Some Contributions to Stochastic Differential Equations

Yuhan Yao

Mansfield College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas term 2021

To my parents,

Lijun Zhai and Jianxin Yao.

To my supervisor,

Professor Zhongmin Qian.

In memory of my grandparents,

Huiyuan Xiao and Guangming Zhai.

Acknowledgements

I would like to express my deepest gratitude to my supervisor Professor Zhongmin Qian for his guidance. I am very grateful to him for sharing his intuitions and knowledge, and building me up as a researcher through my doctoral studies. Without his patience, encouragement and support, this work would not have been possible.

I am very thankful to Professor Shige Peng for discussing the forward backward stochastic differential equations with me during his visit to Oxford. I would also like to thank Professor Ben Hambly and Professor Terry Lyons for valuable feedback and insightful comments in my Confirmation of Status.

I would like to thank my dear friends and colleagues Panpan Ren, Wolfgang Stockinger, Jing Ye, Mingrui Zhou, Christina Zou, with whom I got a chance to share my happiness and low spirits alike. I am so grateful to have met them, who make these years of DPhil life full of joy. I would like to thank Jiawei Li for her help and suggestions as senior DPhil student. She has been acting as a mentor to me, with her encouragement and support, I have been through the difficult times in my study. I would like to thank Alex Tetlow and Jiawei Li for helping me with the proofreading. I would like to thank Zichen Zhang for sharing his insights in careers in quantitative finance. I would also like to thank Yihuang Zhang for discussing coding and programming with me.

Finally, I would like to thank my parents Lijun Zhai and Jianxin Yao, for their endless love and support. I am very grateful for everything they have done for me, I would not be myself today without their support. I also want to thank my boyfriend Tianyu Xu for always being there with me.

Abstract

This thesis elaborates topics on a type of McKean–Vlasov stochastic differential equations and forward–backward stochastic differential equations arising from physics models.

Mainly, this thesis is divided into three parts. Chapter 2 consists of preliminary definitions and results from the theory of parabolic equations, stochastic differential equations and backward stochastic differential equations. Most of these are standard concepts and known results, which will be found in [46, 65, 75, 129], etc.

Motivated by the random vortex method for the Navier–Stokes equation, Chapter 3 studies a special type of McKean–Vlasov stochastic differential equation (MVSDE), whose kernel function is assumed to be continuous except at few singularities in $\mathbb{R}^d$. With some technical estimates, we establish the existence and uniqueness of both the weak and the strong solutions to the MVSDE theoretically. The numerical simulations are also conducted to approximate the MVSDE and the corresponding velocity field in $\mathbb{R}^3$.

Chapter 4 is devoted to studying a system of forward-backward stochastic differential equations (FBSDEs) arising from the porous medium equation. The generating function of the backward equation for this class of FBSDEs has quadratic growth in the variable Z. We construct the unique solution on a small time interval by using a fixed point argument, and further extend the study to a class of quadratic FBSDEs with more general growth conditions.

Contents

Notations

- $\mathcal{C}(\mathbb{R}^d, \mathbb{R}^k)$ is the space of continuous functions from $\mathbb{R}^d$ to $\mathbb{R}^k$.
- $\mathcal{C}_b(\mathbb{R}^d, \mathbb{R}^k)$ is the space of continuous, bounded functions from $\mathbb{R}^d$ to $\mathbb{R}^k$.
- $\mathcal{C}^\infty(\mathbb{R}^d)$ is the space of smooth functions on $\mathbb{R}^d$.
- $L^p(\mathbb{R}^d)$ for $1 \leq p \leq \infty$ is the collections of measurable functions in $\mathbb{R}^d$ with finite L^p-norm with respect to the Lebesgue measure, that is

$$\|f\|_{L^p} = \left(\int_{\mathbb{R}^d} |f|^p d\mu \right)^{1/p} < \infty, \quad 1 \leq p < \infty,$$

and

$$\|f\|_{L^\infty} = \sup_{x \in \mathbb{R}^d} |f(x)| < \infty, \quad p = \infty.$$

- $\mathscr{L}^2_T(\mathbb{R}^k)$, the vector space formed by progressively measurable processes $Y : \Omega \times [0,T] \to \mathbb{R}^k$ such that

$$\|Y\|^2_{\mathscr{L}^2_T} = \mathbb{E}\left[\sup_{0 \leq t \leq T} |Y_t|^2 \right] < \infty.$$

- $\mathscr{H}^\infty_T(\mathbb{R}^k)$, the vector space formed by bounded progressively measurable processes $Y : \Omega \times [0,T] \to \mathbb{R}^k$ such that there exists $M > 0$ with $Y_t \leq M$ a.s. for all $t \leq T$ and

$$\|Y\|^2_{\mathscr{H}^\infty_T} = \inf \{ M \geq 0 : \mathbb{P}(|Y_t| \leq M) = 1, \forall t \in [0,T] \} < \infty.$$

- $\mathscr{H}^2_T(\mathbb{R}^{k \times d})$, the vector space formed by progressively measurable processes $Z : \Omega \times [0,T] \to \mathbb{R}^{k \times d}$ such that

$$\|Z\|^2_{\mathscr{H}^2_T} = \mathbb{E}\left[\int_0^T |Z_t|^2 dt \right] < \infty.$$

- $\mathscr{X}$, the space of bounded and measurable random variables such that $\xi \in \mathscr{X}$ if and only if $c_1 \leq \xi \leq c_2$ for some constants $c_1, c_2 > 0$.

- $\mathscr{Y} \subset \mathscr{H}_T^\infty(\mathbb{R})$, the space of bounded, progressively measurable processes such that $c_1 \leq Y_t \leq c_2$ a.s. for all $0 \leq t \leq T$, where $c_1, c_2 > 0$ are constant.

- If A is a vector field (in the space of dimension d) dependent on some parameters, then its components are labelled with upper-script indices, that is
$$A = \left(A^i\right) = \left(A^1, ..., A^d\right).$$
For the coordinates, we use lower-script indices so that
$$x = (x_i) = (x_1, ..., x_d).$$

List of Figures

List of Tables

1 Introduction

In this thesis, we study two classes of stochastic differential equations arising from physical problems with original formulation in partial differential equations.

1.1 McKean–Vlasov SDE from Navier–Stokes equation

1.1.1 Navier–Stokes equation and the random vortex method

The Navier–Stokes equations are non-linear partial differential equations of second order

$$\begin{aligned} \frac{\partial}{\partial t}u + u \cdot \nabla u &= \nu \Delta u - \nabla p, \\ \nabla \cdot u &= 0, \end{aligned} \tag{1.1}$$

where $\nu > 0$ is the viscosity of the fluid flow, $u(x,t)$ is the velocity field of the incompressible fluid flow, and $p(x,t)$ is the scalar function describing the pressure at the location x and at the instant t. If the fluid occupies a finite range (of $\mathbb{R}^2$ or $\mathbb{R}^3$), then the Navier–Stokes system (1.1) must be supplied with physical boundary conditions. The Navier–Stokes equations have been studied extensively since the fundamental work by Leray [87], where a weak solution was constructed, see [29, 49, 119, 134] and reference therein for example. Nevertheless, the resolution of the three-dimensional Navier–Stokes equations remains to be an open problem (see [119, 123] for example). Most literature in this research area concentrates on the study of related partial differential equations and numerical solutions.

In particular, the vortex method serves as an important tool for studying the Navier–Stokes equations numerically, which is essentially based on the velocity–vorticity formulation [99]. In this formulation, the vorticity ω, which is defined as the curl of the velocity field, i.e. $\omega = \nabla \wedge u$, satisfies the vorticity equation

$$\frac{\partial}{\partial t}\omega + u \cdot \nabla \omega = \nu \Delta \omega + \omega \cdot \nabla u. \tag{1.2}$$

In case of inviscid incompressible flows, that is, $\nu = 0$, the vorticity $\omega(x, t)$ may be numerically expressed in terms of the particle trajectory path X_t. For example, in the point vortex method, the vorticity is expressed as

$$\omega(x, t) = \sum_{i=1}^{N} \Gamma_i \delta(x - X_t^i),$$

where δ is the Dirac delta function, Γ_i represents the circulation and X_t^i is the location of the vortex at time t for each $i = 1, ..., N$. One may refer to [30, 58, 84, 99], etc and literature therein for a detailed review of different approximations. Nevertheless, the particle locations are obtained by solving the following ordinary differential equations

$$\frac{d}{dt}X_t^i = u(X_t^i, t).$$

On the other hand, with the incompressible condition $\nabla \cdot u = 0$, the velocity field can be recovered from the vorticity field via the Poisson equation

$$\Delta u = -\nabla \wedge \omega.$$

Let G be the Green's function for the Laplace operator, and $K = \nabla G$, then one obtains the velocity field through the Biot–Savart law that

$$u(x, t) = \int K(x, y, t) \wedge \omega(y) dy,$$

where

$$G(x) = \frac{1}{2\pi} \ln(|x|), \quad K(x, y) = \frac{1}{2\pi} \frac{x - y}{|x - y|^2}$$

in $\mathbb{R}^2$, and

$$G(x) = \frac{1}{4\pi|x|}, \quad K(x,y) = -\frac{1}{4\pi}\frac{x-y}{|x-y|^3}$$

in $\mathbb{R}^3$.

In 1973, Chorin [27] proposed the *random vortex method* for viscous incompressible fluids, i.e. $\nu > 0$. This method generalized the inviscid vortex method by adding a random walk so that particle trajectory paths were modelled by stochastic differential equations. This random vortex method brings us the McKean–Vlasov type stochastic differential equation

$$\begin{cases} dX^i(x,t) = \left(\sum_{j=1}^d \int_{\mathbb{R}^d} \mathbb{E}\left[K^{ij}(z - X(y,t))\right]\big|_{z=X(x,t)}\, \omega_0^j(y)dy\right) dt + \sqrt{2\nu}dB^i(t), \\ X(x,t) = x, \quad x \in \mathbb{R}^d \end{cases} \tag{1.3}$$

where $i = 1, \dots, d$, $\nu > 0$ is a constant, $B = (B^1, \dots, B^d)$ is a standard Brownian motion on some probability space and $\omega_0 = (\omega_0^1, \dots, \omega_0^d)$ is the initial data to vorticity equation (1.2). The structure kernel function $K = (K^{ij})$ is a $d \times d$ matrix-valued Borel measurable function on $\mathbb{R}^d$, which is continuous except at several singularities. In particular, this structure kernel is the Biot–Savart kernel in $\mathbb{R}^2$ for studying the two-dimensional Navier–Stokes equations. The equation (1.3) is the stochastic differential equation we are going to study. We will explain the random vortex model and formulate the problem (1.3) more precisely in Chapter 3.

1.1.2 Study of McKean–Vlasov equations

The equation (1.3) is a system of SDEs which involves the distributions of its solutions. This type of distribution-dependent SDEs and SDEs which share the same nature may arise from physics models and applied mathematics, and they have been studied intensively over the past decades, especially in the study of mean field games. Historically, the general expositions on the McKean-Vlasov model and nonlinear processes were proposed in the preprint of [139] in 1938 by Vlasov. For the first time, Kac [73, 74] proposed the equations of the form

$$\begin{cases} dX_t^x = b(t, X_t^x, \mathscr{L}(X_t^x))dt + \sigma(t, X_t^x, \mathscr{L}(X_t^x))dB_t, \\ X_0^x = x \in \mathbb{R}^d, \end{cases} \tag{1.4}$$

where $\mathscr{L}(X^x)$ is the law of the process X^x. Kac brought up the equation (1.4) when studying the Boltzman equation for the particle density in diluted monotonic gases, and the stochastic toy model for the Vlasov kinetic equation of plasma. McKean [97] showed that equations of type (1.4) provided a probabilistic representation to the solution of Boltzman equation. Sznitman [130] provided an introduction to the *propagation of chaos* phenomenon, which dealt with symmetric evolution of particles and indicated that the study of one individual particle would give information on the behaviour of the group. Since then, the McKean–Vlasov SDEs, including their well-posedness, the theory of propagation of chaos, numerical approximations, etc. have been studied in a variety of settings. In recent years, there have been renewed interests in the study of McKean–Vlasov SDEs in the context of the mean-field games theory, which was introduced by Lasry and Lions [83] and Huang et al. [64], independently. The idea was to assign N rational players with global information, independently, then the limiting behaviour, as N tending to infinity, gave rise to an approximation of the Nash equilibrium. One can refer to [21, 24, 56] for a detailed account in this topic including recent progress.

The solvability of (1.4) has been studied under various growth conditions. For instance, under the global Lipschitz continuity assumption on the coefficients, Sznitman [130] proved the existence and uniqueness of the strong solution under the global Lipschitz assumption of the coefficients using a fixed point argument on the Wasserstein space. Jourdain et al. [72] followed this approach to study mean-field SDEs driven by Lévy processes. Carmona and Delarue [23, 22] and Buckdahn et al. [20] coupled the McKean–Vlasov SDEs with backward equations and studied the McKean–Vlasov forward backward stochastic differential equations. Regardless of the complexity of the problem, this fixed point argument takes advantage of the Lipschitz regularity of the coefficients upon the measure argument with respect to the Wasserstein distance.

Another typical work for studying McKean–Vlasov SDEs was conducted by Funaki [48], where a martingale problem was formulated to associate diffusion processes with a kind of non-linear parabolic equations. In this work, Funaki established the existence of the solution under the continuity assumption and Lyapunov type condition. Using the same methodology of martingale problem, Gärtner [50] showed the existence and uniqueness of the weak solution of (1.4) under mild growth conditions and smoothness assumptions on the coefficients. Jourdain [71] established the unique weak solution under the assumption that b is bounded and Lipschitz continuous in the law variable. The application of martingale problems in the study of McKean–Vlasov SDEs has also been discussed in other works, for example [26]. If we compare these works, it is easily seen that the martingale problem provides a powerful tool in studying weak solutions for both stochastic differential equations and McKean–Vlasov type equations. However, this approach relies heavily on the continuity of the coefficients.

We remark that in the study of Itô's stochastic differential equations, a typical approach for constructing strong solutions is via the Yamada–Watanabe principle (see [142]). Specifically, with the pathwise uniqueness, a weak solution is a strong solution. Bauer et al. [10] used this property to study the one-dimensional McKean–Vlasov equations under the condition that σ was a constant, and that

b was continuous in the third variable and of at most linear growth. The authors constructed the weak solutions by a fixed point argument in addition to the Girsanov's theorem. With an analogy as Itô's stochastic differential equations, the strong existence followed after showing pathwise uniqueness held for the weak solution. Later in 2019, Bauer and Meyer-Brandis [9] extended the existence and uniqueness result to the multi-dimensional case under additional continuity assumptions on b. Mishura and Veretennikov [100] also used this Yamada–Watanabe argument for the time-inhomogeneous problem, where the drift was required to satisfy a linear growth condition. Whereas the weak existence was established with an analogy as Krylov's [79, 80] for Itô's SDEs.

Some other approaches were also employed to weaken the assumptions and study the solvability of the McKean–Vlasov equation (1.4). One can refer to [89, 34, 61, 120, 131, 136] and literature therein for more recent progress and other approaches for studying McKean–Vlasov equations.

1.1.3 Main contribution

In the study of the McKean–Vlasov tpye SDE (1.3), the kernel function K is singular, for example, the Biot–Savart kernel $-\frac{1}{4\pi}\frac{x}{|x|^3}$ (where $d = 3$). In this case, the above approaches are not appropriate. In our study, we overcome these difficulties by devising a new and powerful approach which allows us to establish the existence and uniqueness of both strong and weak solutions of (1.3) under very weak conditions on the singular integral kernel K. In particular, our results apply to the the Biot–Savart kernel in two and three dimensions, and also apply to the Green kernels (such as $1/|x|^{d-2}$ for $d > 2$), the Riesz kernels $1/|x|^\gamma$ where $\gamma \in [0, d)$ on $\mathbb{R}^d$ and many other singular integral kernels.

Our novel approach is based on the following simple observation. If K is singular, then the mapping

$$(x, \mu) \mapsto \sum_{j=1}^{d} \int_{\mathbb{R}^d} \mathbb{E}\left[K^{ij}(x-\xi)\right] \omega_0^j(y) dy,$$

where ξ has a distribution μ, is unlikely to be Lipschitz continuous with respect to the variational or the Wasserstein metric on the space of distributions. However, we recognise that the distributions of possible solutions to (1.3), even if K is singular, have much higher regularity. In fact, if $\{X(x,t) : x \in \mathbb{R}^d, t \geq 0\}$ is a (weak) solution to (1.3) then

$$b^i(x,t) = \sum_{j=1}^{d} \int_{\mathbb{R}^d} \mathbb{E}\left[K^{ij}(x - X(y,t))\right] \omega_0^j(y) dy$$

defines a vector field (although the vector field $b(x,t)$ is defined via the solution of the SDE), and $X(x,t)$ must be a (weak) solution to the diffusion process defined by ordinary SDE

$$dX_t = b(X_t, t)dt + \sqrt{2\nu} dB_t.$$

Therefore, the distribution of $X(x,t)$ can be represented by the Cameron-Martin formula in terms of the Wiener measure and many results from diffusion processes thus can be brought in to the study of SDE (1.3). The major technical tool is the sharp heat kernel estimates obtained in [117, 115].

1.2 FBSDEs from the porous medium equation

1.2.1 Overview of the porous medium equations and FBSDEs

The porous medium equation is a quasi-linear parabolic equation

$$\frac{\partial u}{\partial t} = \Delta u^m, \quad x \in \mathbb{R}^d, \tag{1.5}$$

where $m > 1$ is a constant and $\Delta = \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2}$ is the Laplace operator with respect to space variables, with given initial data

$$u(x, 0) = u_0(x).$$

This models gas flows in a porous medium, where the non-negative solution $u(x, t)$ represents the density of the gas (see [7, 137] for example). The formulation (1.5) has also been used in describing the evolution of some biological population in [57], where u describes the population density. For more extended applications of the porous medium equation, one may refer to [137] and literature therein for a detailed account.

Though this non-linear parabolic equation has been studied extensively, for example [8, 13, 31], our goal is to find a probabilistic representation and study it using probability method. Historically, the celebrated Feynman–Kac formula gives a probability interpretation for linear second order PDEs of elliptic or parabolic types, and has been generalised to the case of semi-linear second order PDE, see [32, 109, 107], etc. In particular, a system of quasi-linear parabolic partial differential equations:

$$\begin{cases} \frac{\partial u^l}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{d} a^{ij}(t, x, u)\frac{\partial^2 u^l}{\partial x_i \partial x_j} + \sum_{i=1}^{d} f^i(t, x, u, \nabla u \sigma(t, x, u))\frac{\partial u^l}{\partial x_i} \\ \qquad +g^l(t, x, u, \nabla u \sigma(t, x, u) = 0, \quad l = 1, ..., k, t \in (0, T), x \in \mathbb{R}^d, \\ u^l(x, T) = h^l(x), \quad l = 1, ..., k, x \in \mathbb{R}^d, \end{cases} \tag{1.6}$$

with $a^{ij}(t, x, u) = (\sigma\sigma^*)^{ij}(t, x, u)$ for any $1 \leq i, j \leq d, 0 \leq t \leq T, x \in \mathbb{R}^d, y \in \mathbb{R}^k$, is closely related with the following system of forward–backward stochastic differential equations:

$$\begin{cases} X_t = x + \int_0^t f(s, X_s, Y_s, Z_s)ds + \int_0^t \sigma(s, X_s, Y_s, Z_s)dB_s, \\ Y_t = h(X_T) + \int_t^T g(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dB_s, \end{cases} \tag{1.7}$$

when the coefficients f, g, σ are deterministic functions and σ is independent of z. Several papers have studied this connection between forward–backward stochastic differential equations and quasi-linear PDEs, for example [93, 108]. So knowing properties of the system of PDEs (1.6), one can derive some results on the FBSDEs (1.7), and vise versa.

If either of the following cases happens

(i) f, σ do not depend on the processes Y, Z,

(ii) the terminal data of the process Y does not depend on X, i.e. $h(X_T) = \xi$ is a given process,

then we say the system (1.7) is decoupled. In these two cases, the problem can then be treated by solving one of the equations first and then plugging the solution process obtained into the other equation. However, the theory of decoupled problems is much more extensive than the theory of general coupled FBSDE, in which many questions remain unanswered.

Coupled FBSDEs like (1.7) were first considered by Antonelli [1]. In his work, the coefficients f, σ, g were assumed to be measurable and uniformly Lipschitz. In addition, these coefficients were independent of the variable z. Then the existence and uniqueness of the solution were established for $T > 0$ sufficiently small via the contraction mapping theorem. In fact, Antonelli also gave an counterexample in [1] where the adapted solution may not exist when $T > 0$ is large. Pardoux and Tang [108] used this idea and imposed some monotonicity conditions to extend T to a larger time scale, and continuous dependence of the solution on a parameter was obtained. Nevertheless, the main restriction of this *Contraction Method* is that only local solutions could be obtained.

There are two other popular approaches in studying (1.7), which extends T to be an arbitrary time scale under the Markovian setting (i.e. the coefficients are deterministic). One is the so-called *Four-Step-Scheme*, initiated by Ma et al. [93]. This was the first solution method to remove time constraints for Markovian FBSDEs, which combined the methods of partial differential equations and stochastic optimal control. However, the limitation was that the matrix-valued coefficient σ was required to be non-degenerate so that the decoupling quasi-linear PDE had a classical solution. To overcome this limitation, Delarue [35] applied the contraction method to construct a local solution, then extended this local solution to a global one with the help of the PDE method so that the regularity assumptions on the coefficients were relaxed.

Another approach is the *Method of Continuation*, which was introduced by Hu and Peng [63] and Peng and Wu [110], then developed by Yong [60, 143, 144]. This method could be used to treat non-Markovian FBSDEs with arbitrary time duration. However, the monotonicity assumptions on the coefficients were imposed, which made this method restrictive.

More recently, Ma et al. [94] combined these methodologies and established a unified scheme to address non-Markovian FBSDEs. The authors introduced the concept of *Decoupling Fields* to patch solutions on small time intervals into a solution on arbitrary time duration. The authors also pointed out that the solutions of the PDEs in [35, 93] both served as deterministic decoupling functions. Followed by Fromm and Imkeller [47], the authors extended this method to study multidimensional problems.

1.2.2 Main contribution

In our study, we write the porous medium equation in terms of a system of forward–backward SDE

$$\begin{cases} dX_t = \sqrt{2Y_t}dB_t, & X_0 = x \\ dY_t = -\frac{1}{2(m-1)}\frac{|Z_t|^2}{Y_t}dt + Z_t \cdot dB_t, & Y_T = \phi(X_T), \end{cases} \tag{1.8}$$

where $B = (B^1, ..., B^d)$ is a d-dimensional Brownian motion on some probability space, and ϕ is a given function corresponds to u_0 in the porous medium equation. With the assumption that ϕ is bounded and Lipschitz continuous, we show the solvability of (1.8). We are concerned with a probabilistic method for studying this FBSDE and avoid converting this system back to partial differential equations as exploited in [93, 35]. Our approach involves decoupling the system and then establishing the existence and uniqueness of the solution by using the fixed point theorem on the space of bounded random variables. Specifically, we consider the following non-linear mapping

$$\xi \mapsto \phi\left(x + \int_0^T \sqrt{2Y_s^\xi}dB_s\right),$$

where ξ is a real-valued bounded, measurable random variable such that $c_1 \leq \xi \leq c_2$ for some $c_1, c_2 > 0$, and

$$Y_s^\xi = \xi + \int_s^T \frac{1}{2(m-1)}\frac{|Z_r^\xi|^2}{Y_r^\xi}dr - \int_s^T Z_r dB_r.$$

In this way, the backward equation in (1.8) can be studied independently for any given terminal data ξ with $c_1 \leq \xi \leq c_2$. Then the fixed point of this mapping is the solution to this system of forward–backward stochastic differential equations. Moreover, this approach can be employed to study quadratic FBSDEs with more general growth conditions. However, the restriction is that the solution is only obtained for small time intervals.

The thesis is organized as follows: In Chapter 2, we present a brief discussion of the connection between parabolic equations and diffusion processes, and summarize some key results and techniques in the theory of backward stochastic differential equations, which will be needed for later discussion. Chapter 3 is a submitted paper, see [116], which focuses on the study of the McKean–Vlasov SDE (1.3) - we explain the random vortex method for the Navier–Stokes equation, and prove the existence and uniqueness of the solution. Then by devising a numerical scheme based on an existing method, we simulate the velocity field for the 3D problem. In Chapter 4, we show the existence and uniqueness of the system of forward–backward stochastic differential equations (1.8) in a small time interval using a fixed point argument. Following this approach, we extend the solvability to a class of general quadratic FBSDE.

2 PDEs, SDEs and BSDEs

This chapter is devoted to introducing all the background materials that will be used throughout the thesis. Basically, we will review some fundamental results from the theory of parabolic equations and diffusion processes, stochastic differential equations, as well as some classical techniques from backward stochastic differential equations.

In section 1, we investigate elliptic operators of second order. In particular, we are interested in how the fundamental solutions to parabolic equations are related to transition probability densities of diffusion processes. In section 2, we emphasize some existence and uniqueness results from the theory of stochastic differential equations. In section 3, we discuss some key techniques in studying backward stochastic differential equations, which will play an important role in Chapter 4.

For the rest of the discussion, we assume $(\Omega, \mathscr{F}, \mathbb{P})$ is a complete probability space and $B = (B_t)_{t\geq 0}$ is a d-dimensional Brownian motion on this space. Let $(\mathscr{F}_t)_{t\geq 0}$ be the natural augmented filtration generated by this Brownian motion. By this we mean, if $\mathscr{F}_t^B = \sigma(B_s : s \leq t)$ and $\mathscr{F}_\infty^B = \sigma(B_s : s \geq 0)$ is the natural filtration generated by this Brownian motion B, and

$$\mathcal{N} = \{N \subseteq \Omega : \exists G \in \mathscr{F}_\infty^B \text{ with } N \subseteq G, \text{ and } \mathbb{P}(G) = 0\}$$

is the collection of null sets, then

$$\mathscr{F}_t = \sigma\left(\mathscr{F}_t^B \cup \mathcal{N}\right), \quad \mathscr{F}_\infty = \sigma\left(\cup_{t\geq 0}\mathscr{F}_t\right).$$

2.1 Elliptic operator of second order

Throughout this section, let L be an elliptic operator of second-order:

$$L = \sum_{i,j=1}^{d} a^{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b^i(x,t)\frac{\partial}{\partial x_i} + c(x,t)$$

where the coefficients a, b, c are defined on $\bar{D} \times [T_0, T_1]$, with $\bar{D}$ being the closure of a domain $D \subseteq \mathbb{R}^d$. We always take $a = (a^{ij})$ to be a symmetric matrix, $b = (b^i)$ is a time-dependent vector field, and c is a scalar function. Additionally, we make the following assumptions:

(A_1) L is uniformly elliptic, i.e. there exist positive constants $\lambda_0, \lambda_1 > 0$ such that for any real vector $\xi \in \mathbb{R}^d$,

$$\lambda_0 |\xi|^2 \leq \sum_{i,j=1}^{d} a^{ij}(x,t)\xi_i \xi_j \leq \lambda_1 |\xi|^2$$

for all $(x,t) \in \bar{D} \times [T_0, T_1]$.

(A_2) The coefficients of L are continuous for all $(x,t) \in \bar{D} \times [T_0, T_1]$.

The following convention will apply unless otherwise specified: when L applies to a function $u(x,t)$, then

$$Lu(x,t) = \sum_{i,j=1}^{d} a^{ij}(x,t)\frac{\partial^2 u}{\partial x_i \partial x_j}(x,t) + \sum_{i=1}^{d} b^i(x,t)\frac{\partial u}{\partial x_i}(x,t) + c(x,t)u(x,t)$$

for any $(x,t) \in \bar{D} \times [0, T]$.

2.1.1 Fundamental solutions for forward and backward problems

Here we follow the terminologies established in Friedman [46], but we do not stick with the notations established therein.

Let L be an elliptic operator defined as above. The formal adjoint $L^\star$ of L is again an elliptic operator of second order such that

$$L^\star = \sum_{i,j=1}^{d} a^{ij}(x,t)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d}(b^\star)^i(x,t)\frac{\partial}{\partial x_i} + c^\star(x,t),$$

where

$$(b^\star)^i = -b^i + 2\sum_{j=1}^{d}\frac{\partial a^{ij}}{\partial x_j}, \quad i = 1, ..., d,$$
$$c^\star = c - \sum_{i=1}^{d}\frac{\partial b^i}{\partial x_i} + \sum_{i,j=1}^{d}\frac{\partial^2 a^{ij}}{\partial x_i \partial x_j}.$$

When we speak of $L^\star$, we may always assume that $\frac{\partial a^{ij}}{\partial x_h}, \frac{\partial^2 a^{ij}}{\partial x_h \partial x_k}, \frac{\partial b^i}{\partial x_h}$ exist and are continuous functions.

Suppose ϕ, ψ are of class $\mathcal{C}^{2,1}(\bar{D} \times [T_0, T_1])$. Then the two elliptic operators L and $L^\star$ are related via the following identity

$$\varphi L\psi - \psi L^\star \varphi = \sum_{i=1}^{d}\frac{\partial}{\partial x_i}\left[\sum_{j=1}^{d}\left(\varphi a^{ij}\frac{\partial \psi}{\partial x_j} - \psi a^{ij}\frac{\partial \varphi}{\partial x_j} - \varphi\psi\frac{\partial a^{ij}}{\partial x_j}\right) + b^i\varphi\psi\right] - \frac{\partial}{\partial t}(\phi\psi),$$

which is obtained by using formally integration by parts, so that the zero order terms c and $c^\star$ do not appear in this equality. Moreover, here we emphasize the fact that L and $L^\star$ play an equal part and $(L^\star)^\star = L$ (see [46, Theorem 13, page 27]).

Given an elliptic operator L, there are two types of fundamental solutions associated with L (allowed having zero order term), namely fundamental solution to the forward parabolic equations, denoted by $\Gamma_{F,L}$, and fundamental solution to the backward parabolic equations, often denoted by $\Gamma_{B,L}$.

Definition 2.1. Let $D \subseteq \mathbb{R}^d$, $\bar{D}$ be the closure of D in $\mathbb{R}^d$. For any $T_0 \leq \tau < t \leq T_1$ and $x, \xi \in \bar{D}$, a function $\Gamma_{F,L}(x,t;\xi,\tau)$ is called a fundamental solution to the forward parabolic equation

$$\left(\frac{\partial}{\partial_t} - L\right) u(x,t) = 0, \quad (x,t) \in \bar{D} \times [T_0, T_1] \tag{2.1}$$

if the following conditions are satisfied:

(i) for every $(\xi, \tau) \in \bar{D} \times [T_0, T_1]$ fixed, $\mathit{\Gamma}_{F,L}(x,t;\xi,\tau)$ which is seen as a function of (x,t), say $u(x,t) = \mathit{\Gamma}_{F,L}(x,t;\xi,\tau)$, satisfies the equation (2.1) for $\tau < t \leq T_1$ and $x \in \bar{D}$;

(ii) for every continuous function f in $\bar{D}$ such that

$$|f(x)| \leq C \exp\left(h|x|^2\right) \tag{2.2}$$

for some constant $C, h > 0$, we have

$$\lim_{t \downarrow \tau} \int_D f(\xi) \mathit{\Gamma}_{F,L}(x,t;\xi,\tau) d\xi = f(x)$$

for all $x \in \bar{D}$.

Similarly, for $T_0 \leq t < \tau \leq T_1, x, \xi \in \mathbb{R}^d$, $\mathit{\Gamma}_{B,L}(\xi,\tau;x,t)$ is called a fundamental solution to the backward parabolic equation

$$\left(\frac{\partial}{\partial_t} + L\right) u(x,t) = 0, \quad (x,t) \in \bar{D} \times [T_0, T_1] \tag{2.3}$$

if it meets the following conditions:

(i) for every $(\xi, \tau) \in \bar{D} \times [T_0, T_1]$ fixed, $\mathit{\Gamma}_{B,L}(\xi,\tau;x,t)$ which is seen as a function of (x,t), say $u(x,t) = \mathit{\Gamma}_{F,L}(x,t;\xi,\tau)$, satisfies the equation (2.3) for $T_0 \leq t < \tau$ and $x \in \bar{D}$;

(ii) for every continuous function f satisfying (2.2) in $\bar{D}$, it holds that

$$\lim_{t\uparrow\tau}\int_D f(\xi)\Gamma_{B,L}(\xi,\tau;x,t)d\xi = f(x),$$

for all $x\in\bar{D}$

Similarly, we may define fundamental solutions for forward and backward parabolic equations generated by the dual operator $L^\star$ in the same manner.

Fundamental solutions for second-order parabolic equations were first constructed by Gevrey [51] for $a^{ij}=\delta^{ij}$, and Rothe [122] for time independent coefficients by using parametrix method introduced by Levi [88]. Since then, a large amount of work has been done to relax the assumptions on the coefficients when studying the existence of the fundamental solutions, for example [39, 40] for sufficiently smooth coefficients, and Pogorzelski [112] and Aronson [3] for Hölder continuous coefficients. One may refer to [4, 66] for a detailed description for this theory. Here we quote without proof from [46, Theorem 10, Page 23] for the existence theorem of fundamental solutions.

Proposition 2.2. *Let D be any domain in $\mathbb{R}^d$. In addition to $(A_1),(A_2)$, the coefficients satisfy the following assumption:*

(A_3) the coefficients are bounded functions and Hölder continuous such that for all $x_1,x_2\in\bar{D},t_1,t_2\in[T_0,T_1]$,

$$\begin{aligned}\left|a^{ij}(x_1,t_1)-a^{ij}(x_2,t_2)\right| &\le A(|x_1-x_2|^\alpha+|t_1-t_2|^{\alpha/2}),\\ \left|b^{i}(x_1,t_1)-b^{i}(x_2,t_2)\right| &\le A|x_1-x_2|^\alpha,\\ |c(x_1,t_1)-c(x_2,t_2)| &\le A|x_1-x_2|^\alpha,\end{aligned}$$

for any $1\le i,j\le d$, where $A>0$ is a constant and $0<\alpha\le 1$.

Then there exists a fundamental solution $\Gamma_{F,L}(x,t;\xi,\tau)$ of $(\frac{\partial}{\partial t}-L)u=0$.

For the uniqueness of the fundamental solution, we note that

(a) If the boundary ∂D is non-empty, then fundamental solutions are not necessarily unique, so certain boundary conditions to the forward/backward parabolic equations (2.1) and (2.3) must be assigned to achieve a unique fundamental solution.

(b) If $D = \mathbb{R}^d$, and the coefficients satisfy the additional assumption that

(A_4) the functions

$$a^{ij}, \quad b^i, \quad c, \quad \frac{\partial a^{ij}}{\partial x_h}, \quad \frac{\partial^2 a^{ij}}{\partial x_h \partial x_k}, \quad \frac{\partial b^i}{\partial x_h}$$

are bounded and continuous in $\mathbb{R}^d \times [T_0, T_1]$ for any $1 \le i, j, h, k \le d$.

Then the fundamental solutions for both the forward and the backward parabolic equations exist uniquely. The proof of this uniqueness result can be found in [46, Theorem 16, page 29]. In this case,

$$\Gamma_{F,L}(x, t; \xi, \tau) = \Gamma_{B,L^\star}(\xi, \tau; x, t)$$

for any $t > \tau, x \in \mathbb{R}^d$. Since $(L^\star)^\star = L$, we also have

$$\Gamma_{F,L^\star}(x, t; \xi, \tau) = \Gamma_{B,L}(\xi, \tau; x, t)$$

for $t > \tau, x \in \mathbb{R}^d$.

If L is an elliptic operator, the importance of a fundamental solution lies in the resolution of the Cauchy initial value problem

$$\begin{aligned} \left(L - \frac{\partial}{\partial t}\right) u(x, t) &= f(x, t), \quad (x, t) \in \mathbb{R}^d \times (0, T], \\ u(x, 0) &= \phi(x), \qquad x \in \mathbb{R}^d. \end{aligned} \tag{2.4}$$

Proposition 2.3. *Suppose L is an elliptic operator in $\mathbb{R}^d \times [0, T]$ such that it satisfies the assumptions (A_1) – (A_3). Let $f(x, t), \phi(x)$ in (2.4) be bounded continuous functions in $\mathbb{R}^d \times [0, T], \mathbb{R}^d$, respectively. Assume also that $f(x, t)$ is locally Hölder continuous in $x \in \mathbb{R}^d$, uniformly with respect to t. Then the function*

$$u(x,t) = \int_{\mathbb{R}^d} \mathit{\Gamma}_{F,L}(x,t;\xi,0)\phi(\xi)d\xi - \int_0^t \int_{\mathbb{R}^d} \mathit{\Gamma}_{F,L}(x,t;\xi,\tau)f(x,\tau)d\xi d\tau$$

is a solution of the Cauchy problem (2.4), and u is also bounded. In addition, if (A_4) is satisfied, i.e. the functions

$$a^{ij}, \quad b^i, \quad c, \quad \frac{\partial a^{ij}}{\partial x_h}, \quad \frac{\partial^2 a^{ij}}{\partial x_h \partial x_k}, \quad \frac{\partial b^i}{\partial x_h}$$

are continuous and bounded for any $1 \leq i, j, h, k \leq d$, then such solution is unique and satisfies the boundedness condition

$$\int_0^T \int_{\mathbb{R}^d} |u(x,t)| \exp\left(-k|x|\right) dxdt < \infty$$

for some positive number $k > 0$.

Here we state this representation formula without proof from [46, Theorem 12 and Theorem 16, Chapter 1]. In fact, the proof relies on the parametrix method, and we are not interested in this method in this thesis. But a good derivation can be found in the reference cited.

2.1.2 Diffusion processes generated by differential operators

This subsection focuses on the most interesting example for the study in this work, where L is a diffusion operator, i.e. L is a uniformly elliptic operator of second order without zero order term:

$$L=\sum_{i,j=1}^{d}a^{ij}(x,t)\frac{\partial^2}{\partial x_i\partial x_j}+\sum_{i=1}^{d}b^i(x,t)\frac{\partial}{\partial x_i},$$

where (a^{ij}) is a symmetric matrix and (b^i) is a vector field in $\mathbb{R}^d$. Here we assume the assumptions $(A_1),(A_2)$ are satisfied. We may then define diffusion processes via this operator.

Formally, by a diffusion with infinitesimal generator L (or L-diffusion), we mean a family of strong Markovian probability measures $(\mathbb{P}_{\tau,\xi})_{\tau\geq 0,\xi\in\mathbb{R}^d}$ on the path space $\mathcal{C}(\mathbb{R}_+,\mathbb{R}^d)$ of all continuous functions $\omega:\mathbb{R}_+\to\mathbb{R}^d$, such that

(i) for every $\tau\geq 0,\xi\in\mathbb{R}^d$

$$\mathbb{P}_{\tau,\xi}\left(\omega\in\mathcal{C}(\mathbb{R}_+,\mathbb{R}^d):\omega(t)=\xi,\forall 0\leq t\leq\tau\right)=1,$$

(ii) for every $f\in\mathcal{C}^{2,1}(\mathbb{R}^d\times\mathbb{R}_+)$, let

$$M_t^{[f]}=f(\omega(t),t)-f(\omega(\tau),\tau)-\int_\tau^t(L+\partial_s)f(\omega(s),s)ds$$

for any $t\geq\tau\geq 0$ and $M_t^{[f]}=0$ for any $0\leq t\leq\tau$. Then $M^{[f]}$ is a local martingale under the probability measure $\mathbb{P}_{\tau,\xi}$.

Such an L-diffusion is unique if it further satisfies the uniqueness condition:

(iii) if $(\mathbb{P}'_{\tau,\xi})_{\tau\geq 0,\xi\in\mathbb{R}^d}$ is another system of probability measures on the path space $\mathcal{C}(\mathbb{R}_+,\mathbb{R}^d)$ satisfying the above conditions (i) and (ii), then

$$\mathbb{P}_{\tau,\xi}=\mathbb{P}'_{\tau,\xi}$$

for every $\xi\in\mathbb{R}^d$.

In this sense, we say the probability measure $\mathbb{P}_{\tau,\xi}$ solves the L-martingale problem (see [127]).

In the case where (a^{ij}) has a good square root, that is,

$$a^{ij}(x,t)=\sum_{k=1}^{r}\sigma^{ik}(x,t)\sigma^{kj}(x,t),\quad i,j=1,...,d,$$

where $\sigma=(\sigma^{ik})$ is a continuous matrix in $\mathbb{R}^{d\times r}$, we may consider the following stochastic differential equations

$$dX_t^i=\sqrt{2}\sum_{k=1}^{r}\sigma^{ik}(X_t,t)dB_t^k+b^i(X_t,t)dt,\quad X_\tau=\xi, \tag{2.5}$$

where $B=(B^1,...,B^r)$ is a Brownian motion on $(\Omega,\mathscr{F},\mathbb{P})$, and $i=1,...,d$. For example, if σ and b are global Lipschitz continuous, then this SDE has a unique strong solution; if σ is Lipschitz and uniformly elliptic, and b is bounded Borel measurable, then (2.5) has a unique weak solution. Nevertheless, the distribution of X, which solves the SDE (2.5), gives us the probability measure $\mathbb{P}_{\tau,\xi}$. The standard references are Stroock and Varadhan [129] for the martingale approach, and Ikeda and Watanabe [65, Chapter IV] for an excellent exposition of diffusion processes and stochastic differential equations.

Assume $a(x,t)$ is continuous, bounded and uniformly elliptic for any $t\geq 0, x\in\mathbb{R}^d$, and $b(x,t)$ is bounded and measurable. Then [127, 128] indicates that for any $\xi\in\mathbb{R}^d,\tau\geq 0$, the solution $\mathbb{P}_{\tau,\xi}$ to the martingale problem, starting from the point ξ at time $\tau\geq 0$ exists, and is unique. Define

$$P(\tau,\xi,t,\Gamma):=\mathbb{P}_{\tau,\xi}\big(\omega(t)\in\Gamma\big),$$

for any $t\geq\tau,\Gamma\in\mathscr{B}(\mathbb{R}^d)$, then $P(\tau,\xi,t,\Gamma)$, seeing as a function of (τ,ξ), is measurable on $[0,t)\times\mathbb{R}^d$. In particular, $P(\tau,\xi,t,\cdot)$ is the transition probability function of $(\mathbb{P}_{\tau,\xi})_{\tau\geq 0,\xi\in\mathbb{R}^d}$ (see [129, Theorem 5.1.5, page 128]). Under the above assumptions on the coefficients $a(x,t)$ and $b(x,t)$, $P(\tau,\xi,t,x)$ has a density with respect to the Lebesgue measure, that is,

$$P(\tau,\xi,t,dx) = p_L(\tau,\xi,t,x)dx,$$

which defines the transition density function $p_L(\tau,\xi,t,x)$ (for $t > \tau \geq 0$) associated with the L-diffusion (see [129, Theorem 9.1.9, page 220]). Also, $p(\tau,\xi,t,x)$ is continuous for $t > \tau \geq 0$ and $\xi, x \in \mathbb{R}^d$. Moreover, it is assumed that

$$P(\tau,\xi,t,dx) \to \delta_\xi(dx) \text{ as } t \downarrow \tau$$

weakly.

Fix $t \geq 0$, let f be any bounded continuous function on $\mathbb{R}^d$. Define the function

$$v(y,s) := \int_{\mathbb{R}^d} P(s,y,t,dx)f(x)$$

for $y \in \mathbb{R}^d, t > s \geq 0$. Then $v(y,s)$ satisfies the backward parabolic equation

$$(\partial_s + L)v(y,s) = 0 \tag{2.6}$$

for any $(y,s) \in \mathbb{R}^d \times (0,t)$, subject to the terminal condition

$$\lim_{s\uparrow t} v(y,s) = f(y).$$

In particular, for $t > s \geq 0$ and $x, y \in \mathbb{R}^d$, (2.6) implies that the function $u(y,s) := p(s,y,t,x)$ is a solution to

$$(\partial_s + L)u(y,s) = 0$$

for any $(y,s) \in \mathbb{R}^d \times (0,t)$, and

$$\lim_{s \uparrow t} \int_{\mathbb{R}^d} p(s,y,t,x) f(x) dx = f(y)$$

for any bounded continuous function f. Therefore, the transition probability density function for the L-diffusion coincides with the fundamental solution to the backward parabolic equation. We may summarize this, together with the fact from earlier discussion as the following lemma.

Lemma 2.4. *If L is a diffusion operator given as (2.1.2), and satisfies the assumptions (A_1) and (A_2), and also that $a(x,t)$ is continuous and bounded, $b(x,t)$ is bounded and measurable then*

$$\Gamma_{F,L^\star}(x,t;\xi,\tau) = \Gamma_{B,L}(\xi,\tau;x,t) = p_L(\tau,\xi;t,x)$$

for all $0 \le \tau < t$, and $x, \xi \in \mathbb{R}^d$.

2.1.3 Heat kernel estimates

This subsection is devoted to estimating the fundamental solution to the Cauchy problem

$$\left(\frac{\partial}{\partial t} - L \right) u(x,t) = 0, \tag{2.7}$$

where L is an elliptic operator defined as

$$L = \sum_{i,j=1}^{d} a^{ij}(x,t) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b^i(x,t) \frac{\partial}{\partial x_i}. \tag{2.8}$$

To simplify the notation, we write $a = (a^{ij})_{1 \le i,j \le d}$ and $b = (b^i)_{1 \le i \le d}$. Here we assume a is symmetric, and the coefficients a, b satisfy the assumptions (A_1)-(A_4). Then the fundamental solution to (2.7) exists uniquely, and is denoted by

$\Gamma(\tau, \xi; t, x)$. We may now present our first estimate for the fundamental solutions, which is the so-called Aronson estimate. Here we quote from [5, Theorem 1]:

Lemma 2.5 (Aronson estimates)**.** *There exist positive constants* α_1, α_2 *and* $M > 0$ *such that*

$$M^{-1}\gamma_1(x - \xi, t - \tau) \le \Gamma(\tau, \xi; t, x) \le M\gamma_2(x - \xi, t - \tau) \tag{2.9}$$

for all $(x, t), (\xi, \tau) \in \mathbb{R}^d \times (0, T)$ *with* $t > \tau$*, where* $\gamma_i(x, t)$ *is the fundamental solution to the equation*

$$\frac{\partial}{\partial t}u - \alpha_i \Delta u = 0, \quad i = 1, 2.$$

The constants depend only on ν, d, T *and the bounds for the coefficient* a, b*.*

This heat kernel estimate (2.9) for parabolic equations constitutes a summary of the results from De Giorgi [33] and Nash [102], culminated by Moser [101]. In fact, Aronson [4, 5] was the first to to prove the two-sided Gaussian estimates for the fundamental solutions. [5] presented a detailed proof of this inequality (2.9) for the special case where $b \equiv 0$. Then [5, 6] indicated that this heat kernel estimate can be extended to a general parabolic equation

$$\frac{\partial u}{\partial t} - \sum_{i,j=1}^{d} \frac{\partial}{\partial x_j}\left(a^{ij}(x,t)\frac{\partial u}{\partial x_i}\right) - \sum_{j=1}^{d} b^j(x,t)\frac{\partial u}{\partial x_j} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j}\left(\hat{b}^j(x,t)u\right) + c(x,t)u = 0,$$

where a is symmetric and uniformly elliptic, and $b, \hat{b}$ are bounded measurable functions. One may refer to [43, 103, 125] for a proof of the estimates.

In 2003, Qian et al. [115] established a sharp bound estimate for the fundamental solutions of a class of parabolic equations using probabilistic method. Let

L_b be the elliptic operator defined by

$$L_b = \frac{1}{2}\sum_{i,j=1}^{d}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b^i(x,t)\frac{\partial}{\partial x_i},$$

where b is a bounded and measurable vector field on $\mathbb{R}^d$. Then the L_b-diffusion process associated with this infinitesimal generator solves the following stochastic differential equation

$$dX_t = b(X_t, t)dt + dB_t, \quad X_0 = x,$$

where $(B_t)_{t\geq 0}$ is a standard d-dimensional Brownian motion on some probability space. If $p_b(0,x,t,y)$ is the transition probability density function (with respect to the Lebesgue measure), then it coincides with the fundamental solution to the parabolic equation

$$\left(\frac{\partial}{\partial t} + L_b\right) u(x,t) = 0.$$

In the case of bounded drift, that is, $\|b\|_\infty \leq C_0$ for some $C_0 > 0$, Qian et al. [115] presented a sharp bound estimate for $p_b(0,x,t,y)$. Qian and Zheng [118] further extended the two-side estimates to the case where b grows not faster than linear growth. But here we only focus on the case of bounded drift.

Proposition 2.6 (Qian et al.). *For $d \geq 2$, if $\|b\|_\infty \leq C_0$, and $p_b(0,x,t,y)$ is the the transition probability density of the Brownian motion with bounded drift b, that is, the diffusion process*

$$dX_t = b(X_t)dt + dB_t, \quad X_0 = x.$$

Then we have the following lower bound

$$p_b(0,x,t,y) \geq \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{(|x-y|+C_0 t)^2}{2t} - C_0 \left(\zeta_d \sqrt{t} + |x-y| \right) \right),$$

where

$$\zeta_d = 2(a_{d-1,1} + a_{1,1}),$$

and

$$a_{d,p} = \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} |x|^p e^{-\frac{|x|^2}{2}} dx.$$

For the upper bound, for any $1 < q < \frac{d}{d-1}$,

$$\begin{aligned} p_b(0,x,t,y) \leq\ & \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|x-y|^2}{2t}} \\ & \times \left(1 + c_{d,q} C_0 \left(\kappa_q \sqrt{t} + 2|x-y| \right) e^{-\frac{q-1}{2qt}|x-y|^2 + \frac{C_0^2}{2(q-1)} t} \right) \end{aligned}$$

for every $x, y \in \mathbb{R}^d$ *and* $t > 0$, *where*

$$c_{d,q} = \frac{2\sqrt{2}q}{d - q(d-1)}, \quad \kappa_q = \sqrt[2q]{a_{d-1,q} + 2^{2q-1} a_{1,q}}.$$

This sharp bound estimate will play a crucial role in our study for the McKean–Vlasov type SDE in Chapter 3. Without explicit calculation, we give a sketch of the proof followed by [115, 117].

Sketch of proof. The idea is to utilise the comparison theorem for transition probability density functions, which is established in [115, Theorem 4]. Then the estimates for $p_b(0,x,t,y)$ may be passed to the estimates for the transition probability density of Bessel processes.

Specifically, given any $\beta \in \mathbb{R}$ and $x, y \in \mathbb{R}^d$, let $((X_t)_{t\geq 0}, \mathbb{P}^x)$ be the weak solution to the stochastic differential equation

$$dX_t = \beta\left(\nabla|X_t - y|\right)dt + dB_t, \quad X_0 = x.$$

If $p_{\beta,y}(x,t,z)$ is the transition probability density function with respect to the Lebesgue measure, then the comparison theorem established in [115] indicates that

$$p_{\|b\|,y}(x,t,y) \le p_b(x,t,y) \le p_{-\|b\|,y}(x,t,y)$$

for all $x, y \in \mathbb{R}^d$ and $t > 0$, where $\|b\| = \sup_{x\in\mathbb{R}^d}\sqrt{\sum_{i=1}^d |b^i(x)|^2}$ is the supremum norm of Euclidean length of the vector field b.

On the other hand, let $((\gamma_t)_{0\le t\le T}, \mathbb{P}^r)$ be a one-dimensional diffusion process such that

$$d\gamma_t = \frac{d-1}{2\gamma_t}dt + \beta t + dW_t, \quad \gamma_0 = r \ge 0,$$

where $(W_t)_{0\le t\le T}$ is a standard Brownian motion. Then $(\gamma_t)_{0\le t\le T}$ is a Bessel process with parameter d and constant drift β. This process was first introduced by Kendall in [77], and is called a Bessel process with a naive constant drift in [111]. On an application of the Itô's formula to the process $|X - y|$, it can be easily seen that the process $((|X_t - y|)_{t\ge 0}, \mathbb{P}^x)$ has the same distribution as $\left((\gamma_t)_{t\ge 0}, \mathbb{P}^{|x-y|}\right)$ for every $x \in \mathbb{R}^d$. Hence, in order to find the upper and the lower bounds for p_b, it is sufficient to estimate the transition probability density for the Bessel process, which is straightforward to compute. A detailed computation can be found in [115]. □

Another topic related to the Aronson estimates is the regularity of solutions to the parabolic equation $(\frac{\partial}{\partial t} - L_b)\omega = 0$ (see [82] for a complete survey of classical results).

Lemma 2.7 (J. Nash)**.** *Let L be an elliptic operator defined in* (2.8) *such that the*

$a(x,t)$ is a measurable, symmetric matrix elliptic, and the coefficients satisfy the assumptions $A(1)$ and $A(4)$. If λ is the elliptic constant, then for each $\delta \in (0,1)$, there exist $\alpha \in (0,1], C \in (0,\infty)$ depending on $\|b\|_\infty, d, \lambda$ and δ such that

$$\left|p_b(s,y,t,x) - p_b(s,y,\tilde{t},\tilde{x})\right| \leq \frac{C}{\delta^d}\left(\frac{|t-\tilde{t}|^{1/2} \vee |x-\tilde{x}|}{\delta}\right)^{\alpha}$$

for all $s \geq 0, \delta^2 \leq t, \tilde{t} \leq 1/\delta^2, |x-\tilde{x}| \leq \delta$.

We take this from [125, Theorem II.2.12, page 340]. This is the so-called Nash's theorem on continuity. It demonstrates the Hölder continuity of the fundamental solutions, and proves that the Hölder exponent depends only on the dimension, the bound of the coefficient b and the elliptic constant. Within the same setting as in Nash [102], Moser [101] established the Harnack inequality for positive solutions of the parabolic equation $(\frac{\partial}{\partial t} - L)u = 0$. Fabes and Stroock [43] showed that Moser's Harnack inequality could be derived via Nash's idea by using the two-sided estimate (2.9). Stroock [125] further demonstrated that both Nash's theorem and Harnack inequality could be derived via the Aronson estimates on the fundamental solutions.

2.2 Stochastic differential equations

Given a diffusion operator L with coefficients $(a^{ij}(x,t))_{1\leq i,j\leq d}$ and $(b^i(x,t))_{1\leq i\leq d}$, we have discussed how to generate a diffusion process from L. In this section, we study the diffusion processes via Itô's theory of SDEs.

The term *stochastic differential equation* was introduced by Bernštein [14] in the limiting study of a sequence of Markov chains arising in a stochastic difference scheme. However, it was debatable whether to regard Bernštein as the founder of this theory (see [54]). Nevertheless, the study of stochastic differential equations was carried out in a masterly way by Itô [67, 68, 69]. Independently, Gihman

[52, 53] developed a theory of stochastic differential equations, complete with results on the existence, uniqueness, smooth dependence on the initial conditions, and Kolmogorov's equations for the transition density. Since then, the study of stochastic differential equation has enjoyed remarkable success, and many works have been carried out to seek strong solutions under weaker assumptions on b, σ (see [75] for a detailed account).

This section is devoted to reviewing theories of stochastic differential equations (SDEs) driven by Brownian motions. For the study of SDEs driven by general semimartingales, one may refer to [38, 114] for example. Mainly, in this section we will revisit some existence and uniqueness results of stochastic differential equations of the form

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dB_t, \quad X_0 = \xi, \tag{2.10}$$

or written in componentwise as

$$dX_t^i = b^i(X_t, t)dt + \sum_{j=1}^{r} \sigma^{ij}(X_t, t)dB_t^j, \quad X_0^i = \xi^i,$$

for each $1 \leq i \leq d$, where $b : \mathbb{R}^d \times [0, \infty) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times [0, \infty) \to \mathbb{R}^{d \times r}$ are Borel measurable functions, and $B = (B_t)_{t \geq 0}$ is a standard r-dimensional Brownian motion on some probability space. Here we call the d-dimensional vector $b = (b^i)_{1 \leq i \leq d}$ a drift vector, and the $d \times r$ matrix $\sigma = (\sigma^{ij})_{1 \leq i \leq d, 1 \leq j \leq r}$ a dispersion matrix.

2.2.1 Strong solutions

In order to develop the concept of strong solution, we choose a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and $B = (B_t)_{t \geq 0}$ an r-dimensional Brownian motion on this space. Let $(\mathscr{F}_t)_{t \geq 0}$ be the natural augmented filtration generated by this

Brownian motion.

Definition 2.8. A strong solution to the stochastic differential equation (2.10) with respect to the fixed Brownian motion $(B_t)_{t\geq 0}$ and the initial condition ξ is a d-dimensional process $(X_t)_{t\geq 0}$ with continuous sample paths such that it satisfies

(i) $(X)_{t\geq 0}$ is adapted to the filtration $(\mathscr{F}_t)_{t\geq 0}$,

(ii) $\mathbb{P}(X_0 = \xi) = 1$,

(iii) $\int_0^t (|b^i(X_s, s)| + |\sigma^{ij}(X_s, s)|^2)\, ds < \infty$, $\mathbb{P}$-a.s. for every $1 \leq i \leq d, 1 \leq j \leq r$ and $t \geq 0$,

(iv) the integral version of (2.10)

$$X_t = X_0 + \int_0^t b(X_s, s)ds + \int_0^t \sigma(X_s, s)dB_s,$$

holds almost surely for any $t \geq 0$.

Definition 2.9. Given the drift vector b and dispersion matrix σ. Suppose $B = (B_t)_{t\geq 0}$ is an r-dimensional Brownian motion on $(\Omega, \mathscr{F}, \mathbb{P})$, ξ is an independent, d-dimensional random vector, and $(\mathscr{F}_t)_{0\leq t\leq T}$ is an augmented filtration generated by B. Let $(X)_{t\geq 0}, (\tilde{X})_{t\geq 0}$ be two adapted strong solutions to (2.10) relative to B with the initial condition ξ. If $\mathbb{P}(X_t = \tilde{X}_t, t \geq 0) = 1$, then we say pathwise uniqueness (or strong uniqueness) holds for the SDE (2.10).

With a proper definition of strong solutions of SDEs, it is natural to ask when the stochastic differential equation (2.10) has a solution, and whether such a solution is unique. These studies were first carried out by Itô [67, 68] for uniformly continuous coefficients b and σ.

Proposition 2.10. *Suppose that the coefficients $b(t, \cdot), \sigma(t, \cdot)$ are both uniformly Lipschitz continuous and satisfy the linear growth conditions, i.e. there exist pos-*

itive constants $K_1, K_2 > 0$ such that

$$\|b(t,x) - b(t,y)\| + \|\sigma(t,x) - \sigma(t,y)\| \leq K_1\|x - y\|,$$
$$\|b(t,x)\|^2 + \|\sigma(t,x)\|^2 \leq K_2\left(1 + \|x\|^2\right)$$

for every $t \geq 0$ and $x, y \in \mathbb{R}^d$. Let ξ be an $\mathbb{R}^d$-valued square-integrable random vector on $(\Omega, \mathscr{F}, \mathbb{P})$ and independent of the Brownian motion $B = (B_t)_{t\geq 0}$. If $(\mathscr{F}_t)_{t\geq 0}$ is an augmented filtration generated by B, then there exists a continuous, adapted process $(X_t)_{t\geq 0}$, which is a strong solution of (2.10) *relative to B with the initial condition ξ. Moreover, this process is also square-integrable, and is pathwise unique.*

This proposition establishes the existence and uniqueness of the strong solution of an SDE (2.10) under the condition that the drift and dispersion coefficients satisfy the Lipschitz and linear growth conditions. One may refer to [75, Chapter 5.2, page 288] for a detailed discussion of this proposition and related results.

In particular, we are interested in ordinary stochastic differential equations of the form

$$dX_t = b(X_t, t)dt + dB_t, \quad X_0 = x, \tag{2.11}$$

where b is bounded and Borel measurable. This type of SDEs will play an important role in later discussion in Chapter 3.

Lemma 2.11. *Assume $b(x,t)$ is bounded and Borel measurable, then* (2.11) *has a strong solution, and the solution is pathwise unique.*

This lemma is a special case of [138]. In fact, the existence and uniqueness result of the strong solution holds under more general conditions, which can be found in [138, Theorem 1, page 388] or in Zvonkin-Krylov [147] for one-dimensional problem.

2.2.2 Weak solutions

Under certain regularity conditions on the drift and dispersion coefficients, we have seen some results on the existence and uniqueness of the strong solutions. However, for coefficients under more general settings, a strong solution to the stochastic differential equation may not exist. Thus, the concept of weak solutions is introduced, which aims to find a process with the right law. The principal method for constructing weak solutions to stochastic differential equations is by utilising the Girsanov theorem, where the drift term will be transformed. Stroock and Varadhan [127, 128] developed another approach, which formulated the law of a diffusion process in terms of a martingale problem. One may refer to [129] for a detailed discussion.

Definition 2.12. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space. Then $(X_t, B_t)_{t\geq 0}$ is called a weak solution to the stochastic differential equation (2.10) with initial distribution μ if:

(i) $(\mathscr{F}_t)_{t\geq 0}$ is a filtration of sub-σ-fields of $\mathscr{F}$ satisfying the usual conditions such that $(B_t)_{t\geq 0}$ is adapted to $(\mathscr{F}_t)_{t\geq 0}$,

(ii) $(X_t)_{t\geq 0}$ is a continuous, adapted $\mathbb{R}^d$-valued process,

(iii) for any $\Gamma \in \mathscr{B}(\mathbb{R}^d)$,

$$\mu(\Gamma) = \mathbb{P}(X_0 \in \Gamma),$$

and (iii), (iv) of Definition 2.8 are satisfied.

It is clear that a strong solution of the stochastic differential equation (2.10) is constructed on a given probability space with respect to a given filtration and a given Brownian motion, whereas in the notion of a weak solution, the probability space, the filtration, and the driving Brownian motion are part of the solution rather than the statement of the problem.

Here we make a remark: it can be easily checked by the Itô's formula that the weak solution to the SDE (2.10) will give rise to the L-martingale problem, where $\mathbb{P}$ is the distribution of the law of X. The converse is also true in the sense that if b, σ are measurable and bounded, then the weak solution to the SDE (2.10) exists whenever the L-martingale problem exists a solution (see [42, Theorem 5.3.3]).

Definition 2.13. Let the drift vector b and dispersion matrix σ be measurable functions, and consider the stochastic differential equation (2.10).

(i) Let $(X_t, B_t)_{t\geq 0}, (\tilde{X}_t, B_t)_{t\geq 0}$ be weak solutions to (2.10) with the same Brownian motion $(B_t)_{t\geq 0}$ (relative to possibly different filtrations) on a common probability space $(\Omega, \mathscr{F}, \mathbb{P})$. If they have the same initial value, i.e. $\mathbb{P}(X_0 = \tilde{X}_0) = 1$ and they are indistinguishable, i.e. $\mathbb{P}(X_t = \tilde{X}_t, 0 \leq t < \infty) = 1$, then we say that there is a pathwise uniqueness for the weak solution of (2.10).

(ii) Let $(X_t, B_t)_{t\geq 0}, (\tilde{X}_t, \tilde{B}_t)_{t\geq 0}$ be weak solutions to (2.10) possibly defined for different Brownian motions on different probability spaces. If $X_0 = \tilde{X}_0$ in distribution, and the two processes $X, \tilde{X}$ have the same law, then we say there is a uniqueness in law for (2.10).

In particular, Yamada and Watanabe [142] showed that if pathwise uniqueness holds for weak solutions, then every weak solution was a strong solution.

Proposition 2.14. *Fix $T > 0$, consider the stochastic differential equation*

$$dX_t = b(X_t, t)dt + dB_t,$$

for $0 \leq t \leq T$, where $b(x,t)$ is a progressively measurable, $\mathbb{R}^d$-valued function on $\mathbb{R}^d \times [0, T]$ which satisfies

$$\|b(x,t)\| \leq K(1 + \|x\|)$$

for any $0 \leq t \leq T, x \in \mathbb{R}^d$ *and for some* $K > 0$*. Given the initial condition* X_0 *with probability measure* μ *on* $\mathscr{B}(\mathbb{R}^d)$*, then this stochastic differential equation has a weak solution with initial distribution* μ*. Furthermore, the weak solution is unique in law if* $b(x,t)$ *is uniformly bounded.*

This existence and uniqueness result can be found in [75, Proposition 3.6, Corollary 3.11], whose proof heavily relies on the Girsanov's theorem. On an application of the Girsanov's theorem, the underlying measure has been changed so that the process X defines a new Brownian motion under the new measure. We do not go into details here, but will explain the Girsanov's theorem in the next subsection. The proof of this proposition can be found in [75, Chapter 5.3].

A variation of this approach is the study of martingale problems introduced by [127, 128]. The idea is to formulate the search for the law of a diffusion process with given drift and dispersion coefficients in terms of a martingale problem [75]. With the martingale problem, Stroock and Varadhan [127] established the following existence and uniqueness for weak solutions of stochastic differential equations with bounded, continuous coefficients.

Proposition 2.15. *Consider the following stochastic differential equation*

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dB_t,$$

where the coefficients b, σ *are continuous and bounded. Then for any probability measure* μ *on* $\mathscr{B}(\mathbb{R}^d)$ *with*

$$\int_{\mathbb{R}^d} |x|^{2m}\mu(dx) < \infty \quad \text{for some } m > 1,$$

this SDE admits a weak solution $(X_t)_{t\geq 0}$ *such that the law of* X_0 *coincides with* μ*.*

2.2.3 Girsanov's Theorem

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space and $B = (B_t^1, ..., B_t^d)_{t\geq 0}$ be a d-dimensional Brownian motion defined on this space. Let $(\mathscr{F}_t)_{t\geq 0}$ be the augmented filtration generated by the Brownian motion B.

Suppose $X = (X_t^1, ..., X_t^d)_{t\geq 0}$ is a d-dimensional stochastic process, which is assumed to be adapted and measurable, and also that

$$\mathbb{P}\left(\int_0^t \left|X_s^i\right|^2 ds < \infty\right) = 1$$

for any $0 \leq t < \infty, i = 1, ..., d$, then the stochastic integral $\int_0^t X_s dB_s$, is well-defined and is a continuous local martingale. Set

$$Z_t(X) := \exp\left(\int_0^t X_s dB_s - \frac{1}{2}\int_0^t |X_s|^2 ds\right), \tag{2.12}$$

then it follows from the Itô formula that

$$\begin{aligned} Z_t(X) &= 1 + \int_0^t Z_s(X) X_s dB_s \\ &= 1 + \sum_{i=1}^d \int_0^t Z_s(X) X_s^i dB_s^i. \end{aligned}$$

So $Z(X)$ is a continuous local martingale with $Z_0(X) = 1$.

Proposition 2.16 (Girsanov(1960), Cameron and Martin (1946))**.** *Let $\mathbb{P}$ be a probability measure on a probability space $(\Omega, \mathscr{F})$, and B be a d-dimensional Brownian motion with respect to $\mathbb{P}$. If $Z(X)$ defined in* (2.12) *is a martingale with $\mathbb{E}[Z_t(X)] = 1$ for any $t \geq 0$, then there is a unique probability measure $\tilde{\mathbb{P}}$ on $(\Omega, \mathscr{F})$ such that $\tilde{\mathbb{P}}$ is equivalent to $\mathbb{P}$ on $\mathscr{F}_t$ for all $t \geq 0$, and we have the*

Radon–Nikodyn derivative defined as

$$\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathscr{F}_t} = Z_t(X).$$

Moreover, the process $\tilde{B} = (\tilde{B}_t^1, ..., \tilde{B}_t^d)_{t\geq 0}$ *defined by*

$$\tilde{B}_t^i = B_t^i - \int_0^t X_s^i ds$$

for $i = 1, ..., d, t \geq 0$*, is a Brownian motion under the probability measure* $\tilde{\mathbb{P}}$*.*

This is the so-called Cameron-Martin-Girsanov formula. One may refer to [129, Chapter 6.4] or [75, Chapter 3.5] for a detailed discussion.

This proposition indicates that on the transformation of probability measures from $\mathbb{P}$ to $\tilde{\mathbb{P}}$, the Brownian motion B is transformed to $\tilde{B}$. Therefore, this transformation of probability measures induces a drift term $\int_0^t X_s^i ds$ for Brownian motions. For this reason, this transformation from $\mathbb{P}$ to $\tilde{\mathbb{P}}$ is also called transformation of drift (see [65, 75]). As a important application, this method of transformation of drift provides a device for solving stochastic differential equations driven by Brownian motions. Specifically, with a change of probability measure, the driving Brownian motion becomes, under the new probability measure, the solution to the differential equation. This technique will play an important role in our study in Chapter 3.

In order to use the Cameron-Martin-Girsanov's formula effectively, we need some general conditions ensuring that the local martingale $Z(X)$ is a true martingale. The most well-known results are the Novikov's condition [104] and Kazamaki's condition [76].

Proposition 2.17 (Novikov's criteria)**.** *Let* X *be a vector-valued stochastic pro-*

cess described as above. If X satisfies the condition

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^t |X_s|^2 ds\right)\right] < \infty$$

for any $0 \leq t < \infty$, then $Z(X)$ is a martingale, and $\mathbb{E}[Z_t(X)] = 1$.

Proposition 2.18 (Kazamaki's criteria)**.** *With the above notations, if the process*

$$\exp\left(\frac{1}{2}\int_0^t X_s dB_s\right)$$

is uniformly integrable for any $t \geq 0$, then the process $Z(X)$ is a martingale.

2.3 Backward stochastic differential equations

Backward stochastic differential equations (BSDEs) are equations of the following form

$$dY_t = -f(t, Y_t, Z_t)dt + Z_t dB_t, \quad Y_T = \xi, \tag{2.13}$$

where $(B_t)_{0\leq t\leq T}$ is a standard d-dimensional Brownian motion on some probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{0\leq t\leq T}, \mathbb{P})$, with $(\mathscr{F}_t)_{0\leq t\leq T}$ being the natural Brownian filtration. Here $f : [0, T] \times \mathbb{R}^k \times \mathbb{R}^{k\times d} \to \mathbb{R}^k$ is the generator of the BSDE (2.13) and ξ is the terminal condition. Sometimes we say that (2.13) is a backward stochastic differential equation generated by the pair (f, ξ). Formally, the couple $(Y_t, Z_t)_{0\leq t\leq T}$ adapted to the filtration $(\mathscr{F}_t)_{0\leq t\leq T}$ is called a solution to the backward stochastic differential equation (2.13) if

(i) $(Y_t, Z_t)_{0\leq t\leq T}$ is $\mathscr{F}_t$-progressively measurable and taking values in $\mathbb{R}^k \times \mathbb{R}^{k\times d}$ for all $0 \leq t \leq T$, where the integer k is known as the dimension of the BSDE;

(ii) $\int_0^T (|f(s, Y_s, Z_s)| + |Z_s|^2)\, ds < \infty$, $\mathbb{P}$-a.s.;

(iii) $Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dB_s$, $0 \leq t \leq T$, holds almost surely.

For the convenience of notations, we write (Y, Z) for the process $(Y_t, Z_t)_{0 \leq t \leq T}$. In particular, the terminal time could be a finite stopping time, see [16, 28, 32, 113] for instance. But in our study, we are only concerned with $T > 0$ being deterministic.

The concept of *backward stochastic differential equation* was first introduced by Bismut [15], when studying the stochastic version of the Pontryagin maximum principle, and the adjoint process was written as a linear backward stochastic differential equation. Pardoux and Peng [106] were the first to consider BSDEs in general settings, and proved the existence and uniqueness result under the condition that f was Lipschitz continuous in the variables (Y, Z), and that the terminal data ξ was square-integrable. Then in the seminal work [41], El Karoui et al. explained the applications of backward stochastic differential equations in mathematical finance. Since then, the topic of backward stochastic differential equations has drawn great interest in the research of this area. In particular, Lepeltier and San Martín [85] showed the existence of solutions for one dimensional BSDE when the generator f was continuous with linear growth. Inspired by this work, Kobylanski [78] studied the one-dimensional BSDEs, where f had quadratic growth in Z and the terminal data was bounded a.s.. In their studies, the comparison theorem for Lipschitz backward stochastic differential equations played a crucial role. One may refer to [18, 19, 25, 37, 70, 91, 135, 141] for more recent progress in showing the solvability of backward stochastic differential equations under various growth conditions.

This section is devoted to revisiting some existence and uniqueness results, as well as several key techniques, from the theory of backward stochastic differential equations, which will be used in Chapter 4.

Firstly, we define the following spaces:

- $\mathscr{L}_T^2(\mathbb{R}^k)$, the vector space formed by progressively measurable processes Y :

$\Omega \times [0,T] \to \mathbb{R}^k$ such that

$$\|Y\|^2_{\mathscr{L}^2_T} = \mathbb{E}\left[\sup_{0\le t\le T} |Y_t|^2\right] < \infty.$$

- $\mathscr{H}^\infty_T(\mathbb{R}^k)$, the vector space formed by bounded progressively measurable processes $Y : \Omega \times [0,T] \to \mathbb{R}^k$ such that there exists $M > 0$ with $Y_t \le M$ a.s. for all $t \le T$ and

$$\|Y\|^2_{\mathscr{H}^\infty_T} = \inf\{M \ge 0 : \mathbb{P}(|Y_t| \le M) = 1, \forall t \in [0,T]\} < \infty$$

- $\mathscr{H}^2_T(\mathbb{R}^{k\times d})$, the vector space formed by progressively measurable processes $Z : \Omega \times [0,T] \to \mathbb{R}^{k\times d}$ such that

$$\|Z\|^2_{\mathscr{H}^2_T} = \mathbb{E}\left[\int_0^T |Z_t|^2 dt\right] < \infty.$$

Before moving on to the classical existence and uniqueness theorems, we here emphasise the Burkholder-Davis-Gundy inequality (see [65, Theorem 3.1, page 110] for example), which is a fundamental inequality for continuous local martingales, and will be used several times in later discussion.

Proposition 2.19 (Burkholder-Davis-Gundy inequality). *For every continuous local martingale $(M_t)_{t\ge 0}$, there exist universal constants $c_p, C_p > 0$ for $0 < p < \infty$ such that*

$$c_p \mathbb{E}\left[\left(\max_{0\le s\le t} |M_s|\right)^{2p}\right] \le \mathbb{E}\left[\langle M\rangle_t^p\right] \le C_p \mathbb{E}\left[\left(\max_{0\le s\le t} |M_s|\right)^{2p}\right]$$

for every $t \ge 0$, where $\langle M\rangle$ denotes the quadratic variation process of M.

The following proposition is from the seminal paper [41], which states the existence and uniqueness of a backward stochastic differential equation when the coefficients are uniformly Lipschitz. We also give a proof of this result to illustrate how a fixed point argument is applied in proving the existence and uniqueness of the BSDE.

Proposition 2.20. *Assume the generator* $f : [0,T]\times\mathbb{R}^k\times\mathbb{R}^{k\times d} \to \mathbb{R}^k$ *is uniformly Lipschitz with respect to* (y,z)*, i.e.*

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

for any $0 \leq t \leq T, y, y' \in \mathbb{R}^k, z, z' \in \mathbb{R}^{k\times d}$*, where* $C_L > 0$ *is the Lipschitz constant. Also, assume* f *satisfies the integrability condition that*

$$\mathbb{E}\left[\int_0^T |f(s,0,0)|^2 ds\right] < \infty. \tag{2.14}$$

If the terminal data ξ *is square-integrable, i.e.* $\mathbb{E}[|\xi|^2] < \infty$*, then the BSDE* (2.13) *admits a unique solution* $(Y_t, Z_t)_{0\leq t\leq T} \in \mathscr{L}_T^2(\mathbb{R}^k) \times \mathscr{H}_T^2(\mathbb{R}^{k\times d})$.

Proof. We equip the space $\mathscr{L}_T^2(\mathbb{R}^k) \times \mathscr{H}_T^2(\mathbb{R}^{k\times d})$ with the norm $\|\cdot\|_\beta$, which is defined as

$$\|(Y,Z)\|_\beta = \mathbb{E}\left[\sup_{0\leq t\leq T} e^{\beta t}|Y_t|^2 + \int_0^T e^{\beta s}|Z_s|^2 ds\right]^{1/2}$$

for any $\beta > 0$ and $(Y,Z) \in \mathscr{L}_T^2(\mathbb{R}^k) \times \mathscr{H}_T^2(\mathbb{R}^{k\times d})$. We aim to use a fixed point argument in this space by constructing a non-linear mapping Ψ such that $(Y,Z) \in \mathscr{L}_T^2(\mathbb{R}^k) \times \mathscr{H}_T^2(\mathbb{R}^{k\times d})$ is a solution of the BSDE (2.13) if and only if it is a fixed point of Ψ.

Given $(U,V) \in \mathscr{L}_T^2(\mathbb{R}^k) \times \mathscr{H}_T^2(\mathbb{R}^{k\times d})$, define $(Y,Z) = \Psi(U,V)$ to be the solution of the following BSDE

$$Y_t = \xi + \int_t^T f(s,U_s,V_s)ds - \int_t^T Z_s dB_s. \tag{2.15}$$

Since $f(s, U_s, V_s)$ is progressively measurable, then $\left(\int_0^t f(s, U_s, V_s)ds\right)_{0\le t\le T}$ is adapted. If we take the conditional expectations in (2.15) with respect to $\mathscr{F}_t$, we may write

$$\begin{aligned}Y_t &= \mathbb{E}\left[\xi + \int_t^T f(s, U_s, V_s)ds\middle|\mathscr{F}_t\right]\\&= \mathbb{E}\left[\xi + \int_0^T f(s, U_s, V_s)ds\middle|\mathscr{F}_t\right] - \int_0^t f(s, U_s, V_s)ds.\end{aligned}$$

Apply the martingale representation theorem ([75, Theorem 3.4.15 page 182]) to the martingale part, we construct a predictable process $Z \in \mathscr{H}_T^2(\mathbb{R}^{k\times d})$ such that

$$\mathbb{E}\left[\xi + \int_0^T f(s, U_s, V_s)ds\middle|\mathscr{F}_t\right] = \mathbb{E}\left[\xi|\,\mathscr{F}_0\right] + \int_0^t Z_s dB_s,$$

so that

$$\begin{aligned}Y_t &= \mathbb{E}\left[\xi|\,\mathscr{F}_0\right] - \int_0^t f(s, U_s, V_s)ds + \int_0^t Z_s dB_s\\&= \xi + \int_t^T f(s, U_s, V_s)ds - \int_t^T Z_s dB_s.\end{aligned}$$

Therefore, the processes $(Y_t, Z_t)_{0\le t\le T}$ constructed is a solution of the BSDE

$$Y_t = \xi + \int_t^T f(s, U_s, V_s)ds - \int_t^T Z_s dB_s.$$

It follows from the the square-integrable assumptions on f, ξ and (U, V) that $(Y, Z) \in \mathscr{L}_T^2(\mathbb{R}^k) \times \mathscr{H}_T^2(\mathbb{R}^{k\times d})$.

It remains to show the mapping $(U, V) \mapsto (Y, Z)$ is a contraction. Let the processes $(U, V), (U', V')$ be in the space $\mathscr{L}_T^2(\mathbb{R}^k) \times \mathscr{H}_T^2(\mathbb{R}^{k\times d})$. Define

$$(Y, Z) = \Phi(U, V), \quad (Y', Z') = \Phi(U', V').$$

To simplify the notation, we write

$$
\begin{aligned}
y &:= Y - Y', & z &:= Z - Z', \\
u &:= U - U', & v &:= V - V'.
\end{aligned}
$$

Then

$$
dy_t = -\left(f(t, U, V) - f(t, U', V')\right) dt + z_t dB_t, \quad y_T = 0.
$$

Take any $\beta > 0$, apply the Itô's formula to $e^{\beta t}|y_t|^2$, then

$$
\begin{aligned}
& e^{\beta t}|y_t|^2 + \beta \int_t^T e^{\beta s}|y_s|^2 ds + \int_t^T e^{\beta s}|z_s|^2 ds \\
= {} & 2 \int_t^T e^{\beta s} y_s \left(f(s, U_s, V_s) - f(s, U_s', V_s')\right) ds - 2 \int_t^T e^{\beta s} y_s z_s dB_s \\
\leq {} & 2C_L \int_t^T e^{\beta s}|y_s| \left(|u_s| + |v_s|\right) ds - 2 \int_t^T e^{\beta s} y_s z_s dB_s \\
\leq {} & \frac{2C_L^2}{\epsilon} \int_t^T e^{\beta s}|y_s|^2 ds + \epsilon \int_t^T e^{\beta s} \left(|u_s|^2 + |v_s|^2\right) ds - 2 \int_t^T e^{\beta s} y_s z_s dB_s, \quad (2.16)
\end{aligned}
$$

for any $\epsilon > 0$, where the last inequality comes from the inequality $2ab \leq a^2/\epsilon + \epsilon b^2$. By the Burkholder-Davis-Gundy inequality, there exists a constant $C_B > 0$ such that

$$
\begin{aligned}
\mathbb{E}\left[\sup_{0 \leq t \leq T} \int_0^t e^{\beta s} y_s z_s dB_s\right] &\leq e^{\beta T} C_B \mathbb{E}\left[\left(\int_0^t |y_s|^2 |z_s|^2 ds\right)^{1/2}\right] \\
&\leq e^{\beta T} C_B \mathbb{E}\left[\sup_{0 \leq t \leq T} |y_t| \left(\int_0^t |z_s|^2 ds\right)^{1/2}\right] \\
&\leq \frac{e^{\beta T} C_B}{2} \left(\mathbb{E}\left[\sup_{0 \leq t \leq T} |y_t|^2\right] + \mathbb{E}\left[\int_0^T |z_s|^2 ds\right]\right),
\end{aligned}
$$

where the last inequality follows as $ab \leq a^2/2 + b^2/2$. Therefore, it follows from the integrability assumption of y, z that

$$\mathbb{E}\left[\sup_{0\le t\le T}\int_0^t e^{\beta s} y_s z_s dB_s\right] < \infty.$$

So the local martingale $\int_0^t e^{\beta s} y_s z_s dB_s$ is a true martingale, which is null at $t = 0$. Choose $\beta = 2C_L^2/\epsilon$ and take expectations on both sides of (2.16), it yields

$$\mathbb{E}\left[\int_0^T e^{\beta s}|z_s|^2 ds\right] \le \epsilon\mathbb{E}\left[\int_0^T e^{\beta s}\left(|u_s|^2 + |v_s|^2\right) ds\right].$$

On the other hand, take expectations on the supremum of (2.16) and rearranging to obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t}|y_t|^2\right] \\
&\le \epsilon\int_t^T e^{\beta s}\left(|u_s|^2 + |v_s|^2\right) ds - 2\mathbb{E}\left[\sup_{0\le t\le T}\int_t^T e^{\beta s} y_s z_s dB_s\right] \\
&= \epsilon\int_t^T e^{\beta s}\left(|u_s|^2 + |v_s|^2\right) ds + 2\mathbb{E}\left[\sup_{0\le t\le T}\int_0^t e^{\beta s} y_s z_s dB_s\right] \\
&\le \epsilon\int_t^T e^{\beta s}\left(|u_s|^2 + |v_s|^2\right) ds + 2C_B\mathbb{E}\left[\left(\int_0^T e^{2\beta s}|y_s|^2|z_s|^2 ds\right)^{1/2}\right] \\
&\le \epsilon\int_t^T e^{\beta s}\left(|u_s|^2 + |v_s|^2\right) ds + 2C_B\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t}|y_s|\left(\int_0^T |z_s|^2 ds\right)^{1/2}\right] \\
&\le \epsilon\int_t^T e^{\beta s}\left(|u_s|^2 + |v_s|^2\right) ds + \frac{1}{2}\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t}|y_t|^2\right] + 2C_B^2\mathbb{E}\left[\int_0^T e^{\beta s}|z_s|^2 ds\right],
\end{aligned}$$

where the second inequality follows from the Burkholder-Davis-Gundy inequality. So rearranging and adding the two derived inequalities together to obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T} e^{\beta t}|y_t|^2 + \int_0^T e^{\beta s}|z_s|^2 ds\right] \\
&\le \epsilon(3 + 4C_B^2)\mathbb{E}\left[\int_0^T e^{\beta s}\left(|u_s|^2 + |v_s|^2\right) ds\right]
\end{aligned}$$

$$\leq \epsilon(3+4C_B^2)(1\vee T)\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\beta t}|u_t|^2+\int_0^T e^{\beta s}|v_s|^2 ds\right].$$

Choose $\epsilon>0$ small so that $\epsilon(3+4C_B^2)(1\vee T)=\frac{1}{2}$. Then

$$\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\beta t}|y_t|^2+\int_0^T e^{\beta s}|z_s|^2 ds\right]\leq \frac{1}{2}\mathbb{E}\left[\sup_{0\leq t\leq T} e^{\beta t}|u_t|^2+\int_0^T e^{\beta s}|v_s|^2 ds\right],$$

which implies that the mapping $(U,V)\mapsto(Y,Z)$ is a strict contraction from $\mathscr{L}_T^2(\mathbb{R}^k)\times\mathscr{H}_T^2(\mathbb{R}^{k\times d})$ to itself, when equipped with the norm $\|\cdot\|_\beta$. Thus, the mapping Ψ has a unique fixed point, which ensures the existence and uniqueness of a solution of the BSDE (2.13) in $\mathscr{L}_T^2(\mathbb{R}^k)\times\mathscr{H}_T^2(\mathbb{R}^{k\times d})$. $\square$

Another important result in the theory of backward stochastic differential equations is the comparison theorem, which permits to compare the solutions of two one-dimensional BSDEs as soon as one can compare the terminal conditions and the generators (see [41, 106] for example).

Proposition 2.21 (Comparison Theorem)**.** *Suppose $k=1$, that is,* (2.13) *is a one-dimensional BSDE. Let $(f^1,\xi^1),(f^2,\xi^2)$ be parameters of the BSDEs* (2.13) *such that f^i is globally Lipschitz satisfying the integrability condition* (2.14) *and ξ^i is square-integrable for $i=1,2$. Let (Y^1,Z^1) and (Y^2,Z^2) be the corresponding square-integrable solutions, respectively. Suppose also that $\xi^1\leq\xi^2$, $\mathbb{P}$-a.s., and that $f^1(t,Y_t^2,Z_t^2)\leq f^2(t,Y_t^2,Z_t^2)$, $d\mathbb{P}\otimes dt$-a.s.. Then $Y_t^1\leq Y_t^2$ a.s. for any time $t\in[0,T]$.*

This comparison theorem plays a key role in the approximation strategy for showing the existence of the solutions, which involves constructing an increasing sequence of BSDEs with Lipschitz generators. In particular, Lepeltier and San Martin [85] applied this strategy to study BSDEs whose generator was continuous with a linear growth. Kobylanski [78] also used this method to study quadratic BSDEs, whose generating function had a quadratic growth in the parameter Z.

Here we revisit the main argument of this approximation method, which will be used in Chapter 4.

Proposition 2.22. *Let ξ be a real-valued square-integrable process, and $f : [0,T] \times \Omega \times \mathbb{R} \times \mathbb{R}^d$ be measurable such that it satisfies the following conditions:*

(i) linear growth: there exists $K < \infty$ such that $|f(t,\omega,y,z)| \leq K(1+|y|+|z|)$ for all $t \in [0,T], y \in \mathbb{R}, z \in \mathbb{R}^d$;

(ii) $f(t,\omega,\cdot,\cdot)$ is continuous for any given $t \in [0,T], \omega \in \Omega$.

Then the BSDE

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

admits an adapted solution $(Y,Z) \in \mathscr{H}_T^2(\mathbb{R}) \times \mathscr{H}_T^2(\mathbb{R}^d)$.

Outline of proof. Here we outline the proof of the existence of the solution using approximation method. One may refer to [85] for a rigorous proof.

Step 1: Construct a sequence of approximating functions that are Lipschitz continuous.

Suppose f satisfies the above assumptions, we may construct a sequence of functions

$$f_n(y,z) = \inf_{(y',z') \in \mathbb{Q}^{d+1}} \{f(y,z) + n|y-y'| + n|z-z'|\}.$$

Then f_n is well-defined for any $n \geq K$ and it satisfies the following properties:

(i) Linear growth: $|f_n(y,z)| \leq K(1+|y|+|z|)$ for all $y \in \mathbb{R}, z \in \mathbb{R}^d$;

(ii) Monotone in n: $f_n(y,z) \leq f_{n+1}(y,z)$ for any given $(y,z) \in \mathbb{R} \times \mathbb{R}^d$;

(iii) Lipschitz: $|f_n(y,z) - f_n(y',z')| \leq n\,(|y-y'| + |z-z'|)$ for any $(y,z) \in \mathbb{R} \times \mathbb{R}^d$;

(iv) Strong convergence: if $y_n \to y, z_n \to z$ then $f_n(y_n,z_n) \to f(y,z)$ as $n \to \infty$.

With the Lipschitz condition, the BSDE generated by (f_n, ξ) admits a unique solution $(Y^n, Z^n) \in \mathscr{L}_T^2(\mathbb{R}) \times \mathscr{H}_T^2(\mathbb{R}^d)$ for each $n \in \mathbb{N}$. In particular, $(Y^n, Z^n) \in \mathscr{H}_T^2(\mathbb{R}) \times \mathscr{H}_T^2(\mathbb{R}^d)$.

Step 2: Show $(Y^n)_n, (Z^n)_n$ are uniformly bounded in $\mathscr{H}_T^2(\mathbb{R}) \times \mathscr{H}_T^2(\mathbb{R}^d)$.

If one defines

$$h(t, \omega, y, z) = K(1 + |y| + |z|)$$

for any $y \in \mathbb{R}, z \in \mathbb{R}^d$, then h is Lipschitz continuous. So the BSDE generated by (h, ξ) has a unique solution, say $(U, V) \in \mathscr{H}_T^2(\mathbb{R}) \times \mathscr{H}_T^2(\mathbb{R}^d)$. Then by the comparison Theorem,

$$Y^m \leq Y^n \leq U,$$

almost surely for any $n \geq m \geq K$. So one may conclude that there exists a constant $B > 0$ depending on K, T and $\mathbb{E}[\xi]$ such that

$$\sup_{n \geq K} \|Y^n\|_{\mathscr{H}_T^2} \leq B, \quad \|U\|_{\mathscr{H}_T^2} \leq B.$$

By applying the Itô's formula to the process $|Y^n|^2$, it yields

$$|Y_t^n|^2 = |\xi|^2 + 2\int_t^T Y_s^n f_n(s, Y_s^n, Z_s^n) ds - \int_t^T |Z_s^n|^2 ds - 2\int_t^T Y_s^n Z_s^n dB_s.$$

Taking expectations on both sides, together with the linear growth condition,

$$\begin{aligned}
\mathbb{E}\left[\int_0^T |Z_s^n|^2 ds\right] &\leq \mathbb{E}\left[|\xi|^2\right] + 2K\mathbb{E}\left[\int_0^T |Y_s^n|^2 ds\right] + 2K\mathbb{E}\left[\int_0^T |Y_s^n|(1 + |Z_s^n|) ds\right] \\
&\leq \mathbb{E}\left[|\xi|^2\right] + 2K\mathbb{E}\left[\int_0^T |Y_s^n|^2 ds\right] \\
&\quad + K\mathbb{E}\left[\int_0^T \left(\frac{1}{\epsilon^2} + 2\epsilon^2 |Y_s^n|^2 + \frac{1}{\epsilon^2}|Z_s^n|^2\right) ds\right]
\end{aligned}$$

for any $\epsilon > 0$. By choosing $\epsilon^2 > K$, we deduce for any $n \geq K$ that

$$\|Z^n\|^2_{\mathscr{H}^2_T} \leq \frac{\mathbb{E}\left[|\xi|^2\right] + TK/\epsilon^2 + 2K(1+\epsilon^2)B^2}{1 - K/\epsilon^2},$$

that is, the sequence $(Z^n)_n$ is uniformly bounded in $\mathscr{H}^2_T(\mathbb{R}^d)$.

Step 3: Show the sequence $(Y^n, Z^n)_n$ converges in $\mathscr{H}^2_T(\mathbb{R}) \times \mathscr{H}^2_T(\mathbb{R}^d)$.
By the comparison theorem, $(Y^n)_n$ is an increasing sequence of processes. Since it is also bounded in $\mathscr{H}^2_T(\mathbb{R})$, then it follows from the dominated convergence theorem that $(Y^n)_n$ converges in $\mathscr{H}^2_T(\mathbb{R})$ to some process Y. Again, by applying the Itô's formula to the process $|Y^n - Y^m|^2$ for $n, m \geq K$, it yields

$$\begin{aligned}
&\mathbb{E}\left[|Y^n_t - Y^m_t|^2\right] + \mathbb{E}\left[\int_t^T |Z^n_s - Z^m_s|^2\, ds\right] \\
&\leq 2\left(\mathbb{E}\left[\int_t^T |Y^n_s - Y^m_s|^2\, ds\right]\right)^{1/2} \\
&\quad \times \left(\mathbb{E}\left[\int_t^T |f_n(s, Y^n_s, Z^n_s) - f_m(s, Y^m_s, Z^m_s)|^2\, ds\right]\right)^{1/2}.
\end{aligned} \tag{2.17}$$

Take $t = 0$, then by the uniform linear growth condition of the generators f_n and the fact that $(Y^n, Z^n)_n$ is uniformly bounded in $\mathscr{H}^2_T(\mathbb{R}) \times \mathscr{H}^2_T(\mathbb{R}^d)$, there exists a constant C depending on K, T such that

$$\|Z^n - Z^m\|^2_{\mathscr{H}^2_T(\mathbb{R}^d)} \leq C\,\|Y^n - Y^m\|^2_{\mathscr{H}^2_T(\mathbb{R})}.$$

Thus, (Z^n) is a Cauchy sequence in $\mathscr{H}^2_T(\mathbb{R}^d)$. Hence, it converges in $\mathscr{H}^2_T(\mathbb{R}^d)$.

Step 4: The last step is to pass on the limit by showing $(Y^n)_n$ converges uniformly in t to Y.
From (2.17) we have

$$\mathbb{E}\left[|Y^n_t - Y^m_t|^2\right] + \mathbb{E}\left[\int_t^T |Z^n_s - Z^m_s|^2\, ds\right] \leq C\,\|Y^n - Y^m\|^2_{\mathscr{H}^2_T(\mathbb{R})}$$

for some positive constant C depending on K, T, that is, $(Y^n)_n$ converges to Y in mean square. As a consequence, there exists a subsequence $(Y^{n_k})_k$ that converges to Y a.s.. Due to the monotonicity of $(Y^n)_n$, the whole sequence $(Y^n)_n$ converges to Y a.s., and $\sup_n |Y^n|$ is $dt \otimes d\mathbb{P}$-integrable.

On the other hand, since $(Z^n)_n$ converges to some Z in $\mathscr{H}_T^2(\mathbb{R})$, then $\sup_n |Z^n|$ is $dt \otimes d\mathbb{P}$-integrable, and there exists a subsequence of $(Z^n)_n$ such that $(Z^{n_k})_k$ converges to Z a.s.. Hence, we have

$$\lim_{k\to\infty} f_k(t, Y_t^{n_k}, Z_t^{n_k}) = f(t, Y_t, Z_t) \quad a.s.,$$

and

$$|f_k(t, Y_t^{n_k}, Z_t^{n_k})| \leq K(1 + \sup_k |Y_t^{n_k}| + \sup_k |Z_t^{n_k}|).$$

Then by the dominated convergence theorem,

$$\lim_{k\to\infty} \int_t^T f_k(s, Y_s^{n_k}, Z_s^{n_k}) ds = \int_t^T f(s, Y_s, Z_s) ds$$

for almost all $\omega \in \Omega$ and uniformly in $t \in [0, T]$. From the continuity property of the stochastic integral, we have

$$\lim_{k\to\infty} \sup_{0\leq t\leq T} \left| \int_t^T Z_s^{n_k} dB_s - \int_t^T Z_s dB_s \right| = 0 \quad \text{in probability.}$$

By passing to a subsequence, say $(Z^{n_{k_j}})_j$, we may assume that this convergence is $\mathbb{P}$-a.s.. Finally, we consider the dynamics of $|Y^n - Y^{n_{k_j}}|$, that is,

$$\begin{aligned} \left| Y_t^n - Y_t^{n_{k_j}} \right| \leq & \int_t^T \left| f_n(s, Y_s^n, Z_s^n) - f_{n_{k_j}}(s, Y_s^{n_{k_j}}, Z_s^{n_{k_j}}) \right| ds \\ & + \left| \int_t^T Z_s^n dB_s - \int_t^T Z_s^{n_{k_j}} dB_s \right|. \end{aligned}$$

Let $j \to \infty$ and taking the supremum over t, we get from the dominated convergence theorem and the almost sure convergence of the stochastic integral that

$$\begin{aligned}\sup_{0\le t\le T}|Y_t^n - Y_t| &\le \int_0^T |f_n(s,Y_s^n,Z_s^n)-f(s,Y_s,Z_s)|\,ds\\ &\quad+\sup_{0\le t\le T}\left|\int_t^T Z_s^n dB_s-\int_t^T Z_s dB_s\right|\\ &\to 0\end{aligned}$$

almost surely as $n\to\infty$. Hence, Y^n converges to Y uniformly in t. Taking the limit of $n\to\infty$ in the following sequence

$$Y_t^n=\xi+\int_t^T f_n(s,Y_s^n,Z_s^n)ds-\int_t^T Z_s^n dB_s,$$

one concludes that (Y,Z) is an adapted solution in $\mathscr{H}_T^2(\mathbb{R})\times\mathscr{H}_T^2(\mathbb{R}^d)$. $\square$

With the same philosophy, Kobylanski [78, Theorem 2.3, page 565] established the existence of the solution for a class of quadratic BSDEs using a similar approximation method.

Proposition 2.23 (Quadratic BSDE)**.** *Let $\alpha_0,\beta_0,b\in\mathbb{R}$, c be a continuous increasing function. Suppose the continuous function $F:[0,T]\times\mathbb{R}\times\mathbb{R}^d\to\mathbb{R}$ satisfies the following condition:*

$$F(t,y,z)=a_0(t,y,z)y+F_0(t,y,z),$$

with

$$\beta_0\le a_0(t,y,z)\le\alpha_0,\quad a.s.,$$

$$|F_0(t,y,z)|\le b+c(|y|)|z|^2,\quad a.s..$$

If ξ is a bounded real-valued process, then the BSDE generated by (F, ξ), i.e.

$$dY_t = -F(t, Y_t, Z_t)dt + Z_t dB_t, \quad Y_T = \xi,$$

has at least one solution (Y, Z) such that the process $Y \in \mathscr{H}_T^\infty(\mathbb{R})$ and $Z \in \mathscr{H}_T^2(\mathbb{R}^d)$.

3 McKean–Vlasov type SDE arising from random vortex method

This chapter studies the solvability and the numerical approximation of the following McKean–Vlasov type stochastic differential equation

$$\begin{cases} dX^i(x,t) = \left(\sum_{j=1}^d \int_{\mathbb{R}^d} \mathbb{E}\left[K^{ij}(z - X(y,t))\right]\big|_{z=X(x,t)}\, \omega_0^j(y)dy\right) dt + \sqrt{2\nu}dB^i(t), \\ X(x,t) = x, \quad x \in \mathbb{R}^d \end{cases} \tag{3.1}$$

where $i = 1, \ldots, d$, $\nu > 0$ is a constant (which has its origin in fluid mechanics, namely the kinetic viscosity), $B = (B^1, \ldots, B^d)$ is a standard Brownian motion on some probability space and $\omega_0 = (\omega_0^1, \ldots, \omega_0^d)$ is the initial data to the corresponding non-linear (and non-local) vorticity equation. The structure kernel function $K = (K^{ij})$, which defines the SDE (3.1), is a $d \times d$ matrix-valued Borel measurable function. It is assumed to be continuous except at several singularities.

This distribution-dependent SDE (3.1) has its origin from Chorin [27] when applying the random vortex method to study the Navier–Stokes equations in dimension 2. Chorin suggested convolving the singular kernel K with some cutoff functions and assumed the convergence of the cutoff model to the original model. This convergence was not mathematically proved until Marchioro and Pulvirenti [96]. Followed by Osada [105], a propagation of chaos result was established for large viscosities.

As written at the end of [96], Marchioro and Pulvirenti wished to give a unified approach to the Navier–Stokes and Euler equations. So they did not fully exploit the stochastic nature of the Navier–Stokes equation. Méléard [98] followed their framework and proved the existence and uniqueness of the weak solution to the equation (3.1) when K is the two-dimensional Biot–Savart kernel. In addition, a propagation of chaos result for this particle system was established.

Fontbona [45] then extended the study to the three-dimensional problems, where a probabilistic interpretation of local mild solutions in the L^p space was established. Then the author constructed trajectorial stochastic particle approximations to the mild solutions of ω and u. For detailed discussion on various approaches for the random vortex method, one may refer to [95, 140], etc.

Though a few papers discussed the approximation schemes for this McKean–Vlasov type equation (3.1) in dimension 2 with Biot–Savart kernel (for example [44, 90, 124]), and a weak solution was constructed in [98, 145, 146], a strong solution has not yet been established. This chapter devices a new approach to show the existence and uniqueness of the strong solution as well as the weak solution. In particular, our results apply to the Green kernels (such as $1/|x|^{d-1}$ for $d > 2$), the Riesz kernels $1/|x|^{\gamma}$ where $\gamma \in [0, d)$ on $\mathbb{R}^d$ and many other singular integral kernels.

This chapter is a submitted paper, see [116], and it is organised as the following. In section 1, we will explain the random vortex method and formulate the problem (3.1) more precisely. In section 2, we revisit some key properties of the Taylor's diffusions along with the heat kernel estimates, and prove several technical estimates which will be used in proving our main results. In section 3, we define a non-linear mapping, which is crucial to construct both the strong and the weak solution of (3.1). We will then show this non-linear mapping is a contraction so that the SDE admits a unique weak solution. Section 4 shows that (3.1) admits a strong solution and the drift vector field

$$b^i(x,t) = \sum_{j=1}^{d} \int_{\mathbb{R}^d} \mathbb{E}\left[K_j^i(x - X(y,t))\right] \omega_0^j(y)dy$$

is Hölder continuous under certain growth conditions on the kernel function K.

Section 5 recovers the solution to the corresponding non-linear PDEs by using the solutions to (3.1), which can be considered as a probabilistic representation for this class of non-local and non-linear PDEs. Section 6 discusses the numerical approximation schemes for the SDE (3.1). Then the velocity fields are simulated in the 3D space.

The following set of conventions are employed throughout this Chapter. Firstly Einstein's convention on summation on repeated indices through their ranges is assumed, unless otherwise specified. If u is a vector field on $\mathbb{R}^3$ then $\nabla \wedge u$ denotes the curl of u which is again a vector field on $\mathbb{R}^3$ with its component $\varepsilon^{ijk}\frac{\partial u^k}{\partial x_j}$. If $f(x,t)$ is a function on $\mathbb{R}^d \times [0,T]$ then $\|f\|_{L^p(\mathbb{R}^d \times [0,T])}$ or, if no confusion is possible, $\|f\|_p$ denotes the L^p-norm with respect to the Lebesgue measure on the product space $\mathbb{R}^d \times [0,T]$.

3.1 Random vortex method – from PDE to SDE

Particle formulations for fluid flows have been studied as a tool for understanding fluid dynamics of turbulence. The underlying idea is simple, originally due to Taylor [132]. Instead of considering the velocity vector field $u(x,t)$ of the flow, one may study the dynamics of trajectories $X(x,t)$ of the fluid particles emitting from x at the moment 0, i.e. the dynamical equation

$$\frac{d}{dt}X(x,t) = u(X(x,t),t), \quad X(x,0) = x$$

and reformulate the equation of motion of the vorticity $\omega = \nabla \wedge u$ into an evolution equation for $X(x,t)$. This is the vortex method, which works well for certain inviscid fluids (see [11, 12, 95] for example).

For viscous incompressible fluid with constant viscosity $\nu > 0$, a natural idea is to consider Brownian particles instead, i.e. $X(x,t)$ is modelled by the Taylor diffusion

$$dX(x,t) = u(X(x,t),t)dt + \sqrt{2\nu}dB_t, \quad X(x,0) = x, \tag{3.2}$$

where B is a standard Brownian motion. Then the vorticity ω may be reritten in terms of the distribution of the Taylor diffusion (see [132, 133]). This approach is called the random vortex method, originally from Chorin [27], developed by [45, 96, 105, 98], etc. For incompressible fluid flows, $u(x,t)$ satisfies the Navier–Stokes equations

$$\frac{\partial}{\partial t}u + u \cdot \nabla u = \nu \Delta u - \nabla p, \tag{3.3}$$

$$\nabla \cdot u = 0 \tag{3.4}$$

where $p(x,t)$ is a scalar function representing the pressure at (x,t). If the fluid is constrained in a finite region, then certain boundary conditions must be identified, but for simplicity, we consider the case where the evolution of the fluid can take place without physical boundary and also without external force. In this case, the implicit boundary condition at infinity is applied: both $u(x,t)$ and $p(x,t)$ tend to zero sufficiently fast as $|x| \to \infty$. This is the model used in the homogeneous turbulence for example. The incompressible condition (3.4) allows to reformulate the first equation (3.3) in terms of the fluid vorticity $\omega = \nabla \wedge u$. So the equation of vorticity motion satisfies the following vorticity equation

$$\frac{\partial}{\partial t}\omega + u \cdot \nabla \omega = \nu \Delta \omega + \omega \cdot \nabla u,$$

where the velocity field u can be recovered from the Poisson equation

$$\Delta u = -\nabla \wedge \omega.$$

At this stage, it is worth to point out a significant difference between the two-dimensional and three-dimensional problems in the random vortex method. We note that the 3D formulation may still apply to the 2D case by identifying the velocity u as a 3D vector $(u^1, u^2, 0)$. So the vorticity, which is identified with a scalar field, may be written as $\omega = (0, 0, \omega)$ with the formulation that $\omega = \frac{\partial u^2}{\partial x_1} - \frac{\partial u^1}{\partial x_2}$. So the term $\omega \cdot \nabla u$ vanishes identically, and the vorticity equation then appears as a linear equation in ω.

3.1.1 Taylor's diffusions

Clearly in our approach, Taylor's diffusions will play a crucial role due to the formulation (3.2). We have discussed this matter in Chapter 2.1, but here we summarize and emphasise some key properties along with the vortex dynamics.

We rewrite the vorticity equation in terms of an elliptic operator

$$\left(\frac{\partial}{\partial t} - L_{-u}\right)\omega = \omega \cdot \nabla u,$$

where

$$L_{-u} = \nu\Delta - u(x,t) \cdot \nabla$$

with $\nu > 0$. In general, if $b(x,t)$ is a time-dependent vector field (here t is the time variable), then the L_b operator is defined as

$$L_b = \nu\Delta + b(x,t) \cdot \nabla.$$

This is a uniformly elliptic operator of second order and is time inhomogeneous in general. This convention will be applied to any time dependent vector field $b(x,t)$ on $\mathbb{R}^d$ where d is not necessary to be 3. If no confusion may arise, the argument (x,t) will be suppressed.

Moreover, L_b is the infinitesimal generator of the Taylor diffusion describing the motion of Brownian particles $(X_t)_{t\geq\tau}$, which can be defined by the Itô's stochastic differential equation

$$dX_t = b(X_t, t)dt + \sqrt{2\nu}dB_t,$$

where $(B)_{t\geq\tau}$ is a d-dimensional standard Brownian motion on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Denote $\mathbb{P}_{\tau,x}$ to be the probability measure of the process X, then X has a transition probability defined as

$$P(\tau, x, t, dy) := \mathbb{P}_{\tau,x}(X_t \in dy)$$

for any $0 \leq \tau < t$. Moreover, if the vector field $b(x,t)$ is bounded and measurable, then [125, Theorem 2.2.2, page 35] indicates that the probability measure has a density (with respect to the Lebesgue measure) in the sense that

$$\mathbb{P}_{\tau,x}\left(X_t \in dy\right) = p_b(\tau, x, t, y)dy$$

for any $0 \leq \tau < t$. This function $p_b(\tau, x, t, y)$ is the transition probability density function of the L_b-diffusion, for $t > \tau \geq 0$ and $x, y \in \mathbb{R}^d$.

On the other hand, the formal adjoint operator of L_b is given by

$$L_b^\star = \nu\Delta - b\cdot\nabla - \nabla\cdot b = L_{-b} - \nabla\cdot b.$$

This is again a diffusion operator if and only if the vector field b is divergence-free. In particular, if $b(\cdot, t)$ is solenoidal then $L_b^\star = L_{-b}$.

The following lemma summarizes the representation formula in Proposition 2.3 and the relation between fundamental solutions of parabolic equations and transition probability density functions in Lemma 2.4, which will be used throughout this chapter.

Lemma 3.1. *Let $b(x,t)$ be a time-dependent Borel measurable and bounded vector field on $\mathbb{R}^d$. Let $\mathit{\Gamma}_{F,L_b^\star}$ be the fundamental solution to the forward parabolic equation $\left(\frac{\partial}{\partial t} - L_b^\star\right) u = 0$ and $\mathit{\Gamma}_{B,L_b}$ the fundamental solution to the backward parabolic equation $\left(\frac{\partial}{\partial t} + L_b\right) u = 0$.*

(i) The following holds:

$$p_b(\tau, \xi, t, x) = \mathit{\Gamma}_{B,L_b}(\xi, \tau; x, t) = \mathit{\Gamma}_{F,L_{-b}-\nabla\cdot b}(x, t; \xi, \tau)$$

for all $0 \leq \tau < t$ and $\xi, x \in \mathbb{R}^d$.

(ii) For given $\tau \geq 0$, φ and g, the function

$$w(x,t) = \int p_b(\tau, \xi, t, x)\varphi(\xi)d\xi + \int_\tau^t \int p_b(s, \xi, t, x)g(\xi, s)d\xi ds,$$

for $x \in \mathbb{R}^d, t \geq \tau$, solves the initial value problem of the parabolic equation:

$$\left(\frac{\partial}{\partial t} - L_{-b} - \nabla \cdot b\right) w = g, \quad w(\cdot, \tau) = \varphi.$$

(iii) If in addition $b(x,t)$ is solenoidal, i.e. $\nabla \cdot b = 0$, then

$$p_b(\tau, \xi, t, x) = \mathit{\Gamma}_{F,L_{-b}}(x, t, \xi, \tau)$$

for all $0 \leq \tau < t$ and $\xi, x \in \mathbb{R}^d$.

In fact, the results in the previous lemma hold under much weaker conditions on b and can be generalised to a large class of elliptic operators, see [4, 5, 46, 126] and other standard literature on parabolic equations for details.

3.1.2 An archetypical example

Suppose the vorticity $\omega(x,t)$ of an incompressible fluid flow with velocity $u(x,t)$, without applying external force, always lies in the kernel of the rate-of-strain tensor, so that $\omega \cdot \nabla u = 0$ identically. Then the vorticity equation becomes

$$\left(\frac{\partial}{\partial t} - L_{-u(x,t)}\right)\omega(x,t) = 0 \tag{3.5}$$

with initial data $\omega(\cdot,0) = \omega_0$. According to Lemma 3.1, the vorticity equation may be represented by the fundamental solution to the forward parabolic equation (3.5), denoted by $p_u(0,y,t,x)$, that

$$\omega(x,t) = \int_{\mathbb{R}^3} p_u(0,y,t,x)\omega_0(y)dy. \tag{3.6}$$

In particular, p_u is identical to the transition probability density function of the L_u-diffusion. On the other hand, the velocity field u can be recovered from the Poisson equation $\Delta u = -\nabla \wedge \omega$. If $G_0(x) = \frac{1}{4\pi|x|}$ is the Green's function for the Laplace operator in 3D, then according to the Biot–Savart law,

$$u(x,t) = \int_{\mathbb{R}^3} G(x-z) \wedge \omega(z,t)dz, \tag{3.7}$$

where $G(x) = \nabla G_0(x) = -\frac{1}{4\pi}\frac{x}{|x|^3}$ is the vector valued singular kernel in $\mathbb{R}^3$. Since ω is a solution to (3.5) and can be represented by (3.6), we are able to rewrite the velocity field (3.7) in terms of the fundamental solution p_u, that is,

$$\begin{aligned} u(x,t) &= \int\int G(x-z) \wedge \omega_0(y)p_u(0,y,t,z)dzdy \\ &= \int\left(\int G(x-z)p_u(0,y,t,z)dz\right) \wedge \omega_0(y)dy \\ &= \int \mathbb{E}\left[G(x-X(y,t))\right] \wedge \omega_0(y)dy, \end{aligned} \tag{3.8}$$

where $X(y,t)$ is the Taylor diffusion process with infinitesimal generator L_u starting at y at $t=0$. That is, $X(y,t)$ is the solution to the stochastic differential equation

$$dX(y,t) = u(X(y,t),t)dt + \sqrt{2\nu}dB_t, \quad X(y,0) = y, \tag{3.9}$$

where $(B_t)_{t\geq 0}$ is a standard Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Substituting (3.8) into (3.9), we may rewrite the previous stochastic differential equation as

$$\begin{cases} dX(x,t) = \left(\int \mathbb{E}\left[G\left(z - X(y,t)\right)\right] \wedge \omega_0(y)dy\right)\Big|_{z=X(x,t)} dt + \sqrt{2\nu}dB, \\ X(x,0) = x, \end{cases}$$

where x runs through the state space $\mathbb{R}^3$. This is the archetypical example of the SDEs we are going to study in the present section.

3.1.3 Formulation of the problem

Although our main examples come from the study of fluid dynamics, it will be beneficial formulating the problem in a more general setting. Still, we restrict our study to vector fields on the Euclidean space $\mathbb{R}^d$, though the methods and the results can be generalised to tensor fields with certain modifications.

Let $K(x) = (K^{ij}(x))$ be a $d \times d$ matrix-valued 'singular integral' kernel, where K^{ij} are Borel measurable and locally integrable. We are interested in the following stochastic differential equation, written in componentwise,

$$\begin{cases} dX^i(x,t) = \left(\sum_{j=1}^d \int_{\mathbb{R}^d} \mathbb{E}\left[K^{ij}\left(z - X(y,t)\right)\right] \omega_0^j(y)dy \Big|_{z=X(x,t)} \right) dt + dB_t^i, \\ X(x,0) = x, \end{cases} \tag{3.10}$$

where $i = 1, \ldots, d$, $\omega_0 = (\omega_0^i)$ is the initial data, and $B = (B^i)$ is a d-dimensional standard Brownian motion on some probability space. Before we carry out a study of this class of SDEs, let us reformulate (3.10) in a different form to facilitate our approach.

If μ is a measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$, then $K \star \mu = (K^{ij} \star \mu)$ denotes the convolution of K and the measure μ such that

$$K^{ij} \star \mu(x) = \int_{\mathbb{R}^d} K^{ij}(x-y)\mu(dy),$$

for $i, j = 1, \ldots, d$, as long as the right-hand side is well defined.

If U is an $\mathbb{R}^d$-valued random variable on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and its distribution is denoted by $\mathscr{L}(U)$, then by definition

$$K^{ij} \star \mathscr{L}(U)(x) = \mathbb{E}\left[K^{ij}(x-U)\right].$$

If, in addition, the law of U has a probability density function $p(x)$, then

$$K^{ij} \star \mathscr{L}(U)(x) = \int_{\mathbb{R}^d} K^{ij}(x-y)p(y)dy,$$

where the right-hand side is the convolution of K and the function p.

After having introduced the basic data K and ω_0 and the notations, we are now in a position to reformulate the SDE we are going to study:

$$dX^i(x,t) = \left[\sum_{j=1}^{d} \int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(X(y,t))\right)(X(x,t))\omega_0^j(y)dy\right] dt + dB_t^i \tag{3.11}$$

with initial value $X(x,0) = x$ for $x \in \mathbb{R}^d$ and $i = 1, \ldots, d$. The concepts of strong and weak solutions to (3.11) may be defined accordingly as in Chapter 2:

Definition 3.2. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space. Then the pair $(X_t, B_t)_{t\geq 0}$ is called a solution of the distribution-dependent stochastic differential equation (3.11) if

(i) $(\mathscr{F}_t)_{t\geq 0}$ is a filtration of sub-σ-fields of $\mathscr{F}$ satisfying the usual conditions and right-continuity such that $(B_t)_{t\geq 0}$ is adapted to $(\mathscr{F}_t)_{t\geq 0}$,

(ii) $(X_t)_{t\geq 0}$ is a continuous, adapted $\mathbb{R}^d$-valued process,

(iii) the integral equation

$$X^i(x,t) = x^i + \sum_{j=1}^{d} \int_0^t \left[\int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(X(y,s)) \right) (X(x,s)) \omega_0^j(y) dy \right] ds + B_t^i$$

holds $\mathbb{P}$-a.s. for any $i = 1, ..., d$ and $t \geq 0$, where $\mathscr{L}(X(y,s))$ denotes the law of $X(y,s)$ with respect to $\mathbb{P}$.

We do not repeat the definition for strong solutions here, but note that a weak solution is a strong solution if the $(B_t)_{t\geq 0}$ is any fixed Brownian motion on $(\Omega, \mathscr{F}, \mathbb{P})$, and $(\mathscr{F}_t)_{t\geq 0}$ is the filtration generated by the Brownian motion $(B_t)_{t\geq 0}$ and augmented with the $\mathbb{P}$-null sets.

In order to better illustrate our idea of showing the solvability of the McKean-Vlasov equation (3.10), it will be convenient to introduce the following notations. If $Z = (Z(x,t))_{t\geq 0}$ is a family of continuous processes on some probability space, which is jointly continuous in (x,t), then we may define a vector field, denoted by b_Z, whose components are given by

$$b_Z^i(x,t) = \sum_{j=1}^{d} \int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(Z(y,t)) \right) (x) \omega_0^j(y) dy,$$

for $i = 1, \ldots, d$. Notice that by definition, b_Z depends only on the one-dimensional marginal distributions of the process $(Z(y,t))_{t\geq 0}$.

Suppose $b(x,t)$ is a time dependent vector field on $\mathbb{R}^d$, we may define another t-dependent vector field on $\mathbb{R}^d$, denoted by $K \diamond b(x,t)$, such that its i-th component is given by

$$\sum_{j=1}^{d} \int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(Z(y,t)) \right)(x) \omega_0^j(y) dy,$$

where $Z = (Z(y,t))_{t \geq 0}$ is the L_b-diffusion started at y at the moment $t = 0$, so that $K \diamond b = b_Z$. Since $Z(y,t)$ has a transition probability density $p_b(0,y,t,z)$, we can write

$$(K \diamond b)^i(x,t) = \int_{\mathbb{R}^d} \left[\sum_{j=1}^{d} \int_{\mathbb{R}^d} K^{ij}(x-z) \omega_0^j(y) p_b(0,y,t,z) dz \right] dy.$$

We therefore have constructed the mapping $b \mapsto K \diamond b$ which sends a vector field $b(x,t)$ to the vector field $K \diamond b(x,t)$. This non-linear mapping will play a crucial role in our study.

Under the above notations, we may rewrite SDE (3.11) as

$$dX(x,t) = b_X(X(x,t),t)dt + dB_t, \quad X(x,0) = x,$$

for $x \in \mathbb{R}^d$. The following simple observation indeed leads to the approach we are going to develop in what follows.

Lemma 3.3. *Let $b(x,t)$ be a bounded and Borel measurable vector field on $\mathbb{R}^d$, depending on time t. Suppose $K \diamond b = b$ and (X,B) is a weak solution to the following SDE*

$$dX^i(x,t) = b^i(X(x,t),t)dt + dB_t^i, \quad X(x,0) = x, \tag{3.12}$$

for $i = 1, \ldots, d$, where B is a d-dimensional standard Brownian motion on a probability space. Then

$$dX^i(x,t) = \left[\sum_{j=1}^{d} \int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(X(y,t)) \right)(X(x,t)) \omega_0^j(y) dy \right] dt + dB_t^i, \tag{3.13}$$

for $i=1,\ldots,d$. That is, (X,B) is a weak solution to (3.11).

Similarly, if X is a strong solution to the SDE (3.12)*, then it is also a strong solution to* (3.11).

This lemma follows by definition: since $K\diamond b=b$, then $b=b_X$. As a consequence,

$$\sum_{j=1}^{d}\int_{\mathbb{R}^d}\left(K^{ij}\star\mathscr{L}(X(y,t))\right)(x)\omega_0^j(y)dy=b^i(x,t)$$

and (3.13) follows from (3.12) immediately.

Example 3.4. If $d=3$ and $K=(K^1,K^2,K^3)$, then we set $K^{ij}=\varepsilon^{ikj}K^k$ and SDE (3.13) becomes

$$dX(x,t)=\left(\int_{\mathbb{R}^3}\mathbb{E}\left[K\left(z-X(y,t)\right)\right]\wedge\omega_0(y)dy\right)\Bigg|_{z=X(x,t)}dt+dB_t,\quad X(x,0)=x,$$

which is the random vortex dynamical model, where ω_0 represents the initial vorticity.

3.2 Several facts about diffusions with bounded drifts

In this section, we collect a few facts on diffusion processes and prove several technical estimates which will be used in later discussion.

Let $b(x,t)$ be a Borel measurable vector field on the Euclidean space $\mathbb{R}^d$, dependent on the time parameter $t\geq 0$. It is also assumed that $b(x,t)$ is bounded, i.e. $|b(x,t)|\leq A$ for every x and t, where A is a non-negative constant. Then the unique L_b-diffusion (in the sense of weak solutions) may be constructed by using Cameron-Martin formula. Specifically, let $B=(B_t)_{t\geq 0}$ be a d-dimensional standard Brownian motion on some probability space $(\Omega,\mathscr{F},\mathbb{P})$. Let $\mathscr{F}_t=\sigma\left\{B_s:s\leq t\right\}$ and $\mathscr{F}_\infty=\sigma\left\{B_s:s\geq 0\right\}$ be the natural filtration generated by this Brownian motion. Given x and $\tau \geq 0$, define the process

$$N_b(\tau, x, t) := \int_\tau^t b(B_r - B_\tau + x, r) dB_r - \frac{1}{2} \int_\tau^t |b|^2 (B_r - B_\tau + x, r) dr \tag{3.14}$$

for $t \geq \tau$. Clearly, the process $\big(b(B_t - B_\tau + x, t)\big)_{t \geq \tau}$ is adapted, measurable and bounded a.s. for any $t \geq \tau$. So the process N_b is well-defined. Also, it is trivial to see

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\int_\tau^t |b|^2(B_r - B_\tau + x, r) dr\right)\right] \leq \exp\left(\frac{1}{2}A^2(t-\tau)\right) < \infty$$

for any $t \geq \tau$. So by the Novikov's criteria (see Proposition 2.17), the process

$$R_b(\tau, x, t) := e^{N_b(\tau, x, t)} \tag{3.15}$$

is an exponential martingale, which is called the Cameron-Martin density. If $\tau = 0$, then the symbol τ will be suppressed from the notations.

In the next step, we construct the probability $\mathbb{P}_{\tau,x}$ on $(\Omega, \mathscr{F}_\infty)$ such that

$$\left.\frac{d\mathbb{P}_{\tau,x}}{d\mathbb{P}}\right|_{\mathscr{F}_t} = R_b(\tau, x, t)$$

for all $t \geq \tau$. Then by the Girsanov's theorem (see Proposition 2.16), the process

$$W_t := B_t - B_\tau - \int_\tau^t b(B_r - B_\tau + x, r) dr \tag{3.16}$$

is a Brownian motion under the new probability measure $\mathbb{P}_{\tau,x}$. Notice that if we define

$$Y_t := B_t - B_\tau + x$$

for any $t \geq \tau$, with the help of (3.16), we may then rewrite the process $(Y_t)_{t\geq\tau}$ in terms of $(W_t)_{t\geq\tau}$ such that

$$Y_t = x + \int_\tau^t b(Y_r, r)dr + W_t.$$

This implies that $(Y_t)_{t\geq\tau}$ is a diffusion process with infinitesimal generator L_b, whose underlying Brownian motion is $(W_t)_{t\geq\tau}$. Therefore, the family $\{\mathbb{P}_{\tau,x} : \tau \geq 0, x \in \mathbb{R}^d\}$ on $(\Omega, \mathscr{F}_\infty)$ is a diffusion family with generator L_b (see for example [75, 129]). In particular, for any Borel measurable function f,

$$\int_{\mathbb{R}^d} f(y)p_b(\tau, x, t, y)dy = \mathbb{E}\left[R_b(\tau, x, t)f(B_t - B_\tau + x)\right] \tag{3.17}$$

as long as one of the integrals in the equation makes sense, where $p_b(\tau, x, t, y)$ is the transition probability density function of the L_b-diffusion, and the expectation on the right hand side is taken under the probability measure $\mathbb{P}$.

As we have discussed in Chapter 2, under the bounded assumption of the vector field $b(x,t)$, $p_b(\tau, x, t, y)$ is positive and continuous as a function of (t, y). Moreover, the Aronson estimate (see Lemma 2.5) indicates that for every $T > 0$, there is a constant M depending on A, d and T only, such that

$$\frac{1}{Mt^{d/2}} e^{-M\frac{|y-x|^2}{t}} \leq p_b(\tau, x, \tau + t, y) \leq \frac{M}{t^{d/2}} e^{-\frac{|y-x|^2}{Mt}}$$

for any $\tau \geq 0$ and $\tau < t \leq T$. In our study, we need more precise information about the upper bound, which was obtained in [117, 115]. With the notations established in this chapter, we may rewrite Proposition 2.6 as follows:

Lemma 3.5. *There is a positive universal constant κ, depending only on the* *dimension d and $1 < q < \frac{d}{d-1}$, such that*

$$p_b(\tau, x, \tau+t, y) \leq \frac{1}{(2\pi t)^{d/2}} e^{-\frac{|x-y|^2}{2t}} \left(1 + \kappa A \left(\sqrt{t} + |x-y|\right) e^{\frac{q-1}{2qt}|x-y|^2 + \frac{A^2 t}{2(q-1)}}\right) \tag{3.18}$$

for all $x, y \in \mathbb{R}^d$, $\tau \geq 0$ and $t > 0$.

One may refer to Proposition 2.6 for the precise expression for κ. As a consequence, we establish the following estimates, which will play a crucial role in the proof of our main theorem.

Lemma 3.6. *Let $f \in L^\infty(\mathbb{R}^d)$, $\gamma \in [0, d)$ and $\rho > 0$. Define*

$$I(f, x, t, \rho, \gamma) = \int_{\mathbb{R}^d} \int_{\{|z|<\rho\}} \frac{1}{|z|^\gamma} |f(y)| p_b(0, y, t, x-z) dz dy$$

for any $x \in \mathbb{R}^d$ and $t > 0$. Then there exists a universal positive constant κ_1 depending only on d such that

$$I(f, x, t, \rho, \gamma) \leq \frac{\rho^{d-\gamma}}{d-\gamma} \kappa_1 \|f\|_\infty \left(1 + A\sqrt{t} e^{\frac{A^2 t}{2(q-1)}}\right)$$

for all $x \in \mathbb{R}^d$ and $t > 0$.

Proof. Without losing generality, we may assume that $f \geq 0$. Using the sharp estimate (3.18), we have

$$\begin{aligned} I(f, x, t, \rho, \gamma) \leq & \int_{\mathbb{R}^d} \int_{\{|z|<\rho\}} \frac{f(y)}{|z|^\gamma} \frac{e^{-\frac{|y-x+z|^2}{2t}}}{(2\pi t)^{d/2}} \\ & \times \left(1 + \kappa A \left(\sqrt{t} + |y-x+z|\right) e^{\frac{q-1}{2qt}|y-x+z|^2 + \frac{A^2 t}{2(q-1)}}\right) dz dy \\ = & \int_{\mathbb{R}^d} \int_{\{|z|<\rho\}} \frac{f(y+x-z)}{|z|^\gamma} \frac{e^{-\frac{|y|^2}{2t}}}{(2\pi t)^{d/2}} \\ & \times \left(1 + \kappa A \left(\sqrt{t} + |y|\right) e^{\frac{q-1}{2qt}|y|^2 + \frac{A^2 t}{2(q-1)}}\right) dz dy \end{aligned}$$

$$
\begin{aligned}
&\leq \|f\|_\infty \left(\int_{\{|z|<\rho\}} \frac{1}{|z|^\gamma} dz \right) \\
&\quad \times \left(\int_{\mathbb{R}^d} \frac{e^{-\frac{|y|^2}{2t}}}{(2\pi t)^{d/2}} \left(1 + \kappa A \left(\sqrt{t} + |y| \right) e^{\frac{q-1}{2qt}|y|^2 + \frac{A^2 t}{2(q-1)}}) \right) dy \right) \\
&= \|f\|_\infty \frac{\rho^{d-\gamma}}{d-\gamma} \mathrm{Vol}(S^{d-1}) \\
&\quad \times \left(1 + \kappa A \sqrt{t} e^{\frac{A^2 t}{2(q-1)}} \int_{\mathbb{R}^d} \frac{e^{-\frac{|y|^2}{2t}}}{(2\pi t)^{d/2}} \left(1 + \frac{|y|}{\sqrt{t}} \right) e^{\frac{q-1}{2q} \frac{|y|^2}{t}} dy \right) \\
&= \|f\|_\infty \frac{\rho^{d-\gamma}}{d-\gamma} \mathrm{Vol}(S^{d-1}) \\
&\quad \times \left(1 + \kappa A \sqrt{t} e^{\frac{A^2 t}{2(q-1)}} \int_{\mathbb{R}^d} \frac{e^{-\frac{|y|^2}{2}}}{(2\pi)^{d/2}} (1 + |y|) e^{\frac{q-1}{2q}|y|^2} dy \right) \\
&\leq \frac{\rho^{d-\gamma}}{d-\gamma} \kappa_1 \|f\|_\infty \left(1 + A\sqrt{t} e^{\frac{A^2 t}{2(q-1)}} \right)
\end{aligned}
$$

where

$$
\kappa_1 = \max \left\{ \mathrm{Vol}(S^{d-1}), \kappa \int_{\mathbb{R}^d} (1 + |y|) \frac{e^{-\frac{|y|^2}{2q}}}{(2\pi)^{d/2}} dy \right\},
$$

with $\mathrm{Vol}(S^{d-1})$ being the volume of the unit $(d-1)$-sphere, and the proof is complete. □

From the proof of this lemma, it is easy to understand why we have restricted the power of the kernel to be strictly less than the dimension: we must ensure that the kernel function is locally integrable within the disk $\{z \in \mathbb{R}^d : |z| < \rho\}$ so that the computation for the upper bound of $I(f, x, t, \rho, \gamma)$ could carry on.

We also need the following estimate, which is completely elementary though.

Lemma 3.7. *Let $f \in L^1(\mathbb{R}^d)$ and $\rho > 0$, $\gamma \geq 0$ be two constants. Define*

$$
J(f, x, t, \rho, \gamma) = \int_{\mathbb{R}^d} \int_{\{|z| \geq \rho\}} \frac{1}{|z|^\gamma} |f(y)| p_b(0, y, t, x - z) dz dy
$$

for all $x \in \mathbb{R}^d$ *and* $t > 0$. *Then*

$$J(f, x, t, \rho, \gamma) \leq \frac{1}{\rho^\gamma} \|f\|_{L^1}$$

for all x *and* $t > 0$, $\rho > 0$ *and* $\gamma \geq 0$.

Proof. Note that it follows from the fact that p_b is a density function that

$$\int_{\mathbb{R}^d} p_b(0, y, t, z) dz = 1,$$

and $p_b(0, y, t, z) \geq 0$ for any $t \geq 0, z \in \mathbb{R}^d$. Therefore, with $\gamma \geq 0$,

$$\begin{aligned}
J(f, x, t, \rho, \gamma) &\leq \frac{1}{\rho^\gamma} \int_{\mathbb{R}^d} \int_{\{|z| \geq \rho\}} |f(y)| p_b(0, y, t, x - z) dz dy \\
&\leq \frac{1}{\rho^\gamma} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(y)| p_b(0, y, t, x - z) dz dy \\
&= \frac{1}{\rho^\gamma} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} |f(y)| p_b(0, y, t, z) dz dy \\
&= \frac{1}{\rho^\gamma} \left(\int_{\mathbb{R}^d} |f(y)| dy \right) \left(\int_{\mathbb{R}^d} p_b(0, y, t, z) dz \right) \\
&= \frac{1}{\rho^\gamma} \|f\|_{L^1}
\end{aligned}$$

and the proof is complete. □

3.3 Weak solutions

In this section, we prove under certain conditions that there is a unique weak solution to (3.11). To this end, we make several assumptions on ω_0 and K which will be in force throughout the remainder of the discussion.

Let C_0, C_1 and C_∞ be three non-negative constants. In addition to the continuous and measurable conditions, it is assumed that $K = (K^{ij})$ satisfies the following growth condition: there are two constants $\gamma_1 \in [0, d)$ and $\gamma_2 \geq 0$ such that

$$|K(x)| \leq \frac{C_0}{|x|^{\gamma_1}} \quad \text{for all } x \neq 0 \text{ and } |x| < 1,$$

and

$$|K(x)| \leq \frac{C_0}{|x|^{\gamma_2}} \quad \text{for all } |x| \geq 1.$$

In addition, we assume that the initial vorticity ω_0 is bounded and integrable such that $\|\omega_0\|_{L^1} \leq C_1$ and $\|\omega_0\|_\infty \leq C_\infty$. Choose and fix a number $q \in (1, \frac{d}{d-1})$ and set

$$C_K := C_0 \left(\frac{\kappa_1 C_\infty}{d-\gamma_1} \left(1 + e^{\frac{1}{2(q-1)}} \right) + C_1 \right), \quad T_K := \frac{1}{C_K^2}.$$

The crucial fact about C_K and T_K is that they depend on C_0, C_1, C_∞ and γ_1 only.

Firstly, we show whenever the vector field is bounded, the induced vector field $K \diamond b$ is also bounded, and they have the same bounds.

Lemma 3.8. *If $b(x,t)$ is a time-dependent vector field such that $|b(x,t)| \leq C_K$ for all $x \in \mathbb{R}^d$ and $t \leq T_K$, then $K \diamond b$ is also bounded such that $|K \diamond b(x,t)| \leq C_K$ for all $x \in \mathbb{R}^d$ and $t \leq T_K$.*

Proof. Let $B = \left\{ z \in \mathbb{R}^d : |z| < 1 \right\}$. Then

$$\begin{aligned}
|K \diamond b(x,t)| &= \left| \int_{\mathbb{R}^d} \int_B K(z) \wedge \omega_0(y) p_b(0,y,t,x-z) dz dy \right. \\
&\quad \left. + \int_{\mathbb{R}^d} \int_{\mathbb{R}^d \setminus B} K(z) \wedge \omega_0(y) p_b(0,y,t,x-z) dz dy \right| \\
&\leq \int_{\mathbb{R}^d} \int_B |K(z)| \, |\omega_0(y)| \, p_b(0,y,t,x-z) dz dy \\
&\quad + \int_{\mathbb{R}^d} |\omega_0(y)| \left(\int_{\mathbb{R}^d \setminus B} |K(z)| p_b(0,y,t,x-z) dz \right) dy \\
&=: I_1 + I_2
\end{aligned}$$

The estimate for I_1 follows directly from Lemma 3.6:

$$
\begin{aligned}
I_1 &\leq C_0 \int_{\mathbb{R}^d} \int_{\{|z|<1\}} \frac{1}{|z|^{\gamma_1}} \left|\omega_0(y)\right| p_b(0,y,t,x-z) dz dy \\
&= C_0 I(\omega_0, x, t, 1, \gamma_1) \\
&\leq \frac{\kappa_1 C_0 C_\infty}{d-\gamma_1} \left(1 + C_K \sqrt{t} e^{\frac{C_K^2 t}{2(q-1)}}\right)
\end{aligned}
$$

And by a similar argument as Lemma 3.7, we deduce that

$$
\begin{aligned}
I_2 &\leq C_0 \int_{\mathbb{R}^d} |\omega_0(y)| \left(\int_{\{|z|\geq 1\}} \frac{1}{|z|^{\gamma_1}} p_b(0,y,t,x-z) dz \right) dy \\
&\leq C_0 \int_{\mathbb{R}^d} |\omega_0(y)| \left(\int_{\mathbb{R}^d} p_b(0,y,t,x-z) dz \right) dy \\
&\leq C_0 C_1.
\end{aligned}
$$

Putting the estimates for I_1 and I_2 together, we may conclude that

$$
|K \diamond b(x,t)| \leq \frac{\kappa_1 C_0 C_\infty}{d-\gamma_1} \left(1 + C_K \sqrt{t} e^{\frac{C_K^2 t}{2(q-1)}}\right) + C_0 C_1
$$

for all $x \in \mathbb{R}^d$ and $t \geq 0$. Since $C_K \sqrt{t} \leq 1$ for any $t \leq T_K$, we therefore have

$$
|K \diamond b(x,t)| \leq \frac{\kappa_1 C_0 C_\infty}{d-\gamma_1} \left(1 + e^{\frac{1}{2(q-1)}}\right) + C_0 C_1
$$

for any x and $t \leq T_K$. The conclusion then follows immediately from the definition of C_K and T_K. $\square$

Lemma 3.9. *Let $b(x,t), \tilde{b}(x,t)$ be bounded vector fields such that $|b(x,t)| \leq C_K$ and $|\tilde{b}(x,t)| \leq C_K$ for all $x \in \mathbb{R}^d$ and $t \leq T_K$. Let B be a d-dimensional standard Brownian motion on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, define*

$$
R_b(x,t) = e^{N_b(x,t)}, \quad R_{\tilde{b}}(x,t) = e^{N_{\tilde{b}}(x,t)}
$$

to be the Cameron-Martin density (see (3.14) *and* (3.15)*) with respect to the vector field* $b, \tilde{b}$ *starting at* x *at the moment* $\tau = 0$*, respectively. Then*

$$|N_b(x,t) - N_{\tilde{b}}(x,t)| \leq |M_t| + C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d,[0,t])}, \tag{3.19}$$

where

$$M(x,t) := \int_0^t \left(b - \tilde{b} \right) (B_r + x, r) dB_r. \tag{3.20}$$

Also,

$$R_b(x,t) - R_{\tilde{b}}(x,t) \leq \int_0^1 R_{b_s}(x,t)(N_b - N_{\tilde{b}})(x,t) ds \tag{3.21}$$

for any $x \in \mathbb{R}^d, t \leq T_K$

Proof. By the construction of $N_b, N_{\tilde{b}}$, we may write

$$\begin{aligned} N_b(x,t) - N_{\tilde{b}}(x,t) &= \int_0^t \left(b - \tilde{b} \right) (B_r + x, r) dB_r \\ &\quad - \frac{1}{2} \int_0^t \left(|b|^2 - |\tilde{b}|^2 \right) (B_r + x, r) dr \\ &= M(x,t) + A(x,t), \end{aligned}$$

where $M(x,t)$ is defined as (3.20) and

$$\begin{aligned} A(x,t) &:= -\frac{1}{2} \int_0^t \left(|b|^2 - |\tilde{b}|^2 \right) (B_r + x, r) dr \\ &= -\frac{1}{2} \int_0^t \left(b - \tilde{b} \right) \left(b + \tilde{b} \right) (B_r + x, r) dr. \end{aligned}$$

It is clear that

$$|A(x,t)| \leq C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d,[0,t])}.$$

Therefore, we can write

$$|N_b(x,t) - N_{\tilde{b}}(x,t)| \le |M_t| + C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d,[0,t])}$$

for all $x \in \mathbb{R}^d, t \le T_K$.

By the fundamental theorem of calculus, we may write

$$\begin{aligned} R_b - R_{\tilde{b}} &= \int_0^1 \frac{d}{ds} e^{(1-s)N_{\tilde{b}} + sN_b} ds \\ &= \int_0^1 e^{(1-s)N_{\tilde{b}} + sN_b} (N_b - N_{\tilde{b}}) ds. \end{aligned} \tag{3.22}$$

Define $b_s := (1-s)\tilde{b} + sb$ for $s \in [0,1]$, then $|b_s(x,t)| \le C_K$ for any $x \in \mathbb{R}^d, t \le T_K$, and

$$\begin{aligned} \big((1-s)N_{\tilde{b}} + sN_b\big)(x,t) &= \int_0^t \left((1-s)\tilde{b} + sb\right)(B_r + x, r) dB_r \\ &\quad - \frac{1}{2}\int_0^t \left((1-s)|\tilde{b}|^2 + s|b|^2\right)(B_r + x, r) dr \\ &= N_{b_s}(x,t) - \frac{1}{2}s(1-s)\int_0^t |b - \tilde{b}|^2 (B_r + x, r) dr. \end{aligned}$$

Substituting this equality into (3.22) to obtain

$$\begin{aligned} R_b - R_{\tilde{b}} &= \int_0^1 e^{-\frac{1}{2}s(1-s)\int_0^t |b-\tilde{b}|^2(B_r+x,r)dr} R_{b_s}(N_b - N_{\tilde{b}}) ds \\ &\le \int_0^1 R_{b_s}(N_b - N_{\tilde{b}}) ds. \end{aligned}$$

□

With the estimates on the Cameron-Martin density, we are going to establish another key estimate for the non-linear mapping $b \mapsto K \diamond b$, where $b(x,t)$ is any bounded vector field such that $|b(x,t)| \le C_K$ for all $t \le T_K$.

Lemma 3.10. *There exists a positive constant C_L depending only* on C_0, C_1, C_∞, *such that for any $b(x,t)$ and $\tilde{b}(x,t)$ satisfying that $|b(x,t)| \le C_K$ and $|\tilde{b}(x,t)| \le C_K$ for all x and $t \le T_K$, we have*

$$|K \diamond b(x,t) - K \diamond \tilde{b}(x,t)| \le \left(t + \sqrt{t}\right) C_L \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d,[0,t])} \tag{3.23}$$

for all $x \in \mathbb{R}^d$ and $t \le T_K$.

Proof. We prove this inequality by using the Cameron-Martin formula (see Proposition 2.16). According to the expression (3.17), we may write

$$\begin{aligned} K \diamond b(x,t) &= \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} K(x-z) p_b(0,y,t,z) dz \right] \wedge \omega_0(y) dy \\ &= \int_{\mathbb{R}^d} \mathbb{E}\left[R_b(y,t) K(x - B_t - y) \right] \wedge \omega_0(y) dy. \end{aligned}$$

Split K into a sum $K = K_1 + K_2$ where

$$K_1(z) := \mathbb{1}_{\{|z|<1\}} K(z), \qquad K_2(z) := \mathbb{1}_{\{|z|\ge 1\}} K(z).$$

Then

$$\begin{aligned} K \diamond b(x,t) = & \int_{\mathbb{R}^d} \mathbb{E}\left[R_b(y,t) K_1(x - B_t - y) \right] \wedge \omega_0(y) dy \\ & + \int_{\mathbb{R}^d} \mathbb{E}\left[R_b(y,t) K_2(x - B_t - y) \right] \wedge \omega_0(y) dy. \end{aligned}$$

Let

$$D(x,t) := K \diamond b(x,t) - K \diamond \tilde{b}(x,t).$$

Then by using the previous formula for $K \diamond b$, we have

$$D(x,t) = \int_{\mathbb{R}^d} \mathbb{E}\left[\left(R_b(y,t) - R_{\tilde{b}}(y,t) \right) K_1(x - B_t - y) \right] \wedge \omega_0(y) dy$$

$$
\begin{aligned}
&+\int_{\mathbb{R}^d}\mathbb{E}\left[\left(R_b(y,t)-R_{\tilde{b}}(y,t)\right)K_2(x-B_t-y)\right]\wedge\omega_0(y)dy\\
&=: J_1+J_2,
\end{aligned}
$$

With the inequality (3.21) from the previous lemma, it yields

$$
J_1\leq\int_{\mathbb{R}^d}\mathbb{E}\left[\left(\int_0^1 R_{b_s}(y,t)(N_b(y,t)-N_{\tilde{b}}(y,t))ds\right)K_1(x-B_t-y)\right]\wedge\omega_0(y)dy,
$$

and

$$
J_2\leq\int_{\mathbb{R}^d}\mathbb{E}\left[\left(\int_0^1 R_{b_s}(y,t)(N_b(y,t)-N_{\tilde{b}}(y,t))ds\right)K_2(x-B_t-y)\right]\wedge\omega_0(y)dy.
$$

Therefore, with the Fubini's theorem, we may change the order of the integrals to write

$$
\begin{aligned}
|J_1|&\leq\int_0^1\int_{\mathbb{R}^d}\mathbb{E}\left[R_{b_s}(y,t)\left|N_b(y,t)-N_{\tilde{b}}(y,t)\right|\left|K_1(x-B_t-y)\right|\right]\left|\omega_0(y)\right|dyds\\
&\leq C_0\int_0^1\int_{\mathbb{R}^d}\mathbb{E}\left[R_{b_s}(y,t)\frac{\mathbb{1}_{\{|x-B_t-y|<1\}}}{|x-B_t-y|^{\gamma_1}}|M_t|\right]\left|\omega_0(y)\right|dyds\\
&\quad+C_0C_Kt\left\|b-\tilde{b}\right\|_{L^\infty(\mathbb{R}^d,[0,t])}\int_0^1\int_{\mathbb{R}^d}\mathbb{E}\left[R_{b_s}(y,t)\frac{\mathbb{1}_{\{|x-B_t-y|<1\}}}{|x-B_t-y|^{\gamma_1}}\right]\left|\omega_0(y)\right|dyds\\
&=: J_{1,1}+J_{1,2},
\end{aligned}
$$

where the second inequality comes from the inequality (3.19).

To deal with $J_{1,1}$, choose and fix $\alpha,\beta>1$ such that $\alpha\gamma_1<d$ and $\alpha^{-1}+\beta^{-1}=1$. Then by construction,

$$
\begin{aligned}
R_{b_s}^{\alpha}(x,t)&=\exp\left(\alpha\int_0^t b_s(B_r+x,r)dB_r-\frac{1}{2}\alpha\int_0^t|b_s|^2(B_r+x,r)dr\right)\\
&=R_{\alpha b_s}\exp\left(\frac{\alpha(\alpha-1)}{2}\int_0^t|b_s|^2(B_r+x,r)dr\right)\\
&\leq e^{\frac{\alpha(\alpha-1)}{2}C_K^2t}R_{\alpha b_s}.
\end{aligned}\tag{3.24}
$$

Also, by the boundedness of the vector field $b, \tilde{b}$, the process M is a martingale with quadratic process

$$\langle M\rangle_t = \int_0^t |b-\tilde{b}|^2(B_r+x,r)dr \le t\left\|b-\tilde{b}\right\|^2_{L^\infty(\mathbb{R}^d,[0,t])}.$$

So by using Burkholder-Davis-Gundy inequality (see Proposition 2.19), there exists a constant $C_\beta>0$ such that for any $t\le T_K$,

$$\mathbb{E}\left[|M_t|^\beta\right] \le C_\beta \mathbb{E}\left[\langle M\rangle_t^{\frac{\beta}{2}}\right] \le C_\beta\left(\sqrt{t}\left\|b-\tilde{b}\right\|_{L^\infty(\mathbb{R}^d\times[0,t])}\right)^\beta. \tag{3.25}$$

Thus, by applying Hölder's inequality in $J_{1,1}$, we deduce that

$$\begin{aligned}
J_{1,1} &= C_0\int_0^1\int_{\mathbb{R}^d}\mathbb{E}\left[R_{b_s}(y,t)\frac{1_{\{|x-B_t-y|<1\}}}{|x-B_t-y|^{\gamma_1}}|M_t|\right]|\omega_0(y)|\,dyds\\
&\le C_0\int_0^1\sqrt[\alpha]{\int_{\mathbb{R}^d}\mathbb{E}\left[R^\alpha_{b_s}(y,t)\frac{1_{\{|x-B_t-y|<1\}}}{|x-B_t-y|^{\alpha\gamma_1}}\right]|\omega_0(y)|\,dy}\\
&\quad\times\sqrt[\beta]{\int_{\mathbb{R}^d}\mathbb{E}\left[|M_t|^\beta\right]|\omega_0(y)|\,dyds}\\
&\le C_0e^{\frac{\alpha(\alpha-1)}{2\alpha}C_K^2t}\int_0^1\sqrt[\alpha]{\int_{\mathbb{R}^d}\mathbb{E}\left[R_{\alpha b_s}(y,t)\frac{1_{\{|x-B_t-y|<1\}}}{|x-B_t-y|^{\alpha\gamma_1}}\right]|\omega_0(y)|\,dyds}\\
&\quad\times\sqrt[\beta]{C_1C_\beta}\sqrt{t}\left\|b-\tilde{b}\right\|_{L^\infty(\mathbb{R}^d,[0,t])}\\
&= C_0\sqrt[\beta]{C_1C_\beta}e^{\frac{\alpha(\alpha-1)}{2\alpha}C_K^2t}\sqrt{t}\left\|b-\tilde{b}\right\|_{L^\infty(\mathbb{R}^d,[0,t])}\\
&\quad\times\int_0^1\sqrt[\alpha]{\int_{\mathbb{R}^d}\int_{\{|z|<1\}}\frac{1}{|z|^{\alpha\gamma_1}}p_{\alpha b_s}(0,y,t,x-z)dz\,|\omega_0(y)|\,dyds}\\
&\le C_0\sqrt[\beta]{C_1C_\beta}e^{\frac{\alpha(\alpha-1)}{2\alpha}C_K^2t}\sqrt{t}\left\|b-\tilde{b}\right\|_{L^\infty(\mathbb{R}^d\times[0,t])}\\
&\quad\times\sqrt[\alpha]{\frac{\kappa_1C_\infty\left(1+\alpha C_K\sqrt{t}e^{\frac{\alpha^2C_K^2t}{2(q-1)}}\right)}{d-\alpha\gamma_1}}\\
&\le C_0\sqrt[\beta]{C_1C_\beta}e^{\frac{\alpha(\alpha-1)}{2\alpha}}\sqrt[\alpha]{\frac{\kappa_1C_\infty\left(1+\alpha e^{\frac{\alpha^2}{2(q-1)}}\right)}{d-\alpha\gamma_1}}\sqrt{t}\left\|b-\tilde{b}\right\|_{L^\infty(\mathbb{R}^d\times[0,t])},
\end{aligned}$$

for all $t \leq T_K$, where the second inequality follows from (3.24) and (3.25), the third inequality follows from Lemma 3.6, and the last inequality follows from the fact that $C_K t \leq 1$ for all $t \leq T_K$.

To deal with $J_{1,2}$, we use (3.17) to rewrite the expectation in terms of an integral so that

$$
\begin{aligned}
J_{1,2} &= C_0 C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d,[0,t])} \\
&\quad \times \int_0^1 \int_{\mathbb{R}^d} \mathbb{E}\left[R_{b_s}(y,t) \frac{1_{\{|x-B_t-y|<1\}}}{|x-B_t-y|^{\gamma_1}} \right] |\omega_0(y)|\, dyds \\
&= C_0 C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d,[0,t])} \\
&\quad \times \int_0^1 \left(\int_{\mathbb{R}^d} \int_{\{|z|<1\}} \frac{1}{|z|^{\gamma_1}} p_{b_s}(0,y,t,x-z) dz\, |\omega_0(y)|\, dy \right) ds \\
&\leq \frac{C_0 C_\infty C_K \kappa_1}{d-\gamma_1} \left(1 + C_K \sqrt{t} e^{\frac{C_K^2 t}{2(q-1)}} \right) t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d \times [0,t])} \\
&\leq \frac{C_0 C_\infty C_K \kappa_1}{d-\gamma_1} \left(1 + e^{\frac{1}{2(q-1)}} \right) t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d \times [0,t])}
\end{aligned}
$$

for any $t \leq T_K$, where the first inequality follows from the estimate in Lemma 3.6.

Now we treat with J_2. Since

$$
J_2 = \int_{\mathbb{R}^d} \mathbb{E}\left[\left(\int_0^1 R_{b_s}(y,t)(N_b(y,t) - N_{\tilde{b}}(y,t)) ds \right) K_2(x-B_t-y) \right] \wedge \omega_0(y) dy
$$

and $|K_2(z)| \leq C_0$, we may plug in (3.19) to write

$$
\begin{aligned}
|J_2| &\leq C_0 \int_0^1 \int_{\mathbb{R}^d} \mathbb{E}\left[R_{b_s}(y,t) \left(|M_t| + C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d \times [0,t])} \right) \right] |\omega_0(y)|\, dyds \\
&\leq C_0 \int_0^1 \int_{\mathbb{R}^d} \mathbb{E}\left[R_{b_s}(y,t) |M_t| \right] |\omega_0(y)|\, dyds \\
&\quad + C_0 C_1 C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d \times [0,t])} \\
&\leq C_0 \int_0^1 \int_{\mathbb{R}^d} \sqrt{\mathbb{E}\left[|M_t|^2 \right]} \sqrt{\mathbb{E}\left[R_{b_s}^2(y,t) \right]}\, |\omega_0(y)|\, dyds
\end{aligned}
$$

$$
\begin{aligned}
&+ C_0 C_1 C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d \times [0,t])} \\
&\leq C_0 \sqrt{t} \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d \times [0,t])} \int_0^1 \int_{\mathbb{R}^d} \sqrt{\mathbb{E}\left[e^{C_K^2 t} R_{\frac{1}{2} b_s} \right]} \left| \omega_0(y) \right| dy ds \\
&\quad + C_0 C_1 C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d \times [0,t])} \\
&\leq C_0 C_1 e^{\frac{1}{2}} \sqrt{t} \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d \times [0,t])} + C_0 C_1 C_K t \left\| b - \tilde{b} \right\|_{L^\infty(\mathbb{R}^d \times [0,t])},
\end{aligned}
$$

where the third inequality follows fromm Hölder's inequality with $\alpha = \beta = \frac{1}{2}$, the second last inequality comes from (3.24), (3.25) and the last inequality follows from $C_K t \leq 1$ for all $t \leq T_K$ and the fact that

$$
\mathbb{E}\left[R_b(x,t) \right] = 1
$$

for any Cameron-Martin density R_b. Putting these estimates for J_1 and J_2 together, we deduce (3.23) with a positive constant C_L which depends only on the structure constants C_0, C_1, C_∞,γ_1 and d (as α and q are constants depending only on d and γ_1). For example,

$$
\begin{aligned}
C_L = {} & C_0 C_1 e^{\frac{1}{2}} (1 + C_K) \\
& + C_0 C_\infty C_K \frac{1}{d - \gamma_1} \kappa_1 \left(1 + e^{\frac{1}{2(q-1)}} \right) \\
& + C_0 \sqrt[\beta]{C_1 C_\beta} e^{\frac{\alpha(\alpha-1)}{2\alpha}} \sqrt[\alpha]{\frac{C_\infty}{d - \alpha\gamma_1} \kappa_1 \left(1 + \alpha e^{\frac{\alpha^2}{2(q-1)}} \right)}
\end{aligned}
\tag{3.26}
$$

will do. □

We are now in a position to prove the main result about weak solutions to the distribution-dependent stochastic differential equation (3.11).

Theorem 3.11. *There exist two positive constants T_L and C_K depending on C_0, C_1, C_∞, $\gamma_1 \in [0,d)$ and d only such that the followings hold:*

(i) The (non-linear) mapping $b \mapsto K \diamond b$ is a contraction on the space of bounded time-dependent vector fields. More precisely,

$$\left\| K \diamond b - K \diamond \tilde{b} \right\|_{\infty} \leq \frac{1}{2} \left\| b - \tilde{b} \right\|_{\infty} \tag{3.27}$$

for any bounded vector fields b and $\tilde{b}$ such that $\|b\|_{\infty} \leq C_K$, $\|\tilde{b}\|_{\infty} \leq C_K$, where $\|b\|_{\infty} = \|b\|_{L^{\infty}(\mathbb{R}^d \times [0,T_L])}$. Hence, there is a unique vector field b such that $K \diamond b = b$.

(ii) There is a unique weak solution (X, B) on some probability space to the SDE (3.11) up to time T_L, where B is a Brownian motion and X satisfies (3.11), and

$$b^i(\cdot, t) = \sum_{j=1}^{d} \int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(X(y,t)) \right) \omega_0^j(y) dy$$

is bounded for $i = 1, \cdots, d$.

Proof. Choose $T_L = \frac{1}{4} C_L \wedge 1$ where C_L is given by (3.26). Then (3.27) follows immediately.

Since $b(x,t)$ is bounded and Borel measurable, Proposition 2.14 indicates that the Itô's stochastic differential equation

$$dX(x,t) = b(X(x,t),t)dt + dB_t$$

has a unique weak solution. Consequently, the second part follows from Lemma 3.3. □

We finish this section with a comment on the global solutions of (3.11). As long as K is a singular integral kernel, bounded at infinity, we have shown that there is a unique weak solution to (3.11) with initial data $\omega_0 \in L^1(\mathbb{R}^d) \cap L^{\infty}(\mathbb{R}^d)$ for the time duration $[0, T_L]$, where T_L depends only on the structure constants C_i

($i = 0, 1, \infty$) and $\gamma_1 \in [0, d)$. However, we are unable to conclude that the weak solution exists for all time t. The reason is that the SDE (3.11) does not define a dynamical system, which is not proposed as an initial value problem.

Finally, we should point out that we do not claim, although we strongly believe it is not the case, if ω_0 and K are regular enough, there are other fixed vector fields c in the sense that $K \diamond c = c$ but $c(x, t)$ is unbounded on some time interval $[0, T]$.

3.4 Strong solutions

With the same assumptions on K and ω_0 as in the earlier section, we aim to show that there is a strong solution to (3.11) by using the result in [138] for multi-dimensional diffusion processes with bounded drifts. Moreover, under a growth condition on K, we are able to show the Hölder continuity of the vector field $K \diamond b$.

Theorem 3.12. *Let $B = (B_t)_{t\geq 0}$ be a d-dimensional standard Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. There is a unique family of stochastic processes $X(x, t)$ which is jointly continuous in (x, t) almost surely, and satisfies the stochastic integral equations*

$$X^i(x,t) = x + \int_0^t \left[\sum_{j=1}^d \int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(X(y,s)) \right) (X(x,s)) \omega_0^j(y) dy \right] ds + B_t^i$$

for $t \in [0, T_L]$, and

$$b^i(\cdot, t) = \sum_{j=1}^d \int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(X(y,s)) \right) \omega_0^j(y) dy \tag{3.28}$$

where $i = 1, \ldots, d$, are bounded.

Proof. According to Theorem 3.11, for any $t \leq T_L$ (and extend it to be zero for $t > T_L$), there is a unique bounded vector field $b(x, t)$ satisfying $K \diamond b = b$. Since

b is bounded and Borel measurable, Lemma 2.11 indicates that there is a unique strong solution $X(x,t)$ to the ordinary stochastic differential equation

$$dX(x,t) = b(X(x,t),t)dt + dB_t, \quad X(x,0) = x.$$

Hence, by Lemma 3.3, $X(x,t)$ is the unique strong solution to (3.11). $\square$

We are going to show that the vector field (3.28) is in fact Hölder continuous. To this end, we recall Lemma 2.7 for the Hölder continuity result of the transition probability density function, proved originally by Nash [102] and later by Aronson [5], Fabes and Stroock [43] that

$$|p_b(s,y,t,x) - p_b(s,y,\tilde{t},\tilde{x})| \le \frac{C_H}{\delta^d} \left(\frac{|t-\tilde{t}|^{\frac{1}{2}} \vee |x-\tilde{x}|}{\delta} \right)^{\alpha} \tag{3.29}$$

for all $s \ge 0$, $\delta^2 \le t, \tilde{t} \le \frac{1}{\delta^2}$,$|x-\tilde{x}| \le \delta$, for any $\delta > 0$, where $C_H > 0$ and $\alpha \in (0,1)$ depending only $\|b\|_\infty$ and the dimension d.

Lemma 3.13. *Under the same assumptions for K and ω_0 as in the previous section, we further assume $\gamma_1 = \gamma_2 \equiv \gamma$, which belongs to $[0,d)$. Suppose $|b(x,t)| \le C_K$ for all $x \in \mathbb{R}^d$ and $t \le T_K$. Then $K \diamond b(x,t)$ is Hölder continuous on any compact subset of $\mathbb{R}^d \times (0,T_K]$, where the Hölder exponent and Hölder norm depend only on C_K and the dimension d.*

Proof. Let $\delta^2 \le t, \tilde{t} \le T_K$ and $|x - \tilde{x}| < \delta$ (for $\delta > 0$ small enough), and set

$$H = \left| K \diamond b(x,t) - K \diamond b(\tilde{x},\tilde{t}) \right|.$$

for simplicity. By the inequality (3.29), it yields

$$H = \left| \int_{\mathbb{R}^d} \left[\int_{\mathbb{R}^d} K(z) \wedge \omega_0(y) \left(p_b(0,y,t,x-z) - p_b(0,y,\tilde{t},\tilde{x}-z) \right) dz \right] dy \right|$$

$$
\begin{aligned}
&\le \int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d}|K(z)||\omega_0(y)||p_b(0,y,t,x-z)-p_b(0,y,\tilde{t},\tilde{x}-z)|dz\right]dy\\
&\le \int_{\mathbb{R}^d}\left[\int_{\{|z|<\rho\}}|K(z)||\omega_0(y)||p_b(0,y,t,x-z)-p_b(0,y,\tilde{t},\tilde{x}-z)|dz\right]dy\\
&\quad+\int_{\mathbb{R}^d}\left[\int_{\{|z|\ge\rho\}}|K(z)||\omega_0(y)||p_b(0,y,t,x-z)-p_b(0,y,\tilde{t},\tilde{x}-z)|dz\right]dy\\
&=: J_1+J_2.
\end{aligned}
$$

Since the transition probability function is Hölder continuous and (3.29) holds, then

$$
\begin{aligned}
J_1 &= \int_{\mathbb{R}^d}\left[\int_{\{|z|<\rho\}}|K(z)||\omega_0(y)||p_b(0,y,t,x-z)-p_b(0,y,\tilde{t},\tilde{x}-z)|dz\right]dy\\
&\le \frac{C_H}{\delta^d}\int_{\mathbb{R}^d}\left[\int_{\{|z|<\rho\}}|K(z)||\omega_0(y)|\left(\frac{|t-\tilde{t}|^{\frac{1}{2}}\vee|x-\tilde{x}|}{\delta}\right)^{\alpha}dz\right]dy\\
&= \frac{C_H}{\delta^d}\left(\frac{|t-\tilde{t}|^{\frac{1}{2}}\vee|x-\tilde{x}|}{\delta}\right)^{\alpha}\left(\int_{\mathbb{R}^d}|\omega_0(y)|dy\right)\left(\int_{\{|z|<\rho\}}|K(z)|dz\right)\\
&\le \frac{C_H}{\delta^d}\left(\frac{|t-\tilde{t}|^{\frac{1}{2}}\vee|x-\tilde{x}|}{\delta}\right)^{\alpha}\frac{C_1C_0}{d-\gamma}\rho^{d-\gamma}\mathrm{Vol}(S^{d-1}).
\end{aligned}
$$

And it is trivial to check

$$
\begin{aligned}
J_2 &= \int_{\mathbb{R}^d}\left[\int_{\{|z|\ge\rho\}}|K(z)||\omega_0(y)||p_b(0,y,t,x-z)-p_b(0,y,\tilde{t},\tilde{x}-z)|dz\right]dy\\
&\le \frac{C_0}{\rho^\gamma}\int_{\mathbb{R}^d}\left[\int_{\{|z|\ge\rho\}}|\omega_0(y)||p_b(0,y,t,x-z)-p_b(0,y,\tilde{t},\tilde{x}-z)|dz\right]dy\\
&\le \frac{C_0}{\rho^\gamma}\int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d}|\omega_0(y)|p_b(0,y,t,z)dz\right]dy+\frac{C_0}{\rho^\gamma}\int_{\mathbb{R}^d}\left[\int_{\mathbb{R}^d}|\omega_0(y)|p_b(0,y,\tilde{t},z)dz\right]dy\\
&= 2\frac{C_0C_1}{\rho^\gamma}
\end{aligned}
$$

for any $\rho>0$. So putting J_1, J_2 together to obtain

$$
H \le C_1\left\{\frac{C_H}{\delta^d}\left(\frac{|t-\tilde{t}|^{\frac{1}{2}}\vee|x-\tilde{x}|}{\delta}\right)^{\alpha}\right\}\frac{C_0}{d-\gamma}\rho^{d-\gamma}\mathrm{Vol}(S^{d-1})+2\frac{C_0C_1}{\rho^\gamma},
$$

where $\mathrm{Vol}(S^{d-1})$ is the volume of the unit $(d-1)$-sphere. Choose $\rho > 0$ such that

$$C_1 \left\{ \frac{C_H}{\delta^d} \left(\frac{|t-\tilde{t}|^{\frac{1}{2}} \vee |x-\tilde{x}|}{\delta} \right)^{\alpha} \right\} \frac{C_0}{d-\gamma} \rho^{d-\gamma} \mathrm{Vol}(S^{d-1}) = 2 \frac{C_0 C_1}{\rho^\gamma}$$

that is

$$\frac{1}{\rho} = \left\{ \frac{C_H \mathrm{Vol}(S^{d-1})}{2(d-\gamma)\delta^d} \left(\frac{|t-\tilde{t}|^{\frac{1}{2}} \vee |x-\tilde{x}|}{\delta} \right)^{\alpha} \right\}^{1/d}.$$

Then

$$H \le 4 C_0 C_1 \left\{ \frac{C_H \mathrm{Vol}(S^{d-1})}{2(d-\gamma)\delta^d} \left(\frac{|t-\tilde{t}|^{\frac{1}{2}} \vee |x-\tilde{x}|}{\delta} \right)^{\alpha} \right\}^{\gamma/d},$$

which proofs the claim. □

Corollary 3.14. *Under the same conditions for ω_0 as in the previous section. Suppose the kernel $K = (K^{ij})$ satisfies the following condition:*

$$|K(x)| \le C_0 \frac{1}{|x|^\gamma} \quad \textit{for all } x \neq 0$$

where $0 \le \gamma < d$ and $C_0 > 0$. Then there is a unique strong solution $X(x,t)$ to (3.11) *for any $t \le T_L$, such that*

$$b^i(x,t) = \sum_{j=1}^{d} \int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(X(y,s)) \right) \omega_0^j(y) dy$$

$i = 1, \cdots, d$, are bounded and Hölder continuous on any compact subset of $\mathbb{R}^d \times (0, T_L]$.

3.5 From SDE to PDE

In this section, we recover the partial differential equation from the distribution-dependent stochastic differential equation

$$dX^i(x,t) = \left[\sum_{j=1}^{d}\int_{\mathbb{R}^d}\left(K^{ij}\star\mathscr{L}(X(y,t))\right)(X(x,t))\omega_0^j(y)dy\right]dt + dB_t^i \tag{3.30}$$

Theorem 3.15. *Let K and ω_0 satisfy the assumptions in Section 3. For any $x\in\mathbb{R}^d$ and $t\geq 0$, let $(X(x,t),B_t)$ be the unique weak solution of SDE* (3.30) *on a probability space $(\Omega,\mathscr{F},\mathbb{P})$ for $t\in[0,T_L]$. Then for any $y\in\mathbb{R}^d$ and $t>0$, the distribution of $X(y,t)$ has a positive and continuous density denoted by $p(0,y,t,\cdot)$. Let $b(x,t)$ be defined by*

$$b^i(x,t) = \sum_{j=1}^{d}\int_{\mathbb{R}^d}\left(K^{ij}\star\mathscr{L}(X(y,s))\right)\omega_0^j(y)dy$$

for $i=1,\dots,d$, and

$$\omega^i(x,t) = \int_{\mathbb{R}^d}p(0,y,t,x)\omega_0^i(y)dy \tag{3.31}$$

for any $x\in\mathbb{R}^d$ and $t\in[0,T_L]$. Then the pair (b,ω) is the solution to the following non-local partial differential equation

$$\frac{\partial}{\partial t}\omega^i + \left(\sum_{j=1}^{d}b^j\frac{\partial}{\partial x^j}\right)\omega^i = \frac{1}{2}\Delta\omega^i - \left(\sum_{j=1}^{d}\frac{\partial b^j}{\partial x^j}\right)\omega^i, \quad \omega^i(\cdot,0)=\omega_0^i$$

and

$$b^i(x,t) = \sum_{j=1}^{d}\int_{\mathbb{R}^d}K^{ij}(x-y)\omega^j(y,t)dy$$

for any x any $t\in[0,T_L]$, where $i=1,\dots,d$.

Proof. According to our construction, $b(x,t)$ is the unique bounded vector field such that $K \diamond b = b$, and $X(x,t)$ is the unique weak solution of the SDE

$$dX(x,t) = b\left(X(x,t),t\right)dt + dB_t, \quad X(x,0) = x.$$

Then $p(0,y,t,x) = p_b(0,y,t,x)$ is the transition probability density for the diffusion with its generator L_b. Thus, by considering $p_b(0,y,t,x)$ as a function of (t,x), p_b is the fundamental solution to the forward adjoint equation

$$\left(\frac{\partial}{\partial t} - L_b^\star\right) p_b = 0$$

where $L_b^\star = \frac{1}{2}\Delta - b\cdot\nabla - \nabla\cdot b$. Hence, according to Lemma 3.1, $\omega(x,t)$ given by (3.31) is the solution to

$$\left(\frac{\partial}{\partial t} - L_b^\star\right)\omega = 0, \quad \omega(\cdot,0) = \omega_0,$$

that is

$$\frac{\partial}{\partial t}\omega + b\cdot\nabla\omega = \frac{1}{2}\Delta\omega - (\nabla\cdot b)\,\omega, \quad \omega(\cdot,0) = \omega_0.$$

Moreover,

$$\begin{aligned}
b^i(x,t) &= \sum_{j=1}^d \int_{\mathbb{R}^d} \left(K^{ij} \star \mathscr{L}(X(y,s))\right)\omega_0^j(y)dy \\
&= \sum_{j=1}^d \int_{\mathbb{R}^d} \left(\int K^{ij}\left(x-z\right)p_b(0,y,t,z)dz\right)\omega_0^j(y)dy \\
&= \sum_{j=1}^d \int_{\mathbb{R}^d} K^{ij}(x-z)\omega^j(z,t)dz
\end{aligned}$$

which completes the proof. □

As an example, we may apply this representation theorem to the Biot–Savart kernel $G(x) = -\frac{1}{4\pi}\frac{x}{|x|^3}$ on $\mathbb{R}^3$ so that $K^{ij} = \varepsilon^{ikj}G^k$. Then there is a unique weak solution to the following SDE

$$dX(x,t) = \left(\int_{\mathbb{R}^3} \mathbb{E}\left[G(z - X(y,t))\right]\big|_{z=X(x,t)} \wedge \omega_0(y)dy \right) dt + \sqrt{2\nu} dB_t, \qquad (3.32)$$

where $\omega_0 \in L^1(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$ is the initial vorticity. In this case, we define

$$\omega(x,t) = \int_{\mathbb{R}^3} p(0,y,t,x)\omega_0(y)dy,$$

where $p(0,y,t,\cdot)$ is the probability density function of the law of $X(y,t)$ to (3.32), and define

$$\begin{aligned} u^i(x,t) &= \int_{\mathbb{R}^3} K^{ij}(x-z)\omega^j(z,t)dz \\ &= -\frac{1}{4\pi}\int_{\mathbb{R}^3} \varepsilon^{ikj} \frac{x^k - z^k}{|x-z|^3} \omega^j(z,t)dz. \end{aligned}$$

That is,

$$u(x,t) = -\frac{1}{4\pi}\int_{\mathbb{R}^3} \frac{x-z}{|x-z|^3} \wedge \omega(z,t)dz.$$

By the previous theorem, (u,ω) satisfies the following PDE

$$\frac{\partial}{\partial t}\omega + u\cdot\nabla\omega = \nu\Delta\omega - (\nabla\cdot u)\,\omega, \quad \omega(\cdot,0) = \omega_0.$$

Moreover, by a direct computation,

$$\frac{\partial}{\partial x_i} u^i(x,t) = \frac{3}{4\pi}\int_{\mathbb{R}^3} \varepsilon^{ikj} \frac{(z_k - x_k)(z_i - x_i)}{|z-x|^5} \omega^l(z,t)dz,$$

so that

$$\nabla\cdot u(x,t) = 0, \quad \text{and} \quad \Delta u(x,t) = -\nabla\wedge\omega(x,t),$$

where the second equation follows from the Green formula which can be also written as

$$\nabla \wedge (\nabla \wedge u - \omega) = 0.$$

Therefore, it yields that

$$\omega = \nabla \wedge u - \nabla f$$

for some scalar function f. If one also imposes a constraint that $\nabla \cdot \omega_0 = 0$, then $\omega = \nabla \wedge u$. In this case,

$$\frac{\partial}{\partial t}\omega + u \cdot \nabla \omega = \nu \Delta \omega, \quad \omega(\cdot, 0) = \omega_0$$

and

$$\omega = \nabla \wedge u.$$

Hence, $(X(x,t), B_t)$ is the probability representation of the solution to the above vorticity equation.

3.6 Numerical approximation

This section aims to refine the numerical approximation for the random vortex method developed by Long [90]. Then we will simulate the velocity field to compare the two methods.

3.6.1 Approximation scheme

In [90], Long proposed a numerical approximation scheme for the two-dimensional Navier–Stokes equation and proved the effectiveness of the random vortex method by computing the rate of convergence explicitly. Specifically, in his argument, particle trajectory paths $X(x,t)$ are assumed to be solutions to the following diffusion process

$$dX(x,t) = u(x,t)dt + \sqrt{2\nu}dB_t, \quad X(x,0) = x,$$

where $u(x,t)$ is the velocity field for the Navier–Stokes equation, and $B = (B_t)_{t\geq 0}$ is a standard 2-dimensional Brownian motion on some probability space. Then with a similar argument as in the archetypical example, this process $X(x,t)$ is the solution to the following stochastic differential equation

$$dX(x,t) = \left(\int_{\mathbb{R}^2} \mathbb{E}\left[G\left(X(x,t) - X(y,t)\right)\right]\omega_0(y)dy\right)dt + \sqrt{2\nu}dB_t, \quad X(x,0) = x, \tag{3.33}$$

where G is the Biot–Saward kernel $\frac{1}{2\pi}(-\frac{x_2}{|x|^2}, \frac{x_1}{|x|^2})$, and ω_0 is the initial vorticity, which is assumed to be continuous and bounded with compact support. Then the velocity field is modelled by

$$u(x,t) = \int_{\mathbb{R}^2} \mathbb{E}\left[G\left(x - X(y,t)\right)\right]\omega_0(y)dy, \tag{3.34}$$

We point out that in the two-dimensional problem, the vorticity equation is a scalar function defined as $\omega = \frac{\partial u^2}{\partial x_1} - \frac{\partial u^1}{\partial x_2}$.

In [90], Long worked our an approximation for the process $X(x,t)$, and the velocity field followed accordingly. Specifically, trajectory paths are simulated for each starting position $x_{\boldsymbol{i}}$ on the lattice point. For the drift function, the integral in (3.33) is replaced by a discrete sum, and the expectation is dropped. In this way, each discretized path is read as the following

$$d\widetilde{X}_t^{\boldsymbol{i}} = \left(\sum_{\boldsymbol{j}\neq\boldsymbol{i}} G_\delta\left(\widetilde{X}_t^{\boldsymbol{i}} - \widetilde{X}_t^{\boldsymbol{j}}\right)\omega_0(x_{\boldsymbol{j}})h^2\right)dt + \sqrt{2\nu}dB_t^{\boldsymbol{i}}, \quad \widetilde{X}_0^{\boldsymbol{i}} = x_{\boldsymbol{i}}, \tag{3.35}$$

where $(B_t^{\boldsymbol{i}})_{t\geq 0}$ are independent copies of Brownian motions for each $\boldsymbol{i}$, G_δ is the

Biot–Savart kernel convolved with a cutoff function $\psi_\delta(x) = \frac{1}{\delta^2}\psi(\frac{1}{\delta}x)$ for any $\delta > 0$, and $x_{\boldsymbol{i}}$ is a lattice point on the lattice space with spacing $h > 0$, i.e. $x_{\boldsymbol{i}} \in \Lambda_2^h$, where

$$\Lambda_2^h = \{x_{\boldsymbol{i}} : x_{\boldsymbol{i}} = h \cdot \boldsymbol{i} = h \cdot (i_1, i_2), \boldsymbol{i} \in \mathbb{Z}^2\}.$$

Since the coefficient is globally Lipschitz, (3.35) admits a unique strong solution $X_t^{\boldsymbol{i}}$ defined for all $t \geq 0$ (see [130] for example). Then the discrete velocity field is defined as

$$\tilde{u}(x,t) = \sum_{\boldsymbol{j}} G_\delta\left(x - \widetilde{X}_t^{\boldsymbol{j}}\right)\omega_0(x_{\boldsymbol{j}})h^2. \tag{3.36}$$

It is obvious that the discrete velocity field in (3.36) is a random process. Long indicated that except for an event of small probability, this discrete velocity converges to the velocity field $u(x,t)$ defined in (3.34) uniformly in the discrete L^p space, and the rate of convergence depends on the size of h and the choice of the cutoff function.

We make a remark that Long assumed in [90] that the SDE (3.33) has a solution without explicitly working out its solvability. We have filled this gap in earlier discussion by showing it indeed has a unique strong solution in a small time interval. Particularly, the solvability of this type of McKean–Vlasov SDEs holds in higher dimensions with general kernel functions. In fact, one can extend the approximation scheme developed in [90] to approximate the McKean–Vlasov equation (3.11) in higher dimensions. With an explicit computation, this discretization method may yield a different of the rate of convergence. Specifically, if we define the space of lattice points as

$$\Lambda_d^h = \{x_{\boldsymbol{i}} : x_{\boldsymbol{i}} = h \cdot \boldsymbol{i} = h \cdot (i_1, ..., i_d), \boldsymbol{i} \in \mathbb{Z}^d\},$$

the discretized stochastic differential equation $\widetilde{X}^{\boldsymbol{i}}$ is read as

$$d\widetilde{X}_t^{\boldsymbol{i}} = \left(\sum_{\boldsymbol{j} \neq \boldsymbol{i}} K_\delta \left(\widetilde{X}_t^{\boldsymbol{i}} - \widetilde{X}_t^{\boldsymbol{j}} \right) \wedge \omega_0(x_{\boldsymbol{j}}) h^d \right) dt + \sqrt{2\nu} dB_t^{\boldsymbol{i}}, \quad \widetilde{X}_0^{\boldsymbol{i}} = x_{\boldsymbol{i}},$$

where K_δ is the singular kernel convolved with the cutoff function ψ_δ for any $\delta > 0$, and $B^{\boldsymbol{i}}$ are independent copies of Brownian motions for each starting point, and ω_0 is the given initial vorticity, which is assumed to be bounded in $L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$. But still, the problem is that the simulated result might blow up, especially when the viscosity $\nu > 0$ is not small.

In order to reduce this stochastic noise, we modified this approach by splitting each particle into N copies of the same particle, and run independent copies of Brownian motions for each particle. Hence, we propose the following stochastic differential equations for each particle path

$$\begin{cases} d\widetilde{X}_n^{\boldsymbol{i}}(t) = \frac{1}{N} \sum_{m=1}^N \left(\sum_{\boldsymbol{j} \neq \boldsymbol{i}} K_\delta \left(\widetilde{X}_n^{\boldsymbol{i}}(t) - \widetilde{X}_m^{\boldsymbol{j}}(t) \right) \wedge \omega_0(x_{\boldsymbol{j}}) h^d \right) dt + \sqrt{2\nu} dB_n^{\boldsymbol{i}}(t), \\ \widetilde{X}_n^{\boldsymbol{i}}(0) = x_{\boldsymbol{i}}, \end{cases}$$

where $B_n^{\boldsymbol{i}}$ are independent copies of d-dimensional of Brownian motions for each starting point $x_{\boldsymbol{i}}$ and each path n. Again, the drift function is Lipschitz so that this SDE has a unique strong solution. Then for each particle starting at the position $x_{\boldsymbol{i}}$, its trajectory path is modelled by

$$\widetilde{X}^{\boldsymbol{i},N}(t) = \frac{1}{N} \sum_{n=1}^N \widetilde{X}_n^{\boldsymbol{i}}(t)$$

with the initial data $\widetilde{X}^{\boldsymbol{i},N}(0) = x_{\boldsymbol{i}}$.

To simplify the notation, we write

$$\tilde{u}_n(x,t) := \sum_{\boldsymbol{j}\neq \boldsymbol{i}} K_\delta\left(x - \widetilde{X}_n^{\boldsymbol{j}}(t)\right) \wedge \omega_0(x_{\boldsymbol{j}})h^d.$$

Then the particle path with the same starting lattice points is modelled by the following stochastic differential equation

$$d\widetilde{X}^{\boldsymbol{i},N}(t) = \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{N}\sum_{m=1}^{N}\tilde{u}_m(\widetilde{X}_n^{\boldsymbol{i}}(t),t)dt + \sqrt{2\nu}dB_n^{\boldsymbol{i}}(t)\right), \tag{3.37}$$

so that the velocity field is approximated by

$$\tilde{u}^N(x,t) = \sum_{\boldsymbol{j}\neq \boldsymbol{i}} K_\delta\left(x - \widetilde{X}^{\boldsymbol{i},N}(t)\right) \wedge \omega_0(x_{\boldsymbol{j}})h^d.$$

When $N = 1$, this is exactly the same as Long's scheme.

Here we have used independent copies of Brownian motions for each particle path at different locations. So by the strong law of large numbers, the martingale part in (3.37) vanishes almost surely as $N \to \infty$. On this account, the stochastic error is reduced for the approximated velocity field $\tilde{u}^N$.

3.6.2 Simulation for 3D case

With the numerical scheme established in earlier subsection, we show some figures of the approximated velocity field for 3D flows and compare these two schemes. For the simulation, we choose the kernel K to be the 3D Biot–Savart kernel, which is defined as $K(x) = -\frac{1}{4\pi}\frac{x}{|x|^3}$. For illustration purpose, we choose the spacing $h = 0.5$ and set $\nu = 1$.

Before carrying out the simulation, we need the initial vorticity ω_0. In literature, there are several methods for approximating the vorticity equation, and one may refer to [30, 84, 99] and literature therein for a detailed discussion. Here we emphasize the point vortex method originated from Rosenhead [121], where the vorticity field $\omega(x,t)$ is represented by

$$\omega(x,t) = \sum_{i=1}^{N} \Gamma_i \delta(x - x_i(t)),$$

where δ is the Dirac delta function, $x_i(t)$ is the location of the vortex at time t, and Γ_i is the corresponding circulation for each $i = 1, ..., N$. With an analogy as this point vortex method, we construct the initial vorticity as a sum of Dirac delta functions. The locations of the vortices and the corresponding vorticity are listed in Table 1, otherwise the vorticity is null.

x	$\omega_0(x)$
$(1,1,0)$ $(0,1,1)$ $(1,-1,0)$ $(-1,-1,0)$	$(0,0,10)$
$(0.5,1,-1)$ $(1.5,0,0)$ $(-1,0.5,0)$ $(0.5,-1.5,0.5)$	$(10,20,0)$

Table 1: Initial vorticity $\omega_0(x)$ at different positions

With these setups, we may simulate trajectory paths of the particles for each given initial lattice point $x_{\boldsymbol{i}}$. Ideally, $\boldsymbol{i}$ is taken on the whole space of $\mathbb{Z}^3$ so that the integration over $\mathbb{R}^3$ could be approximated by the summation over Λ_3^h. Indeed, the rate of convergence in [90] is based on this assumption. However, this is impossible in the simulation. We set $M > 0$ to be the mesh size instead so that the lattice points are taken on the set

$$\Lambda_3^{h,M} := \{x_{\boldsymbol{i}} : x_{\boldsymbol{i}} = h \cdot \boldsymbol{i} = h \cdot (i_1, i_2, i_3), |i_1|, |i_2|, |i_3| \leq M\}.$$

Then the computational complexity is proportional to M^6. Favorably, under the point vortex assumption, this complexity can be reduced to M^3.

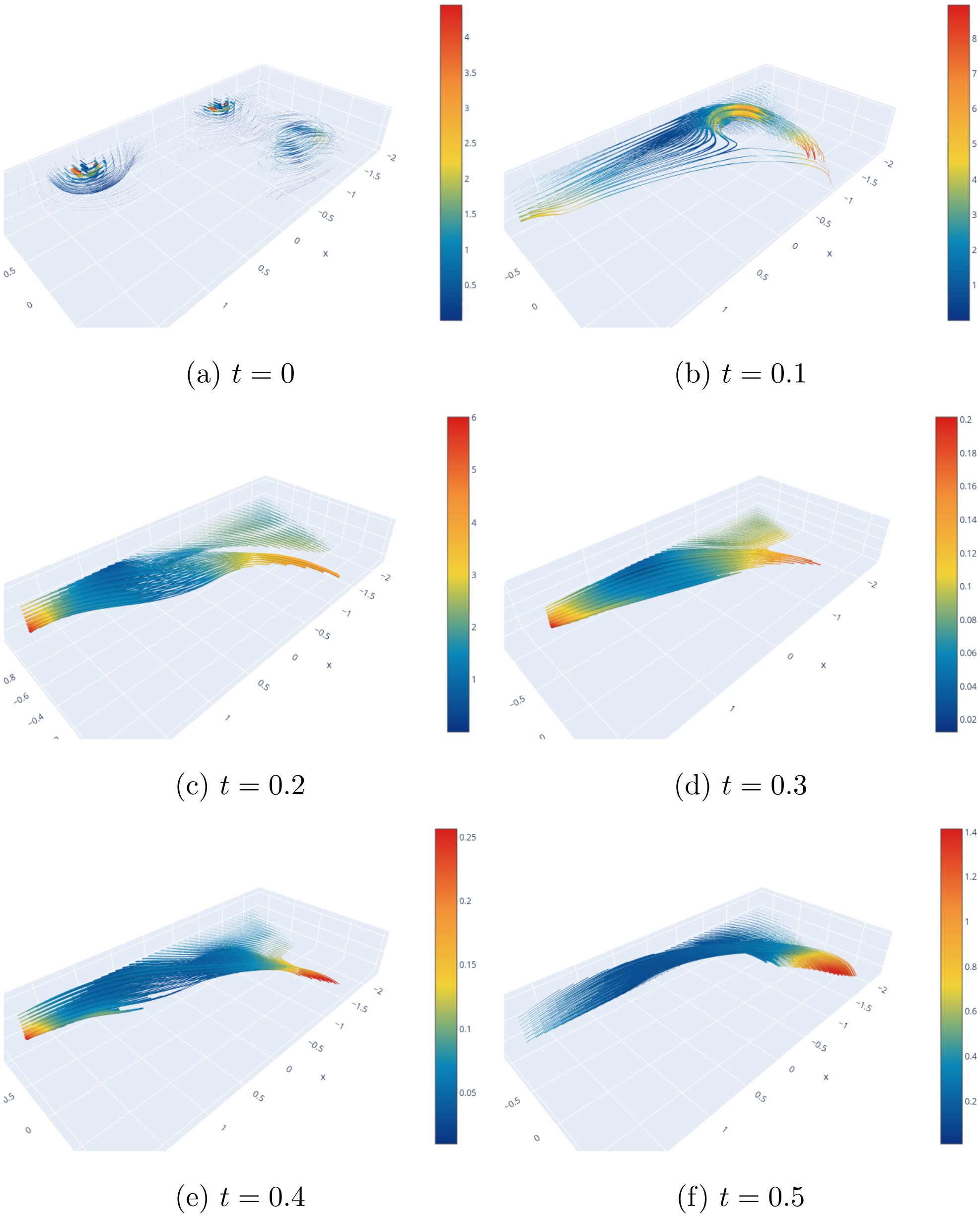
In the modified scheme, we have repeated N paths for each initial lattice point. By allowing this repetition, this scheme becomes N^2-times more complex than without repeating. Table 2 shows the computation time for simulating trajectory paths under different M and N values.

	$M = 10$	$M = 20$
$N = 1$	72s	512s
$N = 10$	7326s	57044s

Table 2: Computation time for simulating trajectory paths

Now we are ready to plot the velocity field for different schemes under different lattice sets. Figure 1 – Figure 4 plot the velocity field on the space $[-2, 2] \times [-1, 1] \times [-0.5, 0.5]$ under different M and N values. In these plots, directions of the velocity field are illustrated by the stream tubes, whereas their magnitude are marked by different colours. Each figure visualizes how velocities evolve along with time in the 3D space.

If we compare Figure 1 and Figure 2, it is obvious to see that the velocity fields have been refined for each time step. For example, at time $t = 0.3$, Figure 1 only displays a rough direction of the velocity field whereas Figure 2 reveals the vortices behind. This observation can also be justified from other time steps in these two figures. However, under the modified approach shown in Figure 3 and Figure 4, we cannot see a big refinement in terms of underlying vortices. So we conclude that without repetition of the trajectory paths, an increment of the lattice size will significantly boost the accuracy of the approximation for the velocity field. However, in the modified approach with $N = 10$, the simulated velocity field tends to stabilize. Therefore, an increase of lattice size will not make a significant difference. Nevertheless, restricted by the hardware of the computer, we cannot test if this still holds true when the lattice size M has been increased

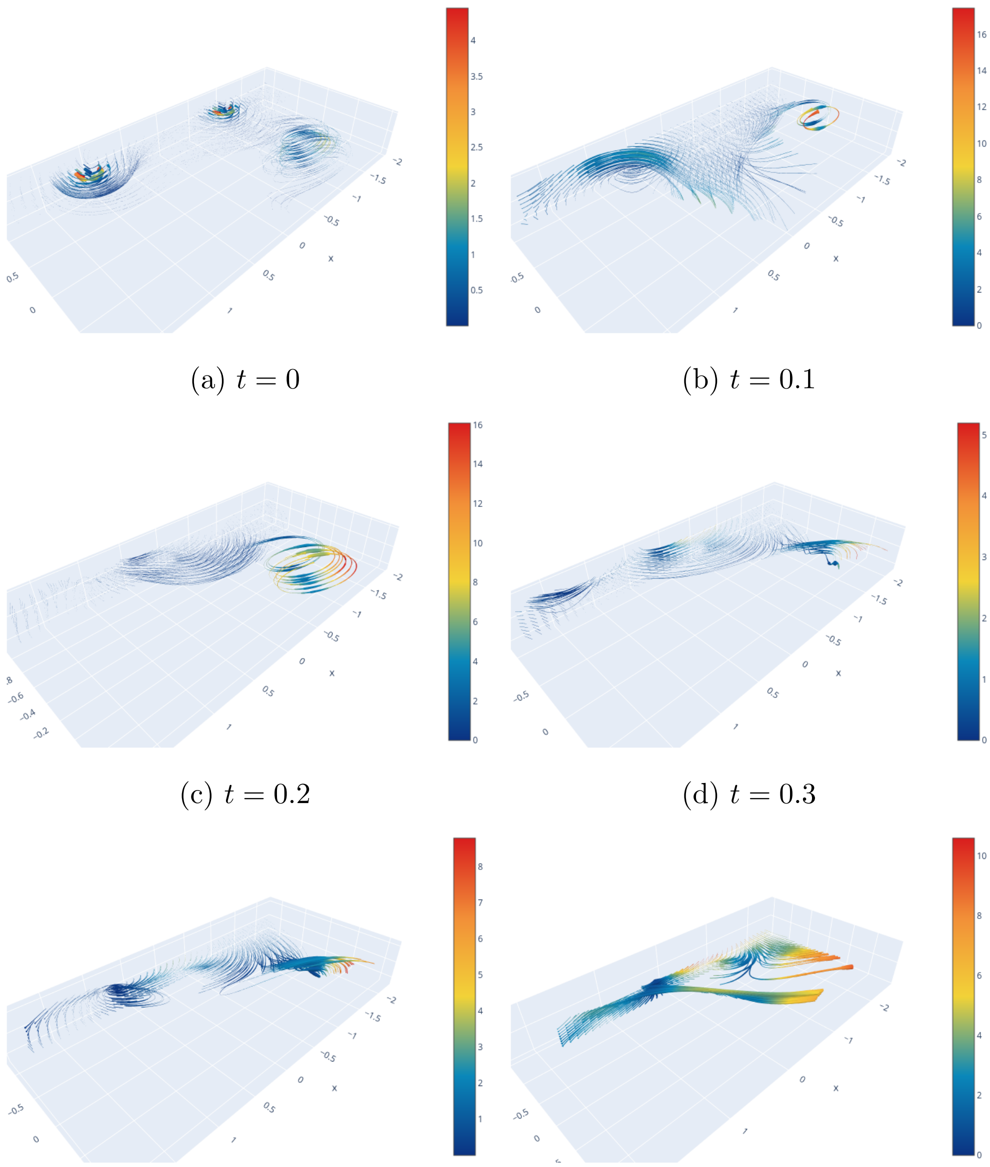


(a) $t = 0$ (b) $t = 0.1$

(c) $t = 0.2$ (d) $t = 0.3$

(e) $t = 0.4$ (f) $t = 0.5$

Figure 1: Plot of velocity field on 3D space, $N = 1, M = 10$.

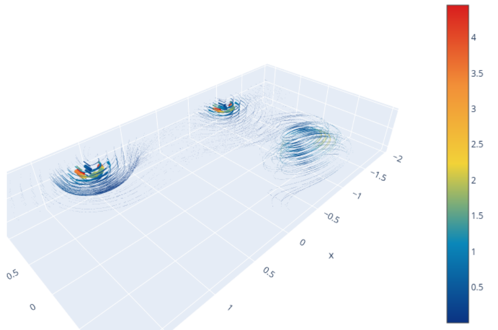


(a) $t = 0$ (b) $t = 0.1$

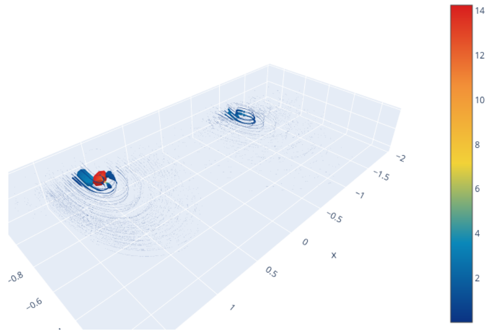
(c) $t = 0.2$ (d) $t = 0.3$

(e) $t = 0.4$ (f) $t = 0.5$

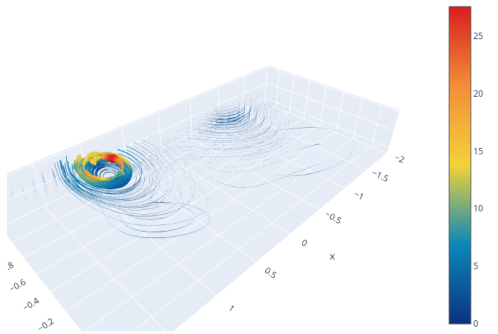
Figure 2: Plot of velocity field on 3D space, $N = 1, M = 20$.



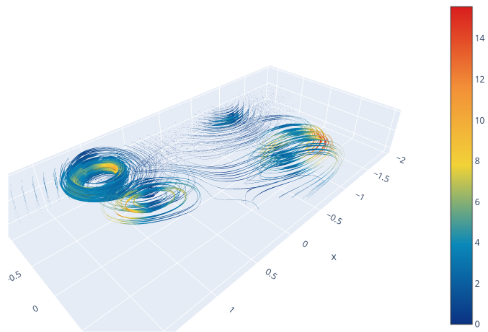
(a) $t = 0$



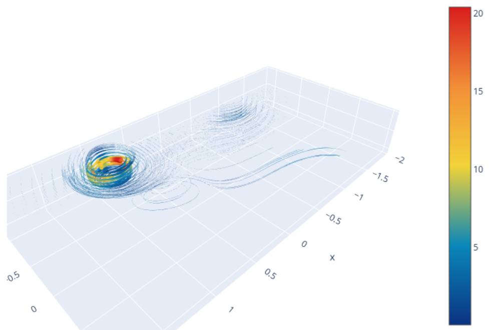
(b) $t = 0.1$



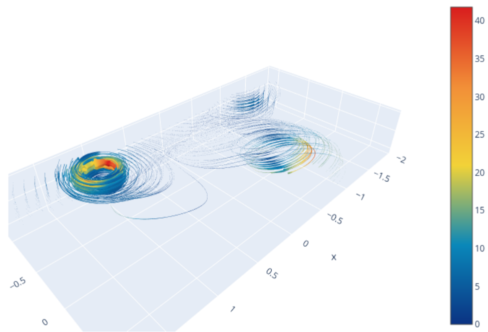
(c) $t = 0.2$



(d) $t = 0.3$

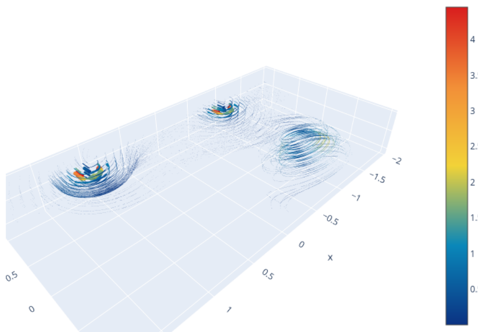


(e) $t = 0.4$

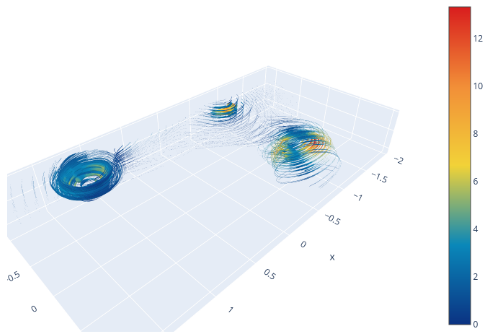


(f) $t = 0.5$

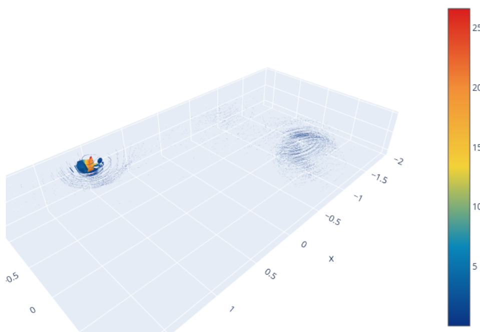
Figure 3: Plot of velocity field on 3D space, $N = 10, M = 10$.



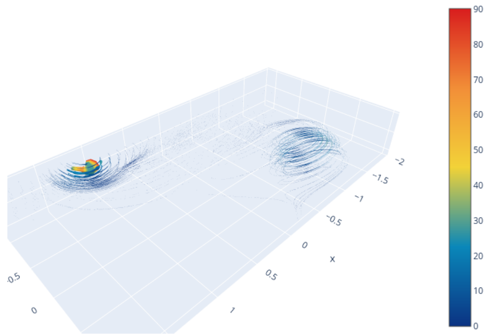
(a) $t = 0$



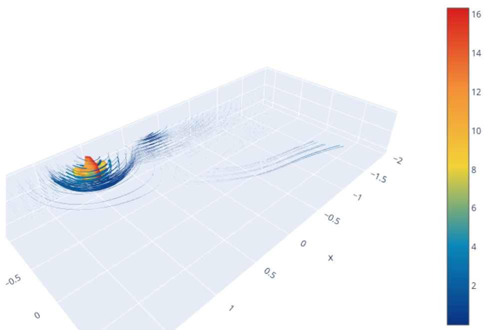
(b) $t = 0.1$



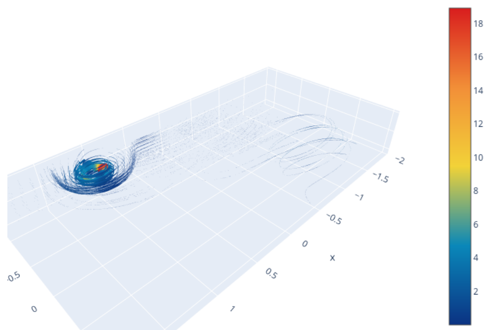
(c) $t = 0.2$



(d) $t = 0.3$

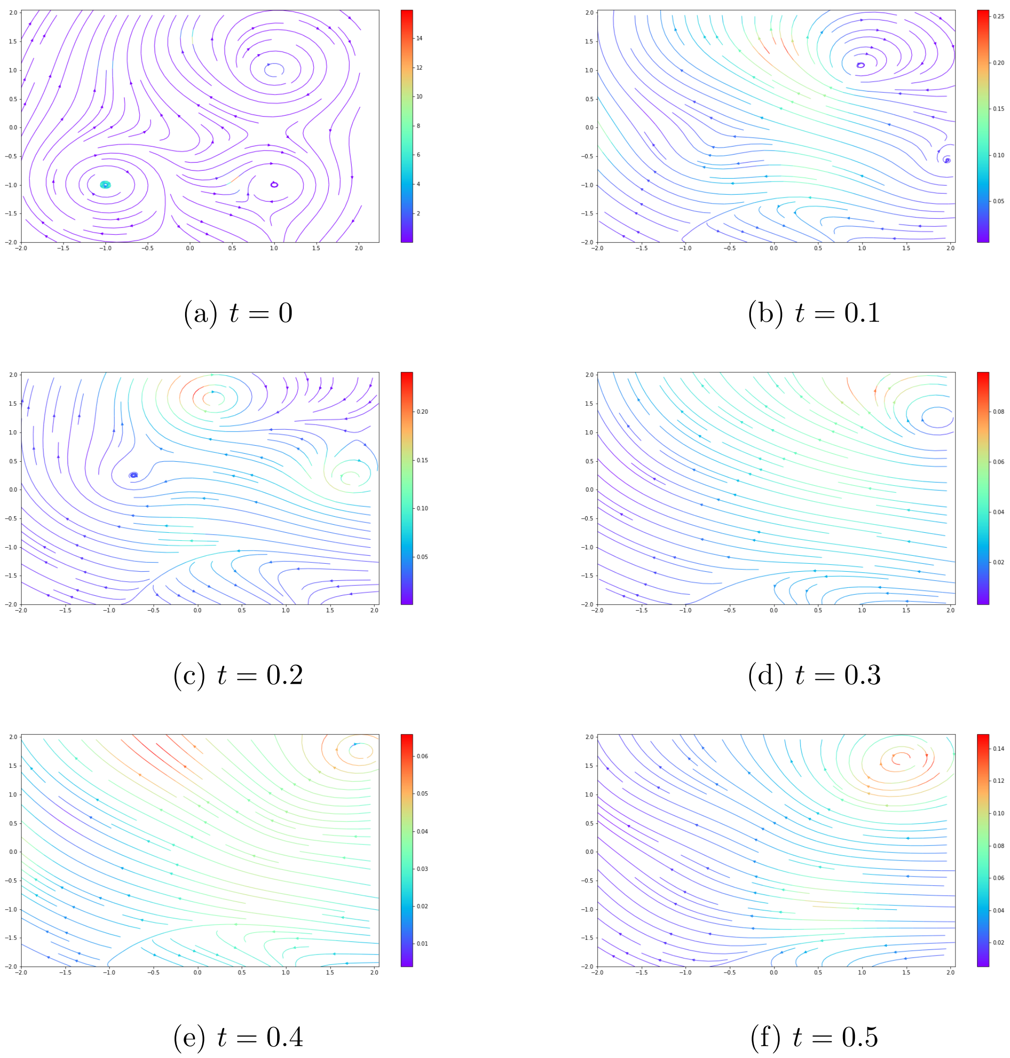


(e) $t = 0.4$



(f) $t = 0.5$

Figure 4: Plot of velocity field on 3D space, $N = 10, M = 20$.



(a) $t = 0$

(b) $t = 0.1$

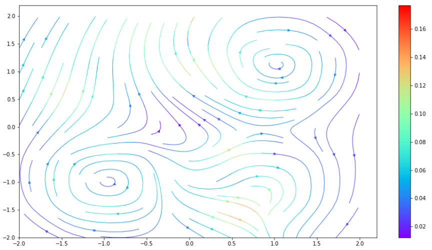
(c) $t = 0.2$

(d) $t = 0.3$

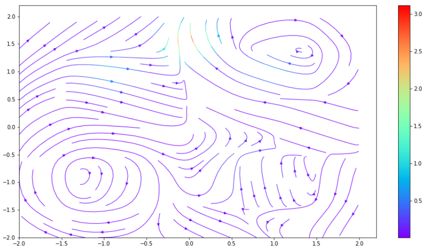
(e) $t = 0.4$

(f) $t = 0.5$

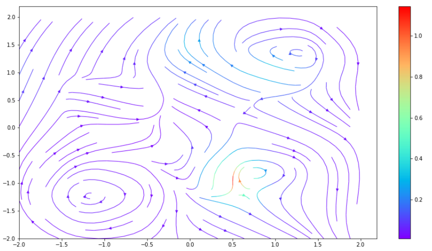
Figure 5: Projected velocity field on the plane $z = 0$, $N = 1,, M = 10$.



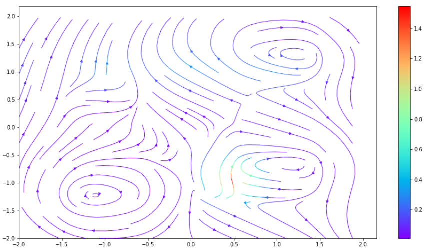
(a) $t = 0$



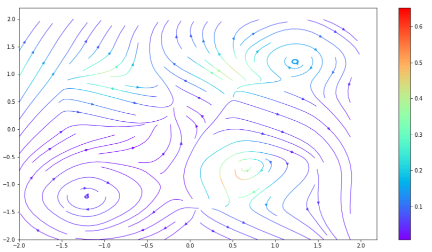
(b) $t = 0.1$



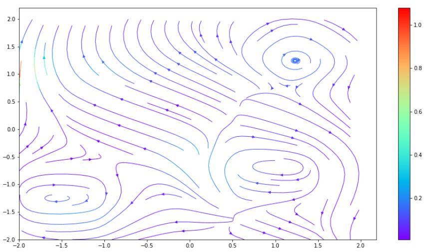
(c) $t = 0.2$



(d) $t = 0.3$

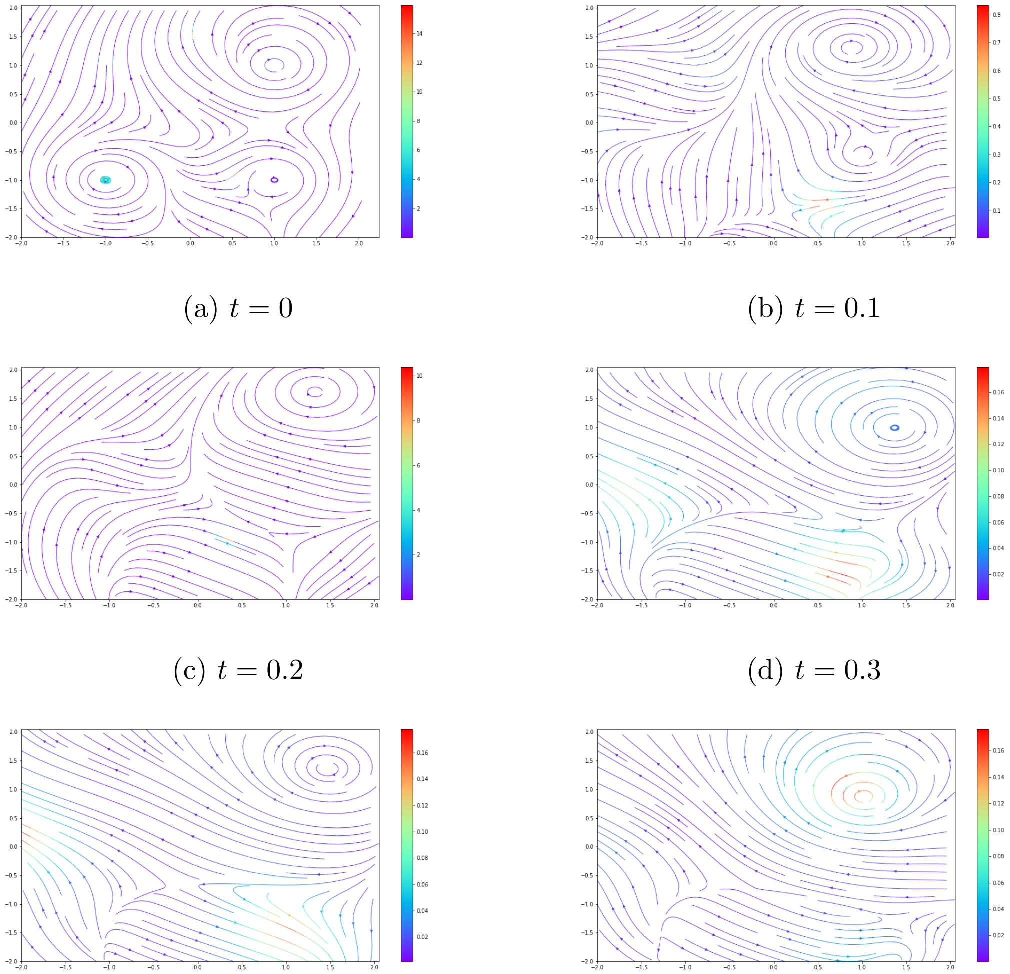


(e) $t = 0.4$



(f) $t = 0.5$

Figure 6: Projected velocity field on the plane $z = 0$, $N = 10, M = 10$.



(a) $t = 0$

(b) $t = 0.1$

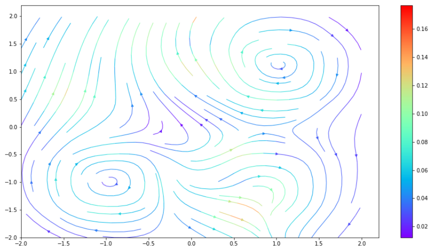
(c) $t = 0.2$

(d) $t = 0.3$

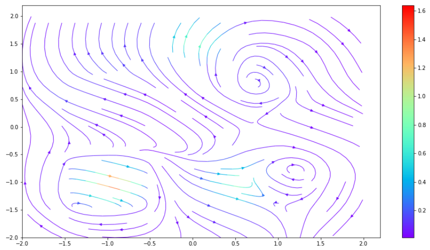
(e) $t = 0.4$

(f) $t = 0.5$

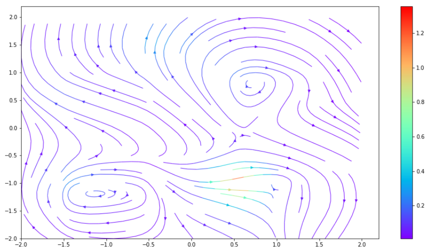
Figure 7: Projected velocity field on the plane $z = 0$, $N = 1, M = 20$.



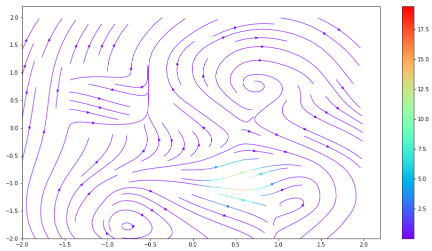
(a) $t = 0$



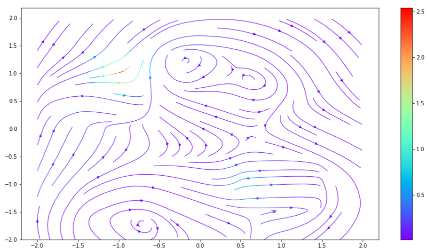
(b) $t = 0.1$



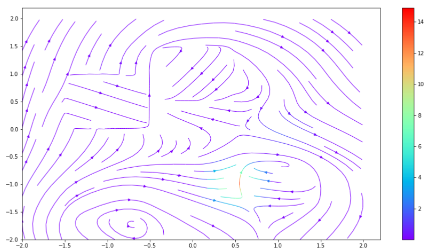
(c) $t = 0.2$



(d) $t = 0.3$



(e) $t = 0.4$



(f) $t = 0.5$

Figure 8: Projected velocity field on the plane $z = 0$, $N = 10, M = 20$.

to a larger scale, say $M = 10000$. Without any doubt, $M = 10000$ will give a better approximated result even with repetitions. But we believe the size of M will not matter as much up to some point.

If we compare Figure 2 and Figure 4, we see a difference in the shape of the velocity field. This is due to the fact that only stream tubes that passes the plane $y = -1$ will be plotted. So a slight change in the position of the vortex around $y = -1$ might result in an obvious change in the velocity field in the figure. So it is better to compare the velocity field in a full picture on the whole space. For this reason, we plot the projection of the velocity on the plane $z = 0$ in Figure 5 – Figure 8. The direction of the velocity field is marked with arrows, whereas the magnitude is represented by different colours.

By comparing Figure 5 with Figure 6, and Figure 7 with Figure 8, we have two observations. Firstly, if we focus on the shape of the velocity field, Figure 6 and Figure 8 seem to be more consistent comparing with Figure 5 and 7, respectively. For example, at $t = 0$, there is a vortex at the location $(-1, -1)$, which can be observed from the figures or by checking the initial vorticity in the Table 1. Theoretically, we would expect this vortex to move away or disappear gradually along with time. This feature can be observed in Figure 6 and Figure 8. However, this vortex disappears in Figure 5, and moves to somewhere around $(-0.5, -2)$ in Figure 7. So we conclude that with a repetition of N (here we have chosen $N = 10$) times on the particle path, the velocity converges faster.

Secondly, less spikes appear in the velocity field under the modified approach. This might not be obvious from the plot as velocities are plotted in terms of stream lines. But if we examine the magnitude of the velocity, we see a difference. At time $t = 0$, most of the stream lines are coloured in violet in Figure 5 and Figure 7, which means the magnitude of the velocity is between 0 and 1. However, the maximum value reaches up to 14 in both figures, whose corresponding velocity streamline can be barely seen in the figures. The reason behind is the potential spike on the velocity fields. At some locations, the velocity might shoot up due to the stochastic error in the simulation. On the other hand, the highest magnitude of the velocity field in both Figure 6 and Figure 8 is 0.16, and the colour map seem to be more consistent.

In conclusion, the modified approach, which splits each particle trajectory paths into N copies, provides a better approximation from two aspects: Firstly, the simulated velocity field converges faster. Even the mesh size is as small as $M = 10$, it provides a more reliable approximation. Secondly, few stochastic errors occur, so the simulated velocity field is more stable.

4 Study of FBSDEs from the porous medium equation

In this chapter, we study the following system of quadratic forward–backward stochastic differential equations (FBSDEs)

$$\begin{cases} dX_t = \sqrt{2Y_t}dB_t, & X_0 = x, \\ dY_t = -\frac{1}{2}\kappa\frac{|Z_t|^2}{Y_t}dt + Z_t dB_t, & Y_T = \phi(X_T), \end{cases} \tag{4.1}$$

where $\kappa > 0$ is a universal constant, $B = (B_t)_{0\leq t\leq T}$ is a standard d-dimensional Brownian motion on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and $\phi : \mathbb{R}^d \to \mathbb{R}$ is uniformly Lipschitz continuous and bounded such that for all $x \in \mathbb{R}^d$, it yields that $c_1 \leq \phi(x) \leq c_2$ for some $c_1, c_2 > 0$. This system of forward–backward equations arises from the porous medium equation

$$\begin{cases} \frac{\partial u}{\partial t}(x,t) = \Delta u^m(x,t), & (x,t) \in \mathbb{R}^d \times (0,\infty), \\ u(x,0) = u_0(x), & x \in \mathbb{R}^d \end{cases}$$

We will explore the connection between the quasi-linear PDE and the FBSDE in later discussion.

In fact, FBSDEs appearing in the study of stochastic control problems are typically of quadratic growth in Z. In particular, Horst et al. [62] showed that this class of quadratic forward–backward stochastic differential equations may characterize solutions of utility maximization problems with non-trivial terminal endowment. To our best knowledge, Antonelli and Hamadène [2], Fromm and Imkeller [47], Luo and Tangqi [92], and Kupper et al. [81] are the only works studying the coupled forward–backward stochastic differential equations whose backward equation has quadratic growth in the variable in SZ, that is, the system of equations

$$
\begin{cases}
X_t = x + \int_0^t b(s, X_s, Y_s, Z_s)ds + \int_0^t \sigma(s, X_s, Y_s, Z_s)dB_s, \\
Y_t = h(X_T) + \int_t^T g(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dB_s,
\end{cases}
\tag{4.2}
$$

where g has quadratic growth in Z.

Antonelli and Hamadène [2] studied the one-dimensional problem with continuous coefficients and assumed that b is independent of Z and σ depends only on X. In addition, the terminal data of the backward equation was assumed to be independent of the solution of the forward equation, that it, $h(X_T) = \xi$ is a random variable independent of the process X. Then the authors constructed a monotone sequence of processes to approximate the solution of the FBSDE. They also pointed out that this method could be applied to the quadratic FBSDEs whose backward equation satisfied the same growth conditions as in [78]. In this scheme, the solution was shown to exist for any arbitrary time interval. However, the uniqueness was not concluded.

Fromm and Imkeller [47] discussed the fully coupled equations under the Markovian setting, where the coefficients were deterministic functions and independent of the variable $\omega \in \Omega$. By assuming the coefficients to be locally Lipschitz in (Y, Z), the authors established the existence of the local solution using the Picard iteration. Then with the decoupling fields, the local solution was extended to a global one via concatenation.

Luo and Tangqi [92] studied the multidimensional FBSDEs under a non-Markovian setting, where the coefficients of (4.2) are allowed to be random. The authors restricted h to be a bounded and Lipschitz function of X_T. In addition, b was Lipschitz, σ was dependent on (t, ω) for any $(t, \omega) \in [0, T] \times \Omega$, and g could be separated into a quadratic and a subquadratic part in the control variable Z.

Under this setting, the authors applied a fixed point argument with the help of BMO-martingales and obtained the existence and uniqueness of the solution on a small time interval. With the same growth conditions but the coefficients being restricted to be deterministic, Kupper et al. [81] established a local existence and uniqueness result for this system of Markovian equations based on a Malliavin calculus argument. Under additional monotonicity conditions on the coefficients, global solutions were constructed by a pasting technique based on PDE solutions.

In our study, we are concerned with the solvability of the system of FBSDEs in (4.1), where the generator for the backward equation has a quadratic growth in Z and an inverse growth in Y. Different from [2], the backward equation in our problem is coupled with the forward equation via both the coefficients and the terminal data. Therefore, the classical approximation method, which relies on constructing a monotone sequence of processes, fails for studying (4.1).

This chapter is structured as follows. In section 1, we will formally formulate the porous medium equation into a system of forward–backward stochastic differential equations. Section 2 gives an *a priori* estimate on the decoupled backward equation and studies the solvability of this decoupled BSDE. Section 3 establishes the existence and uniqueness of the FBSDE (4.1) on a small time interval by using a fixed point argument. Then we use this methodology to study a class of quadratic FBSDEs with more general growth conditions and prove its solvability.

4.1 Formulation of the FBSDE system

The porous medium equation (PME)

$$\frac{\partial u}{\partial t} = \Delta u^m, \tag{4.3}$$

in the Euclidean space, where $m > 1$, is a quasi-linear parabolic equation with initial condition $u(0, x) = u_0(x)$ for all $x \in \mathbb{R}^d$. This equation has been studied as a useful model for describing physical processes of liquid flows passing through a concrete medium. Due to this physical meaning, we are only interested in its positive solutions. For positive u, (4.3) can be written as

$$\frac{\partial u}{\partial t} - mu^{m-1}\Delta u - m(m-1)u^{m-2}|\nabla u|^2 = 0.$$

This differential operator $mu^{m-1}\Delta$ is uniformly elliptic if u is bounded, and is away from 0. For this case, the theory of fully non-linear equations may be applied, and the regularity of positive solutions follows from the Krylov-Safonov's theory [55].

To exploit the positivity of u, we may apply the Hopf transformation to introduce

$$f = \frac{m(u^{m-1} - 1)}{m - 1},$$

and rewrite the PME in terms of f so that

$$\frac{\partial f}{\partial t} = \big((m-1)f + m\big)\Delta f + |\nabla f|^2.$$

Notice that

$$(m-1)f + m = mu^{m-1} > 0.$$

Given any $T > 0$, we run a diffusion process $(X_t)_{0 \le t \le T}$ defined as the solution to the following stochastic differential equation:

$$dX_t = \sqrt{2\big((m-1)f(X_t, T-t) + m\big)}dB_t,$$

where $(B_t)_{0 \le t \le T}$ is a standard d-dimensional Brownian motion on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$. If there is no risk of confusion, we omit the specification of variables on $(X_t, T-t)$ in our computations below. It follows from the Itô's formula that

$$\begin{aligned}
& f(X_T, 0) - f(X_t, T-t) \\
= & -\int_t^T \frac{\partial f}{\partial s} ds + \int_t^T \sqrt{2\left((m-1)f + m\right)} \nabla f dB_s \\
& + \int_t^T \left((m-1)f + m\right) \Delta f ds \\
= & -\int_t^T |\nabla f|^2 ds + \int_t^T \sqrt{2\left((m-1)f + m\right)} \nabla f dB_s
\end{aligned}$$

for all $0 \le t \le T$. Let

$$Y_t = (m-1) f(X_t, T-t) + m,$$

and

$$Z_t = (m-1)\sqrt{2\left((m-1) f(X_t, T-t) + m\right)} \nabla f(X_t, T-t).$$

Then (Y, Z) satisfies the following backward stochastic differential equations

$$Y_T - Y_t = -\int_t^T \frac{|Z_s|^2}{2(m-1)Y_s} ds + \int_t^T Z_s dB_s, \quad Y_T = m\big(u_0(X_T)\big)^{m-1}.$$

Let $\phi(x) = (m-1)f(x,0) + m = m\big(u_0(x)\big)^{m-1}$. Then (Y, Z) is the solution to the following system of forward–backward stochastic differential equations

$$\begin{cases} dX_t = \sqrt{2Y_t} dB_t, & X_0 = x, \\ dY_t = -\frac{1}{2}\kappa \frac{|Z_t|^2}{Y_t} dt + Z_t dB_t, & Y_T = \phi(X_T), \end{cases} \tag{4.4}$$

where $\kappa = \frac{1}{m-1}$ is considered as a parameter.

In our study, we consider the problem (4.4) as an example of FBSDEs and solve this problem by using probabilistic method. Here we assume the function $\phi : \mathbb{R}^d \to \mathbb{R}_+$ is globally Lipschitz and $c_1 \leq \phi(x) \leq c_2$ for every $x \in \mathbb{R}^d$, where $c_1, c_2 > 0$ are two positive constants. For the rest of the discussion, whenever we mention c_1, c_2, we mean the bounds of the positive function ϕ.

Then the main difficulty of solving this system lies in the coupling between the forward equation and the backward equation. This leads to a circular dependence in the solution of both the forward and the backward equation.

Let $\mathscr{X}$ be the space of all $\mathscr{F}_T$-measurable random variables ξ such that $c_1 \leq \xi \leq c_2$. Let $\xi \in \mathscr{X}$. If $(U_t, V_t)_{0 \leq t \leq T}$ solves the following BSDE

$$dU_t = -\frac{1}{2}\kappa \frac{|V_t|^2}{U_t} dt + V_t dB_t, \quad U_T = \xi,$$

set

$$W_T = x + \int_0^T \sqrt{2U_t} dB_t,$$

and define $\xi \mapsto \phi(W_T)$. This defines a mapping from $\mathscr{X}$ to itself. Then the solution to the problem (4.4) are fixed points of this mapping.

4.2 Decoupled quadratic backward stochastic differential equations

Based on the observation described in Section 1, the first step is to study the solvability of the decoupled backward stochastic differential equation

$$dY_t = -\frac{1}{2}\kappa \frac{|Z_t|^2}{Y_t} dt + Z_t dB_t, \quad Y_T = \xi, \tag{4.5}$$

where $\kappa > 0, (B_t)_{0 \leq t \leq T}$ is a standard d-dimensional Brownian motion, and $\xi \in \mathscr{X}$. This is an example of BSDEs whose non-linear drive is quadratic in Z, and therefore the existing theories in BSDEs do not follow at least directly.

Kobylanski [78] was the first to study BSDEs with quadratic growth in the variable Z, then followed by [86]. In [78], Kobylanski applied an exponential change of variable and proved the existence of the solution using the comparison theorem for one-dimensional BSDEs (with Lipschitz coefficients). Briand and Hu [17, 18] relaxed the boundedness assumption on the terminal condition and obtained the existence of the solutions. Additionally, by assuming the generating function was convex in Z, the solution was unique. Moreover, the authors obtained a comparison theorem for solutions to unbounded quadratic BSDEs in [18]. Delbarn et al. [37] studied the uniqueness under these conditions using the stochastic control method. Tevzadze [135] studied a general quadratic BSDE driven by martingales. The existence and uniqueness of the solution was constructed by using Picard iteration method, which relied on the theory of BMO-martingales. In particular, Tevzadze showed that this argument could be extended to multidimensional quadratic BSDEs when the terminal data was small enough. One may also refer to [19, 25, 70, 91, 141] for more recent progress, especially extensions to higher dimensions.

This section is devoted to showing the BSDE (4.5) admits a unique solution on the space $\mathscr{H}_T^\infty(\mathbb{R}) \times \mathscr{H}_T^2(\mathbb{R}^d)$ for any terminal data $\xi \in \mathscr{X}$. Here we adapt an approximation method inspired by [78, 85], while our arguments or notations will be different from their settings in order to solve (4.5).

4.2.1 An a priori estimate

Firstly, we aim to establish an *a priori* estimate on a solution $(Y_t)_{0\leq t\leq T}$ to the BSDE (4.5). Specifically, we aim to show that if the terminal data ξ is bounded, that is, $c_1 \leq \xi \leq c_2$, then the process $(Y_t)_{0\leq t\leq T}$ is also bounded, and shares the same bounds for any $0 \leq t \leq T$. This *a priori* estimate will play a key part in showing the solvability of the BSDE (4.5), and also the coupled forward–backward equation (4.4). However, this boundedness property for a general BSDE does not hold in general. In particular, Delbaen et al. [36] indicated that bounded terminal data may not necessarily lead to bounded solution for quadratic BSDEs.

Lemma 4.1. *Define* $h : \mathbb{R} \to \mathbb{R}$ *such that*

$$h(x) = \begin{cases} e^x - 1 - x - \frac{x^2}{2}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

Then $h \in \mathcal{C}^2(\mathbb{R})$ *with* $h \geq 0, h' \geq 0, h'' \geq 0$. *For any* $\kappa \neq 0$, *define*

$$\psi(x) := h\left(|\kappa|(x-1)\right). \tag{4.6}$$

Then, with the convention that $0/0 = 0$,

$$\psi''(x) - \kappa \frac{\psi'(x)}{x} \geq 0, \tag{4.7}$$

for all $x \in \mathbb{R}$.

Proof. Clearly, h is twice differentiable for all $x > 0$ or $x < 0$. Also, it is easily checked that

$$h(0) = \lim_{x \downarrow 0} h'(x) = \lim_{x \downarrow 0} h''(x) = 0.$$

So $h \in \mathcal{C}^2(\mathbb{R})$ with $h' \geq 0, h'' \geq 0$. This also implies $\psi \in \mathcal{C}^2(\mathbb{R})$. So we may write

$$\psi''(x) - \kappa \frac{\psi'(x)}{x} = \kappa^2 \left(h''\left(|\kappa|(x-1)\right) - \frac{|\kappa|}{\kappa} \frac{h'\left(|\kappa|(x-1)\right)}{x} \right).$$

If $x < 1$, then with the convention that $0/0 = 0$, (4.7) holds true for any $\kappa \neq 0$. Now suppose $x \geq 1$. If $\kappa < 0$, it is trivial to see that (4.7) holds. So for now we assume $\kappa > 0$. Then

$$\begin{aligned}
\psi''(x) - \kappa\frac{\psi'(x)}{x} &= \kappa^2\left(h''\left(\kappa(x-1)\right) - \frac{h'\left(\kappa(x-1)\right)}{x}\right) \\
&= \kappa^2\left(e^{\kappa(x-1)} - 1 - \frac{1}{x}\left(e^{\kappa(x-1)} - 1 - \kappa(x-1)\right)\right) \\
&= \kappa^2\left(1 - \frac{1}{x}\right)\left(e^{\kappa(x-1)} - 1 + \kappa\right) \\
&\geq 0.
\end{aligned}$$

Hence, the inequality (4.7) follows for all $x \in \mathbb{R}$ and $\kappa \neq 0$. □

With the help of this function, we are ready to give an *a priori* estimate for the process Y defined in (4.5).

Lemma 4.2. *Given any $\kappa > 0$, let ξ be a real-valued process in the space $\mathscr{X}$. If the pair of processes $(Y, Z) \in \mathscr{H}_T^\infty(\mathbb{R}) \times \mathscr{H}_T^2(\mathbb{R}^d)$ is a solution to the BSDE* (4.5)*, i.e.,*

$$dY_t = -\frac{1}{2}\kappa\frac{|Z_t|^2}{Y_t}dt + Z_t dB_t, \quad Y_T = \xi,$$

then $c_1 \leq Y_t \leq c_2$ a.e. for all $0 \leq t \leq T$.

Proof. Let $\lambda > 0$ be any positive constant. If (Y, Z) satisfies the backward stochastic differential equation (4.5), then scaled process $(\lambda Y, \lambda Z)$ satisfies the following BSDE

$$d(\lambda Y_t) = -\frac{1}{2}\kappa\frac{|\lambda Z_t|^2}{\lambda Y_t}dt + \lambda Z_t dB_t, \quad \lambda Y_T = \lambda\xi.$$

Therefore, the pair of processes $(\lambda Y, \lambda Z)$ solves the BSDE (4.5) with terminal data $\lambda\xi$. If we choose $\lambda = \frac{1}{c_1}$ or $\lambda = \frac{1}{c_2}$, we may rescale the lower bound and the upper bound so that it is sufficient to show the following:

(i) if $0 < \xi \leq 1$ a.s., then $0 < Y_t \leq 1$ a.s. for all $0 \leq t \leq T$;

(ii) if $\xi \geq 1$ a.s., then $Y_t \geq 1$ a.s. for all $0 \leq t \leq T$.

To show (i), we apply the Itô's formula to $e^t\psi(Y_t)$, where ψ is a non-negative function of class $\mathcal{C}^2(\mathbb{R})$, then

$$\begin{aligned}
& e^T\psi(\xi) - e^t\psi(Y_t) \\
= & \int_t^T e^s\psi(Y_s)ds - \frac{1}{2}\kappa\int_t^T e^s\psi'(Y_s)\frac{|Z_s|^2}{Y_s}ds + \int_t^T e^s\psi'(Y_s)Z_s dB_s \\
& + \frac{1}{2}\int_t^T e^s\psi''(Y_s)|Z_s|^2 ds \\
= & \int_t^T e^s\psi(Y_s)ds + \frac{1}{2}\int_t^T e^s|Z_s|^2\left(\psi''(Y_s) - \kappa\frac{\psi'(Y_s)}{Y_s}\right)ds \\
& + \int_t^T e^s\psi'(Y_s)Z_s dB_s. \qquad (4.8)
\end{aligned}$$

With the Burkholder-Davis-Gundy inequality, we have

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\le t\le T}\int_t^T e^s\psi'(Y_s)Z_s dB_s\right] &\le 2\mathbb{E}\left[\left(\int_0^T \left|e^s\psi'(Y_s)Z_s\right|^2 ds\right)^{1/2}\right] \\
&\le 2e^T\left(\sup_{0\le s\le T}|\psi'(Y_s)|\right)\mathbb{E}\left[\left(\int_0^T |Z_s|^2 ds\right)^{1/2}\right].
\end{aligned}$$

Choose ψ to be the function defined as (4.6), then it follows from the assumption that $(Y,Z)\in\mathscr{H}_T^\infty(\mathbb{R})\times\mathscr{H}_T^2(\mathbb{R}^d)$ that the term $\int_t^T e^s\psi'(Y_s)Z_s dB_s$ is a martingale.

Taking expectations on both sides of (4.8), it follows from the fact that $\psi\ge 0$ and the inequality (4.7) that

$$\mathbb{E}\left[e^T\psi(\xi) - e^t\psi(Y_t)\right] \ge 0.$$

By construction, $\psi(x)=0$ if and only if $x\le 1$. So if $\xi\le 1$ a.e., then

$$\mathbb{E}\left[e^T\psi(\xi)\right] = 0.$$

Therefore, it follows from the positivity of the function ψ that

$$0 \leq \mathbb{E}\left[e^t\psi(Y_t)\right] \leq \mathbb{E}\left[e^T\psi(\xi)\right] = 0$$

for any $t \in [0,T]$. Hence,

$$\mathbb{E}\left[e^t\psi(Y_t)\right] = 0.$$

Since $e^t\psi(x) \geq 0$, then for any $0 \leq t \leq T$,

$$e^t\psi(Y_t) = 0 \quad a.e..$$

Again, $e^t\psi(x) = 0$ if and only if $x \leq 1$, it suffices to show that

$$Y_t \leq 1 \quad a.e.$$

for all $0 \leq t \leq T$.

Now we assume $\xi \geq 1$ a.s., we aim to prove $Y_t \geq 1$ a.s. for all $0 \leq t \leq T$. Define the paired processes $(H_t, \tilde{Z}_t)_{0\leq t\leq T}$ such that

$$H_t := \frac{1}{Y_t}, \qquad \tilde{Z}_t := -\frac{Z_t}{(Y_t)^2}.$$

Then an application of the Itô's formula shows that

$$dH_t = -\frac{1}{2}\tilde{\kappa}\frac{|\tilde{Z}_t|^2}{H_t}dt + \tilde{Z}_t dB_t, \quad H_T = \frac{1}{\xi}$$

with $\tilde{\kappa} = -2 - \kappa < 0$, and the terminal data $H_T \leq 1$. We apply the same argument as above with the parameter $\tilde{\kappa} < 0$. Thanks to Lemma 4.1, we draw the same conclusion that $H_t \leq 1$ a.e. for all $0 \leq t \leq T$. Hence, $Y_t \geq 1$ a.e. for all $0 \leq t \leq T$. $\square$

4.2.2 Existence and uniqueness result

With the boundedness property, we may define a subspace $\mathscr{Y} \subset \mathscr{H}_T^\infty(\mathbb{R})$ of those $Y = (Y_t)_{0 \le t \le T}$ such that $c_1 \le Y_t \le c_2$ a.s. for all $0 \le t \le T$. Lemma 4.2 implies that whenever $\xi \in \mathscr{X}$, any solution of the problem (4.5) lies in the space $\mathscr{Y}$.

Now we aim to show that the BSDE

$$dY_t = -\frac{1}{2}\kappa \frac{|Z_t|^2}{Y_t} dt + Z_t dB_t, \quad Y_T = \xi \tag{4.9}$$

admits a unique solution for any given terminal data $\xi \in \mathscr{X}$.

Define the function $F : \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ as

$$F(y,z) = \frac{1}{2}\kappa \frac{|z|^2}{\big(y \mathbb{1}_{[c_1,c_2]}(y) + c_1 \mathbb{1}_{(-\infty,c_1)}(y) + c_2 \mathbb{1}_{(c_2,\infty)}(y)\big)},$$

then (4.9) is a backward stochastic differential equation generated by (F, ξ):

$$dY_t = -F(Y_t, Z_t)dt + Z_t dB_t, \quad Y_T = \xi.$$

In other words,

$$Y_t = \xi + \int_t^T F(Y_s, Z_s) ds - \int_t^T Z_s ds.$$

We aim to prove the existence and uniqueness of the solution using a suitable approximation of the generator and a comparison theorem for Lipschitz BSDEs. This technique has been illustrated in the proof of Proposition 2.22, originally from [85].

Lemma 4.3. *For each $n \in \mathbb{N}$, define $f_n : \mathbb{R}^d \to \mathbb{R}_+$ by*

$$f_n(z) = \begin{cases} |z|^2, & \text{if } |z| \le n, \\ n^2, & \text{if } |z| > n. \end{cases}$$

Then we may define a sequence of functions $F_n : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ *such that*

$$F_n(y,z) := \frac{1}{2}\kappa \frac{f_n(z)}{y\mathbb{1}_{[c_1,c_2]}(y) + c_1\mathbb{1}_{(-\infty,c_1)}(y) + c_2\mathbb{1}_{(c_2,\infty)}(y)}. \tag{4.10}$$

Then F_n *is bounded and* $F_n \leq F_{n+1}$ *for each* n. *Also,* F_n *is globally Lipschitz continuous such that for any* $y^1, y^2 \in \mathbb{R}_+, z^1, z^2 \in \mathbb{R}^d$, *it yields that*

$$\left|F_n(y^1,z^1) - F_n(y^2,z^2)\right| \leq C_n \left(|y^1 - y^2| + |z^1 - z^2|\right),$$

where $C_n = \max\left(\frac{n^2\kappa}{2|c_1|^2}, \frac{n\kappa}{c_1}\right)$ *is the Lipschitz constant, depending on* n, c_1 *and* κ. *Moreover,*

$$\lim_{n\to} F_n(y^n, z^n) = F(y,z) \tag{4.11}$$

for any converging sequence $\lim_{n\to\infty} y^n = y, \lim_{n\to\infty} z^n = z$.

Proof. The boundedness and monotone properties are trivial. Now we show the function F_n is Lipschitz for each n. To simplify the notation, we define the function $g : \mathbb{R}_+ \to [c_1, c_2]$ such that

$$g(y) = y\mathbb{1}_{[c_1,c_2]}(y) + c_1\mathbb{1}_{(-\infty,c_1)}(y) + c_2\mathbb{1}_{(c_2,\infty)}(y).$$

Then for each n, we may write

$$\begin{aligned}\left|F_n(y^1,z^1) - F_n(y^2,z^2)\right| &\leq \left|F_n(y^1,z^1) - F_n(y^1,z^2)\right| + \left|F_n(y^1,z^2) - F_n(y^2,z^2)\right| \\ &= \frac{1}{2}\kappa\left(\frac{|f_n(z^1) - f_n(z^1)|}{g(y^1)} + f_n(z^2)\left|\frac{1}{g(y^1)} - \frac{1}{g(y^2)}\right|\right)\end{aligned}$$

for any $y^1, y^2 \in \mathbb{R}_+, z^1, z^2 \in \mathbb{R}^d$. Without loss of generality, we assume $y^1 \leq y^2, |z^1| \leq |z^2|$, then

$$\frac{|f_n(z^1) - f_n(z^2)|}{g(y^1)} \leq \frac{|f_n(z^1) - f_n(z^2)|}{c_1}$$

$$\leq \begin{cases} \frac{|z^2|^2-|z^1|^2}{c_1}, & \text{if } |z^1| \leq |z^2| \leq n^2, \\ \frac{n^2-|z^1|^2}{c_1}, & \text{if } |z^1| \leq n^2 \leq |z^2|, \\ 0, & \text{if } n^2 \leq |z^1| \leq |z^2|. \end{cases}$$

$$\leq \frac{2n}{c_1}|z^1 - z^2|.$$

On the other hand,

$$\begin{aligned} f_n(z^2)\left|\frac{1}{g(y^1)} - \frac{1}{g(y^2)}\right| &= f_n(z^2)\frac{|g(y^1) - g(y^2)|}{g(y^1)g(y^2)} \\ &\leq n^2\frac{|g(y^1) - g(y^2)|}{|c_1|^2} \\ &\leq \frac{n^2}{|c_1|^2}|y^1 - y^2|. \end{aligned}$$

Hence, we combine these two terms to get

$$\left|F_n(y^1, z^1) - F(y^2, z^2)\right| \leq \frac{n^2\kappa}{2|c_1|^2}|y^1 - y^2| + \frac{n\kappa}{c_1}|z^1 - z^2|.$$

Finally, the convergence of (4.11) follows from the continuity of the function F_n and the fact that f_n converges pointwise to the function $f(z) = z^2$. □

Lemma 4.4. *Let F_n be defined as in* (4.10) *and ξ be a measureable process with $c_1 \leq \xi \leq c_2$. Then for each $n \in \mathbb{N}$, the BSDE generated by (F_n, ξ) has a unique pair of solution $(Y^n, Z^n) \in \mathscr{Y} \times \mathscr{H}_T^2(\mathbb{R}^d)$. Moreover, the norm of $(Z^n)_n$ in the space $\mathscr{H}_T^2(\mathbb{R}^d)$ is independent of n such that*

$$\|Z^n\|^2_{\mathscr{H}_T^2} \leq \frac{4c_1^2}{\kappa^2}\left(e^{\frac{\kappa c_2}{c_1}} - \frac{\kappa c_2}{c_1} - 1\right). \tag{4.12}$$

Proof. Since F_n is globally Lipschitz, then it follows from Proposition 2.20 that the BSDE generated by (F_n, ξ) admits a unique pair of solution $(Y_t^n, Z_t^n)_{0\leq t\leq T}$

for each n. Moreover, thanks to the comparison theorem in Proposition 2.21, the sequence of processes $(Y^n)_n$ is monotone increasing, that is for each n, $Y_t^n \leq Y_t^{n+1}$ a.s. for all $0 \leq t \leq T$.

Let ψ be a function in $\mathcal{C}^2(\mathbb{R})$ with $\psi' > 0$. Then we may apply the Itô's formula to the process $\psi(Y^n)$ to obtain

$$\begin{aligned}\psi(Y_t^n) &= \psi(\xi) + \frac{\kappa}{2}\int_t^T \frac{\psi'(Y_s^n) f_n(Z_s^n)}{Y_s^n \mathbb{1}_{[c_1,c_2]}(Y_s^n) + c_1 \mathbb{1}_{(-\infty,c_1)}(Y_s^n) + c_2 \mathbb{1}_{(c_2,\infty)}(Y_s^n)} ds \\ &\quad - \frac{1}{2}\int_t^T \psi''(Y_s^n)|Z_s^n|^2 ds - \int_t^T \psi'(Y_s^n) Z_s^n dB_s \\ &\leq \psi(\xi) + \int_t^T \left(\frac{\kappa}{2c_1}\psi'(Y_s^n) - \frac{1}{2}\psi''(Y_s^n)\right)|Z_s^n|^2 ds - \int_t^T \psi'(Y_s^n) Z_s^n dB_s \end{aligned} \tag{4.13}$$

Choose $\psi(x) = e^{\frac{\kappa x}{c_1}}$, then ψ is monotone increasing with $\psi > 0$ and also

$$\frac{\kappa}{2c_1}\psi'(x) - \frac{\psi''(x)}{2} = 0.$$

So taking the conditional expectation on both sides of (4.13) at time t, and by the fact that $\int_t^T \psi'(Y_s^n) Z_s^n dB_s$ is a martingale, we deduce that

$$\psi(Y_t^n) = \mathbb{E}\left[\psi(Y_t^n)\middle|\mathscr{F}_t\right] \leq \mathbb{E}\left[\psi(\xi)\middle|\mathscr{F}_t\right] \leq \psi(c_2)$$

for any $\xi \in \mathscr{X}$. On the other hand, we also have

$$\psi(Y_t^n) \geq \psi(\xi) + \int_t^T \left(\frac{\kappa}{2c_2}\psi'(Y_s^n) - \frac{1}{2}\psi''(Y_s^n)\right)|Z_s^n|^2 ds - \int_t^T \psi'(Y_s^n) Z_s^n dB_s.$$

By choosing $\psi(x) = e^{\frac{\kappa x}{c_2}}$ and taking conditional expectations, it yields

$$\psi(Y_t^n) \geq \mathbb{E}\left[\psi(\xi)\middle|\mathscr{F}_t\right] \geq \psi(c_1).$$

By the construction of the function ψ, we may recover the bounds for Y^n and conclude that $Y^n \in \mathscr{Y}$ for each n.

Now we come back to the inequality (4.13) and choose

$$\psi(x) = \frac{2c_1^2}{\kappa^2}\left(e^{\frac{\kappa x}{c_1}} - \frac{\kappa}{c_1}x - 1\right),$$

so that $\psi'(x) \geq 0$ for all $x \geq 0$, and

$$\frac{\kappa}{2c_1}\psi'(x) - \frac{\psi''(x)}{2} = -1.$$

So (4.13) may be written as

$$\int_t^T |Z_s^n|^2\, ds \leq \psi(\xi) - \psi\left(Y_t^n\right) - \int_t^T \psi'\left(Y_s^n\right) Z_s^n dB_s.$$

Again, taking expectations on both sides yields

$$\|Z^n\|_{\mathscr{H}_T^2}^2 = \mathbb{E}\left[\int_0^T |Z_s^n|^2\, ds\right] = \mathbb{E}\left[\psi(\xi) - \psi\left(Y_t^n\right)\right] \leq \frac{4c_1^2}{\kappa^2}\left(e^{\frac{\kappa c_2}{c_1}} - \frac{\kappa c_2}{c_1} - 1\right) < \infty.$$

□

Before proceeding to showing the convergence of the sequences $(Y^n)_n, (Z^n)_n$, we need the following lemma.

Lemma 4.5. *Let $(Z^n)_n$ be a uniformly bounded sequence in $\mathscr{H}_T^2(\mathbb{R}^d)$ such that $(Z^n)_n$ converges weakly to some $Z \in \mathscr{H}_T^2(\mathbb{R}^d)$. If $(H^n)_n$ is another bounded sequence in $\mathscr{H}_T^2(\mathbb{R}^d)$ and $H^n \to H$ in $\mathscr{H}_T^2(\mathbb{R}^d)$, then $H^n Z^n \to HZ$ weakly.*

Proof. Let $g \in \mathscr{H}_T^2(\mathbb{R}^d)$ be bounded. Since $(Z^n)_n$ converges weakly to Z, so that

$$\mathbb{E}\left[\int_0^T H_s Z_s^n g(s) ds\right] \to \mathbb{E}\left[\int_0^T H_s Z_s g(s) ds\right]$$

as $n \to \infty$. In particular,

$$\mathbb{E}\left[\int_0^T |Z_s|^2 ds\right] \leq \liminf_n \mathbb{E}\left[\int_0^T |Z_s^n|^2 ds\right].$$

On the other hand, since $H^n \to H$ in $\mathscr{H}_T^2(\mathbb{R}^d)$, then

$$\begin{aligned}
&\left|\mathbb{E}\int_0^T (H^n - H_s) Z_s^n g(s) ds\right| \\
&\leq \|g\|_\infty \sup_n \|Z^n\|_{\mathscr{H}_T^2(\mathbb{R}^d)} \|H^n - H\|_{\mathscr{H}_T^2(\mathbb{R}^d)} \to 0.
\end{aligned}$$

Therefore, by using triangle inequality, we may conclude that

$$\mathbb{E}\left[\int_0^T H^n Z_s^n g(s) ds\right] \to \mathbb{E}\left[\int_0^T H_s Z_s g(s) ds\right]$$

as $n \to \infty$ for any bounded processes $g \in \mathscr{H}_T^2(\mathbb{R}^d)$.

For a general $g \in \mathscr{H}_T^2(\mathbb{R}^d)$, consider

$$J_n := \left|\mathbb{E}\left[\int_0^T H_s^n Z_s^n g(s) ds\right] - \mathbb{E}\left[\int_0^T H_s Z_s g(s) ds\right]\right|.$$

For every $\varepsilon > 0$, we may find an N such that

$$\mathbb{E}\left[\int_0^T |g_N(s) - g(s)|^2 ds\right] < \varepsilon,$$

where

$$g_N(s) = -N \vee g(s) \wedge N.$$

Thus, by the triangle inequality,

$$J_n \leq \left|\mathbb{E}\left[\int_0^T H_s^n Z_s^n g_N(s) ds\right] - \mathbb{E}\left[\int_0^T H_s Z_s g_N(s) ds\right]\right|$$

$$
\begin{aligned}
&+\left|\mathbb{E}\left[\int_0^T H_s^n Z_s^n\left(g_N(s)-g(s)\right)ds\right]\right| \\
&+\left|\mathbb{E}\left[\int_0^T H_s Z_s\left(g_N(s)-g(s)\right)ds\right]\right| \\
\leq &\left|\mathbb{E}\left[\int_0^T H_s^n Z_s^n g_N(s)ds\right]-\mathbb{E}\left[\int_0^T H_s Z_s g_N(s)ds\right]\right| \\
&+2\sup_n |H^n|\sup_n \|Z^n\|_{\mathscr{H}_T^2}\|g_N-g\|_{\mathscr{H}_T^2}.
\end{aligned}
$$

By sending $n\to\infty$ to obtain

$$
\limsup_{n\to\infty} J_k \leq 2\sup_n |H_n|\sup_n \|Z^n\|_{\mathscr{H}_T^2}\,\varepsilon.
$$

Since ε can be any given positive number, we therefore can conclude that

$$
\limsup_{n\to\infty} J_n \leq 0.
$$

So we must have

$$
\mathbb{E}\left[\int_0^T H^n Z_s^n g(s)ds\right]\to\mathbb{E}\left[\int_0^T H_s Z_s g(s)ds\right]
$$

as $n\to 0$ for any $g\in\mathscr{H}_T^2(\mathbb{R}^d)$, which by definition shows that

$$
H^n Z_s^n \to H_s Z_s
$$

weakly as $n\to\infty$. □

Lemma 4.6. *For each $n\in\mathbb{N}$, let $(Y^n,Z^n)_n$ be the unique solution to the BSDE generated by (F_n,ξ), where F_n is defined in* (4.10) *and $\xi\in\mathscr{X}$. Then there exists $(Y,Z)\in\mathscr{Y}\times\mathscr{H}_T^2(\mathbb{R}^d)$ such that $(Y^n)_{n\in\mathbb{N}}$ converges to Y uniformly, and $(Z^n)_{n\in\mathbb{N}}$ converges to Z in $\mathscr{H}_T^2(\mathbb{R}^d)$. As a consequence, (Y,Z) is a solution of the BSDE* (4.9)*.*

Proof. This proof is inspired by [85], and some of the arguments below are modifications in [78].

Since $Y^n \in \mathscr{Y}$ is monotone in n, so that $Y^n \uparrow Y \in \mathscr{Y}$, and $\lim_{n\to\infty} \mathbb{E}\left[Y_t^n\right] = \mathbb{E}\left[Y_t\right]$ for every $0 \leq t \leq T$. Since Y^n are continuous semimartingales, so that there exists a subsequence $(Y^{n_k})_k$ of $(Y^n)_n$ such that $(Y_t^{n_k})_k$ converges to Y_t almost surely for all $t \leq T$ (see [59] for example). Since Y^n is monotone in n, so that $Y_t^n \to Y_t$ almost surely for all $t \leq T$.

According to (4.12), $(Z^n)_n$ are bounded uniformly in $\mathscr{H}_T^2(\mathbb{R}^d)$, which is a separable Hilbert space, and the sequence is weakly compact. Hence, there is a subsequence $(Z^{n_k})_k$ of $(Z^n)_n$ such that $(Z^{n_k})_k$ converges weakly to some $Z \in \mathscr{H}_T^2(\mathbb{R}^d)$, that is,

$$\mathbb{E}\left[\int_0^T H_s \cdot (Z_s^{n_k} - Z_s)\, ds\right] \to 0$$

as $k \to \infty$ for any $H \in \mathscr{H}_T^2(\mathbb{R}^d)$. In particular,

$$\mathbb{E}\left[\int_0^T |Z_s|^2\, ds\right] \leq \liminf_{k\to\infty} \mathbb{E}\left[\int_0^T |Z_s^{n_k}|^2\, ds\right].$$

Therefore, Z satisfies (4.12) too.

Step 1. We show $(Z^n)_n$ converges to Z in $\mathscr{H}_T^2(\mathbb{R}^d)$.

For any $n, m \in \mathbb{N}$, we can write

$$\begin{aligned}
|F_n(y_1, z_1) - F_m(y_2, z_2)| &\leq |F_n(y_1, z_1)| + |F_m(y_2, z_2)| \\
&\leq \frac{\kappa}{2c_1}|z_1|^2 + \frac{\kappa}{2c_1}|z_2|^2 \\
&\leq \frac{\kappa}{2c_1}|z_1 - z_2|^2 + \frac{\kappa}{c_1}|z_2 - z|^2 + \frac{\kappa}{c_1}|z|^2
\end{aligned}$$

for any $c_1 \leq y_1, y_2 \leq c_2, z_1, z_2, z \in \mathbb{R}^d$.

Let ψ be a $\mathcal{C}^2$-function such that $\psi'(x) \geq 0$ for every $x \geq 0$, and $\psi(0) = 0$. Let

$n_k > m$ so that $Y^{n_k} \geq Y^m$. Apply the Itô's formula, we have

$$
\begin{aligned}
&\psi\left(Y_t^{n_k} - Y_t^m\right) \\
&= \int_t^T \psi'(Y_s^{n_k} - Y_s^m)\left(F_{n_k}\left(Y_s^{n_k}, Z_s^{n_k}\right) - F_m\left(Y_s^m, Z_s^m\right)\right) ds \\
&\quad - \frac{1}{2}\int_t^T \psi''\left(Y_s^{n_k} - Y_s^m\right)\left|Z_s^{n_k} - Z_s^m\right|^2 ds \\
&\quad - \int_t^T \psi'\left(Y_s^{n_k} - Y_s^m\right)\left(Z_s^{n_k} - Z_s^m\right) dB_s \\
&\leq \int_t^T \psi'(Y_s^{n_k} - Y_s^m)\left(\frac{\kappa}{2c_1}|Z^{n_k} - Z^m|^2 + \frac{\kappa}{c_1}|Z^m - Z|^2 + \frac{\kappa}{c_1}|Z|^2\right) ds \\
&\quad - \frac{1}{2}\int_t^T \psi''\left(Y_s^{n_k} - Y_s^m\right)\left|Z_s^{n_k} - Z_s^m\right|^2 ds \\
&\quad - \int_t^T \psi'\left(Y_s^{n_k} - Y_s^m\right)\left(Z_s^{n_k} - Z_s^m\right) dB_s.
\end{aligned}
$$

Rearrange the terms on the right-hand side to obtain

$$
\begin{aligned}
\psi\left(Y_t^{n_k} - Y_t^m\right) \leq & -\int_t^T \left(\frac{1}{2}\psi'' - \frac{\kappa}{2c_1}\psi'\right)\left(Y_s^{n_k} - Y_s^m\right)\left|Z_s^{n_k} - Z_s^m\right|^2 ds \\
& + \frac{\kappa}{c_1}\int_t^T \psi'(Y_s^{n_k} - Y_s^m)|Z^m - Z|^2 ds \\
& + \frac{\kappa}{c_1}\int_t^T \psi'(Y_s^{n_k} - Y_s^m)|Z|^2 ds \\
& - \int_t^T \psi'\left(Y_s^{n_k} - Y_s^m\right)\left(Z_s^{n_k} - Z_s^m\right) dB_s.
\end{aligned}
$$

Now taking expectations on both sides, to deduce that

$$
\begin{aligned}
\mathbb{E}\left[\psi\left(Y_0^{n_k} - Y_0^m\right)\right] \leq & -\mathbb{E}\left[\int_0^T \left(\frac{1}{2}\psi'' - \frac{\kappa}{2c_1}\psi'\right)\left(Y_s^{n_k} - Y_s^m\right)\left|Z_s^{n_k} - Z_s^m\right|^2 ds\right] \\
& + \frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s^{n_k} - Y_s^m)|Z^m - Z|^2 ds\right] \\
& + \frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s^{n_k} - Y_s^m)|Z|^2 ds\right].
\end{aligned}
$$

Suppose $\psi'(x) \geq 0$ for $x \geq 0$, then from $Y^{n_k} \leq Y$, we deduce that

$$
\begin{aligned}
\mathbb{E}\left[\psi\left(Y_0^{n_k} - Y_0^m\right)\right] \leq \mathbb{E}&\left[\int_0^T \left(-\frac{1}{2}\psi'' + \frac{\kappa}{2c_1}\psi'\right)\left(Y_s^{n_k} - Y_s^m\right)\left|Z_s^{n_k} - Z_s^m\right|^2 ds\right] \\
&+ \frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s - Y_s^m)|Z^m - Z|^2 ds\right] \\
&+ \frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s - Y_s^m)|Z|^2 ds\right].
\end{aligned}
$$

Choose $\psi(x) = \frac{2c_1^2}{9\kappa^2}(e^{\frac{3\kappa}{c_1}x} - \frac{3\kappa}{c_1}x - 1)$ for $x \geq 0$ so that

$$
-\frac{1}{2}\psi'' + \frac{3\kappa}{2c_1}\psi' = -1, \quad \psi'(0) = \psi(0) = 0.
$$

Then

$$
\begin{aligned}
&\mathbb{E}\left[\psi\left(Y_0^{n_k} - Y_0^m\right)\right] + \mathbb{E}\left[\int_0^T \left|Z_s^{n_k} - Z_s^m\right|^2 ds\right] \\
&\leq -\frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s^{n_k} - Y_s^m)\left|Z_s^{n_k} - Z_s^m\right|^2 ds\right] \\
&\quad + \frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s - Y_s^m)|Z^m - Z|^2 ds\right] \\
&\quad + \frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s - Y_s^m)|Z|^2 ds\right].
\end{aligned}
$$

By taking limit on $k \to \infty$, one gets from the previous estimate that

$$
\begin{aligned}
&\mathbb{E}\left[\psi\left(Y_0 - Y_0^m\right)\right] + \liminf_k \mathbb{E}\left[\int_0^T \left|Z_s^{n_k} - Z_s^m\right|^2 ds\right] \\
&\leq \liminf_k -\mathbb{E}\left[\int_0^T \frac{\kappa}{c_1}\psi'\left(Y_s^{n_k} - Y_s^m\right)\left|Z_s^{n_k} - Z_s^m\right|^2 ds\right] \\
&\quad + \frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s - Y_s^m)|Z^m - Z|^2 ds\right] \\
&\quad + \frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s - Y_s^m)|Z|^2 ds\right].
\end{aligned}
$$

Since $\psi'(Y_s - Y_s^m) \geq 0$ is bounded, and $(Z^{n_k} - Z_m)_k \to Z - Z^m$ weakly, we may apply Lemma 4.5 with $H_s^k = \sqrt{\psi'(Y_s^{n_k} - Y_s^m)}$ for $0 \leq s \leq T$, so that for each m,

$$\sqrt{\psi'(Y_s^{n_k} - Y_s^m)}(Z^m - Z^{n_k}) \to \sqrt{\psi'(Y_s - Y_s^m)}(Z^m - Z) \quad \text{weakly as } k \to \infty.$$

In particular,

$$\left\| \sqrt{\psi'(Y_s - Y_s^m)}(Z^m - Z) \right\|_{\mathscr{H}_T^2(\mathbb{R}^d)} \leq \liminf_k \left\| \sqrt{\psi'(Y_s^{n_k} - Y_s^m)}(Z^m - Z^{n_k}) \right\|_{\mathscr{H}_T^2(\mathbb{R}^d)}.$$

This inequality is equivalent to that

$$\mathbb{E}\left[\int_0^T \psi'(Y_s - Y_s^m)|Z^m - Z|^2 ds\right] \leq \liminf_k \mathbb{E}\left[\int_0^T \psi'(Y_s^{n_k} - Y_s^m)|Z^m - Z^{n_k}|^2 ds\right].$$

Putting together these estimates we thus have the following key estimate

$$\mathbb{E}\left[\psi\left(Y_0 - Y_0^m\right)\right] + \liminf_k \mathbb{E}\left[\int_0^T \left|Z_s^{n_k} - Z_s^m\right|^2 ds\right] \leq \frac{\kappa}{c_1}\mathbb{E}\left[\int_0^T \psi'(Y_s - Y_s^m)|Z|^2 ds\right].$$

Since $Z^{n_k} - Z^m \to Z - Z^m$ weakly, so that

$$\mathbb{E}\left[\int_0^T \left|Z_s - Z_s^m\right|^2 ds\right] \leq \liminf_k \mathbb{E}\left[\int_0^T \left|Z_s^{n_k} - Z_s^m\right|^2 ds\right]$$

and therefore

$$\mathbb{E}\left[\psi\left(Y_0 - Y_0^m\right)\right] + \mathbb{E}\left[\int_0^T \left|Z_s - Z_s^m\right|^2 ds \leq \frac{\kappa}{c_1}\mathbb{E}\int_0^T \psi'(Y_s - Y_s^m)|Z|^2 ds\right]$$

for any m. Letting $m \to \infty$, we conclude that

$$\mathbb{E}\left[\int_0^T \left|Z_s - Z_s^m\right|^2 ds\right] \to 0.$$

Thus $Z^m \to Z$ in $\mathscr{H}_T^2(\mathbb{R}^d)$.

Step 2. We show the uniform convergence of the process $(Y^n)_n$ to Y, that is,

$$\lim_{n\to\infty} \sup_{0\le t\le T} |Y_t^n - Y_t| = 0.$$

Since $(Z^n)_n$ converges to Z in $\mathscr{H}_T^2(\mathbb{R}^d)$, then it follows from the continuity property of stochastic integrals (see [75, Proposition 2.26, page 147]) that

$$\lim_{n\to\infty} \sup_{0\le t\le T} \left| \int_t^T Z_s^n dB_s - \int_t^T Z_s dB_s \right| = 0 \quad \text{in probability.}$$

By passing to a subsequence, we may obtain an almost sure convergence, that is

$$\lim_{k\to\infty} \sup_{0\le t\le T} \left| \int_t^T Z_s^{n_k} dB_s - \int_t^T Z_s dB_s \right| = 0 \quad \text{a.s..} \tag{4.14}$$

On the other hand, it has been shown in Lemma 4.3 that

$$\lim_{n\to\infty} F_n(y^n, z^n) = F(y, z)$$

for any $y^n \to y, z^n \to z$. Therefore, it follows from the almost sure convergence of the sequence $(Y^{n_k}, Z^{n_k})_k$ that

$$\lim_{k\to\infty} F_{n_k}\left(Y_t^{n_k}, Z_t^{n_k}\right) = F\left(Y_t, Z_t\right)$$

almost surely for all $0 \le t \le T$. Also, since

$$\left|F_n\left(Y^n, Z^n\right)\right| \le \frac{\kappa}{2c_1} \left(\sup_n |Z^n| \right)^2,$$

with $\sup_n |Z^n| \in \mathscr{H}_T^2(\mathbb{R}^d)$ (see [78, Lemma 2.5, page 569] for example), it follows from the Lebesgue's dominated convergence theorem that

$$\lim_{k\to\infty}\int_t^T F_{n_k}\left(Y_s^{n_k}, Z_s^{n_k}\right) ds = \int_t^T F\left(Y_s, Z_s\right) ds$$

for all $0 \leq t \leq T$ and almost all $\omega \in \Omega$.

If we take any two subsequences such that the almost sure convergence (4.14) holds for $(Z^{n_k})_k, (Z^{n_l})_l$, then

$$\begin{aligned}|Y_t^{n_k} - Y_t^{n_l}| \leq &\int_t^T \left|F_{n_k}\left(Y_s^{n_k}, Z_s^{n_k}\right) - F_{n_l}\left(Y_s^{n_l}, Z_s^{n_l}\right)\right| ds \\ &+ \left|\int_t^T Z_s^{n_k} dB_s - \int_t^T Z_s^{n_l} dB_s\right|.\end{aligned}$$

We consider the supremum over $t \in [0,T]$ on both sides and take the limit on $k \to \infty$, then it yields that

$$\begin{aligned}\sup_{0\leq t\leq T}|Y_t - Y_t^{n_l}| \leq &\int_t^T \left|F\left(Y_s, Z_s\right) - F_{n_l}\left(Y_s^{n_l}, Z_s^{n_l}\right)\right| ds \\ &+ \left|\int_t^T Z_s dB_s - \int_t^T Z_s^{n_l} dB_s\right| \\ \to 0&\end{aligned}$$

as $l \to \infty$. This concludes the uniform convergence of the sequence $(Y^{n_l})_l$ to the process Y. Again, by the monotonicity of the sequence, this uniform convergence holds for the entire sequence. By passing to the limit in the following BSDE

$$Y_t^n = \xi + \int_t^T F^n(Y_s^n, Z_s^n) ds - \int_t^T Z_s^n dB_s,$$

we conclude that $(Y, Z) \in \mathscr{Y} \times \mathscr{H}_T^2(\mathbb{R}^d)$ is a solution of the BSDE generated by the parameter (F, ξ). □

Lemma 4.7. *Let $\xi \in \mathscr{X}$ be a measurable with $c_1 \leq \xi \leq c_2$. Then the backward* *stochastic differential equation*

$$dY_t = -\frac{1}{2}\kappa\frac{|Z_t|^2}{Y_t}dt + Z_t dB_t, \quad Y_T = \xi \tag{4.15}$$

admits a unique pair of solution $(Y, Z) \in \mathscr{Y} \times \mathscr{H}_T^2(\mathbb{R}^d)$.

Proof. The existence of the solution follows immediately from Lemma 4.6. We are are now concerned with the uniqueness.

Suppose both (Y_t^1, Z_t^1) and (Y_t^2, Z_t^2) are solutions to the BSDE (4.15). To simplify the notation, we define

$$Y := Y^1 - Y^2, \qquad Z := Z^1 - Z^2,$$

and also define the process Y^+ to be

$$Y_t^+ := \max(Y_t, 0) = \max\left(Y_t^1 - Y_t^2, 0\right)$$

for any $0 \leq t \leq T$. With an application of the Tanaka's formula to Y^+ (see [75, Chapter 3.6.A]),

$$\begin{aligned}
dY_t^+ &= \mathbb{1}_{\{Y_t>0\}}dY_t - dK_t \\
&= -\frac{\kappa}{2}\mathbb{1}_{\{Y_t>0\}}\left(\frac{|Z_t^1|^2}{Y_t^1} - \frac{|Z_t^2|^2}{Y_t^2}\right)dt + \mathbb{1}_{\{Y_t>0\}}Z_t dB_t - dK_t,
\end{aligned}$$

where K_t is a continuous increasing process with $K_0 = 0$ and K_t increases only when $Y_t^+ = 0$.

Define $f : \mathbb{R}\backslash\{0\} \times \mathbb{R}^d \to \mathbb{R}$ by

$$f(y, z) := \frac{\kappa|z|^2}{2y}.$$

Then for any $y \neq 0, z \in \mathbb{R}^d$, and $\epsilon > 0$,

$$\frac{\partial f}{\partial y}(y,z) + \epsilon \left| \frac{\partial f}{\partial z} \right|^2 (y,z) = \left(\epsilon \kappa^2 - \frac{\kappa}{2} \right) \frac{|z|^2}{y^2}. \tag{4.16}$$

On the other hand, by the Mean-Value Theorem,

$$\begin{aligned} & f(y^1, z^1) - f(y^2, z^2) \\ & = \left(\int_0^1 \frac{\partial f}{\partial y}(y_\lambda, z_\lambda) d\lambda \right) \left(y^1 - y^2 \right) + \left(\int_0^1 \frac{\partial f}{\partial z}(y_\lambda, z_\lambda) d\lambda \right) \left(z^1 - z^2 \right), \end{aligned}$$

for any $y^1, y^2 \in [c_1, c_2]$ and $z^1, z^2 \in \mathbb{R}^d$, where

$$(y_\lambda, z_\lambda) = \left(\lambda y^1 + (1-\lambda) y^2, \lambda z^1 + (1-\lambda) z^2 \right).$$

We may also define Y_t^λ, Z_t^λ in a similar manner such that

$$(Y_t^\lambda, Z_t^\lambda) = \left(\lambda Y_t^1 + (1-\lambda) Y_t^2, \lambda Z_t^1 + (1-\lambda) Z_t^2 \right),$$

so that for any $0 \leq t \leq T$,

$$\begin{aligned} & f(Y_t^1, Z_t^1) - f(Y_t^2, Z_t^2) \\ & = \left(\int_0^1 \frac{\partial f}{\partial y}(Y_t^\lambda, Z_t^\lambda) d\lambda \right) \left(Y_t^1 - Y_t^2 \right) + \left(\int_0^1 \frac{\partial f}{\partial z}(Y_t^\lambda, Z_t^\lambda) d\lambda \right) \left(Z_t^1 - Z_t^2 \right). \end{aligned}$$

Let $p \geq 2$ and apply the Itô's formula to $\left(Y_t^+ \right)^p$. Then for any $\epsilon > 0$, it follows from the inequality $ab \leq \epsilon a^2 + \frac{1}{4\epsilon} b^2$ that

$$\begin{aligned} & \left(Y_t^+ \right)^p + \frac{1}{2} p(p-1) \int_t^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left(Y_s^+ \right)^{p-2} |Z_s|^2 ds \\ = & \; p \int_t^T \left(Y_s^+ \right)^{p-1} \left(f\left(Y_s^1, Z_s^1 \right) - f\left(Y_s^2, Z_s^2 \right) \right) ds \end{aligned}$$

$$
\begin{aligned}
&- p\int_t^T \left(Y_s^+\right)^{p-1} Z_s dB_s + p\underbrace{\int_t^T \left(Y_s^+\right)^{p-1} dK_s}_{=0}\\
&= p\int_t^T \left(Y_s^+\right)^{p-1}\left(\int_0^1 \frac{\partial f}{\partial y}(Y_s^\lambda, Z_s^\lambda)d\lambda\right)\left(Y_s^1 - Y_s^2\right) ds\\
&\quad + p\int_t^T \left(Y_s^+\right)^{p-1}\left(\int_0^1 \frac{\partial f}{\partial z}(Y_s^\lambda, Z_s^\lambda)d\lambda\right)\left(Z_s^1 - Z_s^2\right) ds\\
&\quad - p\int_t^T \left(Y_s^+\right)^{p-1} Z_s dB_s\\
&\le p\int_t^T \left(Y_s^+\right)^{p}\left(\int_0^1 \frac{\partial f}{\partial y}(Y_s^\lambda, Z_s^\lambda)d\lambda\right) ds\\
&\quad + p\epsilon\int_t^T \left(Y_s^+\right)^{p}\left|\int_0^1 \frac{\partial f}{\partial z}(Y_s^\lambda, Z_s^\lambda)d\lambda\right|^2 ds\\
&\quad + \frac{p}{4\epsilon}\int_t^T \mathbb{1}_{\{\Delta Y_s \ge 0\}}\left(Y_s^+\right)^{p-2}|Z_s|^2 ds - p\int_t^T \left(Y_s^+\right)^{p-1} Z_s dB_s\\
&= p\int_t^T \left(Y_s^+\right)^{p}\left(\int_0^1 \left(\frac{\partial f}{\partial y}(Y_s^\lambda, Z_s^\lambda) + \epsilon\left|\frac{\partial f}{\partial z}(Y_s^\lambda, Z_s^\lambda)\right|^2\right) d\lambda\right) ds\\
&\quad + \frac{p}{4\epsilon}\int_t^T \mathbb{1}_{\{\Delta Y_s \ge 0\}}\left(Y_s^+\right)^{p-2}|Z_s|^2 ds - p\int_t^T \left(Y_s^+\right)^{p-1} Z_s dB_s\\
&= p\int_t^T \left(Y_s^+\right)^{p}\left(\int_0^1 \left(\left(\epsilon\kappa^2 - \frac{\kappa}{2}\right)\frac{|Y_s|^2}{|Z_s|^2}\right) d\lambda\right) ds\\
&\quad + \frac{p}{4\epsilon}\int_t^T \mathbb{1}_{\{\Delta Y_s \ge 0\}}\left(Y_s^+\right)^{p-2}|Z_s|^2 ds - p\int_t^T \left(Y_s^+\right)^{p-1} Z_s dB_s,
\end{aligned}
$$

where the last equation follows from the identity (4.16).

Since Y is a bounded process, then $\int_t^T \left(Y_s^+\right)^{p-1} Z_s dB_s$ is a martingale. Choose $\epsilon = \frac{1}{2\kappa} > 0$ and take expectations on both sides, we obtain

$$
\mathbb{E}\left[(Y_t^+)^p\right] + \frac{p}{2}\left(p - 1 - \kappa\right)\mathbb{E}\left[\int_t^T \mathbb{1}_{\{\Delta Y_s \ge 0\}}\left(Y_s^+\right)^{p-2}|Z_s|^2 ds\right] = 0.
$$

So for any $\kappa > 0$, we may choose $p > \max(1+\kappa, 2)$ so that

$$
\mathbb{E}\left[(Y_t^+)^p\right] \le 0.
$$

Together with the fact that $Y_t^+ \geq 0$ for any $0 \leq t \leq T$, we may conclude that

$$Y_t^+ = 0 \quad a.s.$$

for any $0 \leq t \leq T$, i.e.

$$Y_t^1 \leq Y_t^2 \quad a.s..$$

for any $0 \leq t \leq T$.

Conversely, if we define $Y = Y^2 - Y^1$, $Z = Z^2 - Z^1$ and with the same argument as above, we have

$$Y_t^2 \leq Y_t^1 \quad a.s.$$

for any $0 \leq t \leq T$. Hence, for any $t \in [0, T]$,

$$Y_t^1 = Y_t^2 \quad a.s.,$$

which implies the uniqueness of the solution. □

4.3 Study of the FBSDE for small time duration

Given $\phi : \mathbb{R}^d \to \mathbb{R}$ and $\xi \in \mathscr{X}$, we have constructed a non-linear mapping $\Phi : \xi \mapsto \phi(X_T)$ by

$$\Phi(\xi) = \phi\left(x + \int_0^T \sqrt{2Y_s} dB_s\right),$$

where the process $(Y_t)_{0 \leq t \leq T}$ is the unique solution to the following backward stochastic differential equation

$$Y_t = \xi + \frac{1}{2}\kappa \int_t^T \frac{|Z_s|^2}{Y_s} ds - \int_t^T Z_s dB_s. \tag{4.17}$$

We assume that ϕ is bounded such that $c_1 \leq \phi(x) \leq c_2$ for all $x \in \mathbb{R}^d$, then the random variable $\Phi(\xi)$ always lies in the space $\mathscr{X}$. In particular, we have shown in earlier section that given any terminal data $\xi \in \mathscr{X}$, the decoupled BSDE (4.17) admits a unique pair of solution $(Y, Z) \in \mathscr{X} \times \mathscr{H}_T^2(\mathbb{R}^d)$. So Φ is well-defined from $\mathscr{X}$ to itself.

This section aims to show that this mapping has a unique fixed point so that the system of forward–backward stochastic differential equations (4.4) admits a unique solution.

Lemma 4.8. *Let $\xi, \eta \in \mathscr{X}$ be bounded measurable random variables. If (Y^ξ, Z^ξ), (Y^η, Z^η) are the solutions of the BSDE defined in (4.17) with terminal data ξ, η, respectively. Then*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|Y_t^\xi - Y_t^\eta\right|^p\right] \leq \frac{16p}{p-1-\kappa}\mathbb{E}\left[|\xi-\eta|^p\right], \tag{4.18}$$

for any $p > \max(1+\kappa, 2)$.

Proof. Define $f : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ by

$$f(y, z) := \frac{1}{2}\kappa\frac{|z|^2}{y}.$$

Then $(Y^\xi, Z^\xi), (Y^\eta, Z^\eta)$ are the solutions of the BSDEs generated by $(f, \xi), (f, \eta)$, respectively. To simplify the notation, we define

$$\Delta Y = Y^\xi - Y^\eta, \quad \Delta Z = Z^\xi - Z^\eta.$$

Also, we define

$$(Y_t^\lambda, Z_t^\lambda) = \left(\lambda Y_t^\xi + (1-\lambda)Y_t^\eta, \lambda Z_t^\xi + (1-\lambda)Z_t^\eta\right)$$

for every $0 \leq t \leq T$. Then by the Mean Value Theorem,

$$
\begin{aligned}
& f(Y_t^\xi, Z_t^\xi) - f(Y_t^\eta, Z_t^\eta) \\
&= \left(\int_0^1 \frac{\partial f}{\partial y}(Y_t^\lambda, Z_t^\lambda) d\lambda\right) \left(Y_t^\xi - Y_t^\eta\right) + \left(\int_0^1 \frac{\partial f}{\partial z}(Y_t^\lambda, Z_t^\lambda) d\lambda\right) \left(Z_t^\xi - Z_t^\eta\right).
\end{aligned}
$$

Let $\left(Y^\xi - Y^\eta\right)^+ = \max\left(Y^\xi - Y^\eta, 0\right)$, then by the Tanaka's formula,

$$
\begin{aligned}
d\left(Y_t^\xi - Y_t^\eta\right)^+ &= \mathbb{1}_{\{\Delta Y_t \geq 0\}} dY_t - dK_t \\
&= -\mathbb{1}_{\{\Delta Y_t \geq 0\}} \left(f(Y_t^\xi, Z_t^\xi) - f(Y_t^\eta, Z_t^\eta)\right) dt \\
&\quad + \mathbb{1}_{\{\Delta Y_t \geq 0\}} \left(Z_t^\xi - Z_t^\eta\right) dB_t - dK_t,
\end{aligned}
$$

where $K_t \geq 0$ is a non-decreasing process such that it increases only when $Y_t^\xi - Y_t^\eta \leq 0$. Therefore,

$$
\int_t^T \Delta Y_s dK_s = 0
$$

for every $0 \leq t \leq T$.

Apply the Itô's formula to the process $\left((Y^\xi - Y^\eta)^+\right)^p$ for some $p \geq 2$, then

$$
\begin{aligned}
& \left((\Delta Y_t)^+\right)^p + \frac{1}{2} p(p-1) \int_t^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} |\Delta Z_s|^2 ds \\
&= \left((\xi - \eta)^+\right)^p - p \int_t^T \left((\Delta Y_s)^+\right)^{p-1} \Delta Z_s dB_s \\
&\quad + p \int_t^T \left((\Delta Y_s)^+\right)^{p-1} \left(f(Y_s^\xi, Z_s^\xi) - f(Y_s^\eta, Z_s^\eta)\right) ds \\
&= \left((\xi - \eta)^+\right)^p - p \int_t^T \left((\Delta Y_s)^+\right)^{p-1} \Delta Z_s dB_s \\
&\quad + p \int_t^T \left((\Delta Y_s)^+\right)^{p-1} \left(\int_0^1 \frac{\partial f}{\partial y}(Y_s^\lambda, Z_s^\lambda) d\lambda\right) \Delta Y_s ds \\
&\quad + p \int_t^T \left((\Delta Y_s)^+\right)^{p-1} \left(\int_0^1 \frac{\partial f}{\partial z}(Y_s^\lambda, Z_s^\lambda) d\lambda\right) \Delta Z_s ds \\
&\leq \left((\xi - \eta)^+\right)^p - p \int_t^T \left((\Delta Y_s)^+\right)^{p-1} \Delta Z_s dB_s
\end{aligned}
$$

$$
\begin{aligned}
&+ p\int_t^T \left((\Delta Y_s)^+\right)^p \left(\int_0^1 \frac{\partial f}{\partial y}(Y_s^\lambda, Z_s^\lambda) + \epsilon\Big|\frac{\partial f}{\partial z}(Y_s^\lambda, Z_s^\lambda)\Big|^2 d\lambda\right) ds \\
&+ \frac{p}{4\epsilon}\int_t^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} \left|\Delta Z_s\right|^2 ds,
\end{aligned} \tag{4.19}
$$

where the inequality comes from the fact that $ab \leq \epsilon a^2 + \frac{1}{4\epsilon}b^2$ for any $\epsilon > 0$. Choose $\epsilon = \frac{1}{2\kappa} > 0$, so that

$$
\frac{\partial f}{\partial y} + \epsilon \left|\frac{\partial f}{\partial z}\right|^2 = 0.
$$

Then (4.19) is read as

$$
\begin{aligned}
&\left((\Delta Y_t)^+\right)^p + \frac{1}{2}p(p-1)\int_t^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} |\Delta Z_s|^2 ds \\
\leq& \left((\xi - \eta)^+\right)^p - p\int_t^T \left((\Delta Y_s)^+\right)^{p-1} \Delta Z_s dB_s \\
&+ \frac{p}{4\epsilon}\int_t^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} |\Delta Z_s|^2 ds.
\end{aligned} \tag{4.20}
$$

Taking expectations on both sides with $t = 0$ and choosing $p > \max(1+\kappa, 2)$, we obtain that

$$
\mathbb{E}\left[\int_0^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} |\Delta Z_s|^2 ds\right] \leq \frac{2}{p(p-1-\kappa)}\mathbb{E}\left[\left((\xi - \eta)^+\right)^p\right] \tag{4.21}
$$

By the Burkholder-Davis-Gundy's inequality and the inequality $ab \leq \beta a^2 + \frac{b^2}{4\beta}$, it yields for any $\beta > 0$ that

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0\leq t\leq T}\left(-p\int_t^T \left((\Delta Y_s)^+\right)^{p-1} \Delta Z_s dB_s\right)\right] \\
=& p\mathbb{E}\left[\sup_{0\leq t\leq T}\int_0^t \left((\Delta Y_s)^+\right)^{p-1} \Delta Z_s dB_s\right] \\
\leq& 4p\mathbb{E}\left[\left(\int_0^T \left((\Delta Y_s)^+\right)^{2p-2} |\Delta Z_s|^2 ds\right)^{1/2}\right] \\
\leq& 4p\mathbb{E}\left[\sup_{0\leq t\leq T}\left((\Delta Y_t)^+\right)^{p/2}\left(\int_0^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} |\Delta Z_s|^2 ds\right)^{1/2}\right]
\end{aligned}
$$

$$\leq \beta p \mathbb{E}\left[\sup_{0\leq t\leq T}\left((\Delta Y_t)^+\right)^p\right] + \frac{p}{\beta}\mathbb{E}\left[\int_0^T \left((\Delta Y_s)^+\right)^{p-2}|\Delta Z_s|^2\,ds\right].$$

So taking expectations of the supremum on both sides of (4.20) gives

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\leq t\leq T}\left((\Delta Y_t)^+\right)^p\right] + \frac{1}{2}p(p-1)\mathbb{E}\left[\sup_{0\leq t\leq T}\int_t^T \left((\Delta Y_s)^+\right)^{p-2}|\Delta Z_s|^2\,ds\right]\\
&\leq \mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right] + \mathbb{E}\left[\sup_{0\leq t\leq T}\left(-p\int_t^T \mathbb{1}_{\{\Delta Y_s\geq 0\}}\left((\Delta Y_s)^+\right)^{p-1}\Delta Z_s dB_s\right)\right]\\
&\quad + \frac{p}{4\epsilon}\mathbb{E}\left[\sup_{0\leq t\leq T}\int_t^T \mathbb{1}_{\{\Delta Y_s\geq 0\}}\left((\Delta Y_s)^+\right)^{p-2}|\Delta Z_s|^2\,ds\right]\\
&\leq \mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right] + \frac{\kappa p}{2}\mathbb{E}\left[\sup_{0\leq t\leq T}\int_t^T \mathbb{1}_{\{\Delta Y_s\geq 0\}}\left((\Delta Y_s)^+\right)^{p-2}|\Delta Z_s|^2\,ds\right]\\
&\quad + \beta p\mathbb{E}\left[\sup_{0\leq t\leq T}\left((\Delta Y_s)^+\right)^p\right] + \frac{p}{\beta}\mathbb{E}\left[\int_0^T \mathbb{1}_{\{\Delta Y_s\geq 0\}}\left((\Delta Y_s)^+\right)^{p-2}|\Delta Z_s|^2\,ds\right]\\
&= \mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right] + \beta p\mathbb{E}\left[\sup_{0\leq t\leq T}\left((\Delta Y_s)^+\right)^p\right]\\
&\quad + \left(\frac{\kappa p}{2} + \frac{p}{\beta}\right)\mathbb{E}\left[\int_0^T \mathbb{1}_{\{\Delta Y_s\geq 0\}}\left((\Delta Y_s)^+\right)^{p-2}|\Delta Z_s|^2\,ds\right].
\end{aligned}$$

Rearranging to get

$$\begin{aligned}
&(1-\beta p)\,\mathbb{E}\left[\sup_{0\leq t\leq T}\left((\Delta Y_t)^+\right)^p\right]\\
&\leq \mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right] + \frac{1}{2}p\left(\kappa + \frac{2}{\beta} + 1 - p\right)\\
&\quad \times \mathbb{E}\left[\int_0^T \mathbb{1}_{\{\Delta Y_s\geq 0\}}\left((\Delta Y_s)^+\right)^{p-2}|\Delta Z_s|^2\,ds\right]\\
&\leq \frac{2}{\beta(p-1-\kappa)}\mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right],
\end{aligned}$$

where the second inequality follows from (4.21). In particular, choose $\beta = \frac{1}{2p}$, we obtain the following inequality

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left(\left(Y_t^\xi - Y_t^\eta\right)^+\right)^p\right] \leq \frac{8p}{p-1-k}\mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right]. \tag{4.22}$$

By symmetry, we also have

$$\mathbb{E}\left[\sup_{0\le t\le T}\left(\left(Y_t^\eta - Y_t^\xi\right)^+\right)^p\right] \le \frac{8p}{p-1-k}\mathbb{E}\left[\left((\eta-\xi)^+\right)^p\right]. \tag{4.23}$$

Hence, with the identity

$$\left|Y_t^\xi - Y_t^\eta\right|^p = \left(\left(Y_t^\xi - Y_t^\eta\right)^+\right)^p + \left(\left(Y_t^\eta - Y_t^\xi\right)^+\right)^p,$$

we may add (4.22), (4.23) together so that

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}\left|Y_t^\xi - Y_t^\eta\right|^p\right] &\le \mathbb{E}\left[\sup_{0\le t\le T}\left(\left(Y_t^\xi - Y_t^\eta\right)^+\right)^p\right] + \mathbb{E}\left[\sup_{0\le t\le T}\left(\left(Y_t^\eta - Y_t^\xi\right)^+\right)^p\right]\\ &\le \frac{16p}{p-1-k}\mathbb{E}\left[|\xi-\eta|^p\right].\end{aligned}$$

□

Theorem 4.9. *For any random variable $Y \in \mathscr{Y}$, and $p > \max(1+\kappa, 2)$, define the p-norm as*

$$\|Y\|_p = \mathbb{E}\left[|Y|^p\right]^{1/p}.$$

Suppose $\phi : \mathbb{R}^d \to \mathbb{R}$ is a bounded, Lipschitz function, that is, there exist some positive constants $c_1, c_2, C_L > 0$ such that

$$\begin{aligned}&c_1 \le \phi(x) \le c_2,\\ &\left|\phi(x^1) - \phi(x^2)\right| \le C_L\left|x^1 - x^2\right|\end{aligned}$$

for all $x, x^1, x^2 \in \mathbb{R}^d$. Define the non-linear mapping $\Phi : \mathscr{X} \to \mathscr{X}$ by

$$\Phi(\xi) = \phi\left(x + \int_0^T \sqrt{2Y_s^\xi}\,dB_s\right),$$

where

$$Y_s^\xi = \xi + \int_s^T \frac{\kappa}{2} \frac{\left|Z_r^\xi\right|^2}{Y_r^\xi} dr - \int_s^T Z_r^\xi dB_r s \tag{4.24}$$

and

$$T \leq \frac{p-1-\kappa}{16p} \left(\frac{\sqrt{c_1}(p-1)}{\sqrt{2} C_L p} \right)^p.$$

Then given any $\xi, \eta \in \mathscr{Y}$, it yields that

$$\|\Phi(\xi) - \Phi(\eta)\|_p \leq \frac{1}{2} \|\xi - \eta\|_p.$$

Moreover, the system of forward-backward stochastic differential equation (4.4) *admits a unique solution $(X, Y, Z) \in \mathscr{L}_T^2(\mathbb{R}^d) \times \mathscr{Y} \times \mathscr{H}_T^2(\mathbb{R}^d)$.*

Proof. Given any terminal data $\xi \in \mathscr{X}$, Lemma 4.7 shows that (4.24) is uniquely defined, and $Y^\xi \in \mathscr{Y}$. Thus, from the theory of SDE, the mapping Φ is well-defined.

Let $\xi, \eta \in \mathscr{X}$, define Y^ξ, Y^η via (4.24), respectively. Since ϕ is globally Lipschitz, then for any $p > \max(1+\kappa, 2)$, we have

$$\begin{aligned}
&\left| \phi\left(x + \int_0^T \sqrt{2Y_s^\xi} dB_s \right) - \phi\left(x + \int_0^T \sqrt{2Y_s^\eta} dB_s \right) \right|^p \\
&\leq \left| C_L \left(\int_0^T \sqrt{2Y_s^\xi} dB_s - \int_0^T \sqrt{2Y_s^\eta} dB_s \right) \right|^p \\
&\leq C_L^p \left| \frac{1}{\sqrt{2c_1}} \int_0^T \left| Y_s^\xi - Y_s^\eta \right| dB_s \right|^p \\
&= (2c_1)^{-p/2} C_L^p \left| \int_0^T \left| Y_s^\xi - Y_s^\eta \right| dB_s \right|^p,
\end{aligned}$$

where the second inequality follows from the Mean Value Theorem applied to the function $x \mapsto \sqrt{2x}$ on the interval $x \in [c_1, c_2]$. Taking expectations on both sides, it then follows from the Itô isometry that

$$
\begin{aligned}
&\mathbb{E}\left[\left|\phi\left(x+\int_0^T\sqrt{2Y_s^\xi}dB_s\right)-\phi\left(x+\int_0^T\sqrt{2Y_s^\eta}dB_s\right)\right|^p\right]\\
&\leq (2c_1)^{-p/2}C_L^p\mathbb{E}\left[\left|\int_0^T\left|Y_s^\xi-Y_s^\eta\right|dB_s\right|^p\right]\\
&\leq (2c_1)^{-p/2}C_L^p\left(\frac{p}{p-1}\right)^p\mathbb{E}\left[\left(\int_0^T\left|Y_s^\xi-Y_s^\eta\right|^2ds\right)^{p/2}\right]\\
&\leq (2c_1)^{-p/2}C_L^p\left(\frac{p}{p-1}\right)^p\mathbb{E}\left[\int_0^T\left|Y_s^\xi-Y_s^\eta\right|^pds\right]\\
&\leq (2c_1)^{-p/2}C_L^p\left(\frac{p}{p-1}\right)^pT\mathbb{E}\left[\sup_{0\leq s\leq T}\left|Y_s^\xi-Y_s^\eta\right|^p\right]\\
&\leq (2c_1)^{-p/2}C_L^p\left(\frac{p}{p-1}\right)^p\frac{16pT}{p-1-\kappa}\mathbb{E}\left[|\xi-\eta|^p\right],
\end{aligned}
$$

where the last inequality follows from (4.18) in Lemma 4.8. Choose

$$
T^*=\frac{p-1-\kappa}{16p}\left(\frac{\sqrt{c_1}(p-1)}{\sqrt{2}C_Lp}\right)^p>0
$$

so that

$$
(2c_1)^{-p/2}C_L^p\left(\frac{p}{p-1}\right)^p\frac{16pT^*}{p-1-\kappa}=\frac{1}{2^p}.
$$

So for any $T\leq T^*$, we have

$$
\left\|\phi\left(x+\int_0^T\sqrt{2Y_s^\xi}dB_s\right)-\phi\left(x+\int_0^T\sqrt{2Y_s^\eta}dB_s\right)\right\|_p\leq\frac{1}{2}\left\|\xi-\eta\right\|_p.
$$

This implies that the mapping Φ is a contraction on the space $\mathscr{X}$ when equipped with the p-norm. As a consequence, it has a unique fixed point, denoted by ξ^*. Then the corresponding triple $(X^{\xi^*},Y^{\xi^*},Z^{\xi^*})$ is the unique solution to the FBSDE (4.4), and $(X^{\xi^*},Y^{\xi^*},Z^{\xi^*})\in\mathscr{L}_T^2(\mathbb{R}^d)\times\mathscr{Y}\times\mathscr{H}_T^2(\mathbb{R}^d)$. □

Inspired by this decoupling method, we may extend the existence and uniqueness result to FBSDEs under more general settings. Particularly, our approach strongly relies on the solvability of the decoupled BSDE. Here we study the solvability of a class of FBSDEs of the form (4.2), whose decoupled backward equation satisfies the growth conditions described in [78].

Lemma 4.10. *Consider the following one-dimensional BSDE*

$$dY_t = -g(t, Y_t, Z_t)dt + Z_t dB_t, \quad Y_T = \xi, \tag{4.25}$$

where $(B_t)_{0\le t\le T}$ is a standard d-dimensional standard Brownian motion on some probability space, ξ is a bounded, measurable real-valued process, and the generator $g : [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ satisfies the growth conditions as described in Proposition 2.23. In addition, it is assumed that

$$\frac{\partial g}{\partial y}(t,y,z) + \epsilon \left|\frac{\partial g}{\partial z}\right|^2 (t,y,z) \le A(t)$$

for some constant $\epsilon > 0$ and $A : [0,T] \to \mathbb{R}$. If A is bounded on $[0,T]$, then for any bounded measurable terminal data ξ, η, the corresponding processes Y^ξ, Y^η satisfies

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|Y_t^\xi - Y_t^\eta\right|^p\right] \le C_{A,\epsilon,p} e^T \mathbb{E}\left[|\xi-\eta|^p\right], \tag{4.26}$$

where

$$C_{A,\epsilon,p} = \frac{32\epsilon p}{2\epsilon p - 2\epsilon - 1} \exp\left(\frac{16\epsilon p^2}{2\epsilon p - 2\epsilon - 1}\|A\|_{L^\infty([0,T])}\right).$$

depends only on ϵ, the bound of $A(t)$ on $[0,T]$, and the choice of p with $p > \max(1+\frac{1}{2\epsilon}, 2)$.

Proof. Given any bounded measurable process ξ, Proposition 2.23 implies that the backward stochastic differential equation (4.25) has at least one solution. Denote this solution by (Y,Z), then $(Y,Z) \in \mathscr{H}_T^\infty(\mathbb{R}) \times \mathscr{H}_T^2(\mathbb{R}^d)$.

Now suppose $(Y^\xi, Z^\xi), (Y^\eta, Z^\eta)$ are both the solutions to (4.25) with corresponding terminal data ξ, η, respectively. We argue similarly as in Lemma 4.8. Define

$$\Delta Y = Y^{\xi} - Y^{\eta}, \quad \Delta Z = Z^{\xi} - Z^{\eta},$$

and

$$(Y_t^{\lambda}, Z_t^{\lambda}) = \left(\lambda Y_t^{\xi} + (1-\lambda) Y_t^{\eta}, \lambda Z_t^{\xi} + (1-\lambda) Z_t^{\eta}\right)$$

for any $\lambda \in [0,1]$. Then we have

$$\begin{aligned}
&\left((\Delta Y_t)^+\right)^p + \frac{1}{2}p(p-1)\int_t^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} |\Delta Z_s|^2 \, ds \\
&\leq \left((\xi-\eta)^+\right)^p - p\int_t^T \left((\Delta Y_s)^+\right)^{p-1} \Delta Z_s dB_s \\
&\quad + p\int_t^T \left((\Delta Y_s)^+\right)^p \left(\int_0^1 \frac{\partial f}{\partial y}(Y_s^{\lambda}, Z_s^{\lambda}) + \epsilon \Big|\frac{\partial f}{\partial z}(Y_s^{\lambda}, Z_s^{\lambda})\Big|^2 d\lambda\right) ds \\
&\quad + \frac{p}{4\epsilon}\int_t^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} \big|\Delta Z_s\big|^2 ds \\
&\leq \left((\xi-\eta)^+\right)^p - p\int_t^T \left((\Delta Y_s)^+\right)^{p-1} \Delta Z_s dB_s \\
&\quad + p\int_t^T A(s)\left((\Delta Y_s)^+\right)^p ds + \frac{p}{4\epsilon}\int_t^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} \big|\Delta Z_s\big|^2 ds \\
&\leq \left((\xi-\eta)^+\right)^p - p\int_t^T \left((\Delta Y_s)^+\right)^{p-1} \Delta Z_s dB_s \\
&\quad + p\|A\|_{L^\infty([t,T])}\int_t^T \left((\Delta Y_s)^+\right)^p ds + \frac{p}{4\epsilon}\int_t^T \mathbb{1}_{\{\Delta Y_s \geq 0\}} \left((\Delta Y_s)^+\right)^{p-2} \big|\Delta Z_s\big|^2 ds
\end{aligned} \tag{4.27}$$

for any $0 \leq t \leq T$. For any $p > \max(1 + \frac{1}{2\epsilon}, 2)$, we take expectations on both side. Rearrange to get

$$\begin{aligned}
&\mathbb{E}\left[\int_t^T \left((\Delta Y_s)^+\right)^{p-2} |\Delta Z_s|^2 \, ds\right] \\
&\leq \frac{2}{p(p-1-\frac{1}{2\epsilon})}\mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right] + \frac{2\|A\|_{L^\infty([t,T])}}{p-1-\frac{1}{2\epsilon}}\mathbb{E}\left[\int_t^T \left((\Delta Y_s)^+\right)^p ds\right]
\end{aligned}$$

Similarly, with the Burkholder-Davis-Gundy inequality,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}\left(-p\int_t^T \left((\Delta Y_s)^+\right)^{p-1}\Delta Z_s dB_s\right)\right]\\
&\le \beta p\mathbb{E}\left[\sup_{0\le t\le T}\left((\Delta Y_t)^+\right)^p\right]+\frac{p}{\beta}\mathbb{E}\left[\int_0^T \left((\Delta Y_s)^+\right)^{p-2}\left|\Delta Z_s\right|^2 ds\right]
\end{aligned}$$

for any $\beta > 0$.

Hence, take expectations of the supremum of both sides of (4.27) and rearrange to give

$$\begin{aligned}
&(1-\beta p)\mathbb{E}\left[\sup_{0\le t\le T}\left((\Delta Y_t)^+\right)^p\right]\\
&\le \mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right]+p\|A\|_{L^\infty([0,T])}\mathbb{E}\left[\int_0^T\left((\Delta Y_s)^+\right)^p ds\right]\\
&\quad+\frac{1}{2}p\left(\frac{2}{\beta}+\frac{1}{2\epsilon}+1-p\right)\mathbb{E}\left[\int_0^T\left((\Delta Y_s)^+\right)^{p-2}|\Delta Z_s|^2 ds\right]\\
&\le \mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right]+p\|A\|_{L^\infty([0,T])}\mathbb{E}\left[\int_0^T\left((\Delta Y_s)^+\right)^p ds\right]\\
&\quad+\left(\frac{2}{\beta(p-1-\frac{1}{2\epsilon})}-1\right)\mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right]\\
&\quad+\left(\frac{2}{\beta(p-1-\frac{1}{2\epsilon})}-1\right)p\,\|A\|_{L^\infty[0,T]}\,\mathbb{E}\left[\int_0^T\left((\Delta Y_s)^+\right)^p ds\right]\\
&=\frac{2}{\beta(p-1-\frac{1}{2\epsilon})}\mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right]\\
&\quad+\frac{2}{\beta(p-1-\frac{1}{2\epsilon})}p\,\|A\|_{L^\infty[0,T]}\,\mathbb{E}\left[\int_0^T\left((\Delta Y_s)^+\right)^p ds\right]\\
&\le\frac{2}{\beta(p-1-\frac{1}{2\epsilon})}\mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right]\\
&\quad+\frac{2}{\beta(p-1-\frac{1}{2\epsilon})}p\,\|A\|_{L^\infty[0,T]}\,\mathbb{E}\left[\int_0^T\sup_{0\le u\le s}\left((\Delta Y_u)^+\right)^p ds\right]
\end{aligned}$$

for any $\beta > 0$. In particular, choose $\beta = \frac{1}{2p}$ for any $p > \max(1+\frac{1}{2\epsilon}, 2)$, then rearrange to get

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}\left(\left(Y_t^{\xi}-Y_t^{\eta}\right)^+\right)^p\right]\\
\le\ &\frac{8p}{p-1-\frac{1}{2\epsilon}}\mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right]\\
&+\frac{8p^2}{(p-1-\frac{1}{2\epsilon})}\|A\|_{L^\infty[0,T]}\,\mathbb{E}\left[\int_0^T\sup_{0\le u\le s}\left(\left(Y_u^{\xi}-Y_u^{\eta}\right)^+\right)^p ds\right].
\end{aligned}$$

So by the Grönwall's inequality, one obtains

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}\left(\left(Y_t^{\xi}-Y_t^{\eta}\right)^+\right)^p\right]\\
\le\ &\frac{8p}{p-1-\frac{1}{2\epsilon}}\exp\left(\frac{8p^2}{(p-1-\frac{1}{2\epsilon})}\|A\|_{L^\infty[0,T]}\right)\mathbb{E}\left[\left((\xi-\eta)^+\right)^p\right].
\end{aligned}$$

Hence, by the identity

$$\left|Y^{\xi}-Y^{\eta}\right|^p=\left(\left(Y^{\xi}-Y^{\eta}\right)^+\right)^p+\left(\left(Y^{\eta}-Y^{\xi}\right)^+\right)^p,$$

we conclude the desired inequality. □

Theorem 4.11. *Let b,σ,h be Lipschitz continuous functions such that there exist constant $C_1,C_2,C_3>0$ satisfying*

$$\begin{aligned}
|b(t,x^1,y^1)-b(t,x^2,y^2)|&\le C_1\left(|x^1-x^2|+|y^1-y^2|\right),\\
|\sigma(t,x^1,y^1)-\sigma(t,x^2,y^2)|&\le C_2\left(|x^1-x^2|+|y^1-y^2|\right),\\
|h(y^1)-h(y^2)|&\le C_3\left(|y^1-y^2|\right)
\end{aligned}$$

for any $0\le t\le T, x^1,x^2\in\mathbb{R}^d, y^1,y^2\in\mathbb{R}$. In addition, we assume h is bounded and $g\in\mathcal{C}([0,T]\times\mathbb{R}\times\mathbb{R}^d)$ satisfies the growth conditions described in Lemma

4.10. Then there exists a $T_L > 0$ with

$$T_L < \left(\frac{\sqrt{\left(\frac{4C_2 p}{p-1}\right)^{2p} + 4(4C_1)^p} - \left(\frac{4C_2 p}{p-1}\right)^p}{2(4C_1)^p} \right)^{2/p}$$

such that the following system of forward–backward stochastic differential equations

$$\begin{cases} dX_t = b(t, X_t, Y_t)dt + \sigma(t, X_t, Y_t)dB_t, & X_0 = x \\ dY_t = -g(t, Y_t, Z_t)dt + Z_t dB_t, & Y_T = h(X_T) \end{cases} \tag{4.28}$$

has a unique solution $(X, Y, Z) \in \mathscr{L}^2_{T_L}(\mathbb{R}^d) \times \mathscr{H}^\infty_{T_L}(\mathbb{R}) \times \mathscr{H}^2_{T_L}(\mathbb{R}^d)$ on the time duration $[0, T_L]$.

Proof. Choose $\eta = \xi$ in Lemma 4.10, then it yields that the decoupled backward stochastic differential equation

$$dY_t = -g(t, Y_t, Z_t)dt + Z_t dB_t, \quad Y_T = \xi \tag{4.29}$$

admits a unique solution for any $T > 0$, where ξ is $\mathscr{F}_T$-measurable and bounded process. We denote this solution by $(Y^\xi_t, Z^\xi_t)_{0 \le t \le T}$, then $(Y^\xi, Z^\xi) \in \mathscr{H}^\infty_T(\mathbb{R}) \times \mathscr{H}^2_T(\mathbb{R}^d)$.

Under the global Lipschitz condition of the coefficients, and that the process Y^ξ is adapted and bounded a.s., the stochastic differential equation

$$dX_t = b(t, X_t, Y^\xi_t)dt + \sigma(t, X_t, Y^\xi_t)dB_t, \quad X_0 = x \tag{4.30}$$

admits a unique strong solution that is square-integrable. We denote this solution by $(X^\xi_t)_{0 \le t \le T}$. Then whenever h is bounded, one may construct the following non-linear mapping Φ from the space of bounded random variables to itself such that

$$\Phi(\xi) = h\left(x + \int_0^T b(s, X_s^\xi, Y_s^\xi)ds + \int_0^T \sigma(s, X_s^\xi, Y_s^\xi)dB_s\right).$$

Now we aim to show this mapping is a contraction. So we need the following estimate.

Let ξ, η be bounded processes, and $(X^\xi, Y^\xi, Z^\xi), (X^\eta, Y^\eta, Z^\eta)$ be the corresponding processes generated by (4.29) and (4.30), respectively. Then by the global Lipschitz condition, we have

$$\begin{aligned}
\left|X_t^\xi - X_t^\eta\right| &= \int_0^t \left|f(s, X_s^\xi, Y_s^\xi) - f(s, X_s^\eta, Y_s^\eta)\right| ds \\
&\quad + \int_0^t \left|\sigma(s, X_s^\xi, Y_s^\xi) - \sigma(s, X_s^\eta, Y_s^\eta)\right| dB_s \\
&\le C_1 \int_0^t \left(\left|X_s^\xi - X_s^\eta\right| + \left|Y_s^\xi - Y_s^\eta\right|\right) ds \\
&\quad + C_2 \int_0^t \left(\left|X_s^\xi - X_s^\eta\right| + \left|Y_s^\xi - Y_s^\eta\right|\right) dB_s.
\end{aligned}$$

Take the p-th power on both sides with $p > \max(1+\frac{1}{2\epsilon}, 2)$, then with the inequality $|a+b|^p \le 2^p|a^p + b^p|$, it yields that

$$\begin{aligned}
\left|X_t^\xi - X_t^\eta\right|^p &\le 2^p \left(C_1 \int_0^t \left(\left|X_s^\xi - X_s^\eta\right| + \left|Y_s^\xi - Y_s^\eta\right|\right) ds\right)^p \\
&\quad + 2^p \left(C_2 \int_0^t \left(\left|X_s^\xi - X_s^\eta\right| + \left|Y_s^\xi - Y_s^\eta\right|\right) dB_s\right)^p.
\end{aligned}$$

Take expectations on the supremum over $0 \le t \le T$ on both sides, then by the Burkholder-Davis-Gundy inequality, we obtain that

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\le t\le T}\left|X_t^\xi - X_t^\eta\right|^p\right] \\
&\le (2C_1)^p \mathbb{E}\left[\sup_{0\le t\le T}\left(\int_0^t \left(\left|X_s^\xi - X_s^\eta\right| + \left|Y_s^\xi - Y_s^\eta\right|\right) ds\right)^p\right] \\
&\quad + (2C_2)^p \mathbb{E}\left[\sup_{0\le t\le T}\left(\int_0^t \left(\left|X_s^\xi - X_s^\eta\right| + \left|Y_s^\xi - Y_s^\eta\right|\right) dB_s\right)^p\right]
\end{aligned}$$

$$
\begin{aligned}
&\leq (2C_1)^p \mathbb{E}\left[\left(\int_0^T \left(\left|X_s^\xi - X_s^\eta\right| + \left|Y_s^\xi - Y_s^\eta\right|\right) ds\right)^p\right] \\
&\quad + (2C_2)^p \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[\left(\int_0^T 2\left(\left|X_s^\xi - X_s^\eta\right|^2 + \left|Y_s^\xi - Y_s^\eta\right|^2\right) ds\right)^{p/2}\right] \\
&\leq (4C_1T)^p \mathbb{E}\left[\sup_{0\leq t\leq T}\left|X_t^\xi - X_t^\eta\right|^p\right] + (4C_1T)^p \mathbb{E}\left[\sup_{0\leq t\leq T}\left|Y_t^\xi - Y_t^\eta\right|^p\right] \\
&\quad + \left(\frac{4C_2p\sqrt{T}}{p-1}\right)^p \mathbb{E}\left[\sup_{0\leq t\leq T}\left|X_t^\xi - X_t^\eta\right|^p\right] + \left(\frac{4C_2p\sqrt{T}}{p-1}\right)^p \mathbb{E}\left[\sup_{0\leq t\leq T}\left|Y_t^\xi - Y_t^\eta\right|^p\right].
\end{aligned}
$$

So rearrange to get

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0\leq t\leq T}\left|X_t^\xi - X_t^\eta\right|^p\right] \\
&\leq \frac{(4C_1T)^p + \left(\frac{4C_2p\sqrt{T}}{p-1}\right)^p}{1-(4C_1T)^p - \left(\frac{4C_2p\sqrt{T}}{p-1}\right)^p} \mathbb{E}\left[\sup_{0\leq t\leq T}\left|Y_t^\xi - Y_t^\eta\right|^p\right] \\
&\leq \frac{(4C_1T)^p + \left(\frac{4C_2p\sqrt{T}}{p-1}\right)^p}{1-(4C_1T)^p - \left(\frac{4C_2p\sqrt{T}}{p-1}\right)^p} C_{A,\epsilon,p} e^T \mathbb{E}\left[|\xi-\eta|^p\right],
\end{aligned}
\tag{4.31}
$$

where the last inequality follows from the inequality (4.26).

Now we are ready to show the mapping Φ is a contraction under the p-norm, where $p > \max(1+\frac{1}{2\epsilon}, 2)$. Again, by the Lipschitz assumption, we have

$$
\begin{aligned}
&|\Phi(\xi) - \Phi(\eta)|^p \\
&= \left| h\left(x + \int_0^T b(s, X_s^\xi, Y_s^\xi)ds + \int_0^T \sigma(s, X_s^\xi, Y_s^\xi)dB_s\right)\right. \\
&\quad \left. -h\left(x + \int_0^T b(s, X_s^\eta, Y_s^\eta)ds + \int_0^T \sigma(s, X_s^\eta, Y_s^\eta)dB_s\right)\right|^p \\
&\leq \left| C_3 \int_0^T \left|b(s, X_s^\xi, Y_s^\xi) - b(s, X_s^\eta, Y_s^\eta)\right| ds\right. \\
&\quad \left. +C_3 \int_0^T \left|\sigma(s, X_s^\xi, Y_s^\xi) - \sigma(s, X_s^\eta, Y_s^\eta)\right| dB_s\right|^p \\
&\leq 2^p \left| C_1C_3 \int_0^T \left(\left|X_s^\xi - X_s^\eta\right| + \left|Y_s^\xi - Y_s^\eta\right|\right) ds\right|^p
\end{aligned}
$$

$$+ 2^p \left| C_2 C_3 \int_0^T \left(\left| X_s^\xi - X_s^\eta \right| + \left| Y_s^\xi - Y_s^\eta \right| \right) dB_s \right|^p,$$

where the last inequality comes from the inequality $|a+b|^p \leq 2^p |a^p + b^p|$. Since the process $X^\xi - X^\eta$ is square-integrable and $Y^\xi - Y^\eta$ is bounded a.s., then the stochastic integral $\int_0^T \left(\left| X_s^\xi - X_s^\eta \right| + \left| Y_s^\xi - Y_s^\eta \right| \right) dB_s$ is a martingale. So taking expectations on both sides to obtain

$$\begin{aligned}
&\mathbb{E}\left[|\Phi(\xi) - \Phi(\eta)|^p \right] \\
&\leq (2C_1C_3)^p \mathbb{E}\left[\left| \int_0^T \left(\left| X_s^\xi - X_s^\eta \right| + \left| Y_s^\xi - Y_s^\eta \right| \right) ds \right|^p \right] \\
&\leq (2C_1C_3T)^p \mathbb{E}\left[\left| \sup_{0\leq t\leq T} \left| X_t^\xi - X_t^\eta \right| + \sup_{0\leq t\leq T} \left| Y_t^\xi - Y_t^\eta \right| \right|^p \right] \\
&\leq (4C_1C_3T)^p \left(\mathbb{E}\left[\sup_{0\leq t\leq T} \left| X_t^\xi - X_t^\eta \right|^p \right] + \mathbb{E}\left[\sup_{0\leq t\leq T} \left| Y_t^\xi - Y_t^\eta \right|^p \right] \right) \\
&\leq (4C_1C_3T)^p \frac{(4C_1T)^p + \left(\frac{4C_2p\sqrt{T}}{p-1} \right)^p}{1 - (4C_1T)^p - \left(\frac{4C_2p\sqrt{T}}{p-1} \right)^p} C_{A,\epsilon,p} e^T \mathbb{E}\left[|\xi - \eta|^p \right] \\
&\quad + (4C_1C_3T)^p C_{A,\epsilon,p} e^T \mathbb{E}\left[|\xi - \eta|^p \right] \\
&= \frac{C_{A,\epsilon,p} (4C_1C_3T)^p e^T}{1 - (4C_1T)^p - \left(\frac{4C_2p\sqrt{T}}{p-1} \right)^p} \mathbb{E}\left[|\xi - \eta|^p \right],
\end{aligned}$$

where the last inequality follows from (4.31) and (4.26). To simplify the notation, we define the function f to be

$$f(T) = \frac{C_{A,\epsilon,p} (4C_1C_3T)^p e^T}{1 - (4C_1T)^p - \left(\frac{4C_2p\sqrt{T}}{p-1} \right)^p}.$$

Also, define

$$T^* := \left(\frac{\sqrt{\left(\frac{4C_2p}{p-1} \right)^{2p} + 4(4C_1)^p} - \left(\frac{4C_2p}{p-1} \right)^p}{2(4C_1)^p} \right)^{2/p}.$$

Since f is continuous with $\lim_{T\to 0} f(T) = 0$ and $\lim_{T\to T^*} f(T) = \infty$, then we can find a T_L between 0 and T^* such that

$$\frac{C_{A,\epsilon,p}(4C_1C_3T_L)^p e^T}{1-(4C_1T_L)^p - \left(\frac{4C_2p\sqrt{T_L}}{p-1}\right)^p} = \frac{1}{2^p}.$$

If $f(T_L) = \frac{1}{2}$ has more than one root, then we choose the smallest one. Therefore, we have

$$\mathbb{E}\left[|\Phi(\xi) - \Phi(\eta)|^p\right] \leq \frac{1}{2}\mathbb{E}\left[|\xi - \eta|^p\right].$$

This implies that the mapping Φ is a contraction, so it has a unique fixed point. In this case, the system of forward–backward stochastic differential equation (4.28) has a unique solution on the time duration $[0, T_L]$. $\square$

A Simulation of velocity fields

In this section, we include the `Python` code for simulating the velocity field with repetitions so that the reader may repeat the experiment.

Listing 1 lists all the packages we have used. Mainly, the package `plotly` is applied to plot the 3D stream tubes of the velocity fields and the package `matplotlib` is used for the stream lines in the 2D projection.

```
import numpy as np
import matplotlib.pyplot as plt
import math
import time
import plotly.graph_objects as go
```

Listing 1: Packages in use

We enumerate the important parameters in Listing 2 and record some functions for facilitating further calculations in Listing 3.

```
# M is the size of the mesh and h is the spacing. These two
    concepts are explicitly defined in Chapter 3.
M = 20
h = 0.5
# For each given initial value xj, we use the Euler scheme to
    simulate the particle paths.
# Here T is the entire time interval and N is the number of time
    steps.
T = 0.5
N = 5
dt = T/N
# v is the viscosity constant
v = 1
delta = 1e-6
# This path_number is the number of repetitions explained in the
```

```
    modified approach.
path_number = 10
```

Listing 2: Initial data

```
# This is the kernel function after removing singularities
def G_delta(x):
    x = np.array(x)
    sumSquare = sum(x ** 2)
    if sumSquare < delta:
        sumSquare = delta
    return 1/(2 * math.pi * sumSquare ** 1.5) * x

# This is the initial vorticity function as has been defined in
    Chapter 3.6.
def w0(x):
    x = list(x)
    if x in [[1, 1, 0], [0, 1, 1], [1, -1, 0], [-1, -1, 0]]:
        return np.array([0, 0, 10])
    elif x in [[0.5, 1, -1], [1.5, 0, 0], [-1, 0.5, 0], [0.5,
    -1.5, 0.5]]:
        return np.array([10, 20, 0])
    return np.zeros(3)

# This function here computes the cross product of two vectors a,
    b.
def CrossProduct(a, b):
    return np.array([a[1]*b[2] - a[2]*b[1], a[2]*b[0] - a[0]*b
    [2], a[0]*b[1] - a[1]*b[0]])

# With specifications on location, particle path and time, this
    function returns the velocity at (x,t_n).
def u_h(x,X,t_n):
    x = np.array(x)
```

```python
    res = np.zeros(3)
    for j in j_eff:
        res += CrossProduct(G_delta(x - X[t_n][j]), w0(x0[j])) * (h ** 3) * dt
    return res
```

Listing 3: Initial functions

As we have mentioned in Chapter 3.6, choosing the initial vorticity function to be a sum of Dirac delta function could substantially reduce the computational complexity. This is because the initial vorticity ω_0 only takes values on the lattice set $\Lambda_3^{h,M}$. Therefore, it is enough to pin down all the lattice points with non-zero vorticity. Listing 4 shows how we select the valid lattice points, that is, lattice points corresponding to non-zero vorticities.

```python
# Creating the lattice set.
x0 = np.mgrid[-M:M+0.1:h, -M:M+0.1:h, -M:M+0.1:h].reshape(3,-1).T

# Fiding the indices of valid lattice points whose corresponding vorticity is non-zero.
j_eff = []
for j in range(len(x0)):
    if np.any(w0(x0[j]) != np.zeros(3)):
        j_eff.append(j)
```

Listing 4: Indices of valid lattice points

With all the initial data and the self-defined functions, we are then able to simulate the particle trajectory paths (see Listing 5).

```python
# Creating particle trajectory paths with repetitions.
X = np.zeros(shape = (N+1, path_number, int(2*M/h+1) ** 3, 3))
for p in range(path_number):
    X[0][p] = x0
for t in range(N):
```

```
    dX = np.zeros(shape = (path_number, int(2*M/h+1)**3, 3))
    for n in range(path_number):
        for m in range(path_number):
            for i in range(len(x0)):
                for j in j_eff:
                    dX[n][i] += CrossProduct(G_delta(X[t][n][i] - X[t][m][j]), w0(x0[j])) * (h**3) * dt / path_number
        X[t+1][n] = X[t][n] + dX[n] + (2 * v) ** 0.5 * np.random.normal(0, dt**0.5, dX[n].shape)
# Particle paths with the same starting lattice points.
Y = X.mean(axis=1)
```

Listing 5: Simulating particle trajectory paths

Then the velocity at each point (x, t) can be computed from the function `u_h`. But note that the time t is restricted to be the time step defined in Listing 1. So the 3D velocity field is plotted via the function `Streamtube` in the package `plotly` (see Listing 6).

```
spacing = 0.05
n=2

x1, x2, x3 = np.meshgrid(np.arange(-2, 2, spacing), np.arange(-1, 1, spacing), np.arange(-0.5, 0.5, 0.1), indexing='ij')
x1 = x1.flatten()
x2 = x2.flatten()
x3 = x3.flatten()
# Creating the velocity field.
u1 = np.zeros(len(x1))
u2 = np.zeros(len(x1))
u3 = np.zeros(len(x1))
for i in range(len(x1)):
    velocity = u_h([x1[i], x2[i], x3[i]], Y, n)
    u1[i] = velocity[0]
```

```
    u2[i] = velocity[1]
    u3[i] = velocity[2]
fig = go.Figure(data=go.Streamtube(x=x1, y=x2, z=x3, u=u1, v=u2, w=u3,
                sizeref = 1.5,
                colorscale = 'Portland',
                showscale = True,
                maxdisplayed = 3000
               ))
fig.update_layout(
    scene = dict(
        aspectratio = dict(
            x = 2,
            y = 1,
            z = 0.3
        )
    ),
    margin = dict(
        t = 20,
        b = 20,
        l = 20,
        r = 20
    )
)
fig.show()
```

Listing 6: 3D plot of the velocity field

To implement the `Streamtube` function, one needs to specify a starting position. Here we have used the default position which is the plane $y = -1$. But one can use `starts` to determine the stream tubes' starting position.

For the 2D plot, we have used the function `streamplot` from the package `mathplotlib`. Then the velocity field on the plane $z = 0$ is computed at each step. Finally, the projection of the velocity field on the plane is plotted, that is, we plot the velocity field $(\tilde{u}^1((x_1, x_2, 0)), \tilde{u}^2((x_1, x_2, 0)))$ in the 2D plot (see Listing 7).

```
spacing = 0.5

# Creating the velocity field for all the time steps.
vf_proj1 = []
vf_proj2 = []
vf_proj3 = []
vf_abs = []
for n in range(N+1):
    x_axis_2d = np.arange(-2, 2.1, spacing)
    y_axis_2d = np.arange(-2, 2.1, spacing)
    X_axis_2d, Y_axis_2d = np.meshgrid(x_axis_2d, y_axis_2d, indexing='ij')
    vf_proj1_n = np.zeros(X_axis_2d.shape)
    vf_proj2_n = np.zeros(X_axis_2d.shape)
    vf_proj3_n = np.zeros(X_axis_2d.shape)
    vf_abs_n = np.zeros(X_axis_2d.shape)
    for i in range(len(x_axis_2d)):
        for j in range(len(y_axis_2d)):
            vf_proj1_n[i][j], vf_proj2_n[i][j], vf_proj3_n[i][j] = u_h([x_axis_2d[i], y_axis_2d[j], 0], Y, n)
            vf_abs_n[i][j] = np.sqrt(vf_proj1_n[i][j] ** 2 + vf_proj2_n[i][j] ** 2 + vf_proj3_n[i][j] ** 2)
    vf_proj1.append(vf_proj1_n)
    vf_proj2.append(vf_proj2_n)
    vf_proj3.append(vf_proj3_n)
    vf_abs.append(vf_abs_n)

# Plotting the 2D stream lines in for the n-th time step.
n = 2
```

```
fig = plt.figure(figsize=(15, 8))
strm = plt.streamplot(x_axis_2d, y_axis_2d, vf_proj1[n].T,
    vf_proj2[n].T, color=vf_abs[n], linewidth=1, cmap='rainbow')
fig.colorbar(strm.lines)
```

Listing 7: 2D projection plot of the velocity field

We have chosen the plane $z = 0$ for illustration, but one may vary the projection plane for different perspectives.

References

[1] F. Antonelli. Backward–forward stochastic differential equations. *The Annals of Applied Probability*, 3(3):777–793, 1993.

[2] F. Antonelli and S. Hamadène. Existence of the solutions of backward–forward SDE's with continuous monotone coefficients. *Statistics & probability letters*, 76(14):1559–1569, 2006.

[3] D. G. Aronson. The fundamental solution of a linear parabolic equation containing a small parameter. *Illinois Journal of Mathematics*, (4):580–619, 1959.

[4] D. G. Aronson. Uniqueness of positive weak solutions of second order parabolic equations. In *Annales Polonici Mathematici*, volume 3, pages 285–303, 1965.

[5] D. G. Aronson. Bounds for the fundamental solution of a parabolic equation. *Bulletin of the American Mathematical society*, 73(6):890–896, 1967.

[6] D. G. Aronson. Non-negative solutions of linear parabolic equations. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 22(4):607–694, 1968.

[7] D. G. Aronson. The porous medium equation. In *Nonlinear diffusion problems*, pages 1–46. Springer, 1986.

[8] D. G. Aronson and L. A. Caffarelli. The initial trace of a solution of the porous medium equation. *Transactions of the American Mathematical Society*, 280(1):351–366, 1983.

[9] M. Bauer and T. Meyer-Brandis. Existence and regularity of solutions to multi-dimensional mean-field stochastic differential equations with irregular drift. *arXiv preprint arXiv:1912.05932*, 2019.

[10] M. Bauer, T. Meyer-Brandis, and F. Proske. Strong solutions of mean-field stochastic differential equations with irregular drift. *Electronic Journal of Probability*, 23, 2018.

[11] J. T. Beale and A. Majda. Vortex methods I. convergence in three dimensions. *Mathematics of computation*, 39(159):1–27, 1982.

[12] J. T. Beale and A. Majda. Vortex methods II. higher order accuracy in two and three dimensions. *Mathematics of Computation*, 39(159):29–52, 1982.

[13] P. Bénilan, M. G. Crandall, and M. Pierre. Solutions of the porous medium equation in $\mathbb{R}^n$ under optimal conditions on initial values. *Indiana University mathematics journal*, 33(1):51–87, 1984.

[14] S. Bernstein. Principes de la théorie des équations différentielles stochastiques. *Trudy Fiz.-Mat., Steklov Inst., Akad. Nauk*, 5:95–124, 1934.

[15] J. M. Bismut. Conjugate convex functions in optimal stochastic control. *Journal of Mathematical Analysis and Applications*, 44(2):384–404, 1973.

[16] P. Briand and Y. Hu. Stability of BSDEs with random terminal time and homogenization of semilinear elliptic PDEs. *Journal of Functional Analysis*, 155(2):455–494, 1998.

[17] P. Briand and Y. Hu. BSDE with quadratic growth and unbounded terminal value. *Probability Theory and Related Fields*, 136(4):604–618, 2006.

[18] 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.

[19] P. Briand and A. Richou. On the uniqueness of solutions to quadratic BSDEs with non-convex generators. In *International Symposium on BSDEs*, pages 89–107. Springer, 2017.

[20] R. Buckdahn, B. Djehiche, J. Li, and S. Peng. Mean-field backward stochastic differential equations: a limit approach. *The Annals of Probability*, 37(4):1524–1565, 2009.

[21] P. Cardaliaguet, F. Delarue, J. M. Lasry, and P. L. Lions. *The Master Equation and the Convergence Problem in Mean Field Games:(AMS-201)*, volume 201. Princeton University Press, 2019.

[22] R. Carmona and F. Delarue. Mean field forward-backward stochastic differential equations. *Electronic Communications in Probability*, 18:1–15, 2013.

[23] R. Carmona and F. Delarue. Probabilistic analysis of mean-field games. *SIAM Journal on Control and Optimization*, 51(4):2705–2734, 2013.

[24] R. Carmona and F. Delarue. *Probabilistic Theory of Mean Field Games with Applications I-II.* Springer, 2018.

[25] P. Cheridito and K. Nam. Multidimensional quadratic and subquadratic BSDEs with special structure. *Stochastics An International Journal of Probability and Stochastic Processes*, 87(5):871–884, 2015.

[26] T. S. Chiang. McKean–Vlasov equations with discontinuous coefficients. *Soochow J. Math*, 20(4):507–526, 1994.

[27] A. J. Chorin. Numerical study of slightly viscous flow. *J. Fluid Mech*, 57(4):785–796, 1973.

[28] F. Confortola and P. Briand. Quadratic BSDEs with random terminal time and elliptic PDEs in infinite dimension. *Electronic Journal of Probability*, 13:1529–1561, 2008.

[29] P. Constantin and C. Foias. *Navier–Stokes Equations*. University of Chicago Press, 1988.

[30] G. H. Cottet and P. D. Koumoutsakos. *Vortex Methods: Theory and Practice*, volume 8. Cambridge University Press, Cambridge, 2000.

[31] B. E. J. Dahlberg and C. E. Kenig. Non-negative solutions of the porous medium equation. *Communications in partial differential equations*, 9(5):409–437, 1984.

[32] R. W. R. Darling and E. Pardoux. Backwards SDE with random terminal time and applications to semilinear elliptic PDE. *The Annals of Probability*, 25(3):1135–1159, 1997.

[33] M. E De Giorgi. Sulla differenziabilitae l'analiticita delle estremali degli integrali multipli regolari. *Ennio De Giorgi*, page 167, 1957.

[34] P. E. C. de Raynal. Strong well posedness of McKean–Vlasov stochastic differential equations with Hölder drift. *Stochastic Processes and their Applications*, 130(1):79–107, 2020.

[35] F. Delarue. On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case. *Stochastic processes and their applications*, 99(2):209–286, 2002.

[36] F. Delbaen, Y. Hu, and X. Bao. Backward SDEs with superquadratic growth. *Probability Theory and Related Fields*, 150(1):145–192, 2011.

[37] F. Delbaen, Y. Hu, and A. Richou. On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions. In *Annales de l'IHP Probabilités et statistiques*, volume 47, pages 559–574, 2011.

[38] C. Doléans-Dade. On the existence and unicity of solutions of stochastic integral equations. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 36(2):93–101, 1976.

[39] F. G. Dressel. The fundamental solution of the parabolic equation. *Duke Mathematical Journal*, 7(1):186–203, 1940.

[40] F. G. Dressel. The fundamental solution of the parabolic equation, ii. *Duke Mathematical Journal*, 13(1):61–70, 1946.

[41] N. El Karoui, S. Peng, and M.C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.

[42] S. N. Ethier and T. G. Kurtz. *Markov Processes: Characterization and Convergence*, volume 282. John Wiley & Sons, 2009.

[43] E. B. Fabes and D. W. Stroock. A new proof of Moser's parabolic Harnack inequality via the old ideas of Nash. *Arch. Rational Mech. Anal.*, 96:327–338, 1986.

[44] D. Fishelov. A new vortex scheme for viscous flows. *Journal of computational physics*, 86(1):211–224, 1990.

[45] J. Fontbona. A probabilistic interpretation and stochastic particle approximations of the 3-dimensional Navier–Stokes equations. *Probability Theory and Related Fields*, 136(1):102, 2006.

[46] A. Friedman. *Partial Differential Equations of Parabolic Type*. Courier Dover Publications, 2008.

[47] A. Fromm and P. Imkeller. Existence, uniqueness and regularity of decoupling fields to multidimensional fully coupled FBSDEs. *arXiv preprint arXiv:1310.0499*, 2013.

[48] T. Funaki. A certain class of diffusion processes associated with nonlinear parabolic equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 67(3):331–348, 1984.

[49] G. Galdi. *An Introduction to the Mathematical Theory of the Navier–Stokes Equations: Steady-state Problems*. Springer Science & Business Media, 2011.

[50] J. Gärtner. On the McKean–Vlasov limit for interacting diffusions. *Mathematische Nachrichten*, 137(1):197–248, 1988.

[51] M. Gevrey. Sur les équations aux dérivées partielles du type parabolique. *J. de Math.*, 10(6):105–148, 1913.

[52] I. I. Gihman. A method of constructing random processes (in russian). *Akad. Nauk SSSR*, 58:961–964, 1947.

[53] I. I. Gihman. Certain differential equations with random functions (in russian). *Uskr. Math. Z.*, 2(3):45–69, 1950.

[54] I. I. Gihman and A. V. Skorohod. *Stochastic Differential Equations*. Springer-Verlag, 1972.

[55] D. Gilbarg and N. S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag, Berlin, 2001.

[56] O. Guéant, J. M. Lasry, and P. L. Lions. Mean field games and applications. In *Paris-Princeton lectures on mathematical finance 2010*, pages 205–266. Springer, 2011.

[57] M. E. Gurtin and R. C. MacCamy. On the diffusion of biological populations. *Mathematical biosciences*, 33(1-2):35–49, 1977.

[58] K. E. Gustafson and J. A. Sethian. *Vortex Methods and Vortex Motion.* SIAM: Philadelphia, PA, USA, 1991.

[59] A. Gut. *Probability: A Graduate Course.* Springer, 2005.

[60] S. Hamadene. Backward-forward SDE's and stochastic differential games. *Stochastic processes and their applications*, 77(1):1–15, 1998.

[61] W. R. P. Hammersley, D. Šiška, and Ł. Szpruch. McKean–Vlasov SDEs under measure dependent Lyapunov conditions. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 57, pages 1032–1057. Institut Henri Poincaré, 2021.

[62] U. Horst, Y. Hu, P. Imkeller, A. Réveillac, and J. Zhang. Forward–backward systems for expected utility maximization. *Stochastic Processes and their Applications*, 124(5):1813–1848, 2014.

[63] Y. Hu and S. Peng. Solution of forward-backward stochastic differential equations. *Probability Theory and Related Fields*, 103(2):273–283, 1995.

[64] M. Huang, R. P. Malhamé, and P. E. Caines. Large population stochastic dynamic games: closed-loop McKean–Vlasov systems and the nash certainty

equivalence principle. *Communications in Information & Systems*, 6(3):221–252, 2006.

[65] N. Ikeda and S. Watanabe. *Stochastic Differential Equations and Diffusion Processes*. Elsevier, 2014.

[66] A. M. Il'in, A. S. Kalashnikov, and O. A. Oleinik. Linear equations of the second order of parabolic type. *Uspekhi Matematicheskikh Nauk*, 17(3):1–143, 1962.

[67] K. Itô. Differential equations determining a markoff process (in japanese). *J. Pan-Japan Math. Coll.*, 1077:1352–1400, 1942.

[68] K. Itô. On a stochastic integral equation. *Proceedings of the Japan Academy*, 22(1-4):32–35, 1946.

[69] K. Itô. *On stochastic differential equations*. Number 4. American Mathematical Soc., 1951.

[70] A. Jamneshan, M. Kupper, and P. Luo. Multidimensional quadratic BSDEs with separated generators. *Electronic Communications in Probability*, 22, 2017.

[71] B. Jourdain. Diffusions with a nonlinear irregular drift coefficient and probabilistic interpretation of generalized burgers' equations. *ESAIM: Probability and Statistics*, 1:339–355, 1997.

[72] B. Jourdain, S. Méléard, and W. A. Woyczynski. Nonlinear SDEs driven by Lévy processes and related PDEs. *ALEA Lat. Am. J. Probab.*, 4:1–29, 2008.

[73] M. Kac. *Probability and Related Topics in Physical Sciences*, volume 1. American Mathematical Soc., 1959.

[74] M. Kac. *Foundations of kinetic theory*, pages 173–200. University of California Press, 2020.

[75] I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*, volume 113. Springer Science & Business Media, 2012.

[76] N. Kazamaki. On a problem of Girsanov. *Tohoku Mathematical Journal*, 29(4):597–600, 1977.

[77] D. G. Kendall. Pole-seeking brownian motion and bird navigation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(3):365–402, 1974.

[78] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of probability*, 28(2):558–602, 2000.

[79] N. V. Krylov. On ito's stochastic integral equations. *Theory of Probability & Its Applications*, 14(2):330–336, 1969.

[80] N. V. Krylov. *Controlled Diffusion Processes*, volume 14. Springer–Verlag, Berlin, 2009.

[81] M. Kupper, P. Luo, and L. Tangpi. Multidimensional Markovian FBSDEs with super-quadratic growth. *Stochastic Processes and their Applications*, 129(3):902–923, 2019.

[82] O. A. Ladyzhenskaia, V. A. Solonnikov, and N. N. Ural'tseva. *Linear and Quasi-linear Equations of Parabolic Type*, volume 23. American Mathematical Soc., 1968.

[83] J. M. Lasry and P. L. Lions. Mean field games. *Japanese journal of mathematics*, 2(1):229–260, 2007.

[84] A. Leonard. Vortex methods for flow simulation. *Journal of Computational Physics*, 37(3):289–335, 1980.

[85] J. P. Lepeltier and J. San Martín. Backward stochastic differential equations with continuous coefficient. *Statistics & Probability Letters*, 32(4):425–430, 1997.

[86] J. P. Lepeltier and J. San Martín. Existence for BSDE with superlinear–quadratic coefficient. *Stochastics: An International Journal of Probability and Stochastic Processes*, 63(3-4):227–240, 1998.

[87] J. Leray. Sur le mouvement d'un liquide visqueux emplissant l'espace. *Acta mathematica*, 63(1):193–248, 1934.

[88] E. E. Levi. Sulle equazioni lineari totalmente ellittiche alle derivate parziali. *Rend. del. Circ. Mat. Palermo*, 24(1):275–317, 1907.

[89] J. Li and H. Min. Weak solutions of mean-field stochastic differential equations and application to zero-sum stochastic differential games. *SIAM Journal on Control and Optimization*, 54(3):1826–1858, 2016.

[90] D. G. Long. Convergence of the random vortex method in two dimensions. *Journal of the American Mathematical Society*, 1(4):779–804, 1988.

[91] P. Luo. A type of globally solvable BSDEs with triangularly quadratic generators. *Electronic Journal of Probability*, 25, 2020.

[92] P. Luo and L. Tangpi. Solvability of coupled FBSDEs with diagonally quadratic generators. *Stochastics and Dynamics*, 17(06):1750043, 2017.

[93] J. Ma, P. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitly – a four step scheme. *Probability Theory and Related Fields*, 98(3):339–359, 1994.

[94] J. Ma, Z. Wu, D. Zhang, and J. Zhang. On well-posedness of forward-backward SDEs—a unified approach. *Annals of Applied Probability*, 25(4):2168–2214, 2015.

[95] A. J. Majda, A. L. Bertozzi, and A. Ogawa. Vorticity and incompressible flow. Cambridge texts in applied mathematics. *Appl. Mech. Rev.*, 55(4):B77–B78, 2002.

[96] C. Marchioro and M. Pulvirenti. Hydrodynamics in two dimensions and vortex theory. *Communications in Mathematical Physics*, 84(4):483–503, 1982.

[97] H. P. McKean Jr. A class of markov processes associated with nonlinear parabolic equations. *Proceedings of the National Academy of Sciences of the United States of America*, 56(6):1907, 1966.

[98] S. Méléard. A trajectorial proof of the vortex method for the two-dimensional Navier–Stokes equation. *Annals of Applied Probability*, 10(4):1197–1211, 2000.

[99] C. Mimeau and I. Mortazavi. A review of vortex methods and their applications: From creation to recent advances. *Fluids*, 6(2), 2021.

[100] Y. Mishura and A. Veretennikov. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Theor. Probability and Math. Statist.*, 103:59–101, 2020.

[101] J. Moser. A Harnack inequality for parabolic differential equations. *Communications on pure and applied mathematics*, 17(1):101–134, 1964.

[102] J. Nash. Continuity of solutions of parabolic and elliptic equations. *American Journal of Mathematics*, 80(4):931–954, 1958.

[103] J. R. Norris and D. W. Stroock. Estimates on the fundamental solution to heat flows with uniformly elliptic coefficients. *Proceedings of the London Mathematical Society*, 3(2):373–402, 1991.

[104] A. A. Novikov. On an identity for stochastic integrals. *Theory of Probability & Its Applications*, 17(4):717–720, 1973.

[105] H. Osada. Propagation of chaos for the two dimensional Navier–Stokes equation. *Proceedings of the Japan Academy, Series A, Mathematical Sciences*, 62(1):8–11, 1986.

[106] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.

[107] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications*, pages 200–217. Springer, 1992.

[108] E. Pardoux and S. Tang. Forward-backward stochastic differential equations and quasilinear parabolic PDEs. *Probability Theory and Related Fields*, 114(2):123–150, 1999.

[109] S. Peng. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and stochastics reports (Print)*, 37(1-2):61–74, 1991.

[110] S. Peng and Z. Wu. Fully coupled forward-backward stochastic differential equations and applications to optimal control. *SIAM Journal on Control and Optimization*, 37(3):825–843, 1999.

[111] J. Pitman and M. Yor. Bessel processes and infinitely divisible laws. In *Stochastic integrals*, pages 285–370. Springer, 1981.

[112] W. Pogorzelski. Etude de la solution fondamentale de l'équation parabolique. *Richerche di Mat.*, 5:25–57, 1956.

[113] A. Popier. Backward stochastic differential equations with random stopping time and singular final condition. *The Annals of Probability*, pages 1071–1117, 2007.

[114] P. E. Protter. On the existence, uniqueness, convergence and explosions of solutions of systems of stochastic integral equations. *The Annals of Probability*, pages 243–261, 1977.

[115] Z. Qian, F. Russo, and W. Zheng. Comparison theorem and estimates for transition probability densities of diffusion processes. *Probability Theory and Related Fields*, 127(3):388–406, 2003.

[116] Z. Qian and Y. Yao. McKean–Vlasov type stochastic differential equations arising from the random vortex method. *arXiv preprint arXiv:2104.05100*, 2021.

[117] Z. Qian and W. Zheng. Sharp bounds for transition probability densities of a class of diffusions. *Comptes Rendus Mathematique*, 335(11):953–957, 2002.

[118] Z. Qian and W. Zheng. A representation formula for transition probability densities of diffusions and applications. *Stochastic Processes and their Applications*, 111(1):57–76, 2004.

[119] P. G. Rieusset-Lemarié. *The Navier Stokes Problem in the 21st Century.* Chapman and Hall/CRC, 2016.

[120] M. Röckner and X. Zhang. Well-posedness of distribution dependent SDEs with singular drifts. *Bernoulli*, 27(2):1131–1158, 2021.

[121] L. Rosenhead. The formation of vortices from a surface of discontinuity. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 134(823):170–192, 1931.

[122] E. Rothe. Über die grundlösung bei parabolischen gleichungen. *Math. Zeit.*, 33:488–504, 1931.

[123] G. Seregin. *Lecture notes on regularity theory for the Navier–Stokes equations.* World Scientific, 2015.

[124] S. Shankar and L. Van Dommelen. A new diffusion procedure for vortex methods. *Journal of Computational Physics*, 127(1):88–109, 1996.

[125] D. W. Stroock. Diffusion semigroups corresponding to uniformly elliptic divergence form operators. In *Séminaire de probabilités XXII*, pages 316–347. Springer, 1988.

[126] D. W. Stroock. *Partial Differential Equations for Probabilists.* Cambridge Univ. Press, 2008.

[127] D. W. Stroock and S. R. S. Varadhan. Diffusion processes with continuous coefficients, I. *Communications on Pure and Applied Mathematics*, 22(3):345–400, 1969.

[128] D. W. Stroock and S. R. S. Varadhan. Diffusion processes with continuous coefficients, II. *Communications on Pure and Applied Mathematics*, 22(4):479–530, 1969.

[129] D. W. Stroock and S. R. S. Varadhan. *Multidimensional Diffusion Processes*, volume 233. Springer Science & Business Media, 1997.

[130] A. S. Sznitman. Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX—1989*, pages 165–251. Springer, 1991.

[131] D. Talay and M. Tomašević. A new McKean–Vlasov stochastic interpretation of the parabolic–parabolic Keller–Segel model: The one-dimensional case. *Bernoulli*, 26(2):1323–1353, 2020.

[132] G. I. Taylor. Diffusion by continuous movements. *Proceedings of the London Mathematical Society*, 2(1):196–212, 1922.

[133] G. I. Taylor. Statistical theory of turbulence. *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, 151(873):421–444, 1935.

[134] R. Temam. *Navier–Stokes Equations: Theory and Numerical Analysis*, volume 343. American Mathematical Soc., 2001.

[135] R. Tevzadze. Solvability of backward stochastic differential equations with quadratic growth. *Stochastic processes and their Applications*, 118(3):503–515, 2008.

[136] M. Tomašević. A new McKean–Vlasov stochastic interpretation of the parabolic-parabolic Keller–Segel model: The two-dimensional case. *The Annals of Applied Probability*, 31(1):432–459, 2021.

[137] J. L. Vázquez. *The Porous Medium Equation: Mathematical Theory*. Oxford University Press on Demand, 2007.

[138] A. J. Veretennikov. On strong solutions and explicit formulas for solutions of stochastic integral equations. *Mathematics of the USSR-Sbornik*, 39(3):387, 1981.

[139] A. A. Vlasov. The vibrational properties of an electron gas. *Soviet Physics Uspekhi*, 10(6):721, 1968.

[140] E. C. Waymire. Probability & incompressible Navier–Stokes equations: An overview of some recent developments. *Probability Surveys*, 2:1–32, 2005.

[141] H. Xing and G. Žitković. A class of globally solvable Markovian quadratic BSDE systems and applications. *The Annals of Probability*, 46(1):491–550, 2018.

[142] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University*, 11(1):155–167, 1971.

[143] J. Yong. Finding adapted solutions of forward-backward stochastic differential equations: method of continuation. *Probability Theory and Related Fields*, 107(4):537–572, 1997.

[144] J. Yong. Forward-backward stochastic differential equations with mixed initial-terminal conditions. *Transactions of the American Mathematical Society*, 362(2):1047–1096, 2010.

[145] X. Zhang. Weak solutions of McKean–Vlasov SDEs with supercritical drifts. *arXiv preprint arXiv:2010.15330*, 2020.

[146] X. Zhang and G. Zhao. Stochastic Lagrangian path for Leray's solutions of 3D Navier–Stokes equations. *Communications in Mathematical Physics*, 381(2):491–525, 2021.

[147] A. K. Zvonkin and N. V. Krylov. Strong solutions of stochastic differential equations. In *Proceedings of the school and seminar on the theory of random processes (Druskininkai, 1974), Part II (Russian)*, pages 9–88, 1975.