IMA Journal of Applied Mathematics (2025) **00**, 1–28
https://doi.org/10.1093/imamat/hxaf015

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

ABDUL-LATEEF HAJI-ALI*
Maxwell Institute and Heriot-Watt University, Edinburgh, UK
*Corresponding author: a.hajiali@hw.ac.uk

HÅKON HOEL
University of Oslo, Oslo, Norway

AND

RAÚL TEMPONE
RWTH Aachen University, Germany and King Abdullah University of Science & Technology (KAUST), Thuwal, Saudi Arabia

[Received on 13 October 2023; revised on 16 April 2025; accepted on 3 July 2025]

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 $\mathcal{O}(n^{-1})$. 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 $\sigma : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^{d\times d'}$, $t \geq 0$ 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:

$$\begin{aligned} Z(t) = \xi &+ \int_0^t a\left(Z(s), \int_{\mathbb{R}^d} \kappa_1(Z(s), z)\mu_s(\mathrm{d}z)\right)\mathrm{d}s \\ &+ \int_0^t \sigma\left(Z(s), \int_{\mathbb{R}^d} \kappa_2(Z(s), z)\mu_s(\mathrm{d}z)\right)\mathrm{d}W(s), \end{aligned} \tag{1}$$

where μ_s denotes the law of $Z(s)$ for all $s \geq 0$, and ξ denotes a random initial state whose law is μ_0 and is assumed to be independent of the Wiener process, W. We focus on one-dimensional interaction kernels $\kappa_1, \kappa_2 : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ 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 σ being constant and κ_1 being bounded and Lipschitz, (Sznitman, 1991, Theorem I.1.1) shows the existence and uniqueness of a strong solution to Equation (1). A recent analysis yielded the same result in (Mishura & Veretennikov, 2021, Proposition 2) when the initial condition ξ has a finite fourth moment, $a(x, y) = \sigma(x, y) = y$, under uniform non-degeneracy conditions on κ_2, and when for all $\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{y} \in \mathbb{R}^d$, there exists a constant $C > 0$ such that

$$|\kappa_1(\boldsymbol{x},\boldsymbol{y})| + |\kappa_2(\boldsymbol{x},\boldsymbol{y})| \le C(1 + \|\boldsymbol{x}\|),$$
$$|\kappa_2(\boldsymbol{x},\boldsymbol{y}) - \kappa_2\big(\boldsymbol{x}',\boldsymbol{y}\big)| \le C\,(1 + \|\boldsymbol{y}\|^2)\,\|\boldsymbol{x} - \boldsymbol{x}'\|.$$

Crisan & Xiong (2010) also shows the existence and uniqueness result for $\kappa_2 = 0$ and a particular form of a. Existence of weak solutions was also shown by Hammersley *et al.* (2021) under certain measure-dependent Lyapunov conditions. De Raynal (2020); Raynal & Frikha (2021b) further showed strong and weak existence for less smooth, but bounded, drift and diffusion coefficients. In the current work, we do not focus on the existence of solutions to Equation (1) and instead assume the existence of weak solutions and consider strong 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:

$$\begin{aligned} X_i^n(t) = \xi_i \quad &+ \int_0^t a\left(X_i^n(s), \frac{1}{n}\sum_{j=1}^n \kappa_1\left(X_i^n(s), X_j^n(s)\right)\right) \mathrm{d}s \\ &+ \int_0^t \sigma\left(X_i^n(s), \frac{1}{n}\sum_{j=1}^n \kappa_2(X_i^n(s), X_j^n(s))\right) \mathrm{d}W_i(s), \end{aligned} \tag{2}$$

for $i \in \{1, 2, \ldots, n\}$, where ξ_i are i.i.d. and have the same law, μ_0, and $\{W_i(s)\}_{s\ge0}$ are independent d'-dimensional Wiener processes and independent of $\{\xi_i\}_{i=1}^n$. In other words, the law μ_t for $t \ge 0$ is approximated by an empirical measure based on the particles $\boldsymbol{X}^n(t) := \{X_i^n(t)\}_{i=1}^n$. It should be noted that these particles are identically distributed but not independent.

For $\boldsymbol{Z}^n := (Z_i)_{i=1}^n$ being n independent samples of the solution to the McKean–Vlasov equation (1) and a function $g : \mathbb{R}^{d\times n} \to \mathbb{R}$, the weak error at time t is defined as the absolute difference $|\mathsf{E}\left[g(\boldsymbol{X}^n(t))\right] - \mathsf{E}\left[g(\boldsymbol{Z}^n(t))\right]|$. The weak error was established to be $\mathcal{O}(1/n)$ in, e.g. (Kolokoltsov, 2010, Chapter 9) and (Mischler *et al.*, 2013, 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, Bencheikh & Jourdain (2019) employ a similar methodology to the current work but assumes that κ_1 and κ_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 (Szpruch & Tse, 2021, Theorem B.2), (Chassagneux *et al.*, 2022a, Theorem 2.17) and Raynal & Frikha (2021a) use Lions-derivatives as defined by Cardaliaguet (2010) to bound derivatives with respect to measures in a master equation for probability measure, see also Lasry & Lions (2007); Chassagneux *et al.* (2022b). For a more exhaustive literature review, see Chassagneux *et al.* (2022a).

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 & Tubaro (1990), who estimated the time discretization error for uniform deterministic time-steps. Bally & Talay (1995, 1996), extended the analysis to approximations with non-smooth observables and the probability density of the solution at a given time. Kloeden & Platen (1992) generalized the analysis by

Talay & Tubaro (1990) to weak approximations of a higher order. Later, in a series of works inspired by Talay & Tubaro (1990), 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 Szepessy *et al.* (2001); Moon *et al.* (2005); Mordecki *et al.* (2008); Bayer *et al.* (2010). Furthermore, the same analysis line was also used for adaptive Multilevel Monte Carlo, Hoel *et al.* (2014, 2016).

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, Katsoulakis & Szepessy (2006) employed master equations with long-range interaction potentials as a base model as the stochastic Ising model with Glauber dynamics, whereas Schwerin & Szepessy (2010) 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 behaviour. For example, Bayer *et al.* (2022) 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 Bayer *et al.* (2022) 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 Bencheikh & Jourdain (2019) for a special case of (1) in which κ_1 and κ_2 do not depend on the state Z, i.e. $\kappa_\cdot(x, y) = \kappa_\cdot(y)$. A key difference to our methodology is that in Bencheikh & Jourdain (2019), the Kolmogorov backward equation involves μ_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 existing results.

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 Equation (2), such that $A \le cB$. For a multi-index $\boldsymbol{\ell} \in \mathbb{N}^n$, $n \in \mathbb{N}$, define the derivative

$$\frac{\partial^{|\boldsymbol{\ell}|}}{\partial x^{\boldsymbol{\ell}}} := \frac{\partial^{|\boldsymbol{\ell}|}}{\prod_{j=1}^n \partial x_j^{\ell_j}},$$

where $|\boldsymbol{\ell}| = \sum_{i=1}^n \ell_i$. Let $\|\cdot\|$ denote the Euclidean norm and let $C_{\mathrm{b}}(\mathbb{R}^n; \mathbb{R}^m)$ denote the space of bounded continuous functions $u \equiv (u_i)_{i=1}^m : \mathbb{R}^n \to \mathbb{R}^m$ with the usual norm

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

Let also $C_{\mathrm{b}}(\mathbb{R}^n;\mathbb{R}) = C_{\mathrm{b}}(\mathbb{R}^n)$. When u has continuous, bounded derivatives up to order k, define the extended semi-norm

$$|u|_{C^k_{\mathrm{bd}}(\mathbb{R}^n;\mathbb{R}^m)} = \sum_{j=1}^{m} \sum_{\boldsymbol{\ell}\in\mathbb{N}^n, 1\le|\boldsymbol{\ell}|\le k} \left\| \frac{\partial^{|\boldsymbol{\ell}|} u_j}{\partial \boldsymbol{x}^{\boldsymbol{\ell}}} \right\|_{C_{\mathrm{b}}(\mathbb{R}^n\,;\,\mathbb{R}^m)}.$$

and the norm $\|u\|_{C^k_{\mathrm{b}}(\mathbb{R}^n;\mathbb{R}^m)} = \|u\|_{C_{\mathrm{b}}(\mathbb{R}^n;\mathbb{R}^m)} + |u|_{C^k_{\mathrm{bd}}(\mathbb{R}^n;\mathbb{R}^m)}$. For a matrix $\boldsymbol{x} \in \mathbb{R}^{d\times n}$, we denote its components as $\boldsymbol{x} = \left(x_{i,j}\right)_{i=\{1,\dots,n\},j=\{1,\dots,d\}} \in \mathbb{R}^{d\times n}$. Similarly for a function $u : \mathbb{R}^{d\times n} \to \mathbb{R}$, we will use the notation $\nabla u = \left(\frac{\partial u}{\partial x_{i,j}}\right)_{i=\{1,\dots,n\},j=\{1,\dots,d\}}$ 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 $\mathcal{O}(1/n)$. We start by stating boundedness and convergence results involving only samples of Z, the solution to the McKean-Vlasov equation Equation (1).

PROPOSITION 2.1. Assume that weak solutions to Equation (1) exist and let $\{Z_i\}_{i=1}^n$ be n independent processes each satisfying Equation (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

$$|\kappa(\boldsymbol{x},\boldsymbol{y}) - \kappa(\boldsymbol{x}',\boldsymbol{y}')| \le C\big(\|\boldsymbol{x}-\boldsymbol{x}'\| + \|\boldsymbol{y}-\boldsymbol{y}'\|\big) \qquad \text{for all } \boldsymbol{x},\boldsymbol{x}',\boldsymbol{y},\boldsymbol{y}' \in \mathbb{R}^d.$$

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

$$\mathsf{E}\left[\left|\frac{1}{n}\sum_{j=1}^{n}\kappa(Z_i(t),Z_j(t)) - \int_{\mathbb{R}^d}\kappa(Z_i(t),z)\mu_t(\mathrm{d}z)\right|^{2p}\right] \lesssim n^{-p}\,\mathsf{E}\left[\left\|Z(t)\right\|^{2p}\right]. \tag{3}$$

Moreover, assuming that a, σ, κ_1 and κ_2 in Equation (1) are Lipschitz continuous, we have

$$\sup_{0\le s\le t}\mathsf{E}\left[\|Z(s)\|^{2p}\right] \lesssim 1 + \mathsf{E}\left[\|\xi\|^{2p}\right]. \tag{4}$$

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

Proof. The moment boundedness result in Equation (4) with Lipschitz assumptions is classical Sznitman (1991) using Itô's formula, Young's and Grönwall's inequalities. The proof of Equation (3) for general p is fairly technical and we include it in Appendix A. We list here the proof for $p = 1$ to showcase the main ideas. First, we set, without loss of generality, $i = 1$ and let

$$\Delta_j\kappa := \int_{\mathbb{R}^d}\kappa(Z_1(t),z)\mu_t(\mathrm{d}z) - \kappa(Z_1(t),Z_j(t)). \tag{5}$$

Expanding the square

$$\mathsf{E}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\Delta_j\kappa\right)^2\right] = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{j'=1}^{n}\mathsf{E}\left[\Delta_j\kappa\,\Delta_{j'}\kappa\right].$$

When $j \neq j'$ and are both different from 1, we have

$$\mathsf{E}\left[\Delta_j\kappa\,\Delta_{j'}\kappa\right] = \mathsf{E}\left[\mathsf{E}\left[\Delta_j\kappa \mid Z_1(t)\right]\,\mathsf{E}\left[\Delta_{j'}\kappa \mid Z_1(t)\right]\right] = 0,$$

since, for a given $Z_1(t)$, $\Delta_j\kappa$ and $\Delta_{j'}\kappa$ are conditionally independent and $\mathsf{E}\left[\Delta_j\kappa \mid Z_1(t)\right] = 0$ when $j \neq 1$. When $j = j'$, we bound using the Lipschitz assumption on κ,

$$\begin{aligned}\mathsf{E}\left[(\Delta_j\kappa)^2\right] &\leq \mathsf{E}\left[\left(\int_{\mathbb{R}^d} C\left\|z - Z_j(t)\right\|\mu_t(\mathrm{d}z)\right)^2\right]\\ &\leq 4C^2\,\mathsf{E}\left[\left\|Z(t)\right\|^2\right].\end{aligned}$$

A similar bound can be obtained when $1 \in \{j, j'\}$ by using Hölder's inequality. Substituting back yields the claim. □

LEMMA 2.2. Assume that weak solutions to Equation (1) exist and let $\boldsymbol{Z}^n = \{Z_i\}_{i=1}^n$ be n independent processes each satisfying Equation (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 twice differentiable in the second argument and satisfying

$$\left|\frac{\partial f(\boldsymbol{x}, y)}{\partial y}\right| + \left|\frac{\partial^2 f(\boldsymbol{x}, y)}{\partial y^2}\right| \leq \widetilde{C}(1 + \|\boldsymbol{x}\| + |y|), \tag{6}$$

for all $\boldsymbol{x} \in \mathbb{R}^d$ and $y \in \mathbb{R}$. Then, for any $g \in C^1_{\mathrm{b}}(\mathbb{R}^{d\times n})$ and any $i \in \{1, \ldots, n\}$ we have

$$\begin{aligned}&\left|\mathsf{E}\left[\left(f\left(Z_i(t), \int_{\mathbb{R}^d}\kappa(Z_i(t), z)\mu_t(\mathrm{d}z)\right) - f\left(Z_i(t), \frac{1}{n}\sum_{j=1}^{n}\kappa(Z_i(t), Z_j(t))\right)\right)g\left(\boldsymbol{Z}^n(t)\right)\right]\right|\\ &\lesssim n^{-1}\,\|g\|_{C^1_{\mathrm{b}}(\mathbb{R}^{d\times n})}\left(1 + \mathsf{E}\left[\|\xi\|^4\right]\right).\end{aligned} \tag{7}$$

The hidden constant in Equation (7) depend only on $\widetilde{C}, d, t$ and the Lipschitz constant of κ.

Proof. Without loss of generality, we fix $i = 1$ and define

$$\Delta f := f\left(Z_1(t), \int_{\mathbb{R}^d}\kappa(Z_1(t), z)\mu_t(\mathrm{d}z)\right) - f\left(Z_1(t), \frac{1}{n}\sum_{j=1}^{n}\kappa(Z_1(t), Z_j(t))\right).$$

For $\Delta_j\kappa$ as defined in Equation (5), let

$$\mathfrak{D}_n = \frac{1}{n}\sum_{j=1}^{n} \Delta_j\kappa,$$

then by Taylor expanding, we can bound

$$\left|\mathsf{E}\left[\Delta f\, g(\boldsymbol{Z}^n(t))\right]\right| \le \left|\mathsf{E}\left[\bar{f}(Z_1(t))\,\mathfrak{D}_n\, g(\boldsymbol{Z}^n(t))\right]\right| + \|g\|_{C_b(\mathbb{R}^{d\times n})}\mathsf{E}\left[\left|\bar{\bar{f}}(Z_1(t))\right|\mathfrak{D}_n^2\right], \tag{8}$$

where, for all $\boldsymbol{x} \in \mathbb{R}^d$,

$$\bar{f}(\boldsymbol{x}) := \frac{\partial f}{\partial y}\left(\boldsymbol{x}, \int_{\mathbb{R}^d} \kappa(\boldsymbol{x}, z)\mu_t(\mathrm{d}z)\right),$$

$$\text{and} \qquad \bar{\bar{f}}(\boldsymbol{x}) := \int_0^1 (1-s)\,\frac{\partial^2 f}{\partial y^2}\left(\boldsymbol{x}, s\,\mathfrak{D}_n + \int_{\mathbb{R}^d} \kappa(\boldsymbol{x}, z)\mu_t(\mathrm{d}z)\right)\mathrm{d}s.$$

Since κ is Lipschitz continuous, implying linear growth, we have for any $\boldsymbol{x} \in \mathbb{R}^d$,

$$\begin{aligned}\left|\int_{\mathbb{R}^d} \kappa(\boldsymbol{x}, z)\mu_t(\mathrm{d}z)\right| &\lesssim \left|\int_{\mathbb{R}^d} (1 + \|\boldsymbol{x}\| + \|z\|)\mu_t(\mathrm{d}z)\right| \\ &\lesssim 1 + \|\boldsymbol{x}\| + \mathsf{E}\left[\|Z(t)\|\right].\end{aligned} \tag{9}$$

Using Equation (9), we can also conclude that, for any $j \in \{1, \ldots, n\}$,

$$\left|\Delta_j\kappa\right| \lesssim 1 + \|Z_1(t)\| + \|Z_j(t)\| + \mathsf{E}\left[\|Z(t)\|\right]. \tag{10}$$

Using Equations (6) and (9) yields the following bounds for any $\boldsymbol{x} \in \mathbb{R}^d$,

$$|\bar{f}(\boldsymbol{x})| \lesssim 1 + \|\boldsymbol{x}\| + \mathsf{E}\left[\|Z(t)\|\right], \tag{11}$$

$$|\bar{\bar{f}}(\boldsymbol{x})| \lesssim 1 + \left|\mathfrak{D}_n\right| + \|\boldsymbol{x}\| + \mathsf{E}\left[\|Z(t)\|\right]. \tag{12}$$

Using Equation (12), Hölder's inequality for probability measures, Equation (3) for $p = 2$ and the fact that $x^r \le 1 + x$ for $0 \le r \le 1$ and $x \in \mathbb{R}_+$ yields

$$\begin{aligned}\mathsf{E}\left[\left|\bar{\bar{f}}(Z_1(t))\right|\mathfrak{D}_n^2\right] &\lesssim (1 + \mathsf{E}\left[\|Z(t)\|\right])\,\mathsf{E}\left[\mathfrak{D}_n^2\right] + \mathsf{E}\left[|\mathfrak{D}_n|^3\right] + \mathsf{E}\left[\|Z_1(t)\|\,\mathfrak{D}_n^2\right] \\ &\le (1 + \mathsf{E}\left[\|Z(t)\|^4\right]^{1/4})\left(\mathsf{E}\left[\mathfrak{D}_n^4\right]\right)^{1/2} + \left(\mathsf{E}\left[\mathfrak{D}_n^4\right]\right)^{3/4} + \left(\mathsf{E}\left[\|Z(t)\|^4\right]\right)^{1/4}\left(\mathsf{E}\left[\mathfrak{D}_n^4\right]\right)^{1/2} \\ &\lesssim n^{-1}\left(1 + \left(\mathsf{E}\left[\|Z(t)\|^4\right]\right)^{1/4}\right)\mathsf{E}\left[\|Z(t)\|^4\right]^{1/2} + n^{-3/2}\,\mathsf{E}\left[\|Z(t)\|^4\right]^{3/4} + n^{-1}\,\mathsf{E}\left[\|Z(t)\|^4\right]^{3/4} \\ &\lesssim n^{-1}\left(1 + \mathsf{E}\left[\|Z(t)\|^4\right]\right).\end{aligned}$$

Furthermore, using the definition of $\mathfrak{D}_n$,

$$\left|\mathsf{E}\left[\bar{f}(Z_1(t))\,\mathfrak{D}_n\,g\big(\boldsymbol{Z}^n(t)\big)\right]\right| \le \frac{1}{n}\left|\mathsf{E}\left[\bar{f}(Z_1(t))\,\big(\Delta_1\kappa\big)\,g(\boldsymbol{Z}^n(t))\right]\right| + \frac{1}{n}\sum_{j=2}^{n}\left|\mathsf{E}\left[\bar{f}(Z_1(t))\,\Delta_j\kappa\,g(\boldsymbol{Z}^n(t))\right]\right|. \tag{13}$$

Here, using Equations (10) and (11), Jensen's inequality and Hölder's inequality for probability measures yields

$$\begin{aligned}\left|\mathsf{E}\left[\bar{f}(Z_1(t))\,\big(\Delta_1\kappa\big)\,g\big(\boldsymbol{Z}^n(t)\big)\right]\right| &\le \|g\|_{C_b(\mathbb{R}^{d\times n})}\,\mathsf{E}\left[|\bar{f}(Z_1(t))|\,\left|\Delta_1\kappa\right|\right]\\ &\lesssim \|g\|_{C_b(\mathbb{R}^{d\times n})}\,\mathsf{E}\left[(1+\|Z_1(t)\|+\mathsf{E}\left[\|Z_1(t)\|\right])^2\right]\\ &\lesssim \|g\|_{C_b(\mathbb{R}^{d\times n})}\Big(1+\mathsf{E}\left[\|Z(t)\|^2\right]+(\mathsf{E}\left[\|Z(t)\|\right])^2\Big)\\ &\lesssim \|g\|_{C_b(\mathbb{R}^{d\times n})}\Big(1+\mathsf{E}\left[\|Z(t)\|^4\right]\Big),\end{aligned} \tag{14}$$

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

$$\begin{aligned}\left|\mathsf{E}\left[\bar{f}(Z_1(t))\,\Delta_j\kappa\,g\big(\boldsymbol{Z}^n(t)\big)\right]\right| &\le \left|\mathsf{E}\left[\bar{f}(Z_1(t))\,\Delta_j\kappa\,g(\boldsymbol{Z}^n_{-j}(t))\right]\right|\\ &\quad+\left|\mathsf{E}\left[\bar{f}(Z_1(t))\,\Delta_j\kappa\,(\boldsymbol{Z}^n(t)-\boldsymbol{Z}^n_{-j}(t))^T\int_0^1\nabla g\big(s\boldsymbol{Z}^n(t)-(1-s)\boldsymbol{Z}^n_{-j}(t)\big)\mathrm{d}s\right]\right|.\end{aligned} \tag{15}$$

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

$$\mathsf{E}\left[\bar{f}(Z_1(t))\,\Delta_j\kappa\,g(\boldsymbol{Z}^n_{-j}(t))\right]=\mathsf{E}\left[\bar{f}(Z_1(t))\,\mathsf{E}\left[[]\right]\Delta_j\kappa\mid Z_1(t)\,g(\boldsymbol{Z}^n_{-j}(t))\right]\quad=0, \tag{16}$$

as $Z_j(t)$ has law μ_t and is independent of $\boldsymbol{Z}^n_{-j}$ and of $Z_1(t)$. Using Equations (10) and (11) and Hölder's inequality for probability measures, we bound

$$\begin{aligned}&\left|\mathsf{E}\left[\bar{f}(Z_1(t))\,\Delta_j\kappa\,(\boldsymbol{Z}^n(t)-\boldsymbol{Z}^n_{-j}(t))^T\int_0^1\nabla g\big(s\boldsymbol{Z}^n(t)-(1-s)\boldsymbol{Z}^n_{-j}(t)\big)\mathrm{d}s\right]\right|\\ &\le\sum_{i=1}^{d}\left\|\frac{\partial g}{\partial x_{j,i}}\right\|_{C_b(\mathbb{R}^{d\times n})}\mathsf{E}\left[\left|\bar{f}(Z_1(t))\right|\,|\Delta_j\kappa|\,\|Z_j(t)\|\right]\\ &\lesssim\sum_{i=1}^{d}\left\|\frac{\partial g}{\partial x_{j,i}}\right\|_{C_b(\mathbb{R}^{d\times n})}\Big(1+\mathsf{E}\left[\|Z(t)\|^4\right]\Big).\end{aligned} \tag{17}$$

Substituting Equations (16) and (17) into Equation (15), and the result and Equation (14) into Equation (13), and then substituting the result into Equation (8) and using Equation (4), 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 $\boldsymbol{Z}^n = \{Z_i\}_{i=1}^n$ be n independent processes each satisfying Equation (1) with independent underlying Wiener processes. Assume that

$$|a|_{C^3_{\mathrm{bd}}(\mathbb{R}^d\times\mathbb{R}\,;\,\mathbb{R}^d)} + |\sigma|_{C^3_{\mathrm{bd}}(\mathbb{R}^d\times\mathbb{R}\,;\,\mathbb{R}^{d\times d'})} + |\kappa_1|_{C^3_{\mathrm{bd}}(\mathbb{R}^d\times\mathbb{R}^d)} + |\kappa_2|_{C^3_{\mathrm{bd}}(\mathbb{R}^d\times\mathbb{R}^d)} < \infty, \tag{18}$$

and let $\boldsymbol{X}^n = \{X_i^n\}_{i=1}^n$ satisfy Equation (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

$$\left|\mathsf{E}\left[g(\boldsymbol{X}^n(T)) - g(\boldsymbol{Z}^n(T))\right]\right| \lesssim \left(1 + \mathsf{E}\left[\|\xi\|^4\right]\right)\, n^{-1}\ \ |g|_{C^3_{\mathrm{bd}}(\mathbb{R}^{d\times n})}. \tag{19}$$

Proof. Let $a \equiv (a_j)_{j=1}^d$ for $a_j : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ and $\varSigma \equiv (\varSigma_{j,j'})_{j,j'=1}^d := \sigma^T\sigma$ for $\varSigma_{j,j'} : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$. For $\boldsymbol{x} \equiv (\boldsymbol{x}_i)_{i=1}^n \equiv (x_{i,j})_{i\in\{1,\dots,n\},j\in\{1,\dots,d\}}$, define the operators (recall that μ_t is the law of $Z(t)$)

$$\begin{aligned}\mathcal{L}_n = \sum_{i=1}^n\sum_{j=1}^d \Bigg[& a_j\Bigg(\boldsymbol{x}_i, \frac{1}{n}\sum_{i'=1}^n \kappa_1(\boldsymbol{x}_i,\boldsymbol{x}_{i'})\Bigg)\frac{\partial}{\partial x_{i,j}} \\ & + \frac{1}{2}\sum_{j'=1}^d \varSigma_{j,j'}\Bigg(\boldsymbol{x}_i, \frac{1}{n}\sum_{i'=1}^n \kappa_2(\boldsymbol{x}_i,\boldsymbol{x}_{i'})\Bigg)\frac{\partial^2}{\partial x_{i,j}\partial x_{i,j'}}\Bigg]\end{aligned}$$

and

$$\begin{aligned}\mathcal{L}_\infty = \sum_{i=1}^n\sum_{j=1}^d \Bigg[& a_j\Bigg(\boldsymbol{x}_i, \int_{\mathbb{R}^d} \kappa_1(\boldsymbol{x}_i,z)\mu_t(\mathrm{d}z)\Bigg)\frac{\partial}{\partial x_{i,j}} \\ & + \frac{1}{2}\sum_{j'=1}^d \varSigma_{j,j'}\Bigg(\boldsymbol{x}_i, \int_{\mathbb{R}^d} \kappa_2(\boldsymbol{x}_i,z)\mu_t(\mathrm{d}z)\Bigg)\frac{\partial^2}{\partial x_{i,j}\partial x_{i,j'}}\Bigg].\end{aligned}$$

Consider the value function u satisfying the PDE

$$\begin{aligned}&\frac{\partial u}{\partial t}(t,\boldsymbol{x}) + \mathcal{L}_n u(t,\boldsymbol{x}) = 0, \text{ for } 0 \le t < T \text{ and } \boldsymbol{x} \in \mathbb{R}^{d\times n} \\ \text{and}\quad & u(T,\boldsymbol{x}) = g(\boldsymbol{x}) \text{ for } \boldsymbol{x} \in \mathbb{R}^{d\times n}.\end{aligned} \tag{20}$$

Under Equation (18), a strong solution for Equation (2) exists and is unique (Friedman, 1975, Theorem 1.1 in Chapter 5) and $u(t,\boldsymbol{x}) = \mathsf{E}\left[g(\boldsymbol{X}^n(T)) \,|\, \boldsymbol{X}^n(t) = \boldsymbol{x}\right]$, see (Friedman, 1975, Theorem 6.1 in Chapter 5). Given the existence of a solution to Equation (1) and its law, and recalling that the coefficients a and σ in Equation (1) are integrable and square-integrable, respectively, due to Equation (18) and

Equation (4), we define $U(t) := u(t, \boldsymbol{Z}^n(t))$ and apply Itô's formula (Friedman, 1975, Theorem 5.3 in Chapter 4) to arrive at

$$\begin{aligned} \mathsf{E}\left[g\big(\boldsymbol{Z}^n(T)\big)\right] - \mathsf{E}\left[g\big(\boldsymbol{X}^n(T)\big)\right] &= \mathsf{E}\left[U(T) - U(0)\right] \\ &= \int_0^T \mathsf{E}\left[\big(\mathcal{L}_\infty - \mathcal{L}_n\big)u(t, \boldsymbol{Z}^n(t))\right] \mathrm{d}t. \end{aligned} \tag{21}$$

The last equality is satisfied under the boundedness conditions in Equation (18), the integrability of Z, and the boundedness of the derivatives of u, which we will establish later. Then

$$\begin{aligned} \mathsf{E}\left[\big(\mathcal{L}_\infty - \mathcal{L}_n\big)u(t, \boldsymbol{Z}^n(t))\right] = &\sum_{i=1}^n \sum_{j=1}^d \mathsf{E}\left[\Delta_i a_j \frac{\partial u}{\partial x_{i,j}}(t, \boldsymbol{Z}^n(t))\right] \\ &+ \frac{1}{2}\sum_{j'=1}^d \mathsf{E}\left[\Delta_i \Sigma_{j,j'} \frac{\partial^2 u}{\partial x_{i,j}\partial x_{i,j'}}(t, \boldsymbol{Z}^n(t))\right], \end{aligned} \tag{22}$$

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

$$\Delta_i f := f\left(Z_i(t), \int_{\mathbb{R}^d} \kappa(Z_i(t), z)\mu_t(\mathrm{d}z)\right) - f\left(Z_i(t), \frac{1}{n}\sum_{j=1}^n \kappa(Z_i(t), Z_j(t))\right). \tag{23}$$

Using the triangle inequality, Lemma 2.2 and Equation (4), we bound

$$\begin{aligned} &\left|\mathsf{E}\left[\big(\mathcal{L}_\infty - \mathcal{L}_n\big)u(t, \boldsymbol{Z}^n(t))\right]\right| \\ &\lesssim n^{-1}\left(1 + \mathsf{E}\left[\|Z(t)\|^4\right]\right)\left(\sum_{i=1}^n \sum_{j=1}^d \left(\left\|\frac{\partial u(t,\cdot)}{\partial x_{i,j}}\right\|_{C_\mathrm{b}^1(\mathbb{R}^{d\times n})} + \sum_{j'=1}^d \left\|\frac{\partial^2 u(t,\cdot)}{\partial x_{i,j}\partial x_{i,j'}}\right\|_{C_\mathrm{b}^1(\mathbb{R}^{d\times n})}\right)\right) \\ &\lesssim n^{-1}\left(1 + \mathsf{E}\left[\|\xi\|^4\right]\right) |u(t,\cdot)|_{C_\mathrm{bd}^3(\mathbb{R}^{d\times n})}. \end{aligned}$$

It remains to show the bound $|u(t,\cdot)|_{C_\mathrm{bd}^3(\mathbb{R}^{d\times n})} \lesssim |g|_{C_\mathrm{bd}^3(\mathbb{R}^{d\times n})}$ for $t \leq T$. To that end, we use Proposition 3.3 in the following section with an appropriate definition of ν and ς in terms of a and σ, respectively, since Assumption 3.1 is satisfied for $q = 3$ given Equation (18); 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

$$\begin{aligned} &\left|\mathsf{E}\left[g(X_1^n(T), X_2^n(T), \ldots, X_k^n(T)) - g(Z_1(T), Z_2(T), \ldots, Z_k(T))\right]\right| \\ &\lesssim \frac{1}{n}\left(1 + \mathsf{E}\left[\|\xi\|^4\right]\right)\binom{dk+2}{3} \max_{\boldsymbol{\ell}\in\mathbb{N}^{dk}, 1\leq|\boldsymbol{\ell}|\leq 3}\left(\left\|\frac{\partial^{|\boldsymbol{\ell}|} g}{\partial \boldsymbol{x}^{\boldsymbol{\ell}}}\right\|_{C_\mathrm{b}(\mathbb{R}^{k\times d})}\right). \end{aligned}$$

Corollary 2.4 is useful when k is independent of n. For example, for $\tilde{g} : \mathbb{R}^d \to \mathbb{R}$, let

$$g(\boldsymbol{X}^n(T)) := \frac{1}{n}\sum_{i=1}^{n} \tilde{g}(X_i^n(T)).$$

Then, assuming that $|\tilde{g}|_{C^3_{\text{bd}}(\mathbb{R}^d)}$ and $\mathsf{E}\left[\|\xi\|^4\right]$ are finite, Corollary 2.4 immediately implies that

$$\left|\mathsf{E}\left[g(\boldsymbol{X}^n(T)) - g(\boldsymbol{Z}^n(T))\right]\right| = \left|\mathsf{E}\left[\tilde{g}(X_1^n(T)) - \tilde{g}(Z(T))\right]\right| = \mathcal{O}(n^{-1}).$$

REMARK 2.5. In the special case when $\kappa_2 = 0$, we can relax Equation (18) and only assume that

$$|a|_{C^2_{\text{bd}}(\mathbb{R}^d\times\mathbb{R};\mathbb{R}^d)} + |\sigma|_{C^2_{\text{bd}}(\mathbb{R}^d\times\mathbb{R};\mathbb{R}^{d\times d'})} + |\kappa_1|_{C^2_{\text{bd}}(\mathbb{R}^d\times\mathbb{R}^d)} < \infty. \tag{24}$$

The result would then also involve only the first and second derivatives of g,

$$\left|\mathsf{E}\left[g(\boldsymbol{X}^n(T)) - g(\boldsymbol{Z}^n(T))\right]\right| \lesssim n^{-1} \left(1 + \mathsf{E}\left[\|\xi\|^4\right]\right) |g|_{C^2_{\text{bd}}(\mathbb{R}^{d\times n})},$$

thus recovering a similar result to the one obtained, e.g. in (Kolokoltsov, 2010, 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

$$|a|_{C^3_{\text{bd}}(\mathbb{R}^d\times\mathbb{R}^m;\mathbb{R}^d)} + |\sigma|_{C^3_{\text{bd}}(\mathbb{R}^d\times\mathbb{R}^m;\mathbb{R}^{d\times d'})} + |\kappa_1|_{C^3_{\text{bd}}(\mathbb{R}^d\times\mathbb{R}^d;\mathbb{R}^m)} + |\kappa_2|_{C^3_{\text{bd}}(\mathbb{R}^d\times\mathbb{R}^d;\mathbb{R}^m)} < \infty,$$

The proof would follow the same steps by extending the proof of Lemma 2.2 to $f : \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}$. Such an extension can be done by either considering Euclidean norms in the proof of Lemma 2.2, or by adding and subtracting appropriate terms in Equation (7), so that we have m differences each having only one component of the interaction kernel being approximated, and apply Lemma 2.2 directly to each difference.

3. Moments Bounds for SDE Variations with Sobolev-Bounded Coefficients

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

$$\boldsymbol{X}^{t,\boldsymbol{x}}(s) = \boldsymbol{x} + \int_t^s \boldsymbol{\nu}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\mathrm{d}\tau + \int_t^s \boldsymbol{\varsigma}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\mathrm{d}\boldsymbol{W}(\tau), \quad s \in [t, T], \tag{25}$$

with drift coefficient $\boldsymbol{\nu} \equiv (\nu_i)_{i=1}^n : [0, T]\times\mathbb{R}^n \to \mathbb{R}^n$, a diffusion coefficient $\boldsymbol{\varsigma} \equiv (\varsigma_{im})_{i\in\{1,\ldots,n\},m\in\{1,\ldots,n'\}} : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n\times n'}$, and $\boldsymbol{W}$ is a vector of n' independent standard Wiener processes over a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ 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 Equation (2) which is a specific example of Equation (25) with $n \leftarrow nd$. However, we prove the results more generally for Equation (25) to emphasize that the particular structure of Equation (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

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

For brevity of presentation, we will define for the coefficients ν and ς in Equation (25) and for an integer $r \ge 0$,

$$\|\nu\|_{\mathrm{D}^r,\infty} := \|\partial\nu\|_{\mathrm{D}^{r-1},\infty} + \sum_{\boldsymbol{\ell}\in\mathbb{N}^n, |\boldsymbol{\ell}|=r} \max_{\substack{i\in\{1,\dots,n\}\\ \text{s.t. } \ell_i=0}} \left\| \frac{\partial^r \nu_i}{\partial \boldsymbol{x}^{\boldsymbol{\ell}}} \right\|_\infty \tag{26}$$

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

$$\|\varsigma\|_{\mathrm{D}^r,\ell^2_\infty} := \|\partial\varsigma\|_{\mathrm{D}^{r-1},\ell^2_\infty} + \sum_{\boldsymbol{\ell}\in\mathbb{N}^n, |\boldsymbol{\ell}|=r} \max_{\substack{i\in\{1,\dots,n\}\\ \text{s.t. } \ell_i=0}} \left(\sum_{m=1}^{n'} \left\| \frac{\partial^r \varsigma_{im}}{\partial \boldsymbol{x}^{\boldsymbol{\ell}}} \right\|_\infty^2 \right)^{1/2} \tag{27}$$

where $\partial\varsigma = \left(\frac{\partial \varsigma_{im}}{\partial x_i}\right)_{i=\{1,\dots,n\}, m=\{1,\dots,n'\}}$ and $\|\varsigma\|_{\mathrm{D}^{-1},\ell^2_\infty} := 0$. Finally, for the process $\boldsymbol{X}^{t,\boldsymbol{x}}$ in Equation (25), we define for any $p \ge 1$,

$$\begin{aligned} \|\boldsymbol{X}^{t,\boldsymbol{x}}\|_{\mathrm{D}^r,L^\infty([t,T];L^p(\Omega,\mathbb{P}))} &:= \|\partial \boldsymbol{X}^{t,\boldsymbol{x}}\|_{\mathrm{D}^{r-1},L^\infty([t,T];L^p(\Omega,\mathbb{P}))} \\ &\quad + \sum_{\boldsymbol{\ell}\in\mathbb{N}^n, |\boldsymbol{\ell}|=r} \max_{\substack{i\in\{1,\dots,n\}\\ \text{s.t. } \ell_i=0}} \sup_{t\le s\le T} \mathsf{E}\left[\left| \frac{\partial^r X_i^{t,\boldsymbol{x}}}{\partial \boldsymbol{x}^{\boldsymbol{\ell}}}(s) \right|^p \right]^{1/p} \end{aligned} \tag{28}$$

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

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

$$\sum_{r=1}^{q} \|\nu\|_{\mathrm{D}^r,\infty} + \|\varsigma\|_{\mathrm{D}^r,\ell^2_\infty} < C_q,$$

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

The previous assumption deserves some explanation. For example, focusing on the drift coefficient ν, the definition in Equation (26) for $r = 1$ simplifies to

$$\|\nu\|_{\mathrm{D}^1,\infty} = \max_i \left\| \frac{\partial \nu_i}{\partial x_i} \right\|_\infty + \sum_{\ell=1}^{n} \max_{i\ne\ell} \left\| \frac{\partial \nu_i}{\partial x_\ell} \right\|_\infty,$$

and $\|\varsigma\|_{\mathrm{D}^1,\ell^2_\infty}$ expands similarly. Hence, a sufficient condition for Assumption 3.1 when $q = 1$ is to bound

$$\left(\sum_{m=1}^{n'} \left\| \frac{\partial \varsigma_{im}}{\partial x_\ell} \right\|_\infty^2 \right)^{1/2} + \left\| \frac{\partial \nu_i}{\partial x_\ell} \right\|_\infty \le \begin{cases} \tilde{C}_1 & i = \ell, \\ \tilde{C}_1\, n^{-1} & i \ne \ell, \end{cases}$$

for all $i \in \{1,\dots,n\}$ and some constant $\tilde{C}_1 > 0$. For $r = 2$, the definition Equation (26) simplifies to

$$\|\nu\|_{\mathrm{D}^2,\infty} = \max_i \left\| \frac{\partial^2 \nu_i}{\partial x_i^2} \right\|_\infty + \sum_{\ell=1}^{n} \max_{i\neq \ell} \left\| \frac{\partial^2 \nu_i}{\partial x_i \partial x_\ell} \right\|_\infty + \sum_{\ell=1}^{n}\sum_{\ell'=1}^{n} \max_{i\notin\{\ell,\ell'\}} \left\| \frac{\partial^2 \nu_i}{\partial x_\ell \partial x_{\ell'}} \right\|_\infty,$$

and an additional condition on the second derivatives would be required Assumption 3.1 when $q = 2$, e.g.

$$\left(\sum_{m=1}^{n'} \left\| \frac{\partial^r \varsigma_{im}}{\partial \boldsymbol{x}^{\boldsymbol{\ell}}} \right\|_\infty^2 \right)^{1/2} + \left\| \frac{\partial^2 \nu_i}{\partial x_\ell \partial x_{\ell'}} \right\|_\infty \le \begin{cases} \tilde{C}_2 & i = \ell = \ell' \\ \tilde{C}_2 n^{-1} & i = \ell \neq \ell' \text{ or } i \neq \ell = \ell' \\ \tilde{C}_2 n^{-2} & i, \ell, \ell' \text{ are distinct,} \end{cases}$$

for all $i \in \{1,\dots,n\}$ and some constant $\tilde{C}_2 > 0$. In general, for any integer $q > 0$, a sufficient condition for Assumption 3.1 is

$$\left(\sum_{m=1}^{n'} \left\| \frac{\partial^{|\boldsymbol{\ell}|} \varsigma_{im}}{\partial \boldsymbol{x}^{\boldsymbol{\ell}}} \right\|_\infty^2 \right)^{1/2} + \left\| \frac{\partial^{|\boldsymbol{\ell}|} \nu_i}{\partial \boldsymbol{x}^{\boldsymbol{\ell}}} \right\|_\infty \le \tilde{C}_q\, n^{1-|\boldsymbol{\ell}+\boldsymbol{e}_i|_0} \quad \text{for all } i \in \{1,\dots,n\} \text{ and } \boldsymbol{\ell} \in \mathbb{N}^n \ : \ |\boldsymbol{\ell}| \le q,$$

for a constant $\tilde{C}_q > 0$ and where $\boldsymbol{e}_i$ is the i'th unit vector and $|\boldsymbol{\ell}|_0$ denotes the number of non-zero elements of $\boldsymbol{\ell}$.

LEMMA 3.2 (L^p bound of stochastic flows). Let Assumption 3.1 be satisfied for $q \in \{1, 2, 3\}$ and for $\boldsymbol{X}^{t,\boldsymbol{x}}$ in Equation (25) and any $p \ge 2$ then there exists constants $K_{q,p}$, independent of n and $\boldsymbol{x}$, such that

$$\|\boldsymbol{X}^{t,\boldsymbol{x}}\|_{\mathrm{D}^q, L^\infty([t,T];L^p(\Omega,\mathbb{P}))} \le K_{q,p}.$$

Proof. The proof is in Appendix B. □

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

$$\frac{\partial u}{\partial t}(t,\boldsymbol{x}) + \sum_{i=1}^{n} \nu_i(t,\boldsymbol{x}) \frac{\partial u}{\partial x_i}(t,\boldsymbol{x}) + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \left(\sum_{m=1}^{n'} \varsigma_{im}(t,\boldsymbol{x}) \varsigma_{jm}(t,\boldsymbol{x}) \right) \frac{\partial^2 u}{\partial x_i \partial x_j}(t,\boldsymbol{x}) = 0 \quad \text{and} \quad u(T,\boldsymbol{x}) = g(\boldsymbol{x}). \tag{29}$$

Assume that the coefficients, $\{\nu_i\}_{i=1}^n$ and $\{\varsigma_{im}\}_{i\in 1,\dots,n, m\in 1,\dots,n'}$, 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

$$|u(t,\cdot)|_{C^q_{\mathrm{bd}}(\mathbb{R}^n)} \le D_q\, |g|_{C^q_{\mathrm{bd}}(\mathbb{R}^n)}.$$

Proof. First note that under Assumption 3.1 for $q = 2$, u satisfies for all $t \in [0,T]$ and $\boldsymbol{x} \in \mathbb{R}^n$ (Friedman, 1975, Theorem 6.1 in Chapter 5)

$$u(t,\boldsymbol{x}) = \mathsf{E}\left[g(\boldsymbol{X}^{t,\boldsymbol{x}}(T)) \right]. \tag{30}$$

Next, we differentiate u with respect to the initial conditions and exchange the differentiation with the expectation in Equation (30) (Friedman, 1975, Theorem 5.5 in Chapter 5). Then, we bound

$$\begin{aligned}
\sum_{j=1}^{n}\left|\frac{\partial u}{\partial x_j}(t,\boldsymbol{x})\right| &= \sum_{j=1}^{n}\left|\mathsf{E}\left[\sum_{i=1}^{n}\frac{\partial g}{\partial x_i}(\boldsymbol{X}^{t,\boldsymbol{x}}(T))\,\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(T)\right]\right| \\
&\le \sum_{j=1}^{n}\sum_{i=1}^{n}\left\|\frac{\partial g}{\partial x_i}\right\|_{C_{\mathrm{b}}(\mathbb{R}^n)}\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(T)\right|\right] \\
&\le \sum_{i=1}^{n}\left\|\frac{\partial g}{\partial x_i}\right\|_{C_{\mathrm{b}}(\mathbb{R}^n)}\left(\sum_{j=1}^{n}\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(T)\right|\right]\right) \\
&\le \left(\sum_{i=1}^{n}\left\|\frac{\partial g}{\partial x_i}\right\|_{C_{\mathrm{b}}(\mathbb{R}^n)}\right)\left(\max_i\sum_{j=1}^{n}\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(T)\right|\right]\right) \\
&\le K_{1,2}\left(\sum_{i=1}^{n}\left\|\frac{\partial g}{\partial x_i}\right\|_{C_{\mathrm{b}}(\mathbb{R}^n)}\right),
\end{aligned}$$

by Lemma 3.2. Similarly,

$$\begin{aligned}
\sum_{j=1}^{n}\sum_{j'=1}^{n}\left|\frac{\partial^2 u}{\partial x_j\partial x_{j'}}(t,\boldsymbol{x})\right| &\le \sum_{j=1}^{n}\sum_{j'=1}^{n}\left|\mathsf{E}\left[\sum_{i=1}^{n}\frac{\partial g}{\partial x_i}(\boldsymbol{X}^{t,\boldsymbol{x}}(T))\,\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j\partial x_{j'}}(T)\right]\right| \\
&\quad+\sum_{j=1}^{n}\sum_{j'=1}^{n}\left|\mathsf{E}\left[\sum_{i=1}^{n}\sum_{i'=1}^{n}\frac{\partial^2 g}{\partial x_i\partial x_{i'}}(\boldsymbol{X}^{t,\boldsymbol{x}}(T))\,\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(T)\frac{\partial X_{i'}^{t,\boldsymbol{x}}}{\partial x_j}(T)\right]\right| \\
&\le \left(\sum_{i=1}^{n}\left\|\frac{\partial g}{\partial x_i}\right\|_{C_{\mathrm{b}}(\mathbb{R}^n)}\right)\left(\max_i\sum_{j=1}^{n}\sum_{j'=1}^{n}\mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j\partial x_{j'}}(T)\right|\right]\right) \\
&\quad+\left(\sum_{i=1}^{n}\sum_{i'=1}^{n}\left\|\frac{\partial^2 g}{\partial x_i\partial x_{i'}}\right\|_{C_{\mathrm{b}}(\mathbb{R}^n)}\right)\left(\max_i\sum_{j=1}^{n}\mathsf{E}\left[\left(\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(T)\right)^2\right]^{1/2}\right)^2 \\
&\le (K_{1,2}^2+K_{2,2})\left(\sum_{i=1}^{n}\left\|\frac{\partial g}{\partial x_i}\right\|_{C_{\mathrm{b}}(\mathbb{R}^n)}+\sum_{i=1}^{n}\sum_{i'=1}^{n}\left\|\frac{\partial^2 g}{\partial x_i\partial x_{i'}}\right\|_{C_{\mathrm{b}}(\mathbb{R}^n)}\right).
\end{aligned}$$

It is easy to see that the previous proof extends to $q = 3$ as well. □

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

$$\psi_r(x) := \begin{cases} \left(1-x^2\right)^r & |x| \le 1, \\ 0 & |x| > 1. \end{cases}$$

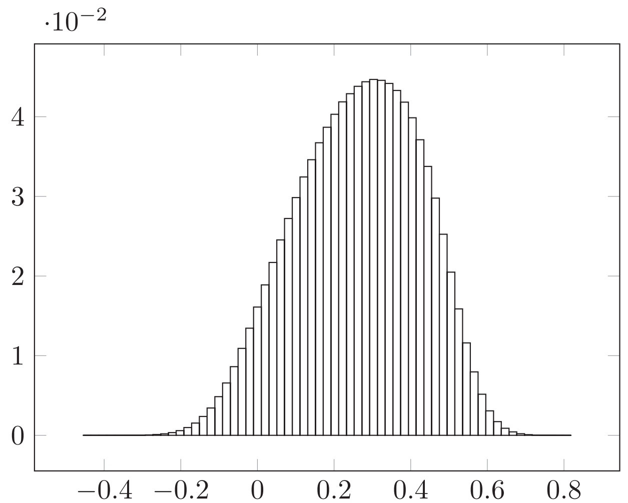

FIG. 1. A histogram of the values of $\{X_i^n(1)\}_{i=1}^n$ which follows the system of SDEs in Equation (2) with Equation (31). Here, the values were approximated using the Euler–Maruyama time-stepping scheme with 64 uniform time-steps and $n = 2,048$.

and note that ψ_0 is discontinuous while for $r > 0$ and all $k < r$ the k'th derivative of ψ_r exists is uniformly bounded, and the $(r-1)$'th derivative is Lipschitz continuous. Subsequently, consider the particle system Equation (2) with $d = 1$ and $X_i(0)$ being uniformly distributed in $[-1, 1]$, and for $x, y \in \mathbb{R}$ set

$$\begin{aligned} a(x,y) &:= 2(x-0.2)+y, \\ \sigma(x,y) &:= 0.2(1+y), \\ \kappa_1(x,y) &:= \psi_1(10|x-y|) \\ \text{and} \qquad \kappa_2(x,y) &:= \psi_1(5|x-y|). \end{aligned} \tag{31}$$

Note that the previous a, σ, κ_1, and κ_2, the latter two being only Lipschitz continuous, do not satisfy the conditions of Theorem 2.3. Nevertheless, existence and uniqueness of solutions to Equation (1) with the coefficient in Equation (31) can be established using, e.g. the results in De Raynal (2020); Raynal & Frikha (2021b). To approximate solutions to Equation (2), we use an Euler–Maruyama time-stepping scheme with a fixed number of time-steps.

We consider the sequence of systems, denoted by $\boldsymbol{X}^n$, satisfying Equation (2) with an increasing number of particles, n. See Fig. 1 for a histogram of the values of $X_i^n(1)$ 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 $\left(\mathsf{E}\left[(X_i^{2n} - X_i^n)^2\right]\right)^{1/2}$ and $|\mathsf{E}\left[g(X_i^{2n}) - g(X_i^n)\right]|$. The same convergence behaviour of these quantities was obtained with different numbers of uniform time-steps. Even though κ_1 and κ_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 κ_1, κ_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, e.g. the smoothness of the probability density.

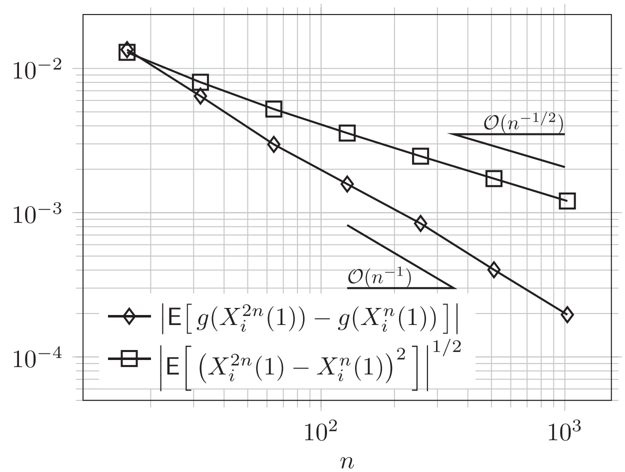

FIG. 2. Convergence rates $X_i^n(1)$ which follows the system of SDEs in Equation (2) with Equation (31) and $g(x) = \psi_0(10|x-0.2|)$. Here, the values were approximated using the Euler–Maruyama time-stepping scheme with $N = 64$ time-steps. Note that the rates are consistent with the predicted rates in Theorem 2.3 even though the coefficients of the SDE do not have sufficient smoothness as required by the theorem.

Funding

KAUST Office of Sponsored Research (OSR), under Award numbers (URF/1/2281-01-01, URF/1/2584-01-01 to R.T.) in the KAUST Competitive Research Grants Program Round 8; the Alexander von Humboldt Foundation, through the Alexander von Humboldt Professorship award.

REFERENCES

BALLY, V. & TALAY, D. (1995) The Euler scheme for stochastic differential equations: Error analysis with Malliavin calculus. *Math. Comput. Simulation*, **38**, 35–41.

BALLY, V. & TALAY, D. (1996) The law of the Euler scheme for stochastic differential equations: II. Convergence rate of the density. *Monte Carlo Methods Appl.*, **2**, 93–128.

BAYER, C., HALL, E. J. & TEMPONE, R. (2022) Weak error rates for option pricing under linear rough volatility. *Int. J. Theor. Appl. Finance*, **25**, 2250029.

BAYER, C., SZEPESSY, A. & TEMPONE, R. (2010) Adaptive weak approximation of reflected and stopped diffusions. *Monte Carlo Methods Appl.*, **16**, 1–67.

BENCHEIKH, O. & JOURDAIN, B. (2019) Bias behaviour and antithetic sampling in mean-field particle approximations of SDEs nonlinear in the sense of McKean. *ESAIM: Proc. Surv.* EDP Sciences, **65**, 219–235.

CARDALIAGUET, P. (2010) Notes on mean field games. *Technical Report*, Technical report.

CHASSAGNEUX, J.-F., SZPRUCH, L. & TSE, A. (2022a) Weak quantitative propagation of chaos via differential calculus on the space of measures. *Ann. Appl. Probab.*, **32**, 6.

CHASSAGNEUX, J.-F., CRISAN, D. & DELARUE, F. (2022b) *A Probabilistic Approach to Classical Solutions of the Master Equation for Large Population Equilibria*, vol. **280**. American Mathematical Society.

CRISAN, D. & XIONG, J. (2010) Approximate Mckean–Vlasov representations for a class of SPDEs. *STOCHASTICS*. Taylor & Francis, **82**, 53–68.

DE RAYNAL & FRIKHA, N. (2021) From the backward Kolmogorov PDE on the Wasserstein space to propagation of chaos for McKean-Vlasov SDEs. *J. Math. Pures Appl.*, **156**, 1–124.

DE RAYNAL & FRIKHA, N. (2022) Well-posedness for some non-linear diffusion processes and related PDE on the Wasserstein space. *Journal de Mathématiques Pures et Appliquées*, **159**, 1–167.

DE RAYNAL, P. E. CHAUDRU (2020) Strong well posedness of McKean–Vlasov stochastic differential equations with hölder drift. *Stochastic Process. Appl.*, **130**, 79–107.

FRIEDMAN, A. (1975) *Stochastic Differential Equations and Applications,* Volume **1**. Academic Press, New York.

HAMMERSLEY, W. R. P., ŠIŠKA, D. & SZPRUCH, Ł. (2021) McKean–Vlasov SDEs under measure dependent Lyapunov conditions. *Ann. I. H. Poincaré PR*, **57**, 1032–1057.

HOEL, H., HÄPPÖLÄ, J. & TEMPONE, R. (2016) Construction of a mean square error adaptive Euler–Maruyama method with applications in multilevel Monte Carlo. *Springer Proceedings in Mathematics & Statistics* (R. COOLS & D. NUYENS eds), Springer International Publishing, pp. 29–86.

HOEL, H., VON SCHWERIN, E., SZEPESSY, A. & TEMPONE, R. (2014) Implementation and analysis of an adaptive multilevel Monte Carlo algorithm. *Monte Carlo Methods Appl.* Walter de Gruyter GmbH, **20**, 1–41.

KATSOULAKIS, M. A. & SZEPESSY, A. (2006) Stochastic hydrodynamical limits of particle systems. *Communications in Mathematical Sciences*, **4**, 513–549.

KLOEDEN, P. E. & PLATEN, E. (1992) *Numerical Solution of Stochastic Differential Equations, Volume 23 of Applications of Mathematics (New York)*. Berlin, Heidelberg: Springer.

KOLOKOLTSOV, V. N. (2010) *Nonlinear Markov Processes and Kinetic Equations*, vol. **182**. Cambridge University Press p. 7.

LASRY, J.-M. & LIONS, P.-L. (2007) Mean field games. *Japan. J. Math. (N.S.)*, **2**, 229–260.

MISCHLER, S., MOUHOT, C. & WENNBERG, B. (2013) A new approach to quantitative propagation of chaos for drift, diffusion and jump processes. *Probab. Theory Relat. Fields*, **161**, 1–59.

MISHURA, Y. & VERETENNIKOV, A. (2021) Existence and uniqueness theorems for solutions of McKean–Vlasov stochastic equations. *Theory Probab. Math. Statist.*, **103**, 59–101.

MOON, K.-S., SZEPESSY, A., TEMPONE, R. & ZOURARIS, G. E. (2005) Convergence rates for adaptive weak approximation of stochastic differential equations. *Stochastic Anal. Appl.*, **23**, 511–558.

MORDECKI, E., SZEPESSY, A., TEMPONE, R. & ZOURARIS, G. E. (2008) Adaptive weak approximation of diffusions with jumps. *SIAM J. Numer. Anal.*, **46**, 1732–1768.

SZEPESSY, A., TEMPONE, R. & ZOURARIS, G. E. (2001) Adaptive weak approximation of stochastic differential equations. *Comm. Pure Appl. Math.*, **54**, 1169–1214.

SZNITMAN, A.-S. (1991) Topics in propagation of chaos. *Lecture Notes in Mathematics*. Berlin, Heidelberg: Springer, pp. 165–251.

SZPRUCH, Ł. & TSE, A. (2021) Antithetic multilevel sampling method for nonlinear functionals of measure. *Ann. Appl. Probab.*, **31**, 06.

TALAY, D. & TUBARO, L. (1990) Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, **8**, 483–509.

VON SCHWERIN, E. & SZEPESSY, A. (2010) A stochastic phase-field model determined from molecular dynamics. *M2AN. Math. Model. Numer. Anal.*, **44**, 627–646.

A. Proof of Proposition 2.1

We include a technical proof showing Equation (3) in Proposition 2.1 for general $p \in \{1, 2, \ldots\}$. We first state a result bounding the cardinality of sets of multi-indices that do not contain unique indices.

LEMMA A1 (Multi-indices with no unique indices). Given a set $\mathcal{I}$ and integer $p \geq 2$, let $\mathcal{I}^p$ be the p-ary Cartesian power of $\mathcal{I}$. Define the set of multi-indices with repeated indices as:

$$\mathcal{J}_{\mathcal{I}^p} = \Big\{(k_1, k_2, \ldots, k_p) \in \mathcal{I}^p \ : \ \forall j \in \{1, \ldots, p\} \ \exists j' \in \{1, \ldots, p\} \setminus \{j\} \text{ s.t. } k_j = k_{j'}\Big\}. \tag{A.1}$$

Then, there exists a constant c_p such that

$$|\mathcal{J}_{\mathcal{I}^p}| \leq c_p \, |\mathcal{I}|^{\lfloor p/2 \rfloor}. \tag{A.2}$$

Proof. We can directly compute that $|\mathcal{J}_{\mathcal{I}^2}| = |\mathcal{J}_{\mathcal{I}^3}| = |\mathcal{I}|$. Then, assuming that the result is true for p, we consider $\mathcal{J}_{\mathcal{I}^{p+2}}$ by counting cases where any index, i, is repeated j times,

$$
\begin{aligned}
|\mathcal{J}_{\mathcal{I}^{p+2}}| &= |\mathcal{I}| + \sum_{i=1}^{|\mathcal{I}|}\sum_{j=2}^{p}\binom{p+2}{j}|\mathcal{J}_{(\mathcal{I}\setminus\{i\})^{p+2-j}}| \\
&= |\mathcal{I}| + \sum_{i=1}^{|\mathcal{I}|}\sum_{j=0}^{p-2}\binom{p+2}{j+2}|\mathcal{J}_{(\mathcal{I}\setminus\{i\})^{p-j}}| \\
&\le |\mathcal{I}| + \sum_{i=1}^{|\mathcal{I}|}\sum_{j=0}^{p-2}\binom{p+2}{j+2}c_{p-j}(|\mathcal{I}|-1)^{\lfloor\frac{p-j}{2}\rfloor} \\
&\le |\mathcal{I}| + |\mathcal{I}|^{1+\lfloor p/2\rfloor}\sum_{j=0}^{p-2}\binom{p+2}{j+2}c_{p-j}(|\mathcal{I}|-1)^{\lfloor\frac{p-j}{2}\rfloor-\lfloor p/2\rfloor} =: c_{p+2}\,|\mathcal{I}|^{\lfloor 1+p/2\rfloor}.
\end{aligned}
$$

□

Proof of Equation (3) in Proposition 2.1. We again set, without loss of generality, $i = 1$, denote $\mathcal{I} := \{1,\ldots,n\}$ and recall the definition in Equation (5) to write

$$
\begin{aligned}
\mathsf{E}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\kappa(Z_1(t),Z_j(t)) - \int_{\mathbb{R}^d}\kappa(Z_1(t),z)\mu_t(\mathrm{d}z)\right)^{2p}\right] &= \mathsf{E}\left[\left(\frac{1}{n}\sum_{j\in\mathcal{I}}\Delta_j\kappa\right)^{2p}\right] \\
&= n^{-2p}\sum_{\boldsymbol{k}\in\mathcal{I}^{2p}}\mathsf{E}\left[\prod_{j=1}^{2p}\Delta_{k_j}\kappa\right]
\end{aligned}
$$

Let $\mathcal{I}^{2p} = \mathcal{U}_{1,-1}\cup\mathcal{U}_{1,1}\cup\mathcal{J}_{\mathcal{I}^{2p}}$ with $\mathcal{J}_{\mathcal{I}^{2p}}$ is as defined in Equation (A.1), i.e. the set of all multi-indices from $\mathcal{I}^{2p}$ with no unique indices, $\mathcal{U}_{1,1}$ is the set of multi-indices with exactly one unique index equal to 1, and $\mathcal{U}_{1,-1}$ is the set of multi-indices with at least one unique index different from 1. For $\boldsymbol{k}\in\mathcal{U}_{1,-1}$, i.e. there is an $\ell\in\mathcal{I}\setminus\{1\}$ such that $k_\ell \ne k_j$ for all $j\ne\ell$, we have

$$
\mathsf{E}\left[\prod_{j=1}^{2p}\Delta_{k_j}\kappa\right] = \mathsf{E}\left[\Delta_{k_\ell}\kappa\prod_{j=1,j\ne\ell}^{2p}\Delta_{k_j}\kappa\right] = \mathsf{E}\left[\mathsf{E}\left[\Delta_{k_\ell}\kappa\mid Z_1(t)\right]\mathsf{E}\left[\prod_{j=1,j\ne\ell}^{2p}\Delta_{k_j}\kappa\mid Z_1(t)\right]\right] = 0,
$$

since, given $Z_1(t)$, $\Delta_{k_\ell}\kappa$ is independent of $\{\Delta_{k_j}\kappa\}_{j=1,j\ne\ell}^{n}$ and $\mathsf{E}\left[\Delta_\ell\kappa\mid Z_1(t)\right] = 0$. For other $\boldsymbol{k}\in\mathcal{J}_{\mathcal{I}^{2p}}\cup\mathcal{U}_{1,1}$, we bound using Hölder's inequality and the Lipschitz assumption on κ,

$$
\mathsf{E}\left[\prod_{j=1}^{2p}\Delta_{k_j}\kappa\right] \le \prod_{j=1}^{2p}\mathsf{E}\left[(\Delta_j\kappa)^{2p}\right]^{1/(2p)} \le 2^{2p}\,C^{2p}\,\mathsf{E}\left[\|Z(t)\|^{2p}\right].
$$

Finally noting, by Lemma A.1, that

$$
\begin{aligned}
&\left|\mathcal{J}_{\mathcal{I}^{2p}}\right| \le c_{2p}\,n^p, \\
&\left|\mathcal{U}_{1,1}\right| = 2p\left|\mathcal{J}_{(\mathcal{I}\setminus\{1\})^{2p-1}}\right| \le 2p\,c_{2p-1}\,(n-1)^{(2p-1)/2} \le 2p\,c_{2p-1}\,n^p,
\end{aligned}
$$

we conclude that

$$\mathsf{E}\left[\left(\frac{1}{n}\sum_{j=1}^{n}\Delta_j\kappa\right)^{2p}\right] \lesssim n^{-p} \sup_{0\le t\le T} \mathsf{E}\left[\|Z(t)\|^{2p}\right].$$

□

B. Proof of Lemma 3.2

We first note the following inequality for any index sets $\mathcal{I}$ and $\mathcal{J}$ and any sequence $\{a_{i,j}\}_{i\in\mathcal{I},j\in\mathcal{J}}$,

$$\left(\sum_{j\in\mathcal{J}}\left(\sum_{i\in\mathcal{I}} a_{i,j}\right)^2\right)^{1/2} \le \sum_{i\in\mathcal{I}}\left(\sum_{j\in\mathcal{J}} a_{i,j}^2\right)^{1/2}, \tag{B.1}$$

which can be shown by expanding the square and using Hölder's inequality. Using Equation (B.1) and Jensen's inequality, we can show the following inequality for any random variables $(Y_i)_{i\in\mathcal{I}}$ and measurable sequence $\{a_{i,j}\}_{i\in\mathcal{I},j\in\mathcal{J}}$ and $p \ge 1$

$$\begin{aligned}\mathsf{E}\left[\left(\sum_{j\in\mathcal{J}}\left(\sum_{i\in\mathcal{I}} a_{i,j}Y_i\right)^2\right)^{p/2}\right] &\le \mathsf{E}\left[\left(\sum_{i\in\mathcal{I}}\left(\sum_{j\in\mathcal{J}} a_{i,j}^2\right)^{1/2}|Y_i|\right)^{p}\right]\\ &\le \left(\sum_{i\in\mathcal{I}}\left(\sum_{j\in\mathcal{J}} a_{i,j}^2\right)^{1/2}\right)^{p-1} \sum_{i\in\mathcal{I}}\left(\sum_{j\in\mathcal{J}} a_{i,j}^2\right)^{1/2} \mathsf{E}\left[|Y_i|^p\right] \qquad (B.2)\\ &\le \left(\sum_{i\in\mathcal{I}}\left(\sum_{j\in\mathcal{J}} a_{i,j}^2\right)^{1/2}\right)^{p} \max_{i\in\mathcal{I}} \mathsf{E}\left[|Y_i|^p\right].\end{aligned}$$

In addition, note that for a positive sequence $\{a_{i,j}\}_{i\in\mathcal{I},j\in\mathcal{J}}$, we have

$$\max_{i\in\mathcal{I}}\sum_{j\in\mathcal{J}} a_{i,j} \le \max_{i\in\mathcal{I}} a_{i,i} + \max_{i\in\mathcal{I}}\sum_{j\in\mathcal{J},i\ne j}^{n} a_{i,j} \le \max_{i\in\mathcal{I}} a_{i,i} + \sum_{j\in\mathcal{J}}\max_{i\in\mathcal{I},j\ne i} a_{i,j} \tag{B.3}$$

B.1 *First variation*

First, note that the process $\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}$ exists under Assumption 3.1 for $q = 1$ and satisfies for $s \ge t$ the SDE

$$\begin{aligned}\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(s) = \Delta_{i,j} &+ \int_t^s \sum_{k=1}^{n} \frac{\partial \nu_i}{\partial x_k}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\,\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\,\mathrm{d}\tau\\ &+ \sum_{m=1}^{n'}\int_t^s \sum_{k=1}^{n} \frac{\partial \varsigma_{im}}{\partial x_k}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\,\mathrm{d}W_m(\tau),\end{aligned}$$

where $\Delta_{i,j} = 1$ whenever $i = j$ and zero otherwise, cf. (Friedman, 1975, Theorem 5.3 in Chapter 5). By Itô's formula

$$\begin{aligned}\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(s)\right|^p\right] &\le \Delta_{i,j} + p\int_t^s \mathsf{E}\left[\left|\sum_{k=1}^n \frac{\partial \nu_i}{\partial x_k}(\tau,\boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\,\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^{p-1}\right]\mathrm{d}\tau \\ &+ \frac{p(p-1)}{2}\int_t^s \sum_{m=1}^{n'} \mathsf{E}\left[\left(\sum_{k=1}^n \frac{\partial \varsigma_{im}}{\partial x_k}(\tau,\boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right)^2\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^{p-2}\right]\mathrm{d}\tau.\end{aligned} \tag{B.4}$$

For the term involving ν_i, using Young's inequality, we can bound

$$\begin{aligned}&\mathsf{E}\left[\left|\sum_{k=1}^n \frac{\partial \nu_i}{\partial x_k}(\tau,\boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\,\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^{p-1}\right] \\ &\le \mathsf{E}\left[\left(\sum_{k=1}^n\left\|\frac{\partial \nu_i}{\partial x_k}\right\|_\infty\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|\right)\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^{p-1}\right] \\ &\le \frac{1}{p}\mathsf{E}\left[\left(\sum_{k=1}^n\left\|\frac{\partial \nu_i}{\partial x_k}\right\|_\infty\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|\right)^p\right] + \frac{p-1}{p}\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right].\end{aligned} \tag{B.5}$$

By using Jensen's inequality, Equation (B.2) and Assumption 3.1, we can further bound

$$\begin{aligned}&\mathsf{E}\left[\left(\sum_{k=1}^n\left\|\frac{\partial \nu_i}{\partial x_k}\right\|_\infty\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|\right)^p\right] \\ &\le 2^{p-1}\left\|\frac{\partial \nu_i}{\partial x_j}\right\|_\infty^p \mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right] + 2^{p-1}\mathsf{E}\left[\left(\sum_{k=1,k\neq j}^n\left\|\frac{\partial \nu_i}{\partial x_k}\right\|_\infty\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|\right)^p\right] \\ &\le 2^{p-1}\left\|\frac{\partial \nu_i}{\partial x_j}\right\|_\infty^p \mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right] + 2^{p-1}C_1^p\max_{k\neq j}\mathsf{E}\left[\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right].\end{aligned} \tag{B.6}$$

For the term involving $\varsigma_{i\cdot}$, using Young's inequality and bounding the derivatives of $\varsigma_{i\cdot}$,

$$\begin{aligned}&\mathsf{E}\left[\left(\sum_{k=1}^n \frac{\partial \varsigma_{im}}{\partial x_k}(\tau,\boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right)^2\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^{p-2}\right] \\ &\le \frac{2}{p}\mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k=1}^n\left\|\frac{\partial \varsigma_{im}}{\partial x_k}\right\|_\infty\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|\right)^2\right)^{p/2}\right] + \frac{p-2}{p}\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right].\end{aligned} \tag{B.7}$$

Using Jensen's inequality and Equation (B.2) and Assumption 3.1, we can further bound

$$
\begin{aligned}
&\mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k=1}^{n}\left\|\frac{\partial\varsigma_{im}}{\partial x_k}\right\|_\infty\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|\right)^2\right)^{p/2}\right]\\
&\le 2^{p-1}\left(\sum_{m=1}^{n'}\left\|\frac{\partial\varsigma_{im}}{\partial x_j}\right\|_\infty^2\right)^{p/2}\mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right]\\
&\qquad+2^{p-1}\mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k=1,k\neq j}^{n}\left\|\frac{\partial\varsigma_{im}}{\partial x_k}\right\|_\infty\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|\right)^2\right)^{p/2}\right]\\
&\le 2^{p-1}\left(\sum_{m=1}^{n'}\left\|\frac{\partial\varsigma_{im}}{\partial x_j}\right\|_\infty^2\right)^{p/2}\mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right]+2^{p-1}C_1^p\max_{k\neq j}\mathsf{E}\left[\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right].
\end{aligned}
\tag{B.8}
$$

Combining Equation (B.6) with Equation (B.5) and Equation (B.8) with Equation (B.7) and substituting the result into Equation (B.4) and simplifying yield

$$
\begin{aligned}
\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(s)\right|^p\right]\le\ &\Delta_{i,j}+c_1\int_t^s\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right]\mathrm{d}\tau\\
&+c_2\,C_1^p\int_t^s\max_{k\neq j}\mathsf{E}\left[\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right]\mathrm{d}\tau\\
&+c_2\,a_{i,j}^p\int_t^s\mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right]\mathrm{d}\tau,
\end{aligned}
\tag{B.9}
$$

where

$$
\begin{aligned}
c_1 &:= \frac{p(p-1)}{2},\\
c_2 &:= p\,2^{p-1}\\
\text{and}\qquad a_{i,j}^p &:= \left\|\frac{\partial\nu_i}{\partial x_j}\right\|_\infty^p+\left(\sum_{m=1}^{n'}\left\|\frac{\partial\varsigma_{im}}{\partial x_j}\right\|_\infty^2\right)^{p/2},
\end{aligned}
$$

so that $\max_i a_{i,i}+\sum_{j=1}^n\max_{i\neq j}a_{i,j}\le C_1$ by Assumption 3.1. Then, taking the maximum over all i in Equation (B.9), we arrive at

$$
\max_i\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(s)\right|^p\right]\le 1+\left(c_1+2C_1^pc_2\right)\int_t^s\max_i\mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right]\mathrm{d}\tau,
$$

and using Grönwall's inequality yields

$$\max_i \mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(s)\right|^p\right] \leq \exp\big((c_1 + 2C_1^p c_2)(s-t)\big) =: D_1, \tag{B.10}$$

for all $t \leq s \leq T$. Taking the maximum over all $i \neq j$ in Equation (B.9), we arrive at

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

Using Grönwall's inequality and Equation (B.10) yields

$$\begin{aligned}\max_{i\neq j} \mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(s)\right|^p\right] &\leq c_2\left(\max_{i\neq j} a_{i,j}^p\right)\exp\big((c_1 + c_2 C_1^p)(s-t)\big)\int_t^s \mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right]\mathrm{d}\tau \\ &\leq c_2\left(\max_{i\neq j} a_{i,j}^p\right)\exp\big((c_1 + c_2 C_1^p)(s-t)\big)\left(\int_t^s D_1 \mathrm{d}\tau\right) \\ &=: \widetilde{D}_1\left(\max_{i\neq j} a_{i,j}^p\right).\end{aligned} \tag{B.11}$$

Finally, using Equations (B.10) and (B.11) we arrive at:

$$\begin{aligned}\|\boldsymbol{X}^{t,\boldsymbol{x}}\|_{\mathrm{D}^1, L^\infty([t,T];L^p(\Omega,\mathbb{P}))} &= \max_i \sup_{t\leq s\leq T} \mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_i}(s)\right|^p\right]^{1/p} + \sum_{j=1}^n \max_{i\neq j}\sup_{t\leq s\leq T} \mathsf{E}\left[\left|\frac{\partial X_i^{t,\boldsymbol{x}}}{\partial x_j}(s)\right|^p\right]^{1/p} \\ &\leq D_1^{1/p} + \widetilde{D}_1^{1/p}\sum_{j=1}^n \max_{i\neq j} a_{i,j} \\ &\leq D_1^{1/p} + \widetilde{D}_1^{1/p} C_1 \\ &=: K_{1,p}.\end{aligned} \tag{B.12}$$

B.2 *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 $\left\{\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(s)\right\}_{s\in[t,T]}$ exists under Assumption 3.1 for $q = 2$ and satisfies for $s \in [t,T]$ the SDE

$$\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(s) = \int_t^s \sum_{k=1}^n \frac{\partial \nu_i}{\partial x_k}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau)) \frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\,\mathrm{d}\tau$$
$$+ \int_t^s \sum_{k=1}^n \sum_{k'=1}^n \frac{\partial^2 \nu_i}{\partial x_k \partial x_{k'}}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau)) \frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau) \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\,\mathrm{d}\tau$$
$$+ \sum_{m=1}^{n'} \int_t^s \sum_{k=1}^n \frac{\partial \varsigma_{im}}{\partial x_k}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau)) \frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\,\mathrm{d}W_m(\tau)$$
$$+ \sum_{m=1}^{n'} \int_t^s \sum_{k=1}^n \sum_{k'=1}^n \frac{\partial^2 \varsigma_{im}}{\partial x_k \partial x_{k'}}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau)) \frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau) \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\,\mathrm{d}W_m(\tau),$$

cf. (Friedman, 1975, Theorem 5.4 in Chapter 5). By Itô's formula,

$$\mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(s)\right|^p\right] \leq \; p \int_t^s \mathsf{E}\left[(f_1 + f_3)\, \left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^{p-1}\right] \mathrm{d}\tau$$
$$+ \frac{p(p-1)}{2} \int_t^s \mathsf{E}\left[(f_2 + f_4)\, \left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^{p-2}\right] \mathrm{d}\tau, \tag{B.13}$$

where

$$f_1 := \left|\sum_{k=1}^n \frac{\partial \nu_i}{\partial x_k}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau)) \frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|,$$
$$f_2 := \sum_{m=1}^{n'} \left(\sum_{k=1}^n \left(\frac{\partial \varsigma_{im}}{\partial x_k}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\right) \frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right)^2,$$
$$f_3 := \left|\sum_{k=1}^n \sum_{k'=1}^n \frac{\partial^2 \nu_i}{\partial x_k \partial x_{k'}}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau)) \frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau) \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right|$$
$$\text{and} \qquad f_4 := \sum_{m=1}^{n'} \left(\sum_{k=1}^n \sum_{k'=1}^n \left(\frac{\partial^2 \varsigma_{im}}{\partial x_k \partial x_{k'}}(\tau, \boldsymbol{X}^{t,\boldsymbol{x}}(\tau))\right) \frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau) \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right)^2.$$

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

$$\mathsf{E}\left[f_u \left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^{p-q_u}\right] \leq \frac{q_u}{p} \mathsf{E}\left[f_u^{p/q_u}\right] + \frac{p-q_u}{p} \mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right]. \tag{B.14}$$

We now turn our attention to bounding $\mathsf{E}\left[f_u^{p/q_u}\right]$ by primarily using Jensen's inequality and Equation (B.2). We present the proof bounding $\mathsf{E}\left[f_2^{p/2}\right]$ and $\mathsf{E}\left[f_4^{p/2}\right]$, Bounding $\mathsf{E}\left[f_1^{p}\right]$ and $\mathsf{E}\left[f_3^{p}\right]$ is analogous,

$$\begin{aligned}
\mathsf{E}\left[f_2^{p/2}\right] &\leq \mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k=1}^{n}\left\|\frac{\partial \varsigma_{im}}{\partial x_k}\right\|_\infty\left|\frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|\right)^2\right)^{p/2}\right] \\
&\leq 3^{p-1}\sum_{k\in\{j,j'\}}\left(\sum_{m=1}^{n'}\left\|\frac{\partial \varsigma_{im}}{\partial x_k}\right\|_\infty^2\right)^{p/2}\mathsf{E}\left[\left|\frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right] \\
&\quad + 3^{p-1}\mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k\in\{1,2,\dots,n\}\setminus\{j,j'\}}\left\|\frac{\partial \varsigma_{im}}{\partial x_k}\right\|_\infty\left|\frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|\right)^2\right)^{p/2}\right] \\
&\leq 3^{p-1}\sum_{k\in\{j,j'\}}\left(\sum_{m=1}^{n'}\left\|\frac{\partial \varsigma_{im}}{\partial x_k}\right\|_\infty^2\right)^{p/2}\left(\mathsf{E}\left[\left|\frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right]\right) \\
&\quad + 3^{p-1}C_1^p\max_{k\in\{1,2,\dots,n\}\setminus\{j,j'\}}\mathsf{E}\left[\left|\frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right],
\end{aligned} \tag{B.15}$$

where we used Equation (B.2) in the last step. On the other hand,

$$\begin{aligned}
\mathsf{E}\left[f_4^{p/2}\right] &\leq \mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k=1}^{n}\sum_{k'=1}^{n}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_k \partial x_{k'}}\right\|_\infty\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right)^2\right)^{p/2}\right] \\
&\leq 4^{p-1}\left(\sum_{m=1}^{n'}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{j'}}\right\|_\infty^2\right)^{p/2}\mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\frac{\partial X_{j'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right|^p\right] \\
&\quad + 4^{p-1}\mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k'=1,k'\neq j'}^{n}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{k'}}\right\|_\infty\frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right)^2\right)^{p/2}\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right] \\
&\quad + 4^{p-1}\mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k=1,k\neq j}^{n}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_k \partial x_{j'}}\right\|_\infty\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_k}(\tau)\right)^2\right)^{p/2}\left|\frac{\partial X_{j'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right|^p\right] \\
&\quad + 4^{p-1}\mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k=1,k\neq j}^{n}\sum_{k'=1,k'\neq j'}^{n}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_k \partial x_{k'}}\right\|_\infty\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right)^2\right)^{p/2}\right].
\end{aligned} \tag{B.16}$$

Looking at each term separately and using the bound on the first variation Equation (B.12),

$$\mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\frac{\partial X_{j'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right|^p\right] \leq \left(\mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^{2p}\right]\right)^{1/2}\left(\mathsf{E}\left[\left|\frac{\partial X_{j'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right|^{2p}\right]\right)^{1/2}$$
$$\leq K_{1,2p}^{2p}.$$

Moreover, using Hölder's inequality, Equation (B.2) and the bound on the first variation Equation (B.12), we write

$$\mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k'=1,k'\neq j'}^{n}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{k'}}\right\|_\infty \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right)^2\right)^{p/2}\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^p\right]$$
$$\leq \left(\sum_{k'=1}^{n}\left(\sum_{m=1}^{n'}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{k'}}\right\|_\infty^2\right)^{1/2}\right)^p\left(\max_{k'\neq j'}\mathsf{E}\left[\left|\frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right|^{2p}\right]\right)^{1/2}\mathsf{E}\left[\left|\frac{\partial X_j^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^{2p}\right]^{1/2}$$
$$\leq K_{1,2p}^p\left(\sum_{k'=1}^{n}\left(\sum_{m=1}^{n'}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{k'}}\right\|_\infty^2\right)^{1/2}\right)^p\left(\max_{k'\neq j'}\mathsf{E}\left[\left|\frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right|^{2p}\right]\right)^{1/2},$$

and

$$\mathsf{E}\left[\left(\sum_{m=1}^{n'}\left(\sum_{k=1,k\neq j}^{n}\sum_{k'=1,k'\neq j'}^{n}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_k \partial x_{k'}}\right\|_\infty \frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right)^2\right)^{p/2}\right]$$
$$\leq \left(\sum_{k=1,k\neq j}^{n}\sum_{k'=1,k'\neq j'}^{n}\left(\sum_{m=1}^{n'}\left\|\frac{\partial^2 \varsigma_{im}}{\partial x_k \partial x_{k'}}\right\|_\infty^2\right)^{1/2}\right)^p \max_{\substack{k,k'\\k\neq j,k'\neq j'}}\mathsf{E}\left[\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right|^p\right]$$
$$\leq C_2^p\left(\max_{k,k\neq j}\mathsf{E}\left[\left|\frac{\partial X_k^{t,\boldsymbol{x}}}{\partial x_j}(\tau)\right|^{2p}\right]\right)^{1/2}\left(\max_{k',k'\neq j'}\mathsf{E}\left[\left|\frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau)\right|^{2p}\right]\right)^{1/2}.$$

Hence

$$\mathsf{E}\left[f_4^{p/2}\right] \leq 4^{p-1}\left(K_{1,2p}^{2p}F_{\varsigma 1,i,j,j'} + K_{1,2p}^p F_{\varsigma 2,i,j,j'} + K_{1,2p}^p F_{\varsigma 2,i,j',j} + C_2^p F_{3,j,j'}\right)$$
$$\leq D_2\left(F_{\varsigma 1,i,j,j'} + F_{\varsigma 2,i,j,j'} + F_{\varsigma 2,i,j',j} + F_{3,j,j'}\right),$$

where

$$F_{\varsigma 1,i,j,j'} := \left(\sum_{m=1}^{n'} \left\| \frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{j'}} \right\|_\infty^2 \right)^{p/2},$$

$$F_{\varsigma 2,i,j,j'} := \left(\sum_{k'=1}^{n} \left(\sum_{m=1}^{n'} \left\| \frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{k'}} \right\|_\infty^2 \right)^{1/2} \right)^{p} \left(\max_{k' \neq j'} \mathsf{E}\left[\left| \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau) \right|^{2p} \right] \right)^{1/2},$$

and $$F_{3,j,j'} := \left(\max_{k \neq j} \mathsf{E}\left[\left| \frac{\partial X_{k}^{t,\boldsymbol{x}}}{\partial x_{j}}(\tau) \right|^{2p} \right] \right)^{1/2} \left(\max_{k' \neq j'} \mathsf{E}\left[\left| \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau) \right|^{2p} \right] \right)^{1/2}.$$

Note that, using Assumption 3.1,

$$\begin{aligned}
&\sum_{j=1}^{n} \max_{i \neq j} F_{\varsigma 1,i,j,i} + \sum_{j=1}^{n} \sum_{j'=1}^{n} \max_{i \notin \{j,j'\}} F_{\varsigma 1,i,j,j'} \\
&\leq \sum_{j=1}^{n} \max_{i \neq j} \left(\sum_{m=1}^{n'} \left\| \frac{\partial^2 \varsigma_{im}}{\partial x_i \partial x_j} \right\|_\infty^2 \right)^{p/2} + \sum_{j=1}^{n} \sum_{j'=1}^{n} \max_{i \notin \{j,j'\}} \left(\sum_{m=1}^{n'} \left\| \frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{j'}} \right\|_\infty^2 \right)^{p/2} \\
&\leq C_2^p.
\end{aligned}$$

Similarly, using Assumption 3.1, Equations (B.3) and (B.12),

$$\begin{aligned}
&\sum_{j=1}^{n} \max_{i \neq j} F_{\varsigma 2,i,j,i} + \sum_{j=1}^{n} \sum_{j'=1}^{n} \max_{i \notin \{j,j'\}} F_{\varsigma 2,i,j,j'} \\
&= \left(\sum_{j=1}^{n} \max_{i \neq j} \sum_{k'=1}^{n} \left(\sum_{m=1}^{n'} \left\| \frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{k'}} \right\|_\infty^2 \right)^{1/2} \right)^{p} \times \left(\max_{i,k'} \left(\mathsf{E}\left[\left| \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{i}}(\tau) \right|^{2p} \right] \right)^{1/2} \right) \\
&\quad + \left(\sum_{j=1}^{n} \max_{i \neq j} \sum_{k'=1}^{n} \left(\sum_{m=1}^{n'} \left\| \frac{\partial^2 \varsigma_{im}}{\partial x_j \partial x_{k'}} \right\|_\infty^2 \right)^{1/2} \right)^{p} \times \left(\sum_{j'=1}^{n} \max_{k' \neq j'} \left(\mathsf{E}\left[\left| \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau) \right|^{2p} \right] \right)^{1/2} \right) \\
&\leq 2 C_2^p K_{1,2p}^p.
\end{aligned}$$

Additionally, using Equation (B.12),

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

Similarly, we bound $\mathsf{E}\left[f_3^p\right]$ to arrive at

$$\mathsf{E}\left[f_3^p\right] \leq D_2\Big(F_{\nu_1,i,j,j'} + F_{\nu_2,i,j,j'} + F_{\nu_2,i,j',j} + F_{3,j,j'}\Big),$$

where

$$F_{\nu_1,i,j,j'} := \left\| \frac{\partial^2 \nu_i}{\partial x_j \partial x_{j'}} \right\|_\infty^p$$

and

$$F_{\nu_2,i,j,j'} := \left(\sum_{k'=1}^n \left\| \frac{\partial^2 \nu_i}{\partial x_j \partial x_{k'}} \right\|_\infty \right)^p \left(\max_{k' \neq j'} \mathsf{E}\left[\left| \frac{\partial X_{k'}^{t,\boldsymbol{x}}}{\partial x_{j'}}(\tau) \right|^{2p} \right] \right)^{1/2}.$$

Therefore we can find $b_{i,j,j'}$ such that

$$\int_t^s \mathsf{E}\left[f_3^p\right] + (p-1)\mathsf{E}\left[f_4^{p/2}\right] \mathrm{d}\tau \leq b_{i,j,j'}^p, \tag{B.17}$$

which satisfy

$$\sum_{j=1}^n \left(\max_{i \neq j} b_{i,j,i} \right) + \sum_{j=1}^n \sum_{j'=1}^n \left(\max_{i \notin \{j,j'\}} b_{i,j,j'} \right) \leq D_2. \tag{B.18}$$

We use Equations (B.15) and (B.17) in Equation (B.14) and the result in Equation (B.13) and simplify to arrive at

$$\begin{aligned}
\mathsf{E}\left[\left| \frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(s) \right|^p \right] \leq{}& D_2 \int_t^s \mathsf{E}\left[\left| \frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau) \right|^p \right] \mathrm{d}\tau \\
&+ D_2 \int_t^s \max_{k \in \{1,2,\ldots,n\} \setminus \{j,j'\}} \mathsf{E}\left[\left| \frac{\partial^2 X_k^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau) \right|^p \right] \mathrm{d}\tau \\
&+ D_2\, d_{i,j}^p \int_t^s \mathsf{E}\left[\left| \frac{\partial^2 X_j^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau) \right|^p \right] \mathrm{d}\tau \\
&+ D_2 d_{i,j'}^p \int_t^s \mathsf{E}\left[\left| \frac{\partial^2 X_{j'}^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau) \right|^p \right] \mathrm{d}\tau \\
&+ b_{i,j,j'}.
\end{aligned} \tag{B.19}$$

Recall that

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

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 Equation (B.19) and using Grönwall's inequality yield

$$\max_i \mathsf{E}\left[\left(\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(s)\right)^p\right] \leq D_2, \tag{B.20}$$

for all $j, j' \in \{1, \dots, n\}$. Setting $j' = i$ and taking the maximum over $i \neq j$ in Equation (B.19) yield

$$\begin{aligned}\max_{i\neq j} \mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_i}(s)\right|^p\right] \leq \quad & D_2 \int_t^s \max_{i\neq j} \mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_i}(\tau)\right|^p\right] \mathrm{d}\tau \\ & + D_2 \left(\max_{i\neq j} a_{i,j}^p\right) \int_t^s \max_i \mathsf{E}\left[\left|\frac{\partial^2 X_{j'}^{t,\boldsymbol{x}}}{\partial x_{j'} \partial x_i}(\tau)\right|^p\right] \mathrm{d}\tau \\ & + D_2 \int_t^s \max_{i\neq j} \mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_i}(\tau)\right|^p\right] \mathrm{d}\tau \\ & + \max_{i\neq j} b_{i,j,i}^p.\end{aligned}$$

Using Equation (B.20) to bound the second term, followed by Grönwall's inequality and then taking the p'th root and summing over j yields

$$\sum_{j=1}^n \max_{i\neq j} \mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_i \partial x_j}(s)\right|^p\right]^{1/p} \leq \quad D_2 \left(\sum_{j=1}^n \left(\max_{i\neq j} a_{i,j}\right) + \sum_{j=1}^n \left(\max_{i\neq j} b_{i,j,i}\right)\right) \leq D_2, \tag{B.21}$$

where we used Equation (B.18). Finally, taking the maximum over $i \notin \{j, j'\}$ in Equation (B.19),

$$\begin{aligned}\max_{i\notin\{j,j'\}} \mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(s)\right|^p\right] \leq D_2 & \int_t^s \max_{i\notin\{j,j'\}} \mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right] \mathrm{d}\tau \\ & + D_2 \left(\max_{i\neq j} a_{i,j}^p\right) \int_t^s \mathsf{E}\left[\left|\frac{\partial^2 X_j^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right] \mathrm{d}\tau \\ & + D_2 \left(\max_{i\neq j'} a_{i,j'}^p\right) \int_t^s \mathsf{E}\left[\left|\frac{\partial^2 X_{j'}^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right] \mathrm{d}\tau \\ & + \max_{i\notin\{j,j'\}} b_{i,j,j'}^p.\end{aligned}$$

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

$$\begin{aligned}
&\sum_{j=1}^{n}\sum_{j'=1}^{n}\max_{i\notin\{j,j'\}} \mathsf{E}\left[\left|\frac{\partial^2 X_i^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(s)\right|^p\right]^{1/p}\\
&\le D_2\left(\sum_{j=1}^{n}\max_{i\neq j} a_{i,j}\right)\left(\max_{j}\sum_{j'=1}^{n}\sup_{t\le\tau\le s}\mathsf{E}\left[\left|\frac{\partial^2 X_j^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right]^{1/p}\right)\\
&\quad + D_2\left(\sum_{j'=1}^{n}\max_{i\neq j'} a_{i,j'}\right)\left(\max_{j'}\sum_{j=1}^{n}\sup_{t\le\tau\le s}\mathsf{E}\left[\left|\frac{\partial^2 X_{j'}^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right]^{1/p}\right)\\
&\quad + D_2\sum_{j=1}^{n}\sum_{j'=1}^{n}\left(\max_{i\notin\{j,j'\}} b_{i,j,j'}\right).
\end{aligned}$$

The result follows by Equation (B.18), and since $\sum_{j=1}^{n}\max_{i\neq j} a_{i,j}\le C_1$ and, by Equation (B.3), Equation (B.20) and Equation (B.21),

$$\begin{aligned}
&\max_{j}\sum_{j'=1}^{n}\sup_{t\le\tau\le s}\mathsf{E}\left[\left|\frac{\partial^2 X_j^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right]\\
&\le \max_{j}\sup_{t\le\tau\le s}\mathsf{E}\left[\left|\frac{\partial^2 X_j^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j}}(\tau)\right|^p\right] + \sum_{j'=1}^{n}\max_{j\neq j'}\sup_{t\le\tau\le s}\mathsf{E}\left[\left|\frac{\partial^2 X_j^{t,\boldsymbol{x}}}{\partial x_j \partial x_{j'}}(\tau)\right|^p\right]\\
&\le D_2.
\end{aligned}$$

B.3 *Third variation*

The result for the third variations can be proven similarly and is omitted here.