Weak convergence analysis in the particle limit of the McKean–Vlasov equations using stochastic flows of particle systems

Abdul-Lateef Haji-Ali∗  Håkon Hoel†  Raúl Tempone ‡

August 7, 2024

Abstract

We present a proof showing that the weak error of a system of n interacting stochastic particles approximating the solution of the McKean–Vlasov equation is $O(n^{-1})$. Our proof is based on the Kolmogorov backward equation for the particle system and bounds on the derivatives of its solution which we derive more generally using the variations of the stochastic particle system. The convergence rate is verified by numerical experiments which also indicate that the assumptions made here and in the literature can be relaxed.

Keywords: Interacting stochastic particle systems, McKean–Vlasov, Stochastic mean-field limit, Weak convergence rates, stochastic flows.

AMS Classification: 65C05, 62P05

1 Introduction

For $a : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$, $\kappa_1, \kappa_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ and $\sigma : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d \times \mathbb{R}^d$, $t \geq 0$ and $\{W(s)\}_{s \geq 0}$ being a $d'$-dimensional Wiener process defined on some probability space with the natural filtration, we consider the following McKean–Vlasov equation:

$$Z(t) = \xi + \int_0^t a\left(Z(s), \int_{\mathbb{R}^d} \kappa_1(Z(s), z) \mu_s(dz)\right) ds$$

$$+ \int_0^t \sigma\left(Z(s), \int_{\mathbb{R}^d} \kappa_2(Z(s), z) \mu_s(dz)\right) dW(s),$$

(1)

where $\mu_s$ denotes the law of $Z(s)$ for all $s \geq 0$, and $\xi$ denotes a random initial state whose law is $\mu_0$ and is assumed to be independent of the Wiener process, $W$. We focus on one-dimensional interaction kernels $\kappa_1, \kappa_2 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ for clarity of presentation since high-dimensional kernels can be treated in a similar way; see Remark 2.6. For $a(x, y) = y$, with $\sigma$ being constant and $\kappa_1$ being bounded and Lipschitz, [24, Theorem I.1.1] shows the existence and uniqueness of a strong solution to (1). A recent analysis yielded the same result in [20] when the initial condition $\xi$ has a finite fourth moment, $a(x, y) = \sigma(x, y) = y$, under non-degeneracy conditions on $\kappa_2$, and when for all $x, x', y \in \mathbb{R}^d$, there exists a constant $C > 0$ such that

$$|\kappa_2(x, y)| + |\kappa_2(x', y)| \leq C(1 + \|x\|),$$

$$|\kappa_2(x, y) - \kappa_2(x', y)| \leq C (1 + \|y\|^2) \|x - x'\|.$$

The work [9] also shows the existence and uniqueness result for $\kappa_2 = 0$ and a particular form of $a$. Existence of weak solutions was also shown in [12] under certain measure-dependent Lyapunov conditions. In the current work, we do not focus on existence of solutions to (1) and instead

∗Maxwell Institute, Heriot-Watt University, Edinburgh, UK. (a.hajiali@hw.ac.uk).
†University of Oslo, Norway. (haakonah@math.uio.no).
‡RWTH Aachen University, Germany and King Abdullah University of Science & Technology (KAUST), Thuwal, Saudi Arabia (raul.tempone@kaust.edu.sa).
assume the existence of weak solutions and consider approximations of \( Z \) using a system of \( n \) Itô stochastic differential equations (SDEs), also known as an interacting stochastic particle system, with pairwise interaction kernels:

\[
X_i^n(t) = \xi_i + \int_0^t a \left( X_i^n(s), \frac{1}{n} \sum_{j=1}^n \kappa_1(X_i^n(s), X_j^n(s)) \right) \, ds \\
+ \int_0^t \sigma \left( X_i^n(s), \frac{1}{n} \sum_{j=1}^n \kappa_2(X_i^n(s), X_j^n(s)) \right) \, dW_i(s),
\]

for \( i \in \{1, 2, \ldots, n\} \), where \( \xi_i \) are i.i.d. and have the same law, \( \mu_0 \), and \( \{W_i(s)\}_{s \geq 0} \) are independent \( d \)-dimensional Wiener processes and independent of \( \{\xi_i\}_{i=1}^n \). In other words, the law \( \mu_t \) for \( t \geq 0 \) is approximated by an empirical measure based on the particles \( X^n(t) := \{X_i^n(t)\}_{i=1}^n \). It should be noted that these particles are identically distributed but not independent.

For \( Z^n := (Z_i)_{i=1}^n \) being \( n \) independent samples of the solution to the McKean–Vlasov equation \( g(Z^n(0)) = 0 \) and a function \( g : \mathbb{R}^{d \times n} \to \mathbb{R} \), the weak error at time \( t \) is defined as the absolute difference \( |E[g(X^n(t))] - E[g(Z^n(t))]| \). The weak error was established to be \( O(1/n) \) in, e.g., [17, Chapter 9] and [19, Theorem 6.1]. These works assume that \( \kappa_2 \equiv 0 \) and build upon semigroup theory in measure-valued function spaces to prove their results. On the other hand, the work [5] employs a similar methodology to the current work but assumes that \( \kappa_1 \) and \( \kappa_2 \) in (1) do not depend on the state \( Z \). There is an increasing interest in extending proofs of strong and weak convergence in more general settings with nonlinear drift/diffusion coefficients. To that end, works such as [25, Theorem B.2], [7, Theorem 2.17] and [10] use Lions-derivatives [6] to bound derivatives with respect to measures and a master equation for probability measures [18, 8]. For a more exhaustive literature review, see [7].

In Section 2, we present a new method to show the rate of weak convergence. The principal steps involve using the Kolmogorov backward equation to represent the weak error and the stochastic flows and the dual functions to bound the weights in the resulting dual weighted residual representation. Using the Kolmogorov backward equation to estimate the weak error in SDEs goes back to the ideas of Talay and Tubaro [26], who estimated the time discretization error for uniform deterministic time-steps. The works [1, 2], extended the analysis to approximations with non-smooth observables and the probability density of the solution at a given time. Kloeden and Platen [16] generalized the analysis in [26] to weak approximations of a higher order. Later, in a series of works inspired by [26], the authors developed methods based on stochastic flows and dual functions to bound the weights in the resulting dual-weighted residual representation. This approach provided the analysis for the weak approximation of SDEs using non-uniform, possibly stochastic, time-steps, see [23, 21, 22, 4]. Furthermore, the same analysis line was also used for adaptive Multilevel Monte Carlo, [14, 13].

The closest inspiration for deriving the weak convergence rate goes back to the use of the mentioned techniques in the context of multiscale approximation. Those works derived macroscopic SDEs continuum models by choosing their drift and diffusion functions to minimize the weak error in the given macroscopic observables when compared with a given base model. Particularly, [15] employed master equations with long-range interaction potentials as a base model as the stochastic Ising model with Glauber dynamics, whereas [27] determined the stochastic phase-field models from atomistic formulations by coarse-graining molecular dynamics to model the dendritic growth of a crystal in an under-cooled melt. Systems of coupled SDEs with increasing size could also be useful for approximating a non-Markovian behavior. For example, the recent work [3] investigated weak convergence rates for a rough stochastic volatility model emerging in mathematical finance, namely, the rough Bergomi model. As in this study, the analysis in [3] also employed a dual-weighted representation of the weak error, yielding an error expansion that characterizes the weak convergence rate. A similar method was also used in [5] for a special case of (1) in which \( \kappa_1 \) and \( \kappa_2 \) do not depend on the state \( Z \), i.e., \( \kappa(x, y) = \kappa(y) \). A key difference to our methodology is that in [5], the Kolmogorov backward equation involves \( \mu_t \), the law of \( Z(t) \) while we instead utilize the Kolmogorov backward equation for the particle system (2).

In Section 3 we prove the technical results that are needed for the preceding analysis. In particular, we determine sufficient conditions to bound derivatives of the solution to the Kolmogorov backward equation for a generic multidimensional SDE by bounding moments of the first, second,
and third variations of the SDE. Finally, in Section 4 we numerically study the weak error of a particle approximation to the solution of the McKean-Vlasov equation. In particular, we show numerically that the weak convergence rate is the same for an example stochastic particle system that does not satisfy the regularity conditions of our theory or those of others in the literature cited above. Therefore, further work is necessary to extend the current approaches.

In what follows, we will use the notation $A \lesssim B$ to denote that there is a constant $0 < c < \infty$ which is independent of $n$, the size of the particle system (2), such that $A \leq cB$. For a multi-index $\ell \in \mathbb{N}^n$, $n \in \mathbb{N}$, define the derivative

$$\frac{\partial^{|\ell|}}{\partial x^\ell} := \frac{\partial^{|\ell|}}{\partial x_1^{\ell_1} \cdots \partial x_n^{\ell_n}},$$

where $|\ell| = \sum_{i=1}^n \ell_i$. Let $\|\cdot\|$ denote the Euclidean norm and let $C(\mathbb{R}^n; \mathbb{R}^m)$ denote the space of continuous functions $u \equiv (u_i)_{i=1}^m : \mathbb{R}^n \to \mathbb{R}^m$ for which the extended norm

$$\|u\|_{C(\mathbb{R}^n; \mathbb{R}^m)} := \sum_{j=1}^m \sup_{x \in \mathbb{R}^n} |u_j(x)|$$

is finite. Let also $C(\mathbb{R}^n; \mathbb{R}) = C(\mathbb{R}^n)$.

When $u$ has continuous derivatives up to order $k$, define the extended semi-norm

$$|u|_{C^k(\mathbb{R}^n; \mathbb{R}^m)} := \sum_{j=1}^m \sum_{|\ell| \leq k} \left\| \frac{\partial^{|\ell|} u_i}{\partial x^{\ell}} \right\|_{C(\mathbb{R}^n; \mathbb{R}^m)}.$$

and the norm $\|u\|_{C^k(\mathbb{R}^n; \mathbb{R}^m)} = \|u\|_{C(\mathbb{R}^n; \mathbb{R}^m)} + |u|_{C^k(\mathbb{R}^n; \mathbb{R}^m)}$. For a vector $x \in \mathbb{R}^{d \times n}$, we denote its components as $x = (x_i)_{i=1, \ldots, n, j=1, \ldots, d} \in \mathbb{R}^{d \times n}$. Similarly for a function $u : \mathbb{R}^{d \times n} \to \mathbb{R}$, we will use the notation $\nabla u = \left( \frac{\partial u}{\partial x_{i,j}} \right)_{i=1, \ldots, n, j=1, \ldots, d}$ for its gradient.

## 2 A Bound on the Weak Error

In this section, we prove that the weak error as defined in the introduction is $O(1/n)$. We start by stating boundedness and convergence results involving only samples of $Z$, the solution to the McKean-Vlasov equation (1).

**Proposition 2.1.** Assume that weak solutions to (1) exist and let $\{Z_i\}_{i=1}^n$ be $n$ independent processes each satisfying (1) with independent underlying Wiener processes. Let $\kappa : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz continuous function, i.e., there exists a constant $C$ such that

$$|\kappa(x, y) - \kappa(x', y')| \leq C(\|x - x'\| + \|y - y'\|)$$

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

Let $p \in \{1, 2, \ldots\}$, then for any $i \in \{1, \ldots, n\}$, we have

$$E\left[ \frac{1}{n} \sum_{j=1}^n \kappa(Z_i(t), Z_j(t)) - \int_{\mathbb{R}^d} \kappa(Z_1(t), z) \mu_1(dz) \right]^{2p} \lesssim n^{-p} E\left[ \|Z(t)\|^{2p} \right].$$

Moreover, assuming that $a, \sigma, \kappa_1$ and $\kappa_2$ in (1) are Lipschitz continuous, we have

$$\sup_{0 \leq s \leq t} E[\|Z(s)\|^{2p}] \lesssim 1 + E[\|\xi\|^{2p}].$$

The hidden constants in (3) and (4) depend only on $d, p, t$, and the Lipschitz constants.

**Proof.** The moment boundedness result (4) with Lipschitz assumptions is classical [24] using Itô’s formula, Young’s and Grönwall’s inequalities. For $p = 1$, we set, without loss of generality, $i = 1$ and let

$$\Delta_1 \kappa := \int_{\mathbb{R}^d} \kappa(Z_1(t), z) \mu_1(dz) - \kappa(Z_1(t), Z_1(t)).$$
Expanding the square
\[
\mathbb{E} \left[ \left( \frac{1}{n} \sum_{j=1}^{n} \Delta_j \right)^2 \right] = \frac{1}{n^2} \sum_{j=1}^{n} \sum_{j'=1}^{n} \mathbb{E}[\Delta_j \Delta_{j'}].
\]

When \( j \neq j' \) and are both different from 1, we have
\[
\mathbb{E}[\Delta_j \Delta_{j'}] = \mathbb{E}[\Delta_j] \mathbb{E}[\Delta_{j'}] = 0,
\]

since, for a given \( Z_j(t) \), \( \Delta_j \) and \( \Delta_{j'} \) are conditionally independent and \( \mathbb{E}[\Delta_j | Z_j(t)] = 0 \) when \( j \neq 1 \). When \( j = j' \) or when \( 1 \in \{j, j'\} \), we bound using Hölder’s inequality and the Lipschitz assumption on \( \kappa \),
\[
|\mathbb{E}[\Delta_j \Delta_{j'}]| \leq \mathbb{E}[(\Delta_j)^2] \\
\leq \mathbb{E} \left[ \left( \int_{\mathbb{R}^d} C \|z - Z_j(t)\| \mu_t(dz) \right)^2 \right] \\
\leq 4C^2 \mathbb{E}[\|Z(t)\|^2].
\]

Substituting back yields the claim. The proof for \( p > 1 \) is fundamentally similar using the multinomial theorem. \( \square \)

**Lemma 2.2.** Assume that weak solutions to (1) exist and let \( Z^n = \{Z_i\}_{i=1}^n \) be \( n \) independent processes each satisfying (1) with independent underlying Wiener processes, \( \kappa : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \) be a Lipschitz continuous function and \( f : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R} \) be such that
\[
\left| \frac{\partial f(x, y)}{\partial y} \right| + \left| \frac{\partial^2 f(x, y)}{\partial y^2} \right| \leq C(1 + \|x\| + |y|), \tag{6}
\]

for all \( x \in \mathbb{R}^d \) and \( y \in \mathbb{R} \). Then, for any \( g \in C^1(\mathbb{R}^d \times \mathbb{R}) \) and any \( i \in \{1, \ldots, n\} \) we have
\[
\left| \mathbb{E} \left[ \left( f\left(Z_i(t), \int_{\mathbb{R}^d} \kappa(Z_i(t), z) \mu_t(dz) \right) - \int_{\mathbb{R}^d} \kappa(Z_i(t), Z_j(t)) g(Z^n(t)) \right) \mathbb{E}^n \right] \right| \lesssim n^{-1} \|g\|_{C^1(\mathbb{R}^d \times \mathbb{R})} (1 + \mathbb{E}[\|Z(t)\|^4]) \tag{7}
\]

**Proof.** Without loss of generality, we fix \( i = 1 \) and define
\[
\Delta f := f\left(Z_1(t), \int_{\mathbb{R}^d} \kappa(Z_1(t), z) \mu_t(dz) \right) - f\left(Z_1(t), \frac{1}{n} \sum_{j=1}^{n} \kappa(Z_1(t), Z_j(t)) \right).
\]

By Taylor expanding, we can bound
\[
\left| \mathbb{E}[\Delta f g(Z^n(t))] \right| \leq \left| \mathbb{E} \left[ \mathcal{F}(Z_1(t)) \left( \frac{1}{n} \sum_{j=1}^{n} \Delta_j \kappa \right) g(Z^n(t)) \right] \right| \\
+ \|g\|_{C^1(\mathbb{R}^d \times \mathbb{R})} \mathbb{E} \left[ \left| \mathcal{F}(Z_1(t)) \left( \frac{1}{n} \sum_{j=1}^{n} \Delta_j \kappa \right)^2 \right| \right], \tag{8}
\]

where \( \Delta_j \kappa \) is as defined in (5) and, for all \( x \in \mathbb{R}^d \),
\[
\mathcal{F}(x) := \frac{\partial f}{\partial y} \left( x, \int_{\mathbb{R}^d} \kappa(x, z) \mu_t(dz) \right),
\]

and
\[
\mathcal{F}(x) := \int_0^1 \frac{\partial^2 f}{\partial y^2} \left( x, s \left( \frac{1}{n} \sum_{j=1}^{n} \Delta_j \kappa \right) + \int_{\mathbb{R}^d} \kappa(x, z) \mu_t(dz) \right) (1 - s) \, ds.
\]
By (6) and $\kappa$ being Lipschitz continuous, implying linear growth, we have for $x \in \mathbb{R}^d$,

$$\|f(x)\| \lesssim 1 + \|x\| + E[\|Z(t)\|]$$

$$\|f(x)\| \lesssim 1 + \left| \frac{1}{n} \sum_{j=1}^n \Delta_j \kappa \right| + \|x\| + E[\|Z(t)\|] .$$

By Hölder’s inequality and (3), we have

$$E \left[ f(Z_1(t)) \left( \frac{1}{n} \sum_{j=1}^n \Delta_j \kappa \right)^2 \right] \lesssim \left( 1 + E[\|Z(t)\|] \right) E \left[ \left( \frac{1}{n} \sum_{j=1}^n \Delta_j \kappa \right)^2 \right] + \left( E \left[ \|Z(t)\|^2 \right] \right)^{3/4} \left( E \left[ \left( \frac{1}{n} \sum_{j=1}^n \Delta_j \kappa \right)^4 \right] \right)^{1/2}$$

$$+ \left( E \left[ \|Z(t)\|^2 \right] \right)^{1/2} \left( E \left[ \left( \frac{1}{n} \sum_{j=1}^n \Delta_j \kappa \right)^4 \right] \right)^{1/2}$$

$$\lesssim \left( 1 + E[\|Z(t)\|^4] \right) n^{-1} .$$

Furthermore,

$$\left| E \left[ f(Z_1(t)) \left( \frac{1}{n} \sum_{j=1}^n \Delta_j \kappa \right) g(Z^n(t)) \right] \right| \leq \frac{1}{n} \left| E \left[ f(Z_1(t)) (\Delta_1 \kappa) g(Z^n(t)) \right] \right|$$

$$+ \frac{1}{n} \sum_{j=2}^n \left| E \left[ f(Z_1(t)) \Delta_j \kappa g(Z^n(t)) \right] \right| .$$

(9)

Here by again using (6) and $\kappa$ being Lipschitz continuous we have,

$$\left| E \left[ f(Z_1(t)) (\Delta_1 \kappa) g(Z^n(t)) \right] \right| \leq \|g\|_{C(\mathbb{R}^{d\times n})} E \left[ |f(Z_1(t))| |\Delta_1 \kappa| \right]$$

$$\lesssim \|g\|_{C(\mathbb{R}^{d\times n})} \left( 1 + E[\|Z(t)\|^4] \right) ,$$

(10)

and for $j = \{2, \ldots, n\}$, using a Taylor expansion of $g$,

$$\left| E \left[ f(Z_1(t)) \Delta_j \kappa g(Z^n(t)) \right] \right| \leq \left| E \left[ f(Z_1(t)) \Delta_j \kappa g(Z^n_{-j}(t)) \right] \right|$$

$$+ \left| E \left[ f(Z_1(t)) \Delta_j \kappa (Z^n(t) - Z^n_{-j}(t))^T \int_0^1 \nabla g(sZ^n(t) - (1-s)Z^n_{-j}(t)) \, ds \right] \right| .$$

(11)

Here, $Z^n_{-j}(t) = (Z_1(t), \ldots, Z_{j-1}(t), 0, Z_{j+1}(t), \ldots, Z_n(t)) \in \mathbb{R}^{d\times n}$ is the same as $Z^n(t)$ but with the $j$’th entry replaced by 0. Note that

$$E \left[ f(Z_1(t)) \Delta_j \kappa g(Z^n_{-j}(t)) \right] = E \left[ f(Z_1(t)) \right] E \left[ \Delta_j \kappa \right] E \left[ \left| Z_1(t) \right| g(Z^n_{-j}(t)) \right] = 0 ,$$

(12)

as $Z_j(t)$ has law $\mu_t$ and is independent of $Z^n_{-j}$ and of $Z_1(t)$. Using that $\kappa$ is Lipschitz continuous, we bound

$$\left| E \left[ f(Z_1(t)) \Delta_j \kappa (Z^n(t) - Z^n_{-j}(t))^T \int_0^1 \nabla g(sZ^n(t) - (1-s)Z^n_{-j}(t)) \, ds \right] \right|$$

$$\leq \sum_{i=1}^d \left\| \frac{\partial g}{\partial x_{i,j}} \right\|_{C(\mathbb{R}^{d\times n})} E \left[ \left| f(Z_1(t)) \right| \left| \Delta_j \kappa \right| \left| Z_j(t) \right| \right]$$

$$\lesssim \sum_{i=1}^d \left\| \frac{\partial g}{\partial x_{i,j}} \right\|_{C(\mathbb{R}^{d\times n})} \left( 1 + E[\|Z(t)\|^4] \right) .$$

(13)

Substituting (12) and (13) into (11), and the result and (10) into (9), and then substituting the result into (8), we arrive at the claimed result.
We now state the main result of the paper, which will also depend on the technical bounds that will be derived in Section 3.

**Theorem 2.3 (Weak convergence result).** Assume that weak solutions to (1) exist and let \( Z^n = (Z_i^n)_{i=1}^n \) be \( n \) independent processes each satisfying (1) with independent underlying Wiener processes. Assume that

\[
|a|_{C^1_b(\mathbb{R}^d; \mathbb{R}^d)} + |\sigma|_{C^1_b(\mathbb{R}^d; \mathbb{R}^{d \times d})} + |\kappa_1|_{C^1_b(\mathbb{R}^d)} + |\kappa_2|_{C^1_b(\mathbb{R}^d)} < \infty ,
\]

and let \( X^n = (X_i^n)_{i=1}^n \) satisfy (2). Then for \( g : \mathbb{R}^{d \times n} \to \mathbb{R} \) with continuous bounded derivatives up to the third order and any \( T > 0 \), we have

\[
|E[g(X^n(T)) - g(Z^n(T))]| \leq (1 + E[\|\xi\|^4]) n^{-1/2} |g|_{C^3_b(\mathbb{R}^{d \times n})} .
\]

**Proof.** Let \( a \equiv (a_j)_j=1^d \) for \( a_j : \mathbb{R}^d \to \mathbb{R} \) and \( \Sigma \equiv (\Sigma_{j,j'})_{j,j'=1}^d := \sigma^T \sigma \) for \( \Sigma_{j,j'} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R} \). For \( x \equiv (x_i)_i=1^n \equiv (x_{i,j})_{i \in \{1, \ldots, n\}, j \in \{1, \ldots, d\}} \), define the operators (recall that \( \mu_t \) is the law of \( Z(t) \))

\[
\mathcal{L}_n = \sum_{i=1}^n \sum_{j=1}^d a_j \left( x_i, \frac{1}{n} \sum_{i'=1}^n \kappa_1(x_i, x_{i'}) \right) \frac{\partial}{\partial x_i,j} + \frac{1}{2} \sum_{j=1}^d \Sigma_{j,j'} \left( x_i, \frac{1}{n} \sum_{i'=1}^n \kappa_2(x_i, x_{i'}) \right) \frac{\partial^2}{\partial x_i,j \partial x_i,j'}
\]

\[
\mathcal{L}_\infty = \sum_{i=1}^n \sum_{j=1}^d a_j \left( x_i, \int_{\mathbb{R}^d} \kappa_1(x_i, z) \mu_t(dz) \right) \frac{\partial}{\partial x_i,j} + \frac{1}{2} \sum_{j=1}^d \Sigma_{j,j'} \left( x_i, \int_{\mathbb{R}^d} \kappa_2(x_i, z) \mu_t(dz) \right) \frac{\partial^2}{\partial x_i,j \partial x_i,j'}
\]

Consider the value function \( u \) satisfying the PDE

\[
\frac{\partial u}{\partial t}(t, x) + \mathcal{L}_n u(t, x) = 0, \quad 0 \leq t < T \text{ and } x \in \mathbb{R}^{d \times n}
\]

and \( u(T, x) = g(x) \) for \( x \in \mathbb{R}^{d \times n} \).

Under (14), a strong solution for (2) exists and is unique [11, Theorem 1.1 in Chapter 5] and \( u(t, x) = E[g(X^n(T)) | X^n(t) = x] \), see [11, Theorem 6.1 in Chapter 5]. Given the existence of a solution to (1) and its law, and recalling that the coefficients \( a \) and \( \sigma \) in (1) are integrable and square-integrable, respectively, due to (14) and (4), we define \( U(t) := u(t, Z^n(t)) \) and apply Itô’s formula [11, Theorem 5.3 in Chapter 4] to arrive at

\[
E[g(Z^n(T))] - E[g(X^n(T))] = E[U(T) - U(0)] = \int_0^T E[(\mathcal{L}_\infty - \mathcal{L}_n)u(t, Z^n(t))] dt .
\]

The last equality is satisfied under the boundedness conditions in (14), the integrability of \( Z \), and the boundedness of the derivatives of \( u \), which we will establish later. Then

\[
E[(\mathcal{L}_\infty - \mathcal{L}_n)u(t, Z^n(t))] = \sum_{i=1}^n \sum_{j=1}^d E \left[ \Delta_i a_j \frac{\partial u}{\partial x_i,j}(t, Z^n(t)) \right] + \frac{1}{2} \sum_{j=1}^d E \left[ \Delta_i \Sigma_{j,j'} \frac{\partial^2 u}{\partial x_i,j \partial x_i,j'}(t, Z^n(t)) \right],
\]

where for \( f \equiv a_j, \kappa \equiv \kappa_1 \) and \( f \equiv \Sigma_{j,j'}, \kappa \equiv \kappa_2 \), we define

\[
\Delta_i f := f \left( Z_i(t), \int_{\mathbb{R}^d} \kappa(Z_i(t), z) \mu_t(dz) \right) - f \left( Z_i(t), \frac{1}{n} \sum_{j=1}^n \kappa(Z_i(t), Z_j(t)) \right).
\]
Using the triangle inequality, Lemma 2.2 and (4), we bound
\[ |E[(\mathcal{L}_\infty - \mathcal{L}_n)u(t, Z^n(t))]| \leq n^{-1} (1 + E[\|Z(t)\|^4]) \left( \sum_{i=1}^n \sum_{j=1}^d \left\| \frac{\partial u(t, \cdot)}{\partial x_{i,j}} \right\|_{C^1(R^d \times \mathbb{R})} + \sum_{j=1}^d \left\| \frac{\partial^2 u(t, \cdot)}{\partial x_{i,j} \partial x_{i,j'}} \right\|_{C^1(R^d \times \mathbb{R})} \right) \]
\[ \leq n^{-1} (1 + E[\|\xi\|^4]) |u(t, \cdot)|_{C^1(R^d \times \mathbb{R})}. \]

It remains to show the bound $|u(t, \cdot)|_{C^1(R^d \times \mathbb{R})} \lesssim |g|_{C^4(R^d \times \mathbb{R})}$ for $t \leq T$. To that end, we use Proposition 3.3 in the following section with an appropriate definition of $\nu$ and $\zeta$ in terms of $a$ and $\sigma$, respectively, since Assumption 3.1 is satisfied for $q = 3$ given (14); see the discussion after Assumption 3.1.

**Corollary 2.4.** From the previous theorem, we can readily deduce that under the same conditions and for an integer $k \leq n$ and $g : \mathbb{R}^{k \times d} \to \mathbb{R}$, we have
\[ |E[g(X^n_1(T), X^n_2(T), \ldots, X^n_k(T)) - g(Z_1(T), Z_2(T), \ldots, Z_k(T))]| \leq \frac{1}{n} (1 + E[\|\xi\|^4]) \left( \frac{d k + 2}{3} \max_{\ell \in \mathbb{N}_{\max 1 \leq |\ell| \leq 3}} \left( \left\| \frac{\partial^\ell g}{\partial x^\ell} \right\|_{C^1(R^d \times \mathbb{R})} \right) \right). \]

Corollary 2.4 is useful when $k$ is independent of $n$. For example, consider for $\tilde{g} : \mathbb{R}^d \to \mathbb{R}$,
\[ g(X^n(T)) := \frac{1}{n} \sum_{i=1}^n \tilde{g}(X^n_i(T)), \]
and note that $|g|_{C^4_k(R^d \times \mathbb{R})} = |\tilde{g}|_{C^4_k(R^d)}$ and hence Theorem 2.3 implies that
\[ |E[g(X^n(T)) - g(Z^n(T))]| \lesssim n^{-1} |\tilde{g}|_{C^4_k(R^d)} = O(n^{-1}). \]

We would not get the same result if we apply Corollary 2.4 directly to $g$ with $k = n$. Instead, in this particular case, we may apply Corollary 2.4 to $\tilde{g}$ with $k = 1$ to see that $E[\tilde{g}(X^n_1(T)) - \tilde{g}(Z(T))] = O(n^{-1})$ and conclude by noting that
\[ E[g(X^n(T)) - g(Z^n(T))] = E[\tilde{g}(X^n_1(T)) - \tilde{g}(Z(T))]. \]

**Remark 2.5.** In the special case when $\kappa_2 = 0$, we can relax (14) and only assume that
\[ |\alpha|_{C^4_k(R^d \times \mathbb{R} \times \mathbb{R}^d)} + |\sigma|_{C^4_k(R^d \times \mathbb{R} \times \mathbb{R}^d)} + |\kappa_1|_{C^4_k(R^d \times \mathbb{R}^d)} < \infty. \]

The result would then also involve only the first and second derivatives of $g$,
\[ |E[g(X^n(T)) - g(Z^n(T))]| \lesssim n^{-1} (1 + E[\|\xi\|^4]) |g|_{C^4_k(R^d \times \mathbb{R})}, \]
thus recovering a similar result to the one obtained, for example, in [17, Chapter 9].

**Remark 2.6** (Multi-dimensional interaction kernels). The result of Theorem 2.3 can be extended to multi-dimensional kernels, $\kappa_1, \kappa_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^m$, for some integer $m$, assuming that
\[ |\kappa_1|_{C^4_k(R^d \times \mathbb{R}^d \times \mathbb{R}^d)} + |\kappa_2|_{C^4_k(R^d \times \mathbb{R}^d \times \mathbb{R}^d)} < \infty. \]

The proof would follow the same steps by adding and subtracting appropriate terms in (19) and treating each component separately.
3 Moments Bounds for SDE Variations with Sobolev-Bounded Coefficients

In this section, for $T > 0$ and $(t, x) \in [0, T] \times \mathbb{R}^n$, we consider a general SDE of the form

$$X^{t,x}(s) = x + \int_t^s \nu(\tau, X^{t,x}(\tau)) \, d\tau + \int_t^s \varsigma(\tau, X^{t,x}(\tau)) \, dW(\tau), \quad s \in [t, T],$$

(21)

with drift coefficient $\nu \equiv (\nu_i)_{i=1}^n : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$, a diffusion coefficient $\varsigma \equiv (\varsigma_{i,m})_{i \in \{1, \ldots, n\}, m \in \{1, \ldots, n'\}} : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times n'}$, and $W$ is a vector of $n'$ independent standard Wiener processes over a probability space $(\Omega, \mathcal{F}, P)$ with the natural filtration. The main results of this section are Lemma 3.2 and Proposition 3.3 with the latter being used in the final step of the proof of Theorem 2.3. Note that the system that we consider in Theorem 2.3 is (2) which is a specific example of (21) with $n \leftarrow nd$. However, we prove the results more generally for (21) to emphasize that the particular structure of (2) is irrelevant as long as Assumption 3.1 is satisfied. Additionally, the results in this section could be useful beyond the current work. In what follows, for any $f : [0, T] \times \mathbb{R}^n \to \mathbb{R}$, define the extended norm

$$\|f\|_\infty := \|f\|_{L^\infty([0,T]; C(\mathbb{R}^n))} := \sup_{0 \leq s \leq T} \|f(s, \cdot)\|_{C(\mathbb{R}^n)}.$$ 

For brevity of presentation, we will define for the coefficients $\nu$ and $\varsigma$ in (21) and for an integer $r \geq 0$,

$$\|\nu\|_{D^r, \infty} := \|\partial^r \nu\|_{D^{r-1}, \infty} + \sum_{\ell \in \mathbb{N}, |\ell| = r} \max_{i \in \{1, \ldots, n\}} \max_{s.t. \, \ell_i = 0} \left( \sum_{m=1}^{n'} \|\partial^r \varsigma_{im}\|_{\mathbb{R}^n} \right)^2 \frac{1}{2}$$

(22)

where $\partial^r \nu = \left( \frac{\partial^r \nu_i}{\partial x_i} \right)_{i=1}^n$ and $\|\partial^r \nu\|_{D^{r-1}, \infty} := 0$. Similarity, we define

$$\|\varsigma\|_{D^r, \ell_\infty} := \|\partial^r \varsigma\|_{D^{r-1}, \ell_\infty} + \sum_{\ell \in \mathbb{N}, |\ell| = r} \max_{i \in \{1, \ldots, n\}} \max_{s.t. \, \ell_i = 0} \left( \sum_{m=1}^{n'} \|\partial^r \varsigma_{im}\|_{\mathbb{R}^n} \right)^2 \frac{1}{2}$$

(23)

where $\partial^r \varsigma = \left( \frac{\partial^r \varsigma_{im}}{\partial x_i} \right)_{i \in \{1, \ldots, n\}, m \in \{1, \ldots, n'\}}$ and $\|\partial^r \varsigma\|_{D^{r-1}, \ell_\infty} := 0$. Finally, for the process $X^{t,x}$ in (21), we define for any $p \geq 1$,

$$\|X^{t,x}\|_{D^r, L^p([t,T]; L^p(\Omega, P))} := \|\partial X^{t,x}\|_{D^{r-1, L^p([t,T]; L^p(\Omega, P))}} + \sum_{\ell \in \mathbb{N}, |\ell| = r} \max_{i \in \{1, \ldots, n\}} \max_{s.t. \, \ell_i = 0} \sup_{t \leq s \leq T} \mathbb{E} \left[ \left| \frac{\partial^r X^{t,x}}{\partial x^\ell} (s) \right|^p \right]^{1/p}$$

(24)

where $\partial X^{t,x} = \left( \frac{\partial X^{t,x}}{\partial x_i} \right)_{i=1}^n$ and $\|X^{t,x}\|_{D^{r-1, L^p([t,T]; L^p(\Omega, P))}} := 0$.

Assumption 3.1 (Bounded derivatives). For an integer $q$ we assume that $\nu : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\varsigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times n'}$ satisfy

$$\sum_{r=1}^q \|\nu\|_{D^r, \infty} + \|\varsigma\|_{D^r, \ell_\infty} < C_q,$$

for some constant $C_q > 0$ independent of $n$.

The previous assumption deserves some explanation. For example, focusing on the drift coefficient $\nu$, the definition in (22) for $r = 1$ simplifies to

$$\|\nu\|_{D^1, \infty} = \max_i \left\| \frac{\partial \nu_i}{\partial x_i} \right\|_{\infty} + \sum_{\ell \neq \ell'} \max_{i \neq \ell} \left\| \frac{\partial \nu_i}{\partial x_\ell} \right\|_{\infty},$$
and \( \|\varsigma\|_{D^1,\ell_\infty} \) expands similarly. Hence, a sufficient condition for Assumption 3.1 when \( q = 1 \) is to bound
\[
\left( \sum_{m=1}^{n'} \left\| \frac{\partial \varsigma_{\text{sim}}}{\partial x^m} \right\|_2^2 \right)^{1/2} + \left\| \frac{\partial \nu_i}{\partial x\ell} \right\|_\infty \leq \begin{cases} \tilde{C}_1 \quad & i = \ell, \\ \tilde{C}_1 \quad & i \neq \ell, \end{cases}
\]
for all \( i \in \{1, \ldots, n\} \) and some constant \( \tilde{C}_1 > 0 \). For \( r = 2 \), the definition \( (22) \) simplifies to
\[
\|\nu\|_{D^2,\infty} = \max_i \left\| \frac{\partial^2 \nu_i}{\partial x_i x_i} \right\|_\infty + \sum_{i=1}^n \max_{i \neq \ell} \left\| \frac{\partial^2 \nu_i}{\partial x_i x_\ell} \right\|_\infty + \sum_{\ell=1}^n \sum_{r=1}^n \max_{i \neq \ell', \ell} \left\| \frac{\partial^2 \nu_i}{\partial x_\ell \partial x_{\ell'}} \right\|_\infty,
\]
and an additional condition on the second derivatives would be required Assumption 3.1 when \( q = 2 \), for example
\[
\left( \sum_{m=1}^{n'} \left\| \frac{\partial^r \varsigma_{\text{sim}}}{\partial x^m} \right\|_2^2 \right)^{1/2} + \left\| \frac{\partial^2 \nu_i}{\partial x_i x_\ell} \right\|_\infty \leq \begin{cases} \tilde{C}_2 \quad & i = \ell = \ell' \\ \tilde{C}_2 n^{-1} \quad & i = \ell \neq \ell' \\ \tilde{C}_2 n^{-2} \quad & i, \ell, \ell' \text{ are distinct}, \end{cases}
\]
for all \( i \in \{1, \ldots, n\} \) and some constant \( \tilde{C}_2 > 0 \). In general, for any integer \( q > 0 \), a sufficient condition for Assumption 3.1 is
\[
\left( \sum_{m=1}^{n'} \left\| \frac{\partial^q \varsigma_{\text{sim}}}{\partial x^m} \right\|_2^2 \right)^{1/2} + \left\| \frac{\partial^q \nu_i}{\partial x_\ell} \right\|_\infty \leq \tilde{C}_q n^{-|\ell|+e_\ell|_0} \quad \text{for all } i \in \{1, \ldots, n\} \text{ and } \ell \in \mathbb{N}^n : |\ell| \leq q,
\]
for a constant \( \tilde{C}_q > 0 \) and where \( e_i \) is the \( i \)th unit vector and \( |\ell|_0 \) denotes the number of non-zero elements of \( \ell \).

**Lemma 3.2 (\( L^p \) bound of stochastic flows).** Let Assumption 3.1 be satisfied for \( q \in \{1, 2, 3\} \) and for \( X^{t,x} \) in \( (21) \) and any \( p \geq 2 \) then there exists constants \( K_{q,p} \), independent of \( n \) and \( x \), such that
\[
\|X^{t,x}\|_{D^q,L^\infty((t,T);L^p(\Omega,\mathbb{P}))} \leq K_{q,p}.
\]

**Proof.** We first note the following inequality for any index sets \( \mathcal{I} \) and \( \mathcal{J} \) and any sequence \( \{a_{i,j}\}_{i \in \mathcal{I}, j \in \mathcal{J}} \),
\[
\left( \sum_{j \in \mathcal{J}} \left( \sum_{i \in \mathcal{I}} a_{i,j} \right)^2 \right)^{1/2} \leq \sum_{i \in \mathcal{I}} \left( \sum_{j \in \mathcal{J}} a_{i,j}^2 \right)^{1/2},
\]
which can be shown by expanding the square and using Hölder’s inequality. Using \( (25) \) and Jensen’s inequality, we can show the following inequality for any random variables \( \{Y_i\}_{i \in \mathcal{I}} \) and measurable sequence \( \{a_{i,j}\}_{i \in \mathcal{I}, j \in \mathcal{J}} \)
\[
E \left[ \left( \sum_{j \in \mathcal{J}} \left( \sum_{i \in \mathcal{I}} a_{i,j} Y_i \right)^2 \right)^{p/2} \right] \leq \left( \sum_{i \in \mathcal{I}} \left( \sum_{j \in \mathcal{J}} a_{i,j}^2 \right)^{1/2} \right)^p \max_{i \in \mathcal{I}} E[|Y_i|^p].
\]

In addition, note that for a positive sequence \( \{a_{i,j}\}_{i \in \mathcal{I}, j \in \mathcal{J}} \), we have
\[
\max_{i \in \mathcal{I}} \sum_{j \in \mathcal{J}} a_{i,j} \leq \max_{i \in \mathcal{I}} a_{i,i} + \sum_{j \in \mathcal{J}, j \neq i} a_{i,j} \leq \max_{i \in \mathcal{I}} a_{i,i} + \sum_{j \in \mathcal{I}, j \neq i} \max_{i \in \mathcal{J}} a_{i,j}
\]

**First variation** First, note that the process \( \frac{\partial X^{t,x}}{\partial x_j} \) exists under Assumption 3.1 for \( q = 1 \) and satisfies for \( s \geq t \) the SDE
\[
\frac{\partial X^{t,x}}{\partial x_j}(s) = \delta_{i,j} + \int_t^s \sum_{k=1}^n \frac{\partial \nu_i}{\partial x_k}(\tau, X^{t,x}(\tau)) \frac{\partial X^{t,x}}{\partial x_j}(\tau) \, d\tau
\]
\[
+ \frac{n'}{n} \sum_{m=1}^{n'} \int_t^s \frac{\partial \varsigma_{\text{sim}}}{\partial x_k}(\tau, X^{t,x}(\tau)) \frac{\partial X^{t,x}}{\partial x_j}(\tau) \, dW_m(\tau),
\]
For the term involving $\nu_t$, using Young’s inequality, we can bound

$$\begin{align*}
\mathbb{E}\left[ \left( \sum_{k=1}^{n} \left( \frac{\partial \nu_t}{\partial x_k} \left( \mathbb{E} \left[ \frac{\partial X_1^{t,x}(\tau)}{\partial x_j}(\tau) \right] \right) \right) \right)^p \right] \\
\leq 2^{p-1} \left\| \frac{\partial \nu_t}{\partial x_j} \right\|_{\infty}^{p} \mathbb{E}\left[ \left\| \frac{\partial X_1^{t,x}(\tau)}{\partial x_j}(\tau) \right\|_{\infty}^{p} \right] + 2^{p-1} \mathbb{E}\left[ \left( \sum_{k=1, k \neq j}^{n} \left\| \frac{\partial \nu_t}{\partial x_k} \left( \mathbb{E} \left[ \frac{\partial X_1^{t,x}(\tau)}{\partial x_j}(\tau) \right] \right) \right) \right) \right]^{p}
\end{align*}$$

For the term involving $\zeta_t$, using Young’s inequality and bounding the derivatives of $\zeta_t$,

$$\begin{align*}
\mathbb{E}\left[ \left( \sum_{k=1}^{n} \left( \frac{\partial \zeta_t}{\partial x_k} \left( \mathbb{E} \left[ \frac{\partial X_1^{t,x}(\tau)}{\partial x_j}(\tau) \right] \right) \right) \right) \right]^{p/2} \\
\leq \frac{2}{p} \mathbb{E}\left[ \left( \sum_{k=1, k \neq j}^{n} \left\| \frac{\partial \zeta_t}{\partial x_k} \right\|_{\infty}^{2} \left( \mathbb{E} \left[ \frac{\partial X_1^{t,x}(\tau)}{\partial x_j}(\tau) \right] \right)^{2} \right) \right]^{p/2} + \frac{p-2}{p} \mathbb{E}\left[ \left\| \frac{\partial X_1^{t,x}(\tau)}{\partial x_j}(\tau) \right\|_{\infty}^{p} \right].
\end{align*}$$

Using Jensen’s inequality and (26) and Assumption 3.1, we can further bound

$$\begin{align*}
\mathbb{E}\left[ \left( \sum_{m=1}^{n'} \left( \sum_{k=1}^{n} \left( \frac{\partial \zeta_t}{\partial x_k} \left( \mathbb{E} \left[ \frac{\partial X_1^{t,x}(\tau)}{\partial x_j}(\tau) \right] \right) \right) \right) \right)^{p/2} \right] \\
\leq 2^{p-1} \left( \sum_{m=1}^{n'} \left\| \frac{\partial \zeta_t}{\partial x_j} \right\|_{\infty}^{2} \mathbb{E}\left[ \left\| \frac{\partial X_1^{t,x}(\tau)}{\partial x_j}(\tau) \right\|_{\infty}^{p} \right] \right)^{p/2} + 2^{p-1} C_1^{p} \max_{k \neq j} \mathbb{E}\left[ \left\| \frac{\partial X_1^{t,x}(\tau)}{\partial x_j}(\tau) \right\|_{\infty}^{p} \right].
\end{align*}$$
Combining (30) with (29) and (32) with (31) and substituting into (28) and simplifying yield

\[
E \left[ \left| \frac{\partial X^{t,x}_i}{\partial x_j} (s) \right|^p \right] \leq \delta_{ij} + c_1 \int_t^s E \left[ \left| \frac{\partial X^{t,x}_i}{\partial x_j} (\tau) \right|^p \right] d\tau + c_2 C_1^p \int_t^s \max_{k \neq j} E \left[ \left| \frac{\partial X^{t,x}_k}{\partial x_j} (\tau) \right|^p \right] d\tau + c_2 a_{i,j}^p \int_t^s E \left[ \left| \frac{\partial X^{t,x}_{i,j}}{\partial x_j} (\tau) \right|^p \right] d\tau ,
\]

where

\[
c_1 := \frac{p(p-1)}{2}, \quad c_2 := p2^{p-1},
\]

and \( a_{i,j}^p := \left\| \frac{\partial \nu_i}{\partial x_j} \right\|_\infty^p + \left( \sum_{n=1}^{n'} \left\| \frac{\partial k_{im}}{\partial x_j} \right\|_\infty^2 \right)^{p/2} \),

so that \( \max_i a_{i,i} + \sum_{j=1}^n \max_{i \neq j} a_{i,j} \leq C_1 \) by Assumption 3.1. Then, taking the maximum over all \( i \) in (33), we arrive at

\[
\max_i E \left[ \left| \frac{\partial X^{t,x}_i}{\partial x_j} (s) \right|^p \right] \leq 1 + (c_1 + 2C_1^p c_2) \int_t^s \max_i E \left[ \left| \frac{\partial X^{t,x}_i}{\partial x_j} (\tau) \right|^p \right] d\tau ,
\]

and using Grönwall’s inequality yields

\[
\max_i E \left[ \left| \frac{\partial X^{t,x}_i}{\partial x_j} (s) \right|^p \right] \leq \exp((c_1 + 2C_1^p c_2)(s-t)) := D_1 ,
\]

for all \( t \leq s \leq T \). Taking the maximum over all \( i \neq j \) in (33), we arrive at

\[
\max_{i \neq j} E \left[ \left| \frac{\partial X^{t,x}_{i,j}}{\partial x_j} (s) \right|^p \right] \leq (c_1 + c_2 C_1^p) \int_t^s \max_{i \neq j} E \left[ \left| \frac{\partial X^{t,x}_{i,j}}{\partial x_j} (\tau) \right|^p \right] d\tau + c_2 \left( \max_{i \neq j} a_{i,j}^p \right) \int_t^s E \left[ \left| \frac{\partial X^{t,x}_{i,j}}{\partial x_j} (\tau) \right|^p \right] d\tau .
\]

Using Grönwall’s inequality and (34) yields

\[
\max_{i \neq j} E \left[ \left| \frac{\partial X^{t,x}_{i,j}}{\partial x_j} (s) \right|^p \right] \leq c_2 \left( \max_{i \neq j} a_{i,j}^p \right) \exp((c_1 + c_2 C_1^p)(s-t)) \int_t^s E \left[ \left| \frac{\partial X^{t,x}_{i,j}}{\partial x_j} (\tau) \right|^p \right] d\tau \leq c_2 \left( \max_{i \neq j} a_{i,j}^p \right) \exp((c_1 + c_2 C_1^p)(s-t)) \left( \int_t^s D_1 d\tau \right) := \widetilde{D}_1 \left( \max_{i \neq j} a_{i,j}^p \right) .
\]

Finally, using (34) and (35) we arrive at:

\[
\| X^{t,x} \|_{D^1, L^\infty([t,T]; L^p(\Omega, P))} = \max \sup_{t \leq s \leq T} E \left[ \left| \frac{\partial X^{t,x}_i}{\partial x_j} (s) \right|^p \right]^{1/p} + \sum_{j=1}^n \max_{i \neq j} \sup_{t \leq s \leq T} E \left[ \left| \frac{\partial X^{t,x}_{i,j}}{\partial x_j} (s) \right|^p \right]^{1/p} \leq D_1^{1/p} + \widetilde{D}_1^{1/p} \sum_{j=1}^n \max_{i \neq j} a_{i,j} \leq D_1^{1/p} + \widetilde{D}_1^{1/p} C_1 =: K_{1,p} .
\]
Second variation In this section, we simplify the presentation by using $D_2$ to denote constants depending only on $t, T, p$, and $C_2$ and independent of $n$. Observe that these constants might change their values from one line to the next. Again, note that the process $\left\{ \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(s) \right\}_{s \in [t,T]}$ exists under Assumption 3.1 for $q = 2$ and satisfies for $s \in [t,T]$ the SDE

$$
\frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(s) = \int_t^s \sum_{k=1}^{n} \frac{\partial v_k}{\partial x_k}(\tau, X^t_{x,\tau}(\tau)) \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(\tau) \, d\tau \\
+ \int_t^s \sum_{k=1}^{n} \sum_{k'=1}^{n} \frac{\partial^2 v_k}{\partial x_k \partial x_{k'}}(\tau, X^t_{x,\tau}(\tau)) \frac{\partial X^t_{i,x}}{\partial x_j}(\tau) \frac{\partial X^t_{i,x}}{\partial x_{j'}}(\tau) \, d\tau \\
+ \sum_{m=1}^{n'} \int_t^s \sum_{k=1}^{n} \frac{\partial \nu_m}{\partial x_k}(\tau, X^t_{x,\tau}(\tau)) \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(\tau) \, dW_m(\tau) \\
+ \sum_{m=1}^{n'} \int_t^s \sum_{k=1}^{n} \sum_{k'=1}^{n} \frac{\partial^2 \nu_m}{\partial x_k \partial x_{k'}}(\tau, X^t_{x,\tau}(\tau)) \frac{\partial X^t_{i,x}}{\partial x_j}(\tau) \frac{\partial X^t_{i,x}}{\partial x_{j'}}(\tau) \, dW_m(\tau),
$$

we refer to [11, Theorem 5.4 in Chapter 5]. By Itô’s formula,

$$
E\left[ \left\| \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(s) \right\|_p^p \right] \leq p \int_t^s E\left[ (f_1 + f_3) \left\| \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(\tau) \right\|_{p-1}^{p-1} \right] \, d\tau \\
+ \frac{p(p-1)}{2} \int_t^s E\left[ (f_2 + f_4) \left\| \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(\tau) \right\|_{p-2}^{p-2} \right] \, d\tau,
$$

where

$$
f_1 := \left\| \sum_{k=1}^{n} \frac{\partial v_k}{\partial x_k}(\tau, X^t_{x,\tau}(\tau)) \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(\tau) \right\|,
$$

$$
f_2 := \left\{ \sum_{m=1}^{n'} \left( \sum_{k=1}^{n} \frac{\partial \nu_m}{\partial x_k}(\tau, X^t_{x,\tau}(\tau)) \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(\tau) \right) \right\}^2,
$$

$$
f_3 := \left\{ \sum_{k=1}^{n} \sum_{k'=1}^{n} \frac{\partial^2 v_k}{\partial x_k \partial x_{k'}}(\tau, X^t_{x,\tau}(\tau)) \frac{\partial X^t_{i,x}}{\partial x_j}(\tau) \frac{\partial X^t_{i,x}}{\partial x_{j'}}(\tau) \right\},
$$

and

$$
f_4 := \left\{ \sum_{m=1}^{n'} \left( \sum_{k=1}^{n} \sum_{k'=1}^{n} \frac{\partial^2 \nu_m}{\partial x_k \partial x_{k'}}(\tau, X^t_{x,\tau}(\tau)) \frac{\partial X^t_{i,x}}{\partial x_j}(\tau) \frac{\partial X^t_{i,x}}{\partial x_{j'}}(\tau) \right) \right\}^2.
$$

As before, the first step is to apply Young’s inequality to each of the previous integrands. For integers $q_u \in \{1, 2\}$ and $u \in \{1, 2, 3, 4\}$, we have

$$
E\left[ f_u \left\| \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(\tau) \right\|^{p-q_u} \right] \leq \frac{q_u}{p} E\left[ f_u^{p/q_u} \right] + \frac{p - q_u}{p} E\left[ \left\| \frac{\partial^2 X^t_{i,x}}{\partial x_j \partial x_{j'}}(\tau) \right\|^{p} \right].
$$

We now turn our attention to bounding $E[f_2^{p/2}]$ by primarily using Jensen’s inequality and (26). We present the proof bounding $E[f_2^{p/2}]$ and $E[f_4^{p/2}]$. Bounding $E[f_3]$ and $E[f_4]$ is analogous,

$$
E[f_2^{p/2}] \leq \left( \sum_{k=1}^{n} \left\| \frac{\partial \nu_m}{\partial x_k} \right\|_{p} \left( \sum_{k=1}^{n} \left\| \frac{\partial^2 \nu_m}{\partial x_k \partial x_{j'}} \right\|_{p} \right) \right)^{\frac{p}{2}} \leq 3^{p-1} \sum_{k \in \{i, j'\}} \left( \sum_{k=1}^{n} \left( \left\| \frac{\partial \nu_m}{\partial x_k} \right\|_{p} \right) \right)^{\frac{p}{2}} E\left[ \left\| \frac{\partial^2 \nu_m}{\partial x_k \partial x_{j'}} \right\|_{p} \right] + 3^{p-1} E\left[ \left( \sum_{k=1}^{n} \left( \left\| \frac{\partial \nu_m}{\partial x_k} \right\|_{p} \right) \right)^{2} \right].
$$

12
\[
\leq 3^{p-1} \sum_{k \in \{j,j'\}} \left( \sum_{m=1}^{n'} \left\| \frac{\partial^2 \sim_m}{\partial x_k} \right\|_\infty \right)^{p/2} \left( \mathbb{E} \left[ \left| \frac{\partial^2 X_{t,x}^i}{\partial x_j \partial x_{j'}^i}(\tau) \right|^p \right] \right)
\]
\[
+ 3^{p-1} C_p \max_{k \in \{1,2,\ldots,n\} \setminus \{j,j'\}} \mathbb{E} \left[ \left| \frac{\partial^2 X_{t,x}^i}{\partial x_j \partial x_{j'}^i}(\tau) \right|^p \right],
\]
where we used (26) in the last step. On the other hand,
\[
\mathbb{E} \left[ f_4^{p/2} \right] \leq \mathbb{E} \left[ \left( \sum_{m=1}^{n'} \left( \sum_{k=1}^n \sum_{k'=1, k' \neq j, k' \neq j'} \left| \frac{\partial^2 \sim_m}{\partial x_k} \right|_\infty \frac{\partial X_{t,x}^i}{\partial x_j} \frac{\partial X_{t,x}^i}{\partial x_{j'}^i}(\tau) \right) \right]^{p/2} \]
\[
\leq 4^{p-1} \left( \sum_{m=1}^{n'} \left\| \frac{\partial^2 \sim_m}{\partial x_j \partial x_{j'}^i} \right\|_\infty^2 \right)^{p/2} \mathbb{E} \left[ \left| \frac{\partial X_{t,x}^i}{\partial x_j}(\tau) \frac{\partial X_{t,x}^i}{\partial x_{j'}^i}(\tau) \right|^p \right]
\]
\[
+ 4^{p-1} \mathbb{E} \left( \left( \sum_{m=1}^{n'} \left( \sum_{k=1, k' \neq j, k' \neq j'} \left| \frac{\partial^2 \sim_m}{\partial x_j \partial x_{k}^i} \right|_\infty \frac{\partial X_{t,x}^i}{\partial x_k}(\tau) \right) \right)^{p/2} \left| \frac{\partial X_{t,x}^i}{\partial x_j}(\tau) \right|^p \right]
\]
\[
+ 4^{p-1} \mathbb{E} \left( \left( \sum_{m=1}^{n'} \left( \sum_{k=1, k \neq j, k' \neq j} \left| \frac{\partial^2 \sim_m}{\partial x_k \partial x_{k'}^i} \right|_\infty \frac{\partial X_{t,x}^i}{\partial x_k}(\tau) \right) \right)^{p/2} \left| \frac{\partial X_{t,x}^i}{\partial x_{j'}^i}(\tau) \right|^p \right]
\]
\[
+ 4^{p-1} \mathbb{E} \left( \left( \sum_{m=1}^{n'} \left( \sum_{k=1, k \neq j} \sum_{k'=1, k' \neq j'} \left| \frac{\partial^2 \sim_m}{\partial x_k \partial x_{k'}^i} \right|_\infty \frac{\partial X_{t,x}^i}{\partial x_k}(\tau) \right) \right)^{p/2} \left| \frac{\partial X_{t,x}^i}{\partial x_{j'}^i}(\tau) \right|^p \right].
\]
Looking at each term separately and using the bound on the first variation (36),
\[
\mathbb{E} \left[ \left| \frac{\partial X_{t,x}^i}{\partial x_j}(\tau) \right|^{2p} \right] \leq \left( \mathbb{E} \left[ \left| \frac{\partial X_{t,x}^i}{\partial x_j}(\tau) \right|^{2p} \right] \right)^{1/2} \left( \mathbb{E} \left[ \left| \frac{\partial X_{t,x}^i}{\partial x_{j'}^i}(\tau) \right|^{2p} \right] \right)^{1/2}
\]
\[
\leq K_{1,2p}^{2p}.
\]
Moreover, using Hölder’s inequality, (26) and the bound on the first variation (36),
\[
\mathbb{E} \left[ \left( \sum_{m=1}^{n'} \left( \sum_{k=1, k' \neq j} \left| \frac{\partial^2 \sim_m}{\partial x_j \partial x_{k'}^i} \right|_\infty \frac{\partial X_{t,x}^i}{\partial x_{j'}^i}(\tau) \right) \right)^{p/2} \left| \frac{\partial X_{t,x}^i}{\partial x_j}(\tau) \right|^p \right]
\]
\[
\leq \left( \sum_{k=1}^{n'} \left( \sum_{m=1}^{n'} \left| \frac{\partial^2 \sim_m}{\partial x_j \partial x_{k'}^i} \right|_\infty^{2p} \right)^{1/2} \right)^p \left( \max_{k' \neq j'} \mathbb{E} \left[ \left| \frac{\partial X_{t,x}^i}{\partial x_{j'}^i}(\tau) \right|^{2p} \right] \right)^{1/2} \mathbb{E} \left[ \left| \frac{\partial X_{t,x}^i}{\partial x_{j'}^i}(\tau) \right|^{2p} \right]^{1/2}
\]
\[
\leq K_{1,2p}^p \left( \sum_{k=1}^{n'} \left( \sum_{m=1}^{n'} \left| \frac{\partial^2 \sim_m}{\partial x_j \partial x_{k'}^i} \right|_\infty^{2p} \right)^{1/2} \right)^{1/2} \left( \max_{k' \neq j'} \mathbb{E} \left[ \left| \frac{\partial X_{t,x}^i}{\partial x_{j'}^i}(\tau) \right|^{2p} \right] \right)^{1/2}.
\]
and
\[
E \left[ \left( \sum_{m=1}^{n'} \left( \sum_{k=1, k \neq j}^{n} \sum_{k' = 1, k' \neq j}^{n} \left\| \frac{\partial^2 \xi_{\text{sim}}}{\partial x_k \partial x_{k'}} \right\|_{\infty} \right)^2 \right)^{p/2} \right] \leq \left( \sum_{k=1, k \neq j}^{n} \sum_{k' = 1, k' \neq j}^{n} \left( \sum_{m=1}^{n'} \left\| \frac{\partial^2 \xi_{\text{sim}}}{\partial x_k \partial x_{k'}} \right\|_{\infty} \right)^2 \right)^{1/2} \left( \max_{k', \ k' \neq j} \mathbb{E} \left[ \left\| \frac{\partial X_{k'}^{t,x}}{\partial x_{j}}(\tau) \right\|_{\infty} \right] \right)^{2p}.
\]

Hence
\[
E \left[ J_4^{p/2} \right] \leq 4^{p-1} \left( K_{1,2p}^{2p} F_{c_1,i,j,j'} + K_{2,2p}^{2p} F_{c_2,i,j,j'} + K_{1,2p}^{2p} F_{c_2,i,j,j'} + C_2^p F_{3,j,j'} \right),
\]
where
\[
F_{c_1,i,j,j'} := \left( \sum_{m=1}^{n'} \left\| \frac{\partial^2 \xi_{\text{sim}}}{\partial x_k \partial x_{j'}} \right\|_{\infty} \right)^{p/2},
\]
\[
F_{c_2,i,j,j'} := \left( \sum_{k=1}^{n} \left( \sum_{m=1}^{n'} \left\| \frac{\partial^2 \xi_{\text{sim}}}{\partial x_j \partial x_{k'}} \right\|_{\infty} \right)^{2p} \right)^{1/2} \left( \max_{k', \ k' \neq j} \mathbb{E} \left[ \left\| \frac{\partial X_{k'}^{t,x}}{\partial x_{j}}(\tau) \right\|_{\infty} \right] \right)^{2p} \left( \max_{k', \ k' \neq j} \mathbb{E} \left[ \left\| \frac{\partial X_{k'}^{t,x}}{\partial x_{j'}}(\tau) \right\|_{\infty} \right] \right)^{1/2},
\]
and
\[
F_{3,j,j'} := \left( \max_{k \neq j} \mathbb{E} \left[ \left\| \frac{\partial X_{k}^{t,x}}{\partial x_{j}}(\tau) \right\|_{\infty} \right] \right)^{1/2} \left( \max_{k', \ k' \neq j} \mathbb{E} \left[ \left\| \frac{\partial X_{k'}^{t,x}}{\partial x_{j'}}(\tau) \right\|_{\infty} \right] \right)^{1/2}.
\]

Note that, using Assumption 3.1,
\[
\sum_{j=1}^{n} \max_{i \neq j} F_{c_1,i,j,j'} + \sum_{j=1}^{n} \max_{i \neq j, j' \neq j} F_{c_2,i,j,j'} \leq \sum_{j=1}^{n} \max_{i \neq j} \left( \sum_{m=1}^{n'} \left\| \frac{\partial^2 \xi_{\text{sim}}}{\partial x_i \partial x_{j'}} \right\|_{\infty} \right)^{p/2} \sum_{j=1}^{n} \sum_{j'=1}^{n} \max_{i \neq j', j' \neq j} \left( \sum_{m=1}^{n'} \left\| \frac{\partial^2 \xi_{\text{sim}}}{\partial x_j \partial x_{k'}} \right\|_{\infty} \right)^{2p} \leq C_2^p.
\]

Similarly, using Assumption 3.1, (27) and (36),
\[
\sum_{j=1}^{n} \max_{i \neq j} F_{c_2,i,j,j} + \sum_{j=1}^{n} \max_{i \neq j, j' \neq j} F_{c_2,i,j,j'} = \left( \sum_{j=1}^{n} \max_{i \neq j} \left( \sum_{m=1}^{n'} \left\| \frac{\partial^2 \xi_{\text{sim}}}{\partial x_i \partial x_{j'}} \right\|_{\infty} \right)^{1/2} \right)^{p} \times \left( \max_{i,k'} \mathbb{E} \left[ \left\| \frac{\partial X_{k'}^{t,x}}{\partial x_i}(\tau) \right\|_{\infty} \right] \right)^{1/2} \times \left( \max_{j=1}^{n} \mathbb{E} \left[ \left\| \frac{\partial X_{j}^{t,x}}{\partial x_{j}}(\tau) \right\|_{\infty} \right] \right)^{1/2} \leq 2C_2^p K_{1,2p}^{2p}.
\]
Additionally, using (36),

\[
\sum_{j=1}^{n} \sum_{j'=1}^{n} F_{3,j,j'} \leq \left( \sum_{j=1}^{n} \max_{k \neq j} \left( \mathbb{E} \left[ \left| \frac{\partial X_{k}^{t,x}(\tau)}{\partial x_{j}} \right|^{2p} \right] \right)^{1/2} \right)^{p} \left( \sum_{j'=1}^{n} \max_{k' \neq j'} \left( \mathbb{E} \left[ \left| \frac{\partial X_{k'}^{t,x}(\tau)}{\partial x_{j'}} \right|^{2p} \right] \right)^{1/2} \right)^{p} \\
\leq K_{1,2p}^{p} 
\]

Similarly, we bound \( \mathbb{E} [f_{3}^{p}] \) to arrive at

\[
\mathbb{E} [f_{3}^{p}] \leq D_{2}(F_{v,1,i,j,j'} + F_{v,2,i,j,j'} + F_{v,3,i,j,j'} + F_{3,j,j'}) ,
\]

where

\[
F_{v,1,i,j,j'} := \left\| \frac{\partial^{2} \nu_{i}}{\partial x_{j} \partial x_{j'}} \right\|_{\infty}^{p}
\]

and

\[
F_{v,2,i,j,j'} := \left( \sum_{k=1}^{n} \left\| \frac{\partial^{2} \nu_{i}}{\partial x_{j} \partial x_{k'}} \right\|_{\infty} \right) \left( \max_{k' \neq j'} \mathbb{E} \left[ \left| \frac{\partial X_{k'}^{t,x}(\tau)}{\partial x_{j'}} \right|^{2p} \right] \right)^{1/2}.
\]

Therefore we can find \( b_{i,j,j'} \) such that

\[
\int_{t}^{s} \mathbb{E} [f_{3}^{p}] + (p-1)\mathbb{E} [f_{4}^{p/2}] \, d\tau \leq \mathbb{E} [f_{3}^{p}] ,
\]

which satisfy

\[
\sum_{j=1}^{n} \left( \max_{i \neq j} b_{i,j,i} \right) + \sum_{j=1}^{n} \sum_{j'=1}^{n} \left( \max_{i \neq (j,j')} b_{i,j,j'} \right) \leq D_{2} .
\]

We use (39) and (41) in (38) and the result in (37) and simplify to arrive at

\[
\mathbb{E} \left[ \left| \frac{\partial^{2} X_{i}^{t,x}(s)}{\partial x_{j} \partial x_{j'}} \right| \right]^{p} \leq D_{2} \int_{t}^{s} \mathbb{E} \left[ \left| \frac{\partial^{2} X_{i}^{t,x}(\tau)}{\partial x_{j} \partial x_{j'}} \right| \right]^{p} \, d\tau \\
+ D_{2} \int_{t}^{s} \max_{k \in \{1, 2, ..., n\} \setminus \{j, j'\}} \mathbb{E} \left[ \left| \frac{\partial^{2} X_{k}^{t,x}(\tau)}{\partial x_{j} \partial x_{j'}} \right| \right]^{p} \, d\tau \\
+ D_{2} a_{i,j}^{p} \int_{t}^{s} \mathbb{E} \left[ \left| \frac{\partial^{2} X_{i}^{t,x}(\tau)}{\partial x_{j} \partial x_{j}} \right| \right]^{p} \, d\tau \\
+ D_{2} a_{i,j'}^{p} \int_{t}^{s} \mathbb{E} \left[ \left| \frac{\partial^{2} X_{i}^{t,x}(\tau)}{\partial x_{j} \partial x_{j'}} \right| \right]^{p} \, d\tau \\
+ b_{i,j,j'} 
\]

Recall that

\[
a_{i,j}^{p} := \left\| \frac{\partial \nu_{i}}{\partial x_{j}} \right\|_{\infty}^{p} + \left( \sum_{n=1}^{n'} \left( \frac{\partial \kappa_{im}}{\partial x_{j}} \right) \right)^{p/2} ,
\]

and \( \max_{i} a_{i,i} + \sum_{j=1}^{n} \max_{i \neq j} a_{i,j} \leq 2C_{1} \). Then, taking the maximum over all \( i \) in (43) and using Grönwall’s inequality yield

\[
\max_{i} \mathbb{E} \left[ \left( \frac{\partial^{2} X_{i}^{t,x}(s)}{\partial x_{j} \partial x_{j'}} \right) \right]^{p} \leq D_{2} ,
\]

15
for all \( j, j' \in \{1, \ldots, n\} \). Setting \( j' = i \) and taking the maximum over \( i \neq j \) in (43) yield

\[
\max_{i \neq j} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_i} (s) \right]^p \leq D_2 \int_t^s \max_{i \neq j} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_i} (\tau) \right]^p \, d\tau \\
+ D_2 \left( \max_{i \neq j} a_{i,j}^p \right) \int_t^s \max_{i \neq j} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_i} (\tau) \right]^p \, d\tau \\
+ D_2 \sum_{i \neq j} \max_{i \neq j} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_i} (\tau) \right]^p \\
+ \frac{b_{i,j,i}^p}{\max_{i \neq j} a_{i,j}}. 
\]

Using (44) to bound the second term, then Grönwall’s inequality and then taking the \( p \)’th root and summing over \( j \) yields

\[
\sum_{j=1}^n \max_{i \neq j} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_j} (s) \right]^p \leq D_2 \left( \sum_{j=1}^n \max_{i \neq j} a_{i,j} + \sum_{j=1}^n \max_{i \neq j} b_{i,j,i} \right) \leq D_2, \tag{45}
\]

where we used (42). Finally, taking the maximum over \( i \notin \{j, j'\} \) in (43),

\[
\max_{i \notin \{j, j'\}} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}} (s) \right]^p \leq D_2 \int_t^s \max_{i \notin \{j, j'\}} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}} (\tau) \right]^p \, d\tau \\
+ D_2 \left( \max_{i \notin \{j, j'\}} a_{i,j'}^p \right) \int_t^s \max_{i \notin \{j, j'\}} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}} (\tau) \right]^p \, d\tau \\
+ D_2 \left( \max_{i \notin \{j, j'\}} a_{i,j'}^p \right) \int_t^s \max_{i \notin \{j, j'\}} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}} (\tau) \right]^p \, d\tau \\
+ \frac{b_{i,j,i'}^p}{\max_{i \notin \{j, j'\}} a_{i,j'}}. 
\]

Then, using Grönwall’s inequality, taking the \( p \)’th root and summing over \( j \) and \( j' \) yield

\[
\sum_{j=1}^n \sum_{j'=1}^n \max_{i \notin \{j, j'\}} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}} (s) \right]^p \leq D_2 \left( \sum_{j=1}^n \max_{i \notin \{j, j'\}} a_{i,j} \right) \left( \max_{j'=1} \sum_{j' \neq j} \max_{i \notin \{j, j'\}} \sup_{s \leq \tau \leq s} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}} (\tau) \right]^p \right) \]

\[
+ D_2 \left( \sum_{j'=1}^n \max_{i \notin \{j, j'\}} a_{i,j'} \right) \left( \max_{j \neq j'} \sum_{j=1}^n \sup_{s \leq \tau \leq s} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}} (\tau) \right]^p \right) \\
+ D_2 \sum_{j=1}^n \sum_{j'=1}^n \left( \max_{i \notin \{j, j'\}} b_{i,j,i'} \right). 
\]

The result follows by (42), and since \( \sum_{j=1}^n \max_{i \notin \{j, j'\}} a_{i,j} \leq C_1 \) and, by (27), (44) and (45),

\[
\max_{j} \sup_{\tau \leq \tau \leq s} E \left[ \frac{\partial^2 X_{j,x}^{t,x}}{\partial x_j \partial x_{j}} (\tau) \right]^p \\
\leq \max_{j} \sup_{\tau \leq \tau \leq s} E \left[ \frac{\partial^2 X_{j,x}^{t,x}}{\partial x_j \partial x_{j}} (\tau) \right]^p + \sum_{j'=1}^n \max_{i \notin \{j, j'\}} \sup_{\tau \leq \tau \leq s} E \left[ \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}} (\tau) \right]^p \\
\leq D_2. 
\]
Third variation The result for the third variations can be proven similarly and is omitted here. 

Proposition 3.3 (Bounds on derivatives of the value function). Let \( u : [0, T] \times \mathbb{R}^n \rightarrow \mathbb{R} \) satisfy the Kolmogorov backward equation on \((t, x) \in [0, T) \times \mathbb{R}^n \),

\[
\frac{\partial u}{\partial t}(t, x) + \sum_{i=1}^{n} \nu_i(t, x) \frac{\partial u}{\partial x_i}(t, x) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left( \sum_{m=1}^{n} \sigma_{im}(t, x) \sigma_{jm}(t, x) \right) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) = 0
\]  

(46)

and \( u(T, x) = g(x) \).

Assume that the coefficients, \( \{\nu_i\}_{i=1}^{n} \) and \( \{\sigma_{i,m}\}_{i=1,...,n,m=1,...,n} \), satisfy Assumption 3.1 for \( q \in \{2, 3\} \) and that \( g \) has continuous bounded derivatives up to order \( q \). Then, for some constant \( D_q \), independent of \( n \), there holds

\[
u(t, \cdot) \|_{C^{q}(\mathbb{R}^n)} \leq D_q \|g\|_{C^{q}(\mathbb{R}^n)}.
\]

Proof. First note that under Assumption 3.1 for \( q = 2 \), \( u \) satisfies for all \( t \in [0, T] \) and \( x \in \mathbb{R}^n \) [11, Theorem 6.1 in Chapter 5]

\[
u(t, x) = E \left[ g(X^{t,x}(T)) \right].
\]  

(47)

Next, we differentiate \( u \) with respect to the initial conditions and exchange the differentiation with the expectation in (47) [11, Theorem 5.5 in Chapter 5]. Then, we bound

\[
\sum_{j=1}^{n} \left| \frac{\partial u}{\partial x_j}(t, x) \right| = \sum_{j=1}^{n} E \left[ \sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(X^{t,x}(T)) \frac{\partial X_{i,x}^{t,x}}{\partial x_j}(T) \right]
\]

\[
\leq \sum_{i=1}^{n} \left( \max_{j=1}^{n} E \left[ \left| \frac{\partial X_{i,x}^{t,x}}{\partial x_j}(T) \right| \right] \right)
\]

\[
\leq K_{1,2} \left( \max_{j=1}^{n} E \left[ \left| \frac{\partial X_{i,x}^{t,x}}{\partial x_j}(T) \right| \right] \right)
\]

by Lemma 3.2. Similarly,

\[
\sum_{j=1}^{n} \sum_{j' = 1}^{n} \left| \frac{\partial^2 u}{\partial x_j \partial x_{j'}}(t, x) \right| \leq \sum_{i=1}^{n} \sum_{i' = 1}^{n} E \left[ \sum_{j=1}^{n} \sum_{j' = 1}^{n} \frac{\partial^2 g}{\partial x_i \partial x_{i'}}(X_{i,x}^{t,x}(T)) \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}}(T) \right]
\]

\[
\leq \left( \max_{i=1}^{n} \sum_{j=1}^{n} \sum_{j' = 1}^{n} E \left[ \left| \frac{\partial^2 X_{i,x}^{t,x}}{\partial x_j \partial x_{j'}}(T) \right| \right] \right)^{1/2}
\]

\[
\leq (K_{1,2}^2 + K_{2,2}) \left( \max_{i=1}^{n} \left( \sum_{j=1}^{n} \sum_{j' = 1}^{n} \left| \frac{\partial^2 g}{\partial x_i \partial x_{i'}}(X_{i,x}^{t,x}(T)) \right| \right) \right).
\]

It is easy to see that the previous proof extends to \( q = 3 \) as well.
Figure 1: A histogram of the values of \( \{X^n_i(1)\}_{i=1}^n \) which follows the system of SDEs in (2) with (48). Here, the values were approximated using the Euler–Maruyama time-stepping scheme with 64 uniform time-steps and \( n = 2,048 \).

4 Numerical Verification

In this section, we present a numerical study of the weak error of a particle approximation of the solution of a simple McKean-Vlasov equation. For \( r \in \mathbb{N} \) and \( x \in \mathbb{R} \), consider the function

\[
\psi_r(x) := \begin{cases} 
(1 - x^2)^r & |x| \leq 1, \\
0 & |x| > 1.
\end{cases}
\]

and note that \( \psi_0 \) is discontinuous while for \( r > 0 \) and all \( k < r \) the \( k \)’th derivative of \( \psi_r \) exists is uniformly bounded, and the \( (r-1) \)'th derivative is Lipschitz continuous. Subsequently, consider the particle system (2) with \( d = 1 \) and \( X_i(0) \) being uniformly distributed in \([-1, 1]\), and for \( x, y \in \mathbb{R} \) set

\[
\begin{aligned}
a(x, y) &:= 2(x - 0.2) + y, \\
\sigma(x, y) &:= 0.2(1 + y), \\
\kappa_1(x, y) &:= \psi_1(10|x - y|), \\
\kappa_2(x, y) &:= \psi_1(5|x - y|).
\end{aligned}
\]

(48)

Note that the previous \( a, \sigma, \kappa_1, \) and \( \kappa_2 \), the latter two being only Lipschitz continuous, do not satisfy the conditions the conditions of Theorem 2.3. To approximate solutions to (2), we use an Euler–Maruyama time-stepping scheme with a fixed number of time-steps.

To illustrate the convergence of \( \mathbf{X}^n \) to \( Z \), the corresponding solution of the McKean–Vlasov equation (1), we consider the sequence of systems, denoted by \( \mathbf{X}^n \), satisfying (2) with an increasing number of particles, \( n \). See Fig. 1 for a histogram of the values of \( X^n_i(1) \) for \( n = 2,048 \) and using 64 uniform time-steps in an Euler-Maruyama scheme. We also consider the discontinuous function \( g(x) = \psi_0(10|x - 0.2|) \) and plot in Fig. 2 the quantities \( \mathbb{E}[(X_i^{2n} - X_i^n)^2] \) and \( \mathbb{E}|g(X_i^{2n}) - g(X_i^n)| \). The same convergence behaviour of these quantities was obtained with different numbers of uniform time-steps. Even though \( \kappa_1 \) and \( \kappa_2 \) are only Lipschitz continuous and \( g \) is discontinuous, the observed weak convergence rate is still \( \mathcal{O}(n^{-1}) \), as predicted by Theorem 2.3 when \( \kappa_1, \kappa_2, \) and \( g \) were assumed to be three-times differentiable. Hence, it may be that the assumptions required by Theorem 2.3 and similar proofs in the literature can be relaxed by exploiting, for example, the smoothness of the probability measure.
Figure 2: Convergence rates $X^n_i(1)$ which follows the system of SDEs in (2) with (48) and $g(x) = \psi_0(10|x - 0.2|)$. Here, the values were approximated using the Euler–Maruyama time-stepping scheme with $N = 64$ time-steps. Note that the rates are consistent with the predicted rates in Theorem 2.3 even though the coefficients of the SDE do not have sufficient smoothness as required by the theorem.

References

[1] Vlad Bally and Denis Talay. The Euler scheme for stochastic differential equations: error analysis with Malliavin calculus. Mathematics and Computers in Simulation, 38(1):35 – 41, 1995.

[2] Vlad Bally and Denis Talay. The law of the Euler scheme for Stochastic Differential Equations: II. Convergence rate of the density. Monte Carlo Methods and Applications, 2(2):93 – 128, 1996.

[3] Christian Bayer, Eric Joseph Hall, and Raúl Tempone. Weak error rates for option pricing under linear rough volatility, 2021.

[4] Christian Bayer, Anders Szepessy, and Raúl Tempone. Adaptive weak approximation of reflected and stopped diffusions. Monte Carlo Methods and Applications, 16(1):1–67, January 2010.

[5] Oumaima Bencheikh and Benjamin Jourdain. Bias behaviour and antithetic sampling in mean-field particle approximations of SDEs nonlinear in the sense of McKean. ESAIM: Proceedings and Surveys, 65:219–235, 2019.

[6] Pierre Cardaliaguet. Notes on mean field games. Technical report, Technical report, 2010.

[7] Jean-François Chassagneux, Lukasz Szpruch, and Alvin Tse. Weak quantitative propagation of chaos via differential calculus on the space of measures. The Annals of Applied Probability, 32(3), June 2022.

[8] Jean-François Chassagneux, Dan Crisan, and François Delarue. A probabilistic approach to classical solutions of the master equation for large population equilibria, volume 280. American Mathematical Society, 2022.

[9] Dan Crisan and Jie Xiong. Approximate Mckean–Vlasov representations for a class of SPDEs. Stochastics An International Journal of Probability and Stochastics Processes, 82(1):53–68, 2010.

[10] Paul-Eric Chaudru de Raynal and Noufel Frikha. From the backward Kolmogorov PDE on the Wasserstein space to propagation of chaos for McKean-Vlasov SDEs. Journal de Mathématiques Pures et Appliquées, 156:1–124, 2021.
[11] Avner Friedman. *Stochastic Differential Equations and Applications*. Elsevier, 1975.

[12] William R. P. Hammersley, David Šiška, and Łukasz Szpruch. McKean–Vlasov SDEs under measure dependent Lyapunov conditions. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 57(2):1032 – 1057, 2021.

[13] Håkon Hoel, Juho Häppölä, and Raúl Tempone. Construction of a mean square error adaptive Euler–Maruyama method with applications in multilevel Monte Carlo. In Ronald Cools and Dirk Nuyens, editors, *Springer Proceedings in Mathematics & Statistics*, pages 29–86. Springer International Publishing, 2016.

[14] Håkon Hoel, Erik von Schwerin, Anders Szepessy, and Raúl Tempone. Implementation and analysis of an adaptive multilevel Monte Carlo algorithm. *Monte Carlo Methods and Applications*, 20(1):1–41, January 2014.

[15] Markos A. Katsoulakis and Anders Szepessy. Stochastic hydrodynamical limits of particle systems. *Communications in Mathematical Sciences*, 4(3):513–549, 2006.

[16] Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23 of *Applications of Mathematics (New York)*. Springer Berlin Heidelberg, 1992.

[17] Vassili N. Kolokoltsov. *Nonlinear Markov Processes and Kinetic Equations*, volume 182. Cambridge University Press, July 2010.

[18] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Japanese Journal of Mathematics*, 2(1):229–260, March 2007.

[19] Stéphane Mischler, Clément Mouhot, and Bernt Wennberg. A new approach to quantitative propagation of chaos for drift, diffusion and jump processes. *Probability Theory and Related Fields*, 161(1-2):1–59, December 2013.

[20] Yuliya Mishura and Alexander Veretennikov. Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Theory of Probability and Mathematical Statistics*, 103:59–101, June 2021.

[21] Kyoung-Sook Moon, Anders Szepessy, Raúl Tempone, and Georgios E. Zouraris. Convergence rates for adaptive weak approximation of stochastic differential equations. *Stochastic Analysis and Applications*, 23(3):511–558, May 2005.

[22] E. Mordecki, A. Szepessy, R. Tempone, and G. E. Zouraris. Adaptive weak approximation of diffusions with jumps. *SIAM Journal on Numerical Analysis*, 46(4):1732–1768, January 2008.

[23] Anders Szepessy, Raúl Tempone, and Georgios E. Zouraris. Adaptive weak approximation of stochastic differential equations. *Communications on Pure and Applied Mathematics*, 54(10):1169–1214, 2001.

[24] Alain-Sol Sznitman. Topics in propagation of chaos. In *Lecture Notes in Mathematics*, pages 165–251. Springer Berlin Heidelberg, 1991.

[25] Łukasz Szpruch and Alvin Tse. Antithetic multilevel particle system sampling method for McKean-Vlasov SDEs. 2019.

[26] Denis Talay and Luciano Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications*, 8(4):483–509, 1990.

[27] Erik von Schwerin and Anders Szepessy. A stochastic phase-field model determined from molecular dynamics. *M2AN Math. Model. Numer. Anal.*, 44(4):627–646, 2010.

**Acknowledgements** R. Tempone was partially supported by the KAUST Office of Sponsored Research (OSR), under Award numbers URF/1/2281-01-01, URF/1/2584-01-01 in the KAUST Competitive Research Grants Program Round 8, and the Alexander von Humboldt Foundation, through the Alexander von Humboldt Professorship award.