PATHWISE MCKEAN-VLASOV THEORY WITH ADDITIVE NOISE

MICHELE COGHI, JEAN-DOMINIQUE DEUSCHEL, PETER K. FRIZ, MARIO MAURELLI

Abstract. We take a pathwise approach to classical McKean-Vlasov stochastic differential equations with additive noise, as e.g. exposed in Sznitmann [38]. Our study was prompted by some concrete problems in battery modelling [23], and also by recent progrss on rough-pathwise McKean-Vlasov theory, notably Cass–Lyons [10], and then Bailleul, Catellier and Delarue [4]. Such a "pathwise McKean-Vlasov theory" can be traced back to Tanaka [40]. This paper can be seen as an attempt to advertize the ideas, power and simplicity of the pathwise appproach, not so easily extracted from [4, 10, 40], together with a number of novel applications. These include mean field convergence without a priori independence and exchangeability assumption; common noise, càdlàg noise, and reflecting boundaries. Last not least, we generalize Dawson–Gärtner large deviations and the central limit theorem to a non-Brownian noise setting.

1. Introduction

We consider the following *generalized* McKean-Vlasov stochastic differential equation (SDE) on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$,

$$\begin{cases} dX_t = b(t, X_t, \mathcal{L}(X_t))dt + dW_t \\ X_0 = \zeta. \end{cases} \tag{1.1}$$

The input data to the problem is the random initial data and noise

$$(\zeta, W): \quad \Omega \quad \rightarrow \quad \mathbb{R}^d \times C_T,$$

and

$$X : \Omega \rightarrow C_T := C([0, T], \mathbb{R}^d)$$

is the solution (process). We denote by $\mathcal{L}(Y)$ the law of a random variable Y. Classically, one takes W as a Brownian Motion. For us, it will be crucial to avoid any a priori specification of the noise. Indeed, we are not even asking for any filtration on the space Ω and equation (1.1) will be studied pathwise. For $p \in [1, \infty)$, let $\mathcal{P}_p(\mathbb{R}^d)$ be the space of probability measures on $\mathbb{R}^d$ with finite p-moment endowed with the p-Wasserstein metric. The drift is a function

$$b : [0, T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \rightarrow \mathbb{R}^d,$$

which is assumed uniformly Lipschitz continuous in the last two variables, cf. Assumption 4 below.

In a nutshell, McKean-Vlasov equations are SDEs which depend on the law of the solution. They have been extensively studied in the literature, for a comprehensive introduction we refer to [38]. They arise in many applications as limit of systems of interacting particles,

Date: September 16, 2019.

for instance in the theory of mean field games developed by Lasry and Lions [27–29]. Other interesting applications arise in fluid-dynamics [6, 21, 32], also with common noise features, neuroscience [16, 31, 41] and macroeconomics [33], also involving general driving signals. Last not least, our motivation comes from a recent battery model, cf. (1.6) below, taken from [23], which is of the form (1.1) but with reflecting boundary as given in (1.5).

Closely related to the McKean-Vlasov equation is the system of particles (classically) driven by independent Brownian motions W^i, with independent identically distributed (i.i.d.) initial conditions ζ^i,

$$\begin{cases} dX_t^{i,N} = b(t, X_t^{i,N}, L^N(X_t^{(N)}))dt + dW_t^i \\ X_0^{i,N} = \zeta^i, \end{cases} \qquad i = 1, \dots, N. \tag{1.2}$$

The particles interact with each other through the empirical measure, which is defined as follows. Given a space E (such as $\mathbb{R}^d$ or C_T) and a vector $x^{(N)} = (x^1, \dots, x^N) \in E^N$, we call $\mathcal{P}(E)$ the space of probability measures over E and we define

$$L^N(x^{(N)}) := \frac{1}{N} \sum_{i=1}^N \delta_{x^i} \in \mathcal{P}(E).$$

Let X be a solution to equation (1.1) with inputs (ζ, W) distributed as (ζ^1, W^1). When the number of particles, N, grows to infinity, we have the following a.s. convergence in $\mathcal{P}(C_T)$ equipped with the usual weak-$*$ topology,

$$L^N(X^{(N)}(\omega)) \overset{*}{\rightharpoonup} \mathcal{L}(X), \quad \text{for } \mathbb{P} - \text{a.e. } \omega. \tag{1.3}$$

This result, as well as the well-posedness of equation (1.1) is proved in [38] when the particles are exchangeable and subjects to independent inputs. This approach can be generalized to more general diffusion coefficients [13,25,26] using standard semi-martingale theory.

Rough paths: Cass and Lyons [10] study McKean-Vlasov equations in the framework of rough paths. That is, they construct (rough) pathwise solutions to the McKean-Vlasov equation driven by suitable random rough paths, which lets them go beyond the classical case when W is a semi-martingale under $\mathbb{P}$. They can treat the case multiplicative noise, that is with our dW replaced by $\sigma(X)d\mathbf{W}$, but with mean field dependence only in the drift. This problem is revisited by Bailleul [3] in the case of a Lipschitz dependence of b on the measure. Finally, Bailleul et al. [4, 5] study the general case when both b and σ are (Lipschitz) dependent on the law of the solution. This requires extra assumptions of differentiability with respect to the measure argument. The rough path case is technically more involved, it especially requires more care when studying the mean-field convergence since the solution map $(\mathcal{L}(\zeta, W) \mapsto \mathcal{L}(X))$ is continuous, but not Lipschitz (cf. [5, Rmk 4.4.]), in contrast our Lipschitz estimate in Theorem 7 below. For a different approach to rough differential equations with common noise, we refer also [14].

Tanaka: As already mentioned, in the context of battery modelling with additive noise [23], no rough path machinery is necessary, leave alone some formidable difficulties for rough differential equations to deal with reflecting boundaries [1, 18]. This was the initial motivation for our pathwise study, which soon turned out informative and rather pleasing in the generality displayed here. As our work neared completion we realized that we were not the first to go in this direction: the basic idea can be found (somewhat hidden) in a paper by Tanaka, [40, Sec.2]. (There is no shortage of citations to [40], but we are unaware of any particular work that makes use of the, for us, crucial Section 2 in that paper.) May that be as it is, advertising this aspect of Tanaka's work, as pathwise ancestor to [3,4,10], is another goal of this note, and in any case there is no significant overlap of our results with [40].

The main intuition of Tanaka [40] and subsequent works is that equation (1.2) can be interpreted as equation (1.1) by using a transformation of the probability space and the input data. We explain this connection between the equations in Section 3.1. This approach makes it possible to reduce the study of the mean field limit to a stability result for equation (1.1). This implies in particular that there is no need for asymptotical independence or exchangeability of the particles in order to obtain convergence (1.3). Indeed, one can show that the solution map

$$\mathcal{L}(\zeta, W) \mapsto \mathcal{L}(X)$$

that associates the law of the solution to the law of the inputs is continuous, and as soon as there is convergence for the law of the input data there is also convergence for the law of the solution. No independence, nor identical distributions (or even exchangeability) for the inputs are required, as we explain in Sections 3.1 and 3.3.

Main ideas. Given a Polish space E, we work on the space of probability measures with finite p-th moment, $\mathcal{P}_p(E)$, endowed with the Wasserstein distance $\mathcal{W}_p$ (see Section 1.1 for the precise definition). The idea is to construct the solution map of equation (1.1), for a generic probability measure μ,

$$\Phi : \mathcal{P}_p(\mathbb{R}^d \times C_T) \times \mathcal{P}_p(C_T) \to \mathcal{P}_p(C_T), \qquad (\mathcal{L}(\zeta, W), \mu) \mapsto \mathcal{L}(X^\mu). \tag{1.4}$$

Here X^μ is the pathwise solution to equation (1.1) when the inputs are ζ, W and the measure in the drift is given as μ, instead of the law of X. Existence and uniqueness of the solutions of the McKean-Vlasov equation (1.1) follow as a fixed point argument of the parameter dependent map Φ. Indeed, one can prove that, for fixed (ζ, W), the map $\Phi(\mathcal{L}(\zeta, W), \cdot)$ is a contraction on the space $\mathcal{P}_p(C_T)$. Hence, there is a unique fixed point $\bar{\mu} := \bar{\mu}(\mathcal{L}(\zeta, W)) = \Phi(\mathcal{L}(\zeta, W), \bar{\mu})$. This fixed point uniquely determines a pathwise solution $X^{\bar{\mu}}$ to equation (1.1).

Since Φ is Lipschitz continuous in all its arguments, it follows from Proposition 6 that also the map that associates the parameter to the fixed point, namely Ψ defined in (2.7) is Lipschitz continuous. This is the stability result that we need in order to prove convergence of the particle system.

Main results. In this setting, we obtain the following list of results.

Theorem (see Theorem 7)**.** *Let $p \in [1,\infty)$ and assume b Lipschitz. For $i = 1, 2$, let $(\zeta^i, W^i) \in L^p(\mathbb{R}^d \times C_T, \mathbb{P}^i)$ be two sets of input data. There exist unique pathwise solutions $X^i \in L^p(C_T)$ to equation* (1.1)*, driven by the respective input data. Moreover,*

$$\mathcal{W}_p(\mathcal{L}(\zeta^1, W^1, X^1), \mathcal{L}(\zeta^2, W^2, X^2)) \leq C\mathcal{W}_p(\mathcal{L}(\zeta^1, W^1), \mathcal{L}(\zeta^2, W^2)),$$

for some constant $C = C(p, T, b) > 0$.

We obtain a similar results for the case when the driver W is a random variable over the càdlàg space D_T.

Theorem (see Lemmas 19 and 20)**.** *Assume b Lipschitz and bounded. For every $(\zeta, W) : \Omega \to \mathbb{R}^d \times D_T$ measurable, there exists a unique pathwise solution $X : \Omega \to D_T$ to equation* (1.1) *driven by (ζ, W). Moreover, the map*

$$\Phi : \mathcal{P}(\mathbb{R}^d \times D_T) \to \mathcal{P}(D_T), \qquad \mathcal{L}(\zeta, W) \mapsto \mathcal{L}(X),$$

is continuous with respect to the weak topology.

We note that, in the case of jump processes, we have only weak continuity of the law of the solution with respect to the law of the inputs. We don't prove Lipschitz continuity with respect to the stronger Wasserstein norm $\mathcal{W}_p$.

As application off the main result, we have

Corollary (see Theorem 21)**.** *Consider the N-particle system (1.2) with (not necessarily Brownian! not necessarily independent!) random driving noise $W^{(N)} := (W^{1,N}, \dots, W^{N,N})$ and initial data $\zeta^{(N)} := (\zeta^{1,N}, \dots, \zeta^{N,N})$. Assume convergence (in p-Wasserstein sense) of the empirical measure*

$$L^N(\zeta^{(N)}(\omega), W^{(N)}(\omega)) \to \nu \in \mathcal{P}_p(\mathbb{R}^d \times C_T)$$

for a.e. ω (resp. in probability) w.r.t. $\mathbb{P}$. Then the empirical measure $L^N(X^{(N)})$ of the particle system converges in the same sense and the limiting law is characterized by a generalized McKean-Vlasov equation, with input data distributed like ν.

Natural non-i.i.d. situations arises in presence of common noise, cf. Section 3.3, or in the presence of heterogeneous inputs, cf. Section 3.4. In an i.i.d. setting, the required assumption is (essentially trivially) verified by the law of large number. Independent driving fractional Brownian motions, for instance, are immediately covered. Another consequence concerns the large deviations.

Definition 1. Let E be a Polish space and $(\mu^N)_{N\in\mathbb{N}}$ a sequence of Borel probability measures on E. Let $(a_N)_{N\in\mathbb{N}}$ be a sequence of positive real numbers with $\lim_{N\to\infty} a_N = \infty$. Given a lower semicontinuous function $I : E \to [0,\infty]$, the sequence μ^N is said to satisfy a *large deviations principle* with rate I if, for each Borel measurable set $A \subset E$,

$$-\inf_{x\in A^\circ} I(x) \leq \liminf a_N^{-1} \log(\mu^N(A)) \leq \limsup_{N\to\infty} a_N^{-1} \log(\mu^N(A)) \leq -\inf_{x\in \bar{A}} I(x).$$

Here A° is the interior of A and $\bar{A}$ its closure. Moreover, if the sublevel sets of I are compact, then I is said to be a *good* rate function.

We say that a sequence of random variables $(X^N)_{N\in\mathbb{N}}$ on E satisfies a large deviations principle, if the sequence of the distributions $(\mathcal{L}(X^N))_{N\in\mathbb{N}}$ does.

The following generalizes a classical result of Dawson–Gärtner [15], see also Deuschel et al. [17].

Corollary (see Theorem 34)**.** *In the i.i.d. case, the empirical measure $L^N(X^{(N)})$ satisfies a large deviations principle with rate function, defined on a suitable Wasserstein space over C_T,*

$$\mu \mapsto H(\mu \mid \Phi(\mathcal{L}(\zeta, W), \mu)),$$

where H is the relative entropy and Φ is introduced below.

This result is consistent with the one obtained in [40, Theorem 5.1], for the case of drivers given as i.i.d. Brownian motions.

One can easily drop the i.i.d. assumption, and replace H by an "assumed" large deviations principle I for the convergence of the input laws. In this case the outputs satisfy a large deviations principle.

Corollary (see Lemma 32)**.** *If the empirical measure of the inputs $L^N(\zeta^{(N)}, W^{(N)})$ satisfies a large deviations principle with (good) rate function I, then the empirical measure $L^N(X^{(N)})$ satisfies a large deviations principle with (good) rate function $\mu \mapsto I(f^\mu_\# \mu)$, defined on a suitable Wasserstein space over C_T. Here f^μ is defined in* (4.1)*.*

Think of f^μ as the function that reconstruct the inputs (initial condition, driving path) from the solution of an ordinary differential equation (ODE).

Moreover, we study the fluctuations of the empirical measure. We can prove the following central limit theorem type of result

Corollary (see Corollary 43)**.** *Let φ be a test function. Assume that the drift b is differentiable in both the spacial and the measure variable (see Assumption 38). The following converges in distribution to a Gaussian random variable, as $N \to \infty$,*

$$Y^N := \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}\varphi(X^{i,N}) - \mu(\varphi)\right).$$

The method presented here can be also applied to SDE defined in a domain $D \subset \mathbb{R}^d$, assumed to be a convex polyhedron for simplicity, and with reflection at the boundary. We consider the generalized McKean-Vlasov Skorokhod problem

$$\begin{cases} dX_t = b(t, X_t, \mathcal{L}(X_t, k_t))dt + dW_t - dk_t, \quad X_0 = \zeta, \\ d|k|_t = 1_{X_t\in\partial D} d|k|_t, \quad dk_t = n(X_t)d|k|_t. \end{cases} \tag{1.5}$$

We have the following:

Theorem (see Theorem 47)**.** *Let $p \in [1,\infty)$ and assume b Lipschitz. For $i = 1, 2$, let $(\zeta^i, W^i) \in L^p(\bar{D} \times C_T, \mathbb{P}^i)$ be two sets of input data. Then there exist unique pathwise solutions (X^i, k^i) to the generalized McKean-Vlasov Skorokhod problem* (1.5)*, driven by the respective input data. Moreover,*

$$\mathcal{W}_p(\mathcal{L}(\zeta^1, W^1, X^1, k^1), \mathcal{L}(\zeta^2, W^2, X^2, k^2)) \le C\mathcal{W}_p(\mathcal{L}(\zeta^1, W^1), \mathcal{L}(\zeta^2, W^2)).$$

with $C = C(p, T, b) > 0$.

Battery modelling. Our initial motivation for the heterogeneous particles case comes from modeling lithium-ion batteries. The numerical simulations of [23] indicate that the capacity of the battery and its efficiency is mainly determined by the size distribution of the lithium iron phosphate particles. It is thus important to allow for the particles to be of fixed different, predetermined sizes.

Lithium-ion batteries are the most promising storage devices to store and convert chemical energy into electrical energy and vice versa. In [23] lithium-ion batteries are studied where at least one of the two electrodes stores lithium within a many-particle ensemble, for example each particle of the electrode is made of Lithium-iron-phosphate. One of the practical achievements of [23] consists of the conclusions that the capacity of the battery and its efficiency as well is dominantly determined by the size distribution of the storage particles, ranging from 20 to 1000 nanometers. The radii r^i of the particles in the battery are distributed according to a distribution $\lambda \in \mathcal{P}([20, 1000])$. However, in the numerical simulations, it leads to better accuracy to artificially choose the radii in advance, instead of randomly sample them. For instance, assume that we want to simulate 1000 particles, whose radii can be of exactly two given sizes, r_1, r_2, with equal probability. It is much more convenient to choose 500 particles of radius r_1 and 500 of radius r_2, instead of sampling them from a binomial, as this could lead to imbalanced simulations and introduce an extra source of error. For this reason it is important that the theoretical results support the use of carefully chosen radii r^i of different length, such their empirical measure converges, as the number of particles grows, to a desired distribution λ. The radii so chosen, are deterministic (hence, independent), but not identically distributed.

The dynamics of the charging/discharging process is modeled in [23] by a coupled system of SDEs for the evolution of the lithium mole fractions $Y^{i,N} \in [0, 1]$ of particles $i = 1, 2, \ldots, N$ of the particle ensemble. The evolution of $Y^{i,N}$ over a time interval $[0, T]$ is described by the following system of SDEs

$$
\begin{cases} dY_t^{i,N} = \frac{1}{\tau_i}(\Lambda_{t,Li} - \mu_{Li}(Y_t^{i,N}))dt + \sigma_i dW_t^i - dk_t^{i,N} \\ d|k_t^{i,N}| = 1_{Y_t^{i,N} \in \{0,1\}} d|k_t^{i,N}|, \quad dk_t^{i,N} = n(Y_t^{i,N}) d|k_t^{i,N}|, \qquad i = 1, \ldots, N, \\ Y_0^{i,N} = a \in [0, 1]. \end{cases} \tag{1.6}
$$

We assume that all the particle have the same amount of lithium mole fraction $a \in [0, 1]$ at time $t = 0$. In practice, this initial condition is very close to 1, when the battery is empty and very close to 0, when the battery is charged. The particles are driven by a family of independent Brownian motions $W^{(N)} := (W^i)_{1 \le i \le N}$, which account for random fluctuations that can occur within the system during charging and discharging. The quantity $\tau_i \equiv \tau(r_i)$, which is related to the relaxation time and to the particle active surface, is a function of the radius r_i of the particle. As discussed earlier, the radii can only have values in a fixed range $I := [r_{min}, r_{max}] \subset (0, \infty)$. We assume that $\tau : I \to \mathbb{R}$ and $\tau^{-1} : I \to \mathbb{R}$ is Lipschitz and bounded. We also assume that $\sigma_i = \sigma(r_i)$ for a Lipschitz and bounded function $\sigma : I \to \mathbb{R}$. The term $\mu_{Li} : \mathbb{R} \to \mathbb{R}$ is the chemical potential of the Lithium and, in this framework, it is also taken Lipschitz and bounded. The interaction between particles is encoded in the surface chemical potential

$$\Lambda_{t,Li} := \frac{\sum_{j=1}^{N}\left[V_j\dot{q}_t + \mu_{Li}(x)\frac{V_j}{\tau^j}\right]}{\int \frac{V_j}{\tau_j}},$$

where q_t is a given C^1 function characterizing the state of charge of the battery at time $t \in [0,T]$ and V_j is the volume of the j-th particle. By the assumptions on μ_{Li} and τ and the bounds on the radii, the surface chemical potential is a bounded and Lipschitz continuous function of the empirical distribution of the Lithium mole fractions and radii, $\mu^N := \frac{1}{N}\sum_{i=1}^{N}\delta_{(Y_t^i,r^i)}$.

Moreover, we impose on the particles Skorokhod-type boundary conditions, of the same type as the ones described in Section 6. We call $n(x)$ the outer unit normal vector, which, in this case, reduces to $n(x) = (-1)^{x+1}$, for $x \in \{0,1\}$. This will force the mole fraction of each particle to remain in $[0,1]$. Reflecting boundary conditions are imposed also in [19], which considers the PDE counterpart of the model (1.6) here in the case of τ_j independent of j. Those boundary condtions are similar but not identical to the boundary conditions here: in [19], the surface chemical potential $\Lambda_{t,Li}$ accounts also for mean field interactions from the boundaries, we disgard those interactions here.

Under the previous assumptions, the particle system (1.6) can be essentially treated combining the results of Theorem 47 and Corollary 27, as follows. To unify the notation to the rest of the paper, we define

$$b : [0,T]\times\mathbb{R}\times\mathbb{R}\times\mathcal{P}(\mathbb{R}\times\mathbb{R}) \to \mathbb{R}, \qquad (t,x,r,\nu) \mapsto \frac{1}{\tau(r)}(-\mu_{Li}(x) + \Lambda t, Li),$$

where (calling $V(r) = 4\pi r^3/3$ the volume of the particles of radius r),

$$\Lambda_{t,Li} = \frac{\int\left[V(r)\dot{q}_t + \mu_{Li}(x)\frac{V(r)}{\tau(r)}\right]\nu(d(x,r))}{\int \frac{V(r)}{\tau(r)}\nu(d(x,r))},$$

and we consider the following *generalized* McKean-Vlasov Skorokhod equation

$$\left\{\begin{array}{l} dX_t = b(t,X_t,R,\mathcal{L}(X_t,R))dt + \sigma(R)dW_t - dk_t, \quad X_0 = a, \\ d|k|_t = 1_{X_t\in\partial D}d|k|_t, \quad dk_t = n(X_t)d|k|_t, \end{array}\right. \tag{1.7}$$

The input data are given by $a \in [0,1]$, $R \in L^p(I)$ and $W \in L^p(C_T)$, for $p \in [1,\infty)$. The solution is a couple $(X,k) \in C_T([0,1]) \times C_T$. When $W := W^{(N)} \in L^p(\Omega_N, C_T)$ and $R = R^{(N)} = (r^1,\cdots,r^N) \in L^p(\Omega_N, C_T(I))$ (the radii are constant path in I), we recover the system (1.6). We assume that the radii r^i are sampled from a distribution $\lambda \in \mathcal{P}_p(C_T(I))$ in such a way that $L^N(R^{(N)}) \overset{*}{\rightharpoonup} \lambda$, this gives the limit process (X,k) solution to (1.7), driven by $(W,R) \in C_T \times C_T(I)$ with law $\mu_{\mathcal{W}} \otimes \lambda$ (here $\mu_{\mathcal{W}}$ is the Wiener measure). We summarize this in the following proposition.

Proposition 2. *Let $p \in [1,\infty)$ and let $(W^i)_{i\in\mathbb{N}}$ be a family of independent Brownian motions on $\mathbb{R}$. Assume $I \subset (0,\infty)$ is a closed interval and let $(r^i)_{i\in\mathbb{N}} \subset I$ be a sequence in I, such that*

$$L^N(R^{(N)}) \overset{*}{\rightharpoonup} \lambda \in \mathcal{P}_p(I).$$

Then, for every $N \in \mathbb{N}$, equation (1.6) *admits a unique solution* $(Y^{(N)}, k^{(N)}) := (Y^{i,N}, k^{i,N})_{i=1,\dots,N}$. *Moreover,*

$$L^{(N)}(Y^{(N)}, k^{(N)}) \overset{*}{\rightharpoonup} \mathcal{L}(X, k),$$

where (X, k) is a solution to equation (1.7)*, with input data $(W, R) \in C_T \times C_T(I)$ with law $\mu_{\mathcal{W}} \otimes \lambda$.*

The proof of this proposition (which we will not give in full details) follows exactly as the proof of Corollary 27 and Remark 30, with the difference that, instead of Theorem 7, one applies Theorem 47.

Structure of the paper. In Section 2 we prove the well-posedness for the *generalized* McKean-Vlasov equation (1.1). In Section 3 we present applications to classical mean field particle approximation, heterogeneous mean field and mean field with common noise as corollaries of the main result. Then, we study other (classical) asymptotic for the particles as a straightforward applications: a large deviations result in Section 4; a central limit theorem in Section 5. Finally, we adapt the result to study McKean-Vlasov equations with reflection at the boundary, see Section 6.

1.1. **Notation.** Given p in $[1, +\infty)$ and a Polish space E, with metric induced by a norm $\|\cdot\|_E$, we denote by $\mathcal{P}_p(E)$ the space of probability measures on E with finite p-moment, namely the measures μ such that

$$\int_E \|x\|_E^p d\mu(x) < +\infty.$$

For $T > 0$, we denote by $C_T(\mathbb{R}^d) := C([0, T], \mathbb{R}^d)$ (the space of continuous functions from $[0, T]$ to $\mathbb{R}^d$), endowed with the supremum norm $\|f\|_{\infty:T} := \sup_{t\in[0,T]} |f(t)|$, for $f \in C_T(\mathbb{R}^d)$. When there is no risk of confusion about the codomain, we denote the space of continuous functions by C_T. Moreover, when there is non risk of confusion about the time interval, we use the lighter notation $\|\cdot\|_\infty$. Moreover, we call $C_{T,0} = \{\gamma \in C_T \mid \gamma_0 = 0\}$, the subsets of paths that vanish at time 0.

For a domain $\bar{D}$ in $\mathbb{R}^d$, we denote by $C_T(\bar{D}) := C([0, T], \bar{D})$ (continuous functions from $[0, T]$ to $\bar{D}$), endowed with the supremum norm $\|\cdot\|_\infty$.

Given $t \in [0, T]$, the projection π_t is defined as the function $\pi_t : C_T \to \mathbb{R}^d$ as $\pi_t(\gamma) := \gamma(t)$. We define the marginal at time t of $\mu \in \mathcal{P}_p(C_T)$ as $\mu_t := (\pi_t)_{\#}\mu \in \mathcal{P}_p(\mathbb{R}^d)$. We also denote by $\mu|_{[0,t]}$ the push forward of μ with respect to the restriction on the subinterval $[0, t]$.

Given a Polish space (E, d), the p-Wasserstein metric on $\mathcal{P}_p(E)$ is defined as

$$\mathcal{W}_{E,p}(\mu, \nu)^p = \inf_{m\in\Gamma(\mu,\nu)} \iint_{E\times E} d(x, y)^p m(dx, dy), \quad \mu, \nu \in \mathcal{P}_p(E), \tag{1.8}$$

where $\Gamma(\mu, \nu)$ is the space of probability measures on $E \times E$ with first marginal equal to μ and second marginal equal to ν. We will omit the space E from the notation when there is no confusion.

We denote by $\mathcal{L}(X)$ the law of a random variable X.

We use C_p to denote constants depending only on p.

Let C_c^∞ and C^n be the set of infinitely differentiable differentiable real-valued functions of compact support defined on $\mathbb{R}^d$ and the set of n times continuously differentiable functions on $\mathbb{R}^d$ such that

$$\|\varphi\|_{C^n} := \sum_{|\alpha|\le n} \sup_{x\in\mathbb{R}^d} |D^\alpha \varphi| < +\infty.$$

Let Lip_1 be the space of Lipschitz continuous functions in C^0, such that

$$\|\varphi\|_{C^0}, \sup_{x\neq y\in\mathbb{R}^d} \frac{|\varphi(x)-\varphi(y)|}{|x-y|} \le 1.$$

For $T>0$, we denote by $D_T(\mathbb{R}^d) := D([0,T],\mathbb{R}^d)$, the space of càdlàg functions (right-continuous with left limit) from $[0,T]$ to $\mathbb{R}^d$. When there is no risk of confusion about the codomain, we denote the space of cadlag functions by D_T. For $\gamma \in D_T(\mathbb{R}^d)$, $\|\gamma\|_{\infty:T} := \sup_{t\in[0,T]} |f(t)|$. Moreover, when there is no risk of confusion about the time interval, we use the lighter notation $\|\cdot\|_\infty$. We endow D_T with the Skohorod metric, defined as follows

$$\sigma(\gamma,\gamma') = \inf\left\{\lambda\in\Lambda \mid \|\lambda\| + \|\gamma-\gamma'\circ\lambda\|_\infty\right\}, \qquad \gamma,\gamma'\in D_T, \tag{1.9}$$

where Λ is the space of strictly increasing bijections on $[0,T]$ and

$$\|\lambda\| := \sup_{s\neq t}\left|\log\left(\frac{\lambda_{s,t}}{t-s}\right)\right|, \quad \lambda\in\Lambda.$$

The space (D_T,σ) is a Polish space.

Acknowledgements. Support from the Einstein Center Berlin, ECMath Project "Stochastic methods for the analysis of lithium-ion batteries" is gratefully acknowledged. J.-D. Deuschel and P.K. Friz are partially supported by DFG research unit FOR 2402. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 683164, PI P.K.Friz.) The authors thank Nikolai Bobenko for carefully reading the paper and pointing out a mistake in the proof of Lemma 13 in an earlier version.

2. The main result

In this section we study the *generalized* McKean-Vlasov SDE on a probability space $(\Omega,\mathcal{A},\mathbb{P})$,

$$\begin{cases} dX_t = b(t,X_t,\mathcal{L}(X_t))dt + dW_t \\ X_0 = \zeta. \end{cases} \tag{2.1}$$

Here the drift $b:[0,T]\times\mathbb{R}^d\times\mathcal{P}_p(\mathbb{R}^d)\to\mathbb{R}^d$ is a given Borel function, the input to the problem is the random variable

$$(\zeta,W):\quad \Omega \quad\to\quad \mathbb{R}^d\times C_T,$$

and $X:\Omega\to C_T$ is the solution. As we will see later, the law $\mathcal{L}(X)$ of the solution depends only on the law $\mathcal{L}(\zeta,W)$, for this reason we refer also to $\mathcal{L}(\zeta,W)$ as input.

Note two differences here with respect to classical SDEs: the drift depends on the solution X also through its law and W is merely a random continuous paths; in particular, it does not have to be a Brownian motion. For these differences, it is worth giving the precise definition of solution.

Definition 3. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\zeta : \Omega \to \mathbb{R}^d$, $W : \Omega \to C_T$ be random variables on it. A solution to equation (2.1) with input (ζ, W) is a random variable $X : \Omega \to C_T$ such that, for a.e. ω, the function $X(\omega)$ satisfies the following integral equality

$$X_t(\omega) = \zeta(\omega) + \int_0^t b(s, X_s(\omega), \mathcal{L}(X_s))ds + W_t(\omega).$$

We assume the following conditions on b:

Assumption 4. *Let $p \in [1, \infty)$. The drift $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}^d$ is a measurable function and there exists a constant K_b such that,*

$$|b(t,x,\mu) - b(t,x',\mu')|^p \leq K_b \left(|x-x'|^p + \mathcal{W}_{\mathbb{R}^d,p}(\mu,\mu')^p\right),$$

$\forall t \in [0,T], x, x' \in \mathbb{R}^d, \mu, \mu' \in \mathcal{P}_p(\mathbb{R}^d)$.

Before giving the main result, we introduce some notation. For a given μ in $\mathcal{P}_p(C_T)$, we consider the SDE

$$\left\{ \begin{array}{l} dY_t^\mu = b(t, Y_t^\mu, \mu_t)dt + dW_t \\ Y_0^\mu = \zeta. \end{array} \right. \tag{2.2}$$

We have the following well-posedness result

Lemma 5. *Under Assumption 4, for every input $(\zeta, W) \in L^p(\mathbb{R}^d \times C_T)$ and $\mu \in \mathcal{P}_p(C_T)$, there exists a unique $Y^\mu \in L^p(C_T)$ which satisfies, $\forall \omega \in \Omega$,*

$$Y_t^\mu(\omega) = \zeta(\omega) + \int_0^t b(s, Y_s^\mu(\omega), \mu_s)ds + W_t(\omega).$$

Moreover, denote by

$$\begin{array}{rccc} S^\mu : & \mathbb{R}^d \times C_T & \to & C_T \\ & (x_0, \gamma) & \mapsto & S^\mu(x_0, \gamma), \end{array} \tag{2.3}$$

where $S^\mu(x_0, \gamma)$ is a solution to the ODE

$$x_t = x_0 + \int_0^t b(s, x_s, \mu_s)ds + \gamma_t. \tag{2.4}$$

Then, $Y^\mu = S^\mu(\zeta, W)$.

Proof. For every couple $(x_0, \gamma) \in \mathbb{R}^d \times C_T$ the ODE (2.4) classically admits a solution $S^\mu(x_0, \gamma)$, which is continuous with respect to the inputs (x_0, γ). It is easy to verify that $S^\mu(\zeta, W)$ solves equation (2.2). We only verify that Y^μ has finite p-moments. There exists a constant $C(p, b, T)$ such that

$$\mathbb{E}\|Y^\mu\|_\infty^p \leq \mathbb{E}|\zeta|^p + C\left(1 + \int_0^T \mathbb{E} \sup_{s\in[0,t]} |Y^\mu|_\infty^p dt + \int_0^T \int_{\mathbb{R}^d} |x|^p d\mu_t(x) dt\right) + \mathbb{E}\|W\|_\infty^p.$$

We notice that $\int_{\mathbb{R}^d} |x|^p \mu_t(dx) \leq \int_{C_T} \|\gamma\|_\infty^p d\mu(\gamma) < +\infty$. Gronwall's inequality and the assumptions on (ζ, W) conclude the proof. □

We call

$$\begin{array}{rccc} \Phi: & \mathcal{P}_p(\mathbb{R}^d \times C_T) \times \mathcal{P}_p(C_T) & \to & \mathcal{P}_p(C_T) \\ & (\mathcal{L}(\zeta, W), \mu) & \mapsto & \mathcal{L}(Y^\mu) = (S^\mu)_\# \mathcal{L}(\zeta, W), \end{array} \tag{2.5}$$

the push forward of a probability measure $\mathcal{L}(\zeta, W)$ under the solution map S^μ defined in (2.3).

Note that X uniquely solves the McKean-Vlasov equation (2.1) with input (ζ, W), if and only if $\mathcal{L}(X)$ is a fixed point of $\Phi(\mathcal{L}(\zeta, W), \cdot)$:

- if X solves (2.1), then, by uniqueness for fixed $\mu = \mathcal{L}(X)$, $X = S^{\mathcal{L}(X)}(\zeta, W)$ $\mathbb{P}$-a.s. and so $\mathcal{L}(X)$ is a fixed point of $\Phi(\mathcal{L}(\zeta, W), \cdot)$;
- conversely, if $\mu^{\mathcal{L}(\zeta,W)}$ is a fixed point of $\Phi(\mathcal{L}(\zeta, W), \cdot)$, then $X = S^{\mu^{\mathcal{L}(\zeta,W)}}(\zeta, W)$ has finite p-moment and solves (2.1).

Hence existence and uniqueness for (2.1) in Theorem 7 follow from existence and uniqueness for fixed points of $\Phi^{\mathcal{L}(\zeta,W)}$, for any law $\mathcal{L}(\zeta, W)$.

For this reason, the main ingredient in the proof of Theorem 7 is the following general proposition, a version of the contraction principle with parameters. The proof is postponed to the appendix.

Proposition 6. *Let (E, d_E) and (F, d_F) be two complete metric spaces. Consider a function $\Phi : F \times E \to E$ with the following properties:*

1) *(uniform Lipschitz continuity) there exists $L > 0$ such that*
$$d_E(\Phi(Q, P), \Phi(Q', P')) \leq L \left[d_E(P, P') + d_F(Q, Q') \right].$$
2) *(contraction) There exist a constant $0 < c < 1$ and a natural number $k \in \mathbb{N}$ such that*
$$d_E((\Phi^Q)^k(P), (\Phi^Q)^k(P')) \leq c d_E(P, P') \quad \forall Q \in F, \ \forall P, P' \in E,$$
with $\Phi^Q(P) := \Phi(Q, P)$.

Then for every $Q \in F$ there exists a unique $P_Q \in E$ such that
$$\Phi(Q, P_Q) = P_Q.$$
Moreover,
$$\forall Q, Q' \in F, \quad d_E(P_Q, P_{Q'}) \leq \tilde{C} d_F(Q, Q'), \tag{2.6}$$
where $\tilde{C} := \left(\sum_{i=1}^k L^i \right) (1 - c)^{-1}$.

We give now the main result, from which most of the applications follow. It states well-posedness of the generalized McKean-Vlasov equation and Lipschitz continuity with respect to the driving signal.

Theorem 7. *Let $T > 0$ be fixed and let $p \in [1, \infty)$, assume Assumption 4.*

(i) *For every input $(\zeta, W) \in L^p(\mathbb{R}^d \times C_T)$, the map $\Phi^{\mathcal{L}(\zeta,W)}$ has a unique fixed point, $\mu^{\mathcal{L}(\zeta,W)}$.*

(ii) The map that associates the law of the inputs to the fixed point, namely

$$\begin{array}{rccc} \Psi: & \mathcal{P}_p(\mathbb{R}^d \times C_T) & \to & \mathcal{P}_p(C_T) \\ & \nu & \mapsto & \mu^\nu \end{array} \tag{2.7}$$

is well-defined and Lipschitz continuous.

(iii) For every input (ζ, W), there exists a unique solution X to the generalized McKean-Vlasov (2.1), given by $X = S^{\Psi(\mathcal{L}(\zeta,W))}(\zeta, W)$.

(iv) There exists a constant $\tilde{C} = \tilde{C}(p,T,b) > 0$ such that: for every two inputs (ζ^i, W^i), $i = 1,2$ (defined possibly on different probability spaces) with finite p-moments, the following is satisfied

$$\mathcal{W}_{C_T,p}(\mathcal{L}(X^1), \mathcal{L}(X^2)) \le \tilde{C}\mathcal{W}_{\mathbb{R}^d\times C_T,p}(\mathcal{L}(\zeta^1, W^1), \mathcal{L}(\zeta^2, W^2)).$$

In particular, the law of a solution X depends only on the law of (ζ, W).

Proof. The result follows from Proposition 6, applied to the spaces $E := \mathcal{P}_p(C_T)$, $F := \mathcal{P}_p(\mathbb{R}^d \times C_T)$ and the map Φ defined in (2.5), provided we verify conditions 1) and 2).

Let now $\mu \in E$ be fixed, let ν^1 and ν^2 be in $\mathcal{P}_p(\mathbb{R}^d \times C_T)$ and let m be an optimal plan on $(\mathbb{R}^d \times C_T)^2$ for these two measures. We call optimal plan a measure m that satisfies the minimum in the Wasserstein distance, see (B.1). On the probability space $((\mathbb{R}^d \times C_T)^2, m)$, we call ζ^i, W^i the r.v. defined by the canonical projections and $Y^i = S^\mu(\zeta^i, W^i)$ the solution to equation (2.2) with input (ζ^i, W^i). By definition of the Wasserstein metric, we have that

$$\mathcal{W}_{C_T,p}(\Phi(\nu^1,\mu), \Phi(\nu^2,\mu))^p = \mathcal{W}_{C_T,p}(\mathcal{L}(Y^1), \mathcal{L}(Y^2))^p \le C_p\mathbb{E}_m\|Y^1 - Y^2\|^p_{\infty:T}.$$

The right hand side can be estimated using the equation,

$$\begin{aligned} \mathbb{E}_m\|Y^1 - Y^2\|^p_{\infty:T} \le & C_p\mathbb{E}_m|\zeta^1 - \zeta^2|^p + C_p\mathbb{E}_m\|W^1 - W^2\|^p_{\infty:T} \\ & + K_bC_p\int_0^T \mathbb{E}_m\|Y^1 - Y^2\|^p_{\infty:t}dt. \end{aligned}$$

Using Gronwall's inequality we obtain

$$\begin{aligned} \mathcal{W}_{C_T,p}(\mathcal{L}(Y^1), \mathcal{L}(Y^2))^p \le & C_pe^{TK_bC_p}\left(\mathbb{E}_m|\zeta^1 - \zeta^2|^p + \mathbb{E}_m\|W^1 - W^2\|^p_{\infty:T}\right) \\ = & \tilde{L}\mathcal{W}_{\mathbb{R}^d\times C_T,p}(\nu^1,\nu^2)^p, \end{aligned} \tag{2.8}$$

where $\tilde{L} := C_pe^{TK_bC_p}$.

Let now (ζ, W) be fixed with law $\nu := \mathcal{L}(\zeta, W)$. Consider $\mu^1, \mu^2 \in E$ and call S^{μ^i}, for $i = 1,2$, the corresponding solution map as defined in (2.3) (driven by the initial datum ζ and the path W). Let $t \in [0,T]$ be fixed. Using equation (2.2) again, we get that

$$\begin{aligned} \int_{\mathbb{R}^d\times C_T} \|S^{\mu^1}(x_0,\gamma) - S^{\mu^2}(x_0,\gamma)\|^p_{\infty:t}\, d\nu(x_0,\gamma) \le & K_pC_p\int_0^t \mathcal{W}_{C_s,p}(\mu^1|_{[0,s]}, \mu^2|_{[0,s]})^pds \\ & + K_pC_p\int_0^t\int_{\mathbb{R}^d\times C_T} \|S^{\mu^1}(x_0,\gamma) - S^{\mu^2}(x_0,\gamma)\|^p_{\infty:s}\, d\nu(x_0,\gamma)ds. \end{aligned}$$

We deduce by the definition of $\Phi^\nu := \Phi(\nu, \cdot)$ and Wasserstein distance and applying Gronwall's lemma that

$$\mathcal{W}_{C_t,p}(\Phi^\nu(\mu^1)|_{[0,t]}, \Phi^\nu(\mu^2)|_{[0,t]})^p \le \int_{\mathbb{R}^d \times C_T} \|S^{\mu^1}(x_0,\gamma) - S^{\mu^2}(x_0,\gamma)\|^p_{\infty:t}\, d\nu(x_0,\gamma)$$

$$\le C_p K_b e^{tK_bC_p} \int_0^t \mathcal{W}_{C_s,p}(\mu^1|_{[0,s]}, \mu^2|_{[0,s]})^p ds. \tag{2.9}$$

Taking $t = T$, we have that

$$\mathcal{W}_{C_T,p}(\Phi^\nu(\mu^1), \Phi^\nu(\mu^2))^p \le \tilde{L}\mathcal{W}_{C_T,p}(\mu^1, \mu^2)^p. \tag{2.10}$$

With estimates (2.8) and (2.10) we have shown that Φ satisfies 1).
To prove 2), we reiterate k times the application Φ^ν and we use (2.9) to obtain

$$\begin{aligned}\mathcal{W}_{C_T,p}((\Phi^\nu)^k(\mu^1), (\Phi^\nu)^k(\mu^2))^p &\le \tilde{L}^k \int_0^T \int_0^{t_k} \cdots \int_0^{t_2} \mathcal{W}_{C_{t_1},p}(\mu^1|_{[0,t_1]}, \mu^2|_{[0,t_1]})^p dt_1 \dots dt_k \\ &\le \tilde{L}^k \mathcal{W}_{C_T,p}(\mu^1,\mu^2)^p \int_0^T \int_0^{t_k} \cdots \int_0^{t_2} dt_1 \dots dt_k \\ &\le \frac{(T\tilde{L})^k}{k!} \mathcal{W}_{C_T,p}(\mu^1,\mu^2)^p.\end{aligned}$$

By choosing $k > 0$ large enough, we have that $c := \frac{(T\tilde{L})^k}{k!} < 1$. This shows point 2) and concludes the proof. ☐

If the driving process is progressively measurable, then so is the solution:

Proposition 8. *Let $(\mathcal{F}_t)_{t\ge 0}$ be a right-continuous, complete filtration on $(\Omega, \mathcal{A}, \mathbb{P})$ such that ζ is $\mathcal{F}_0$-measurable and W is $(\mathcal{F}_t)_{t\ge0}$-progressively measurable. Then the solution X to* (2.1) *is also $(\mathcal{F}_t)_{t\ge0}$-progressively measurable.*

Proof. The proof is classical. Fix t in $[0,T]$, then, $\mathbb{P}$-a.e., the restriction $X|_{[0,t]} = X|_{[0,t]}(\omega)$ on $[0,t]$ of the solution X also solves (2.2) on $[0,t]$ with inputs ζ and $W|_{[0,t]}$ (restriction of W on $[0,t]$) and input measure $\mu|_{[0,t]}$ (pushforward of $\mu = \mathcal{L}(X)$ by the restriction on $[0,t]$). Therefore $X|_{[0,t]}(\omega) = S_t^{\mu|_{[0,t]}}(\zeta, W|_{[0,t]})$. Since $S_t^{\mu|_{[0,t]}}$ is $\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{B}(C_t)$-measurable and ζ and $W|_{[0,t]}$ are $\mathcal{F}_t$-measurable, also $X|_{[0,t]}$ is $\mathcal{F}_t$-measurable, in particular $X|_{[0,t]}$ is $\mathcal{F}_t$-measurable. Hence X is adapted and therefore progressively measurable by continuity of its paths. ☐

2.1. **Weak continuity.** In this note we are generally interested in proving quantitative convergence in the Wasserstein distance. However, one can show that the law of the solution of the mean field equation (2.1) is continuous in the weak topology of measures, with respect to the law of the inputs, in the spirit of [40].

Assumption 9. *Given a Polish space (E,d), we endow the space $\mathcal{P}(E)$ with a metric Π_E, with the following properties*

(i) The metric Π_E is complete and metrizes the weak convergence of measures.
(ii) For any two random variables $X, X' : \Omega \to E$, we have

$$\Pi_E(\mathcal{L}(X), \mathcal{L}(X')) \le \mathbb{E}d(X,X').$$

Remark 10. Let Lip_1 be the space of bounded and Lipschitz functions on E, as defined in Section 1.1. Define the Kantorovich-Rubinstein metric as

$$\Pi_E(\mu,\nu) := \sup_{\varphi\in \mathrm{Lip}_1} \int_E \varphi d(\mu-\nu),$$

This metric satisfies Assumption 9. Note that 9 (i) follows from [8, Theorem 8.3.2 and Theorem 8.9.4]

For the drift we assume the following.

Assumption 11. *The drift $b:[0,T]\times\mathbb{R}^d\times\mathcal{P}(\mathbb{R}^d)\to\mathbb{R}^d$ is a measurable function and there exists a constant K such that,*

- *(Lipschitz continuity)*
$$|b(t,x,\mu)-b(t,x',\mu')| \le K\left(|x-x'|+\Pi_{\mathbb{R}^d}(\mu,\mu')\right),$$
$\forall t\in[0,T], x,x'\in\mathbb{R}^d, \mu,\mu'\in\mathcal{P}(\mathbb{R}^d)$.
- *(boundedness)*
$$|b(t,x,\mu)|\le K,$$
$\forall t\in[0,T], x\in\mathbb{R}^d, \mu\in\mathcal{P}(\mathbb{R}^d)$.

Remark 12. Assume that there exists a function $B:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}^d$ such that there exists a constant $C>0$,

$$|B(x,y)|\le C, \quad |B(x,y)-B(x',y')|\le C\left(|x-x'|+|y-y'|\right), \quad \forall x,x',y,y'\in\mathbb{R}^d,$$

and the drift satisfies $b(t,x,\mu):=\int_{\mathbb{R}^d} B(x,y)\mu(dy)$. Then b satisfies Assumptions 11, with $K=3C$. This is the case treated in [40].

Lemma 13. *Given $\nu\in\mathcal{P}(\mathbb{R}^d\times C_T)$, the solution map*

$$\begin{array}{rccc} S^\nu: & \mathbb{R}^d\times C_T & \to & C_T \\ & (x_0,\gamma) & \mapsto & S^\nu(x_0,\gamma), \end{array} \tag{2.11}$$

to the ODE

$$x_t = x_0 + \int_0^t b(s,x_s,(x_s)_\#\nu)ds + \gamma_t. \tag{2.12}$$

is well defined.

Proof. We prove the lemma by iteration. For a fixed $x_0,\gamma\in\mathbb{R}^d\times C_T$, define $x^0_t:=x_0+\gamma_t$, and x^{n+1}_t defined implicitly as $x^{n+1}_t = x_0+\int_0^t b(s,x^{n+1}_s,(x^n_s)_\#\nu)ds+\gamma_t$. Clearly, for every $n\in\mathbb{N}$, the function $(x_0,\gamma)\mapsto x^n$ is well defined and measurable.

We compute the following, for $t\in[0,T]$, using Assumption 11, Gronwall's Lemma and Assumption 9 (ii)

$$|x^n_t - x^{n+1}_t| \le Ke^{Kt}\int_0^t \Pi_{\mathbb{R}^d}((x^{n-1}_s)_\#\nu,(x^n_s)_\#\nu)ds \le Ke^{Kt}\int_0^t\int_{\mathbb{R}^d\times C_T}|x^{n-1}_s-x^n_s|d\nu ds.$$

Iterating this inequality down to $n=0$, we obtain that there exists a positive constant $C(T,K)$, independent of n, such that

$$|x^n_t - x^{n+1}_t| \le \frac{C(T,K)^n}{n!}.$$

Hence, we have that, for every $x_0, \gamma \in \mathbb{R}^d \times C_T$, the sequence $(x^n(x_0, \gamma))_{n\geq 0}$ is Cauchy in $(C_T, \|\cdot\|_\infty)$. Indeed, for $\epsilon > 0$, there exists $m > 0$ big enough, such that for every $n \geq m$,

$$\|x^m - x^n\|_\infty \leq \sum_{i=m}^{n-1} \|x^i - x^{i+1}\|_\infty \leq \sum_{i=m}^{\infty} \frac{C(T,K)^i}{i!} < \epsilon.$$

We call $x(x_0, \gamma) \in C_T$ its limit as $n \to \infty$. The pointwise limit of Borel measurable functions is measurable, hence $(x_0, \gamma) \mapsto x$ is also measurable and $(x_s)_{\#}\nu$ is well-defined. We can thus pass to the limit in equation (2.12) to show that x is a solution to it.

To prove uniqueness, let x and y be two solutions with the same inputs x_0, γ and ν. As before, we can compute, for $t \in [0, T]$,

$$|x_t - y_t| \leq K e^{Kt} \int_0^t \int_{\mathbb{R}^d \times C_T} |x_s - y_s| d\nu ds.$$

Integrating in $d\nu(\gamma)$ and applying Gronwall's lemma we get that the righ hand side vanishes. Hence, x and y are the same for all $t \in [0, T]$ and $\gamma \in C_T$. □

Lemma 14. *The function*

$$\begin{array}{rccc} \Psi: & (\mathcal{P}(\mathbb{R}^d \times C_T), \Pi_{\mathbb{R}^d \times C_T}) & \to & (\mathcal{P}(C_T), \Pi_{C_T}) \\ & \nu & \mapsto & (S^\nu)_{\#}\nu, \end{array} \tag{2.13}$$

is continuous. By Assumption 9 (i), this is equivalent to continuity with respect to the topology induced by the weak convergence of measures.

Proof. Let $(\nu^n)_{n\geq 0} \subset \mathcal{P}(\mathbb{R}^d \times C_T)$ be a sequence of probability measures that converges weakly to $\nu \in \mathcal{P}(\mathbb{R}^d \times C_T)$. From Skohorokhod representation theorem, there exists a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a sequence $(\zeta^n, W^n) : \Omega \to \mathbb{R}^d \times C_T$ of random variables distributed as ν^n that converges almost surely to a random variable (ζ, W) distributed as ν.

Let $X^n := S^{\nu^n}(\zeta^n, W^n)$. By definition, $\mu^n := \mathcal{L}(X^n) = \Psi(\nu^n)$ and X^n solves the following SDE in the sense of Definition 3,

$$X_t^n = \zeta^n + \int_0^t b(s, X_s^n, \mathcal{L}(X_s^n)) ds + W_t^n.$$

It is easy to check that the random variables X^n are equicontinuous and equibounded and deduce that the family μ^n is tight in C_T. With an abuse of notation, assume that $(\mu^n)_{n\geq 0}$ is a subsequence that converges weakly to some $\mu \in \mathcal{P}(C_T)$, and $(X^n)_{n\geq 0}$ such that $\mathcal{L}(X^n) = \mu^n$. By using the equation, one can check that $(X^n(\omega))_{n\geq 0}$ is a Cauchy sequence in C_T for $\mathbb{P} - a.e.\ \omega$. Let X be the almost sure limit of X^n, as $n \to \infty$. Clearly, μ^n converges weakly to $\mathcal{L}(X)$, hence $\mathcal{L}(X) = \mu$. Passing to the limit in the equation, we can see that $\mu = \mathcal{L}(X) = \Psi(\nu)$. This concludes the proof. □

2.2. **Càdlàg drivers.** In this section we follow the same reasoning as Section 2.1 to study the case when the drivers are discontinuous processes in (D_T, σ). We first set some notation and recall some results about càdlàg functions.

Given $t \in [0, T]$, the projection π_t is defined, analogously to the continuous case, as the function $\pi_t : D_T \to \mathbb{R}^d$ as $\pi_t(\gamma) := \gamma(t)$.

Definition 15. For a function $\gamma \in D_T$, we define its càdlàg modulus as a function of $\delta \in (0,1)$,

$$w_\gamma(\delta) = \inf_{\Pi} \max_{1 \le i \le n} \sup_{t_{i-1} \le s \le t < t_i} |\gamma_{s,t}|,$$

where the infimum is taken over all the partitions Π with mash size bigger than δ.

Then we have the following lemma, from [7, equation (13.3)].

Lemma 16. *Let $(\nu^n)_{n\ge 0} \subset \mathcal{P}(D_T)$ be a sequence of probability measures converging weakly to $\nu \in \mathcal{P}(D_T)$, then there exists a set $T_\nu \subset [0,T]$ of full Lebesgue measure (actually T_ν^c is at most countable) such that ν_t^n converges weakly to ν_t, for all $t \in T_\nu$.*

Given a Polish space (E,d), we use once again the notation Π_E to denote a distance on $\mathcal{P}(E)$ that satisfies Assumption 9.
For the drift we assume the following.

Assumption 17. *The drift $b : \mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d) \to \mathbb{R}^d$ is a measurable function and there exists a constant K such that,*

- *(Lipschitz continuity)*
$$|b(x,\mu) - b(x',\mu')| \le K\left(|x-x'| + \Pi_{\mathbb{R}^d}(\mu,\mu')\right),$$
$\forall x, x' \in \mathbb{R}^d, \mu, \mu' \in \mathcal{P}(\mathbb{R}^d)$.
- *(boundedness)*
$$|b(x,\mu)| \le K,$$
$\forall x \in \mathbb{R}^d, \mu \in \mathcal{P}(\mathbb{R}^d)$.

Remark 18. The function b defined in Remark 12 also satisfies Assumptions 17.

2.2.1. *Well-posedness and continuity.* We have the following results, analogously to Section 2.1.

Lemma 19. *Let b satisfy Assumptions 17. Given $\nu \in \mathcal{P}(\mathbb{R}^d \times D_T)$, the solution map*

$$\begin{array}{cccc} S^\nu : & \mathbb{R}^d \times D_T & \to & D_T \\ & (x_0,\gamma) & \mapsto & S^\nu(x_0,\gamma), \end{array} \tag{2.14}$$

to the ODE

$$x_t = x_0 + \int_0^t b(x_s, (x_s)_\# \nu) ds + \gamma_t. \tag{2.15}$$

is well defined.

Proof. The proof of this lemma follows exactly the proof of Lemma 13. We define the sequence $(x^n)_{n\in\mathbb{N}} \subset D_T$ and show that it is a Cauchy sequence in the uniform norm $\|\cdot\|_\infty$. By taking $\lambda(t) = t$ in the definition of σ, equation (1.9), one notices immediately that σ is bounded by the distance induced by $\|\cdot\|_\infty$. Hence, $(x^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in σ and the conclusion follows as in Lemma 13. □

Lemma 20. *The function*

$$\Psi: \quad (\mathcal{P}(\mathbb{R}^d \times D_T), \Pi_{\mathbb{R}^d\times D_T}) \;\to\; (\mathcal{P}(D_T), \Pi_{D_T}) \qquad \nu \;\mapsto\; (S^\nu)_\# \nu, \tag{2.16}$$

is continuous. By assumption, this is equivalent to continuity with respect to the topology induced by the weak convergence of measures.

Proof. Let $(\nu^n)_{n\geq 0} \subset \mathcal{P}(\mathbb{R}^d \times D_T)$ be a sequence of probability measure that converges weakly to $\nu \in \mathcal{P}(\mathbb{R}^d \times D_T)$. From Skohorokhod representation theorem, there exists a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a sequence $(\zeta^n, W^n) : \Omega \to \mathbb{R}^d \times D_T$, for $n \geq 0$, of random variables distributed as ν^n which converges almost surely to a random variable (ζ, W) distributed as ν.

Let $X^n := S^{\nu^n}(\zeta^n, W^n)$. By definition, $\mu^n := \mathcal{L}(X^n) = \Psi(\nu^n)$ and X^n solves the following SDE pathwise,

$$X^n_t = \zeta^n + \int_0^t b(X^n_s, \mathcal{L}(X^n_s))ds + W^n_t.$$

By construction, the laws of W^n are tight. Equivalently, by [7, Theorem 13.2], they satisfy

$$\lim_{a\to\infty} \limsup_{n\to\infty} \mathbb{P}\{\|W^n\|_\infty \geq a\} = 0, \tag{2.17}$$

$$\forall \epsilon > 0, \qquad \lim_{\delta\to 0} \limsup_{n\in\mathbb{N}} \mathbb{P}\{w_{W^n}(\delta) \geq \epsilon\} = 0. \tag{2.18}$$

It follows from Assumption 17 that the random variables X^n also satisfy (2.17) and (2.18). Thus, we deduce that the family μ^n is tight in $\mathcal{P}(\mathbb{R}^d \times D_T)$.

With an abuse of notation, assume that $(\mu^n)_{n\geq 0}$ is a subsequence that converges weakly to some $\mu \in \mathcal{P}(D_T)$, and $(X^n)_{n\geq 0}$ such that $\mathcal{L}(X^n) = \mu^n$. By using the equation, we now check that $(X^n(\omega))_{n\geq 0}$ is a Cauchy sequence in (D_T, σ) for $\mathbb{P} - a.e.\ \omega$.

First observe that Lemma 16 and Lebesgue dominated convergence imply that $\int_0^T \Pi_{\mathbb{R}^d}(\mu^n_s, \mu_s)ds \to 0$, as $n \to \infty$. Hence $(\mu^n)_{n\in\mathbb{N}} \subset L^1([0,T], \mathcal{P}(\mathbb{R}^d))$ is a Cauchy sequence. Let now $\Omega^0 \subset \Omega$ be a set of full measure such that $(\zeta^n(\omega), W^n(\omega)) \to (\zeta(\omega), W(\omega))$, for all $\omega \in \Omega^0$, as $n \to \infty$.

Fix $\omega \in \Omega^0$, $\epsilon > 0$ there exists $N > 0$, such that for all $m, n \geq N$, we have

$$\sigma(W^n(\omega), W^m(\omega)) < \epsilon, \quad |\zeta^n(\omega) - \zeta^m(\omega)| < \epsilon, \quad \int_0^T \Pi_{\mathbb{R}^d}(\mu^n_s, \mu^m_s)ds < \epsilon.$$

Moreover, since the sequence $(W^n(\omega))$ converges, it is pre-compact in D_T. It follows from [7, Theorem 12.3] that

$$\lim_{\delta\to 0} \sup_n w(W^n(\omega), \delta) = 0. \tag{2.19}$$

It follows from Assumption 17 that $|X^n_{s,t}(\omega)| \leq |W^n_{s,t}(\omega)| + K|t-s|$, for all $t, s \in [0,T]$ and $n \in \mathbb{N}$. Hence, one can replace W^n with X^n in (2.19). We omit now the dependence of the random variables from ω. There exists $\bar{\delta} > 0$, such that for every $0 < \delta < \bar{\delta}$, $\sup_n w(X^n, \delta) < \epsilon$. We can choose $\delta = \epsilon \wedge \bar{\delta}$, and $\lambda := \lambda(\omega, \delta, m, n)$ such that

$$\|\lambda\| + \|W^n - W^m \circ \lambda\|_\infty < \delta < \epsilon. \tag{2.20}$$

It follows from [7, equation (12.17)] that, for all $t \in [0,T]$,

$$|\lambda_t - t| \le e^{\|\lambda\|} - 1 \le e^{\delta} - 1 \approx \delta,$$

Hence, for any $t \in [0,T]$ and for any partition Π of $[0,T]$ of mesh size bigger than δ we find at most one point of the partition between t and λ_t, which gives

$$\|X^m - X^m \circ \lambda\|_\infty < 2w(X^m,\delta) < 2\sup_n w(X^n,\delta) < 2\epsilon.$$

We note that, for all $t \in [0,T]$,

$$\left|\int_{\lambda_t \vee t}^{\lambda_t \wedge t} ds\right| \le |\lambda_t - t| \le e^{\|\lambda\|} - 1,$$

where the last inequality follows from [7, equation (12.17)]. We can thus compute the following, for $t \in [0,T]$,

$$\begin{aligned}
\left|\int_0^t b(X_s^n,\mu_s^n)ds - \int_0^{\lambda_t} b(X_s^m,\mu_s^m)ds\right| \le & K\left|\int_{\lambda_t \vee t}^{\lambda_t \wedge t} ds\right| + \int_0^{\lambda_t \wedge t} |b(X_s^n,\mu_s^n) - b(X_s^m,\mu_s^m)|\, ds \\
\le & K(e^{\|\lambda\|}-1) + K\int_0^T \Pi_{\mathbb{R}^d}(\mu_s^n,\mu_s^m)ds + K\int_0^T |X_s^n - X_s^m| ds \\
\le & K(e^{\|\lambda\|}-1) + K\epsilon + K\int_0^T |X_s^n - (X^m\circ\lambda)_s| ds \\
& + K\int_0^T |X_s^m - (X^m\circ\lambda)_s| ds \\
\lesssim & 4\epsilon + \int_0^T |X_s^n - (X^m\circ\lambda)_s| ds
\end{aligned}$$

From which we deduce

$$\begin{aligned}
|X_t^n - (X^m\circ\lambda)_t| \le & |\zeta^n - \zeta^m| + |W_t^n - (W^m\circ\lambda)_t| + \left|\int_0^t b(X_s^n,\mu_s^n)ds - \int_0^{\lambda_t} b(X_s^m,\mu_s^m)ds\right| \\
\lesssim & \epsilon + \|W^n - W^m\circ\lambda\| + \int_0^T |X_s^n - (X^m\circ\lambda)_s| ds
\end{aligned}$$

We add $\|\lambda\|$ on both sides, apply Gronwall's Lemma and inequality (2.20) to obtain

$$\sigma(X^n(\omega), X^m(\omega)) < C(T,K)\epsilon.$$

Hence, we have that $X^n(\omega)$ is a Cauchy sequence in (D_T,σ), for $\omega \in \Omega^0$.

Let X be the almost sure limit of X^n, as $n \to \infty$. The laws μ^n converge weakly to $\mathcal{L}(X)$, hence $\mathcal{L}(X) = \mu$. Passing to the limit in the equation, we can see that $\mu = \mathcal{L}(X) = \Psi(\nu)$. This concludes the proof. □

3. Applications

3.1. **Particle approximation.** In this section we show how the results in Section 2 yield a convergence result for a particle system associated with the McKean-Vlasov equation.

Given inputs $\bar\zeta$ and $\bar W$ (on a probability space $(\Omega,\mathcal{A},\mathbb{P})$), we consider the following McKean-Vlasov equation

$$\begin{cases} d\bar X_t = b(t,\bar X_t,\mathcal{L}(\bar X_t))dt + d\bar W_t \\ X_0 = \bar\zeta. \end{cases} \tag{3.1}$$

To this, given $N\in\mathbb{N}$, we associate the corresponding interacting particle system (on a probability space $(\Omega,\mathcal{A},\mathbb{P})$),

$$\begin{cases} dX_t^{i,N} = b(t,X_t^{i,N},\frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}})dt + dW_t^{i,N}, \\ X_0^{i,N} = \zeta^{i,N}, \end{cases} \qquad i=1,\dots,N \tag{3.2}$$

with given input

$$\begin{array}{rccc} (\zeta^{(N)},W^{(N)}): & \Omega & \to & (\mathbb{R}^d\times C_T)^N \\ & \omega & \mapsto & (\zeta^{i,N}(\omega),W^{i,N}(\omega))_{1\le i\le N}. \end{array}$$

For a given $N\in\mathbb{N}$ and an N-dimensional vector $Y^{(N)}=(Y^1,\cdots,Y^N)$ with entries in a Polish space E, we define the empirical measure associated with $Y^{(N)}$ as

$$L^N(Y^{(N)}) := \frac{1}{N}\sum_{i=1}^N \delta_{Y^i}.$$

As pointed out in the introduction, the main argument of Cass-Lyons/Tanaka approach is that the particle system (1.2) can be interpreted as the limiting McKean-Vlasov equation (1.1) by using a transformation of the probability space and the input data. The main result Theorem 7 not only implies well-posedness of both McKean-Vlasov and particle approximation, but also allows to deduce convergence of the particle system from convergence of the corresponding signals, something which is usually easy to verify, for example, if the signals are empirical measures of independent noises.

Now we show how to interpret equations (3.1) and (3.2) as generalized McKean-Vlasov equation (2.1). Clearly (3.1) is (2.1) with inputs $\bar\zeta$ and $\bar W$. For (3.2), for fixed $N\in\mathbb{N}$, we consider the space $(\Omega_N,\mathcal{A}_N,\mathbb{P}_N)$, where $\Omega_N := \{1,\dots,N\}$, $\mathcal{A}_N := 2^{\Omega_N}$ and $\mathbb{P}_N := \frac{1}{N}\sum_{i=1}^N \delta_i$. On this space, we can identify any N-uple $Y^{(N)}=(Y^1,\dots,Y^N)\in E^N$, as a random variable $\Omega_N\ni i\mapsto Y^i\in E$. With this identification, the law of $Y^{(N)}$ on Ω_0 is precisely the empirical measure associated with $Y^{(N)}$, namely $L^N(Y^{(N)})$. Indeed, for each continuous and bounded function φ on E, we have

$$\mathbb{E}_{\mathbb{P}_N}[\varphi(Y^{(N)})] = \sum_{i=1}^N \frac{1}{N}\varphi(Y^i) = L^N(Y^{(N)})(\varphi).$$

We assume that $(\zeta^{(N)}(\omega),W^{(N)}(\omega))$ is valued in $(\mathbb{R}^d\times C_T)^N$ for every N and for every $\omega\in\Omega$. We fix $\omega\in\Omega$ and N and we apply the previous argument to the N-uples

$$\begin{aligned} (\zeta^{(N)},W^{(N)})(\omega) &= ((\zeta^{1,N},W^{1,N})(\omega),\dots,(\zeta^{N,N},W^{N,N})(\omega)), \\ X^{(N)}(\omega) &= (X^{1,N}(\omega),\dots,X^{N,N}(\omega)). \end{aligned}$$

For fixed $\omega\in\Omega$, the law of $(\zeta^{(N)}(\omega),W^{(N)}(\omega))$ on Ω_N is the empirical measure $L^N(\zeta^{(N)},W^{(N)})(\omega)$ and the law of $X^{(N)}(\omega)$ on Ω_N is the empirical measure $L^N(X^{(N)})(\omega)$, which appears exactly in (3.2), projected at time t. Hence, for fixed ω in Ω, the interacting particle system

(3.2) is the generalized McKean-Vlasov equation (2.1), defined on the space $(\Omega_N, \mathcal{A}_N, \mathbb{P}_N)$ and driven by the empirical measure $L^N(\zeta^{(N)}, W^{(N)})(\omega)$.

We are ready to apply Theorem 7 to obtain the following result, which ties the convergence of the particles to the convergence of the inputs. An immediate consequence is that the empirical measure of the particle system converges if the input converges: no independence or exchangeability are required.

Theorem 21. *Let $p \in [1, \infty)$ and assume 4. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. For a fixed $N \in \mathbb{N}$, let $(\zeta^{(N)}, W^{(N)}) = (\zeta^{i,N}, W^{i,N})_{1 \le i \le N} : \Omega \to (\mathbb{R}^d \times C_T)^N$ be a family of random variables. Let $\bar{\zeta} \in L^p(\Omega, \mathbb{R}^d)$ and $\bar{W} \in L^p(\Omega, C_T)$. Then,*

i for every $\omega \in \Omega$, there exists a unique pathwise solution $X^{(N)}(\omega)$ in the sense of Definition 3 to the interacting particle system (3.2). Moreover, $\omega \mapsto X^{(N)}(\omega)$ is $\mathcal{A}$-measurable.
ii there exists a unique pathwise solution $\bar{X}$ in the sense of Definition 3 to equation (3.1).
iii there exists a constant C depending on b such that for all $N \ge 1$, for $\mathbb{P}$-a.e. $\omega \in \Omega$,

$$\mathcal{W}_{C_T,p}(L^N(X^{(N)}(\omega)), \mathcal{L}(\bar{X}))^p \le C \mathcal{W}_{\mathbb{R}^d \times C_T, p}(L^N(\zeta^{(N)}(\omega), W^{(N)}(\omega)), \mathcal{L}(\bar{\zeta}, \bar{W}))^p. \tag{3.3}$$

Proof. Let $N \in \mathbb{N}$. Fix $\omega \in \Omega$, we apply Theorem 7 in the following setting

$$(\Omega^1, \mathcal{A}^1, \mathbb{P}^1) := (\Omega_N, \mathcal{A}_N, \mathbb{P}_N), \quad (\zeta^1, W^1)(\omega) := (\zeta^{(N)}(\omega), W^{(N)}(\omega)),$$

$$(\Omega^2, \mathcal{A}^2, \mathbb{P}^2) := (\Omega, \mathcal{A}, \mathbb{P}), \quad (\zeta^2, W^2) := (\bar{\zeta}, \bar{W}).$$

The finite p-moment condition is satisfied by $(\bar{\zeta}, \bar{W})$ by assumption and also by $(\zeta^{(N)}(\omega), W^{(N)}(\omega))$, since

$$\begin{aligned} \|(\zeta^1, W^1)(\omega)\|^p_{L^p(\Omega^1)} =& \mathbb{E}_{\mathbb{P}_N}\left[|\zeta^{(N)}(\omega)|^p + \|W^{(N)}(\omega)\|^p_\infty\right] \\ =& \frac{1}{N}\sum_{i=1}^N |\zeta^i(\omega)|^p + \frac{1}{N}\sum_{i=1}^N \|W^i(\omega)\|^p_\infty < +\infty. \end{aligned}$$

Since the assumptions on the drift b are also satisfied, Theorem 7 establishes the existence of solutions $X^1(\omega) =: X^{(N)}(\omega)$ and $X^2 =: \bar{X}$. Moreover the map Ψ is continuous, hence $\omega \mapsto L^{(N)}(X^{(N)})(\omega)$ is $\mathcal{A}$-measurable, which makes $X^{(N)}(\omega) := S^{L^{(N)}(X^{(N)})(\omega)}(\zeta(\omega), W^{(N)}(\omega))$ measurable. This gives (i) and (ii). Theorem (7) also gives exactly the inequality in (iii). The proof is complete. □

Remark 22. We stress out that, when looking at the particle system, we are applying Theorem 7 on the discrete space, for a fixed ω, and the law that appears on the drift is the empirical measure at fixed ω.

Remark 23. In the proof of point iii of Theorem 21, we can actually get the bound for *every* ω if we use the pathwise solution $X^{(N)}(\omega)$ (in the sense of Definition 3), as this satisfies (3.2) for *every* ω. However, the "$\mathbb{P}$-a.s." is required when dealing with a solution to the interacting particle system (3.2) in the usual probabilistic sense, where (3.2) is required to hold only $\mathbb{P}$-a.s..

3.2. **Classical mean field limit.** Now we specialize the previous result in the case of i.i.d. inputs, recovering the classical result by Sznitman [38]:

Corollary 24. *Given a filtered probability space $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ (with the standard assumptions) and $p \in (1,\infty)$ let $(\zeta^i)_{i\geq 1} \subset L^p(\Omega, \mathbb{R}^d)$, be a family of i.i.d. random variables which are $\mathcal{F}_0$-measurable and $(W^i)_{i\geq 1}$ be a family of independent adapted Brownian motions. Moreover, let $(\bar{\zeta}, \bar{W}) \in L^p(\Omega, \mathbb{R}^d \times C_T)$ be an independent copy of (ζ^1, W^1). Then the solutions $X^{(N)}$ and $\bar{X}$ to the interacting particles system* (3.2) *and the McKean-Vlasov SDE* (3.1)*, respectively, given by Theorem 21, are progressively measurable and we have the following convergence*

$$L^N(X^{(N)}) \overset{*}{\rightharpoonup} \mathcal{L}(\bar{X}), \quad \mathbb{P} - a.s. \tag{3.4}$$

Remark 25. The classical case when b is a convolution with a regular kernel, say $b(t,x,\mu) = (K * \mu)(x)$, is treated here, as b in this case satisfies the assumption of Theorem 21.

Proof of Corollary 24. Progressive measurability for the particle system (3.2) follows from (7 (ii)) of Theorem 7 and is a consequence of Proposition 8 for the McKean-Vlasov SDE (3.1).

We prove now the convergence. First recall that Theorem 21, and in particular inequality (3.3), applies in this case. Hence, if we can prove that the right-hand-side of (3.3) goes to zero, we have the desired convergence (3.4).

Hence, by Lemma 54, we deduce the convergence in p'-Wasserstein, for every $p' \in (1,p)$. This is the convergence of the right-hand-side of (3.3). The proof is complete. ☐

3.3. **Mean field with common noise.** In this section we study a system of interacting particles with common noise. We consider the following system on the space $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$,

$$\begin{cases} dX_t^{i,N} = b(t, X_t^{i,N}, \frac{1}{N}\sum_{i=1}^N \delta_{X_t^{i,N}})dt + dW_t^i + dB_t \\ X_0^{i,N} = \zeta^i. \end{cases} \quad i = 1,\dots,N \tag{3.5}$$

Here $(\zeta^i)_{i=1,\dots,N} \subset L^p(\bar{\Omega}, \mathbb{R}^d)$ is a family of i.i.d. random variables. This system represents N interacting particles where each particle is subject to the interaction with the others as well as some randomness. There are two sources of randomness, one which acts independently on each particle and is represented by the independent family of identically distributed random variables $W^{(N)} = (W^i)_{1\leq i\leq N} \subset L^p(\bar{\Omega}, C_T)$. The second source of randomness is the same for each particle and is represented by the random variable $B \in L^p(\bar{\Omega}, C_T)$, which is assumed to be independent from the W^i. Usually W^i and B are Brownian motions, but it is not necessary to assume it here. The Brownian motion case was considered in [12].

Our aim is to prove that the empirical measure associate to the system converges, as $N \to \infty$, to the conditional law, given B, of the solution of the following McKean-Vlasov SDE

$$\begin{cases} d\bar{X}_t = b(t, \bar{X}_t, \mathcal{L}(\bar{X}_t|B))dt + d\bar{W}_t + dB_t \\ \bar{X}_0 = \bar{\zeta}. \end{cases} \tag{3.6}$$

Here $\bar{\zeta}$ is a random variable on $\mathbb{R}^d$ and $\bar{W}$ is random variables on C_T distributed as ζ^1 and W^1 respectively. We denote by $\mathcal{L}(X|B)$ the conditional law of X given B. Our result is the following.

Corollary 26. *Let $p \in [1, \infty)$, $p' \in (p, \infty)$, and assume 4. Let $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}})$ be a probability space. On this space we consider independent families $\zeta^{(N)} = (\zeta^i)_{1\le i\le N} \subset L^{p'}(\bar{\Omega}, \mathbb{R}^d)$, $W^{(N)} = (W^i)_{1\le i\le N} \in L^{p'}(\bar{\Omega}, C_T)$ of i.i.d. random variables. Let $\bar{\zeta}$ be distributed as $\zeta^{i,N}$ and let $\bar{W}$ be distributed as $W^{i,N}$ and independent of $\bar{\zeta}$. Moreover, assume that $B \in L^p(\bar{\Omega}, C_T)$ is a random variable independent from the others. Then there exists a solution $X^{(N)} \in L^p(\bar{\Omega}, (C_T)^N)$ to equation (3.5) and a solution $\bar{X} \in L^p(\bar{\Omega}, C_T)$ to equation (3.6). Moreover, we have*

$$\mathcal{W}_{C_T,p}\left(L^N(X^{(N)}), \mathcal{L}(\bar{X}|B)\right) \to 0, \quad \bar{\mathbb{P}} - a.s. \quad as\ N \to \infty.$$

Proof. Since B is independent from the other variables, we can assume, without loss of generality, that our probability space is of the form $(\bar{\Omega}, \bar{\mathcal{A}}, \bar{\mathbb{P}}) := (\Omega \times \Omega', \mathcal{A} \otimes \mathcal{A}', \mathbb{P} \otimes \mathbb{P}')$, that the random variables $\zeta^i, \bar{\zeta}, W^i$ and $\bar{W}$ are defined on a space $(\Omega, \mathcal{A}, \mathbb{P})$ and the random variable B is defined on the space $(\Omega', \mathcal{A}', \mathbb{P}')$.

For a fixed path $\beta \in C_T$, we consider the modified inputs, on $(\Omega, \mathcal{A}, \mathbb{P})$, $W^{i,\beta} := W^i + \beta$ and $\bar{W}^\beta := \bar{W} + \beta$. Let $X^{(N),\beta}$ (respectively X^β) be the solution to equation (3.2) (resp. equation (3.1)) with input $(\zeta^{(N)}, W^{(N),\beta})$ (resp. $\bar{\zeta}, \bar{W}^\beta$) given by Theorem 21. The Lipschitz bound in Theorem 21 and the independence of ζ^i and $W^{i,\beta}$, via Lemma 54, imply that, for $\mathbb{P}$-a.e. ω,

$$\mathcal{W}_{C_T,p}(L^N(X^{(N),\beta}(\omega)), \mathcal{L}(X^\beta)) \to 0.$$

Now we build the solution $\bar{X}$ and $X^{(N)}$ resp. to (3.6) and to (3.5). We claim that the maps

$$\Omega \times C_T \ni (\omega, \beta) \mapsto X^\beta(\omega) \in C_T, \quad \Omega \times C_T \ni (\omega, \beta) \mapsto X^{(N),\beta}(\omega)$$

have versions that are jointly measurable and, for such versions, we define $\bar{X}(\omega, \omega') = X^{B(\omega')}(\omega)$ and $X^{(N)}(\omega, \omega') = X^{(N),B(\omega')}(\omega)$. Note that, by the definition of X^B, for every fixed $\omega' \in \Omega'$, we have $\mathbb{P}$-a.s.

$$dX^{B(\omega')} = b(t, X^{B(\omega')}, \mathcal{L}_{\mathbb{P}}(X^{B(\omega')}))dt + dW_t + dB_t(\omega'),$$

where the law is taken with respect to the space $(\Omega, \mathcal{A}, \mathbb{P})$. But the independence of B from the other variables implies that, $\bar{\mathbb{P}}$-a.s.,

$$\mathcal{L}_{\mathbb{P}}(X^B) = \mathcal{L}_{\mathbb{P}\otimes\mathbb{P}'}(X^B|B).$$

Hence $\bar{X}$ is a solution to equation (3.6) on the product space $\Omega \times \Omega'$. Similarly $X^{(N)}$ is a solution to (3.5) on $\Omega \times \Omega'$. Therefore we have, for $\bar{\mathbb{P}}$-a.e. (ω, ω'),

$$\mathcal{W}_{C_T,p}\left(L^N(X^{(N)})(\omega, \omega'), \mathcal{L}(\bar{X}|B)(\omega')\right) = \mathcal{W}_{C_T,p}(L^N(X^{(N),\beta})(\omega), \mathcal{L}(X^\beta)) \mid_{\beta=B(\omega')} \to 0,$$

which is the desired convergence.

It remains to prove the measurability claim on X^β and $X^{(N),\beta}$. We prove it for $X^{(N),\beta}$, the proof for $\bar{X}$ being analogous. Recall the notation in Section 2 and note that the following maps are Borel measurable

$$F_1 : \mathcal{P}_p(C_T) \times \mathbb{R}^d \times C_T \ni (\mu, x_0, \gamma) \mapsto S^\mu(x_0, \gamma) \in C_T,$$
$$F_2 : \mathcal{P}_p(\mathbb{R}^d \times C_T) \times C_T \ni (\nu, \beta) \mapsto (\cdot + (0, \beta))_\# \nu \in \mathcal{P}_p(C_T),$$

(where $\cdot + (0,\beta)$ is the map on $\mathbb{R}^d \times C_T$ defined by $(x,\gamma) + (0,\beta) = (x, \gamma+\beta)$). Indeed, F_1 is continuous (because the solution of (2.2) depends continuously on the drift, the initial data and the signal), F_2 is also Lipschitz-continuous (indeed, for any (β,ν) and (β',ν'), if m is an optimal plan between ν and ν', then $((\cdot + (0,\beta), \cdot + (0,\beta'))_\# m$ is an admissible plan between $F_2(\beta,\nu)$ and $F_2(\beta',\nu')$ and standard bounds give the Lipschitz property). Moreover let Ψ the map defined in (2.7). It is continuous, hence measurable. Now we can write, for every β in C_T, for every $i = 1, \dots N$,

$$X^{(N),\beta,i}(\omega) = F_1(\Psi(F_2(L^N(\zeta^{(N)}(\omega), W^{(N)}(\omega)), \beta)), \zeta^i(\omega), W^i(\omega) + \beta), \quad \mathbb{P} - \text{a.s.}$$

and the right-hand side above is composition of measurable maps, hence measurable. Therefore the right-hand side is a measurable version of $X^{(N),\beta}$. The proof is complete. □

3.4. **Heterogeneous mean field.** As a further application of Theorem 21 we want to consider the case of heterogeneous mean field. We will show the convergence even when the drivers are not identically distributed. This applies in particular to the results of the physical system studied in [23] as was discussed in the introduction. In that model, it is assumed that the state of each particle is influenced by its radius. Particle i has a radius r^i, which is deterministic, and it is known that the radii are distributed according to a distribution λ. We allow here for the radii to be stochastic and not necessarily identically distributed, but still independent. Moreover, we will assume the volume to change in time.

Heterogeneous mean field systems appear also in other contexts, see for example (among many others) [41], [11], which work with semimartingale inputs and use a coupling à la Sznitman [38].

On the probability space $(\Omega, \mathcal{A}, \mathbb{P})$, we consider a family $(\zeta^{(N)}, W^{(N)}) = (\zeta^i, W^i)_{i\geq 1} \subset L^p(\Omega, \mathbb{R}^d \times C_T(\mathbb{R}^d))$. This family is taken i.i.d.

In addition, for each $N \in \mathbb{N}$, we consider a family $R^{(N)} = (R^{i,N})_{1\leq i\leq N} \subset L^p(C_T(\mathbb{R}^n)^N)$.

We construct the following interacting particle system

$$\left\{ \begin{array}{l} dX_t^{i,N} = b(t, X_t^{i,N}, R_t^{i,N}, L^N(X_t^{(N)}, R_t^{(N)}))dt + dW_t^i \\ X_0^{i,N} = \zeta^i. \end{array} \right. \tag{3.7}$$

We call this an heterogeneous particle system because the particles are not exchangeable anymore, if the $R^{i,N}$ are not exchangeable.

We assume that the $R^{i,N}$ are independent of the ζ^i and W^i and that there exists a measure $\lambda \in \mathcal{P}_p(C_T(\mathbb{R}^n))$ such that

$$L^N(R^{(N)})(\omega) \stackrel{*}{\rightharpoonup} \lambda, \quad \mathbb{P} - \text{a.s.}$$

and actually in p'-Wasserstein distance for $p' > p$. We also consider the following mean field equation (on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$):

$$\begin{cases} d\bar{X}_t = b(t, \bar{X}_t, \bar{R}_t, \mathcal{L}(\bar{X}_t, \bar{R}_t))dt + d\bar{W}_t \\ \bar{X}_0 = \bar{\zeta} \end{cases} \tag{3.8}$$

where $\bar{\zeta}$, $\bar{W}$ and $\bar{R}$ are independent random variables distributed resp. as ζ^i, W^i and λ. The following result is a corollary of Theorem 21. We also use Lemma 28 and Lemma 29 to deal with the convergence of the input data.

Corollary 27. *Let $p \in [1, \infty)$, $p' \in (p, \infty)$. Assume that $b : [0, T] \times \mathbb{R}^{d+n} \times \mathcal{P}_p(\mathbb{R}^{d+n}) \to \mathbb{R}^d$ is a measurable function and there exists a constant K_b such that,*

$$|b(t, x, \mu) - b(t, x', \mu')|^p \le K_b \left(|x - x'|^p + \mathcal{W}_{\mathbb{R}^{d+n},p}(\mu, \mu')^p\right),$$

$\forall t \in [0, T], x, x' \in \mathbb{R}^{d+n}, \mu, \mu' \in \mathcal{P}_p(\mathbb{R}^{d+n})$.
Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. On this space we consider independent families $\zeta^{(N)} = (\zeta^i)_{i\ge1} \subset L^{p'}(\Omega, \mathbb{R}^d)$, $W^{(N)} = (W^i)_{i\ge1} \in L^{p'}(\Omega, C_T)$ of i.i.d. random variables. Let $\bar{\zeta}$ be distributed as ζ^1 and let $\bar{W}$ be distributed as W^1 and independent of $\bar{\zeta}$. Moreover, assume that $R^{(N)} = (R^{i,N})_{1\le i\le N}$ is a family of independent random variables in $L^{p'}(\Omega, \mathbb{R}^n)$ which are independent from the others. If there is convergence of the heterogeneous part (in p'-Wasserstein distance),

$$\mathcal{W}_{C_T(\mathbb{R}^n),p'}(L^N(R^{(N)}), \mathcal{L}(\bar{R})) \to 0 \quad \mathbb{P} - a.s. \quad as\ N \to \infty,$$

then also the solution converges (in p-Wasserstein distance),

$$\mathcal{W}_{C_T(\mathbb{R}^{d+n}),p}(L^N(X^{(N)}, R^{(N)}), \mathcal{L}(\bar{X}, \bar{R})) \quad \mathbb{P} - a.s. \quad as\ N \to \infty,$$

Proof. We start by rewriting the system (3.7) so that we can invoke Theorem 21. We change the state space of the system from $\mathbb{R}^d$ to $\mathbb{R}^d \times \mathbb{R}^n$ and we define on this new space the process $Y^{i,N}_t := (X^{i,N}_t, R^{i,N}_t)$. Clearly, $X^{i,N}$ is a solution to system (3.7) if and only if $Y^{i,N}$ solves

$$\begin{cases} dY^{i,N}_t = \begin{pmatrix} b(t, Y^{i,N}_t, L^N(Y^{(N)}_t)) \\ 0 \end{pmatrix} dt + d\begin{pmatrix} W^i_t \\ R^{i,N}_t \end{pmatrix} \\ Y^{i,N}_0 = \begin{pmatrix} \zeta^i \\ R^{i,N}_0 \end{pmatrix}. \end{cases} \tag{3.9}$$

A similar transformation can be applied to the McKean-Vlasov equation to obtain that $\bar{Y}_t = (\bar{X}_t, \bar{R}_t)$ solves

$$\begin{cases} d\bar{Y}_t = \begin{pmatrix} b(t, \bar{Y}_t, \mathcal{L}(\bar{Y}_t)) \\ 0 \end{pmatrix} dt + d\begin{pmatrix} \bar{W}_t \\ \bar{R}_t \end{pmatrix} \\ \bar{Y}_0 = \begin{pmatrix} \bar{\zeta} \\ \bar{R}_0 \end{pmatrix}. \end{cases}$$

In this setting the inputs satisfy the assumption of Theorem 21. Hence, we obtain the following inequality. $\forall \omega \in \Omega$,

$$\begin{aligned} &\mathcal{W}_{C_T(\mathbb{R}^{d+n}),p}(L^N(X^{(N)}, R^{(N)}), \mathcal{L}(\bar{X}, \bar{R}))^p \\ &\qquad \le C\mathcal{W}_{\mathbb{R}^d\times C_T(\mathbb{R}^{d+n}),p}(L^N(\zeta^{(N)}, R^{(N)}, W^{(N)}), \mathcal{L}(\bar{\zeta}, \bar{R}, \bar{W}))^p. \end{aligned}$$

By Lemma 29 (with $X_i := (\zeta^{i,N}, W^{i,N})$ and $Y_{i,N} := (R^{i,N})$ on the spaces $E := \mathbb{R}^d \times C_T$ and $F := \mathbb{R}^n$), $L^N(\zeta^{(N)}, R^{(N)}, W^{(N)})$ converges weakly to $\mathcal{L}(\bar{\zeta}, \bar{R}, \bar{W})$ $\mathbb{P}$-a.s.. Now, for every q with $p < q < p'$, $L^N(\zeta^{(N)}, W^{(N)}$ converges in q-Wasserstein distance, $\mathbb{P}$-a.s., by Lemma 54 and $L^N(R^{(N)})$ converges also in q-Wasserstein distance, $\mathbb{P}$-a.s., by assumption. In particular, $\mathbb{P}$-a.s., $L^N(\zeta^{(N)}, R^{(N)}, W^{(N)})$ have uniformly (in N) bounded q-th moments. Hence, by Lemma 53, $L^N(\zeta^{(N)}, R^{(N)}, W^{(N)})$ converges also in p-Wasserstein distance, $\mathbb{P}$-a.s., and so $\mathcal{W}_{C_T(\mathbb{R}^{d+n}),p}(L^N(X^{(N)}, R^{(N)}), \mathcal{L}(\bar{X}, \bar{R}))$ tends to 0. The proof is complete. □

The following variant of the strong law of large numbers will be useful to prove Lemma 29.

Lemma 28. *Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d. real-valued centered random variables and let $(Y_{i,N})_{1\leq i\leq N}$ be an independent family of real-valued independent random variables. Moreover, assume that there exists $C > 0$ such that*

$$\|X_i\|_{L^4(\mathbb{R})} \leq C, \quad \|Y_{i,N}\|_{L^4(\mathbb{R})} \leq C, \quad \forall i, N \geq 1.$$

Then,

$$S^N := \frac{1}{N}\sum_{i=1}^N X_i Y_{i,N} \to 0, \quad \mathbb{P} - a.s.$$

Proof. We first establish a bound on the fourth moment of the empirical sum S^N.

$$\mathbb{E}|S^N|^4 = \frac{1}{N^4}\sum_{i=1}^N \mathbb{E}\left[X_i^4\right]\mathbb{E}\left[Y_{i,N}^4\right] + \frac{6}{N^4}\sum_{i,j=1}^N \mathbb{E}\left[X_i^2\right]\mathbb{E}\left[X_j^2\right]\mathbb{E}\left[Y_{i,N}^2\right]\mathbb{E}\left[Y_{j,N}^2\right] \leq \frac{C}{N^2}.$$

Only those two terms in the sum do not vanish, because the X_i's are centered. The constant C depends on the upper bounds of the random variables. Let $p < \frac{1}{4}$,

$$E_N := \left\{|S^N| > \frac{1}{N^p}\right\}.$$

Using Chebychev inequality, we have the following

$$\sum_{N=1}^\infty \mathbb{P}\{E^N\} \leq \sum_{N=1}^\infty N^{4p}\mathbb{E}[S^N] \leq C\sum_{N=1}^\infty N^{4p-2}.$$

For our choice of p, we have convergence of the series. Borel Cantelli's Lemma implies that

$$\mathbb{P}\{\limsup_{N\to\infty} E^N\} = 0,$$

which in turn implies almost sure convergence of S^N. □

Lemma 29. *Let $p \in [1, \infty)$ be fixed. Let $(X_i)_{i\geq 1}$ be a sequence of i.i.d. random variables on a space $(\Omega, \mathcal{A}, \mathbb{P})$ taking values in a Polish space E, with law $\mu \in \mathcal{P}_p(E)$. Let $(Y_{i,N})_{1\leq i\leq N}$ be another sequence of random variables taking values on a Polish space F, which is independent from $(X_i)_{i\geq 1}$. Assume that there exists a probability measure $\lambda \in \mathcal{P}_p(F)$ such that*

$$L^N(Y^{(N)}) := \frac{1}{N}\sum_{i=1}^N \delta_{Y_{i,N}} \overset{*}{\rightharpoonup} \lambda, \quad \mathbb{P} - a.s. \tag{3.10}$$

Then,

$$L^N(X^{(N)}, Y^{(N)}) \overset{*}{\rightharpoonup} \mu \otimes \lambda, \quad \mathbb{P} - a.s.$$

Proof. Since $(X_i)_{i\geq 1}$ are a sequence of i.i.d. random variables, there exists a set of full measure $\Omega^x \subset \Omega$, such that $L^N(X^{(N)}(\omega)) \overset{*}{\rightharpoonup} \mu$, for every $\omega \in \Omega^x$. Weak convergence implies tightness of the sequence $(L^N(X^{(N)})(\omega))$, thus, for every $\epsilon > 0$, there exists a compact set $E^\omega_\epsilon \subset E$, such that

$$L^N(X^{(N)}(\omega))((E^\omega_\epsilon)^c) < \frac{\epsilon}{2}, \quad \omega \in \Omega^x.$$

In a similar way, there exists a set of full measure $\Omega^y \subset \Omega$ such that for every $\epsilon > 0$ there exists a compact $F^\omega_\epsilon \subset F$ that satisfies $L^N(Y^{(N)}(\omega))((F^\omega_\epsilon)^c) < \frac{\epsilon}{2}$, $\omega \in \Omega^y$. For every $\omega \in \Omega^x \cap \Omega^y$, we can consider the compact $K^\omega_\epsilon = E^\omega_\epsilon \times F^\omega_\epsilon \subset E \times F$ and compute the following

$$L^N(X^{(N)}(\omega), Y^{(N)}(\omega))((K^\omega_\epsilon)^c) \leq L^N(X^{(N)}(\omega))((E^\omega_\epsilon)^c) + L^N(Y^{(N)}(\omega))((F^\omega_\epsilon)^c) < \epsilon.$$

We have thus shown that the sequence $L^N(X^{(N)}, Y^{(N)})$ is almost surely tight. With an abuse of notation, we call L^N a converging subsequence and we take a continuous and bounded test function of the form $\varphi(x,y) := \varphi_1(x)\varphi_2(y)$ on $E \times F$. We compute the following

$$\begin{aligned} L^N(X^{(N)}, Y^{(N)})(\varphi) - (\mu \otimes \lambda)(\varphi) =& \frac{1}{N}\sum_{i=1}^N \varphi_2(Y_{i,N}) \left[\varphi_1(X_i) - \int_E \varphi_1(x) d\mu(x)\right] \\ &+ \frac{1}{N}\sum_{i=1}^N \int_E \varphi_1(x) d\mu(x) \left[\varphi_2(Y_{i,N}) - \int_F \varphi_2(y) d\lambda(y)\right]. \end{aligned}$$

The first term on the right hand side converges to zero thanks to Lemma 28, since the term in the brackets is a collection of bounded centered i.i.d. random variables. The second term on the right-hand side converges by assumption (3.10). □

Remark 30. The same result Corollary 27 holds actually in a slightly different context of heterogeneous noises, namely when, in equation (3.7), the noise dW^i_t in equation is replaced by $d[\sigma(R^i_t)W^i_t]$ and, in equation (3.8), the noise $d\bar{W}$ is replaced by $d[\sigma(\bar{R}^i_t)\bar{W}^i_t]$, for a continuous bounded function $\sigma : \mathbb{R} \to \mathbb{R}$. Indeed, one can repeat the proof of Corollary 27 replacing the noise in equation (3.9) by

$$d\begin{pmatrix} \sigma(R^{i,N}_t)W^i_t \\ R^{i,N}_t \end{pmatrix}$$

and similarly for corresponding McKean-Vlasov SDE, and one gets

$$\begin{aligned} &\mathcal{W}_{C_T(\mathbb{R}^{d+n}),p}(L^N(X^{(N)}, R^{(N)}), \mathcal{L}(\bar{X}, \bar{R}))^p \\ &\qquad \leq C\mathcal{W}_{\mathbb{R}^d \times C_T(\mathbb{R}^{d+n}),p}(L^N(\zeta^{(N)}, R^{(N)}, [\sigma(R)W]^{(N)}), \mathcal{L}(\bar{\zeta}, \bar{R}, [\sigma(\bar{R})\bar{W}]))^p. \end{aligned}$$

Then one notes that $(\sigma(R^{i,N})W^i, R^{i,N})$ is a continuous function of $(W^i, R^{i,N})$, hence, since $L^N(\zeta^{(N)}, R^{(N)}, W^{(N)})$ converges weakly $\mathbb{P}$-a.s., then also $L^N(\zeta^{(N)}, R^{(N)}, [\sigma(R)W]^{(N)})$ converges weakly $\mathbb{P}$-a.s.. Moreover, the convergence in q-Wasserstein distance, for $p <$

$q < p'$, of $L^N(\zeta^{(N)}, W^{(N)})$ and of $L^N(R^{(N)})$ and the boundedness of σ imply the $\mathbb{P}$-a.s. uniform (in N) bound on the q-th moments of $L^N(\zeta^{(N)}, R^{(N)}, [\sigma(R)W]^{(N)})$. Hence $L^N(\zeta^{(N)}, R^{(N)}, [\sigma(R)W]^{(N)})$ converges also in p-Wasserstein distance $\mathbb{P}$-a.s..

4. Large Deviations

In this section we assume that the driving paths W of equation (2.1) live on the space $C_{T,0}$ of continuous functions starting at 0. The results of Sections 2 and 3 apply also in this case.

Let $p \in [1, \infty)$. Let $b : [0,T] \times \mathbb{R}^d \times \mathcal{P}_p(\mathbb{R}^d) \to \mathbb{R}^d$ be a drift as before and such that it satisfies 4.

As in Section 2, we define the function

$$\begin{array}{rccc} \Phi: & \mathcal{P}_p(\mathbb{R}^d \times C_{T,0}) \times \mathcal{P}_p(C_T) & \to & \mathcal{P}_p(C_T) \\ & (\mathcal{L}(\zeta, W), \mu) & \mapsto & \mathcal{L}(X^\mu) = (S^\mu)_\# \mathcal{L}(\zeta, W), \end{array}$$

where S^μ is the solution map of ODE (2.4), as defined in (2.3), with $\mathbb{R}^d \times C_{T,0}$ instead of $\mathbb{R}^d \times C_T$ as a domain. Similarly, we consider the map Ψ defined as in (2.7), replacing C_T with $C_{T,0}$.

We introduce, for every μ in $\mathcal{P}_p(C_T)$, the map

$$f^\mu : C_T \ni \gamma \mapsto \left(\gamma_0, \gamma_\cdot - \gamma_0 - \int_0^\cdot b(s, \gamma_s, \mu_s) ds \right) \in \mathbb{R}^d \times C_{T,0}. \tag{4.1}$$

Note that $f^\mu = (S^\mu)^{-1}$ and f^μ is continuous, in particular measurable.

Lemma 31. *Let $T > 0$ be fixed and let $p \in [1, \infty)$, assume 4. The function Ψ is a bijection, with inverse given by $\Psi^{-1}(\mu) = f^\mu_\# \mu$.*

Proof. For every ν in $\mathcal{P}_p(\mathbb{R}^d \times C_T)$ and η in $\mathcal{P}_p(C_T)$, we have

$$\Phi(\nu, \mu) = (S^\mu)_\# \nu = \eta \text{ if and only if } \nu = f^\mu_\# \eta.$$

In particular, with $\eta = \mu$, we get that $\Psi(\nu) = \mu$ if and only if $\nu = f^\mu_\# \mu$. Hence Ψ is invertible, with inverse given by $\Psi^{-1}(\mu) = f^\mu_\# \mu$ (one can also show that Ψ^{-1} is continuous). □

For $N \in \mathbb{N}$, let $(\zeta^{(N)}, W^{(N)}) = (\zeta^{i,N}, W^{i,N})_{1 \le i \le N} : \Omega \to (\mathbb{R}^d \times C_{T,0})^N$ be a family of random variables. We consider the system of interacting particles on $\mathbb{R}^d$ as defined in (3.2), namely

$$\begin{cases} dX^{i,N} = b(t, X^{i,N}, L^N(X^{(N)}))dt + dW^{i,N}_t \\ X^{i,N}_0 = \zeta^{i,N}. \end{cases} \tag{4.2}$$

with solution $X^{(N)} := (X^{i,N})_{i=1,\cdots,N}$. We have seen in Section 3.2 that we can define a suitable probability space $(\Omega_N, \mathcal{A}_N, \mathbb{P}_N)$, such that

$$\mathcal{L}_{\mathbb{P}_N}(\zeta^{(N)}, W^{(N)}) = L^N(\zeta^{(N)}, W^{(N)}) := \frac{1}{N} \sum_{i=1}^N \delta_{(\zeta^{i,N}, W^{i,N})},$$

and equation (2.1) is exactly the interacting particle system (4.2). Let $(\bar{\zeta}, \bar{W}) \in L^p(\mathbb{R}^d \times C_{T,0})$, we call $\bar{X} \in L^p(C_T)$ the solution to the related McKean-Vlasov equation (3.1).
This construction shows that Ψ is a continuous function that maps the empirical measure of the inputs into the empirical measure of the particles, namely

$$\Psi\left(L^N(\zeta^{(N)}, W^{(N)})\right) = L^N(X^{(N)}), \quad \forall N \in \mathbb{N}.$$

This suggests the following immediate application to the contraction principle for large deviations.

Lemma 32. *Let $(\zeta^{(N)}, W^{(N)}) = (\zeta^{i,N}, W^{i,N})_{1\le i\le N} \subset L^p(\mathbb{R}^d \times C_{T,0})$ be a sequence of random variables and let $I : \mathcal{P}_p(\mathbb{R}^d \times C_{T,0}) \to [0,+\infty]$ be a lower semi-continuous function. Assume that that $L^N(\zeta^{(N)}, W^{(N)})$ satisfies a large deviations principle with (good) rate function I, in the sense of Definition 1.*
Let $X^{(N)} = (X^{i,N})_{i=1,\dots,N}$ be the solution to the interacting particle system (4.2) with inputs $(\zeta^{i,N}, W^{i,N})_{i=1,\dots,N}$. Then the empirical law $L^N(X^{(N)})$ satisfies a large deviations principle with (good) rate function

$$J(\mu) := I(\Psi^{-1}(\mu)) = I(f^{\mu}_{\#}\mu), \quad \forall \mu \in \mathcal{P}_p(C_T).$$

Proof. We know that the function Ψ is a continuous function, we can thus apply the contraction principle for large deviations which ensures that $L^N(X^{(N)})$ satisfies a large deviations principle with rate function

$$J(\mu) := \inf\left\{I(\nu) \mid \forall \nu \in \mathcal{P}_p(\mathbb{R}^d \times C_{T,0}), \quad \Psi(\nu) = \mu\right\}, \quad \mu \in \mathcal{P}_p(C_T).$$

From the bijectivity of Ψ, given by Lemma 31, we deduce that

$$J(\mu) = I(\Psi^{-1}(\mu)) = I(f^{\mu}_{\#}\mu), \quad \mu \in \mathcal{P}_p(C_T).$$

□

Given a Polish space E, the relative entropy between two measures $\mu, \mu' \in \mathcal{P}_p(E)$ is defined as

$$H(\mu \mid \mu') := \begin{cases} \int_E \log(\frac{d\mu}{d\mu'})d\mu, & \mu << \mu', \\ +\infty, & \textit{otherwise}. \end{cases}$$

We can specialize Lemma 32 to the case when the rate function of the inputs is the entropy with respect to a specific measure. In this case we obtain an even more explicit rate function for the convergence of the empirical measure of the particles.

Lemma 33. *Let $(\zeta^{(N)}, W^{(N)}) = (\zeta^{i,N}, W^{i,N})_{1\le i\le N} : \Omega \to (\mathbb{R}^d \times C_{T,0})^N$ be a sequence of random variables such that: There exists $\bar{\nu} \in \mathcal{P}_p(\mathbb{R}^d \times C_{T,0})$ such that $L^N(\zeta^{N)}, W^{(N)})$ satisfies a large deviations principle with good rate function*

$$H(\nu \mid \bar{\nu}), \quad \forall \nu \in \mathcal{P}_p(\mathbb{R}^d \times C_{T,0}).$$

Let $X^{(N)} = (X^{i,N})_{i=1,\dots,N}$ be the solution to the interacting particle system (4.2) with inputs $(\zeta^{i,N}, W^{i,N})_{i=1,\cdots,N}$. Then the empirical law $L^N(X^{(N)})$ satisfies a large deviations principle with good rate function

$$H(\mu \mid \Phi(\bar{\nu}, \mu)), \quad \forall \mu \in \mathcal{P}_p(C_T).$$

Proof. We can apply Lemma 32 to obtain that $L^N(X^{(N)})$ satisfies a large deviations principle with rate function

$$I(\mu) := H(\Psi^{-1}(\mu) \mid \bar{\nu}), \quad \mu \in \mathcal{P}_p(C_T).$$

We show now that $H(\Psi^{-1}(\mu) \mid \bar{\nu}) = H(\mu \mid \Phi(\bar{\nu}, \mu))$. For this, note that, by Lemma 31 and by the definition of Φ,

$$\Psi^{-1}(\mu) = f^{\mu}_{\#}\mu, \quad \bar{\nu} = f^{\mu}_{\#}\Phi(\bar{\nu}, \mu).$$

Here $f^{\mu}_{\#}$ is a push-forward via a measurable map f^{μ} with measurable inverse S^{μ}. Hence, by standard facts in measure theory, $\Psi^{-1}(\mu) \ll \bar{\nu}$ if and only if $\mu \ll \Phi(\bar{\nu}, \mu)$, in which case we have

$$\frac{d\Psi^{-1}(\mu)}{d\bar{\nu}} = \frac{d\mu}{d\Phi(\bar{\nu}, \mu)} \circ S^{\mu}.$$

Hence, in the case that $\Psi^{-1}(\mu)$ is not absolutely continuous with respect to $\bar{\nu}$, we have $H(\Psi^{-1}(\mu) \mid \bar{\nu}) = H(\mu \mid \Phi(\bar{\nu}, \mu)) = +\infty$. In the case that $\Psi^{-1}(\mu)$ is absolutely continuous with respect to $\bar{\nu}$, we have

$$H(\Psi^{-1}(\mu) \mid \bar{\nu}) = \int \frac{d\Psi^{-1}(\mu)}{d\bar{\nu}} \log \frac{d\Psi^{-1}(\mu)}{d\bar{\nu}} d\bar{\nu} = \int \frac{d\mu}{d\Phi(\bar{\nu}, \mu)} d(S^{\mu}_{\#}\bar{\nu}) = H(\mu \mid \Phi(\bar{\nu}, \mu)).$$

The proof is complete. $\square$

We will now apply Sanov's Theorem to i.i.d. inputs. The case when the convergence happens in the Wasserstein metric was proved in [42], and it requires an exponential integrability assumption on the law of the inputs.

Theorem 34. *Let $(\zeta^i, W^i)_{i\geq 1} \subset L^p(\mathbb{R}^d \times C_{T,0})$ be a sequence of i.i.d. random variables with law $\bar{\nu} := \mathcal{L}(\zeta^1, W^1)$. Assume that there exists $(x^0, \gamma^0) \in \mathbb{R}^d \times C_{T,0}$ such that*

$$\log \int_{\mathbb{R}^d \times C_{T,0}} \exp(\lambda(|x - x^0| + \|\gamma - \gamma^0\|_\infty)^p) d\bar{\nu}(x, \gamma) < +\infty, \quad \forall \lambda > 0.$$

Let $X^{(N)} := (X^{i,N})_{i=1,\dots,N}$ be the solution to the interacting particle system (4.2) with inputs $(\zeta^{(N)}, W^{(N)}) := (\zeta^i, W^i)_{i=1,\cdots,N}$. Then the empirical law $L^N(X^{(N)})$ satisfies a large deviations principle with good rate function

$$H(\mu \mid \Phi(\bar{\nu}, \mu)), \quad \forall \mu \in \mathcal{P}_p(C_T).$$

Proof. Sanov's theorem, as in [42, Theorem 1.1], gives that the empirical measure $L^N(\zeta^{(N)}, W^{(N)})$ satisfies a large deviations principle with good rate function

$$I(\nu) = H(\nu \mid \bar{\nu}), \quad \forall \nu \in \mathcal{P}_p(\mathbb{R}^d \times C_{T,0}).$$

The proof then follows from Lemma 33. $\square$

5. Central limit theorem

In this section we study the fluctuations of the empirical measure around the limit. In order to do so, we apply an abstract result of Tanaka. In its original paper [40], Tanaka studied McKean-Vlasov stochastic differential equations with linear drift, here we show how this can be also applied to more general drifts.

We restate now [40, Theorem 1.1]. Let E be a Polish space and $\mathcal{M}(E)$ (resp. $\mathcal{P}(E)$) the space of signed (resp. probability) measures on E. In this section, given a function $f(x)$ on E, we use the notation $f(\mu)$ to denote $\int_E f(x)d\mu(x)$, for $\mu \in \mathcal{M}(E)$.

Theorem 35 (Tanaka)**.** *Let $f : E \times \mathcal{P}(E) \to \mathbb{R}$ be a bounded function such that there exists*

$$f' : E \times E \times \mathcal{P}(E) \to \mathbb{R}$$

such that

(i) f' is bounded.

(ii) There exists a constant $C > 0$ such that, for all $\mu, \nu \in \mathcal{P}_p(E)$,

$$\sup_{x,y \in E} |f'(x,y,\mu) - f'(x,y,\nu)| \leq \mathcal{W}_{p,E}(\mu,\nu).$$

(iii) for all $x \in E$, $\mu, \nu \in \mathcal{P}(E)$,

$$f(x,\nu) - f(x,\mu) = \int_0^1 f'(x, \nu - \mu, \mu + \theta[\nu - \mu])d\theta,$$

where we used the notation $f'(x,\rho,\mu) = \int_E f'(x,y,\mu)\rho(dy)$, for $\rho \in \mathcal{M}$.

Assume that $(X^i)_{i\in\mathbb{N}}$ is a sequence of independent and identically distributed random variables on E with distribution μ. We define

$$\mu^N := L^N(X^{(N)}) := \frac{1}{N}\sum_{i=1}^N \delta_{X^i},$$

$$Y^N := \sqrt{N}[f(\mu^N, \mu^N) - f(\mu,\mu)],$$

where we used the notation $f(\nu,\mu) = \int_E f(y,\mu)\nu(dy)$.

Then, the probability distribution of Y^N converges to a Gaussian distribution with mean 0 and variance σ^2, where

$$\sigma^2 = \int_E [f(x,\mu) + f'(x,\mu,\mu) - m]^2 \mu(dx)$$

$$m = \int_E [f(x,\mu) + f'(x,\mu,\mu)]\mu(dx).$$

Remark 36. We changed slightly the conditions, the proof of the Theorem is exactly the same in this case as in [40].

The main idea behind Theorem 35 is that one needs to linearize the solution map with respect to the measure. Hence, we introduce the following definition of differentiability with respect to a probability measure.

Definition 37. Let E be a Polish space. A function $b : E \times \mathcal{P}_p(E) \to \mathbb{R}^d$ is said to have a linear functional derivative if there exists a function:

$$\partial_\mu b : E \times E \times \mathcal{P}_p(E) \ni (x, y, \mu) \to \partial_\mu b(x, y, \mu) \in \mathbb{R}^d,$$

continuous for the product topology, such that, for any $x \in E$ and any bounded subset $\mathcal{K} \subset \mathcal{P}_p(E)$, the function $y \to \partial_\mu b(x, y, \mu)$ is at most of p-growth in y, uniformly in $\mu \in \mathcal{K}$, and

$$b(x, \mu') - b(x, \mu) = \int_0^1 \partial_\mu b(x, \mu' - \mu, \mu + \theta[\mu' - \mu]) d\theta, \qquad \forall x \in \mathbb{R}^d,\ \mu, \mu' \in \mathcal{P}_p(\mathbb{R}^d).$$

We prove the central limit theorem under suitable differentiability assumptions on the drift with respect to the measure argument. In this section we assume the following assumption.

Assumption 38. *Let $b : \mathcal{P}_p(\mathbb{R}^d) \times \mathbb{R}^d \to \mathbb{R}^d$ and $K > 0$, assume*

(i) b differentiable in the spatial variable x with derivative $\partial_x b$.
(ii) b differentiable in the sense of Definition 37, with derivative $\partial_\mu b$.
(iii) $\partial_\mu b$ differentiable in the spatial variable y with derivative $\partial_y \partial_\mu b$.
(iv) (uniform Lipschitz continuity) For all $x, x', y, y' \in \mathbb{R}^d, \mu, \mu' \in \mathcal{P}_p(\mathbb{R}^d)$,

$$|b(\mu, x) - b(\mu', x')| \le K \left(\mathcal{W}_{\mathbb{R}^d,p}(\mu, \mu') + |x - x'| \right),$$

$$|\partial_x b(\mu, x) - \partial_x b(\mu', x')| \le K \left(\mathcal{W}_{\mathbb{R}^d,p}(\mu, \mu') + |x - x'| \right),$$

$$|\partial_\mu b(\mu, x, y) - \partial_\mu b(\mu', x', y')| \le K \left(\mathcal{W}_{\mathbb{R}^d,p}(\mu, \mu') + |x - x'| + |y - y'| \right),$$

$$|\partial_y \partial_\mu b(\mu, x, y) - \partial_y \partial_\mu b(\mu', x', y')| \le K \left(\mathcal{W}_{\mathbb{R}^d,p}(\mu, \mu') + |x - x'| + |y - y'| \right).$$

(v) (uniform boundedness) For all $x, y \in \mathbb{R}^d, \mu \in \mathcal{P}_p(\mathbb{R}^d)$.

$$|b(x, \mu)|, |\partial_x b(x, \mu)|, |\partial_\mu b(\mu, x, y)|, |\partial_y \partial_\mu b(\mu, x, y)| \le K.$$

Remark 39. Let $f \in C^1(\mathbb{R}^d)$, and $g \in C_b^1(\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d)$, then

$$b(x, \mu) := f\left(g(x, \mu, \mu)\right) = f\left(\int_{\mathbb{R}^d \times \mathbb{R}^d} g(x, y, z) \mu(dy) \mu(dz) \right)$$

satisfies Assumption 38. The standard, linear case is when $f(x) = x$ and $g(x, y, z) = g(x, y)$.

To significantly simplify the notation in this section, we assume without loss of generality that all the particles start at 0 and we remove the dependence of the solution on the initial condition. With the previous simplification, we have that, for $\mu \in C_T$, the solution map defined in (2.3) is $S(\mu, \gamma) = S^\mu : C_T \to C_T$. Moreover, for $p \in [1, \infty)$, $\Psi : \mathcal{P}_p(C_T) \to \mathcal{P}_p(C_T)$ is the fixed point map defined in (2.7). We first look at the derivative of $F(\gamma, \nu) := S(\Psi(\nu), \gamma)$ with respect to ν, denoted $F'(\gamma, \bar{\gamma}, \nu)$. For $f \in B = C_b(C_T; \mathbb{R}^d)$, define

$$(A_t(\nu) f)(\gamma) := \partial_x b(F_t(\gamma, \nu), \Psi(\nu)_t)\ f(\gamma) + \int_{C_T} \partial_y \partial_\mu b(S(F_t(\gamma, \nu), F_t(\tilde{\gamma}, \nu), \Psi(\nu)_t)\ f(\tilde{\gamma})\ d\nu(\tilde{\gamma}),$$

$$G_t(\gamma, \bar{\gamma}, \nu) := \partial_\mu b(F_t(\gamma, \nu), F_t(\bar{\gamma}, \nu), \Psi(\nu)_t).$$

The derivative F' formally satisfies the following linear differential equation in the Banach space B, with parameters $\gamma \in C_T$ and $\nu \in \mathcal{P}_p(C_T)$,

$$\frac{d}{dt}F'_t(\cdot,\bar{\gamma},\nu) = A_t(\nu)F'_t(\cdot,\bar{\gamma},\nu) + G_t(\cdot,\bar{\gamma},\nu), \quad F'_t|_{t=0} = 0, \tag{5.1}$$

It follows from Assumption 38 that the linear operator A and the forcing term G are bounded, uniformly in t, γ, ν.

Lemma 40. *Assume that b satisfies Assumption 38. Then, for every $\gamma \in C_T$ and $\nu \in \mathcal{P}(C_T)$, equation* (5.1) *admits a unique solution F'. Moreover,*

(i) $\|F'_t(\gamma,\nu)\|_B \le C(K)$, *for all* $\gamma \in C_T$, $\nu \in \mathcal{P}_p(C_T)$, $t \in [0,T]$.
(ii) $\|F'_t(\gamma,\mu) - F'_t(\gamma,\nu)\|_B \le C(K)\mathcal{W}_{p,C_T}(\mu,\nu)$, *for* $\gamma \in C_T$, $\nu,\mu \in \mathcal{P}_p(C_T)$, $t \in [0,T]$.

Proof. Since A is a bounded linear operator, we know from standard theory of ordinary differential equation that equation 5.1 admits a unique solution F' that satisfies

$$\|F'_t\|_B \le \|G_t\|_B e^{\|A_T\|_{\mathcal{L}(B;B)}}, \qquad t \in [0,T].$$

The proof of 40 (i) and 40 (ii) follows now form Assumption 38. □

It is now left to verify that the derivative F' of $F = S(\Psi)$ satisfies equation (5.1). We first need the following properties of the solution map.

Lemma 41. *Let $\nu, \nu' \in \mathcal{P}(C_T)$. For $\epsilon \in [0,1]$, we define $\nu^\epsilon = \nu + \epsilon[\nu' - \nu]$. We have the following,*

(i) $\mathcal{W}_p(\nu^\epsilon,\nu) \to 0$, *as* $\epsilon \to 0$.
(ii) $\mathcal{W}_p(\mu^\epsilon,\mu) = \mathcal{W}_p(\Psi(\nu^\epsilon),\Psi(\nu)) \to 0$, *as* $\epsilon \to 0$.
(iii) $\sup_{\gamma\in C_T}\|S(\mu^\epsilon,\gamma) - S(\mu,\gamma)\|_{C_T} \to 0$, *as* $\epsilon \to 0$.

Proof. 41 (i) follows from the tightness of ν, ν' and iii.
41 (ii) follows from 41 (i) and the Lipschitz continuity of Ψ, 7 (ii).
41 (iii) is implied by 41 (ii) and straight-forward computations. □

Lemma 42. *For every $\gamma \in C_T$, the function $\mathcal{P}_p(C_T) : \nu \to F_t(\gamma,\nu) = S(\Psi(\nu),\gamma)$ is differentiable in the sense of Definition 37 and its derivative satisfies equation* (5.1).

Proof. Let $\nu, \nu' \in \mathcal{P}_p(C_T)$ and let $F'(\gamma,\bar{\gamma},\nu)$ be a solution to equation (5.1). Using the equations for S, Lemma 41 and standard (but lengthy) computations it can be proved that

$$\lim_{\epsilon\to 0}\frac{S(\Psi(\nu^\epsilon),\gamma) - S(\Psi(\nu),\gamma)}{\epsilon} - \int_{C_T} F'_t(\gamma,\bar{\gamma},\nu)d[\nu'-\nu](\bar{\gamma}) = 0.$$

□

The main result of this section is the following, which is a corollary of Theorem 35.

Corollary 43. *Let $(W^i)_{i\in\mathbb{N}}$ be a family of independent and identically distributed random variables on the Banach space C_T with law ν and let $X^{(N)} = (X^{i,N})_{i=1,\dots,N}$ be the solution of the interacting particle system* (1.2) *with input $(W^i)_{i\in\mathbb{N}}$. Let W be a random variable on C_T with law ν, we call X the solution to the McKean-Vlasov equation* (1.1) *driven by W. Define $\mu := \mathcal{L}(X) = \Psi(\nu)$.*

Let $\varphi : C_T \to \mathbb{R}$ be a bounded Fréchet-differentiable test function with bounded derivative φ'. We have that

$$Y^N := \sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N}\varphi(X^{i,N}) - \mu(\varphi)\right)$$

converges, as $N \to \infty$, to a Gaussian $N(0, \sigma(\varphi))$ with

$$\sigma^2(\varphi) = \int_{C_T}[\varphi(F(\gamma,\nu)) + \int_{C_T}\varphi'(F(\bar{\gamma},\nu))F'(\nu,\bar{\gamma},\gamma)\nu(d\bar{\gamma}) - m]^2\nu(d\gamma),$$

$$m(\varphi) = \int_{C_T}[\varphi(F(\gamma,\nu)) + \int_{C_T}\varphi'(F(\bar{\gamma},\nu))F'(\nu,\bar{\gamma},\gamma)\nu(d\bar{\gamma})]\nu(d\gamma),$$

where $F = S(\Psi)$ and F' is the solution to equation (5.1).

Proof. The function $f(\gamma,\nu) := \varphi(S(\Psi(\nu),\gamma))$ satisfies the assumption of Theorem 35 with derivative $f'(\gamma,\bar{\gamma},\nu) = \varphi'(S(\Psi(\nu),\gamma))F(\gamma,\bar{\gamma},\nu)$, where φ' is the Fréchet-derivative of φ and F is a solution to equation (5.1).

Assumptions 35 (i) and 35 (ii) follow from Lemma 41. Assumption 35 (iii) follows from Lemma 42. □

6. Reflection at the boundary

The problem of SDEs in a domain with reflection has been considered since the works by Skorokhod [35], [36]. The literature is vast and we mention the works by Tanaka [39], Lions and Sznitman [30] as two of the most important papers. The case of mean field SDEs with reflection has also been studied, see for example the works by Sznitman [37], Graham and Metivier [22], which establish well-posedness under general conditions and particle approximation for independent inputs and with Brownian motion as driving signal (possibly with a diffusion coefficient). Also other types of SDEs with mean field interactions and in domains have been studied (with different kind of reflections), see for example [24], [9].

Here we show how to adapt the main result, Theorem 7, and the argument to the case of reflecting boundary conditions. With respect to the previously cited works, we can allow general continuous paths as inputs, we do not need to assume independence nor exchengeability of particles for particle approximation.

Throughout this section, we assume that D is a bounded convex polyhedron in $\mathbb{R}^d$ with nonempty interior (see Remark 52 below for extensions).

We are given a Borel vector field b that satisfies the following

Assumption 44. *Let $p \in [1, +\infty)$. The function $b : [0,T] \times \bar{D} \times \mathcal{P}_p(\bar{D} \times \mathbb{R}^d) \to \mathbb{R}^d$ is a measurable function and there exists a constant K_b such that,*

$$|b(t,x,\mu) - b(t,x',\mu')|^p \leq K_b\left(|x-x'|^p + \mathcal{W}_{\mathbb{R}^d,p}(\mu,\mu')^p\right),$$

$\forall t \in [0,T], x, x' \in \mathbb{R}^d, \mu, \mu' \in \mathcal{P}_p(\bar{D} \times \mathbb{R}^d)$.

We consider the generalized McKean-Vlasov Skorokhod problem

$$\begin{cases} dX_t = b(t, X_t, \mathcal{L}(X_t, k_t))dt + dW_t - dk_t \\ X \in C_T(\bar{D}), \quad X_0 = \zeta, \\ k \in BV_T, \quad d|k|_t = 1_{X_t \in \partial D} d|k|_t, \quad dk_t = n(X_t)d|k|_t. \end{cases} \tag{6.1}$$

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, the input to equation 6.1 is a random variable (ζ, W) with values in $\bar{D} \times C_T$, the solution is the couple (X, k) of random variables satisfying the equation above, $|k|$ denotes the total variation process of k (not the modulus of k) and $n(x)$ is the outer normal at x, for x in ∂D, see Remark 46 below for the precise meaning. A short explanation on the meaning of the k term is given later after the main result.

We give now the precise definition of solution:

Definition 45. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\zeta : \Omega \to D$, $W : \Omega \to C_T$ be random variables on it. A solution to the generalized McKean-Vlasov Skorokhod problem with input (ζ, W) is a couple of random variables $X : \Omega \to C_T(\bar{D})$ and $k : \Omega \to C_T$ such that, for Lebesgue-a.e. t, $\mathcal{L}(X_t, k_t)$ is in $\mathcal{P}_p(\bar{D} \times \mathbb{R}^d)$ and, for a.e. ω, equation (6.1) is satisfied (where $X \in C_T(\bar{D})$ means that X is $C_T(\bar{D})$-valued, $k \in BV_T$ means that $k \in BV_T := BV([0, T]; \mathbb{R}^d)$ $\mathbb{P}$-a.s. and where the last line is understood in the sense of Remark 46 below).

Remark 46. Actually the last condition is only valid for smooth domains, which is not the case for D convex polyhedron (it is not smooth at the intersections of the faces of the polyhedron). For simplicity of notation, here and in what follows (also for the particle system), we keep the formulation above, with the understanding that the precise condition should be: for a.e. ω there exists a Borel function $\gamma = \gamma^\omega : [0, T] \to \mathbb{R}^d$ such that $dk_t = \gamma_t d|k|_t$ and, for $d|k|$-a.e. t, γ_t belongs to $d(X_t)$, where

$$d(x) = \left\{ \sum_{i, x \in \partial D_i} \alpha_i n_i \mid \alpha_i \geq 0, \left| \sum_{i, x \in \partial D_i} \alpha_i n_i \right| = 1 \right\}$$

and where ∂D_i are the faces of the polyhedron with outer normals n_i.

Our main result is, as before, well-posedness of the generalized McKean-Vlasov Skorokhod problem and Lipschitz continuity with respect to law of the input.

Theorem 47. *Let $T > 0$ be fixed and let $p \in [1, \infty)$. Assume that b satisfies 44.*

(1) For every input (ζ, W) (random variable in $L^p(\bar{D} \times C_T)$) with finite p-moment, there exists a unique solution (X, k) to the generalized McKean-Vlasov Skorokhod problem (6.1).

(2) There exists a constant $\tilde{C} = \tilde{C}(p, T, b) > 0$ such that: for every two inputs (ζ^i, W^i), $i = 1, 2$ (defined possibly on different probability spaces) with finite p-moments, the following is satisfied

$$W_{C_T(\bar{D}) \times C_T, p}(\mathcal{L}(X^1, k^1), \mathcal{L}(X^2, k^2)) \leq \tilde{C} W_{\bar{D} \times C_T, p}(\mathcal{L}(\zeta^1, W^1), \mathcal{L}(\zeta^2, W^2)).$$

In particular, the law of a solution (X, k) depends only on the law of (ζ, W).

To prove this result, we regard the generalized McKean-Vlasov Skorokhod problem as a fixed point problem with parameter. For this, we introduce the following Skorokhod problem, for fixed μ in $\mathcal{P}_p(C_T(\bar{D}) \times C_T)$ (calling μ_t the marginal at time t):

$$\begin{cases} dY_t^\mu = b(t, Y_t^\mu, \mu_t)dt + dW_t - dh_t^\mu \\ Y^\mu \in C_T(\bar{D}), \quad Y_0^\mu = \zeta, \\ h^\mu \in BV_T, \quad d|h^\mu|_t = 1_{Y_t \in \partial D} d|h^\mu|_t, \quad dh_t^\mu = n(Y_t) d|h^\mu|_t. \end{cases} \tag{6.2}$$

We recall the following well-posedness result for μ fixed:

Lemma 48. *Fix μ in $\mathcal{P}_p(C_T(\bar{D}) \times C_T)$ and assume that b is Lipschitz and bounded as in Theorem 47. Then, for every $T > 0$, for every deterministic initial datum $\zeta \equiv x_0$ in $\bar{D}$ and for every deterministic path $W \equiv \gamma$ in C_T, there exists a unique solution $(Y, h) = (Y^\mu, h^\mu)$ in $C_T(\bar{D}) \times C_T$ to the above equation.*

This result is classical and one can see it as a consequence of well-posedness for Skorokhod problem without drift, via Lemma 49 below, in the same line of the proof of Theorem 47 (see in particular the bound (6.3)). We call $S^\mu : \bar{D} \times C_T \to C_T(\bar{D}) \times C_T$ the solution map to (6.2), that is, $S^\mu(x_0, \gamma) = (Y^\mu, h^\mu)$ where (Y^μ, h^μ) solves (6.2) with deterministic input $(x_0, \gamma) \in \bar{D} \times C_T$.

For a general random input (ζ, W) in $L^p(\bar{D} \times C_T)$, this result, applied to $(\zeta(\omega), W(\omega))$ for a.e. ω, gives existence and pathwise uniqueness of a solution (Y^μ, h^μ) to (6.2) and the representation formula $(Y^\mu, h^\mu) = S^\mu(\zeta, W)$. Moreover, again from Lemma 49 below, if the input (ζ, W) has finite p-moment, then also the solution (Y^μ, h^μ) has finite p-moment. We call

$$\begin{array}{rccc} \Phi : & \mathcal{P}_p(\bar{D} \times C_T) \times \mathcal{P}_p(C_T(\bar{D}) \times C_T) & \to & \mathcal{P}_p(C_T(\bar{D}) \times C_T), \\ & (\mathcal{L}(\zeta, W), \mu) & \mapsto & (S^\mu)_\# \mathcal{L}(\zeta, W), \end{array}$$

the law of a probability measure $\mathcal{L}(\zeta, W)$, under the solution map S_T^μ of the Skorokhod problem with μ fixed.

As in the case without boundaries, note that (X, k) solves the McKean-Vlasov Skorokhod problem if and only if $\mathcal{L}(X, k)$ is a fixed point of $\Phi(\mathcal{L}(\zeta, W), \cdot)$. Hence, Theorem 47 reduces to a fixed point problem with parameter.

A key tool in the proof of this result is the Lipschitz dependence of the boundary term k on the given path in the Skorokhod problem. The precise statement follows from [20, Theorem 2.2] (there the Skorokhod problem is formulated in the space of cadlag functions, but continuity of the solution is ensured by [39, Lemma 2.4]).

Lemma 49. *Fix $T > 0$. For x_0 in $\bar{D}$, z in C_T. Then there exists a unique solution $(y, k) = (y^{x_0, z}, k^{x_0, z})$ in $C_T(\bar{D}) \times C_T$ to the Skorokhod problem driven by z, namely*

$$\begin{cases} dy = dz - dk, \\ y \in C_T(D), \quad y_0 = x_0, \\ k \in BV_T, \quad d|k| = 1_{y \in \partial D} d|k|, \quad dk = n(y) d|k|. \end{cases}$$

Moreover there exists $C \geq 0$ (which is locally bounded in T) such that, for every x_0^1, x_0^2 in D, for every z^1, z^2 in C_T,

$$\|y^{x_0^1, z^1} - y^{x_0^2, z^2}\|_\infty + \|k^{x_0^1, z^1} - k^{x_0^2, z^2}\|_\infty \leq C|x_0^1 - x_0^2| + C\|z^1 - z^2\|_\infty,$$

$$\|y^{x_0^1, z^1} - x_0^1\|_\infty + \|k^{x_0^1, z^1}\|_\infty \leq C\|z^1\|_\infty.$$

Proof of Theorem 47. The result follows from the abstract Proposition 6, provided we verify conditions 1) and 2) on Φ.

Let $\mu \in \mathcal{P}_p(C_T(\bar{D}) \times C_T)$ be fixed, let ν^1 and ν^2 be in $\mathcal{P}_p(\bar{D} \times C_T)$ and let m be an optimal plan on $(\mathbb{R}^d \times C_T)^2$ for these two measures. On the probability space $((\bar{D} \times C_T)^2, m)$, we call ζ^i, W^i, $i = 1, 2$, the r.v. defined by the canonical projections and $(Y^i, h^i) = S^\mu(\zeta^i, W^i)$ the solution to the Skorokhod problem (6.2) with input (ζ^i, W^i). We have

$$W_p(\Phi(\nu^1, \mu), \Phi(\nu^2, \mu))^p \leq \mathcal{E}^m(\|Y^1 - Y^2\|_\infty + \|h^1 - h^2\|_\infty)^p,$$

so it is enough to bound the right-hand side. We can apply Lemma 49 to $z^i = \int_0^t b(t, Y^i_r, \mu)dr + W^i$, $x^i_0 = \zeta^i$ and so $y^i = Y^i$, $k^i = h^i$: we get

$$\|Y^1 - Y^2\|_\infty + \|h^1 - h^2\|_\infty \leq C|\zeta^1 - \zeta^2| + C\int_0^T |b(t, Y^1_t, \mu) - b(t, Y^2_t, \mu)|dt + C\|W^1 - W^2\|_\infty.$$

Using the Lipschitz property of b in x (uniformly in μ), we get

$$\|Y^1 - Y^2\|_\infty + \|h^1 - h^2\|_\infty \leq C|\zeta^1 - \zeta^2| + C\int_0^T \|Y^1 - Y^2\|_\infty dt + C\|W^1 - W^2\|_\infty.$$

By Gronwall inequality

$$\|Y^1 - Y^2\|_\infty + \|h^1 - h^2\|_\infty \leq C|\zeta^1 - \zeta^2| + C\|W^1 - W^2\|_\infty.$$

We take expectation (with respect to m) of the p-power and use the optimality of m, to obtain

$$W_p(\Phi(\nu^1, \mu), \Phi(\nu^1, \mu))^p \leq CW_p(\nu^1, \nu^2)^p.$$

This ends the proof of condition 1) of Proposition 6.

Let now (ζ, W) be fixed with law $\nu := \mathcal{L}(\zeta, W)$. Consider $\mu^1, \mu^2 \in \mathcal{P}_p(C_T(\bar{D}) \times C_T)$ and call $(Y^i, h^i) = (Y^{\mu^i}, h^{\mu^i})$, $i = 1, 2$ the corresponding solutions to the Skorokhod problem (6.2) (driven by the initial datum ζ and the path W). We can apply Lemma 49 to $z^i = \int_0^t b(t, Y^{\mu^i}_r, \mu^i)dr + W$, $x^i_0 = \zeta$ and so $y^i = Y^{\mu^i}$, $k^i = h^i$: we get

$$\|Y^1 - Y^2\|_\infty + \|h^1 - h^2\|_\infty \leq C\int_0^T |b(t, X^{\mu^1}_r, \mu^1) - b(t, X^{\mu^2}_r, \mu^2)|dr.$$

Taking the p-power and arguing as without boundaries, we get

$$\|Y^1 - Y^2\|^p_\infty + \|h^1 - h^2\|^p_\infty \leq C\int_0^T \|Y^1 - Y^2\|^p_\infty dt + C\int_0^T W_{C_t,p}(\mu^1, \mu^2)^p dt$$

and so, by Gronwall inequality,

$$\|Y^1 - Y^2\|^p_\infty + \|h^1 - h^2\|^p_\infty \leq C\int_0^T W_{C_t,p}(\mu^1, \mu^2)^p dt. \tag{6.3}$$

Taking expectation, we conclude

$$W_p(\Phi(\nu, \mu^1), \Phi(\nu, \mu^2))^p \leq C\int_0^T W_{C_t,p}(\mu^1, \mu^2)^p dt.$$

As for without boundaries, iterating this inequality k times for k large enough (such that $(CT)^k/k! < 1$), we get condition 2) in Proposition 6. The proof is complete. □

As in the case without boundary, if the driving process is adapted, then so is the solution to the McKean-Vlasov Skorokhod problem. We omit the proof as it is completely analogous to the one without boundary.

Proposition 50. *Let $(\mathcal{F}_t)_t$ be a right-continuous, complete filtration on $(\Omega, \mathcal{A}, \mathbb{P})$ such that ζ is $\mathcal{F}_0$-measurable and W is $(\mathcal{F}_t)_t$-progressively measurable. Then the solution (X, k) to* (6.1) *is also $(\mathcal{F}_t)_t$-progressively measurable.*

Finally, following Section 3.1, we can obtain a particle approximation to the McKean-Vlasov Skorokhod problem (6.1), just as corollary of the main result Theorem 47. Here the corresponding particle system reads

$$\begin{cases} dX_t^{i,N} = b(t, X_t^{i,N}, L^N(X_t^{(N)}, k_t^{(N)}))dt + dW_t^{i,N} - dk_t^{i,N} \\ X^{i,N} \in C_T(\bar{D}), \quad X_0^{i,N} = \zeta^{i,N}, \\ k^{i,N} \in BV_T, \quad d|k^{i,N}|_t = 1_{X_t^{i,N} \in \partial D} d|k^{i,N}|_t, \quad dk_t^{i,N} = n(X_t^{i,N}) d|k^{i,N}|_t. \end{cases} \tag{6.4}$$

Again the solution is an N-uple of couples $(X^{i,N}, k^{i,N})_{i=1,\dots N}$ (and again $|k^{i,N}|$ denotes the total variation process of $k^{i,N}$ and $k^{i,N} \in BV_T$ means that $k^{i,N}$ belongs to BV_T $\mathbb{P}$-a.s.). The following result can be proven exactly as Theorem 21 (here we use a notation analogous to that theorem).

Theorem 51. *Let $p \in [1, \infty)$ and assume b satisfies Assumption 44. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. On this space we consider, for $N \in \mathbb{N}$, a family of random variables $(\zeta^{(N)}, W^{(N)}) = (\zeta^{i,N}, W^{i,N})_{1 \le i \le N}$ taking values on $\bar{D} \times C_T$. Let $\bar{\zeta} \in L^p(\Omega, \bar{D})$ and $\bar{W} \in L^p(\Omega, C_T)$. Then:*

- *i There exists a unique pathwise solution $(X^{(N)}, k^{(N)})$ (resp. $(\bar{X}, \bar{k})$) to the interacting particle system* (6.4) *(resp. equation* (6.1)*).*
- *ii There exists a constant C depending on b such that for all $N \ge 1$, for a.e. $\omega \in \Omega$,*

$$\begin{aligned} &W_{C_T(\bar{D}) \times C_T, p}(L^N(X^{(N)}(\omega), k^{(N)}(\omega)), \mathcal{L}(\bar{X}, \bar{k}))^p \\ &\qquad \le C W_{\bar{D} \times C_T, p}(L^N(\zeta^{(N)}(\omega), W^{(N)}(\omega)), \mathcal{L}(\bar{\zeta}, \bar{W}))^p. \end{aligned}$$

- *iii If the empirical $L^N(\zeta^{(N)}, W^{(N)})$ converges to $\mathcal{L}(\bar{\zeta}, \bar{W})$ $\mathbb{P}$-a.s., then also the emprical measure of the solution converges.*

Remark 52. More general cases can be treated, for example oblique reflection or even more general domains D, possibly with some extra assumptions: as one can see from the proof, it is enough to have an estimate as in Lemma 49 for the boundary term. The case of oblique reflection (still with D convex polyhedron) is treated in [20] (see Assumptions 2.1 and Theorem 2.1 there). The case of more general domains is treated for example in [34, 39], though the Lipschitz constant in Lemma 49 seems in this case to depend also on z.

Appendix A. Proof of Proposition 6

In this section we prove proposition 6.

First, we must show that Φ^Q has a unique fixed point. If $k=1$, it is exactly the contraction principle, so we will assume $k>1$. Clearly $(\Phi^Q)^k$ is a contraction, hence it is has a unique fixed point P_Q. Hence,

$$d_E(\Phi^Q(P_Q), P_Q) = d_E((\Phi^Q)^{k+1}(P_Q), (\Phi^Q)^k(P_Q)) \leq c d_E(\Phi^Q(P_Q), P_Q).$$

Since $c<1$, this implies $d_E(\Phi^Q(P_Q), P_Q) = 0$ and therefore P_Q is also a fixed point for Φ^Q. Every fixed point of Φ^Q is also a fixed point for $(\Phi^Q)^k$, hence P_Q is the only fixed point of Φ^Q.

We are left to prove (2.6). By induction, one can show that

$$\forall Q, Q' \in F, \forall P \in E \quad d_E((\Phi^Q)^k(P), (\Phi^{Q'})^k(P)) \leq \left(\sum_{i=1}^k L^i\right) d_F(Q, Q').$$

Using a triangular inequality as well as assumption 2) and the previous inequality we obtain

$$\begin{aligned} d_E(P_Q, P_{Q'}) =& d_E((\Phi^Q)^k(P), (\Phi^{Q'})^k(P')) \\ \leq& d_E((\Phi^Q)^k(P), (\Phi^Q)^k(P')) + d_E(\Phi^k_Q(P'), (\Phi^{Q'})^k(P')) \\ \leq& c d_E(P_Q, P_{Q'}) + \left(\sum_{i=1}^k L^i\right) d_F(Q, Q'). \end{aligned}$$

The proof is complete.

Appendix B. Wasserstein Metric

We now recall some useful information on the Wasserstein metric, which we defined in (1.8). For more details the reader can refer to [2]. Let $p \in [1, \infty)$.

i The infimum in the definition of Wasserstein metric is a minimum. For each couple $\mu, \nu \in \mathcal{P}_p(E)$ there exists a measure $m \in \Gamma(\mu, \nu)$ such that

$$\mathcal{W}_{E,p}(\mu, \nu)^p = \iint_{E \times E} d(x,y)^p m(dx, dy). \tag{B.1}$$

ii The Wasserstein distance of two measures on the space of paths is larger than the distance of the corresponding one-time marginals at t, for any t. Indeed, note that, for any $\mu, \nu \in \mathcal{P}_p(C_T)$, if m is in $\Gamma(\mu, \nu)$, then $m_t \in \Gamma(\mu_t, \nu_t)$, therefore we have

$$\mathcal{W}_{\mathbb{R}^d,p}(\mu_t, \nu_t)^p \leq \iint_{\mathbb{R}^d \times \mathbb{R}^d} |x - x'|^p m_t(dx, dx') = \iint_{C_T \times C_T} |\gamma_t - \gamma'_t|^p m(d\gamma, d\gamma') \leq \mathcal{W}_{C_T,p}(\mu, \nu)^p.$$

iii Let E be a Polish space. For any given sequence $(\mu^n)_{n \geq 1} \in \mathcal{P}_p(E)$ the following are equivalent
  (a) (The sequence converges in Wassertein sense) $\lim_{n\to\infty} \mathcal{W}_{E,p}(\mu_n, \mu) = 0$.
  (b) (The sequence converges weakly and is uniformly integrable) There exists $x_0 \in E$ such that,

$$\begin{cases} \mu_n \overset{*}{\rightharpoonup} \mu, \quad \text{as } n \to \infty \\ \lim_{k\to\infty} \int_{E \setminus B_k(x_0)} d^p(x, x_0) d\mu^n(x) = 0, \quad \text{uniformly in } n. \end{cases}$$

Cf. [2, Proposition 7.1.5].

As a consequence of point (iii), we give a sufficient condition to pass from weak convergence of measures to convergence in the p-Wasserstein distance.

Lemma 53. *Let (E,d) be a Polish space and μ_n, $n \in \mathbb{N}$, μ be probability measures on E, fix $q \in [1,\infty)$. If the sequence $(\mu_n)_{n\in\mathbb{N}}$ converges to μ in the weak topology on probability measures and if, for some $p \in (q,\infty)$ and some x_0 in E,*

$$\sup_n \int_E d(x,x_0)^p \mu_n(dx) < \infty, \tag{B.2}$$

then μ_n converges in q-Wasserstein metric to $\mu \in \mathcal{P}_q(E)$.

Proof. By property (iii), it is enough to show that the map $x \mapsto d(x,x_0)^q$ is uniformly integrable with respect to $(\mu_n)_n$. For this, we have, for any $R > 0$, for any n,

$$\int_{d(x,x_0)>R} d(x,x_0)^q \mu_n(dx) \leq R^{p-q} \int_E d(x,x_0)^p \mu_n(dx).$$

By the uniform bound (B.2), we can choose R large enough to make the right-hand side above small for all n. This shows that $x \mapsto d(x,x_0)^q$ is uniformly integrable. □

Lemma 54. *Given $p \in (1,\infty)$ and a separable Banach space $(E,|\cdot|)$, let $(X^i)_{i\geq 1} \in L^p(\Omega,E)$ be a family of i.i.d. random variables on this space with law μ. Then,*

$$\lim_{N\to\infty} \mathcal{W}_{E,q}(L^N(X^{(N)}),\mu) = 0, \quad q \in (1,p), \quad \mathbb{P}-a.s.$$

Proof. Since (X^i) are i.i.d., $\mathbb{P}$-a.s. convergence in the weak topology

$$L^N(X^{(N)}(\omega)) \overset{*}{\rightharpoonup} \mathcal{L}(X^1), \quad \mathbb{P}-a.s.$$

is a classical result, see for example [43] and references therein. Moreover, by the law of large numbers, we have, for a.e. ω,

$$\int_E |x|^p dL^N(X^{(N)}(\omega))(x) = \frac{1}{N}\sum_{i=1}^N |X^i(\omega)|^p \to \mathbb{E}|X^1|^p < \infty.$$

We obtain condition (B.2) in Lemma 53, which concludes the proof. □

References

[1] S. Aida. Reflected rough differential equations. *Stochastic Processes and their Applications*, 125(9):3570–3595, 2015.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

[3] I. Bailleul. Flows driven by rough paths. *Rev. Mat. Iberoam.*, 31(3):901–934, 2015.

[4] I. Bailleul, R. Catellier, and F. Delarue. Mean field rough differential equations. *arXiv preprint arXiv:1802.05882*, 2018.

[5] I. Bailleul, R. Catellier, and F. Delarue. Propagation of chaos for mean field rough differential equations. *arXiv preprint arXiv:1907.00578*, 2019.

[6] H. Bessaih, M. Coghi, and F. Flandoli. Mean field limit of interacting filaments for 3D Euler equations. *J. Stat. Phys.*, 174(3):562–578, 2019.

[7] P. Billingsley. *Convergence of probability measures.* Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[8] V. I. Bogachev. *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007.

[9] P. Briand, P. Éric Chaudru de Raynal, A. Guillin, and C. Labart. Particles systems and numerical schemes for mean reflected stochastic differential equations, 2016.

[10] T. Cass and T. Lyons. Evolving communities with individual preferences. *Proc. Lond. Math. Soc. (3)*, 110(1):83–107, 2015.

[11] C. Chong and C. Klueppelberg. Partial mean field limits in heterogeneous networks, 2015.

[12] M. Coghi and F. Flandoli. Propagation of chaos for interacting particles subject to environmental noise. *Ann. Appl. Probab.*, 26(3):1407–1442, 2016.

[13] M. Coghi and B. Gess. Stochastic nonlinear Fokker-Planck equations. *Nonlinear Anal.*, 187:259–278, 2019.

[14] M. Coghi and T. Nilssen. Rough nonlocal diffusions. *arXiv preprint arXiv:1905.07270*, 2019.

[15] D. A. Dawson and J. Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

[16] F. Delarue, J. Inglis, S. Rubenthaler, and E. Tanré. Global solvability of a networked integrate-and-fire model of McKean-Vlasov type. *Ann. Appl. Probab.*, 25(4):2096–2133, 2015.

[17] J.-D. Deuschel, P. K. Friz, M. Maurelli, and M. Slowik. The enhanced Sanov theorem and propagation of chaos. *Stochastic Process. Appl.*, 128(7):2228–2269, 2018.

[18] A. Deya, M. Gubinelli, M. Hofmanova, and S. Tindel. One-dimensional reflected rough differential equations. *Stochastic Processes and their Applications*, 2018.

[19] W. Dreyer, R. Huth, A. Mielke, J. Rehberg, and M. Winkler. Global existence for a nonlocal and nonlinear Fokker-Planck equation. *Z. Angew. Math. Phys.*, 66(2):293–315, 2015.

[20] P. Dupuis and H. Ishii. On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications. *Stochastics Stochastics Rep.*, 35(1):31–62, 1991.

[21] N. Fournier, M. Hauray, and S. Mischler. Propagation of chaos for the 2D viscous vortex model. *J. Eur. Math. Soc. (JEMS)*, 16(7):1423–1466, 2014.

[22] C. Graham and M. Métivier. System of interacting particles and nonlinear diffusion reflecting in a domain with sticky boundary. *Probab. Theory Related Fields*, 82(2):225–240, 1989.

[23] C. Guhlke, P. Gajewski, M. Maurelli, P. K. Friz, and W. Dreyer. Stochastic many-particle model for lfp electrodes. *Continuum Mechanics and Thermodynamics*, 30(3):593–628, 2018.

[24] B. Hambly and S. Ledger. A stochastic McKean-Vlasov equation for absorbing diffusions on the half-line. *Ann. Appl. Probab.*, 27(5):2698–2752, 2017.

[25] T. G. Kurtz and J. Xiong. Particle representations for a class of nonlinear spdes. *Stochastic Processes and their Applications*, 83(1):103–126, 1999.

[26] T. G. Kurtz and J. Xiong. Numerical solutions for a class of SPDEs with application to filtering. In *Stochastics in finite and infinite dimensions*, Trends Math., pages 233–258. Birkhäuser Boston, Boston, MA, 2001.

[27] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006.

[28] J.-M. Lasry and P.-L. Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006.

[29] J.-M. Lasry and P.-L. Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007.

[30] P.-L. Lions and A.-S. Sznitman. Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.*, 37(4):511–537, 1984.

[31] E. Luçon and W. Stannat. Mean field limit for disordered diffusions with singular interactions. *Ann. Appl. Probab.*, 24(5):1946–1993, 2014.

[32] C. Marchioro and M. Pulvirenti. *Vortex methods in two-dimensional fluid dynamics*, volume 203 of *Lecture Notes in Physics*. Springer-Verlag, Berlin, 1984.

[33] S. Nadtochiy and M. Shkolnikov. Mean field systems on networks, with singular interaction through hitting times. *arXiv preprint arXiv:1807.02015*, 2018.

[34] Y. Saisho. Stochastic differential equations for multidimensional domain with reflecting boundary. *Probab. Theory Related Fields*, 74(3):455–477, 1987.
[35] A. V. Skorohod. Stochastic equations for diffusion processes with a boundary. *Teor. Verojatnost. i Primenen.*, 6:287–298, 1961.
[36] A. V. Skorohod. Stochastic equations for diffusion processes with boundaries. II. *Teor. Verojatnost. i Primenen.*, 7:5–25, 1962.
[37] A.-S. Sznitman. Nonlinear reflecting diffusion process, and the propagation of chaos and fluctuations associated. *J. Funct. Anal.*, 56(3):311–336, 1984.
[38] A.-S. Sznitman. Topics in propagation of chaos. In *Ecole d'été de probabilités de Saint-Flour XIX—1989*, pages 165–251. Springer, 1991.
[39] H. Tanaka. Stochastic differential equations with reflecting boundary condition in convex regions. *Hiroshima Math. J.*, 9(1):163–177, 1979.
[40] H. Tanaka. Limit theorems for certain diffusion processes with interaction. In *Stochastic analysis (Katata/Kyoto, 1982)*, volume 32 of *North-Holland Math. Library*, pages 469–488. North-Holland, Amsterdam, 1984.
[41] J. Touboul. Propagation of chaos in neural fields. *Ann. Appl. Probab.*, 24(3):1298–1328, 2014.
[42] R. Wang, X. Wang, and L. Wu. Sanov's theorem in the Wasserstein distance: a necessary and sufficient condition. *Statist. Probab. Lett.*, 80(5-6):505–512, 2010.
[43] J. A. Wellner. A Glivenko-Cantelli theorem for empirical measures of independent but nonidentically distributed random variables. *Stochastic Process. Appl.*, 11(3):309–312, 1981.

(M. Coghi) Weierstrass Institute for Applied Stochastic, Mohrenstrasse 39, 10117 Berlin and Faculty of Mathematics, University of Bielefeld, 33615 Bielefeld, Germany

Email address: `michele.coghi@gmail.com`

(J.-D. Deuschel) Technische Universität Berlin, Str. des 17. Juni 136, 10587 Berlin, Germany

Email address: `deuschel@math.tu-berlin.de`

(P.K. Friz) Technische Universität Berlin, Str. des 17. Juni 136, 10587 Berlin and Weierstrass Institute for Applied Stochastic, Mohrenstrasse 39, 10117 Berlin, Germany

Email address: `friz@math.tu-berlin.de`

(M. Maurelli) University of York and University of Edinburgh

Email address: `mario.maurelli@york.ac.uk`