

Contributions to the Numerics of Quadratic Backward SDEs and Probability on Turbulence



Mingrui Zhou
St Hugh's College
University of Oxford

A confirmation thesis submitted for the degree of
Doctor of Philosophy

Trinity 2021

Acknowledgements

I would like to thank my supervisor, Prof. Zhongmin Qian, for guiding me through my research and leading me into the field of BSDEs and Turbulence. His suggestions are invaluable. This thesis cannot be completed without his patience and profound understanding of mathematics.

I would like to express my gratitude to Prof. Ben Hambly and Prof. Shizan Fang for agreeing to be my examiners. I am grateful to Prof. Terry Lyons and Prof. David Prömel for providing many suggestions on my transfer. I would also thank Prof. Gui-Qiang Chen and Prof. Dmitry Belyaev for agreeing to become the assessor of my confirmation.

The mathematics community in Oxford benefit me a lot in my DPhil studies. Dr. Jiawei Li's precious ideas and suggestions are vital to the turbulence projects in my final thesis. Collaboration with her is pure enjoyment and I wish her all the best for her career in academia. Communications with Dr. Yufei Zhang inspired me on several results, while Dr. Siyuan Li's humour and his knowledge on Stochastic control theory supported me through the darkest times in my research. Despite having different research topics, Christina Ziyang Zou and Wolfgang Stockinger provided helpful comments on my projects. I am indebted to Dr. Wei Fang, Dr. Zhenru Wang, Dr. Guangyu Xi and Dr. Yixuan Wang who helped me since I was a Master student in Oxford. I am also happy to work with my colleagues Christina, Jing, Wolfgang, Yuhan in the same office for the past few years.

I would like to express my sincere gratitude to my parents, for being supportive on all the important decisions in my life. A special thanks to Zhangling, who encouraged me to apply for the DPhil in Mathematics when I was hesitating about my future. Her optimism and confidence guide me to tackle numerous obstacles and bring light to my life.

Lastly, I would like to thank all the other people that have helped me in my studies at Oxford. This journey is filled with laughter and tears, but I will always keep it in my heart.

Statement of Originality

I hereby declare that this thesis has been written by me and has not been submitted or is currently being submitted for any previous degree or other qualifications. Based on my knowledge, all material that has been published elsewhere and used in this thesis has a clear reference directing to the original result.

I claim that the thesis is an original report of my research, except where work that has formed part of jointly-authored publications has been included. My contribution and those of the other authors to this work have been explicitly indicated below. I confirm that appropriate credit has been given within this thesis where reference has been made to the work of others.

The majority of the contributions in this thesis are included in Chapter 3, 4, 5. Chapter 3 is based on an unpublished paper, which is completed under the supervision of my supervisor Prof. Zhongmin Qian. The next two Chapters are based on two Arxiv papers Li et al. [[57](#), [58](#)] respectively, which are both co-authored by Prof. Zhongmin Qian and Dr. Jiawei Li. All authors, who revised these two articles carefully, have made substantial contributions to the structure of these two papers.

Abstract

In this thesis, we first study the numerical scheme of path-dependent quadratic backward stochastic differential equations (quadratic BSDE or QBSDE for short). Under the assumption that the terminal condition is of the form $\Phi(X)$, where Φ is a Lipschitz function on $\tilde{\mathbb{D}}(\mathbb{R}^m)$ (the space of càdlàg paths) and X is the solution to a Lipschitz SDE satisfying linear growth conditions, we derive the path-regularity result which characterises the solution (Y, Z) . Motivated by the *modified implicit BTZ scheme* in [12], we propose the *modified (inner) explicit BTZ scheme* and conduct a perturbation analysis on Z -part of the scheme, followed by concluding the error of the scheme is $O(N^{-\frac{1}{2}})$ with N as the number of timesteps instead of $O(N^{-\frac{1}{2}+\varepsilon})$ for some $\varepsilon > 0$ in [12]. Numerical experiments support our conclusion on the convergence rate.

The second purpose of this thesis is to derive a partial differential equation (PDE) for the single-time single-point probability density function (PDF) of the velocity field of a turbulent flow. The PDF PDE is a highly non-linear parabolic-transport equation, which depends on two conditional statistical characteristics of important physical meaning. The PDF PDE is a general form of the classical Reynolds mean flow equation [80], and is a precise formulation of the PDF transport equation [77]. The PDF PDE provides us with a new method for modelling turbulence. An explicit example is constructed, though the example is seemingly artificial, it demonstrates the PDF method based on the new PDF PDE.

Lastly, vorticity random fields of turbulent flows are singled out as the main dynamical variables for the description of turbulence, and the evolution equation of the PDF of the vorticity field has been obtained. This PDF evolution equation is a mixed type PDE of the second order which depends only on the conditional mean (which is a first order statistic) of the underlying turbulent flow. This is in contrast with Reynolds' mean flow equation which relies on a quadratic statistic. The PDF equation is interesting on its own and is worthy of study as a PDE of second order.

Contents

1	Introduction	1
1.1	Background	1
1.2	Outline	4
1.3	Notations and conventions	4
2	Preliminaries	7
2.1	Basic setup	7
2.2	Malliavin derivative	9
2.3	Martingale theory	13
2.3.1	Conditional inequalities	13
2.3.2	BMO martingales	18
2.4	BSDE	21
2.4.1	Lipschitz BSDEs and quadratic BSDEs	23
2.4.2	BTZ scheme	25
2.5	Turbulence	26
2.5.1	Velocity field	27
2.5.2	Vorticity field	28
3	Path-regularity and Numerical Simulation of Path-dependent Quadratic BSDEs	29
3.1	Introduction	29
3.1.1	Preliminaries and assumptions	31
3.1.2	Contribution and main results	33
3.2	Path regularity to solutions of a class of BSDE	35
3.2.1	Classical differentiability of BSDE	35
3.2.2	Proof of results on path regularity	43
3.2.3	Averaging Z_t in time	57
3.3	The Modified (inner) explicit scheme	58

3.3.1	Properties of the explicit scheme	59
3.3.2	Perturbation of the centred scheme	65
3.4	Convergence of modified BTZ scheme	71
3.4.1	Modification on the generators	71
3.4.2	Estimating different versions of the Z component	72
3.4.3	Proof of Theorem 3.1.3	76
3.5	Implementation of the numerical scheme	83
3.5.1	Quantization and the lattice grid	84
3.5.2	Discretisation of forward processes	85
3.5.3	Impact on the backward component	89
3.6	Numerical experiments	93
3.6.1	Quantization	95
3.6.2	Regression basis	97
4	On distributions of velocity random fields in turbulent flows	102
4.1	Introduction	102
4.2	PDF equation of velocity fields	104
4.2.1	Random fields and their statistical characteristics	104
4.2.2	The evolution equation for the velocity distribution	107
4.3	Application to turbulent flows	113
4.3.1	Weakly homogeneous and weakly isotropic flows	113
4.3.2	Inviscid flows	120
4.4	Modelling the PDF: a concrete example	122
4.4.1	Space homogeneous density with perturbation	122
4.4.2	Motivation for the construction	124
4.5	Concluding remarks	127
5	On the transport equation for probability density functions of turbulent vorticity fields	128
5.1	Introduction	128
5.2	PDF equation for the vorticity	129
5.2.1	Prelims and assumptions	130
5.2.2	PDF equation and its derivation	131
5.3	Inviscid flows	135
5.4	Weakly homogeneous and weakly isotropic flows	137
5.5	Modeling the PDF of weakly isotropic flows	140
5.6	Heat flow method	143

A	Technical Proofs and Auxiliary results	148
A.1	Proofs on conditional inequalities	148
A.2	Proofs on BSDEs and related numerics	151
B	Supplement Material	166
B.1	Malliavin calculus and Watanabe theory	166
B.2	A note on energy inequalities	167
B.3	Remarks on BTZ scheme for Lipschitz BSDEs	168
B.4	Utility pricing in incomplete market	169
B.5	Other useful results	172
	Bibliography	173

List of Figures

3.1	Error analysis for <i>sin</i> type	96
3.2	Estimation of Y_0, Z_0^1, Z_0^2 under sin-type payoff using quantization methods	96
3.3	Error analysis for Lipschitz payoff, log-log	97
3.4	Estimation of Y_0, Z_0^1, Z_0^2 under Put type payoff using quantization methods	97
3.5	Error analysis for (Y_0, Z_0) with Asian Put type payoff	98
3.6	Estimation of Y_0, Z_0^1, Z_0^2 for Asian Put type payoff, using regression techniques	98
3.7	Error analysis for (Y_0, Z_0) with lookback Put type payoff	99
3.8	Estimation of Y_0, Z_0^1, Z_0^2 with lookback Put type payoff, with regression techniques	99
3.9	Utility indifference pricing and hedging strategy for 2 units of the claim	100
4.1	$x = (0, 0, 0), t = \frac{1}{2}$	124
4.2	$x = (0, 0, 0), t = 40$	125
4.3	$x = (12, 12, 12), t = 40$	126

List of Tables

1.1	Summary of basic notations in this thesis	6
3.1	Summary of parameter choices corresponding to Example 3.1	94
3.2	Summary of parameter choices corresponding to Example 3.2	100
3.3	Hedging strategy and utility indifference price of Example 3.2	101
3.4	Difference of computational time of the explicit and implicit scheme	101

Chapter 1

Introduction

1.1 Background

Stochastic differential equations (SDE), which are usually based on the theory of Itô's calculus and serve as natural models of time-dependent objects with random noise, have applications in a variety of areas, involving physics, engineering, biology and finance. An important aspect of stochastic analysis is developing a bridge between probability and partial differential equations (PDE), in the sense that PDE problems such as the existence and uniqueness of a parabolic equation, is solved by some probabilistic techniques.

To demonstrate how deterministic functions in PDE are associated with random variables in probability, we recall the classical homogeneous heat equation

$$\begin{aligned}\partial_t u(t, x) &= \frac{1}{2} \Delta_x u(t, x), & (t, x) &\in (0, T] \times \mathbb{R}^d, \\ u(0, x) &= g(x), & x &\in \mathbb{R}^d,\end{aligned}$$

where $u : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}$ for some $d \geq 1$. In physics, this equation plays an important role in describing the diffusion through a medium over time, such as the evolution of temperature in reality when $d = 3$. Without going into the technical details, we find the solution (see for example, Evans [26, Section 2.3])

$$\begin{aligned}u(t, x) &= \int_{\mathbb{R}^d} \frac{1}{(2\pi t)^{\frac{d}{2}}} g(y) \exp\left(-\frac{|y - x|^2}{2t}\right) dy \\ &= \mathbb{E}[g(X(t, x))],\end{aligned}$$

where $X(t, x) = x + B_t$ and B_t is the standard Brownian motion on $\mathbb{R}^d$.

It is not surprising that the phenomenon above was discovered in other PDEs. As a generalisation, the theory of Markov processes was developed in the 1940s to provide

a probabilistic approach to a class of linear parabolic PDEs, while the Feynman-Kac theorem indicated the solution to linear parabolic PDE can be written as an expectation of a function of a Markovian SDE, if the solution exists and is regular enough.

Suppose a new term f , which is a function of $u(t, x)$ and $\nabla_x u(t, x)$ with ∇_x as the gradient operator on x , is added to the heat equation. Instead of the initial condition, we assume the terminal condition, in the sense that

$$\begin{aligned}\partial_t u(t, x) &= -\frac{1}{2}\Delta_x u(t, x) - f(u(t, x), \nabla_x u(t, x)), \quad (t, x) \in [0, T) \times \mathbb{R}^d, \\ u(T, x) &= g(x), \quad x \in \mathbb{R}^d.\end{aligned}\tag{1.1}$$

This equation is semi-linear and cannot be solved using the same technique. However, the backward stochastic differential equation (BSDE), introduced at the end of the last century, is indeed a probabilistic counterpart of a class of semi-linear parabolic PDE. Instead of a forward SDE, a pair of adapted processes $(Y_t, Z_t)_{t \in [0, T]}$ is introduced

$$Y_t = \xi + \int_t^T f(Y_s, Z_s) \, ds - \int_t^T Z_s \cdot dB_s \tag{1.2}$$

for a terminal condition ξ , such that the pair (Y, Z) matches up with the solution u and its derivative $\nabla_x u$ respectively. Given $\xi = g(X(T, x))$ with $X(t, x) = x + B_t$, we have $Y_t = u(t, X(t, x))$ and $Z_t = (\nabla_x u)(t, X(t, x))$, which is called the non-linear Feynman-Kac formula.

From a probabilistic point of view, BSDE on its own is a problem worth discussing even if we don't focus on the PDE aspect. Suppose $f = 0$ and $d = 1$ in equation (1.2), the BSDE turns into

$$Y_t = \xi - \int_t^T Z_s \, dB_s,$$

where the process Y_t is a martingale and Z_t exists by the martingale representation theorem. Therefore, BSDE can be regarded as a “non-linear martingale” to some extent. Moreover, suppose ξ is nice enough, the Clark-Ocone formula (see Proposition B.1.2) suggests that Z component is obtained through

$$Z_t = \mathbb{E}[D_t F | \mathcal{F}_t] = D_t \mathbb{E}[F | \mathcal{F}_t] = D_t Y_t,$$

where D is the Malliavin derivative operator. If ξ depends on the paths of Brownian motion, such as $\xi = \sup_{t \in [0, T]} |B_t|$, the PDE counterpart is unknown, but the

associated BSDE is well-defined and corresponds to problems in, for example, mathematical finance. This motivated us to conclude a numerical scheme for BSDEs where the terminal condition is a function of the path of the forward process.

Apart from the heat equation, the Navier-Stokes equation is also a parabolic equation. It has the form

$$\begin{aligned}\frac{\partial U^i}{\partial t} + U(x, t) \cdot \nabla_x U^i(x, t) &= \nu \Delta_x U^i - \frac{\partial P}{\partial x^i}, & 1 \leq i \leq d, \quad (x, t) \in \mathbb{R}^d \times (0, T], \\ \nabla_x \cdot U(x, t) &= 0, & (x, t) \in \mathbb{R}^d \times [0, T], \\ U(x, 0) &= U_0(x), & x \in \mathbb{R}^d,\end{aligned}$$

where P is known as the “pressure term” and $\nu \geq 0$ is the “viscosity constant”. This equation is particularly important in physics. Flows are divided into laminar flows, which are more smooth, and turbulent flows where the characteristics of such flows vary in a disordered manner in space-time. Compared to laminar flows, turbulent flows are more common in practice and much more effective in terms of transporting and mixing the fluid. Unfortunately, they are also much more unpredictable in contrast to laminar flows. The Navier-Stokes equation is particularly useful since it accurately describes the evolution of the turbulent velocity U in time and space.

In order to evaluate the velocity at time $T > 0$, we can plug in the (deterministic) initial velocity followed by applying direct numerical simulation (DNS) of the Navier-Stokes equation. The DNS is intractable for high Reynolds number flows, whereas turbulent flows often have high Reynolds numbers.

This motivates us to switch to the approach in the spirit of statistical physics and to consider the initial velocity as a random field rather than a deterministic field. When the initial velocity is random, the solution $U(x, t)$ at (x, t) becomes a random variable, which has a law (or called distribution) defined as the probability measure on the Borel σ -algebra of $\mathbb{R}^3$. Moreover, the distribution of the random field $\{U(x, t)\}_{x, t}$ consists of, by definition, all possible finite-dimensional marginal joint distributions of

$$U(x_1, t_1), \dots, U(x_n, t_n)$$

where $x_i \in \mathbb{R}^3$, $t_i \geq 0$ and any positive integer n . We are keen to investigate its law, in particular, we will focus on the probability density function (PDF) of $U(x, t)$ at some fixed (x, t) as it is sufficient to describe the law. An alternative way to model turbulence is to analyse the vorticity W , which is defined as $W = \nabla \wedge U$. The PDF of the vorticity, which is expected to be extracted from the vorticity equation together with the random initial data, will be our main purpose in this scenario.

1.2 Outline

In Chapter 2, we list basic definitions and results covering Mallavin calculus, martingale theory, classical numerical schemes for BSDE and the theory of turbulence.

The BSDE with path-dependent terminal condition and driver of quadratic growth in the Z component is the main subject of study in Chapter 3. A path-regularity theorem is established, followed by the analysis on the *modified (inner) explicit BTZ scheme* which is motivated by the modified implicit scheme in [12]. As an application, we illustrate the convergence of the numerical scheme for such type of BSDE with a specified convergence rate.

The next two chapters include two projects on the study of velocity field Li et al. [57] and vorticity field Li et al. [58] respectively. In particular, these projects are based on the joint work of Professor Zhongmin Qian, Dr. Jiawei Li and the author.

Chapter 4 is devoted to characterising the PDF of the velocity field $U(x, t)$ which satisfies the Navier-Stokes equation. The PDF is single-time and single-space, which satisfies a PDE, the so-called PDF PDE. A stochastic representation formula is given to solve the PDE under some assumptions.

In Chapter 5, we derive the PDE that is satisfied by the PDF of vorticity random field. In contrast to the velocity field, only the conditional mean of the vorticity field is involved in the equation. The PDF equation is interesting on its own and is worthy of study as a PDE of second order.

1.3 Notations and conventions

We summarise the notations and list assumptions throughout this thesis, for the convenience of the reader.

$|\cdot|$ is the Euclidean norm for vectors or scalars. For a vector $I = (i_1, \dots, i_k, \dots)$ (or multi-index) with finite or infinite length, $|I|_1 = \sum_{k=1}^{\infty} |i_k|$. For an $A \in \mathbb{R}^{N \times d}$, $|A|$ refers to the Frobenius norm i.e. $|A| := \text{Tr}(AA^T) = \sqrt{\sum_{i=1}^N \sum_{j=1}^d |A_{ij}|^2}$. Vectors are assumed to be column vectors, hence for any vectors $a, b \in \mathbb{R}^N$ and $\sigma \in \mathbb{R}^{N \times d}$, $a \cdot b = \sum_{i=1}^N a^i b^i$ and $a \cdot \sigma = \sigma \cdot a = (a^\top \sigma)^\top = \sigma^\top a \in \mathbb{R}^d$. Given a scalar function $u : \mathbb{R}^N \rightarrow \mathbb{R}$ which maps $x \in \mathbb{R}^N$ to $u(x)$, we write the operator $\partial_x := \nabla_x$. For a vector field $u : \mathbb{R}^N \rightarrow \mathbb{R}^M$, the Jacobian matrix is defined as

$$\nabla_x u = \begin{bmatrix} \partial_{x^1} u^1 & \cdots & \partial_{x^N} u^1 \\ \vdots & \ddots & \vdots \\ \partial_{x^1} u^M & \cdots & \partial_{x^N} u^M \end{bmatrix} \in \mathbb{R}^{M \times N}$$

i.e. operators ∂_{x^i} are applied column-wise. Given a function $f : \mathbb{R}^N \rightarrow \mathbb{R}$, we write $\partial_i f$ as $\frac{\partial f}{\partial x^i}$. Given a multi-index $I = (i_1, \dots, i_k)$, we denote $\partial_I f := \partial_{i_1} \cdots \partial_{i_k} f$. For a vector $x = (x^1, x^2, \dots, x^N)$, we write $\int_{\mathbb{R}^N} f(x) dx = \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} f(x^1, \dots, x^N) dx^1 \cdots dx^N$, whereas $\int_0^t V_s \cdot X_s := \sum_{i=1}^N \int_0^t V_s^i dX_s^i$ for vector-valued processes $X_s, V_s \in \mathbb{R}^N$. If H refers to a Hilbert space, $H^{\otimes j} = \bigotimes_{i=1}^j H$ for all integer $j \geq 1$. If f, g are functions in some Hilbert space H , then we write $f \otimes g \in H^{\otimes 2}$ and $f^{\otimes m} = \bigotimes_{i=1}^m f$.

A table of notations is also presented in the following. The assumptions, such as [A6], unless we specified, are **not** assumed to be satisfied for all statements throughout this thesis. **An assumption is imposed if and only if we mention it in the description of the result (such as lemmas, propositions, theorems, corollaries, remarks).**

Constants

d	The dimension of the Brownian motion $(B_t)_{t \geq 0}$
m	In Chapter 3, it refers to the dimension of the forward SDE X_t
N	In Chapter 3, it refers to the number of time-steps
n	In Chapter 3, it refers to the truncation parameter of the driver f
T	A deterministic positive constant, which denotes the terminal time
C	In Chapter 3, it is a generic constant that depends on some irrelevant (or fixed) parameters

Spaces and Norms

$\mathcal{C}(A; \hat{A})$	The space of continuous functions $f : A \rightarrow \hat{A}$
$\mathcal{C}^{a_1, a_2}(A_1 \times A_2; \hat{A})$	The space of functions $f : A_1 \times A_2 \rightarrow \hat{A}$ such that a_1 -continuously differentiable in A_1 and a_2 -continuously differentiable in A_2

Operators and functions

D	Malliavin derivative operator, see Definition 2.2.2
$\int dX_t$	Itô's integral with respect to a $\mathbb{R}$ -valued process X_t
$\mathcal{E}(X)$	Stochastic exponential ¹ with respect to a process X

¹For a càdlàg martingale X , the stochastic exponential satisfies $\mathcal{E}(X)_t = 1 + \int_0^t \mathcal{E}(X)_{s-} dX_s$. Moreover, $\mathcal{E}(X)_t = \exp(X_t - \langle X^c \rangle_t) \prod_{0 \leq s \leq t} ((1 + \Delta X_s) \exp(-\Delta X_s))$, provided X^c referring to the continuous part of X and $\Delta X_t := X_t - X_{t-}$.

$(\mathcal{Y}_i^\pi, \mathcal{Z}_i^\pi)$	Numerical scheme (3.4) for QBSDE with truncated driver and approximated path-dependent data
(Y_i^π, Z_i^π)	Solution to the (general) modified BTZ numerical scheme (3.22)
$(\tilde{Y}_i, \tilde{Z}_i)$	Perturbed scheme (3.27) with a (general) perturbation ζ^Y on the Y component
$(\check{Y}_i, \check{Z}_i)$	Perturbed scheme (3.33) with a (general) perturbation ζ^Z on the Z component
$(\hat{\mathcal{Y}}_i^\pi, \hat{\mathcal{Z}}_i^\pi)$	Implementation of scheme (3.4) with quantization techniques, see (3.47)

Assumptions

Assumption [A1]	See page 24
Assumption [A2]	See page 31
Assumption [A3]	See page 31
Assumption [A4]	See page 31
Assumption [A5]	See page 60
Assumption [A6]	See page 60
Assumption [A7]	See page 62
Assumption [A8]	See page 113
Assumption [A9]	See page 113

Others

$X \simeq Y$	$(X, \ \cdot\ _X)$ is isometrically isomorphic ² to $(Y, \ \cdot\ _Y)$
$[X]_{\mathbb{P}}$	The law of the random variable X w.r.t $\mathbb{P}$
$A^\top$	The transpose of matrix A

Table 1.1: Summary of basic notations in this thesis

²That is, there exist a bijective (or surjective, since an isometry is injective by construction) linear map $T : X \rightarrow Y$ such that $\|Tx\|_Y = \|x\|_X$ for all $x \in X$.

Chapter 2

Preliminaries

In this chapter, we quote classical results in the literature and present some introductory material which is helpful throughout the subsequent chapters. The notation and terminologies used in this chapter will be consistent in the whole thesis unless we specify.

2.1 Basic setup

Definition 2.1.1 (Function space on $\mathbb{R}^N$). *Let $f : \mathbb{R}^N \rightarrow \mathbb{R}$ be a function.*

- $\mathcal{C}(\mathbb{R}^N; \mathbb{R})$ is the space of continuous functions $f : \mathbb{R}^N \rightarrow \mathbb{R}$. $\mathcal{C}^n(\mathbb{R}^N; \mathbb{R})$ denotes the space of n -times continuously differentiable functions such that for any multi-index $I = (i_1, \dots, i_k) \in \{1, \dots, N\}^k$ such that $k \leq n$, $\partial_I f$ exists and is continuous. At last, $\mathcal{C}^\infty(\mathbb{R}^N; \mathbb{R})$ is the space of infinitely differentiable functions i.e. $f \in \mathcal{C}^n(\mathbb{R}^N; \mathbb{R})$ for all $n \geq 1$.
- $\mathcal{C}_p^\infty(\mathbb{R}^N; \mathbb{R})$ refers to the space of infinitely differentiable functions $f : \mathbb{R}^N \rightarrow \mathbb{R}$ such that partial derivatives $\partial_I f$ with multi-indices I have at most polynomial growth.

The spaces $\mathcal{C}(\mathbb{R}^N; \mathbb{R}^N)$, $\mathcal{C}^n(\mathbb{R}^N; \mathbb{R}^N)$, $\mathcal{C}^\infty(\mathbb{R}^N; \mathbb{R}^N)$, $\mathcal{C}_p^\infty(\mathbb{R}^N; \mathbb{R}^N)$ of vector fields $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$ are defined accordingly.

Definition 2.1.2 (Random variables on the probability space). *Given the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ where $(\mathcal{F}_t)_t$ satisfied the usual conditions, setting $1 \leq p < \infty$, we define the following spaces:*

- $L^p(\mathbb{R}^d, \mathcal{F}_T, \mathbb{P}) = \{X : \Omega \rightarrow \mathbb{R}^d \mid X \text{ is } \mathcal{F}_T\text{-measurable, } \|X\|_{L^p(\mathbb{P})}^p := \mathbb{E}[|X|^p] < \infty\},$

- $L^\infty(\mathbb{R}^d) = \{X : \Omega \rightarrow \mathbb{R}^d \mid X \text{ is } \mathcal{F}_T\text{-measurable, } \|X\|_\infty := \text{ess sup}_{\omega \in \Omega} |X(\omega)| < \infty\},$
- $\mathbb{S}_T^p(\mathbb{R}, \mathbb{P}) := \{Y : (t, \omega) \in [0, T] \times \Omega \rightarrow \mathbb{R}^d \mid (Y_t)_t = (Y(t, \cdot))_t \text{ be an adapted càdlàg process such that } \|Y\|_{\mathbb{S}_T^p}^p := \mathbb{E}[\sup_{[0, T]} |Y_t|^p]\},$
- $\mathbb{H}_T^p(\mathbb{R}^{d \times m}, \mathbb{P}) := \{X : (t, \omega) \in [0, T] \times \Omega \rightarrow \mathbb{R}^{d \times m} \mid (X_t)_t = (X(t, \cdot))_t \text{ be a predictable matrix-valued process such that } \|X\|_{\mathbb{H}_T^p}^p := \mathbb{E} \left[\left(\int_0^T |X_t|^2 dt \right)^{\frac{p}{2}} \right]\},$
- $BMO_{p,T}(\mathbb{R}^d, \mathbb{P}) := \{X : (t, \omega) \in [0, T] \times \Omega \rightarrow \mathbb{R}^d \mid (X_t)_t = (X(t, \cdot))_t \text{ be an adapted càdlàg martingale such that given any bounded stopping time } 0 \leq \tau \leq T, \|X\|_{BMO_{p,T}(\mathbb{P})}^p := \sup_\tau \|E[|X_T - X_{\tau-}|^p \mid \mathcal{F}_\tau]\|_\infty < \infty\},$
- $\tilde{\mathbb{D}}_T(\mathbb{R}^m)$ is the space of m -dimensional càdlàg functions defined in $[0, T]$. We equip this space with the norm $\|\cdot\|_{\tilde{\mathbb{D}}_T(\mathbb{R}^m)}$ either by $\|\cdot\|_{\tilde{\mathbb{D}}_T^1(\mathbb{R}^m)}$, which are defined as $\|\Theta\|_{\tilde{\mathbb{D}}_T^\infty(\mathbb{R}^m)} = \sup_{t \in [0, T]} |\Theta_t|$ and $\|\Theta\|_{\tilde{\mathbb{D}}_T^1(\mathbb{R}^m)} = \int_0^T |\Theta_t| dt,$
- $\tilde{\mathbb{D}}_{[s, t]}(\mathbb{R}^m)$ is the space of m -dimensional càdlàg functions defined in $[s, t]$, which can be embedded into $\tilde{\mathbb{D}}_T(\mathbb{R}^m)$ by defining $\hat{\Theta} = \Theta \mathbb{1}_{[s, t]} \in \tilde{\mathbb{D}}_T(\mathbb{R}^m)$ for any $\Theta \in \tilde{\mathbb{D}}_{[s, t]}(\mathbb{R}^m),$
- $\mathbb{D}^{1,p}(\mathbb{R}, \mathbb{P}) := \{F \mid \|F\|_{\mathbb{D}^{1,p}(\mathbb{R})} = (\|F\|_{L^p(\mathbb{R}, \mathbb{P})}^p + \|DF\|_H^p)^{\frac{1}{p}} < \infty, \text{ where } D \text{ is the Malliavin derivative operator and } H = L^2([0, T]; \mathbb{R}^d)\}, \text{ for some } p \geq 2,$
- $\mathbb{H}_{BMO, T}^p(\mathbb{R}^d, \mathbb{P}) := \{\alpha : (t, \omega) \in [0, T] \times \Omega \rightarrow \mathbb{R}^d \mid (\alpha_t)_{t \in [0, T]} \text{ is an adapted process such that the process } (M_t)_{t \in [0, T]} \in BMO_{p, T}(\mathbb{P}, \mathbb{R}), \text{ where } M_t := \int_0^t \alpha_s \cdot dB_s\}. \text{ The space is equipped with the norm}^1$

$$\begin{aligned} \|\alpha\|_{\mathbb{H}_{BMO, T}^p(\mathbb{P}, \mathbb{R}^d)} &:= \left\| \int_0^\cdot \alpha_s \cdot dB_s \right\|_{BMO_{p, T}(\mathbb{P}, \mathbb{R})} \\ &= \sup_{t \in [0, T]} \left\{ \mathbb{E} \left[\left(\int_t^T |\alpha_s|^2 ds \right)^{\frac{p}{2}} \mid \mathcal{F}_t \right]^{\frac{1}{p}} \right\}. \end{aligned}$$

Remark 2.1 (Notation and conventions). Throughout this thesis, we would like to use the following notation.

- (i) We focus on the filtered atom-less probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ for some $T > 0$ unless we specify otherwise, while the filtration $(\mathcal{F}_t)_{t \in [0, T]}$ supports

¹We are allowed to consider deterministic times only. This will be explained in Corollary 2.3.13.

a d -dimensional Brownian motion $(B_t)_{t \in [0, T]}$, whose components are $(B_t^i)_{t \in [0, T]}$ for all $1 \leq i \leq d$. We denote the augmented filtration generated by $(B_t)_{t \in [0, T]}$ as $\mathcal{F}_t^B$. In general, we assume $(\mathcal{F}_t)_{t \in [0, T]} = \mathcal{F}_t^B$ unless we specified.

- (ii) As all norms are taken under $\mathbb{P}$, we omit $\mathbb{P}$ unless we specify. Moreover, we also omit T , such as $\mathbb{S}^p(\mathbb{R}^d)$, $\mathbb{H}^p(\mathbb{R}^d)$, $BMO_p(\mathbb{R}^d)$ and $\mathbb{H}_{BMO}^p(\mathbb{R}^d)$. As the BMO norms are equivalent (see Lemma 2.3.11), we sometimes write $\mathbb{H}_{BMO}(\mathbb{R}^d)$ instead of $\mathbb{H}_{BMO}^p(\mathbb{R}^d)$.
- (iii) Let $\Theta = (\Theta_t)_{t \in [0, T]}$ be a càdlàg process in $[0, T] \times \Omega$. For any $\omega \in \Omega$, $\Theta(\omega) \in \tilde{\mathbb{D}}(\mathbb{R}^m)$ is a càdlàg path such that $\Theta(\omega)(t) = \Theta_t(\omega)$. For any $\Phi : \tilde{\mathbb{D}}(\mathbb{R}^m) \rightarrow E_1$ (for example, $E_1 = \mathbb{R}^{m_1}$ with $m_1 > 0$), we write $\Phi(\Theta) : \Omega \rightarrow E_1$ as the E_1 valued random variable such that $\Phi(\Theta)(\omega) = \Phi(\Theta(\omega))$.

2.2 Malliavin derivative

As an infinite-dimensional differential calculus on Wiener space, Malliavin calculus was first established to provide a proof of Hörmander's theorem. Our aim is to introduce the basic definitions of the derivative operator adapted from [69], which will be heavily used in a numerical scheme for BSDE. Let us consider the Hilbert space $H = L^2([0, T]; \mathbb{R}^d)$. We observe that $L^2(\Omega; H) \simeq L^2(T \times \Omega; \mathbb{R}^d)$ i.e. they are isometrically isomorphic, hence for every $u \in L^2(\Omega; H)$, it can be treated as an element in $\tilde{u} \in L^2([0, T] \times \Omega; \mathbb{R}^d)$ by writing $\tilde{u}(t, \omega) := u(\omega)(t)$. For simplicity, we do not distinguish between u and $\tilde{u}$, and the same applies for $L^2(\Omega; H^{\otimes k})$ and $L^2([0, T]^k \times \Omega; (\mathbb{R}^d)^{\otimes k})$, for any $k \geq 1$. Now, we define the mapping which connects H to the space of Gaussian random variables.

Definition 2.2.1 (White noise/Wiener integral). *Let $H = L^2([0, T]; \mathbb{R}^d)$, with associated inner product*

$$\langle h_1, h_2 \rangle_H = \int_0^t h_1(s) \cdot h_2(s) \, ds \quad \forall h_1, h_2 \in H.$$

The white noise is the linear isometry $W : h \in H \rightarrow W(h) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ such that for every $h_1, h_2 \in H$, $W(h_1)$ is a centered Gaussian random variable and the function $\langle W(h_1), W(h_2) \rangle_{L^2} = \langle h_1, h_2 \rangle_H$. Equivalently,

$$W(h) := \int_0^T \sum_{i=1}^d h^i(s) \, dB_s^i,$$

which is known as the Wiener integral.

Definition 2.2.2 (The Malliavin derivative). *Let $\mathcal{S}$ is the space of $\mathbb{R}$ -valued smooth random variables such that*

$$\mathcal{S} := \left\{ f(W(h_1), \dots, W(h_k)) \mid \begin{array}{l} h_i \in H \text{ for all } 1 \leq i \leq k, \\ f \in \mathcal{C}_p^\infty(\mathbb{R}^k, \mathbb{R}) \text{ for some } k \geq 1 \end{array} \right\}.$$

The derivative operator D is first defined on $\mathcal{S}$, such that the derivative DF with $F \in \mathcal{S}$ is given by

$$DF := \sum_{i=1}^n \partial_i f(W(h_1), \dots, W(h_n)) h_i.$$

Notice that DF is an H -valued smoothed random variable. Moreover, the operator D can be applied iteratively on smooth random variable F , in the sense that

$$D^k F := \sum_{i_1, \dots, i_k=1}^n (\partial_{i_1} \cdots \partial_{i_k} f)(W(h_1), \dots, W(h_n)) h_{i_1} \otimes \cdots \otimes h_{i_k} \in L^2(\Omega; H^{\otimes k}),$$

for any $k \geq 1$. By convention, we treat $D^0 F = F$.

The derivative can be also treated as the stochastic process $\{D_{t_1, \dots, t_k}^k F\}_{t_1, \dots, t_k \in [0, T]}$, valued in $(\mathbb{R}^d)^{\otimes k}$. In particular, for smooth random variable F , it has the form

$$D_{t_1, \dots, t_k}^k F = \sum_{i_1, \dots, i_k=1}^n (\partial_{i_1} \cdots \partial_{i_k} f)(W(h_1), \dots, W(h_n)) h_{i_1}(t_1) \otimes \cdots \otimes h_{i_k}(t_k).$$

As $D_t F$ is a $\mathbb{R}^d$ -valued random variable, we write it as a vector with components $D^{(i)} F$, namely $D_t F = (D_t^{(1)} F, \dots, D_t^{(d)} F)$.

Lastly, we define the spaces of random variables which has certain degree of Malliavin differentiability.

- 1) *The space $\mathbb{D}^{k,p}$, which can be regarded as the domain of D^k in $L^p(\Omega)$, is defined as the closure of $\mathcal{S}$ with respect to $\|\cdot\|_{k,p}$, whose definition is*

$$\|F\|_{k,p} := \left(\mathbb{E}[|F|^p] + \sum_{j=1}^k \mathbb{E}[\|D^j F\|_{H^{\otimes j}}^p] \right)^{\frac{1}{p}}, \quad \forall F.$$

Finally, the space of smooth Wiener functions is defined as

$$\mathbb{D}^\infty := \bigcap_{p \geq 1} \bigcap_{k \in \mathbb{N}} \mathbb{D}^{k,p}.$$

- 2) *Similarly, $\mathbb{D}^{k,p}(V)$ with some Hilbert space V denotes the V -valued random variables that are bounded under the norm*

$$\|F\|_{k,p,V} := \left(\mathbb{E}[\|F\|_V^p] + \sum_{j=1}^k \mathbb{E}[\|D^j F\|_{H^{\otimes j} \otimes V}^p] \right)^{\frac{1}{p}}.$$

Remark 2.2. The space $(\mathbb{D}^{1,2})^d$ consists of random vectors $F = (F^1, \dots, F^d)$ with every component belongs $\mathbb{D}^{1,2}$. By construction, $(\mathbb{D}^{1,2})^d \simeq \mathbb{D}^{1,2}(\mathbb{R}^d)$.

Similar to the classical Fréchet derivative, the Malliavin derivative possesses the chain rule [69, Proposition 1.2.3, Proposition 1.2.4], at least for Lipschitz functions. The operator D_t satisfies the property

$$D_t (\mathbb{E} [F \mid \mathcal{F}_s]) = \mathbb{E} [D_t F \mid \mathcal{F}_s] \mathbb{1}_{[0,s]}(t),$$

by an application of [69, Proposition 1.2.8]. Moreover, we are able to view the space $\mathbb{D}^{k,2}(V)$ for all $k \geq 1$ and some Hilbert space V , as a Hilbert space associated with inner product

$$\langle F, G \rangle_{\mathbb{D}^{k,2}(V)} = \sum_{j=0}^k \langle D^j F, D^j G \rangle_{L^2(\Omega, H^{\otimes j} \otimes V)}, \quad \forall F, G \in \mathbb{D}^{k,2}(V).$$

The adjoint of the Malliavin derivative operator D is known as the divergence operator.

Definition 2.2.3 (The divergence operator). δ is the adjoint of the Malliavin derivative operator D , such that $\delta : L^2(\Omega; H) \rightarrow L^2(\Omega)$ and it satisfies the following.

1) The domain of δ , denoted by $\text{Dom}(\delta)$, is defined as

$$\text{Dom}(\delta) := \{u \in L^2(\Omega; H) \mid \mathbb{E} [\langle DF, u \rangle_H] \leq c(u) \|F\|_2, \\ \text{with some constant } 0 < c(u) < \infty, \quad \forall F \in \mathbb{D}^{1,2}\}.$$

2) The operator δ is characterised by

$$\mathbb{E} [F \delta(u)] = \mathbb{E} [\langle DF, u \rangle_H].$$

for all $u \in \text{Dom}(\delta)$ and $F \in \mathbb{D}^{1,2}$.

For $u \in \text{Dom}(\delta)$, we also use the following notation²

$$\delta(u) = \int_0^T u(s) \delta B_t,$$

which is known as the Skorohod integral.

²The sign δB_t can be roughly viewed as $\cdot dB_t$, but it should not be confused with Itô's integral.

The Skorohod integral has zero-mean as in the case of Itô's integral, since

$$\mathbb{E}[\delta(u)] = \mathbb{E}[\langle D1, u \rangle_H] = 0.$$

Moreover, both operators are closed, see [70, Proposition 1.30] or Proposition B.1.1. The space $\text{Dom}(\delta) \subseteq L^2(\Omega; H)$ is rather vague, we are often interested in the following objects.

Definition 2.2.4 (Subspace of $\text{Dom}(\delta)$). *Let $p \geq 2$.*

- 1) *We define $\mathbb{L}^{1,p}(\mathbb{R}^d) := \mathbb{D}^{1,p}(H) \subseteq \text{Dom}(\delta)$, which associates with the norm defined in Definition 2.2.2. Given $u = (u_s)_{s \in [0,T]} \in \mathbb{L}^{1,2}(\mathbb{R}^d)$ and $u_s = (u_s^1, \dots, u_s^d)$, we write the $\mathbb{R}^{d \times d}$ -valued matrix $D_t u_s$ with entries $(D_t u_s)_{ij} = D_t^{(j)} u_s^i$.*
- 2) *$\mathbb{L}_a^p(\mathbb{R}^d)$ is the closed subspace of adapted processes $\{u(t, \cdot)\}_{t \geq 0} \in L^p(\Omega; H)$.*
- 3) *$\mathbb{L}_a^{1,p}(\mathbb{R}^d) := \{(u \in \mathbb{L}_a^2(\mathbb{R}^d) \mid u(t, \cdot) \in \mathbb{D}^{1,p}(\mathbb{R}^d) \text{ a.s. } \forall t \in [0, T] \text{ and } \|u\|_{\mathbb{L}_a^{1,p}} < \infty\}$, where*

$$\|u\|_{\mathbb{L}_a^{1,p}} := \left(\mathbb{E} \left[\left(\int_0^T |u(t)|^2 dt \right)^{\frac{p}{2}} \right] + \mathbb{E} \left[\left(\int_0^T \int_0^T |D_s u(t)|^2 ds dt \right)^{\frac{p}{2}} \right] \right)^{\frac{1}{p}}.$$

By [69, Proposition 1.3.11], the Skorohod integral coincides with Itô's integral when $u \in \mathbb{L}_a^{1,2}(\mathbb{R}^d)$, such that it is understood as an extension of Itô's integral when the integrand is not adapted. The Malliavin derivative for such integral is given in the following proposition.

Proposition 2.2.5 (Mallavin derivatives for Skorohod integrals, Proposition 1.3.8 in [69]). *Given $u \in \mathbb{L}^{1,2}(\mathbb{R}^d)$, we denote $u_s = (u_s^1, \dots, u_s^d)$ and the derivative operator $D_t u_s^i = (D_t^{(1)} u_s^i, \dots, D_t^{(d)} u_s^i)$. We write*

$$\left(\int_0^T D_t u_s \delta B_s \right)^j := \delta \left(D_t^{(j)} u \right)$$

and suppose $\{D_t u_s\}_{s \in [0,T]}$ is Skorohod integrable for almost all $t \in [0, T]$. If there is a version of the process $\{\int_0^T D_t u_s \delta B_s\}_{t \in [0,T]}$ which belongs to $L^2([0, T] \times \Omega; \mathbb{R}^d)$, then $\delta(u) \in \mathbb{D}^{1,2}$ and

$$D_t(\delta(u)) = u_t + \int_0^T D_t u_s \delta B_s.$$

In particular, if $X_t = \int_0^t u_s \cdot dB_s$ for some square integrable adapted process $\{u_t\}_t \in \mathbb{L}^{1,2}(\mathbb{R}^d)$, then

$$D_s X_t = \begin{cases} u_s + \sum_{i=1}^d \int_s^t D_s u_r^i dB_r^i, & 0 \leq s \leq t \\ 0, & s > t \end{cases}$$

almost surely on $T \times \Omega$.

Lemma 2.2.6 (Malliavin derivative for diffusion processes, Section 2.3.1, equation (2.59) in [69]). *Let $X_t \in \mathbb{R}^N$ be a solution to the SDE*

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s$$

where $b : [0, T] \times \mathbb{R}^N \rightarrow \mathbb{R}^N$ and $\sigma : [0, T] \times \mathbb{R}^N \rightarrow \mathbb{R}^{N \times d}$ are measurable functions that belongs to class $\mathcal{C}^{1+\alpha}$ for some $\alpha > 0$.

Denoting the Jacobian matrix $\nabla_x X_t$ with entries $(\nabla_x X_t)_{ij} = \partial_{x^j} X_t^i$, we write $(D_\theta X_t)_{ij} = D^{(j)} X_t^i$ with $1 \leq i \leq N, 1 \leq j \leq d$ for any $\theta, t \in [0, T]$, which possesses the representation formula

$$D_\theta X_t = \nabla_x X_t (\nabla_x X_\theta)^{-1} \sigma(\theta, X_\theta) \mathbb{1}_{\theta \leq t}.$$

For further details corresponding to Malliavin derivative, we refer to Nualart's comprehensive book on Mallavin calculus [69].

2.3 Martingale theory

Sharp estimates and characterisations of semi-martingales are vital in probability theory, which are commonly used to show the stability of numerical schemes. Extending classical inequalities to conditional expectations and exploring properties of BMO martingales with applications to Itô's integrals are the two aims of this section.

2.3.1 Conditional inequalities

Within the following chapters of the thesis, we might encounter some circumstances where we need to give an estimate of a conditional expectation. This motivates us to develop some classical inequalities such as Markov inequality, Doob's L^p inequality and Burkholder-Davis-Gundy (BDG) inequality, in the context of conditional expectations. The proof will generally follow the strategy of the proof of these classical inequalities within Revuz and Yor [79], hence the result is not very surprising, in the sense that the conditional expectation preserves the monotonicity relation as well as constants within those inequalities.

We defer most of the technical proofs in this subsection to section A.1.

Lemma 2.3.1 (Conditional Markov inequality). *Let X be an integrable random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Suppose $\mathcal{G} \subseteq \mathcal{F}$ is a sub- σ -algebra. The conditional Markov inequality is referring to, given $a > 0$,*

$$\mathbb{P}(|X| \geq a \mid \mathcal{G}) := \mathbb{E}[\mathbb{1}_{\{|X| \geq a\}} \mid \mathcal{G}] \leq \frac{\mathbb{E}[|X| \mid \mathcal{G}]}{a} \quad a.s.$$

Proof. See section A.1. □

We first give a discrete version of Doob's maximal inequality.

Lemma 2.3.2. *Suppose $(X_n)_{n=0}^N$ is a L^1 -submartingale defined on probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$. For any $0 \leq m \leq N$ and any $\lambda > 0$, we have*

$$\lambda \mathbb{P}(X_{m,N}^* \geq \lambda \mid \mathcal{F}_m) \leq \mathbb{E} \left[X_N \mathbb{1}_{\{X_{m,N}^* \geq \lambda\}} \mid \mathcal{F}_m \right] \leq \mathbb{E} \left[|X_N| \mathbb{1}_{\{X_{m,N}^* \geq \lambda\}} \mid \mathcal{F}_m \right] \text{ a.s.},$$

where $X_{m,N}^* := \sup_{m \leq n \leq N} X_n$.

Proof. See section A.1. □

Lemma 2.3.3. *Suppose $(X_n)_{n=0}^N$ be a L^1 -martingale or a positive submartingale defined on probability space $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n \geq 0}, \mathbb{P})$. Let $X_{m,N}^* := \sup_{m \leq n \leq N} |X_n| \geq 0$. For any $0 \leq m \leq N$, any $\lambda > 0$ and any $p \geq 1$, we obtain*

$$\begin{aligned} \lambda^p \mathbb{P}(X_{m,N}^* \geq \lambda \mid \mathcal{F}_m) &\leq \mathbb{E}[|X_N|^p \mid \mathcal{F}_m] \leq \mathbb{E}[(X_{m,N}^*)^p \mid \mathcal{F}_m] \\ &\leq \left(\frac{p}{p-1} \right)^p \mathbb{E}[|X_N|^p \mid \mathcal{F}_m] \text{ a.s.}, \end{aligned}$$

and for any $p > 1$,

$$\mathbb{E}[|X_N|^p \mid \mathcal{F}_m] \leq \mathbb{E}[(X_{m,N}^*)^p \mid \mathcal{F}_m] \leq \left(\frac{p}{p-1} \right)^p \mathbb{E}[|X_N|^p \mid \mathcal{F}_m] \text{ a.s.}$$

Proof. See section A.1. □

Accordingly, the càdlàg version of Doob's inequalities are extended from the discrete version, even when the conditional expectation is involved.

Theorem 2.3.4 (Càdlàg version of conditional Doob's inequalities). *Let $(X_s)_{s \geq 0}$ be a right-continuous martingale or right-continuous positive submartingale defined on the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. We choose an arbitrary fixed time $T < \infty$ and define $X_{t,T}^* := \sup_{t \leq s \leq T} |X_s| \geq 0$.*

(i) *For any $t \in [0, T]$, $\lambda > 0$ and $p \geq 1$, the conditional Doob's maximal inequality is*

$$\lambda^p \mathbb{P}(X_{t,T}^* \geq \lambda \mid \mathcal{F}_t) \leq \mathbb{E}[|X_T|^p \mid \mathcal{F}_t].$$

(ii) *For any $t \in [0, T]$ and $p > 1$, we get the conditional Doob's L^p -inequality*

$$\mathbb{E}[(X_{t,T}^*)^p \mid \mathcal{F}_t] \leq \left(\frac{p}{p-1} \right)^p \mathbb{E}[|X_T|^p \mid \mathcal{F}_t].$$

The last conditional inequality in which we are interested is conditional Burkholder-Davis-Gundy (BDG) inequality. .

Definition 2.3.5 (Domination relation, P162 in [79]). *Let X_t be a non-negative, integrable and adapted right-continuous process with $X_s = 0$ for all $s \leq t$. Let A_t be an integrable process with $A_s = 0$ for all $s \leq t$. Then X_t is dominated by A_t if for all bounded stopping times $\tau \geq t$,*

$$\mathbb{E}[X_\tau \mid \mathcal{F}_t] \leq \mathbb{E}[A_\tau \mid \mathcal{F}_t].$$

Lemma 2.3.6 (Probability inequality with domination relation). *If X is dominated by a continuous process A , then for all $x, y > 0$,*

$$\mathbb{P}(X_{t,\infty}^* > x, A_\infty \leq y \mid \mathcal{F}_t) \leq \frac{1}{x} \mathbb{E}[A_\infty \wedge y \mid \mathcal{F}_t] \quad a.s.,$$

where $X_{t,\infty}^* := \sup_{s \geq t} X_s$.

Proof. See section A.1. □

Lemma 2.3.7 (L^k inequality with domination relation with $k \in [0, 1)$). *Under the hypothesis of Lemma 2.3.6, for any $k \in [0, 1)$,*

$$\mathbb{E}[(X_{t,\infty}^*)^k \mid \mathcal{F}_t] \leq \frac{2-k}{1-k} \mathbb{E}[(A_\infty)^k \mid \mathcal{F}_t] \quad a.s.$$

Proof. See section A.1. □

Theorem 2.3.8 (Conditional BDG). *Let $(M_t)_{t \geq 0}$ be a continuous local martingale vanishing at zero, defined in the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Given any stopping time $\tau \geq t$, we have constants $C_{1,p}$ and $C_{2,p}$ such that for any $p > 0$,*

$$\begin{aligned} C_{1,p} \mathbb{E} \left[(\langle M \rangle_\tau - \langle M \rangle_t)^{\frac{p}{2}} \mid \mathcal{F}_t \right] &\leq \mathbb{E} \left[\sup_{t \leq s \leq \tau} |M_s - M_t|^p \mid \mathcal{F}_t \right] \\ &\leq C_{2,p} \mathbb{E} \left[(\langle M \rangle_\tau - \langle M \rangle_t)^{\frac{p}{2}} \mid \mathcal{F}_t \right] \quad a.s. \end{aligned}$$

Proof. If we can prove the special case $\tau = \infty$, it is possible to obtain the general case by replacing M by M^τ . Furthermore, provided that the BDG inequality holds for bounded M and the constants $C_{1,p}$ and $C_{2,p}$ are independent of the bound on M , we can derive the same result for unbounded M by using the following trick. We replace M by M^{τ_n} , where $\tau_n := \inf\{s \geq t \mid |M_s - M_t| = n\}$. Consequently M^{τ_n} satisfies the conditional BDG inequality, while we let $n \rightarrow \infty$ and obtain the desired result by the conditional monotone convergence theorem.

The remaining task is to prove for the following inequalities (which holds almost surely),

$$C_{1,p}\mathbb{E}\left[\left(\langle M \rangle_\infty - \langle M \rangle_t\right)^{\frac{p}{2}} \middle| \mathcal{F}_t\right] \leq \mathbb{E}\left[\sup_{t \leq s \leq \infty} |M_s - M_t|^p \middle| \mathcal{F}_t\right], \quad p \geq 4, \quad (2.1)$$

$$C_{1,p}\mathbb{E}\left[\left(\langle M \rangle_\infty - \langle M \rangle_t\right)^{\frac{p}{2}} \middle| \mathcal{F}_t\right] \leq \mathbb{E}\left[\sup_{t \leq s \leq \infty} |M_s - M_t|^p \middle| \mathcal{F}_t\right], \quad 0 \leq p \leq 4, \quad (2.2)$$

$$\mathbb{E}\left[\sup_{t \leq s \leq \infty} |M_s - M_t|^p \middle| \mathcal{F}_t\right] \leq C_{2,p}\mathbb{E}\left[\left(\langle M \rangle_\infty - \langle M \rangle_t\right)^{\frac{p}{2}} \middle| \mathcal{F}_t\right], \quad p \geq 2, \quad (2.3)$$

$$\mathbb{E}\left[\sup_{t \leq s \leq \infty} |M_s - M_t|^p \middle| \mathcal{F}_t\right] \leq C_{2,p}\mathbb{E}\left[\left(\langle M \rangle_\infty - \langle M \rangle_t\right)^{\frac{p}{2}} \middle| \mathcal{F}_t\right], \quad 0 \leq p \leq 2. \quad (2.4)$$

where M is a continuous bounded local martingale with $M_0 = 0$.

To deal with (2.3), we let $p \geq 2$ and fix $s \geq t$. As an application of Itô's lemma and Hölder's inequality, we have

$$\begin{aligned} \mathbb{E}[|M_s - M_t|^p \mid \mathcal{F}_t] &\leq \frac{p(p-1)}{2} \mathbb{E}\left[\sup_{t \leq u \leq s} |M_u - M_t|^p \middle| \mathcal{F}_t\right]^{\frac{p-2}{p}} \\ &\quad \times \mathbb{E}\left[\left(\langle M \rangle_s - \langle M \rangle_t\right)^{\frac{p}{2}} \middle| \mathcal{F}_t\right]^{\frac{2}{p}}, \end{aligned}$$

while the conditional Doob's L^p inequality implies

$$\mathbb{E}\left[\sup_{t \leq u \leq s} |M_u - M_t|^p \middle| \mathcal{F}_t\right] \leq \left(\frac{p(p-1)}{2} \left(\frac{p}{p-1}\right)^p\right)^{\frac{p}{2}} \mathbb{E}\left[\left(\langle M \rangle_s - \langle M \rangle_t\right)^{\frac{p}{2}} \middle| \mathcal{F}_t\right].$$

Finally, we force $s \rightarrow \infty$ and interchange expectation and limit, as the supremum process and quadratic variation are increasing processes.

Next, we focus on (2.1) and fix $p \geq 4$. It is sufficient to prove this inequality, in the case where $\langle M \rangle$ is bounded and $C_{1,p}$ is independent of the bound on $\langle M \rangle$. As a result, the general case can be extended by considering a stopping time $r_n := \inf\{s \geq t \mid \langle M \rangle_s - \langle M \rangle_t = n\}$ and replacing M by M^{r_n} . Rewriting $\langle M \rangle_s - \langle M \rangle_t = (M_s - M_t)^2 - 2 \int_t^s (M_u - M_t) dM_t$ for all $s \geq t$, we can bound the quadratic variation part by

$$\begin{aligned} \mathbb{E}\left[\left(\langle M \rangle_\infty - \langle M \rangle_t\right)^{\frac{p}{2}} \middle| \mathcal{F}_t\right] &\leq a_p \left(\mathbb{E}\left[\sup_{s \geq t} |M_s - M_t|^p \middle| \mathcal{F}_t\right]^{\frac{1}{2}} \mathbb{E}\left[\left(\langle M \rangle_\infty - \langle M \rangle_t\right)^{\frac{p}{2}} \middle| \mathcal{F}_t\right]^{\frac{1}{2}} \right. \\ &\quad \left. + \mathbb{E}\left[\sup_{s \geq t} |M_s - M_t|^p \middle| \mathcal{F}_t\right] \right), \end{aligned}$$

where a_p is a universal constant depending on p only. Consider the equation $x^2 - a_p xy \leq a_p y^2$, we can deduce that $(x - \frac{1}{2}a_p y)^2 \leq (a_p + \frac{a_p^2}{4})y^2$. Plugging in

$$x = \mathbb{E}\left[\left(\langle M \rangle_\infty - \langle M \rangle_t\right)^{\frac{p}{2}} \middle| \mathcal{F}_t\right]^{\frac{1}{2}}, \quad y = \mathbb{E}\left[\sup_{s \geq t} |M_s - M_t|^p \middle| \mathcal{F}_t\right]^{\frac{1}{2}},$$

we get

$$\mathbb{E} \left[(\langle M \rangle_\infty - \langle M \rangle_t)^{\frac{p}{2}} \mid \mathcal{F}_t \right] \leq \left(\sqrt{a_p + \frac{a_p^2}{4}} + \frac{a_p}{2} \right) \mathbb{E} \left[\sup_{s \geq t} |M_s - M_t|^p \mid \mathcal{F}_t \right].$$

When $0 < p < 2$, we choose $X_s := |M_s - M_t|^2$, $A_s := C_{2,2}(\langle M \rangle_s - \langle M \rangle_t)$ and set $k = \frac{p}{2}$. The inequality (2.3) has been able to show the domination relation between X and A , hence we apply Lemma 2.3.7 to conclude the (2.4). Finally, we apply the similar procedure but with $X_s := C_{1,4}(\langle M \rangle_s - \langle M \rangle_t)^2$, $A_s := |M_s - M_t|^4$ and $k = \frac{p}{4}$. The inequality (2.1) has shown that X is dominated by A , therefore we conclude (2.2) by plugging in X, A into Lemma 2.3.7. $\square$

Remark 2.3. *An alternative strategy to show the conditional BDG inequality relies on changing the martingale and the filtration. Given any stopping time τ_1 , we choose an arbitrary set $A \in \mathcal{F}_{\tau_1}$. Rewriting $M'_s := M_{s+\tau_1} - M_{\tau_1}$ and $\mathcal{F}'_s := \mathcal{F}_{s+\tau_1}$ yields M'_s is a martingale with respect to filtration $\mathcal{F}'_s$ under the measure $\mathbb{P}'$, where $\mathbb{P}'$ is defined as $d\mathbb{P}' = \mathbb{P}(A)^{-1} \mathbb{1}_A d\mathbb{P}$. More explicitly, we implement Bayes' rule to derive*

$$\mathbb{E}'[M'_s \mid \mathcal{F}'_u] = \mathbb{E}[M_{s+\tau_1} - M_{\tau_1} \mid \mathcal{F}_{u+\tau_1}] = M_{u+\tau_1} - M_{\tau_1}$$

provided $s \geq u \geq 0$. By observing $\langle M' \rangle_s = \langle M \rangle_{s+\tau_1} - \langle M \rangle_{\tau_1}$, we conclude that

$$\begin{aligned} \mathbb{E} \left[\mathbb{E}[(\langle M \rangle_{\tau_1+\tau_2} - \langle M \rangle_{\tau_1})^{\frac{p}{2}} \mid \mathcal{F}_{\tau_1}] \mathbb{1}_A \right] &= \mathbb{E}' \left[\frac{\mathbb{P}(A)}{\mathbb{1}_A} \mathbb{E}'[\langle M' \rangle_{\tau_2}^{\frac{p}{2}} \mid \mathcal{F}'_0] \mathbb{1}_A \right] \\ &= \mathbb{E}' \left[\langle M' \rangle_{\tau_2}^{\frac{p}{2}} \right] \mathbb{P}(A), \end{aligned}$$

while for the supremum process, we are able to deduce

$$\begin{aligned} \mathbb{E} \left[\mathbb{E} \left[\sup_{\tau_1 \leq s \leq \tau_1+\tau_2} |M_{\tau_1+s} - M_{\tau_1}|^p \mid \mathcal{F}_{\tau_1} \right] \mathbb{1}_A \right] &= \mathbb{E}' \left[\frac{\mathbb{P}(A)}{\mathbb{1}_A} \mathbb{E}' \left[\sup_{\tau_1 \leq s \leq \tau_1+\tau_2} |M'_s|^p \mid \mathcal{F}'_0 \right] \mathbb{1}_A \right] \\ &= \mathbb{E}' \left[\sup_{\tau_1 \leq s \leq \tau_1+\tau_2} |M'_s|^p \right] \mathbb{P}(A). \end{aligned}$$

The classical BDG implies

$$C_{1,p} \mathbb{E}' \left[\langle M' \rangle_{\tau_2}^{\frac{p}{2}} \right] \mathbb{P}(A) \leq \mathbb{E}' \left[\sup_{\tau_1 \leq s \leq \tau_1+\tau_2} |M'_s|^p \right] \mathbb{P}(A) \leq C_{2,p} \mathbb{E}' \left[\langle M' \rangle_{\tau_2}^{\frac{p}{2}} \right] \mathbb{P}(A),$$

hence the result follows for conditional BDG.

³Due to the nature of conditional expectations, it is suffice to consider A such that $\mathbb{P}(A) > 0$

Remark 2.4. *An important observation is the condition imposed on the process A within Lemma 2.3.6 can be relaxed to being increasing and càdlàg, as it is sufficient to guarantee the existence of $A_\infty = \lim_{t \rightarrow \infty} A_t$ [66]. Among its proof, $\{A_\infty \leq y\} = \{\tau_y = \infty\}$ up to null sets, which implies Theorem 2.3.8 holds for càdlàg local martingales M if the quadratic variation is suitably defined for such M . In fact, [39] has included the standard BDG inequality for càdlàg martingales, thus the conditional version can also be deduced from remark 2.3.*

2.3.2 BMO martingales

To begin with, we state the celebrated John-Nirenberg inequality in the theory of BMO martingales.

Theorem 2.3.9 (John-Nirenberg (1), Theorem 2.1 [50]). *If M is a càdlàg martingale and $\|M\|_{BMO_1(\mathbb{R})} < \frac{1}{4}$, then*

$$\mathbb{E}[\exp(|M_T - M_{\tau_-}|) \mid \mathcal{F}_\tau] \leq \frac{1}{1 - 4\|M\|_{BMO_1(\mathbb{R})}},$$

where $\tau \leq T$ is a stopping time.

Meyer gave a bound on the exponential of the quadratic variation of M , while the following version is sometimes also known as the John-Nirenberg inequality.

Theorem 2.3.10 (John-Nirenberg (2), Theorem 2.2 in [50]). *If M is a càdlàg martingale and $\|M\|_{BMO_2(\mathbb{R})} < 1$, then*

$$\mathbb{E}[\exp(\langle M \rangle_T - \langle M \rangle_{\tau_-}) \mid \mathcal{F}_\tau] \leq \frac{1}{1 - \|M\|_{BMO_2(\mathbb{R})}^2},$$

where $\tau \leq T$ is a stopping time.

As an application of John-Nirenberg inequality, the BMO norms can be shown that they are equivalent for all $1 \leq p < \infty$, suggesting that the parameter p is not important.

Lemma 2.3.11 (Equivalence of BMO norms, Corollary 2.1 [50]). *Suppose $1 \leq p < \infty$. For any uniformly integrable martingale M , there is a positive constant $C_p = 8 \cdot (2(p!))^{\frac{1}{p}}$ such that*

$$\|M\|_{BMO_1(\mathbb{R})} \leq \|M\|_{BMO_p(\mathbb{R})} \leq C_p \|M\|_{BMO_1(\mathbb{R})}.$$

There is more than one approach to verify whether M is a $BMO_p(\mathbb{R})$ martingale.

Lemma 2.3.12 (Deterministic time for BMO, Remark A.8.3 in [15] or Remark 76.4 in [21]). *A càdlàg martingale M_t lies in space $BMO_p(\mathbb{R})$ if and only if*

$$\sup_{t \in [0, T]} |M_t - M_{t-}| \leq C, \quad \sup_{t \in [0, T]} \mathbb{E} [|M_T - M_t|^p \mid \mathcal{F}_t] \leq C,$$

where we set $M_{0-} = M_0$ as a convention.

In other words, we can replace the random stopping time by a deterministic time at a cost of controlled jump size. By using the energy inequality arising from the theory of potentials (see section B.2, and for references therein [66]),

$$\mathbb{E} [(\langle M \rangle_T - \langle M \rangle_{t-})^n \mid \mathcal{F}_t] \leq n! \|M\|_{BMO_2(\mathbb{R})}^{2n}, \quad n \in \mathbb{N}, \quad (2.5)$$

we conclude that estimating the conditional increment of the quadratic variation is equivalent to estimating the BMO norm (by the conditional BDG for càdlàg martingales, see remark 2.4). Therefore, to characterise a BMO càdlàg martingale, we essentially require:

Corollary 2.3.13 (Quadratic variation of càdlàg BMO martingale). *A càdlàg martingale M is in $BMO_p(\mathbb{R})$ for some $1 \leq p < \infty$ (hence for all such p) if and only if there exist some constant $C > 0$ such that*

$$\sup_{t \in [0, T]} |M_t - M_{t-}| + \sup_{t \in [0, T]} \mathbb{E} \left[(\langle M \rangle_T - \langle M \rangle_t)^{\frac{p}{2}} \mid \mathcal{F}_t \right] \leq C \quad (2.6)$$

We next state a useful result called the reverse Hölder inequality, which controls the L^q norm of the stochastic exponential of BMO martingales.

Theorem 2.3.14 (Reverse Hölder inequality, Theorem 3.1 and Remark 3.5 in [50]). *Let M be a continuous local martingale. If $\|M\|_{BMO_2(\mathbb{R})} < \hat{\Xi}(p)$ where $p \in (1, \infty)$ and $\hat{\Xi}$ defined below, then for every stopping time τ*

$$\mathbb{E}[\mathcal{E}(M)_\tau^p \mid \mathcal{F}_\tau] \leq C_p \mathcal{E}(M)_\tau^p, \quad (R_p)$$

subject to $\hat{\Xi}(p) = \left(1 + \frac{1}{p^2} \log \left(\frac{2p-1}{2(p-1)} \right) \right)^{\frac{1}{2}} - 1$.

Alternatively, if M is only a local martingale, we suppose $\inf_{t \in [0, T]} (M_t - M_{t-}) > -1 + \epsilon$ for some $\epsilon \in (0, 1]$. Then $\mathcal{E}(M)$ is a uniformly integrable martingale and equation (R_p) holds and the maximum value of such p depends on $\|M\|_{BMO_2(\mathbb{R})}$ only⁴.

⁴A review of the relationship between the reverse Hölder inequality and the BMO property of M can be found in Grandits and Rheinländer [35, Appendix].

It is worth noticing that $\hat{\Xi} : (1, \infty) \mapsto \mathbb{R}_{++}$ is a positive, surjective and strictly decreasing function, which implies that there must exist a suitable $p > 1$ such that $\|M\|_{\text{BMO}(\mathbb{R})} < \hat{\Xi}(p)$ is satisfied. If $\|M\|_{\text{BMO}_2(\mathbb{R})} < \hat{\Xi}(p)$ for some p , then $\mathbb{E}[\frac{\mathcal{E}(M)_T^r}{\mathcal{E}(M)_t^r} \mid \mathcal{F}_t] \leq C_p^{\frac{r}{p}}$ for all $1 < r < p$. However, the maximum value p that can be taken is usually quite small, as $\hat{\Xi}(2) = \sqrt{1 + \frac{1}{4} \log(\frac{3}{2})} - 1 = 0.049$ and $\hat{\Xi}(1.1) = 0.57$. As a consequence of the reverse Hölder's inequality, we are allowed to switch to a new measure $\mathbb{Q}$ at a cost of some constant.

Corollary 2.3.15 (Switching to new a measure). *Suppose M is a BMO martingale, let measure $\mathbb{Q}$ characterised by $\frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} = \mathcal{E}(M)_t$. Given an integrable random variable X which is $\mathcal{F}_T$ -measurable, it satisfies*

$$\mathbb{E}^{\mathbb{Q}}[X \mid \mathcal{F}_t] \leq C_q \mathbb{E}^{\mathbb{P}}[|X|^{q^*} \mid \mathcal{F}_t]^{\frac{1}{q^*}},$$

where q^* is the conjugate of q and q is a value satisfying condition (R_p) .

Moreover, $N_t := \int_0^t Y_s \cdot dB_s^{\mathbb{Q}}$ is a BMO martingale under $\mathbb{Q}$ with $B_t^{\mathbb{Q}}$ as the $\mathbb{Q}$ -Brownian motion, provided $\tilde{N}_t := \int_0^t Y_s \cdot dB_s$ is a BMO martingale under $\mathbb{P}$.

Proof. As a simple application of reverse Hölder inequality and Hölder's inequality,

$$\mathbb{E} \left[\frac{\mathcal{E}(M)_T}{\mathcal{E}(M)_t} X \mid \mathcal{F}_t \right] \leq \mathbb{E} \left[\left(\frac{\mathcal{E}(M)_T}{\mathcal{E}(M)_t} \right)^q \mid \mathcal{F}_t \right] \mathbb{E}[|X|^{q^*} \mid \mathcal{F}_t]^{\frac{1}{q^*}} \leq C_q \mathbb{E}^{\mathbb{P}}[|X|^{q^*} \mid \mathcal{F}_t]^{\frac{1}{q^*}},$$

while the second statement can be derived by implementing Theorem 2.3.8

$$\mathbb{E}^{\mathbb{Q}} \left[(N_T - N_t)^2 \mid \mathcal{F}_t \right] \leq C_q \mathbb{E}^{\mathbb{P}} \left[\left(\int_t^T |Y_s|^2 ds \right)^{q^*} \mid \mathcal{F}_t \right]^{\frac{1}{q^*}} \leq C_q \|\tilde{N}\|_{\text{BMO}_{2q^*}(\mathbb{P}, \mathbb{R})}^2.$$

□

By the martingale representation theorem, all L^2 martingales which are adapted to $\mathcal{F}_t$ can be written in terms of Itô's stochastic integrals. This motivates us to focus on the integrand rather than the martingale itself. If the martingale M is a stochastic integral with a discrete integrand, we have the following characterisation.

Lemma 2.3.16 (BMO for Itô's integrals with discrete integrands). *Let $M_t = \int_0^t K_s \cdot dB_s$ be a L^2 -martingale, with $K_s \in \mathbb{R}^d$. Suppose there exists a sequence of $\mathbb{R}^d$ -valued random variables K^i which are $\mathcal{F}_{t_i}$ -measurable respectively, such that*

$$K_s = K^{(0)} \mathbb{1}_{\{0\}}(s) + \sum_{i=0}^{N-1} K^{(i)} \mathbb{1}_{(t_i, t_{i+1}]}(s).$$

where $0 = t_0 < t_1 < \dots < t_N = T$. Then M is a BMO martingale if and only if

$$\sup_{0 \leq i \leq N-1} \left\| \mathbb{E} \left[\sum_{j=i}^{N-1} |K^{(j)}|^2 (t_{j+1} - t_j) \mid \mathcal{F}_{t_i} \right] \right\|_{L^\infty} \leq C.$$

for some $C > 0$.

Proof. We only need to prove the if part. Observing that M is a continuous martingale, we only need to consider the conditional quadratic variation. Given a fixed $t \in [t_k, t_{k+1})$, we make use of the fact that $K^{(k)}$ is $\mathcal{F}_{t_k}$ -measurable to obtain

$$\begin{aligned} \mathbb{E} \left[\int_t^T |K_s|^2 ds \mid \mathcal{F}_t \right] &= \mathbb{E} \left[\int_t^{t_{k+1}} |K^{(k)}|^2 ds \mid \mathcal{F}_t \right] + \mathbb{E} \left[\sum_{j=k+1}^{N-1} |K^{(j)}|^2 h_j ds \mid \mathcal{F}_t \right] \\ &\leq \mathbb{E} \left[|K^{(k)}|^2 h_k \mid \mathcal{F}_{t_k} \right] + \mathbb{E} \left[\mathbb{E} \left[\sum_{j=k+1}^{N-1} |K^{(j)}|^2 h_j ds \mid \mathcal{F}_{t_{k+1}} \right] \mid \mathcal{F}_t \right] \leq 2C. \end{aligned}$$

Therefore, $\|M\|_{\text{BMO}_2(\mathbb{R})} = \|K\|_{\mathbb{H}_{\text{BMO}}^2(\mathbb{R}^d)} \leq 2C$. $\square$

We close this section by the following remark on BMO martingales.

Remark 2.5. Given a process $K_t \in \mathbb{R}$ with $K_t \in \mathbb{H}_{\text{BMO}}^2(\mathbb{R})$, for any $u \in \mathbb{R}^d$ with $|u| = 1$, $Ku \in \mathbb{H}_{\text{BMO}}^2(\mathbb{R}^d)$ since $\|Ku\|_{\mathbb{H}_{\text{BMO}}^2(\mathbb{R}^d)} = \|K_t\|_{\mathbb{H}_{\text{BMO}}^2(\mathbb{R})}$. We denote $M_t(u) := \int_0^t K_s u \cdot dB_s$ and $\|M(u)\|_{\text{BMO}_2(\mathbb{R})}$ remains constant for all $u \in \mathbb{R}^d$ with $|u| = 1$. Observing that the constraint corresponds to the reverse Hölder inequality (Theorem 2.3.14 for $\mathcal{E}(M(u))$ depends on $\|M_t(u)\|_{\text{BMO}_2(\mathbb{R})}$ only, the equation (R_p) holds for $M_t(u)$ uniformly in u with $|u| = 1$.

2.4 BSDE

Given a the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^B)_{t \in [0, T]}, \mathbb{P})^5$, a BSDE, in its integral form, is

$$Y_t = \xi + \int_t^T f(t, Y_t, Z_t) dt + \int_t^T (Z_t)^\top dB_t \quad (2.7)$$

where f is a $\mathbb{R}^l$ -valued, $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}^l) \otimes \mathcal{B}(\mathbb{R}^{d \times l})$ measurable map and ξ is a $\mathbb{R}^l$ -valued $\mathcal{F}_T$ -measurable random variable. The solution to BSDE (2.7) is a pair (Y, Z) such that $(Y_t)_{t \in [0, T]}$ is a $\mathbb{R}^l$ -valued continuous adapted process, while $(Z_t)_{t \in [0, T]}$ is a predictable $\mathbb{R}^{d \times l}$ -valued process satisfying $\int_0^T |Z_s|^2 ds < \infty$ ⁶ $\mathbb{P}$ -almost surely. The

⁵Here we specify $\mathcal{F}_t = \mathcal{F}_t^B$

⁶ Z has value in $\mathbb{R}^{d \times l}$, hence $|Z_s|$ is the Frobenius norm of the matrix Z .

terminal condition ξ and the function f characterise the solution (Y, Z) , hence (ξ, f) is also known as the generator of BSDE (2.7).

Bismut [4] first considered linear BSDEs to study stochastic control problems, while Pardoux and Peng brought BSDE to the attention of many scholars in their seminal paper [73], where they implemented the fixed point method to prove the well-posedness of BSDE. In 1991, Peng [75] established the connection between decoupled forward-backward stochastic differential equation (FBSDE) and PDE. The decoupled FBSDE is essentially BSDE with terminal condition $\xi = g(X_T)$ and the driver of the form $f(t, X_t, Y_t, Z_t)$, where $f : [0, T] \times \mathbb{R}^m \times \mathbb{R}^l \times \mathbb{R}^{d \times l} \rightarrow \mathbb{R}^l$ is a deterministic function and X_t is a forward SDE with values in $\mathbb{R}^m$. More precisely, given a decoupled FBSDE

$$\begin{aligned} X_s^{t,x} &= x + \int_t^s b(u, X_u^{t,x}) \, ds + \int_t^s \sigma(u, X_u^{t,x}) \, dB_u, \quad X_t^{t,x} = x \in \mathbb{R}^m, \\ Y_s^{t,x} &= g(X_T^{t,x}) + \int_t^T f(u, X_u^{t,x}, Y_u^{t,x}, Z_u^{t,x}) \, du - \int_t^T (Z_u^{t,x})^\top \, dB_u, \end{aligned} \quad (2.8)$$

where $b : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, $\sigma : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$ with $b = (b^1, \dots, b^m)$ and $\sigma = (\sigma_{ij})_{1 \leq i \leq m, 1 \leq j \leq d}$, its corresponding semi-linear PDE problem is

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + \mathcal{L}u(t, x) + f\left(t, x, u(t, x), (\nabla_x u \sigma)^\top(t, x)\right) = 0, \\ u(T, x) = g(x), \end{cases} \quad (2.9)$$

where $u : \mathbb{R}_+ \times \mathbb{R}^m \mapsto \mathbb{R}^l$, $u = (u^1, \dots, u^l) : \mathbb{R}^m \rightarrow \mathbb{R}^l$, $g = (g^1, \dots, g^l)$ and the operator $\mathcal{L}$ is defined as

$$\mathcal{L} := \sum_{i=1}^m b_i(t, x) \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^m (\sigma \sigma^\top)_{ij}(t, x) \frac{\partial^2}{\partial x_i \partial x_j}, \quad \mathcal{L}u = (\mathcal{L}u^1, \dots, \mathcal{L}u^l).$$

Under the standard Lipschitz condition of functions f, g, σ and b , Peng [75] demonstrated $u(t, x) = Y_t^{t,x}$, provided $u \in C^{1,2}([0, T] \times \mathbb{R}^m; \mathbb{R}^l)$ solves the PDE (2.9). Regarding the reverse side, Pardoux and Peng [74] proved that, under sufficient regularity of the terminal function g , driver f , the drift b and the volatility σ (namely, being a member of C^3), the solution to the decoupled FBSDE is exactly the classical solution to a system of quasi-linear (possibly degenerate) parabolic equation. Meanwhile, if the regularity assumptions on g, f, b and σ are reduced to being Lipschitz and $l = 1$ i.e. we focus on one equation instead of a system of PDEs, it is also shown in Pardoux and Peng [74] that $u(t, x) := Y_t^{t,x}$ becomes a unique viscosity solution to the PDE problem associated to the FBSDE.

Numerous efforts has been made to explore BSDE with new constraints on the generator (ξ, f) . Ma and Yong [62] introduced the coupled FBSDE, where the drift

and volatility of forward SDE depend on the backward component (Y, Z) . Another extension to Pardoux and Peng [73] is relaxing the Lipschitz condition of the driver f . Lepeltier and San Martin [56] studied the equation (2.7) in dimension one with continuous driver and linear growth in (Y, Z) , while Darling et al. [19] focused on the BSDE with driver satisfying a monotonicity condition. Finally, quadratic (or super-quadratic) backward stochastic differential equation has attracted lots of attention in the past two decades (see reference therein [9, 8, 20, 88]).

2.4.1 Lipschitz BSDEs and quadratic BSDEs

Regarding the Lipschitz BSDE, we recall the seminal work on the existence and uniqueness.

Theorem 2.4.1 (Theorem 2.1 in [25]). *Suppose $f(\cdot, 0, 0) \in \mathbb{H}^2(\mathbb{R}^l)$, $\xi \in L^2(\mathbb{R}^l)$ and f is uniformly Lipschitz in y, z i.e. there exists $L_f > 0$ such that*

$$|f(t, y, z) - f(t, y', z')| \leq L_f (|y - y'| + |z - z'|),$$

for all $y, y' \in \mathbb{R}^l, z, z' \in \mathbb{R}^{d \times l}, (t, \omega) \in [0, T] \times \Omega$. Then there exists a unique solution $(Y, Z) \in \mathbb{S}^2(\mathbb{R}^l) \times \mathbb{H}^2(\mathbb{R}^{d \times l})$ to the BSDE (2.7).

A particular case is the linear BSDE, where a closed form solution can be derived based on this problem.

Proposition 2.4.2 (Proposition 2.2 in [25]). *Let (ξ, f) be the generator of (2.7) such that $f(\cdot, 0, 0) \in \mathbb{H}^2(\mathbb{R}^l), \xi \in L^2(\mathbb{R}^l)$. If $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ are of the form*

$$f(t, y, z) = \phi_t + \alpha_t + z \cdot \beta_t, \text{ i.e. } f(\omega, t, y, z) = \phi_t(\omega) + \alpha_t(\omega) + z \cdot \beta_t(\omega), \quad \forall \omega \in \Omega, \quad (2.10)$$

such that $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}^d$ are bounded processes and $\phi_t \in \mathbb{H}^2(\mathbb{R}^d)$, then there exists a unique solution $(Y, Z) \in \mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}^d)$ to the BSDE (2.7) and

$$Y_t = \mathbb{E} \left[\xi \Gamma_T^t + \int_t^T \Gamma_s^t \phi_s \, ds \mid \mathcal{F}_t \right]$$

with Γ_s^t satisfying the SDE

$$d\Gamma_s^t = \Gamma_s^t \alpha_s \, ds + \beta_s^* \, dB_s, \quad \Gamma_t^t = 1, \quad s \geq t.$$

As we mentioned in the first chapter, the Z component can be recovered as version of Malliavin derivative, in the sense that there exists a version of $Z_t = D_t Y_t$ in $[0, T] \times \Omega$, by [24, Proposition 5.3].

We are now in the position to review quadratic backward stochastic differential equations (QBSDE or quadratic BSDE). This type of equation received a lot of attention, because it appears in many stochastic control problems when the utility functions is in the form of an exponential or the cost is quadratic. Most of results in this area restrict the dimension of Y to be one, such that a version of maximum principle or a comparison theorem is allowed to be applied to establish a bound on Y . Kobylanski [52] first developed a proof of the existence and uniqueness to such BSDE via PDE-based methods. More explicitly, she added the following constraint on the generators (ξ, f) with $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$.

Assumption [A1]. $\|\xi\|_{L^\infty(\mathbb{R})} + \|f(\cdot, 0, 0)\|_{S^\infty(\mathbb{R})} \leq K_0$ for some $K_0 \geq 0$. Moreover, f is uniformly Lipschitz in y and quadratic growth in z , namely

$$|f(t, y, z) - f(t, y', z')| \leq L_f |y - y'| + K_f (1 + |z| + |z'|) |z - z'|$$

for some constants $L_f > 0$ and $K_f > 0$ independent of y, z , for all $y, y' \in \mathbb{R}$ and $z, z' \in \mathbb{R}^d$.

After people have realised the solution Y is uniformly bounded, Hu et al. [45] utilised the theory of BMO martingales and first characterised the Z component with a sharp estimate. Briand and Hu [8] proved the existence of solution when the terminal condition ξ is unbounded but has bounded exponential moments, while they demonstrated the uniqueness of the solution if the driver is convex in [9]. For multidimensional BSDE with quadratic growth in Z , all relevant works required a special structure, for example Hu and Tang [44]. Meanwhile, Frei and Dos Reis [28] gave an example based on [50], illustrating that the solution may not exist in the multidimensional case even when bounded terminal conditions are imposed.

In this thesis, we will spend most of the time on the BSDE satisfying assumption [A1] in dimension one. We quote the results on QBSDEs before we move on to the numerics.

Theorem 2.4.3 (Well-posedness for QBSDE, [7]). *Suppose (ξ, f) satisfies assumption [A1]. Let (ξ, f) be the generator of the following BSDE*

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s \cdot B_s, \quad (2.11)$$

then the solution (Y, Z) to equation (2.11) exists and it is unique. Moreover, $(Y, Z) \in S^\infty(\mathbb{R}) \times \mathbb{H}_{BMO}^2(\mathbb{R}^d)$ and

$$\|Z\|_{\mathbb{H}_{BMO}^2(\mathbb{R}^d)} \leq \frac{1 \vee T}{3} \left(1 + \frac{4\|f(\cdot, 0, 0)\|_{S^\infty(\mathbb{R})} + 2L_f \|Y\|_{S^\infty(\mathbb{R})}}{K_f} \right) \exp(3K_f \|Y\|_{S^\infty(\mathbb{R})}),$$

$$\|Y\|_{\mathbb{S}^\infty(\mathbb{R})} \leq \exp(L_f T) (\|\xi\|_{L^\infty} + \|f(\cdot, 0, 0)\|_{\mathbb{S}^\infty(\mathbb{R})} T).$$

The comparison theorem is also listed below.

Theorem 2.4.4 (Comparison theorem for QBSDE, Theorem 2.6 in [52] or Theorem 7.3.1 in [91]). *Let (ξ_i, f_i) satisfies assumption [A1] for all $i = 1, 2$, suppose (Y^i, Z^i) is the corresponding solution for equation (2.11). If $f_1(t, Y_t^i, Z_t^i) \geq f_2(t, Y_t^i, Z_t^i)$ for either $i = 1$ or $i = 2$ a.s. and $\xi_1 \geq \xi_2$ a.s, then $Y_t^1 \geq Y_t^2$ for all $0 \leq t \leq T$ a.s.*

Other than using the comparison theorem stated above, we are in a position to introduce a stability theorem, which also characterises the uniqueness of such BSDE as well as specifies the error for further numerical schemes. It has first discovered in [47], where we prove a similar result.

Proposition 2.4.5 (Stability on $\mathbb{S}^2(\mathbb{R}) \times \mathbb{H}^2(\mathbb{R}^d)$, modification of Theorem 6.2 in [47]). *Let the pair (ξ_i, f_i) satisfies assumption [A1] for all $i = 1, 2$, suppose (Y^i, Z^i) is the corresponding solution to (2.11). Denoting $\Delta\xi = \xi^1 - \xi^2$, $\Delta f(y, z) = f^1(t, y, z) - f^2(t, y, z)$, we have, for all $p > 1$,*

$$\|Y^1 - Y^2\|_{\mathbb{S}^p(\mathbb{R})}^p + \|Z^1 - Z^2\|_{\mathbb{H}^p(\mathbb{R}^d)}^p \leq C \mathbb{E} \left[|\Delta\xi|^{p\check{q}} + \left(\int_0^T |\Delta f(s, Y_s^2, Z_s^2)| \, ds \right)^{p\check{q}} \right]^{\frac{1}{\check{q}}}$$

for some constants $C > 0$ and $\check{q} > 1$.

Proof. See section A.2. □

2.4.2 BTZ scheme

From the numerical point of view⁷, we are mostly interested in one-dimensional BSDE

$$\begin{aligned} dY_t &= -f(t, Y_t, Z_t) \, dt + Z_t \cdot dB_t \\ Y_T &= \xi, \end{aligned} \tag{2.12}$$

such that the terminal condition $\xi : \Omega \mapsto \mathbb{R}$ is an $\mathcal{F}_T$ -measurable random variable and generator $f : \Omega \times \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable map. Therefore, $(Y, Z) \in \mathbb{R} \times \mathbb{R}^d$ and we will stick with such problems when we discuss numerical schemes for BSDEs.

⁷In general, most numerical schemes for one-dimensional BSDE can easily be extended to multi-dimensional BSDEs, as long as the multi-dimensional BSDE is well-posed

Under the standard Lipschitz condition, the BTZ scheme was developed by Bouchard and Touzi [5] for decoupled FBSDE and Zhang [90] for BSDE with terminal condition as a functional, such that the pair (Y, Z) was approximated by the discrete version

$$Y_t^\pi = \mathbb{E}[Y_{t+h}^\pi \mid \mathcal{F}_t] + hf(t, Y_t^\pi, Z_t^\pi), \quad Z_t^\pi = \frac{1}{h} \mathbb{E}[Y_{t+h}^\pi (B_{t+h} - B_t) \mid \mathcal{F}_t]$$

for some partition π and time interval $h > 0$. The scheme itself is implicit in Y and explicit in Z , hence we have to solve an equation, usually by Picard iteration, for Y_t^π for every t . A sketch of the motivation as well as the idea of deriving the BTZ scheme are deferred to Section B.3 for the interest of the reader.

Under standard Lipschitz assumptions, most of the numerical scheme aims to solve decoupled FBSDE (2.8). The main difficulty of solving such BSDE numerically is the pair (Y, Z) can only be represented as a form of conditional expectation in the BTZ scheme. Various methods are developed based on such scheme, including binomial tree methods [10], quantization methods [72, 71], Malliavin calculus [5, 33], non-parametric regression [32, 5], cubature formulas [18] and some high order schemes such as Runge-Kuta methods [11]. Finally, we remark that there are new methods [31, 37, 41, 87, 38, 46] which are based on machine learning.

2.5 Turbulence

Numerous efforts has been done in the theory of turbulence, including E and Van- den Eijnden [23] on passive scalar advection and Jan and Raimond on [49] on classifying different types of flows. Meanwhile, Constantin and Iyer [16] provided an probabilistic representation formula for the Navier-Stokes equation on $\mathbb{R}^n$, which is extended by Fang and Luo [27] to the Navier-Stokes equation on manifolds. Apart from these, the study of statistical characteristics of turbulence also brings a lot of attentions. The research on the statistical properties of turbulent flows can be traced back to the semi-empirical theories of turbulence in the 1920's and 1930's, while the seminal advances in the area include Prandtl [78], von Kármán [86] and Taylor [84, 85]. The goal of statistical fluid mechanics is to provide good descriptions and computational tools for understanding the distributions of the velocity random fields of turbulent fluid flows. Unlike some other unsolved problems in theoretical physics, the equations of motion for fluid dynamics, even for turbulent flows, have been known for over a century. These equations are highly non-linear and non-local partial differential equations, and it is difficult to extract information about the evolution of fluid flows in a deterministic manner. Thus, as a matter of fact, the velocity and vorticity field

of a turbulent flow is better to be considered as a random field arising from either the random initial data or a random external force, or both.

To understand the statistics of turbulent flows, it is desired to know, if it is possible, the evolution of some distributional characteristics of fluid flows. The distribution of a random field such as the velocity field is rather complicated and determining the distribution of turbulent flows is a challenging task even when the initial distribution is known. In 1950's Hopf [43] (see also [68]) derived a functional differential equation for the law of the velocity random field, but his equation involves functional derivatives and is infinite dimensional, which becomes too difficult to extract useful information about turbulence. The dominating approach in the statistical theory of turbulence, initiated in G. I. Taylor's seminal work [84, 85], has been based on the analysis of moment structural functions and relies on the spectral method, for this aspect see for example [3, 68].

Convention and notation. As a general rule and to be consistent with popular conventions in the research of turbulence, we apply Einstein convention i.e. repeated indices are assumed to be summed, when we focus on topics that are *related to turbulence only*.

2.5.1 Velocity field

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, the turbulent flow is described by its velocity $U = (U^1, U^2, U^3)$ and the pressure P , which are random fields and the three dimensional Navier-Stokes equations

$$\begin{aligned} \frac{\partial U^i}{\partial t} + U^j \frac{\partial U^i}{\partial x^j} &= \nu \Delta_x U^i - \frac{\partial P}{\partial x^i}, \\ U(x, 0) &= U_0(x), \end{aligned} \tag{2.13}$$

where $i = 1, 2, 3$ and $\nu \geq 0$ is the viscosity constant, together with the constraint

$$\frac{\partial U^j}{\partial x^j} = 0. \tag{2.14}$$

The initial condition is also treated as a random field on $\mathbb{R}^3$ and each sample path $\omega \in \Omega$ corresponds to a deterministic function $U(x, t; \omega)$, which serves as a solution to equation (2.13) with initial data $U_0(x; \omega)$. We will discuss an ideal case, for the purpose of understanding the local properties of turbulent flows, where the region occupied by the fluid is the entire space $\mathbb{R}^3$. Moreover, without further qualifications, the dynamical variables such as $U(x, t)$ and $P(x, t)$ decay to zero sufficiently fast as $|x|$ tends to infinity. In addition, to avoid technical difficulties, but not in any way

implying that these issues are not important, we will assume that the dynamical variables $U(x, t)$ and $P(x, t)$ are sufficiently smooth functions of (x, t) .

Due to the divergence-free condition (2.14), the pressure term satisfies the following Poisson equation

$$\Delta_x P = -\frac{\partial^2(U^i U^j)}{\partial x^j \partial x^i}.$$

Therefore, according to the Green formula, we have the integral representation

$$P(x, t) = \int_{\mathbb{R}^3} \frac{1}{4\pi|y-x|} \frac{\partial^2(U^i U^j)}{\partial y^j \partial y^i} dy, \quad (2.15)$$

which implies in particular that the distribution of P is completely determined by the distribution of the velocity random field.

2.5.2 Vorticity field

The equation of the vorticity motion for $W = \nabla \wedge U$ is the vorticity equation

$$\frac{\partial W^i}{\partial t} + U^j \frac{\partial W^i}{\partial x^j} = \nu \Delta_x W^i + W^j \frac{\partial U^i}{\partial x^j} \quad (2.16)$$

for $i = 1, 2, 3$. Moreover, $W^i = \varepsilon^{ijk} \frac{\partial}{\partial x^j} U^k$, where ε^{ijk} are the Levi-Civita symbols⁸.

As a standard assumption which is enforced throughout the paper, both $U(x, t)$ and $W(x, t)$ have derivatives in space variable x of any order which decay to zero at infinity sufficiently fast, so that possible boundary terms arising in applications of the Stokes' formula have no contributions in computations below. Therefore, since U is divergence-free, $\Delta_x U = -\nabla \wedge W$, according to Green's formula we may write

$$U^i(x, t) = \int_{\mathbb{R}^3} \frac{1}{4\pi|x-y|} \varepsilon^{ijk} \frac{\partial}{\partial y^j} W^k(y, t) dy,$$

which yields the Biot-Savart law:

$$U^i(x, t) = - \int_{\mathbb{R}^3} \varepsilon^{ijk} \frac{x^j - y^j}{4\pi|x-y|^3} W^k(y, t) dy. \quad (2.17)$$

for $i = 1, 2, 3$.

⁸That is, if (i, j, k) is an even permutation of $(1, 2, 3)$, then $\varepsilon^{ijk} = 1$; if (i, j, k) is an odd permutation of $(1, 2, 3)$, $\varepsilon^{ijk} = -1$; $\varepsilon^{ijk} = 0$ in other cases.

Chapter 3

Path-regularity and Numerical Simulation of Path-dependent Quadratic BSDEs

3.1 Introduction

In this chapter, we study the following quadratic BSDE on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$

$$\begin{aligned} X_t &= x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s, \\ Y_t &= \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s \cdot dB_s, \quad \xi = \Phi(X) \end{aligned} \tag{3.1}$$

where $(B_t)_{t \in [0, T]}$ is the Brownian motion in $\mathbb{R}^d$, X satisfies the SDE with drift $b : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ and diffusion $\sigma : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times d}$, $f : [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ is a deterministic function with quadratic growth on Z , Φ is a bounded functional on $\tilde{\mathbb{D}}(\mathbb{R}^m)$, $Y_t : \Omega \rightarrow \mathbb{R}$ and $Z_t : \Omega \rightarrow \mathbb{R}^d$.

To the best of our knowledge, most of the numerics of quadratic BSDE relies on the BTZ method, while the terminal condition of such BSDE is $g(X_T)$, where g is a Lipschitz function and X is a forward diffusion process. This is in fact a decoupled FBSDE. Richou et al. [82] considered such Markovian decoupled quadratic FBSDE with a bounded terminal condition. Richou forced the diffusion term in the forward process to be bounded and to depend on t only, while the regularity of g is loosen from Lipschitz to α -Hölder continuous. After reviewing the regularity bound on Z developed by Delbaen et al. [20], Richou claimed the Z component for FBSDE with quadratic growth satisfies

$$|Z_t| \leq C(T - t)^{-\frac{1}{2}},$$

which suggested Z_t had a large upper bound when t tends to T . He therefore constructed an irregular timestep and truncated the driver based on time t .

Chassagneux et al. [12] focused on the decoupled Markovian FBSDE with a terminal condition $\xi = g(X_T)$ as well, but the assumption of the diffusion is more standard, in the sense that g is Lipschitz and X is a standard Lipschitz SDE. They estimated the driver by a Lipschitz function and use a regular timestep, while Y is computed via an implicit scheme and Z is derived by an explicit scheme. Referring to this numerical scheme as the *modified implicit BTZ scheme*, only the convergence rate of the Y component, which is $O(N^{-\frac{1}{2}+\varepsilon})$ for all $\varepsilon > 0$ with N representing the number of discretisation, has been included in [12]. We also mentioned that [59] worked independently, considered the case of BSDE that is one-side Lipschitz in y and *sub-quadratic* in Z i.e. growth in the form of $|Z|^{2-\epsilon}$ for some $\epsilon \in (0, 2]$ instead of $|Z|^2$. Based on this assumption, they introduced a similar numerical scheme, which had the convergence rate of $O(N^{-\frac{\epsilon}{4}})$ in a stronger topology $\mathbb{S}^\infty(\mathbb{R}) \times \mathbb{H}_{\text{BMO}}^2(\mathbb{R}^d)$.

Compared to conventional BSDE problems, our terminal condition depends on the path of $(X_t)_{t \geq 0}$ and the driver has a quadratic growth in Z . More explicitly, our objective is to establish a path-regularity theorem on Y_t and Z_t and to prove the convergence of the *modified (inner) explicit BTZ scheme* on approximating the solution to (Y_t, Z_t) with a suitable norm, based on the analysis of the modified implicit BTZ scheme developed by Chassagneux et al. [12]. The path-regularity result, which is interesting on its own, is a consequence of discretising the path $X(\omega)$ into $(X_{t_0}(\omega), \dots, X_{t_N}(\omega))$, naturally leading to the study of BSDE that depends on a large system of forward SDEs. Taking advantage of the fact that Z_t serves as a Malliavin derivative of Y_t , we then show that $|Z_t|$ is bounded by $C(1 + \sup_{0 \leq s \leq t} |X_t|)$ almost surely for some constant C . Subsequently, we review the modified (inner) explicit BTZ scheme and conduct a perturbation analysis on the Z component, followed by concluding the error of the numerical scheme on estimating (Y, Z) (with a suitable norm) is $O(N^{-\frac{1}{2}})$ rather than $O(N^{-\frac{1}{2}+\varepsilon})$.

This chapter will be organised as in the following. After the assumptions have been listed, we present the main results that demonstrate our contribution, followed by the proof of the path-regularity result in Section 3.2. In Section 3.3, we check the essential properties of the modified explicit BTZ scheme, then a perturbation technique will be applied to the Z component. Section 3.4 and Section 3.5 are devoted to proving the convergence of the numerical scheme and to adapt the quantization method for so-called constructible functionals. Several numerical experiments which

are conducted by quantization technique and regression basis method, are included in the last section to illustrate the performance of the numerical scheme.

3.1.1 Preliminaries and assumptions

Before we discuss the main contribution, we would like to write down the assumptions, which are frequently used in this chapter. We will quote these conditions only when they are used.

Assumption [A2] (Lipschitz SDE). X_t is a diffusion process satisfying the SDE

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s.$$

Meanwhile, the drift term $b : [0, T] \times \mathbb{R}^m \mapsto \mathbb{R}^m$ and volatility term $\sigma : [0, T] \times \mathbb{R}^m \mapsto \mathbb{R}^{m \times d}$ are measurable functions in (t, x) such that there exists $L_X > 0$ with

- $|b(t, 0)| + |\sigma(t, 0)| \leq L_X,$
- $|b(t_1, x_1) - b(t_2, x_2)| + |\sigma(t_1, x_1) - \sigma(t_2, x_2)| \leq L_X(|t_1 - t_2|^{\frac{1}{2}} + |x_1 - x_2|),$

for all $t, t_1, t_2 \in [0, T], x_1, x_2 \in \mathbb{R}^m$.

Assumption [A3] (Path-dependent functionals). $\xi = \Phi(X) \in L^\infty(\mathbb{R})$ where

- Φ is Lipschitz with respect to either $\|\cdot\|_{\mathbb{D}^\infty(\mathbb{R}^m)}$ or $\|\cdot\|_{\mathbb{D}^1(\mathbb{R}^m)}$ with Lipschitz constant $L_\Phi,$
- X satisfies the SDE and assumption [A2].

The regularity of the driver f in the t variable is required in the numerical simulation.

Assumption [A4] (Quadratic growth and regularity on f). The driver f satisfies

$$\begin{aligned} |f(t_1, y_1, z_1) - f(t_2, y_2, z_2)| &\leq L_f \left(|t_1 - t_2|^{\frac{1}{2}} + |y_1 - y_2| \right) \\ &\quad + K_f (1 + |z_1| + |z_2|) |z_1 - z_2|, \\ |f(t_1, 0, 0)| &\leq L_0, \end{aligned}$$

for any $t_1, t_2 \in [0, T], y_1, y_2 \in \mathbb{R}$ and $z_1, z_2 \in \mathbb{R}^d$, for some $L_0, L_f, K_f > 0$.

The target of this chapter is to find a numerical scheme for equation (3.1) with the generator satisfying assumption **[A3]** and **[A4]**. For simplicity, we assume that there exists an uniform bound on $\|\xi\|_{L^\infty}$, $\|f(\cdot, 0, 0)\|_{\mathbb{S}^\infty(\mathbb{R})}$, constants L_X , L_f , K_f which are corresponding to the SDE X and the driver f . The terminal time T and the initial condition X_0 are fixed, regarding the equation (3.1). Therefore, we will omit the dependence of $\|\xi\|_{L^\infty}$, $\|f(\cdot, 0, 0)\|_{\mathbb{S}^\infty(\mathbb{R})}$, L_X , L_f , K_f , T , X_0 , when we mention the generic constant C .

We are now in the position of discussing the numerical scheme. Define $f^{(n)} : [0, T] \times \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$ by

$$f^{(n)}(t, y, z) = f(t, y, \phi_n(z)), \quad \phi_n(z) = \frac{\min(|z|, n)}{|z|}z, \quad (3.2)$$

where $\phi_n : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $n > 0$. By construction, $f^{(n)}$ is Lipschitz in z with Lipschitz constant $L_z(n)$ depending on n i.e.

$$|f^{(n)}(t, y, z) - f^{(n)}(t, y, \tilde{z})| \leq K_f(1 + 2n)|z - \tilde{z}|,$$

for all $t, y, z, \tilde{z}$. We recall the Euler-Maruyama scheme for the process $(X_t)_t$ with some partition π

$$X_{i+1}^\pi = X_i^\pi + b(t_i, X_i^\pi)h_i + \sigma(t_i, X_i^\pi)\Delta B_{t_i}, \quad X_0^\pi = x$$

where $\Delta B_{t_i} = B_{t_{i+1}} - B_{t_i}$ and $h_i = t_{i+1} - t_i$ for all $0 \leq i \leq N-1$. It is natural to define the path

$$X^\pi(t) = X_i^\pi \text{ if } t \in [t_i, t_{i+1}), \quad 1 \leq i \leq N \quad (3.3)$$

with the convention $[t_N, t_N) = \{T\}$, so that $X^\pi \in \tilde{\mathbb{D}}(\mathbb{R}^m)$ and we introduce the following explicit scheme

$$\begin{aligned} \mathcal{Y}_i^\pi &= \mathbb{E}[\mathcal{Y}_{i+1}^\pi + f^{(n)}(t_i, \mathbb{E}[\mathcal{Y}_{i+1}^\pi | \mathcal{F}_{t_i}], \mathcal{Z}_i^\pi)h_i | \mathcal{F}_{t_i}], \\ \mathcal{Z}_i^\pi &= \mathbb{E}[\mathcal{Y}_{i+1}^\pi H_i^R | \mathcal{F}_{t_i}], \\ H_i^R &:= \left(\frac{R}{\sqrt{h_i}} \wedge \frac{\Delta B_{t_i}^1}{h_i} \vee \frac{-R}{\sqrt{h_i}}, \dots, \frac{R}{\sqrt{h_i}} \wedge \frac{\Delta B_{t_i}^d}{h_i} \vee \frac{-R}{\sqrt{h_i}} \right), \\ \mathcal{Y}_N^\pi &= \Phi(X^\pi), \quad \mathcal{Z}_N^\pi = 0. \end{aligned} \quad (3.4)$$

Comparing to the implicit scheme in [12], $\mathcal{Y}_i^\pi$ is obtained directly.

The most important element in the numerical scheme (3.4) is there exist some $\epsilon \in (0, 1)$, independent of N, n, R , such that

$$hL_f + \sup_{0 \leq i \leq N-1} |h_i H_i^R L_z(n)| \leq 1 - \epsilon, \quad (3.5)$$

where $h = \sup_{0 \leq i \leq N-1} h_i$. This can be achieved, provided

$$R = \ln(h^{-1}), \quad n = h^{-\alpha} \text{ for some } \alpha \in \left(0, \frac{1}{2}\right) \quad (3.6)$$

and h small enough. This is vital for assumption [A7], which will be introduced in Section 3.3.

3.1.2 Contribution and main results

The main results and the contribution of this project will be presented in the current subsection. To begin with, we introduce the following notation. Given a Lipschitz function $g : \mathbb{R}^{m \times r} \rightarrow \mathbb{R}$ which satisfies

$$|g(x_1, x_2, \dots, x_r) - g(y_1, y_2, \dots, y_r)| \leq \sum_{i=1}^r L_g^{(i)} |x_i - y_i|,$$

for any $x_1, x_2, \dots, x_r \in \mathbb{R}^m$ and $y_1, y_2, \dots, y_r \in \mathbb{R}^m$, we write

$$S(g) := \sum_{i=1}^r L_g^{(i)} \quad (3.7)$$

as the sum of Lipschitz coefficients.

Theorem 3.1.1. *Consider the BSDE (2.11) with the generator (ξ, f) such that*

- *Let $\pi_0 : 0 = s_0 < \dots < s_{r-1} < s_r = T$ be a partition, we set $\xi = g(X_{s_0}, \dots, X_{s_r})$, where $g : \mathbb{R}^{m \times r} \rightarrow \mathbb{R}$ is Lipschitz and bounded, X is the solution to the SDE which satisfies [A2];*
- *f satisfies [A4].*

Consider the partition $\pi_1 : 0 = t_0 < t_1 < \dots < t_N = T$ and $h = \sup_j |t_{j+1} - t_j|$. Then we have, for all $p \geq 1$,

$$\mathbb{E} \left[\sup_{u \in [s, t]} |Y_u - Y_s|^{2p} \right] \leq C_{p, S(g)} |t - s|^p,$$

where $0 \leq s \leq t \leq T$ and

$$\sum_{j=0}^{N-1} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |Z_t - Z_{t_j}|^2 dt \right)^p \right] \leq C_{p, S(g)} h^p, \quad (3.8)$$

for a constant $C_{p, S(g)} > 0$. Moreover, $C_{p, S(g)}$ is independent of the partitions π_0, π_1 .

We can regard the terminal condition $\xi = g(X_{s_0}, \dots, X_{s_r})$ as a special case of the path-dependent terminal $\Phi(X)$. The more generic result will be deduced in the next theorem.

Theorem 3.1.2. *Suppose ξ satisfies [A3] and f satisfies [A4]. Suppose (Y, Z) is the solution to BSDE (3.1), associated with the generator (ξ, f) . Let $\pi_1 : 0 = t_0 < t_1 \leq \dots \leq t_N = T$ be a partition and we define $h := \sup_j |t_{j+1} - t_j|$. Then, for any $p \geq 1$, there exists a positive constant C_{p, L_Φ} such that*

$$\mathbb{E} \left[\sup_{u \in [s, t]} |Y_u - Y_s|^{2p} \right] \leq C_{p, L_\Phi} |t - s|^p$$

where $0 \leq s \leq t \leq T$ and

$$\sum_{j=0}^{N-1} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |Z_t - Z_{t_j}|^2 dt \right)^p \right] \leq C_{p, L_\Phi} h^p.$$

Moreover, C_{p, L_Φ} is independent of the partition π_1 .

As an application of the path-regularity theorem, we consider the scheme (3.4) and specify the error.

Theorem 3.1.3. *Let ξ satisfy [A3] and f satisfy [A4]. Suppose (Y, Z) is the solution to BSDE (3.1), associated with the generator (ξ, f) .*

Suppose the partition $\pi : 0 = t_0 < \dots < t_N = T$ with $h := \sup_j (t_{j+1} - t_j) \leq KN^{-1}$ for some $K > 0$. Recall X^π is given by (3.3) and $(\mathcal{Y}^\pi, \mathcal{Z}^\pi)$ is the solution to the numerical scheme (3.4). Then we have, for some $q^ > 1$,*

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq i \leq N} |Y_{t_i} - \mathcal{Y}_i^\pi|^2 \right] + \mathbb{E} \left[\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} |Z_s - \mathcal{Z}_j^\pi|^2 ds \right] \\ \leq C_{\alpha, K, L_\Phi, q^*, \epsilon} \left(\mathbb{E} [|\Phi(X) - \Phi(X^\pi)|^{8q^*}]^{\frac{1}{4q^*}} + h \right), \end{aligned}$$

where α, ϵ appear in (3.5) and (3.6).

If we treat $(\mathcal{Y}_i^\pi, \mathcal{Z}_i^\pi)_{0 \leq i \leq N}$ as process by writing

$$\mathcal{Y}^\pi(t) = \sum_{i=0}^{N-1} \mathcal{Y}_i^\pi \mathbb{1}_{[t_i, t_{i+1})}(t) + \Phi(X^\pi) \mathbb{1}_{\{t=T\}}(t), \quad \mathcal{Z}^\pi(t) = \sum_{i=0}^{N-1} \mathcal{Z}_i^\pi \mathbb{1}_{[t_i, t_{i+1})}(t), \quad (3.9)$$

Theorem 3.1.3 essentially controls $\|Z - \mathcal{Z}^\pi(\cdot)\|_{\mathbb{H}^2(\mathbb{R}^d)}^2$. Meanwhile, a convergence in Z that is *point-wise in t* can be achieved through a simple analysis.

Corollary 3.1.4. *Let $\pi(N)$ refer to the partition with N time-steps with maximum timestep $h(N)$. Given a sequence of partitions $\{\pi(N)\}_{N \geq 1}$ such that $h(N) \leq KN^{-1}$ for some $K > 0$, we let $Z^{\pi(N)}(t)$ be the Z -component that is obtained in the numerical scheme (3.4) when equations (3.5), (3.6) are satisfied.*

Then $Z^{\pi(N)}(t) \rightarrow Z_t$ in L^2 a.s in t , provided that

$$\mathbb{E} \left[|\Phi(X) - \Phi(X^{\pi(N)})|^{8q^*} \right]^{\frac{1}{8q^*}} < CN^{-\beta}$$

for some $C, q^, \beta > 0$.*

Proof. Let $\hat{\beta} = 2 \max\{\lceil \frac{1}{\beta} \rceil, 1\}$ and consider the sum of errors after applying Fubini theorem, we apply Theorem 3.1.3 to obtain

$$\int_0^T \sum_{N=1}^{\infty} \mathbb{E} \left[|Z_s - Z^{\pi(N\hat{\beta})}(s)|^2 \right] ds \leq C_{\alpha, K, L_{\Phi}, \epsilon} \sum_{N=1}^{\infty} \left(N^{-\beta\hat{\beta}} + N^{-\hat{\beta}} \right) < \infty,$$

leading us to $\sum_{N=1}^{\infty} \mathbb{E} \left[|Z_t - Z^{\pi(N\hat{\beta})}(t)|^2 \right]$ is finite a.s in t . Therefore,

$$\mathbb{E} \left[|Z_t - Z^{\pi(N)}(t)|^2 \right] \rightarrow 0, \text{ as } N \rightarrow \infty$$

a.s in $t \in [0, T]$. □

Remark 3.1. *It is possible to add a further dependence on the forward process X on the driver f by assuming that $f(t, x, y, z)$ is Lipschitz and bounded in x . This will lead to a modification of the proof of Theorem 3.1.2 and Theorem 3.1.3. In addition, the error of the numerical scheme depends on the largest interval h , hence the scheme might be extended to other potential non-equidistant partitions.*

3.2 Path regularity to solutions of a class of BSDE

3.2.1 Classical differentiability of BSDE

The purpose of this subsection is to establish an expression for the Z component, through the Malliavin derivative of Y . To begin with, we need to review the sensitivity of Y with respect to the initial condition x of the forward SDE X . The classical derivative of Y should satisfy a BSDE whose driver is Lipschitz in z with stochastic Lipschitz coefficients. Briand and Confortola [6] studied the existence and uniqueness of such type of BSDE.

Theorem 3.2.1 (Theorem 10 in [6]). *Suppose the pair (Y, Z) is the solution to BSDE (2.11) where we make the following assumptions on the generator ξ, f :*

(i) The map $(y, z) \mapsto f(t, y, z)$ is continuous and satisfies

$$\begin{aligned} (y - \tilde{y})(f(t, y, z) - f(t, \tilde{y}, z)) &\leq K_t^{2\alpha} |y - \tilde{y}|^2, \\ |f(t, y, z) - f(t, y, \tilde{z})| &\leq K_t |z - \tilde{z}|, \\ |f(t, y, z)| &\leq |f(t, 0, 0)| + K_t^{2\alpha} |y| + K_t |z|, \end{aligned}$$

for some $\alpha \in (0, 1)$, where $K_t \in \mathbb{H}_{BMO}^2(\mathbb{R})$. In addition, there exists $\tilde{q} > 1$ such that for all $1 < q < \tilde{q}$, $\mathcal{E}(\int_0^\cdot K_t u \cdot dB_t)$ satisfies the reverse Hölder inequality (Theorem 2.3.14) for all¹ $u \in \mathbb{R}^d$ with $|u| = 1$.

(ii) There exists some $p^* > \frac{\tilde{q}}{\tilde{q}-1}$ such that $\|\xi\|_{\mathbb{S}^{p^*}(\mathbb{R})} + \|f(\cdot, 0, 0)\|_{\mathbb{H}^{p^*}(\mathbb{R})} < \infty$.

Then there exists a unique solution $(Y, Z) \in \mathbb{S}^p(\mathbb{R}) \times \mathbb{H}^p(\mathbb{R}^d)$ for all $p < p^*$. Moreover, the norm of (Y, Z) is controlled in the sense that with $p_1 = \frac{pp^*}{p^*-p}$,

$$\begin{aligned} \|Y\|_{\mathbb{S}^p(\mathbb{R})} &\leq C_1 \|\xi\|_{\mathbb{S}^{p^*}(\mathbb{R})} + \|f(\cdot, 0, 0)\|_{\mathbb{H}^{p^*}(\mathbb{R})}, \\ \|Z\|_{\mathbb{H}^p(\mathbb{R}^d)} &\leq C_2 (\|Y\|_{\mathbb{S}^p(\mathbb{R})} + \|f(\cdot, 0, 0)\|_{\mathbb{H}^p(\mathbb{R})} + \|Y\|_{\mathbb{S}^p(\mathbb{R})} (\|K^\alpha\|_{\mathbb{H}^{2p_1}(\mathbb{R})} + \|K\|_{\mathbb{H}^{2p_1}(\mathbb{R})})), \end{aligned}$$

where C_1, C_2 depends on p^*, p .

The random variable under L^{p_1} norm can be controlled by an increasing function of $\|K\|_{\mathbb{H}_{BMO}^2(\mathbb{R})}$. The above theorem replaces the global Lipschitz property of the driver on Y by a monotonicity condition with a positive stochastic Lipschitz condition. Meanwhile, it is not difficult to deduce a comparison theorem for this type of BSDE, with the global Lipschitz assumption on the driver of Y .

Lemma 3.2.2. Suppose (Y^1, Z^1) and (Y^2, Z^2) are solutions to a BSDE of the type

$$Y_t^i = \xi^i + \int_t^T f^i(s, Y_s^i, Z_s^i) ds - \int_t^T Z_s^i \cdot dB_s,$$

satisfying the same conditions as Theorem 3.2.1, except

$$|f^i(t, y, z) - f^i(t, \tilde{y}, z)| \leq L_f |y - \tilde{y}|$$

for $i = 1, 2$. If we further assume $\xi^1 \geq \xi^2$ a.s. and $\Delta f(t, Y^2, Z^2) := f^1(t, Y^2, Z^2) - f^2(t, Y^2, Z^2) \geq 0$ a.s., then $Y_t^1 \geq Y_t^2$ almost surely.

Proof. The proof follows a standard argument, which has been deferred to Section A.2. □

¹See Remark 2.5.

We would also quote the following result.

Theorem 3.2.3 (Theorem 3.1.3 in [22]). *Let ξ be a $\mathcal{F}_T$ -measurable random variable which depends on $x \in \mathbb{R}^m$ and we write $\xi(x)$ to emphasize this dependence. Suppose $(\xi(x), f)$ is the generator of BSDE (2.11). We let $p \geq 2$ and impose the following conditions.*

- $\xi(x)$ is bounded in $(x, \omega) \in \mathbb{R}^m \times \Omega$ and satisfies

$$\sup_{x \in \mathbb{R}^m} \|\nabla_x \xi\|_{L^p(\mathbb{R}^m)} < \infty, \quad x \mapsto \nabla_x \xi \text{ is continuous.}$$

- f satisfies [A4] and $(y, z) \rightarrow f(t, y, z)$ is $\mathcal{C}^{1,1}(\mathbb{R} \times \mathbb{R}^d; \mathbb{R})$.

We denote $(Y^x, Z^x) \in \mathbb{S}^p(\mathbb{R}) \times \mathbb{H}^p(\mathbb{R}^d)$ as the solution to (2.11). Then there exists a unique solution $(\nabla_x Y^x, \nabla_x Z^x)$ to the following BSDE

$$\begin{aligned} \nabla_x Y_t^x &= \nabla_x \xi(x) + \int_t^T (\nabla_y f(s, Y_s^x, Z_s^x) \nabla_x Y_s^x + \nabla_z f(s, Y_s^x, Z_s^x) \cdot \nabla_x Z_s^x) ds \\ &\quad - \int_t^T \nabla_x Z_s^x \cdot dB_s \end{aligned}$$

where $\nabla_y f(t, y, z) = \partial_y f(t, y, z)$ and $\nabla_z f(t, y, z) = (\partial_{z^1} f(t, y, z), \dots, \partial_{z^d} f(t, y, z))$.

Therefore, we immediately have

Proposition 3.2.4. *Suppose (ξ, f) is the generator of BSDE (2.11). We let $p \geq 2$, $x \in \mathbb{R}^m$ and impose the following conditions.*

(i) $\xi = g(X_T^{(1)}, X_T^{(2)}, \dots, X_T^{(r)})$. Moreover,

- $g \in \mathcal{C}^1(\mathbb{R}^{m \times r}; \mathbb{R})$ is bounded;
- $(X^{(i)})_{1 \leq i \leq r}$ are diffusion processes with drift coefficients $\{b^{(i)}\}_i$ and diffusion coefficients $\{\sigma^{(i)}\}_i$;
- $X^{(i)}$ satisfies [A2] for some constant L_X which is independent of i , for all $1 \leq i \leq r$;
- the drift $b^{(i)}(t, x)$ and diffusion term $\sigma^{(i)}(t, x)$ satisfy $x \mapsto b^{(i)}(t, x) \in \mathcal{C}^{1+\alpha}$ and $x \mapsto \sigma^{(i)}(t, x) \in \mathcal{C}^{1+\alpha}$, for some $\alpha > 0$ and for all $t \in [0, T]$, $1 \leq i \leq r$.

(ii) f satisfies [A4] and $(y, z) \rightarrow f(t, y, z)$ is $\mathcal{C}^{1,1}(\mathbb{R} \times \mathbb{R}^d; \mathbb{R})$.

Let $(Y^{\mathbf{x}}, Z^{\mathbf{x}})$ where $\mathbf{x} = (x^{(1)}, \dots, x^{(r)})$, be the solution to (2.11) with generator (ξ, f) . Then for all $1 \leq i \leq r$,

$$\begin{aligned} \nabla_{x^{(i)}} Y_t^{\mathbf{x}} &= (\nabla_{x^{(i)}} g)(X_T^{(1)}, X_T^{(2)}, \dots, X_T^{(r)}) \cdot \nabla_{x^{(i)}} X_T^{(i)} - \int_t^T \nabla_{x^{(i)}} Z_s^{\mathbf{x}} \cdot dB_s \\ &\quad + \int_t^T (\nabla_y f(s, Y_s^{\mathbf{x}}, Z_s^{\mathbf{x}}) \nabla_{x^{(i)}} Y_s^{\mathbf{x}} + \nabla_z f(s, Y_s^{\mathbf{x}}, Z_s^{\mathbf{x}}) \cdot \nabla_{x^{(i)}} Z_s^{\mathbf{x}}) ds. \end{aligned}$$

Moreover, the solution $(X^{(i)}, \nabla_{x^{(i)}} Y^{\mathbf{x}}, \nabla_{x^{(i)}} Z^{\mathbf{x}}) \in \mathbb{S}^p(\mathbb{R}^m) \times \mathbb{S}^p(\mathbb{R}^m) \times \mathbb{H}^p(\mathbb{R}^{d \times m})$ for all $p \geq 1$.

Proof. According to Kunita [55, Theorem 4.6.5], the above assumption guarantees the map $x \mapsto X(t, x)$ is continuously differentiable and the Jacobian matrix $\frac{\partial X}{\partial x}(t, x)$ exists (together with its inverse, see Kunita [55, P176]). The conditions in Theorem 3.2.3 are satisfied and the result follows directly. $\square$

Regarding the Malliavin differentiability of quadratic BSDEs, Ankirchner et al. [2] proved such a result. They assumed in $\xi \in \mathbb{D}^{1,\infty}(\mathbb{R})$ without using the excess integrability of ξ . Their work was generalised by Fujii and Takahashi [30, Theorem 5.1] on QBSDEs with jumps, but we only state the simpler version here.

Theorem 3.2.5 (Theorem 8.4 in [2]). *Let (ξ, f) be the generator for the BSDE (2.11) with the solution (Y, Z) . Suppose*

- $\xi \in \mathbb{D}^{1,2}(\mathbb{R}) \cap L^\infty(\mathbb{R})$;
- f satisfies [A4] and $(y, z) \rightarrow f(t, y, z)$ is $\mathcal{C}^{1,1}(\mathbb{R} \times \mathbb{R}^d; \mathbb{R})$.

Then we have $(Y, Z) \in \mathbb{D}^{1,2}(\mathbb{R}) \times (\mathbb{D}^{1,2}(\mathbb{R}))^d$ and their derivative process satisfies

$$D_u Y_t = D_u \xi + \int_t^T (\nabla_y f(s, Y_s, Z_s) D_u Y_s + \nabla_z f(s, Y_s, Z_s) \cdot D_u Z_s) ds - \int_t^T D_u Z_s \cdot dB_s,$$

for $0 \leq u \leq t \leq T$. For $u > t$, the $D_u Y_t = 0$ and $D_u Z_t = 0$. In addition, $\{D_t Y_t\}_{t \in [0, T]}$ is a version of $\{Z_t\}_{t \in [0, T]}$.

We next apply this result to the BSDE which has the terminal function depending on a system of SDEs. For simplicity, we write

$$\nabla_{x^{(i)}} g(x^{(1)}, \dots, x^{(r)}) := \left(\frac{\partial g}{\partial x_1^{(i)}}(x^{(1)}, \dots, x^{(r)}), \dots, \frac{\partial g}{\partial x_m^{(i)}}(x^{(1)}, \dots, x^{(r)}) \right),$$

where $x^{(i)} = (x_1^{(i)}, \dots, x_m^{(i)})$ is a vector.

Proposition 3.2.6. *Let (Y, Z) be the BSDE satisfying the same condition as in Proposition 3.2.4. Then the Malliavin derivative of Y satisfies the BSDE*

$$\begin{aligned} D_u Y_t &= \sum_{i=1}^r (\nabla_{x^{(i)}} g)(X_T^{(1)}, X_T^{(2)}, \dots, X_T^{(r)}) \cdot D_u X_T^{(i)} - \int_t^T D_u Z_s \cdot dB_s \\ &\quad + \int_t^T (\nabla_y f(s, Y_s, Z_s) D_u Y_s + \nabla_z f(s, Y_s, Z_s) \cdot D_u Z_s) ds \end{aligned}$$

while $0 \leq u \leq t \leq T$. For $u > t$, the $D_u Y_t = 0$ and $D_u Z_t = 0$. In addition, $\{D_t Y_t\}_{t \in [0, T]}$ is a version of $\{Z_t\}_{t \in [0, T]}$.

Inspired by the last two propositions, we obtain a non-surprising result which formulates the Z_t a.s. for all $t \in [0, T]$, together with a bound for Z consisting of forward SDEs.

Lemma 3.2.7. *Let (Y, Z) be the BSDE satisfying the same condition as in Proposition 3.2.4. Then we are able to obtain the explicit representation of $D_\theta Y_t$ and Z_t*

$$\begin{aligned} D_\theta Y_t &= \sum_{i=1}^r (\nabla_{x^{(i)}} Y_t^{\mathbf{x}})^\top \left(\nabla_{x^{(i)}} X_\theta^{(i)} \right)^{-1} \sigma^{(i)}(\theta, X_\theta^{(i)}), \\ Z_t &= D_t Y_t = \sum_{i=1}^r (\nabla_{x^{(i)}} Y_t^{\mathbf{x}})^\top (\nabla X_t^{(i)})^{-1} \sigma^{(i)}(t, X_t^{(i)}) \quad a.s, \end{aligned}$$

for some fixed $\theta \in [0, t]$, where $\left(\nabla_{x^{(i)}} X_\theta^{(i)} \right)^{-1}$ is the inverse of $\nabla_{x^{(i)}} X_\theta^{(i)}$. Moreover, there exists $C > 0$ which is independent of r and $\{X_0^i\}_{1 \leq i \leq r}$ such that

$$|Z_t| \leq C \left(\sum_{i=1}^r L_g^{(i)} \left(1 + |X_t^{(i)}| \right) \right). \quad (3.10)$$

Proof. The first part of the proof is an extension of [2, Theorem 8.5]. Utilizing [69, equation (2.59)], we are allowed to replace $D_\theta X_t^{(i)}$ with

$$D_\theta X_t^{(i)} = \nabla_{x^{(i)}} X_t^{(i)} (\nabla_{x^{(i)}} X_\theta^{(i)})^{-1} \sigma^{(i)}(\theta, X_\theta^{(i)}),$$

for any $1 \leq i \leq r$ with $\theta \leq t$. Setting $I_\theta(t)$ and $M_\theta^{(i)}$ as

$$I_\theta(t) := \sum_{i=1}^r \nabla_{x^{(i)}} Y_t^{\mathbf{x}} \cdot M_\theta^{(i)}, \quad M_\theta^{(i)} := (\nabla_{x^{(i)}} X_\theta^{(i)})^{-1} \sigma^{(i)}(\theta, X_\theta^{(i)}),$$

we now apply Itô's formula to $I_\theta(t)$ to derive

$$I_\theta(t) = I_\theta(T) - \sum_{i=1}^r \int_t^T M_\theta^{(i)} \cdot d(\nabla_{x^{(i)}} Y_s^{\mathbf{x}})$$

$$\begin{aligned}
&= \sum_{i=1}^r \nabla_{x^{(i)}} Y_T^{\mathbf{x}} \cdot M_{\theta}^{(i)} - \sum_{i=1}^r \int_t^T M_{\theta}^{(i)} \cdot d(\nabla_{x^{(i)}} Y_s^{\mathbf{x}}) \\
&= \sum_{i=1}^r \left[(\nabla_{X^{(i)}} g) \left(X_T^{(1)}, X_T^{(2)}, \dots, X_T^{(r)} \right) \right]^{\top} D_u X_T^{(i)} + \int_t^T \nabla_y f(s, Y_s, Z_s) I_{\theta}(s) ds \\
&\quad + \int_t^T \nabla_z f(s, Y_s, Z_s) \cdot \left(\sum_{i=1}^r \nabla_{x^{(i)}} Z_s^{\mathbf{x}} M_{\theta}^{(i)} \right) ds - \int_t^T \left(\sum_{i=1}^r \nabla_{x^{(i)}} Z_s^{\mathbf{x}} M_{\theta}^{(i)} \right) \cdot dB_s.
\end{aligned}$$

Thanks to the integrability of the Jacobian of X , the terminal data of the above BSDE is a member of L^p for any $p \geq 1$. By the uniqueness of solution to such linear BSDE with stochastic Lipschitz coefficients (see Theorem 3.2.1), we conclude the first statement

$$D_{\theta} Y_t = \sum_{i=1}^r (\nabla_{x^{(i)}} Y_t^{\mathbf{x}})^{\top} \left(\nabla_{x^{(i)}} X_{\theta}^{(i)} \right)^{-1} \sigma^{(i)}(\theta, X_{\theta}^{(i)}),$$

whereas the second statement is a result of $Z_t = D_t Y_t$.

Regarding the bound of Z_t , we use a classical strategy, which reformulates $\nabla_{x^{(i)}} Y_t^{\mathbf{x}}$ to derive

$$\begin{aligned}
\nabla_{x^{(i)}} Y_t^{\mathbf{x}} &= (\nabla_{x^{(i)}} g)(X_T^{(1)}, X_T^{(2)}, \dots, X_T^{(i)}) \cdot \nabla_{x^{(i)}} X_T^{(r)} \\
&\quad + \int_t^T \nabla_y f(s, Y_s^{\mathbf{x}}, Z_s^{\mathbf{x}}) \nabla_{x^{(i)}} Y_s^{\mathbf{x}} ds - \int_t^T \nabla_{x^{(i)}} Z_s^{\mathbf{x}} \cdot dB_s^{\mathbb{Q}}.
\end{aligned}$$

Here, the measure $\mathbb{Q}$ and the associated Radon-Nikodym derivative η_t are defined as

$$dB_s^{\mathbb{Q}} = dB_s - \beta_s ds, \quad \beta_s = \nabla_z f(s, Y_s^{\mathbf{x}}, Z_s^{\mathbf{x}}), \quad \eta_t = \frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = \mathcal{E} \left(\int_0^t \beta_s \cdot dB_s \right).$$

This measure is well-defined as $|\beta_s| \leq C(1+|Z_s|)$ and $\int_0^t Z_s \cdot dB_s$ is BMO. Furthermore, if we apply Itô's formula to $\Gamma_t \nabla_{x^{(i)}} Y_t$, where

$$\Gamma_t = \exp \left(\int_0^t \nabla_y f(s, Y_s^{\mathbf{x}}, Z_s^{\mathbf{x}}) ds \right),$$

followed by taking conditional expectation under $\mathbb{Q}$, we derive

$$\Gamma_t \nabla_{x^{(i)}} Y_t = \mathbb{E}^{\mathbb{Q}} \left[\Gamma_T (\nabla_{x^{(i)}} g) \left(X_T^{(1)}, X_T^{(2)}, \dots, X_T^{(r)} \right) \nabla_{x^{(i)}} X_T^{(i)} \mid \mathcal{F}_t \right].$$

As Γ_t is positive and uniformly bounded in t , we can control a piece of Z_t using

$$\begin{aligned}
&\left| \nabla_{x^{(i)}} Y_t^{\mathbf{x}} \left(\nabla_{x^{(i)}} X_t^{(i)} \right)^{-1} \sigma^{(i)} \left(t, X_t^{(i)} \right) \right| \\
&= \mathbb{E}^{\mathbb{Q}} \left[\frac{\Gamma_T}{\Gamma_t} (\nabla_{x^{(i)}} g) \left(X_T^{(1)}, X_T^{(2)}, \dots, X_T^{(r)} \right) \nabla_{x^{(i)}} X_T^{(i)} \left(\nabla_{x^{(i)}} X_t^{(i)} \right)^{-1} \sigma^{(i)} \left(t, X_t^{(i)} \right) \mid \mathcal{F}_t \right]
\end{aligned}$$

$$\leq L_g^{(i)} \exp(L_f(T-t)) L_x \mathbb{E}^{\mathbb{Q}} \left[\left| \nabla_{x^{(i)}} X_T^{(i)} (\nabla_{x^{(i)}} X_t^{(i)})^{-1} \right| \middle| \mathcal{F}_t \right] \left(1 + |X_t^{(i)}| \right),$$

where we have used the linear growth assumption of $\sigma^{(i)}$. In order to bound the conditional expectation, we applied the reverse Hölder inequality to η_t together with the L^{q^*} norm of the Jacobian $\nabla_{x^{(i)}} X_t^{(i)}$ to conclude

$$\mathbb{E}^{\mathbb{Q}} \left[\left| \nabla_{x^{(i)}} X_T^{(i)} (\nabla_{x^{(i)}} X_t^{(i)})^{-1} \right| \middle| \mathcal{F}_t \right] \leq c_q \mathbb{E} \left[\left| \nabla_{x^{(i)}} X_T^{(i)} (\nabla_{x^{(i)}} X_t^{(i)})^{-1} \right|^{q^*} \middle| \mathcal{F}_t \right]^{\frac{1}{q^*}} \leq C_q,$$

for some constants c_q, C_q .

Notice that the above argument is valid for all $1 \leq i \leq r$, since our assumption ensures the growth rate of all $X^{(i)}$ is bounded by a uniform constant, i.e.

$$|Z_t| \leq \sum_{i=1}^r \left| \nabla_{x^{(i)}} Y_t^x (\nabla_{x^{(i)}} X_t^{(i)})^{-1} \sigma^{(i)}(t, X_t^{(i)}) \right| \leq C \sum_{i=1}^r L_g^{(i)} \left(1 + |X_t^{(i)}| \right).$$

□

In general, a sequence of random variable $(\xi_r)_{r \geq 0}$ of the type $g(X_T^{(1)}, \dots, X_T^{(r)})$ may not converge to a random variable. We impose a more realistic assumption on the terminal condition.

Lemma 3.2.8. *Let (Y, Z) be the BSDE satisfying the same condition as in Proposition 3.2.4, except ξ takes the special form $\xi = g(X_{s_0}^{(0)}, \dots, X_{s_{r-1}}^{(r-1)}, X_{s_r}^{(r)})$ with $0 = s_0 \leq s_1 \leq \dots \leq s_{r-1} \leq s_r = T$. If for all $0 \leq i \leq r$, we set $(y_i, z_i) \in \mathbb{R}^m \times \mathbb{R}^{d \times m}$ as the solution to the BSDE*

$$\begin{aligned} y_i(t) &= \xi_i - \int_t^T z_i(s) \cdot dB_s \\ &\quad + \int_t^T (\nabla_y f(s, Y_s, Z_s) y_i(s) + \nabla_z f(s, Y_s, Z_s) \cdot z_i(s)) ds \end{aligned}$$

where $\xi_i = \left((\nabla_{x^{(i)}} g) \left(X_{s_0}^{(0)}, \dots, X_{s_{r-1}}^{(r-1)}, X_{s_r}^{(r)} \right) \right)^\top \nabla_{x^{(i)}} X_{s_i}^{(i)} \in \mathbb{R}^m$. Then we deduce that

$$\begin{aligned} D_\theta Y_t &= \sum_{i=0}^r y_i(t) \cdot (\nabla_{x^{(i)}} X_\theta^{(i)})^{-1} \sigma^{(i)}(\theta, X_\theta^{(i)}) \mathbb{1}_{\{\theta \geq s_i\}}, \quad 0 \leq \theta \leq t, \\ Z_t &= D_t Y_t = \sum_{i=1}^r y_i(t) \cdot (\nabla_{x^{(i)}} X_t^{(i)})^{-1} \sigma^{(i)}(t, X_t^{(i)}) \mathbb{1}_{\{t \geq s_i\}} \quad a.s. \end{aligned}$$

In particular, if $X^{(i)} = X^{(j)}$ for all $i, j \in \{1, \dots, r\}$ (and we simply write $X^i = X$), there exists a unique solution $\nabla^j Y_t$ to the following BSDE

$$\begin{aligned} \nabla^j Y_t = & \sum_{i \geq j}^r (\nabla_{x^{(i)}} g)(X_{s_0}, X_{s_1}, \dots, X_{s_r}) \cdot \nabla_x X_{s_i} - \int_t^T \nabla^j Z_t \cdot dB_s \\ & + \int_t^T (\nabla_y f)(s, Y_s^x, Z_s^x) \nabla^j Y_t + (\nabla_z f)(s, Y_s^x, Z_s^x) \cdot \nabla^j Z_t ds. \end{aligned} \quad (3.11)$$

Given a fixed time $t \in [s_j, s_{j+1})$, $0 \leq j \leq N$ where we set $[s_r, s_{r+1}) = \{T\}$ by convention,

$$D_\theta Y_t = (\nabla^j Y_t)^\top (\nabla_x X_\theta)^{-1} \sigma(\theta, X_\theta), \quad Z_t = (\nabla^j Y_t)^\top (\nabla_x X_t)^{-1} \sigma(t, X_t), \quad (3.12)$$

which implies the bound of Z

$$|Z_t| \leq C_{S(g)} \left(1 + \sup_{0 \leq s \leq t} |X_s| \right) \quad t\text{-a.s.}$$

Proof. The first part of the claim is straightforward if we notice that $D_\theta X_{s_i} = 0$ when $\theta > s_i$, hence we focus on the second part, starting with verifying the conditions in Theorem 3.2.3. Given $\nabla_x \xi^{(j)} := \sum_{i=j}^r (\nabla_{x^{(i)}} g)(X_{s_0}, \dots, X_{s_r}) \cdot \nabla_x X_{s_i}$, we immediately have

$$\begin{aligned} \sup_{0 \leq j \leq r} \sup_{x \in \mathbb{R}^m} \|\nabla_x^{(j)} \xi\|_{L^p(\mathbb{R}^m)} & \leq \left(\sum_{i=0}^r L_g(x^{(i)}) \right) \sup_{0 \leq i \leq r} \|\nabla_x X\|_{\mathbb{S}^p(\mathbb{R}^m)} \\ & \leq S(g) \sup_{0 \leq i \leq r} \|\nabla_x X\|_{\mathbb{S}^p(\mathbb{R}^m \times m)}. \end{aligned}$$

Consequently, we rewrite the previous proof to obtain the Mallavin derivative of Y , which is

$$D_\theta Y_t = \sum_{i=0}^r y_i(t) \cdot (\nabla_x X_\theta)^{-1} \sigma(\theta, X_\theta) \mathbb{1}_{\{\theta \geq s_i\}}, \quad 0 \leq \theta \leq t,$$

where $y_i(t) \in \mathbb{R}^m$ is the solution to

$$\begin{aligned} y_i(t) = & ((\nabla_{x^{(i)}} g)(X_{s_0}, \dots, X_{s_r}))^\top \nabla_x X_{s_i} - \int_t^T z_i(t) \cdot dB_t \\ & + \int_t^T (\nabla_y f(s, Y_s, Z_s) y_i(s) + \nabla_z f(s, Y_s, Z_s) \cdot z_i(s)) ds. \end{aligned} \quad (3.13)$$

Therefore, the bound of Z_t follows directly and the result² has been derived as we required. $\square$

²This result can be viewed as an extension of [90, Lemma 2.5] in the context of quadratic BSDE.

Remark 3.2. We are allowed to deduce a similar bound on Y when we replace the terminal condition in Lemma 3.2.8 by a Lipschitz function $\Phi : \tilde{\mathbb{D}}(\mathbb{R}^m) \rightarrow \mathbb{R}$ with respect to $\|\cdot\|_{\tilde{\mathbb{D}}^\infty(\mathbb{R}^m)}$ or $\|\cdot\|_{\tilde{\mathbb{D}}^1(\mathbb{R}^m)}$ of a forward SDE, provided the forward diffusion X satisfies [A2].

If Z is the solution corresponding to (3.1), we set Z^π as the solution to BSDE in Lemma 3.2.8. Z^π can be viewed as the approximation of Z . By the stability theorem of QBSDEs, there exists a subsequence $Z_t^{\pi_1} \rightarrow Z_t$ $t \times \omega$ -almost surely. Such density argument implies $|Z_t| \leq C_{L_\Phi}(1 + \sup_{0 \leq s \leq t} |X_s|)$, leading us to Z_t is bounded in L^p t -almost surely for any $p \geq 1$, with the bound depending on L_Φ and p .

3.2.2 Proof of results on path regularity

The goal of this subsection is to show the path-regularity of Y and Z in equation (3.1). We begin with stating two technical results, which are building blocks in the proof of the main theorems. The first lemma is very helpful in transforming the sum of differences in BSDE to sum of Riemann integrals.

Lemma 3.2.9. Let X_t, Y_t be adapted processes, where X_t is a martingale and both $X_t, Y_t \in L^2$. Then we have

$$\mathbb{E}[X_{t_i} Y_{t_i}] = \mathbb{E}[X_T Y_{t_i}], \quad \sum_{i=1}^n |a_i|^p \leq \left(\sum_{i=1}^n |a_i| \right)^p, \quad (3.14)$$

while the latter inequality³ remains valid for all $p \geq 1$.

Proof. See Section A.2. □

The following result allows us to control a sum of quadratic variation process by the sum of the terminal value of martingales.

Lemma 3.2.10. Given some $p \geq 1$, let $\xi_i(t) = (\xi_i^{(1)}(t), \dots, \xi_i^{(m)}(t))$ be a $\mathbb{R}^m$ -valued adapted martingale⁴ in $L^{2p}(\Omega; \mathbb{R}^m)$ for all $0 \leq i \leq r$, such that $\xi_i(t) = \alpha_i + \int_0^t \eta_i(u) \cdot dB_u$ for some $\eta_i \in \mathbb{R}^{d \times m}$ (by martingale representation theorem). For any increasing function $\phi : \{0, 1, \dots, N-1\} \rightarrow \{0, \dots, r\}$, we define $\tilde{\phi} : \{0, \dots, r\} \rightarrow \{-1, 0, \dots, N-1\}$ as $\tilde{\phi}(i) = \sup\{j \mid \phi(j) \leq i\}$ with the convention $\sup \emptyset = -1$. Then,

$$\mathbb{E} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left| \sum_{i=\phi(j)}^r \eta_i(u) \right|^2 du \right)^p \right] \leq C_{p,m} \mathbb{E} \left[\left| \sum_{i=0}^r \xi_i(t_{\tilde{\phi}(i)+1}) - \xi_i(0) \right|^{2p} \right],$$

where $C_{p,m}$ is a constant depending on p, m and independent of r .

³When $p \in (0, 1]$, the inequality becomes $\sum_{i=1}^n |a_i|^p \geq (\sum_{i=1}^n |a_i|)^p$.

⁴That is, $\mathbb{E}[\xi_i^{(j)}(t) \mid \mathcal{F}_s] = \xi_s^{(j)}$ for all $1 \leq j \leq m$.

Proof. We first define a square integrable progressively measurable process H_s ,

$$H_u = (H_u^{(1)}, \dots, H_u^{(m)}) := \sum_{j=0}^{N-1} \left(\sum_{i=\phi(j)}^r \eta_i(u) \right) \mathbb{1}_{[t_j, t_{j+1})}(u),$$

together with the corresponding Itô's integral $M_t := (M_t^{(1)}, \dots, M_t^{(m)}) := \int_0^t H_s \cdot dB_s \in \mathbb{R}^m$. Utilizing Jensen's inequality, Burkholder-Davis-Gundy (BDG) inequality and Doob's L^p inequality, we deduce

$$\begin{aligned} \mathbb{E} \left[\left(\int_0^T |H_u|^2 du \right)^p \right] &\leq m^{p-1} \sum_{j=1}^m \mathbb{E} \left[\left(\int_0^T |H_u^{(j)}|^2 du \right)^p \right] \\ &\leq C_{p,m} \sum_{j=1}^m \mathbb{E} \left[|M_T^{(j)}|^{2p} \right] \\ &\leq C_{p,m} \mathbb{E} \left[\left(\sum_{j=1}^m |M_T^{(j)}|^2 \right)^p \right] = C_{p,m} \mathbb{E} [|M_T|^{2p}], \end{aligned}$$

where Lemma 3.2.9 was used in the last line. As the left hand side is an integral of H_u , we have

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left| \sum_{i=\phi(j)}^r \eta_i(u) \right|^2 du \right)^p \right] &= \mathbb{E} \left[\left(\int_0^T |H_u|^2 du \right)^p \right] \\ &\leq C_{p,m} \mathbb{E} \left[\left| \int_0^T \sum_{j=0}^{N-1} \left(\sum_{i=\phi(j)}^r \eta_i(u) \right) \mathbb{1}_{[t_j, t_{j+1})}(u) \cdot dB_u \right|^{2p} \right] \\ &\leq C_{p,m} \mathbb{E} \left[\left(\sum_{i=0}^r \int_0^{t_{\tilde{\phi}(i)+1}} \eta_i(u) \cdot dB_u \right)^{2p} \right] \leq C_{p,m} \mathbb{E} \left[\left| \sum_{i=0}^r \xi_i(t_{\tilde{\phi}(i)+1}) - \xi_i(0) \right|^{2p} \right], \end{aligned}$$

after rearranging the sum and plugging in the definition of ϕ and $\tilde{\phi}$. $\square$

Taking the advantage of the expression of Z in Lemma 3.2.8, we now derive the regularity results, in particular, Theorem 3.1.1 and Theorem 3.1.2. The regularity of Y is not very surprising, as Y is a semi-martingale. Without adding any assumption on the partition, we have the second inequality regarding the expectation of the supremum on maximum Y increments. On the contrary, proving the regularity of Z is a much more valuable result. The main difficulty within the proof is the lack of integrability of the Radon-Nikodym derivative. After applying the explicit expression of Z which is derived in Lemma 3.2.8, this obstacle is solved by establishing control on the difference to y_i within the expression through Lemma 3.2.10.

Remark 3.3. Let Φ be either Lipschitz with respect to $\|\cdot\|_{\tilde{\mathbb{D}}^\infty(\mathbb{R}^m)}$ or $\|\cdot\|_{\tilde{\mathbb{D}}^1(\mathbb{R}^m)}$. We further assume that $\Theta \in \tilde{\mathbb{D}}(\mathbb{R}^m)$ is a continuous paths. By Ma and Zhang [63, Lemma 4.1, Lemma 5.2], $\Phi(\Theta)$ may be gradually estimated by a family of functions $\{g_\pi\}_\pi$, where $\pi : 0 = s_0 < \dots < s_{r-1} < s_r = T$, such that

$$\lim_{|\pi| \rightarrow 0} |\Phi(\Theta_\pi) - g_\pi(\Theta(s_0), \dots, \Theta(s_{r-1}), \Theta(s_r))| = 0, \quad |\pi| = \sup_{0 \leq i \leq r-1} |s_{i+1} - s_i|,$$

where $\Theta_\pi(t) := \frac{1}{h_i}((t_{i+1} - t)\Theta_{t_i} + (t - t_i)\Theta_{t_{i+1}})$ for any $t \in [t_i, t_{i+1}]$. Moreover, $g_\pi : \mathbb{R}^{m \times r} \mapsto \mathbb{R}$ is a smooth function with bounded partial derivatives such that $S(g_\pi) < CL_\Phi$ with some constant $C > 0$ which is independent of π . To ensure the approximating function is bounded, we consider a sequence of mollifiers $\{\eta_r\}_r : \mathbb{R}^{m \times r} \rightarrow \mathbb{R}$ with $\int \phi_r dx_1 \cdots dx_r = 1$ and define

$$\hat{g}_\pi = (\|\xi\|_{L^\infty} \wedge g_\pi \vee (-\|\xi\|_{L^\infty})) * \phi_r.$$

Therefore, there exist a bounded smooth approximation of $\Phi(\Theta)$. If X is a continuous process in the probability space $(\Omega, \mathbb{P}, (\mathcal{F}_t)_t)$, then

$$\hat{g}_\pi(X_{s_0}, \dots, X_{s_r}) \rightarrow \Phi(X) \text{ as } |\pi| \rightarrow 0.$$

The strategy of the proof of Theorem 3.1.1 is summarised in the following. As the Z -component can be controlled by $|Z_t| \leq C(1 + \sup_{0 \leq s \leq t} |X_s|)$, the regularity of Z uses the fact that $\sup_{0 \leq s \leq t} |X_s|$ is L^p integrable for all p . The regularity of the Y -component is a simply consequence of Jensen's inequality combined with the BDG inequality. The Z -component needs a bit more careful analysis on the partitions.

1. We assume the terminal data is smooth enough, aiming to extend the result by a density argument.
2. The case π_1 is wider than π_0 and π_1 is finer than π_0 must be discussed separately.
3. In both cases, we focus on the expression $Z_t = (\nabla^k Y_t)^\top (\nabla_x X_t)^{-1} \sigma(t, X_t)$, followed by analysing $\nabla^k Y_t$, $(\nabla_x X_t)^{-1}$ and $\sigma(t, X_t)$.
4. We utilise the fact that the inverse of the Jacobian and the volatility σ has sufficient regularity.
5. Finally, $\nabla^k Y_t$ is a solution to the BSDE with stochastic Lipschitz coefficients, which motivates us to change the measure and to apply the BMO property that is related to the new measure $\mathbb{Q}$.

Proof of Theorem 3.1.1. We first show (i), followed by dividing the proof of part (ii) into different sections which associates to various scenarios.

Step 1. Regarding the Y component, we apply Jensen's inequality, BDG inequality as well as Fubini theorem to obtain

$$\begin{aligned}
& \mathbb{E} \left[\sup_{u \in [s, t]} |Y_u - Y_s|^{2p} \right] \\
& \leq C_p \mathbb{E} \left[\sup_{u \in [s, t]} \left(\left(\int_s^u |f(r, Y_r, Z_r)| \, d\tau \right)^{2p} + \left(\int_s^u Z_\tau \cdot dB_\tau \right)^{2p} \right) \right] \\
& \leq C_p \mathbb{E} \left[\left(\int_s^t |f(\tau, 0, 0)| + L_f |Y_\tau| + \frac{K_f^2}{2} |Z_\tau|^2 \, d\tau \right)^{2p} + \left(\int_s^t |Z_\tau|^2 \, d\tau \right)^p \right] \\
& \leq C_p \left((t-s)^{2p} + (t-s)^{2p-1} \int_s^t \mathbb{E} [|Z_\tau|^{4p}] \, d\tau + (t-s)^{p-1} \int_s^t \mathbb{E} [|Z_\tau|^{2p}] \, d\tau \right).
\end{aligned}$$

By Theorem 2.4.3, Y is uniformly bounded, hence the constant C_p is independent of $\tau \in [s, t]$. What remains to bound Z_t . By Lemma 3.2.8, we have

$$\mathbb{E} [|Z_\tau|^{2p}] \leq C_p \mathbb{E} \left[\left(1 + \sup_{0 \leq s \leq \tau} |X_s| \right)^{2p} \right] \leq C_{p, S(g)}.$$

whereas $\mathbb{E} [|Z_\tau|^{4p}] \leq C_{p, S(g)}$ are deduced in a similar way.

Step 2. The remaining proof will start with assuming the function g , the driver $(y, z) \mapsto f(t, y, z)$, the drift (of the SDE X), $x \mapsto b(t, x)$ and diffusion (of the SDE X) $x \mapsto \sigma(t, x)$ are all continuously differentiable. We focus on $\mathbb{E} [|Z_t - Z_{t_j}|^{2p}]$ since, by applying Jensen's inequality,

$$\sum_{j=0}^{N-1} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |Z_t - Z_{t_j}|^2 \, dt \right)^p \right] \leq h^{p-1} \sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |Z_t - Z_{t_j}|^{2p} \, dt \right]. \quad (3.15)$$

Recall that the Z component has the expression by in in Lemma 3.2.8, where $\nabla^k Y$ and (y_i, z_i) are defined in (3.11) and (3.13) respectively. In the first scenario (in step 2 to step 4), we assume π_1 is a partition wider than (or equivalent to) π_0 . Without loss of generality, we suppose $t_i = s_{l_i}$ for some increasing sequence $(l_i)_{i=1}^N$. For any $t \in [s_k, s_{k+1}) \in [s_{l_j}, s_{l_{j+1}}) = [t_j, t_{j+1})$, we are able to deduce

$$\begin{aligned}
Z_t - Z_{t_j} &= (\nabla^k Y_t)^\top (\nabla_x X_t)^{-1} \sigma(t, X_t) - \left(\nabla^{l_j} Y_{s_{l_j}} \right)^\top \left(\nabla_x X_{s_{l_j}} \right)^{-1} \sigma(s_{l_j}, X_{s_{l_j}}) \\
&= I_{1,j} + I_{2,j} + I_{3,j}, \\
I_{1,j}(t) &:= (\nabla^k Y_t)^\top (\nabla_x X_t)^{-1} (\sigma(t, X_t) - \sigma(t_j, X_{t_j})), \\
I_{2,j}(t) &:= (\nabla^{l_j} Y_t)^\top \left((\nabla_x X_t)^{-1} - (\nabla_x X_{t_j})^{-1} \right) \sigma(t_j, X_{t_j}),
\end{aligned}$$

$$I_{3,j}(t) := (\nabla^k Y_t - \nabla^{l_j} Y_{t_j})^\top (\nabla_x X_{t_j})^{-1} \sigma(t_j, X_{t_j}),$$

by decomposing $Z_t - Z_{t_j}$ through using their explicit expression. Utilising Jensen's inequality, we get

$$\begin{aligned} \sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |Z_t - Z_{t_j}|^{2p} dt \right] \\ \leq C_p \sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} (|I_{1,j}(t)|^{2p} + |I_{2,j}(t)|^{2p} + |I_{3,j}(t)|^{2p}) dt \right]. \end{aligned}$$

Denoting $h_j = t_{j+1} - t_j$, we employ Hölder's inequality and the assumption on $\sigma(t, X_t)$ to compute the contribution of $I_{1,j}(t)$ and $I_{2,j}(t)$

$$\begin{aligned} \mathbb{E} [|I_{1,j}(t)|^{2p}] &\leq \mathbb{E} [|\nabla^k Y_t|^{6p}]^{\frac{1}{3}} \mathbb{E} [|(\nabla_x X_t)^{-1}|^{6p}]^{\frac{1}{3}} \mathbb{E} \left[\left((t - t_j)^{\frac{1}{2}} + L_X |X_t - X_{t_j}| \right)^{6p} \right]^{\frac{1}{3}} \\ &\leq C_{p,S(g)} (t - t_j)^p \leq C_{p,S(g)} h_j^p, \\ \mathbb{E} [|I_{2,j}(t)|^{2p}] &\leq \mathbb{E} [|\nabla^k Y_t|^{6p}]^{\frac{1}{3}} \mathbb{E} \left[\left| (\nabla_x X_t)^{-1} - (\nabla_x X_{t_j})^{-1} \right|^{6p} \right]^{\frac{1}{3}} \mathbb{E} [|\sigma(t_j, X_{t_j})|^{6p}]^{\frac{1}{3}} \\ &\leq C_{p,S(g)} (t - t_j)^p \leq C_{p,S(g)} h_j^p, \end{aligned}$$

where we have taken advantage of the standard path-regularity estimate for the forward SDE. Theorem 3.2.1 together with Lemma 3.2.2 allow us to establish a uniform bound on $\mathbb{E} [|\nabla^k Y_s|^{6p}]$, since $|\sum_{i=k}^r (\nabla_{x^{(i)}} g) \cdot \nabla_x X_{t_i}| \leq U$ for all $k \geq 0$, where

$$U := S(g) \sup_{0 \leq s \leq T} |\nabla_x X_s| \quad (3.16)$$

and $U \in L^p(\mathbb{R})$ for all $p > 0$.

Step 3. Regarding the last component, we make use of equation (3.14) that to derive

$$\sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |I_{3,j}(t)|^{2p} dt \right] \leq \sum_{j=0}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} \mathbb{E} \left[\int_{s_k}^{s_{k+1}} \mathbb{E}[|\nabla^k Y_t - \nabla^{l_j} Y_{t_j}|^{2p} | \mathcal{F}_{t_j}] |\Lambda_{t_j}|^{2p} dt \right],$$

where $\Lambda_t := (\nabla_x X_t)^{-1} \sigma(t, X_t)$ is a L^p integrable random variable for all $p > 0$. Meanwhile,

$$\nabla^k Y_t - \nabla^{l_j} Y_{t_j} = \sum_{i=k}^r y_i(t) - \sum_{i \geq l_j} y_i(t_j) = \sum_{i=k}^r (y_i(t) - y_i(t_j)) - \sum_{i=l_j}^{k-1} y_i(t_j).$$

Define $\Gamma_t = \exp(\int_0^t \alpha_s ds)$ with $\alpha_s := (\partial_y f)(s, Y_s, Z_s)$, which is a bounded process. $\Gamma_t y_i(t)$, according to Itô's lemma, satisfies the BSDE

$$\Gamma_t y_i(t) = \Gamma_T y_i(T) + \int_t^T (\partial_z f)(s, Y_s, Z_s) \cdot \eta_i(s) ds - \int_t^T \eta_i(s) \cdot dB_s \quad (3.17)$$

$$= \Gamma_T y_i(T) - \int_t^T \eta_i(s) \cdot dB_s^{\mathbb{Q}}, \quad (3.18)$$

where $\eta_i(s) = \Gamma_s z_i(s)$ for all $0 \leq i \leq m$ and $dB_s^{\mathbb{Q}} = dB_s - \beta_s ds$ with $\beta_s = (\partial_z f)(s, Y_s, Z_s)$. By the mean-value theorem,

$$\begin{aligned} y_i(t) - y_i(t_j) &= \Gamma_t^{-1}[(\Gamma_t y_i(t) - \Gamma_{t_j} y_i(t_j)) + (\Gamma_{t_j} y_i(t_j) - \Gamma_t y_i(t_j))] \\ &= \Gamma_t^{-1}(\Gamma_t y_i(t) - \Gamma_{t_j} y_i(t_j)) + \frac{\alpha_\tau \Gamma_\tau}{\Gamma_t} y_i(t_j)(t_j - t), \end{aligned}$$

where $\tau \in [t_j, t]$. Now we take the advantage of Theorem 2.3.8 and the boundness of α_s and Γ_s to conclude

$$\begin{aligned} \mathbb{E} \left[|\nabla^k Y_t - \nabla^{l_j} Y_{t_j}|^{2p} \mid \mathcal{F}_{t_j} \right] &\leq \mathbb{E} \left[\left| \Gamma_t^{-1} \sum_{i=k}^r \left(\int_{t_j}^t \eta_i(u) \cdot dB_u - \int_{t_j}^t \eta_i(u) \cdot \beta(u) du \right) \right. \right. \\ &\quad \left. \left. + \frac{\alpha_\tau \Gamma_\tau}{\Gamma_t} (t_j - t) \sum_{i=k}^r y_i(t_j) + \sum_{i \geq l_j}^{k-1} y_i(t_j) \right|^{2p} \mid \mathcal{F}_{t_j} \right] \\ &\leq C_p \mathbb{E} \left[\left(\int_{t_j}^t \left(\sum_{i=k}^r \eta_i(u) \right)^2 du \right)^p + \left| (t_j - t) \sum_{i=k}^r y_i(t_j) \right|^{2p} + \left| \sum_{i \geq l_j}^{k-1} y_i(t_j) \right|^{2p} \right. \\ &\quad \left. + \left(\int_{t_j}^t (\beta(u))^2 + \left(\sum_{i=k}^r \eta_i(u) \right)^2 du \right)^{2p} \mid \mathcal{F}_{t_j} \right] \\ &\leq C_p \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^p + \left| (t_j - t_{j+1}) \sum_{i=k}^r y_i(t_j) \right|^{2p} + \left| \sum_{i \geq l_j}^{k-1} y_i(t_j) \right|^{2p} \right. \\ &\quad \left. + \left(\int_{t_j}^{t_{j+1}} |\beta(u)|^2 du \right)^{2p} + \left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^{2p} \mid \mathcal{F}_{t_j} \right], \quad (3.19) \end{aligned}$$

where we used the elementary inequality $|a \cdot b| \leq C_{m,d} (|a|^2 + |b|^2)$ for any $a \in \mathbb{R}^d, b \in \mathbb{R}^{d \times m}$. Therefore, we conclude that

$$\sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |I_{3,j}(t)|^{2p} dt \right] \leq C_p \left(\sum_{i=1}^5 A_i \right),$$

where A_i for $1 \leq i \leq 5$ are defined as

$$\begin{aligned}
A_1 &:= \sum_{j=0}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} \mathbb{E} \left[\int_{s_k}^{s_{k+1}} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^{2p} \middle| \mathcal{F}_{t_j} \right] |\Lambda_{t_j}|^{2p} dt \right], \\
A_2 &:= \sum_{j=0}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} \mathbb{E} \left[\int_{s_k}^{s_{k+1}} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^p \middle| \mathcal{F}_{t_j} \right] |\Lambda_{t_j}|^{2p} dt \right], \\
A_3 &:= \sum_{j=0}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} \mathbb{E} \left[\int_{s_k}^{s_{k+1}} \mathbb{E} \left[\left| (t_j - t_{j+1}) \sum_{i=k}^r y_i(t_j) \right|^{2p} \middle| \mathcal{F}_{t_j} \right] |\Lambda_{t_j}|^{2p} dt \right], \\
A_4 &:= \sum_{j=0}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} \mathbb{E} \left[\int_{s_k}^{s_{k+1}} \mathbb{E} \left[\left| \sum_{i \geq l_j}^{k-1} y_i(t_j) \right|^{2p} \middle| \mathcal{F}_{t_j} \right] |\Lambda_{t_j}|^{2p} dt \right], \\
A_5 &:= \sum_{j=0}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} \mathbb{E} \left[\int_{s_k}^{s_{k+1}} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |\beta(u)|^2 du \right)^{2p} \middle| \mathcal{F}_{t_j} \right] |\Lambda_{t_j}|^{2p} dt \right].
\end{aligned}$$

Step 4. We aim to estimate A_1 to A_5 . We consider a partition finer than π_1 , namely $\pi_2 : 0 = v_0 < v_1 < \dots < v_{\hat{\rho}_1} < \dots < v_{\hat{\rho}_2} < \dots < v_{\hat{\rho}_N} = T$ such that

- $v_{\hat{\rho}_i} = s_{l_i} = t_i$ with $\hat{\rho}_{i+1} - \hat{\rho}_i = \gamma$ for some $\gamma > 0$ large enough, for all $0 \leq i \leq r$;
- there exists $\rho_k := \frac{v_{\hat{\rho}_i+k+1} - v_{\hat{\rho}_i+k}}{t_{i+1} - t_i}$ which is independent of i for all $0 \leq k \leq \gamma - 1$;
- π belongs to π_2 and there exist a surjective function $\phi(i, k) : \{0, \dots, N-1\} \times \{0, \dots, \gamma-1\} \rightarrow \{0, 1, \dots, r\}$ such that $v_{\hat{\rho}_i+k} = s_{\phi(i,k)}$ for all $0 \leq i \leq N-1$ and $0 \leq k \leq \gamma-1$;
- moreover, $l_i \leq \phi(i, k) \leq l_{j+1} - 1$ with $\phi(i, k+1) \geq \phi(i, k)$, $\phi(i+1, k) \geq \phi(i, \iota)$ for all i, k, ι .

Therefore, the original partition π_0 is included in the new partition and the definition of π_2 implies $\sum_{k=0}^{\gamma-1} \rho_k = 1$. Consequently, if we focus on one of the components and use Hölder's inequality together with the elementary fact (3.14),

$$\begin{aligned}
A_1 &= \sum_{j=0}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} (s_{k+1} - s_k) \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^{2p} |\Lambda_{t_j}|^{2p} \right] \\
&= \sum_{j=0}^{N-1} \sum_{k=0}^{\gamma-1} \frac{v_{k+1} - v_k}{t_{j+1} - t_j} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=\phi(j,k)}^k \eta_i(u) \right|^2 du \right)^{2p} |\Lambda_{t_j}|^{2p} \right] h_j
\end{aligned}$$

$$\begin{aligned}
&\leq \sum_{k=0}^{\gamma-1} \rho_k \mathbb{E} \left[\sum_{j=0}^{N-1} \left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=\phi(j,k)}^k \eta_i(u) \right|^2 du \right)^{2p} \sup_j |\Lambda_{t_j}|^{2p} \right] h \\
&\leq \sum_{k=0}^{\gamma-1} \rho_k \mathbb{E} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left| \sum_{i=\phi(j,k)}^k \eta_i(u) \right|^2 du \right)^{4p} \right]^{\frac{1}{2}} \mathbb{E} \left[\sup_j |\Lambda_{t_j}|^{4p} \right]^{\frac{1}{2}} h.
\end{aligned}$$

Meanwhile, equation (3.18) together with Lemma 3.2.10 suggest that we can replace the Riemann integral by conditional expectations of some random variables under $\mathbb{Q}$ -measure. Therefore, we define $\tilde{\phi}(i, k) = \sup\{j \mid \phi(j, k) \leq i\}$ with the convention $\sup \emptyset = -1$ and $\tilde{\xi}_i(t) := \Gamma_t y_i(t)$ for all $0 \leq t \leq T$.

Moreover, there exists a $\tilde{q} > 1$ such that $\|\beta\|_{\mathbb{H}_{\text{BMO}}^2(\mathbb{R}^d)} \leq \hat{\Xi}(\tilde{q})$, where $\hat{\Xi}$ is the function defined in Theorem 2.3.14. We fix q such that $q \in (1, \tilde{q})$ and close enough to $\tilde{q}$, while we denote q^* as the conjugate of q . By employing the reverse Hölder inequality,

$$\begin{aligned}
&\mathbb{E} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left| \sum_{i=\phi(j,k)}^k \eta_i(u) \right|^2 du \right)^{4p} \right] \\
&\leq C_{q^*} \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left| \sum_{i=\phi(j,k)}^k \eta_i(u) \right|^2 du \right)^{8pq^*} \right]^{\frac{1}{q^*}} \\
&\leq C_{p,q^*} \mathbb{E}^{\mathbb{Q}} \left[\left| \sum_{i=0}^r \tilde{\xi}_i(t_{\tilde{\phi}(i,k)+1}) - \tilde{\xi}_i(0) \right|^{8pq^*} \right]^{\frac{1}{q^*}} \\
&\leq C_{p,q} \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{i=0}^r \mathbb{E} \left[\|\nabla_{x^{(i)}} g\|_{\infty} \sup_{0 \leq i \leq r} |\nabla_x X_{t_i}| \mid \mathcal{F}_{t_{\tilde{\phi}(i,k)+1}} \right] \right)^{8pq^*} + \mathbb{E}^{\mathbb{Q}} [\nabla^r Y_T]^{8pq^*} \right]^{\frac{1}{q^*}} \\
&\leq C_{p,q} \left(S(g)^{8p(\theta(8pq^*))^{\frac{1}{q^*}}} + \mathbb{E}^{\mathbb{Q}} [U]^{8p} \right),
\end{aligned}$$

where U is defined as in (3.16) and $\theta : \mathbb{R}_{++} \mapsto \mathbb{R}$, irrelevant with $\tilde{\phi}(i, k)$, is defined as

$$\theta(y) := \mathbb{E}^{\mathbb{Q}} \left[\left(\sup_{0 \leq s \leq T} \mathbb{E}^{\mathbb{Q}} \left[\sup_{0 \leq u \leq T} |\nabla_x X_u| \mid \mathcal{F}_s \right] \right)^y \right].$$

Applying Doob's maximal inequality and Hölder's inequality, we can provide an upper bound of $\theta(y)$ depending on y for all $y > 1$, namely

$$\theta(y) \leq \left(\frac{y}{y-1} \right)^y \mathbb{E}^{\mathbb{Q}} \left[\sup_{0 \leq u \leq T} |\nabla_x X_u|^y \right] \leq C_{p,q,y},$$

where the last inequality is a consequence of inverting the measure back to $\mathbb{P}$ via applying Corollary 2.3.15. As this bound is uniform in k and r , we get .

$$A_1 \leq \sum_{k=0}^{\gamma-1} \rho_k C(p, q, S(g)) \mathbb{E} \left[\sup_j |\Lambda_{t_j}|^{4p} \right]^{\frac{1}{2}} h \leq C(p, q, S(g)) h.$$

We are left with terms A_i , $i \in \{2, 3, 4, 5\}$ to obtain the bound. Recall ξ_i is the terminal condition for the BSDE (y_i, z_i) in equation (3.13). By implementing the same procedure as estimating A_1 , it is possible to conclude $A_2 \leq C_{p,q,S(g)} h$. Regarding the bound for A_3 , we first notice that

$$\left| (t_j - t_{j+1}) \sum_{i=k}^r y_i(t_j) \right| \leq \exp(L_f T) \left| (t_j - t_{j+1}) \sum_{i=k}^r \Gamma_{t_j} y_i(t_j) \right|,$$

so we are able to deduce the following bound using the new partition π_2

$$\begin{aligned} A_3 &\leq C \sum_{k=0}^{\gamma-1} \rho_k \mathbb{E} \left[\sum_{j=0}^{N-1} \left(h_j \left| \Gamma_{t_j} \sum_{i=\phi(j,k)}^k y_i(t_j) \right| \right)^{2p} \sup_j |\Lambda_{t_j}|^{2p} dt \right] h_j \\ &\leq C \sum_{k=0}^{\gamma-1} \rho_k \mathbb{E} \left[\left(\sum_{j=0}^{N-1} h_j \left| \mathbb{E}^{\mathbb{Q}} \left[\Gamma_T \sum_{i=\phi(j,k)}^k \xi_i \mid \mathcal{F}_{t_j} \right] \right| \right)^{2p} \sup_j |\Lambda_{t_j}|^{2p} dt \right] h \\ &\leq C_{q^*} \sum_{k=0}^{\gamma-1} \rho_k (A_{3,k})^{\frac{1}{2q^*}} \mathbb{E} \left[\sup_j |\Lambda_{t_j}|^{4p} dt \right]^{\frac{1}{2}} h \\ A_{3,k} &:= \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=0}^{N-1} h_j \mathbb{E}^{\mathbb{Q}} \left[\left| \Gamma_T \sum_{i=\phi(j,k)}^k \xi_i \right| \mid \mathcal{F}_{t_j} \right] \right)^{4pq^*} \right], \end{aligned}$$

where we have implemented equation (3.14), the Cauchy-Schwarz inequality and the reverse Hölder's inequality accordingly. Observing that we are allowed to control the expectation under $\mathbb{Q}$ by

$$\begin{aligned} (A_{3,k})^{\frac{1}{2q^*}} &\leq \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=0}^{N-1} h_j \sum_{i=\phi(j,k)}^k \left(\|\nabla_{x^{(i)}} g\|_{L^\infty} \mathbb{E}^{\mathbb{Q}} \left[\sup_{0 \leq s \leq T} |\nabla_x X_s| \mid \mathcal{F}_{t_j} \right] \right) \right)^{4pq^*} \right]^{\frac{1}{2q^*}} \\ &\leq \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=0}^{N-1} h_j \left(\sum_{i \geq 0} \|\nabla_{x^{(i)}} g\|_{L^\infty} \right) \sup_{0 \leq s \leq T} \mathbb{E}^{\mathbb{Q}} \left[\sup_{0 \leq s \leq T} |\nabla_x X_s| \mid \mathcal{F}_s \right] \right)^{4pq^*} \right]^{\frac{1}{2q^*}} \\ &\leq (S(g)T)^{2p} (\theta(4pq^*))^{\frac{1}{2q^*}} \leq C_{p,,} \end{aligned}$$

$(A_{3,k})^{\frac{1}{2q^*}}$ is bounded by a constant independent of k and r , so the contribution of A_3 is at most of order $O(h)$. It now comes to estimate A_4 , using the same type of technique as how we establish a bound on A_3 :

$$\begin{aligned}
A_4 &= \sum_{k=0}^{\gamma-1} \rho_k \mathbb{E} \left[\sum_{j=0}^{N-1} \left| \sum_{i=l_j}^{\phi(j,k)-1} y_i(t_j) \right|^{2p} |\Lambda_{t_j}|^{2p} \right] h_j \\
&\leq C \sum_{k=0}^{\gamma-1} \rho_k \mathbb{E} \left[\sum_{j=0}^{N-1} \left| \sum_{i=l_j}^{\phi(j,k)-1} \Gamma_{t_j} y_i(t_j) \right|^{2p} \sup_j |\Lambda_{t_j}|^{2p} \right] h \\
&\leq C_{q^*} \sum_{k=0}^{\gamma-1} \rho_k \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=0}^{N-1} \left| \mathbb{E}^{\mathbb{Q}} \left[\sum_{i=l_j}^{\phi(j,k)-1} \Gamma_T \xi_i \mid \mathcal{F}_{t_j} \right] \right| \right)^{4pq^*} \right]^{\frac{1}{2q^*}} \mathbb{E} \left[\sup_j |\Lambda_{t_j}|^{4p} \right]^{\frac{1}{2}} h \\
&\leq C_{p,q^*} \sum_{k=0}^{\gamma-1} \rho_k \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=0}^{N-1} \sum_{i=l_j}^{\phi(j,k)-1} \|\nabla_{x^{(i)}} g\|_{L^\infty} \sup_{0 \leq s \leq T} \mathbb{E}^{\mathbb{Q}} \left[\sup_{0 \leq s \leq T} |\nabla_x X_s| \mid \mathcal{F}_s \right] \right)^{4pq^*} \right]^{\frac{1}{2q^*}} h \\
&\leq C_{p,q^*,S(g)} h.
\end{aligned}$$

Since $|\beta_s| \leq K_f(1 + 2|Z_s|)$, we conclude that $\int_0^\cdot \beta_s \cdot dB_s$ is a BMO-martingale. Using the equivalence of $\text{BMO}_p(\mathbb{R})$ norms of $\int_0^\cdot \beta_s \cdot dB_s$, the upper bound which is independent of r for A_5 is constructed as the following:

$$\begin{aligned}
A_5 &= \sum_{j=0}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} (s_{k+1} - s_k) \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |\beta(u)|^2 du \right)^{2p} |\Lambda_{t_j}|^{2p} \right] \\
&\leq \mathbb{E} \left[\sum_{j=0}^{N-1} \left(\int_{t_j}^{t_{j+1}} |\beta(u)|^2 du \right)^{2p} \sup_j |\Lambda_{t_j}|^{2p} \right] h \\
&\leq \mathbb{E} \left[\left(\int_0^T |\beta(u)|^2 du \right)^{4p} \right]^{\frac{1}{2}} \mathbb{E} \left[\sup_j |\Lambda_{t_j}|^{4p} \right]^{\frac{1}{2}} h \leq C_p h.
\end{aligned}$$

Finally, by realising q is the conjugate of q^* , we summarise the contribution of all terms

$$\begin{aligned}
\sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |Z_t - Z_{t_j}|^{2p} dt \right] &\leq C_p \sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |I_{1,j}(t)|^{2p} + |I_{2,j}(t)|^{2p} + |I_{3,j}(t)|^{2p} dt \right] \\
&\leq C_p \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} 2h_j^{2p} dt + C_p \sum_{j=0}^{N-1} \mathbb{E} [|I_{3,j}(t)|^{2p}] \\
&\leq C_{p,q,S(g)} h.
\end{aligned}$$

Recall (3.15), we get

$$\sum_{j=0}^{N-1} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |Z_t - Z_{t_j}|^2 dt \right)^p \right] \leq C_{p,S(g)} h^p$$

in this scenario.

Step 5. We are now in a position to consider π_1 is a finer partition than π_0 , assuming the function g , $x \mapsto \sigma(t, x)$ and $x \mapsto b(t, x)$ are smooth functions. In this scenario, however, we have $s_k = t_{l_k}$ for some increasing sequence $(l_k)_{k=0}^m$. Noticing that equation (3.15) still holds, what remains is to decompose $Z_t - Z_{t_j}$ for a given $t \in [s_k, s_{k+1})$,

$$\begin{aligned} Z_t - Z_{t_j} &= (\nabla^k Y_t)^\top (\nabla_x X_t)^{-1} \sigma(t, X_t) - (\nabla^k Y_{t_j})^\top (\nabla_x X_{t_j})^{-1} \sigma(t_j, X_{t_j}) \\ &= I_{1,j}(t) + I_{2,j}(t) + I_{3,j}(t), \\ I_{1,j}(t) &:= (\nabla^k Y_t)^\top (\nabla_x X_t)^{-1} (\sigma(t, X_t) - \sigma(t_j, X_{t_j})), \\ I_{2,j}(t) &:= (\nabla^k Y_t)^\top \left((\nabla_x X_t)^{-1} - (\nabla_x X_{t_j})^{-1} \right) \sigma(t_j, X_{t_j}), \\ I_{3,j}(t) &:= (\nabla^k Y_t - \nabla^k Y_{t_j})^\top (\nabla_x X_{t_j})^{-1} \sigma(t_j, X_{t_j}). \end{aligned}$$

Apparently, we are able to construct the bound for $\mathbb{E}[|I_1|^{2p}]$ and $\mathbb{E}[|I_1|^{2p}]$, namely $C_p h_j^p$, if $t \in [t_j, t_{j+1}]$. Similar to step 2, we bound the contribution of I_3 by

$$\sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |I_{3,j}(t)|^{2p} dt \right] \leq \sum_{k=0}^{r-1} \sum_{j=l_k}^{l_{k+1}-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} \mathbb{E} \left[|\nabla^k Y_t - \nabla^k Y_{t_j}|^{2p} \mid \mathcal{F}_{t_j} \right] \Lambda_{t_j}^{2p} dt \right]$$

and introduce a bound on the conditional expectation in the integrand:

$$\begin{aligned} &\mathbb{E} \left[|\nabla^k Y_t - \nabla^k Y_{t_j}|^{2p} \mid \mathcal{F}_{t_j} \right] \\ &\leq C_p \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^p + \left| (t_j - t_{j+1}) \sum_{i=k}^r y_i(t_j) \right|^{2p} \right. \\ &\quad \left. + \left(\int_{t_j}^{t_{j+1}} |\beta(u)|^2 du \right)^{2p} + \left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^{2p} \mid \mathcal{F}_{t_j} \right]. \end{aligned}$$

Therefore, we have the same type of estimation using terms A_i

$$\begin{aligned} &\sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |I_{3,j}(t)|^{2p} dt \right] \leq C_p \sum_{i=1}^4 A_i, \\ A_1 &:= \sum_{k=0}^{r-1} \sum_{j=l_k}^{l_{k+1}-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^{2p} \mid \mathcal{F}_{t_j} \right] |\Lambda_{t_j}|^{2p} dt \right], \end{aligned}$$

$$\begin{aligned}
A_2 &:= \sum_{k=0}^{r-1} \sum_{j=l_k}^{l_{k+1}-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^p \middle| \mathcal{F}_{t_j} \right] |\Lambda_{t_j}|^{2p} dt \right], \\
A_3 &:= \sum_{k=0}^{r-1} \sum_{j=l_k}^{l_{k+1}-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} \mathbb{E} \left[\left| (t_j - t_{j+1}) \sum_{i=k}^r y_i(t_j) \right|^{2p} \middle| \mathcal{F}_{t_j} \right] |\Lambda_{t_j}|^{2p} dt \right], \\
A_4 &:= \sum_{k=0}^{r-1} \sum_{j=l_k}^{l_{k+1}-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |\beta(u)|^2 du \right)^{2p} \middle| \mathcal{F}_{t_j} \right] |\Lambda_{t_j}|^{2p} dt \right].
\end{aligned}$$

By applying the BDG inequality and the reverse Hölder inequality, we first give a upper bound on A_1 ,

$$\begin{aligned}
A_1 &\leq \mathbb{E} \left[\sum_{k=0}^{r-1} \sum_{j=l_k}^{l_{k+1}-1} \left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=k}^r \eta_i(u) \right|^2 du \right)^{2p} \sup_j |\Lambda_{t_j}|^{2p} \right] h \\
&\leq C_q \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left| \sum_{i \geq \phi(j)} \eta_i(u) \right|^2 du \right)^{4pq^*} \right]^{\frac{1}{2q^*}} \mathbb{E} \left[\sup_j |\Lambda_{t_j}|^{4p} \right]^{\frac{1}{2}} h,
\end{aligned}$$

where $\phi : \{0, \dots, N-1\} \mapsto \{0, \dots, r-1\}$ defined as $\phi(j) = k$ for all $l_k \leq j \leq l_{k+1}-1$. Define $\tilde{\phi} = \sup_j \{j \mid \phi(j) \leq i\}$ with the convention that $\sup \emptyset = -1$. As a consequence, we get

$$\begin{aligned}
A_1 &\leq C_{p,q} \mathbb{E}^{\mathbb{Q}} \left[\left| \sum_{k=0}^{r-1} \tilde{\xi} \left(t_{\tilde{\phi}(k)+1} \right) - \tilde{\xi}(0) \right|^{8pq^*} \right]^{\frac{1}{2q^*}} h \\
&\leq C_{p,q} h \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{k=0}^r \|\nabla_{x^i} g\|_{L^\infty} \sup_{0 \leq s \leq T} \mathbb{E}^{\mathbb{Q}} \left[\sup_{0 \leq s \leq T} |\nabla_x X_s| \middle| \mathcal{F}_s \right] \right)^{8pq^*} + \left| \sum_{k=0}^r \tilde{\xi}(0) \right|^{8pq^*} \right]^{\frac{1}{2q^*}} \\
&\leq C_{p,q,S(g)} h
\end{aligned}$$

by rewriting the first expectation via Lemma 3.2.10.

A by-product of above calculation is $A_2 \leq C_{p,q} h$, while $A_4 \leq C_p h$ by similar approach as in step 3. Finally, we aim to get a bound on A_3 , by transforming $y_i(t_j)$ to $\Gamma_{t_j} y_i(t_j)$ and rewriting this random variable as a conditional expectation of $\Gamma_T y_i(T)$ under $\mathbb{Q}$. More explicitly,

$$\begin{aligned}
A_3 &\leq \sum_{k=0}^{r-1} \sum_{j=l_k}^{l_{k+1}-1} \mathbb{E} \left[\left| (t_j - t_{j+1}) \sum_{i=k}^r y_i(t_j) \right|^{2p} |\Lambda_{t_j}|^{2p} dt \right] h \\
&\leq \mathbb{E} \left[\left(\sum_{k=0}^{r-1} \sum_{j=l_k}^{l_{k+1}-1} \left| (t_j - t_{j+1}) \sum_{i=k}^r y_i(t_j) \right| \right)^{4p} \right]^{\frac{1}{2}} \mathbb{E} \left[\sup_j |\Lambda_{t_j}|^{4p} dt \right]^{\frac{1}{2}} h
\end{aligned}$$

$$\begin{aligned}
&\leq C_{p,q} \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=0}^{N-1} h_j \left| \mathbb{E}^{\mathbb{Q}} \left[\Gamma_T \sum_{i \geq \phi(j)} y_i(T) \mid \mathcal{F}_{t_j} \right] \right| \right)^{4pq^*} \right]^{\frac{1}{2q^*}} h \\
&\leq C_{p,q} \mathbb{E}^{\mathbb{Q}} \left[\left(\sum_{j=0}^{N-1} h_j \sum_{i \geq \phi(j)} \|\nabla_{x^{(i)}} g\|_{L^\infty} \sup_{0 \leq s \leq T} \mathbb{E}^{\mathbb{Q}} \left[\sup_{0 \leq s \leq T} |\nabla_x X_s| \mid \mathcal{F}_s \right] \right)^{4pq^*} \right]^{\frac{1}{2q^*}} h \\
&\leq C_{p,q,S(g)} h.
\end{aligned}$$

as we required. Summarising the contribution of each term, we obtain our desirable result.

Step 6. The final task is to remove the continuity condition on g, f, b and σ . Let g, f, b, σ satisfy the usual assumption, we assume g^ϵ is the mollifier of g and $g^\epsilon \rightarrow g$ as $\epsilon \rightarrow 0$. We define, similarly to g^ϵ , f^ϵ , b^ϵ and σ^ϵ , while we denote X^ϵ as the solution to SDE corresponding to b^ϵ and σ^ϵ . For all $p \geq 1$, we can derive

$$\begin{aligned}
&\lim_{\epsilon \rightarrow 0} \mathbb{E} [|g^\epsilon(X_{t_0}^\epsilon, \dots, X_{t_r}^\epsilon) - g(X_{t_0}, \dots, X_{t_r})|^p] \\
&\leq C_p \lim_{\epsilon \rightarrow 0} (\mathbb{E} [|g^\epsilon(X_{t_0}^\epsilon, \dots, X_{t_r}^\epsilon) - g(X_{t_0}^\epsilon, \dots, X_{t_r}^\epsilon)|^p] \\
&\quad - \mathbb{E} [|g(X_{t_0}^\epsilon, \dots, X_{t_r}^\epsilon) - g(X_{t_0}, \dots, X_{t_r})|^p]) \\
&= 0,
\end{aligned}$$

by the dominated convergence theorem (DCT) together with the assumption on g . Meanwhile, the stability theorem suggests that the solution (Y^ϵ, Z^ϵ) to the BSDE with smooth coefficients, will converge to the original solution (Y, Z) in $\mathbb{S}^{2p}(\mathbb{R}) \times \mathbb{H}^{2p}(\mathbb{R}^d)$ for all $p \geq 1$, because we further have

$$\lim_{\epsilon \rightarrow 0} \mathbb{E} \left[\left(\int_0^T |f^\epsilon(t, Y_t^\epsilon, Z_t^\epsilon) - f(t, Y_t^\epsilon, Z_t^\epsilon)| dt \right)^{2pq^*} \right] = 0.$$

By Fubini theorem, we obtain $\lim_{\epsilon \rightarrow 0} \int_0^T \mathbb{E}[(Z_t^\epsilon - Z_t)^2] dt = 0$. Hence there exists a subsequence of $(Z^\epsilon)_\epsilon$, namely $(Z^{\tilde{\epsilon}})_{\tilde{\epsilon}}$, such that $\lim_{\tilde{\epsilon} \rightarrow 0} \mathbb{E}[(Z_t^{\tilde{\epsilon}} - Z_t)^2] = 0$ t -almost surely. A similar argument allows us to conclude that there exists a further subsequence $(Z^{\epsilon_1})_{\epsilon_1}$ of $(Z^{\tilde{\epsilon}})_{\tilde{\epsilon}}$, such that $Z_t^{\epsilon_1} \rightarrow Z_t$ t -almost surely and ω -almost surely. For simplicity, we set the $(Z^\epsilon)_\epsilon$ as the new sequence $(Z^{\epsilon_1})_{\epsilon_1}$. As there exists an $L^{2p}(\Omega; \mathbb{R}^d)$ bound for both $(Z^\epsilon)_\epsilon$ and Z_t , we apply DCT to obtain

$$F^\epsilon(t) := \mathbb{E}[|Z_t^\epsilon - Z_t|^{2p}] \rightarrow 0, \text{ as } \epsilon \rightarrow 0, \quad t\text{-almost surely},$$

and $F^\epsilon(t)$ is bounded t -almost surely and uniformly in ϵ . Back to our proof on path-regularity, we easily see that for any $t \in [t_j, t_{j+1}]$,

$$|Z_t - Z_{t_i}|^{2p} \leq |Z_t^\epsilon - Z_t|^{2p} + |Z_t^\epsilon - Z_{t_j}^\epsilon|^{2p} + |Z_{t_j}^\epsilon - Z_{t_j}|^{2p}. \quad (3.20)$$

Regarding the right hand side of the inequality, we take limit on $\epsilon \rightarrow 0$ to kill first and last term by the DCT. We have managed to show our result for Z^ϵ in step 2 to step 5. Moreover, we observe the constant $C_{p,q}$ is independent of ϵ , thus

$$\sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |Z_t - Z_{t_j}|^{2p} dt \right] \leq C_{p,S(g)} h,$$

which implies the conclusion of this theorem. $\square$

Theorem 3.1.1 can be viewed as an extension of Theorem 5.6 in [47] for QBSDEs that have terminal conditions depend on multiple points on the paths. The next theorem is a consequence of Theorem 3.1.1, which allows us to prove Zhang's path regularity theorem in the context of quadratic BSDE.

Proof of Theorem 3.1.2. Following a similar procedure as step 1 in the proof of Theorem 3.1.1, we consider the uniform L^p bound of Z_t to derive

$$\begin{aligned} \mathbb{E} \left[\sup_{u \in [s,t]} |Y_u - Y_s|^{2p} \right] &\leq C_p \left((t-s)^{2p} + (t-s)^{2p-1} \int_s^t \mathbb{E} [|Z_\tau|^{4p}] d\tau \right. \\ &\quad \left. + (t-s)^{p-1} \int_s^t \mathbb{E} [|Z_\tau|^{2p}] d\tau \right) \\ &\leq C_p (t-s)^p. \end{aligned}$$

The proof of the path-regularity of Z is essentially a density argument. Let $(\pi(r))_{r \geq 1}$ with $\pi(r) : 0 = s_0 \leq s_1 \leq \dots \leq s_r = T$ be a sequence of partitions such that $\lim_{r \rightarrow \infty} |\pi(r)| = 0$, we consider the smooth function $g_{\pi(r)}(X(s_0), \dots, X(s_r))$ which is bounded and serves as the approximation to g . Moreover, $g_{\pi(r)}$ satisfies $S(g_{\pi(r)}) < C$ for some C being independent of r . We denote $(Y^{\pi(r)}, Z^{\pi(r)})$ as the solution to BSDE with the terminal condition $g_{\pi(r)}(X_{s_0}, \dots, X_{s_r})$. According to remark 3.2 and the DCT, there exists a subsequence of $(\pi(r))_r$, namely $\tilde{\pi}(r)$ such that $Z_t^{\tilde{\pi}(r)} \rightarrow Z_t$ in L^p for any $p \geq 2$. The result follows directly. $\square$

3.2.3 Averaging Z_t in time

We claim that the BMO property is not lost if we consider the “average of Z over a small time interval” , whereas the proof has been deferred to Section A.2.

Lemma 3.2.11 (Lemma 2.1 in [12]). *Let (Y, Z) be the solution for BSDE (2.11) satisfying the same condition as Theorem 2.4.3. Given a partition $\pi_1 : 0 = t_0 < t_1 < \dots < t_N = T$, we define the process $\bar{Z}(t)$ as*

$$\bar{Z}(t) := \bar{Z}_j := \frac{1}{h_j} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} Z_s \, ds \, \middle| \, \mathcal{F}_{t_j} \right], \quad \forall t \in [t_j, t_{j+1}), \quad 0 \leq j \leq N-1,$$

together with $\bar{Z}(T) := 0$. We get $\|\bar{Z}\|_{\mathbb{H}_{BMO}^2(\mathbb{R}^d)} < 2\|Z\|_{\mathbb{H}_{BMO}^2(\mathbb{R}^d)}$.

Due to the nature of BMO norm, providing a sharp estimate on the difference of two BMO martingales is a challenge, while the relationship between various L^p norms of BMO martingales are unknown. This is crucial, because the following term appears within the error analysis:

$$A = \mathbb{E} \left[\left(\sum_{j=0}^N \int_{t_j}^{t_{j+1}} |Z_s - \bar{Z}_j|^2 \, ds \right)^{pq^*} \right] = \mathbb{E} \left[\left(\int_0^T |Z_s - \bar{Z}_s|^2 \, ds \right)^{pq^*} \right].$$

We aim to show A is of order $O(h^p)$ where $h = \sup_j |h_j|$. A sufficient condition is, provided $s \in [t_j, t_{j+1})$,

$$\mathbb{E} [|Z_s - Z_{t_j}|^{2p}] \leq Ch_j^p,$$

which requires the path-regularity of $\mathbb{E}[|\nabla_x Y_s - \nabla_x Y_{t_j}|^{pq^*}]$. The other approach is controlling the expectation by the BMO_{pq^*} norm, followed by implementing the equivalence of BMO norms. Unfortunately, it remains as an open question whether taking conditional expectations preserve the regularity for QBSDEs:

$$\mathbb{E} \left[\int_t^T |Z_s - \bar{Z}(s)|^2 \, ds \, \middle| \, \mathcal{F}_t \right] \leq Ch \quad a.s.?$$

We try to overcome this obstacle by following the proof of [12, Proposition 3.2] i.e. removing the impact of q^* by re-introducing a new measure $\mathbb{Q}_N$.

Proposition 3.2.12. *Consider a partition $\pi_1 = 0 = t_0 < t_1 < \dots < t_N = T$ with $h = \sup_i |t_{i+1} - t_i|$. Let $\{\chi_i\}_{i=0}^{N-1}$ be a sequence of random variables such that $\chi_i \in L^\infty(\Omega, \mathcal{F}_{t_{i+1}}, \mathbb{P})$, $\mathbb{E}[\chi_i | \mathcal{F}_{t_i}] = 1$ and $\chi_i \in [\epsilon, 2 - \epsilon]$ for all $0 \leq i \leq N-1$*

with some $\epsilon \in (0, 1)$ independent of N . Define the process $\hat{\chi}(t)$ and the probability measure⁵ $\mathbb{Q}_N$ by

$$\hat{\chi}(t) = \prod_{t_i \leq t} \chi_i, \quad \frac{d\mathbb{Q}_N}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = \hat{\chi}(t).$$

Suppose (Y, Z) is the solution to the QBSDE which satisfies the same condition as in Theorem 3.1.2. If $\hat{\chi}(t)$ satisfies the reverse Hölder inequality (Theorem 2.3.14) for some q that is independent of N , then

$$\mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \left(\int_{t_j}^{t_{j+1}} |Z_t - \bar{Z}(t_j)|^2 dt \right)^v \middle| \mathcal{F}_{t_i} \right]^p \right] \leq C(p, q, L_\Phi, v, \epsilon) h^{pv},$$

for any $v, p \geq 1$.

Proof. See Appendix A.2. □

Remark 3.4. In fact, $\hat{\chi}(t) = \mathcal{E}(M)_t$, where $M_t = \sum_{t_{i+1} \leq t} (\chi_i - 1)$ is a càdlàg martingale. Within the proof of Proposition 3.2.12, we repeatedly use the following inequality,

$$\mathbb{E} [|F(t_{i+1})| \mid \mathcal{F}_{t_i}] = \mathbb{E}^{\mathbb{Q}_N} \left[\frac{\hat{\chi}(t_i)}{\hat{\chi}(t_{i+1})} |F(t_{i+1})| \mid \mathcal{F}_{t_i} \right] \leq \frac{1}{\epsilon} \mathbb{E}^{\mathbb{Q}_N} [|F(t_{i+1})| \mid \mathcal{F}_{t_i}],$$

where $F(t_{i+1})$ is an $\mathcal{F}_{t_{i+1}}$ -measurable random variable. That is, we can trade $\mathbb{Q}_N$ with $\mathbb{P}$ at a cost of some constant depending on ϵ . This phenomenon will be revisited in the next few sections.

3.3 The Modified (inner) explicit scheme

Cheridito et al. [13] provided an example such that comparison theorem for the classical BTZ scheme failed under the quadratic assumption, supporting the fact that the truncation of the driver is inevitable. As the challenge is the instability caused by the growth condition on the Z component, the main strategy involves restricting the growth of Z component by truncating the driver Z with a suitable truncation parameter. In addition, the variable H_i is introduced to replace the Brownian increment ΔB_{t_i} in the classical BTZ scheme to improve the stability of the scheme. These two techniques preserves many essential properties, such as the comparison principle.

More explicitly, we let $\pi : 0 = t_0 < t_1 < \dots < t_N = T$ be a partition, where $h_i = t_{i+1} - t_i$ and $h = \sup_{0 \leq i \leq N-1} h_i$. Our numerical scheme involves a sequence of independent random vectors $(H_i)_{i=0}^{N-1}$ such that

⁵By definition, $\mathbb{Q}_N(\Omega) = \mathbb{E} \left[\frac{d\mathbb{Q}_N}{d\mathbb{P}} \right] = 1$

- H_i is a $\mathbb{R}^d$ -valued random vector and is $\mathcal{F}_{t_{i+1}}$ -measurable, for all $0 \leq i \leq N-1$;
- for all $0 \leq i \leq N-1$, H_i satisfies

$$\mathbb{E}[H_i \mid \mathcal{F}_{t_i}] = 0 \in \mathbb{R}^d, \quad \mathbb{E}[H_i(H_i)^\top \mid \mathcal{F}_{t_i}] = \mathbb{E}[H_i(H_i)^\top] = \frac{c_i}{h_i} I_{d \times d}, \quad (3.21)$$

where $\lambda < c_i < \Lambda$ for some constants (independent of i) $\lambda, \Lambda > 0$.

The *modified inner (centred) explicit BTZ scheme* is

$$\begin{aligned} Y_i^\pi &= \mathbb{E}[Y_{i+1}^\pi + f_i(t_i, \mathbb{E}[Y_{i+1}^\pi \mid \mathcal{F}_{t_i}], Z_i^\pi) h_i \mid \mathcal{F}_{t_i}], \\ Z_i^\pi &= \mathbb{E}[Y_{i+1}^\pi H_i \mid \mathcal{F}_{t_i}] \\ Y_N^\pi &= \xi, \quad Z_N^\pi = 0, \end{aligned} \quad (3.22)$$

such that $(Y^\pi, Z^\pi)^6$ are defined iteratively.

3.3.1 Properties of the explicit scheme

In most cases, H_i is the scaled Brownian increment, but the numerical approximation of the Brownian increment might not be centered. This motivated us to consider a more general scheme, where the original collection of variables $\{H_i\}_i$ is replaced by $\mathbb{R}^d$ -valued $\{\hat{H}_i\}_i$ with mean not necessarily zero. Therefore, for all $i \geq 1$, $\hat{H}_i$ is $\mathcal{F}_{t_{i+1}}$ -measurable and independent of $\mathcal{F}_{t_i}$, such that

$$\hat{\mu}_i := \mathbb{E}[\hat{H}_i \mid \mathcal{F}_{t_i}] = \mathbb{E}[\hat{H}_i], \quad \mathbb{V}[\hat{H}_i(\hat{H}_i)^\top \mid \mathcal{F}_{t_i}] = \mathbb{V}[\hat{H}_i(\hat{H}_i)^\top] = \frac{c_i}{h_i} I_{d \times d},$$

where $\hat{\mu}_i, c_i$ satisfies that there exist constants C, λ, Λ such that

$$\hat{\mu} := \sup_i |\hat{\mu}_i| \leq C, \quad 0 < \lambda \leq c_i \leq \Lambda < \infty.$$

To distinguish with the centred scheme, we write

$$\begin{aligned} \hat{Y}_i^\pi &= \mathbb{E}[\hat{Y}_{i+1}^\pi + f_i(t_i, \mathbb{E}[\hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i}], \hat{Z}_i^\pi) h_i \mid \mathcal{F}_{t_i}], \\ \hat{Z}_i^\pi &= \mathbb{E}[\hat{Y}_{i+1}^\pi \hat{H}_i \mid \mathcal{F}_{t_i}], \\ \hat{Y}_N^\pi &= \xi, \quad \hat{Z}_N^\pi = 0. \end{aligned} \quad (3.23)$$

Apparently, the centred scheme is a special case of $(\hat{Y}^\pi, \hat{Z}^\pi)$. The aim of this subsection is to conclude the boundedness in Y and BMO property in Z , as long as $\mathbb{E}[\hat{H}_i]$ is small enough.

⁶Since the numerical solution (Y^π, Z^π) depends on the triplet $(\xi, \{f_i\}_i, \{H_i\}_i)$ and the partition π , we say (Y^π, Z^π) is the solution that corresponds to the triplet $(\xi, \{f_i\}_i, \{H_i\}_i)$.

By the definition, we are allowed to write

$$\hat{Z}_i^\pi = \mathbb{E} \left[A_i \hat{H}_i \mid \mathcal{F}_{t_i} \right] + \mathbb{E} \left[\hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] \hat{\mu}_i, \quad A_i := \hat{Y}_{i+1}^\pi - \mathbb{E} \left[\hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right],$$

which suggests that there exists an $\mathcal{F}_{t_{i+1}}$ -measurable variable v_i satisfying

$$\begin{aligned} \hat{Y}_{i+1}^\pi &= \hat{Y}_i^\pi - f_i \left(t_i, \mathbb{E} \left[\hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right], \hat{Z}_i^\pi \right) h_i + A_i, \\ v_i &= A_i - \frac{h_i}{c_i} \left(\hat{Z}_i^\pi - \mathbb{E} \left[\hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] \hat{\mu}_i \right) \cdot \left(\hat{H}_i - \hat{\mu}_i \right). \end{aligned} \quad (3.24)$$

If $\hat{Y}_{i+1}^\pi$ is bounded, then $\mathbb{E}[v_i | \mathcal{F}_{t_i}] = 0$, $\mathbb{E}[(v_i)^2 | \mathcal{F}_{t_i}] < \infty$ and v_i is orthogonal with $\hat{H}_i - \hat{\mu}_i$ i.e.

$$\begin{aligned} &\mathbb{E} \left[v_i \left(\hat{H}_i - \hat{\mu}_i \right) \mid \mathcal{F}_{t_i} \right] \\ &= \mathbb{E} \left[A_i \left(\hat{H}_i - \hat{\mu}_i \right) \mid \mathcal{F}_{t_i} \right] \\ &\quad - \frac{h_i}{c_i} \mathbb{E} \left[\left(\left(\hat{Z}_i^\pi - \mathbb{E} \left[\hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] \hat{\mu}_i \right)^\top \left(\hat{H}_i - \hat{\mu}_i \right) \right) \left(\hat{H}_i - \hat{\mu}_i \right) \mid \mathcal{F}_{t_i} \right] \\ &= \hat{Z}_i^\pi - \mathbb{E} \left[\hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] \hat{\mu}_i - \frac{h_i}{c_i} \left(\hat{Z}_i^\pi - \mathbb{E} \left[\hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] \hat{\mu}_i \right) \cdot \frac{c_i}{h_i} I_{d \times d} \\ &= 0 \in \mathbb{R}^d. \end{aligned}$$

We are now in the position to state the relevant assumptions.

Assumption [A5] (The driver being locally Lipschitz in Z). *Given $\mathcal{F}_{t_i} \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d)$ measurable functions $f_i : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ for all $1 \leq i \leq N$, we assume the following:*

- (i) $f_i(t, 0, 0) \in L^2(\mathcal{F}_{t_i})$ for all $t \in [0, T]$ and there exist $L_f \geq 0$ and $L_z(n) \geq 0$ where $L_z(n)$ depends on some constant $n > 0$, such that for all y_1, y_2, z_1, z_2

$$|f_i(t, y_1, z_1) - f_i(t, y_2, z_2)| \leq L_f |y_1 - y_2| + L_z(n) |z_1 - z_2|,$$

- (ii) recall $h = \sup_i h_i$, there exists a constant $\epsilon \in (0, 1]$, independent of N, n , such that

$$hL_f + \sup_{0 \leq i \leq N-1} |h_i \hat{H}_i L_z(n)| \leq 1 - \epsilon.$$

Assumption [A6] (Bounds on ξ, f_i , and $\hat{\mu}$). *Given $\mathcal{F}_{t_i} \otimes \mathcal{B}([0, T]) \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^d)$ -measurable functions $f_i : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ satisfying assumption [A5] for all $1 \leq i \leq N$.*

(i) In addition, $\xi \in L^\infty(\mathcal{F}_T)$ and $\sup_{1 \leq i \leq N} |f_i(t, 0, 0)| \leq L_0$ where L_0 is independent of N .

(ii) Moreover, $\hat{\mu}L_z(n) < 1$, where $L_z(n)$ is the Lipschitz constant of z which depends on n .

Lemma 3.3.1. Let $\xi^l \in L^2$ and f^l satisfy assumption [A5] (i) for all $l = 1, 2$, suppose $(\hat{Y}^{\pi,l}, \hat{Z}^{\pi,l})_l$ be their corresponding solutions to the numerical scheme (3.23). Moreover, we assume $f_k^i(t_k, \hat{Y}_k^\pi, \hat{Z}_k) = f^i(t_k, \hat{Y}_k^\pi, \hat{Z}_k)$ for all $0 \leq k \leq N$. We write $\Delta\xi := \xi^1 - \xi^2$, $\Delta f_k := f_k^1(t_k, \hat{Y}_k^{\pi,2}, \hat{Z}_k^{\pi,2}) - f_k^2(t_k, \hat{Y}_k^{\pi,2}, \hat{Z}_k^{\pi,2})$, $\Delta\hat{Y}^\pi := \hat{Y}^{\pi,1} - \hat{Y}^{\pi,2}$ and $\Delta\hat{Z}^\pi := \hat{Z}^{\pi,1} - \hat{Z}^{\pi,2}$.

(i) For all $0 \leq i \leq N$, we have,

$$\Delta\hat{Y}_i^\pi = \mathbb{E} \left[\Upsilon_i \Delta\hat{Y}_{t_N}^\pi \mid \mathcal{F}_{t_i} \right] + \mathbb{E} \left[\sum_{k=i}^{N-1} h_k \Delta f_k \frac{\Upsilon_i}{\Upsilon_k} \mid \mathcal{F}_{t_i} \right], \quad (3.25)$$

where Υ is defined as

$$\begin{aligned} \Upsilon_i &:= \prod_{j=i}^{N-1} \psi_j, \quad \psi_i := 1 + h_i \alpha_i + h_i \beta_i \cdot \hat{H}_i, \\ \alpha_j &= \frac{f_j^1 \left(t_j, \mathbb{E} \left[\hat{Y}_{j+1}^{\pi,1} \mid \mathcal{F}_{t_j} \right], \hat{Z}_j^{\pi,1} \right) - f_j^1 \left(t_j, \mathbb{E} \left[\hat{Y}_{j+1}^{\pi,2} \mid \mathcal{F}_{t_j} \right], \hat{Z}_j^{\pi,1} \right)}{\mathbb{E} \left[\Delta\hat{Y}_{j+1}^\pi \mid \mathcal{F}_{t_j} \right]} \\ &\quad \times \mathbb{1}_{\{\mathbb{E}[\Delta\hat{Y}_{j+1}^\pi \mid \mathcal{F}_{t_j}] \neq 0\}}, \\ \beta_j &= \frac{f_j^1 \left(t_j, \mathbb{E} \left[\hat{Y}_{j+1}^{\pi,2} \mid \mathcal{F}_{t_j} \right], \hat{Z}_j^{\pi,1} \right) - f_j^1 \left(t_j, \mathbb{E} \left[\hat{Y}_{j+1}^{\pi,2} \mid \mathcal{F}_{t_j} \right], \hat{Z}_j^{\pi,2} \right)}{|\Delta\hat{Z}_j^\pi|^2} \Delta\hat{Z}_j^\pi \mathbb{1}_{\{\Delta\hat{Z}_j^\pi \neq 0\}}. \end{aligned}$$

(ii) If we further assume [A5] (ii) is imposed, we can conclude $\hat{Y}_i^{\pi,1} \geq \hat{Y}_i^{\pi,2}$ for a $0 \leq i \leq N$, when $\hat{Y}_N^{\pi,1} \geq \hat{Y}_N^{\pi,2}$ and $f_j^1(t_j, \hat{Y}_j^{\pi,2}, \hat{Z}_j^{\pi,2}) \geq f_j^2(t_j, \hat{Y}_j^{\pi,2}, \hat{Z}_j^{\pi,2})$ for all $N \geq j \geq i$.

(iii) Moreover, let assumption [A6] be imposed. Then the bound on Y is given as

$$\sup_{0 \leq i \leq N} |\hat{Y}_i^\pi| \leq (\|\xi\|_{L^\infty} + C_0 T) \exp(1 + L_f T).$$

Proof. We make use of the expression $\Delta\hat{Z}_i^\pi = \mathbb{E} \left[\Delta\hat{Y}_{i+1}^\pi \hat{H}_i \mid \mathcal{F}_{t_i} \right]$ to obtain the iterative principle

$$\Delta\hat{Y}_i^\pi = \mathbb{E} \left[\Delta\hat{Y}_{i+1}^\pi + h_i \alpha_i \mathbb{E} \left[\Delta\hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] + h_i \Delta\hat{Z}_i^\pi \cdot \beta_i + h_i \Delta f_i \mid \mathcal{F}_{t_i} \right]$$

$$= \mathbb{E} \left[\psi_i \Delta \hat{Y}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] + \mathbb{E} [h_i \Delta f_i \mid \mathcal{F}_{t_i}]. \quad (3.26)$$

An application of induction yields the first result of the Lemma.

The comparison principle is still valid, as all $\psi_i > 0$ by our assumption. Eventually, the upper bound of $\hat{Y}^\pi$ can be obtained by comparing $\hat{Y}_i^{\pi,2}$, where $\hat{Y}_i^{\pi,2}$ is the solution to the numerical scheme with generators $\|\xi\|_{L^\infty}$ and F_i , where

$$F_i(t, y, z) = L_0 + L_f |y| + L_z(n) |z|.$$

Here $L_z(n)$ is the Lipschitz constant on z which corresponds to $\{f_i\}_i$. As ξ is deterministic, we follow a simple induction argument to observe that $\hat{Y}^{\pi,2}$ is deterministic and $\hat{Z}_i^{\pi,2} = \mathbb{E}[\hat{Y}_{i+1}^{\pi,2} \hat{H}_i \mid \mathcal{F}_{t_i}] = \hat{Y}_{i+1}^{\pi,2} \hat{\mu}_i$, which is also non-random. By definition,

$$f_i(t, y, z) \leq L_0 + L_f |y| + L_z(n) |z| = F_i(t, y, z)$$

for all t, y, z , which implies $\hat{Y}^{\pi,1} \leq \hat{Y}^{\pi,2}$ by the comparison result.

It remains to bound $\hat{Y}_i^{\pi,2}$ with a constant independent of N . As $\hat{Y}_i^{\pi,2}$ satisfies the following difference equation, we see that

$$\begin{aligned} \hat{Y}_i^{\pi,2} &= \hat{Y}_{i+1}^{\pi,2} + \left(L_0 + L_f \hat{Y}_{i+1}^{\pi,2} + L_z(n) |\hat{\mu}_i| \hat{Y}_{i+1}^{\pi,2} \right) h_i \\ &\leq \hat{Y}_{i+1}^{\pi,2} (1 + (L_f + 1) h_i) + L_0 h_i \\ &\leq \prod_{i=0}^{N-1} (1 + (L_f + 1) h_i) \|\xi\|_{L^\infty} + \sum_{i=0}^{N-1} \left(\prod_{j=0}^{i-1} (1 + (L_f + 1) h_j) \right) L_0 h_i \\ &\leq (\|\xi\|_{L^\infty} + L_0 T) \exp((1 + L_f)T). \end{aligned}$$

Meanwhile, the lower bound can be obtained by using symmetry. □

The scheme is rather robust if we want to construct a bound on Y , since the error $\hat{\mu}$ is allowed to be of size $\hat{\mu} \sim O(\sup_i h_i |\hat{H}_i|)$. To check the BMO property of Z , we impose the quadratic structure, which is the last ingredient.

Assumption [A7] (Quadratic structure). *Given measurable functions $f_i : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ satisfying assumption [A6] for all $0 \leq i \leq N-1$ (and hence they satisfy [A5]), we further assume*

(i) *For all t, y_1, y_2, z_1, z_2 , there exists a constant $K_f > 0$ such that*

$$|f_i(t, y_1, z_1) - f_i(t, y_2, z_2)| \leq L_f |y_1 - y_2| + K_f (1 + |z_1| + |z_2|) (|z_1 - z_2|);$$

(ii) *$\hat{\mu} L_z(n) < 1$, where $L_z(n)$ corresponds to the Lipschitz constant for Z .*

Proposition 3.3.2 (Preservation of BMO). *Suppose assumption [A7] is imposed on the triplet $(\xi, \{f_i\}_i, \{H_i\}_i)$. Let (Y^π, Z^π) be the solution which corresponds to $(\xi, \{f_i\}_i, \{H_i\}_i)$.*

We define $\hat{Z}^\pi(s) := \hat{Z}_i^\pi$ for all $s \in [t_i, t_{i+1})$, $0 \leq i \leq N$. Then $\|\hat{Z}^\pi(\cdot)\|_{\mathbb{H}_{BMO}^2(\mathbb{R}^d)} < C$ for some C independent of π .

Proof. We adapt the strategy in the proof of [12, Proposition 2.7]. We divide the terminal condition as well as $f_i(t_i, 0, 0)$ into K -segments, while we consider numerical schemes $(\hat{Y}^{(l),\pi}, \hat{Z}^{(l),\pi})_{l=1}^K$ with each $(\hat{Y}^{(l),\pi}, \hat{Z}^{(l),\pi})$ corresponding to the terminal condition $\xi^l = \frac{\xi}{K}$ and the driver

$$\begin{aligned} \tilde{f}_i^l(t, y, z) &= \frac{f_i(t, 0, 0)}{K} + f_i \left(t, y + \sum_{m=1}^{l-1} \mathbb{E} \left[\hat{Y}_{i+1}^{(m),\pi} \mid \mathcal{F}_{t_i} \right], z + \sum_{m=1}^{l-1} \hat{Z}_i^{(m),\pi} \right) \\ &\quad - f_i \left(t, \sum_{m=1}^{l-1} \mathbb{E} \left[\hat{Y}_{i+1}^{(m),\pi} \mid \mathcal{F}_{t_i} \right], \sum_{m=1}^{l-1} \hat{Z}_i^{(m),\pi} \right), \end{aligned}$$

where $\sum_{i=1}^0 = 0$ by convention. By the linearity and the uniqueness of the scheme, we have $\hat{Y}_i^\pi = \sum_{l=1}^K \hat{Y}_i^{(l),\pi}$ and $\hat{Z}_i^\pi = \sum_{l=1}^K \hat{Z}_i^{(l),\pi}$ i.e.

$$\begin{aligned} \sum_{l=1}^K \hat{Y}_i^{(l),\pi} &= \mathbb{E} \left[\sum_{l=1}^K \hat{Y}_{i+1}^{(l),\pi} \mid \mathcal{F}_{t_i} \right] + \sum_{l=1}^K \tilde{f}_i^l \left(t_i, \mathbb{E} \left[\hat{Y}_{i+1}^{(l),\pi} \mid \mathcal{F}_{t_i} \right], \hat{Z}_i^{(l),\pi} \right), \\ \sum_{l=1}^K \hat{Z}_i^{(l),\pi} &= \mathbb{E} \left[\sum_{l=1}^K \hat{Y}_{i+1}^{(l),\pi} \hat{H}_i \mid \mathcal{F}_{t_i} \right], \\ \sum_{l=1}^K \hat{Y}_N^{(l),\pi} &= \xi, \quad \sum_{l=1}^K \hat{Z}_N^{(l),\pi} = 0. \end{aligned}$$

Step 1. The remaining goal is to choose an appropriate K which is independent N and prove the BMO property for each $\hat{Z}^{(l),\pi}$. A linearisation argument leads to the equation of the type (3.24)

$$\begin{aligned} \hat{Y}_{i+1}^{(l),\pi} &= \hat{Y}_i^{(l),\pi} - \tilde{f}_i^l \left(t_i, \mathbb{E} \left[\hat{Y}_{i+1}^{(l),\pi} \mid \mathcal{F}_{t_i} \right], \hat{Z}_i^{(l),\pi} \right) h_i + A_i, \quad A_i = \hat{Y}_{i+1}^{(l),\pi} - \mathbb{E} \left[\hat{Y}_{i+1}^{(l),\pi} \mid \mathcal{F}_{t_i} \right], \\ v_i &= A_i - \frac{h_i}{c_i} \left(\hat{Z}_i^{(l),\pi} - \mathbb{E} \left[\hat{Y}_{i+1}^{(l),\pi} \mid \mathcal{F}_{t_i} \right] \hat{\mu}_i \right) \cdot \left(\hat{H}_i - \hat{\mu}_i \right), \end{aligned}$$

where v_i is a $\mathcal{F}_{t_{i+1}}$ -measurable and L^2 integrable random variable. More explicitly, we find v_i obeys

$$\mathbb{E} [v_i \mid \mathcal{F}_{t_i}] = \mathbb{E} \left[v_i \left(\hat{H}_i - \hat{\mu}_i \right) \mid \mathcal{F}_{t_i} \right] = 0, \quad \mathbb{E} [v_i^2 \mid \mathcal{F}_{t_i}] < \infty$$

Squaring both sides and taking conditional expectations, we have

$$\begin{aligned}\mathbb{E} \left[\left| \hat{Y}_{i+1}^{(l),\pi} \right|^2 \middle| \mathcal{F}_{t_i} \right] &= \mathbb{E} \left[\left| \hat{Y}_i^{(l),\pi} \right|^2 \middle| \mathcal{F}_{t_i} \right] + \frac{h_i}{c_i} \left| \hat{Z}_i^{(l),\pi} - \mathbb{E} \left[\hat{Y}_{i+1}^{(l),\pi} \middle| \mathcal{F}_{t_i} \right] \mathbb{E} \left[\hat{H}_i \right] \right|^2 \\ &\quad + \mathbb{E} \left[|v_i|^2 \middle| \mathcal{F}_{t_i} \right] - 2 \hat{Y}_i^{(l),\pi} \tilde{f}_i^l \left(t_i, \mathbb{E} \left[\hat{Y}_{i+1}^{(l),\pi} \middle| \mathcal{F}_{t_i} \right], \hat{Z}_i^{(l),\pi} \right) h_i \\ &\quad + \left(\tilde{f}_i^l \left(t_i, \mathbb{E} \left[\hat{Y}_{i+1}^{(l),\pi} \middle| \mathcal{F}_{t_i} \right], \hat{Z}_i^{(l),\pi} \right) h_i \right)^2,\end{aligned}$$

which leads to the bound on Z :

$$\begin{aligned}\frac{1}{2} \frac{h_i}{c_i} \left| \hat{Z}_i^{(l),\pi} \right|^2 &\leq \mathbb{E} \left[\left(\hat{Y}_{i+1}^{(l),\pi} \right)^2 \middle| \mathcal{F}_{t_i} \right] - \left(\hat{Y}_i^{(l),\pi} \right)^2 + 2 \hat{Y}_i^{(l),\pi} h_i \tilde{f}_i^l \left(t_i, \mathbb{E} \left[\hat{Y}_{i+1}^{(l),\pi} \middle| \mathcal{F}_{t_i} \right], \hat{Z}_i^{(l),\pi} \right) \\ &\quad + \frac{h_i}{c_i} \mathbb{E} \left[\hat{Y}_{i+1}^{(l),\pi} \middle| \mathcal{F}_{t_i} \right]^2 |\hat{\mu}_i|^2.\end{aligned}$$

By the quadratic growth assumption on the driver f , we see that

$$\begin{aligned}|f_i(t_i, y, z)| &\leq L_0 + \frac{1}{2} K_f + L_f |y| + \frac{3}{2} K_f |z|^2, \\ |\tilde{f}_i^l(t_i, y, z)| &\leq \frac{L_0}{K} + \frac{1}{2} K_f + (l-1) \sum_{m=1}^{l-1} \left| \hat{Z}_i^{(m),\pi} \right|^2 + L_f |y| + \frac{5}{2} K_f |z|^2.\end{aligned}$$

We write $\|\hat{Y}^{(l),\pi}\|_{\mathbb{S}^\infty(\mathbb{R})} := \text{ess sup}_{0 \leq i \leq N} |\hat{Y}_i^{(l),\pi}|$ and recall $\lambda \leq c_i \leq \Lambda$. The above equation motivates us to sum over i and to take conditional expectations to get

$$\begin{aligned}&\mathbb{E} \left[\sum_{j=i}^{N-1} \frac{1}{2\Lambda} \left| \hat{Z}_j^{(l),\pi} \right|^2 h_j \middle| \mathcal{F}_{t_i} \right] \\ &\leq \mathbb{E} \left[\left(\hat{Y}_N^{(l),\pi} \right)^2 - \left(\hat{Y}_i^{(l),\pi} \right)^2 \middle| \mathcal{F}_{t_i} \right] + \|\hat{Y}^{(l),\pi}\|_{\mathbb{S}^\infty(\mathbb{R})}^2 \sum_{j=i}^{N-1} \frac{h_j}{\lambda} |\hat{\mu}_j|^2 \\ &\quad + 2 \|\hat{Y}^{(l),\pi}\|_{\mathbb{S}^\infty(\mathbb{R})} \left(\sum_{j=i}^{N-1} \mathbb{E} \left[\left(\frac{L_0}{K} + \frac{1}{2} K_f + L_f \|\hat{Y}^{(l),\pi}\|_{\mathbb{S}^\infty(\mathbb{R})} \right) h_j \middle| \mathcal{F}_{t_i} \right] \right. \\ &\quad \left. + \sum_{j=i}^{N-1} \mathbb{E} \left[(l-1) \sum_{m=1}^{l-1} \left| \hat{Z}_j^{(m),\pi} \right|^2 h_j \middle| \mathcal{F}_{t_i} \right] + \sum_{j=i}^{N-1} \mathbb{E} \left[\frac{5}{2} K_f \left| \hat{Z}_j^{(l),\pi} \right|^2 h_j \middle| \mathcal{F}_{t_i} \right] \right).\end{aligned}$$

Rearranging, we have

$$\begin{aligned}&\left(\frac{1}{2\Lambda} - 5 K_f \|\hat{Y}^{(l),\pi}\|_{\mathbb{S}^\infty(\mathbb{R})} \right) \mathbb{E} \left[\sum_{j=i}^{N-1} \left| \hat{Z}_j^{(l),\pi} \right|^2 h_j \middle| \mathcal{F}_{t_i} \right] \\ &\leq \|\hat{Y}^{(l),\pi}\|_{\mathbb{S}^\infty(\mathbb{R})} \left(\left(\frac{2L_0}{K} + K_f \right) T + 2(l-1) \sum_{m=1}^{l-1} \mathbb{E} \left[\sum_{j=i}^{N-1} \left| \hat{Z}_j^{(m),\pi} \right|^2 h_j \middle| \mathcal{F}_{t_i} \right] \right) \\ &\quad + \|\hat{Y}^{(l),\pi}\|_{\mathbb{S}^\infty(\mathbb{R})}^2 \left(1 + 2L_f T + \frac{\hat{\mu}^2 T}{\lambda} \right),\end{aligned}$$

where $\hat{\mu}$ is in fact bounded by the definition of $\hat{H}_i$.

Step 2. In order to construct a bound on $\hat{Z}^{(l),\pi}$, we need to let

$$A(l) := \frac{1}{2\Lambda} - 5K_f \left\| \hat{Y}^{(l),\pi} \right\|_{\mathbb{S}^\infty(\mathbb{R})} > 0.$$

By definition of $\tilde{f}_i^l$, we observe $\tilde{f}_i^l(t, 0, 0) = \frac{f_i(t, 0, 0)}{K} \leq \frac{L_0}{K}$ and

$$\begin{aligned} \left| \tilde{f}_i^l(t, y, z) - \tilde{f}_i^l(t, y', z') \right| &= |f_i(t, y + \theta_1(l), z + \theta_2(l)) - f_i(t, y' + \theta_1(l), z' + \theta_2(l))| \\ &\leq L_f |y - y'| + L_z(n) |z - z'|, \\ \theta_1(l) &= \sum_{m=1}^{l-1} \mathbb{E} \left[\hat{Y}_{i+1}^{(m),\pi} \mid \mathcal{F}_{t_i} \right], \quad \theta_2(l) = \sum_{m=1}^{l-1} \hat{Z}_{i+1}^{(m),\pi}, \end{aligned}$$

for all $t \in [0, T]$, $y, y' \in \mathbb{R}$ and $z, z' \in \mathbb{R}^d$, such that $\{\tilde{f}_i^l\}_i$ satisfies [A6]. By Lemma 3.3.1 (iii),

$$\begin{aligned} \left\| \hat{Y}^{(l),\pi} \right\|_{\mathbb{S}^\infty(\mathbb{R})} &\leq \left(\|\xi^l\|_{L^\infty(\mathbb{R})} + \sup_i \{\|\tilde{f}_i^l\|_{\mathbb{S}^\infty(\mathbb{R})}\} T \right) \exp((1 + L_f)T) \\ &\leq \frac{1}{K} (\|\xi\|_{L^\infty(\mathbb{R})} + L_0 T) \exp((1 + L_f)T). \end{aligned}$$

Therefore, we set $K > 10\Lambda K_f (\|\xi\|_{L^\infty(\mathbb{R})} + L_0 T) \exp((1 + L_f)T)$ which is independent of N to get

$$\mathbb{E} \left[\sum_{j=i}^{N-1} \left| \hat{Z}_j^{(l),\pi} \right|^2 h_j \mid \mathcal{F}_{t_i} \right] \leq C_l \left(1 + \sum_{m=1}^{l-1} \mathbb{E} \left[\sum_{j=i}^{N-1} \left| \hat{Z}_j^{(m),\pi} \right|^2 h_j \mid \mathcal{F}_{t_i} \right] \right),$$

for some C_l being independent of N . An induction argument on $l = 1, \dots, K$ will show that for all $1 \leq l \leq K$,

$$\mathbb{E} \left[\sum_{j=i}^{N-1} \left| \hat{Z}_j^{(l),\pi} \right|^2 h_j \mid \mathcal{F}_{t_i} \right] \leq C_l$$

for some constant C_l being independent of N . What remains is to apply Lemma 2.3.16. $\square$

3.3.2 Perturbation of the centred scheme

We will discuss the impact of perturbation of the Y component of the centred scheme, followed by quantifying the impact of a change in Z . The perturbation analysis is vital for evaluating the accuracy of the numerical scheme. In particular, one version of the perturbed scheme is exactly the non-centred scheme.

Let (Y^π, Z^π) be the solution which corresponds to the triplet $(\xi, \{f_i\}_i, \{H_i\}_i)$ and suppose $(\xi, \{f_i\}_i, \{H_i\}_i)$ satisfies **[A5]**. The perturbed version on Y with ζ^Y of the scheme (3.22), $(\tilde{Y}, \tilde{Z})$, is defined to be

$$\begin{aligned}\tilde{Y}_i &= \mathbb{E} \left[\tilde{Y}_{i+1} + F_i \left(t_i, \mathbb{E} \left[\tilde{Y}_i \mid \mathcal{F}_{t_i} \right], \tilde{Z}_i \right) h_i \mid \mathcal{F}_{t_i} \right], \\ \tilde{Z}_i &= \mathbb{E} \left[\tilde{Y}_{i+1} H_i \mid \mathcal{F}_{t_i} \right], \\ \tilde{Y}_N &= \xi, \quad \tilde{Z}_N = 0, \\ F_i(t, y, z) &:= f_i(t, y, z) + \frac{1}{h_i} \zeta_i^Y, \quad \zeta_i^Y \in L^2(\mathcal{F}_{t_i})\end{aligned}\tag{3.27}$$

for all $1 \leq i \leq N$. Therefore, $(\tilde{Y}, \tilde{Z})$ corresponds to the triplet $(\xi, \{F_i\}_i, \{H_i\}_i)$, which also satisfies **[A5]**.

Given the perturbed scheme (3.27), the difference to the original scheme (3.22), for all $1 \leq i \leq N$, is referred as

$$\tilde{\Delta} Y_i^\pi := \tilde{Y}_i - Y_i^\pi, \quad \tilde{\Delta} Z_i^\pi = \tilde{Z}_i - Z_i^\pi.$$

Then we introduce $\tilde{\beta}_j^\pi$ and $\tilde{\eta}_i^\pi$, given by

$$\begin{aligned}\tilde{\beta}_j^\pi &:= \frac{f_j \left(t_j, \mathbb{E} [Y_{j+1}^\pi \mid \mathcal{F}_{t_j}], \tilde{Z}_j \right) - f_j \left(t_j, \mathbb{E} [Y_{j+1}^\pi \mid \mathcal{F}_{t_j}], Z_j^\pi \right)}{|Z_j^\pi - \tilde{Z}_j|^2} \\ &\quad (\tilde{Z}_j - Z_j^\pi) \mathbb{1}_{\{\tilde{\Delta} Z_j^\pi \neq 0\}}, \\ \tilde{\eta}_i^\pi &:= \prod_{j=i}^{N-1} (1 + h_j H_j \cdot \tilde{\beta}_j^\pi).\end{aligned}\tag{3.28}$$

By **[A5]** (ii), we have, for all $0 \leq i \leq N-1$, $\tilde{\eta}_i^\pi > 0$ and

$$\mathbb{E} [\tilde{\eta}_i^\pi \mid \mathcal{F}_{t_i}] = \mathbb{E} \left[\prod_{j=i}^{N-1} (1 + h_j H_j \cdot \tilde{\beta}_j^\pi) \mid \mathcal{F}_{t_i} \right] = 1.$$

It is valid to define a process $\{\tilde{\eta}^\pi(t)\}_{0 \leq t \leq T}$ and a probability measure $\mathbb{Q}^\pi$

$$\tilde{\eta}^\pi(t) := \prod_{\substack{0 \leq j \leq i-1, \\ \text{s.t. } t_{i+1} \leq t}} \left(1 + h_j \tilde{\beta}_j^\pi \cdot H_j \right), \quad \left. \frac{d\mathbb{Q}^\pi}{d\mathbb{P}} \right|_{\mathcal{F}_t} := \tilde{\eta}^\pi(t),\tag{3.29}$$

since $\mathbb{E} \left[\frac{d\mathbb{Q}^\pi}{d\mathbb{P}} \right] = 1$.

We aim to compare the scheme (3.27) and (3.22). We accommodate α_j and ψ_j in Lemma 3.3.1 to be defined as in this context, by treating $Y^{\pi,1}$ as $\tilde{Y}$. The key is to control ψ_j by finding I_j such that

$$I_j > \frac{\psi_j}{1 + b_j} = \frac{1 + a_j + b_j}{1 + b_j} = 1 + \frac{a_j}{1 + b_j},$$

where we write $a_j = \alpha_j h_j$ and $b_j = \tilde{\beta}_j^\pi \cdot H_j h_j$ for simplicity. If $a_j b_j \geq 0$, we can choose $I_j = 1 + a_j$ such that $I_j(1 + b_j) = \phi_j + a_j b_j$. If $a_j b_j < 0$ and $b_j > 0$, we set $I_j = 1$. Lastly, when $a_j > 0$ but $b_j < 0$, we pick $I_j = (1 + a_j)^\frac{1}{\epsilon}$, where ϵ appears in [A5], such that $I_j > 1 + \frac{1}{\epsilon} a_j \geq 1 + \frac{a_j}{1+b_j}$. We will write I_j^ϵ to emphasize the dependence on ϵ , which satisfies

$$0 < I_j^\epsilon < (1 + L_f h_j)^\frac{1}{\epsilon}.$$

Based on equation (3.25), we realise that the error is controlled by

$$\begin{aligned} |\tilde{\Delta} Y_i^\pi| &\leq \mathbb{E} [\Upsilon_i | \Delta Y_{t_N}^\pi | \mid \mathcal{F}_{t_i}] + \mathbb{E} \left[\sum_{k=i}^{N-1} |\zeta_k^Y| \frac{\Upsilon_i}{\Upsilon_k} \mid \mathcal{F}_{t_i} \right] \\ &\leq \mathbb{E} \left[\left(\prod_{j=i}^{N-1} I_j^\epsilon \right) \tilde{\eta}_i^\pi |\Delta Y_{t_N}^\pi| \mid \mathcal{F}_{t_i} \right] \\ &\quad + \mathbb{E} \left[\sum_{k=i}^{N-1} |\zeta_k^Y| \frac{\tilde{\eta}_i^\pi}{\tilde{\eta}_k^\pi} \prod_{j=i}^{k-1} I_j^\epsilon \mid \mathcal{F}_{t_i} \right]. \end{aligned}$$

As both ζ_k^Y and $\frac{\tilde{\eta}_i^\pi}{\tilde{\eta}_k^\pi}$ are $\mathcal{F}_{t_k}$ -measurable for all $k \geq i$, we apply Tower property to obtain

$$\mathbb{E} \left[|\zeta_k^Y| \frac{\tilde{\eta}_i^\pi}{\tilde{\eta}_k^\pi} \mid \mathcal{F}_{t_i} \right] = \mathbb{E} \left[\mathbb{E} \left[|\zeta_k^Y| \frac{\tilde{\eta}_i^\pi}{\tilde{\eta}_k^\pi} \mathbb{E} [\tilde{\eta}_k^\pi \mid \mathcal{F}_{t_k}] \mid \mathcal{F}_{t_k} \right] \mid \mathcal{F}_{t_i} \right] = \mathbb{E} [|\zeta_k^Y| \tilde{\eta}_i^\pi \mid \mathcal{F}_{t_i}],$$

which leads to the bound on the accumulated error

$$|\tilde{\Delta} Y_i^\pi| \leq \exp \left(\frac{L_f T}{\epsilon} \right) \mathbb{E} \left[|\eta_i^\pi| \left(|\tilde{\Delta} Y_{t_N}^\pi| + \sum_{k=i}^{N-1} |\zeta_k^Y| \right) \mid \mathcal{F}_{t_i} \right]. \quad (3.30)$$

Treating $\tilde{\eta}_i^\pi = \frac{\tilde{\eta}_i^\pi(T)}{\tilde{\eta}_i^\pi(t_i)}$ will enable us to replace $\tilde{\eta}_i^\pi$ by a $\mathbb{Q}^\pi$. The difference is now given under $\mathbb{Q}^\pi$ -measure

$$|\tilde{\Delta} Y_i^\pi| \leq C_\epsilon \mathbb{E}^{\mathbb{Q}^\pi} \left[\left(|\tilde{\Delta} Y_N^\pi| + \sum_{j=i}^{N-1} |\zeta_j^Y| \right) \mid \mathcal{F}_{t_i} \right]. \quad (3.31)$$

What remains is to control the measure $\mathbb{Q}^\pi$, which formally it is the “discrete version” of measure⁷ $\mathbb{Q}$ in the continuous case. The measure $\mathbb{Q}^\pi$ must be absolute continuous with respect to $\mathbb{P}$ for all such π , which motivates us verify the following claim.

⁷If we consider the difference of two solutions of two QBSDE, say $(Y^1, Z^1), (Y^2, Z^2)$, we will end up in a equation of the similar format where the Radon-Nikodym derivative of the measure $\mathbb{Q}$ depends on Z^1, Z^2 .

Proposition 3.3.3 (Analogue of Proposition 2.10 and 2.11 in [12]). *Suppose the triplet $(\xi, \{f_i\}_i, \{H_i\}_i)$ satisfies [A7]. Let (Y^π, Z^π) be the solution which corresponds to $(\xi, \{f_i\}_i, \{H_i\}_i)$.*

Consider the perturbed scheme $(\tilde{Y}, \tilde{Z})$ of (Y^π, Z^π) as in (3.27) and recall $\tilde{\beta}^\pi$ as in (3.28). We further assume $\|\tilde{Z}(\cdot)\|_{\mathbb{H}_{BMO}^2(\mathbb{R}^d)} < C_1$ where $\tilde{Z}(s) = \tilde{Z}_j$ for all $s \in [t_j, t_{j+1})$, $0 \leq j \leq N$ and C_1 is independent of N .

Then we have

$$M_t := \sum_{t_{i+1} \leq t} \tilde{\beta}_j^\pi \cdot H_j h_i \in BMO_{p,T}(\mathbb{P}, \mathbb{R}^d), \quad \|M\|_{BMO_{p,T}(\mathbb{P}, \mathbb{R}^d)} < C_2$$

for some constant C_2 being independent of N . Moreover, $\tilde{\eta}^\pi(t)$ satisfies the reverse Hölder inequality (Theorem 2.3.14) for some $q > 1$ that is independent of π .

Proof. As the proof is similar to the proof of [12, Proposition 2.10 and 2.11], we defer it to Section A.2. $\square$

Consequently, the error is of the form

$$|\tilde{\Delta} Y_i^\pi| \leq C_\epsilon \left(\mathbb{E} \left[|\tilde{\Delta} Y_N^\pi|^{q^*} \mid \mathcal{F}_{t_i} \right]^{\frac{1}{q^*}} + \mathbb{E} \left[\left(\sum_{j=i}^{N-1} |\zeta_j^Y| \right)^{q^*} \mid \mathcal{F}_{t_i} \right]^{\frac{1}{q^*}} \right), \quad (3.32)$$

where q^* is the parameter arising from the reverse Hölder inequality of $\eta^\pi(\cdot)$.

We are now in a position to specify the impact of a perturbation ζ^Y to Z^π . The $\mathbb{H}^2(\mathbb{R}^d)$ difference between the $\tilde{Z}$ and Z^π is bounded by a formula involving ΔY_i^π and ζ_i^Y .

Proposition 3.3.4 (Modification of Proposition 2.12 in [12]). *Under the same condition as in Proposition 3.3.3, we get*

$$\mathbb{E} \left[\sum_{i=0}^{N-1} |\tilde{\Delta} Z_i^\pi|^2 h_i \right] \leq C \left(\mathbb{E} \left[(\tilde{\Delta} Y_N^\pi)^2 \right] + \mathbb{E} \left[\sup_{0 \leq i \leq N-1} |\tilde{\Delta} Y_i^\pi|^4 \right]^{\frac{1}{2}} + \mathbb{E} \left[\sum_{i=0}^N \tilde{\Delta} Y_i^\pi(\zeta_i^Y) \right] \right),$$

where C depends on the BMO norm of N_t^π , which is defined as

$$N_t^\pi := \int_0^t \sum_{i=0}^{N-1} \tilde{\beta}_j^\pi \mathbb{1}_{[t_j, t_{j+1})}(s) \cdot dB_s.$$

Proof. See Section A.2. $\square$

Suppose the (Y^π, Z^π) is the solution that corresponds to the triplet $(\xi, \{f_i\}_i, \{H_i\}_i)$ and suppose $(\xi, \{f_i\}_i, \{H_i\}_i)$ satisfies [A5]. Consider perturbations $\{\zeta_i^Z\}_i$ which are $\mathcal{F}_{t_i}$ -measurable for each i and square integrable for all i . When the perturbation $\{\zeta_i^Z\}_i$ appears in the Z component, the numerical scheme boils down to

$$\begin{aligned}\check{Y}_i &= \mathbb{E} [\check{Y}_{i+1} \mid \mathcal{F}_{t_i}] + h_i f(t_i, \check{Y}_i, \check{Z}_i) \\ \check{Z}_i &= \mathbb{E} [\check{Y}_{i+1} H_i \mid \mathcal{F}_{t_i}] + \zeta_i^Z.\end{aligned}\tag{3.33}$$

We write

$$\check{Y}(s) = \sum_{i=0}^{N-1} \check{Y}_i \mathbb{1}_{[t_i, t_{i+1})}(s) + \check{Y}_N \mathbb{1}_{s=T}, \quad \check{Z}(s) = \sum_{i=0}^{N-1} \check{Z}_i \mathbb{1}_{[t_i, t_{i+1})}(s)$$

such that $\check{Y}, \check{Z}$ can be viewed as càdlàg processes. $(\check{Y}_i, \check{Z}_i)$ has a different format than

Theorem 3.3.5 (Error by perturbation on Z). *Suppose the triplet $(\xi, \{f_i\}_i, \{H_i\}_i)$ satisfies [A7]. Let (Y^π, Z^π) be the solution which corresponds to $(\xi, \{f_i\}_i, \{H_i\}_i)$.*

Consider the perturbed scheme $(\check{Y}, \check{Z})$ of (Y^π, Z^π) as in (3.33). We denote

$$\begin{aligned}\check{\Delta} Y_i^\pi &:= \check{Y}_i - Y_i^\pi, \quad \check{\Delta} Z_i^\pi = \check{Z}_i - Z_i^\pi, \\ \check{\alpha}_j^\pi &:= \frac{f_j(t_j, \mathbb{E}[\check{Y}_{j+1} \mid \mathcal{F}_{t_j}], \check{Z}_j) - f_j(t_j, \mathbb{E}[Y_{j+1}^\pi \mid \mathcal{F}_{t_j}], \check{Z}_j)}{\mathbb{E}[\check{Y}_j - Y_j^\pi \mid \mathcal{F}_{t_j}]} \mathbb{1}_{\{\check{Y}_j \neq Y_j^\pi\}}, \\ \check{\beta}_j^\pi &:= \frac{f_j(t_j, Y_j^\pi, \check{Z}_j) - f_j(t_j, Y_j^\pi, Z_j^\pi)}{|\check{Z}_j - Z_j^\pi|^2} (\check{Z}_j - Z_j^\pi) \mathbb{1}_{\{\check{Z}_j \neq Z_j^\pi\}},\end{aligned}$$

to define a process $\eta^\pi(t)$ and a (slightly abusing the notation) measure $\mathbb{Q}^\pi$

$$\check{\eta}^\pi(t) := \prod_{\substack{0 \leq j \leq i-1, \\ t_{i+1} \leq t}} (1 + h_j \check{\beta}_j^\pi \cdot H_j), \quad \left. \frac{d\mathbb{Q}^\pi}{d\mathbb{P}} \right|_{\mathcal{F}_t} := \check{\eta}^\pi(t).$$

Suppose $\check{Z}(\cdot) \in \mathbb{H}_{BMO}^2(\mathbb{R}^d)$ and $\|\check{Z}(\cdot)\|_{\mathbb{H}_{BMO}^2(\mathbb{R}^d)}$ are independent of N , we get

$$\mathbb{E} \left[\sum_{i=0}^{N-1} |\check{\Delta} Z_i^\pi|^2 h_i \right] \leq C \left(\mathbb{E} [|\check{\Delta} Y_N^\pi|^2] + \mathbb{E} \left[\sup_{0 \leq i \leq N-1} |\check{\Delta} Y_i^\pi|^4 \right]^{\frac{1}{2}} + \mathbb{E} \left[\sum_{i=0}^N (\zeta_i^Z)^2 h_i \right] \right)$$

where C depends on the BMO norm of N_t^π , which is defined as

$$N_t^\pi := \int_0^t \sum_{i=0}^{N-1} \check{\beta}_i^\pi \mathbb{1}_{[t_i, t_{i+1})}(s) \cdot dB_s.$$

Moreover, the impact of perturbation ζ_i^Z on Y is controlled by

$$|\check{\Delta} Y_i^\pi| \leq C_\epsilon \left(\mathbb{E}^{\mathbb{Q}^\pi} [|\check{\Delta} Y_N^\pi| \mid \mathcal{F}_{t_i}] + \mathbb{E}^{\mathbb{Q}^\pi} \left[\sum_{j=i}^{N-1} |\zeta_j^Z|^2 h_j \mid \mathcal{F}_{t_i} \right]^{\frac{1}{2}} \right).$$

Proof. By rewriting $\tilde{A}_i := \check{\Delta}Y_{i+1}^\pi - \mathbb{E}[\check{\Delta}Y_{i+1}^\pi \mid \mathcal{F}_{t_i}]$ as

$$\begin{aligned}\tilde{A}_i &= \mathbb{E} \left[\left(\check{\Delta}Y_{i+1}^\pi - \mathbb{E}[\check{\Delta}Y_{i+1}^\pi \mid \mathcal{F}_{t_i}] \right) H_i \mid \mathcal{F}_{t_i} \right] \cdot H_i \frac{h_i}{c_i} + \check{\Delta}v_i \\ &= \mathbb{E}[\check{\Delta}Y_{i+1}^\pi H_i \mid \mathcal{F}_{t_i}] \cdot H_i \frac{h_i}{c_i} + \check{\Delta}v_i = (\check{\Delta}Z_i^\pi - \zeta_i^Z) \cdot H_i \frac{h_i}{c_i} + \check{\Delta}v_i,\end{aligned}$$

where $\check{\Delta}v_i$ is some square-integrable $\mathcal{F}_{t_{i+1}}$ -measurable random variable such that $\mathbb{E}[\check{\Delta}v_i \mid \mathcal{F}_{t_i}] = \mathbb{E}[\check{\Delta}v_i H_i \mid \mathcal{F}_{t_i}] = 0 \in \mathbb{R}^d$, the difference of two schemes in Y is equal to

$$\begin{aligned}\check{\Delta}Y_i^\pi &= \check{\Delta}Y_{i+1}^\pi + A_i - \tilde{A}_i, \\ A_i &:= (f_i(t_i, \mathbb{E}[\check{Y}_{i+1} \mid \mathcal{F}_{t_i}], \check{Z}_i) - f_i(t_i, \mathbb{E}[Y_{i+1}^\pi \mid \mathcal{F}_{t_i}], Z_i^\pi)) h_i \\ &= h_i \check{\alpha}_i^\pi \mathbb{E}[\check{\Delta}Y_{i+1}^\pi \mid \mathcal{F}_{t_i}] + h_i \check{\beta}_i^\pi \cdot \check{\Delta}Z_i^\pi,\end{aligned}$$

Interpreting the error as $\check{\Delta}Y_{i+1}^\pi = \check{\Delta}Y_i^\pi - A_i + \tilde{A}_i$, we square both sides using the equality followed by taking conditional expectations, which results in

$$\begin{aligned}\mathbb{E}[(\check{\Delta}Y_{i+1}^\pi)^2 \mid \mathcal{F}_{t_i}] &= (\check{\Delta}Y_i^\pi)^2 - 2\check{\Delta}Y_i^\pi A_i + A_i^2 + \mathbb{E}[\tilde{A}_i^2 \mid \mathcal{F}_{t_i}] \\ &\geq (\check{\Delta}Y_i^\pi)^2 + 2\check{\Delta}Y_i^\pi (h_i \check{\alpha}_i^\pi \mathbb{E}[\check{\Delta}Y_{i+1}^\pi \mid \mathcal{F}_{t_i}] + h_i \check{\beta}_i^\pi \cdot \check{\Delta}Z_i^\pi) \\ &\quad + |\check{\Delta}Z_i| \frac{h_i}{c_i} + |\zeta_i^Z|^2 \frac{h_i}{c_i} - 2\check{\Delta}Z_i \cdot \zeta_i^Z \frac{h_i}{c_i}.\end{aligned}$$

Rearranging the equation and the inequality $2ab \leq ca^2 + c^{-1}b^2$ for any positive c , lead us to

$$\begin{aligned}\frac{1}{c_i} |\check{\Delta}Z_i^\pi|^2 h_i &\leq \mathbb{E} \left[|\check{\Delta}Y_{i+1}^\pi|^2 \mid \mathcal{F}_{t_i} \right] - |\check{\Delta}Y_i^\pi|^2 + 2L_f |\check{\Delta}Y_i^\pi| \mathbb{E} \left[|\check{\Delta}Y_{i+1}^\pi| \mid \mathcal{F}_{t_i} \right] h_i \\ &\quad + 3 \frac{h_i}{c_i} |\zeta_i^Z|^2 + \frac{1}{3c_i} |\check{\Delta}Z_i^\pi|^2 h_i + 3\Lambda h_i |\check{\beta}_i^\pi|^2 |\check{\Delta}Y_i^\pi|^2 + \frac{1}{3\Lambda} |\check{\Delta}Z_i^\pi|^2 h_i,\end{aligned}$$

followed by summing up all i and taking expectation to get

$$\begin{aligned}\mathbb{E} \left[\sum_{i=0}^{N-1} |\check{\Delta}Z_i^\pi|^2 h_i \right] &\leq C_\Lambda \left(\mathbb{E} \left[|\check{\Delta}Y_N^\pi|^2 - |\check{\Delta}Y_0^\pi|^2 \right] + 2L_f T \mathbb{E} \left[\sup_{0 \leq i \leq N-1} |\check{\Delta}Y_i^\pi|^2 \right] \right. \\ &\quad \left. + \mathbb{E} \left[\left(\sum_{i=0}^{N-1} h_i |\check{\beta}_i^\pi|^2 \right)^2 \right]^{\frac{1}{2}} \mathbb{E} \left[\sup_{0 \leq i \leq N} |\Delta Y_i^\pi|^4 \right]^{\frac{1}{2}} \right. \\ &\quad \left. + \mathbb{E} \left[\sum_{i=0}^{N-1} \frac{h_i}{\lambda} |\zeta_i^Z|^2 \right] \right).\end{aligned}$$

Due to the fact that $|\check{\beta}_j^\pi| \leq K_f(1 + |Z_j^\pi| + |\check{Z}_j|)$, N_t^π is a BMO martingale by Lemma 2.3.16 and Proposition 3.3.2. By the equivalence of BMO norms, $\mathbb{E}[(\sum_{i=0}^{N-1} h_i |\check{\beta}_i^\pi|^2)^2]^{\frac{1}{2}}$ is a constant independent of N .

We cannot directly apply Lemma 3.3.1 as the scheme (3.33) has a different format as in (3.23). Nevertheless, we can employ the same strategy to show that

$$\begin{aligned}\check{\Delta}Y_i^\pi &= \mathbb{E} [\check{\Delta}Y_{i+1}^\pi \mid \mathcal{F}_{t_i}] + \check{\alpha}_i \mathbb{E} [\check{\Delta}Y_{i+1}^\pi \mid \mathcal{F}_{t_i}] h_i + \check{\beta}_i \cdot (\mathbb{E} [\check{\Delta}Y_{i+1}^\pi H_i \mid \mathcal{F}_{t_i}] + \xi_i^Z) h_i \\ &= \mathbb{E} [\check{\Delta}Y_{i+1}^\pi \check{\psi}_i \mid \mathcal{F}_{t_i}] + \check{\beta}_i \cdot \zeta_i^Z h_i\end{aligned}$$

for all $0 \leq i \leq N-1$, where $\check{\psi}_i$ is defined as

$$\psi_i := 1 + h_i \check{\alpha}_i + h_i \check{\beta}_i \cdot H_i.$$

This result has a similar form as (3.26), so we can treat Δf_k in (3.26) as $\check{\beta}_i \xi_i^Z$. Since $\{f_i\}_i$ satisfies [A5], we can conclude similar results as in (3.28) to (3.31), except we replace ζ_i^Y by $\zeta_i^Z \cdot \check{\beta}_i h_i$ and the measure $\mathbb{Q}^\pi$ is defined slightly different⁸ (as its Radon-Nikodym derivative depends on $\check{\beta}$ rather than $\tilde{\beta}$). In fact, we will end up in the error of the form

$$\begin{aligned}|\check{\Delta}Y_i^\pi| &\leq C_\epsilon \mathbb{E}^{\mathbb{Q}^\pi} \left[\left(|\check{\Delta}Y_N^\pi| + \sum_{j=i}^{N-1} |\zeta_j^Z \check{\beta}_j| h_j \right) \mid \mathcal{F}_{t_i} \right] \\ &\leq C_\epsilon \left(\mathbb{E}^{\mathbb{Q}^\pi} \left[\sum_{j=i}^{N-1} |\zeta_j^Z|^2 h_j \mid \mathcal{F}_{t_i} \right]^{\frac{1}{2}} \mathbb{E}^{\mathbb{Q}^\pi} \left[\sum_{j=i}^{N-1} |\check{\beta}_j|^2 h_j \mid \mathcal{F}_{t_i} \right]^{\frac{1}{2}} \right. \\ &\quad \left. + \mathbb{E}^{\mathbb{Q}^\pi} [|\check{\Delta}Y_N^\pi| \mid \mathcal{F}_{t_i}] \right)\end{aligned}$$

which is exactly what we are looking for. $\square$

3.4 Convergence of modified BTZ scheme

We are now ready to specify the convergence of the numerical approximation. As a sketch of the proof, the solution of the QBSDE (Y, Z) will first be replaced by the truncated version (Y^n, Z^n) , where the error of truncation is controlled by Proposition 3.4.2. The perturbed scheme with ζ^Y , chosen in equation (3.35), will act as the bridge between the numerical solution $(\mathcal{Y}^\pi, \mathcal{Z}^\pi)$ and (Y^n, Z^n) . We fill the final gap by introducing a few processes which estimates the Z -component.

3.4.1 Modification on the generators

The first step is discretising the terminal data and truncating the driver. The proof of the following two results which are the generalisation of some conclusions in [90, 81], are deferred to Section A.2.

⁸We will not try to differentiate different $\mathbb{Q}^\pi$ - after all, it will be replaced by raising the power q^* .

Lemma 3.4.1 (L^p version of Corollary 4.4 in [90]). *Let X be the solution to the SDE which satisfies the [A2] and choose $\pi : 0 = t_0 < t_1 < \dots < t_N = T$ as a partition with $h = \sup_i |t_{i+1} - t_i|$. Then,*

$$\mathbb{E} [|\Phi(X) - \Phi(X^\pi)|^{2p}] \leq \begin{cases} C_p (h \ln(\frac{1}{h}))^p & \text{if } \Phi \text{ is Lipschitz w.r.t } \|\cdot\|_{\tilde{\mathbb{D}}^\infty(\mathbb{R}^m)}, \\ C_p h^p & \text{if } \Phi \text{ is Lipschitz w.r.t } \|\cdot\|_{\tilde{\mathbb{D}}^1(\mathbb{R}^m)}. \end{cases}$$

We then present a result on the difference of solutions to the BSDE having the same terminal condition, but one with a quadratic driver and the other one with a Lipschitz driver. The Lipschitz version of a quadratic driver f is constructed by projecting Z in a Ball centred at origin with radius n , as in equation (3.2). We let $(\xi, f^{(n)})$ be the generator associated with pair of solution (Y^n, Z^n) .

Proposition 3.4.2 (Generalisation of Remark 5.5 in [81]). *Let the pair (ξ, f) satisfy [A1] and let $(\xi, f^{(n)})$ be the truncated generator, which also satisfies [A1]. Suppose (Y, Z) be the solution to the BSDE with generator (ξ, f) and (Y^n, Z^n) represents the solution to the BSDE with generators $(\xi, f^{(n)})$.*

We further assume there exists a constant $p_1 > 1$, such that

$$\mathbb{E} \left[\left(\int_0^T |Z_s^n|^{2p_1} ds \right)^{p\ddot{q}} \right] \leq C_{p,p_1}$$

where $\ddot{q}, C_{p_1}$ are constants independent of n . Then,

$$\|Y - Y^n\|_{\mathbb{S}^p(\mathbb{R})}^p + \|Z - Z^n\|_{\mathbb{H}^p(\mathbb{R}^d)}^p \leq C_{p,p_1,k} n^{-k} \wedge C_{p,p_1} \exp(-C_p n^2),$$

for any $p \geq 1$ and any $k \geq 0$.

Proof. See Section A.2. □

As a consequence of the John-Nirenberg inequality, we can always establish a sharper convergence rate if we choose a larger k . Meanwhile, truncating the quadratic driver with a Lipschitz driver only works when there exist control of $\mathbb{E}[(\int_0^T |Z_t|^{2p_1} dt)^{p\ddot{q}}]$ with $p_1 > 1$. Based on our control on the Z component in Remark 3.2, the requirement in the above assumption is obviously satisfied.

3.4.2 Estimating different versions of the Z component

Before we implement the modified BTZ scheme on our problem, we aim to characterize the properties of Z in order to verify the assumptions in the perturbation results. Motivated by standard arguments in [12], we introduce processes $(\bar{Z}(t))_t, (\hat{Z}(t))_t$

and $(\tilde{Z}(t))_t$ to transform the solution to the original problem $(Z_t)_t$ to the numerical approximation $(Z_t^\pi)_t$. Formally, we discuss whether the BMO property and the L^p boundedness are preserved during this sequence of approximations.

Lemma 3.4.3 (Extension of Lemma 2.2 in [12]). *Let (Y, Z) be the solution to equation (2.11), where the generator (ξ, f) satisfies [A1]. Given a partition $\pi : 0 = t_0 < t_1 < \dots < t_N = T$, we define the following processes*

$$\begin{aligned}\bar{Z}(t) &:= \bar{Z}_j \quad \forall t \in [t_j, t_{j+1}), \quad \bar{Z}_j := \frac{1}{h_j} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} Z_s \, ds \mid \mathcal{F}_{t_j} \right], \quad \bar{Z}_T = 0, \\ \hat{Z}(t) &:= \hat{Z}_j \quad \forall t \in [t_j, t_{j+1}), \quad \hat{Z}_j := \frac{1}{h_j} \mathbb{E} [Y_{t_{j+1}} \Delta B_{t_j} \mid \mathcal{F}_{t_j}], \quad \hat{Z}_T = 0, \\ \tilde{Z}(t) &:= \tilde{Z}_j \quad \forall t \in [t_j, t_{j+1}), \quad \tilde{Z}_j = \mathbb{E} [Y_{t_{j+1}} H_j^R \mid \mathcal{F}_{t_j}], \quad \tilde{Z}_T = 0,\end{aligned}\tag{3.34}$$

respectively where H_j^R is defined in equation (3.4). They all belongs to the space $\mathbb{H}_{BMO}^2(\mathbb{R}^d)$ and their $\mathbb{H}_{BMO}^2(\mathbb{R}^d)$ norm depends on $\|Z\|_{\mathbb{H}_{BMO}^2(\mathbb{R}^d)}$, but is independent of π .

Proof. The $\|\bar{Z}\|_{\mathbb{H}_{BMO}^2(\mathbb{R}^d)}$ has been controlled by Lemma 3.2.11. Similarly, if we apply Hölder's inequality,

$$\begin{aligned}(\hat{Z}_j - \bar{Z}_j)^2 h_j &= \mathbb{E} \left[\left(Y_{t_j} - \int_{t_j}^{t_{j+1}} f(t, Y_s, Z_s) \, ds + \int_{t_j}^{t_{j+1}} Z_s \cdot dB_s \right) \frac{\Delta B_{t_j}}{h_j} \right. \\ &\quad \left. - \frac{1}{h_j} \int_{t_j}^{t_{j+1}} Z_s \, ds \mid \mathcal{F}_{t_j} \right]^2 h_j \\ &= \mathbb{E} \left[- \int_{t_j}^{t_{j+1}} f(t, Y_s, Z_s) \, ds \frac{\Delta B_{t_j}}{h_j} \mid \mathcal{F}_{t_j} \right]^2 h_j \\ &\leq \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |f(t, Y_s, Z_s)| \, ds \right)^2 \mid \mathcal{F}_{t_j} \right]\end{aligned}$$

Based on the assumption of the driver, we observe that for any $p \geq 1$

$$\begin{aligned}\mathbb{E} \left[\left(\int_{t_k}^T |f(t, Y_s, Z_s)| \, ds \right)^p \mid \mathcal{F}_{t_k} \right] &\leq C_p (T - t_k)^p + C_p \mathbb{E} \left[\left(\int_{t_k}^T |Z_s|^2 \, ds \right)^p \mid \mathcal{F}_{t_k} \right] \\ &\leq C_p \left((T - t_k)^p + \|Z_s\|_{\mathbb{H}_{BMO}^2(\mathbb{R}^d)}^{2p} \right),\end{aligned}$$

where we have used Theorem 2.3.8 together with Lemma 2.3.11. We are able to conclude that $\hat{Z} \in \mathbb{H}_{BMO}^2(\mathbb{R}^d)$ using equation (3.14), as

$$\mathbb{E} \left[\sum_{j=k}^{N-1} (\hat{Z}_j)^2 h_j \, ds \mid \mathcal{F}_{t_k} \right] = \mathbb{E} \left[\sum_{j=k}^{N-1} (\bar{Z}_j)^2 h_j \mid \mathcal{F}_{t_k} \right] + \mathbb{E} \left[\sum_{j=k}^{N-1} (\hat{Z}_j - \bar{Z}_j)^2 h_j \mid \mathcal{F}_{t_k} \right]$$

$$\begin{aligned}
&\leq C + \mathbb{E} \left[\sum_{j=k}^{N-1} \left(\int_{t_j}^{t_{j+1}} |f(t, Y_s, Z_s)| \, ds \right)^2 \middle| \mathcal{F}_{t_k} \right] \\
&\leq C + \mathbb{E} \left[\left(\int_{t_k}^T |f(t, Y_s, Z_s)| \, ds \right)^2 \middle| \mathcal{F}_{t_k} \right] \leq C + C(T - t_k)^2,
\end{aligned}$$

by using the inequality in (3.14) with $p = 2$.

What remains is to verify is the BMO property of $\tilde{Z}$. If we choose the perturbation term ζ_i^Y in (3.27) as

$$\zeta_i^Y = \mathbb{E} \left[\int_{t_i}^{t_{i+1}} \left(f(s, Y_s, Z_s) - f \left(t_i, \mathbb{E} [\tilde{Y}_{i+1} \mid \mathcal{F}_{t_i}], \tilde{Z}_i \right) \right) ds \middle| \mathcal{F}_{t_i} \right], \quad (3.35)$$

the perturbed version $(\tilde{Y}_j, \tilde{Z}_j)$ that is associated with the driver $f_i(t, y, z) = f(t, y, z)$ in (3.27) will coincide with $(Y_{t_j}, \tilde{Z}_j)$. By standard estimates on normal random variables,

$$\mathbb{E} \left[\left| H_j^R - \frac{\Delta B_{t_j}}{h_j} \right| \right] = \frac{\sqrt{2d}}{\sqrt{\pi}} \exp \left(\frac{-R^2}{2} \right) h_j^{-\frac{1}{2}} =: R_1(j), \quad (3.36)$$

$$\mathbb{E} \left[\left(H_j^R - \frac{\Delta B_{t_j}}{h_j} \right)^2 \right] \leq \frac{d\sqrt{2}}{\sqrt{\pi}} \left(R + \frac{1}{R} \right) \exp \left(\frac{-R^2}{2} \right) h_j^{-1} =: R_2(j), \quad (3.37)$$

where $(R_2(j))h_j < C$ for some $C > 0$ ⁹. Comparing $\hat{Z}_j$ and $\tilde{Z}_j$, we realise

$$\begin{aligned}
(\tilde{Z}_j - \hat{Z}_j)^2 h_j &= \mathbb{E} \left[Y_{t_{j+1}} \left(H_j^R - \frac{\Delta B_{t_j}}{h_j} \right) \middle| \mathcal{F}_{t_j} \right]^2 h_j \\
&= \mathbb{E} \left[(Y_{t_{j+1}} - Y_{t_j}) \left(H_j^R - \frac{\Delta B_{t_j}}{h_j} \right) \middle| \mathcal{F}_{t_j} \right]^2 h_j.
\end{aligned}$$

Applying Hölder's inequality, we arrive at

$$\begin{aligned}
\mathbb{E} \left[\sum_{j \geq k} (\tilde{Z}_j - \hat{Z}_j)^2 h_j \middle| \mathcal{F}_{t_k} \right] &\leq C \mathbb{E} \left[\sum_{j \geq k} \mathbb{E} [(Y_{t_{j+1}} - Y_{t_j})^2 \mid \mathcal{F}_{t_j}] \middle| \mathcal{F}_{t_k} \right] \\
&\leq C \mathbb{E} \left[\sum_{j \geq k} \left(\int_{t_j}^{t_{j+1}} f(t, Y_s, Z_s) \, ds + \int_{t_j}^{t_{j+1}} Z_s \cdot dB_s \right)^2 \middle| \mathcal{F}_{t_k} \right] \\
&\leq C \mathbb{E} \left[\left(\int_{t_k}^T |f(t, Y_s, Z_s)| \, ds \right)^2 + \int_{t_k}^T |Z_s|^2 \, ds \middle| \mathcal{F}_{t_k} \right] \leq C,
\end{aligned}$$

where we also utilized the conditional BDG inequality. $\square$

⁹In fact, if we choose $R = c \log(N)$ for some $c > 0$, we can show that $\|(\tilde{Z} - \hat{Z})(\cdot)\|_{\mathbb{H}_{\text{BMO}}^2(\mathbb{R}^d)} = o(N^{-\lambda})$ for any $\lambda > 0$.

$\bar{Z}(t)$ and $\hat{Z}(t)$ are not directly related to the numerical scheme, whereas $\tilde{Z}_i$ can be realised as a result of the perturbed scheme (3.27) for some ζ^Y . Nevertheless, they do not lost the BMO property. The definition of ξ is not used in the above Lemma, whereas the next result which based on the control of Z_s by the forward process X_s , relies on the form of $\xi = \Phi(X)$.

Lemma 3.4.4. *Let ξ satisfy [A3] and f satisfy [A4]. Let (Y, Z) be the solution to equation (3.1), with the generator (ξ, f) . Let $\pi : 0 = t_0 < t_1 < \dots < t_N = T$ be a partition. Given $\bar{Z}, \hat{Z}, \tilde{Z}$ which are defined in equation (3.34), we have*

$$\begin{aligned} |\bar{Z}_j| &\leq C_{L_\Phi} \left(1 + \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s| \mid \mathcal{F}_{t_j} \right] \right), \\ |\hat{Z}_j| &\leq C_{L_\Phi} \left(1 + \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s| \mid \mathcal{F}_{t_j} \right] + \sqrt{h_j} \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s|^4 \mid \mathcal{F}_{t_j} \right]^{\frac{1}{2}} \right), \\ |\tilde{Z}_j| &\leq C_{L_\Phi} \left(1 + \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s| \mid \mathcal{F}_{t_j} \right] + \sqrt{h_j} \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s|^4 \mid \mathcal{F}_{t_j} \right]^{\frac{1}{2}} \right), \end{aligned}$$

where we recall, L_Φ is the Lipschitz constant for Φ .

Proof. The L^p -type of bound on $\bar{Z}$ can be established by considering Lemma 3.2.8, Remark 3.2 as well as

$$|\bar{Z}_j| \leq \frac{1}{h_j} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} C_{L_\Phi} (1 + \sup_{0 \leq s \leq t} |X_s|) dt \mid \mathcal{F}_{t_j} \right] \leq C_{L_\Phi} \mathbb{E} \left[1 + \sup_{0 \leq s \leq t_{j+1}} |X_s| \mid \mathcal{F}_{t_j} \right],$$

whereas the difference between $\hat{Z}$ and $\bar{Z}$ is controlled by

$$\begin{aligned} |\hat{Z}_j - \bar{Z}_j| &\leq \mathbb{E} \left[\left| \int_{t_j}^{t_{j+1}} f(t, Y_s, Z_s) ds \frac{\Delta B_{t_j}}{h_j} \right| \mid \mathcal{F}_{t_j} \right] \\ &\leq \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |f(t, Y_s, Z_s)|^2 ds \mid \mathcal{F}_{t_j} \right]^{\frac{1}{2}} \\ &\leq \mathbb{E} \left[\int_{t_j}^{t_{j+1}} C + C|Z_s|^4 ds \mid \mathcal{F}_{t_j} \right]^{\frac{1}{2}} \\ &\leq C\sqrt{h_j} + C_{L_\Phi} \sqrt{h_j} \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s|^4 \mid \mathcal{F}_{t_j} \right]^{\frac{1}{2}}. \end{aligned} \quad (3.38)$$

Meanwhile, it is not difficult to show that

$$|\tilde{Z}_j - \hat{Z}_j| \leq \mathbb{E} \left[(Y_{t_{j+1}} - Y_{t_j})^2 \mid \mathcal{F}_{t_j} \right]^{\frac{1}{2}} \sqrt{R_2(j)}$$

$$\begin{aligned}
&\leq 2\mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |f(t, Y_s, Z_s)| \, ds \right)^2 + \left(\int_{t_j}^{t_{j+1}} Z_s \cdot dB_s \right)^2 \middle| \mathcal{F}_{t_j} \right]^{\frac{1}{2}} \sqrt{R_2(j)} \\
&\leq 2\mathbb{E} \left[h_j \int_{t_j}^{t_{j+1}} C + C|Z_s|^4 \, ds + \int_{t_j}^{t_{j+1}} |Z_s|^2 \, ds \middle| \mathcal{F}_{t_j} \right]^{\frac{1}{2}} \sqrt{R_2(j)} \\
&\leq 2\tilde{G}_j \left(R + \frac{1}{R} \right)^{\frac{1}{2}} \exp \left(\frac{-R^2}{4} \right),
\end{aligned}$$

where $\tilde{G}_j$ is defined as

$$\tilde{G}_j := 1 + \sqrt{h_j} \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s|^4 \middle| \mathcal{F}_{t_j} \right]^{\frac{1}{2}} + \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s|^2 \middle| \mathcal{F}_{t_j} \right]^{\frac{1}{2}}. \quad (3.39)$$

□

Remark 3.5. Given the truncation of the quadratic driver f , denoted by $f^{(n)}$, we recall the solution to this new BSDE is given as (Y^n, Z^n) . The truncation preserves the norm $\|\int_0^\cdot Z_t^n \, dB_t\|_{BMO_2(\mathbb{R}^d)} \leq C$ and $|Z_t^n| \leq C_{L_\Phi}(1 + \sup_{0 \leq s \leq t} |X_s|)$, where the constant C, C_{L_Φ} are independent of n in both inequalities. Similar to equation (3.34), we are allowed to define the sequence $\bar{Z}_i^n, \hat{Z}_i^n, \tilde{Z}_i^n$ and their corresponding processes accordingly, while all Lemmas within this section are still valid if we replace Z^n by Z .

Unlike Chassagneux et al. [12], we never used the relation between n and N to obtain the results in this section. This is mainly due to the process Z in this section is closely related to Z^n . However, when it comes to the numerical solution Z_i^π , the truncation is necessary to maintain the comparison theorem in Y and BMO property of Z . In addition, it is an essential part to characterise the propagation error under a new measure $\mathbb{Q}^\pi$.

3.4.3 Proof of Theorem 3.1.3

The claim in Theorem 3.1.3 will be divided into two parts, including Proposition 3.4.5 and 3.4.6. We will adapt the strategy of [12, Proposition 3.3] to estimate the error in both Y and Z .

First of all, the choice of $f^{(n)}, \{H_i^R\}_{1 \leq i \leq R}$ satisfy the assumption [A7], since $\hat{\mu}_i = 0$ for all i and

$$\begin{aligned}
0 &\leq hL_f + \sup_{0 \leq i \leq N-1} h_i \frac{R}{\sqrt{h_i}} (3 \max\{K_f, 1\}) h^{-\alpha} \\
&\leq hL_f + 3 \max\{K_f, 1\} h^{\frac{1}{2}-\alpha} \ln(h^{-1}) \rightarrow 0
\end{aligned}$$

when $h \rightarrow 0$. Recall equation (3.34), we observe that if

$$\zeta_i^Y = \mathbb{E} \left[\int_{t_i}^{t_{i+1}} f^{(n)}(s, Y_s^n, Z_s^n) - f^{(n)}(t_i, \mathbb{E}[\tilde{Y}_{i+1} \mid \mathcal{F}_{t_i}], \tilde{Z}_i) \, ds \mid \mathcal{F}_{t_i} \right], \quad (3.40)$$

the perturbed scheme (3.27) with $f_i(t, y, z) = f^{(n)}(t, y, z)$ and $H_i = H_i^R$ i.e. truncating the driver and the Brownian increment, by a induction argument, boils down to

$$\begin{aligned} \tilde{Y}_i &= \mathbb{E} \left[\tilde{Y}_{i+1} + \int_{t_i}^{t_{i+1}} f^{(n)}(s, Y_s^n, Z_s^n) \, ds \mid \mathcal{F}_{t_i} \right] = Y_{t_i}^n, \\ \tilde{Z}_i &= \mathbb{E} \left[\tilde{Y}_{i+1} H_i^R \mid \mathcal{F}_{t_i} \right] = \tilde{Z}_i^n, \\ \tilde{Y}_N &= \xi = \Phi(X), \quad \tilde{Z}_N = 0. \end{aligned}$$

We write

$$\tilde{\Delta} \mathcal{Y}_i^\pi := Y_{t_i}^n - \mathcal{Y}_i^\pi, \quad \tilde{\Delta} \mathcal{Y}^\pi(t) = \tilde{\Delta} \mathcal{Y}_i^\pi \quad \forall t \in [t_i, t_{i+1}), \quad \tilde{\Delta} \mathcal{Y}^\pi(T) = \tilde{\Delta} \mathcal{Y}_N^\pi, \quad (3.41)$$

to differentiate with $\tilde{\Delta} Y_i^\pi$, which refers to increment of the generic perturbed scheme in Y compared to scheme (3.22). Moreover, in this case, we specify the Radon-Nikodym derivative of $\mathbb{Q}^\pi$ (slightly abusing the notation) with respect to $\mathbb{P}$ as

$$\begin{aligned} \eta_i^{(n)} &:= \frac{d\mathbb{Q}^\pi}{d\mathbb{P}} \Big|_{\mathcal{F}_{t_i}} = \prod_{j=0}^{i-1} (1 + h_j H_j^R \cdot \beta_j^{(n)}), \\ \eta^{(n)}(t) &:= \prod_{\substack{0 \leq j \leq i-1, \\ t_{i+1} \leq t}} (1 + h_j \beta_j^{(n)} \cdot H_j^R), \quad \frac{d\mathbb{Q}^\pi}{d\mathbb{P}} \Big|_{\mathcal{F}_t} := \eta^{(n)}(t), \\ \beta_j^{(n)} &= \frac{f^{(n)}(t_j, Y_{t_j}^n, Z_j^\pi) - f^{(n)}(t_j, Y_{t_j}^n, \tilde{Z}_j)}{|Z_{t_j}^\pi - \tilde{Z}_j|^2} (Z_j^\pi - \tilde{Z}_j) \mathbb{1}_{\{Z_j^\pi - \tilde{Z}_j \neq 0\}}, \end{aligned} \quad (3.42)$$

which is a specification of (3.29).

Proposition 3.4.5. *Under the same condition as in Theorem 3.1.3, then there exist some $q^* > 1$ such that for all $p \geq 1$,*

$$\mathbb{E} \left[\sup_{0 \leq i \leq N} |Y_{t_i} - \mathcal{Y}_i^\pi|^{2p} \right] \leq C_{p, q^*} \mathbb{E} [|\Phi(X) - \Phi(X^\pi)|^{2pq^*}]^{\frac{1}{q^*}} + C_{\alpha, p, K, \epsilon} h^p.$$

Proof. Recall that we require $\alpha \in (0, \frac{1}{2})$ and $h \leq \frac{K}{N}$ in Theorem 3.1.3. We divided the proof into two steps.

Step 1. By truncation, we notice that the total error can be divided into

$$\mathbb{E} \left[\sup_{0 \leq i \leq N} |Y_{t_i} - \mathcal{Y}_i^\pi|^{2p} \right] \leq C_p \left(\mathbb{E} \left[\sup_{0 \leq i \leq N} |Y_{t_i} - Y_{t_i}^n|^{2p} \right] + \mathbb{E} \left[\sup_{0 \leq i \leq N} |Y_{t_i}^n - \mathcal{Y}_i^\pi|^{2p} \right] \right)$$

where the first component shrinks arbitrary fast in terms of n by Proposition 3.4.2. We fix $\alpha \in (0, \frac{1}{2})$ and conclude that $\mathbb{E} [\sup_{0 \leq i \leq n} |Y_{t_i} - Y_{t_i}^n|^{2p}] \leq C_{\alpha,p} h^p$.

We recall the definition for $\bar{Z}_i^n, \hat{Z}_i^n, \tilde{Z}_i^n$ in equation (3.34). If we focus on the second component in the above inequality, we notice that equation (3.35) can be written as

$$\zeta_j^Y = \mathbb{E} \left[\tilde{\zeta}_j^{Y,t} + \tilde{\zeta}_j^{Y,y} + \tilde{\zeta}_j^{Y,\bar{z}} + \tilde{\zeta}_j^{Y,y,E} + \tilde{\zeta}_j^{Y,\hat{z}} + \tilde{\zeta}_j^{Y,\tilde{z}} \mid \mathcal{F}_{t_j} \right],$$

where $\tilde{\zeta}_j^{Y,t}, \tilde{\zeta}_j^{Y,y}, \tilde{\zeta}_j^{Y,y,E}, \tilde{\zeta}_j^{Y,\bar{z}}, \tilde{\zeta}_j^{Y,\hat{z}}$ and $\tilde{\zeta}_j^{Y,\tilde{z}}$ are $\mathcal{F}_{t_{j+1}}$ -measurable and defined as

$$\begin{aligned} \tilde{\zeta}_j^{Y,t} &:= \int_{t_j}^{t_{j+1}} f^{(n)}(s, Y_s^n, Z_s^n) - f^{(n)}(t_j, Y_s^n, Z_s^n) \, ds, \\ \tilde{\zeta}_j^{Y,y} &:= \int_{t_j}^{t_{j+1}} f^{(n)}(t_j, Y_s^n, Z_s^n) - f^{(n)}(t_j, Y_{t_j}^n, Z_s^n) \, ds, \\ \tilde{\zeta}_j^{Y,y,E} &:= \int_{t_j}^{t_{j+1}} f^{(n)}(t_j, Y_{t_j}^n, Z_s^n) - f^{(n)}(t_j, \mathbb{E}[Y_{t_{j+1}}^n \mid \mathcal{F}_{t_j}], Z_s^n) \, ds, \\ \tilde{\zeta}_j^{Y,\bar{z}} &:= \int_{t_j}^{t_{j+1}} f^{(n)}(t_j, \mathbb{E}[Y_{t_{j+1}}^n \mid \mathcal{F}_{t_j}], Z_s^n) - f^{(n)}(t_j, \mathbb{E}[Y_{t_{j+1}}^n \mid \mathcal{F}_{t_j}], \bar{Z}_j^n) \, ds, \\ \tilde{\zeta}_j^{Y,\hat{z}} &:= \int_{t_j}^{t_{j+1}} f^{(n)}(t_j, \mathbb{E}[Y_{t_{j+1}}^n \mid \mathcal{F}_{t_j}], \bar{Z}_j^n) - f^{(n)}(t_j, \mathbb{E}[Y_{t_{j+1}}^n \mid \mathcal{F}_{t_j}], \hat{Z}_j^n) \, ds, \\ \tilde{\zeta}_j^{Y,\tilde{z}} &:= \int_{t_j}^{t_{j+1}} f^{(n)}(t_j, \mathbb{E}[Y_{t_{j+1}}^n \mid \mathcal{F}_{t_j}], \hat{Z}_j^n) - f^{(n)}(t_j, \mathbb{E}[Y_{t_{j+1}}^n \mid \mathcal{F}_{t_j}], \tilde{Z}_j^n) \, ds, \end{aligned}$$

Using Doob's L^p inequality together with inequality (3.31), we change the measure from $\mathbb{P}$ to $\mathbb{Q}^\pi$ at a cost of $\frac{1}{\epsilon}$ to derive

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq i \leq N} |Y_{t_i}^n - \mathcal{Y}_i^\pi|^{2p} \right] \\ & \leq C_{p,\epsilon} \mathbb{E} \left[\sup_{0 \leq i \leq N} \mathbb{E}^{\mathbb{Q}^\pi} \left[\left(\sum_{j=i}^{N-1} \mathbb{E}^{\mathbb{Q}^\pi} \left[\frac{|\tilde{\zeta}_j^{Y,t} + \tilde{\zeta}_j^{Y,y} + \tilde{\zeta}_j^{Y,y,E} + \tilde{\zeta}_j^{Y,\bar{z}} + \tilde{\zeta}_j^{Y,\hat{z}} + \tilde{\zeta}_j^{Y,\tilde{z}}|}{1 + h_j H_j^R \cdot \beta_j^n} \mid \mathcal{F}_{t_j} \right] \right) \right. \right. \\ & \quad \left. \left. + |\tilde{\Delta} \mathcal{Y}_N^\pi| \mid \mathcal{F}_{t_i} \right]^{2p} \right] \\ & \leq C_{p,\epsilon} \left(\mathbb{E} \left[|\tilde{\Delta} \mathcal{Y}_N^\pi|^{2pq^*} + \left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,y}| \right)^{2pq^*} + \left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,\bar{z}}| \right)^{2pq^*} + \left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,\tilde{z}}| \right)^{2pq^*} \right]^{\frac{1}{q^*}} \right. \\ & \quad \left. + \mathbb{E} \left[\sup_{0 \leq i \leq N} \mathbb{E}^{\mathbb{Q}^\pi} \left[\left(\sum_{j=i}^{N-1} |\tilde{\zeta}_j^{Y,t}| \right) + \left(\sum_{j=i}^{N-1} |\tilde{\zeta}_j^{Y,y,E}| \right) \mid \mathcal{F}_{t_i} \right]^{2p} \right] \right. \\ & \quad \left. + \mathbb{E} \left[\sup_{0 \leq i \leq N} \mathbb{E}^{\mathbb{Q}^\pi} \left[\left(\sum_{j=i}^{N-1} |\tilde{\zeta}_j^{Y,\bar{z}}| \right) \mid \mathcal{F}_{t_i} \right]^{2p} \right] \right), \end{aligned}$$

because $1 + H_j^R \cdot \beta_j^{(n)} h_j \in (\epsilon, 2 - \epsilon)$. Thanks to Lemma 3.4.1, we control $\tilde{\Delta} \mathcal{Y}_N = \Phi(X) - \Phi(\hat{X}^\pi)$ and arrive at the first term on the right hand side of the inequality.

Step 2. We aim to analyse all the terms involving the exponent $2pq^*$. By the assumption of the driver, we deduce $|\tilde{\zeta}_j^{Y,y}| \leq \int_{t_j}^{t_{j+1}} L_f |Y_s^n - Y_{t_i}^n| ds$ as well as

$$\begin{aligned} \mathbb{E} \left[\left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,y}| \right)^{2pq^*} \right]^{\frac{1}{q^*}} &\leq \left(C_p \mathbb{E} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} |Y_s^n - Y_{t_j}^n| ds \right)^{2pq^*} \right] \right)^{\frac{1}{q^*}} \\ &\leq \left(C_{p,q} \mathbb{E} \left[\left(\sum_{j=0}^{N-1} \sup_{s \in [t_j, t_{j+1}]} |Y_s^n - Y_{t_i}^n| h_j \right)^{2pq^*} \right] \right)^{\frac{1}{q^*}} \\ &\leq \left(C_{p,q} \mathbb{E} \left[\left(\sum_{j=0}^{N-1} \sup_{s \in [t_j, t_{j+1}]} |Y_s^n - Y_{t_i}^n|^{2pq^*} h_j \right) T^{2pq^*-1} \right] \right)^{\frac{1}{q^*}} \\ &\leq \left(C_{p,q} \left(\sum_{j=0}^{N-1} h_j \right) \sup_{0 \leq j \leq N-1} \mathbb{E} \left[\sup_{s \in [t_j, t_{j+1}]} |Y_s^n - Y_{t_i}^n|^{2pq^*} \right] \right)^{\frac{1}{q^*}} \\ &= C_{p,q,L_\Phi} h^{2p}, \end{aligned}$$

where we have used Doob's L^p inequality to remove the conditional expectation, have applied Hölder's inequality to split the sum of products and have utilised Theorem 3.1.2 to obtain the final line. Furthermore, the assumption on the driver implies

$$\begin{aligned} &\mathbb{E} \left[\left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,\bar{z}}| \right)^{2pq^*} \right]^{\frac{1}{q^*}} \\ &\leq C_{p,q^*} \mathbb{E} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} K_f (1 + |\bar{Z}^n(s)| + |\hat{Z}^n(s)|) |\bar{Z}^n(s) - \hat{Z}^n(s)| ds \right)^{2pq^*} \right]^{\frac{1}{q^*}} \\ &\leq C_{p,q^*} \|K_f(1 + |\hat{Z}^n(s)| + |\bar{Z}^n(s)|)\|_{\mathbb{H}_{\text{BMO}}^{4pq^*}(\mathbb{R})}^{2p} \mathbb{E} \left[\left(\int_0^T |\bar{Z}^n(s) - \hat{Z}^n(s)|^2 ds \right)^{2pq^*} \right]^{\frac{1}{2q^*}}, \end{aligned}$$

where Hölder's inequality is applied in the last line. Meanwhile,

$$\begin{aligned} &\mathbb{E} \left[\left(\int_0^T |\bar{Z}^n(s) - \hat{Z}^n(s)|^2 ds \right)^{2pq^*} \right]^{\frac{1}{2q^*}} \\ &\leq \mathbb{E} \left[\left(\sum_{j=0}^{N-1} h_j C_{L_\Phi} \left(1 + \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s|^4 \mid \mathcal{F}_{t_j} \right] \right) h_j \right)^{2pq^*} \right]^{\frac{1}{2q^*}} \end{aligned}$$

$$\begin{aligned}
&\leq C_{p,L_\Phi} h^p \mathbb{E} \left[\sum_{j=0}^{N-1} \left(1 + \mathbb{E} \left[\sup_{0 \leq s \leq t_{j+1}} |X_s|^4 \mid \mathcal{F}_{t_j} \right] \right)^{2pq^*} h_j \right]^{\frac{1}{2q^*}} \left(\sum_{j=0}^{N-1} h_j \right)^{\frac{2pq^*-1}{2q^*}} \\
&\leq C_{p,L_\Phi} \mathbb{E} \left[\sup_{0 \leq s \leq T} |X_s|^{8pq^*} \right]^{\frac{1}{2q^*}} h^p,
\end{aligned} \tag{3.43}$$

where we have applied equation (3.38) together with Hölder's inequality within the calculation. Recall the definition for $R_1(j)$ and $R_2(j)$ defined in equation (3.36) and (3.37), we proceed our verification using a similar procedure to the above inequality. Due to the fact that $\|K_f(1 + |\hat{Z}^n(s)| + |\bar{Z}^n(s)|)\|_{\mathbb{H}_{\text{BMO}}^{4pq^*}(\mathbb{R})} < C$ for some C independent of π, n , we obtain

$$\begin{aligned}
&\mathbb{E} \left[\left(\int_0^T |\hat{Z}^n(s) - \tilde{Z}_s^n|^2 ds \right)^{2pq^*} \right]^{\frac{1}{2q^*}} \\
&\leq C_p \mathbb{E} \left[\sum_{j=0}^{N-1} \left(\tilde{G}_j \left(R + \frac{1}{R} \right)^{\frac{1}{2}} \exp \left(\frac{-R^2}{4} \right) \right)^{2pq^*} h_j \right]^{\frac{1}{2q^*}} \left(\sum_{j=0}^{N-1} h_j \right)^{\frac{2pq^*-1}{2q^*}} \\
&\leq C_p \mathbb{E} \left[\sup_{0 \leq s \leq T} |X_s|^{8pq^*} \right]^{\frac{1}{2q^*}} \left(\left(R + \frac{1}{R} \right) \exp \left(\frac{-R^2}{2} \right) \right)^p
\end{aligned} \tag{3.44}$$

where $\tilde{G}_j$ is defined as in equation (3.39). Provided $R = \ln(h^{-1})$, $(R + \frac{1}{R}) \exp \left(\frac{-R^2}{2} \right) \leq 2h^{m(h)}$ where $m(h) = \frac{1}{2} \ln(h^{-1}) + \frac{\ln(\ln(h^{-1}))}{\ln(h)} \rightarrow \infty$ as $h \rightarrow 0$. The contribution of error for replacing $\hat{Z}$ by $\tilde{Z}$ is of order $o(h^{-v})$ for all $v \geq 0$.

Step 3. Now we deal with the term involving $\mathbb{Q}^\pi$. Using the assumption on the driver f , $|\zeta_j^{Y,t}| \leq \frac{2}{3} L_f(h_j)^{\frac{3}{2}}$, hence the contribution of $\zeta_j^{Y,t}$ in the total error is controlled by

$$\mathbb{E} \left[\sup_{0 \leq i \leq N} \mathbb{E}^{\mathbb{Q}^\pi} \left[\left(\sum_{j=i}^{N-1} |\tilde{\zeta}_j^{Y,t}| \right) \mid \mathcal{F}_{t_i} \right]^{2p} \right] \leq \left(\sum_{i=0}^{N-1} \left(\frac{2}{3} L_f(h_i)^{\frac{3}{2}} \right) \right)^{2p} \leq C_{p,T} h^p.$$

Utilising the Lipschitz property of f , we have

$$\begin{aligned}
A_1(i) &:= \mathbb{E}^{\mathbb{Q}^\pi} \left[\sum_{j=i}^{N-1} \left| \tilde{\zeta}_j^{Y,y,E} \right| \mid \mathcal{F}_{t_i} \right] \\
&\leq \mathbb{E}^{\mathbb{Q}^\pi} \left[\sum_{j=i}^{N-1} L_f h_j \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |f(s, Y_s^n, Z_s^n)| ds \mid \mathcal{F}_{t_j} \right] \mid \mathcal{F}_{t_i} \right] \\
&\leq C_\epsilon \mathbb{E}^{\mathbb{Q}^\pi} \left[\int_{t_j}^T |f(s, Y_s^n, Z_s^n)| ds \mid \mathcal{F}_{t_i} \right] h
\end{aligned}$$

$$\leq C_\epsilon \mathbb{E}^{\mathbb{Q}^\pi} \left[C \left(1 + \sup_{s \in [t_j, T]} |Y_s^n| + \int_{t_j}^T |Z_s^n|^2 ds \right) \middle| \mathcal{F}_{t_i} \right] h \leq C_{T, \epsilon} h,$$

by the Corollary 2.3.15. Therefore,

$$\mathbb{E} \left[\sup_{0 \leq i \leq N} |A_1(i)|^{2p} \right] \leq C_{T, p, \epsilon} h^{2p}.$$

By observing $(1 + h_j H_j \cdot \beta_j^{(n)}) \in (\epsilon, 2 - \epsilon)$, we analyse the last component

$$\begin{aligned} A_2(i) &:= \mathbb{E}^{\mathbb{Q}^\pi} \left[\left(\sum_{j=i}^{N-1} |\tilde{\zeta}_j^{Y, \bar{z}}| \right) \middle| \mathcal{F}_{t_i} \right]^2 \\ &\leq \mathbb{E}^{\mathbb{Q}^\pi} \left[\int_{t_i}^T K_f (1 + |Z_s^n| + |\bar{Z}^n(s)|) |Z_s^n - \bar{Z}^n(s)| ds \middle| \mathcal{F}_{t_i} \right]^2 \\ &\leq \mathbb{E}^{\mathbb{Q}^\pi} \left[\int_{t_i}^T K_f^2 (1 + |Z_s^n| + |\bar{Z}^n(s)|)^2 ds \middle| \mathcal{F}_{t_i} \right] \mathbb{E}^{\mathbb{Q}^\pi} \left[\int_{t_i}^T |Z_s^n - \bar{Z}^n(s)|^2 ds \middle| \mathcal{F}_{t_i} \right] \\ &\leq \|K_f(1 + |Z^n| + |\bar{Z}^n(\cdot)|)\|_{\mathbb{H}_{\text{BMO}}^{2q^*}(\mathbb{P}, \mathbb{R})}^{2q^*} \mathbb{E}^{\mathbb{Q}^\pi} \left[\sum_{j=i}^N \int_{t_j}^{t_{j+1}} |Z_s^n - \bar{Z}^n(s)|^2 ds \middle| \mathcal{F}_{t_i} \right], \end{aligned}$$

where the last line is deduced by Corollary 2.3.15. Lastly, the following inequality is a consequence of Proposition 3.2.12:

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq i \leq N} |A_2(i)|^p \right] &\leq C(p, q) \mathbb{E} \left[\sup_{0 \leq i \leq N} \mathbb{E}^{\mathbb{Q}^\pi} \left[\int_{t_j}^T |Z_s^n - \bar{Z}^n(s)|^2 ds \middle| \mathcal{F}_{t_i} \right]^p \right] \\ &\leq C(p, q, L_\Phi, \epsilon) h^p. \end{aligned}$$

Summarising all error will lead us to our claim. $\square$

Proposition 3.4.6. *Under the same condition as Theorem 3.1.3, there exists a constant $q^* > 1$ such that*

$$\mathbb{E} \left[\int_0^T |Z_s - Z^\pi(s)|^2 ds \right] \leq C_{\alpha, K, L_\Phi, q^*, \epsilon} \left(\mathbb{E} [|\Phi(X) - \Phi(X^\pi)|^{8q^*}]^{\frac{1}{8q^*}} h^{\frac{1}{2}} + h \right)$$

Proof. Considering the elementary inequality $|Z_s - Z^\pi|^2 \leq 5(|Z_s - Z_s^n|^2 + |Z_s^n - \bar{Z}_s^n|^2 + |\bar{Z}_s^n(s) - \hat{Z}_s^n|^2 + |\hat{Z}_s^n(s) - \tilde{Z}_s^n|^2 + |\tilde{Z}_s^n - Z_s^\pi|^2)$, we first bound

$$\mathbb{E} \left[\int_0^T |Z_s - Z_s^n|^2 ds \right] \leq C \left(p, \frac{1}{\alpha} \right) n^{-\frac{1}{\alpha}} = C(p, \alpha) N^{-1} \leq C(p, \alpha) \frac{h}{T},$$

by Proposition 3.4.2, the choice of n . The last inequality is the consequence of $h = \sup_i h_i$, as $h < \frac{T}{N}$ leads to the contradiction $\sum_i h_i < T$. Considering the inequality

$$|Z_s^n - \bar{Z}^n(s)|^2 \leq 2|Z_s^n - \bar{Z}_{t_j}^n|^2 + 2 \left| Z_{t_j}^n - \frac{1}{h_j} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} Z_u^n du \middle| \mathcal{F}_{t_j} \right] \right|^2$$

$$\leq 2|Z_s^n - \bar{Z}_{t_j}^n|^2 + \frac{2}{h_j} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |Z_u^n - Z_{t_j}^n|^2 du \mid \mathcal{F}_{t_j} \right],$$

we implement Theorem 3.1.2 and Proposition 3.2.12 to deduce

$$\mathbb{E} \left[\int_0^T |Z_s^n - \bar{Z}^n(s)|^2 ds \right] \leq 4 \mathbb{E} \left[\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} |Z_s - Z_{t_j}|^2 ds \right] \leq C_{L_\Phi} h.$$

As we have proved inequality (3.43) and (3.44), we easily seen that

$$\mathbb{E} \left[\int_0^T |\bar{Z}^n(s) - \hat{Z}^n(s)|^2 ds \right] + \mathbb{E} \left[\int_0^T |\hat{Z}^n(s) - \tilde{Z}_s^n|^2 ds \right] \leq C_{L_\Phi} h.$$

Recall that there exists a uniform bound, which is independent of N but depends on ϵ , for $\tilde{Y}_i^n := Y_{t_i}^n$ and $\mathcal{Y}_i^\pi$. By Proposition 3.3.4, we focus on the term

$$\begin{aligned} \mathbb{E} \left[\sum_{j=0}^{N-1} \Delta \mathcal{Y}_j^\pi \zeta_j^Y \right] &= \mathbb{E} \left[\sum_{j=0}^{N-1} \mathbb{E} \left[\Delta \mathcal{Y}_j^\pi \left(\tilde{\zeta}_j^{Y,t} + \tilde{\zeta}_j^{Y,y} + \tilde{\zeta}_j^{Y,y,E} + \tilde{\zeta}_j^{Y,\bar{z}} + \tilde{\zeta}_j^{Y,\hat{z}} + \tilde{\zeta}_j^{Y,\tilde{z}} \right) \mid \mathcal{F}_{t_j} \right] \right] \\ &= \left(\mathbb{E} \left[\left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,t}| \right)^2 \right]^{\frac{1}{2}} + \mathbb{E} \left[\left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,y}| \right)^2 \right]^{\frac{1}{2}} + \mathbb{E} \left[\left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,y,E}| \right)^2 \right]^{\frac{1}{2}} \right. \\ &\quad \left. + \mathbb{E} \left[\left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,\hat{z}}| \right)^2 \right]^{\frac{1}{2}} + \mathbb{E} \left[\left(\sum_{j=0}^{N-1} |\tilde{\zeta}_j^{Y,\tilde{z}}| \right)^2 \right]^{\frac{1}{2}} \right) \mathbb{E} \left[\sup_{0 \leq j \leq N-1} |\Delta \mathcal{Y}_j^\pi|^2 \right]^{\frac{1}{2}} \\ &\quad + \mathbb{E} \left[\sum_{j=0}^{N-1} \Delta \mathcal{Y}_j^\pi \tilde{\zeta}_j^{Y,\bar{z}} \right]. \end{aligned}$$

We have seen the estimate of the term $\mathbb{E} \left[\sup_{0 \leq i \leq N-1} |\Delta \mathcal{Y}_i^\pi|^2 \right]^{\frac{1}{2}}$. By step 2 within the proof of Proposition 3.4.5, we only need to estimate the last element. Writing $V := \mathbb{E} \left[\sum_{j=0}^{N-1} |\Delta \mathcal{Y}_j^\pi \tilde{\zeta}_j^{Y,\bar{z}}| \right]$, we apply the assumption on the driver f to derive

$$\begin{aligned} V &\leq \mathbb{E} \left[\sum_{j=0}^{N-1} |\Delta \mathcal{Y}_j^\pi| \int_{t_j}^{t_{j+1}} K_f(1 + |Z_s^n| + |\bar{Z}^n(s)|) |Z_s^n - \bar{Z}^n(s)| ds \right] \\ &\leq \mathbb{E} \left[\int_0^T K_f(1 + |Z_s^n| + |\bar{Z}^n(s)|) |\Delta \mathcal{Y}^\pi(s)| |Z_s^n - \bar{Z}^n(s)| ds \right] \\ &\leq \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E} \left[\int_{t_i}^T K_f(1 + |Z_s^n| + |\bar{Z}^n(s)|) |\Delta \mathcal{Y}^\pi(s)| |Z_s^n - \bar{Z}^n(s)| ds \mid \mathcal{F}_{t_i} \right] \right], \end{aligned}$$

Meanwhile, we make use of conditional Hölder's inequality to deduce

$$\mathbb{E} \left[\int_{t_i}^T K_f(1 + |Z_s^n| + |\bar{Z}^n(s)|) |\Delta \mathcal{Y}^\pi(s)| |Z_s^n - \bar{Z}^n(s)| ds \mid \mathcal{F}_{t_i} \right]$$

$$\leq A(t_i) \mathbb{E} \left[\sup_{i \leq j \leq N-1} |\Delta \mathcal{Y}_j^\pi|^4 \mid \mathcal{F}_{t_i} \right]^{\frac{1}{4}} \mathbb{E} \left[\int_{t_i}^T |Z_s^n - \bar{Z}^n(s)|^2 ds \mid \mathcal{F}_{t_i} \right]^{\frac{1}{2}},$$

where $A(t_i)$ denotes

$$\begin{aligned} A(t_i) &:= \mathbb{E} \left[\left(\int_{t_i}^T K_f^2 (1 + |Z_s^n| + |\bar{Z}^n(s)|)^2 ds \right)^2 \mid \mathcal{F}_{t_i} \right]^{\frac{1}{4}} \\ &\leq K_f \|1 + |Z^n| + |\bar{Z}^n|\|_{\mathbb{H}_{\text{BMO}}^4(\mathbb{R})} \end{aligned}$$

for all i . When we apply Proposition 3.2.12 and Doob's martingale inequality, V can be controlled by

$$\begin{aligned} V &\leq C \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \left(\mathbb{E} \left[\sup_{i \leq j \leq N-1} |\Delta \mathcal{Y}_j^\pi|^4 \mid \mathcal{F}_{t_i} \right]^{\frac{1}{4}} \mathbb{E} \left[\int_{t_i}^T |Z_s^n - \bar{Z}^n(s)|^2 ds \mid \mathcal{F}_{t_i} \right]^{\frac{1}{2}} \right) \right] \\ &\leq C \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E} \left[\sup_{0 \leq j \leq N-1} |\Delta \mathcal{Y}_j^\pi|^4 \mid \mathcal{F}_{t_i} \right]^{\frac{1}{2}} \right]^{\frac{1}{2}} \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E} \left[\int_{t_i}^T |Z_s^n - \bar{Z}^n(s)|^2 ds \mid \mathcal{F}_{t_i} \right] \right]^{\frac{1}{2}} \\ &\leq C \mathbb{E} \left[\sup_{0 \leq j \leq N-1} |\Delta \mathcal{Y}_j^\pi|^8 \right]^{\frac{1}{8}} h^{\frac{1}{2}}. \end{aligned}$$

We reconsider Proposition 3.3.4 to deduce

$$\begin{aligned} &\mathbb{E} \left[\int_0^T (\tilde{Z}^n(t) - Z^\pi(t))^2 dt \right] \\ &\leq C \left(\mathbb{E} \left[\sup_{0 \leq j \leq N} |\Delta \mathcal{Y}_j^\pi|^4 \right]^{\frac{1}{2}} + \mathbb{E} \left[\sup_{0 \leq j \leq N} |\Delta \mathcal{Y}_j^\pi|^2 \right]^{\frac{1}{2}} h^{\frac{1}{2}} + \mathbb{E} \left[\sup_{0 \leq j \leq N} |\Delta \mathcal{Y}_j^\pi|^8 \right]^{\frac{1}{8}} h^{\frac{1}{2}} \right), \end{aligned}$$

as we required. $\square$

Proof of Theorem 3.1.3. This theorem is a direct consequence of Proposition 3.4.5 and 3.4.6, if we observe

$$2\mathbb{E} [|\Phi(X) - \Phi(X^\pi)|^{8q^*}]^{\frac{1}{8q^*}} a \leq \mathbb{E} [|\Phi(X) - \Phi(X^\pi)|^{8q^*}]^{\frac{1}{4q^*}} + a^2.$$

$\square$

3.5 Implementation of the numerical scheme

As we have seen, the modified BTZ scheme requires a further approximation to estimate the conditional expectation at each step. [12] considered numerical integration by taking advantage of the Markovian property of the problem. We aim to extend their work to our generalised setting, but an extra assumption is required.

Definition 3.5.1. A functional $\Phi : \tilde{\mathbb{D}}(\mathbb{R}^m) \mapsto \mathbb{R}$ which is Lipschitz with respect to $\|\cdot\|_{\tilde{\mathbb{D}}^\infty(\mathbb{R}^m)}$ or $\|\cdot\|_{\tilde{\mathbb{D}}^1(\mathbb{R}^m)}$ is called *constructible* (with function $g, \tilde{g}$) if there exists a family of functionals $(\Phi_t)_{0 \leq t \leq T}$ with $\Phi_t : \tilde{\mathbb{D}}(\mathbb{R}^m) \mapsto \mathbb{R}^{d_1}$ for some $d_1 \in \mathbb{N}$ and a function $g : [0, T] \times [0, T] \times \mathbb{R}^{d_1} \times \tilde{\mathbb{D}}_{[s,t]}(\mathbb{R}^m) \mapsto \mathbb{R}$ such that

$$(i) \quad \Phi_t(\Theta) = \Phi_t(\Theta \mathbb{1}_{[0,t]}),$$

$$(ii) \quad \Phi_t(\Theta) = g(s, t, \Phi_s(\Theta), \Theta \mathbb{1}_{[s,t]}),$$

(iii) $x \mapsto g(s, t, x, y)$ is uniformly Lipschitz and $y \mapsto g(s, t, x, y)$ is Lipschitz with respect to $\|\cdot\|_{\tilde{\mathbb{D}}^{m,\infty}}$ or $\|\cdot\|_{\tilde{\mathbb{D}}^{m,1}}$, with the Lipschitz constant being independent of s, t ,

$$(iv) \quad \Phi(\Theta) = \tilde{g}(\Phi_T(\Theta), \Theta(T)) \text{ and } \tilde{g} \text{ is Lipschitz in } \mathbb{R}^{d_1} \times \mathbb{R}^m,$$

for any $\Theta \in \tilde{\mathbb{D}}(\mathbb{R}^m)$ and any $0 \leq s \leq t \leq T$.

As a direct consequence of the definition, $(\Phi_t)_t$ is uniformly Lipschitz in t with respect to $\|\cdot\|_{\tilde{\mathbb{D}}^\infty(\mathbb{R}^m)}$ or $\|\cdot\|_{\tilde{\mathbb{D}}^1(\mathbb{R}^m)}$. If Φ is constructible, then it is possible to introduce a new random variable to account for the dependence of the path. As an example, if $\Phi(\Theta) = \int_0^T \Theta(t) dt$, we pick $\Phi_t(\Theta) = \int_0^t \Theta(s) ds$ and $g(s, t, \Phi_s(\Theta), \Theta \mathbb{1}_{[s,t]}) = \Phi_s(\Theta) + \int_s^t (\Theta \mathbb{1}_{[s,t]})(u) du$. In other words, the role of $\Phi_t(\Theta)$ is to carry the history of the path, while the function g is designed to update the $\Phi_t(\Theta)$.

3.5.1 Quantization and the lattice grid

Quantization method [34] provides a cheap algorithm to estimate expectation, by establishing a grid associated with probability based on the distribution of the random variable. We apply quantization to the Brownian increments, by replacing ΔB_{t_i} by $\Delta \hat{B}_{t_i}$.

Lemma 3.5.2 (Zador's theorem, [34]). *Given $p \geq 1$, there exists an function $G_M : \mathbb{R}^d \mapsto \Gamma_M$ where Γ_M is the optimal grid of standard d -dimensional normal distribution with $|\Gamma_M| = M$ (i.e. number of grid points in Γ_M is M). Moreover, G_M is defined as the closest neighbour rule projection on Γ_M . Denoting $\Delta \hat{B}_{t_i} = \sqrt{h_i} G_M(h_i^{-\frac{1}{2}} \Delta B_{t_i})$, we obtain*

$$\mathbb{E} \left[|\Delta B_{t_i} - \Delta \hat{B}_{t_i}|^p \right]^{\frac{1}{p}} \leq C_{p,d} \sqrt{h_i} M^{-\frac{1}{d}}.$$

Remark 3.6. According to [34], the optimal grid of a random variable is unique for $d = 1$ and not necessarily unique for $d > 1$. Moreover, the quantization version of a symmetrical random variable does not necessarily need to be centred for $d > 1$. For $d = 1$, we can use symmetry to conclude the symmetry of the grid.

We discretise $\mathbb{R}$ into $2\kappa + 2$ points to form a grid in $\mathbb{R}^m$

$$\Gamma = \{x = (x^1, \dots, x^m) \in \mathbb{R}^m \mid |x - X_0| \in \delta\mathbb{Z}^m, |x^j - X_0^j| \leq \kappa\delta, \forall 1 \leq j \leq d\}.$$

Meanwhile, we define a projection $\text{Proj}_{X_0} : \mathbb{R}^m \mapsto \Gamma$, centred at $X_0 \in \Gamma$, such that the j -th entry of $\text{Proj}_{X_0}(x)$ is defined as

$$\begin{aligned} \text{Proj}_{X_0}(x)^j &= \sum_{i=0}^{\kappa} (X_0^j + i\delta) \mathbb{1}_{[X_0^j + (i-\frac{1}{2})\delta, X_0^j + (i+\frac{1}{2})\delta)}(x^j) \\ &\quad + \sum_{i=-1}^{-\kappa} (X_0^j + i\delta) \mathbb{1}_{[X_0^j + (i-\frac{1}{2})\delta, X_0^j + (i+\frac{1}{2})\delta)}(x^j) \\ &\quad + (X_0^j + \kappa\delta) \mathbb{1}_{[X_0^j + (i+\frac{1}{2})\delta, \infty)}(x^j) + (X_0^j - \kappa\delta) \mathbb{1}_{(-\infty, X_0^j - (\kappa-\frac{1}{2})\delta)}(x^j). \end{aligned}$$

Remark 3.7. If the timesteps are not equidistant, naturally we will choose an adaptive projection, where each δ depends on the timestep h_i .

We are now in a position to discretise the path-dependent term $\Phi_s(\hat{X}^\pi)$. As the discretisation for the forward SDE, we consider a fixed grid centred at $\Phi_0(X)$, namely $\hat{\Gamma} = \{x = (x^1, \dots, x^{d_1}) \in \mathbb{R}^{d_1} \mid |x - \Phi_0(X)| \in \hat{\delta}\mathbb{Z}^{d_1}, |x^j - \Phi_0(X)^j| \leq \hat{\kappa}\hat{\delta}, \forall 1 \leq j \leq d_1\}$.

Similarly, the projection operator is defined as

$$\begin{aligned} \text{Proj}_{\Phi_0(X)}^{(1)}(x)^j &= \sum_{i=0}^{\hat{\kappa}} (\Phi_0(X)^j + i\hat{\delta}) \mathbb{1}_{[\Phi_0(X)^j + (i-\frac{1}{2})\hat{\delta}, \Phi_0(X)^j + (i+\frac{1}{2})\hat{\delta})}(x^j) \\ &\quad + \sum_{i=-1}^{-\hat{\kappa}} (\Phi_0(X)^j + i\hat{\delta}) \mathbb{1}_{[\Phi_0(X)^j + (i-\frac{1}{2})\hat{\delta}, \Phi_0(X)^j + (i+\frac{1}{2})\hat{\delta})}(x^j) \\ &\quad + (\Phi_0(X)^j + \hat{\kappa}\hat{\delta}) \mathbb{1}_{[\Phi_0(X)^j + (i+\frac{1}{2})\hat{\delta}, \infty)}(x^j) \\ &\quad + (\Phi_0(X)^j - \hat{\kappa}\hat{\delta}) \mathbb{1}_{(-\infty, \Phi_0(X)^j - (\hat{\kappa}-\frac{1}{2})\hat{\delta})}(x^j). \end{aligned}$$

3.5.2 Discretisation of forward processes

Suppose the forward SDE $(X_t)_{t \geq 0}$ satisfies the assumption [A2]. We quote the stability theorem which corresponds to the numerical scheme of Euler-Maruyama type in [12] and slightly generalise their result, which relies on BDG inequalities for càdlàg martingales [39] and discrete Gronwall lemma.

Proposition 3.5.3 (Stability of Euler scheme, modification of Lemma A.1 [12]). *Let $(t, x) \mapsto b(t, x)$ and $(t, x) \mapsto \sigma(t, x)$ be functions that are $\frac{1}{2}$ -Hölder continuous in t and Lipschitz in x with $|b(0, 0)| + |\sigma(0, 0)| < \infty$. Fix $p \geq 1$ and set $(\xi_i)_{i=0}^N$ as a sequence of L^{2p} integrable random variables. Assume $(\gamma_i)_{0 \leq i \leq N-1} \in L^{2p}(\mathbb{R}^d)$ is a sequence of independent random variables such that for all $0 \leq i \leq N-1$*

$$\hat{\gamma} = \sup_j \frac{|\mathbb{E}[\gamma_j]|}{\sqrt{h_j}}, \quad \mathbb{E} \left[|\gamma_i|^{2p} \mid \mathcal{F}_{t_i} \right] = \mathbb{E} [|\gamma_i|^{2p}] \leq C$$

for some $C \geq 0$. Given two schemes $(\mathbf{X}_i)_{0 \leq i \leq N}$ and $(\tilde{\mathbf{X}}_i)_{0 \leq i \leq N}$ defined as

$$\begin{aligned} \mathbf{X}_{i+1} &= \mathbf{X}_i + b(t_i, \mathbf{X}_i)h_i + \sigma(t_i, \mathbf{X}_i)\sqrt{h_i}\gamma_i, \\ \tilde{\mathbf{X}}_{i+1} &= \tilde{\mathbf{X}}_i + b(t_i, \tilde{\mathbf{X}}_i)h_i + \sigma(t_i, \tilde{\mathbf{X}}_i)\sqrt{h_i}\gamma_i + \zeta_i, \end{aligned}$$

their difference can be controlled by

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq i \leq N} |\mathbf{X}_i - \tilde{\mathbf{X}}_i|^{2p} \right] \\ & \leq C_p \exp(C_{p,T} \max(\hat{\gamma}, 1)^{2p}) \left(\mathbb{E} [|\mathbf{X}_0 - \tilde{\mathbf{X}}_0|^{2p}] + \mathbb{E} \left[\left(\sum_{j=0}^{i-1} |\zeta_j| \right)^{2p} \right] \right). \end{aligned}$$

Proof. Deferred to Section A.2. □

We are ready to state error due to the projection of the forward process on the grid. As we have seen, the error of the scheme (on both SDE and BSDE) relies on the largest interval h for any partition π .

Lemma 3.5.4 (Section 4.1 in [12]). *Suppose $(X_t)_{t \geq 0}$ satisfies the assumption [A2] and $\pi : 0 = t_0 < t_1 < \dots < t_N = T$ is a partition with $h_i = t_{i+1} - t_i$ and $h = \sup_i |h_i| \leq KN^{-1}$ for some $K > 0$. Recall the Euler-Maruyama estimation of X ,*

$$X_{i+1}^\pi = X_{t_i}^\pi + b(t_i, X_{t_i}^\pi)h_i + \sigma(t_i, X_{t_i}^\pi)\Delta B_{t_i}, \quad X_0^\pi = X_0,$$

we propose the version with projection

$$\hat{X}_{i+1}^\pi = \text{Proj}_{X_0} \left(\hat{X}_i^\pi + b(t_i, \hat{X}_i^\pi)h_i + \sigma(t_i, \hat{X}_i^\pi)\Delta \hat{B}_{t_i} \right), \quad \hat{X}_0 = X_0. \quad (3.45)$$

For any $p \geq 2$, the difference of the two schemes is given by

$$\mathbb{E} \left[\sup_{0 \leq i \leq N} |\hat{X}_i^\pi - X_{t_i}^\pi|^p \right] \leq C \sup_{0 \leq i \leq N} h_i^{-\frac{p}{2}} M^{-\frac{p}{d}} + CN^p \delta^p + C_r N (\kappa \delta)^{-rp}$$

provided $\sup_{0 \leq i \leq N} M^{-\frac{1}{d}} h_i^{-\frac{1}{2}} \leq C$ and $\sup_i (\delta h_i) \leq C$ for some $C > 0$. Moreover, when we define

$$\hat{X}^\pi(t) = \hat{X}_i^\pi \quad \forall t \in [t_i, t_{i+1}), \quad \hat{X}^\pi(T) = \hat{X}_N^\pi, \quad (3.46)$$

for all $0 \leq i \leq N-1$, we also have

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{X}^\pi(t) - X^\pi(t)|^p \right] \leq C \sup_{0 \leq i \leq N} h_i^{-\frac{p}{2}} M^{-\frac{p}{d}} + CN^p \delta^p + C_r N (\kappa \delta)^{-rp}.$$

where $X^\pi(t)$ is defined in (3.3).

Proof. See Section A.2. □

The following convergence is only based on two examples, as the choice of Φ_t depends on the choice of Φ and the choice of Φ_t is not unique in many cases. We will consider the following cases for the path-dependent element $\Phi_t(X)$ to illustrate our scheme.

- (i) Basket lookback put options: $\xi = (K - \sum_{i=1}^m w_i \inf_{0 \leq t \leq T} X_s^i)^+$, where $w_i > 0$ represents the weight of each $\inf_{0 \leq t \leq T} X_s^i$ in the basket, such that $\sum_{1 \leq i \leq m} w_i = 1$. A natural choice is $(\Phi_t(X))^i = \inf_{0 \leq s \leq t} X_s^i$ for all $0 \leq t \leq T$ and $1 \leq i \leq m$, with corresponding function $(g(s, t, \Phi_s(X), X \mathbb{1}_{[s, t]}))^i = \min\{(\Phi_s(X))^i, (X \mathbb{1}_{[s, t]})^i\}$, while one can define the version where $\sup_{0 \leq t \leq T}$ is used instead.
- (ii) Basket Asian put options: $\xi = (K - \sum_{i=1}^m w_i \int_0^T X_s^i ds)^+$, where $w_i > 0$ represents the weight of each X_s^i in the basket, such that $\sum_{1 \leq i \leq m} w_i = 1$. We can choose $\Phi_t(X) = \sum_{i=1}^m w_i \int_0^T X_s^i ds$ for all $0 \leq t \leq T$ to ensure $\Phi_t(X)$ is one-dimensional, while we are allowed to choose $g(s, t, \Phi_s(X), X \mathbb{1}_{[s, t]}) = \Phi_s(X) + \int_s^t X(u) \mathbb{1}_{[s, t]}(u) du$.

In order to choose appropriate $\hat{\kappa}$ and $\hat{\delta}$ to estimate $\Phi_{t_i}(\hat{X}^\pi)$, we try to show:

Lemma 3.5.5. *Let $(\hat{X}^\pi)$ defined as equation (3.46) and let the assumptions in Lemma 3.5.4 be satisfied. Let Φ be the payoff of either Basket lookback put option or basket asian put options, with $\hat{\Phi}_{t_i}(\hat{X}^\pi)$ defined as*

$$\hat{\Phi}_{t_{i+1}}(\hat{X}^\pi) = Proj_{\Phi_0(X)}^{(1)}(g(t_i, t_{i+1}, \hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}^\pi \mathbb{1}_{[t_i, t_{i+1}]}) , \quad \hat{\Phi}_0(\hat{X}^\pi) = \Phi_0(X).$$

Fixing an arbitrary $p \geq 2$. For basket lookback options, we have

$$\mathbb{E} \left[\sup_{0 \leq i \leq N} \left| \hat{\Phi}_{t_i}(\hat{X}^\pi) - \Phi_{t_i}(X^\pi) \right|^p \right] \leq C_p \mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{X}^\pi(t) - X^\pi(t)|^p \right],$$

while for basket Asian options, there exists a constant C_r such that

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq i \leq N} \left| \hat{\Phi}_{t_i}(\hat{X}^\pi) - \Phi_{t_i}(X^\pi) \right|^p \right] \\ \leq C_p(N\hat{\delta})^p + C_{p,r}N^p(\hat{\kappa}\hat{\delta})^{-rp} + C_p \mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{X}^\pi(t) - X^\pi(t)|^p \right]. \end{aligned}$$

Proof. Splitting the error and applying the definition of Φ_t as the following

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq i \leq N} \left| \hat{\Phi}_{t_i}(\hat{X}^\pi) - \Phi_{t_i}(X^\pi) \right|^p \right] \\ \leq C_p \left(\mathbb{E} \left[\sup_{0 \leq i \leq N} \left| \hat{\Phi}_{t_i}(\hat{X}^\pi) - \Phi_{t_i}(\hat{X}^\pi) \right|^p \right] + \mathbb{E} \left[\sup_{0 \leq i \leq N} \left| \Phi_{t_i}(\hat{X}^\pi) - \Phi_{t_i}(X^\pi) \right|^p \right] \right), \end{aligned}$$

it remains to calculate the error for the first term. Regarding the basket lookback options, we can set $\hat{\kappa} = \kappa$ and $\hat{\delta} = \delta$, since the value of $\Phi_{t_i}(\hat{X}^\pi)$ also lies on the grid Γ . As a result, for any i and $\hat{X}^\pi$, $\hat{\Phi}_{t_i} = \Phi_{t_i}$ and error is

$$\mathbb{E} \left[\sup_{0 \leq i \leq N} \left| \hat{\Phi}_{t_i}(\hat{X}^\pi) - \Phi_{t_i}(X^\pi) \right|^p \right] \leq C_p \mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{X}^\pi(t) - X^\pi(t)|^p \right].$$

The error corresponding to basket Asian options requires more computation, starting with rewriting the scheme as

$$\begin{aligned} \hat{\Phi}_{t_{i+1}}(\hat{X}^\pi) &= \hat{\Phi}_{t_i}(\hat{X}^\pi) + \sum_{j=1}^m w_j \int_{t_i}^{t_{i+1}} \hat{X}^{\pi,j}(s) ds + \zeta_i, \quad \hat{\Phi}_0(\hat{X}^\pi) = \Phi_0(X), \\ \zeta_i &= \text{Proj}_{\Phi_0(X)}^{(1)}(\tilde{\Phi}_{t_i}) - \tilde{\Phi}_{t_i}, \quad \tilde{\Phi}_{t_i} := \hat{\Phi}_{t_i}(\hat{X}^\pi) + \sum_{j=1}^m w_j \int_{t_i}^{t_{i+1}} \hat{X}^{\pi,j}(s) ds. \end{aligned}$$

As $\zeta_i \leq \tilde{\delta} + \frac{(\tilde{\Phi}_{t_{i+1}} - \Phi_0(X))^{r+1}}{(\kappa\delta)^r} \mathbb{1}_{\{|\tilde{\Phi}_{t_i} - \Phi_0(X)| > \kappa\delta\}}$, we are able to show that

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq i \leq N} \left| \hat{\Phi}_{t_i}(\hat{X}^\pi) - \Phi_{t_i}(\hat{X}^\pi) \right|^p \right] \\ \leq C_p(N\hat{\delta})^p + C_p N^p(\hat{\kappa}\hat{\delta})^{-rp} \mathbb{E} \left[\sup_{0 \leq i \leq N} |\tilde{\Phi}_{t_i} - \Phi_0(X)|^{p(r+1)} \right], \end{aligned}$$

whereas the moments of $\tilde{\Phi}_{t_i} - \Phi_0(X)$ is bounded if we consider the inequality

$$\begin{aligned} \left| \tilde{\Phi}_{t_{i+1}} - \Phi_0(X) \right| &\leq \left| \text{Proj}_0^{(1)} \left(\tilde{\Phi}_{t_{i+1}} - \Phi_0(X) + \sum_{j=1}^m w_j \int_{t_i}^{t_{i+1}} \hat{X}^{\pi,j}(s) ds \right) \right| \\ &\leq \sum_{j=1}^m w_j \int_{t_0}^{t_{i+1}} \left| \hat{X}^{\pi,j}(s) \right| ds. \end{aligned}$$

□

3.5.3 Impact on the backward component

Lemma 3.5.6. *Let Φ be constructible with functions g and $\tilde{g}$ and let $\pi : 0 = t_0 < t_1 < \dots < t_N = T$ be a partition with $h_i = t_{i+1} - t_i$ and $h = \sup_i |h_i| \leq KN^{-1}$ for some $K > 0$. Let X satisfies [A2], we introduce the following forward processes*

$$\begin{aligned}\hat{\Phi}_{t_{i+1}}(\hat{X}^\pi) &:= \text{Proj}_{\Phi_0(X)}^{(1)}(g(t_i, t_{i+1}, \Phi_{t_{i+1}}(\hat{X}^\pi), \hat{X}^\pi \mathbb{1}_{[t_1, t_{i+1}]}) , \quad \hat{\Phi}_0(\hat{X}^\pi) = \Phi_0(X), \\ \hat{X}_{i+1}^\pi &= \text{Proj}_{X_0} \left(\hat{X}^\pi + b(t_i, \hat{X}_i^\pi) h_i + \sigma(t_i, \hat{X}_i^\pi) \Delta \hat{B}_{t_i} \right), \quad \hat{X}_0^\pi = X_0,\end{aligned}$$

we obtain the following implicit scheme based on quantization

$$\begin{aligned}\hat{\mathcal{Y}}_i^\pi &:= \mathbb{E} \left[\hat{\mathcal{Y}}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] + h_i f^{(n)} \left(t_i, \mathbb{E} \left[\hat{\mathcal{Y}}_{i+1}^\pi \mid \mathcal{F}_{t_i} \right], \hat{\mathcal{Z}}_i^\pi \right), \\ \hat{\mathcal{Z}}_i^\pi &:= \mathbb{E} \left[\hat{\mathcal{Y}}_{i+1}^\pi \hat{H}_i^R \mid \mathcal{F}_{t_i} \right], \quad \hat{H}_i^R := \phi_{Rh_i^{-\frac{1}{2}}} \left(\frac{\Delta \hat{B}_{t_i}}{h_i} \right) \\ \hat{\mathcal{Y}}_N^\pi &= \tilde{g} \left(\hat{\Phi}_{t_N}(\hat{X}^\pi), \hat{X}_N^\pi \right), \quad \hat{\mathcal{Z}}_N^\pi = 0.\end{aligned} \tag{3.47}$$

Then, the process $(\hat{\mathcal{Y}}_i^\pi, \hat{\mathcal{Z}}_i^\pi, \hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi)$ defined in equation (3.47) is a Markov process.

Proof. $\hat{X}_i^\pi$ is a Markov process by construction. Meanwhile, there exists a sequence of transition kernels $(P_i)_{i=0}^{N-1}$ such that

$$\begin{aligned}& \mathbb{E} \left[F \left(\hat{\Phi}_{t_{i+1}}(\hat{X}^\pi), \hat{X}_{i+1}^\pi \right) \mid \mathcal{F}_{t_i} \right] \\ &= \mathbb{E} \left[F \left(\text{Proj}_{\Phi_0(X)}^{(1)}(g(t_i, t_{i+1}, \Phi_{t_{i+1}}(\hat{X}^\pi), \hat{X}^\pi \mathbb{1}_{[t_1, t_{i+1}]}) , \hat{X}_{i+1}^\pi \right) \mid \mathcal{F}_{t_i} \right] \\ &= \int_{\mathbb{R}^{d_1} \times \mathbb{R}^d} F(y_1, y_2) P_i((\hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi), d(y_1, y_2)),\end{aligned}$$

for any Borel measurable $F : \mathbb{R}^{d_1} \times \mathbb{R}^d \mapsto \mathbb{R}$.

Regarding the backward components, we write $\hat{\mathcal{Y}}_i^\pi = u^\pi(t_i, \hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi)$ and $\hat{\mathcal{Z}}_i^\pi = v^\pi(t_i, \hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi)$ and apply induction starting with N . Given $(\hat{\mathcal{Y}}_{i+1}^\pi, \hat{\mathcal{Z}}_{i+1}^\pi)$ that can be written in to borel functions of $(t_i, \Phi_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi)$, the representation of $\hat{\mathcal{Z}}_i^\pi$ in terms of Borel measurable functions v^π can be proved easily. By definition,

$$\begin{aligned}\hat{\mathcal{Y}}_i^\pi &= h_i f^{(n)} \left(t_i, \tilde{u}^\pi(t_i, \hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}_{i+1}^\pi), v^\pi(t_i, \hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi) \right), \\ &+ \tilde{u}^\pi(t_i, \hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi) \\ \tilde{u}^\pi(t_i, \hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}_{i+1}^\pi) &:= \mathbb{E} \left[u^\pi(t_{i+1}, \hat{\Phi}_{t_{i+1}}(\hat{X}^\pi), \hat{X}_{i+1}^\pi) \mid \hat{\Phi}_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi \right].\end{aligned} \tag{3.48}$$

as we required. $\square$

By observing that $\Phi_{t_i}(\hat{X}^\pi)$ and $\hat{X}_i^\pi$ only take discrete values, the above lemma allows us to compute the conditional expectation explicitly via computing the values v^π and u^π on the grid. Implementing the discrete structure of $\Delta\hat{B}_{t_i}$, (3.47) is equivalent to

Corollary 3.5.7. *Under the same condition as in Lemma 3.5.6, $(\hat{\mathcal{Y}}_i^\pi, \hat{\mathcal{Z}}_i^\pi)$ within the numerical scheme (3.47) has the representation*

$$\left(\hat{\mathcal{Y}}_i^\pi, \hat{\mathcal{Z}}_i^\pi\right) = \left(u^\pi\left(t_i, \Phi_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi\right), v^\pi\left(t_i, \Phi_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi\right)\right)$$

for some Borel-measurable functions u^π, v^π . For any $(x, y) \in \hat{\Gamma} \times \Gamma$,

$$\left(\hat{\mathcal{Y}}_i^\pi|_{\Phi_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi=(x,y)}, \hat{\mathcal{Z}}_i^\pi|_{\Phi_{t_i}(\hat{X}^\pi), \hat{X}_i^\pi=(x,y)}\right) = (u^\pi(t_i, x, y), v^\pi(t_i, x, y))$$

such that $(u^\pi(t_i, x, y), v^\pi(t_i, x, y))$ is computed by

$$\begin{aligned} u^\pi(t_i, x, y) &= \tilde{u}^\pi(t_i, x, y) + h_i f^{(n)}(t_i, \tilde{u}^\pi(t_i, x, y), v^\pi(t_i, x, y)), \\ v^\pi(t_i, x, y) &= \sum_{j=1}^M u^\pi\left(t_{i+1}, Proj_{\Phi_0}^{(1)}(g(t_i, t_{i+1}, x, y, w^\pi(t_i, y, h_i, z_j)))\right), \\ &\quad w^\pi(t_i, y, h_i, z_j)) \phi_{\frac{R}{h_i}}(\sqrt{h_i} z_j) p_j, \\ \tilde{u}^\pi(t_i, x, y) &= \sum_{j=1}^M u^\pi\left(t_{i+1}, Proj_{\Phi_0}^{(1)}(g(t_i, t_{i+1}, x, y, w^\pi(t_i, y, h_i, z_j)))\right), \\ &\quad w^\pi(t_i, y, h_i, z_j)) p_j, \\ u^\pi(t_N, x, y) &= \tilde{g}(x, y), \quad v^\pi(t_N, x, y) = 0, \end{aligned} \tag{3.49}$$

where $p_j = \mathbb{P}\left(G_M\left(\frac{\Delta B_{t_i}}{\sqrt{h_i}}\right) = z_j\right)$ and

$$w^\pi(t_i, y, h_i, z_j) := Proj_{X_0}(y + b(t_i, y)h_i + \sigma(t_i, y)\sqrt{h_i}z_j).$$

We revisit the case of basket lookback option and the case of Asian options to conclude the convergence rate.

Theorem 3.5.8. *Assume $\pi : 0 = t_0 < t_1 < \dots < t_N = T$ is a partition with $h_i = t_{i+1} - t_i$ and $h = \sup_i |h_i| \leq KN^{-1}$ for some $K > 0$. Moreover, the assumptions and the choices of parameters R, n, α are inherited from Theorem 3.1.3. Suppose the number of grid points (of discretising Brownian motion) M is at least of order*

$$M \sim O\left(N^{\frac{d}{2}} \sup_i |h_i|^{-\frac{d}{2}}\right).$$

Let $p \geq 1$ be fixed and consider $(\hat{\mathcal{Y}}^\pi, \hat{\mathcal{Z}}^\pi)$ defined in (3.47).

(i) For basket lookback options, we have

$$\begin{aligned}\mathbb{E} \left[\sup_{0 \leq i \leq N} |\hat{\mathcal{Y}}_i^\pi - Y_{t_i}|^{2p} \right] &\leq C \sup_{0 \leq i \leq N} h_i^{-p} M^{-\frac{2p}{d}} + CN^{2p} \delta^{2p} + C_r N^{2p} (\kappa \delta)^{-2rp} \\ &\quad + C_p h^p \ln \left(\frac{1}{h} \right)^p + C_{\alpha, p, K, \epsilon} h^p, \\ \mathbb{E} \left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} |\hat{\mathcal{Z}}_i^\pi - Z_s|^2 ds \right] &\leq C \sup_{0 \leq i \leq N} h_i^{-1} M^{-\frac{2}{d}} + CN^2 \delta^2 + C_r N^2 (\kappa \delta)^{-2r} \\ &\quad + Ch \ln \left(\frac{1}{h} \right) + C_{\alpha, K, \epsilon} h.\end{aligned}$$

(ii) For basket Asian-type options,

$$\begin{aligned}\mathbb{E} \left[\sup_{0 \leq i \leq N} |\hat{\mathcal{Y}}_i^\pi - Y_{t_i}|^{2p} \right] &\leq C \sup_{0 \leq i \leq N} h_i^{-p} M^{-\frac{2p}{d}} + CN^{2p} \delta^{2p} + C_r N^{2p} (\kappa \delta)^{-2rp} \\ &\quad + C_p (N \hat{\delta})^{2p} + C_{p, r} N^{2p} (\hat{\kappa} \hat{\delta})^{-2rp} + C_{\alpha, p, K, \epsilon} h^p, \\ \mathbb{E} \left[\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} |\hat{\mathcal{Z}}_i^\pi - Z_s|^2 ds \right] &\leq C \sup_{0 \leq i \leq N} h_i^{-1} M^{-\frac{2}{d}} + CN^2 \delta^2 + C_r N^2 (\kappa \delta)^{-2r} \\ &\quad + C (N \hat{\delta})^2 + C_r N^2 (\hat{\kappa} \hat{\delta})^{-2r} + C_{\alpha, K, \epsilon} h.\end{aligned}$$

Proof. This result can be obtained by the application of Theorem 3.3.5, Theorem 3.1.3 and Lemma 3.5.5. The mean of $\hat{H}_i^R$ is controlled by

$$\mu_i := \left| \mathbb{E} \left[\hat{H}_i^R \right] \right| \leq \frac{1}{h_i} \mathbb{E} \left[\left| \Delta \hat{B}_{t_i} - \Delta B_{t_i} \right| \right] \leq C_{1, d} \frac{1}{\sqrt{h_i}} M^{-\frac{1}{d}},$$

where $C_{1, d}$ is the constant related to the quantization of Gaussian random variables. A sufficient condition on M to satisfy **[A6](ii)** is

$$\hat{\mu}_i L_z(n) \leq C_{1, d} \frac{1}{\sqrt{h_i}} M^{-\frac{1}{d}} C N^\alpha \leq 1 \implies M^{\frac{1}{d}} \geq \frac{N^\alpha}{C_{1, d} C \sqrt{h_i}}.$$

Moreover, for $M \sim O \left(N^{\frac{d}{2}} \sup_i |h_i|^{-\frac{d}{2}} \right)$, $M^{-\frac{1}{d}} \sup_i (h_i^{-\frac{1}{2}}) = O(N^{-\frac{1}{2}})$, which satisfies the the constraint in Lemma 3.5.4.

As the error can be split into

$$\mathbb{E} \left[\sup_{0 \leq i \leq N} |\hat{\mathcal{Y}}_i^\pi - Y_{t_i}|^{2p} \right] \leq C_p \left(\mathbb{E} \left[\sup_{0 \leq i \leq N} |Y_i^\pi - Y_{t_i}|^{2p} \right] + \mathbb{E} \left[\sup_{0 \leq i \leq N} |\hat{\mathcal{Y}}_i^\pi - \mathcal{Y}_i^\pi|^{2p} \right] \right),$$

we focus on the latter term and plug in the expression of $\hat{\Delta} \mathcal{Y}_N^\pi, \zeta_j^Z$ as in

$$\zeta_i^Z = \mathbb{E} \left[\hat{Y}_{i+1}^\pi (\hat{H}_i^R - H_i^R) \mid \mathcal{F}_{t_i} \right] \quad (3.50)$$

into Theorem 3.3.5 to derive

$$\begin{aligned}
\mathbb{E} \left[\sup_{0 \leq i \leq N} |\hat{\mathcal{Y}}_i^\pi - \mathcal{Y}_i^\pi|^{2p} \right] &\leq C_p \left(\mathbb{E} [|\Delta Y_N|^{2pq^*}]^{\frac{1}{q^*}} + \mathbb{E} \left[\left(\sum_{i=0}^{N-1} |\zeta_i^Z|^2 h_i \right)^{2pq^*} \right]^{\frac{1}{2q^*}} \right) \\
&\leq C_p \left(L_{\hat{g}}^{2p} \mathbb{E} \left[\sup_{0 \leq t \leq T} |\hat{X}^\pi(t) - X^\pi(t)|^{2pq^*} \right]^{\frac{1}{q^*}} \right. \\
&\quad \left. + \mathbb{E} \left[\left(\sum_{i=0}^{N-1} \|\hat{\mathcal{Y}}^\pi\|_{\mathbb{S}^\infty(\mathbb{R})}^2 \mathbb{E} [|\hat{H}_i^R - H_i^R|]^2 h_i \right)^{2pq^*} \right]^{\frac{1}{2q^*}} \right) \\
&\leq C_p \left(\sup_{0 \leq i \leq N} h_i^{-p} M^{-\frac{2p}{d}} + CN^{2p} \delta^{2p} + C_r N^{2p} (\kappa \delta)^{-2rp} + \sup_{0 \leq i \leq N} h_i^{-p} M^{-\frac{2p}{d}} \right) \\
&\leq C_p \left(\sup_{0 \leq i \leq N} h_i^{-p} M^{-\frac{2p}{d}} + CN^{2p} \delta^{2p} + C_r N^{2p} (\kappa \delta)^{-2rp} \right),
\end{aligned}$$

where we have used the fact that

$$\mathbb{E} [|\hat{H}_i^R - H_i^R|] \leq \mathbb{E} \left[\left| \frac{\Delta \hat{B}_{t_i}}{h_i} - \frac{\Delta \hat{B}_{t_i}}{h_i} \right| \right] \leq C_{1,d} \sqrt{h_i} M^{\frac{1}{d}}.$$

Meanwhile, we apply our result on Y to derive the error appears in Z -component

$$\begin{aligned}
\mathbb{E} \left[\sum_{i=0}^{N-1} |\hat{Z}_i^\pi - Z_i^\pi|^2 h_i \right] &\leq C \left(\sup_{0 \leq i \leq N} h_i^{-1} M^{-\frac{2}{d}} + CN^2 \delta^2 + C_r N^2 (\kappa \delta)^{-2r} \right) \\
&\quad + \mathbb{E} \left[\left(\sum_{i=0}^{N-1} \|\hat{\mathcal{Y}}^\pi\|_{\mathbb{S}^\infty}^2 \mathbb{E} [|\hat{H}_i^R - H_i^R|]^2 h_i \right)^2 \right] \\
&\leq C \left(\sup_{0 \leq i \leq N} h_i^{-1} M^{-\frac{2}{d}} + CN^2 \delta^2 + C_r N^2 (\kappa \delta)^{-2r} \right).
\end{aligned}$$

Regarding the convergence error on Asian-type options, the only difference within the argument is adding the error of new grid. $\square$

Assume we have a uniform grid, then the requirement on M is it is at least of order $O(N^d)$. We can see that the optimal choice of M is of order $O(N^d)$, $\delta = N^{-\frac{3}{2}}$ and $\hat{\delta} = N^{-\frac{3}{2}}$. Suppose $r = \frac{3}{2}$, we then choose $\kappa = N^{\frac{5}{2}}$, and $\hat{\kappa} = N^{\frac{5}{2}}$. The total number of grid is $\kappa^m \times \hat{\kappa}$ for Asian options, while the grid size changed to κ^{2m} for lookback types of payoff (depending on the definition of the option). As $O(M)$ computations is required to compute the value for one grid at time t_i , the overall complexity for Asian option is $O(\text{number of grid} \times M \times N)$, where

$$\text{complexity} = O \left(N^{\frac{5}{2}(m+1)} \right) \times O(N^d) \times O(N).$$

In particular, if $d = m$, the computation cost is of order $= O\left(N^{\frac{7}{2}(d+1)}\right)$, which is enormous especially when d is large. Due to the restriction on κ , the complexity must be at least of order $O\left(N^{\frac{5+\epsilon}{2}(d+1)}\right)$, provided $\kappa\delta = O(N^{\frac{1}{2}\epsilon})$ for some $\epsilon > 0$.

3.6 Numerical experiments

This section is devoted to demonstrating the effectiveness of our scheme, through a series of numerical examples. For simplicity, $m = d = 2$ will be fixed. Given a standard Brownian motion $B_t = (B_t^1, B_t^2)$, we consider the forward process as a geometric Brownian motion

$$\begin{aligned} dX_t^1 &= r_1 X_t^1 dt + \sigma_1 X_t^1 dB_t^1, \\ dX_t^2 &= r_2 X_t^2 dt + \sigma_2 X_t^2 (\rho dB_t^1 + \sqrt{1 - \rho^2} dB_t^2), \\ X_0^1 &= x_0^1 \in \mathbb{R}, \quad X_0^2 = x_0^2 \in \mathbb{R}, \end{aligned} \tag{3.51}$$

such that the explicit expression for X_t is given by

$$\begin{aligned} X_t^1 &= X_0^1 \exp\left(\left(r_1 - \frac{1}{2}\sigma_1^2\right)t + \sigma_1 B_t^1\right), \\ X_t^2 &= X_0^2 \exp\left(\left(r_2 - \frac{1}{2}\sigma_2^2\right)t + \sigma_2\left(\rho B_t^1 + \sqrt{1 - \rho^2} B_t^2\right)\right). \end{aligned}$$

As a classic example, we consider f to be a quadratic function of Z .

Example 3.1 (Square driver f). *Let $f(t, y, z) = \frac{a}{2}|z|^2$ for some $a \in \mathbb{R}$, the BSDE is then*

$$Y_t = \xi + \int_t^T \frac{a}{2} |Z_s|^2 ds - \int_t^T Z_s \cdot dB_s \tag{3.52}$$

Three types of bounded terminal conditions are considered, namely

- 1) (sin payoff with an integral) $\xi = \hat{a}(\sin(X_T^1 + X_T^2) + \sin(\frac{1}{2T} \int_0^T X_u^1 + X_u^2 du))$,
- 2) (Put payoff with an integral) $\xi = \hat{a}(K - \frac{1}{2T} \int_0^T X_u^1 + X_u^2 du)^+$,
- 3) (Put payoff with infimum) $\xi = \hat{a}(K - \inf_{0 \leq t \leq T} \{X_t^1, X_t^2\})^+$.

As a consequence of the Cole-Hopf transformation, we convert equation (3.52) in Example 3.1 to $(\tilde{Y}, \tilde{Z}) = (\exp(aY), a \exp(aY)Z)$, where

$$d\tilde{Y}_t = a\tilde{Y}_t \left(-\frac{a}{2}|Z_s|^2 ds + Z_s dB_s\right) + \frac{a^2}{2}\tilde{Y}_t |Z_s|^2 ds = \tilde{Z}_s \cdot dB_s, \quad \tilde{Y}_T = \exp(a\xi).$$

This is a linear BSDE such that Y is given by

$$\tilde{Y}_t = \mathbb{E}[\exp(a\xi) \mid \mathcal{F}_t], \quad Y_t = \frac{1}{a} \ln(\mathbb{E}[\exp(a\xi) \mid \mathcal{F}_t]), \quad Y_0 = \frac{1}{a} \ln(\mathbb{E}[\exp(a\xi)]).$$

Meanwhile, assuming $\xi \in \mathbb{D}^{1,2}$, Theorem 3.2.5 implies

$$Z_t = \frac{\mathbb{E}[\exp(a\xi) D_t \xi \mid \mathcal{F}_t]}{\mathbb{E}[\exp(a\xi) \mid \mathcal{F}_t]}, \quad Z_0 = \frac{\mathbb{E}[\exp(a\xi) D_0 \xi]}{\mathbb{E}[\exp(a\xi)]}.$$

Hence the exact solution can be simulated directly by Monte-Carlo methods.

Figure	Figure 3.1, 3.2	Figure 3.3, 3.4	Figure 3.5, 3.6	Figure 3.7, 3.8
(X_0^1, X_0^2)	(1, 1.2)	(5, 6)	(5, 6)	(1, 1.2)
(r_1, r_2)	(0.1, 0.05)	(0.1, 0.05)	(0.1, 0.05)	(0.2, 0.4)
(σ_1, σ_2)	(0.5, 0.4)	(0.4, 0.3)	(0.4, 0.3)	(0.6, 0.3)
ρ	0.6	0.4	0.4	-0.3
a	-1	-0.2	-0.2	2
Type	sin	Put, $K = 8$	Put, $K = 8$	Put, $K = 3$
Path	Integral	Integral	Integral	Minimum
$\hat{a}$	2	2	2	-2
T	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
Method	Quantization	Quantization	Regression	Regression
α	0.4	0.4	0.4	0.4
Timestep	$N = 4, 8, \dots, 20$	$N = 4, 8, \dots, 20$	$N = 2^1, \dots, 2^6$	$N = 2^2, \dots, 2^8$

Table 3.1: Summary of parameter choices corresponding to Example 3.1

Example 3.2 (Utility example). *Consider a world with three different type of assets, including a tradable risk-free bond with zero interest rate, a tradable financial index P_t and a set of observable but non-tradable assets $S = (S^1, S^2)$. For all $i \in \{1, 2\}$, their dynamics are given by*

$$\begin{aligned} P_t &= P_0 + \int_0^t \mu^P P_s \, ds + \int_0^t (\sigma^P P_s) \cdot dB_s, \\ &= P_0 + \int_0^t \mu^P P_s \, ds + \int_0^t P_s \tilde{\sigma}^P \cdot \Lambda \cdot dB_s, \\ S_t^1 &= S_0^1 + \int_0^t r_1 S_s^1 \, ds + \int_0^t S_s^1 \tilde{\sigma}^1 \, dB_s^1, \end{aligned}$$

$$S_t^2 = S_0^2 + \int_0^t r_2 S_s^2 ds + \int_0^t S_s^2 \tilde{\sigma}^2 \left(\rho dB_s^1 + \sqrt{1 - \rho^2} dB_s^2 \right),$$

where $\Lambda = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1 - \rho^2} \end{pmatrix}$.

Let $\Phi(S) = \left(K - \frac{1}{2T} \int_0^T S_u^1 + S_u^2 du \right)$, which is the payoff of a claim written on S and (Y^λ, Z^λ) is the pair which solves

$$Y_t^\lambda = \lambda \Phi(S) + \int_0^T \theta_s ds + \int_t^T f(s, Z_s) ds - \int_t^T (Z_s^\lambda) \cdot dB_s,$$

$$f(s, z) = -\frac{\gamma}{2} \|z\|^2 + \frac{\gamma}{2 \|\sigma_s^P\|^2} \left| (\sigma_s^P) \cdot z - \frac{\mu_s^P}{\gamma} \right|^2 - \theta_s, \quad \theta_s = \frac{|\mu_s^P|^2}{2\gamma \|\sigma_s^P\|^2}.$$

Suppose the claims pays $\Phi(S)$ at time $T > 0$, the utility indifference price of such a claim at time t , denoted by Ψ_t^λ , is the difference $Y_t^\lambda - Y_t^0$. The hedging strategy Π_t^λ , which demonstrates the value of extra investments due to purchasing λ -units of the derivative at time t , is $-\frac{(\sigma_t^P) \cdot (Z_t^\lambda - Z_t^0)}{\|\sigma_t^P\|^2}$.

We remark that the driver in the Example 3.2 is Lipschitz when there is only one asset S and quadratic in Z when $d \geq 2$. For further details on utility indifference pricing when the utility is an exponential, we refer to [40]. The relevant parameters that are used in testing with Example 3.1 and 3.2 are listed in Table 3.1 and Table 3.2 respectively.

Both quantization and regression methods are implemented for the numerical scheme for QBSDE. Moreover, we set the partition $\pi : 0 = t_0 < \dots < t_N = T$ to be uniform i.e. $t_j = \frac{jT}{N}$ for all $0 \leq j \leq N$. We will plot the error for computing (Y_0, Z_0) when the explicit solution is known.

3.6.1 Quantization

Two tests have been performed to validate the quantization scheme. In both scenarios, we consider Example 3.1 and pick $M = 1000$ as the number of supported points¹⁰ in the discrete random variable that serves as the replacement to the 2-dimensional Gaussian random variable. Moreover, we set $\kappa\delta = \hat{\kappa}\hat{\delta} = \frac{2}{3}$ and choose $r = \frac{9}{4}$.

Setting $(X_0^1, X_0^2) = (1, 1.2)$ with $(r_1, r_2) = (0.1, 0.05)$, $\sigma = (0.5, 0.4)$ and $\rho = 0.6$ for the forward SDE, the terminal condition ξ of the BSDE includes a sin-type payoff

¹⁰The optimal quantization grids for multi-dimensional Gaussian random variables are listed on http://www.quantize.maths-fi.com/gaussian_database, without extra costs. However, for multi-dimensional random variables, the size of quantization grids is less than 1450.

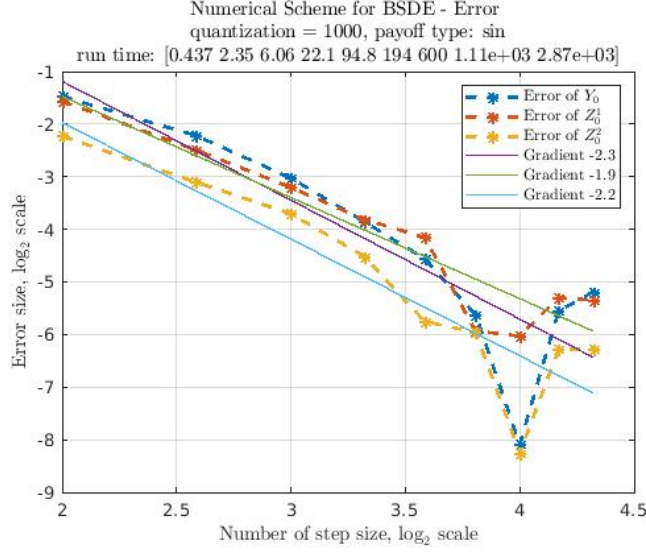


Figure 3.1: Error analysis for *sin* type

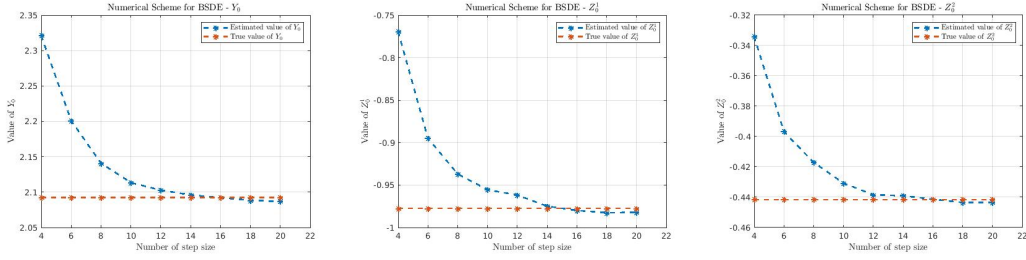


Figure 3.2: Estimation of Y_0, Z_0^1, Z_0^2 under sin-type payoff using quantization methods

with $\hat{a} = 2$, $T = 1$ and depends on the integral of the trajectory. The truncation parameter is $\alpha = 0.4$, which satisfies the requirement, leading us to the convergence rate of roughly $O(N^{-2})$ in (Y, Z) , as in the $\log_2 - \log_2$ plot (Figure 3.1). The result of our experiment over-performs the predicted rate, $O(N^{-\frac{1}{2}})$. As the terminal condition is a smooth functional and belongs to $\mathbb{D}^\infty$, it might be possible to apply Taylor expansion carefully to obtain a sharper rate.

A maximum of 20 timesteps is used when we carried out such experiment, since the computational time rises from 0.4 seconds to 2870 seconds, as we increase the number of timesteps N from 4 to 20. The rise in computational time is attributed to the rise in the number of points in the space discretisation. Figure 3.1 suggests the error starts increasing from 2^{-5} after the $N = 16$ instead of reducing as in the theory. However, this phenomenon is caused by the quantization error, which is of order $O(\frac{1}{\sqrt{1000}})$. When $N = 16$, the quantization error accidentally offsets the discretisation error, which improves the final precision.

Now we consider the same setting as the last example, except $\rho = 0.4$, $T = \frac{1}{2}$, $a =$

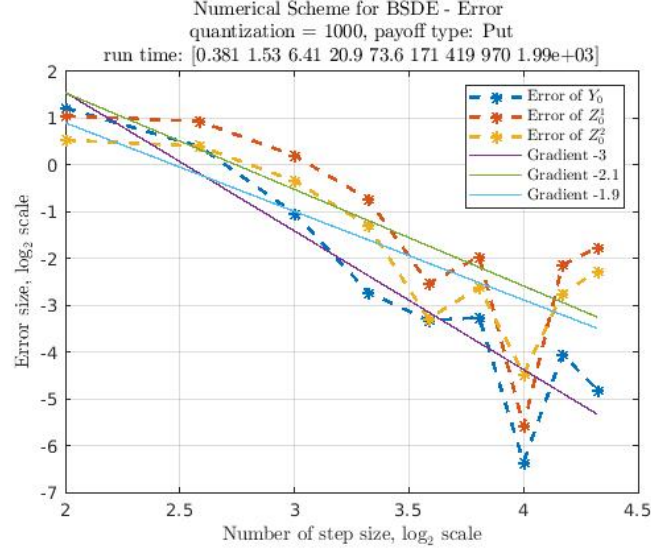


Figure 3.3: Error analysis for Lipschitz payoff, log-log

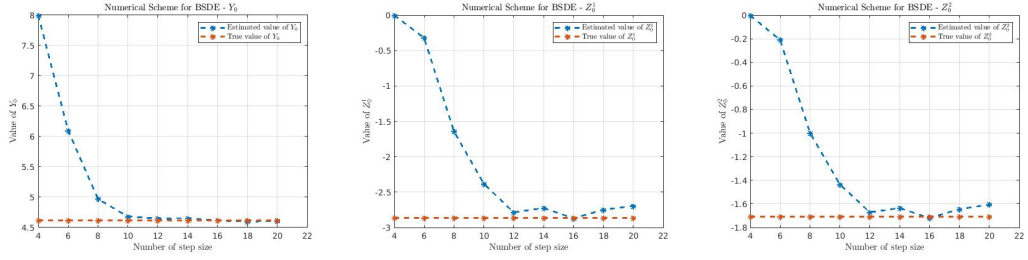


Figure 3.4: Estimation of Y_0, Z_0^1, Z_0^2 under Put type payoff using quantization methods

-0.2 , $X_0 = (5, 6)$ and the terminal data is a Put payoff with an integral. Fixing $K = 8$, the convergence is even faster than the *sin* example, but the fluctuation appears in approximating the Z -component when N is large. The variance of X with $X_0 = (5, 6)$, $\sigma = (0, 4, 0.3)$ is significantly larger in contrast to the first example, therefore we need to trade-off between capturing the heavy tail of X and reducing the gap within the grid. A finer grid is more efficient when the number of timesteps is large, as $\kappa\delta = N^{\frac{2}{3}}$ is sufficient to cover most of the density of X . On the contrary, the tails will contribute a significant error, when N is small.

3.6.2 Regression basis

The regression of basis method, which is widely accepted in this area (see [61, 91]), provides an alternative approach to approximate the conditional expectation. The overarching idea is based on the fact that the conditional expectation $\mathbb{E}[F | \mathcal{G}]$ where $F \in L^2$ and $\mathcal{G} \subseteq \mathcal{F}$ is sub- σ -algebra, is equivalent to the projection of F in

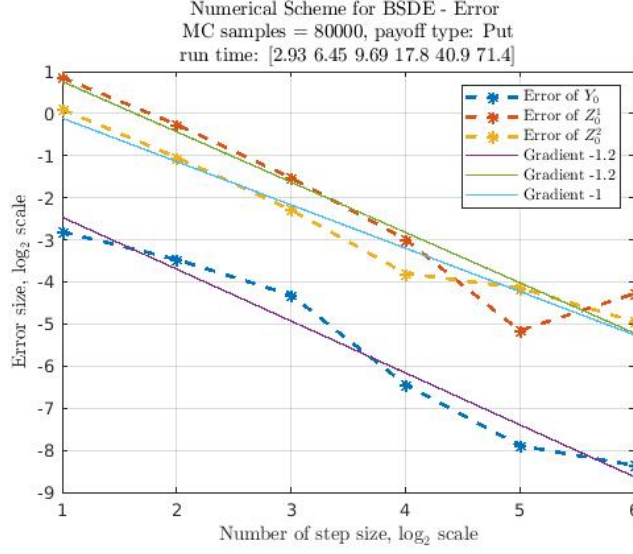


Figure 3.5: Error analysis for (Y_0, Z_0) with Asian Put type payoff

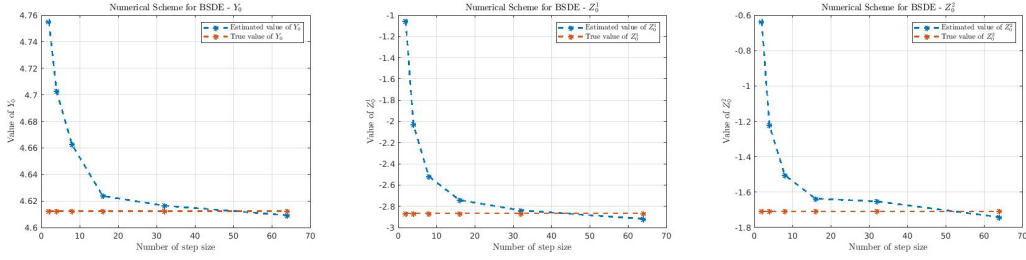


Figure 3.6: Estimation of Y_0, Z_0^1, Z_0^2 for Asian Put type payoff, using regression techniques

$L^2(\Omega, \mathcal{G}, \mathbb{P})$. If there exists a basis $\{e_j\}_j$ for $L^2(\Omega, \mathcal{G}, \mathbb{P})$, the conditional expectation $\mathbb{E}[F | \mathcal{G}]$ is a linear combination of the basis $\{e_j\}_j$, which minimises $\|F - \mathbb{E}[F | \mathcal{G}]\|_{L^2}$. Replacing the expectation operator $\mathbb{E}$ by the average of M Monte-Carlo samples, a numerical scheme is obtained accordingly. To compute the conditional expectation, the coefficients associated with the basis is deduced by regression, which minimises the L^2 -error for samples $1, \dots, M$. A recent development in this technique is [1], where the author proposed two correction methods to reduce the variance of estimating the Z -component in the context of smooth Markovian BSDEs.

The quantities (Y_i^π, Z_i^π) are both adapted to the filtration $\mathcal{F}_{t_i}$, which suggests that a desirable basis must be able to span $L^2(\Omega, \mathcal{F}_{t_i}, \mathbb{P})$. Verifying the feasibility and the performance of the basis is an obstacle except for a few cases, such as the indicator basis which discretises $\mathbb{R}^{m+1}$ by cubes. The polynomial basis, which is a popular choice, is used in this subsection and we assume that the error due to the basis is negligible compared to other forms of the error. Throughout this subsection,

we fix $M = 80000$ as the number of sample paths, such that the Monte-Carlo error is of order $O(2^{-8})$.

We test the regression method with the second example that is used in the quantization technique: $X_0 = (5, 6)$, $r = (0.1, 0.05)$, $\sigma = (0.4, 0.3)$, $\rho = 0.4$, $a = -0.2$, $\hat{a} = 2$, $T = \frac{1}{2}$ and $K = 8$. To obtain an error of $O(2^{-6})$ in the Y component, the regression method requires 17.8 seconds as shown in Figure 3.5, compared to a minimum of 970 seconds needed in the quantization method. There is also an advantage in approximating the Z component in terms of computational time.

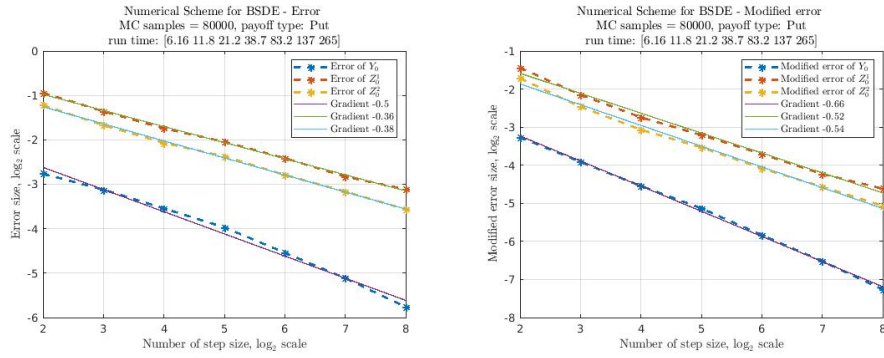


Figure 3.7: Error analysis for (Y_0, Z_0) with lookback Put type payoff

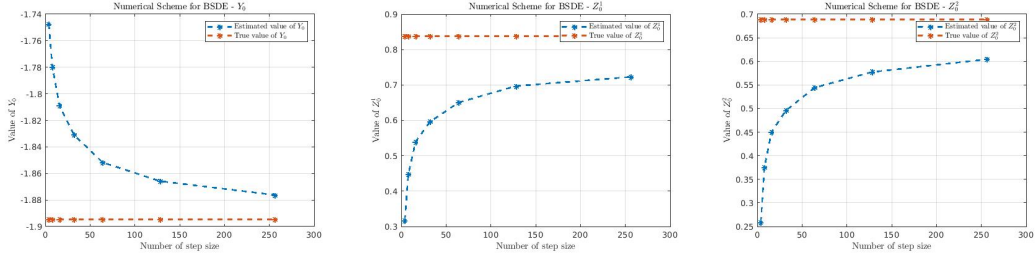


Figure 3.8: Estimation of Y_0, Z_0^1, Z_0^2 with lookback Put type payoff, with regression techniques

Setting $X_0 = (1, 1.2)$, $r = (0.2, 0.4)$, $\sigma = (0.6, 0.3)$, $\rho = -0.3$, $a = 2$, $\hat{a} = -2$, $T = \frac{1}{2}$ and $K = 3$, we move to the next scenario, which is a Put-payoff associating with the infimum of the forward process. Figure 3.7 includes the $\log_2$ - $\log_2$ plot for the error and the modified error against the number of timesteps N . By Theorem 3.1.3, the error of Y is controlled by $O(N^{-\frac{1}{2}}(\log(N))^{\frac{1}{2}})$, indicating the $\log_2$ error is of the form $-\frac{1}{2}\log_2(N) + \frac{1}{2}\log_2(\log_2(N)) + C$ for some constant C . If we naively plot the $\log_2$ error against the time-step, convergence rate in Y is $O(N^{-0.4})$. Modifying the $\log_2$ error by taking away the $\frac{1}{2}\log_2(\log_2(N))$ term, we get the $\log_2$ modified error¹¹, which leads

¹¹The modified error does not reflect the true error. The aim of introducing this term is to verify the order $O(N^{-\frac{1}{2}}(\log(N))^{\frac{1}{2}})$.

us to the correct plot and coincides with our theory. The drop in convergence rate is due to the drop in the regularity of ξ . As we expected, approximating the solution (Y, Z) with a terminal condition being Lipschitz with respect to $\|\cdot\|_{\mathbb{D}^\infty(\mathbb{R}^m)}$ is more challenging than estimating the one which is Lipschitz with respect to $\|\cdot\|_{\mathbb{D}^1(\mathbb{R}^m)}$.

λ	1, 2, 3	μ^p	0.15	$(\tilde{\sigma}_1^p, \tilde{\sigma}_2^p)$	(0.6, 0.15)
(S_0^1, S_0^2)	(10, 14)	(r_1, r_2)	(0.05, 0.03)	$(\tilde{\sigma}_1, \tilde{\sigma}_2)$	(0.4, 0.3)
ρ	0.4	γ	1.2	T	$\frac{1}{2}$
K	14	M	80000	α	0.45

Table 3.2: Summary of parameter choices corresponding to Example 3.2

As an application to the utility indifference pricing theory, we consider Example 3.2 with parameters listed in Table 3.2. To evaluate the utility indifference price, two numerical experiments need to be conducted to evaluate Y_t^λ and Y_t^0 respectively, hence the computational cost is doubled¹². Figure 3.9 demonstrates the impact of discretising the interval into finer time intervals. In order to simulate the price for 2-units of the claim, the computational cost rises from 20.7 seconds to 1070 seconds as we boost the number of intervals $N = 2^3, 2^4, \dots, 2^9$. However, we observe $N = 128$ is sufficient to estimate the value of the 2 units of the claim.

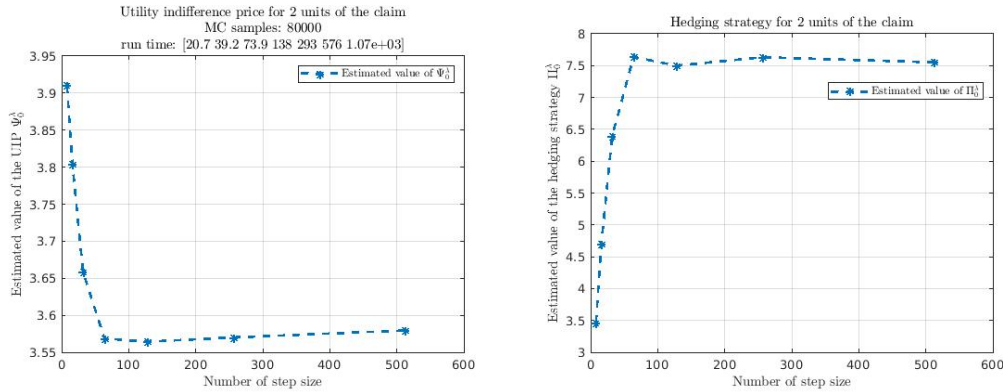


Figure 3.9: Utility indifference pricing and hedging strategy for 2 units of the claim

Table 3.3 compares the utility indifference price and the hedging strategy¹³ for 1, 2, 3 units of the claim. The price for a European derivative with terminal payoff $\lambda\Phi(S)$ in the classical Black-Scholes theory is linear in the λ , since the price is a

¹²The computational time can be reduced by parallel computing.

¹³See Example 3.2

solution to a linear PDE. On the contrary, the utility indifference price is non-linear in λ , while the price difference of an extra claim is decreasing with respect to λ .

Units of Claim	1	2	3
Utility indifference price	1.97	3.58	4.73
Hedging strategy	4.1	7.6	9.8

Table 3.3: Hedging strategy and utility indifference price of Example 3.2

Finally, we consider the difference between explicit and implicit scheme. Similar to the choice corresponding to Figure 3.7, we inherit most of the parameters, except choosing $f(t, y, z) = 2 \max\{y + |z - 1|^2, 3\} + \frac{a}{2}|z|^2$ and set $\hat{a} = 1$, $\rho = -0.6$. We still observe a reduction in the computational time, as in Table 3.4, even though most of the computational power is devoted to computing conditional expectations.

N	2	4	8	16	32
Time saved in the explicit scheme (s)	0.1043	0.1756	0.1375	0.4398	0.7327

Table 3.4: Difference of computational time of the explicit and implicit scheme

Chapter 4

On distributions of velocity random fields in turbulent flows

4.1 Introduction

In the past decades, the probability density function (PDF) method, based on the transport equation, a formal adjoint equation of the Navier-Stokes equation, has been developed into a useful tool for modelling turbulent flows. This method focuses on evaluating the one-point one-time PDF $p(u; x, t)$ of the velocity field $U(x, t)$ or equivalently the centred field $U(x, t) - \overline{U(x, t)}$. The exact transport equation for the PDF, which involves the mean of the pressure term as well as the conditional expectation of the pressure term, has been derived by Pope and can be found in [76, 77] for details. However, only few features can be extracted from the formal PDF transport equation for the purpose of modelling turbulent flows. Therefore, applications of the PDF method have been based on the generalised Langevin model, where the time-dependent velocity $U(t)$ of a particle at position $X(t)$ is assumed to satisfy a stochastic differential equation (SDE).

The main contribution in this chapter is the derivation of the PDF partial differential equation (PDE) for the velocity random field which is much more explicit than the formal PDF transport equation. This PDE is a generalisation of Reynolds mean flow equations, which can be closed by introducing Reynolds stress tensor field. Having 6 dimensions in space (u, x) and 1 dimension in time t , our PDF PDE can be regarded as a parabolic-transport equation which has a parabolic operator $\frac{1}{2}\partial_t + u \cdot \nabla_x + \nu \Delta_x$ in x and transport operator $\frac{1}{2}\partial_t + \nabla_u \cdot$ in u . However, the PDF PDE for velocity fields is a second order partial differential equation which is in general not parabolic due to the appearance of a mixed derivative term $\partial_{u^i} \partial_{x^k} p$. Even this mixed derivative does not appear in the PDF PDE (which is the case for some turbulent flows which will be

explained below), the parabolic part in the PDF PDE involves only the variable x , and therefore even for this case the PDF PDE is highly degenerate. This feature of the PDF PDE distinguishes itself from the prevalent parabolic PDEs or other types of PDE theories in the literature.

The PDF PDE that we have obtained, relies on two conditional structure functions, which are the conditional average increment

$$\rho^i(x, y, u, t) = \mathbb{E} [U^i(y, t) - U^i(x, t) | U(x, t) = u]$$

and the conditional covariance

$$\sigma^{ij}(x, y, u, t) = \mathbb{E} [U^i(y, t)U^j(y, t) | U(x, t) = u].$$

These conditional structure functions describe the interactions of the velocity random field at different positions, hence they are natural to appear in the PDF PDE. The fact that the distribution of velocity random fields is characterised by the conditional first and second moments is an interesting feature revealed in this Chapter. These statistical characteristics are local, which have the capacity to determine the PDF PDE. Moreover, these local statistical characteristics can localise many concepts, such as homogeneity, isotropy etc, which were introduced firstly by Taylor and Kolmogorov [53, 54, 85], allowing us to generalise such concepts to their weak versions.

This chapter, which is based on [57], has the following structure. In Section 4.2, we introduce definitions related to random fields, which are cornerstones of our main results. The evolution equation for the distribution of the velocity random field of turbulence over time will be derived, under the assumption that the random field is regular. The PDE is going to be applied to various types of flows in Section 4.3, including both the viscous and inviscid cases. We also obtain a stochastic representation formula for the solution of the PDF PDE, together with the constraint that ensures the solution is indeed a PDF for all time $t \geq 0$ and position $x \in \mathbb{R}^3$. These theoretical results are important when we apply the PDF PDE for modelling turbulent flows. Section 4.4 is thus devoted to an example of modelling the PDF of turbulence, which has the ability to demonstrate the change of distribution at a fixed position x over time. This Chapter will be closed by a few remarks in the last section.

Conventions on notation. The following set of conventions is employed throughout the Chapter. Firstly Einstein's convention on summation on repeated indices through their ranges is assumed, unless otherwise specified. If A is a vector or a vector field (usually in the space of dimension three) dependent on some parameters, then its components are labelled with upper-script indices so that $A := (A^i) = (A^1, A^2, A^3)$.

The same convention applies to coordinates too. The derivative operators ∇ and Δ are labelled with subscripts to indicate the variable to which the operator is applied, such as $\Delta_x := \partial_{x^i} \partial_{x^i}$ and $\nabla_x \cdot A := \partial_{x^i} A^i$. Finally, the velocity vector field will be denoted by $U = (U^i)$, unless we specified.

4.2 PDF equation of velocity fields

In this section, we aim to introduce some fundamental concepts on random fields and to derive the evolution equation for the random velocity field $\{U(x, t)\}_{x,t}$, where $x \in \mathbb{R}^3$, $t \geq 0$ and $U(x, t)$ takes values in $\mathbb{R}^3$.

4.2.1 Random fields and their statistical characteristics

Given a random field $\{U(x, t)\}_{x,t}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, $U(x, t)$ is, by definition, an $\mathbb{R}^3$ -valued random variable for every $x \in \mathbb{R}^3$ and $t \geq 0$. The most important statistical property for understanding a random field is the correlation function of two random variables, which plays the dominant role in the study of turbulence [3, 68]. In this Chapter, we however emphasize the use of a few statistical characteristics based on the conditional distribution. Let us introduce these statistical characteristics, which we believe are of the most importance.

Definition 4.2.1. *Given a time-dependent random field $\{U(x, t)\}_{x,t}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, for $x, y, u \in \mathbb{R}^3$ and $t \geq 0$, and $i \in \{1, 2, 3\}$,*

1) *the conditional average increment function ρ^i is defined as*

$$\rho^i(x, y, u, t) = \mathbb{E} [U^i(y, t) - U^i(x, t) | U(x, t) = u], \quad (4.1)$$

2) *the conditional covariance function $\sigma(x, y, u, t)$ is defined to be the covariance of $U(y, t) - U(x, t)$ given $U(x, t) = u$,*

$$\sigma^{ij}(x, y, u, t) = \text{cov} [(U^i(y, t), U^j(y, t)) | U(x, t) = u]. \quad (4.2)$$

From the definition, it is clear that for every i , the *conditional mean function* is of the form

$$b^i(x, y, u, t) := \mathbb{E} [U^i(y, t) | U(x, t) = u] = \rho^i(x, y, u, t) + u^i \quad (4.3)$$

and $\rho(x, x, u, t) = 0$ for all u, x and t . As the Reynolds stress is given by

$$r^{ij}(x, t) := \mathbb{E} [(U^i(x, t) - \mathbb{E}[U^i(x, t)])(U^j(x, t) - \mathbb{E}[U^j(x, t)])]$$

the conditional covariance function $\sigma^{ij}(x, y, u, t)$ can be treated as the conditional Reynolds stress. These statistical characteristics have explicit representations in terms of the two-point joint distribution. For our purpose, we assume that the distribution of $U(x, t)$ for every (x, t) has a probability density function (PDF) with respect to Lebesgue measure on $\mathbb{R}^3$, denoted by $p(u; x, t)$, in the sense that

$$\mathbb{E} [\mathbb{1}_{\{U(x, t) \in E\}}] = \int_E p(u; x, t) du \quad \text{for any Borel set } E \in \mathcal{B}(\mathbb{R}^3).$$

Similarly, the joint distribution of $U(x, t)$ and $U(y, t)$ at two distinct points $x \neq y$ has a joint PDF, denoted by $p_2(u_1, u_2; x, y, t)$. It follows that the conditional law of $U(y, t)$ given that $U(x, t) = u$ possesses the following conditional PDF

$$p_{2|1}(v; u, x, y, t) := \mathbb{P}(U(y, t) = v \mid U(x, t) = u) = \frac{p_2(u, v; x, y, t)}{p(u; x, t)}$$

with $p_{2|1}(v; u, x, y, t) = 0$ if $p(u; x, t) = 0$. In terms of the conditional law, the joint PDF of $U(x, t)$ and $U(y, t)$ may be split into a product

$$\begin{aligned} p_2(u_1, u_2; x, y, t) &= p(u_1; x, t) p_{2|1}(u_2; u_1, x, y, t) \\ &= p(u_2; y, t) p_{2|1}(u_1; u_2, y, x, t). \end{aligned}$$

As a result, we are allowed to represent the conditional average difference function (4.1) and covariance function (4.2) as an integral of the conditional density, namely

$$\rho^i(x, y, u, t) = \int_{\mathbb{R}^3} (v^i - u^i) p_{2|1}(v; u, x, y, t) dv \quad (4.4)$$

and

$$\sigma^{ij}(x, y, u, t) = \int_{\mathbb{R}^3} (v^i - u^i) (v^j - u^j) p_{2|1}(v; u, x, y, t) dv.$$

The use of the conditioning techniques is in fact the main reason for advocating the foundation of statistical fluid mechanics based on the probability theory, rather than on an average procedure, which was first explicitly proposed by Kolmogorov [53, 54]. The homogeneity and the isotropy can be defined in general for random fields indexed by a space variable $x \in \mathbb{R}^3$, which has been introduced into the study of turbulence by G. I. Taylor. The local homogeneous and local isotropic flows were introduced by Kolmogorov for formulating K41 theory (and its improved version K61 theory). According to Kolmogorov [53, 54], a random field $\{U(x, t)\}_{x, t}$ is locally homogeneous if for any $x, y \in \mathbb{R}^3$, the conditional distribution of $U(y, t) - U(x, s)$ given $U(x, s) = u$ depends on $y - x$ and u , and further it is locally isotropic if the conditional distribution depends only on $|y - x|$ and u . By using the conditional

average and the conditional covariance functions, it is possible to generalise these terminologies to their weak versions. We are now in a position to state technical assumptions on the random field.

Definition 4.2.2. *The random field $\{U(x, t)\}_{x, t}$ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is*

- 1) *regular if the conditional average increment function ρ has derivatives up to second order and $\rho(x, y, u, t)$ has a Taylor expansion (for every x, u and t fixed) about y :*

$$\rho(x, y, u, t) = B_k(x, u, t)(y^k - x^k) + \frac{1}{2}A_{jk}(x, u, t)(y^j - x^j)(y^k - x^k) + o(|y - x|^2) \quad (4.5)$$

as $|y - x| \rightarrow 0$, where

$$B_k(x, u, t) = \left. \frac{\partial}{\partial y^k} \rho(x, y, u, t) \right|_{y=x}, \quad A_{jk}(x, u, t) = \left. \frac{\partial^2}{\partial y^j \partial y^k} \rho(x, y, u, t) \right|_{y=x}, \quad (4.6)$$

and B and A are differentiable in x^i, u^i, t for all $i \in \{1, 2, 3\}$;

- 2) *weakly homogeneous if given $u \in \mathbb{R}^3, t \geq 0$, $\rho^i(x, y, u, t) = o(|y - x|)$ as $|y - x| \rightarrow 0$ for all i, y, x ;*
- 3) *weakly isotropic if both $\rho(x, y, u, t)$ and $\sigma(x, y, u, t)$ depend only on $|y - x|$, u and t , and A_{kk}^i only depends on u and t .*

The functions B_k^i and A_{jk}^i in the Taylor expansion of the conditional average increment $\rho^i(x, y, u, t)$ also have the form

$$\begin{cases} B_k^i(x, u, t) = \mathbb{E} \left[\frac{\partial}{\partial x^k} U^i(x, t) \mid U(x, t) = u \right], \\ A_{jk}^i(x, u, t) = \mathbb{E} \left[\frac{\partial^2}{\partial x^j \partial x^k} U^i(x, t) \mid U(x, t) = u \right] \end{cases}$$

for $i, j, k \in \{1, 2, 3\}$. Moreover, if the $\{U(x, t)\}_{x, t}$ is weakly homogeneous, we have the equivalent characterisation

$$B_k^i(x, u, t) = \lim_{\varepsilon \rightarrow 0} \frac{\rho^i(x, x + \varepsilon e^{(k)}, u, t)}{\varepsilon} = 0$$

for all $i, k \in \{1, 2, 3\}$, where $e^{(k)}$ is the unit vector at the k -th direction.

Unlike Kolmogorov's definition of isotropic flows and homogeneous flows, our concept of weak isotropy has no direct relationship to weak homogeneity. Nevertheless, a regular locally homogeneous turbulent flow in the sense of Kolmogorov satisfies the

condition that A_{kk}^i depends only on u , since the conditional average increment of such a flow must obey $\rho^i(x, y, u, t) = g^i(y - x, u)$ for some function g^i and

$$A_{jk}^i(x, u, t) = \frac{\partial^2 g^i}{\partial y^j \partial y^k}(0, u).$$

Moreover, if we further assume the flow is locally isotropic, the conditional average increment function satisfies $\rho^i(x, y, u, t) = g^i(|y - x|, u)$ and

$$B_k^i(x, u, t) = \lim_{\varepsilon \rightarrow 0^+} \frac{\rho^i(x, x + \varepsilon e^{(k)}, u, t)}{\varepsilon} = \lim_{\varepsilon \rightarrow 0^+} \frac{-\rho^i(x, x - \varepsilon e^{(k)}, u, t)}{\varepsilon}$$

is well-defined if and only if $B_k^i = 0$. Therefore, this turbulent flow is both weakly homogeneous and weakly isotropic in our sense.

Apart from extending Kolmogorov's definitions on homogeneity and isotropy, the significance of introducing these terminologies is they will eliminate the mixed-terms in the PDF PDE, which will be thoroughly explained in section 4.3.

4.2.2 The evolution equation for the velocity distribution

In this subsection, we derive the main theoretical result, which provides the theoretical foundation of modelling PDFs of turbulent flows based on two statistical characteristics. We consider an incompressible turbulent flow, inviscid or viscous, with kinetic viscosity constant ν which is positive for viscous fluid, or ν reads as zero for inviscid fluid.

We assume that $\{U(x, t)\}_{x,t}$ is a regular random field. Since $U(x, t)$ is divergence-free as in equation (2.14), we have for all i, k ,

$$\frac{\partial}{\partial y^i} \rho^i(x, y, u, t) = 0, \quad B_i^i = 0, \quad A_{ik}^i = A_{ki}^i = 0,$$

as well as the following integral condition for PDF of $U(x, t)$

$$\frac{\partial}{\partial x^i} \int_{\mathbb{R}^3} u^i p(u; x, t) du = 0, \quad (4.7)$$

which will appear as a natural constraint for the PDF PDE we will derive.

Recall that the Reynolds equation (see [80]) is obtained by taking the average in (2.13), more explicitly

$$\begin{aligned} \frac{\partial \mathbb{E}[U^i]}{\partial t} + \mathbb{E} \left[U^j \frac{\partial U^i}{\partial x^j} \right] &= \nu \Delta_x \mathbb{E}[U^i] - \frac{\partial \mathbb{E}[P]}{\partial x^i}, \\ \frac{\partial \mathbb{E}[U^i]}{\partial x^i} &= 0. \end{aligned}$$

The conventional treatment for the non-linear term on the left-hand side is to write

$$\mathbb{E} \left[U^j \frac{\partial U^i}{\partial x^j} \right] = \mathbb{E} [U^j] \frac{\partial \mathbb{E} [U^i]}{\partial x^j} + \frac{\partial}{\partial x^i} r^{ij}.$$

The PDF equation can be obtained by carrying out this computation for the average $\mathbb{E} [F(U)]$ where $F : \mathbb{R}^3 \rightarrow \mathbb{R}$ is set as a smooth function with a compact support, instead of choosing $F(x) = x$ for each velocity component in the case of Reynold stress. We are now in a position to establish the most important work in this Chapter:

Theorem 4.2.3. *Let $\{U(x, t)\}_{(x, t) \in \mathbb{R}^3 \times \mathbb{R}_+}$ be a regular random field defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and take values in $\mathbb{R}^3$. Suppose $\{U(x, t)\}_{x, t}$ satisfies the Navier-Stokes equation (2.13) and the divergence-free condition (2.14), then the PDF $p(u; x, t)$ of the velocity $U(x, t)$ satisfies the evolution equation:*

$$\begin{aligned} \frac{\partial p}{\partial t} + u^i \frac{\partial p}{\partial x^i} &= \nu \Delta_x p + \frac{\partial}{\partial u^i} \left(\nu \frac{\partial (p B_k^i)}{\partial x^k} - \nu p A^i + p Q^i \right), \\ p(u; x, 0) &= p_0(u; x), \end{aligned} \quad (4.8)$$

where $p_0(u; x)$ is the PDF of the $U_0(x)$ in the random field $\{U_0(x)\}_x$ and

$$Q^i(x, u, t) = \int_{\mathbb{R}^3} \frac{y^i - x^i}{4\pi|y - x|^3} \frac{\partial^2 (\sigma^{jk} + b^j b^k)}{\partial y^k \partial y^j} dy, \quad (4.9)$$

$$B_k^i = \left. \frac{\partial \rho^i(x, y, u, t)}{\partial y^k} \right|_{y=x}, \text{ and } A^i(x, u, t) = \left. \Delta_y \rho^i(x, y, u, t) \right|_{y=x} = A_{kk}^i(x, u, t). \quad (4.10)$$

Proof. Let $F \in D(\mathbb{R}^3)$ be a test function, which is a smooth function taking values in $\mathbb{R}$ with a compact support. For simplicity, we denote $F_i(u) := \partial_{u^i} F(u) \in D(\mathbb{R}^3)$. Applying ∂_t on the average $\mathbb{E} [F(u)]$ followed by exchanging integral and their derivative operator, we have

$$\frac{d}{dt} \mathbb{E} [F(U(x, t))] = \int F(u) \frac{\partial}{\partial t} p(u; x, t) du.$$

On the other hand

$$\frac{d}{dt} \mathbb{E} [F(U(x, t))] = \mathbb{E} \left[\frac{\partial}{\partial t} F(U(x, t)) \right] \quad (4.11)$$

together with utilizing the Navier-Stokes equations (2.13), we get

$$\left(\frac{\partial}{\partial t} - \nu \Delta_x \right) F(U) = - \frac{\partial (U^i F(U))}{\partial x^i} - \nu \frac{\partial F_i(U)}{\partial x^k} \frac{\partial U^i}{\partial x^k} - F_i(U) \frac{\partial P}{\partial x^i}.$$

Substituting this into (4.11), we obtain that

$$\mathbb{E} \left[\frac{\partial}{\partial t} F(U) \right] = \nu \Delta_x \mathbb{E} [F(U)] - \frac{\partial}{\partial x^i} \mathbb{E} [U^i F(U)] + J_1 + J_2,$$

where the first two terms are equivalent to

$$\nu \Delta_x \mathbb{E} [F(U)] = \int_{\mathbb{R}^3} F(u) \nu \Delta_x p(u; x, t) \, du,$$

and

$$-\frac{\partial}{\partial x^i} \mathbb{E} [U^i F(U)] = \int_{\mathbb{R}^3} F(u) \left(-u^i \frac{\partial}{\partial x^i} p(u; x, t) \right) \, du.$$

Subsequently, the remaining J_1, J_2 are of the form

$$J_1 = -\nu \mathbb{E} \left[\frac{\partial F_i(U)}{\partial x^k} \frac{\partial U^i}{\partial x^k} \right], \quad J_2 = -\mathbb{E} \left[F_i(U) \frac{\partial P}{\partial x^i} \right].$$

The evaluation of J_1 and J_2 requires invoking the joint distribution at two points together with taking limits. Here we depart from this approach by expressing this term via the PDF, which allows us to perform similar computations for a general case. The partial derivative $\frac{\partial U^i}{\partial x^j}$ may be written as a limit

$$U^j(x, t) \frac{\partial}{\partial x^j} U^i(x, t) = \lim_{h \rightarrow 0} \frac{1}{h} U^j(x, t) (U^i(x + h e^{(j)}, t) - U^i(x, t)).$$

Assuming that we are able to take the average under the limit i.e. the dominated convergence theorem can be applied, we are able to write the non-linear term in terms of

$$\mathbb{E} \left[U^j \frac{\partial U^i}{\partial x^j} \right] = \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E} [U^j(x, t) (U^i(x + h e^{(j)}, t) - U^i(x, t))] =: \lim_{h \rightarrow 0} L^i(h).$$

The average appearing on the right-hand side, denoted by $L^i(h)$, may be evaluated in terms of the two-point joint distribution

$$\begin{aligned} L^i(h) &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} u^j u_1^i p_2(u, u_1; x, x + h e^{(j)}, t) \, du_1 \, du - \int u^j u^i p(u; x, t) \, du \\ &= \int u^j p(u; x, t) \left[\int_{\mathbb{R}^3} u_1^i p_{2|1}(u_1; u, x, x + h e^{(j)}, t) \, du_1 \right] \, du \\ &\quad - \int_{\mathbb{R}^3} u^j u^i p(u; x, t) \, du \\ &= \int_{\mathbb{R}^3} u^j p(u; x, t) \left[\int_{\mathbb{R}^3} (u_1^i - u^i) p_{2|1}(u_1; u, x, x + h e^{(j)}, t) \, du_1 \right] \, du \\ &= \int_{\mathbb{R}^3} u^j p(u; x, t) \rho^i(x, x + h e^{(j)}, u, t) \, du, \end{aligned}$$

which follows that

$$\frac{\partial}{\partial x^j} \mathbb{E} [U^j U^i] = \int_{\mathbb{R}^3} u^j B_j^i(x, u, t) p(u; x, t) \, du.$$

Now we deal with J_1 . Writing the space derivatives $\frac{\partial F_i(U)}{\partial x^k} \frac{\partial U^i}{\partial x^k}$ as the following limits

$$\lim_{h \rightarrow 0} \frac{1}{h^2} (F_i(U(x + he^{(k)}, t)) - F_i(U(x, t))) (U^i(x + he^{(k)}, t) - U^i(x, t)),$$

where $e^{(1)} = (1, 0, 0)$ and so on. Taking expectation first, we obtain that

$$\begin{aligned} J_1 &= -\nu \lim_{h \rightarrow 0} \frac{1}{h^2} \sum_{k=1}^3 \int_{\mathbb{R}^3} (F_i(u_2) - F_i(u_1)) (u_2^i - u_1^i) p_2(u_1, u_2; x, x + he^{(k)}, t) du_1 du_2 \\ &= \nu \int_{\mathbb{R}^3} F_i(u) \lim_{h \rightarrow 0} \frac{I^{i,k}(h)}{h^2} du, \end{aligned}$$

where

$$\begin{aligned} I^{i,k}(h) &:= \int_{\mathbb{R}^3} (u_1^i - u^i) p_2(u_1; u, x, x + he^{(k)}, t) du_1 \\ &\quad + \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} (u_1^i - u^i) p_2(u; u_1, x, x + he^{(k)}, t) du_1. \end{aligned}$$

Using the conditional probability notation that we introduced, we may write this as

$$\begin{aligned} I^{i,k}(h) &:= p(u; x + he^{(k)}, t) \int_{\mathbb{R}^3} (u_1^i - u^i) p_{2|1}(u_1; u, x + he^{(k)}, x, t) du_1 \\ &\quad + p(u; x, t) \int_{\mathbb{R}^3} (u_1^i - u^i) p_{2|1}(u_1; u, x, x + he^{(k)}, t) du_1 \\ &= p(u; x + he^{(k)}, t) (\rho^i(x + he^{(k)}, x, u, t) + \rho^i(x, x + he^{(k)}, u, t)) \\ &\quad - (p(u; x + he^{(k)}, t) - p(u; x, t)) \rho^i(x, x + he^{(k)}, u, t). \end{aligned}$$

The last equality is a result of applying (4.4), which converts integrals into conditional average increments ρ^i . As a consequence of the regularity condition on the random field, we make use of (4.5) to deduce

$$\begin{aligned} &\rho^i(x + he^{(k)}, x, u, t) + \rho^i(x, x + he^{(k)}, u, t) \\ &= \left\{ B_k^i(x, u, t)h + \frac{1}{2} A_{kk}^i(x, u, t)h^2 + o(h^2) \right\} \\ &\quad + \left\{ -B_k^i(x + he^{(k)}, u, t)h + \frac{1}{2} A_{kk}^i(x + he^{(k)}, u, t)h^2 + o(h^2) \right\} \\ &= (B_k^i(x, u, t) - B_k^i(x + he^{(k)}, u, t))h + \frac{1}{2} (A_{kk}^i(x, u, t) + A_{kk}^i(x + he^{(k)}, u, t))h^2 \\ &\quad + o(h^2) \end{aligned}$$

and therefore

$$\lim_{h \rightarrow 0} \frac{1}{h^2} I^{i,k}(h) = -p \frac{\partial}{\partial x^k} B_k^i + p A^i - \frac{\partial p}{\partial x^k} B_k^i$$

$$= -\frac{\partial}{\partial x^k}(pB_k^i) + pA^i.$$

We perform integration by parts to derive

$$J_1 = \int_{\mathbb{R}^3} F(u) \left[-\frac{\partial}{\partial u^i} \left(\nu \left[-\frac{\partial}{\partial x^k}(pB_k^i) + pA^i \right] \right) \right] du.$$

Next we handle J_2 . Applying the representation (2.15), we arrive at

$$\frac{\partial P}{\partial x^i} = \int_{\mathbb{R}^3} \frac{y^i - x^i}{4\pi|y - x|^3} \frac{\partial^2(U^j U^k)}{\partial y^k \partial y^j} dy,$$

which implies

$$J_2 = -\mathbb{E} \left[F_i(U) \int_{\mathbb{R}^3} \frac{y^i - x^i}{4\pi|y - x|^3} \frac{\partial^2(U^j U^k)}{\partial y^k \partial y^j} dy \right].$$

Writing the derivative in terms of

$$\begin{aligned} \frac{\partial^2(U^j U^k)}{\partial y^k \partial y^j} = \lim_{h \rightarrow 0} \frac{1}{h^2} & \left\{ U^j(y + h(e^{(k)} + e^{(j)}), t) U^k(y + h(e^{(k)} + e^{(j)}), t) \right. \\ & - U^j(y + h e^{(k)}, t) U^k(y + h e^{(k)}, t) \\ & \left. - U^j(y + h e^{(j)}, t) U^k(y + h e^{(j)}, t) + U^j(y, t) U^k(y, t) \right\} \end{aligned}$$

and using the two-point PDF by integrating then taking limit as $h \rightarrow 0$, lead us to

$$J_2 = - \int_{\mathbb{R}^3} F_i(u) \left[\int_{\mathbb{R}^3} \frac{y^i - x^i}{4\pi|y - x|^3} \frac{\partial^2}{\partial y^k \partial y^j} J_2^{jk} dy \right] du,$$

where the integral J_2^{jk} has the following form

$$\begin{aligned} J_2^{jk} &:= \int_{\mathbb{R}^3} u_1^j u_1^k p_2(u, u_1; x, y, t) du_1 \\ &= p(u; x, t) \int_{\mathbb{R}^3} u_1^j u_1^k p_{2|1}(u_1; u, x, y, t) du_1 \\ &= p(u; x, t) (\sigma^{jk} + b^j b^k)(x, y, u, t), \end{aligned}$$

through using equations (4.1, 4.2). Therefore, substituting this into the equation for J_2 yields

$$\begin{aligned} J_2 &= - \int_{\mathbb{R}^3} F_i(u) p(u; x, t) \int_{\mathbb{R}^3} \frac{y^i - x^i}{4\pi|y - x|^3} \frac{\partial^2 (\sigma^{jk} + b^j b^k)}{\partial y^k \partial y^j} dy du \\ &= \int_{\mathbb{R}^3} F(u) \frac{\partial}{\partial u^i} [p(u; x, t) Q^i(x, u, t)] du. \end{aligned}$$

Putting all terms together, we deduce that

$$\begin{aligned} \int_{\mathbb{R}^3} F(u) & \left(-\frac{\partial}{\partial u^i} \left(\nu \frac{\partial (pB_k^i)}{\partial x^k}(x, u, t) - \nu p(u; x, t) A^i(x, u, t) + p(u; x, t) Q^i(x, u, t) \right) \right. \\ & \left. + (\partial_t + u \cdot \nabla_x - \nu \Delta_x) p(u; x, t) \right) du = 0 \end{aligned}$$

for all such $F \in D(\mathbb{R}^3)$. Therefore, we must have (4.8). $\square$

We finish this section by adding several comments. The PDF PDE (4.8) may be written as

$$\left(\frac{\partial}{\partial t} + u^i \frac{\partial}{\partial x^i} - \nu \Delta_x\right) p = \frac{\partial}{\partial u^i} \left(\nu B_k^i \frac{\partial p}{\partial x^k} + p C^i \right), \quad (4.12)$$

where for simplicity we introduce the following vector field

$$C^i = \nu \frac{\partial B_k^i}{\partial x^k} - \nu A^i + Q^i \quad (4.13)$$

for $i = 1, 2, 3$. The equation (4.12) is a mixed type of parabolic and transport PDE. The parabolic operator in variables (x, t)

$$\frac{\partial}{\partial t} + u^i \frac{\partial}{\partial x^i} - \nu \Delta_x$$

is independent of fluid flows, which is a significant feature.

Although the PDF PDE (4.12) appears to be linear in the PDF $p(u; x, t)$, it is much more complicated than it looks. In particular, the coefficients A, B and Q are functionals of the conditional average and covariance functions, which are in general not determined by the PDF $p(u; x, t)$ alone. Therefore the PDF PDE (4.12) is not a closed partial differential equation. The significance of the PDF PDE lies in the fact that if the statistical characteristics ρ and σ are considered as given, which will be the case for modelling turbulent flows, then the PDF PDE is a partial differential equation of second order, though mixed type of parabolic and transport in general.

Nevertheless, the PDE (4.8) is a challenging obstacle even if A, B, Q are all considered as given. The function B can be understood as the mean velocity gradient at (x, t) conditioning on the velocity vector at (x, t) , which brings the mixed derivatives $\partial_{u^i} \partial_{x^k} p$, while the corresponding diffusion matrix $(D_{ij})_{1 \leq i, j \leq 6}$ collecting the second order terms is of the form

$$D_{ij} = \begin{cases} 0 & 1 \leq i, j \leq 3, \\ \nu B_{j-3}^i & 1 \leq i \leq 3, 4 \leq j \leq 6, \\ \nu B_{i-3}^j & 1 \leq j \leq 3, 4 \leq i \leq 6, \\ \nu \mathbb{1}_{\{j=k\}} & 4 \leq i, j \leq 6, \end{cases}$$

if we consider (u, x) as a whole. The matrix D_{ij} is not necessarily symmetric and not non-negative definite even if it is symmetric. It poses a challenging mathematical problem developing a theory of this kind of mixed type partial differential equations to facilitate the modelling of turbulent flows based on the PDF PDE.

4.3 Application to turbulent flows

As we have seen, our PDE (4.8) does not fit into any existing categories of PDE theories. However, functionals A, B, Q will be derived, if the conditional statistics can be obtained or estimated through practical experiments. Therefore, tracking the PDF of the turbulent flow is equivalent to measuring or modelling the conditional mean and conditional variance, followed by solving the PDF PDE (4.8) using some numerical methods. This brings a new approach to the modelling of turbulent flows.

In this part, we establish some mathematical tools for the purpose of modelling turbulence based on the PDF PDE.

For convenience, let us introduce the following technical assumptions on a function f .

Assumption [A8]. *For a function $f(x, u, t)$ which is uniformly continuous in t , there exist constants $K_1, K_2 \geq 0$ such that*

$$\begin{aligned} |f(x, u, t) - f(y, w, t)| &\leq K_1(|x - y| + |u - w|), \\ |f(x, u, t)| &\leq K_2(1 + |u|), \end{aligned}$$

for all $x, y, u, w \in \mathbb{R}^3$ and $t \geq 0$.

Assumption [A9]. *The function $f(u; x)$ with $u \in \mathbb{R}^3$ and $x \in \mathbb{R}^3$ is continuous and has up to polynomial growth in (u, x) .*

4.3.1 Weakly homogeneous and weakly isotropic flows

When the viscous incompressible flow is weakly homogeneous, the mixed-derivative term disappears and the PDF PDE is simplified to

$$\begin{aligned} \left(\frac{\partial}{\partial t} + u^i \frac{\partial}{\partial x^i} - \nu \Delta_x \right) p &= \frac{\partial}{\partial u^i} (p C^i), \\ p(u; x, 0) &= p_0(u; x), \end{aligned} \tag{4.14}$$

where $C = Q - \nu A$, A and Q are given by equations (4.9) and (4.10) respectively. By the definition of B , a weakly homogeneous flow has the property that the velocity gradient condition on the velocity vector at the same location is a centred random vector. The weak homogeneity allows us to state the representation formula, which provides a useful tool when we model weakly homogeneous turbulent flows.

Theorem 4.3.1. *Given the explicit form of C , we suppose that C satisfies assumption [A8] and $p_0(u; x)$ satisfies assumption [A9] respectively.*

1) Let $p(u; x, t)$ be the solution to equation (4.14). Assuming p has up to polynomial growth in (u, x) and $p \in \mathcal{C}^{1,2,1}(\mathbb{R}^3 \times \mathbb{R}^3 \times [0, T])$ with some fixed $T > 0$, the solution p is unique and possesses the following representation form

$$p(u; x, t) = \mathbb{E} [p_0(Y(t); X(t))q(t)], \quad (4.15)$$

where for any given $t \in [0, T]$ and (x, u) , (X, Y, q) is the unique strong solution to the system of SDEs

$$dX^i(s) = -Y^i(s) ds + \sqrt{2\nu} dM^i(s), \quad X(0) = x,$$

$$dY^i(s) = C^i(X(s), Y(s), t-s) ds, \quad Y(0) = u,$$

and

$$dq(s) = q(s) \frac{\partial C^k}{\partial u^k}(X(s), Y(s), t-s) ds, \quad q(0) = 1,$$

for $i = 1, 2, 3$ and $s \in [0, t]$, associated with $M = (M^1, M^2, M^3)$ being the Brownian motion in $\mathbb{R}^3$ with $M(0) = 0$ defined on some probability space.

2) The equation (4.15) is the unique viscosity solution to the PDF PDE (4.14).

Proof. If C is Lipschitz, the previous system of SDEs for (X, Y) has a unique solution (X, Y) and q is given via the exponential function. Notice (X, Y, q) depends on (x, u) as well. Let $\theta(s) = (Y(s); X(s), t-s)$, $\eta(s) = (X(s), Y(s), t-s)$ and denote $Z(s) = p(\theta(s))q(s)$. According to Itô's formula,

$$\begin{aligned} dZ(s) &= q(s) \frac{\partial p}{\partial u^i}(\theta(s)) dY^i(s) + q(s) \frac{\partial p}{\partial x^i}(\theta(s)) dX^i(s) + q(s) \nu \Delta_x p(\theta(s)) ds \\ &\quad - q(s) \frac{\partial p}{\partial s}(\theta(s)) ds + p(\theta(s)) dq(s) \\ &= \left(\frac{\partial(pC^i)}{\partial u^i}(\eta(s)) - p(\theta(s)) \frac{\partial C^i}{\partial u^i}(\eta(s)) - \frac{\partial p}{\partial x^i}(\theta(s)) Y^i(s) \right. \\ &\quad \left. + \nu(\Delta_x p)(\theta(s)) - \frac{\partial p}{\partial s}(\theta(s)) \right) q(s) ds + \sqrt{2\nu} q(s) \frac{\partial p}{\partial x^i}(\theta(s)) dM^i(s) \\ &= \sqrt{2\nu} q(s) \frac{\partial p}{\partial x^i}(\theta(s)) dM^i(s), \end{aligned}$$

it follows that $(Z(s))_{s \in [0, t]}$ is a local martingale

$$Z(s) = Z(0) + \sqrt{2\nu} \int_0^s q(r) \frac{\partial p}{\partial x^i}(\theta(r)) dM^i(r),$$

which associates with an increasing sequence of stopping times $\{\tau_n\}_{n \geq 0}$

$$\tau_n = \min \{t, \inf \{s \geq 0 : |(Y(s), X(s)) - (u, x)| \geq n\}\}.$$

After taking expectations and utilising the continuity and growth condition of p , we apply the dominated convergence theorem to obtain

$$p(u; x, t) = Z(0) = \lim_{n \rightarrow \infty} \mathbb{E}[Z(\tau_n)] = \mathbb{E}[Z(t)],$$

which coincides with the representation formula.

Regarding the second part 2), we introduce functions $C_\varepsilon^i(x, u, t)$ and $p_{0,\varepsilon}(u; x)$, which are smooth functions (e.g. obtained from convolution with mollifiers) and converge to C, p_0 uniformly on compact sets. We further define the system of non-degenerate SDEs on the time interval $[0, t]$

$$\begin{aligned} dX_\varepsilon^i(s) &= -Y_\varepsilon^i(s) dt + \sqrt{2\nu} dM^i(s), & X_\varepsilon(0) &= x, \\ dY_\varepsilon^i(s) &= C_\varepsilon^i(X_\varepsilon^i(s), Y_\varepsilon^i(s), t-s) ds + \sqrt{\varepsilon} dM^i(s), & Y_\varepsilon(0) &= u, \end{aligned}$$

and bounded process $(q_\varepsilon(s))_{s \in [0, t]}$

$$dq_\varepsilon(s) = q_\varepsilon(s) \frac{\partial C_\varepsilon^k}{\partial u^k}(X(s), Y(s), t-s) ds, \quad q_\varepsilon(0) = 1.$$

Consider the following parabolic PDE

$$\begin{aligned} \left(\frac{\partial}{\partial t} + u^i \frac{\partial}{\partial x^i} - \nu \Delta_x - \varepsilon \Delta_u \right) p_\varepsilon &= \frac{\partial}{\partial u^i} (p_\varepsilon C_\varepsilon^i), \\ p_\varepsilon(u; x, t) &= p_{0,\varepsilon}(u; x), \end{aligned}$$

it admits a unique classical smooth solution p_ε with $|p_\varepsilon|(u; x, t) \leq \gamma_0 \exp(\gamma_1(|u|^2 + |x|^2))$ for some positive constants γ_0, γ_1 by [29, Chapter 1, Theorem 12]. Moreover, the solution possesses the representation

$$p_\varepsilon(u; x, t) = \mathbb{E}[p_{0,\varepsilon}(Y_\varepsilon(t), X_\varepsilon(t))q_\varepsilon(t)],$$

if we make use of 1). Applying Burkholder-Davis-Gundy and Gronwall inequalities and routine arguments for stability of SDEs, we have

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |(Y_\varepsilon(s), X_\varepsilon(s)) - (Y(s), X(s))|^2 \right] \rightarrow 0$$

as $\varepsilon \rightarrow 0$. Therefore, at least through a subsequence, $(Y_\varepsilon(s), X_\varepsilon(s)) \rightarrow (Y(s), X(s))$ almost surely, leading to

$$p_\varepsilon(u; x, t) = \mathbb{E}[p_{0,\varepsilon}(Y_\varepsilon(t), X_\varepsilon(t))q_\varepsilon(t)] \rightarrow \mathbb{E}[p_0(Y(t), X(t))q(t)] = p(u; x, t),$$

uniformly on compact sets. Therefore equation (4.15) is a viscosity solution by Proposition 5.9 in [89], whereas the uniqueness follows from [48]. $\square$

The PDF PDE (4.14) boils down to a degenerate parabolic PDE after the weak homogeneity has been applied. Apart from solving the PDE, the stochastic representation (4.15) offers a route to numerically solving the PDE. The PDE (4.14) has 6 dimensions in space and 1 dimension in time, which is a challenging task for classical finite difference methods due to the size of grid in space. Instead, we can simulate the solution based on Monte-Carlo methods directly.

Remark 4.1. *Moreover, the stochastic representation formula can be extended to non-weakly homogeneous turbulent flows. Indeed,*

$$p(u; x, t) = \mathbb{E} [p_0(Y(t); X(t))q(t)],$$

is a solution to (4.8) subject to the system of SDEs

$$\begin{aligned} dY^i(s) &= C^i(X(s), Y(s), t-s) ds, \quad Y(0) = u, \\ dX^i(s) &= -Y^i(s) ds + \sqrt{2\nu} dM^i(s), \quad X(0) = x, \end{aligned}$$

with $C^i = \nu \frac{\partial B_k^i}{\partial x^k} - \nu A^i + Q^i$ for all $i \in \{1, 2, 3\}$ if we impose the following condition on p :

$$\partial_{u^i} (B^i(x, u, t) \cdot \nabla_x p(u; x, t)) = 0.$$

We are now in a position to verify the solution of our PDF PDE (4.14) is indeed a PDF. That is, it must carry two properties, including positivity and the mass preservation property. It turns out that under some technical assumptions, the mass preservation property is equivalent to the divergence-free condition.

Lemma 4.3.2. *Let C be a given function which satisfies assumption [A8], while $p_0(u; x)$ satisfies assumption [A9]. Let $p(u; x, t)$ be the solution to equation (4.14). We have the following statements:*

- 1) *If $p_0 \geq 0$ then $p(u; x, t) \geq 0$.*
- 2) *If we further assume there exists $m \geq 1$ such that $\lim_{|u| \rightarrow \infty} p_0(u; x)|u|^m = 0$ uniformly in x as $|u| \rightarrow \infty$, then $\lim_{|u| \rightarrow \infty} p(u; x, t)|u|^m = 0$ uniformly in x . Moreover, if $m > q + 3$ for some $q \geq 1$, then $\int_{\mathbb{R}^3} |u|^q p(u; x, t) du < \infty$.*
- 3) *Suppose p_0 also satisfy the conditions in 2) and $\int_{\mathbb{R}^3} p_0(u; x) du = 1$, we have $\int_{\mathbb{R}^3} p(u; x, t) du = 1$ for all x, t if and only if*

$$\frac{\partial}{\partial x^i} \int_{\mathbb{R}^3} u^i p(u; x, t) du = 0. \tag{4.16}$$

Proof. 1) follows directly from the stochastic representation formula (4.15). To deal with 2), let $Y(s; u)$ and $X(s; u)$ denote the solutions to the SDEs in Theorem 4.3.1, while we put u in the representation of X, Y to emphasize their dependence on the initial data u . As a result of the Lipschitz condition, we deduce that $|C(X(s; u), Y(s; u), t - s) \cdot Y(s; u)| \leq \frac{1}{2}K(1 + |Y(s; u)|^2)$ for some $K > 0$. Introducing scalar processes $Z(s; u) := |Y(s; u)|^2$, $\alpha(s; u)$ and $\beta(s; u)$:

$$\begin{aligned} dZ(s; u) &= 2C(Y(s; u), X(s; u), t - s) \cdot Y(s; u) ds, \\ d\alpha(s; u) &= -K(1 + \alpha(s; u)) ds, \\ d\beta(s; u) &= K(1 + \beta(s; u)) ds, \end{aligned}$$

with $\alpha(0; u) = \beta(0; u) = |Y(0; u)|^2 = |u|^2$, $Z(s; u)$ must satisfy the inequality $\alpha(s; u) \leq Z(s; u) \leq \beta(s; u)$, since we are able to make use of the comparison theorem (see for example, [65]) and

$$-K(1 + Z(s; u)) \leq 2C(Y(s; u), X(s; u), t - s) \cdot Y(s; u) \leq K(1 + Z(s; u))$$

for all $Z(s; u) \geq 0$. Therefore

$$Z(s; u) \in [\exp(-Ks)(|u|^2 + 1) - 1, \exp(Ks)(|u|^2 + 1) - 1].$$

Applying the dominated convergence theorem, we get

$$\lim_{|u| \rightarrow \infty} |p(u; x, t)| |u|^m \leq \exp(Kt) \mathbb{E} \left[\lim_{|u| \rightarrow \infty} |p_0(Y(t; u); X(t, u))| |u|^m \right] = 0.$$

Recall that $p_0(u; x)$ has at most polynomial growth in (u, x) , i.e. $p_0(u; x) \leq L(|u|^l + |x|^l)$ for some $l, L \geq 1$. Denoting $B(0, R)$ as the ball centred at the origin with radius $R > 0$, the moment bound can be obtained by

$$\begin{aligned} \int_{\mathbb{R}^3} |u|^q p(u; x, t) du &\leq \int_{B(0, R)} |u|^q \exp(Kt) \mathbb{E} [|Y(s; u)|^l + |X(s; u)|^l] du \\ &\quad + \int_{\mathbb{R}^3 \setminus B(0, R)} |u|^m p(u; x, t) \frac{1}{|u|^{m-q}} du \\ &< \infty, \end{aligned}$$

where R is chosen large enough such that $|u|^m p(u; x, t) \leq K_1$ for all $|u| > R$ and some constant $K_1 > 0$.

In order to check 3), we consider $f(x, t) := \int_{\mathbb{R}^3} p(u; x, t) du$. Integrating the PDF PDE with respect to the variable u , we obtain that

$$\left(\frac{\partial}{\partial t} - \nu \Delta_x \right) f(x, t) = - \frac{\partial}{\partial x^i} \int_{\mathbb{R}^3} u^i p(u; x, t) du + \int_{\mathbb{R}^3} \nabla_u \cdot (pC) du,$$

$$f(x, 0) = 1.$$

Therefore, $|pC| \rightarrow 0$ and the conclusion follows immediately. $\square$

Remark 4.2. *There exists a (viscosity) solution to equation (4.14) when we assume C is Lipschitz in (x, u) and uniformly continuous in t , but the fact that C^i is bounded in x for all $i \in \{1, 2, 3\}$ is crucial to part 2) of Lemma 4.3.2. Assuming $C^i(y, x, t) = 2a^i y^i + b^i x^i + c^i(t)$ for some $a^i, b^i \in \mathbb{R}$, $(a^i)^2 > b^i$ and bounded function c^i , where no Einstein's convention is applied, the solution to the system of SDEs $(Y^i, X^i)^T$ is given by*

$$\begin{aligned} \begin{bmatrix} Y^i(s) \\ X^i(s) \end{bmatrix} &= \exp(L^i s) \begin{bmatrix} u^i \\ x^i \end{bmatrix} + \int_0^s \exp(L^i(s-r)) \begin{bmatrix} c^i(r) \\ 0 \end{bmatrix} dr \\ &\quad + \int_0^s \exp(M^i(s-r)) \begin{bmatrix} 0 \\ \sqrt{2\nu} \end{bmatrix} dM_r^i, \\ L^i &:= \begin{bmatrix} 2a^i & b^i \\ -1 & 0 \end{bmatrix} \end{aligned}$$

The corresponding eigenvalues of L^i are $\lambda_{1,i} = a^i + \sqrt{(a^i)^2 - b^i}$ and $\lambda_{2,i} = a^i - \sqrt{(a^i)^2 - b^i}$, which implies

$$\exp(L^i t) = \frac{1}{\lambda_{2,i} - \lambda_{1,i}} \begin{bmatrix} -\lambda_{1,i} \exp(\lambda_{1,i} t) + \lambda_{2,i} \exp(\lambda_{2,i} t) & b^i (-\exp(\lambda_{1,i} t) + \exp(\lambda_{2,i} t)) \\ \exp(\lambda_{1,i} t) - \exp(\lambda_{2,i} t) & \lambda_{2,i} \exp(\lambda_{1,i} t) - \lambda_{1,i} \exp(\lambda_{2,i} t) \end{bmatrix}$$

and $Y^i(s)$ is independent of u^i provided $s = \frac{\ln(\lambda_{2,i}/\lambda_{1,i})}{\lambda_{1,i} - \lambda_{2,i}} > 0$. In particular, $a^i < 0$ and $b^i > 0$ with $(a^i)^2 > b^i$ leads to $s > 0$, and therefore

$$p(u; x, t) = \mathbb{E}[p_0(Y(t); X(t))q(t)]$$

not necessarily vanishes at $|u| \rightarrow \infty$ for all (x, t) if we only assume $p_0(u; x) \rightarrow 0$ as $|u| \rightarrow \infty$ i.e. $\int_{\mathbb{R}^3} p(u; x, t) du = \infty$ for some $t > 0$.

We introduce a sufficient condition for the mass-preservation property when C^i satisfies the constraints in the following Lemma.

Lemma 4.3.3. *Suppose C, p_0 satisfy condition 2) in Lemma 4.3.2 with $m \geq 2$ and define $\{U(x, t)\}_{x,t}$ to be the corresponding random field such that $U(x, t)$ has the PDF $p(u; x, t)$ for all (x, t) . If equation (4.16) is satisfied, we have for all $(x, t) \in \mathbb{R}^3 \times \mathbb{R}_+$,*

$$\nabla_x \cdot \mathbb{E}[C(x, U(x, t), t)] = -\partial_{x^i} \partial_{x^j} \mathbb{E}[U^i(x, t) U^j(x, t)].$$

Proof. We apply $u \cdot \nabla_x$ on both sides of the PDE (4.14) followed by integrating with respect to u , resulting

$$\int_{\mathbb{R}^3} (u \cdot \nabla_x) (\partial_t - \nu \Delta_x) p(u; x, t) \, du = (\partial_t - \nu \Delta_x) \nabla_x \cdot \int_{\mathbb{R}^3} u p(u; x, t) \, du = 0$$

as well as

$$\begin{aligned} \int_{\mathbb{R}^3} (u \cdot \nabla_x) u \cdot \nabla_x p(u; x, t) \, du &= \int_{\mathbb{R}^3} u^i u^j \partial_{x^i} \partial_{x^j} p(u; x, t) \, du \\ &= \partial_{x^i} \partial_{x^j} \mathbb{E} [U^i(x, t) U^j(x, t)]. \end{aligned}$$

Eventually, the right-hand side can be written in terms of

$$\begin{aligned} \int_{\mathbb{R}^3} (u \cdot \nabla_x) \nabla_u \cdot (p(u; x, t) C(x, u, t)) \, du &= \int_{\mathbb{R}^3} \partial_{x^i} \nabla_u \cdot (u^i p(u; x, t) C(x, u, t)) \, du \\ &\quad - \int_{\mathbb{R}^3} \nabla_u \cdot (p(u; x, t) C(x, u, t)) \, du \\ &= -\nabla_x \cdot \mathbb{E} [C(x, U(x, t), t)], \end{aligned}$$

provided $|p(u; x, t) C(x, u, t) u^i| \rightarrow 0$ as $|u| \rightarrow \infty$. This has been guaranteed by the growth condition on C^i . $\square$

Remark 4.3. As we only used the growth condition of C and the fact that $p(u; x, t)|u|^m \rightarrow 0$ as $|u| \rightarrow \infty$, Lemma 4.3.3 is also valid different types flows, other than weakly homogeneous flow or weakly isotropic flows.

If we further assume the incompressible viscous turbulent fluid flow is both weakly homogeneous and weakly isotropic, the PDF PDE is a parabolic-transport equation

$$\begin{aligned} \left(\frac{\partial}{\partial t} + u^i \frac{\partial}{\partial x^i} - \nu \Delta_x \right) p &= \frac{\partial}{\partial u^i} (p C^i), \\ p(u; x, 0) &= p_0(u; x), \end{aligned} \tag{4.17}$$

where C is essentially

$$\begin{aligned} C^i(x, u, t) &= -\nu A^i(u, t) + Q^i(x, u, t) \\ &= -\nu A^i(u, t) + \int_{\mathbb{R}^3} \frac{y^i}{4\pi|y|} \frac{\partial^2 (\sigma^{jk} + b^j b^k)}{\partial y^k \partial y^j} (y, u, t) \, dy \end{aligned}$$

depending only on u and t but not on x .

Corollary 4.3.4. Let $C^i(u, t)$ be Lipschitz continuous in u and uniformly continuous in t , for all $i \in \{1, 2, 3\}$. For $t > 0$ and $(x, u) \in \mathbb{R}^3 \times \mathbb{R}^3$, (X, Y, q) denotes the unique solution to the system of equations for all $i \in \{1, 2, 3\}$ and $s \in [0, t]$:

$$dX^i(s) = -Y^i(s) \, ds + \sqrt{2\nu} \, dM_s^i, \quad X(0) = x,$$

$$dY^i(s) = C^i(Y(s), t-s) ds, \quad Y(0) = u,$$

and

$$dq(s) = -q(s) \frac{\partial C^k}{\partial u^k}(Y(s), t-s) ds, \quad q(0) = 1.$$

Suppose $p(u; x, t)$ is a $\mathcal{C}^{1,1,1}(\mathbb{R}^3 \times \mathbb{R}^3 \times [0, T])$ solution to (4.17) then

$$p(u; x, t) = q(t) \int_{\mathbb{R}^3} H(x, t, z; u) p_0(Y(t); z) dz,$$

where

$$H(x, t, y; u) = \frac{1}{(4\pi\nu t)^{3/2}} \exp\left(-\frac{\left|y - x + \int_0^t Y(s) du\right|^2}{4\nu t}\right).$$

Therefore if $p_0(u; x) \geq 0$ for every x , then so is $p(u; x, t)$ for any (x, t) . If $p_0(u; x)$ is a PDF for all x , then $p(u; x, t)$ is again a PDF for all (x, t) if only if the following constraint holds:

$$\frac{\partial}{\partial x^i} \int_{\mathbb{R}^3} u^i p(u; x, t) du = 0.$$

Proof. By definition, Y and q are deterministic processes, while X is a 3-dimensional Gaussian process such that for all $E \in \mathcal{B}(\mathbb{R}^3)$ and $s \in [0, t]$,

$$\mathbb{P}(X_s \in E) = \int_E H(x, s, z; u) dz.$$

As a result of Theorem 4.3.1,

$$p(u; x, t) = q(t) \mathbb{E}[p(Y(t); X(t), 0)] = q(t) \int_{\mathbb{R}^3} H(x, t, z; u) p_0(Y(t); z) dz.$$

The remaining part is a consequence of Lemma 4.3.2. □

4.3.2 Inviscid flows

The modelling of inviscid incompressible flows is significantly simplified comparing to modelling viscous flow, since the PDF PDE can be solved without imposing the weak homogeneity or the weak isotropy conditions. As $\nu = 0$, the velocity $U(x, t)$ fulfils the Euler equations, while the PDF PDE becomes the following transport equation

$$\begin{aligned} \frac{\partial p}{\partial t} + u^i \frac{\partial p}{\partial x^i} &= \frac{\partial}{\partial u^i} (p Q^i), \\ p(u; x, 0) &= p_0(u; x). \end{aligned} \tag{4.18}$$

Theorem 4.3.5. *Suppose that $Q^i(x, u, t)$ satisfies assumption [A8], $p_0(u; x)$ satisfies assumption [A9] and $p \in \mathcal{C}^{1,1,1}(\mathbb{R}^3, \mathbb{R}^3, [0, T])$ for some fixed $T > 0$ is a solution to the transport PDE (4.18), we have*

$$p(u; x, t) = p_0(Y(t); X(t))q(t) \quad (4.19)$$

for every $t > 0$, x and u , where (X, Y, q) is the unique solution to the following system of ODEs:

$$\begin{aligned} dX^i(s) &= -Y^i(s) ds, \quad X(0) = x, \\ dY^i(s) &= Q^i(X(s), Y(s), t-s) ds, \quad Y(0) = u \\ dq(s) &= q(s)(\nabla_u \cdot Q)(X(s), Y(s), t-s) ds, \quad q(0) = 1. \end{aligned} \quad (4.20)$$

for $i = 1, 2, 3$. Moreover, if $p(u; x, 0) \geq 0$ for every x , then so is $p(u; x, t)$ for any (x, t) . If $p(u; x, 0)$ is a PDF for all x and $p_0(u; x)|u|^m \rightarrow 0$ uniformly in x for some $m \geq 1$, then $p(u; x, t)$ is again a PDF for all (x, t) if only if the the following constraint is satisfied:

$$\frac{\partial}{\partial x^i} \int_{\mathbb{R}^3} u^i p(u; x, t) du = 0.$$

Proof. The system of ODEs (4.20) has a unique solution pair (X, Y) and

$$q(s) = \exp \left(\int_0^s \frac{\partial Q^i}{\partial u^i}(X(r), Y(r), t-r) dr \right),$$

which is a bounded process by the Lipschitz assumption. Let $\theta(s) = (Y(s); X(s), t-s)$, $\eta(s) = (X(s), Y(s), t-s)$ and define

$$h(s) = p(\theta(s))q(s),$$

then $h(0) = p(u; x, t)$ and $h(t) = p(Y(t); X(t), 0)q(t)$. Moreover for $s \in [0, t]$, we have

$$\begin{aligned} h'(s) &= q(s) \left(\frac{\partial p}{\partial u^i}(\theta(s)) \frac{\partial Y^i}{\partial s} + \frac{\partial p}{\partial x^i}(\theta(s)) \frac{\partial X^i}{\partial s} - \frac{\partial p}{\partial s}(\theta(s)) \right) + p(\theta(s)) \frac{\partial q}{\partial s} \\ &= q(s) \left(\frac{\partial p}{\partial u^i}(\theta(s)) Q^i(\eta(s)) + \frac{\partial p}{\partial x^i}(\theta(s)) Y^i(s) - \frac{\partial p}{\partial s}(\theta(s)) + p(\theta(s)) \frac{\partial Q^i}{\partial u^i}(\eta(s)) \right) \\ &= 0, \end{aligned}$$

hence h is constant on $[0, t]$ and we make use of $p(u; x, t) = h(t)$ to deduce (4.19). Regarding the positivity and mass preservation properties, the proof is almost the same as the proof of Lemma 4.3.2, except (4.20) is not stochastic: if $p_0 \geq 0$ then $p(u; x, t) \geq 0$ directly and

$$\lim_{|u| \rightarrow \infty} |p(u; x, t)| |u|^m \leq \sup_{s \in [0, t]} |q(s)| \lim_{|u| \rightarrow \infty} |p_0(Y(t; u); X(t; u))| |u|^m = 0.$$

Let $f(x, t) = \int_{\mathbb{R}^3} p(u; x, t) du$. If p_0 is a probability density, then $f(x, 0) = 1$ for all x . By integrating the equation (4.18) with respect to u , we obtain

$$\frac{\partial}{\partial t} f(x, t) - \frac{\partial}{\partial x^i} \int_{\mathbb{R}^3} u^i p(u; x, t) du = 0,$$

which leads us to the conclusion. $\square$

4.4 Modelling the PDF: a concrete example

On one hand, from the modelling point of view, only those solutions to the PDF PDE which satisfy the natural mass conservation condition (4.16) can be used as models of distributions of turbulent velocity fields. On the other hand, from view-point of PDEs, the mass conservation property of solutions to the PDF PDE imposes a strong constraint on its solutions. As a matter of fact, most solutions of PDF PDE with given A, B, Q do not satisfy this constraint. In general, solutions to the PDF PDE do not have an explicit expression, although our stochastic representations established in the previous sections may be helpful in dealing with the mass conservation. Consequently, it brings a next level of difficulty to verify the constraint (4.16). It turns out that the natural condition that $\int_{\mathbb{R}^3} p(u; x, t) du = 1$ for a solution $p(u; x, t)$ to the PDF PDE, which is equivalent to the mass conservation, is a very strict constraint for whatever the coefficients A, B and Q which may be modelled or measured. At least our experience demonstrates that the explicit PDF solutions to the PDF PDE are rare and difficult to find. In addition, we see that the PDF solution $p(u; x, t)$ is not so sensitive for the choices of A, B and Q , in the sense that different choices of A, B, Q might leads to the same solution. Therefore the PDF PDE together with the mass conservation constraint is very rigid, and hence is good for modelling the PDF of turbulent flows. We hope to see further exploration in this direction in the future.

In this section, we study an explicit example to the PDF PDE, where the mixed-derivative term vanishes in the PDE i.e. the turbulent flow is weakly homogeneous. The example seems artificial, but as we have explained above, we believe that this example has relevance to real turbulent flows.

4.4.1 Space homogeneous density with perturbation

For simplicity, the viscosity parameter in this example is set to be $\nu = 1$. As the solution to the PDE (4.14) is solely determined by the initial data $p_0(u; x)$ as well as the function $C(x, u, t)$, we consider the simplest scenario $C = 0$. Meanwhile, instead

of setting up a common distribution such as Gaussian or exponential distribution for the initial data, we introduce the following non-negative function: for every $x \in \mathbb{R}^3$, let

$$p_0(u; x) = \alpha(u) + \beta(u)\gamma(x), \quad (4.21)$$

where $\alpha(u)$ is a PDF given by

$$\alpha(u) = \frac{1}{(\sqrt{2\pi})^3 \sqrt{\det \sigma}} \exp\left(-\frac{1}{2}u^T \sigma^{-1}u\right)$$

with $\sigma_{ij} = (\frac{3}{2})^{i-2} \mathbb{1}_{\{i=j\}}$, corresponding to the PDF of a centred Gaussian vector with independent components. Here, β satisfies $\int_{\mathbb{R}^3} \beta(u) du = 0$, and is chosen to be the product of reciprocals

$$\beta(u) = \begin{cases} \prod_i \frac{1}{u^i}, & (u^1, u^2, u^3) \in I, \\ 0, & \text{otherwise,} \end{cases}$$

and truncated if u is not in the region I , where I is defined as

$$I := \left(\left[\frac{1}{4}, 1\right] \cup \left[-1, -\frac{1}{4}\right]\right) \times \left(\left[\frac{1}{4}, 2\right] \cup \left[-2, -\frac{1}{4}\right]\right) \times \left(\left[\frac{2}{7}, 3\right] \cup \left[-3, -\frac{2}{7}\right]\right).$$

We further set the last function γ as

$$\gamma(x) = \left\{ \left[\frac{30(x^1 - 1)}{30 + 3(x^2)^2 + 2(x^3)^2} \exp\left(-\frac{(x^1)^2}{3} + \frac{1}{3} \sin\left(\sum_i x^i\right)\right) + \frac{\cos(x^2 + x^3)}{(|x|^2 + 1)} \right] \right. \\ \left. + 2 \exp\left(-\frac{|x|^2}{200}\right) \right\} \frac{1}{36} \frac{1}{(\sqrt{2\pi})^3},$$

which vanishes at $|x| \rightarrow \infty$. We select these functions so that the positivity and mass preserving properties are fulfilled, and the initial velocity is a random field whose marginal density at x is given by (4.21).

By the stochastic representation formula, the PDF of the random field at (x, t) has the form

$$p(u; x, t) = \alpha(u) + \beta(u) \mathbb{E} \left[\gamma \left(x - ut + \sqrt{2\nu} M_t \right) \right]. \quad (4.22)$$

We focus on the PDF in the random field $\{U(x, t)\}_{x,t}$ at $x = (0, 0, 0)$ and plot the graph of $p(u; x, t)$ against u at $u_3 = 0.3$ for different time t . At $t = \frac{1}{2}$, $p(u; x, t)$ is discontinuous on the boundary of I as in Figure 4.1. If we compare the density of $p(u; x, \frac{1}{2})$ and $p(u; x, 0)$ by evaluating $p(u; x, \frac{1}{2}) - p(u; x, 0)$, we can see the change of density in the region I . Meanwhile, the discontinuity becomes less apparent on

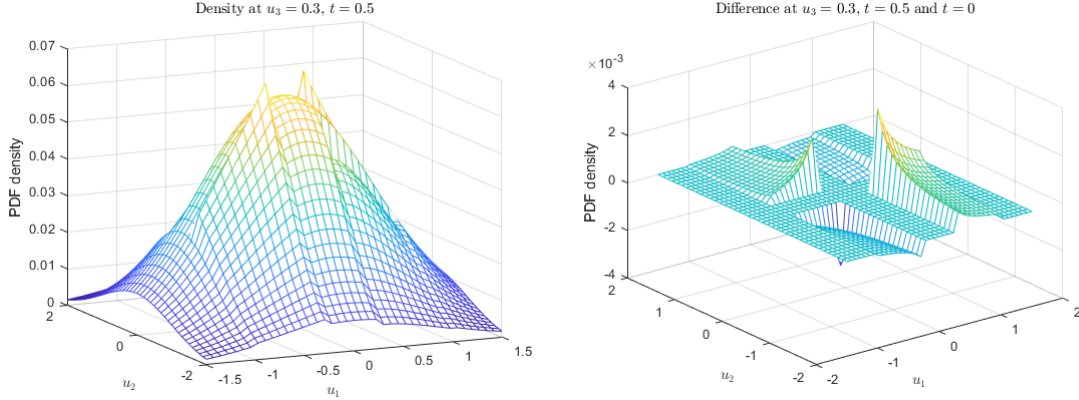


Figure 4.1: $x = (0, 0, 0)$, $t = \frac{1}{2}$

the plot when we increase the time to $t = 40$. From Figure 4.2, the PDF of $U(x, t)$ is close to the density of a Gaussian random variable with density function α , even if the discontinuity still exists. This is due to the fact that γ vanishes at infinity and $\mathbb{E}[\gamma(X_t)] \rightarrow 0$ as $t \rightarrow \infty$. However, the impact of β does not disappear from the velocity field. There is a strong discontinuity near $(\frac{1}{4}, \frac{1}{4}, \frac{2}{7})$ on $p(u; x, t)$ when $x = (12, 12, 12)$ and $t = 40$. The PDF at $x = (12, 12, 12)$ is asymmetric and has a different evolution than the density at the origin, which demonstrates that the impact β gradually shifts from the origin to somewhere far away as time changes.

4.4.2 Motivation for the construction

The mass-preserving property of the PDE (4.14) corresponds to the divergence-free condition (4.16), which is difficult to verify explicitly even when we have the stochastic representation. Apart from describing the motivation for choosing α, β and γ , we will demonstrate our solution to the PDF PDE (4.22) satisfies the divergence-free constraint.

Assuming $C = 0$ does not imply that the turbulence flow associated to solution $p(u; x, t)$ is weakly isotropic. For example, we force the conditional average increment to satisfy

$$\rho^i(x, y, u, t) = O(|y - x|^3)$$

when $|y - x| \rightarrow 0$ and to vanish sufficiently fast when $|y| \rightarrow \infty$, which naturally leads to $A_{jk}^i = 0$. In addition, we let

$$\rho^i(x, y + x, u, t) = \rho^i(x, -y + x, u, t)$$

for all i . If the conditional variance is of the form $\sigma^{jk}(x, y, u, t) = c + \prod_i f_{i,j,k}(y_i - x_i, t)\lambda(x, u, t)$, we conclude that

$$\begin{aligned} Q^i(x, u, t) &= \sum_{j,k} \int \frac{1}{4\pi|y-x|^3} \frac{\partial}{\partial y^i} \frac{\partial^2 (\sigma^{kl} + b^k b^l)}{\partial y^k \partial y^l} dy \\ &= \sum_{j,k} \int \frac{1}{4\pi|y|^3} \frac{\partial^3}{\partial y^i \partial y^j \partial y^k} \left(\prod_i f_{i,j,k}(y_i) \lambda(x, u, t) \right. \\ &\quad \left. + (\rho^j(x, y+x, u, t) \rho^k(x, y+x, u, t)) \right) dy = 0, \end{aligned}$$

provided that $f_{i,j,k}$ is an even function in z with $f_{i,j,k}(z, t) \rightarrow 0$ sufficiently fast as $z \rightarrow 0$ and $z \rightarrow \infty$ for all t . In particular $f_{i,j,k}(z, t) = \frac{1}{8} z^4 \exp\left(-\frac{z^2(1+t)}{2}\right)$ fits the criterion that we required. A pair of the conditional statistics ρ, σ obeying the above constraints is a reasonable choice leading to $C = 0$.

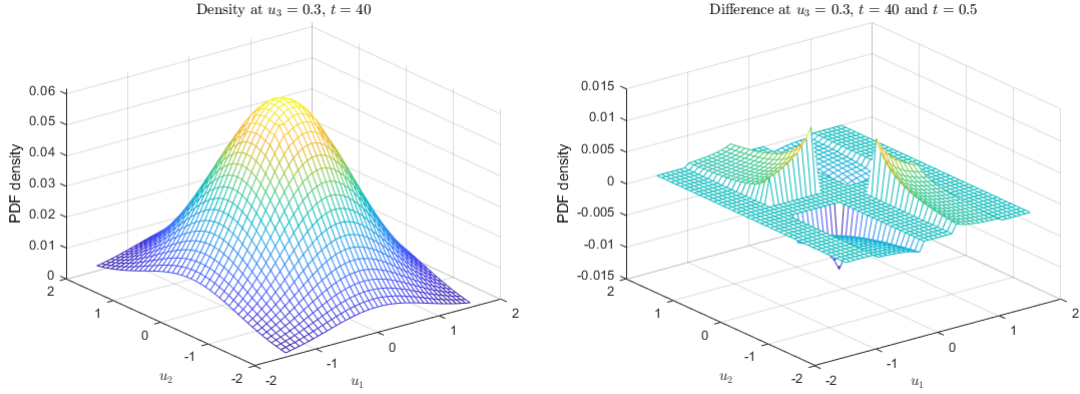


Figure 4.2: $x = (0, 0, 0)$, $t = 40$

Regarding the initial data $p_0(u; x)$ of the form (4.21), there is no strong restriction on α , hence α is allowed to be replaced by another PDF, which does not necessarily correspond to a Gaussian vector. The crucial ingredients in $p_0(u; x)$ are β and γ , which ensure the solution $p(u; x, t)$ satisfy the mass conservation property. As a consequence of the fact that $C = 0$, the right-hand side of equation (4.17) vanishes, ending up an equation which depends solely on the derivatives of p with respect to (t, x) . Moreover, the relevant SDEs in the stochastic representation (4.15) have the explicit form $Y_t = u$ and $X_t = x - ut + \sqrt{2\nu}M_t$, while the divergence-free constraint reads that

$$\nabla_x \cdot \int_{\mathbb{R}^3} u p(u; x, t) du = \mathbb{E} \left[\int_{\mathbb{R}^3} u \cdot \nabla_x p(u, x + \sqrt{2\nu}M_t - ut, 0) du \right] = 0.$$

for all $x \in \mathbb{R}^3, t > 0$. In particular, if

$$\int_{\mathbb{R}^3} u \cdot \nabla_x p(u, x - ut, 0) du = 0 \quad (4.23)$$

is ensured for all (x, t) , $\int_{\mathbb{R}^3} p(u; x, t) du = 1$ is guaranteed.

When $t > 0$, the left-hand side of the equation (4.23) has the following form

$$\begin{aligned} \int_{\mathbb{R}^3} u \cdot \nabla_x p(u; x - ut, 0) du &= \int_{\mathbb{R}^3} u \cdot \nabla_x \gamma(x - ut) \beta(u) du \\ &= -\frac{1}{t} \int_{\mathbb{R}^3} u \cdot \nabla_u (\gamma(x - ut)) \beta(u) du \\ &= -\frac{1}{t} \int_{\mathbb{R}^3} \left(\nabla_u \cdot (u \gamma(x - ut) \beta(u)) \right. \\ &\quad \left. - \nabla_u \cdot (u \beta(u)) \gamma(x - ut) \right) du. \end{aligned}$$

If γ and β decay sufficiently fast as $|u| \rightarrow \infty$ for all x , the first term in the integrand has zero contribution after integrating with respect to u . Apart from our consideration on the mathematical side, our choice of γ guarantees the impact of β vanish as $|x| \rightarrow \infty$, but the speed of decay is slow enough such that the impact of β is still observable when t is rather large. Last but not least, it must satisfy the following constraint

$$\|\gamma(x)\|_{\infty} \leq \sup_{u \in \mathbb{R}^3} \left| \frac{\beta(u)}{\alpha(u)} \right|.$$

We remark that under the current form of β , γ decays as $|x| \rightarrow 0$ is not a necessity.

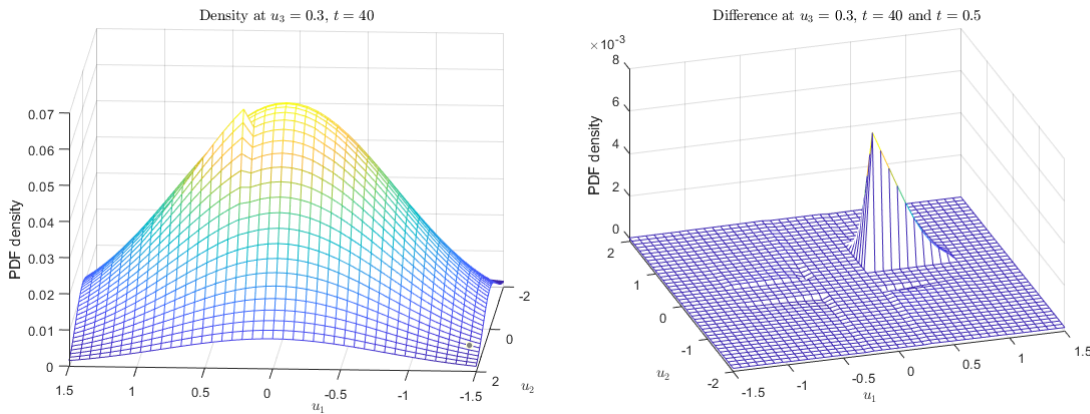


Figure 4.3: $x = (12, 12, 12), t = 40$

The remaining problem turns into finding the right β such that

$$\int_{\mathbb{R}^3} \nabla_u \cdot (u \beta(u)) \gamma(x - ut) du = 0$$

for all $t > 0$ and $\int_{\mathbb{R}^3} u\beta(u) du \cdot \nabla_x \gamma(x) = 0$, which ensure the initial data must also satisfy the divergence-free condition (4.16). Our choice on β is motivated by the fact that equation (4.23) is satisfied provided we impose the constraint $\nabla_u \cdot (u\beta(u)) = 0$ on β . β is truncated to this form, in order to guarantee the integrability as well as the positivity of $p_0(u; x)$. Moreover, we can ensure $\int_{\mathbb{R}^3} \beta(u) du = 0$, $\int_{\mathbb{R}^3} u^i \beta(u) du = 0$ and $\partial_{u^i}(u^i \beta(u)) = \partial_{u^i}(u^j u^k)^{-1} = 0$ for all $i = 1, 2, 3$ and i, j, k being distinct. Nevertheless, the purpose of symmetry of the interval I and β is to simplify our example and β can be asymmetrical.

The PDF (4.22) at (x, t) is discontinuous in the variable u , but it is still a strong solution to the PDE, because the $\partial_{u^i} p$ disappears in the PDE (4.17) in this circumstance.

4.5 Concluding remarks

This Chapter derives a new PDE which describes the evolution of the one-time one-point PDF of the velocity random field of a turbulent flow. The PDF PDE, which is highly non-linear and is determined by two conditional statistics of a turbulent flow, should be a useful tool in modelling distributions of turbulence velocity fields.

The modelling of viscous turbulence in various environments by solving numerically the PDF PDE (4.12) with measured data or based on the priori determination of A, B and Q should be beneficial in understanding turbulent flows. To implement good models of PDFs for turbulent flows, we need to numerically calculate solutions of the PDF PDE, with fed data which determine the functions A, B and Q . The solution has to satisfy the natural constraint, that the mass must be preserved throughout the evolution of the PDF. The conservation of the total mass of the solution is an important topic itself and is worthy of further study. Finally, we would like to point out that the coefficients A, B and Q defined in equation (4.10), which determine the statistics of the turbulence at one-time one-point, must have significant physical meaning in turbulence. These coefficients, which are considered as turbulent flow parameters, should play their roles in further research.

Chapter 5

On the transport equation for probability density functions of turbulent vorticity fields

5.1 Introduction

The vortex dynamics in turbulence have been studied long ago by many researchers (see for example [14, 17, 64, 67, 83] and the literature therein). For incompressible fluid flows, the Navier-Stokes equation is equivalent to the vorticity equation (see (2.16)), and therefore it is natural to consider the vorticity $W = \nabla \wedge U$ as the main fluid dynamical variable in the study of turbulence [67]. There are good reasons why we should concentrate on the vorticity. The velocity of a turbulent flow is unlikely to be independent or possess Markov property (i.e. conditional independence) with respect to the spatial variable, while the vorticity motions of many turbulent flows that can be observed in nature (such as vortex lines, vortex rings) are close to being conditionally independent. In fact there is good evidence, which is already applied in vortex methods (see [17, 64] for a comprehensive account), demonstrating some sort of superposition property of vorticity may be maintained, although not exactly due to the highly non-linear and non-local nature of the fluid dynamics. These observations are valuable in modeling turbulence via the vorticity, especially through the PDF method based on the vorticity.

The main contribution of the present work is the partial differential equation of the probability density function of the vorticity (which will be called the PDF PDE, or called the PDF equation for short) which will be derived in the main body of the chapter. The PDF PDE we discovered is highly non-linear, and its most striking aspect is that the PDF PDE for the vorticity depends only on one single first order

statistical characteristic of the turbulent flow. More precisely, we identify the main statistical characteristic for the PDF PDE, that is the conditional mean function of the increment of the vorticity given the current vorticity:

$$\mu^i(x, y, w, t) = \mathbb{E} [W^i(y, t) - W^i(x, t) | W(x, t) = w] \quad (5.1)$$

for $i = 1, 2, 3$, where $x, y, w \in \mathbb{R}^3$ and $t \geq 0$. The PDF of the vorticity W is a solution to the PDF PDE, which is a second-order partial differential equation where coefficients appearing in the PDF PDE depend on μ only. Besides its theoretical interest, this PDF PDE paves the way towards practical modeling of turbulent flows based on the vorticity PDF.

The chapter, which is based on [58], is organized in the following way. In Section 5.2, several notions and notation together with several standard assumptions about fluid dynamical random fields will be introduced, and the main result, namely the PDF PDE for vorticity will be derived. In Sections 5.3 and 5.4, the theoretical foundation, based on our PDF PDE for vorticity, will be laid for the purpose of modeling various turbulence flows. In Section 5.5, we propose the most direct way of modeling the PDF of the vorticity by specifying the conditional mean function μ^i . The PDF PDE considered purely as a PDE theory is over-determined, due to its non-linearity in the sense that the coefficients appearing in the PDF PDE are not independent of its solutions. Therefore care is needed to ensure that the additional constraint is satisfied. In the last section, we propose the heat flow method to model statistical quantities needed for closing the PDF PDE and obtain concrete PDF examples.

Conventions on notation. The following set of conventions are used throughout this Chapter. We follow the notation (such as the Einstein convention) that is used in Chapter 4. Moreover, the velocity vector field will be denoted by U . $W = \nabla \wedge U$ is its vorticity so that $W^i = \varepsilon^{ijk} \frac{\partial}{\partial x^j} U^k$, where ε^{ijk} are the Levi-Civita symbols.

5.2 PDF equation for the vorticity

In this section we consider the vorticity $W(x, t)$ of an incompressible fluid turbulent flow as the main dynamical variable. Therefore for every pair (x, t) , $W(x, t)$ is a random variable on some probability space. The goal is to derive an evolution equation for the PDF of $W(x, t)$.

5.2.1 Prelims and assumptions

At each point x and instance $t \geq 0$, $W(x, t)$ is a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in $\mathbb{R}^3$, and therefore its law (or called its distribution) is a probability measure on $(\mathbb{R}^3, \mathcal{B}(\mathbb{R}^3))$. In order to discuss the statistics of the turbulent flow, certain assumptions must be imposed, which will be made explicit in the following.

We assume that $W(x, t)$ has a positive and smooth probability density function (PDF), denoted by $f(w; x, t)$, so that

$$\mathbb{P}(W(x, t) \in A) = \int_A f(w, x, t) dw.$$

The joint distribution of $W(x, t)$ and $W(y, t)$ for any pair $x \neq y$ has a positive and smooth PDF, denoted by $f_2(w_1, w_2; x, y, t)$. Let

$$f_{2|1}(w; w_1, x, y, t) = \frac{f_2(w_1, w; x, y, t)}{f(w_1; x, t)}$$

be the conditional PDF of $W(y, t)$ given $W(x, t) = w_1$, and

$$f_{2|1}(w; w_2, y, x, t) = \frac{f_2(w, w_2; x, y, t)}{f(w_2; y, t)}$$

is the conditional PDF of $W(x, t)$ given $W(y, t) = w_2$. Then we have the following integration splitting equations:

$$f_2(w_1, w_2; x, y, t) = f(w_1; x, t)f_{2|1}(w_2; w_1, x, y, t) = f(w_2; y, t)f_{2|1}(w_1; w_2, x, y, t).$$

This helps us to define the conditional mean function $\mu^i(x, y, w, t)$ is defined by

$$\begin{aligned} \mu^i(x, y, w, t) &= \mathbb{E} [W^i(y, t) - W^i(x, t) | W(x, t) = w] \\ &= \int_{\mathbb{R}^3} (w_1^i - w^i) f_{2|1}(w_1; w, x, y, t) dw_1, \end{aligned} \quad (5.2)$$

for $i = 1, 2, 3$, which will play an important role in the sequel.

Regarding the regularity of the conditional mean function μ , we assume the derivatives of μ of any order exist and decay to zero sufficiently fast at infinity. Consequently, μ has an asymptotic expansion

$$\mu^i(x, y, w, t) = a_k^i(x, w, t)(y^k - x^k) + b_{jk}^i(x, w, t)(y^k - x^k)(y^j - x^j) + o(|y - x|^2) \quad (5.3)$$

as $|y - x| \rightarrow 0$, where a_k^i 's and b_{jk}^i 's are continuous with respect to all of their arguments. We denote $b^i = \sum_{k=1}^3 b_{kk}^i$ for $i = 1, 2, 3$. By definition, we observe that

$$a_k^i(x, w, t) = \mathbb{E} \left[\frac{\partial W^i}{\partial x^k}(x, t) \middle| W(x, t) = w \right] = \frac{\partial}{\partial y^k} \mu^i(x, y, w, t) \bigg|_{y=x} \quad (5.4)$$

which represents the local change rate of the vortex motion over the turbulent region, whereas

$$b^i(x, w, t) = \frac{1}{2} \mathbb{E} [\Delta_x W^i(x, t) | W(x, t) = w] = \frac{1}{2} \Delta_y \mu^i(x, y, w, t) \big|_{y=x}. \quad (5.5)$$

5.2.2 PDF equation and its derivation

In this section, we derive the main result of this chapter.

Theorem 5.2.1. *Under the assumptions and notation established in subsection 5.2.1, suppose the PDF $f(w; x, t)$ of the vorticity is smooth in (w, x, t) and has finite moments, that is $\int_{\mathbb{R}^3} |w|^n f(w; x, t) dw < \infty$ for every $n \geq 1$. Then f satisfies the following partial differential equation:*

$$\left(\frac{\partial}{\partial t} + \frac{\partial B^i}{\partial x^i} + B^i \frac{\partial}{\partial x^i} - \nu \Delta_x \right) f = \nu \frac{\partial}{\partial w^i} \left(\frac{\partial}{\partial x^k} (f a_k^i) - 2b^i f \right) + \frac{\partial}{\partial w^i} (f D^i), \quad (5.6)$$

where Δ_x is the Laplacian with respect to the space variable x , a_k and b are given as in (5.4) and (5.5),

$$B^i(x, w, t) = \int_{\mathbb{R}^3} \frac{1}{4\pi|y-x|} \varepsilon^{ijk} \frac{\partial}{\partial y^j} \mu^k(x, y, w, t) dy, \quad (5.7)$$

and

$$D^i(x, w, t) = w^l \int_{\mathbb{R}^3} \frac{y^l - x^l}{4\pi|y-x|^3} \varepsilon^{ijk} \frac{\partial}{\partial y^j} \mu^k(x, y, w, t) dy. \quad (5.8)$$

One direct consequence of the above theorem is the following result.

Corollary 5.2.2. *Assume that the derivatives of a_k^i , b^i , B^i and D^i are continuous and decay to zero sufficiently fast as $|w| \rightarrow \infty$. Then*

$$\frac{\partial}{\partial x^i} \int_{\mathbb{R}^3} B^i(x, w, t) f(w; x, t) dw = 0 \quad (5.9)$$

for all $x \in \mathbb{R}^3, t > 0$.

Proof. This follows easily by integrating (5.6) over $\mathbb{R}^3$ and the fact that

$$\int_{\mathbb{R}^3} f(w; x, t) dw = 1.$$

□

The PDF equation (5.6) can be regarded as the adjoint equation¹ of the vorticity equation (2.16) which is however non-linear, and therefore the PDF equation requires a careful derivation which will be the subject of the remaining part of this section.

¹This is the adjoint equation in the sense that there exist an “adjoint operator” $\mathcal{L}$ such that $\int_{\mathbb{R}^3} F(w) \partial_t f(w; x, t) dw = \int_{\mathbb{R}^3} F(w) (\mathcal{L} f)(w; x, t) dw$.

Proof. (Proof of Theorem 5.2.1) Let F be a smooth function on $\mathbb{R}^3$ with a compact support, and let us consider the quantity

$$\frac{\partial}{\partial t} \mathbb{E} [F(W(x, t))].$$

On one hand we have

$$\begin{aligned} \frac{\partial}{\partial t} \mathbb{E} [F(W(x, t))] &= \frac{\partial}{\partial t} \int_{\mathbb{R}^3} F(w) f(w; x, t) dw \\ &= \int_{\mathbb{R}^3} F(w) \frac{\partial}{\partial t} f(w; x, t) dw \end{aligned}$$

and on the other hand

$$\frac{\partial}{\partial t} \mathbb{E} [F(W(x, t))] = \mathbb{E} \left[\frac{\partial}{\partial t} F(W(x, t)) \right].$$

Hence, we have

$$\int_{\mathbb{R}^3} F(w) \frac{\partial}{\partial t} f(w; x, t) dw = \mathbb{E} \left[\frac{\partial}{\partial t} F(W(x, t)) \right].$$

We need to express the right-hand side expectation in terms of the PDF $f(w; x, t)$ together with the statistical numerical characteristics. This will be done by using the equation

$$\frac{\partial}{\partial t} F(W(x, t)) = \nu \Delta_x F(W) - \frac{\partial (U^i F(W))}{\partial x^i} + F_j(W) W^i \frac{\partial U^j}{\partial x^i} - \nu \frac{\partial F_j(W)}{\partial x^k} \frac{\partial W^j}{\partial x^k},$$

where, for simplicity, F_j denote the partial derivatives $\frac{\partial F}{\partial x^j}$ of F , $j = 1, 2, 3$. This equation follows directly from the vorticity equation. From this equation we obtain that

$$\begin{aligned} \mathbb{E} \left[\frac{\partial}{\partial t} F(W(x, t)) \right] &= \nu \mathbb{E} [\Delta_x F(W)] - \mathbb{E} \left[\frac{\partial (U^i F(W))}{\partial x^i} \right] \\ &\quad + \mathbb{E} \left[F_j(W) W^i \frac{\partial U^j}{\partial x^i} \right] - \mathbb{E} \left[\nu \frac{\partial F_j(W)}{\partial x^k} \frac{\partial W^j}{\partial x^k} \right] \\ &:= I_1 + I_2 + I_3 + I_4 \end{aligned} \tag{5.10}$$

Let us calculate I_i 's on the right-hand side of the equation (5.10). First, it is easy to see that by definition

$$I_1 = \nu \Delta_x \mathbb{E} [F(W)] = \int_{\mathbb{R}^3} F(w) \nu \Delta_x f(w; x, t) dw.$$

For the second term in the sum, we recall (2.17) is equivalent to

$$U^i(x, t) = \int_{\mathbb{R}^3} \frac{1}{4\pi|y-x|} \varepsilon^{ijk} \frac{\partial}{\partial y^j} W^k(y, t) dy,$$

for $i = 1, 2, 3$. We thus obtain that

$$\begin{aligned} I_2 &= -\mathbb{E} \left[\frac{\partial(U^i F(W))}{\partial x^i} \right] = -\frac{\partial}{\partial x^i} \mathbb{E} [F(W)U^i] \\ &= -\frac{\partial}{\partial x^i} \mathbb{E} \left[F(W) \varepsilon^{ijk} \int_{\mathbb{R}^3} \frac{1}{4\pi|y-x|} \frac{\partial}{\partial y^j} W^k(y, t) dy \right]. \end{aligned}$$

To work out I_2 , we rewrite the partial derivative as a limit:

$$\frac{\partial}{\partial y^j} W^k(y, t) = \lim_{h \rightarrow 0} \frac{1}{h} (W^k(y + h e^{(j)}, t) - W^k(y, t)),$$

where $e^{(j)}$ represents the unit vector with j -th component equal to 1, and the other two components equal to 0, e.g. $e^{(1)} = (1, 0, 0)$ and so on. Using the above limit definition, the second term I_2 turns into

$$I_2 = -\frac{\partial}{\partial x^i} \lim_{h \rightarrow 0} \frac{1}{h} \mathbb{E} \left(F(W) \varepsilon^{ijk} \int_{\mathbb{R}^3} \frac{1}{4\pi|y-x|} (W^k(y + h e^{(j)}, t) - W^k(y, t)) dy \right).$$

Before taking the limit, the expectation part in this expression can be written in terms of the two point joint distribution of the random field W as follows:

$$\int_{\mathbb{R}^3} F(w) \left[\int \frac{\varepsilon^{ijk}}{4\pi|y-x|} \left(\int_{\mathbb{R}^3} w_1^k \left(f_2(w, w_1; x, y + h e^{(j)}, t) - f_2(w, w_1; x, y, t) \right) dw_1 \right) dy \right] dw,$$

where the inner integral against the variable w_1 equals

$$f(w; x, t) (\mu(x, y + h e^{(j)}, w, t) - \mu(x, y, w, t)).$$

After substituting this in and sending $h \rightarrow 0$, we obtain that

$$I_2 = \int_{\mathbb{R}^3} F(w) \left[-\frac{\partial}{\partial x^i} \left(f(w; x, t) \int_{\mathbb{R}^3} \frac{\varepsilon^{ijk}}{4\pi|y-x|} \frac{\partial}{\partial y^j} \mu^k(x, y, w, t) dy \right) \right] dw.$$

Next we deal with I_3 . Again, using (2.17), we may write

$$\frac{\partial}{\partial x^j} U^i(x, t) = \int_{\mathbb{R}^3} \frac{y^j - x^j}{4\pi|y-x|^3} \varepsilon^{ilk} \frac{\partial}{\partial y^l} W^k(y, t) dy,$$

and thus

$$\begin{aligned} I_3 &= \mathbb{E} \left[F_i(W) W^j \frac{\partial U^i}{\partial x^j} \right] \\ &= \mathbb{E} \left[F_i(W) W^j \int_{\mathbb{R}^3} \frac{y^j - x^j}{4\pi|y-x|^3} \varepsilon^{ilk} \frac{\partial}{\partial y^l} W^k(y, t) dy \right], \end{aligned}$$

which can be evaluated by using the two-point joint PDF. Indeed, we may repeat the same idea as in the computation of the second term I_2 , and this leads to

$$I_3 = \int_{\mathbb{R}^3} F(w) \frac{\partial}{\partial w^i} \left[-f(w; x, t) \varepsilon^{ilk} w^j \int_{\mathbb{R}^3} \frac{y^j - x^j}{4\pi|y - x|^3} \frac{\partial}{\partial y^l} \mu^k(x, y, w, t) dy \right] dw.$$

Similarly, for the last term I_4 , we write

$$\begin{aligned} I_4 &= -\nu \sum_{k=1}^3 \mathbb{E} \left[\frac{\partial F_i(W)}{\partial x^k} \frac{\partial W^i}{\partial x^k} \right] \\ &= -\nu \sum_{k=1}^3 \lim_{h \rightarrow 0} \frac{1}{h^2} \mathbb{E} \left[(F_i(W(x + he^{(k)}), t) - F_i(W(x, t))) (W^i(x + he^{(k)}), t) - W^i(x, t)) \right] \\ &= -\nu \sum_{k=1}^3 \lim_{h \rightarrow 0} \frac{1}{h^2} \iint_{\mathbb{R}^3 \times \mathbb{R}^3} (F_i(w_2) - F_i(w_1)) (w_2^i - w_1^i) f_2(w_1, w_2, x, x + he^{(k)}, t) dw_1 dw_2 \\ &= \nu \int_{\mathbb{R}^3} F_i(w) \lim_{h \rightarrow 0} \frac{1}{h^2} \sum_{k=1}^3 \left\{ f(w; x + he^{(k)}, t) \int_{\mathbb{R}^3} (w_1^i - w^i) f_{2|1}(w_1; w, x + he^{(k)}, x, t) dw_1 \right. \\ &\quad \left. + f(w; x, t) \int_{\mathbb{R}^3} (w_1^i - w^i) f_{2|1}(w_1; w, x, x + he^{(k)}, t) dw_1 \right\} dw \\ &= \nu \int_{\mathbb{R}^3} F_i(w) \lim_{h \rightarrow 0} \frac{1}{h^2} \sum_{k=1}^3 \left\{ (f(w; x + he^{(k)}, t) - f(w; x, t)) \mu^i(x + he^{(k)}, x, w, t) \right. \\ &\quad \left. + f(w; x, t) (\mu^i(x, x + he^{(k)}, w, t) + \mu^i(x + he^{(k)}, x, w, t)) \right\} dw. \end{aligned}$$

For convenience, let us denote the sum by I_4^h , i.e.

$$\begin{aligned} I_4^h &:= \sum_{k=1}^3 \left\{ (f(w; x + he^{(k)}, t) - f(w; x, t)) \mu^i(x + he^{(k)}, x, w, t) \right. \\ &\quad \left. + f(w; x, t) (\mu^i(x, x + he^{(k)}, w, t) + \mu^i(x + he^{(k)}, x, w, t)) \right\}. \end{aligned}$$

Notice that I_4^h can be evaluated by using the regularity condition (5.3) when h is sufficiently small, by which we mean that

$$\mu^i(x, x + he^{(k)}, w, t) = a_k^i(x, w, t)h + b_{kk}^i(x, w, t)h^2 + o(h^2)$$

and

$$\mu^i(x + he^{(k)}, x, w, t) = -a_k^i(x + he^{(k)}, w, t)h + b_{kk}^i(x + he^{(k)}, w, t)h^2 + o(h^2).$$

Consequently, we have

$$\mu^i(x, x + he^{(k)}, w, t) + \mu^i(x + he^{(k)}, x, w, t)$$

$$\begin{aligned}
&= - \left(a_k^i(x + he^{(k)}, w, t) - a_k^i(x, w, t) \right) h \\
&\quad + \left(b_{kk}^i(x, w, t) + b_{kk}^i(x + he^{(k)}, w, t) \right) h^2 + o(h^2),
\end{aligned}$$

which yields that

$$\lim_{h \rightarrow 0} \frac{1}{h^2} I_4^h = - \frac{\partial}{\partial x^k} (f a_k^i) + 2f \sum_k b_{kk}^i.$$

Then, we conclude that

$$\begin{aligned}
I_4 &= -\nu \int_{\mathbb{R}^3} F_i(w) \left[\frac{\partial}{\partial x^k} (f a_k^i) - 2f \sum_k b_{kk}^i \right] dw \\
&= \int_{\mathbb{R}^3} F(w) \frac{\partial}{\partial w^i} \left[\nu \frac{\partial}{\partial x^k} (f a_k^i) - 2\nu f \sum_k b_{kk}^i \right] dw.
\end{aligned}$$

Putting all together the results we just derived for I_1 to I_4 , we deduce the PDF equation (5.6). $\square$

Remark 5.1. *The partial differential equation (5.6) is in fact, although it appears linear in the PDF f , highly non-linear, more importantly, the coefficients a, b, B and D cannot be in general determined by f solely. We have seen that these functions are functionals of the conditional mean function μ , hence the PDF f cannot be obtained solely through the PDF equation. However, the PDF PDE provides an appealing and new method for modeling the PDF f based on the modeling of μ only, we will provide in this work several results on modeling the PDF f based on the PDF equation we just derived.*

Remark 5.2. *By feeding in the data on the conditional mean function μ , which is equivalent to that all coefficients are considered as given, we can pose the initial value problem for solving the PDF PDE. The good news is that then the PDF PDE is a scalar linear partial differential equation of second order, but it is, in general, a mixed type of partial differential equations of second order on a 6 dimensional space, which is a kind of mixture of parabolic and transport equations. To the best knowledge of the authors of the present chapter, it seems that a theory for this kind of partial differential equation is not yet available.*

5.3 Inviscid flows

In this section we consider an ideal inviscid “turbulent” flow. Suppose $\{U(x, t)\}$ is a random field satisfying the technical assumptions listed in Section 5.2.1, and $U(x, t)$

satisfies the Euler equation

$$\frac{\partial}{\partial t} U^i + U^j \frac{\partial U^i}{\partial x^j} = -\nabla_x P,$$

and

$$\frac{\partial U^i}{\partial x^i} = 0.$$

Let $W = \nabla \wedge U$. Then PDF $f(w; x, t)$ now satisfies the (non-linear) transport differential equation

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial x^i} (f B^i) = \frac{\partial}{\partial w^i} (f D^i), \quad (5.11)$$

where B and D are given as in (5.7) and (5.8) respectively.

The following theorem provides the necessary mathematical tool for modeling the PDF of an inviscid turbulent flow.

Theorem 5.3.1. *Consider PDE (5.11) where the data B and D are assumed as given. Assume that B and D are Lipschitz continuous in (x, w) uniformly in $t > 0$. Suppose $f(w; x, t)$ is a solution to (5.11) with initial data $f(w; x, 0)$ which is continuous, and suppose fD decays to zero sufficiently fast as $|w| \rightarrow \infty$. Then we have the following.*

- 1) $f(w; x, 0) \geq 0$ for all $x \in \mathbb{R}^3$, then $f(w; x, t) \geq 0$ for all $x \in \mathbb{R}^3$ and $t > 0$.
- 2) Suppose $\int_{\mathbb{R}^3} f(w; x, 0) dw = 1$ for all $x \in \mathbb{R}^3$, then

$$\int_{\mathbb{R}^3} f(w; x, t) dw = 1$$

for all $x \in \mathbb{R}^3$ and $t > 0$, if and only if

$$\frac{\partial}{\partial x^i} \int_{\mathbb{R}^3} B^i(x, w, t) f(w; x, t) dw = 0 \quad (5.12)$$

for all $x \in \mathbb{R}^3$ and $t > 0$.

Proof. Let us define the integral curves X and Y by the following ordinary differential equation system:

$$\begin{cases} \frac{\partial}{\partial s} X^i(t, s) = -B^i(X(t, s), Y(t, s), t - s), & X(t, 0) = x, \\ \frac{\partial}{\partial s} Y^i(t, s) = D^i(X(t, s), Y(t, s), t - s), & Y(t, 0) = w. \end{cases}$$

Consider the function defined by

$$h(s) = f(Y(t, s); X(t, s), t - s) e^{N(w; x, s)}$$

for all $s \in [0, t]$, where for convenience

$$N(w; x, s) := \int_0^s \left(\frac{\partial D^i}{\partial w^i} - \frac{\partial B^i}{\partial x^i} \right) (Y(t, r); X(t, r), t - r) dr.$$

Then $h(0) = f(w; x, t)$ and

$$h(t) = f(Y(t, t); X(t, t), 0)e^{N(w; x, t)}.$$

It is clear that

$$\begin{aligned} \frac{\partial}{\partial s} h(s) &= e^N \frac{\partial f}{\partial w^i} \frac{\partial}{\partial s} Y^i(t, s) + e^N \frac{\partial f}{\partial x^i} \frac{\partial}{\partial s} X^i(t, s) \\ &\quad - e^N \frac{\partial}{\partial t} f + e^N f \frac{\partial D^i}{\partial w^i} - e^N f \frac{\partial B^i}{\partial x^i} \\ &= e^N \left\{ \frac{\partial(f D^i)}{\partial w^i} - \frac{\partial(B^i f)}{\partial x^i} - \frac{\partial}{\partial t} f \right\} \\ &= 0, \end{aligned}$$

so by integrating the variable s over $[0, t]$, we may conclude that

$$f(w; x, t) = f(Y(t, t); X(t, t), 0) \exp \left[\int_0^t \left(\frac{\partial D^i}{\partial w^i} - \frac{\partial B^i}{\partial x^i} \right) (Y(t, s); X(t, s), t - s) dr \right].$$

The conservation of positivity property 1) follows immediately from this representation. To show 2) we consider the function $u(x, t) = \int_{\mathbb{R}^3} f(w; x, t) dw$ for $x \in \mathbb{R}^3$ and $t \geq 0$. If $u(x, t) = 1$ for all x and $t > 0$, then by integrating (5.11) over $\mathbb{R}^3$ we obtain (5.12). Conversely, if (5.12) holds, by integrating (5.11) over $\mathbb{R}^3$ then

$$\frac{\partial}{\partial t} u(x, t) = 0,$$

which yields that $u(x, t) = 1$. □

5.4 Weakly homogeneous and weakly isotropic flows

As it has been pointed out in the section 2 and 3, the coefficients in the PDF PDE are determined by the conditional mean function $\mu(x, y, w, t)$ defined in (5.1). The significance of this statistical quantity of a turbulent flow has been demonstrated in the derivation of the PDF PDE for vorticity. Due to its clear meaning, this statistical characteristic can be used in the classification of turbulent flows. In section, we define weak homogeneous and weak isotropic conditions, then study the PDF equation under these assumptions.

The homogeneity and the isotropy can be defined in general for random fields indexed by a space variable $x \in \mathbb{R}^d$, which has been introduced into the study of turbulence by G. I. Taylor. The local homogeneous and local isotropic flows were introduced by Kolmogorov for formulating K41 theory (and its improved version

K61 theory). According to Kolmogorov [54, 53], a random field $\{Z(x, t)\}$ is locally homogeneous if the conditional distribution of $Z(y, t) - Z(x, s)$ given $Z(x, s) = z$ is independent of (x, s, z) for any $t \geq s$, and further it is locally isotropic if the conditional distribution of $Z(y, t) - Z(x, s)$ depends only on $|y - x|$. The usage of conditional mean function makes it possible to generalize these concepts to their weak versions.

We may simplify the PDF PDE for turbulent flows when the vorticity random field $W(x, t)$ is *weakly homogeneous* in the sense that for every pair (y, x) and $t > 0$, the conditional mean of $W^i(y, t) - W^i(x, t)$ given $W(x, t) = w$ depends only on $y - x$, w and $t > 0$ but independent of x , so that we may write

$$\mu^i(x, y, w, t) = \beta^i(y - x, w, t),$$

where $\beta^i(0, w, t) = 0$. Then we can simplify expressions in equation (5.4) and (5.5), and write that

$$a_k^i(x, w, t) = \frac{\partial \beta^i}{\partial x^k}(0, w, t),$$

and

$$b^i(x, w, t) = \Delta_x \beta^i(0, w, t),$$

which are still denoted by $a_k^i(w, t)$ and $b^i(w, t)$, though independent of x , and similarly now we have

$$B^i(w, t) = \int_{\mathbb{R}^3} \frac{1}{4\pi|y|} \varepsilon^{ijk} \frac{\partial \beta^k}{\partial y^j}(y, w, t) dy,$$

and

$$D^i(w, t) = w^l \int_{\mathbb{R}^3} \frac{y^l}{4\pi|y|^3} \varepsilon^{ijk} \frac{\partial \beta^k}{\partial y^j}(y, w, t) dy,$$

which again, are functions of (w, t) only. Therefore the PDF PDE can be simplified to be the following mixed type PDE

$$\left(\frac{\partial}{\partial t} + B^i \frac{\partial}{\partial x^i} - \nu \Delta_x \right) f = \nu \frac{\partial}{\partial w^i} \left(a_k^i \frac{\partial f}{\partial x^k} \right) + \frac{\partial}{\partial w^i} (f A^i),$$

where

$$A^i = D^i - 2\nu b^i,$$

with $i = 1, 2, 3$.

We say a random field $\{Z(x, t)\}$ indexed by (x, t) is *weakly isotropic*, if

$$\mathbb{E} [\nabla_x Z(x, t) | Z(x, t) = w] = 0$$

holds for all x, w and t . By definition, a random field $\{Z(x, t)\}$ is locally isotropic in Kolmogorov's sense, then it is weakly isotropic.

If W is weakly homogeneous and weakly isotropic, then $a = 0$, B and A depend on (w, t) only, so for this case, the PDF PDE becomes a parabolic-transport equation

$$\left(\frac{\partial}{\partial t} + B^i \frac{\partial}{\partial x^i} - \nu \Delta_x \right) f = \frac{\partial}{\partial w^i} (f A^i). \quad (5.13)$$

After we have deduced the PDE (5.13), then there is no need to assume that B and A are independent of x . Therefore we may consider the following PDE

$$\left(\frac{\partial}{\partial t} + B^i \frac{\partial}{\partial x^i} + \frac{\partial B^i}{\partial x^i} - \nu \Delta_x \right) f = \frac{\partial}{\partial w^i} (f A^i). \quad (5.14)$$

where both B and A are functions of three variables x , w and t .

The following theorem provides the foundation for modeling weakly isotropic turbulent flows based on the vorticity PDF.

Theorem 5.4.1. *Consider the parabolic-transport differential equation (5.14), where A and B are considered as given. Assume that $A(x, w, t)$ and $B(x, w, t)$ are Lipschitz continuous in the space variable (x, w) uniformly in t . Suppose $f(w; x, t)$ is a solution to (5.14) such that fA is continuous and decays to zero sufficiently fast as $|w| \rightarrow \infty$. Then the following holds:*

- 1) *If $f(w; x, 0) \geq 0$ for all x and w , then $f(w; x, t) \geq 0$ for all $t \geq 0$, x and w .*
- 2) *If $f(w; x, 0)$ is a PDF for every x , i.e. it is non-negative and $\int_{\mathbb{R}^3} f(w; x, 0) dw = 1$ for every x , then so are $f(w; x, t)$ for all x and $t > 0$ if and only if*

$$\frac{\partial}{\partial x^i} \int_{\mathbb{R}^3} B^i(x, w, t) f(x; w, t) dw = 0 \quad (5.15)$$

for all x and $t > 0$.

- 3) *Suppose in addition B is independent of w and $\frac{\partial B^i}{\partial x^i} = 0$. Then if $f(w; x, 0)$ is a PDF for every x , $f(w; x, t)$ is also a PDF for every x and for every $t > 0$.*

The proof of this result, is rather similar to the proof of Theorem 5.3.1, so we only provide an outline of its proof below.

Proof. Let $(M_t^i)_{t \geq 0}$, $i = 1, 2, 3$ be a standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and we assume that M^i and M^j are independent if $i \neq j$. In this case for every fixed $t > 0$ we run the following SDE system:

$$\begin{cases} dX^i(t, s) = \sqrt{2\nu} dM_s^i - B^i(X(t, s), Y(t, s), t - s) ds, & X(t, 0) = x, \\ dY^i(t, s) = A^i(X(t, s), Y(t, s), t - s) ds, & Y(t, 0) = w \end{cases}$$

which has a unique solution running up to t . Apply Itô's formula to

$$H_s = f(Y(t, s); X(t, s), t - s)e^{P(s)},$$

where

$$P(s) = \int_0^s \frac{\partial A^i}{\partial w^i}(X(t, r), Y(t, r), t - r) dr.$$

Then

$$\begin{aligned} dH_s &= e^P \left\{ \frac{\partial f}{\partial w^i} dY_s^i + \frac{\partial f}{\partial x^i} dX_s^i + f \frac{\partial A^i}{\partial w^i} - \frac{\partial f}{\partial s} ds + \nu \Delta_x f ds \right\} \\ &= \sqrt{2\nu} e^P \frac{\partial f}{\partial x^i} dM_s^i \end{aligned}$$

so that

$$H_0 = H_t - \sqrt{2\nu} e^P \frac{\partial f}{\partial x^i} dM_s^i.$$

By taking expectation, one may deduce that the solution can be expressed as

$$f(w; x, t) = \mathbb{E} \left[f(Y(t, t); X(t, t), 0) \exp \left(\int_0^t \frac{\partial A^i}{\partial w^i}(X(t, s), Y(t, s), t - s) ds \right) \right].$$

Statements 1) - 3) can be easily shown using this expression, and the argument in the proof of Theorem 5.3.1. If B is independent of w and divergence-free in x , then, by integrating (5.13),

$$\left(\frac{\partial}{\partial t} + B^i \frac{\partial}{\partial x^i} - \nu \Delta_x \right) u(x, t) = 0,$$

where $u(x, t) = \int_{\mathbb{R}^3} f(w; x, t) dw$. Thus 3) follows from the uniqueness of the previous parabolic equation. $\square$

5.5 Modeling the PDF of weakly isotropic flows

In this and the next sections, several simple models based on our PDF equation for modeling vorticity distributions in turbulence are discussed. Clearly, the most straightforward way for modeling the distribution of vorticity is to assign the conditional mean function $\mu(x, y, w, t)$. The other parameters in the PDF PDE may be determined accordingly.

For practical reasons which will be clarified in our computations below, the simplest yet not trivial model for $\mu(x, y, w, t)$ should be a function of $|y - x|^2$ only and μ decays sufficiently fast at infinity. For such a model, the parameters B^i and D^i appearing in the PDF PDE (5.6) vanish identically. The other parameters a_k^i and

b_{jk}^i are determined by the asymptotic condition as $|y - x| \rightarrow 0$. For this reason, we assign for all $i \in \{1, 2, 3\}$,

$$\mu^i(x, y, w, t) = \begin{cases} -\frac{C}{6\nu}|y - x|^2, & \text{if } |x - y| < \delta, \\ h(|y - x|^2), & \text{if } |x - y| \geq \delta, \end{cases} \quad (5.16)$$

for some $\delta > 0$, where h is a $\mathcal{C}^1$ -function such that h and its derivative decay sufficiently fast at infinity. $C \geq 0$ is a model parameter which should reflect certain physical properties of the underlying flows. The other parameters can be read out from this model easily: $a_k^i = 0$, $b_{jk}^i = -\delta_{jk}\frac{C}{6\nu}$ and therefore $b^i = -\frac{C}{2\nu}$. With this model of μ , the PDF PDE is simplified to be the following parabolic-transport equation:

$$\left(\frac{\partial}{\partial t} - \nu \Delta_x\right) f = C \frac{\partial f}{\partial w^i}. \quad (5.17)$$

The solution to (5.17) has a nice probabilistic representation which may be read out from the proof of Theorem 5.4.1. In fact

$$f(w; x, t) = \mathbb{E}[f(Y(t, t); X(t, t), 0)] \quad (5.18)$$

where $f(w; x, 0)$ is the PDF of the initial vorticity, and X and Y are solutions to the SDE:

$$\begin{cases} dX^i(t, s) = \sqrt{2\nu} dM_s^i, & X(t, 0) = x, \\ dY^i(t, s) = C ds, & Y(t, 0) = w, \end{cases}$$

where M is a standard 3D Brownian motion. The solutions when $s = t$ are given by

$$Y^i(t, t) = w^i + Ct$$

and

$$X(t, t) = x + \sqrt{2\nu}M_t.$$

Substituting these into (5.18) we obtain

$$\begin{aligned} f(w; x, t) &= \mathbb{E}\left[f(w + Ct; x + \sqrt{2\nu}M_t, 0)\right] \\ &= \mathbb{E}\left[f(w + Ct; x + \sqrt{2\nu}t\xi, 0)\right], \end{aligned} \quad (5.19)$$

where ξ is a random vector with the standard 3D normal distribution $N(0, I_3)$. Clearly the representation (5.19) may also be written in terms of Gaussian density

$$\begin{aligned} f(w; x, t) &= \frac{1}{(4\pi\nu t)^{3/2}} \int_{\mathbb{R}^3} f(w + Ct, z, 0) e^{-\frac{|z-x|^2}{4\nu t}} dz \\ &= \frac{1}{(2\pi)^{3/2}} \int_{\mathbb{R}^3} f(w + Ct, x + \sqrt{2\nu}tz, 0) e^{-\frac{|z|^2}{2}} dz. \end{aligned} \quad (5.20)$$

Both representations (5.19) and (5.20) may be used to evaluate $f(w; x, t)$ by using for example Monte-Carlo scheme by sampling Gaussian random variables.

The initial PDF $f(w; x, 0)$ is determined by the initial vorticity distribution. For a turbulent flow, the initial vorticity may be written as a sum:

$$W(x, 0) = \omega_0(x) + \varepsilon(x),$$

where $\omega_0(x)$ is the initial mean vorticity at the location x , and $\varepsilon(x)$ represents a small random perturbation. It is reasonable, therefore, to assume that $\varepsilon(x)$ has a normal distribution $N(0, \sigma(x)^2)$, where the variance $\sigma(x)$ may or may not depend on the location x . However, if $\varepsilon(x)$ is allowed to depend on the location x , then the noise $\varepsilon(x)$ has to satisfy the divergence-free condition as well². It is reasonable to assume that the initial vorticity mean ω_0 is distributed in a small region of the space. In particular ω_0 decays to zero sufficiently fast near infinity. Under this assumption on the initial vorticity,

$$f(w; x, 0) = \frac{1}{(2\pi\sigma(x)^2)^{3/2}} \exp \left[-\frac{|w - \omega_0(x)|^2}{2\sigma(x)^2} \right]. \quad (5.21)$$

By substituting this into the representation (5.19), we obtain

$$f(w; x, t) = \mathbb{E} \left[\frac{1}{\left(2\pi\sigma(x + \sqrt{2\nu t}\xi)\right)^{3/2}} e^{-\frac{|w + Ct - \omega_0(x + \sqrt{2\nu t}\xi)|^2}{2\sigma(x + \sqrt{2\nu t}\xi)^2}} \right]. \quad (5.22)$$

It is easy to see that $f(w; x, t)$ is no longer Gaussian except for special cases which we would like to discuss below.

If ω_0 vanishes and if σ is a positive small constant, the initial distribution is homogeneous and equation (5.22) leads to a simple expression

$$f(w; x, t) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp \left(-\frac{|w + Ct|^2}{2\sigma^2} \right),$$

which is independent of the location x and remains a Gaussian density. The interesting feature about this model is that the variance stays as the constant σ^2 but new mean vorticity $-Ct$ is created evenly after duration t . This is the case of a turbulence with a small constant random perturbation. For turbulent flows observed in nature, the initial vorticity mean ω_0 does exist and does not vanish, and the random noise

² $\omega_o(x) = \mathbb{E}[W(x, 0)]$ is divergence-free, hence $\varepsilon(x) = W(x, 0) - \omega_0(x)$ must be divergence free.

$\varepsilon(x)$, for simplicity, may be modeled by a Gaussian random variable independent of x . Therefore when $\sigma > 0$ is a small constant, we have

$$f(w; x, t) = \frac{1}{(2\pi\sigma^2)^{3/2}} \mathbb{E} \left[e^{-\frac{|w+Ct-\omega_0(x+\sqrt{2\nu t}\xi)|^2}{2\sigma^2}} \right], \quad (5.23)$$

where $\xi \sim N(0, I_3)$. For this case the mean vorticity $\omega(x, t)$ at (x, t) can be evaluated. Indeed

$$\begin{aligned} \omega(x, t) &= \int_{\mathbb{R}^3} \frac{w}{(2\pi\sigma^2)^{3/2}} \mathbb{E} \left[e^{-\frac{|w+Ct-\omega_0(x+\sqrt{2\nu t}\xi)|^2}{2\sigma^2}} \right] dw \\ &= \mathbb{E} \left[\int_{\mathbb{R}^3} \frac{w}{(2\pi\sigma^2)^{3/2}} e^{-\frac{|w+Ct-\omega_0(x+\sqrt{2\nu t}\xi)|^2}{2\sigma^2}} dw \right] \\ &= \mathbb{E} \left[\omega_0 \left(x + \sqrt{2\nu t}\xi \right) \right] - Ct, \end{aligned}$$

and therefore

$$\omega^i(x, t) = \int_{\mathbb{R}^3} \frac{\omega_0^i(y)}{(4\pi\nu t)^{3/2}} e^{-\frac{|y-x|^2}{4\nu t}} dy - Ct. \quad (5.24)$$

This equality shows that the vorticity mean $\omega(x, t)$ under this simple model is independent of the noise parameter σ^2 , and $\omega(x, t)$ evolves according to the heat type equation

$$\frac{\partial}{\partial t} \omega(x, t) - C = \nu \Delta_x \omega(x, t), \quad \omega(x, 0) = \omega_0(x),$$

which is a rather crude approximation to the mean vorticity equation.

5.6 Heat flow method

In this section we propose another model for the PDF of the vorticity, based on the heat flow method, in which the conditional mean function μ is generated by a random field $R(x, \tau)$. The random field $R(x, \tau)$ evolves according to the heat flow

$$\left(\frac{\partial}{\partial \tau} - \nu \Delta_x \right) R^i(x, \tau) = 0, \quad (5.25)$$

where $R(x, \tau) = (R^1(x, \tau), R^2(x, \tau), R^3(x, \tau))$. The initial value $R(x, 0) = \xi(x) = (\xi^i(x))$ is a centered Gaussian noise³ white in space in the sense that the co-variance of ξ^i at two locations u and v in $\mathbb{R}^3$ is given by

$$\mathbb{E} [\xi^i(u) \xi^i(v)] = \delta(u - v). \quad (5.26)$$

³We remark that equation (5.26) is an informal representation. For a proper definition of White noise, see Hairer [36, Example 3.56] and Hairer [36, Section 3] on cylindrical Wiener processes.

Also, we assume that ξ^i 's are i.i.d. random variables. For simplicity, we assume that R^i is centered here, but our argument can definitely be generalized to the case when it is non-centered.

The solution to equation (5.25) is given by

$$R^i(x, \tau) = \int_{\mathbb{R}^3} \frac{1}{(4\pi\nu\tau)^{\frac{3}{2}}} e^{-\frac{|x-y|^2}{4\nu\tau}} \xi^i(y) dy, \quad (5.27)$$

so that $R(x, \tau)$ is a centered Gaussian random field with its co-variance

$$\begin{aligned} \sigma_\tau(x, y) &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{1}{(4\pi\nu\tau)^3} e^{-\frac{|x-u|^2 + |y-v|^2}{4\nu\tau}} \mathbb{E} [\xi^i(u) \xi^i(v)] du dv \\ &= \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{1}{(4\pi\nu\tau)^3} e^{-\frac{|x-u|^2 + |y-v|^2}{4\nu\tau}} \delta(u - v) du dv, \\ &= e^{-\frac{|x-y|^2}{8\nu\tau}} \int_{\mathbb{R}^3} \frac{1}{(4\pi\nu\tau)^3} e^{-\frac{|u - \frac{(x+y)}{2}|^2}{2\nu\tau}} du \\ &= \frac{1}{8(2\pi\nu\tau)^{\frac{3}{2}}} e^{-\frac{|x-y|^2}{8\nu\tau}}. \end{aligned}$$

It follows that the conditional distribution of $R(y, \tau)$ given $R(x, \tau) = w$ has a normal distribution with mean $e^{-\frac{|x-y|^2}{8\nu\tau}} w$ and co-variance matrix

$$\frac{1}{8(2\pi\nu\tau)^{\frac{3}{2}}} \left(1 - e^{-\frac{|x-y|^2}{4\nu\tau}} \right) I_3.$$

Therefore the conditional mean of $R(y, \tau) - R(x, \tau)$ given $R(x, \tau) = w$ can be easily found to be

$$- \left(1 - e^{-\frac{|x-y|^2}{8\nu\tau}} \right) w,$$

which will be our $\mu(x, y, w, t)$ with $\tau = \varphi(t)$ is a reparametrization as part of the model. For simplicity let us consider the power law model, that is,

$$\tau = \lambda_1(t + \lambda_2)^\alpha, \quad \alpha > 0, \quad (5.28)$$

where λ_1 and λ_2 are two positive numbers which then become our model parameters. Since

$$\mu(x, y, w, t) = -\frac{|x-y|^2}{8\nu\tau} w + o(|x-y|^2) w,$$

therefore $a_k^i = 0$ and $b_{jk}^i = -\delta_{jk} \frac{1}{8\nu\tau} w^i$. In particular,

$$b^i(w, t) = -\frac{3w^i}{8\nu\lambda_1(t + \lambda_2)^\alpha} \quad (5.29)$$

for $i = 1, 2, 3$. Since the conditional mean function μ depends only on $|x - y|^2$ which decays exponentially fast at infinity, therefore $B^i = 0$ and $D^i = 0$, which implies that

$$A^i(w, t) = \frac{3w^i}{4\lambda_1(t + \lambda_2)^\alpha}. \quad (5.30)$$

Thus the PDF PDE with these parameters is reduced to the simple parabolic-transport equation

$$\left(\frac{\partial}{\partial t} - \nu \Delta_x \right) f = \frac{\partial}{\partial w^i} (A^i f).$$

The divergence of A is given by

$$\frac{\partial A^i}{\partial w^i}(x, w, t) = \frac{9}{4\lambda_1(t + \lambda_2)^\alpha}.$$

If $\alpha \neq 1$ but $\alpha > 0$, by using the Feynman-Kac formula we have

$$f(w; x, t) = \theta(\alpha, t)^3 \mathbb{E} [f(Y(t, t), X(t, t), 0)], \quad (5.31)$$

where we have introduced

$$\theta(\alpha, t) = \exp \left[\frac{3((t + \lambda_2)^{1-\alpha} - \lambda_2^{1-\alpha})}{4\lambda_1(1 - \alpha)} \right] \quad (5.32)$$

for simplicity, and X and Y are solutions to the SDE

$$\begin{cases} dX^i(t, s) = \sqrt{2\nu} dM_s^i, & X(t, 0) = x, \\ dY^i(t, s) = \frac{3Y^i(t, s)}{4\lambda_1(t-s+\lambda_2)^\alpha} ds, & Y(t, 0) = w. \end{cases}$$

If $\alpha = 1$ then

$$f(w; x, t) = \lambda(t)^3 \mathbb{E} [f(Y(t, t), X(t, t), 0)], \quad (5.33)$$

where

$$\lambda(t) = \left(\frac{t + \lambda_2}{\lambda_2} \right)^{\frac{3}{4\lambda_1}}. \quad (5.34)$$

The previous SDEs have explicit solutions:

$$X(t, t) = x + \sqrt{2\nu} M_t$$

and

$$Y(t, t) = w\theta(\alpha, t)$$

if $\alpha \neq 1$, and

$$Y(t, t) = \lambda(t)w$$

if $\alpha = 1$.

Therefore, if $\alpha > 0$ and $\alpha \neq 1$, by plugging these explicit solutions X and Y into (5.31) we have

$$\begin{aligned} f(w; x, t) &= \theta(\alpha, t)^3 \mathbb{E} \left[f \left(w\theta(\alpha, t), x + \sqrt{2\nu}M_t, 0 \right) \right] \\ &= \theta(\alpha, t)^3 \mathbb{E} \left[f \left(w\theta(\alpha, t), x + \sqrt{2\nu}t\xi, 0 \right) \right] \end{aligned} \quad (5.35)$$

where $\xi \sim N(0, I_3)$.

As in the previous section, the initial vorticity is assumed to be of the form

$$W(x, 0) = \omega(x) + \varepsilon(x),$$

where $\varepsilon(x)$ has a normal distribution $N(0, \sigma^2)$ and $\omega(x)$ is the mean vorticity at x , so that

$$f(w; x, 0) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp \left[-\frac{|w - \omega(x)|^2}{2\sigma^2} \right]$$

and therefore, according to (5.35),

$$f(w; x, t) = \theta(\alpha, t)^3 \mathbb{E} \left[\frac{1}{(2\pi\sigma(x + \sqrt{2\nu}t\xi)^2)^{3/2}} e^{-\frac{|w\theta(\alpha, t) - \omega(x + \sqrt{2\nu}t\xi)|^2}{2\sigma(x + \sqrt{2\nu}t\xi)^2}} \right],$$

where $\xi \sim N(0, I_3)$. Hence $f(w; x, t)$ is no longer a Gaussian density unless $\sigma(x)$ is independent of x and $\omega = 0$ identically.

Let us discuss a special case where $\omega = 0$ identically and $\sigma > 0$ is constant. Then

$$f(w; x, t) = \frac{\theta(\alpha, t)^3}{(2\pi\sigma^2)^{3/2}} \exp \left[-\frac{|w\theta(\alpha, t)|^2}{2\sigma^2} \right]$$

is a Gaussian density with mean 0 and variance $\rho^2 I_3$, where

$$\rho^2(x, t) = \frac{1}{\sqrt{\theta(\alpha, t)}} \sigma^2. \quad (5.36)$$

If $\sigma > 0$ is a constant and ω does not vanish identically, then $f(w; x, t)$ is not Gaussian, but has a nice representation:

$$f(w; x, t) = \frac{\theta(\alpha, t)^3}{(2\pi\sigma^2)^{3/2}} \mathbb{E} \left[e^{-\frac{|w\theta(\alpha, t) - \omega(x + \sqrt{2\nu}t\xi)|^2}{2\sigma^2}} \right], \quad (5.37)$$

where $\xi \sim N(0, I_3)$, and $\theta(\alpha, t)$ is defined by (5.32).

A similar discussion applies to the model where $\alpha = 1$, which is certainly interesting too. For this model the PDF $f(w; x, t)$ is given by

$$f(w; x, t) = \lambda(t)^3 \mathbb{E} \left[f \left(w\lambda(t), x + \sqrt{2\nu}t\xi, 0 \right) \right] \quad (5.38)$$

where $\xi \sim N(0, I_3)$ and $\lambda(t)$ is defined by (5.34).

Suppose again the initial vorticity

$$W(x, 0) = \omega(x) + \varepsilon(x)$$

where $\varepsilon(x)$ has a normal distribution $N(0, \sigma^2(x))$ and $\omega(x)$ is the mean. Then

$$f(w; x, t) = \mathbb{E} \left[\frac{\lambda(t)^3}{(2\pi\sigma(x + \sqrt{2\nu t}\xi)^2)^{3/2}} e^{-\frac{|w\lambda(t) - \omega(x + \sqrt{2\nu t}\xi)|^2}{2\sigma(x + \sqrt{2\nu t}\xi)^2}} \right]. \quad (5.39)$$

which is not Gaussian except for the following special case.

If $\sigma(x) = \sigma > 0$ is a constant and $\omega(x) = 0$, then

$$f(w; x, t) = \frac{\lambda(t)^3}{(2\pi\sigma^2)^{3/2}} e^{-\frac{|w\lambda(t)|^2}{2\sigma^2}}$$

is Gaussian with mean zero and variance $\rho^2 I_3$, where

$$\rho^2(x, t) = \sigma^2 \left(\frac{\lambda_2}{t + \lambda_2} \right)^{\frac{3}{2\lambda_1}}.$$

The most interesting case for the purpose of modeling turbulent flows is the model where $\sigma > 0$ is a small parameter and ω is not zero, so that

$$f(w; x, t) = \frac{\lambda(t)^3}{(2\pi\sigma^2)^{3/2}} \mathbb{E} \left[e^{-\frac{|w\lambda(t) - \omega(x + \sqrt{2\nu t}\xi)|^2}{2\sigma^2}} \right]$$

where $\xi \sim N(0, I_3)$. Although it is no longer Gaussian, some its features can be extracted by doing Monto-Carlo simulations.

We may conclude that this model provides some nice features of the propagation of the vorticity which may be helpful for the understanding of the energy dissipation in turbulence, and we will explore this in the future work.

Appendix A

Technical Proofs and Auxiliary results

A.1 Proofs on conditional inequalities

Proof of Lemma 2.3.1. For any $A \in \mathcal{G}$, we have

$$\begin{aligned} \mathbb{E} [\mathbb{P}(|X| \geq a \mid \mathcal{G}) \mathbb{1}_A] &= \mathbb{E} [\mathbb{1}_{\{|X| \geq a\}} \mathbb{1}_A] \leq \int_{\Omega} \frac{|X|}{a} \mathbb{1}_{\{|X| \geq a\}} \mathbb{1}_A \, d\mathbb{P} \\ &\leq \frac{1}{a} \mathbb{E} [|X| \mathbb{1}_A] = \frac{1}{a} \mathbb{E} [\mathbb{E} [|X| \mid \mathcal{G}] \mathbb{1}_A]. \end{aligned}$$

□

Proof of Lemma 2.3.2. Let $\tau := \inf\{n \in [m, N] \mid X_n \geq \lambda\}$, while $\tau = N$ if $X_n \leq \lambda$ for all n and $\tau = m$ if $X_m \geq \lambda$. Since τ is a bounded stopping time, we are allowed to get

$$\begin{aligned} \mathbb{E} [X_N \mid \mathcal{F}_m] &\geq \mathbb{E} [X_\tau \mid \mathcal{F}_m] = \mathbb{E} \left[X_\tau \mathbb{1}_{\{X_{m,N}^* \geq \lambda\}} \mid \mathcal{F}_m \right] + \mathbb{E} \left[X_\tau \mathbb{1}_{\{X_{m,N}^* < \lambda\}} \mid \mathcal{F}_m \right] \\ &\geq \lambda \mathbb{E} \left[\mathbb{1}_{\{X_{m,N}^* \geq \lambda\}} \mid \mathcal{F}_m \right] + \mathbb{E} \left[X_N \mathbb{1}_{\{X_{m,N}^* < \lambda\}} \mid \mathcal{F}_m \right], \end{aligned}$$

where we applied optional stopping theorem and use the fact that $X_\tau = X_N$ on the set $\{X_{m,N}^* < \lambda\}$. The first inequality can be deduced by subtracting both sides by $\mathbb{E} [X_N \mathbb{1}_{\{X_{m,N}^* < \lambda\}} \mid \mathcal{F}_m]$, whereas the second inequality follows from basic properties of the conditional expectation. □

Proof of Lemma 2.3.3. As X_n is either a martingale or positive submartingale, $|X_n|^p$ is again a submartingale. By Lemma 2.3.2, we are able to bound the conditional probability by

$$\mathbb{P} ((X_{m,N}^*)^p \geq \lambda^p \mid \mathcal{F}_m) \leq \lambda^{-p} \mathbb{E} [|X_N|^p \mathbb{1}_{\{X_{m,N}^* \geq \lambda\}} \mid \mathcal{F}_m] \leq \lambda^{-p} \mathbb{E} [|X_N|^p \mid \mathcal{F}_m] \quad a.s..$$

If we truncate $X_{m,N}^*$ by $X_{m,N}^* \wedge k$ for some positive k , we are able to bound the conditional expectation of $(X_{m,N}^* \wedge k)^p$ by

$$\begin{aligned}
\mathbb{E} \left[(X_{m,N}^* \wedge k)^p \mid \mathcal{F}_m \right] &= \mathbb{E} \left[\int_0^{X_{m,N}^* \wedge k} py^{p-1} dy \mid \mathcal{F}_m \right] \\
&= \int_0^k py^{p-1} \mathbb{E} \left[\mathbb{1}_{\{X_{m,N}^* > y\}} \mid \mathcal{F}_m \right] dy \\
&= \mathbb{E} \left[\int_0^k py^{p-2} |X_N| \mathbb{1}_{\{X_{m,N}^* > y\}} dy \mid \mathcal{F}_m \right] \\
&= \mathbb{E} \left[|X_N| \frac{p}{p-1} (X_{m,N}^* \wedge k)^{p-1} \mid \mathcal{F}_m \right] \\
&= \frac{p}{p-1} \mathbb{E} [|X_N|^p \mid \mathcal{F}_m]^{\frac{1}{p}} \mathbb{E} \left[(X_{m,N}^* \wedge k)^p \mid \mathcal{F}_m \right]^{\frac{p-1}{p}}.
\end{aligned}$$

Hence, we obtain

$$\mathbb{E} \left[(X_{m,N}^* \wedge k)^p \mid \mathcal{F}_m \right]^p \leq \left(\frac{p}{p-1} \right)^p \mathbb{E} [|X_N|^p \mid \mathcal{F}_m],$$

while we apply conditional Monotone convergence theorem to remove the truncation k . $\square$

Proof of Theorem 2.3.4. Let $D = \mathbb{Q} \cap [t, T]$ where $\mathbb{Q}$ denotes the set of rational numbers, such that D is countable. We can construct a sequences of sets $(D_m)_{m \geq 2}$ with $D_m = \{s_0^m, \dots, s_m^m\}$ such that $t = s_0^m \leq \dots \leq s_m^m = T$, $D = \cup_{m \geq 2} D_m$ and $D_m \uparrow D$. By the previous Lemma, we already seen that

$$\lambda^p \mathbb{P} \left(\sup_{s \in D_m} |X_s| \geq \lambda \mid \mathcal{F}_t \right) \leq \left(\frac{p}{p-1} \right)^p \mathbb{E} [|X_T|^p \mid \mathcal{F}_t] \quad a.s.$$

Meanwhile, by conditional Montonic convergence theorem,

$$\mathbb{P} \left(\sup_{s \in D_m} |X_s| \geq \lambda \mid \mathcal{F}_t \right) \uparrow \mathbb{P} \left(\sup_{s \in D} |X_s| \geq \lambda \mid \mathcal{F}_t \right) \quad a.s \quad m \rightarrow \infty.$$

We observe that $\sup_{s \in D} |X_s| = \sup_{s \in [t, T]} |X_s|$, as X_s has right continuous sample paths.

The proof to part (ii) is very similar to above, as we are able to show

$$\begin{aligned}
\mathbb{E} \left[\sup_{s \in D_m} |X_s|^p \mid \mathcal{F}_m \right] &\leq \left(\frac{p}{p-1} \right)^p \mathbb{E} [|X_N|^p \mid \mathcal{F}_m] \quad a.s, \\
\mathbb{E} \left[\sup_{s \in D_m} |X_s|^p \mid \mathcal{F}_m \right] &\uparrow \mathbb{E} \left[\sup_{s \in D} |X_s|^p \mid \mathcal{F}_m \right] \quad a.s \quad m \rightarrow \infty.
\end{aligned}$$

$\square$

Proof of Lemma 2.3.6. Consider stopping times $\tau_y = \inf\{s \geq t \mid A_s > y\}$ and $S_x = \inf\{s \geq t \mid X_s > y\}$, while $\inf\{\emptyset\} = \infty$ by convention. By the continuity of A , we first observe that the event $\{A_\infty \leq y\}$ is equivalent to $\{\tau_y = \infty\}$. Moreover, the event has the inclusive relation $\{X_{t,\infty}^* > x, A_\infty \leq y\} \subset \{X_{S_x} \geq x, (\tau_y = \infty) \cap (S_x < \infty)\} \subset \{X_{S_x} \geq x, (\tau_y \wedge S_x < \infty)\} \subset \{X_{\tau_y \wedge S_x} \geq x\}$. As a consequence,

$$\mathbb{P}(X_{t,\infty}^* > x, A_\infty \leq y \mid \mathcal{F}_t) \leq \mathbb{P}(X_{\tau_y \wedge S_x} \geq x \mid \mathcal{F}_t) \leq \frac{1}{x} \mathbb{E}[X_{\tau_y \wedge S_x} \mid \mathcal{F}_t],$$

if we apply Lemma 2.3.1. Up to replacing $\tau_y \wedge S_x$ by $\tau_y \wedge S_x^u$ with $u \geq t$ and set $u \rightarrow \infty$, we are able to extend the domination relation $\mathbb{E}[X_{\tau_y \wedge S_x \wedge u} \mid \mathcal{F}_t] \leq \mathbb{E}[A_{\tau_y \wedge S_x \wedge u} \mid \mathcal{F}_t]$ to

$$\begin{aligned} \mathbb{E}[X_{\tau_y \wedge S_x} \mid \mathcal{F}_t] &\leq \liminf_{u \rightarrow \infty} \mathbb{E}[X_{\tau_y \wedge S_x \wedge u} \mid \mathcal{F}_t], \\ \mathbb{E}[X_{\tau_y \wedge S_x \wedge u} \mid \mathcal{F}_t] &\leq \lim_{u \rightarrow \infty} \mathbb{E}[A_{\tau_y \wedge S_x \wedge u} \mid \mathcal{F}_t] = \mathbb{E}[A_{\tau_y \wedge S_x} \mid \mathcal{F}_t], \end{aligned}$$

if we apply the conditional Fatou's lemma and conditional Monotonic convergence theorem. Therefore,

$$\begin{aligned} \mathbb{P}(X_{t,\infty}^* > x, A_\infty \leq y \mid \mathcal{F}_t) &\leq \frac{\mathbb{E}[A_{\tau_y \wedge S_x} \mid \mathcal{F}_t]}{x} \\ &= \frac{\mathbb{E}[A_{\tau_y} \wedge A_{S_x} \mid \mathcal{F}_t]}{x} \leq \frac{\mathbb{E}[A_\infty \wedge y \mid \mathcal{F}_t]}{x}. \end{aligned}$$

□

Proof of Lemma 2.3.7. It is sufficient to focus on the case $k \in (0, 1)$. Let $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a continuous and increasing function with $f(0) = 0$. As $\{X_{t,\infty}^* > x\} \subset (\{X_{t,\infty}^* > x, A_\infty \leq x\} \cup \{A_\infty > x\})$, we obtain

$$\begin{aligned} \mathbb{E}[f(X_{t,\infty}^*) \mid \mathcal{F}_t] &\leq \int_0^\infty \mathbb{P}(X_{t,\infty}^* > x, A_\infty \leq x \mid \mathcal{F}_t) + \mathbb{P}(A_\infty > x \mid \mathcal{F}_t) \, df(x) \\ &\leq \int_0^\infty \frac{1}{x} \mathbb{E}[A_\infty \wedge x (\mathbb{1}_{\{A_\infty \leq x\}} + \mathbb{1}_{\{A_\infty > x\}}) \mid \mathcal{F}_t] + \mathbb{P}(A_\infty > x \mid \mathcal{F}_t) \, df(x) \\ &\leq \int_0^\infty \frac{1}{x} \mathbb{E}[A_\infty (\mathbb{1}_{\{A_\infty \leq x\}}) \mid \mathcal{F}_t] + 2\mathbb{P}(A_\infty > x \mid \mathcal{F}_t) \, df(x) \\ &\leq \mathbb{E}\left[\int_0^\infty \frac{1}{x} A_\infty (\mathbb{1}_{\{A_\infty \leq x\}}) \, df(x) \mid \mathcal{F}_t\right] + \mathbb{E}\left[\int_0^\infty 2\mathbb{1}_{\{x < A_\infty\}} \, df(x) \mid \mathcal{F}_t\right] \\ &\leq \mathbb{E}\left[A_\infty \int_{A_\infty}^\infty \frac{1}{x} \, df(x) \mid \mathcal{F}_t\right] + 2\mathbb{E}[f(A_\infty) \mid \mathcal{F}_t], \end{aligned}$$

where we used Fubini theorem for conditional expectations as well as Lemma 2.3.6. What remains is to take $f(x) = x^k$, hence

$$\mathbb{E}[(X_{t,\infty}^*)^k \mid \mathcal{F}_t] \leq \mathbb{E}\left[A_\infty \int_{A_\infty}^\infty kx^{k-2} \, dx \mid \mathcal{F}_t\right] + 2\mathbb{E}[(A_\infty)^k \mid \mathcal{F}_t],$$

$$\leq \mathbb{E} \left[\frac{k}{1-k} (A_\infty)^k \mid \mathcal{F}_t \right] + 2\mathbb{E} [(A_\infty)^k \mid \mathcal{F}_t] \leq \frac{2-k}{1-k} \mathbb{E} [(A_\infty)^k \mid \mathcal{F}_t].$$

□

A.2 Proofs on BSDEs and related numerics

Proof of Proposition 2.4.5. Let $\Delta Y = Y^1 - Y^2$, $\Delta Z = Z^1 - Z^2$. The difference in Y satisfies the following BSDE

$$\begin{aligned} \Delta Y_t &= \Delta \xi + \int_t^T \Delta f(s, Y_s^2, Z_s^2) + f^1(s, Y_s^1, Z_s^1) - f^1(s, Y_s^2, Z_s^2) ds - \int_t^T \Delta Z_s \cdot dB_s \\ &= \Delta \xi + \int_t^T \Delta f(s, Y_s^2, Z_s^2) + \alpha_s \Delta Y_s + \beta_s \Delta Z_s ds - \int_t^T \Delta Z_s \cdot dB_s, \\ \alpha_s &= \frac{f^1(s, Y_s^1, Z_s^1) - f^1(s, Y_s^2, Z_s^2)}{\Delta Y_s} \mathbb{1}_{\{\Delta Y_s \neq 0\}}, \\ \beta_s &= \frac{f^1(s, Y_s^2, Z_s^1) - f^1(s, Y_s^2, Z_s^2)}{|\Delta Z_s|^2} \Delta Z_s \mathbb{1}_{\{\Delta Z_s \neq 0\}}. \end{aligned}$$

The assumption on f provides bounds on α_s and β_s :

$$|\alpha_s| \leq L_f, \quad |\beta_s| \leq K_f(1 + |Z_s^1| + |Z_s^2|),$$

hence $\Gamma_t := \exp(\int_0^t \alpha_s ds)$ lies in region $[\exp(-L_f T), \exp(L_f T)]$ and $\int_0^\cdot \beta_s \cdot dB_s$ is also a BMO martingale. Applying Itô's formula on $\Gamma_t \Delta Y_t$ yields

$$\Gamma_t \Delta Y_t = \Gamma_T \Delta \xi + \int_t^T \Gamma_s \Delta f(s, Y_s^2, Z_s^2) ds - \int_t^T \Gamma_s \Delta Z_s \cdot dB_s^\mathbb{Q}, \quad (\text{A.1})$$

where $dB_s^\mathbb{Q} = dB_s - \beta_s$ and measure $\mathbb{Q}$ is characterized by the Radon-Nikodym derivative $\eta_t := \frac{d\mathbb{Q}}{d\mathbb{P}}|_{\mathcal{F}_t} = \mathcal{E}(\int_0^t \beta_s \cdot dB_s)$. As $\int_0^\cdot \Gamma_s \Delta Z_s \cdot dB_s^\mathbb{Q}$ is BMO under $\mathbb{Q}$, we are allowed to take the conditional expectation with respect to measure $\mathbb{Q}$,

$$|\Gamma_t \Delta Y_t| \leq \mathbb{E}^\mathbb{Q} \left[\Gamma_T |\Delta \xi| + \int_t^T \Gamma_s |\Delta f(s, Y_s^2, Z_s^2)| ds \mid \mathcal{F}_t \right],$$

followed by

$$|\Delta Y_t| \leq C_q \exp(L_f T) \mathbb{E} \left[|\Delta \xi|^{q^*} + \left(\int_t^T |\Delta f(s, Y_s^2, Z_s^2)| ds \right)^{q^*} \mid \mathcal{F}_t \right]^{\frac{1}{q^*}},$$

where q is the coefficient appears in the reverse Hölder inequality and q^* is the conjugate of q . For all $p > 1$, applying Doob's maximum inequality yields,

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |\Delta Y_t|^p \right] \leq \mathbb{E} \left[\sup_{0 \leq t \leq T} |\Delta Y_t|^{pq^*} \right]^{\frac{1}{q^*}}$$

$$\leq C_{p,q} \mathbb{E} \left[|\Delta \xi|^{pq^*} + \left(\int_0^T |\Delta f(s, Y_s^2, Z_s^2)| \, ds \right)^{pq^*} \right]^{\frac{1}{q^*}}.$$

To establish a control on $\|\Delta Z\|_{H_T^p}$, we rearrange the equation (A.1), raise both sides to the power of p and set $t = 0$,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\left(\int_0^T \Gamma_s^2 |\Delta Z_s|^2 \, ds \right)^{\frac{p}{2}} \right] &= C_p \mathbb{E}^{\mathbb{Q}} \left[\left(\int_0^T \Gamma_s \Delta Z_s \cdot dB_s^{\mathbb{Q}} \right)^p \right] \\ &= C_p \mathbb{E}^{\mathbb{Q}} \left[\left(\Gamma_T \Delta \xi + \int_0^T \Gamma_s \Delta f(s, Y_s^2, Z_s^2) \, ds - \Gamma_0 \Delta Y_0 \right)^p \right] \end{aligned}$$

where we used Burkholder-Davis-Gundy's inequality. As an application of the bound on Γ_t and $\Gamma_0 Y_0$, we get

$$\begin{aligned} &\mathbb{E}^{\mathbb{Q}} \left[\left(\int_0^T |\Delta Z_s|^2 \, ds \right)^{\frac{p}{2}} \right] \\ &\leq C_p \mathbb{E}^{\mathbb{Q}} \left[|\Delta \xi|^p + \left(\int_0^T |\Delta f(s, Y_s^2, Z_s^2)| \, ds \right)^p \right] + \mathbb{E}^{\mathbb{Q}} \left[\sup_{0 \leq t \leq T} |Y_t|^p \right]^{\frac{1}{q^*}} \\ &\leq C_{p,q} \mathbb{E} \left[\left(|\Delta \xi| + \int_0^T |\Delta f(s, Y_s^2, Z_s^2)| \, ds \right)^{pq^*} \right] + \mathbb{E} \left[\sup_{0 \leq t \leq T} |Y_t|^{pq^*} \right]^{\frac{1}{q^*}}. \end{aligned}$$

In the next stage, we estimate the following

$$\begin{aligned} \|\Delta Z\|_{\mathbb{H}^p(\mathbb{P}, \mathbb{R}^d)}^p &= \mathbb{E}^{\mathbb{Q}} \left[\mathcal{E} \left(- \int_0^T \beta_s \cdot dB_s^{\mathbb{Q}} \right) \left(\int_0^T |\Delta Z_s|^2 \, ds \right)^{\frac{p}{2}} \right] \\ &= C_q \mathbb{E}^{\mathbb{Q}} \left[\left(\int_0^T |\Delta Z_s|^2 \, ds \right)^{\frac{pq^*}{2}} \right]^{\frac{1}{q^*}} \\ &\leq C_{p,q} \mathbb{E} \left[\left(|\Delta \xi| + \int_0^T |\Delta f(s, Y_s^2, Z_s^2)| \, ds \right)^{p(q^*)^2} \right]^{\frac{1}{(q^*)^2}}. \end{aligned}$$

Alternatively, $\|\Delta Z\|_{\mathbb{H}^{2p}(\mathbb{R}^d)}$ might be controlled by applying Itô's formula on ΔY_0^2 to derive

$$\begin{aligned} \Delta Y_0^2 &= \Delta \xi^2 + 2 \int_0^T \Delta Y_s (\Delta f(s, Y_s^2, Z_s^2) + \alpha_s \Delta Y_s + \beta_s \Delta Z_s) \, ds \\ &\quad - 2 \int_0^T \Delta Y_s \Delta Z_s \cdot dB_s - \int_0^T |\Delta Z_s|^2 \, ds. \end{aligned}$$

By the assumption on the driver, we are allowed to control the following term

$$\int_0^T \Delta Y_s (\alpha_s \Delta Y_s + \beta_s \cdot \Delta Z_s) \, ds \leq \int_0^T L_f \Delta Y_s^2 + K_f |\Delta Y_s| |\beta_s| |\Delta Z_s| \, ds$$

$$\leq \sup_{0 \leq s \leq T} (\Delta Y_s^2) \left(L_f T + \int_0^T K_f^2 (|\Delta Y_s|^2 (1 + |Z_s^1| + |Z_s^2|)^2) ds \right) + \frac{1}{4} \int_0^T |\Delta Z_s|^2 ds.$$

We rearrange the above equation and raise two sides by power $\frac{p}{2}$ to obtain

$$\begin{aligned} \frac{1}{2} \mathbb{E} \left[\left(\int_0^T |\Delta Z_s|^2 ds \right)^{\frac{p}{2}} \right] &\leq C_p \left(\mathbb{E}[|\Delta \xi|^p] + \mathbb{E} \left[\left(\int_0^T \Delta Y_s (\Delta f(s, Y_s^2, Z_s^2)) ds \right)^{\frac{p}{2}} \right] \right. \\ &\quad + \mathbb{E} \left[\sup_{0 \leq s \leq T} (\Delta Y_s)^2 L_f T \right] + \mathbb{E} \left[\left(\int_0^T |\Delta Y_s|^2 |\beta_s|^2 ds \right)^{\frac{p}{2}} \right] \\ &\quad \left. + \mathbb{E} \left[\sup_{0 \leq s \leq T} (\Delta Y_s)^2 \right] + \mathbb{E} \left[\left(\int_0^T (\Delta Y_s \Delta Z_s)^2 ds \right)^{\frac{p}{2}} \right] \right) \\ &\leq C_p \left(\mathbb{E}[|\Delta \xi|^p] + \mathbb{E} \left[\sup_{0 \leq t \leq T} |\Delta Y_s|^p + \left(\int_0^T \Delta f(s, Y_s^2, Z_s^2) ds \right)^{\frac{p}{2}} \right] \right) + \|Y\|_{\mathbb{S}^p(\mathbb{R})}^p \\ &\quad + \mathbb{E} \left[\left(\int_0^T |\beta_s|^2 ds \right)^p \right]^{\frac{1}{2}} \mathbb{E} \left[\sup_{0 \leq s \leq T} |\Delta Y_s|^{2p} \right]^{\frac{1}{2}} \\ &\quad + \mathbb{E} \left[\left(\int_0^T |\Delta Z_s|^2 ds \right)^{\frac{p}{2}} \right]^{\frac{1}{2}} \mathbb{E} \left[\sup_{0 \leq s \leq T} |\Delta Y_s|^p \right]^{\frac{1}{2}} \end{aligned}$$

Subsequently, using $ab < \frac{1}{\lambda} a^2 + \lambda b^2$ with $\lambda = \frac{C_p}{4}$, we bound the LHS by

$$\begin{aligned} \|\Delta Z\|_{\mathbb{H}^p(\mathbb{R}^d)}^p &\leq C_p \left(\|\xi\|_{L^p}^p + \mathbb{E} \left[\left(\int_0^T |\Delta f(s, Y_s^2, Z_s^2)| ds \right)^p \right] \right) + \|Y\|_{\mathbb{S}^p(\mathbb{R})}^p + \|Y\|_{\mathbb{S}^{2p}(\mathbb{R})}^p \\ &\leq C_{p,q} \mathbb{E} \left[|\Delta \xi|^{2pq^*} + \left(\int_0^T |\Delta f(s, Y_s^2, Z_s^2)| ds \right)^{2pq^*} \right]^{\frac{1}{2q^*}} \end{aligned}$$

Finally, combining the estimation of ΔY and ΔZ , we obtain the desired result. $\square$

Proof of Lemma 3.2.2. Set $\Delta Y_t = Y^1 - Y^2$, $\Delta \xi = \xi^1 - \xi^2$ and define $\Gamma_t = \exp(\int_0^t \alpha_s ds)$ where

$$\begin{aligned} \alpha_s &= \frac{f^1(s, Y_s^1, Z_s^1) - f^1(s, Y_s^2, Z_s^1)}{\Delta Y_s} \mathbb{1}_{\{\Delta Y_s \neq 0\}}, \\ \beta_s &= \frac{f^1(s, Y_s^2, Z_s^1) - f^1(s, Y_s^2, Z_s^2)}{|\Delta Z_s|^2} \Delta Z_s \mathbb{1}_{\{\Delta Z_s \neq 0\}}. \end{aligned}$$

As a consequence of equation (A.1), we take conditional expectations on both sides under measure $\mathbb{Q}$ to deduce

$$\Gamma_t \Delta Y_t = \mathbb{E}^{\mathbb{Q}} \left[\Gamma_T \Delta \xi + \int_t^T \Gamma_s \Delta f(s, Y_s^2, Z_s^2) ds \mid \mathcal{F}_t \right] \geq 0,$$

where measure $\mathbb{Q}$ is characterised by the Radon-Nikodym derivative

$$\left. \frac{\partial \mathbb{Q}}{\partial \mathbb{P}} \right|_{\mathcal{F}_t} = \mathcal{E} \left(\int_0^t \beta_s \cdots dB_s \right).$$

□

Proof of Lemma 3.2.9. The equality is a direct consequence of Tower property, hence we focus on the inequality. Suppose $\sum_{i=1}^n |b_i| = 1$, then $|b_i| \leq 1$ for all $1 \leq i \leq n$ and $|b_i|^p \leq |b_i| \leq |b_i|^{\frac{1}{p}}$ for any $p \geq 1$.

$$\sum_{i=1}^n |b_i|^p \leq \sum_{i=1}^n |b_i| \leq \sum_{i=1}^n |b_i|^{\frac{1}{p}}.$$

Substituting $b_i = \frac{|a_i|}{\sum_{i=1}^n |a_i|} \geq 0$, we arrive at

$$\frac{\sum_{i=1}^n |a_i|^p}{(\sum_{i=1}^n |a_i|)^p} \leq 1 \leq \frac{\sum_{i=1}^n |a_i|^{\frac{1}{p}}}{(\sum_{i=1}^n |a_i|)^{\frac{1}{p}}}$$

as we requested. □

Proof of Lemma 3.2.11. By Lemma 2.3.16, it is suffice to derive a bound on the following random variable,

$$\mathbb{E} \left[\sum_{j=k}^{N-1} |\bar{Z}_j|^2 h_j \mid \mathcal{F}_{t_k} \right] = \mathbb{E} \left[\int_{t_k}^T |Z_s|^2 ds \mid \mathcal{F}_{t_k} \right] \leq \left\| \int_0^\cdot Z_s \cdot dB_s \right\|_{\text{BMO}_2(\mathbb{R}^d)} \leq C.$$

□

Proof of Proposition 3.2.12. We prove this theorem in a similar approach to the proof of Theorem 3.1.1 and 3.1.2.

Step 1. Assume the terminal condition is of same type with ξ in the condition of Theorem 3.1.1. As $\chi_i \in [\epsilon, 2 - \epsilon]$ for all $0 \leq i \leq N - 1$, we are allowed to replace $|Z_s - \bar{Z}(s)|^2$ by

$$\begin{aligned} |Z_s - \bar{Z}(s)|^2 &\leq 2|Z_s - Z_{t_j}|^2 + 2 \left| Z_{t_j} - \frac{1}{h_j} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} Z_u^n du \mid \mathcal{F}_{t_j} \right] \right|^2 \\ &\leq 2|Z_s - Z_{t_j}|^2 + \frac{2}{h_j \epsilon} \mathbb{E}^{\mathbb{Q}_N} \left[\int_{t_j}^{t_{j+1}} |Z_u^n - Z_{t_j}|^2 du \mid \mathcal{F}_{t_j} \right], \end{aligned}$$

since the integral within the conditional expectation is $\mathcal{F}_{t_{j+1}}$ -measurable. Jensen's inequality implies

$$\begin{aligned} I &:= \mathbb{E} \left[\sup_{1 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \left(\int_{t_j}^{t_{j+1}} |Z_t - \bar{Z}(t_j)|^2 dt \right)^v \middle| \mathcal{F}_{t_i} \right]^p \right] \\ &\leq C(\epsilon, p, v) h^{p(v-1)} F, \\ F &:= \mathbb{E} \left[\sup_{1 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \int_{t_j}^{t_{j+1}} |Z_t - Z_{t_j}|^{2v} dt \middle| \mathcal{F}_{t_i} \right]^p \right]. \end{aligned} \quad (\text{A.2})$$

It is suffice to assume the function g , the drift $(y, z) \mapsto f(t, y, z)$, the drift $x \mapsto b(t, x)$ and volatility term $x \mapsto \sigma(t, x)$ are all continuously differentiable. In the first scenario, we assume π_1 is a partition wider than (or equivalent to) π_0 . Without loss of generality, we suppose $t_i = s_{l_i}$ for some increasing sequence $(l_i)_{i=1}^N$. For any $t \in [s_k, s_{k+1}) \in [s_{l_j}, s_{l_{j+1}}) = [t_j, t_{j+1})$, we are able to deduce

$$\begin{aligned} Z_t - Z_{t_j} &= (\nabla^k Y_t)^\top (\nabla_x X_t)^{-1} \sigma(t, X_t) - (\nabla^{l_j} Y_{s_{l_j}})^\top (\nabla_x X_{s_{l_j}})^{-1} \sigma(s_{l_j}, X_{s_{l_j}}) \\ &= I_{1,j}(t) + I_{2,j}(t) + I_{3,j}(t), \\ I_{1,j}(t) &:= (\nabla^k Y_t)^\top (\nabla_x X_t)^{-1} (\sigma(t, X_t) - \sigma(t_j, X_{t_j})), \\ I_{2,j}(t) &:= (\nabla^{l_j} Y_t)^\top \left((\nabla_x X_t)^{-1} - (\nabla_x X_{t_j})^{-1} \right) \sigma(t_j, X_{t_j}), \\ I_{3,j}(t) &:= (\nabla^k Y_t - \nabla^{l_j} Y_{t_j})^\top (\nabla_x X_{t_j})^{-1} \sigma(t_j, X_{t_j}). \end{aligned}$$

Therefore, we have

$$\begin{aligned} F &\leq C_{p,v} (F_1 + F_2 + F_3), \\ F_1 &= \mathbb{E} \left[\sup_{1 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \int_{t_j}^{t_{j+1}} |I_{1,j}(t)|^{2v} dt \middle| \mathcal{F}_{t_i} \right]^p \right], \\ F_2 &= \mathbb{E} \left[\sup_{1 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \int_{t_j}^{t_{j+1}} |I_{2,j}(t)|^{2v} dt \middle| \mathcal{F}_{t_i} \right]^p \right], \\ F_3 &= \mathbb{E} \left[\sup_{1 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \int_{t_j}^{t_{j+1}} |I_{3,j}(t)|^{2v} dt \middle| \mathcal{F}_{t_i} \right]^p \right]. \end{aligned}$$

Denoting $h_j := t_{j+1} - t_j$, we apply Hölder's inequality to compute the contribution of $I_1(t)$

$$F_1 \leq C_{p,q} \mathbb{E} \left[\sup_{1 \leq i \leq N-1} \mathbb{E} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} |I_{1,j}(t)|^{2v} dt \right)^{q^*} \middle| \mathcal{F}_{t_i} \right]^{2p} \right]^{\frac{1}{2q^*}}$$

$$\begin{aligned}
&\leq C_{p,q} \mathbb{E} \left[\left(\int_0^T \sum_{j=0}^{N-1} |I_{1,j}(t)|^{2v} \mathbb{1}_{[t_j, t_{j+1})}(t) dt \right)^{2pq^*} \right]^{\frac{1}{2q^*}} \\
&\leq C_{p,q} \left(\int_0^T \mathbb{E} \left[\sum_{j=0}^{N-1} |I_{1,j}(t)|^{4vpq^*} \mathbb{1}_{[t_j, t_{j+1})}(t) \right] dt \right)^{\frac{1}{2q^*}} \leq C_{p,q} h^{vp},
\end{aligned}$$

where the last line has made use of step 2 in the proof of Theorem 3.1.1. The contribution of $I_{2,j}(t)$ can be derived accordingly.

We consider a partition finer than π_1 , namely $\pi_2 : 0 = v_0 < \dots < v_{\hat{\rho}_1} < \dots < v_{\hat{\rho}_N} = T$ and the corresponding function $(j, k) \mapsto \phi(j, k)$ such that they are defined using the same way as in step 4 of proof to Theorem 3.1.1. Denoting $\Lambda_t = (\nabla_x X_t)^{-1} \sigma(t, X_t)$ which is a L^v integrable random variable for all $v > 0$, we recall the our notation in equation (3.17) and (3.18). The last component can be regarded as

$$\begin{aligned}
I_3(i) &:= \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \int_{t_j}^{t_{j+1}} |I_{3,j}(t)|^{2v} dt \mid \mathcal{F}_{t_i} \right] \\
&\leq \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\int_{t_j}^{t_{j+1}} |I_{3,j}(t)|^{2v} dt \mid \mathcal{F}_{t_j} \right] \mid \mathcal{F}_{t_i} \right] \\
&\leq (2 - \epsilon) \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} \int_{s_k}^{s_{k+1}} \mathbb{E} [|\nabla^k Y_t - \nabla^{l_j} Y_{t_j}|^{2v} \mid \mathcal{F}_{t_j}] dt \mid \Lambda_{t_j}|^{2v} \mid \mathcal{F}_{t_i} \right] \\
&\leq (2 - \epsilon) \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \sum_{k=l_j}^{l_{j+1}-1} (s_{k+1} - s_k) \mathbb{E} [|\nabla^k Y_t - \nabla^{l_j} Y_{t_j}|^{2v} \mid \mathcal{F}_{t_j}] \mid \Lambda_{t_j}|^{2v} \mid \mathcal{F}_{t_i} \right] \\
&\leq (2 - \epsilon) \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{k=0}^{\gamma-1} \rho_k \sum_{j=i}^{N-1} \mathbb{E} \left[\sum_{l=1}^4 A_{l,j,k} \mid \mathcal{F}_{t_j} \right] \mid \Lambda_{t_j}|^{2v} \mid \mathcal{F}_{t_i} \right] h \\
&\quad + (2 - \epsilon) \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \mathbb{E} \left[\left(\int_{t_j}^{t_{j+1}} |\beta(u)|^2 du \right)^{2v} \mid \Lambda_{t_j}|^{2v} \mid \mathcal{F}_{t_j} \right] \mid \mathcal{F}_{t_i} \right] h
\end{aligned}$$

where we applied inequality (3.19) and $A_{l,i}$ are defined as

$$\begin{aligned}
A_{1,j,k} &:= \left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=\phi(j,k)}^r \eta_i(u) \right|^2 du \right)^{2v}, & A_{2,j,k} &:= \left(\int_{t_j}^{t_{j+1}} \left| \sum_{i=\phi(j,k)}^r \eta_i(u) \right|^2 du \right)^v, \\
A_{3,j,k} &:= \left| (t_j - t_{j+1}) \sum_{i=\phi(j,k)}^r y_i(t_j) \right|^{2v}, & A_{4,j,k} &:= \left| \sum_{i=l_j}^{\phi(j,k)-1} y_i(t_j) \right|^{2p}.
\end{aligned}$$

Consider the consequence of Hölder's inequality

$$\left(\sum_{k=0}^{\gamma-1} M_k \rho_k \right)^p \leq \left(\sum_{k=0}^{\gamma-1} (M_k)^p \rho_k \right) \left(\sum_{k=0}^{\gamma-1} \rho_k \right)^{p-1}$$

which implies the following control on the I_3 term

$$\begin{aligned} F_3 &= \mathbb{E} \left[\sup_{0 \leq i \leq N-1} I_3(i)^p \right] \\ &\leq C_p h^p \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{k=0}^{\gamma-1} \rho_k \sum_{j=i}^{N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\frac{1}{\epsilon} \sum_{l=1}^4 A_{l,j,k} \mid \mathcal{F}_{t_j} \right] \Lambda_{t_j}^{2v} \mid \mathcal{F}_{t_i} \right]^p \right] \\ &\quad + C_p h^p \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\frac{1}{\epsilon} \left(\int_{t_j}^{t_{j+1}} |\beta(u)|^2 du \right)^{2v} |\Lambda_{t_j}|^{2v} \mid \mathcal{F}_{t_j} \right] \mid \mathcal{F}_{t_i} \right]^p \right] \\ &\leq C_{p,\epsilon} h^p \sum_{l=1}^4 \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{k=0}^{\gamma-1} \rho_k \left(\sum_{j=i}^{N-1} A_{l,j,k} \right) \sup_{0 \leq j \leq N-1} |\Lambda_{t_j}|^{2v} \mid \mathcal{F}_{t_i} \right]^p \right] \\ &\quad + C_{p,\epsilon} h^p \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\left(\int_{t_i}^T |\beta(u)|^2 du \right)^{2v} \sup_{0 \leq j \leq N-1} |\Lambda_{t_j}|^{2v} \mid \mathcal{F}_{t_i} \right]^p \right] \\ &\leq C_{p,\epsilon} h^p \sum_{l=1}^4 \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E} \left[\left(\sum_{k=0}^{\gamma-1} \rho_k \left(\sum_{j=0}^{N-1} A_{l,j,k} \right) \sup_{0 \leq j \leq N-1} |\Lambda_{t_j}|^{2v} \right)^{q^*} \mid \mathcal{F}_{t_i} \right]^{2p} \right]^{\frac{1}{2}} \\ &\quad + C_{p,\epsilon} h^p \mathbb{E} \left[\left(\int_{t_i}^T |\beta(u)|^2 du \right)^{2vq^*} \sup_{0 \leq j \leq N-1} |\Lambda_{t_j}|^{2vq^*2p} \right]^{\frac{1}{2}} \\ &\leq C_{p,\epsilon} h^p \sum_{l=1}^4 \mathbb{E} \left[\left(\sum_{k=0}^{\gamma-1} \rho_k \left(\sum_{j=0}^{N-1} A_{l,j,k} \right) \sup_{0 \leq j \leq N-1} |\Lambda_{t_j}|^{2v} \right)^{2pq^*} \right]^{\frac{1}{2}} + C_{p,\epsilon,q,v} h^p \\ &\leq C_{p,\epsilon} h^p \sum_{l=1}^4 \mathbb{E} \left[\sum_{k=0}^{\gamma-1} \rho_k \left(\sum_{j=0}^{N-1} A_{l,j,k} \right)^{4pq^*} \right]^{\frac{1}{4}} \mathbb{E} \left[\sup_{0 \leq j \leq N-1} |\Lambda_{t_j}|^{4vpq^*} \right]^{\frac{1}{4}} + C_{p,\epsilon,q,v} h^p \end{aligned}$$

Within the proof of Theorem 3.1.1, the bound has already been established for

$$\sum_{k=0}^{\gamma-1} \rho_k \mathbb{E} \left[\left(\sum_{j=0}^{N-1} A_{l,j,k} \right)^{4pq^*} \right].$$

Hence we have shown that, if π_1 is finer than π ,

$$I \leq C(\epsilon, p, v) h^{p(v-1)} (C(p, q) h^{vp} + C(\epsilon, p, q, v) h^p) \leq C(\epsilon, p, q, v) h^{pv}.$$

The proof to the situation where π_1 is wider than π , is very much similar to the proof to the first case, hence we omit here.

Step 2. We briefly review the density argument. By step 6 within Theorem 3.1.1, there exist $Z_t^\epsilon \rightarrow Z_t$ $t \times \omega$ -a.s such that the argument within step 1 is applicable to Z^ϵ . Inequality (3.20) showcases that it is sufficient to consider the differences term

$$\begin{aligned} & \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E}^{\mathbb{Q}_N} \left[\sum_{j=i}^{N-1} \int_{t_j}^{t_{j+1}} |Z_{t_j}^\epsilon - Z_{t_j}|^{2v} + |Z_t^\epsilon - Z_t|^{2v} dt \mid \mathcal{F}_{t_i} \right]^p \right] \\ & \leq \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \mathbb{E} \left[\left(\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} |Z_{t_j}^\epsilon - Z_{t_j}|^{2v} + |Z_t^\epsilon - Z_t|^{2v} dt \right)^{q^*} \mid \mathcal{F}_{t_i} \right]^{2p} \right]^{\frac{1}{2q^*}} \\ & \leq N^{2pq^*-1} h^{2pq^*-1} \sum_{j=0}^{N-1} \mathbb{E} \left[\int_{t_j}^{t_{j+1}} |Z_{t_j}^\epsilon - Z_{t_j}|^{4vpq^*} + |Z_t^\epsilon - Z_t|^{4vpq^*} dt \right]^{\frac{1}{2q^*}} \rightarrow 0 \end{aligned}$$

as $\epsilon \rightarrow 0$, by applying Dominated convergence theorem as well as Jensen's inequality. Apply a similar argument in the proof of Theorem 3.1.2, we obtain the inequality as you required. $\square$

Proof of proposition 3.3.3. By construction, $\tilde{\beta}_j^\pi \leq K_f(1 + |\tilde{Z}_j| + |Z_j^\pi|)$, hence $\|M_t\| \leq C(1 + \|\tilde{Z}(\cdot)\|_{\mathbb{H}_{\text{BMO}}^2(\mathbb{R}^d)} + \|Z^\pi(\cdot)\|_{\mathbb{H}_{\text{BMO}}^2(\mathbb{R}^d)})$ for some C depending on K_f . Meanwhile, M_t is a discrete process, which suggests $M_t = \sum_{0 \leq s \leq t} (M_s - M_{s-})$ and the stochastic exponential is of the form

$$\mathcal{E}(M)_t = \exp(M_t) \prod_{0 \leq s \leq t} ((1 + M_s - M_{s-}) \exp(-(M_s - M_{s-}))) = \tilde{\eta}^\pi(t).$$

By construction, we conclude our result by Theorem 2.3.14. $\square$

Proof of Proposition 3.3.4. The proof has adopted the strategy of the proof of [12, Proposition 2.12]. Writing $\tilde{Y}_i$ without the conditional expectation, we have

$$\tilde{Y}_i = \tilde{Y}_{i+1} + h_i f_i \left(t_i, \mathbb{E} \left[\tilde{Y}_{i+1} \mid \mathcal{F}_{t_i} \right], \tilde{Z}_i \right) + \zeta_i^Y - \frac{h_i}{c_i} (\tilde{Z}_i)^\top H_i - \tilde{v}_i.$$

$\tilde{v}_i$ is an $\mathcal{F}_{t_{i+1}}$ -measurable random variable such that $\mathbb{E}[\tilde{v}_i | \mathcal{F}_{t_i}] = 0$, $\mathbb{E}[(\tilde{v}_i)^2 | \mathcal{F}_{t_i}] < \infty$ and $\mathbb{E}[\tilde{v}_i H_i | \mathcal{F}_{t_i}] = 0$. Together with equation (3.24), we deduce that

$$\begin{aligned} \tilde{\Delta} Y_i^\pi &:= Y_i^\pi - \tilde{Y}_i = \tilde{\Delta} Y_{i+1}^\pi + A_i - \frac{h_i}{c_i} \tilde{\Delta} Z_i^\pi \cdot H_i - \tilde{\Delta} v_i, \\ A_i &:= F_i \left(t_i, \mathbb{E} \left[\tilde{Y}_{i+1} \mid \mathcal{F}_{t_i} \right], \tilde{Z}_i \right) h_i - f_i \left(t_i, \mathbb{E} \left[Y_{i+1}^\pi \mid \mathcal{F}_{t_i} \right], Z_i^\pi \right) h_i \\ &= h_i \tilde{\alpha}_j^\pi \mathbb{E} \left[\tilde{\Delta} Y_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] + h_i \tilde{\beta}_j^\pi \tilde{\Delta} Z_i^\pi + \zeta_i^Y, \end{aligned}$$

where $\tilde{\Delta}v_i := \tilde{v}_i - v_i$ and $\tilde{\alpha}_i^\pi$ is the change in Y component of f_i , associated with $\tilde{\beta}_i^\pi$. We get

$$\begin{aligned} (\tilde{\Delta}Y_{i+1}^\pi)^2 &= (\tilde{\Delta}Y_i^\pi)^2 + 2\tilde{\Delta}Y_i^\pi(\tilde{\Delta}Y_{i+1}^\pi - \tilde{\Delta}Y_i^\pi) + (\tilde{\Delta}Y_{i+1}^\pi - \tilde{\Delta}Y_i^\pi)^2 \\ &= (\tilde{\Delta}Y_i^\pi)^2 + 2\tilde{\Delta}Y_i^\pi \left(\frac{h_i}{c_i} \tilde{\Delta}Z_i^\pi \cdot H_i + \tilde{\Delta}v_i - A_i \right) + \left(\frac{h_i}{c_i} \tilde{\Delta}Z_i^\pi \cdot H_i \right)^2 + (\tilde{\Delta}v_i)^2 \\ &\quad + (A_i)^2 - 2A_i \frac{h_i}{c_i} \tilde{\Delta}Z_i^\pi \cdot H_i - 2A_i \tilde{\Delta}v_i + 2 \frac{h_i}{c_i} \tilde{\Delta}Z_i^\pi \cdot H_i \tilde{\Delta}v_i \end{aligned}$$

while taking the conditional expectation yields

$$\begin{aligned} \mathbb{E} \left[(\tilde{\Delta}Y_{i+1}^\pi)^2 \mid \mathcal{F}_{t_i} \right] &\geq \mathbb{E} \left[(\tilde{\Delta}Y_i^\pi)^2 + 2\tilde{\Delta}Y_i^\pi (-A_i) + \left(\frac{h_i}{c_i} \tilde{\Delta}Z_i^\pi \cdot H_i \right)^2 \mid \mathcal{F}_{t_i} \right] \\ &= (\tilde{\Delta}Y_i^\pi)^2 - 2\tilde{\Delta}Y_i^\pi \left(h_i \tilde{\alpha}_j^\pi \mathbb{E} \left[\tilde{\Delta}Y_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] + h_i \tilde{\beta}_j^\pi \cdot \tilde{\Delta}Z_i^\pi \right) \\ &\quad - 2\tilde{\Delta}Y_i^\pi (\zeta_i^Y) + \frac{h_i}{c_i} \left| \tilde{\Delta}Z_i^\pi \right|^2. \end{aligned}$$

Consider the inequality $2ab \leq ca^2 + c^{-1}b^2$ for any positive c , the bound on c_i yields

$$\begin{aligned} \frac{1}{\Lambda} \left| \tilde{\Delta}Z_i^\pi \right|^2 h_i &\leq \mathbb{E} \left[(\tilde{\Delta}Y_{i+1}^\pi)^2 \mid \mathcal{F}_{t_i} \right] - (\tilde{\Delta}Y_i^\pi)^2 + 2\tilde{\Delta}Y_i^\pi (\zeta_i^Y) \\ &\quad + 2L_f \tilde{\Delta}Y_i^\pi \mathbb{E} \left[\tilde{\Delta}Y_{i+1}^\pi \mid \mathcal{F}_{t_i} \right] h_i + \Lambda h_i \left| \tilde{\beta}_j^\pi \right|^2 (\tilde{\Delta}Y_i^\pi)^2 + \frac{1}{2\Lambda} \left| \tilde{\Delta}Z_i^\pi \right|^2 h_i, \end{aligned}$$

followed by summing up all i and taking expectation

$$\begin{aligned} \frac{1}{2\Lambda} \mathbb{E} \left[\sum_{i=0}^{N-1} (\tilde{\Delta}Z_i^\pi)^2 h_i \right] &\leq \mathbb{E} \left[(\tilde{\Delta}Y_N^\pi)^2 - (\tilde{\Delta}Y_0^\pi)^2 \right] + 2L_f T \mathbb{E} \left[\sup_{0 \leq i \leq N-1} \left| \tilde{\Delta}Y_i^\pi \right|^2 \right] \\ &\quad + \mathbb{E} \left[\sum_{i=0}^{N-1} \Lambda h_i \left| \tilde{\beta}_j^\pi \right|^2 (\tilde{\Delta}Y_i^\pi)^2 \right] + 2\mathbb{E} \left[\sum_{i=0}^{N-1} \tilde{\Delta}Y_i^\pi \zeta_i^Y \right] \end{aligned}$$

Due to the fact that $|\tilde{\beta}_j^\pi| \leq K_f(1 + |Z_j^\pi| + |\tilde{Z}_j|)$, N_t^π is a BMO martingale by Lemma 2.3.16 and Proposition 3.3.2. Finally, we apply conditional BDG to bound

$$\begin{aligned} \mathbb{E} \left[\sum_{i=0}^{N-1} h_i \left| \tilde{\beta}_j^\pi \right|^2 (\tilde{\Delta}Y_i^\pi)^2 \right] &\leq \mathbb{E} \left[\left(\sum_{i=0}^{N-1} h_i \left| \tilde{\beta}_j^\pi \right|^2 \right)^2 \right]^{\frac{1}{2}} \mathbb{E} \left[\sup_{0 \leq i \leq N-1} (\tilde{\Delta}Y_i^\pi)^4 \right]^{\frac{1}{2}} \\ &\leq C_4 \|N^\pi\|_{\text{BMO}_4(\mathbb{R})}^2 \mathbb{E} \left[\sup_{0 \leq i \leq N-1} (\tilde{\Delta}Y_i^\pi)^4 \right]^{\frac{1}{2}} \end{aligned}$$

with some constant $C_4 \geq 0$. □

Proof of Lemma 3.4.1. Suppose we are focusing on the SDE satisfying assumption [A2] (iii). Let us introduce the Euler-Maruyama approximation $X^{E,\pi}$ and the step process X_t^π corresponding to a partition $\pi : 0 = s_0 < s_1 < \dots, < s_r = T$, defined as

$$\begin{aligned} X^{E,\pi}(t) &= x + \int_0^t b\left(\pi(u), X_{\pi(u)}^{E,\pi} du\right) + \int_0^t \sigma\left(\pi(u), X_{\pi(u)}^{E,\pi}\right) dB_u \\ X^{E,\pi}(t) &= X_{\pi(t)}^{E,\pi}, \\ \pi(u) &:= \sum_{i=0}^{r-1} t_i \mathbb{1}_{[t_i, t_{i+1})}(u) + \mathbb{1}_{\{T\}}(u). \end{aligned}$$

The estimation on the strong error of Euler-Maruyama scheme is well-known,

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t - X^{E,\pi}(t)|^{2p} \right] \leq C_p (1 + |x|^{2p}) h^p,$$

which can be found in Kloeden and Platen [51, Proposition 10.2.2]. By following a similar procedure to Zhang [90], we estimate the error of the step process

$$\begin{aligned} \sup_{0 \leq t \leq T} \mathbb{E} [|X_t - X^\pi(t)|^{2p}] &\leq C_p \sup_{0 \leq i \leq r-1} \sup_{t_i \leq t \leq t_{i+1}} \mathbb{E} [|X_t - X^{E,\pi}(t)|^{2p} + |X^{E,\pi}(t_i) - X^{E,\pi}(t)|^{2p}] \\ &\leq C_p (1 + |x|^{2p}) h^p. \end{aligned}$$

We observe that $\sup_{0 \leq s \leq 1} |B_s| = \sup_{0 \leq s \leq 1} (B_s^+ + B_s^-) \leq \sup_{0 \leq s \leq 1} B_s^+ + \sup_{0 \leq s \leq 1} B_s^-$, where $\sup_{0 \leq s \leq 1} B_s^+ = \sup_{0 \leq s \leq 1} B_s$ has the same distribution as $|B_1|$ by the reflection principle. Let us denote $(N_i)_{i \geq 1}$ are standard i.i.d Normal random variables valued in $\mathbb{R}$, classical results on N_i indicate

$$\mathbb{E} \left[|N_i|^{2p} \mathbb{1}_{\{N_i^2 \geq a^2\}} \right] \leq \frac{2}{\sqrt{2\pi}} \frac{(2p)!}{2^p p!} \frac{a^{2p} - \frac{1}{a}}{a - \frac{1}{a}}$$

for all integers $p \geq 1$, while it is valid to take $a = h \ln \left(\frac{1}{h}\right)$ in the following analysis. If we swap the supremum and expectation, we obtain a similar sharp bound at a cost of $\ln \left(\frac{1}{h}\right)^p$,

$$\begin{aligned} \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t - X^\pi(t)|^{2p} \right] &\leq \mathbb{E} \left[\sup_{0 \leq t \leq T} |X_t - X^{E,\pi}(t_i)|^{2p} + \sup_{0 \leq i \leq r-1} \sup_{t_i \leq t \leq t_{i+1}} |b(t_i, X^{E,\pi}(t_i))|^{2p} (h_i)^{2p} \right. \\ &\quad \left. + \sup_{0 \leq i \leq r-1} \sup_{t_i \leq t \leq t_{i+1}} |\sigma(t_i, X^{E,\pi}(t_i))|^{2p} |B_t - B_{t_i}|^{2p} \right] \\ &\leq C_p (1 + |x|^{2p}) \left(h^p + \mathbb{E} \left[\sup_{0 \leq i \leq r-1} \sup_{t_i \leq t \leq t_{i+1}} |B_t - B_{t_i}|^{4p} \right]^{\frac{1}{2}} \right) \end{aligned}$$

$$\begin{aligned}
&\leq C_p(1 + |x|^{2p}) \left(h^p + \mathbb{E} \left[\sup_{0 \leq i \leq r-1} h_i^{2p} |N_i|^{4p} \right]^{\frac{1}{2}} \right) \\
&\leq C_p(1 + |x|^{2p}) \left(h \ln \left(\frac{1}{h} \right) \right)^p.
\end{aligned}$$

Notice that for the last inequality holds for all $p \geq 1$, as we can raise p -norm to $[p]$ -norm, as we required. $\square$

Proof of Proposition 3.4.2. By Proposition 2.4.5, it remains to estimate

$$\begin{aligned}
&\mathbb{E} \left[\left(\int_0^T |f(s, Y_s^n, Z_s^n) - f^{(n)}(s, Y_s^n, Z_s^n)| \, ds \right)^{p\ddot{q}} \right]^{\frac{1}{\ddot{q}}} \\
&\leq \mathbb{E} \left[\left(\int_0^T K_f(1 + |Z_s^n| + |\phi_n(Z_s^n)|) |Z_s^n| \mathbb{1}_{\{|Z_s^n| > n\}} \, ds \right)^{p\ddot{q}} \right]^{\frac{1}{\ddot{q}}} \\
&\leq \mathbb{E} \left[\left(\int_0^T C(1 + |Z_s^n|^{2p_1}) \, ds \right)^{p\ddot{q}} \right]^{\frac{1}{p_1\ddot{q}}} \mathbb{E} \left[\left(\int_0^T \mathbb{1}_{\{|Z_s^n| > n\}} \, ds \right)^{p\ddot{q}} \right]^{\frac{1}{p_1^*\ddot{q}}} \\
&\leq C \mathbb{E} \left[\left(\int_0^T C(1 + |Z_s^n|^{2p_1}) \, ds \right)^{p\ddot{q}} \right]^{\frac{1}{p_1\ddot{q}}} \left(\int_0^T \mathbb{E} \left[\mathbb{1}_{\{(\int_0^T |Z_s^n|^2 \, ds)^k > (n^2T)^k\}} \right] \, ds \right)^{\frac{1}{p_1^*\ddot{q}}},
\end{aligned}$$

where the above inequality holds for any $k \geq 0$. Recall Lemma 2.3.11 and Markov inequality, we are allowed to obtain

$$\mathbb{E} \left[\mathbb{1}_{\{(\int_0^T |Z_s^n|^2 \, ds)^k > (n^2T)^k\}} \right] \leq \frac{1}{(n^2T)^k} C_k (C(Z^n, 2))^k, \quad C(Z^n, 2) := \|Z_s^n\|_{\mathbb{H}_{\text{BMO}}^2(\mathbb{R}^d)}^2$$

By the arbitrary choice of k , we derive the following proposition. In fact, the error decreases exponentially as

$$\begin{aligned}
\mathbb{E} [\mathbb{1}_{\{|Z_s^n| > n\}}] &\leq \mathbb{P} \left(\exp \left(\frac{1}{2C(Z^n, 2)} \int_0^T |Z_s^n|^2 \, ds \right) \geq \exp \left(\frac{n^2T}{2C(Z^n, 2)} \right) \right) \\
&\leq \mathbb{E} \left[\exp \left(\frac{1}{2C(Z^n, 2)} \int_0^T |Z_s^n|^2 \, ds \right) \right] \exp \left(-\frac{n^2T}{2C(Z^n, 2)} \right) \\
&\leq \frac{4}{3} \exp(-Cn^2),
\end{aligned}$$

while Theorem 2.3.10 was applied to the last line. $\square$

Proof of Proposition 3.5.3. By plugging the expression of $\mathbf{X}_i$ and $\tilde{\mathbf{X}}_i$ repeatedly, we have

$$\mathbf{X}_i - \tilde{\mathbf{X}}_i = \mathbf{X}_{i-1} - \tilde{\mathbf{X}}_{i-1} + (b(t_{i-1}, \mathbf{X}_{i-1}) - b(t_{i-1}, \tilde{\mathbf{X}}_{i-1}))h_i$$

$$\begin{aligned}
& + (\sigma(t_{i-1}, \mathbf{X}_{i-1}) - \sigma(t_{i-1}, \tilde{\mathbf{X}}_{i-1}))\sqrt{h_{i-1}}\gamma_{i-1} - \zeta_{i-1} \\
& = \mathbf{X}_0 - \tilde{\mathbf{X}}_0 + \sum_{j=0}^{i-1} (b(t_j, \mathbf{X}_j) - b(t_j, \tilde{\mathbf{X}}_j))h_j \\
& \quad + \sum_{j=0}^{i-1} (\sigma(t_j, \mathbf{X}_j) - \sigma(t_j, \tilde{\mathbf{X}}_j))\sqrt{h_j}\mathbb{E}[\gamma_j] \\
& \quad + \sum_{j=0}^{i-1} (\sigma(t_j, \mathbf{X}_j) - \sigma(t_j, \tilde{\mathbf{X}}_j))\sqrt{h_j}(\gamma_j - \mathbb{E}[\gamma_j]) - \sum_{j=0}^{i-1} \zeta_j.
\end{aligned}$$

Taking suitable norm followed by apply Jensen's inequality yields

$$\mathbb{E} \left[\sup_{0 \leq k \leq i} |\mathbf{X}_k - \tilde{\mathbf{X}}_k|^{2p} \right] \leq C_p \left(\mathbb{E} [|\mathbf{X}_0 - \tilde{\mathbf{X}}_0|^{2p}] + \mathbb{E} \left[\sup_{0 \leq k \leq i} \left| \sum_{j=0}^{i-1} \zeta_j \right|^{2p} \right] + I_1 + I_2 + I_3 \right)$$

where I_1, I_2, I_3 are defined as

$$\begin{aligned}
I_1 &:= \mathbb{E} \left[\sup_{0 \leq k \leq i} \left| \sum_{j=0}^{k-1} (\sigma(t_j, \mathbf{X}_j) - \sigma(t_j, \tilde{\mathbf{X}}_j))\sqrt{h_j}\mathbb{E}[\gamma_j] \right|^{2p} \right], \\
I_2 &:= \mathbb{E} \left[\sup_{0 \leq k \leq i} \left| \sum_{j=0}^{k-1} (b(t_j, \mathbf{X}_j) - b(t_j, \tilde{\mathbf{X}}_j))h_j \right|^{2p} \right], \\
I_3 &:= \mathbb{E} \left[\sup_{0 \leq k \leq i} \left| \sum_{j=0}^{k-1} (\sigma(t_j, \mathbf{X}_j) - \sigma(t_j, \tilde{\mathbf{X}}_j))\sqrt{h_j}(\gamma_j - \mathbb{E}[\gamma_j]) \right|^{2p} \right].
\end{aligned}$$

By the assumptions on σ , we bounded I_1 using Hölder's inequality

$$I_1 \leq \mathbb{E} \left[\left(\sum_{j=0}^{i-1} C_1 L_X |\mathbf{X}_j - \tilde{\mathbf{X}}_j| \hat{\gamma} h_j \right)^{2p} \right] \leq C_{p,T} \hat{\gamma}^{2p} \mathbb{E} \left[\sum_{j=0}^{i-1} \sup_{0 \leq k \leq j} |\mathbf{X}_k - \tilde{\mathbf{X}}_k|^{2p} h_j \right],$$

whereas I_2 can be bounded in a similar way. An application of BDG inequality for càdlàg martingales implies

$$\begin{aligned}
I_3 &\leq C_p \mathbb{E} \left[\left(\sum_{j=0}^{i-1} (\mathbf{X}_j - \tilde{\mathbf{X}}_j)^2 (\gamma_j - \mathbb{E}[\gamma_j])^2 h_j \right)^p \right] \\
&\leq C_{p,T} \mathbb{E} \left[\sum_{j=0}^{i-1} |\mathbf{X}_j - \tilde{\mathbf{X}}_j|^{2p} (\gamma_j - \mathbb{E}[\gamma_j])^{2p} h_j \right] \leq C_{p,T} \mathbb{E} \left[\sum_{j=0}^{i-1} \sup_{0 \leq k \leq j} |\mathbf{X}_k - \tilde{\mathbf{X}}_k|^{2p} h_j \right].
\end{aligned}$$

Consequently,

$$\mathbb{E} \left[\sup_{0 \leq k \leq i} |\mathbf{X}_k - \tilde{\mathbf{X}}_k|^{2p} \right] \leq C_p \left(\mathbb{E} [|\mathbf{X}_0 - \tilde{\mathbf{X}}_0|^{2p}] + \mathbb{E} \left[\left(\sum_{j=0}^{i-1} |\zeta_j| \right)^{2p} \right] \right)$$

$$\begin{aligned}
& + C_{p,T} \max(\hat{\gamma}, 1)^{2p} \sum_{j=0}^{i-1} \mathbb{E} \left[\sup_{0 \leq k \leq j} |\mathbf{X}_k - \tilde{\mathbf{X}}_k|^{2p} \right] h_j \\
& \leq C_p \exp(C_{p,T} \max(\hat{\gamma}, 1)^{2p}) \left(\mathbb{E} [|\mathbf{X}_0 - \tilde{\mathbf{X}}_0|^{2p}] + \mathbb{E} \left[\left(\sum_{j=0}^{i-1} |\zeta_j| \right)^{2p} \right] \right).
\end{aligned}$$

□

Proof of Lemma 3.5.4. Define a function $\phi^E(b, \sigma, t_i, h_i, x, y) = x + b(t_i, x)h_i + \sigma(t_i, x)y$. Beginning with replacing the Brownian increment ΔB_{t_i} by $\Delta \hat{B}_{t_i}$, we get a new scheme

$$\begin{aligned}
X_{i+1}^{\pi, (1)} &= \phi^E(b, \sigma, t_i, h_i, X_i^{\pi, (1)}, \Delta \hat{B}_{t_i}) = \phi^E(b, \sigma, t_i, h_i, X_i^{\pi, (1)}, \Delta B_{t_i}) + \zeta_i^{X, (1)}, \\
X_0^{\pi, (1)} &= X_0,
\end{aligned}$$

where $\zeta_i^{X, (1)} = \sigma(t_i, X_i^{\pi, (1)}) (\Delta \hat{B}_{t_i} - \Delta B_{t_i})$. Meanwhile, we are allowed to write $\hat{X}_i^\pi$ as

$$\begin{aligned}
\hat{X}_{i+1}^\pi &= \phi^E(b, \sigma, t_i, h_i, \hat{X}_i^\pi, \Delta \hat{B}_{t_i}) + \zeta_i^{\pi, (2)}, \quad \hat{X}_0 = X_0, \\
\zeta_i^{\pi, (2)} &= \text{Proj}_{X_0}(X_{i+1}^{\pi, (2)}) - X_{i+1}^{\pi, (2)}, \quad X_{i+1}^{\pi, (2)} := \phi^E(b, \sigma, t_i, h_i, \hat{X}_i^\pi, \Delta \hat{B}_{t_i}).
\end{aligned}$$

Since the error can be computed as

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq i \leq N} |\hat{X}_i^\pi - X_i^\pi|^p \right] \\
& \leq C_p \left(\mathbb{E} \left[\sup_{0 \leq i \leq N} |X_i^{\pi, (1)} - X_i^\pi|^p \right] + \mathbb{E} \left[\sup_{0 \leq i \leq N} |\hat{X}_i^\pi - X_i^{\pi, (1)}|^p \right] \right),
\end{aligned}$$

we apply Lemma 3.5.3, plug in the expression of $\zeta_i^{X, (1)}$, together with BDG inequality to get

$$\begin{aligned}
& \mathbb{E} \left[\sup_{0 \leq j \leq i} |X_j^{\pi, (1)} - X_j^\pi|^p \right] \\
& \leq C_p \mathbb{E} \left[\left(\sum_{j=0}^{i-1} \left| \sigma(t_j, X_j^{\pi, (1)}) \left((\Delta \hat{B}_{t_j} - \Delta B_{t_j} - \mathbb{E}[\Delta \hat{B}_{t_j}]) + \mathbb{E}[\Delta \hat{B}_{t_j}] \right) \right| \right)^p \right] \\
& \leq C_p \mathbb{E} \left[\left(\sum_{j=0}^{i-1} \left| \sigma(t_j, X_j^{\pi, (1)}) \frac{(\Delta \hat{B}_{t_j} - \Delta B_{t_j} - \mathbb{E}[\Delta \hat{B}_{t_j}])^2}{h_j} h_j \right| \right)^{\frac{p}{2}} \right] \\
& \quad + C_{p,d} \mathbb{E} \left[\left(\sum_{j=0}^{i-1} \left| \sigma(t_j, X_j^{\pi, (1)}) \frac{\mathbb{E}[\Delta \hat{B}_{t_j}]}{h_j} h_j \right| \right)^p \right]
\end{aligned}$$

$$\begin{aligned}
&\leq C_{p,T} \mathbb{E} \left[\sum_{j=0}^{i-1} \left(C_p + \sup_{0 \leq k \leq j} |X_{t_k}^\pi|^p + \sup_{0 \leq k \leq j} |\hat{X}_k^\pi - X_k^\pi|^p \right) \frac{\mathbb{E} [|\Delta B_{t_j} - \Delta \hat{B}_{t_j}|^p]}{(h_j)^p} h_j \right] \\
&\leq C_{p,T} \sup_{0 \leq k \leq i} h_k^{-\frac{p}{2}} M^{-\frac{p}{d}} \exp \left(C_{p,T} \sum_{k=0}^{i-1} h_k^{-\frac{p}{2}} M^{-\frac{p}{d}} h_k \right)
\end{aligned}$$

where we also used Hölder's inequality as well as the fact that

$$\left| \mathbb{E} [\Delta \hat{B}_{t_j}] \right| \leq \mathbb{E} \left[|\Delta B_{t_j} - \Delta \hat{B}_{t_j}|^p \right]^{\frac{1}{p}}.$$

Observing the perturbation is controlled by

$$\zeta_i^{\pi,(2)} \leq \delta + \frac{(X_i^{\pi,(2)} - X_0)^{r+1}}{\kappa^r \delta^r} \mathbb{1}_{\{|X_i^{\pi,(2)} - X_0| > \kappa \delta\}}$$

for all $r > 0$, we directly get

$$\begin{aligned}
&\mathbb{E} \left[\sup_{0 \leq i \leq N} |\hat{X}_i^\pi - X_i^{\pi,(1)}|^p \right] \\
&\leq C_p \exp(C_{d,p,T} \tilde{\gamma}) \left(N^p \delta^p + N^p \mathbb{E} \left[\sup_{0 \leq i \leq N} |X_i^{\pi,(2)} - X_0|^{p(r+1)} \right] (\kappa \delta)^{-rp} \right),
\end{aligned}$$

where $\tilde{\gamma} = \max(\sup_{0 \leq i \leq N} M^{-\frac{1}{d}} h_i^{-\frac{1}{2}}, 1)$. In order to control the error, it is sufficient to prove the moments of $\sup_{0 \leq i \leq N} |\hat{X}_i^\pi|$ are bounded by a value independent of N .

Consider $X_i^{\pi,(3)} := \hat{X}_i^\pi - X_0$ which satisfies the numerical scheme

$$\begin{aligned}
X_{i+1}^{\pi,(3)} &= \text{Proj}_0 \left(\phi^E(\hat{b}, \hat{\sigma}, t_i, h_i, X_i^{\pi,(3)}, \Delta \hat{B}_{t_i}) \right), \quad X_0^{\pi,(3)} = 0, \\
\hat{b}(t, x) &= b(t, x + X_0), \quad \hat{\sigma}(t, x) = \sigma(t, x + X_0).
\end{aligned}$$

By definition of projection, we define $\tilde{b}(t_i, x) = \hat{b}(t_i, x) + \frac{\sqrt{d}\delta}{2h_i}$ and derive

$$\begin{aligned}
|X_{i+1}^{\pi,(3)}|^2 &\leq \left| \phi^E(\tilde{b}, \hat{\sigma}, t_i, h_i, X_i^{\pi,(3)}, \Delta \tilde{B}_{t_i}) + \frac{\sqrt{d}}{2} \delta \right|^2 \\
&\leq \sum_{j=0}^i |\tilde{b}(t_j, X_j^{\pi,(3)}) h_j|^2 + \sum_{j=0}^i |\hat{\sigma}(t_j, X_j^{\pi,(3)}) \Delta \tilde{B}_{t_j}|^2 + \sum_{j=0}^i 2X_j^{\pi,(3)} \tilde{b}(t_j, X_j^{\pi,(3)}) h_j \\
&\quad + \sum_{j=0}^i 2X_j^{\pi,(3)} \hat{\sigma}(t_j, X_j^{\pi,(3)}) \Delta \tilde{B}_{t_j} + \sum_{j=0}^i 2\tilde{b}(t_j, X_j^{\pi,(3)}) h_j \hat{\sigma}(t_j, X_j^{\pi,(3)}) \Delta \tilde{B}_{t_j}.
\end{aligned}$$

Following a similar approach as the proof of Proposition 3.5.3, we raise the power to $\frac{p}{2}$ to have

$$\mathbb{E} \left[\sup_{0 \leq j \leq i} |X_j^{\pi,(3)}|^p \right]$$

$$\begin{aligned}
&\leq C_{p,T} \left(C + \sup_{0 \leq i \leq N-1} (\delta h_i^{-1})^{\frac{p}{2}} + \sup_{0 \leq i \leq N-1} (\mathbb{E}[\Delta \hat{B}_{t_i}] h_i^{-1})^{\frac{p}{2}} + \sup_{0 \leq i \leq N-1} (\delta \mathbb{E}[\Delta \hat{B}_{t_i}] h_i^{-2})^{\frac{p}{2}} \right) \\
&\quad \times \sum_{j=0}^{N-1} \mathbb{E} \left[\sup_{0 \leq i \leq j} |X_j^{\pi, (3)}|^p \right] C_p \left(C + \sup_{0 \leq i \leq N-1} (\mathbb{E}[\Delta \hat{B}_{t_i}] h_i^{-1})^{\frac{p}{2}} \right) h_i.
\end{aligned}$$

What remains is to apply discrete Gronwall inequality. □

Appendix B

Supplement Material

B.1 Malliavin calculus and Watanabe theory

Proposition B.1.1 (Closability of D and δ). $D : L^2(\Omega) \rightarrow L^2(\Omega; H)$ and $\delta : L^2(\Omega; H) \rightarrow L^2(\Omega)$ are both closed operators.

Proof. Recall that an operator $A : B_1 \rightarrow B_2$ where A is defined on $B_0 \subseteq B_1$ is closed if given a sequence of $\{x_n\}_n \in B_0$ such that

$$x_n \rightarrow x \text{ in } B_1, \quad Ax_n \rightarrow y \text{ in } B_2,$$

we must have $x \in B_0$ and $Ax = y$. By the linearity of D and δ , it is sufficient to consider a sequence $x_n \rightarrow 0$.

Given a sequence $X_n \rightarrow 0$ in $L^2(\Omega)$ with $X_n \in \mathcal{S}$ such that $DX_n \rightarrow Y$ in $L^2(\Omega; H)$, we have, for any $Z \in \mathcal{S}$ and $h \in H$

$$\begin{aligned} \mathbb{E}[Z\langle Y, h \rangle_H] &= \lim_{n \rightarrow \infty} \mathbb{E}[Z\langle DX_n, h \rangle_H] \\ &= \lim_{n \rightarrow \infty} \mathbb{E}[X_n ZW(h) - X_n \langle DZ, h \rangle_H] \rightarrow 0, \end{aligned}$$

if we apply integration by parts formula for smooth random variables together with the dominated convergence theorem. Hence, $Y = 0$ by the arbitrariness of Z, h . The similar strategy applies to $\delta(u)$: let u_n be a sequence of smooth random variables in $L^2([0, T] \times \Omega; \mathbb{R}^d)$, such that $u_n \rightarrow 0$ in $L^2([0, T] \times \Omega; \mathbb{R}^d)$ and $\delta(u_n) \rightarrow Y \in L^2(\Omega)$. By definition and repeatedly using the dominated convergence theorem (DCT), for any $F \in \mathbb{D}^{1,2}$

$$\mathbb{E}[F\delta(u)] = \lim_{n \rightarrow \infty} \mathbb{E}[F\delta(u_n)] = \lim_{n \rightarrow \infty} \mathbb{E}[\langle DF, u_n \rangle_H] \rightarrow 0,$$

which implies $Y = 0$. □

Proposition B.1.2 (Clark-Ocone, Proposition 1.3.14 in [69]). *Assuming $d = 1$ such that the filtration $(\mathcal{F}_t)_t$ is generated by one-dimensional Brownian motion B . For all $F \in \mathbb{D}^{1,2}$, we have*

$$F = \mathbb{E}[F] + \int_0^T \mathbb{E}[D_t F \mid \mathcal{F}_t] dB_t.$$

Remark B.1. *The Skorohod integral is linear*

$$\delta(c_1 u^1 + c_2 u^2) = c_1 \delta(u^1) + c_2 \delta(u^2),$$

but does not satisfy

$$\int_0^T X u_t \delta B_t \neq X \int_0^T u_t \delta B_t.$$

B.2 A note on energy inequalities

We review some classical results within potential theory, which can be found in Meyer [66].

Definition B.2.1. *A potential $(X_t)_{t \geq 0}$ is a càdlàg positive supermartingale such that*

$$\lim_{t \rightarrow \infty} \mathbb{E}[X_t] = 0,$$

i.e. $X_t \rightarrow 0$ a.s. and in the sense of L^1 .

By Doob-Meyer decomposition, we have a unique expression $X_t = M_t - A_t$ where M_t is a right-continuous martingale and A_t is a right-continuous increasing finite variation process. If X_t is uniformly integrable, then A_t is integrable as $\mathbb{E}[M_t]$ and $\mathbb{E}[X_t]$ are both bounded. As integrability of A is equivalent to the integrability of A_∞ . Therefore M_∞ exists and M_t is also a uniform integrable martingale. Writing $X_t = \mathbb{E}[M_\infty \mid \mathcal{F}_t] - A_t$, we deduce that $M_\infty = A_\infty$ by the fact that $X_\infty = 0$. This motivates us to define

Definition B.2.2. *A potential $(X_t)_{t \geq 0}$ is generated by a right-continuous increasing and integrable process A_t if $X_t = \mathbb{E}[A_\infty - A_t \mid \mathcal{F}_t]$. The generalised energy of such potential X generated by A is, for some integer $p > 1$,*

$$e_p(X) := \frac{1}{p!} \mathbb{E}[(A_\infty)^p].$$

The classical definition for energy is the above equation but choosing $p = 2$. We are now in a position to introduce the energy inequality.

Theorem B.2.3. *Given a potential X generated by a right-continuous increasing and integrable process A , the generalised energy is bounded by, provided $X_t \leq c$ a.s.,*

$$e_p(X) := \frac{1}{p!} \mathbb{E}[(A_\infty)^p] \leq c^p.$$

Given M_t is a BMO martingale, an example of application is choosing $A_t := \sup_{0 \leq s \leq t} |M_s|$, which leads to the John-Nirenberg inequality. Alternatively, setting $A_t := \langle M \rangle_t$ yields Theorem 2.3.10.

B.3 Remarks on BTZ scheme for Lipschitz BSDEs

We aim to sketch the idea of deriving the BTZ scheme. Meanwhile, there are many variants of the BTZ scheme (see for example, Lionnet et al. [60]). Suppose the terminal condition ξ can be generated, for example, $\xi = g((X_t)_{0 \leq t \leq T})$ where X satisfies an SDE. Let $\pi : 0 = t_0 < t_1 < \dots < t_N = T$ be a partition with path $(B_{t_i})_{0 \leq i \leq N}$, the approximation of terminal condition is denoted by $\tilde{\xi}^\pi$. As we are mainly interested in the backward component Y and we observe that

$$Y_{t_i} = Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(s, Y_s, Z_s) ds - \int_{t_i}^{t_{i+1}} Z_s \cdot dB_s. \quad (\text{B.1})$$

A naive Euler discretisation of BSDE is, setting $h_i := t_{i+1} - t_i$

$$Y_{t_i}^\pi = Y_{t_{i+1}}^\pi + f(t_i, Y_{t_i}^\pi, Z_{t_i}^\pi) h_i - Z_{t_i}^\pi (B_{t_{i+1}} - B_{t_i}).$$

However, the pair (Y^π, Z^π) is not necessarily adapted. To avoid this issue, we take the conditional expectation to analyse $Y_{t_i} = \mathbb{E}[Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} f(s, Y_s, Z_s) ds \mid \mathcal{F}_{t_i}]$, together with approximating the integral by a θ_1 -scheme, similar to [92] to conclude:

$$Y_{t_i}^\pi = \mathbb{E} \left[Y_{t_{i+1}}^\pi + \theta_1 f(t_i, Y_{t_i}^\pi, Z_{t_i}^\pi) h_i + (1 - \theta_1) f(t_i, Y_{t_{i+1}}^\pi, Z_{t_{i+1}}^\pi) h_i \mid \mathcal{F}_{t_{i+1}} \right].$$

Therefore, $\theta_1 = 1$ keeps the implicit term $f(t_i, Y_{t_i}^\pi, Z_{t_i}^\pi)$, whereas $\theta_1 = 0$ keeps the explicit term $f(t_i, Y_{t_{i+1}}^\pi, Z_{t_{i+1}}^\pi)$, which requires Z_N^π .

Regarding to estimating the Z component, we estimate the k -th component Z^k where $k \in \{1, \dots, d\}$. Multiplying the Brownian increment $\Delta B_{t_i}^k := B_{t_{i+1}}^k - B_{t_i}^k$ on both sides of equation (B.1) and taking conditional expectation yields

$$\begin{aligned} 0 &= \mathbb{E} [Y_{t_{i+1}} \Delta B_{t_i}^k \mid \mathcal{F}_{t_i}] + \mathbb{E} \left[\Delta B_{t_i}^k \int_{t_i}^{t_{i+1}} f(s, Y_s, Z_s) ds \mid \mathcal{F}_{t_i} \right] \\ &\quad - \mathbb{E} \left[\Delta B_{t_i}^k \int_{t_i}^{t_{i+1}} Z_s \cdot dB_s \mid \mathcal{F}_{t_i} \right] \end{aligned}$$

$$= \mathbb{E} [Y_{t_{i+1}} \Delta B_{t_i}^k \mid \mathcal{F}_{t_i}] + \mathbb{E} \left[\int_{t_i}^{t_{i+1}} f(s, Y_s, Z_s) \Delta B_{t_i}^k \, ds \mid \mathcal{F}_{t_i} \right] - \mathbb{E} \left[\int_{t_i}^{t_{i+1}} Z_s^k \, ds \mid \mathcal{F}_{t_i} \right].$$

After replacing the integral by θ_2 -approximation, we have

$$\begin{aligned} Z_{t_i}^k &\approx \mathbb{E} \left[\frac{1}{h_i} \int_{t_i}^{t_{i+1}} Z_s^k \, ds \mid \mathcal{F}_{t_i} \right] \\ &\approx \mathbb{E} \left[Y_{t_{i+1}} \frac{\Delta B_{t_i}^k}{h_i} \mid \mathcal{F}_{t_i} \right] + \mathbb{E} \left[f(t_{i+1}, Y_{t_{i+1}}, Z_{t_{i+1}})(1 - \theta_2) \Delta B_{t_i}^k \mid \mathcal{F}_{t_i} \right]. \end{aligned}$$

To summarise, numerical scheme of BSDE of backward Euler type is

$$\begin{aligned} Y_i^\pi &= \mathbb{E} [Y_{i+1}^\pi + \theta_1 f(t_i, Y_i^\pi, Z_i^\pi) h_i + (1 - \theta_1) f(t_i, Y_{i+1}^\pi, Z_i^\pi) h_i \mid \mathcal{F}_{t_i}], \\ Z_i^{\pi, k} &= \mathbb{E} \left[(Y_{i+1}^\pi + f(t_{i+1}, Y_{i+1}^\pi, Z_{i+1}^\pi)(1 - \theta_2) h_i) \frac{\Delta B_{t_i}^k}{h_i} \mid \mathcal{F}_{t_i} \right], \quad k = 1, \dots, d; \quad (\text{B.2}) \\ Y_N^\pi &= \tilde{\xi}^\pi, \quad Z_N^\pi = (D_T \xi)^\pi \mathbb{1}_{\{\theta_2 \neq 1\}}, \end{aligned}$$

where the operator D_T is relevant to the Malliavin derivative, as in Definition 2.2.2. As a dynamic programming algorithm, the scheme is partially implicit in Y and explicit in Z for any values of θ_2 . Selecting $\theta_1 = 1$ makes a implicit scheme (in Y), whereas taking $\theta_1 = 0$ imposes an explicit scheme.

This type of scheme appears in [60, 61], while some modification of such backward Euler scheme exists. One can develop an explicit inner scheme:

$$Y_i^\pi = \mathbb{E}[Y_{i+1}^\pi \mid \mathcal{F}_{t_i}] + f(t_i, \mathbb{E}[Y_{i+1}^\pi \mid \mathcal{F}_{t_i}], Z_i^\pi) h_i,$$

which avoids computing Z_N^π and can be found in Pagès and Sagna [71].

We add a remark on BTZ scheme on QBSDEs. The original BTZ scheme is designed for Lipschitz BSDEs. An intuitive idea on numerically simulating QBSDEs is to approximate the driver f by a Lipschitz version, followed by numerical simulation based on the estimation of f . However, if we attempt to apply any previous methods on Lipschitz approximation, we cannot control the rise of the Lipschitz constant L_f . Hence the rate of convergence will be $\frac{1}{\sqrt{\log(N)}}$ if N is the number of timestep.

B.4 Utility pricing in incomplete market

We introduce briefly a problem which is specified by Henderson and Liang [40]. Consider a world with three type different assets, including a tradable risk-free bond

with zero interest rate, a tradable financial index P_t and a set of observable but non-tradable asset $S = (S^1, \dots, S^d)$. For all $i \in \{1, \dots, d\}$, their dynamics are given by

$$\begin{aligned} P_t &= P_0 + \int_0^t \mu_s^P P_s \, ds + \int_0^t (\sigma_s^P P_s) \cdot dB_s, \\ S_t^i &= S_0^i + \int_0^t \mu_s^i S_s^i \, ds + \int_0^t (\sigma_s^i S_s^i) \cdot dB_s, \end{aligned}$$

where $\mu_s^P, \mu_s^i \in \mathbb{R}$ and $\sigma_s^i, \sigma_s^P \in \mathbb{R}^d$. We focusing on path-dependent contingent claims written on non-tradable asset S , where the terminal payoff is given as $\Phi(S)$ with no early exercise rights. The market is incomplete in the sense that this contingent claim cannot be replicated by asset S and bond.

Setting the $t \in [0, T]$ as the starting time, the holder of the derivative has the initial wealth X_t , while the value corresponding to λ -units of contingent claim with payoff $\Phi(S)$ is given by V_t^λ . The investor is not allowed to trade the claim before maturity, but he/she is able to invest in the tradable index with price P and riskless bond freely. Subject to purchasing V_t^λ at time t , we denote $(\mathcal{X}_s^{(\mathcal{X}_t - V_t^\lambda)})_{s \in [t, T]}$ an $\mathcal{F}_t$ -adapted process to indicate the wealth of process of the holder of the claim. Regarding the risk arising from the contingent claim written on non-tradable assets, the holder will follow some admissible strategy which allows him/her to partially hedge the risk. The investor will choose the optimal strategy in the sense that he/she maximises the expected utility of the terminal wealth under this strategy, hence we are looking for the pricing formula and hedging strategy of the financial derivative in this circumstance.

In order to solve this problem, some assumptions on the asset and the payoff has to be made on the price dynamics and utility function of the holder of the claim.

Assumption [A10]. P_t, S_t and $\Phi(S)$ satisfy

- (i) The coefficients μ_s^P, μ_s^i and σ_s^i, σ_s^P are $\mathcal{F}_t$ -predictable processes which are uniformly bounded. Moreover, the volatility term σ_s^P is uniformly elliptic, which is equivalently imposing $\|\sigma_s^P\| > \epsilon$ for some $\epsilon > 0$ independent of (s, ω) ;
- (ii) The terminal payoff $\Phi(S)$ is a bounded and non-negative functional of S .

Assumption [A11]. The investor has an exponential utility function with parameter $\gamma > 0$ with respect to the terminal wealth, where

$$U_T(x) = -\exp(-\gamma x).$$

The investor of the claim will choose an admissible strategy π to spend on risky asset P , which is a process $\pi \in \mathcal{A}_{\text{ad}}[0, T]$ with

$$\mathcal{A}_{\text{ad}}[0, T] := \left\{ (\pi_t)_{t \in [0, T]} \text{ } \mathcal{F}_t\text{-predictable process, } \pi : [0, T] \times \Omega \rightarrow \mathbb{R} \mid \pi \in \mathbb{H}_{\text{BMO}}^2(\mathbb{R}), \right. \\ \left. \exists K > \epsilon > 0 \text{ such that } \epsilon \leq \mathbb{E} \left[\exp \left(-\gamma \int_{\tau}^T \frac{\pi_s}{P_s} dP_s \right) \right] \leq K \right\}.$$

As a result, the wealth process corresponds to a fixed admissible strategy π and satisfies the following SDE:

$$\begin{aligned} \mathcal{X}_s^{(\mathcal{X}_t - V_t^\lambda)}(\pi) &= \mathcal{X}_t - V_t^\lambda + \int_t^s \frac{\pi_u}{P_u} dP_u \\ &= \mathcal{X}_t - V_t^\lambda + \int_t^s \pi_u \mu_s^P ds + \int_t^s \pi_u (\sigma_u^P P_u) \cdot dB_u. \end{aligned}$$

As the total wealth of the holder of the claim at maturity is $\mathcal{X}_T^{(\mathcal{X}_t - V_t^\lambda)} + \Phi(S)$, optimal admissible strategy subject to purchasing λ units of the claim at initial time t , $\pi^{*, \lambda}$, is the one which maximises the expected terminal utility

$$\text{ess sup}_{\pi \in \mathcal{A}_{\text{ad}}[t, T]} \mathbb{E}^{\mathbb{P}} \left[-\exp \left(\gamma \mathcal{X}_T^{(\mathcal{X}_t - V_t^\lambda)} + \gamma \Phi(S) \right) \mid \mathcal{F}_t \right].$$

Meanwhile, the investor may decide not to invest in the financial derivative, while he/she will follow the optimal strategy, $\pi^{*, 0}$ satisfying

$$\text{ess sup}_{\pi \in \mathcal{A}_{\text{ad}}[t, T]} \mathbb{E}^{\mathbb{P}} \left[-\exp \left(\gamma \mathcal{X}_T^{(\mathcal{X}_t)} \right) \mid \mathcal{F}_t \right].$$

By no-arbitrage condition, we realise $V_t^0 = 0$, hence that not investing in the claim is equivalent to taking $\lambda = 0$.

We are now in a position to state the definition for utility indifference price of the financial derivative. Formally, it is treated as the price where the purchasing the claim does not offer any expected utility at maturity than not investing in it.

Definition B.4.1. *Given λ -units of contingent claim with payoff $\Phi(S)$ and investor's utility function of terminal wealth U_T , its utility indifference price V_t^λ is the solution to*

$$\text{ess sup}_{\pi \in \mathcal{A}_{\text{ad}}[t, T]} \mathbb{E}^{\mathbb{P}} \left[U_T \left(\mathcal{X}_T^{(\mathcal{X}_t - V_t^\lambda)} + \Phi(S) \right) \mid \mathcal{F}_t \right] = \text{ess sup}_{\pi \in \mathcal{A}_{\text{ad}}[t, T]} \mathbb{E}^{\mathbb{P}} \left[U_T \left(\mathcal{X}_T^{(\mathcal{X}_t)} \right) \mid \mathcal{F}_t \right].$$

Moreover, we refer $\pi^{*, \lambda} - \pi^{*, 0}$ to the hedging strategy corresponds to the λ -units of claim with price V_t^λ .

Corresponding to this definition, Henderson and Liang [40] proposed that the problem can be formulated in terms of BSDE, while the utility indifference price V_t^λ and hedging strategy $\pi^{*,\lambda} - \pi^{*,0}$ are associated to solution a to BSDE.

Theorem B.4.2. *We impose assumption [A10] and [A11]. Consider (Y^λ, Z^λ) is the pair which solves the quadratic BSDE*

$$Y_t^\lambda = \lambda\Phi(S) + \int_0^T \theta_s \, ds + \int_t^T f(s, Z_s) \, ds - \int_t^T (Z_s^\lambda) \cdot dB_s,$$

where the driver $f(s, z)$ and θ_s is defined as

$$f(s, z) = -\frac{\gamma}{2}\|z\|^2 + \frac{\gamma}{2\|\sigma_s^P\|^2} \left| (\sigma_s^P) \cdot z - \frac{\mu_s^P}{\gamma} \right|^2 - \theta_s, \quad \theta_s = \frac{|\mu_s^P|^2}{2\gamma\|\sigma_s^P\|^2}. \quad (\text{B.3})$$

The utility difference price V_t^λ is the difference to the solution of (B.3)

$$V_t^\lambda = Y_t^\lambda - Y_t^0,$$

whereas the optimal admissible strategy for the investor who invest in λ -units of claim and not purchasing the claim are

$$\pi_t^{*,\lambda} = -\frac{(\sigma_s^P) \cdot Z_s^\lambda}{\|\sigma_s^P\|^2} + \frac{\mu_s^P}{\gamma\|\sigma_s^P\|^2}, \quad \pi_t^{*,0} = -\frac{(\sigma_s^P) \cdot Z_s^0}{\|\sigma_s^P\|^2} + \frac{\mu_s^P}{\gamma\|\sigma_s^P\|^2},$$

respectively. Consequently, the hedging strategy which corresponds to the λ -units of claim has the value

$$\pi_t^{*,\lambda} - \pi_t^{*,0} = -\frac{(\sigma_s^P) \cdot (Z_s^\lambda - Z_s^0)}{\|\sigma_s^P\|^2}$$

B.5 Other useful results

Lemma B.5.1 (Discrete Gronwall, [42]). *Let $(x_n)_{n \geq 0}$ and $(y_n)_{n \geq 0}$ be non-negative sequences. Assuming there exist $c > 0$ such that*

$$x_n \leq c + \sum_{0 \leq i \leq n-1} x_i y_i$$

for all $n \geq 0$, then every x_n satisfies

$$x_n \leq c \prod_{i=0}^{n-1} (1 + y_i) \leq c \exp \left(\sum_{i=0}^{n-1} y_i \right).$$

Bibliography

- [1] S. Alanko. Stability of Regression-Based Monte Carlo Methods for Solving Non-linear PDEs. *Communications on Pure and Applied Mathematics*, 69(5):958–980, 2016.
- [2] S. Ankirchner, P. Imkeller, G. Dos Reis, et al. Classical and variational differentiability of BSDEs with quadratic growth. *Electronic Journal of Probability*, 12: 1418–1453, 2007.
- [3] G. K. Batchelor. *The Theory of Homogeneous Turbulence*. Cambridge science classics. Cambridge University Press, Cambridge ; New York, 1982. ISBN 9780521041171.
- [4] J.-M. Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.
- [5] B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Processes and their applications*, 111(2):175–206, 2004.
- [6] P. Briand and F. Confortola. BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces. *Stochastic Processes and their Applications*, 118(5):818–838, 2008.
- [7] P. Briand and R. Elie. A simple constructive approach to quadratic BSDEs with or without delay. *Stochastic processes and their applications*, 123(8):2921–2939, 2013.
- [8] P. Briand and Y. Hu. BSDE with quadratic growth and unbounded terminal value. *Probability Theory and Related Fields*, 136(4):604–618, 2006.
- [9] P. Briand and Y. Hu. Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probability Theory and Related Fields*, 141(3-4):543–567, 2008.

- [10] P. Briand, B. Delyon, J. Mémin, et al. Donsker-type theorem for BSDEs. *Electronic Communications in Probability*, 6:1–14, 2001.
- [11] J.-F. Chassagneux, D. Crisan, et al. Runge–Kutta schemes for backward stochastic differential equations. *The Annals of Applied Probability*, 24(2):679–720, 2014.
- [12] J.-F. Chassagneux, A. Richou, et al. Numerical simulation of quadratic BSDEs. *The Annals of Applied Probability*, 26(1):262–304, 2016.
- [13] P. Cheridito, M. Stadje, et al. BSDEs and BSDEs with non-Lipschitz drivers: comparison, convergence and robustness. *Bernoulli*, 19(3):1047–1085, 2013.
- [14] A. J. Chorin. *Vorticity and Turbulence*, volume 103 of *Applied mathematical sciences*. Springer-Verlag, New York; London, 1994. ISBN 9780387941974.
- [15] S. N. Cohen and R. J. Elliott. *Stochastic Calculus and Applications*, volume 2 of *Probability and its applications*. Birkhäuser, New York, NY, New York, 2015. ISBN 9781493928668 (hardback).
- [16] P. Constantin and G. Iyer. A stochastic Lagrangian representation of the three-dimensional incompressible Navier-Stokes equations. *Communications on pure and applied mathematics*, 61(3):330–345, 2008. ISSN 0010-3640.
- [17] G.-H. Cottet and P. D. Koumoutsakos. *Vortex Methods: Theory and Practice*. Cambridge University Press, Cambridge, 2000. ISBN 9780521621861.
- [18] D. Crisan and K. Manolarakis. Solving backward stochastic differential equations using the cubature method: application to nonlinear pricing. *SIAM Journal on Financial Mathematics*, 3(1):534–571, 2012.
- [19] R. W. Darling, E. Pardoux, et al. Backwards SDE with random terminal time and applications to semilinear elliptic PDE. *The Annals of Probability*, 25(3):1135–1159, 1997.
- [20] F. Delbaen, Y. Hu, and X. Bao. Backward SDEs with superquadratic growth. *Probability Theory and Related Fields*, 150(1-2):145–192, 2011.
- [21] C. Dellacherie and P.-A. Meyer. *Probabilités et Potentiel. Chapitres Va VIII*, volume 1385 of *Actualités Scientifiques et Industrielles [Current Scientific and Industrial Topics]. Théorie des martingales. [Martingale theory]*. Hermann, Paris, revised edition, 1980.

- [22] G. Dos Reis. *On some properties of solutions of quadratic growth BSDE and applications in finance and insurance*. PhD thesis, Citeseer, 2010.
- [23] W. E and E. Vanden Eijnden. Generalized flows, intrinsic stochasticity, and turbulent transport. *Proceedings of the National Academy of Sciences - PNAS*, 97(15):8200–8205, 2000. ISSN 0027-8424.
- [24] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.
- [25] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.
- [26] L. C. Evans. *Partial Differential Equations*. Graduate studies in mathematics ; 19. American Mathematical Society, Providence, R.I., 2nd ed. edition, 2010. ISBN 9780821849743.
- [27] S. Fang and D. Luo. Constantin and Iyer’s representation formula for the Navier–Stokes equations on manifolds. *Potential analysis*, 48(2):181–206, 2017. ISSN 0926-2601.
- [28] C. Frei and G. Dos Reis. A financial market with interacting investors: does an equilibrium exist? *Mathematics and financial economics*, 4(3):161–182, 2011.
- [29] A. Friedman. *Partial Differential Equations of Parabolic Type*. Prentice-Hall, Englewood Cliffs, 1964.
- [30] M. Fujii and A. Takahashi. Quadratic–exponential growth BSDEs with jumps and their Malliavin’s differentiability. *Stochastic Processes and their Applications*, 128(6):2083–2130, 2018. ISSN 0304-4149. doi: <https://doi.org/10.1016/j.spa.2017.09.002>. URL <https://www.sciencedirect.com/science/article/pii/S0304414917302211>.
- [31] M. Fujii, A. Takahashi, and M. Takahashi. Asymptotic expansion as prior knowledge in deep learning method for high dimensional bsdes. *Asia-Pacific Financial Markets*, pages 1–18, 2017.
- [32] E. Gobet, J.-P. Lemor, X. Warin, et al. A regression-based Monte Carlo method to solve backward stochastic differential equations. *The Annals of Applied Probability*, 15(3):2172–2202, 2005.

- [33] E. Gobet, P. Turkedjiev, et al. Approximation of backward stochastic differential equations using Malliavin weights and least-squares regression. *Bernoulli*, 22(1): 530–562, 2016.
- [34] S. Graf and H. Luschgy. *Foundations of Quantization for Probability Distributions*. Lecture notes in mathematics (Springer-Verlag); 1730. Springer, Berlin; London, 2000. ISBN 9783540673941.
- [35] P. Grandits and T. Rheinländer. On the minimal entropy martingale measure. *The Annals of probability*, 30(3):1003–1038, 2002.
- [36] M. Hairer. An introduction to stochastic PDEs. *arXiv preprint arXiv:0907.4178*, 2009.
- [37] J. Han, A. Jentzen, and W. E. Overcoming the curse of dimensionality: solving high-dimensional partial differential equations using deep learning. *arXiv preprint arXiv:1707.02568*, pages 1–13, 2017.
- [38] J. Han, A. Jentzen, and W. E. Solving high-dimensional partial differential equations using deep learning. *Proceedings of the National Academy of Sciences*, 115(34):8505–8510, 2018.
- [39] S.-w. He, J.-g. Wang, and J.-a. Yan. *Semimartingale Theory and Stochastic Calculus*. Science Press; CRC Press, Beijing; Boca Raton, 1992. ISBN 9780849377150.
- [40] V. Henderson and G. Liang. Pseudo linear pricing rule for utility indifference valuation. *Finance and Stochastics*, 18(3):593–615, 2014.
- [41] P. Henry-Labordere. Deep primal-dual algorithm for BSDEs: Applications of machine learning to CVA and IM. *Available at SSRN 3071506*, 2017.
- [42] J. M. Holte. Discrete gronwall lemma and applications. In *MAA-NCS meeting at the University of North Dakota*, volume 24, pages 1–7, 2009.
- [43] E. Hopf. Statistical hydromechanics and functional calculus. *Journal of Rational Mechanics and Analysis*, 1:87–123, 1952.
- [44] Y. Hu and S. Tang. Multi-dimensional BSDE with oblique reflection and optimal switching. *Probability Theory and Related Fields*, 147(1-2):89–121, 2010.

- [45] Y. Hu, P. Imkeller, M. Müller, et al. Utility maximization in incomplete markets. *The Annals of Applied Probability*, 15(3):1691–1712, 2005.
- [46] C. Huré, H. Pham, and X. Warin. Some machine learning schemes for high-dimensional nonlinear PDEs. *arXiv preprint arXiv:1902.01599*, 2019.
- [47] P. Imkeller and G. Dos Reis. Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stochastic Processes and their Applications*, 120(3):348–379, 2010.
- [48] H. Ishii and P.-L. Lions. Viscosity solutions of fully nonlinear second-order elliptic partial differential equations. *Journal of Differential Equations*, 83(1):26–78, 1990.
- [49] Y. L. Jan and O. Raimond. Flows, coalescence and noise. *The Annals of probability*, 32(2):1247–1315, 2004. ISSN 0091-1798.
- [50] N. Kazamaki. *Continuous Exponential Martingales and BMO*. Lecture notes in mathematics (Springer-Verlag); 1579. Springer-Verlag, Berlin; London, 1994. ISBN 9783540580423.
- [51] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Applications of mathematics. Springer Science & Business Media, Berlin, corrected third printing 1999. edition, 2010. ISBN 9783642081071 (paperback).
- [52] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability*, pages 558–602, 2000.
- [53] A. N. Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. *Comptes Rendus de l’Académie des Sciences de l’URSS*, 30:301–305, 1941.
- [54] A. N. Kolmogorov. Dissipation of energy in the locally isotropic turbulence. *Comptes Rendus de l’Académie des Sciences de l’URSS*, 32:16–18, 1941.
- [55] H. Kunita. *Stochastic Flows and Stochastic Differential Equations*. Cambridge studies in advanced mathematics ; 24. Cambridge University Press, Cambridge, 1990. ISBN 9780521350501.
- [56] J.-P. Lepeltier and J. San Martin. Backward stochastic differential equations with continuous coefficient. *Statistics & Probability Letters*, 32(4):425–430, 1997.

- [57] J. Li, Z. Qian, and M. Zhou. On distributions of velocity random fields in turbulent flows. *arXiv preprint arXiv:2107.03219*, 2021.
- [58] J. Li, Z. Qian, and M. Zhou. On the transport equation for probability density functions of turbulent vorticity fields. *arXiv preprint arXiv:2104.12217*, 2021.
- [59] Y. Li and S. Tang. Approximation of bsdes with super-linearly growing generators by euler’s polygonal line method: A simple proof of the existence. *Systems & Control Letters*, 153:104952, 2021.
- [60] A. Lionnet, G. Dos Reis, L. Szpruch, et al. Time discretization of FBSDE with polynomial growth drivers and reaction–diffusion PDEs. *The Annals of Applied Probability*, 25(5):2563–2625, 2015.
- [61] A. Lionnet, G. Dos Reis, L. Szpruch, et al. Convergence and qualitative properties of modified explicit schemes for BSDEs with polynomial growth. *The Annals of Applied Probability*, 28(4):2544–2591, 2018.
- [62] J. Ma and J. Yong. Solvability of Forward-Backward SDEs and the Nodal Set of Hamilton-Jacobi-Bellman Equations. *Chinese Annals of Mathematics*, 16B(3): 279–298, 1995. URL <https://conservancy.umn.edu/handle/11299/2397>.
- [63] J. Ma and J. Zhang. Path regularity for solutions of backward stochastic differential equations. *Probability Theory and Related Fields*, 122(2):163–190, 2002.
- [64] A. Majda and A. L. Bertozzi. *Vorticity and Incompressible Flow*. Cambridge texts in applied mathematics. Cambridge University Press, Cambridge, 2002. ISBN 9780521630573.
- [65] A. McNabb. Comparison theorems for differential equations. *Journal of Mathematical Analysis and Applications*, 119(1-2):417–428, 1986.
- [66] P. A. Meyer. *Probability and Potentials*, volume 14 of *Blaisdell book in pure and applied mathematics*. Blaisdell Pub. Co., Waltham, Mass., 1966.
- [67] H. K. Moffatt. A brief introduction to vortex dynamics and turbulence. In *Environmental Hazards: The Fluid Dynamics and Geophysics of Extreme Events*, pages 1–27. World Scientific, 2011.
- [68] A. S. Monin and A. M. Yaglom. *Statistical Fluid Mechanics: Mechanics of Turbulence*, volume 2. MIT Press, 1975.

- [69] D. Nualart. *The Malliavin Calculus and Related Topics*. Probability and its applications (Springer-Verlag). Springer, Berlin, 2nd ed. edition, 2006. ISBN 9783540283287.
- [70] E. Nualart. Lectures on malliavin calculus and its applications to finance. <https://people.math.wisc.edu/~kurtz/NualartLectureNotes.pdf>, 2009.
- [71] G. Pagès and A. Sagna. Improved error bounds for quantization based numerical schemes for BSDE and nonlinear filtering. *Stochastic Processes and their Applications*, 128(3):847–883, 2018.
- [72] G. Pagès, H. Pham, and J. Printems. Optimal quantization methods and applications to numerical problems in finance. In *Handbook of computational and numerical methods in finance*, pages 253–297. Springer, 2004.
- [73] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.
- [74] E. Pardoux and S. Peng. Backward stochastic differential equations and quasi-linear parabolic partial differential equations. In *Stochastic partial differential equations and their applications*, pages 200–217. Springer, 1992.
- [75] S. Peng. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and stochastics reports*, 37(1-2):61–74, 1991.
- [76] S. B. Pope. PDF methods for turbulent reactive flows. *Progress in Energy and Combustion Science*, 11(2):119–192, 1985.
- [77] S. B. Pope. *Turbulent flows*. Cambridge University Press, Cambridge, 2000. ISBN 9780521591256.
- [78] L. Prandtl. 7. Bericht über untersuchungen zur ausgebildeten turbulenz. *Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 5(2):136–139, 1925.
- [79] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*, volume 293 of *Grundlehren der mathematischen Wissenschaften, A series of comprehensive studies in mathematics*. Springer Science & Business Media, 2005.

- [80] O. Reynolds. IV. On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Philosophical Transactions of the Royal Society of London.(A.)*, 186:123–164, 1895. doi: <http://doi.org/10.1098/rsta.1895.0004>.
- [81] A. Richou. Markovian quadratic and superquadratic BSDEs with an unbounded terminal condition. *Stochastic Processes and their Applications*, 122(9):3173–3208, 2012.
- [82] A. Richou et al. Numerical simulation of BSDEs with drivers of quadratic growth. *The Annals of Applied Probability*, 21(5):1933–1964, 2011.
- [83] P. G. Saffman. *Vortex Dynamics*. Cambridge monographs on mechanics and applied mathematics. Cambridge University Press, Cambridge, 1992. ISBN 9780521420587.
- [84] G. I. Taylor. Diffusion by continuous movements. *Proceedings of the London Mathematical Society*, 2(1):196–212, 1922.
- [85] G. I. Taylor. Statistical theory of turbulence-II. *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, 151(873):444–454, 1935. doi: <https://doi.org/10.1098/rspa.1935.0159>.
- [86] T. Von Kármán. *Mechanical Similitude and Turbulence*. National Advisory Committee for Aeronautics, 1931.
- [87] E. Weinan, J. Han, and A. Jentzen. Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, 5(4):349–380, 2017.
- [88] H. Xing and G. Zitkovic. A class of globally solvable markovian quadratic BSDE systems and applications. *The Annals of probability*, 46(1):491–550, 2018. ISSN 0091-1798.
- [89] J. Yong and X. Y. Zhou. *Stochastic Controls: Hamiltonian Systems and HJB Equations*, volume 43 of *Applications of Mathematics*. Springer Science & Business Media, New York; London, 1999. ISBN 9780387987231.
- [90] J. Zhang. A numerical scheme for BSDEs. *The Annals of Applied Probability*, 14(1):459–488, 2004.

- [91] J. Zhang. *Backward Stochastic Differential Equations: from Linear to Fully Non-linear Theory*, volume 86 of *Probability theory and stochastic modelling*. Springer New York, New York, NY, 2017. ISBN 9781493972548.
- [92] W. Zhao, Y. Li, and G. Zhang. A generalized θ -scheme for solving backward stochastic differential equations. *Discrete & Continuous Dynamical Systems-B*, 17(5):1585, 2012.