Swansea University
Prifysgol Abertawe

Convergence of Numerical Solutions of Stochastic Differential Delay Equations

Ulises Botija Muñoz

student no. [redacted]

supervisor: Prof Chenggui Yuan

Submitted to Swansea University in fulfilment of the requirements for the Degree of Doctor of Philosophy

Swansea University

September 2024

Dedication

To my mother, in loving memory.

Acknowledgements

I would like to express my appreciation to my supervisor, Professor Yuan, for his valuable support and guidance during my PhD study. It has been a privilege to be his student.

I also would like to thank Professor Lytvynov for his kind advice and the mathematics department for its support.

In addition, I want to express my gratitittude to my wife. Without her help, this "chapter" would not have been possible.

Abstract

In this thesis we investigate explicit numerical approximations for stochastic differential delay equations (SDDEs) under a local Lipschitz condition by employing the adaptive Euler-Maruyama (EM) method. Working in both finite and infinite horizons, we achieve strong convergence results of the adaptive EM solution. We also obtain the order of convergence in finite horizon. In addition, we show almost sure exponential stability of the adaptive approximate solution for both SDEs and SDDEs. Further, we prove strong convergence of the adaptive solution for McKean-Vlasov SDDEs (MV-SDDEs). In the second part of the thesis, we estimate the variance of two coupled paths derived with the Multilevel Monte Carlo method combined with the EM discretization scheme for the simulation of MV-SDEs with small noise first and for MV-SDDEs later. The result often translates into a more efficient method than the standard Monte Carlo method combined with algorithms tailored to the small noise setting.

Key Words: Adaptive Euler Maruyama scheme; McKean-Vlasov Stochastic differential delay equations (MV-SDDEs); Strong convergence; Boundedness of the pth-moments; Almost sure exponential stability; Multilevel Monte Carlo simulation; Variance of two coupled paths

DECLARATIONS

STATEMENT 1

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

[redacted] 14/09/2024

Signed Date

STATEMENT 2

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

[redacted] 14/09/2024

Signed Date

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for electronic sharing.

[redacted] 14/09/2024

Signed Date

STATEMENT 4

The University's ethical procedures have been followed and, where appropriate, that ethical approval has been granted.

[redacted]

Signed

14/09/2024

Date

Contents

Chapter 1

Introduction

In 2020, Wei and Giles [14] obtained the boundedness of the pth moments of the numerical solution using the adaptive Euler Maruyama (EM) method in a finite horizon under local Lipschitz and one-sided linear growth conditions for a standard SDE of the type

$$X_t = f(X_t)dt + g(X_t)dW_t, \quad t \geq 0.$$

In the adaptive EM scheme, the time step is not a constant, but a function of the solution at that point in time. They also, under more restrictive conditions, showed strong convergence in infinite horizon. In Chapter 3, we extend their work to SDDEs in both, finite and infinite horizons. Following [14], we will show the boundedness of the pth moments but in our case, this is not enough to prove strong convergence. The main difficulty is that the delay times might not match the times where the numerical solution is computed. To solve the issue, we introduced an auxiliary piecewise constant process on the delay times. This varies from the standard EM method for SDDEs and requires a new proof of convergence.

In [14], the almost sure (a.s.) exponential stability of the adaptive-EM solution was not studied. Here we studied it first for SDEs and later for SDDEs. Moment stability for numerical solutions of SDDEs has been studied extensively, see for example [3], [39]. Almost sure exponential stability is usually derived from moment stability by means of the Borel-Cantelli lemma and Markov's inequality (see [25]). In Wu et al. [49], using the EM and the Backward EM (BEM) methods, a.s. exponential stability was studied for SDDEs without using moment stability. Their approach was based on the martingale convergence theorem. They required the linear growth condition when dealing with the standard EM scheme. When they weaken the linear growth to the one-sided linear growth condition for the drift function, they showed how the standard EM approximate solution loses the stability of the exact solution. Then they showed that under the one-sided linear growth condition, the a.s. exponential stablility can be achieved by using the BEM method. This method is implicit and therefore more computationally expensive than explicit methods like the adaptive EM. In the last three sections of Chapter 3, under similar conditions to the ones used in [49], we obtained a.s. exponential stability using the EM-adaptive method. At the end of the chapter, in Section 3.7, we present some simulations to ilustrate the ideas discussed in the stability sections. Chapter 3 is based on the paper "Explicit Numerical Approximations for SDDEs in Finite and Infinite Horizons using the Adaptive EM Method: Strong Convergence and Almost Sure Exponential Stability", which has already been accepted for publication in the journal of Applied Mathematics and Computation.

In 2021 [46], Reisinger and Stokinger extended the work on the adaptive method from [14] to MV-SDEs. For two different particles the value of the processes at time t_n may differ resulting in two different values (one for each particle) of the random variable t_{n+1}. This, unlike the standard EM method, presents a challenge when computing $\frac{1}{M}\sum_{j=1}^{M}\delta_{\hat{X}^{j,M}(t_1)}$. In [46], they proposed two different schemes, which both deal with this issue. In Chapter 4, we extend these two schemes to the delay case and prove the strong convergence of the adaptive numerical solutions for MV-SDDEs.

An important problem in mathematics is to compute $\mathbb{E}[\Psi(X_T)]$ where $\{X_t\}_{0\leq t\leq T}$ is the solution to an SDE and $\Psi : \mathbb{R}^d \to \mathbb{R}$. This is a very significant problem in financial mathematics where financial derivatives are priced by computing the above the expectation. Among all the methods that allow us to compute the previous expectation, Monte Carlo simulation is arguably the more flexible. Its drawback is the high computational cost. Therefore a lot of effort has been placed to reduce this cost. In 2008, Giles, in a very relevant paper, [15], proposed the multilevel Monte Carlo (MLMC) method which greatly reduces the computational cost with respect to the standard Monte Carlo (MC) method. Since [15], numerous papers have appeared to customize, adapt and extend the principles of multilevel Monte Carlo method to specific problems. One of these papers is [1], where the authors applied the multilevel Monte Carlo framework to standard SDEs with small noise. They compare the computation cost derived from the standard Monte Carlo method (combined with discretization algorithms tailored to the small noise setting) versus the multilevel Monte Carlo method combined with the Euler-Maruyama (EM) scheme. In Chapter 5, we extend the work from [1] to McKean-Vlasov SDEs (MV-SDEs) with small noise and we obtained the same estimate for the variance of two coupled paths. This presents some challenges since we have to deal with the measures approximation $\frac{1}{M}\sum_{j=1}^{M}\delta_{\hat{X}^{j,M}(t_1)}$. The conclusion we arrived at is that the additional McKean-Vlasov component does not add computational complexity (per equation in the system of particles). Chapter 5 is based on the paper "Multilevel Monte Carlo EM scheme for MV-SDEs with small noise", which has been published in the journal Numerical Algebra, Control and Optimization. In Chapter 6 we extend the work from Chapter 5 to MV-SDDEs.

Chapter 2

Preliminaries

In this chapter we review some results from stochastic processes and stochastic differential equation (SDEs) that will be used later throughout the thesis.

2.1 Notation and basic definitions

For a $\mathbb{R}^m$-vector v, we denote the Euclidean norm by $|v| := (|v_1|^2 + ... + |v_m|^2)^{\frac{1}{2}}$ and the inner product of two $\mathbb{R}^m$-vectors v and w by $\langle v, w \rangle := v_1 w_1 + ... + v_m w_m$. For a $m \times d$ matrix A, we denote the Frobenius matrix norm by $||A|| := \sqrt{\text{trace}(A^T A)}$.

Definition 2.1.1. *A probability space $(\Omega, \mathcal{F}, P)$ is said to be complete, if for all $B \in \mathcal{F}$ such that $P(B) = 0$, we have that if $A \subset B$, then $A \in \mathcal{F}$.*

Definition 2.1.2. *A filtration $\{\mathcal{F}_t; 0 \leq t \leq \infty\}$ on $(\Omega, \mathcal{F}, P)$ is a collection of sub-σ-algebras of $\mathcal{F}$, such that $\mathcal{F}_s \subset \mathcal{F}_t$ for every $s \leq t \leq \infty$.*

Remark 2.1.1. The time parameter set can also be $[0, \infty)$ or a finite set $[0, T]$ for some $T > 0$. Sometimes, for convenience, we will not specify the time parameter set and we will just write $\{\mathcal{F}_t\}$ to denote the filtration.

Definition 2.1.3. *For a filtration $\{\mathcal{F}_t\}$, we define the σ-algebras*

$$\mathcal{F}_{t+} := \cap_{s>t} \mathcal{F}_s; \quad \mathcal{F}_{t-} := \sigma(\cup_{s<t} \mathcal{F}_s).$$

We say that the filtration is right-continuous (left continuous), if $\mathcal{F}_t = \mathcal{F}_{t+}$ *(*$\mathcal{F}_t = \mathcal{F}_{t-}$*) for all* t*.*

Definition 2.1.4. *A filtered complete probability space* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$*, is said to satisfy the usual conditions if*

(i) The filtration $\{\mathcal{F}_t\}$ *is right-continuous.*

(ii) $\mathcal{F}_0$ *contains all* P*-null sets of* $\mathcal{F}$*.*

Definition 2.1.5. *A stochastic process* $X := \{X_t; 0 \le t < \infty\}$*, is a collection of* $\mathbb{R}^d$*-valued random variables. The process is said to be* $\mathscr{F}_t$*-adapted if* $X_t \in \mathcal{F}_t$ *(i.e.* X_t *is* $\mathcal{F}_t$*-measurable) for each* t*.*

Remark 2.1.2 (Notation)*.* When it is important to emphasise the time parameter set, we denote a stochastic process by $\{X_t; 0 \le t < \infty\}$ or $\{X_t; 0 \le t \le T\}$. But when there is no place for confusion, we will denote the process just by a capital letter, in this case, X. When we write X_t we are referring to a random variable. For example, we will denote by $\{W_t\}_{0 \le t \le T}$ or just W a standard d-dimensional Brownian motion.

Definition 2.1.6. *A stochastic process* X *is said to be bounded if there exists a constant* $K > 0$ *such that for almost all* $\omega \in \Omega$ *and all* $t \in [0, \infty)$*,* $|X_t(\omega)| \le K$*.*

Definition 2.1.7. *Two stochastic processes* X *and* Y*, are modifications of each other if*

$$P(X_t = Y_t) = 1 \quad \text{for all } t.$$

We say that X *and* Y *are indistinguishable if*

$$P(X_t = Y_t, \text{ for all } t) = 1.$$

Definition 2.1.8 (Martingale)**.** *A stochastic process* M *is a martingale with respect to the filtration* $\{\mathcal{F}_t\}$ *(or a* $\{\mathcal{F}_t\}$*-martingale) if:*

(i) M *is adapted to* $\{\mathcal{F}_t\}$*;*

(ii) $E|M_t| < \infty$ *for all* $t \geq 0$*;*

(iii) $E[M_t|\mathcal{F}_s] = M_s$ *a.s., for all* $0 \leq s \leq t$*.*

Definition 2.1.9 (Stopping time)**.** *The random variable* τ*, taking values in* $[0, \infty]$*, is an* $\{\mathscr{F}_t\}$*-stopping time if*

$$\{\tau \leq t\} = \{\omega : \tau(\omega) \leq t\} \in \mathscr{F}_t$$

for all $t \leq \infty$*.*

Definition 2.1.10 (Stopped process)**.** *Given a stochastic process* X *and a stopping time* τ*, the stopped process* X^τ *is defined by*

$$X_t^\tau := X_{t \wedge \tau}(\omega).$$

Definition 2.1.11 (Local martingale)**.** *Let* M *be an adapted process null at 0. Then* M *is called a local martigale null at 0, and we write* $M \in \mathscr{M}_{0,loc}$*, if there exists an increasing sequence* $\{\tau_n\}$ *of stopping times with* $\tau_n \uparrow \infty$ *a.s. (i.e. for each* $T > 0$ *and each* $\omega \in \Omega$*, there exists* $N(\omega)$*, so that* $n \geq N(\omega)$ *implies* $\tau_n(\omega) \geq T$*) such that each stopped process* M^{τ_n} *is a martingale (null at 0). If* M *is also continuous we write* $M \in c\mathscr{M}_{0,loc}$*. The sequence* $\{\tau_n\}$ *is referred to as a reducing sequence for* M *(into* $\mathscr{M}_0$*).*

Definition 2.1.12 (Cadlag process)**.** *A function is said to be cadlag if it is right-continuous with left limits. We say that a stochastic process is cadlag if for almost every* $\omega \in \Omega$*,* $t \to X_t(\omega)$ *is a cadlag function.*

Definition 2.1.13 (Semimartingale)**.** *A process* X *is called a semimartingale if it is an adapted process that can be written in the form*

$$X = X_0 + M + A, \tag{2.1.1}$$

where X_0 *is an* $\mathscr{F}_0$*-measurable random variable,* M *is a local martingale null at zero and* A *is an adapted cadlag process, also null at zero, having paths of finite variation. We denote by* $\mathcal{S}$ *the space of semimartingales and by* $c\mathcal{S}$ *the subspace of continuous semimartingales.*

Theorem 2.1.14. *[49][30](Discrete semimartingale convergence theorem)*
Let $\{A_i\}, \{U_i\}$ be two sequences of nonnegative random variables such that both A_i and U_i are both $\mathcal{F}_{i-1}$-measurable for $i = 1, 2, ...$ and $A_0 = U_0 = 0$ a.s. Let M_i be a real-value local martingale with $M_0 = 0$ a.s. Let ζ be a nonnegative $\mathcal{F}_0$-measurable random variable. Assume that $\{X_i\}$ is a nonnegative semimartingale with the Doob-Mayer decomposition

$$X_i = \zeta + A_i - U_i + M_i.$$

If $\lim_{i\to\infty} A_i < \infty$ a.s., then for almost all $\omega \in \Omega$,

$$\lim_{i\to\infty} X_i < \infty \quad \text{and} \lim_{i\to\infty} U_i < \infty,$$

that is, both X_i and U_i converge a.s. to finite random variables.

Definition 2.1.15. *Let $0 \le a < b < \infty$. Denote by $\mathcal{L}^i([a,b], \mathbb{R}^{d\times\bar{d}}), i = 1, 2,$ the space of all $\mathbb{R}^{d\times\bar{d}}$-valued measurable $\{\mathcal{F}_t\}$-adapted processes $f := \{f_t, a \le t \le b\}$ such that*

$$\int_a^b ||f_t||^i dt < \infty \quad a.s.$$

Denote by $\mathcal{M}^2([a,b], \mathbb{R}^{d\times\bar{d}})$ the space of all processes $f \in \mathcal{L}^2([a,b], \mathbb{R}^{d\times\bar{d}})$ such that

$$\mathbb{E}\left[\int_a^b ||f_t||^2 dt\right] < \infty.$$

Definition 2.1.16. *A d-dimensional Ito process is an $\mathbb{R}^d$-valued continuous adapted process $x_t = (x_t^{(1)}, ..., x_t^{(d)})^T$ on $t \ge 0$ of the form*

$$x_t = x_0 + \int_0^t f_s ds + \int_0^t g_s dW_s,$$

where $f_t = (f_t^{(1)}, ..., f_t^{(d)})^T \in \mathcal{L}^1(\mathbb{R}_+; \mathbb{R}^d)$ and $g = (g_{ij})_{d\times\bar{d}} \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{d\times\bar{d}})$. we shall say that x_t has stochastic differential dx_t on $t \ge 0$ given by

$$dx_t = f_t dt + g_t dW_t.$$

Theorem 2.1.17. *(Ito's formula)*
Let x_t be a d-dimensional Ito process on $t \ge 0$ with the stochastic differential

$$dx_t = f_t dt + g_t dW_t.$$

Let $V \in C^{2,1}(\mathbb{R}^d \times \mathbb{R}_+; \mathbb{R})$. *Then* $V(x_t, t)$ *is again an Ito process with the stochastic differential given by*

$$dV(x_t, t) = [V_t(x_t, t) + V_x(x_t, t)f_t + \frac{1}{2}trace(g_t^T V_{xx}(x_t, t)g_t)]dt + V_x(x_t, t)g_t dW_t \ a.s.$$

2.2 Delay McKean Vlasov SDEs (MV-SDDEs) with small noise

A MV-SDE is a type of SDE where the coefficients depend on the law (or distribution) of the solution process itself. This is a generalization of classical SDEs where the dynamics of the system are influenced by the collective behavior of all agents in a population. The pionering work on MV-SDEs is due to McKean on his work on the Boltzmann equation in thermodynamics, [42], [43]. Since then, MV-SDEs have been used extensively in in biological systems, financial engineering and physics, [4],[7], [19], [13].

A SDE with small noise is a type of SDE where the drift coefficient function is multiplied by a small positive constant, which in the rest of the thesis will be denoted by ε, i.e. $(0 < \varepsilon << 1)$. In standard SDEs, the noise term (i.e. the diffusion part of the SDE) can be significant and comparable to the drift term. In the small noise case, the diffusion term is scaled down by the small parameter ε. The presence of small noise affects the behavior of the solution and the methods used for analysis. Some examples of fields in science where these equations are used are biochemestry, economics and fluid dynamics, see [1] and references therein.

2.2.1 Wasserstein distance

For any $q > 0$, let $L^q = L^q(\Omega; \mathbb{R}^d)$ be the family of $\mathbb{R}^d$-valued random variables Z with $\mathbb{E}[|Z|^q] < +\infty$. Let $\mathcal{L}^Z$ denote the probability law (or distribution) of a random variable Z. $\delta_x(\cdot)$ denotes the Dirac delta measure concentrated at a point $x \in \mathbb{R}^d$. For $q \geq 1$, we denote by $\mathcal{P}_q(\mathbb{R}^d)$ the set of probability measures on $\mathbb{R}^d$ with finite qth moments, and define

$$W_q(\mu) := \left(\int_{\mathbb{R}^d} |x|^q \mu(\mathrm{d}x) \right)^{\frac{1}{q}}, \quad \forall \mu \in \mathcal{P}_q(\mathbb{R}^d). \tag{2.2.1}$$

Lemma 2.2.1. *[9] (Wasserstein Distance) Let $q \geq 1$. Define*

$$\mathbb{W}_q(\mu, \nu) := \inf_{\pi \in \mathcal{D}(\mu,\nu)} \left\{ \int_{\mathbb{R}^d} |x - y|^q \pi(\mathrm{d}x, \mathrm{d}y) \right\}^{\frac{1}{q}}, \ \mu, \nu \in \mathcal{P}_q(\mathbb{R}^d), \tag{2.2.2}$$

where $\mathcal{D}(\mu, \nu)$ is the set of all couplings for μ and ν. Then $\mathbb{W}_q$ is a distance on $\mathcal{P}_q(\mathbb{R}^d)$.

Lemma 2.2.2. *[9] For any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $\mathbb{W}_2(\mu, \delta_0) = W_2(\mu)$.*

2.2.2 MV-SDDEs with small noise

Let $W(t) = (W_1(t), \ldots, W_{\bar{d}}(t))^T$ be an $\bar{d}$-dimensional Brownian motion defined on the probability space and let $\tau > 0$. The most general type of SDE that we will work with in this thesis, has the form

$$\mathrm{d}X^\varepsilon(t) = f(X^\varepsilon(t), X^\varepsilon(t-\tau), \mathcal{L}_t^X)dt + \varepsilon g(X^\varepsilon(t), X^\varepsilon(t-\tau), \mathcal{L}_t^X)dW(t), t \geq 0 \tag{2.2.3}$$

where $\varepsilon \in (0, 1)$, $\mathcal{L}_t^X$ is the law (or distribution) of $X(t)$,

$$f : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \text{ and } g : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d \times \bar{d}}$$

and the initial data satisfies the following condition: for any $p \geq 2$

$$\{X(\theta) : -\tau \leq \theta \leq 0\} := \Xi \in L^p_{\mathcal{F}_0}([-\tau, 0]; \mathbb{R}^d),$$

that is, Ξ is a $\mathcal{F}_0$-measurable $C([-\tau, 0]; \mathbb{R}^d)$-valued random variable such that $E||\Xi||^p < \infty$.

Definition 2.2.3. *(Strong solution and uniqueness)*
An $\mathbb{R}^d$-valued stochastic process $\{X_t; -\tau \leq t \leq T\}$ is called a strong solution to (2.2.3) *if it satisfies*

(i) *X is continuous and* $\{\mathcal{F}_t\}$*-adapted;*

(ii) $\{f(X^\varepsilon(t), X^\varepsilon(t-\tau), \mathcal{L}_t^X)\} \in L^1(W; [0,T]; \mathbb{R}^d)$ *and* $\{g(X^\varepsilon(t), X^\varepsilon(t-\tau), \mathcal{L}_t^X)\} \in L^2(W; [0,T]; \mathbb{R}^{d\times \bar{d}})$;

(iii) *Equation* (2.2.3) *holds for every* $t \in [0,T]$ *with probability 1.*

A solution X is said to be unique, if any other solution $\widehat{X}$ *is indistinguishable from X.*

Theorem 2.2.4. *[33] Assume that the coefficient functions f and g satisfy:*

(i) *(Lipschitz condition on g) There exists a positive constant L such that*

$$||g(x,y,\mu) - g(\bar{x},\bar{y},\bar{\mu})||^2 \leq L(|x-\bar{x}|^2 + |y-\bar{y}|^2 + \mathbb{W}_2^2(\mu,\bar{\mu})) \tag{2.2.4}$$

for all $x, \bar{x}, y, \bar{y} \in \mathbb{R}^d$ *and* $\mu, \bar{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$.

(ii) *(one-sided Lipschitz condition on f) There exists a positive constant L such that*

$$\langle x-\bar{x}, f(x,y,\mu) - f(\bar{x},\bar{y},\mu)\rangle \leq L(|x-\bar{x}|^2 + |y-\bar{y}|^2) \tag{2.2.5}$$

for all $x, \bar{x}, y, \bar{y} \in \mathbb{R}^d$ *and* $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

(iii) *(Lipschitz measure dependence condition on f) There exists a positive constant L such that*

$$|f(x,y,\mu) - f(x,y,\bar{\mu})| \leq L\mathbb{W}_2(\mu,\bar{\mu}) \tag{2.2.6}$$

for all $x, y \in \mathbb{R}^d$ *and* $\mu, \bar{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$.

(iv) *polynomial growth Lipschitz condition on f, i.e. there exist constants* $\gamma, \lambda, q > 0$ *such that for all* $x, y, \bar{x}, \bar{y} \in \mathbb{R}^m$

$$|f(x,y,\mu) - f(\bar{x},\bar{y},\mu)| \leq (\gamma(|x|^q + |y|^q + |\bar{x}|^q + |\bar{y}|^q) + \lambda)(|x-\bar{x}| + |y-\bar{y}|). \tag{2.2.7}$$

Then the equation (2.2.3) *has a unique strong solution and the solution belongs to* $\mathcal{M}^2([t_0,T]; \mathbb{R}^{d\times\bar{d}})$ *[33].*

2.2.3 Lions derivative

Now we will give the definition of the Lions derivative (or L-derivative) for a function $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ as introduced in [8],[9].

Definition 2.2.5. *Given* $(\Omega, \mathcal{F}, \mathbb{P})$*, an atom is a set* $A \in \mathcal{F}$ *such that* $\mathbb{P}(A) > 0$ *and for any* $B \in \mathcal{F}, B \subset A, \mathbb{P}(A) > \mathbb{P}(B)$*, we have that* $\mathbb{P}(B) = 0$*. We say a probability space is atomless if it does not have any atoms.*

Definition 2.2.6. *We say that* $u : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ *is L-differentiable at* $\mu \in \mathbb{P}_2(\mathbb{R}^d)$ *if there is an atomless probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and an* $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ *such that* $\mu = \mathcal{L}^X$ *and the lifted function* $U : L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d) \to \mathbb{R}$ *given by* $U(X) := u(\mathcal{L}^X)$ *is Frechet differentiable at* X*.*

The following two propositions taken from [9] are key in order to define the L-derivative later.

Proposition 2.2.7. *[9] Let* u *be a real valued function on* $\mathbb{P}_2(\mathbb{R}^d)$ *and* U *be its lifting to* $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$*. If* u *is L-differentiable at* $\mu_0 \in \mathbb{P}_2(\mathbb{R}^d)$*, then the lifting* U *is differentiable at each* $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ *such that* $\mu_0 = \mathcal{L}^X$*, and the law of the pair* $(X, [DU](X))$ *does not depend upon the random variable* X *as long as* $\mu_0 = \mathcal{L}^X$*.*

Proposition 2.2.7 implies that the distribution of the L-derivative of u at μ_0, when viewed as a random variable, depends only upon the law μ_0, an not upon the particular X_0 having distribution μ_0. The Frechet derivative $[DU](X_0)$ is called the representation of the L-derivative of u at μ_0 along the variable X_0. Since it is viewed as an element of $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$, by definition,

$$U(X) = U(X_0) + [DU](X_0)(X - X_0) + o(||X - X_0||_2),$$

whenever X and X_0 are random variables with distributions μ and μ_0 respectively.

Proposition 2.2.8. *[9] Let* u *be a real valued continuously L-differentiable function on* $\mathbb{P}_2(\mathbb{R}^d)$ *and* U *its lifting to* $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$*. Then for any* $\mu \in \mathbb{P}_2(\mathbb{R}^d)$*, there exists* *a measurable function* $h : \mathbb{R}^d \to \mathbb{R}^d$ *such that for all* $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ *with* $\mathcal{L}^X = \mu$, *it holds that* $[DU](X) = h(X)$ *almost surely.*

Quoting [9], when we say that u is continuously L-differentiable, we mean that the Fréchet derivative $[DU](X)$ of the lifting u is a continuous function of X from the space $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ into itself. Proposition 2.2.8 implies that, as a random variable, this Fréchet derivative is of the form $h(X_0)$ for some deterministic measurable function $h : \mathbb{R}^d \to \mathbb{R}^d$, which is uniquely defined μ_0-almost everywhere on $\mathbb{R}^d$. The equivalence class of h in $L^2(\mathbb{R}^d, \mu_0; \mathbb{R}^d)$ is uniquely defined and we denoted by $\partial_\mu u(\mu_0)$. We say that $\partial_\mu u(\mu_0)$ is the L-derivative of u at μ_0 and identify it with the function $\partial_\mu u(\mu_0)(\cdot) : \mathbb{R}^d \ni x \to \partial_\mu u(\mu_0)(x) \in \mathbb{R}^d$.

By the mean value theorem (see chapter 5 in [9]), for any two d-dimensional random variables X and X', there exists a $\theta \in [0, 1]$ such that

$$u(\mathcal{L}^X) - u(\mathcal{L}^{X'}) = \mathbb{E}[\langle \partial_\mu u(\mathcal{L}^{\theta X + (1-\theta)X'})(\theta X + (1-\theta)X'), (X - X')\rangle]. \tag{2.2.8}$$

2.2.4 Stochastic Particle Method for MV-SDDEs with small noise

This method, known in the literature as the propagation of chaos result [42], allows us to approximate the MV-SDDE (2.2.3) by a system of particles. Each one of these particles satisfy a MV-SDDEs and the system is constructed in a way such that the particles are uncorrelated with each other. In Theorem 2.2.9 we will see that the solution of the system converges to the solution of (2.2.3). The benefit of this is that the system is more tractable and facilitates the construction of numerical solutions that will still converge to the exact solution of (2.2.3). Now we provide the details of the construction process. For all $i \in \mathbb{N}$, let $\{W^i(t)\}_{t\in[0,T]}$ be a $\bar{d}$-dimensional Brownian motion. Assume $\{W^1(t)\}, \{W^2(t)\}, \cdots$ are independent and $x^1, x^2, \cdots$ are independent and identically distributed (*i.i.d.*) $L^p_{\mathcal{F}_0}([-\tau, 0]; \mathbb{R}^d)$ random variables with the same distribution as Ξ. Let $\{X^{\varepsilon,i}(t)\}_{t\in[0,T]}$ be the unique solution to the MV-SDDE

$$\mathrm{d}X^{\varepsilon,i}(t) = f(X^{\varepsilon,i}(t), X^{\varepsilon,i}(t-\tau), \mathcal{L}_t^{X^{\varepsilon,i}})\mathrm{d}t + \varepsilon g(X^{\varepsilon,i}(t), X^{\varepsilon,i}(t-\tau), \mathcal{L}_t^{X^{\varepsilon,i}})W^i(t), \tag{2.2.9}$$

where $\varepsilon \in (0,1), \mathcal{L}_t^{X^{\varepsilon,i}}$ is the law (or distribution) of $X^{\varepsilon,i}(t)$, $X_0^{\varepsilon,i} = x^i$ and

$$f : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \text{ and } g : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times \bar{d}}.$$

One can see that $X^{\varepsilon,1}(t), X^{\varepsilon,2}(t), ...$ are *i.i.d.* for $t \geq 0$.

Now we define the M-particles systems of MV-SDDEs. For any $M \in \mathbb{N}$, $1 \leq i \leq M$, let $X^{\varepsilon,i,M}(t)$ be the solution of the MV-SDDE

$$\mathrm{d}X^{\varepsilon,i,M}(t) = f(X^{\varepsilon,i,M}(t), X^{\varepsilon,i,M}(t-\tau), \mathcal{L}_t^{\varepsilon,X,M})dt + \varepsilon g(X^{\varepsilon,i,M}(t), X^{\varepsilon,i,M}(t), \mathcal{L}_t^{\varepsilon,X,M})dW^i(t) \tag{2.2.10}$$

with the initial condition $X_0^{\varepsilon,i,M} = x^i$, where $\mathcal{L}_t^{\varepsilon,X,M} := \frac{1}{M}\sum_{j=1}^{M} \delta_{X^{\varepsilon,j,M}(t)}$. The next theorem is known in the literature as the propagation of chaos proves the convergence between $X^{\varepsilon,i}$ and $X^{\varepsilon,i,M}$.

Theorem 2.2.9. *[42][10] [33] If the assumptions of Theorem 2.2.4 hold, then*

$$\lim_{M\to\infty} \sup_{1\leq i\leq M} \mathbb{E}\Big[\sup_{0\leq t\leq T} |X^{\varepsilon,i}(t) - X^{\varepsilon,i,M}(t)|^2\Big] = 0.$$

2.3 Some definitions of convergence and stability of numerical solutions of SDEs

Here we summarize a few general definitions about convergence and stability of numerical solutions to SDEs. Let $\{Y_t, 0 \leq t \leq T\}$ and $\{X_t, 0 \leq t \leq T\}$ be the exact and the numerical solution to a SDE respectively. Let Δ be the stepsize of the numerical solution.

Definition 2.3.1. *(Strong convergence)*
We say that X converges strongly to Y if

$$\lim_{\Delta\to 0} \mathbb{E}[\sup_{0\leq t\leq T} |Y_t - X_t|] = 0.$$

Definition 2.3.2. *(Order of convergence)*

We say that X converges strongly to Y with order p if there exist a constant C independent of Δ, such that

$$\mathbb{E}[\sup_{0 \leq t \leq T} |Y_t - X_t|] = C\Delta^p.$$

Definition 2.3.3. *(Weak convergence)*

Let g be a Borel-measurable function. We say that X converges weakly to Y at time T with respect to a class $\mathcal{C}$ of test functions g if we have

$$\lim_{\Delta \to 0} |\mathbb{E}[g(Y_T)] - \mathbb{E}[g(X_T)]| = 0$$

for all $g \in \mathcal{C}$.

Definition 2.3.4. *(Almost sure exponential stability)*

We say that the solution to a SDE, X, is almost surely exponentially stable if there is a $\Delta^* \in (0,1)$ and a positive constant η such that

$$\limsup_{k \to \infty} \frac{1}{k\Delta} \log |X_k| \leq -\eta \quad a.s.$$

for any $\Delta \leq \Delta^*$.

Definition 2.3.5. *(Moment exponential stability) Let $p \in (0,2]$.)We say that X is p-moment exponentially stable if there is a $\Delta^* \in (0,1)$ and a positive constant η such that*

$$\lim_{k \to \infty} \frac{1}{k\Delta} \log \mathbb{E}[|X_k|^p] \leq -\eta$$

for any $\Delta \leq \Delta^*$.

Chapter 3

Numerical Approximations for SDDEs using the Adaptive EM Method: Strong Convergence and Almost Sure Exponential Stability

3.1 Introduction

The classical existence-and-uniqueness theorem for SDDEs requires the drift and diffusion functions to satisfy a local Lipschitz condition and a linear growth condition (see [37]). However, in applications there are many SDDEs which do not satisfy the linear growth condition on the drift coefficient. The Khasminskii-type theorem (monotone condition) in [38] enables to prove existence-and-uniquess for a class of SDDEs using a weaker condition than the linear growth one. Since there is no explicit solutions for most SDEs, it is desirable, under these weaker conditions, to find numerical approximate solutions that converge strongly to the exact solution. In 2003, Mao [40] proved strong convergence using the EM scheme and assuming the boundedness of the pth moments for both the exact and the numerical solution. It is well-known that the linear growth condition implies the boundedness of the pth moments for the EM approximate solution. But when the drift function grows faster than linear, the standard EM scheme fails. We provide some examples from [28] with polynomial growth:

- Stochastic Ginzburg-Landau equation

$$dX_t = \left((\eta + 2\sigma^2)X_t - \lambda X_t^3)\right) dt + \sigma X_t dWt, \quad X_0 = x_0 \in (0, \infty),$$

  where $\eta \geq 0$ and $\lambda, \sigma > 0$.

- Stochastic Verhulst equation

$$dX_t = \left((\eta + \frac{1}{2}\sigma^2)X_t - \lambda X_t^2)\right) dt + \sigma X_t dWt, \quad X_0 = x_0 \in (0, \infty),$$

  where $\eta, \lambda, \sigma > 0$.

- Feller diffusion with logistic growth

$$dX_t = \lambda X_t(K - X_t)dt\sigma\sqrt{X_t}dW_t,$$

  where $\lambda, K, \sigma > 0$.

. Therefore, modifications of the EM scheme which provide explicit approximate solutions, have appeared in the last few years to account for this issue. Examples of these are the Tamed [29] and the Truncated [21] methods.

In 2020, Wei and Giles [14] obtained the boundedness of the pth moments of the numerical solution using the adaptive-EM method in a finite horizon under local Lipschitz and one-sided linear growth conditions. This, by the previous work of Higham in 2002 [23], automatically implies strong convergence. In the adaptive EM scheme, the time step is not a constant, but a function of the solution at that point in time. They also, under more restrictive conditions, showed strong convergence in infinite horizon. Here, in the first part of this chapter we extend their work to SDDEs in both, finite and infinite horizons. Following [14], we will show the boundedness of the pth moments but in our case, this is not enough to prove strong convergence. The main difficulty is that delay times might not match the times where the numerical solution is computed. To solve the issue, we introduced an auxiliary piecewise constant process on the delay times. This varies from the standard EM method for SDDEs and requires a new proof of convergence.

In [14], the almost sure (a.s.) exponential stability of the adaptive-EM solution was not studied. Here we studied it first for SDEs and later for SDDEs. Moment stability for SDDEs has been studied extensively, see for example [3], [39]. A.s. exponential stability is usually derived from moment stability by means of the Borel-Cantelli lemma and Markov's inequality (see [25]). In Wu et al. [49], using the EM and the Backward EM (BEM) methods, a.s. exponential stability was studied for SDDEs without using moment stability. Their approach was based on the martingale convergence theorem. They required the linear growth condition when dealing with the standard EM scheme. When they weaken the linear growth to the one-sided linear growth condition for the drift function, they showed how the standard EM approximate solution loses the stability of the exact solution. Then they showed that under the one-sided linear growth condition, the a.s. exponential stablility can be achieved by using the BEM method. This method is implicit and therefore more computationally expensive than explicit methods like the adaptive EM. Here, under similar conditions to the ones used in [49], we obtained a.s. exponential stability using the EM-adaptive method.

The rest of the chapter is structured as follows. Section 3.2 describes the adaptive EM method. Section 3.3 deals with strong convergence and order of convergence in finite horizon. In Section 3.4 we obtained the boundedness of the pth moments for the adaptive EM approximate solution in infinite horizon. In Section 3.5 we show a.s. exponential stability of the adaptive EM solution for SDEs and provide a counterexample in which standard EM fails. Section 3.6 follows closely to Section 3.5, but this time we work with SDDEs.

3.2 Adaptive-EM Method for SDDEs

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ be a filtered complete probability space where the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfies the usual conditions. Let $W(t) = (W_1(t), \ldots, W_{\bar{d}}(t))^T$ be an $\bar{d}$-dimensional Brownian motion defined with respect to $\{\mathcal{F}_t\}_{t\geq 0}$. Let $\tau > 0$ and $T > 0$ be constants and denote $C([-\tau, 0]; \mathbb{R}^d)$ the space of all continuous functions from $[-\tau, 0]$ to $\mathbb{R}^d$ with the norm $||\phi|| = \sup_{-\tau\leq\theta\leq 0} |\phi(\theta)|$. Consider an d-dimensional SDDE of the form

$$dY_t = f(Y_t, Y_{t-\tau})dt + g(Y_t, Y_{t-\tau})dW_t \tag{3.2.1}$$

on $t \in [0, T]$, where $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times\bar{d}}$ are Borel-measurable functions, and the initial data satisfies the following condition: for any $p \geq 2$

$$\{Y(\theta) : -\tau \leq \theta \leq 0\} = \xi \in L^p_{\mathcal{F}_0}([-\tau, 0]; \mathbb{R}^d),$$

that is ξ is a $\mathcal{F}_0$-measurable $C([-\tau, 0]; \mathbb{R}^d)$-valued random variable such that $E||\xi||^p < \infty$.

Now we define the numerical solution based on the adaptive method. The time step is determined by a function $h^\delta : \mathbb{R}^d \to \mathbb{R}^+$ with $\delta \in (0, 1)$. The family of functions $\{h^\delta\}_{0<\delta<1}$ is not specifically defined, it just has to satisfy certain conditions that we will describe later in the next assumption. To see concrete examples where the function h^δ is fully specified, see the example (3.6.11) at the end of this chapter. We now define the adaptive method for SDDEs. Set

$$\hat{X}_0 := \xi(0),\ h^\delta_0 := h^\delta(\hat{X}_0),\ t_1 := h^\delta_0.$$

We introduce the continuous-time step (auxiliary) process $\bar{X}$. Define

$$\bar{X}_t := \xi(t), t \in [-\tau, 0), \quad \bar{X}_t := \xi(0), t \in [0, t_1).$$

For t_1 we define the discrete-time approximate solution $\hat{X}$ as

$$\hat{X}_{t_1} := \hat{X}_0 + f(\bar{X}_0, \bar{X}_{-\tau})h^\delta_0 + g(\bar{X}_0, \bar{X}_{-\tau})\Delta W_0,$$
$$h^\delta_1 := h^\delta(\hat{X}_{t_1}),\ t_2 = t_1 + h^\delta_1,$$

$$\bar{X}_t := \hat{X}_{t_1}, t \in [t_1, t_2),$$

where $\Delta W_0 := W_{t_1} - W_0$. Then for a generic t_n we define

$$\begin{aligned} \hat{X}_{t_{n+1}} &:= \hat{X}_{t_n} + f(\bar{X}_{t_n}, \bar{X}_{t_n-\tau})h_n^\delta + g(\bar{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n, \\ h_{n+1}^\delta &:= h^\delta(\hat{X}_{t_{n+1}}),\ t_{n+2} = t_{n+1} + h_{n+1}^\delta, \\ \bar{X}_t &:= \hat{X}_{t_{n+1}}, t \in [t_{n+1}, t_{n+2}), \end{aligned} \tag{3.2.2}$$

where $\Delta W_n := W_{t_{n+1}} - W_{t_n}$. For every path $\omega \in \Omega$, we continue the recursion (3.1) until $n = N(\omega) := \inf\{n \in \mathbb{Z}^+ : t_n(\omega) \geq T\}$. Note that t_n and h_n^δ are random variables. We now introduce a second auxiliary step process. For every ω, let $r = r(\omega)$ be such $t_r \leq \tau \leq t_{r+1}$. Then we define the process $\tilde{X}$ as

$$\begin{aligned} &\tilde{X}_t := \bar{X}_{-\tau}, t \in [-\tau, t_1 - \tau), \quad \tilde{X}_t := \bar{X}_{t_1-\tau}, t \in [t_1 - \tau, t_2 - \tau), ..., \\ &\tilde{X}_t := \bar{X}_{t_r-\tau}, t \in [t_r - \tau, t_{r+1} - \tau), \tilde{X}_t := \bar{X}_{t_{r+1}-\tau}, t \in [t_{r+1} - \tau, t_{r+2} - \tau), \\ &\tilde{X}_t := \bar{X}_{t_{r+n}-\tau}, t \in [t_{r+n} - \tau, t_{r+n+1} - \tau) \end{aligned} \tag{3.2.3}$$

for $n = 1, ..., N - r$. We now define the continuous approximate solution

$$\begin{aligned} &X_t := \xi(t), \quad t \in [-\tau, 0]; \\ &X_t := X_0 + \int_0^t f(\bar{X}_s, \tilde{X}_{s-\tau})ds + \int_0^t g(\bar{X}_s, \tilde{X}_{s-\tau})dW_s, \quad t \in [0, T]. \end{aligned} \tag{3.2.4}$$

Note that $\hat{X}_{t_n} = \bar{X}_{t_n} = X_{t_n}$ for $n = 0, 1, ..., N$.

Remark 3.2.1. The reason to introduce the second step process $\tilde{X}$ is that we can not use the process $\bar{X}_{t-\tau}, t \in [t_n, t_{n+1}]$ to construct the continuous approximation. This is because $\bar{X}_{t-\tau}$ may not be constant in the intervals $[t_n, t_{n+1}]$ which implies that the desired equality $\hat{X}_{t_n} = X_{t_n}$, might not hold. This equality is crucial later to prove convergence. Unlike the case for SDEs in [14], the fact that we can not use $\bar{X}_{t-\tau}, t \in [t_n, t_{n+1}]$, has the added difficulty that in order to prove convergence is not not enough to just show the boundedness of the pth moments and then refer to [40]. In our case a new proof of convergence is required.

3.3 Convergence of the numerical solutions on finite time interval

In this section we will work on a finite time interval $[-\tau, T], T > 0$, and investigate the convergence of the numerical solutions to the exact solution on $[0, T]$.

Assumption 3.3.1. *The functions f and g satisfy the local Lipschitz condition: for every $R > 0$ there exists a positive constant C_R such that*

$$|f(x,y) - f(\bar{x},\bar{y})| + ||g(x,y) - g(\bar{x},\bar{y})|| \le C_R(|x-\bar{x}| + |y-\bar{y}|) \tag{3.3.1}$$

for all $x, y, \bar{x}, \bar{y} \in \mathbb{R}^d$ with $|x| \vee |y| \vee |\bar{x}| \vee |\bar{y}| \le R$. Furthermore, there exist two constants $\alpha, \beta \ge 0$ such that for all $x, y \in \mathbb{R}^d$, f satisfies the one-sided linear growth condition:

$$\langle x, f(x,y)\rangle \le \alpha(|x|^2 + |y|^2) + \beta \tag{3.3.2}$$

and g satisfies the linear growth condition:

$$||g(x,y)||^2 \le \alpha(|x|^2 + |y|^2) + \beta. \tag{3.3.3}$$

Assumption 3.3.2. *The time step function $h^\delta : \mathbb{R}^d \to \mathbb{R}^+,\ \delta \in (0,1)$, is continuous, strictly positive and bounded by δT, i.e.*

$$0 < h^\delta(x) \le \delta T \quad \textit{for all } x \in \mathbb{R}^d. \tag{3.3.4}$$

Furthermore, there exist constants $\alpha, \beta > 0$ such that for all $x, y \in \mathbb{R}^d$.

$$\langle x, f(x,y)\rangle + \frac{1}{2}h^\delta(x)|f(x,y)|^2 \le \alpha(|x|^2 + |y|^2) + \beta. \tag{3.3.5}$$

Note that condition (3.3.5) implies condition (3.3.2) with the same values of α and β.

Remark 3.3.1. In practice, the theory of this section can be applied in the following way. Assume we are giving a SDDE which satisfies Assumption 3.3.1. After, knowing the specific definition of the SDDE we are working with, we define a timestep function h^δ that must satisfy Assumption 3.3.2. Then as we will see later in Theorem 3.3.9, we can assure that the numerical solution converges to the exact solution.

3.3.1 The boundedness of the pth moments of the exact solution and the numerical solutions

The next lemma shows the boundedness of the pth moments of the exact solution.

Lemma 3.3.3. *If the SDDE* (3.2.1) *satisfies Assumption* (3.3.1), *then any* $p > 0$ *there exists a positive constant* C *such that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|Y_t|^p\right] \leq C. \tag{3.3.6}$$

Proof. The proof is given in Lemma 3.2 in [29]. □

Now, the pth moments of numerical solution will be investigated. In the standard Euler-Maruyama method the discretisation times $\{t_n\}$ are built using a constant time step Δ and a fixed number of steps $N \in \mathbb{N}$, i.e. $t_N = N\Delta = T$. However, in the adaptive method, $\{t_n\}$ is a sequence of random variables and there is no guarantee that it reaches T in a finite number of steps. Thus, we have the following definition.

Definition 3.3.4. *We say that the time horizon* T *is attainable if* $\{t_n\}$ *reaches* T *in a finite number of steps* N, *i.e. for almost all* $\omega \in \Omega$, *there exists a* $N(\omega)$ *such that* $t_{N(\omega)} = \sum_{n=0}^{N(\omega)} h^\delta(X_{t_n}) \geq T$.

Theorem 3.3.5. *If the SDDE* (3.2.1) *and the function* h^δ *satisfy Assumption 3.3.1 and 3.3.2 respectively, then* T *is attainable and for all* $p > 0$ *there exists a constant* $C > 0$ *dependent on* T *and* p, *but independent of* h_n^δ, *such that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|X_t|^p\right] \leq C. \tag{3.3.7}$$

The discrete-time approximate solution defined in (3.2.2) need not be bounded. In order to show that T is attainable and prove Theorem 3.3.5, we need to work with a bounded approximate solution. To this end we now introduce the following auxiliary scheme. Let $K > ||\xi||$. Set $\hat{X}_0^K := \xi(0), h_0^\delta := h^\delta(\hat{X}_0), t_1 := h_0^\delta$ and $\bar{X}_t^K := \xi(t), t \in [-\tau, 0), \bar{X}_t^K := \xi(0), t \in [0, t_1)$. Consider the function $\Phi_K : \mathbb{R}^d \to \mathbb{R}^d, \Phi(x) =$

$\min(1, K/|x|)x$. Then for every $\omega \in \Omega$ and for $n = 0, 1, ...N(\omega)$, we define

$$\begin{aligned}\hat{X}^K_{t_{n+1}} &:= \Phi_K(\hat{X}^K_{t_n} + f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})h^\delta_n + g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n) \\ h^\delta_{n+1} &:= h^\delta(X^K_{t_{n+1}}), t_{n+2} := t_{n+1} + h^\delta_{n+1}, \\ \bar{X}^K_t &:= \hat{X}^K_{t_{n+1}}, t \in [t_{n+1}, t_{n+2}).\end{aligned} \tag{3.3.8}$$

where $N(\omega) := \inf\{n \in \mathbb{Z}^+ : t_n(\omega) \geq T\}$. Define for $n = 0, ..., N - r$

$$\tilde{X}^K_t := \bar{X}^K_{t_n-\tau}, t \in [t_n - \tau, t_{n+1} - \tau), \tag{3.3.9}$$

where $r = r(\omega)$ is such that $t_r \leq \tau \leq t_{r+1}$. We now define the the continuous approximate solution

$$\begin{aligned}X^K_t &:= \xi(t), \quad t \in [-\tau, 0]; \\ X^K_t &:= \Phi_K\left(\hat{X}^K_{\underline{t}} + f(\hat{X}^K_{\underline{t}}, \bar{X}^K_{\underline{t}-\tau})(t - \underline{t}) + g(\hat{X}^K_{\underline{t}}, \bar{X}^K_{\underline{t}-\tau})(W_t - W_{\underline{t}})\right) \quad t \in [0, T],\end{aligned} \tag{3.3.10}$$

where $\underline{t} := \max\{t_n : t_n \leq t\}$. Note that $X^K_{t_n} = \hat{X}^K_{t_n} = \bar{X}^K_{t_n}$.

Lemma 3.3.6. *Let the SDDE* (3.2.1) *satisfy Assumption 3.3.1 and the function h^δ satisfy Assumption 3.3.2. Then, for the auxiliary scheme defined by* (3.3.10)*, T is attainable and for all $p > 0$ there exists a constant C dependent on T and p, but independent of h^δ_n and K such that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X^K_t|^p\right] \leq C. \tag{3.3.11}$$

Proof. Let $p \geq 4$ and fix $\delta \in (0, 1)$. Since h^δ is continuous and strictly positive, $\inf_{|x|\leq K} h^\delta(x) > 0$. This implies that for every $\omega \in \Omega$

$$\liminf_{n\to\infty} h^\delta_n(\omega) = \liminf_{n\to\infty} h^\delta(\hat{X}^K_{t_n}(\omega)) > 0,$$

so $\lim_{n\to\infty} t_n(\omega) = \sum_{n=0}^{\infty} h^\delta_n(\omega) = \infty$ for all $\omega \in \Omega$ and T is attainable in the bounded scheme.

Now we will prove the boundedness of the pth moments and the upper bound will be independent of h_n^δ and K. Let $t \in [0,T]$. Define $\underline{t} := \max\{t_n : t_n \le t\}$, and $n_t := \max\{n : t_n \le t\}$. Using (3.3.8) and since for any $x \in \mathbb{R}^d$, $|\Phi(x)|^2 \le |x|^2$, we have that for $n = 0$ to $n = n_t - 1$,

$$\begin{aligned}
|\hat{X}^K_{t_{n+1}}|^2 &\le |\hat{X}^K_{t_n} + f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})h_n + g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n|^2 \\
&= \langle \hat{X}^K_{t_n}, \hat{X}^K_{t_n}\rangle + 2\langle \hat{X}^K_{t_n}, f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})h_n\rangle \\
&+ \langle f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})h_n, f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})h_n\rangle \\
&+ 2\langle \hat{X}^K_{t_n} + f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})h_n, g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n\rangle \\
&+ \langle g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n, g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n\rangle \\
&= |\hat{X}^K_{t_n}|^2 + 2h_n\left[\langle \hat{X}^K_{t_n}, f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\rangle + \frac{1}{2}h_n|f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})|^2\right] \\
&+ 2\langle \hat{X}^K_{t_n} + f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})h_n, g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n\rangle + |g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n|^2 \\
&\le |\hat{X}^K_{t_n}|^2 + 2h_n\alpha(|\hat{X}^K_{t_n}|^2 + |\bar{X}^K_{t_n-\tau}|^2) + 2h_n\beta \\
&+ 2\langle \hat{X}^K_{t_n} + f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})h_n, g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n\rangle + |g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n|^2,
\end{aligned}$$

where in the last step we have used condition (3.3.5). Note that, since it is irrelevant in this proof, we have dropped the symbol "δ" in the adaptive time-step "h_n^δ" to ease the notation. Solving the recurrence relation, we get

$$\begin{aligned}
|\hat{X}^K_{\underline{t}}|^2 &\le |\hat{X}^K_0|^2 + 2\alpha\left(\sum_{n=0}^{n_t-1}|\hat{X}^K_{t_n}|^2h_n + |\bar{X}^K_{t_n-\tau}|^2h_n\right) + 2\beta\underline{t} \\
&+ 2\sum_{n=0}^{n_t-1}\langle \hat{X}^K_{t_n} + f(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})h_n, g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n\rangle \\
&+ \sum_{n=0}^{n_t-1}|g(\hat{X}^K_{t_n}, \bar{X}^K_{t_n-\tau})\Delta W_n|^2. \qquad (3.3.12)
\end{aligned}$$

Similarly, the continuous approximate solution verifies

$$\begin{aligned}
|X^K_t|^2 &\le |\hat{X}^K_{\underline{t}}|^2 + 2(t-\underline{t})\alpha(|\hat{X}^K_{\underline{t}}|^2 + |\bar{X}^K_{\underline{t}-\tau}|^2) + 2(t-\underline{t})\beta \\
&+ 2\langle \hat{X}^K_{\underline{t}} + f(\hat{X}^K_{\underline{t}}, \bar{X}^K_{\underline{t}-\tau})(t-\underline{t}), g(\hat{X}^K_{\underline{t}}, \bar{X}^K_{\underline{t}-\tau})(W_t - W_{\underline{t}})\rangle \\
&+ |g(\hat{X}^K_{\underline{t}}, \bar{X}^K_{\underline{t}-\tau})(W_t - W_{\underline{t}})|^2. \qquad (3.3.13)
\end{aligned}$$

Substituting (3.3.12) into (3.3.13) yields

$$
\begin{aligned}
|X_t^K|^2 &\le |\hat{X}_0^K|^2 \\
&+ 2\alpha \left(\sum_{n=0}^{n_t-1} |\hat{X}_{t_n}^K|^2 h_n + |\bar{X}_{t_n-\tau}^K|^2 h_n + |\hat{X}_{\underline{t}}^K|^2 (t-\underline{t}) + |\bar{X}_{\underline{t}-\tau}^K|^2 (t-\underline{t}) \right) \\
&+ 2\beta t + 2 \sum_{n=0}^{n_t-1} \langle \hat{X}_{t_n}^K + f(\hat{X}_{t_n}^K, \bar{X}_{t_n-\tau}^K) h_n, g(\hat{X}_{t_n}^K, \bar{X}_{t_n-\tau}^K) \Delta W_n \rangle \\
&+ 2 \langle \hat{X}_{\underline{t}}^K + f(\hat{X}_{\underline{t}}^K, \bar{X}_{\underline{t}-\tau}^K)(t-\underline{t}), g(\hat{X}_{\underline{t}}^K, \bar{X}_{\underline{t}-\tau}^K)(W_t - W_{\underline{t}}) \rangle \\
&+ \sum_{n=0}^{n_t-1} |g(\hat{X}_{t_n}^K, \bar{X}_{t_n-\tau}^K) \Delta W_n|^2 + |g(\hat{X}_{\underline{t}}^K, \bar{X}_{\underline{t}-\tau}^K)(W_t - W_{\underline{t}})|^2.
\end{aligned}
$$

Using the step processes $\bar{X}^K$ and $\tilde{X}^K$ defined previously, the second summand on the RHS of the equation above, can be expressed as a Riemann integral. Similarly the sixth and the seventh terms can be written as an Itô integral, i.e.

$$
\begin{aligned}
|X_t^K|^2 &\le |X_0^K|^2 + 2\alpha \int_0^t (|\bar{X}_s^K|^2 + |\tilde{X}_{s-\tau}^K|^2) ds + 2\beta t \\
&+ 2 \int_0^t \langle \bar{X}_s^K + f(\bar{X}_s^K, \tilde{X}_{s-\tau}^K)[h(\bar{X}_u^K) I_{[0,\underline{t})}(u) \\
&+ (t-\underline{t}) I_{[\underline{t},t]}(u)], g(\bar{X}_s^K, \tilde{X}_{s-\tau}^K) dW_s \rangle \\
&+ \sum_{n=0}^{n_t-1} |g(\bar{X}_{t_n}^K, \tilde{X}_{t_n-\tau}^K) \Delta W_n|^2 + |g(\bar{X}_t^K, \tilde{X}_{t-\tau}^K)(W_t - W_{\underline{t}})|^2.
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
|X_t^K|^p &\le 6^{p/2-1} \Bigg\{ |X_0^K|^p + \left(2\alpha \int_0^t (|\bar{X}_s^K|^2 + |\tilde{X}_{s-\tau}^K|^2) ds \right)^{p/2} + (2\beta t)^{p/2} \\
&+ \Big| 2 \int_0^t \langle \bar{X}_s^K + f(\bar{X}_s^K, \tilde{X}_{s-\tau}^K)[h(\bar{X}_u^K) I_{[0,\underline{t})}(u) \\
&+ (t-\underline{t}) I_{[\underline{t},t]}(u)], g(\bar{X}_s^K, \tilde{X}_{s-\tau}^K) dW_s \rangle \Big|^{p/2} \\
&+ \left(\sum_{n=0}^{n_t-1} |g(\bar{X}_{t_n}^K, \tilde{X}_{t_n-\tau}^K) \Delta W_n|^2 \right)^{p/2} + |g(\bar{X}_t^K, \tilde{X}_{t-\tau}^K)(W_t - W_{\underline{t}})|^p \Bigg\}.
\end{aligned}
$$

Taking the expectation of the supremum, one has

$$
\mathbb{E}\left[\sup_{0\le s\le t} |X_s^K|^p \right] \le 6^{p/2-1}(I_1 + I_2 + I_3 + I_4),
$$

where

$$I_1 := \mathbb{E}|X_0^K|^p + \mathbb{E}\left[\left(2\alpha\int_0^t (|\bar{X}_s^K|^2 + |\tilde{X}_{s-\tau}^K|^2)ds\right)^{p/2}\right] + (2\beta t)^{p/2};$$

$$I_2 := \mathbb{E}\Big[\sup_{0\le s\le t}\Big|2\int_0^s \langle \bar{X}_u^K + f(\bar{X}_u^K, \tilde{X}_{u-\tau}^K)[h(\bar{X}_u^K)I_{[0,\underline{s})}(u)$$
$$+ (s-\underline{s})I_{[\underline{s},s]}(u)], g(\bar{X}_u^K, \tilde{X}_{u-\tau}^K)dW_u\rangle\Big|^{p/2}\Big];$$

$$I_3 := \mathbb{E}\left[\left(\sum_{n=0}^{n_t-1}|g(\bar{X}_{t_n}^K, \tilde{X}_{t_n-\tau}^K)\Delta W_n|^2\right)^{p/2}\right];$$

$$I_4 := \mathbb{E}\left[\sup_{0\le s\le t}|g(\bar{X}_s^K, \tilde{X}_{s-\tau}^K)(W_s - W_{\underline{s}})|^p\right].$$

Now we will establish bounds for each of the four terms above. In the remainder of the proof, C is positive constants, independent of K, that may change from line to line.

Using Hölder's inequality, we have

$$I_1 \le \mathbb{E}|X_0^K|^p + (2\alpha)^{p/2}T^{p/2-1}2^{p/2-1}\int_0^t \mathbb{E}[|\bar{X}_s^K|^p + |\tilde{X}_{s-\tau}^K|^p]ds + (2\beta T)^{p/2}$$
$$\le C\int_0^t \mathbb{E}\left[\sup_{0\le u\le s}|X_u^K|^p\right]ds + C.$$

By the Burkholder-Davis-Gundy (BDG) inequality we obtain

$$I_2 \le 2^{p/2}C\mathbb{E}\Big[\Big(\int_0^t |(\bar{X}_u^K + f(\bar{X}_u^K, \tilde{X}_{u-\tau}^K)[h(\bar{X}_u^K)I_{[0,\underline{t})}(u)$$
$$+ (t-\underline{t})I_{[\underline{t},t]}(u)])g(\bar{X}_u^K, \tilde{X}_{u-\tau}^K)|^2du\Big)^{p/4}\Big]$$

An application of the Hölder inequality yields that

$$I_2 \le 2^{\frac{p}{2}}T^{\frac{p}{4}-1}C\mathbb{E}\Big[\int_0^t \Big|\bar{X}_u^K + f(\bar{X}_u^K, \tilde{X}_{u-\tau}^K)[h(\bar{X}_u^K)I_{[0,\underline{t})}(u)$$
$$+ (t-\underline{t})I_{[\underline{t},t]}(u)]\Big|^{\frac{p}{2}}||g(\bar{X}_u^K, \tilde{X}_{u-\tau}^K)||^{\frac{p}{2}}du\Big] \tag{3.3.14}$$

Now, we bound the integrand of the integral above. Using condition (3.3.5) we obtain

$$|\bar{X}_u^K + f(\bar{X}_u^K, \tilde{X}_{u-\tau}^K)[h(\bar{X}_u^K)I_{[0,\underline{t})}(u) + (t-\underline{t})I_{[\underline{t},t]}(u)]|^2 =$$
$$= |\bar{X}_u^K|^2 + 2[h(\bar{X}_u^K)I_{[0,\underline{t})}(u) + (t-\underline{t})I_{[\underline{t},t]}(u)]\Big[\langle \bar{X}_u^K, f(\bar{X}_u^K, \tilde{X}_{u-\tau}^K)\rangle$$

$$
\begin{aligned}
&+\frac{1}{2}[h(\bar{X}_u^K)I_{[0,\underline{t})}(u)+(t-\underline{t})I_{[\underline{t},t]}(u)]|f(\bar{X}_u^K,\tilde{X}_{u-\tau}^K)|^2\Big]\\
&\le |\bar{X}_u^K|^2+2[h(\bar{X}_u^K)I_{[0,\underline{t})}(u)+(t-\underline{t})I_{[\underline{t},t]}(u)]\left[\alpha\left(|\bar{X}_u^K|^2+|\tilde{X}_{u-\tau}^K|^2\right)+\beta\right]\\
&=(1+2\alpha T)|\bar{X}_u^K|^2+2\alpha T|\tilde{X}_{u-\tau,K}^{1,i,M}|^2+2\beta T.
\end{aligned}
$$

This implies

$$
\begin{aligned}
|\bar{X}_u^K &+ f(\bar{X}_u^K,\tilde{X}_{u-\tau}^K)[h(\bar{X}_u^K)I_{[0,\underline{t})}(u)+(t-\underline{t})I_{[\underline{t},t]}(u)]|^{p/2}\\
&\le 3^{p/4-1}\left[(1+2\alpha T)^{p/4}|\bar{X}_u^K|^{p/2}+(2\alpha T)^{p/4}|\tilde{X}_{u-\tau,K}^{1,i,M}|^{p/2}+(2\beta T)^{p/4}\right]\\
&\le C\left(|\bar{X}_u^K|^{p/2}+|\tilde{X}_{u-\tau,K}^{1,i,M}|^{p/2}+1\right).
\end{aligned}
$$

Also by condition (3.3.3) one can see that

$$
\begin{aligned}
||g(\bar{X}_u^K,\tilde{X}_{u-\tau}^K)||^{p/2}=\left(||g(\bar{X}_u^K,\tilde{X}_{u-\tau}^K)||^2\right)^{p/4}&\le\left[\alpha\left(|\bar{X}_u^K|^2+|\tilde{X}_{u-\tau}^K|^2\right)+\beta\right]^{p/4}\\
&\le C\left(|\bar{X}_u^K|^{p/2}+|\tilde{X}_{u-\tau}^K|^{p/2}+1\right).
\end{aligned}
$$

Substituting the last two inequalities into (3.3.14), we obtain

$$
\begin{aligned}
I_2&\le C\mathbb{E}\left[\int_0^t\left(1+|\bar{X}_u^K|^p+|\tilde{X}_{u-\tau}^K|^p\right)du\right]\\
&\le C+C\left(\int_0^t\mathbb{E}\left[\sup_{0\le u\le s}|X_u^K|^p\right]ds\right).
\end{aligned}
$$

Now we will bound I_3. Note that t_n is a stopping time of the filtration $\{\mathcal{F}_t^W\}$. Define

$$
\mathcal{F}_{t_n}:=\{A\in\mathcal{F}:A\cap\{t_n\le t\}\in\mathcal{F}_t^W\}.
$$

By the strong Markov property of the Brownian motion, $\{B_u:=W_{t_n+u}-W_{t_n},u\ge 0\}$ is a standard Brownian motion independent of $\mathcal{F}_{t_n}$ (page 86, Theorem 6.16 in [30]). Thus

$$
\mathbb{E}[\sup_{0\le u\le s}|W_{t_n+u}-W_{t_n}|^p|\mathcal{F}_{t_n}]=\mathbb{E}[\sup_{0\le u\le s}|B_u|^p]\le Cs^{p/2}.
$$

This implies

$$
\mathbb{E}[\sup_{t_n\le u\le t_{n+1}}|W_u-W_{t_n}|^p|\mathcal{F}_{t_n}]\le Ch_n^{p/2}. \tag{3.3.15}
$$

Combining Jensen's inequality and equation (3.3.15), we arrive at

$$
\begin{aligned}
I_3 &\leq \mathbb{E}\left[\left(\sum_{n=0}^{n_t-1} ||g(\bar{X}_{t_n}^K, \tilde{X}_{t_n-\tau}^K)||^2 |\Delta W_n|^2\right)^{p/2}\right] \\
&= \mathbb{E}\left[\left(\sum_{n=0}^{n_t-1} h_n ||g(\bar{X}_{t_n}^K, \tilde{X}_{t_n-\tau}^K)||^2 \frac{|\Delta W_n|^2}{h_n}\right)^{p/2}\right] \\
&\leq T^{p/2-1}\mathbb{E}\left[\sum_{n=0}^{n_t-1} h_n ||g(\bar{X}_{t_n}^K, \tilde{X}_{t_n-\tau}^K)||^p \frac{E[|\Delta W_n|^p|\mathcal{F}_{t_n}]}{h_n^{p/2}}\right] \\
&\leq CT^{p/2-1}\mathbb{E}\left[\sum_{n=0}^{n_t-1} h_n ||g(\bar{X}_{t_n}^K, \tilde{X}_{t_n-\tau}^K)||^p\right] \\
&\leq CT^{\frac{p}{2}-1}\mathbb{E}\left[\int_0^{\underline{t}} ||g(\bar{X}_s^K, \tilde{X}_{s-\tau}^K)||^p ds\right] \leq CT^{\frac{p}{2}-1}\mathbb{E}\left[\int_0^{t} ||g(\bar{X}_s^K, \tilde{X}_{s-\tau}^K)||^p ds\right].
\end{aligned}
$$

Using condition (3.3.3) and Hölder's inequality, we have

$$
\begin{aligned}
I_3 &\leq CT^{p/2-1}\mathbb{E}\left[\int_0^t \left(||g(\bar{X}_s^K, \tilde{X}_{s-\tau}^K)||^2\right)^{p/2} ds\right] \\
&\leq CT^{p/2-1}\mathbb{E}\left[\int_0^t \left(\alpha(|\bar{X}_s^K|^2 + |\tilde{X}_{s-\tau}^K|^2) + \beta\right)^{p/2} ds\right] \\
&\leq T^{p/2-1}2^{p-2}C\mathbb{E}\left[\int_0^t \left(\alpha^{p/2}(|\bar{X}_s^K|^p + |\bar{X}_{s-\tau}^K|^p) + \beta^{p/2}\right) ds\right] \\
&\leq C + C\int_0^t \mathbb{E}\left[\sup_{0\leq u\leq s} |X_u^K|^p\right] ds.
\end{aligned}
$$

For I_4, using the linear condition (3.3.3), we obtain

$$
\begin{aligned}
I_4 &\leq \mathbb{E}\left[\sup_{0\leq s\leq t} |g(\bar{X}_s^K, \tilde{X}_{s-\tau}^K)(W_s - W_{\underline{s}})|^p\right] \\
&\leq \mathbb{E}\left[\sup_{0\leq s\leq t} \left\{[(\alpha(|\bar{X}_s^K|^p + |\tilde{X}_{s-\tau}^K)|^p) + \beta]\ |(W_s - W_{\underline{s}})|^p\right\}\right] \\
&\leq \mathbb{E}\left[\sum_{n=0}^{n_t-1} [\alpha(|\bar{X}_{t_n}^K|^p + |\tilde{X}_{t_n-\tau}^K|^p) + \beta]\mathbb{E}\left[\sup_{t_n\leq s\leq t_{n+1}} |(W_s - W_{t_n})|^{p/2}|\mathcal{F}_{t_n}\right]\right. \\
&\left. + [\alpha(|\bar{X}_t^K|^p + |\tilde{X}_{t-\tau}^K|^p) + \beta]\mathbb{E}\left[\sup_{\underline{t}\leq s\leq t} |(W_s - W_{\underline{t}})|^{p/2}|\mathcal{F}_{\underline{t}}\right]\right] \\
&\leq C + C\int_0^t E\left[\sup_{0\leq u\leq s} |X_u^K|^p\right] ds.
\end{aligned}
$$

Adding all the bounds for I_1 to I_4, we have that for all $t \in [0,T]$

$$\mathbb{E}\left[\sup_{0\leq s\leq t} |X_s^K|^p\right] \leq C + C\int_0^t E\left[\sup_{0\leq u\leq s} |X_u^K|^p\right] ds,$$

and by the Gronwall inequality we obtain

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t^K|^p\right] \leq C.$$

The result has been proved for $p \geq 4$. For $0 < p < 4$, note that

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t^K|^p\right] I_{\{\sup_{0\leq t\leq T}|X_t^K|\leq 1\}} \leq 1$$

and

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t^K|^p\right] I_{\{\sup_{0\leq t\leq T}|X_t^K|>1\}} \leq \mathbb{E}\left[\sup_{0\leq t\leq T} |X_t^K|^4\right] I_{\{\sup_{0\leq t\leq T}|X_t^K|>1\}} \leq C,$$

where I_A es the indicator function of the set A. Therefore,

$$\begin{aligned}
&\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t^K|^p\right] \\
&= \mathbb{E}\left[\sup_{0\leq t\leq T} |X_t^K|^p\right] I_{\{\sup_{0\leq t\leq T}|X_t^K|\leq 1\}} + \mathbb{E}\left[\sup_{0\leq t\leq T} |X_t^K|^p\right] I_{\{\sup_{0\leq t\leq T}|X_t^K|>1\}} \leq C.
\end{aligned}$$

□

Remark 3.3.2. Note that assuming that T was attainable, we have proved the boundedness of the pth moments without using the auxiliary scheme. The only reason why we needed to work with a bounded scheme was to show that $\inf_{|x|\leq K} h^\delta(x)$ is strictly positive and therefore T is attainable.

Proof of Theorem 3.3.5. By Lemma 3.3.6 and the Markov inequality

$$\mathbb{P}(\sup_{0\leq t\leq T} |X_t| < K) = 1 - \mathbb{P}(\sup_{0\leq t\leq T} |X_t^K| \geq K) \geq 1 - \frac{\mathbb{E}[\sup_{0\leq t\leq T}|X_t^K|^4}{K^4} \geq 1 - \frac{C}{K^4}.$$

Thus

$$\lim_{K\to\infty} \mathbb{P}(\sup_{0\leq t\leq T} |X_t| < K) = 1,$$

This means that $\sup_{0\le t\le T}|X_t| < \infty$ a.s., i.e. for almost all $\omega \in \Omega$ there exist a K_ω such that

$$\sup_{0\le t\le T}|X_t(\omega)| \le K_\omega. \tag{3.3.16}$$

Since h^δ is continuous and strictly positive, $\inf_{|x|\le K_\omega} h^\delta(x) > 0$. This implies that for almost every $\omega \in \Omega$

$$\liminf_{n\to\infty} h_n^\delta(\omega) = \liminf_{n\to\infty} h^\delta(X_{t_n}(\omega)) \ne 0,$$

so $\lim_{n\to\infty} t_n(\omega) = \sum_{n=0}^\infty h_n^\delta(\omega) = \infty$ a.s. and T is attainable. Also, for all ω and all $0 < K_1 \le K_2$, we have

$$\begin{aligned}\sup_{0\le t\le T}|X_t^{K_1}(\omega)| &= \min(\sup_{0\le t\le T}|X_t(\omega)|, K_1) \le \min(\sup_{0\le t\le T}|X_t(\omega)|, K_2)\\ &= \sup_{0\le t\le T}|X_t^{K_2}(\omega).| \end{aligned}\tag{3.3.17}$$

Equations (3.3.16) and (3.3.17) imply that

$$\lim_{K\to\infty}\sup_{0\le t\le T}|X_t^K| = \sup_{0\le t\le T}|X_t| \quad \text{a.s.} \tag{3.3.18}$$

This together with Lemma 3.3.6, yields

$$\mathbb{E}\left[\sup_{0\le t\le T}|X_t|^p\right] = \lim_{K\to\infty}\mathbb{E}\left[\sup_{0\le t\le T}|X_t^K|^p\right] \le C.$$

The proof is complete for $p \ge 4$. For $0 \le p < 4$, The required assertion follows from the Hölder inequality. □

3.3.2 Strong convergence of the numerical solutions

In order to prove the strong convergence of the approximate solution (3.2.4) to the exact solution of the SDDE (3.2.1), we need the following lemma and corollary.

Lemma 3.3.7. *Let the SDDE* (3.2.1) *and the function h^δ satisfy Assumption 3.3.1 and 3.3.2 respectively. Assume also that the function f satisfies the (global) linear growth condition, i.e. there exist a constant $C_1 \ge 0$ such that for all $x, y \in \mathbb{R}^d$,*

$$|f(x,y)|^2 \le C_1(|x|^2 + |y|^2 + 1). \tag{3.3.19}$$

Then there exists a positive constant C such that for all $t \in [0,T]$.

$$\mathbb{E}|X_t - \bar{X}_t|^2 \leq C\delta T, \tag{3.3.20}$$

$$\mathbb{E}|X_t - \tilde{X}_t|^2 \leq C\delta T. \tag{3.3.21}$$

Proof. Let $t \in [0,T]$. Let r be such that $t_r \leq t \leq t_{r+1}$. Then by definition we have $X_{t_r} = \bar{X}_{t_r} = \bar{X}_t$. Thus

$$X_t = \bar{X}_t + \int_{t_r}^{t} f(\bar{X}_s, \tilde{X}_s)ds + \int_{t_r}^{t} g(\bar{X}_s, \tilde{X}_s)dW_s.$$

This together with (3.3.19),(3.3.3), Assumption 3.3.2 and Theorem 3.3.5 imply that

$$\begin{aligned}
\mathbb{E}|X_t - \bar{X}_t|^2 &\leq 2\mathbb{E}\left|\int_{t_r}^{t} f(\bar{X}_s, \tilde{X}_s)ds\right|^2 + 2\mathbb{E}\left|\int_{t_r}^{t} g(\bar{X}_s, \tilde{X}_s)dW_s\right|^2 \\
&\leq 2\mathbb{E}[C_1(h_r^\delta)^2(1 + 2\sup_{t_r \leq s \leq t}|X_s|^2 + ||\xi||)] + 2\mathbb{E}[\alpha h_r^\delta(2\sup_{t_r \leq s \leq t}|X_s|^2 + ||\xi||) + \beta] \\
&\leq 4(\delta T)^2(1 + \mathbb{E}[\sup_{t_r \leq s \leq t}|X_s|^2] + \mathbb{E}||\xi||) + 4\alpha\delta T(\mathbb{E}[\sup_{t_r \leq s \leq t}|X_s|^2] + \mathbb{E}||\xi||) + \beta] \\
&\leq C\delta T.
\end{aligned}$$

To prove assertion (3.3.21), we first prove that there is a constant C such that for all $t \in [0,T]$

$$\mathbb{E}|\tilde{X}_t - \bar{X}_t|^2 \leq C\delta T. \tag{3.3.22}$$

Let $t \in [0,T]$. Let k and n be such that $t_k \leq t < t_{k+1}$ and $t_n - \tau \leq t \leq t_{n+1} - \tau$ respectively. Let $r, 0 \leq r \leq k$ be such that $t_{k-r} \leq t_n - \tau \leq t_{k-r+1}$. From (3.2.2) and the definitions of the step processes $\bar{X}$ and $\tilde{X}$, one can see that

$$\begin{aligned}
\hat{X}_{t_k} &= \hat{X}_{t_{k-r}} + \sum_{i=0}^{r-1}[f(\bar{X}_{t_{k-r+i}}, \bar{X}_{t_{k-r+i}-\tau})h_{k-r+i} + g(\bar{X}_{t_{k-r+i}}, \bar{X}_{t_{k-r+i}-\tau})\Delta W_{k-r+i}] \\
&= \hat{X}_{t_{k-r}} + \sum_{i=0}^{r-1}\int_{t_{k-r+i}}^{t_{k-r+i+1}} f(\bar{X}_s, \tilde{X}_{s-\tau})ds + \sum_{i=0}^{r-1}\int_{t_{k-r+i}}^{t_{k-r+i+1}} g(\bar{X}_s, \tilde{X}_{s-\tau})dW_s \\
&= \hat{X}_{t_{k-r}} + \int_{t_{k-r}}^{t_k} f(\bar{X}_s, \tilde{X}_{s-\tau})ds + \int_{t_{k-r}}^{t_k} g(\bar{X}_s, \tilde{X}_{s-\tau})dW_s.
\end{aligned}$$

Note that $\bar{X}_t = \hat{X}_{t_k}$ and $\hat{X}_{t_{k-r}} = \bar{X}_{t_{k-r}} = \bar{X}_{t_n-\tau} = \tilde{X}_{t_n-\tau} = \tilde{X}_t$, we have that

$$\bar{X}_t = \tilde{X}_t + \int_{t_{k-r}}^{t_k} f(\bar{X}_s, \tilde{X}_{s-\tau})ds + \int_{t_{k-r}}^{t_k} g(\bar{X}_s, \tilde{X}_{s-\tau})dW_s.$$

Also, we have that

$$t_k - t_{k-r} \leq (t_{n+1} - \tau) - (t_n - \tau) + h^\delta_{k-r} = h^\delta_n + h^\delta_{k-r} \leq 2\delta T.$$

Therefore, by (3.3.19),(3.3.3), Assumption 3.3.2 and Theorem 3.3.5 we have that

$$\begin{aligned}
\mathbb{E}|\bar{X}_t - \tilde{X}_t|^2 &\leq 2\mathbb{E}\left|\int_{t_{k-r}}^{t_k} f(\bar{X}_s, \tilde{X}_{s-\tau})ds\right|^2 + 2\mathbb{E}\left|\int_{t_{k-r}}^{t_k} g(\bar{X}_s, \tilde{X}_{s-\tau})dW_s\right|^2 \\
&\leq 2\mathbb{E}[C_1(t_k - t_{k-r})^2(1 + 2\sup_{t_k \leq s \leq t}|X_s|^2 + ||\xi||)] \\
&+ 2\mathbb{E}[\alpha(t_k - t_{k-r})(2\sup_{t_k \leq s \leq t}|X_s|^2 + ||\xi||) + \beta] \\
&\leq 4(\delta T)^2(1 + \mathbb{E}[\sup_{t_k \leq s \leq t}|X_s|^2] + \mathbb{E}||\xi||) + 4\alpha\delta T(\mathbb{E}[\sup_{t_k \leq s \leq t}|X_s|^2] + \mathbb{E}||\xi||) + \beta] \\
&\leq C\delta T.
\end{aligned}$$

This together with (3.3.20) imply that

$$\mathbb{E}|X_t - \tilde{X}_t|^2 = \mathbb{E}|X_t - \bar{X}_t|^2 + \mathbb{E}|\bar{X}_t - \tilde{X}_t|^2 \leq C\delta T.$$

□

In our attempt to prove the strong convergence using the local Lipschitz condition instead of the global one, we introduce the stopping times

$$\tau_m := \inf\{t \geq 0 : |Y_t| \geq m\}, \quad \sigma_m := \inf\{t \geq 0 : |X_t| \geq m\}$$

and $\upsilon_m := \tau_m \wedge \sigma_m$. As usual we set $\inf \emptyset = \infty$. In the next corollary, we relax the global linear condition imposed to f in the previous lemma and use instead the local Lipschitz condition.

Corollary 3.3.8. *Let the SDDE* (3.2.1) *and the function h^δ satisfy Assumption 3.3.1 and 3.3.2 respectively. Then there exists a positive constant C_m such that for all $t \in [0, T]$.*

$$\mathbb{E}|X_{t\wedge\upsilon_m} - \bar{X}_{t\wedge\upsilon_m}|^2 \leq C_m\delta T, \tag{3.3.23}$$

$$\mathbb{E}|X_{t\wedge\upsilon_m-\tau} - \tilde{X}_{t\wedge\upsilon_m-\tau}|^2 \leq C_m\delta T. \tag{3.3.24}$$

Proof. The processes $X_{t\wedge\upsilon_m}, \bar{X}_{t\wedge\upsilon_m}$ and $\tilde{X}_{t\wedge\upsilon_m}$ are bounded by m. Thus, the local Lipschitz condition (3.3.1) implies condition (3.3.19). Therefore the corollary follows directly from Lemma 3.3.7. □

Theorem 3.3.9. *If the SDDE* (3.2.1) *and the function* h^δ *satisfy Assumption 3.3.1 and 3.3.2 respectively, then for all* $p > 0$

$$\lim_{\delta\to 0}\mathbb{E}\left[\sup_{0\le t\le T}|X_t - Y_t|^p\right] = 0.$$

Proof. One can see that

$$\mathbb{E}[\sup_{0\le t\le T}|Y_t - X_t|^2] = \mathbb{E}[\sup_{0\le t\le T}|Y_t - X_t|^2 I_{\{\tau_m > T \text{ and } \sigma_m > T\}}] \qquad (3.3.25)$$
$$+ \mathbb{E}[\sup_{0\le t\le T}|Y_t - X_t|^2 I_{\{\tau_m \le T \text{ or } \sigma_m \le T\}}] =: R_1 + R_2,$$

where I_A es the indicator function of the set A. In order to bound R_1, we combine the definitions of the continuous-time approximation (3.2.4) and the exact solution (3.2.1) to obtain

$$\begin{aligned}
&|Y_{t\wedge\upsilon_m} - X_{t\wedge\upsilon_m}|^2\\
&= \left|\int_0^{t\wedge\upsilon_m}[f(Y_s, Y_{s-\tau}) - f(\hat{X}_s, \tilde{X}_{s-\tau})]ds + \int_0^{t\wedge\upsilon_m}[g(Y_s, Y_{s-\tau}) - g(\hat{X}_s, \tilde{X}_{s-\tau})]dW_s\right|^2\\
&\le 2T\int_0^{t\wedge\upsilon_m}|f(Y_s, Y_{s-\tau}) - f(\hat{X}_s, \tilde{X}_{s-\tau})|^2 ds\\
&+ 2\left|\int_0^{t\wedge\upsilon_m}[g(Y_s, Y_{s-\tau}) - g(\hat{X}_s, \tilde{X}_{s-\tau})]dW_s\right|^2
\end{aligned}$$

Thus, for any $t_1 \le T$,

$$\begin{aligned}
&\mathbb{E}[\sup_{0\le t\le t_1}|Y_{t\wedge\upsilon_m} - X_{t\wedge\upsilon_m}|^2]\\
&\le 2T\mathbb{E}\left[\int_0^{t\wedge\upsilon_m}|f(Y_s, Y_{s-\tau}) - f(\hat{X}_s, \tilde{X}_{s-\tau})|^2 ds\right]\\
&+ 8\mathbb{E}\left[\int_0^{t\wedge\upsilon_m}|g(Y_s, Y_{s-\tau}) - g(\hat{X}_s, \tilde{X}_{s-\tau})|^2 ds\right],
\end{aligned}$$

where we have used the Doob martingale inequality in the second summand. Using the local Lipschitz condition (3.3.1) in the RHS of the previous equation and then, adding and subtracting X_t twice yields

$$
\begin{aligned}
&\mathbb{E}[\sup_{0\le t\le t_1} |Y_{t\wedge \upsilon_m} - X_{t\wedge \upsilon_m}|^2] \\
&\le C_m \left(\int_0^{t_1} \mathbb{E}|Y_{s\wedge \upsilon_m} - X_{s\wedge \upsilon_m}|^2 ds + \int_0^{t_1} \mathbb{E}|Y_{s\wedge \upsilon_m - \tau} - X_{s\wedge \upsilon_m - \tau}|^2 ds \right) \\
&+ C_m \left(\int_0^{t_1} \mathbb{E}|X_{s\wedge \upsilon_m} - \bar{X}_{s\wedge \upsilon_m}|^2 ds + \int_0^{t_1} \mathbb{E}|X_{s\wedge \upsilon_m - \tau} - \tilde{X}_{s\wedge \upsilon_m - \tau}|^2 ds \right),
\end{aligned}
$$

where C_m is a positive constant that depends on T and m. By Corollary 3.3.8, we obtain

$$
\begin{aligned}
&\mathbb{E}[\sup_{0\le t\le t_1} |Y_{t\wedge \upsilon_m} - X_{t\wedge \upsilon_m}|^2] \\
&\le C_m \left(\int_0^{t_1} \mathbb{E}|Y_{s\wedge \upsilon_m} - X_{s\wedge \upsilon_m}|^2 ds + \int_0^{t_1} \mathbb{E}|Y_{s\wedge \upsilon_m - \tau} - X_{s\wedge \upsilon_m - \tau}|^2 ds \right) + C_m \delta.
\end{aligned}
$$

The Gronwall inequality yields

$$
R_1 = \mathbb{E}[\sup_{0\le t\le T} |Y_{t\wedge \upsilon_m} - X_{t\wedge \upsilon_m}|^2] \le C_m \delta.
$$

Proceeding in exactly the same way as in [23], one can see that for all $\alpha, \beta, \eta, \mu > 0$ we have

$$
R_2 \le \frac{2^{p+1}\eta C}{p} + \frac{2(p-2)C}{p\eta^{2/(p-2)} m^p}
$$

where $\bar{C}$ is a positive constant. Substituting the estimates of R_1 and R_2 into (3.3.25), we obtain

$$
\mathbb{E}[\sup_{0\le t\le T} |Y_t - X_t|^2] \le C_m \delta + \frac{2^{p+1}\eta C}{p} + \frac{2(p-2)C}{p\eta^{2/(p-2)} m^p}.
$$

Now, given any $\epsilon > 0$, we can find an η sufficiently small so

$$
\frac{2^{p+1}\eta C}{p} < \frac{\epsilon}{3},
$$

and then m large enough so

$$
\frac{2(p-2)C}{p\eta^{2/(p-2)} m^p} < \frac{\epsilon}{3},
$$

and finally δ small enough such that

$$
\delta C_m < \frac{\epsilon}{3}.
$$

The proof is complete. □

3.3.3 Order of convergence

Now we investigate the order of convergence of the adaptive EM numerical solutions.

Assumption 3.3.10. *There exists a constant $L > 0$ such that for all $x, y, \bar{x}, \bar{y} \in \mathbb{R}^m$, f satisfies the one-sided Lipschitz condition*

$$2\langle x - \bar{x}, f(x, y) - f(\bar{x}, \bar{y})\rangle \leq L(|x - \bar{x}|^2 + |y - \bar{y}|^2) \tag{3.3.26}$$

and g satisfies the (global) Lipschitz condition

$$||g(x, y) - g(\bar{x}, \bar{y})||^2 \leq L(|x - \bar{x}|^2 + |y - \bar{y}|^2). \tag{3.3.27}$$

In addition f satisfies the polynomial growth Lipschitz condition: there exist constants $\gamma, \lambda, q > 0$ such that for all $x, y, \bar{x}, \bar{y} \in \mathbb{R}^m$

$$|f(x, y) - f(\bar{x}, \bar{y})| \leq (\gamma(|x|^q + |y|^q + |\bar{x}|^q + |\bar{y}|^q) + \lambda)(|x - \bar{x}| + |y - \bar{y}|). \tag{3.3.28}$$

Furthermore, for any $s, t \in [-\tau, 0]$ and $q > 0$, there exists a positive constant Λ such that

$$\mathbb{E}||\xi(t) - \xi(s)|| \leq \Lambda|t - s|^q. \tag{3.3.29}$$

Theorem 3.3.11. *If the SDDE* (3.2.1) *satisfies Assumption 3.3.10 and the time-step function h satisfies Assumption 3.3.2, then for all $p > 0$, there exists a positive constant C independent of δ such that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X_t - Y_t|^p\right] \leq C\delta^{p/2}.$$

Proof. The proof is similar to that of SDEs given in [14]. We only give the proof for $p \geq 4$; the result for $0 \leq p < 4$ follows from Hölder's inequality. Define $e_t := Y_t - X_t, 0 \leq t \leq T$. Hence

$$e_t = \int_0^t [f(Y_s, Y_{s-\tau}) - f(\bar{X}_s, \tilde{X}_{s-\tau})]ds + \int_0^t [g(Y_s, Y_{s-\tau}) - g(\bar{X}_s, \tilde{X}_{s-\tau})]dW_s.$$

Applying Itô's formula we obtain

$$|e_t|^2 \leq 2\int_0^t \langle e_s, f(Y_s, Y_{s-\tau}) - f(\bar{X}_s, \tilde{X}_{s-\tau})\rangle ds + \int_0^t |g(Y_s, Y_{s-\tau}) - g(\bar{X}_s, \tilde{X}_{s-\tau})|^2 ds$$

$$
\begin{aligned}
&+ 2\int_0^t \langle e_s, (g(Y_s, Y_{s-\tau}) - g(\bar{X}_s, \tilde{X}_{s-\tau}))dW_s\rangle \\
&\le 2\int_0^t \langle e_s, f(Y_s, Y_{s-\tau}) - f(X_s, X_{s-\tau})\rangle ds + 2\int_0^t \langle e_s, f(X_s, X_{s-\tau}) - f(\bar{X}_s, \tilde{X}_{s-\tau})\rangle ds \\
&+ \int_0^t |g(Y_s, Y_{s-\tau}) - g(\bar{X}_s, \tilde{X}_{s-\tau})|^2 ds + 2\int_0^t \langle e_s, (g(Y_s, Y_{s-\tau}) - g(\bar{X}_s, \tilde{X}_{s-\tau}))dW_s\rangle.
\end{aligned} \tag{3.3.30}
$$

Using condition (3.3.26) we get

$$
\begin{aligned}
2\langle e_s, f(Y_s, Y_{s-\tau}) - f(X_s, X_{s-\tau})\rangle &\le L(|Y_s - X_s|^2 + |Y_{s-\tau} - X_{s-\tau}|^2) \\
&= L(|e_s|^2 + |e_{s-\tau}|^2).
\end{aligned} \tag{3.3.31}
$$

Condition (3.3.28) implies that

$$
\begin{aligned}
&|\langle e_s, f(X_s, X_{s-\tau}) - f(\bar{X}_s, \tilde{X}_{s-\tau})\rangle| \le |e_s| \; |f(X_s, X_{s-\tau}) - f(\bar{X}_s, \tilde{X}_{s-\tau})| \\
&\le |e_s| Q(X_s, X_{s-\tau}, \bar{X}_s, \tilde{X}_{s-\tau})(|X_s - \bar{X}_s| + |X_{s-\tau} - \tilde{X}_{s-\tau}|) \\
&\le \frac{1}{2}|e_s|^2 + \frac{1}{2}Q(X_s, X_{s-\tau}, \bar{X}_s, \tilde{X}_{s-\tau})^2 \; 2(|X_s - \bar{X}_s|^2 + |X_{s-\tau} - \tilde{X}_{s-\tau}|^2),
\end{aligned} \tag{3.3.32}
$$

where $Q(x, y, \bar{x}, \bar{y}) := \gamma(|x|^q + |y|^q + |\bar{x}|^q + |\bar{y}|^q) + \lambda$. In addition, condition (3.3.27) implies that

$$
\begin{aligned}
||g(Y_s, Y_{s-\tau}) - g(\bar{X}_s, \tilde{X}_{s-\tau})||^2 &\le L(|Y_s - \bar{X}_s|^2 + |Y_{s-\tau} - \tilde{X}_{s-\tau}|^2) \\
&= L(|Y_s - X_s + X_s - \bar{X}_s|^2 + |Y_{s-\tau} - X_{s-\tau} + X_{s-\tau} - \tilde{X}_{s-\tau}|^2) \\
&\le 2L(|e_s|^2 + |e_{s-\tau}|^2 + |X_s - \bar{X}_s|^2 + |X_{s-\tau} - \tilde{X}_{s-\tau}|^2).
\end{aligned} \tag{3.3.33}
$$

Substituting (3.3.31), (3.3.32) and (3.3.33) in (3.3.30), we have

$$
\begin{aligned}
|e_t|^2 \le &\int_0^t \left[(3L+1)|e_s|^2 + 3L|e_{s-\tau}|^2\right] ds \\
&+ 2\int_0^t [Q(X_s, X_{s-\tau}, \bar{X}_s, \tilde{X}_{s-\tau})^2 + L](|X_s - \bar{X}_s|^2 + |X_{s-\tau} - \tilde{X}_{s-\tau}|^2) ds \\
&+ 2\int_0^t \langle e_s, (g(Y_s, Y_{s-\tau}) - g(\bar{X}_s, \tilde{X}_{s-\tau}))dW_s\rangle.
\end{aligned}
$$

Using Hölder's inequality yields

$$
|e_t|^p \le (6T)^{p/2-1} \int_0^t ((3L+1)^{p/2}|e_s|^p + (2L)^{p/2}|e_{s-\tau}|^p) ds
$$

$$+ (6T)^{p/2-1}2^{p/2}\int_0^t [Q(X_s, X_{s-\tau}, \bar{X}_s, \tilde{X}_{s-\tau}) + L]^{p/2}(|X_s - \bar{X}_s|^p + |X_{s-\tau} - \tilde{X}_{s-\tau}|^p)ds$$

$$+ 3^{p/2-1}2^{p/2}\left|\int_0^t \langle e_s, (g(Y_s, Y_{s-\tau}) - g(\bar{X}_s, \tilde{X}_{s-\tau}))dW_s\rangle\right|^{p/2}.$$

In the remainder of the proof, C is positive constant, independent of δ, that may change from line to line.

Taking the supremum on each side of the previous inequality and then the expectation yields

$$\mathbb{E}\left[\sup_{0\le s\le t}|e_s|^p\right] \le J_1 + J_2 + J_3,$$

where

$$J_1 := C\int_0^t \mathbb{E}\left[\sup_{0\le u\le s}|e_u|^p\right]ds;$$

$$J_2 := C\int_0^t \mathbb{E}\left[[Q(X_s, X_{s-\tau}, \bar{X}_s, \tilde{X}_{s-\tau}) + L]^{p/2}(|X_s - \bar{X}_s|^p + |X_{s-\tau} - \tilde{X}_{s-\tau}|^p)\right]ds;$$

$$J_3 := C\mathbb{E}\left[\sup_{0\le s\le t}\left|\int_0^s \langle e_u, (g(Y_u, Y_{u-\tau}) - g(\bar{X}_u, \tilde{X}_{u-\tau}))dW_u\rangle\right|^{p/2}\right].$$

For J_2, by Hölder's inequality one has

$$J_2 \le C\int_0^t \Big(\mathbb{E}\left[[Q(X_s, X_{s-\tau}, \bar{X}_s, \tilde{X}_{s-\tau}) + L]^p\right] \times \mathbb{E}\left[(|X_s - \bar{X}_s|^{2p} + |X_{s-\tau} - \tilde{X}_{s-\tau}|^{2p})\right]\Big)^{1/2}ds. \tag{3.3.34}$$

By Theorem 3.3.5 there exists a constant C such that

$$\mathbb{E}\left[[Q(X_s, X_{s-\tau}, \bar{X}_s, \tilde{X}_{s-\tau}) + L]^p\right] \le C. \tag{3.3.35}$$

Let $\underline{s} := \max\{t_n : t_n \le s\}$. From (3.2.4), we can write

$$X_s - \bar{X}_s = f(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})(s - \underline{s}) + g(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})(W_s - W_{\underline{s}}).$$

Thus, by Hölder inequality

$$\mathbb{E}|X_s - \bar{X}_s|^{2p} = \mathbb{E}|f(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})(s - \underline{s}) + g(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})(W_s - W_{\underline{s}})|^{2p}$$
$$\le 2^{2p-1}\mathbb{E}|f(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})(s - \underline{s})|^{2p} + 2^{2p-1}\mathbb{E}|g(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})(W_s - W_{\underline{s}})|^{2p}$$

$$\leq 2^{2p-1}(\mathbb{E}[f(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})]^{4p}\mathbb{E}[(s-\underline{s})^{4p})^{1/2}$$
$$+ 2^{2p-1}(\mathbb{E}[g(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})]^{4p}\mathbb{E}[(W_s - W_{\underline{s}})^{4p}])^{1/2}. \tag{3.3.36}$$

By Assumption 3.3.2 we have

$$E[(s-\underline{s})^{4p}] \leq \mathbb{E}[(h^{\delta}_{\underline{s}})^{4p}] \leq (\delta T)^{4p} \leq \delta^{2p}T^{4p} \tag{3.3.37}$$

and by condition (3.3.15), we get

$$\mathbb{E}[(W_s - W_{\underline{s}})^{4p}] \leq C(\delta T)^{2p}. \tag{3.3.38}$$

Also it follows from the global Lipschitz condition 3.3.27 that

$$||g(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})||^{4p} \leq \frac{1}{2^{2p}}K^{2p}(|\bar{X}_{\underline{s}}|^2 + |\tilde{X}_{\underline{s}-\tau}|^2)^{2p} + C \tag{3.3.39}$$
$$\leq C(|\bar{X}_{\underline{s}}|^{4p} + |\tilde{X}_{\underline{s}-\tau}|^{4p} + 1)$$

and from the polynomial growth condition that

$$|f(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})|^{4p} \leq \Big[(\gamma(|\bar{X}_{\underline{s}}|^q + |\tilde{X}_{\underline{s}-\tau})|^q) + \mu)(|\bar{X}_{\underline{s}}| + |\tilde{X}_{\underline{s}-\tau})|) + f(0,0)\Big]^{4p} \tag{3.3.40}$$
$$\leq C(|\bar{X}_{\underline{s}}|^{4p(q+1)} + |\tilde{X}_{\underline{s}-\tau})|^{4p(q+1)} + 1),$$

so by Theorem 3.3.5, there exists a constant C such that

$$\mathbb{E}[|f(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})|^{4p}] \leq C \text{ and } \mathbb{E}[|g(\bar{X}_{\underline{s}}, \tilde{X}_{\underline{s}-\tau})|^{4p}] \leq C.$$

Substituting these last two expressions together with (3.3.37) and (3.3.38) into (3.3.36), we obtain

$$\mathbb{E}|X_s - \bar{X}_s|^{2p} \leq C\delta^p. \tag{3.3.41}$$

Using (3.3.39) and (3.3.40), and proceeding in exactly the same way as in Lemma 3.3.7, yields $\mathbb{E}|X_{s-\tau} - \tilde{X}_{s-\tau}|^{2p} \leq C\delta^p$. Using this fact together with (3.3.41) and (3.3.35) in (3.3.34), we obtain that $J_2 \leq C\delta^{p/2}$.

Now we estimate J_3. By the BDG and Hölder's inequalities one can see that

$$J_3 \leq C\mathbb{E}\left[\left(\int_0^t |e_s|^2 \ |(g(Y_s, Y_{s-\tau}) - g(\bar{X}_s, \tilde{X}_{s-\tau}))|^2 ds\right)^{p/4}\right]$$

$$
\begin{aligned}
&\leq C\mathbb{E}\left[\int_0^t |e_s|^{p/2}(|\bar{X}_s - Y_s|^{p/2} + |\tilde{X}_{s-\tau} - Y_{s-\tau}|^{p/2})ds\right] \\
&\leq C\mathbb{E}\left[\int_0^t \frac{1}{2}|e_s|^p + |\bar{X}_s - Y_s|^p + |\tilde{X}_{s-\tau} - Y_{s-\tau}|^p ds\right] \\
&\leq C\mathbb{E}\Big[\int_0^t |e_s|^p + (|\bar{X}_s - X_s|^p + |X_s - Y_s|^p + |\tilde{X}_{s-\tau} - X_{s-\tau}|^p \\
&\qquad\qquad\qquad\qquad\qquad\qquad + |X_{s-\tau} - Y_{s-\tau}|^p)ds\Big] \\
&\leq C\mathbb{E}\left[\int_0^t |e_s|^p + |e_{s-\tau}|^p + (|\bar{X}_s - X_s|^p + |\tilde{X}_{s-\tau} - X_{s-\tau}|^p)ds\right].
\end{aligned}
$$

By the same argument we used with J_2 we know that

$$
\mathbb{E}\left[(|\bar{X}_s - X_s|^p + |\tilde{X}_{s-\tau} - X_{s-\tau}|^p)\right] \leq C\delta^{p/2}.
$$

Thus

$$
J_3 \leq C\int_0^t \mathbb{E}\left[\sup_{0\leq u\leq s} |e_u|^p\right] ds + C\delta^{p/2}.
$$

Collecting the bounds for J_1, J_2 and J_3, we conclude that there exist a constant C such that

$$
\mathbb{E}\left[\sup_{0\leq t\leq T} |e_t|^p\right] \leq C\int_0^t \mathbb{E}\left[\sup_{0\leq u\leq s} |e_u|^p\right] ds + C\delta^{p/2}.
$$

The required assertion follows from the Gronwall inequality. □

3.4 Convergence of the numerical solutions on infinite time interval

In this section we will study the convergence of the numerical solutions on the time interval $[0, \infty)$. The assumptions will be stronger than the ones on the finite time interval.

Assumption 3.4.1. *The functions f and g satisfy the local Lipschitz condition: for every $R > 0$ there exists a positive constant C_R such that*

$$
|f(x,y) - f(\bar{x},\bar{y})| + ||g(x,y) - g(\bar{x},\bar{y})|| \leq C_R(|x-\bar{x}| + |y-\bar{y}|) \tag{3.4.1}
$$

for all $x, y, \bar{x}, \bar{y} \in \mathbb{R}^m$ *with* $|x|, |y|, |\bar{x}|, |\bar{y}| \leq R$*. Furthermore, there exists constants* $\alpha_1 > \alpha_2 \geq 0$ *and* $\beta > 0$*, such that for all* $x, y \in \mathbb{R}^d$*,* f *satisfies the dissipative one-sided linear growth condition:*

$$\langle x, f(x, y)\rangle \leq -\alpha_1|x|^2 + \alpha_2|y|^2 + \beta, \tag{3.4.2}$$

and g *is globally bounded:*

$$||g(x, y)||^2 \leq \beta. \tag{3.4.3}$$

Assumption 3.4.2. *For every* δ*, the time step function* $h^\delta : \mathbb{R}^d \to \mathbb{R}^+$*, is continuous and uniformly bounded by* h^δ_{max}*, where* $h^\delta_{max} \in (0, \infty)$*.*
Furthermore, there exist constants $\alpha_1 > \alpha_2 \geq 0$ *and* $\beta > 0$*, such that for all* $x, y \in \mathbb{R}^d$*.*

$$\langle x, f(x, y)\rangle + \frac{1}{2}h^\delta(x)|f(x, y)|^2 \leq -\alpha_1|x|^2 + \alpha_2|y|^2 + \beta. \tag{3.4.4}$$

3.4.1 The boundedness of the pth moments of the exact and the numerical solutions

The next lemma shows the boundedness of the pth moments of the exact solution on a non-bounded time interval.

Lemma 3.4.3. *If the SDDE* (3.2.1) *satisfies Assumption 3.4.1, then for every* $p > 0$ *there exists a positive constant* C *(which depends on* p*) such that for all* $t \geq 0$

$$\mathbb{E}\left[|Y_t|^p\right] \leq C. \tag{3.4.5}$$

Proof. A proof can be found in [41]. □

Now, we investigate the pth moments of numerical solution. The proof about attainability given for the finite time interval, is valid for the infinite time interval $[-\tau, \infty)$.

Theorem 3.4.4. *If the SDDE* (3.2.1) *and the function* h^δ *satisfy Assumption 3.4.1 and 3.4.2 respectively, then for all* $p > 0$ *there exists a constant* C *dependent on* $h_{max}, \beta, \alpha_1, \alpha_2$ *and* p*, but independent of* δ *and* t*, such that for all* $t \geq 0$*,*

$$\mathbb{E}\left[|X_t|^p\right] \leq C. \tag{3.4.6}$$

Proof. The proof is given for $p \geq 4$. For $0 < p < 4$, the result holds from Hölder's inequality. Fix t and define $\underline{t} := \max\{t_n : t_n \leq t\}$, $\hat{\underline{t}} := \max\{t_n : t_n \leq t - \tau\}$ and $n_t := \max\{n : t_n \leq t\}$. Taking squared norms in (3.2.2), we have that for $n = 0$ to $n = n_t$,

$$|\hat{X}_{t_{n+1}}|^2 = |\hat{X}_{t_n}|^2 + 2h_n(\langle \hat{X}_{t_n}, f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\rangle + \frac{1}{2}h_n|f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2) \\ + 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle + |g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n|^2.$$

Note that, since it is irrelevant in this proof, we have dropped the term "δ" in the adaptive time-step "h_n^δ" to ease the notation.Using conditions (3.4.4) and (3.4.3), we obtain

$$|\hat{X}_{t_{n+1}}|^2 \leq |\hat{X}_{t_n}|^2 - 2h_n\alpha_1|\hat{X}_{t_n}|^2 + 2h_n\alpha_2|\bar{X}_{t_n-\tau}|^2 + 2h_n\beta \\ + 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle + \beta|\Delta W_n|^2.$$

Multiplying both sides by $e^{2\alpha_1 t_{n+1}}$ yields

$$e^{2\alpha_1 t_{n+1}}|\hat{X}_{t_{n+1}}|^2 \leq e^{2\alpha_1 t_{n+1}}|\hat{X}_{t_n}|^2 - 2h_n\alpha_1 e^{2\alpha_1 t_{n+1}}|\hat{X}_{t_n}|^2 + 2h_n\alpha_2 e^{2\alpha_1 t_{n+1}}|\bar{X}_{t_n-\tau}|^2 \\ + 2h_n\beta e^{2\alpha_1 t_{n+1}} + 2e^{2\alpha_1 t_{n+1}}\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle \\ + e^{2\alpha_1 t_{n+1}}\beta|\Delta W_n|^2.$$

Now, taking into account that $t_{n+1} = t_n + h_n$ and using the fact that for all $x \in \mathbb{R}$, $1 + x \leq e^x$ with $x = -2h_n\alpha_1$, we obtain

$$e^{2\alpha_1 t_{n+1}}|\hat{X}_{t_{n+1}}|^2 \leq e^{2\alpha_1 t_n}|\hat{X}_{t_n}|^2 + 2h_n\alpha_2 e^{2\alpha_1 t_{n+1}}|\bar{X}_{t_n-\tau}|^2 + 2h_n\beta e^{2\alpha_1 t_{n+1}} \\ + 2e^{2\alpha_1 t_{n+1}}\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle + e^{2\alpha_1 t_{n+1}}\beta|\Delta W_n|^2.$$

Solving the recurrence, we have

$$e^{2\alpha_1 \underline{t}}|\hat{X}_{\underline{t}}|^2 \leq |\hat{X}_0|^2 + 2\alpha_2 \sum_{n=0}^{n_t-1} e^{2\alpha_1 t_{n+1}}|\bar{X}_{t_n-\tau}|^2 h_n + 2\beta \sum_{n=0}^{n_t-1} e^{2\alpha_1 t_{n+1}} h_n \\ + 2\sum_{n=0}^{n_t-1} e^{2\alpha_1 t_{n+1}}\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle + \beta \sum_{n=0}^{n_t-1} e^{2\alpha_1 t_{n+1}}|\Delta W_n|^2. \tag{3.4.7}$$

Similarly for the partial time step from $\underline{t}$ to t, we get

$$e^{2\alpha_1 t}|X_t|^2 \leq e^{2\alpha_1 \underline{t}}|\hat{X}_{\underline{t}}|^2 + 2(t-\underline{t})\alpha_2 e^{2\alpha_1 t}|\bar{X}_{\underline{t}-\tau}|^2 + 2(t-\underline{t})\beta e^{2\alpha_1 t}$$
$$+ 2e^{2\alpha_1 t}\langle \hat{X}_{\underline{t}} + f(\hat{X}_{\underline{t}}, \bar{X}_{\underline{t}-\tau})h_n, g(\hat{X}_{\underline{t}}, \bar{X}_{\underline{t}-\tau})(W_t - W_{\underline{t}})\rangle + e^{2\alpha_1 t}\beta|(W_t - W_{\underline{t}})|^2. \tag{3.4.8}$$

Substituting the penultimate inequality into the last one, we obtain

$$\begin{aligned} e^{2\alpha_1 t}|X_t|^2 \leq |X_0|^2 &+ 2\alpha_2 \sum_{n=0}^{n_t-1} e^{2\alpha_1 t_{n+1}}|\bar{X}_{t_n-\tau}|^2|h_n + 2\alpha_2 e^{2\alpha_1 t}|\bar{X}_{t_n-\tau}|^2(t-\underline{t}) \\ &+ 2\beta \sum_{n=0}^{n_t-1} e^{2\alpha_1 t_{n+1}}h_n + 2\beta e^{2\alpha_1 t}(t-\underline{t}) \\ &+ 2\sum_{n=0}^{n_t-1} e^{2\alpha_1 t_{n+1}}\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle \\ &+\beta \sum_{n=0}^{n_t-1} e^{2\alpha_1 t_{n+1}}|\Delta W_n|^2 + e^{2\alpha_1 t}\beta|(W_t - W_{\underline{t}})|^2 \\ &+ 2e^{2\alpha_1 t}\langle \hat{X}_{\underline{t}} + f(\hat{X}_{\underline{t}}, \bar{X}_{\underline{t}-\tau})(t-\underline{t}), g(\hat{X}_{\underline{t}}, \bar{X}_{\underline{t}-\tau})(W_t - W_{\underline{t}})\rangle. \end{aligned}$$

Since $t_{n+1} \leq t_n + h_{max}$ and $t \leq \underline{t} + h_{max}$, we can take the common factor $e^{2\alpha_1 h_{max}}$ out in the equation above. The processes $\bar{X}$ and $\tilde{X}$, defined in (3.2.2) and (3.2.3) respectively, are a simple processes, so we express the second and the third terms in the RHS of the previous equation as a Riemann integral. The same for the fourth and fifth terms. Similarly, the sixth and ninth terms can be written together as a (pathwise) Itô integral,

$$\begin{aligned} e^{2\alpha_1 t}|X_t|^2 \leq |X_0|^2 + e^{2\alpha_1 h_{max}}\Bigg\{ &\int_0^t e^{2\alpha_1 s}|\tilde{X}_{s-\tau}|^2 ds + 2\beta \int_0^t e^{2\alpha_1 s} ds \\ &+ 2\int_0^t e^{2\alpha_1 s}\langle \bar{X}_s + f(\bar{X}_s, \tilde{X}_{s-\tau})[h(\bar{X}_s)I_{[0,\underline{t})}(s) + (t-\underline{t})I_{[\underline{t},t]}(s)], g(\bar{X}_s, \tilde{X}_{s-\tau})dW_s\rangle \\ &+ \beta \sum_{n=0}^{n_t-1} e^{2\alpha_1 t_n}|\Delta W_n|^2 + e^{2\alpha_1 \underline{t}}\beta|(W_t - W_{\underline{t}})|^2 \Bigg\}. \end{aligned}$$

Now, raising to the power $p/2$, using Hölder's inequality and taking the expectation of the supremum, we obtain

$$e^{p\alpha_1 t}\mathbb{E}\left[\sup_{0\leq s\leq t} |X_t|^p\right] \leq 6^{p/2-1}e^{p\alpha_1 h_{max}}(H_1 + H_2 + H_3 + H_4), \tag{3.4.9}$$

where

$$H_1 := \mathbb{E}|X_0|^p + \mathbb{E}\left[\left(2\alpha_2 \int_0^t e^{2\alpha_1 s}|\tilde{X}_{s-\tau}|^2 ds\right)^{p/2}\right] + \left(2\beta \int_0^t e^{2\alpha_1 s} ds\right)^{p/2};$$

$$H_2 := \mathbb{E}\Bigg[\sup_{0\le s\le t}\Bigg|2\int_0^s e^{2\alpha_1 u}\langle \bar{X}_u + f(\bar{X}_u, \tilde{X}_{u-\tau})[h(\bar{X}_u)I_{[0,\underline{s})}(u)$$

$$+ (s-\underline{s})I_{[\underline{s},s]}(u)], g(\bar{X}_s, \tilde{X}_{u-\tau})dW_u\rangle\Bigg|^{p/2}\Bigg];$$

$$H_3 := \mathbb{E}\left[\left(\beta \sum_{n=0}^{n_t-1} e^{2\alpha_1 t_n}|\Delta W_n|^2\right)^{p/2}\right];$$

$$H_4 := \beta^{p/2} e^{p\alpha_1 t}\mathbb{E}[\sup_{0\le s\le t}|(W_s - W_{\underline{s}})|^p].$$

Now we will establish bounds for each of the four terms above. In the remainder of the proof, C is a positive constant that may depend on $\beta, \alpha_1, \alpha_2, h_{max}$ and p, but independent of t, that may change from line to line. We start by bounding H_1.

$$\begin{aligned}
H_1 &\le \mathbb{E}|X_0|^p + \mathbb{E}\left[\left(2\alpha_2 \sup_{-\tau\le s\le t}|X_s|^2 \int_0^t e^{2\alpha_1 s} ds\right)^{p/2}\right] + \left(2\beta\int_0^t e^{2\alpha_1 s} ds\right)^{p/2} \\
&\le \mathbb{E}|X_0|^p + \left(\frac{\alpha_2}{\alpha_1}\right)^{p/2}\mathbb{E}\left[\sup_{-\tau\le s\le t}|X_s|^p\right]e^{\alpha_1 pt} + \left(\frac{2\beta}{2\alpha_1}\right)^{p/2} e^{\alpha_1 pt} \\
&\le e^{\alpha_1 pt}\left(C + \left(\frac{\alpha_2}{\alpha_1}\right)^{p/2}\mathbb{E}\left[\sup_{0\le s\le t}|X_s|^p\right]\right).
\end{aligned}$$

For H_2, the BDG inequality and condition (3.4.3) yields

$$H_2 \le 2^{p/2}\beta^{p/4} C\mathbb{E}\Bigg[\Big(\int_0^t e^{4(\alpha_1-\alpha_2)s}|(\bar{X}_s + f(\bar{X}_s, \tilde{X}_{s-\tau})[h(\bar{X}_s)I_{[0,\underline{t})}(s)$$

$$+ (t-\underline{t})I_{[\underline{t},t]}(s)])|^2 ds\Big)^{p/4}\Bigg].$$

Since $e^{4(\alpha_1-\alpha_2)s} = e^{2(\alpha_1-\alpha_2)\frac{p-4}{p}s} e^{2(\alpha_1-\alpha_2)(1+\frac{4}{p})s}$, by Hölder's inequality, we get

$$\left(\int_0^t e^{4(\alpha_1-\alpha_2)s}|(\bar{X}_s + f(\bar{X}_s, \tilde{X}_{s-\tau})[h(\bar{X}_s)I_{[0,\underline{t})}(s) + (t-\underline{t})I_{[\underline{t},t]}(s)])|^2 ds\right)^{p/4}$$

$$\le \left(\int_0^t e^{2(\alpha_1-\alpha_2)s} ds\right)^{\frac{p-4}{4}}$$

$$\times \int_0^t e^{(\alpha_1-\alpha_2)\frac{p+4}{2}s}|(\bar{X}_s + f(\bar{X}_s, \tilde{X}_{s-\tau})[h(\bar{X}_s)I_{[0,\underline{t})}(s) + (t-\underline{t})I_{[\underline{t},t]}(s)])|^{p/2}ds.$$

Using Assumption (3.4.2), we obtain

$$\begin{aligned}
&|\bar{X}_s + f(\bar{X}_s, \tilde{X}_{s-\tau})[h(\bar{X}_s)I_{[0,\underline{t})}(s) + (t-\underline{t})I_{[\underline{t},t]}(s)]|^2 \\
&\quad \leq |\bar{X}_s|^2 + 2[h(\bar{X}_s)I_{[0,\underline{t})}(s) + (t-\underline{t})I_{[\underline{t},t]}(s)]\left(-\alpha_1|\bar{X}_s|^2 + \alpha_2|\tilde{X}_{s-\tau}|^2 + \beta\right) \\
&\quad \leq |\bar{X}_s|^2 + 2h_{max}\left(\alpha_2|\tilde{X}_{s-\tau}|^2 + \beta\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
H_2 \leq \mathbb{E}\Bigg[&C\left(\int_0^t e^{2\alpha_1 s}ds\right)^{\frac{p-4}{4}} \\
&\times \int_0^t e^{\alpha_1\frac{p+4}{2}s}\left\{|\bar{X}_s|^{p/2} + (2h_{max}\alpha_2)^{p/4}|\tilde{X}_{s-\tau}|^{p/2} + (2\beta h_{max})^{p/4}\right\}ds\Bigg].
\end{aligned}$$

We can write the previous inequality as $H_2 \leq H_{21} + H_{22} + H_{23}$, where

$$\begin{aligned}
H_{21} &:= C\mathbb{E}[\sup_{0\leq s\leq t}|X_s|^{p/2}]\left(\int_0^t e^{2\alpha_1 s}ds\right)^{\frac{p-4}{4}}\int_0^t e^{\alpha_1\frac{p+4}{2}s}ds; \\
H_{22} &:= C(2h_{max}\alpha_2)^{p/4}\mathbb{E}[\sup_{-\tau\leq s\leq t}|X_s|^{p/2}]\left(\int_0^t e^{2\alpha_1 s}ds\right)^{\frac{p-4}{4}}\left(\int_0^t e^{\alpha_1\frac{p+4}{2}s}ds\right); \\
H_{23} &:= C(2h_{max}\alpha_2)^{p/4}\left(\int_0^t e^{2\alpha_1 s}ds\right)^{\frac{p-4}{4}}\left(\int_0^t e^{\alpha_1\frac{p+4}{2}s}ds\right).
\end{aligned}$$

Since,

$$\begin{aligned}
\left(\int_0^t e^{2\alpha_1 s}ds\right)^{\frac{p-4}{4}}\int_0^t e^{\alpha_1\frac{p+4}{2}s} &= \frac{e^{\alpha_1(p-4)t}-1}{(2\alpha_1)^{\frac{p-4}{4}}} \cdot \frac{e^{\alpha_1\frac{p+4}{2}}t-1}{\alpha_1\frac{p+4}{2}} \\
&\leq \frac{e^{\alpha_1 pt}}{\alpha_1\frac{p+4}{2}(2\alpha_1)^{\frac{p-4}{4}}} \leq Ce^{\alpha_1 pt},
\end{aligned}$$

we arrive at

$$\begin{aligned}
H_2 &\leq C\mathbb{E}[\sup_{0\leq s\leq t}|X_s|^{p/2}]e^{\alpha_1 pt} + C\mathbb{E}[\sup_{-\tau\leq s\leq t}|X_s|^{p/2}]e^{\alpha_1 pt} + Ce^{\alpha_1 pt} \\
&= e^{\alpha_1 pt}(C\mathbb{E}[\sup_{0\leq s\leq t}|X_s|^{p/2}] + C).
\end{aligned}$$

Using the elementary inequality $ab \le \frac{1}{2\gamma}a^2 + \frac{\gamma}{2}b^2$ for all $\gamma \in \mathbb{R}^+$ and all $a, b \in \mathbb{R}$ with $a = C$ and $b = \mathbb{E}[\sup_{0\le s\le t} |X_s|^{p/2}$, and later Jensen's inequality, we get

$$C\mathbb{E}[\sup_{0\le s\le t} |X_s|^{p/2}] \le \frac{1}{2\gamma}C^2 + \frac{\gamma}{2}(\mathbb{E}[\sup_{0\le s\le t} |X_s|^{p/2}])^2 \le \frac{1}{2\gamma}C^2 + \frac{\gamma}{2}\mathbb{E}[\sup_{0\le s\le t} |X_s|^{p}].$$

Therefore,

$$H_2 \le e^{\alpha_1 pt}(\frac{\gamma}{2}\mathbb{E}[\sup_{0\le s\le t} |X_s|^p] + C_\gamma), \tag{3.4.10}$$

where the "γ" in C_γ is to emphasise that this constant depends also on γ and is not fixed yet.

Now we will estimate H_3. By the discrete Hölder's inequality we obtain

$$\begin{aligned}\left|\sum_{n=0}^{n_t-1} e^{2\alpha_1 t_n}|\Delta W_n|^2\right| &= \left|\sum_{n=0}^{n_t-1}\left(h_n^{\frac{p-2}{p}} e^{2\alpha_1 t_n \frac{p-2}{p}}\right)\left(h_n^{\frac{2}{p}} e^{\frac{4\alpha_1 t_n}{p}}\frac{|\Delta W_n|^2}{h_n}\right)\right| \\ &\le \left(\sum_{n=0}^{n_t-1} h_n e^{2\alpha_1 t_n}\right)^{\frac{p-2}{p}}\left(\sum_{n=0}^{n_t-1} h_n e^{\frac{2\alpha_1 t_n}{p}}\frac{|\Delta W_n|^p}{h_n^{p/2}}\right)^{\frac{2}{p}}.\end{aligned}$$

By (3.3.15) we can derive that

$$\begin{aligned}H_3 &\le \mathbb{E}\left[\beta^{p/2}\left(\sum_{n=0}^{n_t-1} h_n e^{2\alpha_1 t_n}\right)^{\frac{p-2}{2}}\sum_{n=0}^{n_t-1} h_n e^{2\alpha_1 t_n}\frac{|\Delta W_n|^p}{h_n^{p/2}}\right] \\ &\le \beta^{p/2}\left(\int_0^t e^{2\alpha_1 s}ds\right)^{\frac{p-2}{2}} C\int_0^t e^{2\alpha_1 s}ds \le Ce^{2\alpha_1 t}.\end{aligned}$$

Using (3.3.15) again, we have that

$$H_4 \le \beta^{p/2}e^{\alpha_1 pt}Ch_{max}^{p/2} \le Ce^{\alpha_1 pt}.$$

Collecting together the bounds for H_1, H_2 H_3 and H_4, we obtain

$$e^{p\alpha_1 t}\mathbb{E}[\sup_{0\le s\le t}|X_s|^p] \le e^{p\alpha_1 t}(C_\gamma + \frac{\gamma}{2}\mathbb{E}[\sup_{0\le s\le t}|X_s|^p]) + \left(\frac{\alpha_2}{\alpha_1}\right)^{p/2}\mathbb{E}[\sup_{0\le s\le t}|X_s|^p]).$$

Noting that the constant C is independent of t, $0 \le (\alpha_2/\alpha_1)^{p/2} < 1$ and taking γ small enough such that $\frac{\gamma}{2} < 1 - (\alpha_2/\alpha_1)^{p/2}$, the required assertion follows. □

3.5 Almost sure exponential stability for SDEs

In Wei and Giles [14], the almost sure exponential stability of the adaptive EM solution has not been investigated. In this section we switch momentarily to SDEs to cover this topic. Let $\{W_t\}_{t\geq 0}$ be a $\bar{d}$-dimensional Brownian motion. Consider the d-dimensional SDE

$$dY_t = f(Y_t)dt + g(Y_t)dW_t \tag{3.5.1}$$

for $t \geq 0$ where $f : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^{d\times\bar{d}}$ are Borel-measurable functions, and initial data $Y_0 = \xi \in L^2_{\mathcal{F}_0}(\Omega; \mathbb{R}^d)$, i.e. ξ is a $\mathcal{F}_0$-measurable $\mathbb{R}^d$-valued random variable with $E|\xi|^2 < \infty$.

It was shown in [24] that among other condtions, if the drift function satisfies the linear growth condition, then the Euler-Maruyama approximate solution is a.s. exponentially stable. However, if the drift function satisfies the less restrictive one-sided linear growth condition, the EM solution is not longer stable. It was proved in the same paper that the backward EM solution mantains the stability. But it's well known that the BEM method is much more computationally expensive than explicit methods. Therefore, it is deserable to find explicit methods that are exponetially stable. Our goal in this section is to show that the adaptive EM solution can be a.s. exponentially stable for some SDEs where the standard EM breaks down.

We will impose the following assumption of the SDE (3.5.1)

Assumption 3.5.1. *The functions f and g satisfy the local Lipschitz condition: for every $R > 0$ there exists a positive constant C_R such that*

$$|f(x) - f(y)| + ||g(x) - g(y)|| \leq C_R(|x - y| \tag{3.5.2}$$

for all $x, y \in \mathbb{R}^d$ with $|x|, |y| \leq R$. Furthermore, there exists a constant $\alpha \geq 0$ such that for all $x \in \mathbb{R}^d$, f and g satisfy

$$\langle x, f(x)\rangle + \frac{1}{2}|g(x)|^2 \leq -\alpha|x|^2, \quad \alpha > 0. \tag{3.5.3}$$

Under the conditions (3.5.2) and (3.5.3), the SDE (3.5.1) has a unique solution (Theorem 2.3.6 in [36]).

3.5.1 Example

Consider the following SDE:

$$dY_t = (-2Y_t - Y_t^3)dt + \sqrt{2}Y_t dW_t \quad Y_0 = c \in \mathbb{R}/\{0\}. \tag{3.5.4}$$

Using [25, Theorem 5.1], we can show that the exact solution of the SDE (3.5.4) is almost sure exponentially stable, i.e.

$$\limsup_{t\to\infty} \frac{1}{t} \log |Y_t| \le -\lambda \text{ a.s.}, \qquad \lambda > 0.$$

However, the discrete (standard) EM approximate solution

$$X_{k+1} = X_k(1 - 2\Delta - X_k^2\Delta + \sqrt{2}\Delta W_k), \quad X_0 = Y_0, \quad \Delta \in (0,1) \tag{3.5.5}$$

where $\Delta = 1/m, m \in \mathbb{N}$, is not almost sure exponentially stable. This means that it does not exist a constant $\eta > 0$ and a $\Delta^* \in (0,1)$ such that for all $\Delta \in (0, \Delta^*)$

$$\limsup_{k\to\infty} \frac{1}{k\Delta} \log |X_k| \le -\eta \ \text{ a.s. } .$$

One the contrary, as we will see in Section 3.5.2, the adaptive EM approximate solution to Equation (3.5.4) is almost sure exponentially stable. The following lemma proves a much stronger result that implies the above. It shows that the set in which the EM solution grows at a geometric rate, has positive probability.

Lemma 3.5.2. *Consider the EM aproximate solution* (3.6.5) *to the SDE* (3.5.4). *Then*

$$\mathbb{P}\left(|X_k| \ge \frac{2^{k+3}}{\sqrt{\Delta}}, \ \forall k \ge 1 \right) > 0. \tag{3.5.6}$$

The proof is based on the counterexample's proof given in [24].

Proof. First we show that if $|X_1| \ge 2^4/\sqrt{\Delta}$, then

$$\mathbb{P}\left(|X_k| \ge \frac{2^{k+3}}{\sqrt{\Delta}}, \ \forall k \ge 1 \right) \ge \exp\left(-4e^{-2/\sqrt{\Delta}} \right). \tag{3.5.7}$$

We start by proving the following fact:

$$|X_k| \ge \frac{2^{k+3}}{\sqrt{\Delta}} \quad \text{and} \quad |\Delta W_k| \le 2^k \quad \text{imply} \quad |X_{k+1}| \ge \frac{2^{k+4}}{\sqrt{\Delta}}. \tag{3.5.8}$$

To prove (3.5.8), assume that $|X_k| \geq \frac{2^{k+3}}{\sqrt{\Delta}}$. Then

$$\begin{aligned} |X_{k+1}| &\geq |X_k|(\Delta|X_k|^2 - 1 - 2\Delta - \sqrt{2}|\Delta W_k|) \\ &\geq \frac{2^{k+3}}{\sqrt{\Delta}}(2^{2k+6} - 1 - 2\Delta - \sqrt{2}2^k) \\ &\geq \frac{2^{k+4}}{\sqrt{\Delta}}(2^{2k+5} - 1 - 2\Delta - \sqrt{2}2^{k-1}) \\ &\geq \frac{2^{k+4}}{\sqrt{\Delta}}. \end{aligned}$$

Now, from (3.5.8), given that $|X_1| \geq 2^4/\sqrt{\Delta}$, for any integer $K \geq 0$, the event that $\{|X_k| \geq 2^{k+3}/\sqrt{\Delta}, \forall 1 \leq k \leq K\}$ contains the event that $\{|W_k| \leq 2^k, \forall 1 \leq k \leq K\}$. So since the $\{\Delta W_k\}$ are independent, we have

$$\mathbb{P}\left(|X_k| \geq \frac{2^{k+3}}{\sqrt{\Delta}},\ \forall 1 \leq k \leq K\right) \geq \prod_{k=1}^{K} \mathbb{P}(|\Delta W_k| \leq 2^k).$$

In order to prove (3.5.7), the rest of the proof is identical to the one in Lemma 3.1 in [25]. To obtain the final result, Equation (3.6.7), we need to prove that $\mathbb{P}(|X_1| \geq 2^4/\sqrt{\Delta}) > 0$. But this is true since X_1 is a normal random variable and for a normal random variable X with density function f, we have that for all $a \in \mathbb{R}$, $\mathbb{P}(X \geq a) = \int_a^\infty f(x)dx > 0$. □

In constrast to the EM solution, now we will see that the adaptive approximate solution of the SDE (3.5.4) preserves the stability of the exact solution.

3.5.2 Adaptive Euler-Maruyama method for SDEs and main result

We now define the adaptive-EM method for SDEs. In the same way as for SDDEs, section 3.2, the time step is determined by a function $h^\delta : \mathbb{R}^d \to \mathbb{R}^+$ with $\delta \in (0,1)$. Set $\hat{X}_0 := \xi, h_0^\delta := h^\delta(\hat{X}_0)$ and for $n = 0, 1, 2, ...$ define

$$\hat{X}_{t_{n+1}} := \hat{X}_{t_n} + f(\hat{X}_{t_n})h_n^\delta + g(\hat{X}_{t_n})\Delta W_n, \tag{3.5.9}$$

$$h_n^\delta := h^\delta(\hat{X}_{t_n}), \quad t_{n+1} := t_n + h_n^\delta, \tag{3.5.10}$$

where $\Delta W_n := W_{t_{n+1}} - W_{t_n}$.
We now define the the continuous-time approximate solution. For every $t \geq 0$, let

$$\bar{X}_t := \hat{X}_{t_n} \quad \text{for} \quad t \in [t_n, t_{n+1}) \tag{3.5.11}$$

and define

$$X_t := X_0 + \int_0^t f(\bar{X}_s)ds + \int_0^t g(\bar{X}_s)dW_s. \tag{3.5.12}$$

Assumption 3.5.3. *For every δ, the time step function $h^\delta : \mathbb{R}^d \to \mathbb{R}^+$, is continuous and there exists a constant $\alpha > 0$ such that for all $x \in \mathbb{R}^m$,*

$$\langle x, f(x) \rangle + \frac{1}{2}|g(x)|^2 + \frac{d}{2}h^\delta(x)|f(x)|^2 \leq -\alpha|x|^2, \tag{3.5.13}$$

where $\bar{d}$ is the dimension of the Brownian motion in the SDEs (3.5.1)*. Furthermore, h^δ is uniformly bounded by the real number $h^\delta_{\max} \in (0, \infty)$.*

Given SDE (3.5.4), we define (as an example) the following timestep function:

$$h^\delta(x) := \left(\frac{1}{25} I_{\{|x|<1\}} + 0.25 I_{\{|x|\geq 1\}} \frac{\max(1, |x|)}{\max(1, |f(x)|)} \right) \delta, \tag{3.5.14}$$

which satisfies condition (3.5.13). There is not an automatic procedure to find h^δ, it must be found manually and customized to the specific SDE we are working with. The function (3.5.14) is just and example. As we will see in the next theorem, any function h^δ that satisfies condition (3.5.13) would serve the same purpose.

which satisfies condition (3.5.13) and therefore (as it is proved in the next theorem) ensures the almost sure exponential stability of the numerical solution. There is not automatic procedure to find h^δ, it must be found manually and customized to the specific SDE we are working with.
Under Assumptions 3.5.1 and 3.5.13 the adaptive-EM approximate solution (3.2.4) converges strongly to the exact solution of the SDEs (3.5.1) (see [14]). Now we formulate the main result of the chapter.

Theorem 3.5.4. *Consider the SDEs* (3.5.1). *If f and g satisfy Assumption 3.5.1 and h^{δ} satisfies Assumption 3.5.3, then the adaptive approximate solution* (3.5.9) *(or* (3.5.12)*) is almost sure exponentially stable, i.e. there exists $\lambda > 0$ such that*

$$\limsup_{n\to\infty} \frac{\log|\hat{X}_{t_n}|}{t_n} \leq -\lambda \ a.s.$$

Before proving Theorem 3.5.4, we show that the SDEs (3.5.4) satisfies Assumption 3.5.1

$$\langle x, f(x)\rangle + \frac{1}{2}|g(x)|^2 = -2x^2 - x^4 + \frac{1}{2}2x^2 = -x^2 - x^4 \leq -x^2.$$

Thus the adaptive approximate solution of the SDE (3.5.4) is almost sure exponentially stable.

Proof. From (3.5.9), by using the linearity property of the inner product, we have that

$$|\hat{X}_{t_{n+1}}|^2 \leq |\hat{X}_{t_n}|^2 + 2h_n(\langle \hat{X}_{t_n}, f(\hat{X}_{t_n})\rangle + \frac{1}{2}h_n|f(\hat{X}_{t_n})|^2)$$
$$+ 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n})h_n, g(\hat{X}_{t_n})\Delta W_n\rangle + |g(\hat{X}_{t_n})\Delta W_n|^2.$$

Adding and subtracting $|g(\hat{X}_{t_n})|^2 h_n \bar{d}$ to the RHS of the previous inequality gives

$$|\hat{X}_{t_{n+1}}|^2 \leq |\hat{X}_{t_n}|^2 + 2h_n(\langle \hat{X}_{t_n}, f(\hat{X}_{t_n})\rangle + \frac{1}{2}h_n|f(\hat{X}_{t_n})|^2 + \frac{\bar{d}}{2}|g(\hat{X}_{t_n})|^2)$$
$$+ 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n})h_n, g(\hat{X}_{t_n})\Delta W_n\rangle + |g(\hat{X}_{t_n})|^2(|\Delta W_n|^2 - h_n\bar{d})$$

Using (3.5.13), we obtain

$$|\hat{X}_{t_{n+1}}|^2 \leq |\hat{X}_{t_n}|^2 - 2\alpha h_n|\hat{X}_{t_n}|^2 + 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n})h_n, g(\hat{X}_{t_n})\Delta W_n\rangle$$
$$+ |g(\hat{X}_{t_n})|^2(|\Delta W_n|^2 - h_n\bar{d})$$

Multiplying by $e^{\alpha t_{n+1}}$ and using the fact that $1 + x \leq e^x$ with $x = -2h_n\alpha$ yields

$$e^{\alpha t_{n+1}}|\hat{X}_{t_{n+1}}|^2 \leq e^{\alpha t_n}|\hat{X}_{t_n}|^2 + e^{\alpha t_{n+1}}|g(\hat{X}_{t_n})|^2(|\Delta W_n|^2 - h_n\bar{d})$$
$$+ 2e^{\alpha t_{n+1}}\langle \hat{X}_{t_n} + f(\hat{X}_{t_n})h_n, g(\hat{X}_{t_n})\Delta W_n\rangle.$$

Solving the recurrence and using the bound $h_{\max}$ we have

$$e^{\alpha t_n}|\hat{X}_{t_n}|^2 \leq |\hat{X}_0|^2 + e^{\alpha h_{\max}}\Bigg\{\sum_{k=0}^{n-1} e^{\alpha t_k}|g(\hat{X}_{t_k})|^2(|\Delta W_k|^2 - h_k\bar{d})$$

$$
\begin{aligned}
&+ 2\sum_{k=0}^{n-1} e^{\alpha t_k}\langle \hat{X}_{t_k} + f(\hat{X}_{t_k})h_n, g(\hat{X}_{t_k})\Delta W_k\rangle \Bigg\} \\
&= |\hat{X}_0|^2 + e^{\alpha h_{\max}}\{M_n + N_n\}, \qquad (3.5.15)
\end{aligned}
$$

where:

- $M_n := \sum_{k=0}^{n-1} e^{\alpha t_k}|g(\hat{X}_{t_k})|^2(|\Delta W_k|^2 - h_k\bar{d})$;
- $N_n := 2\sum_{k=0}^{n-1} e^{\alpha t_k}\langle \hat{X}_{t_k} + f(\hat{X}_{t_k})h_k, g(\hat{X}_{t_k})\Delta W_k\rangle$.

Taking logarithms and dividing by t_n, it follows that

$$
\frac{1}{t_n}\log(e^{\lambda t_n}|\hat{X}_{t_n}|^2) \leq \frac{1}{t_n}\log\left(C + \bar{C}\{M_n + N_n\}\right),
$$

where C and $\bar{C}$ are positive constants dependent on $\omega \in \Omega$ and on the constants α and h_{max}, but not on t_n. Since

$$
\begin{aligned}
\mathbb{E}[M_{n+1}|\mathcal{F}_{t_n}] &= \mathbb{E}[e^{\alpha t_n}|g(\hat{X}_{t_n})|^2(|\Delta W_n|^2 - h_n\bar{d}) + M_n|\mathcal{F}_{t_n}] \\
&= e^{\alpha t_n}|g(\hat{X}_{t_n})|^2(\mathbb{E}[|\Delta W_n|^2] - h_n\bar{d}) + M_n = M_n
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}[N_{n+1}|\mathcal{F}_{t_n}] &= \mathbb{E}[2e^{\alpha t_n}\langle \hat{X}_{t_n} + f(\hat{X}_{t_n})h_n, g(\hat{X}_{t_n})\Delta W_n\rangle + N_n|\mathcal{F}_{t_n}] \\
&= 2e^{\alpha t_n}\langle \hat{X}_{t_n} + f(\hat{X}_{t_n})h_n, g(\hat{X}_{t_n})\mathbb{E}[\Delta W_n]\rangle + N_n = N_n,
\end{aligned}
$$

$M+N$ is a local martingale with respect to $\{\mathcal{F}_{t_n}\}$. Thus by the discrete semimartingale convergence theorem (Theorem 2.1.14), we obtain

$$
\lim_{n\to\infty}(M_n + N_n) < \infty \quad \text{a.s.}
$$

Therefore,

$$
\limsup_{n\to\infty}\frac{1}{t_n}\log(e^{\alpha t_n}|\hat{X}_{t_n}|^2) \leq 0 \text{ a.s.}
$$

This is

$$
\limsup_{n\to\infty}\frac{\log|\hat{X}_{t_n}|}{t_n} \leq -\frac{\alpha}{2} \text{ a.s.}
$$

and the proof is complete. □

3.6 Almost sure exponential stability for SDDEs

Here we extend to work done in section 3.5 to SDDEs. The delay component adds some difficulty and additional conditions (although not very restrictive) on the coefficient functions will be needed. It was shown in [49] that among other conditions, when the drift function satisfy the linear growth condition, the Euler-Maruyama approximate solution is a.s. exponentially stable. However, when the drift function satisfies the less restrictive one-sided linear growth condition, the EM solution needs not longer to be stable. It was proved in the same paper that the BEM solution maintains the stability. But as we said in the previous section, the BEM method is implicit and therefore, much more computationally expensive than explicit methods such as the adaptive EM method. Our goal in this section is to show that the adaptive solution can be a.s. exponentially stable for some SDDEs where the EM breaks down.

Assumption 3.6.1. *The functions f and g satisfy the local Lipschitz condition: for every $R > 0$ there exists a positive constant C_R such that*

$$|f(x,y) - f(\bar{x},\bar{y})| + ||g(x,y) - g(\bar{x},\bar{y})|| \leq C_R(|x-\bar{x}| + |y-\bar{y}|) \tag{3.6.1}$$

for all $x, y, \bar{x}, \bar{y} \in \mathbb{R}^m$ with $|x|, |y|, |\bar{x}|, |\bar{y}| \leq R$. Furthermore, there exist constants α_1, α_2 and β satisfying

$$\alpha_1 > 2\alpha_2 \geq 0 \text{ and } \beta > 0, \tag{3.6.2}$$

such that for all $x, y \in \mathbb{R}^m$, f satisfies

$$\langle x, f(x,y)\rangle + \frac{1}{2}||g(x,y)||^2 \leq -\alpha_1|x|^2 + \alpha_2|y|^2. \tag{3.6.3}$$

Under this assumption, the SDDE (3.2.1) has a unique solution.

3.6.1 Counterexample (SDDE)

Consider the following SDDEs

$$dY_t = (-2Y_t - Y_t^3 + \frac{1}{2}Y_t\sin(Y_{t-1}))dt + \sqrt{2}Y_t\cos(Y_{t-1})dW_t \tag{3.6.4}$$

with initial data $\xi \in C([-1,0];\mathbb{R}), \xi(0) = c \in \mathbb{R}/\{0\}$. Using [49, Theorem 1], we can show that the exact solution of the SDDE (3.6.4) is almost sure exponentially stable, i.e.

$$\limsup_{t\to\infty} \frac{1}{t} \log |Y_t| \le -\lambda \text{ a.s.}, \qquad \lambda > 0.$$

However, the discrete (standard) EM approximate solution

$$\begin{aligned} X_k &= \xi(k\Delta) \quad k = -m, -m+1, ..., 0, \\ X_{k+1} &= X_k - X_k[(2 + X_k^2 - \frac{1}{2} X_k \sin(X_{k-1}))\Delta + \sqrt{2}\cos(X_{k-1})\Delta W_k], \\ k &= 0, 1, \dots \end{aligned} \tag{3.6.5}$$

where $\Delta = 1/m, m \in \mathbb{N}$, is not almost sure exponentially stable. This means that it does not exist a constant $\eta > 0$ and a $\Delta^* \in (0,1)$ such that for all $\Delta \in (0, \Delta^*)$

$$\limsup_{k\to\infty} \frac{1}{k\Delta} \log |X_k| \le -\eta \ \text{ a.s. } .$$

On the contrary, as we will see later, the adaptive EM approximate solution to equation (3.6.4) is almost sure exponentially stable.

Let X_k be defined by (3.6.5) The following lemma proves a much stronger result that X_k is not almost sure exponential stable. It shows that the set in which the EM solution grows at a geometric rate has positive probability.

Lemma 3.6.2. *Consider the EM approximate solution* (3.6.5) *to the SDE* (3.6.4). *Then*

$$\mathbb{P}\left(|X_k| \ge \frac{2^{k+3}}{\sqrt{\Delta}}, \ \forall k \ge 1 \right) > 0. \tag{3.6.6}$$

The following proof is based on the counterexample's proof given in [24].

Proof. First we show that if $|X_1| \ge 2^4/\sqrt{\Delta}$, then

$$\mathbb{P}\left(|X_k| \ge \frac{2^{k+3}}{\sqrt{\Delta}}, \ \forall k \ge 1 \right) \ge \exp\left(-4e^{-2/\sqrt{\Delta}} \right). \tag{3.6.7}$$

We start by proving the following fact:

$$|X_k| \ge \frac{2^{k+3}}{\sqrt{\Delta}} \quad \text{and} \quad |\Delta W_k| \le 2^k \quad \text{imply} \quad |X_{k+1}| \ge \frac{2^{k+4}}{\sqrt{\Delta}}. \tag{3.6.8}$$

To prove (3.6.8), assume that $|X_k| \geq \frac{2^{k+3}}{\sqrt{\Delta}}$. Then

$$
\begin{aligned}
|X_{k+1}| &\geq |X_k| \left| |X_k|^2\Delta - |1 + 2\Delta + 1/2\sin(X_{k-1})\Delta + \sqrt{2}\cos(X_{k-1})\Delta W_k| \right| \\
&\geq |X_k| \left| |X_k|^2\Delta - (|1| + |2\Delta| + |1/2\Delta| + |\sqrt{2}\Delta W_k|) \right| \\
&\geq \frac{2^{k+3}}{\sqrt{\Delta}}(2^{2k+6} - 6 - \sqrt{2}2^k) \geq \frac{2^{k+4}}{\sqrt{\Delta}}(2^{2k+5} - 3 - \sqrt{2}2^{k-1}) \\
&\geq \frac{2^{k+4}}{\sqrt{\Delta}}.
\end{aligned}
$$

Now, from (3.6.8), given that $|X_1| \geq 2^4/\sqrt{\Delta}$, for any integer $K \geq 0$, the event that $\{|X_k| \geq 2^{k+3}/\sqrt{\Delta}, \forall 1 \leq k \leq K\}$ contains the event that $\{|W_k| \leq 2^k, \forall 1 \leq k \leq K\}$. Since $\{\Delta W_k\}$ are independent, we have

$$
\mathbb{P}\left(|X_k| \geq \frac{2^{k+3}}{\sqrt{\Delta}},\ \forall 1 \leq k \leq K\right) \geq \prod_{k=1}^{K} \mathbb{P}(|\Delta W_k| \leq 2^k).
$$

In order to prove (3.6.7), the rest of the proof is identical to the one in Lemma 3.1 in [24]. To obtain the final result, Equation (3.6.7), we need to prove that $\mathbb{P}(|X_1| \geq 2^4/\sqrt{\Delta}) > 0$. But this is true since X_1 is a normal random variable and for a normal random variable X with density function f, we have that for all $a \in \mathbb{R}$, $\mathbb{P}(X \geq a) = \int_a^\infty f(x)dx > 0$. □

In contrast to the standard EM solution, now we will see that the adaptive EM solution, maintains the stability of the exact solution of SDDE (3.6.4). We need the following assumption.

Assumption 3.6.3. *For every δ, the time step function $h^\delta : \mathbb{R}^{\bar{d}} \to \mathbb{R}^+$, is continuous and there exist constants $\alpha_1 > \alpha_2 \geq 0$ and $\beta > 0$, such that for all $x, y \in \mathbb{R}^m$,*

$$
\langle x, f(x,y)\rangle + \frac{1}{2}h^\delta(x)|f(x,y)|^2 + \frac{d}{2}||g(x,y)||^2 \leq -\alpha_1|x|^2 + \alpha_2\frac{\min(h^\delta(y), h^\delta(x))}{h^\delta(x)}|y|^2, \tag{3.6.9}
$$

where $\bar{d}$ is the dimension of the Brownian motion in the SDDE (3.2.1). *Furthermore, the function h^δ is uniformly bounded by the real numbers $0 < h^\delta_{\min} < h^\delta_{\max} < 1$, where $h^\delta_{\max}$ is small enough such that*

$$
2\alpha_2 e^{2\alpha_1 h_{\max}} < \alpha_1. \tag{3.6.10}
$$

Note that condition (3.6.9) implies condition (3.6.3) with the same values of α_1 and α_2. An example of function h^δ that satisfies condition (3.6.9) for the SDDE (3.6.4) is

$$h^\delta(x) := \left(\frac{1}{25} I_{\{|x|<1\}} + 0.25 I_{\{|x|\geq 1\}} \frac{|x|^2}{\max(1, |f(x,y)|^2)}\right)\delta. \tag{3.6.11}$$

The following is the main result of this section.

Theorem 3.6.4. *Consider the SDDE* (3.2.1) *with a d-dimensional Brownian motion. If f and g satisfy Assumption 3.6.1 and h^δ satisfies Assumption 3.6.3, then the adaptive approximate solution* (3.2.2) *is almost sure exponentially stable, i.e. there exists a $\lambda > 0$ such that*

$$\limsup_{n\to\infty} \frac{\log|\hat{X}_{t_n}|}{t_n} \leq -\lambda \ a.s.$$

Before proving Theorem 3.6.4, we show that the SDDE (3.6.4) satisfies Assumption 3.6.1

$$\langle x, f(x,y)\rangle + \frac{1}{2}|g(x,y)|^2 = -2x^2 - x^4 + \frac{1}{2}\sin(y)x^2 + x^2\cos^2(y) \leq -\frac{1}{2}x^2.$$

In order to show that h^δ satisfies (3.6.9) for the SDDE (3.6.4), we substitute (3.6.11) into (3.6.9) and differentiate between the cases $|x| < 1$ and $|x| \geq 1$. For $|x| < 1$ we have

$$\begin{aligned}\langle x, & f(x,y)\rangle + \frac{1}{2}h^\delta(x)|f(x,y)|^2 + \frac{d}{2}||g(x,y)||^2 = -2x^2 - x^4 + \frac{1}{2}x^2\sin(y) \\ &+ \frac{1}{2}\frac{1}{25}\delta(4x^2 + 4x^4 - 2x^2\sin(y) + x^6 - x^4\sin(y) + \frac{1}{4}x^2\sin(y)) + \frac{1}{2}2x^2\cos^2(y) \\ &\leq \frac{-3x^2}{10}\end{aligned}$$

and for $|x| \geq 1$ we have

$$\begin{aligned}&\langle x, f(x,y)\rangle + \frac{1}{2}h^\delta(x)|f(x,y)|^2 + \frac{d}{2}||g(x,y)||^2 \\ &= -2x^2 - x^4 + \frac{1}{2}x^2\sin(y) + \frac{1}{2}\frac{1}{4}\delta|x|^2 + \frac{1}{2}2x^2\cos^2(y) \leq \frac{-3x^2}{8}.\end{aligned}$$

Thus the adaptive approximate solution of the SDDE (3.6.4) implemented with h^δ defined as (3.6.11) is almost sure exponentially stable.

We will prove the theorem, but first we need the following lemma.

Lemma 3.6.5. *Consider the SDDE* (3.2.1) *with a $\bar{d}$-dimensional Brownian motion. Suppose f and g satisfy Assumption 3.6.1 and h^δ satisfies Assumption 3.6.3. Let l be a positive integer. Then there exists $\lambda \in (0, \alpha_1)$ such that*

$$\sum_{n=1}^{l} e^{\lambda t_n} |\hat{X}_{t_n}|^2 h_n \leq C + C \sum_{n=1}^{l} e^{\lambda t_n} |g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau})|^2 (|\Delta W_n|^2 - h_n \bar{d})$$
$$+ C \sum_{n=1}^{l} e^{\lambda t_n} \langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) \Delta W_n \rangle \ \textit{a.s.}, \tag{3.6.12}$$

where C is a positive constant dependent on $\omega \in \Omega$, the constants $\alpha_1, \alpha_2, h_{max}$ and λ, but independent of l or t_n.

Proof. From (3.2.2) and (3.6.9), we have

$$|\hat{X}_{t_{n+1}}|^2 = |\hat{X}_{t_n}|^2 + 2h_n(\langle \hat{X}_{t_n}, f(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) \rangle + \frac{1}{2} h_n |f(\hat{X}_{t_n}, \bar{X}_{t_n - \tau})|^2)$$
$$+ 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) \Delta W_n \rangle + |g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) \Delta W_n|^2$$
$$\leq |\hat{X}_{t_n}|^2 + 2h_n(\langle \hat{X}_{t_n}, f(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) \rangle + \frac{1}{2} h_n |f(\hat{X}_{t_n}, \bar{X}_{t_n - \tau})|^2 + \frac{\bar{d}}{2} |g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau})|^2)$$
$$+ 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) \Delta W_n \rangle + |g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau})|^2 (|\Delta W_n|^2 - h_n \bar{d})$$
$$\leq |\hat{X}_{t_n}|^2 - 2\alpha_1 h_n |\hat{X}_{t_n}|^2 + 2\alpha_2 h^\delta(\bar{X}_{t_n - \tau}) |\bar{X}_{t_n - \tau}|^2$$
$$+ 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) \Delta W_n \rangle + |g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau})|^2 (|\Delta W_n|^2 - h_n \bar{d}).$$

Multiplying by $e^{\alpha_1 t_{n+1}}$ and using the fact that $1 + x \leq e^x$ with $x = -h_n \alpha_1$, yields

$$e^{\alpha_1 t_{n+1}} |\hat{X}_{t_{n+1}}|^2 \leq e^{\alpha_1 t_n} |\hat{X}_{t_n}|^2 + 2\alpha_2 h^\delta(\bar{X}_{t_n - \tau}) e^{\alpha_1 t_{n+1}} |\bar{X}_{t_n - \tau}|^2$$
$$+ e^{\alpha_1 t_{n+1}} |g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau})|^2 (|\Delta W_n|^2 - h_n \bar{d})$$
$$+ 2e^{\alpha_1 t_{n+1}} \langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n - \tau}) \Delta W_n \rangle.$$

Solving the recurrence and using the bound $h_{\max}$, one can see that

$$e^{\alpha_1 t_n} |\hat{X}_{t_n}|^2 \leq |X_0|^2 + e^{\alpha_1 h_{\max}} \Bigg\{ \sum_{k=0}^{n-1} e^{\alpha_1 t_k} |g(\hat{X}_{t_k}, \bar{X}_{t_k - \tau})|^2 (|\Delta W_k|^2 - h_k \bar{d})$$
$$+ 2\alpha_2 \sum_{k=0}^{n-1} e^{\alpha_1 t_k} |\bar{X}_{t_k - \tau}|^2 h^\delta(\bar{X}_{t_k - \tau})$$

$$+ 2\sum_{k=0}^{n-1} e^{\alpha_1 t_k}\langle \hat{X}_{t_k} + f(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})h_k, g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})\Delta W_k\rangle \Bigg\}.$$

Thus,

$$|\hat{X}_{t_n}|^2 \le e^{-\alpha_1 t_n}|X_0|^2 + e^{\alpha_1 h_{\max}}\Bigg\{ e^{-\alpha_1 t_n}\sum_{k=0}^{n-1} e^{\alpha_1 t_k}|g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})|^2(|\Delta W_k|^2 - h_k\bar{d})$$
$$+ 2\alpha_2 e^{-\alpha_1 t_n}\sum_{k=0}^{n-1} e^{\alpha_1 t_k}|\bar{X}_{t_k-\tau}|^2 h^{\delta}(\bar{X}_{t_k-\tau})$$
$$+ 2e^{-\alpha_1 t_n}\sum_{k=0}^{n-1} e^{\alpha_1 t_k}\langle \hat{X}_{t_k} + f(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})h_k, g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})\Delta W_k\rangle \Bigg\}.$$

So, for any $\lambda \in (0, \alpha_1)$ we have

$$\sum_{n=1}^{l} e^{\lambda t_n}|\hat{X}_{t_n}|^2 h_n \le e^{-(\alpha_1-\lambda)t_n}|X_0|^2 h_n \tag{3.6.13}$$
$$+ e^{\alpha_1 h_{\max}}\Bigg\{ \sum_{n=0}^{l} e^{-(\alpha_1-\lambda)t_n} h_n \sum_{k=0}^{n-1} e^{\alpha_1 t_k}|g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})|^2(|\Delta W_k|^2 - h_k\bar{d})$$
$$+ 2\alpha_2 \sum_{n=0}^{l} e^{-(\alpha_1-\lambda)t_n} h_n \sum_{k=0}^{n-1} e^{\alpha_1 t_k}|\bar{X}_{t_k-\tau}|^2 h^{\delta}(\bar{X}_{t_k-\tau})$$
$$+ 2\sum_{n=0}^{l} e^{-(\alpha_1-\lambda)t_n} h_n \sum_{k=0}^{n-1} e^{\alpha_1 t_k}\langle \hat{X}_{t_k} + f(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})h_k, g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})\Delta W_k\rangle \Bigg\}. \tag{3.6.14}$$

Moreover, we can see that

$$2\alpha_2 e^{\alpha_1 h_{\max}}\sum_{n=1}^{l} e^{-(\alpha_1-\lambda)t_n} h_n \sum_{k=0}^{n-1} e^{\alpha_1 t_k}|\bar{X}_{t_k-\tau}|^2 h^{\delta}(\bar{X}_{t_k-\tau})$$
$$= 2\alpha_2 e^{\alpha_1 h_{\max}}\sum_{n=1}^{l} e^{\alpha_1 t_n}|\bar{X}_{t_n-\tau}|^2 h^{\delta}(\bar{X}_{t_n-\tau})\sum_{k=n}^{l} e^{-(\alpha_1-\lambda)t_k} h_k.$$

Now since the function $e^{-(\alpha_1-\lambda)s}$ is decreasing on s, we see that

$$\sum_{k=n}^{l} e^{-(\alpha_1-\lambda)t_k} h_k = \sum_{k=n}^{l} e^{(\alpha_1-\lambda)h_k} e^{-(\alpha_1-\lambda)t_{k+1}} h_k$$
$$\le e^{(\alpha_1-\lambda)h_{\max}}\int_{t_n}^{t_l} e^{-(\alpha_1-\lambda)s}ds \le \frac{e^{\alpha_1 h_{\max}}}{\alpha_1-\lambda}e^{-(\alpha_1-\lambda)t_n}.$$

Thus

$$2\alpha_2 e^{\alpha_1 h_{\max}} \sum_{n=1}^{l} e^{-(\alpha_1-\lambda)t_n} h_n \sum_{k=0}^{n-1} e^{\alpha_1 t_k} |\bar{X}_{t_k-\tau}|^2 h^{\delta}(\bar{X}_{t_k-\tau})$$
$$\leq \frac{2\alpha_2 e^{2\alpha_1 h_{max}}}{\alpha_1-\lambda} \left(\sum_{n=1}^{l} e^{\lambda t_n} |\bar{X}_{t_n-\tau}|^2 h^{\delta}(\bar{X}_{t_n-\tau}) \right). \quad (3.6.15)$$

Let $M = M(\omega)$ be such that $t_M \leq \tau < t_{M+1}$. Then we can write

$$\sum_{n=1}^{l} e^{\lambda t_n} |\bar{X}_{t_n-\tau}|^2 h^{\delta}(\bar{X}_{t_n-\tau})$$
$$= \sum_{n=1}^{M} e^{\lambda t_n} |\bar{X}_{t_n-\tau}|^2 h^{\delta}(\bar{X}_{t_n-\tau}) + \sum_{n=M+1}^{l} e^{\lambda t_n} |\bar{X}_{t_n-\tau}|^2 h^{\delta}(\bar{X}_{t_n-\tau})$$
$$\leq C + e^{\lambda h_{\max} M} \sum_{n=1}^{l} e^{\lambda t_n} |\hat{X}_{t_n}|^2 h_n, \quad (3.6.16)$$

Substituting Equation (3.6.16) into (3.6.15), we obtain

$$2\alpha_2 e^{\alpha_1 h_{\max}} \sum_{n=1}^{l} e^{-(\alpha_1-\lambda)t_n} h_n \sum_{k=0}^{n-1} e^{\alpha_1 t_k} |\bar{X}_{t_k-\tau}|^2 h_k$$
$$\leq C + \frac{2\alpha_2 e^{2\alpha_1 h_{max}} e^{\lambda h_{max} M}}{\alpha_1-\lambda} \sum_{n=1}^{l} e^{\lambda t_n} |\hat{X}_{t_n}|^2 h_n, \quad (3.6.17)$$

Similarly we obtain

$$e^{\alpha_1 h_{\max}} \sum_{n=0}^{l} e^{-(\alpha_1-\lambda)t_n} h_n \sum_{k=0}^{n-1} e^{\alpha_1 t_k} |g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})|^2 (|\Delta W_k|^2 - h_k \bar{d})$$
$$\leq \frac{2e^{\alpha_1 h_{\max}}}{\alpha_1-\lambda} \sum_{n=1}^{l} e^{\lambda t_n} |g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2 (|\Delta W_n|^2 - h_n \bar{d}). \quad (3.6.18)$$

and

$$e^{\alpha_1 h_{\max}} 2 \sum_{n=0}^{l} e^{-(\alpha_1-\lambda)t_n} h_n \sum_{k=0}^{n-1} e^{\alpha_1 t_k} \langle \hat{X}_{t_k} + f(\hat{X}_{t_k}, \bar{X}_{t_k-\tau}) h_n, g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau}) \Delta W_k \rangle$$
$$\leq \frac{2e^{2\alpha_1 h_{\max}}}{\alpha_1-\lambda} \sum_{n=1}^{l} e^{\lambda t_n} \langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau}) h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau}) \Delta W_n \rangle.$$
$$(3.6.19)$$

We observe that by condition (3.6.10), h_{max} is such that $0 < 2\alpha_2 e^{2\alpha_1 h_{max}} < \alpha_1$. Then by choosing λ small enough so $0 < \frac{2\alpha_2 e^{2\alpha_1 h_{max}} e^{\lambda h_{max} M}}{\alpha_1 - \lambda} < 1$ and by substituting Equations (3.6.17), (3.6.18) and (3.6.19) into (3.6.13), we obtain the final result.

□

We are now in the position to give

Proof of Theorem 3.6.4. From (3.2.2) and (3.6.9), we have

$$
\begin{aligned}
&|\hat{X}_{t_{n+1}}|^2 = |\hat{X}_{t_n}|^2 + 2h_n(\langle \hat{X}_{t_n}, f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\rangle + \frac{1}{2}h_n|f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2) \\
&+ 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle + |g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n|^2 \\
&\leq |\hat{X}_{t_n}|^2 + 2h_n(\langle \hat{X}_{t_n}, f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\rangle + \frac{1}{2}h_n|f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2 + \frac{\bar{d}}{2}|g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2) \\
&+ 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle + |g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2(|\Delta W_n|^2 - h_n\bar{d}) \\
&\leq |\hat{X}_{t_n}|^2 - 2\alpha_1 h_n|\hat{X}_{t_n}|^2 + 2\alpha_2 h^\delta(\bar{X}_{t_n-\tau})|\bar{X}_{t_n-\tau}|^2 \\
&+ 2\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle + |g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2(|\Delta W_n|^2 - h_n\bar{d}).
\end{aligned}
$$

Now we multiply by $e^{\lambda t_{n+1}}$, where $\lambda \in (0, \alpha_1)$ is the one from Lemma 3.6.5, which makes equation (3.6.12) to hold true. Then using the fact that $1 + x \leq e^x$ with $x = -2h_n\alpha_1$, yields

$$
\begin{aligned}
e^{\lambda t_{n+1}}|\hat{X}_{t_{n+1}}|^2 \leq\ & e^{\lambda t_n}|\hat{X}_{t_n}|^2 + 2\alpha_2 e^{\lambda t_{n+1}}|\bar{X}_{t_n-\tau}|^2 h_n \\
&+ e^{\lambda t_{n+1}}|g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2(|\Delta W_n|^2 - h_n\bar{d}) \\
&+ 2e^{\lambda t_{n+1}}\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\Delta W_n\rangle.
\end{aligned}
$$

Note that in the equation above we have used the fact that $e^{-h_n\alpha_1} \leq e^{-h_n\lambda}$. Solving the recurrence and using the bound $h_{\max}$ we have

$$
\begin{aligned}
e^{\lambda t_n}|\hat{X}_{t_n}|^2 \leq |X_0|^2 + e^{\lambda h_{\max}} &\left\{ \sum_{k=0}^{n-1} e^{\lambda t_k}|g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})|^2(|\Delta W_k|^2 - h_k\bar{d}) \right. \\
&+ 2\alpha_2 \sum_{k=0}^{n-1} e^{\lambda t_k}|\bar{X}_{t_k-\tau}|^2 h^\delta(\bar{X}_{t_k-\tau}) \\
&\left. + 2\sum_{k=0}^{n-1} e^{\lambda t_k}\langle \hat{X}_{t_k} + f(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})h_k, g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})\Delta W_k\rangle \right\}.
\end{aligned}
$$

Using (3.6.16), we obtain

$$
\begin{aligned}
e^{\lambda t_n}|\hat{X}_{t_n}|^2 \leq |X_0|^2 + e^{\lambda h_{\max}} \Bigg\{ & \sum_{k=0}^{n-1} e^{\lambda t_k} |g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})|^2 (|\Delta W_k|^2 - h_k \bar{d}) + C \\
& + e^{\lambda h_{max} M} \sum_{k=1}^{n-1} e^{\lambda t_k} |\hat{X}_{t_k}|^2 h_k \\
& + 2 \sum_{k=0}^{n-1} e^{\lambda t_k} \langle \hat{X}_{t_k} + f(\hat{X}_{t_k}, \bar{X}_{t_k-\tau}) h_k, g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau}) \Delta W_k \rangle \Bigg\}.
\end{aligned} \tag{3.6.20}
$$

Substituting Equation (3.6.12) (from Lemma 3.6.5) into (3.6.20) yields

$$
\begin{aligned}
e^{\lambda t_n}|\hat{X}_{t_n}|^2 \leq |X_0|^2 + C + C & \sum_{k=0}^{n-1} e^{\lambda t_k} |g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})|^2 (|\Delta W_k|^2 - h_k \bar{d}) \\
& + C \sum_{k=0}^{n-1} e^{\lambda t_k} \langle \hat{X}_{t_k} + f(\hat{X}_{t_k}, \bar{X}_{t_k-\tau}) h_k, g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau}) \Delta W_k \rangle \Bigg\} \\
& \leq C + C\{M_n + N_n\},
\end{aligned}
$$

where:

- $M_n := \sum_{k=0}^{n-1} e^{\lambda t_k} |g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau})|^2 (|\Delta W_k|^2 - h_k \bar{d})$;
- $N_n := \sum_{k=0}^{n-1} e^{\lambda t_k} \langle \hat{X}_{t_k} + f(\hat{X}_{t_k}, \bar{X}_{t_k-\tau}) h_k, g(\hat{X}_{t_k}, \bar{X}_{t_k-\tau}) \Delta W_k \rangle$;
- C is a positive constant (that changed from the second to the last line) dependent on $\omega \in \Omega$ and on the constants $\alpha_1, \alpha_2, h_{max}$ and λ, but not on t_n.

Taking logarithms and dividing by t_n, it follows that

$$
\frac{1}{t_n} \log(e^{\lambda t_n} |X_{t_n}|^2) \leq \frac{1}{t_n} \log \left(C + C\{M_n + N_n\} \right).
$$

We observe that

$$
\begin{aligned}
\mathbb{E}[M_{n+1}|\mathcal{F}_{t_n}] &= \mathbb{E}[e^{\lambda t_n} |g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2 (|\Delta W_n|^2 - h_n \bar{d}) + M_n | \mathcal{F}_{t_n}] \\
&= e^{\lambda t_n} |g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})|^2 (\mathbb{E}[|\Delta W_n|^2] - h_n \bar{d}) + M_n = M_n
\end{aligned}
$$

and

$$
\mathbb{E}[N_{n+1}|\mathcal{F}_{t_n}] = \mathbb{E}[2e^{\lambda t_n} \langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau}) h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau}) \Delta W_n \rangle + N_n | \mathcal{F}_{t_n}]
$$

$$= 2e^{\lambda t_n}\langle \hat{X}_{t_n} + f(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})h_n, g(\hat{X}_{t_n}, \bar{X}_{t_n-\tau})\mathbb{E}[\Delta W_n]\rangle + N_n = N_n.$$

Hence $M + N$ is a local martingale with respect to $\{\mathcal{F}_{t_n}\}$. Thus by the discrete semi-martingale convergence theorem (Theorem 2.1.14), one can see that

$$\lim_{n\to\infty}(M_n + N_n) < \infty \quad \text{a.s.}$$

Therefore,

$$\limsup_{n\to\infty} \frac{1}{t_n}\log(e^{\lambda t_n}|\hat{X}_{t_n}|^2) \leq 0 \text{ a.s.}$$

This is

$$\limsup_{n\to\infty} \frac{\log|\hat{X}_{t_n}|}{t_n} \leq -\frac{\lambda}{2} \text{ a.s.}$$

The proof is therefore complete. □

3.7 Simulations

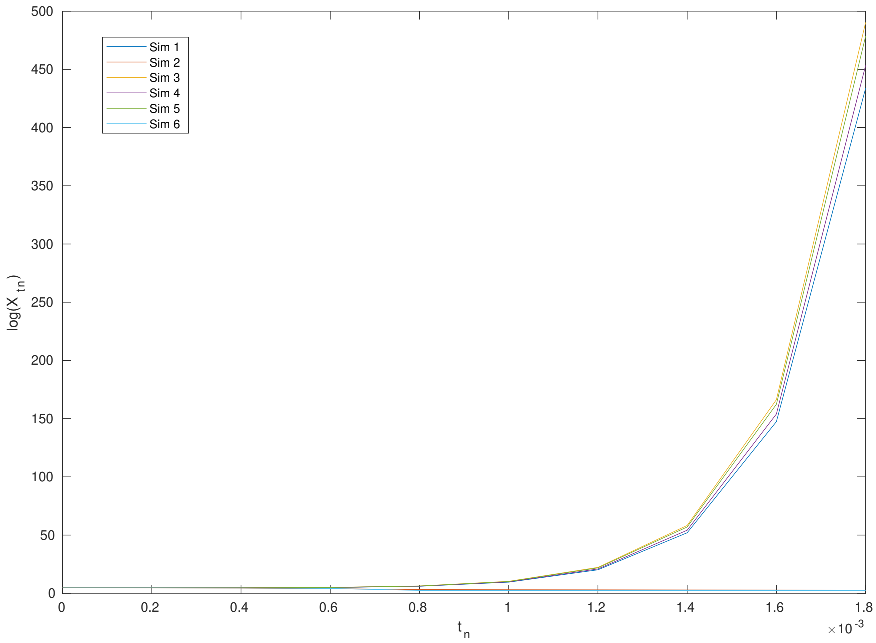
In this section we present simulations which illustrate the results discussed in Section 3.6. Consider the SDDE (3.6.4) with $\tau = 1$ and initial condition $Y(t) = 100, -1 \leq t \leq 0$. We simulated in Matlab paths of the EM solution of the SDDE (3.6.4) using different step sizes, Δ. As we have seen in section 3.6 there is a positive probability that the EM solution explodes. In Table 3.1 we present six different simulations of the EM solution for $\Delta = 2\text{e}{-3}$. We observe in simulations 1,3,4 and 5 the EM solution explodes.

Table 3.1: Six simulations of the EM solution for $\Delta = 2\text{e}{-3}$

Time	0	2e−3	4e−3	6e−3	8e−3	10e−3	12e−3	14e−3	16e−3	18e−3	20e−3
Sim 1	100	101.1	107.4	-141.1	418.1	−1.4e4	5.7e8	−3.7e22	1.1e64	−2.3e188	Inf
Sim 2	100	-98	88.97	-50.99	-24.51	-21.33	-19.37	-17.29	-16.15	-15.13	-14.87
Sim 3	100	-101.3	109.6	-150.1	525.68	−2.8e4	4.6e9	−2e25	1.6e72	−8.3e212	Inf
Sim 4	100	-101.9	108.5	-143.9	452.6	−1.8e4	1.2e9	−3.3e23	7.3e66	−7.9e196	Inf
Sim 5	100	-101.9	108.5	-143.9	452.6	−1.8e4	1.2e9	−3.3e23	7.3e66	−7.9e196	Inf
Sim 6	100	-99	91.8	-63.44	-11.65	-11.03	-10.87	-10.27	-10.17	-9.91	-10

In Figure 3.1, we graphed the logarithm of EM solution presented in Table 3.1.

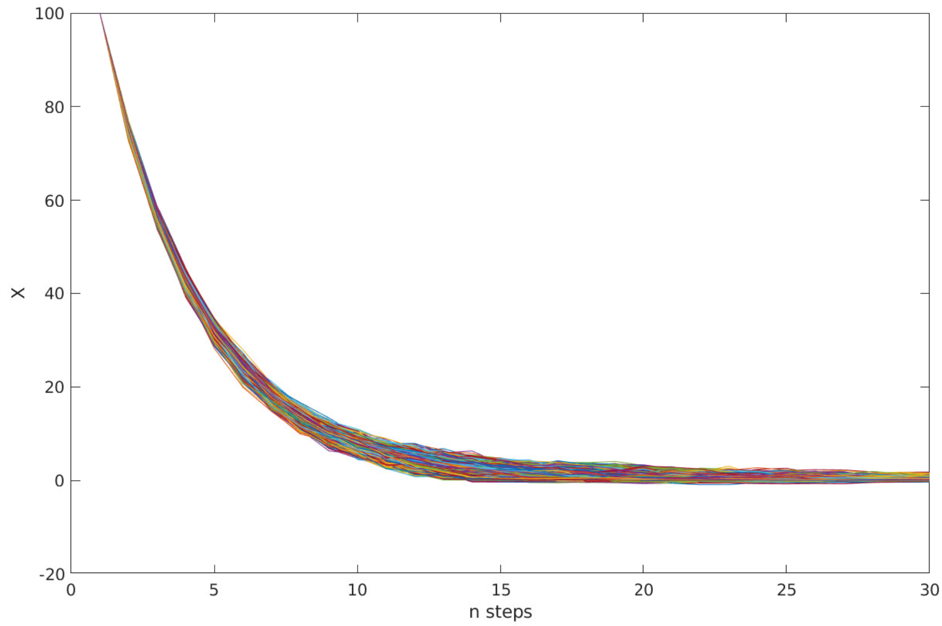
Figure 3.1: Simulations of the logarithm of the EM solution for $\Delta = 2\text{e}{-3}$



Note: From Lemma 3.6.2 we know that as Δ decreases, the probability of explosion decreases. Thus, for "very small" Δ (say less than 10^{-4}) we couldn't find one explosion in 100,000 simulations.

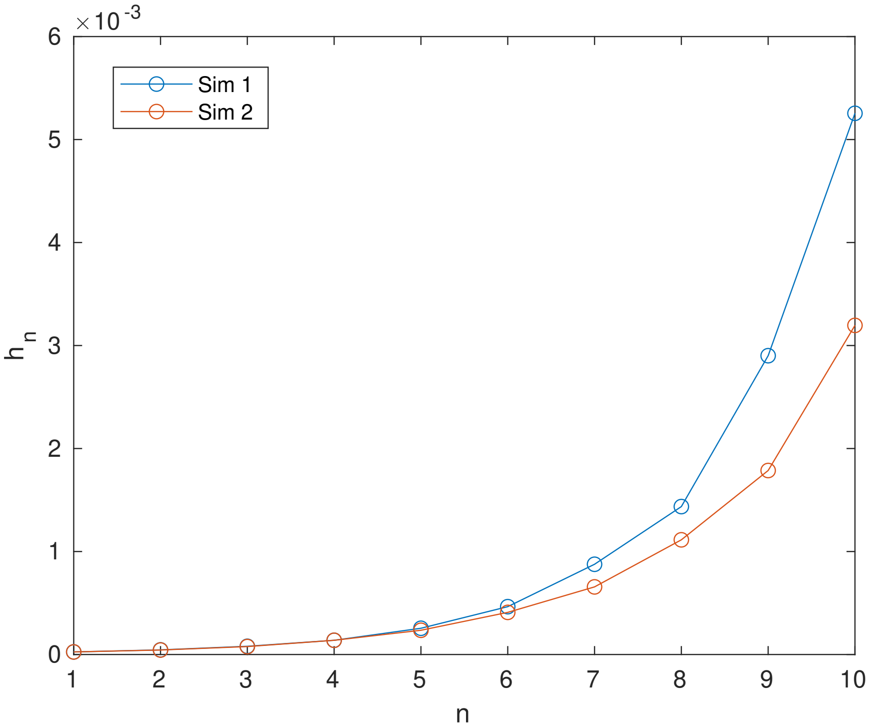
In addition, we simulated the adaptive-EM solution of the SDDE (3.6.4) using the function h^{δ} defined in (3.6.11). As we proved in Section 3.6, the solution is a.s. exponentially stable. Figure 3.2 shows 10,000 paths of the adaptive-EM solution.

Figure 3.2: Simulations of adaptive-EM solution



The next graph shows the first 10 values of $h^\delta(\hat{X}_{t_n})$ for two different simulations. At the start, $\hat{X}_0 = 100$, so the term $-\hat{X}^3_{t_n}$ dominates the equation, making the diffusion term very "big" (in absolute value) in comparison with $\hat{X}_{t_n}$. Therefore, the adaptive step is very "small" at the beginning and increases progressively as the ratio $f(\hat{X}_{t_n}, \hat{X}_{\hat{t}_n})/\hat{X}_{t_n}$ decreases. This ensures all the simulated paths to decay exponentially in a "small" number of steps.

Figure 3.3: The first ten adaptive steps for two different simulations



Chapter 4

Numerical solutions for McKean-Vlasov SDDEs using the adaptive method

In 2021 [46], Reisinger and Stokinger extended the work from [14] to MV-SDEs. In this chapter we extend [46] to MV-SDDEs.

4.1 The EM-adaptive scheme for McKean-Vlasov SDDEs

Let τ and T be positive constants and denote $C([-\tau, 0]; \mathbb{R}^d)$ the space of all continuous functions from $[-\tau, 0]$ to $\mathbb{R}^d$ with the norm $||\phi|| = \sup_{-\tau\le\theta\le 0} |\phi(\theta)|$. Let W be an $\bar{d}$-dimensional Brownian motion defined on the a complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, \mathbb{P})$. Consider the d dimensional MV-SDDE of the form

$$\mathrm{d}Y_t = f(Y_t, Y_{t-\tau}, \mathcal{L}_t^Y)dt + g(Y_t, Y_{t-\tau}, \mathcal{L}_t^Y)dW(t), t \in [0, T] \tag{4.1.1}$$

where $\mathcal{L}_t^Y$ is the law (or distribution) of Y_t,

$$f : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \text{ and } g : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times\bar{d}}$$

and the initial data satisfies the following condition: for any $p \geq 2$

$$\{Y(\theta) : -\tau \leq \theta \leq 0\} := \xi \in L^p_{\mathcal{F}_0}([-\tau, 0]; \mathbb{R}^d),$$

that is ξ is a $\mathcal{F}_0$-measurable $C([-\tau, 0]; \mathbb{R}^d)$-valued random variable such that $E||\xi||^p < \infty$.

By the propagation of chaos result, Theorem 2.2.9, the MV-SDDEs (4.1.1) can be regarded as the limit the M-particle system of d-dimensional MV-SDDEs

$$dY_t^{i,M} = f(Y_t^{i,M}, Y_{t-\tau}^{i,M}, \mathcal{L}_t^{Y^M})dt + g(Y_t^{i,M}, Y_{t-\tau}^{i,M}, \mathcal{L}_t^{Y^M})dW_t^i, \quad t \in [0, T], \tag{4.1.2}$$

with the initial condition $X_0^{i,M} = \xi$ and $\mathcal{L}_t^{Y^M} := \frac{1}{M}\sum_{j=1}^{M} \delta_{Y^{j,M}(t)}$.

We will imposed the following conditions on the coefficient functions f and g.

Assumption 4.1.1. *The functions f and g satisfy:*

(i) (Lipschitz condition on g) There exists a positive constant L such that

$$||g(x, y, \mu) - g(\bar{x}, \bar{y}, \bar{\mu})||^2 \leq L(|x - \bar{x}|^2 + |y - \bar{y}|^2 + \mathbb{W}_2^2(\mu, \bar{\mu})) \tag{4.1.3}$$

for all $x, \bar{x}, y, \bar{y} \in \mathbb{R}^d$ and $\mu, \bar{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$.

(ii) (one-sided Lipschitz condition on f) There exists a positive constant L such that

$$\langle x - \bar{x}, f(x, y, \mu) - f(\bar{x}, \bar{y}, \mu)|| \leq L(|x - \bar{x}|^2 + |y - \bar{y}|^2) \tag{4.1.4}$$

for all $x, \bar{x}, y, \bar{y} \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

(iii) (Lipschitz measure dependence condition on f) There exists a positive constant L such that

$$|f(x, y, \mu) - f(x, y, \bar{\mu})| \leq L\mathbb{W}_2(\mu, \bar{\mu}) \tag{4.1.5}$$

for all $x, y \in \mathbb{R}^d$ and $\mu, \bar{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$.

(iv) polynomial growth Lipschitz condition on f, i.e. there exist constants $\gamma, \lambda, q > 0$ such that for all $x, y, \bar{x}, \bar{y} \in \mathbb{R}^m$

$$|f(x, y, \mu) - f(\bar{x}, \bar{y}, \mu)| \leq (\gamma(|x|^q + |y|^q + |\bar{x}|^q + |\bar{y}|^q) + \lambda)(|x - \bar{x}| + |y - \bar{y}|). \tag{4.1.6}$$

Remark 4.1.1. We note that:

- Conditions (4.1.4) and (4.1.6) are uniform on the measure μ. These conditions are standard to guarantee existence and uniqueness of the exact solution, [33], [12].
- The one-sided condition (4.1.4) allows for a larger class of models than the standard globally Lipschitz drift assumption, [11]. Some of these models are the adjusted Ginzburg Landau equation [11], Kinetic models e.g. in Gomes et al. (2020) [18] and Self-stabilizing diffusions Bolley et al. (2011) [5], Malrieu (2003) [34].

Remark 4.1.2. Condition 4.1.3 implies the linear growth condition on g, i.e. there are positive constants α and β such that

$$||g(x, y, \mu)||^2 \leq \alpha(|x|^2 + |y|^2) + \beta \tag{4.1.7}$$

for all $x, y \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. In addition, condition 4.1.6 implies polynomial growth on f, i.e., there exist constants $\gamma, \lambda, q > 0$ such that

$$|f(x, y, \mu)| \leq (\gamma(|x|^q + |y|^q) + \lambda)(|x| + |y|) \tag{4.1.8}$$

for all $x, y \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

Now we define the numerical solution based on the adaptive method. In the same way as in chapter 3, the time step is determined by a function $h^\delta : \mathbb{R}^d \to \mathbb{R}^+$ with $\delta \in (0, 1)$. The family of functions $\{h^\delta\}_{0<\delta<1}$ is not specifically defined, it just has to satisfy certain conditions that we will describe later in Assumption 4.1.2. Note that for two different particles, the value of the processes at time t_1 may differ resulting in two different values (one for each particle) of the random variable $t_2 = h_1$. This, unlike the standard EM method, presents a challenge when computing $\frac{1}{M}\sum_{j=1}^{M} \delta_{\hat{X}^{j,M}(t_1)}$. In [46], they proposed two different schemes, which both deal with this issue. Here, using our ideas from chapter 3, we extend these two schemes to the delay case. Although the times t_n are different in each scheme, we will use the same notation, t_n, for both. Also given $t > 0$, we define for both schemes $\underline{t} := \max\{t_n : t_n \leq t\}$ and $n_t := \max\{n : t_n \leq t\}$.

Scheme 1

Set $\hat{X}_0^{1,i,M} := \xi(0), h_0^\delta := h^\delta(\hat{X}_0), t_1 := h_0^\delta$ and $\bar{X}_t^{1,i,M} := \xi(t), t \in [-\tau, 0), \bar{X}_t^{1,i,M} := \xi(0), t \in [0, t_1)$. For every $\omega \in \Omega$ and for $n = 0, 1, ...N(\omega)$, we define

$$\begin{aligned}
\hat{X}_{t_{n+1}}^{1,i,M} &:= \hat{X}_{t_n}^{1,i,M} + f(\hat{X}_{t_n}^{1,i,M}, \bar{X}_{t_n-\tau}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}^{1,M}})h_n^\delta + g(\hat{X}_{t_n}^{1,i,M}, \bar{X}_{t_n-\tau}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}^{1,M}})\Delta W_n \\
h_{n+1}^\delta &:= \min_{i\in\{1,\dots,M\}} \left(h^\delta(\hat{X}_{t_{n+1}}^{1,i,M})\right), t_{n+2} := t_{n+1} + h_{n+1}^\delta, \qquad (4.1.9) \\
\bar{X}_t^{1,i,M} &:= \hat{X}_{t_{n+1}}^{1,i,M}, t \in [t_{n+1}, t_{n+2}),
\end{aligned}$$

where $\mathcal{L}_{t_n}^{\hat{X}^{1,M}} := \frac{1}{M}\sum_{i=1}^M \delta_{\hat{X}_{t_n}^{1,i,M}}$, $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ and $N(\omega) := \inf\{n \in \mathbb{Z}^+ : t_n(\omega) \geq T\}$. Define for $n = 0, ..., N - r$

$$\tilde{X}_t^{1,i,M} := \bar{X}_{t_n-\tau}^{1,i,M}, t \in [t_n - \tau, t_{n+1} - \tau), \qquad (4.1.10)$$

where $r = r(\omega)$ is such that $t_r \leq \tau \leq t_{r+1}$. We now define the the continuous approximate solution

$$\begin{aligned}
X_t^{1,i,M} &:= \xi(t), \quad t \in [-\tau, 0]; \\
X_t^{1,i,M} &:= \hat{X}_{\underline{t}}^{1,i,M} + f(\hat{X}_{\underline{t}}^{1,i,M}, \bar{X}_{\underline{t}-\tau}^{1,i,M}, \mathcal{L}_{\underline{t}}^{\hat{X}^{1,M}})(t - \underline{t}) + g(\hat{X}_{\underline{t}}^{1,i,M}, \bar{X}_{\underline{t}-\tau}^{1,i,M}, \mathcal{L}_{\underline{t}}^{\hat{X}^{1,M}})(W_t - W_{\underline{t}}) \\
&\quad t \in [0, T], \qquad (4.1.11)
\end{aligned}$$

which solve the equation

$$X_t^{1,i,M} = \xi(0) + \int_0^t f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})ds + \int_0^t g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})dW_s,$$

for $t \in [0, T]$, where $\mathcal{L}_t^{\bar{X}^{1,M}} := \mathcal{L}_{\underline{t}}^{\hat{X}^{1,M}}$.

Note that $X_{t_n}^{1,i,M} = \hat{X}_{t_n}^{1,i,M} = \bar{X}_{t_n}^{1,i,M}$.

Scheme 2

For a given $\delta \in (0, 1)$ let k_n be the integer such that $t_n \in [k_n\delta T, (k_n + 1)\delta T)$. Set $\hat{X}_0^{2,i,M} := \xi(0), h_0^{i,\delta} := h^\delta(\hat{X}_0), t_1 := h_0^\delta$ and $\bar{X}_t^{2,i,M} := \xi(t), t \in [-\tau, 0), \bar{X}_t^{2,i,M} :=$

$\xi(0), t \in [0, t_1)$. For every $\omega \in \Omega$ and for $n = 0, 1, ...N(\omega)$, we define

$$\begin{aligned}\hat{X}^{2,i,M}_{t_{n+1}} &:= \hat{X}^{2,i,M}_{t_n} + f(\hat{X}^{2,i,M}_{t_n}, \bar{X}^{2,i,M}_{t_n-\tau}, \mathcal{L}^{\hat{X}^{2,M}}_{k_n\delta T})h^{i,\delta}_n + g(\hat{X}^{2,i,M}_{t_n}, \bar{X}^{2,i,M}_{t_n-\tau}, \mathcal{L}^{\hat{X}^{2,M}}_{k_n\delta T})\Delta W_n \\ h^{i,\delta}_{n+1} &:= \min(h^\delta(\hat{X}^{2,i,M}_{t_{n+1}}), (k_n+1)\delta T - t_n), t_{n+2} := t_{n+1} + h^{i,\delta}_{n+1}, \\ \bar{X}^{2,i,M}_t &:= \hat{X}^{2,i,M}_{t_{n+1}}, t \in [t_{n+1}, t_{n+2}).\end{aligned} \tag{4.1.12}$$

where $\mathcal{L}^{\hat{X}^{2,M}}_{k_n\delta T} := \frac{1}{M}\sum_{i=1}^M \delta_{\hat{X}^{2,i,M}_{k_n\delta T}}$, $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ and $N(\omega) := \inf\{n \in \mathbb{Z}^+ : t_n(\omega) \geq T\}$. Define for $n = 0, ..., N - r$

$$\tilde{X}^{2,i,M}_t := \bar{X}^{2,i,M}_{t_n-\tau}, t \in [t_n - \tau, t_{n+1} - \tau), \tag{4.1.13}$$

where $r = r(\omega)$ is such that $t_r \leq \tau \leq t_{r+1}$. We now define the the continuous approximate solution.

$$\begin{aligned}X^{2,i,M}_t &:= \xi(t), \quad t \in [-\tau, 0]; \\ X^{2,i,M}_t &:= \hat{X}^{2,i,M}_{\underline{t}} + f(\hat{X}^{2,i,M}_{\underline{t}}, \bar{X}^{2,i,M}_{\underline{t}-\tau}, \mathcal{L}^{\hat{X}^{2,M}}_{k_{n_t}\delta T})(t - \underline{t}) + g(\hat{X}^{2,i,M}_{\underline{t}}, \bar{X}^{2,i,M}_{\underline{t}-\tau}, \mathcal{L}^{\hat{X}^{2,M}}_{k_{n_t}\delta T})(W_t - W_{\underline{t}}), \\ &\quad t \in [0, T],\end{aligned} \tag{4.1.14}$$

which solves the equation

$$\begin{aligned}X^{2,i,M}_t &= \xi(0) + \int_0^t f(\hat{X}^{2,i,M}_s, \tilde{X}^{2,i,M}_{s-\tau}, \mathcal{L}^{\bar{X}^{2,M}}_s)ds + \int_0^t g(\hat{X}^{2,i,M}_s, \tilde{X}^{2,i,M}_{s-\tau}, \mathcal{L}^{\bar{X}^{2,M}}_s)dW_s, \\ &\quad t \in [0, T],\end{aligned} \tag{4.1.15}$$

where $\mathcal{L}^{\bar{X}^{2,M}}_t := \mathcal{L}^{\hat{X}^{2,M}}_{k_{n_t}\delta T}$.

Note that $X^{2,i,M}_{t_n} = \hat{X}^{2,i,M}_{t_n} = \bar{X}^{2,i,M}_{t_n}$.

Remark 4.1.3.

- In scheme 1, for each n we compute $h^\delta(\hat{X}^{1,i,M}_{t_n})$ for every particle, then we choose the smallest and set it as the common step-size for every particle. This scheme is theoretically convenient to prove strong convergence but in practice it may be appropriate only if we are simulating a system with a small number of particles.

If the system is large, the step-size (being the smallest for every particle) is going to be quite "small", which makes this scheme computationally expensive. This is the reason why in [46], they proposed also a second, less expensive scheme.

- In scheme 2, the measure is kept constant in the intervals $[k_n\delta T, (k_n+1)\delta T)$. Note that k_n may have the same value for different n. For example we could have for some $n = \bar{n}$ that $k_{\bar{n}} = k_{\bar{n}+1} = k_{\bar{n}+2} = k_{\bar{n}+3}$, so $t_{\bar{n}}, t_{\bar{n}+1}, t_{\bar{n}+2}, t_{\bar{n}+3} \in [k_{\bar{n}}\delta T, (k_{\bar{n}} + 1)\delta T) = [k_{\bar{n}+i}\delta T, (k_{\bar{n}+i} + 1)\delta T), i = 1, 2, 3$. Note also that by definition, the sequence of times $\{k_n\delta T\}_{n\in\mathbb{N}}$ is a subsequence of the sequence of times $\{t_n\}_{n\in\mathbb{N}}$.

- The reason to introduce the second step process $\tilde{X}$, is to ensure that the equalities $\hat{X}^{1,i,M}_{t_n} = X^{1,i,M}_{t_n}$ and $\hat{X}^{2,i,M}_{t_n} = X^{2,i,M}_{t_n}$ hold true. This is because $\bar{X}_{t-\tau}$ may not be constant in the intervals $[t_n, t_{n+1}]$, due to the variability of the adaptive stepsize.

We will impose the following conditions on the time-step function h^δ.

Assumption 4.1.2. *For each $\delta \in (0, 1)$, the time step function $h^\delta : \mathbb{R}^d \to \mathbb{R}^+$ is continuous, strictly positive and bounded by δT, i.e.*

$$0 < h^\delta(x) \leq \delta T \quad \text{for all } x \in \mathbb{R}^d. \tag{4.1.16}$$

Furthermore, there exist constants $\alpha, \beta > 0$ such that

$$\langle x, f(x, y, \mu)\rangle + \frac{1}{2}h^\delta(x)|f(x, y, \mu)|^2 \leq \alpha(|x|^2 + |y|^2) + \beta, \tag{4.1.17}$$

for all $x, y \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

4.2 Convergence of the numerical solutions

In this section we will prove the strong convergence of the two numerical schemes which were defined in the previous section.

4.2.1 The boundedness of the pth moments of the numerical solutions

The discrete-time approximate solution defined in (4.1.9) and (4.1.12) are not necessarily bounded. In order to prove the the boundedness of the p-moments of the numerical solutions $X^{j,i,M}, j = 1, 2$, we will construct K-bounded schemes $X_K^{j,i,M}, j = 1, 2$ such that for any $K > ||\xi||$, we have $X_K^{j,i,M}, j = 1, 2 < K$. Then we will show that the p-moments of $X_K^{j,i,M}, j = 1, 2$ are bounded by a constant independent of K. Then the boundedness of the pth moments yields from letting K go to infinity and using monotone convergence theorem. Now we make the above explanation rigorous. Let $K > ||\xi||$. Set $\hat{X}_{0,K}^{1,i,M} := \xi(0), h_0^\delta := h^\delta(\hat{X}_{0,K}^{1,i,M}), t_1 := h_0^\delta$ and $\bar{X}_{t,K}^{1,i,M} := \xi(t), t \in [-\tau, 0), \bar{X}_{t,K}^{1,i,M} := \xi(0), t \in [0, t_1)$. Consider the function $\Phi_K : \mathbb{R}^d \to \mathbb{R}^d, \Phi(x) = \min(1, K/|x|)x$. Then for every $\omega \in \Omega$ and for $n = 0, 1, ...N(\omega)$, we define

$$\begin{aligned}\hat{X}_{t_{n+1},K}^{1,i,M} &:= \Phi_K(\hat{X}_{t_n,K}^{1,i,M} + f(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})h_n^\delta + g(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})\Delta W_n)\\ h_{n+1}^\delta &:= \min_{i\in\{1,...,M\}} \left(h^\delta(\hat{X}_{t_{n+1},K}^{1,i,M})\right), t_{n+2} := t_{n+1} + h_{n+1}^\delta,\\ \bar{X}_{t,K}^{1,i,M} &:= \hat{X}_{t_{n+1},K}^{1,i,M}, t \in [t_{n+1}, t_{n+2}).\end{aligned} \tag{4.2.1}$$

where $\mathcal{L}_{t_n}^{\hat{X}_K^{1,i,M}} := \frac{1}{M}\sum_{i=1}^M \delta_{\hat{X}_{t_n,K}^{1,i,M}}$, $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ and $N(\omega) := \inf\{n \in \mathbb{Z}^+ : t_n(\omega) \geq T\}$. Define for $n = 0, ..., N - r$

$$\tilde{X}_{t,K}^{1,i,M} := \bar{X}_{t_n-\tau,K}^{1,i,M}, t \in [t_n - \tau, t_{n+1} - \tau), \tag{4.2.2}$$

where $r = r(\omega)$ is such that $t_r \leq \tau \leq t_{r+1}$. We now define the the continuous approximate solution

$$\begin{aligned} X_{t,K}^{1,i,M} &:= \xi(t), \quad t \in [-\tau, 0];\\ X_{t,K}^{1,i,M} &:= \Phi_K\big(\hat{X}_{\underline{t},K}^{1,i,M} + f(\hat{X}_{\underline{t},K}^{1,i,M}, \bar{X}_{\underline{t}-\tau,K}^{1,i,M}, \mathcal{L}_{\underline{t}}^{\hat{X}_K^{1,M}})(t - \underline{t})\\ &\quad + g(\hat{X}_{\underline{t},K}^{1,i,M}, \bar{X}_{\underline{t}-\tau,K}^{1,i,M}, \mathcal{L}_{\underline{t}}^{\hat{X}_K^{1,M}})(W_t - W_{\underline{t}})\big), \quad t \in [0, T], \end{aligned} \tag{4.2.3}$$

Note that $X_{t_n,K}^{1,i,M} = \hat{X}_{t_n,K}^{1,i,M} = \bar{X}_{t_n,K}^{1,i,M}$. In the same way, we construct the K-bounded schemes $\hat{X}_K^{2,i,M}, \bar{X}_K^{2,i,M}, \tilde{X}_K^{2,i,M}$ and $X_K^{2,i,M}$.

Lemma 4.2.1. *Let $p \geq 4$, the MV-SDDE (4.1.2) satisfy Assumption 4.1.1 and the function h^δ satisfy Assumption 4.1.2. Then, for the K-bounded schemes, T is attainable and for all $p > 0$ there exists a constant C dependent on T and p, but independent of h_n^δ and K such that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|X_{t,K}^{1,i,M}|^p\right]\bigvee\mathbb{E}\left[\sup_{0\leq t\leq T}|X_{t,K}^{2,i,M}|^p\right]\leq C. \tag{4.2.4}$$

Proof. We prove first attainability. Since h^δ is continuous and strictly positive, $\inf_{|x|\leq K} h^\delta(x) > 0$. This implies that for every $\omega \in \Omega$

$$\liminf_{n\to\infty} h_n^\delta(\omega) = \liminf_{n\to\infty} h^\delta(\hat{X}_{t_n}^K(\omega)) > 0,$$

so $\lim_{n\to\infty} t_n(\omega) = \sum_{n=0}^\infty h_n^\delta(\omega) = \infty$ for all $\omega \in \Omega$ and T is attainable in the bounded scheme.

Now we will prove the boundedness of the pth moments for scheme 1, where the upper bound will be a positive constant independent of h_n^δ and K. Since it is irrelevant in this proof, we shall drop the symbol "δ" in the adaptive time-step "h_n^δ" to ease the notation. Let $p \geq 4$ and let $t \in [0,T]$. Define $\underline{t} := \max\{t_n : t_n \leq t\}$, and $n_t := \max\{n : t_n \leq t\}$. Using (4.2.1) and since for any $x \in \mathbb{R}^m$, $|\Phi(x)|^2 \leq |x|^2$, we have that for $n = 0$ to $n = n_t - 1$,

$$\begin{aligned}
|\hat{X}_{t_{n+1},K}^{1,i,M}|^2 &\leq |\hat{X}_{t_n,K}^{1,i,M} + f(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})h_n + g(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})\Delta W_n|^2 \\
&= \langle \hat{X}_{t_n,K}^{1,i,M}, \hat{X}_{t_n,K}^{1,i,M}\rangle + 2\langle \hat{X}_{t_n,K}^{1,i,M}, f(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})h_n\rangle \\
&+ \langle f(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})h_n, f(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})h_n\rangle \\
&+ 2\langle \hat{X}_{t_n,K}^{1,i,M} + f(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})h_n, g(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})\Delta W_n\rangle \\
&+ \langle g(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})\Delta W_n, g(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})\Delta W_n\rangle \\
&= |\hat{X}_{t_n,K}^{1,i,M}|^2 + 2h_n\Big[\langle \hat{X}_{t_n,K}^{1,i,M}, f(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})\rangle \\
&+ \frac{1}{2}h_n|f(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})|^2\Big] \\
&+ 2\langle \hat{X}_{t_n,K}^{1,i,M} + f(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})h_n, g(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})\Delta W_n\rangle \\
&+ |g(\hat{X}_{t_n,K}^{1,i,M}, \bar{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})\Delta W_n|^2
\end{aligned}$$

$$
\begin{aligned}
&\leq |\hat{X}^{1,i,M}_{t_n,K}|^2 + 2h_n\alpha(|\hat{X}^{1,i,M}_{t_n,K}|^2 + |\bar{X}^{1,i,M}_{t_n-\tau,K}|^2) + 2h_n\beta \\
&+ 2\langle \hat{X}^{1,i,M}_{t_n,K} + f(\hat{X}^{1,i,M}_{t_n,K}, \bar{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})h_n, g(\hat{X}^{1,i,M}_{t_n,K}, \bar{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})\Delta W_n\rangle \\
&+ |g(\hat{X}^{1,i,M}_{t_n,K}, \bar{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})\Delta W_n|^2,
\end{aligned}
$$

where in the last step we have used condition (4.1.17). Solving the recurrence relation, we get

$$
\begin{aligned}
|\hat{X}^{1,i,M}_{\underline{t},K}|^2 &\leq |\hat{X}^K_0|^2 + 2\alpha\left(\sum_{n=0}^{n_t-1} |\hat{X}^{1,i,M}_{t_n,K}|^2 h_n + |\bar{X}^{1,i,M}_{t_n-\tau,K}|^2 h_n\right) + 2\beta\underline{t} \\
&+ 2\sum_{n=0}^{n_t-1} \langle \hat{X}^{1,i,M}_{t_n,K} + f(\hat{X}^{1,i,M}_{t_n,K}, \bar{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})h_n, g(\hat{X}^{1,i,M}_{t_n,K}, \bar{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})\Delta W_n\rangle \\
&+ \sum_{n=0}^{n_t-1} |g(\hat{X}^{1,i,M}_{t_n,K}, \bar{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})\Delta W_n|^2. \qquad (4.2.5)
\end{aligned}
$$

Similarly, the continuous approximate solution verifies

$$
\begin{aligned}
&|X^{1,i,M}_{t,K}|^2 \leq |\hat{X}^{1,i,M}_{\underline{t},K}|^2 + 2(t-\underline{t})\alpha(|\hat{X}^{1,i,M}_{\underline{t},K}|^2 + |\bar{X}^{1,i,M}_{\underline{t}-\tau,K}|^2) + 2(t-\underline{t})\beta \\
&+ 2\langle \hat{X}^{1,i,M}_{\underline{t},K} + f(\hat{X}^{1,i,M}_{\underline{t},K}, \bar{X}^{1,i,M}_{\underline{t}-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{\underline{t}})(t-\underline{t}), g(\hat{X}^{1,i,M}_{\underline{t},K}, \bar{X}^{1,i,M}_{\underline{t}-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{\underline{t}})(W_t - W_{\underline{t}})\rangle \\
&+ |g(\hat{X}^{1,i,M}_{\underline{t},K}, \bar{X}^{1,i,M}_{\underline{t}-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{\underline{t}})(W_t - W_{\underline{t}})|^2. \qquad (4.2.6)
\end{aligned}
$$

Substituting (4.2.5) into (4.2.6) yields

$$
\begin{aligned}
&|X^{1,i,M}_{t,K}|^2 \leq |\hat{X}^{1,i,M}_{0,K}|^2 \\
&+ 2\alpha\left(\sum_{n=0}^{n_t-1} |\hat{X}^{1,i,M}_{t_n,K}|^2 h_n + |\bar{X}^{1,i,M}_{t_n-\tau,K}|^2 h_n + |\hat{X}^{1,i,M}_{\underline{t},K}|^2(t-\underline{t}) + |\bar{X}^{1,i,M}_{\underline{t}-\tau,K}|^2(t-\underline{t})\right) + 2\beta t \\
&+ 2\sum_{n=0}^{n_t-1} \langle \hat{X}^{1,i,M}_{t_n,K} + f(\hat{X}^{1,i,M}_{t_n,K}, \bar{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})h_n, g(\hat{X}^{1,i,M}_{t_n,K}, \bar{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})\Delta W_n\rangle \\
&+ 2\langle \hat{X}^{1,i,M}_{\underline{t},K} + f(\hat{X}^{1,i,M}_{\underline{t},K}, \bar{X}^{1,i,M}_{\underline{t}-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{\underline{t}})(t-\underline{t}), g(\hat{X}^{1,i,M}_{\underline{t},K}, \bar{X}^{1,i,M}_{\underline{t}-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{\underline{t}})(W_t - W_{\underline{t}})\rangle \\
&+ \sum_{n=0}^{n_t-1} |g(\hat{X}^{1,i,M}_{t_n,K}, \bar{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})\Delta W_n|^2 + |g(\hat{X}^{1,i,M}_{\underline{t},K}, \bar{X}^{1,i,M}_{\underline{t}-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})(W_t - W_{\underline{t}})|^2.
\end{aligned}
$$

Using the step processes $\bar{X}^K$ and $\tilde{X}^K$ defined previously, the second summand on the RHS of the equation above, can be expressed as a Riemann integral. Similarly the sixth and the seventh terms can be written as an Itô integral, i.e.

$$
\begin{aligned}
|X_{t,K}^{1,i,M}|^2 \le |X_{0,K}^{1,i,M}| + 2\alpha \int_0^t (|\bar{X}_{s,K}^{1,i,M}|^2 + |\tilde{X}_{s-\tau,K}^{1,i,M}|^2) ds + 2\beta t \\
+ 2\int_0^t \langle \bar{X}_{s,K}^{1,i,M} + f(\bar{X}_{s,K}^{1,i,M}, \tilde{X}_{s-\tau,K}^{1,i,M}, \mathcal{L}_s^{\bar{X}_K^{1,M}})[h(\bar{X}_{u,K}^{1,i,M}) I_{[0,\underline{t})}(u) \\
+ (t-\underline{t}) I_{[\underline{t},t]}(u)], g(\bar{X}_{s,K}^{1,i,M}, \tilde{X}_{s-\tau,K}^{1,i,M}, \mathcal{L}_s^{\bar{X}_K^{1,M}}) dW_s \rangle \\
+ \sum_{n=0}^{n_t-1} |g(\bar{X}_{t_n,K}^{1,i,M}, \tilde{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}}) \Delta W_n|^2 + |g(\bar{X}_{t,K}^{1,i,M}, \tilde{X}_{t-\tau,K}^{1,i,M}, \mathcal{L}_t^{\bar{X}_K^{1,M}})(W_t - W_{\underline{t}})|^2.
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
|X_{t,K}^{1,i,M}|^p \le 6^{p/2-1} \Bigg\{ |X_{0,K}^{1,i,M}|^p + \left(2\alpha \int_0^t (|\bar{X}_{s,K}^{1,i,M}|^2 + |\tilde{X}_{s-\tau,K}^{1,i,M}|^2) ds\right)^{p/2} + (2\beta t)^{p/2} \\
+ \Big| 2\int_0^t \langle \bar{X}_{s,K}^{1,i,M} + f(\bar{X}_{s,K}^{1,i,M}, \tilde{X}_{s-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}})[h(\bar{X}_{u,K}^{1,i,M}) I_{[0,\underline{t})}(u) \\
+ (t-\underline{t}) I_{[\underline{t},t]}(u)], g(\bar{X}_{s,K}^{1,i,M}, \tilde{X}_{s-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}}) dW_s \rangle \Big|^{p/2} \\
+ \left(\sum_{n=0}^{n_t-1} |g(\bar{X}_{t_n,K}^{1,i,M}, \tilde{X}_{s-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}}) \Delta W_n|^2 \right)^{p/2} + |g(\bar{X}_{t,K}^{1,i,M}, \tilde{X}_{t-\tau,K}^{1,i,M}, \mathcal{L}_t^{\bar{X}_K^{1,M}})(W_t - W_{\underline{t}})|^p \Bigg\}.
\end{aligned}
$$

Taking the expectation of the supremum, one has

$$
\mathbb{E}\left[\sup_{0\le s\le t} |X_{s,K}^{1,i,M}|^p\right] \le 6^{p/2-1}(I_1 + I_2 + I_3 + I_4),
$$

where

$$
\begin{aligned}
I_1 &:= \mathbb{E}|X_{0,K}^{1,i,M}|^p + \mathbb{E}\left[\left(2\alpha \int_0^t (|\bar{X}_{s,K}^{1,i,M}|^2 + |\tilde{X}_{s-\tau,K}^{1,i,M}|^2) ds\right)^{p/2}\right] + (2\beta t)^{p/2}; \\
I_2 &:= \mathbb{E}\Big[\sup_{0\le s\le t} \Big| 2\int_0^s \langle \bar{X}_{u,K}^{1,i,M} + f(\bar{X}_{u,K}^{1,i,M}, \tilde{X}_{u-\tau,K}^{1,i,M}, \mathcal{L}_u^{\bar{X}_K^{1,M}})[h(\bar{X}_{u,K}^{1,i,M}) I_{[0,\underline{s})}(u) \\
&\quad + (s-\underline{s}) I_{[\underline{s},s]}(u)], g(\bar{X}_{u,K}^{1,i,M}, \tilde{X}_{u-\tau,K}^{1,i,M}, \mathcal{L}_u^{\bar{X}_K^{1,M}}) dW_u \rangle \Big|^{p/2}\Big]; \\
I_3 &:= \mathbb{E}\left[\left(\sum_{n=0}^{n_t-1} |g(\bar{X}_{t_n,K}^{1,i,M}, \tilde{X}_{t_n-\tau,K}^{1,i,M}, \mathcal{L}_{t_n}^{\hat{X}_K^{1,M}}) \Delta W_n|^2 \right)^{p/2}\right]; \\
I_4 &:= \mathbb{E}\left[\sup_{0\le s\le t} |g(\bar{X}_{s,K}^{1,i,M}, \tilde{X}_{s-\tau,K}^{1,i,M}, \mathcal{L}_s^{\bar{X}_K^{1,M}})(W_s - W_{\underline{s}})|^p\right].
\end{aligned}
$$

Now we will establish bounds for each of the four terms above. In the remainder of the proof, C is positive constants, independent of K, that may change from line to line.

Using Hölder's inequality, we have

$$I_1 \le \mathbb{E}|X^{1,i,M}_{0,K}|^p + (2\alpha)^{p/2}T^{p/2-1}2^{p/2-1}\int_0^t \mathbb{E}[|\bar{X}^{1,i,M}_{s,K}|^p + |\tilde{X}^{1,i,M}_{s-\tau,K}|^p]ds + (2\beta T)^{p/2}$$
$$\le C\int_0^t \mathbb{E}\left[\sup_{0\le u\le s}|X^{1,i,M}_{u,K}|^p\right]ds + C.$$

By the Burkholder-Davis-Gundy (BDG) inequality we obtain

$$I_2 \le 2^{p/2}C\mathbb{E}\Big[\Big(\int_0^t |(\bar{X}^{1,i,M}_{u,K} + f(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K}, \mathcal{L}_u^{\bar{X}^{1,M}_K})[h(\bar{X}^{1,i,M}_{u,K})I_{[0,\underline{t})}(u)$$
$$+ (t-\underline{t})I_{[\underline{t},t]}(u)])g(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K}, \mathcal{L}_u^{\bar{X}^{1,M}_K})|^2du\Big)^{p/4}\Big]$$

An application of the Hölder inequality yields that

$$I_2 \le 2^{\frac{p}{2}}T^{\frac{p}{4}-1}C\mathbb{E}\Big[\int_0^t \Big|\bar{X}^{1,i,M}_{u,K} + f(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K}, \mathcal{L}_u^{\bar{X}^{1,M}_K})[h(\bar{X}^{1,i,M}_{u,K})I_{[0,\underline{t})}(u)$$
$$+ (t-\underline{t})I_{[\underline{t},t]}(u)]\Big|^{\frac{p}{2}}||g(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K}, \mathcal{L}_u^{\bar{X}^{1,M}_K})||^{\frac{p}{2}}du\Big] \tag{4.2.7}$$

Now, we bound the integrand of the integral above. Using condition (4.1.17) we obtain

$$|\bar{X}^{1,i,M}_{u,K} + f(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K}, \mathcal{L}_u^{\bar{X}^{1,M}_K})[h(\bar{X}^{1,i,M}_{u,K})I_{[0,\underline{t})}(u) + (t-\underline{t})I_{[\underline{t},t]}(u)]|^2 =$$
$$= |\bar{X}^{1,i,M}_{u,K}|^2 + 2[h(\bar{X}^{1,i,M}_{u,K})I_{[0,\underline{t})}(u) + (t-\underline{t})I_{[\underline{t},t]}(u)]\Big[\langle\bar{X}^{1,i,M}_{u,K}, f(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K}, \mathcal{L}_u^{\bar{X}^{1,M}_K})\rangle$$
$$+ \frac{1}{2}[h(\bar{X}^{1,i,M}_{u,K})I_{[0,\underline{t})}(u) + (t-\underline{t})I_{[\underline{t},t]}(u)]|f(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K}, \mathcal{L}_u^{\bar{X}^{1,M}_K})|^2\Big]$$
$$\le |\bar{X}^{1,i,M}_{u,K}|^2 + 2[h(\bar{X}^{1,i,M}_{u,K})I_{[0,\underline{t})}(u) + (t-\underline{t})I_{[\underline{t},t]}(u)]\left[\alpha\left(|\bar{X}^{1,i,M}_{u,K}|^2 + |\tilde{X}^{1,i,M}_{u-\tau,K}|^2\right) + \beta\right]$$
$$= (1+2\alpha T)|\bar{X}^{1,i,M}_{u,K}|^2 + 2\alpha T|\tilde{X}^{1,i,M}_{u-\tau,K}|^2 + 2\beta T.$$

This implies

$$|\bar{X}^{1,i,M}_{u,K} + f(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K}, \mathcal{L}_u^{\bar{X}^{1,M}_K})[h(\bar{X}^{1,i,M}_{u,K})I_{[0,\underline{t})}(u) + (t-\underline{t})I_{[\underline{t},t]}(u)]|^{p/2}$$
$$\le 3^{p/4-1}\left[(1+2\alpha T)^{p/4}|\bar{X}^{1,i,M}_{u,K}|^{p/2} + (2\alpha T)^{p/4}|\tilde{X}^{1,i,M}_{u-\tau,K}|^{p/2} + (2\beta T)^{p/4}\right]$$
$$\le C\left(|\bar{X}^{1,i,M}_{u,K}|^{p/2} + |\tilde{X}^{1,i,M}_{u-\tau,K}|^{p/2} + 1\right).$$

Also by condition (4.1.7) one can see that

$$||g(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K})||^{p/2} = \left(||g(\bar{X}^{1,i,M}_{u,K}, \tilde{X}^{1,i,M}_{u-\tau,K}, \mathcal{L}_u^{\bar{X}^{1,M}_K})||^2\right)^{p/4}$$

$$\leq \left[\alpha\left(|\bar{X}^{1,i,M}_{u,K}|^2 + |\tilde{X}^{1,i,M}_{u-\tau,K}|^2\right) + \beta\right]^{p/4}$$
$$\leq C\left(|\bar{X}^{1,i,M}_{u,K}|^{p/2} + |\tilde{X}^{1,i,M}_{u-\tau,K}|^{p/2} + 1\right).$$

Substituting the last two inequalities into (4.2.7), we obtain

$$I_2 \leq C\mathbb{E}\left[\int_0^t \left(1 + |\bar{X}^{1,i,M}_{u,K}|^p + |\tilde{X}^{1,i,M}_{u-\tau,K}|^p\right) du\right]$$
$$\leq C + C\left(\int_0^t \mathbb{E}\left[\sup_{0\leq u\leq s} |X^{1,i,M}_{u,K}|^p\right] ds\right).$$

Now we will bound I_3. Note that t_n is a stopping time of the filtration $\{\mathcal{F}^W_t\}$. Define

$$\mathcal{F}_{t_n} := \{A \in \mathcal{F} : A \cap \{t_n \leq t\} \in \mathcal{F}^W_t\}.$$

By the strong Markov property of the Brownian motion, $\{B_u := W_{t_n+u} - W_{t_n}, u \geq 0\}$ is a standard Brownian motion independent of $\mathcal{F}_{t_n}$ (page 86, Theorem 6.16 in [30]). Thus

$$\mathbb{E}[\sup_{0\leq u\leq s} |W_{t_n+u} - W_{t_n}|^p | \mathcal{F}_{t_n}] = \mathbb{E}[\sup_{0\leq u\leq s} |B_u|^p] \leq Cs^{p/2}.$$

This implies

$$\mathbb{E}[\sup_{t_n\leq u\leq t_{n+1}} |W_u - W_{t_n}|^p | \mathcal{F}_{t_n}] \leq Ch_n^{p/2}. \tag{4.2.8}$$

Combining Jensen's inequality and equation (4.2.8), we arrive at

$$I_3 \leq \mathbb{E}\left[\left(\sum_{n=0}^{n_t-1} ||g(\bar{X}^{1,i,M}_{t_n,K}, \tilde{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})||^2 |\Delta W_n|^2\right)^{p/2}\right]$$
$$= \mathbb{E}\left[\left(\sum_{n=0}^{n_t-1} h_n ||g(\bar{X}^{1,i,M}_{t_n,K}, \tilde{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})||^2 \frac{|\Delta W_n|^2}{h_n}\right)^{p/2}\right]$$
$$\leq T^{p/2-1}\mathbb{E}\left[\sum_{n=0}^{n_t-1} h_n ||g(\bar{X}^{1,i,M}_{t_n,K}, \tilde{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})||^p \frac{E[|\Delta W_n|^p | \mathcal{F}_{t_n}]}{h_n^{p/2}}\right]$$
$$\leq CT^{p/2-1}\mathbb{E}\left[\sum_{n=0}^{n_t-1} h_n ||g(\bar{X}^{1,i,M}_{t_n,K}, \tilde{X}^{1,i,M}_{t_n-\tau,K}, \mathcal{L}^{\hat{X}^{1,M}_K}_{t_n})||^p\right]$$
$$\leq CT^{\frac{p}{2}-1}\mathbb{E}\left[\int_0^{\underline{t}} ||g(\bar{X}^{1,i,M}_{s,K}, \tilde{X}^{1,i,M}_{s-\tau,K}, \mathcal{L}^{\bar{X}^{1,M}_K}_s)||^p ds\right]$$
$$\leq CT^{\frac{p}{2}-1}\mathbb{E}\left[\int_0^{t} ||g(\bar{X}^{1,i,M}_{s,K}, \tilde{X}^{1,i,M}_{s-\tau,K}, \mathcal{L}^{\bar{X}^{1,M}_K}_s)||^p ds\right].$$

Using condition (4.1.7) and Hölder's inequality, we have

$$\begin{aligned}
I_3 &\leq CT^{p/2-1}\mathbb{E}\left[\int_0^t \left(||g(\bar{X}^{1,i,M}_{s,K}, \tilde{X}^{1,i,M}_{s-\tau,K}, \mathcal{L}_s^{\bar{X}_K^{1,M}})||^2\right)^{p/2} ds\right] \\
&\leq CT^{p/2-1}\mathbb{E}\left[\int_0^t \left(\alpha(|\bar{X}^{1,i,M}_{s,K}|^2 + |\tilde{X}^{1,i,M}_{s-\tau,K}|^2) + \beta\right)^{p/2} ds\right] \\
&\leq T^{p/2-1}2^{p-2}C\mathbb{E}\left[\int_0^t \left(\alpha^{p/2}(|\bar{X}^{1,i,M}_{s,K}|^p + |\tilde{X}^{1,i,M}_{s-\tau,K}|^p) + \beta^{p/2}\right) ds\right] \\
&\leq C + C\int_0^t \mathbb{E}\left[\sup_{0\leq u\leq s} |X^{1,i,M}_{u,K}|^p\right] ds.
\end{aligned}$$

For I_4, using the linear condition (3.3.3), we obtain

$$\begin{aligned}
I_4 &\leq \mathbb{E}\left[\sup_{0\leq s\leq t} |g(\bar{X}^{1,i,M}_{s,K}, \tilde{X}^{1,i,M}_{s-\tau,K}, \mathcal{L}_s^{\bar{X}_K^{1,M}})(W_s - W_{\underline{s}})|^p\right] \\
&\leq \mathbb{E}\left[\sup_{0\leq s\leq t}\left\{[(\alpha(|\bar{X}^{1,i,M}_{s,K}|^p + |\tilde{X}^{1,i,M}_{s-\tau,K})|^p) + \beta]\ |(W_s - W_{\underline{s}})|^p\right\}\right] \\
&\leq \mathbb{E}\Bigg[\sum_{n=0}^{n_t-1}[\alpha(|\bar{X}^{1,i,M}_{t_n,K}|^p + |\tilde{X}^{1,i,M}_{s-\tau,K}|^p) + \beta]\mathbb{E}\left[\sup_{t_n\leq s\leq t_{n+1}} |(W_s - W_{t_n})|^{p/2}|\mathcal{F}_{t_n}\right] \\
&\quad + [\alpha(|\bar{X}^{1,i,M}_{t,K}|^p + |\tilde{X}^{1,i,M}_{t-\tau,K}|^p) + \beta]\mathbb{E}\left[\sup_{\underline{t}\leq s\leq t} |(W_s - W_{\underline{t}})|^{p/2}|\mathcal{F}_{\underline{t}}\right]\Bigg] \\
&\leq C + C\int_0^t E\left[\sup_{0\leq u\leq s} |X^{1,i,M}_{u,K}|^p\right] ds.
\end{aligned}$$

Adding all the bounds for I_1 to I_4, we have that for all $t \in [0, T]$

$$\mathbb{E}\left[\sup_{0\leq s\leq t} |X^{1,i,M}_{s,K}|^p\right] \leq C + C\int_0^t E\left[\sup_{0\leq u\leq s} |X^{1,i,M}_{u,K}|^p\right]$$

and by the Gronwall inequality we obtain

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X^{1,i,M}_{t,K}|^p\right] \leq C.$$

Similarly we can show that $\mathbb{E}\left[\sup_{0\leq t\leq T} |X^{2,i,M}_t|^p\right] \leq C$. Thus, the result is proved for $p \geq 4$. For $0 < p < 4$, the result follows proceeding similarly as in the last part of the proof of Lemma 3.3.6. □

Theorem 4.2.2. *If the SDDE* (4.1.2) *satisfies Assumption 4.1.1 and the function* h^δ *satisfies Assumption 4.1.2, then* T *is attainable and for all* $p > 0$ *there exists a constant*

$C > 0$ dependent on T and p, but independent of h_n^δ, such that

$$\mathbb{E}\left[\sup_{0\le t\le T}|X_t^{1,i,M}|^p\right]\bigvee\mathbb{E}\left[\sup_{0\le t\le T}|X_t^{2,i,M}|^p\right]\le C. \tag{4.2.9}$$

Proof. Since the proof is similar to that of Theorem 3.3.5, we omit here. □

4.2.2 Strong convergence of the numerical solutions

Now we will prove the strong convergence for both schemes.

Convergence of scheme 1

We will need the following lemma.

Lemma 4.2.3. *Let $p \ge 1$. Let the MV-SDDE (4.1.2) and the function h^δ satisfy Assumption 4.1.1 and 4.1.2 respectively. Then there exists a positive constant C such that for all $t \in [0,T]$.*

$$\mathbb{E}|X_t^{1,i,M}-\bar{X}_t^{1,i,M}|^{2p}\le C\delta^p, \tag{4.2.10}$$

$$\mathbb{E}|X_t^{1,i,M}-\tilde{X}_t^{1,i,M}|^{2p}\le C\delta^p. \tag{4.2.11}$$

Proof. Let $t\in[0,T]$. Let r be the integer such that $t_r\le t\le t_{r+1}$. Then by definition we have $X_{t_r}^{1,i,M}=\bar{X}_{t_r}^{1,i,M}=\bar{X}_t^{1,i,M}$. Thus

$$X_t^{1,i,M}=\bar{X}_t^{1,i,M}+\int_{t_r}^t f(\bar{X}_s^{1,i,M},\tilde{X}_s^{1,i,M},\mathcal{L}_s^{\bar{X}^{1,M}})ds+\int_{t_r}^t g(\bar{X}_s^{1,i,M},\tilde{X}_s^{1,i,M},\mathcal{L}_s^{\bar{X}^{1,M}})dW_s,$$

which by the Hölder and BDG inequalities yields

$$\begin{aligned}\mathbb{E}|X_t^{1,i,M}-\bar{X}_t^{1,i,M}|^{2p}&\le 2^{2p-1}\mathbb{E}\left|\int_{t_r}^t f(\bar{X}_s^{1,i,M},\tilde{X}_s^{1,i,M},\mathcal{L}_s^{\bar{X}^{1,M}})ds\right|^{2p}\\&\quad+2^{2p-1}\mathbb{E}\left|\int_{t_r}^t g(\bar{X}_s^{1,i,M},\tilde{X}_s^{1,i,M},\mathcal{L}_s^{\bar{X}^{1,M}})dW_s\right|^{2p}\\&\le 2^{2p-1}\mathbb{E}\left[(h_r^\delta)^{2p-1}\int_{t_r}^t|f(\bar{X}_s^{1,i,M},\tilde{X}_s^{1,i,M},\mathcal{L}_s^{\bar{X}^{1,M}})|^{2p}ds\right]\\&\quad+2^{2p-1}\mathbb{E}\left[\left(\int_{t_r}^t|g(\bar{X}_s^{1,i,M},\tilde{X}_s^{1,i,M},\mathcal{L}_s^{\bar{X}^{1,M}})|^2ds\right)^p\right].\end{aligned}$$

From the polynomial growth condition 4.1.8 we have that

$$|f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})|^{2p} \leq \Big[(\gamma(|\bar{X}_s^{1,i,M}|^q + |\tilde{X}_{s-\tau}^{1,i,M})|^q) + \lambda)(|\bar{X}_s^{1,i,M} + |\tilde{X}_{s-\tau}^{1,i,M})|)\Big]^{2p} \tag{4.2.12}$$

$$\leq C(|\bar{X}_s^{1,i,M}|^{2p(q+1)} + |\tilde{X}_{s-\tau}^{1,i,M})|^{2p(q+1)} + 1),$$

and from the linear growth condition 4.1.7 we have that

$$||g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}))||^2 \leq \alpha(|\bar{X}_s^{1,i,M}|^2 + |\tilde{X}_{s-\tau}^{1,i,M}|^2) + \beta. \tag{4.2.13}$$

Hence, by Theorem 4.2.2 and Assumption 4.1.2, we obtain

$$\begin{aligned}\mathbb{E}|X_t^{1,i,M} - \bar{X}_t^{1,i,M}|^{2p} &\leq C\mathbb{E}[(h_r^\delta T)^{2p}(\sup_{t_r\leq s\leq t}|X_s^{1,i,M}|^{2p} + ||\xi||^{2p} + C)] \\ &+ C\mathbb{E}[(h_r^\delta T)^p(\sup_{t_r\leq s\leq t}|X_s^{1,i,M}|^{2p} + ||\xi||^{2p}) + C] \\ &\leq C\delta^p.\end{aligned}$$

To prove assertion (4.2.11), we first prove that there is a constant C such that for all $t \in [0, T]$

$$\mathbb{E}|\tilde{X}_t - \bar{X}_t|^{2p} \leq C\delta^p. \tag{4.2.14}$$

Let $t \in [0, T]$. Let k and n be integers such that $t_k \leq t < t_{k+1}$ and $t_n - \tau \leq t \leq t_{n+1} - \tau$, respectively. Let $r, 0 \leq r \leq k$ be such that $t_{k-r} \leq t_n - \tau \leq t_{k-r+1}$. From (4.1.9) and the definitions of the step processes $\bar{X}$ and $\tilde{X}$, one can see that

$$\begin{aligned}\hat{X}_{t_k}^{1,i,M} &= \hat{X}_{t_{k-r}}^{1,i,M} + \sum_{i=0}^{r-1}[f(\bar{X}_{t_{k-r+i}}^{1,i,M}, \bar{X}_{t_{k-r+i}-\tau}^{1,i,M}, \mathcal{L}_{t_{k-r+i}}^{\hat{X}^{1,M}})h_{k-r+i} \\ &+ g(\bar{X}_{t_{k-r+i}}^{1,i,M}, \bar{X}_{t_{k-r+i}-\tau}^{1,i,M}, \mathcal{L}_{t_{k-r+i}}^{\hat{X}^{1,M}})\Delta W_{k-r+i}] \\ &= \hat{X}_{t_{k-r}}^{1,i,M} + \sum_{i=0}^{r-1}\int_{t_{k-r+i}}^{t_{k-r+i+1}} f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})ds \\ &+ \sum_{i=0}^{r-1}\int_{t_{k-r+i}}^{t_{k-r+i+1}} g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})dW_s \\ &= \hat{X}_{t_{k-r}}^{1,i,M} + \int_{t_{k-r}}^{t_k} f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})ds + \int_{t_{k-r}}^{t_k} g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})dW_s.\end{aligned}$$

Note that since $\bar{X}_t^{1,i,M} = \hat{X}_{t_k}^{1,i,M}$ and $\hat{X}_{t_{k-r}}^{1,i,M} = \bar{X}_{t_{k-r}}^{1,i,M} = \bar{X}_{t_n-\tau}^{1,i,M} = \tilde{X}_{t_n-\tau}^{1,i,M} = \tilde{X}_t^{1,i,M}$, we have that

$$\bar{X}_t^{1,i,M} = \tilde{X}_t^{1,i,M} + \int_{t_{k-r}}^{t_k} f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})ds + \int_{t_{k-r}}^{t_k} g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})dW_s.$$

Also, we have that

$$t_k - t_{k-r} \leq (t_{n+1} - \tau) - (t_n - \tau) + h_{k-r}^\delta = h_n^\delta + h_{k-r}^\delta \leq 2\delta T.$$

Therefore, by (4.2.12),(4.2.13), Assumption 4.1.2 and Theorem 4.2.2 we have that

$$\begin{aligned}
\mathbb{E}|\bar{X}_t^{1,i,M} - \tilde{X}_t^{1,i,M}|^{2p} &\leq 2\mathbb{E}\left|\int_{t_{k-r}}^{t_k} f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})ds\right|^{2p} \\
&+ 2\mathbb{E}\left|\int_{t_{k-r}}^{t_k} g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})dW_s\right|^{2p} \\
&\leq C\mathbb{E}[(t_k - t_{k-r})^{2p}(\sup_{t_k \leq s \leq t} |X_s^{1,i,M}|^{2p} + ||\xi||^{2p} + 1)] \\
&+ C\mathbb{E}[(t_k - t_{k-r})^{p}(\sup_{t_k \leq s \leq t} |X_s^{1,i,M}|^{2p} + ||\xi||^{2p} + 1)] \\
&\leq C\delta^p.
\end{aligned}$$

This together with (4.2.10) imply that

$$\mathbb{E}|X_t^{1,i,M} - \tilde{X}_t^{1,i,M}|^{2p} \leq C\mathbb{E}|X_t^{1,i,M} - \bar{X}_t^{1,i,M}|^{2p} + C\mathbb{E}|\bar{X}_t^{1,i,M} - \tilde{X}_t^{1,i,M}|^{2p} \leq C\delta.$$

□

Now we show the convergence of the scheme 1.

Theorem 4.2.4. *If the SDDE* (4.1.2) *satisfies Assumption 4.1.1 and the time-step function* h *satisfies Assumption 4.1.2, then for all* $p > 0$, *there exists a positive constant* C *independent of* δ *such that*

$$\mathbb{E}\left[\sup_{0 \leq t \leq T} |Y_t^{i,M} - X_t^{1,i,M}|^p\right] \leq C\delta^{p/2}.$$

Proof. Let $p \geq 4$; the result for $0 \leq p < 4$ follows from Hölder's inequality. Define $e_t := Y_t^{i,M} - X_t^{1,i,M}, 0 \leq t \leq T$. Hence

$$\begin{aligned} e_t = &\int_0^t [f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})]ds \\ &+ \int_0^t [g(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})]dW_s. \end{aligned}$$

Applying Itô's formula we obtain

$$\begin{aligned} |e_t|^2 \leq\, & 2\int_0^t \langle e_s, f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})\rangle ds \\ &+ \int_0^t |g(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})|^2 ds \\ &+ 2\int_0^t \langle e_s, (g(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}))dW_s\rangle. \end{aligned} \tag{4.2.15}$$

Note that

$$\begin{aligned} &2\int_0^t \langle e_s, f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})\rangle ds \\ &= 2\int_0^t \langle e_s, f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{X^{1,M}})\rangle ds \\ &+ 2\int_0^t \langle e_s, f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{X^{1,M}}) - f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})\rangle ds \\ &+ 2\int_0^t \langle e_s, f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}) - f(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})\rangle ds \\ &+ 2\int_0^t \langle e_s, f(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}) - f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})\rangle ds. \end{aligned}$$

By condition (4.1.5) we have

$$\begin{aligned} &2\langle e_s, f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{X^{1,M}})\rangle \leq 2|e_s| L \mathbb{W}_2(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{X^{1,M}}) \\ &\leq |e_s|^2 + L^2 \mathbb{W}_2^2(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{X^{1,M}}) \end{aligned} \tag{4.2.16}$$

and

$$\begin{aligned} &2\langle e_s, f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{X^{1,M}}) - f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})\rangle \leq 2|e_s| L \mathbb{W}_2(\mathcal{L}_s^{X^{1,M}}, \mathcal{L}_s^{\bar{X}^{1,M}}) \\ &\leq |e_s|^2 + L^2 \mathbb{W}_2^2(\mathcal{L}_s^{X^{1,M}}, \mathcal{L}_s^{\bar{X}^{1,M}}). \end{aligned} \tag{4.2.17}$$

Using condition (4.1.4) we get

$$
\begin{aligned}
2\langle e_s, & f(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}) - f(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})\rangle \\
& \leq L(|Y_s^{i,M} - X_s^{1,i,M}|^2 + |Y_{s-\tau}^{i,M} - X_{s-\tau}^{1,i,M}|^2) = L(|e_s|^2 + |e_{s-\tau}|^2). \qquad (4.2.18)
\end{aligned}
$$

Condition (4.1.6) implies that

$$
\begin{aligned}
2\langle e_s, & f(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}) - f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})\rangle \\
& \leq |e_s| \; |f(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}) - f(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})| \\
& \leq 2|e_s| Q(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M})(|X_s^{1,i,M} - \bar{X}_s^{1,i,M}| + |X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|) \\
& \leq |e_s|^2 + Q(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M})^2 \; 2(|X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^2 + |X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^2), \\
& \qquad (4.2.19)
\end{aligned}
$$

where $Q(x, y, \bar{x}, \bar{y}) := \gamma(|x|^q + |y|^q + |\bar{x}|^q + |\bar{y}|^q) + \lambda$. In addition, condition (4.1.3) implies that

$$
\begin{aligned}
& ||g(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathbb{W}(\mathcal{L}_s^{Y^M}) - g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})||^2 \\
& \leq L(|Y_s^{i,M} - \bar{X}_s^{1,i,M}|^2 + |Y_{s-\tau}^{i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^2 + \mathbb{W}_2^2(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{\bar{X}^{1,M}})) \\
& = L(|Y_s^{i,M} - X_s^{1,i,M} + X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^2 + |Y_{s-\tau}^{i,M} - X_{s-\tau}^{1,i,M} + X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^2 \\
& + \mathbb{W}_2^2(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{X^{1,M}}) + \mathbb{W}_2^2(\mathcal{L}_s^{X^{1,M}}, \mathcal{L}_s^{\bar{X}^{1,M}})) \\
& \leq 2L(|e_s|^2 + |e_{s-\tau}|^2 + |X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^2 + |X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^2) \\
& + L(\mathbb{W}_2^2(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{X^{1,M}}) + \mathbb{W}_2^2(\mathcal{L}_s^{X^{1,M}}, \mathcal{L}_s^{\bar{X}^{1,M}})). \qquad (4.2.20)
\end{aligned}
$$

Substituting (4.2.16), (4.2.17), (4.2.18), (4.2.19) and (4.2.20) in (4.2.15), we have

$$
\begin{aligned}
|e_t|^2 \leq & \int_0^t \big[(3L+1)|e_s|^2 + 3L|e_{s-\tau}|^2\big] ds \\
& + 2\int_0^t [Q(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M})^2 + L](|X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^2 + |X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^2) ds \\
& + C\int_0^t \big(\mathbb{W}_2^2(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{X^{1,M}}) + \mathbb{W}_2^2(\mathcal{L}_s^{X^{1,M}}, \mathcal{L}_s^{\bar{X}^{1,M}})\big) ds \\
& + 2\int_0^t \langle e_s, (g(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}})) dW_s\rangle.
\end{aligned}
$$

Using Hölder's inequality yields

$$
\begin{aligned}
|e_t|^p &\le (8T)^{p/2-1}\int_0^t ((3L+1)^{p/2}|e_s|^p + (2L)^{p/2}|e_{s-\tau}|^p)ds \\
&+ (8T)^{p/2-1}2^{p/2}\int_0^t [Q(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}) + L]^{p/2} \\
&\qquad\qquad \times (|X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^p + |X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^p)ds \\
&+ (8T)^{p/2-1}C^{p/2}\int_0^t \big(\mathbb{W}_2^p(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{X^{1,M}}) + \mathbb{W}_2^p(\mathcal{L}_s^{X^{1,M}}, \mathcal{L}_s^{\bar{X}^{1,M}})\big)ds \\
&+ 4^{p/2-1}2^{p/2}\left|\int_0^t \langle e_s, (g(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}))dW_s\rangle\right|^{p/2}.
\end{aligned}
$$

Taking the supremum on each side of the previous inequality and then the expectation, we obtain

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{0\le s\le t}|e_s|^p\right] \le C\int_0^t \mathbb{E}\left[\sup_{0\le u\le s}|e_u|^p\right]ds \\
&+ C\int_0^t \big(\mathbb{E}\big[\mathbb{W}_2^p(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{X^{1,M}})\big] + \mathbb{E}\big[\mathbb{W}_2^p(\mathcal{L}_s^{X^{1,M}}, \mathcal{L}_s^{\bar{X}^{1,M}})\big]\big)ds \\
&+ C\int_0^t \mathbb{E}\big[[Q(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}) + L]^{p/2} \\
&\qquad\qquad \times (|X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^p + |X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^p)\big]ds \\
&+ C\mathbb{E}\left[\sup_{0\le s\le t}\left|\int_0^s \langle e_u, (g(Y_u^{i,M}, Y_{u-\tau}^{i,M}, \mathcal{L}_u^{Y^M}) - g(\bar{X}_u^{1,i,M}, \tilde{X}_{u-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}))dW_u\rangle\right|^{p/2}\right].
\end{aligned}
$$

Applying the definition of Wasserstein distance yields

$$
\mathbb{E}\left[\sup_{0\le s\le t}|e_s|^p\right] \le C\int_0^t \mathbb{E}\left[\sup_{0\le u\le s}|e_u|^p\right]ds + J_1 + J_2 + J_3,
$$

where

$$
\begin{aligned}
J_1 &:= C\int_0^t \mathbb{E}[|X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^p]ds; \\
J_2 &:= C\int_0^t \mathbb{E}\big[[Q(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}) + L]^{p/2} \\
&\quad \times (|X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^p + |X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^p)\big]ds; \\
J_3 &:= C\mathbb{E}\left[\sup_{0\le s\le t}\left|\int_0^s \langle e_u, (g(Y_u^{i,M}, Y_{u-\tau}^{i,M}, \mathcal{L}_u^{Y^M}) - g(\bar{X}_u^{1,i,M}, \tilde{X}_{u-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}))dW_u\rangle\right|^{p/2}\right].
\end{aligned}
$$

By Equation (4.2.10), we obtain $J_1 \leq C\delta^{p/2}$. For J_2, by Hölder's inequality one has

$$J_2 \leq C \int_0^t \Big(\mathbb{E}\left[[Q(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}) + L]^p \right] \\ \times \mathbb{E}\left[(|X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^{2p} + |X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^{2p}) \right] \Big)^{1/2} ds. \quad (4.2.21)$$

By Theorem 4.2.2 there exists a constant C such that

$$\mathbb{E}\left[[Q(X_s^{1,i,M}, X_{s-\tau}^{1,i,M}, \bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}) + L]^p \right] \leq C \quad (4.2.22)$$

and by Lemma 4.2.3

$$\mathbb{E}|X_s^{1,i,M} - \bar{X}_s^{1,i,M}|^{2p} \leq C\delta^p \quad \text{and} \quad \mathbb{E}|X_{s-\tau}^{1,i,M} - \tilde{X}_{s-\tau}^{1,i,M}|^{2p} \leq C\delta^p. \quad (4.2.23)$$

Subtituting (4.2.23) and (4.2.22) in (4.2.21), we obtain that $J_2 \leq C\delta^{p/2}$.

Now we estimate J_3. By the definition of Wasserstein distance and the BDG and Hölder's inequalities, one can see that

$$J_3 \leq C\mathbb{E}\left[\left(\int_0^t |e_s|^2 \, |(g(Y_s^{i,M}, Y_{s-\tau}^{i,M}, \mathcal{L}_s^{Y^M}) - g(\bar{X}_s^{1,i,M}, \tilde{X}_{s-\tau}^{1,i,M}, \mathcal{L}_s^{\bar{X}^{1,M}}))|^2 ds \right)^{p/4} \right]$$
$$\leq C\mathbb{E}\left[\int_0^t |e_s|^{p/2} (|\bar{X}_s^{1,i,M} - Y_s^{i,M}|^{p/2} + |\tilde{X}_{s-\tau}^{1,i,M} - Y_{s-\tau}^{i,M}|^{p/2} + \mathbb{W}_2^{p/2}(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{\bar{X}^{1,M}})) ds \right]$$
$$\leq C\left[\int_0^t (\frac{1}{2}\mathbb{E}[|e_s|^p] + 2\mathbb{E}[|\bar{X}_s^{1,i,M} - Y_s^{i,M}|^p] + \mathbb{E}[|\tilde{X}_{s-\tau}^{1,i,M} - Y_{s-\tau}^{i,M}|^p]) ds \right]$$
$$\leq C\mathbb{E}\Big[\int_0^t \mathbb{E}[|e_s|^p] + (\mathbb{E}[|\bar{X}_s^{1,i,M} - X_s^{1,i,M}|^p] + \mathbb{E}[|X_s^{1,i,M} - Y_s^{i,M}|^p] + \mathbb{E}[|\tilde{X}_{s-\tau}^{1,i,M} - X_{s-\tau}^{1,i,M}|^p] \\ + \mathbb{E}[|X_{s-\tau}^{1,i,M} - Y_{s-\tau}^{i,M}|^p]) ds \Big]$$
$$\leq C\mathbb{E}\left[\int_0^t \mathbb{E}[|e_s|^p] + \mathbb{E}[|e_{s-\tau}|^p] + (\mathbb{E}[|\bar{X}_s^{1,i,M} - X_s^{1,i,M}|^p] + \mathbb{E}[\tilde{X}_{s-\tau}^{1,i,M} - X_{s-\tau}^{1,i,M}|^p]) ds \right].$$

By Lemma 4.2.3 we have that

$$J_3 \leq C \int_0^t \mathbb{E}\left[\sup_{0 \leq u \leq s} |e_u|^p \right] ds + C\delta^{p/2}.$$

Collecting the bounds for J_1, J_2 and J_3, we conclude that for all $0 \leq \bar{t} \leq T$, there exists a constant C such that

$$\mathbb{E}\left[\sup_{0 \leq t \leq \bar{t}} |e_t|^p \right] \leq C \int_0^{\bar{t}} \mathbb{E}\left[\sup_{0 \leq u \leq s} |e_u|^p \right] ds + C\delta^{p/2}.$$

The required assertion follows from the Gronwall inequality. □

Convergence of scheme 2

Lemma 4.2.5. *Let the MV-SDDE* (4.1.2) *and the function* h^δ *satisfy Assumption 4.1.1 and 4.1.2 respectively. Let* $t \in [0, T]$ *and* k_n *be the integer such that* $t \in [k_n\delta T, (k_n + 1)\delta T)$*. Then for every* $p > 0$*, there exists a positive constant* C *such that*

$$\mathbb{E}|X_t^{2,i,M} - X_{k_n\delta T}^{2,i,M}|^{2p} \leq C\delta, \tag{4.2.24}$$

$$\mathbb{E}|X_t^{2,i,M} - \bar{X}_t^{2,i,M}|^{2p} \leq C\delta, \tag{4.2.25}$$

$$\mathbb{E}|X_t^{2,i,M} - \tilde{X}_t^{2,i,M}|^{2p} \leq C\delta. \tag{4.2.26}$$

Proof. By the definition of scheme 2, we can write

$$X_t^{2,i,M} = X_{k_n\delta T}^{2,i,M} + \int_{k_n\delta T}^t f(\bar{X}_s^{2,i,M}, \tilde{X}_{s-\tau}^{2,i,M}, \mathcal{L}_s^{\bar{X}^{2,M}})ds + \int_{k_n\delta T}^t g(\bar{X}_s^{2,i,M}, \tilde{X}_{s-\tau}^{2,i,M}, \mathcal{L}_s^{\bar{X}^{2,M}})dW_s.$$

Since $t \in [k_n\delta T, (k_n+1)\delta T)$ we have that $t - K_n\delta T \leq (K_n+1)\delta T - K_n\delta T \leq \delta T$. Thus, by the Hölder and BDG inequalities, conditions (4.1.7) and (4.1.8) and Theorem 4.2.2, we obtain

$$\begin{aligned}
\mathbb{E}[|X_t^{2,i,M} - X_{k_n\delta T}^{2,i,M}|^{2p}] &\leq 2^{2p-1}\mathbb{E}\left[\left|\int_{k_n\delta T}^t f(\bar{X}_s^{2,i,M}, \tilde{X}_{s-\tau}^{2,i,M}, \mathcal{L}_s^{\bar{X}^{2,M}})ds\right|^{2p}\right] \\
&\quad + 2^{2p-1}\mathbb{E}\left[\left|\int_{k_n\delta T}^t g(\bar{X}_s^{2,i,M}, \tilde{X}_{s-\tau}^{2,i,M}, \mathcal{L}_s^{\bar{X}^{2,M}})dW_s\right|^{2p}\right] \\
&\leq (2\delta)^{2p-1}\int_{k_n\delta T}^t \mathbb{E}[|f(\bar{X}_s^{2,i,M}, \tilde{X}_{s-\tau}^{2,i,M}, \mathcal{L}_s^{\bar{X}^{2,M}})|^2]ds \\
&\quad + 2^{2p-1}\mathbb{E}\left[\left(\int_{k_n\delta T}^t |g(\bar{X}_s^{2,i,M}, \tilde{X}_{s-\tau}^{2,i,M}, \mathcal{L}_s^{\bar{X}^{2,M}})|^2 ds\right)^p\right] \\
&\leq (2\delta)^{2p}\mathbb{E}\left[\sup_{k_n\delta T \leq s \leq t}\left\{(\gamma(|\bar{X}_s^{2,i,M}|^q + |\tilde{X}_{s-\tau}^{2,i,M}|^q) + \lambda)(|\bar{X}_s^{2,i,M}| + |\tilde{X}_{s-\tau}^{2,i,M}|)^{2p}\right\}\right] \\
&\quad + C\delta^p\mathbb{E}\left[\left(\sup_{k_n\delta T \leq s \leq t}|\bar{X}_s^{2,i,M}|^{2p} + ||\xi||^{2p} + 1\right)\right] \\
&\leq C\delta^p.
\end{aligned}$$

Since the proof for claims (4.2.25) and (4.2.26) is similar to that of Lemma 4.2.3, we omit it here. □

Theorem 4.2.6. *If the SDDE (4.1.2) satisfies Assumption 4.1.1 and the time-step function h satisfies Assumption 4.1.2, then for all* $p > 0$, *there exists a positive constant* C *independent of* δ *such that*

$$\mathbb{E}\left[\sup_{0\le t\le T}|Y_t^{i,M} - X_t^{2,i,M}|^p\right] \le C\delta^{p/2}.$$

Proof. Using the same arguments as in the proof of Theorem 4.2.4, we obtain

$$\mathbb{E}\left[\sup_{0\le s\le t}|e_s|^p\right] \le C\int_0^t \mathbb{E}\left[\sup_{0\le u\le s}|e_u|^p\right]ds + J_1 + J_2 + J_3,$$

where

$$J_1 := C\int_0^t \big(\mathbb{E}\big[\mathbb{W}_2^p(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{X^{2,M}})\big] + \mathbb{E}\big[\mathbb{W}_2^p(\mathcal{L}_s^{X^{2,M}}, \mathcal{L}_s^{\bar{X}^{2,M}})\big]\big)ds;$$

$$J_2 := C\int_0^t \mathbb{E}\big[[Q(X_s^{2,i,M}, X_{s-\tau}^{2,i,M}, \bar{X}_s^{2,i,M}, \tilde{X}_{s-\tau}^{2,i,M}) + L]^{p/2}$$

$$\times (|X_s^{2,i,M} - \bar{X}_s^{2,i,M}|^p + |X_{s-\tau}^{2,i,M} - \tilde{X}_{s-\tau}^{2,i,M}|^p)\big]ds;$$

$$J_3 := C\mathbb{E}\left[\sup_{0\le s\le t}\left|\int_0^s \langle e_u, (g(Y_u^{i,M}, Y_{u-\tau}^{i,M}, \mathcal{L}_u^{Y^M}) - g(\bar{X}_u^{2,i,M}, \tilde{X}_{u-\tau}^{2,i,M}, \mathcal{L}_s^{\bar{X}^{2,M}}))dW_u\rangle\right|^{p/2}\right].$$

For any $s \in [0, T]$ there is a k_n such that $s \in [k_n\delta T, (k_n + 1)\delta T)$. Hence, by (4.2.24)

$$\begin{aligned}
\mathbb{E}[\mathbb{W}_2^p(\mathcal{L}_s^{Y^M}, \mathcal{L}_s^{X^{2,i,M}})] &\le \mathbb{E}[|Y_s^{i,M} - X_{K_n\delta T}^{2,i,M}|^p] \\
&= \mathbb{E}[|Y_s^{i,M} - X_{K_n\delta T}^{2,i,M} - (X_s^{2,i,M} - X_{K_n\delta T}^{2,i,M}) + (X_s^{2,i,M} - X_{K_n\delta T}^{2,i,M})|^p] \\
&\le 2^{p-1}\mathbb{E}[|Y_s^{i,M} - X_s^{2,i,M}|^p] + 2^{p-1}\mathbb{E}[|X_s^{2,i,M} - X_{K_n\delta T}^{2,i,M}|^p] \\
&\le C\mathbb{E}[\sup_{0\le u\le s}|e_u|^p] + C\delta^{p/2}.
\end{aligned}$$

Also, since the sequence of times $\{k_n\delta T\}_{n\in\mathbb{N}}$ is contained in the sequence of times $\{t_n\}_{n\in\mathbb{N}}$, we have that $X_\delta^{2,i,M} = \bar{X}_\delta^{2,i,M}$. Thus,

$$\mathbb{E}\big[\mathbb{W}_2^p(\mathcal{L}_s^{X^{2,M}}, \mathcal{L}_s^{\bar{X}^{2,M}})\big] \le \mathbb{E}[|X_{K_n\delta T}^{2,i,M} - \bar{X}_{K_n\delta T}^{2,i,M}|^2] = 0.$$

Thus,

$$J_1 \le C\int_0^t \mathbb{E}\left[\sup_{0\le u\le s}|e_u|^p\right]ds + C\delta^{p/2}.$$

The terms J_2 and J_3 are bounded in identical way as in Theorem 4.2.4. An application of the Gronwall inequality yields the result. □

Chapter 5

Multilevel Monte Carlo EM scheme for MV-SDEs with small noise

5.1 Introduction

An important problem in science is to compute $\mathbb{E}[\Psi(X_T)]$ where $\{X_t\}_{0\leq t\leq T}$ is the solution to an SDE and $\Psi : \mathbb{R}^d \to \mathbb{R}$. Among all the methods that allow us to compute the previous expectation, Monte Carlo simulation is arguably the more flexible. Its drawback is the high computational cost. Therefore a lot of effort has been placed to reduce this cost. In 2008, Giles, in a very relevant paper, [15], proposed the multilevel Monte Carlo (MLMC) method which greatly reduces the computational cost with respect to the standard Monte Carlo (MC) method. In the standard MC method, if δ is the accuracy in terms of confidence intervals, the computation of $\mathbb{E}[\Psi(X_T)]$, where X_T is simulated using the Euler-Maruyama (EM) method, has a computation cost (measured as the number of times that the random number generator is called) that scales as δ^{-3}. However, following [15], we can see the MLMC method combined with the EM scheme, scales like $\delta^{-2}(\log\delta)^2$ (see next section for an overview explanation). Since [15], numerous papers have appeared to customize, adapt and extend the principles of multilevel Monte Carlo method to specific problems. One of these papers is [1], where the authors applied the multilevel Monte Carlo framework to SDEs with small noise. They compare the computation cost derived from the standard Monte Carlo method (combined with discretization algorithms tailored to the small noise setting) versus the multilevel Monte Carlo method combined with the Euler-Maruyama (EM) scheme. They found that when $\delta \leq \varepsilon^2$, there is not benefit from using discretization methods customized for the small noise case. Moreover, if $\delta \geq e^{-\frac{1}{\varepsilon}}$, the EM scheme combined with the MLMC method leads to a cost $O(1)$. This is the same cost we would have with the standard MC method if we had X_T as a formula of W_T, so no discretization method was required. Here, we extend the work from [1] to McKean-Vlasov SDEs (MV-SDEs) with small noise and we obtained the same estimate for the variance of two coupled paths. This means that the additional McKean-Vlasov component does not add computational complexity (per equation in the system of particles) and the conclusion we mentioned above about the computational cost of the method remains valid in our case.

5.2 Computational complexity of the standard Monte Carlo and the Multilevel Monte Carlo methods

In this section, we will discuss the computational complexity of the Monte Carlo and the multilevel Monte Carlo methods in the context of solving the following problem. We want to obtain and approximation for

$$\mathbb{E}[\Psi(X_T)], \tag{5.2.1}$$

where $\{X_t\}_{0 \leq t \leq T}$ is the solution to an SDE and $\Psi : \mathbb{R}^d \to \mathbb{R}$ is a globally Lipschitz function. The problem (5.2.1) is very relevant in financial mathematics, since this is one of the ways used to price financial options. In this case Ψ is the payoff function and X_T is the underlying asset price at time T. But it is also a common/important problem in other areas of science. The problem 5.2.1 can be solved applying finite differences methods to the PDE resulting from applying Feynman-Kac theorem to the

SDE (for which X is the solution). A drawback of this method is that the cost of the computation depends on dimension of the PDE. In practice this method becomes too expensive for equations of higher order than three. One of the advantages of the Monte Carlo method is that its computational cost does not depend on the dimension of the SDE. Another advantage is its flexibility and robusticity to handle all types of SDEs and Ψ functions. However, Monte Carlo simulation is an expensive method. Therefore, a lot of research have been conducted to try to reduce this cost. In 2008, Giles, in a very relevant paper, [15], proposed the multilevel Monte Carlo method which greatly reduces the computational cost with respect to the standard Monte Carlo method. Now, we analyze the computational complexity of the MLMC method versus the MC method when solving (5.2.1), where X_T is discretized using the Euler-Maruyama method. In the rest of this section, we will use the simplest example of SDE to ilustrate in a clear way the usefulness of the next section. Let $\Delta \in (0,1)$ and K a positive integer such that $\Delta = T/K$. Consider an SDE of the type

$$X_t = f(X_t)dt + g(X_t)dW_t, \quad t \in [0,T], \quad X_0 = x_0 \in \mathbb{R}, \tag{5.2.2}$$

where W is a *one*-dimensional Brownian motion. The EM approximate solution of the previous SDE is defined as

$$Y_0 := x_0, \quad Y_k := f(Y_{k-1})\Delta + g(Y_{k-1})\Delta W_k, k = 1, ..., K \tag{5.2.3}$$

where $\Delta W_k = W_{t_{k+1}} - W_{t_k}$. We want to generate M samples of Y_K. In order to do that, we generate $M \cdot K$ independent and identically distributed (i.i.d.) standard normal random variables $Z_k^i, k = 1, .., K, i = 1, ..., M$. Define $\Delta W_k^i := \sqrt{\Delta} Z_k^i$. Then for $i = 1, ..., M$, using Equation (5.2.3) we generate the paths $Y_k^i, k = 1, ..., K$ of the EM approximate solution. We define the standard MC approximation, Q_M, of problem (5.2.1) as the sample mean

$$Q_M := \frac{1}{M}\sum_{i=1}^{M} \Psi(Y_K^i).$$

Let δ be defined as the accuracy of the approximation in terms of confidence intervals, i.e. the total error, $e := \mathbb{E}[\Psi(X_T)] - Q_M$, will be within the interval $(-\delta, \delta)$ with a confidence level of α (usually in practice $\alpha = 0.99$ or $\alpha = 0.95$). We also define the mean square error, MSE, as

$$MSE := \mathbb{E}[e^2] = \mathbb{E}[(Q_M - \mathbb{E}[\Psi(X_T)])^2].$$

Now, by adding and subtracting $\mathbb{E}[\Psi(Y_K)]$ to Q_M in the expression above, we obtain

$$\begin{aligned}\mathbb{E}[(Q_M - \mathbb{E}[\Psi(X_T)])^2] &= \mathbb{E}[(Q_M - \mathbb{E}[\Psi(Y_K)] + \mathbb{E}[\Psi(Y_K)] - \mathbb{E}[\Psi(X_T)])^2] \\ &\leq 2\mathbb{E}[(Q_M - \mathbb{E}[\Psi(Y_K)])^2] + 2(\mathbb{E}[\Psi(Y_K)] - \mathbb{E}[\Psi(X_T)])^2. \quad (5.2.4)\end{aligned}$$

We note that

$$\mathbb{E}[(Q_M - \mathbb{E}[\Psi(Y_K)])^2] = \mathrm{Var}(Q_M) = \mathrm{Var}\left(\frac{1}{M}\sum_{i=1}^{M}\Psi(Y_K^i)\right) = M^{-1}\mathrm{Var}(\Psi(Y_K)).$$

Since $\mathrm{Var}(\Psi(Y_K))$ is a constant that does not depend of Δ nor M, we have $\mathbb{E}[(Q_M - \mathbb{E}[\Psi(Y_K)])^2] = \mathcal{O}(M^{-1})$. To estimate $(\mathbb{E}[\Psi(Y_K)] - \mathbb{E}[\Psi(X_T)])^2$ we realize that the EM method has weak order one, i.e.

$$\mathbb{E}[\Psi(X_T)] - \mathbb{E}[\Psi(Y_K)] = \mathcal{O}(\Delta).$$

Hence, $(\mathbb{E}[\Psi(X_T)] - \mathbb{E}[\Psi(Y_K)])^2 = \mathcal{O}(\Delta^2)$. Thus,

$$MSE = \mathbb{E}[(Q_M - \mathbb{E}[\Psi(X_T)^2] = \mathcal{O}(M^{-1}) + \mathcal{O}(\Delta^2).$$

By definition δ^2 scales like the MSE, so in order to achieve accuracy of δ we must have

$$\delta = \sqrt{O(M^{-1})} + \sqrt{\mathcal{O}(\Delta^2)},$$

and this holds true if $M = \mathcal{O}(\delta^{-2})$ and $\Delta = \mathcal{O}(\delta)$.

We define the computational complexity, CC_{MC}, of solving problem (5.2.1) as the number of times the random number generator is called in order to compute Q_M, i.e. $CC_{MC} = MK = M(T/\Delta)$. Therefore,

$$CC_{MC} = \mathcal{O}(M/\Delta) = \mathcal{O}(\delta^{-3}).$$

Now we estimate the computational complexity of the MLMC method. We will consider samples of the EM approximate solution at different discretization levels $l =$

$0, 1, ..., L$. At level l, the stepsize is defined as $\Delta_l := \tilde{K}^{-l}T$, where $\tilde{K} > 1$ is a constant, so for level l we have steps $1, 2, ..., \tilde{K}^{-l} =: K_l$. This means that for level l, we reach T at step K_l. Let Y_{l,K_l} denote the EM approximation at time $\Delta_l K_l$ for level l. Then $\mathbb{E}[\Psi(Y_{L,K_L})]$ can be written as

$$\mathbb{E}[\Psi(Y_{L,K_L})] = \mathbb{E}[\Psi(Y_{0,K_0})] + \sum_{l=1}^{L} \mathbb{E}[\Psi(Y_{l,K_l}) - \Psi(Y_{l-1,K_{l-1}})].$$

We define the MLMC approximation of problem (5.2.1) as

$$Q_{ML} := \frac{1}{N_0} \sum_{i=1}^{N_0} \Psi(Y^i_{0,K_0}) + \sum_{l=1}^{L} \frac{1}{N_l} \sum_{i=1}^{N_l} (\Psi(Y^i_{l,K_l}) - \Psi(Y^i_{l-1,K_{l-1}})) \tag{5.2.5}$$

where Y^i_{l,K_l} is the sample i of the EM approximate solution Y_{l,K_l} and N_l is the number of simulations for the paths generated at level l. The samples Y^i_{l,K_l} and $Y^i_{l-1,K_{l-1}}$ are built using the same discretized Brownian paths with the different stepsizes Δ_l and Δ_{l-1} respectively. We say that the paths $Y^i_{l,k_l}, k_l = 1, .., K_l$ and $Y^i_{l-1,k_{l-1}}, k = 1, .., K_{l-1}$ are coupled. The next level coupled paths $Y^i_{l+1,k_{l+1}}, k = 1, .., K_{l+1}$ and $Y^i_{l,k_l}, k = 1, .., K_l$ are generated using new Brownian paths (so new i.i.d. standard normal random variables will be required). We shall decompose the MSE in the same way as we did before for the standard MC method. We have that

$$\begin{aligned}\mathbb{E}[(Q_{ML} - \mathbb{E}[\Psi(X_T)])^2] &= (Q_{ML} - \mathbb{E}[\Psi(Y_{L,K_L})] + \mathbb{E}[\Psi(Y_{L,K_L})] - \mathbb{E}[\Psi(X_T)])^2 \\ &\leq 2\mathbb{E}[(Q_{ML} - \mathbb{E}[\Psi(Y_{L,K_L})])^2] + 2(\mathbb{E}[\Psi(Y_{L,K_L})] - \mathbb{E}[\Psi(X_T)])^2.\end{aligned} \tag{5.2.6}$$

We want to find out the values of L and $N_l, l = 1, ..., L$, that will provide the target accuracy δ. Set

$$L = \frac{\log \delta^{-1}}{\log \tilde{K}}. \tag{5.2.7}$$

Then

$$\Delta_L = \frac{T}{\tilde{K}^L} = \frac{T}{e^{L \log \tilde{K}}} = T\delta = \mathcal{O}(\delta). \tag{5.2.8}$$

Since the EM method has a weak convergence order of one, we obtain

$$(\mathbb{E}[\Psi(Y_{L,K_L})] - \mathbb{E}[\Psi(X_T)])^2 = \mathcal{O}(\delta^2), \tag{5.2.9}$$

which is in line with the required target accuracy. Now we estimate the term $\mathbb{E}[(Q_{ML} - \mathbb{E}[\Psi(Y_{L,K_L})])^2]$. Using the definition of Q_{ML}, we have that

$$\begin{aligned}
\mathbb{E}[(Q_{ML} - \mathbb{E}[\Psi(Y_{L,K_L})])^2] &= \mathrm{Var}(Q_{ML}) \\
&= \mathrm{Var}\left(\frac{1}{N_0}\sum_{i=1}^{N_0}\Psi(Y^i_{0,K_0}) + \sum_{l=1}^{L}\frac{1}{N_l}\sum_{i=1}^{N_l}(\Psi(Y^i_{l,K_l}) - \Psi(Y^i_{l-1,K_{l-1}}))\right) \\
&= \frac{1}{N_0}\mathrm{Var}(\Psi(Y_{0,K_0})) + \sum_{l=1}^{L}\frac{1}{N_l}\mathrm{Var}(\Psi(Y_{l,K_l}) - \Psi(Y_{l-1,K_{l-1}})).
\end{aligned}$$

We will see in Lemma 5.2.1 that $\mathrm{Var}(\Psi(Y_{l,K}) - \Psi(Y_{l-1,K_{l-1}})) = \mathcal{O}(\Delta)$. We also know that $\mathrm{Var}(Y_{0,K_0})$ is a constant that does not depend of Δ nor N_0. Hence,

$$\mathbb{E}[(Q_{ML} - \mathbb{E}[\Psi(Y_{L,K_L})])^2] = \frac{1}{N_0}C + \sum_{l=1}^{L}\frac{1}{N_l}\mathcal{O}(\Delta_l).$$

Setting $N_l = \delta^{-2}L\Delta_l$, we obtain

$$\mathbb{E}[(Q_{ML} - \mathbb{E}[\Psi(Y_{L,K_L})])^2] = \mathcal{O}(\delta^2) + \sum_{l=1}^{L} O(\delta^2 L^{-1}) = \mathcal{O}(\delta^2),$$

which is the accuracy required. Therefore, the mathematical complexity of the MLMC method, CC_{MLMC}, is

$$CC_{MLMC} = \sum_{l=0}^{L} N_l\Delta_l^{-1} = \sum_{l=0}^{L}\delta^{-2}L\Delta_l\Delta_l^{-1} = \delta^{-2}L^2.$$

Substituting Equation (5.2.7) into the last one, we obtain

$$CC_{MLMC} = \mathcal{O}(\delta^{-2}(\log\delta)^2).$$

Lemma 5.2.1. *Consider the EM aproximate solutions Y_{l,K_l} and $Y_{l-1,K_{l-1}}$. Then*

$$\mathrm{Var}(\Psi(Y_{l,K}) - \Psi(Y_{l-1,K_{l-1}})) = \mathcal{O}(\Delta).$$

Proof. The EM method has a strong convergence order of 1/2 and the function Ψ satisfies de global Lipstitz conditions, we have

$$\mathrm{Var}(\Psi(Y_{l,K_l}) - \Psi(X_T)) \leq \mathbb{E}[|\Psi(Y_{l,K_l}) - \Psi(X_T)|^2]$$

$$
\begin{aligned}
&\leq C\mathbb{E}[|Y_{l,K_l} - X_T|^2] \\
&\leq \mathcal{O}(\Delta_l).
\end{aligned}
$$

Using the above inequality and (B.0.2) yields

$$
\begin{aligned}
\text{Var}(\Psi(Y_{l,K_l}) - \Psi(Y_{l-1,K_{l-1}})) &= \text{Var}(\Psi(Y_{l,K_l}) - \Psi(X_T) + \Psi(X_T) - \Psi(Y_{l-1,K_{l-1}})) \\
&\leq 2\text{Var}(\Psi(Y_{l,K_l}) - \Psi(X_T)) + 2\text{Var}(\Psi(X_T) - \Psi(Y_{l-1,K_{l-1}})) \\
&\leq \mathcal{O}(\Delta_l).
\end{aligned}
$$

□

Note that in the computational complexity analysis, in the standard Monte Carlo method is not required the strong convergence order of the discretization scheme, it only uses the weak convergence order. However in the MLMC methods, both weak and strong convergence orders are used. The Milstein scheme has weak order of converge of 1, the same as the EM scheme, but it has a strong order of convergence of 1. This implies that using the Milstein scheme does not affect the computational cost of the standard MC method. However, if we use the Milstein scheme combined with the MLMC method, we can obtain a computational complexity of $\mathcal{O}(\delta^{-2})$, [16]. This is the same computational complexity that we would obtain with the standard MC method if we did not have any discretization error at all, i.e. if we could compute the solution Y_K as function of W_T, so no discretization scheme was needed. The MLMC method represent a huge improvement in computation cost with respect to the standard MC method. For example for an accuracy of $\delta = 0.01$, the MLMC method is 100 times faster than the standard MC method. The prior anaylsis can be done with other types of SDEs. Note that the key differential component in the computational complexity analysis of the MLMC method for different SDEs is to estimate the variance between two coupled paths (Lemma 5.2.1). The aim of this chapter is to estimate the variance between two couple paths for MV-SDEs with small noise.

5.3 The EM Scheme for MV-SDEs with small noise

Let W be a $\bar{d}$-dimensional Brownian motion. The MV-SDE with small noise that we will be working on in this chapter, has the form

$$dX^\varepsilon(t) = f(X^\varepsilon(t), \mathcal{L}_t^X)dt + \varepsilon g(X^\varepsilon(t), \mathcal{L}_t^X)dW(t), \quad t \geq 0 \tag{5.3.1}$$

with initial data $X^\varepsilon(0) = x_0 \in L^2_{\mathcal{F}_0}(\Omega, \mathbb{R}^d)$, where $\varepsilon \in (0,1)$, $\mathcal{L}_t^X$ is the law (or distribution) of $X^\varepsilon(t)$, and

$$f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \text{ and } g : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times\bar{d}}.$$

We assume that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ is atomless so that, for any $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, there exists a random variable $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ such that $\mu = \mathcal{L}^X$. Let f_i be the i^{th} component of f. Then for $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we denote

$$\nabla f_i(x,\mu) := \left(\frac{\partial f_i(x,\mu)}{\partial x_1}, ..., \frac{\partial f_i(x,\mu)}{\partial x_d}\right),$$
$$\nabla^2 f_i(x,\mu) := \begin{bmatrix} \frac{\partial^2 f_i(x,\mu)}{\partial x_1^2} & \cdots & \frac{\partial^2 f_i(x,\mu)}{\partial x_1 x_d} \\ \vdots & \vdots & \vdots \\ \frac{\partial^2 f_i(x,\mu)}{\partial x_d x_1} & \cdots & \frac{\partial^2 f_i(x,\mu)}{\partial x_d^2} \end{bmatrix}.$$

As we have seen in the preliminaries, due to the propagation of chaos result 2.2.9, Equation (5.3.1) can be regarded as the limit of the following interacting particle system

$$\mathrm{d}X^{\varepsilon,i,M}(t) = f(X^{\varepsilon,i,M}(t), \mathcal{L}_t^{\varepsilon,X,M})dt + \varepsilon g(X^{\varepsilon,i,M}(t), \mathcal{L}_t^{\varepsilon,X,M})dW^i(t), \quad t \in [0,T]. \tag{5.3.2}$$

where $\mathcal{L}_t^{\varepsilon,X,M} := \frac{1}{M}\sum_{i=1}^{M} \delta_{X^{\varepsilon,i,M}(t)}$. Our main task in the rest of the chapter is to discretize (5.3.2) using the EM scheme and estimate the variance of two coupled paths in the Multilevel Monte Carlo setting. This directly translate into the computational cost of solving $\mathbb{E}[\Psi((X^{\varepsilon,i,M}(T)]$, see the previous section and [1] for details.

By solving the expectation $\mathbb{E}[\Psi((X^{\varepsilon,i,M}(T)]$ by MC simulation, where X is the solution to an MV-SDEs, we can solve nonlinear PDEs for the function Ψ numerically, see [50] and [51]. This can be quite useful when we are dealing with a high order PDE for which specific numerical methods for PDEs become too expensive.

We shall impose the following hypothesis on the functions f and g:

Assumption 5.3.1. *There exists a positive constant* $K > 0$ *such that*

$$|f(x,\mu) - f(y,\nu)|^2 \vee |g(x,\mu) - g(y,\nu)|^2 \leq K\big(|x-y|^2 + \mathbb{W}_2^2(\mu,\nu)\big), \tag{5.3.3}$$

hold for any $x,\ y \in \mathbb{R}^d$, $\mu,\ \nu \in \mathcal{P}_2(\mathbb{R}^d)$*. Furthermore there exists a positive constant* $\bar{K}$ *such that*

$$|\nabla f(x,\mu)|^2 \vee |\nabla^2 f(x,\mu)|^2 \vee |\partial_\mu f(x,\mu)(y)|^2 \vee |\partial_\mu^2 f(x,\mu)(y)|^2 \vee |\partial_\mu \nabla f(x,\mu)|^2 \leq \bar{K}$$

for all $x,\ y \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$*. In addition, there exists a positive constant* K *such that*

$$|\partial_\mu f(x,\mu)(y) - \partial_\mu f(\bar{x},\nu)(\bar{y})|^2 \leq K\big(|x-\bar{x}|^2 + |y-\bar{y}|^2 + \mathbb{W}_2^2(\mu,\nu)\big). \tag{5.3.4}$$

for all $x, y, \bar{x}, \bar{y} \in \mathbb{R}^d$, $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$.

Remark 5.3.1. Assumption 5.3.1 implies the existence and uniqueness of equation (5.3.1), Theorem 2.2.4. Moreover, if Assumption 5.3.1, then

$$|f(x,\mu)|^2 \vee |g(x,\mu)|^2 \leq \beta(1 + |x|^2 + W_2^2(\mu)),$$

where $\beta = 2\max\{1, |f(0,\delta_0)|, |g(0,\delta_0)|\}$, and for any $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

$$\langle x - y, f(x,\mu) - f(y,\nu)\rangle \leq \alpha(|x-\bar{x}|^2 + \mathbb{W}_2^2(\mu,\nu)),$$

where $\bar{\alpha} = \frac{1}{2}(1+K)$.

Lemma 5.3.2. *Let Assumption 5.3.1 hold. Then, for any* $T > 0$ *and* $p > 0$*, we have*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X^\varepsilon(t)|^p\right] \leq C.$$

Proof. Let $p \geq 4$. From 5.3.1 we have that

$$|X^\varepsilon(t)|^p = \left| x_0 + \int_0^t f(X^\varepsilon(s), \mathcal{L}_s^X) ds + \varepsilon \int_0^t g(X^\varepsilon(s), \mathcal{L}_s^X) dW(s) \right|^p.$$

By the Hölder and the BDG inequalities we have that for every $\hat{t} \leq T$

$$\mathbb{E}[\sup_{0 \leq t \leq \hat{t}} |X_t^\varepsilon|^p] \leq 3^{p-1}|x_0|^p + (3T)^{p-1}\mathbb{E}\int_0^{\bar{t}} |f(X^\varepsilon(s), \mathcal{L}_s^X)|^p ds$$
$$+ \varepsilon 3^{p-1} C \mathbb{E}\left[\left(\int_0^{\bar{t}} |g(X^\varepsilon(s), \mathcal{L}_s^X)|^2 ds\right)^{p/2}\right].$$

By Remark 5.3.1, one can see that for every $\hat{t} \leq T$

$$\mathbb{E}[\sup_{0 \leq t \leq \hat{t}} |X_t^\varepsilon|^p] \leq C + C\int_0^{\bar{t}} \mathbb{E}[\sup_{0 \ \leq s \leq t} |X_s^\varepsilon|^p] dt.$$

The required assertion follows from the Gronwall inequality. Thus, the result is proved for $p \geq 4$. For $0 < p < 4$, the result follows proceeding similarly as in the last part of the proof of Lemma 3.3.6. □

. We now introduce the EM scheme for (5.3.1). Given any time $T > 0$, assume that there exists a positive integer such that $h = \frac{T}{m}$, where $h \in (0,1)$ is the step size. Let $t_n = nh$ for $n \geq 0$. Compute the discrete approximations $Y_{h,n}^{\varepsilon,i,M} = Y_h^{\varepsilon,i,M}(t_n)$ by setting $Y_h^{\varepsilon,i,M}(0) = x_0$ and forming

$$Y_{h,n+1}^{\varepsilon,i,M} = Y_{h,n}^{\varepsilon,i,M} + f(Y_{h,n}^{\varepsilon,i,M}, \mathcal{L}_h^{\varepsilon,Y_n,M})h + \varepsilon g(Y_{h,n}^{\varepsilon,i,M}, \mathcal{L}_h^{\varepsilon,Y_n,M})\Delta W^i(t_n), \tag{5.3.5}$$

where $\mathcal{L}_h^{\varepsilon,Y_n,M} = \frac{1}{M}\sum_{j=1}^{M} \delta_{Y_{h,n}^{\varepsilon,j,M}}$ and $\Delta W(t_n) = W(t_{n+1}) - W(t_n)$.

Let

$$Y_h^{\varepsilon,i,M}(t) = Y_{h,k}^{\varepsilon,i,M}, \quad t \in [t_k, t_{k+1}). \tag{5.3.6}$$

For convenience, we define $\mathcal{L}_{h,t}^{\varepsilon,Y,M} = \frac{1}{M}\sum_{j=1}^{M} \delta_{Y_h^{\varepsilon,i,M}(t)}$ and $\eta_h(t) := \lfloor t/h \rfloor h$ for $t \geq 0$. Then one observes $\mathcal{L}_{h,t}^{\varepsilon,Y,M} = \mathcal{L}_{h,\eta_h(t)}^{\varepsilon,Y,M} = \mathcal{L}_h^{\varepsilon,Y_k,M}$, for $t \in [t_k, t_{k+1})$. We now define the EM continuous approximate solution as follows:

$$\bar{Y}_h^{\varepsilon,i,M}(t) = x_0^i + \int_0^t f(Y_h^{\varepsilon,i,M}(s), \mathcal{L}_{h,s}^{\varepsilon,Y,M})\mathrm{d}s + \varepsilon\int_0^t g(Y_h^{\varepsilon,i,M}(s), \mathcal{L}_{h,s}^{\varepsilon,Y,M})\mathrm{d}W^i(s), \ t \geq 0. \tag{5.3.7}$$

Lemma 5.3.3. *Let Assumption 5.3.1 hold. Then, for any $T > 0$ and $p \geq 2$, we have*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |\bar{Y}_h^{\varepsilon,i,M}(t)|^p\right] \leq C.$$

Proof. The proof is the same as the one in Lemma 5.3.2. □

Lemma 5.3.4. *Let Assumption 5.3.1 hold. Then, for any $p \geq 2$, we have*

$$\sup_{0\leq t\leq T} \mathbb{E}[|\bar{Y}_h^{\varepsilon,i,M}(t) - Y_h^{\varepsilon,i,M}(t)|^p] \leq Ch^p + C\varepsilon^p h^{p/2}.$$

Proof. Let n be such that $t_n \leq t \leq t_{n+1}$. From (5.3.7) we have

$$\bar{Y}_h^{\varepsilon,i,M}(t) - Y_h^{\varepsilon,i,M}(t) = \int_{t_n}^t f(Y_h^{\varepsilon,i,M}(s), \mathcal{L}_{h,s}^{\varepsilon,Y,M})ds + \varepsilon \int_{t_n}^t g(Y_h^{\varepsilon,i,M}(s), \mathcal{L}_{h,s}^{\varepsilon,Y,M})dW^i(s)$$

By Remark 5.3.1 and the BDG inequality, one has

$$\begin{aligned}\mathbb{E}|\bar{Y}_h^{\varepsilon,i,M}(t) - Y_h^{\varepsilon,i,M}(t)|^p \leq & 2^{p-1}h^{p-1}\mathbb{E}\int_{t_n}^t |f(Y_h^{\varepsilon,i,M}(s), \mathcal{L}_{h,s}^{\varepsilon,Y,M})|^p ds \\ & + \varepsilon^p h^{\frac{p}{2}-1}\mathbb{E}\int_{t_n}^t |g(Y_h^{\varepsilon,i,M}(s), \mathcal{L}_{h,s}^{\varepsilon,Y,M})|^p ds \\ \leq & Ch^p + C\varepsilon^p h^{\frac{p}{2}}.\end{aligned}$$

The proof is therefore complete. □

We now reveal the error between the numerical solution (5.3.7) and the exact solution (5.3.1).

Theorem 5.3.5. *Let Assumption 5.3.1 hold, assume that $\Psi : \mathbb{R}^d \to \mathbb{R}$ has continuous second order derivative and there exists a constant C such that*

$$\left|\frac{\partial \Psi}{\partial x_i}\right| \leq C$$

for any $i = 1, 2, \cdots, d$. Then we have

$$\sup_{0\leq t\leq T} \mathbb{E}|\Psi(X^{\varepsilon,i,M}(t)) - \Psi(\bar{Y}_h^{\varepsilon,i,M}(t))|^2 = Ch^2 + Ch\varepsilon^2.$$

Proof. By Assumption 5.3.1 and Lemma 5.3.4, one can see that

$$
\begin{aligned}
&\sup_{0\le t\le T}\mathbb{E}|X^{\varepsilon,i,M}(t)-\bar{Y}_h^{\varepsilon,i,M}(t)|^2\\
&\le 2T\mathbb{E}\int_0^T |f(X^{\varepsilon,i,M}(s),\mathcal{L}_s^{\varepsilon,X,M})-f(Y_h^{\varepsilon,i,M}(s),\mathcal{L}_{h,s}^{\varepsilon,Y,M})|^2ds\\
&+8\sqrt{T}\varepsilon^2\mathbb{E}\int_0^T |g(X^{\varepsilon,i,M}(s),\mathcal{L}_s^{\varepsilon,X,M})-g(Y_h^{\varepsilon,i,M}(s),\mathcal{L}_{h,s}^{\varepsilon,Y,M})|^2ds\\
&\le 2TK\mathbb{E}\int_0^T \big(|X^{\varepsilon,i,M}(s)-Y_h^{\varepsilon,i,M}(s)|^2+W_2^2(\mathcal{L}_s^{\varepsilon,X,M},\mathcal{L}_{h,s}^{\varepsilon,Y,M}\big)\\
&+8K\sqrt{T}\varepsilon^2\mathbb{E}\int_0^T \big(|X^{\varepsilon,i,M}(s)-Y_h^{\varepsilon,i,M}(s)|^2+W_2^2(\mathcal{L}_s^{\varepsilon,X,M},\mathcal{L}_{h,s}^{\varepsilon,Y,M}\big)\\
&\le 4TK\mathbb{E}\int_0^T |X^{\varepsilon,i,M}(s)-\bar{Y}_h^{\varepsilon,i,M}(s)|^2ds+4TK\mathbb{E}\int_0^T|\bar{Y}^{\varepsilon,i,M}(s)-Y_h^{\varepsilon,i,M}(s)|^2ds\\
&+16K\sqrt{T}\varepsilon^2\mathbb{E}\int_0^T |X^{\varepsilon,i,M}(s)-\bar{Y}_h^{\varepsilon,i,M}(s)|^2ds+16K\sqrt{T}\varepsilon^2\mathbb{E}\int_0^T|\bar{Y}^{\varepsilon,i,M}(s)-Y_h^{\varepsilon,i,M}(s)|^2ds\\
&\le Ch^2+C\varepsilon^2h+C\varepsilon^2\int_0^T\sup_{0\le t\le s}\mathbb{E}|X^{\varepsilon,i,M}(s)-\bar{Y}_h^{\varepsilon,i,M}(s)|^2ds+C\varepsilon^2h^2+C\varepsilon^4h.
\end{aligned}
\tag{5.3.8}
$$

The Gronwall inequality implies that

$$\sup_{0\le t\le T}\mathbb{E}|X^{\varepsilon,i,M}(t)-\bar{Y}_h^{\varepsilon,i,M}(t)|^2\le Ch^2+\varepsilon^2h.$$

Since Ψ has continuous bounded first order derivative, we immediately get

$$\sup_{0\le t\le T}\mathbb{E}|\Psi(X^{\varepsilon,i,M}(t))-\Psi(\bar{Y}_h^{\varepsilon,i,M}(t))|^2\le C\sup_{0\le t\le T}\mathbb{E}|X^{\varepsilon,i,M}(t)-\bar{Y}_h^{\varepsilon,i,M}(t)|^2.$$

The desired result then follows. □

In the next corollary, we are going to use different stepsize to define the numerical solutions.

Corollary 5.3.6. *Assume that the conditions of Theorem 5.3.5 hold. Let* $M\ge 2, l\ge 1$, $h_l=T\cdot M^{-l}, h_{l-1}=T\cdot M^{-(l-1)}$. *Then*

$$\max_{0\le n<M^{l-1}}\mathrm{Var}(\Psi(\bar{Y}_{h_l}^{\varepsilon,i,M}(t_n))-\Psi(\bar{Y}_{h_{l-1}}^{\varepsilon,i,M}(t_n))\le Ch_{l-1}^2+C\varepsilon^2h_{l-1}.$$

Proof. For $0 \le n \le M^{l-1}-1$, by Theorem 5.3.5,

$$\begin{aligned}
&\mathrm{Var}(\Psi(\bar{Y}^{\varepsilon,i,M}_{h_l}(t_n)) - \Psi(\bar{Y}^{\varepsilon,i,M}_{h_{l-1}}(t_n)) \le 2\mathbb{E}|\Psi(\bar{Y}^{\varepsilon,i,M}_{h_l}(t_n)) - \Psi(\bar{Y}^{\varepsilon,i,M}_{h_{l-1}}(t_n))|^2\\
\le& 4\mathbb{E}|\Psi(\bar{Y}^{\varepsilon,i,M}_{h_l}(t_n)) - \Psi(X^{\varepsilon,i,M}(t))|^2 + 2\mathbb{E}|\Psi(X^{\varepsilon,i,M}(t)) - \Psi(\bar{Y}^{\varepsilon,i,M}_{h_{l-1}}(t_n))|^2\\
\le& Ch^2_{l-1} + C\varepsilon^2 h_{l-1}.
\end{aligned}$$

□

The following lemma is presented here because it applies to any EM scheme but it will only be use later when estimating the variance of coupled processes in the Multilevel Monte Carlo setting.

Define $\eta_h(s) := \lfloor s/h \rfloor$ where $\lfloor \cdot \rfloor$ is the integer-part function. Let z_h be the deterministic solution to

$$z_h(t) = X(0) + \int_0^t f(z_h(\eta_h(s)), \delta_{z_h(s)})ds, \tag{5.3.9}$$

which is the Euler approximation to the ODE obtained from (5.3.1) when ε is set to zero.

Lemma 5.3.7. *For any $T > 0$ we have*

$$\mathbb{E}[\sup_{0\le s\le T} |\bar{Y}^{\varepsilon,i,M}_h(s) - z_h(s)|^2] \le C\varepsilon^2. \tag{5.3.10}$$

Proof. Using (5.3.7) and (5.3.10), using the fact that $|a+b|^2 \le a^2 + b^2$ and the Cauchy-Schwarz inequality we have that for every $t \le T$

$$\begin{aligned}
&|\bar{Y}^{\varepsilon,i,M}_h(t) - z_h(t)|^2\\
&= \left| \int_0^t (f(Y^{\varepsilon,i,M}_h(s), \mathcal{L}^{\varepsilon,Y,M}_{h,s}) - f(z_h(\eta_h(s)), \delta_{z_h(s)})))ds + \varepsilon \int_0^t g(Y^{\varepsilon,i,M}_h(s), \mathcal{L}^{\varepsilon,Y,M}_{h,s})\mathrm{d}W^i(s) \right|^2\\
&\le 2T \int_0^t |f(Y^{\varepsilon,i,M}_h(s), \mathcal{L}^{\varepsilon,Y,M}_{h,s}) - f(z_h(\eta_h(s)), \delta_{z_h(s)}))|^2 ds\\
&+ 2\varepsilon^2 \left| \int_0^t g(Y^{\varepsilon,i,M}_h(s), \mathcal{L}^{\varepsilon,Y,M}_{h,s})\mathrm{d}W^i(s) \right|^2.
\end{aligned}$$

By the BDG inequality we have that

$$\mathbb{E}\left[\sup_{0\le s\le t} \left| \int_0^t g(Y^{\varepsilon,i,M}_h(s), \mathcal{L}^{\varepsilon,Y,M}_{h,s})\mathrm{d}W^i(s) \right|^2 \right] \le 4 \int_0^t \mathbb{E}[|g(Y^{\varepsilon,i,M}_h(s), \mathcal{L}^{\varepsilon,Y,M}_{h,s})|^2]ds.$$

Thus by Assumption 5.3.1 one can see that

$$\mathbb{E}[\sup_{0\le s\le t}|\bar{Y}_h^{\varepsilon,i,M}(t)-z_h(t)|^2]\le 2TK\int_0^t(\mathbb{E}[\sup_{0\le s\le r}|\bar{Y}_h^{\varepsilon,i,M}(s)-z_h(s)|^2]$$
$$+\sup_{0\le s\le r}\mathbb{W}_2^2(\mathcal{L}_{h,s}^{\varepsilon,Y,M},\delta_{z_h(s)}))dr+8T\varepsilon^2\beta\int_0^t\mathbb{E}[(1+|\bar{Y}_h^{\varepsilon,i,M}(s)|^2+W_2^2(\mathcal{L}_{h,s}^{\varepsilon,Y,M})]ds.$$

Using (2.2.1), (2.2.2) and Lemma 5.3.3 we have that for all $0\le t\le T$

$$\mathbb{E}[\sup_{0\le s\le t}|\bar{Y}_h^{\varepsilon,i,M}(t)-z_h(t)|^2]\le C\varepsilon^2+C\int_0^t\mathbb{E}[\sup_{0\le s\le r}|\bar{Y}_h^{\varepsilon,i,M}(s)-z_h(s)|^2]dr.$$

The final result is obtained by applying the Gronwall inequality. □

5.4 The Multilevel Monte Carlo-EM Scheme

We now define the multilevel Monte Carlo EM scheme. Given any $T>0$, let $N\ge 2, l\in\{0,...,L\}$, where L is a positive integer that will be determined later. Let $h_l=T\cdot N^{-l}, h_{l-1}=T\cdot N^{-(l-1)}$.

For step sizes h_l and h_{l-1} the EM continuous approximate solutions are respectively

$$\bar{Y}_{h_l}^{\varepsilon,i,M}(t)=x_0^i+\int_0^t f(Y_{h_l}^{\varepsilon,i,M}(s),\mathcal{L}_{h_l,s}^{\varepsilon,Y,M})\mathrm{d}s+\int_0^t g(Y_{h_l}^{\varepsilon,i,M}(s),\mathcal{L}_{h_l,s}^{\varepsilon,Y,M})\mathrm{d}W^i(s), \quad (5.4.1)$$

and

$$\bar{Y}_{h_{l-1}}^{\varepsilon,i,M}(t)=x_0^i+\int_0^t f(Y_{h_{l-1}}^{\varepsilon,i,M}(s),\mathcal{L}_{h_{l-1},s}^{\varepsilon,Y,M})\mathrm{d}s+\int_0^t g(Y_{h_{l-1}}^{\varepsilon,i,M}(s),\mathcal{L}_{h_{l-1},s}^{\varepsilon,Y,M})\mathrm{d}W^i(s). \quad (5.4.2)$$

We now construct the discrete version of the previous approximate solutions using the same Brownian motion for both processes. We say that the two processes are coupled. For $n\in\{0,1,\dots,N^{l-1}-1\}$ and $k\in\{0,\dots,N\}$, let

$$t_n=nh_{l-1}\text{ and }t_n^k=nh_{l-1}+kh_l.$$

This means we divide the interval $[t_n,t_{n+1}]$ into N equal parts by h_l with $t_n^0=t_n, t_n^N=t_{n+1}$. For $n\in\{0,1,\dots,N^{l-1}-1\}$ and $k\in\{0,\dots,N-1\}$, let

$$Y_{h_l}^{\varepsilon,i,M}(t_n^{k+1})=Y_{h_l}^{\varepsilon,i,M}(t_n^k)+f(Y_{h_l}^{\varepsilon,i,M}(t_n^k),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})h_l+\varepsilon\sqrt{h_l}g(Y_{h_l}^{\varepsilon,i,M}(t_n^k),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})\Delta\xi_n^k, \quad (5.4.3)$$

where $\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M} = \frac{1}{M}\sum_{j=1}^{M}\delta_{Y_{h_l}^{\varepsilon,j,M}(t_n^k)}$, the random vector $\Delta\xi_n^k \in \mathbb{R}^{\bar{d}}$ has independent components, and each component is distributed as $\mathcal{N}(0,1)$. Therefore, to simulate $Y_{h_l}^{\varepsilon,i,M}$, we use

$$\begin{aligned} Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) = Y_{h_l}^{\varepsilon,i,M}(t_n) + \sum_{k=0}^{N-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})h_l \\ + \varepsilon\sqrt{h_l}\sum_{k=0}^{N-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})\Delta\xi_n^k. \end{aligned} \tag{5.4.4}$$

To simulate $Y_{h_{l-1}}^{\varepsilon,i,M}$, we use

$$\begin{aligned} Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1}) = Y_{h_{l-1}}^{\varepsilon,i,M}(t_n) + f(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})h_{l-1} \\ + \varepsilon\sqrt{h_l}g(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\sum_{k=0}^{N-1}\Delta\xi_n^k, \end{aligned} \tag{5.4.5}$$

where $\mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M} = \frac{1}{M}\sum_{j=1}^{M}\delta_{Y_{h_{l-1}}^{\varepsilon,j,M}(t_n)}$.

The following theorem is the main result of this section.

Theorem 5.4.1. *Let Assumption 5.3.1 hold. Then it holds that*

$$\max_{0\le n<M^{l-1}} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] \le CN^2h_l^2 + \bar{C}\varepsilon^4 Nh_l.$$

In order to prove Theorem 5.4.1, we need a few lemmas.

Lemma 5.4.2. *Let $p \ge 2$. Then*

$$\max_{\substack{0\le n\le N^{l-1} \\ 1\le k\le N}} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^p] \le C_1N^p h_l^p + C_2N^{p/2}h_l^{p/2}\varepsilon^p,$$

where C and C are positive constants that only depend on β, T, m and $X^\varepsilon(0)$ (β from Remark 5.3.1).

Proof. From (5.4.3) we have that

$$Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n) = \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})h_l + \varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j. \tag{5.4.6}$$

Hence, we obtain

$$\mathbb{E}[|Y^{\varepsilon,i,M}_{h_l}(t^k_n) - Y^{\varepsilon,i,M}_{h_l}(t_n)|^p] \tag{5.4.7}$$
$$\leq 2^{p-1}\mathbb{E}\left|\sum_{j=0}^{k-1} f(Y^{\varepsilon,i,M}_{h_l}(t^j_n), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})h_l\right|^p + 2^{p-1}\mathbb{E}\left|\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y^{\varepsilon,i,M}_{h_l}(t^j_n), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})\Delta\xi^j_n\right|^p.$$

By Remark 5.3.1 and Lemma 5.3.3 one can see that

$$\mathbb{E}\left|\sum_{j=0}^{k-1} f(Y^{\varepsilon,i,M}_{h_l}(t^j_n), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})h_l\right|^p \leq N^{p-1}\sum_{j=0}^{k-1}\mathbb{E}\left|f(Y^{\varepsilon,i,M}_{h_l}(t^j_n), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})h_l\right|^p$$
$$\leq N^{p-1}\sum_{j=0}^{k-1}\mathbb{E}\left[\left(\beta\left(1+|Y^{\varepsilon,i,M}_{h_l}(t^j_n)|^2 + W_2^2(\mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})\right)\right)^{p/2}\right]$$
$$\leq CN^{p-1}h_l^p\sum_{j=0}^{k-1}\left(1+2\mathbb{E}[|Y^{\varepsilon,i,M}_{h_l}(t^j_n)|^2]\right)^{p/2} \leq CN^p h_l^p. \tag{5.4.8}$$

Using the BDG inequality, Remark 5.3.1 and Lemma 5.3.3, we obtain

$$\mathbb{E}\left|\varepsilon\sum_{j=0}^{k-1} g(Y^{\varepsilon,i,M}_{h_l}(t^j_n), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})\sqrt{h_l}\Delta\xi^j_n\right|^p \leq C\varepsilon^p\mathbb{E}\left[\left|\sum_{j=0}^{k-1}|g(Y^{\varepsilon,i,M}_{h_l}(t^j_n), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})|^2 h_l\right|^{p/2}\right]$$
$$\leq C\varepsilon^p N^{p/2-1}h_l^{p/2}\mathbb{E}\left[\sum_{j=0}^{k-1}(|g(Y^{\varepsilon,i,M}_{h_l}(t^j_n), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})|^2)^{p/2}\right]$$
$$\leq C\varepsilon^p N^{p/2-1}h_l^{p/2}\sum_{j=0}^{k-1}\mathbb{E}\left[\left(\beta\left(1+|Y^{\varepsilon,i,M}_{h_l}(t^j_n)|^2 + W_2^2(\mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})\right)\right)^{p/2}\right] \leq CN^{p/2}h_l^{p/2}\varepsilon^p \tag{5.4.9}$$

The result follows from substituting (5.4.8) and (5.4.9) into (5.4.7). □

Lemma 5.4.3. *Let f_m be the m^{th} component of f. Then there exist $s, r \in [0,1]$ such that*

$$f(Y^{\varepsilon,i,M}_{h_l}(t^k_n), \mathcal{L}^{\varepsilon,Y^k_n,M}_{h_l}) - f(Y^{\varepsilon,i,M}_{h_l}(t_n), \mathcal{L}^{\varepsilon,Y^k_n,M}_{h_l}) = A_k + B_k + E_k,$$

where

$$A_k = (A^1_k, ..., A^d_k)', B_k = (B^1_k, ..., B^d_k)', E_k = (E^1_k, ..., E^d_k)'$$
$$A^m_k := \langle\nabla f_m(sY^{\varepsilon,i,M}_{h_l}(t^k_n) + (1-s)Y^{\varepsilon,i,M}_{h_l}(t_n)), \mathcal{L}^{\varepsilon,Y^k_n,M}_{h_l}), h_l\sum_{j=0}^{k-1} f(Y^{\varepsilon,i,M}_{h_l}(t^j_n), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})\rangle,$$

$$B_k^m := \langle \nabla f_m(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}), \varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\rangle,$$

$$E_k^m := \langle \nabla^2 f_m(rs(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)) + Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))s,$$
$$\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\rangle$$
$$+ \langle \mathbb{E}[\langle \partial_\mu \nabla f_m(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^w,M})(Y_w^s), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)},$$
$$\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\rangle, m \in \{1, ..., d\}.$$

Proof. By the mean value theorem (Lemma C.0.1) with $y = Y_{h_l}^{\varepsilon,i,M}(t_n^k), x = Y_{h_l}^{\varepsilon,i,M}(t_n)$ and $g(z) = f_m(z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$, there exists a $s \in [0,1]$ such that

$$f_m(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_m(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$$
$$= \langle \nabla f_m(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle.$$

Substituting (5.4.6) in the equation above yields

$$f_m(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_m(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \tag{5.4.10}$$
$$= \langle \nabla f_m(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}), \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})h_l\rangle$$
$$+ \langle \nabla f_m(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}), \varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\rangle.$$

Let $\nabla_q f_m$ denote the q^{th} component of the vector function ∇f_m. Applying the mean value theorem again with $y = sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n), x = Y_{h_l}^{\varepsilon,i,M}(t_n)$ and $g(z) = \nabla_q f_m(z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ ensures that there exists a $r \in [0,1]$ such that

$$\nabla_q f_m(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) = \nabla_q f_m(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$$
$$+ \langle \nabla(\nabla_q f_m)(rs(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)) + Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))s\rangle.$$

An application of Equation (2.2.8) to $u(\mathcal{L}^\xi) = \nabla_q f_m(Y_{h_l}^{\varepsilon,i,M}, \mathcal{L}_{h_l}^{\varepsilon,\xi,M})$
with $X = Y_{h_l}^{\varepsilon,i,M}(t_n^k), X' = Y_{h_l}^{\varepsilon,i,M}(t_n)$ implies that there exists a random variable w :

$\Omega \to [0,1]$ such that

$$\nabla_q f_m(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) = \nabla_q f_m(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \quad (5.4.11)$$
$$+ \mathbb{E}[\langle \partial_\mu \nabla_q f_m(Z,\mathcal{L}_{h_l}^{\varepsilon,Y_n^w,M})(Y_w^s),(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)},$$

where $Y_n^w := wY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-w)Y_{h_l}^{\varepsilon,i,M}(t_n)$. Thus

$$\nabla f_m(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) = \nabla f_m(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})$$
$$+ \mathbb{E}[\langle \partial_\mu \nabla f_m(Z,\mathcal{L}_{h_l}^{\varepsilon,Y_n^w,M})(Y_w^s),(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}$$
$$+ \nabla^2 f_m(rs(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)) + Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))s.$$

Substituting the last equation into the second summand of the RHS of (5.4.10) completes the proof. □

Lemma 5.4.4. *There exist random variables $s, r : \Omega \to [0,1]$ such that*

$$f(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) = \bar{A}_k + \bar{E}_k,$$

where

$$\bar{A}_k = (\bar{A}_k^1,...,\bar{A}_k^d)', \bar{E}_k = (\bar{E}_k^1,...,\bar{E}_k^d)'$$
$$\bar{A}_k^m := \mathbb{E}[\langle \partial_\mu f_m(Z,\mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s), h_l \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}$$
$$\bar{E}_k^m := \mathbb{E}[\langle \partial_\mu^2 f_m(Z,\mathcal{L}_{h_l}^{\varepsilon,Y_n^{s,r},M})(Y_n^{s,r})(((Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))s,$$
$$\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)},$$
$$Y_n^s := sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n),$$
$$Y_n^{s,r} := sr(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)) + Y_{h_l}^{\varepsilon,i,M}(t_n).$$

Proof. Let f_m be the m^{th} component of f. A direct application of Equation (2.2.8) with $X = Y_{h_l}^{\varepsilon,i,M}(t_n^k), X' = Y_{h_l}^{\varepsilon,i,M}(t_n)$ and $\bar{u}(\mathcal{L}(\xi)) = f_m(Y_{h_l}^{\varepsilon,i,M},\mathcal{L}_{h_l}^{\varepsilon,\xi,M})$ implies that there exists a random variable $s : \Omega \to [0,1]$ such that

$$f_m(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_m(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \quad (5.4.12)$$

$$= \mathbb{E}[\langle \partial_\mu f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n^s, M})(Y_n^s), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}$$

$$= \mathbb{E}[\langle \partial_\mu f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n^s, M})(Y_n^s), \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon, Y_n^j, M}) h_l\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}$$

$$+ \mathbb{E}[\langle \partial_\mu f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n^s, M})(Y_n^s), \varepsilon\sqrt{h_l} \sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon, Y_n^j, M}) \Delta\xi_n^j\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}.$$

Let $\partial_{\mu,q} f_m$ be the q^{th} component of the vector function $\partial_\mu f_m$. Applying Equation (2.2.8) again with $X = sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n) =: Y_n^s, X' = Y_{h_l}^{\varepsilon,i,M}(t_n)$ and $\bar{u}(\mathcal{L}(\xi)) = \partial_{\mu,q} f_m(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,\xi,M})(\xi)$, we find that there exists a random variable $r : \Omega \to [0,1]$ such that

$$\partial_{\mu,q} f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n^s, M})(Y_n^s) = \partial_{\mu,q} f_m(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon, Y_n, M})(Y_{h_l}^{\varepsilon,i,M}(t_n))$$
$$+ \mathbb{E}[\langle \partial_\mu(\partial_{\mu,q} f_m)(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n^{s,r}, M})(Y_n^{s,r}), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))s\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}.$$

Thus

$$\partial_\mu f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n^s, M})(Y_n^s) = \partial_\mu f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n, M})(Y_{h_l}^{\varepsilon,i,M}(t_n))$$
$$+ \mathbb{E}[\partial_\mu^2 f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n^{s,r}, M})(Y_n^{s,r}), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))s]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}.$$

Substituting the last equation into the second summand of the RHS of Equation (5.4.12) yields

$$f_m(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon, Y_n^k, M}) - f_m(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon, Y_n, M})$$
$$= \mathbb{E}[\langle \partial_\mu f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n^s, M})(Y_n^s), \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon, Y_n^j, M}) h_l\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}$$
$$+ \mathbb{E}[\langle \partial_\mu f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n, M})(Y_{h_l}^{\varepsilon,i,M}(t_n)), \varepsilon\sqrt{h_l} \sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon, Y_n^j, M}) \Delta\xi_n^j\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}$$
$$+ \mathbb{E}[\langle \partial_\mu^2 f_m(Z, \mathcal{L}_{h_l}^{\varepsilon, Y_n^{s,r}, M})(Y_n^{s,r})(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))s,$$
$$\varepsilon\sqrt{h_l} \sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon, Y_n^j, M}) \Delta\xi_n^j\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}.$$

By independence the second expectation above is zero, therefore the proof is complete.

□

Proof of Theorem 5.4.1 From (5.4.4) and (5.4.5) we have that for $n \le N^{l-1}-1$

$$
\begin{aligned}
Y^{\varepsilon,i,M}_{h_l}(t_{n+1}) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_{n+1}) &= Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n) \\
&+ h_l \sum_{k=0}^{N-1} \Big(f(Y^{\varepsilon,i,M}_{h_l}(t^k_n), \mathcal{L}^{\varepsilon,Y^k_n,M}_{h_l}) - f(Y^{\varepsilon,i,M}_{h_l}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \Big) \\
&+ h_l \sum_{k=0}^{N-1} \Big(f(Y^{\varepsilon,i,M}_{h_l}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) - f(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_{l-1}}) \Big) \\
&+ \varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \Big(g(Y^{\varepsilon,i,M}_{h_l}(t^k_n), \mathcal{L}^{\varepsilon,Y^k_n,M}_{h_l}) - g(Y^{\varepsilon,i,M}_{h_l}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \Big) \Delta\xi^k_n \\
&+ \varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \Big(g(Y^{\varepsilon,i,M}_{h_l}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) - g(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_{l-1}}) \Big) \Delta\xi^k_n \\
&=: Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n) + R_N.
\end{aligned}
$$

By using the linearity property of the inner product, we obtain

$$
\begin{aligned}
&|Y^{\varepsilon,i,M}_{h_l}(t_{n+1}) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_{n+1})|^2 \\
&= \langle Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n) + R_N, Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n) + R_N \rangle \\
&= |Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)|^2 + |R_N|^2 + 2\langle Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), R_N \rangle.
\end{aligned}
$$

Applying the elementary inequality $|a+b+c+d|^2 \le 4|a|^2+4|b|^2+4|c|^2+4|d|^2$ to the term $|R_N|^2$ above, we derive that

$$
\begin{aligned}
&|Y^{\varepsilon,i,M}_{h_l}(t_{n+1}) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_{n+1})|^2 \le |Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)|^2 \\
&+ 4h_l^2 \left| \sum_{k=0}^{N-1} \Big(f(Y^{\varepsilon,i,M}_{h_l}(t^k_n), \mathcal{L}^{\varepsilon,Y^k_n,M}_{h_l}) - f(Y^{\varepsilon,i,M}_{h_l}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \Big) \right|^2 \\
&+ 4h_l^2 \left| \sum_{k=0}^{N-1} \Big(f(Y^{\varepsilon,i,M}_{h_l}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) - f(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_{l-1}}) \Big) \right|^2 \\
&+ 4\varepsilon^2 \left| \sum_{k=0}^{N-1} \Big(g(Y^{\varepsilon,i,M}_{h_l}(t^k_n), \mathcal{L}^{\varepsilon,Y^k_n,M}_{h_l}) - g(Y^{\varepsilon,i,M}_{h_l}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \Big) \sqrt{h_l}\Delta\xi^k_n \right|^2 \\
&+ 4\varepsilon^2 \left| \sum_{k=0}^{N-1} \Big(g(Y^{\varepsilon,i,M}_{h_l}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) - g(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), \mathcal{L}^{\varepsilon,Y_n,M}_{h_{l-1}}) \Big) \sqrt{h_l}\Delta\xi^k_n \right|^2
\end{aligned}
$$

$$+ 2h_l \sum_{k=0}^{N-1} \langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\rangle$$

$$+ 2h_l \sum_{k=0}^{N-1} \langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), f(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\rangle$$

$$+ 2\varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \big(g(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - g(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\big)\Delta\xi_n^k\rangle$$

$$+ 2\varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \big(g(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - g(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\big)\Delta\xi_n^k\rangle.$$

Now, we take expectations on both sides of the previous inequality. Since $\Delta\xi_n^k$ is independent of $Y_{h_l}^{\varepsilon,i,M}(t_n^k)$ and $Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)$, the expectation of the last two summands in the equation above is zero. Thus,

$$\begin{aligned}
&\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})|^2] \le \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] \qquad (5.4.13)\\
&\quad + 4Nh_l^2 \sum_{k=0}^{N-1} \mathbb{E}\left|f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\right|^2\\
&\quad + 4Nh_l^2 \sum_{k=0}^{N-1} \mathbb{E}\left|f(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\right|^2\\
&\quad + 4\varepsilon^2\mathbb{E}\left[\left|\sum_{k=0}^{N-1}\Big(g(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - g(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\Big)\sqrt{h_l}\Delta\xi_n^k\right|^2\right]\\
&\quad + 4\varepsilon^2\mathbb{E}\left[\left|\sum_{k=0}^{N-1}\Big(g(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - g(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\Big)\sqrt{h_l}\Delta\xi_n^k\right|^2\right]\\
&\quad + 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\rangle]\\
&\quad + 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), f(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\rangle].\\
&\quad =: \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] + I_1 + I_2 + I_3 + I_4 + I_5 + I_6. \qquad (5.4.14)
\end{aligned}$$

By Assumption 5.3.1 and Lemma 5.4.2, one can see that

$$I_1 \le 4KNh_l^2 \sum_{k=0}^{N-1} (\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^2 + \mathbb{W}_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}, \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}))$$

$$\leq 8KNh_l^2\sum_{k=0}^{N-1}\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n)|^2\leq 8KN^2h_l^2(CN^2h_l^2+CN\varepsilon^2h_l).$$

Also, by Assumption 5.3.1

$$\begin{aligned}I_2&\leq 4KNh_l^2\sum_{k=0}^{N-1}(\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2+\mathbb{W}_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n,M},\mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M}))\\&\leq 8KN^2h_l^2\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2].\end{aligned}$$

By the BDG inequality, Assumption 5.3.1 and Lemma 5.4.2, we obtain

$$\begin{aligned}I_3&\leq C\varepsilon^2\sum_{k=0}^{N-1}\mathbb{E}[|g(Y_{h_l}^{\varepsilon,i,M}(t_n^k),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})-g(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})|^2]h_l\\&=Ch_l\varepsilon^2\sum_{k=0}^{N-1}(\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n)|^2]+\mathbb{W}_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M},\mathcal{L}_{h_l}^{\varepsilon,Y_n,M}))\\&\leq CN^3h_l^3\varepsilon^2+CN^2h_l^2\varepsilon^4.\end{aligned}$$

Similarly to I_3,

$$I_4\leq CNh_l\varepsilon^2\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2].$$

For I_5 note that

$$\begin{aligned}I_5&=2h_l\sum_{k=0}^{N-1}\mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),f(Y_{h_l}^{\varepsilon,i,M}(t_n^k),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})-f(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\rangle]\\&=2h_l\sum_{k=0}^{N-1}\mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),f(Y_{h_l}^{\varepsilon,i,M}(t_n^k),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})-f(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})\rangle]\\&+2h_l\sum_{k=0}^{N-1}\mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),f(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})-f(Y_{h_l}^{\varepsilon,i,M}(t_n),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\rangle]\\&=:I_{5A}+I_{5B}.\end{aligned}$$

Applying Lemma 5.4.3 we have

$$\begin{aligned}I_{5A}&\leq 2h_l\sum_{k=0}^{N-1}\mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),A_k\rangle]+2h_l\sum_{k=0}^{N-1}\mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),B_k\rangle]\\&+2h_l\sum_{k=0}^{N-1}\mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),E_k\rangle].\end{aligned}$$

By independence, the second summand above is zero. We note that

$$\begin{aligned}\mathbb{E}[|A_k|^2] &= \sum_{m=1}^{\bar{d}} \mathbb{E}[|A_k^m|]^2 \leq \bar{d}\bar{K}\mathbb{E}\left| h_l \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\right|^2 \\ &\leq \bar{d}\bar{K}h_j^2 N \sum_{j=0}^{k-1} \mathbb{E}\left[\left(\beta\left(1+|Y_{h_l}^{\varepsilon,i,M}(t_n^j)|^2 + W_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\right)\right)^2\right] \\ &\leq \bar{K}h_l^2 N^2 C\end{aligned}$$

Also, using the Cauchy-Schwarz and the BDG inequalities, Assumption 5.3.1 and Lemma 5.4.2, we obtain

$$\mathbb{E}[|E_k|^2] = \sum_{m=1}^{\bar{d}} \mathbb{E}[(E_k^m)^2]$$

$$= \sum_{m=1}^{\bar{d}} \mathbb{E}\Bigg[\Big\{\big\langle \nabla^2 f_m(rs(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)) + Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \tag{5.4.15}$$

$$\times (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))s, \varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\big\rangle$$

$$+ \big\langle \mathbb{E}[\langle \partial_\mu \nabla f_m(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^w,M})(Y_w^s), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)},$$

$$\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\big\rangle\Big\}^2\Bigg]$$

$$\leq 4\bar{d}\bar{K}\varepsilon^2\mathbb{E}\left[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^2\left|\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\sqrt{h_l}\Delta\xi_n^j\right|^2\right]$$

$$\leq 4\bar{d}\bar{K}\varepsilon^2(\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^4])^{1/2}\left(\mathbb{E}\left[\left|\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\sqrt{h_l}\Delta\xi_n^j\right|^4\right]\right)^{1/2}$$

$$\leq \bar{K}\varepsilon^2 CN^3h_l^3 + \bar{K}\varepsilon^4 CN^2h_l^2. \tag{5.4.16}$$

Therefore, applying the Cauchy-Schwarz inequality first and the elementary inequality $2ab \leq a^2 + b^2$ later yields

$$I_{5A} \leq 2h_l\sum_{k=0}^{N-1}\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)||A_k|] + 2h_l\sum_{k=0}^{N-1}\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)||E_k|]$$

$$\leq 2h_l \sum_{k=0}^{N-1} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + h_l \sum_{k=0}^{N-1} \mathbb{E}[|A_k|^2] + h_l \sum_{k=0}^{N-1} \mathbb{E}[|E_k|^2]$$
$$\leq 2h_l N\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + \bar{K}h_l^3 N^3 C + \bar{K}CN^4 h_l^4 \varepsilon^2 + \bar{K}CN^3 h_l^3 \varepsilon^4$$

Similarly, using Lemma 5.4.4 one can see that

$$I_{5B} \leq 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \bar{A}_k\rangle] + 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \bar{E}_k\rangle]$$

Also, we have $\mathbb{E}[|\bar{A}_k|^2] \leq h_l^2 N^2 C$ and

$$\mathbb{E}[|\bar{E}_k|^2] \leq \bar{K}\varepsilon^2 CN^3 h_l^3 + \bar{K}\varepsilon^4 CN^2 h_l^2. \tag{5.4.17}$$

Thus,

$$I_{5B} \leq 2h_l \sum_{k=0}^{N-1} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|\mathbb{E}[|\bar{A}_k|]]$$
$$+ 2h_l \sum_{k=0}^{N-1} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|\mathbb{E}[|\bar{E}_k|]]$$
$$\leq 2h_l \sum_{k=0}^{N-1} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + h_l \sum_{k=0}^{N-1} \mathbb{E}[|\bar{A}_k|^2] + h_l \sum_{k=0}^{N-1} \mathbb{E}[|\bar{E}_k|^2]$$
$$\leq 2h_l N\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + \bar{K}h_l^3 N^3 C + \bar{K}CN^4 h_l^4 \varepsilon^2 + \bar{K}CN^3 h_l^3 \varepsilon^4$$

Additionally, we have

$$I_6 \leq h_l N\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + h_l N\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2$$
$$+ h_l N\mathbb{W}_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}, \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n^k,M})$$
$$\leq 3h_l N\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2.$$

Substituting the bounds for the terms I_1 to I_6 into Equation (5.4.13) yields that for $n \leq N^{l-1} - 1$

$$\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})|^2] \leq \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2]$$
$$+ \hat{C}\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] + CN^3 h_l^3 + CN^2 h_l^2 \varepsilon^4,$$

which implies that that for $n \leq N^{l-1} - 1$

$$\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})|^2] \leq \hat{C} \sum_{k=1}^{n} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_k) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_k)|^2] + CN^2 h_l^2 + CNh_l\varepsilon^4.$$

An application of the discrete Gronwall inequality yields the result.

□

5.5 Variance estimate of two coupled paths of the MLMC-EM scheme

In this section we provide an estimate for the variance of two coupled paths which is the main result of the paper and will be presented in Theorem 5.5.4. We will need the following lemma taken from [1]. Proof of this theorem can be found in [2].

Lemma 5.5.1. *Suppose that $A^{\varepsilon,h}$ and $B^{\varepsilon,h}$ are families of random variables determined by scaling parameters ε and h. Further, suppose that there are $C_1 > 0, C_2 > 0$ and $C_3 > 0$ such that for all $\varepsilon \in (0,1)$ the following three conditions hold:*

(i) $\mathrm{Var}(A^{\varepsilon,h}) \leq C_1\varepsilon^2$ *uniformly in h,*

(ii) $|A^{\varepsilon,h}| \leq C_2$ *uniformly in h,*

(iii) $|\mathbb{E}[B^{\varepsilon,h}]| \leq C_3 h$.

Then

$$\mathrm{Var}(A^{\varepsilon,h}B^{\varepsilon,h}) \leq 3C_3^2 C_1 h^2 \varepsilon^2 + 15C_2^2 \mathrm{Var}(B^{\varepsilon,h}).$$

The following two lemmas that will be needed to prove Theorem 5.5.4.

Lemma 5.5.2. *Assume that $\gamma : \mathbb{R}^d \to \mathbb{R}$ satisfies the Lipschitz condition, i.e. for all $x, y \in \mathbb{R}^d$ there exists a positive constant L, such that $|\gamma(x) - \gamma(y)|^2 \leq L|x - y|^2$. Then* *for $s \in [0,1]$ one has*

$$\max_{\substack{0\le n\le N^{l-1}\\ 1\le k\le N}} \operatorname{Var}(\gamma(sY_{h_l}^{\varepsilon,i,M}(t_n^k)+(1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n)))) \le C\varepsilon^2; \tag{5.5.1}$$

$$\max_{0\le n\le N^{l-1}} \operatorname{Var}(\gamma(sY_{h_l}^{\varepsilon,i,M}(t_n)+(1-s)(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)))) \le C\varepsilon^2. \tag{5.5.2}$$

Proof. We will only prove (5.5.1), the prove for (5.5.2) is very similar. Let z_{h_l} and $z_{h_{l-1}}$ be defined by (5.3.9). Using the fact that for a random variable X and a constant a, $\operatorname{Var}(X+a)=\operatorname{Var}(X)$ and the fact that γ is Lipschitz, we have that

$$\begin{aligned}
&\max_{\substack{0\le n\le N^{l-1}\\ 1\le k\le N}} \operatorname{Var}(\gamma(sY_{h_l}^{\varepsilon,i,M}(t_n^k)+(1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n))))\\
&=\max_{\substack{0\le n\le N^{l-1}\\ 1\le k\le N}} \operatorname{Var}(\gamma(sY_{h_l}^{\varepsilon,i,M}(t_n^k)+(1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n)))-\gamma(sz_{h_l}(t_n^k)+(1-s)(z_{h_l}(t_n))))\\
&\le\max_{\substack{0\le n\le N^{l-1}\\ 1\le k\le N}} \mathbb{E}[|(\gamma(sY_{h_l}^{\varepsilon,i,M}(t_n^k)+(1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n)))-\gamma(sz_{h_l}(t_n^k)+(1-s)(z_{h_l}(t_n)))|^2]\\
&=\max_{\substack{0\le n\le N^{l-1}\\ 1\le k\le N}} L\mathbb{E}[|sY_{h_l}^{\varepsilon,i,M}(t_n^k)+(1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n))-sz_{h_l}(t_n^k)-(1-s)(z_{h_l}(t_n))|^2]\\
&\le\max_{\substack{0\le n\le N^{l-1}\\ 1\le k\le N}} sL\mathbb{E}[|(Y_{h_l}^{\varepsilon,i,M}(t_n^k)-z_{h_l}(t_n^k)|^2]+(1-s)L\mathbb{E}[|(Y_{h_l}^{\varepsilon,i,M}(t_n)-(z_{h_l}(t_n)|^2].
\end{aligned}$$

The required assertion follows by Lemma 5.3.7. □

Lemma 5.5.3. *Let Assumption 5.3.1 hold. Then there exists a positive constant C such that*

$$\max_{\substack{0\le n\le N^{l-1}\\ 1\le k\le N}} |\mathbb{E}[Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n)]| \le CNh_l.$$

Proof. From (5.4.3) we have that

$$\begin{aligned}
&|\mathbb{E}[Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n)]|\\
&\qquad=\left|\sum_{j=0}^{k-1}\mathbb{E}[f(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})]h_l+\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1}\mathbb{E}[g(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j]\right|.
\end{aligned}$$

By independence the second summand of RHS in above is zero. Thus using Jensen's inequality and Remark 5.3.1 yields

$$|\mathbb{E}[Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n)]| \le \sum_{j=0}^{k-1}\mathbb{E}[|f(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})|]h_l$$

$$\leq h_l \sum_{j=0}^{k-1} \mathbb{E}[\sqrt{\beta}\big(1 + |Y_{h_l}^{\varepsilon,i,M}(t_n^j)|^2 + W_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\big)^{1/2}]$$

$$\leq \sqrt{\beta} h_l \sum_{j=0}^{k-1} \Big(1 + 2\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^j)|^2]\Big)^{1/2}.$$

An application of Lemma 5.3.3 and the fact that $k \leq N$, completes the proof. □

Now, we can formulate the main result of the paper.

Theorem 5.5.4. *Let Assumption 5.3.1 hold, assume that $\Psi : \mathbb{R}^d \to \mathbb{R}$ has continuous second order derivative and there exists a constant C such that*

$$\left|\frac{\partial \Psi}{\partial x_i}\right| \leq C \quad and \quad \left|\frac{\partial^2 \Psi}{\partial x_i \partial x_j}\right| \leq C$$

for any $i, j = 1, 2, \cdots, a$. Then, we have

$$\max_{0 \leq n < M^{l-1}} \operatorname{Var}(\Psi(Y_{h_l}^{\varepsilon,i,M}(t_{n+1})) - \Psi(Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})) \leq C\varepsilon^2 h_{l-1}^2 + C\varepsilon^4 h_{l-1}.$$

Proof. From (5.4.4) and (5.4.5) we have that for $n \leq N^{l-1} - 1$

$$\begin{aligned}
[Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j &= [Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j \\
&+ h_l \sum_{k=0}^{N-1} \Big(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\Big) \\
&+ h_l \sum_{k=0}^{N-1} \Big(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\Big) \\
&+ \varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \Big(g_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\Big) \Delta\xi_n^k \\
&+ \varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \Big(g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - g_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\Big) \Delta\xi_n^k \\
&=: [Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j + h_l T_1 + h_l T_2 + \varepsilon\sqrt{h_l} T_3 + \varepsilon\sqrt{h_l} T_4,
\end{aligned}$$

where f_j is the jth component of f and g_j is the jth row of g. Taking variances on both sides of the previous inequality and using the fact that a finite sequence of random variables $X_i, i = 1, ..., n$, satisfies that $\operatorname{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \operatorname{Var}(X_i) + 2\sum_{i<j} \operatorname{Cov}(X_i, X_j)$,

we obtain

$$\begin{aligned}
&\mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j) = \mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j) + h_l^2\mathrm{Var}(T_1)\\
&+ h_l^2\mathrm{Var}(T_2) + \varepsilon^2 h_l\mathrm{Var}(T_3) + \varepsilon^2 h_l\mathrm{Var}(T_4) + 2\mathrm{Cov}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j, h_lT_1)\\
&+ 2\mathrm{Cov}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j, h_lT_2)\\
&+ 2\mathrm{Cov}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j, \varepsilon\sqrt{h_l}T_3)\\
&+ 2\mathrm{Cov}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j, \varepsilon\sqrt{h_l}T_4) + 2\mathrm{Cov}(h_lT_1, h_lT_2) + 2\mathrm{Cov}(h_lT_1, \varepsilon\sqrt{h_l}T_3)\\
&+ 2\mathrm{Cov}(h_lT_1, \varepsilon\sqrt{h_l}T_4) + 2\mathrm{Cov}(h_lT_2, \varepsilon\sqrt{h_l}T_3) + 2\mathrm{Cov}(h_lT_2, \varepsilon\sqrt{h_l}T_4)\\
&+ 2\mathrm{Cov}(\varepsilon\sqrt{h_l}T_3, \varepsilon\sqrt{h_l}T_4).
\end{aligned} \tag{5.5.3}$$

By the independence of $\Delta\xi_n^k$ with respect to $[Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j, g_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})$ and $g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - g_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})$, we have that

$$\begin{aligned}
&2\mathrm{Cov}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j, \varepsilon\sqrt{h_l}T_3)\\
&= 2\mathrm{Cov}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j, \varepsilon\sqrt{h_l}T_4) = 0.
\end{aligned} \tag{5.5.4}$$

Also note that

$$\begin{aligned}
&2\mathrm{Cov}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j, h_lT_2)\\
&= 2Nh_l\mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j, (f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\\
&\qquad - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M}))\Big).
\end{aligned} \tag{5.5.5}$$

Substituting (5.5.4) and (5.5.5) into (5.5.3) and using the fact that for two random variables X, Y, we have that $2\mathrm{Cov}(X,Y) \leq \mathrm{Var}(X) + \mathrm{Var}(Y)$, yields

$$\begin{aligned}
&\mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j) \leq (1 + Nh_l)\mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j)\\
&+ 4h_l^2N\sum_{k=0}^{N-1}\mathrm{Var}\left(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\right)\\
&+ (4Nh_l + 1)Nh_l\mathrm{Var}\left(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\right)\\
&+ 4\varepsilon^2h_l\sum_{k=0}^{N-1}\mathrm{Var}\left(g_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\right)\Delta\xi_n^k
\end{aligned}$$

$$
\begin{aligned}
&+ 4\varepsilon^2 h_l \sum_{k=0}^{N-1} \mathrm{Var}\left(g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - g_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\right) \Delta\xi_n^k \\
&+ 2\mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, \\
&\qquad\qquad h_l \sum_{k=0}^{N-1} f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\Big) \\
&=: I_1 + I_2 + I_3 + I_4 + I_5 + I_6.
\end{aligned}
$$

In order to complete the proof of the theorem, we give estimates for $I_i, i = 2, ..., 6$, which will be shown in the following lemmas.

Lemma 5.5.5. *There exists a positive constant C such that*

$$I_2 \le CN^3 h_l^3 \varepsilon^2.$$

Proof. Using (B.0.2), we have that

$$
\begin{aligned}
&\mathrm{Var}(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})) \\
&\quad \le 2\mathrm{Var}(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})) \\
&\quad + 2\mathrm{Var}(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})) =: I_{2A} + I_{2B}.
\end{aligned}
$$

First we estimate I_{2A}. By the mean value theorem there exists an $s \in [0,1]$ such that

$$
\begin{aligned}
&f_m(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
&\qquad = \langle \nabla f_j(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle.
\end{aligned}
$$

Let $\nabla_q f_j(sY_{h_l}^{\varepsilon,i,M}(t_n^k)+(1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $[(Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n))]_q$ be the q components of $\nabla f_j(sY_{h_l}^{\varepsilon,i,M}(t_n^k)+(1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $(Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n))$ respectively. We want to apply Lemma 5.5.1 with $A^{\varepsilon,h} = \nabla_q f_j(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1 - s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $B^{\varepsilon,h} = [(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))]_q$ so we check that the three conditions are satisfied. By Assumption 5.3.1, the function $\nabla_q^2 f_j$ is bounded, so $\nabla_q f_j$ is Lipschitz on the first argument. Applying Lemma 5.5.2 with $\gamma = \nabla_q f_j(\cdot, \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $h_{l_1} = h_{l_2} = h_l$, we obtain

$$\mathrm{Var}(\nabla_q f_j(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})) \le C_1\varepsilon^2, \tag{5.5.6}$$

so the first condition of Lemma 5.5.1 is satisfied. Conditions 2 and 3 are satisfied by Assumption 5.3.1 and Lemma 5.5.3 respectively. Thus by Lemma 5.5.1 we have that

$$\begin{aligned}\mathrm{Var}(\nabla_q f_j(sY_{h_l}^{\varepsilon,i,M}(t_n^k)+(1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})[(Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n))]_q)\\ \leq 3C_3^2C_1N^2h_l^2\varepsilon^2+15C_2^2\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n))]_q).\end{aligned}$$

In order to estimate $\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n))]_q)$ we use Equation (5.4.3) to obtain

$$\begin{aligned}&\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n))]_q)\\ &\quad\leq 2\mathrm{Var}(\sum_{j=0}^{k-1}f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})h_l)+2\mathrm{Var}(\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1}g_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j).\end{aligned}$$

By Asumption 5.3.1 and Lemma 5.3.7 we have that

$$\begin{aligned}\mathrm{Var}(\sum_{j=0}^{k-1}f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})h_l)&=\mathrm{Var}(h_l\sum_{j=0}^{k-1}f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})-f_q(z_h(t_n^j),\delta_{z_h(t_n^j)}))\\ &\leq h_l^2\mathbb{E}[|(\sum_{j=0}^{k-1}f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})-f_q(z_h(t_n^j),\delta_{z_h(t_n^j)}))|^2]\leq CN^2h_l^2\varepsilon^2.\end{aligned}$$

From (5.4.9) we have that

$$\mathrm{Var}(\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1}g_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j),\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j)\leq CNh_l\varepsilon^2.$$

Thus

$$\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n))]_q)\leq CN^2h_l^2\varepsilon^2+CNh_l\varepsilon^2.$$

Using the formula $\mathrm{Var}(\sum_{i=1}^d X_i)\leq d\sum_{i=1}^d\mathrm{Var}(X_i)$ with $i=q, X_i=[Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n)]_q$ yields

$$\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k)-Y_{h_l}^{\varepsilon,i,M}(t_n))]_q)\leq d^2CN^2h_l^2\varepsilon^2+d^2CNh_l\varepsilon^2\leq CNh_l\varepsilon^2.$$

Thus,

$$I_{2A}\leq CNh_l\varepsilon^2.$$

Next, we estimate I_{2B}. By Equation (2.2.8) there exists a random variable $s : \Omega \to [0,1]$ such that

$$f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})$$
$$= \mathbb{E}[\langle \partial_\mu f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}.$$

where $Y_n^s := sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)$. Let $\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q$ be the q-components of $\partial_\mu f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)$ respectively. Then

$$\begin{aligned}
&\mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)})\\
&= \mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}\\
&- \mathbb{E}[\partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q])\\
&= \mathrm{Var}(\mathbb{E}[(\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n)))\\
&\times [Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)})\\
&\le \mathbb{E}[(\mathbb{E}[(\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n)))\\
&\times [Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)})^2]\\
&\le \mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))|^2]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}\\
&\times \mathbb{E}[|[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q|^2]],
\end{aligned}$$

where we have use the Cauchy-Schwarz inequality in the penultimate step. By condition (5.3.4) and Lemma 5.3.7

$$\mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))|^2]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)} \le C\varepsilon^2$$

and by Lemma 5.4.2

$$\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^2] \le CN^2h_l^2 + CNh_l\varepsilon^2.$$

Therefore

$$I_{2B} \le CN^2h_l^2\varepsilon^2 + CNh_l\varepsilon^4,$$

and the proof is complete. $\square$

Lemma 5.5.6. *There exists positive constants C and $\bar{C}$ such that*

$$I_3 \leq CNh_l \sum_{q=1}^{d} \mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))]_q + CN^3h_l^3\varepsilon^2.$$

Proof. Note that

$$\begin{aligned}
&\mathrm{Var}(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M}))\\
&\quad \leq 2\mathrm{Var}(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}))\\
&\quad + 2\mathrm{Var}(f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})) =: I_{3A} + I_{3B}.
\end{aligned}$$

First, we estimate I_{3B}. By the mean value theorem there exists an $s \in [0,1]$ such that

$$\begin{aligned}
&f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\\
&\quad = \langle \nabla f_j(sY_{h_l}^{\varepsilon,i,M}(t_n) + (1-s)Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}), (Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))\rangle.
\end{aligned}$$

Let $\nabla_q f_j(sY_{h_l}^{\varepsilon,i,M}(t_n)+(1-s)Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})$ and $[(Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))]_q$ be the q components of $\nabla f_j(sY_{h_l}^{\varepsilon,i,M}(t_n)+(1-s)Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})$ and $(Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))$ respectively. We want to apply Lemma 5.5.1 with $A^{\varepsilon,h} = \nabla_q f_j(sY_{h_l}^{\varepsilon,i,M}(t_n) + (1-s)Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})$ and $B^{\varepsilon,h} = [(Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))]_q$ so we check that the three conditions are satisfied. Applying Lemma 5.5.2 with $\gamma = \nabla_q f_j, k = 0, h_{l_1} = h_{l-1}$ and $h_{l_2} = h_l$, we obtain

$$\mathrm{Var}(\nabla_q f_j(sY_{h_l}^{\varepsilon,i,M}(t_n) + (1-s)Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})) \leq C_1\varepsilon^2,$$

so the first condition of Lemma 5.5.1 is satisfied. Conditions 2 and 3 are satisfied by Assumption 5.3.1 and Lemma 5.5.3 respectively. Thus by Lemma 5.5.1 we have that

$$\begin{aligned}
&\mathrm{Var}(\nabla_q f_j(sY_{h_l}^{\varepsilon,i,M}(t_n) + (1-s)Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})[(Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))]_q)\\
&\qquad \leq 3C_3^2C_1N^2h_l^2\varepsilon^2 + 15C_2^2\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))]_q).
\end{aligned}$$

Using the formula $\mathrm{Var}(\sum_{i=1}^d X_i) \leq d\sum_{i=1}^d \mathrm{Var}(X_i)$ with $i = q, X_i = [Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q$ yields

$$\mathrm{Var}((Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))) \leq C\sum_{q=1}^{d} \mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))]_q + CN^2h_l^2\varepsilon^2.$$

Therefore,

$$I_{3A} \leq CN^2 h_l^2 \varepsilon^2 + C \sum_{q=1}^{d} \mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q).$$

Next we estimate I_{3B}. By Equation (2.2.8) there exists a random variable $s : \Omega \to [0,1]$ such that

$$\begin{aligned} f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - & f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M}) \\ &= \mathbb{E}[\langle \partial_\mu f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s), (Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))\rangle]_{Z=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)}. \end{aligned}$$

where $Y_n^s := sY_{h_l}^{\varepsilon,i,M}(t_n) + (1-s)Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)$. Let $\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $[Y_{h_l}^{\varepsilon,i,M}(t^n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q$ be the q-components of $\partial_\mu f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)$ respectively. Then

$$\begin{aligned} &\mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q]_{Z=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)}) \\ &= \mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]_{Z=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)} \\ &- \mathbb{E}[\partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n))[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q]) \\ &= \mathrm{Var}(\mathbb{E}[(\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n))) \\ &\times [Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q]]_{Z=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)}) \\ &\leq \mathbb{E}[(\mathbb{E}[(\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n))) \\ &\times [Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q]]_{Z=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)})^2] \\ &\leq \mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n))|^2]_{Z=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)} \\ &\times \mathbb{E}[|[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q|^2]], \end{aligned}$$

where we have use the Cauchy-Schwarz inequality in the penultimate step. By condition (5.3.4) and Lemma 5.3.7

$$\mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n))|^2]_{Z=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)} \leq C\varepsilon^2$$

and by Theorem 5.4.1

$$\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] \leq CN^2 h_l^2 + C\varepsilon^4 N h_l.$$

Therefore,

$$I_{3B} \leq CN^2 h_l^2 \varepsilon^2 + C\varepsilon^6 N h_l,$$

and the proof is complete.

$\square$

Lemma 5.5.7. *There exists a positive constant C such that*

$$I_4 \leq C\varepsilon^2 h_{l-1}^3 + C\varepsilon^4 h_{l-1}^2.$$

Proof. By Lemma 5.4.2 and Assumption 5.3.1 one can see that

$$\begin{aligned} I_4 &\leq 4\varepsilon^2 h_l \sum_{k=0}^{N-1} \mathbb{E}[|g_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})|^2] \\ &\leq 8\varepsilon^2 h_l N K(Ch_{l-1}^2 + C\varepsilon^2 h_{l-1}) = C\varepsilon^2 h_{l-1}^3 + C\varepsilon^4 h_{l-1}^2. \end{aligned}$$

$\square$

Lemma 5.5.8. *There exists a positive constant C such that*

$$I_5 \leq C\varepsilon^2 h_{l-1}^3 + C\varepsilon^6 h_{l-1}^2.$$

Proof. By Assumption 5.3.1 and Theorem 5.4.1 we have that

$$\begin{aligned} I_5 &\leq 4\varepsilon^2 h_l \sum_{k=0}^{N-1} \mathbb{E}[|g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - g_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})|^2] \\ &\leq 4\varepsilon^2 h_l N K(Ch_{l-1}^2 + C\varepsilon^4 h_{l-1}) = C\varepsilon^2 h_{l-1}^3 + C\varepsilon^6 h_{l-1}^2. \end{aligned}$$

$\square$

Lemma 5.5.9. *There exists a positive constant C such that*

$$I_6 \leq 2Nh_l \mathrm{Var}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j\Big) + CN^3 h_l^3 \varepsilon^2.$$

Proof. Since the covariance is a linear function, by subtracting and adding $f(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ to $f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})$ we have that

$$I_6 = 2\mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j,$$

$$h_l \sum_{k=0}^{N-1} [f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})]\Big)$$

$$+ 2\mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j,$$

$$h_l \sum_{k=0}^{N-1} [f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})]\Big)$$

$$=: I_{6A} + I_{6B}.$$

By Lemma 5.4.3, we obtain

$$I_{6A} = 2\mathrm{Cov}\left([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, h_l \sum_{k=0}^{N-1} (A_k^j + B_k^j + E_k^j)\right)$$

Using property (B.0.4) from the appendix we have

$$I_{6A} = 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, A_k^j\Big)$$
$$+ 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, B_k^j\Big)$$
$$+ 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, E_k^j\Big).$$

Using the definition of covariance and since the increments ξ_n^j in B_k^j are independent, we find that

$$\mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, B_k^j\Big)$$
$$= \mathbb{E}[[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j B_k^j] - \mathbb{E}[[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j]\mathbb{E}[B_k^j] = 0.$$

Then using (B.0.3) yields

$$I_{6A} \leq 2Nh_l \mathrm{Var}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j\Big) + h_l \sum_{k=0}^{N-1} \mathrm{Var}(A_k^j) + h_l \sum_{k=0}^{N-1} \mathrm{Var}(E_k^j). \quad (5.5.7)$$

Recall from Lemma 5.4.3 that

$$A_k^j = \langle \nabla f_j(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}), h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})\rangle.$$

In order to estimate $\mathrm{Var}(A_k^j)$ we use Lemma 5.5.1 with $A^{\varepsilon,h} = \nabla_q f_j(sY_{h_l}^{\varepsilon,i,M}(t_n^k) + (1-s)Y_{h_l}^{\varepsilon,i,M}(t_n)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $B^{\varepsilon,h} = [h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_q$ so we check that the three conditions are satisfied. The first and second conditions are satisfied by (5.5.6) and Assumption 5.3.1 respectively. By Lemma 5.3.3 and Assumption 5.3.1 we have that

$$|\mathbb{E}[[h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_q]| \leq CNh_l,$$

so the third condition is also satisfied. Thus Lemma 5.5.1 implies that

$$\mathrm{Var}(A_k^j) \leq CN^2h^2\varepsilon^2 + C\mathrm{Var}([h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_q).$$

Lemma 5.3.7 yields

$$\begin{aligned}
&\mathrm{Var}([h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_q) \\
&\quad = \mathrm{Var}([h_l \sum_{r=0}^{k-1} \{f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M}) - f(z_{h_l}(t_n^r), \delta_{z_{h_l}(t_n^r)})\}]_q) \\
&\quad \leq \mathbb{E}[|([h_l \sum_{r=0}^{k-1} \{f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M}) - f(z_{h_l}(t_n^r), \delta_{z_{h_l}(t_n^r)})\}]_q)|^2] \leq CN^2h_l^2\varepsilon^2.
\end{aligned}$$

Therefore

$$\mathrm{Var}(A_k^j) \leq CN^2h^2\varepsilon^2 + CN^2h_l^2\varepsilon^2. \tag{5.5.8}$$

From (5.4.15) we have

$$\mathrm{Var}(E_k^j) \leq \mathbb{E}[|E_k^j|^2] \leq CN^3h_l^3\varepsilon^2 + CN^2h_l^2\varepsilon^4. \tag{5.5.9}$$

Substituting (5.5.8) and (5.5.9) into (5.5.7) we obtain

$$I_{6A} \leq 2Nh_l\mathrm{Var}\big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j\big) + CN^3h_l^3\varepsilon^2.$$

Using Lemma 5.4.4, (B.0.3) and (B.0.4), yields

$$I_{6B} = 2\mathrm{Cov}\left([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, h_l \sum_{k=0}^{N-1} (\bar{A}_k^j + \bar{E}_k^j)\right)$$

$$
\begin{aligned}
&\le 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, \bar{A}_k^j\Big) \\
&+ 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, \bar{E}_k^j\Big) \\
&\le 2Nh_l \mathrm{Var}\big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j\big) + h_l \sum_{k=0}^{N-1} \mathrm{Var}(\bar{A}_k^j) + h_l \sum_{k=0}^{N-1} \mathrm{Var}(\bar{E}_k^j).
\end{aligned}
$$

Recall from Lemma 5.4.4 that

$$
\bar{A}_k^j = \mathbb{E}[\langle \partial_\mu f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s), h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})\rangle]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}.
$$

Let $\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})$ be the the q-components of $\partial_\mu f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})$ respectively. Then

$$
\begin{aligned}
&\mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}) \\
&= \mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)} \\
&- \mathbb{E}[\partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n)) h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]) \\
&= \mathrm{Var}(\mathbb{E}[(\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))) \\
&\times h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)}) \\
&\le \mathbb{E}[(\mathbb{E}[(\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))) \\
&\times h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)})^2] \\
&\le \mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))|^2]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)} \\
&\times \mathbb{E}[|h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})|^2]],
\end{aligned}
$$

where we have used the Cauchy-Schwarz inequality in the last step. By condition (5.3.4) and Lemma 5.3.7

$$
\mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))|^2]_{Z=Y_{h_l}^{\varepsilon,i,M}(t_n)} \le C\varepsilon^2
$$

and by Lemma 5.3.3 and Remark 5.3.1

$$\mathbb{E}[|h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})|^2]] \leq CN^2 h_l^2.$$

Thus,

$$\mathrm{Var}(\bar{A}_k^j) \leq CN^2 h_l^2 \varepsilon^2.$$

From (5.4.17) we have

$$\mathrm{Var}(\bar{E}_k^j) \leq \mathbb{E}[|\bar{E}_k^j|^2] \leq \bar{K}\varepsilon^2 C_1 N^3 h_l^3 + \bar{K}\varepsilon^4 CN^2 h_l^2.$$

Therefore,

$$I_{6B} \leq 2Nh_l \mathrm{Var}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j\Big) + CN^3 h_l^3 \varepsilon^2$$

and the proof is complete. □

Continuation of the proof of Theorem 5.5.4 By Lemmas 5.5.5-5.5.9, we have

$$\begin{aligned}\mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j) &\leq \mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j) \\ &+ CNh_l \sum_{q=1}^{d} \mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q) + CN^3 h_l^3 \varepsilon^2 + CN^2 h_l^2 \varepsilon^4.\end{aligned}$$

Taking the maximum in both sides yields that for $n \leq N^{l-1} - 1$

$$\begin{aligned}\max_{1\leq j\leq d} \mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j) &= \max_{1\leq j\leq d} \mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j) \\ &+ CNh_l \max_{1\leq j\leq d} \mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j) + CN^3 h_l^3 \varepsilon^2 + CN^2 h_l^2 \varepsilon^4.\end{aligned}$$

An application of the Grownwall inequality produces

$$\max_{\substack{0\leq n\leq N^{l-1} \\ 1\leq j\leq N}} \mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j) \leq CN^2 h_l^2 \varepsilon^2 + CNh_l \varepsilon^4. \tag{5.5.10}$$

In order to estimate $\mathrm{Var}(\Psi(Y_{h_l}^{\varepsilon,i,M}(t_n)) - \Psi(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))$ we apply the mean value theorem, so there exists $s \in [0,1]$ such that

$$\Psi(Y_{h_l}^{\varepsilon,i,M}(t_n)) - \Psi(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n) = \nabla\Psi(sY_{h_l}^{\varepsilon,i,M}(t_n) + (1-s)Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))(Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)).$$

We shall apply Lemma 5.5.1 with $A^{\varepsilon,h} = \nabla_q \Psi(sY^{\varepsilon,i,M}_{h_l}(t_n) + (1-s)Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))$ and $B^{\varepsilon,h} = [(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q$. Applying Lemma 5.5.2 with $\gamma = \nabla_q\Psi, k = 0, h_{l_1} = h_{l-1}$ and $h_{l_2} = h_l$, we obtain

$$\mathrm{Var}(\nabla_q \Psi(sY^{\varepsilon,i,M}_{h_l}(t_n) + (1-s)Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))) \leq C\varepsilon^2,$$

so the first condition of Lemma 5.5.1 is satisfied. Conditions 2 and 3 are satisfied by Assumption 5.3.1 and Lemma 5.5.3 respectively. Thus by Lemma 5.5.1 we have that

$$\begin{aligned}\mathrm{Var}(\nabla_q \Psi(sY^{\varepsilon,i,M}_{h_l}(t_n) + (1-s)Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)))[(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q)\\ \leq CN^2 h_l^2 \varepsilon^2 + C\mathrm{Var}([(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q).\end{aligned}$$

Thus

$$\mathrm{Var}(\Psi(Y^{\varepsilon,i,M}_{h_l}(t_n)) - \Psi(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)) \leq CN^2 h_l^2 \varepsilon^2 + C\mathrm{Var}((Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)). \tag{5.5.11}$$

Sustituting (5.5.10) into (5.5.11) we obtain the desire result. □

5.6 Simulations

In this section we provide an example (based on [23]) to ilustrate the main result of the previous section, Theorem 5.5.4. Consider the following MV-SDE with small noise

$$dX(t) = (-X(t) - \frac{1}{2}\mathbb{E}[X(t)])dt + \varepsilon X(t)dW(t), \quad X(0) = 1, \quad t \in [0, 1/2]. \tag{5.6.1}$$

As we have seen in section 5.3, SDE (5.6.1) can be regarded as the limit of the following interacting particle system

$$dX^{\varepsilon,i,M}(t) = (-X^{\varepsilon,i,M}(t) - \frac{1}{2}\frac{1}{M}\sum_{i=1}^{M} X^{\varepsilon,i,M}(t))dt + \varepsilon X^{\varepsilon,i,M}(t)dW^i(t). \tag{5.6.2}$$

Assume $M = 50$ and $\Psi(x) = x$. We simulated two coupled paths of SDE (5.6.2) with timesteps h_l and h_{l-1} by the MLMC method. We computed the simulated paths following section 5.3, by forming

$$\begin{aligned} X^{\varepsilon,i,M}_{h_l}(t_{n+1}) = X^{\varepsilon,i,M}_{h_l}(t_n) + \sum_{k=0}^{N-1} \left(-X^{\varepsilon,i,M}_{h_l}(t_n^k) - \frac{1}{2}\frac{1}{M}\sum_{i=1}^{M} X^{\varepsilon,i,M}(t_n^k) \right) h_l \\ + \varepsilon\sqrt{h_l}\sum_{k=0}^{N-1} X^{\varepsilon,i,M}_{h_l}\Delta\xi_n^k \end{aligned} \tag{5.6.3}$$

and

$$\begin{aligned} X^{\varepsilon,i,M}_{h_{l-1}}(t_{n+1}) = X^{\varepsilon,i,M}_{h_{l-1}}(t_n) + (-X^{\varepsilon,i,M}_{h_{l-1}}(t_n) - \frac{1}{2}\frac{1}{M}\sum_{i=1}^{M} X^{\varepsilon,i,M}(t_n))h_{l-1} \\ + \varepsilon\sqrt{h_l}X^{\varepsilon,i,M}_{h_{l-1}}(t_n)\sum_{k=0}^{N-1}\Delta\xi_n^k, \end{aligned} \tag{5.6.4}$$

where ξ_n^k samples of the standard normal random variable. We simulated in Matlab samples of the Equations (5.6.3) and (5.6.4) to test numerically the sharpness of the bound obtained in Theorem 5.5.4, i.e.,

$$\mathrm{Var}(X^{\varepsilon,i,M}_{h_l}(t_{n+1}) - X^{\varepsilon,i,M}_{h_{l-1}}(t_{n+1})) \leq \mathcal{O}(\varepsilon^2 h_{l-1}^2 + \varepsilon^4 h_{l-1}). \tag{5.6.5}$$

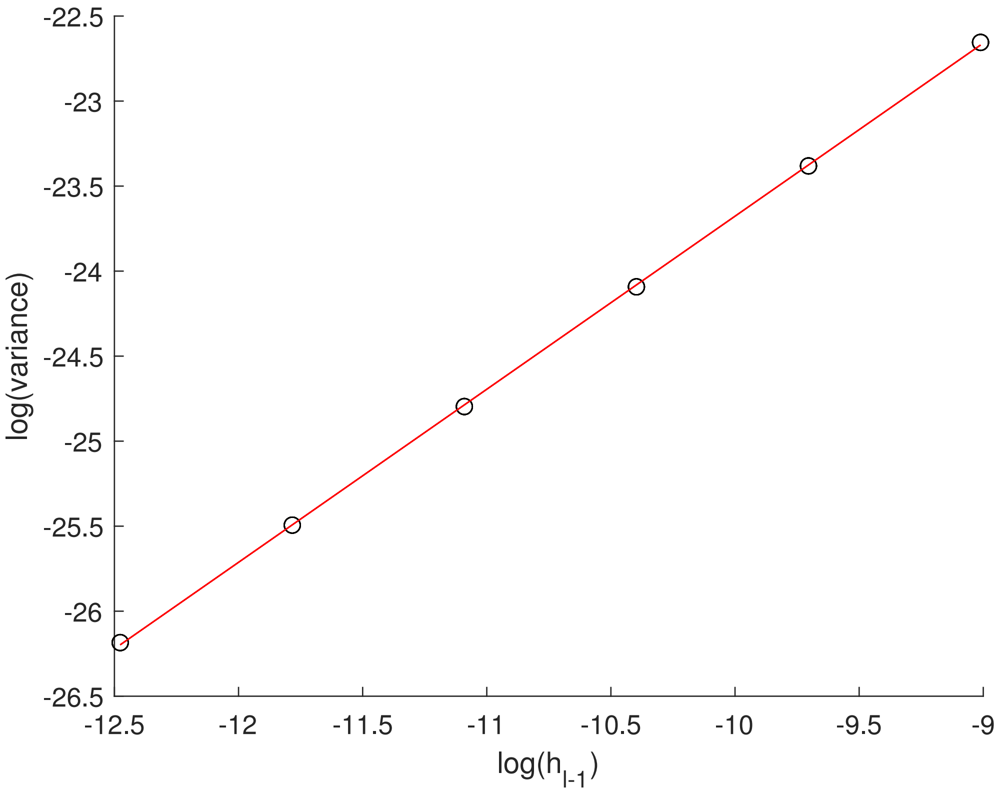
Note that

$$h_{l-1}^2\varepsilon^2 \text{ is the dominant term in } \varepsilon^2 h_{l-1}^2 + \varepsilon^4 h_{l-1} \text{ if and only if } h_{l-1} \geq \varepsilon^2. \tag{5.6.6}$$

In (5.6.5) we see that the variance is $\mathcal{O}(\varepsilon^2 h_{l-1}^2 + \varepsilon^4 h_{l-1})$. We formed the following 4 cases by choosing the parameters h_{l-1} and ε in a way that allows to study the dependency of the bound with respect to the terms $\varepsilon^2, h_{l-1}^2, \varepsilon^4$ and h_l individually.

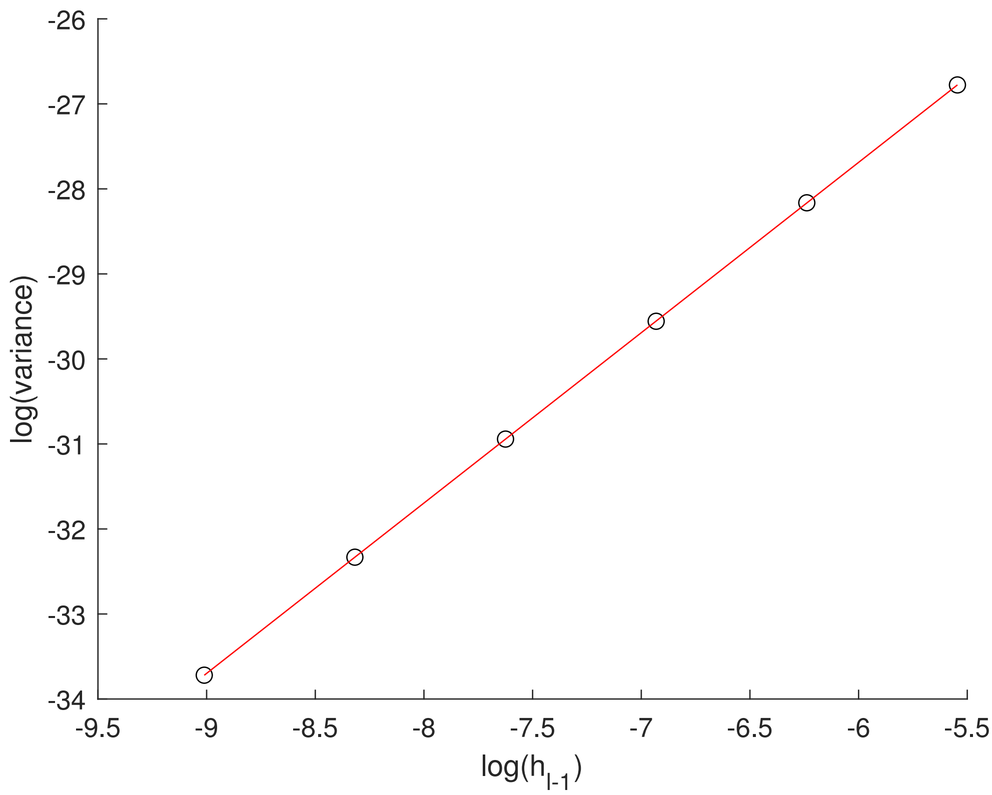
- The exponent of h_{l-1} in $\varepsilon^4 h_{l-1}$: we fix $\varepsilon = 2^{-4}$ and let $h_{l-1} \in \{2^{-13}, 2^{-14}, 2^{-15}, 2^{-16}, 2^{-17}, 2^{-18}\}$. With this choice of parameters we have that $h_{l-1} < \varepsilon^2$, so by (5.6.6) we know that $\varepsilon^4 h_{l-1}$ is likely to be the dominant term of the bound. We simulated the two coupled paths (5.6.2) and (5.6.3) six times, where ε is fixed and h_{l-1} is changing as described above. We plot the results in a 'log − log graph' where $\log(h_{l-1})$ and $\log(\mathrm{Var}(X^{\varepsilon,i,M}_{h_l}(t_{n+1}) - X^{\varepsilon,i,M}_{h_{l-1}}(t_{n+1})))$ are represented in the x and y-axis respectively, Figure 5.1. The black dots are the values of the variance obtained in the simulations for each h_{l-1}. The red line is the function $\boldsymbol{f(x) = 1.02x - 13.50}$, which is the best fit curve (linear regression) computed with the data (black dots). We observe that the slope is close to 1 in agreement with the exponent of h_{l-1} in $\varepsilon^4 h_{l-1}$.

Figure 5.1: Log of variance of two simulated coupled paths where $\varepsilon = 2^{-4}$ and $h_{l-1} \in \{2^{-13}, 2^{-14}, 2^{-15}, 2^{-16}, 2^{-17}, 2^{-18}\}$



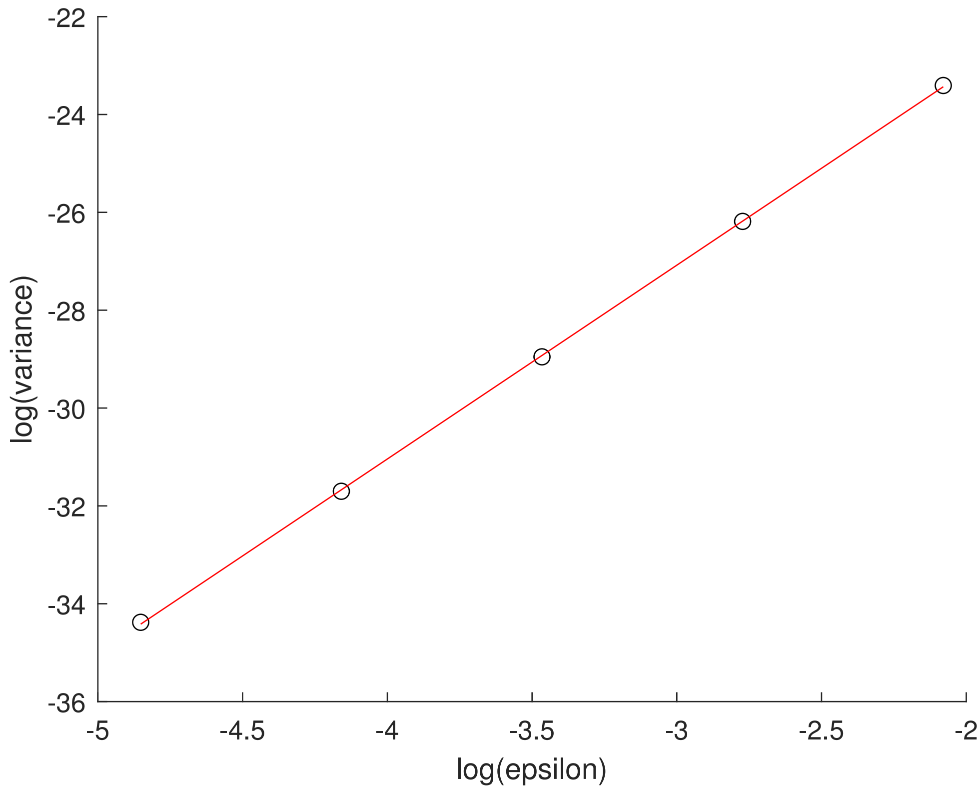
- <u>The exponent of h_{l-1} in $\varepsilon^2 h_{l-1}^2$</u>: we fix $\varepsilon = 2^{-10}$ and let $h_{l-1} \in \{2^{-8}, 2^{-9}, 2^{-10}, 2^{-11}, 2^{-12}, 2^{-13}\}$. With this choice of parameters we have that $h_{l-1} > \varepsilon^2$, so by (5.6.6) we know that $\varepsilon^2 h_{l-1}^2$ is likely to be the dominant term of the bound. We simulated the two coupled paths (5.6.2) and (5.6.3) six times, where ε is fixed and h_{l-1} is changing as described above. We plot the results in a 'log − log graph' where $\log(\varepsilon)$ and $\log(\mathrm{Var}(X_{h_l}^{\varepsilon,i,M}(t_{n+1}) - X_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})))$ are represented in the x and y-axis respectively, Figure 5.2. The black dots are the values of the variance obtained in the simulations for each h_{l-1}. The red line is the function $\boldsymbol{f(x) = 2.00x - 15.67}$, which is the best fit curve (linear regression) computed with the data (black dots). We observe that the slope is close to 2 in agreement with the exponent of h_{l-1} in $\varepsilon^2 h_{l-1}^2$.

Figure 5.2: Log of variance of two simulated coupled paths where $\varepsilon = 2^{-10}$ and $h_{l-1} \in \{2^{-8}, 2^{-9}, 2^{-10}, 2^{-11}, 2^{-12}, 2^{-13}\}$



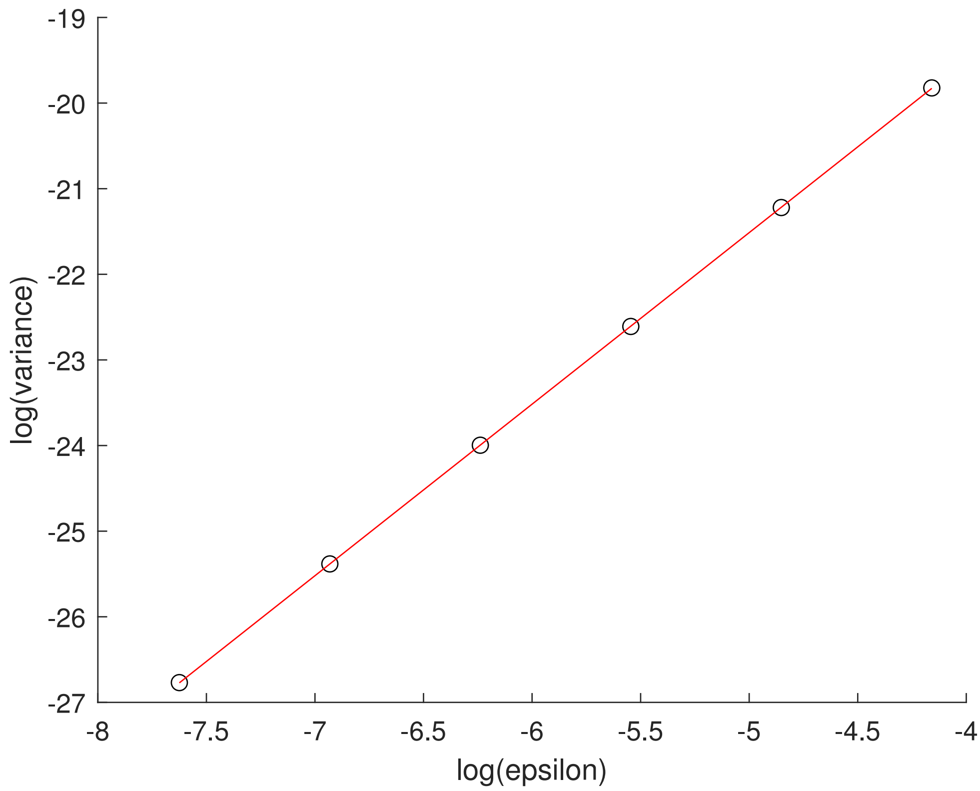
- <u>The exponent of ε in $\varepsilon^4 h_{l-1}$</u>: we fix $h_{l-1} = 2^{-18}$ and let $\varepsilon \in \{2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}\}$. With this choice of parameters we have that $h_{l-1} < \varepsilon^2$, so by (5.6.6) we know that $\varepsilon^4 h_{l-1}$ is likely to be the dominant term of the bound. We simulated the two coupled paths (5.6.2) and (5.6.3) five times, where h_{l-1} is fixed and ε is changing as described above. We plot the results in a 'log − log graph' where $\log(\varepsilon)$ and $\log(\text{Var}(X_{h_l}^{\varepsilon,i,M}(t_{n+1}) - X_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})))$ are represented in the x and y-axis respectively, Figure 5.3. The black dots are the values of the variance obtained in the simulations for each h_{l-1}. The red line is the function $\boldsymbol{f(x) = 3.96x - 15.19}$, which is the best fit curve (linear regression) computed with the data (black dots). We observe that the slope is close to 4 in agreement with the exponent of h_{l-1} in $\varepsilon^4 h_{l-1}$.

Figure 5.3: Log of variance of two simulated coupled paths where $h_{l-1} = 2^{-18}$ and $\varepsilon \in \{2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}\}$



- The exponent of ε in $\varepsilon^2 h_{l-1}^2$: we fix $h_{l-1} = 2^{-7}$ and let $\varepsilon \in \{2^{-6}, 2^{-7}, 2^{-8}, 2^{-9}, 2^{-10}, 2^{-11}\}$. With this choice of parameters we have that $h_{l-1} > \varepsilon^2$, so by (5.6.6) we know that $\varepsilon^2 h_{l-1}^2$ is likely to be the dominant term of the bound. We simulated the two coupled paths (5.6.2) and (5.6.3) five times, where h_{l-1} is fixed and ε is changing as described above. We plot the results in a 'log − log graph' where $\log(h_{l-1})$ and $\log(\mathrm{Var}(X_{h_l}^{\varepsilon,i,M}(t_{n+1}) - X_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})))$ are represented in the x and y-axis respectively, Figure 5.4. The black dots are the values of the variance obtained in the simulations for each h_{l-1}. The red line is the function $\boldsymbol{f(x) = 2.00x - 11.50}$, which is the best fit curve (linear regression) computed with the data (black dots). We observe that the slope is close to 2 in agreement with the exponent of ε in $\varepsilon^2 h_{l-1}^2$.

Figure 5.4: Log of variance of two simulated coupled paths where $h_{l-1} = 2^{-7}$ and $\varepsilon \in \{2^{-6}, 2^{-7}, 2^{-8}, 2^{-9}, 2^{-10}, 2^{-11}\}$



Chapter 6

Multilevel Monte Carlo EM scheme for MV-SDDEs with small noise

Here we extend the previous chapter to MV-SDDEs with small noise.

6.1 The EM Scheme for delay MV-SDDEs with small noise

Let W be $\bar{d}$-dimensional Brownian motion defined on a complete probability space and let $\tau > 0$. Consider the MV-SDDE with small noise of the form

$$\mathrm{d}X^{\varepsilon}(t) = f(X^{\varepsilon}(t), X^{\varepsilon}(t-\tau), \mathcal{L}_t^X)\mathrm{d}t + \varepsilon g(X^{\varepsilon}(t), X^{\varepsilon}(t-\tau), \mathcal{L}_t^X)\mathrm{d}W(t), t \in [0, T] \quad (6.1.1)$$

where $\varepsilon \in (0, 1)$, $\mathcal{L}_t^X$ is the law (or distribution) of $X(t)$,

$$f : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d \text{ and } g : \mathbb{R}^d \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d \times \bar{d}}$$

and the initial data satisfies the following condition: for any $p \geq 2$

$$\{X(\theta) : -\tau \leq \theta \leq 0\} := \Xi \in L^p_{\mathcal{F}_0}([-\tau, 0]; \mathbb{R}^d),$$

that is Ξ is a $\mathcal{F}_0$-measurable $C([-\tau, 0]; \mathbb{R}^d)$-valued random variable such that $E||\Xi||^p < \infty$.

As we explained in the prelimiaries, by the propagation of chaos result, Theorem 2.2.9, Equation (6.1.1) can be regarded as the limit of the following interacting M-particle system of $\mathbb{R}^d$-valued MV-SDDEs

$$\begin{aligned} \mathrm{d}X^{\varepsilon,i,M}(t) &= f(X^{\varepsilon,i,M}(t), X^{\varepsilon,i,M}(t-\tau), \mathcal{L}_t^{\varepsilon,X,M})dt \\ &\quad + \varepsilon g(X^{\varepsilon,i,M}(t), X^{\varepsilon,i,M}(t-\tau), \mathcal{L}_t^{\varepsilon,X,M})dW^i(t), \quad t \in [0,T], \end{aligned} \tag{6.1.2}$$

with the initial condition $X^{\varepsilon,i,M}(0) = x_0^i$, where $\mathcal{L}_t^{\varepsilon,X,M} := \frac{1}{M}\sum_{i=1}^{M} \delta_{X^{\varepsilon,i,M}(t)}$.

Our main task in this chapter is to discretize (6.1.2) using the EM scheme and estimate the variance of two coupled paths in the Multilevel Monte Carlo setting. As we discussed in the previous chapter, this directly translates into the computational cost of solving $\mathbb{E}[\Psi((X^{\varepsilon,i,M}(T)]$.

We shall impose the following hypothesis on the functions f and g:

Assumption 6.1.1. *There exists a positive constant* $K > 0$ *such that*

$$|f(x,y,\mu) - f(\bar{x},\bar{y},\nu)|^2 \vee |g(x,y,\mu) - g(\bar{x},\bar{y},\nu)|^2 \le K\big(|x-y|^2 + |\bar{x}-\bar{y}|^2 + \mathbb{W}_2^2(\mu,\nu)\big), \tag{6.1.3}$$

holds for any $x, y, \bar{x}, \bar{y} \in \mathbb{R}^d$, $\mu,\ \nu \in \mathcal{P}_2(\mathbb{R}^d)$. *Furthermore there exists a positive constant* $\bar{K}$ *such that*

$$|\nabla f(x,y,\mu)|^2 \vee |\nabla^2 f(x,y,\mu)|^2 \vee |\partial_\mu f(x,y,\mu)(z)|^2 \vee |\partial_\mu^2 f(x,y,\mu)(z)|^2 \vee |\partial_\mu^2 f(x,y,\mu)(z)|^2$$
$$\vee\, |\partial_\mu \nabla f(x,y,\mu)|^2 \le \bar{K}$$

for all $x,\ y \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. *In addition, there exists a positive constant* K *such that*

$$|\partial_\mu f(x,y,\mu)(z) - \partial_\mu f(\bar{x},\bar{y},\nu)(\bar{z})|^2 \le K\big(|x-\bar{x}|^2 + |y-\bar{y}|^2 + |z-\bar{z}|^2 + \mathbb{W}_2^2(\mu,\nu)\big). \tag{6.1.4}$$

for all $x, y, z, \bar{x}, \bar{y}, \bar{z} \in \mathbb{R}^d$, $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^d)$.

Remark 6.1.1. Assumption 6.1.1 implies the existence and uniqueness of the solution to equation (6.1.1), see Theorem 2.2.4. Moreover, under Assumption 6.1.1, we have

$$|f(x,y,\mu)|^2 \vee |g(x,y,\mu)|^2 \le \beta(1 + |x|^2 + |y|^2 + W_2^2(\mu)),$$

where $\beta = 2\max\{1, |f(0,0,\delta_0)|, |g(0,0,\delta_0)|\}$, and for any $x, y \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$.

We now introduce the EM scheme for (6.1.1). Given any time $T > 0$, assume that there exists a positive integer $h \in (0,1)$ such that $h = \frac{T}{N} = \frac{\tau}{m}$. Let $t_n = nh$ for $n \geq 0$. For $n = -m, ..., 0$, we compute the discrete approximations by setting $Y_h^\varepsilon(t_n) = \Xi(t_n)$. For $n = 1, ..., N$ we define

$$Y_{h,n+1}^{\varepsilon,i,M} = Y_{h,n}^{\varepsilon,i,M} + f(Y_{h,n}^{\varepsilon,i,M}, Y_{h,n-m}^{\varepsilon,i,M}, \mathcal{L}_h^{\varepsilon,Y_n,M})h + \varepsilon g(Y_{h,n}^{\varepsilon,i,M}, Y_{h,n-m}^{\varepsilon,i,M}, \mathcal{L}_h^{\varepsilon,Y_n,M})\Delta W^i(t_n), \tag{6.1.5}$$

where $\mathcal{L}_h^{\varepsilon,Y_n,M} = \frac{1}{M}\sum_{j=1}^{M} \delta_{Y_{h,n}^{\varepsilon,j,M}}$ and $\Delta W^i(t_n) = W^i(t_{n+1}) - W^i(t_n)$.

Let

$$Y_h^{\varepsilon,i,M}(t) = Y_{h,n}^{\varepsilon,i,M}, \qquad t \in [t_n, t_{n+1}). \tag{6.1.6}$$

For convenience, we define $\mathcal{L}_{h,t}^{\varepsilon,Y,M} = \frac{1}{M}\sum_{j=1}^{M} \delta_{Y_h^{\varepsilon,i,M}(t)}$ and $\eta_h(t) := \lfloor t/h \rfloor h$ for $t \geq 0$. Then one observes $\mathcal{L}_{h,t}^{\varepsilon,Y,M} = \mathcal{L}_{h,\eta_h(t)}^{\varepsilon,Y,M} = \mathcal{L}_h^{\varepsilon,Y_n,M}$, for $t \in [t_n, t_{n+1})$. We now define the EM continuous approximate solution as follows:

$$\begin{aligned}\bar{Y}_h^{\varepsilon,i,M}(t) = x_0^i &+ \int_0^t f(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M})ds \\ &+ \varepsilon \int_0^t g(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M})dW^i(s),\ t \in [0,T].\end{aligned} \tag{6.1.7}$$

The next lemma show the boundedness of the pth-moments of the EM approximate solution.

Lemma 6.1.2. *Let Assumption 6.1.1 hold. Given $T > 0$ and $p > 0$, we have that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |\bar{Y}_h^{\varepsilon,i,M}(t)|^p\right] \leq C.$$

Proof. Let $p \geq 4$. From (6.1.7) we have that

$$\begin{aligned}|\bar{Y}_h^{\varepsilon,i,M}(t)|^p = \Big| x_0^i &+ \int_0^t f(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M})ds \\ &+ \varepsilon \int_0^t g(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M})dW^i(s)\Big|^p.\end{aligned}$$

Using the Hölder and the BDG inequalities we obtain that for every $\hat{t} \leq T$

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |\bar{Y}_h^{\varepsilon,i,M}(t)|^p\right] \leq 3^{p-1}|x_0|^p + (3T)^{p-1}\mathbb{E}\int_0^{\bar{t}} |f(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M})|^p ds$$
$$+ \varepsilon 3^{p-1}C\mathbb{E}\left[\left(\int_0^{\bar{t}} |g(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M})|^2 ds\right)^{p/2}\right].$$

By Remark 6.1.1 and the Wasserstein distance definition 2.2.2, one can see that for every $\hat{t} \leq T$

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |\bar{Y}_h^{\varepsilon,i,M}(t)|^p\right] \leq C + C\int_0^{\bar{t}} \mathbb{E}[\sup_{0\leq s\leq t} |\bar{Y}_h^{\varepsilon,i,M}(s)|^p]dt.$$

The required assertion follows from the Gronwall inequality. Thus, the result is proved for $p \geq 4$. For $0 < p < 4$, the result follows proceeding similarly as in the last part of the proof of Lemma 3.3.6. □

The following lemma will be used later when estimating the variance of two coupled processes in the MLMC setting. Let $z_h(t) := \Xi(t)$ for $t \in [-\tau, 0]$ and for $t \in [0, T]$, let $z_h(t)$ be the solution to

$$z_h(t) = X(0) + \int_0^t f(z_h(\eta_h(s)), z_h(\eta_h(s-\tau)), \delta_{z_h(s)})ds, \tag{6.1.8}$$

Lemma 6.1.3. *For any $T > 0$ we have*

$$\mathbb{E}[\sup_{0\leq s\leq T} |\bar{Y}_h^{\varepsilon,i,M}(s) - z_h(s)|^2] \leq C\varepsilon^2. \tag{6.1.9}$$

Proof. Using (6.1.7) and (6.1.9), using the fact that $|a+b|^2 \leq 2a^2 + 2b^2$ and the Cauchy-Schwarz inequality we have that for every $t \leq T$

$$|\bar{Y}_h^{\varepsilon,i,M}(t) - z_h(t)|^2$$
$$= \Big|\int_0^t (f(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M}) - f(z_h(\eta_h(s)), z_h(\eta_h(s-\tau)), \delta_{z_h(s)})))ds$$
$$+ \varepsilon\int_0^t g(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M})dW^i(s)\Big|^2$$
$$\leq 2T\int_0^t |f(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M}) - f(z_h(\eta_h(s)), z_h(\eta_h(s-\tau)), \delta_{z_h(s)}))|^2 ds$$

$$+ 2\varepsilon^2 \left| \int_0^t g(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M}) dW^i(s) \right|^2.$$

By the BDG inequality we have that

$$\mathbb{E}\left[\sup_{0\le s\le t} \left| \int_0^t g(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M}) dW^i(s) \right|^2 \right]$$
$$\le 4 \int_0^t \mathbb{E}[|g(Y_h^{\varepsilon,i,M}(s), Y_h^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h,s}^{\varepsilon,Y,M})|^2] ds.$$

Thus by Assumption 6.1.1 one can see that

$$\mathbb{E}[\sup_{0\le s\le t} |\bar{Y}_h^{\varepsilon,i,M}(t) - z_h(t)|^2] \le 2TK \int_0^t (2\mathbb{E}[\sup_{0\le s\le r} |\bar{Y}_h^{\varepsilon,i,M}(s) - z_h(s)|^2]$$
$$+ \sup_{0\le s\le r} \mathbb{W}_2^2(\mathcal{L}_{h,s}^{\varepsilon,Y,M}, \delta_{z_h(s)}))dr + 8T\varepsilon^2\beta \int_0^t (1 + 2\mathbb{E}[\sup_{0\le s\le r} |\bar{Y}_h^{\varepsilon,i,M}(s)|^2] + W_2^2(\mathcal{L}_{h,s}^{\varepsilon,Y,M})ds.$$

Using (2.2.1), (2.2.2) and Lemma 6.1.2 we have that for all $0 \le t \le T$

$$\mathbb{E}[\sup_{0\le s\le t} |\bar{Y}_h^{\varepsilon,i,M}(t) - z_h(t)|^2] \le C\varepsilon^2 + C \int_0^t \mathbb{E}[\sup_{0\le s\le r} |\bar{Y}_h^{\varepsilon,i,M}(s) - z_h(s)|^2] dr.$$

The final result is obtained by applying the Gronwall inequality. □

6.1.1 The Multilevel Monte Carlo-EM Scheme

We now define the multilevel Monte Carlo EM scheme. Given any $T > 0$, let $N \ge 2, l \in \{0, ..., L\}$, where L is a positive integer that will be determined later. Let $h_l = T \cdot N^{-l}, h_{l-1} = T \cdot N^{-(l-1)}$. Assume there exist positive integers m_l and m_{l-1} such that $h_l = \tau/m_l$ and $h_{l-1} = \tau/m_{l-1}$ respectively.

For step sizes h_l and h_{l-1} the EM continuous approximate solutions are respectively

$$\bar{Y}_{h_l}^{\varepsilon,i,M}(t) = x^i + \int_0^t f(Y_{h_l}^{\varepsilon,i,M}(s), Y_{h_l}^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h_l,s}^{\varepsilon,Y,M}) ds$$
$$+ \int_0^t g(Y_{h_l}^{\varepsilon,i,M}(s), Y_{h_l}^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h_l,s}^{\varepsilon,Y,M}) dW^i(s),$$

and

$$\bar{Y}_{h_{l-1}}^{\varepsilon,i,M}(t) = x^i + \int_0^t f(Y_{h_{l-1}}^{\varepsilon,i,M}(s), Y_{h_{l-1}}^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h_{l-1},s}^{\varepsilon,Y,M}) ds$$
$$+ \int_0^t g(Y_{h_{l-1}}^{\varepsilon,i,M}(s), Y_{h_{l-1}}^{\varepsilon,i,M}(s-\tau), \mathcal{L}_{h_{l-1},s}^{\varepsilon,Y,M}) dW^i(s).$$

We now construct the discrete version of the previous approximate solutions using the same Brownian motion for both processes. We say that the two processes are coupled. For $n \in \{0, 1, \ldots, N^{l-1} - 1\}$ and $k \in \{0, \ldots, N\}$, let

$$t_n = nh_{l-1} \text{ and } t_n^k = nh_{l-1} + kh_l.$$

This means we divide the interval $[t_n, t_{n+1}]$ by h_l into N equal parts with $t_n^0 = t_n, t_n^N = t_{n+1}$. For $n \in \{0, 1, \ldots, N^{l-1} - 1\}$ and $k \in \{0, \ldots, N - 1\}$, let

$$\begin{aligned} Y_{h_l}^{\varepsilon,i,M}(t_n^{k+1}) = Y_{h_l}^{\varepsilon,i,M}(t_n^k) &+ f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})h_l \\ &+ \varepsilon\sqrt{h_l} g(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})\Delta\xi_n^k, \end{aligned} \tag{6.1.10}$$

where $\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M} = \frac{1}{M}\sum_{j=1}^M \delta_{Y_{h_l}^{\varepsilon,j,M}(t_n^k)}$, the random vector $\Delta\xi_n^k \in \mathbb{R}^{\bar{d}}$ has independent components, and each component is distributed as $\mathcal{N}(0,1)$. Therefore, to simulate $Y_{h_l}^{\varepsilon,i,M}$, we use

$$\begin{aligned} Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) = Y_{h_l}^{\varepsilon,i,M}(t_n) &+ \sum_{k=0}^{N-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})h_l \\ &+ \varepsilon\sqrt{h_l}\sum_{k=0}^{N-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})\Delta\xi_n^k. \end{aligned} \tag{6.1.11}$$

To simulate $Y_{h_{l-1}}^{\varepsilon,i,M}$, we use

$$\begin{aligned} Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1}) = Y_{h_{l-1}}^{\varepsilon,i,M}(t_n) &+ f(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n - m_{l-1}h_{l-1}), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})h_{l-1} \\ &+ \varepsilon\sqrt{h_l} g(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n - m_{l-1}h_{l-1}), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\sum_{k=0}^{N-1}\Delta\xi_n^k, \end{aligned} \tag{6.1.12}$$

where $\mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M} = \frac{1}{M}\sum_{j=1}^M \delta_{Y_{h_{l-1}}^{\varepsilon,j,M}(t_n)}$.

The following theorem is the main result of this section.

Theorem 6.1.4. *Let Assumption 6.1.1 hold. Then it holds that*

$$\max_{0 \le n < M^{l-1}} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] \le CN^2h_l^2 + \bar{C}\varepsilon^4 Nh_l.$$

In order to prove Theorem 6.1.4, we need a few lemmas.

Lemma 6.1.5. *Let $p \geq 2$. Then*

$$\max_{\substack{0\leq n\leq N^{l-1}\\ 1\leq k\leq N}} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^p] \leq CN^p h_l^p + CN^{p/2} h_l^{p/2} \varepsilon^p,$$

where C is a positive constant that only depend on β, T, m and $X^\varepsilon(0)$ (β from Remark 6.1.1).

Proof. Let $p = 4$. From (6.1.10) we have that

$$\begin{aligned} Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n) = &\sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) h_l \\ &+ \varepsilon\sqrt{h_l} \sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) \Delta\xi_n^j. \end{aligned} \tag{6.1.13}$$

Hence, we obtain

$$\begin{aligned} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^p] \leq\ &2^{p-1}\mathbb{E}\left|\sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) h_l\right|^p \\ &+ 2^{p-1}\mathbb{E}\left|\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) \Delta\xi_n^j\right|^p. \end{aligned} \tag{6.1.14}$$

By Remark 6.1.1 and Lemma 6.1.2 one can see that

$$\begin{aligned} &\mathbb{E}\left|\sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) h_l\right|^p \\ &\leq N^{p-1} \sum_{j=0}^{k-1} \mathbb{E}\left| f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) h_l\right|^p \\ &\leq N^{p-1} \sum_{j=0}^{k-1} \mathbb{E}\left[\left(\beta\left(1 + |Y_{h_l}^{\varepsilon,i,M}(t_n^j)|^2 + |Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l)|^2 + W_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\right)\right)^{p/2}\right] \\ &\leq CN^{p-1} h_l^p \sum_{j=0}^{k-1}\left(1 + 2\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^j)|^p] + \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l)|^p]\right) \leq CN^p h_l^p. \end{aligned} \tag{6.1.15}$$

Using the BDG inequality, Remark 6.1.1 and Lemma 6.1.2, we obtain

$$
\begin{aligned}
&\mathbb{E}\left|\varepsilon\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\sqrt{h_l}\Delta\xi_n^j\right|^p \\
&\leq C\varepsilon^p \mathbb{E}\left[\left|\sum_{j=0}^{k-1} |g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})|^2 h_l\right|^{p/2}\right] \\
&\leq C\varepsilon^p N^{p/2-1} h_l^{p/2} \mathbb{E}\left[\sum_{j=0}^{k-1} (|g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})|^2)^{p/2}\right] \\
&\leq C\varepsilon^p N^{p/2-1} h_l^{p/2} \sum_{j=0}^{k-1} \mathbb{E}\left[\left(\left(1 + |Y_{h_l}^{\varepsilon,i,M}(t_n^j)|^2 + |Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l)|^2 + W_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\right)\right)^{p/2}\right] \\
&\leq C N^{p/2} h_l^{p/2} \varepsilon^p.
\end{aligned}
\tag{6.1.16}
$$

The result follows from substituting (6.1.15) and (6.1.16) into (6.1.14). □

Lemma 6.1.6. *Let f_m be the m^{th} component of f. Then there exist $s, r \in [0,1]$ such that*

$$
\begin{aligned}
f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
= A_k + B_k + E_k,
\end{aligned}
$$

where

$$
\begin{aligned}
A_k &= (A_k^1, ..., A_k^d)', B_k = (B_k^1, ..., B_k^d)', E_k = (E_k^1, ..., E_k^d)' \\
A_k^m &:= \langle \nabla f_m(s(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l)) \\
&\quad + (1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - m_l h_l)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}), (H_k^{11}, H_k^{1,2})\rangle, \\
B_k^m &:= \langle \nabla f_m(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}), (H_k^{21}, H_k^{2,2})\rangle, \\
E_k^m &:= \langle \nabla^2 f_m\big(rs((Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l)) - (Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - m_l h_l))) \\
&\quad + (Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - m_l h_l)), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}\big) \\
&\quad \times \big((Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l)) - (Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - m_l h_l))\big)s, \\
&\quad (H_k^{21}, H_k^{2,2})\rangle \\
&\quad + \langle \mathbb{E}[\langle \partial_\mu \nabla f_m(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^w,M})(Y_w^s), \\
&\quad \times \big((Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l)) \\
&\qquad\qquad - (Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - m_l h_l))\big)\rangle]_{Z_1 = Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2 = Y_{h_l}^{\varepsilon,i,M}(t_n - m_l h_l)}, \\
&\quad (H_k^{21}, H_k^{2,2})\rangle, m \in \{1, ..., d\}.
\end{aligned}
$$

$$
\begin{aligned}
H_k^{11} &:= h_l \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}), \\
H_k^{1,2} &:= h_l \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), Y_{h_l}^{\varepsilon,i,M}(t_n^j - 2m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}), \\
H_k^{21} &:= \varepsilon\sqrt{h_l} \sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j, \\
H_k^{22} &:= \varepsilon\sqrt{h_l} \sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j - m_l h_l), Y_{h_l}^{\varepsilon,i,M}(t_n^j - 2m_l h_l), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j.
\end{aligned}
$$

Proof. The proof is similar to the one in Lemma 5.4.3, we omit it here. □

Remark 6.1.2. $Y_{h_l}^{\varepsilon,i,M}(t_n^k)$ is a d-dimensional vector. $(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l))$ is a $2d$-dimensional vector. H_k^{11} is a d-dimensional vector. (H_k^{11}, H_k^{12}) is a $2d$-dimensional vector. $\nabla^2 f_m\big(rs((Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - m_l h_l)) - (Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - m_l h_l)))$ is a $2d \times 2d$-matrix.

Lemma 6.1.7. *There exist random variables $s, r : \Omega \to [0,1]$ such that*

$$f(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - m_l h_l), \mathcal{L}^{\varepsilon,Y^k_n,M}_{h_l}) - f(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - m_l h_l). \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) = \bar{A}_k + \bar{E}_k,$$

where

$$\begin{aligned}
\bar{A}_k &= (\bar{A}^1_k, ..., \bar{A}^d_k)', \bar{E}_k = (\bar{E}^1_k, ..., \bar{E}^d_k)' \\
\bar{A}^m_k &:= \mathbb{E}[\langle \partial_\mu f_m(Z_1, Z_2, \mathcal{L}^{\varepsilon,Y^s_n,M}_{h_l})(Y^s_n), h_l \\
&\quad \times \sum_{j=0}^{k-1} f(Y^{\varepsilon,i,M}_{h_l}(t^j_n), Y^{\varepsilon,i,M}_{h_l}(t^j_n - m_l h_l), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l}))\rangle]_{Z_1 = Y^{\varepsilon,i,M}_{h_l}(t_n), Z_2 = Y^{\varepsilon,i,M}_{h_l}(t_n - m_l h_l)} \\
\bar{E}^m_k &:= \mathbb{E}[\langle \partial^2_\mu f_m(Z_1, Z_2, \mathcal{L}^{\varepsilon,Y^{s,r}_n,M}_{h_l})(Y^{s,r}_n) \\
&\quad \times ((Y^{\varepsilon,i,M}_{h_l}(t^k_n), Y^{\varepsilon,i,M}_{h_l}(t^k_n - m_l h_l)) - (Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - m_l h_l)))s, \varepsilon\sqrt{h_l} \\
&\quad \times \sum_{j=0}^{k-1} g(Y^{\varepsilon,i,M}_{h_l}(t^j_n), Y^{\varepsilon,i,M}_{h_l}(t^j_n - m_l h_l), \mathcal{L}^{\varepsilon,Y^j_n,M}_{h_l})\Delta\xi^j_n\rangle]_{Z_1 = Y^{\varepsilon,i,M}_{h_l}(t_n), Z_2 = Y^{\varepsilon,i,M}_{h_l}(t_n - m_l h_l)}, \\
Y^s_n &:= s(Y^{\varepsilon,i,M}_{h_l}(t^k_n), Y^{\varepsilon,i,M}_{h_l}(t^k_n - m_l h_l)) + (1-s)(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - m_l h_l)), \\
Y^{s,r}_n &:= sr\big((Y^{\varepsilon,i,M}_{h_l}(t^k_n), Y^{\varepsilon,i,M}_{h_l}(t^k_n - m_l h_l)) - (Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - m_l h_l))\big) \\
&\quad + (Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - m_l h_l)).
\end{aligned}$$

Proof. The proof is the same as the one in Lemma 5.4.4. □

Proof of Theorem 6.1.4 Recall that $h_l = \tau/m_l$, so for notational convenience we will write $Y^{\varepsilon,i,M}_{h_{l-1}}(t_n - \tau)$ instead of $Y^{\varepsilon,i,M}_{h_{l-1}}(t_n - m_l h_l)$. From (6.1.11) and (6.1.12) we have that for $n \leq N^{l-1} - 1$

$$\begin{aligned}
Y^{\varepsilon,i,M}_{h_l}(t_{n+1}) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_{n+1}) &= Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n) \\
&\quad + h_l \sum_{k=0}^{N-1} \Big(f(Y^{\varepsilon,i,M}_{h_l}(t^k_n), Y^{\varepsilon,i,M}_{h_l}(t^k_n), Y^{\varepsilon,i,M}_{h_l}(t^k_n - \tau), \mathcal{L}^{\varepsilon,Y^k_n,M}_{h_l}) \\
&\quad - f(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - \tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \Big) \\
&\quad + h_l \sum_{k=0}^{N-1} \Big(f(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - \tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \\
&\quad - f(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n - \tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_{l-1}}) \Big)
\end{aligned}$$

$$
\begin{aligned}
&+\varepsilon\sqrt{h_l}\sum_{k=0}^{N-1}\Big(g(Y_{h_l}^{\varepsilon,i,M}(t_n^k),Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})\\
&-g(Y_{h_l}^{\varepsilon,i,M}(t_n),Y_{h_l}^{\varepsilon,i,M}(t_n-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\Big)\Delta\xi_n^k\\
&+\varepsilon\sqrt{h_l}\sum_{k=0}^{N-1}\Big(g(Y_{h_l}^{\varepsilon,i,M}(t_n),Y_{h_l}^{\varepsilon,i,M}(t_n-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\\
&-g(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau),\mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\Big)\Delta\xi_n^k\\
&=:Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)+R_N.
\end{aligned}
$$

By using the linearity property of the inner product, we obtain

$$
\begin{aligned}
&|Y_{h_l}^{\varepsilon,i,M}(t_{n+1})-Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})|^2=\langle Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)\\
&\quad+R_N,Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)+R_N\rangle\\
&=|Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2+|R_N|^2+2\langle Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),R_N\rangle.
\end{aligned}
$$

Applying the elementary inequality $|a+b+c+d|^2\leq 4|a|^2+4|b|^2+4|c|^2+4|d|^2$ to the term $|R_N|^2$ above, we derive that

$$
\begin{aligned}
&|Y_{h_l}^{\varepsilon,i,M}(t_{n+1})-Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})|^2\leq|Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2\\
&+4h_l^2\left|\sum_{k=0}^{N-1}\Big(f(Y_{h_l}^{\varepsilon,i,M}(t_n^k),Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})-f(Y_{h_l}^{\varepsilon,i,M}(t_n),Y_{h_l}^{\varepsilon,i,M}(t_n-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\Big)\right|^2\\
&+4h_l^2\left|\sum_{k=0}^{N-1}\Big(f(Y_{h_l}^{\varepsilon,i,M}(t_n),Y_{h_l}^{\varepsilon,i,M}(t_n-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})-f(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau),\mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\Big)\right|^2\\
&+4\varepsilon^2\Big|\sum_{k=0}^{N-1}\Big(g(Y_{h_l}^{\varepsilon,i,M}(t_n^k),Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})\\
&\qquad-g(Y_{h_l}^{\varepsilon,i,M}(t_n),Y_{h_l}^{\varepsilon,i,M}(t_n-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\Big)\sqrt{h_l}\Delta\xi_n^k\Big|^2\\
&+4\varepsilon^2\Big|\sum_{k=0}^{N-1}\Big(g(Y_{h_l}^{\varepsilon,i,M}(t_n),Y_{h_l}^{\varepsilon,i,M}(t_n-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\\
&\qquad-g(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau),\mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\Big)\sqrt{h_l}\Delta\xi_n^k\Big|^2\\
&+2h_l\sum_{k=0}^{N-1}\langle Y_{h_l}^{\varepsilon,i,M}(t_n)-Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),f(Y_{h_l}^{\varepsilon,i,M}(t_n^k),t_n^k-\tau),\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})
\end{aligned}
$$

$$
\begin{aligned}
&\quad - f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\rangle \\
&+ 2h_l \sum_{k=0}^{N-1} \langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), Y_{h_l}^{\varepsilon,i,M}(t_n), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \\
&\quad - f(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\rangle \\
&+ 2\varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \big(g(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
&\quad - g(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\big)\Delta\xi_n^k\rangle \\
&+ 2\varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \big(g(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \\
&\quad - g(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\big)\Delta\xi_n^k\rangle.
\end{aligned}
$$

Now, we take expectations on both sides of the previous inequality. Since $\Delta\xi_n^k$ is independent of $Y_{h_l}^{\varepsilon,i,M}(t_n^k)$ and $Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)$, the expectation of the last two summands in the equation above is zero. Thus,

$$
\begin{aligned}
&\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})|^2] \leq \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] \qquad (6.1.17)\\
&+ 4Nh_l^2 \sum_{k=0}^{N-1} \mathbb{E}\left| f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\right|^2 \\
&+ 4Nh_l^2 \sum_{k=0}^{N-1} \mathbb{E}\left| f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\right|^2 \\
&+ 4\varepsilon^2 \mathbb{E}\Big[\Big| \sum_{k=0}^{N-1} \Big(g(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
&\qquad\qquad - g(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)\mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\Big)\sqrt{h_l}\Delta\xi_n^k\Big|^2\Big] \\
&+ 4\varepsilon^2 \mathbb{E}\Big[\Big| \sum_{k=0}^{N-1} \Big(g(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \\
&\qquad\qquad - g(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\Big)\sqrt{h_l}\Delta\xi_n^k\Big|^2\Big] \\
&+ 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
&\qquad\qquad - f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\rangle]
\end{aligned}
$$

$$+ 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})$$

$$- f(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M}))\rangle].$$

$$=: \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] + I_1 + I_2 + I_3 + I_4 + I_5 + I_6. \tag{6.1.18}$$

By Assumption 6.1.1 and Lemma 6.1.5, one can see that

$$I_1 \le 4KNh_l^2 \Big(\sum_{k=0}^{N-1} (\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^2 + \mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau) - Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)|^2$$

$$+ \mathbb{W}_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}, \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})) \Big) \le CN^4h_l^4 + CN^3\varepsilon^2h_l^3.$$

Also, by Assumption 6.1.1

$$I_2 \le 4KNh_l^2 \Big(\sum_{k=0}^{N-1} (\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + \mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n-\tau) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau)|^2$$

$$+ \mathbb{W}_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n,M}, \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})) \Big)$$

$$\le CN^2h_l^2 \Big(\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] + \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n-\tau) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau)|^2] \Big).$$

By the BDG inequality, Assumption 6.1.1 and Lemma 6.1.5, we obtain

$$I_3 \le C\varepsilon^2 \sum_{k=0}^{N-1} \mathbb{E}[|g(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$$

$$- g(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})|^2]h_l$$

$$= Ch_l\varepsilon^2 \sum_{k=0}^{N-1} \Big(\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^2] + \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau) - Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)|^2]$$

$$+ \mathbb{W}_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}, \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \Big) \le CN^3h_l^3\varepsilon^2 + CN^2h_l^2\varepsilon^4.$$

Similarly to I_3,

$$I_4 \le CNh_l\varepsilon^2 \Big(\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] + \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n-\tau) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau)|^2] \Big).$$

An application of the Cauchy-Schwarz inequality and Assumption 6.1.1 gives

$$I_5 = 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\big\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),$$

$$f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\rangle]$$

$$= 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),$$

$$f(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})\rangle]$$

$$+ 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n),$$

$$f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})\rangle]$$

$$=: I_{5A} + I_{5B}.$$

Applying Lemma 6.1.6 we have

$$I_{5A} \le 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), A_k\rangle] + 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), B_k\rangle]$$
$$+ 2h_l \sum_{k=0}^{N-1} \mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), E_k\rangle].$$

By independence, the second summand above is zero. Also, we note that

$$\mathbb{E}[|A_k|^2] = \sum_{m=1}^{d} \mathbb{E}[|A_k^m|]^2 \le d\bar{K}\mathbb{E}\left|h_l \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\right|^2$$
$$+ d\bar{K}\mathbb{E}\left|h_l \sum_{j=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^j-\tau), Y_{h_l}^{\varepsilon,i,M}(t_n^j-2\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\right|^2$$
$$\le d\bar{K}h_j^2 N \sum_{j=0}^{k-1} \mathbb{E}\Big[\Big(\beta\Big(1 + |Y_{h_l}^{\varepsilon,i,M}(t_n^j)|^2$$
$$+ 2|Y_{h_l}^{\varepsilon,i,M}(t_n^j-\tau)|^2 + |Y_{h_l}^{\varepsilon,i,M}(t_n^j-2\tau)|^2 + 2W_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Big)\Big)^2\Big]$$
$$\le \bar{K}h_l^2 N^2 C.$$

and

$$\mathbb{E}[|E_k|^2] = \sum_{m=1}^{d} \mathbb{E}[(E_k^m)^2]$$
$$\le 2d\bar{K}\varepsilon^2 h_l \mathbb{E}\Big[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^2$$

$$\times \left|\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\right|^2\Bigg]$$

$$+ 2d\bar{K}\varepsilon^2 h_l\mathbb{E}\Bigg[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^2$$

$$\times \left|\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j-\tau), Y_{h_l}^{\varepsilon,i,M}(t_n^j-2\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\right|^2\Bigg]$$

$$\leq 2d\bar{K}\varepsilon^2 h_l(\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^4])^{1/2}\left(\mathbb{E}\left[\left|\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\right|^4\right]\right)^{1/2}$$

$$+ 2d\bar{K}\varepsilon^2 h_l(\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^4])^{1/2}$$

$$\times\left(\mathbb{E}\left[\left|\sum_{j=0}^{k-1} g(Y_{h_l}^{\varepsilon,i,M}(t_n^j-\tau), Y_{h_l}^{\varepsilon,i,M}(t_n^j-2\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j\right|^4\right]\right)^{1/2}$$

$$\leq \varepsilon^2 CN^3h_l^3 + \varepsilon^4 CN^2h_l^2, \tag{6.1.19}$$

where Lemma 6.1.5 is used in the last inequality. Therefore, applying the Cauchy-Schwartz inequality first and the elementary inequality $2ab \leq a^2+b^2$ later yields

$$I_{5A} \leq 2h_l\sum_{k=0}^{N-1}\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)||A_k|] + 2h_l\sum_{k=0}^{N-1}\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)||E_k|]$$

$$\leq 2h_l\sum_{k=0}^{N-1}\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + h_l\sum_{k=0}^{N-1}\mathbb{E}[|A_k|^2] + h_l\sum_{k=0}^{N-1}\mathbb{E}[|E_k|^2]$$

$$\leq 2h_lN\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + h_l^3N^3C + CN^4h_l^4\varepsilon^2 + CN^3h_l^3\varepsilon^4$$

Similarly, using Lemma 6.1.7 one can see that

$$I_{5B} \leq 2h_l\sum_{k=0}^{N-1}\mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \bar{A}_k\rangle] + 2h_l\sum_{k=0}^{N-1}\mathbb{E}[\langle Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), \bar{E}_k\rangle]$$

Also, we have $\mathbb{E}[|\bar{A}_k|^2] \leq h_l^2N^2C$ and

$$\mathbb{E}[|\bar{E}_k|^2] \leq \varepsilon^2CN^3h_l^3 + \varepsilon^4CN^2h_l^2. \tag{6.1.20}$$

Thus,

$$I_{5B} \leq 2h_l\sum_{k=0}^{N-1}\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|\mathbb{E}[|\bar{A}_k|]]$$

$$+ 2h_l \sum_{k=0}^{N-1} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|\mathbb{E}[|\bar{E}_k|]]$$

$$\leq 2h_l \sum_{k=0}^{N-1} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + h_l \sum_{k=0}^{N-1} \mathbb{E}[|\bar{A}_k|^2] + h_l \sum_{k=0}^{N-1} \mathbb{E}[|\bar{E}_k|^2]$$

$$\leq 2h_l N\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + h_l^3 N^3 C + CN^4 h_l^4 \varepsilon^2 + CN^3 h_l^3 \varepsilon^4$$

Additionally, we have

$$\begin{aligned}
I_6 &\leq h_l N\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + h_l N\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 \\
&+ h_l N\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n - \tau) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n - \tau)|^2 + h_l N\mathbb{W}_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}, \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n^k,M}) \\
&\leq 3h_l N\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2 + h_l N\mathbb{E}|Y_{h_l}^{\varepsilon,i,M}(t_n - \tau) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n - \tau)|^2.
\end{aligned}$$

Substituting the bounds for the terms I_1 to I_6 into Equation (6.1.17) yields that for $n \leq N^{l-1} - 1$

$$\begin{aligned}
&\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})|^2] \leq \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] \\
&+ \hat{C}\left(\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] + \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n - \tau) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n - \tau)|^2]\right) \\
&+ CN^3 h_l^3 + CN^2 h_l^2 \varepsilon^4,
\end{aligned}$$

which implies that that for all $0 \leq n_0 \leq N^{l-1} - 1$

$$\begin{aligned}
\sup_{0 \leq n \leq n_0} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})|^2] &\leq \hat{C} \sum_{k=1}^{n_0} \sup_{0 \leq n \leq k} \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] \\
&+ CN^2 h_l^2 + CNh_l \varepsilon^4.
\end{aligned}$$

An application of the discrete Gronwall inequality yields the result. □

6.1.2 Variance estimate of two coupled paths of the MLMC-EM scheme

In this section we provide an estimate for the variance of two coupled paths which is the main result of the paper and will be presented in Theorem 6.1.10.

The following two lemmas that will be needed to prove Theorem 6.1.10.

Lemma 6.1.8. *Assume that $\gamma : \mathbb{R}^d \to \mathbb{R}$ satisfies the Lipschitz condition, i.e. for all $x, y \in \mathbb{R}^d$ there exists a positive constant L, such that $|\gamma(x) - \gamma(y)|^2 \leq L|x - y|^2$. Then for $s \in [0,1]$ one has*

$$\max_{\substack{0\leq n\leq N^{l-1}\\ 1\leq k\leq N}} \mathrm{Var}[\gamma\{s(Y_{h_{l_2}}^{\varepsilon,i,M}(t_n^k), Y_{h_{l_2}}^{\varepsilon,i,M}(t_n^k - \tau)) + (1-s)(Y_{h_{l_1}}^{\varepsilon,i,M}(t_n), Y_{h_{l_1}}^{\varepsilon,i,M}(t_n - \tau)))\}] \leq C\varepsilon^2.$$

Proof. The proof is similar to that of Lemma 5.5.2, we omit here. □

Lemma 6.1.9. *Let Assumption 6.1.1 hold. Then there exists a positive constant C such that*

$$\max_{\substack{0\leq n\leq N^{l-1}\\ 1\leq k\leq N}} |\mathbb{E}[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]| \leq CNh_l.$$

Proof. From (6.1.10) we have that

$$\begin{aligned}
&|\mathbb{E}[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]| \\
&\qquad = \Bigg| \sum_{j=0}^{k-1} \mathbb{E}[f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})]h_l \\
&\qquad + \varepsilon\sqrt{h_l} \sum_{j=0}^{k-1} \mathbb{E}[g(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j]\Bigg|.
\end{aligned}$$

By independence the second summand of RHS in above is zero. Thus using Jensen's inequality and Remark 6.1.1 yields

$$\begin{aligned}
|\mathbb{E}[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]| &\leq \sum_{j=0}^{k-1} \mathbb{E}[|f(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})|]h_l \\
&\leq h_l \sum_{j=0}^{k-1} \mathbb{E}[\sqrt{\beta}\big(1 + |Y_{h_l}^{\varepsilon,i,M}(t_n^j)|^2 + |Y_{h_l}^{\varepsilon,i,M}(t_n^j - \tau)|^2 + W_2^2(\mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\big)^{1/2}] \\
&\leq \sqrt{\beta}h_l \sum_{j=0}^{k-1} \Big(1 + 2\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^j)|^2] + \mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^j - \tau)|^2]\Big)^{1/2}.
\end{aligned}$$

An application of Lemma 6.1.2 and the fact that $k \leq N$, completes the proof. □

Now, we can formulate the main result of the paper.

Theorem 6.1.10. *Let Assumption 6.1.1 hold, assume that* $\Psi : \mathbb{R}^d \to \mathbb{R}$ *has continuous second order derivative and there exists a constant* C *such that*

$$\left|\frac{\partial \Psi}{\partial x_i}\right| \leq C \quad \text{and} \quad \left|\frac{\partial^2 \Psi}{\partial x_i \partial x_j}\right| \leq C$$

for any $i, j = 1, 2, \cdots, a$*. Then, we have*

$$\max_{0 \leq n < M^{l-1}} \mathrm{Var}(\Psi(Y_{h_l}^{\varepsilon,i,M}(t_{n+1})) - \Psi(Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})) \leq C\varepsilon^2 h_{l-1}^2 + C\varepsilon^4 h_{l-1}.$$

Proof. From (6.1.11) and (6.1.12) we have that for $n \leq N^{l-1} - 1$

$$\begin{aligned}
&[Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j = [Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j \\
&+ h_l \sum_{k=0}^{N-1} \Big(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \Big) \\
&+ h_l \sum_{k=0}^{N-1} \Big(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M}) \Big) \\
&+ \varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \Big(g_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
&\qquad\qquad - g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \Big) \Delta\xi_n^k \\
&+ \varepsilon\sqrt{h_l} \sum_{k=0}^{N-1} \Big(g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \\
&\qquad\qquad - g_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M}) \Big) \Delta\xi_n^k,
\end{aligned}$$

where f_j is the jth component of f and g_j is the jth row of g. Taking variances on both sides of the previous inequality and using (B.0.1), (B.0.3) and (B.0.4) from the Appendix, we obtain

$$\begin{aligned}
&\mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j) \leq (1 + Nh_l)\mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j) \\
&+ 4h_l^2 N \sum_{k=0}^{N-1} \mathrm{Var}\Big(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
&\qquad\qquad - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \Big) \\
&+ (4Nh_l + 1)Nh_l \mathrm{Var}\Big(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})
\end{aligned}$$

$$
\begin{aligned}
&\qquad\qquad\qquad\qquad - f_j(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_{l-1}})\Big) \\
&+ 4\varepsilon^2 h_l \sum_{k=0}^{N-1} \mathrm{Var}\Big(g_j(Y^{\varepsilon,i,M}_{h_l}(t_n^k), Y^{\varepsilon,i,M}_{h_l}(t_n^k-\tau), \mathcal{L}^{\varepsilon,Y_n^k,M}_{h_l}) \\
&\qquad\qquad\qquad\qquad - g_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l})\Big)\Delta\xi_n^k \\
&+ 4\varepsilon^2 h_l \sum_{k=0}^{N-1} \mathrm{Var}\Big(g_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \\
&\qquad\qquad\qquad\qquad - g_j(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_{l-1}})\Big)\Delta\xi_n^k \\
&+ 2\mathrm{Cov}\Big([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_j, \\
&\quad h_l \sum_{k=0}^{N-1} f_j(Y^{\varepsilon,i,M}_{h_l}(t_n^k), Y^{\varepsilon,i,M}_{h_l}(t_n^k-\tau), \mathcal{L}^{\varepsilon,Y_n^k,M}_{h_l}) - f_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l})\Big) \\
&=: I_1 + I_2 + I_3 + I_4 + I_5 + I_6.
\end{aligned}
$$

In order to complete the proof of the theorem, we give estimates for $I_i, i = 2, ..., 6$, which will be shown in the following lemmas.

Lemma 6.1.11. *There exists a positive constant C such that*

$$
I_2 \le CN^3 h_l^3 \varepsilon^2.
$$

Proof. Using the fact that for two random variables X, Y, $\mathrm{Var}(X+Y) \le 2\mathrm{Var}(X) + 2\mathrm{Var}(Y)$, we have that

$$
\begin{aligned}
&\mathrm{Var}(f_j(Y^{\varepsilon,i,M}_{h_l}(t_n^k), Y^{\varepsilon,i,M}_{h_l}(t_n^k-\tau), \mathcal{L}^{\varepsilon,Y_n^k,M}_{h_l}) - f_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l})) \\
&\le 2\mathrm{Var}(f_j(Y^{\varepsilon,i,M}_{h_l}(t_n^k), Y^{\varepsilon,i,M}_{h_l}(t_n^k-\tau), \mathcal{L}^{\varepsilon,Y_n^k,M}_{h_l}) - f_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n^k,M}_{h_l})) \\
&+ 2\mathrm{Var}(f_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n^k,M}_{h_l}) - f_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l})) \\
&=: I_{2A} + I_{2B}.
\end{aligned}
$$

First we estimate I_{2A}. By the mean value theorem there exists an $s \in [0,1]$ such that

$$
\begin{aligned}
&f_j(Y^{\varepsilon,i,M}_{h_l}(t_n^k), Y^{\varepsilon,i,M}_{h_l}(t_n^k-\tau), \mathcal{L}^{\varepsilon,Y_n^k,M}_{h_l}) - f_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n^k,M}_{h_l}) \\
&= \big\langle \nabla f_j(s(Y^{\varepsilon,i,M}_{h_l}(t_n^k), Y^{\varepsilon,i,M}_{h_l}(t_n^k-\tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau)), \mathcal{L}^{\varepsilon,Y_n^k,M}_{h_l}),
\end{aligned}
$$

$$(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle.$$

Let $\nabla_q f_j(s(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau)) + (1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $[(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))]_q$ be the q component of $\nabla f_j(s(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau)) + (1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))$ respectively. We want to apply Lemma 5.5.1 with $A^{\varepsilon,h} = \nabla_q f_j(s(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau)) + (1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $B^{\varepsilon,h} = [(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))]_q$ so we check that the three conditions are satisfied. By Assumption 6.1.1, the function $\nabla_q^2 f_j$ is bounded, so $\nabla_q f_j$ is Lipschitz on the first and second arguments. Applying Lemma 6.1.8 with $\gamma = \nabla_q f_j(\cdot,\cdot,\mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $h_{l_1} = h_{l_2} = h_l$, we obtain

$$\mathrm{Var}(\nabla_q f_j(s(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau)) + (1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})) \leq C_1 \varepsilon^2, \tag{6.1.21}$$

so the first condition of Lemma 5.5.1 is satisfied. Conditions 2 and 3 are satisfied by Assumption 6.1.1 and Lemma 6.1.9 respectively. Thus by Lemma 5.5.1 we have that

$$\begin{aligned}
&\mathrm{Var}(\nabla_q f_j(s(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau)) + (1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
&\quad \times [(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))]_q) \\
&\quad \leq 3C_3^2 C_1 N^2 h_l^2 \varepsilon^2 + 15 C_2^2 \mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))]_q).
\end{aligned}$$

In order to estimate $\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))]_q)$ we use Equation (6.1.10) to obtain

$$\begin{aligned}
&\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))]_q) \\
&\quad \leq 2\mathrm{Var}(\sum_{j=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) h_l) \\
&\quad + 2\mathrm{Var}(\varepsilon\sqrt{h_l} \sum_{j=0}^{k-1} g_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) \Delta\xi_n^j).
\end{aligned}$$

By Asumption 6.1.1 and Lemma 6.1.3 we have that

$$\mathrm{Var}(\sum_{j=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) h_l)$$

$$= \mathrm{Var}(h_l \sum_{j=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) - f_q(z_h(t_n^j), z_h(t_n^j-\tau), \delta_{z_h(t_n^j)}))$$

$$\leq h_l^2 \mathbb{E}[|(\sum_{j=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M}) - f_q(z_h(t_n^j), z_h(t_n^j-\tau), \delta_{z_h(t_n^j)}))|^2]$$

$$\leq CN^2 h_l^2 \varepsilon^2.$$

From (6.1.16) we have that

$$\mathrm{Var}(\varepsilon\sqrt{h_l}\sum_{j=0}^{k-1} g_q(Y_{h_l}^{\varepsilon,i,M}(t_n^j), Y_{h_l}^{\varepsilon,i,M}(t_n^j-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^j,M})\Delta\xi_n^j) \leq CNh_l\varepsilon^2.$$

Thus

$$\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))]_q) \leq CN^2h_l^2\varepsilon^2 + CNh_l\varepsilon^2.$$

Using the formula $\mathrm{Var}(\sum_{i=1}^d X_i) \leq d\sum_{i=1}^d \mathrm{Var}(X_i)$ with $i = q, X_i = [Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q$ yields

$$\mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))]_q) \leq d^2CN^2h_l^2\varepsilon^2 + d^2CNh_l\varepsilon^2 \leq CNh_l\varepsilon^2.$$

Thus,

$$I_{2A} \leq CNh_l\varepsilon^2.$$

Next, we estimate I_{2B}. By Equation (2.2.8) there exists a random variable $s : \Omega \to [0,1]$ such that

$$\begin{aligned}
&f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) - f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \\
&\quad = \mathbb{E}[\langle \partial_\mu f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s), (Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n))\rangle]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)}.
\end{aligned}$$

where $Y_n^s := s(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k-\tau)) + (1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau))$.
Let $\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q$ be the q-components of $\partial_\mu f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)$ respectively. Then

$$\mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)})$$

$$= \mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)}$$

$$
\begin{aligned}
&- \mathbb{E}[\partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n-\tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]) \\
&= \mathrm{Var}(\mathbb{E}[(\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n-\tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))) \\
&\times [Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n)}) \\
&\leq \mathbb{E}[(\mathbb{E}[(\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n-\tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))) \\
&\times [Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)})^2] \\
&\leq \mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) \\
&- \partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n-\tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))|^2]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)} \\
&\times \mathbb{E}[|[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q|^2]],
\end{aligned}
$$

where we have use the Cauchy-Schwarz inequality in the penultimate step. By condition (6.1.4) and Lemma 6.1.3

$$
\begin{aligned}
&\mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) \\
&\qquad - \partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n-\tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))|^2]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)} \\
&\qquad \leq C\varepsilon^2.
\end{aligned}
$$

and by Lemma 6.1.5

$$
\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)|^2] \leq CN^2h_l^2 + CNh_l\varepsilon^2.
$$

Therefore

$$
I_{2B} \leq CN^2 h_l^2 \varepsilon^2 + CNh_l \varepsilon^4,
$$

and the proof is complete. □

Lemma 6.1.12. *There exists positive constants C and $\bar{C}$ such that*

$$
I_3 \leq CNh_l \sum_{q=1}^{d} \mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))]_q + CN^3 h_l^3 \varepsilon^2.
$$

Proof. Note that

$$
\mathrm{Var}(f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M}))
$$

$$\leq 2\mathrm{Var}(f_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) - f_j(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}))$$
$$+ 2\mathrm{Var}(f_j(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) - f_j(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_{l-1}}))$$
$$=: I_{3A} + I_{3B}.$$

First, we estimate I_{3B}. By the mean value theorem there exists an $s \in [0,1]$ such that

$$\begin{aligned} &f_j(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) - f_j(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \\ &\quad = \langle \nabla f_j(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau)), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}), \\ &\qquad\qquad (Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))\rangle. \end{aligned}$$

Let $\nabla_q f_j(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau)), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l})$ and $[(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q$ be the q components of $\nabla f_j(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau)), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l})$ and $(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))$ respectively. We want to apply Lemma 5.5.1 with
$A^{\varepsilon,h} = \nabla_q f_j(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau)), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l})$ and $B^{\varepsilon,h} = [(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q$ so we check that the three conditions are satisfied. Applying Lemma 6.1.8 with $\gamma = \nabla_q f_j, k = 0, h_{l_1} = h_{l-1}$ and $h_{l_2} = h_l$, we obtain

$$\mathrm{Var}(\nabla_q f_j(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau)), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \leq C_1\varepsilon^2,$$

so the first condition of Lemma 5.5.1 is satisfied. Conditions 2 and 3 are satisfied by Assumption 6.1.1 and Lemma 6.1.9 respectively. Thus by Lemma 5.5.1 we have that

$$\begin{aligned} &\mathrm{Var}(\nabla_q f_j(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n-\tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n-\tau)), \mathcal{L}^{\varepsilon,Y_n,M}_{h_l}) \\ &\qquad \times [(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q) \\ &\qquad \leq 3C_3^2 C_1 N^2 h_l^2 \varepsilon^2 + 15 C_2^2 \mathrm{Var}([(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q). \end{aligned}$$

Using the formula $\mathrm{Var}(\sum_{i=1}^d X_i) \leq d\sum_{i=1}^d \mathrm{Var}(X_i)$ with $i = q, X_i = [Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_q$ yields

$$\mathrm{Var}((Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))) \leq C\sum_{q=1}^d \mathrm{Var}([(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q + CN^2 h_l^2 \varepsilon^2.$$

Therefore,

$$I_{3A} \le CN^2h_l^2\varepsilon^2 + C\sum_{q=1}^{d} \mathrm{Var}([(Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q).$$

Next we estimate I_{3B}. By Equation (2.2.8) there exists a random variable $s : \Omega \to [0,1]$ such that

$$\begin{aligned}
&f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) - f_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})\\
&\quad = \mathbb{E}[\langle \partial_\mu f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s), (Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n))\rangle]_{Z_1=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau)}.
\end{aligned}$$

where $Y_n^s := s(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)) + (1-s)(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau))$.
Let $\partial_{\mu,q} f_j(Z_1, Z_2\mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $[Y_{h_l}^{\varepsilon,i,M}(t^n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q$ be the q-components of $\partial_\mu f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$ and $Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)$ respectively. Then

$$\begin{aligned}
&\mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q]_{Z_1=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau)})\\
&= \mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)[Y_{h_l}^{\varepsilon,i,M}(t_n^k) - Y_{h_l}^{\varepsilon,i,M}(t_n)]_q]_{Z_1=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau)}\\
&- \mathbb{E}[\partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), z_{h_{l-1}}(t_n-\tau), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n))[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q])\\
&= \mathrm{Var}(\mathbb{E}[(\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), z_{h_{l-1}}(t_n-\tau), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n)))\\
&\times [Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q]]_{Z_1=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau)})\\
&\le \mathbb{E}[(\mathbb{E}[(\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), z_{h_{l-1}}(t_n-\tau), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n)))\\
&\times [Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q]]_{Z_1=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau)})^2]\\
&\le \mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)\\
&- \partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), z_{h_{l-1}}(t_n-\tau), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n))|^2]_{Z_1=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n-\tau)}\\
&\times \mathbb{E}[|[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_q|^2]],
\end{aligned}$$

where we have use the Cauchy-Schwarz inequality in the penultimate step. By condition (6.1.4) and Lemma 6.1.3

$$\begin{aligned}
&\mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)\\
&\qquad - \partial_{\mu,q} f_j(z_{h_{l-1}}(t_n), z_{h_{l-1}}(t_n-\tau), \delta_{z_{h_{l-1}}(t_n)})(z_{h_{l-1}}(t_n))|^2]_{Z_1=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)}\\
&\qquad \le C\varepsilon^2
\end{aligned}$$

and by Theorem 6.1.4

$$\mathbb{E}[|Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)|^2] \leq CN^2h_l^2 + C\varepsilon^4 Nh_l.$$

Therefore,

$$I_{3B} \leq CN^2h_l^2\varepsilon^2 + C\varepsilon^6 Nh_l,$$

and the proof is complete.

□

Lemma 6.1.13. *There exists a positive constant C such that*

$$I_4 \leq C\varepsilon^2 h_{l-1}^3 + C\varepsilon^4 h_{l-1}^2.$$

Proof. By Lemma 6.1.5 and Assumption 6.1.1 one can see that

$$\begin{aligned} I_4 \leq 4\varepsilon^2 h_l \sum_{k=0}^{N-1} \mathbb{E}[|g_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\ - g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})|^2] \\ \leq 8\varepsilon^2 h_l NK(Ch_{l-1}^2 + C\varepsilon^2 h_{l-1}) = C\varepsilon^2 h_{l-1}^3 + C\varepsilon^4 h_{l-1}^2. \end{aligned}$$

□

Lemma 6.1.14. *There exists a positive constant C such that*

$$I_5 \leq C\varepsilon^2 h_{l-1}^3 + C\varepsilon^6 h_{l-1}^2.$$

Proof. By Assumption 6.1.1 and Theorem 6.1.4 we have that

$$\begin{aligned} I_5 \leq 4\varepsilon^2 h_l \sum_{k=0}^{N-1} \mathbb{E}[|g_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M}) \\ - g_j(Y_{h_{l-1}}^{\varepsilon,i,M}(t_n), Y_{h_{l-1}}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_{l-1}}^{\varepsilon,Y_n,M})|^2] \\ \leq 4\varepsilon^2 h_l NK(Ch_{l-1}^2 + C\varepsilon^4 h_{l-1}) \\ = C\varepsilon^2 h_{l-1}^3 + C\varepsilon^6 h_{l-1}^2. \end{aligned}$$

□

Lemma 6.1.15. *There exists a positive constant C such that*

$$I_6 \leq 2Nh_l \mathrm{Var}\big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j\big) + CN^3h_l^3\varepsilon^2.$$

Proof. Since the covariance is a linear function, we have that

$$\begin{aligned}
I_6 &= 2\mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, h_l \sum_{k=0}^{N-1}[f_j(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
&- f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})]\Big) \\
&+ 2\mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, h_l \sum_{k=0}^{N-1}[f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}) \\
&- f_j(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n-\tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n,M})]\Big) \\
&=: I_{6A} + I_{6B}.
\end{aligned}$$

By Lemma 6.1.6, we obtain

$$I_{6A} = 2\mathrm{Cov}\left([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, h_l \sum_{k=0}^{N-1}(A_k^j + B_k^j + E_k^j)\right)$$

Using property (B.0.4) from the appendix we have

$$\begin{aligned}
I_{6A} &= 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, A_k^j\Big) \\
&+ 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, B_k^j\Big) \\
&+ 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, E_k^j\Big).
\end{aligned}$$

Using the definition of covariance and since the increments ξ_n^j in B_k^j are independent, we find that

$$\begin{aligned}
&\mathrm{Cov}\Big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j, B_k^j\Big) \\
&\qquad = \mathbb{E}[[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j B_k^j] - \mathbb{E}[[Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j]\mathbb{E}[B_k^j] = 0.
\end{aligned}$$

Then using (B.0.3) yields

$$I_{6A} \leq 2Nh_l \mathrm{Var}\big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j\big) + h_l \sum_{k=0}^{N-1} \mathrm{Var}(A_k^j) + h_l \sum_{k=0}^{N-1} \mathrm{Var}(E_k^j). \quad (6.1.22)$$

Recall from Lemma 6.1.6 that

$$A_k^j = \langle \nabla f_j(s(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau)) + (1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M}),$$
$$h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})\rangle.$$

In order to estimate $\mathrm{Var}(A_k^j)$ we use Lemma 5.5.1 with $A^{\varepsilon,h} = \nabla_q f_j(s(Y_{h_l}^{\varepsilon,i,M}(t_n^k), Y_{h_l}^{\varepsilon,i,M}(t_n^k - \tau)) + (1-s)(Y_{h_l}^{\varepsilon,i,M}(t_n), Y_{h_l}^{\varepsilon,i,M}(t_n - \tau)), \mathcal{L}_{h_l}^{\varepsilon,Y_n^k,M})$ and $B^{\varepsilon,h} = [h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_q$ so we check that the three conditions are satisfied. The first and second conditions are satisfied by (6.1.21) and Assumption 6.1.1 respectively. By Lemma 6.1.2 and Assumption 6.1.1 we have that

$$|\mathbb{E}[[h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_q]| \leq CNh_l,$$

so the third condition is also satisfied. Thus Lemma 5.5.1 implies that

$$\mathrm{Var}(A_k^j) \leq CN^2h^2\varepsilon^2 + C\mathrm{Var}([h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_q).$$

Lemma 6.1.3 yields

$$\begin{aligned}
&\mathrm{Var}([h_l \sum_{r=0}^{k-1} f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_q) \\
&= \mathrm{Var}([h_l \sum_{r=0}^{k-1} \{f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M}) - f(z_{h_l}(t_n^r), z_{h_l}(t_n^r - \tau), \delta_{z_{h_l}(t_n^r)})\}]_q) \\
&\leq \mathbb{E}[|([h_l \sum_{r=0}^{k-1} \{f(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M}) - f(z_{h_l}(t_n^r), z_{h_l}(t_n^r - \tau), \delta_{z_{h_l}(t_n^r)})\}]_q)|^2] \\
&\leq CN^2h_l^2\varepsilon^2.
\end{aligned}$$

Therefore

$$\mathrm{Var}(A_k^j) \leq CN^2h^2\varepsilon^2 + CN^2h_l^2\varepsilon^2. \tag{6.1.23}$$

From (6.1.19) we have

$$\mathrm{Var}(E_k^j) \leq \mathbb{E}[|E_k^j|^2] \leq CN^3h_l^3\varepsilon^2 + CN^2h_l^2\varepsilon^4. \tag{6.1.24}$$

Substituting (6.1.23) and (6.1.24) into (6.1.22) we obtain

$$I_{6A} \leq 2Nh_l \mathrm{Var}\big([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_j\big) + CN^3 h_l^3 \varepsilon^2.$$

Using Lemma 6.1.7, (B.0.3) and (B.0.4), yields

$$\begin{aligned}
I_{6B} &= 2\mathrm{Cov}\left([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_j, h_l \sum_{k=0}^{N-1} (\bar{A}^j_k + \bar{E}^j_k)\right) \\
&\leq 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_j, \bar{A}^j_k\Big) \\
&+ 2h_l \sum_{k=0}^{N-1} \mathrm{Cov}\Big([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_j, \bar{E}^j_k\Big) \\
&\leq 2Nh_l \mathrm{Var}\big([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_j\big) + h_l \sum_{k=0}^{N-1} \mathrm{Var}(\bar{A}^j_k) + h_l \sum_{k=0}^{N-1} \mathrm{Var}(\bar{E}^j_k).
\end{aligned}$$

Recall from Lemma 6.1.7 that

$$\begin{aligned}
\bar{A}^j_k &= \mathbb{E}[\langle \partial_\mu f_j(Z_1, Z_2, \mathcal{L}^{\varepsilon, Y^s_n, M}_{h_l})(Y^s_n), \\
&\quad h_l \sum_{r=0}^{k-1} f(Y^{\varepsilon,i,M}_{h_l}(t^r_n), Y^{\varepsilon,i,M}_{h_l}(t^r_n - \tau), \mathcal{L}^{\varepsilon, Y^r_n, M}_{h_l}))\rangle]_{Z_1 = Y^{\varepsilon,i,M}_{h_l}(t_n), Z_2 = Y^{\varepsilon,i,M}_{h_l}(t_n - \tau)}.
\end{aligned}$$

Let $\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}^{\varepsilon, Y^s_n, M}_{h_l})(Y^s_n)$ and $f_q(Y^{\varepsilon,i,M}_{h_l}(t^r_n), Y^{\varepsilon,i,M}_{h_l}(t^r_n - \tau), \mathcal{L}^{\varepsilon, Y^r_n, M}_{h_l})$ be the the q-components of
$\partial_\mu f_j(Z_1, Z_2, \mathcal{L}^{\varepsilon, Y^s_n, M}_{h_l})(Y^s_n)$ and $f(Y^{\varepsilon,i,M}_{h_l}(t^r_n), Y^{\varepsilon,i,M}_{h_l}(t^r_n - \tau), \mathcal{L}^{\varepsilon, Y^r_n, M}_{h_l})$ respectively. Then

$$\begin{aligned}
&\mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}^{\varepsilon, Y^s_n, M}_{h_l})(Y^s_n) \\
&\times h_l \sum_{r=0}^{k-1} f_q(Y^{\varepsilon,i,M}_{h_l}(t^r_n), Y^{\varepsilon,i,M}_{h_l}(t^r_n - \tau), \mathcal{L}^{\varepsilon, Y^r_n, M}_{h_l})]_{Z_1 = Y^{\varepsilon,i,M}_{h_l}(t_n), Z_2 = Y^{\varepsilon,i,M}_{h_l}(t_n - \tau)}\big) \\
&= \mathrm{Var}(\mathbb{E}[\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}^{\varepsilon, Y^s_n, M}_{h_l})(Y^s_n) \\
&\times h_l \sum_{r=0}^{k-1} f_q(Y^{\varepsilon,i,M}_{h_l}(t^r_n), Y^{\varepsilon,i,M}_{h_l}(t^r_n - \tau), \mathcal{L}^{\varepsilon, Y^r_n, M}_{h_l})]_{Z_1 = Y^{\varepsilon,i,M}_{h_l}(t_n), Z_2 = Y^{\varepsilon,i,M}_{h_l}(t_n - \tau)} \\
&- \mathbb{E}[\partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n - \tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n)) h_l \sum_{r=0}^{k-1} f_q(Y^{\varepsilon,i,M}_{h_l}(t^r_n), Y^{\varepsilon,i,M}_{h_l}(t^r_n - \tau), \mathcal{L}^{\varepsilon, Y^r_n, M}_{h_l})]) \\
&= \mathrm{Var}(\mathbb{E}[(\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}^{\varepsilon, Y^s_n, M}_{h_l})(Y^s_n) - \partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n - \tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n)))
\end{aligned}$$

$$\times h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)})$$

$$\leq \mathbb{E}[(\mathbb{E}[(\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s) - \partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n - \tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n)))$$

$$\times h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n)})^2]$$

$$\leq \mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$$

$$- \partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n - \tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))|^2]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n-\tau)}$$

$$\times \mathbb{E}[|h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})|^2]],$$

where we have used the Cauchy-Schwarz inequality in the last step. By condition (6.1.4) and Lemma 6.1.3

$$\mathbb{E}[\mathbb{E}[|\partial_{\mu,q} f_j(Z_1, Z_2, \mathcal{L}_{h_l}^{\varepsilon,Y_n^s,M})(Y_n^s)$$
$$- \partial_{\mu,q} f_j(z_{h_l}(t_n), z_{h_l}(t_n - \tau), \delta_{z_{h_l}(t_n)})(z_{h_l}(t_n))|^2]_{Z_1=Y_{h_l}^{\varepsilon,i,M}(t_n), Z_2=Y_{h_l}^{\varepsilon,i,M}(t_n)} \leq C\varepsilon^2$$

and by Lemma 6.1.2 and Remark 6.1.1

$$\mathbb{E}[|h_l \sum_{r=0}^{k-1} f_q(Y_{h_l}^{\varepsilon,i,M}(t_n^r), Y_{h_l}^{\varepsilon,i,M}(t_n^r - \tau), \mathcal{L}_{h_l}^{\varepsilon,Y_n^r,M})|^2]] \leq CN^2 h_l^2.$$

Thus,

$$\mathrm{Var}(\bar{A}_k^j) \leq CN^2 h_l^2 \varepsilon^2.$$

From (6.1.20) we have

$$\mathrm{Var}(\bar{E}_k^j) \leq \mathbb{E}[|\bar{E}_k^j|^2] \leq \bar{K}\varepsilon^2 C_1 N^3 h_l^3 + \bar{K}\varepsilon^4 CN^2 h_l^2.$$

Therefore,

$$I_{6B} \leq 2Nh_l \mathrm{Var}\big([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j\big) + CN^3 h_l^3 \varepsilon^2$$

and the proof is complete. □

Continuation of the proof of Theorem 6.1.10 By Lemmas 6.1.11-6.1.15, we have

$$\mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_{n+1}) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_{n+1})]_j) \leq \mathrm{Var}([Y_{h_l}^{\varepsilon,i,M}(t_n) - Y_{h_{l-1}}^{\varepsilon,i,M}(t_n)]_j)$$

$$+ CNh_l \sum_{q=1}^{d} \mathrm{Var}([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_q) + CN^3 h_l^3 \varepsilon^2 + CN^2 h_l^2 \varepsilon^4.$$

Taking the maximum in both sides yields that for $n \leq N^{l-1} - 1$

$$\max_{1\leq j\leq d} \mathrm{Var}([Y^{\varepsilon,i,M}_{h_l}(t_{n+1}) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_{n+1})]_j) = \max_{1\leq j\leq d} \mathrm{Var}([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_j)$$
$$+ CNh_l \max_{1\leq j\leq d} \mathrm{Var}([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_j) + CN^3 h_l^3 \varepsilon^2 + CN^2 h_l^2 \varepsilon^4.$$

An application of the Grownwall inequality produces

$$\max_{\substack{0\leq n\leq N^{l-1}\\ 1\leq j\leq N}} \mathrm{Var}([Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)]_j) \leq CN^2 h_l^2 \varepsilon^2 + CNh_l \varepsilon^4. \tag{6.1.25}$$

In order to estimate $\mathrm{Var}(\Psi(Y^{\varepsilon,i,M}_{h_l}(t_n)) - \Psi(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))$ we apply the mean value theorem, so there exists $s \in [0,1]$ such that

$$\Psi(Y^{\varepsilon,i,M}_{h_l}(t_n)) - \Psi(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n) = \nabla\Psi(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - \tau))$$
$$+ (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n - \tau))(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)).$$

We shall apply Lemma 5.5.1 with
$A^{\varepsilon,h} = \nabla_q\Psi(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - \tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n - \tau))$ and $B^{\varepsilon,h} = [(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q$. Applying Lemma 6.1.8 with $\gamma = \nabla_q\Psi, k = 0, h_{l_1} = h_{l-1}$ and $h_{l_2} = h_l$, we obtain

$$\mathrm{Var}(\nabla_q\Psi(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - \tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n - \tau)))) \leq C\varepsilon^2,$$

so the first condition of Lemma 5.5.1 is satisfied. Conditions 2 and 3 are satisfied by Assumption 6.1.1 and Lemma 6.1.9 respectively. Thus by Lemma 5.5.1 we have that

$$\mathrm{Var}(\nabla_q\Psi(s(Y^{\varepsilon,i,M}_{h_l}(t_n), Y^{\varepsilon,i,M}_{h_l}(t_n - \tau)) + (1-s)(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n), Y^{\varepsilon,i,M}_{h_{l-1}}(t_n - \tau)))$$
$$\times [(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q)$$
$$\leq CN^2 h_l^2 \varepsilon^2 + C\mathrm{Var}([(Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n))]_q).$$

Thus

$$\mathrm{Var}(\Psi(Y^{\varepsilon,i,M}_{h_l}(t_n)) - \Psi(Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)) \leq CN^2 h_l^2 \varepsilon^2 + C\mathrm{Var}((Y^{\varepsilon,i,M}_{h_l}(t_n) - Y^{\varepsilon,i,M}_{h_{l-1}}(t_n)). \tag{6.1.26}$$

Sustituting (6.1.25) into (6.1.26) we obtain the desire result.

□

Appendices

Appendix A

Some inequalities

Theorem A.0.1. *(Burkholder–Davis–Gundy's (BDG) inequality)*
Let $g \in \mathcal{L}^2(\mathbb{R}_+; \mathbb{R}^{d\times m})$*. Define for* $t \geq 0$*,*

$$x_t = \int_0^t g_s dW_s \quad \text{and} \quad A_t = \int_0^t |g_s|^2 ds.$$

Then for every $p > 0$*, there exist positive constants (depending only on* p*), such that*

$$c_p \mathbb{E}[|A_t|^{p/2}] \leq \mathbb{E}[\sup_{0\leq s\leq t} |x_s|^p] \leq C_p \mathbb{E}[|A_t|^{p/2}]$$

for all $t \geq 0$*.*

Theorem A.0.2. *(Gronwall's inequality)*
Let $T > 0$ *and* $C \geq 0$*. Let* u *be a Borel measurable bounded nonnegative function on* $[0, T]$*, and let* v *be a nonnegative integrable function on* $[0, T]$*. If*

$$u(t) \leq C + \int_0^t v(s)u(s)ds \quad \text{for all} \quad 0 \leq t \leq T,$$

then

$$u(t) \leq C \exp\left(\int_0^t v(s)ds\right) \quad \text{for all} \quad 0 \leq t \leq T.$$

Proposition A.0.3. *Let* $p > 1, \varepsilon > 0$ *and* $a, b \in \mathbb{R}$*. Then*

$$|a + b|^p \leq \left[1 + \varepsilon^{\frac{1}{p-1}}\right]^{p-1} \left(|a|^p + \frac{|b|^p}{\varepsilon}\right).$$

Appendix B

Basic variance and covariance properties

Here, we summary a few useful properties and inequalities regarding the variance and covariance functions. Let X, Y, V, W be random variables and $a, b, c, d \in \mathbb{R}$.

Definition B.0.1. *(Variance)*

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Definition B.0.2. *(Covariance)*

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

By definition of variance and covariance we have the following identity

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X, Y). \tag{B.0.1}$$

Lemma B.0.3.

$$\boxed{\mathrm{Var}\left(\sum_{i=1}^{m} X_i\right) \le m \sum_{i=1}^{m} \mathrm{Var}(X_i).} \tag{B.0.2}$$

Proof Using the definition of variance, linearity of expectation and Cauchy-Schwartz inequality, we obtain

$$\mathrm{Var}\left(\sum_{i=1}^{m} X_i\right) = \mathbb{E}\left[\left(\sum_{i=1}^{m} X_i - \mathbb{E}[\sum_{i=1}^{m} X_i]\right)^2\right] = \mathbb{E}\left[\left(\sum_{i=1}^{m} X_i - \mathbb{E}[X_i]\right)^2\right]$$

$$\leq \mathbb{E}\left[m\sum_{i=1}^{m}(X_i - \mathbb{E}[X_i])^2\right] = m\sum_{i=1}^{m}\mathrm{Var}(X_i).$$

□

Lemma B.0.4.

$$\boxed{\mathrm{Cov}(X,Y) \leq \frac{1}{2}\mathrm{Var}(X) + \frac{1}{2}\mathrm{Var}(Y).} \tag{B.0.3}$$

Proof Substituting (B.0.1) into (B.0.2) while setting m equal to 2.

□

Lemma B.0.5. *The covariance function is bilinear, i.e.*

$$\boxed{\mathrm{Cov}(aX+bY, cW+dV) = ac\mathrm{Cov}(X,W) + ad\mathrm{Cov}(X,V) + bc\mathrm{Cov}(Y,W) + bd\mathrm{Cov}(Y,V).} \tag{B.0.4}$$

Appendix C

Mean value theorem

Lemma C.0.1. *(Mean value theorem for a function of several variables)*

Let G be an open subset of $\mathbb{R}^n$ and let $g : G \to \mathbb{R}$ be a differentiable function. Then there exists a $t \in [0, 1]$ such that

$$g(y) - g(x) = \langle \nabla g((1-t)x + ty), (y - x) \rangle.$$

Appendix D

MATLAB code

MATLAB code that implements the simulation tests described in section 3.7.

```
%macro that simulates the numerical adaptive-EM solution of ...
    the SDDE (1)
clear

rng('default');
rng(1);
s_0 = 100;
T = 0.5;

dt = 0.000001;
nSims = 10000;

times = (0:dt:T);
times = [times T];
times = transpose(times);
numSaltos = length(times)-1;
dW = zeros(nSims,1);

S = s_0 + zeros(nSims,1);
%Mpaths = zeros(nSims,numSaltos);
```

```
    T_adap = 0;
    hs=zeros(nSims,1);
    values = zeros(nSims,1);
    Sprev = S;
    values(:,1) = S;

    tic

    while min(T_adap) < T

    Sprev = S;
    h_adap = 1/25 * (abs(Sprev)<1)...
    + 0.25*(abs(Sprev)≥1).*max(1,abs(Sprev))./...
    max(1,abs(-2*Sprev-Sprev.^3+0.5*Sprev*sin(100)));

    dW= sqrt(h_adap).*randn(nSims,1);
    S=Sprev+(-2*Sprev-Sprev.^3+0.5*Sprev*sin(100)).*...
    h_adap+sqrt(2)*cos(100)*Sprev.*dW;

    T_adap = T_adap + h_adap;

    hs = [hs h_adap];
    values = [values S];

    end

    toc
```

```
%macro that simulates the numerical (standard) EM solution of the ...
    SDDE (1)
clear

rng('default');
```

```
rng(1);
s_0 = 100;
T = 0.0022;

dt = 0.0002;
nSims = 10;

times = (0:dt:T);
times = [times T];
times = transpose(times);
numSaltos = length(times)-1;
dW = zeros(nSims,1);

S = s_0 + zeros(nSims,1);
values = zeros(nSims,1);
values(:,1) = S;
Sprev = S;

tic

for i=1:numSaltos

Sprev = S;

dW= sqrt(dt)*randn(nSims,1);
S=Sprev+(-2*Sprev-Sprev.^3+0.5*Sprev*sin(100))*dt+sqrt(2)*cos(100)*Sprev.*dW;

values = [values S];

end


toc

values(6:9,:)=[];
```

```
values(:,11:13)=[];
values = log(abs(values));
values = values';
x = (0.0:0.0002:0.0018);
plot(x,values)
```

MATLAB code that implements the simulation tests described in section 5.6.

```
clear

rng('default');
rng(1);
s_0 = 1;
T = 0.5;

epsilon = 2^-11;
h_l = 2^-7;
nSims = 2000;
numParticles = 50;

times = (0:h_l:T+h_l);
times = [times T+h_l];
times = transpose(times);
numSaltos = length(times)-1;
dW = zeros(nSims,numParticles);

S = s_0 + zeros(nSims,numParticles);
S_minus = s_0 + zeros(nSims,numParticles);


Sprev = S;
Sprev_minus = S_minus;
expected_Sprev = s_0;
expected_Sprev_minus = s_0;
```

```
tic

for i=1:numSaltos % numSaltos HAS TO BE an even number

expected_S = zeros(nSims,1);
expected_S_minus = zeros(nSims,1);

for j=1:numParticles

Sprev(:,j) = S(:,j);
dW(:,j) = sqrt(h_l)*randn(nSims,1);
S(:,j)=Sprev(:,j) + (-Sprev(:,j)-0.5*expected_Sprev)*h_l + ...
    epsilon*(Sprev(:,j)).*dW(:,j);


if mod(i,2) == 0
Sprev_minus(:,j) = S_minus(:,j);
S_minus(:,j)=Sprev_minus(:,j)+(-Sprev_minus(:,j)-0.5*expected_Sprev)*(h_l+h_l) ...
    + epsilon * (Sprev_minus(:,j)).*(dW(:,j) + dW_prev(:,j));
%values_minus = [values_minus S];
end
dW_prev(:,j) = dW(:,j);

expected_S = expected_S + S(:,j);
expected_S_minus = expected_S_minus + S_minus(:,j);

end
expected_Sprev = expected_S / numParticles;
expected_Sprev_minus = expected_S_minus / numParticles;


end
```

```
toc

diff = S - S_minus;
V = var(diff);
variance = mean(V);
```

```
h = zeros(6,0);
h(1) = 2^-13; h(2) = 2^-14; h(3) = 2^-15; h(4) = 2^-16;
h(5) = 2^-17; h(6) = 2^-18;
clear
var = zeros(6,0);
var(1) = 1.44924236792382E-10; var(2) = 7.00796703487318E-11; ...
    var(3) = 3.44170640135815E-11;
var(4) = 1.70202378303428E-11; var(5) = 8.47381928597234E-12; ...
    var(6) = 4.24656270332934E-12;

h = log(h);
var = log(var);

scatter(h,var,'black')

function res =f(x)

a = 1.01782587505298;
b = -13.4987407417417;

res = a*x + b;
end

hold on
fplot(@(x)f(x),[h(6),h(1)],'red');
hold on
xlabel('log(h_{l-1})')
```

```
ylabel('log(variance)')
```

```
clear
h = zeros(6,0);
h(1) = 2^-8; h(2) = 2^-9; h(3) = 2^-10; h(4) = 2^-11;
h(5) = 2^-12; h(6) = 2^-13;

var = zeros(6,0);
var(1) = 2.34889611913546E-12; var(2) = 5.87202344934886E-13; ...
    var(3) = 1.45746201473501E-13;
var(4) = 3.64562088338735E-14; var(5) = 9.08706974096245E-15; ...
    var(6) = 2.26812526205551E-15;

h = log(h);
var = log(var);
hold off
scatter(h,var,'black')

function res =f(x)

a = 2.003492929;
b = -15.6671214234075;

res = a*x + b;
end

hold on
fplot(@(x)f(x),[h(6),h(1)],'red');
hold on
xlabel('log(h_{l-1})')
ylabel('log(variance)')
```

```
clear
e = zeros(5,0);
e(1) = 2^-3; e(2) = 2^-4; e(3) = 2^-5; e(4) = 2^-6;
e(5) = 2^-7;

var = zeros(5,0);
var(1) = 6.82952513950179E-11; var(2) = 4.24656270332934E-12; ...
    var(3) = 2.66599756200494E-13;
var(4) = 1.70743520978924E-14; var(5) = 1.1727764029295E-15;

e = log(e);
var = log(var);
hold off
scatter(e,var,'black')

function res =f(x)

a = 3.96174602457879;
b = -15.1947788474293;

res = a*x + b;
end

hold on
fplot(@(x)f(x),[e(5),e(1)],'red');
hold on
xlabel('log(epsilon)')
ylabel('log(variance)')
```

```
clear
e = zeros(6,0);
e(1) = 2^-6; e(2) = 2^-7; e(3) = 2^-8; e(4) = 2^-9;
e(5) = 2^-10; e(6) = 2^-11;
```

```
var = zeros(6,0);
var(1) = 2.46028511616704E-09; var(2) = 6.08421725877772E-10; ...
   var(3) = 1.51689746473825E-10;
var(4) = 3.78964401334516E-11; var(5) = 9.47248315902495E-12; ...
   var(6) = 2.36801884749758E-12;

e = log(e);
var = log(var);
hold off
scatter(e,var,'black')

function res =f(x)

a = 2.00346260244457;
b = -11.496111583537;

res = a*x + b;
end

hold on
fplot(@(x)f(x),[e(6),e(1)],'red');
hold on
xlabel('log(epsilon)')
ylabel('log(variance)')
```

Bibliography

[1] Anderson D.F., Higham D.J., Sun Y., Multilevel Monte Carlo for stochastic differential equations with small noise. *SIAM J. Numer. Anal.*, 54, 505-529, 2016.

[2] D.F. Anderson, D.J. Higham, and Y.Sun, Complexity of multilevel Monte Carlo tau-leaping, SIAM J. Numer. Anal., 52 (2014), pp. 3106–3127.

[3] Baker, C.T.H., Buckwar, E., Exponential stability in p-th mean of solutions, and of convergent Euler- type solutions, of stochastic delay differential equations. J. Comput. Appl. Math. 184 (2005), 404-427.

[4] Baladron, J., Fasoli, D., Faugeras, O. and Touboul, J., Mean-field description and propagation of chaos in networks of Hodgkin-Huxley and FitzHugh-Nagumo neurons, The Journal of Mathematical Neuroscience, 2 (1) (2012), 10.

[5] Bolley, F., Canizo, J. A. and Carrillo, J. A. (2011) Stochastic mean-field limit: non-Lipschitz forces and swarming. Mathematical Models and Methods in Applied Sciences, 21, 2179–2210.

[6] Bossy, M. and Talay, D., A stochastic particle method for the McKean-Vlasov and the Burgers equation, Mathematics of Computation, 66 (217) (1997), 157-192

[7] Buckdahn, R., Li, J. and Ma, J., A mean-field stochastic control problem with partial observations, Annals of Applied Probability, 27 (5) (2017), 3201-3245.

[8] Cardaliaguet P., Notes on Mean Field Games (from P.-L. Lions' lectures at Collège de France), http://www.science.unitn.it/∼ bagagiol /Notes by Cardaliaguet.pdf.

[9] Carmona R, Delarue F., Probabilistic Theory of Mean Field Games with Applications I, Springer, 2018.

[10] Carmona R, Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Finantial Applications, Princeton University, 2016.

[11] Dos Reis, G., Engelhardt, S. and Smith, G., Simulation of McKean Vlasov SDEs with super linear growth, IMA Journal of Numerical Analysis, Volume 42, Issue 1, January 2022, Pages 874–922,

[12] Dos Reis G, Salkeld W, and Tugaut J., Freidlin-Wentzell LDPs in path space for McKean-Vlasov equations and the functional iterated logarithm law, Ann. Appl. Probab., Vol. 29(3), 2017.

[13] R. Erban and J. Haskovec, From individual to collective behaviour of coupled velocity jump processes: a locust example. Kinetic and Related Models 5, 4 (December 2012), 817-842.

[14] Fang, W., Giles, M.B. Adaptive Euler-Maruyama method for SDEs with nonglobally Lipschitz drift. Ann. Appl. Probab. 30 (2020), 526-560.

[15] Giles M.B., Multi-level Monte Carlo path simulation. *Oper. Res.*, 56(3), 607-617, 2008.

[16] Giles M.B., Improved multilevel Monte Carlo convergence using the Milstein scheme, Monte Carlo and Quasi-Monte Carlo Methods, Springer, 2006, pp. 343-358.

[17] Giles M.B., *Improved Multilevel Monte Carlo Convergence using the Milstein Scheme, Monte Carlo and Quasi-Monte Carlo Methods*. Springer, Berlin, 2008.

[18] Gomes, S., Pavliotis, G. and Vaes, U. (2020) Mean field limits for interacting diffusions with colored noise: phase transitions and spectral numerical methods, Multiscale Modeling and Simulation, 18, 1343–1370.

[19] Guhlke, C., Gajewski, P., Maurelli, M., Friz, P. K. and Dreyer, W., Stochastic many-particle model for LFP electrodes, Continuum Mechanics and Thermodynamics, 30 (3) (2018), 593-628.

[20] Guo Q., Liu W., Mao X. and Zhan, W., Multi-level Monte Carlo methods with the truncated Euler-Maruyama scheme for stochastic differential equation. *International Journal of Computer Mathematics*, 95 (9) (2018), 1715-1726.

[21] Guo, Q., Mao, X., Yue, R., The Truncated Euler-Maruyama Method for Stochastic Differential Delay Equations, Numer. Algorithms, 78 (2018), 599-624.

[22] Higham D.J., Mao X.R., Yuan C.G., Almost Sure and Moment Exponential Stability in the Numerical Simulation of Stochastic Differential Equations. *SIAM Numer. Anal.*, 45, 592-609, 2007.

[23] Higham, D.J., Mao, X., Stuart, A. M., Strong convergence of Euler-type methods for nonlinear stochastic differential equations. SIAM J. Numer. Anal. 40 (2002), 1041-1063.

[24] Higham, D.J., Mao, X., Yuan, C., Almost sure and moment exponential stability in the numerical simulation of stochastic differential equations. SIAM J. Numer. Anal. 45 (2002), 592-609.

[25] Higham, D.J., Mao, X., Yuan, C., Almost sure and Moment exponential stability in the numerical simulation of stochastic differential equations. SIAM J. Numer. Anal. 45 (2007), 592-607.

[26] Hutzenthaler M., Jentzen A., Kloeden P E., Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc R Soc Lond Ser A Math Phys Eng Sci*, 467, 1563-1576, 2011.

[27] Hutzenthaler M., Jentzen A., Kloeden P.E., Devergence of the multilevel monte carlo Euler method for nonlinear stochastic differential equations. *Annals Appl. Probab.*, 23, 1913-1966, 2013.

[28] Hutzenthaler, M., Jentzen, A., Non-globally Lipschitz Counterexamples for the stochastic Euler scheme, Proceedings of The Royal Society A: Mathematical, Physical and Engineering Sciences, 467.

[29] Ji, Y., Yuan, C., Tamed EM scheme of Neutral Stochastic Differential Delay Equations, J. Comput. Appl. Math. 326 (2017), 337-357.

[30] Karatzas, I., Shreve, S.E., Brownian Motion and Stochastic Calculus, Springer, 1988,

[31] Kloeden P.E., Platen E., Schurz H., The numerical solution of non-linear stochastic dynamical systems: A brief introduction. *Int. J. Bif. Chaos.*, 1, 277-286, 1991.

[32] Kloeden P. E., Platen E., *Numerical Solution of Stochastic Differential Equation.* Springer-Verlag, Berlin Heidelberg, 1992.

[33] Li X., Yi L., Yuan C., Explicit numerical approximations for McKean-Vlasov neutral differential delay equations, arXiv:2105.04175v1.

[34] Malrieu, F. (2003) Convergence to equilibrium for granular media equations and their Euler schemes. Ann. Appl. Probab., 13, 540–560.

[35] Mao X., The truncated Euler-Maruyama method for stochastic differential equations. *J Comput Appl Math*, 290, 370-383, 2015.

[36] Mao, X., Stochastic Differential Equations and Applications, Hoorwood Publishing, 2007.

[37] Mao, X., Exponential Stability of Stochastic Differential Equations, Marcel Dekker, New York, 1995

[38] Mao, X., A note on the LaSalle-type theorems for stochastic differential delay equations, J. Math. Anal. Appl. 268 (2002), 125-142.

[39] Mao, X, Exponential stability of equidistant Euler-Maruyama approximations of stochastic differential delay equations. J. Comput. Appl. Math. 200 (2007), 297-316.

[40] Mao, X., Sabanis, S., Numerical solutions of stochastic differential delay equations under local Lipschitz condition. J. Comput. Appl. Math. 151 (2003), 215-227.

[41] Mao, X, Yuan C., Stochastic Differential Equations with Markovian Switching, Imperial College Press, 2006.

[42] McKean, H. P., Propagation of chaos for a class of non-linear parabolic equations, In:Lecture Series in Differential Equations, 2 (1) (1967), 41-57.

[43] McKean, H. P., Fluctuations in the kinetic theory of gases, Communications on Pure and Applied Mathematics, 28 (4) (1975), 435-455.

[44] Milstein G.N., Tret'yakov M.V., Mean-square numerical methods for stochastic differential equations with small noises. *SIAM J. Sci. Comput.*, 18, 1067-1087, 1997.

[45] Platen E., Bruti-Liberati N., *Numerical Solution of Stochastic Differential Equations with Jumps in Finance*. Springer, 2011.

[46] Reisinger C., Stockinger W., An adaptive Euler-Maruyama scheme for McKean-Vlasov SDEs with super-linear growth and application to the mean-field FitzHugh-Nagumo model, Journal of Computational and Applied Mathematics, Volume 400, Jan 2022.

[47] Römisch W., Winkler R., Stepsize control for mean-square numerical methods for stochastic differential equations with small noise. *SIAM J. Sci. Comput.*, 28, 604-625, 2006.

[48] Sznitman, A.S., Topics in Propagation of Chaos, Ecole D'été de Probabilités de Saint-Flour XIX - 1989, in: Lect. Notes in Math., vol. 1464, Springer-Verlag, 1991.

[49] Wu, F., Mao, X., Szpruch, L., Almost sure exponential stability of numerical solutions for stochastic delay differential equations, Numerische Mathematik 115(2010), 681-697.

[50] Wang, F., Distribution Dependent SDEs for Navier-Stokes type equations, Electronic Communications in Volatility, vol. 27, 2022.

[51] Wang, F.,Distribution dependent SDEs for Landau type equations, Stochastic Processes and their Applications, Volume 128, Issue 2, February 2018, Pages 595-621.