Electronic Journal of Probability

Electron. J. Probab. **26** (2021), article no. 111, 1–32.
ISSN: 1083-6489 https://doi.org/10.1214/21-EJP676

On explicit Milstein-type scheme for McKean–Vlasov stochastic differential equations with super-linear drift coefficient*

Chaman Kumar[†] Neelima[‡]

Abstract

We introduce an explicit Milstein-type scheme for McKean–Vlasov stochastic differential equations using the notion of a measure derivative given by P.-L. Lions in his lectures at the *Collège de France* and presented in [9]. We show that the proposed Milstein-type scheme converges to the true solution in strong sense with a rate equal to 1.0. The drift coefficient is allowed to grow super-linearly in the state variable and both the drift and the diffusion coefficients are assumed to be only once differentiable in variables corresponding to state and measure. Furthermore, derivatives of drift and diffusion coefficients with respect to the measure component are uniformly bounded. The challenges arising due to the dependence of coefficients on measures are tackled and our findings are consistent with the analogous results for stochastic differential equations.

Keywords: McKean–Vlasov SDE; explicit Milstein scheme; super-linear coefficient; propagation of chaos; rate of strong convergence.
MSC2020 subject classifications: 65C30; 65C35; 65C05; 60H35.
Submitted to EJP on June 25, 2020, final version accepted on July 7, 2021.

1 Introduction

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{\{t\geq 0\}}, P)$ be a filtered probability space satisfying the usual conditions, i.e., the filtration is right continuous and complete. Assume that $\{W_t\}_{\{t\geq 0\}}$ is an m-dimensional Brownian motion adapted to the filtration $\{\mathcal{F}_t\}_{\{t\geq 0\}}$. For a fixed $T > 0$, consider the following d-dimensional McKean–Vlasov stochastic differential equation (MV–SDE) on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{\{t\geq 0\}}, P)$,

$$X_t = X_0 + \int_0^t b(X_s, \mu_s^X)ds + \int_0^t \sigma(X_s, \mu_s^X)dW_s \tag{1.1}$$

*The first named author gratefully acknowledges the financial support provided by Science and Engineering Research Board (SERB) under its MATRICS program through grant number SER-1329-MTD.

[†]Department of Mathematics, Indian Institute of Technology Roorkee, Uttarakhand, India, 247 667.
E-mail: chaman.kumar@ma.iitr.ac.in

[‡]Department of Mathematics, Ramjas College, University of Delhi, Delhi, India, 110 007.
E-mail: neelima_maths@ramjas.du.ac.in

almost surely for any $t \in [0,T]$, where μ_t^X denotes the law of the random variable X_t. If the law μ_t^X is known, then the coefficients b and σ are functions of time and space variables and hence the MV–SDE reduces to an SDE. The work on MV–SDEs was originated in [30] and since then such equations have found wide applications in physics, biology and neural activities, see for example, [3], [6], [7], [10], [11], [16], [19], and references therein. The existence and uniqueness of the solution of MV–SDE (1.1) in strong and weak sense has been thoroughly investigated in the literature, see for example, [10], [11], [13], [21], [31], [32], [35], [39] and references therein. In particular, authors in [35] have established the existence and uniqueness of the strong solution of MV–SDE (1.1) where drift grows super-linearly and diffusion satisfies global Lipschitz condition in the state component. In [31], authors prove the result under the monotonicity and coercivity conditions on the coefficients and thus both drift and diffusion coefficients can grow super-linearly in the state component. In [10], [11], [31] and [35], the drift and diffusion coefficients are assumed to satisfy global Lipschitz condition in the variable corresponding to the measure component.

Many researchers have shown their interest in the strong approximation of SDEs since it is needed in the Multilevel Monte Carlo path simulations for SDEs, see [18]. The numerical approximation of SDEs in strong sense is well understood in the literature for global and non-global Lipschitz coefficients, see for example [5], [15], [22], [24], [26], [27], [28], [37], [38], [41], [43] and references therein. The earliest work on the numerical approximation for the one-dimensional particle system can be found in [8] where the rate of strong convergence is shown to be equal to $1/2$, both for the time-step and for the number of particles under the global Lipschitz conditions on the coefficients in both state and measure components. The convergence in particles is popularly known as the *Propagation of Chaos* in the literature.

It is well known that the classical Euler scheme for SDEs with super-linear coefficients diverges in finite time, see for example [23] and hence such divergence can also be observed in the case of MV–SDEs. Further, in the case of MV–SDEs, a single particle may become influential over all other particles which can destroy the setting of the weakly interacting particle system connected with the MV–SDE. This phenomenon is called as the *particle corruption* in [34] and is tackled by proposing an explicit tamed Euler scheme and an implicit Euler scheme. The strong rate of convergence of the interacting particle system connected with the MV–SDE is shown to be equal to $1/2$ when the drift coefficient satisfies one-sided global Lipschitz condition (and hence is allowed to grow super-linearly) in the state variable while the diffusion coefficient satisfies global Lipschitz condition. Furthermore, both the coefficients are assumed to be global Lipschitz continuous in the measure component.

The main purpose of this article is to develop an explicit Milstein-type numerical scheme for the interacting particle system associated with the MV–SDE (1.1) and investigate its rate of convergence in strong sense. For this, we require more regularity assumptions on the coefficients as compared to [34] while the drift coefficient is still allowed to grow super-linearly in the state variable, the diffusion coefficient satisfies global Lipschitz continuity condition in the state variable and both the coefficients are assumed to be global Lipschitz continuous in the measure component. Mainly, the drift and diffusion coefficients are assumed to be once differentiable in variables corresponding to the state and measure components. Moreover, their derivatives with respect to the measure component are uniformly bounded. For differentiability with respect to measure, we use the concept of Lions' derivative, see [9] and [10] for details. The precise assumptions on the coefficients are listed in Section 2. Under these assumptions, we establish that the rate of strong convergence in $\mathcal{L}^2$-sense of the proposed tamed Milstein-type scheme for the interacting particle system connected with the MV–SDE

(1.1) is equal to 1.0 (see Theorem 2.7), which is consistent with the analogous result available in the literature for SDEs. Novel techniques have been developed to tackle challenges arising due to the presence of the law μ_t in the coefficients.

An independent work [4] on tamed Milstein scheme for MV–SDEs with drift having super-linear growth in the state variable appeared concurrently, but under stronger regularity assumptions on the coefficients. Authors in [4] use the Itô's formula in the proof of their main result and thus require the coefficients to be twice differentiable in both state and measure components. This manuscript, on the other hand, uses mean-value type results (see Lemma 4.2 and Lemma 4.3 below) which only require once differentiability of the coefficients in the state and measure components. Thus, our results are obtained under more relaxed conditions on the coefficients than those made in [4]. Furthermore, we use $b^n(x,\mu) = b(x,\mu)(1+n^{-1}|x|^{\rho+2})^{-1}$ (see equation (2.3) below) whereas authors in [4] use $b^n(x,\mu) = b(x,\mu)(1+n^{-1}|b(x)|)^{-1}$ to tame the drift coefficient $b(x,\mu)$, however either way, the tamed coefficient $b^n(x,\mu)$ converges to $b(x,\mu)$ point-wise with the same rate equal to 1. Though, our taming seems stronger, but authors in [4] require the existence of all moments of the initial random variable X_0 whereas we only require moments up to order p for some fixed $p > 4$.

Some recent developments on MV–SDEs with non-global Lipschitz coefficients can be found in [25], [29], [33] and references therein. In [25], well-posedness, propagation of chaos, tamed Euler scheme and tamed Milstein scheme for MV–SDE with common noise have been investigated when drift and diffusion coefficients grow super-linearly in the state variable. In [33], well-posedness, propagation of chaos and the tamed Euler scheme for MV–SDE driven by Lévy noise have been studied when the drift, diffusion and jump coefficients can grow super-linearly in the state variable. In [29], well-posedness and Euler scheme for one-dimensional MV–SDE with discontinuous drift coefficients have been discussed. Further, some more recent works on the numerical approximation of the MV–SDE are [14] and [36].

We now list some examples, by adding a mean-field term to the well-known SDEs, which fit in our framework.

1. Mean-field stochastic Ginzburg–Landau equation:
$$X_t = X_0 + \int_0^t \big(\alpha X_s - \beta X_s^3 + cE[X_s]\big)ds + \int_0^t \gamma X_s dW_s$$
almost surely for any $t \in [0,T]$, where $\alpha, \beta, \gamma \in (0,\infty)$ and $c \in \mathbb{R}$ are fixed constants. This equation has been investigated in [34] and its variant (without the mean-field term) in [22], [40] and references therein.
2. Mean-field stochastic Verhulst equation:
$$X_t = X_0 + \int_0^t \big(\alpha X_s - \beta X_s^2 + c_1 E[X_s]\big)ds + \int_0^t \big(\gamma X_s + c_2 E[X_s]\big) \circ dW_s$$
almost surely for any $t \in [0,T]$, where $\alpha, \beta \in (0,\infty)$ and $c, \gamma \in \mathbb{R}$ are fixed constants. The integral in the diffusion term should be understood in the Stratonovich sense. Without the mean-field term, the above equation is studied in [22] and references therein.

Further applications include the following two popular models:

1. Kuramoto Model: The phase $X_t^{i,N}$ of the i-th Kuramoto oscillator at time t is given by the following SDE,
$$X_t^{i,N} = X_0^{i,N} + \int_0^t \big(\nu^i + \frac{\kappa}{N}\sum_{k=1}^N \sin(X_s^{k,N} - X_s^{i,N})\big)ds + \frac{\sqrt{2\sigma}}{N}\int_0^t \sum_{k=1}^N \sin(X_s^{k,N} - X_s^{i,N})dW_s^i$$

almost surely for any $t \in [0,T]$ and $i \in \{1,\dots,N\}$ for some $N \in \mathbb{N}$, where ν^i, κ and σ are constants. This type of model has been investigated in [1], [20] and references therein.

2. Cucker–Smale Model: The position $X_t^{i,N}$ and the velocity $V_t^{i,N}$ of the i-th agent in the Cucker–Smale model is given by the following SDE

$$X_t^{i,N} = X_0^{i,N} + \int_0^t V_s^{i,N} ds,$$
$$V_t^{i,N} = V_0^{i,N} + \frac{\lambda}{N}\int_0^t \sum_{k=1}^N (V_s^{k,N} - V_s^{i,N})ds + \frac{\sigma^{i,N}}{N}\int_0^t \sum_{k=1}^N (V_s^{k,N} - V_s^{i,N})dW_s^i$$

almost surely for any $t \in [0,T]$ and $i \in \{1,\dots,N\}$ for some $N \in \mathbb{N}$, where parameters $\lambda > 0$ and $\sigma^{i,N} \in \mathbb{R}$ are coupling and noise strength respectively. This type of model has been studied in [17] and reference therein.

We finish this section by introducing the notations and the notion of Lion's derivative used in this article.

1.1 Notations

The notation $\langle \cdot, \cdot \rangle$ stands for the inner product in $\mathbb{R}^d$. We use the same notation $|\cdot|$ for both Euclidean and Hilbert–Schmidt norms and its meaning should be clear from the context. Also, σx denotes the usual matrix multiplication of $\sigma \in \mathbb{R}^{d\times m}$ and $x \in \mathbb{R}^d$. With a slight abuse of notation, $b^{(l)}$ and $\sigma^{(l)}$ are used to denote the l-th element of $b \in \mathbb{R}^d$ and the l-th column vector of $\sigma \in \mathbb{R}^{d\times m}$ respectively which is clear from the context and should not cause any confusion in the reader's mind. Further, $\sigma^{(k,l)}$ stands for the (k,l)-th element of $\sigma \in \mathbb{R}^{d\times m}$. For a function $f : \mathbb{R}^d \to \mathbb{R}$, we use $\partial_x f$ to denote the gradient of f. The symbol $\lfloor \cdot \rfloor$ stands for the floor function and $\delta_x(\cdot)$ denotes the Dirac measure at point $x \in \mathbb{R}^d$. We use $\mathcal{B}(\mathbb{R}^d)$ to denote the Borel σ-field on $\mathbb{R}^d$. Moreover, $\mathcal{P}_2(\mathbb{R}^d)$ denotes the space of probability measures μ on the measurable space $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ satisfying $\int_{\mathbb{R}^d} |x|^2 \mu(dx) < \infty$. It is well-known that $\mathcal{P}_2(\mathbb{R}^d)$ is a Polish space under the $\mathcal{L}^2$-Wasserstein metric given by

$$\mathcal{W}_2(\mu_1,\mu_2) := \inf_{\pi \in \Pi(\mu_1,\mu_2)} \Big(\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |x-y|^2 \pi(dx,dy) \Big)^{1/2}$$

where $\Pi(\mu_1,\mu_2)$ is the set of all couplings of $\mu_1, \mu_2 \in \mathcal{P}_2(\mathbb{R}^d)$. Further, EX is used for the expectation of a random variable X. Throughout this article, K stands for a generic constant which may vary from place to place.

1.2 Differentiability of functions of measures

There are many different notions for differentiating functions of measures and we refer the reader to [2], [10] and [42] for a nice exposition on this topic. In this article, we use the notion of differentiability introduced by Lions in his lectures at the Collège de France which has been reproduced in [9]. We give a brief description of the concept of measure derivative for functions defined on the Wasserstein space $\mathcal{P}_2(\mathbb{R}^d)$. A function $f : \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is said to be differentiable at $\nu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ if there exists an atomless, Polish probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and a random variable $Y_0 \in \mathcal{L}^2(\tilde{\Omega}; \mathbb{R}^d)$ such that its law $L_{Y_0} := \tilde{P} \circ Y_0^{-1} = \nu_0$ and the function $F : \mathcal{L}^2(\tilde{\Omega}; \mathbb{R}^d) \to \mathbb{R}$ defined by $F(Z) := f(L_Z)$ has Fréchet derivative at $Y_0 \in \mathcal{L}^2(\tilde{\Omega}; \mathbb{R}^d)$, which we denote by $F'[Y_0]$. The function F is called the "lift" of f. Further, f is said to be of class C^1 if its "lift" F is of class C^1. Since $F'[Y_0] : \mathcal{L}^2(\tilde{\Omega}; \mathbb{R}^d) \to \mathbb{R}$ is a bounded linear operator, by Riesz representation theorem, there exists an element $DF(Y_0) \in \mathcal{L}^2(\tilde{\Omega}; \mathbb{R}^d)$ such that $F'[Y_0](Z) = \tilde{E}\langle DF[Y_0], Z\rangle$ for all

$Z \in \mathcal{L}^2(\tilde{\Omega}; \mathbb{R}^d)$. By Theorem 6.5 (structure of the gradient) in [9], if f is of class C^1, then there exists a function $\partial_\mu f(\nu_0) : \mathbb{R}^d \to \mathbb{R}^d$ satisfying $\int_{\mathbb{R}^d} |\partial_\mu f(\nu_0)(x)|^2 \nu_0(dx) < \infty$ such that $DF(Y_0) = \partial_\mu f(\nu_0)(Y_0)$. Also, $\partial_\mu f(\nu_0)$ is independent of the choice of the probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ and the random variable Y_0. The function $\partial_\mu f(\nu_0)$ is called the *Lions' derivative* of f at $\nu_0 = L_{Y_0}$. Moreover, $\partial_\mu f : \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ is defined as $\partial_\mu f(\nu, z) = \partial_\mu f(\nu)(z)$ for any $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ and $z \in \mathbb{R}^d$.

2 Assumptions and main results

Let $(\Omega, \{\mathcal{F}_t\}_{\{t\geq 0\}}, \mathcal{F}, P)$ be a filtered probability space satisfying the usual conditions, *i.e.*, the probability space $(\Omega, \mathcal{F}, P)$ is complete, $\mathcal{F}_0$ contains all P – null sets of $\mathcal{F}$ and filtration is right continuous. Let $\{W_t\}_{\{t\geq 0\}}$ be an m-dimensional Brownian motion adapted to the filtration $\{\mathcal{F}_t\}_{\{t\geq 0\}}$. Assume that $b : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^{d\times m}$ are measurable functions. We consider the following McKean–Vlasov Stochastic Differential Equation (MV–SDE) defined on $(\Omega, \{\mathcal{F}_t\}_{\{t\geq 0\}}, \mathcal{F}, P)$,

$$X_t = X_0 + \int_0^t b(X_s, \mu_s^X) ds + \sum_{l=1}^{m} \int_0^t \sigma^{(l)}(X_s, \mu_s^X) dW_s^{(l)} \tag{2.1}$$

almost surely for any $t \in [0, T]$ where μ_s^X denotes the law of X_s, *i.e.* $\mu_s^X := P \circ X_s^{-1}$ for every $s \in [0, T]$ and X_0 stands for an $\mathbb{R}^d$-valued and $\mathcal{F}_0$-measurable random variable.

We make the following assumptions on the coefficients and the initial value.

Assumption 2.1. *$E|X_0|^p < \infty$ for a fixed constant $p \geq 2$.*

Assumption 2.2. *There exist constants $L > 0$ and $\rho > 0$ such that,*

$$\begin{aligned} \big\langle x - \bar{x}, b(x,\mu) - b(\bar{x},\mu)\big\rangle &\leq L|x-\bar{x}|^2, \\ |b(x,\mu) - b(\bar{x},\bar{\mu})| &\leq L\big\{(1+|x|+|\bar{x}|)^{\rho/2+1}|x-\bar{x}| + \mathcal{W}_2(\mu,\bar{\mu})\big\}, \\ |b(0,\mu)| &\leq L, \end{aligned}$$

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

Assumption 2.3. *There exists a constant $L > 0$ such that,*

$$\begin{aligned} |\sigma(x,\mu) - \sigma(\bar{x},\bar{\mu})| &\leq L\big\{|x-\bar{x}| + \mathcal{W}_2(\mu,\bar{\mu})\big\}, \\ |\sigma(0,\mu)| &\leq L, \end{aligned}$$

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

Assumption 2.4. *There exists a constant $L > 0$ such that for every $k \in \{1, \ldots, d\}$, the derivatives $\partial_x b^{(k)} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ satisfy,*

$$\begin{aligned} |\partial_x b^{(k)}(x,\mu) - \partial_x b^{(k)}(\bar{x},\bar{\mu})| &\leq L\big\{(1+|x|+|\bar{x}|)^{\rho/2}|x-\bar{x}| + \mathcal{W}_2(\mu,\bar{\mu})\big\}, \\ |\partial_x b^{(k)}(0,\mu)| &\leq L, \end{aligned}$$

and the measure derivatives $\partial_\mu b^{(k)} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ satisfy,

$$\begin{aligned} |\partial_\mu b^{(k)}(x,\mu,y) - \partial_\mu b^{(k)}(\bar{x},\bar{\mu},\bar{y})| &\leq L\big\{(1+|x|+|\bar{x}|)^{\rho/2+1}|x-\bar{x}| + \mathcal{W}_2(\mu,\bar{\mu}) + |y-\bar{y}|\big\}, \\ |\partial_\mu b^{(k)}(0,\mu,0)| &\leq L, \end{aligned}$$

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

Assumption 2.5. *There exists a constant $L > 0$ such that for every $k \in \{1, \ldots, d\}$ and $l \in \{1, \ldots, m\}$, the derivatives $\partial_x \sigma^{(k,l)} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ satisfy,*

$$|\partial_x \sigma^{(k,l)}(x,\mu) - \partial_x \sigma^{(k,l)}(\bar{x},\bar{\mu})| \leq L\big\{|x-\bar{x}| + \mathcal{W}_2(\mu,\bar{\mu})\big\},$$

$$|\partial_x \sigma^{(k,l)}(0,\mu)| \leq L,$$

and the measure derivatives $\partial_\mu \sigma^{(k,l)} : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ *satisfy,*

$$|\partial_\mu \sigma^{(k,l)}(x,\mu,y) - \partial_\mu \sigma^{(k,l)}(\bar{x},\bar{\mu},\bar{y})| \leq L\big\{|x-\bar{x}| + \mathcal{W}_2(\mu,\bar{\mu}) + |y-\bar{y}|\big\},$$
$$|\partial_\mu \sigma^{(k,l)}(0,\mu,0)| \leq L,$$

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

2.1 Wellposedness, propagation of chaos and interacting particle system

For a fixed $N \in \mathbb{N}$, let $\{W^i\}_{i\in\{1,\ldots,N\}}$ be N independent Brownian motions that are also independent of W. Consider the N-dimensional system of interacting particles given by,

$$X_t^{i,N} = X_0^i + \int_0^t b\big(X_s^{i,N}, \mu_s^{X,N}\big)ds + \sum_{l=1}^m \int_0^t \sigma^{(l)}\big(X_s^{i,N}, \mu_s^{X,N}\big)dW_s^{(l),i} \tag{2.2}$$

almost surely for any $t \in [0,T]$ and $i \in \{1,\ldots,N\}$, where

$$\mu_s^{X,N} := \frac{1}{N}\sum_{j=1}^N \delta_{X_s^{j,N}}$$

for any $s \in [0,T]$. For the propagation of chaos result, consider the system of non-interacting particles given by,

$$X_t^i = X_0^i + \int_0^t b\big(X_s^i, \mu_s^{X^i}\big)ds + \sum_{l=1}^m \int_0^t \sigma^{(l)}\big(X_s^i, \mu_s^{X^i}\big)dW_s^{(l),i}$$

almost surely for any $t \in [0,T]$ and $i \in \{1,\ldots,N\}$, where $\mu_s^{X^i} = \mu_s^X$ for every $i \in \{1,\ldots,N\}$ because X^i's are independent.

The proof of the following proposition can be found in [31], [34] and [35].

Proposition 2.1. Let Assumptions 2.1, 2.2 and 2.3 hold for some $p > 4$. Then, there exists a unique solution to MV–SDE (2.1) and the following holds,

$$E\sup_{t\in[0,T]}\big|X_t\big|^p \leq K,$$

where $K := K(m,d,L,p,T,E|X_0|^p) > 0$ is a constant. Also, the interacting particle system (2.2) has a unique solution and

$$\sup_{i\in\{1,\ldots,N\}} E\sup_{t\in[0,T]}\big|X_t^i - X_t^{i,N}\big|^2 \leq K\begin{cases} N^{-1/2} & \text{if } d<4, \\ N^{-1/2}\ln(N) & \text{if } d=4, \\ N^{-2/d} & \text{if } d>4, \end{cases}$$

where the constant $K > 0$ does not depend on N.

2.2 Explicit Milstein-type scheme

For introducing Milstein-type scheme for MV–SDE (2.1), we partition the interval $[0,T]$ into n sub-intervals each of length $h = T/n$ and define $\kappa_n(s) := \lfloor ns \rfloor / n$ for any $s \in [0,T]$. For any $x \in \mathbb{R}^d$ and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, define

$$b_n\big(x,\mu\big) := \frac{b\big(x,\mu\big)}{1+n^{-1}|x|^{\rho+2}} \tag{2.3}$$

for every $n \in \mathbb{N}$. We propose the following explicit Milstein-type scheme for MV–SDE (2.1),

$$X_t^{i,N,n} = X_0^i + \int_0^t b_n\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)ds + \sum_{l=1}^{m} \int_0^t \tilde{\sigma}^{(l)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)dW_s^{(l),i} \tag{2.4}$$

almost surely for any $t \in [0,T]$ and $n, N \in \mathbb{N}$, where

$$\mu_{\kappa_n(s)}^{X,N} := \frac{1}{N}\sum_{j=1}^{N} \delta_{X_{\kappa_n(s)}^{j,N,n}}$$

for any $s \in [0,T]$ and $n, N \in \mathbb{N}$. Also, for every $l \in \{1, \ldots, m\}$, $\tilde{\sigma}^{(l)}$ is given by,

$$\begin{aligned}\tilde{\sigma}^{(l)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big) &:= \sigma^{(l)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big) + \Lambda_1^{(l)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big) \\ &\quad + \Lambda_2^{(l)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)\end{aligned} \tag{2.5}$$

where $\Lambda_1^{(l)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)$ and $\Lambda_2^{(l)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)$ are the l-th column of $d \times m$-matrices whose (k,l)-th elements are respectively given by,

$$\begin{aligned}
&\Lambda_1^{(k,l)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big) \\
&\quad := \Big\langle \partial_x \sigma^{(k,l)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big), \sum_{l_1=1}^{m} \int_{\kappa_n(s)}^{s} \sigma^{(l_1)}\big(X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^{(l_1),i} \Big\rangle \\
&\quad = \sum_{l_1=1}^{m} \int_{\kappa_n(s)}^{s} \Big\langle \partial_x \sigma^{(k,l)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big), \sigma^{(l_1)}\big(X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big) \Big\rangle dW_r^{(l_1),i}, \\
&\Lambda_2^{(k,l)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big) \\
&\quad := \frac{1}{N}\sum_{j=1}^{N} \Big\langle \partial_\mu \sigma^{(k,l)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}, X_{\kappa_n(s)}^{j,N,n}\big), \sum_{l_1=1}^{m} \int_{\kappa_n(s)}^{s} \sigma^{(l_1)}\big(X_{\kappa_n(r)}^{j,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^{(l_1),j} \Big\rangle \\
&\quad = \frac{1}{N}\sum_{j=1}^{N}\sum_{l_1=1}^{m} \int_{\kappa_n(s)}^{s} \Big\langle \partial_\mu \sigma^{(k,l)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}, X_{\kappa_n(s)}^{j,N,n}\big), \sigma^{(l_1)}\big(X_{\kappa_n(r)}^{j,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big) \Big\rangle dW_r^{(l_1),j}
\end{aligned}$$

for any $s \in [0,T]$ and $n, N \in \mathbb{N}$ and for every $k \in \{1, \ldots, d\}$ and $l \in \{1, \ldots, m\}$.

Remark 2.6. It is easy to observe that the term Λ_2 may be ignored for sufficiently large values of N. Indeed, due to Remark 2.10 and Lemma 3.4 (mentioned below),

$$E\big|\Lambda_2\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)\big|^2 \le K(n^{-1}/N) \to 0 \text{ as } N \to \infty$$

and hence

$$\Lambda_2\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big) = 0 \text{ a.s. as } N \to \infty.$$

This observation can be used in implementing the scheme on the computer, see Subsection 2.3.

The following proposition and theorem are the main results of this article.

Proposition 2.2. Let Assumptions 2.1 to 2.5 hold for some $p \ge 8(\rho/2 + 2)$. Then, the Milstein-type scheme (2.4) converges to the interacting particle system (2.2) with the rate of convergence given by,

$$\sup_{i \in \{1,\ldots,N\}} \sup_{t \in [0,T]} E\big|X_t^{i,N} - X_t^{i,N,n}\big|^2 \le Kn^{-2},$$

for any $n, N \in \mathbb{N}$ where the constant $K > 0$ does not depend on $n, N \in \mathbb{N}$.

By combining Propositions 2.1 and 2.2, we obtain the following theorem.

Theorem 2.7. *Let Assumptions 2.1 to 2.5 hold for some* $p \geq 8(\rho/2+2)$. *Then, the Milstein-type scheme* (2.4) *converges to the true solution of MV–SDE* (2.1) *with the rate of convergence given by,*

$$\sup_{i\in\{1,\ldots,N\}}\sup_{t\in[0,T]} E\big|X_t^i - X_t^{i,N,n}\big|^2 \leq K \begin{cases} N^{-1/2}+n^{-2} & \text{if } d<4,\\ N^{-1/2}\ln(N)+n^{-2} & \text{if } d=4,\\ N^{-2/d}+n^{-2} & \text{if } d>4, \end{cases}$$

for any $n, N \in \mathbb{N}$ *where constant* $K>0$ *does not depend on* $n, N \in \mathbb{N}$.

We conclude this section by listing following remarks which are consequences of the assumptions mentioned above.

Remark 2.8. From Assumptions 2.2 and 2.4,

$$\begin{aligned} \langle x, b(x,\mu)\rangle &\leq K(1+|x|)^2,\\ |b(x,\mu)| &\leq K(1+|x|)^{\rho/2+2},\\ |\partial_x b^{(k)}(x,\mu)| &\leq K(1+|x|)^{\rho/2+1},\\ |\partial_\mu b^{(k)}(x,\mu,y)| &\leq K, \end{aligned}$$

for any $x, y \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $k \in \{1,\ldots,d\}$.

Remark 2.9. From Remark 2.8 and equation (2.3),

$$|b_n(x,\mu)| \leq K\min\left\{n^{1/2}(1+|x|),(1+|x|)^{\rho/2+2}\right\},$$

for all $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $n \in \mathbb{N}$ where the constant $K>0$ does not depend on $n \in \mathbb{N}$.

Remark 2.10. From Assumptions 2.3 and 2.5,

$$\begin{aligned} |\sigma(x,\mu)| &\leq K(1+|x|),\\ |\partial_x \sigma^{(k,l)}(x,\mu)| &\leq K,\\ |\partial_\mu \sigma^{(k,l)}(x,\mu,y)| &\leq K, \end{aligned}$$

for any $x, y \in \mathbb{R}^d$, $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, $k \in \{1,\ldots,d\}$ and $l \in \{1,\ldots,m\}$.

2.3 Example

We now present some examples of the Milstein-type scheme (2.4). Let $0<h<2h<\ldots<nh$ be a partition of the interval $[0,T]$ where $h=T/n$. Consider the case when $\sigma(x,\mu)=\sigma(x)$ for all $x\in\mathbb{R}^d$ and $\mu\in\mathcal{P}_2(\mathbb{R}^d)$ which implies $\Lambda_2=0$. Also notice that,

$$\begin{aligned} &\sum_{l=1}^m \int_{jh}^{(j+1)h} \Lambda_1^{(k,l)}\big(s, X_{\kappa_n(s)}^{i,N,n}\big)dW_s^{(l),i}\\ &= \sum_{l=1}^m\sum_{l_1=1}^m \int_{jh}^{(j+1)h}\int_{\kappa_n(s)}^s \Big\langle \partial_x\sigma^{(k,l)}\big(X_{\kappa_n(r)}^{i,N,n}\big), \sigma^{(l_1)}\big(X_{\kappa_n(r)}^{i,N,n}\big)\Big\rangle dW_r^{(l_1),i}dW_s^{(l),i}\\ &= \sum_{l=1}^m \int_{jh}^{(j+1)h}\int_{jh}^s \Big\langle \partial_x\sigma^{(k,l)}\big(X_{jh}^{i,N,n}\big), \sigma^{(l)}\big(X_{jh}^{i,N,n}\big)\Big\rangle dW_r^{(l),i}dW_s^{(l),i}\\ &\quad+\sum_{l=1}^m\sum_{l_1>l} \int_{jh}^{(j+1)h}\int_{jh}^s \Big\{\Big\langle \partial_x\sigma^{(k,l)}\big(X_{jh}^{i,N,n}\big), \sigma^{(l_1)}\big(X_{jh}^{i,N,n}\big)\Big\rangle dW_r^{(l_1),i}dW_s^{(l),i} \end{aligned}$$

$$+\Big\langle \partial_x\sigma^{(k,l_1)}\big(X_{jh}^{i,N,n}\big),\sigma^{(l)}\big(X_{jh}^{i,N,n}\big)\Big\rangle dW_r^{(l),i}dW_s^{(l_1),i}\Big\}$$

and if the commutative condition,

$$\big\langle \partial_x\sigma^{(k,l)}(x),\sigma^{(l_1)}(x)\big\rangle=\big\langle \partial_x\sigma^{(k,l_1)}(x),\sigma^{(l)}(x)\big\rangle$$

is assumed for all $k=1,\dots,d$ and $l,l_1=1,\dots,m$, then the k-th component is given by,

$$\begin{aligned}
&\sum_{l=1}^{m}\int_{jh}^{(j+1)h}\Lambda_1^{(k,l)}\big(s,X_{jh}^{i,N,n}\big)dW_s^{(l),i}\\
&\quad=\sum_{l=1}^{m}\Big\langle \partial_x\sigma^{(k,l)}\big(X_{jh}^{i,N,n}\big),\sigma^{(l)}\big(X_{jh}^{i,N,n}\big)\Big\rangle\int_{jh}^{(j+1)h}\int_{jh}^{s}dW_r^{(l),i}dW_s^{(l),i}\\
&\qquad+\sum_{l=1}^{m}\sum_{l_1>l}^{m}\Big\langle \partial_x\sigma^{(k,l)}\big(X_{jh}^{i,N,n}\big),\sigma^{(l_1)}\big(X_{jh}^{i,N,n}\big)\Big\rangle\\
&\qquad\times\Big\{\int_{jh}^{(j+1)h}\int_{jh}^{s}dW_r^{(l_1),i}dW_s^{(l),i}+\int_{jh}^{(j+1)h}\int_{jh}^{s}dW_r^{(l),i}dW_s^{(l_1),i}\Big\}\\
&\quad=\sum_{l=1}^{m}\Big\langle \partial_x\sigma^{(k,l)}\big(X_{jh}^{i,N,n}\big),\sigma^{(l)}\big(X_{jh}^{i,N,n}\big)\Big\rangle\frac{1}{2}\big\{(\Delta W_{(j+1)h}^{(l),i})^2-\Delta t\big\}\\
&\qquad+\sum_{l=1}^{m}\sum_{l_1>l}^{m}\Big\langle \partial_x\sigma^{(k,l)}\big(X_{jh}^{i,N,n}\big),\sigma^{(l_1)}\big(X_{jh}^{i,N,n}\big)\Big\rangle\Delta W_{(j+1)h}^{(l_1),i}\Delta W_{(j+1)h}^{(l),i}\\
&\quad=\frac{1}{2}\sum_{l=1}^{m}\sum_{l_1=1}^{m}\Big\langle \partial_x\sigma^{(k,l)}\big(X_{jh}^{i,N,n}\big),\sigma^{(l_1)}\big(X_{jh}^{i,N,n}\big)\Big\rangle\big\{\Delta W_{(j+1)h}^{(l_1),i}\Delta W_{(j+1)h}^{(l),i}-hI_{\{l=l_1\}}\big\}
\end{aligned}$$

almost surely for any $j=0,\dots,n-1$, $i\in\{1,\dots,N\}$, $n,N\in\mathbb{N}$, where $I_{\{l=l_1\}}$ is the indicator function of the set $\{l=l_1\}$. Thus, the Milstein-type scheme (2.4) becomes

$$\begin{aligned}
X_{(j+1)h}^{i,N,n}&=X_{jh}^{i,N,n}+b_n(X_{jh}^{i,N,n},\mu_{jh}^{X,N,n})h+\sum_{l=1}^{m}\sigma^{(l)}(X_{jh}^{i,N,n})\Delta W_{(j+1)h}^{(l),i}\\
&\quad+\frac{1}{2}\sum_{l=1}^{m}\sum_{l_1=1}^{m}\Big\langle \partial_x\sigma^{(l)}\big(X_{jh}^{i,N,n}\big),\sigma^{(l_1)}\big(X_{jh}^{i,N,n}\big)\Big\rangle\big\{\Delta W_{(j+1)h}^{(l_1),i}\Delta W_{(j+1)h}^{(l),i}-hI_{\{l=l_1\}}\big\}
\end{aligned}$$

almost surely for any $j=0,\dots,n-1$, $i\in\{1,\dots,N\}$, $n,N\in\mathbb{N}$.

Notice that when σ is a function of both $x\in\mathbb{R}^d$ and $\mu\in\mathcal{P}_2(\mathbb{R}^d)$, then a commutative condition similar to the one mentioned above is needed when N is small. However, such commutative condition is not required when N is sufficiently large due to Remark 2.6.

Remark 2.11. The results of this manuscript can also be extended to the case when b and σ additionally depend on the time parameter, *i.e* when $b:[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\mapsto\mathbb{R}^d$ and $\sigma:[0,T]\times\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\mapsto\mathbb{R}^{d\times m}$. However, the proofs become rather cumbersome. In this case, we require the following additional regularity conditions,

$$\begin{aligned}
&|b(t,x,\mu)-b(\bar t,x,\mu)|+|\sigma(t,x,\mu)-\sigma(\bar t,x,\mu)|\le L|t-\bar t|,\\
&\Big|\frac{\partial b(t,x,\mu)}{\partial t}\Big|\le L(1+|x|)^{\rho/2+2},\ \Big|\frac{\partial \sigma(t,x,\mu)}{\partial t}\Big|\le L(1+|x|),
\end{aligned}$$

for all $t,\bar t\in[0,T]$, $x\in\mathbb{R}^d$ and $\mu\in\mathcal{P}_2(\mathbb{R}^d)$. The explicit Milstein-type scheme is given by,

$$X_t^{i,N,n}=X_0^i+\int_0^t b_n\big(\kappa_n(s),X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)ds+\sum_{l=1}^{m}\int_0^t\tilde\sigma^{(l)}\big(s,X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)dW_s^{(l),i}$$

where b_n and k-th element of $\tilde{\sigma}^{(l)}$ are given by,

$$\begin{aligned}
&b_n\big(\kappa_n(s), X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big) := \frac{b\big(\kappa_n(s), X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)}{1+n^{-1}|X^{i,N,n}_{\kappa_n(s)}|^{\rho+2}}\,,\\
&\tilde{\sigma}^{(k,l)}\big(s, X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big) = \sigma^{(k,l)}\big(\kappa_n(s), X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)\\
&+\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\big\langle \partial_x\sigma^{(k,l)}\big(\kappa_n(s), X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big), \sigma^{(l_1)}\big(\kappa_n(s), X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)\big\rangle dW^{(l_1),i}_r\\
&+\frac{1}{N}\sum_{j=1}^{N}\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\big\langle \partial_\mu\sigma^{(k,l)}\big(\kappa_n(s), X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}, X^{j,N,n}_{\kappa_n(s)}\big),\\
&\qquad\qquad\qquad \sigma^{(l_1)}\big(\kappa_n(s), X^{j,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)\big\rangle dW^{(l_1),j}_r
\end{aligned}$$

almost surely for any $s\in[0,T]$, $n,N\in\mathbb{N}$, $k\in\{1,\dots,d\}$ and $l\in\{1,\dots,m\}$.

3 Moment bounds

Before establishing the moment bound of the Milstein-type scheme (2.4) in Lemma 3.4, we first establish following lemmas and corollaries.

Lemma 3.1. *Let Assumptions 2.3 and 2.5 be satisfied. Then,*

$$E|\Lambda_1^{(u,v)}\big(s, X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)|^p \le Kn^{-\frac{p}{2}}E(1+|X^{i,N,n}_{\kappa_n(s)}|)^p,$$

for any $s\in[0,T]$, $i\in\{1,\cdots,N\}$, n, $N\in\mathbb{N}$, $u\in\{1,\dots,d\}$ and $v\in\{1,\dots,m\}$, where constant $K>0$ does not depend on n and N.

Proof. By Burkholder–Gundy–Davis inequality, Cauchy–Schwarz inequality and Hölder's inequality,

$$\begin{aligned}
&E|\Lambda_1^{(u,v)}\big(s, X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)|^p\\
&= E\Big|\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\Big\langle \partial_x\sigma^{(u,v)}\big(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big), \sigma^{(l_1)}\big(X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\big)\Big\rangle dW^{(l_1),i}_r\Big|^p\\
&\le Kn^{-\frac{p}{2}+1}E\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\big|\partial_x\sigma^{(u,v)}\big(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)\big|^p\big|\sigma^{(l_1)}\big(X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\big)\big|^p dr
\end{aligned}$$

and then the application of Remark 2.10 completes the proof. □

Lemma 3.2. *Let Assumptions 2.3 and 2.5 be satisfied. Then,*

$$E|\Lambda_2^{(u,v)}\big(s, X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)|^p \le Kn^{-\frac{p}{2}}\frac{1}{N}\sum_{j=1}^{N}E\big(1+|X^{j,N,n}_{\kappa_n(s)}|\big)^p,$$

for any $s\in[0,T]$, $i\in\{1,\cdots,N\}$, n, $N\in\mathbb{N}$, $u\in\{1,\dots,d\}$ and $v\in\{1,\dots,m\}$, where constant $K>0$ does not depend on n and N.

Proof. On using Burkholder–Gundy–Davis inequality, Cauchy–Schwarz inequality and Hölder's inequality,

$$\begin{aligned}
&E|\Lambda_2^{(u,v)}\big(s, X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)|^p\\
&= E\Big|\frac{1}{N}\sum_{j=1}^{N}\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\Big\langle \partial_\mu\sigma^{(u,v)}(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}, X^{j,N,n}_{\kappa_n(s)}), \sigma^{(l_1)}\big(X^{j,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\big)\Big\rangle dW^{(l_1),j}_r\Big|^p
\end{aligned}$$

$$\leq Kn^{-\frac{p}{2}+1}E\frac{1}{N}\sum_{j=1}^{N}\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\big|\partial_\mu\sigma^{(u,v)}(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)})\big|^p\big|\sigma^{(l_1)}\big(X^{j,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^p dr$$

and then the proof is completed by using Remark 2.10. □

As a consequence of Remark 2.10, Lemma 3.1 and Lemma 3.2, one obtains the following corollary.

Corollary 3.1. Let Assumptions 2.3 and 2.5 be satisfied. Then,

$$E\big|\tilde\sigma\big(s,X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big)\big|^p \leq KE\big(1+|X^{i,N,n}_{\kappa_n(s)}|\big)^p+Kn^{-\frac{p}{2}}E\big(1+|X^{i,N,n}_{\kappa_n(s)}|\big)^p$$
$$+Kn^{-\frac{p}{2}}\frac{1}{N}\sum_{j=1}^{N}E\big(1+|X^{j,N,n}_{\kappa_n(s)}|\big)^p,$$

for any $s\in[0,T]$, $i\in\{1,\cdots,N\}$ and n, $N\in\mathbb{N}$, where the constant $K>0$ does not depend on n and N.

Lemma 3.3. *Let Assumptions 2.2 to 2.5 be satisfied. Then,*

$$E|X^{i,N,n}_s-X^{i,N,n}_{\kappa_n(s)}|^p\leq Kn^{-\frac{p}{2}}E\big(1+|X^{i,N,n}_{\kappa_n(s)}|\big)^p+Kn^{-\frac{p}{2}}\frac{1}{N}\sum_{j=1}^{N}E\big(1+|X^{j,N,n}_{\kappa_n(s)}|\big)^p,$$

for any $s\in[0,T]$, $i\in\{1,\cdots,N\}$ and n, $N\in\mathbb{N}$ where the constant $K>0$ does not depend on n and N.

Proof. From equation (2.4), one can get the following estimate,

$$E|X^{i,N,n}_s-X^{i,N,n}_{\kappa_n(s)}|^p\leq KE\Big|\int_{\kappa_n(s)}^{s}b_n\big(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dr\Big|^p$$
$$+KE\Big|\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\tilde\sigma^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^{(l),i}_r\Big|^p$$

and then the application of Hölder's inequality and Burkholder–Gundy–Davis inequality gives,

$$E|X^{i,N,n}_s-X^{i,N,n}_{\kappa_n(s)}|^p\leq Kn^{-p+1}E\int_{\kappa_n(s)}^{s}|b_n\big(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)|^p dr$$
$$+Kn^{-\frac{p}{2}+1}E\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}|\tilde\sigma^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)|^p dr$$

which on using Remark 2.9 and Corollary 3.1 completes the proof. □

Lemma 3.4. *Let Assumptions 2.1 to 2.5 be satisfied. Then,*

$$\sup_{i\in\{1,\dots,N\}}E\sup_{t\in[0,T]}(1+|X^{i,N,n}_t|^2)^{p/2}\leq K,$$

for any n, $N\in\mathbb{N}$ where the constant $K>0$ does not depend on n and N.

Proof. By the application of Itô's formula,

$$\big(1+|X^{i,N,n}_t|^2\big)^{p/2}=\big(1+|X^i_0|^2\big)^{p/2}+p\int_0^t\big(1+|X^{i,N,n}_s|^2\big)^{p/2-1}\big\langle X^{i,N,n}_s,b_n\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big)\big\rangle ds$$
$$+p\int_0^t\big(1+|X^{i,N,n}_s|^2\big)^{p/2-1}\big\langle X^{i,N,n}_s,\tilde\sigma\big(s,X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big)dW^i_s\big\rangle$$

$$+\frac{p(p-2)}{2}\int_0^t \big(1+|X_s^{i,N,n}|^2\big)^{p/2-2}\big|\tilde{\sigma}^*\big(s,X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)X_s^{i,N,n}\big|^2ds$$
$$+\frac{p}{2}\int_0^t \big(1+|X_s^{i,N,n}|^2\big)^{p/2-1}\big|\tilde{\sigma}\big(s,X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big|^2ds$$

almost surely, which on the application of Burkholder–Gundy–Davis inequality and Cauchy–Schwarz inequality yields the following estimate,

$$E\sup_{t\in[0,u]}\big(1+|X_t^{i,N,n}|^2\big)^{p/2}\le E\big(1+|X_0^i|^2\big)^{p/2}$$
$$+KE\sup_{t\in[0,u]}\Big|\int_0^t \big(1+|X_s^{i,N,n}|^2\big)^{p/2-1}\big\langle X_{\kappa_n(s)}^{i,N,n},b_n\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big\rangle ds\Big|$$
$$+KE\sup_{t\in[0,u]}\Big|\int_0^t \big(1+|X_s^{i,N,n}|^2\big)^{p/2-1}\big\langle X_s^{i,N,n}-X_{\kappa_n(s)}^{i,N,n},b_n\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big\rangle ds\Big|$$
$$+KE\Big\{\int_0^u \big(1+|X_s^{i,N,n}|^2\big)^{p-2}|\tilde{\sigma}\big(s,X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)|^2ds\Big\}^{\frac{1}{2}}$$
$$+KE\int_0^u \big(1+|X_s^{i,N,n}|^2\big)^{p/2-1}\big|\tilde{\sigma}\big(s,X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big|^2ds$$

for any $i\in\{1,\dots,N\}$, $n,N\in\mathbb{N}$, and $u\in[0,T]$. Also, one uses Remark 2.8, Cauchy–Schwarz inequality and Young's inequality to obtain the following estimate,

$$E\sup_{t\in[0,u]}\big(1+|X_t^{i,N,n}|^2\big)^{p/2}\le E\big(1+|X_0^i|^2\big)^{p/2}+K\int_0^u E\sup_{r\in[0,s]}\big(1+|X_r^{i,N,n}|^2\big)^{p/2}ds$$
$$+KE\int_0^u \big(1+|X_s^{i,N,n}|^2\big)^{p/2-1}\big(1+|X_{\kappa_n(s)}^{i,N,n}|^2\big)ds$$
$$+KE\int_0^u n^{p/4}\big|X_s^{i,N,n}-X_{\kappa_n(s)}^{i,N,n}\big|^{p/2}n^{-p/4}\big|b_n\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big|^{p/2}ds$$
$$+KE\Big\{\int_0^u \big|\tilde{\sigma}\big(s,X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big|^2ds\Big\}^{p/2}+KE\int_0^u \big|\tilde{\sigma}\big(s,X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big|^pds$$

and then due to Young's inequality and Hölder's inequality, one gets

$$E\sup_{t\in[0,u]}\big(1+|X_t^{i,N,n}|^2\big)^{p/2}\le E\big(1+|X_0^i|^2\big)^{p/2}+K\int_0^u E\sup_{r\in[0,s]}\big(1+|X_r^{i,N,n}|^2\big)^{p/2}ds$$
$$+KE\int_0^u n^{\frac{p}{2}}|X_s^{i,N,n}-X_{\kappa_n(s)}^{i,N,n}|^pds+KE\int_0^u n^{-\frac{p}{2}}|b_n\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)|^pds$$
$$+KE\int_0^u |\tilde{\sigma}\big(s,X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)|^pds$$

for any $i\in\{1,\dots,N\}$, $n,N\in\mathbb{N}$ and $u\in[0,T]$. The application of Remark 2.9, Lemma 3.3 and Corollary 3.1 yields,

$$\sup_{i\in\{1,\dots,N\}}E\sup_{t\in[0,u]}(1+|X_t^{i,N,n}|^2)^{p/2}\le E(1+|X_0^2)^{p/2}$$
$$+K\int_0^u \sup_{i\in\{1,\dots,N\}}E\sup_{r\in[0,s]}(1+|X_r^{i,N,n}|^2)^{p/2}ds$$

for any $n,N\in\mathbb{N}$ and $u\in[0,T]$. Finally, the proof is completed by using Grönwall's inequality. □

4 Rate of convergence

In this section, we shall prove Proposition 2.2. For this, we require some lemmas and corollaries which are proved below.

Notice that as a consequence of Lemmas 3.1, 3.2 and 3.4, we obtain the following corollary.

Corollary 4.1. Let $q \geq 2$ and Assumptions 2.1 to 2.5 hold for some $p \geq (\rho/2+2)q$. Then,

$$\begin{aligned} &E|\Lambda_1\big(s, X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)|^q \leq Kn^{-\frac{q}{2}}, \\ &E|\Lambda_2\big(s, X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)|^q \leq Kn^{-\frac{q}{2}}, \\ &E|\tilde{\sigma}\big(s, X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)|^q \leq K, \\ \text{and} \quad &E|X^{i,N,n}_s - X^{i,N,n}_{\kappa_n(s)}|^q \leq Kn^{-\frac{q}{2}}, \end{aligned}$$

for any $s \in [0,T]$, $i \in \{1, \cdots, N\}$ and n, $N \in \mathbb{N}$, where K is a constant that does not depend on n and N.

The following lemma is very useful in this article.

Lemma 4.1. *Let $f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a continuous function such that its derivative $\partial_x f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and its measure derivative $\partial_\mu f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ exists. Then, there exists a $\theta \in (0,1)$ such that,*

$$\begin{aligned} f(y,\mu) - f(\bar{y},\bar{\mu}) = &\big\langle \partial_x f(\bar{y} + \theta(y-\bar{y}), \mu), y - \bar{y} \big\rangle \\ &+ \tilde{E}\big\langle \partial_\mu f(\bar{y}, L_{\bar{Z}+\theta(Z-\bar{Z})}, \bar{Z} + \theta(Z-\bar{Z})), Z - \bar{Z} \big\rangle, \end{aligned}$$

for any $y, \bar{y} \in \mathbb{R}^d$, $\mu, \bar{\mu} \in \mathcal{P}_2(\mathbb{R}^d)$, and random variables $Z, \bar{Z}$ defined on an atomless, Polish probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$ such that $L_Z = \mu$ and $L_{\bar{Z}} = \bar{\mu}$.

Proof. Define $\phi : [0,1] \to \mathbb{R}$ by,

$$\phi(t) := f(\bar{y} + t(y-\bar{y}), \mu) + F(\bar{y}, \bar{Z} + t(Z - \bar{Z}))$$

where $F(\bar{y}, \cdot)$ is an “extension” of $f(\bar{y}, \cdot)$ on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$. For any $t_0 \in (0,1)$, we have,

$$\begin{aligned} \lim_{h\to 0} \frac{\phi(t_0+h) - \phi(t_0)}{h} = &\lim_{h\to 0} \frac{f(\bar{y} + (t_0+h)(y-\bar{y}), \mu) - f(\bar{y} + t_0(y-\bar{y}), \mu)}{h} \\ &+ \lim_{h\to 0} \frac{F(\bar{y}, \bar{Z} + (t_0+h)(Z-\bar{Z})) - F(\bar{y}, \bar{Z} + t_0(Z-\bar{Z}))}{h} < \infty \end{aligned}$$

i.e., $\phi'(t_0)$ exists. Notice that the second term on the right hand side of the above expression is the Gateaux derivative of F at $\bar{Z} + t_0(Z-\bar{Z})$ in the direction of $Z - \bar{Z}$. Also, ϕ is continuous on $[0,1]$. Hence, by the mean value theorem, there exists a $\theta \in [0,1]$ such that $\phi(1) - \phi(0) = \phi'(\theta)$. Thus,

$$f(y,\mu) - f(\bar{y},\bar{\mu}) = f(y,\mu) + F(\bar{y}, Z) - f(\bar{y},\mu) - F(\bar{y},\bar{Z}) = \phi'(\theta)$$

completes the proof. □

As a special case of the above lemma, we obtain the following corollary.

Corollary 4.2. Let $f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}$ be a function such that its derivative $\partial_x f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \to \mathbb{R}^d$ and its measure derivative $\partial_\mu f : \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ exists. Then, there exists a $\theta \in (0,1)$ such that,

$$f\Big(y, \frac{1}{N}\sum_{j=1}^N \delta_{z^j}\Big) - f\Big(\bar{y}, \frac{1}{N}\sum_{j=1}^N \delta_{\bar{z}^j}\Big) = \Big\langle \partial_x f\Big(\bar{y} + \theta(y-\bar{y}), \frac{1}{N}\sum_{j=1}^N \delta_{z^j}\Big), y - \bar{y} \Big\rangle$$

$$+\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu f\Big(\bar{y},\frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{z}^j+\theta(z^j-\bar{z}^j)},\bar{z}^j+\theta(z^j-\bar{z}^j)\Big),z^j-\bar{z}^j\Big\rangle,$$

for all $y,\bar{y},z^j,\bar{z}^j\in\mathbb{R}^d$.

Proof. For the measures $\mu=\frac{1}{N}\sum_{j=1}^{N}\delta_{z^j}$ and $\bar{\mu}=\frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{z}^j}$, consider an atomless, Polish probability space $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{P})$ and define random variables $Z:\tilde{\Omega}\to\mathbb{R}^d$ and $\bar{Z}:\tilde{\Omega}\to\mathbb{R}^d$ by,

$$Z:=\sum_{j=1}^{N}z^j\mathbb{1}_{\tilde{\Omega}_j},\ \bar{Z}:=\sum_{j=1}^{N}\bar{z}^j\mathbb{1}_{\tilde{\Omega}_j}$$

where $\{\tilde{\Omega}_1,\ldots,\tilde{\Omega}_N\}$ is a partition of $\tilde{\Omega}$ satisfying $\tilde{P}(\tilde{\Omega}_j)=\frac{1}{N}$ for any $j=1,\ldots,N$. Clearly, laws of Z and $\bar{Z}$ satisfy $L_Z=\mu$ and $L_{\bar{Z}}=\bar{\mu}$. Also,

$$L_{\bar{Z}+\theta(Z-\bar{Z})}=\frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{z}^j+\theta(z^j-\bar{z}^j)}.$$

By Lemma 4.1,

$$\begin{aligned}
f\Big(y,\frac{1}{N}\sum_{j=1}^{N}\delta_{z^j}\Big)-f\Big(\bar{y},\frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{z}^j}\Big)&=\Big\langle\partial_x f\Big(\bar{y}+\theta(y-\bar{y}),\frac{1}{N}\sum_{j=1}^{N}\delta_{z^j}\Big),y-\bar{y}\Big\rangle\\
&\quad+\tilde{E}\Big\langle\partial_\mu f\Big(\bar{y},\frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{z}^j+\theta(z^j-\bar{z}^j)},\sum_{j=1}^{N}\{\bar{z}_j+\theta(z^j-\bar{z}^j)\}\mathbb{1}_{\tilde{\Omega}_j}\Big),\sum_{j=1}^{N}(z^j-\bar{z}^j)\mathbb{1}_{\tilde{\Omega}_j}\Big\rangle\\
&=\Big\langle\partial_x f\Big(\bar{y}+\theta(y-\bar{y}),\frac{1}{N}\sum_{j=1}^{N}\delta_{z^j}\Big),y-\bar{y}\Big\rangle\\
&\quad+\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu f\Big(\bar{y},\frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{z}^j+\theta(z^j-\bar{z}^j)},\bar{z}^j+\theta(z^j-\bar{z}^j)\Big),z^j-\bar{z}^j\Big\rangle
\end{aligned}$$

which completes the proof. □

Lemma 4.2. *Let $f:\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ be a function such that its derivative $\partial_x f:\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}^d$ and measure derivative $\partial_\mu f:\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\to\mathbb{R}^d$ satisfy Lipschitz conditions i.e., there exists a constant $L>0$ such that,*

$$\begin{aligned}
|\partial_x f(x,\mu)-\partial_x f(\bar{x},\bar{\mu})|&\le L\big\{|x-\bar{x}|+\mathcal{W}_2(\mu,\bar{\mu})\big\},\\
|\partial_\mu f(x,\mu,y)-\partial_\mu f(\bar{x},\bar{\mu},\bar{y})|&\le L\big\{|x-\bar{x}|+\mathcal{W}_2(\mu,\bar{\mu})+|y-\bar{y}|\big\},
\end{aligned}$$

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

$$\begin{aligned}
&\Big|f\Big(x^i,\frac{1}{N}\sum_{j=1}^{N}\delta_{x^j}\Big)-f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{x}^j}\Big)-\Big\langle\partial_x f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{x}^j}\Big),x^i-\bar{x}^i\Big\rangle\\
&\quad-\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^{N}\delta_{\bar{x}^j},\bar{x}^j\Big),x^j-\bar{x}^j\Big\rangle\Big|\le K|x^i-\bar{x}^i|^2+K\frac{1}{N}\sum_{j=1}^{N}|x^j-\bar{x}^j|^2,
\end{aligned}$$

for all x^i, $\bar{x}^i\in\mathbb{R}^d$, $i\in\{1,\ldots,N\}$, where the constant $K>0$ does not depend on $N\in\mathbb{N}$.

Proof. By Corollary 4.2, Cauchy–Schwarz inequality, assumptions on f and Young's inequality, one obtains,

$$\begin{aligned}
&\Big|f\Big(x^i,\frac{1}{N}\sum_{j=1}^N\delta_{x^j}\Big)-f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j}\Big)-\Big\langle\partial_x f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j}\Big),x^i-\bar{x}^i\Big\rangle\\
&-\frac{1}{N}\sum_{j=1}^N\Big\langle\partial_\mu f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j},\bar{x}^j\Big),x^j-\bar{x}^j\Big\rangle\Big|\\
=&\Big|\Big\langle\partial_x f\Big(\bar{x}^i+\theta(x^i-\bar{x}^i),\frac{1}{N}\sum_{j=1}^N\delta_{x^j}\Big)-\partial_x f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j}\Big),x^i-\bar{x}^i\Big\rangle\\
&+\frac{1}{N}\sum_{j=1}^N\Big\langle\partial_\mu f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j+\theta(x^j-\bar{x}^j)},\bar{x}^j+\theta(x^j-\bar{x}^j)\Big)-\partial_\mu f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j},\bar{x}^j\Big),x^j-\bar{x}^j\Big\rangle\Big|\\
\leq&K|x^i-\bar{x}^i|^2+K\mathcal{W}_2\Big(\frac{1}{N}\sum_{j=1}^N\delta_{x^j},\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j}\Big)^2\\
&\quad+K\mathcal{W}_2\Big(\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j+\theta(x^j-\bar{x}^j)},\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j}\Big)^2+K\frac{1}{N}\sum_{j=1}^N|x^j-\bar{x}^j|^2.
\end{aligned}$$

The proof is completed by the following estimate on Wasserstein metric,

$$\begin{aligned}
&\mathcal{W}_2\Big(\frac{1}{N}\sum_{j=1}^N\delta_{x^j},\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j}\Big)^2\\
&=\inf_\pi\Big\{\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|x-y|^2\pi(dx,dy):\frac{1}{N}\sum_{j=1}^N\delta_{x^j}\text{ and }\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j}\text{ are marginals of }\pi\Big\}\\
&\leq\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}|x-y|^2\frac{1}{N}\sum_{j=1}^N\delta_{x^j}(dx)\delta_{\bar{x}^j}(dy)=\frac{1}{N}\sum_{j=1}^N|x^j-\bar{x}^j|^2.
\end{aligned}$$

□

Lemma 4.3. *Let $f:\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}$ be a function such that its derivative $\partial_x f:\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\to\mathbb{R}^d$ and measure derivative $\partial_\mu f:\mathbb{R}^d\times\mathcal{P}_2(\mathbb{R}^d)\times\mathbb{R}^d\to\mathbb{R}^d$ satisfy polynomial Lipschitz condition i.e., there exists a constant $L>0$ such that,*

$$\begin{aligned}
|\partial_x f(x,\mu)-\partial_x f(\bar{x},\bar{\mu})|&\leq L\big\{(1+|x|+|\bar{x}|)^{\rho/2}|x-\bar{x}|+\mathcal{W}_2(\mu,\bar{\mu})\big\},\\
|\partial_\mu f(x,\mu,y)-\partial_\mu f(\bar{x},\bar{\mu},\bar{y})|&\leq L\big\{(1+|x|+|\bar{x}|)^{\rho/2+1}|x-\bar{x}|+\mathcal{W}_2(\mu,\bar{\mu})+|y-\bar{y}|\big\},
\end{aligned}$$

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

$$\begin{aligned}
\Big|f\Big(x^i,\frac{1}{N}\sum_{j=1}^N\delta_{x^j}\Big)-f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j}\Big)-\Big\langle\partial_x f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j}\Big),x^i-\bar{x}^i\Big\rangle&\\
-\frac{1}{N}\sum_{j=1}^N\Big\langle\partial_\mu f\Big(\bar{x}^i,\frac{1}{N}\sum_{j=1}^N\delta_{\bar{x}^j},\bar{x}^j\Big),x^j-\bar{x}^j\Big\rangle\Big|&\\
\leq K(1+|x^i|+|\bar{x}^i|)^{\rho/2}|x^i-\bar{x}^i|^2+K\frac{1}{N}\sum_{j=1}^N|x^j-\bar{x}^j|^2,&
\end{aligned}$$

for all x^i, $\bar{x}^i\in\mathbb{R}^d$, $i\in\{1,\dots,N\}$, where the constant $K>0$ does not depend on $N\in\mathbb{N}$.

Proof. The proof follows by adapting the arguments of Lemma 4.2. □

Lemma 4.4. *Let Assumptions 2.1 to 2.5 hold for some $p \geq 4(\rho/2+2)$. Then, for each $i \in \{1, \ldots, N\}$,*

$$E|\sigma(X_s^{i,N,n}, \mu_s^{X,N,n}) - \tilde{\sigma}(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n})|^2 \leq Kn^{-2},$$

for any $s \in [0,T]$ and $n, N \in \mathbb{N}$ where the constant $K > 0$ does not depend on n, N.

Proof. From equation (2.5),

$$\begin{aligned}
&\sigma^{(u,v)}(X_s^{i,N,n}, \mu_s^{X,N,n}) - \tilde{\sigma}^{(u,v)}(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}) \\
&\quad= \sigma^{(u,v)}(X_s^{i,N,n}, \mu_s^{X,N,n}) - \sigma^{(u,v)}(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}) \\
&\qquad - \Big\langle \partial_x \sigma^{(u,v)}(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}), X_s^{i,N,n} - X_{\kappa_n(s)}^{i,N,n} \Big\rangle \\
&\qquad - \frac{1}{N}\sum_{j=1}^{N} \Big\langle \partial_\mu \sigma^{(u,v)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}, X_{\kappa_n(s)}^{j,N,n}\big), X_s^{j,N,n} - X_{\kappa_n(s)}^{j,N,n} \Big\rangle \\
&\qquad + \Big\langle \partial_x \sigma^{(u,v)}(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}), X_s^{i,N,n} - X_{\kappa_n(s)}^{i,N,n} \Big\rangle \\
&\qquad + \frac{1}{N}\sum_{j=1}^{N} \Big\langle \partial_\mu \sigma^{(u,v)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}, X_{\kappa_n(s)}^{j,N,n}\big), X_s^{j,N,n} - X_{\kappa_n(s)}^{j,N,n} \Big\rangle \\
&\qquad - \Lambda_1^{(u,v)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big) - \Lambda_2^{(u,v)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)
\end{aligned} \tag{4.1}$$

almost surely for any $s \in [0,T]$ and $n, N \in \mathbb{N}$. Further, from equation (2.4),

$$\begin{aligned}
X_s^{i,N,n} - X_{\kappa_n(s)}^{i,N,n} &= \int_{\kappa_n(s)}^s b_n\big(X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dr + \sum_{l=1}^m \int_{\kappa_n(s)}^s \sigma^{(l)}\big(X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^{(l),i} \\
&\quad + \sum_{l=1}^m \int_{\kappa_n(s)}^s \Lambda_1^{(l)}\big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^{(l),i} + \sum_{l=1}^m \int_{\kappa_n(s)}^s \Lambda_2^{(l)}\big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^{(l),i}
\end{aligned}$$

which gives the following expression,

$$\begin{aligned}
&\Big\langle \partial_x \sigma^{(u,v)}(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}), X_s^{i,N,n} - X_{\kappa_n(s)}^{i,N,n} \Big\rangle = \Lambda_1^{(u,v)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big) \\
&\quad + \Big\langle \partial_x \sigma^{(u,v)}(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}), \int_{\kappa_n(s)}^s b_n\big(X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dr \Big\rangle \\
&\quad + \Big\langle \partial_x \sigma^{(u,v)}(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}), \sum_{l=1}^m \int_{\kappa_n(s)}^s \Lambda_1^{(l)}\big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^{(l),i} \Big\rangle \\
&\quad + \Big\langle \partial_x \sigma^{(u,v)}(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}), \sum_{l=1}^m \int_{\kappa_n(s)}^s \Lambda_2^{(l)}\big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^{(l),i} \Big\rangle
\end{aligned} \tag{4.2}$$

and similarly,

$$\begin{aligned}
&\frac{1}{N}\sum_{j=1}^{N} \Big\langle \partial_\mu \sigma^{(u,v)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}, X_{\kappa_n(s)}^{j,N,n}\big), X_s^{j,N,n} - X_{\kappa_n(s)}^{j,N,n} \Big\rangle = \Lambda_2^{(u,v)}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big) \\
&+ \frac{1}{N}\sum_{j=1}^{N} \Big\langle \partial_\mu \sigma^{(u,v)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}, X_{\kappa_n(s)}^{j,N,n}\big), \int_{\kappa_n(s)}^s b_n\big(X_{\kappa_n(r)}^{j,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dr \Big\rangle \\
&+ \frac{1}{N}\sum_{j=1}^{N} \Big\langle \partial_\mu \sigma^{(u,v)}\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}, X_{\kappa_n(s)}^{j,N,n}\big), \sum_{l=1}^m \int_{\kappa_n(s)}^s \Lambda_1^{(l)}\big(r, X_{\kappa_n(r)}^{j,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^{(l),j} \Big\rangle
\end{aligned}$$

$$+\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu\sigma^{(u,v)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\Lambda_2^{(l)}\big(r,X^{j,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^{(l),j}_r\Big\rangle \tag{4.3}$$

almost surely for any $s\in[0,T]$ and $n,N\in\mathbb{N}$. We now use equations (4.2) and (4.3) in equation (4.1) which on taking expectation on both the sides and then using Lemma 4.2 along with Cauchy–Schwarz inequality yields,

$$\begin{aligned}
&E|\sigma^{(u,v)}(X^{i,N,n}_s,\mu^{X,N,n}_s)-\tilde\sigma^{(u,v)}(s,X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)})|^2\\
&\le KE|X^{i,N,n}_s-X^{i,N,n}_{\kappa_n(s)}|^4+K\frac{1}{N^2}\sum_{j,k=1}^{N}E|X^{j,N,n}_s-X^{j,N,n}_{\kappa_n(s)}|^2|X^{k,N,n}_s-X^{k,N,n}_{\kappa_n(s)}|^2\\
&+KE\Big\{|\partial_x\sigma^{(u,v)}(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)})|\Big|\int_{\kappa_n(s)}^{s}b_n\big(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dr\Big|\Big\}^2\\
&+KE\Big\{|\partial_x\sigma^{(u,v)}(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)})|\sum_{l=1}^{m}\Big|\int_{\kappa_n(s)}^{s}\Lambda_1^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^{(l),i}_r\Big|\Big\}^2\\
&+KE\Big\{|\partial_x\sigma^{(u,v)}(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)})|\sum_{l=1}^{m}\Big|\int_{\kappa_n(s)}^{s}\Lambda_2^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^{(l),i}_r\Big|\Big\}^2\\
&+KE\Big\{\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu\sigma^{(u,v)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big)\big|\Big|\int_{\kappa_n(s)}^{s}b_n\big(X^{j,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dr\Big|\Big\}^2\\
&+KE\Big\{\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu\sigma^{(u,v)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big)\big|\sum_{l=1}^{m}\Big|\int_{\kappa_n(s)}^{s}\Lambda_1^{(l)}\big(r,X^{j,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^{(l),j}_r\Big|\Big\}^2\\
&+KE\Big\{\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu\sigma^{(u,v)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big)\big|\sum_{l=1}^{m}\Big|\int_{\kappa_n(s)}^{s}\Lambda_2^{(l)}\big(r,X^{j,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^{(l),j}_r\Big|\Big\}^2
\end{aligned}$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$. Moreover, the application of Corollary 4.1, Young's inequality, Remarks 2.8 and 2.10 yields

$$\begin{aligned}
&E|\sigma^{(u,v)}(X^{i,N,n}_s,\mu^{X,N,n}_s)-\tilde\sigma^{(u,v)}(s,X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)})|^2\le Kn^{-2}+Kn^{-2}E(1+|X^{i,N,n}_{\kappa_n(s)}|)^{\rho+4}\\
&\quad+Kn^{-2}\frac{1}{N}\sum_{j=1}^{N}E(1+|X^{j,N,n}_{\kappa_n(s)}|)^{\rho+4}+KE\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big|\Lambda_1^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^2dr\\
&\quad+KE\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}|\Lambda_2^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)|^2dr+KE\frac{1}{N}\sum_{j=1}^{N}\int_{\kappa_n(s)}^{s}\big|\Lambda_1^{(l)}\big(r,X^{j,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^2dr\\
&\quad+KE\frac{1}{N}\sum_{j=1}^{N}\int_{\kappa_n(s)}^{s}\big|\Lambda_2^{(l)}\big(r,X^{j,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^2dr
\end{aligned}$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$. The proof is completed by the application of Lemma 3.4 and Corollary 4.1. □

Lemma 4.5. *Let Assumptions 2.1 to 2.5 hold for some* $p\ge 4(\rho/2+2)$*. Then, for each* $i\in\{1,\dots,N\}$*,*

$$E|b(X^{i,N,n}_s,\mu^{X,N,n}_s)-b_n(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)})|^2\le Kn^{-1},$$

for any $s\in[0,T]$ *and* $n,N\in\mathbb{N}$ *where the constant* $K>0$ *does not depend on* n,N*.*

Proof. By Assumption 2.2,

$$\begin{aligned}
E|b(X_s^{i,N,n},\mu_s^{X,N,n})-b_n(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n})|^2 &\leq KE|b(X_s^{i,N,n},\mu_s^{X,N,n})-b(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n})|^2\\
&+KE|b(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n})-b_n(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n})|^2\\
&\leq KE\big(1+|X_s^{i,N,n}|+|X_{\kappa_n(s)}^{i,N,n}|\big)^{\rho+2}|X_s^{i,N,n}-X_{\kappa_n(s)}^{i,N,n}|^2+KE\mathcal{W}_2\big(\mu_s^{X,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)^2
\end{aligned}$$

which on using Hölder's inequality and Corollary 4.1 completes the proof. □

Lemma 4.6. *Let Assumptions 2.1 to 2.5 hold for some $p\geq 8(\rho/2+2)$. Then, for each $i\in\{1,\ldots,N\}$,*

$$\begin{aligned}
&E\langle X_s^{i,N}-X_s^{i,N,n},b(X_s^{i,N,n},\mu_s^{X,N,n})-b(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n})\rangle\\
&\qquad\leq Kn^{-2}+K\sup_{i\in\{1,\cdots,N\}}\sup_{r\in[0,s]}E|X_r^{i,N}-X_r^{i,N,n}|^2,
\end{aligned}$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$, where the constant $K>0$ does not depend on n,N.

Proof. Note that,

$$\begin{aligned}
&E\Big\langle X_s^{i,N}-X_s^{i,N,n},b\Big(X_s^{i,N,n},\mu_s^{X,N,n}\Big)-b\Big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\Big)\Big\rangle\\
&=E\sum_{k=1}^{d}\big(X_s^{(k),i,N}-X_s^{(k),i,N,n}\big)\Big\{b^{(k)}\big(X_s^{i,N,n},\mu_s^{X,N,n}\big)-b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\\
&\quad-\Big\langle\partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\Big),X_s^{i,N,n}-X_{\kappa_n(s)}^{i,N,n}\Big\rangle\\
&\quad-\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}),X_s^{j,N,n}-X_{\kappa_n(s)}^{j,N,n}\Big\rangle\Big\}\\
&\quad+E\sum_{k=1}^{d}\big(X_s^{(k),i,N}-X_s^{(k),i,N,n}\big)\Big\langle\partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\Big),X_s^{i,N,n}-X_{\kappa_n(s)}^{i,N,n}\Big\rangle\\
&\quad+E\sum_{k=1}^{d}\big(X_s^{(k),i,N}-X_s^{(k),i,N,n}\big)\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}),X_s^{j,N,n}-X_{\kappa_n(s)}^{j,N,n}\Big\rangle\\
&=:T_1+T_2+T_3
\end{aligned}\tag{4.4}$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$. By using Young's inequality, T_1 is estimated as,

$$\begin{aligned}
T_1:=&E\sum_{k=1}^{d}\big(X_s^{(k),i,N}-X_s^{(k),i,N,n}\big)\Big\{b^{(k)}\big(X_s^{i,N,n},\mu_s^{X,N,n}\big)-b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\\
&-\Big\langle\partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\Big),X_s^{i,N,n}-X_{\kappa_n(s)}^{i,N,n}\Big\rangle\\
&-\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}),X_s^{j,N,n}-X_{\kappa_n(s)}^{j,N,n}\Big\rangle\Big\}\\
\leq&KE\sum_{k=1}^{d}\big|X_s^{(k),i,N}-X_s^{(k),i,N,n}\big|^2+KE\sum_{k=1}^{d}\Big|b^{(k)}\big(X_s^{i,N,n},\mu_s^{X,N,n}\big)-b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\\
&-\Big\langle\partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\Big),X_s^{i,N,n}-X_{\kappa_n(s)}^{i,N,n}\Big\rangle\\
&-\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}),X_s^{j,N,n}-X_{\kappa_n(s)}^{j,N,n}\Big\rangle\Big|^2
\end{aligned}$$

which on using Lemma 4.3, Corollary 4.1, Lemma 3.4 and Hölder's inequality yields,

$$T_1 \leq KE\big|X_s^{i,N} - X_s^{i,N,n}\big|^2 + KE(1 + |X_s^{i,N,n}| + |X_{\kappa_n(s)}^{i,N,n}|)^{\rho}|X_s^{i,N,n} - X_{\kappa_n(s)}^{i,N,n}|^4$$
$$+ K\frac{1}{N}\sum_{j=1}^{N} E|X_s^{j,N,n} - X_{\kappa_n(s)}^{j,N,n}|^4 \leq K \sup_{i\in\{1,\ldots,N\}} \sup_{r\in[0,s]} E\big|X_r^{i,N} - X_r^{i,N,n}\big|^2 + Kn^{-2} \tag{4.5}$$

for any $s \in [0, T]$ and $n, N \in \mathbb{N}$. Further, using equation (2.4),

$$T_2 := E\sum_{k=1}^{d}\big(X_s^{(k),i,N} - X_s^{(k),i,N,n}\big)\Big\langle \partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\Big), X_s^{i,N,n} - X_{\kappa_n(s)}^{i,N,n}\Big\rangle$$
$$= E\sum_{k=1}^{d}\big(X_s^{(k),i,N} - X_s^{(k),i,N,n}\big)\Big\langle \partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\Big), \int_{\kappa_n(s)}^{s} b_n\Big(X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\Big)dr\Big\rangle$$
$$+ E\sum_{k=1}^{d}\big(X_s^{(k),i,N} - X_s^{(k),i,N,n}\big)\Big\langle \partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\Big)dW_r^i\Big\rangle$$
$$= E\sum_{k=1}^{d}\big(X_s^{(k),i,N} - X_s^{(k),i,N,n}\big)\Big\langle \partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\Big), \int_{\kappa_n(s)}^{s} b_n\Big(X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\Big)dr\Big\rangle$$
$$+ E\sum_{k=1}^{d}\big(X_{\kappa_n(s)}^{(k),i,N} - X_{\kappa_n(s)}^{(k),i,N,n}\big)\Big\langle \partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\Big)dW_r^i\Big\rangle$$
$$+ E\sum_{k=1}^{d}\big(X_s^{(k),i,N} - X_{\kappa_n(s)}^{(k),i,N} - X_s^{(k),i,N,n} + X_{\kappa_n(s)}^{(k),i,N,n}\big)$$
$$\Big\langle \partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\Big)dW_r^i\Big\rangle$$

for any $s \in [0, T]$ and $n, N \in \mathbb{N}$. Notice that the second term on the right hand side of the above expression is zero. Thus, from Young's inequality, Remark 2.9 and equations (2.2) and (2.4), one obtains

$$T_2 \leq KE\big|X_s^{i,N} - X_s^{i,N,n}\big|^2 + Kn^{-2}E\big(1 + \big|X_{\kappa_n(s)}^{i,N,n}\big|\big)^{2\rho+6}$$
$$+ E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}(X_r^{i,N}, \mu_r^{X,N}) - b_n^{(k)}(X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n})\big\}dr$$
$$\Big\langle \partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\Big)dW_r^i\Big\rangle$$
$$+ E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}(X_r^{i,N}, \mu_r^{X,N}) - \tilde{\sigma}^{(k,l)}(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n})\big\}dW_r^{(l),i}$$
$$\Big\langle \partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\Big)dW_r^i\Big\rangle$$

for any $s \in [0, T]$ and $n, N \in \mathbb{N}$. By using Lemma 3.4, one can write

$$T_2 \leq KE\big|X_s^{i,N} - X_s^{i,N,n}\big|^2 + Kn^{-2} + E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}(X_r^{i,N}, \mu_r^{X,N}) - b^{(k)}(X_r^{i,N,n}, \mu_r^{X,N})\big\}dr$$
$$\Big\langle \partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X_{\kappa_n(r)}^{i,N,n}, \mu_{\kappa_n(r)}^{X,N,n}\Big)dW_r^i\Big\rangle$$
$$+E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}(X_r^{i,N,n}, \mu_r^{X,N}) - b^{(k)}(X_r^{i,N,n}, \mu_r^{X,N,n})\big\}dr$$

$$
\begin{aligned}
&\quad \Big\langle \partial_x b^{(k)}\Big(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\Big) dW^i_r \Big\rangle \\
&+E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s} \big\{b^{(k)}(X^{i,N,n}_r, \mu^{X,N,n}_r) - b^{(k)}_n(X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)})\big\} dr \\
&\quad \Big\langle \partial_x b^{(k)}\Big(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\Big) dW^i_r \Big\rangle \\
&+E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s} \big\{\sigma^{(k,l)}(X^{i,N}_r, \mu^{X,N}_r) - \sigma^{(k,l)}(X^{i,N,n}_r, \mu^{X,N}_r)\big\} dW^{(l),i}_r \\
&\quad \Big\langle \partial_x b^{(k)}\Big(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\Big) dW^i_r \Big\rangle \\
&+E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s} \big\{\sigma^{(k,l)}(X^{i,N,n}_r, \mu^{X,N}_r) - \sigma^{(k,l)}(X^{i,N,n}_r, \mu^{X,N,n}_r)\big\} dW^{(l),i}_r \\
&\quad \Big\langle \partial_x b^{(k)}\Big(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\Big) dW^i_r \Big\rangle \\
&+E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s} \big\{\sigma^{(k,l)}(X^{i,N,n}_r, \mu^{X,N,n}_r) - \tilde{\sigma}^{(k,l)}(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)})\big\} dW^{(l),i}_r \\
&\quad \Big\langle \partial_x b^{(k)}\Big(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\Big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\Big(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\Big) dW^i_r \Big\rangle \\
&=: \sup_{i\in\{1,\cdots,N\}}\sup_{r\in[0,s]} E\big|X^{i,N}_s - X^{i,N,n}_s\big|^2 + Kn^{-2} + T_{21} + T_{22} + T_{23} + T_{24} + T_{25} + T_{26}
\end{aligned} \tag{4.6}
$$

for any $s \in [0,T]$ and $n, N \in \mathbb{N}$.

Using Cauchy–Schwarz inequality, Young's inequality, Assumption 2.2 and Remark 2.8, T_{21} can be estimated by,

$$
\begin{aligned}
T_{21} &:= E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s} \big\{b^{(k)}(X^{i,N}_r, \mu^{X,N}_r) - b^{(k)}(X^{i,N,n}_r, \mu^{X,N}_r)\big\} dr \\
&\qquad \Big\langle \partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big), \int_{\kappa_n(s)}^{s} \tilde{\sigma}\big(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\big) dW^i_r \Big\rangle \\
&\le E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s} \big|b^{(k)}(X^{i,N}_r, \mu^{X,N}_r) - b^{(k)}(X^{i,N,n}_r, \mu^{X,N}_r)\big| dr \\
&\qquad \big|\partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)}, \mu^{X,N,n}_{\kappa_n(s)}\big)\big| \Big|\int_{\kappa_n(s)}^{s} \tilde{\sigma}\big(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\big) dW^i_r\Big| \\
&\le KE\int_{\kappa_n(s)}^{s} \big(1 + |X^{i,N}_r| + |X^{i,N,n}_r|\big)^{\rho/2+1} \big|X^{i,N}_r - X^{i,N,n}_r\big| dr \\
&\qquad \Big|\int_{\kappa_n(s)}^{s} \big(1 + |X^{i,N,n}_{\kappa_n(s)}|\big)^{\rho/2+1} \tilde{\sigma}\big(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\big) dW^i_r\Big| \\
&\le KE\int_{\kappa_n(s)}^{s} n^{-\frac{1}{2}}\big(1 + |X^{i,N}_r| + |X^{i,N,n}_r|\big)^{\rho/2+1} \\
&\qquad \Big|\int_{\kappa_n(s)}^{s} \big(1 + |X^{i,N,n}_{\kappa_n(s)}|\big)^{\rho/2+1} \tilde{\sigma}\big(r, X^{i,N,n}_{\kappa_n(r)}, \mu^{X,N,n}_{\kappa_n(r)}\big) dW^i_r\Big| n^{\frac{1}{2}} \big|X^{i,N}_r - X^{i,N,n}_r\big| dr
\end{aligned}
$$

which on using Young's inequality and Hölder's inequality gives,

$$T_{21} \leq KE\int_{\kappa_n(s)}^{s} n^{-1}\big(1+|X_r^{i,N}|+|X_r^{i,N,n}|\big)^{\rho+2}$$
$$\Big|\int_{\kappa_n(s)}^{s}\big(1+|X_{\kappa_n(s)}^{i,N,n}|\big)^{\rho/2+1}\tilde{\sigma}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^i\Big|^2 dr + KE\int_{\kappa_n(s)}^{s} n\big|X_r^{i,N}-X_r^{i,N,n}\big|^2 dr$$
$$\leq K\int_{\kappa_n(s)}^{s} n^{-1}\big\{E\big(1+|X_r^{i,N}|+|X_r^{i,N,n}|\big)^{2\rho+2}\big\}^{\frac{1}{2}}$$
$$\Big\{E\Big|\int_{\kappa_n(s)}^{s}\big(1+|X_{\kappa_n(s)}^{i,N,n}|\big)^{\rho/2+1}\tilde{\sigma}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^i\Big|^4\Big\}^{\frac{1}{2}} dr + K\sup_{r\in[0,s]}E\big|X_r^{i,N}-X_r^{i,N,n}\big|^2$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$. Further use of Lemma 3.4, Corollary 4.1 and Hölder's inequality yields,

$$\begin{aligned}T_{21} &\leq K\int_{\kappa_n(s)}^{s} n^{-1}\Big\{E\Big(\int_{\kappa_n(s)}^{s}\big(1+|X_{\kappa_n(s)}^{i,N,n}|\big)^{\rho+2}\big|\tilde{\sigma}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)\big|^2 dr\Big)^2\Big\}^{\frac{1}{2}} dr \\ &\quad + K\sup_{r\in[0,s]}E\big|X_r^{i,N}-X_r^{i,N,n}\big|^2 \\ &\leq K\int_{\kappa_n(s)}^{s} n^{-1}\Big\{n^{-1}E\int_{\kappa_n(s)}^{s}\big(1+|X_{\kappa_n(s)}^{i,N,n}|\big)^{2\rho+4}\big|\tilde{\sigma}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)\big|^4 dr\Big\}^{\frac{1}{2}} dr \\ &\quad + K\sup_{r\in[0,s]}E\big|X_r^{i,N}-X_r^{i,N,n}\big|^2 \\ &\leq K\int_{\kappa_n(s)}^{s} n^{-1}\Big\{n^{-1}\int_{\kappa_n(s)}^{s}\big\{E\big(1+|X_{\kappa_n(s)}^{i,N,n}|\big)^{4\rho+8}\big\}^{\frac{1}{2}}\big\{E\big|\tilde{\sigma}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)\big|^8\big\}^{\frac{1}{2}} dr\Big\}^{\frac{1}{2}} dr \\ &\quad + K\sup_{r\in[0,s]}E\big|X_r^{i,N}-X_r^{i,N,n}\big|^2 \leq K\sup_{i\in\{1,\ldots,N\}}\sup_{r\in[0,s]}E\big|X_r^{i,N}-X_r^{i,N,n}\big|^2 + Kn^{-3} \end{aligned} \tag{4.7}$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$.

Using Cauchy–Schwarz inequality, Assumption 2.2 and Remark 2.8, T_{22} can be estimated as,

$$\begin{aligned}T_{22} &:= E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}(X_r^{i,N,n},\mu_r^{X,N})-b^{(k)}(X_r^{i,N,n},\mu_r^{X,N,n})\big\}dr \\ &\qquad \Big\langle\partial_x b^{(k)}\Big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\Big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^i\Big\rangle \\ &\leq KE\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big|b^{(k)}(X_r^{i,N,n},\mu_r^{X,N})-b^{(k)}(X_r^{i,N,n},\mu_r^{X,N,n})\big|dr \\ &\qquad \big|\partial_x b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big|\Big|\int_{\kappa_n(s)}^{s}\tilde{\sigma}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^i\Big| \\ &\leq KE\int_{\kappa_n(s)}^{s} n^{\frac{1}{2}}\Big\{\frac{1}{N}\sum_{j=1}^{N}\big|X_r^{j,N}-X_r^{j,N,n}\big|^2\Big\}^{\frac{1}{2}} \\ &\qquad n^{-\frac{1}{2}}\Big|\int_{\kappa_n(s)}^{s}\big(1+|X_{\kappa_n(s)}^{i,N,n}|\big)^{\rho/2+1}\tilde{\sigma}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)dW_r^i\Big|dr\end{aligned}$$

which on the application of Young's inequality, Lemma 3.4, Hölder's inequality and Corollary 4.1 yields,

$$T_{22} \leq KE\int_{\kappa_n(s)}^{s} n\frac{1}{N}\sum_{j=1}^{N}\big|X_r^{j,N,n}-X_r^{j,N,n}\big|^2 dr$$

$$
\begin{aligned}
&\quad + Kn^{-2}E\Big|\int_{\kappa_n(s)}^{s}\big(1+\big|X^{i,N,n}_{\kappa_n(s)}\big|\big)^{\rho/2+1}\tilde{\sigma}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^i_r\Big|^2\\
&\leq K\sup_{i\in\{1,\cdots,N\}}\sup_{r\in[0,s]}E\big|X^{i,N}_r-X^{i,N,n}_r\big|^2\\
&\quad + Kn^{-2}E\int_{\kappa_n(s)}^{s}\big(1+\big|X^{i,N,n}_{\kappa_n(s)}\big|\big)^{\rho+2}\big|\tilde{\sigma}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^2dr\\
&\leq K\sup_{i\in\{1,\cdots,N\}}\sup_{r\in[0,s]}E\big|X^{i,N}_r-X^{i,N,n}_r\big|^2+Kn^{-3}
\end{aligned}
\tag{4.8}
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$.

Further, using Cauchy–Schwarz inequality, Hölder's inequality, Lemma 4.5, Remark 2.8, Corollary 4.1 and Lemma 3.4, T_{23} can be estimated by,

$$
\begin{aligned}
T_{23}&:=E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}(X^{i,N,n}_r,\mu^{X,N,n}_r)-b^{(k)}_n(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\big\}dr\\
&\qquad\Big\langle\partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^i_r\Big\rangle\\
&\leq E\int_{\kappa_n(s)}^{s}\big|b(X^{i,N,n}_r,\mu^{X,N,n}_r)-b_n(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\big|\\
&\qquad\Big|\int_{\kappa_n(s)}^{s}\big(1+\big|X^{i,N,n}_{\kappa_n(s)}\big|\big)^{\rho/2+1}\tilde{\sigma}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^i_r\Big|dr\\
&\leq\int_{\kappa_n(s)}^{s}\big\{E\big|b(X^{i,N,n}_r,\mu^{X,N,n}_r)-b_n(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\big|^2\big\}^{\frac{1}{2}}\\
&\qquad\Big\{E\Big|\int_{\kappa_n(s)}^{s}\big(1+\big|X^{i,N,n}_{\kappa_n(s)}\big|\big)^{\rho/2+1}\tilde{\sigma}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^i_r\Big|^2\Big\}^{\frac{1}{2}}dr\\
&\leq Kn^{-3/2}\Big\{E\int_{\kappa_n(s)}^{s}\big(1+\big|X^{i,N,n}_{\kappa_n(s)}\big|\big)^{\rho+2}\big|\tilde{\sigma}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^2dr\Big\}^{\frac{1}{2}}dr\leq Kn^{-2}
\end{aligned}
\tag{4.9}
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$.

For estimating T_{24}, one applies Cauchy–Schwarz inequality, Young's inequality and Remark 2.8 to get,

$$
\begin{aligned}
T_{24}&:=E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}(X^{i,N}_r,\mu^{X,N}_r)-\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N}_r)\big\}dW^{(l),i}_r\\
&\qquad\Big\langle\partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^i_r\Big\rangle\\
&=E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}(X^{i,N}_r,\mu^{X,N}_r)-\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N}_r)\big\}dW^{(l),i}_r\\
&\qquad\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\Big\langle\partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big),\tilde{\sigma}^{(l_1)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\Big\rangle dW^{(l_1),i}_r\\
&=E\sum_{k=1}^{d}\sum_{l=1}^{d}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}(X^{i,N}_r,\mu^{X,N}_r)-\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N}_r)\big\}\\
&\qquad\Big\langle\partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big),\tilde{\sigma}^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\Big\rangle dr\\
&\leq KE\sum_{k=1}^{d}\sum_{l=1}^{d}\int_{\kappa_n(s)}^{s}n^{1/2}\big|\sigma^{(k,l)}(X^{i,N}_r,\mu^{X,N}_r)-\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N}_r)\big|
\end{aligned}
$$

$$
\begin{aligned}
&n^{-1/2}\big|\partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big)\big|\big|\tilde\sigma^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|dr\\
\leq{}& KE\int_{\kappa_n(s)}^s n\big|X^{i,N}_r-X^{i,N,n}_r\big|^2dr\\
&+KE\sum_{l=1}^d\int_{\kappa_n(s)}^s n^{-1}\big(1+\big|X^{i,N,n}_{\kappa_n(r)}\big|\big)^{\rho+2}\big|\tilde\sigma^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^2dr
\end{aligned}
$$

which on using Hölder's inequality, Corollary 4.1 and Lemma 3.4 yields,

$$
T_{24}\leq K\sup_{i\in\{1,\cdots,N\}}\sup_{r\in[0,s]}E\big|X^{i,N}_r-X^{i,N,n}_r\big|^2+Kn^{-2} \tag{4.10}
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$.

Again notice that,

$$
\begin{aligned}
T_{25}:={}&E\sum_{k=1}^d\sum_{l=1}^m\int_{\kappa_n(s)}^s\big\{\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N}_r)-\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N,n}_r)\big\}dW^{(l),i}_r\\
&\Big\langle\partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^s\tilde\sigma\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^i_r\Big\rangle\\
={}&E\sum_{k=1}^d\sum_{l=1}^m\int_{\kappa_n(s)}^s\big\{\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N}_r)-\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N,n}_r)\big\}dW^{(l),i}_r\\
&\sum_{l_1=1}^m\int_{\kappa_n(s)}^s\Big\langle\partial_x b^{(k)}\Big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\Big),\tilde\sigma^{(l_1)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\Big\rangle dW^{(l_1),i}_r\\
={}&E\sum_{k=1}^d\sum_{l=1}^m\int_{\kappa_n(s)}^s\big\{\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N}_r)-\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N,n}_r)\big\}\\
&\Big\langle\partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big),\tilde\sigma^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\Big\rangle dr
\end{aligned}
$$

which by applying Cauchy–Schwarz inequality, Assumption 2.3, Remark 2.8, Young's inequality, Hölder's inequality, Lemma 3.4 and Corollary 4.1 gives,

$$
\begin{aligned}
T_{25}\leq{}&KE\sum_{l=1}^m\int_{\kappa_n(s)}^s\mathcal{W}_2\big(\mu^{X,N}_r,\mu^{X,N,n}_r\big)\big(1+\big|X^{i,N,n}_{\kappa_n(s)}\big|\big)^{\rho/2+1}\big|\tilde\sigma^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|dr\\
\leq{}&KE\sum_{l=1}^m\int_{\kappa_n(s)}^s n^{\frac12}\Big\{\frac1N\sum_{j=1}^N\big|X^{i,N}_r-X^{i,N,n}_r\big|^2\Big\}^{\frac12}\\
&n^{-\frac12}\big(1+\big|X^{i,N,n}_{\kappa_n(s)}\big|\big)^{\rho/2+1}\big|\tilde\sigma^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|dr\\
\leq{}&KE\sum_{l=1}^m\int_{\kappa_n(s)}^s n\frac1N\sum_{j=1}^N\big|X^{i,N}_r-X^{i,N,n}_r\big|^2dr\\
&+KE\sum_{l=1}^m\int_{\kappa_n(s)}^s n^{-1}\big(1+\big|X^{i,N,n}_{\kappa_n(s)}\big|\big)^{\rho+2}\big|\tilde\sigma^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^2dr\\
\leq{}&K\sup_{i\in\{1,\cdots,N\}}\sup_{r\in[0,s]}E\big|X^{i,N}_r-X^{i,N,n}_r\big|^2+Kn^{-2}
\end{aligned} \tag{4.11}
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$.

Proceeding as before, by Cauchy–Schwarz inequality and Remark 2.8,

$$
T_{26}:=E\sum_{k=1}^d\sum_{l=1}^m\int_{\kappa_n(s)}^s\big\{\sigma^{(k,l)}(X^{i,N,n}_r,\mu^{X,N,n}_r)-\tilde\sigma^{(k,l)}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\big\}dW^{(l),i}_r
$$

$$
\begin{aligned}
&\Big\langle \partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)dW^{i}_{r}\Big\rangle\\
&=E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}(X^{i,N,n}_{r},\mu^{X,N,n}_{r})-\tilde{\sigma}^{(k,l)}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\big\}dW^{(l),i}_{r}\\
&\quad\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\Big\langle \partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big),\tilde{\sigma}^{(l_1)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\Big\rangle dW^{(l_1),i}_{r}\\
&=E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}(X^{i,N,n}_{r},\mu^{X,N,n}_{r})-\tilde{\sigma}^{(k,l)}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\big\}\\
&\quad\Big\langle \partial_x b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)}\big),\tilde{\sigma}^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\Big\rangle dr\\
&\le KE\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big|\sigma^{(k,l)}(X^{i,N,n}_{r},\mu^{X,N,n}_{r})-\tilde{\sigma}^{(k,l)}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\big|\\
&\quad\big(1+|X^{i,N,n}_{\kappa_n(s)}|\big)^{\frac{\rho}{2}}\big|\tilde{\sigma}^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|dr
\end{aligned}
$$

which on using Hölder's inequality, Lemma 4.4, Lemma 3.4 and Corollary 4.1 gives,

$$
\begin{aligned}
T_{26}&\le KE\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{E\big|\sigma^{(k,l)}(X^{i,N,n}_{r},\mu^{X,N,n}_{r})d-\tilde{\sigma}^{(k,l)}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\big|^{2}\big\}^{\frac{1}{2}}dr\\
&\quad\big\{E\big(1+|X^{i,N,n}_{\kappa_n(s)}|\big)^{2\rho}\big\}^{\frac{1}{4}}\big\{E\big|\tilde{\sigma}^{(l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^{4}\big\}^{\frac{1}{4}}dr\le Kn^{-2} \qquad (4.12)
\end{aligned}
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$. Substituting values from (4.7) to (4.12) in (4.6) gives,

$$
T_2\le \sup_{i\in\{1,\cdots,N\}}\sup_{r\in[0,s]}E|X^{i,N}_{r}-X^{i,N,n}_{r}|^{2}+Kn^{-2} \qquad (4.13)
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$.

Finally, we proceed with estimation of T_3 of equation (4.4). By Young's inequality and Cauchy–Schwarz inequality,

$$
\begin{aligned}
T_3&:=E\sum_{k=1}^{d}\big(X^{(k),i,N}_{s}-X^{(k),i,N,n}_{s}\big)\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),X^{j,N,n}_{s}-X^{j,N,n}_{\kappa_n(s)}\Big\rangle\\
&=E\sum_{k=1}^{d}\big(X^{(k),i,N}_{s}-X^{(k),i,N,n}_{s}\big)\\
&\quad\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}b_n(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dr\Big\rangle\\
&\quad+E\sum_{k=1}^{d}\big(X^{(k),i,N}_{s}-X^{(k),i,N,n}_{s}\big)\\
&\quad\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^{i}_{r}\Big\rangle\\
&\le KE\big|X^{i,N}_{s}-X^{i,N,n}_{s}\big|^{2}+K\sum_{k=1}^{d}E\frac{1}{N^2}N\sum_{j=1}^{N}\big|\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big)\big|^{2}\\
&\qquad\Big|\int_{\kappa_n(s)}^{s}b_n(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dr\Big|^{2}+E\sum_{k=1}^{d}\big(X^{(k),i,N}_{\kappa_n(s)}-X^{(k),i,N,n}_{\kappa_n(s)}\big)
\end{aligned}
$$

$$
\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle
$$

$$
+E\sum_{k=1}^{d}\big(X^{(k),i,N}_s-X^{(k),i,N,n}_s-X^{(k),i,N}_{\kappa_n(s)}+X^{(k),i,N,n}_{\kappa_n(s)}\big)
$$

$$
\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$. Notice that third term on the right hand side of the above expression is zero. Further, by Remark 2.8, Remark 2.9 and Lemma 3.4, one obtains

$$
T_3\le KE\big|X^{i,N}_s-X^{i,N,n}_s\big|^2+Kn^{-2}
$$

$$
+E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}\big(X^{i,N}_r,\mu^{X,N}_r\big)-b^{(k)}_n\big(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big\}dr
$$

$$
\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle
$$

$$
+E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X^{i,N}_r,\mu^{X,N}_r\big)-\tilde{\sigma}^{(k,l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big\}dW^{(l),i}_r
$$

$$
\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$. The above expression can further be written as,

$$
T_3\le KE\big|X^{i,N}_s-X^{i,N,n}_s\big|^2+Kn^{-2}+E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}\big(X^{i,N}_r,\mu^{X,N}_r\big)-b^{(k)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)\big\}dr
$$

$$
\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle
$$

$$
+E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)-b^{(k)}\big(X^{i,N,n}_r,\mu^{X,N,n}_r\big)\big\}dr
$$

$$
\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle
$$

$$
+E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}\big(X^{i,N,n}_r,\mu^{X,N,n}_r\big)-b^{(k)}_n\big(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big\}dr
$$

$$
\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle
$$

$$
+E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X^{i,N}_r,\mu^{X,N}_r\big)-\sigma^{(k,l)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)\big\}dW^{(l),i}_r
$$

$$
\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle
$$

$$
+E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)-\sigma^{(k,l)}\big(X^{i,N,n}_r,\mu^{X,N,n}_r\big)\big\}dW^{(l),i}_r
$$

$$
\begin{aligned}
&\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle\\
&+E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X^{i,N,n}_r,\mu^{X,N,n}_r\big)-\tilde{\sigma}^{(k,l)}\big(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big\}dW^{(l),i}_r\\
&\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle\\
&=:KE\big|X^{i,N}_s-X^{i,N,n}_s\big|^2+Kn^{-2}+T_{31}+T_{32}+T_{33}+T_{34}+T_{35}+T_{36}
\end{aligned}
\tag{4.14}
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$.

By Cauchy–Schwarz inequality, Assumption 2.2, Remark 2.8 and Young's inequality,

$$
\begin{aligned}
T_{31}:=&E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}\big(X^{i,N}_r,\mu^{X,N}_r\big)-b^{(k)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)\big\}dr\\
&\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle\\
\le& KE\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}n^{\frac12}\big(1+|X^{i,N}_r|+|X^{i,N,n}_r|\big)^{\rho/2+1}\big|X^{i,N}_r-X^{i,N,n}_r\big|\\
&n^{-\frac12}\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big)\big|\Big|\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big|dr\\
\le& KE\int_{\kappa_n(s)}^{s}n\big|X^{i,N}_r-X^{i,N,n}_r\big|^2dr+KE\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}n^{-1}\big(1+|X^{i,N}_r|+|X^{i,N,n}_r|\big)^{\rho+2}\\
&\frac{1}{N}\sum_{j=1}^{N}\big(\big(1+|X^{i,N,n}_{\kappa_n(s)}|\big)^{\rho+4}+1+|X^{j,N,n}_{\kappa_n(s)}|\big)^2\Big|\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big|^2dr
\end{aligned}
$$

which on using Hölder's inequality, Corollary 4.1 and Lemma 3.4 gives,

$$
T_{31}\le K\sup_{r\in[0,s]}E\big|X^{i,N}_r-X^{i,N,n}_r\big|^2+Kn^{-3}
\tag{4.15}
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$.

Moreover, by using Assumption 2.2 and Cauchy–Schwarz inequality, Hölder's inequality and Remark 2.8, one obtains,

$$
\begin{aligned}
T_{32}:=&E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)-b^{(k)}\big(X^{i,N,n}_r,\mu^{X,N,n}_r\big)\big\}dr\\
&\frac{1}{N}\sum_{j=1}^{N}\Big\langle \partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle\\
\le& KE\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}n^{\frac12}\mathcal{W}_2\big(\mu^{X,N}_r,\mu^{X,N,n}_r\big)\\
&n^{-\frac12}\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big)\big|\Big|\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big|dr\\
\le& KE\int_{\kappa_n(s)}^{s}n\mathcal{W}_2\big(\mu^{X,N}_r,\mu^{X,N,n}_r\big)^2dr
\end{aligned}
$$

$$+ KE\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s} n^{-1}\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big)\big|^2$$

$$\Big|\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big|^2 dr \leq KE\int_{\kappa_n(s)}^{s} n\frac{1}{N}\sum_{j=1}^{N}|X^{j,N}_r - X^{j,N,n}_r|^2 dr + Kn^{-3}$$

$$\leq \sup_{i\{1,\dots,N\}}\sup_{r\in[0,s]} E|X^{j,N}_r - X^{j,N,n}_r|^2 + Kn^{-3}$$

for any $s \in [0,T]$ and $n, N \in \mathbb{N}$.

To estimate T_{33}, we use Cauchy–Schwarz inequality, Hölder's inequality, Lemma 4.5, Remark 2.8 and Corollary 4.1 to obtain,

$$T_{33} := E\sum_{k=1}^{d}\int_{\kappa_n(s)}^{s}\big\{b^{(k)}\big(X^{i,N,n}_r,\mu^{X,N,n}_r\big) - b^{(k)}_n\big(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big\}dr$$

$$\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle$$

$$\leq \sum_{k=1}^{d}\frac{1}{N}\sum_{j=1}^{N}\int_{\kappa_n(s)}^{s}\big\{E\big|b^{(k)}\big(X^{i,N,n}_r,\mu^{X,N,n}_r\big) - b^{(k)}_n\big(X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)}\big)\big|^2\big\}^{\frac{1}{2}}$$

$$\Big\{E\big|\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big)\big|^2\Big|\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big|^2\Big\}^{\frac{1}{2}} dr \leq Kn^{-2}$$

for any $s \in [0,T]$ and $n, N \in \mathbb{N}$.

Notice that,

$$T_{34} := E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X^{i,N}_r,\mu^{X,N}_r\big) - \sigma^{(k,l)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)\big\}dW^{(l),i}_r$$

$$\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})dW^i_r\Big\rangle$$

$$= E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X^{i,N}_r,\mu^{X,N}_r\big) - \sigma^{(k,l)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)\big\}dW^{(l),i}_r$$

$$\frac{1}{N}\sum_{j=1}^{N}\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\Big\langle\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\tilde{\sigma}^{(l_1)}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\Big\rangle dW^{(l_1),i}_r$$

$$= E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X^{i,N}_r,\mu^{X,N}_r\big) - \sigma^{(k,l)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)\big\}$$

$$\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big),\tilde{\sigma}^{(l)}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\Big\rangle dr$$

which on the application of Cauchy–Schwarz inequality, Young's inequality, Remark 2.8 and Corollary 4.1 gives,

$$T_{34} \leq E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big|\sigma^{(k,l)}\big(X^{i,N}_r,\mu^{X,N}_r\big) - \sigma^{(k,l)}\big(X^{i,N,n}_r,\mu^{X,N}_r\big)\big|$$

$$\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu b^{(k)}\big(X^{i,N,n}_{\kappa_n(s)},\mu^{X,N,n}_{\kappa_n(s)},X^{j,N,n}_{\kappa_n(s)}\big)\big|\big|\tilde{\sigma}^{(l)}(r,X^{i,N,n}_{\kappa_n(r)},\mu^{X,N,n}_{\kappa_n(r)})\big|dr$$

$$
\begin{aligned}
&\leq KE\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s} n^{\frac{1}{2}}\big|X_r^{i,N}-X_r^{i,N,n}\big| \\
&n^{-\frac{1}{2}}\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big)\big|\big|\tilde{\sigma}^{(l)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\big|dr \\
&\leq KE\int_{\kappa_n(s)}^{s} n\big|X_r^{i,N}-X_r^{i,N,n}\big|^2 dr \\
&\quad + KE\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\frac{n^{-1}}{N}\sum_{j=1}^{N}\big|\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big)\big|^2\big|\tilde{\sigma}^{(l)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\big|^2 dr \\
&\leq K\sup_{i\in\{1,\ldots,N\}}\sup_{r\in[0,s]}E\big|X_r^{i,N}-X_r^{i,N,n}\big|^2+Kn^{-2}
\end{aligned}
\tag{4.16}
$$

for any $s\in[0,T]$ and $n,N\in\mathbb{N}$.

Also notice that,

$$
\begin{aligned}
T_{35} &:= E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N}\big)-\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N,n}\big)\big\}dW_r^{(l),i} \\
&\quad \frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})dW_r^i\Big\rangle \\
&= E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N}\big)-\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N,n}\big)\big\}dW_r^{(l),i} \\
&\quad \frac{1}{N}\sum_{j=1}^{N}\sum_{l_1=1}^{m}\Big\langle\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}^{(l_1)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\Big\rangle dW_r^{(l_1),i} \\
&= E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N}\big)-\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N,n}\big)\big\} \\
&\quad \frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}^{(l)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\Big\rangle dr
\end{aligned}
$$

and then one uses Assumption 2.3, Remark 2.8, Cauchy–Schwarz inequality, Young's inequality, Hölder's inequality, Corollary 4.1 and Lemma 3.4 to obtain,

$$
\begin{aligned}
T_{35} &\leq KE\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s} n^{\frac{1}{2}}\mathcal{W}_2\big(\mu_r^{X,N},\mu_r^{X,N,n}\big) \\
&n^{-\frac{1}{2}}\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big)\big|\Big|\int_{\kappa_n(s)}^{s}\tilde{\sigma}^{(l)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\Big|dr \\
&\leq KE\int_{\kappa_n(s)}^{s} n\frac{1}{N}\sum_{j=1}^{N}|X_r^{i,N}-X_r^{i,N,n}|^2 dr \\
&+ K\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s} n^{-1}\frac{1}{N^2}N\sum_{j=1}^{N}\big\{E\big|\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big)\big|^4\big\}^{\frac{1}{4}} \\
&\Big\{E\Big|\int_{\kappa_n(s)}^{s}\tilde{\sigma}^{(l)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\Big|^4\Big\}^{\frac{1}{4}}dr \leq K\sup_{i\in\{1,\cdots,N\}}\sup_{r\in[0,s]}E|X_r^{i,N}-X_r^{i,N,n}|^2+Kn^{-2}
\end{aligned}
\tag{4.17}
$$

for any $s \in [0,T]$ and $n, N \in \mathbb{N}$.

Similarly,

$$T_{36} := E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N,n}\big)-\tilde{\sigma}^{(k,l)}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)\big\}dW_r^{(l),i}$$
$$\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big),\int_{\kappa_n(s)}^{s}\tilde{\sigma}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})dW_r^i\Big\rangle$$
$$=E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N,n}\big)-\tilde{\sigma}^{(k,l)}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)\big\}dW_r^{(l),i}$$
$$\frac{1}{N}\sum_{j=1}^{N}\sum_{l_1=1}^{m}\int_{\kappa_n(s)}^{s}\Big\langle\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big),\tilde{\sigma}^{(l_1)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\Big\rangle dW_r^{(l_1),i}$$
$$=E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N,n}\big)-\tilde{\sigma}^{(k,l)}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)\big\}$$
$$\frac{1}{N}\sum_{j=1}^{N}\Big\langle\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big),\tilde{\sigma}^{(l)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\Big\rangle dr$$

and then using Hölder's inequality, Cauchy–Schwarz inequality, Remark 2.8, Lemma 4.4 and Corollary 4.1, one obtains,

$$T_{36} \leq E\sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big|\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N,n}\big)-\tilde{\sigma}^{(k,l)}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)\big|$$
$$\frac{1}{N}\sum_{j=1}^{N}\big|\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big)\big|\big|\tilde{\sigma}^{(l)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\big|dr$$
$$\leq \sum_{k=1}^{d}\sum_{l=1}^{m}\int_{\kappa_n(s)}^{s}\big\{E\big|\sigma^{(k,l)}\big(X_r^{i,N,n},\mu_r^{X,N,n}\big)-\tilde{\sigma}^{(k,l)}\big(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n}\big)\big|^2\big\}^{\frac{1}{2}}$$
$$\frac{1}{N}\sum_{j=1}^{N}\big\{E\big|\partial_\mu b^{(k)}\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n},X_{\kappa_n(s)}^{j,N,n}\big)\big|^4\big\}^{\frac{1}{4}}\big\{E\big|\tilde{\sigma}^{(l)}(r,X_{\kappa_n(r)}^{i,N,n},\mu_{\kappa_n(r)}^{X,N,n})\big|^4\big\}^{\frac{1}{4}}dr$$
$$\leq Kn^{-2} \tag{4.18}$$

for any $s \in [0,T]$ and $n, N \in \mathbb{N}$. Substituting values from (4.15) to (4.18) in equation (4.14), we get

$$T_3 \leq K \sup_{i\in\{1,\cdots,N\}}\sup_{r\in[0,s]} E|X_r^{i,N}-X_r^{i,N,n}|^2 + Kn^{-2} \tag{4.19}$$

for any $s \in [0,T]$ and $n, N \in \mathbb{N}$. Substituting estimates from (4.5), (4.13) and (4.19) in equation (4.4) completes the proof. ☐

Proof of Proposition 2.2

Recall equations (2.2) and (2.4) and use Itô's formula to obtain,

$$\big|X_t^{i,N}-X_t^{i,N,n}\big|^2 = 2\int_0^t\big\langle X_s^{i,N}-X_s^{i,N,n}, b\big(X_s^{i,N},\mu_s^{X,N}\big)-b_n\big(X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big\rangle ds$$
$$+\int_0^t\big\langle X_s^{i,N}-X_s^{i,N,n},\big\{\sigma\big(X_s^{i,N},\mu_s^{X,N}\big)-\tilde{\sigma}\big(s,X_{\kappa_n(s)}^{i,N,n},\mu_{\kappa_n(s)}^{X,N,n}\big)\big\}dW_s^i\big\rangle$$

$$+ \int_0^t \big|\sigma\big(X_s^{i,N}, \mu_s^{X,N}\big) - \tilde{\sigma}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)\big|^2 ds$$

almost surely for any $t \in [0,T]$ and $n, N \in \mathbb{N}$. Hence,

$$\begin{aligned}
E\big|X_t^{i,N} - X_t^{i,N,n}\big|^2 &\le 2E\int_0^t \big\langle X_s^{i,N} - X_s^{i,N,n}, b\big(X_s^{i,N}, \mu_s^{X,N}\big) - b\big(X_s^{i,N,n}, \mu_s^{X,N}\big)\big\rangle ds \\
&+ 2E\int_0^t \big\langle X_s^{i,N} - X_s^{i,N,n}, b\big(X_s^{i,N,n}, \mu_s^{X,N}\big) - b\big(X_s^{i,N,n}, \mu_s^{X,N,n}\big)\big\rangle ds \\
&+ 2E\int_0^t \big\langle X_s^{i,N} - X_s^{i,N,n}, b\big(X_s^{i,N,n}, \mu_s^{X,N,n}\big) - b_n\big(X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)\big\rangle ds \\
&+ KE\int_0^t \big|\sigma\big(X_s^{i,N}, \mu_s^{X,N}\big) - \sigma\big(X_s^{i,N,n}, \mu_s^{X,N}\big)\big|^2 ds \\
&+ KE\int_0^t \big|\sigma\big(X_s^{i,N,n}, \mu_s^{X,N}\big) - \sigma\big(X_s^{i,N,n}, \mu_s^{X,N,n}\big)\big|^2 ds \\
&+ KE\int_0^t \big|\sigma\big(X_s^{i,N,n}, \mu_s^{X,N,n}\big) - \tilde{\sigma}\big(s, X_{\kappa_n(s)}^{i,N,n}, \mu_{\kappa_n(s)}^{X,N,n}\big)\big|^2 ds
\end{aligned}$$

which on using Assumption 2.2, Assumption 2.3, Young's inequality, Lemma 4.4 and Lemma 4.6 yields,

$$E\big|X_t^{i,N} - X_t^{i,N,n}\big|^2 \le Kn^{-2} + E\int_0^t \big|X_s^{i,N} - X_s^{i,N,n}\big|^2 ds + KE\int_0^t \mathcal{W}_2\big(\mu_s^{X,N}, \mu_s^{X,N,n}\big)^2 ds$$

which further implies,

$$\begin{aligned}
\sup_{i\in\{1,\dots,N\}} \sup_{r\in[0,t]} E\big|X_r^{i,N} - X_r^{i,N,n}\big|^2 &\le Kn^{-2} \\
&+ \int_0^t \sup_{i\in\{1,\dots,N\}} \sup_{r\in[0,s]} E\big|X_r^{i,N} - X_r^{i,N,n}\big|^2 ds
\end{aligned}$$

for any $t \in [0,T]$ and $n, N \in \mathbb{N}$. Thus, Grönwall's inequality completes the proof.

References

[1] Acebrón, J.A., Bonilla, L.L., Vicente, C.J.P., Ritort, F. and Spigler, R. (2005). The Kuramoto model: A simple paradigm for synchronization phenomena. *Rev. Mod. Phys.* **77(1)**.

[2] Ambrosio, L., Gigli, N. and Savaré, G. (2008). *Gradient Flows in Metric Spaces and in the Space of Probability Measures.* 2nd ed., Birkhäuser, Basel. MR-2401600

[3] Baladron, J., Fasoli, D., Faugeras, O. and Touboul, J. (2012). Mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons. *J. Math. Neurosci.* **2(1)**:10. MR-2974499

[4] Bao, J., Reisinger, C., Ren, P. and Stockinger, W. (2021). First-order convergence of Milstein schemes for McKean-Vlasov equations and interacting particle systems. *Proc. R. Soc. A.* ***477**:20200258*, http://doi.org/10.1098/rspa.2020.0258. MR-4212406

[5] Beyn, W.-J., Isaak, E. and Kruse, R. (2017). Stochastic C-stability and B-consistency of explicit and implicit Milstein-type schemes. *J. Sci. Comput.* **70(3)** 1042–1077. MR-3608332

[6] Bolley, F., Cañizo, J.A. and Carrillo, J.A. (2011). Stochastic mean-field limit: non-Lipschitz forces and swarming. *Math. Models Methods Appl. Sci.* **21** 2179–2210. MR-2860672

[7] Bossy, M., Faugeras, O. and Talay, D. (2015). Clarification and complement to "mean-field description and propagation of chaos in networks of Hodgkin–Huxley and FitzHugh–Nagumo neurons". *J. Math. Neurosci.* **5(1)**:19. MR-3392551

[8] Bossy, M. and Talay, D. (1997). A stochastic particle method for the McKean–Vlasov and the Burgers equation. *Mathematics of Computation* **66(217)** 157–192. MR-1370849

[9] Cardaliaguet, P. (2013). *Notes on Mean-field games, notes from P. L. Lions lectures at Collège de France*. https://www.ceremade.dauphine.fr/cardalia/MFG100629.pdf.

[10] Carmona, R. and Delarue, F. (2018a). *Probabilistic theory of mean field games with applications I: Mean-field FBSDEs, control, and games*. Springer International Publishing, Switzerland. MR-3752669

[11] Carmona, R. and Delarue, F. (2018b). *Probabilistic theory of mean field games with applications II: Mean field games with common noise and master equations*. Springer International Publishing, Switzerland. MR-3753660

[12] Chassagneux, J.-F., Crisan, D. and Delarue, F. A Probabilistic approach to classical solutions of the master equation for large population equilibria. Forthcoming in *Mem. Amer. Math. Soc.*

[13] Chaudru de Raynal, P.-E. (2020). Strong well-posedness of McKean–Vlasov stochastic differential equation with Hölder drift. *Stochastic Process. Appl.* **130(1)** 79–107. MR-4035024

[14] Chen, X. and Reis, G. dos (2021). A flexible split-step scheme for MV–SDEs. arXiv:2105.09688 [math.NA].

[15] Dareiotis, K., Kumar, C. and Sabanis, S. (2016). On tamed Euler approximations of SDEs driven by Lévy noise with application to delay equations. *SIAM J. Numer. Anal.* **54(3)** 1840–1872. MR-3513865

[16] Dreyer, W., Gaberšček, M., Guhlke, C., Huth, R. and Jamnik, J. (2011). Phase transition in a rechargeable lithium battery. *Eur. J. Appl. Math.* **22(3)** 267–290. MR-2795141

[17] Erban, R., Haškovec, J. and Sun, Y. (2016). A Cucker–Smale Model with Noise and Delay. *SIAM J. Appl. Math.* **76(4)** 1535–1557. MR-3534479

[18] Giles, M.B. (2008). Multilevel Monte Carlo path simulation. *Oper. Res.* **56** 607–617. MR-2436856

[19] Guhlke, C., Gajewski, P., Maurelli, M., Friz, P.K. and Dreyer, W. (2018). Stochastic many-particle model for LFP electrodes. *Contin. Mech. Thermodyn.* **30(3)** 593–628. MR-3783112

[20] Ha, S.-Y., Ko, D., Min, C. and Zhang, X. (2020). Emergent collective behaviors of stochastic kuramoto oscillators. *Discrete Cont. Dyn. – B* **25(3)** 1059–1081. MR-4064020

[21] Hammersley, W., Šiška, D. and Szpruch, Ł. (2018). McKean–Vlasov SDEs under measure dependent Lyapunov conditions. arXiv:1802.03974.

[22] Hutzenthaler, M. and Jentzen, A. (2015). Numerical approximations of stochastic differential equations with non globally Lipschitz continuous coefficients. *Mem. Amer. Math. Soc.* **236** 1112. MR-3364862

[23] Hutzenthaler, M., Jentzen, A. and Kloeden, P. E. (2010). 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. MR-2795791

[24] Hutzenthaler, M., Jentzen, A. and Kloeden, P.E. (2012). Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab.* **22(4)** 1611–1641. MR-2985171

[25] Kumar, C., Neelima, Reisinger, C. and Stockinger, W. (2020). Well-posedness and tamed schemes for McKean–Vlasov Equations with Common Noise. arXiv:2006.00463 [math.PR].

[26] Kumar, C. and Sabanis, S. (2017a). On tamed Milstein scheme of SDEs driven by Lévy noise. *Discrete Contin. Dyn. Syst. Ser. B* **22(2)** 421–463. MR-3639123

[27] Kumar, C. and Sabanis, S. (2017b). On explicit approximations for Lévy driven SDEs with super-linear diffusion coefficients. *Electron. J. Probab.* **22(73)** 1–19. MR-3698742

[28] Kumar, C. and Sabanis, S. (2019). On Milstein approximations with varying coefficients: the case of super-linear diffusion coefficients. *BIT* **59(4)** 929–96. MR-4032893

[29] Leobacher, G., Reisinger, C. and Stockinger, W. (2020). Well-posedness and numerical schemes for McKean–Vlasov equations and interacting particle systems with discontinuous drift. arXiv:2006.14892v2 [math.PR].

[30] McKean Jr., H.P. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. U.S.A.* **56** 1907–1911. MR-0221595

[31] Mehri, S., Scheutzow, M. Stannat, W. and Zangeneh, B.Z. (2020). Propagation of chaos for stochastic spatially structured neuronal networks with delay driven by jump diffusions. *Ann. Appl. Probab.* **30(1)** 175–207. MR-4068309

[32] Mishura, Y.S. and Veretennikov, A.Y. (2020). Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. arXiv:1603.02212v12 [math.PR].

[33] Neelima, Biswas, S., Kumar, C., Reis, G. dos and Reisinger, C. (2020). Well-posedness and tamed Euler schemes for McKean-Vlasov equations driven by Lévy noise. arXiv:2010.08585 [math.PR].

[34] Reis, G. dos, Engelhardt, S. and Smith, G. (2021). Simulation of McKean–Vlasov SDEs with super linear Growth. *IMA Journal of Numerical Analysis*, doi:10.1093/imanum/draa099.

[35] Reis, G. dos, Salkeld, W. and Tugaut, J. (2019). Freidlin–Wentzell LDP in path space for McKean–Vlasov equations and the functional iterated logarithm law. *Ann. Appl. Probab.* **29(3)** 1487–1540. MR-3914550

[36] Reisinger, C. and Stockinger, W. (2021). An adaptive Euler-Maruyama scheme for McKean-Vlasov SDEs with super-linear growth and application to the mean-field FitzHugh-Nagumo model. arXiv:2005.06034v3 [math.NA]. MR-4293705

[37] Sabanis, S. (2013). A note on tamed Euler approximations. *Electron. Comm. Probab.* **18** 1–10. MR-3070913

[38] Sabanis, S. (2016). Euler approximations with varying coefficients: the case of superlinearly growing diffusion coefficients. *Ann. Appl. Probab.* **26(4)** 2083–2105. MR-3543890

[39] Sznitman, A.-S. (1991). *Topics in propagation of chaos.* Springer, Berlin, Heidelberg. MR-1108185

[40] Tien, D.N. (2013). A stochastic Ginzburg–Landau equation with impulsive effects. *Phys. A* **392** 1962–1971. MR-3038269

[41] Tretyakov, M.V. and Zhang, Z. (2013). A fundamental mean-square convergence theorem for SDEs with locally Lipschitz coefficients and its applications. *SIAM J. Numer. Anal.* **51** 3135–3162. MR-3129758

[42] Villani, C. (2009). *Optimal Transport: Old and New*. Springer, Berlin. MR-2914136

[43] Wang, X. and Gan, S. (2013). The tamed Milstein method for commutative stochastic differential equations with non-globally Lipschitz continuous coefficients. *J. Difference Equ. Appl.* **19(3)** 466–490. MR-3037286

Acknowledgments. We are thankful to the anonymous referees for several insightful comments and suggestions which improved the presentation of the manuscript.