J. Appl. Probab. 1–25 (2026)
*doi:*10.1017/jpr.2026.10091

MEAN-FIELD STOCHASTIC VOLTERRA EQUATIONS

DAVID J. PRÖMEL,* ** AND
DAVID SCHEFFELS,* *** *University of Mannheim*

Abstract

Well-posedness is established for multi-dimensional mean-field stochastic Volterra equations with Lipschitz-continuous coefficients, allowing for singular kernels as well as for one-dimensional mean-field stochastic Volterra equations with Hölder-continuous diffusion coefficients and sufficiently regular kernels. In these different settings, quantitative, pointwise propagation of chaos results are derived for the associated Volterra-type interacting particle systems.

Keywords: Mean-field SDE; McKean–Vlasov process; pathwise uniqueness; propagation of chaos; stochastic Volterra equation; strong solution; Yamada–Watanabe theorem

2020 Mathematics Subject Classification: Primary 60H20
Secondary 60K35; 45D05

1. Introduction

Mean-field stochastic differential equations (mean-field SDEs), also known as McKean–Vlasov stochastic differential equations, provide mathematical descriptions of random systems of interacting particles whose time evolutions depend, in some manner, on the probability distribution of the entire systems. A crucial reason for the frequent use of mean-field SDEs in applied mathematics is the fact that they allow the modeling of the 'propagation of chaos' within large interacting particle systems. Recall that, on a microscopic scale, the trajectory of each individual particle can often be appropriately modeled by a stochastic process. However, when the number of particles becomes very large, the microscopic scale usually contains too much information, making the interactions of individual particles intractable. Fortunately, sending the number of particles to infinity, the propagation of chaos states that the behavior of an individual particle depends only on the probability distribution of the entire system, i.e. on the macroscopic scale the interaction of individual particles becomes negligible.

Mean-field SDEs, as well as the propagation of chaos, originated in statistical physics and were first studied by Kac [20], McKean [23] and Vlasov [34]. Since then, these concepts have found a wide range of applications in a variety of fields such as physics, finance, and data science. We refer, e.g., to [12–15, 19, 33] for comprehensive introductions to mean-field SDEs and their numerous applications. Except for a very small number of publications, like the rough path-based approaches to mean-field SDEs [5, 6, 16], the vast majority of literature on

Received 4 February 2024; accepted 28 February 2026.
* Postal address: University of Mannheim, Institute of Mathematics, B 6, 26, 68159 Mannheim, Germany.
** Email address: proemel@uni-mannheim.de
*** Email address: dscheffe@mail.uni-mannheim.de

mean-field SDEs and the propagation of chaos is restricted to Markovian systems of interacting particles, i.e. the behavior of each particle has to be independent of all past states of the system. On the contrary, it is often observed that many real-world dynamical systems do have memory effects, and thus do indeed depend on past states of the underlying systems. Well-known examples of such systems are the growth of populations, the spread of epidemics, and turbulence flows.

Classical mathematical models for random dynamical systems with memory effects are given by stochastic Volterra equations (SVEs), as introduced in the seminal works of Berger and Mizel [7, 8]; see also, e.g., [26, 27]. While SVEs allow for generating non-Markovian stochastic processes, the solutions of SVEs, in contrast to mean-field SDEs, do not depend directly on the probability distributions of the generated random systems.

In the present paper we aim to unify the theories of mean-field stochastic differential equations and stochastic Volterra equations, enabling us to combine the desirable modeling advantages of both classes of equations. More precisely, we introduce mean-field stochastic Volterra equations (mean-field SVEs),

$$X_t = X_0 + \int_0^t K_\mu(s,t)\mu(s, X_s, \mathcal{L}(X_s))\,\mathrm{d}s + \int_0^t K_\sigma(s,t)\sigma(s, X_s, \mathcal{L}(X_s))\,\mathrm{d}B_s, \qquad t \in [0,T], \quad (1)$$

where X_0 is a random variable, B is a Brownian motion, and the coefficients μ, σ, as well as the kernels K_μ, K_σ, are measurable functions. Here, $\mathcal{L}(X_s)$ denotes the law of the random variable X_s. In words, mean-field SVEs are a class of stochastic integral equations that describe the dynamics of random systems with both nonlinear interactions and memory effects. They constitute a generalization of mean-field SDEs and of classical SVEs. Notice that a solution to the mean-field SVE (1) is, in general, neither a Markov process nor a semimartingale.

Our first contribution is to establish the (strong) well-posedness of the mean-field SVE (1), meaning that there exists a unique strong solution to (1), under two sets of assumptions. On the one hand, we show the existence of a unique solution to the mean-field SVE (1) in a multi-dimensional setting with standard assumptions on the kernels and coefficients, i.e. we assume some integrability of the kernels as well as Lipschitz continuity and a linear growth condition for the coefficients, cf., e.g., [11, 35]. The proof is based on a classical fixed-point argument in combination with techniques from the theories of mean-field SDEs and SVEs. On the other hand, we show the existence of a unique solution to the mean-field SVE (1) in a one-dimensional setting, assuming sufficiently smooth kernels and Hölder-continuous diffusion coefficients that are independent of the law of the solution. To that end, we rely on a Yamada–Watanabe approach [36] to SVEs with sufficiently smooth kernels, as recently generalized in [3, 30]. As comparison, for well-posedness results in the case of mean-field SDEs we refer to [4, 18, 21], and in the case of SVEs to [3, 30, 35]. Furthermore, we remark that a specific type of mean-field SVEs was studied in [31], where the coefficients may depend on the law of the solution but only through an expectation operator.

Our second contribution is to establish quantitative, pointwise propagation of chaos results of Volterra-type systems of interacting particles. In words, sending the number of Volterra-type interacting particles to infinity, we obtain a macroscopic description of the systems based on a mean-field stochastic Volterra equation. The approach developed is based on a synchronous coupling method; it was initiated in [24] and extended in [33]. In the case of mean-field SDEs, synchronous coupling methods are widely used for systems that are described by systems of McKean–Vlasov diffusions, and often lead to pathwise propagation of chaos; see, e.g., [14, Theorem 3.20], [11, Theorem 1.10], and [18]. In the present case of mean-field SVEs,

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

implementing a synchronous coupling method becomes more challenging as the underlying McKean–Vlasov processes are of Volterra type and, thus, in general, lack the semimartingale and Markov property. As for our well-posedness theory of mean-field SVEs, we distinguish between the aforementioned multi- and one-dimensional settings. The pointwise nature of the our propagation of chaos results for mean-field SVEs is caused by the non-availability of a Burkholder–Davis–Gundy inequality in the multi-dimensional setting and by the Hölder continuity of the diffusion coefficients in the one-dimensional setting. The latter setting requires us to combine the synchronous coupling method with a Yamada–Watanabe approach.

1.1. Organization of the paper

In Section 2 we present the main results regarding the well-posedness and propagation of chaos for mean-field stochastic Volterra equations. Section 3 provides some necessary well-posedness results for ordinary stochastic Volterra equations. The proofs of the main results are contained in Sections 4, 5, and 6.

2. Main results: Well-posedness and propagation of chaos

Let $T \in (0, \infty)$, $d, m \in \mathbb{N}$, and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space that satisfies the usual conditions. Suppose $B = (B_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion with respect to $(\mathcal{F}_t)_{t\in[0,T]}$. The law of a random variable X is denoted by $\mathcal{L}(X)$ and, for $p \geq 1$, the space of probability measures on $\mathbb{R}^d$ with finite pth moments by $\mathcal{P}_p(\mathbb{R}^d)$. Let $C([0, T]; \mathbb{R}^d)$ be the space of continuous functions from $[0,T]$ to $\mathbb{R}^d$, equipped with the supremum norm $\|\cdot\|_\infty$, and $\mathcal{P}_p(C([0, T]; \mathbb{R}^d))$ be the space of probability measures on $C([0, T];\mathbb{R}^d)$ with finite pth moment. For $\rho, \tilde{\rho} \in \mathcal{P}_p(\mathbb{R}^d)$, we write $W_p(\rho, \tilde{\rho})$ for the p-Wasserstein distance between ρ and $\tilde{\rho}$ (see [12, Chapter 5] for its definition) and, with a slight abuse of notation, for $\rho, \tilde{\rho} \in \mathcal{P}_p(C([0, T]; \mathbb{R}^d))$ we define the p-Wasserstein distance by

$$W_p(\rho, \tilde{\rho}) = \inf_{\pi\in\Pi(\rho,\tilde{\rho})} \left[\int_{C([0,T];\, \mathbb{R}^d)^2} \|x - y\|_\infty^p \, \mathrm{d}\pi(x, y) \right]^{1/p},$$

where $\Pi(\rho, \tilde{\rho})$ denotes the set of all probability measures on $C([0, T]; \mathbb{R}^d)^2$ with marginal distributions given by ρ and $\tilde{\rho}$, respectively. The space $\mathbb{R}^d$ is always equipped with the Euclidean norm $|\cdot|$, and on the space $\mathbb{R}^{d\times m}$ we use the Frobenius norm, also denoted by $|\cdot|$. Moreover, we set $\Delta_T := \{(s, t) \in [0, T] \times [0, T] \colon 0 \leq s \leq t \leq T\}$ and use the notation $A_\eta \lesssim B_\eta$ for a generic parameter η, meaning that $A_\eta \leq CB_\eta$ for some constant $C > 0$ independent of η.

We consider the d-dimensional mean-field stochastic Volterra equation

$$X_t = X_0 + \int_0^t K_\mu(s, t)\mu(s, X_s, \mathcal{L}(X_s))\, \mathrm{d}s + \int_0^t K_\sigma(s, t)\sigma(s, X_s, \mathcal{L}(X_s))\, \mathrm{d}B_s, \qquad t \in [0, T], \quad (2)$$

where X_0 is a d-dimensional, $\mathcal{F}_0$-measurable random variable that is independent of B, and the coefficients $\mu\colon [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}^d$, $\sigma\colon [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}^{d\times m}$ and the kernels $K_\mu, K_\sigma\colon \Delta_T \to \mathbb{R}$ are measurable functions. The integral $\int_0^t K_\sigma(s, t)\sigma(s, X_s, \mathcal{L}(X_s))\, \mathrm{d}B_s$ is defined as a stochastic Itô integral.

Let us briefly recall the concepts of well-posedness, strong solutions, and pathwise uniqueness. We use, for measure spaces $\mathcal{X}, \mathcal{Y}$ and $p \geq 1$, the notation $L^p = L^p(\mathcal{X}; \mathcal{Y})$ for the space of all $\mathcal{Y}$-valued, measurable, p-integrable functions on $\mathcal{X}$, and, for two Banach spaces $\mathcal{X}, \mathcal{Y}$, $C(\mathcal{X}; \mathcal{Y})$ for the space of all $\mathcal{Y}$-valued, continuous functions on $\mathcal{X}$. An $(\mathcal{F}_t)_{t\in[0,T]}$-progressively

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

measurable stochastic process $(X_t)_{t\in[0,T]}$ in $L^p(\Omega\times[0,T];\mathbb{R}^d)$, on the given probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,T]},\mathbb{P})$, is called a (strong) L^p-solution of the mean-field SVE (2) if

$$\int_0^t (|K_\mu(s,t)\mu(s,X_s,\mathcal{L}(X_s))| + |K_\sigma(s,t)\sigma(s,X_s,\mathcal{L}(X_s))|^2)\,\mathrm{d}s < \infty \quad \text{for all } t\in[0,T],$$

and the integral equation (2) holds $\mathbb{P}$-almost surely. We say that *pathwise uniqueness in L^p* holds for the mean-field SVE (2) if $\mathbb{P}(X_t=\tilde{X}_t \text{ for all } t\in[0,T])=1$ for any two L^p-solutions $(X_t)_{t\in[0,T]}$ and $(\tilde{X}_t)_{t\in[0,T]}$ of (2) defined on the same probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,T]},\mathbb{P})$. We say that the mean-field SVE (2) is well-posed in L^p (or that there exists a unique L^p-solution) for $p\geq 1$ if there exists a strong L^p-solution to (2) and pathwise uniqueness in L^p holds.

In the following we distinguish between multi-dimensional and one-dimensional settings, since these settings allow us to establish the well-posedness of the mean-field SVE (2) with different regularity assumptions on the kernels and coefficients. The main existence and uniqueness results regarding mean-field SVEs as well as propagation of chaos are stated in Subsections 2.1 and 2.2. In the multi-dimensional setting (Subsection 2.1) we make standard Lipschitz assumptions on the coefficients μ,σ, whereas in the one-dimensional setting (Subsection 2.2) we assume that μ is Lipschitz continuous but allow σ to be only Hölder continuous. We prove the corresponding results in Sections 4, 5, and 6.

2.1. Mean-field SVEs with Lipschitz-continuous coefficients

In this subsection we consider the multi-dimensional stochastic Volterra equation (2) with dimensions $d,m\in\mathbb{N}$ and coefficients μ,σ that are Lipschitz continuous in the space and distributional component, uniformly in the time component, allowing for potentially singular kernels. We start by stating the assumptions on the kernels.

Assumption 1. *Assume there are constants $\gamma\in\left(0,\frac{1}{2}\right]$, $\epsilon>0$, and $L>0$ such that $K_\mu,K_\sigma\colon\Delta_T\to\mathbb{R}$ are measurable functions fulfilling*

$$\int_0^t |K_\mu(s,t')-K_\mu(s,t)|^{1+\epsilon}\,\mathrm{d}s + \int_t^{t'} |K_\mu(s,t')|^{1+\epsilon}\,\mathrm{d}s \leq L|t'-t|^{\gamma(1+\epsilon)},$$
$$\int_0^t |K_\sigma(s,t')-K_\sigma(s,t)|^{2+\epsilon}\,\mathrm{d}s + \int_t^{t'} |K_\sigma(s,t')|^{2+\epsilon}\,\mathrm{d}s \leq L|t'-t|^{\gamma(2+\epsilon)}$$

for all $(t,t')\in\Delta_T$.

Note that Assumption 1 allows for singular kernels, like the fractional convolutional kernel $K(s,t)=(t-s)^{-\alpha}$ for $\alpha\in(0,1/2)$ and the examples provided in [1, Example 1.3]. Moreover, for $\epsilon>0$ given by Assumption 1, let the fixed parameter $\delta>2$ be defined by

$$\delta := \frac{4+2\epsilon}{\epsilon}, \tag{3}$$

such that

$$\frac{2}{2+\epsilon}+\frac{2}{\delta}=1. \tag{4}$$

In the following we use the δ-Wasserstein distance on the space $\mathcal{P}_\delta(\mathbb{R}^d)$ of probability measures on $\mathbb{R}^d$ with finite δth moments. Relying on the δ-Wasserstein distance, we specify the assumptions on the regularity of the coefficients μ and σ, which are a classical linear growth condition and a Lipschitz assumption.

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

Assumption 2. *Let* $\mu\colon [0,T]\times\mathbb{R}^d\times\mathcal{P}_\delta(\mathbb{R}^d)\to\mathbb{R}^d$ *and* $\sigma\colon [0,T]\times\mathbb{R}^d\times\mathcal{P}_\delta(\mathbb{R}^d)\to\mathbb{R}^{d\times m}$ *be measurable functions such that:*

(i) *for any bounded set* $\mathcal{K}\subset\mathcal{P}_\delta(\mathbb{R}^d)$*, there is a constant* $C_{\mathcal{K}}>0$ *such that the linear growth condition* $|\mu(t,x,\rho)|+|\sigma(t,x,\rho)|\le C_{\mathcal{K}}(1+|x|)$ *holds for all* $\rho\in\mathcal{K}$*,* $t\in[0,T]$*, and* $x\in\mathbb{R}^d$*;*

(ii) μ *and* σ *are Lipschitz continuous in* x *and in* ρ *with respect to the* δ*-Wasserstein distance uniformly in t, i.e. there is a constant* $C_{\mu,\sigma}>0$ *such that*

$$|\mu(t,x,\rho)-\mu(t,\tilde{x},\tilde{\rho})|+|\sigma(t,x,\rho)-\sigma(t,\tilde{x},\tilde{\rho})|\le C_{\mu,\sigma}\big(|x-\tilde{x}|+W_\delta(\rho,\tilde{\rho})\big)$$

holds for all $t\in[0,T]$*,* $x,\tilde{x}\in\mathbb{R}^d$*, and* $\rho,\tilde{\rho}\in\mathcal{P}_\delta(\mathbb{R}^d)$*.*

Our first result is the well-posedness of the mean-field stochastic Volterra equation (2).

Theorem 1. *Suppose that the initial value* X_0 *is in* $L^p(\Omega;\mathbb{R}^d)$*, the kernels* K_μ, K_σ *fulfill Assumption 1, the coefficients* μ,σ *fulfill Assumption 2, and* $p>\max\{1/\gamma, 1+2/\epsilon\}$*, where* $\gamma\in\big(0,\frac{1}{2}\big]$ *and* $\epsilon>0$ *are given by Assumption 1. Then, the mean-field stochastic Volterra equation (2) is well posed in* L^p*. Moreover, for any* $q\ge p$*, if* $X_0\in L^q(\Omega;\mathbb{R}^d)$*, the unique* L^p*-solution X of (2) satisfies*

$$\sup_{t\in[0,T]}\mathbb{E}[|X_t|^q]<\infty. \tag{5}$$

Our second result is the propagation of chaos for mean-field stochastic Volterra equations, i.e. we show that the unique L^p-solution to the mean-field stochastic Volterra equation (2) is the limit $N\to\infty$ of the solutions to the following system of N mean-field stochastic Volterra equations:

$$X_t^{N,i}=X_0^i+\int_0^t K_\mu(s,t)\mu(s,X_s^{N,i},\bar{\rho}_s^N)\,\mathrm{d}s+\int_0^t K_\sigma(s,t)\sigma(s,X_s^{N,i},\bar{\rho}_s^N)\,\mathrm{d}B_s^i,\qquad t\in[0,T], \tag{6}$$

for $i\in\{1,\dots,N\}$, where $\bar{\rho}_t^N:=(1/N)\sum_{i=1}^N\delta_{X_t^{N,i}}$ is the empirical distribution of $(X_t^{N,i})_{i=1,\dots,N}$, $(X_0^i)_{i\in\mathbb{N}}\subset L^q(\Omega;\mathbb{R}^d)$ is a sequence of $\mathcal{F}_0$-measurable, independent and identically distributed (i.i.d.) random variables for some $q>4$, and $(B^i)_{i\in\mathbb{N}}$ is a sequence of independent m-dimensional Brownian motions, which are all defined on the given probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,T]},\mathbb{P})$. Strong L^p-solutions, pathwise uniqueness in L^p, and well-posedness in L^p for the system (6) of mean-field SVEs are defined analogously to (2), and δ_x denotes the Dirac measure at x for $x\in\mathbb{R}^d$. Moreover, for $i\in\mathbb{N}$, let $\underline{\mathbf{X}}^i$ be the solution of the mean-field SVE (2) with the initial condition X_0^i and driving Brownian motion B^i. In the present multi-dimensional setting, we obtain the following convergence result.

Theorem 2. (Volterra propagation of chaos). *Suppose Assumptions 1 and 2, and that the sequence of initial conditions* $(X_0^i)_{i\in\mathbb{N}}\subset L^q(\Omega;\mathbb{R}^d)$ *for some* $q>\max\{p,2\delta\}$ *and* $p>\max\{1/\gamma,1+2/\epsilon\}$*, where* δ *is defined in (3). Then, the system (6) of mean-field SVEs is well posed in* L^p *for every* $N\ge 1$*, where the unique* L^p*-solution is denoted by* $(X_t^{N,i})_{i=1,\dots,N}$*. Moreover, we have*

$$\lim_{N\to\infty}\left(\max_{1\le i\le N}\left(\sup_{t\in[0,T]}\mathbb{E}\big[|X_t^{N,i}-\underline{X}_t^i|^\delta\big]\right)+\sup_{t\in[0,T]}\mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{i=1}^N\delta_{X_t^{N,i}},\mathcal{L}(\underline{X}_t^1)\right)^\delta\right]\right)=0. \tag{7}$$

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

The rate of convergence in (7) is explicitly stated in the next lemma.

Lemma 1. *With the assumptions and notation of Theorem 2, we have*

$$\max_{1\leq i\leq N}\left(\sup_{t\in[0,T]}\mathbb{E}[|X_t^{N,i}-\underline{X}_t^{i}|^{\delta}]\right)+\sup_{t\in[0,T]}\mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^{N,i}},\mathcal{L}(\underline{X}_t^{1})\right)^{\delta}\right]\lesssim\varepsilon_N, \tag{8}$$

where $(\varepsilon_N)_{N\in\mathbb{N}}$ *is given by*

$$\varepsilon_N=\begin{cases}N^{-1/2} & \text{if } d<2\delta,\\ N^{-1/2}\log_2(1+N) & \text{if } d=2\delta,\\ N^{-\delta/d} & \text{if } d>2\delta.\end{cases} \tag{9}$$

Remark 1. The rates of convergence obtained in (9) are analogous to the classical rates for ordinary mean-field SDEs with Lipschitz coefficients (see [14, Theorem 3.20]), using $W_\delta(\cdots)^\delta$ instead of $W_2(\cdots)^2$ and, consequently, replacing the exponent $2/d$ by δ/d in (9). Note that in the case of ordinary mean-field SDEs we obtain a pathwise propagation of chaos result (meaning that the sup in (8) is inside the expectation operators), which is a stronger type of convergence than the pointwise convergence presented in Theorem 2. This weaker type of convergence is caused by the missing availability of the standard Burkholder–Davis–Gundy inequality for the solutions of stochastic Volterra equations since they are, in general, not semi-martingales. However, the rates of convergence provided in Lemma 1 seem to be optimal for synchronous coupling methods, since it is shown in [17, Theorem 1ff] that for terms of the form $\mathbb{E}[W_\delta(\bar{\rho}_N,\rho)^\delta]$ the rates in (9) are sharp. Consequently, optimality could only be lost in the inequalities (47) or (48), which, at least in general, appears not to be the case.

2.2. Mean-field SVEs with Hölder-continuous diffusion coefficients

In this subsection we consider mean-field SVEs in a one-dimensional setting, i.e. we assume $d=m=1$. This allows us to relax the Lipschitz assumption on the diffusion coefficient σ to Hölder continuity in the space variable, provided that σ is independent of the distribution of the solution and that the kernels are sufficiently regular. More precisely, we consider the one-dimensional mean-field stochastic Volterra equation

$$X_t=X_0+\int_0^t K_\mu(s,t)\mu(s,X_s,\mathcal{L}(X_s))\,\mathrm{d}s+\int_0^t K_\sigma(s,t)\sigma(s,X_s)\,\mathrm{d}B_s,\qquad t\in[0,T], \tag{10}$$

where $(B_t)_{t\in[0,T]}$ is a one-dimensional Brownian motion, X_0 is an $\mathcal{F}_0$-measurable random variable, the coefficients $\mu\colon[0,T]\times\mathbb{R}\times\mathcal{P}_p(\mathbb{R})\to\mathbb{R}$, $\sigma\colon[0,T]\times\mathbb{R}\to\mathbb{R}$ and the kernels $K_\mu,K_\sigma\colon\Delta_T\to\mathbb{R}$ are measurable functions. We consider two different sets of assumptions on the kernels and on the initial condition.

Assumption 3. *Let* $\gamma\in\left(0,\frac{1}{2}\right]$ *and* $\epsilon>0$. *Let* X_0 *be an* $\mathcal{F}_0$*-measurable random variable and* $K_\mu,K_\sigma\colon\Delta_T\to\mathbb{R}$ *be continuous functions such that:*

(i) $K_\mu(s,\cdot)$ *is absolutely continuous for every* $s\in[0,T]$, *and* $\partial_2 K_\mu$ *is bounded on* Δ_T*;*

(ii) $K_\sigma(\cdot,t)$ *is absolutely continuous for every* $t\in[0,T]$, $K_\sigma(s,\cdot)$ *is absolutely continuous for every* $s\in[0,T]$ *with* $\partial_2 K_\sigma\in L^2(\Delta_T)$, *and* $\partial_2 K_\sigma(\cdot,t)$ *is absolutely continuous for*

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

every $t \in [0, T]$. Furthermore, there is a constant $C_1 > 0$ such that $|K_\sigma(t, t)| \geq C_1$ for any $t \in [0, T]$, and there exists $C_2 > 0$ such that

$$\int_0^s |K_\sigma(u, t) - K_\sigma(u, s)|^{2+\epsilon} \, \mathrm{d}u \leq C_2 |t - s|^{\gamma(2+\epsilon)}$$

and $|\partial_1 K_\sigma(s, t)| + |\partial_2 K_\sigma(s, s)| + \int_s^t |\partial_{21} K_\sigma(s, u)| \, \mathrm{d}u \leq C_2$ hold for any $(s, t) \in \Delta_T$;

(iii) $X_0 \in L^p(\Omega; \mathbb{R})$ *for* $p > \max\{1/\gamma, 1 + 2/\epsilon\}$.

Instead of Assumption 3, we can alternatively require K_μ, K_σ, and X_0 to fulfill the following assumption, where the kernels are supposed to be convolutional.

Assumption 4. *Let X_0 be an $\mathcal{F}_0$-measurable random variable and $K_\mu, K_\sigma \colon \Delta_T \to \mathbb{R}$ be continuous functions such that:*

(i) $K_\mu(s, t) = K_\sigma(s, t) = \tilde{K}(t - s)$ *for some* $\tilde{K} \in C^1([0, T]; \mathbb{R})$*;*

(ii) $X_0 \in L^p(\Omega; \mathbb{R})$ *for* $p > 2$.

Next, we formulate the assumptions on the coefficients.

Assumption 5. *Let $\mu \colon [0, T] \times \mathbb{R} \times \mathcal{P}_1(\mathbb{R}) \to \mathbb{R}$ and $\sigma \colon [0, T] \times \mathbb{R} \to \mathbb{R}$ be measurable functions such that:*

(i) *for any bounded set $\mathcal{K} \subset \mathcal{P}_1(\mathbb{R})$, there is a constant $C_{\mathcal{K}} > 0$ such that the linear growth condition $|\mu(t, x, \rho)| + |\sigma(t, x)| \leq C_{\mathcal{K}} \rho(1 + |x|)$ holds for all $\rho \in \mathcal{K}$, $t \in [0, T]$, and $x \in \mathbb{R}$;*

(ii) *μ is Lipschitz continuous in x and ρ with respect to the 1-Wasserstein distance, uniformly in t, i.e. there is a constant $C_\mu > 0$ such that*

$$|\mu(t, x, \rho) - \mu(t, \tilde{x}, \tilde{\rho})| \leq C_\mu \big(|x - \tilde{x}| + W_1(\rho, \tilde{\rho}) \big)$$

holds for all $t \in [0, T]$, $x, \tilde{x} \in \mathbb{R}$, and $\rho, \tilde{\rho} \in \mathcal{P}_1(\mathbb{R})$, and σ is Hölder continuous of order $\frac{1}{2} + \xi$ for some $\xi \in \big[0, \frac{1}{2}\big]$ in x uniformly in t, i.e. there is a constant $C_\sigma > 0$ such that $|\sigma(t, x) - \sigma(t, \tilde{x})| \leq C_\sigma |x - \tilde{x}|^{(1/2)+\xi}$ holds for all $t \in [0, T]$ and $x, \tilde{x} \in \mathbb{R}$.

First, we establish the well-posedness of the mean-field stochastic Volterra equation (10) with Hölder-continuous diffusion coefficients. Its proof is based on a Yamada–Watanabe-type approach [36], which requires essentially a one-dimensional setting and leads to the stronger assumptions on the kernels. Moreover, note that the Hölder-continuous diffusion coefficients are required to be independent of the law of the solution, which is essentially a standard assumption for ordinary mean-field stochastic differential equations as it appears to be a necessary assumption to implement a Yamada–Watanabe type approach, cf. [21].

Theorem 3. *Suppose Assumption 5, and that the kernels K_μ, K_σ and the initial condition X_0 satisfy Assumption 3 or 4 with p given therein. Then the mean-field stochastic Volterra equation (10) is well posed in L^p. Moreover, for any $q \geq p$, if $X_0 \in L^q(\Omega; \mathbb{R}^d)$, the unique solution X of (10) satisfies $\sup_{t \in [0,T]} \mathbb{E}[|X_t|^q] < \infty$.*

Second, we establish the propagation of chaos for one-dimensional stochastic mean-field SVEs with Hölder-continuous diffusion coefficients. To that end, we consider the symmetric

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

system of N mean-field stochastic Volterra equations

$$X_t^{N,i} = X_0^i + \int_0^t K_\mu(s,t)\mu(s, X_s^{N,i}, \bar{\rho}_s^N)\,\mathrm{d}s + \int_0^t K_\sigma(s,t)\sigma(s, X_s^{N,i})\,\mathrm{d}B_s^i, \qquad t \in [0,T], \tag{11}$$

for $i \in \{1, \ldots, N\}$, where $(X_0^i)_{i\in\mathbb{N}} \subset L^p(\Omega;\mathbb{R})$ is an i.i.d. sequence of initial conditions, and $(B^i)_{i\in\mathbb{N}}$ is a sequence of independent one-dimensional Brownian motions. Moreover, for $i \in \mathbb{N}$, $\underline{X}^i$ denotes the solution of the mean-field SVE (10) with initial condition X_0^i and driving Brownian motion B^i. In the present one-dimensional setting, we obtain the following convergence result.

Theorem 4. (Volterra propagation of chaos) *Suppose Assumption 5, and that the kernels K_μ, K_σ and the initial conditions X_0^i for $i \in \mathbb{N}$ satisfy Assumption 3 or 4 with p given therein. Then, the system (11) of mean-field SVEs is well posed in L^p, where the unique L^p-solution is denoted by $(X_t^{N,i})_{i=1,\ldots,N}$ for every $N \geq 1$. Moreover,*

$$\lim_{N\to\infty}\left(\max_{1\le i\le N}\left(\sup_{t\in[0,T]}\mathbb{E}[|X_t^{N,i} - \underline{X}_t^i|]\right) + \sup_{t\in[0,T]}\mathbb{E}\left[W_1\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_t^{N,i}}, \mathcal{L}(\underline{X}_t^1)\right)\right]\right) = 0. \tag{12}$$

The rate of convergence in (12) is explicitly stated in the next lemma.

Lemma 2. *Supposing the assumptions and notation of Theorem 4, we have*

$$\max_{1\le i\le N}\left(\sup_{t\in[0,T]}\mathbb{E}[|X_t^{N,i} - \underline{X}_t^i|]\right) + \sup_{t\in[0,T]}\mathbb{E}\left[W_1\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_t^{N,i}}, \mathcal{L}(\underline{X}_t^1)\right)\right] \lesssim N^{-1/2}. \tag{13}$$

Remark 2. The rate of convergence in (13) is expected to be optimal for synchronous coupling methods, cf. Remark 1, since it is shown in [17, Theorem 1ff] that for terms of the form $\mathbb{E}[W_1(\bar{\rho}_N, \rho)]$ the rate is sharp. Consequently, optimality could only be lost in the inequalities (36) or (46).

3. On the well-posedness of ordinary stochastic Volterra equations

In this section we provide various well-posedness results for ordinary stochastic Volterra equations with random initial conditions that are needed to prove the well-posedness results for mean-field stochastic Volterra equations presented in Section 2. We start with SVEs with Lipschitz-continuous coefficients, which is a slight modification of [35, Theorem 1.1].

Lemma 3. *Let the kernels K_μ, K_σ fulfill Assumption 1, $p > \max\{1/\gamma, 1 + 2/\epsilon\}$ with $\gamma \in \left(0, \frac{1}{2}\right]$ and $\epsilon > 0$ from Assumption 1, the initial value $X_0 \in L^p(\Omega;\mathbb{R}^d)$ be adapted, and the measurable coefficients $\mu\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon [0,T]\times\mathbb{R}^d \to \mathbb{R}^{d\times m}$ for some $d, m \in \mathbb{N}$ fulfill the linear growth condition $|\mu(t,x)| + |\sigma(t,x)| \le C_{\mu,\sigma}(1+|x|)$ for some $C_{\mu,\sigma} > 0$ and all $t \in [0,T]$, $x \in \mathbb{R}^d$, and the Lipschitz condition*

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| \le C_{\mu,\sigma}|x-y|$$

for some $C_{\mu,\sigma} > 0$ and all $t \in [0,T]$, $x, y \in \mathbb{R}^d$. Then, the d-dimensional stochastic Volterra equation

$$X_t = X_0 + \int_0^t K_\mu(s,t)\mu(s,X_s)\,\mathrm{d}s + \int_0^t K_\sigma(s,t)\sigma(s,X_s)\,\mathrm{d}B_s, \qquad t \in [0,T],$$

is well posed in L^p, where $(B_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion.

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

Proof. With the assumed integrability on X_0, it is straightforward to adapt the Picard iteration and the Grönwall type estimates in proof of [35, Theorem 1.1] to allow for random initial conditions X_0, as stated in Lemma 3. □

For one-dimensional ordinary stochastic Volterra equations the Lipschitz assumption on the diffusion coefficients can be relaxed to Hölder continuity, provided the kernels are sufficiently regular or have a convolutional structure. The next results are a slight modification of [30, Theorem 2.3], allowing for SVEs with random initial conditions.

Lemma 4. *Let the kernels K_μ, K_σ fulfill Assumption 3, $p > \max\{1/\gamma, 1 + 2/\epsilon\}$ with $\gamma \in \left(0, \frac{1}{2}\right]$ and $\epsilon > 0$ from Assumption 3, the initial value $X_0 \in L^p(\Omega; \mathbb{R})$, and the measurable coefficients $\mu\colon [0, T] \times \mathbb{R} \to \mathbb{R}$ and $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ fulfill the linear growth condition*

$$|\mu(t, x)| + |\sigma(t, x)| \le C_{\mu,\sigma}(1 + |x|)$$

for some $C_{\mu,\sigma} > 0$ and all $t \in [0, T]$, $x \in \mathbb{R}$, μ be the Lipschitz condition

$$|\mu(t, x) - \mu(t, y)| \le C_\mu |x - y|$$

for some $C_\mu > 0$ and all $t \in [0, T]$, $x, y \in \mathbb{R}$, and σ be the Hölder condition

$$|\sigma(t, x) - \sigma(t, y)| \le C_\sigma |x - y|^{(1/2)+\xi}$$

for $\xi \in \left[0, \frac{1}{2}\right]$, some $C_\sigma > 0$, and all $t \in [0, T]$, $x, y \in \mathbb{R}$. Then the stochastic Volterra equation

$$X_t = X_0 + \int_0^t K_\mu(s, t)\mu(s, X_s)\,\mathrm{d}s + \int_0^t K_\sigma(s, t)\sigma(s, X_s)\,\mathrm{d}B_s, \quad t \in [0, T],$$

is well posed in L^p, where $(B_t)_{t\in[0,T]}$ is a one-dimensional Brownian motion.

Proof. With the assumed integrability on X_0, it is straightforward to adapt the proof of [30, Theorem 2.3] to the case that X_0 is a random variable. □

The next lemma is a slight generalization of [2, Proposition B.3], providing the well-posedness of one-dimensional SVEs with convolutional kernels and random initial conditions.

Lemma 5. *Suppose that $X_0 \in L^p(\Omega; \mathbb{R})$ for some $p > 2$, the kernels are of the form $K_\mu(s, t) = K_\sigma(s, t) = \tilde{K}(t - s)$ for some $\tilde{K} \in C^1([0, T]; \mathbb{R})$, and the measurable coefficients $\mu\colon [0, T] \times \mathbb{R} \to \mathbb{R}$ and $\sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$ fulfill the linear growth condition*

$$|\mu(t, x)| + |\sigma(t, x)| \le C_{\mu,\sigma}(1 + |x|)$$

for some $C_{\mu,\sigma} > 0$ and all $t \in [0, T]$, $x \in \mathbb{R}$, μ satisfies the Lipschitz condition

$$|\mu(t, x) - \mu(t, y)| \le C_\mu |x - y|$$

for some $C_\mu > 0$ and all $t \in [0, T]$, $x, y \in \mathbb{R}$, and σ satisfies the Hölder condition

$$|\sigma(t, x) - \sigma(t, y)| \le C_\sigma |x - y|^{(1/2)+\xi}$$

for $\xi \in \left[0, \frac{1}{2}\right]$, some $C_\sigma > 0$, and all $t \in [0, T]$, $x, y \in \mathbb{R}$. Then the stochastic Volterra equation

$$X_t = X_0 + \int_0^t \tilde{K}(t - s)\mu(s, X_s)\,\mathrm{d}s + \int_0^t \tilde{K}(t - s)\sigma(s, X_s)\,\mathrm{d}B_s, \quad t \in [0, T], \tag{14}$$

is well posed in L^p, where $(B_t)_{t\in[0,T]}$ is a one-dimensional Brownian motion.

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

Proof. The weak existence of some L^p-solution to the SVE (14) follows from [29, Theorem 3.3] with the straightforward adaptation to random initial conditions X_0. For the pathwise uniqueness, we can adapt the proof from [2, Proposition B.3] using the Lipschitz and Hölder continuity of μ, σ uniformly in t. □

Moreover, for the well-posedness results of mean-field SVEs we need a multi-dimensional well-posedness result for stochastic Volterra equations where the Hölder-continuous coefficient σ is a diagonal matrix, where each entry only depends on the component of the solution of the respective dimension, as provided in the next remark.

Remark 3. For $N \in \mathbb{N}$ let us consider the N-dimensional stochastic Volterra equation

$$X_t = X_0 + \int_0^t K_\mu(s, t)\mu(s, X_s)\,\mathrm{d}s + \int_0^t K_\sigma(s, t)\sigma(s, X_s)\,\mathrm{d}B_s, \qquad t \in [0, T], \tag{15}$$

where $(B_t)_{t\in[0,T]}$ is an N-dimensional Brownian motion, and

$$X_t = \begin{pmatrix} X_t^1 \\ \vdots \\ X_t^N \end{pmatrix}, \qquad X_0 = \begin{pmatrix} X_0^1 \\ \vdots \\ X_0^N \end{pmatrix},$$

$$\mu(s, X_s) = \begin{pmatrix} \mu_1(s, X_s) \\ \vdots \\ \mu_N(s, X_s) \end{pmatrix}, \qquad \sigma(s, X_s) = \begin{pmatrix} \sigma_1(s, X_s^1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_N(s, X_s^N) \end{pmatrix}.$$

Suppose that the kernels K_μ, K_σ and the initial value X_0 fulfill Assumption 3 or 4 with p as defined there, that $\mu\colon [0, T] \times \mathbb{R}^N \to \mathbb{R}^N$ is Lipschitz continuous in the space variable, uniformly in the time variable, and each $\sigma_i\colon [0, T] \times \mathbb{R} \to \mathbb{R}$ for $i \in \{1, \ldots, N\}$ is $(1/2 + \xi)$-Hölder continuous in the space variable, uniformly in the time variable for some $\xi \in \left[0, \frac{1}{2}\right]$. By considering each dimension separately, e.g. as done for SDEs in [36, Theorem 1], it is straightforward to conclude the well-posedness in L^p of the SVE (15) from the corresponding one-dimensional results in Lemmas 4 and 5.

We conclude this section with a remark on the path regularity of solutions and one on the notion of L^p-well-posedness.

Remark 4. (Path regularity). Let X be the unique (d-, 1-, or N-dimensional) solution to the stochastic Volterra equation in any of the settings in Lemmas 3, 4, 5 or Remark 3 with $p > \max\{1/\gamma, 1 + 2/\epsilon\}$. In the case of Assumption 4, we can set $\gamma = \frac{1}{2}$ and $p > 2$ as given there. Assuming $X_0 \in L^q$ for $q \geq p$, by adapting [30, Lemmas 3.1 and 3.4] to the multi-dimensional setting, it follows that $\sup_{t\in[0,T]} \mathbb{E}[|X_t|^q] < \infty$ and $\mathbb{E}[|X_t - X_s|^q] \lesssim |t - s|^{\beta q}$ for any $q \geq 1$, $\beta \in (0, \gamma - 1/p)$, $s, t \in [0, T]$, and, hence, that the solution X has a modification with β-Hölder-continuous sample paths.

Remark 5. The notion of L^p-well-posedness, as used in Lemmas 3, 4, 5 and Remark 3, appears to be necessary to prove the existence of a strong solution and pathwise uniqueness. First, we need to assume that a solution X is in $L^p(\Omega \times [0, T]; \mathbb{R}^d)$ to conclude continuity of its sample paths with standard estimates, as in [30, Lemma 3.1]. Second, in order to be able to apply Grönwall's lemma to an inequality of the form $\mathbb{E}[|X_t - Y_t|^p] \lesssim \int_0^t \mathbb{E}[|X_s - Y_s|^p]\,\mathrm{d}s$, we need to assume that both solutions X,Y are in $L^p(\Omega \times [0, T]; \mathbb{R}^d)$ to guarantee finiteness of the expectations $\sup_{s\in[0,t]} \mathbb{E}[|X_s|^p]$ and $\sup_{s\in[0,t]} \mathbb{E}[|Y_s|^p]$ by standard estimates, as in [30, Lemma 3.4].

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

4. Well-posedness: Proofs of Theorems 1 and 3

Proof of Theorem 1. We define the solution map Φ by

$$\Phi\colon C\big([0,T];\mathcal{P}_\delta(\mathbb{R}^d)\big)\to C\big([0,T];\mathcal{P}_\delta(\mathbb{R}^d)\big),\qquad \rho\mapsto\Phi(\rho):=\big(\mathcal{L}(X_t^\rho)\big)_{t\in[0,T]},\tag{16}$$

where X^ρ is the unique L^p-solution to the stochastic Volterra equation

$$X_t=X_0+\int_0^t K_\mu(s,t)\mu(s,X_s,\rho_s)\,\mathrm{d}s+\int_0^t K_\sigma(s,t)\sigma(s,X_s,\rho_s)\,\mathrm{d}B_s,\qquad t\in[0,T].\tag{17}$$

Note that a unique fixed point of the solution map Φ implies the existence of a unique L^p-solution $X=(X_t)_{t\in[0,T]}$ to the mean-field SVE (2) satisfying $\sup_{t\in[0,T]}\mathbb{E}[|X_t|^q]<\infty$ for every $q\geq 1$; cf. Step 1 below. Hence, it is sufficient to prove that the solution map Φ has a unique fixed point.

Step 1. We show the well-definedness of the solution map Φ.

For a fixed $\rho=(\rho_t)_{t\in[0,T]}\in C([0,T];\mathcal{P}_\delta(\mathbb{R}^d))$, the integral equation (17) is an ordinary stochastic Volterra equation. Due to Assumption 2, the linear growth and Lipschitz condition of Lemma 3 are satisfied. Hence, there exists a unique strong L^p-solution $X^\rho=(X_t^\rho)_{t\in[0,T]}$ to the SVE (17) and, by Remark 4, $\sup_{t\in[0,T]}\mathbb{E}[|X_t^\rho|^q]<\infty$ for $q\geq q$, provided $X_0\in L^q$, and the sample paths of X^ρ are almost surely continuous. Moreover, note that $(\mathcal{L}(X_t^\rho))_{t\in[0,T]}\in C([0,T];\mathcal{P}_\delta(\mathbb{R}^d))$, since, by the representation of the Wasserstein distance in terms of random variables (see [12, (5.14)]) and by Remark 4, we have

$$W_\delta\big(\mathcal{L}(X_t^\rho),\mathcal{L}(X_s^\rho)\big)\leq\mathbb{E}\big[|X_t^\rho-X_s^\rho|^\delta\big]^{1/\delta}\lesssim|t-s|^\beta,\qquad s,t\in[0,T],$$

for any $\beta\in(0,\gamma-1/p)$ with $\gamma\in\big(0,1/2\big]$, where the parameters are given in Assumption 1.

Step 2: For $\rho,\tilde\rho\in C\big([0,T];\mathcal{P}_\delta(\mathbb{R}^d)\big)$, we show that

$$\sup_{s\in[0,t]}W_\delta(\Phi(\rho)_s,\Phi(\tilde\rho)_s)^\delta\lesssim\int_0^t W_\delta(\rho_s,\tilde\rho_s)^\delta\,\mathrm{d}s,\qquad t\in[0,T].\tag{18}$$

We have

$$\begin{aligned}
&\mathbb{E}\big[|X_t^\rho-X_t^{\tilde\rho}|^\delta\big]\\
&\quad\lesssim\mathbb{E}\bigg[\bigg|\int_0^t K_\mu(s,t)\big(\mu(s,X_s^\rho,\rho_s)-\mu(s,X_s^{\tilde\rho},\tilde\rho_s)\big)\,\mathrm{d}s\bigg|^\delta\bigg]\\
&\qquad+\mathbb{E}\bigg[\bigg|\int_0^t K_\sigma(s,t)\big(\sigma(s,X_s^\rho,\rho_s)-\sigma(s,X_s^{\tilde\rho},\tilde\rho_s)\big)\,\mathrm{d}B_s\bigg|^\delta\bigg]\\
&\quad\lesssim\bigg(\int_0^t|K_\mu(s,t)|^{(4+2\epsilon)/(4+\epsilon)}\,\mathrm{d}s\bigg)^{(4+\epsilon)/\epsilon}\int_0^t\mathbb{E}\big[|\mu(s,X_s^\rho,\rho_s)-\mu(s,X_s^{\tilde\rho},\tilde\rho_s)|^\delta\big]\,\mathrm{d}s\\
&\qquad+\mathbb{E}\bigg[\bigg(\int_0^t\big|K_\sigma(s,t)\big(\sigma(s,X_s^\rho,\rho_s)-\sigma(s,X_s^{\tilde\rho},\tilde\rho_s)\big)\big|^2\,\mathrm{d}s\bigg)^{\delta/2}\bigg]\\
&\quad\lesssim\int_0^t\mathbb{E}\big[\big|\mu(s,X_s^\rho,\rho_s)-\mu(s,X_s^{\tilde\rho},\tilde\rho_s)\big|^\delta\big]\,\mathrm{d}s\\
&\qquad+\bigg(\int_0^t|K_\sigma(s,t)|^{2+\epsilon}\,\mathrm{d}s\bigg)^{(4+2\epsilon)/(\epsilon(2+\epsilon))}\int_0^t\mathbb{E}\big[\big|\sigma(s,X_s^\rho,\rho_s)-\sigma(s,X_s^{\tilde\rho},\tilde\rho_s)\big|^\delta\big]\,\mathrm{d}s\\
&\quad\lesssim\int_0^t\big(\mathbb{E}\big[\big|X_s^\rho-X_s^{\tilde\rho}\big|^\delta\big]+W_\delta(\rho_s,\tilde\rho_s)^\delta\big)\,\mathrm{d}s
\end{aligned}\tag{19}$$

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

for $t \in [0, T]$, where we used Hölder's inequality in the drift integral with

$$\frac{4+2\epsilon}{4+\epsilon} < 1+\epsilon$$

(noting that $(4+2\epsilon)/(4+\epsilon)$ is the conjugate of $\delta/2$) such that, by the choice of δ in (3),

$$\frac{4+\epsilon}{4+2\epsilon} + \frac{1}{\delta} = 1,$$

and in the diffusion integral with $(2+\epsilon)/2$ such that (4) holds, Burkholder–Davis–Gundy's inequality applied to the stochastic processes

$$\left(\int_0^r K_\sigma(s,t)\big(\sigma(s, X_s^\rho, \rho_s) - \sigma(s, X_s^{\tilde{\rho}}, \tilde{\rho}_s)\big)\, \mathrm{d}B_s \right)_{r\in[0,t]},$$

Fubini's theorem, the integrability of the kernels from Assumption 1, and the Lipschitz continuity of μ and σ from Assumption 2. Since $\sup_{s\in[0,T]} \mathbb{E}[|X_s^\rho - X_s^{\tilde{\rho}}|^\delta] < \infty$, we can apply Grönwall's inequality to conclude that

$$\mathbb{E}\big[|X_t^\rho - X_t^{\tilde{\rho}}|^\delta\big] \lesssim \int_0^t W_\delta(\rho_s, \tilde{\rho}_s)^\delta \, \mathrm{d}s. \tag{20}$$

Since, by assumption, $\rho, \tilde{\rho} \in C\big([0,T]; \mathcal{P}_\delta(\mathbb{R}^d)\big)$, we can bound the Wasserstein distance by

$$W_\delta(\Phi(\rho)_t, \Phi(\tilde{\rho})_t) = W_\delta(\mathcal{L}(X_t^\rho), \mathcal{L}(X_t^{\tilde{\rho}})) \le \mathbb{E}[|X_t^\rho - X_t^{\tilde{\rho}}|^\delta]^{1/\delta},$$

cf. [12, (5.14)], and plugging this into (20) and taking the supremum, we obtain (18).

Step 3: We show that the solution map Φ has a unique fixed point.

First note that it is sufficient to show that Φ^k is a contraction (see [10, Theorem]), since the Wasserstein space $C([0,T]; \mathcal{P}_\delta(\mathbb{R}^d))$ is a complete metric space (see, e.g., [25, Proposition 2.2.8]), where Φ^k denotes the kth composition of Φ with itself. Let $C > 0$ denote the generic constant in (18). Then, iteratively for $k \in \mathbb{N}$,

$$\begin{aligned}\sup_{s\in[0,T]} W_\delta(\Phi^k(\rho)_s, \Phi^k(\tilde{\rho})_s)^\delta &\le C^k \int_0^T \frac{(T-s)^{k-1}}{(k-1)!} W_\delta(\rho_s, \tilde{\rho}_s)^\delta \, \mathrm{d}s \\ &\le \frac{C^k T^k}{k!} \sup_{s\in[0,T]} W_\delta(\rho_s, \tilde{\rho}_s)^\delta.\end{aligned}$$

Thus, choosing k large enough that $C^k T^k / k! < 1$, we see that the mapping Φ^k is a contraction and, hence, Φ admits a unique fixed point, which completes the proof. □

Next, we provide the proof of Theorem 3. We keep its presentation fairly short since it is in parts similar to the proof of Theorem 1.

Proof of Theorem 3. We again consider the solution map Φ, as defined in (16), but choose $\delta = 1$ and $d = 1$, that is,

$$\Phi\colon C\big([0,T]; \mathcal{P}_1(\mathbb{R})\big) \to C\big([0,T]; \mathcal{P}_1(\mathbb{R})\big), \qquad \rho \mapsto \Phi(\rho) := \big(\mathcal{L}(X_t^\rho)\big)_{t\in[0,T]},$$

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

where X^ρ is the unique L^p-solution to the stochastic Volterra equation

$$X_t = X_0 + \int_0^t K_\mu(s, t)\mu(s, X_s, \rho_s)\,\mathrm{d}s + \int_0^t K_\sigma(s, t)\sigma(s, X_s)\,\mathrm{d}B_s, \qquad t \in [0, T].$$

In the following we show that the solution map Φ possesses a unique fixed point. We proceed as in the proof of Theorem 1. Step 1 works exactly the same, using Lemmas 4 and 5, respectively, instead of Lemma 3, and Step 3 works exactly the same. That means we only need to show Step 2 or, more precisely, estimate (20) with $\delta = 1$. To do that, we treat separately the cases that Assumption 3 or Assumption 4 holds.

First, suppose the kernels K_μ, K_σ and initial condition X_0 satisfy Assumption 3. To get an analogous estimate as in (20), we use the semimartingale property of a solution $(X_t^\rho)_{t\in[0,T]}$ to (2) with fixed $\rho \in C([0, T]; \mathcal{P}_1(\mathbb{R}))$ (cf. [30, Lemma 3.6] or [27, Theorem 3.3]),

$$\begin{aligned}X_t^\rho - X_0 &= \int_0^t K_\sigma(s, s)\sigma(s, X_s^\rho)\,\mathrm{d}B_s + \int_0^t K_\mu(s, s)\mu(s, X_s^\rho, \rho_s)\,\mathrm{d}s\\ &\quad + \int_0^t \left(\int_0^s \partial_2 K_\mu(u, s)\mu(u, X_u^\rho, \rho_u)\,\mathrm{d}u + \int_0^s \partial_2 K_\sigma(u, s)\sigma(u, X_u^\rho)\,\mathrm{d}B_u \right) \mathrm{d}s,\end{aligned}$$

and the Yamada–Watanabe functions ϕ_n for $n \in \mathbb{N}$ (cf. [30, Proof of Theorem 5.3] or the original work, [36]) that approximate the absolute value function in the following way. Let $(a_n)_{n\in\mathbb{N}}$ be a strictly decreasing sequence with $a_0 = 1$ such that $a_n \to 0$ as $n \to \infty$ and

$$\int_{a_n}^{a_{n-1}} \frac{1}{|x|^{1+2\xi}}\,\mathrm{d}x = n,$$

where $\frac{1}{2} + \xi$ is the Hölder regularity of σ. Furthermore, we define a sequence of mollifiers: let $(\psi_n)_{n\in\mathbb{N}} \in C_0^\infty(\mathbb{R})$ be smooth functions with compact support such that $\operatorname{supp}(\psi_n) \subset (a_n, a_{n-1})$, and with the properties

$$0 \le \psi_n(x) \le \frac{2}{n|x|^{1+2\xi}} \text{ for all } x \in \mathbb{R}, \quad \text{and} \quad \int_{a_n}^{a_{n-1}} \psi_n(x)\,\mathrm{d}x = 1. \tag{21}$$

We set $\phi_n(x) := \int_0^{|x|} \left(\int_0^y \psi_n(z)\,\mathrm{d}z \right) \mathrm{d}y$, $x \in \mathbb{R}$. By (21) and the compact support of ψ_n, it follows that $\phi_n(\,\cdot\,) \to |\cdot|$ uniformly as $n \to \infty$. Since every ψ_n, and thus every ϕ_n, is zero in a neighborhood around zero, the functions ϕ_n are smooth with $\|\phi_n'\|_\infty \le 1$, $\phi_n'(x) = \operatorname{sgn}(x) \int_0^{|x|} \psi_n(y)\,\mathrm{d}y$, and $\phi_n''(x) = \psi_n(|x|)$ for $x \in \mathbb{R}$, where $\|\cdot\|_\infty$ denotes the sup-norm on $\mathbb{R}$.

Using ϕ_n, we apply Itô's formula to $\tilde{X}_t := X_t^\rho - X_t^{\tilde{\rho}}$, with the notation

$$\tilde{Z}_t := \int_0^t \left(\mu(s, X_s^\rho, \rho_s) - \mu(s, X_s^{\tilde{\rho}}, \tilde{\rho}_s) \right) \mathrm{d}s, \qquad Y_t^\rho := \int_0^t \sigma(s, X_s^\rho)\,\mathrm{d}B_s,$$

$H_t^\rho := \int_0^t \partial_2 K_\sigma(s, t)\,\mathrm{d}Y_s^\rho$, and $Y_t^{\tilde{\rho}}$ and $H_t^{\tilde{\rho}}$ analogously, as well as $\tilde{Y}_t := Y_t^\rho - Y_t^{\tilde{\rho}}$, and $\tilde{H}_t := H_t^\rho - H_t^{\tilde{\rho}}$, for $t \in [0, T]$, to obtain

$$\begin{aligned}\phi_n(\tilde{X}_t) &= \int_0^t \phi_n'(\tilde{X}_s)\,\mathrm{d}\tilde{X}_s + \frac{1}{2}\int_0^t \phi_n''(\tilde{X}_s)\,\mathrm{d}\langle \tilde{X} \rangle_s\\ &= \int_0^t \phi_n'(\tilde{X}_s) K_\mu(s, s)(\mu(s, X_s^\rho, \rho_s) - \mu(s, X_s^{\tilde{\rho}}, \tilde{\rho}_s))\,\mathrm{d}s\\ &\quad + \int_0^t \phi_n'(\tilde{X}_s) \left(\int_0^s \partial_2 K_\mu(u, s)\,\mathrm{d}\tilde{Z}_u \right) \mathrm{d}s\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

$$
\begin{aligned}
&+\int_0^t \phi_n'(\tilde{X}_s)\tilde{H}_s \,\mathrm{d}s+\int_0^t \phi_n'(\tilde{X}_s)K_\sigma(s,s)\,\mathrm{d}\tilde{Y}_s\\
&+\frac{1}{2}\int_0^t \phi_n''(\tilde{X}_s)K_\sigma(s,s)^2\big(\sigma(s,X_s^\rho)-\sigma(s,X_s^{\tilde{\rho}})\big)^2\,\mathrm{d}s\\
&=: I_{1,t}^n+I_{2,t}^n+I_{3,t}^n+I_{4,t}^n+I_{5,t}^n.
\end{aligned}
\tag{22}
$$

Note that H_t^ρ and $H_t^{\tilde{\rho}}$ are well-defined stochastic Itô integrals due to Assumption 3.

For $I_{1,t}^n$, the bound $\|\phi_n'\|_\infty \le 1$, the boundedness of K_μ, the Lipschitz continuity of μ, and Jensen's inequality yield

$$
\mathbb{E}[I_{1,t}^n] \lesssim \int_0^t \big(\mathbb{E}[|\tilde{X}_s|]+W_1(\rho_s,\tilde{\rho}_s)\big)\,\mathrm{d}s. \tag{23}
$$

For $I_{2,t}^n$, we additionally use the boundedness of $\partial_2 K_\mu(u,s)$ on Δ_T to obtain

$$
\mathbb{E}[I_{2,t}^n] \lesssim \int_0^t \big(\mathbb{E}[|\tilde{X}_s|]+W_1(\rho_s,\tilde{\rho}_s)\big)\,\mathrm{d}s. \tag{24}
$$

For $I_{3,t}^n$, we use $\|\phi_n'\|_\infty \le 1$ and the integration by parts formula to estimate

$$
\begin{aligned}
\mathbb{E}[I_{3,t}^n] &\le \int_0^t \mathbb{E}[|\tilde{H}_s|]\,\mathrm{d}s\\
&\le \int_0^t |\partial_2 K_\sigma(s,s)|\,\mathbb{E}[|\tilde{Y}_s|]\,\mathrm{d}s+\int_0^t\int_0^s |\partial_{21}K_\sigma(u,s)|\,\mathbb{E}[|\tilde{Y}_u|]\,\mathrm{d}u\,\mathrm{d}s\\
&\le \int_0^t \mathbb{E}[|\tilde{Y}_s|]\bigg(\partial_2 K_\sigma(s,s)+\int_s^t |\partial_{21}K_\sigma(s,u)|\,\mathrm{d}u\bigg)\,\mathrm{d}s \lesssim \int_0^t \mathbb{E}[|\tilde{Y}_s|]\,\mathrm{d}s,
\end{aligned}
\tag{25}
$$

with the boundedness of $\partial_2 K_\sigma(s,s)$ and $\int_s^t \partial_{21}K_\sigma(s,u)\,\mathrm{d}u$ from Assumption 3. For $I_{4,t}^n$, since $I_{4,t}^n$ is a martingale by [28, p. 73, Corollary 3] due to the boundedness of K_σ, the growth bound on σ and the finiteness of the moments of X^ρ and $X^{\tilde{\rho}}$ (cf. [30, Theorem 2.3]), we get

$$
\mathbb{E}[I_{4,t}^n] = \mathbb{E}\bigg[\int_0^t \phi_n'(\tilde{X}_s)K_\sigma(s,s)(\sigma(s,X_s^\rho)-\sigma(s,X_s^{\tilde{\rho}}))\,\mathrm{d}B_s\bigg] = 0. \tag{26}
$$

For $I_{5,t}^n$, by using the boundedness of K_σ, the Hölder continuity of σ, and the inequality $\phi_n''(x) \le 2/n|x|^{1+2\xi}$, we get

$$
\mathbb{E}[I_{5,t}^n] \lesssim \mathbb{E}\bigg[\int_0^t \phi_n''(\tilde{X}_s)|\tilde{X}_s|^{1+2\xi}\,\mathrm{d}s\bigg] \le \mathbb{E}\bigg[\int_0^t \frac{2}{n|\tilde{X}_s|^{1+2\xi}}|\tilde{X}_s|^{1+2\xi}\,\mathrm{d}s\bigg] \lesssim \frac{1}{n}. \tag{27}
$$

Sending $n \to \infty$ and combining the five previous estimates (23), (24), (25), (26), and (27) with (22) yields

$$
\mathbb{E}[|\tilde{X}_t|] \lesssim \int_0^t \big(\mathbb{E}[|\tilde{X}_s|]+\mathbb{E}[|\tilde{Y}_s|]+W_1(\rho_s,\tilde{\rho}_s)\big)\,\mathrm{d}s. \tag{28}
$$

To apply Grönwall's lemma, we set $M(t) := \mathbb{E}[|\tilde{X}_t|]+\mathbb{E}[|\tilde{Y}_t|]$ for $t \in [0,T]$. To find a bound for $\mathbb{E}[|\tilde{Y}_t|]$, we apply the integration by parts formula to obtain

$$
\begin{aligned}
\tilde{X}_t &= \int_0^t K_\mu(s,t)(\mu(s,X_s^\rho,\rho_s)-\mu(s,X_s^{\tilde{\rho}},\tilde{\rho}_s))\,\mathrm{d}s+\int_0^t K_\sigma(s,t)\,\mathrm{d}\tilde{Y}_s\\
&= \int_0^t K_\mu(s,t)(\mu(s,X_s^\rho,\rho_s)-\mu(s,X_s^{\tilde{\rho}},\tilde{\rho}_s))\,\mathrm{d}s+K_\sigma(t,t)\tilde{Y}_t-\int_0^t \partial_1 K_\sigma(s,t)\tilde{Y}_s\,\mathrm{d}s,
\end{aligned}
\tag{29}
$$

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

keeping in mind that $K_\sigma(\cdot, t)$ is absolutely continuous for every $t \in [0, T]$. Due to $|K_\sigma(t, t)| > C$ for some constant $C > 0$, we can rearrange (29) and use (28) to get

$$\begin{aligned}\mathbb{E}[|\tilde{Y}_t|] &\leq C\bigg(\int_0^t \mathbb{E}\big[|\mu(s, X_s^\rho, \rho_s) - \mu(s, X_s^{\tilde\rho}, \tilde\rho_s)|\big]\,\mathrm{d}s \\ &\qquad + \mathbb{E}[|\tilde{X}_t|] + \int_0^t |\partial_1 K_\sigma(s, t)|\mathbb{E}[|\tilde{Y}_s|]\,\mathrm{d}s\bigg) \\ &\lesssim \int_0^t \big(\mathbb{E}[|\tilde{X}_s|] + \mathbb{E}[|\tilde{Y}_s|] + W_1(\rho_s, \tilde\rho_s)\big)\,\mathrm{d}s. \end{aligned} \tag{30}$$

Now, Grönwall's lemma applied to (28) and (30) yields $M(t) \lesssim \int_0^t W_1(\rho_s, \tilde\rho_s)\,\mathrm{d}s$ and hence $\mathbb{E}[|X_t^\rho - X_t^{\tilde\rho}|] \lesssim \int_0^t W_1(\rho_s, \tilde\rho_s)\,\mathrm{d}s$, which is the analogous estimate of (20).

For the second case, suppose the kernels K_μ, K_σ and initial condition X_0 satisfy Assumption 4. We need to find an analogue to estimate (20). By using the notation $\tilde{X}_t := X_t^\rho - X_t^{\tilde\rho}$ and $Y_t^\rho := \int_0^t \mu(s, X_s^\rho, \rho_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s^\rho)\,\mathrm{d}B_s$, $Y_t^{\tilde\rho}$ analoguously, $\tilde{Y}_t := Y_t^\rho - Y_t^{\tilde\rho}$, and the semimartingale property

$$X_t^\rho - X_0 = \int_0^t \tilde{K}(0)\,\mathrm{d}Y_s^\rho + \int_0^t \int_0^s \tilde{K}'(s-u)\,\mathrm{d}Y_u^\rho\,\mathrm{d}s,$$

we can implement the Yamada–Watanabe approach with

$$\begin{aligned}\phi_n(\tilde{X}_t) &= \int_0^t \phi_n'(\tilde{X}_s)\tilde{K}(0)\,\mathrm{d}\tilde{Y}_s + \int_0^t \phi_n'(\tilde{X}_s)\int_0^s \tilde{K}'(s-u)\,\mathrm{d}\tilde{Y}_u\,\mathrm{d}s \\ &\quad + \frac{1}{2}\int_0^t \phi_n''(\tilde{X}_s)\tilde{K}(0)^2\big(\sigma(s, X_s^\rho) - \sigma(s, X_s^{\tilde\rho})\big)^2\,\mathrm{d}s \\ &=: I_{1,t}^n + I_{2,t}^n + I_{3,t}^n. \end{aligned} \tag{31}$$

Now, the Lipschitz assumption on μ applied to $I_{1,t}^n$ and $I_{2,t}^n$, the Hölder assumption on σ applied to $I_{3,t}^n$, the boundedness of $\tilde{K}$ and $\tilde{K}'$, the inequalities $\|\phi_n\|_\infty \leq 1$ and $\phi_n''(x) \leq 2/n|x|^{1+2\xi}$, and sending $n \to \infty$ yields, as in the first case, with Grönwall's lemma the inequality $\mathbb{E}[|X_t^\rho - X_t^{\tilde\rho}|] \lesssim \int_0^t W_1(\rho_s, \tilde\rho_s)\,\mathrm{d}s$, which implies the estimate (20) and, hence, yields the claimed well-posedness of the mean-field SVE (10). □

Remark 6. The well-posedness from Theorems 1 and 3, together with a general version of the classical Yamada–Watanabe result (see, e.g., [22, Theorem 1.5]; see also [22, Example 2.14]), implies that there is some measurable map $G\colon \mathbb{R}^d \times C([0, T]; \mathbb{R}^m) \to C([0, T]; \mathbb{R}^d)$ such that any solution X of (2) and (10), respectively, given some initial value X_0 and Brownian motion B, can be represented as $X = G(X_0, B)$. Hence, if $X, \tilde{X}$ are solutions of (2) and (10), respectively, for initial values $X_0, \tilde{X}_0$ with the same law and Brownian motions $B, \tilde{B}$, it is straightforward that $\mathcal{L}(X_t) = \mathcal{L}(\tilde{X}_t)$ almost surely for all $t \in [0, T]$.

5. Propagation of chaos: Proofs of Theorems 2 and 4

An important argument in the proofs of the propagation of chaos results will be to show that the coupled processes $((X^{N,i}, \underline{X}^i))_{1 \leq i \leq N}$ are identically distributed. The following lemma plays a crucial role. Recall that a sequence of random variables $(\zeta^1, \zeta^2, \ldots)$ is called

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

exchangeable if, for any $N \in \mathbb{N}$, the vectors $(\zeta^1, \ldots, \zeta^N)$ and $(\zeta^{\sigma(1)}, \ldots, \zeta^{\sigma(N)})$ have the same joint distribution, where $\{\sigma(1), \ldots, \sigma(N)\}$ is an arbitrary permutation of $\{1, \ldots, N\}$.

Lemma 6. *Let $(A, \mathcal{F}_A)$ and $(B, \mathcal{F}_B)$ be measurable spaces and for some fixed $N \in \mathbb{N}$, let $(\zeta^1, \ldots, \zeta^N)$ be an exchangeable family of A-valued random variables. Let $F\colon A \to B$ be a measurable function and define the family of random variables $(X^1, \ldots, X^N)$ by $X^i := F(\zeta^i)$ for $i \in \{1, \ldots, N\}$. Further, let $G\colon A^N \to B^N$ be a measurable function that fulfills the exchangeability property*

$$(y_1, \ldots, y_N) = G((x_1, \ldots, x_N)) \Rightarrow (y_{\sigma(1)}, \ldots, y_{\sigma(N)}) = G((x_{\sigma(1)}, \ldots, x_{\sigma(N)})) \tag{32}$$

for arbitrary $x_1, \ldots, x_N \in A$ and any permutation $\{\sigma(1), \ldots, \sigma(N)\}$ of $\{1, \ldots, N\}$. Define the family of random variables $(Y^1, \ldots, Y^N)$ by $(Y^1, \ldots, Y^N) := G((\zeta^1, \ldots, \zeta^N))$. Then, the coupled family of random variables $((X^i, Y^i))_{1 \le i \le N}$ is exchangeable.

Proof. Let $\{\sigma(1), \ldots, \sigma(N)\}$ be an arbitrary permutation of $\{1, \ldots, N\}$. By (32), we have

$$\begin{aligned}
Y^{\sigma(1)} &= G_1\big((\zeta^{\sigma(1)}, \zeta^{\sigma(2)}, \ldots, \zeta^{\sigma(N-1)}, \zeta^{\sigma(N)})\big),\\
Y^{\sigma(2)} &= G_1\big((\zeta^{\sigma(2)}, \zeta^{\sigma(3)}, \ldots, \zeta^{\sigma(N)}, \zeta^{\sigma(1)})\big),\\
&\vdots\\
Y^{\sigma(N)} &= G_1\big((\zeta^{\sigma(N)}, \zeta^{\sigma(1)}, \ldots, \zeta^{\sigma(N-2)}, \zeta^{\sigma(N-1)})\big),
\end{aligned} \tag{33}$$

where G_1 denotes the first component of the N-dimensional mapping G. Define $W^i := (X^i, Y^i)$ for $i \in \{1, \ldots, N\}$. Then, by the definition of X^i and (33),

$$\begin{aligned}
W^{\sigma(1)} &= \big(F(\zeta^{\sigma(1)}), G_1\big((\zeta^{\sigma(1)}, \zeta^{\sigma(2)}, \ldots, \zeta^{\sigma(N-1)}, \zeta^{\sigma(N)})\big)\big),\\
W^{\sigma(2)} &= \big(F(\zeta^{\sigma(2)}), G_1\big((\zeta^{\sigma(2)}, \zeta^{\sigma(3)}, \ldots, \zeta^{\sigma(N)}, \zeta^{\sigma(1)})\big)\big),\\
&\vdots\\
W^{\sigma(N)} &= \big(F(\zeta^{\sigma(N)}), G_1\big((\zeta^{\sigma(N)}, \zeta^{\sigma(1)}, \ldots, \zeta^{\sigma(N-2)}, \zeta^{\sigma(N-1)})\big)\big).
\end{aligned} \tag{34}$$

Analogously, we have

$$\begin{aligned}
W^1 &= \big(F(\zeta^1), G_1\big((\zeta^1, \zeta^2, \ldots, \zeta^{N-1}, \zeta^N)\big)\big),\\
W^2 &= \big(F(\zeta^2), G_1\big((\zeta^2, \zeta^3, \ldots, \zeta^N, \zeta^1)\big)\big),\\
&\vdots\\
W^N &= \big(F(\zeta^N), G_1\big((\zeta^N, \zeta^1, \ldots, \zeta^{N-2}, \zeta^{N-1})\big)\big).
\end{aligned} \tag{35}$$

Now, since, by assumption, $(\zeta^1, \ldots, \zeta^N)$ and $(\zeta^{\sigma(1)}, \ldots, \zeta^{\sigma(N)})$ have the same joint distribution, (34) and (35) yield that $(W^1, \ldots, W^N)$ and $(W^{\sigma(1)}, \ldots, W^{\sigma(N)})$ also have the same joint distribution, which proves the claimed exchangeability. □

We start with the proof of Theorem 2.

Proof of Theorem 2. Let us briefly outline the main steps of the proof:

Step 1. We show the existence of the system of processes $(X^{N,i})_{i=1,\ldots,N}$ uniquely solving (6), for every $N \in \mathbb{N}$.

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

Step 2. We prove the inequality

$$\mathbb{E}[|X_t^{N,i} - \underline{X}_t^i|^\delta] \lesssim \int_0^t \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^i) \right)^\delta \right] \mathrm{d}s, \qquad t \in [0, T], \tag{36}$$

for any $1 \le i \le N$. Recall that $X^{N,i}$ is defined in (6) and $\underline{X}^i$ is defined as the solution of the mean-field SVE (2) with initial condition X_0^i and driving Brownian motion B^i.

Step 3. We prove that the right-hand side of (36) tends to zero.

Step 4. We show that Steps 2 and 3 imply the statement.

Step 1: By the Lipschitz continuity of μ and σ, and the observation that $W_\delta(\bar{\rho}_x^N, \bar{\rho}_y^N)^\delta \le (1/N)\sum_{j=1}^N |x_j - y_j|^\delta$ for $x, y \in \mathbb{R}^{N\times d}$ with the notation $\bar{\rho}_x^N = (1/N)\sum_{j=1}^N \delta_{x_j} \in \mathcal{P}_\delta(\mathbb{R}^d)$, we obtain, for every $i \in \{1, \dots, N\}$, the Lipschitz condition

$$\begin{aligned}\left|\mu(t, x_i, \bar{\rho}_x^N) - \mu(t, y_i, \bar{\rho}_y^N)\right|^\delta + \left|\sigma(t, x_i, \bar{\rho}_x^N) - \sigma(t, y_i, \bar{\rho}_y^N)\right|^\delta &\lesssim |x_i - y_i|^\delta + \frac{1}{N}\sum_{j=1}^N |x_j - y_j|^\delta \\ &\lesssim \|x - y\|_{N\times d}^\delta,\end{aligned}$$

where $\|\cdot\|_{N\times d}$ denotes the row sum norm on $\mathbb{R}^{N\times d}$. With the notation $\tilde{\mu}_i(t, x) := \mu(t, x_i, \bar{\rho}_x^N)$ and $\tilde{\sigma}_i(t, x)$ analogously for any $1 \le i \le N$, we directly conclude that the growth condition is fulfilled by

$$\begin{aligned}\left|\tilde{\mu}_i(t, x)\right| + \left|\tilde{\sigma}_i(t, x)\right| &\le \left|\tilde{\mu}_i(t, x) - \tilde{\mu}_i(t, 0)\right| + \left|\tilde{\sigma}_i(t, x) - \tilde{\sigma}_i(t, 0)\right| + \left|\tilde{\mu}_i(t, 0)\right| + \left|\tilde{\sigma}_i(t, 0)\right| \\ &\lesssim \|x\|_{N\times d} + \left|\mu(t, 0, \delta_0)\right| + \left|\sigma(t, 0, \delta_0)\right| \\ &\lesssim \|x\|_{N\times d} + C_{\delta_0} \lesssim 1 + \|x\|_{N\times d}\end{aligned}$$

for all $t \in [0, T]$, $x \in \mathbb{R}^{N\times d}$. Thus, due to the equivalence of all norms on the finite-dimensional vector space $\mathbb{R}^{N\times d}$, we can apply the standard Volterra well-posedness result for Lipschitz coefficients from Lemma 3 to obtain the system of processes $(X^{N,i})_{i=1,\dots,N,}$ that uniquely solves (6) for every $N \in \mathbb{N}$.

Step 2: We consider the first summand on the left-hand side of (7), i.e. $\mathbb{E}[|X_t^{N,i} - \underline{X}_t^i|^\delta]$. Using Hölder's inequality as in (19), Fubini's theorem, and the Burkholder–Davis–Gundy inequality such as the Lipschitz continuity of μ and σ, we can bound, for $1 \le i \le N$,

$$\begin{aligned}&\mathbb{E}[|X_t^{N,i} - \underline{X}_t^i|^\delta] \\ &\quad= \mathbb{E}\bigg[\bigg| \int_0^t K_\mu(s, t)\big(\mu(s, X_s^{N,i}, \bar{\rho}_s^N) - \mu(s, \underline{X}_s^i, \mathcal{L}(\underline{X}_s^i))\big)\,\mathrm{d}s \\ &\qquad\qquad + \int_0^t K_\sigma(s, t)\big(\sigma(s, X_s^{N,i}, \bar{\rho}_s^N) - \sigma(s, \underline{X}_s^i, \mathcal{L}(\underline{X}_s^i))\big)\,\mathrm{d}B_s^i\bigg|^\delta\bigg] \\ &\quad\lesssim \left(\int_0^t |K_\mu(s, t)|^{(4+2\epsilon)/(4+\epsilon)}\,\mathrm{d}s \right)^{(4+\epsilon)/\epsilon} \int_0^t \mathbb{E}\big[\big|\mu(s, X_s^{N,i}, \bar{\rho}_s^N) - \mu(s, \underline{X}_s^i, \mathcal{L}(\underline{X}_s^i))\big|^\delta\big]\,\mathrm{d}s \\ &\qquad + \mathbb{E}\bigg[\bigg(\int_0^t \big|K_\sigma(s, t)\big(\sigma(s, X_s^{N,i}, \bar{\rho}_s^N) - \sigma(s, \underline{X}_s^i, \mathcal{L}(\underline{X}_s^i))\big|^2\,\mathrm{d}s\bigg)^{\delta/2}\bigg]\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

$$\lesssim \int_0^t \mathbb{E}\big[\big|\mu(s, X_s^{N,i}, \bar{\rho}_s^N) - \mu(s, \underline{X}_s^i, \mathcal{L}(\underline{X}_s^i))\big|^\delta\big]\,\mathrm{d}s$$
$$+ \bigg(\int_0^t |K_\sigma(s,t)|^{2+\epsilon}\,\mathrm{d}s\bigg)^{4/\epsilon} \int_0^t \mathbb{E}\big[\big|\sigma(s, X_s^{N,i}, \bar{\rho}_s^N) - \sigma(s, \underline{X}_s^i, \mathcal{L}(\underline{X}_s^i))\big|^\delta\big]\,\mathrm{d}s$$
$$\lesssim \int_0^t \mathbb{E}\big[|X_s^{N,i} - \underline{X}_s^i|^\delta + W_\delta(\bar{\rho}_s^N, \mathcal{L}(\underline{X}_s^i))^\delta\big]\,\mathrm{d}s \tag{37}$$

for any $t \in [0, T]$. By Remark 6, we obtain that $\mathcal{L}(\underline{X}_s^i) = \mathcal{L}(\underline{X}_s^1)$. Hence, we get

$$W_\delta\big(\bar{\rho}_s^N, \mathcal{L}(\underline{X}_s^i)\big)^\delta = W_\delta\Big(\bar{\rho}_s^N, \mathcal{L}(\underline{X}_s^1)\Big)^\delta$$
$$\leq 2^\delta W_\delta\Big(\bar{\rho}_s^N, \frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}\Big)^\delta + 2^\delta W_\delta\Big(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\Big)^\delta$$
$$\lesssim \frac{1}{N}\sum_{j=1}^N |X_s^{N,j} - \underline{X}_s^j|^\delta + W_\delta\Big(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\Big)^\delta. \tag{38}$$

Moreover, by Remark 6, we can find a measurable map $G\colon \mathbb{R}^d \times C([0,T];\mathbb{R}^m) \to C([0,T];\mathbb{R}^d)$ such that, for any $1 \leq i \leq N$, $\underline{X}^i = G(X_0^i, B^i)$. In the same way, there is a measurable map $G_N\colon (\mathbb{R}^d \times C([0,T];\mathbb{R}^m))^N \to C([0,T];\mathbb{R}^d)^N$, such that

$$(X^{N,1}, \ldots, X^{N,N}) = G_N\big((X_0^1, \ldots, X_0^N), (B^1, \ldots, B^N)\big).$$

More generally, by the symmetry of the system (6), for any permutation ς of $\{1, \ldots, N\}$,

$$(X^{N,\varsigma(1)}, \ldots, X^{N,\varsigma(N)}) = G_N\big((X_0^{\varsigma(1)}, \ldots, X_0^{\varsigma(N)}), (B^{\varsigma(1)}, \ldots, B^{\varsigma(N)})\big).$$

Hence, since the random variables $((X_0^i, B^i))_{1\leq i\leq N}$ are i.i.d. and, in particular, exchangeable, we can apply Lemma 6 to obtain that the coupled processes $((X^{N,i}, \underline{X}^i))_{1\leq i\leq N}$ are exchangeable and hence, in particular, are identically distributed. For $i = 1$ we can insert (38) into (37) and conclude by Jensen's inequality that

$$\mathbb{E}[|X_t^{N,1} - \underline{X}_t^1|^\delta]$$
$$\lesssim \int_0^t \mathbb{E}\bigg[|X_s^{N,1} - \underline{X}_s^1|^\delta + \frac{1}{N}\sum_{j=1}^N |X_s^{N,j} - \underline{X}_s^j|^\delta + W_\delta\Big(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\Big)^\delta\bigg]\,\mathrm{d}s$$
$$= \int_0^t \mathbb{E}\bigg[2|X_s^{N,1} - \underline{X}_s^1|^\delta + W_\delta\Big(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\Big)^\delta\bigg]\,\mathrm{d}s.$$

Using Grönwall's lemma, we deduce that

$$\mathbb{E}[|X_t^{N,1} - \underline{X}_t^1|^\delta] \lesssim \int_0^t \mathbb{E}\bigg[W_\delta\Big(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\Big)^\delta\bigg]\,\mathrm{d}s,$$

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

and since the processes $((X^{N,i}, \underline{X}^i))_{1\le i\le N}$ are identically distributed, this completes Step 2.

Step 3: First, we show that

$$\lim_{N\to\infty} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\right)^\delta \right] = 0 \tag{39}$$

for any $s \in [0, T]$ by showing convergence in probability and uniform integrability. By the Glivenko–Cantelli theorem (see [32, Chapter 26, Theorem 1] for a general version) and since the $\underline{X}^j$ are i.i.d., we get the convergence $(1/N)\sum_{j=1}^N \delta_{\underline{X}_s^j} \to \mathcal{L}(\underline{X}_s^1)$ as $N \to \infty$ almost surely, and hence in probability. Furthermore, again using the notation $\bar{\rho}_s^N = (1/N)\sum_{j=1}^N \delta_{\underline{X}_s^j}$, we can bound using Hölder's inequality and the boundedness of all moments of $\underline{X}_s^i$, $1 \le i \le N$, in (5), to get

$$\begin{aligned}
&\sup_{N\in\mathbb{N}} \mathbb{E}\big[W_\delta(\bar{\rho}_s^N, \mathcal{L}(\underline{X}_s^1))^\delta \mathbf{1}_{\{W_\delta(\bar{\rho}_s^N,\mathcal{L}(\underline{X}_s^1))>K\}}\big] \\
&\quad \le K^{-1} \sup_{N\in\mathbb{N}} \mathbb{E}\big[W_\delta(\bar{\rho}_s^N, \mathcal{L}(\underline{X}_s^1))^{\delta+1}\big] \\
&\quad \le K^{-1} \sup_{N\in\mathbb{N}} \mathbb{E}\big[W_{\delta+1}(\bar{\rho}_s^N, \mathcal{L}(\underline{X}_s^1))^{\delta+1}\big] \\
&\quad \le K^{-1} \sup_{N\in\mathbb{N}} \mathbb{E}\big[W_{\delta+1}(\bar{\rho}_s^N, \delta_0)^{\delta+1} + W_{\delta+1}(\delta_0, \mathcal{L}(\underline{X}_s^1))^{\delta+1}\big] \\
&\quad = K^{-1} \sup_{N\in\mathbb{N}} \mathbb{E}\left[\frac{1}{N}\left(\sum_{i=1}^N |\underline{X}_s^i|^{\delta+1}\right) + |\underline{X}_s^1|^{\delta+1}\right] \\
&\quad = 2K^{-1}\mathbb{E}[|\underline{X}_s^1|^{\delta+1}] \to 0
\end{aligned} \tag{40}$$

as $K \to \infty$, which shows uniform δ-integrability of the family of random variables

$$\left(W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\right)\right)_{N\in\mathbb{N}}.$$

Hence, Vitali's convergence theorem (see [9, Theorem 4.5.4]) reveals the L^δ-convergence as claimed in (39).

To conclude Step 3, it remains to show that the convergence (39) is uniform in s. Therefore, we first notice that, for any $p \ge \delta$,

$$\begin{aligned}
\mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\right)^p\right] &\le \mathbb{E}\left[W_p\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\right)^p\right] \\
&\lesssim \mathbb{E}\left[W_p\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \delta_0\right)^p\right] + W_p(\delta_0, \mathcal{L}(\underline{X}_s^1))^p \\
&= \frac{1}{N}\mathbb{E}\left[\sum_{j=1}^N |\underline{X}_s^j|^p\right] + \mathbb{E}[|\underline{X}_s^1|^p] = 2\mathbb{E}[|\underline{X}_s^1|^p] < \infty,
\end{aligned} \tag{41}$$

by (5). With Jensen's inequality, (41) also follows for $1 \le p < \delta$.

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

Let $k := \lceil \delta \rceil \geq \delta$ denote the smallest integer greater than or equal to δ. Notice that with the same argument as in (40), by substituting the exponent δ by k and then bounding from above using the $(k+1)$-Wasserstein distance and again by Vitali's convergence theorem, the L^k-convergence of the δ-Wasserstein distance in (39) also follows. Once we show that this L^k-convergence is uniform in s, then it will follow that

$$\lim_{N\to\infty} \sup_{s\in[0,T]} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right)^\delta \right] \leq \lim_{N\to\infty} \sup_{s\in[0,T]} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right)^k \right]^{\delta/k} = 0. \quad (42)$$

Therefore, using the factorization $a^k - b^k = (a-b)\sum_{r=0}^{k-1} a^{k-1-r} b^r$ and Hölder's inequality with δ and $q = (4+2\epsilon)/(4+\epsilon)$ such that $1/\delta + 1/q = 1$, we get

$$\begin{aligned}
&\left| \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_t^j}, \mathcal{L}(\underline{\mathbf{X}}_t^1)\right)^k \right] - \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right)^k \right] \right| \\
&\quad = \Bigg| \mathbb{E}\Bigg[\left(W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_t^j}, \mathcal{L}(\underline{\mathbf{X}}_t^1)\right) - W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right) \right) \\
&\qquad\qquad \times \sum_{r=0}^{k-1} W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_t^j}, \mathcal{L}(\underline{\mathbf{X}}_t^1)\right)^{k-1-r} W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right)^r \Bigg] \Bigg| \\
&\quad \leq \Bigg| \mathbb{E}\Bigg[\left(W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_t^j}, \mathcal{L}(\underline{\mathbf{X}}_t^1)\right) - W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right) \right)^\delta \Bigg]^{1/\delta} \\
&\qquad \times \mathbb{E}\Bigg[\left(\sum_{r=0}^{k-1} W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_t^j}, \mathcal{L}(\underline{\mathbf{X}}_t^1)\right)^{k-1-r} W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right)^r \right)^q \Bigg]^{1/q} \Bigg|. \quad (43)
\end{aligned}$$

Again using Hölder's inequality such as (41), we can bound the second expectation by

$$\begin{aligned}
&\mathbb{E}\Bigg[\left(\sum_{r=0}^{k-1} W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_t^j}, \mathcal{L}(\underline{\mathbf{X}}_t^1)\right)^{k-1-r} W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right)^r \right)^q \Bigg]^{1/q} \\
&\lesssim \left(\sum_{r=0}^{k-1} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_t^j}, \mathcal{L}(\underline{\mathbf{X}}_t^1)\right)^{q(k-1-r)} W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right)^{qr} \right] \right)^{1/q} \\
&\lesssim \left(\sum_{r=0}^{k-1} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_t^j}, \mathcal{L}(\underline{\mathbf{X}}_t^1)\right)^{2q(k-1-r)} \right]^{1/2} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{\mathbf{X}}_s^j}, \mathcal{L}(\underline{\mathbf{X}}_s^1)\right)^{2qr} \right]^{1/2} \right)^{1/q} \\
&< \infty. \quad (44)
\end{aligned}$$

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

Inserting (44) into (43) and using the triangle inequality

$$
\begin{aligned}
&W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_t^j}, \mathcal{L}(\underline{X}_t^1)\right) \\
&\quad\le W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_t^j}, \frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}\right) + W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\right) + W_\delta\big(\mathcal{L}(\underline{X}_s^1), \mathcal{L}(\underline{X}_t^1)\big),
\end{aligned}
$$

which also holds if we switch s and t, we arrive at

$$
\begin{aligned}
&\left|\mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_t^j}, \mathcal{L}(\underline{X}_t^1)\right)^k\right] - \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\right)^k\right]\right| \\
&\quad\lesssim \mathbb{E}\left[\left|\left(W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_t^j}, \mathcal{L}(\underline{X}_t^1)\right) - W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^1)\right)\right)\right|^\delta\right]^{1/\delta} \\
&\quad\lesssim \mathbb{E}\left[\left(W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_t^j}, \frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}\right) + W_\delta\big(\mathcal{L}(\underline{X}_t^1), \mathcal{L}(\underline{X}_s^1)\big)\right)^\delta\right]^{1/\delta} \\
&\quad\lesssim \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_t^j}, \frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}\right)^\delta\right]^{1/\delta} + \mathbb{E}\big[W_\delta\big(\mathcal{L}(\underline{X}_t^1), \mathcal{L}(\underline{X}_s^1)\big)^\delta\big]^{1/\delta} \\
&\quad\lesssim \mathbb{E}[|\underline{X}_t^1 - \underline{X}_s^1|^\delta]^{1/\delta} \\
&\quad\lesssim |t-s|^\beta,
\end{aligned}
$$

where the last line holds by Remark 4 for any $\beta \in (0, \gamma - 1/p)$ with $\gamma \in \big(0, \frac{1}{2}\big]$ from Assumption 1. Hence, we obtain that (42) holds, which together with (36) shows that

$$
\lim_{N\to\infty} \sup_{t\in[0,T]} \mathbb{E}[|X_t^{N,1} - \underline{X}_t^1|^\delta] = 0, \tag{45}
$$

and knowing that $((X^{N,i}, \underline{X}^i))_{1\le i\le N}$ are identically distributed, this completes Step 3.

Step 4: We already know from Step 3 that the first summand in (7) converges to zero. For the second summand, we use the triangle inequality and Jensen's inequality to obtain

$$
\begin{aligned}
&\sup_{t\in[0,T]} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_t^{N,i}}, \mathcal{L}(\underline{X}_t^1)\right)^\delta\right] \\
&\quad\lesssim \sup_{t\in[0,T]} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{i=1}^N \delta_{X_t^{N,i}}, \frac{1}{N}\sum_{i=1}^N \delta_{\underline{X}_t^i}\right)^\delta\right] + \sup_{t\in[0,T]} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{i=1}^N \delta_{\underline{X}_t^i}, \mathcal{L}(\underline{X}_t^1)\right)^\delta\right] \\
&\quad\lesssim \sup_{t\in[0,T]} \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^N |X_t^{N,i} - \underline{X}_t^i|^\delta\right] + \sup_{t\in[0,T]} \mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{i=1}^N \delta_{\underline{X}_t^i}, \mathcal{L}(\underline{X}_t^1)\right)^\delta\right],
\end{aligned} \tag{46}
$$

which also tends to 0 as $N \to \infty$ by (42) and (45). □

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

We continue with the proof of Theorem 4. Since the proof is similar to that of Theorem 4, we focus, for the sake of brevity, on the main differences.

Proof of Theorem 4. We prove the statement by using the same Steps 1–4 as in the proof of Theorem 2, but with $\delta = 1$. For Step 2, though, we need to consider separately the cases where Assumption 3, the first case in the proof of Theorem 1, and Assumption 4, the second case, hold.

Step 1: By Remark 3, we obtain as in the proof of Theorem 2 the unique system of stochastic processes $(X^{N,i})_{i=1,\dots,N}$ that solves (6).

Step 2: In the first case, suppose the kernels K_μ, K_σ and initial condition X_0 satisfy Assumption 3. To mimic the inequality (36), we use the semimartingale property

$$
\begin{aligned}
X_t^{N,i} - \underline{X}_t^i = & \int_0^t K_\sigma(s,s)\big(\sigma(s, X_s^{N,i}) - \sigma(s, \underline{X}_s^i)\big)\,\mathrm{d}B_s \\
& + \int_0^t K_\mu(s,s)\big(\mu(s, X_s^{N,i}, \bar{\rho}_s^N) - \mu(s, \underline{X}_s^i, \mathcal{L}(\underline{X}_s^i))\big)\,\mathrm{d}s \\
& + \int_0^t \bigg(\int_0^s \partial_2 K_\mu(u,s)\big(\mu(u, X_u^{N,i}, \bar{\rho}_u^N) - \mu(u, \underline{X}_u^i, \mathcal{L}(\underline{X}_u^i))\big)\,\mathrm{d}u \\
& \qquad + \int_0^s \partial_2 K_\sigma(u,s)\big(\sigma(u, X_u^{N,i}) - \sigma(u, \underline{X}_u^i)\big)\,\mathrm{d}B_u \bigg)\,\mathrm{d}s
\end{aligned}
$$

to perform a Yamada–Watanabe approach exactly as we did around equality (22), and obtain, for fixed $i \in \{1, \dots, N\}$ with the notation $M^{N,i}(t) := \mathbb{E}[|X_t^{N,i} - \underline{X}_t^i|] + \mathbb{E}[|\tilde{Y}_t|]$, where $\tilde{Y}_t := \int_0^t \sigma(s, X_s^{N,i})\,\mathrm{d}B_s^i - \int_0^t \sigma(s, \underline{X}_s^i)\,\mathrm{d}B_s^i$, that $M^{N,i}(t) \lesssim \int_0^t \big(M^{N,i}(s) + \mathbb{E}[W_1(\bar{\rho}_s^N, \mathcal{L}(\underline{X}_s^i))]\big)\,\mathrm{d}s$, such that, proceeding as in the proof of Theorem 2, including applying Grönwall's inequality, we obtain

$$
\mathbb{E}\Big[|X_t^{N,i} - \underline{X}_t^i|\Big] \lesssim \int_0^t \mathbb{E}\bigg[W_1\bigg(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^i) \bigg) \bigg]\,\mathrm{d}s. \tag{47}
$$

In the second case, suppose the kernels K_μ, K_σ and initial condition X_0 satisfy Assumption 4. As in the previous case, to mimic inequality (36) we use the semimartingale property

$$
\begin{aligned}
X_t^{N,i} - \underline{X}_t^i = & \int_0^t \tilde{K}(0)\big(\sigma(s, X_s^{N,i}) - \sigma(s, \underline{X}_s^i)\big)\,\mathrm{d}B_s \\
& + \int_0^t \tilde{K}(0)\big(\mu(s, X_s^{N,i}, \bar{\rho}_s^N) - \mu(s, \underline{X}_s^i, \mathcal{L}(\underline{X}_s^i))\big)\,\mathrm{d}s \\
& + \int_0^t \bigg(\int_0^s \tilde{K}'(s-u)\big(\mu(u, X_u^{N,i}, \bar{\rho}_u^N) - \mu(u, \underline{X}_u^i, \mathcal{L}(\underline{X}_u^i))\big)\,\mathrm{d}u \\
& \qquad + \int_0^s \tilde{K}'(s-u)\big(\sigma(u, X_u^{N,i}) - \sigma(u, \underline{X}_u^i)\big)\,\mathrm{d}B_u \bigg)\,\mathrm{d}s
\end{aligned}
$$

to perform a Yamada–Watanabe approach and apply Grönwall's inequality as in (31), which yields

$$
\mathbb{E}[|X_t^{N,i} - \underline{X}_t^i|] \lesssim \int_0^t \mathbb{E}\bigg[W_1\bigg(\frac{1}{N}\sum_{j=1}^N \delta_{\underline{X}_s^j}, \mathcal{L}(\underline{X}_s^i) \bigg) \bigg]\,\mathrm{d}s.
$$

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

Step 3: Obtaining the convergence to zero uniformly in s of the right-hand side of (47) now follows easily by using

$$\mathbb{E}\left[W_1\left(\frac{1}{N}\sum_{j=1}^{N}\delta_{\underline{X}_s^j},\mathcal{L}(\underline{X}_s^1)\right)\right]\leq \mathbb{E}\left[W_2\left(\frac{1}{N}\sum_{j=1}^{N}\delta_{\underline{X}_s^j},\mathcal{L}(\underline{X}_s^1)\right)^2\right]^{1/2},$$

and then using [12, (5.19)], and proceeding as in [13, Proof of Theorem 2.12].

Step 4: As in (46), we obtain

$$\begin{aligned}&\sup_{0\leq t\leq T}\mathbb{E}\left[W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{X_t^{N,i}},\mathcal{L}(\underline{X}_t^1)\right)\right]\\&\quad\lesssim \sup_{0\leq t\leq T}\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}|X_t^{N,i}-\underline{X}_t^i|\right]+\sup_{0\leq t\leq T}\mathbb{E}\left[W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{\underline{X}_t^i},\mathcal{L}(\underline{X}_t^1)\right)\right],\end{aligned}\tag{48}$$

which tends to zero by the uniform convergence to zero of the right-hand side of (47), and finishes the proof. □

6. Rate of convergence: Proofs of Lemmas 1 and 2

The proofs of Lemmas 1 and 2 rely on a quantitative Glivenko–Cantelli theorem due to Fournier and Guillin [17], which provides a sharp estimate of the δ-Wasserstein distance. For the sake of completeness, we recall [17, Theorem 1] in the following lemma.

Lemma 7. *Let $\delta>0$ and $\bar{\rho}^N:=(1/N)\sum_{i=1}^N\delta_{X^i}$ be the empirical distribution of i.i.d. random variables $(X^i)_{i=1,\dots,N}$ with common distribution ρ such that $\rho\in\mathcal{P}_p(\mathbb{R}^d)$ for every $p\geq 1$. Then $\mathbb{E}[W_\delta(\bar{\rho}^N,\rho)^\delta]\lesssim\varepsilon_N$, where $(\varepsilon_N)_{N\in\mathbb{N}}$ is given by (9), i.e.*

$$\varepsilon_N=\begin{cases}N^{-1/2} & \text{if } d<2\delta,\\ N^{-1/2}\log_2(1+N) & \text{if } d=2\delta,\\ N^{-\delta/d} & \text{if } d>2\delta,\end{cases}$$

and $\mathbb{E}[W_1(\bar{\rho}^N,\rho)]\lesssim N^{-1/2}$.

With this lemma at hand, we can prove Lemmas 1 and 2.

Proof of Lemma 1. By Lemma 7, for any $t\in[0,T]$,

$$\mathbb{E}\left[W_\delta\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{\underline{X}_t^i},\mathcal{L}(\underline{X}_t^1)\right)^\delta\right]\lesssim\varepsilon_N,\tag{49}$$

where $(\varepsilon_N)_{N\in\mathbb{N}}$ is given by (9) and the right-hand side does not depend on t. Plugging (49) into (36) and taking the supremum over $[0,T]$ and maximum over $1,\dots,N$ shows the desired convergence rate of the first term in (8). Then, using this and plugging (49) into (46) gives the desired rate for the second term. □

Proof of Lemma 2. First, suppose the kernels K_μ,K_σ and initial condition X_0 satisfy Assumption 3. By Lemma 7 we obtain

$$\mathbb{E}\left[W_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{\underline{X}_t^i},\mathcal{L}(\underline{X}_t^1)\right)\right]\lesssim N^{-1/2}\tag{50}$$

independent of $t\in[0,T]$. Plugging (50) into (47) and (48) yields the statement.

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press

Otherwise, suppose the kernels K_μ, K_σ and initial condition X_0 satisfy Assumption 4. Plugging (50) into the analogues of (47) and (48) yields the statement. □

Acknowledgements

D. J. Prömel and D. Scheffels would like to thank P. Nikolaev for fruitful discussions which helped to improve the present work.

Funding information

D. Scheffels gratefully acknowledges financial support by the Research Training Group 'Statistical Modeling of Complex Systems' (RTG 1953) funded by the German Science Foundation (DFG).

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] ABI JABER, E., CUCHIERO, C., LARSSON, M. AND PULIDO, S. (2021). A weak solution theory for stochastic Volterra equations of convolution type, *Ann. Appl. Prob.* **31**, 2924–2952.

[2] ABI JABER, E. AND EL EUCH, O. (2019). Multifactor approximation of rough volatility models, *SIAM J. Financial Math.* **10**, 309–349.

[3] ABI JABER, E., LARSSON, M. AND PULIDO, S. (2019). Affine Volterra processes, *Ann. Appl. Prob.* **29**, 3155–3200.

[4] BAHLALI, K., MEZERDI, M. A. AND MEZERDI, B. (2020). Stability of McKean–Vlasov stochastic differential equations and applications, *Stoch. Dynam.* **20**, 2050007.

[5] BAILLEUL, I., CATELLIER, R. AND DELARUE, F. (2020). Solving mean field rough differential equations, *Electron. J. Prob.* **25**, 21.

[6] BAILLEUL, I., CATELLIER, R, and DELARUE, F. (2021). Propagation of chaos for mean field rough differential equations, *Ann. Prob.* **49**, 944–996.

[7] BERGER, M. A. AND MIZEL, V. J. (1980). Volterra equations with Itô integrals. I, *J. Integral Equat*. **2**, 187–245.

[8] BERGER, M. A. AND MIZEL, V. J. (1980). Volterra equations with Itô integrals. II, *J. Integral Equat*. **2**, 319–337.

[9] BOGACHEV, V. I. (2007). *Measure Theory*, Vol. I. Springer, Berlin.

[10] BRYANT, V. W. (1968). A remark on a fixed-point theorem for iterated mappings, *Amer. Math. Monthly* **75**, 399–400.

[11] CARMONA, R. (2016). *Lectures on BSDEs, Stochastic Control, and Stochastic Differential Games with Financial Applications*. SIAM, Philadelphia, PA.

[12] CARMONA, R. AND DELARUE, F. (2018). *Probabilistic Theory of Mean Field Games with Applications I* (Prob. Theory Stoch. Model. **83**). Springer, Cham.

[13] CARMONA, R. AND DELARUE, F. (2018). *Probabilistic Theory of Mean Field Games with Applications. II* (Prob. Theory Stoch. Model. **84**). Springer, Cham.

[14] CHAINTRON, L.-P. AND DIEZ, A. (2022). Propagation of chaos: A review of models, methods and applications. I. Models and methods. *Kinet. Relat. Model*. **15**, 895–1015.

[15] CHAINTRON, L.-P. AND DIEZ, A. (2022). Propagation of chaos: A review of models, methods and applications. II. Applications. *Kinet. Relat. Model*. **15**, 1017–1173.

[16] COGHI, M., DEUSCHEL, J.-D., FRIZ, P. K. AND MAURELLI, M. (2020). Pathwise McKean–Vlasov theory with additive noise. *Ann. Appl. Prob.* **30**, 2355–2392.

[17] FOURNIER, N. AND GUILLIN, A. (2015). On the rate of convergence in Wasserstein distance of the empirical measure. *Prob. Theory Relat. Fields* **162**, 707–738.

[18] HUANG, X. AND WANG, X. (2023). Path dependent McKean–Vlasov SDEs with Hölder continuous diffusion. *Discrete Continuous Dynam. Syst. – S* **16**, 982–998.

[19] JABIN, P. E. AND WANG, Z. (2017). Mean field limit for stochastic particle systems. In *Active Particles*, Vol. 1, eds P. Degond, N. Bellomo and E. Tadmor. Birkhäuser, Boston, MA, pp. 379–402.

[20] KAC, M. (1956). Foundations of kinetic theory. In *Proc. 3rd Berkeley Symp. Math. Statist. Prob. 1954–1955,* Vol. III, ed. J. Neyman. University of California Press, Berkeley, CA, pp. 171–197.

[21] KALININ, A., MEYER-BRANDIS, T. AND PROSKE, F. (2024). Stability, uniqueness and existence of solutions to McKean–Vlasov stochastic differential equations in arbitrary moments. *J. Theoret. Prob.* **37**, 2941–2989.

[22] KURTZ, T. G. (2014). Weak and strong solutions of general stochastic models, *Electron. Commun. Prob.* **19**, 58.

[23] MCKEAN, H. P. (1966). A class of Markov processes associated with nonlinear parabolic equations. *Proc. Nat. Acad. Sci. USA* **56**, 1907–1911.

[24] MCKEAN, Jr., H. P. (1967). Propagation of chaos for a class of non-linear parabolic equations. In *Stochastic Differential Equations* (Lecture Ser. Diff. Equat., Session 7, Catholic University). Air Force Office Sci. Res., Arlington, VA, pp. 41–57.

[25] PANARETOS, V. M. AND ZEMEL, Y. (2020). *An Invitation to Statistics in Wasserstein Space*. Springer, Cham.

[26] PARDOUX, É. AND PROTTER, P. (1990). Stochastic Volterra equations with anticipating coefficients. *Ann. Prob.* **18**, 1635–1655.

[27] PROTTER, P. (1985). Volterra equations driven by semimartingales, *Ann. Prob.* **13**, 519–530.

[28] PROTTER, P. E. (2004). *Stochastic Integration and Differential Equations*, 2nd ed. (Appl. Math. (New York) **21**). Springer, New York.

[29] PRÖMEL, D. J. AND SCHEFFELS, D. (2023). On the existence of weak solutions to stochastic Volterra equations. *Electron. Commun. Prob.* **28**, 52.

[30] PRÖMEL, D. J. AND SCHEFFELS, D. (2023). Stochastic Volterra equations with Hölder diffusion coefficients. *Stoch. Process. Appl.* **161**, 291–315.

[31] SHI, Y., WANG, T. AND YONG, J. (2013). Mean-field backward stochastic Volterra integral equations. *Discrete Contin. Dynam. Syst. Ser. B* **18**, 1929–1967.

[32] SHORACK, G. R. AND WELLNER, J. A. (2009). *Empirical Processes with Applications to Statistics*. SIAM, Philadelphia, PA.

[33] SZNITMAN, A.-S. (1991). Topics in propagation of chaos. In *École d'Été de Probabilités de Saint-Flour XIX—1989* (Lect. Notes Math. **1464**), ed. P.-L. Hennequin. Springer, Berlin, pp. 165–251.

[34] VLASOV, A. A. (1968). The vibrational properties of an electron gas. *Sov. Phys. Uspekhi* **10**, 721.

[35] WANG, Z. (2008). Existence and uniqueness of solutions to stochastic Volterra equations with singular kernels and non-Lipschitz coefficients. *Statist. Prob. Lett.* **78**, 1062–1071.

[36] YAMADA, T. AND WATANABE, S. (1971). On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11**, 155–167.

https://doi.org/10.1017/jpr.2026.10091 Published online by Cambridge University Press