Studies on Numerical Methods for Backward Stochastic Differential Equations with Spatially/Temporally Discrete Structures

Akihiro Kaneko

March 2024

Studies on Numerical Methods for Backward Stochastic Differential Equations with Spatially/Temporally Discrete Structures

A dissertation submitted to
The Graduate School of Engineering Science
Osaka University
in partial fulfillment of the requirements for the degree of
Doctor of Phillosophy in Science

BY
Akihiro Kaneko

March 2024

Abstract

This thesis is devoted to studying numerical methods for solving Markov backward stochastic differential equations (BSDEs) with discrete features.

Chapter 2 considers BSDEs driven by continuous-time finite state Markov chains (CTMCs), a class of BSDEs with discrete state space. The contributions in this chapter are divided into the following three results: (1) We construct multi-stage Euler-Maruyama methods for Markov BSDEs driven by CTMCs and observe that the methods are equivalent to exponential integrators for solving associated systems of ordinary differential equations (ODEs). (2) Motivated by the feature that exponential integrators avoid the stiffness of equations, we propose to use the multi-stage Euler-Maruyama methods for solving stiff BSDEs driven by CTMCs arising from spatial discretizations of BSDEs driven by Brownian motion. We also illustrate the effectiveness of the presented methods with a number of numerical experiments in which we treat nonlinear BSDEs arising from option pricing problems in finance.

Chapter 3 is an extension of the results presented in Chapter 2. The previous chapter considers the BSDEs with the terminal times being deterministic, whereas the target of this chapter is BSDEs with bounded stopping terminal times. In this case, multi-stage Euler-Maruyama methods for Markov BSDEs driven by CTMCs result in exponential integrators with a slight modification. Numerical experiments are also presented; there, nonlinear BSDEs arising from pricing barrier options are considered.

Chapter 4 is devoted to constructing a sparse grid-based multilevel spatial discretization for solving Markov BSDEs driven by Brownian motion. Utilizing the idea of the sparse grid combination technique, the method efficiently approximates high-dimensional Makov BSDEs (driven by Brownian motion) with a combination of multiple Markov BSDEs driven by CTMCs on grids with different resolutions. Similar to the previous chapters, we present numerical experiments for solving nonlinear BSDEs arising from option pricing problems, and both BSDEs with deterministic and bounded stopping terminal times are considered.

Chapter 5 considers BSDEs whose temporal structures are discrete, namely, back-

ward stochastic difference equations (BSΔEs). Because an arbitrary noise distribution is allowed for BSΔEs, the equations are advantageous to model various situations in discrete time. We construct a sparse grid-based numerical method for solving high-dimensional Markov BSΔEs. There, conditional expectations are approximated with sparse grid quadratures and the nestings of conditional expectations and nonlinear functions are approximated with sparse grid interpolants. These approximations result in a significant reduction of the computational cost. The performance is also confirmed through a simple example numerically.

Contents

Chapter 1

Introduction

1.1 Background

Backward stochastic differential equations (BSDEs) have been playing a powerful tool in various fields including optimal controls, partial differential equations and mathematical finance. They were originally introduced by Bismut [7, 8] in the context of stochastic control problems. Pardoux and Peng later introduced general nonlinear BSDEs driven by Brownian motions as noise process, typically written as

$$\mathcal{Y}_t = \xi + \int_t^T f(s, \mathcal{Y}_s, \mathcal{Z}_s)ds - \int_t^T \mathcal{Z}_s^* dW_s, \quad t \in [0, T]. \tag{1.1}$$

Here, $(W_t)_{t\in[0,T]}$ is a Brownian motion, $\int_t^T \mathcal{Z}_s^* dW_s$ is Itô's stochastic integral, and the data (ξ, f), a pair of a *terminal condition* ξ (a random variable) and a *driver* f (a function), is given in advance. A solution of (1.1) means a pair $(\mathcal{Y}_t, \mathcal{Z}_t)_{t\in[0,T]}$ of adapted processes that satisfies (1.1). Since the later 1990s, the study of BSDEs has been highly connected to mathematical finance. It has provided a multitude of research topics leading to the development of BSDE theory and its applications. For example, hedging derivative securities under nonlinear wealth process dynamics (e.g., different interest rates for borrowing and lending), dynamic risk measures and recursive utilities are successful applications of BSDEs. For details, see El Karoui et al. [45].

Numerical methods for solving such BSDEs have been developed in different directions. One may employ Euler-Maruyama methods to approximate them with a stochastic difference equation. Taking its conditional expectation and together with an approximation of conditional expectations, a numerical formula that is evaluated in a backward manner is obtained. For example, Gobet et al. [32] proposed least

square Monte-Carlo (LSMC) methods that approximates conditional expectations on a finite linear combination of (predetermined) basis functions. Another direction is connected to partial differential equations (PDEs): Solutions of BSDEs being *Markov*, written in the form

$$\mathcal{Y}_t = g(\mathcal{X}_T) + \int_t^T f(s, \mathcal{X}_s, \mathcal{Y}_s, \mathcal{Z}_s)ds - \int_t^T \mathcal{Z}_s dW_s, \tag{1.2}$$

$$\mathcal{X}_t = x_0 + \int_0^t \mu(s, \mathcal{X}_s)ds + \int_0^t \sigma(s, \mathcal{X}_s)dW_s, \tag{1.3}$$

can be represented using solutions of second order semilinear parabolic PDEs due to the nonlinear Feynman-Kac formula. Here, $\mu : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0,T] \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, $f : [0,T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $g : \mathbb{R}^d \to \mathbb{R}$ are deterministic functions, $x_0 \in \mathbb{R}^d$, and process $(\mathcal{X}_t)_{t\in[0,T]}$ is interpreted as the state variable of a system. Using the nonlinear Feynman-Kac formula, we can calculate solutions of BSDEs through PDE solvers such as finite difference methods or finite element methods. For details on such approach, we refer to Douglas et al. [27] and Milstein and Tretyakov [54].

BSDEs driven by Brownian motion and numerical methods for solving them have been studied in details from both theoretical and application points of view, but meanwhile, BSDEs with different features has also been considered and studied. In this thesis, we are interested in the following two types of BSDEs with "discrete structures".

BSDEs Driven by Continuos-Time Finite State Markov Chains Continuous-time finite state Markov chains (CTMCs) are stochastic processes that take values in discrete state space. In [15], Cohen and Elliott presented a result on the existence and uniqueness of solutions of BSDEs driven by CTMCs:

$$Y_t = \xi + \int_{]t,T]} h(t, Y_{s-}, Z_s)ds - \int_{]t,T]} Z_s^* dM_s, \quad t \in [0,T]. \tag{1.4}$$

Here, $(M_s)_{s\in[0,T]}$ is a martingale related to a CTMC $(X_s)_{s\in[0,T]}$ and the data (h, ξ) is given in advance. Further studies on such kind of BSDEs can be found in [16, 21, 20, 14], for example. In [21], a nonlinear Feynman-Kac type result is obtained; they revealed that solutions of Markov BSDEs driven by CTMCs written as

$$Y_t = X_T^* G + \int_{]t,T]} h(X_{s-}, s, Y_{s-}, Z_s)ds - \int_{]t,T]} Z_s^* dM_s,$$

can be represented using solutions of systems of ordinary differential equations (ODEs).

Backward Stochastic Difference Equations Backward stochastic difference equations (BSΔEs) are counterparts of BSDEs in discrete time. In [17], Cohen and Elliott considered BSΔEs driven by arbitrary finite-state processes:

$$Y_t = \xi + \sum_{s=t}^{T-1} F(s, Y_s, Z_{s+1}) - \sum_{s=t}^{T-1} Z_s(M_{s+1} - M_s), \quad t \in \{0, 1, \dots, T-1\}. \tag{1.5}$$

Here, the driving process M_s is allowed to an arbitrary martingale with independent increments. In the cases of the driving process being infinite-state, the corresponding BSΔEs are written as

$$Y_t = \xi + \sum_{s=t}^{T-1} F(s, Y_s, Z_{s+1}) - \sum_{s=t}^{T-1} Z_s(M_{s+1} - M_s) + N_s - N_T, \quad t \in \{0, 1, \dots, T-1\}, \tag{1.6}$$

where an additional orthogonal martingale N is required to be solved together with Y and Z; it is a direct consequencce of the Galtchouk-Kunita-Watanabe decomposition [29]. We can also choose M as an arbitrary (square integrable) martingale with independent increments. Compared to the continuous-time framework, BSΔEs can be used to model more various situations in applications. For details, see [17, 6].

1.2 Summary

In this thesis, we study and develop numerical methods for BSDEs driven by CTMCs and BSΔEs. As seen later, we can exploit their distinctive features to improve the effectiveness of the numerical computation. Before moving on to the next chapter, we present a summary of contributions.

Chapter 2

Numerical methods for computing the solutions of Markov BSDEs driven by CTMCs are explored. We construct multi-stage Euler-Maruyama methods for them and observe that they are equivalent to exponential integrators for solving an associated system of ODEs. Taking advantage of this observation, we propose to use these multi-stage Euler-Maruyama methods for effectively solving "stiff" Markov BSDEs driven by CTMCs arising from the spatial discretization of Markov BSDEs driven by Brownian motion. We also illustrate the effectiveness of the presented methods with several numerical experiments in which we treat nonlinear BSDEs arising from pricing problems of European options in finance.

Chapter 3

The results obtained in Chapter 2 are extended to Markov BSDEs with terminal times being bounded stopping times. Similarly to the case of BSDEs with deterministic terminal times, which we treated in the previous chapter, multi-stage Euler-Maruyama methods are equivalent to exponential integrators with slight modifications, and together with spatial discretizations, we can utilize them to solve BSDEs with bounded stopping terminal times driven by Brownian motion effectively. In numerical experiments, we focus on BSDEs driven by Brownian motion arising from pricing barrier options and confirm the efficiency of our methods.

Chapter 4

In Chapters 2 and 3, we proposed to apply spatial discretizations to BSDEs driven by Brownian motion and solve the obtained BSDEs driven by CTMCs using the multi-stage Euler-Maruyama method we constructed. With an argument based on the idea of sparse grid methods, we present a multilevel spatial discretization methods in which high-dimensional Markov BSDEs driven by Brownian motion are approximated a combination of multiple Markov BSDEs driven by CTMCs on grids with different resolutions. Through several numerical experiments, we illustrate the efficiency.

Chapter 5

Chapter 5 studies numerical methods for solving Markov BSΔEs. We focus on those whose state space is high dimensional and present sparse grid-based numerical methods for solving them. Specifically, we calculate conditional expectations appeared as integrals on high dimensional domains with sparse grid quadratures and replace nestings of them and nonlinear functions with sparse grid interpolations. The presented method can calculate solutions accurately and efficiently. We also present an error estimate and demonstrate them through a numerical experiment of a simple case.

Chapter 2

Multi-Stage Euler-Maruyama Methods for Backward Stochastic Differential Equations Driven by Continuous-Time Markov Chains

2.1 Introduction

2.1.1 Overview

In this chapter, we are interested in a different class of BSDEs, that is, (Markov) BSDEs driven by continuous-time Markov chains (CTMCs), written as

$$Y_t = X_T^* G + \int_{]t,T]} h(X_{s-}, s, Y_{s-}, Z_s) ds - \int_{]t,T]} dM_s^* Z_s,$$

and study numerical schemes for solving them. Here, $(X_t)_{t\in[0,T]}$ is a CTMC having a finite state space $\mathcal{I}$, $N \coloneqq \#\mathcal{I}$, $G \in \mathbb{R}^N$, $h : \mathcal{I} \times [0,T] \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$, and $(M_t)_{t\in[0,T]}$ is the associate martingale with X (ses Section 2.3 for the details). A "nonlinear Feynman-Kac type" formula for such Markov BSDEs driven by CTMCs has been derived in [21, 26]; the solutions of the BSDEs can be represented using solutions of the associated systems of ordinary differential equations (ODEs).

The main contributions of this chapter are summerized as follows.

1. We observe that Euler-Maruyama temporal discritization methods for solving a Markov BSDE driven by a CTMC is equivalent to exponential integrators [41] for solving the associated system of ODEs. (See Section 2.3.)

2. We introduce multi-stage Euler-Maruyama methods for efficiently solving "stiff" BSDEs driven by CTMCs. Together with a spatial discretization, they can be applied to solve BSDEs driven by Brownian motion. (See Section 2.3 and Section 2.4.)

2.1.2 Organization

The chapter is organized as follows: At the end of this section, we introduce notations frequently used in the chapter. In Section 2.2, we present preliminary results on BSDEs driven by CTMCs and ones driven by Brownian motion that are required for the subsequent arguments. In Section 2.3, we construct multi-stage Euler-Maruyama methods for BSDEs driven by CTMCs and observe that they are equivalent to exponential integrators, solvers that calculate stiff systems of ODEs successfully. Section 2.4 presents an application of the multi-stage Euler-Maruyama methods to BSDEs driven by Brownian motion, in which we present a concrete discretization and the resulting BSDE driven by a CTMC. Section 2.5 provides experiments highlighting the effectiveness of our schemes. Specifically, we treat option pricing problems under nonlinear wealth dynamics with several asset price process models, such as the Black-Scholes model and stochastic volatility models including the SABR model.

2.1.3 Notations

For $N \in \mathbb{N}$, e_i means the i-th unit vector in the Euclidean space $\mathbb{R}^N$ whose i-th element is 1. The notations $|\cdot|$ and $\|\cdot\|$ represent the absolute value and the Euclidean norm, respectively. Note that we will also use a stochastic seminorm represented as $\|\cdot\|_v$; for the definition, see (2.5). For any matrix Q, Q^* denotes the matrix transposition, Q^+ denotes the Moore-Penrose inverse, and $\mathrm{Tr}(Q)$ denotes the trace of Q. For any vector v, $\mathrm{diag}(v)$ is a diagonal matrix whose i-th diagonal element is $e_i^* v$. For any two vectors $v, w \in \mathbb{R}^N$, denote

$$v \leq w \iff e_i^* v \leq e_i^* w, \quad i = 1, \ldots, N.$$

We set as follows.

- $C([0,T] \times \mathbb{R})$ and $C(\mathbb{R})$ are the sets of $\mathbb{R}$-valued continuous functions defined on $[0,T] \times \mathbb{R}$ and $\mathbb{R}$, respectively.
- $C_b([0,T] \times \mathbb{R})$ and $C_b(\mathbb{R})$ are the sets of $\mathbb{R}$-valued bounded continuous functions defined on $[0,T] \times \mathbb{R}$ and $\mathbb{R}$, respectively.

- $C_b^2(\mathbb{R})$ is is the set of $\mathbb{R}$-valued, twice continuously differentiable functions u such that $\partial_x u$, $\partial_{xx} u$ as well as u are in $C_b(\mathbb{R})$.
- $C_b^{1,2}([0,T]\times\mathbb{R})$ is the set of $\mathbb{R}$-valued functions u, which is once continuously differentiable in its first argument, twice continuously differentiable in its second, and $\partial_t u$, $\partial_x u$, $\partial_{xx} u$ as well as u are in $C_b([0,T]\times\mathbb{R})$.

Here, $\partial_t u(t,x) \coloneqq \dfrac{\partial u}{\partial t}(t,x)$, $\partial_x u(t,x) \coloneqq \dfrac{\partial u}{\partial x}(t,x)$ and $\partial_{xx} u(t,x) \coloneqq \dfrac{\partial^2 u}{\partial x^2}(t,x)$. For a vector-valued càdlàg stochastic process X_t, X_{t-} denotes the left limit and $\Delta X_t \coloneqq X_t - X_{t-}$. Additinally, if X_t is a semimartingale, $\langle X, X\rangle$ denotes the predictable quadratic variation matrix. Throughout the chapter, we will work on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a finite time horizon $T > 0$. For $k, m \in \mathbb{N}$, a filtration $\mathbb{F}$ with the usual conditions, and a square-integrable càdlàg $\mathbb{F}$-martingale M, we define the following spaces of stochastic processes.

- $L^2(\mathbb{F}, \mathbb{R}^k)$ is the set of càdlàg $\mathbb{F}$-adapted processes $\mathcal{X} : [0,T]\times\Omega \to \mathbb{R}^k$ with $\mathbb{E}\left[\int_{[0,T]} \|\mathcal{X}_t\|^2 dt\right] < \infty$.
- $\mathbb{S}^2(\mathbb{F}, \mathbb{R}^k)$ is the set of càdlàg $\mathbb{F}$-adapted processes $\mathcal{Y} : [0,T]\times\Omega \to \mathbb{R}^k$ with $\mathbb{E}\left[\sup_{0\le t\le T} |\mathcal{Y}_t|^2\right] < \infty$.
- $L^2(\langle M\rangle, \mathbb{F}, \mathbb{R}^{k\times m})$ is the set of $\mathbb{F}$-predictable processes $\mathcal{Z} : [0,T]\times\Omega \to \mathbb{R}^{k\times m}$ with
$$\mathbb{E}\left[\left|\int_{[0,T]} \mathcal{Z}_t dM_t\right|^2\right] = \mathbb{E}\left[\int_{[0,T]} \mathrm{Tr}(\mathcal{Z}_t d\langle M, M\rangle_t \mathcal{Z}_t^*)\right] < \infty.$$

2.2 Setups and Preliminary Results

2.2.1 BSDEs driven by a CTMC

Let $X = (X_t)_{t\in[0,T]}$ be a continuous-time, finite-state Markov chain with state space $\mathcal{I} = \{e_1, \dots, e_N\}$, for some $N \in \mathbb{N}$. Suppose that X is defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{G})$ where $\mathbb{G} \coloneqq (\mathcal{G}_t)_{t\in[0,T]}$ is the completion of the filtration generated by X. Note that X is a càdlàg pure jump process in this case.

X is associated with a family of Q-matrices; recall that $N\times N$ matrices $Q_t, t \in [0,T]$ are called Q-matrices on $\mathcal{I}$ if $e_i^* Q_t e_j \ge 0$ for all $i \ne j$ and $\sum_j e_i^* Q_t e_j = 0$ for all i. We suppose Q-matrices appeared in this chapter are uniformly bounded

in time t. Note that some literature refer to its transpose as Q-matrix such as [28, 15]. We sometimes call Q_t as the transition rate matrix for X. For $p_t = (\mathbb{P}(X_t = e_1), \dots, \mathbb{P}(X_t = e_N))$ and $t \geq 0$, we see that it satisfies the following Kolmogorov's forward equation

$$\frac{dp_t}{dt} = p_t Q_t. \tag{2.1}$$

Hence, the transition probability matrix of X, given by

$$\Phi(t,s) \coloneqq \begin{pmatrix} \mathbb{P}(X_t = e_1 | X_s = e_1) & \dots & \mathbb{P}(X_t = e_N | X_s = e_1) \\ \vdots & \ddots & \vdots \\ \mathbb{P}(X_t = e_1 | X_s = e_N) & \dots & \mathbb{P}(X_t = e_N | X_s = e_N) \end{pmatrix}$$

for $t \geq s$ satisfies the following equations

$$\frac{d\Phi(t,s)}{dt} = \Phi(t,s) Q_t, \quad \Phi(s,s) = I, \tag{2.2}$$

$$\frac{d\Phi(t,s)}{ds} = -Q_s \Phi(t,s), \quad \Phi(t,t) = I, \tag{2.3}$$

for $t \geq s \geq 0$ where I is the $N \times N$ identity matrix. (2.2) and (2.3) are referred to as the forward and backward Kolmogorov equation, respectively. X is time-(in)homogeneous if Q_t does (not) depend on $t \in [0,T]$. Note also that the transition probability of the time-homogeneous chain X with a transition rate matrix Q is the matrix exponential $\Phi(t,s) = \exp((t-s)Q)$.

From Appendix B in [28], X has the following semi-martingale representation

$$X_t = x_0 + \int_{]0,t]} Q_s^* X_{s-} ds + M_t. \tag{2.4}$$

Here, $x_0 \in \{e_1, \dots, e_N\}$ and M_t is an $\mathbb{R}^N$-valued $\mathbb{G}$-martingale. The predictable quadratic covariation matrix of M is given by

$$\langle M, M \rangle_t = \int_{]0,t]} (\operatorname{diag}(Q_s^* X_{s-}) - \operatorname{diag}(X_{s-}) Q_s - Q_s^* \operatorname{diag}(X_{s-})) ds,$$

which is also shown in Appendix B in [28]. Let $\psi_t \coloneqq \operatorname{diag}(Q_t^* X_{t-}) - \operatorname{diag}(X_{t-}) Q_t - Q_t^* \operatorname{diag}(X_{t-})$. Note that ψ_t is a predictable process, valued in $N \times N$ real symmetric nonnnegative semi-definite matrices. For later use, define the seminorm for $z \in \mathbb{R}^N$ by

$$\|z\|_v^2 \coloneqq z^* (\operatorname{diag}(Q_t^* v) - \operatorname{diag}(v) Q_t - Q_t^* \operatorname{diag}(v)) z, \tag{2.5}$$

where $v \in \{e_1, \dots, e_N\}$. Note that $\|\cdot\|_v$ depends on t when Q_t is time-dependent. The following Itô's isometry is a key property of this seminorm for $v = X_{t-}$. That is, for any $\mathbb{R}^N$-valued predictable process Z, it holds

$$\mathbb{E}\left[\left(\int_{]s,t]} Z_u^* dM_u\right)^2\right] = \mathbb{E}\left[\int_{]s,t]} \|Z_u\|_{X_{u-}}^2 du\right] \text{ for } t > s \geq 0. \tag{2.6}$$

The proof is given in [17]. We also define the equivalence relation $Z \sim_M Z'$ on $\mathbb{R}^N$-valued predictable processes as $\|Z_t - Z'_t\|_{X_{t-}} = 0, dt \otimes d\mathbb{P}$-a.s.

In [15], Cohen and Elliott treat BSDEs driven by a CTMC in the form of

$$Y_t = \xi + \int_{]t,T]} h(t, Y_{s-}, Z_s) ds - \int_{]t,T]} dM_s^* Z_s, \tag{2.7}$$

where ξ is an $\mathcal{G}_T$-measurable square-integrable random variable, $h : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^N \ni (\omega, t, y, z) \mapsto h(\omega, t, y, z) \in \mathbb{R}$ is $\mathbb{G}$-predictable in (ω, t) and Borel measurable in (y, z). The following result on the existence and a uniqueness of the solution $(Y, Z) \in \mathbb{S}^2(\mathbb{G}, \mathbb{R}) \times L^2(\mathbb{G}, \langle M \rangle, \mathbb{R}^N)$ of (2.7) has been established.

Theorem 2.2.1 ([15]). *Assume that,*

$$\mathbb{E}\left[\int_{]0,T]} h(t, 0, 0)^2 dt\right] < \infty,$$

and that for some constant $L > 0$,

$$|h(t, y, z) - h(t, y', z')|^2 \leq L(|y - y'|^2 + \|z - z'\|_{X_{t-}}^2), \quad dt \otimes d\mathbb{P}\text{-a.s.} \tag{2.8}$$

for all $y, y' \in \mathbb{R}$ and $z, z' \in \mathbb{R}^N$. Then, it admits a unique solution $(Y, Z) \in \mathbb{S}^2(\mathbb{G}, \mathbb{R}) \times L^2(\mathbb{G}, \langle M \rangle, \mathbb{R}^N)$. We remark that it is unique up to indistinguishability for Y and up to $\sim_M$ equivalence for Z.

Remark 2.2.1. *Since $\|z\|_{e_i}^2 \leq 3 \max_{j,k=1,\dots,N} |e_j^* Q_t e_k| \cdot \|z\|^2$ for $z \in \mathbb{R}^N$, (2.8) leads to the usual Lipschitz continuity as $|h(t, y, z) - h(t, y', z')|^2 \leq L(|y - y'|^2 + \|z - z'\|^2)$, $dt \otimes \mathbb{P}$-a.s. Note that, however, the converse does not hold; there does not exist $C > 0$ such that $\|z\| \leq C\|z\|_{e_i}$ for all z. Taking $z_1 = (1, 1, \dots, 1) \in \mathbb{R}^N$, it is easy to see $\|z_1\|_{e_i}^2 = 0 < N = \|z_1\|^2$.*

Next, we consider the Markov BSDE driven by a CTMC of the form

$$Y_t = X_T^* G + \int_{]t,T]} h(X_{s-}, s, Y_{s-}, Z_s) ds - \int_{]t,T]} dM_s^* Z_s, \tag{2.9}$$

where $G \in \mathbb{R}^N$ and $h : \{e_1, \dots, e_N\} \times [0,T] \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ is a Borel measurable function. Associated with (2.9), setting $t \in [0,T]$ as the starting time for the BSDE, we consider

$$\begin{cases} X_s^{t,e_i} = e_i + \displaystyle\int_{]t,s]} Q_u^* X_{u-}^{t,e_i} du + M_s - M_t, \quad s > t, \\ X_s^{t,e_i} = e_i, \quad s \le t, \\ Y_s^{t,e_i} = (X_T^{t,e_i})^* G + \displaystyle\int_{]s,T]} h(X_{u-}^{t,e_i}, u, Y_{u-}^{t,e_i}, Z_u^{t,e_i}) du - \int_{]s,T]} dM_u^* Z_u^{t,e_i}, \quad s \in [0,T]. \end{cases} \tag{2.10}$$

Then, we give the following nonlinear Feynman-Kac type result. Recall that, similar statements can be found in [21, 26].

Theorem 2.2.2. *Assume that there exists a constant $L > 0$ such that*

$$|h(e_i,t,y,z) - h(e_i,t,y',z')|^2 \le L^2(|y-y'|^2 + \|z-z'\|_{e_i}^2), \tag{2.11}$$

for any $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^N$, $t \in [0,T]$ and $i = 1, \dots, N$, and $\int_0^T h(e_i,u,0,0)^2 du < \infty$. Define $H : [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$ such that

$$e_i^* H(t,z) = h(e_i, t, e_i^* z, z) \quad \text{for } t \in [0,T], z \in \mathbb{R}^N, i = 1, \dots, N.$$

1. *For a solution U_t of the system of ODEs*

$$\frac{dU_t}{dt} + Q_t U_t + H(t, U_t) = 0, \quad U_T = G, \tag{2.12}$$

$(Y_s^{t,e_i}, Z_s^{t,e_i}) = ((X_s^{t,e_i})^* U_s, U_s) \in \mathbb{S}^2(\mathbb{G}, \mathbb{R}) \times L^2(\mathbb{G}, \langle M \rangle, \mathbb{R}^N)$ uniquely solves (2.10).

2. *Conversely, for a unique solution $(Y_s^{t,e_i}, Z_s^{t,e_i}) \in \mathbb{S}^2(\mathbb{G}, \mathbb{R}) \times L^2(\mathbb{G}, \langle M \rangle, \mathbb{R}^N)$ of (2.10), a continuous function $V_t = (Y_t^{t,e_1}, \dots, Y_t^{t,e_N})^*$ satisfies $V \sim_M Z^{t,e_i}$ for $i = 1, \dots, N$ and $t \in [0,T]$, and solves (2.12).*

Proof. See Section 3.7.1. □

Remark 2.2.2. *If we assume the continuity of $t \mapsto h(e_i,t,y,z)$ for all i, y, z additionally, a uniqueness of (2.12) immediately holds from the well-known Picard-Lindelöf theorem (e.g. Theorem 110C, P.23 in [12]).*

Corollary 2.2.1. *Under the square integrability $t \mapsto h(e_i, t, y, z)$ in $[0, T]$ and the uniform Lipschitz continuity* (2.11)*, a unique solution (Y, Z) of* (2.9) *is also a unique solution of*

$$Y_t = X_T^* G + \int_{]t,T]} X_{s-}^* H(s, Z_s) ds - \int_{]t,T]} dM_s^* Z_s. \tag{2.13}$$

Moreover, the relation

$$Y_t = X_t^* U_t \quad \textit{up to indistinguishability and} \quad Z \sim_M U$$

holds, where U is a solution of (2.12)*.*

2.2.2 BSDEs Driven by a Brownian Motion

Let $W = (W_t)_{t \geq 0}$ be a d-dimensional standard Brownian motion. Let $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ be the completion of the filtration generated by W. We consider the following Markov BSDE driven by Brownian motion.

$$\begin{cases} \mathcal{X}_t = x_0 + \displaystyle\int_0^t \mu(s, \mathcal{X}_s) ds + \int_0^t \sigma(s, \mathcal{X}_s) dW_s, \\ \mathcal{Y}_t = g(\mathcal{X}_T) + \displaystyle\int_t^T f(s, \mathcal{X}_s, \mathcal{Y}_s, \mathcal{Z}_s) ds - \int_t^T \mathcal{Z}_s^* dW_s, \end{cases} \tag{2.14}$$

where $\mu : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$, $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, $g : \mathbb{R}^d \to \mathbb{R}$ are Borel measurable, and referred to as the drift coefficient, the diffusion coefficient, the driver and the terminal condition, respectively. Assuming that, there exists $L > 0$ and $p \in \mathbb{N}$ such that

$$\begin{gathered} \|\mu(t, x) - \mu(t, x')\| + \|\sigma(t, x) - \sigma(t, x')\| \leq L \|x - x'\|, \\ |f(t, x, y, z) - f(t, x, y', z')| \leq L(|y - y'| + \|z - z'\|), \\ \|\mu(t, x)\| + \|\sigma(t, x)\| \leq L(1 + \|x\|^2), \\ |f(t, x, y, z)| + |g(x)| \leq L(1 + \|x\|^p), \end{gathered} \tag{2.15}$$

for all $t \in [0, T]$, $x, x', z, z' \in \mathbb{R}^d$ and $y, y' \in \mathbb{R}$, (2.14) has a unique solution $((\mathcal{X}_t)_{t \in [0,T]}, (\mathcal{Y}_t)_{t \in [0,T]}, (\mathcal{Z}_t)_{t \in [0,T]}) \in \mathbb{S}^2(\mathbb{F}, \mathbb{R}^d) \times \mathbb{S}^2(\mathbb{F}, \mathbb{R}) \times L^2(\langle W \rangle, \mathbb{F}, \mathbb{R}^d)$. $\mathcal{X}$ is sometimes referred to as the state process, and it is solvable independently of $(\mathcal{Y}, \mathcal{Z})$.

The nonlinear Feynman-Kac formula describes the relation between (2.14) and

$$\begin{cases} \partial_t u(t, x) + \mathcal{L}_t u(t, x) + f(t, x, u(t, x), \sigma^*(x) \nabla_x u(t, x)) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}^d, \\ u(T, x) = g(x), \quad x \in \mathbb{R}^d. \end{cases} \tag{2.16}$$

Here,

$$\mathcal{L}_t u(t,x) = \sum_{i=1}^{d} \mu^{(i)}(t,x) \frac{\partial u}{\partial x_i}(t,x) + \frac{1}{2} \sum_{i,j=1}^{d} (\sigma\sigma^*)^{(i,j)}(t,x) \frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) \tag{2.17}$$

is the infinitesimal generator of the Markov process $\mathcal{X}$,

$$\nabla_x u(t,x) = \left(\frac{\partial u}{\partial x_1}(t,x), \ldots, \frac{\partial u}{\partial x_d}(t,x) \right)^* \in \mathbb{R}^d$$

is the gradient vector, $\mu^{(i)}(t,x)$ is the i-th component of $\mu(t,x)$, and $(\sigma\sigma^*)^{(i,j)}(t,x)$ is the (i,j)-th component of $\sigma(t,x)\sigma^*(t,x)$. The precise statement is as follows.

Theorem 2.2.3 (The nonlinear Feynman-Kac formula (e.g. pp.487-489 in [19]))**.** *Suppose that μ, σ, f and g are defined as above. For $(t,x) \in [0,T] \times \mathbb{R}^d$, let $(\mathcal{X}^{t,x}, \mathcal{Y}^{t,x}, \mathcal{Z}^{t,x})$ be a unique solution of the Markov BSDE*

$$\begin{cases} \mathcal{X}_s^{t,x} = x + \displaystyle\int_t^s \mu(\tau, \mathcal{X}_\tau^{t,x}) d\tau + \int_t^s \sigma(\tau, \mathcal{X}_\tau^{t,x}) dW_\tau \quad \textit{for} \quad s \geq t, \\ \mathcal{X}_s^{t,x} = x \in \mathbb{R}^d \quad \textit{for} \quad s \leq t, \\ \mathcal{Y}_s^{t,x} = g(\mathcal{X}_T^{t,x}) + \displaystyle\int_s^T f(\tau, \mathcal{X}_\tau^{t,x}, \mathcal{Y}_\tau^{t,x}, \mathcal{Z}_\tau^{t,x}) d\tau - \int_s^T (\mathcal{Z}_\tau^{t,x})^* dW_\tau \quad \textit{for} \quad s \in [0,T]. \end{cases} \tag{2.18}$$

Then,

1. *for every classical solution $u \in C^{1,2}([0,T] \times \mathbb{R}^d; \mathbb{R})$ of (2.16), such that, for some $K > 0$,*

$$|u(t,x)| + \|\nabla_x u(t,x)\| \leq K(1 + \|x\|) \quad \textit{for} \quad (t,x) \in [0,T] \times \mathbb{R}^d \tag{2.19}$$

a unique solution of BSDE (2.18) is represented as

$$\mathcal{Y}_s^{t,x} = u(s, \mathcal{X}_s^{t,x}), \quad \mathcal{Z}_s^{t,x} = \sigma^*(\mathcal{X}_s^{t,x}) \nabla_x u(s, \mathcal{X}_s^{t,x}) \quad \textit{for} \quad s \geq t. \tag{2.20}$$

(The inequality (2.19) is sufficient for showing $(\mathcal{Y}_s^{t,x}, \mathcal{Z}_s^{t,x})$ is of the class $\mathbb{S}^2(\mathbb{F}, \mathbb{R}) \times L^2(\langle W \rangle, \mathbb{F}, \mathbb{R}^d)$.)

2. *Suppose further that f and g are Lipschitz continuous and uniformly continuous with respect to x, uniformly in t. Then, $u(t,x) := \mathcal{Y}_t^{t,x}$ is a viscosity solution of (2.16).*

3. *Additionally, if for each $R > 0$ there exists a continuous function $m_R : [0, \infty) \to [0, \infty)$ such that $m_R(0) = 0$, and*

$$|f(t, x, y, z) - f(t, x', y, z)| \leq m_R(\|x - x'\|(1 + \|z\|))$$

holds for $y \in \mathbb{R}, x, x', z \in \mathbb{R}^d$ such that $\max\{\|x\|, \|x'\|, \|z\|\} < R$, then a uniqueness of u also holds.

Remark 2.2.3. *Viscosity solution is a weak solution of PDEs, which allows us to regard a continuous function, not smooth enough, as a solution. It coincides with the classical solution if the coefficients of PDEs satisfy some regularity conditions. For the details, see [19, 75, 60, 23] and references therein.*

2.3 Multi-Stage Euler-Maruyama Methods

In this section, we introduce several (multi-stage) Euler-Maruyama methods for solving BSDEs driven by CTMCs. Hereafter, we always assume that $(X_t)_{t\in[0,T]}$ is time-homogeneous, namely, the transition rate matrix Q_t equals some constant matrix Q for all t. The transition expectation is then represented as the action of a matrix exponential on the present state X_t. This can be seen from

$$\mathbb{E}[X_s|X_t] = \sum_{i=1}^{N} e_i \mathbb{P}(X_s = e_i|X_t) = \begin{pmatrix} \mathbb{P}(X_s = e_1|X_t) \\ \vdots \\ \mathbb{P}(X_s = e_N|X_t) \end{pmatrix} = e^{(s-t)Q} X_t, \tag{2.21}$$

for all $t \leq s \leq T$.

Euler-Maruyama methods are constructed in the following two steps:

1. Slice the time interval $[0, T]$ into a temporal grid $\{0 = t_0 < t_1 < \cdots < t_M = T\}$ and derive a stochastic difference equation on the grid.

2. Take conditional expectations and suitably approximate the (Riemann) integral part that appeared.

Let $(Y, Z) = (X^* Z, Z)$ be a unique solution of a BSDE driven by a CTMC (2.13). Discretize $[0, T]$ on a uniform grid $t_m = m\Delta t$ for $m = 0, 1, \ldots, M$, where $\Delta t = T/M$. We immediately see that $(Y_{t_m})_{m=0}^M$ satisfies the following stochastic difference equation

$$Y_{t_m} = Y_{t_{m+1}} + \int_{]t_m, t_{m+1}]} X^*_{s-} H(s, Z_s) ds - \int_{]t_m, t_{m+1}]} dM^*_s Z_s, \tag{2.22}$$

for $m = 0, 1, \ldots, M-1$. Taking the conditional expectations $\mathbb{E}[\cdots|X_{t_m}]$ in both-hand sides of (2.22), we observe that

$$Y_{t_m} = \mathbb{E}[X_{t_{m+1}}|X_{t_m}]^* Z_{t_{m+1}} + \int_{t_m}^{t_{m+1}} \mathbb{E}[X_s|X_{t_m}]^* H(s, Z_s) ds. \tag{2.23}$$

We can consider the following Euler-Maruyama-type approximations from (2.23):

(The Lawson-Euler Method). The simplest would be

$$Y_{t_m} \approx \mathbb{E}[X_{t_{m+1}}|X_{t_m}]^* Z_{t_{m+1}} + \Delta t \mathbb{E}[X_{t_{m+1}}|X_{t_m}]^* H(t_{m+1}, Z_{t_{m+1}}). \tag{2.24}$$

From (2.21), (2.24) is reduced to

$$Y_{t_m} \approx X_{t_m}^* \left(e^{\Delta t Q} Z_{t_{m+1}} + \Delta t e^{\Delta t Q} H(t_{m+1}, Z_{t_{m+1}}) \right). \tag{2.25}$$

As a consequence, we have the following 1-stage Euler-Maruyama scheme.

$$\begin{cases} Z_{t_M}^M \coloneqq G, \\ Z_{t_m}^M \coloneqq e^{\Delta t Q}(Z_{t_{m+1}}^M + \Delta t H(t_{m+1}, Z_{t_{m+1}}^M)), \quad m = 0, 1 \ldots, M-1, \end{cases} \tag{2.26}$$

which is known as the Lawson-Euler method [47] for solving system of ODEs (2.12). Note that we take $Y_{t_m}^M \coloneqq X_{t_m}^* Z_{t_m}^M$ for $m = 0, 1 \ldots, M$.

(The Nørsett-Euler Method). We can consider another (1-stage) Euler-Maruyama approximation as

$$Y_{t_m} \approx \mathbb{E}[X_{t_{m+1}}|X_{t_m}]^* Z_{t_{m+1}} + \left(\int_{t_m}^{t_{m+1}} \mathbb{E}[X_s|X_{t_m}]^* ds \right) H(t_{m+1}, Z_{t_{m+1}}). \tag{2.27}$$

From (2.21) as before, (2.27) is reduced

$$Y_{t_m} \approx X_{t_m}^* \left(e^{\Delta t Q} Z_{t_{m+1}} + \Delta t \left(\int_0^1 e^{(1-\theta)\Delta t Q} d\theta \right) H(t_{m+1}, Z_{t_{m+1}}) \right), \tag{2.28}$$

and we have

$$\begin{cases} Z_{t_M}^M \coloneqq G, \\ Z_{t_m}^M \coloneqq e^{\Delta t Q} Z_{t_{m+1}}^M + \Delta t \left(\int_0^1 e^{(1-\theta)\Delta t Q} d\theta \right) H(t_{m+1}, Z_{t_{m+1}}^M), \quad m = 0, 1 \ldots, M-1. \end{cases} \tag{2.29}$$

It is known as the Nørsett-Euler method [22] for solving system of ODEs (2.12).

(The ETD2RK Method). We can consider 2-stage Euler-Maruyama methods, as well as 1-stage ones. For example, a 2-stage Nørsett-Euler-Maruyama method is described as

$$\begin{aligned} Y_{t_m} &\approx \mathbb{E}[X_{t_{m+1}}|X_{t_m}]^* Z_{t_{m+1}} + \int_{t_m}^{t_{m+1}} \mathbb{E}[X_s|X_{t_m}]^* \\ &\qquad \left(H(t_{m+1}, Z_{t_{m+1}}) + \frac{H(t_m, Z_{t_m}) - H(t_{m+1}, Z_{t_{m+1}})}{\Delta t}(t_{m+1} - s) \right) ds \\ &\approx \mathbb{E}[X_{t_{m+1}}|X_{t_m}]^* Z_{t_{m+1}} + \left(\int_{t_m}^{t_{m+1}} \mathbb{E}[X_s|X_{t_m}]^* ds \right) H(t_{m+1}, Z_{t_{m+1}}) \\ &\qquad + \left(\int_{t_m}^{t_{m+1}} (t_{m+1} - s)\mathbb{E}[X_s|X_{t_m}]^* ds \right) \frac{H(t_m, \zeta_m) - H(t_{m+1}, Z_{t_{m+1}})}{\Delta t}, \end{aligned}$$

where ζ_m is defined by

$$\zeta_m = \mathbb{E}[X_{t_{m+1}}|X_{t_m}]^* Z_{t_{m+1}} + \left(\int_{t_m}^{t_{m+1}} \mathbb{E}[X_s|X_{t_m}]^* ds \right) H(t_{m+1}, Z_{t_{m+1}}).$$

From (2.21), it results in

$$\begin{cases} Z_{t_M}^M \coloneqq G, \\ Z_{t_m}^M \coloneqq e^{\Delta t Q} Z_{t_{m+1}}^M + \Delta t \left(\displaystyle\int_0^1 e^{(1-\theta)\Delta t Q} d\theta - \int_0^1 e^{(1-\theta)\Delta t Q} \theta d\theta \right) H(t_{m+1}, Z_{t_{m+1}}^M) \\ \qquad\qquad + \Delta t \left(\displaystyle\int_0^1 e^{(1-\theta)\Delta t Q} \theta d\theta \right) H(t_{m+1}, \zeta_m^M) \\ \zeta_m^M \coloneqq e^{\Delta t Q} Z_{t_{m+1}}^M + \Delta t \left(\displaystyle\int_0^1 e^{(1-\theta)\Delta t Q} d\theta \right) H(t_{m+1}, Z_{t_{m+1}}^M), \quad m = 0, 1 \ldots, M-1, \end{cases} \tag{2.30}$$

which is known as the second-order exponential time differencing Runge-Kutta (ETD2RK) method [22] for solving system of ODEs (2.12).

(General s-Stage Exponential Integrators). Furthermore, we consider general s-stage Euler-Maruyama methods taking the form of

$$\begin{aligned} Z_{t_m}^M &= \chi_0(\Delta t Q) Z_{t_{m+1}}^M + \Delta t \sum_{i=1}^{s} b_i(\Delta t Q) G_{mi}, \\ G_{mi} &= H(t_{m+1} - c_i \Delta t, \zeta_{ni}^M), \quad \text{for} \quad i = 1, \dots, s, \\ \zeta_{mi}^M &= \chi_i(\Delta t Q) Z_{m+1}^M + \Delta t \sum_{j=1}^{s} a_{ij}(\Delta t Q) G_{nj}, \quad \text{for} \quad i = 1, \dots, s, \end{aligned} \tag{2.31}$$

for $m = 0, \dots, M-1$. Here, $Z_{t_m}^M$ approximates Z_{t_m}, G_{mi} is the i-th internal stage, $s \in \mathbb{N}$ is the number of stages, c_i are real numbers, and χ_i, a_{ij} and b_i are functions constructed from "ϕ-functions" defined by

$$\phi_l(Q) = \int_0^1 e^{(1-\theta)Q} \frac{\theta^{l-1}}{(l-1)!} d\theta \quad \text{for} \quad l \in \mathbb{N}, \quad \text{and} \quad \phi_0(Q) = e^Q.$$

For example, the three methods mentioned above are obtained from the settings:

- Lawson-Euler: $s = 1$, $\chi_0(z) = e^z$, $\chi_1(z) = 1$, $a_{11}(z) = 0$, $b_1(z) = e^z$ and $c_1 = 0$.
- Nørsett-Euler: $s = 1$, $\chi_0(z) = e^z$, $\chi_1(z) = 1$, $a_{11}(z) = 0$, $b_1(z) = \phi_1(z)$ and $c_1 = 0$.
- ETD2RK: $s = 2$, $\chi_0(z) = e^z$, $\chi_1(z) = 1$, $\chi_2(z) = e^z$, $a_{11}(z) = a_{12}(z) = a_{22}(z) = 0$, $a_{21}(z) = \phi_1(z)$, $b_1(z) = \phi_1(z) - \phi_2(z)$, $b_2(z) = \phi_2(z)$, $c_1 = 0$ and $c_2 = 1$.

Note that the multi-stage Euler-Maruyama method (2.31) is the same as exponential integrators (exponential Runge-Kutta methods) for solving systems of ODEs (2.12). For details on exponential integrators, we refer to [41], a comprehensive survey. We remark on the following:

Remark 2.3.1. *It is known that using exponential integrators for numerical calculations of "stiff" systems of ODEs work well. Here, a differential equation is called stiff if explicit methods become numerically unstable unless the step size is taken to be extremely small. It seems to be difficult to formulate a precise mathematical definition of stiffness, but such an equation includes some terms that can lead to rapid variation in the solution. As for the historical development of the notion of stiffness, we refer to [37, 67]. As described in [37, 68, 69, 70, 73, 12], it is known that the method-of-lines approach for solving parabolic PDEs (that is, the discretization of the spatial variable) often results in large stiff systems of ODEs, and exponential integrators are considered to be effective for these systems as seen in [38, 39, 55], for example.*

Remark 2.3.2. *χ_i, a_{ij}, b_i and c_i in* (2.31) *are prescribed parameters to be set so that we can obtain various schemes that have (stiff / nonstiff) orders of convergence.*

Remark 2.3.3. *Exponential integrators exploits matrix functions ϕ_l. However, evaluating them numerically is not straightforward and has been studied in numerical literature. A standard approach that is widely used is a combination of Padè approximations and scaling-and-squaring methods. Although it enables efficient evaluation, note that it is only applicable to ϕ_l of moderate dimension. For solving ODE systems whose dimension is large, it is advantageous to apply Krylov subspace methods; instead of evaluating ϕ_l itself, its action on a state vector is approximated with a vector on a Krylov subspace whose dimension is small.*

2.4 Application to BSDEs Driven by Brownian Motion

In this section, we are interested in computing Markov BSDEs driven by Brownian motion (2.14) with an appropriate spatial discretization. From a probabilistic point of view, it can be seen as approximating a BSDE driven by a Brownian motion with a BSDE driven by a CTMC. From a differential equation point of view, on the other hand, it can be seen as the method of lines, approximating a second-order parabolic PDE with a system of ODEs. As mentioned in remark 2.3.1, the method of lines discretization of parabolic PDEs leads to stiff systems of ODE, and our multi-stage Euler-Maruyama methods efficiently work for them. One can represent this situation as Figure 2.1.

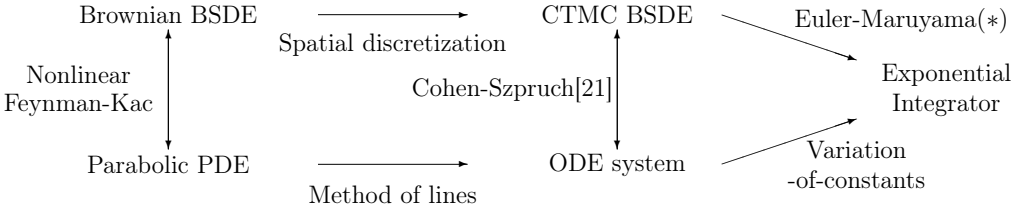


Figure 2.1: A diagram representing the relation between key ingredients for the argument up to here on this chapter. It is based on the diagram given in [21]. The arrow denoted as $(*)$ has been newly drawn by the arguments in section 2.3.

We present a construction of Markov BSDEs driven by CTMCs from spatial discretization of Markov BSDEs driven by Brownian motion. Through the method of lines discretization to the associated second-order parabolic PDEs, systems of

ODEs are obtained. We see that the systems of ODEs are equivalent to Markov BSDEs driven by CTMCs. Hereafter, we focus on (2.14) such that $(\mathcal{X}_t)_{t\in[0,T]}$ is time homogeneous, namely, $\mu(t,x)$ and $\sigma(t,x)$ do not depend on t. Then, we can simply write $\mu(t,x)=\mu(x)$ and $\sigma(t,x)=\sigma(x)$, and the subscript of the infinitesimal generator can also be omitted: $\mathcal{L}_t=\mathcal{L}$.

2.4.1 The Case of 1-dimensional State Space

First, we shall discuss the case of $d=1$ for simplicity. Let a strictly increasing sequence $\Pi=\{x_i\}_{i=-N_0}^{N_0}$ of length $N:=2N_0+1$ be the set of nodes on $\mathbb{R}$ and define $\delta x_i := x_{i+1}-x_i$ for $-N_0\le i<N_0$. For any function $v:\mathbb{R}\to\mathbb{R}$, let $v^\Pi=(v(t,x_{-N_0}),v(t,x_{-N_0+1}),\dots,v(t,x_{N_0}))^*\in\mathbb{R}^N$ be the evaluation of v over Π. Then, derivatives evaluated at nodes of Π are replaced by

$$\begin{pmatrix}\dfrac{\partial v}{\partial x}(x_{-N_0})\\ \vdots \\ \dfrac{\partial v}{\partial x}(x_{N_0}))\end{pmatrix}\approx D_1u^\Pi,\quad \begin{pmatrix}\dfrac{\partial^2 v}{\partial x^2}(x_{-N_0})\\ \vdots \\ \dfrac{\partial^2 v}{\partial x^2}(x_{N_0}))\end{pmatrix}\approx D_2u^\Pi.$$

Here, $N\times N$-matrices D_1 and D_2 are defined by

$$e_i^*D_1e_j=\begin{cases}\dfrac{-\delta x_i}{\delta x_{i-1}(\delta x_{i-1}+\delta x_i)}, & j=i-1,\\ \dfrac{\delta x_i-\delta x_{i-1}}{\delta x_i\delta x_{i-1}}, & j=i,\\ \dfrac{\delta x_{i-1}}{\delta x_i(\delta x_{i-1}+\delta x_i)}, & j=i+1,\\ 0 & \text{otherwise,}\end{cases}\quad \text{for}\quad -N_0<i<N_0,$$
$$e_{-N_0}^*D_1e_i=e_{N_0}^*D_1e_i=0,\quad \text{for}\quad -N_0\le i\le N_0, \tag{2.32}$$

$$e_i^*D_2e_j=\begin{cases}\dfrac{2}{\delta x_{i-1}(\delta x_{i-1}+\delta x_i)}, & j=i-1,\\ \dfrac{-2}{\delta x_i\delta x_{i-1}}, & j=i,\\ \dfrac{2}{\delta x_i(\delta x_{i-1}+\delta x_i)}, & j=i+1,\\ 0 & \text{otherwise,}\end{cases}\quad \text{for}\quad -N_0<i<N_0,$$
$$e_{-N_0}^*D_2e_i=e_{N_0}^*D_2e_i=0,\quad \text{for}\quad -N_0\le i\le N_0, \tag{2.33}$$

that result from the central difference scheme, and we denote e_i as the $(i+N_0+1)$-th unit vector in $\mathbb{R}^N$ whose $(i+N_0+1)$-th element is 1. We then solve, in place of PDE (2.16), a system of N ODEs in what follows:

$$\frac{dU_t^\Pi}{dt} + QU_t^\Pi + F(t, U_t^\Pi) = 0 \quad \text{for} \quad t \in [0,T], \quad U_T^\Pi = G. \tag{2.34}$$

Here, $G = (g(x_{-N_0}), \dots, g(x_{N_0}))^* \in \mathbb{R}^N$, $F : [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$ is defined by

$$F(t,z) = \begin{pmatrix} f(t, x_{-N_0}, e_{-N_0}^* z, \sigma(x_{-N_0}) e_{-N_0}^* D_1 z) \\ \vdots \\ f(t, x_{N_0}, e_{N_0}^* z, \sigma(x_{N_0}) e_{N_0}^* D_1 z) \end{pmatrix}, \tag{2.35}$$

and

$$Q = \operatorname{diag}(\mu^\Pi) D_1 + \frac{1}{2} \operatorname{diag}((\sigma^2)^\Pi) D_2$$

approximates the infinitesimal generator $\mathcal{L}$ of $\mathcal{X}$. Using (2.32) and (2.33), each element of Q is

$$e_i^* Q e_j = \begin{cases} \dfrac{\sigma^2(x_i) - \delta x_i \mu(x_i)}{\delta x_{i-1}(\delta x_{i-1} + \delta x_i)}, & j = i-1, \\ \dfrac{(\delta x_i - \delta x_{i-1})\mu(x_i) - \sigma^2(x_i)}{\delta x_i \delta x_{i-1}}, & j = i, \\ \dfrac{\sigma^2(x_i) + \delta x_{i-1}\mu(x_i)}{\delta x_i(\delta x_{i-1} + \delta x_i)}, & j = i+1, \\ 0 & \text{otherwise}, \end{cases} \quad \text{for} \quad -N_0 < i < N_0, \tag{2.36}$$

$$e_{-N_0}^* Q e_i = e_{N_0}^* Q e_i = 0, \quad \text{for} \quad -N_0 \le i \le N_0.$$

In Section 2.8, we give a convergence result of (2.34) to (2.16) in a case which a unique classical solution of (2.16) exists.

Theorem 2.4.1 (Convergence)**.** *Consider the case of $d = 1$ and take a spatial grid $\Pi^{(N,\Delta x)} := \{i\Delta x\}_{i=-N_0}^{N_0}$ for some $\Delta x > 0$. Suppose that Assumption 2.8.1, 2.8.2 in Section 2.8, and* (2.16) *admits a unique solution u. (For example,* (2.16) *is uniquely solvable in the case of Lemma 2.8.1.) Denote $U_t^{(N,\Delta x)}$ as a unique solution of* (2.34) *in the case of $\Pi^{(N,\Delta x)}$. For any compact set $K \subset \mathbb{R}$, it holds*

$$\lim_{\Delta x \to 0} \lim_{N \to \infty} \sup_{\substack{-N_0 \le i \le N_0 \\ i \in \mathbb{Z}, i\Delta x \in K}} |u(t, i\Delta x) - e_i^* U_t^{(N,\Delta x)}| = 0,$$

where e_i is the $(i+N_0+1)$-th unit vector in $\mathbb{R}^N$ whose $(i+N_0+1)$-th element is 1.

Thus, for obtaining the system of ODE which approximate the PDE with a small error in given bounded K, we should take a sufficiently small $k > 0$ and expand the spatial grid for sufficiently large N that depends on k.

2.4.2 The Case of d-dimensional State Space

Using the Kronecker product "$\otimes$", the argument in the case of $d = 1$ can be carried out in multidimensional cases, that is, for any $d \in \mathbb{N}$. For $p = 1, \dots, d$, let a strictly increasing sequence $\Pi^{(p)} = (x_i^{(p)})_{i=-N_0^{(p)}}^{N_0^{(p)}}$ of length $N^{(p)} = 2N_0^{(p)} + 1$ be the set of nodes on the p-th axis in $\mathbb{R}^d$, and let $D_1^{(p)}$ and $D_2^{(p)}$ are the corresponding $N^{(p)} \times N^{(p)}$ difference matrices defined by (2.32) and (2.33), constructed on $\Pi^{(p)}$. Consider the grid on $\mathbb{R}^d$ by

$$\Pi = \Pi^{(1)} \otimes \Pi^{(2)} \otimes \cdots \otimes \Pi^{(d)} = (x_i = (x_{i_1}^{(1)}, x_{i_2}^{(2)}, \dots, x_{i_d}^{(d)}) : i = 1, \dots, N),$$

where $N \coloneqq \prod_{p=1}^{d} N^{(p)}$ is the total size of Π, and multi-indices $(i_1, i_2, \dots, i_d)$ are ordered lexicographically. For $v : \mathbb{R}^d \to \mathbb{R}$, first and second derivatives along the p-th axis are approximated by

$$\frac{\partial v}{\partial x^{(p)}}(x_i) \approx e_i^* \widetilde{D}_1^{(p)} v^\Pi \quad \text{and} \quad \frac{\partial^2 v}{(\partial x^{(p)})^2}(x_i) \approx e_i^* \widetilde{D}_2^{(p)} v^\Pi,$$

where matrix $\widetilde{D}_k^{(p)}$ for $k = 1, 2$ and $p = 1, \dots, d$ is given by

$$\widetilde{D}_k^{(p)} \coloneqq I_{N^{(1)}} \otimes \cdots \otimes I_{N^{(p-1)}} \otimes D_k^{(p)} \otimes I_{N^{(p+1)}} \otimes \cdots \otimes I_{N^{(d)}}.$$

In multidimensional cases, we additionally need to specify the approximation of cross derivatives since $\mathcal{L}$ possibly contains them. In this work, we approximate the cross derivative along the p-th and q-th axes as

$$\frac{\partial^2 v}{\partial x^{(p)} \partial x^{(q)}}(x_i) \approx e_i^* D_1^{(p,q)} v^\Pi,$$

where

$$D_1^{(p,q)} \coloneqq I_{N^{(1)}} \otimes \cdots \otimes I_{N^{(p-1)}} \otimes D_1^{(p)} \otimes I_{N^{(p+1)}} \otimes \cdots \otimes I_{N^{(q-1)}} \otimes D_1^{(q)} \otimes I_{N^{(q+1)}} \otimes \cdots \otimes I_{N^{(d)}},$$

for $p < q$. In this situation, Q is defined by

$$Q = \sum_{p=1}^{d} \mathrm{diag}((\mu^{(p)})^{\Pi})\widetilde{D}_1^{(p)} + \sum_{p=1}^{d-1}\sum_{q=p+1}^{d} \mathrm{diag}(((\sigma\sigma^*)^{(p,q)})^{\Pi})\widetilde{D}_1^{(p,q)} + \frac{1}{2}\sum_{p=1}^{d} \mathrm{diag}(((\sigma\sigma^*)^{(p,p)})^{\Pi})\widetilde{D}_2^{(p)} \quad (2.37)$$

Defining $F : [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$ by

$$F(t,z) = \begin{pmatrix} f(t, x_1, e_1^* z, \sigma^*(x_1)(e_1^*\widetilde{D}_1^{(1)} z, \ldots, e_1^*\widetilde{D}_1^{(d)} z)^*) \\ \vdots \\ f(t, x_N, e_N^* z, \sigma^*(x_N)(e_N^*\widetilde{D}_1^{(1)} z, \ldots, e_N^*\widetilde{D}_1^{(d)} z)^*) \end{pmatrix}$$

and $G = (g(x_1), \ldots, g(x_N))^*$, the system of ODEs results in the same form as (2.34):

$$\frac{dU_t^{\Pi}}{dt} + QU_t^{\Pi} + F(t, U_t^{\Pi}) = 0, \quad U_T^{\Pi} = G. \quad (2.38)$$

2.4.3 Probabilistic Interpretation

Recall that Q in (2.36) or (2.37) is constructed from the spatial discretization of the infinitesimal generator $\mathcal{L}$. In the probabilistic manner, it is natural to interpret Q as the Q-matrix of a time-homogeneous CTMC. Since Q might no longer be a Q-matrix, it is required to see the "validity" conditions of Q to be the Q-matrix. In the case of $d = 1$, one can easily give the following sufficiency condition. It guarantees the validity of a CTMC constructed by Q provided the spatial difference is sufficiently fine.

Proposition 2.4.1 (Validity)**.** *Q defined by* (2.36) *is the transition rate matrix of a continuous-time Markov chain if*

$$0 < \max_{-N_0 \le i \le N_0 - 1} \{\delta x_i\} \le \min_{\substack{-N_0 \le i \le N_0 \\ \mu(x_i) \neq 0}} \left\{ \frac{\sigma^2(x_i)}{|\mu(x_i)|} \right\}. \quad (2.39)$$

Additionally, if the above inequality is strict, $e_{i-1}^* Q e_i$ and $e_{i+1}^* Q e_i$ are positive for all $i = -N_0 + 1, \ldots, N_0 - 1$.

Proof. See Section 2.7.2. □

Note that validity conditions in multi-dimensional settings would be more complicated; the existence of terms that involve cross derivatives occasionally violates the definition of Q-matrices. Our approach presented in Section 2.3, that substantially solves systems of ODEs using exponential integrators, works without any issues at least numerically, regardless of the validity. However, in certain situations when we need to simulate the CTMC, we should instead employ several discretization schemes that avoid the invalidity issues [10, 52].

Assuming the validity of Q, let X be a finite-state Markov chain with Q as its Q-matrix. We now consider the Markov BSDE arising from (2.34) as

$$Y_t = X_T^* G + \int_{]t,T]} X_{s-}^* F(s, Z_s) ds - \int_{]t,T]} dM_s^* Z_s. \tag{2.40}$$

Thus, we can regard (2.40) as spatially discretized counterpart of (2.14), and applying exponential integrators to (2.34) is equivalent to the (multi-stage) Euler-Maruyama methods of (2.40).

We give a result on a uniqueness of (2.40) under standard conditions. Theoretical justification of this probabilistic interpretation is completed with it.

Proposition 2.4.2. *Suppose that* $\int_0^T f(t,x,0,0)^2 dt < \infty$ *for any* $x \in \mathbb{R}^d$*, and that for some* $L > 0$*,*

$$|f(t,x,y,z) - f(t,x,y',z')|^2 \le L(|y-y'|^2 + \|z-z'\|^2)$$

for all $t \in [0,T]$*,* $x \in \mathbb{R}^d$*,* $y, y' \in \mathbb{R}$*, and* $z, z' \in \mathbb{R}^d$*. Suppose further that* Q *is valid,* $e_i^* \widetilde{D}_1^{(p)} 1 = 0$*, and that*

$$e_j^* Q e_i = 0 \Longrightarrow e_j^* \widetilde{D}_1^{(p)} e_i = 0 \quad \textit{for} \quad p = 1, \dots, d, \tag{2.41}$$

for $i, j = 1, \dots, N$*. Then,* (2.40)*, which is derived from the Markov chain approximation of* (2.16)*, has a unique solution.*

Proof. See Section 2.7.3. □

Remark 2.4.1. *As a related study on CTMC approximation of SDEs, we refer to [53, 24, 49, 46, 25].*

Remark 2.4.2. *In [21], the authors presented a CTMC version of least-squares Monte Carlo methods. Although this type of method is also a natural counterpart of numerical solutions of Markov BSDEs based on the Euler-Maruyama temporal discretization, we note that it is not suitable to solve CTMC-driven Markov BSDEs*

arising from spatial discretization of Brownian motion-driven Markov BSDEs. A major bottleneck is that the resulting Q-matrix may contain quite larger absolute values. To illustrate it, suppose that $\mu(x) = \mu x$, $\sigma(x) = \sigma x$, *and* $\delta x_i \equiv \Delta x > 0$. *Plugging them into* (2.36), *we obtain*

$$e_i^* Q e_{i-1} = \frac{\sigma^2 i^2 - \mu i}{2}, \quad e_i^* Q e_{i+1} = \frac{\sigma^2 i^2 + \mu i}{2}.$$

If i *is nearby* N *and* N *is large, these elements take large values, which implies the resulting Markov chain jumps too rapidly. It interferes with us simulating CTMCs naively using Gillespie's exact simulation or the 1st-order approximation of the transition probability matrix; the former suffers from a tremendously large number of jumps, and the latter method requires the temporal step small enough for each row of the resulting matrix to represent probabilities.*

2.5 Numerical Results

In this section, we demonstrate the efficiency and stability of the numerical approach presented in Section 2.4, using several examples. We apply spatial discretization to the BSDEs driven by Brownian motions, obtain BSDEs driven by CTMCs (i.e. a system of ODEs), and calculate numerical solutions using multi-stage Euler-Maruyama methods (i.e. exponential integrators.) Before moving on to specific results, we explain the details on settings in what follows:

Spatial Discretization We approximate the unbounded spatial domain of the problem at hand with a bounded one. Because all of the BSDEs we solve in this section have $(0, \infty)^d$ as spatial domains of them, we approximate as

$$(0, \infty)^d \approx [0, 2x_1] \times \cdots \times [0, 2x_d] \tag{2.42}$$

for some $x_1, \ldots, x_d > 0$, or sometimes as

$$(0, \infty)^d \approx [\Delta_{x_1}, 2x_1 - \Delta_{x_1}] \times \cdots \times [\Delta_{x_d}, 2x_d - \Delta_{x_d}] \tag{2.43}$$

for small $\Delta_{x_1}, \ldots, \Delta_{x_d} > 0$. The latter is applied to the problems that should be evaluated only at points on the first quadrant. Here, $(x_1, \ldots, x_d)$ is a point at which we want to evaluate the numerical solution.

Throughout the section, $\Pi_x^{\mathrm{Unif}}(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_{x,0}) = (x_i)_{i=-N_{x,0}}^{N_{x,0}}$ means the standard one-dimensional uniform grid such that $x_{-N_{x,0}} = x_{\mathrm{left}}$, $x_0 = x_{\mathrm{center}}$ and

$x_{N_{x,0}} = x_{\text{right}}$. We occasionally omit its arguments and simply write Π_x^{Unif}. In addition to the uniform grid, a non-uniform grid is also employed for spatial discretization. We use a version of the (one-dimensional) Tavella-Randall-type grids [63] in what follows:

$$x_k = \begin{cases} x_{\text{center}} + g_1 \sinh\left(\operatorname{arc\,sinh}\left(\dfrac{x_{\text{center}} - x_{\text{left}}}{g_1}\right) \cdot \dfrac{k}{N_{x,0}}\right), & k = -N_{x,0}, \dots, -1, 0, \\ x_{\text{center}} + g_2 \sinh\left(\operatorname{arc\,sinh}\left(\dfrac{x_{\text{right}} - x_{\text{center}}}{g_2}\right) \cdot \dfrac{k}{N_{x,0}}\right), & k = 1, 2, \dots, N_{x,0}, \end{cases} \tag{2.44}$$

where $N_x \coloneqq 2N_{x,0} + 1$ is the grid size, x_{left} and x_{right} are the leftmost and rightmost points of the domain, $x_{\text{center}} \in (x_{\text{left}}, x_{\text{right}})$ is the central point of the grid, and g_1 and g_2 are parameters for the left- and right-side of the grid, respectively. Note that $x_{-N_{x,0}} = x_{\text{left}}$, $x_0 = x_{\text{center}}$ and $x_{N_{x,0}} = x_{\text{right}}$. Intuitively, setting $g_1, g_2 \ll x_{\text{right}} - x_{\text{left}}$ leads to the grid that is highly concentrated around x_{center}. It is commonly used in numerical computation for pricing options to mitigate the effect of the non-linearity of the payoff function [9, 53, 63]. Similarly to the uniform grid, denote $\Pi_x^{\text{TR}}(x_{\text{left}}, x_{\text{center}}, x_{\text{right}}, N_{x,0}, g_1, g_2)$ as the Tavella-Randall grid (2.44) whose parameters are $(x_{\text{left}}, x_{\text{center}}, x_{\text{right}}, N_{x,0}, g_1, g_2)$.

Figure 2.2: $g_1 = g_2 = 50.0$

Figure 2.3: $g_1 = g_2 = 5.0$

Figure 2.4: $g_1 = g_2 = 0.5$

Figure 2.5: Examples of Tavella-Randall grids with $x_{\text{left}} = 0$, $x_{\text{center}} = 100$, $x_{\text{right}} = 200$, $N_{x,0} = 100$ and $g_1 = g_2 \in \{50, 5, 0.5\}$.

Temporal Discretization We employ solvers implemented in `DifferentialEquations.jl` [62], listed below:

- `LawsonEuler` : A single-stage method of classical/stiff order 1/1, referred to as the Lawson-Euler method [47].

- `NorsettEuler` : A single-stage method of classical/stiff order 1/1, referred to as the Nørsett-Euler method or ETD1RK method [22].
- `ETDRK2` : A 2-stage method of classical/stiff order 2/2 [22].
- `ETDRK3` : A 3-stage method of classical/stiff order 3/3 [22].
- `ETDRK4` : A 4-stage method of classical/stiff order 4/2 [22].
- `HochOst4` : A 5-stage method of classical/stiff order 4/4, developed by Hochbruck and Ostermann [39].

Taking temporal grid size $N_t \in \mathbb{N}$, we calculate solutions on the grid $\Pi_t^{\mathrm{Unif}}(N_t) = (i\Delta t)_{i=0}^{N_t}$ using these exponential integrators. Here, $\Delta t = T/N_t$ is the step size. Note that a large-scale system of ODEs is obtained from the spatial discretization in each experiment. In this case, employing Krylov subspace methods in evaluating matrix exponentials and related ϕ functions is more effective, as described in Remark 2.3.3. In all the experiments, we use the Arnoldi iteration with a size-m Krylov subspace, which is readily available on all the solvers above. For simplicity, we always take $m = 100$.

A Least-Squares Monte Carlo Method: A Reference BSDEs (driven by Brownian motion) that appeared in this section include those whose analytical solutions are unknown. In experiments of those BSDEs, as a reference, we shall report numerical solutions using the least squares Monte Carlo (LSMC) method [50, 32, 3] using Laguerre polynomials as

$$\mathrm{poly}_p^{\mathrm{Lag}}(x) = \sum_{k=0}^{p} \frac{(-1)^k}{k!} \binom{p}{k} x^k, \quad p = 0, 1, 2, \ldots. \tag{2.45}$$

Note that for the multi-dimensional case, the basis function corresponds to the Cartesian product of (2.45). Since LSMC methods include randomness, we independently calculate solutions for 50 times, and report the mean values, the standard deviations, and the total runtimes.

Implementation All of our experiments were performed on a 3.70 GHz, 64-GB RAM Linux workstation. Our code was written entirely in Julia [5] and all the plots were produced using `Plot.jl` [13]. The full code for the experiments is available at https://github.com/kanekoakihiro/EMCTMCBSDE.

2.5.1 European Call Option under the Black-Scholes Model

First, we consider a linear BSDE arising from pricing a European call option under the Black-Scholes model:

$$\begin{aligned} \mathcal{S}_t &= s_0 + \int_0^t \mu \mathcal{S}_s ds + \int_0^t \sigma \mathcal{S}_s dW_s, \\ \mathcal{Y}_t &= (\mathcal{S}_T - K)^+ - \int_t^T r\mathcal{Y}_t dt - \int_t^T \mathcal{Z}_t dW_t. \end{aligned} \tag{2.46}$$

Here, K is the strike price of the European call option, T is the maturity, r is the interest rate, and $\mathcal{S}_t$ is the spot price of the underlying risky asset with initial price s_0, appreciation μ and volatility σ. Then, terminal condition $\mathcal{Y}_T = (\mathcal{S}_T - K)^+ = \max\{S_T - K, 0\}$ is the payoff of the European call opton, and the solution $\mathcal{Y}_t = \mathcal{Y}_t^{t,\mathcal{S}_t}$ means the price of the option at time $t \in [0, T]$ and spot price $\mathcal{S}_t$. For details on the derivation of (2.46), see Section 4.5.1. (p.91) in [75]. Using the cumulative distribution function of the standard Gaussian distribution $\Psi(x)$, the solution can be evaluated as

$$\begin{cases} \mathcal{Y}_t = \mathcal{S}_t \cdot \Psi(d_1) - K \exp(-r(T-t)) \cdot \Psi(d_2), \\ \mathcal{Z}_t = \mathcal{S}_t \cdot \Psi(d_1)\sigma, \end{cases} \tag{2.47}$$

where d_1 and d_2 are constants as follows

$$d_1 = \frac{\ln\left(\frac{\mathcal{S}_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = \frac{\ln\left(\frac{\mathcal{S}_t}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}.$$

We choose the parameters of (2.46) as follows:

T	K	r	μ	σ
1	100	0.03	0.03	0.2

Despite a pretty simple case, its spatial discretization leads to a stiff system of ODEs. To this end, we discretize (2.46) on the Tavella-Randall grid Π_x^{TR} with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_{x,0}, g_1, g_2) = (0, 100, 200, 1000, 50, 50)$ and calculate solutions of the resulting system of ODEs using `DP5`, an implementation of the Dormand-Prince explicit solver in Julia, for different time steps N_t. Consequently, we observed that it requires approximately 58005 steps along the temporal direction to achieve the "stable" solution; otherwise, terribly large and rapid oscillations occur in some parts

of the numerical solution. Fig. 2.6 shows the surface plots of the numerical solutions of $(t, x) \mapsto \mathcal{Y}_t^{t,x}$ for 57990, 57995, 58000, and 58005 temporal steps, respectively. For visibility, the surfaces displayed are the 30×30 arrays uniformly sampled from numerical solutions whose values heve been clipped in the $[0, 250]$ range. Absolute errors of $\mathcal{Y}_t^{t,x}$ at $(t, x) = (0, 100)$, maximum absolute errors of $\mathcal{Y}_t^{t,x}$ in $(t, x) \in \Pi_t^{\mathrm{Unif}} \times ([80, 120] \cap \Pi_x^{\mathrm{TR}})$ and runtime in seconds are reported in Table 2.1. These results epitomize how stiff systems arise from the spatial discretization of parabolic PDEs and prevent explicit solvers from calculating solutions efficiently.

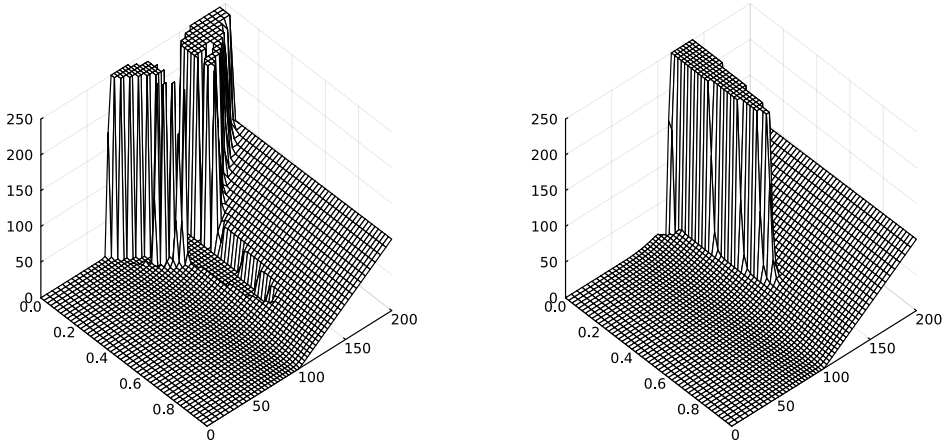
	$N_t = 57990$	$N_t = 57995$	$N_t = 58000$	$N_t = 58005$
Sup Error	3.304e+31	1.727e+18	1.218e+5	1.326e−2
Abs Error	1.989e+12	1.324e+3	4.612e−4	4.612e−4
Runtime [s]	600.51	629.94	665.08	668.86

Table 2.1: Results on numerical solutions of (2.46) using DP5 for different N_t. Here, we spatially discretize (2.46) on Π_x^{TR} and solve the resulting system of ODEs. Here, the parameters of Π_x^{TR} are $x_{\mathrm{left}} = 0$, $x_{\mathrm{center}} = 100$, $x_{\mathrm{right}} = 200$, $N_{x,0} = 1000$, and $g_1 = g_2 = 50$. For each N_t, the numerical solution is evaluated on the grid $\Pi_t^{\mathrm{Unif}}(N_t) \times \Pi_x^{\mathrm{TR}}$. The maximum absolute errors in $\Pi_t^{\mathrm{Unif}} \times ([80, 120] \cap \Pi_x^{\mathrm{TR}})$ are reported on the row of "Sup Error", the absolute errors at $(t, x) = (0, 100)$ are on the row of "Abs Error", and the runtimes in seconds are at the bottom line.

Let us solve the system of ODEs constructed from above using exponential integrators. Table 2.2 reports the results of `HochOst4`, a 4-stage exponential integrator of order 4, for different temporal steps N_t. It successfully calculates solutions without suffering from huge errors as appearing in `DP5`, and provides accurate solutions with fewer temporal steps; for example, the solution calculated using `HochOst4` with $N_t \geq 200$ has achieved the same level of accuracy as the one using `DP5` with $N_t = 58005$.

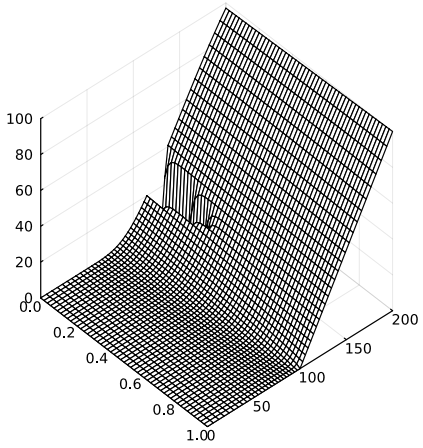
	$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$	$N_t = 500$	$N_t = 1000$
Sup Error	5.102e−1	1.374e−1	1.384e−2	1.326e−2	1.326e−2	1.326e−2	1.326e−2
Abs Error	2.611e−1	2.638e−2	1.092e−3	4.693e−4	4.612e−4	4.612e−4	4.612e−4
Runtime[s]	1.56	2.33	5.43	12.21	19.73	49.26	96.63

Table 2.2: Results on numerical solutions of (2.5.1) using HochOst4 in the same situation in Table 2.1.

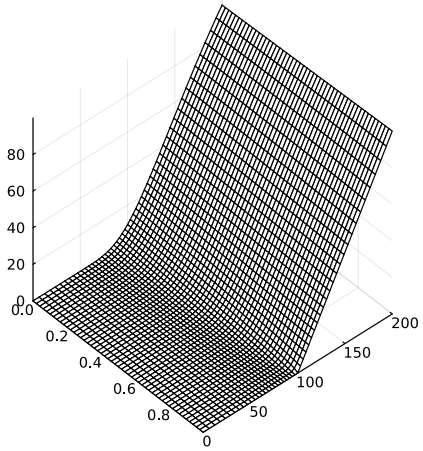


(a) $N_t = 57990$.

(b) $N_t = 57995$.



(c) $N_t = 58000$.



(d) $N_t = 58005$.

Figure 2.6: Plots of numerical solutions of (2.5.1) using DP5 for different N_t. The spatial grid is Π_x^{TR} whose parameters are $x_{\mathrm{left}} = 0$, $x_{\mathrm{center}} = 100$, $x_{\mathrm{right}} = 200$, $N_{x,0} = 1000$, and $g_1 = g_2 = 50$. For visibility, data plotted are 50×50 arrays uniformly sampled and clipped into range $[0, 250]$.

2.5.2 European Call Option with Higher Interest Rate for Borrowing

Next, consider a nonlinear version of (2.46):

$$\begin{aligned}\mathcal{S}_t &= s_0 + \int_0^t \mu \mathcal{S}_s ds + \int_0^t \sigma \mathcal{S}_s dW_s, \\ \mathcal{Y}_t &= (\mathcal{S}_T - K)^+ - \int_t^T [r(\mathcal{Y}_s - \frac{\mathcal{Z}_s}{\sigma})^+ - R(\mathcal{Y}_s - \frac{\mathcal{Z}_s}{\sigma})^- + \frac{\mu}{\sigma}\mathcal{Z}_s]ds - \int_t^T \mathcal{Z}_s dW_s,\end{aligned} \tag{2.48}$$

which comes from considering a more practical case that the borrowing interest rate R is higher than the lending interest rate r. Here, $(x)^+ = \max\{x, 0\}$ and $(x)^- = \min\{x, 0\}$ for $x \in \mathbb{R}$. For details on the derivation of (2.48), Section 4.5.1. (p.91) in [75]. We choose the parameters of (2.48) as follows:

T	μ	σ	R	r	K
1.0	0.03	0.2	0.3	0.01	100

We discretize 2.48 on the Tavella-Randall grid Π_x^{TR} with $x_{\mathrm{left}} = 0$, $x_{\mathrm{center}} = 100$, $x_{\mathrm{right}} = 200$, $N_{x,0} = 1000$, and $g_1 = g_2 = 50$, which is the same as before. Unlike the previous experiment, however, we cannot evaluate numerical errors, since we do not know the analytical solution of (2.48). We only focus on numerical solutions of $\mathcal{Y}_t = \mathcal{Y}_t^{t,s}$ at $(t, s) = (0, 100)$. Table 2.3 reports numerical solutions using different exponential integrators and the runtimes in second. For each scheme, when N_t increases, the numerical solution seems to converge approximately $\mathcal{Y}_0^{0,100} \approx 26.3305$. We observe that multi-stage methods (ETD2RK, ETD3RK, ETD4RK, and HochOst4) converge faster than single-stage methods (Lawson-Euler and Nørsett-Euler). As a reference, we calculate the solution $\mathcal{Y}_0^{0,100}$ using the LSMC method with Laguerre polynomials (2.45) up to p-th order for different parameters. Here, the number of Monte Carlo simulations is $M = 2^{22}$, and the pair (N_t, p) of the number of temporal steps N_t and the maximal order p are taken from $\{(6, 6), (7, 7), (8, 8), (9, 9)\}$. The LSMC solution of $\mathcal{Y}_0^{0,100}$ is evaluated 50 times independently, and we compute the sample mean, the unbiased standard deviation and the total runtime. The results of the numerical solution of the LSMC method are reported in Table 2.4.

		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$
Lawson-Euler [47]	$\mathcal{Y}_0^{0,s_0}$	25.0377	26.3522	26.3545	26.3437	26.3372
	Runtime [s]	5.26	0.61	1.52	2.81	5.40
Nørsett-Euler [22]	$\mathcal{Y}_0^{0,s_0}$	26.0777	26.4527	26.3881	26.3597	26.3452
	Runtime [s]	1.25	0.53	1.30	2.51	5.23
ETD2RK [22]	$\mathcal{Y}_0^{0,s_0}$	25.8658	26.3182	26.3310	26.3307	26.3306
	Runtime [s]	1.51	1.27	2.89	5.85	10.78
ETD3RK [22]	$\mathcal{Y}_0^{0,s_0}$	26.1212	26.3137	26.3302	26.3305	26.3305
	Runtime [s]	1.78	1.71	3.81	8.02	16.68
ETD4RK [22]	$\mathcal{Y}_0^{0,s_0}$	26.0947	26.3136	26.3302	26.3305	26.3305
	Runtime [s]	3.10	2.72	6.47	14.35	27.32
HochOst4 [39]	$\mathcal{Y}_0^{0,s_0}$	26.1080	26.3136	26.3302	26.3305	26.3305
	Runtime [s]	2.28	2.79	6.84	14.27	29.64

Table 2.3: Numerical solutions $\mathcal{Y}_0^{0,100}$ and its runtime in seconds using different exponential integrators. Here, we spatially discretize (2.48) on Π_x^{TR} and solve the resulting system of ODEs. Here, the parameters of Π_x^{TR} are $x_{\mathrm{left}} = 0$, $x_{\mathrm{center}} = 100$, $x_{\mathrm{right}} = 200$, $N_{x,0} = 1000$, and $g_1 = g_2 = 50$.

M		$(N_t, p) = (6, 6)$	$(N_t, p) = (7, 7)$	$(N_t, p) = (8, 8)$	$(N_t, p) = (9, 9)$
4194304	mean	26.1462	26.2239	26.3001	26.3404
	std	0.0457	0.0386	0.0611	0.0580
	Runtime [s]	1436.79	1706.81	2021.94	2340.43

Table 2.4: Results on numerical solutions $\mathcal{Y}_0^{0,100}$ of (2.48) using the LSMC methods with Laguerre polynomials up to p-th order for N_t time steps. Here, $M = 4194304$ is the number of samples for the Monte-Carlo approximation. We run the LSMC algorithm 50 times independently and collect each $\mathcal{Y}_0^{0,100}$. The row of "mean" and "std" reports their sample means and sample (unbiased) standard deviations, respectively. The bottom of the table reports the total CPU times required for the 50 experments.

2.5.3 European Options under Stochastic Local Volatility Models

In this section, we consider BSDEs arising from the valuation of European options in stochastic local volatility (SLV) models, a class of Markov BSDEs with 2-dimensional

state processes. A general form of stochastic volatility models is given in [24] as

$$\begin{cases} \mathcal{S}_t = S_0 + \int_0^t \omega(\mathcal{S}_s, v_s)ds + \int_0^t m(v_s)\Gamma(\mathcal{S}_s)dW_s^{(1)}, \\ v_t = v_0 + \int_0^t \mu(v_s)ds + \int_0^t \sigma(v_s)dW_s^{(2)}, \end{cases} \tag{2.49}$$

where $\langle W^{(1)}, W^{(2)} \rangle_t = \rho t$ with $\rho \in (-1, 1)$. Let $\mathbf{L} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}$ be the lower triangular matrix constructed from the Cholesky decomposition $\mathbf{C} = \mathbf{L}\mathbf{L}^*$ of $\mathbf{C} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. Since $\mathbf{W} = \mathbf{L}^{-1} \begin{pmatrix} W^{(1)} \\ W^{(2)} \end{pmatrix}$ is a 2-dimensional standard Brownian motion, we can reformulate (2.49) as

$$\mathcal{X}_t = \mathcal{X}_0 + \int_0^t \boldsymbol{\mu}(\mathcal{X}_s)ds + \int_0^t \boldsymbol{\sigma}(\mathcal{X}_s)d\mathbf{W}_s \quad \text{for} \quad t \in [0, T],$$

where

$$\mathcal{X}_t = \begin{pmatrix} \mathcal{S}_t \\ v_t \end{pmatrix}, \quad \boldsymbol{\mu}(x) = \begin{pmatrix} \omega(x_1, x_2) \\ \mu(x_2) \end{pmatrix}, \quad \boldsymbol{\sigma}(x) = \begin{pmatrix} m(x_2)\Gamma(x_1) & 0 \\ 0 & \sigma(x_2) \end{pmatrix} \mathbf{L}.$$

Similarly to the arguments in Section 4.5.1. (p.91) in [75], we formulate Markov BSDEs describing the price of European options. Consider the self-financing portfolio $\boldsymbol{\Delta}_t = \begin{pmatrix} \Delta_t^1 \\ \Delta_t^2 \end{pmatrix}$ consisting of Δ_t^1 assets of $\mathcal{S}_t$ and Δ_t^2 assets of v_t and bonds with borrowing rate R and lending rate r. Let $\mathcal{V}_t$ be the wealth dynamics of $\boldsymbol{\Delta}_t$. The self-financing condition reads

$$d\mathcal{V}_t = [r(\mathcal{V}_t - \boldsymbol{\Delta}_t^* \mathcal{X}_t)^+ - R(\mathcal{V}_t - \boldsymbol{\Delta}_t^* \mathcal{X}_t)^-]dt + \boldsymbol{\Delta}_t^* d\mathcal{X}_t.$$

Consider that $\boldsymbol{\Delta}_t$ hedges the European option with payoff $g(\mathcal{S}_T)$. Denoting $\mathcal{Y}_t \coloneqq \mathcal{V}_t$ and $\mathcal{Z}_t \coloneqq \boldsymbol{\Delta}_t^* \boldsymbol{\sigma}(\mathcal{X}_t)$, we obtain

$$\mathcal{Y}_t = g(\mathcal{S}_T) - \int_t^T f(s, \mathcal{X}_s, \mathcal{Y}_s, \mathcal{Z}_s)ds - \int_t^T \mathcal{Z}_s^* d\mathbf{W}_s, \tag{2.50}$$

where the driver is

$$f(t, x, y, z) = r(y - z^* \boldsymbol{\sigma}(x)^{-1} x)^+ - R(y - z^* \boldsymbol{\sigma}(x)^{-1} x)^- + z^* \boldsymbol{\sigma}(x)^{-1} \boldsymbol{\mu}(x).$$

Then, $\mathcal{Y}_t^{t, \mathcal{X}_t}$ is the price of the European option.

Before moving on to the specific results, we note what follows: (1) Unlike the cases considered in the previous two experiments, we need to discretize the two-dimensional process $\mathcal{X}_t = (\mathcal{S}_t, v_t)$. To this end, we construct the Kronecker product of two grids as described in Section 2.4.1. Specifically, $\mathcal{S}_t$ is discretized on the Tavella-Randall grid Π_x^{TR} and v_t is on the uniform grid Π_x^{Unif}. (2) BSDE (2.50) contains the evaluation of $\boldsymbol{\sigma}(x)^{-1}$, and we should set each grid carefully; if $\Gamma(x)$, $m(x)$ or $\sigma(x)$ take zero at $x = 0$, grid (2.43) should be chosen rather than grid (2.42). (3) We calculate numerical solutions using the LSMC method for reference. Since the Euler-Maruyama discretization paths can take negative values, coefficient functions defined only on $(0, \infty)$ (e.g. the square root) may fail to evaluate. To avoid it, the absolute value is taken under such coefficient functions.

2.5.3.1 European Put Opton under the Heston-SABR Model with Higher Interest Rate for Borrowing

The Heston-SABR SLV model [71] takes the following form.

$$\begin{aligned} \mathcal{S}_t &= \mathcal{S}_0 + \int_0^t b \cdot \mathcal{S}_s ds + \int_0^t \sqrt{v_s}\mathcal{S}_s^{\beta} dW_s^{(1)} \\ v_t &= v_0 + \int_0^t \eta(\theta - v_s) ds + \int_0^t \alpha\sqrt{v_s} dW_s^{(2)} \end{aligned} \tag{2.51}$$

We apply (2.5.3.1) to (2.49), and consider BSDE (2.50) with $g(s) = (K - s)^+$ a payoff function of a European put option. We choose the parameters of (2.5.3.1) as follows:

T	β	η	θ	α	ρ	b	K	R	r
1.0	0.7	4.0	0.035	0.15	-0.75	0.01	100	0.07	0.01

In this experiment, we are interested in numerical solutions of $\mathcal{Y}_0 = \mathcal{Y}_0^{0,(s_0,v_0)}$ at $(s_0, v_0) = (100, 0.4)$. As noted before, since the driver of (2.50) contains the evaluation of inverses of $\sqrt{v_t}\mathcal{S}_t^{\beta}$ and $\alpha\sqrt{v_s}$, we need to design the spatial grid to contain only points in the first quadrant. To this end, we set the spatial grid onto $[\Delta_1, 2s_0 - \Delta_1] \times [\Delta_2, 2v_0 - \Delta_2]$, where $\Delta_1 = s_0/N_0^{(1)}$, $\Delta_2 = v_0/N_0^{(2)}$. We take $(N_0^{(1)}, N_0^{(2)}) = (100, 15)$, and the resulting spatial grid $\Pi_x^{\mathrm{TR}} \otimes \Pi_x^{\mathrm{Unif}}$ contains $(2 \times 100 + 1) \times (2 \times 15 + 1) = 6231$ points in the reculangular domain $[100/100, 200 - 100/100] \times [0.4/15, 0.8 - 0.4/15] (\approx [1, 199] \times [0.0267, 0.7733].)$ For g_1 and g_2, we take $g_1 = g_2 = 1$.

Numerical solutions using exponential integrators are reported in Table 2.5. As N_t increases, the solutions seem to converge approximately $\mathcal{Y}_0^{0,(100,0.4)} \approx 5.6394$. As

before, faster convergence of multi-stage methods than 1-stage methods has been confirmed. Numerical solutions using the LSMC method is given in Table 2.6.

		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$
Lawson-Euler [47]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	4.7985	5.5605	5.6015	5.6211	5.6302
	Runtime [s]	22.79	29.70	66.15	129.13	255.58
Nørsett-Euler [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	4.9076	5.5662	5.6234	5.6326	5.6360
	Runtime [s]	14.71	26.26	61.46	116.11	231.01
ETD2RK [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	4.9348	5.6006	5.6368	5.6393	5.6394
	Runtime [s]	26.15	43.70	112.08	221.11	443.11
ETD3RK [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	4.9329	5.6014	5.6369	5.6393	5.6394
	Runtime [s]	33.04	66.37	159.77	341.57	724.57
ETD4RK [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	4.9521	5.6015	5.6369	5.6393	5.6394
	Runtime [s]	63.90	121.81	293.59	566.32	1150.64
HochOst4 [39]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	4.9328	5.6013	5.6369	5.6393	5.6394
	Runtime [s]	60.23	114.64	285.99	551.47	1117.35

Table 2.5: Numerical solutions $\mathcal{Y}_0^{0,(100,0.4)}$ and its runtime in seconds using different exponential integrators. Here, we spatially discretize (2.50) on $\Pi_x^{\mathrm{TR}} \otimes \Pi_x^{\mathrm{Unif}}$ and solve the resulting system of ODEs. Here, the parameters of Π_x^{TR} are $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_{x,0}, g_1, g_2) = (1, 100, 199, 100, 1, 1)$, and of Π_x^{Unif} are $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_{x,0}) \approx (0.0267, 0.4, 0.7733, 15)$.

M		$(N,p) = (6,6)$	$(N,p) = (7,7)$	$(N,p) = (8,8)$	$(N,p) = (9,9)$
1048576	mean	5.5956	5.6132	5.6247	5.6249
	std	0.0153	0.0153	0.0179	0.0192
	Runtime [s]	3250.84	4737.51	6788.49	9863.74

Table 2.6: Results on numerical solutions $\mathcal{Y}_0^{0,(100,0.4)}$ of (2.50) under the Heston-SABR model (2.51) using the LSMC methods with Laguerre polynomials up to p-th order for N_t time steps. Here, $M = 2^20 = 1048576$ is the number of samples for the Monte-Carlo approximation. We run the LSMC algorithm 50 times independently and collect each $\mathcal{Y}_0^{0,(100,0.4)}$. The row of "mean" and "std" reports their sample means and sample (unbiased) standard deviations, respectively. The bottom of the table reports the total CPU times required for the 50 experments.

2.5.3.2 Calls Combination with Different Interest Rates under the Hyp Hyp SLV Model

Consider the Hyp Hyp SLV model [43] defined as:

$$\begin{aligned} \mathcal{S}_t &= \mathcal{S}_0 + \int_0^t b \cdot \mathcal{S}_s ds + \int_0^t \sigma_0 \cdot F(\mathcal{S}_s) G(v_s) dW_s^{(1)}, \\ v_t &= v_0 - \int_0^t \kappa \cdot v_s ds + \int_0^t \alpha \sqrt{2\kappa} dW_s^{(2)}, \end{aligned} \tag{2.52}$$

where

$$\begin{aligned} F(x) &= \left[(1 - \beta + \beta^2) \cdot x + (\beta - 1) \cdot (\sqrt{x^2 + \beta^2 (1-x)^2} - \beta) \right] / \beta, \\ G(v) &= v + \sqrt{v^2 + 1}. \end{aligned}$$

We apply (2.52) to (2.49), and consider BSDE (2.50) with

$$g(s) = (s - 95)^+ - 2(s - 105)^+,$$

which is a payoff function of a combination of two European call options. We choose the parameters of (2.52) as follows:

T	β	κ	σ_0	α	ρ	b	K	R	r
1.0	0.25	0.5	0.25	0.3	0.8	0.04	100	0.06	0.006

We discretize $\mathcal{X}_t$ in the same way of Section 2.5.3.1 except the discretization of $\mathcal{S}_t$; because the payoff function g is non-smooth at $s = 95$ and $s = 105$, we should employ the concatenation of two Tavella-Randall grids Π_x^{TR} with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}) = (1, 95, 100)$ and $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}) = (100, 105, 199)$; the other parameters are commonly taken as $N_{x,0} = N_0^{(1)}/2 = 50$ and $g_1 = g_2 = 1$.

Results on numerical solutions using exponential integrators are shown in Table 2.7. Numerical solutions of $\mathcal{Y}_0^{0,(s_0,v_0)}$ at $(s_0, v_0) = (95, 0.4), (100, 0.4), (105, 0.4)$ have been reported; as N_t increases, they seem to converge towards

$$\mathcal{Y}_0^{0,(95,0.4)} \approx 4.4061, \quad \mathcal{Y}_0^{0,(100,0.4)} \approx 5.8218, \quad \text{and} \quad \mathcal{Y}_0^{0,(105,0.4)} \approx 5.6394.$$

Numerical solutions using the LSMC method are presented in Table **??**.

		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$
Lawson-Euler [47]	$\mathcal{Y}_0^{0,(95,0.4)}$	4.5589	4.4829	4.4368	4.4214	4.4137
	$\mathcal{Y}_0^{0,(100,0.4)}$	6.0971	5.9595	5.8768	5.8492	5.8355
	$\mathcal{Y}_0^{0,(105,0.4)}$	5.9113	5.7756	5.6938	5.6665	5.6529
	Runtime[s]	15.32	24.69	57.58	113.83	231.95
Nørsett-Euler [22]	$\mathcal{Y}_0^{0,(95,0.4)}$	4.5561	4.4787	4.4343	4.4200	4.4130
	$\mathcal{Y}_0^{0,(100,0.4)}$	5.9936	5.9035	5.8531	5.8371	5.8293
	$\mathcal{Y}_0^{0,(105,0.4)}$	5.7821	5.707	5.6652	5.6520	5.6456
	Runtime[s]	13.52	22.99	52.94	108.18	198.73
ETD2RK [22]	$\mathcal{Y}_0^{0,(95,0.4)}$	4.4217	4.4115	4.4074	4.4065	4.4063
	$\mathcal{Y}_0^{0,(100,0.4)}$	5.8444	5.8294	5.8236	5.8224	5.8220
	$\mathcal{Y}_0^{0,(105,0.4)}$	5.6631	5.6473	5.6412	5.6400	5.6396
	Runtime[s]	22.80	41.33	99.87	206.08	419.02
ETD3RK [22]	$\mathcal{Y}_0^{0,(95,0.4)}$	4.4082	4.4068	4.4063	4.4062	4.4061
	$\mathcal{Y}_0^{0,(100,0.4)}$	5.8248	5.8227	5.8220	5.8218	5.8218
	$\mathcal{Y}_0^{0,(105,0.4)}$	5.6424	5.6403	5.6396	5.6395	5.6394
	Runtime[s]	36.45	64.68	157.76	327.25	631.46
ETD4RK [22]	$\mathcal{Y}_0^{0,(95,0.4)}$	4.4080	4.4068	4.4063	4.4062	4.4061
	$\mathcal{Y}_0^{0,(100,0.4)}$	5.8243	5.8226	5.8220	5.8218	5.8218
	$\mathcal{Y}_0^{0,(105,0.4)}$	5.6421	5.6403	5.6396	5.6395	5.6394
	Runtime[s]	59.71	113.26	257.03	499.43	1060.09
HochOst4 [39]	$\mathcal{Y}_0^{0,(95,0.4)}$	4.4080	4.4068	4.4063	4.4062	4.4061
	$\mathcal{Y}_0^{0,(100,0.4)}$	5.8244	5.8226	5.8220	5.8218	5.8218
	$\mathcal{Y}_0^{0,(105,0.4)}$	5.6421	5.6403	5.6396	5.6395	5.6394
	Runtime[s]	59.85	113.38	275.02	544.55	1043.06

Table 2.7: Numerical solutions $\mathcal{Y}_0^{0,(s_0,v_0)}$ at $(s_0, v_0) = (95, 0.4), (100, 0.4), (105, 0.4)$ of (2.52) and its runtime in seconds using different exponential integrators. Here, we spatially discretize (2.50) on $\Pi_x^{\mathrm{TR}} \otimes \Pi_x^{\mathrm{Unif}}$, and solve the resulting system of ODEs. We employ concatenation of two Tavella-Randall grids Π_x^{TR} with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}) = (1, 95, 100)$ and $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}) = (100, 105, 199)$ in the $\mathcal{S}_t$-direction. In both grids, we take $N_{x,0} = 50$ and $g_1 = g_2 = 1$.

M	$(\mathcal{S}_0, v_0)$		$(N,p)=(6,6)$	$(N,p)=(7,7)$	$(N,p)=(8,8)$	$(N,p)=(9,9)$
1048576	$(95, 0.4)$	mean	4.4666	4.4503	4.4448	4.4369
		std	0.0138	0.0140	0.0145	0.0178
	$(100, 0.4)$	mean	5.7952	5.7897	5.7851	5.7824
		std	0.0200	0.0207	0.0196	0.0253
	$(105, 0.4)$	mean	5.6619	5.6384	5.6293	5.6384
		std	0.0250	0.0299	0.0350	0.0331
Avg Runtime [s]			3614.21	5123.17	7363.07	8898.37

Table 2.8: Results on numerical solutions $\mathcal{Y}_0^{0,(100,0.4)}$ of (2.50) under the HypHyp SLV model (2.52) at $(s_0, v_0) = (95, 0.4), (100, 0.4), (105, 0.4)$ using the LSMC methods with Laguerre polynomials up to p-th order for N_t time steps. Here, $M = 2^{20} = 1048576$ is the number of samples for the Monte-Carlo approximation. For each (s_0, v_0), we run the LSMC algorithm 50 times independently and collect each $\mathcal{Y}_0^{0,(s_0,v_0)}$. The row of "mean" and "std" reports their sample means and sample (unbiased) standard deviations, respectively. The bottom of the table reports the average of the total CPU times for the three cases.

2.6 Conclusion

A Markov BSDE driven by a CTMC associates with a system of ODEs. With arguments based on this observation, we proposed the multi-stage Euler-Maruyama methods for the BSDE, directly related to exponential integrators for solving the system of ODEs. Together with a suitable spatial discretization, these methods can be applied to solve BSDEs driven by Brownian motion. The efficiency of our numerical methods has been confirmed through numerical experiments using derivative pricing problems in mathematical finance.

2.7 Proofs

2.7.1 Proof of Theorem 2.2.2

Proof. For a solution U_t to (2.12), the Itô formula immediately implies $(Y_t, Z_t) = (X_t^* U_t, U_t)$ solves (2.9).

Let t_1 and t_2 be fixed. Without loss of generality, assume that $t_2 > t_1$.

$$\begin{aligned} Y_{t_1}^{t_1,e_i} - Y_{t_2}^{t_2,e_i} &= \mathbb{E}[Y_{t_1}^{t_1,e_i} - Y_{t_2}^{t_1,e_i} + (X_{t_2}^{t_1,e_i})^* V_{t_2} - e_i^* X_{t_2}^{t_2,e_i}] \\ &= \mathbb{E}\left[\int_{]t_1,t_2]} h(X_{u-}^{t_1,e_i}, u, Y_{u-}^{t_1,e_i}, Z_u^{t_1,e_i})du - \int_{]t_1,t_2]} dM_u^* Z_u^{t_1,e_i} \right. \\ &\qquad \left. + \left(\int_{]t_1,t_2]} Q_u^* X_{u-}^{t_1,e_i} - M_{t_2} + M_{t_1}\right)^* V_{t_2}\right] \\ &= \mathbb{E}\left[\int_{]t_1,t_2]} [h(X_{u-}^{t_1,e_i}, u, Y_{u-}^{t_1,e_i}, Z_u^{t_1,e_i}) + (X_{u-}^{t_1,e_i})^* Q_u V_{t_2}]du\right] \end{aligned}$$

Hence,

$$|Y_{t_1}^{t_1,e_i} - Y_{t_2}^{t_2,e_i}| \leq C\sqrt{t_2 - t_1}\sqrt{\int_{]t_1,t_2]} \mathbb{E}[|h(X_{u-}^{t_1,e_i}, u, Y_{u-}^{t_1,e_i}, Z_u^{t_1,e_i}) + (X_{u-}^{t_1,e_i})^* Q_u V_{t_2}|^2]du}$$

Using the uniform boundedness of Q_u, the Lipschitz continuity of h, evaluate the integrand as

$$\begin{aligned} &|h(X_{u-}^{t_1,e_i}, u, Y_{u-}^{t_1,e_i}, Z_u^{t_1,e_i}) + (X_{u-}^{t_1,e_i})^* Q_u V_{t_2}|^2 \\ &\qquad \leq C(|Y_{u-}^{t_1,e_i}|^2 + \|Z_u^{t_1,e_i}\|^2_{X_{u-}^{t_1,e_i}} + |h(X_{u-}^{t_1,e_i}, u, 0, 0)|^2 + \sup_{\substack{0\leq s,u\leq T \\ i=1,\dots,N}} |e_j^* Q_s V_u|^2). \end{aligned}$$

Recall that Q_s is assumed to be uniform bounded and that

$$\begin{aligned} \mathbb{E}\int_{]t_1,t_2]} |h(X_{u-}^{t_1,e_i}, u, 0, 0)|^2 du &= \sum_{j=1}^N \int_{]t_1,t_2]} |h(e_j, u, 0, 0)|^2 \mathbb{P}(X_u^{t_1,e_i} = e_j | X_{t_1}^{t_1,e_i} = e_i)du \\ &\leq C \sup_{j=1,\dots,N} \int_{]0,T]} |h(e_j, u, 0, 0)|^2 du. \end{aligned}$$

Hence,

$$\begin{aligned} |Y_{t_1}^{t_1,e_i} - Y_{t_2}^{t_2,e_i}| \leq C\sqrt{t_2 - t_1} \left(\mathbb{E}\left[\sup_{0\leq u\leq T} |Y_u^{t_1,e_i}|^2 + \int_{]0,T]} \|Z_u^{t_1,e_i}\|^2_{X_{u-}^{t_1,e_i}} \right. \right. \\ \left. \left. + \sup_{i=1,\dots,N} \int_{]0,T]} |h(e_i, u, 0, 0)|^2 du + 1\right] du\right), \end{aligned}$$

from which the continuity of $t \mapsto Y_t^{t,e_i}$ directly follows, as well as of V_t. Using

$$(\Delta X_u^{t,e_i})^* V_u = \Delta((X_u^{t,e_i})^* V_u) = \Delta Y_u^{u, X_u^{t,e_i}} = \Delta Y_u^{t,e_i} = \Delta M_u^* Z_u^{t,e_i} = (\Delta X_u^{t,e_i})^* Z_u^{t,e_i},$$

we obtain $\int_{]0,t]}(dX_u^{t,e_i})^*(Z_u^{t,e_i}-V_u)=0$. Notice that

$$\int_{]0,t]} dM_u^*(Z_u^{t,e_i}-V_u) = -\int_{]0,t]}(X_{u-}^{t,e_i})^* Q_u(Z_u^{t,e_i}-V_u)du = 0,$$

since any predictable finite variation martingales starting at 0 takes zero constantly (e.g. Corollary 8.2.14, p.204 in [19].) Hence

$$\mathbb{E}\int_{]0,T]}\|Z_u^{t,e_i}-V_u\|^2_{X_{u-}^{t,e_i}}du = \mathbb{E}\left|\int_{]0,T]} dM_u^*(Z_u^{t,e_i}-V_u)\right|^2 = 0,$$

which means $Z_u^{t,e_i}\sim_M V_u$. Together with the Lipschitz continuity,

$$h(X_{u-}^{t,e_i},u,(X_{u-}^{t,e_i})^*V_u,Z_u^{t,e_i}) = h(X_{u-}^{t,e_i},u,(X_{u-}^{t,e_i})^*V_u,V_u),\quad du\otimes d\mathbb{P}\text{-a.s.}$$

Plugging it into the conditional expectation representation of Y_t^{t,e_i}, we obtain for $t\in[0,T]$,

$$e_i^*V_t = Y_t^{t,e_i} = e_i^*\Phi(T,t)G + e_i^*\int_{]t,T]}\Phi(u,t)H(u,V_s)du,$$

which results in the variation-of-constants of (2.12) in what follows:

$$V_t = \Phi(T,t)G + \int_t^T \Phi(s,t)H(s,V_s)ds.$$

□

2.7.2 Proof of Proposition 2.4.1

Proof. We show $e_j^*Qe_i\geq 0$ for $i\neq j$ and $\sum_i e_j^*Qe_i=0$ for all $j=1,\dots,N$. It is trivial for $j=1$ and N since $e_1^*Q=e_N^*Q=0$. Let $j=2,\dots,N-1$ be fixed. The condition $\sum_{i=1}^N e_j^*Qe_i=0$ clearly holds since

$$\begin{aligned} e_i^*Qe_{i-1}+e_i^*Qe_{i+1}^* &= \frac{\sigma^2(x_i)-\delta x_i\mu(x_i)}{\delta x_{i-1}(\delta x_{i-1}+\delta x_i)} + \frac{\sigma^2(x_i)+\delta x_{i-1}\mu(x_i)}{\delta x_i(\delta x_{i-1}+\delta x_i)} \\ &= -\frac{(\delta x_i-\delta x_{i-1})\mu(x_i)-\sigma^2(t,x_i)}{\delta x_i\delta x_{i-1}} = -e_i^*Qe_i. \end{aligned}$$

We remain to prove the nonnegativity of off-diagonal elements of Q. Denote $\mu(x_i)=\mu(x_i)^+-\mu(x_i)^-$, where

$$\mu(x)^+=\max\{\mu(x),0\}\,(\geq 0)\quad\text{and}\quad \mu(x)^-=-\min\{\mu(x),0\}\,(\geq 0).$$

We obtain

$$\frac{\sigma^2(x_i) - \delta x_i \mu(x_i)}{\delta x_{i-1}(\delta x_{i-1} + \delta x_i)} = \frac{\mu(x_i)^-}{\delta x_{i-1}} + \frac{\sigma^2(s_i) - (\delta x_{i-1}\mu(x_i)^- + \delta x_i \mu(x_i)^+)}{\delta x_{i-1}(\delta x_{i-1} + \delta x_i)} \tag{2.53}$$

and

$$\frac{\sigma^2(x_i) + \delta x_{i-1} \mu(x_i)}{\delta x_i(\delta x_{i-1} + \delta x_i)} = \frac{\mu(x_i)^+}{\delta x_i} + \frac{\sigma^2(x_i) - (\delta x_{i-1}\mu(x_i)^- + \delta x_i \mu(x_i)^+)}{\delta x_i(\delta x_{i-1} + \delta x_i)}. \tag{2.54}$$

The first terms on the right-hand side of (2.53) and (2.54) are clearly nonnnegative. Under the condition (2.39), we obtain

$$\begin{aligned}\sigma^2(x_i) \geq \max_{1\leq j\leq N-1}\{\delta x_j\} \cdot |\mu(x_i)| &= \max_{1\leq j\leq N-1}\{\delta x_j\}(\mu^+(x_i) + \mu^-(x_i)) \\ &\geq \delta x_{i-1}\mu^-(x_i) + \delta x_i \mu^+(x_i),\end{aligned}$$

so that the second terms in (2.53) and (2.54) are nonnegative. Moreover, if (2.39) is strict, (2.53) and (2.54) are positive. □

2.7.3 Proof of Proposition 2.4.2

In this subsection, I_N is the $N \times N$ identity matrix, δ_{ij} is the Kronecker's delta, $\mathbb{1} = (1, 1, \ldots, 1)^* \in \mathbb{R}^N$, and

$$\psi_{M_1, e_i} = \operatorname{diag}(M_1^* e_i) - M_1^* \operatorname{diag}(e_i) - \operatorname{diag}(e_i) M_1. \tag{2.55}$$

for $N \times N$ matrix M_1. We sometimes omit subscripts when they can be unambiguously determined from the context.

Step 1. Let $h(e_i, t, y, z) \coloneqq f(t, x_i, e_i^* z, \sigma^*(x_i)(e_i^* \widetilde{D}_1^{(1)} z, \ldots, e_i^* \widetilde{D}_1^{(d)} z)^*)$ for $t \in [0, T]$, $i = 1, \ldots, N$, $y \in \mathbb{R}$ and $z \in \mathbb{R}^N$. The Lipschitz continuity for f implies

$$\begin{aligned}&|h(e_i, t, y, z) - h(e_i, t, y', z')|^2 \\ &\qquad \leq L(|y - y'|^2 + \|\sigma^*(x_i)(e_i^* \widetilde{D}_1^{(1)}(z - z'), \ldots, e_i^* \widetilde{D}_1^{(d)}(z - z'))^*\|^2).\end{aligned}$$

To obtain the desired result, it is sufficient to show

$$\|\sigma^*(x_i)(e_i^* \widetilde{D}_1^{(1)}(z - z'), \ldots, e_i^* \widetilde{D}_1^{(d)}(z - z'))^*\|^2 \leq C\|z - z'\|_{e_i}^2 \tag{2.56}$$

for some constant $C > 0$ for any $i = 1, \dots, N$. As the left-hand side of (2.56) can be represented as a quadratic form of symmetric matrix

$$M_0 \coloneqq \sum_{p,q=1}^{d} (\sigma\sigma^*)^{(p,q)}(x_i)(\widetilde{D}_1^{(q)})^* e_i e_i^* \widetilde{D}_1^{(p)}, \tag{2.57}$$

(2.56) is equivalent to the positive semi-definiteness of $M_{i,C} \coloneqq C\psi_{Q,e_i} - M_0$. Before showing this, we require several lemmas in Step 2.

Step 2. In this step, for any $N \times N$ matrix M_1, we denote $M_1^{(N-1)}$ as the $(N-1) \times (N-1)$ matrix obtained by removing the last row and column vector of M_1.

Lemma 2.7.1. *Let M_1, M_2 be $N \times N$ real symmetric matrices satisfying $M_1\mathbb{1} = 0$ and $M_2\mathbb{1} = 0$. If $M_1^{(N-1)}$ is positive definite, $cM_1 - M_2$ is positive semi-definite for sufficiently large $c > 0$.*

Proof. Let $\lambda_1 > 0$ be the minimum eigenvalue of $M_1^{(N-1)}$, and λ_2 be the maximum eigenvalue of $M_2^{(N-1)}$. $cM_1^{(N-1)} - M_2^{(N-1)}$ is positive definite for $c > (\lambda_2/\lambda_1) \vee 0$. Indeed, for any non-zero vector z, we see that

$$z^*(cM_1^{(N-1)} - M_2^{(N-1)})z \geq c\lambda_1\|z\|^2 - \lambda_2\|z\|^2 > 0.$$

Since any real symmetric matrix M_0 satisfying $M_0\mathbb{1} = 0$ has the following block matrix representation

$$M_0 = \begin{pmatrix} M_0^{(N-1)} & -M_0^{(N-1)}\mathbb{1} \\ -\mathbb{1}^* M_0^{(N-1)} & \mathbb{1}^* M_0^{(N-1)}\mathbb{1} \end{pmatrix},$$

the quadratic form of M can be written as

$$z^* M_0 z = (x^*, y) \begin{pmatrix} M_0^{(N-1)} & -M_0^{(N-1)}\mathbb{1} \\ -\mathbb{1}^* M_0^{(N-1)} & \mathbb{1}^* M_0^{(N-1)}\mathbb{1} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x - y\mathbb{1})^* M_0^{(N-1)}(x - y\mathbb{1}), \tag{2.58}$$

for $z = (x^*, y)^* \in \mathbb{R}^N$. Applying $M_0 = cM_1 - M_2$ to (2.58), the quadratic form takes a positive value for $z = (x^*, y)$ except $x - y\mathbb{1} = 0$, and takes zero if $x - y\mathbb{1} = 0$. Hence the positive semi-definiteness of $cM_1 - M_2$ is obtained. □

Lemma 2.7.2. *For $i, j = 1, \dots, N$ and any matrix M_1,*

$$\psi_{M_1,e_i} e_j = (e_i^* M_1 e_j)(e_j - e_i) - \delta_{ij} M_1^* e_j. \tag{2.59}$$

Proof. We can see it directly as:

$$\begin{aligned}\psi_{M_1,e_i}e_j &= \mathrm{diag}(M_1^*e_i)e_j - M_1^*e_ie_i^*e_j - e_ie_i^*M_1e_j \\ &= e_i^*M_1e_j\cdot e_j - \delta_{ij}M_1^*e_i - e_i^*M_1e_j\cdot e_i = (e_i^*M_1e_j)(e_j-e_i) - \delta_{ij}M_1^*e_j.\end{aligned}$$

□

Lemma 2.7.3. *Let $i=1,\dots,N$ be fixed, and M_1 be a matrix satisfying $e_i^*M_1e_j>0$ for all $j\in\{1,\dots,N\}\setminus\{i\}$ and $e_i^*M_1\mathbb{1}=0$ (Note that $e_i^*M_1e_i=-\sum_{j=1,j\neq i}^N e_i^*M_1e_j$.) Then, $\psi^{(N-1)}_{M_1,e_i}$ is positive definite.*

Proof. Since (2.59), $e_i^*M_1e_j$ for $j\neq i$ is strictly positive under the assumption of this lemma. Thus $\psi^{(N-1)}_{M_1,e_N}=\mathrm{diag}(e_N^*M_1e_1,\dots,e_N^*M_1e_{N-1})$ is positive definite obviously. Let $x\in\mathbb{R}^{N-1}\setminus\{0\}$ and $i<N$ be fixed.

$$x^*\psi^{(N-1)}_{M_1,e_i}x = \sum_{j=1}^{N-1}\sum_{k=1}^{N-1}[(\delta_{jk}-\delta_{ki})e_i^*M_1e_j - \delta_{ij}e_i^*M_1e_k]e_k^*xe_j^*x \tag{2.60}$$

$$\begin{aligned}&= [(\delta_{ii}-\delta_{ii})e_i^*M_1e_i - \delta_{ii}e_i^*M_1e_i]e_i^*xe_i^*x \\ &\quad + \sum_{\substack{j=1\\j\neq i}}^{N-1}[(\delta_{ji}-\delta_{ii})e_i^*M_1e_j - \delta_{ij}e_i^*M_1e_i]e_i^*xe_j^*x \\ &\quad + \sum_{\substack{k=1\\k\neq i}}^{N-1}[(\delta_{ik}-\delta_{ki})e_i^*M_1e_i - \delta_{ii}e_i^*M_1e_k]e_k^*xe_i^*x \\ &\quad + \sum_{\substack{j=1\\j\neq i}}^{N-1}\sum_{\substack{k=1\\k\neq i}}^{N-1}[(\delta_{jk}-\delta_{ki})e_i^*M_1e_j - \delta_{ij}e_i^*M_1e_k]e_k^*xe_j^*x\end{aligned} \tag{2.61}$$

$$= -e_i^*M_1e_i(e_i^*x)^2 - 2\left[\sum_{\substack{j=1\\j\neq i}}^{N-1}e_i^*M_1e_je_j^*x\right]e_i^*x + \sum_{\substack{j=1\\j\neq i}}^{N-1}e_i^*M_1e_j(e_j^*x)^2. \tag{2.62}$$

Note that $-e_i^*M_1e_i>0$. The discriminant of (2.62) as a quadratic polynomial of e_i^*x

can be evaluated in what follows:

$$\begin{aligned}
&\left[\sum_{\substack{j=1\\j\neq i}}^{N-1}\sqrt{e_i^*M_1e_j}(\sqrt{e_i^*M_1e_j}e_j^*x)\right]^2+e_i^*M_1e_i\sum_{\substack{j=1\\j\neq i}}^{N-1}e_i^*M_1e_j(e_j^*x)^2\\
&\leq\sum_{\substack{j=1\\j\neq i}}^{N-1}e_i^*M_1e_j\sum_{\substack{k=1\\k\neq i}}^{N-1}e_i^*M_1e_j(e_j^*x)^2+e_i^*M_1e_i\sum_{\substack{j=1\\j\neq i}}^{N-1}e_i^*M_1e_j(e_j^*x)^2 \qquad (2.63)\\
&=\sum_{j=1}^{N-1}e_i^*M_1e_j\sum_{\substack{k=1\\k\neq i}}^{N-1}e_i^*M_1e_j(e_j^*x)^2,
\end{aligned}$$

where we have applied the Cauchy-Schwarz inequality to obtain the first inequality. The assumptions on M_1 leads to $\sum_{j=1}^{N-1}e_i^*M_1e_j=-e_i^*M_1e_N<0$, and the discriminant (2.63) is negative for any x. Hence (2.62) is always positive which amounts to the positive definiteness of $\psi^{(N-1)}_{M_1,e_i}$. □

Step 3. Suppose that e_i^*Q contains no elements that equal 0. Then, Q satisfies the assumptions of Lemma 2.7.3, and the positive definiteness of $\psi^{(N-1)}_{Q,e_i}$ is obtained. $\psi_{Q,e_i}\mathbb{1}=0$ is clear, and $M_0\mathbb{1}=0$ follows from assumption 2.41. Applying $M_1=\psi_{Q,e_i}\mathbb{1}$ and $M_2=M_0$ into Lemma 2.7.1, the positive semi-definiteness of $M_{i,C}$ is obtained for sufficiently large $C>0$.

Step 4. In the case of e_i^*Q possibly containing element 0, the following arguments are required for obtaining the desired result. To this end, We additionally introduce some notations: Denote $e_{i,N}$ as the i-th unit vector in $\mathbb{R}^N$ whose i-th element is 1. Note that we sometimes omit subscripts N and simply write e_i when they can be unambiguously determined from the context. For N-dimensional vector v, denote $\mathcal{I}(v)=\{n_1,\ldots,n_K\}\subset\{1,\ldots,N\}$ as the collection of indices of the elements that are non-zero. Similarly, for a $N\times N$ real symmetric matrix M_1, denote $\mathcal{I}(M_1)=\{n_1,\ldots,n_K\}\subset\{1,\ldots,N\}$ as the collection of indices of non-zero row/column vectors in M_1. In both cases, n_k is sorted in ascending order, and K means the total number. For a $N\times N$ matrix M_1 and a collection of indices $\mathcal{J}=\{n_1,\ldots,n_K\}\subset\{1,\ldots,N\}$, denote $M_1^{\mathcal{J}}$ as a $K\times K$ matrix obtained by removing the rows and column vectors that do not belong to $\mathcal{J}$. Equivalently, it can be defined by

$$M_1^{\mathcal{J}}=I_{\mathcal{J},N}M_1I_{\mathcal{J},N}^*, \qquad (2.64)$$

where $I_{\mathcal{J},N}$ is a $N \times K$ matrix as

$$I_{\mathcal{J},N} = \begin{pmatrix} e_{n_1,N} & \dots & e_{n_K,N} \end{pmatrix}.$$

Note that $I_{\mathcal{J},N} I_{\mathcal{J},N}^* = I_K$. First, we confirm the follwing two lemmas.

Lemma 2.7.4. *Let $\mathcal{J} = \{n_1, \dots, n_K\} \subset \{1, \dots, N\}$ be a collection of indices. For any $N \times N$ matrix M_1,*

$$e_{k,K}^* M_1^{\mathcal{J}} e_{l,K} = e_{n_k,N}^* M_1 e_{n_l,N} \quad \textit{for} \quad k, l = 1, \dots, K.$$

Proof. It immediately follows from (2.64):

$$\begin{aligned} e_{k,K}^* M_1^{\mathcal{J}} e_{l,K} &= e_{k,K}^* I_{\mathcal{J},N}^* M_1 I_{\mathcal{J},N} e_{l,K} \\ &= e_{k,K}^* \begin{pmatrix} e_{n_1,N}^* \\ \vdots \\ e_{n_K,N}^* \end{pmatrix} M_1 \begin{pmatrix} e_{n_1,N} & \dots & e_{n_K,N} \end{pmatrix} e_{l,K} = e_{n_k,N}^* M_1 e_{n_l,N}, \end{aligned}$$

for $k = 1, \dots, K$. □

Lemma 2.7.5. *Let $\mathcal{J} = \{n_1, \dots, n_K\} \subset \{1, \dots, N\}$ be a collection of indices and M_1 a $N \times N$ matrix satisfying $M_1 \mathbb{1} = 0$. If $\mathcal{I}(M_1) \subset \mathcal{J}$, then $M_1^{\mathcal{J}} \mathbb{1} = 0$ holds.*

Proof. Note that $M_1 e_{i,N} = 0$ for $i \notin \mathcal{J}$ since $\mathcal{I}(M_1) \subset \mathcal{J}$. Hence,

$$\begin{aligned} e_{k,K}^* M_1^{\mathcal{J}} \mathbb{1} &= \sum_{l=1}^{K} e_{k,K}^* M_1^{\mathcal{J}} e_{l,K} = \sum_{i \in \mathcal{J}} e_{n_k,N}^* M_1 e_{i,N} + 0 \\ &= \sum_{i \in \mathcal{J}} e_{n_k,N}^* M_1 e_{i,N} + \sum_{i \notin \mathcal{J}} e_{n_k,N}^* M_1 e_{i,N} = \sum_{i=1}^{N} e_{n_k,N}^* M_1 e_{i,N} = e_{n_k,N}^* M_1 \mathbb{1} = 0, \end{aligned}$$

for $k = 1, \dots, K$. □

Now, we proceed to show the positive semi-definiteness of $M_{i,C}$. If $e_i^* Q = 0$, it is trivial since $M_{i,C}$ equals the zero matrix. Hereafter, we suppose that $e_i^* Q \neq 0$. Clearly, $M_{i,C}$ is positive semi-definite if $M_{i,C}^{\mathcal{I}(M_{i,C})}$ is positive semi-definite. We confirm the relationship between $\mathcal{I}(M_{i,C})$, $\mathcal{I}(e_i^* Q)$, $\mathcal{I}(\psi_{Q,e_i})$ as well as $\mathcal{I}(M_0)$.

Lemma 2.7.6. *Let $N \in \mathbb{N}$ be fixed. ($e_{i,N}$ is abbreviated to e_i in this lemma and its proof.)*

1. *For all $C > 0$, $i, j = 1, \dots, N$,*

$$\psi_{Q,e_i} e_j = 0 \iff e_i^* Q e_j = 0 \Longrightarrow M_0 e_j = 0, \tag{2.65}$$

$$e_i^* Q e_j = 0 \Longrightarrow M_{i,C} e_j = 0. \tag{2.66}$$

2. *For sufficiently large $C > 0$,*

$$M_{i,C} e_j = 0 \Longrightarrow \psi_{Q,e_i} e_j = 0, \tag{2.67}$$

for $i, j = 1, \dots, N$.

Therefore $\mathcal{I}(M_{i,C}) = \mathcal{I}(e_i^* Q) = \mathcal{I}(\psi_{Q,e_i}) \supset \mathcal{I}(M_0)$ holds for sufficiently large $C > 0$.

Proof. 1. For $i \neq j$, $\psi_{Q,e_i} e_j = (e_i^* Q e_j)(e_j - e_i)$ implies that $e_i^* Q e_j = 0$ if and only if $\psi_{Q,e_i} e_j = 0$. For $i = j$, $0 = \psi_{Q,e_i} e_i = -Q^* e_i \iff e_i^* Q e_i = 0$. Hence $\psi_{Q,e_i} e_j = 0 \iff e_i^* Q e_j = 0$ is obtained for any i, j.

Next, assume that $e_i^* Q e_j = 0$. (2.41) leads to $e_i^* \widetilde{D}_1^{(p)} e_j = 0$ for all $p = 1, \dots, d$, which yields $M_0 e_j = 0$.

Finally, (2.66) can be seen using (2.65).

2. Take

$$C > \sup \left\{ \frac{e_k^* M_0 e_j}{e_k^* \psi_{Q,e_i} e_j} : k = 1, \dots, N \quad \text{and} \quad e_k^* \psi_{Q,e_i} e_j \neq 0 \right\} \vee 0,$$

where $\sup \emptyset = -\infty$. Assuming that $e_k^* \psi_{e_i} e_j \neq 0$ for some $k \in \{1, \dots, N\}$,

$$e_k^* M_{i,C} e_j = C e_k^* \psi_{Q,e_i} e_j - e_k^* M_0 e_j > 0.$$

It contradicts that $e_k^* M_{i,C} e_j = 0$.

□

Denote $\mathcal{I} = \{n_1, \dots, n_K\} \coloneqq \mathcal{I}(M_{i,C}) = \mathcal{I}(\psi_{Q,e_{i,N}}) = \mathcal{I}(Q^* e_{i,N})$, and denote K as its total number. Observe that

$$M_{i,C}^{\mathcal{I}} = C \psi_{Q,e_{i,N}}^{\mathcal{I}} - M_0^{\mathcal{I}}. \tag{2.68}$$

As we suppose $e_i^* Q \neq 0$, $e_i^* Q e_i < 0$ and hence i belongs to $\mathcal{I}$. Take $i \in \{1, \dots, K\}$ that satisfies $i = n_q$. For any $k, l = 1, \dots, K$,

$$\begin{aligned} e_{l,K}^* \psi_{Q,e_p}^{\mathcal{I}} e_{k,K} &= e_{n_l,N}^* \psi_{Q,e_p} e_{n_k,N} = e_{n_l,N}^* [e_{p,N}^* Q e_{n_k,N} (e_{n_k,N} - e_{p,N}) - \delta_{p,n_k} Q^* e_{n_k,N}] \\ &= e_{q,K}^* Q^{\mathcal{I}} e_{k,K} e_{l,K}^* (e_{k,K} - e_{q,K}) - \delta_{e,k} e_{l,K}^* (Q^{\mathcal{I}})^* e_{k,K} \\ &= e_{l,K}^* [e_{q,K}^* Q^{\mathcal{I}} e_{k,K} (e_{k,K} - e_{q,K}) - \delta_{e,k} (Q^{\mathcal{I}})^* e_{k,K}] = e_{l,K}^* \psi_{Q^{\mathcal{I}},e_{q,K}} e_{k,K}, \end{aligned}$$

where we used Lemma 2.7.4. As a result, we obtain $\psi_{Q,e_{i,N}}^{\mathcal{I}} = \psi_{Q^{\mathcal{I}},e_{q,K}}$. Notice what follows:

- $\psi_{Q^{\mathcal{I}},e_{q,K}}$ and $M_0^{\mathcal{I}}$ are symmetric.
- Since $\mathcal{I}(\psi_{Q,e_{i,N}}), \mathcal{I}(M_0) \subset \mathcal{I}$, Lemma 2.7.5 implies $\psi^{\mathcal{I}}_{Q,e_{i,N}}\mathbb{1} = 0$ (i.e. $\psi_{Q^{\mathcal{I}},e_{q,K}}\mathbb{1}$) and $M_0^{\mathcal{I}}\mathbb{1} = 0$.
- Since $Q^{\mathcal{I}} = Q^{\mathcal{I}(e_i^* Q)}$, $e_q^* Q^{\mathcal{I}} e_l > 0$ for $k \in \{1, \dots, K\} \setminus \{q\}$ and $e_q^* Q^{\mathcal{I}}\mathbb{1} = 0$. Applying $M_1 = Q^{\mathcal{I}}$ to Lemma 2.7.3, we obtain the positive definiteness of $\psi^{(K-1)}_{Q^{\mathcal{I}},e_{q,K}}$.

Thus we can apply $M_1 = \psi_{Q^{\mathcal{I}},e_{q,K}}$ and $M_2 = M_0^{\mathcal{I}}$ to Lemma 2.7.1, and we obtain the positive semi-definiteness of $M^{\mathcal{I}}_{i,C}$.

2.8 Convergence Results

In this section, we establish a convergence result for the numerical solution discussed in Section 4 to the true solution of BSDE (2.14). For simplicity, we only consider a situation with (i) one-dimensional space variable, (ii) the corresponding PDE being uniquely solvable in the classical sense, as well as (iii) a spatial discretization using central difference with constant step size i.e. $\delta x_i \equiv \Delta x > 0$.

Throughout the section, the following notations are introduced;

$$\begin{aligned} \breve{f} : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \ni (t,x,z,p,r) \\ \mapsto \mu(x) \cdot p + \frac{\sigma^2(x)}{2} \cdot r + f(t,x,z,\sigma(x) \cdot p) \in \mathbb{R}. \end{aligned} \tag{2.69}$$

$$U_t^{(N,k)} \text{ is a unique solutions of (2.34) for } N \text{ and } k. \tag{2.70}$$

Denote (2.16) as

$$\begin{cases} \partial_t u(t,x) + \breve{f}(t,x,u(t,x),\partial_x u(t,x),\partial_{xx} u(t,x)) = 0, \quad (t,x) \in [0,T] \times \mathbb{R}, \\ u(T,x) = g(x), \quad x \in \mathbb{R}. \end{cases} \tag{2.71}$$

The present analysis in this section is mainly based on the textbook written by Walter et al.[72]. Suppose that the following conditions are satisfied.

Assumption 2.8.1 ([72], pp.287 and 302). *1. $\mu(x), \sigma(x)$ and $f(t,x,y,z)$ are twice continuously differentiable in all variables. f and its first- and second-order derivatives are bounded and uniformly continuous in $(t,x,y,z) \in [0,T] \times \mathbb{R} \times B$ for any bounded set $B \subset \mathbb{R}^2$.*

2. *For some constant* $C > 0$ *and a continuous function* $\lambda : [0,\infty) \to (0,\infty)$ *satisfying*

$$\lim_{s\to\infty} \lambda(s) = \infty \quad \textit{and} \quad \int_0^\infty \frac{ds}{s\cdot\lambda(s)} = \infty, \tag{2.72}$$

it holds

$$f(t,x,y,0)\vee(-f(t,x,-y,0)) \le C + y\cdot\lambda(y), \quad \textit{for any} \quad (t,x,y) \in [0,T]\times\mathbb{R}\times[0,\infty).$$

3. $\sigma^2(x) > 0$ *for any* $x \in \mathbb{R}$.

4. *For any* $M > 0$, *there exists* $C_0 > 0$ *and* $\lambda_0 : [0,\infty) \to (0,\infty)$ *satisfying the following conditions:*

   - λ_0 *is continuous and satisfies* (2.72).
   - *For* $(t,x) \in [0,T]\times\mathbb{R}$, $|y| \le M$ *and* $z \ge 0$,

$$|z|\partial_y f(t,x,y,\sigma(x)z) \le |z|\lambda_0(z) + C_0.$$

5. $g \in C_b^2(\mathbb{R})$.

We obtain the following lemmas. For details, see [72], Chapter IV, Section 36.

Lemma 2.8.1 ([72], p.292)**.** *Cauchy problem* (2.71) *with* $\mu(x), \sigma(x), f(t,x,y,z)$ *and* $g(x)$ *satisfying Assumption 2.8.1 is uniquely solvable if* g *is three times continuously differentiable and* g, $\partial_x g$, $\partial_{xx} g$ *and* $\partial_{xxx} g$ *are bounded and Lipschitz continuous.*

Lemma 2.8.2 ([72], p.301)**.** *If* (2.71) *satisfying Assumption 2.8.1 (not necessarily the case in Lemma 2.8.1) admits a unique solution* u, *then for every* $\Delta x > 0$, *the infinite system of ODEs,*

$$\begin{cases} \dfrac{de_i^* U_t^{(\infty,\Delta x)}}{dt} = \breve{f}(t, i\Delta x, e_i^* U_t^{(\infty,\Delta x)}, e_i^* D_1 U_t^{(\infty,\Delta x)}, e_i^* D_2 U_t^{(\infty,\Delta x)}), \\ e_i^* U_T^{(\infty,\Delta x)} = g(i\Delta x), \end{cases} \quad (t,i) \in [0,T]\times\mathbb{Z}, \tag{2.73}$$

which is derived from the spatial discretization described in Section 4 with step size Δx, *admits a unique solution* $U^{(\infty,\Delta x)} : [0,T] \ni t \mapsto U_t^{(\infty,\Delta x)} = (e_i^* U_t^{(\infty,\Delta x)})_{i\in\mathbb{Z}} \in l^\infty$. *Here,* l^∞ *is the Banach space consisting of all real sequences* $x = (x_i)_{i\in\mathbb{Z}}$ *with finite supremum norm* $\|x\|_\infty = \sup_{i\in\mathbb{Z}} |x_i|$, *and*

$$\begin{aligned} e_i^* D_1 U_t^{(\infty,\Delta x)} &:= \frac{-1}{2\Delta x} e_{i-1}^* U_t^{(\infty,\Delta x)} + \frac{1}{2\Delta x} e_{i+1}^* U_t^{(\infty,\Delta x)}, \\ e_i^* D_2 U_t^{(\infty,\Delta x)} &:= \frac{1}{\Delta x^2} e_{i-1}^* U_t^{(\infty,\Delta x)} + \frac{-2}{\Delta x^2} e_i^* U_t^{(\infty,\Delta x)} + \frac{1}{\Delta x^2} e_{i+1}^* U_t^{(\infty,\Delta x)}, \end{aligned}$$

for $i \in \mathbb{Z}$. Furthermore, $U^{(\infty,\Delta x)}$ converges to u in the following sense; for any compact set $K \subset \mathbb{R}$,

$$\lim_{k\to 0} \sup_{\substack{t\in[0,T], i\Delta x\in\mathbb{Z} \\ i\Delta x\in K}} |u(t, i\Delta x) - e_i^* U_t^{(\infty,\Delta x)}| = 0.$$

Note that Walter et al. [72] do not consider results on the convergence of "truncated" finite systems of ODEs (2.34) to (2.73). Fortunately, it can be carried out by the standard diagonalization argument using the Arzelà-Ascoli theorem. To this end, suppose the following Lipschitz condition additionally.

Assumption 2.8.2. *There exists $L > 0$ such that*

$$|f(t, x, y, z) - f(t, x, y', z')| \le L(|y - y'| + |z - z'|)$$

for any $(t, x) \in [0, T] \times \mathbb{R}, y, y', z, z' \in \mathbb{R}$.

It leads to the Lipschitz continuity of F defined by (3.31), and (2.34) admits a unique solution $U^{(N,\Delta x)}$. Then, we define $\overline{U}^{(N,\Delta x)} = (e_i^* \overline{U}^{(N,\Delta x)})_{i\in\mathbb{Z}} \in l^\infty$ for $N \in \mathbb{N}$ as

$$e_i^* \overline{U}^{(N,\Delta x)} = \begin{cases} e_N^* U^{(N,\Delta x)}, & \text{if} \quad i = N+1, N+2, \dots, \\ e_i^* U^{(N,\Delta x)}, & \text{if} \quad i = -N, -N+1, \dots, N-1, N, \\ e_{-N}^* U^{(N,\Delta x)}, & \text{if} \quad i = -N-1, -N-2, \dots. \end{cases}$$

Lemma 2.8.3. *Suppose that Assumption 2.8.1 and 2.8.2 hold, and that Cauchy problem* (2.71) *admits a unique solution u. Then, for any $\Delta x > 0$,*

$$\lim_{N\to\infty} \sup_{t\in[0,T], -N\le i\le N} |e_i^* U_t^{(\infty,\Delta x)} - e_i^* \overline{U}_t^{(N,\Delta x)}|_\infty = 0.$$

Here, $U_t^{(\infty,\Delta x)} \in l^\infty$ is a unique solution of (2.73).

Proof. Let $\Delta x > 0$ be fixed.

Uniform Boundedness of $(\overline{U}_t^{(N,\Delta x)})_{N=1}^\infty$. Note that

$$\sup_{N\in\mathbb{N}} |e_i^* U_T^{(N,\Delta x)}| \le \|g(x)\|_\infty, \qquad \sup_{(t,x)\in[0,T]\times\mathbb{R}} |f(t, x, 0, 0)| < C, \tag{2.74}$$

which follow from 5 and 2 in Assumption 2.8.1, respectively. For $N \in \mathbb{N}$ and $i = \pm N$, 2 in Assumption 2.8.1 and the Lipschitz continuity 2.8.2 lead to

$$|f(t, i\Delta x, e_i^* U_t^{(N,\Delta x)}, \sigma(i\Delta x) e_i^* D_1 U_t^{(N\Delta x)}))| = |f(t, i\Delta x, e_i^* U_t^{(N,\Delta x)}, 0)| \le C + L|e_i^* U_t^{(N,\Delta x)}|.$$

These estimates imply

$$|e_i^* U_t^{(N,\Delta x)}| \le C + \|g(x)\|_\infty + L\int_t^T |e_i^* U_s^{(N,\Delta x)}|ds \quad \text{for} \quad i = \pm N. \tag{2.75}$$

Applying the Grönwall inequality to (2.75), we obtain

$$|e_{-N}^* U_t^{(N,\Delta x)}| \vee |e_N^* U_t^{(N,\Delta x)}| \le e^{LT}\cdot[C + \|g(x)\|_\infty] =: c_1, \tag{2.76}$$

for any $N \in \mathbb{N}$ and $t \in [0,T]$. For $N > i > -N$, notice that using (2.74) and the Lipschitz continuity,

$$\begin{aligned}
&|f(t, i\Delta x, e_i^* U_t^{(N,\Delta x)}, \sigma(i\Delta x) e_i^* D_1 U_t^{(N,\Delta x)})| \\
&\qquad \le C + L\left[|e_i^* U_t^{(N,\Delta x)}| + \|\sigma\|_\infty \frac{|e_{i+1}^* U_t^{(N,\Delta x)}| + |e_{i-1}^* U_t^{(N,\Delta x)}|}{2\Delta x}\right] \\
&\qquad \le (C \vee L \vee L\frac{\|\sigma\|_\infty}{2\Delta x})[1 + |e_i^* U_t^{(N,\Delta x)}| + |e_{i+1}^* U_t^{(N,\Delta x)}| + |e_{i-1}^* U_t^{(N,\Delta x)}|],
\end{aligned}$$

and

$$\begin{aligned}
&\mu(ki)\cdot\frac{|e_{i+1}^* U_t^{(N,\Delta x)}| + |e_{i-1}^* U_t^{(N,\Delta x)}|}{2k} \\
&\qquad + \frac{\sigma^2(i\Delta x)}{2}\frac{2|e_{i+1}^* U_t^{(N,\Delta x)}| + |e_i^* U_t^{(N,\Delta x)}| + 2|e_{i-1}^* U_t^{(N,\Delta x)}|}{2k^2} \\
&\qquad \le 2\left(\frac{\|\mu\|_\infty}{2\Delta x} \vee \frac{\|\sigma\|^2}{4\Delta x^2}\right)\sum_{j=i-1}^{i+1} |e_j^* U_t^{(N,\Delta x)}|.
\end{aligned}$$

Let $C' := C \vee L \vee L\frac{\|\sigma\|_\infty}{2\Delta x}$, $C'' = 2\left(\frac{\|\mu\|_\infty}{2k} \vee \frac{\|\sigma\|^2}{4\Delta x^2}\right)$, and $c_2 := C' + C''$. Since

$$\begin{aligned}
|e_i^* U_t^{(N,\Delta x)}| &\le \|g(x)\|_\infty + C'T + c_2\int_t^T \sum_{m=i-1}^{i+1} |e_m^* U_s^{(N,\Delta x)}|ds \\
&\le \|g(x)\|_\infty + C'T + c_2\int_t^T \sup_{|j|<N}\sum_{m=j-1}^{j+1} |e_m^* U_s^{(N,\Delta x)}|ds \\
&\le \|g(x)\|_\infty + C'T + 2c_2c_1T + 3c_2\int_t^T \sup_{|j|<N} |e_j^* U_s^{(N,\Delta x)}|ds
\end{aligned}$$

for $N > i > -N$, we obtain

$$\sup_{t\in[0,T],|i|<N} |e_i^* U_t^{(N,\Delta x)}| \le e^{3c_2T}[\|g(x)\|_\infty + C'T + 2c_2c_1T] =: c_3.$$

Therefore, $\sup_{t\in[0,T]} \|\overline{U}_t^{(N,\Delta x)}\|_\infty = \sup_{|i|\le N} |e_i^* U_t^{(N,\Delta x)}| \le c_1 \vee c_3 =: c_4$.

Uniform Boundedness of $(\frac{dU_t^{(N,\Delta x)}}{dt})_{N=1}^\infty$ It follows from

$$\left|\frac{de_i^* U_t^{(N,\Delta x)}}{dt}\right| \le C + L|e_i^* U_t^{(N,\Delta x)}| \le C + Lc_4$$

for $i = \pm N$, and

$$\left|\frac{de_i^* U_t^{(N,\Delta x)}}{dt}\right| \le C' + 2c_2c_1 + 3c_2 \sup_{|i|<N} |e_i^* U_s^{(N,\Delta x)}| \le C' + 2c_2c_1 + 3c_2c_4$$

for $N > i > -N$. $(\frac{dU_t^{(N,\Delta x)}}{dt})_{N=1}^\infty$ has an uniform upper bound $c_5 \coloneqq (C + Lc_4) \vee (C' + 2c_2c_1 + 3c_2c_4)$.

Equicontinuity of $(U_t^{(N,\Delta x)})_{N=1}^\infty$ It follows from the mean value theorem and the uniform boundedness of $(\frac{dU_t^{(N,\Delta x)}}{dt})_{N=1}^\infty$. Precisely, for any $t, t' \in [0,T]$, $-N \le i \le N$, and $N \in \mathbb{N}$,

$$|U_t^{(N,\Delta x)} e_i - U_{t'}^{(N,\Delta x)} e_i| \le \sup_{s\in[0,T]} \left|\frac{dU_s^{(N,\Delta x)} e_i}{ds}\right| |t - t'| < c_5|t - t'|,$$

which immediately leads to the equicontinuity.

Equicontinuity of $(\frac{dU_t^{(N,\Delta x)}}{dt})_{N=1}^\infty$**.** Let $s, t \in [0,T]$ be fixed. Since $(\frac{dU_t^{(N,\Delta x)}}{dt})_{N=1}^\infty$ is uniformly bounded, using the mean value theorem,

$$\begin{aligned}
&|f(s, i\Delta x, U_s^{(N,\Delta x)} e_i, \sigma(i\Delta x) e_i^* D_1 U_s^{(N,\Delta x)}) - f(s, i\Delta x, e_i^* U_t^{(N,\Delta x)}, \sigma(i\Delta x) e_i^* D_1 U_t^{(N,\Delta x)})| \\
&\qquad \le L\left(1 \vee \frac{\|\sigma\|_\infty}{\Delta} x\right) \sum_{j=i-1}^{i+1} |e_j^* U_s^{(N,\Delta x)} - e_j^* U_t^{(N,\Delta x)}| \\
&\qquad = \underbrace{3Lc_5\left(1 \vee \frac{\|\sigma\|_\infty}{\Delta} x\right)}_{=:c_6} \cdot |s - t| = c_6|s - t|
\end{aligned} \tag{2.77}$$

and

$$
\begin{aligned}
&|f(s, i\Delta x, e_i^* U_t^{(N,\Delta x)}, \sigma(i\Delta x) e_i^* D_1 U_t^{(N,\Delta x)}) - f(t, i\Delta x, e_i^* U_t^{(N,\Delta x)}, \sigma(i\Delta x) e_i^* D_1 U_t^{(N,\Delta x)})| \\
&\leq \sup_{t' \in [0,T]} |\partial_t f(t', i\Delta x, e_i^* U_t^{(N,\Delta x)}, \sigma(i\Delta x) e_i^* D_1 U_t^{(N,\Delta x)})| \cdot |s - t| \\
&\leq \sup_{(t',x,y,z) \in [0,T] \times \mathbb{R}^3} |\partial_t f(t', x, y, z)| \cdot |s - t| < c_7 |s - t|.
\end{aligned}
\tag{2.78}
$$

Using (2.77), (2.78) and the triangle inequality,

$$
\begin{aligned}
&\left| \frac{de_i^* U_s^{(N,\Delta x)}}{dt} - \frac{de_i^* U_t^{(N,\Delta x)}}{dt} \right| \\
&\leq |f(s, i\Delta x, e_i^* U_s^{(N,\Delta x)}, \sigma(i\Delta x) e_i^* D_1 U_s^{(N,\Delta x)}) - f(t, i\Delta x, e_i^* U_t^{(N,\Delta x)}, \sigma(i\Delta x) e_i^* D_1 U_t^{(N,\Delta x)})| \\
&\qquad + C'' \left[|e_{i+1}^* U_s^{(N,\Delta x)}| + |e_i^* U_s^{(N,\Delta x)}| + |e_{i-1}^* U_s^{(N,\Delta x)}| \right] \\
&\leq [C'' + c_6 + c_7] \cdot |s - t|,
\end{aligned}
$$

from which $(\frac{dU_t^{(N,\Delta x)}}{dt})_{N=1}^{\infty}$ enjoys the equicontinuity.

Convergence. Let $(N_n)_{n=1}^{\infty}$ be an arbitrary subsequence of $\mathbb{N}$. Applying the Arzelà-Ascoli theorem ([56], P.290) guarantees the existence of convergent subsequences $U_t^{(N_{n(j)},\Delta x)}$ and $\dfrac{dU_t^{(N_{n(j)},\Delta x)}}{dt}$. Let $U_t^{(N_{n(\infty)},\Delta x)}$ and $\dfrac{dU_t^{(N_{n(\infty)},\Delta x)}}{dt}$ be the limit functions of t in the sense as

$$
\begin{aligned}
\lim_{j \to \infty} \sup_{t \in [0,T]} \| U_t^{(N_{n(\infty)},\Delta x)} - U_t^{(N_{n(j)},\Delta x)} \|_\infty &= 0, \\
\lim_{j \to \infty} \sup_{t \in [0,T]} \left\| \frac{dU_t^{(N_{n(\infty)},\Delta x)}}{dt} - \frac{dU_t^{(N_{n(j)},\Delta x)}}{dt} \right\|_\infty &= 0.
\end{aligned}
$$

Using the triangle inequality,

$$\begin{aligned}
&\left|\frac{de_i^* U_t^{(N_{n(\infty)},\Delta x)}}{dt} - \frac{de_i^* U_t^{(N_{n(j)},\Delta x)}}{dt}\right| \\
&= \left|\frac{de_i^* U_t^{(N_{n(\infty)},\Delta x)}}{dt} - \breve{f}(t, i\Delta x, e_i^* U_t^{(N_{n(j)})}, e_i^* D_1 U_t^{(N_{n(j)})}, e_i^* D_2 U_t^{(N_{n(j)})})\right| \\
&\geq -c_8 \sup_{t\in[0,T]} \|U_t^{(N_{n(\infty)})} - U_t^{(N_{n(j)})}\|_\infty \\
&\quad + \left|\frac{de_i^* U_t^{(N_{n(\infty)},\Delta x)}}{dt} - \breve{f}(t, i\Delta x, e_i^* U_t^{(N_{n(\infty)})}, e_i^* D_1 U_t^{(N_{n(\infty)})}, e_i^* D_2 U_t^{(N_{n(\infty)})})\right|.
\end{aligned}$$

Considering $j \to \infty$ yields

$$\frac{de_i^* U_t^{(N_{n(\infty)},\Delta x)}}{dt} = \breve{f}(t, i\Delta x, e_i^* U_t^{(N_{n(\infty)})}, \sigma(i\Delta x) e_i^* D_1 U_t^{(N_{n(\infty)})}, e_i^* D_2 U_t^{(N_{n(\infty)},\Delta x)})$$

for $t \in [0,T]$ and $i \in \mathbb{Z}$, which implies that $U_t^{(N_{n(\infty)},\Delta x)}$ solves (2.73). Using a uniqueness result in Lemma 2.8.1, we obtain $U^{(\infty,\Delta x)} = U^{(N_{n(\infty)},\Delta x)}$. As $(N_j)_{j=1}^\infty$ is arbitrary, we obtain the desired result. □

Theorem 2.8.1. *Suppose that Assumption 2.8.1, 2.8.2 and* (2.71) *admits a unique solution u. For any compact set $K \subset \mathbb{R}$, it holds*

$$\lim_{\Delta x\to 0} \lim_{N\to\infty} \sup_{\substack{-N_0\leq i\leq N_0 \\ i\in\mathbb{Z}, i\Delta x\in K}} |u(t, i\Delta x) - e_i^* U_t^{(N,\Delta x)}| = 0.$$

Proof. The argument follows from Lemma 2.8.2, 2.8.3 and the triangle inequality;

$$\begin{aligned}
\sup_{\substack{i\in\mathbb{Z} \\ i\Delta x\in K}} |u(t, i\Delta x) - e_i^* U_t^{(N,\Delta x)}| &\leq \sup_{\substack{i\in\mathbb{Z} \\ i\Delta x\in K}} |u(t, i\Delta x) - e_i^* U_t^{(\infty,\Delta x)}| + \|U_t^{(\infty,\Delta x)} - U_t^{(N,\Delta x)}\|_\infty \\
&\to \sup_{\substack{i\in\mathbb{Z} \\ i\Delta x\in K}} |u(t, i\Delta x) - e_i^* U_t^{(\infty,\Delta x)}|, \quad N \to \infty. \\
&\to 0, \quad \Delta x \to 0.
\end{aligned}$$

□

Chapter 3

Multi-Stage Euler-Maruyama Methods for Backward Stochastic Differential Equations Driven by Continuous-Time Markov Chains with Bounded Stopping Terminal Times

3.1 Introduction

3.1.1 Overview

In the previous chapter, we studied numerical methods for solving Markov BSDEs driven by time-homogeneous CTMC $(X_t)_{t\in[0,T]}$ in what follows:

$$\begin{aligned} X_t &= x_0 + \int_{]0,t]} Q^* X_{s-} ds + M_t, \\ Y_s &= X_T^* \boldsymbol{\zeta} + \int_{]s,T]} H(X_{r-}, r, \mathcal{Y}_{r-}, Z_r) dr - \int_{]t,T]} dM_r^* Z_r. \end{aligned} \tag{3.1}$$

We focused on that the solution satisfies $Y_t = X_t^* V_t$ up to indistinguishability and $Z \sim_M V$ for a deterministic process V_t that solves an associated system of ODEs and constructed multi-stage Euler-Maruyama methods for solving (3.1) in Section 2.3;

interestingly, we observed that the methods are equivalent to exponential integrators for solving the systems of ODEs associated with (3.1). Noticing that exponential integrators are known as solvers that mitigate numerical instabilities referred to as the "stiffness", Section 2.4 proposed to use these methods to solve (3.1) that comes from a spatial discretization of Markov BSDEs driven by Brownian motions; such (CTMC-)BSDEs typically associate with stiff systems of ODEs.

In this chapter, we extend these results into a stopping terminal time counterpart of (3.1), that is,

$$\begin{aligned} X_t &= x_0 + \int_{]0,t]} Q_s^* X_{s-} ds + M_t, \\ Y_s &= X^*_{T\wedge\tau}\boldsymbol{\zeta}_{T\wedge\tau} + \int_{]s,T]} 1_{\{r\leq\tau\}} H(X_{r-}, r, \mathcal{Y}_{r-}, Z_r) dr - \int_{]t,T]} dM_r^* Z_r. \end{aligned} \tag{3.2}$$

Here, τ is a stopping time defined by $\tau \coloneqq \inf\{t \geq 0 : X_t \notin \Xi\}$ for some subset $\Xi \subset \mathcal{I}$, and ζ_t is deterministic. As seen later, similarly to the case of the terminal times being deterministic, the solution of (3.2) can be expressed using the solution V_t of an associated system of ODEs as $Y_t = X^*_{t\wedge\tau} V_{t\wedge\tau}$ and $Z_t \sim_M V_t 1_{\{t\leq\tau\}}$, and the multi-stage Euler-Maruyama methods results in the calculation of V_t using exponential integrators. In the same way of Section 2.4, we also can utilize the methods to solve stiff BSDEs driven by CTMCs arising from the spatial discretization of BSDEs driven by Brownian motion with bounded stopping terminal times.

3.1.2 Motivation : BSDEs with Bounded Stopping Terminal Times

All of the BSDEs we treated in Chapter 2 have deterministic terminal time, denoted as $T > 0$, whereas we can also consider counterparts of BSDEs with non-deterministic terminal times. Among them, BSDEs with bounded stopping terminal times can be solved as BSDEs with deterministic terminal times. In Brownian case, for example, they are written as follows:

$$\mathcal{Y}_t = \xi + \int_t^T 1_{\{s\leq\tau\}} f(s, \mathcal{Y}_s, \mathcal{Z}_s) ds - \int_t^T \mathcal{Z}_s^* dW_s, \quad t \in [0, T], \tag{3.3}$$

where τ is a $\{\mathcal{F}_t\}_{t\in[0,T]}$-stopping time, $\{\mathcal{F}_t\}_{t\in[0,T]}$ is the completion of the filtration generated by Brownian motion $(W_t)_{t\in[0,T]}$, ξ is $\mathcal{F}_{T\wedge\tau}$-measurable and all other conditions remain the same as the determinisitc case. As the solution of (3.3) satisfies $\mathcal{Y}_t = \mathcal{Y}_{t\wedge\tau}$ and $Z_t = Z_t 1_{\{t\leq\tau\}}$, it can be regarded as BSDEs whose terminal time is a bounded stopping time $T \wedge \tau (= \min\{T, \tau\})$.

Let us now motivate (3.3) with an example in mathematical finance. Consider a European knock-out option, a barrier option. It is paid if the underlying asset price $\mathcal{S}_t = (\mathcal{S}_t^{(1)}, \dots, \mathcal{S}_t^{(d)})$ has stayed in a prescribed open domain G. Once $\mathcal{S}_t$ hits ∂G, the option becomes worthless and the rebate described by a function $\phi_0(t,x)$ is paid. We can represent the payoff of the option as random variable $\phi(T \wedge \tau, \mathcal{S}_{T\wedge\tau})$ where

$$\phi(t,s) \coloneqq \begin{cases} \phi_0(t,s), & (t,s) \in [0,T] \times \partial G, \\ (s-K)^+, & (t,s) \in \{T\} \times G, \end{cases}$$

and $\tau = \inf\{t \geq 0 : \mathcal{S}_t \in \partial G\}$. We are now interested in evaluating the fair price of the option at time 0 under different interest rates for borrowing and lending. That is accomplished by considering its hedge portfolio. Suppose that $\mathcal{S}_t$ satisfies

$$\mathcal{S}_t = s_0 + \int_0^t \mu(\mathcal{S}_s)ds + \int_0^t \sigma(\mathcal{S}_s)dW_s,$$

for some $s_0 > 0$, $\mu : \mathbb{R}^d \to \mathbb{R}^d$, and $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ being invertible, and consider a portfolio $\boldsymbol{\Delta}_t = (\Delta_t^{(1)}, \dots, \Delta_t^{(d)})^*$ consisting $\Delta_t^{(i)}$ assets of $\mathcal{S}_t^{(i)}$ and bonds with borrowing rate R and lending rate r. Let $\mathcal{V}_t$ be the wealth process of $\boldsymbol{\Delta}_t$. The hedge portfolio satisfies $\mathcal{V}_T = \phi(T \wedge \tau, \mathcal{S}_{T\wedge\tau})$ and the self-financing condition that reads

$$d\mathcal{V}_t = 1_{\{t\leq\tau\}}[r(\mathcal{V}_t - \boldsymbol{\Delta}_t^*\mathcal{S}_t)^+ - R(\mathcal{V}_t - \boldsymbol{\Delta}_t^*\mathcal{S}_t)^-]dt + 1_{\{t\leq\tau\}}\Delta_t^* d\mathcal{S}_t.$$

Here, $(x)^+ = \max\{x, 0\}$ and $(x)^- = \max\{-x, 0\}$ for $x \in \mathbb{R}$. As a result, writing $\mathcal{Y}_t \coloneqq \mathcal{V}_t$ and $\mathcal{Z}_t \coloneqq \Delta_t^*\sigma(\mathcal{S}_t)1_{\{s\leq\tau\}}$, we obtain the (Markov) BSDE in what follows:

$$\mathcal{Y}_t = \phi(T \wedge \tau, \mathcal{S}_{T\wedge\tau}) - \int_t^T 1_{\{s\leq\tau\}} f(s, \mathcal{S}_s, \mathcal{Y}_s, \mathcal{Z}_s)ds - \int_t^T \mathcal{Z}_s^* dW_s \tag{3.4}$$

where

$$f(t,x,y,z) \coloneqq r(y - z^*\sigma(x)^{-1}x)^+ - R(y - z^*\sigma(x)^{-1}x)^- + z^*\sigma(x)^{-1}\mu(x).$$

The solution $\mathcal{Y}_0$ is then the fair price of the option at time 0.

3.1.3 Organization

This chapter is organized as follows: The end of this section is devoted to notations. In Section 3.2, we introduce BSDEs with terminal times being stopping times driven by CTMCs / Brownian motion. Section 3.3 is devoted to constructing multi-stage Euler-Maruyama methods for solving BSDEs with stopping terminal times (3.1). We present an application of these methods to BSDEs driven by Brownian motion in Section 2.4, and its efficiency is confirmed by numerical results in Section 2.5; we there focus on pricing European barrier options under nonlinear wealth dynamics.

3.1.4 Notations

Throughout this chapter, the same notations as Chapter 2 are used.

3.2 Preliminary Results

3.2.1 BSDEs Driven by CTMCs with Bounded Stopping Terminal Times

Let $X = (X_t)_{t\in[0,T]}$ be a continuous-time, finite-state Markov chain with state space $\mathcal{I} = \{e_1, \dots, e_N\}$, for some $N \in \mathbb{N}$. Suppose that X is defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{G})$ where $\mathbb{G} \coloneqq (\mathcal{G}_t)_{t\in[0,T]}$ is the completion of the filtration generated by X. Denote Q_t as the Q-matrix of X. Then, X has the following semi-martingale representation

$$X_t = x_0 + \int_{]0,t]} Q_s^* X_{s-} ds + M_t.$$

Here, $x_0 \in \{e_1, \dots, e_N\}$ and M_t is an $\mathbb{R}^N$-valued $\mathbb{G}$-martingale.

Generally, the BSDE driven by CTMC X with deterministic terminal time $T > 0$ is written as

$$Y_t = \xi + \int_{]t,T]} h(s, Y_{s-}, Z_s) ds - \int_{]t,T]} dM_s^* Z_s, \tag{3.5}$$

where ξ is an $\mathcal{G}_T$-measurable square-integrable random variable, $h : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^N \ni (\omega, t, y, z) \mapsto h(\omega, t, y, z) \in \mathbb{R}$ is $\mathbb{G}$-predictable in (ω, t) and Borel measurable in (y, z). Recall that the following result on the existence and a uniqueness of the solution $(Y, Z) \in \mathbb{S}^2(\mathbb{G}, \mathbb{R}) \times L^2(\mathbb{G}, \langle M \rangle, \mathbb{R}^N)$ of (3.5).

Theorem 3.2.1 ([15]). *Assume that,*

$$\mathbb{E}\left[\int_{]0,T]} h(t, 0, 0)^2 dt\right] < \infty,$$

and that for some constant $L > 0$,

$$|h(t, y, z) - h(t, y', z')|^2 \le L(|y - y'|^2 + \|z - z'\|^2_{X_{t-}}), \quad dt \otimes d\mathbb{P}\text{-a.s.}$$

for all $y, y' \in \mathbb{R}$ and $z, z' \in \mathbb{R}^N$. Then, it admits a unique solution $(Y, Z) \in \mathbb{S}^2(\mathbb{G}, \mathbb{R}) \times L^2(\mathbb{G}, \langle M \rangle, \mathbb{R}^N)$. We remark that it is unique up to indistinguishability for Y and up to $\sim_M$ equivalence for Z.

The main target of this chapter is the following.

$$Y_t = \xi + \int_{]t,T]} 1_{\{s\leq\tau\}} h(s, Y_{s-}, Z_s)ds - \int_{]t,T]} dM_s^* Z_s, \tag{3.6}$$

where $T > 0$ is a determinisitic time horizon, τ is a $\mathbb{G}$-stopping time, ξ is a $\mathcal{G}_{T\wedge\tau}$-measurable square-integrable random variable and $1_{\{s\leq\tau\}}$ is an indicator function. Note that, taking f as above, Theorem 3.2.1 directly implies the existence and uniqueness of the solution of (3.6). Moreover, the following proposition holds.

Proposition 3.2.1. *The solution* (Y, Z) *of* (3.6) *satisfies* $Y_t = Y_{t\wedge\tau}$ *for* $t \in [0, T]$ *and* $Z \sim_M Z1_{]0,T\wedge\tau]}$.

Proof. Since $Y_{T\wedge\tau} = \xi - \int_{]T\wedge\tau,T]} dM_s^* Z_s$, we obtain $Y_{T\wedge\tau} = \mathbb{E}[Y_{T\wedge\tau}|\mathcal{G}_{T\wedge\tau}] = \xi$. Note that for any $t \in [0, T]$,

$$Y_{t\wedge\tau}1_{\{t\wedge\tau<t\}} = Y_\tau 1_{\{t\wedge\tau<t\}} = Y_{T\wedge\tau}1_{\{t\wedge\tau<t\}} = \xi 1_{\{t\wedge\tau<t\}},$$

and

$$\begin{aligned}
Y_t 1_{\{t\wedge\tau<t\}} &= \mathbb{E}[Y_t 1_{\{t\wedge\tau<t\}}|\mathcal{G}_t] \\
&= \mathbb{E}\left[\left(\xi + \int_{]t,T]} 1_{\{s\leq\tau\}} h(s, Y_{s-}, Z_s)ds - \int_{]t,T]} dM_s^* Z_s\right) 1_{\{t\wedge\tau<t\}} \middle| \mathcal{G}_t\right] \\
&= \mathbb{E}\left[Y_{t\wedge\tau}1_{\{t\wedge\tau<t\}} - \int_{]t,T]} dM_s^* Z_s 1_{\{t\wedge\tau<t\}} \middle| \mathcal{G}_t\right] \\
&= Y_{t\wedge\tau}1_{\{t\wedge\tau<t\}} - \mathbb{E}\left[\int_{]t,T]} dM_s^* Z_s \middle| \mathcal{G}_t\right] 1_{\{t\wedge\tau<t\}} = Y_{t\wedge\tau}1_{\{t\wedge\tau<t\}}.
\end{aligned}$$

$Y_t = Y_{t\wedge\tau}$ follows from

$$Y_t - Y_{t\wedge\tau} = (Y_t - Y_{t\wedge\tau})1_{\{t\wedge\tau<t\}} + (Y_t - Y_t)1_{\{t\wedge\tau=t\}} = (Y_t - Y_{t\wedge\tau})1_{\{t\wedge\tau<t\}} = 0$$

Squaring $Y_{T\wedge\tau}$ leads to

$$\xi^2 = Y_{T\wedge\tau}^2 = \xi^2 - 2\xi \int_{]T\wedge\tau,T]} dM_s^* Z_s + \left|\int_{]T\wedge\tau,T]} dM_s^* Z_s\right|^2.$$

Taking conditional expectation $\mathbb{E}[\cdot|\mathcal{G}_{T\wedge\tau}]$, we obtain $\mathbb{E}\left[\left|\int_{]T\wedge\tau,T]} dM_s^* Z_s\right|^2 \middle| \mathcal{G}_{T\wedge\tau}\right] = 0$. Considering the predictable quadratic variation of $\int_{]T\wedge\tau,T]} dM_s^* Z_s$,

$$0 = \mathbb{E}\left[\int_{]T\wedge\tau,T]} \|Z_s\|_{X_{s-}}^2 ds \middle| \mathcal{G}_{T\wedge\tau}\right] = \mathbb{E}\left[\int_{]0,T]} \|Z_s 1_{]T\wedge\tau,T]}\|_{X_{s-}}^2 ds \middle| \mathcal{G}_{T\wedge\tau}\right],$$

which implies $\mathcal{Z}1_{]T\wedge\tau,T]} \sim_M 0$. □

In this chapter, we mainly focus on the Markov case. Let $\Xi \subset \mathcal{I}$ and $\tau = \inf\{s \geq 0 : X_s \notin \Xi\}$ with the usual convention that $\inf \emptyset = \infty$. Let $\boldsymbol{\zeta} : [0,T] \to \mathbb{R}^N$ and $h : \{e_1, \ldots, e_N\} \times [0,T] \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ be Borel measurable functions. Then, we consider the following Markov BSDE.

$$Y_t = X^*_{T\wedge\tau}\boldsymbol{\zeta}_{T\wedge\tau} + \int_{]t,T]} 1_{\{r\leq\tau\}} h(X_{r-}, r, Y_{r-}, Z_r)dr - \int_{]t,T]} dM^*_r Z_r$$

Associated with it, setting $t \in [0,T]$ as the starting time for the BSDE, we consider

$$\begin{cases} X^{t,e_i}_s = e_i + \displaystyle\int_{]t,s]} Q^*_u X^{t,e_i}_{u-} du + M_s - M_t, \quad s > t, \\ X^{t,e_i}_s = e_i, \quad s \leq t, \\ Y^{t,e_i}_s = (X^{t,e_i}_{T\wedge\tau_{t,e_i}})^* \boldsymbol{\zeta}_{T\wedge\tau_{t,e_i}} + \displaystyle\int_{]s,T]} 1_{\{r\leq\tau_{t,e_i}\}} h(X^{t,e_i}_{r-}, r, Y^{t,e_i}_{r-}, Z^{t,e_i}_r)dr - \int_{]s,T]} dM^*_r Z^{t,e_i}_r, \end{cases} \tag{3.7}$$

for $s \in [0,T]$. Then, we give the following nonlinear Feynman-Kac type result.

Theorem 3.2.2. *Assume that there exists a constant $L > 0$ such that*

$$|h(e_i,t,y,z) - h(e_i,t,y',z')|^2 \leq L^2(|y-y'|^2 + \|z-z'\|^2_{e_i}), \tag{3.8}$$

for any $y, y' \in \mathbb{R}$, $z, z' \in \mathbb{R}^N$, $t \in [0,T]$ and $i \in \Xi$, and $\mathbb{E}\int_0^T h(X^{t,e_i}_{u-}, u, 0, 0)^2 du < \infty$ for $i \in \Xi$. Define $H : [0,T] \times \mathbb{R}^N \to \mathbb{R}^N$ such that

$$e_i^* H(t,z) = h(e_i, t, e_i^* z, z) \quad \textit{for} \quad t \in [0,T], \quad z \in \mathbb{R}^N, \quad i \in \Xi.$$

1. *For a solution U_t of the system of ODEs*

$$\begin{cases} e_i^* \left(\dfrac{dU_t}{dt} + Q_t U_t + H(t, U_t) \right) = 0, & (t, e_i) \in [0,T) \times \Xi, \\ e_i^* U_t = e_i^* \boldsymbol{\zeta}_t, & (t,i) \in (\{T\} \times \{e_1, \ldots, e_N\}) \cup ([0,T] \times \Xi^c), \end{cases} \tag{3.9}$$

   $(Y^{t,e_i}_s, Z^{t,e_i}_s) = ((X^{t,e_i}_{s\wedge\tau_{t,e_i}})^* U_{s\wedge\tau_{t,e_i}}, U_s 1_{\{s\leq\tau_{t,x}\}}) \in \mathbb{S}^2(\mathbb{G},\mathbb{R}) \times L^2(\mathbb{G}, \langle M\rangle, \mathbb{R}^N)$ is the unique solution to the BSDE (3.7).

2. *Conversely, for a unique solution $(Y^{t,e_i}_s, Z^{t,e_i}_s) \in \mathbb{S}^2(\mathbb{G},\mathbb{R}) \times L^2(\mathbb{G}, \langle M\rangle, \mathbb{R}^N)$ of* (3.7)*, a continuous function $V_t = (Y^{t,e_1}_t, \ldots, Y^{t,e_N}_t)^*$ satisfies $V_s 1_{\{s\leq\tau_{t,e_i}\}} \sim_M Z^{t,e_i}_s$ for $i = 1, \ldots, N$ and $t \in [0,T]$, and solves* (3.9).

Corollary 3.2.1. *Under the square integrability $t \mapsto h(e_i, t, y, z)$ in $[0,T]$ and the uniform Lipschitz continuity* (3.8)*, a unique solution (Y^{t,e_i}, Z^{t,e_i}) of* (3.7) *is also a unique solution of*

$$Y_s^{t,e_i} = X^*_{T\wedge\tau_{t,e_i}}\boldsymbol{\zeta}_{T\wedge\tau_{t,e_i}} + \int_{]s,T]} (X^{t,e_i}_{r-})^* H(r, Z_r^{t,e_i}) 1_{\{r\le\tau_{t,e_i}\}} dr - \int_{]t,T]} dM_r^* Z_r^{t,e_i}. \quad (3.10)$$

Moreover, the relation

$$Y_s^{t,e_i} = (X^{t,e_i}_{s\wedge\tau_{t,e_i}})^* U_{t\wedge\tau_{t,e_i}} \quad \textit{up to indistinguishability and} \quad Z_t^{t,e_i} \sim_M U_t 1_{\{t\le\tau_{t,e_i}\}}$$

holds, where U is a solution of (3.9)*.*

3.2.2 BSDEs Driven by Brownian motion with Bounded Stopping Terminal Times

Let us introduce the Brownian case of BSDEs with bounded stopping terminal times. Let $T > 0$ be a (fixed, deterministic) time horizon, $W = (W_t)_{t\in[0,T]}$ be a d-dimensional standard Brownian motion, and $\mathbb{F} = (\mathcal{F}_t)_{t\ge0}$ be the completion of the filtration generated by W. Recall that a result on the uniqueness and existence of BSDE

$$\mathcal{Y}_t = \xi + \int_t^T f(s, \mathcal{Y}_s, \mathcal{Z}_s)ds - \int_t^T \mathcal{Z}_s^* dW_s, \quad t \in [0,T], \quad (3.11)$$

is given in what follows:

Theorem 3.2.3. *BSDE* (3.11) *admits a unique solution $(\mathcal{Y}, \mathcal{Z}) \in \mathbb{S}^2(\mathbb{F}, \mathbb{R}) \times \mathbb{L}^2(\langle\mathcal{W}\rangle, \mathbb{F}, \mathbb{R}^d)$ under the following conditions:*

- *ξ is $\mathbb{R}$-valued $\mathcal{F}_T$-measurable random variable and satisfies $\mathbb{E}[\xi^2] < \infty$.*
- *$f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is a $\mathbb{F}$-measurable function that satisfies (i) there exists $K > 0$ such that for all $y, y' \in \mathbb{R}$ and $z, z' \in \mathbb{R}^d$,*

  $$\|f(\omega, t, y, z) - f(\omega, t, y', z')\|^2 \le K(|y - y'|^2 + \|z - z'\|^2) \quad d\mathbb{P} \otimes dt\text{-a.s.},$$

  and (ii) $\mathbb{E}[\int_0^T |f(t, 0, 0)|^2 dt] < \infty$.

Next, we consider the following BSDE.

$$\mathcal{Y}_t = \xi + \int_t^T 1_{\{s\le\tau\}} f(s, \mathcal{Y}_s, \mathcal{Z}_s)ds - \int_t^T \mathcal{Z}_s^* dW_s, \quad t \in [0,T]. \quad (3.12)$$

Here, τ is a $\mathbb{F}$-stopping time, ξ is $\mathcal{F}_{T\wedge\tau}$-measurable, and $1_{\{s\leq\tau\}}$ is an indicator function. If we take f satisfying the conditions above and ξ such that $\mathbb{E}[\xi^2]$, we can apply Theorem 3.2.3 and see that BSDE (3.12) admits a unique solution. Moreover, we note that the following proposition.

Proposition 3.2.2 (Proposition 2.6 in [58]). *A solution* $(\mathcal{Y},\mathcal{Z})$ *of* (3.12) *satisfies* $\mathcal{Y}_t=\mathcal{Y}_{t\wedge\tau}$ *for* $t\in[0,T]$ *and* $\mathcal{Z}=0$ $dt\otimes\mathbb{P}(d\omega)$*-a.e. in* $]T\wedge\tau,T]$.

Next, we introduce a "Markov" counterpart of (3.12). To this end, let $G\subset\mathbb{R}^d$ is a connected open set whose boundary is of class C^1, and define a stopping time

$$\tau=\inf\{s\geq 0:\mathcal{X}_s\notin\overline{G}\}.$$

Then, we can formulate Markov BSDEs in what follows:

$$\begin{cases}\mathcal{X}_t=x_0+\displaystyle\int_0^t\mu(s,\mathcal{X}_s)ds+\int_0^t\sigma(s,\mathcal{X}_s)dW_s,\\ \mathcal{Y}_{t\wedge\tau}=\chi(T\wedge\tau,\mathcal{X}_{T\wedge\tau})+\displaystyle\int_t^T 1_{\{s<\tau\}}f(s,\mathcal{X}_s,\mathcal{Y}_s,\mathcal{Z}_s)ds-\int_t^T(\mathcal{Z}_s)^*dW_s.\end{cases}\tag{3.13}$$

where $\mu:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$, $\sigma:[0,T]\times\mathbb{R}^d\to\mathbb{R}^{d\times d}$, $f:[0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$, $\chi:[0,T]\times\mathbb{R}^d\to\mathbb{R}$ are Borel measurable. Assuming that, there exists $L>0$ and $p\in\mathbb{N}$ such that

$$\begin{gathered}\|\mu(t,x)-\mu(t,x')\|+\|\sigma(t,x)-\sigma(t,x')\|\leq L\|x-x'\|,\\ |f(t,x,y,z)-f(t,x,y',z')|\leq L(|y-y'|+\|z-z'\|),\\ \|\mu(t,x)\|+\|\sigma(t,x)\|\leq L(1+\|x\|^2),\\ |f(t,x,y,z)|+|\chi(t,x)|\leq L(1+\|x\|^p),\end{gathered}\tag{3.14}$$

for all $t\in[0,T]$, $x,x',z,z'\in\mathbb{R}^d$ and $y,y'\in\mathbb{R}$, (3.13) has a unique solution $((\mathcal{X}_t)_{t\in[0,T]},(\mathcal{Y}_t)_{t\in[0,T]},(\mathcal{Z}_t)_{t\in[0,T]})\in\mathbb{S}^2(\mathbb{F},\mathbb{R}^d)\times\mathbb{S}^2(\mathbb{F},\mathbb{R})\times L^2(\langle W\rangle,\mathbb{F},\mathbb{R}^d)$. It is worth noting that Markov BSDE (3.13) is related to a Cauchy-Dirichlet problem of second order parabolic partial differential equations in what follows:

$$\begin{cases}\frac{\partial u}{\partial t}(t,x)+\mathcal{L}_t u(t,x)+f(t,x,u(t,x),\sigma^*(x)\nabla_x u(t,x))=0,\quad (t,x)\in[0,T]\times G,\\ u(T,x)=\kappa(x),\quad x\in\overline{G},\\ u(t,x)=\chi(t,x),\quad (t,x)\in[0,T]\times\partial G.\end{cases}\tag{3.15}$$

Here, $\kappa(x) \coloneqq \chi(T,x)$ for $x \in \overline{G}$,

$$\mathcal{L}_t u(t,x) = \sum_{i=1}^{d} \mu^{(i)}(t,x) \frac{\partial u}{\partial x_i}(t,x) + \frac{1}{2} \sum_{i,j=1}^{d} (\sigma\sigma^*)^{(i,j)}(t,x) \frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) \tag{3.16}$$

is the infinitesimal generator of the Markov process $\mathcal{X}$,

$$\nabla_x u(t,x) = \left(\frac{\partial u}{\partial x_1}(t,x), \ldots, \frac{\partial u}{\partial x_d}(t,x) \right)^* \in \mathbb{R}^d$$

is the gradient vector, $\mu^{(i)}(t,x)$ is the i-th component of $\mu(t,x)$, and $(\sigma\sigma^*)^{(i,j)}(t,x)$ is the (i,j)-th component of $\sigma(t,x)\sigma^*(t,x)$. The precise statement is as follows.

Theorem 3.2.4 (The nonlinear Feynman-Kac formula (pp. 421–422 in [60].))**.** *Denote $\tau_{t,x} = \inf\{s \geq t : \mathcal{X}_s^{t,x} \notin \overline{G}\}$ for $(t,x) \in [0,T] \times \overline{G}$. Suppose that $\mu, \sigma, f, \chi, \kappa, G$ are defined as above, and also that G is bounded, $\chi \in C([0,T] \times \partial G)$, and*

$$\Lambda = \{(t,x) \in [0,T] \times \partial G : \mathbb{P}(\tau_{t,x} > t) = 0\} \tag{3.17}$$

is a closed set. For $(t,x) \in [0,T] \times \overline{G}$, let $(\mathcal{X}^{t,x}, \mathcal{Y}^{t,x}, \mathcal{Z}^{t,x})$ be a unique solution of the Markov BSDE

$$\begin{cases} \mathcal{X}_s^{t,x} = x + \displaystyle\int_t^s \mu(r, \mathcal{X}_r^{t,x}) dr + \int_t^s \sigma(r, \mathcal{X}_r^{t,x}) dW_r \quad \textit{for} \quad s \geq t, \\ \mathcal{X}_s^{t,x} = x \in \mathbb{R}^d \quad \textit{for} \quad s \leq t, \\ \mathcal{Y}_{s \wedge \tau_{t,x}}^{t,x} = \chi(T \wedge \tau_{t,x}, \mathcal{X}_{T \wedge \tau_{t,x}}^{t,x}) + \displaystyle\int_s^T 1_{\{r < \tau_{t,x}\}} f(r, \mathcal{X}_r^{t,x}, \mathcal{Y}_r^{t,x}, \mathcal{Z}_r^{t,x}) dr \\ \qquad\qquad\qquad\qquad - \displaystyle\int_s^T (\mathcal{Z}_r^{t,x})^* dW_r \quad \textit{for} \quad s \in [0,T]. \end{cases} \tag{3.18}$$

Then,

1. *for every classical solution $u \in C^{1,2}([0,T] \times G; \mathbb{R}) \cap C([0,T] \times \overline{G}; \mathbb{R})$ of (3.15), the unique solution of BSDE (3.18) is represented as*

$$\mathcal{Y}_s^{t,x} = u(s \wedge \tau_{t,x}, \mathcal{X}_{s \wedge \tau_{t,x}}^{t,x}), \quad \mathcal{Z}_s^{t,x} = \sigma^*(\mathcal{X}_s^{t,x}) \nabla_x u(s, \mathcal{X}_s^{t,x}) 1_{\{s < \tau_{t,x}\}} \quad \textit{for} \quad s \geq t. \tag{3.19}$$

(The above inequality is required for $(\mathcal{Y}_s^{t,x}, \mathcal{Z}_s^{t,x})$ to be of the class $\mathbb{S}^2(\mathbb{F}, \mathbb{R}) \times L^2(\langle W \rangle, \mathbb{F}, \mathbb{R}^d)$.)

2. *Suppose further that f and g are uniformly continuous with respect to x. Then, $v : [0,T] \times \overline{G} \to \mathbb{R}$ defined by $v(t,x) \coloneqq \mathcal{Y}_t^{t,x}$ is a viscosity solution of* (3.15).

3. *Additionally, if for each $R > 0$ there exists a continuous function $m_R : [0,\infty) \to [0,\infty)$ such that $m_R(0) = 0$, and*

$$|f(t,x,y,z) - f(t,x',y,z)| \leq m_R(|x-x'|(1+\|z\|))$$

holds for $x, x' \in \overline{G}, y \in \mathbb{R}, z \in \mathbb{R}^d$ such that $\max\{\|x\|, \|x'\|, \|z\|\} < R$, then the uniqueness of v also holds.

3.3 Multi-Stage Euler-Maruyama Methods

We are interested in multi-stage Euler-Maruyama methods for (3.10). In the deterministic terminal time case, the solution (Y^{t,e_i}, Z^{t,e_i}) satisfies $Y_s^{t,e_i} = (X_s^{t,e_i})^* V_t$ and $Z^{t,e_i} \sim_M V_t$ for $V_t = (Y_t^{t,e_1}, \ldots, Y_t^{t,e_N})^*$, and the Euler-Maruyama methods result in exponential integrators for solving ODE system (3.9) which V_t satisfies. In the stopping terminal time case, on the other hand, the solution (Y^{t,e_i}, Z^{t,e_i}) satisfies $Y_s^{t,e_i} = (X_{s\wedge\tau_{t,e_i}}^{t,e_i})^* V_{s\wedge\tau_{t,e_i}}$ and $Z^{t,e_i} \sim_M V 1_{]0,T\wedge\tau]}$. Here, $V_t = (Y_t^{t,e_1}, \ldots, Y_t^{t,e_N})^*$ solves ODE system (3.9). As seen later, the Euler-Maruyama methods in this case are accomplished by using exponential integrators for solving (3.9).

For $M \in \mathbb{N}$, consider a uniform time partition $t_m = m\Delta t$ $(m = 0, 1, \ldots, M)$ of interval $[0,T]$, where $\Delta t = T/M$ is the step size. For later use, denote $I_\Xi = \mathrm{diag}(1_\Xi(e_1), \ldots, 1_\Xi(e_N))$ and $I_{\Xi^c} = I_N - I_\Xi$. Let (Y^{t,e_i}, Z^{t,e_i}) be the solution of (3.10) and $V_t = (Y_t^{t,e_1}, \ldots, Y_t^{t,e_N})$. Suppose that V_t uniquely solves ODE (3.9). $Y_{t_m}^{t_m,e_i}$ satisfies the following backward stochastic difference equation.

$$Y_{t_m}^{t_m,e_i} = Y_{t_{m+1}\wedge\tau_{t_m,e_i}}^{t_m,e_i} + \int_{]t_m,t_{m+1}]} (X_{s-}^{t_m,e_i})^* H(s, Z_s^{t_m,e_i}) 1_{\{s\leq\tau_{t_m,e_i}\}} ds - \int_{]t_m,t_{m+1}]} dM_s^* Z_s^{t_m,e_i}.$$

Plugging $(Y_t^{t_m,e_i}, Z_t^{t_m,e_i}) = ((X_{t\wedge\tau_{t_m,e_i}}^{t_m,e_i})^* V_{t\wedge\tau_{t_m,e_i}}, V_t 1_{\{t\leq\tau_{t_m,e_i}\}})$,

$$\begin{aligned}
Y_{t_m}^{t_m,e_i} &= \mathbb{E}[Y_{t_{m+1}}^{t_m,e_i} | X_{t_m}^{t_m,e_i}] + \int_{t_m}^{t_{m+1}} \mathbb{E}[1_{\{r\leq\tau_{t_m,e_i}\}} (X_{r-}^{t_m,e_i})^* H(r, Z_r^{t_m,e_i}) | X_{t_m}^{t_m,e_i}] dr \\
&= \mathbb{E}[(X_{t_{m+1}\wedge\tau_{t_m,e_i}}^{t_m,e_i})^* V_{t_{m+1}\wedge\tau_{t_m,e_i}} | X_{t_m}^{t_m,e_i}] \\
&\quad + \int_{t_m}^{t_{m+1}} \mathbb{E}[1_{\{r\leq\tau_{t_m,e_i}\}} (X_r^{t_m,e_i})^* H(r, V_r 1_{\{r\leq\tau_{t_m,e_i}\}}) | X_{t_m}^{t_m,e_i}] dr.
\end{aligned}$$

Since

$$1_{\{r\le\tau_{t_m,e_i}\}}\cdot(X_r^{t_m,e_i})^*H(r,V_r1_{\{r\le\tau_{t_m,e_i}\}}) \\ =1_{\{r\le\tau\}}\cdot(X_r^{t_m,e_i})^*H(r,V_r)=X^*_{r\wedge\tau_{t_m,e_i}}I_{\Xi^c}H(r,V_r),$$

we obtain

$$Y_{t_m}^{t_m,e_i}=\mathbb{E}[(X^{t_m,e_i}_{t_{m+1}\wedge\tau_{t_m,e_i}})^*V_{t_{m+1}\wedge\tau_{t_m,e_i}}|X_{t_m}^{t_m,e_i}]+\int_{t_m}^{t_{m+1}}\mathbb{E}[X^{t_m,e_i}_{r\wedge\tau_{t_m,e_i}}|X_{t_m}^{t_m,e_i}]^*I_{\Xi^c}H(r,V_r)dr \tag{3.20}$$

for $e_i\in\Xi^c$; otherwise $Y_{t_m}^{t_m,e_i}=e_i^*\boldsymbol{\zeta}_{t_m}$. We approximate

$$\mathbb{E}[(X^{t_m,e_i}_{t_{m+1}\wedge\tau_{t_m,e_i}})^*V_{t_{m+1}\wedge\tau_{t_m,e_i}}|X_{t_m}^{t_m,e_i}]\approx\mathbb{E}[X^{t_m,e_i}_{t_{m+1}\wedge\tau_{t_m,e_i}}|X_{t_m}^{t_m,e_i}]^*V_{t_{m+1}} \tag{3.21}$$

Note that the following lemma is useful to proceed:

Lemma 3.3.1. *For* $0\le t<s\le T$,

$$\mathbb{P}(X^{t,e_i}_{s\wedge\tau_{t,e_i}}=e_j|X_t^{t,e_i}=e_i)=e_i^*\exp((s-t)I_\Xi Q)e_j\quad\textit{for}\quad e_i\in\Xi. \tag{3.22}$$

Proof. See Section 3.7.2 □

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

$$Y_{t_m}^{t_m,e_i}\approx e_i^*\left[\exp(\Delta tI_\Xi Q)V_{t_{m+1}}+\int_{t_m}^{t_{m+1}}\exp((r-t_m)I_\Xi Q)I_\Xi H(r,V_r)dr\right] \tag{3.23}$$

for $e_i\in\Xi$. Hence

$$V_{t_m}=\sum_{i\in\Xi}Y_{t_m}^{t_m,e_i}e_i+\sum_{i\in\Xi^c}Y_{t_m}^{t_m,e_i}e_i \\ \approx I_\Xi\left[\exp(\Delta tI_\Xi Q)V_{t_{m+1}}+\int_{t_m}^{t_{m+1}}\exp((r-t_m)I_\Xi Q)I_\Xi H(r,V_r)dr\right]+I_{\Xi^c}\boldsymbol{\zeta}_{t_m}. \tag{3.24}$$

The variation-of-constants formula for a system of ODEs is appeared in the first term of (3.24). Motivated by this observation, we discretize (3.24) using exponential integrators.

(The Nørsett-Euler Method). For example, applying the Nørsett-Euler method to (3.24) in what follows:

$$V_{t_m} \approx I_\Xi \left[e^{\Delta t I_\Xi Q} V_{t_{m+1}} + \Delta t \left(\int_0^1 e^{(1-\theta)\Delta t I_\Xi Q} d\theta \right) I_\Xi H(t_{m+1}, V_{t_{m+1}}) \right] + I_{\Xi^c} \boldsymbol{\zeta}_{t_m}.$$

Then, we obtain the following scheme.

$$\begin{cases} V_{t_M}^M \coloneqq \boldsymbol{\zeta}_{t_M}, \\ V_{t_m}^M \coloneqq I_\Xi \widetilde{V}_{t_m}^M + I_{\Xi^c} \boldsymbol{\zeta}_{t_m}, \\ \widetilde{V}_{t_m}^M \coloneqq e^{\Delta t I_\Xi Q} V_{t_{m+1}}^M + \Delta t \left(\int_0^1 e^{(1-\theta)\Delta t I_\Xi Q} d\theta \right) I_\Xi H(t_{m+1}, V_{t_{m+1}}^M), \end{cases} \tag{3.25}$$

for $m = 0, 1 \ldots, M-1$. Note that $\widetilde{V}_{t_m}^M$ is calculated by using the Nørsett-Euler method for one step.

(General s-Stage Exponential Integrators). We can choose different exponential integrators to calculate U_{t_m}. Specifically, the general s-stage exponential integrators in what follows can be applicable:

$$\begin{aligned} \widetilde{V}_{t_m}^M &= \chi_0(\Delta t I_\Xi Q) \widetilde{V}_{t_{m+1}}^M + \Delta t \sum_{i=1}^s b_i(\Delta t I_\Xi Q) G_{mi}, \\ G_{mi} &= I_\Xi H(t_{m+1} - c_i \Delta t, \zeta_{ni}^M), \quad \text{for} \quad i = 1, \ldots, s, \\ \zeta_{mi}^M &= \chi_i(\Delta t I_\Xi Q) \widetilde{V}_{m+1}^M + \Delta t \sum_{j=1}^s a_{ij}(\Delta t I_\Xi Q) G_{nj}, \quad \text{for} \quad i = 1, \ldots, s, \end{aligned} \tag{3.26}$$

for $m = 0, \ldots, M-1$. Here, G_{mi} is the i-th internal stage, $s \in \mathbb{N}$ is the number of stages, c_i are real numbers, and χ_i, a_{ij} and b_i are functions constructed from "ϕ-functions" defined by

$$\phi_l(Q) = \int_0^1 e^{(1-\theta)Q} \frac{\theta^{l-1}}{(l-1)!} d\theta \quad \text{for} \quad l \in \mathbb{N}, \quad \text{and} \quad \phi_0(Q) = e^Q.$$

As a result, the following scheme is constructed.

$$\begin{cases} V_{t_M}^M \coloneqq \boldsymbol{\zeta}_{t_M}, \\ V_{t_m}^M \coloneqq I_\Xi \widetilde{V}_{t_m}^M + I_{\Xi^c} \boldsymbol{\zeta}_{t_m}, \\ \widetilde{V}_{t_m}^M \coloneqq (\text{calculated using (3.26)}) \end{cases} \tag{3.27}$$

for $m = 0, 1 \ldots, M-1$. If necessary, one can construct Y_{t_m} and Z_{t_m} using $V_{t_m}^M$ following the way in what follows: (i) Simulate $(X_t)_{t \in [0,T]}$ and obtain τ_{0,X_0} and $(X_{t_m})_{m=0}^M$. (ii) $Y_{t_m}^M \coloneqq X_{t_m}^* V_{t_m \wedge \tau_{0,X_0}}^M$ and $Z_{t_m}^M \coloneqq V_{t_m} 1_{\{t_m \le \tau_{0,X_0}\}}$.

3.4 An Application to BSDEs Driven by Brownian Motion

This section is devoted to constructing BSDEs with stopping terminal times driven by CTMCs from suitable spatial discretizations of those who are driven by Brownian motion, and we propose to use the multi-stage Euler-Maruyama methods to solve the obtained BSDEs driven by CTMCs. The arguments are carried out similarly to Section 4 in the previous work with slight modifications. As described there, this approach is motivated by the following observations: (i) Approximating Markov BSDEs driven by Brownian motion with BSDEs driven by CTMCs can be interpreted as the method of lines that approximates second-order parabolic PDEs with systems of ODEs, and vice versa. (ii) Systems of ODEs arising from the method of lines typically result in "stiff" systems that cause numerical instabilities. (iii) Exponential integrators can solve stiff systems successfully. In this section, we briefly explain the contruction. For details, see the previous work.

Hereafter, we focus on Markov BSDEs (3.13) such that $(\mathcal{X}_t)_{t\in[0,T]}$ is time homogeneous, namely, $\mu(t,x)$ and $\sigma(t,x)$ do not depend on t. Then, we can simply write $\mu(t,x) = \mu(x)$ and $\sigma(t,x) = \sigma(x)$, and the subscript of the infinitesimal generator can also be omitted: $\mathcal{L}_t = \mathcal{L}$.

3.4.1 Method of Lines Spatial Discretization

For $p = 1, \dots, d$, let a strictly increasing sequence $\Pi^{(p)} = \{x_n^{(p)}\}_{n=1}^{N^{(p)}}$ of length $N^{(p)}$ be the set of nodes on the p-th axis in $\mathbb{R}^d$, and define $\delta x_i^{(p)} := x_{i+1}^{(p)} - x_i^{(p)}$ for $i = 1, \dots, N^{(p)} - 1$. Using the Kronecker product "$\otimes$", we construct the grid on $\mathbb{R}^d$ by

$$\Pi = \Pi^{(1)} \otimes \cdots \otimes \Pi^{(d)} = (x_i = (x_i^{(1)}, \dots, x_i^{(d)}) : i = 1, \dots, N)$$

where $N := \prod_{p=1}^{d} N^{(p)}$. Note that multi-indices $(i_1, \dots, i_d)$ are then ordered lexicographically. For any function $v : \mathbb{R}^d \to \mathbb{R}$, let $v^\Pi = (v(t,x_1), v(t,x_2), \dots, v(t,x_N))^* \in \mathbb{R}^N$ be the evaluation of v over Π. Then, derivatives evaluated at nodes of Π are replaced by

$$\begin{pmatrix} \dfrac{\partial v}{\partial x^{(p)}}(x_1) \\ \vdots \\ \dfrac{\partial v}{\partial x^{(p)}}(x_N) \end{pmatrix} \approx \widetilde{D}_1^{(p)} u^\Pi, \quad \begin{pmatrix} \dfrac{\partial^2 v}{\partial x^{(p)} \partial x^{(q)}}(x_1) \\ \vdots \\ \dfrac{\partial^2 v}{\partial x^{(p)} \partial x^{(q)}}(x_N) \end{pmatrix} \approx \widetilde{D}_1^{(p,q)} u^\Pi, \quad \begin{pmatrix} \dfrac{\partial^2 v}{(\partial x^{(p)})^2}(x_1) \\ \vdots \\ \dfrac{\partial^2 v}{(\partial x^{(p)})^2}(x_N) \end{pmatrix} \approx \widetilde{D}_2^{(p)} u^\Pi.$$

Here, $N \times N$-matrices $\widetilde{D}_1$ and $\widetilde{D}_2$ are Kronecker products:

$$\begin{aligned}
\widetilde{D}_k^{(p)} &\coloneqq I_{N^{(1)}} \otimes \cdots \otimes I_{N^{(p-1)}} \otimes D_k^{(p)} \otimes I_{N^{(p+1)}} \otimes \cdots \otimes I_{N^{(d)}}, \quad k = 1, 2 \\
\widetilde{D}_1^{(p,q)} &\coloneqq I_{N^{(1)}} \otimes \cdots \otimes I_{N^{(p-1)}} \otimes D_1^{(p)} \otimes I_{N^{(p+1)}} \otimes \cdots \\
&\qquad \otimes I_{N^{(q-1)}} \otimes D_1^{(q)} \otimes I_{N^{(q+1)}} \otimes \cdots \otimes I_{N^{(d)}}, \quad p \neq q,
\end{aligned}$$

where $I_{N^{(p)}}$ is $N^{(p)} \times N^{(p)}$ identity matrix, and $D_1^{(p)}$ and $D_2^{(p)}$ are $N^{(p)} \times N^{(p)}$ matrices arising from central difference methods:

$$\begin{aligned}
e_i^* D_1^{(p)} e_j &= \begin{cases}
\dfrac{-\delta x_i^{(p)}}{\delta x_{i-1}^{(p)} (\delta x_{i-1}^{(p)} + \delta x_i^{(p)})}, & j = i-1, \\
\dfrac{\delta x_i^{(p)} - \delta x_{i-1}^{(p)}}{\delta x_i^{(p)} \delta x_{i-1}^{(p)}}, & j = i, \\
\dfrac{\delta x_{i-1}^{(p)}}{\delta x_i^{(p)} (\delta x_{i-1}^{(p)} + \delta x_i^{(p)})}, & j = i+1, \\
0 & \text{otherwise},
\end{cases} \quad \text{for} \quad 1 < i < N^{(p)}, \\
e_1^* D_1^{(p)} e_i &= e_{N^{(p)}}^* D_1^{(p)} e_i = 0, \quad \text{for} \quad 1 \leq i \leq N^{(p)},
\end{aligned} \tag{3.28}$$

$$\begin{aligned}
e_i^* D_2^{(p)} e_j &= \begin{cases}
\dfrac{2}{\delta x_{i-1}^{(p)} (\delta x_{i-1}^{(p)} + \delta x_i^{(p)})}, & j = i-1, \\
\dfrac{-2}{\delta x_i^{(p)} \delta x_{i-1}^{(p)}}, & j = i, \\
\dfrac{2}{\delta x_i^{(p)} (\delta x_{i-1}^{(p)} + \delta x_i^{(p)})}, & j = i+1, \\
0 & \text{otherwise},
\end{cases} \quad \text{for} \quad 1 < i < N^{(p)}, \\
e_1^* D_2^{(p)} e_i &= e_{N^{(p)}}^* D_2^{(p)} e_i = 0, \quad \text{for} \quad 1 \leq i \leq N^{(p)}.
\end{aligned} \tag{3.29}$$

where e_i is the i-th standard basis vector in $\mathbb{R}^{N^{(p)}}$. We then solve, in place of PDE (3.15), a system of N ODEs in what follows:

$$\begin{cases}
e_i^* \left(\dfrac{dU_t^\Pi}{dt} + QU_t^\Pi + F(t, U_t^\Pi) \right) = 0, & (t, e_i) \in [0, T) \times \Xi, \\
e_i^* U_t^\Pi = e_i^* \boldsymbol{\zeta}_t^\Pi, & (t, i) \in (\{T\} \times \{e_1, \ldots, e_N\}) \cup ([0, T] \times \Xi^c),
\end{cases} \tag{3.30}$$

Here, Ξ and Ξ^c are

$$\Xi \coloneqq \Pi \cap G \quad \text{and} \quad \Xi^c \coloneqq \Pi \cap \partial G,$$

$e_i^*\boldsymbol{\zeta}_t^\Pi = \chi(t,x_i)$ for $(t,x_i) \in [0,T]\times\Xi^c$, $e_i^*\boldsymbol{\zeta}_T^\Pi = (\kappa(x_1),\dots,\kappa(x_N))^*$, $F : [0,T]\times\mathbb{R}^N \to \mathbb{R}^N$ is defined by

$$F(t,z) = \begin{pmatrix} f(t,x_1,e_1^*z,\sigma(x_1)^*(e_1^*\widetilde{D}_1^{(1)}z,\dots,e_1^*\widetilde{D}_1^{(d)}z)^*) \\ \vdots \\ f(t,x_N,e_N^*z,\sigma(x_N)^*(e_N^*\widetilde{D}_1^{(1)}z,\dots,e_N^*\widetilde{D}_1^{(d)}z)^*) \end{pmatrix}, \tag{3.31}$$

and

$$Q = \sum_{p=1}^{d} \mathrm{diag}((\mu^{(p)})^\Pi)\widetilde{D}_1^{(p)} + \sum_{p=1}^{d-1}\sum_{q=p+1}^{d} \mathrm{diag}((\sigma\sigma^*)^{(p,q)})^\Pi)\widetilde{D}_1^{(p,q)} + \frac{1}{2}\sum_{p=1}^{d} \mathrm{diag}((\sigma\sigma^*)^{(p,p)})^\Pi)\widetilde{D}_2^{(p)}$$

approximates the infinitesimal generator $\mathcal{L}$ of $\mathcal{X}$. The following proposition presents us a probabilistic point of view of ODE (3.30). Note that its proof is the same as of Proposition 2.4.2.

Proposition 3.4.1. *Suppose that* $\int_0^T f(t,x,0,0)^2dt < \infty$ *for any* $x \in \mathbb{R}^d$, *and that for some* $L > 0$,

$$|f(t,x,y,z) - f(t,x,y',z')|^2 \le L(|y-y'|^2 + \|z-z'\|^2)$$

for all $t \in [0,T]$, $x \in \mathbb{R}^d$, $y,y' \in \mathbb{R}$, *and* $z,z' \in \mathbb{R}^d$. *Suppose further that* Q *is a* Q*-matrix,* $e_i^*\widetilde{D}_1^{(p)}1 = 0$, *and that*

$$e_j^*Qe_i = 0 \Longrightarrow e_j^*\widetilde{D}_1^{(p)}e_i = 0 \quad \textit{for} \quad p = 1,\dots,d, \tag{3.32}$$

for $i,j = 1,\dots,N$. *Then, for any* $(t,e_i) \in [0,T]\times\mathcal{I}$,

$$Y_s^{t,e_i} = (X_{T\wedge\tau_{t,e_i}}^{t,e_i})^*\boldsymbol{\zeta}_{T\wedge\tau_{t,e_i}} + \int_{]s,T]} 1_{\{r\le\tau\}}(X_{r-}^{t,e_i})^*F(r,Z_r^{t,e_i})ds - \int_{]s,T]} dM_s^*Z_r^{t,e_i}$$

which is derived from the Markov chain approximation of (3.13), *has a unique solution.*

3.5 Numerical Results

In this section, we demonstrate the efficiency and stability of the numerical approach presented in Section 2.4, using BSDEs arising from pricing European barrier options. We apply spatial discretization to the BSDEs driven by Brownian motions, obtain BSDEs driven by CTMCs (i.e. a system of ODEs), and calculate numerical solutions Y_t^{t,e_i} using multi-stage Euler-Maruyama methods (i.e. exponential integrators for the associated ststems of ODEs.) First, we explain the details on settings in what follows:

Spatial Discretization We approximate the domain at hand with a Kronecker product of one dimensional grids $\Pi^{(1)}, \ldots, \Pi^{(d)}$ as described in Section 2.4. Throughout the section, $\Pi_x^{\mathrm{Unif}}(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (x_i)_{i=1}^{N_l+N_r+1}$ means the standard one-dimensional uniform grid such that $x_1 = x_{\mathrm{left}}$, $x_{N_l+1} = x_{\mathrm{center}}$ and $x_{N_l+N_r+1} = x_{\mathrm{right}}$. We occasionally omit its arguments and simply write Π_x^{Unif}. In addition to the uniform grid, a non-uniform grid is also employed for spatial discretization. We use a version of the (one-dimensional) Tavella-Randall-type grids [63] in what follows:

$$x_k = \begin{cases} x_{\mathrm{center}} + g_1 \sinh\left(\operatorname{arc\,sinh}\left(\dfrac{x_{\mathrm{center}} - x_{\mathrm{left}}}{g_1}\right) \cdot \dfrac{k - N_l - 1}{N_l}\right), \\ \qquad\qquad\qquad\qquad 1 \le k \le N_l + 1, \\ x_{\mathrm{center}} + g_2 \sinh\left(\operatorname{arc\,sinh}\left(\dfrac{x_{\mathrm{right}} - x_{\mathrm{center}}}{g_2}\right) \cdot \dfrac{k - N_l - 1}{N_r}\right), \\ \qquad\qquad\qquad\qquad N_l + 1 \le k \le N_l + N_r + 1, \end{cases} \tag{3.33}$$

where $N_l + N_r + 1$ is the grid size, x_{left} and x_{right} are the leftmost and rightmost points of the domain, $x_{\mathrm{center}} \in (x_{\mathrm{left}}, x_{\mathrm{right}})$ is the central point of the grid, and g_1 and g_2 are parameters for the left- and right-side of the grid, respectively. Note that $x_1 = x_{\mathrm{left}}$, $x_{N_l+1} = x_{\mathrm{center}}$ and $x_{N_l+N_r+1} = x_{\mathrm{right}}$. Intuitively, setting $g_1 \ll x_{\mathrm{center}} - x_{\mathrm{left}}$ and $g_2 \ll x_{\mathrm{right}} - x_{\mathrm{center}}$ lead to the grid that is highly concentrated around x_{center}. It is commonly used in numerical computation for pricing options to mitigate the effect of the nonlinearity of the payoff function [9, 53, 63]. Similarly to the uniform grid, denote $\Pi_x^{\mathrm{TR}}(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r, g_1, g_2)$ as the Tavella-Randall grid (3.33) whose parameters are $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r, g_1, g_2)$.

Temporal Discretization We employ solvers implemented in `DifferentialEquations.jl` [62], listed below:

- `LawsonEuler` : A single-stage method of classical/stiff order 1/1, referred to as the Lawson-Euler method [47].
- `NorsettEuler` : A single-stage method of classical/stiff order 1/1, referred to as the Nørsett-Euler method or ETD1RK method [22].
- `ETDRK2` : A 2-stage method of classical/stiff order 2/2 [22].
- `ETDRK3` : A 3-stage method of classical/stiff order 3/3 [22].
- `ETDRK4` : A 4-stage method of classical/stiff order 4/2 [22].

- `HochOst4` : A 5-stage method of classical/stiff order 4/4, developed by Hochbruck and Ostermann [39].

Taking temporal grid size $N_t \in \mathbb{N}$, we calculate solutions on the grid $\Pi_t^{\mathrm{Unif}}(N_t) = (i\Delta t)_{i=0}^{N_t}$ using these exponential integrators. Here, $\Delta t = T/N_t$ is the step size. Note that a large-scale system of ODEs is obtained from the spatial discretization in each experiment. In this case, employing Krylov subspace methods in evaluating matrix exponentials and related ϕ functions is more effective, as described in Remark 2.3.3. In all the experiments, we use the Arnoldi iteration with a size-m Krylov subspace, which is readily available on all the solvers above. For simplicity, we always take $m = 100$.

Implementation All of our experiments were performed on a 3.70 GHz, 64-GB RAM Linux workstation. Our code was written entirely in Julia [5] and all the plots were produced using `Plot.jl` [13]. The full code for the experiments is available at https://github.com/kanekoakihiro/EMCTMCBSDE.

3.5.1 Down-and-Out Call Option under the Black-Scholes Model

First, we consider the linear BSDE arising from pricing an European up-and-out call option under the Black-Scholes model

$$\mathcal{S}_t = s_0 + \int_0^t \mu \mathcal{S}_s ds + \int_0^t \sigma \mathcal{S}_s dW_s, \tag{3.34}$$

where $\mathcal{S}_t$ represents the spot price of the asset with initial price s_0, appreciation μ and volatility σ. Suppose that the option has rebate 10 uniformly. Denote the strike price as K, the barrier as $B(< s_0)$ and the maturity as T. Considering the self-financing portfolio to hedge the option as described in Section 3.1, we obtain

$$\mathcal{Y}_t = \phi(T \wedge \tau, \mathcal{S}_{T\wedge\tau}) - \int_t^T 1_{\{s<\tau\}} r \mathcal{Y}_t dt - \int_t^T \mathcal{Z}_t dW_t, \tag{3.35}$$

Here,

$$\phi(t,x) = \begin{cases} 0, & (t,x) \in [0,T] \times [0,B], \\ (x-K)^+, & (t,x) \in \{T\} \times (B,\infty). \end{cases}$$

is the payoff of the option, $\tau = \tau_{0,s_0} = \inf\{t \geq 0 : \mathcal{S}_t \notin (B,\infty)\}$, and the solution $\mathcal{Y}_t = \mathcal{Y}_{t\wedge\tau}^{t\wedge\tau,\mathcal{S}_{t\wedge\tau}}$ means the price of the option at time $t \in [0,T]$. We choose the

parameters of (3.35) as follows:

T	K	B	r	μ	σ
1	100	110	0.03	0.03	0.2

For the spatial grid, we choose Π_x^{TR} with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}) = (0, K, B)$, $N_l = N_r = 500$), $g_1 = 80.0$, and $g_2 = \frac{x_{\mathrm{right}} - x_{\mathrm{center}}}{x_{\mathrm{center}} - x_{\mathrm{left}}} g_1$.

In this section, we calculate the maximum absolute errors of $\mathcal{Y}_t^{t,s}$ in $(t,s) \in \Pi_t^{\mathrm{Unif}} \times ([80,120] \times \Pi_x^{\mathrm{TR}})$, the absolute error of $\mathcal{Y}_t^{t,s}$ at $(t,s) = (0,100)$ and runtimes in seconds.

Since an analytic formula of $\mathcal{Y}_t^{t,s}$ using the cumurative distribution function of the standard Gaussian distribution $\Psi(x)$ (e.g. p.152 in [36]) is known, we regard it as the true solution. In table 3.1 the maximum absolute errors in $(t,s) \in \Pi_t^{\mathrm{Unif}} \times ([80,120] \cap \Pi_x^{\mathrm{TR}})$, the absolute errors at $(t,s) = (0,100)$ and its runtimes in seconds.

		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$	$N_t = 500$	$N_t = 1000$
Lawson-Euler [47]	Sup Error	1.341e+6	3.956e+0	2.270e+0	8.980e−1	8.416e−2	1.114e−2	1.090e−3
	Abs Error	6.216e+2	3.282e+0	1.559e+0	$5.539E-01$	6.131e−4	2.301e−4	1.055e−4
	Runtime[s]	0.12	0.16	0.35	0.61	1.41	3.13	6.28
Nørsett-Euler [22]	Sup Error	4.586e+0	3.735e+0	2.203e+0	7.149e−1	8.368e−2	1.072e−2	1.089e−3
	Abs Error	4.364e+0	3.386e+0	1.766e+0	4.999e−1	4.593e−3	5.830e−5	3.110e−5
	Runtime[s]	0.09	0.15	0.31	0.57	1.64	3.07	6.66
ETD2RK [22]	Sup Error	4.586e+0	3.735e+0	2.203e+0	7.151e−1	8.368e−2	1.072e−2	1.091e−3
	Abs Error	4.364e+0	3.387e+0	1.766e+0	5.000e−1	4.707e−3	1.039e−4	8.314e−6
	Runtime[s]	0.21	0.36	0.58	1.44	2.42	5.86	11.86
ETD3RK [22]	Sup Error	4.586e+0	3.735e+0	2.203e+0	7.151e−1	8.368e−2	1.072e−2	1.092e−3
	Abs Error	4.364e+0	3.387e+0	1.766e+0	5.000e−1	4.706e−3	1.039e−4	8.309e−6
	Runtime[s]	0.24	0.37	1.01	2.16	4.18	10.15	20.10
ETD4RK [22]	Sup Error	4.586e+0	3.735e+0	2.203e+0	7.151e−1	8.368e−2	1.072e−2	1.092e−3
	Abs Error	4.364e+0	3.387e+0	1.767e+0	5.001e−1	4.707e−3	1.039e−4	8.309e−6
	Runtime[s]	0.40	0.62	1.55	3.22	6.92	16.72	30.22
HochOst4 [39]	Sup Error	4.586e+0	3.735e+0	2.203e+0	7.151e−1	8.368e−2	1.072e−2	1.092e−3
	Abs Error	4.364e+0	3.387e+0	1.766e+0	5.000e−1	4.708e−3	1.039e−4	8.309e−6
	Runtime[s]	0.41	0.67	1.87	3.27	6.35	15.30	30.23

Table 3.1: Results on numerical solutions of (3.35) using exponential integrators. Here, we spatially discretize it on Π_x^{TR} and solve the resulting system of ODEs. Here, the parameters of Π_x^{TR} are $x_{\mathrm{left}} = 0$, $x_{\mathrm{center}} = K = 100$, $x_{\mathrm{right}} = B = 110$, $N_{x,0} = 500$, and $g_1 = 80$, $g_2 = \frac{110-100}{100-0} g_1 = 8$. For each N_t, the numerical solution is evaluated on the grid $\Pi_t^{\mathrm{Unif}}(N_t) \times \Pi_x^{\mathrm{TR}}$. Maximum absolute errors in $\Pi_t^{\mathrm{Unif}} \times ([80,120] \cap \Pi_x^{\mathrm{TR}})$ are reported on the row of "Sup Error", absolute errors at $(t, s_0) = (0,100)$ are on the row of "Abs Error", and runtimes in seconds are at the bottom line.

3.5.2 Barrier Options under Stochastic Local Volatility Models

Hereafter, we will consider barrier options under stochastic local volatility (SLV) models. In general, SLV models are written as

$$\begin{cases} \mathcal{S}_t = S_0 + \int_0^t \omega(\mathcal{S}_s, v_s)ds + \int_0^t m(v_s)\Gamma(\mathcal{S}_s)dW_s^{(1)}, \\ v_t = v_0 + \int_0^t \mu(v_s)ds + \int_0^t \sigma(v_s)dW_s^{(2)}, \end{cases} \tag{3.36}$$

where $\langle W^{(1)}, W^{(2)}\rangle_t = \rho t$ with $\rho \in (-1, 1)$. Let $\mathbf{L} = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}$ be the lower triangular matrix constructed from the Cholesky decomposition $\mathbf{C} = \mathbf{L}\mathbf{L}^*$ of $\mathbf{C} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. Since $\mathbf{W} = \mathbf{L}^{-1}\begin{pmatrix} W^{(1)} \\ W^{(2)} \end{pmatrix}$ is a 2-dimensional standard Brownian motion, we can reformulate (3.36) as

$$\mathcal{X}_t = \mathcal{X}_0 + \int_0^t \boldsymbol{\mu}(\mathcal{X}_s)ds + \int_0^t \boldsymbol{\sigma}(\mathcal{X}_s)d\mathbf{W}_s \quad \text{for} \quad t \in [0, T],$$

where

$$\mathcal{X}_t = \begin{pmatrix} \mathcal{S}_t \\ v_t \end{pmatrix}, \quad \boldsymbol{\mu}(x) = \begin{pmatrix} \omega(x_1, x_2) \\ \mu(x_2) \end{pmatrix}, \quad \boldsymbol{\sigma}(x) = \begin{pmatrix} m(x_2)\Gamma(x_1) & 0 \\ 0 & \sigma(x_2) \end{pmatrix}\mathbf{L}.$$

Under such $\mathcal{S}_t$, we consider European barrier options whose payoff is written as $\phi(T \wedge \tau, \mathcal{S}_{T\wedge\tau})$, where ϕ is a function and τ is the first hitting time of the barrier. Then, the price of the option at $(t, \mathcal{S}_t, v_t) = (0, s_0, v_0)$ is the solution of BSDE

$$\mathcal{Y}_t = \phi(T \wedge \tau, \mathcal{S}_{T\wedge\tau}, v_{T\wedge\tau}) - \int_t^T 1_{\{r\le\tau\}} f(r, \mathcal{S}_r, v_r, \mathcal{Y}_r, \mathcal{Z}_r)dr - \int_t^T \mathcal{Z}_r^* dW_r, \tag{3.37}$$

where

$$f(t, s, v, y, z) = r(y - z^*\mathbf{L}^{-1}\boldsymbol{\sigma}(x)^{-1}x)^+ - R(y - z^*\mathbf{L}^{-1}\boldsymbol{\sigma}(x)^{-1}x)^-.$$

3.5.2.1 Up-and-Out Put Options under the Mean-Reverting SABR Model

Consider the mean-reverting SABR model [24] definied as

$$\begin{aligned} \mathcal{S}_t &= s_0 + \int_0^t \kappa(\zeta - \mathcal{S}_s)ds + \int_0^t \sqrt{v_s}\mathcal{S}_s^\beta dW_s^{(1)} \\ v_t &= v_0 + \int_0^t \eta(\theta - v_s)ds + \int_0^t \alpha\sqrt{v_s}dW_s^{(2)}, \end{aligned} \tag{3.38}$$

where $\langle W^{(1)}, W^{(2)} \rangle_t = \rho t$ with $\rho \in (-1, 1)$, under which consider an up-and-out put option. The payoff is then written using

$$\phi(t, s, v) = \begin{cases} 0, & (t, s, v) \in [0, T] \times [B, \infty) \times [0, \infty), \\ (K - s)^+, & (t, s, v) \in \{T\} \times [0, B) \times [0, \infty), \end{cases}$$

where $B(> s_0)$ is the barrier, K is the strike price and T is the maturity. The corresponding BSDE is written in the form of (3.37). The parameters of (3.38) we have chosen are as follows:

T	β	κ	ζ	ρ	ν	θ	α	B	K	R	r
1.0	0.8	0.3	0.7	-0.9	0.5	1.0	0.4	120	100	0.55	0.05

For the spatial grid $\Pi^{(1)} \otimes \Pi^{(2)}$, we choose $\Pi^{(1)} \coloneqq \Pi_x^{\mathrm{TR}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}) = (1.0, K, B)$, $N_l = N_r = 100$, $g_1 = 5.0$, and $g_2 = \frac{x_{\mathrm{right}} - x_{\mathrm{center}}}{x_{\mathrm{center}} - x_{\mathrm{left}}} g_1$, and $\Pi^{(2)} \coloneqq \Pi_x^{\mathrm{Unif}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (0.0267, 0.4, 0.7733, 15, 15)$.

The result is presented in Table 3.3. As N_t increases, the solutions seem to converge approximately $\mathcal{Y}_0^{0,(100,0.4)} \approx 7.2360$. Note that numerical instabilities are appeared for $N_t = 10$. They are provoked by non-smoothness of f and the boundary condition and different from the stiffness exponential integrators are targetting.

3.5.2.2 Down-and-In Call Option under the 4/2 SABR Model

The following SDE is referred to as the 4/2 SABR model, proposed in Grasselli [33].

$$\begin{aligned} \mathcal{S}_t &= \mathcal{S}_0 + \int_0^t \kappa \mathcal{S}_s ds + \int_0^t (a\sqrt{v_s} + b/\sqrt{v_s}) \mathcal{S}_s^{\beta} dW_s^{(1)} \\ v_t &= v_0 + \int_0^t \eta(\theta - v_s) ds + \int_0^t \alpha \sqrt{v_s} dW_s^{(2)}, \end{aligned} \tag{3.39}$$

where $\langle W^{(1)}, W^{(2)} \rangle_t = \rho t$ with $\rho \in (-1, 1)$. We consider a down-and-in call option under (3.39), a knock-in option. Recall that, at time $t = T$, knock-in options are paid if its underlying asset breaches the predetermined barrier before T; otherwise, a rebate is paid. Denote K as the strike price, and suppose that there is no rebate. In this case, it is represented as a random variable $\phi(T \wedge \tau, \mathcal{S}_{T \wedge \tau}, v_{T \wedge \tau})$ with a function defined as

$$\phi(t, s, v) = \begin{cases} \widehat{\mathcal{Y}}_t^{t,s,v}, & (t, s, v) \in [0, T] \times [0, B] \times [0, \infty), \\ 0, & (t, s, v) \in \{T\} \times (B, \infty) \times [0, \infty), \end{cases}$$

		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$
Lawson-Euler [47]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	0.0097	1.7581	7.1640	7.2172	7.2277
	Runtime[s]	16.38	28.53	66.04	127.67	253.40
Nørsett-Euler [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	6.2364	7.0297	7.2311	7.2335	7.2347
	Runtime[s]	20.09	27.50	63.96	126.26	241.96
ETD2RK [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	7.0322	12.7319	7.2372	7.2364	7.2361
	Runtime[s]	26.18	47.37	112.88	225.22	437.73
ETD3RK [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	314.9722	7.5636	7.2360	7.2360	7.2360
	Runtime[s]	36.10	66.59	160.46	342.93	676.57
ETD4RK [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	-3588.9123	7.2483	7.2358	7.2360	7.2360
	Runtime[s]	61.61	104.54	252.92	524.85	1039.47
HochOst4 [39]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	7494.2639	6.8035	7.2357	7.2360	7.2360
	Runtime[s]	55.99	105.88	260.48	518.88	1023.81

Table 3.2: Results on numerical solutions $\mathcal{Y}_0^{0,(100,0.4)}$ of Markov BSDE (3.37) with (3.38) and its runtime in seconds using different exponential integrators. We spatially discretize the BSDE on $\Pi_x^{\mathrm{TR}} \otimes \Pi_x^{\mathrm{Unif}}$ and solve the resulting systems of ODEs. Here, the parameters of Π_x^{TR} are $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r, g_1, g_2) = (1.0, K, B, 100, 100, 5.0, \frac{x_{\mathrm{right}} - x_{\mathrm{center}}}{x_{\mathrm{center}} - x_{\mathrm{left}}} g_1)$, and of Π_x^{Unif} are $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (0.0267, 0.4, 0.7733, 15, 15)$.

where $\widehat{\mathcal{Y}}_t^{t,(s,v)}$ is the price of the European (vanilla) call option at time t and spot price $(\mathcal{S}_t, v_t) = (s, v)$ satisfying the BSDE (with deterministic terminal time T)

$$\widehat{\mathcal{Y}}_t^{t,(s,v)} = (\mathcal{S}_T^{t,(s,v)} - K)^+ + \int_t^T f(s, \mathcal{S}_s^{t,(s,v)}, \widehat{\mathcal{Y}}_s^{t,(s,v)}, \widehat{\mathcal{Z}}_s^{t,(s,v)}) ds - \int_t^T \widehat{\mathcal{Z}}_s^{t,(s,v)} dW_s. \quad (3.40)$$

Now, we are interested in $\mathcal{Y}_0^{0,(s_0,v_0)}$ of (3.37) for $(s_0, v_0) = (100, 0.4)$. We choose the parameters of (3.39) as follows:

T	β	κ	ρ	ν	θ	α	a	b	B	K	R	r
1.0	0.8	0.7	0.25	0.9	0.02	0.004	0.8	0.05	90	100	0.15	0.09

Since the $\widehat{\mathcal{Y}}_t^{t,(s,v)}$ is also unknown, we first solve (3.40), and then using the numerical solution of $\widehat{\mathcal{Y}}_t^{t,(s,v)}$ for $s = B$, we calculate $\mathcal{Y}_t^{t,(s,v)}$. To obtain $\widehat{\mathcal{Y}}_t^{t,(s,v)}$, we discretize (3.40) on spatial grid $\widehat{\Pi}^{(1)} \otimes \widehat{\Pi}^{(2)}$ where

- $\widehat{\Pi}^{(1)} \coloneqq \Pi_x^{\mathrm{TR}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r, g_1, g_2) = (1, B, 2B-1, 100, 100, 50, 50)$,

- $\widehat{\Pi}^{(2)} := \Pi_x^{\mathrm{Unif}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (0.0267, 0.4, 0.7733, 15, 15)$.

The spatial grid $\Pi^{(1)} \otimes \Pi^{(2)}$ on which (3.37) is discretized is chosen as follows:

- $\Pi^{(1)} := \Pi_x^{\mathrm{TR}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}) = (B, K, 2K)$, $N_l = N_r = 100$, $g_2 = 50$, and $g_1 = \frac{x_{\mathrm{center}} - x_{\mathrm{left}}}{x_{\mathrm{right}} - x_{\mathrm{center}}} g_2 = 1$,
- $\Pi^{(2)} := \Pi_x^{\mathrm{Unif}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (0.0267, 0.4, 0.7733, 15, 15)$.

The result is presented in Table 3.3. As N_t increases, the solutions seem to converge approximately $\mathcal{Y}_0^{0,(100,0.4)} \approx 4.074$.

		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$	$N_t = 500$
Lawson-Euler [47]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	1.0155	2.7035	3.3656	3.6428	3.8525	3.9881
	Runtime[s]	16.27	26.50	62.19	115.91	242.25	563.67
Nørsett-Euler [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	2.6224	3.4995	3.9315	4.0236	4.0540	4.0701
	Runtime[s]	14.32	28.78	58.18	109.45	211.17	522.18
ETD2RK [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	2.8137	3.5893	3.9688	4.0422	4.0633	4.0738
	Runtime[s]	23.40	42.02	105.36	221.58	416.25	1012.08
ETD3RK [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	2.9080	3.6321	3.9803	4.0462	4.0647	4.0742
	Runtime[s]	31.56	59.38	146.88	295.14	639.15	1698.71
ETD4RK [22]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	2.8982	3.6268	3.9804	4.0464	4.0647	4.0742
	Runtime[s]	52.00	103.71	269.59	528.56	990.81	2447.03
HochOst4 [39]	$\mathcal{Y}_0^{0,(s_0,v_0)}$	2.9356	3.6359	3.9811	4.0464	4.0647	4.0742
	Runtime[s]	52.65	101.06	250.44	495.44	980.85	2473.19

Table 3.3: Results on numerical solutions $\mathcal{Y}_0^{0,(100,0.4)}$ of Markov BSDE (3.37) with (3.38) and its runtime in seconds using different exponential integrators. We spatially discretize the BSDE on $\Pi_x^{\mathrm{TR}} \otimes \Pi_x^{\mathrm{Unif}}$ and solve the resulting systems of ODEs. Here, the parameters of Π_x^{TR} are $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r, g_1, g_2) = (1.0, K, B, 100, 100, 5.0, \frac{x_{\mathrm{right}} - x_{\mathrm{center}}}{x_{\mathrm{center}} - x_{\mathrm{left}}} g_1)$, and of Π_x^{Unif} are $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (0.0267, 0.4, 0.7733, 15, 15)$.

3.6 Conclusion

BSDEs with terminal times being bounded stopping times are an extension of BSDEs and have a great interest in applications. In this chapter, we constructed multi-stage Euler-Maruyama methods for solving Markov BSDEs driven by CTMCs with terminal times being bounded stopping times. We conducted it by extending our

previous work in which these methods for solving Markov BSDEs driven by CTMCs with terminal times being deterministic were constructed. The resulting method partially employs exponential integrators to solve the associated systems of ODEs, which can treat "stiff" BSDEs effectively. As a numerical method for solving BSDEs driven by Brownian motions with terminal times being bounded stopping times, we proposed to discretize them spatially and solve the resulting "stiff" BSDEs driven by CTMCs using the presented methods. To illustrate the effectiveness, we presented numerical experiments that treat (Brownian-)BSDEs arising from pricing problems of barrier options.

3.7 Proofs

3.7.1 Proof of Theorem 3.2.2

Proof. For a solution U_t to (3.9) , the Itô formula immediately implies $(Y_t, Z_t) = (X^*_{t\wedge\tau} U_{t\wedge\tau}, U_t 1_{\{t\leq\tau\}})$ solves (3.7).

Let t_1 and t_2 be fixed. Without loss of generality, assume that $t_2 > t_1$.

$$\begin{aligned}
Y^{t_1,e_i}_{t_1} - Y^{t_2,e_i}_{t_2} &= \mathbb{E}[Y^{t_1,e_i}_{t_1} - Y^{t_1,e_i}_{t_2} + (X^{t_1,e_i}_{t_2})^* V_{t_2} - e_i^* V_{t_2}] \\
&= \mathbb{E}\left[\int_{]t_1,t_2]} 1_{\{u\leq\tau_{t_1,e_i}\}} h(X^{t_1,e_i}_{u-}, u, Y^{t_1,e_i}_{u-}, Z^{t_1,e_i}_u)du - \int_{]t_1,t_2]} dM^*_u Z^{t_1,e_i}_u \right. \\
&\qquad \left. + \left(\int_{]t_1,t_2]} Q^*_u X^{t_1,e_i}_{u-} - M_{t_2} + M_{t_1}\right)^* V_{t_2}\right] \\
&= \mathbb{E}\left[\int_{]t_1,t_2]} [1_{\{u\leq\tau_{t_1,e_i}\}} h(X^{t_1,e_i}_{u-}, u, Y^{t_1,e_i}_{u-}, Z^{t_1,e_i}_u) + (X^{t_1,e_i}_{u-})^* Q_u V_{t_2}]du\right]
\end{aligned}$$

Hence,

$$\begin{aligned}
&|Y^{t_1,e_i}_{t_1} - Y^{t_2,e_i}_{t_2}| \leq \\
&\qquad C\sqrt{t_2 - t_1}\sqrt{\int_{]t_1,t_2]} \mathbb{E}[|1_{\{u\leq\tau_{t_1,e_i}\}} h(X^{t_1,e_i}_{u-}, u, Y^{t_1,e_i}_{u-}, Z^{t_1,e_i}_u) + (X^{t_1,e_i}_{u-})^* Q_u V_{t_2}|^2]du}
\end{aligned}$$

Using the uniform boundedness of Q_u, the Lipschitz continuity of h, evaluate the integrand as

$$\begin{aligned}
&|1_{\{u\leq\tau_{t_1,e_i}\}} h(X^{t_1,e_i}_{u-}, u, Y^{t_1,e_i}_{u-}, Z^{t_1,e_i}_u) + (X^{t_1,e_i}_{u-})^* Q_u V_{t_2}|^2 \\
&\quad \leq C(|Y^{t_1,e_i}_{u-}|^2 + \|Z^{t_1,e_i}_u\|^2_{X^{t_1,e_i}_{u-}} + |1_{\{u\leq\tau_{t_1,e_i}\}} h(X^{t_1,e_i}_{u-}, u, 0, 0)|^2 + \sup_{\substack{0\leq s,u\leq T \\ i=1,\ldots,N}} |e_j^* Q_s V_u|^2).
\end{aligned}$$

Since

$$\begin{aligned}
&\mathbb{E}\int_{]t_1,t_2]}|1_{\{u<\tau_{t_1,e_i}\}}h(X_{u-}^{t_1,e_i},u,0,0)|^2du\\
&=\sum_{j=1}^{N}\int_{]t_1,t_2]}|h(e_j,u,0,0)|^2 1_{\{u\le\tau_{t_1,e_i}\}}\mathbb{P}(X_u=e_j|X_{t_1}=e_i)du\\
&=\sum_{e_j\in\Xi}\int_{]t_1,t_2]}|h(e_j,u,0,0)|^2\mathbb{P}(X_u=e_j|X_{t_1}=e_i)du\le C\sup_{e_j\in\Xi}\int_{]0,T]}|h(e_j,u,0,0)|^2du,
\end{aligned}$$

we obtain

$$\begin{aligned}
|Y_{t_1}^{t_1,e_i}-Y_{t_2}^{t_2,e_i}|\le C\sqrt{t_2-t_1}\Bigg(\mathbb{E}\Bigg[&\sup_{0\le u\le T}|Y_u^{t_1,e_i}|^2+\int_{]0,T]}\|Z_u^{t_1,e_i}\|^2_{X_{u-}^{t_1,e_i}}\\
&+\sup_{e_j\in\Xi}\int_{]0,T]}|h(e_i,u,0,0)|^2du+1\Bigg]du\Bigg),
\end{aligned}$$

which implies the continuity of $t\mapsto Y_t^{t,e_i}$. Notice that

$$\begin{aligned}
(\Delta X_u^{t,e_i})^*V_u1_{\{u\le\tau_{t,e_i}\}}&=\Delta((X_u^{t,e_i})^*V_u)1_{\{u\le\tau_{t,e_i}\}}\\
&=\Delta Y_u^{u,X_u^{t,e_i}}\cdot 1_{\{u\le\tau_{t,e_i}\}}=\Delta Y_u^{t,e_i}\cdot 1_{\{u\le\tau_{t,e_i}\}}.
\end{aligned}$$

Together it with $\Delta Y_u^{t,e_i}1_{\{u>\tau_{t,e_i}\}}=\Delta Y_{\tau_{t,e_i}}^{t,e_i}1_{\{u>\tau_{t,e_i}\}}=0$, we obtain

$$(\Delta X_u^{t,e_i})^*V_u1_{\{u\le\tau_{t,e_i}\}}=\Delta Y_u^{t,e_i}=\Delta M_u^*Z_u^{t,e_i}=(\Delta X_u^{t,e_i})^*Z_u^{t,e_i}.$$

Hence $\int_{]0,t]}(dX_u^{t,e_i})^*(Z_u^{t,e_i}-V_u1_{\{u\le\tau_{t,e_i})\}}=0$, which implies

$$\int_{[0,t]}dM_u^*(Z_u^{t,e_i}-V_u1_{\{u\le\tau_{t,e_i})\}}=-\int_{[0,t]}X_{u-}^{t,e_i}Q_u(Z_u^{t,e_i}-V_u1_{\{u\le\tau_{t,e_i})\}}du=0,$$

since any predictable finite variation martingales starting at 0 takes zero constantly (e.g. Corollary 8.2.14 p.204 in [19].) Hence

$$\mathbb{E}\int_{]0,T]}\|Z_u^{t,e_i}-V_u1_{\{u\le\tau_{t,e_i})\}}\|^2_{X_{u-}^{t,e_i}}du=\mathbb{E}\left|\int_{]0,T]}dM_u^*(Z_u^{t,e_i}-V_u1_{\{u\le\tau_{t,e_i})\}})\right|^2=0,$$

which means $Z_u^{t,e_i}\sim_M V_u1_{\{u\le\tau_{t,e_i})\}}$ Together with the Lipschitz continuity,

$$\begin{aligned}
&h(X_{u-}^{t,e_i},u,(X_{u-}^{t,e_i})^*V_u1_{\{u\le\tau_{t,e_i}\}},Z_u^{t,e_i})\\
&\quad=h(X_{u-}^{t,e_i},u,(X_{u-}^{t,e_i})^*V_u1_{\{u\le\tau_{t,e_i}\}},V_u1_{\{u\le\tau_{t,e_i}\}}),\quad du\otimes d\mathbb{P}\text{-a.s.}
\end{aligned}\tag{3.41}$$

Plugging it into the conditional expectation representation of Y_t^{t,e_i}, we obtain for $t \in [0,T]$,

$$e_i^* V_t = Y_t^{t,e_i} = e_i^* \Phi(T,t) G + e_i^* \int_{]t,T]} \Phi(u,t) H(u, V_s) du,$$

which results in the cariation-of-constants of (3.9) in what follows:

$$V_t = \Phi(T,t) G + \int_t^T \Phi(s,t) H(s, V_s) ds.$$

□

3.7.2 Proof of Lemma 3.3.1

Without loss of generality, we assume that $t = 0$ and $s = T$. For $M_1 \in \mathbb{N}$, let $\delta t \coloneqq T/M_1$ and $\tau_{0,e_i}^{M_1} = \inf\{l \cdot \delta t : l = 0, \dots, M_1, X_{l\cdot\delta t}^{0,e_i} \notin \Xi\}$. From the dominated convergence theorem,

$$\mathbb{P}(X_{T\wedge\tau_{0,e_i}}^{0,e_i} = e_j | X_0^{0,e_i} = e_i) = \lim_{M_1\to\infty} \mathbb{P}(X_{T\wedge\tau_{0,e_i}^{M_1}}^{0,e_i} = e_j | X_0^{0,e_i} = e_i).$$

For $r \in [0,T]$, let $P_r^{M_1}$ be the $N \times N$ matrix defined by $e_i^* P_r^{M_1} e_j = \mathbb{P}(X_{r\wedge\tau_{0,e_i}^{M_1}}^{0,e_i} = e_j | X_0^{0,e_i} = e_i)$ for $i,j = 1,\dots,N$. Clearly, $e_i^* P_r^{M_1} e_j = \delta_{ij}$ for $e_i \notin \Xi$ and $r \in [0,T]$. Here, δ_{ij} is the Kronecker's delta. For $e_i \in \Xi$, we obtain

$$\begin{aligned}
e_i^* P_T^{M_1} e_j &= \sum_{k=1}^N \mathbb{P} X_{T\wedge\tau_{0,e_i}^{M_1}}^{0,e_i} = e_j | X_{\delta t}^{0,e_i} = e_k) \mathbb{P}(X_{\delta t}^{0,e_i} = e_k | X_0^{0,e_i} = e_i) \\
&= \sum_{k=1}^N \mathbb{P}(X_{(T-\delta t)\wedge\tau_{0,e_i}^{M_1}}^{0,e_i} = e_j | X_0^{0,e_i} = e_k) \mathbb{P}(X_{\delta t}^{0,e_i} = e_k | X_0^{0,e_i} = e_i) \\
&= e_i^* \Phi_{\delta t} P_{T-\delta t}^{M_1} e_j
\end{aligned}$$

In other words, the two cases mean $I_{\Xi^c} P_r^{M_1} = I_{\Xi^c}$ and $I_\Xi P_r^{M_1} = I_\Xi \Phi_{\delta t} P_{T-\delta t}^M$, respectively. Combining them,

$$P_T^{M_1} = I_{\Xi^c} + I_\Xi \Phi_{\delta t} P_{T-\delta t}^{M_1} = (I_{\Xi^c} + I_\Xi \Phi_{\delta t}) P_{T-\delta t}^{M_1} = (I_{\Xi^c} + I_\Xi \Phi_{\delta t})^{M_1} P_0^{M_1} = (I_{\Xi^c} + I_\Xi \Phi_{\delta t})^{M_1}.$$

Since $\Phi_{\delta t} = I_N + \delta t \cdot Q + o(\Delta t)$,

$$\begin{aligned}
\lim_{M_1\to\infty} (I_\Xi \Phi_{\delta t} + I_{\Xi^c})^{M_1} &= \lim_{\delta t\to 0} (I_\Xi (I_N + \delta t Q + o(\delta t)) + I_{\Xi^c})^{T/\delta t} \\
&= \lim_{\delta t\to 0} (I_N + \delta t \cdot (I_\Xi Q) + o(\delta t))^{T/\delta t} = \exp(T I_\Xi Q).
\end{aligned}$$

Chapter 4

A Sparse Grid-Based Multilevel Spatial Discretization for BSDEs Driven by Brownian Motions

4.1 Introduction

In Chapters 2 and 3, we developed multi-stage Euler-Maruyama methods for BSDEs driven by CTMCs whose terminal times are deterministic or bounded stopping times, respectively. We saw that these methods are equivalent to exponential integrators for solving the associated ODEs. This observation led us to applying them to BSDEs driven by CTMCs arising from the spatial discretization of BSDEs driven by Brownian motion. In multi-dimensional settings, however, spatial discretization typically results in CTMCs with a large state space. That brings us to a numerical limitation known as the curse of dimensionality, and computation in more than 4-dimensional cases is challenging.

This chapter aims to develop a CTMC approximation method of solving high-dimensional BSDEs, which overcomes the curse of dimensionality. Specifically, we present a multilevel spatial discretization on a sparse grid for solving Markov BSDEs driven by Brownian motion, in which we construct a sequence of BSDEs driven by CTMCs on grids with different resolutions and suitably superimpose their solutions. With the help of the idea of sparse grid methods [11, 34], it reduces the computational cost and efficiently mitigates the curse of dimensionality.

The organization of this chapter is as follows: The next section presents preliminaries, in which we recall BSDEs with deterministic terminal times (or bounded stopping terminal times) driven by CTMCs (or Brownian motion). In Section 4.3, the

presented method, the sparse grid-based multilevel spatial discretization is presented. Finally, several numerical experiments are devoted to confirming its efficiency.

4.1.1 Notations

Throughout this chapter, the same notations as Chapters 2 and 3 are used.

4.2 Preliminaries

4.2.1 Spatial Discretization of BSDEs

Let us briefly recall the spatial discretization of BSDEs driven by Brownian motion.

Let $W = (W_t)_{t\geq 0}$ be a d-dimensional standard Brownian motion. Let $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ be the completion of the filtration generated by W. For $t \in [0,T]$ and $x \in \mathbb{R}^d$, we consider the Markov BSDE driven by Brownian motion starting at time t and state x in what follows:

$$\begin{cases} \mathcal{X}_s^{t,x} = x + \displaystyle\int_t^s \mu(r, \mathcal{X}_r^{t,x})dr + \int_t^s \sigma(r, \mathcal{X}_r^{t,x})dW_r \quad \text{for} \quad s \geq t, \\ \mathcal{X}_s^{t,x} = x \in \mathbb{R}^d \quad \text{for} \quad s \leq t, \\ \mathcal{Y}_s^{t,x} = g(\mathcal{X}_T^{t,x}) + \displaystyle\int_s^T f(r, \mathcal{X}_r^{t,x}, \mathcal{Y}_r^{t,x}, \mathcal{Z}_r^{t,x})dr - \int_s^T (\mathcal{Z}_r^{t,x})^* dW_r \quad \text{for} \quad s \in [0,T], \end{cases} \tag{4.1}$$

or its bounded stopping terminal time version in what follows:

$$\begin{cases} \mathcal{X}_s^{t,x} = x + \displaystyle\int_t^s \mu(r, \mathcal{X}_r^{t,x})dr + \int_t^s \sigma(r, \mathcal{X}_r^{t,x})dW_r \quad \text{for} \quad s \geq t, \\ \mathcal{X}_s^{t,x} = x \in \mathbb{R}^d \quad \text{for} \quad s \leq t, \\ \mathcal{Y}_{s\wedge\tau_{t,x}}^{t,x} = \chi(T \wedge \tau_{t,x}, \mathcal{X}_{T\wedge\tau_{t,x}}^{t,x}) + \displaystyle\int_s^T 1_{\{r<\tau_{t,x}\}} f(r, \mathcal{X}_r^{t,x}, \mathcal{Y}_r^{t,x}, \mathcal{Z}_r^{t,x})dr \\ \qquad\qquad\qquad\qquad - \displaystyle\int_s^T (\mathcal{Z}_r^{t,x})^* dW_r \quad \text{for} \quad s \in [0,T]. \end{cases} \tag{4.2}$$

Here, $G \subset \mathbb{R}^d$ is a connected open set whose boundary is of class C^1 and $\tau_{t,x} = \inf\{s \geq t : \mathcal{X}_s^{t,x} \notin \overline{G}\}$. The nonlinear Feynman-Kac formula states connections between these BSDEs and second-order semilinear parabolic PDEs. Let $\mathcal{L}_t$ be the

infinitesimal generator of the Markov process $\mathcal{X}$, that is,

$$\mathcal{L}_t u(t,x) = \sum_{i=1}^{d} \mu^{(i)}(t,x) \frac{\partial u}{\partial x_i}(t,x) + \frac{1}{2} \sum_{i,j=1}^{d} (\sigma\sigma^*)^{(i,j)}(t,x) \frac{\partial^2 u}{\partial x_i \partial x_j}(t,x). \tag{4.3}$$

The PDE associated with (4.1) is

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) + \mathcal{L}_t u(t,x) + f(t,x,u(t,x),\sigma^*(x)\nabla_x u(t,x)) = 0, \quad (t,x) \in [0,T] \times \mathbb{R}^d, \\ u(T,x) = g(x), \quad x \in \mathbb{R}^d, \end{cases} \tag{4.4}$$

and with (4.2) is

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) + \mathcal{L}_t u(t,x) + f(t,x,u(t,x),\sigma^*(x)\nabla_x u(t,x)) = 0, \quad (t,x) \in [0,T] \times G, \\ u(T,x) = \kappa(x), \quad x \in \overline{G}, \\ u(t,x) = \chi(t,x), \quad (t,x) \in [0,T] \times \partial G. \end{cases} \tag{4.5}$$

Suppose that PDE (4.4) (or PDE (4.5)) admits a unique classical solution u. Under some appropriate conditions, we obtain $\mathcal{Y}_t^{t,x} = u(t,x)$ in both cases. (For the precise statements, see Section 2.2.2 and Section 3.2.2.)

Markov BSDEs driven by CTMCs that approximate (4.1) or (4.2) can be derived from spatial discretizations of PDE (4.4) or (4.5). For each axis of the spatial domains of the PDEs, we define a one-dimensional grid on which spatial derivatives appeared in the PDE are approximated. With the Kronecker product, the approximated system of ODEs

$$\frac{dU_t}{dt} + QU_t + F(t,U_t) = 0, \quad U_T = G, \tag{4.6}$$

or

$$\begin{cases} e_i^* \left(\dfrac{dU_t^\Pi}{dt} + QU_t^\Pi + H(t,U_t^\Pi) \right) = 0, & (t,e_i) \in [0,T) \times \Xi, \\ e_i^* U_t^\Pi = e_i^* \boldsymbol{\zeta}_t^\Pi, & (t,i) \in (\{T\} \times \{e_1,\dots,e_N\}) \cup ([0,T] \times \Xi^c), \end{cases} \tag{4.7}$$

are obtained. We can interpret them from the probabilistic viewpoint. Assuming the validity of Q, let X be a finite-state Markov chain with Q as its Q-matrix. Consider the Markov BSDEs driven by CTMC X associated with (4.6) in what follows:

$$Y_t = X_T^* G + \int_{]t,T]} X_{s-}^* F(s,Z_s) ds - \int_{]t,T]} dM_s^* Z_s, \tag{4.8}$$

or ones associated with (4.7) in what follows:

$$Y_s^{t,e_i} = (X_{T\wedge\tau_{t,e_i}}^{t,e_i})^* \boldsymbol{\zeta}_{T\wedge\tau_{t,e_i}} + \int_{]s,T]} 1_{\{r\leq\tau\}} (X_{r-}^{t,e_i})^* F(r, Z_r^{t,e_i}) ds - \int_{]s,T]} dM_s^* Z_r^{t,e_i}. \quad (4.9)$$

A relatively mild condition ensures each BSDEs admits a unique solution (see Proposition 2.4.2). For more details on the spatial discretization, see Section 2.4 and 3.4.

4.2.2 Sparse Grids

Sparse grid methods date back to a study by Smolyak [66], in which he presented an algorithm for constructing multivariate quadrature and interpolation rules from a linear combination of tensor products of univariate ones in a specific manner. It mitigates the curse of dimensionality, which conventional formulas that compute solutions on the "full grid" cannot overcome. Beyond quadratures and interpolations, the construction of sparse grids has applications to various fields such as data mining [31] or differential equations. As introduced in Section 4.5 in [11], different approaches have been presented to apply the idea of sparse grids to solving partial differential equations. The most straightforward one among them will be the sparse grid combination technique. Motivated by the observation that the sparse grid can be decomposed to a combination of several coarser rectangular grids, this approach constructs a numerical solution from a linear combination of numerical solutions of PDEs on those grids. As a result, one can compute solutions by applying the PDE solver at hand to our PDE on each resolution grid without specific treatments.

4.3 Multilevel Spatial Approximation Using Sparse Grids

We first need to approximate the spatial domain of (4.4) (or (4.5)) with a fixed, bounded, and rectangular domain on which we can define sparse grids. For the sake of simplicity, we set a hypercube $[-1,1]^d$ as such a domain; we note that the argument in this section can be applied to arbitrary bounded rectangular domains using dilation and translation. and $p \in \{1,\dots,d\}$, let $\chi_p^l = \{z_1^{l,p},\dots,z_{m_l}^{l,p}\} \subset [-1,1]$ be a prescribed set of nodes parametrized by l and p; in this section, we specifically set equidistant nodes

$$z_i^{l,p} = \frac{2(i-1)}{m_l - 1} - 1, \quad \text{for} \quad i = 1,\dots,m_l,$$

with $m_l = 2^l + 1$. For $t \in [0, T]$, consider the piecewise linear interpolation of the solution $U_t^{(l_1,\dots,l_d)} \coloneqq U_t^{\chi_1^{l_1} \otimes \cdots \otimes \chi_d^{l_d}}$ of (4.6) (or (4.7)) for $\Pi = \chi_1^{l_1} \otimes \cdots \otimes \chi_d^{l_d}$ defined as

$$u^{(l_1,\dots,l_d)}(t, x_1, \dots, x_d) = \sum_{i_1=1}^{m_{l_1}} \cdots \sum_{i_d=1}^{m_{l_d}} U_t^{(l_1,\dots,l_d)}(z_{i_1}^{l_1,1}, \dots, z_{i_d}^{l_d,d}) \prod_{p=1}^{d} a_{i_p}^{l_p,p}(x_p)$$

for $(x_1, \dots, x_d) \in [-1, 1]^d$, where $a_{i_p}^{l_p,p}(x_p)$ are the one-dimensional standard hat functions defined as

$$a_i^{l,p}(x) = (1 - \frac{m_l - 1}{2}|x - z_i^{l,p}|) \vee 0$$

for $x \in [-1, 1]$, $i = 1, \dots, m_l$, $l \in \mathbb{N}$ and $p = 1, \dots, d$, and $U_t^{(l_1,\dots,l_d)}(z_{i_1}^{l_1,1}, \dots, z_{i_d}^{l_d,d})$ means the component of the vector $U_t^{(l_1,\dots,l_d)}$ corresponding to $(z_{i_1}^{l_1,1}, \dots, z_{i_d}^{l_d,d}) \in \chi_1^{l_1} \otimes \cdots \otimes \chi_d^{l_d}$. The sparse grid solution with the level parameter $q \in \mathbb{N}$ is then constructed from $u^{(l_1,\dots,l_d)}(t, x_1, \dots, x_d)$ by

$$u^{\mathrm{SG},q}(t, x_1, \dots, x_d) = \sum_{\substack{\mathbf{l}=(l_1,\dots,l_d) \in \mathbb{N}^d \\ q-d+1 \le |\mathbf{l}|_1 \le q}} (-1)^{q-|\mathbf{l}|_1} \binom{d-1}{q-|\mathbf{l}|_1} \cdot u^{\mathbf{l}}(t, x_1, \dots, x_d) \tag{4.10}$$

for $(x_1, \dots, x_d) \in [-1, 1]^d$, where $|\mathbf{l}|_1 = l_1 + \cdots + l_d$ for $\mathbf{l} = (l_1, \dots, l_d)$ and $\binom{d-1}{q-|\mathbf{l}|_1}$ is the binomial coefficient.

Numerically, we obtain an algorithm in what follows:

1. Solve systems of ODEs for each $\mathbf{l}$ such that $q - d + 1 \le |\mathbf{l}|_1 \le q$.
2. For each multi-index $\mathbf{l}$ and discrete time t_m, construct the d-dimensional piecewise linear interpolant $\mathbf{x} \mapsto u^{\mathbf{l}}(t_m, \mathbf{x})$ of $\{U_{t_m}^{\mathbf{l}}(\mathbf{z}), \mathbf{z} \in \chi_1^{l_1} \otimes \cdots \otimes \chi_d^{l_d}\}$.
3. Combine interpolants according to (4.10). The obtained function $u^{\mathrm{SG},q}(t, \mathbf{x})$ can be evaluated at any $\mathbf{x} \in [-1, 1]^d$ for each discretized time t_m.

In particular, we propose to use the multi-stage Euler-Maruyama methods constructed in Section 2.3 and Section 3.3 for obtaining each solution $U_t^{(l_1,\dots,l_d)}$ of ODEs, which is accomplished by employing exponential integrators. In the next section, we illustrate the efficiency of this approach through several numerical experiments.

The numerical solution $u^{\mathrm{SG},q}$ achieves much less computational cost with a slight deterioration in its quality, in comparison to the corresponding full grid formula i.e. $u^{(n,\dots,n)}$ for $n = q - d + 1$. Specifically, the total number of spatial points that the

sparse grid formula with parameter q contains is $O(2^q q^{d-1})$. Taking that the full grid formula at the same level is comprised of $O(2^{qd})$ spatial points into account, it turns out to be a significant reduction in computational complexity [34]. We can also give an error estimate based on arguments in existing leteratures such as [34, 48, 51] . Let $\alpha > 0$ be the order of accuracy of the spatial discretization scheme. Suppose that we know the exact solutions $U_t^{\mathbf{l}}$ of each systems of ODEs in (4.10), the exact solution u (of the approximated PDE on $[-1,1]^d$) are sufficiently smooth, and an error expansion

$$\begin{aligned} &u(t,x_1,\dots,x_d) - u^{\mathbf{l}}(t,x_1,\dots,x_d) \\ &\qquad = \sum_{i=1}^{d} \sum_{j_1=1}^{d} \sum_{j_2=j_1+1}^{d} \cdots \sum_{j_i=j_{i-1}+1}^{d} C^{(i)}(t,x_1,\dots,x_d,h_{j_1},h_{j_2},\dots,h_{j_i}) h_{j_1}^{\alpha} h_{j_2}^{\alpha} \cdots h_{j_i}^{\alpha} \end{aligned}$$

exists for some bounded functions $C^{(i)}$ for $i = 1,\dots,d$. Here, $h_i = 2^{-i}$ for $i \in \mathbb{N}$. Then, according to the existing literatures aforementioned, the error estimate reads

$$u(t,x_1,\dots,x_d) - u^{\mathrm{SG},q}(t,x_1,\dots,x_d) = O(h_q^{-\alpha}(\log_2(h_q^{-1}))^{d-1}) (= O(2^{-\alpha q} q^{d-1}))$$

as $q \to \infty$, which is slightly worse than $O(h_q^{-d}) (= O(2^{-\alpha q}))$ the estimate in the full grid case.

4.4 Numerical Results

We present numerical experiments in this section. In Section 4.4.1 and 4.4.2, we consider BSDEs driven by Brownian motion arising from pricing European call options under stochastic local volatility models as treated in Section 2.5. First, Section 4.4.1 treats two-dimensional BSDEs, in which we compare the sparse grid-based multilevel spatial discretization developed in this chapter with the "full grid-based" spatial discretization developed in Chapter 2 and confirm the efficiency of the former one. In Section 4.4.2, we solve a four dimensional BSDE arising related to Basket options using the presented method. A BSDE with bounded stopping terminal time is also treated in Section 4.4.3; under the similar situation in Section 4.4.2, we consider a BSDE arising from prising European basket barrier options. All the experiments were conducted in the same environment as in Chapters 2 and 3, and the same notations are used.

4.4.1 A Comparison to the "Full Grid" Method

In this section, we consider two ways of spatial discretizations of BSDEs driven by Brownian motion: (1) the discretization on the Kronecker product of grids (we call it a "full grid"), the same as considered in all the experiments in Chapters 2 and 3, and (2) the multilevel discretization on a sparse grid presented in this chapter. We calculate numerical solutions of discretized systems of ODEs on the two grids using exponential integrators and compare them.

Problem Setting : European Call Option under the SABR Model Consider pricing a European call option under the stochastic-alpha-beta-rho (SABR) model. It is a stochastic local volatility (SLV) model being commonly used, and designed to model price dynamics of a forward contract. Consider the forward price process $F_t = \mathcal{S}_t e^{r(T-t)}$ for some asset price $\mathcal{S}_t$ and interest rate r. The SABR model assumes that F_t satisfies the following SLV model

$$F_t = F_0 + \int_0^t v_s F_s^{\beta} dW_s^{(1)}, \quad v_t = v_0 + \int_0^t \alpha v_s dW_s^{(2)}. \tag{4.11}$$

To consider pricing of the European call option under the SABR model, we additionally suppose that the volatility process v_t also satisfies $v_t = v_t^0 e^{r(T-t)}$ for some underlying asset v_t^0. Taking the hedge portofolio of two risky assets $\mathcal{S}_t$ and v_t^0 and a bond with riskless rate r into account, the corresponding Makov BSDE (4.1) becomes

$$\mathcal{Y}_t = (F_T - K)^+ - \int_t^T r\mathcal{Y}_s ds - \int_t^T \mathcal{Z}_s d\mathbf{W}_s \tag{4.12}$$

with the state process (4.11).

Fortunately, $\mathcal{Y}_t$ has an approximation formula in what follows:

$$\mathcal{Y}_t \approx \exp(-r(T-t))[F_t \cdot \Psi(d_1) - K \cdot \Psi(d_2)], \tag{4.13}$$

where

$$d_1 = \frac{\ln\left(\frac{F_t}{K}\right) + \left(r + \frac{\sigma_B^2}{2}\right)(T-t)}{\sigma_B\sqrt{T-t}}, \quad d_2 = \frac{\ln\left(\frac{F_t}{K}\right) + \left(r - \frac{\sigma_B^2}{2}\right)(T-t)}{\sigma_B\sqrt{T-t}},$$

and σ_B is the approximated implied volatility by Hagan et al. [35] defined as

$$\sigma_B = \frac{v_t\left\{1 + \left[\frac{(1-\beta)^2}{24}\frac{v_t^2}{(F_tK)^{1-\beta}} + \frac{1}{4}\frac{\rho\beta\alpha v_t}{(F_tK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24}\right](T-t)\right\}}{(F_tK)^{(1-\beta)/2}\left\{1 + \frac{(1-\beta)^2}{24}\log^2(F_t/K) + \frac{(1-\beta)^4}{1920}\log^4(F_t/K)\right\}}\frac{z}{\chi(z)},$$

where z and $\chi(z)$ are

$$z = \frac{\alpha}{v_t}(F_t K)^{\frac{1-\beta^2}{2}} \log\left(\frac{F_t}{K}\right), \quad \chi(z) = \log\left(\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}\right).$$

In this section, we regard the solutions calculated using (4.13) as the exact solutions. We choose the parameters as follows:

T	α	β	ρ	r	K
1.0	0.4	0.9	0.3	0.05	100

4.4.1.1 Spatial Discretization on a "Full Grid"

First, we discretize (4.12) on the Kronecker product $\Pi_x^{\mathrm{TR}} \otimes \Pi_x^{\mathrm{Unif}}$ of grids similarly to the previous experiments. Precisely, $\mathcal{S}_t$ is discretized on the Tavella-Randall grid Π_x^{TR} with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_{x,0}, g_1, g_2) = (0, 100, 200, 100, 5, 5)$, and v_t is discretized on the uniform grid Π_x^{Unif} with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_{x,0}) = (0, 0.4, 0.8, 15)$. The size of the resulting grid is 6231. (For definitions of both grids, see Section 2.5)

Table 4.1 reports the maximum absolute errors in $(t, s, v) \in \Pi_t^{\mathrm{Unif}} \times ([80, 120] \cap \Pi_x^{\mathrm{TR}}) \times ([0.32, 0.48] \cap \Pi_x^{\mathrm{Unif}})$, the absolute errors at $(t, s, v) = (0, 100, 0.4)$, and the runtime in seconds for different exponential integrators. In this spatial discretization, when N_t increases, the maximum absolute error and the absolute error seem to converge towards 3.240×10^{-2} and 1.730×10^{-3}, respectively. To further improve the accuracy of the solutions, increasing the number of points of the spatial grid or expanding the approximated domain should be required.

4.4.1.2 The Multilevel Discretization on a Sparse Grid

Next, we discretize (4.12) on a sparse grid using the algorithm presented in Section 4.3. While the discretization on "full grids" results in a single BSDE driven by a CTMC, it approximates BSDE (4.12) (driven by Brownian motion) with a sequence of BSDEs driven by CTMCs on grids with different resolutions. Recalling that the algorithm presented in Section 4.3, the numerical solution of (4.12) can be constructed in the following steps: (1) Let $q \in \mathbb{N}$ be fixed. (2) For each $\mathbf{l} = (l_1, l_2)$ such that $q - 1 \leq |l|_1 \leq q$, discretize $\mathcal{S}_t$ and v_t on $\Pi_x^{\mathrm{TR}}(0, 100, 200, 2^{l_1-1}, 5, 5)$ and $\Pi_x^{\mathrm{Unif}}(0, 0.4, 0.8, 2^{l_2-1})$, respectively. Solve the resulting system of ODEs and obtain piecewise linear interpolants on $[0, 200] \times [0, 0.8]$ for each discrete time t_m. (3) Combine them using (4.10). In Fig. 4.1, we specifically show the spatial grids that result from the multilevel discretization with $q = 7$.

		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$
Lawson-Euler [47]	Sup Error	4.500e−1	9.640e−2	3.264e−2	3.245e−2	3.242e−2
	Abs Error	1.157e−1	1.538e−2	2.136e−3	1.853e−3	1.791e−3
	Runtime[s]	20.21	24.97	58.12	120.34	238.51
Nørsett-Euler [22]	Sup Error	4.354e−1	9.113e−2	3.053e−2	3.140e−2	3.190e−2
	Abs Error	1.003e−1	8.260e−3	3.903e−4	6.003e−4	1.166e−3
	Runtime[s]	14.79	24.19	56.82	106.12	212.07
ETD2RK [22]	Sup Error	4.362e−1	9.127e−2	3.253e−2	3.240e−2	3.240e−2
	Abs Error	1.117e−1	1.394e−2	1.872e−3	1.730e−3	1.730e−3
	Runtime[s]	25.10	44.03	106.28	221.38	418.79
ETD3RK [22]	Sup Error	4.362e−1	9.123e−2	3.253e−2	3.240e−2	3.240e−2
	Abs Error	1.119e−1	1.397e−2	1.876e−3	1.731e−3	1.730e−3
	Runtime[s]	34.49	69.38	166.38	359.62	670.06
ETD4RK [22]	Sup Error	4.364e−1	9.126e−2	3.253e−2	3.240e−2	3.240e−2
	Abs Error	1.120e−1	1.397e−2	1.876e−3	1.731e−3	1.730e−3
	Runtime[s]	57.08	117.35	286.77	550.36	1077.41
HochOst4 [39]	Sup Error	4.362e−1	9.123e−2	3.253e−2	3.240e−2	3.240e−2
	Abs Error	1.119e−1	1.397e−2	1.876e−3	1.731e−3	1.730e−3
	Runtime[s]	56.55	110.74	270.06	541.89	1046.20

Table 4.1: Results on numerical solutions of (4.12). Here, we spatially discretize it on $\Pi_x^{\mathrm{TR}} \otimes \Pi_x^{\mathrm{Unif}}$ and solve the resulting system of ODEs. The parameters of Π_x^{TR} are $x_{\mathrm{left}} = 0$, $x_{\mathrm{center}} = 100$, $x_{\mathrm{right}} = 200$, $N_{x,0} = 1000$, and $g_1 = g_2 = 50$. For each N_t, the numerical solution is evaluated on the grid $\Pi_t^{\mathrm{Unif}}(N_t) \times \Pi_x^{\mathrm{TR}} \times \Pi_x^{\mathrm{Unif}}$. Maximum absolute errors in $\Pi_t^{\mathrm{Unif}} \times ([80, 120] \cap \Pi_x^{\mathrm{TR}}) \times ([0.32, 0.48] \cap \Pi_x^{\mathrm{Unif}})$ are reported on the row of "Sup Error", absolute errors at $(t, s_0, v_0) = (0, 100, 0.4)$ are on the row of "Abs Error", and runtimes in seconds are at the bottom line.

Table 4.2 reports the result on the numerical solutions for different q and temporal steps N_t. Here, we used `HochOst4` to solve the resulting systems of ODEs and have evaluated maximum absolute errors of $\mathcal{Y}_t^{t,(s,v)}$ in the same grid as in Section 4.4.1.1, namely, $(t, s, v) \in \Pi_t^{\mathrm{Unif}} \times ([80, 120] \cap \Pi_x^{\mathrm{TR}}) \times ([0.32, 0.48] \cap \Pi_x^{\mathrm{Unif}})$. We observe that the discretization on the sparse grid can provide solutions more efficiently than on the full grid. Table 4.3 presents the total numbers of spatial points the sparse grid method comsumes in Section 4.2 and the numbers of the corresponding full grids. Although the size of our "full grid" is 6231, the sparse grid approach still outperforms in terms of runtime even if the total number of spatial points exceeds this (i.e. $q > 8$).

q		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$	$N_t = 500$
7	Sup Error	4.521e−2	4.521e−2	4.521e−2	4.521e−2	4.488e−2	4.468e−2
	Abs Error	6.921e−3	$6.921e-3$	6.921e−3	6.921e−3	6.921e−3	6.921e−3
	Runtime[s]	1.24	2.38	5.57	10.53	20.22	48.33
8	Sup Error	3.469e−2	3.465e−2	3.465e−2	3.465e−2	3.423e−2	3.398e−2
	Abs Error	2.454e−3	2.419e−3	2.419e−3	2.419e−3	2.419e−3	2.419e−3
	Runtime[s]	2.20	4.08	9.34	17.38	33.07	78.77
9	Sup Error	3.961e−1	8.479e−2	3.244e−2	3.230e−2	3.189e−2	3.164e−2
	Abs Error	1.193e−1	1.498e−2	1.474e−3	1.309e−3	1.308e−3	1.308e−3
	Runtime[s]	3.35	6.18	14.75	28.49	53.03	127.02
10	Sup Error	5.537e+0	3.187e+0	2.353e−1	6.068e−2	3.152e−2	3.100e−2
	Abs Error	4.942e+0	2.854e+0	5.267e−2	6.812e−3	1.351e−3	1.032e−3
	Runtime[s]	6.75	12.74	30.60	59.37	126.54	327.56

Table 4.2: Results on numerical solutions based on the multilevel discretization on sparse grids. The row of "Sup Error" reports the maximum absolute errors on the same points (t, s, v) in Table 4.1 (i.e. $\Pi_t^{\mathrm{Unif}} \times (\Pi_x^{\mathrm{TR}} \cap [80, 120]) \times (\Pi_x^{\mathrm{Unif}} \cap [0.32, 0.48])$.)

q	7	8	9	10
SG	1475	3333	7431	16393
FG	4225	16641	66049	263169

Table 4.3: SG : Total numbers of spatial points required for calculating the numerical solution in Section 4.4.1.2. FG : The size of the corresponding full grid $(2^{q-d+1}+1)^d$.

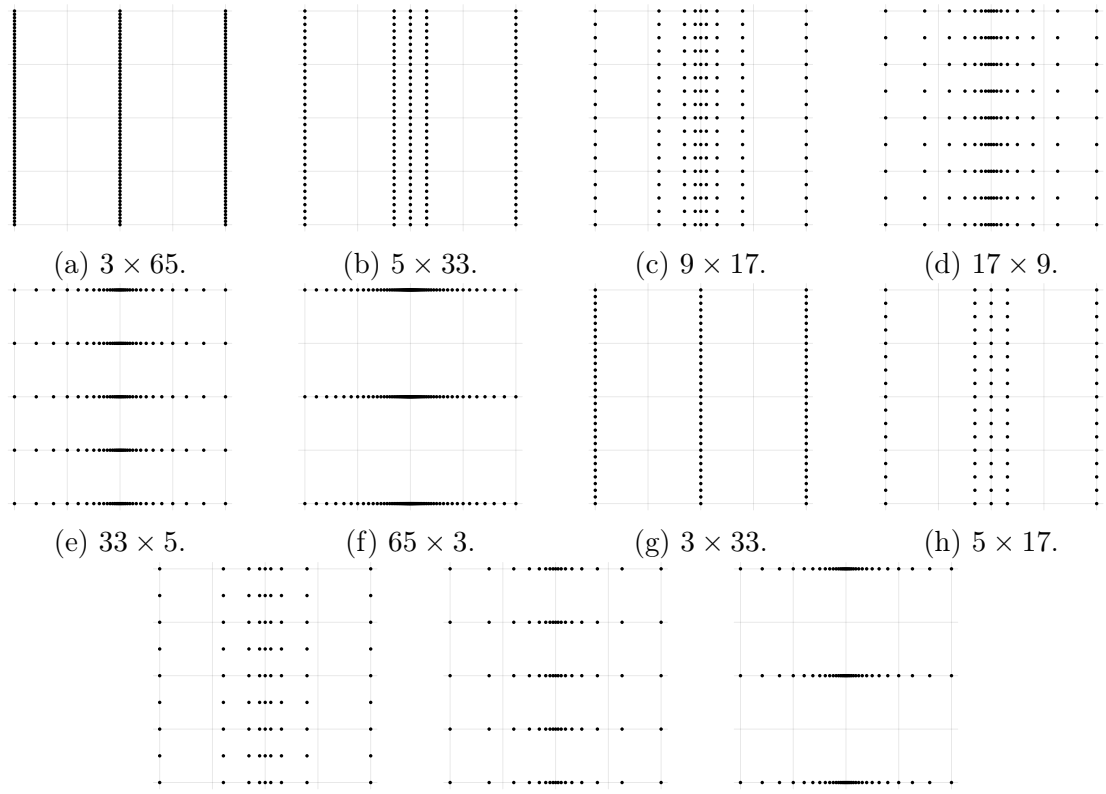
4.4.2 Multi-Asset Option Pricing Using the Multilevel Discretization on Sparse Grids

This section is devoted to solving high-dimensional BSDEs arising from multi-asset option pricing under SLV models. Consider d SLV models, that is, for $i = 1, \dots, d$,

$$\begin{aligned} \mathcal{S}_t^{(i)} &= S_0^{(i)} + \int_0^t \omega^{(i)}(\mathcal{S}_s^{(i)}, v_s^{(i)})ds + \int_0^t m^{(i)}(v_s^{(i)})\Gamma^{(i)}(\mathcal{S}_s^{(i)})dW_s^{(\mathcal{S},i)}, \\ v_t^{(i)} &= v_0^{(i)} + \int_0^t \mu^{(i)}(v_s^{(i)})ds + \int_0^t \sigma^{(i)}(v_s^{(i)})dW_s^{(v,i)}. \end{aligned} \tag{4.14}$$

Here, $W^{(\mathcal{S},i)}$ and $W^{(v,i)}$ are correlated as

$$\langle W^{(\mathcal{S},i)}, W^{(\mathcal{S},j)} \rangle_t = c_{i,j}t, \quad \langle W^{(\mathcal{S},i)}, W^{(v,i)} \rangle_t = \rho_{i,j}t, \quad \langle W^{(v,i)}, W^{(v,i)} \rangle_t = r_{i,j}t.$$



(a) 3×65. (b) 5×33. (c) 9×17. (d) 17×9.

(e) 33×5. (f) 65×3. (g) 3×33. (h) 5×17.

(i) 9×9. (j) 17×5. (k) 33×3.

Figure 4.1: The grids arising from the multilevel discretization for $q = 7$ in the situation of Section 4.4.1.2. (4.12) has been approximated with a combination of 11 BSDEs driven by CTMCs on these grids (a)-(k).

Let $\mathbf{C} = \begin{pmatrix} C_{\mathcal{S}} & C_{\mathcal{S},v} \\ C^*_{\mathcal{S},v} & C_v \end{pmatrix}$ be the correlation matrix, where

$$C_{\mathcal{S}} = \begin{pmatrix} c_{1,1} & \dots & c_{1,d} \\ \vdots & \ddots & \vdots \\ c_{d,1} & \dots & c_{d,d} \end{pmatrix}, \quad C_{\mathcal{S},v} = \begin{pmatrix} \rho_{1,1} & \dots & \rho_{1,d} \\ \vdots & \ddots & \vdots \\ \rho_{d,1} & \dots & \rho_{d,d} \end{pmatrix}, \quad C_v = \begin{pmatrix} r_{1,1} & \dots & r_{1,d} \\ \vdots & \ddots & \vdots \\ r_{d,1} & \dots & r_{d,d} \end{pmatrix}.$$

Using a lower triangular matrix $\mathbf{L}$ constructed from the Cholesky decomposition $\mathbf{C} = \mathbf{L}\mathbf{L}^*$ and $\mathbf{W}$ defined as

$$\mathbf{W} = \mathbf{L}^{-1} \begin{pmatrix} W^{(\mathcal{S},1)} \\ \vdots \\ W^{(\mathcal{S},d)} \\ W^{(v,1)} \\ \vdots \\ W^{(v,d)} \end{pmatrix}$$

is $2d$-dimensional standard Brownian motion.

As in Section 2.5 and Section 3.5, Markov BSDEs describing the price of European options $g(\mathcal{S}_T^{(1)}, \dots, \mathcal{S}_T^{(d)})$ can be formulated by consider the hedge portfolio, which result in

$$\begin{aligned} \mathcal{X}_t &= \mathcal{X}_0 + \int_0^t \boldsymbol{\mu}(\mathcal{X}_s) ds + \int_0^t \boldsymbol{\sigma}(\mathcal{X}_s) d\mathbf{W}_s \quad \text{for} \quad t \in [0,T] \\ \mathcal{Y}_t &= g(\mathcal{S}_T^{(1)}, \dots, \mathcal{S}_T^{(d)}) - \int_t^T f(s, \mathcal{X}_s, \mathcal{Y}_s, \mathcal{Z}_s) ds - \int_t^T \mathcal{Z}_s^* d\mathbf{W}_s, \end{aligned} \tag{4.15}$$

where

$$\mathcal{X}_t = \begin{pmatrix} \mathcal{S}_t^{(1)} \\ \vdots \\ \mathcal{S}_t^{(d)} \\ v_t^{(1)} \\ \vdots \\ v_t^{(d)} \end{pmatrix}, \quad \boldsymbol{\mu}(x) = \begin{pmatrix} \omega^{(1)}(x_1, x_{d+1}) \\ \vdots \\ \omega^{(d)}(x_d, x_{2d}) \\ \mu^{(1)}(x_{d+1}) \\ \vdots \\ \mu^{(d)}(x_{2d}) \end{pmatrix}, \quad \boldsymbol{\sigma}(x) = \operatorname{diag} \begin{pmatrix} m^{(1)}(x_{d+1})\Gamma^{(1)}(x_1) \\ \vdots \\ m^{(d)}(x_{2d})\Gamma^{(d)}(x_{2d}) \\ \sigma^{(1)}(x_{d+1}) \\ \vdots \\ \sigma^{(d)}(x_{2d}) \end{pmatrix} \mathbf{L},$$

the driver is

$$f(t,x,y,z) = r(y - z^* \boldsymbol{\sigma}(x)^{-1} x)^+ - R(y - z^* \boldsymbol{\sigma}(x)^{-1} x)^- + z^* \boldsymbol{\sigma}(x)^{-1} \boldsymbol{\mu}(x),$$

and r and R are the lending rate and the borrowing rate, respectively.

4.4.2.1 Basket Option under Two Heston-SABR Models

Consider pricing of the European basket call option whose basket is comprised of two Heston-SABR models. That is, the coefficient functions in (4.14) are

$$\begin{aligned} \omega^{(i)}(s,v) = b^{(i)} \cdot s, \quad m^{(i)}(v) = \sqrt{v}, \quad \Gamma^{(i)}(s) = s^{\beta^{(i)}}, \\ \mu^{(i)}(v) = \eta^{(i)}(\theta^{(i)} - v), \quad \sigma^{(i)}(v) = \alpha^{(i)}\sqrt{v}. \end{aligned} \tag{4.16}$$

for $i = 1, 2$. The corresponding BSDE is four-dimensional and results in (**??**) with g replaced as

$$g(f_1, f_2) = (\lambda_1 f_1 + \lambda_2 f_2 - K)^+,$$

where λ_1 and λ_2 are constants. Since a non-differentiability of g is appreared in the hyperplane $\{(s^{(1)}, s^{(2)}, v^{(1)}, v^{(2)}) : \lambda_1 s^{(1)} + \lambda_2 s^{(2)} - K = 0\}$, we use a linear coordinate transformation, say

$$\widehat{\mathcal{X}}_t = \begin{pmatrix} \widehat{\mathcal{S}}_t^{(1)} \\ \widehat{\mathcal{S}}_t^{(2)} \\ v_t^{(1)} \\ v_t^{(2)} \end{pmatrix} = B \begin{pmatrix} \mathcal{S}_t^{(1)} \\ \mathcal{S}_t^{(2)} \\ v_t^{(1)} \\ v_t^{(2)} \end{pmatrix}, \quad \text{where} \quad B = \begin{pmatrix} \lambda_1 & \lambda_2 & 0 & 0 \\ -\lambda_1 & \lambda_2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{4.17}$$

before the spatial discretization, that turns out to be

$$\begin{aligned} \widehat{\mathcal{X}}_t &= \widehat{\mathcal{X}}_0 + \int_0^t \mathbf{B}\boldsymbol{\mu}(\mathbf{B}^{-1}\widehat{\mathcal{X}}_s)ds + \int_0^t \mathbf{B}\boldsymbol{\sigma}(\mathbf{B}^{-1}\widehat{\mathcal{X}}_s)d\mathbf{W}_s, \\ \mathcal{Y}_t &= (\widehat{\mathcal{S}}_T^{(1)} - K)^+ - \int_t^T f(s, \mathbf{B}^{-1}\widehat{\mathcal{X}}_s, \mathcal{Y}_s, \mathcal{Z}_s)ds - \int_t^T \mathcal{Z}_s^* d\mathbf{W}_s. \end{aligned} \tag{4.18}$$

The parameters chosen here are:

i	T	K	R	r	$\lambda^{(i)}$	$\beta^{(i)}$	$\eta^{(i)}$	$\theta^{(i)}$	$\alpha^{(i)}$	$b^{(i)}$
1	1.0	100	0.07	0.01	0.5	0.6	0.9	0.02	0.65	0.01
2					0.5	0.07	0.2	0.3	0.3	0.01

The correlation matrices are:

$$C_{\mathcal{S}} = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}, \quad C_{\mathcal{S},v} = \begin{pmatrix} 0.65 & 0.3 \\ -0.1 & 0.05 \end{pmatrix}, \quad C_v = \begin{pmatrix} 1.0 & 0.7 \\ 0.7 & 1.0 \end{pmatrix}.$$

For $(\widehat{s}^{(1)}, \widehat{s}^{(2)}, \widehat{v}^{(1)}, \widehat{v}^{(2)})^* = \mathbf{B}(s^{(1)}, s^{(2)}, v^{(1)}, v^{(2)})^*$, we approximate the spatial domain as $(\widehat{s}^{(1)}, \widehat{s}^{(2)}, \widehat{v}^{(1)}, \widehat{v}^{(2)}) \in [51, 149] \times [-49, 49] \times [0.01, 0.79] \times [0.01, 0.59]$, and apply the multilevel discretization on a sparse grid. We apply a Tavella-Randall grid to the first dimension and the uniform grids to the others. More precisely, using the notation χ_p^l for $p = 1, \ldots, d$ and $l \in \mathbb{N}$ in Section 4.3, we set as follows:

- $\chi_1^l = \Pi_x^{\mathrm{TR}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r, g_1, g_2) = (1, K, 149, 2^{l-1}, 2^{l-1}, 1, 1)$.
- $\chi_2^l = \Pi_x^{\mathrm{Unif}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (-49, 0, 49, 2^{l-1}, 2^{l-1})$.

- $\chi_3^l = \Pi_x^{\mathrm{Unif}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (0.01, 0.4, 0.79, 2^{l-1}, 2^{l-1})$,
- $\chi_4^l = \Pi_x^{\mathrm{Unif}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (0.01, 0.3, 0.59, 2^{l-1}, 2^{l-1})$.

Table 4.6 reports sparse grid solutions $\mathcal{Y}_0^{0,(s^{(1)},s^{(2)},v^{(1)},v^{(2)})}$ at $(s^{(1)}, s^{(2)}, v^{(1)}, v^{(2)}) = (100, 100, 0.4, 0.3)$ calculated using `HochOst4` and their computational times in seconds, for different q and N_t. The numerical solutions seem to converge towards approximately 7.517.

q		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$
8	$\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$	7.50016	7.50013	7.50013	7.50013	7.50013
	Runtime[s]	50.71	57.89	141.68	276.37	552.95
9	$\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$	7.51457	7.51455	7.51455	7.51455	7.51455
	Runtime[s]	119.65	195.74	484.67	931.01	1904.49
10	$\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$	7.51522	7.51656	7.51658	7.51658	7.51658
	Runtime[s]	838.15	1487.51	3376.85	6542.38	12817.34
11	$\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$	7.43057	7.49986	7.51675	7.51713	7.51714
	Runtime[s]	5965.28	10932.85	24935.39	48705.96	97810.97

Table 4.4: Results on numerical solutions $\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$ of (4.20) using a multilevel spatial discretization.

q	8	9	10	11
SG	36901	112105	320675	877655
FG	1185921	17850625	276922881	4362470401

Table 4.5: SG : Total numbers of spatial points of grids comprised of the sparse grid solutions for different q. FG : The size of the corresponding full grid $(2^{q-d+1}+1)^d$.

4.4.3 Down-and-Out Basket Option under Two Heston-SABR Models

Finally, we consider a basket barrier option under two Heston-SABR models. Consider the down-and-out call option where the underlying asset is the basket $\lambda_1 \mathcal{S}_t^{(1)} +$

$\lambda_2 \mathcal{S}_t^{(2)}$. We set

$$\phi(t,x,y) = \begin{cases} 0, & t \in [0,T], \quad \lambda_1 x + \lambda_2 y \le B, \\ (\lambda_1 x + \lambda_2 y - K)^+, & t = T, \quad \lambda_1 x + \lambda_2 y > B. \end{cases}$$

and $\tau = \inf\{t \ge 0 : \lambda_1 \mathcal{S}_t^{(1)} + \lambda_2 \mathcal{S}_t^{(2)} \le B\}$, and the corresponding BSDE is

$$\begin{aligned} \mathcal{X}_t &= \mathcal{X}_0 + \int_0^t \boldsymbol{\mu}(\mathcal{X}_s) ds + \int_0^t \boldsymbol{\sigma}(\mathcal{X}_s) d\mathbf{W}_s \quad \text{for} \quad t \in [0,T] \\ \mathcal{Y}_t &= \phi(T \wedge \tau, \mathcal{S}_{T\wedge\tau}^{(1)}, \mathcal{S}_{T\wedge\tau}^{(2)}) - \int_t^T 1_{\{\tau \le s\}} f(s, \mathcal{X}_s, \mathcal{Y}_s, \mathcal{Z}_s) ds - \int_t^T \mathcal{Z}_s^* d\mathbf{W}_s, \end{aligned} \tag{4.19}$$

where $\boldsymbol{\mu}$, $\boldsymbol{\sigma}$ and f are same as before. The parameters chosen here are:

i	T	K	B	R	r	$\lambda^{(i)}$	$\beta^{(i)}$	$\eta^{(i)}$	$\theta^{(i)}$	$\alpha^{(i)}$	$b^{(i)}$
1	1.0	100	95	0.03	0.001	0.5	0.05	1.0	0.5	0.95	0.04
2						0.5	0.1	0.08	0.03	0.2	0.1

The correlation matrices are:

$$C_{\mathcal{S}} = \begin{pmatrix} 1.0 & 0.5 \\ 0.5 & 1.0 \end{pmatrix}, \quad C_{\mathcal{S},v} = \begin{pmatrix} 0.65 & 0.3 \\ -0.1 & 0.05 \end{pmatrix}, \quad C_v = \begin{pmatrix} 1.0 & 0.7 \\ 0.7 & 1.0 \end{pmatrix}.$$

Applying the linear transformation (4.17) to (4.19), that turns out to be

$$\begin{aligned} \widehat{\mathcal{X}}_t &= \widehat{\mathcal{X}}_0 + \int_0^t \mathbf{B}\boldsymbol{\mu}(\mathbf{B}^{-1}\widehat{\mathcal{X}}_s) ds + \int_0^t \mathbf{B}\boldsymbol{\sigma}(\mathbf{B}^{-1}\widehat{\mathcal{X}}_s) d\mathbf{W}_s, \\ \mathcal{Y}_t &= \widehat{\phi}(T \wedge \tau, \widehat{\mathcal{S}}_{T\wedge\tau}^{(1)}) - \int_t^T 1_{\{s \le \tau\}} f(s, \mathbf{B}^{-1}\widehat{\mathcal{X}}_s, \mathcal{Y}_s, \mathcal{Z}_s) ds - \int_t^T \mathcal{Z}_s^* d\mathbf{W}_s. \end{aligned} \tag{4.20}$$

Here, $\widehat{\phi} : [0,T] \times \mathbb{R} \to \mathbb{R}$ is

$$\widehat{\phi}(t,x) = \begin{cases} 0, & (t,x) \in [0,T] \times [0,B], \\ (x-K)^+, & (t,x) \in \{T\} \times (B, \infty). \end{cases}$$

For (4.20), we discretize on the (multilevel) grids on $(\widehat{s}^{(1)}, \widehat{s}^{(2)}, \widehat{v}^{(1)}, \widehat{v}^{(2)}) \in [B, 150] \times [-49, 49] \times [0.01, 0.79] \times [0.01, 0.59]$, where $(\widehat{s}^{(1)}, \widehat{s}^{(2)}, \widehat{v}^{(1)}, \widehat{v}^{(2)})^* = \mathbf{B}(s^{(1)}, s^{(2)}, v^{(1)}, v^{(2)})^*$. For each subgrid, we employ the Tavella-Randall grid to the first dimension and the uniform grids to the others. More precisely, using the notation χ_p^l for $p = 1, \dots, d$ and $l \in \mathbb{N}$ in Section 4.3, we set as follows:

- $\chi_1^l = \Pi_x^{\mathrm{TR}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (B, K, 150, 2^{l-1}, 2^{l-1})$, $g_2 = 1.0$ and $g_1 = \frac{K-B}{150-K} g_2$,
- $\chi_2^l = \Pi_x^{\mathrm{Unif}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (-49, 0, 49, 2^{l-1}, 2^{l-1})$.
- $\chi_3^l = \Pi_x^{\mathrm{Unif}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (0.01, 0.4, 0.79, 2^{l-1}, 2^{l-1})$,
- $\chi_4^l = \Pi_x^{\mathrm{Unif}}$ with $(x_{\mathrm{left}}, x_{\mathrm{center}}, x_{\mathrm{right}}, N_l, N_r) = (0.01, 0.3, 0.59, 2^{l-1}, 2^{l-1})$.

Under the setting, we calculated the numerical solution $\mathcal{Y}_t^{t,(s^{(1)},s^{(2)},v^{(1)},v^{(2)})}$ using our sparse grid-based multilevel discretization with parameters $q = 8, 9, 10$ and 11. The results are shown in Table 4.6. Here, we solve each ODE using `HochOst4`. The numerical solutions seem to converge towards approximately 2.9554.

q		$N_t = 10$	$N_t = 20$	$N_t = 50$	$N_t = 100$	$N_t = 200$
8	$\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$	2.95842	2.95835	2.95872	2.95873	2.95873
	Runtime[s]	53.02	74.83	183.73	361.43	721.02
9	$\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$	2.95625	2.95554	2.95539	2.95539	2.95539
	Runtime[s]	159.65	251.94	604.26	1188.12	2100.54
10	$\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$	1835.09847	15711.15453	2.95538	2.95539	2.95539
	Runtime[s]	924.82	1555.15	3605.58	6542.38	12817.34
11	$\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$	101.29716	59724.22711	1.9379E+11	2.95543	2.95543
	Runtime[s]	5529.99	9828.60	22814.82	49279.10	95128.83

Table 4.6: Results on numerical solutions $\mathcal{Y}_0^{0,(100,100,0.4,0.3)}$ of (4.20) using a multilevel spatial discretization.

4.5 Conclusion

Motivated by the idea of sparse grid methods, we proposed a multilevel spatial discretization of BSDEs driven by Brownian motion. It approximates the BSDE driven by Brownian motion with a sequence of BSDEs driven by CTMCs on spatial grids with different resolutions, and the solutions are then superimposed along with the sparse grid formula. The method significantly reduces the computational cost while keeping accuracy. As can be seen in the numerical results, it is a promising approach to handle high-dimensional BSDEs driven by Brownian motion efficiently.

Chapter 5

A Numerical Method for Solving High-Dimensional Backward Stochastic Difference Equations Using Sparse Grids

5.1 Introduction

Backward stochastic difference equations (BSΔEs) are discrete-time counterparts of backward stochastic differential equations (BSDEs). Their applications include some kinds of stochastic optimal control problems and dynamic risk measures in mathematical finance. In [18], Cohen and Elliott developed a general theory of BSΔEs, where they suggested that BSΔEs admit unique solutions in a broader degree of generality for the probability distributions of their noise processes compared with BSDEs in continuous time.

In this chapter, we aim to develop a numerical method for a class of BSΔEs with a Markov process in continuous-state space, with a focus specifically on high-dimensional problems. It is a well-known fact that many numerical solutions of BSDEs and BSΔEs suffer from the so-called "curse of dimensionality", which hampers us when computing these solutions in high-dimensional state space. To overcome this drawback, we must appropriately treat the nestings of high-dimensional conditional expectations that appear in these solutions.

We propose a numerical solution using sparse grids. Sparse grids construct multivariate interpolation or quadrature formulae from given univariate ones and utilize the smoothness of the function for computational efficiency. Specifically, the non-

linear functions between conditional expectations are approximated by sparse grid interpolants, and we compute these conditional expectations through sparse grid quadratures. Our method is based on the research of Zhang et al. [74], who proposed a similar method for continuous-time BSDEs driven by Brownian motion.

5.2 Markov BSΔEs

We introduce the class of BSΔEs with a Markov process, namely, Markov BSΔEs. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\mathbb{T} = \{0, 1, \dots, T\}$ be a discrete time set with $T \in \mathbb{N}$. Let $(\xi_t)_{t\in\mathbb{T}\setminus\{0\}}$ be an $\mathbb{R}^d$-valued independent identically distributed sequence with $\mathbb{E}[\xi_1] = 0_d$ and $\mathbb{V}[\xi_1] = I_d$, where 0_d is the d-dimensional zero vector and I_d is the $d \times d$ identity matrix. Let $\mathbb{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ be a filtration generated by $(\xi_t)_{t\in\mathbb{T}\setminus\{0\}}$, and set $\mathcal{F}_0 = \{\emptyset, \Omega\}$. For $p \geq 1$ and $K \geq 1$, let $L^p(\mathbb{R}^K; \mathcal{F}_t)$ denote the set of $\mathbb{R}^K$-valued $\mathcal{F}_t$-measurable p-integrable random variables. We define a Markov process as

$$X_0 = x_0 \in \mathbb{R}^n, \quad X_t = f(t-1, X_{t-1}, \xi_t) \text{ for } t \in \mathbb{T} \setminus \{0\},$$

where $f : (\mathbb{T} \setminus \{T\}) \times \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$ is a function such that $f(t, \cdot, \cdot)$ is Borel measurable for each $t \in \mathbb{T} \setminus \{T\}$. In addition, we define the difference operator Δ as $\Delta U_t \coloneqq U_t - U_{t-1}$ for a process $(U_t)_{t=0}^T$.

We consider the following Markov BSΔE:

$$\begin{aligned} Y_T &= h(X_T), \\ -\Delta Y_t &= g(t-1, X_{t-1}, Y_{t-1}, Z_t) - Z_t^* \xi_t - \Delta M_t \end{aligned} \tag{5.1}$$

for $t = T, \dots, 1$, where $g : (\mathbb{T} \setminus \{T\}) \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ such that $g(t, \cdot, \cdot, \cdot)$ is Borel measurable for each $t \in \mathbb{T} \setminus \{T\}$, $h : \mathbb{R}^n \to \mathbb{R}$ is Borel measurable, and $(\cdot)^*$ is vector transposition. Here, a solution of the Markov BSΔE is a triplet $(Y, Z, M) = ((Y_t)_{t\in\mathbb{T}}, (Z_t)_{t\in\mathbb{T}\setminus\{0\}}, (M_t)_{t\in\mathbb{T}})$ that satisfies (5.1), such that Y is an $\mathbb{R}$-valued square-integrable adapted process, Z is an $\mathbb{R}^d$-valued square-integrable predictable process, and M is an $\mathbb{R}$-valued square-integrable martingale, such that $M_0 = 0$ and $\mathbb{E}[(\Delta M_t)\xi_t | \mathcal{F}_{t-1}] = 0$ for all t. We remark that since our noise process $(\xi_t)_{t\in\mathbb{T}\setminus\{0\}}$ does not have the predictable representation property, the BSΔEs need to be formulated with the additional term M.

Assumption 5.2.1. *We assume the following conditions.*

1. *For any* $(t, A_t, B_t) \in \mathbb{T} \times L^2(\mathbb{R}; \mathcal{F}_t) \times L^2(\mathbb{R}^d; \mathcal{F}_t)$

$$g(t, X_t, A_t, B_t) \in L^2(\mathbb{R}; \mathcal{F}_t).$$

2. For any $(t, x, z) \in (\mathbb{T} \setminus \{T\}) \times \mathbb{R}^n \times \mathbb{R}^d$*, the function*

$$\Phi(\cdot; t, x, z) : \mathbb{R} \ni y \mapsto y - g(t, x, y, z) \in \mathbb{R}$$

is bijective, and for any $(t, A_t, B_t) \in (\mathbb{T} \setminus \{T\}) \times L^2(\mathbb{R}; \mathcal{F}_t) \times L^2(\mathbb{R}^d; \mathcal{F}_t)$*,*

$$\Phi^{-1}(A_t; t, X_t, B_t) \in L^2(\mathbb{R}; \mathcal{F}_t).$$

3. $h(X_T) \in L^2(\mathbb{R}; \mathcal{F}_T)$*.*

Theorem 5.2.1. *Under Assumption 1, the BSΔE* (5.1) *admits a unique solution.*

Proof. The solution can be constructed using the backward induction in time. Clearly, $Y_T = h(X_T)$ is the solution at time T. For $t = T-1, T-2, \ldots, 0$, we set

$$\begin{aligned} Z_{t+1} &= \mathbb{E}[\xi_{t+1} Y_{t+1} | \mathcal{F}_t], \\ \Delta M_{t+1} &= Y_{t+1} - \mathbb{E}[Y_{t+1} | \mathcal{F}_t] - Z_{t+1}^* \xi_{t+1}, \\ Y_t &= \Phi^{-1}(\mathbb{E}[Y_{t+1} | \mathcal{F}_t]; t, X_t, Z_{t+1}). \end{aligned}$$

The uniqueness and L^2-integrability of the solution are clear. □

We define the following continuous linear operators:

$$\begin{aligned} P_{t-1,t} &: \mathcal{B} \ni \phi \mapsto \mathbb{E}[\phi(f(t-1, \cdot, \xi_t))] \in \mathcal{B}, \\ Q_{t-1,t} &: \mathcal{B} \ni \phi \mapsto \mathbb{E}[\phi(f(t-1, \cdot, \xi_t))\xi_t] \in \mathcal{B}^d, \end{aligned}$$

where $\mathcal{B}$ is the space of bounded measurable functions defined on $\mathbb{R}^n$. We impose the following assumption.

Assumption 5.2.2. *1.* h *and* $\Phi^{-1}(y; t, \cdot, z)$ *belong to* $\mathcal{B}$ *for all* y*,* t*, and* z*.*

2. For all t *and* x*,* $\Phi^{-1}(\cdot; t, x, \cdot)$ *is continuous.*

Under Assumption 5.2.2, we define the nonlinear operator $\rho_{t-1,t} : \mathcal{B} \to \mathcal{B}$ by

$$(\rho_{t-1,t}\phi)(x) = \Phi^{-1}((P_{t-1,t}\phi)(x); t-1, x, (Q_{t-1,t}\phi)(x)).$$

Corollary 5.2.1. *The unique solution of BSΔE* (5.1) *can be expressed in the form of*

$$Y_t = \mathcal{Y}_t(X_t), \quad Z_t = \mathcal{Z}_t(X_{t-1}), \quad \Delta M_t = \mathcal{M}_t(X_{t-1}, X_t, \xi_t). \tag{5.2}$$

Here, $\mathcal{Y}_t : \mathbb{R}^n \to \mathbb{R}$*,* $\mathcal{Z}_t : \mathbb{R}^n \to \mathbb{R}^d$*, and* $\mathcal{M}_t : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ *are given by*

$$\begin{aligned} \mathcal{Y}_t(x) &\coloneqq (\rho_{t,t+1} \circ \cdots \circ \rho_{T-1,T}) h(x), \\ \mathcal{Z}_t(x) &\coloneqq Q_{t-1,t} \mathcal{Y}_t(x), \\ \mathcal{M}_t(x, y, e) &\coloneqq \mathcal{Y}_t(y) - P_{t-1,t}\mathcal{Y}_t(x) - Q_{t-1,t}\mathcal{Y}_t(x)^* e. \end{aligned} \tag{5.3}$$

Hence, we can reformulate our problem of solving the BSΔE into an evaluation of the functions $\mathcal{Y}_t$, $\mathcal{Z}_t$, and $\mathcal{M}_t$.

5.3 Computing the Numerical Solution

We are interested in computing solution (5.2) in the situation where the dimension n of the state process X and the dimension d of the noise process ξ are rather large (e.g. $3 \leq n, d \leq 10$). We assume that the random variables $\xi_1^1, \ldots, \xi_1^d$ are mutually independent and that each ξ_1^i has the probability density g_i, where we write $\xi_1 = (\xi_1^1, \ldots, \xi_1^d)$. Defining the linear functional

$$I\phi \coloneqq \int_{\mathbb{R}^d} \phi(w) g_\xi(w) dw_1 \cdots dw_d,$$

where $g_\xi(w) = \prod_{i=1}^d g_i(w_i)$ and $w = (w_1, \ldots, w_d)$, $P_{t-1,t}$ and $Q_{t-1,t}$ can then be expressed as d-dimensional integrals with the density g_ξ:

$$\begin{aligned}(P_{t-1,t}\phi)(x) &= I(\phi \circ f(t-1, x, \cdot)),\\ (Q_{t-1,t}\phi)(x) &= I(\phi \circ f(t-1, x, \cdot)(\cdot)).\end{aligned}$$

Approximating I as some quadrature formula $\widehat{I}$, we obtain the approximations $\widehat{P}_{t-1,t}$ and $\widehat{Q}_{t-1,t}$ of $P_{t-1,t}$ and $Q_{t-1,t}$, respectively. Using these, we also define

$$(\widehat{\rho}_{u,u+1}\phi)(x) = \Phi^{-1}((\widehat{P}_{u,u+1}\phi)(x); u, x, (\widehat{Q}_{u,u+1}\phi)(x)).$$

We now face computational difficulties at two points. First, all the integrations we compute have dimension d. A straightforward way to obtain d-variate quadrature formulae is to use the tensor product construction of univariate quadratures, but its computational cost increases exponentially with d. Second, because of the nestings of integrations in (5.3), the computational cost of the solution increases exponentially with T.

For the first point, we take sparse grid quadrature formulae that can alleviate these costs by using the smoothness of integrands [42]. For the second point, we compute the numerical solution at time t as $\widehat{\mathcal{Y}}_t(x) \coloneqq \widehat{\rho}_{t,t+1}\widetilde{\mathcal{Y}}_{t+1}(x)$, using a sparse grid interpolant $\widetilde{\mathcal{Y}}_{t+1}$ of $\widehat{\mathcal{Y}}_{t+1}$. A detailed explanation of sparse grid formulae is given in Section 4.

5.3.1 Prototype of Our Scheme

Our proposed scheme for finding the numerical solution $\widehat{\mathcal{Y}}_t(x)$ for $x \in \mathbb{R}^n$ is as follows.

- Initialize $\widetilde{\mathcal{Y}}_T = h$.

- For $u = T-1, T-2, \ldots, t+1$, given $\widetilde{\mathcal{Y}}_{u+1}$ at time $u+1$, compute an interpolant $\widetilde{\mathcal{Y}}_u$ of $\widehat{\mathcal{Y}}_u = \widehat{\rho}_{u,u+1}\widetilde{\mathcal{Y}}_{u+1}$.
- Compute $\widehat{\mathcal{Y}}_t(x) = \widehat{\rho}_{t,t+1}\widetilde{\mathcal{Y}}_{t+1}(x)$.

Here, $\widehat{\mathcal{Z}}_t$ and $\widehat{\mathcal{M}}_t$ can be computed using $\widetilde{\mathcal{Y}}_t$ as

$$\begin{aligned}\widehat{\mathcal{Z}}_t(x) &= \widehat{Q}_{t-1,t}\widetilde{\mathcal{Y}}_t(x),\\ \widehat{\mathcal{M}}_t(x,y,e) &= \widetilde{\mathcal{Y}}_t(x) - \widehat{P}_{t-1,t}\widetilde{\mathcal{Y}}_t(x) - \widehat{Q}_{t-1,t}\widetilde{\mathcal{Y}}_t(x)^* e.\end{aligned}$$

5.3.2 Truncation of State Space

The bounded spatial domain $E_t \subset \mathbb{R}^n$ of each interpolant $\widetilde{\mathcal{Y}}_t$ needs to be determined. We design them as follows. If, for example, we would like to compute $\widehat{\mathcal{Y}}_0(0)$, then we evaluate the interpolant $\widetilde{\mathcal{Y}}_1$ on $\{f(0,0,\eta_i)\}_{i=1}^{N_Q}$, where $\eta_1, \ldots, \eta_{N_Q}$ are the quadrature points of $\widehat{I}$. Hence, we set E_1 such that $\{f(0,0,\eta_i)\}_{i=1}^{N_Q}$ lies in E_1. Similarly, for any $t = 0, \ldots, T-1$, since the interpolant $\widetilde{\mathcal{Y}}_t$ is constructed with the N_t-point set $\{\widehat{\mathcal{Y}}_t(x_j^t)\}_{j=1}^{N_t}$, and since we evaluate $\widetilde{\mathcal{Y}}_{t+1}$ on $\{f(t,x_j^t,\eta_i)\}_{i=1}^{N_Q}$ for the computation of $\widehat{\mathcal{Y}}_t(x_j^t)$ for $j = 1, \ldots, N_t$, we set E_{t+1} such that $\bigcup_{j=1}^{N_t}\{f(t,x_j^t,\eta_i)\}_{i=1}^{N_Q}$ is a subset of E_{t+1}.

5.4 Sparse Grids

5.4.1 Sparse Grid Interpolation

For a univariate smooth function $f : [-1,1] \to \mathbb{R}$, we consider a sequence of interpolation formulae $(\mathcal{U}^i)_{i=1}^{\infty}$,

$$\mathcal{U}^i(f) = \sum_{j=1}^{m_i} f(x_j^i)\cdot a_j^i.$$

Here, m_i points $x_j^i \in \mathbb{R}$ and basis functions a_j^i satisfy

$$\lim_{i\to\infty} \|\mathcal{U}^i(f) - f\|_\infty = 0,$$

where $\|\cdot\|_\infty$ denotes the supremum norm. If we define the difference of algorithms as

$$\mathcal{U}^0(f) = 0, \quad \Delta^i = \mathcal{U}^i(f) - \mathcal{U}^{i-1}(f) \text{ for } i \in \mathbb{N},$$

then the sparse grid interpolant for a d-variate smooth function f, defined on $[-1,1]^d$, can be defined as

$$\mathcal{A}^{q,d}(f) = \sum_{\substack{i\in\mathbb{N}^d \\ |i|\leq q}} (\Delta^{i_1} \otimes \cdots \otimes \Delta^{i_d})(f) \tag{5.4}$$

for any $q \geq d$, $|i| \coloneqq i_1 + \cdots + i_d$. We also remark that more interpolating points are used by $\mathcal{A}^{q,d}$ the larger we take q to be.

5.4.1.1 Sparse Grid Piecewise Linear Interpolation

Throughout this chapter, we adopt the sparse grid piecewise linear interpolation as our interpolant. This is based on the following piecewise linear basis functions:

$$a_j^i(x) = \begin{cases} 1 - \frac{m_i - 1}{2}|x - x_j^i|, & |x - x_j^i| < \frac{2}{m_i - 1}, \\ 0, & \text{otherwise}, \end{cases}$$

for $i > 1, j = 1, \ldots, m_i$, and $a_1^1(x) = 1$, where

$$m_1 = 1, \quad m_i = 2^{i-1} + 1, \text{ for } i > 1,$$

and x_j^i are the Newton-Cotes equidistant points given by

$$x_1^1 = 0, \quad x_j^i = \frac{2(j-1)}{m_i - 1} - 1, \quad j = 1, \ldots, m_i, \quad i > 1.$$

For our error analysis, we define the function class

$$F_d^r = \{ f : [-1,1]^d \to \mathbb{R} \mid \mathcal{D}^\alpha f \text{ is continuous if } \alpha_i \leq r \text{ for all i} \}$$

with smoothness r, where $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0$, $|\alpha| = \alpha_1 + \cdots + \alpha_d$, and $\mathcal{D}^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}$. Clearly,

$$\|f - \mathcal{U}^q f\|_\infty \leq c_{1,2} m_q^{-2} \text{ for } f \in F_1^2,$$

where $c_{1,2}$ depends on the upper bound of the second derivative of f. For $d > 1$, the error of the sparse grid piecewise linear interpolation is given as [2]

$$\|f - \mathcal{A}^{q,d} f\|_\infty \leq c_{d,2} N^{-2} (\log N)^{3(d-1)} \tag{5.5}$$

for $f \in F_d^2$, where N is the number of interpolating points and $c_{d,2}$ depends on d and the upper bound of the second derivative of f.

We remark that the sparse grid interpolant based on polynomial interpolation is also known, and its error is

$$\|f - \mathcal{A}^{q,d} f\|_\infty \leq c_{d,r} N^{-r} (\log N)^{(r+1)(d-1)}$$

for $f \in F_d^r$, where $c_{d,r}$ depends on d, r, and the upper bound of the r-th derivative of f, which is superior to (5.5) when the smoothness r of f is larger than 2. However, it requires the evaluation of Lagrange polynomials, which is computationally expensive, and causes the Runge phenomenon, which prevents us from using the Newton-Cotes grid. Hence, we use the sparse grid piecewise linear interpolation because of its ease of implementation and its low computational complexity.

5.4.2 Sparse Grid Quadrature

Analogously, we can construct the sparse grid quadrature from univariate quadrature formulae. The main difference is that each a_j^i is a real number instead of a basis function, and $(\mathcal{U}^i)_{i=1}^\infty$ is a sequence of quadratures for the function $f : D \to \mathbb{R}$ satisfying

$$\lim_{i \to \infty} |\mathcal{U}^i(f) - If| = 0,$$

where If is the integral of f on $D \subset \mathbb{R}$ with a weight w:

$$If \coloneqq \int_D f(x) w(x) dx.$$

5.4.2.1 Sparse Grid Gauss-Hermite Quadrature

The Gauss-Hermite quadrature based on Hermite polynomials computes If over $D = \mathbb{R}$ with $w(x) = e^{-x^2}$. According to Theorem 2 in [65] and Theorem 3.9 in [30], the error of this quadrature, $\mathcal{U}^q$, is

$$|If - \mathcal{U}^q f| \leq c_{1,r} m_q^{-r/2}$$

for $f \in \widetilde{F_1}^r$. Here, $c_{1,r}$ depends on r and the upper bound of the r-th derivative of f, and

$$\widetilde{F_d}^r = \{ f : \mathbb{R}^d \to \mathbb{R} \big| \, \mathcal{D}^\alpha f \text{ is continuous and } \|(\mathcal{D}^\alpha f) W_d\|_\infty < \infty \text{ if } \alpha_i \leq r \text{ for all } i \},$$

where $W_d(x) = \prod_{i=1}^{d} \sqrt{1+x_i^2}\exp(-\frac{x^*x}{2})$. For $d > 1$, using the discussion in [2, 57], we obtain

$$|If - \mathcal{A}^{q,d}f| \leq c_{d,r} N^{-\frac{r}{2}} (\log N)^{(\frac{r}{2}+1)(d-1)}$$

for $f \in \widetilde{F_d}^r$, where $c_{d,r}$ depends on d, r and the upper bound of the r-th derivative of f.

5.5 Error Analysis

Our numerical solutions are computed as discussed in Section 3.1. Recall that we use the sparse grid piecewise linear interpolation $\mathcal{A}^{q_t,n}$ on E_t with parameter q_t at each time t. We then choose the sparse grid quadrature $\mathcal{A}^{q_Q,d}$ with parameter q_Q depending on the noise process of the BSΔE at hand.

5.5.1 The Gaussian Case

Assuming $(\xi_t)_{t=1}^T$ are Gaussian distributed, we choose the sparse grid Gauss-Hermite quadrature formula.

Assumption 5.5.1. *h, $\Phi^{-1}(\cdot;\cdot,t,\cdot)$, and $f(t,\cdot,\cdot)$ have bounded derivatives up to order r with respect to all variables.*

Theorem 5.5.1. *Let $(x_i^t)_{i=1}^{N_t}$ be interpolating points of $\mathcal{A}^{q_t,n}$ for $t \in \mathbb{T}$ and let $(\eta_k)_{k=1}^{N_Q}$ be quadrature points of $\mathcal{A}^{q_Q,d}$. Under Assumption 3, we obtain*

$$\begin{aligned}&\max_{0\leq t\leq T, i=1,\dots,N_t} |\mathcal{Y}_t(x_i^t) - \widehat{\mathcal{Y}}_t(x_i^t)| \\ &\qquad \leq c_{r,n,d}\left\{\max_{1\leq t\leq T}\left[(N_t)^{-2}(\log N_t)^{3(n-1)}\right] + (N_Q)^{-\frac{r}{2}}(\log N_Q)^{(\frac{r}{2}+1)(d-1)}\right\}\sum_{u=1}^{T} L^u,\end{aligned}$$

where L is the Lipschitz coefficient for $\Phi^{-1}(\cdot;\cdot,t,\cdot)$ and $c_{r,n,d}$ depends on r, n, and d. For $\widehat{\mathcal{Z}}_t$ and $\widehat{\mathcal{M}}_t$, similarly,

$$\max_{1\leq t\leq T, i=1,\dots,N_{t-1}} \|\mathcal{Z}_t(x_i^{t-1}) - \widehat{\mathcal{Z}}_t(x_i^{t-1})\| \leq c_{r,n,d}(N_Q)^{-\frac{r}{2}}(\log N_Q)^{(\frac{r}{2}+1)(d-1)},$$

Table 5.1: Relative errors (RE) and computational times (CT) in seconds for our experiment.

	$q_Q = 4$		$q_Q = 5$	
	RE	CT[s]	RE	CT[s]
$q_t = 5$	1.00	0.0	1.50	0.2
$q_t = 6$	1.00	0.5	3.80×10^{-1}	5.0
$q_t = 7$	1.00	8.5	9.24×10^{-2}	108.2
$q_t = 8$	1.00	123.5	2.31×10^{-2}	1538.7
$q_t = 9$	1.00	1341.9	5.77×10^{-3}	16678.3

where $\|\cdot\|$ *denotes the Euclidean norm, and*

$$\max_{\substack{1\le t\le T, i=1,\dots,N_{t-1}\\ j=1,\dots,N_t, k=1,\dots,N_Q}} |\mathcal{M}_t(x_i^{t-1}, x_j^t, \eta_k) - \widehat{\mathcal{M}}_t(x_i^{t-1}, x_j^t, \eta_k)|$$
$$\le c_{r,n,d}\left\{\max_{1\le t\le T}\left[(N_t)^{-2}(\log N_t)^{3(n-1)}\right] + (N_Q)^{-\frac{r}{2}}(\log N_Q)^{(\frac{r}{2}+1)(d-1)}\right\}\sum_{u=1}^{T} L^u$$
$$+ c_{r,d}(N_Q)^{\frac{1}{2}-\frac{r}{2}}(\log N_Q)^{(\frac{r}{2}+1)(d-1)}.$$

5.6 Numerical Results

Let $\xi_1, \dots, \xi_T$ be independent random variables where $\xi_1 \sim N(0_d, I_d)$, and let $(X_t)_{t\in\mathbb{T}}$ be defined as $X_{t+1} \coloneqq X_t + \xi_{t+1}$ with $X_0 = x_0 \in \mathbb{R}^d$. That is, we take a d-dimensional random walk as our Markov process. For the driver g and the terminal condition h, we consider

$$g(x) = \tilde{g}(a^*x), \quad h(x) = \tilde{h}(b^*x),$$

where $\tilde{g}$ and $\tilde{h}$ are functions defined on $\mathbb{R}$, and a and b are d-dimensional real valued vectors. We test our algorithm on the following form of BSΔEs:

$$Y_T = h(X_T), \quad -\Delta Y_t = g(X_{t-1}) + Z_t^*\xi_t + \Delta M_t.$$

The solution $Y_0(x_0)$ has the closed form

$$Y_0(x_0) = \tilde{g}(a^*x_0) + \sum_{j=1}^{T-1}\mathbb{E}[\tilde{g}(a^*x_0 + \sqrt{ja^*\Sigma a}Z)] + \mathbb{E}[\tilde{h}(a^*x_0 + \sqrt{Tb^*\Sigma b}Z)], \quad (5.6)$$

where $Z \sim N(0,1)$. We then compare the true solution calculated by (5.6) and our numerical solution. We use the sparse grid Gauss-Hermite quadrature $\mathcal{A}^{q_Q,d}$ with the parameter $q_Q \geq d$. Also, we set $q_1 = q_2 = \cdots = q_T$ for simplicity. We take $d = 4, T = 5$, $\tilde{g}(x) = \tilde{h}(x) = x^2$, $a = b = (1, \ldots, 1)^*$, and $x_0 = 0_d$. All of our experiments were performed on a 3.70 GHz, 64-GB RAM Linux workstation. Our code was written entirely in python, using numpy and scipy. The results are shown in Table 1. Taking sufficiently large q_Q and q_t (e.g. $(q_Q, q_t) = (5, 9)$) produces a highly accurate solution. On the other hand, we observe that increasing q_t does not improve the performance for $q_Q = 4$. That means that our algorithm had already achieved the best performance under our inaccurate quadrature. Also, when q_t equals 5, increasing q_Q worsens the overall performance because of the accumulation of errors that come from inaccurate interpolations. These observations suggest that increasing both parameters simultaneously is required for efficiently obtaining highly accurate solutions.

Chapter 6

Conclusion

Throughout this thesis, numerical methods for BSDEs with discrete features were explored.

Chapters 2 and 3 were concerned with multi-stage Euler-Maruyama methods for solving Markov BSDEs driven by CTMCs; the case where the terminal time is deterministic was considered in Chapter 2, and the case where the terminal time is a bounded stopping time was considered in Chapter 3. A key observation is that the methods are equivalent to exponential integrators, which are known to work well for solving stiff systems of ODEs. There, we further proposed to apply the methods to solving BSDEs driven by Brownian motion. Employing a suitable spatial discretization for such BSDEs typically leads to "stiff" BSDEs driven by CTMCs, and we can effectively solve them using the Euler-Maruyama methods.

Chapter 4 developed the spatial discretization that had been considered in the previous chapters. Classical schemes of discretizing high-dimensional BSDEs result in BSDEs driven by CTMCs whose state space is too large to calculate numerically. Focusing on such computational issues, we proposed a multilevel spatial discretization method for BSDEs driven by Brownian motion. Employing the idea of sparse grids, it approximates the solution with a superposition of the solutions of BSDEs driven by CTMCs on grids with different resolutions. This construction reduces computational costs drastically and overcomes the curse-of-dimensionality.

In Chapter 5, we employed the sparse grid methods for solving Markov BSΔEs, discrete-time counterparts of BSDEs. The nestings of conditional expectations that appeared in the solution of BSΔEs at each time are approximated with the sparse grid interpolants, and the conditional expectations are evaluated using the sparse grid quadratures. The presented method calculates the solutions of high-dimensional BSΔEs with less computational cost while keeping accuracy.

Future Work

Various directions of further research from this work can be considered.

- In Chapter 2, we pointed out that it is not very effective to employ the Monte-Carlo approach naively. However, in [1], the author presented a Monte-Carlo approach to evaluate an action of a matrix exponential on a vector and observed that it achieves higher performance than the Krylov subspace approach for large scale problems. It might be possible to develop the technique to explore efficient Monte-Carlo-based methods for solving BSDEs driven by CTMCs.

- We proposed to spatially discretize BSDEs driven by Brownian motion and apply the multi-stage Euler-Maruyama methods to the obtained BSDEs driven by CTMCs. For this approach, we can consider various extensions and generalizations such as (1) employing different discretization schemes (e.g. finite element methods or finite volume methods) or (2) considering the case where the driving process of BSDEs is not only a diffusion process (e.g. Lévy processes).

- The multilevel spatial discretization method, developed in Chapter 4, might be further extended. For example, the sparse grid technique, on which the method is based, overcomes the curse of dimensionality to some extent, but not completely. To treat problems with higher dimensions than ten-dimension, dimension reduction techniques such as principal component analysis [64] need to be introduced additionally.

- For Chapter 5, the following directions of further research remains: (i) Extending our results to other cases, for example, where the noise process is not necessarily independent. (ii) Conducting further numerical experiments for the equations that appeared in specific problems, such as those involving dynamic risk measurement and stochastic control problems. (iii) Performing investigations on some computational aspects, such as tuning parameters and selecting the underlying quadrature and interpolation rules.

Acknowledgements

It has been a long journey to complete my PhD research, and many people have kindly supported me.

I would like to express utmost gratitude towards my supervisor Professor Jun Sekine, whose thoughtful and insightful advice helps my study a lot. He constantly encouraged me to explore creative thinking, and his guidance, supports and feedbacks always helped me to grow into an independent researcher. It was a great pleasure for me to conduct research under his supervision.

I am profoundly indebted to Professors Masaaki Fukasawa and Yuko Yano. Despite their busy schedule, they spared their valuable time for examining this thesis as co-examiners.

A debt of gratitude is also owed to Professor Yushi Hamaguchi. His invaluable discussion and encouragement are very helpful in building a foundation for my PhD research.

I would like to express my thanks to Professors Hidehiro Kaise, Nobuaki Naganuma, Dai Taguchi, Masato Hoshino, Yuki Hirai and all the other members of Fukasawa-Sekine-Yano laboratory for their kind support.

This thesis was supported by JST SPRING, Grant Number JPMJSP2138.

Bibliography

[1] Acebrón, J. A. (2019). A monte carlo method for computing the action of a matrix exponential on a vector. *Applied Mathematics and Computation*, 362:124545.

[2] Barthelmann, V., Novak, E., and Ritter, K. (2000). High dimensional polynomial interpolation on sparse grids. *Advances in Computational Mathematics*, 12:273–288.

[3] Bender, C. and Steiner, J. (2012). Least-squares monte carlo for backward sdes. In Carmona, R. A., Del Moral, P., Hu, P., and Oudjane, N., editors, *Numerical Methods in Finance*, pages 257–289, Berlin, Heidelberg. Springer Berlin Heidelberg.

[4] Berland, H., Skaflestad, B., and Wright, W. M. (2007). Expint—a matlab package for exponential integrators. *ACM Trans. Math. Softw.*, 33(1):#4.

[5] Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98.

[6] Bielecki, T. R., Cialenco, I., and Chen, T. (2015). Dynamic conic finance via backward stochastic difference equations. *SIAM Journal on Financial Mathematics*, 6(1):1068–1122.

[7] Bismut, J.-M. (1976). Linear quadratic optimal stochastic control with random coefficients. *SIAM Journal on Control and Optimization*, 14(3):419–444.

[8] Bismut, J.-M. (1978). An introductory approach to duality in optimal stochastic control. *SIAM Review*, 20(1):62–78.

[9] Bodeau, J., Riboulet, G., and Roncalli, T. (2000). Non-uniform grids for pde in finance. *Available at SSRN 1031941*.

[10] Bou-Rabee, N. and Vanden-Eijnden, E. (2018). *Continuous-time random walks for the numerical solution of stochastic differential equations*. Memoirs of the American Mathematical Society. American Mathematical Society, Providence, RI.

[11] Bungartz, H.-J. and Griebel, M. (2004). Sparse grids. *Acta Numerica*, 13:147–69.

[12] Butcher, J. C. (2003). *Numerical Methods for Ordinary Differential Equations*. Wiley.

[13] Christ, S., Schwabeneder, D., Rackauckas, C., Borregaard, M. K., and Breloff, T. (2023). Plots.jl - a user extendable plotting api for the julia programming language. *Journal of Open Research Software*.

[14] Cohen, S. N. (2014). Undiscounted markov chain bsdes to stopping times. *Journal of Applied Probability*, 51(1):262–281.

[15] Cohen, S. N. and Elliott, R. J. (2008). Solutions of backward stochastic differential equations on markov chains. *Communications on Stochastic Analysis*, 2(2):5.

[16] Cohen, S. N. and Elliott, R. J. (2010a). Comparisons for backward stochastic differential equations on markov chains and related no-arbitrage conditions. *The Annals of Applied Probability*, 20(1):267–311.

[17] Cohen, S. N. and Elliott, R. J. (2010b). A general theory of finite state backward stochastic difference equations. *Stochastic Processes and their Applications*, 120(4):442–466.

[18] Cohen, S. N. and Elliott, R. J. (2011). Backward stochastic difference equations and nearly time-consistent nonlinear expectations. *SIAM Journal on Control and Optimization*, 49(1):125–139.

[19] Cohen, S. N. and Elliott, R. J. (2015). *Stochastic Calculus and Applications*. Probability and Its Applications. Springer New York.

[20] Cohen, S. N. and Hu, Y. (2013). Ergodic bsdes driven by markov chains. *SIAM Journal on Control and Optimization*, 51(5):4138–4168.

[21] Cohen, S. N. and Szpruch, L. (2012). On markovian solutions to markov chain bsdes. *Numerical Algebra, Control & Optimization*, 2(2):257.

[22] Cox, S. M. and Matthews, P. C. (2002). Exponential time differencing for stiff systems. *Journal of Computational Physics*, 176(2):430–455.

[23] Crandall, M. G., Ishii, H., and Lions, P.-L. (1992). User' s guide to viscosity solutions of second order partial differential equations. *Bulletin of the American mathematical society*, 27(1):1–67.

[24] Cui, Z., Kirkby, J. L., and Nguyen, D. (2018). A general valuation framework for sabr and stochastic local volatility models. *SIAM Journal on Financial Mathematics*, 9(2):520–563.

[25] Cui, Z., Lars Kirkby, J., and Nguyen, D. (2019). *Continuous-Time Markov Chain and Regime Switching Approximations with Applications to Options Pricing*, pages 115–146. Springer International Publishing, Cham.

[26] Djehiche, B. and Löfdahl, B. (2016). Nonlinear reserving in life insurance: Aggregation and mean-field approximation. *Insurance: Mathematics and Economics*, 69:1–13.

[27] Douglas, J., Ma, J., and Protter, P. (1996). Numerical methods for forward-backward stochastic differential equations. *The Annals of Applied Probability*, 6(3):940–968.

[28] Elliott, R. J., Aggoun, L., and Moore, J. B. (1995). *Hidden Markov Models: Estimation and Control*. Applications of mathematics. Springer.

[29] Föllmer, H. and Schied, A. (2011). *Stochastic finance: an introduction in discrete time*. Walter de Gruyter.

[30] Freud, G. (1972). *A Contribution to the Problem of Weighted Polynomial Approximation*, pages 431–447. Birkhäuser Basel, Basel.

[31] Garcke, J., Griebel, M., and Thess, M. (2001). Data mining with sparse grids. *Computing*, 67:225–253.

[32] Gobet, E., Lemor, J.-P., and Warin, X. (2005). A regression-based Monte Carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability*, 15(3):2172–2202.

[33] Grasselli, M. (2017). The 4/2 stochastic volatility model: A unified approach for the heston and the 3/2 model. *Mathematical Finance*, 27(4):1013–1034.

[34] Griebel, M., Schneider, M., and Zenger, C. (1992). A combination technique for the solution of sparse grid problems. In *Proceedings of the IMACS International Symposium on Iterative Methods in Linear Algebra*. Elsevier, Amsterdam.

[35] Hagan, P. S., Kumar, D., Lesniewski, A. S., and Woodward, D. E. (2002). Managing smile risk. *The Best of Wilmott*, 1:249–296.

[36] Haug, E. G. (1997). *The Complete Guide to Option Pricing Formulas*. Investing/Finance. McGraw-Hill.

[37] Higham, D. J. and Trefethen, L. N. (1993). Stiffness of odes. *BIT Numerical Mathematics*, 33:285–303.

[38] Hochbruck, M., Lubich, C., and Selhofer, H. (1998). Exponential integrators for large systems of differential equations. *SIAM Journal on Scientific Computing*, 19(5):1552–1574.

[39] Hochbruck, M. and Ostermann, A. (2005a). Explicit exponential runge–kutta methods for semilinear parabolic problems. *SIAM Journal on Numerical Analysis*, 43(3):1069–1090.

[40] Hochbruck, M. and Ostermann, A. (2005b). Explicit exponential runge–kutta methods for semilinear parabolic problems. *SIAM Journal on Numerical Analysis*, 43(3):1069–1090.

[41] Hochbruck, M. and Ostermann, A. (2010). Exponential integrators. *Acta Numerica*, 19:209–286.

[42] Holtz, M. (2010). *Sparse grid quadrature in high dimensions with applications in finance and insurance*, volume 77. Springer Science & Business Media.

[43] Jäckel, P. and Kahl, C. (2007). Hyp hyp hooray.

[44] Ji, S. and Liu, H. (2023). Solvability of one kind of forward-backward stochastic difference equations. *Communications in Statistics - Theory and Methods*, 0(0):1–18.

[45] Karoui, N. E., Peng, S., and Quenez, M. C. (1997). Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71.

[46] Kirkby, J. L., Nguyen, D. H., and Nguyen, D. (2020). A general continuous time markov chain approximation for multi-asset option pricing with systems of correlated diffusions. *Applied Mathematics and Computation*, 386:125472.

[47] Lawson, J. D. (1967). Generalized runge-kutta processes for stable systems with large lipschitz constants. *SIAM Journal on Numerical Analysis*, 4(3):372–380.

[48] Leentvaar, C. C. W. (2008). *Pricing multi-asset options with sparse grids*. PhD thesis, Delft University of Technology.

[49] Li, L. and Zhang, G. (2018). Error analysis of finite difference and markov chain approximations for option pricing. *Mathematical Finance*, 28(3):877–919.

[50] Longstaff, F. A. and Schwartz, E. S. (2001). Valuing american options by simulation: a simple least-squares approach. *The review of financial studies*, 14(1):113–147.

[51] López-Salas, J. G. and Cendón, C. V. (2017). *Sparse Grid Combination Technique for Hagan SABR/LIBOR Market Model*, pages 477–500. Springer International Publishing, Cham.

[52] Meier, C., Li, L., and Zhang, G. (2023). Simulation of multidimensional diffusions with sticky boundaries via markov chain approximation. *European Journal of Operational Research*, 305(3):1292–1308.

[53] Mijatović, A. and Pistorius, M. (2013). Continuously monitored barrier options under markov processes. *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 23(1):1–38.

[54] Milstein, G. N. and Tretyakov, M. V. (2006). Numerical algorithms for forward-backward stochastic differential equations. *SIAM Journal on Scientific Computing*, 28(2):561–582.

[55] Minchev, B. V. and Wright, W. M. (2005). A review of exponential integrators for first order semi-linear problems.

[56] Munkres, J. R. (2000). *Topology*. Featured Titles for Topology. Prentice Hall, Incorporated.

[57] Novak, E. and Ritter, K. (1996). High dimensional integration of smooth functions over cubes. *Numerische Mathematik*, 75(1):79–97.

[58] Pardoux, É. (1999). *BSDEs, weak convergence and homogenization of semilinear PDEs*, pages 503–549. Springer Netherlands, Dordrecht.

[59] Pardoux, É. and Peng, S. (1990). Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61.

[60] Pardoux, É. and Răscanu, A. (2014). *Stochastic Differential Equations, Backward SDEs, Partial Differential Equations*. Springer.

[61] Peng, S. (1991). Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and Stochastics Reports*, 37:61–74.

[62] Rackauckas, C. and Nie, Q. (2017). DifferentialEquations.jl–a performant and feature-rich ecosystem for solving differential equations in Julia. *Journal of Open Research Software*, 5(1).

[63] Randall, C. and Tavella, D. A. (2000). *Pricing Financial Instruments: The Finite Difference Method*. Wiley, New York.

[64] Reisinger, C. and Wittum, G. (2007). Efficient hierarchical approximation of high - dimensional option pricing problems. *SIAM Journal on Scientific Computing*, 29(1):440–458.

[65] Smith, W. E., Sloan, I. H., and Opie, A. H. (1983). Product integration over infinite intervals i. rules based on the zeros of hermite polynomials. *Mathematics of Computation*, 40:519–535.

[66] Smolyak, S. A. (1963). Quadrature and interpolation formulas for tensor products of certain classes of functions. In *Doklady Akademii Nauk*, volume 148, pages 1042–1045. Russian Academy of Sciences.

[67] Söderlind, G., Jay, L., and Calvo, M. (2015). Stiffness 1952–2012: Sixty years in search of a definition. *BIT Numerical Mathematics*, 55(2):531–558.

[68] Strang, G. (2007). *Computational Science and Engineering*. Wellesley-Cambridge Press.

[69] Trefethen, L. N. (1996). Finite difference and spectral methods for ordinary and partial differential equations. available at http://people.maths.ox.ac.uk/trefethen/pdetext.html.

[70] Trefethen, L. N. and Embree, M. (2005). *Spectra and Pseudospectra*. Princeton University Press, Princeton.

[71] van der Stoep, A. W., Grzelak, L. A., and Oosterlee, C. W. (2014). The heston stochastic-local volatility model: Efficient monte carlo simulation. *International Journal of Theoretical and Applied Finance*, 17(07):1450045.

[72] Walter, W., Rosenblatt, L., and Shampine, L. (2012). *Differential and Integral Inequalities*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 2. Folge. Springer Berlin Heidelberg.

[73] Wanner, G. and Hairer, E. (1996). *Solving ordinary differential equations II*, volume 375. Springer Berlin Heidelberg New York.

[74] Zhang, G., Gunzburger, M., and Zhao, W. (2013). A sparse-grid method for multi-dimensional backward stochastic differential equations. *Journal of Computational Mathematics*, 31:221–248.

[75] Zhang, J. (2017). *Backward Stochastic Differential Equations: From Linear to Fully Nonlinear Theory*. Probability Theory and Stochastic Modelling. Springer New York.

List of Works

Papers

1. Kaneko, A. (2022). A numerical method for solving high-dimensional backward stochastic difference equations using sparse grids. JSIAM Letters, 14:104-107

2. Kaneko, A. (2023). Multi-stage Euler-Maruyama methods for Backward Stochastic Differential Equations driven by Continuous-time Markov chains, (to appear in Japan Journal of Industrial and Applied Mathematics)

Talks

1. Kaneko, A. (2021). A study on a numerical methods for solving backward stochastic difference equations using sparse grids. The 17th Joint Meeting of JSIAM Activity Groups, Japan SIAM.

2. Kaneko, A. (2021). A study on a numerical methods for solving backward stochastic difference equations using sparse grids. Osaka University Probability Seminar.

3. Kaneko, A. (2021). A study on a numerical methods for solving backward stochastic difference equations using sparse grids. Okayama Probability Seminar.

4. Kaneko, A. (2021). A study on a numerical methods for solving backward stochastic difference equations using sparse grids. Probability Young Seminar Online 2021.

5. Kaneko, A. (2022) Multi-stage Euler-Maruyama methods for backward stochastic differential equations driven by continuous-time Markov chains. The 2022 Annual Meeting, Japan SIAM.

6. Kaneko, A. (2023) Multi-stage Euler-Maruyama methods for backward stochastic differential equations with bounded stopping terminal times driven by continuous-time Markov chains. The 19th Joint Meeting of JSIAM Activity Groups, Japan SIAM.

7. Kaneko, A. (2023) Multi-stage Euler-Maruyama methods for backward stochastic differential equations driven by continuous-time Markov chains. Probability Early Spring Seminar 2023.

Award

1. Best presentation award for young researchers at the 2022 JSIAM Annual Meeting