_{t_{i+1}})\Delta W_i | W_{t_i} = x\right] - e^{-\alpha x}\beta \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathbf{i}\nu x} \widehat{\eta^{\alpha}_{\beta,\kappa}}(\nu)\psi(\nu)d\nu - e^{-\alpha x}\beta.
\end{aligned}$$

Similarly, if $\psi(\nu) = \phi(\nu - \mathbf{i}\alpha)$, we have

$$\begin{aligned}\theta(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\eta^{\alpha}}(\nu)\psi(\nu)d\nu \\ &= e^{-\alpha x}\mathbf{E}\left[\eta(W_{t_{i+1}})|W_{t_i} = x\right] \\ &= e^{-\alpha x}\mathbf{E}\left[\eta_{\beta,\kappa}(W_{t_{i+1}})|W_{t_i} = x\right] - e^{-\alpha x}(\beta x + \kappa) \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha}_{\beta,\kappa}}(\nu)\psi(\nu)d\nu - e^{-\alpha x}(\beta x + \kappa).\end{aligned}$$

□

The solution for the control process $Z^{\pi}_{t_i}$ satisfies

$$Z^{\pi}_{t_i} = \frac{1}{\Delta_i}\mathbf{E}_{t_i}\left[\left(Y^{\pi}_{t_{i+1}} + \beta W_{t_{i+1}} + \kappa\right)\Delta W_i\right] - \beta \tag{2.3.29}$$

for any constants $\beta, \kappa \in \mathbb{R}$ and at any time step t_i, $i = 0, 1, ..., n-1$. Its conditional expectation must be shifted downward by β when using the transform function $\eta^{\alpha}_{\beta,\kappa}$ instead of η^{α}. Thus, equation (2.3.1) can be replaced by

$$\begin{aligned}\theta(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\eta^{\alpha}}(\nu)\psi(\nu)d\nu \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha}_{\beta,\kappa}}(\nu)\psi(\nu)d\nu - e^{-\alpha x}\beta\end{aligned} \tag{2.3.30}$$

whenever $\psi(\nu) = (\alpha + \mathbf{i}\nu)\phi(\nu - \mathbf{i}\alpha)$ and the integral of this last equation is computed by the method presented in equation (2.3.12) after optimizing for α, β and κ.

Moreover, for any $\mathcal{F}_{t_{i+1}}$-measurable random variable Y we have

$$\mathbf{E}_{t_i}[Y] = \mathbf{E}_{t_i}\left[Y + \beta W_{t_{i+1}} + \kappa\right] - \beta W_{t_i} - \kappa. \tag{2.3.31}$$

For both proposed schemes, the conditional expectations involved in the expression of the forward solution $Y^{\pi}_{t_i}$ must be corrected by the linear function $\beta x + \kappa$. Equivalently, the relation of equation (2.3.1) is replaced by

$$\begin{aligned}\theta(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\eta^{\alpha}}(\nu)\psi(\nu)d\nu \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha}_{\beta,\kappa}}(\nu)\psi(\nu)d\nu - e^{-\alpha x}(\beta x + \kappa)\end{aligned} \tag{2.3.32}$$

whenever $\psi(\nu) = \phi(\nu - \mathbf{i}\alpha)$ where the integral of the last equation is computed with the method of equation (2.3.12).

As to the interval length of the frequency domain L, we set it large enough so that the value of the characteristic function ϕ is approximately zero (0) at the boundaries of the truncated frequency domain $[-\frac{L}{2}, \frac{L}{2}]$. Ideally, the truncated real space $[x_0, x_N]$ is centered around zero (0) or more generally around the initial value of the Brownian motion $\{W_t\}_{t\in[0,T]}$.

The structure of the BSDE driver f may have an undesirable effect on the algorithm, specially when it is a non-smooth function. This problem can be solved simply by reducing the time step Δ_i. As a general rule, the larger is the Lipschitz constant of the driver f, the smaller should the time step Δ_i be. Due to the efficiency of the DFT algorithm, the method is well adapted for relatively large values of n, the number of time steps.

Many quadrature rules are available to compute the approximate solution and approximate gradient values. One may use the composite trapezoidal rule with weights of the form

$$w_i = 1 - \frac{1}{2}(\delta_{0,i} + \delta_{N,i}) \ , \ i = 0, 1, ..., N \tag{2.3.33}$$

leading to $\tilde{w}_i = 1$. The composite Simpson rule will improve accuracy in presence of a smooth driver f.

Finally, the implementation of the convolution method gives the approximation values $\{u_{ik}\}_{k=0}^{N}$ and $\{\dot{u}_{ik}\}_{k=0}^{N}$ to the approximate solutions u_i and $\dot{u}_i$ for $i = 0, 1, 2, ..., n-1$. The detailed algorithm is as follows:

Algorithm 2.1. *Convolution Method*

1. *Discretize the restricted real space* $[-\frac{l}{2}, \frac{l}{2}]$ *and the restricted Fourier space* $[-\frac{L}{2}, \frac{L}{2}]$ *with* N *space steps so to have the real space nodes* $\{x_k\}_{k=0}^{N}$ *and the Fourier space nodes* $\{\nu_k\}_{k=0}^{N}$
2. *Set* $u_n(x_k) = g(x_k)$
3. *For any* i *from* $n-1$ *to* 0
   (a) *Compute* α, β *and* κ, *defined in equation (2.3.18), such that*
   $$\eta = (u_{i+1})_{\beta,\kappa}^{\alpha} \tag{2.3.34}$$
   and η *satisfies the boundary conditions.*
   (b) *Compute* $\theta(x_k)$ *through equation (2.3.12) for* $k = 0, 1, ..., N$ *with*
   $$\psi(\nu) = \phi(\nu - \mathbf{i}\alpha) \tag{2.3.35}$$
   and retrieve the values $\tilde{u}_{ik}$ *as*
   $$\tilde{u}_{ik} = e^{\alpha x_k}\theta(x_k) - (\beta x_k + \kappa) \tag{2.3.36}$$
   through Theorem 2.6.
   (c) *Compute* $\theta(x_{ik})$ *through equation (2.3.12) for* $k = 0, 1, ..., N$ *with*
   $$\psi(\nu) = (\alpha + \mathbf{i}\nu)\phi(\nu - \mathbf{i}\alpha) \tag{2.3.37}$$
   and retrieve the values $\dot{u}_{ik}$ *as*
   $$\dot{u}_{ik} = e^{\alpha x_k}\theta(x_k) - \beta \tag{2.3.38}$$
   through Theoreom 2.6.
   (d) *Compute the values* u_{ik} *as*
   $$u_{ik} = \tilde{u}_{ik} + \Delta_i f(t_i, \tilde{u}_{ik}, \dot{u}_{ik}) \tag{2.3.39}$$
   for $k = 0, 1, ..., N_i N$ *through equation (2.2.6) when using the explicit Euler scheme 1 or as*
   $$u_{ik} = \tilde{u}_{ik} + \Delta_i f(t_i, u_{ik}, \dot{u}_{ik}) \tag{2.3.40}$$
   through equation (2.2.12) under the implicit Euler scheme.

Under the implicit Euler scheme, the node values u_{ik} for the approximate solution solve a non-linear system of equation (2.2.12). For small time steps, and particularly if the condition of equation (2.2.13) is satisfied, the node values u_{ik} can be computed iteratively using Picard iterations. When the node values are available under the implicit or explicit Euler scheme, an approximate solution to the BSDE consists of a (linear) interpolation of a Brownian path values through the node values $\{u_{ik}\}_{k=0}^{N}$ and $\{\dot{u}_{ik}\}_{k=0}^{N}$ for $i = 0, 1, 2, ..., n-1$. The problem of simulation is treated with more details in Chapter 3. the next section deals with the problem space discretization errors.

2.4 Local space discretization error

The convolution method induces two (2) main types of error. Aside from the time discretization error E_π that we already discussed in the introduction, we have a space discretization error. We focuses on the study of this last error term. We limit the error analysis to the convolution method on the explicit Euler scheme 1 since equivalent results are easily obtained for explicit Euler scheme 2 and the implicit Euler scheme using the same techniques. When considering the implicit scheme, the condition of equation (2.2.13) must be satisfied to assure existence and uniqueness of the implicit approximate solution u_i.

Throughout the section, $\{\mathbf{u}_{ik}\}_{k=0}^N$, $\{\tilde{\mathbf{u}}_{ik}\}_{k=0}^N$ and $\{\dot{\mathbf{u}}_{ik}\}_{k=0}^N$ denote the numerical solution of equation (2.3.12) obtained from the convolution method at time mesh t_i given the solution u_{i+1} at time t_{i+1}, $i = 0, 1, ..., n-1$. The convolution method induces a space discretization error when approximating the values of $u_i(x_k)$ and $\dot{u}_i(x_k)$ by $\mathbf{u}_{ik}$ and $\dot{\mathbf{u}}_{ik}$ respectively. We will particularly describe the local behavior of this error term. We define it as

$$E_{ik} := |u_i(x_k) - \mathbf{u}_{ik}| + |\dot{u}_i(x_k) - \dot{\mathbf{u}}_{ik}|. \tag{2.4.1}$$

The following theorem gives an error bound for the space discretization error under smoothness conditions on the BSDE coefficients f and g.

Theorem 2.7. *Suppose $f \in \mathcal{C}^{1,2,2}$ and $g \in \mathcal{C}^2$. Then for any $i = 0, 1, ..., n-1$ and $k = 0, 1, ..., N$, the convolution method applied on the truncated interval $\left[-\frac{l}{2}, \frac{l}{2}\right]$ yields a (local) discretization error of the form*

$$E_{ik} = \chi(x_k) + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-K\Delta_i^{-1}l^2}\right) \tag{2.4.2}$$

where the extrapolation error χ satifies

$$|\chi(x_k)| \leq C\left(\int_{\frac{l}{2}-|x_k|}^{\frac{l}{2}} h(y)dy\right)^{\frac{1}{2}} \tag{2.4.3}$$

for some positive constants $C, K > 0$ depending on the driver f, the terminal function g and T when using the trapezoidal quadrature rule.

Proof. Suppose the solution u_{i+1} at time t_{i+1} is known. Since $f \in \mathcal{C}^{1,2,2}$ and $g \in \mathcal{C}^2$, it is easily shown that $u_{i+1} \in \mathcal{C}^2$. Also, we know from Zhang [124] and Bouchard and Touzi [20] that $Y_{t_{i+1}}^\pi = u_{i+1}(W_{t_{i+1}})$ is square integrable so that u_{i+1} is square integrable (with respect to the Gaussian density).

In the light of Theorem 2.6, we can limit ourselves to the case where

$$u_{i+1}\left(-\frac{l}{2}\right) = u_{i+1}\left(\frac{l}{2}\right) \text{ and } \frac{\partial u_{i+1}}{\partial x}\left(-\frac{l}{2}\right) = \frac{\partial u_{i+1}}{\partial x}\left(\frac{l}{2}\right)$$

so that $\alpha = \beta = \kappa = 0$. Let $\{c_k\}_{k=-\infty}^\infty$ be the Fourier coefficients of u_{i+1} on $\left[-\frac{l}{2}, \frac{l}{2}\right]$. We have that

$$\begin{aligned}
\tilde{u}_i(x_k) &= \int_{|y-x_k|\leq\frac{l}{2}} u_{i+1}(y)h(y-x_k)dy + \int_{|y-x_k|>\frac{l}{2}} u_{i+1}(y)h(y-x_k)dy \\
&= \int_{|y|\leq\frac{l}{2}} u_{i+1}(x_k+y)h(y)dy + \int_{|y|>\frac{l}{2}} u_{i+1}(x_k+y)h(y)dy
\end{aligned}$$

where

$$\int_{|y|>\frac{l}{2}} u_{i+1}(x_k+y)h(y)dy = \mathbf{E}\left[u_{i+1}(x_k+\Delta W_{n-1})\mathbf{1}_{\mathbb{R}\setminus[-\frac{l}{2},\frac{l}{2}]}(\Delta W_{n-1})\right]$$

$$= \mathcal{O}\left(e^{-Kl^2}\right)$$

for some constant $K > 0$ which is inversely proportional to Δ_i by Cauchy-Schwartz and Chernoff inequalities since the solution u_{i+1} is square integrable. Hence

$$\begin{aligned}\tilde{u}_i(x_k) &= \int_{|y|\leq\frac{l}{2}} T_\infty(x_k+y)h(y)dy \\ &\quad + \int_{|y|\leq\frac{l}{2}} \left(u_{i+1}(x_k+y) - T_\infty(x_k+y)\right)h(y)dy \\ &\quad + \mathcal{O}\left(e^{-Kl^2}\right) \end{aligned} \tag{2.4.4}$$

where $T_\infty(x) = \sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}k\frac{2\pi}{l}x}$ for $x \in \mathbb{R}$. So that, on one hand, we have

$$\begin{aligned}&\int_{|y|\leq\frac{l}{2}} T_\infty(x_k+y)h(y)dy \\ &= \sum_{j=-\infty}^{\infty} c_j e^{\mathbf{i}j\frac{2\pi}{l}x_k}\phi\left(j\frac{2\pi}{l}\right) - \int_{|y|>\frac{l}{2}} T_\infty(x_k+y)h(y)dy \\ &= \sum_{j=-\frac{N}{2}}^{\frac{N}{2}-1} c_j e^{\mathbf{i}j\frac{2\pi}{l}x_k}\phi\left(j\frac{2\pi}{l}\right) - \int_{|y|>\frac{l}{2}} T_\infty(x_k+y)h(y)dy + \mathcal{O}(\Delta x) \\ &\quad \text{(by Proposition E.6)}, \\ &= \sum_{j=-\frac{N}{2}}^{\frac{N}{2}-1} \phi\left(j\frac{2\pi}{l}\right) c_j e^{\mathbf{i}j\frac{2\pi}{l}x_k} + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right) \\ &\quad \text{(by boundedness of } T_\infty \text{ and Chernoff inequalities)}, \\ &= (-1)^k \sum_{j=0}^{N-1} \phi(\nu_j)(-1)^{j-\frac{N}{2}} c_{j-\frac{N}{2}} e^{\mathbf{i}\frac{2\pi}{N}jk} + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right) \\ &= (-1)^k \sum_{j=0}^{N-1} \phi(\nu_j)\mathcal{D}\left[\{(-1)^s u_{i+1}(x_s)\}_{s=0}^{N-1}\right]_j e^{\mathbf{i}\frac{2\pi}{N}jk} \\ &\quad + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right) \text{ (by Proposition E.8)}, \\ &= \tilde{\mathbf{u}}_{ik} + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right). \end{aligned}$$

On the other hand, assuming $x_k \geq 0$ without loss of generality, let's define χ_0 as

$$\begin{aligned}\chi_0(x_k) &= \int_{|y|\leq\frac{l}{2}} \left(u_{i+1}(x_k+y) - T_\infty(x_k+y)\right)h(y)dy \\ &= \int_{\frac{l}{2}-x_k}^{\frac{l}{2}} \left(u_{i+1}(x_k+y) - u_{i+1}(x_k+y-l)\right)h(y)dy \end{aligned}$$

since T_∞ is periodic and $T_\infty(x) = u_{i+1}(x)$ on the interval $[-\frac{l}{2}, -\frac{l}{2}]$. Equation (2.4.4) them becomes

$$\tilde{u}_i(x_k) = \tilde{\mathbf{u}}_{ik} + \chi_0(x_k) + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right) \tag{2.4.5}$$

and we notice that, by the continuity of u_{i+1},

$$|\chi_0(x_k)| \leq C_0 \int_{\frac{l}{2}-|x_k|}^{\frac{l}{2}} h(y)dy \tag{2.4.6}$$

for some positive constant $C_0 > 0$ which is independent of Δ_i.

Similarly

$$\dot{u}_i(x_k) = \frac{1}{\Delta_i}\int_{|y|\leq\frac{l}{2}} u_{i+1}(x_k+y)yh(y)dy + \frac{1}{\Delta_i}\int_{|y|>\frac{l}{2}} u_{i+1}(x_k+y)yh(y)dy$$

where

$$\begin{aligned}
&\frac{1}{\Delta_i}\int_{|y|>\frac{l}{2}} u_{i+1}(x_k+y)yh(y)dy \\
&= \frac{1}{\Delta_i}\mathbf{E}\left[u_{i+1}(x_k+\Delta W_{n-1})\Delta W_{n-1}\mathbf{1}_{\mathbb{R}\setminus[-\frac{l}{2},\frac{l}{2}]}(\Delta W_{n-1})\right] \\
&\leq \frac{K}{\Delta_i}\mathbf{E}\left[(\Delta W_{n-1})^2\mathbf{1}_{\mathbb{R}\setminus[-\frac{l}{2},\frac{l}{2}]}(\Delta W_{n-1})\right]^{\frac{1}{2}} \\
&\quad \text{(by Chauchy-Schwartz inequality),} \\
&= \mathcal{O}\left({\Delta_i}^{-\frac{1}{2}}e^{-K{\Delta_i}^{-1}l^2}\right) \\
&\quad \text{(by successively applying Cauchy-Schwartz and Chernoff inequalities),} \\
&= \mathcal{O}\left(e^{-\frac{1}{2}K{\Delta_i}^{-1}l^2}\right).
\end{aligned}$$

Hence

$$\begin{aligned}
\dot{u}_i(x_k) &= \frac{1}{\Delta_i}\int_{|y|\leq\frac{l}{2}} T_\infty(x_k+y)yh(y)dy + \mathcal{O}\left(e^{-K{\Delta_i}^{-1}l^2}\right) \\
&\quad + \frac{1}{\Delta_i}\int_{|y|\leq\frac{l}{2}} \left(u_{i+1}(x_k+y)-T_\infty(x_k+y)\right)yh(y)dy. \qquad (2.4.7)
\end{aligned}$$

Letting $c\prime_j = \mathbf{i}j\frac{2\pi}{l}c_j$, we have

$$\begin{aligned}
&\frac{1}{\Delta_i}\int_{|y|\leq\frac{l}{2}} T_\infty(x_k+y)yh(y)dy \\
&= \sum_{j=-\infty}^{\infty} c\prime_j e^{\mathbf{i}j\frac{2\pi}{l}x_k}\phi\left(j\frac{2\pi}{l}\right) - \frac{1}{\Delta_i}\int_{|y|>\frac{l}{2}} T_\infty(x_k+y)yh(y)dy \\
&= \sum_{j=-\frac{N}{2}}^{\frac{N}{2}-1} c\prime_j e^{\mathbf{i}j\frac{2\pi}{l}x_k}\phi\left(j\frac{2\pi}{l}\right) - \frac{1}{\Delta_i}\int_{|y|>\frac{l}{2}} T_\infty(x_k+y)yh(y)dy + \mathcal{O}(\Delta x) \\
&= \sum_{j=-\frac{N}{2}}^{\frac{N}{2}-1} \phi\left(j\frac{2\pi}{l}\right) c\prime_j e^{\mathbf{i}j\frac{2\pi}{l}x_k} + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-K{\Delta_i}^{-1}l^2}\right) \\
&\quad \text{(by boundedness of } T_\infty \text{ and Chernoff inequality),} \\
&= (-1)^k\sum_{j=0}^{N-1}\phi(\nu_j)(-1)^{j-\frac{N}{2}}c\prime_{j-\frac{N}{2}}e^{\mathbf{i}\frac{2\pi}{N}jk} + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-K{\Delta_i}^{-1}l^2}\right) \\
&= (-1)^k\sum_{j=0}^{N-1}\mathbf{i}\nu_j\phi(\nu_j)\mathcal{D}\left[\{(-1)^s u_{i+1}(x_s)\}_{s=0}^{N-1}\right]_j e^{\mathbf{i}\frac{2\pi}{N}jk} \\
&\quad + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-K{\Delta_i}^{-1}l^2}\right) \text{ (by Proposition E.8),} \\
&= \dot{u}_{ik} + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-K{\Delta_i}^{-1}l^2}\right). \qquad (2.4.8)
\end{aligned}$$

By equations (2.4.7) and (2.4.8)

$$\dot{u}_i(x_k) = \dot{u}_{ik} + \chi_1(x_k) + \mathcal{O}\left(\Delta x\right) + \mathcal{O}\left(e^{-K{\Delta_i}^{-1}l^2}\right) \qquad (2.4.9)$$

where $K>0$ and, letting $\upsilon(y) = u_{i+1}(x_k+y) - T_\infty(x_k+y)$,

$$\begin{aligned}
\chi_1(x_k) &= \Delta_i^{-1}\int_{|y|\leq\frac{l}{2}} y\upsilon(y)h(y)dy \\
&= \Delta_i^{-1}\int_{|y|\leq\frac{l}{2}} y^2\frac{\upsilon(y)-\upsilon(0)}{y}h(y)dy \\
&= \Delta_i^{-1}\int_{|y|\leq\frac{l}{2}} y^2\left(\frac{\partial\upsilon}{\partial x}(y) + \frac{\partial^2\upsilon}{\partial x^2}(\xi)y\right)h(y)dy \text{ (for some } \xi\in\left[-\frac{l}{2},\frac{l}{2}\right]),
\end{aligned}$$

$$
\begin{aligned}
&= \Delta_i^{-1} \int_{|y|\leq \frac{l}{2}} y^2 \frac{\partial v}{\partial x}(y) h(y) dy \text{ (by symmetry)}, \\
&= \Delta_i^{-1} \int_{|y|\leq \frac{l}{2}} y^2 \left(\frac{\partial u_{i+1}}{\partial x}(x_k + y) - \frac{\partial T_\infty}{\partial x}(x_k + y) \right) h(y) dy.
\end{aligned}
$$

Since $\frac{\partial T_\infty}{\partial x}$ is the Fourier expansion of $\frac{\partial u_{i+1}}{\partial x}$, we get

$$
\begin{aligned}
|\chi_1(x_k)| &\leq C_1 \Delta_i^{-1} \int_{\frac{l}{2}-|x_k|}^{\frac{l}{2}} y^2 h(y) dy \\
&\leq C_1 \left(\int_{\frac{l}{2}-|x_k|}^{\frac{l}{2}} h(y) dy \right)^{\frac{1}{2}} \qquad (2.4.10)
\end{aligned}
$$

by the boundedness of $\frac{\partial u_{i+1}}{\partial x}$ and Chauchy-Schwartz inequality, for some constant $C_1 > 0$.

The Lipschitz property of the driver f completes the proof from the relations in equations (2.4.5), (2.4.6), (2.4.9) and (2.4.10). □

Theorem 2.7 decomposes the spatial discretization error in three parts: the truncation error, the discretization error and the extrapolation error. Most PDE based and spatial discretization based methods for BSDEs fail in giving a bound for the error due to truncation. The error analysis shows that the truncation error $\mathcal{O}(e^{-K\Delta_i^{-1} l^2})$ has a spectral convergence of index 2 when applying the convolution method. Also, the discretization error $\mathcal{O}(\Delta x)$, of first order, is similar to other PDE based methods such as Douglas et al. [40] or Milstein and Tretyakov [90].

It is important to notice that the convolution method presented here is a (Fourier) spectral as shown in the proof of Theorem 2.7. Hence, the space discretization error is actually also spectral when the BSDE coefficients are smooth $f \in \mathcal{C}^\infty$ and $g \in \mathcal{C}^\infty$. The proof of Theorem 2.7 produces only a first order space discretization error since the smoothness of the BSDE coefficients is restricted to the second order differentiability.

The extrapolation error χ is specific to the convolution method implemented using the DFT. Equation (2.4.3) shows that errors appear and may accumulate around the boundaries of the truncated domain. Nonetheless, the truncation error is mainly time related through the density h and can be confined at the boundaries for fine time discretizations as shown in the following corollary.

Corollary 2.8. *Under the conditions of Theorem 2.7,*

$$
\lim_{|\pi|\to 0} \chi(x_k) = 0 \qquad (2.4.11)
$$

for any $x_k \in \left(-\frac{l}{2}, \frac{l}{2}\right)$.

Proof. If $x_k = 0$ then equation (2.4.3) gives $|\chi(0)| \leq 0$ and the result holds. If $x_k \neq 0$ and $x_k \in \left(-\frac{l}{2}, \frac{l}{2}\right)$, then

$$
\begin{aligned}
\lim_{|\pi|\to 0} \left(\int_{\frac{l}{2}-|x_k|}^{\frac{l}{2}} h(y) dy \right)^{\frac{1}{2}} &= \left(\lim_{|\pi|\to 0} \int_{\frac{l}{2}-|x_k|}^{\frac{l}{2}} h(y) dy \right)^{\frac{1}{2}} \\
&= \left(\lim_{\Delta_i \to 0} \int_{\frac{l}{2}-|x_k|}^{\frac{l}{2}} h(y) dy \right)^{\frac{1}{2}} \\
&= \left(\int_{\frac{l}{2}-|x_k|}^{\frac{l}{2}} \delta(y) dy \right)^{\frac{1}{2}} \\
&\quad \text{(where } \delta \text{ is the Dirac delta function)}, \\
&= 0
\end{aligned}
$$

since $0 \notin \left[\frac{l}{2} - |x_k|, \frac{l}{2}\right]$. Equation (2.4.3) then leads to the result. □

2.5 Extensions

Various simple extensions can be made of the convolution method. One of the most important one is reflected BSDEs. We also consider the convolution method under arithmetic Brownian motion. These cases have interesting applications in mathematical finance, especially for option pricing.

2.5.1 Reflected BSDEs

Euler schemes have been constructed for reflected BSDE with continuous barrier which make it possible the application of the convolution method to such BSDEs. Consider the solution (Y, Z, A) of the system

$$\begin{cases} -dY_t = f(t, Y_t, Z_t,)dt + dA_t - Z_t dW_t \\ Y_t \geq B_t \ , \ dA_t \geq 0 \ , \ \forall t \in [0, T] \\ \int_0^T (Y_t - B_t) dA_t = 0 \ , \ Y_T = g(W_T) \end{cases} \tag{2.5.1}$$

where the lower barrier is a deterministic function $B : [0, T] \times \mathbb{R} \to \mathbb{R}$ of time and the Brownian motion

$$B_t = B(t, W_t). \tag{2.5.2}$$

This RBSDE is associated to the following obstacle problem

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial^2 u}{\partial x^2} + f(t, x, u, \nabla u) = 0, \\ u(t, x) \geq B(t, x), \ (t, x) \in [0, T] \times \mathbb{R} \\ u(T, x) = g(x), \ x \in \mathbb{R} \end{cases} \tag{2.5.3}$$

as established by El Karoui et al. [45]. An adaption of the Euler scheme 1 provides the numerical solution to the reflected BSDE through the equations

$$\begin{cases} Z_{t_n}^\pi = 0, \ Y_{t_n}^\pi = \xi^\pi \\ Z_{t_i}^\pi = \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^\pi \Delta W_i | \mathcal{F}_{t_i}\right] \\ \Delta A_{t_i}^\pi = \left(\mathbf{E}\left[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\right] + f(t_i, \mathbf{E}\left[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\right], Z_{t_i}^\pi)\Delta_i - B(t_i, W_{t_i})\right)^- \\ Y_{t_i}^\pi = \mathbf{E}\left[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\right] + f(t_i, \mathbf{E}\left[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\right], Z_{t_i}^\pi)\Delta_i + \Delta A_{t_i}^\pi \end{cases} \tag{2.5.4}$$

where for any number $x \in \mathbb{R}$, $x^- = \max(0, -x)$.

The problems of time discretization of RBSDEs and their convergence were treated in Bouchard and Chassagneux [18] for the implicit Euler scheme. Peng and Xu [99] proposed an equivalent scheme with a discrete filtration and proved its convergence under a binomial method. This scheme is easily solved with a convolution method by noticing that the approximate solution u_i, the approximate gradient $\dot{u}_i$ and the approximate reflection $\bar{u}_i$ at mesh time t_i can be written as

$$u_i(x) = \tilde{u}_i(x) + \Delta_i f(t_i, \tilde{u}_i(x), \dot{u}_i(x)) + \Delta\bar{u}_i(x) \tag{2.5.5}$$

where

$$\dot{u}_i(x) = \frac{1}{\Delta_i}\int_{-\infty}^{\infty}(y - x)u_{i+1}(y)h(y|x)dy \tag{2.5.6}$$

$$\tilde{u}_i(x) = \int_{-\infty}^{\infty} u_{i+1}(y)h(y|x)dy \tag{2.5.7}$$

$$\Delta\bar{u}_i(x) := \bar{u}_{i+1}(x) - \bar{u}_i(x)$$

$$= \quad [\tilde{u}_i(x) + \Delta_i f(t_i, \tilde{u}_i(x), \dot{u}_i(x)) - B(t_i, x)]^- \tag{2.5.8}$$

for $i = 0, 1, ..., n-1$ and $u_n(x) = g(x)$. The computation of the approximated gradient $\dot{u}_i$ and the intermediate solution $\tilde{u}_i$ is identical to the non-reflected case exposed in Section 2.3.

One can also build an alternative scheme from the explicit Euler scheme 2 as follows

$$\begin{cases} Z^\pi_{t_n} = 0,\ Y^\pi_{t_n} = \xi^\pi \\ Z^\pi_{t_i} = \frac{1}{\Delta_i}\mathbf{E}\left[Y^\pi_{t_{i+1}}\Delta W_i | \mathcal{F}_{t_i}\right] \\ \Delta A^\pi_{t_i} = \left(\mathbf{E}\left[Y^\pi_{t_{i+1}} + f(t_i, Y^\pi_{t_{i+1}}, Z^\pi_{t_i})\Delta_i | \mathcal{F}_{t_i}\right] - B(t_i, W_{t_i})\right)^- \\ Y^\pi_{t_i} = \mathbf{E}\left[Y^\pi_{t_{i+1}} + f(t_i, Y^\pi_{t_{i+1}}, Z^\pi_{t_i})\Delta_i | \mathcal{F}_{t_i}\right] + \Delta A^\pi_{t_i}. \end{cases} \tag{2.5.9}$$

and define the approximate solution v_i, the approximate gradient $\dot{v}_i$ and the approximate reflection $\bar{v}_i$ as

$$v_i(x) = \int_{-\infty}^{\infty} \tilde{v}_{i+1}(y)h(y-x)dy + \Delta\bar{v}_i(x) \tag{2.5.10}$$

where

$$\begin{aligned} \tilde{v}_{i+1}(x) &= v_{i+1}(x) + \Delta_i f(t_i, v_{i+1}(x), \dot{v}_i(x)) \text{ and} && (2.5.11) \\ \dot{v}_i(x) &= \int_{-\infty}^{\infty} (y-x)v_{i+1}(y)h(y-x)dy && (2.5.12) \\ \Delta\bar{v}_i(x) &:= \bar{v}_{i+1}(x) - \bar{v}_i(x) \\ &= [\tilde{v}_{i+1}(x) - B(t_i, x)]^- && (2.5.13) \end{aligned}$$

for $i = 0, 1, ..., n-1$ and $v_n(x) = g(x)$.

One may even consider another scheme proposed by Peng and Xu [99] based on the penalization method used by El Karoui et al.[45] to proved reflected BSDEs well-posedness. But since the penalization method relies on estimates that approaches the RBSDE solution from below, the scheme necessarily under-estimates the RBSDE solution.

Alternative to the Euler scheme itself can be found in the $\theta-$schemes of Zhao, Shen and Peng [125]. We already gave a description of these latter schemes in equations (1.3.2) and (1.3.3) to which a reflected feature can easily be added.

2.5.2 Arithmetic Brownian motion

We can extend the convolution method to consider an arithmetic Brownian motion

$$X_t = x_0 + \mu t + \sigma W_t \tag{2.5.14}$$

as the forward process and solve for a Cauchy problem to an advection-diffusion equation

$$\begin{cases} \frac{\partial u}{\partial t} + \mu\frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 u}{\partial x^2} + f(t, x, u, \sigma\nabla u) = 0\ ,\ (t, x) \in [0, T) \times \mathbb{R} \\ u(T, x) = g(x),\ x \in \mathbb{R}. \end{cases} \tag{2.5.15}$$

to which an obstacle can be added when in the presence of a reflected BSDE. The forward process increments are indeed stationary, independent and normally distributed with density

$$h(x) = \frac{1}{(2\pi\Delta_i)^{\frac{1}{2}}\sigma}\exp\left(-\frac{(x-\mu\Delta_i)^2}{2\sigma^2\Delta_i}\right). \tag{2.5.16}$$

and characteristic function

$$\phi(\nu) = e^{\Delta_i\left(\mathbf{i}\mu\nu - \frac{1}{2}\sigma^2\nu^2\right)}. \tag{2.5.17}$$

The development of the convolution method in this case also leads to transforms identical to equation (2.3.1) with $\psi(\nu) = \phi(\nu - \mathbf{i}\alpha)$ when computing the approximate solutions v_i or the

intermediate solutions $\tilde{u}_i$ and $\psi(\nu) = \sigma(\alpha + \mathbf{i}\nu)\phi(\nu - \mathbf{i}\alpha)$ when computing the approximate gradients $\dot{u}_i$ and $\dot{v}_i$. In our codes, the approximate gradients $\dot{u}$ and $\dot{v}$ are actually estimates for $\sigma\nabla u = \sigma\frac{\partial u}{\partial x}$ but the schemes can easily be modified so as to estimate the gradient ∇u directly.

The equivalences

$$\begin{aligned}\theta(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\eta^{\alpha}}(\nu)\psi(\nu)d\nu \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha}_{\beta,\kappa}}(\nu)\psi(\nu)d\nu - e^{-\alpha x}(\beta(x+\mu\Delta_i)+\kappa)\end{aligned} \tag{2.5.18}$$

when $\psi(\nu) = \phi(\nu - \mathbf{i}\alpha)$ and

$$\begin{aligned}\theta(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\eta^{\alpha}}(\nu)\psi(\nu)d\nu \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha}_{\beta,\kappa}}(\nu)\psi(\nu)d\nu - e^{-\alpha x}\beta\sigma\end{aligned} \tag{2.5.19}$$

when $\psi(\nu) = \sigma(\alpha + \mathbf{i}\nu)\phi(\nu - \mathbf{i}\alpha)$ of Theorem 2.6 still holds.

2.6 Numerical results

The convolution approach to BSDEs presented in this chapter can be used in various applications where a (numerical) solution to a PDE or a BSDE is needed. Here, we give examples of applications in numerical solution to PDEs, simulation of (R)BSDEs and option pricing in a one-dimensional framework.

2.6.1 Simulation of (R)BSDEs

The availability of a numerical solution to the PDE underlying the BSDE clearly makes it easy to simulate the BSDE itself. Many authors, including Douglas, Ma and Protter [40] and Milstein and Tretyakov [89, 90], put forth a PDE approach to solve (coupled) FBSDEs numerically as discussed in Chapter 1. The convolution method, along with the binomial tree method of Peng and Xu [99], is a lighter and more suitable method for the less general case of BSDEs.

We shall first consider the one-dimensional linear BSDE with generator

$$f(t, y, z) = ay + bz + c \tag{2.6.1}$$

with a, b and c being real numbers. These BSDEs were already treated in Peng and Xu [99] where the authors indicate that the initial value for the forward process Y of such a BSDE is given by

$$Y_0 = e^{(a-\frac{1}{2}b^2)T}\mathbf{E}\left[g(W_T)e^{bW_T}\right] + \frac{c}{a}\left(e^{aT} - 1\right). \tag{2.6.2}$$

where we take a maturity $T = 1$,

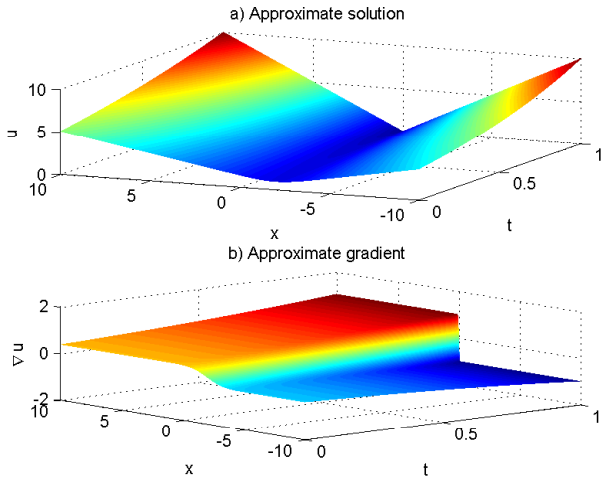
$$g(x) = |x|, \tag{2.6.3}$$

as the terminal condition and $a = -1$, $b = 2$ and $c = 1$ for the sake of this example.

The PDE is solved on the restricted, real space domain $[x_0, x_N] = [-10, 10]$ with $N = 2^{12}$ grid points. The number of time steps n is set to 1000 and the minimal slope to $\epsilon = 5$.

When simulating the BSDE, the restricted real space domain plays a key role since it has to contain the path values of the Brownian motion. One way to select the domain is by taking it large enough so that the probability that the Brownian motion finishes in the restricted domain approaches one (1). Indeed, the larger the restricted domain, the larger is the number

Figure 2.6.1: Numerical solution to the linear PDE.



The surfaces are obtained with the convolution method applied on Scheme 1 on the restricted domain $[x_0, x_N] = [-10, 10]$ with $N = 2^{12}$ grid points, $n = 500$ time steps and a minimal slope of $\epsilon = 5$.

Table 2.6.1: Estimates for the initial value of the linear BSDE forward process.

n (number of time steps)	100	500	1000	2000
Convolution (Scheme 1)	1.3785	1.3750	1.3746	1.3743
Convolution (Scheme 2)	1.3777	1.3748	1.3745	1.3743
Trinomial tree (Scheme 1)	1.3785	1.3750	1.3746	1.3743

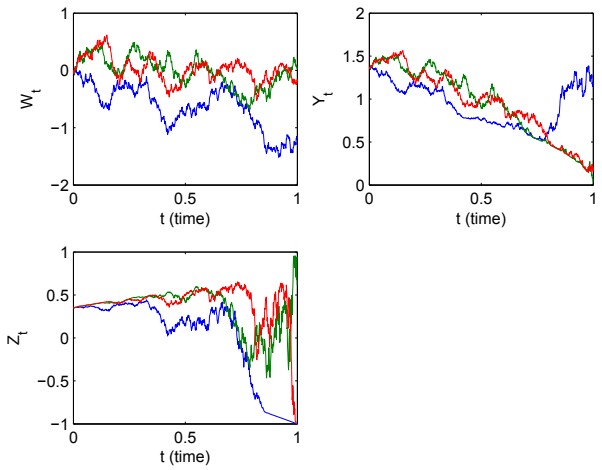
For the convolution method, the estimates are valued on the restricted domain $[x_0, x_N] = [-10, 10]$ with 2^{12} grid points and a minimal slope of $\epsilon = 5$.

of grid points N since the space step Δx must be small enough for the numerical solution of the PDE to detect the variation in the Brownian motion paths. When the Brownian motion takes intermediate values that are not on the space grids, the corresponding values for the BSDE solution can be interpolated (linearly) from the PDE solutions. The numerical solutions for the PDE (approximate solution and gradient) are presented in Figure 2.6.1.

A Monte Carlo method gives an estimate of 1.3745 with a standard deviation 0.0045 for the initial value of the forward process Y_0 using equation (2.6.2) and 5×10^6 trajectories. With this estimate as a benchmark, Table 2.6.1 displays the values obtained with the convolution method on the second scheme where only the number of time steps is changed among the specified inputs. Additional results obtained with a trinomial tree method are presented for comparison. As one can see, both the convolution and the trinomial tree estimates close up on finer time grids.

Figure 2.6.2 shows three (3) simulated paths for the Brownian motion W and the corresponding simulated paths for the forward process Y and the control process Z. Paths for the backward and control processes (Y_t and Z_t respectively) are simulated using the solution from the convolution method applied on Scheme 1 on the restricted domain $[x_0, x_N] = [-10, 10]$ with $N = 2^{12}$ grid points, $n = 1000$ time steps and a minimal slope of $\epsilon = 5$. The same number of time steps is used to simulate the Brownian paths (W_t). An advantage of the convolution method over tree based methods is the simplification of the simulation procedure since the Brownian path does not have to be approximated by scaled random walk as in Peng and Xu [99].

Figure 2.6.2: Path simulation for the BSDE solution.



It is possible to simulate reflected BSDEs in a similar manner. If we consider the reflected BSDE with the same driver and terminal condition as in the previous non-reflected and linear case and set the reflecting barrier to be

$$B(t,x) = g(x) = |x| \ , \ (t,x) \in [0,T] \times \mathbb{R}, \tag{2.6.4}$$

then we get the approximate solution and the approximate gradient displayed in Figure 2.6.3.

In order to simulate the reflecting process A, we compute the reflection increments from equation (2.5.8) at each time step along the Brownian path. Summing those increments then gives the corresponding path values for the reflecting process. Three (3) simulated Brownian paths and their counterparts for the RBSDE solution are plotted in Figure 2.6.6. Paths for the backward, control and reflecting processes (Y_t, Z_t and A_t respectively) are simulated using the solution from the convolution method applied on Scheme 1 on the restricted domain $[x_0, x_N] = [-10, 10]$ with $N = 2^{12}$ grid points, $n = 1000$ time steps and a minimal slope of $\epsilon = 5$. We naturally use $n = 1000$ time steps to simulate the Brownian paths (W_t).

The convolution method returns a forward process initial values of $Y_0 = 1.3820$ for both Schemes 1 and 2 with the specified inputs. Those estimates are identical to the approximation given by the trinomial tree method (Scheme 2) with the same number of time steps $n = 1000$ as the convolution methods.

2.6.2 Option pricing under Black-Scholes model

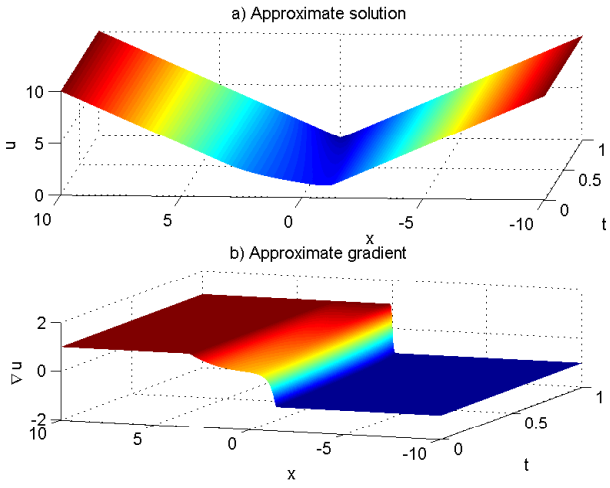
Through option pricing problems, we will particularly treat the case of BSDEs with non-linear drivers that was not considered in the previous examples. An introduction to financial applications of BSDEs, particularly to imperfect markets and American option problems, can be found in El Karoui and Quenez [48], El Karoui, Pardoux and Quenez [46] or El Karoui, Peng and Quenez [47]. Also the celebrated papers of Black and Scholes [17] and Merton [88] constitute the financial basis in this section.

For the market model consisting of a single risky asset (or stock) $\{S_t\}_{t\in[0,T]}$ with the dynamic

$$S_t = e^{X_t} \tag{2.6.5}$$

where the process $\{X_t\}_{t\in[0,T]}$ represents the stock return, we price an European call option with maturity $T = 1$ and strike price K under a lending rate of $r = 0.01$ and a borrowing rate R.

Figure 2.6.3: Numerical solution to the linear PDE with obstacle.



The surfaces are obtained with the convolution method applied on Scheme 1 on the restricted domain $[x_0, x_N] = [-10, 10]$ with $N = 2^{12}$ grid points, $n = 500$ time steps and a minimal slope of $\epsilon = 5$.

Figure 2.6.4: Path simulation for the reflected BSDE solution.

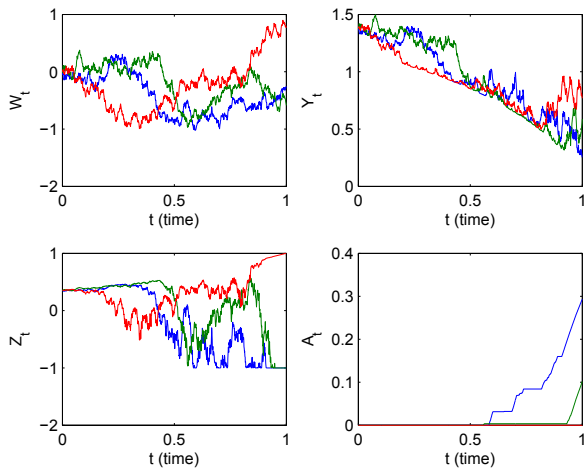
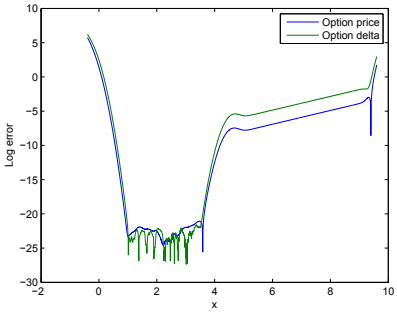


Figure 2.6.5: Absolute errors on American call option prices and deltas.



Estimates were obtained by applying the convolution method (Scheme 1) on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points, $n = 1000$ time steps and a minimal slope of $\epsilon = 5$. The option has stike price $K = S_0 = 100$ with $R = r = 0.01$.

The return process is an arithmetic Brownian motion

$$X_t = X_0 + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma W_t \tag{2.6.6}$$

such that the stock has an initial value of $S_0 = e^{X_0} = 100$, an expected return rate of $\mu = 0.05$ and a volatility of $\sigma = 0.2$.

The call option price then follows a BSDE with the return process $\{X_t\}_{t\in[0,T]}$ as the forward process, the driver

$$f(t, y, z) = -ry - \left(\frac{\mu - r}{\sigma}\right) z + (R - r)\left(y - \frac{z}{\sigma}\right)^- \tag{2.6.7}$$

and the terminal function

$$g(x) = (e^x - K)^+ \tag{2.6.8}$$

under those imperfect market conditions.

When the borrowing rate equals the lending rate $R = r = 0.01$, the European and American call options have the same price. Figure 2.6.5 shows the structure of the absolute error on stock option prices and deltas where the true values are computed using the Black-Scholes formula. As expected the errors are amplified at the boundaries of the truncated domain, but also for around-the-money options in a lesser extend due to the non-smoothness of the terminal function g. In addition, out-of-the-money options have smaller absolute errors compared to in-the-money options and option prices also presents smaller absolute errors compare to option deltas.

The Black-Scholes formula gives call option prices of 4.6101, 8.4333 and 14.1929 at strike prices $K = 110$, 100 and 90 respectively. Also, the true values for the option deltas are 0.7507, 0.5596 and 0.3720 when the strike price is $K = 90, 100$ and 110 respectively. Table 2.6.2 gives the price estimates with both convolution schemes and the trinomial tree method using different time steps and the indicated strike prices. Also, Table 2.6.3 contains the relative errors for the option deltas obtained from the approximate gradient by

$$\text{Delta} = \frac{\dot{u}_0(X_0)}{\sigma S_0} \tag{2.6.9}$$

when using Scheme 1 or

$$\text{Delta} = \frac{\dot{v}_0(X_0)}{\sigma S_0} \tag{2.6.10}$$

for Scheme 2. A similar computation allows to obtain the option deltas from the trinomial tree approach.

Table 2.6.2: Relative errors (in percentage) for American call option prices on non-dividend-paying stock with no market frictions.

	K (Strike)	n=500	n=1000	n=2000	n=5000
Convolution (Scheme 1)	110	0.0456	0.0217	0.0108	0.0043
	100	0.0178	0.0095	0.0047	0.0024
	90	0.0049	0.0028	0.0014	0.0007
Convolution (Scheme 2)	110	0.0087	0.0239	0.0022	0.0001
	100	0.0059	0.0024	0.0012	0.0007
	90	0.0028	0.0014	0.0007	0.0004
Trinomial tree (Scheme 1)	110	0.0065	0.0369	0.0087	0.0022
	100	0.0356	0.0012	0.0047	0.0024
	90	0.0007	0.0028	0.0021	0.0007

For the convolution method, the option prices are valued on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points and a minimal slope of $\epsilon = 5$. Both the lending and the borrowing rates are taken equal $R = r = 0.01$.

Table 2.6.3: Relative errors (in percentage) for the American call option deltas on non-dividend-paying stock with no market frictions.

K (Strike)	90	100	110
Convolution (Scheme 1)	0.0133	0.0010	0.2414
Convolution (Scheme 2)	0.0133	0.0010	0.2414
Trinomial tree (Scheme 1)	0.0133	0.0010	0.2414

For the convolution method, the option deltas are valued on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points and a minimal slope of $\epsilon = 5$. The number of time steps is set to $n = 2000$ for all three methods. The borrowing and lending rates are equal $R = r = 0.01$.

The results of Table 2.6.2 and 2.6.3 show the accuracy of the convolution method on a RBSDE with a smooth linear driver. Indeed, the relative error percentages remain low (less than 0.3%) for the estimated option prices and deltas. However, out-the-money option estimates seem to display the largest relative errors. In the same order of idea, option deltas have larger relative errors compared to option prices which confirms the observations on Figure 2.6.5. Overall, the precision of the convolution method is similar to the trinomial method since both methods display similar relative errors.

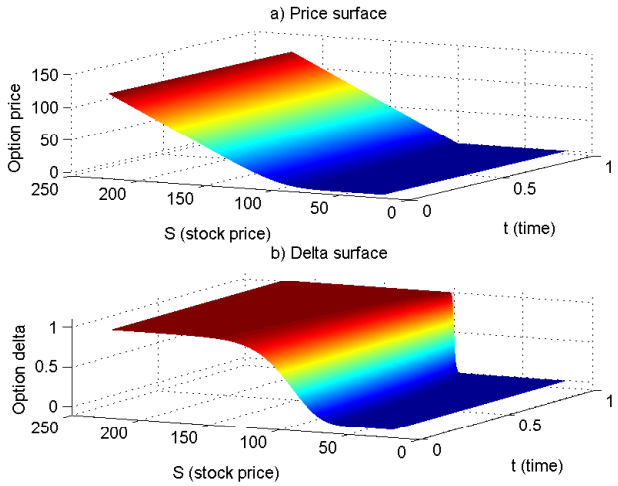
For a borrowing rate of $R = 0.03$ (different from the lending rate $r = 0.01$), the Black-Scholes formula does not apply but the convolution method is able to produce the price and delta surfaces for option along with estimates. Surfaces are displayed in Figure 2.6.6 for the at-the-money European call option. Table 2.6.4 gives the at-the-money European option price estimates with different time discretizations. Table 2.6.5 completes the information on option price with out-of-the-money European option prices. All three methods (convolution method on Scheme 1 and 2 and the trinomial method) return identical delta values as can be seen on Table 2.6.6 when applied with $n = 2000$ time steps. The similarity in the estimates given by the convolution method and the trinomial method is an indication of the good performance of the convolution in non-smooth and non-linear driver cases.

Table 2.6.4: At-the-money call option prices under imperfect market conditions.

n (number of time steps)	500	1000	2000	5000
Convolution (Scheme 1)	9.4132	9.4133	9.4133	9.4134
Convolution (Scheme 2)	9.4127	9.4131	9.4132	9.4133
Trinomial tree (Scheme 1)	9.4107	9.4136	9.4130	9.4132

For the convolution method, the estimates are valued on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points and a minimal slope of $\epsilon = 5$. The risk free rates are $R = 0.03$ when borrowing and $r = 0.01$ when lending and the option strike price is $K = S_0 = 100$.

Figure 2.6.6: At-the-money European call option price and delta surfaces.



Surfaces were obtained by truncating the approximate solution and grandient of the convolution method (Scheme 1) on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points, $n = 1000$ time steps and a minimal slope of $\epsilon = 5$. The option has stike price $K = S_0 = 100$ with $R = 0.03$ and $r = 0.01$.

The option pricing can be made using a Monte Carlo method such as the forward scheme of Bender and Denk [10]. But in the context of uni-dimensional BSDEs, Monte Carlo methods will generally be heavier than space discretization methods. As an illustration, the convolution method on both Scheme 1 and 2 runs in approximately 4.4 seconds when pricing the option of Table 2.6.4 with $n = 1000$ time steps. On the other hand, the trinomial tree method runs in 0.25 second. As to the forward scheme, it runs in 18 seconds with only $n = 20$ time steps[4]. Fifty (50) independent valuations with the Monte Carlo method give a 95% confidence interval of $[9.3972, 9.4222]$ which includes all estimates of Table 2.6.4. Hence, the convolution method is faster that Monte Carlo methods but slower than the trinomial (or binomial) method.

The American call options, which solves a reflected BSDE with the barrier function

$$B(t,x) = g(x) = (e^x - K)^+ \ , \ (t,x) \in [0,T] \times \mathbb{R} \tag{2.6.11}$$

essentially have the same price as their European counterparts under the market conditions stated above since our risky asset $\{S_t\}_{t\in[0,T]}$ pays no dividend and, hence, the early exercise premium is null. This can be seem numerically by simulating sample paths for the corresponding

[4]We also used the 7 first power functions and 100000 paths to generate the estimates. The Picard iterations are stopped whenever the difference in two consecutive prices is less than 10^{-4} for a maximum number of 10 integrations.

Table 2.6.5: Out-of-the-money European call option prices under imperfect market conditions.

	K (Strike)	$n = 500$	$n = 1000$	$n = 2000$	$n = 5000$
Convolution	110	5.2932	5.2933	5.2933	5.2934
(Scheme 1)	90	15.4290	15.4291	15.4291	15.4292
Convolution	110	5.2924	5.2929	5.2931	5.2933
(Scheme 2)	90	15.4289	15.4291	15.4292	15.4292
Trinomial tree	110	5.2933	5.2918	5.2938	5.2935
(Scheme 1)	90	15.4295	15.4297	15.4295	15.4293

For the convolution method, the option prices are valued on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with 2^{12} grid points and a minimal slope of $\epsilon = 5$. The risk free rates are $R = 0.03$ and $r = 0.01$.

Table 2.6.6: European call option deltas under imperfect market conditions.

K (Strike)	90	100	110
Convolution (Scheme 1)	0.7814	0.5987	0.4104
Convolution (Scheme 2)	0.7814	0.5987	0.4104
Trinomial tree (Scheme 1)	0.7814	0.5987	0.4104

For the convolution method, the option deltas are valued on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points and a minimal slope of $\epsilon = 5$. The number of time steps is set to $n = 2000$ for all three methods. The risk free rates are $R = 0.03$ and $r = 0.01$.

RBSDE solution. On Figure 2.6.7, one notices that the cost for hedging the (at-the-money) American option is negligible either the option finishes in or out the money. Similar results are obtained for in and out-of-the-money American call options.

If we introduce a dividend rate of $\delta = 0.035$ under imperfect market conditions ($R = 0.03$ and $r = 0.01$), the forward (return) process takes the form

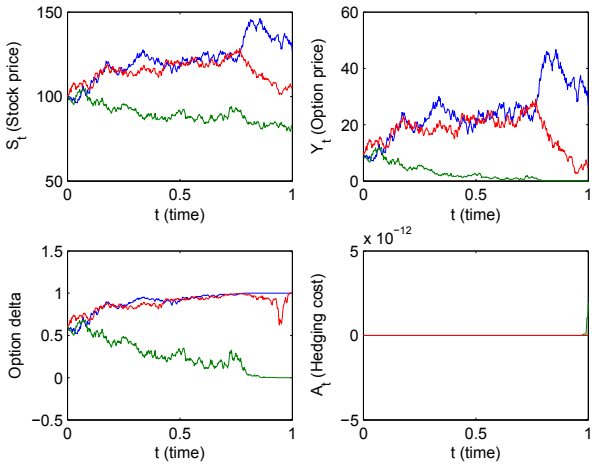
$$X_t = X_0 + \left(\mu - \delta - \frac{1}{2}\sigma^2\right) t + \sigma W_t. \tag{2.6.12}$$

The European and American option prices differ and the Black-Scholes formula does not apply. Table 2.6.7 compares the European and American call option prices under the convolution and the trinomial method at different strike prices. Table 2.6.8 does the same exercise for option deltas.

If the deltas computed with the convolution and the trinomial method are identical, the convolution method gives slightly higher option prices compared to the trinomial method. This difference in option prices between both methods may find its explanation in the extrapolation errors generated by the convolution method or in the non-smoothness of the option price and delta functions.

Nonetheless, the difference between European and American option prices shows that the convolution method captures the reflecting effect. This difference between both option prices can be visualized on Figure 2.6.8. Finally, Figure 2.6.9 shows the typical sample paths for the American option where the reflecting process A_t (hedging cost) is now non-zero for in-the-money path indicating a difference in price with the European call option.

Figure 2.6.7: Sample paths for the American call option on non-dividend-paying stock.



Paths are simulated using the solution from the convolution method applied on Scheme 1 on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points, $n = 1000$ time steps and $\epsilon = 5$. We used $n = 1000$ time steps to simulate the stock price (S_t). Also, the American option, with stike $K = S_0 = 100$, was priced under imperfect market conditions: $R = 0.03$ and $r = 0.01$.

Table 2.6.7: European and American call option prices on dividend-paying stock.

	K (Strike)	European	American
Convolution (Scheme 1)	110	3.9963	4.0322
	100	7.4712	7.5610
	90	12.8339	13.0505
Convolution (Scheme 2)	110	3.9962	4.0321
	100	7.4712	7.5609
	90	12.8339	13.0505
Trinomial tree (Scheme 1)	110	3.9958	4.0317
	100	7.4716	7.5614
	90	12.8333	13.0500

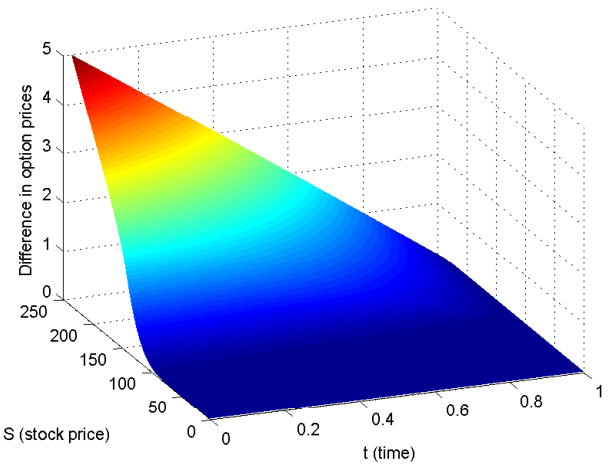
For the convolution method, the option deltas are valued on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points and a minimal slope of $\epsilon = 5$. The number of time steps is set to $n = 2000$ for all three methods. The risk free rates are $R = 0.03$ and $r = 0.01$. The dividend yield is $\delta = 0.035$.

Table 2.6.8: European and American call option deltas on dividend-paying stock.

	K (Strike)	European	American
Convolution (Scheme 1)	110	0.3322	0.3362
	100	0.5117	0.5207
	90	0.7014	0.7203
Convolution (Scheme 2)	110	0.3322	0.3363
	100	0.5117	0.5207
	90	0.7014	0.7204
Trinomial tree (Scheme 1)	110	0.3322	0.3362
	100	0.5117	0.5207
	90	0.7014	0.7204

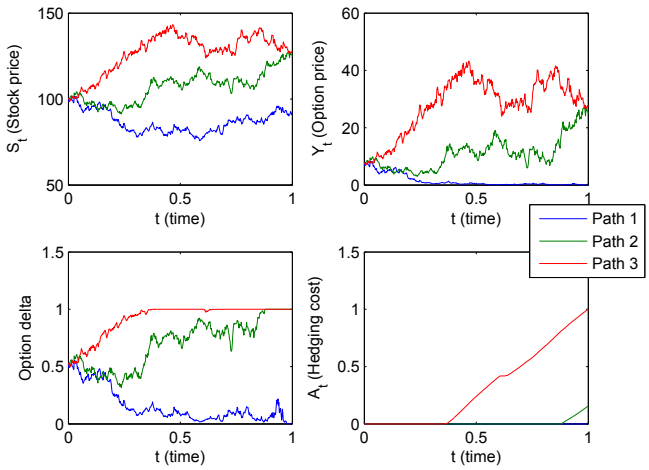
For the convolution method, the option deltas are valued on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points and a minimal slope of $\epsilon = 5$. The number of time steps is set to $n = 2000$ for all three methods. The risk free rates are $R = 0.03$ and $r = 0.01$. The dividend yield is $\delta = 0.035$.

Figure 2.6.8: Difference in price between the American and European call options on dividend-paying stock.



Surfaces was obtained by truncating the approximate solutions of the convolution method (Scheme 1) on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points, $n = 1000$ time steps and a minimal slope of $\epsilon = 5$. The option has stike price $K = S_0 = 100$ with $R = 0.03$, $r = 0.01$ and $\delta = 0.035$.

Figure 2.6.9: Sample paths for the American call option on dividend-paying stock.



Paths are simulated using the solution from the convolution method applied on Scheme 1 on the restricted domain $[x_0, x_N] = X_0 + [-5, 5]$ with $N = 2^{12}$ grid points, $n = 1000$ time steps and a minimal slope of $\epsilon = 5$. The option has stike price $K = S_0 = 100$ with $R = 0.03$, $r = 0.01$ and $\delta = 0.035$.

Overall, the convolution method implemented on a uniform grid gives satisfactory results. The numerical results shows the method's accuracy even on BSDEs with unbounded terminal conditions and non-smooth coefficients. Nonetheless, the error analysis indicates the presence of a truncation error. Since the truncation error depends on the time discretization and not on the space discretization, a more suitable space discretization can suppress it. The following chapter investigates the issue.

Chapter 3

Alternative discretization of convolution

This chapter proposes an alternative discretization of the convolution method developed in Chapter 2. This alternative discretization is motivated by the absence of convergence of the convolution scheme in the previous chapter. The non-converging error term is due to an extrapolation that occurs for most space nodes. This error term is larger for the space nodes around the boundaries of the restricted domain as shown in the error analysis of the previous chapter.

In order to avoid extrapolation, we work with two different restricted domains. The first domain discretizes the Brownian increment and the second discretizes the spatial domain on which the BSDE solution is defined. The scheme then combines both discretizations by choosing the same space step for each of them. The FFT algorithm is once again used to perform the quadratures so to maintain the algorithm's efficiency. If the procedure forces the contraction of the space grid through times steps, we are able to build a space grid suitable for simulation which assures convergence.

3.1 Alternative discretization

We shall illustrate the alternative discretization with the explicit Euler scheme 1 with the approximate solutions of equations (2.2.6), (2.2.7) and (2.2.8). On the time mesh $\pi = \{t_0 = 0 < t_1 < ... < t_n = T\}$ such that

$$\Delta_i = t_{i+1} - t_i,\ i = 0, 1, ..., n-1, \tag{3.1.1}$$

these equations may be written as

$$u_i(x) = \tilde{u}_i(x) + \Delta_i f(t_i, \tilde{u}_i(x), \dot{u}_i(x)) \tag{3.1.2}$$

where

$$\dot{u}_i(x) = \frac{1}{\Delta_i} \int_{-\infty}^{\infty} y u_{i+1}(x+y) h(y) dy \tag{3.1.3}$$

$$\tilde{u}_i(x) = \int_{-\infty}^{\infty} u_{i+1}(x+y) h(y) dy \tag{3.1.4}$$

after a change of variable for $i = 0, 1, ..., n-1$ and $u_n(x) = g(x)$. Under the implicit Euler scheme, equation (3.1.2) may be replaced by

$$u_i(x) = \tilde{u}_i(x) + \Delta_i f(t_i, u_i(x), \dot{u}_i(x)) \tag{3.1.5}$$

with the additional the condition of equation (2.2.13) on the time discretization.

3.1.1 Alternative transform

In order to compute the convolutions of equations (3.1.3) and (3.1.4) with the FFT algorithm, the main requirement was that the function values and the derivative values match at the boundaries of the truncated domain. In Chapter 2, we used a transform to meet that requirement. Given a function $\eta : [a, b] \to \mathbb{R}$ and $\eta \in \mathcal{C}^1$, we considered the transform

$$\eta^{\alpha}_{\beta,\kappa}(x) = e^{-\alpha x}(\eta(x) + \beta x + \kappa) \tag{3.1.6}$$

and the coefficients α, β and κ were chosen such that

$$\begin{cases} \eta^{\alpha}_{\beta,\kappa}(a) = \eta^{\alpha}_{\beta,\kappa}(b) \\ \frac{\partial \eta^{\alpha}_{\beta,\kappa}}{\partial x}(a) = \frac{\partial \eta^{\alpha}_{\beta,\kappa}}{\partial x}(b). \end{cases} \tag{3.1.7}$$

The transform in equation (3.1.6) presents two main disadvantages. First, it uses three coefficients when only two conditions need to be satisfied, leaving the third coefficient almost free. Hence, we require a simplified transform with only two coefficients. Also, the transform depends exponentially on the dampening coefficient α. A linear dependence in the coefficients is more suitable so that the error induced is also linear.

We propose the alternative transform

$$\eta^{\alpha,\beta}(x) \quad := \quad \eta^{0}_{\beta,0} + \alpha x^2 = \eta(x) + \alpha x^2 + \beta x \tag{3.1.8}$$

satisfying

$$\begin{cases} \eta^{\alpha,\beta}(a) = \eta^{\alpha,\beta}(b) \\ \frac{\partial \eta^{\alpha,\beta}}{\partial x}(a) = \frac{\partial \eta^{\alpha,\beta}}{\partial x}(b). \end{cases} \tag{3.1.9}$$

The following lemma gives a method to select the coefficients α and β for the transform of equation (3.1.8).

Lemma 3.1. *Suppose the real function $\eta \in \mathcal{C}^1[a, b]$ is differentiable and let $\eta^{\alpha,\beta}$ be its transformed function as defined in equation (3.1.8). Then*

$$\alpha \quad = \quad \frac{\frac{\partial \eta}{\partial x}(a) - \frac{\partial \eta}{\partial x}(b)}{2(b-a)}\ , \tag{3.1.10}$$

$$\beta \quad = \quad \frac{\eta(a) - \eta(b)}{(b-a)} - \alpha(b+a) \tag{3.1.11}$$

solve the system of linear equations defined by the conditions of equation (3.1.9).

Proof. The first equation of the system (3.1.9) gives (3.1.10) in a straightforward manner. Equation (3.1.11) is given by the second equation of the system. □

A major feature of the transform in equation (3.1.8) is the absence of dampening. Thus, the convolutions on Scheme I are represented as

$$\tilde{u}_i(x) \quad = \quad \mathfrak{F}^{-1}\left[\mathfrak{F}[u_{i+1}](\nu)\phi(\nu)\right](x) \tag{3.1.12}$$

$$\dot{u}_i(x) \quad = \quad \mathfrak{F}^{-1}\left[\mathbf{i}\nu\mathfrak{F}[u_{i+1}](\nu)\phi(\nu)\right](x) \tag{3.1.13}$$

whenever u_{i+1}, $i = 1, 2, ..., n$ is integrable from equations (2.2.26) and (2.2.27). Numerically, the truncation solves the integrability problem so that dampening is not needed. Hence, non-integrable functions can be treated as already indicated in Subsection 2.3.2.

The next theorem gives the representation of convolution under the transform of equation (3.1.8).

Theorem 3.2. *Let $\eta : [a,b] \to \mathbb{R}$ be an integrable and differentiable function and let $\eta^{\alpha,\beta}$ be its transformed function as defined in equation (3.1.8). Then the function $\theta : [a,b] \to \mathbb{R}$ given by*

$$\theta(x) = \mathfrak{F}^{-1}\left[\mathfrak{F}[\eta](\nu)\psi(\nu)\right](x) \tag{3.1.14}$$

admits the alternative representation

$$\theta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha,\beta}}(\nu)\psi(\nu)d\nu - (2\alpha x + \beta) \tag{3.1.15}$$

if $\psi(\nu) = \mathbf{i}\nu\phi(\nu)$ or

$$\theta(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha}_{\beta,\kappa}}(\nu)\psi(\nu)d\nu - \alpha(x^2 + \Delta_i) - \beta x \tag{3.1.16}$$

if $\psi(\nu) = \phi(\nu)$.

Proof. First, let $\psi(\nu) = \mathbf{i}\nu\phi(\nu)$. By definition, we know that

$$\begin{aligned}
\theta(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\eta}(\nu)\psi(\nu)d\nu \\
&= \frac{1}{\Delta_i}\mathbf{E}\left[\eta(W_{t_{i+1}})\Delta W_i | W_{t_i} = x\right] \\
&= \frac{1}{\Delta_i}\mathbf{E}\left[\left(\eta^{\alpha,\beta}(W_{t_{i+1}}) - \alpha W_{t_{i+1}}^2 - \beta W_{t_{i+1}}\right)\Delta W_i | W_{t_i} = x\right] \\
&= \frac{1}{\Delta_i}\mathbf{E}\left[\left(\eta^{\alpha,\beta}(W_{t_{i+1}}) - \alpha(W_{t_i} + \Delta W_i)^2 - \beta W_{t_{i+1}}\right)\Delta W_i | W_{t_i} = x\right] \\
&= \frac{1}{\Delta_i}\mathbf{E}\left[\eta^{\alpha,\beta}(W_{t_{i+1}})\Delta W_i | W_{t_i} = x\right] - (2\alpha x + \beta) \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha,\beta}}(\nu)\psi(\nu)d\nu - (2\alpha x + \beta).
\end{aligned}$$

Similarly, if $\psi(\nu) = \phi(\nu)$, we have

$$\begin{aligned}
\theta(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\eta}(\nu)\psi(\nu)d\nu \\
&= \mathbf{E}\left[\eta(W_{t_{i+1}}) | W_{t_i} = x\right] \\
&= \mathbf{E}\left[\eta^{\alpha,\beta}(W_{t_{i+1}}) | W_{t_i} = x\right] - \alpha(x^2 + \Delta_i) - \beta x \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha,\beta}}(\nu)\psi(\nu)d\nu - \alpha(x^2 + \Delta_i) - \beta x.
\end{aligned}$$

□

As in Chapter 2, the values of the derivatives can be approximated by finite differences. Another approach is to use the approximation of the derivative given by the convolution method.

3.1.2 Alternative grid

The discretization proposed in this chapter considers a fixed space grid for the integrated variable y in equations (3.1.3) and (3.1.4) which represents the Brownian increment. This variable is restricted on the interval of length $l > 0$ centered at zero (0) with an even number $N \in \mathbb{N}^*$ of steps. Hence, the space step is given by

$$\Delta x = \frac{l}{N}. \tag{3.1.17}$$

This space is then used for the grid discretizing the domain variable x which represents the Brownian process itself. At each time node t_i, $i = 0, 1, 2, ..., n$, the spatial domain is restricted on an interval of length $N_i l$ where $N_i \in \mathbb{N}$ is a positive integer with

$$N_i = N_0 + i. \tag{3.1.18}$$

In order to maintain a space step of Δx at each time step, we discretize the spatial domain with $N_i N$ space steps. Since we center the spatial domain at $W_0 = 0$, we get the space nodes

$$x_{ik} = -\frac{N_i l}{2} + k\Delta x,\ k = 0, 1, ..., N_i N \tag{3.1.19}$$

at mesh time t_i. In particular, if $N_0 = 0$ then the space grid at mesh time t_0 is composed of the single point

$$x_{00} = W_0. \tag{3.1.20}$$

More precisely, we start the space grid at mesh time t_0 with $(N_0 N + 1)$ points. Then at each times step, we add N new points to the space grid of the previous time step. Moreover, those new N points are equally distributed at both boundaries of the previous restricted space domain. These ideas are reflected in the relationship between space nodes of consecutive space grids since

$$x_{ik} = x_{i+1,k+\frac{N}{2}},\ k = 0, 1, ..., N_i N. \tag{3.1.21}$$

Figure 3.1.1 gives examples of alternative grids using different parameter values.

This approach already underlies tree based methods to a lesser extent however the present discretization offers more flexibility. Whereas the number of space grid points at mesh time t_0 is limited to 1 in multinomial methods, it can be selected almost freely in this alternate discretization. Also, the number $N + 1$ of nodes on the Brownian increment restricted domain can be compared to the number of branches in a multinomial method. Hence, this alternative discretization can be seen as a recombining tree with $N + 1$ branches and $N_0 N + 1$ initial grid points since we use a fixed space step Δx.

The Fourier relations of equations (3.1.12) and (3.1.13) call for a discretization of the Fourier space as well. At each mesh time t_i, $i = 1, 2, ..., n$, the Fourier space is restricted on an interval of length L centered at zero (0) and discretized with $N_i N$ space steps. The equidistant nodes are thus of the form

$$\nu_{ik} = -\frac{L}{2} + k\Delta\nu_i,\ k = 0, 1, ..., N_i N \tag{3.1.22}$$

where $\Delta\nu_i = \frac{L}{N_i N}$. The Nyquist relation holds whenever L is such that

$$Ll = 2\pi N. \tag{3.1.23}$$

For a fixed time mesh, the length of the Brownian increment restricted interval l, the number of space steps of this interval N and the number N_0 completely defines the space grid described above.

3.1.3 Numerical implementation

In this chapter, we seek numerical approximations of equations (3.1.12) and (3.1.13) at each mesh time t_i, $i = 0, 1, ..., n-1$. At time t_i, we use the generic functions $\theta_i : \mathbb{R} \to \mathbb{R}$, $\psi : \mathbb{R} \to \mathbb{C}$ and $\theta_{i+1} : \mathbb{R} \to \mathbb{R}$ such that

$$\theta_i(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{\mathbf{i}\nu x} \hat{\theta}_{i+1}(\nu)\psi(\nu)d\nu. \tag{3.1.24}$$

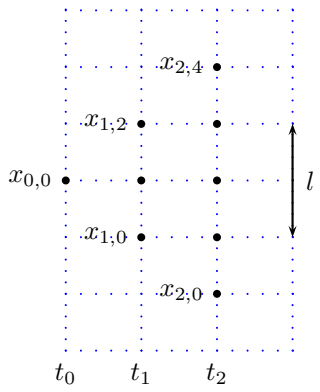
We assume that the function θ_{i+1} satisfies the boundary value equalities

$$\theta_{i+1}\left(-\frac{N_{i+1}l}{2}\right) = \theta_{i+1}\left(\frac{N_{i+1}l}{2}\right) \tag{3.1.25}$$

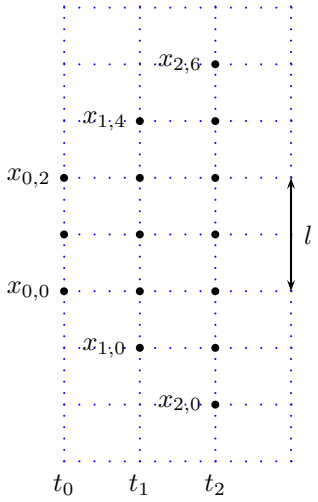
$$\frac{\partial\theta_{i+1}}{\partial x}\left(-\frac{N_{i+1}l}{2}\right) = \frac{\partial\theta_{i+1}}{\partial x}\left(\frac{N_{i+1}l}{2}\right). \tag{3.1.26}$$

Figure 3.1.1: Examples of alternative grids with different values of N_0 and N.

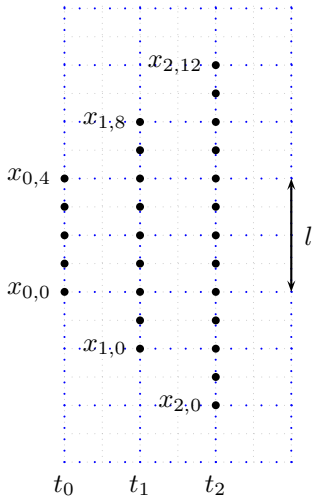
(a) $N_0 = 0$ and $N = 2$.



(b) $N_0 = 1$ and $N = 2$



(c) $N_0 = 1$ and $N = 4$.



Hence, the Fourier integral

$$\hat{\theta}_{i+1}(\nu) = \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu x}\theta_{i+1}(x)dx \tag{3.1.27}$$

is restricted on the interval $[-\frac{N_{i+1}l}{2}, \frac{N_{i+1}l}{2}]$ and discretized using the grid points $\{x_{i+1,k}\}_{k=0}^{N_{i+1}N}$ with a quadrature rule with weights $\{w_k\}_{k=0}^{N_{i+1}N}$. As to the inverse Fourier integral of equation (3.1.24) we restrict it on the interval $[-\frac{L}{2}, \frac{L}{2}]$ and discretize it with lower Riemann sums.

Following the steps of Section 2.3.1, the values of the function θ_i on the grid points $\{x_{i+1,k}\}_{k=0}^{N_{i+1}N-1}$ are given by

$$\theta_i(x_{i+1,k}) \approx (-1)^k \mathfrak{D}^{-1}\left[\{\psi(\nu_{i+1,j})\mathbb{D}[\theta_{i+1}]_j\}_{j=0}^{N_{i+1}N-1}\right]_k \tag{3.1.28}$$

where

$$\mathbb{D}[\theta_{i+1}]_j = \mathfrak{D}\left[\{(-1)^s \tilde{w}_s \theta_{i+1}(x_{i+1,s})\}_{s=0}^{N_{i+1}N-1}\right]_j \tag{3.1.29}$$

and the weights $\{\tilde{w}_j\}_{j=0}^{N_{i+1}N-1}$ are defined as in equation (2.3.10). Consequently, equation (3.1.21) gives

$$\theta_i(x_{ik}) \approx (-1)^{k+\frac{N}{2}} \mathfrak{D}^{-1}\left[\{\psi(\nu_{i+1,j})\mathbb{D}[\theta_{i+1}]_j\}_{j=0}^{N_{i+1}N-1}\right]_{k+\frac{N}{2}} \tag{3.1.30}$$

for $k = 0, 1, ..., N_i N$.

Due to the absence of dampening, we chose

$$\psi(\nu) = \phi(\nu) \tag{3.1.31}$$

when computing the approximate solution u_i and

$$\psi(\nu) = \mathbf{i}\nu\phi(\nu) \tag{3.1.32}$$

when computing the approximate gradient $\dot{u}_i$ for $i = 0, 1, ..., n-1$.

The backward algorithm on the alternative grid is not significantly more complex than the regular grid of Chapter 2. One simply needs to take into account the domain contraction through time steps and discretize the Fourier space accordingly. The following algorithm details the numerical procedure on the alternative grid.

Algorithm 3.1. *Convolution Method on Alternative Grid*

1. *Discretize the restricted real space* $[-\frac{N_n l}{2}, \frac{N_n l}{2}]$ *and the restricted Fourier space* $[-\frac{L}{2}, \frac{L}{2}]$ *with* $N_n N$ *space steps so to have the real space nodes* $\{x_{nk}\}_{k=0}^{N_n N}$ *and* $\{\nu_{nk}\}_{k=0}^{N_n N}$
2. *Value* $u_n(x_{nk}) = g(x_{nk})$
3. *For any* i *from* $n-1$ *to* 0
   (a) *Compute* α *and* β *defining the transform of equation (3.1.8), such that*

$$\theta_{i+1} = (u_{i+1})^{\alpha,\beta} \tag{3.1.33}$$

   and θ_{i+1} *satisfies the boundary conditions of equations (3.1.25) and (3.1.26).*

   (b) *Compute* $\theta_i(x_{ik})$ *through equation (3.1.30) for* $k = 0, 1, ..., N_i N$ *with*

$$\psi(\nu) = \phi(\nu) \tag{3.1.34}$$

   and retrieve the values $\tilde{u}_{ik}$ *as*

$$\tilde{u}_{ik} = \theta_i(x_{ik}) - \alpha(x_{i,k}^2 - \Delta_i)\beta x_{i,k}. \tag{3.1.35}$$

(c) Compute $\theta_i(x_{ik})$ through equation (3.1.30) for $k = 0, 1, ..., N_iN$ with

$$\psi(\nu) = \mathbf{i}\nu\phi(\nu) \tag{3.1.36}$$

and retrieve the values $\dot{u}_{ik}$ as

$$\dot{u}_{ik} = \theta_i(x_{ik}) - (2\alpha x_{ik} + \beta)\,. \tag{3.1.37}$$

(d) Compute the values u_{ik} as

$$u_{ik} = \tilde{u}_{ik} + \Delta_i f(t_i, \tilde{u}_{ik}, \dot{u}_{ik}) \tag{3.1.38}$$

for $k = 0, 1, ..., N_iN$ through equation (3.1.2) when using the explicit Euler scheme 1 or as

$$u_{ik} = \tilde{u}_{ik} + \Delta_i f(t_i, u_{ik}, \dot{u}_{ik}) \tag{3.1.39}$$

through equation (3.1.5) under the implicit Euler scheme.

(e) Update the real space grid with equation (3.1.21) and the Fourier space grid by discretizing the interval $[-\frac{L}{2}, \frac{L}{2}]$ with N_iN space steps so to have the real space nodes $\{x_{ik}\}_{k=0}^{N_iN}$ and $\{\nu_{ik}\}_{k=0}^{N_iN}$.

The algorithm produces the numerical solutions $\{u_{ik}\}_{k=0}^{N_iN}$, $\{\tilde{u}_{ik}\}_{k=0}^{N_iN}$ and $\{\dot{u}_{ik}\}_{k=0}^{N_iN}$, $i = 0, 1, ..., n-1$. The next section deals with error considerations under the alternative discretization.

3.2 Error analysis

First, we give a bound for the local discretization error. Let $\{\mathbf{u}_{ik}\}_{k=0}^{N_iN}$, $\{\tilde{\mathbf{u}}_{ik}\}_{k=0}^{N_iN}$ and $\{\dot{\mathbf{u}}_{ik}\}_{k=0}^{N_iN}$ denote the numerical solutions obtained from the convolution method at time mesh t_i given the solution u_{i+1} at time t_{i+1}. For the convolution method on the alternative grid, we defined the local discretization error as

$$E_{ik} := |u_i(x_k) - \mathbf{u}_{ik}| + |\dot{u}_i(x_k) - \dot{\mathbf{u}}_{ik}| \tag{3.2.1}$$

for $i = 0, 1, ..., n-1$ and $k = 0, 1, ..., N_iN$.

Theorem 3.3. *Suppose that the driver $f \in \mathcal{C}^{1,2,2}$ and the terminal condition $g \in \mathcal{C}^2$. Then the convolution method yields a discretization error of the form*

$$E_{ik} = \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-K|\Delta_i|^{-1}l^2}\right) \tag{3.2.2}$$

for some constant $K > 0$ on the alternative grid and under the trapezoidal quadrature rule.

Proof. The proof of this result is similar to the proof of Theorem 2.7 and strongly relies on the ideas developped in Appendix E.3. We suppose the solution u_{i+1} at time t_{i+1} is known. The solution $u_{i+1} \in \mathcal{C}^2$ is twice differentiable since $f \in \mathcal{C}^{1,2,2}$ and $g \in \mathcal{C}^2$. Also, u_{i+1} is square integrable with respect to the Gaussian density.

As in Theorem 2.7, we limit ourselves to the case where

$$u_{i+1}\left(-\frac{N_{i+1}l}{2}\right) = u_{i+1}\left(\frac{N_{i+1}l}{2}\right) \text{ and } \frac{\partial u_{i+1}}{\partial x}\left(-\frac{N_{i+1}l}{2}\right) = \frac{\partial u_{i+1}}{\partial x}\left(\frac{N_{i+1}l}{2}\right)$$

so that the coefficients of the transform are $\alpha = \beta = 0$. Let T_i be the Fourier polynomial interpolating u_{i+1} on $\left[-\frac{N_{i+1}l}{2}, \frac{N_{i+1}l}{2}\right]$ such that

$$T_i(x) := T_{NN_{i+1}}[u_{i+1}](x),\ x \in \mathbb{R}. \tag{3.2.3}$$

We have that

$$\tilde{u}_i(x_{ik}) \quad = \quad \int_{|y|\leq\frac{l}{2}} u_{i+1}(x_{ik}+y)h(y)dy + \int_{|y|>\frac{l}{2}} u_{i+1}(x_{ik}+y)h(y)dy$$

where

$$\int_{|y|>\frac{l}{2}} u_{i+1}(x_{ik}+y)h(y)dy \quad = \quad \mathcal{O}\left(e^{-Kl^2}\right)$$

for some constant $K > 0$ which is inversely proportional to Δ_i by the Cauchy-Schwartz and Chernoff inequalities since the solution u_{i+1} is square integrable. Hence

$$\begin{aligned}
\tilde{u}_i(x_{ik}) &= \int_{|y|\leq\frac{l}{2}} u_{i+1}(x_{ik}+y)h(y)dy + \mathcal{O}\left(e^{-Kl^2}\right) \\
&= \int_{|y|\leq\frac{l}{2}} T_i(x_{ik}+y)h(y)dy + \mathcal{O}\left(\Delta x\right) + \mathcal{O}\left(e^{-Kl^2}\right) \\
&\quad \text{(by Proposition E.10),} \\
&= \int_{\mathbb{R}} T_i(x_{ik}+y)h(y)dy - \int_{|y|>\frac{l}{2}} T_i(x_{ik}+y)h(y)dy \\
&\quad + \mathcal{O}\left(\Delta x\right) + \mathcal{O}\left(e^{-Kl^2}\right) \\
&= \int_{\mathbb{R}} T_i(x_{ik}+y)h(y)dy + \mathcal{O}\left(\Delta x\right) + \mathcal{O}\left(e^{-Kl^2}\right) \\
&\quad \text{(by Chernoff's inequality, since } T_i \text{ is bounded),} \\
&= \int_{\mathbb{R}} \sum_{j=-\frac{N_{i+1}N}{2}}^{\frac{N_{i+1}N}{2}-1} d_j e^{\mathbf{i}j\frac{2\pi}{N_{i+1}l}(x_{i,k}+y)} h(y)dy + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right) \\
&= \sum_{j=-\frac{N_{i+1}N}{2}}^{\frac{N_{i+1}N}{2}-1} d_j e^{\mathbf{i}j\frac{2\pi}{N_{i+1}l}x_{i,k}} \phi\left(j\frac{2\pi}{N_{i+1}l}\right) + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right) \\
&= \sum_{j=-\frac{N_{i+1}N}{2}}^{\frac{N_{i+1}N}{2}-1} d_j e^{\mathbf{i}j\frac{2\pi}{N_{i+1}l}x_{i+1,k+\frac{N}{2}}} \phi\left(j\frac{2\pi}{N_{i+1}l}\right) + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right) \\
&= (-1)^{k+\frac{N}{2}} \sum_{j=0}^{N_{i+1}N-1} \phi(\nu_{i+1,j})(-1)^{j-\frac{N_{i+1}N}{2}} d_{j-\frac{N_{i+1}N}{2}} e^{\mathbf{i}\frac{2\pi}{N_{i+1}N}j(k+\frac{N}{2})} \\
&\quad + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right) \\
&= (-1)^{k+\frac{N}{2}} \sum_{j=0}^{N_{i+1}N-1} \phi(\nu_{i+1,j})\mathbb{D}[u_{i+1}]_j e^{\mathbf{i}\frac{2\pi}{N_{i+1}N}j(k+\frac{N}{2})} \\
&\quad + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right) \\
&\quad \text{(by Proposition E.9 when using the trapezoidal quadrature rule),} \\
&= \tilde{\mathbf{u}}_{ik} + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl^2}\right).
\end{aligned}$$

Similar techniques show that

$$\dot{u}_i(x_k) = \dot{\mathbf{u}}_{ik} + \mathcal{O}\left(\Delta x\right) + \mathcal{O}\left(e^{-Kl^2}\right) \tag{3.2.4}$$

where $K > 0$ is inversely proportional to Δ_i. The Lipschitz property of the driver f completes the proof. □

As expected, the alternative discretization improves the local error bound by eliminating extrapolation errors. The result of Theorem 3.3 establishes the consistency of the convolution

method with respect to the approximate functions u_i and gradients $\dot{u}_i$. Hence, the convolution method is consistent to the PDE solution u and its derivative $\frac{\partial u}{\partial x}$ since the time discretization is itself convergent.

Furthermore, the absence of extrapolation errors in the local discretization allows us to develop a bound for the global discretization error. The following corollary proves helpful when deriving the global discretization error bound.

Corollary 3.4. *Under the conditions of Theorem 3.3,*

$$\sup_{i,k} E_{i,k} = \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right) \tag{3.2.5}$$

where $C > 0$ and $|\pi| = \sup_i \Delta_i$.

We define the global error as

$$E_{l,\Delta x} := \sup_{i,k} e_{ik} + \sup_{i,k} \dot{e}_{ik} \tag{3.2.6}$$

where

$$e_{ik} = |u_{n-i}(x_k) - u_{n-i,k}| \tag{3.2.7}$$

and

$$\dot{e}_{ik} = |\dot{u}_{n-i}(x_k) - \dot{u}_{n-i,k}| \tag{3.2.8}$$

for $i = 1, ..., n$ with $e_{0,k} = \dot{e}_{0,k} = 0$. The next theorem describes the stability and convergence properties of the convolution method.

Theorem 3.5. *Suppose the conditions of Theorem 3.3 are satisfied. If the discretization is such that*

$$\sup_i \max\left(\frac{\Delta x}{\sqrt{2\pi\Delta_i}}, \frac{\Delta x}{\pi\Delta_i}\right) \leq 1 \tag{3.2.9}$$

then the convolution method is stable and the global discretization error $E_{l,\Delta x}$ satisfies

$$E_{l,\Delta x} = \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right) \tag{3.2.10}$$

where $C > 0$.

Proof. First note that from the definitions of equations (3.2.1) and (3.2.7)

$$\begin{aligned} e_{ik} &\leq E_{n-i,k} + |\mathbf{u}_{n-i,k} - u_{n-i,k}| \\ &\leq E_{n-i,k} + (1 + \Delta_i K)\,|\tilde{\mathbf{u}}_{n-i,k} - \tilde{u}_{n-i,k}| \\ &\quad + \Delta_i K\,|\dot{\mathbf{u}}_{n-i,k} - \dot{u}_{n-i,k}| \end{aligned} \tag{3.2.11}$$

where $K > 0$ is the Lipschitz constant of the driver f. Also, we have that

$$\dot{e}_{ik} \leq E_{n-i,k} + |\dot{\mathbf{u}}_{n-i,k} - \dot{u}_{n-i,k}|\,. \tag{3.2.12}$$

from equations (3.2.1) and (3.2.8).

Furthermore, the construction of the convolution method gives

$$\begin{aligned} |\tilde{\mathbf{u}}_{i,k} - \tilde{u}_{i,k}| &\leq \left|\mathfrak{D}^{-1}\left[\{\phi(\nu_{i+1,j})\mathbb{D}[u_{i+1} - u_{i+1,s}]_j\}_{j=0}^{N_{i+1}N-1}\right]_{k+\frac{N}{2}}\right| + E_{i,k} \\ &\quad \text{(by Theorem 3.3 since the transform function is given),} \\ &\leq \frac{1}{N_{i+1}N}\left(\sum_{j=0}^{N_{i+1}N-1} \phi(\nu_{i+1,j})\right) \sup_k |u_{i+1}(x_{i,k}) - u_{i+1,k}| + E_{i,k} \\ &\quad \text{(using the matrix-vector representation of DFTs),} \end{aligned}$$

$$
\begin{aligned}
&\leq \frac{1}{N_{i+1}N}\left(\sum_{j=0}^{N_{i+1}N-1}\phi(\nu_{i+1,j})\right)\sup_k e_{n-i-1,k}+E_{i,k}\\
&\leq \frac{(\Delta\nu_{i+1})^{-1}}{N_{i+1}N}\left(\int_{\mathbb{R}}\phi(x)dx\right)\sup_k e_{n-i-1,k}+E_{i,k}\\
&= \frac{\Delta x}{(2\pi\Delta_i)^{\frac{1}{2}}}\sup_k e_{n-i-1,k}+E_{i,k}. \qquad (3.2.13)
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
|\dot{\mathbf{u}}_{i,k}-\dot{u}_{i,k}| &\leq \left|\mathfrak{D}^{-1}\left[\{\psi(\nu_{i+1,j})\mathbb{D}[u_{i+1}-u_{i+1,s}]_j\}_{j=0}^{N_{i+1}N-1}\right]_{k+\frac{N}{2}}\right|+E_{i,k}\\
&\quad \text{(by Theorem 3.3 since the transform function is given),}\\
&\leq \frac{1}{N_{i+1}N}\left(\sum_{j=0}^{N_{i+1}N-1}|\nu_{i+1,j}|\,\phi(\nu_{i+1,j})\right)\sup_k e_{n-i-1,k}+E_{i,k}\\
&\quad \text{(using the matrix representation of DFTs),}\\
&\leq \frac{(\Delta\nu_{i+1})^{-1}}{N_{i+1}N}\left(\int_{\mathbb{R}}|x|\,\phi(x)dx\right)\sup_k e_{n-i-1,k}+E_{i,k}\\
&= \frac{\Delta x}{\pi\Delta_i}\sup_k e_{n-i-1,k}+E_{i,k}. \qquad (3.2.14)
\end{aligned}
$$

Then, combining the inequalities of equations (3.2.11), (3.2.13) and (3.2.14) leads to

$$
\begin{aligned}
e_{i,k} &\leq C_0E_{i,k}+(1+2\Delta_iK)\max\left(\frac{\Delta x}{\sqrt{2\pi\Delta_i}},\frac{\Delta x}{\pi\Delta_i}\right)\sup_k e_{i-1,k}\\
&\leq C_0\sup_{i,k}E_{i,k}+(1+2\Delta_iK)\max\left(\frac{\Delta x}{\sqrt{2\pi\Delta_i}},\frac{\Delta x}{\pi\Delta_i}\right)\sup_k e_{i-1,k}
\end{aligned}
$$

where $C_0>0$ and $K>0$ is the Lipschitz constant of the driver f. So that

$$
\begin{aligned}
\sup_k e_{i,k} &\leq C_0\sup_{i,k}E_{i,k}+(1+2\Delta_iK)\max\left(\frac{\Delta x}{\sqrt{2\pi\Delta_i}},\frac{\Delta x}{\pi\Delta_i}\right)\sup_k e_{i-1,k}\\
&\leq C_0\sup_{i,k}E_{i,k}+(1+2\Delta_iK)\zeta\sup_k e_{i-1,k} \qquad (3.2.15)
\end{aligned}
$$

for some positive number ζ satisfying

$$
\sup_i\max\left(\frac{\Delta x}{\sqrt{2\pi\Delta_i}},\frac{\Delta x}{\pi\Delta_i}\right)\leq\zeta\leq 1.
$$

From the inequality of equation (3.2.15), Gronwall's Lemma yields

$$
\sup_k e_{i,k}\leq C_0e^{2TK}\sup_{i,k}E_{i,k} \qquad (3.2.16)
$$

for $i=0,1,...,n$ knowing that $e_{0,k}=0$. Hence, the convolution method is stable for the approximate solution u_i since its error at any time step is absolutely bounded.

The inequalities of equations (3.2.12), (3.2.14) and (3.2.16) lead to

$$
\begin{aligned}
\sup_k \dot{e}_{i,k} &\leq \left(C_1+\frac{\Delta x}{\pi\Delta_i}C_0e^{2TK}\right)\sup_{i,k}E_{i,k}\\
&\leq \left(C_1+C_0e^{2TK}\right)\sup_{i,k}E_{i,k} \qquad (3.2.17)
\end{aligned}
$$

for a positive constant $C_1>0$. Hence, the convolution method is also stable for the approximate gradient $\dot{u}_i$.

The result of equation (3.2.10) follows by taking the supremum on the left hand sides of equations (3.2.16) and (3.2.17) other time steps and applying Corollary 3.4. □

Similar to most explicit methods for PDEs[1], the convolution method displays a stability condition described in equation (3.2.9). This condition is actually weaker compared to other methods, especially the explicit finite difference method since the condition is easily satisfied. In general, Theorem 3.5 shows that the convolution method converges when the space discretization is relatively as fine as the time discretization and/or the square root of the time discretization.

Other numerical methods for BSDEs, and particularly Monte Carlo based method, have a stability and convergence condition. Indeed, error explosion occurs for fine time discretizations in the backward methods of Gobet et al. [53] and Bouchard and Touzi [20]. In order to maintain stability and convergence, the space discretization has to be refined by increasing the number of simulated paths.

The next corollary gives a method of discretization to produce a convergent scheme. It also confirms that the major variables that impact convergence are the maximal time step $|\pi|$ and the space step Δx. The length of the truncated domain l has but a negligible effect on the global error $E_{l,\Delta x}$ for fine time discretizations. Nonetheless, large values of l are preferable to improve speed of convergence but also for simulation reasons as we shall see in Section 3.3.

Corollary 3.6. *If the space discretization is such that*

$$\Delta x \leq \sqrt{2\pi} \min_{0 \leq i < n} \Delta_i \tag{3.2.18}$$

then the condition of equation (3.2.9) holds and the convolution method applied on Scheme 1 converges on the alternative grid as $|\pi| \to 0$.

Proof. Assume without loss of generality that $|\pi| < 1$ then we have that

$$\begin{aligned} \max\left(\frac{\Delta x}{\sqrt{2\pi\Delta_i}}, \frac{\Delta x}{\pi\Delta_i}\right) &< \frac{\Delta x}{\sqrt{2\pi}\Delta_i} \\ &\leq 1 \end{aligned}$$

for $i = 0, 1, 2, ..., n-1$ where the last inequality holds by equation (5.3.9). Hence, the stability condition of equation (3.2.9) holds.

Clearly $|\pi| + \Delta x \to 0$, and consequently $E_{l,\Delta x} \to 0$ by Theorem 3.5, as $|\pi| \to 0$. □

As already mentioned in Chapter 2, the implicit Euler scheme will provide numerical solutions with similar properties as those developed in Theorem 3.3 and Theorem 3.5. Indeed, the convergence properties under the implicit Euler scheme can be established when the time discretization satisfy the condition

$$|\pi| K < 1 \tag{3.2.19}$$

where K is the Lipschitz constant of the driver f. The numerical solution to the BSDE is explicitly defined in the next section.

3.3 Simulation of BSDEs

The availability of approximations for the functions u_i and $\dot{u}_i$, $i = 0, 1, 2, .., n$ allows us to simulate the BSDE. A numerical approximation of the BSDE solution can indeed be constructed

[1] For instance, the explicit finite difference method on the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

The space step Δx and the time step Δ have to satisfy an a stability and convergence condition of the form

$$\Delta \leq \frac{1}{2}\Delta x^2$$

as shown by the Von Neumann stability analysis of the scheme in Tveito and Winther [110], page 132.

from the numerical solution of the approximate solution $\{u_{ik}\}_{k=0}^{N_iN}$ and the approximate gradient $\{\dot{u}_{ik}\}_{k=0}^{N_iN}$, $i = 0, 1, 2, ..., n$. The approach is similar to most four step scheme (or PDE) based methods for solving BSDEs.

Let $(U_i, \dot{U}_i)$ be the extended solution at time mesh t_i defined on the truncated interval

$$\mathcal{I}_i = [-\frac{1}{2}N_i l, \frac{1}{2}N_i l]. \tag{3.3.1}$$

More precisely, $U_i : \mathcal{I}_i \to \mathbb{R}$ (resp. $\dot{U}_i : \mathcal{I}_i \to \mathbb{R}$) is the function obtained by linearly interpolating the approximate solution $\{u_{ik}\}_{k=0}^{N_iN}$ (resp. the approximate gradient $\{\dot{u}_{ik}\}_{k=0}^{N_iN}$) on the space grid[2]. Hence, if $x \in [x_{ik}, x_{i,k+1}]$ and $N_0 > 0$ then

$$U_i(x) = u_{ik} + \frac{u_{i,k+1} - u_{ik}}{x_{i,k+1} - x_{ik}}(x - x_{ik}) \tag{3.3.2}$$

and similarly

$$\dot{U}_i(x) = \dot{u}_{ik} + \frac{\dot{u}_{i,k+1} - \dot{u}_{ik}}{x_{i,k+1} - x_{ik}}(x - x_{ik}) \tag{3.3.3}$$

for $i = 0, 1, ..., n-1$ and $k = 0, 1, ..., N_iN - 1$. In the particular case where $N_0 = 0$, we set

$$U_0(x) = u_{00}\delta_{x,0} \tag{3.3.4}$$

$$\dot{U}_0(x) = \dot{u}_{00}\delta_{x,0} \tag{3.3.5}$$

so that there is no interpolation at time $t_0 = 0$.

An extended solution for the approximate solution was also defined by Douglas et al. [40] using linear interpolation. Nonetheless, Douglas et al. [40] consider extensions on both time and space. Also, the truncation inherent to the convolution method forces us to set the extended solution to a *graveyard value* outside the space grid. Here, we set it as zero (0) but the boundary values of the numerical solution can be used.

An important feature of the extension is that it does not introduce an additional error term since the interpolation error is of a lesser order (in space) than the global discretization error $E_{l,\Delta x}$. The following theorem gives the interpolation (quadratic) error bound.

Theorem 3.7. *Suppose the driver $f \in \mathcal{C}^{1,2,2}$, the terminal condition $g \in \mathcal{C}^2$ and the stability and convergence condition of equation (3.2.9) is satisfied. Then*

$$\begin{aligned} \sup_i \sup_{x \in \mathcal{I}_i} |u_i(x) - U_i(x)|^2 \quad + \quad & \sup_i \sup_{x \in \mathcal{I}_i} \left|\dot{u}_i(x) - \dot{U}_i(x)\right|^2 \\ & = \mathcal{O}(\Delta x^2) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right) \end{aligned} \tag{3.3.6}$$

where $C > 0$.

Proof. First note that since $f \in \mathcal{C}^{1,2,2}$ and $g \in \mathcal{C}^2$, we have that $u_i \in \mathcal{C}^2$ and $\dot{u}_i \in \mathcal{C}^2$ are both twice differentiable (using the Leibniz integral rule recursively). Hence, for $x \in [x_{ik}, x_{i,k+1}]$

$$\begin{aligned} u_i(x) - U_i(x) &= u_i(x) - u_{ik} - \frac{u_{i,k+1} - u_{ik}}{x_{i,k+1} - x_{ik}}(x - x_{ik}) \\ &= u_i(x) - u_i(x_{ik}) - \frac{u_i(x_{i,k+1}) - u_i(x_{ik})}{x_{i,k+1} - x_{ik}}(x - x_{ik}) \\ &\quad + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right) \end{aligned}$$

[2]The numerical solutions $\{u_{ik}\}_{k=0}^{N_iN}$ and $\{\dot{u}_{ik}\}_{k=0}^{N_iN}$ are actually complex numbers. Since they approximate real values, their imaginary parts are of the order of the global discretization error $E_{l,\Delta x}$ and hence negligible. For this reason, we define the functions U_i and $\dot{U}_i$ as real-valued functions using the projection of the approximate solutions onto the real line.

$$
\begin{aligned}
&\text{(by Theorem 3.5)},\\
= \;& \mathcal{O}(\Delta x^2) + \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right)\\
&\text{(since } u_i \text{ is twice differentiable)},\\
= \;& \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right).
\end{aligned}
$$

If $N_0 = 0$, the last equation obviously holds at $x = 0$.

Clearly,

$$\sup_i \sup_{x\in\mathcal{I}_i} |u_i(x) - U_i(x)|^2 = \mathcal{O}(\Delta x^2) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right) \tag{3.3.7}$$

since the $\mathcal{I}_i$ are bounded intervals.

The same techniques show that

$$\sup_i \sup_{x\in\mathcal{I}_i} \left|\dot{u}_i(x) - \dot{U}_i(x)\right|^2 = \mathcal{O}(\Delta x^2) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right) \tag{3.3.8}$$

and the result follows. □

The approximation of the BSDE solution (Y, Z) is then given by the couple (y, z) defined on $[0, T)$ as

$$y_t = \sum_{i=0}^{n-1} U_i(W_{t_i})\mathbf{1}_{[t_i,t_{i+1})}(t) \tag{3.3.9}$$

$$z_t = \sum_{i=0}^{n-1} \dot{U}_i(W_{t_i})\mathbf{1}_{[t_i,t_{i+1})}(t) \tag{3.3.10}$$

where $\{W_t\}_{t\in[0,T]}$ is a standard Brownian motion. By the definition of (y, z) and the continuity of linear interpolations, and hence of the extended functions $U_i : \mathcal{I}_i \to \mathbb{R}$ and $\dot{U}_i : \mathcal{I}_i \to \mathbb{R}$, we obtain the following corollary.

Corollary 3.8. *The processes (y, z) are both $\mathbb{F}-$adapted, càdlàg and constant on any interval $[t_i, t_{i+1})$, $i = 0, 1, ..., n-1$. Moreover,*

$$(y, z) \in L_S^\infty(\mathbb{R}^2) \tag{3.3.11}$$

i.e the processes are bounded.

The (quadratic) error on the BSDE solution is defined as

$$E^2_{\pi,l,\Delta x} := \max_{0\le i<n} \mathbf{E}\left[\sup_{t\in[t_i,t_{i+1})} |Y_t - y_t|^2\right] + \sum_{i=0}^{n-1} \mathbf{E}\left[\int_{t_i}^{t_{i+1}} |Z_s - z_s|^2 \, ds\right] \tag{3.3.12}$$

and characterized in the next theorem.

Theorem 3.9. *Suppose the driver $f \in \mathcal{C}^{1,2,2}$, the terminal condition $g \in \mathcal{C}^2$, the stability and convergence condition of equation (3.2.9) is satisfied and*

$$\xi = g(W_T) \in L^4 \tag{3.3.13}$$

then

$$E^2_{\pi,l,\Delta x} = \mathcal{O}(|\pi|) + \mathcal{O}(\Delta x^2) + \mathcal{O}\left(e^{-C(N_0+1)^2l^2} + e^{-C|\pi|^{-1}l^2}\right) \tag{3.3.14}$$

where $C > 0$.

Proof. We have that

$$
\begin{aligned}
E^2_{\pi,l,\Delta x} &\leq 2E^2_\pi + 2 \max_{0\leq i<n} \mathbf{E}\left[\sup_{t\in[t_i,t_{i+1}]} \left|Y^\pi_{t_i} - y_t\right|^2\right] \\
&\quad + 2\sum_{i=0}^{n-1} \mathbf{E}\left[\int_{t_i}^{t_{i+1}} \left|Z^\pi_{t_i} - z_s\right|^2 ds\right] \\
&\leq 2 \max_{0\leq i<n} \mathbf{E}\left[\sup_{t\in[t_i,t_{i+1})} \left|Y^\pi_{t_i} - y_t\right|^2\right] + 2\sum_{i=0}^{n-1} \mathbf{E}\left[\int_{t_i}^{t_{i+1}} \left|Z^\pi_{t_i} - z_s\right|^2 ds\right] \\
&\quad + C_1 |\pi| \text{ (by the result of equation I.16)}, \\
&= 2 \max_{0\leq i<n} \mathbf{E}\left[\left|Y^\pi_{t_i} - y_{t_i}\right|^2\right] + 2\sum_{i=0}^{n-1} \Delta_i \mathbf{E}\left[\left|Z^\pi_{t_i} - z_{t_i}\right|^2\right] + C_1 |\pi| \\
&\leq 2 \max_{0\leq i<n-1} \mathbf{E}\left[|u_i(W_{t_i}) - U_i(W_{t_i})|^2 \mathbf{1}_{\mathcal{I}_i}(W_{t_i})\right] \\
&\quad + 2\sum_{i=0}^{n-1} \Delta_i \mathbf{E}\left[\left|\dot{u}_i(W_{t_i}) - \dot{U}_i(W_{t_i})\right|^2 \mathbf{1}_{\mathcal{I}_i}(W_{t_i})\right] \\
&\quad + 2 \max_{0\leq i<n-1} \mathbf{E}\left[\left|Y^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] + 2\sum_{i=0}^{n-1} \Delta_i \mathbf{E}\left[\left|Z^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] \\
&\quad + C_1 |\pi| \\
&\leq 2 \max_{0\leq i<n} \sup_{x\in\mathcal{I}_i} |u_i(x) - U_i(x)|^2 + 2T \max_{0\leq i<n-1} \sup_{x\in\mathcal{I}_i} \left|\dot{u}_i(x) - \dot{U}_i(x)\right|^2 \\
&\quad + 2 \max_{0\leq i<n} \mathbf{E}\left[\left|Y^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] + 2\sum_{i=0}^{n-1} \Delta_i \mathbf{E}\left[\left|Z^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] \\
&\quad + C_1 |\pi| \\
&= 2 \max_{0\leq i<n} \mathbf{E}\left[\left|Y^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] + 2\sum_{i=0}^{n-1} \Delta_i \mathbf{E}\left[\left|Z^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] \\
&\quad + C_1 |\pi| + C_2 \Delta x^2 + C_3 e^{-C_4 |\pi|^{-1} l^2} \\
&\quad \text{(by Theorem 3.7)}, \\
&\leq 2 \max_{0\leq i<n} \mathbf{E}\left[\left|Y^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] + 2T \max_{0\leq i<n} \mathbf{E}\left[\left|Z^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] \\
&\quad + C_1 |\pi| + C_2 \Delta x^2 + C_3 e^{-C_4 |\pi|^{-1} l^2} \\
&\leq 2 \max_{1\leq i<n} \mathbf{E}\left[\left|Y^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] + 2T \max_{1\leq i<n} \mathbf{E}\left[\left|Z^\pi_{t_i}\right|^2 \mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right] \\
&\quad + C_1 |\pi| + C_2 \Delta x^2 + C_3 e^{-C_4 |\pi|^{-1} l^2} \\
&\quad \text{(since } W_0 \in \mathcal{I}_0), \\
&\leq C_1 |\pi| + C_2 \Delta x^2 + C_3 e^{-C_4 |\pi|^{-1} l^2} + C_5 \max_{1\leq i<n} \mathbf{E}\left[\mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(W_{t_i})\right]^{\frac{1}{2}} \\
&\quad \text{(by the Cauchy-Schwartz inequality and Lemma 2.1)}, \\
&\leq C_1 |\pi| + C_2 \Delta x^2 + C_3 e^{-C_4 |\pi|^{-1} l^2} + C_5 \max_{1\leq i<n} e^{-C_6 {t_i}^{-1} N_i^2 l^2} \\
&\quad \text{(by Chernoff's inequality)}, \\
&\leq C_1 |\pi| + C_2 \Delta x^2 + C_3 e^{-C_4 |\pi|^{-1} l^2} + C_5 e^{-C_6 T^{-1} (N_0+1)^2 l^2}.
\end{aligned}
$$

This last inequality is equivalent to the result. □

As shown in Theorem 3.9, three different error terms compose the simulation (quadratic) error $E^2_{|\pi|,l,\Delta x}$: the time discretization error, the space discretization error and the truncation error. The time discretization error appears naturally knowing that the convolution method is based on an explicit Euler scheme and is of first order as is the original Euler scheme. The space

discretization error, of second order, results from both the global discretization error $E_{l,\Delta x}$ and the quadratic form of the simulation error. The truncation error is influenced by the probability that the Brownian motion path exits the space grid at a given time step. It remains spectral of index 2 with respect to the truncation length l, and to some extent with respect to the length of the truncated domain at the first time step $(N_0+1)l$. Thus, increasing N_0 improves the speed of convergence. However, convergence is assured for large values of l as long as the assumptions of Theorem 3.9 are satisfied.

In particular, the condition of equation (3.3.13) and its use in the proof of Theorem 3.9 indicate that the simulation error $E^2_{|\pi|,l,\Delta x}$ is controllable outside the space grid when the approximate backward process $Y^\pi_{t_i}$ and the approximate control process $Z^\pi_{t_i}$ are in L^4. These conditions are satisfied for a large range of BSDEs including BSDEs with terminal function g of exponential growth such that

$$|g(x)| \leq Ce^{p|x|},\ x \in \mathbb{R} \tag{3.3.15}$$

for some constants $C > 0$ and $p \geq 0$. Unfortunately, the BSDE well-posedness does not require the condition of equation (3.3.13). Thus, the convolution method may not converge for some well-posed BSDEs in the sense of the simulation error $E^2_{|\pi|,l,\Delta x}$ in equation (3.3.12).

We can define an alternate simulation error by discarding the occurrences of the Brownian motion outside the alternative grid. Then, the alternate simulation (quadratic) error $\bar{E}^2_{|\pi|,l,\Delta x}$ is

$$\begin{aligned}\bar{E}^2_{|\pi|,l,\Delta x} \quad := \quad & \max_{0\leq i<n} \mathbf{E}\left[\sup_{t\in[t_i,t_{i+1})} |Y_t - y_t|^2 \mathbf{1}_{\mathcal{I}_i}(W_{t_i})\right] \\ & + \sum_{i=0}^{n-1} \mathbf{E}\left[\int_{t_i}^{t_{i+1}} |Z_s - z_s|^2 \mathbf{1}_{\mathcal{I}_i}(W_{t_i})ds\right].\end{aligned} \tag{3.3.16}$$

From the proof of Theorem 3.9, the following corollary stands.

Corollary 3.10. *Suppose the driver $f \in \mathcal{C}^{1,2,2}$, the terminal condition $g \in \mathcal{C}^2$ and the stability and convergence condition of equation (3.2.9) is satisfied, then*

$$\bar{E}^2_{\pi,l,\Delta x} = \mathcal{O}(|\pi|) + \mathcal{O}(\Delta x^2) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right) \tag{3.3.17}$$

where $C > 0$.

3.4 Extensions

In this section, we discuss further extensions of the convolution method on the alternative grid. These extensions include Reflected BSDEs and also BSDES based on an arithmetic Brownian motion.

3.4.1 Simulation of RBSDEs

We defined, in Section 2.5, explicit schemes for the RBSDEs

$$\begin{cases} -dY_t = f(t,Y_t,Z_t,)dt + dA_t - Z_t dW_t \\ Y_t \geq B_t\ ,\ dA_t \geq 0\ ,\ \forall t \in [0,T] \\ \int_0^T (Y_t - B_t)dA_t = 0\ ,\ Y_T = g(X_T) \end{cases} \tag{3.4.1}$$

where

$$B_t = B(t,W_t) \tag{3.4.2}$$

for a deterministic function $B : [0,T] \times \mathbb{R} \to \mathbb{R}$. The time discretization of the RBSDE through Euler scheme 1 is essentially equivalent to

$$u_i(x) = \tilde{u}_i(x) + \Delta_i f(t_i, \tilde{u}_i(x), \dot{u}_i(x)) + \Delta\bar{u}_i(x) \tag{3.4.3}$$

with

$$\dot{u}_i(x) = \frac{1}{\Delta_i}\int_{-\infty}^{\infty}(y-x)u_{i+1}(x+y)h(y)dy \tag{3.4.4}$$

$$\tilde{u}_i(x) = \int_{-\infty}^{\infty}u_{i+1}(x+y)h(y)dy \tag{3.4.5}$$

$$\Delta\bar{u}_i(x) = [\tilde{u}_i(x) + \Delta_i f(t_i, \tilde{u}_i(x), \dot{u}_i(x)) - B(t_i,x)]^- \tag{3.4.6}$$

for $i = 0, 1, ..., n-1$ and $u_n(x) = g(x)$.

The convolution method on the alternative grid provides numerical estimates for the approximate solution $\{u_{ik}\}_{k=0}^{N_iN}$, the approximate gradient $\{\dot{u}_{ik}\}_{k=0}^{N_iN}$ and the approximate increment $\{\Delta\bar{u}_{ik}\}_{k=0}^{N_iN}$, $i = 0, 1, ..., n-1$. We define the extended functions U_i and $\dot{U}_i$ as in equations (3.3.2) and (3.3.3). An additional extended function $\Delta\bar{U}_i : \mathcal{I}_i \to \mathbb{R}$ is defined by linearly interpolating the values $\{\Delta\bar{u}_{ik}\}_{k=0}^{N_iN}$ on the space grid at time mesh t_i. Hence, for any $x \in [x_{ik}, x_{i,k+1}]$,

$$\Delta\bar{U}_i(x) = \Delta\bar{u}_{ik} + \frac{\Delta\bar{u}_{i,k+1} - \Delta\bar{u}_{ik}}{x_{i,k+1} - x_{ik}}(x - x_{ik}) \tag{3.4.7}$$

for $i = 0, 1, .., n-1$. When $N_0 = 0$, we simply have

$$\Delta\bar{U}_0(x) = \Delta\bar{u}_{00}\delta_{x,0}. \tag{3.4.8}$$

A numerical approximation of the RBSDE solution (Y, Z, A) consists of the triplet (y, z, a) where

$$y_t = \sum_{i=0}^{n-1} U_i(W_{t_i})\mathbf{1}_{[t_i,t_{i+1})}(t) \tag{3.4.9}$$

$$z_t = \sum_{i=0}^{n-1} \dot{U}_i(W_{t_i})\mathbf{1}_{[t_i,t_{i+1})}(t) \tag{3.4.10}$$

$$a_t = \sum_{i=0}^{n-1} \Delta\bar{U}_i(W_{t_i})\mathbf{1}_{[t_i,T)}(t) \tag{3.4.11}$$

and $\{W_t\}_{t\in[0,T]}$ is a standard Brownian motion. By definition, the triple of processes $\{(y_t, z_t, a_t)\}_{t\in[0,T)}$ is $\mathbb{F}-$adapted and càdlàg. Both processes $\{y_t\}_{t\in[0,T)}$ and $\{z_t\}_{t\in[0,T)}$ are in L_S^∞ (bounded) whereas the process $\{a_t\}_{t\in[0,T)}$ is non-decreasing.

3.4.2 Arithmetic Brownian motion

When the forward process is the arithmetic Brownian motion X

$$X_t = x_0 + \mu t + \sigma W_t, \tag{3.4.12}$$

the BSDE solution is associated to the Cauchy problem on the advection-diffusion equation

$$\begin{cases} \frac{\partial u}{\partial t} + \mu\frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 u}{\partial x^2} + f(t,x,u,\sigma\nabla u) = 0\ ,\ (t,x) \in [0,T)\times\mathbb{R} \\ u(T,x) = g(x),\ x \in \mathbb{R}. \end{cases} \tag{3.4.13}$$

We already mentioned in Chapter 2 that the forward process increments are indeed stationary, independent and normally distributed with density

$$h(x) = \frac{1}{(2\pi\Delta_i)^{\frac{1}{2}}\sigma}\exp\left(-\frac{(x-\mu\Delta_i)^2}{2\sigma^2\Delta_i}\right). \tag{3.4.14}$$

and characteristic function

$$\phi(\nu) = e^{\Delta_i\left(\mathbf{i}\mu\nu - \frac{1}{2}\sigma^2\nu^2\right)}. \tag{3.4.15}$$

In this case, the convolution method is applied with $\psi(\nu) = \phi(\nu)$ when computing the intermediate solutions $\tilde{u}_i$ and with $\psi(\nu) = \mathbf{i}\nu\phi(\nu)$ when computing the approximate gradient $\dot{u}_i$. Though, we approximate the approximate gradient through the quantity $\sigma\nabla u = \sigma\frac{\partial u}{\partial x}$ and hence use $\psi(\nu) = \mathbf{i}\sigma\nu\phi(\nu)$ in our implementation.

The equivalences of convolution representation under the transform of equation (3.1.8) are given by

$$\begin{aligned}\theta(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\eta}(\nu)\psi(\nu)d\nu \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha,\beta}}(\nu)\psi(\nu)d\nu \\ &\quad -\alpha\left[(x+\mu\Delta_i)^2 + \sigma^2\Delta_i\right] - \beta(x+\mu\Delta_i)\end{aligned} \tag{3.4.16}$$

when $\psi(\nu) = \phi(\nu)$ and

$$\begin{aligned}\theta(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\eta}(\nu)\psi(\nu)d\nu \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\widehat{\eta^{\alpha,\beta}}(\nu)\psi(\nu)d\nu \\ &\quad -\sigma\left[2\alpha(x+\mu\Delta_i) + \beta\right]\end{aligned} \tag{3.4.17}$$

when $\psi(\nu) = \mathbf{i}\sigma\nu\phi(\nu)$. The computation of the approximate solution u_i and the approximate gradient $\dot{u}_i$ as well as the BSDE simulation are done as described in the previous sections.

3.4.3 The Euler scheme 2

As stated in Chapter 2, the approximate solution v_i and the approximate gradient $\dot{v}_i$ satisfy

$$v_i(x) = \int_{-\infty}^{\infty} \tilde{v}_{i+1}(x+y)h(y)dy \tag{3.4.18}$$

where

$$\tilde{v}_{i+1}(x) = v_{i+1}(x) + \Delta_i f(t_i, v_{i+1}(x), \dot{v}_i(x)) \ , \tag{3.4.19}$$

$$\dot{v}_i(x) = \int_{-\infty}^{\infty} (y-x)v_{i+1}(x+y)h(y)dy \tag{3.4.20}$$

for $i = 0, 1, ..., n-1$ and $v_n(x) = g(x)$ under the explicit Euler scheme 2.

In this setting, one notices that two successive conditional expectations have to be computed at each time step through these equations. In order to maintain accuracy as described in Section 3.2, a total of $2N$ points have to be discarded on the space grid at each time step[3]. As a comparison, only N points are lost in the alternative grid presented above. Thus, the implementation of Euler scheme 2 is more computationally demanding compared to the Euler scheme 1.

Nonetheless, it is possible to implement the Euler scheme 2 on the alternative grid of Section 3.1. Indeed, a simple algorithm consists in using all values of the approximate gradient computed through equation (3.4.20) in the intermediate solution of equation (3.4.19). Hence, only N points are discarded from the space grid when computing the approximate solution with equation (3.4.18). As a consequence, the Euler scheme 2 will display an additional extrapolation error,

[3] N points are discarded after computing equation (3.4.20) and N other points when computing equation (3.4.18) which includes the solution of (3.4.20).

especially at the boundaries on the truncated domain. The complete algorithm is given in Algorithm 3.2.

Indeed, simple adaptations of Algorithm 3.2 allow us to solve for RBSDEs and/or for arithmetic Brownian motion. These modification were already discussed in the beginning of this section. Also, the simulation of (R)BSDEs follows naturally in the setting of Scheme 2 from the presentation of the numerical solution for Scheme 1.

Algorithm 3.2. *Convolution Method on Alternative Grid (scheme 2)*

1. *Discretize the restricted real space* $[-\frac{N_n l}{2}, \frac{N_n l}{2}]$ *and the restricted Fourier space* $[-\frac{L}{2}, \frac{L}{2}]$ *with* $N_n N$ *space steps so to have the real space nodes* $\{x_{nk}\}_{k=0}^{N_n N}$ *and* $\{\nu_{nk}\}_{k=0}^{N_n N}$
2. *Value* $v_n(x_{nk}) = g(x_{nk})$
3. *For any* i *from* $n-1$ *to* 0
   (a) *Compute* α *and* β *defining the transform of equation (3.1.8), such that*
   $$\theta_{i+1} = (v_{i+1})^{\alpha,\beta} \tag{3.4.21}$$
   and θ_{i+1} *satisfies the boundary conditions of equations (3.1.25) and (3.1.26).*
   (b) *Compute* $\theta_i(x_{i+1,k})$ *through equation (3.1.28) for* $k = 0, 1, ..., N_{i+1}N$ *with*
   $$\psi(\nu) = \mathbf{i}\nu\phi(\nu) \tag{3.4.22}$$
   and retrieve the values $\dot{v}_{ik}$ *as*
   $$\dot{v}_{ik} = \theta_i(x_{i+1,k}) - (2\alpha x_{i+1,k} + \beta)\,. \tag{3.4.23}$$
   (c) *Compute the values* $\tilde{v}_{i+1,k}$ *as*
   $$\tilde{v}_{i+1,k} = v_{i+1,k} + \Delta_i f(t_i, v_{i+1,k}, \dot{v}_{ik}) \tag{3.4.24}$$
   for $k = 0, 1, ..., N_{i+1}N$ *through equation (3.4.19).*
   (d) *Compute* α *and* β *defining the transform of equation (3.1.8), such that*
   $$\theta_{i+1} = (\tilde{v}_{i+1})^{\alpha,\beta} \tag{3.4.25}$$
   and θ_{i+1} *satisfies the boundary conditions of equations (3.1.25) and (3.1.26).*
   (e) *Compute* $\theta_i(x_{ik})$ *through equation (3.1.30) for* $k = 0, 1, ..., N_i N$ *with*
   $$\psi(\nu) = \phi(\nu) \tag{3.4.26}$$
   and retrieve the values v_{ik} *as*
   $$\tilde{v}_{ik} = \theta_i(x_{ik}) - \alpha(x_{i,k}^2 - \Delta_i)\beta x_{i,k}. \tag{3.4.27}$$
   (f) *Update the real space grid with equation (3.1.21) and the Fourier space grid by discretizing the interval* $[-\frac{L}{2}, \frac{L}{2}]$ *with* $N_i N$ *space steps so to have the real space nodes* $\{x_{ik}\}_{k=0}^{N_i N}$ *and* $\{\nu_{ik}\}_{k=0}^{N_i N}$.

3.5 Numerical results

The numerical results on the alternative grid stands as a complement of the results presented in Section 2.6. We intend to demonstrate the absence of extrapolation error and give illustrations of the (space and time) convergence order on the alternative grid. Since the truncation error is demonstrated to be of spectral order, it can be easily set to the order of epsilon machine[4] and, hence, be considered negligible.

[4]We use a double precision arithmetic with epsilon machine $\epsilon = 2.2204 \times 10^{-16}$.

3.5.1 Space and time convergence order

In Appendix D, we develop closed form expressions for the approximate solution of the BSDE with driver

$$f(t, y, z) = ay + bz \tag{3.5.1}$$

and terminal condition

$$g(x) = e^{\varphi x}. \tag{3.5.2}$$

Equations (D.14) and (D.15) give the backward and control process true solutions and equations (D.16) and (D.17) give the approximate solutions after time discretization. It clearly follows from these equations that the time discretization is convergent as expected. We would like to test the convergence of the numerical solutions to the true solutions u and ∇u and also to the approximate solutions u_i and $\dot{u}_i$, $i = 0, 1, ..., n-1$.

We solve the BSDE with terminal time $T = 0.1$ and coefficient parameters $a = 3$, $b = -5$ and $\varphi = 0.15$ on uniform time meshes with $\Delta = \frac{T}{n}$ and space grids with $N_0 = 0$. For a given time mesh with n time steps and a space grid with interval length l and N space steps, we compute two different errors. On one hand, the error of the numerical solutions with respect to the true solutions is computed as

$$\begin{aligned} E_{True} &= \max_{0\le i<n} \max_{0\le k\le NN_i} |u(t_i, x_{ik}) - u_{ik}| \\ &+ \max_{0\le i<n} \max_{0\le k\le NN_i} |\nabla u(t_i, x_{ik}) - \dot{u}_{ik}|. \end{aligned} \tag{3.5.3}$$

On the other hand, the error of the numerical solutions with respect to the approximate solutions is given by

$$\begin{aligned} E_{App} &= \max_{0\le i<n} \max_{0\le k\le NN_i} |u_i(x_{ik}) - u_{ik}| \\ &+ \max_{0\le i<n} \max_{0\le k\le NN_i} |\dot{u}_i(x_{ik}) - \dot{u}_{ik}|. \end{aligned} \tag{3.5.4}$$

We hence consider the maximal absolute error of the numerical solutions with Scheme 1 over the entire grid on the solution u and its gradient ∇u.

The error of the numerical solutions with respect to the approximate solutions (E_{App}) is certainly the most important one since it indicates the accuracy of the convolution method when computing the conditional expectations appearing in the Euler scheme. On the alternative grid, this error term includes the space discretization error and the truncation error as shown in Theorem 3.5. We will use the error E_{App} to analyze the spatial convergence of the method. As to the error of the numerical solutions with respect to the true solutions (E_{True}), it incorporates the underlying time discretization error and can be used to analyze the effect of the time discretization procedure.

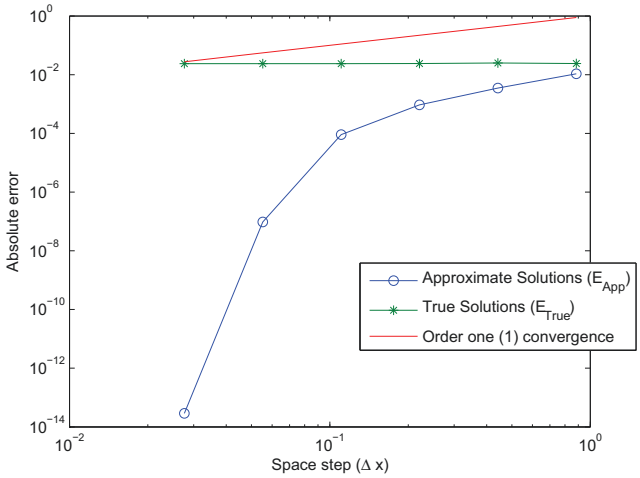
In order to perform a spatial convergence analysis, we set $n = 20$ and we choose l large enough so that the truncation error can be neglected. More precisely, we set

$$l = 25\sqrt{\Delta} \tag{3.5.5}$$

and the truncation error is of the order of $e^{-25^2 C}$ which is expected to be closed to machine error for the constant $C > 0$. The BSDE is then solved for different space discretizations with $N = 2^j$, $j \in \{1, 2, 3, 4, 5, 6\}$. Figure 3.5.1 shows the log-log plot of the maximal errors E_{App} and E_{True} for each of the space discretizations.

A first observation on Figure 3.5.1 is that the error with respect to the true solutions E_{True} does not change throughout the different space discretizations. This indicates that the time discretization error remains constant since we chose a constant time step ($n = 20$) for each of

Figure 3.5.1: Log-log plot for space convergence analysis on alternative grid.



Errors computed using the numerical solutions (Scheme 1) on alternative grids. The time step is kept constant at $\Delta = 0.005$ and the truncation error of the order of epsilon machine for each space discretizations.

the space discretizations. The level of the associated line gives an idea of the time discretization error which is of the order of 10^{-2}.

The error with respect to the approximate solutions E_{App} displays a curve which has a slope of approximately one (1) for coarse space grids indicating an order one (1) convergence rate in space as stated in Theorem 3.5. Nonetheless, the slope of the curve rapidly increases when the space step Δx is reduced so that the error drops to the order of 10^{-14} near the level of machine error. Hence, the convergence order is improved to nearly spectral when the space grid is refined. This is consequence of the smoothness of the BSDE coefficients f and g and the use of Fourier interpolation. Indeed, for infinitely differentiable coefficients f and g, the accuracy of the Fourier interpolation underlying the convolution method[5] is actually greater than any polynomial order. The assumptions of twice differentiability on the coefficients are simply minimal requirements leading to the first order convergence rate in space.

Since the convolution method uses a time discretization of the BSDE, it may be interesting to investigate the effect of the underlying time discretization. For the time convergence analysis, we set the length l to be

$$l = 10\sqrt{\Delta}, \tag{3.5.6}$$

for each time discretization with time step Δ, so that the truncation error is of the order of $e^{-10^2 C}$ for a constant $C > 0$. Then we fix the maximal space step at 0.0025. This is done by choosing N such that

$$N = \left\lceil \frac{l}{0.0025} \right\rceil + \left\lceil \frac{l}{0.0025} \right\rceil \bmod 2 \tag{3.5.7}$$

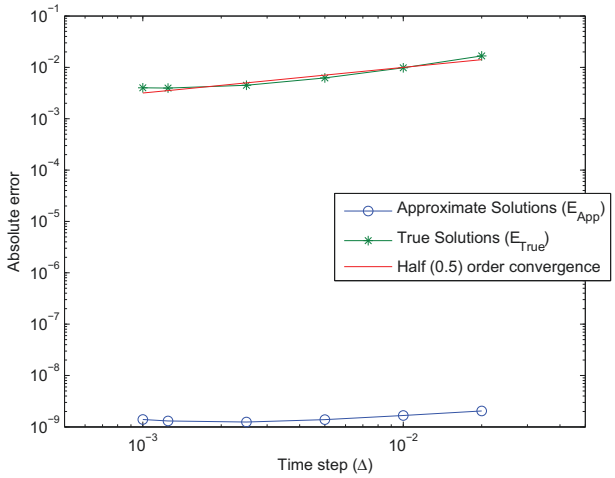
where, for $x \in \mathbb{R}$, $\lceil x \rceil = \min\{m \in \mathbb{Z} : m \geq x\}$ is the ceiling function and, for $m \in \mathbb{Z}$,

$$m \bmod 2 = \begin{cases} 1 \text{ if } m \text{ is odd} \\ 0 \text{ if } m \text{ is even} \end{cases} \tag{3.5.8}$$

is the modulo 2 operator. This procedure ensures that the chosen value of N is even and the

[5] See the proof of Theorem 3.3.

Figure 3.5.2: Log-log plot for time convergence analysis on alternative grid.



Errors computed using the numerical solutions (Scheme 1) on alternative grids. The space step is kept constant at about $\Delta x = 0.0025$ for each time discratization. Also, the length l is set to be $l = 10\sqrt{\Delta}$ so that the truncation errors are of the same order for each time discretization.

alternative grid is constructable with a space step such that

$$\Delta x \leq 0.0025. \tag{3.5.9}$$

Since we have a maximal space step, we are forced to fix a maximal number of time steps such that

$$n < \frac{T\sqrt{2\pi}}{0.0025} \approx 100.2651 \tag{3.5.10}$$

in order to satisfy the condition of Lemma 3.6 and produce a stable algorithm. For this reason, we solve the BSDE with $n \in \{5, 10, 20, 40, 80, 100\}$ leading to different time discretizations. The errors computed with those discretizations are displayed in the log-log plot of Figure 3.5.3.

Since we set a (almost) constant space step and a negligible truncation error, the error E_{App} is (almost) constant for each time discretization. It also gives an indication of the impact of the space discretization which is of the order of 10^{-9}. As to the error with respect to the true solutions E_{True}, it describes a line with slope $\frac{1}{2}$ in the log-log plot. Hence, we may conclude that the time discretization (absolute) error is of half $(\frac{1}{2})$ order[6].

This result is consistent with the order of convergence of BSDE time discretization in Zhang [124] or Bouchard and Touzi [20] and, hence, with the statements of Theorem 3.9 and Corollary 3.10. Indeed, if the time discretization (absolute) error with respect to the true solutions is of half $(\frac{1}{2})$ order on the alternative grid, then we can expect the quadratic errors described in Theorem 3.9 and Corollary 3.10 to be of first order in time.

3.5.2 Application to option pricing

As in Chapter 2, we use option pricing to test the convolution method performance in a non-smooth setting. The market model considered here is identical to the one introduced in Section 2.6.2 with the difference that we price put options.

The risky asset $\{S_t\}_{t\in[0,T]}$ price is given by

$$S_t = e^{X_t} \tag{3.5.11}$$

[6] At least in this case.

Table 3.5.1: Relative errors (in percentage) on the alterntive grid for European put option prices

	K (Strike)	n=50	n=100	n=200	n=500
Convolution (Scheme 1)	110	0.0149	0.0746	0.0335	0.0083
	100	0.6537	0.3269	0.1639	0.0660
	90	1.0993	0.0635	0.2549	0.0416
Convolution (Scheme 2)	110	0.1102	0.0268	0.0095	0.0013
	100	0.7450	0.3730	0.1871	0.0753
	90	1.00745	0.0525	0.2494	0.0398
Trinomial tree (Scheme 1)	110	0.0830	0.0646	0.0200	0.0087
	100	0.1891	0.0990	0.0501	0.0374
	90	0.0388	0.3072	0.1035	0.0324

The convolution method on the alternative grid is performed with a uniform time mesh and parameters $N_0 = 5$, $l = 15\sqrt{\Delta}\sigma$ and $N = 2^3$. All methods use a uniform time mesh.

where the return process $\{X_t\}_{t\in[0,T]}$ satisfies

$$X_t = X_0 + \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t \tag{3.5.12}$$

with initial price $S_0 = e^{X_0} = 100$, volatility $\sigma = 0.2$ and expected return $\mu = 0.05$. In addition, the market offers a lending rate of $r = 0.01$. Under these conditions, a European put option with strike price K solves the BSDE with linear driver

$$f(t, y, z) = -ry - \left(\frac{\mu - r}{\sigma}\right)z \tag{3.5.13}$$

and the non-smooth terminal function

$$g(x) = (K - e^x)^+ . \tag{3.5.14}$$

The maturity of all considered options is $T = 1$.

At strike prices $K = 110$, 100 and 90, the Black-Scholes formula returns put option prices of 13.5156, 7.4383 and 3.2974 respectively and option deltas of -0.2655, -0.4404 and -0.6468 respectively. Table 3.5.1 and Table 3.5.2 give the relative errors for various time discretizations produced by the estimates from the convolution methods on the alternative grid. For comparison purposes, these tables also contain the relative errors of the estimates from the trinomial method as presented in Appendix C. Since the numerical results in Section 2.6.2 indicate satisfactory precision for relatively coarse time grids, we limit the number of time steps to $n = 500$. As expected, the relative error decreases when the number of time steps increases so that the precision of the convolution method is improved on finer time grids when using an alternative space grid.

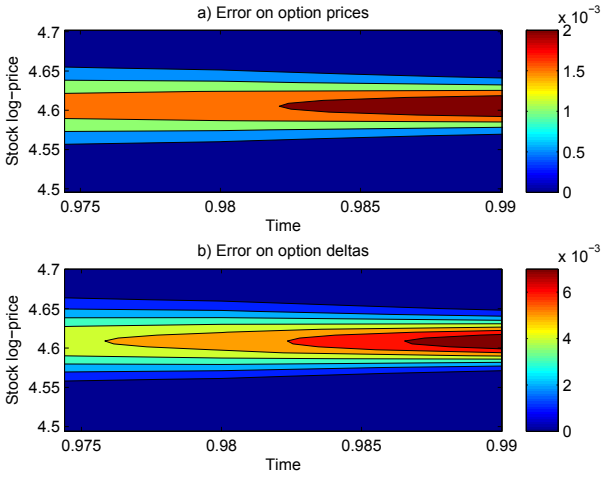
Overall, the convolution and trinomial methods give similar results since they produce similar relative errors particularly for fine time discretizations. Nonetheless, the error on at-the-money option prices are slightly higher than those on out-of-the money and in-the-money option prices under the convolution method. As to option deltas, out-of-the money options have higher errors than in-the-money options. Those observations are consequences of the non-smoothness of the terminal condition g and the way errors diffuse through the approximation solution and the approximate gradient under the convolution method on the alternative grid.

Table 3.5.2: Relative errors (in percentage) on the alternative grid for European put option deltas.

	K (Strike)	n=50	n=100	n=200	n=500
Convolution (Scheme 1)	110	0.1039	0.1063	0.0506	0.0172
	100	0.2244	0.1113	0.0553	0.0219
	90	0.7750	0.2329	01808	0.0332
Convolution (Scheme 2)	110	0.1532	0.1311	0.0633	0.0226
	100	0.3849	0.1924	0.0965	0.0389
	90	0.9876	0.3503	0.2413	0.0583
Trinomial tree (Scheme 1)	110	0.1865	0.0243	0.0424	0.0132
	100	0.2657	0.1264	0.0596	0.0240
	90	0.4370	0.3030	0.0680	0.0385

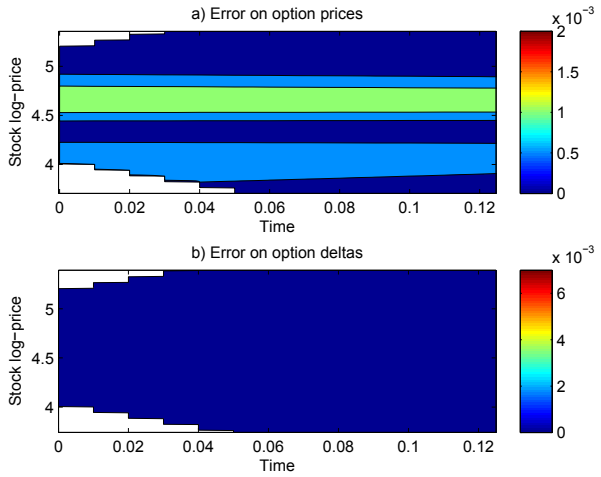
The convolution method on the alternative grid is performed with parameters $N_0 = 5$, $l = 15\sqrt{\Delta}\sigma$ and $N = 2^3$. All methods use a uniform time mesh.

Figure 3.5.3: Absolute errors at the European put option maturity.



Errors were computed using the numerical solutions (Scheme 1) on the alternative grid with parameters $l = 6\sqrt{\Delta}\sigma$, $N_0 = 10$ and $N = 2^5$. The uniform time mesh has $n = 100$ time steps.

Figure 3.5.4: Absolute errors at the European put option issuance.



Errors were computed using the numerical solutions (Scheme 1) on the alternative grid with parameters $l = 6\sqrt{\Delta}\sigma$, $N_0 = 10$ and $N = 2^5$. The uniform time mesh has $n = 100$ time steps.

As already mentioned, the convolution method uses Fourier series interpolation which has a particular behavior when applied to non-smooth functions. It is well-known that trigonometric interpolations display high oscillations around the discontinuities of the interpolated function. Those high oscillations are known as the Gibbs phenomenon and can be observed in the convolution method on the alternative grid. Figure 3.5.3 illustrates the Gibbs phenomenon in option prices and option deltas when the (at-the-money) option is close to maturity. If the oscillations seem to be higher for option deltas, they are more persistent for option prices. Indeed, oscillations disappear in option deltas close to the option issuance as shown in Figure 3.5.4. Also, Figure 3.5.4 indicates that the error induced by the Gibbs phenomenon diffuse along the point of discontinuity through time steps since the errors in option prices at the option issuance are clearly higher when the option is at-the-money.

Of course, the presence of the Gibbs phenomenon is independent of the grid used under the convolution method. Hence, this analysis also holds for the uniform space grid of Chapter 2. However, the presence of the extrapolation error on a uniform grid makes the analysis more complicated since this error globally dominates the space discretization error which contains the error induced by the Gibbs phenomenon.

We conclude this section with the numerical resolution of RBSDEs using the convolution method through American put option pricing. The barrier in this case is given as usual by

$$B(t,x) = g(x),\ (t,x) \in [0,T] \times \mathbb{R}. \tag{3.5.15}$$

Table 3.5.3 shows at-the-money American put option price and delta estimates from the convolution and the trinomial methods. The estimates from both methods remain similar which is an indication of the convolution method accuracy for RBSDEs on the alternative grid. Similarly, the regularity of the price and delta surfaces of Figure 3.5.5 gives an idea of the stability of the numerical solutions for RBSDEs from the convolution method. Finally, this regularity and accuracy allows for a reliable path simulation for (R)BSDEs as illustrated in Figure 3.5.6.

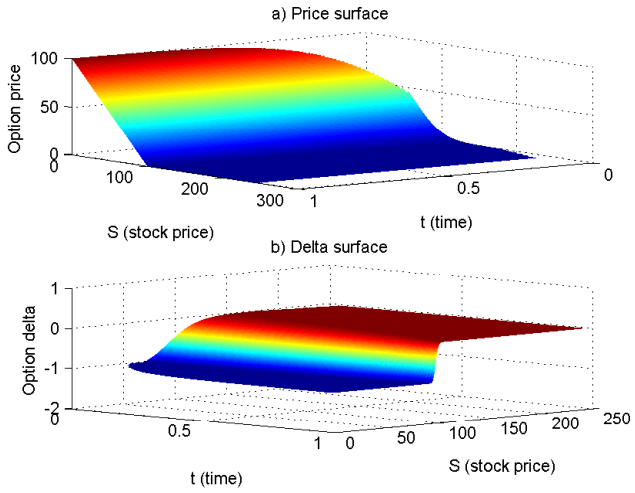
The implementation of the convolution method on the alternative grid produces a local discretization error exempt of truncation errors. The BSDE numerical solution was formally defined and a global error analysis was conducted. The numerical results presented in this sections also

Table 3.5.3: American put option price and delta estimates

	K (Strike)	Price	Delta
Convolution (Scheme 1)	110	13.6860	-0.6402
	100	7.5115	-0.4466
	90	3.3234	-0.2518
Convolution (Scheme 2)	110	13.6855	-0.6402
	100	7.5112	-0.4465
	90	3.3235	-0.2517
Trinomial tree (Scheme 1)	110	13.6860	-0.6402
	100	7.5135	-0.4465
	90	3.3239	-0.2518

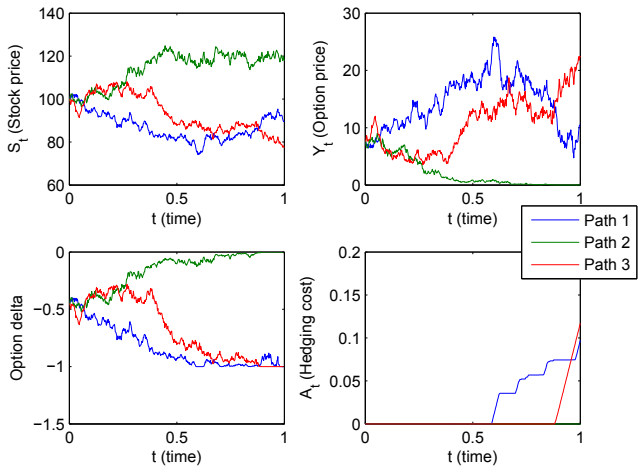
Estimates were computed using the numerical solutions (Scheme 1) on the alternative grid with parameters $l = 15\sqrt{\Delta}\sigma$, $N_0 = 5$ and $N = 2^3$. The uniform time mesh has $n = 1000$ time steps for all methods.

Figure 3.5.5: At-the-money American put option price and delta surfaces.



The surfaces are obtained using the numerical solutions (Scheme 1) on the alternative grid with parameters $l = 6\sqrt{\Delta}\sigma$, $N_0 = 10$ and $N = 2^5$. The uniform time mesh has $n = 100$ time steps.

Figure 3.5.6: Path simulation for the at-the-money American put option.



The paths are interpolated from the numerical solutions (Scheme 1) on the alternative grid with parameters $l = 6\sqrt{\Delta}\sigma$, $N_0 = 10$ and $N = 2^5$. The uniform time mesh has $n = 100$ time steps.

confirm the theoretical results developed in the error analysis. However, the convolution is limited to the BSDE case. Hence, the following section extends the results to the more general FBSDE framework.

Chapter 4

A Fourier interpolation method for FBSDEs

An important and interesting extension of the convolution method consists in applying it to FBSDEs. Unfortunately, the convolution representation of our approximate solutions, as developed and used in Chapters 2 and 3, are unavailable in the FBSDE case. The reason for this inconvenience is that the forward process in a general FBSDE setting, even though a Markovian process, does not have independent increments. Hence the distribution function of the forward process increment does not depend exclusively on the increment, but also on the process's last position.

In this chapter, we shall use some of the ideas introduced in the previous chapters in order to build a numerical method for FBSDEs. The main purpose is to obtain a numerical solution on a uniform space grid as flexible as the one in Chapter 3 and resulting in a convergent scheme. The majority of space discretization and PDE methods for FBSDEs fail in producing a uniform space grid which makes the implementation of those methods quite challenging. Milstein and Tretyakov [89, 90] and Delarue and Menozzi [35, 36] are notable examples. The reason for the extend usage of non-uniform space grids stems from the non-stationarity of the forward process[1]. Nonetheless, a uniform space grid seems to be easier to handle and more suitable for simulation.

4.1 Preliminaries

As already indicated, we work on the complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbf{P})$ where the filtration $\mathbb{F} = \{\mathcal{F}_t : t \in [0,T]\}$ is generated by a d-dimensional Brownian motion $\{W_t\}_{t\in[0,T]}$. The general FBSDE for which we seek a numerical solution is a system of the form

$$\begin{cases} dX_t = a(t, X_t)dt + \sigma(t, X_t)dW_t \\ -dY_t = f(t, X_t, Y_t, Z_t)dt - Z_t^* dW_t \\ X_0 = x_0 \ , \ Y_T = \xi \end{cases} \tag{4.1.1}$$

where the forward drift $a : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, the forward volatility $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, the driver $f : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ are deterministic functions. The initial condition $x_0 \in \mathbb{R}^d$ and the terminal condition takes the Markovian form $\xi = g(X_T)$ where $g : \mathbb{R}^d \to \mathbb{R}$.

We assume the usual Lipschitz and growth conditions on decoupled FBSDE coefficients. The following Assumption gives the details.

[1] In the weak sense.

Assumption 4.1.1. *There exist positive constants K_1, K_2 K_3, and K_4 such that the coefficients of the FBSDE (4.1.1) satisfy*

$$\begin{aligned} |a(t,x_1)-a(t,x_2)| &\leq K_1\,|x_1-x_2| && (4.1.2)\\ \|\sigma(t,x_1)-\sigma(t,x_2)\|_2 &\leq K_1\,|x_1-x_2| && (4.1.3)\\ |a(t,x)|+\|\sigma(t,x)\|_2 &\leq K_2 && (4.1.4)\\ |f(t,x_1,y,z)-f(t,x_2,y,z)| &\leq K_1\,|x_1-x_2| && (4.1.5)\\ |f(t,x,y_1,z_1)-f(t,x,y_2,z_2)| &\leq K_1\,(|y_1-y_2|+|z_1-z_2|) && (4.1.6)\\ |f(t,x,y,z)| &\leq K_3(1+|x|+|y|+|z|) && (4.1.7) \end{aligned}$$

for any $t\in[0,T]$, $x,x_1,x_2\in\mathbb{R}^d$, $y,y_1,y_2\in\mathbb{R}$, $z,z_1,z_2\in\mathbb{R}^d$.

Moreover $\sigma^2:=\sigma\sigma^*$ is (uniformly) invertible, continuous and bounded

$$\left\|(\sigma^2(t,x))^{-1}\right\|_2\leq K_4 \tag{4.1.8}$$

for any $t\in[0,T]$, $x\in\mathbb{R}^d$.

In addition, the terminal value is square integrable

$$\|\xi\|_{L^2}^2:=\boldsymbol{E}\left[|g(X_T)|^2\right]<\infty. \tag{4.1.9}$$

A solution of the system of stochastic differential equations (4.1.1) is a triple of processes (X,Y,Z) where the forward process $X\in L_S^2(\mathbb{R}^d)$ is adapted and square integrable. In addition, the backward process $Y\in L_S^2(\mathbb{R})$ and the control process $Z\in L_I^2(\mathbb{R}^d)$ are also adapted and square integrable.

The problem of well-posedness for the systems of stochastic differential equations (4.1.1) has been widely studied. The existence and uniqueness result of the forward process X is established through SDE theory. Pardoux and Peng [96] proved the well-posedness of the backward SDE.

As indicated by Pardoux and Peng [97], the FBSDE of equation (4.1.1) is associated with the following Cauchy problem on a quasilinear (parabolic) PDE

$$\begin{cases} \frac{\partial u}{\partial t}+\mathcal{L}u+f(t,x,u,\sigma^*(t,x)\nabla u)=0,\ (t,x)\in[0,T)\times\mathbb{R}^d\\ u(T,x)=g(x),\ x\in\mathbb{R}^d \end{cases} \tag{4.1.10}$$

where

$$\mathcal{L}u = \sum_{i=1}^{d}a_i(t,x)\frac{\partial u}{\partial x_i}+\frac{1}{2}\sum_{i,j=1}^{d}\sigma_{ij}^2(t,x)\frac{\partial^2 u}{\partial x_i\partial x_j}. \tag{4.1.11}$$

The FBSDE solution can hence be expressed in terms of the PDE solution u as

$$\begin{aligned} X_t &= x_0+\int_0^t a(s,X_s)ds+\int_0^t\sigma(s,X_s)dW_s && (4.1.12)\\ Y_t &= u(t,X_t) && (4.1.13)\\ Z_t &= \sigma^*(t,X_t)\nabla u(t,X_t). && (4.1.14) \end{aligned}$$

through the *four step scheme* of Ma, Protter and Yong [79].

4.2 Numerical implementation

We develop the numerical implementation of the Fourier interpolation method method. In this section, the basics on the time and space discretizations are given. The presentation gives Fourier representation of numerical solution in the general multidimensional case. The numerical implementation however is restricted to the one-dimensional case.

4.2.1 Time discretization and Fourier representation

As usual, the starting point is the explicit Euler scheme 1 applied to the FBSDE of equation (4.1.1). On the time mesh $\pi = \{t_0 = 0 < t_1 < ... < t_n = T\}$ with time steps

$$\Delta_i = t_{i+1} - t_i,\ i = 0, 1, ..., n-1, \tag{4.2.1}$$

the time discretization takes the form

$$\begin{cases} X_0^\pi = x_0 \\ X_{t_{i+1}}^\pi = X_{t_i}^\pi + a(t_i, X_{t_i}^\pi)\Delta_i + \sigma(t_i, X_{t_i}^\pi)\Delta W_i \\ Z_{t_n}^\pi = 0,\ Y_{t_n}^\pi = \xi^\pi \\ Z_{t_i}^\pi = \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^\pi \Delta W_i | \mathcal{F}_{t_i}\right] \\ Y_{t_i}^\pi = \mathbf{E}\left[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\right] + f(t_i, X_{t_i}^\pi, \mathbf{E}\left[Y_{t_{i+1}}^\pi | \mathcal{F}_{t_i}\right], Z_{t_i}^\pi)\Delta_i \end{cases} \tag{4.2.2}$$

where $\Delta W_i = W_{t_{i+1}} - W_{t_i}$. First note that

$$\begin{aligned} \sigma(t_i, X_{t_i}^\pi)Z_{t_i}^\pi &= \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^\pi \sigma(t_i, X_{t_i}^\pi)\Delta W_i | \mathcal{F}_{t_i}\right] \\ &= \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^\pi (\Delta X_i^\pi - \Delta_i a(t_i, X_{t_i}^\pi)) | \mathcal{F}_{t_i}\right] \end{aligned} \tag{4.2.3}$$

with $\Delta X_i^\pi = X_{t_{i+1}}^\pi - X_{t_i}^\pi$. Hence, the approximate solution u_i at mesh time t_i can be written as

$$u_i(x) = \tilde{u}_i(x) + \Delta_i f(t_i, x, \tilde{u}_i(x), \dot{u}_i(x)) \tag{4.2.4}$$

where the intermediate solution $\tilde{u}_i$ and the approximate gradient $\dot{u}_i$ at mesh time t_i satisfy

$$\begin{aligned} \tilde{u}_i(x) &= \mathbf{E}\left[Y_{t_{i+1}}^\pi | X_{t_i}^\pi = x\right] \\ &= \int_{\mathbb{R}^d} u_{i+1}(x+y)h_i(y|x)dy \end{aligned} \tag{4.2.5}$$

$$\begin{aligned} \sigma(t_i, x)\dot{u}_i(x) &= \frac{1}{\Delta_i}\mathbf{E}\left[Y_{t_{i+1}}^\pi \sigma(t_i, X_{t_i}^\pi)\Delta W_i | X_{t_i}^\pi = x\right] \\ &= \frac{1}{\Delta_i}\int_{\mathbb{R}^d} (y - \Delta_i a(t_i, x))u_{i+1}(x+y)h_i(y|x)dy \end{aligned} \tag{4.2.6}$$

for $i = 0, 1, ..., n-1$ and $u_n(x) = g(x)$. Moreover, the function h_i is the conditional density of the discrete forward increment ΔX_i^π given an initial position of $X_{t_i}^\pi = x$ at mesh time t_i. From the Euler scheme in equation (4.2.2), h_i is the density of a Gaussian random variable with mean $\Delta_i a(t_i, x)$ and variance-covariance matrix $\Delta_i \sigma^2(t_i, x)$.

The density is explicitly given by

$$h_i(y|x) = (2\pi)^{-\frac{d}{2}} \left\|\Delta_i \sigma^2(t_i, x)\right\|_2^{-\frac{1}{2}} \exp\left(-\frac{1}{2\Delta_i}\mathbf{y}^*(\sigma^2(t_i, x))^{-1}\mathbf{y}\right) \tag{4.2.7}$$

where $\mathbf{y} = y - \Delta_i a(t_i, x)$ with characteristic function

$$\phi_i(\nu, x) = e^{\mathbf{i}\Delta_i \nu^* a(t_i, x) - \frac{1}{2}\Delta_i \nu^* \sigma^2(t_i, x)\nu}. \tag{4.2.8}$$

The density function h_i and the characteristic function ϕ_i satisfy the relation

$$h_i(y|x) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-\mathbf{i}\nu^* y}\phi_i(\nu, x)d\nu. \tag{4.2.9}$$

Consequently, they also satisfy

$$yh_i(y|x) = -\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{-\mathbf{i}\nu^* y}\mathbf{i}\nabla_\nu \phi_i(\nu, x)d\nu$$

$$
\begin{aligned}
&= \Delta_i a(t_i,x) h_i(y|x) \\
&\quad + \frac{\Delta_i \sigma^2(t_i,x)}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-\mathbf{i}\nu^* y} \mathbf{i}\nu \phi_i(\nu,x) d\nu
\end{aligned} \tag{4.2.10}
$$

where the first equality holds by Proposition E.4.

Hence, assuming u_{i+1} is (Lebesque) integrable, equation (4.2.5) leads to

$$
\begin{aligned}
\tilde{u}_i(x) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} u_{i+1}(x+y) \int_{\mathbb{R}^d} e^{-\mathbf{i}\nu^* y} \phi_i(\nu,x) d\nu dy \\
&\quad \text{(using the relation of equation (4.2.9))}, \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-\mathbf{i}\nu^* y} u_{i+1}(x+y) \phi_i(\nu,x) dy d\nu \\
&= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{\mathbf{i}\nu^* x} \mathfrak{F}[u_{i+1}](\nu) \phi_i(\nu,x) d\nu \\
&= \mathfrak{F}^{-1}[\mathfrak{F}[u_{i+1}](\nu) \phi_i(\nu,x)](x).
\end{aligned} \tag{4.2.11}
$$

Since $\sigma(t,x)$ is a full rank matrix, equation (4.2.6) is equivalent to

$$
\begin{aligned}
\dot{u}_i(x) &= \frac{\sigma^*(t_i,x)}{(2\pi)^d} \int_{\mathbb{R}^d} u_{i+1}(x+y) \int_{\mathbb{R}^d} e^{-\mathbf{i}\nu^* y} \mathbf{i}\nu \phi_i(\nu,x) d\nu dy \\
&\quad \text{(using the relation of equation (4.2.10))}, \\
&= \frac{\sigma^*(t_i,x)}{(2\pi)^d} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} e^{-\mathbf{i}\nu^* y} u_{i+1}(x+y) \mathbf{i}\nu \phi_i(\nu,x) dy d\nu \\
&= \frac{\sigma^*(t_i,x)}{(2\pi)^d} \int_{\mathbb{R}^d} e^{\mathbf{i}\nu^* x} \mathfrak{F}[u_{i+1}](\nu) \mathbf{i}\nu \phi_i(\nu,x) d\nu \\
&= \sigma^*(t_i,x) \mathfrak{F}^{-1}[\mathfrak{F}[u_{i+1}](\nu) \mathbf{i}\nu \phi_i(\nu,x)](x).
\end{aligned} \tag{4.2.12}
$$

We use the Fourier representations in equations (4.2.4), (4.2.11) and (4.2.12) in the implementation of the method under the explicit Euler scheme 1. Equation (4.2.4) can be replaced by

$$
u_i(x) = \tilde{u}_i(x) + \Delta_i f(t_i, x, u_i(x), \dot{u}_i(x)) \tag{4.2.13}
$$

under the implicit Euler scheme. In this case, the condition on the time discretization of equation (2.2.13) has to be satisfied.

If the forward process X admits the conditional characteristic function

$$
\phi_t(\nu,x,\tau) = \mathbf{E}\left[e^{\mathbf{i}\nu^*(X_{t+\tau} - X_t)} | X_t = x \right], \tag{4.2.14}
$$

then the conditional characteristic function may be used in the convolution method. The procedure leads to the expressions

$$
\tilde{u}_i(x) = \mathfrak{F}^{-1}[\mathfrak{F}[u_{i+1}](\nu) \phi_{t_i}(\nu,x,\Delta_i)](x) \tag{4.2.15}
$$

$$
\dot{u}_i(x) = \sigma^*(t_i,x) \mathfrak{F}^{-1}[\mathfrak{F}[u_{i+1}](\nu) \mathbf{i}\nu \phi_{t_i}(\nu,x,\Delta_i)](x) \tag{4.2.16}
$$

for the intermediate solution and the approximation gradient in place of equations (4.2.11) and (4.2.12). By using the conditional characteristic function, we are considering the true distribution the forward increment $\Delta X_i = X_{t_{i+1}} - X_{t_i}$ instead of the Gaussian distribution of its Euler discretization ΔX_i^π in the conditional expectations.

4.2.2 Space discretization

Before discretizing the Fourier integrals in the one-dimensional case $d = 1$, we first consider the behavior of the relations in equations (4.2.11) and (4.2.12) under the alternative transform. The next theorem gives the result and its proof is essentially similar to the proof of Theorem 3.2.

Theorem 4.1. *Let $u_{i+1}^{\alpha,\beta}$ be the alternative transform defined in equation (3.1.8) of the approximate solution u_{i+1}. Then the intermediate solution $\tilde{u}_i$ and the approximate gradient $\dot{u}_i$ in equations (4.2.5) and (4.2.6) satisfy*

$$\begin{aligned}\tilde{u}_i(x) &= \mathfrak{F}^{-1}[\mathfrak{F}[u_{i+1}^{\alpha,\beta}](\nu)\phi(\nu,x)](x) \\ &\quad -\alpha[(x+\Delta_i a(t_i,x))^2+\Delta_i\sigma^2(t_i,x)] \\ &\quad -\beta(x+\Delta_i a(t_i,x)) \qquad (4.2.17)\\ \dot{u}_i(x) &= \sigma(t_i,x)\mathfrak{F}^{-1}[\mathfrak{F}[u_{i+1}^{\alpha,\beta}](\nu)\mathbf{i}\nu\phi(\nu,x)](x) \\ &\quad -\sigma(t_i,x)[2\alpha(x+\Delta_i a(t_i,x))+\beta]. \qquad (4.2.18)\end{aligned}$$

As usual, we select the parameters α and β such that the function $u_{i+1}^{\alpha,\beta}$ and its derivative values are equal at the boundaries of a given interval through the method of Lemma 3.1. Consequently, it suffices to compute the values of a generic function $\theta_i : \mathbb{R} \to \mathbb{R}$ at each time step t_i verifying

$$\theta_i(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{\mathbf{i}\nu x}\hat{\theta}_{i+1}(\nu)\psi(\nu,x)d\nu. \qquad (4.2.19)$$

with $\psi : \mathbb{R}^2 \to \mathbb{C}$ and $\theta_{i+1} : \mathbb{R} \to \mathbb{R}$. We may assume that θ_{i+1} satisfies the value and derivative conditions at the boundaries of a given interval.

The space discretization is performed with an alternative grid as described in Chapter 3. In this case, the alternative grid discretizes the forward process X and not the Brownian motion as in the previous chapter. The grid is defined by the increment interval length $l > 0$, its (even) number of space steps $N > 0$ and the initial number of intervals $N_0 > 0$ at mesh time t_0. The grid can be easily built such that the initial forward value x_0 is a grid point at the initial time step t_0 i.e

$$x_0 \in \{x_{0,k} : k = 0, 1, ..., NN_0\}. \qquad (4.2.20)$$

We assume that the grid is centered at the initial value of the forward process $X_0 = x_0$. This requirement simplifies the error analysis even though a shifted grid does not alter the convergence results. Moreover, a shifted grid may be useful to take into account the presence of a drift in the forward process X or a known constraint on the support of its actual transition density.

Hence, assuming that

$$\begin{aligned}\theta_{i+1}(x_{i+1,0}) &= \theta_{i+1}\left(x_{i+1,NN_{i+1}}\right) \qquad (4.2.21)\\ \frac{\partial\theta_{i+1}}{\partial x}(x_{i+1,0}) &= \frac{\partial\theta_{i+1}}{\partial x}\left(x_{i+1,NN_{i+1}}\right). \qquad (4.2.22)\end{aligned}$$

and following the discretization steps of Section 1.2 or Section 3.1, we have

$$\theta_i(x_{ik}) \approx (-1)^{k+\frac{N}{2}}\mathfrak{D}^{-1}\left[\{\psi(\nu_{i+1,j},x_{ik})\mathbb{D}[\theta_{i+1}]_j\}_{j=0}^{N_{i+1}N-1}\right]_{k+\frac{N}{2}} \qquad (4.2.23)$$

for $k = 0, 1, ..., N_iN$. In addition,

$$\mathbb{D}[\theta_{i+1}]_j = \mathfrak{D}\left[\{(-1)^s\tilde{w}_s\theta_{i+1}(x_{i+1,s})\}_{s=0}^{N_{i+1}N-1}\right]_j. \qquad (4.2.24)$$

where the weights $\{\tilde{w}_j\}_{j=0}^{N_{i+1}N-1}$ are as in equation (2.3.10).

In equation (4.2.23), the generic function ψ depends on the space node x_{ik}. If the relation generalizes for all space nodes x_{ik}, $k = 0, 1, ..., N_iN$, the function values $\theta_i(x_{ik})$, $k = 0, 1, ..., N_iN$, can not be computed with a single direct FFT procedure. Instead, a separate FFT procedure using the values of the generic function ψ at x_{ik} is needed to compute the function value $\theta_i(x_{ik})$. Nonetheless the vector-matrix representation of the FFT procedure in equation (4.2.23) allows the computation of all function values $\theta_i(x_{ik})$ with a matrix multiplication.

In the vector-matrix representation, equation (4.2.23) write

$$\theta_i(x_{ik}) = (-1)^{k+\frac{N}{2}} \hat{F}_{k+\frac{N}{2}} \Psi(x_{ik}) \mathbb{D}[\theta_{i+1}] \tag{4.2.25}$$

where $\hat{F}_{k+\frac{N}{2}}$ is the $(k+\frac{N}{2})$th row of the $N_{i+1}N$ dimension inverse FFT matrix $\hat{F}$ and $\Psi(x_{ik})$ is the $N_{i+1}N$ dimension diagonal matrix built with the values $\{\psi(\nu_{i+1,j}, x_{ik})\}_{j=0}^{N_{i+1}N-1}$. Let $\Theta^{(i)}$ be the N_iN dimension vector of the function values $\theta_i(x_{ik})$ such that

$$\Theta^{(i)}_{1+k} = \theta_i(x_{ik}) \tag{4.2.26}$$

for $k = 0, 1, ..., N_iN$. The matrix representation gives

$$\Theta^{(i)} = \hat{\Psi}^{(i)} \mathbb{D}[\theta_{i+1}] \tag{4.2.27}$$

where $\hat{\Psi}^{(i)}$ is the $(N_iN+1) \times N_{i+1}N$ matrix such that

$$\hat{\Psi}^{(i)}_{1+k,1+j} = (-1)^{k+\frac{N}{2}} \bar{\omega}_i^{j(k+\frac{N}{2})} \psi(\nu_{i+1,j}, x_{ik}) \tag{4.2.28}$$

with $\bar{\omega}_i = e^{\mathbf{i}2\pi(N_{i+1}N)^{-1}}$, $k = 0, 1, ..., N_iN$ and $j = 0, 1, ..., N_{i+1}N - 1$.

The algorithm for this Fourier interpolation method for (decoupled) FBSDEs is essentially similar to Algorithm 3.1. One just has to adapt the methods of computation by using the equations introduced in this section.

4.3 Error analysis

The error analysis for the Fourier interpolation method follows the ideas of Chapter 3. As in Section 3.2, Fourier interpolation is used to derive a local discretization error which naturally leads to a global discretization error under a stability condition. From the global error bound, the simulation error is obtained using the time discretization error of Zhang [124] or Bouchard and Touzi [20] as in Section 3.3. We focus the analysis as usual on the explicit Euler scheme 1. The next theorem gives a bound for the local discretization error defined in equation (3.2.1).

Theorem 4.2. *Suppose that the driver $f \in \mathcal{C}^{1,2,2}$ and the terminal condition $g \in \mathcal{C}^2$ and Assumption 4.1.1 is satisfied. Then the Fourier interpolation method yields a discretization error of the form*

$$E_{ik} = \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-K|\Delta_i|^{-1}l^2}\right) \tag{4.3.1}$$

for some constant $K > 0$ on the alternative grid and under the trapezoidal quadrature rule.

Proof. Following the steps of Theorem 3.3, the truncation error relies on the expression

$$\begin{aligned}\int_{|y|>\frac{l}{2}} h(y|x_{i,k})dy &= \mathbf{P}\left[|\Delta X_i^\pi| > \frac{l}{2} \left| X_{t_i}^\pi = x_{i,k}\right.\right] \\ &= \mathbf{P}\left[\left(\frac{\Delta X_i^\pi}{\sigma(t_i, x_{i,k})\sqrt{\Delta_i}}\right)^2 > \frac{l^2}{4\sigma^2(t_i, x_{i,k})\Delta_i} \left| X_{t_i}^\pi = x_{i,k}\right.\right].\end{aligned}$$

Let $\zeta = \frac{l^2}{4\sigma^2(t_i,x_{i,k})\Delta_i}$ and knowing that the random variable $\left(\frac{\Delta X_i^\pi}{\sigma(t_i,x_{i,k})\sqrt{\Delta_i}}\right)^2$ follows a non-central chi-square distribution with one (1) degree of freedom and non-centrality parameter $\lambda = \left(\frac{a(t_i,x_{i,k})}{\sigma(t_i,x_{i,k})}\right)^2 \Delta_i$, we have

$$\begin{aligned}\int_{|y|>\frac{l}{2}} h(y|x_{i,k})dy &\leq \inf_{0<s<\frac{1}{2}} (1-2s)^{-\frac{1}{2}} e^{-s\zeta + \frac{\lambda s}{1-2s}} \\ &\qquad \text{(by Chernoff's inequality)}\end{aligned}$$

$$
\begin{aligned}
&< Ce^{-s\zeta} \\
&\quad \text{(since } \lambda \text{ is bounded by Assumption 4.1.1)}, \\
&< Ce^{-\frac{s}{4}(\sigma^2(t_i,x_{i,k})\Delta_i)^{-1}l^2} \\
&< Ce^{-\frac{sK}{4}(\Delta_i)^{-1}l^2}
\end{aligned}
$$

for some constants $C, K > 0$ and $s \in (0, \frac{1}{2})$ where the last inequality holds since σ^{-2} is also bounded from bellow.

This last inequality characterizes the truncation error. The Fourier interpolation then gives the first order space discretization error and the Lipschitz property of the driver f completes the proof. □

Theorem 4.2 implies that the structure of the local discretization error does not change when solving a FBSDE with bounded forward coefficients on the alternative grid. The local space discretization error is still of first order and the space truncation error of spectral order with index 2. The boundedness of the forward drift a and volatility σ plays a key role in maintaining these convergence properties for the Fourier interpolation method. Also, the global discretization error defined in equation (3.2.6) displays the same structure in our FBSDE case under a slightly different stability condition that takes into account the presence of bounded forward process coefficients. The next theorem states the result.

Theorem 4.3. *Suppose the conditions of Theorem 4.2 are satisfied. If the space discretization is such that*

$$
\sup_i \max\left(\frac{K_4^{\frac{1}{2}}\Delta x}{\sqrt{2\pi\Delta_i}}, \frac{K_4\Delta x}{\pi\Delta_i}\right) \leq 1 \tag{4.3.2}
$$

then the Fourier interpolation method is stable and the global discretization error $E_{l,\Delta x}$ satisfies

$$
E_{l,\Delta x} = \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-C|\pi|^{-1}l^2}\right) \tag{4.3.3}
$$

where $C > 0$ and K_4 is the upper bound of equation (4.1.8).

Proof. In the proof of Theorem 3.3, we established that

$$
\begin{aligned}
e_{ik} &\leq E_{n-i,k} + |\mathbf{u}_{n-i,k} - u_{n-i,k}| \\
&\leq E_{n-i,k} + (1 + \Delta_i K)\,|\tilde{\mathbf{u}}_{n-i,k} - \tilde{u}_{n-i,k}| \\
&\quad + \Delta_i K\,|\dot{\mathbf{u}}_{n-i,k} - \dot{u}_{n-i,k}| \qquad (4.3.4) \\
\dot{e}_{ik} &\leq E_{n-i,k} + |\dot{\mathbf{u}}_{n-i,k} - \dot{u}_{n-i,k}|\,. \qquad (4.3.5)
\end{aligned}
$$

In the FBSDE case presented in this chapter, and assuming that the values of the function u_{i+1} and the sequence $\{u_{i+1,s}\}_{s=0}^{N_{i+1}N}$ match at the boundaries of the truncated interval,

$$
\begin{aligned}
|\tilde{\mathbf{u}}_{i,k} - \tilde{u}_{i,k}| &\leq \left|\mathfrak{D}^{-1}\left[\{\phi(\nu_{i+1,j})\mathbb{D}[u_{i+1} - u_{i+1,s}]_j\}_{j=0}^{N_{i+1}N-1}\right]_{k+\frac{N}{2}}\right| \\
&\leq \frac{1}{N_{i+1}N}\left(\sum_{j=0}^{N_{i+1}N-1}|\phi(\nu_{i+1,j}, x_{ik})|\right)\sup_k |u_{i+1}(x_{ik}) - u_{i+1,k}| \\
&\quad \text{(using the matrix-vector representation of DFTs)}, \\
&\leq \frac{1}{N_{i+1}N}\left(\sum_{j=0}^{N_{i+1}N-1}|\phi(\nu_{i+1,j}, x_{ik})|\right)\sup_k e_{n-i-1,k} \\
&\leq \frac{(\Delta\nu_{i+1})^{-1}}{N_{i+1}N}\left(\int_{\mathbb{R}}|\phi(\nu, x_{ik})|\,d\nu\right)\sup_k e_{n-i-1,k}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\Delta x}{(2\pi\sigma^2(t_i,x_{ik})\Delta_i)^{\frac{1}{2}}} \sup_k e_{n-i-1,k} \\
&\leq \frac{K_4^{\frac{1}{2}}\Delta x}{(2\pi\Delta_i)^{\frac{1}{2}}} \sup_k e_{n-i-1,k}. \qquad (4.3.6)
\end{aligned}
$$

where the last inequality holds by Assumption 4.1.1. Similarly,

$$
\begin{aligned}
|\dot{\mathbf{u}}_{i,k} - \dot{u}_{i,k}| &\leq \left| \mathfrak{D}^{-1}\left[\{\psi(\nu_{i+1,j},x_{ik})\mathbb{D}[u_{i+1}-u_{i+1,s}]_j \}_{j=0}^{N_{i+1}N-1} \right]_{k+\frac{N}{2}} \right| \\
&\leq \frac{1}{N_{i+1}N} \left(\sum_{j=0}^{N_{i+1}N-1} |\mathbf{i}\nu_{i+1,j}\phi(\nu_{i+1,j},x_{ik})| \right) \sup_k e_{n-i-1,k} \\
&\quad \text{(using the matrix representation of DFTs)}, \\
&\leq \frac{(\Delta\nu_{i+1})^{-1}}{N_{i+1}N} \left(\int_{\mathbb{R}} |\nu\phi(\nu,x_{ik})|\, d\nu \right) \sup_k e_{n-i-1,k} \\
&= \frac{\Delta x}{\pi\sigma^2(t_i,x_{ik})\Delta_i} \sup_k e_{n-i-1,k} \\
&\leq \frac{K_4\Delta x}{\pi\Delta_i} \sup_k e_{n-i-1,k}. \qquad (4.3.7)
\end{aligned}
$$

The rest of the proof is identical to the proof of Theorem 3.3. The inequalities of equations (4.3.4), (4.3.6) and (4.3.7) lead to

$$
\begin{aligned}
e_{i,k} &\leq C_0 E_{i,k} + (1+2\Delta_i K)\max\left(\frac{\Delta x}{\sqrt{2\pi\Delta_i}}, \frac{\Delta x}{\pi\Delta_i} \right) \sup_k e_{i-1,k} \\
&\leq C_0 \sup_{i,k} E_{i,k} + (1+2\Delta_i K)\max\left(\frac{\Delta x}{\sqrt{2\pi\Delta_i}}, \frac{\Delta x}{\pi\Delta_i} \right) \sup_k e_{i-1,k}
\end{aligned}
$$

where $C_0 > 0$ and $K > 0$ is the Lipschitz constant of the driver f. Consequently,

$$
\begin{aligned}
\sup_k e_{i,k} &\leq C_0 \sup_{i,k} E_{i,k} + (1+2\Delta_i K)\max\left(\frac{\Delta x}{\sqrt{2\pi\Delta_i}}, \frac{\Delta x}{\pi\Delta_i} \right) \sup_k e_{i-1,k} \\
&\leq C_0 \sup_{i,k} E_{i,k} + (1+2\Delta_i K)\zeta \sup_k e_{i-1,k} \qquad (4.3.8)
\end{aligned}
$$

for some positive number ζ satisfying

$$
\sup_i \max\left(\frac{K_4^{\frac{1}{2}}\Delta x}{\sqrt{2\pi\Delta_i}}, \frac{K_4\Delta x}{\pi\Delta_i} \right) \leq \zeta \leq 1.
$$

and the Gronwall's Lemma yields

$$
\sup_k e_{i,k} \leq C_0 e^{2TK} \sup_{i,k} E_{i,k} \qquad (4.3.9)
$$

from the inequality of equation (4.3.8) for $i = 0, 1, ..., n$ knowing that $e_{0,k} = 0$. The last equation establishes the stability of the Fourier interpolation method for the approximate solution u_i since its error at any time step is absolutely bounded.

The inequalities of equations (4.3.5), (4.3.7) and (4.3.9) lead to

$$
\begin{aligned}
\sup_k \dot{e}_{i,k} &\leq \left(C_1 + \frac{\Delta x}{\pi\Delta_i} C_0 e^{2TK} \right) \sup_{i,k} E_{i,k} \\
&\leq \left(C_1 + C_0 e^{2TK} \right) \sup_{i,k} E_{i,k} \qquad (4.3.10)
\end{aligned}
$$

for a positive constant $C_1 > 0$. Hence, the convolution method is also stable for the approximate gradient $\dot{u}_i$.

The result of equation (4.3.3) follows by taking the supremum on the left hand sides of equations (4.3.9) and (4.3.10) other time steps and applying Theorem 4.2. □

Let $(U_i, \dot{U}_i)$ be the extended solution at time mesh t_i as defined at equations (3.3.2) and (3.3.3). The numerical solution $\{(x_t, y_t, z_t)\}_{t\in[0,T)}$ to the FBSDE (4.1.1) takes the form

$$x_t = \sum_{i=0}^{n-1} X_{t_i}^{\pi} \mathbf{1}_{[t_i,t_{i+1})}(t) \tag{4.3.11}$$

$$y_t = \sum_{i=0}^{n-1} U_i(X_{t_i}^{\pi}) \mathbf{1}_{[t_i,t_{i+1})}(t) \tag{4.3.12}$$

$$z_t = \sum_{i=0}^{n-1} \dot{U}_i(X_{t_i}^{\pi}) \mathbf{1}_{[t_i,t_{i+1})}(t) \tag{4.3.13}$$

on the time interval $[0, T)$. The quadratic error on the FBSDE solution is

$$\begin{aligned} E^2_{\pi,l,\Delta x} := & \max_{0\le i<n} \mathbf{E}\left[\sup_{t\in[t_i,t_{i+1})} |X_t - x_t|^2\right] + \max_{0\le i<n} \mathbf{E}\left[\sup_{t\in[t_i,t_{i+1})} |Y_t - y_t|^2\right] \\ & + \sum_{i=0}^{n-1} \mathbf{E}\left[\int_{t_i}^{t_{i+1}} |Z_s - z_s|^2 \, ds\right] \end{aligned} \tag{4.3.14}$$

and the next theorem, inspired by Theorem 3.9, describes its error bound.

Theorem 4.4. *Suppose the driver $f \in \mathcal{C}^{1,2,2}$, the terminal condition $g \in \mathcal{C}^2$, the stability and convergence condition of equation (3.2.9) is satisfied and*

$$\xi = g(X_T) \in L^4 \tag{4.3.15}$$

then

$$E^2_{\pi,l,\Delta x} = \mathcal{O}(|\pi|) + \mathcal{O}(\Delta x^2) + \mathcal{O}\left(e^{-C(N_0+1)^2 l^2} + e^{-C|\pi|^{-1} l^2}\right) \tag{4.3.16}$$

where $C > 0$.

Proof. Since the quadratic error of the Euler scheme on the forward process is of first order in time we clearly have

$$\max_{0\le i<n} \mathbf{E}\left[\sup_{t\in[t_i,t_{i+1})} |X_t - x_t|^2\right] = \mathcal{O}(|\pi|). \tag{4.3.17}$$

As to the backward part, the proof follows the steps of Theorem 3.9. The Euler scheme also gives a quadratic error of first order in time for the backward and control processes. The quadratic space discretization error and the spectral space truncation error are consequences of Theorem 4.3.

The error due the simulation on a finite grid reduces to the expression $\max_{1\le i<n} \mathbf{E}\left[\mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(X_{t_i}^{\pi})\right]$ whenever $g(X_T) \in L^4$. Since the intervals $\mathcal{I}_i$ are centered at X_0, we further have

$$\begin{aligned} \max_{1\le i<n} \mathbf{E}\left[\mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(X_{t_i}^{\pi})\right] \le & \max_{1\le i<n} \mathbf{P}\left[X_{t_i}^{\pi} \ge X_0 + N_i\frac{l}{2}\right] \\ & + \max_{1\le i<n} \mathbf{P}\left[X_{t_i}^{\pi} \le X_0 - N_i\frac{l}{2}\right] \\ \le & \max_{0<i<n} \inf_{s>0} e^{-s\left(X_0 + N_i\frac{l}{2}\right)} M_{t_i}(s) \\ & + \max_{0<i<n} \inf_{s>0} e^{s\left(X_0 - N_i\frac{l}{2}\right)} M_{t_i}(-s) \end{aligned}$$

by Chernoff's inequality where M_{t_i} is the moment generating function of $X_{t_i}^{\pi}$. The moment generating function M_{t_i} satisfies

$$M_{t_i}(s) = \mathbf{E}\left[e^{sX_{t_i}^{\pi}}\right]$$

$$
\begin{aligned}
&= \mathbf{E}\left[e^{sX^\pi_{t_{i-1}}}\mathbf{E}_{t_i}\left[e^{s\left(X^\pi_{t_i}-X^\pi_{t_{i-1}}\right)}\right]\right] \\
&= \mathbf{E}\left[e^{sX^\pi_{t_{i-1}}}\phi_{i-1}(-\mathbf{i}s, X^\pi_{t_{i-1}})\right] \\
&\leq e^{\frac{1}{2}\Delta_{i-1}K_2^2 s^2+\Delta_{i-1}K_2|s|}\mathbf{E}\left[e^{sX^\pi_{t_{i-1}}}\right] \\
&\quad \text{(by Assumption 4.1.1)}, \\
&\leq e^{sX_0}e^{\frac{1}{2}t_iK_2^2s^2+t_iK_2|s|} \\
&\quad \text{(after a recursion)}, \\
&\leq e^{sX_0}e^{\frac{1}{2}TK_2^2s^2+TK_2|s|}
\end{aligned}
$$

for any $s \in \mathbb{R}$, so that

$$
\begin{aligned}
\max_{1\leq i<n}\mathbf{E}\left[\mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(X^\pi_{t_i})\right] &\leq 2\max_{0<i<n}\inf_{s>0}e^{-sN_i\frac{l}{2}}e^{\frac{1}{2}TK_2^2s^2+TK_2s} \\
&\leq 2\max_{0<i<n}e^{-\frac{1}{2TK_2^2}\left(TK_2-N_i\frac{l}{2}\right)^2} \\
&\leq K\max_{0<i<n}e^{-CN_i^2l^2} \\
&\leq Ke^{-C(N_0+1)^2l^2}.
\end{aligned}
$$

□

Overall, the boundedness of the forward drift a and volatility σ is crucial in the Fourier interpolation method. Not only does it allow us to derive a bound for the local truncation error in Theorem 4.2 but it also leads to the method's global stability as shown by equation (4.3.2). The proof of Theorem 4.4 also indicates that the boundedness of the forward coefficients is necessary for the control of the simulation error.

The Fourier interpolation method can obviously be extended to reflected FBSDEs following the procedure of Section 3.4.1. For a reflected FBSDE with a lower barrier of the form

$$
\begin{cases}
dX_t = a(t,X_t)dt + \sigma(t,X_t)dW_t \\
-dY_t = f(t,X_t,Y_t,Z_t,)dt + dA_t - Z_t dW_t \\
Y_t \geq B_t \ , \ dA_t \geq 0 \ , \ \forall t \in [0,T] \\
\int_0^T (Y_t - B_t)dA_t = 0 \\
X_0 = x_0 \ , \ Y_T = g(X_T)
\end{cases} \tag{4.3.18}
$$

where

$$
B_t = B(t,X_t) \tag{4.3.19}
$$

we may define the approximate solutions of the reflected FBSDE as

$$
u_i(x) = \tilde{u}_i(x) + \Delta_i f(t_i,x,\tilde{u}_i(x),\dot{u}_i(x)) + \Delta\bar{u}_i(x) \tag{4.3.20}
$$

$$
\dot{u}_i(x) = \frac{1}{\Delta_i}\int_{-\infty}^{\infty}(y-x)u_{i+1}(x+y)h(y|x)dy \tag{4.3.21}
$$

$$
\tilde{u}_i(x) = \int_{-\infty}^{\infty}u_{i+1}(x+y)h(y|x)dy \tag{4.3.22}
$$

$$
\Delta\bar{u}_i(x) = [\tilde{u}_i(x) + \Delta_i f(t_i,x,\tilde{u}_i(x),\dot{u}_i(x)) - B(t_i,x)]^- \tag{4.3.23}
$$

for $i = 0,1,...,n-1$ and $u_n(x) = g(x)$. The Fourier interpolation method presented in this chapter allows us to compute the values of numerical solutions u_{ik} and $\dot{u}_{ik}$ for the backward and control processes. The values for numerical solution of the reflecting process $\Delta\bar{u}_{ik}$ follow naturally. Defining the extended functions U_i, $\dot{U}_i$ and $\Delta\bar{U}_i$ as in equations (3.3.2), (3.3.3) and

(3.4.7) respectively, the numerical solution (x, y, z, a) of the reflected FBSDE solution (X, Y, Z, A) write

$$x_t = \sum_{i=0}^{n-1} X_{t_i}^{\pi} \mathbf{1}_{[t_i, t_{i+1})}(t) \tag{4.3.24}$$

$$y_t = \sum_{i=0}^{n-1} U_i(X_{t_i}^{\pi}) \mathbf{1}_{[t_i, t_{i+1})}(t) \tag{4.3.25}$$

$$z_t = \sum_{i=0}^{n-1} \dot{U}_i(X_{t_i}^{\pi}) \mathbf{1}_{[t_i, t_{i+1})}(t) \tag{4.3.26}$$

$$a_t = \sum_{i=0}^{n-1} \Delta \bar{U}_i(X_{t_i}^{\pi}) \mathbf{1}_{[t_i, T)}(t). \tag{4.3.27}$$

For this chapter, we will omit the numerical results. The material presented here is a generalization of Chapter 3 . Hence, the numerical results from the previous chapters and the error analysis conducted in this one already give an insight of the Fourier interpolation method performance. Also, the present chapter can be viewed as a special case of the following chapter so that the numerical results provided in Chapter 5 confirm the error analysis on the method presented here.

Chapter 5

Discretization of FBSDEs with Runge-Kutta schemes

Runge-Kutta schemes, introduced by Chassagneux and Crisan [27], are a recent development of time discretization methods for FBSDEs. They are themselves an extension of the well-known Runge-Kutta methods for ODEs to FBSDEs. Hence, the schemes produce numerical solutions of higher order of convergence and their properties are studied in Chassagneux and Crisan [27] on a decoupled FBSDE. More precisely, the authors provide a discretization of the backward SDE assuming that a discretization of the forward SDE is available.

In this chapter, we develop a Fourier method for the numerical solution of one-dimensional BSDEs using Runge-Kutta schemes. The main ideas leading to BSDE numerical solutions were already introduced in Chapter 3. Hence, we shall use the Fourier transform representations of the various conditional expectations involved in the Runge-Kutta schemes to define approximate solutions to the BSDE. These approximate solutions are then computed on an alternative grid using the FFT algorithm in order to retrieve higher order convergent numerical solutions.

5.1 Runge-Kutta schemes

This section gives the formal definition of Runge-Kutta schemes. We also propose simple assumptions on the time discretization of the forward SDE which lead to a simplification of the scheme expressions.

5.1.1 Time discretization

In the general setting, the FBSDE considered is of the form

$$\begin{cases} dX_t = a(X_t)dt + \sigma(X_t)dW_t \\ -dY_t = f(Y_t, Z_t)dt - Z_t^* dW_t \\ X_0 = x_0 \ , \ Y_T = g(X_T) \end{cases} \tag{5.1.1}$$

where W is a $d-$dimensional Brownian motion. The forward drift $a : \mathbb{R}^d \to \mathbb{R}^d$, the forward volatility $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$, the driver $f : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and the terminal function $g : \mathbb{R}^d \to \mathbb{R}$ are all deterministic functions. It is always possible to generalize the system to consider time dependent coefficients.

Assumption 5.1.1. *In addition to the usual Lipschitz and linear growth conditions* [1]*, we suppose that all coefficients are bounded. In particular, the backward drift* $f \in \mathcal{C}_b^m$ *and the terminal function* $g \in \mathcal{C}_b^m$ *with* $m \geq 2$ *are at least twice differentiable.*

The BSDE is discretized on the time partition $\pi = \{0 = t_0 < t_1 < ... < t_n = T\}$ as usual with successive time steps

$$\Delta_i = t_{i+1} - t_i \ , \ i = 0, 1, ..., n-1$$

and maximal time step

$$|\pi| := \max_{0 \leq i < n} \Delta_i. \tag{5.1.2}$$

Let $q \in \mathbb{N}^*$, we consider the q-stage Runge-Kutta scheme giving the following numerical solution at mesh time t_i

$$Z_{t_i}^\pi = \mathbf{E}_{t_i}\left[H_{t_i,\Delta_i}^{\varphi_1} Y_{t_{i+1}}^\pi + \Delta_i \sum_{j=1}^{q} \beta_j H_{t_i,(1-\gamma_j)\Delta_i}^{\varphi_1} f(Y_{i,j}^\pi, Z_{i,j}^\pi)\right] \tag{5.1.3}$$

$$Y_{t_i}^\pi = \mathbf{E}_{t_i}\left[Y_{t_{i+1}}^\pi + \Delta_i \sum_{j=1}^{q+1} \alpha_j f(Y_{i,j}^\pi, Z_{i,j}^\pi)\right] \tag{5.1.4}$$

for a set positive coefficients $\{\gamma_j\}_{j=1}^{q+1}$ such that $0 = \gamma_1 < ... < \gamma_{q+1} = 1$. The intermediate solutions $\{(Y_{i,j}^\pi, Z_{i,j}^\pi)\}_{j=2}^{q}$ take the form

$$Z_{i,j}^\pi = \mathbf{E}_{t_{i,j}}\left[H_{t_{i,j},\gamma_j\Delta_i}^{\varphi_j} Y_{t_{i+1}}^\pi + \Delta_i \sum_{k=1}^{j-1} \beta_{jk} H_{t_{i,j},(\gamma_j-\gamma_k)\Delta_i}^{\varphi_j} f(Y_{i,k}^\pi, Z_{i,k}^\pi)\right] \tag{5.1.5}$$

$$Y_{i,j}^\pi = \mathbf{E}_{t_{i,j}}\left[Y_{t_{i+1}}^\pi + \Delta_i \sum_{k=1}^{j} \alpha_{jk} f(Y_{i,k}^\pi, Z_{i,k}^\pi)\right] \tag{5.1.6}$$

where

$$t_{i,j} = t_i + (1 - \gamma_j)\Delta_i, \ 1 \leq j \leq q+1 \tag{5.1.7}$$

with $(Y_{i,1}^\pi, Z_{i,1}^\pi) = (Y_{t_{i+1}}^\pi, Z_{t_{i+1}}^\pi)$, $(Y_{i,q+1}^\pi, Z_{i,q+1}^\pi) = (Y_{t_i}^\pi, Z_{t_i}^\pi)$ and terminal condition

$$(Y_{t_n}, Z_{t_n}) = (g(X_T), \sigma^*(X_T)\nabla g(X_T)). \tag{5.1.8}$$

The coefficients $\{\alpha_j\}_{j=1}^{q+1}$, $\{\beta_j\}_{j=1}^{q}$, $\{\alpha_{jk} : 1 \leq j \leq q, 1 \leq k \leq j\}$ and $\{\beta_{jk} : 1 \leq j \leq q, 1 \leq k < j\}$ are all positive and satisfy

$$\sum_{j=1}^{q+1} \alpha_j = 1 \tag{5.1.9}$$

$$\beta_{jj} = 0, \ 1 \leq j \leq q, \tag{5.1.10}$$

$$\sum_{k=1}^{j} \alpha_{jk} = \sum_{k=1}^{j-1} \beta_{jk} = \gamma_j, \ 1 < j \leq q. \tag{5.1.11}$$

Let $\mathcal{B}^m$ denote the set of continuous and bounded functions on $[0,1]$ such that

$$\mathcal{B}^m := \{\phi \in \mathcal{C}_b : \int_0^1 s^k \phi(s) ds = \delta_{0,k}, k \leq m \text{ and } k, m \in \mathbb{N}^*\}. \tag{5.1.12}$$

The stochastic coefficient $H_{t,\Delta}^\varphi$ with $t \in [0, T)$ and $\Delta > 0$ is defined as

$$H_{t,\Delta}^\varphi := \frac{1}{\Delta} \int_t^{t+\Delta} \varphi\left(\frac{s-t}{\Delta}\right) dW_s \tag{5.1.13}$$

[1] See Assumption 4.1.1.

with $\varphi \in \mathcal{B}^m$ for some $m \in \mathbb{N}^*$.

The global error of the $q-$stage Runge-Kutta scheme $\mathcal{E}_\pi$ is defined as

$$\begin{aligned}\mathcal{E}_\pi^2 &:= \max_{0\le i<n} \left\|Y_{t_i} - Y_{t_i}^\pi\right\|_{L^2}^2 + \sum_{i=0}^{n-1} \Delta_i \left\|Z_{t_i} - Z_{t_i}^\pi\right\|_{L^2}^2 \\ &= \max_{0\le i<n} \mathbf{E}\left[\left|Y_{t_i} - Y_{t_i}^\pi\right|^2\right] + \sum_{i=0}^{n-1} \Delta_i \mathbf{E}\left[\left|Z_{t_i} - Z_{t_i}^\pi\right|^2\right] \end{aligned} \tag{5.1.14}$$

and is hence weaker than the error E_π considered for the Euler scheme. Nonetheless, the global error $\mathcal{E}_\pi$ is easier to handle since it is strongly related to the local time discretization error.

The scheme can be represented by the following tableau

γ_1	$\alpha_{1,1}$	0	$\dots$	0	0	$\beta_{1,1}$	0	$\dots$	0
γ_2	$\alpha_{2,1}$	$\alpha_{2,2}$	$\dots$	0	0	$\beta_{2,1}$	$\beta_{2,2}$	$\dots$	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
γ_q	$\alpha_{q,1}$	$\alpha_{q,2}$	$\dots$	$\alpha_{q,q}$	0	$\beta_{q,1}$	$\beta_{q,2}$	$\dots$	$\beta_{q,q}$
γ_{q+1}	α_1	α_2	$\dots$	α_q	α_{q+1}	β_1	β_2	$\dots$	β_q

One can observe that if $\alpha_{q+1} = 0$ and $\alpha_{jj} = 0$, $1 < j \le q$, then the q-stage Runge-Kutta scheme is explicit. Otherwise, the scheme is implicit. For instance, the Runge-Kutta schemes with tableau

0	0	0	0
1	0	1	1

and the scheme with tableau

0	0	0	0
1	$\frac{1}{2}$	$\frac{1}{2}$	1

known as the Crank-Nicholson scheme constitute $1-$stage implicit Runge-Kutta schemes. The only $1-$stage explicit Runge-Kutta scheme admits the tableau

0	0	0	0
1	1	0	1

In Chassagneux and Crisan [27], the implicit and the explicit $1-$stage Runge-Kutta schemes are shown to be one-half $(\frac{1}{2})$ order convergent. The Crank-Nicholson scheme, already studied in Crisan and Manolarakis [31], presents a first order of convergence. Notice that the Euler schemes used in the previous chapters are not $1-$stage Runge-Kutta schemes since they do not lead to any consistent tableau. Nonetheless, their structure is equivalent to the explicit $1-$stage Runge-Kutta scheme and both schemes display the same half $(\frac{1}{2})$ order of convergence. The following tableau gives a example of explicit 2-stage Runge-Kutta schemes of first order of convergence for $\gamma_2 \in (0, 1]$ and $\beta_1 \in [0, 1]$.

0	0	0	0	0	0
γ_2	γ_2	0	0	γ_2	0
1	$1 - \frac{1}{2\gamma_2}$	$\frac{1}{2\gamma_2}$	0	β_1	$1 - \beta_1$

5.1.2 Further simplification

From the q-stage Runge-Kutta scheme for BSDEs, one notices that we have at least $2q$ conditional expectations to compute at each time step. These conditional expectations can be simplified and made more suitable for numerical implementation if we consider a reasonable time discretization of the forward SDE. Hence, we make the following assumption that we will use throughout the chapter.

Assumption 5.1.2. *(1). The forward SDE is discretized with the piecewise constant process* X^π *such that for* $t \in [t_i, t_{i+1})$ *we have*

$$X_t^\pi = X_{t_i}^\pi \tag{5.1.15}$$

pathwise.

(2). The forward SDE time discretization with global error $\mathcal{E}_{X,\pi}$ *is of order* $m > 0$ *i.e*

$$\mathcal{E}_{X,\pi}^2 := \max_{0 \le i \le n} \left\| X_{t_i} - X_{t_i}^\pi \right\|_{L^2}^2 = \mathcal{O}(|\pi|^{2m}). \tag{5.1.16}$$

The conditions of Assumption 5.1.2 are not hard to meet. Many higher order time discretizations for forward SDEs satisfying the conditions for a given order $m > 0$ are indeed available. Appendix B stands as an introduction to Itô-Taylor expansion based schemes as an example and a more complete presentation of these schemes can be found in Kloeden and Platen [69] among others. The next theorem gives a simplification of the BSDE time discretization expressions.

Theorem 5.1. *Under Assumption 5.1.2 (1), the solution of the q-stage Runge-Kutta scheme satisfies*

$$\{(Y_{i,j}^\pi, Z_{i,j}^\pi)\}_{j=2}^{q+1} \in \mathcal{F}_{t_i} \tag{5.1.17}$$

for $0 \le i < n$. *Consequently, we can write*

$$Z_{i,j}^\pi = \boldsymbol{E}_{t_i}\left[H_{t_{i,j},\gamma_j \Delta_i}^{\varphi_j} \left(Y_{t_{i+1}}^\pi + \Delta_i \beta_{j,1} f(Y_{t_{i+1}}^\pi, Z_{t_{i+1}}^\pi) \right) \right] \tag{5.1.18}$$

$$\begin{aligned} Y_{i,j}^\pi &= \boldsymbol{E}_{t_i}\left[Y_{t_{i+1}}^\pi + \Delta_i \alpha_{j,1} f(Y_{t_{i+1}}^\pi, Z_{t_{i+1}}^\pi) \right] \\ &\quad + \Delta_i \sum_{k=2}^{j} \alpha_{jk} f(Y_{i,k}^\pi, Z_{i,k}^\pi) \end{aligned} \tag{5.1.19}$$

for $0 \le i < n$ *and* $1 < j \le q+1$ *where* $\varphi_{q+1} = \varphi_1$, $\beta_{q+1,1} = \beta_1$ *and* $\alpha_{q+1,k} = \alpha_k$.

Proof. Clearly $(Y_{i,q+1}^\pi, Z_{i,q+1}^\pi) = (Y_{t_i}^\pi, Z_{t_i}^\pi) \in \mathcal{F}_{t_i}$ from equations (5.1.3) and (5.1.4). For $1 < j \le q$ and $0 \le i < n$, we have

$$\begin{aligned} Y_{i,j}^\pi &= \mathbf{E}\left[Y_{t_{i+1}}^\pi + \Delta_i \sum_{k=1}^{j} \alpha_{jk} f(Y_{i,k}^\pi, Z_{i,k}^\pi) \mid X_{t_{i,j}}^\pi \right] \\ &\quad \text{(starting from equation (5.1.6))}, \\ &= \mathbf{E}\left[Y_{t_{i+1}}^\pi + \Delta_i \sum_{k=1}^{j} \alpha_{jk} f(Y_{i,k}^\pi, Z_{i,k}^\pi) \mid X_{t_i}^\pi \right] \\ &\quad \text{(by Assumption 5.1.2 since } t_{i,j} \in [t_i, t_{i+1})), \\ &= \mathbf{E}_{t_i}\left[Y_{t_{i+1}}^\pi + \Delta_i \sum_{k=1}^{j} \alpha_{jk} f(Y_{i,k}^\pi, Z_{i,k}^\pi) \right] \end{aligned}$$

so that $Y_{i,j}^\pi \in \mathcal{F}_{t_i}$. Similar arguments also show that $Z_{i,j}^\pi \in \mathcal{F}_{t_i}$ starting from equation (5.1.5).

Since$\{(Y_{i,j}^\pi, Z_{i,j}^\pi)\}_{j=2}^{q+1} \in \mathcal{F}_{t_i}$, we naturally get equation (5.1.19) from equations (5.1.6) and (5.1.4) . In addition, knowing that

$$\mathbf{E}_{t_i}\left[H_{t_{i,j},(\gamma_i - \gamma_k)\Delta_i}^{\varphi_j} \right] = 0 \ , \ 1 < k < j \tag{5.1.20}$$

leads to equation (5.1.18) from equations (5.1.5) and (5.1.3). □

As a consequence of Assumption 5.1.2, if the $q-$stage Runge-Kutta scheme and the forward SDE time discretization are of order $m > 0$ then error of the FBSDE numerical solution defined as $\mathcal{E}_{X,\pi} + \mathcal{E}_\pi$ is of order m. We must hence choose the Runge-Kutta scheme and the SDE scheme accordingly.

5.2 Fourier transform representations

We develop the Fourier representation of our approximate solution for the BSDE under various time discretizations for the forward SDE. First we focus on the simple framework of BSDEs then we consider the FBSDE case. For implicit Runge-Kutta schemes, we assume that

$$|\pi|\, K < 1 \tag{5.2.1}$$

where K is the Lipschitz constant of the driver f for well-posedness reasons.

5.2.1 The BSDE case

In the context of BSDEs where the forward process is simply a Brownian motion[2]

$$Y_t = g(W_T) + \int_t^T f(Y_s, Z_s)ds + \int_t^T Z_s dW_s, \tag{5.2.2}$$

the Euler time discretization satisfies Assumption 5.1.2 at any order of convergence $m > 0$. Indeed, the Euler scheme is exact at any time node for a Brownian motion. We can then used the result of Theorem 5.1 with any Runge-Kutta scheme for BSDE without limiting the scheme convergence order.

Following Theorem 5.1, the intermediate solution $\{(u_{i,j}, \dot{u}_{i,j})\}_{j=2}^{q+1}$ at mesh time t_i, $0 \leq i < n$, are given by

$$\dot{u}_{i,j}(x) = \mathbf{E}\left[H^{\varphi_j}_{t_{i,j},\gamma_j\Delta_i}\tilde{u}_{i+1}(W_{t_{i+1}}, \beta_{j,1})|W_{t_i} = x\right] \tag{5.2.3}$$

$$u_{i,j}(x) = \mathbf{E}\left[\tilde{u}_{i+1}(W_{t_{i+1}}, \alpha_{j,1})|W_{t_i} = x\right] + \Delta_i \sum_{k=2}^{j} \alpha_{jk} f(u_{i,k}(x), \dot{u}_{i,k}(x)) \tag{5.2.4}$$

for $1 < j \leq q+1$ with $\varphi_{q+1} = \varphi_1$, $\beta_{q+1,1} = \beta_1$ and $\alpha_{q+1,k} = \alpha_k$. The approximate solution u_i and approximate gradient $\dot{u}_i$ at mesh time t_i, $0 \leq i < n$, are then

$$u_i(x) = u_{i,q+1}(x) \tag{5.2.5}$$

$$\dot{u}_i(x) = \dot{u}_{i,q+1}(x) \tag{5.2.6}$$

with

$$\tilde{u}_{i+1}(x, \alpha) = u_{i+1}(x) + \Delta_i \alpha f(u_{i+1}(x), \dot{u}_{i+1}(x)) \tag{5.2.7}$$

and

$$u_n(x) = g(x) \tag{5.2.8}$$

$$\dot{u}_n(x) = \nabla g(x). \tag{5.2.9}$$

From the analysis performed in Chapter 2 and Chapter 3, equation (5.2.4) naturally leads to

$$u_{i,j}(x) = \mathfrak{F}^{-1}\left[\mathfrak{F}\left[\tilde{u}_{i+1}(., \alpha_{j,1})\right](\nu)\phi(\nu)\right](x) + \Delta_i \sum_{k=2}^{j} \alpha_{jk} f(u_{i,k}(x), \dot{u}_{i,k}(x)) \tag{5.2.10}$$

when $\tilde{u}_{i+1}(., \alpha)$ is Lebesgue integrable where ϕ is characteristic function on the Brownian increment

$$\phi(\nu) = \exp\left(-\frac{1}{2}\Delta_i \nu^* \nu\right) \tag{5.2.11}$$

Moreover, if the function $\tilde{u}(., \beta) \in \mathcal{C}^1$ is differentiable, equation (5.2.3) gives

$$\dot{u}_{i,j}(x) = \frac{1}{\gamma_j \Delta_i}\mathbf{E}\left[\nabla\tilde{u}_{i+1}(W_{t_{i+1}}, \beta_{j,1}) \int_{t_{i,j}}^{t_{i+1}} \varphi_j\left(\frac{s - t_{i,j}}{\gamma_j \Delta_i}\right) ds | W_{t_i} = x\right]$$

[2] This can be extended to arithmetic Brownian motion and Brownian motion with time dependent coefficients.

(by the duality formula),

$$
\begin{aligned}
&= \mathbf{E}\left[\nabla\tilde{u}_{i+1}(W_{t_{i+1}},\beta_{j,1})|W_{t_i}=x\right]\\
&= \mathfrak{F}^{-1}\left[\mathfrak{F}\left[\nabla\tilde{u}_{i+1}(.;\beta_{j,1})\right](\nu)\phi(\nu)\right](x)\\
&\quad \text{(when } \nabla\tilde{u}_{i+1}(.,\beta) \text{ is Lebesque integrable)},\\
&= \mathfrak{F}^{-1}\left[\mathbf{i}\nu\mathfrak{F}[\tilde{u}_{i+1}(.;\beta_{j,1})](\nu)\phi(\nu)\right](x) \qquad (5.2.12)
\end{aligned}
$$

where the last inequality holds by Proposition E.4.

Equations (5.2.10) and (5.2.12) characterize the Fourier transform representation of $q-$stage Runge-Kutta schemes. It is important to note that these expressions do not depend on the function φ_j. As a consequence their choice is irrelevant under Fourier methods in the BSDE case even though they play a key role in the convergence of Runge-Kutta schemes. When using the Fourier representations, one can always assume that the functions φ_j are consistent with the convergence order of the Runge-Kutta scheme. In addition, the expressions for the conditional expectations are essentially similar to those developed in Chapter 3.

5.2.2 The FBSDE case

Runge-Kutta schemes for FBSDEs require higher order time discretizations for SDE for convergence reasons. In order to develop Fourier representations of the FBSDE numerical solutions, another requirement is the availability of explicit conditional characteristic functions. We hence make the following additional assumption on the forward SDE time discretization.

Assumption 5.2.1. *The forward SDE time discretization admits the conditional characteristic functions* $\phi_i : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$

$$\phi_i(\nu,x) = \boldsymbol{E}\left[e^{\mathbf{i}\nu^*\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}|X^\pi_{t_i}=x\right] \qquad (5.2.13)$$

and $\Phi_{i,j} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}^d$

$$\Phi_{i,j}(\nu,x) = \boldsymbol{E}\left[H^{\varphi_j}_{t_{i,j},\gamma_j\Delta_i}e^{\mathbf{i}\nu^*\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}|X^\pi_{t_i}=x\right] \qquad (5.2.14)$$

for $0 \le i < n$ *and* $1 < j \le q+1$ *with* $\varphi_{q+1} = \varphi_1$.

In this setting, and letting the terminal conditions be

$$
\begin{aligned}
u_n(x) &= g(x) && (5.2.15)\\
\dot{u}_n(x) &= \sigma^*(x)\nabla g(x), && (5.2.16)
\end{aligned}
$$

the intermediate solutions $u_{i,j}$ at time step t_i, $0 \le i < n$ and $1 < j \le q+1$ are given by

$$
\begin{aligned}
u_{i,j}(x) &= \mathbf{E}\left[\tilde{u}_{i+1}(X^\pi_{t_{i+1}};\alpha_{j,1})|X^\pi_{t_i}=x\right] + \Delta_i\sum_{k=2}^{j}\alpha_{jk}f(u_{i,k}(x),\dot{u}_{i,k}(x))\\
&= \mathbf{E}^x_{t_i}\left[\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}e^{\mathbf{i}\nu^*X^\pi_{t_{i+1}}}\mathfrak{F}\left[\tilde{u}_{i+1}(.;\alpha_{j,1})\right](\nu)d\nu\right]\\
&\quad +\Delta_i\sum_{k=2}^{j}\alpha_{jk}f(u_{i,k}(x),\dot{u}_{i,k}(x))\\
&= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\mathbf{E}^x_{t_i}\left[e^{\mathbf{i}\nu^*X^\pi_{t_{i+1}}}\right]\mathfrak{F}\left[\tilde{u}_{i+1}(.;\alpha_{j,1})\right](\nu)d\nu\\
&\quad +\Delta_i\sum_{k=2}^{j}\alpha_{jk}f(u_{i,k}(x),\dot{u}_{i,k}(x))
\end{aligned}
$$

$$
\begin{aligned}
& \text{(using Fubini's theorem),} \\
= \quad & \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} e^{\mathbf{i}\nu^* x}\phi_i(\nu,x)\mathfrak{F}\left[\tilde{u}_{i+1}(.;\alpha_{j,1})\right](\nu)d\nu \\
& +\Delta_i\sum_{k=2}^{j}\alpha_{jk}f(u_{i,k}(x),\dot{u}_{i,k}(x)) \\
= \quad & \mathfrak{F}^{-1}\left[\mathfrak{F}\left[\tilde{u}_{i+1}(.;\alpha_{j,1})\right](\nu)\phi_i(\nu,x)\right](x)+\Delta_i\sum_{k=2}^{j}\alpha_{jk}f(u_{i,k}(x),\dot{u}_{i,k}(x)) \quad (5.2.17)
\end{aligned}
$$

whenever $\tilde{u}_{i+1}(.,\alpha)$ is Lebesgue integrable.

As to the intermediate solutions $\dot{u}_{i,j}$, $0 \leq i < n$ and $1 < j \leq q+1$, we have

$$
\begin{aligned}
\dot{u}_{i,j}(x) &= \mathbf{E}\left[H^{\varphi_j}_{t_{i,j},\gamma_j\Delta_i}\tilde{u}_{i+1}(X^\pi_{t_{i+1}};\beta_{j,1})|X^\pi_{t_i}=x\right] \\
&= \mathbf{E}^x_{t_i}\left[H^{\varphi_j}_{t_{i,j},\gamma_j\Delta_i}\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}e^{\mathbf{i}\nu^* X^\pi_{t_{i+1}}}\mathfrak{F}\left[\tilde{u}_{i+1}(.;\beta_{j,1})\right](\nu)d\nu\right] \\
&= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\mathbf{E}^x_{t_i}\left[H^{\varphi_j}_{t_{i,j},\gamma_j\Delta_i}e^{\mathbf{i}\nu^* X^\pi_{t_{i+1}}}\right]\mathfrak{F}\left[\tilde{u}_{i+1}(.;\beta_{j,1})\right](\nu)d\nu \\
&\quad \text{(using Fubini's theorem),} \\
&= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}e^{\mathbf{i}\nu^* x}\Phi_{i,j}(\nu,x)\mathfrak{F}\left[\tilde{u}_{i+1}(.;\beta_{j,1})\right](\nu)d\nu \\
&= \mathfrak{F}^{-1}\left[\mathfrak{F}\left[\tilde{u}_{i+1}(.;\alpha_{j,1})\right](\nu)\Phi_{i,j}(\nu,x)\right](x) \qquad (5.2.18)
\end{aligned}
$$

for an integrable function $\tilde{u}_{i+1}(.,\alpha)$. In addition, letting $D_sX^\pi_{t_{i+1}}$ be the Malliavin derivative of $X^\pi_{t_{i+1}}$ given $X^\pi_{t_i}=x$, we have

$$
\begin{aligned}
\Phi_{i,j}(\nu,x) &= \mathbf{E}\left[H^{\varphi_j}_{t_{i,j},\gamma_j\Delta_i}e^{\mathbf{i}\nu^*\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}|X^\pi_{t_i}=x\right] \\
&= \mathbf{i}\nu\mathbf{E}^x_{t_i}\left[\left(\frac{1}{\gamma_j\Delta_i}\int_{t_{i,j}}^{t_{i+1}}D_sX^\pi_{t_{i+1}}\varphi^j\left(\frac{s-t_{i,j}}{\gamma_j\Delta_i}\right)ds\right)e^{\mathbf{i}\nu^*\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}\right] \\
&\quad \text{(by the duality formula),} \\
&= \mathbf{i}\nu\mathbf{E}^x_{t_i}\left[\mathbf{H}_{i,j}e^{\mathbf{i}\nu^*\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}\right] \qquad (5.2.19)
\end{aligned}
$$

with

$$
\mathbf{H}_{i,j}=\frac{1}{\gamma_j\Delta_i}\int_{t_{i,j}}^{t_{i+1}}D_sX^\pi_{t_{i+1}}\varphi_j\left(\frac{s-t_{i,j}}{\gamma_j\Delta_i}\right)ds. \qquad (5.2.20)
$$

Even if the expressions in equations (5.2.17) and (5.2.18) appear too general, they are implementable with the Fourier interpolation method on the alternative grid in various particular cases. Indeed some SDE time discretizations allow us to retrieve not only the characteristics ϕ_i and $\Phi_{i,j}$ and also the Fourier representation under the alternative transform. In the sequel, we give two notable examples with Itô-Taylor expansion schemes for the forward SDE in the one-dimensional case. As already mentioned, an introduction to these schemes can be found in Appendix B. We shall mainly focus on half order and first order schemes.

Half order Itô-Taylor schemes

The Euler scheme constitute the main example of half order Itô-Taylor scheme since

$$
\begin{aligned}
X^\pi_{t_{i+1}} &= X^\pi_{t_i}+a(X^\pi_{t_i})\Delta_i+\sigma(X^\pi_{t_i})\Delta W_i \\
&= X^\pi_{t_i}+\sum_{\imath\in\mathbb{A}_{\frac{1}{2}}\setminus\{\emptyset\}}F_\imath(X^\pi_{t_i})I^\imath_{t_i,t_{i+1}}
\end{aligned}
$$

where $F_{(0)}(x) = a(x)$, $F_{(1)}(x) = \sigma(x)$. In addition, we have that, for $s \in (t_i, t_{i+1})$,

$$D_s I^{(0)}_{t_i,t_{i+1}} = 0 \tag{5.2.21}$$

and

$$D_s I^{(1)}_{t_i,t_{i+1}} = 1. \tag{5.2.22}$$

Hence,

$$D_s X^{\pi}_{t_{i+1}} = F_{(1)}(x) \tag{5.2.23}$$

so we get, from equation (5.2.19), that

$$\begin{aligned} \Phi_{i,j}(\nu, x) &= \mathbf{i}\nu F_{(1)}(x)\mathbf{E}\left[e^{\mathbf{i}\nu\left(X^{\pi}_{t_{i+1}} - X^{\pi}_{t_i}\right)}\right] \text{ (since } \varphi_j \in \mathcal{B}^0), \\ &= F_{(1)}(x)\mathbf{i}\nu\phi_i(\nu, x). \end{aligned} \tag{5.2.24}$$

The conditional characteristic function is explicitly given by

$$\phi_i(\nu, x) = \exp\left\{\Delta_i\left(\mathbf{i}F_{(0)}(x)\nu - \frac{1}{2}F^2_{(1)}(x)\nu^2\right)\right\} \tag{5.2.25}$$

since the increment has a Gaussian distribution.

Equations (5.2.17) and (5.2.18) along with the characteristics of equations (5.2.25) and (5.2.24) define the Fourier method under half order Itô-Taylor schemes for SDEs and the method is implementable with the procedure given in Chapter 3. The following theorem generalizes the result of Lemma 4.1 to half order Itô-Taylor schemes under Runge-Kutta schemes.

Theorem 5.2. *Let $\tilde{u}^{\alpha,\beta}_{i+1}(., y)$ be the alternative transform defined in equation (3.1.8) of the approximate solution $\tilde{u}_{i+1}(., y)$. Then the intermediate solutions $u_{i,j}$ and $\dot{u}_{i,j}$ in equations (5.2.17) and (5.2.18) satisfy*

$$\begin{aligned} u_{i,j}(x) &= \mathfrak{F}^{-1}[\mathfrak{F}[\tilde{u}^{\alpha,\beta}_{i+1}(., \alpha_{j,1})](\nu)\phi_i(\nu, x)](x) \\ &\quad -\alpha[(x + \Delta_i F_{(0)}(x))^2 + \Delta_i F^2_{(1)}(x)] - \beta(x + \Delta_i F_{(0)}(x)) \\ &\quad +\Delta_i \sum_{k=2}^{j} \alpha_{jk} f(u_{i,k}(x), \dot{u}_{i,k}(x)) \end{aligned} \tag{5.2.26}$$

$$\begin{aligned} \dot{u}_{i,j}(x) &= F_{(1)}(x)\mathfrak{F}^{-1}[\mathfrak{F}[\tilde{u}^{\alpha,\beta}_{i+1}(., \beta_{j,1})](\nu)\mathbf{i}\nu\phi_i(\nu, x)](x) \\ &\quad -F_{(1)}(x)[2\alpha(x + \Delta_i F_{(0)}(x)) + \beta]. \end{aligned} \tag{5.2.27}$$

under a half order Itô-Taylor scheme.

First order Itô-Taylor schemes

Consider the first order scheme

$$\begin{aligned} X^{\pi}_{t_{i+1}} &= X^{\pi}_{t_i} + \sum_{\imath \in \mathbb{A}_1 \setminus \{\emptyset\}} F_\imath(X^{\pi}_{t_i}) I^{\imath}_{t_i,t_{i+1}} \\ &= X^{\pi}_{t_i} + \sum_{\imath \in \mathbb{A}_{\frac{1}{2}} \setminus \{\emptyset\}} F_\imath(X^{\pi}_{t_i}) I^{\imath}_{t_i,t_{i+1}} + F_{(1,1)}(X^{\pi}_{t_i}) I^{(1,1)}_{t_i,t_{i+1}}. \end{aligned}$$

Then knowing that, for $s \in (t_i, t_{i+1})$,

$$\begin{aligned} D_s I^{(1,1)}_{t_i,t_{i+1}} &= I^{(1)}_{t_i,s} + I^{(1)}_{s,t_{i+1}} \\ &\quad \text{(using the fundamental theorem of calculus)}, \\ &= I^{(1)}_{t_i,t_{i+1}}, \end{aligned} \tag{5.2.28}$$

the Malliavin derivative of the discretized forward process is given by

$$D_s X^\pi_{t_{i+1}} = F_{(1)}(x) + F_{(1,1)}(x) I^{(1)}_{t_i,t_{i+1}}. \tag{5.2.29}$$

Equation (5.2.19) leads to

$$\begin{aligned}
\Phi_{i,j}(\nu,x) &= \mathbf{i}\nu F_{(1)}(x)\mathbf{E}\left[e^{\mathbf{i}\nu\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}|X^\pi_{t_i}=x\right] \\
&\quad + \mathbf{i}\nu F_{(1,1)}(x)\mathbf{E}^x_{t_i}\left[I^{(1)}_{t_i,t_{i+1}} e^{\mathbf{i}\nu\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}\right] \\
&\quad (\text{since } \varphi_j \in \mathcal{B}^0), \\
&= F_{(1)}(x)\mathbf{i}\nu\phi_i(\nu,x) \\
&\quad + \mathbf{i}\nu F_{(1,1)}(x)\mathbf{E}^x_{t_i}\left[I^{(1)}_{t_i,t_{i+1}} e^{\mathbf{i}\nu\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}\right] \\
&= F_{(1)}(x)\mathbf{i}\nu(1+\zeta_i(\nu,x))\phi_i(\nu,x)
\end{aligned} \tag{5.2.30}$$

with

$$\zeta_i(\nu,x) = \frac{\mathbf{i}\nu F_{(1,1)}(x)\Delta_i}{1-\mathbf{i}\nu F_{(1,1)}(x)\Delta_i} \tag{5.2.31}$$

since

$$\begin{aligned}
&F_{(1,1)}(x)\mathbf{E}^x_{t_i}\left[I^{(1)}_{t_i,t_{i+1}} e^{\mathbf{i}\nu\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}\right] \\
&= F_{(1)}(x)F_{(1,1)}(x)\Delta_i\mathbf{E}^x_{t_i}\left[\mathbf{i}\nu e^{\mathbf{i}\nu\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}\right] \\
&\quad + F^2_{(1,1)}(x)\Delta_i\mathbf{E}^x_{t_i}\left[\mathbf{i}\nu I^{(1)}_{t_i,t_{i+1}} e^{\mathbf{i}\nu\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}\right]
\end{aligned}$$

using the duality formula, so that

$$\begin{aligned}
F_{(1,1)}(x)\mathbf{E}^x_{t_i}\left[I^{(1)}_{t_i,t_{i+1}} e^{\mathbf{i}\nu\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}\right] &= F_{(1)}(x)\zeta_i(\nu,x)\mathbf{E}^x_{t_i}\left[e^{\mathbf{i}\nu\left(X^\pi_{t_{i+1}}-X^\pi_{t_i}\right)}\right] \\
&= F_{(1)}(x)\zeta_i(\nu,x)\phi_i(\nu,x).
\end{aligned} \tag{5.2.32}$$

Equations (5.2.17) and (5.2.18) along with the expression in equation (5.2.30) characterize the method under first order discretizations on the forward process when the characteristic ϕ_i is available. The procedure introduced in Chapter 4 allows to do the computations given the characteristics ϕ_i and $\Phi_{i,j}$ using the following theorem.

Theorem 5.3. *Let $\tilde{u}^{\alpha,\beta}_{i+1}(.,y)$ be the alternative transform defined in equation (3.1.8) of the approximate solution $\tilde{u}_{i+1}(.,y)$. Then the intermediate solutions $u_{i,j}$ and $\dot{u}_{i,j}$ in equations (5.2.17) and (5.2.18) satisfy*

$$\begin{aligned}
u_{i,j}(x) &= \mathfrak{F}^{-1}[\mathfrak{F}[\tilde{u}^{\alpha,\beta}_{i+1}(.,\alpha_{j,1})](\nu)\phi_i(\nu,x)](x) \\
&\quad -\alpha\left[(x+\Delta_i F_0(x))^2 + \Delta_i F^2_{(1)}(x) + \frac{1}{2}\Delta_i^2 F^2_{(1,1)}(x)\right] \\
&\quad -\beta(x+\Delta_i F_0(x)) \\
&\quad +\Delta_i\sum_{k=2}^{j}\alpha_{jk} f(u_{i,k}(x),\dot{u}_{i,k}(x))
\end{aligned} \tag{5.2.33}$$

$$\begin{aligned}
\dot{u}_{i,j}(x) &= F_{(1)}(x)\mathfrak{F}^{-1}[\mathfrak{F}[\tilde{u}^{\alpha,\beta}_{i+1}(.,\beta_{j,1})](\nu)\mathbf{i}\nu(1+\zeta_i(\nu,x))\phi_i(\nu,x)](x) \\
&\quad -F_{(1)}(x)\left[2\alpha\left(x+\Delta_i F_0(x)+\Delta_i F_{(1,1)}(x)\right)+\beta\right].
\end{aligned} \tag{5.2.34}$$

under a first order Itô-Taylor scheme.

Proof. By the definition of the alternative transform, we must have that

$$\begin{aligned} u_{i,j}(x) &= \mathfrak{F}^{-1}[\mathfrak{F}[\tilde{u}_{i+1}^{\alpha,\beta}(.,\alpha_{j,1})](\nu)\phi_i(\nu,x)](x) \\ &\quad -\mathbf{E}_{t_i}^x\left[\alpha(X_{t_{i+1}}^\pi)^2+\beta X_{t_{i+1}}^\pi\right] \\ &\quad +\Delta_i \mathbf{1}_{\{j>2\}}\sum_{k=2}^{j-1}\alpha_{jk}f(u_{i,k}(x),\dot{u}_{i,k}(x)). \end{aligned} \tag{5.2.35}$$

Notice that

$$\mathbf{E}_{t_i}^x\left[X_{t_{i+1}}^\pi\right] = x+\Delta_i F_{(0)}(x) \tag{5.2.36}$$

and

$$\begin{aligned} \mathbf{E}_{t_i}^x\left[(X_{t_{i+1}}^\pi)^2\right] &= \mathbf{E}_{t_i}^x\left[X_{t_{i+1}}^\pi\right]^2+\mathbf{Var}_{t_i}^x[X_{t_{i+1}}^\pi] \\ &= (x+\Delta_i F_{(0)}(x))^2 \\ &\quad +\mathbf{E}_{t_i}^x\left[\left(F_{(1)}(x)I_{t_i,t_{i+1}}^{(1)}+F_{(1,1)}(x)I_{t_i,t_{i+1}}^{(1,1)}\right)^2\right] \\ &= (x+\Delta_i F_{(0)}(x))^2 \\ &\quad +\Delta_i F_{(1)}^2(x)+\frac{1}{2}\Delta_i^2 F_{(1,1)}^2(x) \end{aligned} \tag{5.2.37}$$

knowing that $\mathbf{E}_{t_i}^x\left[\left(I_{t_i,t_{i+1}}^{(1)}\right)^2\right] = \Delta_i$, $\mathbf{E}_{t_i}^x\left[\left(I_{t_i,t_{i+1}}^{(1,1)}\right)^2\right] = \frac{1}{2}\Delta_i^2$ and $\mathbf{E}_{t_i}^x\left[I_{t_i,t_{i+1}}^{(1)}I_{t_i,t_{i+1}}^{(1,1)}\right] = 0$. Equations (5.2.35), (5.2.36) and (5.2.37) lead to the expression for $u_{i,j}$ in equation (5.2.33).

The definition of the alternative transform also requires

$$\begin{aligned} \dot{u}_{i,j}(x) &= F_{(1)}(x)\mathfrak{F}^{-1}[\mathfrak{F}[\tilde{u}_{i+1}^{\alpha,\beta}(.,\beta_{j,1})](\nu)\mathbf{i}\nu(1+\zeta_i(\nu,x))\phi_1(\nu,x)](x) \\ &\quad -\mathbf{E}_{t_i}^x\left[H_{t_{i,j},\gamma_j\Delta_i}^{\varphi_j}(\alpha(X_{t_{i+1}}^\pi)^2+\beta X_{t_{i+1}}^\pi)\right] \\ &= F_{(1)}(x)\mathfrak{F}^{-1}[\mathfrak{F}[\tilde{u}_{i+1}^{\alpha,\beta}(.,\beta_{j,1})](\nu)\mathbf{i}\nu(1+\zeta_i(\nu,x))\phi_1(\nu,x)](x) \\ &\quad -F_{(1)}(x)\mathbf{E}_{t_i}^x\left[2\alpha(X_{t_{i+1}}^\pi)+\beta\right] \\ &\quad -F_{(1,1)}(x)\mathbf{E}_{t_i}^x\left[I_{t_i,t_{i+1}}^{(1)}\left(2\alpha(X_{t_{i+1}}^\pi)+\beta\right)\right] \\ &\quad \text{(using the duality formula),} \\ &= F_{(1)}(x)\mathfrak{F}^{-1}[\mathfrak{F}[\tilde{u}_{i+1}^{\alpha,\beta}(.,\beta_{j,1})](\nu)\mathbf{i}\nu(1+\zeta_i(\nu,x))\phi_1(\nu,x)](x) \\ &\quad -F_{(1)}(x)\left[2\alpha\left(x+\Delta_i F_{(0)}(x)+\Delta_i F_{(1,1)}(x)\right)+\beta\right] \end{aligned} \tag{5.2.38}$$

using the duality formula once again. □

One notices that when using the half and first order Itô-Taylor schemes, the Fourier representations do not depend on the scaling functions φ_j since the characteristic $\Phi_{i,j}$ do not depend on them. In general, the expression for the characteristic $\Phi_{i,j}$ in equations (5.2.19) and (5.2.20) tells that if the forward SDE time discretization is such that the Malliavin derivative $D_s X_{t_{i+1}}^\pi$ is independent of $s \in [t_i, t_{i+1})$ then the Fourier representations are independent of the scaling functions φ_j. This is due to the fact that the scaling function φ_j are at least in $\mathcal{B}^0$, i.e they integrate to 1.

5.3 Error analysis

We denote by $\{u_{i,j,k}\}_{k=0}^{N_i N}$ and $\{\dot{u}_{i,j,k}\}_{k=0}^{N_i N}$ the intermediate numerical solutions obtained at mesh time t_i, $i = 0, 1, ..., n-1$ and stage j, $1 < j \leq q+1$, from the procedure of Chapter 4 when using

a $q-$stage Runge-Kutta scheme. In addition, $\{\mathbf{u}_{i,j,k}\}_{k=0}^{N_iN}$ and $\{\dot{\mathbf{u}}_{i,j,k}\}_{k=0}^{N_iN}$ are the intermediate numerical solutions obtained at the intermediate stage j, $1 < j \leq q+1$, of the mesh time t_i given the exact solutions u_{i+1} and $\dot{u}_{i+1}$ at time t_{i+1}. We have from the notation previously used that the numerical solution at mesh time t_i is given by

$$u_{i,k} = u_{i,q+1,k} \tag{5.3.1}$$

$$\dot{u}_{i,k} = \dot{u}_{i,q+1,k} \tag{5.3.2}$$

and computed from the intermediate solutions $\{\tilde{u}_{i,k}\}_{k=0}^{N_iN}$, $0 < i \leq n$ where $\tilde{u}_{n,k} = \tilde{u}_n(x_{n,k})$. When the exact solutions u_{i+1} and $\dot{u}_{i+1}$ are known at t_{i+1}, we also write

$$\mathbf{u}_{i,k} = \mathbf{u}_{i,q+1,k} \tag{5.3.3}$$

$$\dot{\mathbf{u}}_{i,k} = \dot{\mathbf{u}}_{i,q+1,k}. \tag{5.3.4}$$

The local (space) discretization error has the form

$$E_{ik} := |u_i(x_k) - \mathbf{u}_{i,k}| + |\dot{u}_i(x_k) - \dot{\mathbf{u}}_{i,k}| \tag{5.3.5}$$

for $i = 0, 1, ..., n-1$ and $k = 0, 1, ..., N_iN$. The following assumptions prove crucial in the error analysis that focuses exclusively on the one-dimensional case $d = 1$ and explicit Runge-Kutta schemes even though the results can be generalized to implicit Runge-Kutta schemes when the condition of equation (5.2.1) is satisfied.

Assumption 5.3.1. *There are positive constants p_0, s_0 ,K_0 and $C_0 > 0$ such that*

$$\max(|\phi_i(\mathbf{i}s_0, x)|, |\phi_i(-\mathbf{i}s_0, x)|) \leq e^{K_0\Delta_i} \tag{5.3.6}$$

and, hence, the discrete version of the forward process has conditional exponential moments. In adddition,

$$\int_{\mathbb{R}^d} |\phi_i(\nu, x)|\, d\nu + \max_{1<j\leq q+1} \int_{\mathbb{R}^d} |\Phi_{i,j}(\nu, x)|\, d\nu \leq C_0\Delta_i^{-p_0}. \tag{5.3.7}$$

The next theorem gives a description of the local (space) discretization error bound where we assume that the time discretization is given.

Theorem 5.4. *Suppose that Assumptions 5.1.1, 5.1.2, 5.2.1 and 5.3.1 are satisfied. Then the Fourier interpolation method yields a local space discretization error of the form*

$$\sup_{i,k} E_{ik} = \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl}\right) \tag{5.3.8}$$

for some constant $K > 0$ on the alternative grid and under the trapezoidal quadrature rule for any explicit q-stage Runge-Kutta scheme.

Proof. Once again, we follow the steps in the proof of Theorem 3.3. The truncation error when computing the numerical solutions $\dot{\mathbf{u}}_{i,j,k}$ is

$$\begin{aligned}
&\mathbf{E}_{t_i}^{x_{ik}}\left[H_{t_{i,j},\gamma_j\Delta_i}^{\varphi_j}\tilde{u}_{i+1}(X_{t_{i+1}}^{\pi};\beta_{j,1})\mathbf{1}_{|\Delta X_i^{\pi}|>\frac{l}{2}}\right] \\
&< K\mathbf{E}_{t_i}^{x_{ik}}\left[\left|H_{t_{i,j},\gamma_j\Delta_i}^{\varphi_j}\right|\mathbf{1}_{|\Delta X_i^{\pi}|>\frac{l}{2}}\right] \\
&\quad \text{(by boundedness of the BSDE coefficients)}, \\
&< K\mathbf{E}_{t_i}^{x_{ik}}\left[\left(H_{t_{i,j},\gamma_j\Delta_i}^{\varphi_j}\right)^2\right]^{\frac{1}{2}}\mathbf{E}_{t_i}^{x_{ik}}\left[\mathbf{1}_{|\Delta X_i^{\pi}|>\frac{l}{2}}\right]^{\frac{1}{2}} \\
&\quad \text{(by the Cauchy-Schwartz inequality)} \\
&= K\mathbf{E}_{t_i}^{x_{ik}}\left[\int_{t_{i,j}}^{t_{i+1}}\left(\frac{1}{\gamma_j\Delta_i}\varphi_j\left(\frac{s-t_{i,j}}{\gamma_j\Delta_i}\right)\right)^2 ds\right]^{\frac{1}{2}}\mathbf{E}_{t_i}^{x_{ik}}\left[\mathbf{1}_{|\Delta X_i^{\pi}|>\frac{l}{2}}\right]^{\frac{1}{2}}
\end{aligned}$$

$$
\begin{aligned}
& \text{(by Itô isometry)} \\
< \quad & K\Delta_i^{-\frac{1}{2}} \mathbf{E}_{t_i}^{x_{ik}} \left[\mathbf{1}_{|\Delta X_i^\pi| > \frac{l}{2}} \right]^{\frac{1}{2}} \\
& \text{(by the boundedness of functions } \varphi_j\text{)}, \\
\leq \quad & K\Delta_i^{-\frac{1}{2}} \left(\inf_{s>0} e^{-s\frac{l}{2}} \phi_i(-\mathbf{i}s) + \inf_{s>0} e^{-s\frac{l}{2}} \phi_i(\mathbf{i}s) \right)^{\frac{1}{2}} \\
& \text{(by Chernoff's inequality)} \\
< \quad & K\Delta_i^{-\frac{1}{2}} e^{-s_0 \frac{l}{4} + \frac{1}{2} K_0 \Delta_i} \\
< \quad & K e^{-s_0 \frac{l}{4}}
\end{aligned}
$$

using Assumption 5.2.1. The Fourier interpolation leads to a first order space discretization error by the twice differentiability of the driver f and the terminal condition g when computing the numerical solutions $\dot{\mathbf{u}}_{i,j,k}$.

The same statements hold for the numerical solutions $\mathbf{u}_{i,2,k}$ using identical arguments. By recursion and using the Lipschitz property of the driver f, the statements hold for $\mathbf{u}_{i,j,k}$, $1 < j \leq q+1$. Since the mesh time t_i and the space node x_{ik} are arbitrary, the space truncation and discretization error bounds hold for any i and k. □

Locally, the truncation error remains spectral. Nonetheless, it is just of index 1 in this general setting where the conditional characteristic function ϕ_i is unspecified. In Chapter 3 and 4, the quadratic exponential form of the characteristic function is the main reason for the spectral convergence of index 2 in the truncation error. The space discretization error though is unchanged with first order due to the second order differentiability of the BSDE coefficients. Indeed, the Fourier interpolation produces a space discretization error with a higher order when the driver f and the terminal function g have the required smoothness. We already illustrated the phenomenon in the numerical results of Section 3.5.1. In general, if $f \in \mathcal{C}_b^{m+1}$ and $g \in \mathcal{C}_b^{m+1}$, we can expect a space discretization error of order m which is the convergence order of the underlying Fourier interpolation.

We now turn to the global space discretization error $E_{l,\Delta x}$ as defined in equation (3.2.6). The next theorem gives its error bound.

Theorem 5.5. *Suppose the conditions of Theorem 5.4 are satisfied. If the discretization is such that*

$$
\sup_i \left\{ \frac{C_0 \Delta x}{\pi \Delta_i^{p_0}} \right\} \leq 1 \tag{5.3.9}
$$

then the Fourier interpolation method is stable and yields a global discretization error $E_{l,\Delta x}$ of the form

$$
E_{l,\Delta x} = \mathcal{O}(\Delta x) + \mathcal{O}\left(e^{-Kl}\right) \tag{5.3.10}
$$

where $K > 0$ for any explicit q-stage Runge-Kutta scheme.

Proof. From the definition of the global space discretization error, we may write

$$
\begin{aligned}
e_{ik} \quad & \leq \quad E_{n-i,k} + |\mathbf{u}_{n-i,k} - u_{n-i,k}| && (5.3.11) \\
\dot{e}_{ik} \quad & \leq \quad E_{n-i,k} + |\dot{\mathbf{u}}_{n-i,k} - \dot{u}_{n-i,k}| \,. && (5.3.12)
\end{aligned}
$$

Let's assume the boundary values of the function $\tilde{u}_{i+1}$ and the sequence $\tilde{u}_{i+1,s}$ are matched on the alternative grid so that we don't have to treat the alternative transform. Under an explicit $q-$stage Runge-Kutta scheme, we have

$$
|\dot{\mathbf{u}}_{i,j,k} - \dot{u}_{i,j,k}| \quad = \quad \left| \mathfrak{D}^{-1} \left[\{ \Phi_{i,j}(\nu_{i+1,m}, x_{ik}) \mathbb{D}[\tilde{u}_{i+1} - \tilde{u}_{i+1,s}]_m \}_{m=0}^{N_{i+1}N-1} \right]_{k+\frac{N}{2}} \right|
$$

$$
\begin{aligned}
&\leq \frac{\sum_{m=0}^{N_{i+1}N-1}|\Phi_{i,j}(\nu_{i+1,m},x_{ik})|}{N_{i+1}N}\sup_k|\tilde{u}_{i+1}(x_{ik},\beta_{1,j})-\tilde{u}_{i+1,k}|\\
&\leq \frac{\Delta x}{2\pi}\left(\int_{\mathbb{R}^d}|\Phi_{i,j}(\nu,x_{i,k})|\,d\nu\right)\sup_k|\tilde{u}_{i+1}(x_{ik},\beta_{1,j})-\tilde{u}_{i+1,k}|\\
&\leq \frac{C_0\Delta x}{2\pi\Delta_i^{p_0}}\sup_k|\tilde{u}_{i+1}(x_{ik},\beta_{1,j})-\tilde{u}_{i+1,k}|\\
&\quad \text{(using Assumption 5.3.1)},\\
&\leq \frac{C_0\Delta x}{2\pi\Delta_i^{p_0}}(1+\Delta_iK)\sup_k e_{n-i-1,k}\\
&\quad +\frac{C_0\Delta x}{2\pi\Delta_i^{p_0}}\Delta_iK\sup_k\dot{e}_{n-i-1,k}\\
&\quad \text{(since } f \text{ is Lipschitz and } \beta_{1,j} \text{ is bounded)},\\
&\leq \frac{C_0\Delta x}{2\pi\Delta_i^{p_0}}(1+\Delta_iK)\sup_k e_{n-i-1,k}\\
&\quad +\frac{C_0\Delta x}{2\pi\Delta_i^{p_0}}(1+\Delta_iK)\sup_k\dot{e}_{n-i-1,k}. \qquad (5.3.13)
\end{aligned}
$$

Similarly, we get

$$
\begin{aligned}
|\mathbf{u}_{i,2,k}-u_{i,2,k}| &\leq \left|\mathfrak{D}^{-1}\left[\{\phi_i(\nu_{i+1,m},x_{ik})\mathbb{D}[\tilde{u}_{i+1}-\tilde{u}_{i+1,s}]_m\}_{m=0}^{N_{i+1}N-1}\right]_{k+\frac{N}{2}}\right|\\
&\leq \frac{\Delta x}{2\pi}\left(\int_{\mathbb{R}^d}|\phi_i(\nu,x_{i,k})|\,d\nu\right)\sup_k|\tilde{u}_{i+1}(x_{ik},\alpha_{1,2})-\tilde{u}_{i+1,k}|\\
&\leq \frac{C_0\Delta x}{2\pi\Delta_i^{p_0}}\sup_k|\tilde{u}_{i+1}(x_{ik},\alpha_{1,2})-\tilde{u}_{i+1,k}|\\
&\quad \text{(using Assumption 5.3.1)},\\
&\leq \frac{C_0\Delta x}{2\pi\Delta_i^{p_0}}(1+\Delta_iK)\sup_k e_{n-i-1,k}\\
&\quad +\frac{C_0\Delta x}{2\pi\Delta_i^{p_0}}(1+\Delta_iK)\sup_k\dot{e}_{n-i-1,k}
\end{aligned}
$$

so that we get

$$
\begin{aligned}
|\mathbf{u}_{i,j,k}-u_{i,j,k}| &\leq \frac{C_0\Delta x}{2\pi\Delta_i^{p_0}}(1+\Delta_iK)\sup_k e_{n-i-1,k}\\
&\quad +\frac{C_0\Delta x}{\pi\Delta_i^{p_0}}(1+\Delta_iK)\sup_k\dot{e}_{n-i-1,k} \qquad (5.3.14)
\end{aligned}
$$

recursively for $1<j\leq q+1$ using the Lipschitz property of the driver f and the boundedness of the Runge-Kutta coefficients. Equations (5.3.11) and (5.3.12) combined with equations (5.3.14) and (5.3.13) lead to

$$
\begin{aligned}
\sup_k e_{i,k}+\sup_k\dot{e}_{i,k} &\leq 2\sup_{i,k}E_{ik}\\
&\quad +\frac{C_0\Delta x}{\pi\Delta_i^{p_0}}(1+\Delta_{n-i}K)\left(\sup_k e_{i-1,k}+\sup_k\dot{e}_{i-1,k}\right)\\
&\leq 2\sup_{i,k}E_{ik}+\zeta(1+\Delta_{n-i}K)\left(\sup_k e_{i-1,k}+\sup_k\dot{e}_{i-1,k}\right)
\end{aligned}
$$

where

$$
\sup_i\left\{\frac{C_0\Delta x}{\pi\Delta_i^{p_0}}\right\}\leq\zeta\leq 1.
$$

Gronwall's Lemma then yields

$$
\sup_k e_{i,k}+\sup_k\dot{e}_{i,k}\leq 2e^{TK}\sup_{i,k}E_{ik} \qquad (5.3.15)
$$

so that the scheme is stable. The result of equation (5.3.10) follows by taking the supremum on the left hand side of equation (5.3.15) other time step and applying Theorem 5.4. □

In this general case, the global discretization error maintains the structure of the local discretization error under a stability condition. Equation (5.3.9) indicates that the space discretization has to be relatively as fine as the time discretization to ensure stability. Hence, stability can always be reached for any time discretization by refining the space discretization. However, the structure of the characteristic functions ϕ_i and Φ_{ij} determines the relative refinement needed for the space discretization.

The simulation of FBSDEs is not different from Chapter 3 and 4. Letting $(U_i, \dot{U}_i)$ be the extended solution at time mesh t_i of equations (3.3.2) and (3.3.3), we define the approximate processes (x, y, z) as

$$x_t = \sum_{i=0}^{n-1} X^\pi_{t_i} \mathbf{1}_{[t_i, t_{i+1})}(t) \tag{5.3.16}$$

$$y_t = \sum_{i=0}^{n-1} U_i(X^\pi_{t_i}) \mathbf{1}_{[t_i, t_{i+1})}(t) \tag{5.3.17}$$

$$z_t = \sum_{i=0}^{n-1} \dot{U}_i(X^\pi_{t_i}) \mathbf{1}_{[t_i, t_{i+1})}(t) \tag{5.3.18}$$

for $t \in [0, T)$. The boundedness of backward and control process solutions $\{y_t\}_{t\in[0,T)}$ and $\{z_t\}_{t\in[0,T)}$ was established in Chapter 3 (Corollary 3.8) and holds in the Runge-Kutta framework. We will instead focus on the simulation error $\mathcal{E}_{\pi,l,\Delta x}$ defined as

$$\mathcal{E}^2_{\pi,l,\Delta x} := \max_{0\le i<n} \|X_{t_i} - x_{t_i}\|^2_{L^2} + \max_{0\le i<n} \|Y_{t_i} - y_{t_i}\|^2_{L^2} + \sum_{i=0}^{n-1} \Delta_i \|Z_{t_i} - z_{t_i}\|^2_{L^2}. \tag{5.3.19}$$

The next theorem describes the error bound.

Theorem 5.6. *Suppose that Assumptions 5.1.1, 5.1.2, 5.2.1 and 5.3.1 are satisfied. If the stability and convergence condition of equation (5.3.9) holds and both the forward discretization and q-stage Runge-Kutta scheme are of order $m > 0$ then*

$$\mathcal{E}^2_{\pi,l,\Delta x} = \mathcal{O}(|\pi|^{2m}) + \mathcal{O}(\Delta x^2) + \mathcal{O}\left(e^{-C(N_0+1)l} + e^{-Cl}\right) \tag{5.3.20}$$

where $C > 0$.

Proof. The proof is essentially similar to the proof of Theorem 3.9. Since we assume that both the forward and the backward schemes are of order $m > 0$ the time discretization error $\mathcal{O}(|\pi|^{2m})$ is obviously of order m. Also the space discretization error $\mathcal{O}(\Delta x^2)$ and the space truncation error $\mathcal{O}(e^{-Cl})$ follow from Theorem 5.5 and the quadratic nature of the simulation error.

Since the driver f and the terminal function g are bounded, the remaining error term is related to

$$\begin{aligned} \max_{0<i<n} \mathbf{E}\left[\mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(X^\pi_{t_i})\right] \le & \max_{0<i<n} \inf_{s>0} e^{-s\left(X_0 + N_i \frac{l}{2}\right)} \phi_{t_i}(-\mathbf{i}s) \\ & + \max_{0<i<n} \inf_{s>0} e^{s\left(X_0 - N_i \frac{l}{2}\right)} \phi_{t_i}(\mathbf{i}s) \end{aligned}$$

by Chernoff's inequality where ϕ_{t_i} is the characteristic function of $X^\pi_{t_i}$. It is shown by recursion and using Assumption 5.3.1 that

$$\begin{aligned} \phi_{t_i}(-\mathbf{i}s_0) &\le e^{K_0 T + s_0 X_0} \\ \phi_{t_i}(\mathbf{i}s_0) &\le e^{K_0 T - s_0 X_0}. \end{aligned}$$

Hence,

$$\begin{aligned}\max_{0<i<n} \mathbf{E}\left[\mathbf{1}_{\mathbb{R}\setminus\mathcal{I}_i}(X^\pi_{t_i})\right] &\leq 2e^{K_0 T} \max_{0<i<n} e^{-s_0 N_i \frac{l}{2}} \\ &\leq 2e^{K_0 T} e^{-s_0(N_0+1)\frac{l}{2}}.\end{aligned}$$

□

A simple extension of q-stage Runge-Kutta schemes to reflected FBSDE consists in applying the reflection at the last stage and for all time nodes t_i. Let's consider the system

$$\begin{cases} dX_t = a(t,X_t)dt + \sigma(t,X_t)dW_t \\ -dY_t = f(t,X_t,Y_t,Z_t,)dt + dA_t - Z_t dW_t \\ Y_t \geq B_t \ , \ dA_t \geq 0 \ , \ \forall t \in [0,T] \\ \int_0^T (Y_t - B_t)dA_t = 0 \\ X_0 = x_0 \ , \ Y_T = g(X_T) \end{cases} \tag{5.3.21}$$

where

$$B_t = B(t,X_t). \tag{5.3.22}$$

The numerical solution of a $q-$stage Runge-Kutta scheme can be defined as

$$u_i(x) = u_{i,q+1}(x) + \Delta\bar{u}_i(x) \tag{5.3.23}$$

$$\dot{u}_i(x) = \dot{u}_{i,q+1}(x) \tag{5.3.24}$$

$$\Delta\bar{u}_i(x) = [u_{i,q+1}(x) - B(t_i,x)]^- \tag{5.3.25}$$

instead of equations (5.2.5) and (5.2.6). The intermediate solutions u_{ij} and $\dot{u}_{ij}$, $1 < j \leq q+1$, are defined as previously and their numerical values may be given by the Fourier method on the alternative grid. From there, the simulation a numerical solution (x,y,z,a) for the reflected FBSDE is conducted through the equations

$$x_t = \sum_{i=0}^{n-1} X^\pi_{t_i} \mathbf{1}_{[t_i,t_{i+1})}(t) \tag{5.3.26}$$

$$y_t = \sum_{i=0}^{n-1} U_i(X^\pi_{t_i}) \mathbf{1}_{[t_i,t_{i+1})}(t) \tag{5.3.27}$$

$$z_t = \sum_{i=0}^{n-1} \dot{U}_i(X^\pi_{t_i}) \mathbf{1}_{[t_i,t_{i+1})}(t) \tag{5.3.28}$$

$$a_t = \sum_{i=0}^{n-1} \Delta\bar{U}_i(X^\pi_{t_i}) \mathbf{1}_{[t_i,T)}(t) \tag{5.3.29}$$

where the extended solutions U_i, $\dot{U}_i$ and $\Delta\bar{U}_i$ are as in equations (3.3.2), (3.3.3) and (3.4.7) respectively.

5.4 Application to commodity derivatives

We test the convergence properties of the Fourier interpolation method on Runge-Kutta schemes with a problem of commodity derivative pricing under a model proposed by Lucia and Schwartz [77]. We shall test the method's convergence and behavior on smooth and unbounded FBSDE coefficients. The non-smoothness of BSDE coefficients affects only the space discretization error and was already studied in Chapter 3. Also, the unbounded coefficient framework includes

the bounded coefficient case so that the ideas developed here also hold for bounded coefficient FBSDEs.

The commodity spot price X is defined by

$$X_t = e^{S(t)+V_t} \tag{5.4.1}$$

where the deterministic function $S : \mathbb{R}^+ \to \mathbb{R}$ represents the seasonality component of the commodity and V is the price diffusion following an Ornstein-Uhlenbeck process according to the Vasicek [111] model

$$dV_t = -\kappa V_t dt + \sigma dW_t. \tag{5.4.2}$$

As indicated by Lucia and Schwartz [77], the commodity spot price X satisfies the stochastic differential equation

$$dX_t = \kappa(\theta(t) - \ln X_t)X_t dt + \sigma X_t dW_t \tag{5.4.3}$$

where

$$\theta(t) = \frac{1}{\kappa}\left(\frac{\sigma^2}{2} + \frac{dS}{dt}(t)\right) + S(t). \tag{5.4.4}$$

We consider the commodity price as our forward process through equation (5.4.3).

When the risk free rate r and the market price of risk λ are both constant, the forward (or future) price $F_{t,T} := Y_t = u(t, X_t)$ with maturity $T > 0$ at time $t < T$ is given by

$$\begin{aligned} Y_t &= \mathbf{E}_t^{\mathbf{Q}}[X_T] \\ &= e^{S(T)+(\ln X_t - S(t))e^{-\kappa(T-t)} - \frac{\sigma\lambda}{\kappa}h(T-t,\kappa) + \frac{\sigma^2}{4\kappa}h(T-t,2\kappa)} \end{aligned} \tag{5.4.5}$$

with

$$h(\tau, \kappa) = 1 - e^{-\kappa\tau} \tag{5.4.6}$$

where the expectation is taken under the equivalent risk measure $\mathbf{Q}$. It can be shown that the forward price solves a BSDE with linear driver

$$f(t, y, z) = -\lambda z \tag{5.4.7}$$

and terminal condition

$$g(x) = x. \tag{5.4.8}$$

Options on forward contracts can also be represented in form of BSDEs in this spot price model but we limit our analysis to forward price estimation. From equation (5.4.5) the control process (or equivalently the forward price delta) is given by

$$\begin{aligned} Z_t &= \sigma X_t \nabla u(t, X_t) \\ &= \sigma e^{-\kappa(T-t)} u(t, X_t). \end{aligned} \tag{5.4.9}$$

The adjustment speed of the diffusion process is $\kappa = 1.5$ and the volatility of the diffusion is set to be $\sigma = 0.065$. The seasonality component is given by

$$S(t) = \ln \bar{P} + 0.05 \sin(2\pi t) \tag{5.4.10}$$

and the initial spot price by

$$X_0 = \bar{P}e^{V_0} = 0.95\bar{P} \tag{5.4.11}$$

where we normalize the real value[3] of the commodity $\bar{P} = 1$. Also, the maturity of the forward contract is $T = 0.25$ and we suppose a market price of risk of $\lambda = 0.25$.

[3] The real value $\bar{P}$ can be considered as the production cost (per unit) of the commodity.

The FBSDE is solved on an alternative grid centered at X_0 with a uniform time mesh. For a given number of time steps n and the initial number $N_0 = 1$ of intervals, the length of an increment interval is set as

$$l = \frac{1.8}{N_0 + n} \tag{5.4.12}$$

so that the truncated interval at time t_n has length 1.8. This restriction keeps the space nodes in the upper half plane knowing that the commodity price is a positive process. Moreover, the number of space steps on an increment interval is $N = 2$.

We numerically solve the BSDE with the explicit $1-$stage Runge-Kutta scheme of half order and an explicit $2-$stage Runge-Kutta scheme of first order. Under the explicit $1-$stage scheme, the commodity price is discretized with an Euler scheme whereas a Milstein scheme is used for the forward process X under the explicit $2-$stage Runge-Kutta scheme. Note that for the Milstein scheme,

$$F_{(1,1)}(x) = \sigma^2 x \tag{5.4.13}$$

and

$$I^{(1,1)}_{t_i,t_{i+1}} = \frac{1}{2}(\Delta W_i^2 - \Delta_i) \tag{5.4.14}$$

and we estimate the characteristic function with a Gaussian characteristic

$$\phi_i(\nu, x) = \exp\left\{\Delta_i \left(\mathbf{i}F_{(0)}(x)\nu - \frac{1}{2}\left(F^2_{(1)}(x) + \frac{1}{2}\Delta_i F^2_{(1,1)}(x)\right)\nu^2\right)\right\}. \tag{5.4.15}$$

In addition, we use an explicit $2-$stage Runge-Kutta scheme with tableau

0	0	0	0	0	0
$\frac{2}{3}$	$\frac{2}{3}$	0	0	$\frac{2}{3}$	0
1	$\frac{1}{4}$	$\frac{3}{4}$	0	1	0

Under both FBSDE discretizations, we compute two different types of error. The first error E_{True} evaluates the maximal absolute error of the numerical solution with respect to the true solution

$$\begin{aligned} E_{True} &= \max_{0\le i<n} \max_{0\le k\le NN_i} |u(t_i, x_{ik}) - u_{ik}| \\ &+ \max_{0\le i<n} \max_{0\le k\le NN_i} |\dot{u}(t_i, x_{ik}) - \dot{u}_{ik}| \end{aligned} \tag{5.4.16}$$

where

$$\dot{u}(t,x) = \sigma x \nabla u(t,x) = \sigma e^{-\kappa(T-t)} u(t,x). \tag{5.4.17}$$

The second error E_{Sim} approximates the simulation error $\mathcal{E}_{\pi,l,\Delta x}$. Given the numerical solution $\{X^\pi_{t_i,j}\}_{j=1}^m$, $i = 0, 1, ..., n-1$ with $m > 0$ simulated paths for the forward process, we compute the numerical solution $\{(y_{t_i,j}, z_{t_i,j})\}_{j=1}^m$ of the backward process through equations (5.3.17) and (5.3.18). The error E_{Sim} hence writes as

$$\begin{aligned} E_{Sim} &= \frac{1}{m}\sum_{j=1}^m \max_{0\le i<n} \left|u(t_i, X^\pi_{t_i,j}) - y_{t_i,j}\right| \\ &+ \frac{1}{m}\sum_{j=1}^m \left(\sum_{i=0}^{n-1} \Delta_i (\dot{u}(t_i, X^\pi_{t_i,j}) - z_{t_i,j})^2\right)^{\frac{1}{2}}. \end{aligned} \tag{5.4.18}$$

We systematically use $m = 1000$ paths. Even if the errors E_{True} and E_{Sim} may be of the same order, they are interpreted differently. The error E_{True} gives the behavior of the maximal approximation error on the grid whereas E_{Sim} gives the behavior of the error on the relevant part of grid when solving the FBSDE numerically. Figure 5.4.1 displays the errors under the

explicit 1−stage Runge-Kutta scheme with $n \in \{5, 10, 20, 50, 100\}$ and Figure 5.4.2 shows the errors under the explicit 2−stage scheme.

The error graphs of Figures 5.4.1 and 5.4.2 look almost identical and confirm that the 2−stage scheme is of first order and the 1−stage scheme of (at least) half order. The extra-efficiency of the 1−stage scheme may be attributed in this particular case to the simplicity of the driver f and the terminal condition g.

In Figure 5.4.3, we present the absolute errors along the simulated paths for the BSDE solution. One notices that the maximal errors occur at the initial time $t_0 = 0$ for the forward price (Y_t) and at maturity $T = 0.25$ for the control process (Z_t). Nonetheless, the simulation errors are of the same order (10^{-4}) for both processes. This information is confirmed by the contour plot of Figure 5.4.4 not only along the simulated paths but on the entire grid.

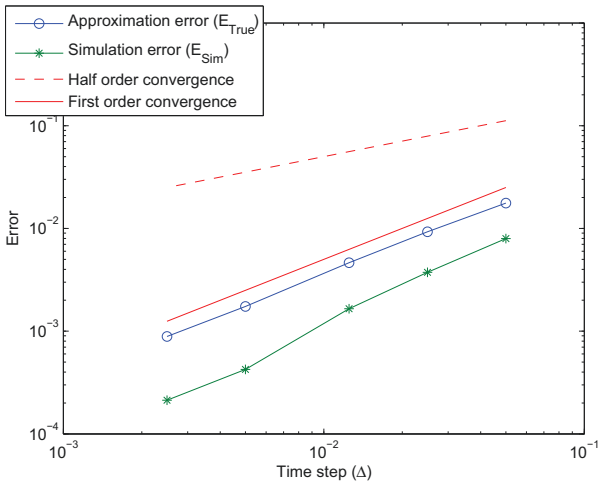
Moreover, the contour plot gives indication on the source of errors. Indeed, Figure 5.4.4 shows that the maximal errors mainly occur for the upper space node values on the alternative grid and they decrease for lower space node values. This is due to the unbounded nature of the spot price process coefficients. Since the volatility of the spot price is a positive and increasing function of the spot price[4], higher spot price values lead to higher local volatility. Hence, the fixed length of increment interval l may not be sufficiently large to ensure accuracy for higher space node values. In general, the phenomenon is amplified with the magnitude of the forward process coefficients as illustrated in the contour plot of Figure 5.4.5 where we choose a higher value for the volatility σ and keep the other parameters unchanged. Similar results can be obtained by selecting a higher value for the speed of adjustment κ as shown in Figure 5.4.6 .

We end this chapter with an efficiency study of our schemes. Using the parameters initially given, the BSDE is solved on a uniform time grid with $n \in \{10, 20, 40, 50, 60, 80, 100\}$ time steps and $N \in \{2, 2^2, 2^3, 2^4\}$ space steps and value the computation time. Figure 5.4.7 displays the results. First note that since the Fourier interpolation method of Chapter 4 performs matrix multiplications, it is much slower than the convolution method of Chapter 3. As a comparison, the convolution method runs in less than half a second of CPU time on the grids considered in Figure 5.4.7 and using the same computing device.

As shown in Figure 5.4.7, the computation time of Fourier interpolation method increases with the number of time steps leading to a tradeoff between computation speed and accuracy. The exponential nature of the curves suggests that preference has to be given to the coarsest time discretization providing a satisfactory level of accuracy. Similarly, the computation time also increase drastically with the number N of space steps. Coarse space grid insuring accuracy are hence also preferable. Since a total number of $2q$ conditional expectations are computed under a q-stage Runge-Kutta scheme, we can expect the 1-stage scheme to run twice as fast as the 2-stage scheme. This is confirmed on Figure 5.4.7, especially when looking at the computation times for $n = 100$.

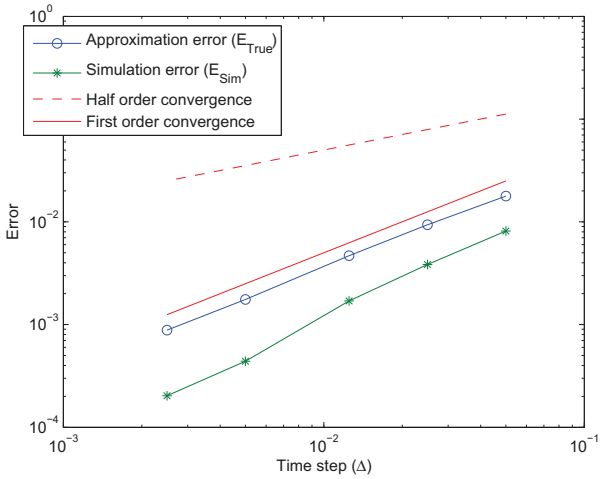
[4]See equation (5.4.3).

Figure 5.4.1: Log-log plot of errors using the 1-stage Runge-Kutta scheme.



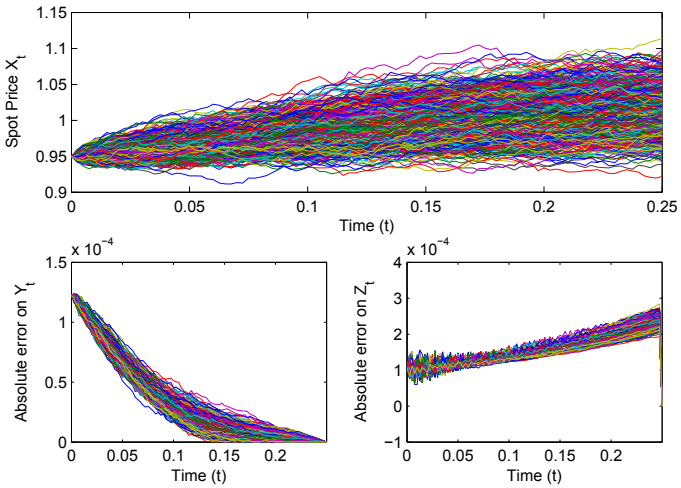
The sample standard deviation of the error E_{Sim} was less than 2×10^{-6} for all time discretizations.

Figure 5.4.2: Log-log plot of errors using the 2-stage Runge-Kutta scheme.



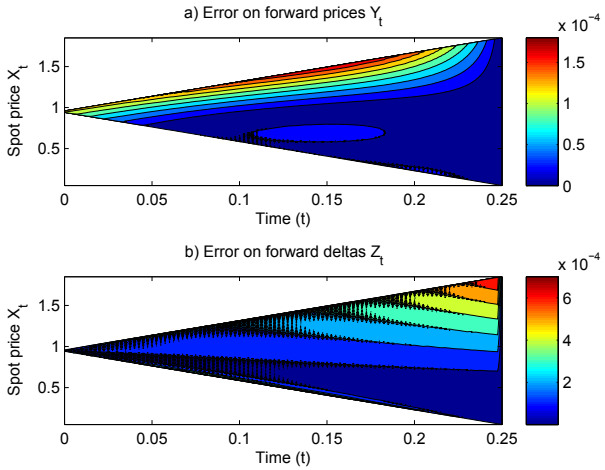
The sample standard deviation of the error E_{Sim} was less than 2×10^{-6} for all time discretizations.

Figure 5.4.3: Simulation errors using the 2-stage Runge-Kutta scheme.



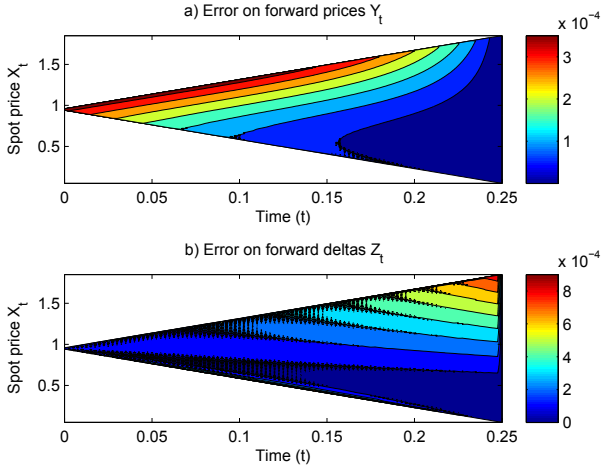
The numerical solution is obtained on a time mesh with $n = 100$ time steps and returns an forward price of 1.0121 and initial value of 0.0453 for the control process. The exact values are 1.0123 and 0.0452 respectively.

Figure 5.4.4: Contour plot of errors using the 2-stage Runge-Kutta scheme.



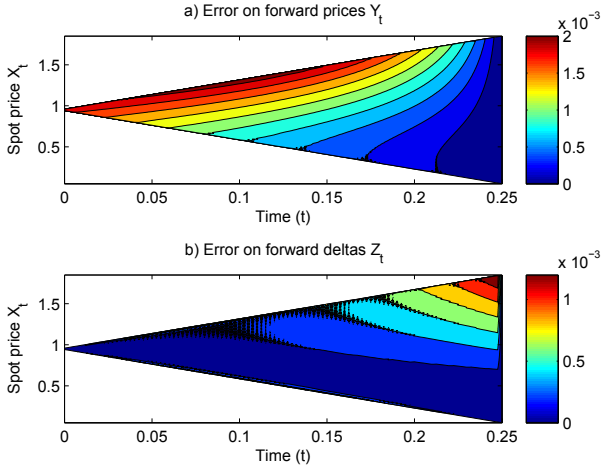
The numerical solution is obtained on a time mesh with $n = 100$ time steps and returns an forward price of 1.0121 and initial value of 0.0453 for the control process. The exact values are 1.0123 and 0.0452 respectively.

Figure 5.4.5: Errors using the 2-stage Runge-Kutta scheme with $\sigma = 0.08$.



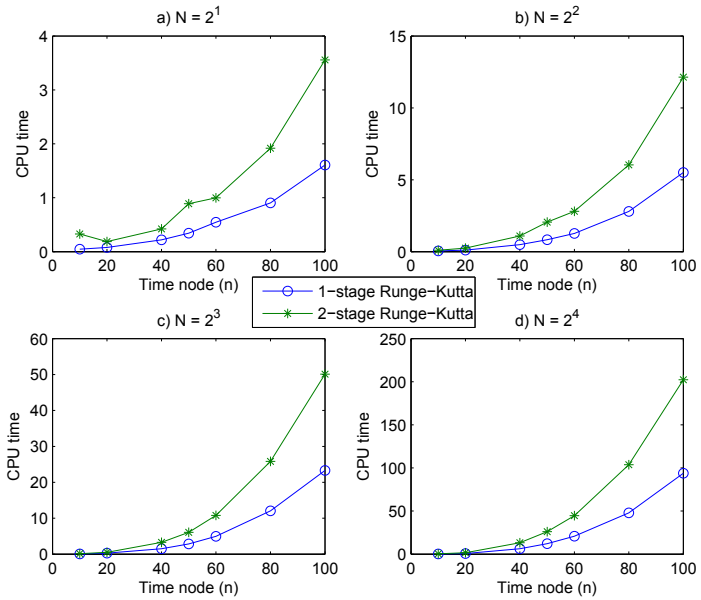
The numerical solution is obtained on a time mesh with $n = 100$ time steps and returns an forward price of 1.0115 and initial value of 0.0558 for the control process. The exact values are 1.0119 and 0.0556 respectively.

Figure 5.4.6: Errors using the 2-stage Runge-Kutta scheme with $\kappa = 3$.



The numerical solution is obtained on a time mesh with $n = 100$ time steps and returns an forward price of 1.0238 and initial value of 0.0316 for the control process. The exact values are 1.0257 and 0.0315 respectively.

Figure 5.4.7: CPU time (in seconds) of Runge-Kutta schemes.



Conclusion

The present thesis investigated the application of Fourier methods to numerical solutions of FBSDEs. In general, the method proposed consists of expressing the solution from a BSDE time discretization in terms of Fourier integrals using some available characteristic functions. The Fourier integrals are then discretized over a uniform space grid. In particular, the alternative grid of Chapter 3 produces a consistent, stable and globally convergent FFT based method for BSDEs using the Euler time discretization.

The results of Chapter 3 were extended to the FBSDE case and, in this framework, boundedness conditions are necessary on the forward process coefficients to ensure consistency, stability and convergence. Even if the method is still based on Fourier analysis, a matrix multiplication is need in the FBSDE case since the increment of the forward process are not necessarily independent. The matrix multiplication may lead to efficiency problems especially on fine time and space grids.

While the Euler scheme constitutes the primary time discretization for BSDEs used in this thesis, we focus on higher order Runge-Kutta schemes in Chapter 5. In this general framework, we explicitly define the characteristic functions and perform the error analysis accordingly. Under some integrability conditions on the characteristics, the Fourier interpolation based method is consistent, stable and globally convergent.

We mainly illustrate the convergence and efficiency properties of the Fourier method with derivative pricing examples from mathematical finance. Option pricing problems under the Black and Scholes model are considered in Chapter 2 and Chapter 3. The numerical example of Chapter 5 deals with a commodity modeling problem.

The thesis proposes a numerical implementation of the method only in the one-dimensional case. Hence, the extension of the method to the multidimensional framework is of importance. The definition of an alternative transform to mimic periodicity seems to be the only requirement for convergence in the multidimensional case. However, efficiency may be problematic especially when the method is applied to FBSDEs.

It may also be interesting to investigate alternative basis functions. Indeed, Fourier basis functions have well known disadvantages such as their lack of localization or their non-causality. As an alternative to Fourier basis functions, wavelets can be used and produce efficient algorithms in the one-dimensional and the multidimensional framework for the BSDE and FBSDE cases.

Since Runge-Kutta methods for BSDEs are quite recent, the problem of their implementation offers many research opportunities. An interesting area of research could be the Fourier representation of BSDE solutions under higher order time discretization for the forward process. Also, Monte-Carlo and spatial discretization based methods can be extended to these schemes with relative ease. Finally, (higher order) time discretizations for FBSDEs with non-Lipschitz coefficients are still an open problem.

Appendix

A Elements of FBSDE theory

Forward backward stochastic differential equations (FBSDEs) are quite recent mathematical objects. Many interesting results have been proved concerning their existence, uniqueness and properties. Some of those properties are particularly important for numerical simulation.

This appendix serves as an introduction to the theory of FBSDEs with Lipschitz coefficients. We cover the existence and uniqueness results in Section A.1 and present major properties of FBSDEs in Section A.2.

A.1 Classification of FBSDEs

From the well-known classical theory of stochastic differential equation (SDE), see Kloeden and Platen [69] or Øksendal [94], we consider a forward process satisfying

$$\begin{cases} dX_t = a(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 = x_0 \end{cases} \tag{A.1}$$

where $a : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ is the forward drift coefficient, $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is the forward diffusion (or volatility) coefficient. Both coefficients are assumed to be deterministic[12] functions. The $\mathcal{F}_0$-measurable random variable $x_0 \in \mathbb{R}^d$ defines the initial condition and is assumed to be constant. The SDE is interpreted as the integral equation

$$X_t = x_0 + \int_0^t a(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \tag{A.2}$$

with $t \in [0,T]$. In addition, we assume that both coefficients a and σ along with the initial condition x_0 satisfy the regularity conditions listed below.

Assumption A.1. *(on the forward SDE coefficients).*

H1. The coefficients a and σ are uniformly Lipschitz continuous in the space variable. Thus, there exists $K > 0$ so that

$$|a(t,x) - a(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq K\,|x - y|$$

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

H2. The coefficients a and σ have linear growth. Thus, there exists $K > 0$ such that

$$|a(t,x)| + |\sigma(t,x)| \leq K(1 + |x|)$$

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

[12] It is indeed possible to consider stochastic coefficients.

For stochastic coefficients, the inequalities of assumptions (H1) and (H2) hold almost surely and the coefficients have to satisfy additional measurability and square integrability conditions. Explicitly, the functional $a : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ are measurable with respect to $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$ (where $\mathcal{P}$ is a σ-algebra on $\Omega \times [0,T]$) and square integrable with respect to $d\mathbf{P} \times dt$, with

$$\mathbf{E}\left[\int_0^t \left\{|a(t,0)|^2 + |\sigma(t,0)|^2\right\} dt\right] < \infty.$$

The solution of the forward SDE of equation (A.1) is an adapted and continuous process $\{X_t\}_{t\in[0,T]} \in L_S^2$. The following proposition, due to Itô and proved in Kloeden and Platen [69], states the forward SDE well-posedness.

Proposition A.1. *Under assumptions (H1) and (H2), the forward stochastic differential equation given by equation (A.1) admits a unique (strong) solution* $X = \{X_t\}_{t\in[0,T]}$.

Instead of an initial condition, we may consider an equation evolving backward in time from a terminal condition formally given by

$$\begin{cases} -dY_t = f(t, Y_t, Z_t)dt - Z_t^* dW_t \\ Y_T = \xi. \end{cases} \tag{A.3}$$

Equation (A.3) is called a backward stochastic differential equation (BSDE). The corresponding integral equation takes the form

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s^* dW_s, \tag{A.4}$$

with $t \in [0,T]$. We shall suppose, for simplicity, that the backward process $\{Y_t\}_{t\in[0,T]}$ is unidimensional. The (deterministic) function

$$f : [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$$

stands for the backward process driver (or generator), the $\mathcal{F}_T$-measurable random variable ξ defines the terminal condition, and the adapted process $\{Z_t\}_{t\in[0,T]}$ taking values in $\mathbb{R}^d$ is called the control process.

In order to ensure existence and uniqueness, the driver f and the terminal condition ξ satisfy the following regularity conditions:

Assumption A.2. *(on the backward SDE coefficients).*

H3. The driver f is uniformly Lipschitz continuous in the space variables, i.e there exists $K > 0$ such that

$$|f(t, y_1, z_1) - f(t, y_2, z_2)| \leq K(|y_1 - y_2| + |z_1 - z_2|).$$

H4. The terminal condition ξ is square integrable, i.e $\xi \in L^2$.

When the driver is stochastic then it must be measurable with respect to $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d)$ and square integrable with respect to $d\mathbf{P} \times dt$

$$\mathbf{E}\left[\int_0^T |f(t,0,0)|^2 dt\right] < \infty.$$

Also, the inequality of assumption (H3) holds $\mathbf{P}-$almost surely.

A pair of adapted processes (Y, Z) is called a solution of the BSDE of equation (A.3) if it satisfies equation (A.3), the process $\{Y_t\}_{t\in[0,T]} \in L_S^2$ is continuous and $\{Z_t\}_{t\in[0,T]} \in L_I^2$.

Pardoux and Peng [96] proved an existence and uniqueness result for the BSDE.

Proposition A.2. *Under assumptions (H3) and (H4), the backward stochastic differential equation given by equation (A.3) admits a unique solution* (Y, Z).

A decoupled forward-backward stochastic differential equation is obtained by combining a forward SDE and a backward SDE in the following manner

$$\begin{cases} dX_t = a(t, X_t)dt + \sigma(t, X_t)dW_t \\ -dY_t = f(t, X_t, Y_t, Z_t)dt - Z_t dW_t \\ X_0 = x_0, Y_T = \xi \end{cases} \tag{A.5}$$

where the forward coefficients a and σ do not depend upon the solution of the backward equation (Y, Z) but the measurable driver

$$f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$$

may depend on the forward process X.

The terminal condition ξ is said to be Markovian if it can be written as a functional of the forward process terminal value, i.e

$$\xi = g(X_T) \tag{A.6}$$

for some function $g : \mathbb{R}^d \to \mathbb{R}$ and non-Markovian if the functional involves other state variables. In this last case we simply write $\xi = G(X)$ with

$$G : \mathcal{S}^d \to \mathbb{R} \tag{A.7}$$

such that $|G(\mathbf{0})| \leq K$ for some constant $K > 0$ where $\mathbf{0}$ is the zero-valued vector function on $[0, T]$. The functional G is said to be L^∞-Lipschitz if there exists a constant $K > 0$ such that

$$|G(x) - G(y)| \leq K \sup_{t \in [0,T]} |x(t) - y(t)| \tag{A.8}$$

and L^1-Lipschitz if there exists a constant $K > 0$ such that

$$|G(x) - G(y)| \leq K \int_0^T |x(t) - y(t)| \, dt \tag{A.9}$$

for $x, y \in \mathcal{S}^d$.

As to the solution to the decoupled FBSDE, it consists of a triple of processes (X, Y, Z) and its existence and uniqueness naturally follow from Propositions A.1 and A.2 after simple adaptations of hypothesis (H3) for a stochastic driver.

When the driver f is independent of the control process $Z = \{Z_t\}_{t \in [0,T]}$, one can rewrite the backward equation using an expectation representation as

$$Y_t = \mathbf{E}\left[\xi + \int_t^T f(s, X_s, Y_s)ds \;\middle|\; \mathcal{F}_t \right] \tag{A.10}$$

with $t \in [0, T]$ and the control process $Z = \{Z_t\}_{t \in [0,T]}$ can then be obtained through the martingale representation theorem.

Also, many special cases of coupled forward-backward differential equations have been considered. The most general case takes the form

$$\begin{cases} dX_t = a(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t, Z_t)dW_t \\ -dY_t = f(t, X_t, Y_t, Z_t)dt - Z_t dW_t \\ X_0 = x_0 \ , \ Y_T = \xi \end{cases} \tag{A.11}$$

where both forward coefficients depend on the pair of processes solution of the backward equation.

Pardoux and Tang [98] have proved the FBSDE well-posedness, i.e, existence and uniqueness for the FBSDE solution (X, Y, Z) in three sub-cases with a probabilistic approach under monotonicity condition and stochastic coefficients. In particular, the Markovian case where the forward volatility $\sigma : [0,T] \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^{d \times d}$ is independent of the control process $Z = \{Z_t\}_{t \in [0,T]}$ is treated.

In the general situation of equation (A.11), the *four step scheme* of Ma, Protter and Yong [79] provides conditions for the well-posedness of a FBSDE with deterministic coefficients. Also, the method of continuation of Yong [114] constitutes an alternative method to show FBSDE well-posedness in the random coefficient case.

An important variation of BSDEs is the class of reflected BSDEs (RBSDEs) introduced by El Karoui et al. [45]. They are obtained by imposing boundary conditions on classical BSDEs and defined by the dynamics

$$\begin{cases} -dY_t = f(t, Y_t, Z_t)dt - Z_t dW_t + dA_t \\ Y_T = \xi \end{cases} \tag{A.12}$$

where the reflecting process $\{A_t\}_{t \in [0,T]}$ is a increasing process allowing the forward process to satisfy the boundary condition.

The solution of the differential equation consists of a square integrable forward process Y satisfying the boundary condition, the control process Z and a bounded variation process A which satisfies minimality conditions. One may refer to Ma and Yong [81] for existence and uniqueness results or El Karoui, Pardoux and Quenez [46] for applications to American option pricing.

A.2 Properties of solutions to FBSDEs

Additional theoretical results, which play a key role in numerical methods for solutions of FBSDEs, are available. We present them for FBSDEs with Lipschitz coefficients but equivalent results exist non-Lipschitz cases. In the following subsections, we present the relation between FBSDEs and quasilinear partial differential equations (PDEs) and also results on moment estimates of FBSDE solutions.

A.2.1 A priori estimates and regularity of solutions

The triple solution (X, Y, Z) of the FBSDE whose coefficients are allowed to be stochastic embeds some regularity properties that are worth mentioning. First, we have the following moment estimates for the forward process.

Proposition A.3. *For any $p \geq 2$, there exists a constant C depending on the time horizon T, the Lipschitz constant K and p such that the unique solution $\{X_t\}_{t \in [0,T]}$ of the forward SDE of equation (A.1) with Lipschitz (and measurable square integrable) coefficients a and b satisfies*

$$\boldsymbol{E}\left[\sup_{t \in [0,T]} |X_t|^p\right] \leq C\boldsymbol{E}\left[|x_0|^p + \int_0^T |a(t,0)|^p + |\sigma(t,0)|^p \, dt\right] \tag{A.13}$$

$$\boldsymbol{E}\left[|X_t - X_s|^p\right] \leq C\boldsymbol{E}\left[|x_0|^p + \sup_{t \in [0,T]} |a(t,0)|^p + \sup_{t \in [0,T]} |b(t,0)|^p\right] |t-s|^{p/2}. \tag{A.14}$$

for any $x \in \mathbb{R}^d$ and $s, t \in [0,T]$.

Moment estimates, from El Karoui, Peng and Quenez [47], are also available for solutions of BSDEs.

Proposition A.4. *For any $p \geq 2$, there exist a constant C depending on the time horizon T, the Lipschitz constant K and p such that the unique solution (Y, Z) of the BSDE of equation (A.3) with Lipschitz (and measurable square integrable) driver verifies, for $s, t \in [0, T]$,*

$$\begin{aligned} & \boldsymbol{E}\left[\sup_{t\in[0,T]} |Y_t|^p\right] + \boldsymbol{E}\left[\left(\int_0^T |Z_t|^2\, dt\right)^{p/2}\right] \\ & \quad \leq \quad C\boldsymbol{E}\left[|\xi|^p + \int_0^T |f(t,0,0)|^p\, dt\right], \end{aligned} \tag{A.15}$$

$$\begin{aligned} & \boldsymbol{E}\left[|Y_t - Y_s|^p\right] \\ & \quad \leq \quad C\boldsymbol{E}\left[\left(|\xi|^p + \sup_{t\in[0,T]} |f(t,0,0)|^p\right)|t-s|^{p-1} + \left(\int_s^t |Z_u|^2\, du\right)^{p/2}\right]. \end{aligned} \tag{A.16}$$

We now state a result on the stability property of the FBSDE solution and its continuous dependence to the FBSDE coefficients.

Proposition A.5. *Suppose $(X^\delta, Y^\delta, Z^\delta)$ is the solution of the perturbed decoupled FBSDE with initial value x_0^δ, terminal condition ξ^δ and coefficients a^δ, σ^δ and f^δ satisfying the Lipschitz conditions such that $x_0^\delta \to x_0$ as $\delta \to 0$ and*

$$\begin{aligned} \lim_{\delta\to 0} \boldsymbol{E}\left[\left|a^\delta(t,x) - a(t,x)\right|^2 + \left|\sigma^\delta(t,x) - \sigma(t,x)\right|^2\right] &= 0, \\ \lim_{\delta\to 0} \boldsymbol{E}\left[\left|\xi^\delta - \xi\right|^2 + \left|f^\delta(t,y,z) - f(t,y,z)\right|^2\right] &= 0. \end{aligned}$$

Then, we have that

$$\lim_{\delta\to 0} \boldsymbol{E}\left[\sup_{t\in[0,T]} \left|X_t^\delta - X_t\right|^2 + \sup_{t\in[0,T]} \left|Y_t^\delta - Y_t\right|^2 + \int_0^T \left|Z_t^\delta - Z_t\right|^2 dt\right] = 0. \tag{A.17}$$

Pardoux and Tang [98] obtained results similar to Propositions A.3, A.4 and A.5 for $p = 2$ in different coupled cases. Proofs of well-posedness of uni-dimensional BSDEs usually rely on an important property which is slightly stronger than the previous proposition and given by the comparison theorem bellow. The proof of this proposition and the proposition itself figure in El Karoui, Peng and Quenez [47].

Proposition A.6. *Let (Y^i, Z^i) be the unique solution to the BSDE of equation (A.3) with driver f^i and terminal value ξ^i for $i = 1, 2$. If the following inequalities hold*

- $\xi^1 \geq \xi^2$, $\mathbf{P}$*-a.s*
- $\delta f_t = f^1(t,y,z) - f^2(t,y,z) \geq 0$, $d\mathbf{P} \times dt$*-a.s and for all $y \in \mathbb{R}$ and $z \in \mathbb{R}^d$*

then $Y_t^1 \geq Y_t^2$ almost surely for any time $t \in [0, T]$.

Also, a path regularity property for the control process Z was proved by Zhang [123, 124]. The result is central when proving convergence of FBSDE time discretizations.

Proposition A.7. *Suppose the terminal function G is L^∞-Lipschitz and the control process Z is càdlàg. Then, there is a constant C depending on T and K only such that, for any partition $\pi = \{0 = t_0 < t_1 < \ldots < t_n = T\}$ of $[0, T]$, we have*

$$\sum_{i=1}^n \boldsymbol{E}\left[\int_{t_{i-1}}^{t_i} \left|Z_t - Z_{t_{i-1}}\right|^2 + \left|Z_t - Z_{t_i}\right|^2 dt\right] \leq C(1 + |x_0|^2)\,|\pi|\,, \tag{A.18}$$

where $|\pi| = \max_{1\leq i\leq n} |t_i - t_{i-1}|$ is the partition maximal time step.

When the FBSDE coefficients are deterministic, the control process Z displays càdlàg paths for L^∞-Lipschitz terminal functions and continuous paths for L^1-Lipschitz terminal functions as shown by Ma and Zhang[82]. A more recent result due to Gobet and Makhlouf [54] extends the control process regularity to irregular Markovian terminal function g lying in the space

$$L_{2,\alpha} = \left\{ g : \mathbf{E}\left[g(X_T)^2\right] + \sup_{t\in[0,T]} \frac{\mathbf{E}\left[g(X_T) - \mathbf{E}_t\left[g(X_T)\right]\right]^2}{(T-t)^\alpha} < \infty \right\} \tag{A.19}$$

for $\alpha \in (0,1]$.

A.2.2 Relation to quasilinear PDEs

The solution (X, Y, Z) of a FBSDE depends on the connections between its components since, for instance, coefficients of one of the FBSDE equations may depend on a component described by the other equation. This connection is clearly established for coupled FBSDEs with deterministic coefficients and Markovian terminal condition, i.e for the general situation of equation (A.11) with the terminal condition of equation (A.6). In this case, the FBSDE is linked to the quasilinear parabolic PDE

$$\begin{cases} \frac{\partial u}{\partial t} + \mathcal{L}u + f(t,x,u,z(t,x,u,\nabla u)) = 0,\ (t,x) \in [0,T)\times\mathbb{R}^d \\ u(T,x) = g(x),\ x \in \mathbb{R}^d \end{cases} \tag{A.20}$$

where $\nabla := \left(\frac{\partial}{\partial x_1}, ..., \frac{\partial}{\partial x_d}\right)$ represents the gradient operator,

$$z : [0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^d$$

is a function such that

$$z(t,x,y,p) = p\sigma(t,x,y,z(t,x,y,p)), \tag{A.21}$$

and

$$\begin{aligned} \mathcal{L}u &= \sum_{i=1}^{d} a_i(t,x,u,z(t,x,u,\nabla u))\frac{\partial u}{\partial x_i} \\ &\quad + \frac{1}{2}\sum_{i,j=1}^{d} b_{ij}(t,x,u,z(t,x,u,\nabla u))\frac{\partial^2 u}{\partial x_i \partial x_j} \end{aligned} \tag{A.22}$$

with $b = \sigma\sigma^*$.

Under various conditions on the existence, uniqueness and regularity of solutions to both equations (A.20) and (A.21), it can be shown[13] that the triple of processes (X, Y, Z) satisfies

$$X_t = x_0 + \int_0^t \tilde{a}(s, X_t)ds + \int_0^t \tilde{\sigma}(t, X_t)dW_t \tag{A.23}$$

where $\tilde{a} : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $\tilde{\sigma} : [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times d}$ are given by

$$\begin{aligned} \tilde{a}(t,x) &= a(t,x,u(t,x),z(t,x,u(t,x),\nabla u(t,x))), \\ \tilde{\sigma}(t,x) &= \sigma(t,x,u(t,x),z(t,x,u(t,x),\nabla u(t,x))) \end{aligned}$$

and the two remaining solution processes are defined as

$$Y_t = u(t, X_t), \tag{A.24}$$

$$Z_t = z(t, X_t, u(t, X_t), \nabla u(t, X_t)). \tag{A.25}$$

This result leads to a procedure, introduced by Ma, Yong and Protter [79], for solving coupled FBSDEs called the *four step scheme* that goes as follows:

[13] See Ma, Yong and Protter [79] or Ma and Yong [81] in Chapter 4, Theorem 1.1 for proof.

Four Step Method.

1. *Find the function* $z : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^d$ *defined by equation (A.21).*
2. *Solve the nonlinear PDE of equation (A.20) using the function* z *to get the function* $u : [0,T] \times \mathbb{R}^d \to \mathbb{R}$.
3. *Solve the forward SDE of equation (A.23) to get the forward process* X.
4. *Define the processes* Y *and* Z *according to equations (A.24) and (A.25) respectively using the functions* u *and* z *and the forward process* X.

B Time discretization of SDEs

The formal definition of higher order time discretizations uses the multi-index notation and iterated Brownian integral presented in Chapter 5 of Kloeden an Platen [69] and also in Chassagneux and Crisan [27]. The set $\mathbb{M}$ of multi-indice with entries in $\{0,...,d\}$ is given by

$$\mathbb{M} := \{\emptyset\} \bigcup \cup_{j=1}^{\infty} \{0,...,d\}^j \tag{B.1}$$

and for any multi-index $\imath \in \mathbb{M}$ the measures ℓ and $\bar{\ell}$ return the length and the number of zero of the multi-index respectively with $\ell(\emptyset) = 0$. Moreover, $-\imath$ (resp. $\imath-$) is the multi-index obtained by deleting the first (resp. last) entry of $\imath$ and $(j)_m$ refers to the multi-index of length $m > 0$ whose entries are identical and equal to $j \in \{0,...,d\}$. A hierarchical set $\mathbb{A} \subset \mathbb{M}$ is a set of multi-indice such that

$$\sup_{\imath} \ell(\imath) < \infty \text{ and } -\imath \in \mathbb{A}, \forall \imath \in \mathbb{A} \backslash \{\emptyset\}.$$

For instance, it is easily shown that the set

$$\mathbb{A}_m = \{\imath : \ell(\imath) + \bar{\ell}(\imath) \leq 2m \text{ or } \ell(\imath) = \bar{\ell}(\imath) = m + \frac{1}{2}, 2m \in \mathbb{N}^*\}$$

is a hierarchical set. We define the iterated Brownian integral $I^{\imath}_{s,t}$ with index $\imath$ of length $\ell(\imath) = l$ recursively as

$$I^{\imath}_{s,t} := \begin{cases} 1 & \text{, if } l = 0 \\ \int_s^t I^{\imath-}_{s,u} du & \text{, if } l > 0 \text{ and } \imath_l = 0 \\ \int_s^t I^{\imath-}_{s,u} dW^j_u & \text{, if } l > 0 \text{ and } \imath_l = j,\ 1 \leq j \leq d. \end{cases} \tag{B.2}$$

On a time partition $\pi = \{0 = t_0 < t_1 < ... < t_n = T\}$ of $[0,T]$, a strong order scheme for the forward process defined by the SDE (A.1) then has the form

$$X^{\pi}_{t_{i+1}} = X^{\pi}_{t_{i+1}} + \sum_{\imath \in \mathbb{A}_m \backslash \{\emptyset\}} F_{\imath}(X^{\pi}_{t_i}) I^{\imath}_{t_i, t_{i+1}} \tag{B.3}$$

for some bounded functions $F_{\imath} : \mathbb{R}^d \to \mathbb{R}^d$ related to the SDE coefficients a and σ. General strong schemes are interesting mainly for their convergence properties. Kloeden and Platen [69] (Theorem 11.5.1, page 391) show that a strong scheme built with the hierarchical set $\mathbb{A}_m$ converges with order m i.e

$$\begin{aligned} \mathcal{E}^2_{X,\pi} &:= \max_{0 \leq i \leq n} \left\| X_{t_i} - X^{\pi}_{t_i} \right\|^2_{L^2} \\ &= \mathcal{O}(|\pi|^{2m}). \end{aligned} \tag{B.4}$$

C Gauss quadrature approach to BSDEs

Instead of the convolution approach of Chapter 2 and Chapter 3, a Gauss quadrature can be used to compute the conditional expectations involved in the numerical solution of BSDEs. In this section, we present the method in the deterministic and stochastic points of view. Nonetheless, both approaches are equivalent in practice since the resolution of the stochastic approach is made through the deterministic approach presented bellow.

C.1 Deterministic approach

Starting from the Euler time discretization, we develop a deterministic algorithm for the approximate solution u_i and the approximate gradient $\dot{u}_i$ in Chapter 2. At time step t_i, the solutions may be written as

$$u_i(x) = \tilde{u}_i(x) + \Delta_i f(t_i, \tilde{u}_i(x), \dot{u}_i(x)) \tag{C.1}$$

where

$$\dot{u}_i(x) = \frac{1}{\Delta_i} \int_{-\infty}^{\infty} y u_{i+1}(x+y) h(y) dy \tag{C.2}$$

$$\tilde{u}_i(x) = \int_{-\infty}^{\infty} u_{i+1}(x+y) h(y) dy \tag{C.3}$$

for $i = 0, 1, ..., n-1$ and $u_n(x) = g(x)$.

The presence of the Gaussian density h in equation C.2 and C.3 allows the usage of Gaussian quadrature based numerical method. Indeed, knowing the Gaussian density h has the form

$$h(x) = (2\pi\Delta_i)^{-\frac{1}{2}} \exp\left(-\frac{x^2}{2\Delta_i}\right), \tag{C.4}$$

equations C.2 and C.3 can be written as

$$\dot{u}_i(x) = (2\pi\Delta_i)^{-\frac{1}{2}} \int_{-\infty}^{\infty} y u_{i+1}\left(x + y\sqrt{\Delta_i}\right) e^{-\frac{1}{2}y^2} dy \tag{C.5}$$

$$\tilde{u}_i(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} u_{i+1}\left(x + y\sqrt{\Delta_i}\right) e^{-\frac{1}{2}y^2} dy \tag{C.6}$$

after a change of variable. The approximate solution and gradient values can hence be approximated locally with a Gauss-Hermite quadrature. For a concise introduction to Gaussian quadratures, one may refer to Chapter 9 of Kress [71], Chapter 6 of Kiusalaas [72], Chapter 4 of Burden and Faires [25], Chapter 3 of Moin [91], Chapter 5 of Sauer [106] or Chapter 8 of Hildebrand [58] among many others.

At a particular space position $x \in \mathbb{R}$, the intermediate solution $\tilde{u}_i$ and the approximate gradient $\dot{u}_i$ at time mesh t_i may be computed with the $N-$points Gauss-Hermite quadrature with $N > 1$ as

$$\tilde{u}_i(x) = \sum_{j=1}^{N} w_j u_{i+1}(x + y_j\sqrt{\Delta_i}) + E_0 \tag{C.7}$$

$$\dot{u}_i(x) = \frac{1}{\sqrt{\Delta_i}} \sum_{j=1}^{N} w_j y_j u_{i+1}(x + y_j\sqrt{\Delta_i}) + E_1. \tag{C.8}$$

where E_0 and E_1 stand for the integration errors. In addition, the integration nodes $\{y_j\}_{j=1}^N$ and weights $\{w_j\}_{j=1}^N$ are retrieved from the Nth order polynomial of the family of orthogonal polynomials with weight function $w(x) = e^{-\frac{1}{2}x^2}$.

Suppose the polynomials H_n of degree $n \in \mathbb{N}$ satisfy

$$\int_{\mathbb{R}} H_n(x)H_m(x)w(x)dx = 0,\ n \neq m. \tag{C.9}$$

Then $\{H_n\}_{n\in\mathbb{N}}$ is said to be a family of orthogonal polynomials with weight function $w : \mathcal{I} \to \mathbb{R}$ for some interval $\mathcal{I}$. The literature on orthogonal polynomials is considerable, we will limit ourselves to Szegö [109] or Stahl and Totik [108] for a theoretical approach and Marcellán and Van Assche [86] or Gautschi, Golub and Opfer [51] for applications. A first property of interest for orthogonal polynomials has to do with their zeros and is stated in the next proposition. The statement may be found in Szegö [109] (Theorem 3.3.1, page 44).

Proposition C.1. *The zeros of the orthogonal polynomials H_n are real, distinct and located in the interior of the interval $\mathcal{I}$.*

In our case, the orthogonal polynomials $\{H_n\}_{n\in\mathbb{N}}$ are the (probabilistic) Hermite polynomials since the weight function is $w(x) = e^{-\frac{1}{2}x^2}$. The Hermite polynomials admits the representation

$$H_n(x) = (-1)^n e^{\frac{1}{2}x^2} \frac{d^n}{dx^n} e^{-\frac{1}{2}x^2} \tag{C.10}$$

and the recurrence formula

$$H_{n+1}(x) = xH_n(x) - nH_{n-1}(x) \tag{C.11}$$

with initial polynomials $H_{-1}(x) = 0$ and $H_0(x) = 1$.

When performing the Gauss-Hermite quadrature, the nodes $\{y_j\}_{j=1}^N$ and weights $\{w_j\}_{j=1}^N$ are chosen such that any polynomial of degree $2N-1$ or less is integrated exactly by the approximation. As a consequence, the following proposition holds and is an adaptation of the results of Hildebrand [58] (pp. 388-390).

Proposition C.2. *The nodes $\{y_j\}_{j=1}^N$ of the Gauss-Hermite quadrature are the N real and distict zeros of the orthogonal polynomial H_N. Also, the weights satisfy*

$$w_j = \frac{(N-1)!}{N[H_{n-1}(y_j)]^2} \tag{C.12}$$

for $j = 1, 2, ..., N$.

Another interesting feature of the weights w_i is that they sum to one (1). Indeed

$$\sum_{j=1}^{N} w_i = \int_{\mathbb{R}} (2\pi)^{-\frac{1}{2}} e^{-\frac{1}{2}x^2} dx = 1$$

since constant functions are integrated exactly. The error terms E_0 and E_1 admits the bound given in the following proposition which is also adapted from Hildebrand [58] (pages 388-390).

Proposition C.3. *If $u_{n+1} \in \mathcal{C}^{2N}$, then*

$$E_0 = u_{i+1}^{(2N)}(\zeta) \frac{\Delta^N \sqrt{2\pi} N!}{(2N)!} \tag{C.13}$$

and

$$E_1 = \left(\Delta^{\frac{1}{2}} \zeta u_{i+1}^{(2N)}(\zeta) + (2N) u_{i+1}^{(2N-1)}(\zeta) \right) \frac{\Delta^{N-1} \sqrt{2\pi} N!}{(2N)!} \tag{C.14}$$

for some $\zeta \in \mathbb{R}$.

Hence, estimates of the approximate solutions u_i and the approximate gradient $\dot{u}_i$, $i = 0, 1, ..., n-1$, may be computed using a classical multinomial tree with N branches through equations C.1, C.7 and C.8. For $N = 2$ and $N = 3$, the multinomial tree recombines which eases the implementation of the Gauss-Hermite method.

C.2 Stochastic approach

Following Briand, Delyon and Mémin [21] and Peng and Xu [99], the BSDE

$$Y_t = g(W_T) + \int_t^T f(s, Y_s, Z_s)ds + \int_t^T Z_s dW_s \tag{C.15}$$

is discretized on the time mesh $\pi = \{t_0 = 0 < t_1 < ... < t_n = T\}$ with a number n of time steps and $t_i = i\Delta$ where $\Delta = \frac{T}{n}$. The discretization then gives the discrete time BSDE

$$Y^\pi_{t_i} = Y^\pi_{t_{i+1}} + \Delta f(t_i, Y^\pi_{t_i}, Z^\pi_{t_i}) - \sqrt{\Delta} Z^\pi_{t_i}\epsilon_{i+1} \ , \ i = 0, 1, ..., n-1 \tag{C.16}$$

where $\{\epsilon_i\}_{i=1}^n$ is a sequence of discrete independent and identically distributed random variables such that $\epsilon_i \in \mathcal{F}_{t_i}$. The Gauss-Hermite quadrature approach gives a systematic way to define the random variable ϵ_i. Indeed, one may select them so that their probability distribution function is given by

$$f_N(y) = \sum_{j=1}^N w_j \delta(y - y_j) \ , \ y \in \mathbb{R} \tag{C.17}$$

leading to the probability measure

$$\mathbf{P}\left[\epsilon_i \in A\right] = \sum_{j=1}^N w_j \delta_{y_j}(A) \ , \ A \subseteq \mathbb{R} \tag{C.18}$$

where δ is the Dirac delta function and δ_x is the Dirac delta measure. Hence, ϵ_i takes the value y_j with probability w_j. The following proposition holds since the Gauss-Hermite quadrature integrates monomials of degree less that $2N$ exactly.

Proposition C.4. *The first $2N-1$ moments of ϵ_i are those of a standard normal distribution.*

Let $\{W^\pi_t\}_{t\in[0,T]}$ be the adapted process defined as

$$W^\pi_t = \sqrt{\Delta} \sum_{i=1}^{\lfloor t\Delta^{-1}\rfloor} \epsilon_i \ , \tag{C.19}$$

then the discrete BSDE terminal value is given by

$$Y^\pi_T = g(W^\pi_T). \tag{C.20}$$

The next proposition describes how the scale random walk $\{W^\pi_t\}_{t\in[0,T]}$ may be used to approximate the standard Brownian motion in this context.

Proposition C.5. *Suppose that for a sequence $\{k_n\}_{n\in\mathbb{N}}$ where $0 \le k_n \le 1$ and $\lim_{n\to\infty} k_n = 0$,*

$$\Delta^{-\frac{1}{2}} \boldsymbol{E}\left[\epsilon_i \Delta W_j\right] = \delta_{i,j}(1 - k_n), \ i, j = 1, 2, ..., n. \tag{C.21}$$

Then the adapted process $\{W^\pi_t\}_{t\in[0,T]}$ is such that

$$\sup_{t\in[0,T]} |W^\pi_t - W_t| \to 0 \tag{C.22}$$

as $n \to \infty$, where the convergence holds in probability.

Proof. First notice that the process $\{W^\pi_t\}_{t\in[0,T]}$ is a martingale, so that $|W^\pi_t - W_t|$ is a sub-martingale by Jensen inequality. Also, for any $\varepsilon > 0$

$$\mathbf{P}\left[\sup_{t\in[0,T]} |W^\pi_t - W_t| > \varepsilon\right] \le \varepsilon^{-2}\mathbf{E}\left[|W^\pi_T - W_T|^2\right]$$

by Doob's inequality, where

$$\begin{aligned}
&\mathbf{E}\left[\left|W_T^\pi - W_T\right|^2\right] \\
&= \mathbf{E}\left[(W_T^\pi)^2\right] - 2\mathbf{E}\left[W_T^\pi W_T\right] + \mathbf{E}\left[(W_T)^2\right] \\
&= \Delta\sum_{i=1}^{n}\mathbf{E}\left[\epsilon_i^2\right] - 2\sqrt{\Delta}\mathbf{E}\left[\left(\sum_{i=1}^{n}\epsilon_i\right)\left(\sum_{j=1}^{n}\Delta W_j\right)\right] + T \\
&= 2T - 2\sqrt{\Delta}\mathbf{E}\left[\sum_{i=1}^{n}\epsilon_i\Delta W_i\right] \\
&= 2T - 2T(1-k_n) \\
&= 2Tk_n.
\end{aligned}$$

From this last equation, we have that $\mathbf{P}\left[\sup_{t\in[0,T]}|W_t^\pi - W_t| > \varepsilon\right] \to 0$ as $n \to \infty$. □

For any $t \in [0,T]$, we extend the discrete solution such that

$$Y_t^\pi = Y_{t_i}^\pi \ , t \in [t_i, t_{i+1}) \tag{C.23}$$

$$Z_t^\pi = Z_{t_i}^\pi \ , t \in [t_i, t_{i+1}). \tag{C.24}$$

The following proposition holds as a consequence of Proposition C.5 and Theorem 2.1 of Briand, Delyon and Mémin [21].

Proposition C.6. *Suppose that Assumption A.2 holds, $g \in \mathcal{C}_b$ and the condition of equation (C.21) holds. Then*

$$\sup_{t\in[0,T]}|Y_t^\pi - Y_t| + \int_0^T |Z_t^\pi - Z_t|^2\,dt \to 0 \tag{C.25}$$

as $n \to \infty$, where the convergence holds in probability.

The boundedness of the terminal condition g is required only to ensure the convergence in mean of $g(W_T^\pi)$ and $g^2(W_T^\pi)$ to $g(W_T)$ and $g^2(W_T)$ respectively. This convergence in mean then allows to meet the requirements of Theorem 2.1 in Briand, Delyon and Mémin [21].

In general, the BSDE is then numerically solved with the explicit scheme of equations (1.2.4) and (1.2.3). A particular case of this general (Gauss-Hermite) multinomial approach is the binomial method of Peng and Xu [99] where ϵ_i has the distribution function

$$f_2(y) = \frac{1}{2}\left(\delta(y-1) + \delta(y+1)\right). \tag{C.26}$$

In the trinomial approach that we use in Chapter 2, the distribution of the discrete increments ϵ_i is given by

$$f_3(y) = \frac{2}{3}\delta(y) + \frac{1}{6}\left(\delta(y-\sqrt{3}) + \delta(y+\sqrt{3})\right). \tag{C.27}$$

D The Gauss-Weierstrass transform

The convolution method starts with expressions of the approximate solutions in integral forms. For the explicit Euler scheme 1 and the implicit Euler scheme, the intermediate solution $\tilde{u}_i$ and the intermediate gradient $\dot{u}_i$ at time mesh t_i are convolution transformations. Convolution transformations are special cases of integral transformations and have been extensively studied and applied. The books of Hirschman and Widder [60] summarizes the theory and applications range from signal processing to probability and statistics.

In our case, the convolution integrals involved in the approximate solutions are particularly Gauss-Weierstrass transformations. Hirschman and Widder [60] give an overview in Chapter 8 of their book. Additional results may be found in authors such as Hille [59], Widder [113], Bilodeau [14, 13], Zemanian [118, 119] among others.

D.1 Definition and connection to BSDEs

Let's consider the Gaussian density h_t with

$$h_t(x) = (2\pi t)^{-\frac{1}{2}} e^{-\frac{x^2}{2t}},\ t > 0,\ x \in \mathbb{R}. \tag{D.1}$$

For a real valued function $g : \mathbb{R} \to \mathbb{R}$, the Weierstrass transform of g can be defined as the function $u : \mathbb{R} \to \mathbb{R}$

$$\begin{aligned} u(x) &= e^{\frac{1}{2}D^2} g(x) \\ &:= \int_{\mathbb{R}} g(y) h_1(x-y) dy \end{aligned} \tag{D.2}$$

whenever the integral converges where $D := \frac{d}{dx}$ is the first order differential operator. Hence, u is the convolution transform of g with kernel h_1. The literature usually defines the transform at scale $t = 2$, we use scale $t = 1$ for convenience purposes to suit our approximate solution expressions.

The generalized Gauss-Weierstrass transform considers kernel at different scales $t > 0$ and may be defined as

$$\begin{aligned} u(t,x) &= e^{\frac{t}{2}D^2} g(x) \\ &:= \int_{\mathbb{R}} g(y) h_t(x-y) dy. \end{aligned} \tag{D.3}$$

The literature has focused mainly on the inversion problem for the Gauss-Weierstrass transform. For entire real valued functions u on $\mathbb{R}$ such that the complex extension $u(x+\mathbf{i}y)$ satisfies some analytical and growth conditions, Hirshman and Widder [60] (Theorem 3.2, page 180) show that the inverse Gauss-Weierstrass transform of u is given by

$$e^{-\frac{t}{2}D^2} u(x) = \int_{\mathbb{R}} h_t(y) u(x+\mathbf{i}y) dy. \tag{D.4}$$

Hence, an entire function u (satisfying analytical and growth conditions) can be represented as the Gauss-Weierstrass transform of g where

$$g(x) = e^{-\frac{1}{2}D^2} u(x). \tag{D.5}$$

Theorem 13.3 (page 207) of Hirshman and Widder [60] gives a growth condition on the inverse Gauss-Weierstrass transform g. One must have that

$$|g(x)| < M e^{ax^2} \tag{D.6}$$

for some constant $M > 0$ and $-\infty < a < \frac{1}{2}$ in our case.

The Gauss-Weierstrass transform has strong links with diffusion PDE. Notice the kernel h_t of equation (D.1) is the Green's function associated to the diffusion operator

$$\mathcal{L} = \frac{\partial}{\partial t} - \frac{1}{2}\frac{\partial^2}{\partial x^2}.$$

Hence, the function $u(T-t,x) = e^{\frac{T-t}{2}D^2} g(x)$ is the solution to the Cauchy problem

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial^2 u}{\partial x^2} = 0,\ (t,x) \in [0,T) \times \mathbb{R} \\ u(T,x) = g(x),\ x \in \mathbb{R}. \end{cases} \tag{D.7}$$

This PDE is itself associated to the BSDE with null driver and terminal condition g.

The approximate solutions for our BSDE admit representations in term of Gauss-Weierstrass transform. The following proposition sums up the idea in the one-dimensional case for the Euler scheme 1 and holds for the implicit Euler scheme. Similar expressions can be developed for the Euler scheme 2.

Proposition D.1. *For the Euler scheme 1, the intermediate solution $\tilde{u}_i$ and the approximate gradient $\dot{u}_i$ at mesh time t_i satisfy*

$$\tilde{u}_i(x) = e^{\frac{\Delta_i}{2}D^2}u_{i+1}(x) \tag{D.8}$$

$$\dot{u}_i(x) = D\tilde{u}_i(x) = \frac{d\tilde{u}_i}{dx}(x) \tag{D.9}$$

for $i = 0, 1, ...n-1$.

Proof. The representation of equation (D.8) follows obviously from the integral representation of the intermediate solution $\tilde{u}_i$ in equation (2.2.8) given equation (2.2.9) holds. Also, differentiating the intermediate solution integral in equation (2.2.8) leads to equation (D.9) assuming the differential and integral operators can be interchanged. □

The proposition shows that the approximation solutions for BSDEs are obtained by solving a diffusion PDE with null driver between consecutive mesh times.

For some input functions g, the Gauss-Weierstrass transform is known in closed form. These input functions include Hermite polynomials, exponential functions and Gaussian functions with scale parameter strictly less than $\frac{1}{2}$. Hirshman and Widder [60] gives a table of different transforms on page 178 using an alternative definition of the Gauss-Weierstrass transform. The orthogonality of Hermite polynomials also leads to interesting series expansions for Gauss-Weierstrass transforms which can be used to represent approximate solutions for BSDEs. Elements on this subject may be found in Hille [59] or Bilodeau [14].

D.2 Closed forms for approximate solutions

We next give an example where the numerical solutions for BSDEs can be developed in closed form. We will limit to linear BSDEs with drivers of the form

$$f(t, y, z) = ay + bz \tag{D.10}$$

and terminal condition $g : \mathbb{R} \to \mathbb{R}$. It is well-known in this linear case that the backward process is given (see El Karoui, Peng and Quenez [47])

$$Y_t = \mathbf{E}_t\left[g(W_T)\Gamma_T^t\right] \tag{D.11}$$

where, for $s \geq t$,

$$\Gamma_s^t = e^{(a-\frac{1}{2}b^2)(s-t)+b(W_s-W_t)}. \tag{D.12}$$

If we set, for any $\varphi \in \mathbb{R}$,

$$g(x) = e^{\varphi x} \tag{D.13}$$

then the BSDE solution can be represented as

$$Y_t = e^{(a+b\varphi+\frac{1}{2}\varphi^2)(T-t)}g(W_t) \tag{D.14}$$

$$Z_t = \varphi e^{(a+b\varphi+\frac{1}{2}\varphi^2)(T-t)}g(W_t). \tag{D.15}$$

The approximate solutions for the BSDE derived from the Euler scheme 1 on an equidistant time grid $\pi = \{0 = t_1 < t_2 < ... < t_n = T\}$ with time step $\Delta = \frac{T}{n}$ is

$$u_i(x) = (1 + a\Delta + b\varphi\Delta)^{n-i}e^{\frac{1}{2}\varphi^2(T-t_i)}g(x) \tag{D.16}$$

$$\dot{u}_i(x) = \varphi(1 + a\Delta + b\varphi\Delta)^{n-i-1}e^{\frac{1}{2}\varphi^2(T-t_i)}g(x) \tag{D.17}$$

$i = 0, 1, ..., n-1$. This numerical solution may be obtained by induction using Proposition D.1 and knowing that the Gauss-Weierstrass transform of g is $e^{\frac{t}{2}D^2}g(x) = e^{\frac{1}{2}\varphi^2 t}g(x)$.

E Elements of Fourier analysis

Fourier analysis studies function representations as trigonometric series. It has been a topic of research since Joseph Fourier's work on the application of Fourier series to the heat equation in the beginning of the nineteenth century. In addition to partial differential equations, applications of Fourier analysis now include signal processing, probability theory or finance among others. For the main theoretical tool in the convolution method of Chapters 2 and 3 is Fourier analysis, this appendix aims to be an overview of the subject. Many books are also available on theoretical or practical aspects of Fourier analysis. Many results in this appendix can be found in the books of Vretblad [112], Edwards [43], Bernatz [12] and Plato [101] for instance.

Section E.1 presents complex Fourier series and their basic properties. Results of the Fourier transform are given in Section E.2 and Section E.3 deals with Fourier series approximation and its relationship with the discrete Fourier transform (DFT).

E.1 Fourier series expansion

Let $h : \mathbb{R} \to \mathbb{C}$ be a complex-valued function defined on the real line. Suppose in addition that the following assumption holds.

Assumption E.1. *The function $h : \mathbb{R} \to \mathbb{C}$ is periodic with period 2π and integrable on $[-\pi, \pi]$ with*

$$\int_{-\pi}^{\pi} |h(t)|\, dt < \infty. \tag{E.1}$$

The complex Fourier series associated with the function h is the expansion of the form

$$\sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}kx} \tag{E.2}$$

where the Fourier coefficients are defined as

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(t) e^{-\mathbf{i}kt} dt \tag{E.3}$$

and $\mathbf{i} = \sqrt{-1}$ is the imaginary unit.

Another Fourier series expansion that is widely used is the real Fourier series which is the expansion of the form

$$\frac{1}{2}a_0 + \sum_{k=1}^{\infty} \left(a_k \cos(kx) + b_k \sin(kx)\right) \tag{E.4}$$

where the (real) Fourier coefficients are

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} h(t) \cos(kt) dt \tag{E.5}$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} h(t) \sin(kt) dt. \tag{E.6}$$

The equivalence between the complex and the real Fourier series expansions can be established by observing the relationship between their respective coefficients. Indeed, it's easily shown that

$$a_k = c_k + c_{-k} \text{ and } b_k = \mathbf{i}(c_k - c_{-k}). \tag{E.7}$$

For this reason, we focus on the complex Fourier series of equation (E.2) in the rest of this appendix.

Some of the most important issues with Fourier series expansion are the study of their convergence and uniqueness. Since the function h is assumed periodic with period 2π, one can

limit the study of its Fourier series on the interval $[-\pi, \pi]$. Unfortunately, the integrability condition of Assumption E.1 gives only very little information on the Fourier coefficients and consequently on the convergence of the Fourier series. The next proposition illustrates this idea and it can be found in Vretblad [112] (Lemma 4.1, page 79).

Proposition E.1. *Suppose h satisfies Assumpion E.1 then the sequence of Fourier coefficients $\{c_k\}_{k\in\mathbb{Z}}$ is bounded with*

$$|c_k| \leq \frac{1}{2\pi} \int_{-\pi}^{\pi} |h(t)|\, dt \tag{E.8}$$

and converges to zero (0)

$$\lim_{|k|\to\infty} c_k = 0. \tag{E.9}$$

Indeed, the convergence of Fourier coefficients to zeros (0) does not ensure the convergence of the series expansion. Moreover, convergence of the Fourier series expansion does not guarantee its convergence to the function h. Hence, further smoothness assumptions have to be made on the function h as shown by the following proposition. The proposition is stated in Vretblad [112] (Theorem 4.2, page 83) and ideas of its proof can be found in the same reference or in Edwards [43].

Proposition E.2. *Assume the function h satisfies Assumption E.1 and admits the Fourier coefficients $\{c_k\}_{k\in\mathbb{Z}}$ such that the serie of Fourier coefficients converges absolutely*

$$\sum_{k=-\infty}^{\infty} |c_k| < \infty. \tag{E.10}$$

If h is continuous then the Fourier series of equation (E.2) converges uniformly to h on $[-\pi, \pi]$.

A result similar to Proposition E.2 (Carleson's theorem) uses less constraints on function h. It can actually be shown that the Fourier series expansion of a continuous and integrable function h converges to h almost everywhere on $[-\pi, \pi]$ in the Lebesque sense. Almost everywhere convergence also holds if $h \in \mathbf{L}^p([-\pi, \pi])$ with $p > 1$.

Continuity also guarantees uniqueness of the Fourier series expansion (See Corollary 4.1, page 84 of Vretblad [112]). Hence, for a continuous function h satisfying Assumption E.1 we can write

$$h(x) = \sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}kx},\ x \in [-\pi, \pi] \tag{E.11}$$

where the equality holds almost everywhere in the Lebesque sense.

An interesting result gives a description of Fourier coefficients when the function h is differentiable. We state it as in Vretblad [112] (Theorem 4.4 page 85).

Proposition E.3. *If $h \in \mathcal{C}^m[-\pi, \pi]$ then*

$$|c_k| \leq M\, |k|^{-m} \tag{E.12}$$

for some constant M.

It is possible to define the Fourier series of the periodic function h with period $b - a$ on a general interval $[a, b]$ by a change of variable. All results exposed previously holds and the expansion takes the form

$$h(x) = \sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}k\frac{2\pi}{b-a}x} \tag{E.13}$$

where the complex coefficients are given by

$$c_k = \frac{1}{b-a} \int_a^b h(t) e^{-\mathbf{i}k\frac{2\pi}{b-a}t} dt. \tag{E.14}$$

E.2 Fourier transform

The Fourier transform of an integrable function $h : \mathbb{R}^d \to \mathbb{C}$ with

$$\int_{\mathbb{R}^d} |h(t)|\, dt < \infty \tag{E.15}$$

is the function $\hat{h} : \mathbb{R}^d \to \mathbb{C}$ defined as

$$\hat{h}(\nu) := \mathfrak{F}[h](\nu) = \int_{\mathbb{R}^d} e^{-\mathbf{i}\nu^* x} h(x) dx. \tag{E.16}$$

The inverse Fourier transform recovers the function h from its Fourier transform $\hat{h}$ through the relation

$$h(x) := \mathfrak{F}^{-1}[\hat{h}](x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{\mathbf{i}\nu^* x} \hat{h}(\nu) d\nu. \tag{E.17}$$

Among the various properties of the Fourier transform (and its inverse transform) such as linearity and relations under shifting and scaling, two of them are of particular importance in Chapter 2 and throughout the thesis. We present these properties in the one-dimensional case since they can easily generalized. The first relates the initial function with the derivatives of its Fourier transform.

Proposition E.4. *Let $\hat{h}$, the Fourier transform of the function h, be differentiable. Then*

$$\mathfrak{F}[xh(x)](\nu) = \mathbf{i}\frac{\partial \hat{h}}{\partial \nu}(\nu) \tag{E.18}$$

and

$$\mathfrak{F}[\frac{\partial h}{\partial x}(x)](\nu) = \mathbf{i}\nu\hat{h}(\nu). \tag{E.19}$$

Proof. For the first part of the proposition, we have that

$$\begin{aligned} \frac{\partial \hat{h}}{\partial \nu}(\nu) &= \frac{\partial}{\partial \nu} \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu x} h(x) dx \\ &= \int_{-\infty}^{\infty} \frac{\partial}{\partial \nu} e^{-\mathbf{i}\nu x} h(x) dx, \end{aligned}$$

assuming the integral and the differential operators can be interchanged, so that

$$\begin{aligned} \frac{\partial \hat{h}}{\partial \nu}(\nu) &= -\mathbf{i} \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu x} x h(x) dx \\ &= -\mathbf{i}\mathfrak{F}\left[xh(x)\right](\nu). \end{aligned}$$

As to the second part, we have

$$\begin{aligned} \mathbf{i}\nu\hat{h}(\nu) &= \int_{-\infty}^{\infty} \mathbf{i}\nu e^{-\mathbf{i}\nu x} h(x) dx \\ &= -\int_{-\infty}^{\infty} \left(\frac{\partial}{\partial x} e^{-\mathbf{i}\nu x}\right) h(x) dx \\ &= \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu x} \frac{\partial h}{\partial x}(x) dx - e^{-\mathbf{i}\nu x} h(x)\Big|_{-\infty}^{\infty} \\ &\quad \text{(applying the intregration by part formula),} \\ &= \mathfrak{F}\left[\frac{\partial h}{\partial x}(x)\right](\nu) \end{aligned}$$

whenever h is integrable. □

The second property is the well known convolution theorem which define the Fourier transform of a convolution of two functions as the product of the individual Fourier transforms.

Proposition E.5. (Convolution theorem) *Let $h * k$ be the convolution of the functions h and k i.e*

$$(h * k)(x) = \int_{-\infty}^{\infty} h(x-t)k(t)dt$$

then

$$\mathfrak{F}[h * k] = \mathfrak{F}[h]\mathfrak{F}[k]. \tag{E.20}$$

Proof. The proof is straightforward,

$$\begin{aligned}\mathfrak{F}[h * k](\nu) &= \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu x} \int_{-\infty}^{\infty} h(x-t)k(t)dtdx \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\mathbf{i}\nu x} h(x-t)k(t)dxdt.\end{aligned}$$

The change of variable $z = x - t$ then gives

$$\begin{aligned}\mathfrak{F}[h * k](\nu) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\mathbf{i}\nu(z+t)} h(z)k(t)dzdt \\ &= \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu z} h(z)dz \int_{-\infty}^{\infty} e^{-\mathbf{i}\nu t} k(t)dt.\end{aligned}$$

□

E.3 Fourier series and the discrete Fourier transform

In order to approximate a function with Fourier series, on may consider truncating the Fourier series expansion. Proposition E.3 naturally leads to a result on this procedure. Hence, the proof of the following theorem is mainly based on the convergence of hyper-harmonic series.

Proposition E.6. *Suppose the function u satisfies Assumpsion E.1 and admits the Fourier series expansion*

$$u(x) = \sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}kx},\ x \in [-\pi, \pi]\,. \tag{E.21}$$

If $h \in \mathcal{C}^{m+1}[-\pi, \pi]$ with $m \geq 1$ then for any $N \in \mathbb{N}^*$

$$u(x) = \sum_{k=-N}^{N-1} c_k e^{\mathbf{i}kx} + \mathcal{O}(N^{-m}). \tag{E.22}$$

Proof. Let e_N be such that $u(x) = \sum_{k=-N}^{N-1} c_k e^{\mathbf{i}kx} + e_N$. Then by Proposition E.3

$$\begin{aligned}|e_N| &\leq 2M \sum_{k=N}^{\infty} k^{-m-1} \\ &= 2MN^{-m-1} \sum_{k=0}^{\infty} \left(1 + \frac{k}{N}\right)^{-m-1} \\ &= 2MN^{-m-1} \sum_{n=1}^{\infty}\sum_{k=0}^{N-1} \left(n + \frac{k}{N}\right)^{-m-1} \\ &\leq 2MN^{-m} \sum_{n=1}^{\infty} n^{-m-1} \\ &= CN^{-m}\end{aligned}$$

for some constant $C > 0$ and for any $m \geq 1$. □

For an approximation by Fourier series truncation to be useful, the Fourier coefficients involved in the truncated series must be available. Deriving analytical formulas for the Fourier coefficients of a given function is usually a very tedious task. Hence, approximating the coefficients appears much more suitable.

The discrete Fourier transform (DFT) transforms a set of real or complex numbers $\{x_j\}_{j=0}^{N-1}$ into another set $\{\hat{x}_j\}_{j=0}^{N-1}$ through the relation

$$\hat{x}_k := \mathfrak{D}[x]_k = \frac{1}{N}\sum_{j=0}^{N-1} e^{-\mathbf{i}jk\frac{2\pi}{N}} x_j \tag{E.23}$$

for $k = 0, 1, ..., N-1$. The inverse DFT performs the reciprocal operation by computing the set of numbers $\{x_j\}_{j=0}^{N-1}$ using the numbers $\{\hat{x}_j\}_{j=0}^{N-1}$ as

$$x_k := \mathfrak{D}^{-1}[\hat{x}]_k = \sum_{j=0}^{N-1} e^{\mathbf{i}jk\frac{2\pi}{N}} \hat{x}_j \tag{E.24}$$

for $k = 0, 1, ..., N-1$.

The DFT and inverse DFT operations can be represented in a matrix and/or vector form. In particular, the DFT takes the form

$$\begin{bmatrix} \hat{x}_0 \\ \vdots \\ \hat{x}_{N-1} \end{bmatrix} = \frac{1}{N} F \begin{bmatrix} x_0 \\ \vdots \\ x_{N-1} \end{bmatrix} \tag{E.25}$$

where F is a Vandermonde matrix,

$$F = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & \omega & \omega^2 & \dots & \omega^{N-1} \\ 1 & \omega^2 & \omega^4 & \dots & \omega^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \dots & \omega^{(N-1)(N-1)} \end{bmatrix} \tag{E.26}$$

with $\omega = e^{-\frac{2\pi}{N}\mathbf{i}}$. The inverse DFT is represented accordingly by

$$\begin{bmatrix} x_0 \\ \vdots \\ x_{N-1} \end{bmatrix} = \hat{F} \begin{bmatrix} \hat{x}_0 \\ \vdots \\ \hat{x}_{N-1} \end{bmatrix} \tag{E.27}$$

where $\hat{F}$ satisfies

$$\hat{F} = \left[\frac{1}{N}F\right]^{-1} = F^H. \tag{E.28}$$

i.e, $\hat{F}$ is the Hermitian transpose of F.

This matrix-vector representation is useful in many cases and especially in the proof of energy conservation properties of the DFT. The following proposition, known as the Parseval's theorem, is an illustration.

Proposition E.7. *Let the senquence* $\hat{\mathbf{x}} = \begin{bmatrix} \hat{x}_0 & \dots & \hat{x}_{N-1} \end{bmatrix}^*$ *be the discrete Fourier transform of* $\mathbf{x} = \begin{bmatrix} x_0 & \dots & x_{N-1} \end{bmatrix}^*$, *then*

$$|\hat{\mathbf{x}}|^2 = \frac{1}{N}|\mathbf{x}|^2. \tag{E.29}$$

Proof. Let $\|.\|_2$ be the induced Euclidean norm on complex matrices $A \in \mathbb{C}^{n\times m}$. We know that

$$\|A\|_2 = r(A^H A) \tag{E.30}$$

where, for a square matrix $M \in \mathbb{C}^{m\times m}$, $r(M)$ is the spectral radius of M.

Note that, from equation (E.25), we have

$$\begin{aligned} |\hat{\mathbf{x}}|^2 &\leq \frac{1}{N^2} \|F\|_2 |\mathbf{x}|^2 \\ &\leq \frac{1}{N^2} r(F^H F) |\mathbf{x}|^2 \\ &\leq \frac{1}{N^2} r(N I_{N\times N}) |\mathbf{x}|^2 \text{ (by equation (E.28))}, \\ &\leq \frac{1}{N} |\mathbf{x}|^2 . \end{aligned} \tag{E.31}$$

Also, from equation (E.27), we have

$$\begin{aligned} \frac{1}{N} |\mathbf{x}|^2 &\leq \frac{1}{N} \left\|\hat{F}\right\|_2 |\hat{\mathbf{x}}|^2 \\ &\leq \frac{1}{N} r(F^H F) |\hat{\mathbf{x}}|^2 \\ &\leq \frac{1}{N} r(N I_{N\times N}) |\hat{\mathbf{x}}|^2 \text{ (by equation (E.28))}, \\ &\leq |\hat{\mathbf{x}}|^2 . \end{aligned} \tag{E.32}$$

Both equations (E.31) and (E.32) then lead to the result. □

The DFT is strongly related to Fourier series and trigonometric approximation. If we consider a integrable real function $h : [a, b] \to \mathbb{R}$ such that $h(a) = h(b)$ and h admits the complex Fourier series expansion

$$h(x) = \sum_{k=-\infty}^{\infty} c_k e^{\mathbf{i}k\frac{2\pi}{b-a}x} \tag{E.33}$$

on the interval $[a, b]$. The DFT allows to approximate the coefficients c_k given a sampling of the function h at equidistant nodes through the following proposition that can be found in Plato [101].

Proposition E.8. *Let $\{h(x_k)\}_{k=0}^{N-1}$ be the values at equidistant nodes $\{x_k\}_{k=0}^{N-1}$ of a real function h. Assume that the function $h \in \mathcal{C}^2([a, b])$ is twice differentiable on $[a, b]$ with $h(a) = h(b)$ and $x_k = a + k\Delta$ where $\Delta = \frac{b-a}{N}$ where N is even. Then*

$$c_{k-\frac{N}{2}} = e^{-\mathbf{i}(k-\frac{N}{2})\frac{2\pi}{b-a}a} \mathfrak{D}[\{(-1)^i h(x_i)\}_{i=0}^{N-1}]_k + \mathcal{O}(\Delta^2) , \tag{E.34}$$

for $k = 0, 1, ..., N-1$.

Proof. Under the conditions of this proposition, the DFT is simply a composite trapezoidal quadrature which yields a second order error. □

Instead of a Fourier series truncation to approximate a function u, a better approach is an interpolation with trigonometric polynomials. For any even integer N, the real valued function $h : [a, b] \to \mathbb{R}$ can be interpolated by a trigonometric polynomial $T_N[h]$ of the form

$$T_N[h](x) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} d_k e^{\mathbf{i}k\frac{2\pi}{b-a}x} \tag{E.35}$$

at equidistant nodes $\{x_k\}_{k=0}^{N-1}$ such that

$$T_N[h](x_k) = h(x_k) \ , \ k = 0, 1, ..., N-1. \tag{E.36}$$

The Fourier coefficients $\{d_k\}_{k=-n}^{n-1}$ are exactly computed by a DFT.

Proposition E.9. *Supposed the trigonometric polynomial $T_N[h]$ interpolates the real valued function $h : [a,b] \to \mathbb{R}$ at the nodes $\{x_k\}_{k=0}^{N-1}$ with $x_0 = a$ and $x_i = x_{i-1} + \frac{b-a}{N}$ such that the relations of equation (E.36) hold. Then, the Fourier coefficients are given by*

$$e^{\mathbf{i}(k-\frac{N}{2})\frac{2\pi}{b-a}a} d_{k-\frac{N}{2}} = \mathfrak{D}[\{(-1)^i h(x_i)\}_{i=0}^{N-1}]_k \ , \tag{E.37}$$

for any $k = 0, ..., N-1$.

Proof. Let's note that for $i = 0, 1, ..., N-1$,

$$\begin{aligned} h(x_i) &= \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} d_k e^{\mathbf{i}k\frac{2\pi}{b-a}x_i} \\ &= \sum_{k=0}^{N-1} d_{k-\frac{N}{2}} e^{\mathbf{i}(k-\frac{N}{2})\frac{2\pi}{b-a}x_i} \\ &= \sum_{k=0}^{N-1} d_{k-\frac{N}{2}} e^{\mathbf{i}(k-\frac{N}{2})\frac{2\pi}{b-a}a} e^{\mathbf{i}ki\frac{2\pi}{N}} e^{-\mathbf{i}i\pi}, \end{aligned}$$

so that

$$(-1)^i h(x_i) = \mathfrak{D}^{-1}\left[\left\{d_{k-\frac{N}{2}} e^{\mathbf{i}(k-\frac{N}{2})\frac{2\pi}{b-a}a}\right\}_{k=0}^{N-1}\right]_i$$

since $e^{-\mathbf{i}i\pi} = (-1)^i$. Taking the DFT gives the result. □

An error bound is available for the trigonometric interpolation previously defined. It suffices to combine the result of Propositions E.6 and E.8 to have a proof of the following proposition.

Proposition E.10. *If the function $h \in \mathcal{C}^2([a,b])$ satisfies Assumtions E.1, then the trigonometric polynomial $T_N[h]$ interpolating h as defined in equations (E.35) and (E.36) satisfies*

$$h(x) = T_N[h](x) + \mathcal{O}(\Delta x) \tag{E.38}$$

for any $x \in [a,b]$ where $\Delta x = \frac{b-a}{N}$.

Let's notice that the error bound in Proposition E.10 is given in the $\mathbf{L}^1$-norm. A variation of this result exists with the $\mathbf{L}^2$-norm and the error term is, in that case, of second order under an additional square integrability condition on the function second derivative (Plato [101], Theorem 3.10, page 46). The previous proposition can be generalized for smooth functions. It can indeed be shown that for any (periodic) function $h \in \mathcal{C}^{m+1}([a,b])$, the trigonometric interpolation error is of order m.

Bibliography

[1] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions (10th ed).* U.S. National Bureau of Standards. Washington D.C., 1972.

[2] S. Ankirchner, P. Imkeller, and G. D. Reis. Classical and variational differentiability of BSDEs with quadratic growth. *Electronic Journal of Probability*, 12:1418–1453, 2007.

[3] S. Ankirchner, P. Imkeller, and G. D. Reis. Pricing and hedging of derivatives based on nontradable underlyings. *Mathematical Finance*, 20 (2):289–312, 2010.

[4] F. Antonelli. Backward-forward stochastic differential equations. *The Annals of Applied Probability*, 3 (3):777–793, 1993.

[5] L. Auslander, J. Johnson, and R. Johnson. Multidimensional Cooley-Tukey algorithms revisited. *Advances in Applied Mathematics*, 17 (4):477–519, 1996.

[6] D. H. Bailey and P. N. Swarztrauber. A fast method for the numerical evaluation of continuous Fourier and Laplace transforms. *SIAM J. Sci. Comput.*, 15 (5):1105–1110, 1994.

[7] V. Bally. Approximation scheme for solutions of BSDE. *Backward stochastic differential equations (N. El Karoui and L. Mazliak)*, 364:177–191, 1997.

[8] V. Bally and G. Pages. A quantization algorithm for solving multidimensional discrete-time optimal stopping problems. *Bernoulli*, 9(6):1003–1049, 2003.

[9] V. Bally, G. Pages, and J. Printems. A quantization tree method for pricing and hedging multidimensional American options. *Mathematical Finance*, 15(1):119–168, 2005.

[10] C. Bender and R. Denk. A forward scheme for backward SDEs. *Stochastic Processes and their Applications*, 117:1793–1812, 2007.

[11] C. Bender and J. Zhang. Time discretization and markovian iteration for coupled FBSDEs. *The Annals of Applied Probability*, 18(1):143–177, 2008.

[12] R. A. Bernatz. *Fourier Series and Numerical Methods for Partial Differential Equations.* John Wiley & Sons Inc. Hoboken, New Jersey, USA, 2010.

[13] G. G. Bilodeau. On the summability of series of Hermite polynomials. *Journal of Mathematical Analysis and Applications*, 8:406–422, 1664.

[14] G. G. Bilodeau. The Weierstrass transform and Hermite polynomials. *Duke Math. J.*, 29 (2):293–308, 1962.

[15] G. G. Bilodeau. The convolution transform and Appell polynomials. *Journal of Mathematical Analysis and Applications*, 34 (2):360–370, 1971.

[16] J. Bismut. Théorie probabiliste du contrôlle des diffusions. *Memoirs of the Amer. Math. Soc.*, 176, 1973.

[17] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. Political Economy*, 81:637–654, 1973.

[18] B. Bouchard and J.-F. Chassagneux. Discrete time approximation for continuously and discretely reflected BSDEs. *Stochastic Processes and their Applications*, 118:2269–2293, 2008.

[19] B. Bouchard, R. Elie, and N. Touzi. Discrete-time approximation of BSDEs and probabilistic schemes for fully nonlinear PDEs. *Radon Series Comp. Appl. Math.*, 8:1–34, 2009.

[20] B. Bouchard and N. Touzi. Discrete-time approximation and monte-carlo simulation of backward stochastic differential equations. *Stochastic Processes and their Applications*, 111:175–206, 2004.

[21] P. Briand, B. Delyon, and J. Memin. Donsker-type theorem for BSDEs. *Elect. Comm. in Probab.*, 6:1–14, 2001.

[22] P. Briand, B. Delyon, and J. Memin. On the robustness of backward stochastic differential equations. *Stochastic Processes and their Applications*, 97:229–253, 2002.

[23] P. Briand and Y. Hu. BSDE with quadratic growth and unbounded terminal value. *Probab. Theory Relat. Fields*, 136(4):604–618, 2006.

[24] P. Briand and Y. Hu. Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Relat. Fields*, 141:543–567, 2008.

[25] R. L. Burden and J. D. Faires. *Numerical Analysis (9th ed).* Brooks/Cole, Cengage Learning. Boston, MA, 2011.

[26] P. Carr and D. Madan. Option valuation using the Fast Fourier Transform. *J. Comp. Finance*, 2:61–73, 1999.

[27] J. Chassagneux and D. Crisan. Runge-Kutta schemes for BSDEs. *to appear in the Annals of Applied Probability.*

[28] D. Chevance. Numerical methods for backward stochastic differential equations. *Numerical methods in finance, Cambridge Univ. Press, Cambridge*, pages 232–244, 1997.

[29] J. W. Cooley, P. A. W. Lewis, and P. D. Welch. The fast Fourier transform and its applications. *IEEE Trans. on Education*, 12 (1):27–34, 1969.

[30] J. W. Cooley and J. W. Tukey. An algorithm for machine calculation of complex Fourier series. *Math. Comp.*, 19:297–301, 1965.

[31] D. Crisan and K. Manolarakis. Second order discretization of backward SDEs. *to appear in the Annals of Applied Probability.*

[32] D. Crisan and K. Manolarakis. Solving backward stochastic differential equations using the cubature method: Application to nonlinear pricing. *SIAM Journal on Financial Mathematics*, 3 (1):534–571, 2012.

[33] D. Crisan, K. Manolarakis, and N. Touzi. On the Monte Carlo simulation of BSDEs: An improvement on the Malliavin weights. *Stochastic Processes and their Applications*, 120 (7):1133–1158, 2010.

[34] J. Cvitanić and J. Ma. Hedging options for a large investor and forward-backward SDE's. *The Annals of Applied Probability*, 6(2):370–398, 1996.

[35] F. Delarue and S. Menozzi. A forward-backward stochastic algorithm for quasi-linear PDEs. *The Annals of Applied Probability*, 16(1):140–184, 2006.

[36] F. Delarue and S. Menozzi. An interpolated stochastic algorithm for quasi-linear PDEs. *Mathematics of Computation*, 77 (261):125–158, 2008.

[37] F. Delbaen, Y. Hu, and X. Bao. Backward SDEs with superquadratic growth. *Probab. Theory Relat. Fields*, 150:145–192, 2011.

[38] G. Di Nunno, B. Øksendal, and F. Proske. *Malliavin Calculus for Lévy Processes with Applications to Finance*. Springer-Verlag Berlin Heidelberg, Germany., 2009.

[39] Z. Ditzian. Generalized functions and convolution transforms. *Journal of Mathematical Analysis and Applications*, 26 (2):345–356, 1969.

[40] J. Douglas JR., J. Ma, and P. Protter. Numerical methods for forward-backward stochastic differential equations. *The Annals of Applied Probability*, 6:940–968, 1996.

[41] D. Duffie and L. G. Epstein. Stochastic differential utility. *Econometrica*, 60 (2):353–394, 1992.

[42] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68 (6):1343–1376, 2000.

[43] R. Edwards. *FourierSeries: A Modern Introduction (2nd ed)*. Graduate Text in Mathematics (64), Springer-Verlag. New York, USA, 1979.

[44] N. El Karou and L. Mazliak. *Backward Stochastic Differential Equations*. Addison Wesley Longman Ltd, England, 1997.

[45] N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng, and M.-C. Quenez. Reflected solutions of backward SDE and related obstacle problems for PDEs. *Annals of Probability*, 25:702–737, 1997.

[46] N. El Karoui, E. Pardoux, and M. Quenez. Reflected backward SDEs and American options. In L. C. G. Rogers and D. Talay, editors, *Numerical Methods in finance*, pages 215–231. Cambridge University Press, 1997.

[47] N. El Karoui, S. Peng, and M.-C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, 7 (1):1–71, 1997.

[48] N. El Karoui and M. Quenez. Imperfect markets and backward stochastic differential equations. In L. C. G. Rogers and D. Talay, editors, *Numerical Methods in finance*, pages 181–214. Cambridge University Press, 1997.

[49] L. C. Evans. *Partial Differential Equations (2nd ed)*. Graduate Studies in Mathematics (19), American Mathematical Society. Providence, Rhode Island, USA, 2010.

[50] J. Fourier. *Théorie Analytique de la Chaleur*. Didot Père et Fils Paris, France, 1822.

[51] W. Gautschi, G. H. Golub, and G. Opfer. *Applications and Computation of Orthogonal Polynomials*. Birkaüser Verlag. Bassel, Switzerland, 1999.

[52] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer. New York, USA, 2003.

[53] E. Gobet, J.-P. Lemor, and X. Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability*, 15:2172–2202, 2005.

[54] E. Gobet and A. Makhlouf. L_2-time regularity of BSDEs with irregular terminal functions. *Stochastic Processes and their Applications*, 120:1105–1132, 2010.

[55] G. R. Grimmett and D. R. Stirzaker. *Probability and Random Processes (3rd ed)*. Oxford University Press. New York, USA, 2001.

[56] R. Haberman. *Applied Partial Differential Equations (4th ed)*. Pearson. New Jesey, USA, 2003.

[57] M. Hegland. An implementation of multiple and multivariate Fourier transforms on vector processors. *SIAM J. Sci. Comput.*, 16 (2):271–288, 1995.

[58] F. B. Hildebrand. *Introduction to Numerical Analysis (2nd ed)*. Dover Publication Inc. New York, USA, 1974.

[59] E. Hille. A class of reciprocal functions. *Annals of Mathematics (Second Series)*, 27 (4):427–464, 1926.

[60] I. I. Hirshman and D. V. Widder. *The Convolution Transform*. Princeton University Press. Princeton, NJ, USA, 1955.

[61] Y. Hu, D. Nualart, and X. Song. Malliavin calculus for backward stochastic differential equations and application to numerical solutions. *The Annals of Applied Probability*, 21 (6):2379–2423, 2011.

[62] J. Hull and A. White. Pricing interest-rate derivative securities. *The Review of Financial Studies*, 3 (4):573–592, 1990.

[63] C. B. Hyndman. A forward-backward SDE approach to affine models. *Math. Finan. Econ.*, 2:107–128, 2009.

[64] C. B. Hyndman and P. Oyono Ngou. A convolution method for numerical solution of backward stochastic differential equations. *arXiv:1304.1783v1*, 2013.

[65] P. Imkeller and G. D. Reis. Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stochastic Processes and their Applications*, 120:348–379, 2010.

[66] G. W. Inverarity. Fast computation of multidimensional Fourier integrals. *SIAM J. Sci. Comput.*, 24(2):645–651, 2002.

[67] V. Karunakaran and T. Venugopal. The Weierstrass transform for a class of generalized functions. *Journal of Mathematical Analysis and Applications*, 220 (2):508–527, 1998.

[68] N. Kazamaki. *Continuous exponential martingales and BMO*. Lecture Notes in Mathematics (1579), Springer-Verlag. Berlin, Germany, 1994.

[69] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag. Berlin, Germany, 1992.

[70] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *The Annals of Probability*, 28 (2):558–602, 2000.

[71] R. Kress. *Numerical analysis*. Graduate Text in Mathematics (181), Springer-Verlag. New York, USA, 1998.

[72] J. Kuisalaas. *Numerical Methods In Engineering With Matlab*. Cambridge University Press. New York, USA, 2005.

[73] J. Lepeltier and J. S. Martin. Backward stochastic differential equations with continuous coefficient. *Statistics and Probability Letters*, 32:425–430, 1997.

[74] Y. Li and W. Zhao. l^p-error estimates for numerical schemes for solving certain kinds of backward stochastic differential equations. *Statistics and Probability Letters*, 80:1612–1617, 2010.

[75] F. Longstaff and E. Schwartz. Valuing American options by simulation: A simple least-squares approach. *Rev. Fin. Studies*, 14:113–147, 2001.

[76] R. Lord, F. Fang, F. Bervoets, and C. Osterlee. A fast and accurate FFT-based method for pricing early-exercise options under Lévy processes. *SIAM J. Sci. Comput.*, 30(4):1678–1705, 2008.

[77] J. Lucia and E. Schwartz. Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research*, 5:5–50, 2002.

[78] J. Ma, P. Protter, J. S. Martin, and S. Torres. Numerical method for backward stochastic differential equations. *The Annals of Applied Probability*, 12 (1):302–316, 2002.

[79] J. Ma, P. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitly - a four step scheme. *Probab. Theory Relat. Fields*, 98:339–359, 1994.

[80] J. Ma, J. Shen, and Y. Zhao. On numerical approximations of forward-backward stochastic differential equations. *Siam J. Numer. Anal.*, 46(5):2636–2661, 2008.

[81] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and their Applications*. Springer-Verlag Berlin Heidelberg. Berlin, Germany, 2007.

[82] J. Ma and J. Zhang. Path regularity for solutions of backward stochastic differential equations. *Probab. Theory Relat. Fields*, 122:163–190, 2002.

[83] J. Ma and J. Zhang. Representations and regularities for solutions of BSDEs with relections. *Stochastic Processes And Their Applications*, 115:539–569, 2005.

[84] D. Madan, F. Milne, and H. Shefrin. The multinomial option pricing model and its Brownian and Poisson limits. *Rev. Fin. Studies*, 2:251–265, 1990.

[85] X. Mao. Adapted solutions of backward stochastic differential equations with non-Lipschitz coefficients. *Stochastic Processes and their Applications*, 58:281–292, 1995.

[86] F. Marcellán and W. van Assche. *Orthogonal polynomials and special functions : computation and applications*. Lecture notes in mathematics, Springer-Verlag New York. New York, USA, 2006.

[87] M. Martinez, J. S. Martin, and S. Torres. Numerical method for reflected backward stochastic differential equations. *Stochastic Analysis and Applications*, 29 (6):1008–1032, 2011.

[88] R. C. Merton. Theory of rational pricing. *The Bell Journal of Economics and Management Science*, 4:141–183, 1973.

[89] G. N. Milstein and M. V. Tretyakov. Numerical algorithms for forward-backward stochastic differential equations. *Siam J. Sci. Comput.*, 28 (2):561–582, 2006.

[90] G. N. Milstein and M. V. Tretyakov. Discretization of forward-backward stochastic differential equations and related quasi-linear parabolic equations. *IMA Journal of Numerical Analysis*, 27:24–44, 2007.

[91] P. Moin. *Fundamentals of engineering numerical analysis (2nd ed)*. Cambridge University Press. New York, USA, 2010.

[92] D. Nualart. *The Malliavin Calculus and Related Topics*. Probability and Its Applications, Springer-Verlag. New york, USA, 1995.

[93] H. J. Nussbaumer. *Fast Fourier Transform and Convolution Algorithms (2nd ed)*. Springer Series in Information Sciences (2), Springer Berlin Heidelberg. Berlin, Germany, 1990.

[94] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications (6th ed)*. Springer-Verlag New York, 2003.

[95] P. Oyono Ngou. Alternative discretization of the convolution method for BSDEs. *Working paper*, 2013.

[96] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Letters*, 14 (1):55–61, 1990.

[97] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Lecture Notes in Control and Inform. Sci.*, 176:200–217, 1992.

[98] E. Pardoux and S. Tang. Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probab. Theory Relat. Fields*, 114:123–150, 1999.

[99] S. Peng and M. Xu. Numerical algorithms for backward stochastic differential equations with 1-d Brownian motion: Convergence and simulations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 45:335–360, 2011.

[100] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer-Verlag Berlin Heidelberg, Germany., 2009.

[101] R. Plato. *Concise Numerical Mathematics*. Graduate Studies in Mathematics (57), American Mathematical Society. Providence, Rhode Island, USA, 2003.

[102] P. Protter. *Stochastic Integration and Differential Equations (2nd ed)*. Applications of Mathematics: Stochastic Modelling and Applied Probability (21), Springer-Verlag. Berlin, Germany, 2004.

[103] A. Richou. Numerical simulation of BSDEs with drivers of quadratic growth. *The Annals of Applied Probability*, 21 (5):1933–1964, 2011.

[104] A. Richter. Explicit solutions to quadratic BSDEs and applications to utility maximization in multivariate affine stochastic volatility models. *arXiv:1201.2877*.

[105] R. Rouge and N. El Karou. Pricing via utility maximization and entropy. *Mathematical finance*, 10 (2):259–276, 2000.

[106] T. Sauer. *Numerical Analysis (2nd ed)*. Peason. Boston, MA, 2012.

[107] E. S. Schwartz. The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance*, 52:922–973, 1997.

[108] H. Stahl and V. Totik. *General Orthogonal Polynomials*. Cambridge University Press. New York, USA, 1992.

[109] G. Szegö. *Orthogonal Polynomials (3rd ed)*. Colloquium Publications (23), American Mathematical Society. Providence, Rhode Island, USA, 1967.

[110] A. Tveito and R. Winther. *Introduction to Partial Differential Equation: A Computational Approach*. Springer-Verlag New York. New York, USA, 1998.

[111] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5 (2), 1977.

[112] A. Vretblad. *Fourier Analysis and Its Applications*. Graduate Text in Mathematics (223), Springer-Verlag. New York, USA., 2003.

[113] D. V. Widder. Necessary and sufficient conditions for the representation of a function by a Weierstrass transform. *Trans. Amer. Math. Soc.*, 71:430–439, 1951.

[114] J. Yong. Finding adapted solutions of forward-backward stochastic differential equations: method of continuation. *Probab. Theory. Relat. Fields*, 107:537–572, 1997.

[115] E. Zauderer. *Partial Differential Equations of Applied Mathematics (2n ed)*. Wiley-Interscience Series in Pure and Applied Mathematics, John Wiley & Sons Inc. New York, USA, 1998.

[116] A. H. Zemanian. The convolution transformation of certain generalized functions and its inversion. *Bull. Amer. Math. Soc.*, 72:725–727, 1966.

[117] A. H. Zemanian. A generalised convolution transformation. *SIAM J. Appl. Math.*, 15:324–346, 1967.

[118] A. H. Zemanian. Generalised Weierstrass transformation. *SIAM J. Appl. Math.*, 15:1088–1105, 1967.

[119] A. H. Zemanian. The Weierstrass transformation of certain generalized functions. *Bull. Amer. Math. Soc.*, 73:682–684, 1967.

[120] A. H. Zemanian. *Generalized Integral Transformations (2n ed)*. Dover Publications. New York, USA, 1987.

[121] A. H. Zemanian and R. P. Tewarson. The numerical evaluation of distributional transforms. *SIAM J. Numer. Anal.*, 4:271–282, 1967.

[122] B. Zhang, L. Grzelak, and C. Oosterlee. Efficient pricing of commodity options with early-exercise under the ornstein-uhlenbeck process. *Applied Numerical Mathematics*, 62:91–111, 2012.

[123] J. Zhang. *Some Fine Properties of Backward Stochastic Differential Equations*. PhD thesis, Purdue University, 2001.

[124] J. Zhang. A numerical scheme for BSDEs. *The Annals of Applied Probability*, 14:459–488, 2004.

[125] W. Zhao, L. Chen, and S. Peng. A new kind of accurate numerical method for backward stochastic differential equations. *Siam J. Sci. Comput.*, 28(4):1563–1581, 2006.

[126] W. Zhao, J. Wang, and S. Peng. Error estimates of the θ-scheme for backward stochastic differential equations. *Discrete Contin. Dyn. Syst. Ser. B*, 12:905–924, 2009.

[127] W. Zhao, G. Zhang, and L. Ju. A stable multistep scheme for solving backward stochastic differential equations. *Siam J. Numer. Anal.*, 48(4):1369–1394, 2010.