Approximations for Stochastic McKean-Vlasov Equations with Non-Lipschitz Coefficients by an Euler-Maruyama Scheme

Emilia Koskela

Master's Thesis

University of Jyväskylä
Department of Mathematics and Statistics
Spring 2023

Tiivistelmä:
Emilia Koskela, *Approximations for Stochastic McKean-Vlasov Equations with Non-Lipschitz Coefficients by an Euler-Maruyama Scheme*, matematiikan pro gradu -tutkielma, 41 s., Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, toukokuu 2023.

Tässä tutkielmassa käsittelemme stokastisia McKean-Vlasov -yhtälöitä. Nämä ovat stokastisia differentiaaliyhtälöitä, joiden kerroinfunktiot riippuvat myös ratkaisun jakaumasta. Tämä riippuvuus lisää yhtälöiden monimutkaisuutta, joten tutkielmassa tutkimme yhtälöitä diskreetin approksimaation kautta.

Keskitymme tässä tutkielmassa tarkastelemaan yksikäsitteisen vahvan ratkaisun olemassaoloa stokastisille McKean-Vlasov -yhtälöille diskreettiä ja rekursiivista Euler-Maruyama -approksimaatiota käyttäen, sekä tämän approksimaation suppenemisnopeutta. Pääasiallisena lähteenä tässä tutkielmassa käytämme Xiaojie Dingin ja Huijie Qiaon artikkelia *Euler-Maruyama Approximations for Stochastic McKean-Vlasov Equations with Non-Lipschitz Coefficients.*

Tutkielmassa esittelemme pohjateoriaa ja joitakin tuloksia stokastisten prosessien sekä stokastisten differentiaaliyhtälöiden ympärillä. Esitämme muun muassa määritelmän McKean-Vlasov -yhtälöiden vahvalle ja heikolle ratkaisulle sekä esittelemme martingaaliongelman. Käymme läpi myös joitakin hyödyllisiä epäyhtälöitä. Asetamme lisäksi oletukset, joiden puitteissa tutkielmassa työskentelemme: oletamme tässä tutkielmassa esimerkiksi, että McKean-Vlasov -yhtälöiden kerroinfunktiot täyttävät tietyt ei-Lipschitz -ehdot.

Yksi tutkielman päätuloksista on näyttää yksikäsitteisen vahvan ratkaisun olemassaolo. Toteutamme tämän kahdessa osassa: ensin näytämme, kuinka diskreetti Euler-Maruyama -approksimaatio voidaan määritellä rekursiivisesti. Tämän approksimaation avulla todistamme, että tarkastelemallemme yhtälölle on olemassa ratkaisu martingaaliongelmaan, ja täten saamme osoitetuksi myös heikon ratkaisun olemassaolon.

Tämän jälkeen näytämme Iton kaavaa hyödyntäen, että poluittainen yksikäsitteisyys pätee oletustemme ollessa voimassa. Näiden vaiheiden jälkeen osoitamme, että vahvan yksikäsitteisen ratkaisun olemassaolo voidaan näyttää. Tarkastelemme Iton kaavan avulla myös ratkaisun olemassaolon osoittamiseen käytetyn Euler-Maruyama -approksimaation suppenemisnopeutta.

Abstract:

In this thesis we study stochastic McKean-Vlasov equations. These are stochastic differential equations where the coefficients depend also on the distribution of the solution. This dependency adds to the complexity of the equation so in this thesis we will study these equations using a discrete approximation.

We focus on considering the existence of a unique strong solution to stochastic McKean-Vlasov equations using a discrete and recursive Euler-Maruyama approximation, as well as the convergence rate of the approximation. Our main source is the article *Euler-Maruyama Approximations for Stochastic McKean-Vlasov Equations with Non-Lipschitz Coefficients* written by Xiaojie Ding and Huijie Qiao, which we follow throughout this thesis.

In the thesis we recall some preliminary theory surrounding stochastic processes and stochastic differential equations and introduce some results. We give the definitions for weak and strong solutions for the McKean-Vlasov equation as well as the definition for the martingale problem. We also introduce some useful inequalities. We give the assumptions under which we work in this thesis, such as the assumption that the coefficients of the McKean-Vlasov equations satisfy some non-Lipschitz conditions.

One of the main results in this thesis is to show the existence of unique strong solutions. We approach this in two steps: first, we show the recursive construction of the Euler-Maruyama approximation. With this approximation we show that there exists a solution to the martingale problem and hence we get the existence of a weak solution.

Then, using Ito's formula we prove that pathwise uniqueness holds under our assumptions. After these two steps we show that the existence of a strong unique solution can be proven. We also investigate with the help of Ito's formula the convergence rate of the Euler-Maruyama approximation used to show the existence of the solution.

Contents

Introduction

Stochastic differential equations have been of interest amongst researchers thanks to their many applications, for example in physics and engineering. Thus the existence of their solutions as well as their uniqueness has also been investigated under different assumptions regarding the regularity of the coefficients. In order to utilize computers in the research numerical approximations have become of use. A useful and relatively simple tool is the Euler-Maruyama approximation which is a discrete, recursively defined method. [**30**]

In this thesis we consider stochastic McKean-Vlasov equations, also known as mean-field equations, and the existence of unique strong solutions using the Euler-Maruyama approximation. McKean-Vlasov equations are stochastic differential equations whose coefficients depend also on the distribution of the solution. In this thesis we consider the coefficients to be non-Lipschitz.

Throughout this thesis we follow the article [**10**] by Xiaojie Ding and Huijie Qiao as our main source. We provide the calculations and deduction from the article in more detail in this thesis. Some basic probability theory is assumed to be known, but some preliminary definitions and results are introduced in the first chapter. In section 1.1. we recall some basic probability theory, from defining the stochastic basis to considering some convergence results as well as probability measures on metric spaces.

In section 1.2. stochastic processes are introduced and we give a definition for the Brownian motion. We introduce the McKean-Vlasov stochastic differential equation and the definition for weak and strong solutions. We also consider the martingale problem and some notions of uniqueness, and give Ito's formula for the multidimensional case. Section 1.3. consists of some useful and fundamental inequalities such as the Burkholder-Davis-Gundy inequality and the Grönwall inequality. In section 1.4. we introduce some notation and in section 1.5. the assumptions under which we will work in this thesis are presented.

As one of our main goals of this thesis, in chapter 2 we focus on proving the existence and uniqueness of a strong solution for the stochastic McKean-Vlasov equation introduced in section 1.2. In section 2.1. we consider the Euler-Maruyama approximation to show the existence of a solution to the martingale problem, which then implies the existence of weak solutions. In section 2.2. we use Ito's formula to prove that pathwise uniqueness holds under our conditions. We then conclude that these together imply the existence of a unique strong solution. Section 2.3. consists of some technical results that are used in chapter 2.

Finally, in chapter 3 we consider the convergence rate of the Euler-Maruyama approximation used to show the existence results. We use Ito's formula to approximate the convergence rate.

CHAPTER 1

Preliminaries

1.1. Probability Theory

First we introduce some basic probability theory and notations. In this thesis some basic theory is assumed to be known.

Unless stated otherwise, we assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space, $\mathcal{F}$ is the σ-algebra and $\mathbb{P}$ the probability measure on $\mathcal{F}$. The probability space is introduced more in depth in [**13**] Chapter 2. Additionally, for $T > 0$ we assume a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ on $(\Omega, \mathcal{F})$ to form a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ ([**14**] Definition 2.1.8.).

In the following we state the conditions, so-called "usual conditions", which we assume to be satisfied throughout this thesis.

Definition 1.1 (The usual conditions, see [**14**] Definition 2.4.11.). The stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ satisfies the usual conditions if

(i) the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, i.e. for all $A \in \mathcal{F}$ with $\mathbb{P}(A) = 0$ it holds that $B \subset A$ implies $B \in \mathcal{F}$,

(ii) for all $A \in \mathcal{F}$ with $\mathbb{P}(A) = 0$ it holds that $A \in \mathcal{F}_t$ for $t \in [0, T]$

(iii) the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is right-continuous, i.e. for $t \in [0, T)$ we have

$$\mathcal{F}_t = \bigcap_{t<s<T} \mathcal{F}_s.$$

For $t = T$ we note that the relation $\mathcal{F}_T = \bigcap_{T<s<T} \mathcal{F}_s$ does not pose any condition on $\mathcal{F}_T$.

We denote by $\mathbb{E}$ the expectation with respect to the probability measure $\mathbb{P}$. To specify that we take the expectation with respect to some probability measure $\tilde{\mathbb{P}}$ we write $\mathbb{E}^{\tilde{\mathbb{P}}}$.

Definition 1.2 (see [**1**] Definitions 2.3.5, 2.3.7). We say that a measurable function f on a measure space $(\Omega, \mathcal{F}, \mu)$ is integrable (with respect to μ) provided that

$$\int_\Omega |f| d\mu < \infty.$$

For $0 < p < \infty$ we denote

$$L^p(\Omega, \mathcal{F}, \mu) := \{f : |f|^p \text{ is measurable and integrable with respect to } \mu\}.$$

The conditional expectation of a measurable and integrable function f given a sub-sigma-algebra $\mathcal{G} \subset \mathcal{F}$ is denoted by $\mathbb{E}[f|\mathcal{G}]$. The conditional expectation is introduced in [**23**] 2.b.- 2.b.0.

Here we give some notions regarding continuity:

Definition 1.3 (Uniform continuity, see [**11**] Definition 5.1.). Let $f : D \to \mathbb{R}$ be a function. Suppose that for all $\varepsilon > 0$ there exists a $\delta > 0$ such that for $x, y \in D$

$$|x - y| < \delta \qquad \text{implies} \qquad |f(x) - f(y)| < \varepsilon.$$

Then the function f is called uniformly continuous.

Proposition 1.4 (See [**25**] Theorem 3.1). *A composition of two continuous functions is continuous.*

In the following we will recall some results and definitions regarding convergence. First we introduce monotone and dominated convergence:

Theorem 1.5 (Monotone Convergence Theorem, see [**29**] Theorem 1.3.5 (i)). *Suppose $f^n \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ for $n \geq 1$ and let $f^n \uparrow f$ as $n \to \infty$ a.s. Then it holds that*

$$\mathbb{E}f = \lim_{n\to\infty} \mathbb{E}f^n.$$

Theorem 1.6 (Dominated Convergence Theorem, see [**1**] Corollary 2.3.12). *Assume that $(\Omega, \mathcal{F}, \mu)$ is a measure space and $f_1, f_2, ..., g : \Omega \to \mathbb{R}$ are measurable functions. Suppose that $|f_n| \leq g$ μ-a.e for all $n \geq 1$, $\int_\Omega g d\mu < \infty$ and $\lim_{n\to\infty} f_n = f$ μ-a.e. Then it holds that $f \in L^1(\Omega, \mathcal{F}, \mu)$ and*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

Next we define different forms of convergence for measurable functions:

Definition 1.7 (See [**23**] 1.a.). Assume $f, f_1, f_2, ...$ to be measurable functions defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

(i) We say that the sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ converges to f in L^1 provided that $\|X_n - X\|_1 = \mathbb{E}(|f_n - f|) \xrightarrow{n} 0$. We write $f_n \xrightarrow{L^1} f$.

(ii) We say that the sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ converges to f ($\mathbb{P}$-) almost surely if it holds that $\mathbb{P}(\omega \in \Omega : f_n(\omega) \xrightarrow{n} f(\omega)) = 1$. We write $f_n \xrightarrow{a.s.} f$.

(iii) We say that the sequence of functions $\{f_n\}_{n\in\mathbb{N}}$ converges to f in probability if for all $\varepsilon > 0$ we have $\mathbb{P}(|f_n - f| > \varepsilon) \xrightarrow{n} 0$. We write $f_n \xrightarrow{\mathbb{P}} f$.

Theorem 1.8 (See [**23**] 1.a.0.). *(i) Almost sure convergence implies convergence in probability.*

(ii) Convergence in L^1 implies convergence in probability.

In addition to having different types of convergence for measurable functions, we also consider convergence for probability measures in the weak sense:

Definition 1.9 (Weak convergence, see [**5**] Chapter 1 Section 1). Let S be a metric space and $\mathcal{S}$ be the Borel σ-algebra generated by S. Let $\{P_n\}_{n\in\mathbb{N}}, P$ be probability measures on $(S, \mathcal{S})$ such that for all continuous bounded real functions f on S it holds that

$$\lim_{n\to\infty} \int f(x)P_n(dx) = \int f(x)P(dx).$$

Then we say that the sequence of probability measures $\{P_n\}_{n\in\mathbb{N}}$ converges weakly to P. We denote this by $P_n \xrightarrow{w} P$.

Next we will consider uniform integrability:

Definition 1.10 (Uniform integrability, see [**7**] Definition 4.5.1). Assume a family of functions $\mathscr{A} \subset L^1(\Omega, \mathcal{F}, \mathbb{P})$. We say that $\mathscr{A}$ is *uniformly integrable* provided that

$$\lim_{c\to+\infty} \sup_{f\in\mathscr{A}} \int_{\{|f|>c\}} |f| d\mathbb{P} = 0.$$

Proposition 1.11 (see [**23**] 1.b.7). *Let* $1 < p < \infty$. *A family of measurable functions* $\{f_n\}_{n\in\mathbb{N}} \subset L^p(\Omega, \mathcal{F}, \mathbb{P})$ *is uniformly integrable provided that* $\sup_{n\in\mathbb{N}} \mathbb{E}|f_n|^p < \infty$.

The following theorem gives us the implication between convergence in probability and in L^1 in the other direction with the help of uniform integrability:

Theorem 1.12 (Lebesgue-Vitali theorem, see [**7**] Theorem 4.5.4.). *Let* f *be a measurable function and let* $\{f_n\}_n$ *be a sequence of integrable functions. Then the following assertions are equivalent:*

(i) the sequence $\{f_n\}_n$ *is uniformly integrable and converges to* f *in probability,*

(ii) the function f *is integrable and* $\{f_n\}_n$ *converges to* f *in* L^1.

As a corollary of Theorem 1.12 we get the following lemma:

Lemma 1.13. *Let* f *be a measurable function and let* $\{f_n\}_{n\in\mathbb{N}}$ *be a uniformly integrable sequence of measurable functions such that* $\{f_n\}_{n\in\mathbb{N}}$ *converges to* f *in probability. Then it holds that*

$$\lim_{n\to\infty} \mathbb{E}f_n = \mathbb{E}f.$$

PROOF. By Theorem 1.12 it holds that $\lim_{n\to\infty} \|f_n - f\|_{L^1} = 0$ so the statement follows from

$$0 \leq |\lim_{n\to\infty} \mathbb{E}(f_n) - \mathbb{E}(f)| = \lim_{n\to\infty} |\mathbb{E}(f_n - f)| \leq \lim_{n\to\infty} \mathbb{E}|f_n - f| = 0.$$

□

Tonelli's Theorem given below is a useful tool for changing the order of integration, for example when taking expectations of integrals.

THEOREM 1.14 (Tonelli's Theorem, see [**2**] Theorem 5.28). *Let $(X, \mathcal{F}, \mu)$ and $(Y, \mathcal{G}, \nu)$ be σ-finite measure spaces and let $f : X \times Y \to [0, \infty]$ be a $\mathcal{F} \otimes \mathcal{G}$-measurable function. Then the following holds:*

(i) $x \mapsto \int_Y f(x,y) d\nu(y)$ is an $\mathcal{F}$-measurable function on X,

(ii) $y \mapsto \int_X f(x,y) d\mu(x)$ is a $\mathcal{G}$-measurable function on Y and

(iii) $$\int_{X\times Y} f d(\mu \times \nu) = \int_X \int_Y f(x,y) d\nu(y) d\mu(x) = \int_Y \int_X f(x,y) d\mu(x) d\nu(y).$$

DEFINITION 1.15 (π-system, see [**6**]). A class $\mathcal{P}$ of subsets of Ω is called a π-system provided that for all $A, B \in \mathcal{P}$ it holds that $A \cap B \in \mathcal{P}$.

THEOREM 1.16 (See [**6**] Theorem 3.3). *Let $\mathcal{P}$ be a π-system and let μ and ν be probability measures on $\sigma(\mathcal{P})$. Assume that $\mu(A) = \nu(A)$ for all $A \in \mathcal{P}$. Then one has that $\mu(B) = \nu(B)$ for all $B \in \sigma(\mathcal{P})$.*

In the following we consider tightness for a family of probability measures:

DEFINITION 1.17 (see [**20**] Definition 4.6). Assume that (S, ρ) is a metric space and Π is a family of probability measures on $(S, \mathcal{B}(S))$, where $\mathcal{B}(S)$ denotes the Borel σ-algebra on S. Then Π is called

(i) *relatively compact* provided that every sequence of its elements contains a weakly convergent subsequence,

(ii) *tight*, if for all $\varepsilon > 0$ there exists a compact set $K \subset S$ satisfying $P(K) \geq 1 - \varepsilon$ for every $P \in \Pi$.

THEOREM 1.18 (see [**20**] Theorem 4.7). *Let S be a complete, separable metric space and let Π be a family of probability measures on S. Then Π is relatively compact if and only if it is tight.*

1.2. Stochastic Processes

In this section we introduce theory regarding stochastic processes, the Brownian motion and the McKean-Vlasov stochastic differential equation as well as the definitions of the solutions.

DEFINITION 1.19 (Stochastic Process, see [**14**] Definition 2.1.1.). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. An $\mathcal{F}$-measurable map $X : \Omega \to \mathbb{R}^d$ is called a d-dimensional random variable. A family of random variables $(X_t)_{t\geq 0}$ with $X_t : \Omega \to \mathbb{R}^d$ is called a stochastic process.

DEFINITION 1.20 (Brownian motion, see [**19**] Definition 2.1.1.). (a) A real-valued adapted stochastic process $(B_t)_{t\geq 0}$ on the stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ is called a one-dimensional standard Brownian motion with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ provided that the following conditions are satisfied:

(i) $B_0 = 0$ $a.s.$

(ii) The process has stationary increments: for $0 \leq s \leq t$ the random variable $B_t - B_s$ has a normal distribution $\mathcal{N}(0, t-s)$, i.e. for $a \in \mathbb{R}$

$$\mathbb{P}(B_t - B_s \leq a) = \frac{1}{2\pi(t-s)} \int_{-\infty}^{a} e^{-\frac{x^2}{2(t-s)}}\, dx.$$

(iii) The process has independent increments: for $0 \leq t_1 < t_2 < ... < t_n$, $n \in \mathbb{N}$, the random variables $B_1 - B_0, ..., B_{t_n} - B_{t_{n-1}}$ are independent.

(iv) The paths $t \mapsto B_t(w)$ are continuous for almost all $w \in \Omega$.

(b) A d-dimensional (standard) Brownian motion is a stochastic process $(B_t)_{t\geq 0} = (B_t^1, ..., B_t^d)_{t\geq 0}$, where the processes $(B_t^1)_{t\geq 0}, ..., (B_t^d)_{t\geq 0}$ are independent real-valued Brownian motions.

The stochastic integral used in the following is assumed to be known in this thesis. The construction and definition of the stochastic integral can be found [**19**] Chapter 5.

In this thesis our main focus is on stochastic differential equations (abbreviated SDEs), and more specifically on so-called McKean-Vlasov equations, also known as mean-field equations. For SDEs which do not have a mean-field term the existence and uniqueness of solutions is shown in [**19**] Chapter 6.

For the McKean-Vlasov type SDEs introduced in [**9**] Section 4.2. the coefficients depend also on the distribution of the solution:

$$\begin{cases} X_t = \xi + \int_0^t b(X_s, \mu_s)\mathrm{d}s + \int_0^t \sigma(X_s, \mu_s)\, dB_s \\ \mu_s = \text{ probability distribution of } X_s. \end{cases} \tag{1.1}$$

Here $(B_s)_{s\geq 0}$ is the d-dimensional Brownian motion on our assumed stochastic basis, ξ is an $\mathcal{F}_0$-measurable random variable and the coefficients $b : \mathbb{R}^d \times \mathcal{M}_{\lambda^2} \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{M}_{\lambda^2} \to \mathbb{R}^d \times \mathbb{R}^d$, where $\mathcal{M}_{\lambda^2}$ is defined in section 1.4., are Borel-measurable.

Now we define what it means to have a weak and a strong solution to Eq. (1.1). The weak and strong solution for SDEs without a mean-field term are defined in [**19**] Chapter 6.

Definition 1.21 (Weak solution, see [**21**] Definition 3.1.). Assume that there exists a stochastic basis $\hat{\mathcal{S}} := (\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}}, (\hat{\mathcal{F}}_t)_{t\in[0,T]})$ which satisfies the usual conditions, a d-dimensional $(\hat{\mathbb{P}}, (\hat{\mathcal{F}}_t)_{t\in[0,T]})$-Brownian motion $\hat{B}$ and a continuous $\mathbb{R}^d$-valued $(\hat{\mathcal{F}}_t)_{t\in[0,T]}$-adapted process $\hat{\mathbf{X}} = (\hat{X}_t)_{t\in[0,T]}$ on $\hat{\mathcal{S}}$. Then $(\hat{\mathcal{S}}; \hat{W}, \hat{\mathbf{X}})$ is called a weak solution to Eq. (1.1) with the initial law $\mu_0 = \mathbb{P} \circ \xi^{-1}$ if

(i) $\hat{\mathbb{P}} \circ \hat{X}_0^{-1} = \mu_0$,

(ii)
$$\int_0^T |b(\hat{X}_s, \hat{\mu}_s)| + \|\sigma(\hat{X}_s, \hat{\mu}_s)\|^2 \, ds < \infty, \qquad a.s.\, \hat{\mathbb{P}},$$
where $\hat{\mu}_t = \hat{\mathbb{P}} \circ \hat{X}_t^{-1}$,

(iii) For all $t \in [0,T]$ it holds that
$$\hat{X}_t = \xi + \int_0^t b(\hat{X}_s, \hat{\mu}_s) ds + \int_0^t \sigma(\hat{X}_s, \hat{\mu}_s) \, d\hat{B}_s \quad a.s.\, \hat{\mathbb{P}}.$$

The existence of a unique strong solution to a McKean-Vlasov equation in a slightly different setting is shown in [**9**] Section 4.2.

Definition 1.22 (Strong solution). Let $\mathbf{X} = (X_t)_{t\in[0,T]}$ be a continuous d-dimensional process on $(\Omega, \mathcal{F}, \mathbb{P})$. We say that $\mathbf{X}$ is a strong solution to Eq. (1.1) with the initial condition $X_0 = \xi$ provided that the following conditions hold:

(i) The process $\mathbf{X}$ is $(\mathcal{F}_t^B)_{t\in[0,T]}$-adapted, where $(\mathcal{F}_t^B)_{t\in[0,T]}$ is the augmented filtration generated by the Brownian motion B,

(ii)
$$\int_0^T \left(|b(X_t), \mu_t)| + \|\sigma(X_t, \mu_t)\|^2 \right) dt < \infty \; a.s.$$

(iii) For all $t \in [0,T]$ it holds that
$$X_t = \xi + \int_0^t b(X_s, \mu_s) ds + \int_0^t \sigma(X_s, \mu_s) \, dB_s \quad a.s.$$

Next we will introduce the martingale problem. Martingales are defined in [**27**] Definition 1, and the martingale problem related to SDEs without a mean-field term are discussed in [**19**] Chapter 7.

Let

$$\mathcal{W} := C([0,T],\mathbb{R}^d), \qquad \mathscr{W} = \mathscr{B}(\mathcal{W}),$$
$$\mathcal{W}_t := \{w(\cdot \wedge t) : w \in \mathcal{W}\}, \qquad \bar{\mathscr{W}}_t = \bigcap_{s>t} \mathscr{B}(\mathcal{W}_s), \qquad t \in [0,T], \tag{1.2}$$

where $\mathscr{B}(A)$ denotes the Borel-σ-algebra on A, i.e. the smallest σ-algebra generated by all open sets in A. By [**28**] Chapter 1.3 (p. 30) it holds that for $Z_u(w) = w_u$, $w \in \mathcal{W}_t$ we have $\mathscr{B}(\mathcal{W}_t) = \sigma(Z_u, 0 \le u \le t)$.

Definition 1.23 (The Martingale Problem, see [**21**] Definition 3.2.). We call a probability measure P on $(\mathcal{W}, \mathscr{W})$ a solution to the martingale problem associated with $\mathscr{A}$, if for $f \in C_0^2(\mathbb{R}^d)$ the process $(M_t^f)_{t\in[0,T]}$ with

$$M_t^f := f(w_t) - f(w_0) - \int_0^t (\mathscr{A}(\mu_s)f)(w_s)\, ds$$

is a continuous $(\bar{\mathscr{W}}_t)_{t\in[0,T]}$-adapted martingale, where $\mu_s := P \circ w_s^{-1}$ and

$$(\mathscr{A}(\mu)f)(x) := \frac{1}{2}\sum_{i,j=1}^d (\sigma(x,\mu)\sigma^T(x,\mu))_{ij}\frac{\partial^2}{\partial x_i \partial x_j} f + \sum_{i=1}^d b_i(x,\mu)\frac{\partial}{\partial x_i} f.$$

In the following we will introduce some notions of uniqueness:

Definition 1.24 (Pathwise uniqueness, see [**20**] Definition 3.2). Assume (X,B), $(\Omega,\mathscr{F},\mathbb{P})$, $\{\mathscr{F}_t\}$ and $(\tilde{X},B)$, $(\Omega,\mathscr{F},\mathbb{P})$, $\{\tilde{\mathscr{F}}_t\}$ to be weak solutions to Eq.(1.1) such that $\mathbb{P}(X_0 = \tilde{X}_0) = 1$. We say that *pathwise uniqueness* holds for Eq(1.1) provided that X and $\tilde{X}$ are indistinguishable, i.e.,

$$\mathbb{P}(X_t = \tilde{X}_t\,;\, 0 \le t \le \infty) = 1.$$

Definition 1.25 (Uniqueness in the sense of probability law, see [**20**] Definition 3.4.). Let (X,B), $(\Omega,\mathscr{F},\mathbb{P})$, $\{\mathscr{F}_t\}$ and $(\tilde{X},\tilde{B})$, $(\tilde{\Omega},\tilde{\mathscr{F}},\tilde{\mathbb{P}})$, $\{\tilde{\mathscr{F}}_t\}$ be weak solutions to Eq.(1.1) with the same initial distribution, i.e.,

$$\mathbb{P}(X_0 \in A) = \tilde{\mathbb{P}}(\tilde{X}_0 \in A) \qquad \text{for all } A \in \mathscr{B}(\mathbb{R}^d).$$

We say that *uniqueness in the sense of probability law* holds for Eq.(1.1) provided that the processes X and $\tilde{X}$ have the same law.

Proposition 1.26 (see [**20**] prop 3.20). *Pathwise uniqueness implies uniqueness in the sense of probability law.*

Ito's formula is a useful tool when handling with processes of a certain type. In the following we introduce Ito's formula for the d-dimensional case:

THEOREM 1.27 (Ito's Formula, see [**29**] Theorem 2.3.3.). *Assume a d-dimensional Brownian motion B and functions $b = (b_1, ..., b_{d_1})^T$ and $\sigma = (\sigma_{i,j})_{1\leq i\leq d_1, 1\leq j\leq d}$ such that $b_i \in L^1_{loc}((\mathcal{F}_t)_{t\in[0,T]})$ and $\sigma_{i,j} \in L^2_{loc}((\mathcal{F}_t)_{t\in[0,T]})$, where for $p = 1, 2$*

$$L^p_{loc}((\mathcal{F}_t)_{t\in[0,T]}) = \{(\mathcal{F}_t)_{t\in[0,T]} \text{ -progressively measurable processes } (X)_{t\in[0,T]} \text{ such that } \int_0^T |X_t|^p dt < \infty \text{ a.s.}\}.$$

Suppose that for $\mathbf{X} = (X_t)_{t\in[0,T]}$ it holds that

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

Let $f : [0;T] \times \mathbb{R}^d_1 \to \mathbb{R}$, $f \in C^{1,2}$. Then we have

$$\begin{aligned} f(t, X_t) =& f(0, X_0) + \int_0^t \frac{\partial f(s, X_s)}{\partial t} ds + \sum_{i=1}^{d_1} \int_0^t \frac{\partial f(s, X_s)}{\partial x_i} b_i(s)ds \\ &+ \frac{1}{2} \sum_{i,j=1}^{d_1} \sum_{k=1}^{d_2} \int_0^t \frac{\partial^2 f(s, X_s)}{\partial x_i \partial x_j} \sigma_{i,k}(s)\sigma_{j,k}(s)ds \\ &+ \sum_{i=1}^{d_1} \sum_{j=1}^{d_2} \int_0^t \frac{\partial f(s, X_s)}{\partial x_i} \sigma_{i,j}(s)dB^i_s. \end{aligned}$$

THEOREM 1.28 (Kolmogorov-Chentsov tighness criterion, see [**18**] Corollary 14.9.). *Let $(\mathbf{X}^n)_{n=1}^\infty$ be a sequence of continuous processes satisfying*

(i) $\sup_{n\geq 1} \mathbb{E}|X^n_0|^\delta < \infty$,

(ii) $\sup_{n\geq 1} \mathbb{E}|X^n_t - X^n_s|^\alpha \leq C_T|t - s|^{1+\beta}$ *for all* $T > 0$ *and* $0 \leq s \leq t \leq T$

for some α, β, $\delta > 0$ and $C_T > 0$. Then it holds that $\mathbb{P}^n := \mathbb{P} \circ (\mathbf{X}^n)^{-1}$, $n \in \mathbb{N}$, form a tight sequence on $(C([0, T], \mathbb{R}^d), \mathcal{C})$.

THEOREM 1.29 (see [**5**] Theorem 6.7.). *Let ξ, ξ_1, ξ_2, ... be random elements in a separable metric space (S, ρ) such that $\xi_n \xrightarrow{d} \xi$. Then, on a suitable probability space, there exist some random elements $\eta \overset{d}{=} \xi$ and $\eta \overset{d_n}{=} \xi_n$, $n \in \mathbb{N}$, with $\eta_n \to \eta$ a.s.*

In the following we define closable martingales and the connection with uniform integrability:

DEFINITION 1.30 (Closable Martingale, see [**26**] Chapter 2). Let $M = (M_t)_{t\geq 0}$ be a martingale. We say that M is closable if there exists a random variable Y such that $\mathbb{E}|Y| < \infty$ and

$$M_t = \mathbb{E}[Y|\mathcal{F}_t] \qquad \text{for all } t \geq 0.$$

THEOREM 1.31 (see [**26**] Chapter 2 Theorem 13). *Assume $M = (M_t)_{t\geq 0}$ to be a right continuous martingale. If M is closable then it is uniformly integrable.*

1.3. Some Inequalities

In this section we give some elementary inequalities. We start with Hölder's inequality, which takes two forms:

THEOREM 1.32 (Hölder's inequality, see [**24**] Theorem 3.9.9.). *Assume X and Y to be two real-valued random variables. For $p > 1$, $q > 1$, such that $\frac{1}{p} + \frac{1}{q} = 1$ and $\mathbb{E}(|X|^p) < \infty$, $\mathbb{E}(|Y|^q) < \infty$, it holds that*

$$\mathbb{E}|(XY)| \leq \big(\mathbb{E}(|X|^p)\big)^{\frac{1}{p}} \big(\mathbb{E}(|Y|^q)\big)^{\frac{1}{q}}.$$

For some cases Hölder's inequality can be presented in the following form:

REMARK 1.33 (see [**15**] p.53). Let $x, y \in \mathbb{R}^n$ and let $p > 1$, $q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\sum_{i=1}^{n} |x_i y_i| \leq \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{\frac{1}{p}} \Big(\sum_{i=1}^{n} |y_i|^q\Big)^{\frac{1}{q}}$$

PROPOSITION 1.34 (Young's Inequality, see [**8**] Section 2.2.4.). *Let $a, b \geq 0$. For $1 < p, q < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$ it holds that*

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

DEFINITION 1.35 (see [**22**] p. xvii). (i) We denote by $\mathcal{L}^2([a,b], \mathbb{R}^{d\times m})$ the family of $\mathbb{R}^{d\times m}$-valued $(\mathcal{F}_t)_{t\geq 0}$-progressively measurable processes $\{f(t)\}_{a\leq t\leq b}$ for which it holds that

$$\int_a^b |f(t)|^2 dt < \infty \ a.s.$$

(ii) We denote by $\mathcal{L}^2(\mathbb{R}_+, \mathbb{R}^{d\times m})$ the family of processes $\{f(t)\}_{t\geq 0}$ which satisfy $\{f(t)\}_{0\leq t\leq s} \in \mathcal{L}^2([0,s], \mathbb{R}^{d\times m})$ for every $s > 0$.

The following theorem is called the Burkholder-Davis-Gundy inequality. It is useful when estimates for norms of stochastic integrals are needed.

THEOREM 1.36 (Burkholder-Davis-Gundy inequality, see [**22**] Theorem 7.3.). *Let $g \in \mathcal{L}^2(\mathbb{R}_+, \mathbb{R}^{d\times m})$ and let $\|\cdot\|$ be the Frobenius norm (see page 12). For every $p > 0$ there exist constants $c_p > 0$, $C_p > 0$ which depend only on p and for which it holds that*

$$c_p \mathbb{E}\Big(\int_0^t \|g(s)\|^2 ds\Big)^{\frac{p}{2}} \leq \mathbb{E}\Big(\sup_{0\leq s\leq t} \Big|\int_0^s g(u) dB_u\Big|^p\Big) \leq C_p \mathbb{E}\Big(\int_0^t \|g(s)\|^2 ds\Big)^{\frac{p}{2}}.$$

THEOREM 1.37 (Grönwall Inequality, see [**18**] Lemma 18.4). *Assume $f(t)$ to be a continuous function for $t \geq 0$. If for some $a, b \geq 0$ it holds that*

$$f(t) \leq a + b \int_0^t f(s) ds$$

for all $t \geq 0$, then one has that

$$f(t) \leq ae^{bt}$$

for all $t \geq 0$.

The following lemma is a generalization of a Grönwall-Bellman type inequality:

LEMMA 1.38 (See [**30**] Lemma 2.1). *Let $0 < \eta < \frac{1}{e}$ and let*

$$\kappa_\eta(x) := \begin{cases} 0 & x = 0, \\ x \log x^{-1}, & 0 < x \leq \eta \\ (\log \eta^{-1} - 1)x + \eta, & x > \eta. \end{cases}$$

Assume that g and q are strictly positive functions on $[0, \infty)$ such that $g(0) < \eta$ and

$$g(t) \leq g(0) + \int_0^t q(s) \kappa_\eta\big(g(s)\big) ds, \qquad t \geq 0.$$

Then it holds that

$$g(t) \leq \big(g(0)\big)^{exp\{-\int_0^t q(s)ds\}}.$$

THEOREM 1.39 (Fatou's Lemma, see [**29**], Theorem 1.3.6 (i)). *For $n \geq 1$, assume that $0 \leq X^n \in L^1(\Omega, \mathscr{F}, \mathbb{P})$. Then one has*

$$\mathbb{E} \liminf_{n \to \infty} X^n \leq \liminf_{n \to \infty} \mathbb{E} X^n.$$

PROPOSITION 1.40 (Jensen's inequality, see [**29**] Proposition 1.3.1). *Assume that $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$, $X : \Omega \to \mathbb{R}^d$, and let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be convex. Then it follows that*

$$\varphi\big(\mathbb{E}[X]\big) \leq \mathbb{E}[\varphi(X)].$$

For a concave function we get the opposite result for Jensen's inequality:

COROLLARY 1.41. *Let $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$, $X : \Omega \to \mathbb{R}^d$, and let $\varphi : \mathbb{R}^d \to \mathbb{R}$ be concave. Then it holds that*

$$\varphi\big(\mathbb{E}[X]\big) \geq \mathbb{E}[\varphi(X)].$$

PROOF. If φ is concave, then $-\varphi$ is convex. Thus by Jensen's inequality

$$-\mathbb{E}(\varphi(X)) = \mathbb{E}(-\varphi(X)) \geq -\varphi((X)),$$

and the statement follows from multiplying with -1. □

1.4. Notation

By $C(\mathbb{R}^d)$ we denote the space of all continuous functions on $\mathbb{R}^d$, and by $C([0,T],\mathbb{R}^d)$ the space of continuous $\mathbb{R}^d$-valued functions defined on $[0,T]$. By $C_0^2(\mathbb{R}^d)$ we mean the collection of continuous functions which have continuous partial derivatives up to order 2 and vanish at infinity.

As for norms, for $x \in \mathbb{R}^d$ we use the Euclidean norm $|x| = \sqrt{\sum_{i=1}^d x_i^2}$ and for $y \in \mathbb{R}^d \times \mathbb{R}^d$ we use the Frobenius norm $\|y\| = \sqrt{\sum_{i,j=1}^d y_{ij}^2}$. For $f \in C([0,T],\mathbb{R}^d)$ we use the sup norm $\|f\|_\infty = \sup_{t\in[0,T]} |f(t)|$. Additionally, we denote with $\langle \cdot, \cdot \rangle$ the scalar product in $\mathbb{R}^d$. The transpose of a matrix A is denoted by A^T.

Set

$$C_\rho(\mathbb{R}^d) := \Big\{ \varphi \in C(\mathbb{R}^d), \|\varphi\|_{C_\rho(\mathbb{R}^d)} = \sup_{x\in\mathbb{R}^d} \frac{|\varphi(x)|}{(1+|x|)^2} + sup_{x\neq y} \frac{|\varphi(x)-\varphi(y)|}{|x-y|} < \infty \Big\}.$$

Let $\mathcal{M}(\mathbb{R}^d)$ be the space of probability measures on $\mathscr{B}(\mathbb{R}^d)$ carrying the usual topology of weak convergence and let $\mathcal{M}_{\lambda^2}^s(\mathbb{R}^d)$ be the Banach space of signed measures m on $\mathscr{B}(\mathbb{R}^d)$ for which

$$\|m\|_{\lambda^2}^2 := \int_{\mathbb{R}^d} (1+|x|)^2 |m|(dx) < \infty,$$

where $|m| = m^+ + m^-$ and $m = m^+ - m^-$ is the Jordan decomposition of m. Let $\mathcal{M}_{\lambda^2}(\mathbb{R}^d) = \mathcal{M}_{\lambda^2}^s(\mathbb{R}^d) \bigcap \mathcal{M}(\mathbb{R}^d)$ be the set of probability measures on $\mathscr{B}(\mathbb{R}^d)$.

Define

$$\rho(\mu,\nu) := \sup_{\|\varphi\|_{C_\rho(\mathbb{R}^d)}\leq 1} \Big| \int_{\mathbb{R}^d} \varphi(x)\mu(dx) - \int_{\mathbb{R}^d} \varphi(x)\nu(dx) \Big|. \tag{1.3}$$

As stated in [**10**] Section 2.1., $(\mathcal{M}_{\lambda^2}(\mathbb{R}^d), \rho)$ is now a complete metric space.

1.5. Assumptions

In this section we will state the assumptions under which we will work in this thesis.

($\mathbf{H}_1$) The functions $b : \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d) \to \mathbb{R}^{d\times d}$ are continuous and satisfy

$$|b(x,\mu)|^2 + \|\sigma(x,\mu)\|^2 \leq L_1(1+|x|^2+\|\mu\|_{\lambda^2}^2)$$

for all $(x,\mu) \in \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d)$, where $L_1 > 0$ is a constant.

($\mathbf{H}_2$) The functions $b : \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d) \to \mathbb{R}^{d\times d}$ satisfy for all (x_1,μ_1), $(x_2,\mu_2) \in \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d)$ that

$$\begin{aligned} &2\langle x_1 - x_2, b(x_1,\mu_1) - b(x_2,\mu_2)\rangle + \|\sigma(x_1,\mu_1) - \sigma(x_2,\mu_2)\|^2 \\ &\leq L_2\Big(\kappa_1(|x_1-x_2|^2) + \kappa_2\big(\rho^2(\mu_1,\mu_2)\big)\Big), \end{aligned}$$

where $L_2 > 0$ is a constant and $\kappa_i(x)$, $i = 1, 2$, are positive, strictly increasing, continuous and concave and satisfy $\kappa_i(0) = 0$, and $\int_{0^+}(\kappa_1(x) + \kappa_2(x))^{-1}dx = \infty$.

($\mathbf{H}_2'$) The functions $b : \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d) \to \mathbb{R}^{d\times d}$ satisfy for all (x_1, μ_1), $(x_2, \mu_2) \in \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d)$ that

$$|b(x_1, \mu_1) - b(x_2, \mu_2)| \le \lambda_1\Big(|x_1 - x_2|\gamma_1(|x_1 - x_2)| + \rho(\mu_1, \mu_2)\Big),$$
$$\|\sigma(x_1, \mu_1) - \sigma(x_2, \mu_2)\|^2 \le \lambda_2\Big(|x_1 - x_2|^2\gamma_2(|x_1 - x_2)| + \rho^2(\mu_1, \mu_2)\Big),$$

where λ_1, $\lambda_2 > 0$ are constants and γ_1, γ_2 are positive and continuous functions such that they are bounded on $[1, \infty)$ and satisfy

$$\lim_{x\downarrow 0} \frac{\gamma_i(x)}{x\log(x^{-1})} = \delta_i < \infty \qquad i = 1, 2.$$

Remark 1.42. Assume $b(x, \mu)$ satisfies ($\mathbf{H}_2'$). Then for (x_1, μ_1), $(x_2, \mu_2) \in \mathbb{R}^d \times \mathcal{M}_{\lambda^2}(\mathbb{R}^d)$ one has that

$$\begin{aligned}
&\langle x_1 - x_2, b(x_1, \mu_1) - b(x_2, \mu_2)\rangle \\
&\qquad \le |x_1 - x_2||b(x_1, \mu_1) - b(x_2, \mu_2)| \\
&\qquad \le \lambda_1\Big(|x_1 - x_2|^2\gamma_1(|x_1 - x_2|) + |x_1 - x_2|\rho(\mu_1, \mu_2)\Big) \\
&\qquad \le \lambda_1\Big(|x_1 - x_2|^2\gamma_1(|x_1 - x_2|) + |x_1 - x_2|^2 + \rho^2(\mu_1, \mu_2)\Big).
\end{aligned}$$

From the proof of Theorem 2.3. in [**30**] one has that since $\lim_{x\downarrow 0} \frac{x^2\gamma_i(x)+x^2}{x^2\log(x^{-2})} = 0$, it follows that there exists an $0 < \eta < \frac{1}{e}$ such that for a constant $C > 0$ we have

$$\begin{aligned}
x^2\gamma_i(x) + x^2 &\le C\kappa_\eta(x^2) \\
x^2\gamma_i^2(x) &\le C\kappa_\eta^2(x) \\
x^2\gamma_i(x) &\le C\kappa_\eta(x^2) \qquad \text{for } i = 1, 2,
\end{aligned} \tag{1.4}$$

where

$$\kappa_\eta(x) := \begin{cases} 0 & x = 0, \\ x\log x^{-1}, & 0 < x \le \eta \\ (\log\eta^{-1} - 1)x + \eta, & x > \eta, \end{cases} \tag{1.5}$$

is a positive, strictly increasing, continuous concave function with $\kappa_\eta(0) = 0$ and $\int_{0^+} \frac{1}{\kappa_\eta(x)+x}dx = \infty$.

Hence we get

$$\langle x_1 - x_2,\, b(x_1, \mu_1) - b(x_2, \mu_2)\rangle \le \hat{\lambda}_1\Big(\kappa_\eta(|x_1 - x_2|^2) + \rho^2(\mu_1, \mu_2)\Big),$$

where $\hat{\lambda}_1$ depends on C and λ_1.

Additionally, assuming that $\sigma(x, \mu)$ satisfies ($\mathbf{H}_2'$) we have with (1.4) that

$$\begin{aligned}\|\sigma(x_1, \mu_1) - \sigma(x_2, \mu_2)\|^2 &\leq \lambda_2\Big(|x_1 - x_2|^2\gamma_2(|x_1 - x_2)| + \rho^2(\mu_1, \mu_2)\Big)\\ &\leq \hat{\lambda}_2\Big(\kappa_\eta(|x_1 - x_2|^2) + \rho^2(\mu_1, \mu_2)\Big),\end{aligned}$$

where $\hat{\lambda}_2$ depends on C and λ_2.

Therefore it holds that ($\mathbf{H}_2'$) implies ($\mathbf{H}_2$).

CHAPTER 2

The Existence and Uniqueness of Strong Solutions

In this chapter we will show that there exists a unique strong solution to Eq.(1.1), which is one of the main results of this thesis. To prove this statement we consider a solution to the martingale problem and pathwise uniqueness and show how these imply a unique strong solution.

2.1. The Martingale Problem

In the first step we use the existence of a solution to the martingale problem to get a weak solution to Eq. (1.1). First we show some techincal results:

LEMMA 2.1. *Assume that $b(x,\mu)$ and $\sigma(x,\mu)$ satisfy $(\mathbf{H}_1)$. Let $(\hat{S};\hat{B},\hat{\mathbf{X}})$ be a weak solution to Eq.*(1.1) *and let $\hat{\mathbb{E}}$ denote the expectation under $\hat{\mathbb{P}}$. Then, for $p\geq 1$, it holds that*

$$\hat{\mathbb{E}}(|\hat{X}_t|^{2p}) \leq C(1+\hat{\mathbb{E}}|\hat{X}_0|^{2p})\, e^{Ct}, \qquad 0\leq t\leq T, \tag{2.1}$$

$$\hat{\mathbb{E}}(|\hat{X}_t-\hat{X}_s|^{2p}) \leq C(1+\hat{\mathbb{E}}|\hat{X}_0|^{2p})\,(t-s)^p, \qquad 0\leq s\leq t\leq T, \tag{2.2}$$

where $C>0$ is a constant depending on T, p and L_1 from $(\mathbf{H}_1)$.

PROOF. Let $\tau_k := \inf\{t\geq 0, |\hat{X}_t|\geq k\}$, $k\in\mathbb{N}$. Since $\liminf_{k\to\infty}|\hat{X}_{\tau_k\wedge t}| = |\hat{X}_t|$ then by Fatou's Lemma (Theorem 1.39) we have that $\hat{\mathbb{E}}|\hat{X}_t|\leq \liminf_{k\to\infty}\hat{\mathbb{E}}|\hat{X}_{\tau_k\wedge t}|$. Thus if $\hat{X}_{\tau_k\wedge t}$ satisfies (2.1) and (2.2) then the inequalities hold also for $\hat{X}_t$.

For $0\leq t\leq T$, by Hölder's Inequality (Remark 1.33) we have

$$\begin{aligned}
\hat{\mathbb{E}}|\hat{X}_t|^{2p} \leq& \hat{\mathbb{E}}\Big(|\hat{X}_0| + \Big|\int_0^t b(\hat{X}_s,\hat{\mu}_s)ds\Big| + \Big|\int_0^t \sigma(\hat{X}_s,\hat{\mu}_s)d\hat{B}_s\Big|\Big)^{2p}\\
\leq& \hat{\mathbb{E}}\Big(3^{\frac{2p-1}{2p}}\Big(|\hat{X}_0|^{2p} + \Big|\int_0^t b(\hat{X}_s,\hat{\mu}_s)ds\Big|^{2p} + \Big|\int_0^t \sigma(\hat{X}_s,\hat{\mu}_s)d\hat{B}_s\Big|^{2p}\Big)^{\frac{1}{2p}}\Big)^{2p}\\
=& 3^{2p-1}\Big(\hat{\mathbb{E}}|\hat{X}_0|^{2p} + \hat{\mathbb{E}}\Big|\int_0^t b(\hat{X}_s,\hat{\mu}_s)ds\Big|^{2p} + \hat{\mathbb{E}}\Big|\int_0^t \sigma(\hat{X}_s,\hat{\mu}_s)d\hat{B}_s\Big|^{2p}\Big).
\end{aligned}$$

By Lemmas 2.8 and 2.9 from section 2.3. it holds that

$$\begin{aligned}
\hat{\mathbb{E}}|\hat{X}_t|^{2p} \leq& 3^{2p-1}\Big(\hat{\mathbb{E}}|\hat{X}_0|^{2p} + t^{2p-1}\hat{\mathbb{E}}\int_0^t |b(\hat{X}_s,\hat{\mu}_s)|^{2p}ds + C_p\, t^{p-1}\,\hat{\mathbb{E}}\int_0^t \|\sigma(\hat{X}_s,\hat{\mu}_s)\|^{2p}ds\Big)\\
\leq& C_1\Big(\hat{\mathbb{E}}|\hat{X}_0|^{2p} + \int_0^t \hat{\mathbb{E}}\big(|b(\hat{X}_s,\hat{\mu}_s)|^{2p} + \|\sigma(\hat{X}_s,\hat{\mu}_s)\|^{2p}\big)ds\Big)
\end{aligned}$$

for $C_1 := 3^{2p-1}(1 + t^{2p-1} + C_p t^{p-1})$.

Furthermore, by Lemma 2.10 from section 2.3. we get that for some positive constants C_2, C_3, C_4

$$\begin{aligned}
\hat{\mathbb{E}}|\hat{X}_t|^{2p} &\leq C_2\Big(\hat{\mathbb{E}}|\hat{X}_0|^{2p} + \int_0^t \hat{\mathbb{E}}(1 + |\hat{X}_s|^{2p} + \hat{\mathbb{E}}(1 + |\hat{X}_s|^{2p}))ds\Big) \\
&\leq C_3\Big(\hat{\mathbb{E}}|\hat{X}_0|^{2p} + \int_0^t 2(1 + \hat{\mathbb{E}}|\hat{X}_s|^{2p})ds\Big) \\
&\leq C_4\Big(1 + \hat{\mathbb{E}}|\hat{X}_0|^{2p} + \int_0^t \hat{\mathbb{E}}|\hat{X}_s|^{2p} ds\Big) \\
&= C_4\big(1 + \hat{\mathbb{E}}|\hat{X}_0|^{2p}\big) + C_4 \int_0^t \hat{\mathbb{E}}|\hat{X}_s|^{2p} ds
\end{aligned}$$

Hence by Grönwall's inequality (Theorem 1.37) we get

$$\hat{\mathbb{E}}|\hat{X}_t|^{2p} \leq C\Big(1 + \hat{\mathbb{E}}|\hat{X}_0|^{2p}\Big)e^{Ct},$$

where C is a constant depending on p, L_1, and T.

To show (2.2) we conduct similar calculations using Hölder's Inequality (Remark 1.33) and Lemmas 2.8, 2.9 and 2.10 from section 2.3.:

$$\begin{aligned}
\hat{\mathbb{E}}|\hat{X}_t - \hat{X}_s|^{2p} &= \hat{\mathbb{E}}\Big|\int_s^t b(\hat{X}_u, \hat{\mu}_u)du + \int_s^t \sigma(\hat{X}_u, \hat{\mu}_u)d\hat{B}_u\Big|^{2p} \\
&\leq \hat{\mathbb{E}}\Big(\Big|\int_s^t b(\hat{X}_u, \hat{\mu}_u)du\Big| + \Big|\int_s^t \sigma(\hat{X}_u, \hat{\mu}_u)d\hat{B}_u\Big|\Big)^{2p} \\
&\leq \hat{\mathbb{E}}\Big(\Big(\Big|\int_s^t b(\hat{X}_u, \hat{\mu}_u)du\Big|^{2p} + \Big|\int_s^t \sigma(\hat{X}_u, \hat{\mu}_u)d\hat{B}_u\Big|^{2p}\Big)^{\frac{1}{2p}} 2^{\frac{2p-1}{2p}}\Big)^{2p} \\
&= \hat{\mathbb{E}}\Big(\Big|\int_s^t b(\hat{X}_u, \hat{\mu}_u)du\Big|^{2p} + \Big|\int_s^t \sigma(\hat{X}_u, \hat{\mu}_u)d\hat{B}_u\Big|^{2p}\Big)2^{2p-1} \\
&\leq \tilde{C}_1(t-s)^{p-1}\hat{\mathbb{E}}\int_s^t |b(\hat{X}_u, \hat{\mu}_u)|^{2p} + \|\sigma(\hat{X}_u, \hat{\mu}_u)\|^{2p} du \\
&\leq \tilde{C}_2(t-s)^{p-1}\int_s^t 1 + \hat{\mathbb{E}}|\hat{X}_u|^{2p} du.
\end{aligned}$$

Now we can use (2.1) and get

$$\begin{aligned}
\hat{\mathbb{E}}|\hat{X}_t - \hat{X}_s|^{2p} &\leq \tilde{C}_3(t-s)^{p-1}\int_s^t C(1 + \hat{\mathbb{E}}|\hat{X}_0|^{2p})e^{Cu}du \\
&\leq \tilde{C}(t-s)^{p-1}(1 + \hat{\mathbb{E}}|\hat{X}_0|^{2p})(t-s) \\
&= \tilde{C}(1 + \hat{\mathbb{E}}|\hat{X}_0|^{2p})(t-s)^p,
\end{aligned}$$

where $\tilde{C}$ is a constant depending on T, p, and L_1. □

PROPOSITION 2.2 (See [**4**] Theorem V.1.1. p.98). *The existence of a solution to the martingale problem implies the existence of weak solutions and vice versa.*

Next we will prove the existence of a solution to the martingale problem:

PROPOSITION 2.3. *Suppose that* $(\mathbf{H}_1)$ *holds and* $\mathbb{E}|\xi|^{2p} < \infty$ *for any* $p > 1$. *Then there exists a solution to the martingale problem related to Eq.* (1.1).

PROOF. Let $n \in \mathbb{N}$ be fixed. For $t \in [0, T]$, we consider the Euler-Maruyama approximation equation

$$X_t^n = X_0^n + \int_0^t b(X_{t_n(s)}^n, \mu_{t_n(s)}^n)ds + \int_0^t \sigma(X_{t_n(s)}^n, \mu_{t_n(s)}^n)dB_s, \tag{2.3}$$

where $X_0^n = \xi$ and $t_n(s) = \frac{kT}{2^n} =: t_k$ for $s \in \left[\frac{kT}{2^n}, \frac{(k+1)T}{2^n}\right)$, $k = 0, 1, ..., 2^n - 1$.

Since $b(X_{t_n(s)}^n, \mu_{t_n(s)}^n)$ and $\sigma(X_{t_n(s)}^n, \mu_{t_n(s)}^n)$ are constant on the interval $[t_k, t_{k+1})$, we have that

$$\begin{aligned} X_{t_1}^n &= X_0^n + \int_0^{t_1} b(X_{t_0}^n, \mu_{t_0}^n)ds + \int_0^{t_1} \sigma(X_{t_0}^n, \mu_{t_0}^n)dB_s \\ &= \xi + b(\xi, \mu_{t_0}^n)t_1 + \sigma(\xi, \mu_{t_0}^n)B_{t_1}, \end{aligned}$$

and again

$$X_{t_2}^n = X_{t_1}^n + b(X_{t_1}^n, \mu_{t_1}^n)(t_2 - t_1) + \sigma(X_{t_1}^n, \mu_{t_1}^n)(B_{t_2} - B_{t_1}).$$

Thus we can construct a recursive definition for X_t^n, for $t \in [t_k, t_{k+1})$, as follows:

$$\begin{aligned} X_t^n &= X_{t_k}^n + \int_{t_k}^t b(X_{t_n(s)}^n, \mu_{t_n(s)}^n)ds + \int_{t_k}^t \sigma(X_{t_n(s)}^n, \mu_{t_n(s)}^n)dB_s \\ &= X_{t_k}^n + b(X_{t_k}^n, \mu_{t_k}^n)(t - t_k) + \sigma(X_{t_k}^n, \mu_{t_k}^n)(B_t - B_{t_k}). \end{aligned} \tag{2.4}$$

Hence there exists a solution $\mathbf{X^n} := (X_t^n)_{t\in[0,T]}$ to Eq.(2.3), and by $(\mathbf{H}_1)$ and Lemma 2.1 we have

$$\begin{aligned} \mathbb{E}(|X_t^n|^{2p}) &\le C(1 + \mathbb{E}|\xi|^{2p})e^{Ct}, \qquad 0 \le t \le T, \\ \mathbb{E}(|X_t^n - X_s^n|^{2p}) &\le C(1 + \mathbb{E}|\xi|^{2p})(t - s)^p, \qquad 0 \le s \le t \le T, \end{aligned} \tag{2.5}$$

where C does not depend on n.

Moreover, since $\mathbb{E}|\xi|^{2p} < \infty$, we get

$$\begin{aligned} \sup_{n\ge1} \mathbb{E}(|X_0^n|^{2p}) &= \mathbb{E}|\xi|^{2p} < \infty, \\ \sup_{n\ge1} \mathbb{E}(|X_t^n - X_s^n|^{2p}) &\le C(1 + \mathbb{E}|\xi|^{2p})(t - s)^p \le C_2(t - s)^p. \end{aligned}$$

Put $P^n := \mathbb{P} \circ (\mathbf{X^n})^{-1}$. By Theorem 1.28 we have that $\{P^n\}$ is tight on $(\mathcal{W}, \mathscr{W})$ defined in (1.2), and by Theorem 1.18 we further have that there exists a subsequence, which we still denote by $\{P^n\}$, and a probability measure P^0 on $(\mathcal{W}, \mathscr{W})$ such that P^n converges weakly to P^0 as $n \to \infty$.

Set

$$\begin{aligned} M_t^{n,f} :=& f(w_t) - f(w_0) - \frac{1}{2} \int_0^t \sum_{i,j=1}^d [\sigma(w_{t_n(s)}, \mu^n_{t_n(s)}) \sigma^T(w_{t_n(s)}, \mu^n_{t_n(s)})]_{ij} \frac{\partial^2}{\partial x_i \partial x_j} f(w_s) ds \\ & - \int_0^t \sum_{i=1}^d b_i(w_{t_n(s)}, \mu^n_{t_n(s)}) \frac{\partial}{\partial x_i} f(w_s) ds, \qquad\qquad f \in C_0^2(\mathbb{R}^d), \end{aligned}$$

where $\mu^n_{t_n(s)} = P^n \circ w^{-1}_{t_n(s)}$.

Since there exists a strong and therefore also a weak solution $\mathbf{X^n}$ to Eq.(2.3), by Proposition 2.2 we have that Eq.(2.3) has a solution P^n to the martingale problem on $(\mathcal{W}, \mathscr{W})$. Thus $(M_t^{n,f})_{t \in [0,T]}$ is a continuous $(\bar{\mathscr{W}}_t)_{t \in [0,T]}$-adapted martingale under P^n, so for a continuous, bounded and $\bar{\mathscr{W}}_s$-measurable functional G_s we have that

$$\mathbb{E}^{P_n}((M_t^{n,f} - M_s^{n,f}) G_s) = 0, \qquad 0 \le s < t \le T.$$

In order to prove that P^0 on $(\mathcal{W}, \mathscr{W})$ is a martingale solution to Eq.(1.1) we use Lemma 2.15 from Section 2.3 to show that $(M_t^f)_{t \in [0,T]}$ is a continuous $\bar{\mathscr{W}}_t$-adapted martingale under P^0, i.e. we want to get

$$\mathbb{E}^{P^0}((M_t^f - M_s^f) G_s) = \int_{\mathcal{W}} \Big(\Big(f(w_t) - f(w_s) - \int_s^t \mathscr{A}(\mu_u) f(w_u) du \Big) G_s(w) \Big) P^0(dw) = 0.$$

We prove this by showing

$$\lim_{n \to \infty} \mathbb{E}^{P_n}((M_t^{n,f} - M_s^{n,f}) G_s) = \mathbb{E}^{P^0}((M_t^f - M_s^f) G_s).$$

By Definition 1.3 the map $w \mapsto w_t$, where $w \in \mathcal{W}$, $t \in [0,T]$, is Lipschitz continuous: Let $\varepsilon > 0$ and put $\delta_\varepsilon = \varepsilon$. If for all $w, \hat{w} \in \mathcal{W}$

$$\|w - \hat{w}\|_\infty = \sup_{s \in [0,T]} |w_s - \hat{w}_s| < \delta_\varepsilon,$$

then one has that

$$|w_t - \hat{w}_t| \le \sup_{s \in [0,T]} |w_s - \hat{w}_s| < \delta_\varepsilon = \varepsilon.$$

Now since both f and G are bounded and continuous, from Proposition 1.4 and the weak convergence of P^n to P^0 we get

$$\lim_{n \to \infty} \int_{\mathcal{W}} \Big(f(w_t) - f(w_s)) G_s(w) \Big) P^n(dw) = \int_{\mathcal{W}} \Big((f(w_t) - f(w_s)) G_s(w) \Big) P^0(dw).$$

Our goal is to now show that

$$\lim_{n\to\infty}\int_{\mathcal{W}}\Big(\Big(\int_s^t\sum_{i=0}^d b_i(w_{t_n(u)},\mu^n_{t_n(u)})\frac{\partial}{\partial x_i}f(w_u)du\Big)G_s(w)\Big)P^n(dw)$$
$$=\int_{\mathcal{W}}\Big(\Big(\int_s^t\sum_{i=0}^d b_i(w_u,\mu_u)\frac{\partial}{\partial x_i}f(w_u)du\Big)G_s(w)\Big)P^0(dw) \tag{2.6}$$

and

$$\lim_{n\to\infty}\int_{\mathcal{W}}\Big(\Big(\int_s^t\sum_{i,j=0}^d\big(\sigma(w_{t_n(u)},\mu^n_{t_n(u)})\sigma^T(w_{t_n(u)},\mu^n_{t_n(u)})\big)_{ij}\frac{\partial^2 f(w_u)}{\partial x_i\partial x_j}du\Big)G_s(w)\Big)P^n(dw)$$
$$=\int_{\mathcal{W}}\Big(\Big(\int_s^t\sum_{i,j=0}^d\big(\sigma(w_u,\mu_u)\sigma^T(w_u,\mu_u)\big)_{ij}\frac{\partial^2 f(w_u)}{\partial x_i\partial x_j}du\Big)G_s(w)\Big)P^0(dw). \tag{2.7}$$

Since we know that P^n converges weakly to P^0, we get from Theorem 1.29 that there exists a probability space $(\tilde{\Omega},\tilde{\mathscr{F}},\tilde{\mathbb{P}})$ and $\mathcal{W}$-valued processes $\tilde{\mathbf{X}}^n := (\tilde{X}^n_t)_{t\in[0,T]}$ and $\tilde{\mathbf{X}} := (\tilde{X}_t)_{t\in[0,T]}$ on that space such that

(i) $P^n = \tilde{\mathbb{P}}\circ(\tilde{\mathbf{X}}^n)^{-1}$ and $P^0 = \tilde{\mathbb{P}}\circ(\tilde{\mathbf{X}})^{-1}$
(ii) $\tilde{\mathbf{X}}^n \xrightarrow{a.s.} \tilde{\mathbf{X}}$ as $n\to\infty$.

With the help of (i), (2.6) and (2.7) can now be written again as

$$\lim_{n\to\infty}\mathbb{E}^{\tilde{\mathbb{P}}}\Big(\Big(\int_s^t\sum_{i=0}^d b_i(\tilde{X}^n_{t_n(u)},\mu^n_{t_n(u)})\frac{\partial}{\partial x_i}f(\tilde{X}^n_u)du\Big)G_s(\tilde{\mathbf{X}}^n)\Big)$$
$$=\mathbb{E}^{\tilde{\mathbb{P}}}\Big(\Big(\int_s^t\sum_{i=0}^d b_i(\tilde{X}_u,\mu_u)\frac{\partial}{\partial x_i}f(\tilde{X}_u)du\Big)G_s(\tilde{\mathbf{X}})\Big) \tag{2.8}$$

and

$$\lim_{n\to\infty}\mathbb{E}^{\tilde{\mathbb{P}}}\Big(\Big(\int_s^t\sum_{i,j=0}^d\big(\sigma(\tilde{X}^n_{t_n(u)},\mu^n_{t_n(u)})\sigma^T(\tilde{X}^n_{t_n(u)},\mu^n_{t_n(u)})\big)_{ij}\frac{\partial^2}{\partial x_i\partial x_j}f(\tilde{X}^n_u)du\Big)G_s(\tilde{\mathbf{X}}^n)\Big)$$
$$=\mathbb{E}^{\tilde{\mathbb{P}}}\Big(\Big(\int_s^t\sum_{i,j=0}^d\big(\sigma(\tilde{X}_u,\mu_u)\sigma^T(\tilde{X}_u,\mu_u)\big)_{ij}\frac{\partial^2}{\partial x_i\partial x_j}f(\tilde{X}_u)du\Big)G_s(\tilde{\mathbf{X}})\Big). \tag{2.9}$$

Moreover, from (ii) it follows that $\tilde{X}^n_{t_n(u)} \xrightarrow{a.s.} \tilde{X}_u$ for $u\in[s,t]$ as $n\to\infty$.

Now we will show that for ρ defined in (1.3) we have $\rho(\mu^n_{t_n(u)},\mu_u)\to 0$ as $n\to\infty$:

Since $\mathbf{X}^n$ and $\tilde{\mathbf{X}}^n$ have the same law, then by (2.5) we have for any $\lambda > 0$

$$\begin{aligned}\int_{|\tilde{X}^n_{t_n(u)}|>\lambda} |\tilde{X}^n_{t_n(u)}| \mathrm{d}\tilde{\mathbb{P}} &\leq \int_{|\tilde{X}^n_{t_n(u)}|>\lambda} |\tilde{X}^n_{t_n(u)}| \left(\frac{|\tilde{X}^n_{t_n(u)}|}{\lambda}\right)^{2p-1} \mathrm{d}\tilde{\mathbb{P}} \\ &= \int_{|\tilde{X}^n_{t_n(u)}|>\lambda} \frac{|\tilde{X}^n_{t_n(u)}|^{2p}}{\lambda^{2p-1}} \mathrm{d}\tilde{\mathbb{P}} \\ &\leq \frac{1}{\lambda^{2p-1}} \mathbb{E}^{\tilde{\mathbb{P}}} |\tilde{X}^n_{t_n(u)}|^{2p} \\ &= \frac{1}{\lambda^{2p-1}} \mathbb{E} |X^n_{t_n(u)}|^{2p} \\ &\leq \frac{1}{\lambda^{2p-1}} C(1+\mathbb{E}|\xi|^{2p}) e^{CT}.\end{aligned}$$

Furthermore, we have that

$$\lim_{\lambda\to\infty} \sup_{n\geq 1} \int_{|\tilde{X}^n_{t_n(u)}|>\lambda} |\tilde{X}^n_{t_n(u)}| \mathrm{d}\tilde{\mathbb{P}} = 0,$$

so by Definition 1.10 we know that $\{\tilde{X}^n_{t_n(u)}\}_{n\geq 1}$ is uniformly integrable. From the almost sure convergence of $\tilde{X}^n_{t_n(u)}$ to $\tilde{X}_u$ and the uniform integrability of $\{\tilde{X}^n_{t_n(u)}\}_{n\geq 1}$ it follows from Theorems 1.8 and 1.12 and Lemma 2.11 from section 2.3. that

$$0 \leq \lim_{n\to\infty} \rho(\mu^n_{t_n(u)}, \mu_u) \leq \lim_{n\to\infty} \mathbb{E}^{\tilde{\mathbb{P}}} \left| \tilde{X}^n_{t_n(u)} - \tilde{X}_u \right| = 0,$$

and thus we get $\lim\limits_{n\to\infty} \rho(\mu^n_{t_n(u)}, \mu_u) = 0$.

Let $1 < \alpha < \infty$. With (i), $(\mathbf{H}_1)$, Hölder's Inequality (Remark 1.33), (2.15) from section 2.3. and (2.5) it holds that

$$\begin{aligned}\mathbb{E}^{\tilde{\mathbb{P}}} \Big| \sum_{i=0}^{d} b_i(\tilde{X}^n_{t_n(u)}, \mu^n_{t_n(u)}) \Big|^{\alpha} &= \mathbb{E} \Big| \sum_{i=0}^{d} b_i(X^n_{t_n(u)}, \mu^n_{t_n(u)}) \Big|^{\alpha} \leq \mathbb{E} \Big(\sum_{i=0}^{d} |b_i(X^n_{t_n(u)}, \mu^n_{t_n(u)}|\Big)^{\alpha} \\ &\leq \mathbb{E} \Big(\sqrt{d}\, |b(X^n_{t_n(u)}, \mu^n_{t_n(u)})|^2 \Big)^{\alpha} \\ &= \sqrt{d^{\alpha}}\ \mathbb{E} |b(X^n_{t_n(u)}, \mu^n_{t_n(u)})|^{2\alpha} \\ &\leq C_1\ \mathbb{E}\big(1 + |X^n_{t_n(u)}|^{2\alpha} + 2^{2\alpha-1}\mathbb{E}(1+|X^n_{t_n(u)}|^{2\alpha})\big) \\ &\leq C_2 \Big(\mathbb{E}(1+|X^n_{t_n(u)}|^{2\alpha}) \Big) \\ &\leq C\Big(1 + C(1+\mathbb{E}|\xi|^{2\alpha})e^{CT}\Big) < \infty, \end{aligned} \tag{2.10}$$

where the constant C does not depend on n. By Proposition 1.11 we have that $\{\sum_{i=0}^{d} b_i(\tilde{X}^n_{t_n(u)}, \mu^n_{t_n(u)})\}_{n\in\mathbb{N}}$ is uniformly integrable.

Furthermore, since b is continuous, $\tilde{X}^n_{t_n(u)} \xrightarrow{a.s.} \tilde{X}_u$, $\lim\limits_{n\to\infty} \rho(\mu^n_{t_n(u)}, \mu_u) = 0$, $f \in C^2_0(\mathbb{R}^d)$, and G_s is bounded and continuous, it holds that

$$\sum_{i=1}^{d} b_i(\tilde{X}^n_{t_n(u)}, \mu^n_{t_n(u)})\frac{\partial}{\partial x_i} f(\tilde{X}^n_u) \xrightarrow{\mathbb{P}} \sum_{i=1}^{d} b_i(\tilde{X}_u, \mu_u)\frac{\partial}{\partial x_i} f(\tilde{X}_u) \qquad \text{and}$$

$$G_s(\tilde{\mathbf{X}}^n) \xrightarrow{a.s.} G_s(\tilde{\mathbf{X}}).$$

Now from Theorem 1.8 and Lemma 1.13 it follows that (2.8) holds.

Similarly, with the help of (2.15) from section 2.3 and (2.5) we have

$$\begin{aligned}
&\mathbb{E}\Big|\sum_{i,j=1}^{d}\big(\sigma(\tilde{X}^n_{t_n(u)}, \mu^n_{t_n(u)})\sigma^T(\tilde{X}^n_{t_n(u)}, \mu^n_{t_n(u)})\big)_{ij}\Big|^{\alpha} \\
&\le \mathbb{E}\Big(\sum_{i,j=1}^{d}\Big|\sum_{k=1}^{d}\sigma_{ik}(X^n_{t_n(u)}, \mu^n_{t_n(u)})\sigma_{jk}(X^n_{t_n(u)}, \mu^n_{t_n(u)})\Big|\Big)^{\alpha} \\
&\le \mathbb{E}\Big(d\sum_{i,j=1}^{d}\Big|\sum_{k=1}^{d}\sigma_{ik}(X^n_{t_n(u)}, \mu^n_{t_n(u)})\sigma_{jk}(X^n_{t_n(u)}, \mu^n_{t_n(u)})\Big|^2\Big)^{\alpha} \\
&= d^{\alpha}\,\mathbb{E}\big(\|\sigma(X^n_{t_n(u)}, \mu^n_{t_n(u)})\sigma^T(X^n_{t_n(u)}, \mu^n_{t_n(u)})\|^2\big)^{\alpha} \\
&\le d^{\alpha}\,\mathbb{E}\big(\|\sigma(X^n_{t_n(u)}, \mu^n_{t_n(u)})\|^{4\alpha}\big) \\
&\le C'\Big(1 + C'(1+\mathbb{E}|\xi|^{4\alpha})e^{C'T}\Big) < \infty,
\end{aligned} \tag{2.11}$$

where the constant C' does not depend on n. Then by Proposition 1.11 it holds that also $\{\sum_{i,j=1}^{d}(\sigma(\tilde{X}^n_{t_n(u)}, \mu^n_{t_n(u)})\sigma^T(\tilde{X}^n_{t_n(u)}, \mu^n_{t_n(u)}))_{ij}\}_{n\in\mathbb{N}}$ is uniformly integrable, and since σ is continuous, we get from Lemma 1.13 that also (2.9) holds. □

Now Propositions 2.2 and 2.3 give us the existence of weak solutions to Eq. (1.1).

2.2. Pathwise Uniqueness

Now that we have the existence of weak solutions, for the second step we will show that the pathwise uniqueness holds under our conditions. First we show a technical lemma:

LEMMA 2.4. *Let $t \ge 0$. Provided that for a Borel-measurable function $y : [0,\infty) \to [0,\infty)$ it holds that*

$$0 \le y_t \le \int_0^t (\kappa_1(y_s) + \kappa_2(y_s))ds < \infty,$$

where $\kappa_1(u)$ and $\kappa_2(u)$ satisfy the conditions in $(\mathbf{H}_2)$, it follows that $y_t \equiv 0$ for all $t \ge 0$.

PROOF. Put $z_t := \int_0^t (\kappa_1(y_s) + \kappa_2(y_s))ds$. From the properties of $\kappa_1(u)$ and $\kappa_2(u)$ we get that z_t is absolutely continuous and nondecreasing. Now, since $y_t \leq z_t$ and κ_1 and κ_2 are increasing, we have for all $t \geq 0$ that

$$\frac{dz_t}{dt} = \kappa_1(y_t) + \kappa_2(y_t) \leq \kappa_1(z_t) + \kappa_2(z_t).$$

Now set $t_0 := \sup\{t \geq 0; z_s = 0 \,\forall s \in [0,t]\}$. If $t_0 < \infty$ then for $t > t_0$ we have $z_t > 0$. Using the properties of κ_1 and κ_2 and substitution we get

$$\begin{aligned}\infty = \int_0^{z_{t_0+\varepsilon}} \frac{du}{\kappa_1(u) + \kappa_2(u)} = \int_{t_0}^{t_0+\varepsilon} \frac{dz_t}{\kappa_1(z_t) + \kappa_2(z_t)} &= \int_{t_0}^{t_0+\varepsilon} \frac{\kappa_1(y_t) + \kappa_2(y_t)}{\kappa_1(z_t) + \kappa_2(z_t)} dt \\ &\leq \int_{t_0}^{t_0+\varepsilon} dt = \varepsilon\end{aligned}$$

for all $\varepsilon > 0$. This is a controdiction, and therefore $t_0 = \infty$ from which it follows that $z_t = 0$.
Now $0 \leq y_t \leq 0$, so we get that $y_t = 0$. □

PROPOSITION 2.5. *Assume that the conditions of* $(\mathbf{H}_2)$ *are satisfied. Then we have pathwise uniqueness for Eq.*(1.1).

PROOF. Assume that for Eq.(1.1) there exist two weak solutions $(\hat{\mathcal{S}}; \hat{W}, (\hat{X}_t^1)_{t\in[0,T]})$ and $(\hat{\mathcal{S}}; \hat{W}, (\hat{X}_t^2)_{t\in[0,T]})$ such that $\hat{X}_0^1 = \hat{X}_0^2$ *a.s.*

Let $Z_t := \hat{X}_t^1 - \hat{X}_t^2$. Now it holds that

$$Z_t = \int_0^t \Big(b(\hat{X}_s^1, \hat{\mu}_s^1) - b(\hat{X}_s^2, \hat{\mu}_s^2)\Big) ds + \int_0^t \Big(\sigma(\hat{X}_s^1, \hat{\mu}_s^1) - \sigma(\hat{X}_s^2, \hat{\mu}_s^2)\Big) d\hat{W}_s.$$

By Ito's formula (Theorem 1.27) we get

$$\begin{aligned}|Z_t|^2 =& \sum_{i=1}^d \int_0^t 2Z_t^i\big(b_i(\hat{X}_s^1, \hat{\mu}_s^1) - b_i(\hat{X}_s^2, \hat{\mu}_s^2)\big)\, ds \\ &+ \frac{1}{2}\sum_{i,k=1}^d \int_0^t 2\big(\sigma_{i,k}(\hat{X}_s^1, \hat{\mu}_s^1) - \sigma_{i,k}(\hat{X}_s^1, \hat{\mu}_s^1)\big)^2\, ds \\ &+ \sum_{i,j=1}^d \int_0^t 2Z_t^i\big(\sigma_{i,j}(\hat{X}_s^1, \hat{\mu}_s^1) - \sigma_{i,j}(\hat{X}_s^1, \hat{\mu}_s^1)\big)\, d\hat{B}_s^i\end{aligned}$$

$$
\begin{aligned}
&= \int_0^t 2\langle Z_s, b(\hat{X}_s^1, \hat{\mu}_s^1) - b(\hat{X}_s^2, \hat{\mu}_s^2)\rangle \mathrm{ds} + \int_0^t \|\sigma(\hat{X}_s^1, \hat{\mu}_s^1) - \sigma(\hat{X}_s^2, \hat{\mu}_s^2)\|^2\, ds \\
&\quad + \int_0^t 2\langle Z_s, (\sigma(\hat{X}_s^1, \hat{\mu}_s^1) - \sigma(\hat{X}_s^2, \hat{\mu}_s^2))d\hat{B}_s\rangle.
\end{aligned}
$$

Since $\int_0^t 2\langle Z_s, (\sigma(\hat{X}_s^1, \hat{\mu}_s^1) - \sigma(\hat{X}_s^2, \hat{\mu}_s^2))d\hat{B}_s\rangle$ is a local martingale, there exits an increasing sequence of stopping times $(\tau_n)_{n=1}^{\infty}$, $\lim_{n\to\infty} \tau_n = \infty$, such that

$$
\int_0^{t\wedge\tau_n} 2\langle Z_s, (\sigma(\hat{X}_s^1, \hat{\mu}_s^1) - \sigma(\hat{X}_s^2, \hat{\mu}_s^2))d\hat{B}_s\rangle
$$

is a martingale for all $n \in \mathbb{N}$. Thus is holds that

$$
\hat{\mathbb{E}} \int_0^{t\wedge\tau_n} 2\langle Z_s, (\sigma(\hat{X}_s^1, \hat{\mu}_s^1) - \sigma(\hat{X}_s^2, \hat{\mu}_s^2))d\hat{B}_s\rangle = 0.
$$

Now, by $(\mathbf{H}_2)$ we get that

$$
\begin{aligned}
\hat{\mathbb{E}}|Z_{t\wedge\tau_n}|^2 &= \hat{\mathbb{E}} \int_0^{t\wedge\tau_n} 2\langle Z_s, b(\hat{X}_s^1, \hat{\mu}_s^1) - b(\hat{X}_s^2, \hat{\mu}_s^2)\rangle \mathrm{ds} + \|\sigma(\hat{X}_s^1, \hat{\mu}_s^1) - \sigma(\hat{X}_s^2, \hat{\mu}_s^2)\|^2\, ds \\
&\le \hat{\mathbb{E}} \int_0^{t\wedge\tau_n} L_2\big(\kappa_1(|Z_s|^2) + \kappa_2(\rho^2(\hat{\mu}_s^1, \hat{\mu}_s^2))\big) ds.
\end{aligned}
$$

Since $a_s := L_2\big(\kappa_1(|Z_s|^2) + \kappa_2(\rho^2(\hat{\mu}_s^1, \hat{\mu}_s^2))\big) \ge 0$ for all $s \ge 0$ and $\tau_n \le \tau_{n+1}$ for all $n \in \mathbb{N}$, then by monotone convergence (Theorem 1.5) we have

$$
\begin{aligned}
\limsup_{n\to\infty} \hat{\mathbb{E}}|Z_{t\wedge\tau_n}|^2 &\le \limsup_{n\to\infty} \hat{\mathbb{E}} \int_0^{t\wedge\tau_n} a_s\, ds = \lim_{n\to\infty} \hat{\mathbb{E}} \int_0^{t\wedge\tau_n} a_s\, ds \\
&= \hat{\mathbb{E}} \lim_{n\to\infty} \int_0^{t\wedge\tau_n} a_s\, ds = \hat{\mathbb{E}} \int_0^t a_s\, ds.
\end{aligned}
$$

Furthermore, by Fatou's Lemma (Theorem 1.39) it holds that

$$
\hat{\mathbb{E}}|Z_t|^2 = \hat{\mathbb{E}} \liminf_{n\to\infty} |Z_{t\wedge\tau_n}|^2 \le \liminf_{n\to\infty} \hat{\mathbb{E}}|Z_{t\wedge\tau_n}|^2 \le \limsup_{n\to\infty} \hat{\mathbb{E}}|Z_{t\wedge\tau_n}|^2 \le \hat{\mathbb{E}} \int_0^t a_s\, ds.
$$

Set $G_t := \hat{\mathbb{E}}|Z_t|^2$. From Lemma 2.11 from section 2.3. and Jensen's inequality (Theorem 1.40) it follows that

$$
\rho^2(\hat{\mu}_s^1, \hat{\mu}_s^2) \le \Big(\hat{\mathbb{E}}\big|\hat{X}_s^1 - \hat{X}_s^2\big|\Big)^2 \le \hat{\mathbb{E}}|Z_s|^2 = G_s.
$$

Set $\tilde{\kappa}_i(u) = L_2\, \kappa_i(u)$, $i = 1, 2$. Then also $\tilde{\kappa}_i(u)$ satisfies the conditions in $(\mathbf{H}_2)$ and thus $\tilde{\kappa}_2(\rho^2(\hat{\mu}_s^1, \hat{\mu}_s^2)) \le \tilde{\kappa}_2(G_s)$.

Since $\tilde{\kappa}_1$ is concave, we have by Jensen's inequality (Proposition 1.40) and Tonelli's Theorem (Theorem 1.14) that

$$G_t \leq \hat{\mathbb{E}} \int_0^t \tilde{\kappa}_1(|Z_s|^2) + \tilde{\kappa}_2(G_s)\, ds \leq \int_0^t \tilde{\kappa}_1(\hat{\mathbb{E}}|Z_s|^2) + \tilde{\kappa}_2(G_s)\, ds$$
$$= \int_0^t \tilde{\kappa}_1(G_s) + \tilde{\kappa}_2(G_s)\, ds.$$

Now by Lemma 2.4 we get that $G_t = \hat{\mathbb{E}}|Z_t|^2 = 0$. Moreover, it follows that $Z_t = \hat{X}_t^1 - \hat{X}_t^2 = 0$ for all $t \geq 0$ a.s. and by Definition 1.24 we have that pathwise uniqueness holds for Eq.(1.1). ☐

Theorem 2.6. *Suppose that for each given initial distribution there exists a solution to Eq.* (1.1) *and that pathwise uniqueness holds. Then Eq.* (1.1) *has a unique strong solution.*

We will not prove Theorem 2.6 rigorously in this thesis but we will provide a brief outline of the proof found in [**17**] in the proof for Theorem IV.1.1. and [**20**] p. 391 (solution to 3.22.):

Definition IV.1.6. in [**17**] gives us an alternative definition for strong solutions by using the existence of a function F with certain measurability conditions for which $X = F(X(0), B)$ a.s., where X is a solution to Eq. (1.1) with a Brownian motion B. Furthermore, we can consider strong solutions as the function F which gives us the solution to Eq. (1.1) when plugging in an initial value and a Brownian motion.

We assume (X, B) and (X', B') to be two weak solutions to Eq. (1.1) which can be defined on different probability spaces. We then construct a new probability space such that there are two solutions (w_1, w_3) and (w_2, w_3) defined on that space and (w_1, w_3) and (X, B) have the same distribution as well as (w_2, w_3) and (X', B').

Now pathwise uniqueness implies that $Q^W \times Q'^W(w_1 = w_2) = 1$ P^W-a.s., where Q^W and Q'^W are regular conditional probabilities given W (defined in [**17**]) and P^W is the probability law of the Brownian motion. Using [**20**] p. 391 we obtain the existence of a function $F_x(w)$ determined uniquely up to measure 0 such that $Q^W = Q'^W = \delta_{(F_x(w))}$ P^W-a.s. Lastly we can see that the function $F_x(w)$ has the desired measurability and that any solution of Eq. (1.1) with the given initial value can be produced by $F_x(w)$ a.s.

Finally, Propositions 2.3 and 2.5 and Theorem 2.6 give us the desired result:

Theorem 2.7. *Provided that* $(\mathbf{H}_1)$ *and* $(\mathbf{H}_2)$ *hold and* $\mathbb{E}|\xi|^{2p} < \infty$ *for any* $p > 1$, *it follows that there exists a unique strong solution for Eq.* (1.1).

2.3. Some Techincal Results

In this section we provide some more techincal results and helpful calculations used above:

Lemma 2.8.

$$\Big|\int_s^t b(X_u,\mu_u)\,du\Big|^{2p} \le (t-s)^{2p-1}\int_s^t |b(X_u,\mu_u)|^{2p}\,du.$$

Proof. By Hölder's inequality (Theorem 1.32) it holds that

$$\begin{aligned}
\Big|\int_s^t b(X_u,\mu_u)\mathrm{d}u\Big|^{2p} &= \Big(\sum_{i=1}^d \Big(\int_s^t b_i(X_u,\mu_u)\mathrm{d}u\Big)^2\Big)^p \\
&\le \Big(\sum_{i=1}^d \Big(\Big(\int_s^t (b_i(X_u,\mu_u))^2\mathrm{d}u\Big)^{\frac{1}{2}}\Big(\int_s^t 1\mathrm{d}u\Big)^{\frac{1}{2}}\Big)^2\Big)^p \\
&= \Big((t-s)\int_s^t \sum_{i=1}^d (b_i(X_u,\mu_u))^2\mathrm{d}u\Big)^p \\
&= \Big((t-s)\int_s^t |b(X_u,\mu_u)|^2\mathrm{d}u\Big)^p \\
&\le \Big((t-s)\Big(\int_s^t |b(X_u,\mu_u)|^{2p}\mathrm{d}u\Big)^{\frac{1}{p}}\Big(\int_s^t 1\mathrm{d}u\Big)^{\frac{p-1}{p}}\Big)^p \\
&= (t-s)^{2p-1}\int_s^t |b(X_u,\mu_u)|^{2p}\mathrm{d}u. \qquad (2.12)
\end{aligned}$$

□

Lemma 2.9.

$$\mathbb{E}\Big|\int_s^t \sigma(X_u,\mu_u)\,dB_u\Big|^{2p} \le C_{2p}(t-s)^{p-1}\mathbb{E}\Big(\int_s^t \|\sigma(X_u,\mu_u)\|^{2p}\,du\Big).$$

Proof. Using Theorems 1.36 and 1.32 we get

$$\begin{aligned}
\mathbb{E}\Big|\int_s^t \sigma(X_u,\mu_u)\mathrm{d}B_u\Big|^{2p} &\le \mathbb{E}\Big(\sup_{r\in[s,t]}\Big|\int_s^r \sigma(X_u,\mu_u)\mathrm{d}B_u\Big|^{2p}\Big) \\
&\le C_{2p}\mathbb{E}\Big(\int_s^t \|\sigma(X_u,\mu_u)\|^2\mathrm{d}u\Big)^p \\
&\le C_{2p}\mathbb{E}\Big(\Big(\int_s^t \|\sigma(X_u,\mu_u)\|^{2p}\mathrm{d}u\Big)^{\frac{1}{p}}\Big(\int_s^t 1\mathrm{d}u\Big)^{\frac{p-1}{p}}\Big)^p \\
&= C_{2p}(t-s)^{p-1}\mathbb{E}\Big(\int_s^t \|\sigma(X_u,\mu_u)\|^{2p}\mathrm{d}u\Big). \qquad (2.13)
\end{aligned}$$

□

LEMMA 2.10. *Assume that* $(\mathbf{H}_1)$ *holds and* $\mu_s := \mathbb{P} \circ (X_s)^{-1}$. *Then*

$$|b(X_s, \mu_s)|^{2p} + \|\sigma(X_s, \mu_s)\|^{2p} \leq 3^{p-1} L_1^p (1 + |X_s|^{2p} + 2^{2p-1}\mathbb{E}(1 + |X_s|^{2p})).$$

PROOF. By Hölder's inequality (Theorem 1.32) we have

$$\begin{aligned}
\|\mu_s\|_{\lambda^2}^{2p} &= \Big(\int_{\mathbb{R}^d} (1 + |x|)^2 \mathrm{d}\mu_s(x) \Big)^p \\
&\leq \Big(\Big(\int_{\mathbb{R}^d} (1 + |x|)^{2p} \mathrm{d}\mu_s(x) \Big)^{\frac{1}{p}} \Big(\int_{\mathbb{R}^d} 1^{\frac{p}{p-1}} \mathrm{d}\mu_s(x) \Big)^{\frac{p-1}{p}} \Big)^p \\
&= \int_{\mathbb{R}^d} (1 + |x|)^{2p} \mathrm{d}\mu_s(x) \\
&= \int_{\mathbb{R}^d} (1 + |x|)^{2p} \mathrm{d}(\mathbb{P} \circ X_s^{-1})(x) \\
&= \mathbb{E}(1 + |X_s|)^{2p} \\
&\leq \mathbb{E}((1 + |X_s|^{2p})^{\frac{1}{2p}} \cdot 2^{\frac{2p-1}{2p}})^{2p} \\
&= 2^{2p-1}\mathbb{E}(1 + |X_s|^{2p})
\end{aligned} \tag{2.14}$$

and thus we get

$$\begin{aligned}
|b(X_s, \mu_s)|^{2p} + \|\sigma(X_s, \mu_s)\|^{2p} &\leq \big(|b(X_s, \mu_s)|^2 + \|\sigma(X_s, \mu_s)\|^2\big)^p \\
&\leq \big(L_1(1 + |X_s|^2 + \|\mu\|_{\lambda^2}^2)\big)^p \\
&\leq \big(L_1(1 + |X_s|^{2p} + \|\mu\|_{\lambda^2}^{2p})^{\frac{1}{p}} \, 3^{\frac{p-1}{p}}\big)^p \\
&= 3^{p-1} L_1^p (1 + |X_s|^{2p} + \|\mu\|_{\lambda^2}^{2p}) \\
&\leq 3^{p-1} L_1^p (1 + |X_s|^{2p} + 2^{2p-1}\mathbb{E}(1 + |X_s|^{2p})).
\end{aligned} \tag{2.15}$$

□

LEMMA 2.11. *Let* $\mu_t^1 := \mathbb{P} \circ (X_t^1)^{-1}$ *and* $\mu_t^2 := \mathbb{P} \circ (X_t^2)^{-1}$. *Then it holds that*

$$\rho(\mu_t^1, \mu_t^2) \leq \mathbb{E}\big|X_t^1 - X_t^2\big|.$$

PROOF. The condition $\|\varphi\|_{C_\rho(\mathbb{R}^d)} \leq 1$ gives that

$$\sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{|x - y|} \leq \|\varphi\|_{C_\rho(\mathbb{R}^d)} \leq 1,$$

so it follows that $|\varphi(x) - \varphi(y)| \leq |x - y|$ for $x, y \in \mathbb{R}^d$, $x \neq y$.

Using the definition of ρ and the previous results we get

$$\begin{aligned}\rho(\mu_t^1,\mu_t^2) &= \sup_{\|\varphi\|_{C_\rho(\mathbb{R}^d)}\le 1}\Big|\int_{\mathbb{R}^d}\varphi(x)\mu_t^1(\mathrm{d}x)-\int_{\mathbb{R}^d}\varphi(x)\mu_t^2(\mathrm{d}x)\Big| \\ &= \sup_{\|\varphi\|_{C_\rho(\mathbb{R}^d)}\le 1}\Big|\mathbb{E}\varphi(X_t^1)-\mathbb{E}\varphi(X_t^2)\Big| \\ &\le \sup_{\|\varphi\|_{C_\rho(\mathbb{R}^d)}\le 1}\mathbb{E}\Big|\varphi(X_t^1)-\varphi(X_t^2)\Big| \\ &\le \mathbb{E}\sup_{\|\varphi\|_{C_\rho(\mathbb{R}^d)}\le 1}\Big|\varphi(X_t^1)-\varphi(X_t^1)\Big| \\ &\le \mathbb{E}\Big|X_t^1-X_t^2\Big|. \qquad (2.16)\end{aligned}$$

□

LEMMA 2.12. *For all* $0<\delta<2$ *there exists a constant* $c_\delta>0$ *such that*

$$x^2(\log x)^2 \le c_\delta x^{2-\delta} \qquad \textit{for all } x\in(0,1].$$

PROOF. Put $c_\delta := \big(\frac{2}{\delta e}\big)^2$, $0<\delta<2$. We show that $x^\delta(\log x)^2\le c_\delta$ for all $x\in(0,1]$:

Looking at the first and second derivatives of $x^\delta(\log x)^2$ we can see that it reaches its maximum at $x=e^{-\frac{2}{\delta}}$. Thus we get

$$x^\delta(\log x)^2 \le \big(e^{-\frac{2}{\delta}}\big)^\delta\big(\log e^{-\frac{2}{\delta}}\big)^2 = e^{-2}\big(-\frac{2}{\delta}\big)^2 = \big(\frac{2}{\delta e}\big)^2 = c_\delta.$$

Moreover, it holds that $(\log x)^2\le c_\delta x^{-\delta}$, so the statement follows from multiplying both sides by x^2. □

In the following we will show how to confirm the martingale property with respect to $(\bar{\mathscr{W}}_t)_{t\in[0,T]}$ defined in (1.2) using continuous, bounded and measurable functionals as test functions instead of indicator functions of measurable sets. We will do this in two steps: first we show the statement for a $(\mathscr{B}(\mathcal{W}_t))_{t\in[0,T]}$-martingale. Then we will consider a right-continous filtration to help prove the final statement.

LEMMA 2.13. *Let* $(M_t)_{t\in[0,T]}$ *be a* $(\mathscr{B}(\mathcal{W}_t))_{t\in[0,T]}$*-adapted, continuous and integrable process. If for a continuous, bounded and* $\mathscr{B}(\mathcal{W}_s)$*-measurable functional* G_s *it holds that*

$$\mathbb{E}\big(M_tG_s\big)=\mathbb{E}\big(M_sG_s\big)$$

for $0\le s<t\le T$*, then* $(M_t)_{t\in[0,T]}$ *is a* $(\mathscr{B}(\mathcal{W}_t))_{t\in[0,T]}$*-martingale.*

Proof. By [**28**] Chapter 1.3 (p. 30) we have that for $X_u(w) = w_u$, $w \in \mathcal{W}_t$ it holds that $\mathscr{B}(\mathcal{W}_t) = \sigma(X_u, 0 \leq u \leq t)$.

Define a π-system $\mathcal{P}$ as $\mathcal{P} := \big\{\{X_{t_1} \in (a_1, b_1), \dots, X_{t_n} \in (a_n, b_n)\}$ where $, n \in \mathbb{N}$, $-\infty \leq a_i < b_i \leq \infty$, $i = 1, ..., n$ and $0 \leq t_1 < ... < t_n \leq s\big\}$.

Now we have that $\mathscr{B}(\mathcal{W}_t) = \sigma(\mathcal{P})$.

Set μ and ν to be finite measures on $\mathscr{B}(\mathcal{W}_t)$ defined as follows:

$$\mu\Big(\bigcap_{i=1}^{n}\{X_{t_i} \in (a_i, b_i)\}\Big) := \mathbb{E}\Big(M_t \prod_{i=1}^{n} \mathbb{1}_{\{X_{t_i} \in (a_i,b_i)\}}\Big)$$
$$\nu\Big(\bigcap_{i=1}^{n}\{X_{t_i} \in (a_i, b_i)\}\Big) := \mathbb{E}\Big(M_s \prod_{i=1}^{n} \mathbb{1}_{\{X_{t_i} \in (a_i,b_i)\}}\Big).$$

Let $\{\varphi_k^i\}_{k\in\mathbb{N}}$ be a family of continuous functions such that $\lim\limits_{k\to\infty} \varphi_k^i = \mathbb{1}_{(a_i,b_i)}$ for $i = 1, ..., n$.

Since $w \mapsto (w_{t_1}, ..., w_{t_n})$ and $(y_1, ..., y_n) \mapsto \prod_{i=1}^n \varphi_k^i(y_i)$ are continuous maps it follows from Proposition 1.4 that $w \mapsto \prod_{i=1}^n \varphi_k^i(w_{t_i})$ is also continuous.

Therefore by our assumption it holds that

$$\mathbb{E}\Big(M_t \prod_{i=1}^{n} \varphi_k^i(w_{t_i})\Big) = \mathbb{E}\Big(M_s \prod_{i=1}^{n} \varphi_k^i(w_{t_i})\Big)$$

for all $k \in \mathbb{N}$.

Furthermore, by Theorem 1.6 we get that

$$\lim_{k\to\infty} \mathbb{E}\Big(M_t \prod_{i=1}^{n} \varphi_k^i(w_{t_i})\Big) = \mathbb{E}\Big(M_t \prod_{i=1}^{n} \mathbb{1}_{(a_i,b_i)}(w_{t_i})\Big),$$
$$\lim_{k\to\infty} \mathbb{E}\Big(M_s \prod_{i=1}^{n} \varphi_k^i(w_{t_i})\Big) = \mathbb{E}\Big(M_s \prod_{i=1}^{n} \mathbb{1}_{(a_i,b_i)}(w_{t_i})\Big),$$

and therefore it follows that

$$\mathbb{E}\Big(M_t \prod_{i=1}^{n} \mathbb{1}_{(a_i,b_i)}(w_{t_i})\Big) = \mathbb{E}\Big(M_s \prod_{i=1}^{n} \mathbb{1}_{(a_i,b_i)}(w_{t_i})\Big).$$

Now we have that $\mu(A) = \nu(A)$ for all $A \in \mathcal{P}$. Then by Theorem 1.16 it holds that $\mu(B) = \nu(B)$ for all $B \in \mathscr{B}(\mathcal{W}_s)$ and hence we conclude that $(M_t)_{t\in[0,T]}$ is a martingale with respect to $(\mathscr{B}(\mathcal{W}_t))_{t\in[0,T]}$. □

Lemma 2.14. *Assume that $(M_t)_{t\geq 0}$ is a continuous $(\mathcal{F}_t)_{t\geq 0}$-martingale. Then it is also a martingale with respect to $(\mathcal{G}_t)_{t\geq 0}$ where $\mathcal{G}_t := \bigcap\limits_{s>t} \mathcal{F}_s$.*

PROOF. Set $X_{\frac{1}{n}} := M_{s+\frac{1}{n}}$, $n \in \mathbb{N}$ and $X := M_{s+1}$. Put $\mathcal{A} := \{\frac{1}{n} : n \in \mathbb{N}\}$ and $\mathcal{H}_r := \mathcal{F}_{s+r}$ for $r \in \mathcal{A}$.

Now $\lim_{n\to\infty} X_{\frac{1}{n}} = M_s$ and since $(M_t)_{t\geq 0}$ is an $(\mathcal{F}_t)_{t\geq 0}$-martingale one has that

$$\mathbb{E}[X_{\frac{1}{n}}|\mathcal{H}_{\frac{1}{n+1}}] = \mathbb{E}[M_{s+\frac{1}{n}}|\mathcal{F}_{s+\frac{1}{n+1}}] = M_{s+\frac{1}{n+1}} = X_{\frac{1}{n+1}} \quad a.s.$$

for all $n \in \mathbb{N}$. Thus $(X_r)_{r\in\mathcal{A}}$ is a $(\mathcal{H}_r)_{r\in\mathcal{A}}$-martingale.

Furthermore, we have

$$\mathbb{E}[X|\mathcal{H}_{\frac{1}{n}}] = \mathbb{E}[M_{s+1}|\mathcal{F}_{s+\frac{1}{n}}] = M_{s+\frac{1}{n}} = X_{\frac{1}{n}} \quad a.s.$$

for all $n \in \mathbb{N}$, so $(X_r)_{r\in\mathcal{A}}$ is a closable martingale. Then by Theorem 1.31 it holds that $(X_r)_{r\in\mathcal{A}}$ is uniformly integrable and from Lemma 1.13 it follows that $\lim_{n\to\infty} \mathbb{E}M_{s+\frac{1}{n}} = \mathbb{E}M_s$.

Let $0 \leq s < t$, $A \in \mathcal{G}_s$ and $\varepsilon \in (0, t-s)$. Then $A \in \mathcal{F}_{s+\varepsilon}$ for all $\varepsilon > 0$, and since $(M_t)_{t\geq 0}$ is a $(\mathcal{F}_t)_{t\geq 0}$-martingale it holds that

$$\mathbb{E}(M_t \mathbb{1}_A) = \mathbb{E}((M_t - M_{s+\varepsilon})\mathbb{1}_A) + \mathbb{E}(M_{s+\varepsilon}\mathbb{1}_A) = 0 + \mathbb{E}(M_{s+\varepsilon}\mathbb{1}_A)$$

for all $\varepsilon \in (0, t-s)$.

Since the above holds for all $\varepsilon \in (0, t-s)$ we can consider a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$ with $\varepsilon_n = \min\{\frac{1}{n}, t-s-\delta\}$ for any $\delta \in (0, t-s-\varepsilon)$. It holds that $\lim_{n\to\infty} \varepsilon_n = 0$.

Now we get

$$\mathbb{E}(M_t \mathbb{1}_A) = \lim_{n\to\infty} \mathbb{E}(M_{s+\varepsilon_n}\mathbb{1}_A) = \mathbb{E}(M_s \mathbb{1}_A)$$

for all $A \in \mathcal{G}_s$. □

With Lemmas 2.13 and 2.14 we obtain the following result:

LEMMA 2.15. *Let $(M_t)_{t\in[0,T]}$ be a $(\bar{\mathscr{W}}_t)_{t\in[0,T]}$-adapted continuous process. If for any continuous, bounded and $\bar{\mathscr{W}}_s$-measurable functional G_s we have that*

$$\mathbb{E}(M_t G_s) = \mathbb{E}(M_s G_s)$$

for $0 \leq s < t \leq T$, then $(M_t)_{t\in[0,T]}$ is a martingale with respect to $(\bar{\mathscr{W}}_t)_{t\in[0,T]}$.

CHAPTER 3

The Convergence Rate

In chapter 2 we proved the existence of a unique strong solution for Equation (1.1) by using the Euler-Maruyama Approximation introduced in (2.3). In this chapter we will consider the convergence rate of this approximation.

Definition 3.1 ("Big-O notation", see [**16**] (5.42) p. 340). Let f and g be real-valued functions such that there exists a $C > 0$ and $a \in \mathbb{R}$ with

$$|f(x)| \le C|g(x)| \qquad \text{for all } x > a.$$

We denote this by

$$f(x) = O(g(x)).$$

Lemma 3.2. *Let κ_η, $0 < \eta < \frac{1}{e}$, be defined as in* (1.5)*. Then it holds that*

$$x + \kappa_\eta(x) \le C_\eta \kappa_\eta(x)$$

for $x \ge 0$, where C_η is a constant depending on η.

Proof. Consider $0 < x \le \eta$. Then

$$\log x^{-1} \ge \log \eta^{-1} \ge \log \Big(\frac{1}{e}\Big)^{-1} = \log e = 1,$$

and hence

$$x + \kappa_\eta(x) = x + x\log x^{-1} \le x\log x^{-1} + x\log x^{-1} = 2x\log x^{-1} = 2\kappa_\eta(x).$$

Next assume that $x > \eta$. Then we have

$$\begin{aligned} x = \eta + x - \eta &\le \eta\log\eta^{-1} + x - \eta \\ &\le \eta\log\eta^{-1} + x - \eta + (\log\eta^{-1} - 1)(x-\eta) + \frac{\eta\log\eta^{-1}}{\log\eta^{-1} - 1} \\ &= \Big(\frac{1}{\log\eta^{-1} - 1} + 1\Big)\big(\eta\log\eta^{-1} + (\log\eta^{-1} - 1)(x-\eta)\big) \\ &= \Big(\frac{1}{\log\eta^{-1} - 1} + 1\Big)\kappa_\eta(x), \end{aligned}$$

and thus we get

$$x + \kappa_\eta(x) \le \Big(\frac{1}{\log\eta^{-1} - 1} + 1\Big)\kappa_\eta(x) + \kappa_\eta(x) \le \hat{C}_\eta\kappa_\eta(x).$$

Now for $x > 0$ it holds that $x + \kappa_\eta(x) \le C_\eta\kappa_\eta(x)$, where $C_\eta := \hat{C}_\eta \vee 2$. □

The following statement regarding the convergence rate of the approximation is a slightly different formulation of that found in [**10**] Section 4:

THEOREM 3.3. *Consider X_t from Eq.* (1.1) *and the Euler-Maruyama Approximation from* (2.3). *Assume that* ($\mathbf{H}_1$) *and* ($\mathbf{H}_2'$) *hold for b and σ and that $\mathbb{E}|\xi|^{2p} < \infty$ for any $p > 1$.*
Then it follows that there exists a $T_0 > 0$ such that

$$\mathbb{E}\Big(\sup_{t\in[0,T]} |X_t^n - X_t|^2\Big) = O((2^{-n}T_0)^{1-\gamma})$$

for any $\gamma \in (0, 1)$.

PROOF. Put $H_t := X_t^n - X_t$. Then we have

$$H_t = \int_0^t \big(b(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - b(X_s, \mu_s)\big)ds + \int_0^t \Big(\sigma(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - \sigma(X_s, \mu_s)\Big)dB_s.$$

By a similar deduction as in Proposition 2.5 we get from Ito's formula (Theorem 1.27) that

$$|H_t|^2 = J_1(t) + J_2(t) + J_3(t),$$

where

$$J_1(t) := \int_0^t 2\langle H_s, b(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - b(X_s, \mu_s)\rangle ds,$$

$$J_2(t) := \int_0^t 2\langle H_s, (\sigma(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - \sigma(X_s, \mu_s))dB_s\rangle,$$

$$J_3(t) := \int_0^t \|\sigma(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - \sigma(X_s, \mu_s)\|^2 ds.$$

We first take a look at $J_1(t)$. By ($\mathbf{H}_2'$) we have

$$\begin{aligned}
|J_1(t)| =& \Big|\int_0^t 2\langle H_s, b(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - b(X_s, \mu_s)\rangle ds\Big| \\
\leq& 2\int_0^t |H_s||b(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - b(X_s, \mu_s)|ds \\
=& 2\int_0^t |H_s||b(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - b(X_s^n, \mu_s^n) + b(X_s^n, \mu_s^n) - b(X_s, \mu_s)|ds \\
\leq& 2\int_0^t |H_s||b(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - b(X_s^n, \mu_s^n)| + |H_s||b(X_s^n, \mu_s^n) - b(X_s, \mu_s)|ds \\
\leq& \int_0^t |H_s|^2 + |b(X_{t_n(s)}^n, \mu_{t_n(s)}^n) - b(X_s^n, \mu_s^n)|^2 ds \\
&+ 2\int_0^t \lambda_1\big(|H_s|^2\gamma_1(|H_s|) + |H_s|\rho(\mu_s^n, \mu_s)\big)ds.
\end{aligned}$$

With $(\mathbf{H}_2')$ and Hölder's inequality (Remark 1.33) we get

$$|J_1(t)| \leq \int_0^t |H_s|^2 ds + 2\int_0^t \lambda_1^2\big(|X^n_{t_n(s)} - X^n_s|^2\gamma_1^2(|X^n_{t_n(s)} - X^n_s|) + \rho^2(\mu^n_{t_n(s)}, \mu^n_s)\big)ds$$
$$+ 2\int_0^t \lambda_1|H_s|^2\gamma_1(|H_s|)ds + \lambda_1\int_0^t |H_s|^2 ds + \lambda_1\int_0^t \rho^2(\mu^n_s, \mu_s)ds.$$

Furthermore, by Remark 1.42 and (2.16) we get for some constant $c > 0$

$$|J_1(t)| \leq (1+\lambda_1)\int_0^t |H_s|^2 ds + 2\lambda_1^2 c\int_0^t \kappa_\eta^2(|X^n_{t_n(s)} - X^n_s|)ds$$
$$+ 2\lambda_1^2\int_0^t |X^n_{t_n(s)} - X^n_s|^2 ds + 2\lambda_1 c\int_0^t \kappa_\eta(|H_s|^2)ds + \lambda_1\int_0^t |X^n_s - X_s|^2 ds.$$

Next we want to consider the supremum. Since all the integrands are positive and κ_η is increasing we have for $0 \leq t \leq T$ that

$$\sup_{s\in[0,t]}|J_1(s)| \leq (1+2\lambda_1)\sup_{s\in[0,t]}\int_0^s |H_r|^2 dr + 2\lambda_1^2 c\sup_{s\in[0,t]}\int_0^s \kappa_\eta^2(|X^n_{t_n(r)} - X^n_r|)dr$$
$$+ 2\lambda_1^2\sup_{s\in[0,t]}\int_0^s |X^n_{t_n(r)} - X^n_r|^2 dr + 2\lambda_1 c\sup_{s\in[0,t]}\int_0^s \kappa_\eta(|H_r|^2)dr$$
$$\leq (1+2\lambda_1)\int_0^t \sup_{s\in[0,r]}|H_s|^2 dr + 2\lambda_1^2 c\int_0^t \kappa_\eta^2(\sup_{s\in[0,r]}|X^n_{t_n(s)} - X^n_s|)dr$$
$$+ 2\lambda_1^2\int_0^t \sup_{s\in[0,r]}|X^n_{t_n(s)} - X^n_s|^2 dr + 2\lambda_1 c\int_0^t \kappa_\eta(\sup_{s\in[0,r]}|H_s|^2)dr.$$

By linearity of expectation and Tonelli's Theorem (Theorem 1.14) we now take the expectation and obtain

$$\mathbb{E}\Big(\sup_{s\in[0,t]}|J_1(s)|\Big) \leq (1+2\lambda_1)\int_0^t \mathbb{E}\big(\sup_{s\in[0,r]}|H_s|^2\big)dr + 2\lambda_1^2 c\int_0^t \mathbb{E}\big(\kappa_\eta^2(\sup_{s\in[0,r]}|X^n_{t_n(s)} - X^n_s|)\big)dr$$
$$+ 2\lambda_1^2\int_0^t \mathbb{E}\big(\sup_{s\in[0,r]}|X^n_{t_n(s)} - X^n_s|^2\big)dr + 2\lambda_1 c\int_0^t \mathbb{E}\big(\kappa_\eta(\sup_{s\in[0,r]}|H_s|^2)\big)dr.$$

It is useful to notice that the map $x \mapsto \kappa_\eta^2(\sqrt{x})$ is concave:
For the case $0 < x \leq \eta$ we have

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\kappa_\eta^2(\sqrt{x}) = \frac{\mathrm{d}^2}{\mathrm{d}x^2}x\big(\log(x^{-\frac{1}{2}})\big)^2 = \frac{\mathrm{d}^2}{\mathrm{d}x^2}\frac{x}{4}(\log x)^2 = \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{1}{4}\big(\log x\big)^2 + \frac{x}{2x}\log x\Big)$$
$$= \frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{1}{4}\big(\log x\big)^2 + \frac{1}{2}\log x\Big) = \frac{1}{2x}\log x + \frac{1}{2x} = \frac{1}{2x}\big(\log x + 1\big) < 0$$

for all $0 < x < \frac{1}{e}$.

Similarly, it holds for all $x > \eta$ that

$$\begin{aligned}\frac{\mathrm{d}^2}{\mathrm{d}x^2}\,\kappa_\eta^2(\sqrt{x}) &= \frac{\mathrm{d}^2}{\mathrm{d}x^2}\,\big((\log\eta^{-1}-1)\sqrt{x}+\eta\big)^2\\ &= \frac{\mathrm{d}^2}{\mathrm{d}x^2}\,(\log\eta^{-1}-1)^2x+2(\log\eta^{-1}-1)\eta\sqrt{x}+\eta^2\\ &= \frac{\mathrm{d}}{\mathrm{d}x}(\log\eta^{-1}-1)^2+(\log\eta^{-1}-1)\eta\frac{1}{\sqrt{x}}\\ &= -\frac{1}{2}(\log\eta^{-1}-1)\eta\frac{1}{x\sqrt{x}} < 0.\end{aligned}$$

Now by Jensen's inequality (Proposition 1.40) we have for $Y := \sup_{s\in[0,r]}|X^n_{t_n(s)} - X^n_s|$ that

$$\mathbb{E}\kappa_\eta^2(Y) = \mathbb{E}\kappa_\eta^2(\sqrt{Y^2}) \le \kappa_\eta^2(\sqrt{\mathbb{E}Y^2}).$$

Since κ_η is also concave, by Jensen's inequality we get

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\in[0,t]}|J_1(s)|\Big) \le{}& C_1\int_0^t \mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big)dr\\ &+ C_1\int_0^t \kappa_\eta^2\Big(\Big(\mathbb{E}\Big(\sup_{s\in[0,r]}|X^n_{t_n(s)}-X^n_s|^2\Big)\Big)^{\frac{1}{2}}\Big)dr\\ &+ C_1\int_0^t \mathbb{E}\Big(\sup_{s\in[0,r]}|X^n_{t_n(s)}-X^n_s|^2\Big)dr\\ &+ C_1\int_0^t \kappa_\eta\Big(\mathbb{E}(\sup_{s\in[0,r]}|H_s|^2)\Big)dr. \end{aligned}\tag{3.1}$$

Next we will consider the term $J_3(t)$. With $(\mathbf{H}_2')$ and Remark 1.42 we get

$$\begin{aligned}|J_3(t)| &= \int_0^t \|(\sigma(X^n_{t_n(s)},\mu^n_{t_n(s)}) - \sigma(X_s,\mu_s)\|^2ds\\ &= \int_0^t \|\sigma(X^n_{t_n(s)},\mu^n_{t_n(s)}) - \sigma(X^n_s,\mu^n_s) + \sigma(X^n_s,\mu^n_s) - \sigma(X_s,\mu_s)\|^2ds\\ &\le 2\int_0^t \|\sigma(X^n_{t_n(s)},\mu^n_{t_n(s)}) - \sigma(X^n_s,\mu^n_s)\|^2 + \|\sigma(X^n_s,\mu^n_s) - \sigma(X_s,\mu_s)\|^2ds\\ &\le 2\hat{\lambda}_2\int_0^t \kappa_\eta(|X^n_{t_n(s)}-X^n_s|^2)ds + 2\hat{\lambda}_2\int_0^t \rho^2(\mu^n_{t_n(s)},\mu^n_s)ds\\ &\quad + 2\hat{\lambda}_2\int_0^t \kappa_\eta(|X^n_s-X_s|^2)ds + 2\hat{\lambda}_2\int_0^t \rho^2(\mu^n_s,\mu_s)ds.\end{aligned}$$

Using similar calculations as for the term $J_1(t)$ we have

$$
\begin{aligned}
|J_3(t)| \leq & 2\hat{\lambda}_2 \int_0^t \kappa_\eta(|X^n_{t_n(s)} - X^n_s|^2)ds + 2\hat{\lambda}_2 \int_0^t |X^n_{t_n(s)} - X^n_s|^2 ds \\
& + 2\hat{\lambda}_2 \int_0^t \kappa_\eta(|H_s|^2)ds + 2\hat{\lambda}_2 \int_0^t |H_s|^2 ds,
\end{aligned}
$$

and since κ_η is increasing and all the integrands are positive we get for $0 \leq t \leq T$

$$
\begin{aligned}
\sup_{s\in[0,t]} |J_3(s)| \leq & C_3 \int_0^t \kappa_\eta(\sup_{s\in[0,r]} |X^n_{t_n(s)} - X^n_s|^2)dr + C_3 \int_0^t \sup_{s\in[0,r]} |X^n_{t_n(s)} - X^n_s|^2 dr \\
& + C_3 \int_0^t \kappa_\eta(\sup_{s\in[0,r]} |H_s|^2)dr + C_3 \int_0^t \sup_{s\in[0,r]} |H_s|^2 dr.
\end{aligned}
$$

Thus it follows from Tonelli's Theorem (Theorem 1.14) and Jensen's inequality (Proposition 1.40) that

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{s\in[0,t]} |J_3(s)|\Big) \leq & C_3 \int_0^t \mathbb{E}\Big(\sup_{s\in[0,r]} |H_s|^2\Big)dr \\
& + C_3 \int_0^t \kappa_\eta\Big(\mathbb{E}\Big(\sup_{s\in[0,r]} |X^n_{t_n(s)} - X^n_s|^2\Big)\Big)dr \\
& + C_3 \int_0^t \mathbb{E}\Big(\sup_{s\in[0,r]} |X^n_{t_n(s)} - X^n_s|^2\Big)dr \\
& + C_3 \int_0^t \kappa_\eta\Big(\mathbb{E}\Big(\sup_{s\in[0,r]} |H_s|^2\Big)\Big)dr. \qquad (3.2)
\end{aligned}
$$

Now we take a look at the term $J_2(t)$. By the Burkholder-Davis-Gundy inequality (Theorem 1.36) we have for $0 \leq t \leq T$ that

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{s\in[0,t]} |J_2(t)|\Big) = & 2\mathbb{E}\Big(\sup_{s\in[0,t]} \Big| \int_0^s \sum_{i,j=1}^d H^i_r\big(\sigma_{i,j}(X^n_{t_n(r)}, \mu^n_{t_n(r)}) - \sigma_{i,j}(X_r, \mu_r)\big)dW^i_r\Big|\Big) \\
\leq & 2 \sum_{i,j=1}^d \mathbb{E}\Big(\sup_{s\in[0,t]} \Big| \int_0^s H^i_r\big(\sigma_{i,j}(X^n_{t_n(r)}, \mu^n_{t_n(r)}) - \sigma_{i,j}(X_r, \mu_r)\big)dW^i_r\Big|\Big) \\
\leq & 2 \sum_{i,j=1}^d \tilde{C}\mathbb{E}\Big(\int_0^t \big|H^i_r\big(\sigma_{i,j}(X^n_{t_n(r)}, \mu^n_{t_n(r)}) - \sigma_{i,j}(X_r, \mu_r)\big)\big|^2 dr\Big)^{\frac{1}{2}} \\
= & 2\tilde{C}\,\mathbb{E}\Big(\sum_{i,j=1}^d \Big(\int_0^t \big|H^i_r\big(\sigma_{i,j}(X^n_{t_n(r)}, \mu^n_{t_n(r)}) - \sigma_{i,j}(X_r, \mu_r)\big)\big|^2 dr\Big)^{\frac{1}{2}}\Big).
\end{aligned}
$$

Hölder's inequality (Remark 1.33) gives us that

$$
\begin{aligned}
&\sum_{i=1}^{d}\sum_{j=1}^{d}\Big(\int_0^t \big|H_r^i\big(\sigma_{i,j}(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma_{i,j}(X_r,\mu_r)\big)\big|^2dr\Big)^{\frac{1}{2}}\\
&\quad\le \sum_{i=1}^{d}\left(\left(\sum_{j=1}^{d}\Big(\int_0^t \big|H_r^i\big(\sigma_{i,j}(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma_{i,j}(X_r,\mu_r)\big)\big|^2dr\Big)^{\frac{1}{2}\cdot 2}\right)^{\frac{1}{2}}\cdot d^{\frac{1}{2}}\right)\\
&\quad=\sqrt{d}\sum_{i=1}^{d}\left(\sum_{j=1}^{d}\int_0^t \big|H_r^i\big(\sigma_{i,j}(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma_{i,j}(X_r,\mu_r)\big)\big|^2dr\right)^{\frac{1}{2}}\\
&\quad\le\ \sqrt{d}\left(\sum_{i=1}^{d}\left(\sum_{j=1}^{d}\int_0^t \big|H_r^i\big(\sigma_{i,j}(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma_{i,j}(X_r,\mu_r)\big)\big|^2dr\right)^{\frac{1}{2}\cdot 2}\right)^{\frac{1}{2}}\cdot\sqrt{d}\\
&\quad=\ d\left(\int_0^t\sum_{i=1}^{d}\sum_{j=1}^{d}\big|H_r^i\big(\sigma_{i,j}(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma_{i,j}(X_r,\mu_r)\big)\big|^2dr\right)^{\frac{1}{2}}\\
&\quad=\ d\left(\int_0^t \|H_r\big(\sigma(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma(X_r,\mu_r)\big)\|^2dr\right)^{\frac{1}{2}}.
\end{aligned}
$$

Hence we have

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{s\in[0,t]}|J_2(s)|\Big)&\le\tilde{C}'\,\mathbb{E}\left(\int_0^t |H_r|^2\|\sigma(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma(X_r,\mu_r)\|^2dr\right)^{\frac{1}{2}}\\
&\le\tilde{C}'\,\mathbb{E}\left(\sup_{s\in[0,t]}|H_s|^2\int_0^t \|\sigma(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma(X_r,\mu_r)\|^2dr\right)^{\frac{1}{2}}\\
&=\mathbb{E}\left(\Big(\sup_{s\in[0,t]}|H_s|^2\Big)^{\frac{1}{2}}\Big((\tilde{C}')^2\int_0^t \|\sigma(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma(X_r,\mu_r)\|^2dr\Big)^{\frac{1}{2}}\right),
\end{aligned}
$$

and furthermore by Young's inequality (Proposition 1.34) it follows that

$$
\mathbb{E}\Big(\sup_{s\in[0,t]}|J_2(s)|\Big)\le\frac{1}{2}\mathbb{E}\Big(\sup_{s\in[0,t]}|H_s|^2\Big)+\ \frac{1}{2}(\tilde{C}')^2\,\mathbb{E}\Big(\int_0^t \|\sigma(X^n_{t_n(r)},\mu^n_{t_n(r)})-\sigma(X_r,\mu_r)\|^2dr\Big).
$$

For the latter term we use (3.2) and get

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\in[0,t]}|J_2(s)|\Big) \leq{}& \frac{1}{2}\,\mathbb{E}\Big(\sup_{s\in[0,t]}|H_s|^2\Big) + C_2\int_0^t \mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big)dr \\ &+ C_2\int_0^t \kappa_\eta\Big(\mathbb{E}\Big(\sup_{s\in[0,r]}|X^n_{t_n(s)}-X^n_s|^2\Big)\Big)dr \\ &+ C_2\int_0^t \mathbb{E}\Big(\sup_{r\in[0,s]}|X^n_{t_n(r)}-X^n_r|^2\Big)ds \\ &+ C_2\int_0^t \kappa_\eta\Big(\mathbb{E}\Big(\sup_{r\in[0,s]}|H_r|^2\Big)\Big)ds. \end{aligned} \tag{3.3}$$

Now from (3.1), (3.2) and (3.3) we conclude that for $0 \leq t \leq T$

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\in[0,t]}|H_s|^2\Big) \leq{}& \mathbb{E}\Big(\sup_{s\in[0,t]}|J_1(s)| + \sup_{s\in[0,t]}|J_2(s)| + \sup_{s\in[0,t]}|J_3(s)|\Big) \\ \leq{}& \frac{1}{2}\,\mathbb{E}\Big(\sup_{s\in[0,t]}|H_s|^2\Big) + C_4\int_0^t \mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big)dr \\ &+ C_4\int_0^t \kappa_\eta^2\Big(\Big(\mathbb{E}\big(\sup_{s\in[0,r]}|X^n_{t_n(s)}-X^n_s|^2\big)\Big)^{\frac{1}{2}}\Big)dr \\ &+ C_4\int_0^t \kappa_\eta\Big(\mathbb{E}\Big(\sup_{s\in[0,r]}|X^n_{t_n(s)}-X^n_s|^2\Big)\Big)dr \\ &+ C_4\int_0^t \mathbb{E}\Big(\sup_{s\in[0,t]}|X^n_{t_n(s)}-X^n_s|^2\Big)dr \\ &+ C_4\int_0^t \kappa_\eta\Big(\mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big)\Big)dr, \end{aligned} \tag{3.4}$$

where C_4 depends on λ_1, λ_2 and d.

By moving the term $\frac{1}{2}\,\mathbb{E}\Big(\sup_{s\in[0,t]}|H_s|^2\Big)$ to the left-hand side and multiplying both sides of (3.4) by 2 we obtain

$$\begin{aligned}\mathbb{E}\Big(\sup_{s\in[0,t]}|H_s|^2\Big) \leq{}& C^*\int_0^t \mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big)dr \\ &+ C^*\int_0^t \kappa_\eta^2\Big(\Big(\mathbb{E}\big(\sup_{s\in[0,r]}|X^n_{t_n(s)}-X^n_s|^2\big)\Big)^{\frac{1}{2}}\Big)dr \\ &+ C^*\int_0^t \kappa_\eta\Big(\mathbb{E}\Big(\sup_{s\in[0,r]}|X^n_{t_n(s)}-X^n_s|^2\Big)\Big)dr \\ &+ C^*\int_0^t \mathbb{E}\Big(\sup_{s\in[0,t]}|X^n_{t_n(s)}-X^n_s|^2\Big)dr \\ &+ C^*\int_0^t \kappa_\eta\Big(\mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big)\Big)dr. \end{aligned} \tag{3.5}$$

Next we will construct an estimate for $\mathbb{E}\Big(\sup_{s\in[0,r]}|X^n_{t_n(s)}-X^n_s|^2\Big)$. By (2.4) and Hölder's inequality we have for $k=0,..,2^n-1$ and $0\le t\le T$ that

$$
\begin{aligned}
&\mathbb{E}\Big(\sup_{\frac{k}{2^n}t\le s<\frac{k+1}{2^n}t}|X^n_s-X^n_{t_k}|^2\Big)\\
&\quad=\mathbb{E}\Big(\sup_{\frac{k}{2^n}t\le s<\frac{k+1}{2^n}t}|b(X^n_{t_k},\mu^n_{t_k})(s-t_k)+\sigma(X^n_{t_k},\mu^n_{t_k})(B_s-B_{t_k}|^2\Big)\\
&\quad\le\ 2\mathbb{E}\Big(|b(X^n_{t_k},\mu^n_{t_k})|^2\Big|\frac{k+1}{2^n}t-\frac{k}{2^n}t\Big|^2+\|\sigma(X^n_{t_k},\mu^n_{t_k})\|^2\sup_{\frac{k}{2^n}t\le s<\frac{k+1}{2^n}t}|B_s-B_{t_k}|^2\Big)\\
&\quad=2\mathbb{E}\big(|b(X^n_{t_k},\mu^n_{t_k})|^2\big)(2^{-n}t)^2+2\mathbb{E}\Big(\|\sigma(X^n_{t_k},\mu^n_{t_k})\|^2\sup_{\frac{k}{2^n}t\le s<\frac{k+1}{2^n}t}|B_s-B_{t_k}|^2\Big).
\end{aligned}
$$

From Hölder's inequality (Theorem 1.32) it follows that

$$
\begin{aligned}
&\mathbb{E}\Big(\|\sigma(X^n_{t_k},\mu^n_{t_k})\|^2\sup_{\frac{k}{2^n}t\le s<\frac{k+1}{2^n}t}|B_s-B_{t_k}|^2\Big)\\
&\qquad\le\Big(\mathbb{E}(\|\sigma(X^n_{t_k},\mu^n_{t_k})\|^4\Big)^{\frac12}\Big(\mathbb{E}\big(\sup_{\frac{k}{2^n}t\le s<\frac{k+1}{2^n}t}|B_s-B_{t_k}|^2\big)^2\Big)^{\frac12}\\
&\qquad=\Big(\mathbb{E}(\|\sigma(X^n_{t_k},\mu^n_{t_k})\|^4\Big)^{\frac12}\Big(\mathbb{E}\Big(\sup_{\frac{k}{2^n}t\le s<\frac{k+1}{2^n}t}\Big|\int_{t_k}^s dB_r\Big|^4\Big)\Big)^{\frac12}.
\end{aligned}
$$

Furthermore, the Burkholder-Davis-Gundy inequality (Theorem 1.36) gives us

$$
\mathbb{E}\Big(\sup_{\frac{k}{2^n}t\le s<\frac{k+1}{2^n}t}\Big|\int_{t_k}^s dB_r\Big|^4\Big)\le\hat{C}_1\,\mathbb{E}\Big(\int_{\frac{k}{2^n}t}^{\frac{k+1}{2^n}t}dr\Big)^{\frac42}=\hat{C}_1\,(2^{-n}t)^2.
$$

By similar calculations as in (2.10) and (2.11) we can see that

$$
\begin{aligned}
\mathbb{E}\big(|b(X^n_{t_k},\mu^n_{t_k})|^2\big)&\le\hat{C}_2<\infty,\\
\mathbb{E}(\|\sigma(X^n_{t_k},\mu^n_{t_k})\|^4&\le\hat{C}_3<\infty,
\end{aligned}
$$

so we get that

$$
\mathbb{E}\Big(\sup_{\frac{k}{2^n}t\le s<\frac{k+1}{2^n}t}|X^n_s-X^n_{t_k}|^2\Big)\le\hat{C}_4(2^{-n}t)^2+\hat{C}_4((2^{-n}t)^2)^{\frac12}\le\hat{C}(2^{-n}t)\le\hat{C}(2^{-n}T).
$$

Now (3.5) can be presented as

$$
\begin{aligned}
\mathbb{E}\Big(\sup_{s\in[0,t]}|H_s|^2\Big)\le C\int_0^t\mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big)+\kappa_\eta\Big(\mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big)\Big)dr\\
+CT\kappa_\eta^2(C(2^{-n}T)^{\frac12})+CT\kappa_\eta(C2^{-n}T)+CT(2^{-n}T).\qquad(3.6)
\end{aligned}
$$

Set $A = CT\kappa_\eta^2(C(2^{-n}T)^{\frac{1}{2}}) + CT\kappa_\eta(C2^{-n}T) + CT(2^{-n}T)$ and for $0 \le t \le T$ put

$$g(t) := \mathbb{E}\Big(\sup_{s\in[0,t]} |H_s|^2\Big).$$

Then g is strictly positive and strictly increasing.

Now with (3.6) and Lemma 3.2 we get

$$\begin{aligned} g(t) &\le C\int_0^t \mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big) + \kappa_\eta\Big(\mathbb{E}\Big(\sup_{s\in[0,r]}|H_s|^2\Big)\Big)dr + A \\ &= C\int_0^t g(r) + \kappa_\eta\Big(g(r)\Big)dr + A \\ &\le C_\eta \int_0^t \kappa_\eta\Big(g(r)\Big)dr + A. \end{aligned}$$

Now Lemma 1.38 gives us that for a constant $\bar{C}$ depending on $\eta, T, \lambda_1, \lambda_2$ and d

$$\mathbb{E}\Big(\sup_{s\in[0,T]}|H_s|^2\Big) \le \bar{C}A^{\exp\{-\bar{C}T\}}.$$

Since it holds that $\lim_{n\to\infty} C(2^{-n}T)^{\frac{1}{2}} = 0$ and $\lim_{n\to\infty} C(2^{-n}T) = 0$, we can assume that $C(2^{-n}T)^{\frac{1}{2}} \le \eta$ and $C(2^{-n}T) \le \eta$ and use the property $\kappa_\eta(x) = -x\log(x)$ for $x \le \eta$.

We can see that for A the term $CT\kappa_\eta^2(C(2^{-n}T)^{\frac{1}{2}})$ converges the slowest as $n \to \infty$, so for large n we have $A \le C^{(1)}T\kappa_\eta^2(C(2^{-n}T)^{\frac{1}{2}})$ and thus

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|H_t|^2\Big) \le \bar{C}^{(1)}\Big(T\kappa_\eta^2(C(2^{-n}T)^{\frac{1}{2}})\Big)^{\exp\{-\bar{C}T\}}.$$

By Lemma 2.12 we have that for all $0 < \delta < 2$ there exists a constant $c_\delta > 0$ such that

$$T\kappa_\eta^2(C(2^{-n}T)^{\frac{1}{2}}) \le Tc_\delta C^{2-\delta}(2^{-n}T)^{\frac{2-\delta}{2}} = T^{2-\frac{\delta}{2}}c_\delta C^{2-\delta}2^{-n(1-\frac{\delta}{2})}.$$

Since $0 < e^{-\bar{C}T} < 1$ and for large n it holds that $0 < A < 1$ we can see that the estimate for $\mathbb{E}\Big(\sup_{t\in[0,T]}|H_t|^2\Big)$ gets worse the smaller $e^{-\bar{C}T}$ is.

Let $\varepsilon > 0$ and set $T_0 \in (0,1)$ such that $e^{-\bar{C}T_0} \geq 1-\varepsilon$. Now it follows that

$$\begin{aligned}\Big(T_0\kappa_\eta^2(C(2^{-n}T_0)^{\frac{1}{2}})\Big)^{\exp(-\bar{C}T_0)} &\leq C^{(2)}T_0^{(2-\frac{\delta}{2})(1-\varepsilon)}\,2^{-n(1-\frac{\delta}{2})(1-\varepsilon)} \\ &\leq C^{(2)}(2^{-n}T_0)^{(1-\frac{\delta}{2})(1-\varepsilon)} \\ &\leq C^{(2)}(2^{-n}T_0)^{1-\gamma}\end{aligned}$$

for any $\gamma \in (0,1)$ since δ and ε can be chosen arbitrarily small.

Therefore we get for $T \in [0, T_0]$

$$\mathbb{E}\Big(\sup_{t\in[0,T_0]}|X_t^n - X_t|^2\Big) = O((2^{-n}T)^{1-\gamma}).$$

For the case $T \leq T_0$ the statement is proved. If $T > T_0$ we consider the intervals $[T_0, 2T_0]$, $[2T_0, 3T_0]$, ..., $[[\frac{T}{T_0}]T_0, T]$ and by similar means as above we get

$$\mathbb{E}\Big(\sup_{t\in[0,T]}|X_t^n - X_t|^2\Big) = O((2^{-n}T)^{1-\gamma}).$$

and thus the proof is finished.

□

Bibliography

[1] Krishna B. Athreya and Soumendra N. Lahiri: *Measure Theory and Probability Theory.* Springer, 2006.

[2] Sheldon Axler: *Measure, Integration & Real Analysis.* Springer, 2020.

[3] Drumi Bainov and Pavel Simeonov: *Integral Inequalities and Applications.* Springer, 1992.

[4] Richard F. Bass: *Diffusions and Elliptic Operators.* Springer, 1998.

[5] Patrick Billingsley: *Convergence of Probability Measures.* Second Edition, John Wiley & Sons, 1999.

[6] Patrick Billingsley: *Probability and Measure.* Anniversary Edition, John Wiley & Sons, 2012.

[7] Vladimir I. Bogachev: *Measure Theory Volume I.* Springer, 2007.

[8] Siegfried Carl, Vy K. Le and Dumitru Motreanu: *Nonsmooth Variational Problems and Their Inequalities: Comparison Principles and Applications.* Springer, 2007.

[9] René Carmona and François Delarue: *Probabilistic Theory of Mean Field Games with Applications I.* Springer, 2018.

[10] Xiaojie Ding and Huijie Qiao: *Euler-Maruyama Approximations for Stochastic McKean-Vlasov Equations with Non-Lipschitz Coefficients.* J Theor Probab 34, 1408–1425, 2021.

[11] J. A. Fridy: *Introductory Analysis: The Theory of Calculus.* Academic Press, 2000.

[12] Bert E. Fristedt and Lawrence F. Gray: *A Modern Approach to Probability Theory.* Birkhäuser, Boston, 1997.

[13] Christel Geiss and Stefan Geiss: *An Introduction to Probability Theory.* University of Jyväskylä, Department of Mathematics and Statistics, 2021.

[14] Stefan Geiss: *Stochastic Differential Equations.* University of Jyväskylä, Department of Mathematics and Statistics, 2020.

[15] Gene H. Golub and Charles F. Van Loan: *Matrix Computations.* Third Edition, The Johns Hopkins University Press, Baltimore, 1996.

[16] James L. Hein: *Discrete Mathematics.* Second Edition, Jones and Bartlett Publishers, 2003.

[17] Nobuyuki Ikeda and Shinzo Watanabe: *Stochastic Differential Equations and Diffusion Processes.* Kodansha Ltd. ,1981.

[18] Olav Kallenberg: *Foundations of Modern Probability.* From the Series *Probability and its Applications*, Springer, 1997.

[19] Gopinath Kallianpur and P. Sundar: *Stochastic Analysis and Diffusion Processes.* Oxford University Press, 2014.

[20] Ioannis Karatzas and Steven E. Shreve: *Brownian Motion and Stochastic Calculus.* Second Edition, Springer, 1991.

[21] Juan Li and Hui Min: *Weak solutions of mean-field stochastic differential equations.* Stochastic Analysis and Applications, 35:3, 542-568, 2017.

[22] Xuerong Mao: *Stochastic Differential Equations and Applications.* Second Edition, Woodhead Publishing Limited, 2007.

[23] Michael Meyer: *Continuous Stochastic Calculus with Applications to Finance.* Chapman & Hall/CRC, 2001.

[24] Nitis Mukhopadhyay: *Probability and Statistical Inference.* Marcel Dekker, New York, 2000.

[25] Lawrence C. Paulson: *Logic and Computation: Interactive Proof with Cambridge LCF.* Cambridge University Press, 1990.

[26] Philip E. Protter: *Stochastic Integration and Differential Equations.* Second Edition, Version 2.1, Springer, 2005.

[27] Rong Situ: *Theory if Stochastic Differential Equations with Jumps and Applications*. Springer, New York, 2005.

[28] Daniel W. Stroock and S.R.S. Varadhan: *Multidimensional Diffusion Processes*. Springer, 2006.

[29] Jianfeng Zhang: *Backward Stochastic Differential Equations*. From the Series *Probability Theory and Stochastic Modelling*, Springer, 2017.

[30] Xicheng Zhang: *Euler-Maruyama approximations for SDEs with non-Lipschitz coefficients and applications*. J. Math. Anal. Appl. 316, 447-458, 2005.