Stochastic Differential Inclusions

Xiaoli Chen

Doctor of Philosophy
University of Edinburgh
2006

To my beloved grandmother,
my mother and father, to whom I owe everything.

Abstract

Stochastic differential inclusions (SDIs) on $\mathbb{R}^d$ have been investigated in this thesis,

$$dx(t) \in a(t, x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j,$$

where a is a maximal monotone mapping, b is a Lipschitz continuous function, and w is a Wiener process.

The principal aim of this work is to present some new results on solvability and approximations of SDIs. Two methods are adapted to obtain our results: the method of minimization and the method of implicit approximation. We interpret the method of monotonicity as a method of constructing minimizers to certain convex functions. Under the monotonicity condition and the usual linear growth condition, the solutions are characterized as the minimizers of convex functionals, and are constructed via implicit approximations. Implicit numerical scheme is given and the result on the rate of convergence is also presented. The ideas of our work are inspired by N.V.Krylov, where stochastic differential equations (SDEs) in $\mathbb{R}^d$ are solved by minimizing convex functions via Euler approximations.

Furthermore, since the linear growth condition is too strong, an approach is proposed for truncating maximal monotone functions to get bounded maximal monotone functions. It is a technical challenge in this thesis. Thus the existence of solutions to SDIs is proved under essentially weaker growth condition than the linear growth.

For a special case of SDEs, a few of recent results from [5] are generalized. Some existing results of the convergence by implicit numerical schemes are proved under the locally Lipschitz condition. We will show that under certain weaker conditions, if the drift coefficient satisfies one-sided Lipschitz condition and the diffusion coefficient is Lipschitz continuous, implicit approximations applied to SDEs, converge almost surely to the solution of SDEs. The rate of convergence we get is 1/4.

Finally it is shown that SDEs can directly be associated to mini-max problems. It is demonstrated that there exists strong solutions which are 'saddle points' of mini-max problems. This technique provides a simple proof of the existence results.

Acknowledgements

Firstly, I would like to thank my supervisor, Professor István Gyöngy. I could not have imagined having a better advisor and mentor for my PhD, and without his enthusiasm, his great efforts to explain things clearly and simply, I would never have finished. He provided me with encouragement, sound advice, good teaching and lots of good ideas. I would have been lost without him.

For financial support, I thank the School of Mathematics at the University of Edinburgh.

Finally, I am forever indebted to my parents and Tiejun for their understanding, endless patience and encouragement when it was most required. I thank for my brother Rui Xu who provided incomparable support to me throughout my PhD study. I thank my friend Jie Ding for the many discussions carried on long-distance via e-mail. I also warmly thank my friends and colleagues in the school of mathematics.

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

Xiaoli Chen

Table of Contents

Notations

$\mathbb{R}^d$	the Euclidean d-dimensional space
$\|x\|$	the norm $\|x\| = (\sum_{k-1}^{d} x_k^2)^{1/2}$ for vector $x = (x_i) \in \mathbb{R}^d$
$\|b\|$	the norm $\|b\| = (\sum_i \sum_j b_{ij}^2)^{1/2}$ for b is a $d \times d_1$ matrix
xy	$xy = \sum_i x_i y_i$ for vectors $x = (x_i), y = (y_i) \in \mathbb{R}^d$
bz	the product of matrices if b is a $d \times d_1$ matrix and $z \in \mathbb{R}^d$ notice that $bz \in \mathbb{R}^d$ and $\|bz\| \leq \|b\|\|z\|$
C	constant, usually without indices that may change line by line in the same proof
$C = C(\cdots)$	C depends only on what are inside the parenthesis
$X := Y$	X is equal to Y by definition
$\mathcal{B}(\mathbb{R}^d)$	the $\sigma-$algebra of the Borel subsets of $\mathbb{R}^d$
$B_r(x)$	the closed ball B_r of radius r centered at x: $\{y \in \mathbb{R}^d, \|x - y\| < r\}$
$conv M$	convex hull of the set M
$\overline{conv} M$	closure of convex hull of the set M
E	mathematical expection
$w = \{w_t : t \geq 0\}$	d_1-dimensional Wiener Process
$e_1, e_2, \cdots, e_d$	an orthonormal basis in $\mathbb{R}^d$
$a.e.$	almost surely
$\longrightarrow$	strong convergence
$\rightharpoonup$	weak convergence
SDE	stochastic differential equation
SDI	stochastic differential inclusion

The following assumptions will be adopted *throughout this thesis.*

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a complete probability space with natural filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$ carrying a d_1-dimensional $\mathcal{F}_t-$Wiener martingale $w = \{w_t : t \geq 0\}$. Fix an integer $d > 1$.

Chapter 1

Introduction

This thesis is devoted mainly to stochastic differential inclusions (SDIs). SDIs represent an important generalization of the notion of stochastic differential equations (SDEs). In the case of an SDE, one wants to find a stochastic process $x = x(t)$, whose stochastic differential $dx(t)$ is given by an equation:

$$dx(t) = a(t, x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j \tag{1.0.1}$$

with a deterministic drift term a, perturbed by a noisy diffusion term b, where $a : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$, $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^{d \times d_1}$ are Borel functions and w is a d_1-dimensional Wiener process. For SDIs we require $dx(t)$ belong to the set of stochastic differential described by the right-hand side of (1.0.1). In the thesis, the following SDI will be investigated on domain of $\mathbb{R}^d$ with a multi-valued drift term:

$$\begin{cases} dx(t) \in a(t, x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j \\ x(0) = x_0. \end{cases} \tag{1.0.2}$$

There is a great variety of motivations that lead us to study SDIs. By the classical results of Itô's SDE (1.0.1) with a specified initial value (independent of w), SDE has a unique solution if a and b are Lipschitz continuous in x and satisfy linear growth condition. It is shown that under the so-called monotonicity condition, the existence and uniqueness of a solution to SDEs is obtained in [20]. In many practical problems, several applications are possible like solving SDEs with discontinuous right-hand side. In such case the existence of a solution is not always guaranteed. For example, let us consider the following Itô's SDE:

$$\begin{cases} dx(t) = a(x)dt + xdw_t \\ x(0) = 0 \end{cases} \tag{1.0.3}$$

with $a(x) = 1$ for $x \leq 0$ and $a(x) = -1$ for $x > 0$. Obviously, there exists no solution. On the other hand, assume that we consider an explicit Euler approximation x^n for the SDE (1.0.3). It is easy to show that x^n converges almost surely to some stochastic process x, but x is not the solution of the given SDE (1.0.3). It is both practical and essential to extend SDE to SDI. However, if a is extended to be multi-valued and SDE to be a stochastic inclusion, we can show that it is a unique solution of the SDI obtained. Nonetheless it can be found that this equation falls into a general class. Generally, we understand SDIs as an enlargement of SDEs. The right-hand side of an SDI is a set rather than a single value. So far we can see that SDIs play a crucial role in the theory of SDEs with a discontinues right-hand side.

It is noticed that the nature of the existence and uniqueness solution for SDIs problem has been extensively studied for long time by using different methods. Articles have appeared recently in which SDIs or SDEs with multi-valued operators are studied [25], [31]. The concept of solutions to SDIs has been introduced in these papers. Numerically approximate methods have also been tackled in [31], [27], and [8]. There has been a strong desire to produce numerical solutions to SDIs. This is also our main interest in this thesis. We shall see that SDIs are solved by a minimization method for some convex functionals via implicit approximations (also known as semi-implicit method or backwark Euler method).

During our study of implicit method for SDIs, some recent results from SDEs, by Higham, D.J., Mao, X. and Stuart,A.M. [5] and Hu, Y. [16] are generalized, from where a rate of convergence is derived. It is shown that, the existing proofs of the convergence results on such numerical schemes in both of these two papers are proved under the locally Lipschitz condition. This is the main reason why these work have a better rate of convergence. In this thesis, the local Lipschitz condition of the drift term is weaken and replaced by one-sided Lipschitz condition comparing with those previous papers. We will show that implicit approximations converge almost surely if the drift satisfies one-sided Lipschitz condition and the diffusion is Lipschitz continuous.

In the last chapter of this thesis, it is demonstrated that SDEs can directly be associated with mini-max problems in suitable infinite dimensional spaces. This adapts an idea of N.V. Krylov. More precisely, if the coefficients of an SDE satisfy the so called monotonicity condition, then one can construct a mini-max problem such that its saddle point is the solution of the given SDE.

A brief description of the chapters contained in this thesis is presented as follows:

- Chapter 2: this chapter presents the background material and some results that are required in later chapters;
- Chapter 3: in this chapter, some new results are generalized from SDEs by

comparing with the paper [16] and [5];

- Chapter 4: this chapter is devoted to the existence and uniqueness of solutions for SDIs by minimization method;
- Chapter 5: this chapter contains the extension of monotone and maximal monotone function that are useful technique needed in the following chapter;
- Chapter 6: this chapter further studies the existence of solution for SDIs by truncation method;
- Chapter 7: this chapter shows that the solution of a SDE can be considered as saddle points of a mini-max problem.

Chapter 2

Preliminaries

The purpose of this chapter is to present some general background and briefly summarize some results from the theory of probability and stochastic differential equations which we will need in later chapters. Such theory is much more extensive than what I present here and can be found in many textbooks. For example we refer the reader to books by N.V.Krylov [23], [24], B. Øksendal [30]. Most definitions and results in this chapter are due from lecture notes [15]. This chapter covers:

- section 2.1: this section gives the background of probability theory;
- section 2.2: this section gives the definition of stochastic process, and the special class Wiener process;
- section 2.3: this section contains some basic inequalities and statements which will be used through whole thesis;
- section 2.4: in this section we discuss the important theory of monotone and maximal monotone mappings.

2.1 Probability Theory Background

Definition 2.1.1. Let Ω be a set. Then a *σ-algebra* $\mathcal{F}$ is a collection of subsets of Ω such that

1. $\Omega \in \mathcal{F}$.
2. If A_i is a sequence of elements of $\mathcal{F}$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.
3. If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$, where $A^c = \Omega \setminus A$.

Definition 2.1.2. A *probability measure* defined on a σ-algebra of Ω is a function $P : \mathcal{F} \to [0, 1]$ that satisfies the following properties

1. $P(\Omega) = 1$.

2. If $A_1, A_2, \cdots$ is a sequence of elements of $\mathcal{F}$ that are *pairwise disjoint* (i.e. $A_i \cap A_j = \emptyset$ for all $i \neq j$), then $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

Definition 2.1.3. Let $(\Omega, \mathcal{F}, P)$ be a probability space. Then, a function $X : \Omega \to \mathbb{R}$ is called a *random variable* if and only if the set $\{X \in [a,b]\} = \{\omega \in \Omega : X(\omega) \in [a,b]\} \in \mathcal{F}$ for all $a < b$.

Definition 2.1.4. The *Borel σ-algebra* $\mathcal{B}(\mathbb{R})$ is defined as the smallest σ-algebra containing all intervals of the form $[a,b]$, where a and b are real numbers $(a < b)$:

$$\mathcal{B}(\mathbb{R}) = \sigma([a,b] : a < b)$$

(in other words, $\mathcal{B}(\mathbb{R})$ is generated by intervals of the above form).

Definition 2.1.5. If X is a random variable defined on the probability space $(\Omega, \mathcal{F}, P)$, then the *expected value* or *mean value* of X is

$$EX = \int_{\Omega} X dP.$$

Expectations satisfy various properties. For example

1. It is a linear functional: If $E|X| < \infty$ and $E|Y| < \infty$ then $E|\alpha X + \beta Y| = \alpha E|X| + \beta E|Y|$ for every $\alpha, \beta \in \mathbb{R}$.

2. If $X \geq 0$, then $E|X| \geq 0$.

Moreover, the following assertions hold:

Lemma 2.1.1 (Borel-Cantelli Lemma). *If A_n are any events with $\sum_n P(A_n) < \infty$, then*

$$P(\limsup_n A_n) = 0.$$

If the A_n are independent and $\sum_n P(A_n) = +\infty$, then $P(\limsup_n A_n) = 1$.

Theorem 2.1.2 (Beppo-Levi's theorem). *If $\{X_n\}_{n\geq 1}$ is a sequence of nonnegative increasing sequence of random variables that converges almost surely to a random variable X, then*

$$EX = \lim_{n\to\infty} EX_n.$$

Lemma 2.1.3 (Fatou's lemma). *Let $\{X_n\}_{n\geq 1}$ be a sequence of non-negative random variables. Then*

$$E(\liminf_{n\to\infty} X_n) \leq \liminf_{n\to\infty} EX_n.$$

Theorem 2.1.4 (Lebesgue's Dominated Convergence Theorem). *Let $\{X_n\}_{n\geq 1}$ be a sequence of random variables, converging(a.s.) to some random variable X. Assume that there exists a (non-negative) random variable Y such that $|X_n| < Y$ for all n(a.s.) and $E|Y| < \infty$, then $E|X| < \infty$, and*

$$E|X| = \lim_{n\to\infty} EX_n.$$

If there is an infinite sequence of random variable, then it is necessary to know the convergence of sequences. The following are given the different modes of convergence:

Definition 2.1.6. A sequence of random variables $\{X_n(\omega)\}$ *converges with probability one* to $X(\omega)$, if $P(\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}) = 1$. This is also called *almost surely convergence.*

Definition 2.1.7. A sequence of random variable $\{X_n(\omega)\}$ *converges in probability* to X if

$$\lim_{n\to\infty} P(|X_n(\omega) - X(\omega)| \geq \varepsilon) = 0, \ \forall \varepsilon > 0.$$

Definition 2.1.8. Suppose that X_n and X are real-valued random variables with distribution functions F_n and F respectively. We say that the distribution of X_n *converges to the distribution* of X as $n \to \infty$, if

$$F_n(X) \to F(X), \text{ as } n \to \infty.$$

Definition 2.1.9. We say that the sequence X_n *converges in r-th mean* or *in the L^r-norm* towards X, if $r \geq 1$, $E|X_n|^r < \infty$, for all n, and

$$\lim_{n\to\infty} \mathrm{E}\left(|X_n - X|^r\right) = 0.$$

2.2 Stochastic Processes and Wiener Processes

Definition 2.2.1. A *stochastic process* $X := \{X_t : t \in \boldsymbol{T}\}$ is a parameterized collection of random variables with index set T. For each fixed $\omega \in \Omega$, the function

$$t \to X_t(\omega); \ t \in \boldsymbol{T}$$

is called a *trajectory*. When $\boldsymbol{T}$ is discrete, then X is called stochastic process in discrete time; When $\boldsymbol{T}$ is an interval or half line of the real line, or the whole real line, e.g., $\boldsymbol{T} = [0,T]$ or $\boldsymbol{T} = D \subset \mathbb{R}^d$ then X is called stochastic process in continuous time. In this case we always assume that the stochastic process $X : \Omega \times \mathbf{T} \to \mathbb{R}^d$ is measurable in (ω, t) with respect to the product $\sigma-$algebra $\mathcal{F} \otimes \mathcal{B}(\mathbb{R})$.

Definition 2.2.2. Let $\boldsymbol{T} = [0, T]$. Then, $X := \{X_t : t \in \boldsymbol{T}\}$ is a *continuous process* if the trajectories are continuous, i.e., each $\omega \in \Omega$ is mapped to a continuous function of time defined on $[0, T]$, $\omega \to X.(\omega) \in C([0, T])$.

Definition 2.2.3. A stochastic process $X := \{X_t : t \in \boldsymbol{T}\}$ is called *cadlag*, if the trajectories are right-continuous with left limits, i.e. for every fixed $\omega \in \Omega$, $X_t(\omega)$ is right-continuous and the $\lim_{s\uparrow t} X_s$ exists for every $t \in [0, T]$.

An important class of stochastic processes is that with independent increments; that is, where the difference $X(t_{k+1}) - X(t_k)$ are independent.

Definition 2.2.4 (Wiener Process). A standard one-dimensional *Wiener process* (Brownian motion) with respect to $\{\mathcal{F}_t\}$ is a continuous $\mathcal{F}_t-$adapted process $W = \{W_t : t \geq 0\}$ defined on $(\Omega, \mathcal{F}, P)$ with properties

1. $W_0 = 0, a.s;$
2. for every $0 \leq s < t$, $W(t) - W(s)$ is independent of $\mathcal{F}_s$
3. $W(t) - W(s)$ is normally distributed with mean zero and variance $t - s$.

Definition 2.2.5 (Martingale). A stochastic process X_t is called a *martingale* with respect to $\mathcal{F}_t$ if it satisfies the following conditions:

1. X_t is $\mathcal{F}_t-$adapted;
2. $E|X_t| < \infty$, for all $t \geq 0$;
3. $E(X_t|X_s) = X_s$, for every $s, t \geq 0$, such that $s \leq t$.

A stochastic process X_t is called a *local martingale* with respect to $\mathcal{F}_t$ if there exists a sequence of stopping time τ_n such that $\tau_n \uparrow \infty$ a.s. and $X_{t\wedge\tau_n}$ is a martingale with respect to $\mathcal{F}_{t\wedge\tau_n}$ for $n \in N$.

2.3 Fundamental Inequalities and Statements

Next we present some important inequalities and statements that will be used in the following chapters:

Proposition 2.3.1 (Chebyshev's inequality). *If ξ is a random variable then*

$$P(|\xi| \geq \lambda) \leq \lambda^{-\alpha} E(|\xi|^{\alpha}),$$

$\forall \lambda > 0,\ \alpha > 0.$

Lemma 2.3.2 (Young's inequality). *If a, b, p, q and δ are positive real number with $1/p + 1/q = 1$, then we have*

$$ab \leq \frac{\delta}{p}a^p + \frac{1}{q\delta^{q/p}}b^q.$$

Theorem 2.3.3 (The Roger-Hölder inequality). *For any random variables ξ, η*

$$E(|\xi\eta|) \leq |\xi|_p|\eta|_q, \forall p \geq 1,$$

where $q := \frac{p}{p-1}$ if $p > 1$ and $q = \infty$ if $p = 1$.

Remark 2.3.1. The special case $p = q = 2$ of the Roger-Hölder inequality is often called the *Cauchy-Bunyakovsky* inequality.

Theorem 2.3.4 (Minkowski's inequality). *For any random variables ξ, η, and for $p \in [1, \infty]$*

$$|\xi + \eta|_p \leq |\xi|_p + |\eta|_p.$$

Theorem 2.3.5 (Jensen's inequality). *If f is a convex function then $f(E\xi) \leq Ef(\xi)$ for every random variable ξ, provided $E\xi$ is finite.*

Theorem 2.3.6 (Burkholder-Davis-Gundy's inequality). *For any $p \in (0, \infty)$, there exists constant $C_p < \infty$ depending only on p, such that for every $T > 0$,*

$$E \sup_{t \leq T} |\int_0^t f_s dw_s|^p \leq C_p E(\int_0^T f_s^2 ds)^{p/2},$$

for every $\mathcal{F}_t-$adapted stochastic process $\{f_t : t \in [0, T]\}$.

Theorem 2.3.7 (Gronwall's inequality). *Let $T > 0$ and $c \geq 0$. Let u be a Borel function on $[0, T]$, such that*

$$0 \leq u(t) \leq c + \int_0^t v(s)u(s)ds,$$

holds for all $0 \leq t \leq T$, where v is a non-negative function having finite integral over $[0, T]$. Then

$$u(t) \leq c \exp(\int_0^t v(s)ds),$$

for all $0 \leq t \leq T$.

We need the following discrete version from the above Gronwall theorem to get some estimates for the discrete time approximations.

Lemma 2.3.8 (Discrete Gronwall inequality). *Let $\{a_i\}$ be a sequence, $i = 0, 1, 2, \cdots, k-1$. If for $k = 1, 2, \cdots, n$ the inequality*

$$a_k \leq C + K \sum_{i=0}^{k-1} a_i$$

holds, where C, $K \geq 0$ are constants. Then

$$|a_k| \leq C(1+K)^k.$$

Proof. Define $b_k := C + K \sum_{i=0}^{k-1} b_i$. Then we claim that

$$a_k \leq b_k.$$

Indeed,

$$a_0 \leq C = b_0; \; a_1 \leq C + K a_0 \leq C + K b_0 = b_1;$$

we get the claim by induction. Since,

$$b_{k+1} - b_k = K b_k, \; b_{k+1} = b_k(K+1).$$

So

$$b_{k+1} = b_k(K+1) = b_{k-1}(K+1)^2$$
$$= b_{k-2}(K+1)^3 = \cdots = b_0(K+1)^{k+1}$$

Finally, we obtain

$$a_{k+1} \leq b_{k+1} = b_0(K+1)^{k+1} = C(1+K)^{k+1}.$$

□

In order to construct (maximal) monotone extensions of (maximal) monotone functions, we will make use of the following well-known result:

Theorem 2.3.9 (Separation of Convex Sets [34]). *Suppose A and B are disjoint, nonempty, convex sets in a topological vector space X. If A is compact, B is closed, and X is locally convex, then there exist $a \in X^*$, $\gamma_1 \in \mathbb{R}$, $\gamma_2 \in \mathbb{R}$, such that*

$$\langle a, x \rangle < \gamma_1 < \gamma_2 < \langle a, y \rangle,$$

for every $x \in A$ and for every $y \in B$, where $\langle x, x^* \rangle$ denotes the duality product of $x \in X$ and $x^* \in X^*$.

2.4 Monotone and Maximal Monotone Mappings

We devote the last part of this chapter to a very important theory of monotone and maximal monotone mappings. The material covered in this part is the base for the rest of the thesis.

2.4.1 Introduction

The theory of monotone and maximal monotone mappings play an important role in several fields of mathematics such as functional analysis, partial differential equations. They have turned out to be very useful in the study of the existence and uniqueness theory for ordinary differential equations, partial differential equations, differential inclusions, stochastic Itô's equations. This is a very simple but important technique. We have seen in many works, that the property of monotonicity and maximal monotonicity make a big contribution to the SDEs problems. For the beginning, let us summarize some general sources and concepts of the monotone and maximal monotone mappings. Then in chapter 5, we will give the extension of monotone and maximal monotone mappings that will lead to our proof.

We start with some basic definitions:

2.4.2 Notation and Some Definitions

Definition 2.4.1. A mapping $a : \mathbb{R}^d \to \mathbb{R}^d$ is called *monotone*, if

$$(a(x) - a(y))(x - y) \leq 0,$$

for all $x, y \in \mathbb{R}^d$.

The definition of monotone mapping can be extended to a multi-valued case. A mapping $a : \mathbb{R}^d \to 2^{\mathbb{R}^d}$ is multi-valued, where $2^{\mathbb{R}^d}$ is all the subsets of $\mathbb{R}^d$, and a will be viewed as the subset of $\mathbb{R}^d \times \mathbb{R}^d$. If $a \subset \mathbb{R}^d \times \mathbb{R}^d$ we define,

$$a(x) = \{y \in \mathbb{R}^d : (x, y) \in a\} \text{ is the image;}$$

$$D(a) = \{x \in \mathbb{R}^d : a(x) \neq \emptyset\} \text{ is the effective domain of } a;$$

$$R(a) = \cup\{a(x), x \in D(a)\} \text{ is called the range of } a;$$

$$a^{-1} = \{(y, x) : (x, y) \in a\} \text{ is the inverse mapping .}$$

If a is multi-valued then the above definition of monotonicity is replaced by,

Definition 2.4.2. A mapping $a : \mathbb{R}^d \to 2^{\mathbb{R}^d}$ is called *monotone*, if $(x^* - y^*)(x - y) \leq 0$, for all $x, y \in \mathbb{R}^d$, $x^* \in a(x), y^* \in a(y)$.

Example 2.4.1. *We give an example of a multi-valued monotone function. Let* $a : \mathbb{R} \to \mathbb{R};\ a(x) = 0,$ *if* $x < 1;$ $a(x) = 1$ *if* $x > 1,$ *and let* $a(1)$ *be any subset of* $[0, 1]$.

In more general case, we give the definition of a $K-$monotone mapping:

Definition 2.4.3. A multi-valued mapping $a : \mathbb{R}^d \longrightarrow 2^{\mathbb{R}^d}$ is called *K–monotone* if

$$(x^* - y^*)(x - y) \leq K|x - y|^2, \ \forall\, x, y \in \mathbb{R}^d, \ x^* \in a(x), \ y^* \in a(y),$$

where K is a constant.

In order to introduce the definition of maximal monotone mapping, we consider their graph first.

We say that the set $\Gamma(a) = \{(x, y) : x \in D(a), y \in a(x)\}$ is called the *graph* of a. The map $\bar{a} : \mathbb{R}^d \times \mathbb{R}^d$ is called an *extension* of a if $\Gamma(a) \subsetneq \Gamma(\bar{a})$.

Definition 2.4.4. If $a : \mathbb{R}^d \longrightarrow \mathbb{R}^{2^d}$ is a K–monotone mapping, such that it does not have a proper K–monotone extension, then it is called a *maximal K–monotone mapping.*

If $K = 0$, in Definition 2.4.3 and 2.4.4, we shall call 0–monotone and 0–maximal monotone simply monotone and maximal monotone.

From Zorn's lemma, the graph of every monotone map is contained in the graph of a multi-valued maximal monotone map. It shows that monotone mapping can get extended to a maximal monotone mapping.

Remark 2.4.1. Notice that monotone (maximal monotone) mapping we defined here is based on monotone decreasing, while most of the definitions in textbooks are monotone increasing.

Remark 2.4.2. We can see that maximal monotonicity is a property for continuous monotone function, i.e., if a continuous function $a : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ is monotone, it is maximal monotone.

In order to get a better understanding of a maximal monotone function, let us give an example.

Example 2.4.2. *We consider a multi-valued function f from $\mathbb{R}$ to $2^{\mathbb{R}}$. If f is monotone discreasing but discontinuous at x, where $x \in \mathbb{R}$. Then f is monotone but not maximal monotone. It is easy to see that the set $\Gamma(F) = \{(x, y) : x \in \mathbb{R},\ f(x+0) \leq y \leq f(x-0)\}$ is a monotone set and that every monotone extension of f is contained by it. Such multi-valued function F is maximal monotone if and only if $F(x) := [f(x + 0), f(x - 0)]$.*

Remark 2.4.3. If a is a K–monotone function, we set $\bar{a} = a(x) - Kx$. Then $\bar{a}(x)$ turns out to be monotone. Furthermore, if a is a maximal K–monotone function, then it is easy to see that $\bar{a}$ is maximal monotone.

Proposition 2.4.1. *A multi-valued map a is monotone (maximal monotone) if and only if its inverse a^{-1} is monotone (maximal monotone).*

Proposition 2.4.2. *Let a be maximal monotone. Then its images are closed and convex.*

These results are consequences of definition of maximal monotone function. We omitted the proofs.

2.4.3 Known Results

The property of maximal monotonicity is closely related to the surjectivity of the operator $I + a$, in which I is the unit operator in $\mathbb{R}^d$. This result is a theorem of Minty, (see e.g., Aubin and Cellina[2])

Theorem 2.4.3 (Minty). *Let a be a monotone set-valued map from X to X. It is maximal if and only if $I + a$ is surjective.*

Here X denotes Hilbert space.

For details of proof of the above theorem, we refer to Aubin and Cellina [2].

We give some well-known properties of maximal monotone mappings. The relevant material can be found in Aubin and Cellina [2] and E. Zeidler [40]. The following lemma is sufficient for our later proofs.

Lemma 2.4.4. *Let a be a mapping on $\mathbb{R}^d$, the following statement are equivalent:*

(a) *a is maximal monotone;*

(b) *a is monotone, and $Im(I + a) = \mathbb{R}^d$,*

(c) *$(I + \lambda a)^{-1}$ is a contraction on $\mathbb{R}^d$.*

A fundamental characterization of maximal monotone mapping is as follows:

2.4.3.1 Yosida Approximation

Let A be a maximal monotone map on $\mathbb{R}^d$. For $\lambda > 0$, the inclusion $y \in x + \lambda A(x)$ has a unique solution x, for any fixed y , denoted by $(I + \lambda A)^{-1}y = J_\lambda y$. The mapping J_λ is a contraction from $\mathbb{R}^d \to \mathbb{R}^d$. We put $A_\lambda := \frac{I - J_\lambda}{\lambda}$. It can be shown that A_λ is maximal monotone and Lipschitz continuous. Such Lipschitz mapping A_λ is known as *Yosida approximation.* We state some important properties of Yosida approximation without proofs due to Aubin and Cellina [2].

Theorem 2.4.5. *Let A be a maximal monotone map on $\mathbb{R}^d$. Then*

1. *For all $x \in D(A)$, there exists a unique point A^0x such that $|A^0x| = \min\{|y| : y \in Ax\}$.*

2. $|J_\lambda x - J_\lambda y| \leq |x - y|$, *for all* $\lambda > 0$ *and* $x, y \in \mathbb{R}^d$.

3. J_λ, A_λ *are Lipschitzean with constants 1 and* $1/\lambda$ *respectively.*

4. A_λ *is maximal monotone.*

5. $A_\lambda x \in AJ_\lambda x$ *for all* $\lambda > 0$ *and* $x \in \mathbb{R}^d$.

6. *For* $x \in D(A)$, $|A_\lambda x| \leq |A^0 x|$, *and* $A_\lambda x \to A^0 x$ *as* $\lambda \to 0$.

7. *For* $x \in D(A)$, $J_\lambda x \to x$.

We can see that in particular, a maximal monotone mapping A can be approximated by single-valued Lipschitz mapping A_λ that are also maximal monotone to some extent.

Chapter 3

Implicit Approximation Schemes for Stochastic Differential Equations

In this chapter, some resent results on stochastic differential equations from [5] are generalized. We will prove that implicit approximations for SDEs on domain of $\mathbb{R}^d$ converge almost surely if the drift satisfies one-sided Lipschitz condition and the diffusion coefficient is Lipschitz continuous.

3.1 Introduction

Let us investigate the following SDE in this chapter, for $x \in \mathbb{R}^d$, and $t \in [0, T]$,

$$dx(t) = a(t, x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j.$$

For simplicity, we consider the time-independent case.

It is well known that numerical methods are extensively applied to solve SDEs problems. There are various types of methods to solve numerically SDEs. An overview of the existing numerical methods is given in Kloeden and Platen [26]. A numerical solution $x^n := \{x^n(t), t \in [0, T]\}$ is a stochastic process that approximates the solution $x := \{x(t), t \in [0, T]\}$ of an SDE. The first step towards the development of numerical solutions to SDEs is Euler's polygonal approximation, which is the simplest discrete approximation, (it is also known as Euler-Maruyama approximations). One can define Euler's polygonal approximations as follows: for every integer $n \geq 1$, given a partition

$$0 = t_0 < t_1 < \cdots < t_k < \cdots = T, \ k = 1, 2, \cdots, n$$

of the time interval $[0, T]$, with a step size $\Delta t = \frac{T}{n}$, it allows us to express the discrete time stochastic process in the form of

$$x^n(t_k) = x^n(t_{k-1}) + a(x^n(t_{k-1}))\Delta t + \sum_{j=1}^{d_1} b^j(x^n(t_{k-1}))\Delta w^j_{k-1},$$

for $\Delta w^j_{k-1} = w^j_{t_k} - w^j_{t_{k-1}}$.

We notice that the convergence of Euler's approximations is discovered by many authors under different conditions. It is worth mentioning that it is first known from Maruyama, G. [28] in the case of the Lipschitz continuity of the drift and diffusion coefficients. In [19] Krylov, N.V. showed that the existence of strong solutions can be constructed under the Euler polygonal line method. Then it is shown in [1] that Euler polygonal lines can be used as a new proof of the existence of solutions under the monotonicity and the linear growth condition by Alyushina, L.A. . She also obtained an estimate for the speed of convergence. Afterward, it is known from Krylov, N.V [20], where a simple proof of solvability by Euler's polygonal line method is presented under monotonicity and under a condition which is weaker than the usual linear growth. Later a proof of solvability, based on very general conditions by Euler's approximations, can be found in [9]. Gyöngy, I. and Krylov, N.V. obtained that Euler's approximations converge in probability to strong solutions, even if the drift term is only measurable and the diffusion term is Lipschtiz, while a description of convergent proof by Euler's approximations, based on the monotonicity of the drift and Lipschitz continuity of the diffusion can be found in Gyöngy, I.'s paper [10]. Higham, D.J., Mao, X. and Stuart,A.M., in [5] gave a strong convergent result for Euler method when the drift and diffusion coefficients are locally Lipschitz. As a further extension, they showed a more widely used implicit variant of the Euler methods by relating two implicit methods. One is split-step extension of the backward Euler method. The other, more naturally extends the backward Euler method (we call it implicit approximations). In this chapter we are interested in *implicit discretization scheme* in the forms of:

$$x^n(t_k) = x^n(t_{k-1}) + a(x^n(t_k))\Delta t + \sum_{j=1}^{d_1} b^j(x^n(t_{k-1}))(w^j_{t_k} - w^j_{t_{k-1}}),\ 1 \le k \le n,$$

for $T_n := \{t_k = k\Delta t\}$, where $\Delta t = \frac{T}{n}$ for $n \ge 1$. In their work [5], under a local Lipschitz condition, and the boundedness of the pth moments of both the exact and numerical approximations for any $p > 2$, combined with a polynomial growth condition, they proved that the backward Euler approximations converge in mean square with a rate of $1/2$. The other published work in this area that we are aware of is Hu, Y [16]. He proved the strong convergent result when the drift coefficient satisfies the one-sided Lipschitz condition and exponential growth, the diffusion coefficient has the bounded derivative. He obtained that implicit scheme has the convergence rate of 1/2. This chapter is influenced by the works [16], [5].

Our main result is to prove the convergence of implicit scheme under somewhat weaker conditions. It is worthwhile to compare our work with that in the above two papers [5] and [16]. As we can see, the existing proofs of the convergence of such numerical schemes in both of these two papers require the locally Lipschitz condition. This is the main reason why they have a better rate of convergence. In our work, we obtain boundedness of the qth moments of both the exact and numerical approximations for any $q \geq 1$, where the diffusion is Lipschitz continuous and the drift satisfies a one-sided Lipschitz condition. Further, we require the drift behave polynomially. We show that implicit approximations converge almost surely to solutions of SDEs. Moreover, the rate of convergence we proved is $n^{-\alpha}$ for every $\alpha < \frac{1}{4}$. (See Theorem 3.2.6 and 3.2.7 below.) Our results are comparable to the results in [5] and [16].

This chapter is organized in the following way:

- section 3.2: in this section we describe our discretization implicit scheme, prove the existence and uniqueness of implicit approximations precisely, and state the main results of this chapter;

- section 3.3: this section gives the preliminary lemmas that will be used to establish the main proofs. Boundedness of the implicit approximation solutions is given in this section;

- section 3.4: in this part we provide the main proofs of main theorems.

3.2 Formulation of the Results

We consider the stochastic differential equation as follows:

$$\begin{cases} dx(t) = a(x(t))dt + \sum_{j=1}^{d_1} b^j(x(t))dw_t^j, \\ x(0) = x_0, \end{cases} \tag{3.2.1}$$

where $a(x), b(x)$ are Borel functions on $\mathbb{R}^d$ taking values in $\mathbb{R}^d$ and $\mathbb{R}^{d \times d_1}$ respectively.

By a solution of SDE (3.2.1) we mean an $\mathcal{F}_t-$adapted $\mathbb{R}^d-$valued stochastic process $x(t) = x_t(\omega)$ satisfying equation (3.2.1) on the interval $[0, T]$ for almost every $\omega \in \Omega$.

The hypothesis on convergence theory are usually sufficient, but not necessary. Some of those are quite strong, but can be weakened in several ways. In what follows we list assumptions that are concerned in this chapter:

Assumption 3.2.1. Let x_0 be an $\mathcal{F}_0$−measurable random variable in $\mathbb{R}^d$ such that $E|x_0| < \infty$.

Assumption 3.2.2. a is continuous in $x \in \mathbb{R}^d$.

Assumption 3.2.3 (Local Monotonicity of a)**.** For any $R > 0$, there exists a constant L_R, such that

$$(x-y)(a(x)-a(y)) \leq L_R|x-y|^2,$$

for any $x, y \in \mathbb{R}^d$, with $|x|, |y| \leq R$.

Assumption 3.2.4 (Local Lipschitz of b)**.** For any $R > 0$, there exists a constant L_R, such that

$$|b(x)-b(y)| \leq L_R|x-y|,$$

for any $x, y \in \mathbb{R}^d$, with $|x|, |y| \leq R$.

Assumption 3.2.5 (One-sided Linear Growth of a, b)**.** There exist non-negative constants C_1, C_2, C_3, such that

$$xa(x) \leq C_1|x|^2 + C_2;$$

$$|b(x)| \leq C_3(1+|x|).$$

Remark 3.2.1. Notice that in the local conditions 3.2.3 and 3.2.4 if $L_R := L$ for all R and L is a non-negative constant, then we get the global conditions as follows:

Assumption 3.2.6. There exist constants $L_1, L_2 > 0$, such that for all $x, y \in \mathbb{R}^d$,

$$\text{(i) } (x-y)(a(x)-a(y)) \leq L_1|x-y|^2,$$

$$\text{(ii) } |b(x)-b(y)| \leq L_2|x-y|.$$

Remark 3.2.2. Inequality (i) in Assumption 3.2.6 is also known as *one-sided Lipschitz condition.* This means, the function a is a K−monotone function.

Remark 3.2.3. Notice that Assumptions 3.2.6 implies Assumption 3.2.5.

Assumption 3.2.7 (Polynomial Condition of a)**.** For any integer $r \geq 1$, there exists a constant $C_4 > 0$, such that

$$|a(x)| \leq C_4(1+|x|^r).$$

We now describe implicit discretization schemes. Let $n \geq 1$ be any integer, and set $t_k = k\Delta t$ for $0 \leq k \leq n$.

3.2.1 Implicit Approximation Schemes

Under the above conditions, we approximate the solution $x(t)$ of the equation (3.2.1) by the process $x^n(t)$ solving the following equation:

$$\begin{cases} x^n(t_k) = x^n(t_{k-1}) + a(x^n(t_k))\Delta t + \sum_{j=1}^{d_1} b^j(x^n(t_{k-1}))(w^j_{t_k} - w^j_{t_{k-1}}), \\ x^n(t_0) = x_0, \end{cases} \tag{3.2.2}$$

and defined as

$$x^n(t) := x^n(t_{k-1}), \text{ for } t \in [t_{k-1}, t_k),$$

for $1 \le k \le n$. Here we denote $\Delta w_{k-1} := w_{t_k} - w_{t_{k-1}}$. This scheme is also known as *backward Euler method* [5] or *semi-implicit Euler-Maruyama time discretization scheme* [16].

The following statements establish the existence and uniqueness of $x^n(t)$ for this system of stochastic equations (3.2.2):

3.2.2 Existence and Uniqueness of Solution of the Implicit Schemes

The lemmas below give the existence and uniqueness for the solution to the equation $f(x) = z$, which deduce the proof of existence and uniqueness of solutions to the implicit schemes.

Lemma 3.2.1. *Let f be a vector field on $\mathbb{R}^d$ and consider the equation*

$$f(x) = z \tag{3.2.3}$$

for a given $z \in \mathbb{R}^d$. If f is monotone, i.e.,

$$(x - y)(f(x) - f(y)) > 0 \tag{3.2.4}$$

for all $x, y \in \mathbb{R}^d$, $x \neq y$, then the equation has at most one solution.

If f is continuous and it is "coercive", i.e., there exist constants $c_1 > 0$, and $c_2 \in \mathbb{R}^d$, such that

$$xf(x) \ge c_1|x|^2 + c_2, \ \forall\ x \in \mathbb{R}^d. \tag{3.2.5}$$

then for every $z \in \mathbb{R}^d$, the equation has a solution $x \in \mathbb{R}^d$, and

$$|x|^2 \le \frac{1}{c_1^2}|z|^2 - \frac{2c_2}{c_1} \tag{3.2.6}$$

with constants depending only on c_1 and c_2.

Proof. If x and y are solutions to equation (3.2.3), then $f(x) - f(y) = 0$. Hence $(x - y)(f(x) - f(y)) = 0$ and condition (3.2.4) implies $x = y$.

If f is continuous and satisfies the "coercivity condition", then the existence of a solution $x \in \mathbb{R}^d$ is a classical result, see, e.g. Zeidler's book [40]. To show the estimate (3.2.6) for a solution x of equation (3.2.3), notice that the coercivity (3.2.5) implies

$$c_1|x|^2 + c_2 \leq xf(x) = xz \leq \frac{c_1}{2}|x|^2 + \frac{1}{2c_1}|z|^2.$$

Hence,

$$|x|^2 \leq \frac{1}{c_1^2}|z|^2 - \frac{2c_2}{c_1},$$

which proves (3.2.6). □

Corollary 3.2.2. *Let a be a continuous vector field on $\mathbb{R}^d$ satisfying local monotonicity and one-sided linear growth conditions: 3.2.3 and 3.2.5. Then for Δt satisfying $C_1\Delta t < 1$, (where C_1 is the coefficient in one-sided linear growth condition) and for every $z \in \mathbb{R}^d$, the equation*

$$x - a(x)\Delta t = z \tag{3.2.7}$$

has a solution $x \in \mathbb{R}^d$ and

$$|x| \leq (1 - C_1\Delta t)^{-1}|z|.$$

If for $R := (1 - C_1\Delta t)^{-1}|z|$, and we have $L_R\Delta t < 1$, then the equation (3.2.7) has a unique solution.

Proof. Notice that $f(x) := x - a(x)\Delta t$ is a continuous vector field on $\mathbb{R}^d$, such that

$$xf(x) = |x|^2 - xa(x)\Delta t \geq |x|^2 - C_1\Delta t|x|^2 = c_1|x|^2,$$

with $c_1 := 1 - C_1\Delta t$.

Clearly, for sufficiently small Δt the constant c_1 is strictly positive and hence f is coercive.

Then by Lemma 3.2.1, (3.2.7) has a solution x and by estimate (3.2.6),

$$|x| \leq C|z|,$$

with $C = (1 - C_1\Delta t)^{-1}$. Let x, y be solutions to (3.2.7), then $|x| \leq R$, $|y| \leq R$, with $R := C(z)$, and by the local monotonicity condition

$$(x - y)(f(x) - f(y)) \geq |x - y|^2 - L_R|x - y|^2\Delta t = (1 - L_R\Delta t)|x - y|^2,$$

which implies $x = y$ if $L_R\Delta t < 1$. □

Corollary 3.2.3. *Let a be a continuous vector field on $\mathbb{R}^d$ satisfying the local monotonicity condition 3.2.3 and the one-sided linear growth condition 3.2.5. Then for Δt satisfying $C_1\Delta t < 1$ there is a Borel function $\rho : \mathbb{R}^d \to \mathbb{R}^d$ such that $x = \rho_{\Delta t}(z)$ is the solution of the equation (3.2.7), for all $z \in \mathbb{R}^d$, and $|\rho_{\Delta t}(z)| \leq (1 - C_1\Delta t)^{-1}|z|$.*

Proof. If $L_R = L$ for all R, i.e., the one-sided Lipschitz condition 3.2.6 holds, then for $L\Delta t < 1$ and $C_1\Delta t < 1$, equation (3.2.7) has a unique solution $x := \rho_{\Delta t}(z)$ for all z, which defines the Borel function $\rho_{\Delta t}$. If $C_1\Delta t < 1$, then there exists a solution x for each $z \in \mathbb{R}^d$, but there may be many solutions. In this case one knows that it is possible to pick up a solution $\rho_{\Delta t}(z)$, such that the function $\rho_{\Delta t}$ is Borel measurable. From the above corollary, $|\rho_{\Delta t}(z)| \leq (1 - C_1\Delta t)^{-1}|z|$. □

Under the above preparation, we go back to the system of equations (3.2.2).

Theorem 3.2.4. *Let Assumptions 3.2.1, 3.2.2 and 3.2.6 hold, then the system of equations (3.2.2) has a unique solution $\{x^n(t_k) : k = 1, \cdots, n\}$, if Δt is sufficiently small, i.e., if n is sufficiently large. Moreover, $x^n(t_k)$ is $\mathcal{F}_{t_k}$−measurable for any k.*

Proof. The system (3.2.2) can be written as

$$f(x^n(t_k)) = x^n(t_k) - a(x^n(t_k))\Delta t = z$$

where

$$z := x^n(t_{k-1}) + \sum_{j=1}^{d_1} b^j(x^n(t_{k-1}))\Delta w^j_{k-1}, k = 1, 2, \cdots, n.$$

Hence we can get the existence and uniqueness of the sequences $\{x^n(t_k) : k = 1, 2 \cdots, n\}$ by induction on k from Lemma 3.2.1. Notice that, the random variables $x^n(t_k)$ are $\mathcal{F}_{t_k}$−measurable.

□

Theorem 3.2.5. *Assume Assumptions 3.2.1, 3.2.2, 3.2.3 and 3.2.5 hold, then for sufficiently small Δt, the system of equations (3.2.2) admits a solution $\{x^n(t_k) : k = 1, 2, \cdots, n\}$ such that $x^n(t_k)$ is $\mathcal{F}_{t_k}$−measurable.*

Proof. Define

$$x^n(t_k) := \rho_{\Delta t}\big(x^n(t_{k-1}) + \sum_{j=1}^{d_1} b^j(x^n(t_{k-1}))\Delta w^j_{k-1}\big), \tag{3.2.8}$$

where $\rho_{\Delta t}$ is $x^n(t_k) - a(x^n(t_k))\Delta t$. Corollary 3.2.3 implies the result.

□

Remark 3.2.4. In the proof of the main theorem 3.2.7 given as below, we will define implicit approximations as (3.2.8). In the point of my practice, under the local conditions $\rho_{\Delta t}$ is bounded by an increasing sequence $C(R)$. Further explanation is given later in this chapter.

Our goal is to show that implicit approximations converge to a stochastic process, which is the solution of equation (3.2.1). The theorems below are the main results of this chapter. Moveover, we obtain a rate of convergence result for the implicit method which is a generalization of [5].

Theorem 3.2.6. *Let $r \geq 1$ and $q \geq 1$ be any real number. Assume $E|x_0|^{qr} < \infty$. Then under Assumptions 3.2.2, 3.2.6 and 3.2.7, there exists a constant C independent of n, such that*

$$E \max_{0<k\leq n} |x^n(t_k) - x(t_k)|^q < Cn^{-q/4}.$$

We can weaken the assumptions of Theorem 3.2.6 as follows:

Theorem 3.2.7. *Under Assumptions 3.2.1 and 3.2.2, let Assumptions 3.2.3, 3.2.4 and 3.2.5 hold. Then $x^n(t)$ defined by (3.2.2), converges to $x(t)$ almost surely for each $t \in [0,T]$. Moreover, for every $\alpha < \frac{1}{4}$, there exists a finite random variable ξ, such that almost surely*

$$\sup_{t\leq T} |x(t) - x^n(t)| \leq \xi n^{-\alpha},$$

for all $n \geq 1$.

3.3 Preliminary Lemmas

In order to establish the main results, we shall need to prepare some lemmas. We will show that under Assumption 3.2.5, the true solution of SDE (3.2.1) has a finite qth moments for each $q \geq 1$, and the qth moments of the numerical solutions is bounded by some constants independent of n.

Lemma 3.3.1. *Let $q \geq 1$. Assume $E|x_0|^q < \infty$. Under Assumptions 3.2.2 and 3.2.5, if $x(t)$ is a solution of SDE (3.2.1), then there exists a constant C, such that*

$$E \sup_{0\leq t\leq T} |x_t|^q \leq C.$$

This statement is well-known, see e.g. in Krylov [22]. For the convenience of the reader, we give a brief proof.

Proof. By Itô's formula, we can derive that for all $t \in [0,T]$,

$$\begin{aligned} d|x(t)|^2 &= (2x(t)a(x(t)) + \sum_{j=1}^{d_1} |b^j(x(t))|^2)dt + \sum_{j=1}^{d_1} 2x(t)b^j(x(t))dw_t^j \\ &\le C(1+|x(t)|^2)dt + \sum_{j=1}^{d_1} 2x(t)b^j(x(t))dw_t^j, \end{aligned}$$

where $C = C(C_1, C_2, C_3)$.

Set $x^R(t) := x(t \wedge \tau_R)$, where τ_R is a stopping time defined by

$$\tau_R := \inf\{t \ge 0 : |x_t| \ge R\} \text{ for each } R > 0.$$

Then

$$\begin{aligned} \sup_{s\le t} |x^R(s)|^2 \mathbf{1}_{\tau_R>0} &\le |x_0|^2 \mathbf{1}_{\tau_R>0} + C\int_0^{t\wedge\tau_R} (1+|x(s)|^2)ds \\ &\quad +2\sup_{s\le t} |\int_0^{s\wedge\tau_R} \sum_{j=1}^{d_1} x(r)b^j(x(r))dw_r^j|. \end{aligned}$$

Hence for any $q \ge 1$,

$$\begin{aligned} & E\sup_{s\le t} |x^R(s)|^q \mathbf{1}_{\tau_R>0} \\ \le\ & C_q\Big(E|x_0|^q + t^{q/2-1}E\int_0^t (1+|x^R(s)|^q)ds \\ & +E\sup_{s\le t} |\int_0^s \sum_{j=1}^{d_1} x^R(r)b^j(x^R(r))dw_r^j|^{q/2}\Big). \end{aligned} \tag{3.3.9}$$

Notice that the right-hand side of the inequality (3.3.9) is finite since $|x^R(s)| \le R$ and

$$\begin{aligned} I &:= E\sup_{s\le t} |\int_0^s \sum_{j=1}^{d_1} x^R(r)b^j(x^R(r))dw_r^j|^{q/2} \\ &\le C_qE(\int_0^t \sum_{j=1}^{d_1} |x^R(r)|^2|b^j(x^R(r))|^2dr)^{q/4} \\ &< \infty, \end{aligned}$$

by the Burkholder-Davis-Gundy inequality.

Moveover, by the Cauchy-Schwarz and the Young's inequalities, also let $C_q = C_q(C_1, C_2, C_3, q)$ be a constant that may change line by line

$$\begin{aligned} I &\le C_qE\Big(\sup_{s\le t} |x^R(s)|^{3q/4}(\int_0^t (1+|x^R(r)|)dr)^{q/4}\Big) \\ &\le \frac{3}{4}E\sup_{s\le t} |x^R(s)|^q + C_qt^{q-1}E\int_0^t (1+|x^R(r)|^q)dr \\ &< \infty. \end{aligned}$$

Thus, with (3.3.9), we have

$$\frac{1}{4}E\sup_{s\le t}|x^R(s)|^q\mathbf{1}_{\tau_R>0}\le C_qE|x_0|^q+C_qt^q+C_qt^{q-1}E\int_0^t\sup_{r\le s}|x_r^R|^q\mathbf{1}_{\tau_R>0}dr,$$

for every $t\in[0,T]$.

By applying the Gronwall's inequality 2.3.7, we get

$$E\sup_{0\le s\le t}|x^R(s)|^q\mathbf{1}_{\tau_R>0}\le C_q(E|x_0|^q+t^q)e^{C_qt^{q-1}}. \tag{3.3.10}$$

Hence

$$E\sup_{0\le s\le t}|x^R(s)|^q\le E\sup_{0\le s\le t}|x^R(s)|^q\mathbf{1}_{\tau_R>0}+E|x_0|^q. \tag{3.3.11}$$

Since $\tau_R\to\infty$ for $R\to\infty$, we have $x^R(s)\to x(s)$, then consequently,

$$\sup_{0\le s\le t}|x(s)|^q=\sup_{0\le s\le t}\liminf_{R\to\infty}|x^R(s)|^q\le\liminf_{R\to\infty}\sup_{0\le s\le t}|x^R(s)|^q.$$

Hence, by Fatou's Lemma, together with (3.3.10), (3.3.11)

$$\begin{aligned}E\sup_{0\le s\le t}|x(s)|^q &\le E\Big(\liminf_{R\to\infty}\sup_{0\le s\le t}|x^R(s)|^q\Big)\\ &\le \liminf_{R\to\infty}E\Big(\sup_{0\le s\le t}|x^R(s)|^q\Big)\mathbf{1}_{\tau_R>0}+E|x_0|^q\\ &\le C_q,\end{aligned}$$

where $C_q:=C_q(q,E|x_0|^q,C_1,C_2,C_3,T)$. The assertion is proved. □

The next lemma provides important bound for implicit approximations.

Lemma 3.3.2. *Let Assumptions 3.2.2 and 3.2.5 hold. Let $q\ge1$ be any real number, and assume that $E|x_0|^q<\infty$. Then there exists a constant C independent of n, such that the solutions of the system of equations (3.2.2) satisfy*

$$E\max_{0\le k\le n}|x^n(t_k)|^q\le C,$$

for all sufficiently large integer n.

Proof. **Step 1**. First we prove that $E|x^n(t_k)|^q<\infty$ for each n and k.

For fixed n we proceed with the proof by induction in $k=0,1,\cdots,n$. For $k=0$, we have $E|x_0|^q<\infty$ by assumption. Assume that $E|x^n(t_{k-1})|^q<\infty$, for $1<k<n$. We want to show that

$$E|x^n(t_k)|^q<\infty.$$

Take a constant $K > 0$, for $l = 1, 2, \cdots, n$, and notice that

$$\begin{aligned}
& e^{-Kt_l}|x^n(t_l)|^2 - e^{-Kt_{l-1}}|x^n(t_{l-1})|^2 \\
= \; & e^{-Kt_{l-1}}(|x^n(t_l)|^2 - |x^n(t_{l-1})|^2) + (e^{-Kt_l} - e^{-Kt_{l-1}})|x^n(t_l)|^2 \\
\leq \; & e^{-Kt_{l-1}}(|x^n(t_l)|^2 - |x^n(t_{l-1})|^2) - C_n K e^{-Kt_{l-1}}|x^n(t_l)|^2 \Delta t,
\end{aligned}$$

where

$$C_n := \frac{1 - e^{-K\Delta t}}{K\Delta t} > \frac{1}{2}, \quad \text{for sufficiently small } \Delta t.$$

Then

$$\begin{aligned}
& e^{-Kt_l}|x^n(t_l)|^2 - e^{-Kt_{l-1}}|x^n(t_{l-1})|^2 \\
\leq \; & e^{-Kt_{l-1}}(|x^n(t_l)|^2 - |x^n(t_{l-1})|^2) - \frac{1}{2} K e^{-Kt_{l-1}}|x^n(t_l)|^2 \Delta t.
\end{aligned}$$

Notice that from (3.2.2) we have

$$x^n(t_l) - x^n(t_{l-1}) = a(x^n(t_l))\Delta t + \sum_{j=1}^{d_1} b^j(x^n(t_{l-1}))\Delta w_{l-1}^j.$$

Furthermore, by using the formula $b^2 - a^2 = 2b(b-a) - (b-a)^2$, we get,

$$\begin{aligned}
& |x^n(t_l)|^2 - |x^n(t_{l-1})|^2 \\
= \; & 2x^n(t_l)(x^n(t_l) - x^n(t_{l-1})) - |x^n(t_l) - x^n(t_{l-1})|^2 \\
= \; & 2x^n(t_l)a(x^n(t_l))\Delta t + \sum_{j=1}^{d_1} 2x^n(t_l)b^j(x^n(t_{l-1}))\Delta w_{l-1}^j - |a(x^n(t_l))|^2|\Delta t|^2 \\
& - \sum_{j=1}^{d_1} |b^j(x^n(t_{l-1}))|^2|\Delta w_{l-1}^j|^2 - \sum_{j=1}^{d_1} 2a(x^n(t_l))b^j(x^n(t_{l-1}))\Delta t \Delta w_{l-1}^j \\
= \; & 2x^n(t_l)a(x^n(t_l))\Delta t + \sum_{j=1}^{d_1} 2\big(x^n(t_l) - x^n(t_{l-1})\big)b^j(x^n(t_{l-1}))\Delta w_{l-1}^j \\
& + \sum_{j=1}^{d_1} 2x^n(t_{l-1})b^j(x^n(t_{l-1}))\Delta w_{l-1}^j - |a(x^n(t_l))|^2|\Delta t|^2 \\
& - \sum_{j=1}^{d_1} |b^j(x^n(t_{l-1}))|^2|\Delta w_{l-1}^j|^2 - \sum_{j=1}^{d_1} 2a(x^n(t_l))b^j(x^n(t_{l-1}))\Delta t \Delta w_{l-1}^j \\
= \; & 2x^n(t_l)a(x^n(t_l))\Delta t + \sum_{j=1}^{d_1} 2x^n(t_{l-1})b^j(x^n(t_{l-1}))\Delta w_{l-1}^j - |a(x^n(t_l))|^2|\Delta t|^2 \\
& + \sum_{j=1}^{d_1} |b^j(x^n(t_{l-1}))|^2|\Delta w_{l-1}^j|^2 \\
\leq \; & 2x^n(t_l)a(x^n(t_l))\Delta t + \sum_{j=1}^{d_1} 2x^n(t_{l-1})b^j(x^n(t_{l-1}))\Delta w_{l-1}^j
\end{aligned}$$

$$+\sum_{j=1}^{d_1} |b^j(x^n(t_{l-1}))|^2 |\Delta w_{l-1}^j|^2.$$

Then

$$\begin{aligned}
& e^{-Kt_l}|x^n(t_l)|^2 - e^{-Kt_{l-1}}|x^n(t_{l-1})|^2 \\
\leq\ & e^{-Kt_{l-1}}\Big\{(2x^n(t_l)a(x^n(t_l))\Delta t + \sum_{j=1}^{d_1} 2x^n(t_{l-1})b^j(x^n(t_{l-1}))\Delta w_{l-1}^j \\
& +\sum_{j=1}^{d_1} |b^j(x^n(t_{l-1}))|^2 |\Delta w_{l-1}^j|^2)\Big\} - \frac{1}{2}Ke^{-Kt_{l-1}}|x^n(t_l)|^2\Delta t \\
\leq\ & e^{-Kt_{l-1}}C_1(1+|x^n(t_l)|^2)\Delta t + e^{-Kt_{l-1}}\sum_{j=1}^{d_1} 2x^n(t_{l-1})b^j(x^n(t_{l-1}))\Delta w_{l-1}^j \\
& +e^{-Kt_{l-1}}\sum_{j=1}^{d_1} |b^j(x^n(t_{l-1}))|^2 |\Delta w_{l-1}^j|^2 - \frac{1}{2}Ke^{-Kt_{l-1}}|x^n(t_l)|^2\Delta t,
\end{aligned}$$

by making use of Assumption 3.2.5. For $K > 2C_1$, we obtain

$$\begin{aligned}
& e^{-Kt_{l-1}}|x^n(t_l)|^2 - e^{-Kt_{l-1}}|x^n(t_{l-1})|^2 \\
\leq\ & C_1 e^{-Kt_l}\Delta t + \sum_{j=1}^{d_1} 2e^{-Kt_{l-1}}x^n(t_{l-1})b^j(x^n(t_{l-1}))\Delta w_{l-1}^j \\
& +\sum_{j=1}^{d_1} e^{-Kt_{l-1}}|b^j(x^n(t_{l-1}))|^2 |\Delta w_{l-1}^j|^2.
\end{aligned}$$

By

$$e^{-Kt_k}|x^n(t_k)|^2 - e^{-Kt_0}|x^n(t_0)|^2 = \sum_{l=1}^{k} \left(e^{-Kt_l}|x^n(t_l)|^2 - e^{-Kt_{l-1}}|x^n(t_{l-1})|^2\right),$$

summing up, we deduce that

$$\begin{aligned}
& e^{-Kt_k}|x^n(t_k)|^2 \\
\leq\ & |x_0|^2 + C_1\sum_{l=1}^{k} e^{-Kt_{l-1}}\Delta t + \sum_{j=1}^{d_1}\sum_{l=1}^{k} e^{-Kt_{l-1}}|b^j(x^n(t_{l-1}))|^2 |\Delta w_{l-1}^j|^2 \\
& +2\sum_{j=1}^{d_1}\sum_{l=1}^{k} e^{-Kt_{l-1}}x^n(t_{l-1})b^j(x^n(t_{l-1}))\Delta w_{l-1}^j \\
\leq\ & |x_0|^2 + C_1\Delta t\frac{e^{-Kt_0}}{1-e^{-K\Delta t}} + \sum_{j=1}^{d_1}\sum_{l=1}^{k} e^{-Kt_{l-1}}|b^j(x^n(t_{l-1}))|^2 |\Delta w_{l-1}^j|^2 \\
& +2\sum_{j=1}^{d_1}\sum_{l=1}^{k} e^{-Kt_{l-1}}x^n(t_{l-1})b^j(x^n(t_{l-1}))\Delta w_{l-1}^j. \qquad (3.3.12)
\end{aligned}$$

For sufficiently small Δt, $0 < \frac{e^{-Kt_0}}{1-e^{-K\Delta t}} \to 2e^{-Kt_0} = 2$, hence

$$\begin{aligned} & e^{-Kt_k}|x^n(t_k)|^2 \\ \leq\ & e^{-Kt_{k-1}}|x^n(t_{k-1})|^2 + \sum_{j=1}^{d_1} e^{-Kt_{k-1}}|b^j(x^n(t_{k-1}))|^2|\Delta w_{k-1}^j|^2 \\ & + \sum_{j=1}^{d_1} 2e^{-Kt_{k-1}} x^n(t_{k-1}) b^j(x^n(t_{k-1}))\Delta w_{k-1}^j. \end{aligned}$$

Then raising both sides to the power $q/2$, we take expectations to give

$$\begin{aligned} & Ee^{-\frac{1}{2}qKt_k}|x^n(t_k)|^q \\ \leq\ & C_q\Big\{E(e^{-\frac{1}{2}qKt_{k-1}}|x^n(t_{k-1})|^q) \\ & + \sum_{j=1}^{d_1} Ee^{-\frac{1}{2}qKt_{k-1}}|b^j(x^n(t_{k-1}))|^q|\Delta w_{k-1}^j|^q \\ & + \sum_{j=1}^{d_1} 2^{q/2}Ee^{-\frac{1}{2}qKt_{k-1}}|x^n(t_{k-1})|^{q/2}|b^j(x^n(t_{k-1}))|^{q/2}|\Delta w_{k-1}^j|^{q/2}\Big\}, \end{aligned}$$

Notice that, by the linear growth condition of b,

$$\begin{aligned} & \sum_{j=1}^{d_1} E|b^j(x^n(t_{k-1}))|^q|\Delta w_{k-1}^j|^q \\ \leq\ & \sum_{j=1}^{d_1} E|b^j(x^n(t_{k-1}))|^q E|\Delta w_{k-1}^j|^q \\ \leq\ & C_3(1+E|x^n(t_{k-1})|^q)(\Delta t)^{q/2}, \end{aligned}$$

and

$$\begin{aligned} & \sum_{j=1}^{d_1} E|x^n(t_{k-1})|^{q/2}|b^j(x^n(t_{k-1}))|^{q/2}|\Delta w_{k-1}^j|^{q/2} \\ \leq\ & C_3(1+E|x^n(t_{k-1})|^q)(\Delta t)^{q/4}. \end{aligned}$$

Both are finite by assumption $E|x^n(t_{k-1})|^q < \infty$, which proves that

$$E|x^n(t_k)|^q < \infty,$$

for $k = 1, 2, \cdots, n$ and finish the induction proof.

Step 2 Now we would like to prove the statement of $E\max_{k\leq n}|x^n(t_k)|^q \leq C$.

From (3.3.12), we raise to the power $q/2$, for $k \leq i \leq n$, then take expectations to get

$$E\max_{0\leq k\leq i} e^{-\frac{qK}{2}t_k}|x^n(t_k)|^q \leq C_q\{E|x_0|^q + (2C_1\Delta t)^{q/2} + I_1 + 2^{q/2}I_2\}, \qquad (3.3.13)$$

with

$$I_1 := \sum_{j=1}^{d_1} E \max_{0\le k\le i} \big(\sum_{l=1}^{k} e^{-Kt_{l-1}} |b^j(x^n(t_{l-1}))|^2 |\Delta w_{l-1}^j|^2\big)^{q/2},$$

and

$$I_2 := \sum_{j=1}^{d_1} E \max_{0\le k\le i} \big(\sum_{l=1}^{k} e^{-Kt_{l-1}} x^n(t_{l-1}) b^j(x^n(t_{l-1}))\Delta w_{l-1}^j\big)^{q/2}.$$

Notice that $I_1 \le C_q(I_1' + I_1'')$, with

$$I_1' := \sum_{j=1}^{d_1} E \max_{0\le k\le i} \big(\sum_{l=1}^{k} e^{-Kt_{l-1}} |b^j(x^n(t_{l-1}))|^2 \Delta t\big)^{q/2};$$

$$I_1'' := \sum_{j=1}^{d_1} E \max_{0\le k\le i} \big(\sum_{l=1}^{k} e^{-Kt_{l-1}} |b^j(x^n(t_{l-1}))|^2 (|\Delta w_{l-1}^j|^2 - \Delta t)\big)^{q/2}$$

Estimation of I_1':

$$\begin{aligned}
I_1' &\le \sum_{j=1}^{d_1} T^{q/2-1} \sum_{l=1}^{i} e^{-\frac{1}{2}Kqt_{l-1}} (E|b^j(x^n(t_{l-1}))|^2)^{q/2}\Delta t \\
&\le C_3 T^{q/2-1} \sum_{l=1}^{i} e^{-\frac{1}{2}qKt_{l-1}} E(1+|x^n(t_{l-1}))|^2)^{q/2}\Delta t \\
&\le 2C_3T^{q/2-1} + CT^{q/2-1}\sum_{l=0}^{i-1} E \max_{0\le k\le l} e^{-\frac{1}{2}qKt_k}|x^n(t_k)|^q \Delta t.
\end{aligned}$$

Estimation of I_1'':

$$I_1' := \sum_{j=1}^{d_1} E \max_{k\le i} \big(\sum_{l=1}^{k} e^{-Kt_{l-1}} |b^j(x^n(t_{l-1}))|^2 (|\Delta w_{l-1}^j|^2 - \Delta t)\big)^{q/2}$$

To estimate I_1'' we define a stochastic process $\{m_n(t) : t \in [0,T]\}$ by

$$m_n(t) = \sum_{j=1}^{d_1} \int_0^t g_n(s) dw_s^j,$$

for each $n \ge 1$, where $g_n(t) := 2(w_t - w_{t_{l-1}})$ for $t \in [t_{l-1}, t_l]$, $l = 1, 2, \cdots, n$. Then m_n is an $\mathcal{F}_t$-martingale for each n, and

$$m_n(t_l) - m_n(t_{l-1}) = \sum_{j=1}^{d_1} |\Delta w_{l-1}^j|^2 - \Delta t.$$

Set $\kappa_1(t) := t_{l-1}$, when $t \in [t_{l-1}, t_l]$. Then

$$\sum_{l=1}^{k} e^{-Kt_{l-1}} |b(x^n(t_{l-1}))|^2 (|\Delta w_{l-1}|^2 - \Delta t) = \sum_{j=1}^{d_1} \int_0^{t_k} e^{-K\kappa_1(t)} |b^j(x^n(\kappa_1(t)))|^2 g_n(t) dw_t^j.$$

Hence by the Burkholder-Davis-Gundy inequality,

$$\begin{aligned} I_1'' &\leq C_q \sum_{j=1}^{d_1} E\{ \int_0^{t_i} e^{-2K\kappa_1(t)} |b^j(x^n(\kappa_1(t)))|^4 g_n^2(t) dt \}^{q/4} \\ &\leq C_q E\{ \int_0^{t_i} e^{-2K\kappa_1(t)} (1 + |x^n(\kappa_1(t))|^4) g_n^2(t) dt \}^{q/4}. \end{aligned}$$

By using Young's inequality: for $\frac{1}{q} + \frac{1}{p} = 1$, with $p = \frac{4}{3}$, $q = 4$, we obtain,

$$\begin{aligned} & I_1'' \\ \leq\ & C_q E\{ \sup_{t \leq t_i} e^{-\frac{3}{8}Kq\kappa_1(t)} (1 + |x^n(\kappa_1(t))|)^{3q/4} (\int_0^{t_i} e^{-\frac{1}{2}K\kappa_1(t)} (1 + |x^n(\kappa_1(t))|) g_n^2(t) dt)^{q/4} \} \\ \leq\ & \frac{3}{4} E \sup_{t \leq t_i} e^{-\frac{1}{2}Kq\kappa_1(t)} (1 + |x^n(\kappa_1(t))|)^q \\ & + \frac{1}{4} CE\{ \int_0^{t_i} e^{-\frac{1}{2}K\kappa_1(t)} (1 + |x^n(\kappa_1(t))|) g_n^2(t) dt \}^{4q/4} \\ \leq\ & \frac{3}{4} E \max_{k \leq i} e^{-\frac{1}{2}Kqt_k} (1 + |x^n(t_k)|^q) + \frac{1}{4} C J_n, \end{aligned}$$

with

$$J_n := E \int_0^{t_i} e^{-\frac{1}{2}K\kappa_1(t)} (1 + |x^n(\kappa_1(t))|) g_n^2(t) dt \}^q,$$

where

$$\begin{aligned} J_n &\leq T^{q-1} E \int_0^{t_i} e^{-\frac{1}{2}Kq\kappa_1(t)} (1 + |x^n(\kappa_1(t))|^q) g_n^{2q}(t) dt \\ &\leq \sum_{j=1}^{d_1} C T^{q-1} E \int_0^{t_i} e^{-\frac{1}{2}Kq\kappa_1(t)} (1 + |x^n(\kappa_1(t))|^q) |w_t^j - w_{\kappa_1(t)}^j|^{2q} dt \\ &\leq \sum_{j=1}^{d_1} C T^{q-1} \sum_{l=1}^{i} \int_{t_{l-1}}^{t_l} E e^{-\frac{1}{2}Kqt_{l-1}} (1 + |x^n(t_{l-1})|^q) |w_t^j - w_{t_{l-1}}^j|^{2q} dt. \end{aligned}$$

Since

$$\sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} E|w_t^j - w_{t_{l-1}}^j|^{2q} dt \leq C \Delta t^{q+1},$$

$$J_n \leq C \sum_{l=1}^{i} E e^{-\frac{1}{2}Kqt_{l-1}} \Delta t + C \sum_{l=1}^{i} E e^{-\frac{1}{2}Kqt_{l-1}} |x(t_{l-1})|^q \Delta t.$$

Hence we get

$$\begin{aligned} I_1'' &\leq \frac{3}{4} E \max_{k \leq i} e^{-\frac{1}{2}Kqt_k} (1 + |x^n(t_k)|^q) \\ &\quad + \frac{1}{4} C (\sum_{l=1}^{i} E e^{-\frac{1}{2}Kqt_{l-1}} \Delta t + \sum_{l=1}^{i} E e^{-\frac{1}{2}Kqt_{l-1}} |x(t_{l-1})|^q \Delta t) \\ &\leq \frac{3}{4} E \max_{k \leq i} e^{-\frac{1}{2}Kqt_k} |x^n(t_k)|^q + \frac{1}{4} C \sum_{l=0}^{i-1} E \max_{k \leq l} e^{-\frac{1}{2}Kqt_k} |x(t_k)|^q \Delta t \\ &\quad + C e^{-\frac{1}{2}Kqt_0}. \end{aligned}$$

where C is a constant changing line by line.

Combining I_1', I_1'', we obtain

$$I_1 \le \frac{3}{4} E \max_{k\le i} e^{-\frac{1}{2}Kqt_k}|x^n(t_k)|^q + Ce^{-\frac{1}{2}qKt_0} + C\sum_{l=0}^{i-1} E\max_{k\le l} e^{-\frac{1}{2}qKt_k}|x^n(t_k)|^q \Delta t,$$

where $C = C(q, T, E|x_0|^q, C_1, C_2, C_3)$.

Estimation of I_2: Also, by applying the Burkholder-Davis-Gundy and Young's inequalities, we get

$$\begin{aligned}
I_2 &= \sum_{j=1}^{d_1} E\max_{k\le i}\Big(\sum_{l=1}^{k} e^{-Kt_{l-1}} x^n(t_{l-1}) b^j(x^n(t_{l-1}))\Delta w_{l-1}^j\Big)^{q/2} \\
&\le C_q E\Big(\int_0^{t_i} e^{-2K\kappa_1(t)}(1+|x^n(\kappa_1(t))|^4)dt\Big)^{q/4} \\
&\le \frac{3}{4} E\sup_{t\le t_i} e^{-\frac{1}{2}Kq\kappa_1(t)}(1+|x^n(\kappa_1(t))|)^q + \frac{1}{4}C\Big\{\int_0^{t_i} e^{-\frac{1}{2}K\kappa_1(t)}(1+|x^n(\kappa_1(t))|)dt\Big\}^{4q/4} \\
&\le \frac{3}{4} E\max_{k\le i} e^{-\frac{1}{2}Kqt_k}|x^n(t_k)|^q + Ce^{-\frac{1}{2}qKt_0} + C\sum_{l=0}^{i-1} E\max_{k\le l} e^{-\frac{1}{2}Kqt_k}|x^n(t_k)|^q\Delta t.
\end{aligned}$$

Together with I_1, putting into (3.3.13), again with a possible different constant $C := C(q, T, E|x_0|^q, C_1, C_2, C_3)$, we obtain,

$$\begin{aligned}
&E\max_{k\le i} e^{-\frac{1}{2}qKt_k}|x^n(t_k)|^q \\
\le\; & C + C\sum_{l=0}^{i-1} E\max_{k\le l} e^{-\frac{1}{2}qKt_k}|x(t_k)|^q\Delta t.
\end{aligned}$$

Consequently by the discrete Gronwall's Lemma (See Theorem 2.3.8), we have

$$\begin{aligned}
&E\max_{k\le i} e^{-\frac{1}{2}qKt_k}|x^n(t_k)|^q \\
\le\; & C(1+C\Delta t)^i \\
\le\; & Ce^{CT}.
\end{aligned}$$

So for sufficiently small Δt, we get $E\max_{k\le i} e^{-\frac{1}{2}qKt_k}|x^n(t_k)|^q \le Ce^{CT}$, where C is a constant dependent of q, T, C_1, C_2, C_3 and $E|x_0|^q$, independent of n.

Hence $e^{-\frac{1}{2}qKT}E\max_{k\le i}|x^n(t_k)|^q \le Ce^{CT}$, i.e., $E\max_{k\le i}|x^n(t_k)|^q \le Ce^{c_1T}$, where $c_1 = c_1(q, K)$. □

Remark 3.3.1. In the work of Higham, Mao and Stuart [5], when they begin to bound the moments of the numerical solutions by the split-step backward Euler method, they firstly deal with the estimates of the discrete approximations, (see Lemma 3.7, [5]). They acquire a fine boundedness i.e., $E\sup_{0\le N\le M}|Y_N|^{2p} \le$

Ce^{CT}, where $C := C(p,T)$. These estimate seems right to us, but in fact here the constant C depends on M as well. So this result is not useful. Indeed, in the proof of their lemma, (see P1052, line 4, [5]) they apply an inequality of

$$(\sum_{j=0}^{N-1} |g(Y_j^*)\Delta W_j|^2)^p \leq N^{p-1} \sum_{j=0}^{N-1} |g(Y_j^*)\Delta W_j|^{2p}.$$

This is not accurate, since N goes to infinity as $\Delta t \to 0$. In our proof, I_1 is not a well-defined stochastic integral, which means Burkholder-Davis-Gundy's inequality can not be applied directly. In order to overcome this problem, we split $|\Delta w_t|^2$ into two terms, i.e., Δt and $|\Delta w_t|^2 - \Delta t$. Obviously, Δt part I_1' is easy to estimate, while regarding I_1'' part, it is natural for us to define a stochastic process m_n. Apparently m_n is an $\mathcal{F}_t$-martingale for each n. Therefore, the problem can be solved.

Recall that for $0 < k \leq n$ and $t \in [t_{k-1}, t_k)$, we set

$$\kappa_1(t) := t_{k-1}, \text{ and } \kappa_2(t) := t_k, \text{ for } t \in [t_{k-1}, t_k).$$

Then equation (3.2.2) can be cast in the integer form

$$x^n(t) = x^n(t_0) + \int_0^t a(x^n(\kappa_2(s)))ds + \sum_{j=1}^{d_1} \int_0^t b^j(x^n(\kappa_1(s)))dw_s^j. \tag{3.3.14}$$

We also give the following useful estimates that will lead to the proof of our main theorem.

Lemma 3.3.3. *Let $q \geq 1$ and $r \geq 1$ be any real number. Assume $E|x_0|^{qr} < \infty$. Then under Assumptions 3.2.5 and 3.2.7, there exists a constant C independent of n, such that*

$$E|x(t) - x(\kappa_1(t))|^q \leq Cn^{-q/2}.$$

Proof. By the inequality $|a+b|^q \leq 2^{q-1}(|a|^q + |b|^q)$, it is easy to see that

$$E|x(t) - x(\kappa_1(t))|^q \leq 2^{q-1}E|\int_{\kappa_1(t)}^t a(x(s))ds|^q + \sum_{j=1}^{d_1} 2^{q-1}E|\int_{\kappa_1(t)}^t b^j(x(s))dw_s^j|^q.$$

Using the Burkholder-Davis-Gundy's inequality, one can derive that

$$E|x(t) - x(\kappa_1(t))|^q \leq I_1 + I_2,$$

with

$$I_1 := 2^{q-1}\Delta t^{q-1} E \int_{\kappa_1(t)}^t |a(x(s))|^q ds;$$

and

$$I_2 := \sum_{j=1}^{d_1} c_q 2^{q-1} E\Big(\int_{\kappa_1(t)}^t |b^j(x(s))|^2 ds\Big)^{q/2}.$$

Clearly,

$$\begin{aligned} I_1 &\le C_4 2^{q-1} \Delta t^{q-1} E \int_{\kappa_1(t)}^t (1+|x(s)|^r)^q ds \\ &\le C_4 2^{q-1} \Delta t^{q-1} \int_{\kappa_1(t)}^t (1+E|x(s)|^{rq}) ds \\ &\le Cn^{-q}, \end{aligned}$$

by the boundedness of $E|x(t)|^{qr}$, where $C := C(r,q,C_4,T)$.

Now

$$\begin{aligned} I_2 &\le C_3 c_q 2^{q-1} E\Big(\int_{\kappa_1(t)}^t (1+|x(s)|^2) ds\Big)^{q/2} \\ &\le C\Delta t^{q/2} + CE\Big(\int_{\kappa_1(t)}^t |x(s)|^2 ds\Big)^{q/2}. \end{aligned}$$

Note that

$$\begin{aligned} & E\Big(\int_{\kappa_1(t)}^t |x(s)|^2 ds\Big)^{q/2} \\ \le\ & E\{\sup_{s\le t} |x(s)|^{q/2} \Big(\int_{\kappa_1(t)}^t |x(s)| ds\Big)^{q/2}\} \\ \le\ & E(\sup_{s\le t} |x(s)|^q)^{1/2} \{E\Big(\int_{\kappa_1(t)}^t |x(s)| ds\Big)^q\}^{1/2} \\ \le\ & C(\Delta t)^{(q-1)/2} \{E\int_{\kappa_1(t)}^t |x(s)|^q ds\}^{1/2} \\ \le\ & C(\Delta t)^{(q-1)/2} \Delta t^{1/2} (E|x(s)|^q)^{1/2} \\ \le\ & Cn^{-q/2}. \end{aligned}$$

Hence we obtain

$$E|x(t) - x(\kappa_1(t))|^q \le Cn^{-q/2},$$

where C is a constant independent of n. □

Remark 3.3.2. By using the similar method, we can get the following estimates as well:

$$E|x(t) - x(\kappa_2(t))|^q \le Cn^{-q/2},$$

$$E|x^n(t) - x^n(\kappa_1(t))|^q \le Cn^{-q/2},$$

$$E|x^n(t) - x^n(\kappa_2(t))|^q \le Cn^{-q/2}.$$

3.4 Proof of the Main Results

Now under the assumptions described above and the additional assumption that a is polynomial, it is possible to prove the main result of Theorem 3.2.6.

Proof of Theorem 3.2.6. **Step 1**. The first step of proving the theorem is to find a formula for $|x(t) - x^n(t)|^2$. We would like to use the formula $b^2 - a^2 = 2b(b-a) - (b-a)^2$ again.

For the equation (3.2.1), it is natural to define $x(t)$ on $[t_{l-1}, t_l)$, which is given by:

$$x(t_l) = x(t_{l-1}) + \int_{t_{l-1}}^{t_l} a(x(s))ds + \sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} b^j(x(s))dw_s^j.$$

With equation (3.2.2), we get the difference,

$$x(t_l) - x^n(t_l)$$

$$= x(t_{l-1}) - x^n(t_{l-1}) + \int_{t_{l-1}}^{t_l} (a(x(s)) - a(x^n(t_l)))ds + \sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} (b^j(x(s)) - b^j(x(t_{l-1})))dw_s^j \tag{3.4.15}$$

Notice that,

$$\begin{aligned} & |x(t_l) - x^n(t_l)|^2 - |x(t_{l-1}) - x^n(t_{l-1})|^2 \\ = \; & 2(x(t_l) - x^n(t_l))\{(x(t_l) - x^n(t_l)) - (x(t_{l-1}) - x^n(t_{l-1}))\} \\ & -\{(x(t_l) - x^n(t_l)) - (x(t_{l-1}) - x^n(t_{l-1})))\}^2. \end{aligned}$$

Putting equation (3.4.15) into it, after simple arithmetic computation, we could get

$$\begin{aligned} & |x(t_l) - x^n(t_l)|^2 - |x(t_{l-1}) - x^n(t_{l-1}))|^2 \\ = \; & 2(x(t_l) - x^n(t_l))\Big(\int_{t_{l-1}}^{t_l} (a(x(s)) - a(x^n(t_l)))ds + \sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} (b^j(x(s)) - b^j(x^n(t_{l-1}))dw_s^j\Big) \\ & -\Big(\int_{t_{l-1}}^{t_l} (a(x(s)) - a(x^n(t_l)))ds + \sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} (b^j(x(s)) - b^j(x^n(t_{l-1}))))dw_s^j\Big)^2 \\ = \; & 2(x(t_l) - x^n(t_l)) \int_{t_{l-1}}^{t_l} \Big(a(x(s)) - a(x^n(t_l))\Big)ds \\ & + \sum_{j=1}^{d_1} 2(x(t_l) - x^n(t_l)) \int_{t_{l-1}}^{t_l} \Big(b^j(x(s)) - b^j(x^n(t_{l-1}))\Big)dw_s^j \\ & -|\int_{t_{l-1}}^{t_l} \Big(a(x(s)) - a(x^n(t_l))\Big)ds|^2 - \sum_{j=1}^{d_1} |\int_{t_{l-1}}^{t_l} \Big(b^j(x(s)) - b^j(x^n(t_{l-1}))\Big)dw_s^j|^2 \\ & -2\int_{t_{l-1}}^{t_l} \Big(a(x(s)) - a(x^n(t_l))\Big)ds \sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} \Big(b^j(x(s)) - b^j(x^n(t_{l-1}))\Big)dw_s^j \end{aligned}$$

$$
\begin{aligned}
= \;& 2(x(t_l) - x^n(t_l)) \int_{t_{l-1}}^{t_l} \Big(a(x(s)) - a(x^n(t_l))\Big) ds \\
& + 2(x(t_{l-1}) - x^n(t_{l-1})) \sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} \Big(b^j(x(s)) - b^j(x^n(t_{l-1}))\Big) dw_s^j \\
& + 2\big((x(t_l) - x^n(t_l)) - (x(t_{l-1}) - x^n(t_{l-1}))\big) \sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} \Big(b^j(x(s)) - b^j(x^n(t_{l-1}))\Big) dw_s^j \\
& - \Big| \int_{t_{l-1}}^{t_l} \Big(a(x(s)) - a(x^n(t_l))\Big) ds \Big|^2 - \sum_{j=1}^{d_1} \Big| \int_{t_{l-1}}^{t_l} \Big(b^j(x(s)) - b^j(x^n(t_{l-1}))\Big) dw_s^j \Big|^2 \\
& - 2 \int_{t_{l-1}}^{t_l} \Big(a(x(s)) - a(x^n(t_l))\Big) ds \sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} \Big(b^j(x(s)) - b^j(x^n(t_{l-1}))\Big) dw_s^j \\
= \;& 2(x(t_l) - x^n(t_l)) \int_{t_{l-1}}^{t_l} \Big(a(x(s)) - a(x^n(t_l))\Big) ds \\
& + 2(x(t_{l-1}) - x^n(t_{l-1})) \sum_{j=1}^{d_1} \int_{t_{l-1}}^{t_l} \Big(b^j(x(s)) - b^j(x^n(t_{l-1}))\Big) dw_s^j \\
& + \sum_{j=1}^{d_1} \Big| \int_{t_{l-1}}^{t_l} \Big(b^j(x(s)) - b^j(x^n(t_{l-1}))\Big) dw_s^j \Big|^2 - \Big| \int_{t_{l-1}}^{t_l} \Big(a(x(s)) - a(x^n(t_l))\Big) ds \Big|^2.
\end{aligned}
$$

By

$$
|x(t_k) - x^n(t_k)|^2 = \sum_{l=1}^{k} (|x(t_l) - x^n(t_l)|^2 - |x(t_{l-1}) - x^n(t_{l-1}))|^2),
$$

summing up, when $l = 1, 2, \cdots, k$, we obtain

$$
\begin{aligned}
& |x(t_k) - x^n(t_k)|^2 \\
\le \;& \int_0^{t_k} 2(x(\kappa_2(s)) - x^n(\kappa_2(s))) \Big(a(x(s)) - a(x^n(\kappa_2(s))\Big) ds \\
& + \sum_{j=1}^{d_1} \Big| \int_0^{t_k} \Big(b^j(x(s)) - b^j(x^n(\kappa_1(s)))\Big) dw_s^j \Big|^2 \\
& + \sum_{j=1}^{d_1} \int_0^{t_k} 2(x(\kappa_1(s)) - x^n(\kappa_1(s))) \Big(b^j(x(s)) - b^j(x^n(\kappa_1(s)))\Big) dw_s^j.
\end{aligned}
$$

Raising both sides to the power $q/2$ and taking expectations, for $k \le i \le n$, we have

$$
E \max_{0 \le k \le i} |x(t_k) - x^n(t_k)|^q \le C_q(I_1 + I_2 + I_3), \tag{3.4.16}
$$

with

$$
I_1 := E \max_{0 \le k \le i} \Big(\int_0^{t_k} 2(x(\kappa_2(s)) - x^n(\kappa_2(s))) \big(a(x(s)) - a(x^n(\kappa_2(s))\big) ds \Big)^{q/2};
$$

$$I_2 := E \max_{0\le k\le i} \sum_{j=1}^{d_1} \Big(\int_0^{t_k} \Big(b^j(x(s)) - b^j(x^n(\kappa_1(s)))\Big) dw_s^j \Big)^q;$$

$$I_3 := E \max_{0\le k\le i} \sum_{j=1}^{d_1} \Big(\int_0^{t_k} 2(x(\kappa_1(s)) - x^n(\kappa_1(s)))(b^j(x(s)) - b^j(x^n(\kappa_1(s))))dw_s^j \Big)^{q/2}.$$

Step 2. Estimation of I_1:

$$I_1 \le t_i^{q/2-1} 2^{q/2} E \int_0^{t_i} \Big((x(\kappa_2(s)) - x^n(\kappa_2(s)))(a(x(s)) - a(x^n(\kappa_2(s)))) \Big)^{q/2} ds$$

Let C be a constant, $C := C(T, q, r, L_1, L_2, C_4)$ that may change line by line, by one-sided Lipschitz condition (i) of Assumption 3.3.11,

$$I_1 \le CE \int_0^{t_i} \Big(|x(s) - x^n(\kappa_2(s))|^2 + I_1' \Big)^{q/2} ds,$$

with $I_1' := |(x(\kappa_2(s) - x(s))(a(x(s)) - a(x^n(\kappa_2(s))|$,

$$\begin{aligned} I_1 &\le CE \int_0^{t_i} \Big(|x(\kappa_2(s)) - x^n(\kappa_2(s))|^2 + |x(s) - x(\kappa_2(s))|^2 + I_1' \Big)^{q/2} ds \\ &\le C \sum_{l=0}^{i-1} E \max_{k\le l} |x(t_k) - x^n(t_k)|^q \Delta t \\ &\quad + CE \int_0^{t_i} |x(s) - x(\kappa_2(s))|^q ds + CE \int_0^{t_i} I_1'^{q/2} ds. \end{aligned}$$

From the known estimate from Lemma 3.3.3, it is easy to get

$$CE \int_0^{t_i} |x(s) - x(\kappa_2(s))|^q ds \le Cn^{-q/2}.$$

Then polynomial condition on a shows that,

$$\begin{aligned} & CE \int_0^{t_i} I_1'^{q/2} ds \\ \le\ & CE \int_0^{t_i} |x(\kappa_2(s) - x(s)|^{q/2}(1 + |x(s)|^{qr/2} + |x^n(\kappa_2(s))|^{qr/2}) ds \\ \le\ & C \int_0^{t_i} E(|x(\kappa_2(s) - x(s)|^q)^{\frac{1}{2}} \{E(1 + |x(s)|^{qr} + |x^n(s)|^{qr})\}^{\frac{1}{2}} ds \\ \le\ & Cn^{-q/4}. \end{aligned}$$

Hence

$$I_1 \le C \sum_{l=0}^{i-1} E \max_{k\le l} |x(t_k) - x^n(t_k)|^q \Delta t + Cn^{-q/4} + Cn^{-q/2}. \qquad (3.4.17)$$

Step 3. Estimation of I_2:

By Burkholder-Davis-Gundy's inequality,

$$\begin{aligned}
I_2 &\leq C\sum_{j=1}^{d_1} E\Big(\int_0^{t_i} \big|b^j(x(s)) - b^j(x^n(\kappa_1(s))\big|^2 ds\Big)^{q/2} \\
&\leq CE\Big(\int_0^{t_i} \big|x(s) - x^n(\kappa_1(s))\big|^2 ds\Big)^{q/2} \\
&\leq CE\Big(\int_0^{t_i} |x(\kappa_1(s)) - x^n(\kappa_1(s))|^2 ds\Big)^{q/2} + CE\Big(\int_0^{t_i} |x(\kappa_1(s)) - x(s)|^2 ds\Big)^{q/2} \\
&\leq C\sum_{l=0}^{i-1} E\max_{k\leq l} |x(t_k) - x^n(t_k)|^q \Delta t + CE\Big(\int_0^{t_i} |x(\kappa_1(s)) - x(s)|^2 ds\Big)^{q/2},
\end{aligned}$$

note that

$$\begin{aligned}
& E\Big(\int_0^{t_i} |x(\kappa_1(s)) - x(s)|^2 ds\Big)^{q/2} \\
\leq\ & E\{\sup_{t\leq t_i} |x(\kappa_1(s)) - x(s)|^{q/2} (\int_0^{t_i} |x(\kappa_1(s)) - x(s)| ds)^{q/2}\} \\
\leq\ & (E\sup_{t\leq t_i} |x(\kappa_1(s)) - x(s)|^q)^{1/2} \{E(\int_0^{t_i} |x(\kappa_1(s)) - x(s)| ds)^q\}^{1/2} \\
\leq\ & C(E\sup_{t\leq t_i} |x(\kappa_1(s)) - x(s)|^q)^{1/2} \{E(\int_0^{t_i} |x(\kappa_1(s)) - x(s)|^q ds)\}^{1/2} \\
\leq\ & Cn^{-q/2}.
\end{aligned}$$

Then

$$I_2 \leq C\sum_{l=0}^{i-1} E\max_{k\leq l} |x(t_k) - x^n(t_k)|^q \Delta t + Cn^{-q/2}. \tag{3.4.18}$$

Step 4. Estimation of I_3:

Finally,

$$\begin{aligned}
I_3 &\leq CE\Big(\int_0^{t_i} \big|x(\kappa_1(s)) - x^n(\kappa_1(s))\big|^2 \big|b^j(x(s)) - b^j(x^n(\kappa_1(s)))\big|^2 ds\Big)^{q/4} \\
&\leq CE\{\sup_{t\leq t_i} \big|x(\kappa_1(t)) - x^n(\kappa_1(t))\big|^{q/2} \Big(\int_0^{t_i} \big|b^j(x(s)) - b^j(x^n(\kappa_1(s)))\big|^2 ds\Big)^{q/4}\} \\
&\leq \frac{1}{2} E\max_{k\leq i} |x(t_k) - x^n(t_k)|^q + CE\Big(\int_0^{t_i} \big|x(s) - x^n(\kappa_1(s))\big|^2 ds\Big)^{q/2}.
\end{aligned}$$

$$\leq \frac{1}{2} E\max_{k\leq i} |x(t_k) - x^n(t_k)|^q + C\sum_{l=0}^{i-1} E\max_{k\leq l} |x(t_k) - x^n(t_k)|^q \Delta t + Cn^{-q/2} \tag{3.4.19}$$

Finally combining 3.4.17, 3.4.18, and 3.4.19, putting into 3.4.16, we obtain

$$E\max_{0\leq k\leq i} |x(t_k) - x^n(t_k)|^q$$

$$\leq C\sum_{l=0}^{i-1} E\max_{k\leq l} |x(t_k) - x^n(t_k)|^q \Delta t + Cn^{-q/2} + Cn^{-q/4}$$

Consequently, by discrete Gronwall's lemma (see Theorem 2.3.8), we obtain

$$\begin{aligned} & E \max_{0\le k\le i} |x(t_k) - x^n(t_k)|^q \\ \le\ & (Cn^{-q/2} + Cn^{-q/4})(1 + C\Delta t)^i \\ \le\ & Ce^{CT} n^{-q/4}, \end{aligned}$$

where C depends on the coefficients in the assumptions, and $T, q, r, E|x_0|^{qr}$, but independent of n. □

We begin to prove Theorem 3.2.7:

Proof of Theorem 3.2.7. Remember in (3.2.8) that $\{x^n(t) : t \in [0,T]\}$ satisfies

$$\begin{aligned} x^n(t) &= x^n(t_{k-1}) + a\Big(\rho_{\Delta t}(x^n(t_{k-1}) + \sum_{j=1}^{d_1} b^j(x^n(t_{k-1}))\Delta w^j_{t_{k-1}})\Big)\Delta t \\ &\quad + \sum_{j=1}^{d_1} b^j(x^n(t_{k-1}))\Delta w^j_{t_{k-1}}, \end{aligned}$$

for $t = t_k$, $k = 1, 2, \cdots, n$. Hence

$$x^n(t) = x^n(t_{k-1}) + a(\theta_n(x^n(t_{k-1})), \Delta w^j_{t_{k-1}}))\Delta t + \sum_{j=1}^{d_1} b^j(x^n(t_{k-1}))\Delta w^j_{t_{k-1}}, \quad (3.4.20)$$

for $t = t_k$, where $\theta_n : \mathbb{R}^d \times \mathbb{R}^{d_1} \to \mathbb{R}^d$ is defined by $\theta_n(x,y) = \rho_{\Delta t}(x + b(x)y)$, for $x \in \mathbb{R}^d$, $y \in \mathbb{R}^{d_1}$.

Notice that

$$\begin{aligned} |\theta_n(x,y)| &\le (1 - C_1\Delta t)^{-1}(|x| + |b(x)||y|) \quad (3.4.21) \\ &\le (1 - C_1\Delta t)^{-1}(|x| + C_3(1+|x|)|y|), \end{aligned}$$

by Corollary 3.2.3, where C_3 is the constant from the linear growth condition 3.2.5 on b. Fix $\epsilon > 0$ and let $n_0 \ge 1$ be an integer such that $1 - C_1T/n_0 > \epsilon$. Then by (3.4.21) there exists an increasing sequence $\{(C(R))\}_{R=1}^{\infty}$, such that $C(R) \uparrow \infty$ and

$$|\theta_n(x,y)| \le C(R),$$

for all $n \ge n_0$, $|x|, |y| \le R$. Such a sequence can be defined for example by

$$C(R) := R + \sup_{n\ge n_0} \sup_{|x|\le R} \sup_{|y|\le R} |\theta_n(x,y)|.$$

Then from the truncation result of functions (a, b), (see, e.g., [13]), for each R there exist bounded Borel functions $a_R : \mathbb{R}^d \to \mathbb{R}^d$, $b_R : \mathbb{R}^d \to \mathbb{R}^{d\times d_1}$ such that

$$a_R(x) = a(x),\ b_R(x) = b(x),$$

for all $x \in \mathbb{R}^d$, $|x| \le C(R)$, and for some constant $L = L_R$,

$$(x-y)(a_R(x) - a_R(y)) \le L|x-y|^2;$$
$$|b_R(x) - b_R(y)|^2 \le L|x-y|^2,$$

for all $x, y \in \mathbb{R}^d$.

Consider the problem

$$dx_R(t) = a_R(x_R(t))dt + \sum_{j=1}^{d_1} b_R^j(x_R(t))dw_t^j \tag{3.4.22}$$

with the initial condition $x_R(0) = x_0 \mathbf{1}_{|x_0| \le R - \frac{1}{2}}$ and let $\{x_R^n(t), t \in [0,T]\}$ denote its implicit approximation defined on the interval $[0,T]$. They applying Theorem 3.2.6 to equation (3.4.22), for every $q \ge 1$ we have a constant $C = C_{qR}$ such that

$$E \max_{0 \le j \le n} |x_R(t_i) - x_R^n(t_i)|^q \le C n^{-\frac{q}{4}},$$

then by Lemma 3.3.3,

$$E \sup_{0 \le t \le T} |x_R(t) - x_R^n(t)|^q \le C n^{-\frac{q}{4}} \tag{3.4.23}$$

holds for all sufficiently large integers $n \ge 1$, where x_R^n is the implicit approximation, and x_R is the solution of equation (3.4.22). Hence by virtue of the Borel-Cantelli lemma (see Lemma 2.1.1) for every $0 < \alpha < \frac{1}{4}$, there exists a finite random variable η_R such that

$$\sup_{t \le T} |x_R^n(t) - x_R(t)| \le \eta_R n^{-\alpha} (a.s.) \tag{3.4.24}$$

for all $n \ge 1$.

Define the stopping times

$$\tau^n = \inf\{t \in [0,T] : |x(t) - x^n(t)| \ge \frac{1}{2}\},$$
$$\tau_R = \inf\{t \in [0,T] : 2|w_t^j| + |x(t)| \ge R - \frac{1}{2}\},$$
$$\sigma_R^n = \tau^n \wedge \tau_R,$$

for all $R \ge 1$, where $x(t)$ is the solution of

$$\begin{cases} dx(t) = a(x(t))dt + \sum_{j=1}^{d_1} b^j(x(t))dw_t^j, \\ x(0) = x_0 \end{cases}$$

and $x^n(t)$ is its implicit approximation on $[0,T]$. Then $x(t)$ satisfies equation (3.4.22) on $[0, \sigma_R^n \wedge T]$, $\omega \in [\tau_R > 0] = [x_0 < R - \frac{1}{2}] \in \mathcal{F}_0$. Let us denote $A_R := [x_0 < R - \frac{1}{2}]$. Hence

$$x(t) = x_R(t), \text{ for } t \in [0, \sigma_R^n \wedge T),$$

for almost all $\omega \in A_R$.

By virtue of (3.4.20) and the definition of a_R, b_R

$$x^n(t) = x^n_R(t), \text{ on } t \in [0, \sigma^n_R \wedge T)$$

for almost all $\omega \in A_R$. Hence by (3.4.24), we have a random variable η_R, such that for almost every $\omega \in A_R$,

$$\sup_{t\in T} |x^n(t \wedge \sigma^n_R) - x(t \wedge \sigma^n_R)| \le \eta_R n^{-\alpha}, \tag{3.4.25}$$

for all $n \ge n_0$. Since $\tau_R \uparrow \infty$ as $R \to \infty$,

$$\Omega = \cup_{R=1}^{\infty} \Omega_R, \tag{3.4.26}$$

where $\Omega_R := \{\omega : \tau_R \ge T\}$.

Notice also that

$$\Omega_R \subset A_R. \tag{3.4.27}$$

Therefore (3.4.25)-(3.4.27) imply that almost surely

$$\sup_{t\in[0,T]} |x^n(t \wedge \tau^n) - x(t \wedge \tau^n)| \le \xi n^{-\alpha},$$

for all $n \ge n_0$, where ξ is a finite random variable defined by

$$\xi(\omega) := \xi_1(\omega), \text{ for } \omega \in \Omega_1;$$

and

$$\xi(\omega) := \eta_R(\omega), \text{ for } \omega \in \Omega_R \backslash \Omega_{R-1}, R \ge 2.$$

Hence by using Lemma A.0.2 in Appendix, we get a finite random variable η such that almost surely

$$\sup_{t\in[0,T]} |x^n(t) - x(t)| \le \eta n^{-\alpha}$$

for all $n \ge n_0$. The proof of the theorem is complete.

□

Theorem 3.4.1. *Let the assumptions of the previous theorem be satisfied. Assume moreover that* $E|x_0|^q < \infty$, *for some* $q > 1$. *Then for every* $0 < p < q$,

$$E \sup_{t\le T} |x^n(t) - x(t)|^p \to 0,$$

as $n \to \infty$.

Proof. From the fact above that $x^n(t)$ converges a.s. to $x(t)$, when $n \to \infty$. Also from the boundedness of $E\sup_{t\le T}|x^n(t)|^q$ and $E\sup_{t\le T}|x(t)|^q$, we know that both $\sup_{t\le T}|x^n(t)|^p$ and $\sup_{t\le T}|x(t)|^p$ are uniformly integrable. So $\sup_{t\le T}|x^n(t)-x(t)|^p$ is also uniformly integrable. Then one can interchange the limit with the expectation, which proves the theorem. □

Chapter 4

Solutions of Stochastic Differential Inclusions

In the next few chapters stochastic differential inclusions (SDIs) on domain of $\mathbb{R}^d$ have been investigated as follows:

$$dx(t) \in a(t, x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j, \ x(0) = x_0,$$

where a is a maximal monotone mapping, b is a Lipschitz continuous function and w is a Wiener process.

In this chapter we are concerned with the existence and uniqueness of solutions for SDIs. The existence of solutions is proved by monotonicity method in terms of minimizing certain convex functionals and in this way solutions are approximated by implicit schemes. Moreover a result on the rate of convergence is presented.

4.1 Introduction

There are a variety of motivations that lead us to study such class of SDIs. If we deal with an SDE, in many practical problems, the existence of a solution is not always guaranteed. There are many cases when there is no solution to the SDEs, even though the Euler approximations converge. For example, let us consider the following Itô's stochastic differential equation on $\mathbb{R}^d$:

$$\begin{cases} dx(t) = a(x(t))dt + x(t)dw_t \\ x(0) = 0 \end{cases} \tag{4.1.1}$$

with $a(x) = 1$ for $x \leq 0$ and $a(x) = -1$ for $x > 0$.

In order to prove the existence for such problems, assume that we consider an explicit Euler approximation x^n for the SDE (4.1.1). Then it is easy to show that x^n converges almost surely to some stochastic process x. On the other hand it does not have any solution of SDE (4.1.1). This may be inconvenient in certain applications. Nonetheless we shall see that this equation falls into a general class. If we extend however, a to be multi-valued and the SDE to be a stochastic inclusion, it can be proved that there exists a unique solution of the SDI obtained. Therefore the equation is no longer a usual SDE type, but the one with a multi-valued drift term. Generally, the right hand side of an SDI is a set of values rather than a single value. So far we can see that SDIs play a crucial role in the theory of SDEs with a discontinues right-hand side.

In the last ten years, it is noticed that existence and approximation of solutions to SDIs have received broad attention. Existence and numerical approximation methods have been tackled in several papers. P. Krée introduced the notion of multi-valued stochastic differential equations in his paper [25] (we call it as stochastic differential inclusions). In his type, he showed SDI in the form as:

$$dx(t) \in a(t, x(t))dt - A(x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dB_t^j, \tag{4.1.2}$$

where A is a so-called multi-valued 'maximal monotone' map, $(B_t)_{t\geq 0}$ is the standard Brownian motion on $\mathbb{R}^d$. Krée's work presented the existence of a solution to (4.1.2) in a product situation. This indicates that a composition of $\mathbb{R}^d$, $\mathbb{R}^d = \mathbb{R}^p \times \mathbb{R}^{d-p}$ can be made so that for $x \in D(A)$, the pth first components of Ax is 0 and $b_{ij} = 0$ for $i = p+1, \cdots, d$ and all j. Such type of SDIs has also been developed by other authors. Pettersson in [31] defined approximate solutions when the maximal monotone map A is replaced by Yosida approximation A_λ. He proved convergence and acquired the existence of a solution to the multi-valued stochastic differential equations under suitable conditions. It is well known that such technique of Yosida approximation was employed before on differential inclusions problems. Aubin defined approximation solutions by replacing the maximal multi-valued monotone map by the Yosida approximation, and proved the existence and uniqueness of solutions to differential inclusions (as we can see in the book [2]). In [38] Y.S.Yong consider the SDIs associated with following form:

$$dx(t) \in a(t, x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dB_t^j,$$

where B is a Brownian motion. He proved the existence of solution for SDIs under the condition that the drift and diffusion terms satisfy the local Lipschitz property and linear growth condition. Our aim is to prove the existences and uniqueness of solutions to the SDIs on $\mathbb{R}^d$ as follows:

$$\begin{cases} dx(t) \in a(t, x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j, \\ x(0) = x_0. \end{cases} \tag{4.1.3}$$

We will use the method of monotonicity, which is interpreted as a method of minimizing certain convex functionals. The idea of our results is inspired by N.V. Krylov [19], where SDEs in $\mathbb{R}^d$ are solved by minimizing convex functionals via Euler approximations. It provides the basis for a promising adaption. Such technique is a straightforward method to obtain the existence results, which are also developed by Gyöngy, I. and Millet, A. in [12]. They made use of such corresponding construction for stochastic partial differential equations to obtain the solution via Euler-Galerkin approximation.

Before we move on, we recall the minimization method for convex functionals. A brief explanation on SDEs is given as follows. The main source of information for this part is [19].

4.1.1 Minimization method

Krylov characterized the solutions to stochastic differential equations

$$dx(t) = a(t, x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j, \ x(0) = x_0 \tag{4.1.4}$$

as the minimizers of a suitable convex functional G. Let $\mathcal{H}$ be the space of pairs (α, β) of $\mathcal{F}_t$−adapted stochastic processes, i.e., $\alpha = \{\alpha_t : t \in [0, T]\}$ is $\mathbb{R}^d$−valued, and $\beta = \{\beta_t : t \in [0, T]\}$ is $\mathbb{R}^{d\times d_1}$−valued with $E\int_0^T |\alpha_s|^2 ds < \infty$, $E\int_0^T \sum_{j=1}^{d_1} |\beta_s^j|^2 ds < \infty$, and such that

$$x_t(\alpha, \beta) = x_0 + \int_0^t \alpha_s ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^j dw_s^j, \tag{4.1.5}$$

is an $\mathcal{F}_t$−adapted stochastic process, where x_0 is an $\mathcal{F}_0$−measurable random variable in $\mathcal{H}$. Let us define $x(t) := x_t(\alpha, \beta)$ as in (4.1.5) and set functional

$$G(\alpha, \beta) = \sup_{y\in Y} E\int_0^T 2(x(t)-y(t))(\alpha_t - a(t, y(t))) + \sum_{j=1}^{d_1} |\beta_t^j - b^j(t, y(t))|^2 dt, \tag{4.1.6}$$

for

$$Y = \{y(t) : E\int_0^T |y(t)|^2 dt < \infty\}.$$

We say that

$$x_t(\bar{\alpha}, \bar{\beta}) = x_0 + \int_0^t \bar{\alpha}_s ds + \sum_{j=1}^{d_1} \int_0^t \bar{\beta}_s^j dw_s^j$$

is a *generalized solution* of equation (4.1.4) in the sense of extremals, if G attains its minimum at $(\bar{\alpha}, \bar{\beta})$.

Our aim in this chapter is to formulate the method of monotonicity in terms of finding extremals of convex functionals. We construct the solutions to SDIs via implicit approximations. Implicit approximate schemes are presented for SDEs in the previous chapter. We will show that implicit approximations defined by $x^n(t)$ converge almost surely to a stochastic process $x(t)$, given as the unique solution of inclusion (4.1.3). Finally, we establish the existence and uniqueness of a solution to SDIs. The convergence of the approximations to the solutions of SDIs is proved, moreover the rate of convergence is also presented.

This chapter is organized as follows:

- section 4.2: this section defines the SDIs, gives the definition of solution to stochastic differential inclusions and states the main results of this chapter;
- section 4.3: this section links the minimization methods to SDIs, defines the implicit approximations, and discusses the existence and uniqueness of solutions to the implicit schemes;
- section 4.4: this section proves the existence of a solution to SDIs via implicit method, and shows the rate of convergence.

4.2 Stochastic Differential Inclusions

We first introduce some notions used in this chapter. Let T be a fixed positive constant. Let $a : [0,\infty) \times \mathbb{R}^d \to 2^{\mathbb{R}^d}$ be a multi-valued function. Let $b : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}^{d\times d_1}$ be a Borel function. Consider the following SDI in $\mathbb{R}^d$ on the time interval $[0,T]$:

$$\begin{cases} dx(t) \in a(t,x(t))dt + \sum_{j=1}^{d_1} b^j(t,x(t))dw_t^j \\ x(0) = x_0. \end{cases} \tag{4.2.7}$$

Definition 4.2.1 (Definition of a solution to stochastic differential inclusions). We say that $x(t) := x_t(\alpha,\beta) = \{x(t) : t \in [0,T]\}$ is a solution to stochastic differential inclusion (4.2.7), if there exist $\mathcal{F}_t-$adapted $\mathbb{R}^d-$valued stochastic process $\alpha = \{\alpha_t : t \in [0,T]\}$, and $\mathbb{R}^{d\times d_1}-$valued stochastic process $\beta = \{\beta_t : t \in [0,T]\}$ such that

$$x(t) = x_0 + \int_0^t \alpha_s ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^j dw_s^j$$

holds almost surely for all $t \in [0,T]$, and $\alpha_t(\omega) \in a(t,x(t))$ for $dt \times dP-$almost every $(t,\omega) \in [0,T] \times \Omega$, where $\beta_t = b(t,x(t))$ for almost every $(t,\omega) \in [0,T] \times \Omega$.

The following assumptions are needed.

Assumption 4.2.1. Let x_0 be an $\mathcal{F}_0-$measurable random variable in $\mathbb{R}^d$, such that $E|x_0| < \infty$.

Assumption 4.2.2. Let a be a maximal $K-$monotone (multi-valued) function in $x \in \mathbb{R}^d$ for $t \in [0, T]$, and b be a Lipschitz continuous function.

It can be seen that maximal monotonicity of a together with the Lipschitz continuity of b imply the following monotonicity condition of pair (a, b).

Assumption 4.2.3 (Monotonicity of (a, b)). There exists a constant K, such that

$$2(x-y)(a(t,x)-a(t,y)) + \sum_{j=1}^{d_1} |b^j(t,x) - b^j(t,y)|^2 \leq K|x-y|^2,$$

for all $x, y \in \mathbb{R}^d$, and $t \in [0, T]$.

Assumption 4.2.4 (Linear growth of (a, b)). There exists a constant L_1, such that for $dP \times dt$ almost every $(\omega, t) \in \Omega \times [0, T]$,

$$|a(t,x)|^2 + \sum_{j=1}^{d_1} |b^j(t,x)|^2 \leq L_1(1+|x|^2),$$

for all $x \in \mathbb{R}^d$.

Sometimes the following weaker condition is considered:

Assumption 4.2.5 (Growth condition of (a, b)). There exists a constant L_2, such that almost surely

$$2xa(t,x) + \sum_{j=1}^{d_1} |b^j(t,x)|^2 \leq L_2(1+|x|^2)$$

for all $x \in \mathbb{R}^d$, and $t \in [0, T]$,

Remark 4.2.1. It can be seen that monotonicity condition 4.2.3 implies that diffusion term b must be single-valued and continuous in $x \in \mathbb{R}^d$.

Remark 4.2.2. Without loss of generality, we may assume $K = 0$ in Assumptions 4.2.2 and 4.2.3. Notice that, if x is a solution to (4.2.7), and $\bar{x}(t) = e^{-Kt/2}x(t)$, then by Itô's formula, x exists if and only if the process $\bar{x}$ solves (4.2.7) with $(\bar{a}, \bar{b})$ instead of (a, b), where

$$\bar{a}(t,x) = e^{-Kt/2}a(t, e^{Kt/2}x) - \frac{K}{2}x,$$
$$\bar{b}(t,x) = e^{-Kt/2}b(t, e^{Kt/2}x),$$

K is the constant in the monotonicity condition. Therefore we will assume $K = 0$. The following monotonicity assumption will be used throughout this chapter.

Assumption 4.2.6 (Monotonicity of (a, b)). For all $x, y \in \mathbb{R}^d$, for $t \in [0, T]$ such that

$$2(x-y)(a(t,x)-a(t,y))+\sum_{j=1}^{d_1}|b^j(t,x)-b^j(t,y)|^2 \leq 0.$$

For simplicity of presentation, we assume that a and b are time-independent. In this chapter, we aim to prove the existence of a solution to the SDI (4.2.7) on the interval $[0, T]$ under certain assumptions . The following theorem is the main theorem of this chapter.

Theorem 4.2.1. *Under Assumptions 4.2.1, 4.2.2, 4.2.4 and 4.2.6, stochastic differential inclusion (4.2.7) has one and only one solution.*

Proof. Now we are going to prove the uniqueness of the solution. Suppose there exist $x^1(t)$ and $x^2(t)$ two possible solutions of inclusions (4.2.7), i.e.

$$x^1(t) = x_0 + \int_0^t \alpha_s^1 ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^{1,j} dw_s^j,$$

$$x^2(t) = x_0 + \int_0^t \alpha_s^2 ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^{2,j} dw_s^j,$$

and $\alpha_t^1 \in a(x^1(t))$, $\alpha_t^2 \in a(x^2(t))$, where $\beta_t^{1,j} = b^j(x^1(t))$, $\beta_t^{2,j} = b^j(x^2(t))$. Therefore, from the monotonicity assumption, by Itô's formula, we obtain

$$d(e^{-kt}|x^1(t)-x^2(t)|^2)$$

$$= e^{-kt}\Big(2(x^1(t)-x^2(t))(\alpha_t^1-\alpha_t^2)+\sum_{j=1}^{d_1}|b^j(x^1(t))-b^j(x^2(t))|^2\Big)\,dt$$

$$-e^{-kt}k|x^1(t)-x^2(t)|^2 dt$$

$$+\sum_{j=1}^{d_1} e^{-kt}2(x^1(t)-x^2(t))(b^j(x^1(t))-b^j(x^2(t)))\,dw_t^j$$

Thus we get,

$$e^{-kt}|x^1(t)-x^2(t)|^2$$

$$= \int_0^t e^{-ks}\Big(2(x^1(s)-x^2(s))(\alpha_s^1-\alpha_s^2)+\sum_{j=1}^{d_1}|b^j(x^1(s))-b^j(x^2(s))|^2$$

$$-k|x^1(s)-x^2(s)|^2\Big)ds + m_t' \text{ (a.s.)},$$

where

$$m_t' = \int_0^t e^{-ks}2(x^1(s)-x^2(s))(b^j(x^1(s))-b^j(x^2(s)))dw_s$$

is a non-negative local martingale, starting from 0. Hence,

$$0 \le e^{-kt}|x^1(t) - x^2(t)|^2 \le m'_t,$$

which implies $m'_t = 0$. Thus almost surely $x^1(t) = x^2(t)$ and the uniqueness of the solution is proved.

The existence of the solution will be followed after some preliminary lemmas: □

4.3 Preliminary Results

4.3.1 Solutions as Extremals

We make the following natural construction.

Let us consider $\mathcal{H}$ be the space of pairs (α, β) of $\mathcal{F}_t-$adapted stochastic processes such that $\alpha = \{\alpha_t : t \in [0,T]\}$ is $\mathbb{R}^d-$valued, $\beta = \{\beta_t : t \in [0,T]\}$ is $\mathbb{R}^{d\times d_1}-$valued with $E\int_0^T |\alpha_s|^2 ds < \infty, E\int_0^T \sum_{j=1}^{d_1} |\beta_s^j|^2 ds < \infty$. For $(\alpha,\beta) \in \mathcal{H}$, we define the process

$$x_t(\alpha,\beta) = x_0 + \int_0^t \alpha_s ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^j dw_s^j, \ t \in [0,T]. \tag{4.3.8}$$

Let us consider also the space of processes

$$Y = \{y(t) : \mathcal{F}_t - \text{adapted process}, \ |y|_Y := (E\int_0^T |y(t)|^2 dt)^{1/2} < \infty\}.$$

Let us define $x(t) := x_t(\alpha,\beta)$ as in (4.3.8), and the functional

$$G(\alpha,\beta) = \sup_{y\in Y} F_y(\alpha,\beta), \tag{4.3.9}$$

where

$$F_y(\alpha,\beta) = E\int_0^T 2(x(t) - y(t))(\alpha_t - a(y(t))) + \sum_{j=1}^{d_1} |\beta_t^j - b^j(y(t))|^2 dt.$$

This section contains preliminary results that are needed in constructive proof of the existence of solutions. First, we explore the properties of maximal monotone operators.

Lemma 4.3.1. *Let $a = a(x)$ be maximal monotone on $\mathbb{R}^d$. Define the multi-valued operator A on Y as follows: For $y \in Y, z \in Y$ belongs to $A(y)$, if $z_t(\omega) \in a(y_t(\omega))$ for $dt \times P-$almost every $(t,\omega) \in [0,T]$, then A is a maximal monotone operator on Y, with respect to the inner product on Y defined by $(u,v) := E\int_0^T u(t)v(t)dt$, for $u, v \in Y$.*

Proof. Obviously, A is monotone. If $z_{t_i} \in A(y_{t_i})$, $i = 1, 2$, then

$$E\int_0^T (z_{t_1} - z_{t_2})(y_{t_1} - y_{t_2})dt \leq 0.$$

In order to prove that A is maximal, we only need to prove that for each $\lambda \geq 0$, and $z \in Y$, there is a solution $y \in Y$ of the equation

$$Ay - \lambda y = z. \tag{4.3.10}$$

Since a is maximal monotone, so $ay - \lambda y = z_t(\omega)$ has a unique solution $y = y_t(\omega)$, for every $\omega \in \Omega$. Because of the uniqueness, y_t is an $\mathcal{F}_t-$measurable in (t, ω). In this way we get an $\mathcal{F}_t-$adapted stochastic process $\{y_t(\omega) : t \in [0, T]\}$, such that $ay_t(\omega) - \lambda y_t(\omega) = z_t(\omega)$ holds for all $t \in [0, T]$ and $\omega \in \Omega$, which means equation (4.3.10) holds. It remains to show that $\{y_t : t \in [0, T]\} \in Y$.

Since $(1 - \lambda a)^{-1}$ is lipschiz continuous, we could get

$$|y(t)| - |(1 - \lambda a)^{-1}(0)| \leq |y(t) - (1 - \lambda a)^{-1}(0)| \leq |z(t)|.$$

Hence $|y(t)| \leq |z(t)| + |(1 - \lambda a)^{-1}(0)|$, which implies
$E\int_0^T |y(t)|^2 dt \leq C + E\int_0^T |z(t)|^2 dt < \infty$. It means $y \in Y$. □

The method to prove the existence consist in characterizing the solutions of SDI (4.2.7) as the extremal values of a suitable convex functional. Below we interpret this method.

Lemma 4.3.2. *G is a convex function, and $G \geq 0$.*

Proof. Notice that, the functional G defined in (4.3.9) can be transformed as

$$G(\alpha, \beta) = E\int_0^T (2x(t)\alpha_t + \sum_{j=1}^{d_1} |\beta_t^j|^2)dt + \sup_{y \in Y} f(\alpha, \beta),$$

where

$$f(\alpha, \beta) = E\int_0^T \Big(-2y(t)\alpha_t - 2x(t)a(y(t))$$

$$+2y(t)a(y(t)) + \sum_{j=1}^{d_1} |b^j(y(t))|^2 - \sum_{j=1}^{d_1} 2\beta_t^j b^j(y(t))\Big)dt. \tag{4.3.11}$$

In this case, from Ito's formula,

$$G(\alpha, \beta) = E|x(T)|^2 - E|x_0|^2 + \sup_{y \in Y} f(\alpha, \beta).$$

Then it remains to note that $x(\alpha, \beta)$ is a linear function of (x_0, α, β), while $|x|^2$ is a convex function. Clearly from the expression above, $f(\alpha, \beta)$ is a convex function. Hence $G(\alpha, \beta)$ is convex. Since one can take $y(t) := \{x(t) : t \in [0, T]\}$ in calculating the supremum, we get $G(\alpha, \beta) \geq 0$.

□

In the following part we summarize the solution of SDI (4.2.7) in terms of extremals of the functional G.

Proposition 4.3.3. *The following statements hold:*

1. *Let Assumptions 4.2.1, 4.2.2, 4.2.4, and 4.2.6 hold, if the stochastic process $x = x(\bar{\alpha}, \bar{\beta})$ is a solution to SDI (4.2.7) for some $(\bar{\alpha}, \bar{\beta}) \in \mathcal{H}$, then $G(\bar{\alpha}, \bar{\beta}) = 0$.*

2. *Let Assumptions 4.2.1, 4.2.2, 4.2.5, and 4.2.6 hold, if for some $(\bar{\alpha}, \bar{\beta}) \in \mathcal{H}$, $G(\bar{\alpha}, \bar{\beta}) \le 0$, then $x(t) = x_t(\bar{\alpha}, \bar{\beta})$ is a solution of SDI (4.2.7).*

Proof. (1) Let $x = x(\bar{\alpha}, \bar{\beta})$ be a solution to SDI (4.2.7), i.e., by the definition of a solution, we say that there exist $\bar{\alpha}, \bar{\beta}$, which are $\mathcal{F}_t-$adapted stochastic processes, such that $\bar{\alpha}_t \in a(x(t))$, where $\bar{\beta}_t^j = b^j(x(t))$, and almost surely

$$x(t) = x_0 + \int_0^t \bar{\alpha}_s ds + \int_0^t \sum_{j=1}^{d_1} \bar{\beta}_s^j dw_s^j$$

holds for all $0 \le t \le T$. Then by monotonicity condition 4.2.6,

$$\int_0^T 2(x(t) - y(t))(\bar{\alpha}_t - a(y(t))) + \sum_{j=1}^{d_1} |\bar{\beta}_t^j - b^j(y(t))|^2 dt \le 0.$$

Hence,

$$E\int_0^T 2(x(t) - y(t))(\bar{\alpha}_t - a(y(t))) + \sum_{j=1}^{d_1} |\bar{\beta}_t^j - b^j(y(t))|^2 dt \le 0. \tag{4.3.12}$$

Since one can take $y := \{x(t) : t \in [0, T]\}$ in calculating the supremum,

$$\begin{aligned} G(\bar{\alpha}, \bar{\beta}) &\ge \sup_{y \in Y} E\int_0^T 2(x(t) - y(t))(\bar{\alpha}_t - a(y(t)))dt \\ &\ge 0. \end{aligned}$$

Consequently, by (4.3.12)

$$G(\bar{\alpha}, \bar{\beta}) = 0.$$

(2) If $G(\bar{\alpha}, \bar{\beta}) \le 0$ hold for some $(\bar{\alpha}, \bar{\beta}) \in \mathcal{H}$, then we have

$$E\int_0^T 2(x(t) - y(t))(\bar{\alpha}_t - a(y(t)) + \sum_{j=1}^{d_1} |\bar{\beta}_t^j - b^j(y(t))|^2 dt \le 0,$$

for all $y \in Y$. For $y := \{x(t) : t \in [0, T]\}$, this gives $\bar{\beta}_t^j = b^j(x(t))$, for $dt \times dP-$almost every $(t, \omega) \in [0, T] \times \Omega$.

Obviously,

$$E\int_0^T 2(x(t)-y(t))(\bar{\alpha}_t - a(y(t)))dt \leq 0.$$

Hence by Lemma 4.3.1 $\bar{\alpha}_t \in a(x(t))\ dt\times dP$ a.e., and the proposition is proved. □

By the above proposition, in order to get the existence result, we need to find the existence of $(\bar{\alpha},\bar{\beta}) \in \mathcal{H}$, such that $G(\bar{\alpha},\bar{\beta}) \leq 0$. The method is to generate a sequence of solutions to SDI (4.2.7) by implicit schemes, and obtain weak limits from this sequence to make sure that the functional G is non-positive. By the end of this section, the implicit time discretization schemes will be described. We will extend the technique of implicit methods introduced in the previous chapter to SDIs.

4.3.2 Implicit Approximation Schemes

For every integer $n \geq 1$, we approximate the solution $x(t)$ of the SDI (4.2.7) by the process $x^n(t)$ solving the following stochastic inclusions:

$$\begin{cases} x^n(t_k) \in x^n(t_{k-1}) + a(x^n(t_k))\Delta t + \sum_{j=1}^{d_1} b^j(x^n(t_{k-1}))(w^j_{t_k} - w^j_{t_{k-1}}) \\ x^n(t_0) = x_0 \end{cases} \tag{4.3.13}$$

for $t_k := k\Delta t,\ k = 1,2,\cdots,n$. The process $x^n(t)$ are defined on the partition $[t_{k-1},t_k)$ as stepwise constant adapted stochastic process, i.e.

$$x^n(t) := x^n(t_{k-1}),\ \text{ for } 1 \leq k < n. \tag{4.3.14}$$

Here we denote $\Delta w^j_{k-1} := w^j_{t_k} - w^j_{t_{k-1}}$.

4.3.3 Existence and Uniqueness of Solutions to the Implicit Schemes

The following theorem gives the existence and uniqueness of solutions $x^n(t)$ to (4.3.13) for sufficiently large n:

Theorem 4.3.4. *Let Assumptions 4.2.1, 4.2.2 and 4.2.6 hold, then the system of inclusions (4.3.13) has a unique solution $\{x^n(t_k) : k = 1,\cdots,n\}$ if n is sufficiently large. This means there exist sequences of random vectors $\{x^n(t_k) : k = 1,2,\cdots,n\}$, and $\{\alpha_k : k = 1,\cdots n\}$, such that*

$$x^n(t_k) = x^n(t_{k-1}) + \alpha_k\Delta t + \sum_{j=1}^{d_1}\beta^j_{k-1}\Delta w^j_{k-1}, \tag{4.3.15}$$

and $\alpha_k \in a(x^n(t_k))$, where $\beta_{k-1} = b(x^n(t_{k-1}))$, for all $\omega \in \Omega$, $k = 1,2,\cdots,n$.

Proof. First we prove the uniqueness. We fix $n \geq 1$, and use the notation $\xi_k = x^n(t_k)$, for $k = 1, 2, \cdots, n$. Then the system (4.3.13) can be written as

$$y \in -\xi_k + a(\xi_k)\Delta t,$$

where

$$y := -\xi_{k-1} - \sum_{j=1}^{d_1} b^j(\xi_{k-1})\Delta w^j_{k-1}, \ k = 1, 2, \cdots, n.$$

Let ξ_k, ξ'_k be two solutions, such that

$$y = -\xi_k + \alpha_k \Delta t = -\xi'_k + \alpha'_k \Delta t,$$

where

$$\alpha_k \in a(\xi_k), \ \alpha'_k \in a(\xi'_k).$$

We note that $\{\alpha_k : k = 1, 2 \cdots, n\}$ with

$$\alpha_k := \frac{1}{\Delta t}\Big(\xi_k - \xi_{k-1} - \sum_{j=1}^{d_1} b^j(\xi_{k-1})(w^j(t_k) - w^j(t_{k-1})\Big).$$

Obviously,

$$\xi_k - \xi'_k = \Delta t(\alpha_k - \alpha'_k).$$

So,

$$|\xi_k - \xi'_k|^2 = \Delta t(\alpha_k - \alpha'_k)(\xi_k - \xi'_k).$$

By the monotonicity condition, we deduce that

$$|\xi_k - \xi'_k|^2 \leq 0, \text{that is, } \xi_k = \xi'_k,$$

and hence $\alpha_k = \alpha'_k$, which proves that ξ_k are uniquely determined for any given y. Moreover by induction we get the uniqueness of the sequences $\{x^n(t_k) : k = 1, 2, \cdots, n\}$.

We now claim that the system of inclusions has a solution $\{\xi_k : k = 1, \cdots, n\}$. We shall show this by induction on k. For $k = 1$, we have

$$y \in -\xi_1 + a(\xi_1)\Delta t, \ \text{where } y = -x_0 - \sum_{j=1}^{d_1} b^j(x_0)\Delta w^j_0 \text{ is fixed.}$$

Because a is maximal monotone, so $\Delta t a$ is maximal monotone in the usual sense. By Minty Theorem 2.4.3, we know that $-I+\Delta t a$ is surjective, i.e., $\mathrm{Im}(-I+\Delta t a) = \mathbb{R}^d$. Hence for any given y, there exists a solution ξ_1 such that $y = -\xi_1 + a(\xi_1)\Delta t$. Set $\alpha_1 := \frac{y+\xi_1}{\Delta t}$. So ξ_1 satisfy inclusion (4.3.13) for $k = 1$. Assume it holds when $1 \leq k < n$. Then $y = -\xi_k - \sum_{j=1}^{d_1} b(\xi_k)\Delta w^j_k$ is given since solution ξ_k exists, and by Lemma 2.4.4 the inclusion $y \in -\xi_{k+1} + a(\xi_{k+1})\Delta t$ admits a solution ξ_{k+1}. That is, the assumption holds for $k + 1$. Hence by induction on k, the system of inclusions has a solution $\{x^n(t_k) : k = 1, 2, \cdots, n\}$.

□

4.3.4 Estimates of the Implicit Approximations

From now on, we reformulated the stochastic inclusion (4.3.13) in an integral form. Recall that for any fixed integer $n \geq 1$, for $0 < k \leq n$, set $t_k := k\Delta t$, we define

$$\kappa_1 := t_{k-1}, \ \kappa_2 := t_k, \text{ for } t \in [t_{k-1}, t_k).$$

Then the pair of processes (α^n, β^n) on each interval is defined as

$$\alpha_t^n := \alpha_k^n \in a(x^n(\kappa_2(t))), \text{ and } \beta_t^n := \beta_{k-1}^n = b(x^n(\kappa_1(t))).$$

Let us define $x^n(t)$ for $t \in [t_{k-1}, t_k)$ as follows,

$$x^n(t) = x^n(t_{k-1}) + (t - t_{k-1})\alpha_t^n + \sum_{j=1}^{d_1} \beta_t^{n,j}(w_t^j - w_{t_{k-1}}^j).$$

Then it is easy to see that $\{x^n(t) : t \in [0,T]\}$ satisfies the following equation

$$x^n(t) = x^n(t_0) + \int_0^t \alpha_s^n ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^{n,j} dw_s^j, \tag{4.3.16}$$

i.e. $x^n(t) := x_t(\alpha^n, \beta^n)$.

The following lemma provides important estimates for the approximations.

Lemma 4.3.5. *Let Assumptions 4.2.1, 4.2.2, and 4.2.5 hold, then the following estimates hold:*

(i) The solutions of the system of stochastic inclusions (4.3.13) satisfy

$$E \max_{0 \leq k \leq n} |x^n(t_k)|^2 \leq C,$$

where C is a constant independent of n.

(ii)Assume moreover that Assumption 4.2.4 hold, then the pair $\{(\alpha^n, \beta^n)\}$ is in a ball of $\mathcal{H}$, i.e.,

$$E \int_0^T (|\alpha_t^n|^2 + \sum_{j=1}^{d_1} |\beta_t^{n,j}|^2) dt \leq R,$$

for some constant R.

(iii)

$$\sup_n \sup_{t \leq T} E|x^n(t)|^2 < \infty.$$

Proof. (i) Lemma 3.3.2 in Chapter 3 implies this result.

(ii) By using linear growth condition 4.2.4, we get

$$
\begin{aligned}
& E\int_0^T (|\alpha_t^n|^2 + \sum_{j=1}^{d_1} |\beta_t^{n,j}|^2)dt \\
= & \sum_{k=1}^{n} E\int_{t_{k-1}}^{t_k} (|\alpha_t^n|^2 + \sum_{j=1}^{d_1} |\beta_t^{n,j}|^2)dt \\
\leq & 2L_1 \sum_{k=1}^{n} E\int_{t_{k-1}}^{t_k} (1 + |x^n(t_k)|^2 + |x^n(t_{k-1})|^2)dt \\
\leq & 2L_1 T + 4L_1 \Delta t \sum_{k=1}^{n} E|x^n(t_{k-1})|^2.
\end{aligned}
$$

Hence by the first statement,

$$
E\int_0^T (|\alpha_t^n|^2 + \sum_{j=1}^{d_1} |\beta_t^{n,j}|^2)dt \leq 2L_1 T + 4L_1 C
$$

where C is the constant from the first estimate. The first assertion is proved when we let $R := 2L_1 T + 4L_1 C$.

(iii) From the definition of the approximations of x^n, repeating the steps in lemma 3.3.2, we are able to obtain

$$
\begin{aligned}
& |x^n(t_k)|^2 - |x^n(t_{k-1})|^2 \\
\leq\ & 2x^n(t_k)\alpha_k^n \Delta t + 2\sum_{j=1}^{d_1} x^n(t_{k-1}))\beta_{k-1}^{n,j}\Delta w_{t_{k-1}}^j + \sum_{j=1}^{d_1} |\beta_{k-1}^{n,j}\Delta w_{t_{k-1}}^j|^2,
\end{aligned}
$$

$k = 1, 2, \cdots, n$. When adding these inequations and taking expectations, we get

$$
E|x^n(t_l)|^2 \leq E|x^n(t_0)|^2 + 2E\int_0^{t_l} \alpha_{\kappa_2(s)}^n x^n(\kappa_2(s))ds + E\int_0^{t_l} \sum_{j=1}^{d_1} |\beta_{\kappa_1(s)}^{n,j}|^2 ds,
$$

with $l = 1, 2, \cdots, n$. Growth condition 4.2.5 implies

$$
\begin{aligned}
& E|x^n(t_l)|^2 \\
\leq\ & E|x^n(t_0)|^2 + 2L_2 E\int_0^{t_l} (1 + |x^n(\kappa_2(s))|^2)ds + L_2 E\int_0^{t_l} (1 + |x^n(\kappa_1(s))|^2)ds \\
\leq\ & E|x^n(t_0)|^2 + 3L_2 T + 3L_2 \int_0^{t_l} E|x^n(\kappa_2(s))|^2 ds.
\end{aligned}
$$

Hence by discrete Gronwall inequality 2.3.8, we can get the existence of a constant $C > 0$ such that

$$
\sup_n \sup_{1\leq l\leq n} E|x^n(t_l)|^2 \leq C < \infty,
$$

where C depends on T, $E|x_0|$, and coefficient in the growth condition.

□

4.4 Proof of the Main Results

We begin to prove the last part of Theorem 4.2.1.

Proof of Theorem 4.2.1. To prove the existence of a solution to SDI (4.2.7), it remains to show the existence of $(\bar{\alpha}, \bar{\beta}) \in \mathcal{H}$ such that $G(\bar{\alpha}, \bar{\beta}) \leq 0$.

Observe that (α^n, β^n) is in a bounded set of $\mathcal{H}$, x^n is bounded in Y and $x^n(T)$ is bounded in $L_2(\Omega; \mathbb{R}^d)$. Thus, by Banach-Alaoglu Theorem (see book [7]) each of their subsequence has a weakly convergent subsequence. Let us denote their subsequences in the same way as the sequences themselves, and their limits denoted by

$$(\alpha^n, \beta^n) \rightharpoonup (\alpha^\infty, \beta^\infty), \text{ in } \mathcal{H} \; ; \tag{4.4.17}$$

$$x^n \rightharpoonup x_\infty, \text{ in } Y \; ; \tag{4.4.18}$$

$$x^n(T) \rightharpoonup x_\infty(T), \text{ in } L_2(\Omega; \mathbb{R}^d) \; . \tag{4.4.19}$$

Here "$\rightharpoonup$" denotes the weak convergence.

Let us define operator $I(\alpha, \beta)(t) := \int_0^t \alpha_s ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^j dw_s^j$. It is easy to see that I is a bounded linear operator from $\mathcal{H}$ into Y. Indeed, for $(\alpha, \beta) \in \mathcal{H}$,

$$\begin{aligned}
|I(\alpha, \beta)|^2 &= E\int_0^T |\int_0^t \alpha_s ds + \sum_{j=1}^{d_1} \beta_s^j dw_s^j|^2 dt \\
&\leq T \sup_{t \leq T} E|\int_0^t \alpha_s ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^j dw_s^j|^2 \\
&\leq 2T(E\int_0^T |\alpha_s|^2 ds + \sum_{j=1}^{d_1} E|\int_0^T \beta_s^j dw_s^j|^2) \\
&\leq 2TE\int_0^T (|\alpha_s|^2 + \sum_{j=1}^{d_1} |\beta_s^j|^2) ds \\
&< \infty,
\end{aligned}$$

where $|\beta| := (\sum_{i,j} \beta_{ij}^2)^{1/2}$ is the Hilbert Schmidt norm for matrix β. Hence, we have that

$$\int_0^t \alpha_s^n ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^{n,j} dw_s^j \rightarrow \int_0^t \alpha_s^\infty ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^{\infty,j} dw_s^j.$$

We define

$$x^\infty(t) := x_0 + \int_0^t \alpha_s^\infty ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^{\infty,j} dw_s^j.$$

Then clearly $x^\infty(t) = x_\infty(t)$ for $dt \times dP-$almost all $(t, \omega) \in [0, T] \times \Omega$. It is easy to see that $x^\infty(T) = x_\infty(T)$ $(a.s.)$.

The final part of is to prove for fixed $y \in Y$,

$$F_y(\alpha^\infty, \beta^\infty) \le \liminf_{n\to\infty} F_y(\alpha^n, \beta^n) \le 0. \tag{4.4.20}$$

We consider the function $G(\alpha^\infty, \beta^\infty) = \sup_{y\in Y} F_y(\alpha^\infty, \beta^\infty)$, where $x(t) = x_t(\alpha^\infty, \beta^\infty)$.

For the first inequality of (4.4.20), observe that

$$F_y(\alpha^n, \beta^n) = E|x^n(T)|^2 - E|x^n(t_0)|^2 + E\int_0^T \Big(-2y(t)\alpha_t^n - 2x^n(t)a(y(t))$$

$$+2y(t)a(y(t)) + \sum_{j=1}^{d_1} |b^j(y(t))|^2 - \sum_{j=1}^{d_1} 2\beta_t^{n,j} b^j(y(t))\Big) dt.$$

When $n \to \infty$, due to (4.4.16) – (4.4.18) we know that

$$E\int_0^T 2y(t)\alpha_t^n dt \to E\int_0^T 2y(t)\alpha_t^\infty dt;$$

$$E\int_0^T 2x^n(t)a(y(t))dt \to E\int_0^T 2x^\infty(t)a(y(t))dt;$$

$$E\int_0^T 2\beta_t^{n,j} b^j(y(t))dt \to E\int_0^T 2\beta_t^{\infty,j} b^j(y(t))dt.$$

For the term $E|x^n(T)|^2$, since $x^n(T)$ converges weakly to $x_\infty(T) = x^\infty(T)$ in $L_2(\Omega;\mathbb{R}^d)$,

$$\liminf_{n\to\infty} E|x^n(T)|^2 \ge E|x^\infty(T)|^2.$$

Thus, for some constant $d \ge 0$,

$$\liminf_{n\to\infty} E|x^n(T)|^2 = d + E|x^\infty(T)|^2.$$

Hence, we get the first inequality of (4.4.20)

$$F_y(\alpha^\infty, \beta^\infty) \le \liminf_{n\to\infty} F_y(\alpha^n, \beta^n).$$

For the second inequality, we know that,

$$F_y(\alpha^n, \beta^n) = E\int_0^T 2(x^n(t) - y(t))(\alpha_t^n - a(y(t))) + \sum_{j=1}^{d_1} |\beta_t^{n,j} - b^j(y(t))|^2 dt,$$

and $\alpha_t^n \in a(x^n(\kappa_2(t)))$, where $\beta_t^n = b(x^n(\kappa_1(t)))$. By monotonicity condition 4.2.3, we get

$$\begin{aligned}
& F_y(\alpha^n, \beta^n) \\
\le\ & E\int_0^T 2(x^n(\kappa_2^n(t)) - y(t))(\alpha_t^n - a(y(t))) + \sum_{j=1}^{d_1} |b^j(x^n(\kappa_2(t))) - b^j(y(t))|^2 dt \\
& + I_1 + I_2 \\
\le\ & I_1 + I_2,
\end{aligned}$$

with

$$I_1 := E\int_0^T 2(x^n(t) - x^n(\kappa_2(t)))(\alpha_t^n - a(y(t)))dt;$$

$$I_2 := \sum_{j=1}^{d_1} E\int_0^T |b^j(x^n(\kappa_1(t))) - b^j(x^n(\kappa_2(t)))|^2 dt.$$

Now we estimate each term:

$$\begin{aligned}
I_1 &\leq C\int_0^T (E\int_t^{\kappa_2(t)} |x^n(r)|^2 dr)^{1/2}(E|y(t)|^2)^{1/2}dt \\
&\leq C\int_0^T \left(E\int_t^{\kappa_2(t)} |\alpha_r^n|^2 dr + \sum_{j=1}^{d_1} dtE\int_t^{\kappa_2(t)} |\beta_r^{n,j}|^2 dr\right)^{1/2}(E|y(t)|^2)^{1/2}dt \\
&\leq Cn^{-1/2};
\end{aligned}$$

and

$$\begin{aligned}
I_2 &\leq CTE\int_{\kappa_1(t)}^{\kappa_2(t)} |x^n(r)|^2 dr \\
&\leq CTE\int_{\kappa_1(t)}^{\kappa_2(t)} |\alpha_r^n|^2 dr + \sum_{j=1}^{d_1} E\int_{\kappa_1(t)}^{\kappa_2(t)} |\beta_r^{n,j}|^2 dr. \\
&\leq Cn^{-1},
\end{aligned}$$

where C is a constant. Combining both estimates, we find that

$$F_y(\alpha^n, \beta^n) \leq Cn^{-1/2},$$

where C is a constant depending on $y \in Y$.

Hence, $F_y(\alpha^\infty, \beta^\infty) \leq 0$, which implies $G(\alpha^\infty, \beta^\infty) = \sup_y F_y(\alpha^\infty, \beta^\infty) \leq 0$. Consequently by Part (ii) of Proposition (4.3.3), we conclude that $\{x^\infty(t) : t \in [0, T]\}$ is a solution of stochastic inclusion (4.2.7). □

We can thus conclude that from the proofs presented above, there exists subsequence of the implicit approximation x^n which converges weakly in Y to a solution x^∞ of the stochastic differential inclusion (4.2.7), and $x^n(T)$ converges strongly in $L_2(\Omega; \mathbb{R}^d)$ to $x^\infty(T)$ with the same sequence. Since the solution to SDI (4.2.7) is unique, we get the convergence results hold for any sequences of approximations $x^n(t)$ and $x^n(T)$ as $n \to \infty$.

The following theorem is given the rate of convergence result:

Theorem 4.4.1. *Let $q \geq 1$ be any real number and assume that $E|x_0|^q < \infty$. Let Assumptions 4.2.2, 4.2.4 and 4.2.6 hold. Then there exists a constant C independent of n, such that*

$$E\sup_{0\leq t\leq T} |x(t) - x^n(t)|^q < Cn^{-q/4},$$

for all sufficiently large integers n, where $x^n(t)$ defined by (4.3.15) and $x(t)$ is a solution to SDI (4.2.7).

Proof. We can easily check that we can repeat the proof of Theorem 3.2.6 in the chapter 3. Remember that for $t \in [0,T]$, the implicit approximations $x^n(t)$ satisfy

$$x^n(t) = x_0 + \int_0^t \alpha_s^n ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^{n,j} dw_s^j,$$

and $\alpha_t^n \in a(x^n(\kappa_2(t)))$, where $\beta_t^n = b(x^n(\kappa_1(t)))$. Then by using the formula $b^2 - a^2 = 2b(b-a) - (b-a)^2$, for $t_i = i\Delta t$ and for any $i = 1, 2, \cdots, n$, we get

$$\begin{aligned}
&|x(t_i) - x^n(t_i)|^2 \\
\le\ & \int_0^{t_i} 2(x(\kappa_2(s)) - x^n(\kappa_2(s)))\Big(\alpha_s - \alpha^n_{\kappa_2(s)}\Big)ds \\
&+\sum_{j=1}^{d_1}\Big|\int_0^{t_i}\Big(\beta_s^j - \beta^{n,j}_{\kappa_1(s)}\Big)dw_s^j\Big|^2 \\
&+\sum_{j=1}^{d_1}\int_0^{t_i} 2(x(\kappa_1(s)) - x^n(\kappa_1(s)))\Big(\beta_s^j - \beta^{n,j}_{\kappa_1(s)}\Big)dw_s^j,
\end{aligned}$$

and $\alpha(t) \in a(x(t)), \alpha^n_{\kappa_2(t)} \in a(x^n(\kappa_2(t)))$, where $\beta_t = b(x(t)), \beta^n_{\kappa_1(t)} = b(x^n(\kappa_1(t)))$.

When raising both side to the power $q/2$ and taking expectations, for $i \le m \le n$, we obtain

$$E \max_{0\le i\le m} |x(t_i) - x^n(t_i)|^q \le C_q(I_1 + I_2 + I_3),$$

with

$$I_1 := E \max_{0\le i\le m}\Big(\int_0^{t_i} 2(x(\kappa_2(s)) - x^n(\kappa_2(s)))\big(\alpha(s) - \alpha^n_{\kappa_2(s)}\big)ds\Big)^{q/2};$$

$$I_2 := E \max_{0\le i\le m}\sum_{j=1}^{d_1}\Big(\int_0^{t_i}\Big(b^j(s) - b^j(x^n(\kappa_1(s)))\Big)dw_s^j\Big)^q;$$

$$I_3 := E \max_{0\le i\le m}\sum_{j=1}^{d_1}\Big(\int_0^{t_i} 2(x(\kappa_1(s)) - x^n(\kappa_1(s)))\big(b^j(s) - b^j(x^n(\kappa_1(s)))\big)dw_s^j\Big)^{q/2}.$$

By making use of assumptions, we can finish the proof in exactly the same methods as we completed the proof of Theorem 3.2.6,

$$I_1 \le C\sum_{k=0}^{m-1} E\max_{i\le k}|x(t_i) - x^n(t_i)|^q\Delta t + Cn^{-q/2} + Cn^{-q/4};$$

$$I_2 \le C\sum_{k=0}^{m-1} E\max_{i\le k}|x(t_i) - x^n(t_i)|^q\Delta t + Cn^{-q/2};$$

$$I_3 \leq \frac{1}{2} E \max_{i \leq k} |x(t_i) - x^n(t_i)|^q + C \sum_{k=0}^{m-1} E \max_{i \leq m} |x(t_i) - x^n(t_j)|^q \Delta t + C n^{-q/2}.$$

Thus by discrete Gronwall's lemma 2.3.8,

$$E \max_{0 \leq i \leq m} |x(t_i) - x^n(t_i)|^q \leq C n^{-q/4},$$

where C is a constant which does not depend on n. Then by Lemma 3.3.3, we obtain

$$E \sup_{0 \leq t \leq T} |x(t) - x^n(t)|^q \leq C n^{-q/4}$$

holds for all sufficiently large integers $n \geq 1$. □

Chapter 5

Extension of Monotone and Maximal Monotone Mappings

The purpose of this chapter is to discover more information about monotone and maximal monotone mappings, which is required for the later proofs. The theory developed in this chapter plays an important role in weakening the linear growth condition when proving the existence of solutions for stochastic differential inclusions. The complete proof will be given in the later chapter.

Our motivation for extending monotone and maximal monotone mappings comes from the Krylov [21]. He further studied the properties of monotone mappings of a finite dimension space.

This chapter covers:

- section 5.1: this section states some properties of monotone mappings. This results generalize an extension of monotone functions obtained by N.V. Krylov [21], by using the similar techniques.
- section 5.2: this section gives the further concept of maximal monotone functions.

5.1 Extension of Monotone Mappings

A general local boundedness result for multi-valued monotone functions will be established, which implies the boundedness of monotone function on every ball in $\mathbb{R}^d$.

Remark 5.1.1. In what follows we note that for convenience we remain to use the definition of monotonicity function from [21]. A multi-valued function a defined on some set $D(a) \subset \mathbb{R}^d$ with values in $\mathbb{R}^d$, is called monotone, if

$$(x - y)(a(x) - a(y)) \geq 0, \text{ for any } x, y \in D(a).$$

We shall assume that D is a fixed bounded convex domain on $\mathbb{R}^d$.

Theorem 5.1.1. *Let a be a multi-valued monotone function on $\mathbb{R}^d$. Then a is bounded on every ball of radius r.*

In proving of this theorem, we will state some useful lemmas which are the main ingredients of the proof for this theorem.

Lemma 5.1.2. *Let $r > 0$, $x_0 \in D(a)$. Then there exists an $\varepsilon > 0$, depending only on d, such that if $y_{\pm 1}, y_{\pm 2}, \cdots, y_{\pm d} \in D(a)$, satisfying the conditions:*

i) the angles between $y_{\pm i} - x_0$ and $\pm e_i$ are smaller than ε,

ii) $r \leq |y_{\pm i} - x_0| < 2r$,

where $|i| = 1, 2, \cdots, d$, then

$$\varepsilon \sup\{|z| : z \in a(x_0)\} \leq \max_{1 \leq i \leq d} \inf\{|v| : v \in a(y_i)\}.$$

Remark 5.1.2. For this proof, it is sufficient for us to show that if $|\alpha| = 1$,

$$2\varepsilon r \leq \max_i\{\alpha(y_i - x_0) : |i| = 1, 2, \cdots d\},$$

for some $\varepsilon > 0$. So we first prove the following lemma.

Lemma 5.1.3. *Let $r > 0$ and $x_0 \in \mathbb{R}^d$. Then there exists an $\varepsilon > 0$ depending only on d, such that if $y_{\pm 1}, y_{\pm 2}, \cdots, y_{\pm d}$ are vectors from $\mathbb{R}^d$ satisfying*

i) the angle between $y_{\pm i} - x_0$ and $\pm e_i$ is smaller than ε,

ii) $r \leq |y_{\pm i} - x_0| < 2r$,

where $|i| = 1, 2, \cdots, d$, then

$$2\varepsilon r \leq \max_i\{\alpha(y_i - x_0) : |i| = 1, 2, \cdots d\} \tag{5.1.1}$$

for any $\alpha \in \mathbb{R}^d$ with $|\alpha| = 1$.

Proof. We may assume that $x_0 = 0$, since we can consider $y_i - x_0$ in place of y_i.

Indeed, let $\alpha = \sum_{i=1}^d \alpha_i e_i$, $|\alpha_l| = \max\{|\alpha_1|, \cdots, |\alpha_d|\}$, where $|\alpha_l| \geq 1/\sqrt{d}$.

Then

$$\alpha y_l = (\sum_{i=1}^d \alpha_i e_i) y_l = \alpha_l e_l y_l + \sum_{i \neq l, i=1}^d \alpha_i e_i y_l.$$

It is easy to see that,

$$\alpha y_l = \sum_{i=1}^d \alpha_i e_i y_l \geq \alpha_l (e_l y_l) - \sum_{i \neq l, i=1}^d |\alpha_i (e_i y_l)|. \tag{5.1.2}$$

Since $e_1, e_2, \cdots, e_d$ is an orthonormal basis,

$$\sum_{i=1}^d |(e_i y_l)|^2 = \sum_{i \neq l, i=1}^d |(e_i y_l)|^2 + |(e_l y_l)|^2 = |y_l|^2,$$

so for fixed l, if the angle between e_l and y_l is smaller than ε, and $\varepsilon > 0$,

$$\sum_{i \neq l, i=1}^{d} |(e_i y_l)|^2 = |y_l|^2 - |(e_l y_l)|^2 \leq \varepsilon^2 |y_l|^2.$$

Since

$$(e_l y_l)^2 \geq |y_l|^2 \cos^2 \varepsilon = |y_l|^2 (1 - \sin^2 \varepsilon) \geq |y_l|^2 (1 - \varepsilon^2).$$

Hence

$$\sum_{i \neq l, i=1}^{d} |(e_i y_l)| \leq \varepsilon |y_l|.$$

According to (5.1.2), we get that

$$\begin{aligned} \alpha y_l &\geq |\alpha_l||y_l|\sqrt{(1-\varepsilon^2)} - \varepsilon \sum_{i \neq l, i=1}^{d} |\alpha_i||y_l| \\ &\geq \frac{r}{\sqrt{d}}\sqrt{(1-\varepsilon^2)} - 2r\sqrt{d}\varepsilon. \end{aligned}$$

Let $2\varepsilon \leq \frac{1}{\sqrt{d}}\sqrt{(1-\varepsilon^2)} - 2\sqrt{d}\varepsilon$, then we get that ε is a fixed constant depending only on d, i.e. $\varepsilon = \varepsilon(d)$, this implies (5.1.1).

□

Proof of Lemma 5.1.2. Apply the result of Lemma 5.1.3 with $\alpha := \frac{z}{|z|}$, where $0 \neq z \in a(x_0)$, and together with the monotonicity condition, we get

$$\begin{aligned} 2\varepsilon r|z| \leq \max_{1 \leq |i| \leq d} z(y_i - x_0) \leq \max_{1 \leq |i| \leq d} v_i(y_i - x_0) \leq \max_{1 \leq |i| \leq d} |v_i||y_i - x_0| \\ \leq 2r \max_{1 \leq |i| \leq d} |v_i|, \end{aligned}$$

for any $v_i \in a(y_i)$.

Hence

$$\varepsilon |z| \leq \max_{1 \leq |i| \leq d} \inf\{|v| : v \in a(y_i)\}.$$

Thus we get

$$\sup\{|z| : z \in a(x_0)\} \leq \frac{1}{\varepsilon} \max_{1 \leq |i| \leq d} \inf\{|v| : v \in a(y_i)\},$$

where $|i| \in \{1, 2, \cdots d\}$. □

This lemma immediately allows us to conclude the following statement:

Lemma 5.1.4. *Let D be a fixed bounded convex set of $\mathbb{R}^d$ such that $D(a)$ is everywhere dense in D. Then for every compact $\Gamma \subset D$, the function a is bounded on $\Gamma \cap D(a)$.*

Proof. Let Y be a set of the form $\{y_{\pm1}, y_{\pm2}, \cdots, y_{\pm d}\}$, where $y_{\pm i} \in D(a)\}$. For $r > 0$ and fixed ε, let

$$U_{r,Y} = \{x \in \mathbb{R}^d : x \text{ satisfies the conditions } (i) \text{ and } (ii) \text{ in Lemma 5.1.2}\}, i.e.,$$

$$r < |y_{\pm i} - x| < 2r, \text{ and}$$

$$\text{the angles between } y_{\pm i} - x \text{ and } \pm e_i \text{ are smaller than } \varepsilon,$$

then $U_{r,Y}$ is an open set, $\Gamma \subset \cup_{r,Y} U_{r,Y}$. For the compactness of Γ, every open covering of Γ has finite subcovering, i.e. , $\Gamma \subset \cup_{l=1}^{N} U_{r_l,Y_l}$. Hence for all $x \in \Gamma$, from the above lemma, we may get that a is bounded on $\Gamma \cap D(a)$. □

Proof of Theorem 5.1.1. This theorem follows immediately from the application of the above lemmas. □

The main sources for the following part is in [21]. We present some results that we will use in our later chapters. From now on we study the monotone function a defined on $\bar{D}$ bounded on $\bar{D}$ instead of on the dense subset of D. Lemma 5.1.4 allows us to extend the monotone function a defined on a dense subset $D(a)$ in D to the D, taking as $a(x)$ for $x \in D \setminus D(a)$ any limit point of $a(x_n)$, where $x_n \to x$, $x_n \in D(a)$. Since the so-extended function remains monotone and will be locally bounded in D, one can reduce the study of an arbitrary monotone function defined on a dense subset of D to a monotone function bounded on $\bar{D}$ defined on $\bar{D}$.

The following functions will be useful in studying the properties of monotone functions.

Let

$$\begin{aligned} R(a, x, y) &= \sup_{x'}\{(a(x') - y)(x - x') + xy, x' \in F\} \\ &= \sup_{x'}\{(x - x')a(x') + x'y, x' \in F\}, \end{aligned} \tag{5.1.3}$$

where F is a dense subset defined on D. Let set $Z \subset \mathbb{R}^d \times \mathbb{R}^d := \{(x, y)\}$ be monotone.

Lemma 5.1.5. *Let a be a bounded monotone function on $\bar{D}$. Then*

1. *$R(a, x, y)$ is a function of (x, y) on $\mathbb{R}^d \times \mathbb{R}^d$.*

2. *$R(a, x, y) \geq xy$ on $\bar{D} \times \mathbb{R}^d$, $R(a, x, a(x)) = xa(x)$ on $\bar{D}$.*

3. *R does not change if in (5.1.3) one replaces F by D.*

4. *Set* $Z(a) = \{(x,y) : x \in \bar{D},\ R(a,x,y) = xy\}$. *If the integer* $n \geq 1$, $\alpha_i \geq 0$, $(x_i, y_i) \in Z(a), i = 1, \cdots, n$, $\sum_i \alpha_i = 1$, *then*

$$\sum_{i=1}^{n} \alpha_i x_i y_i \geq \sum_{i=1}^{n} \alpha_i x_i \sum_{i=1}^{n} \alpha_i y_i. \tag{5.1.4}$$

5. *The set* $Z(a)$ *is closed and monotone.*

6. *If* $(x_n, y_n) \to (x,y), x_n \in \bar{D}$, *and bounded monotone function* a_n *is defined on* $\bar{D}$ *with* $a_n \to a$ *on* F, *then*

$$\liminf_{n\to\infty} R(a_n, x_n, y_n) \geq R(a,x,y).$$

7. *For any* $x \in D$, *the set* $Z(a,x) = \{y : (x,y) \in Z(a)\}$ *is the convex hull of the set of partial limits of* $a(x_n)$ *as* $x_n \to x$.

8. $\{x \in \bar{D} : R(u,x,y) = xy\} \neq \emptyset$, *for any* $y \in \mathbb{R}^d$.

We refer to Appendix for the detailed proofs.

5.2 Properties of Maximal Monotone Mappings

We have seen that Lemma 5.1.4 allows us to extend the monotone function a defined on a dense subset $D(a)$ in D to the D, taking as the convex hull of partial limits of $a(x_n)$ (see result 7 in Lemma 5.1.5.) Such so-extended function remains monotone and will be locally bounded in D. Hence we can extend the result concerning monotone mappings into maximal monotone. A main goal of this section is to examine the consequence of the monotone set $Z(a)$. An easy consequence of the following version of the lemma states that $Z(a)$ is a maximal monotone set.

We make the following observation on the maximal monotone mappings.

Lemma 5.2.1. *From the result 7 in Lemma 5.1.5,* $Z(a)$ *is the convex hull of the set of partial limits of* $a(x_n)$ *as* $x_n \to x$, *then* $Z(a)$ *is a maximal monotone set when* x *is restricted to* $\bar{D}$.

Proof. The monotonicity of $Z(a)$ is obtained from result 5 in the previous Lemma, when one takes $n = 2$, $\alpha_1 = \alpha_2 = \frac{1}{2}$ in (5.1.4).

To show that $Z(a)$ is maximal monotone, assume that for some set $(x,y), x \in \bar{D}$, such that $Z(a) \cup \{x,y\}$ is monotone, this means, in particular, that

$$(a(x') - y)(x - x') \leq 0, \forall\, x' \in F.$$

This gives $\sup_x\{(a(x') - y)(x - x'), \forall\, x' \in F\} = 0$.

Then we get $R(a, x, y) = xy$. From result 4 in the above lemma, it shows $(x, y) \in Z(a)$. In other words, $Z(a)$ is a maximal monotone set. □

Definition 5.2.1. We say that a multi-valued function a is maximal monotone on a set D, if it does not have a proper extension to a monotone function defined on D.

The following Theorem shows how to establish a maximal monotone function from a monotone function.

Theorem 5.2.2. *Let $a(x) : x \in R^d$ be a monotone function. Define $\tilde{a}(x)$ is convex hull of the partial limits $a(x_n)$, when $x_n \to x$. Then $\tilde{a}$ is a maximal monotone function.*

Proof. First, we prove the monotonicity. We have to check that for any $x, y \in R^d$, $(x - y)(\tilde{a}(x) - \tilde{a}(y)) \geq 0$.

Let us take a ball B_R for some sufficiently large R, such that $x, y \in B_R$. Then by the previous result $\tilde{a}|_{B_R}$ is a maximal monotone function restricted to B_R.

Then we show that $\tilde{a}$ is a maximal monotone function. Take $x, y \in \mathbb{R}^d$, such that $\tilde{a} \cup \{(x, y)\}$ is monotone. Then its restriction to any ball B_R containing x is monotone, and it is contained in a maximal monotone set restricted to B_R. Hence $y \in$ $conv\{\lim_n(a(x_n)) :$ for $x_n \to x\}$, which means $\tilde{a}$ is maximal monotone. □

Chapter 6

Solution of Stochastic Differential Inclusion without the Linear Growth Condition

It is the intention of this chapter to keep concentrating on the existence of solutions to stochastic differential inclusions. In chapter 4, the linear growth condition to ensure the existence of solutions are quite strong, but can be weakened in this chapter. The properties of (maximal) monotone functions, applied with truncation methods, yield the existence of solution while under a growth condition which is weaker than the usual linear growth condition.

6.1 Introduction

In this chapter we still consider stochastic differential inclusions:

$$dx(t) \in a(t, x(t)dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j, \tag{6.1.1}$$

with a multi-valued drift term. Hence the existence and uniqueness of solutions taking values in $\mathbb{R}^d$ is proved in chapter 4 under some strong conditions for the coefficients a, b. a, b are required to satisfy the usual linear growth condition of the type

$$|a(t,x)|^2 + |b(t,x)|^2 \leq K(1 + |x|^2),$$

together with the monotonicity condition. In chapter 4 we approximate the solution by implicit approximation schemes. In that case, linear growth condition ensures that the drift term α^n of implicit approximation solutions is still in a bounded set (see, lemma 4.3.5 in chapter 4). Then the SDI admits one and only one solution by means of minimization method. In chapter 5, we further study the extension of monotone functions. It is shown that (multi-valued) monotone function is bounded in the ball over $\mathbb{R}^d$. Here is our motivation for continuous exploring the existence of solutions to SDIs. Our idea is based on the technique of truncating a maximal monotone function. If we can prove that there exists a maximal monotone function a_R such that $a_R = a$ on the ball B_R, then our restriction on growth is essentially weaker. We will make use of the special structure of maximal monotone functions to overcome the lack of boundedness.

Showing the existence of solutions to SDIs stems from N. V. Krylov's paper [19], where he employed truncation method on the monotone pair. This technique provide the inspiration to apply it to maximal monotone functions. It is known that such technique has already been used by many authors such as I. Gyöngy and N.V.Krylov in [13], where they gave a way for truncating monotonic pairs but did not preserve monotonicity afterwards. Later proof in N.V. Krylov [19] indicated truncation method for monotonicity functions, which leads to the monotonicity functions being bounded and still monotonic. In this paper, N.V. Krylov solved the truncation problem under the conditions that if a is a monotone and continuous function, then there exists a continuous bounded monotone function function a_R, such that $a_R = a$ on the ball B_R. Here we can already see the similarity to our situation. Hence our proofs are mainly based on this method used by Krylov. Influenced by his work, we will make use of this approach on truncating the possibly discontinuous maximal monotone drift term, then get the existence of the solution for the SDI. Our assumptions are similar to those in chapter 4, but it should be mentioned that if we apply the truncation method to the maximal monotone function a, then growth condition can be weaker than the usual linear growth condition.

This chapter is organized as follows:

- section 6.2: this section formulates the main theorem and necessary conditions;
- section 6.3: this section presents some preliminary lemmas about the truncation methods;
- section 6.4: this section gives the proof of the main theorem.

6.2 Formulation of the Results

Once again we consider stochastic differential inclusion (SDI) of the following form on domain $\mathbb{R}^d$:

$$\begin{cases} dx(t) \in a(t, x(t)dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j, \\ x(0) = x_0, \end{cases} \tag{6.2.2}$$

where x_0 is an $\mathcal{F}_0$-measurable random vector with values in $\mathbb{R}^d$, and $E|x_0|^2 < \infty$, $a : [0, +\infty) \times \mathbb{R}^d \to 2^{\mathbb{R}^d}$ is a multi-valued function, b is a Borel function on $[0, +\infty) \times \mathbb{R}^d$ taking values in $\mathbb{R}^{d \times d_1}$ and continuous in x, with norm $|b| = (\sum_{i=1}^{d} \sum_{j=1}^{d_1} b_{ij})^{1/2}$.

We use the following assumptions from [9]:

Assumption 6.2.1. Let D be a domain in $\mathbb{R}^d$. Assume there exist an increasing sequence of bounded sub-domains $D_1 \subset D_1 \subset \cdots$ and a non-negative function $V \in C^{1,2}([0, +\infty) \times D;\ \mathbb{R})$ such that the following conditions hold:

(i) $\cup_{k=1}^{\infty} D_k = D$, and for every $k, t \in [0, k]$

$$\sup_{x \in \bar{D}_k} |a(t, x)| \leq M_k, \ \sup_{x \in \bar{D}_k} |b(t, x)|^2 \leq M_k,$$

where M_k is a constant;

(ii)

$$\begin{aligned} LV(t, x) &\leq MV(t, x),\ \forall\, t \in [0, T],\ x \in D, \\ V_k(T) &:= \inf_{x \in \partial D_k,\ t \leq T} V(t, x) \to \infty \end{aligned}$$

as $k \to \infty$ for every finite T, where $M = M(T)$ is a constant, ∂D_k denotes the boundary of D_k and L is the differential operator

$$L := \frac{\partial}{\partial t} + \sum_{i}^{d} a_i(t, x) \frac{\partial}{\partial x^i} + \frac{1}{2} \sum_{i,j}^{d} (bb^T)_{ij}(t, x) \frac{\partial^2}{\partial x^i \partial x^j};$$

(iii) $P(x_0 \in D) = 1$.

Under our assumptions above the solutions of SDI (6.2.2) will never leave D, therefore the values of a, b outside D are irrelevant and for convenience, we define $a(t, x) = 0,\ b(t, x) = 0$ for $x \notin D, t \geq 0$.

Definition 6.2.1. By solution of the SDI (6.2.2) we mean an $\mathcal{F}_t$-adapted process $x(t)$ lives in D, in other words, it does not leave D and satisfies SDI (6.2.2).

An explanation of the definition can be found in the following statement from [9]:

Lemma 6.2.1. *Let $x(t)$ be an $\mathcal{F}_t$-adapted process defined for all $t \geq 0$. Assume that $x(t)$ satisfies SDI (6.2.2) for $t < \tau := \inf\{t : x(t) \notin D\}$, and assume (i) though (iii). Then $\tau = \infty$ (a.s.).*

Proof. Define τ^k as the first exit time of $x(t)$ from D_k. Obviously $\tau^k \uparrow \tau$. Therefore to prove the lemma it suffices to show that for any k and $\delta, T > 0$ we have

$$P(\tau^k \leq T) \leq P(x_0 \notin D_k) + P(V(0,x_0) \geq \log\frac{1}{\delta}) + \frac{1}{\delta V_k(T)} \exp \int_0^T M(t)dt. \tag{6.2.3}$$

Indeed,

$$\begin{aligned} P(\tau^k \leq T) &\leq P(\tau^k \leq T, x_0 \in D_k) + P(x_0 \notin D_k) \\ &\leq P(\tau^k \leq T, x_0 \in D_k, \exp(-V(0,x_0)) \leq \delta) \end{aligned}$$

$$+P(\tau^k \leq T, x_0 \in D_k, \exp(-V(0,x_0)) > \delta) + P(x_0 \notin D_k). \tag{6.2.4}$$

Now the first term of right-hand side of (6.2.4)

$$P(\exp(-V(0,x_0)) \leq \delta) = P(-V(0,x_0) \leq \log\delta) = P(V(0,x_0) \geq \log\frac{1}{\delta}).$$

To estimate the second term on the right-hand side of (6.2.4), we use assumption (ii) and apply Itô's formula to $\gamma(t)V(t,x(t))$ where

$$\gamma := \exp[-\int_0^t M(s)ds - V(0,x_0)].$$

Then it follows that for all $t \leq T$,

$$\gamma(t)V(t\wedge\tau^k, x(t\wedge\tau^k))\chi_{\tau^k>0} \leq \gamma(0)V(0,x_0) + m^k(t),$$

where $m^k(t)$ is a continuous local martingale starting from 0. Hence for any $R > 0$

$$P\{\sup_{t\leq\tau^k} \gamma(t)V(t,x(t))\chi_{\tau^k>0} \geq R\} \leq \frac{1}{R}E(\gamma(0)V(0,x_0)) \leq \frac{1}{R}.$$

As a result, we obtain

$$\begin{aligned} &P(\tau^k \leq T, x_0 \in D_k, \exp(-V(0,x_0)) > \delta) \\ \leq\ & P\Big(\sup_{t\leq\tau^k} \gamma(t)V(t,x(t)) \geq V_k(T)\delta \exp(-\int_0^T M(t)dt)\Big) \\ \leq\ & \frac{\exp(\int_0^T M(t)dt)}{V_k(T)\delta}, \end{aligned}$$

which implies (6.2.3).

Notice that $P(\tau^k \leq T) \to 0$ as $k \to \infty$. Indeed since $V(0,x_0) < \infty$, and $\log\frac{1}{\delta} \to \infty$ as $\delta \to 0$, this gives $P(V(0,x_0) \geq \log\frac{1}{\delta}) \to 0$; by assumption (ii) $V_k(T) \to \infty$ as $k \to \infty$, this gives $P(\frac{1}{\delta V_k(T)} \exp\int_0^T M(t)dt) \to 0$; together with the fact $P(x_0 \notin D_k) = 0$, we obtain the assertion of the lemma. □

Assumption 6.2.2. Assume that a is a maximal K-monotone function such that for each k its restriction to D_k has a maximal monotone extension a_k to the whole $\mathbb{R}^d$ which satisfies the linear growth condition.

Remark 6.2.1. If D_k is a ball then we have proved (see chapter 4) that the extension a_k regenerated in the above Assumption exists. In the later part of this chapter we consider the situation when the set D_k contains a ball.

Assumption 6.2.3 (Local Lipschitz of b). There exists a constant $L_k > 0$, such that

$$|b(t,x) - b(t,y)| \leq L_k|x-y|,$$

for $t \geq 0, x, y \in D_k$.

Theorem 6.2.2. *Assume that a is a maximal K-monotone function defined on D, and b is locally Lipschitz defined on D. Let Assumption 6.2.2 hold. Then SDI (6.2.2) has a unique solution $x(t)$ which lives in D for all $t \geq 0$.*

Proof. Since b is locally Lipschitz in D for every k, we have a bounded measurable function b_k on the whole $[0, +\infty) \times \mathbb{R}^d$, such that b and b_k are agree on $[0,T] \times D_k$, and b_k is locally Lipschitz on $\mathbb{R}^d$. Define the stopping time

$$\tau^k := \inf\{t \geq 0, x^k(t) \notin D_k\} \wedge T.$$

From the proof of Theorem 4.2.1, we can get that for $t \leq \tau^k$

$$x^k(t)\mathbf{1}_{\tau^k>0} = x_0 + \int_0^t \alpha_s ds + \sum_{j=1}^{d_1} \int_0^t b^j(s, x^k(s))dw_s^j,$$

where $\alpha_t \in a(t, x^k(t))$.

Let l, k be integers, such that $k \leq l$. Set $\tau^{kl} := \tau^k \wedge \tau^l$. Then by using Itô's formula,

$$\exp(-Lt)|(x^k - x^l)(t \wedge \tau^{kl})|^2\mathbf{1}_{\tau^{kl}>0} \leq m^{kl}(t),$$

where L is a sufficiently large constant, and m^{kl} is a local martingale starting from 0. Hence $x^k(t) = x^l(t)$ for $t \leq \tau^{kl}$. Since $k \leq l$, then $\tau^k \leq \tau^l$. There exists a stopping time τ with

$$\tau := \lim_{k\to\infty} \tau^k = \inf\{t \geq 0, x(t) \notin D\} \wedge T,$$

such that $x(t) := \lim_{k\to\infty} x^k(t)$, satisfies the SDI (6.2.2). Thus by Lemma 6.2.1, we know that $\tau = T$. The uniqueness can be achieved in the same way as Theorem 4.2.1. The proof is completed. □

Notice that taking $V(t,x) := e^{(-Mt)}(1+|x|^2)$ in the case of $D := \mathbb{R}^d$, $D_k = \{x \in \mathbb{R}^d : |x| < k\}$, condition *(ii)* in Assumption 6.2.2 can be restated as follows:

Assumption 6.2.4 (Growth Condition). There exists a constant M, such that

$$2xa(t,x) + |b(t,x)|^2 \leq M(1+|x|^2), \tag{6.2.5}$$

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

Our goal is to present the existence of solution to SDI without the linear growth condition. More precisely, if the linear growth condition is weakened by Assumption 6.2.4, the SDI is still solvable by means of truncating on the maximal monotone drift term a. Below is the main theorem in this chapter.

Theorem 6.2.3. *Assume that a is a maximal $K-$monotone function defined on $\mathbb{R}^d$, b is locally Lipschitz defined on $\mathbb{R}^d$. Under Assumption 6.2.4, stochastic differential inclusion (6.2.2) has one and only one solution.*

The uniqueness of the solution can be obtained in the usual way. The existence of the solution will be followed after some preliminary lemmas. For simplicity, we consider time-independent case.

6.3 Preliminary Lemmas

Before we prove the main Theorem 6.2.3, let us give the truncation procedure. This is the most technically hard part in this chapter. We will discuss the method of truncating the maximal monotone function a, which leads to be bounded and again maximal monotone. For this purpose we require several propositions, which carry out the facts that will make the realization of truncation method.

Let B_R be the closed ball.

Proposition 6.3.1. *If a is a (possibly multi-valued) monotone function defined on the ball B_R, then for every $\bar{a} \in \mathbb{R}^d$, there exists $x \in B_R$, such that*

$$(x-y)(\bar{a} - a(y)) \leq 0.$$

Proof. It follows directly from the last statement in the Lemma 5.1.5, which is equivalent to the statement $\{x \in B_R : R(a,x,\bar{a}) = x\bar{a}\} \neq \emptyset$, for any $\bar{a} \in \mathbb{R}^d$, with $R(a,x,\bar{a}) = \sup_y\{(-\bar{a} - a(y))(x-y) + x\bar{a},\ y \in B_R\}$ by applying $y := -\bar{a}$ in Lemma 5.1.5. We leave the proof to Lemma 5.1.5 in the Appendix. □

Proposition 6.3.2. *Assume that $x \in \mathbb{R}^d$ is a non-zero vector. Set $Z_x = \{z \in \mathbb{R}^d : zx < 0\}$. Let $h : Z_x \to \mathbb{R}^d$ be a bounded function such that $h(z)z \leq 0$ for every $z \in Z_x$. Then the closure $\bar{H}$ of the convex hull of $\{h(z) : z \in Z_x\}$ contains μx for some $\mu \geq 0$.*

Proof. Assume that $C := \{\mu x : \mu \geq 0\}$ and $\bar{H}$ are disjoint. Then C and $\bar{H}$ are disjoint convex sets, C is closed and $\bar{H}$ is compact. Therefore by the theorem of "strong" separation of convex sets (as we can see Theorem 2.3.9), there exist numbers $\gamma_1 < \gamma_2$ and a continuous linear functional on $\mathbb{R}^d$, i.e., a vector $a \in \mathbb{R}^d$ such that

$$ay < \gamma_1 < \gamma_2 < aw,$$

for all $y \in C$ and $w \in \bar{H}$. In particular, $ax < \gamma_1 < \gamma_2 < aw$ for every $w \in \mathbb{R}^d$. Hence $ax \leq 0$, because otherwise $\mu ax > \gamma_1$ for sufficiently large μ. Hence we have

$$ay \leq 0 < \gamma_1 < \gamma_2 < aw \tag{6.3.6}$$

for all $y \in C$ and $w \in \bar{H}$.

Let $a_\epsilon := a - \epsilon x$ for $\epsilon > 0$. Then $a_\epsilon x = ax - \epsilon |x|^2 < 0$, and from (6.3.6), we obtain

$$a_\epsilon w = (a - \epsilon x)w = aw - \epsilon xw > \gamma_2 - \epsilon K|x| > \gamma_1 > 0, \tag{6.3.7}$$

for all $w \in \bar{H}$ and for sufficient small $\epsilon > 0$, where $K = \sup\{|w| : w \in \bar{H}\} < \infty$.

On the other hand, since $a_\epsilon x < 0$, then for $a_\epsilon \in Z_x$ and $h(a_\epsilon) \in \bar{H}$, we have

$$h(a_\epsilon)a_\epsilon \leq 0,$$

which contradicts (6.3.7). Consequently, $\bar{H}$ contains μx for some $\mu \geq 0$. □

These two propositions motivates the following lemma. The next two lemmas extend Lemma 3 and Lemma 4 by N.V. Krylov [19] to possibly discontinuous maximal monotone function.

Lemma 6.3.3. *Let $\gamma, R > 0$ and let $a(x)$ be a maximal monotone function defined on the ball $B_R =: \{x \in \mathbb{R}^d : |x| \leq R\}$. Assume that $|a(x)|^2 - |a(0)|^2 > \gamma^2$ if $|x| = R$. Then $\{a(x) : |x| < R\} \supset B_\gamma$.*

Proof. Let us take any $\bar{a} \in B_\gamma$. From Proposition 6.3.1, there exists an $x \in B_R$, such that

$$(x - y)(\bar{a} - a(y)) \leq 0, \text{ for } y \in B_R.$$

We want to show that $|x| < R$ and $\bar{a} \in a(x)$. To this end we define

$$Z_x^- := \{z \in \mathbb{R}^d : zx < 0\}$$

and

$$Z_x^+ := \{z \in \mathbb{R}^d : -zx < 0\}.$$

Then for sufficiently small $\lambda > 0$, for every $z \in Z_x^-$ we have $|x + \lambda z| \in B_R$, and for every $z \in Z_x^+$, we have $|x - \lambda z| \in B_R$.

For all $z \in Z_x^-$, taking $y := x + \lambda z$ above, $\lambda z(a(x+\lambda z) - \bar{a}) \leq 0$. This yields $z(a(x+\lambda z) - \bar{a}) \leq 0$,. Then there exists a subsequence $\lambda_n \to 0$, such that $a(x + \lambda_n z) \to \tilde{a}_{z_x^-}$. Because a is maximal monotone on B_R, $\tilde{a}_{z_x^-} \in a(x)$ and

$$z(\tilde{a}_{z_x^-} - \bar{a}) \leq 0, \text{ this is true for any } z \in Z_x^-.$$

Then by the proposition (6.3.2) above, let H be the closed convex hull of the vectors $\{\tilde{a}_{z_x^-}\}$, then there exists $\mu \geq 0$, such that $\mu x \in \bar{H}$. Hence $\tilde{a}_{z_x^-} - \bar{a} = \mu x$, i.e.,

$$\bar{a} = \tilde{a}_{z_x^-} - \mu x.$$

Therefore repeating the above augments with Z_x^+ in place of Z_x^-, we get $\tilde{a}_{z_x^+} \in a(x)$ and there exists $\nu \geq 0$, such that $-\nu x \in \bar{H}$, where H is the closed convex hull of the vectors $\{\tilde{a}_{z_x^+}\}$ i.e.,

$$\bar{a} = \tilde{a}_{z_x^+} + \nu x.$$

Define $\bar{a} = \alpha \tilde{a}_{z_x^+} + \beta \tilde{a}_{z_x^-}$, with $\alpha = \frac{\mu}{\mu+\nu}$, and $\beta = \frac{\nu}{\mu+\nu}$, we see that $\bar{a} = \tilde{a}$, where $\tilde{a}$ is the linear combination of $(\tilde{a}_{z_x^+}, \tilde{a}_{z_x^-})$.

The last step is to show that $|x| < R$. Because of the monotonicity, $x\tilde{a}_{z_x^-} \leq xa(0)$. Assume that $|x| = R$ then

$$\begin{aligned}
\gamma^2 &\geq |\bar{a}|^2 = |\tilde{a}_{z_x^-}|^2 - 2a_{z_x^-}\mu x + \mu^2|x|^2 \\
&\geq \tilde{a}_{z_x^-}^2 - 2\mu x a(0) + \mu^2|x|^2 \\
&= |\tilde{a}_{z_x^-}|^2 - |a(0)|^2 + |\mu x - a(0)|^2 \\
&> \gamma^2,
\end{aligned}$$

which is impossible. Consequently $|x| < R$. We conclude that if $|\bar{a}| \leq \gamma$, then $\bar{a} \in \{a(x) : |x| < R\}$.

□

Now we can see how the truncation method is applied to maximal monotone function to get bounded maximal monotone functions. Notice that the second assumption of the following lemma is needed to ensure that $|\bar{a}| \leq \gamma$.

Lemma 6.3.4. *Let $\gamma, R > 0$, and $a(x)$ be a maximal monotone function defined on B_R. Assume the following conditions hold:*

(i) $(x-y)(a(x) - a(y)) \leq -\epsilon|x-y|^2$, $\forall\ x, y \in B_R$ with a fixed $\epsilon > 0$.

(ii) $|a(x)|^2 - |a(0)|^2 > \gamma^2$, if $|x| = R$, where $\gamma \neq 0$ is a fixed number.

We introduce

$$F(\tilde{a}, x) = \sup_{y \in D}(x-y)(\tilde{a} - a(y)),$$

where $D = \{x : |a(x)| \leq \gamma\} \cap B_R$. Then, *(a) $F(\tilde{a}, x)$ is continuous in $(\tilde{a}, x)$, $F(\tilde{a}, x) \geq 0$ for all $|\tilde{a}| < \gamma$, where $\tilde{a} \in \mathbb{R}^d$.*

(b) for each $x \in \mathbb{R}^d$, there exists $y(x) \in D$, such that $F(\hat{a}, x) = 0$, where $\hat{a}(x) \in a(y(x))$; here if $F(\tilde{a}, x) = 0$, $|\tilde{a}| < \gamma$, then $\tilde{a} = \hat{a}(x)$.

(c) $\hat{a}$ is monotonic in $\mathbb{R}^d$, and $y(x) = x$ for all $x \in D$.

Proof. First, by Lemma 6.3.3, we have $B_\gamma \subset \{a(x) : |x| < R\}$. Hence,

(a) for $|\tilde{a}| \leq \gamma$, we can find y, such that $|y| < R$, $\tilde{a} \in a(y)$. So

$$F(\tilde{a}, x) = \sup_{y \in D} (x - y)(\tilde{a} - a(y)) \geq 0.$$

Due to the boundedness of D and of the function $a = a(y)$ on D, F is continuous in $(\tilde{a}, x)$.

(b) Fix x, set $\Delta = (a(y) : y \in D)$, $\Gamma = \overline{conv}\Delta$. We prove the solvability of $F(\tilde{a}, x) = 0$ on Γ. Since $a = a(y)$ is maximal monotone, for each $y \in D$, the set $a(y)$ is convex and closed. Consequently, D is compact. Further we know that Γ is compact. Since $F \geq 0$, it suffices to show that

$$\min_{\tilde{a} \in \Gamma} F(\tilde{a}, x) \leq 0. \tag{6.3.8}$$

Take in D a countable everywhere dense subset x_i and define

$$\Psi^n = \max_{y \in \{x_1, x_2, \cdots, x_n\}} (x - y)(\tilde{a} - a(y)).$$

Clearly, Ψ^n is continuous in $(\tilde{a}, x)$, $F(\tilde{a}, x)$ is a bounded convex function on every bounded convex set. Hence it is continuous. By Dini's theorem, $\Psi^n \to F$ uniformly on Γ. So in order to prove (6.3.8), it suffices for us to prove that

$$\min_{\tilde{a} \in \Gamma} \Psi^n(\tilde{a}, x) \leq 0, \forall\, n. \tag{6.3.9}$$

For $p \in P := \{p = (p_1, \cdots, p_n) : p_i \geq 0, \sum_{i=1}^n p_i = 1\}$, consider the functions

$$\psi(p, \tilde{a}, x) = \sum_i p_i (x - x_i)(\tilde{a} - a(x_i)).$$

Obviously, $\Psi^n = \max\{\psi : p \in P\}$. By the minimax theorem the left side of (6.3.9) is equal to

$$\max_{p \in P} \min_{\tilde{a} \in \Gamma} \psi(p, \tilde{a}, x). \tag{6.3.10}$$

Noting that for $|\tilde{a}| \leq \gamma$, $F(\tilde{a}, x) \geq 0$. $F(a(x), x) = \sup_y (x - y)(a(x) - a(y)) = 0$, for $|x| \leq R$. Then by using the same method as in Lemma 5.1.5, we deduce that

$$\sum_i p_i x_i a(x_i) \leq \sum_i p_i x_i \sum_i p_i a(x_i). \tag{6.3.11}$$

Hence we get $\psi(p, \tilde{a}, x) \leq 0$. Consequently, (6.3.10) is negative, inequality (6.3.9) is proved, and we get the existence of a solution of the equation $F(\tilde{a}, x) = 0$ on Γ.

Next step is to show that this solution lies in Δ. For it, obviously, $(x - y, \tilde{a} - a(y)) \leq 0$, for all $y \in D$. From Lemma 6.3.3, $\tilde{a} \in a(y)$ for some $y \in D$, we see that $\tilde{a} \in \Delta$, let $y = y(x)$ denote such y.

Now let us prove the uniqueness of the solution of $F(\tilde{a}, x) = 0$ and the uniqueness of y such that $\tilde{a} \in a(y)$ on $U = \{|\tilde{a}| \leq \gamma\} \times \mathcal{L}(H, \mathbb{R}^d)$. Let $F(\tilde{a}_i, x) = 0$, $i = 0, 1$, $|\tilde{a}| \leq \gamma$, $\tilde{a}_s = \tilde{a}_1 s + \tilde{a}_0(1 - s)$. Since F is the sup of functions that are convex in $(\tilde{a}, x)$. But $F \geq 0$ on the convex set U and $(\tilde{a}_s, x) \in U$, therefore $F(\tilde{a}_s, x) = 0$ for $s \in \{0, 1\}$. We know that $\tilde{a}_s \in a(y_s)$ for some $y \in D$ and $(x - y, \tilde{a}_s - a(y)) \leq 0$ for all $y \in D$, $s \in [0, 1]$. This in particular implies that for every $s \in [0, 1]$, the function

$$(x - y_s)(\tilde{a}_r - a(y_s)) \tag{6.3.12}$$

is convex in r on $[0, 1]$, is non positive on $[0, 1]$ and equal to 0 at $r = s$. This is possible only if the function (6.3.12) is 0 for all $r \in [0, 1]$. The derivative of (6.3.12) with respect to r is 0. Since we have

$$\begin{aligned}(x - y_s)(\tilde{a}_r - a(y_s)) &= (x - y_s)(\tilde{a}_1 - a(y_s))r + (x - y_s)(\tilde{a}_0 - a(y_s))(1 - r) \\ &= (x - y_s)(\tilde{a}_1 - \tilde{a}_0)r + (x - y_s)(\tilde{a}_0 - a(y_s)),\end{aligned}$$

which gives

$$(x - y_s)(\tilde{a}_1 - \tilde{a}_0) = 0, \text{ and } (x - y_s)(\tilde{a}_0 - a(y_s)) = 0.$$

Hence for $s, s_1, s_2 \in (0, 1)$, one has

$$(x - y_s)(\tilde{a}_{s_1} - \tilde{a}_{s_2}) = 0, \ (y_{s_1} - y_{s_2})(\tilde{a}_{s_1} - \tilde{a}_{s_2}) = 0,$$

and

$$(y_{s_1} - y_{s_2})(a(y_{s_1}) - a(y_{s_2})) = 0.$$

Finally, we conclude that $y_{s_1} = y_{s_2}$, $\tilde{a}_1 \in a(y_{s_1}) = a(y_{s_2}) \ni \tilde{a}_0$, $\tilde{a}_0 = \tilde{a}_1$.

(c) The equalities $y(x) = x$ comes from the uniqueness of the solution of $\tilde{a} \in a(y)$ on D. The monotonicity of $\tilde{a}$ follows from (6.3.11), if we take $i = 2$, and $p_1 = p_2 = 1/2$. □

Theorem 6.3.5. *Let $\gamma, R > 0$ be constants and let $a(x)$ be a maximal monotone function defined on $\mathbb{R}^d$. Assume the following conditions hold:*

(i) $(x - y)(a(x) - a(y)) \leq -\epsilon|x - y|^2, \forall\, x, y \in B_R$ with a fixed $\epsilon > 0$.

(ii) $|a(x)|^2 - |a(0)|^2 > \gamma^2$, *such that* $|x| = R$, *where* $\gamma \neq 0$ *be a fixed number. Then there exists a maximal monotone bounded function* a_R *on* $\mathbb{R}^d$ *such that*

$$a_R(x) = a(x), \text{ for all } x \in D,$$

where $D = \{x : |a(x)| \leq \gamma\} \cap B_R$.

Proof. From the above lemma, we define that

$$a_R(x) \text{ as the maximal monotone extension of } \{a(y(x)) : x \in \mathbb{R}^d\}.$$

Then for $x \in D$, $a(y(x)) = a(x)$ is a consequence of the uniqueness of the solution of $F(a, x) = 0$ on B_R, and $y(x) = x$ on D follows from the uniqueness of solution of the equation $\tilde{a} \in a(y)$. The theorem is proved.

□

Remark 6.3.1. We will apply this theorem in a situation when the set D defined above contains a ball B_γ for a given $\gamma > 0$.

Lemma 6.3.6. *Let a be a maximal* $K-$*monotone function defined on the* $\mathbb{R}^d$. *Then there exists a maximal* $K-$*monotone function* a_R *such that* $a_R = a$ *on the closed ball* B_R *and* a_R *satisfies the linear growth condition.*

Proof. First, we define $\bar{a}(x) := a(x) - Lx$, where L is a constant. Notice that $\bar{a}(x)$ satisfies the conditions of Theorem 6.3.5 if L is sufficiently large. Indeed *(i)* clearly holds. To show *(ii)* we notice that for $|x| = R$ we have

$$|\bar{a}(x)^2| - |\bar{a}(0)|^2 = |a(x) - Lx|^2 - |a(0)|^2$$

$$\geq |a(x)|^2 + |Lx|^2 - 2Lx|a(x)| - |a(0)|^2$$

$$\geq |a(x)|^2 + L^2R^2 - 2LR|a(x)| - |a(0)|^2$$

$$\geq |a(x)|^2 + L^2R^2 - (\frac{L^2R^2}{2} + 2|a(x)|^2 - |a(0)|^2)$$

$$\geq \frac{1}{2}L^2R^2 - 2M_R^2,$$

where $M_R := \sup_{x \in B_R} |a(x)|$. Clearly, for L large enough *(ii)* is satisfied with $\gamma^2 := \frac{1}{4}L^2R^2 - M_R^2$. Moreover, for R large enough, $\gamma > M_R$, which means $D = B_R$ for such R. By Theorem 6.3.5 there exists a maximal monotone bounded function $\tilde{a}_R$ on $\mathbb{R}^d$ such that $\tilde{a}_R = \bar{a}$ on D. Hence there exists $a_R := \tilde{a}_R + Lx$ on $\mathbb{R}^d$ such that $a_R = a$ on D. Linear growth condition is clearly satisfied. □

6.4 Proof of Theorem 6.2.3

Now we are going to finish the proof of the Theorem 6.2.3.

Proof. For every $n \geq 1$, for sufficiently large $L = L(n)$, by applying the above lemma, we obtain that there exists a bounded maximal monotone function $a^n(x)$ on $\mathbb{R}^d$, such that $a^n(x) = a(x)$ on $D := B_n$.

We now consider SDI

$$dx^n(t) \in a^n(x(t))dt + \sum_{j=1}^{d_1} b^j(x(t))dw_t^j.$$

Moreover by virtue of Theorem 6.2.2, for every n there exists a unique process x^n, such that

$$x^n(t) = x_0 + \int_0^t \alpha_s^n ds + \sum_{j=1}^{d_1} \int_0^t \beta_s^{n,j} dw_s^j,$$

and $\alpha_t^n \in a^n(x) = a(x)$, where $\beta_t^n = b(x(t))$. Let us define the stopping time

$$\tau^n := \inf\{t \geq 0 : |x^n(t)| \geq n\}$$

and $\tau^{nm} := \tau^n \wedge \tau^m$. Then $x^n(t), x^m(t)$ are in the ball of radius n, m. It follows that the processes $x^n(t \wedge \tau^{nm}), x^m(t \wedge \tau^{nm})$ satisfy the same SDI

$$dx(t) \in a(x(t))\mathbf{1}_{t \leq \tau^{nm}} dt + \sum_{j=1}^{d_1} b^j(x(t))\mathbf{1}_{t \leq \tau^{nm}} dw_t^j, \tag{6.4.13}$$

$$x(0) = x_0.$$

The uniqueness of the solution of (6.4.13) implies that $x^n(t) = x^m(t)$ until τ^{nm}. Let m be an integer larger than n, then $\tau^n \leq \tau^m$ a.s. and there exists a stopping time τ with $\tau := \lim_{n\to\infty} \tau^n$ a.s. on $[0, \tau]$, such that

$$\alpha_t := \lim_{n\to\infty} \alpha_t^n,$$

$$\beta_t := \lim_{n\to\infty} \beta_t^n,$$

$$x(t) := \lim_{n\to\infty} x^n(t),$$

satisfies the SDI (6.2.2). Hence,

$$x(t \wedge \tau) = x_0 + \int_{[0,\ t\wedge\tau]} \alpha_s ds + \int_{[0,\ t\wedge\tau]} \sum_{j=1}^{d_1} \beta_s^j dw_s^j,$$

hold. Finally by Lemma 6.2.1, we know that $\tau = \infty$. $x(t)$ is the solution of our SDI. The proof the theorem is completed. □

Let us conclude the proof of the main theorem 6.2.3. As we can see also in [19] and [13], such proofs usually involves two parts. The first part is to show the existence of a solution before the first exit from D; The second part is to deduce the time of departure to infinity of the solution is equal to infinity.

Chapter 7

Solving SDEs via Mini-Max Theorems

It is shown that SDEs can be demonstrated related to mini-max problems in certain infinite dimensional spaces. We will show that the proper mini-max theorem provides a simple proof of the existence of strong solutions to SDEs.

7.1 Introduction

Mini-max theorems are useful and important tools in various fields of mathematics. It is the fundamental theorem of game theory, and was first proved by von Neumann in 1928. We are going to study the connection between SDEs and mini-max problems. Adapting an idea of N.V. Krylov, we show that SDEs can be associated with mini-max problems in suitable infinite dimensional spaces. Specifically, if the coefficients of an SDE satisfy the so called monotonicity condition, then one can construct a mini-max problem such that its saddle point is a solution of the given SDE. This connection between SDEs and mini-max problems can be applicable in many ways. In this chapter, we will give a constructive theorem of mini-max problems, inspired by the well-known mini-max theorem from Fan [6]. This theorem is then applied to give an alternative proof of the existence of (strong) solutions to SDEs. We will combine this with minimization method. Such method has been introduced in chapter 4 on SDIs problems. We would like to present that the suitable mini-max theorem can be applied to give a very simple proof of the traditional SDEs problems.

The organization of this chapter is given as follows:

- section 7.2: this section formulates the problem and the main result;
- section 7.3: this section constructs mini-max theorem, which will be used in the main proof;

- section 7.4: this section gives the finial proof by application of mini-max theorem.

7.2 Formulation of the Results

We consider the following stochastic differential equation:

$$\begin{cases} dx(t) = a(t, x(t))dt + \sum_{j=1}^{d_1} b^j(t, x(t))dw_t^j, \\ x(0) = x_0, \end{cases} \tag{7.2.1}$$

where $a : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^d$ and $b : [0, \infty) \times \mathbb{R}^d \to \mathbb{R}^{d \times d_1}$ are measurable functions, and x_0 is an $\mathcal{F}_t$−measurable random variable with values in $\mathbb{R}^d$, independent of w, and such that $E|x_0|^2 < \infty$.

We first introduce the conditions we will be placing on the drift and diffusion coefficients.

Assumption 7.2.1 (Monotonicity of (a, b)). The pair (a, b) satisfies

$$2(x - y)(a(t, x) - a(t, y)) + \sum_{j=1}^{d_1} |b^j(t, x) - b^j(t, y)|^2 \le M(t)|x - y|^2,$$

for all $x, y \in \mathbb{R}^d$, $t \ge 0$, where M is an $\mathcal{F}_t$−adapted non negative process such that $\int_0^T M(t)dt < \infty$ a.s., for every $T > 0$.

Assumption 7.2.2 (Boundedness of (a, b)). There is an $\mathcal{F}_t$−adapted process R such that

$$|a(t, x)|^2 + \sum_{j=1}^{d_1} |b^j(t, x)|^2 \le R(t) \ (a.s.),$$

for all $t \ge 0$, $x \in \mathbb{R}^d$, and $\int_0^T R(t)dt \le K^2$, for some deterministic constant K.

Assumption 7.2.3 (Continuity). $a(t, x)$ is continuous in x for all $t \ge 0$.

Remark 7.2.1. Notice that by taking $x(t) - x_0$ in place of $x(t)$ we may and will assume that $x_0 = 0$.

The following theorem is well-known:

Theorem 7.2.1. *Let Assumptions 7.2.1, 7.2.2 and 7.2.3 hold. Then stochastic differential equation (7.2.1) has a unique solution on* $[0, T]$.

This theorem was first proved in [18] and then it is generalized in [13]. As we shall see, the proofs given in these two papers are rather complicated. Later Krylov gave a simple proof of the existence of a solution in [20]. We shall give another simple proof based on our extremal approach and on the mini-max theorem. It is the purpose of this chapter to display the connection between SDEs and mini-max problems.

Proof of Theorem 7.2.1. The uniqueness of solution to SDE (7.2.1) follows from monotonicity condition (7.2.1) by Itô's formula. Indeed, if $x^1(t)$ and $x^2(t)$ are two possible solutions of SDE (7.2.1), then we obtain

$$d(e^{-Mt}|x^1(t) - x^2(t)|^2)$$

$$= e^{-Mt}(2(x^1(t) - x^2(t))(a(t, x^1(t) - a(t, x^2(t)) + \sum_{j=1}^{d_1} |b^j(t, x^1(t)) - b^j(t, x^2(t))|^2)\, dt$$

$$-e^{-Mt}M|x^1(t) - x^2(t)|^2\, dt$$

$$+\sum_{j=1}^{d_1} e^{-Mt}2(x^1(t) - x^2(t))(b^j(t, x^1(t)) - b^j(t, x^2(t)))\, dw_t^j$$

Thus we get,

$$0 \leq e^{-Mt}|x^1(t) - x^2(t)|^2$$

$$= \int_0^t e^{-Ms}(2(x^1(s)-x^2(s))(a(s, x^1(s)-a(s, x^2(s))+\sum_{j=1}^{d_1} |b^j(s, x^1(s))-b^j(s, x^2(s))|^2$$

$$-M|x^1(s) - x^2(s)|^2)ds + m_t' \leq m_t' \text{ (a.s.)},$$

where

$$m_t' = \int_0^t \sum_{j=1}^{d_1} e^{-Ms}2(x^1(s) - x^2(s))(b^j(s, x^1(s)) - b^j(s, x^2(s)))dw_s^j$$

is a non-negative local martingale, starting from 0. Hence, $x^1(t) = x^2(t)$ almost surely.

To prove the existence we will use the following well-known mini-max theorem.

□

7.3 Mini-Max Theorems

Let **V** and **U** be convex subsets of some metric vector spaces, such that they are compact metric spaces with respect to some metrics. In the application **V** and **U** will be closed balls in some separable Hilbert spaces considered in the weak topologies. Notice that bounded closed balls in Hilbert spaces are compact sets in the weak topology and that the weak topology restricted onto any ball is metriczable. Then the mini-max theorem of Fan [6] reads as follows:

Theorem 7.3.1. *Let $L : \mathbf{V} \times \mathbf{U} \longrightarrow \mathbb{R}$ be a function which satisfies the following conditions:*

(i) *L is convex and lower semi-continuous in $v \in \mathbf{V}$ for each $u \in \mathbf{U}$.*

(ii) *L is concave and upper semi-continuous in $u \in \mathbf{U}$ for each $v \in \mathbf{V}$.*

Then

$$\min_{v\in\mathbf{V}} \max_{u\in\mathbf{U}} L(u,v) = \max_{u\in\mathbf{U}} \min_{v\in\mathbf{V}} L(v,u).$$

For the proof we refer to [17].

We will develop a mini-max theorem which leads to our proof. It is a generalization of the above well-known mini-max theorem.

Theorem 7.3.2 (Mini-Max Theorem). *Let $L : \mathbf{V} \times \mathbf{U} \to \mathbb{R}$ satisfy the following conditions:*

(i) *L is lower semi-continuous and convex in $v \in \mathbf{V}$ for each $u \in \mathbf{U}$.*

(ii) *L is upper semi-continuous in $u \in \mathbf{U}$ for each $v \in \mathbf{V}$, and there is a countable dense subset $M = \{u_i : i = 1, 2, \cdots, n\}$ of $\mathbf{U}$ such that for every integer $n \geq 1$,*

$$\sum_{i=1}^{n} p_i L(v, u_i) \leq L(v, \sum_{i=1}^{n} p_i u_i)$$

for all $u_1, \cdots, u_n \in M$ and $p \in \mathbf{P}_n := \{(p_1, p_2, \cdots, p_n) \in \mathbb{R}^n : p_i \geq 0, \sum_i^n p_i = 1\}$

Then

$$\min_{v\in\mathbf{V}} \max_{u\in\overline{M}} L(u,v) \leq \max_{u\in\mathbf{U}} \min_{v\in\mathbf{V}} L(v,u),$$

where $\overline{M}$ is the closure of M.

Proof. Let us define

$$L_n(v,p) := \sum_{i=1}^{n} p_i L(v, u_i),$$

for $v \in \mathbf{V}$ and $p \in \mathbf{P}_n$. Then L_n satisfies the condition of Theorem 7.3.1. Hence for some $v^{(n)} \in \mathbf{V}$ and $p^{(n)} \in \mathbf{P}_n$ we have

$$L_n(v^{(n)}, p^{(n)}) = \min_{v\in\mathbf{V}} \max_{p\in\mathbf{P}_n} L_n(v,p) = \max_{p\in\mathbf{P}_n} \min_{v\in\mathbf{V}} L_n(v,p)$$

$$\leq \max_{p\in\mathbf{P}_n} \min_{v\in\mathbf{V}} L(v, \sum_{i=1}^{n} p_i u_i) \leq \max_{u\in\mathbf{U}} \min_{v\in\mathbf{V}} L(v,u) =: \alpha < \infty. \tag{7.3.2}$$

The sequence $\{v^{(n)}\}$ contains a subsequence, denoted also by $\{v^{(n)}\}$ such that $v^{(n)}$ converges to some element $\bar{v} \in \mathbf{V}$. By (7.3.2) and the definition of L_n,

$$\alpha \geq L_n(v^{(n)}, p^{(n)}) = \max_{p\in\mathbf{P}_n} L_n(v_n, p) = \max_{1\leq i\leq n} L(v^{(n)}, u_i).$$

Hence

$$\alpha \geq L(v^{(n)}, u_i),$$

for each $i \leq n$. Consequently, by the lower semi-continuity of L

$$L(\bar{v}, u^i) \leq \liminf_{n\to\infty} L(v^n, u^i) \leq \alpha,$$

for all $i \geq 1$. Therefore

$$\min_{v\in\mathbf{V}} \max_{u\in\overline{M}} L(v,u) \leq \max_{u\in\overline{M}} L(\bar{v}, u) \leq \alpha = \max_{u\in\mathbf{U}} \min_{v\in\mathbf{V}} L(v,u).$$

□

To apply the above theorem we need to introduce the following objects:

Let $\mathbf{V}$ denote the set of $\mathcal{F}_t$-adapted processes $v = (\alpha, \beta)$ on the interval $[0,T]$ such that α is $\mathbb{R}^d$-valued, β is $\mathbb{R}^{d\times d_1}$-valued and

$$|v|^2 := E\int_0^T |\alpha(t)|^2 dt + E\int_0^T \sum_{j=1}^{d_1} |\beta^j(t)|^2 dt \leq K^2.$$

Let $\mathbf{Y}$ denote the set of $\mathbb{R}^d$-valued $\mathcal{F}_t$-adapted processes $y = \{y(t) : t \in [0,T]\}$ such that

$$|y|^2 = E\int_0^T |y(t)|^2 dt \leq 4TK^2.$$

Define also the following functionals:

$$L(v,u) := L((\alpha,\beta),(y,\gamma,\delta)) := E\int_0^T e^{-Mt}\{2(x(t)-y(t))(\alpha(t)-\gamma(t))$$

$$+\sum_{j=1}^{d_1} |\beta^j(t) - \delta^j(t)|^2 - M(t)|x(t)-y(t)|^2\}dt \tag{7.3.3}$$

for $(\alpha,\beta) \in \mathbf{V}$ and $(y,\gamma,\delta) \in \mathbf{Y}\times\mathbf{V} =: \mathbf{U}$, where

$$x(t) := x_{\alpha\beta}(t) := \int_0^t \alpha(s)ds + \sum_{j=1}^{d_1} \beta^j(s)dw_s^j.$$

Let $u(y) := (y, a(y), b(y))$ for $y \in \mathbf{Y}$, where $a(y), b(y)$ are the processes

$$\{a(t,y(t)) : t \in [0,T]\},\ \{b(t,y(t)) : t \in [0,T]\}.$$

Define

$$F(v,y) = F(\alpha,\beta,y) := L(v,u(y)) = E\int_0^T e^{-Mt}\{2(x(t)-y(t))(\alpha(t)-a(y(t)))$$

$$+\sum_{j=1}^{d_1} |\beta^j(t) - b^j(y(t))|^2 - M(t)|x(t)-y(t)|^2\}dt. \tag{7.3.4}$$

Notice that

$$\sup_{y \in Y} F(\alpha, \beta, y) \geq 0,$$

for every $(\alpha, \beta) \in \mathbf{V}$. Indeed if $(\alpha, \beta) \in \mathbf{V}$, then it is easy to see that

$$E|x(t)|^2 < \infty, \ \forall \, t \leq T,$$

and

$$\begin{aligned} E|x(t)|^2 &\leq E\Big(\int_0^t 2x(s)\alpha(s)ds + \sum_{j=1}^{d_1} |\beta^j(s)|^2 ds\Big) \\ &\leq \frac{1}{2} E \int_0^t |x(s)|^2 ds + 2E \int_0^t |\alpha(s)|^2 ds + E \int_0^t \sum_{j=1}^{d_1} |\beta^j(s)|^2 ds \end{aligned}$$

which gives

$$E \int_0^T |x(t)|^2 dt \leq 4TE \int_0^T |\alpha(t)|^2 dt + 2T \int_0^T \sum_{j=1}^{d_1} E|\beta^j(t)|^2 dt$$

$$\leq 4TK^2.$$

Consequently, for $(\alpha, \beta) \in \mathbf{V}$, we can take $y := x_{\alpha\beta} \in \mathbf{Y}$, which gives

$$F(v, y) = 0.$$

The method to prove the existence consists in characterizing the solutions of equation (7.2.1) in terms of extremals of the functional F. Moreover, the following theorem holds,

Theorem 7.3.3. *Assume Assumption 7.2.1, 7.2.2 and 7.2.3 hold. Then the following assertions hold:*

(i) *If $x = \{x(t) : t \in [0, T]\}$ is a solution to SDE (7.2.1), then*

$$\sup_{y \in \mathbf{Y}} F(\bar{\alpha}, \bar{\beta}, y) = 0$$

for $(\bar{\alpha}, \bar{\beta}) \in \mathbf{V}$, where $\bar{\alpha}(t) := a(t, x(t))$, $\bar{\beta}(t) = b(t, x(t))$.

(ii) *If for some $(\bar{\alpha}, \bar{\beta}) \in \mathbf{V}$,*

$$\sup_{y \in \mathbf{Y}} F(\bar{\alpha}, \bar{\beta}, y) = 0, \tag{7.3.5}$$

then

$$x(t) = \int_0^t \bar{\alpha}(s) ds + \int_0^t \sum_{j=1}^{d_1} \bar{\beta}^j(s) dw_s^j, \ t \in [0, T]$$

is a solution to SDE (7.2.1) on $[0, T]$.

Proof. Assume that $x = x(\bar{\alpha}, \bar{\beta})$ is a solution to equation (7.2.1). Then by monotonicity condition 7.2.1, we have

$$F(\bar{\alpha}, \bar{\beta}, y) = E\int_0^T e^{-Mt}\{2(x(t) - y(t))(\bar{\alpha}(t) - a(y(t)))$$

$$+\sum_{j=1}^{d_1} |\bar{\beta}^j(t) - b^j(y(t))|^2 - M(t)|x(t) - y(t)|^2\}dt \leq 0,$$

thus $\sup_{y\in\mathbf{Y}} F(\bar{\alpha}, \bar{\beta}, y) = 0$.

Now we prove part (ii). Let (7.3.5) hold for some $(\bar{\alpha}, \bar{\beta}, y) \in \mathbf{V}$, we have that

$$F(\bar{\alpha}, \bar{\beta}, y) \leq 0,$$

for any $y \in \mathbf{Y}$. Hence it is sufficient for us to show that if $F \leq 0$ for some $\bar{\alpha}, \bar{\beta}, t, \omega$, then $\bar{\alpha}_t = a(t, x)$, and $\bar{\beta}_t = b(t, x)$. For these $\bar{\alpha}, \bar{\beta}, x$ and any $y \in \mathbf{Y}$, we have

$$E\int_0^T e^{-Mt}\{2(x(t) - y(t))(\bar{\alpha}(t) - a(t, y(t)))$$

$$+\sum_{j=1}^{d_1} |\bar{\beta}^j(t) - b^j(t, y(t))|^2 - M(t)|x(t) - y(t)|^2\}dt \leq 0.$$

In particular, taking $y(t) := x(t)$, this gives $\bar{\beta}(t) = b(t, x(t))$ for $dt \times P-$almost every $(t, \omega) \in [0, T] \times \Omega$. Now, let $y = x - \epsilon z$, where $\epsilon > 0$ and small enough, for $z \in \mathbf{Y}$. We get

$$2\epsilon E\int_0^T z(\bar{\alpha}(t) - a(t, x(t) - \epsilon z(t))) - M(t)\epsilon|z|^2 dt \leq 0,$$

then divided by 2ϵ and let $\epsilon \to 0$. By the continuity of $a(t, y)$ in y, we get

$$E\int_0^T z(\bar{\alpha}(t) - a(t, x(t))\; dt \leq 0.$$

This is true for all $z \in \mathbf{Y}$. Therefore $\bar{\alpha}(t) = a(t, x(t))\; dt \times P-$a.e., which, together with $\bar{\beta}(t) = b(t, x(t))$, imply that

$$x(t) = \int_0^t \bar{\alpha}(s)ds + \int_0^t \sum_{j=1}^{d_1} \bar{\beta}^j(s)dw_s^j$$

is a solution to SDE (7.2.1). □

7.4 Proof Theorem 7.2.1

It is also important to notice the following properties of the functional $L = L(v, u)$, for $v \in \mathbf{V}$, $u \in \mathbf{U}$, defined by (7.3.3).

We equip $\mathbf{V}$ with the weak topology of the Hilbert space defined by the inner product

$$(v_1, v_2) = E \int_0^T (\alpha_1(t)\alpha_2(t) + \beta_1(t)\beta_2^{\mathrm{T}}(t))dt$$

for $v_1 = (\alpha_1, \beta_1)$, $v_2 = (\alpha_2, \beta_2)$. We consider $\mathbf{U}$ with the product topology, with the topology introduced by the norm $|y| = (E \int_0^T |y(s)|^2 ds)^{1/2}$ on $\mathbf{Y}$ and with the weak topology on $\mathbf{V}$.

Definition 7.4.1. We say that a sequence $u^n = (y^n, \gamma^n, \delta^n)$ converges to $u = (y, \gamma, \delta)$ in $\mathbf{U}$ if $y^n \to y$ strongly in $\mathbf{Y}$, $\gamma^n \rightharpoonup \gamma$ and $\delta^n \rightharpoonup \delta$ weakly in $\mathbf{V}$.

Before proving the theorem, we will state a theorem which introduce the idea that will lead to the proof of Theorem 7.2.1. The following statements will be crucial for applying mini-max theorem.

Theorem 7.4.1. *Under Assumptions 7.2.1, 7.2.2 and 7.2.3, the following properties hold:*

(1) *$L(v, u)$ is convex and lower semi-continuous in $v = (\alpha, \beta) \in \mathbf{V}$ for fixed $u \in \mathbf{U}$;*

(2) *$L(v, u)$ is continuous in $u \in \mathbf{U}$ for fixed $v \in \mathbf{V}$;*

(3) *For $\overline{M} := \{(y, a(y), b(y)) : y \in \mathbf{Y}\} \subset \mathbf{U}$, it has a dense subset $M := \{u_i : i = 1, 2, \cdots, n\}$*

(4) *For every integer $n \geq 1$,*

$$L(v, \sum_{i=1}^{n} p_i u_i) \geq \sum_{i=1}^{n} L(v, u_i)$$

for every $u_1, \cdots, u_n \in \overline{M}$ and for all $p \in \mathbf{P}_n$, where $\mathbf{P}_n = \{(p_1, p_2, \cdots, p_n) \in \mathbb{R}^n : p_i \geq 0, \sum_i^n p_i = 1\}$.

Proof. **(1)** Recall that $L(v, u)$ is defined by

$$L(v, u) = E \int_0^T e^{-Mt}\{2(x(t) - y(t))(\alpha(t) - \gamma(t)) + \sum_{j=1}^{d_1} |\beta^j(t) - \delta^j(t)|^2$$

$$-M(t)|x(t) - y(t)|^2\}dt,$$

where $v = (\alpha, \beta) \in \mathbf{V}$, $u = (y, \gamma, \delta) \in \mathbf{Y} \times \mathbf{V} = \mathbf{U}$. It follows from Itô's formula that for fixed $u \in \mathbf{U}$,

$$L(v, u) = e^{-Mt} E|x(T)|^2 + H_u(\alpha, \beta),$$

where

$$H_u(\alpha, \beta) = E \int_0^T e^{-Mt} \{-2y(t)\alpha(t) - 2x(t)\gamma(t) + 2y(t)\gamma(t) +$$

$$\sum_{j=1}^{d_1} |\delta^j(t)|^2 - 2\sum_{j=1}^{d_1} \beta^j(t)\delta^j(t) - 2M(t)|y(t)|^2 + 2M(t)x(t)y(t)\} dt.$$

From the expression above, the function $L(v, u)$ is convex in $v = (\alpha, \beta) \in \mathbf{V}$, and is lower semi-continuous of $(\alpha, \beta) \in \mathbf{V}$ in the weak topology.

(2) Let $u^n \to u$ in $\mathbf{U}$. By definition (7.4.1), that is, $y^n \to y$ strongly in $\mathbf{Y}$, $\gamma^n \rightharpoonup \gamma$, and $\delta^n \rightharpoonup \delta$ weakly in $\mathbf{V}$. Then,

$$|L(v, u^n) - L(v, u)| = |-I_1 - I_2 + I_3 - I_4 - I_5 + I_6| \leq |I_1| + |I_2| + |I_3| + \cdots + |I_6| \tag{7.4.6}$$

it goes to 0 as $n \to \infty$, with

$$\begin{aligned}
I_1 &:= E \int_0^T e^{-Mt} 2(y^n(t) - y(t))(\alpha(t) - \gamma^n(t)) dt, \\
I_2 &:= E \int_0^T 2e^{-Mt}(x(t) - y(t))(\gamma^n(t) - \gamma(t)) dt, \\
I_3 &:= E \int_0^T e^{-Mt} \sum_j (|\delta^{n,j}(t)|^2 - |\delta^j(t)|^2) dt, \\
I_4 &:= E \int_0^T 2e^{-Mt} \sum_j \beta^j(t)(\delta^{n,j}(t) - \delta^j(t)) dt, \\
I_5 &:= E \int_0^T 2e^{-Mt} M(t)(|y^n(t)|^2 - |y(t)|^2) dt, \\
I_6 &:= E \int_0^T 2e^{-Mt} M(t)x(t)(y^n(t) - y(t)) dt.
\end{aligned}$$

Indeed for $n \to \infty$, the first term clearly goes to 0 by Hölder inequality;

$$\begin{aligned}
I_2 &\leq 8TK^2 E \int_0^T |\gamma^n(t) - \gamma(t))| dt \to 0; \\
I_3 &\leq KE \int_0^T \sum_j (\delta^{n,j}(t) - \delta^j(t))(\delta^{n,j}(t) + \delta^j(t)) dt \\
&\leq K(E \int_0^T |\delta^{n,j}(t) - \delta^j(t)|^2 dt)^{1/2} (E \int_0^T |\delta^{n,j}(t) + \delta^j(t)|^2 dt)^{1/2} \\
&\to 0,
\end{aligned}$$

since δ^n converges to δ weakly.

Similarly, we can get

$$I_4 \to 0; I_5 \to 0; I_6 \to 0.$$

(3) First, we can easily see that there exists countable everywhere dense subset $\{y_i : i = 1, 2, \cdots,\}$ of $\mathbf{Y}$. Then for all $y \in \mathbf{Y}$, there exists subsequences denoted by y_i converging to y. Since $y_i \to y$ strongly in $\mathbf{Y}$, we have $a(y_i) \rightharpoonup a(y)$ and $b(y_i) \rightharpoonup b(y)$ weakly in $\mathbf{V}$, i.e., there exists subset $M := \{u_i : i = 1, 2, \cdots., n\}$, where $u_i = \{y_i, a(y_i), b(y_i) : y \in \mathbf{Y}\}$

(4) It is sufficient to prove

$$\sum_{i=1}^{n} p_i E \int_0^T e^{-Mt}[2y_i(t)a(y_i(t)) + \sum_{j=1}^{d_1} |b^j(y_i(t))|^2 - M(t)|y_i(t)|^2]dt$$

$$\leq E \int_0^T e^{-Mt}[2\bar{y}(t)\bar{a}(y(t)) + \sum_{j=1}^{d_1} |\bar{b}^j(y(t))|^2 - M(t)|\bar{y}(t)|^2]dt,$$

where $\bar{y} := \sum_{i=1}^{n} p_i y_i$, $\bar{a} := \sum_{i=1}^{n} p_i a(y_i)$, $\bar{b} := \sum_{i=1}^{n} p_i b(y_i)$.

Let us define

$$f_u(\alpha, \beta) := \sup_{y \in \mathbf{Y}} L(v, u(y)) - E \int_0^T e^{-Mt}[2x(t)a(x(t))$$

$$+ \sum_j |\beta^j(t)|^2 - M(t)|x(t)|^2]dt.$$

Notice that, $\sup_{y \in \mathbf{Y}} L(v, u(y)) \geq 0$ and

$$\sup_{y \in \mathbf{Y}} L(a(t,x), b(t,x), u(y)) = 0,$$

where $(a(t,x), b(t,x)) \in \mathbf{V}$, because of the monotonicity condition of (a, b).

Then

$$f_u(\alpha, \beta) \geq -E \int_0^T e^{-Mt}[2x(t)a(x(t)) + \sum_{j=1}^{d_1} |\beta^j(t)|^2 - M(t)|x(t)|^2]dt \ ,$$

and

$$f_u(a(x), b(x)) = -E \int_0^T e^{-Mt}[2x(t)a(x(t))$$

$$+ \sum_{j=1}^{d_1} |\beta^j(t)|^2 - M(t)|x(t)|^2]dt.$$

Furthermore, f_u is a convex function. Hence we get

$$f_u(\bar{a}(y(t)), \bar{b}(y(t)))$$

$$\leq \sum_{i=1}^{n} p_i f_u(a(y_i(t)), b(y_i(t))),$$

where $\bar{y} := \sum_{i=1}^{n} p_i y_i$, $\bar{a} := \sum_{i=1}^{n} p_i a(y_i)$, $\bar{b} := \sum_{i=1}^{n} p_i b(y_i)$.

From above we can obtain that

$$-E\int_0^T e^{-Mt}[2\bar{y}(t)\bar{a}(y(t)) + \sum_{j=1}^{d_1} |\bar{b}^j(y(t))|^2 - 2M(t)|\bar{y}(t)|^2]dt$$

$$\leq f_u(\bar{a}(y(t)), \bar{b}(y(t)))$$

and we also have

$$\sum_{i=1}^{n} p_i f_u(a(y_i(t)), b(y_i(t)))$$

$$= -\sum_{i=1}^{n} p_i E\int_0^T e^{-Mt}[2y_i(t)a(y_i(t)) + \sum_{j=1}^{d_1} |b^j(y_i(t))|^2 - 2M(t)|y_i(t)|^2]dt.$$

The theorem is proved.

□

Now we are in the position to show that the existence of a solution is a consequence of Theorem 7.3.2.

Proof of the existence. By virtue of the previous theorem we can apply Theorem 7.3.2 to $L, F, \mathbf{U}, \mathbf{V}, M$. Recall that $L = L(v, u) := L((\alpha, \beta), (y, \gamma, \delta))$ is defined in (7.3.3) and $F(v, y)$ is given in (7.3.4); $\mathbf{V}$ is defined as $\mathcal{F}_t-$adapted processes $v = (\alpha, \beta)$ on the interval $[0, T]$; $\mathbf{U}$ is defined by $\mathbf{Y} \times \mathbf{V}$, where $\mathbf{Y}$ is the set of $\mathbb{R}^d-$adapted processes $y = \{y(t) : t \in [0, T]\}$.

Consequently, for

$$F(v, y) = L(v, u(y))$$

we have

$$\begin{aligned} 0 &\leq \min_{v \in \mathbf{V}} \max_{y \in \mathbf{Y}} F(v, y) = \min_{v \in \mathbf{V}} \max_{u \in \overline{M}} L(v, u(y)) \\ &\leq \max_{u \in \mathbf{U}} \min_{v \in \mathbf{V}} L(v, u). \end{aligned}$$

Notice that for each $u = (y, \gamma, \delta) \in \mathbf{U}$ we can choose

$$\bar{v} := (\gamma, \delta) \in \mathbf{V},$$

which gives $L(\bar{v}, u) = 0$. Consequently,

$$\min_{v \in \mathbf{V}} L(v, u) \leq 0,$$

for every $u \in \mathbf{U}$. Hence

$$\min_{v \in \mathbf{V}} \max_{y \in \mathbf{Y}} F(v, y) = 0$$

which proves the existence of a solution to SDE (7.2.1) by virtue of Theorem 7.3.3. □

Appendix A

The follow lemma is from the version of Lemma 3.5 from [14].

Lemma A.0.2. *If $Z_n = \{Z_n(t) : t \in [0, T]\}$ is a sequence of cadlag stochastic processes. For a fixed $\epsilon > 0$ define*

$$\tau^n = \inf\{t \in [0,T] : |Z_n(t)| \geq \epsilon.\}$$

Then the following statements hold:

1. *If $Z_{n\epsilon} = \{Z_n(t \wedge \tau^n) : t \in [0,T]\}$ converges in probability to 0, uniformly in $t \in [0,T]$, then Z_n converges to 0 in probability uniformly in $t \in [0,T]$.*

2. *If almost surely $Z_{n\epsilon}$ converges to 0, uniformly in $t \in [0,T]$, then almost surely Z_n converges to 0, uniformly in $t \in [0,T]$.*

3. *If for some sequence $0 < \alpha(n) \to 0$, there is a finite random variable η, such that almost surely*

$$\sup_{t\leq T} |Z_{n\epsilon}(t)| \leq \eta\alpha(n), \forall n,$$

then there is a finite random variable ξ such that

$$\sup_{t\leq T} |Z_n(t)| \leq \xi\alpha(n).$$

For the proof we refer to [14].

Now we give the proof of of Lemma 5.1.5:

Proof of Lemma 5.1.5. 1. Easy to see that R is finite, convex and continuous.

2. If we take the limit point $a(x)$ of any sequence $a(x'_n)$ in (5.1.3), when $x'_n \to x, x'_n \in D(a)$, then we get

$$\sup_{x'}(a(x') - y)(x - x') \geq 0.$$

Hence, $R(a,x,y) \geq xy$ is obtained.

The equality is a consequence of the monotonicity of a, that is,

$$R(a,x,a(x)) = \sup_{x'}\{(a(x') - a(x))(x - x') + xa(x) : x' \in D\} = xa(x).$$

3. From above $R \geq xy$ on $\bar{D} \times \mathbb{R}^d$, we see that the graph of R lies above the tangent plane of xy for any $x' \in \bar{D}$, i.e., $R(a,x,y) \geq (x - x')a(x') + x'y$ for any $x' \in D$.

4. From the convexity of $R(a,x,y)$, by virtue of the second result,

$$\sum_{i=1}^{n} \alpha_i x_i y_i = \sum_{i}^{n} \alpha_i R(a, x_i, y_i) \geq R(a, \sum_{i}^{n} \alpha_i x_i, \sum_{i}^{n} \alpha_i y_i)$$

$$\geq \sum_{i=1}^{n} \alpha_i x_i \sum_{i=1}^{n} \alpha_i y_i.$$

5. The monotonicity is obtained from the result above if one takes $n = 2$, $\alpha_1 = \alpha_2 = 1/2$, and make simple transformations. The closedness of $Z(a)$ is from the continuous of function R.

6. Using the rule for passaging to the limit under limsup sign.

7. We denote the convex hull mentioned by $P(a,x)$. We want to show that $P(a,x) = Z(a,x)$. The set $P(a,x)$ is closed, as the convex hull of a closed set. Since $R(a,x,y) - xy$ for fixed x is a nonnegative convex function of y, $Z(a,x)$ is a convex set. From the boundedness a, there exist bounded subsequence a_n , and $(x_n, y_n) \to (x,y)$ for any $x_n, y_n \in \bar{D}$, such that $a_n \to a$ on D. Then $\liminf_n R(a_n, x_n, y_n) \geq R(a,x,y)$. So any partial limit of $a(x_n)$ as $x_n \to x$ lies in $Z(a,x)$. hence $P(a,x) \subset Z(a,x)$.

   Now we want to show $P(a,x) = Z(a,x)$. Assume there exists $y \in Z(a,x) \setminus P(a,x)$, then one can find $\phi \in \mathbb{R}^d$ and numbers $a < b$ such that $\phi y = b$, $\limsup_{x'} \phi a(x') \leq a$ as x' goes to x. Moreover, $R(a,x,y) = xy$ and $(a(x') - y)(x - x') \leq 0$ for all $x' \in D$. Choosing $x' \to x$, so that $x - x'$ and ϕ have the same direction, we conclude that $0 \leq \limsup(a(x') - y)\phi \leq a - b$. which is impossible.

8. We may assume that $y = 0$, since we can take $a - y$ instead of a. From the fact in result 2, we know that $R(a,x,0) \geq 0$. Hence it suffices for us to prove that

$$\min_{x \in \bar{D}} R(a,x,0) \leq 0. \tag{A.0.1}$$

From (5.1.3), it is easy to see that

$$\min_{x\in\bar{D}} R(a,x,0) = \min_{x\in\bar{D}} \sup_{x'\in D} (xa(x') - x'a(x')).$$

Let L be the convex hull of $\{a(x'), -x'a(x') : x' \in D\} := \{(p,q) \in L\}$, then

$$\sup_{x'\in D} (xa(x') - x'a(x')) = \sup_{(p,q)\in L} \{xp+q\}.$$

By using the Mini-Max theorem, we can prove (A.0.1) by showing that

$$\sup_{(p,q)\in L} \min_{x\in\bar{D}} \{xp+q\} \le 0.$$

Any point $(p,q) \in L$ admits a form $(\sum_{i=1}^n \alpha_i y_i, \sum_{i=1}^n \alpha_i x_i y_i)$, where $\alpha_i \ge 0, \sum \alpha_i = 1, y_i = a(x_i)$. We apply result 4 to get

$$\min_{x\in\bar{D}} x\sum_{i=1}^n \alpha_i y_i - \sum_{i=1}^n \alpha_i x_i y_i \le 0.$$

Consequently, inequality (A.0.1) is proved.

□

Bibliography

[1] Alyushina, L.A., (1987). Euler ploygonal lines for Ito's equations with monotone coeffcients. *Theory Probab. Appl.* **32**, 340-214.

[2] Aubin, J.P. and Cellina, A. , (1984). Differential inclusion. *Springer, Berlin.*

[3] A. Bensoussan, A. Rascanu, (1997). Stochastic variational inequalities in infinte dimensional space.*Numer. Funct. Anal. and Optimiz.*, **18**, 19-54.

[4] Brezis, H. ,(1973). Opérateurs maximaux monotones et semi-groups de contractions dans les espaces de Hilbert. *North Holland Math. Studies,* 5.

[5] Desmond J. Higham, Xuerong Mao, and Andrew M. Stuart, (2002). Strong convergence of Euler-Type methods for nonlinear stochastic differential equations. *SIAM J. Numer Anal.* **40**, No.3, 1041-1063.

[6] Fan, K., Minimax theorems. *Proc. of Nat. Acad, Sci.*, **39**, 42-47.

[7] Francis Hirsch, Gilleslacombe, (1999). Elements of Functional Analysis. *Springer*, 115.

[8] Frederic Bernardin, (2003). Multivalued sochastic differential equations: convergence of a numerical scheme. *Set-Valued Analysis*, **11**, 393-415.

[9] Gyöngy, I. and Krylov, N.V., (1996). Existence of strong solution for Ito's stochastic equations via approximations. *Probab. Theory Relate. Fields* **105**, 143-158.

[10] Gyöngy, I., (1998). A note of Euler's Approximations. *Protential Analysis*, **8**, 205-216.

[11] Gyöngy, I., Martinez, A., (2005). Solutions of stochastic partial differential equations as extremals of convex functionals. *Acta Math. Hungar.*, **109 (1-2)**, 127-145.

[12] Gyöngy, I., Millet, A., (2005). On discretization schemes for stochastic evolution equations. *Potential Analysis*, **23**, 99-134.

[13] Gyöngy, I., Krylov; N.V., (1980). On stochastic equations with respect to semimartingales. *Stochastics*, **I**, 1-21.

[14] Gyöngy, I., Shmalthov, A., Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations, *to appear in Appl. Math. and Optimization.*

[15] Gyöngy, I., (2006) A brief introduction to the theory of random processes. *Postgraduate course on probability and stochastic analysis.*

[16] Y, Hu, (1996). Semi-implicit Euler-Maruyama scheme for stiff stochastic equations. *Stochasic Analysis and Related Topics V: The Silvri Workshop, Propr. Probab. 38, H. Koerezlioglu, ed., Birkhauser, Boston*, 183-192.

[17] J. Szép, F. Forgó, (1985). Introduction to the theory of games. *D. Reidel Publishing Company, Boston* .

[18] Krylov, N.V., Rosovski, B.L., (1979). Itô equations in Banach space. *Itogi nauki, Teor rerojatn*, (Viniti an U.S.S.R), Moscow, 72-147.

[19] Krylov, N.V., (1984). Ėxtremal properties of solutions of stochastic equations. *Theory Probab. Appl.* **5**, 205-346.

[20] Krylov, N.V., (1990). A simple proof of the existence of a solution of Ito's equation with monotone coeffcients. *Theory Probab. Appl.* **3**, 583-587.

[21] Krylov, N.V., (1982). Some properties of monotonic mappings, *Lit. Matem. Sb.* **22** , 2, 80-87.

[22] Krylov, N.V. (1980). Controlled Diffusion Processes , *Springer-Verlag* .

[23] Krylov, N.V., (2002).Introduction to the Theory of Random Processes, *American Mathematical Society.*

[24] Krylov, N.V., (1995). Introduction to the Theory of Diffusion Processes, *American Mathematical Society.*

[25] Krée, P. (1982). Diffusion equation for multivalued stochastic differential equations. *J. Funct. Anul.* **49**, 73-90.

[26] Peter Kloeden, Eckhard Platen, (1992). Numerical Solution of Stochastic Differential Equations, *Springer* .

[27] Dominique Lepingle, Nguyen Thi Thao, . (2004) Approximating and simulation multivalued stochastic differential equations. *Montecarlo methods and Applications* , **10-2**, 129-154(24).

[28] Maruyama, G., (1955). Continuous Markov processes and stochastic equations. *Rend Circ. Mat. Palermo* **4**, 48-90.

[29] Jerzy Motyl, (1995). On the solution of stochastic differential inclusion. *Journal of Mathematical analysis and applications* , **192**, 117-132.

[30] B. Øksendal, (1995). Stochastic differential equations. *Springer-Verlag, New York* .

[31] Roger Pettersson (1994). Yosida approximations for multivalued stochastic differential equations. *Stochastics and Stochastics Reports* **52**, 107-120.

[32] R.R.Phelps (1993). Lectures on maximal monotone operators. *Lectures given at Prague/Paseky Summer School, Crech Republic.*

[33] Eckhard, Platen (1999). An introduction to numerical methods for stochastic differential equations. *Acta Numerica*, 197-246.

[34] W.Rudin (1973). Functional Analysis. *McGraw-Hill, New York.*

[35] R.T. Rockafellar, (1969). Local boundedness of nolinear monotone operators, *Michigan Math. J.* **16**, 397-407.

[36] Schurz, H., (2002). General theorems for numerical approximation of stochastic processes on the Hilbert Space $H_2([0,T]\mu)$, *Report, Maths Dept, Southern Illinois University.*

[37] T. Mikosch, (1998). Elementary Stochastic Caculus. *World Scientific.*

[38] Y. S. Yun, I. Shigekawa, The existence of solutions for stochastic differential inclusion, *to appear in Far East J. Math. Sci.*

[39] Zeidler, E.(1990). Nonlinear functional analysis and its applications. A,B. Nonlinear monotone operators,I *Springer-Verlag, New York* .

[40] Zeidler, E. (1990). Nonlinear functional analysis and its applications. Nonlinear monotone operators, II *Springer-Verlag, New York* .